Contributions to
Signal Processing for MRI

Marcus Björk

UPPSALA
UNIVERSITET

Dissertation presented at Uppsala University to be publicly examined in ITC 2446, Polacksbacken, Lägerhyddsvägen 2, Uppsala, Friday, 8 May 2015 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Andreas Jakobsson (Division of Mathematical Statistics, Lund University).

Abstract
Björk, M. 2015. Contributions to Signal Processing for MRI. *Uppsala Dissertations from the Faculty of Science and Technology* 113. xviii+176 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-554-9204-5.

Magnetic Resonance Imaging (MRI) is an important diagnostic tool for imaging soft tissue without the use of ionizing radiation. Moreover, through advanced signal processing, MRI can provide more than just anatomical information, such as estimates of tissue-specific physical properties.

Signal processing lies at the very core of the MRI process, which involves input design, information encoding, image reconstruction, and advanced filtering. Based on signal modeling and estimation, it is possible to further improve the images, reduce artifacts, mitigate noise, and obtain quantitative tissue information.

In quantitative MRI, different physical quantities are estimated from a set of collected images. The optimization problems solved are typically nonlinear, and require intelligent and application-specific algorithms to avoid suboptimal local minima. This thesis presents several methods for efficiently solving different parameter estimation problems in MRI, such as multi-component T_2 relaxometry, temporal phase correction of complex-valued data, and minimizing banding artifacts due to field inhomogeneity. The performance of the proposed algorithms is evaluated using both simulation and in-vivo data. The results show improvements over previous approaches, while maintaining a relatively low computational complexity. Using new and improved estimation methods enables better tissue characterization and diagnosis.

Furthermore, a sequence design problem is treated, where the radio-frequency excitation is optimized to minimize image artifacts when using amplifiers of limited quality. In turn, obtaining higher fidelity images enables improved diagnosis, and can increase the estimation accuracy in quantitative MRI.

Keywords: Parameter estimation, efficient estimation algorithms, non-convex optimization, multicomponent T2 relaxometry, artifact reduction, T2 mapping, denoising, phase estimation, RF design, MR thermometry, in-vivo brain

Marcus Björk, Department of Information Technology, Division of Systems and Control, Box 337, Uppsala University, SE-75105 Uppsala, Sweden. Department of Information Technology, Automatic control, Box 337, Uppsala University, SE-75105 Uppsala, Sweden.

To Science...
(not the journal)

List of papers

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

I M. Björk, R. R. Ingle, E. Gudmundson, P. Stoica, D. G. Nishimura, and J. K. Barral. Parameter estimation approach to banding artifact reduction in balanced steady-state free precession. *Magnetic Resonance in Medicine*, 72(3):880–892, 2014

II M. Björk, E. Gudmundson, J. K. Barral, and P. Stoica. Signal processing algorithms for removing banding artifacts in MRI. In *Proc. 19th European Signal Processing Conference (EUSIPCO-2011)*, pages 1000–1004, Barcelona, Spain, 2011

III M. Björk, D. Zachariah, J. Kullberg, and P. Stoica. A multicomponent T_2 relaxometry algorithm for myelin water imaging of the brain. *Magnetic Resonance in Medicine*, 2015. DOI: 10.1002/mrm.25583

IV M. Björk and P. Stoica. Fast denoising techniques for transverse relaxation time estimation in MRI. In *Proc. 21st European Signal Processing Conference (EUSIPCO-2013)*, pages 1–5, Marrakech, Morocco, 2013

V M. Björk and P. Stoica. New approach to phase correction in multi-echo T_2 relaxometry. *Journal of Magnetic Resonance*, 249:100–107, 2014

VI M. Björk and P. Stoica. Magnitude-constrained sequence design with application in MRI. In *Proc. 39th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-2014)*, pages 4943–4947, Florence, Italy, 2014

VII M. Björk, J. Berglund, J. Kullberg, and P. Stoica. Signal modeling and the Cramér-Rao Bound for absolute magnetic resonance thermometry in fat tissue. In *Proc. 45th Asilomar Conference on Signals, Systems, and Computers*, pages 80–84, Pacific Grove, CA, USA, 2011

Reprints were made with permission from the publishers.

Acknowledgements

First of all, I would like to thank my main supervisor, Professor Petre Stoica, for sharing his profound knowledge, supporting me throughout my PhD, and believing in me even when there was no reason to do so (I have never called you Peter and I won't start now, even if that is your *official* name). As you once said, you don't need PhD students to do your research for you, as you can do that better yourself, and therefore I am truly grateful that you took me on. You are my scientific role model, and I have really learned a great deal from you during these past years.

Of course, a big thanks also goes out to my co-supervisor Professor Alexander Medvedev. You got me started on biomedical applications, and for this, I am truly grateful. But I would also like to thank all the senior staff at the Division of Systems and Control for their engagement in the PhD students, and the general well-being of the division. Håkan, Kjartan, Hans, Bengt, Torsten, Torbjörn, Kristiaan, and more recently, Thomas and Dave, you have all played an important role in my development as a person, and a as young researcher.

A special thanks goes out to all past and present SysCon PhD student. You were the main reason I was happy to come to the office every day (well, if it wasn't raining or anything...). We have had a lot of fun, both inside and outside the office walls. More specifically, I would like to thank all the, more or less, crazy people who started around the same time as me, Soma, Margarida, Daniel, Olov, Pelle. I will never forget the ski trips, pub sessions, parties, and general freakouts. We sure raised a lot of hell at the division, and the department.

I am also grateful to Prof. Andreas Jacobsson, Lund university, for being my faculty opponent, and to Prof. Mats Viberg, Chalmers, Dr. Robin Strand, Uppsala university, Prof. Irene Guo, Chalmers, and Prof. Magnus Jansson, KTH, for agreeing to serve as committee members during the defense of this thesis.

My brothers and sisters, when we started we were just a bunch of rookie engineering-physics students, and now, we have evolved into something truly remarkable. I would like to thank you all for the years gone by, the many late nights, the beer, the kladdkaka, the best of company. Special thanks to our Rector Magnificus, FreddieP, I wish you and the family all the best in the future; but also to the others that are constantly pushing us forward, Charlii, DtB, re^{iz}, etc., your work is highly appreciated. Till the end of days...

Thanks to all my extraordinary collaborators, Dr. Erik Gudmundson, Dr. Joëlle Barral, Dr. Reeve Ingle, Prof. Dwight Nishimura, Dr. Joel Kullberg, Dr. Johan Berglund, your support and energy has given me strength, and together, we have made this thesis something to be proud of. I would especially like to thank my predecessor Erik, for being a good mentor, and a friend. I can always count on your wisdom.
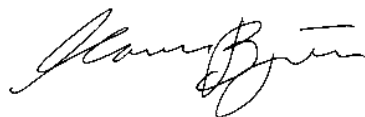
My dear Sandra, you are always supportive of me, and you have really helped me a lot by just being there. I could not have wished for a better partner. Every day, I get to come home to you, which makes me so happy. During my time as a PhD student, we have visited many places together, and had so much fun. You made all those conference trips a thousand times better. I truly love you!

My mother and father have always supported and believed in me, and that is the main reason I managed get this far in life, and write this thesis. I hope you know how much you mean to me, even though I might not express it all that often.

Last and least, thanks to my gray hair. You all must have really liked my thesis, judging from the exponential increase I witnessed after I started writing it. I'm done now, so feel free to go somewhere else. Really...

Marcus Björk
Uppsala
March 2015

Glossary and Notation

Abbreviations

BIC	Bayesian Information Criterion
bSSFP	balanced Steady-State Free Precession
BLUE	Best Linear Unbiased Estimator
CRB	Cramér-Rao Bound
EASI	Exponential Analysis via System Identification
FFT	Fast Fourier Transform
FID	Free Induction Decay
FOS	Feasibility-based Order Selection
FOV	Field Of View
GLS	Generalized Least Squares
GN	Gauss-Newton
GPU	Graphics Processing Unit
IQML	Iterative Quadratic Maximum Likelihood
i.i.d.	Independent and Identically Distributed
LASSO	Least Absolute Shrinkage and Selection Operator
LCQP	Linearly Constrained Quadratic Program
LM	Levenberg-Marquardt
LORE	Linearization for Off-Resonance Estimation
LP	Linear Program
LS	Least Squares
MACO	Magnitude-Constrained Cyclic Optimization
MC	Monte Carlo
ML	Maximum Likelihood
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
MSE	Mean Square Error
MT	Magnetization Transfer
MWF	Myelin Water Fraction
NLS	Nonlinear Least Squares
NNLS	Non-Negative Least Squares

NSA	Number of Signal Averages
qMRI	quantitative Magnetic Resonance Imaging
QP	Quadratic Program
rMSE	root Mean Square Error
RSD	Relative Standard Deviation
PDF	Probability Distribution Function
PRF	Proton Resonance Frequency
SAR	Specific Absorption Rate
SM	Steiglitz-McBride
SNR	Signal-to-Noise Ratio
SPICE	Sparse Covariance-Based Estimation
TE	Echo Time (variable)
TPC	Temporal Phase Correction
TR	Repetition Time (variable)
TV	Total Variation
WELPE	Weighted Linear Phase Estimation

Notation

$\mathbf{a}, \mathbf{b}, \ldots$	boldface lower case letters are used for vectors, for example, $\mathbf{a} = [a_1,\ a_2,\ \cdots]^{\mathrm{T}}$
$\mathbf{A}, \mathbf{B}, \ldots$	boldface upper case (capital) letters are used for matrices
$A, a, \alpha, \ldots$	non-bold letters are generally used to denote scalars
$\hat{\mathbf{A}}, \hat{\mathbf{a}}, \hat{a}, \hat{\alpha}, \ldots$	a hat, $\hat{\cdot}$, is used to denote an estimate
$\mathbf{I}$	the identity matrix (of unspecified dimension)
$\mathbf{I}_n$	the $n \times n$ identity matrix
$\mathbf{0}, \mathbf{1}$	the vector of all zeros or ones, respectively
$(\cdot)^{\mathrm{T}}$	vector or matrix transpose
$(\cdot)^*$	complex conjugate, or for vectors and matrices, the conjugate transpose
i	the imaginary unit, $\sqrt{-1}$, unless otherwise specified
$\mathbb{R}^{n \times m}$	the real-valued $n \times m$-dimensional matrix space
$\mathbb{R}^n$	the real-valued n-dimensional vector space ($\mathbb{R}$ is used for $n = 1$)
$\mathbb{C}^{n \times m}$	the complex-valued $n \times m$-dimensional matrix space
$\mathbb{C}^n$	the complex-valued n-dimensional vector space ($\mathbb{C}$ is used for $n = 1$)
$\mathscr{Z}$	the set if integer numbers
$\mathrm{Re}\{\cdot\}$	real part of a complex number
$\mathrm{Im}\{\cdot\}$	imaginary part of a complex number

$\arg(\cdot)$	phase of a complex number
$\mathrm{diag}(\cdot)$	diagonal; $\mathrm{diag}(\mathbf{a})$ means the matrix with the vector $\mathbf{a}$ on the diagonal and zeros everywhere else, and $\mathrm{diag}(\mathbf{A})$ means the vector containing the elements on the diagonal of the matrix $\mathbf{A}$
$\mathrm{bdiag}\left(\{\mathbf{A}_p\}_{p=1}^P\right)$	block diagonal matrix with P blocks, and the matrix $\mathbf{A}_p$ in block p (the matrices may have different sizes)
$\mathrm{tr}(\cdot)$	trace of a matrix
$\mathrm{vec}(\cdot)$	columnwise vectorized version of a matrix
$\ln(\cdot)$	natural logarithm
$\mathrm{mod}(a,b)$	modulo operation with dividend a and divisor b, with the result defined to be positive
$\mathcal{L}(\cdot)$	the likelihood function
q^{-1}	unit delay operator, $q^{-1}s(k)=s(k-1)$
$\{x_k\}_{k=1}^K$	a set of K elements x_k
$p(x\|y)$	probability distribution function of the variable x, conditioned on y
$\mathrm{Rice}(\eta,\sigma)$	the Rice distribution with parameters η and σ
$\sim$	distributed as; e.g., $\mathbf{x}\sim\mathcal{N}(\boldsymbol{\mu},\mathbf{R})$ means that $\mathbf{x}$ is Gaussian distributed with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{R}$
$\otimes$	the Kronecker product
$\triangleq$	defined as equal to
$\in$	belongs to; e.g., $\mathbf{a}\in\mathbb{C}^n$ means that $\mathbf{a}$ is an n-dimensional complex-valued vector, and $\mathbf{A}\in\mathbb{R}^{n\times m}$ means that $\mathbf{A}$ is a real-valued $m\times n$ matrix
∇	gradient operator
∇^2	Laplacian operator
$\|\cdot\|$	magnitude, or in the case of vectors, element-wise magnitude
$\|\cdot\|_p$	L_p-norm; $\|\mathbf{a}\|_p=\left(\sum_j\|a_j\|^p\right)^{1/p}$
$\|\cdot\|$	L_2-norm (Euclidian norm)
$\|\cdot\|_{\mathbf{W}}$	L_2-norm (Euclidian norm) weighted by the matrix $\mathbf{W}$
$\|\cdot\|_{\mathrm{F}}$	Frobenius norm of a matrix

Contents

Chapter 1

Introduction

Magnetic Resonance Imaging (MRI) is an important diagnostic tool for imaging soft tissue without the use of ionizing radiation. Furthermore, signal processing lies at the very core of the imaging process. These two factors make the interdisciplinary topic of signal processing for MRI especially interesting, as there are many diverse problems that can be tackled from experiment design to image processing. This chapter describes the contribution of the thesis, and provides an outline of the structure, including an overview of the treated signal processing problems.

1.1 Contribution

The contribution of this thesis is mainly in quantitative MRI (qMRI), which is the field of MRI signal processing where different physical quantities are estimated from measured data. Beyond that, one input design problem is included, where the excitation is optimized with the aim of minimizing image artifacts, and in turn, improving diagnosis, or increasing the estimation accuracy in qMRI.

The theme throughout the thesis is optimization, with the goal of finding efficient algorithms for obtaining the desired quantities. Typically, the estimation problems arising in MRI are nonlinear and nonconvex, meaning that approximations or problem reformulations are often needed to make estimation tractable. As a result of this, the algorithms developed are in many ways application specific.

A set of interesting problems has been treated, and in each case, the goal has been to analyze the problem and derive new efficient algorithms for estimation. The work has also involved model development, and validation on data from human subjects. More specific details on the contribution of this thesis is given in the next section.

1.2 Thesis outline

The thesis consists of two parts. The first is a basic introduction to the field of MRI, including magnetic resonance (MR) physics, the imaging process, and the MR scanner. The points of intersection between signal processing and MRI are described, with focus on modeling and parameter estimation for qMRI. Finally, some background to the signal processing techniques used in the thesis, such as, maximum likelihood estimation and image filtering, is given. In the second part, a few specific signal processing problems in MRI are presented. The overall concern is to find efficient estimation algorithms for these typically nonlinear problems. A brief summary of each chapter in Part II is given below, including references to the corresponding papers.

Chapter 4: Off-resonance mapping and banding removal

Banding artifacts causing signal loss and obstructing diagnosis is a major problem for the otherwise efficient bSSFP protocol in MRI. A fast two-step algorithm for 1) estimating the unknowns in the bSSFP signal model from multiple phase-cycled acquisitions, and 2) reconstructing band-free images, is presented. The first step, Linearization for Off-Resonance Estimation (LORE), approximately solves the nonlinear problem by a robust linear approach. The second step applies a Gauss-Newton algorithm (GN), initialized by LORE, to minimize the nonlinear least squares criterion. The full algorithm is named LORE-GN. By deriving the Cramér-Rao bound it is shown that LORE-GN is statistically efficient; and moreover, that simultaneous estimation of T_1 and T_2 from phase-cycled bSSFP is difficult, since the variance is bound to be high at common SNR values. Using simulated, phantom, and in-vivo data, the band-reduction capabilities of LORE-GN are illustrated, and compared to other techniques, such as the sum-of-squares. It is shown that LORE-GN is successfully able to minimize banding artifacts in bSSFP where other methods fail, for example, at high field strengths. This chapter is based on papers I and II.

Chapter 5: Multi-component T_2 relaxometry and myelin-water imaging

Models based on a sum of damped exponentials occur in many applications, particularly in multi-component T_2 relaxometry. The problem of estimating the relaxation parameters and the corresponding amplitudes is known to be difficult, especially as the number of components increases. In this chapter, a parameter estimation algorithm called EASI-SM is compared to the non-negative least squares (NNLS) spectrum approach commonly used in the context of MRI. The performance of the two algorithms is evaluated via simulation using the Cramér-Rao

bound. Furthermore, the algorithms are applied to in-vivo brain multi-echo spin-echo dataset, containing 32 images, to estimate the myelin water fraction and the most significant T_2 relaxation time. EASI-SM is shown to have superior performance when estimating the parameters of multiple relaxation components in simulation, and in vivo, it results in a lower variance of the T_2 point estimates. It provides an efficient and user-parameter-free alternative to NNLS, and gives a new way of estimating the spatial variations of myelin in the brain. This chapter is based on paper III.

Chapter 6: Edge-preserving denoising of T_2 estimates

Estimating the transverse relaxation time, T_2, from magnitude spin echo images is a common problem in MRI. The standard approach is to use voxelwise estimates; however, noise in the data can be a problem when only two images are available. By imposing inter-voxel information it is possible to reduce the variance of the T_2 estimates, but this typically compromises the details in image, especially at tissue boundaries. By developing intelligent algorithms that use data from several pixels, the estimation performance can improved without affecting tissue contrast. An optimal formulation of the global T_2 estimation problem is nonlinear, and typically time consuming to solve. Here, two fast methods to reduce the variance of the T_2 estimates are presented: 1) a simple local least squares method, and 2) a total variation based approach that can be cast as a linear program. The two approaches are evaluated using both simulated and in-vivo data. It is shown that the variance of the proposed T_2 estimates is smaller than the pixelwise estimates, and that the contrast is preserved. This chapter is based on paper IV.

Chapter 7: Temporal phase correction

Estimation of the transverse relaxation time, T_2, from multi-echo spin-echo images is usually performed using the magnitude of the noisy data, and a least squares (LS) approach. The noise in these magnitude images is Rice distributed, which can lead to a considerable bias in the LS-based T_2 estimates. One way to avoid this bias problem is to estimate a real-valued and Gaussian distributed dataset from the complex-valued data, rather than using the magnitude. In this chapter, two algorithms for phase correction, which can be used to generate real-valued data suitable for LS-based parameter estimation approaches, are proposed. The first is a Weighted Linear Phase Estimation algorithm, abbreviated as WELPE. This method provides an improvement over a previously published algorithm, while simplifying the estimation procedure and extending it to support multi-coil input. The second method is a maximum likelihood estimator of the true decaying signal magnitude,

which can be efficiently implemented when the phase variation is linear in time. The performance of the algorithms is demonstrated via Monte Carlo simulations, by comparing the accuracy of the estimated decays. Furthermore, it is shown that using one of the proposed algorithms enables more accurate T_2 estimates in multi-component T_2 relaxometry, compared to when using magnitude data. The practical feasibility of WELPE is illustrated by applying it to a 32-echo in-vivo brain dataset. This chapter is based on paper V.

Chapter 8: Sequence design for excitation

When using amplifiers of limited quality, signal distortion can be a problem, which in turn can result in image artifacts. Here, an algorithm for sequence design with magnitude constraints is presented. Such sequences can, for example, be used to achieve the desired excitation pattern in parallel MRI, when several low-cost amplifiers are used. The formulated non-convex design optimization criterion is minimized locally by means of a cyclic algorithm, consisting of two simple algebraic substeps. Since the proposed algorithm truly minimizes the criterion, the obtained sequence designs are guaranteed to improve upon the estimates provided by a previous method, which is based on the heuristic principle of the Iterative Quadratic Maximum Likelihood algorithm. The performance of the proposed algorithm is illustrated in two numerical examples. This chapter is based on paper VI.

Chapter 9: Magnetic resonance thermometry

Measuring the temperature using MRI has applications in thermal therapy and metabolism research. In tissue containing both fat and water resonances it is possible to obtain an absolute measure of the temperature through parametric modeling. The fat resonance is used as a reference to determine the absolute water resonance frequency, which is linearly related to the temperature of the tissue. In this chapter, the feasibility of using this method to estimate the absolute temperature in fat tissue is investigated. Using the Cramér-Rao bound, it is shown that the highest obtainable accuracy at common SNR is too low for the application in mind, when using a 1.5 T scanner. However, increasing the field strength can improve the bound significantly. Moreover, it is shown that the choice of sampling interval is important to avoid signal cancellation. It is concluded that to make proton resonance frequency-based temperature mapping feasible, a high SNR is typically needed. This chapter is based on paper VII.

1.3 Future work

There are still many open signal processing problems in MRI, and many current techniques that could be optimized and further developed using methods similar to those presented in this thesis.

Array processing has relatively recently been getting more attention in the MR community. Processing of MRI data from multiple receive coils enables significant improvements in image quality, reduction in hardware costs, or speedup of the imaging process. The flexibility provided by trading off these three strengths makes efficient methods for multi-coil signal processing valuable, both for research and clinical use. For example, methods like sensitivity encoding (SENSE) enable a significant acquisition speedup [97]. The method is based on the fact that the receive coils used have slightly different properties. Given accurate measurements of each coil's sensitivity profile in space, the gathered information can be combined into one image, reducing the noise, or in the case of acquisition speedup, removing the aliasing. The SENSE approach has since been refined and extended in several ways [71, 98], and inspired other approaches [60, 21], but the main drawback of SENSE remains, namely that it requires the coil sensitivities to be known. Measuring the coil sensitivities takes time, and as they change over time, this is typically done at the beginning of every scan session, accelerating the acquisition of the following images. Developing algorithms to make the best use of these multi-image datasets, either for speed, image quality, or estimation accuracy, is still a major topic for the future.

To reduce the acquisition time further, compressed sensing can be used, where a fraction of the data samples are collected, without compromising the image quality or the diagnostic capabilities [83]. By collecting fewer samples, the time a patient spends in the scanner can be reduced; however, the image reconstruction problem becomes more complicated and requires advanced signal processing algorithms. Compressed sensing is one of the hottest research topics in MRI today, an will most likely only grow in the years to come.

Previously, many algorithms for qMRI were developed for single coil data, and furthermore, the compressive sensing and image reconstruction was often treated separately from the parameter estimation problem. Attempts to integrate parallel MRI and compressive sensing with qMRI, to fully take advantage of both the data and the problem structure, will lead to many interesting and challenging estimation problems in the future; and by developing efficient algorithms to solve these problems, much is to be gained, in terms of estimation quality and acquisition speedup.

Part I:
Introduction to MRI and signal processing

Chapter 2

MR physics and imaging

This chapter provides a friendly introduction to MR physics, the imaging process, and the MR scanner. The discussion is based on my understanding of the field, and is by no means complete. For a more extensive description of the topic, the reader is referred to, for example, [89, 62, 9, 132, 63], and the references therein.

2.1 Nuclear magnetic resonance

MRI typically utilizes the magnetic moments of hydrogen nuclei, or *protons*; but it is also possible to image based on other substances, such as phosphorus. As the human body consists of a large proportion of water molecules, fat, and other organic molecules, all containing hydrogen, proton-based imaging is particularly useful to obtain soft tissue contrast.

When a proton is placed in an external magnetic field of strength B_0, its magnetic moment $\boldsymbol{\mu}$ aligns with the direction of the field. The size of the magnetization depends on the field strength, which is measured in Tesla (T). By applying a radio frequency (RF) pulse, it is possible to excite the protons, effectively making the magnetization flip a certain angle, α, relative to the B_0 field, see Fig. 2.1. After the pulse, the magnetic moment of the proton rotates freely around the B_0 field according to the Larmor precession law, and eventually returns to its equilibrium position through a process called relaxation, see Fig. 2.2. The changes in the magnetic field during the relaxation can be captured by receiver coils, which produces a free induction decay (FID) signal. The resonance, that is, the frequency at which the protons can absorb energy, is called the Larmor frequency, and it is proportional to the local magnetic

Table 2.1: Approximate proton densities in percent, and relaxation times T_1 and T_2 in milliseconds, for different tissues at $B_0 = 1.5$ T, as stated in [132].

Tissue	PD	T_1 [ms]	T_2 [ms]
White matter	70	780	90
Gray matter	85	920	100
Fat	100	260	80
Cerebrospinal fluid	100	> 4000	> 2000

field strength, that is

$$\omega = -\gamma B_0, \tag{2.1}$$

where γ is a constant called the *gyromagnetic ratio*. Note that γ depends on the nuclei imaged, and for protons $\gamma \approx 42.58$ MHz/Tesla, which implies that the precession about magnetic field vector **B** is clockwise.

The relaxation of the magnetization in a small volume element (voxel) has two components: the first describes the exponential recovery of the longitudinal magnetization and is denoted by the time constant T_1, the second describes the transverse relaxation, which is due to the loss of phase coherence between protons in the voxel, and is denoted by T_2. Using the fact that the proton density (PD) and relaxation times depend on the tissue, it is possible to obtain contrast. It should be noted that the absolute relaxation times also depend on B_0; however, the physical principles are the same regardless of the field strength. The approximate PD, T_1, and T_2 values for different tissue types at $B_0 = 1.5$ T is shown in Table 2.1.

The equations that govern the macroscopic behavior of a magnetic moment in an external magnetic field are called the Bloch equations, and are given by

$$\begin{aligned}
\frac{dM_x(t)}{dt} &= \gamma(M_y(t)B_z(t) - M_z(t)B_y(t)) - \frac{M_x(t)}{T_2}, \\
\frac{dM_y(t)}{dt} &= \gamma(M_z(t)B_x(t) - M_x(t)B_z(t)) - \frac{M_y(t)}{T_2}, \\
\frac{dM_z(t)}{dt} &= \gamma(M_x(t)B_y(t) - M_y(t)B_x(t)) - \frac{M_z(t) - M_0}{T_1},
\end{aligned} \tag{2.2}$$

where the components $M_{x,y,z}(t)$ and $B_{x,y,z}(t)$, describe the time evolution in $\mathbb{R}^3$ of magnetization and the external magnetic field, respectively, and M_0 is the equilibrium magnetization, which depends on the proton density. Often, the transverse components are represented by a single complex-valued quantity, that is, $M_{xy} = M_x + iM_y$ and $B_{xy} = B_x + iB_y$,
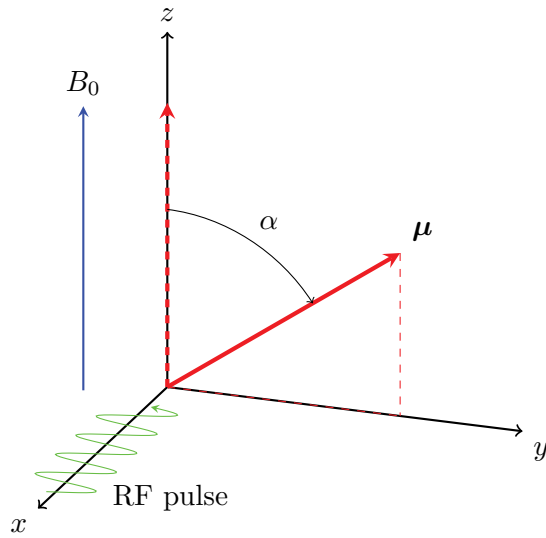
Figure 2.1: Flip of the magnetic moment μ by an angle α relative to the static magnetic field B_0.
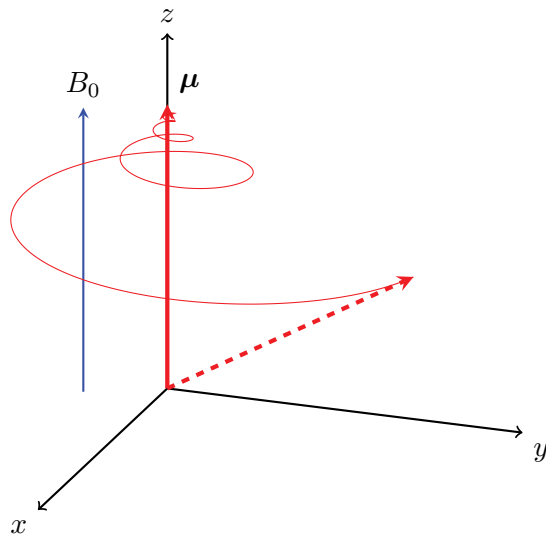

Figure 2.2: The relaxation and precession of a magnetic moment μ in a static magnetic field B_0, after excitation.

which enables a more compact form of (2.2):

$$
\begin{aligned}
\frac{dM_{xy}(t)}{dt} &= -i\gamma(M_{xy}(t)B_z(t) - M_z(t)B_{xy}(t)) - \frac{M_{xy}(t)}{T_2}, \\
\frac{dM_z(t)}{dt} &= i\frac{\gamma}{2}(M_{xy}(t)B_{xy}^*(t) - M_{xy}^*(t)B_{xy}(t)) - \frac{M_z(t) - M_0}{T_1},
\end{aligned}
\tag{2.3}
$$

where the $*$ indicates the complex conjugate. These Bloch equations are nonlinear and coupled, and solving them for arbitrary magnetic field changes can be difficult. However, the solution for a static magnetic field $B_z(t) = B_0$ ($B_x = B_y = 0$) is easily obtained as

$$
M_{xy}(t) = M_{xy}(0)e^{-i\omega_0 t - t/T_2},
\tag{2.4}
$$

$$
M_z(t) = M_0(1 - e^{-t/T_1}) + M_z(0)e^{-t/T_1},
\tag{2.5}
$$

and this type of behavior was illustrated in Fig. 2.2. Ideally, the measured FID in the xy-plane decays exponentially with a time constant T_2 that only depends on the tissue present on the current voxel; in practice, however, inhomogeneities in B_0 will cause additional decoherence, effectively shortening T_2. The observed FID decay rate is therefore denoted T_2^*, where the star should not be confused with the complex conjugate. For a deviation ΔB_0 from the ideal field strength B_0, we can write the following implicit expression for T_2^*:

$$
\frac{1}{T_2^*} = \frac{1}{T_2} + \gamma\Delta B_0,
\tag{2.6}
$$

which implies that $T_2^* \leq T_2$. Even deviations in the order of one percent can have a significant impact on the decay time, especially when the T_2 decay is slow. To counter this rapid signal decay, different pulse sequences or excitation schemes can be used, as is discussed in Section 2.2.3.

As the basal rotation frequency ω_0 is known, it is possible to introduce a rotating frame of reference, which simplifies the Bloch equations as well as the description of the excitation. This formalism will be used in the following, effectively removing ω_0 from all equations, and where applicable, only modeling the off-resonance frequency.

2.2 The imaging process

2.2.1 Excitation

To extract information regarding the subject under study, excitation is needed. An RF pulse is used to excite the sample and rotate the magnetic moments, or *spins*, relative to the B_0 field. There are many

application-specific RF pulses, but this introduction will mainly cover the basic pulse shapes and their uses.

The RF pulses can be designed to flip the magnetization to different angles, but it is also possible to use several consecutive pulses of different types. The spin-echo experiment is a simple example where two excitation pulses are used. The first excites the spins, while the next pulse refocuses the spins to generate a so called *echo*. This technique is used to counter imperfections in the system, and reduce the resulting artifacts in the images. For more details on the spin-echo sequence, see Section 2.2.3.

For 2D imaging, the excitation is typically in the form of sinc pulses, $\text{sinc}(x) = \sin(\pi x)/\pi x$, which are designed to be narrowband. The idea is to excite a specific range of frequencies and get a clear slice profile, and for this, the box-shaped spectrum of the sinc function is useful. In practice, the sinc pulse shape needs to be truncated or windowed, giving a non-ideal spectrum, as shown in Fig. 2.3. In turn, this leads to a non-ideal flip of the in-slice protons, but also excitation leakage into adjacent slices, so called *cross talk*.

Parallel excitation using several transmitter coils can be used to achieve higher fidelity in the excitation profiles. By dividing the load on several units, the need to transmit high quality signals is reduced, adding robustness to the excitation while enabling the use of low-cost amplifiers. Moreover, the range of usable excitation signals is increased, as the protons are affected by the net frequency content of the magnetic field.

It also is possible to design RF pulses for a specific application, taking time limitations and other physical constraints into account. The resulting optimization problems are typically nonlinear, and can require intelligent algorithms to solve. For example, the Shinnar-Le Roux algorithm can be used to design pulses with specific spectral profiles [93]. Input design is a broad topic in signal processing which is discussed more in Section 3.1.3, and a specific design problem is treated in Chapter 8.

Rectangular pulses are also used in some cases, particularly in 3D encoding when no slice selection is required. Usually, some type of windowing is needed, as realizing a rectangular pulse is difficult in practice. The problem lies in creating a pulse with sufficient bandwidth to excite all the frequencies of interest, which in turn requires the transmitted signal to be short in time, according to the time-bandwidth product. As there is only an upper limit to the duration of the excitation in clinical practice, the 3D sequences can typically generate higher quality images than 2D sequences, within a given time frame.
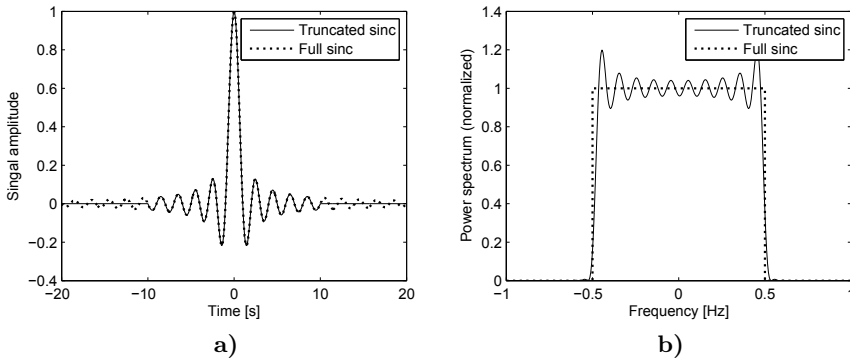
Figure 2.3: a) A sinc-type RF pulse and its truncated version, and b) the corresponding normalized spectra.

2.2.2 Encoding and gradients

To acquire 3D information regarding the anatomy of the scanned subject, the data must be encoded in space. This can be achieved by actively changing the magnetic fields while transmitting one or several excitation pulses. The encoding can either be done in a number of 2D slices, or directly in 3D. Here, we will focus on 2D Fourier encoding with slice selection, which is the classical approach. Initially, a gradient magnetic field across the z-direction is applied during excitation from a narrow-band pulse. According to (2.1), this linear gradient will cause planes orthogonal to the z-axis to have different resonance frequencies. A narrow-band signal can therefore excite a specific slice along the z-axis, see Fig. 2.4. Following that, a gradient across the y-direction is applied for a limited time period, to create a linear offset of the phases of all spins along the y-axis; this is called *phase encoding*. Lastly, a readout gradient across the x-direction is applied while recording the signal. This step is called frequency encoding as the spins across the x-direction will have different frequencies during the data collection. After waiting for the system to return to equilibrium, the experiment is repeated with different phase gradients to obtain 2D information. The time between experiments is denoted *TR*, and is called the *repetition time*. The spatial phase and frequency distribution of the spins after a single 2D encoding step, is illustrated in Fig. 2.5. To collect 3D information with this approach, several slices are collected. For direct 3D image acquisition, the phase encoding is performed along two spatial dimensions.

With the encoding scheme outlined above, the phase of a spin at location (x, y), at time t' during the readout gradient G_x, and for a phase-encoding gradient G_y applied for τ seconds, can be expressed as

$$\phi(t', \tau) = \gamma G_x x t' + \gamma G_y y \tau, \tag{2.7}$$
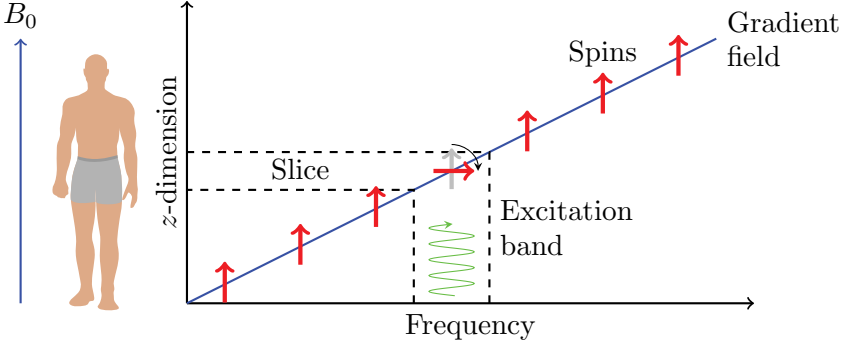
Figure 2.4: In 2D imaging, a narrowband excitation pulse is used to flip the magnetization of a certain slice in the 3D subject, as the resonance frequency depends on the field strength.

where the phase has been demodulated with the known Larmor frequency ω_0. In practice, G_y is varied rather than τ, but to make the description more intuitive, we will assume that τ is the variable here. In sum, t' represents the time index during each readout across the x-dimension, and τ is the time index for each phase-encoding step along the y-dimension. The received time-domain signal is a superposition of the signals from all the spins in the slice, and can be written as

$$S(t', \tau) = \iint \rho(x, y) e^{i\gamma(G_x x t' + G_y y \tau)} \mathrm{d}x\mathrm{d}y, \tag{2.8}$$

where $\rho(x, y)$ corresponds to the PD at the spatial location (x, y). By defining $k_x = -\gamma G_x t'$ and $k_y = -\gamma G_y \tau$, we can write (2.8) as

$$S(k_x, k_y) = \iint \rho(x, y) e^{-i(k_x x + k_y y)} \mathrm{d}x\mathrm{d}y, \tag{2.9}$$

which is a 2D Fourier transform of the image $\rho(x, y)$ in terms of the wave numbers k_x and k_y. This encoding ensures that all spins have different sets of $\{k_x, k_y\}$, and therefore we can reconstruct a spatial image of the subject by a simple inverse Fourier transformation, which can be efficiently computed using the Fast Fourier Transform (FFT). For 3D encoding with two phase encoding steps, an inverse 3D Fourier transform is used. The wave number notation used in (2.9) brings us to the concept of k-space, the image Fourier domain. In general, we can write the received signal as

$$S(t) = \int \rho(\mathbf{x}) e^{-i\mathbf{k}(t)\cdot\mathbf{x}} \mathrm{d}\mathbf{x}, \tag{2.10}$$

where $\mathbf{k}(t)$ can be seen as an arbitrary k-space trajectory in time. This formalism enables us to think of many other ways of obtaining data. By
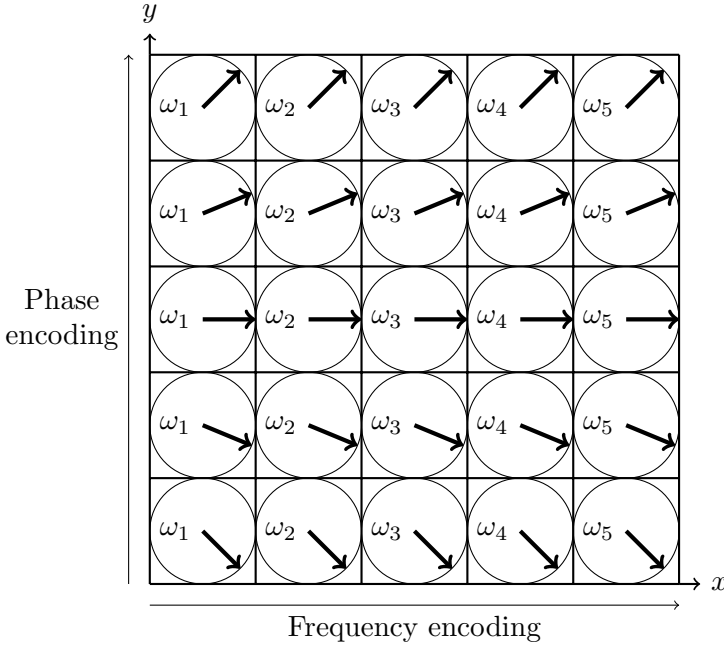
Figure 2.5: Illustration of the achieved phase and frequency distribution in space, after a single 2D-encoding step. Each voxel has a different frequency and phase pair.

changing the gradient fields, $\mathbf{k}(t)$ can trace out almost any curve in k-space. In fact, there are endless possibilities of how to combine gradients and RF pulses, which makes MRI remarkably versatile. In the scenario described above, the readouts are in the form of straight lines from left to right along k_x, and the phase encoding can, for example, start from negative k_y and advance with equal steps to positive k_y, by changing τ or G_y in each acquisition. A few examples of k-space trajectories are given in Section 2.2.3.

From the above description of the data collection, the scan time can be expressed as

$$\text{Scan time} = TR \times N_y \times \text{NSA}, \tag{2.11}$$

where N_y is the number of phase-encoding steps (in the y-direction), NSA is the number of signal averages (collected images), and TR is the time between two consecutive readouts, which is long enough to accommodate both signal decay and necessary k-space traveling. The number of samples in the x-direction does not influence the scan time, as this readout time is much shorter than TR.

There are also a wide range of correction gradients that are often used to minimize the effects of imperfect hardware, physical phenomena such as eddy currents, and motion of the subject. For example, spoiler or

crusher gradients can be used to eliminate residual signal remaining at the end of each excitation cycle, which could otherwise cause the signal to eventually saturate at zero. This is accomplished by using the slice-selection gradient to dephase the spins before the next repetition. The spoiler gradients enable the use of a shorter TR, and therefore a shorter scan time, as there is no need to wait for the magnetization to naturally return to its equilibrium. Although many of the other correction gradients are important to avoid image artifacts, their details are beyond the scope of this discussion.

2.2.3 Pulse sequences and contrast

The acquisition scheme, or the pulse sequence, is usually visualized in a timing diagram. An example of the previously mentioned spin-echo sequence is shown in Fig. 2.6. The timing diagram shows when RF pulses and gradients are applied, as well as when a signal or echo is produced and sampled. By the use of timing diagrams it is easy to illustrate differences between acquisition schemes, and schematically understand how each sequence works.

Regardless of the sequence, an acquired image will have a combination of contrast from PD, T_1 and T_2 (or T_2^*). By modifying the sequence, that is, when and how the magnetization is flipped, it is possible to make one of these contrasts dominate. There is wide range of contrasts that can be achieved, and depending on the application, one can be more useful than the other. Because of this, many pulse sequences have been designed to generate a certain type of contrast, and in the following, three basic sequence types are presented.

As previously mentioned, an echo is sometimes formed to cancel system imperfections, but it also gives time for the spatial encoding to take place. In the basic spin echo sequence, the first pulse flips the magnetization into the xy-plane, and a second pulse is applied $TE/2$ seconds later, flipping the magnetization $180°$. The decay resulting from dephasing due to field inhomogeneity will be reversed by the $180°$ pulse, and after another $TE/2$ seconds, the spins will refocus, giving an echo which is sampled. One spin echo collects data from a line in k-space, and after a time period TR, the sequence is repeated with different phase encodings until a sufficient number of k-space samples has been acquired. By this method, the extra dephasing effect observed in practical FIDs, characterized by T_2^*, can be avoided, to obtain an image whose contrast reflects the T_2-values of the tissue, so called T_2 *weighting*. Because of this, the spin echo typically provides high quality images; however, the sequence is associated with a long scan time.

Another basic pulse sequence is the gradient echo. It is similar to the spin echo except that no refocusing RF pulse is used. Instead,
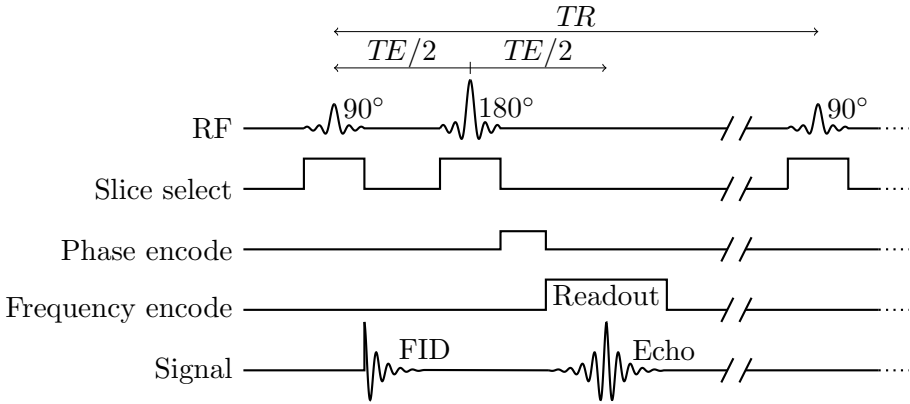
Figure 2.6: A schematic timing diagram for a typical spin echo, showing the RF pulses, gradients applied, and the echo and repetition times *TE* and *TR*, and the readout.

Table 2.2: Relationship between sequence parameters and contrast for the gradient echo.

Weighting:	T_1	T_2^*	PD
TE	Short	Long	Short
TR	Short	Long	Long
α	Large	Small	Small

two gradients of opposite polarity are used to dephase and rephase the spins, giving an echo. This does not cancel the dephasing due to field inhomogeneity, and unlike the spin echo, the signal decays with time constant T_2^*. The gradient echo can be varied by changing the sequence parameters: *TE*, *TR*, and the flip angle of the RF pulse, α. Table 2.2 summarizes the approximate relationship between these parameters and the obtained image weighting, and as can be seen, all main types of contrasts can be achieved. Because the gradient echo only uses a single RF pulse and supports small α:s, it is possible to use a shorter *TE* and *TR*, enabling rapid imaging.

The final basic pulse-sequence type is the inversion-recovery sequence. It is mainly used for T_1-weighted imaging, but can also be designed to generate T_2 predominance. Inversion recovery is basically a spin echo with an initial 180° setup pulse. Flipping the equilibrium magnetization 180° will direct it along the negative z-axis, with no transverse component. According to (2.5), it will start relaxing along the z-axis with a rate dictated by T_1, through the origin, back to the positive equilibrium M_0. After a certain time, called the *inversion time*, a spin echo is typically performed, but there are also other alternatives. By changing this inversion time based on the expected T_1 for a certain tissue type,

the spin echo can be performed when the magnetization of this tissue is zero, effectively canceling its contribution to the signal. This is mainly used the achieve fat suppression or fluid attenuation in the resulting images.

Multi-slice methods can be used to improve the efficiency of otherwise slow sequences, such as the spin echo. Particularly, in sequences with long TR, significant time is spent waiting for the longitudinal magnetization to relax. By exciting and recording other slices during this wait, the protocol can be made more efficient.

Another option to improve the efficiency of the standard spin-echo and gradient-echo sequences is to use multi-echo techniques. In the spin-echo case, the initial echo is followed by a sequence of refocusing pulses, each giving a new smaller echo that is recorded. In this manner, several images can be acquired within the same scan time as a single image with the basic approach. A schematic view of the recorded signals, including the T_2 and T_2^* decay, is shown in Fig. 2.7. Another multi-echo sequence is the echo-planar gradient echo, for which the entire k-space can be sampled during one TR. This is accomplished by successively applying rephasing gradients and alternating the polarity of the readout gradient. The fast scan time enables the capture of rapid physiological processes, such as cardiac motion; however, due to the long readout, the amount of signal decay is not identical for all lines in k-space. This means that the image will have hybrid T_2^* weighting.

The basic sequences often use a Cartesian sampling of k-space, like Fourier encoding, where consecutive lines are collected giving a uniform rectangular grid. This approach requires repeated travel in k-space without collecting any data, for example, to get back to the left side and start a new readout. The idea of echo-planar imaging is to traverse the k-space more efficiently, but still on a Cartesian grid. Figure 2.8 illustrates the k-space trajectories for Fourier and echo planar imaging. It should be noted that, as there is a limit to how fast the k-space can be traversed, different means of obtaining the same samples can lead to different results. This is mainly due to hardware imperfections and signal decay during the acquisition, and these two aspects should be taken into consideration when designing a pulse sequence.

There are also a wide range of trajectories that do not sample the k-space on a uniform grid; for example, it is possible to use a rectangular grid with varying density. In k-space, the central values correspond to the low frequency content, that is, mainly the image contrast; while the outer regions are high frequency details. Looking at an image in k-space, it is clear that most of the energy is located around the center, as is shown in Fig. 2.9, which would support a denser sampling in this region. The spiral sequences typically have this property, as well as
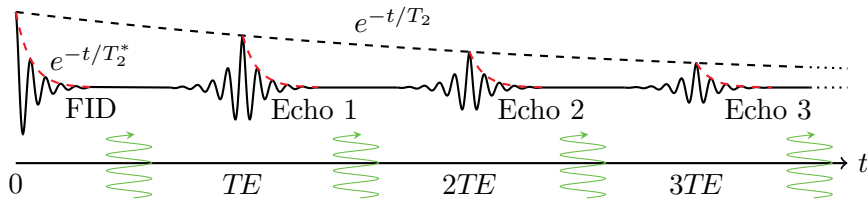
Figure 2.7: The signal from a multi-echo spin-echo sequence, showing the T_2 and T_2^* decay. The sinusoids indicate RF excitation in the form of 180° refocusing pulses.
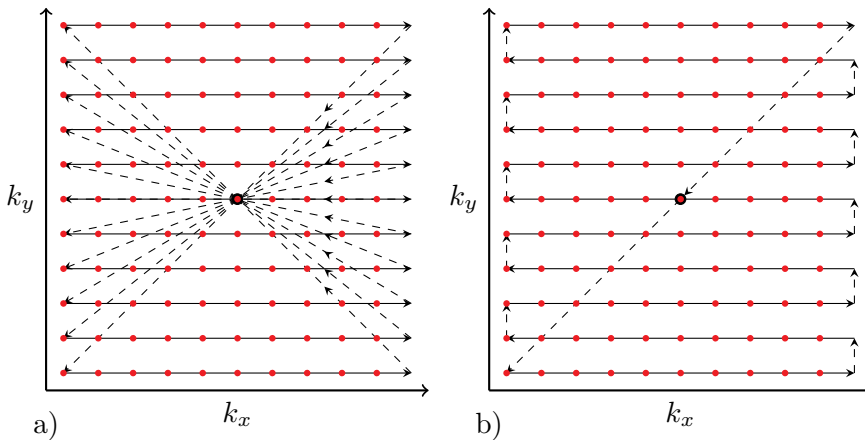
Figure 2.8: Example of two Cartesian k-space trajectories: a) Fourier, and b) echo planar sampling. Dashed lines indicate traveling in the k-space without sampling, and the red dots indicate sampling points.

radial sampling, and Fig. 2.10 shows an example of two such nonuniform k-space trajectories.

Finally, more advanced techniques such as steady-state free precession (SSFP), which is a type of gradient echo, or diffusion weighted sequences, which provides a different type of contrast, are also available. In steady-state imaging, a series of RF pulses and subsequent relaxations eventually lead to an equilibrium of the magnetization from one repetition to the next. This is useful for rapid acquisitions, as there is no need to wait for full T_1 or T_2 recovery before the next pulse is applied. In the balanced SSFP sequence (bSSFP), the goal is to preserve as much magnetization as possible by using balanced gradients that reduces the dephasing during each repetition. This leads to high SNR and fast scans, but the method is sensitive to imperfections in the static magnetic field. The obtained signal depends on T_1, T_2, TE, TR, and α; and the images typically have a T_1/T_2-type contrast.
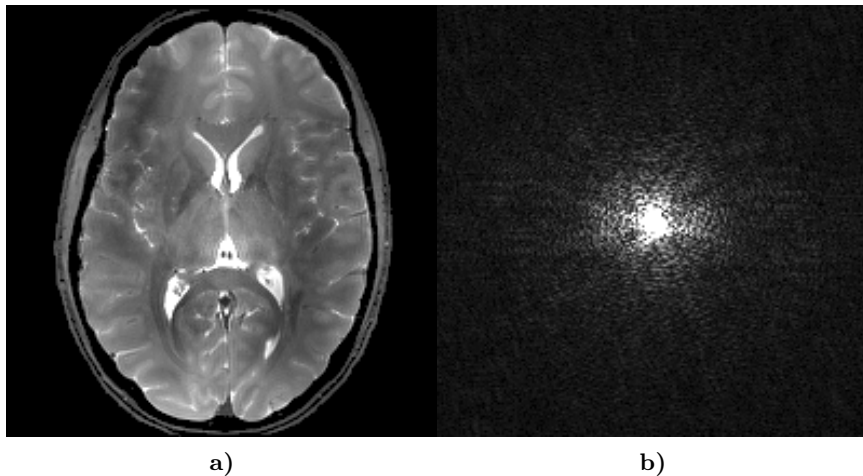
Figure 2.9: a) Example of the magnitude of a brain image, and b) the magnitude of the corresponding k-space.

2.2.4 Reconstruction

If k-space is sampled on a sufficiently dense uniform rectangular grid, the simplest type of image reconstruction, the inverse discrete Fourier transform, can be applied to obtain the sought image. In the Fourier encoding case, the density of the grid corresponds to the field of view (FOV), that is, size of the imaged area. To avoid aliasing, the spacing of k-space samples must be smaller that the inverse of the FOV, that is $\Delta k < 1/\text{FOV}$. The extent of the grid, or the k-space area covered, will determine the resolution of the reconstructed image.

When the sampling is nonuniform, the standard inverse Fourier transform cannot be used. There are two possibilities, either the data is interpolated giving a uniform grid, or a nonlinear reconstruction problem needs to be solved. In MRI, the standard approach is to use the nonuniform fast Fourier transform (NUFFT), which makes use of an efficient algorithm to compute the transform [44]. However, the general reconstruction problem is non-convex, and developing algorithms to find the optimal solution is a signal processing problem.

In parallel MRI, the final reconstructed image is based on data collected from several parallel coils. This enables speedup of any pulse sequence, effectively widening the their applicability. The main idea is to reduce the number of collected k-space samples, in particular, the number of phase-encoding steps, as the scan time is typically proportional to this quantity, as was shown in (2.11). This can be achieved in several ways, for example, by collecting every other line in k-space. Undersampling in the phase-encode direction leads to aliasing, similar
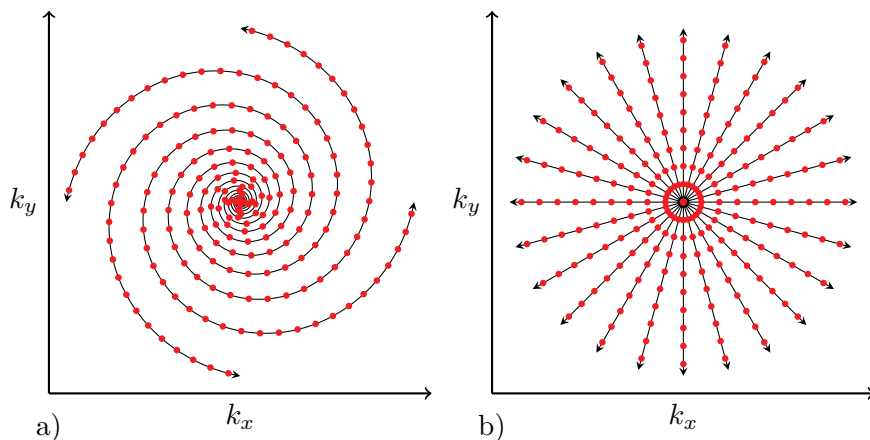
Figure 2.10: Example of two nonuniform k-space trajectories: a) spiral, and b) radial sampling. The red dots indicate sampling points.

to what is shown in Fig. 2.11; however, using data collected by several individual coils, it is possible to remove the aliasing and reconstruct the full image [21].

Another approach is to use sparse signal processing to reconstruct the image from a single set of undersampled data. These compressed sensing techniques enable a significant acquisition speedup, often at virtually no loss in image quality [83]. The basic idea is to use a domain in which the image is sparsely represented, and regularize the reconstruction problem to avoid aliasing. Again, the problem is non-convex, and therefore the reconstruction is often only an approximation of the true optimum, and in some cases, it can even fail. The sparse reconstruction problem is a challenge for signal processing, especially when it is combined with parameter estimation along the lines of Section 2.3.

2.2.5 Image quality and time

Signal-to-noise ratio

The main objective in MRI is to acquire high-SNR images, with sufficient resolution, in a short scan time. However, for a given pulse sequence and hardware, it is only possible to trade one of the above properties for the other, according to the simplified SNR equation [63]:

$$\text{SNR} \propto (\text{voxel volume})\sqrt{\text{acquisition time}} . \qquad (2.12)$$

The background of this equation will be outlined in the following, together with hardware-related means of improving SNR. It should, however, be noted that image artifacts due to, for example, hardware im-
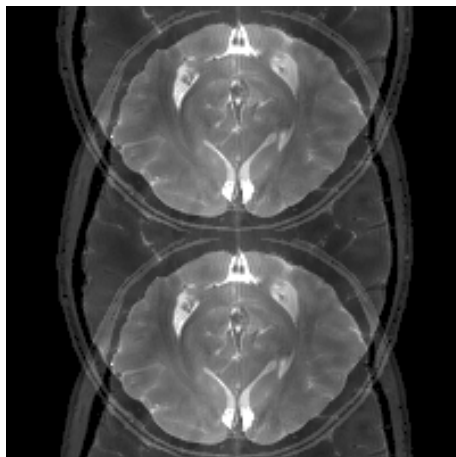
Figure 2.11: Example of the aliasing occurring when collecting every other line in the phase-encode direction.

perfections will also affect the images and the effective SNR. This topic is discussed in the next section.

Even though (2.12) is rather restrictive, there are still several ways in which the acquisition can be adapted to the application at hand, giving priority to the more crucial properties. Overall, the SNR depends on the following factors:

- Scan time
- Resolution
- Number of acquisitions
- Scan parameters (TE, TR, and α)
- Magnetic field strength
- Transmit and receive system

The reasons for prioritizing a quick scan are both financial and practical. For example, motion during the scan can be reduced with a fast scan, and some rapid physiological processes require fast imaging to capture the relevant information. But the scan time is also important for the patient throughput, and to manage costs for the equipment and its maintenance. Furthermore, with more time it is possible to collect k-space samples from a larger area to get a higher resolution, or to collect several images and perform averaging to increase the SNR. On the other hand, with an increased voxel size, that is, lower resolution, more spins will contribute to the signal in each voxel, and hence, the SNR will increase.

As mentioned in Section 2.2.3, the scan parameters will have a major impact on the signal. In general, a shorter TE will improve the SNR, as the signal has less time to decay. Also a larger flip angle will lead to

a higher SNR if the signal is measured in the xy-plane. For example, the rapid T_2^* decay of the gradient-echo signal results in a lower SNR compared to the spin echo. Moreover, using a short TR leaves little time for T_1 recovery, which may saturate the signal, resulting in a lower SNR when large flip angles are used. By applying spoiler gradients, giving a so called *spoiled gradient echo*, the saturation problems can be avoided, which restores the SNR.

Using a higher field strength increases the SNR, as the excited magnetization, and hence the observed signal, is larger. However, it also comes with a few downsides, as is discussed in the next section.

The MR scanner can use different coils depending on the application, and therefore, the choice of coil can effect the SNR without increasing the scan time. Ideally, the coil should be placed close to the imaged area to obtain a high SNR, and there are several coils available to achieve this, for example, volume coils or surface coils. Other options include coil arrays, where several coils together collect data to image a certain area. Each coil is associated with a matrix that describes its spatial sensitivity, and therefore, using high-quality coils can increase the SNR further. Furthermore, it is possible to adjust the receiver-coil bandwidth, which corresponds to the range of frequencies captured during the readout gradient. A higher bandwidth means that more information can be collected in a single readout, which speeds up the acquisition. However, the thermal noise power in the coil is proportional to the bandwidth, meaning that increasing the bandwidth will also increase the noise level.

Image artifacts

Sequences with long scan times, like the standard spin echo, are sensitive to motion during the scan. Motion artifacts are a big problem in MRI, and can lead to severely degraded image quality. Even if the patient under study is keeping all limbs still during the scan, cardiac motion, as well as blood flow, would still be present. Also respiratory motion can be a problem, especially when the imaging is not fast enough to be performed in a breath hold. The resulting images can suffer from both blurring and ghosting, depending on which part of the k-space was collected during the motion. Sequences can be designed to be less sensitive to motion, mainly by shortening the acquisition time. The echo-planar sequence is an example of a fast sequence, which significantly reduces the risk of motion artifacts.

For 2D encoding, the in-band slice profile is never perfectly flat, meaning that the achieved flip angle will vary across the slice. This can lead to image artifacts, and particularly, it is a problem for qMRI, where a constant and known flip angle is often assumed to simplify the model and enable efficient parameter estimation. An additional problem is the potential leakage that can occur between slices. If the RF partially excites

neighboring slices, the resulting images will be distorted. To counter this, it is possible to use an inter-slice gap, which minimizes the cross talk, but also leads to loss of information as some parts of the subject are not sampled. Cross talk is particularly problematic for multi-slice sequences, where several neighboring slices can be intentionally excited.

Gibbs ringing is another common type of artifact, which manifests as oscillations in the image magnitude adjacent to abrupt intensity changes; for example between air and tissue. The ringing is most significant close to tissue boundaries, and decays when moving outwards. Collecting more high frequency samples is the only way to reduce this artifact, as it is intrinsic to the Fourier series for a signal containing a jump discontinuity. In fact, the ringing does not go to zero as the frequency approaches infinity, meaning that there is no way to fully eliminate the Gibbs phenomenon. However, the visual impact in images will be unnoticeable when a sufficient number of k-space samples are collected. Gibbs ringing is illustrated in Fig. 2.12, which shows a phantom image reconstructed with, and without, the high frequency k-space samples, as well as the difference between the two images, which was included to highlight the ringing pattern.

Apart from increasing Gibbs ringing, decreasing the resolution also increases the risk of partial volume effects, which occur when two or several tissue types are present in a single voxel. As a result, the acquired signal is described by several decay constants, which in turn can lead to image artifacts, or complicate parameter estimation.

The proton resonance frequency (PRF) depends on the external field, but also on the local molecular environment, which can shield the magnetic field to a varying extent. For example, protons bound to fat and water have slightly different resonance frequencies, called a chemical shift. This shift can be useful in imaging to separate components with different molecular structure, or to measure the temperature using MRI, as is done in Chapter 9, but can also cause artifacts due to misregistration in space along the readout direction, and signal cancellation. By increasing the magnitude of the readout gradient, and bandwidth of the receiver, the relative size of the chemical frequency shift can be made small compared to the k-space sampling interval, thus leading to a small errors in the spatial registration.

With higher field strengths, it becomes increasingly difficult to construct a homogeneous magnetic field, which can lead to artifacts. Furthermore, problems that arise in the interfaces between tissues with different magnetic susceptibility are also magnified in a higher field. Artifacts due to inhomogeneity of the static field are a major problem for some sequences. In SSFP, off-resonance excitation can lead to a significant loss in the signal magnitude, resulting in dark bands in some parts of the image, as is illustrated in Fig. 2.13. Moreover, as described

a)   b)   c)
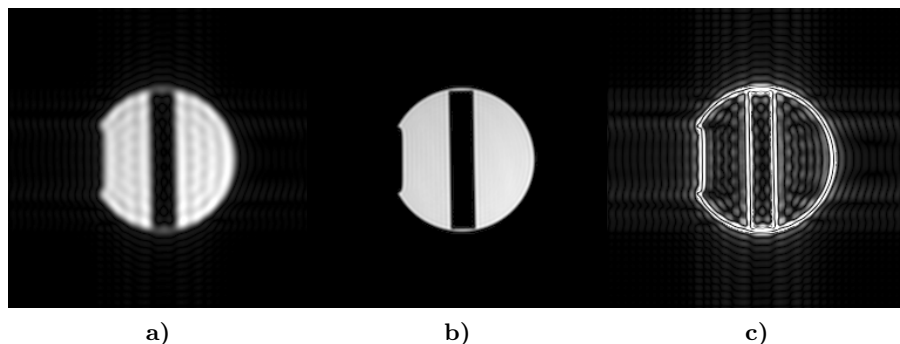
Figure 2.12: a) Phantom image showing ringing due to the Gibbs phenomenon, b) the same image with more high frequency k-space samples collected, and c) the difference between the two previous images, to highlight the ringing pattern.

by Faraday's law, a varying magnetic field will induce a current in a conductor, which in turn will generate a magnetic field. Therefore, small loop currents in the structure of the MR scanner, or *eddy currents*, cause errors in the applied fields. The effects of eddy currents are most significant at high field strengths, and when using sequences with rapidly changing magnetic gradients; and they can result in a wide range of artifacts ranging from blurring to spatial misregistration [1]. To counter these problems, it is possible to use both actively and passively shielded coils, or to compensate for the eddy currents in the pulse sequence [116]. These eddy currents may also occur within the subject which is a problem for patient safety as they can lead to tissue heating or involuntary nerve stimulation, see Section 2.4.2.

2.3 Data Modeling and Quantitative MRI

In qMRI, the goal is typically to estimate some physical quantity, such as T_1 or T_2, or to reduce image artifacts, given a set of images or k-space datasets [24]. Using the Bloch equations given by (2.3), it is possible to derive various closed form expressions for the resulting signals, depending on the pulse sequence used. These parametric models can in turn be fitted to data, to obtain estimates of the model parameters. By visualizing the estimates as a function of space, additional anatomic, chemical, physical, or functional information can be gained; alternatively, the estimates can be used to improve the collected images by eliminating the effects of the modeled artifacts. Signal processing, and particularly estimation theory, plays an important role in qMRI when
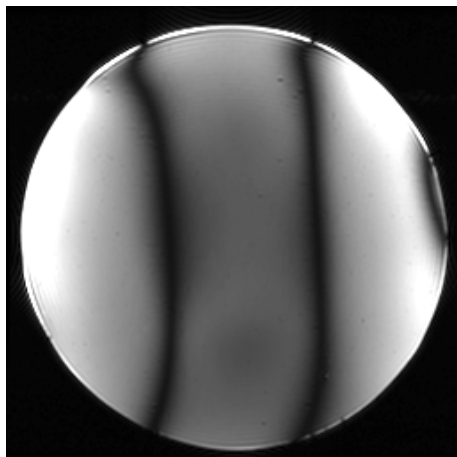
Figure 2.13: Phantom image showing banding artifacts due to inhomogeneities in the static B_0 field, when using a SSFP sequence.

trying to estimate the model parameters and their uncertainties, and more details on this topic are given in Section 3.1.

The model structure can vary dramatically depending on the pulse sequence and the parameters of interest, but typically the resulting estimation problem is nonlinear. To enable efficient optimization, the models are sometimes reduced or simplified, effectively making assumptions on the unknown parameters. This will introduce bias in the estimates, but this bias might be relatively small compared to the reduction in variance. For example, the flip angle set by the pulse sequence is often assumed to be achieved in the subject, an assumption that does not hold in practice. Depending on the model, this assumption can either have an insignificant effect on the fitting, or it can fully compromise the results. In practice, all models are approximate, as it is not possible to model everything; nor is it generally feasible to collect the data needed to accurately estimate all parameters in a complicated model. The challenge is to find a model that is *good enough* for the application at hand.

Performing the spin-echo experiment with different echo times provides samples of the T_2 decay curve in the image domain. The intensity at echo time t, for each voxel in the image, can according to (2.4) be described by the following signal model:

$$s(t) = \rho e^{-t/T_2}. \tag{2.13}$$

Using several datasets with different echo times, we can estimate the decay rate T_2 in every voxel to generate a T_2 map, a problem which is treated in Chapter 6. Similarly, the spin-echo inversion-recovery sequence results in the following voxelwise complex-valued model versus

the inversion time t [6]:

$$\tilde{s}(t) = a + be^{-t/T_1}, \tag{2.14}$$

where $a, b \in \mathbb{C}$, which can be used to estimate a T_1 map. Quantification of T_1 and T_2 gives valuable tissue information used in a wide range of applications [87]. For example, such relaxation maps can be used to detect tumors, which have been shown to possess longer T_1 decay times in general. Also, neurodegenerative diseases such as multiple sclerosis affect the overall relaxation times in the brain [119].

Other quantitative imaging techniques include diffusion MRI, functional MRI, magnetization transfer, and fat and water imaging. In diffusion MRI, the diffusion coefficient, or even the multi-dimensional diffusion tensor, is estimated from data [7]. This gives information regarding water mobility, which in turn can identify swelling as a result of a stroke. The goal in functional MRI is to detect neuron activation based on the local blood oxygen level. Often, linear models are used to detect the small metabolic variations that result from different stimuli [34]. In magnetization transfer, the exchange of magnetization between hydrogen bound to large molecules and more mobile hydrogen is modeled and estimated to reveal details of the chemical structure [68]. Finally, fat and water imaging is a technique where the chemical shift between fat and water-bound hydrogen is estimated to be able to separate the two types of tissue in post-processing [8].

Modeling with the aim of artifact reduction includes both motion correction and field mapping. By modeling the effects of motion on the k-space samples, it is possible to correct for some of the resulting artifacts by post-processing. Another alternative is to measure the motion and use this information to perform more advanced correction [33]. Field mapping, that is, estimating the achieved static and RF magnetic fields in space, does not directly reveal any information regarding the subject under study; however, having accurate estimates of the practical fields enables significant improvement of the images. The result can facilitate diagnosis, or be used to obtain higher accuracy when estimating tissue characteristics [113]. The images can be improved either by post-processing, taking the field knowledge into account, or by calibrating the fields prior to the imaging to obtain the desired flip angle and spatial registration. In Chapter 4, the problem of reducing field inhomogeneity artifacts in bSSFP images is treated.

Modeling the noise is also of interest to obtain accurate estimates. Mainly, the noise properties depend on the type of data used. The k-space data can de accurately modeled as independent and Gaussian distributed, and for linear reconstruction, these properties are transfered to the image domain. Using magnitude images, however, makes the

samples Rice distributed, which can complicate the estimation. This topic is discussed further in Section 3.1.1.

2.4 MRI scan

2.4.1 Hardware

The hardware needed to perform an MRI scan is technically advanced, and expensive to both purchase and run. The scanner mainly consists of a large magnet, which together with smaller shim coils is used to create a homogeneous static field B_0. Commonly, the magnets used are liquid-helium-cooled superconducting electromagnets, which provide high field strengths and good stability. Other parts of the scanner include gradient coils to accurately alter the field strength in space for encoding, transmitter coils to generate the RF excitation, receiver coils to measure the magnetic resonance, amplifiers, a moving bed, and computers for controlling the device and processing the collected signals. An example of a 1.5 T Philips scanner is shown in Fig. 2.14. Apart from the built in body coil, it is possible to choose from a set of different external coils, depending on the application. For example, surface coils, bird-cage head coils, or arrays of small coils can be used.

The field strength is typically in the range of 1.5 T to 7 T, but magnets over 20 T have been evaluated in research and animal studies [107]. The switching performance of the gradient coils is also limited for a given scanner, effectively limiting the type of sequences that can be used.

The scanner is typically kept in a Faraday-cage-like room, to avoid leakage of the generated electromagnetic fields, and to provide shielding from external disturbances. However, the signals generated by the scanner are relatively weak and do not cause much interference.

2.4.2 Safety issues

MRI does not use ionizing radiation, and there are no known side effects of being scanned. However, there are a few limitation due to patient safety. For example, the use varying magnetic fields can lead to tissue heating or involuntary nerve stimulation, due to the induced eddy currents. This particularly applies to fast sequences and steady state imaging, where the gradients are switched at a high rate. The heating and nerve stimulation effects have to be taken into account when designing a clinical acquisition protocol, to make sure the set specific absorption rate (SAR) limits are met, where SAR measures the rate at which energy is absorbed by the human body when exposed to an RF field.

Figure 2.14: A 1.5 T Philips Ingenia MRI scanner. Image courtesy of Philips Healthcare.

Another risk involve ferromagnetic metal objects that are accidentally brought into the vicinity of the scanner. The strong magnetic field can rapidly accelerate rather large objects to high velocities, potentially putting patients, or anyone in its path, at risk. Different metallic implants and other foreign metal objects can also pose problems. Particularly, electronic devices such as pacemakers, insulin pumps, and neurostimulation systems, can cause a health hazard during the MRI exam, if the radiologist is left unaware. However, most orthopedic and neurological implants are not ferromagnetic, and therefore do not pose any danger to the patient, although they might cause susceptibility artifacts in the resulting images [132].

The scanner coil is typically made to be rather narrow to maintain a homogeneous field, and therefore, the physical space inside the scanner is limited. Moreover, in the presence of strong magnetic field gradients the coils try to move against the scanner structure, resulting in loud vibrations that make hearing protection essential. As a consequence, claustrophobia or other types of anxiety symptoms, can be a risk during the scan. There are both wide bore and open scanners that can alleviate these problems, while enabling larger patients to be scanned, but these are relatively uncommon.

The cryogenic liquid used to cool the superconducting electromagnet is not in itself toxic; however, in the event of a involuntary shut down of the magnet the helium will start to boil, and in the absence of proper ventilation, there is a risk for asphyxiation. Although shut downs of this kind are uncommon, extra safety measures are usually in place, such as oxygen monitors, pressure valves, and fans, to make sure the helium gas is efficiently evacuated.

Contrast agents are sometimes administered prior to the scan, to efficiently reduce the T_1 relaxation time of the affected tissues. Unlike when performing an x-ray, the Gadolinium-based contrast liquids typically used do not contain iodine, and comes with a lower risk for allergic reactions. However, all contrast agents are, to some extent, toxic, and both intravenous and oral administration comes with some risk of side effects [48].

Chapter 3

Information processing

3.1 Signal processing

In the following, some background to the algorithms developed in this thesis is given. Signal processing is a broad subject, and the scope is by no means to give a full description of the topic, but only to highlight some signal processing concepts that are particularly useful in MRI. For more details on the topic, see for example [72, 129, 105].

3.1.1 Parameter estimation

Estimation theory is a sub-field of signal processing to which parametric methods and parameter estimation belong. The goal is to find the numerical values of the parameters based on some model of the observations. The solution is typically obtained as the optimal value of some optimization problem, but a wide range of heuristic approaches also exist, both application specific and more general in scope. The most common approach is the least squares (LS) method, where the parameters giving the smallest sum of squared model errors are chosen, that is

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \ \sum_{n=1}^{N} |y(t_n) - g(t_n, \boldsymbol{\theta})|^2 \,, \tag{3.1}$$

where $y(t)$ is the data, and we have a measurement model given by

$$y(t_n) = g(t_n, \boldsymbol{\theta}) + v(t_n), \tag{3.2}$$

where $g(t_n, \boldsymbol{\theta})$ is the model parameterized by the vector $\boldsymbol{\theta}$, and $v(t_n)$ is the measurement noise.

LS is actually the maximum likelihood (ML) estimator if the noise is independent, identically distributed (i.i.d.), and Gaussian. More gener-

ally, the ML problem can be formulated as

$$\hat{\boldsymbol{\theta}}_{\mathrm{ML}} = \underset{\boldsymbol{\theta}}{\arg\max} \ \mathcal{L}(\mathbf{y}, \boldsymbol{\theta}), \tag{3.3}$$

where $\mathcal{L}$ is the likelihood function, that is, the joint probability density of all observations $\mathbf{y}$, given parameters $\boldsymbol{\theta}$. Often, the more convenient log-likelihood is used, and assuming that the measurements are i.i.d., we can write

$$\ln \mathcal{L}(\mathbf{y}, \boldsymbol{\theta}) = \sum_{n=1}^{N} \ln p(y_n | \boldsymbol{\theta}), \tag{3.4}$$

where $p(y_n | \boldsymbol{\theta})$ is the probability density function for each y_n, conditioned on $\boldsymbol{\theta}$, and $\ln(\cdot)$ is the natural logarithm. The ML estimator possesses several important asymptotic properties that hold when the number of samples tends to infinity, namely: consistency, normality, and efficiency. Consistency means that with a sufficiently large number of samples, it is possible to find the true parameter values with arbitrary precision. Furthermore, under some regularity conditions, the errors will in the limit be normally (Gaussian) distributed. Finally, the estimator achieves the Cramér-Rao bound (see Section 3.1.1), which means that there is no other consistent estimator with a lower asymptotic mean squared error (MSE).

The ML problem consists of finding the parameter vector $\boldsymbol{\theta}$ that maximizes the likelihood of the measurements $\mathbf{y}$. There is no closed-form solution to this problem in general, although in specific scenarios, it might be possible to find one. An example of this is the linear Gaussian case, which reduces to linear LS. In matrix form, we can write a linear LS problem as

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\arg\min} \ \|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|^2, \tag{3.5}$$

where $\mathbf{A}$ is a matrix of known regressors, and the problem has a closed form solution given by

$$\hat{\boldsymbol{\theta}} = (\mathbf{A}^* \mathbf{A})^{-1} \mathbf{A}^* \mathbf{y}. \tag{3.6}$$

For a nonlinear model, however, numerical methods of optimization are needed to minimize the nonlinear LS (NLS) criterion. The same applies to the more general ML estimator, which provides a great deal of flexibility in the problem formulation. Regrettably, using this flexibility can lead to a rather involved non-convex criterion function, which makes finding the optimal solution increasingly problematic. Because of this, approximate and suboptimal solutions to the general ML problem in (3.3) are often used, for example, by applying the simple linear LS method to non-Gaussian data.

Generalized LS (GLS) can by used to accommodate a wider class of distributions, with correlation between the observations and varying noise levels. Assuming that the regressors are independent of the noise, the criterion involves a weighting matrix given by the inverse of the noise covariance matrix, that is

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\text{argmin}} \ \|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|_{\mathbf{W}}^2 \, , \tag{3.7}$$

where

$$\mathbf{W} = \text{E}\Big\{ (\mathbf{v} - \text{E}\{\mathbf{v}\}) \, (\mathbf{v} - \text{E}\{\mathbf{v}\})^{\text{T}} \Big\}^{-1} \, , \tag{3.8}$$

and $\text{E}\{\cdot\}$ denotes the expected value. The weighting matrix is effectively whitening the errors, so that the Gauss-Markov theorem applies. Hence, the GLS estimator is the best linear unbiased estimator (BLUE), and when the noise is Gaussian, it is ML. In practice, the covariance matrix of the noise is unknown, and needs to be estimated or approximated using prior information. This can, for example, be achieved by starting from ordinary LS, and using the resulting residuals to find an estimate of the covariance matrix. This estimate can then be inserted into (3.7) to generate new estimates and the corresponding covariance matrix, enabling an iterative refinement process. Note, however, that the method requires some assumptions on the noise properties to enable estimation of the covariance matrix. Clearly, a full matrix of independent elements cannot be estimated from a single measurement vector $\mathbf{y}$.

Further variants of LS include additional terms in the criterion to regularize the problem, such as, Tikhonov regularization, where a L_2-norm penalty is used. For the linear case, the problem can be expressed as

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\text{argmin}} \ \|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|^2 + \|\mathbf{R}\boldsymbol{\theta}\|^2 \, , \tag{3.9}$$

where $\mathbf{R} \in \mathbb{R}^{P \times M}$ is a regularization matrix. For any $\mathbf{R}$ with full column rank, there is a closed form solution to the problem, which is given by

$$\hat{\boldsymbol{\theta}} = (\mathbf{A}^*\mathbf{A} + \mathbf{R}^*\mathbf{R})^{-1}\mathbf{A}^*\mathbf{y}. \tag{3.10}$$

This means that (3.9) will have a unique solution, even if (3.5) is underdetermined.

The regularization is effectively imposing some prior information regarding the model parameters, for example, that parameter vector is small in norm (for which $\mathbf{R} = \lambda\mathbf{I}$, and $\lambda > 0$). For more general types of regularization, such as penalties based on the L_1-norm or nonlinear functions, no closed form solution exists, and more details on this topic are to be found in Section 3.1.2.

The Cramér-Rao bound

The Cramér-Rao bound (CRB) is a lower bound on the covariance matrix of the parameter estimates, for any unbiased estimator, given the model [72, 122]. It captures both how the sampling times affect the estimation errors, as well as how the errors correlate, and relates to the amount of information regarding the parameters that is expected in the data. It should be noted that the CRB is a lower bound that might not be achievable; however, if a specific unbiased estimation algorithm achieves the bound, it is statistically efficient.

The Fisher information matrix (FIM) provides the sought information measure, and in general, it can be derived from the formula

$$\mathbf{F} = \mathrm{E}\left\{ \left(\frac{\partial \ln(\mathcal{L}(\mathbf{y}, \boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} \right)^{\mathrm{T}} \left(\frac{\partial \ln(\mathcal{L}(\mathbf{y}, \boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} \right) \right\}, \qquad (3.11)$$

where again, $\mathcal{L}(\mathbf{y}, \boldsymbol{\theta})$ is the likelihood function. Taking the inverse of the FIM gives the CRB matrix

$$\mathbf{C}_{\mathrm{CRB}}(\boldsymbol{\theta}) = \mathbf{F}^{-1}(\boldsymbol{\theta}), \qquad (3.12)$$

which provides a bound on the covariance matrix of the parameters:

$$\mathrm{cov}\{\hat{\boldsymbol{\theta}}\} \geq \mathbf{C}_{\mathrm{CRB}}(\boldsymbol{\theta}). \qquad (3.13)$$

Specifically, the j:th diagonal element of $\mathbf{C}_{\mathrm{CRB}}$, c_{jj}, gives bounds on the mean square error (MSE) of the corresponding estimate $\hat{\theta}_j$, assuming it is unbiased. This provides a benchmark for the statistical performance of various estimators. Geometrically, the CRB matrix defines an uncertainty ellipsoid in the parameter space to which $\boldsymbol{\theta}$ belongs, and the diagonal elements correspond to uncertainty projected onto each parameter axis.

Under the assumption of zero-mean i.i.d. circular complex-Gaussian noise of variance σ^2, the FIM is given by the Slepian-Bangs formula [122]:

$$\mathbf{F}_{\mathrm{Gauss}}(\boldsymbol{\theta}) = \frac{2}{\sigma^2} \mathrm{Re}\left\{ \left(\frac{\partial \mathbf{g}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^{*} \left(\frac{\partial \mathbf{g}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \right\}, \qquad (3.14)$$

where $\mathbf{g}(\boldsymbol{\theta}) = [g(\boldsymbol{\theta}, t_1), \ldots, g(\boldsymbol{\theta}, t_N)]^{\mathrm{T}}$ is the signal model vectorized over time, and $\partial \mathbf{g}/\partial \boldsymbol{\theta}$ is the Jacobian matrix of the vector-valued function $\mathbf{g}(\boldsymbol{\theta})$, given by

$$\frac{\partial \mathbf{g}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{\partial g_1}{\partial \theta_1} & \cdots & \frac{\partial g_1}{\partial \theta_M} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_N}{\partial \theta_1} & \cdots & \frac{\partial g_N}{\partial \theta_M} \end{bmatrix}. \qquad (3.15)$$

In the real-valued case, this reduces to

$$\mathbf{F}_{\mathrm{Gauss}}(\boldsymbol{\theta}) = \frac{1}{\sigma^2} \left(\frac{\partial \mathbf{g}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^{\mathrm{T}} \left(\frac{\partial \mathbf{g}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right). \tag{3.16}$$

Both (3.14) and (3.16) are easily computed and straightforward to use, while computing the FIM in a more general case can be difficult.

Apart from determining the statistical efficiency of different parameter estimation algorithms, the CRB is useful in experiment design. It is possible to minimize the CRB for a given set of parameters with respect to the sampling times, or other choices made during the data collection, to enable improved results. In MRI, this can be used to optimize pulse sequences by appropriately choosing *TE*, *TR*, and α, based on the CRB, or more fundamentally, to ensure identifiability of the model parameters for a given acquisition protocol. This topic is discussed further in Section 3.1.3.

Noise distributions
As mentioned, the noise in complex-valued MRI data is accurately approximated by a zero-mean i.i.d. circular complex-Gaussian, which means that each sample can be modeled as

$$y = g(\boldsymbol{\theta}) + v(\sigma), \tag{3.17}$$

where $g(\boldsymbol{\theta})$ is the signal (or expected value of y) which parameterized by $\boldsymbol{\theta}$, and $v(\sigma)$ is the independent noise which is described by the PDF [95]

$$p_{\mathrm{G}}(z|\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{|z|^2}{2\sigma^2}}, \tag{3.18}$$

for some complex-valued variable z. Given the model in (3.17), the ML estimator of $\boldsymbol{\theta}$ can easily be derived, and is given by LS. However, magnitude images are often used in practice, in which case each sample can be modeled as an observation, y, of a Rice-distributed stochastic variable S_{R} parameterized by $|g(\boldsymbol{\theta})|$ and σ [26, 112], that is

$$y = Y_{\mathrm{R}} \sim \mathrm{Rice}(|g(\boldsymbol{\theta})|, \sigma). \tag{3.19}$$

The Rician PDF is given by

$$p_{\mathrm{R}}(x \,|\, \eta, \sigma) = \frac{x}{\sigma^2} e^{\frac{-(x^2 + \eta^2)}{2\sigma^2}} I_0\left(\frac{x\eta}{\sigma^2} \right), \tag{3.20}$$

where I_0 is the modified Bessel function of the first kind and order zero. For integers n, these Bessel functions can be defined by

$$I_n(z) = \frac{1}{\pi} \int_0^\pi e^{z\cos(\xi)} \cos(n\xi) \mathrm{d}\xi. \tag{3.21}$$
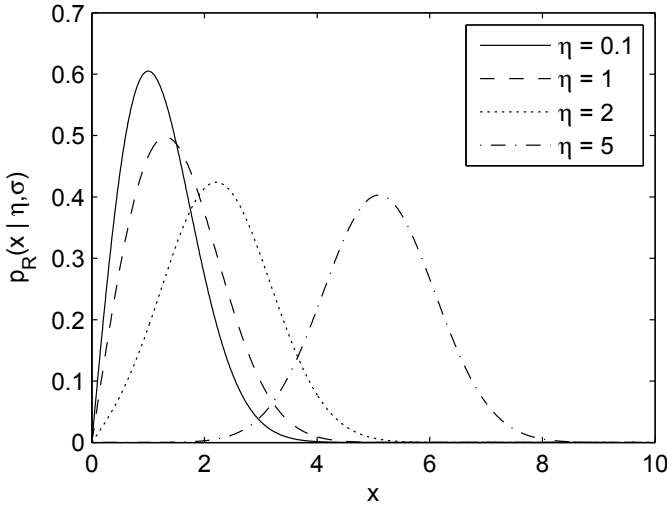
Figure 3.1: Examples of the Rice distribution for different values of η, and a fixed $\sigma = 1$.

For the complex-valued Gaussian model in (3.17), the magnitude would be distributed according to (3.20), where η corresponds to the magnitude of the signal component, $|g(\boldsymbol{\theta})|$, and σ corresponds to the standard deviation of the noise in both the real and imaginary part. The PDF in (3.20) results from projecting the mass of (3.18) to the positive real axis along concentric circles, as these circles correspond to the same magnitude. A few examples of the Rician distribution for different η and $\sigma = 1$, are shown in Fig. 3.1. As can be seen, the distribution is significantly different from the Gaussian when $\eta \leq \sigma$.

The variance of a Rice-distributed variable S_R is given by the expression [110]

$$\mathrm{E}\Big\{(S_R - \mathrm{E}\{S_R\})^2\Big\} = 2\sigma^2 + \eta^2 - \frac{\pi\sigma^2}{2}L_{1/2}^2\left(-\frac{\eta^2}{2\sigma^2}\right), \qquad (3.22)$$

while the mean is found to be

$$\mathrm{E}\{S_R\} = \sigma\sqrt{\frac{\pi}{2}}L_{1/2}\left(-\frac{\eta^2}{2\sigma^2}\right), \qquad (3.23)$$

where $L_{1/2}(\cdot)$ is a Laguerre function of order $1/2$, which can be expressed in terms of the modified Bessel functions according to

$$L_{1/2}(x) = e^{x/2}\left[(1-x)\,I_0\left(-\frac{x}{2}\right) - xI_1\left(-\frac{x}{2}\right)\right]. \qquad (3.24)$$

The noise is no longer independent and additive, as both the mean and the variance depend on $|g(\boldsymbol{\theta})|$. However, the sample model can be
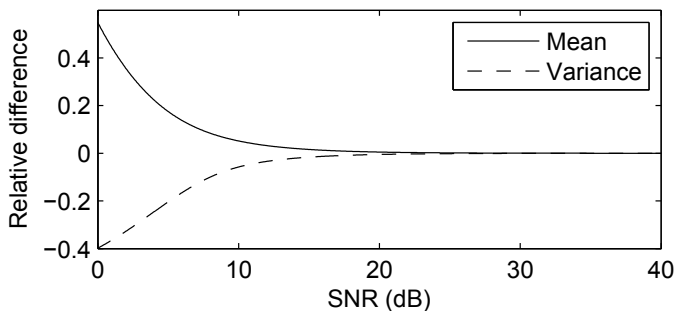
Figure 3.2: The relative difference between the mean and variance for the Rice distribution and the Gaussian distribution, versus the SNR.

written in the same form as (3.17):

$$|y| = |g(\boldsymbol{\theta})| + \tilde{v}(|g(\boldsymbol{\theta})|, \sigma), \qquad (3.25)$$

where the noise term $\tilde{v}$ now depends on the signal magnitude. The Rician noise complicates the model, which makes estimation of the parameters $\boldsymbol{\theta}$ more difficult. However, as the SNR increases, that is, $|g(\boldsymbol{\theta})|$ becomes large compared to σ, the Rician mean and variance converge towards their Gaussian counterparts, as is shown in Fig. 3.2. This reduces the need to model the data as Rice distributed in high SNR applications, and enables approximate ML estimation using LS.

3.1.2 Optimization

In numerical optimization, the goal is to select the best parameters values with respect to a given objective, subject to some constraints. Usually, the problem is formulated as a minimization or maximization of some function of the parameters, the *optimization criterion*. Depending on the structure of this function, the problem can be convex (or concave for maximization), meaning that there is a unique minimum. A problem of this type can typically be efficiently solved. For example, the linear LS problem is convex, and in the unconstrained case there is a closed form solution, which was given in (3.6). Non-convex problems, on the other hand, may have multiple stationary points and many local optima, meaning that there is typically no closed form solution, nor is it generally possible to find a numerical minimization method that obtains the optimal value within a reasonable time frame. For example, exhaustive methods can always find the global optimum in theory, but as the computational complexity grows rapidly with the number of parameters, such approaches are commonly intractable in practice.

   The optimization problems that arise in qMRI are often non-convex. In specific cases, it might be possible to find the global optimum with

high probability, but many times a proof of optimality cannot be given. There are a few ways to sidestep this problem, for example by relaxing the criterion or the constraints and solving a closely related convex problem, or by designing an efficient approximation algorithm that is essentially closed form, that is, does not rely on non-convex optimization. In Chapter 5, an approximate algorithm is used to estimate the parameters of multiple exponential decays, which is a non-convex problem.

Even when the problem to be solved is convex, the vast amount of data usually obtained in MRI can be an issue. To make an algorithm useful in a practical MR setting it must be fast to execute, preferably without requiring special hardware.

Convex problems

The commonly used linear LS approach with linear constraints leads to a convex quadratic program (QP), which can be formulated as

$$
\begin{aligned}
\underset{\mathbf{x}}{\text{minimize}} \quad & \frac{1}{2}\mathbf{x}^{\text{T}}\mathbf{Q}\mathbf{x} + \mathbf{c}^{\text{T}}\mathbf{x} \\
\text{subject to} \quad & \mathbf{B}\mathbf{x} \leq \mathbf{b}
\end{aligned}
, \tag{3.26}
$$

where $\mathbf{Q} \in \mathbb{R}^{M \times M}$ is a positive definite symmetric matrix, $\mathbf{B} \in \mathbb{R}^{P \times M}$ is the constraint matrix, and $\mathbf{c} \in \mathbb{R}^{M \times 1}$, $\mathbf{b} \in \mathbb{R}^{P \times 1}$ are vectors. The matrix $\mathbf{Q}$ and the vector $\mathbf{c}$ are easily obtained from (3.5), by expanding the criterion:

$$
\|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|^2 = \|\mathbf{y}\|^2 + \boldsymbol{\theta}^{\text{T}}\mathbf{A}^{\text{T}}\mathbf{A}\boldsymbol{\theta} - 2\mathbf{y}^{\text{T}}\mathbf{A}\boldsymbol{\theta}, \tag{3.27}
$$

where the problem has been assumed to be real-valued. By identifying $\boldsymbol{\theta}$ with $\mathbf{x}$, and omitting $\|\mathbf{y}\|^2$ since it is a positive constant independent of the optimization variable, we obtain $\mathbf{Q} = \mathbf{A}^{\text{T}}\mathbf{A}$ and $\mathbf{c} = \mathbf{A}^{\text{T}}\mathbf{y}$.

Even though QP solvers are relatively efficient, finding the optimal value can be quite time consuming when the problem size is in the order of 100,000 parameters. Such large problems are not uncommon in qMRI, for example, when simultaneously estimating several parameters per voxel in a 256×256 image. By replacing the L_2 norm used in the LS criterion with an L_1 norm, the problem can be cast in the form of a linear program:

$$
\begin{aligned}
\underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{d}^{\text{T}}\mathbf{x} \\
\text{subject to} \quad & \mathbf{C}\mathbf{x} \leq \mathbf{h}
\end{aligned}
, \tag{3.28}
$$

which typically supports larger problem sizes within a fixed set of hardware and time frame constraints. Using an L_1-norm fitting term is common in applications with outliers, as large residuals are less penalized compared to when using a squared term. In fact, L_1-norm fitting is

the ML estimator in the case of Laplacian noise [23]. However, as mentioned in Section 3.1.1, the LS method is often used even when the noise is not Gaussian, due to its simplicity. Therefore, the L_1 norm could be another a sub-optimal alternative, used to reduce the computation time for large estimation problems.

Many regularized versions of linear L_1 and L_2 minimization problems are convex, for example, the Tikhonov regularization in (3.9). In particular, L_1-norm regularization can be used to obtain sparse estimates, that is, a parameter vector containing only a few nonzero elements, as will be discussed next.

Sparse methods

The original problem in sparse estimation is to find a vector $\boldsymbol{\theta}$ with at most J nonzero elements, which minimizes an LS fitting criterion, that is

$$
\begin{aligned}
&\underset{\mathbf{x}}{\text{minimize}} && \|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|^2 \\
&\text{subject to} && \|\mathbf{R}\boldsymbol{\theta}\|_0 \leq J
\end{aligned}, \tag{3.29}
$$

for some regularization matrix $\mathbf{R}$. However, the L_0 pseudo-norm makes this problem non-convex. By convex relaxation, we replace the L_0 norm by the L_1 norm, which can be shown to induce sparsity, and formulate the regularized problem:

$$
\underset{\boldsymbol{\theta}}{\text{minimize}} \ \|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|^2 + \lambda\|\mathbf{R}\boldsymbol{\theta}\|_1, \tag{3.30}
$$

which can be solved efficiently. The main problem of this so called LASSO (least absolute shrinkage and selection operator) approach is that the λ has to be chosen somehow. There are several methods for this available in the literature, for example, the cross-validation approach [64]. But there are also user-parameter free sparse estimation algorithms, such as SPICE (sparse iterative covariance-based estimation). Actually, SPICE has been proven to be identical to square-root LASSO with $\lambda = 1$, that is, to solve the problem [3]:

$$
\underset{\mathbf{r}}{\text{minimize}} \ \|\mathbf{y} - \mathbf{A}\mathbf{r}\| + \|\mathbf{r}\|_1. \tag{3.31}
$$

Sparse methods for solving problems with both linear and nonlinear parameters are particularly interesting. A simple example of the technique is given in the following, where a sinusoidal model of the form

$$
y(t) = \sum_{m=1}^{M} r_m \sin(\omega_m t) + v(t), \tag{3.32}
$$

where $v(t)$ is a Gaussian noise, is assumed. Both the number of sinusoids M, and the parameters $\{r_m, \omega_m\}_{m=1}^{M}$ are unknown, and need to be

estimated. Although simple in structure, the LS problem

$$\underset{\{r_m,\omega_m\}_{m=1}^M}{\text{minimize}} \sum_{t=1}^{N} \left| y(t) - \sum_{m=1}^{M} r_m \sin(\omega_m t) \right|^2 \qquad (3.33)$$

is highly nonlinear and multimodal, making it a difficult task to find the optimum solution by a general iterative minimization method. Moreover, the problem would have to be solved for each M, and a choice between the resulting candidate solutions would have to be made. With the sparse approach, we assume an interval of potential ω-values and sample it on a grid in K points, giving a set of frequencies $\{\tilde{\omega}_k\}_{k=1}^K$. The corresponding sinusoids provide us with a basis, and the remaining problem in vector form, becomes

$$\underset{\mathbf{r}}{\text{minimize}} \ \|\mathbf{y} - \mathbf{A}\mathbf{r}\|^2, \qquad (3.34)$$

where $\mathbf{r} \in \mathbb{R}^{K \times 1}$ is the vector of amplitudes with $K \gg M$, and $\mathbf{A}$ can be written as

$$\mathbf{A} = \begin{bmatrix} \sin(\tilde{\omega}_1) & \cdots & \sin(\tilde{\omega}_K) \\ \sin(2\tilde{\omega}_1) & \cdots & \sin(2\tilde{\omega}_K) \\ \vdots & & \vdots \\ \sin(N\tilde{\omega}_1) & \cdots & \sin(N\tilde{\omega}_K) \end{bmatrix}. \qquad (3.35)$$

The problem is often underdetermined, meaning that $K > N$, giving an infinite number of solutions to (3.34). However, even in the overdetermined case, for which there is a unique solution to the problem, the LS-estimated vector $\hat{\mathbf{r}}$ is unlikely to be sparse in the presence of noise. By solving a regularized problem similar to (3.30):

$$\underset{\mathbf{r}}{\text{minimize}} \ \|\mathbf{y} - \mathbf{A}\mathbf{r}\|^2 + \lambda\|\mathbf{r}\|_1, \qquad (3.36)$$

where λ is suitably chosen, sparsity will be promoted, giving an estimate $\hat{\mathbf{r}}$ with relatively few nonzero elements. These elements correspond to a set of $\hat{\omega}$:s, and the number of nonzeroes provides an estimate of M. The above approach can be used for any model that has a mixture of linear and a few nonlinear parameters. However, when there are several nonlinear parameters to be gridded, the matrix dimensions grow rapidly, often making the columns of $\mathbf{A}$ significantly correlated.

Optimization algorithms

Large-scale optimization problems are often solved by efficient interior-point methods, although other alternatives are also available, such as the simplex method for LPs, gradient methods, and variants of the augmented Lagrangian method [28]. The interior-point methods often

use a barrier function to enforce the constrains, cf. [73]. Given a general optimization problem of the form:

$$
\begin{aligned}
\underset{\mathbf{x}}{\text{minimize}} \quad & f(\mathbf{x}) \\
\text{subject to} \quad & c_p(\mathbf{x}) \geq 0 \\
& p = 1, \ldots, P
\end{aligned}
\qquad , \qquad (3.37)
$$

we can define the barrier function as

$$
B(\mathbf{x}, \nu) = f(\mathbf{x}) - \nu \sum_{p=1}^{P} \ln(c_p(\mathbf{x})). \qquad (3.38)
$$

For $\nu > 0$, any active constraint would lead to an infinite value of B, and hence the minimum of the barrier function fulfills the constraints. If we let ν converge to zero, the minimum of B should approach a solution of (3.37). An iterative minimization, for example, based on Newton's method, is performed to find the optimum value. In general, the Newton method tries to find a stationary point of a function f by iterating the following equation:

$$
\mathbf{x}_{k+1} = \mathbf{x}_k - \mu \left[\mathcal{H} f(\mathbf{x}_k) \right]^{-1} \nabla f(\mathbf{x}_k), \qquad (3.39)
$$

where $\mathcal{H}$ is the Hessian operator (the Hessian of $f(\mathbf{x}_k)$ is assumed to exist), ∇ is the gradient operator, $\mu > 0$ is the step size, and $k \geq 0$ is the iteration number. The method can easily be derived from the second-order Taylor expansion of f around $\mathbf{x}_k$:

$$
f(\mathbf{x}_k + \Delta \mathbf{x}) \approx f(\mathbf{x}_k) + \Delta \mathbf{x}^{\mathrm{T}} \nabla f(\mathbf{x}_k) + \frac{1}{2} \Delta \mathbf{x}^{\mathrm{T}} \mathcal{H} f(\mathbf{x}_k) \Delta \mathbf{x}, \qquad (3.40)
$$

which after differentiation with respect to $\Delta \mathbf{x}$, and setting the derivative to zero, gives

$$
\nabla f(\mathbf{x}_k) + \mathcal{H} f(\mathbf{x}_k) \Delta \mathbf{x} = 0. \qquad (3.41)
$$

By rearranging and defining the next iterate as $\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta \mathbf{x}$, we arrive at (3.39) for $\mu = 1$. Effectively, Newton's method approximates the criterion surface with a quadratic function around $\mathbf{x}_k$, and takes the minimum of this quadratic as the next estimate $\mathbf{x}_{k+1}$. The iterations can be shown to converge to a stationary point of the criterion [28].

When the iterates are sufficiently close to the optimum, Newton's method converges quadratically. This is typically much faster than the simple gradient descent method, where each iteration makes a step along the negative gradient of the function in the current point. This can be explained by the occurrence of second-order derivative, in the form of the Hessian, in Newton's method, making it a second-order method, while gradient descent is a first-order method. It can be difficult to obtain

an analytical expression for the Hessian, and numerical methods can be too expensive computationally, in which case approximate quasi-Newton methods can be applied. These methods use the gradient to estimate the Hessian matrix, or its inverse, and can provide faster convergence in terms of computation time, compared to the full Newton method, while outperforming the gradient descent.

In practice, the explicit inverse of the Hessian in (3.39) is never computed, but rather a system of linear equations in solved. If the Hessian is close to singular, problems with convergence can arise. An approach to alleviate this problem is to modify the Hessian and make it nonsingular. For example, the Levenberg-Marquardt algorithm uses an approximate Hessian, and adds a diagonal loading $\rho\mathbf{I}$ to stabilize the matrix. By adjusting the value of ρ during the iterations, the algorithm effectively alternates between Newton and gradient descent. This can be useful to stabilize the standard Newton method when the Hessian is noninformative.

Initialization

In nonlinear and multi-modal optimization, finding the *global* optimum is the main problem. Given a smooth criterion surface, an iterative minimization approach will typically converge to a local minimum close to the starting point. To ensure that the found local minimum is also global, an initialization point in the attraction domain of the global minimum needs to be supplied to the minimization algorithm. Either, an initial guess of the parameters can be obtained from knowledge of the problem, or other prior information, or some means of estimating the initial values is needed. By initializing a nonlinear minimization using an approximate, convex or closed form, solution, it is often possible to get close to the optimal parameters, although it can be hard to prove that they are in fact optimal in a practical situation. A few common techniques for this are: linearization, where a linear LS problem is solved to find an approximation; subspace methods, where the structure of the problem leads to a matrix with specific eigenvalues; basis expansion, where the nonlinearity is linearly parameterized in terms of a set of nonlinear functions; and the sparse methods previously mentioned.

3.1.3 Input and experiment design

This topic concerns the design of input signals to obtain the most informative output. Many applications, such as communication and radar, use signal design to find sequences with good correlation properties, but it is also of great interest in system identification where design criteria based on the MSE of the parameters can be used.

In MRI, one problem is to design signals with specific spectral profiles, providing excitation of a specific slice in space while minimizing artifacts; another is to optimize the sequence parameters, for example, to maximize the SNR.

There are several physical limitations to RF design, for example, the hardware used to transmit the signal. The design problem is usually cast as an optimization problem. Often, this problem is nonlinear, and one can only hope to find a local optimum. However, since the local optima typically correspond to a different signals with similar properties, it is possible to obtain a set of candidate designs by repeatedly solving the problem with different initializations. Then, the best signal can be chosen, for example, based on the criterion value, or best in some other non-optimized respect, given that there are several signals with sufficiently low criterion values.

For qMRI, the first step is to establish identifiability, that is, to ensure that a given set of images is sufficient to uniquely estimate the parameters of the signal model. This can usually be done by analyzing the structure of the problem, or by examining the rank of the CRB matrix. A separate problem is to optimize the pulse sequence parameters, such as TE, TR, and α, with the aim of minimizing the variance of the estimated parameters. Given a model, as well as the expected set of parameters, one can compute the CRB as a function of TE, TR, and α, which will provide a hyper-surface that bounds the optimal estimation performance. Minimizing the bound enables improvement of the results, although the actual performance depends on the estimation algorithm used, and whether or not it can achieve the bound in practice.

3.2 Image processing

This section presents a few concepts from image processing that have been useful when developing the estimation algorithms presented in this thesis. For a more general description of the topic, see for example [55].

3.2.1 Image filtering

One of the first image processing examples that usually comes to mind is image denoising. This is typically performed by filtering or smoothing the image across the spatial dimensions, but denoising can also be performed in the frequency domain. The main problem in MRI is that the contrast between tissues, and the fine details of the image, need to be largely preserved during the filtering.

Smoothing and sharpening

Smoothing is mainly used for denoising images. There are numerous ways of smoothing an image, the simplest being averaging, which is performed in some local region around each voxel of the image. As an extension, it is possible to perform a weighted average, giving a linear filter of the form

$$f_{k,p}(\mathbf{B}) = \sum_{m=-r}^{r} \sum_{n=-r}^{r} w_{m,n} b_{k+m,p+n}, \qquad (3.42)$$

where $f_{k,p}(\mathbf{B})$ is the filtered voxel of the image $\mathbf{B}$ with index k, p, the constant r is the size of the filter in space (e.g. $r = 1$ gives a 3×3 region), and $\{w_{m,n}\}$ is a set of weights that together sum to unity. The weights can be chosen in many ways, leading to different filters. One alternative is to obtain the weights from a sampled Gaussian kernel:

$$w_{m,n} = \frac{1}{2\pi\sigma^2} e^{-\frac{m^2+n^2}{2\sigma^2}}, \qquad (3.43)$$

for which the region r in (3.42) is theoretically infinite, and σ determines the spatial extent of the filter. In practice, the kernel can be truncated with good approximation, as the exponential function rapidly decays to zero. In addition to noise reduction, all linear filtering approaches lead to unwanted blurring of the images, to some extent.

Another common method is median filtering, which works in a manner similar to the averaging filter, but instead assigns the median value of the surrounding voxels to the center voxel. This filter is nonlinear and usually results in less blurring than the linear filter, for a given filter size. Particularly, impulse noise, or salt-and-pepper noise, which consists of sparse but large errors, can be efficiently reduced by this method. There are also more advanced nonlinear approaches to smoothing that attempt to preserve the edges in the image, such as the bilateral or guided filter [127, 66].

Smoothing is often performed by averaging, and sharpening can be obtained by the inverse operation, that is, instead of summing the data, it is differentiated. Differentiation emphasizes the sharp variations, such as edges in the image, but also the noise. This can be used for edge enhancement using both the first and second-order differences. The discrete five-point Laplacian, given by

$$\nabla^2 b_{k,p} = b_{k-1,p} + b_{k+1,p} + b_{k,p-1} + b_{k,p+1} - 4b_{k,p}, \qquad (3.44)$$

will produce an image highlighting rapid changes, and by adding the filtered image to the original image data, a new sharpened image is produced. Another alternative is to use a filter based on the magnitude

of the gradient, which is a nonlinear filter that achieves a result similar to that of the Laplacian filter [55].

As mentioned, the differentiation also amplifies the noise, making the sharpening process sensitive to disturbances. In MRI, the problem consists of preserving edge detail while minimizing the noise. Intuitively, somehow combining the properties of smoothing and sharpening could be fruitful in an attempt to achieve this goal. Total-variation-based methods is one class of algorithms that attempt to solve this problem, and they are discussed in more detail in the following.

Fourier domain filtering

Smoothing and sharpening can also be performed in the frequency domain, where the previously mentioned spatial filters would correspond to low-pass or high-pass filters in frequency. In MRI, artifacts due to errors in the data collection can often be identified and mitigated in the frequency domain. This follows from the fact that the data is collected in k-space, and sample outliers are easily distinguished from the expected frequency content. For example, periodic disturbances and ripple artifacts can often be eliminated by nulling the corresponding frequencies in k-space. The inverse Fourier transform then generates a filtered image where the disturbances are suppressed. The impact on the overall image quality is typically minor if the frequency is high enough, as the corresponding k-space samples are generally small. Close to the center of k-space, where the bulk of the image energy is located, it might be favorable to replace the erroneous data points rather than nulling them. This can be performed by averaging over some local region, or using the symmetry properties of the transform to mirror samples in k-space. In sum, simple post-processing of the k-space samples can provide a significant improvement of the data quality.

Total variation techniques

Total variation (TV) refers to estimation methods where the algorithms themselves are allowed to distribute a predefined amount of uncertainty based on a given criterion. By doing so, it is possible to obtain denoising of the resulting estimates without smoothing the tissue boundaries and image details. The main idea is that noisy measurements have a relatively high total variation, in some sense, compared to the noise-free signal. For example, one can formulate a regularized LS filtering problem in 1D as

$$\underset{\mathbf{x}}{\text{minimize}} \ \|\mathbf{y} - \mathbf{x}\|^2 + \lambda V(\mathbf{x}), \qquad (3.45)$$

where $V(\mathbf{x}) \geq 0$ is a measure of the variation, and λ determines the total amount of variation allowed. The problem can be equivalently

formulated as a constrained minimization where $V(\mathbf{x}) \leq \epsilon$, for some $\epsilon > 0$, which explicitly bounds $V(\mathbf{x})$. By finding a signal $\mathbf{x}$ that is close to $\mathbf{y}$ but has a lower $V(\mathbf{x})$, it is possible to denoise the measurements. It is straightforward to extend the method to two or more dimensions, which is useful in image processing.

More generally, LS TV typically leads to an optimization problem of the form

$$\underset{\mathbf{x}}{\text{minimize}} \; \|\mathbf{y} - \mathbf{g}(\mathbf{x})\|^2 + \lambda V(\mathbf{x}), \qquad (3.46)$$

where $\mathbf{g}(\mathbf{x})$ is a vector-valued function of the parameters $\mathbf{x}$. The criterion in (3.46) could, for example, represent a simple linear LS estimation problem, with smoothness imposed on the result. Using this formalism, it is possible to go beyond denoising, and perform image deconvolution, reconstruction, and inpainting [11, 22, 31].

A simple measure of the variation of a vector $\mathbf{x}$ that leads to a convex problem if $\mathbf{g}(\mathbf{x})$ is linear, is given by

$$V(\mathbf{x}) = \sum_{m=1}^{M-1} |x_{m+1} - x_m| = \|\mathbf{D}_1 \mathbf{x}\|_1, \qquad (3.47)$$

where $\mathbf{D}_1$ is the first-order difference matrix. The measure in (3.47) is of L_1-norm type, which enables denoising of MR images where the tissue boundaries are largely preserved. As mentioned, the L_1 norm will promote sparsity of the first-order differences, or equivalently, that the denoised image $\hat{\mathbf{x}}$ is piecewise constant, which is usually a good approximation. An L_2-norm regularization, on the other hand, would smooth over the edges, leaving an image with less detail. Total-variation denoising of T_2 estimates based on the L_1-norm is used in Chapter 6.

In practice, the choice of $V(\mathbf{x})$ depends on the prior knowledge of the problem; however, if the resulting optimization problem is non-convex, it can be difficult to solve. To get reliable performance, it is preferable to formulate a convex problem, as this gives a well defined result for a given λ that does not depend on the minimization procedure. Given a choice of $V(\mathbf{x})$, the problem of choosing λ (or equivalently ϵ) remains, and a suitable value for this parameter is essential to reduce the noise without corrupting the underlying information. This can, for example, be done based on an estimate of the noise level, or by using cross validation [54].

3.2.2 Phase unwrapping

Complex-valued MRI data contains phase information. This information can be used in clinical applications, such as MR angiography [43],
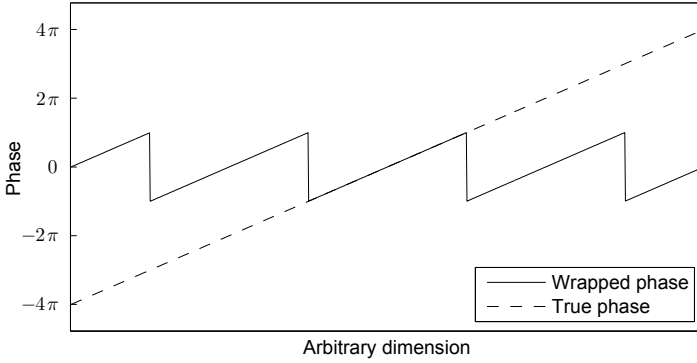
Figure 3.3: An illustration of the phase wrapping that occurs from (3.48).

but the phase can also be a nuisance parameter, arising due to field inhomogeneity. The measured complex number in each voxel has a phase ϕ defined on the interval $(-\pi, \pi]$, which means that a true phase outside this interval will be wrapped according to

$$\phi_{\text{wrapped}} = w(\phi) \triangleq \arg(e^{i\phi}) = \text{mod}(\phi + \pi, 2\pi) - \pi, \qquad (3.48)$$

where $\arg(\cdot)$ is the argument of a complex number, and $\text{mod}(a, b)$ is the modulo operation with dividend a and divisor b, where the result is defined to be positive. An example of a wrapped linear phase is shown in Fig. 3.3. The wrapping problem arises as it is not possible to separate a phase ϕ from $\phi + k2\pi$, $k \in \mathcal{Z}$, based on a single complex number. However, using additional information such as a sequence of measurements, it may be possible to unwrap the phase. In MRI, the assumption that the phase varies smoothly in space or time makes it possible to track phase increments, detect wraps that occur, and unwrap these to obtain a better representation of the underlying physics.

Under the assumption of smoothness, the phase unwrapping problem for 1D sequences is relatively straightforward, and consists of eliminating jumps in the phase larger than approximately 2π. The procedure may differ in details, but the simplest approach is to add or subtract 2π to samples differing by more than $2\pi k$ from the previous sample in the sequence. The constant $k \in [0.5, 1]$ enables unwrapping in the presence of minor noise, and is typically chosen based on the SNR; however, in Gaussian noise there is always a non-zero probability of erroneous unwrapping. The described unwrapping procedure follows from the fact that the derivative (where defined) of the wrapped phase is equal to the derivative of the true phase, given that the derivative of the true phase is less than π everywhere [70]. From this property, it can be seen that integrating the derivative of the wrapped phase will result in

unwrapping. This can be generalized to the sampled case by computing the first-order differences and wrapping the result to be in the interval $(-\pi, \pi]$. However, the differentiation will amplify the high frequency disturbances in the data, making this approach sensitive to noise.

It should be noted that when the SNR decreases, the phase information rapidly deteriorates. This is intuitively clear, as the noise can almost arbitrarily shift the angle of a vector if the amplitude is small enough. With random angular fluctuations in the interval $(-\pi, \pi]$, it is impossible to perform successful phase unwrapping with the heuristic approach above. However, by applying more sophisticated model-based parametric methods it is still possible to obtain a smooth estimate of the unwrapped phase. Parametric methods generally require some prior knowledge to set the model structure. Moreover, even for a given model structure, it can be problematic to estimate the parameters. For example, using the wrapping function defined in (3.48), we can construct the problem

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \ \|\boldsymbol{\phi}_{\text{wrapped}} - w\left(\boldsymbol{\psi}(\boldsymbol{\theta})\right)\|^2, \tag{3.49}$$

where $\boldsymbol{\psi}(\boldsymbol{\theta})$ is some predefined phase function depending on the parameter vector $\boldsymbol{\theta}$, and wrapping function $w(\cdot)$ acts on its argument elementwise. This problem is nonlinear, and might be intractable. Furthermore, the standard LS approach is only optimal if the phase noise is i.i.d. Gaussian, which does not hold when the phases are obtained from -complex-Gaussian distributed samples. The criterion surface corresponding to (3.49) typically has erratic behavior and multiple local minima. A simple example is shown in Fig. 3.4, where the phase varies linearly over the dimension under study d, that is, $\psi_n(\boldsymbol{\theta}) = \theta_1 d_n + \theta_2$, for a set of known values $\{d_n\}_{n=1}^N$. Clearly, the chance of finding the global optimum using a local minimization approach is small, although applying a brute-force method could be an option if it can be executed within the given time constraints. It should also be noted that not much is to be gained from using samples where the phase information is close to zero, and any estimation algorithm using the phase as input should take this fact into account.

An alternative is to estimate the phase based on the complex-valued data directly:

$$\underset{\boldsymbol{\theta}, \mathbf{a}}{\text{minimize}} \ \|\mathbf{y} - \mathbf{a} \odot e^{i\boldsymbol{\psi}(\boldsymbol{\theta})}\|^2, \tag{3.50}$$

where $\mathbf{a} \in \mathbb{R}^+$ is the vector of positive signal amplitudes, and $\odot$ is the Hadamard product. The advantage of this approach is that the LS estimates of the parameters are ML in Gaussian noise. As no structure of $\mathbf{a}$ is assumed, we can substitute its LS estimate $\hat{\mathbf{a}} = \text{Re}\{\mathbf{y} \odot e^{-i\boldsymbol{\psi}(\boldsymbol{\theta})}\}$

into (3.50) to obtain

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \ \left\| \mathbf{y} - \text{Re}\Big\{ \mathbf{y} \odot e^{-i\boldsymbol{\psi}(\boldsymbol{\theta})} \Big\} \odot e^{i\boldsymbol{\psi}(\boldsymbol{\theta})} \right\|^2, \tag{3.51}$$

or equivalently

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \ \left\| \text{Im}\Big\{ \mathbf{y} \odot e^{-i\boldsymbol{\psi}(\boldsymbol{\theta})} \Big\} \right\|^2. \tag{3.52}$$

It should be noted that there is an ambiguity between the sign of each a and a shift of the corresponding ψ by π. However, the smoothness of the function $\boldsymbol{\psi}(\boldsymbol{\theta})$ will ensure that all elements of the LS estimate of $\mathbf{a}$ are of the same sign, which makes this ambiguity easy to resolve.

At first glance, the criterion in (3.52) might seem odd; however, minimizing the imaginary part of the rotated data makes sense, as the phase is defined relative to the real axis. If the phase function $\boldsymbol{\psi}(\boldsymbol{\theta})$ fully explains the phase of the data, the corresponding rotated measurements should have a zero imaginary component. From this reasoning, it is easy to see that maximizing the real part is also equivalent to (3.52). Similarly to (3.49), the obtained criterion surface is relatively complex, even for a linear phase function.

If we normalize the amplitudes of the data to unity and omit the amplitude variable $\mathbf{a}$, we can express (3.50) as

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \ \left\| e^{i\boldsymbol{\phi}} - e^{i\boldsymbol{\psi}(\boldsymbol{\theta})} \right\|^2, \tag{3.53}$$

where $\boldsymbol{\phi}$ is the phase of the data. By expanding the criterion:

$$\left\| e^{i\boldsymbol{\phi}} - e^{i\boldsymbol{\psi}(\boldsymbol{\theta})} \right\|^2 = \left(e^{i\boldsymbol{\phi}} - e^{i\boldsymbol{\psi}(\boldsymbol{\theta})} \right)^* \left(e^{i\boldsymbol{\phi}} - e^{i\boldsymbol{\psi}(\boldsymbol{\theta})} \right) \tag{3.54}$$

$$= 2 \left(N - \mathbf{1}^{\text{T}} \cos\left(\boldsymbol{\phi} - \boldsymbol{\psi}(\boldsymbol{\theta}) \right) \right), \tag{3.55}$$

where N is the number of samples, and introducing the second-order Taylor approximation of $\cos(x)$ around zero:

$$\cos(x) = 1 - \frac{x^2}{2} + \mathcal{O}(x^4), \tag{3.56}$$

the criterion transforms into a sum of squares of the phase errors, and we arrive at the wrapped LS problem given by (3.49). It should be noted that normalizing the data affects the noise properties, and therefore, the estimates obtained from (3.49) can be biased. By introducing a weight matrix $\mathbf{W} = \text{diag}\left(|\mathbf{y}|^2\right)$, where the magnitude is taken elementwise, we can take the varying SNR into account, giving the criterion

$$\left\| e^{i\boldsymbol{\phi}} - e^{i\boldsymbol{\psi}(\boldsymbol{\theta})} \right\|_{\mathbf{W}}^2 = 2 \left(\|\mathbf{y}\|^2 - (|\mathbf{y}|^2)^{\text{T}} \cos\left(\boldsymbol{\phi} - \boldsymbol{\psi}(\boldsymbol{\theta}) \right) \right) \tag{3.57}$$

$$= \left\| \mathbf{y} - |\mathbf{y}| \odot e^{i\boldsymbol{\psi}(\boldsymbol{\theta})} \right\|, \tag{3.58}$$
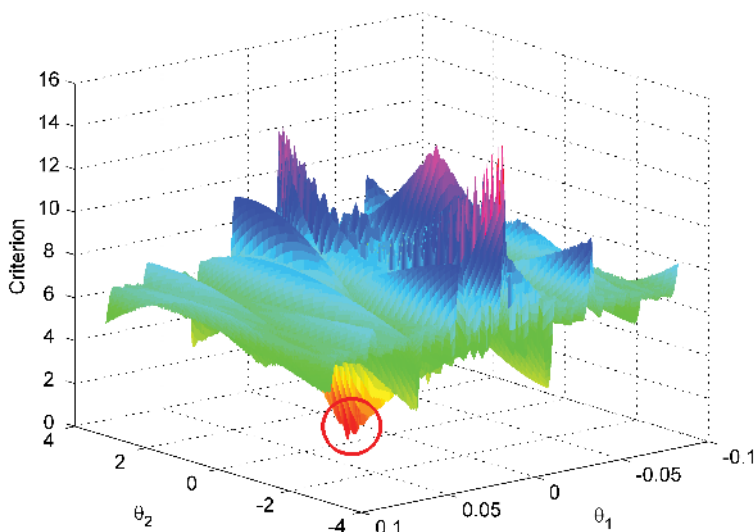
Figure 3.4: The LS phase unwrapping criterion of (3.49) for a simple linear phase model $\psi_n(\boldsymbol{\theta}) = \theta_1 d_n + \theta_2$, and no noise, versus the optimization parameters θ_1 and θ_2. The location of the global minimum is indicated by the red circle.

which gives a higher weight to the phase errors corresponding to samples with large magnitude. As seen from (3.58), the same expression is obtained by setting $\hat{\mathbf{a}} = |\mathbf{y}|$ in (3.50), which is a reasonable, although suboptimal, estimate of $\mathbf{a}$. Again, the expression can be simplified using the Taylor-series expansion of $\cos(\cdot)$, but this also introduces the wrapping function in (3.48). A parametric temporal phase-unwrapping approach is presented in Chapter 7.

Many parametric approaches suffer from nonlinearity, which is why heuristic methods are often used in practice. In 1D unwrapping, the starting point of the simple nonparametric phase unwrapping approach, described below equation (3.48), can only offset the unwrapped phase by a constant. Either a reference point is predefined, meaning that this constant is known, or the point of reference is arbitrary, in which case the constant can be omitted. Therefore, the result of 1D phase unwrapping is essentially unique. In two or more dimensions, the result of the phase unwrapping will depend on the starting point, or points, meaning that there is no unique solution (up to a constant). This is related to the fact that the unwrapping is done relative to another sample, and in higher dimensions, choosing this reference sample can be done in many ways. Moreover, starting from a sample with poor SNR can propagate
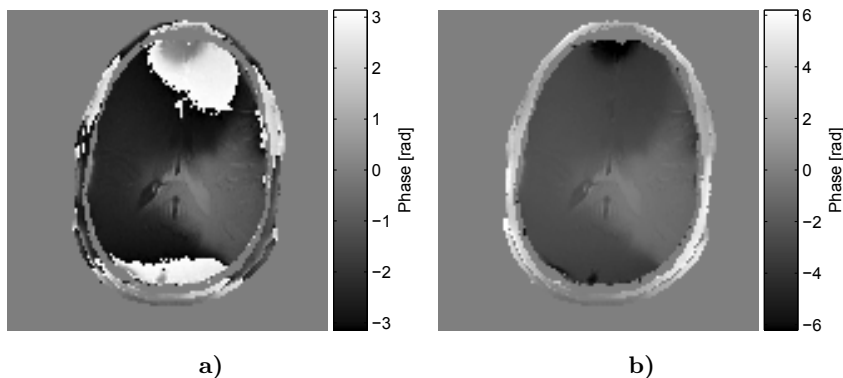
Figure 3.5: a) The wrapped phase of an in-vivo brain image, and b) the resulting unwrapped phase using the quality-guided flood-fill approach. The noisy background was masked in both images.

more errors, compared to using a high SNR starting point. There are many algorithms for multi-dimensional phase unwrapping available in the literature, cf. [49] for the details on several 2D methods, including both global and local unwrapping. Some algorithms are based on optimization, while others use heuristics such as branch cuts or network flow. A simple idea is to estimate the phase quality and start from a high quality sample, then propagate outwards, each time unwrapping the neighbor with highest phase quality. This approach is often called *quality-guided flood-fill* phase unwrapping. It allows for multiple starting points and can be generalized to higher dimensions, however, these options will not be considered here. There are several ways of defining a quality measure, for example, the standard deviation of the local phase derivative can be used. The first-order differences across both the x and y-direction are computed, and at voxel (k, p), the standard deviation of the differences across the four adjoining voxels $(k-1, p)$, $(k+1, p)$, $(k, p-1)$, and $(k, p+1)$ are computed for each direction, and added together. The resulting matrix contains estimates of the local variation of the phase in each pixel. A low variation implies a high-quality phase measure, and by unwrapping the these pixels first, the errors in the resulting 2D phase map can be suppressed. The quality-guided flood-fill method performs well for relatively high SNR, but does not guarantee that the result is free of discontinuities. Moreover, if the true phase variations are so large that they approach the wrapping limits, the smoothness assumption is invalid and the method may fail. An example of successful unwrapping of the phase of an in-vivo brain image using the quality-guided flood-fill approach is shown in Fig. 3.5.

Part II:
Signal processing problems in MRI

Chapter 4

Off-resonance mapping and banding removal

4.1 Introduction

One of the challenges of MRI is to acquire an image with high SNR in a short scan time, and for this the balanced steady-state free precession (bSSFP) sequence has proven to be of great interest. The main drawback of bSSFP is due to off-resonance effects, typically manifesting as banding artifacts [5, 143]. These artifacts are of major concern, especially at high field strengths. Off-resonance effects can lead to signal losses in parts of the image, and techniques for improving image quality are necessary.

When several acquisitions are made with different phase increments of the RF excitation, the resulting images can be combined to minimize these off-resonance artifacts. Commonly, two or four phase-cycled acquisitions are used as a compromise between performance and scan time. Several image-based techniques have been previously proposed, such as *sum-of-squares*, where the square root of the sum of the squared magnitude of the images is used; or *maximum-intensity*, where the maximum magnitude over all images is combined into one image [5]. These methods can in some cases give insufficient banding suppression; for example, when using small flip angles where the passband in the bSSFP signal profile is not flat [5]. Additionally, these techniques do not provide estimates of the model parameters, which can be of interest in qMRI.

Recent works have applied parameter estimation techniques to reduce banding artifacts in bSSFP. The principle is to use a signal model and estimate a parameter that is independent of the off-resonance. This estimate is then used as the band-free image. In [136], the authors treat the special case occurring when setting the echo time, *TE*, to zero and

acquiring data with a specific choice of phase increments. Then, the off-resonance effects can be removed using an analytical solution named the *cross-solution*. The resulting image will, however, have a different contrast compared to the original images. This is due to the fact that the parameter estimated relates to the original images through a function depending on both T_1 and T_2. The approach is also sub-optimal in the least squares sense, since it is derived with the assumption of no noise. This causes problems when the SNR is low, leading to poor estimates. Furthermore, the method does not provide estimates of all unknown parameters in the model equation and cannot be directly generalized to more than four phase-cycled images.

The approach suggested in [104] is to identify some of the unknown model parameters while assuming the others to be constant. Keeping the relaxation parameters constant makes the estimates less reliable, since in practice, the true values can vary significantly over an image. The approach is based on a manually initialized Levenberg-Marquardt (LM) nonlinear minimization algorithm applied to magnitude data. The use of magnitude data makes the estimation less tractable from a mathematical viewpoint, due to the non-differentiability of the absolute value. Furthermore, as discussed in Section 3.1.1, it changes the noise properties from a Gaussian distribution to a Rician distribution, making the NLS criterion sub-optimal (biased).

Another characteristic of bSSFP is the T_2/T_1-weighted image contrast and the subsequent difficulty to sensitize the signal to T_1 or T_2 alone. There are various techniques to estimate T_2 from bSSFP data [41, 10]. One popular technique is the DESPOT2, which was introduced in [41], and later improved to account for off-resonance effects [38]. The DESPOT2 method has the drawback that it needs a T_1 estimate obtained prior to the T_2 estimation, which requires the acquisition of an additional dataset. In [92] an outline of a method for simultaneous estimation of T_1 and T_2 from bSSFP data was proposed. The method is evaluated using 12 images, which is more than what is needed for DESPOT2, and no accuracy analysis has so far been presented for this technique. Both aforementioned methods [38, 92] utilize a variable flip angle in combination with phase cycling.

There are also methods for T_1 estimation as well as simultaneous estimation of T_1 and T_2, using inversion recovery bSSFP [106, 108]. However, neither of these methods take off-resonance effects into account. Furthermore, it is shown that in the presence of off-resonances, the method in [108] can suffer from significant bias.

In this chapter, we first describe a parameter estimation algorithm for the phase-cycled bSSFP signal model with the aim of reducing banding artifacts. We use complex-valued data to estimate all unknown parameters in a model derived from [45]. From the parameter estimates

we can reconstruct band-free images with bSSFP-like contrast. As a first step, we derive a fast and robust linear method based on least squares to approximately solve the estimation problem. This eliminates the need for user-defined parameters, such as manual initialization. In the second step, we fine tune the estimates using a nonlinear iterative minimization algorithm. The obtained estimates can then be used to reconstruct band-free images. The proposed algorithm can be applied to datasets regardless of the used echo time (TE) and repetition time (TR); does not rely on any prior assumption on the flip angle (α); and can be used with any number of phase-cycled images larger than or equal to three. Here, we will focus on four-image datasets, since they enable the parameter estimation approach, and generally provide better banding suppression compared to using two images. We then proceed to generalize the algorithm to simultaneously estimate T_1 and T_2 and the equilibrium magnetization including coil sensitivity (KM_0) from phase-cycled data. We derive the CRB for the bSSFP model to determine the statistical efficiency of the proposed algorithm, as well as the maximum theoretical accuracy we can expect when estimating T_1 and T_2 simultaneously, using phase-cycled bSSFP.

4.2 Theory

4.2.1 Signal model

In bSSFP imaging, the complex signal, S, at an arbitrary voxel of the n:th phase-cycled image can be modeled as [45, 79]

$$S_n = KM e^{-\frac{TE}{T_2}} e^{i\Omega TE} \frac{1 - a e^{-i(\Omega + \Delta\Omega_n)TR}}{1 - b\cos\left[(\Omega + \Delta\Omega_n)TR\right]} + v_n, \qquad (4.1)$$

where we have the following definitions:

$$
\begin{aligned}
M &= iM_0 \frac{(1 - E_1)\sin\alpha}{1 - E_1\cos\alpha - (E_1 - \cos\alpha)E_2^2}, \\
a &= E_2, \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (4.2) \\
b &= E_2 \frac{1 - E_1 - E_1\cos\alpha + \cos\alpha}{1 - E_1\cos\alpha - (E_1 - \cos\alpha)E_2^2}.
\end{aligned}
$$

Furthermore, $E_1 = e^{-TR/T_1}$, $E_2 = e^{-TR/T_2}$, K is the complex-valued coil sensitivity, M_0 the equilibrium magnetization, T_1 and T_2 are the longitudinal and transverse relaxation times, respectively, α is the flip angle, $\Delta\Omega_n TR$ the user-controlled phase increment, TE the echo time, and TR the repetition time. We also define the joint variable KM_0, which describes the equilibrium magnetization perturbed by the coil

sensitivity. The off resonance corresponds to $\Omega = 2\pi\gamma\Delta B_0 = 2\pi f_{OR}$, where γ is the gyromagnetic ratio, ΔB_0 is the effective deviation from the ideal static magnetic field strength, including both tissue susceptibility and inhomogeneities, and f_{OR} is the corresponding off-resonance frequency. Finally, v_n denotes the noise, which is assumed to be i.i.d. complex-Gaussian distributed.

If the data is acquired by changing the center frequency, which mimics phase cycling, an extra shift of $\Delta\Omega_n$ is added in the first exponential term of (4.1), giving

$$\tilde{S}_n = KMe^{-\frac{TE}{T_2}}e^{i(\Omega+\Delta\Omega_n)TE}\frac{1 - ae^{-i(\Omega+\Delta\Omega_n)TR}}{1 - b\cos\left[(\Omega + \Delta\Omega_n)TR\right]} + v_n. \qquad (4.3)$$

This model will not be used here, but as will be shown in the following, it is straightforward to modify the algorithms presented in this chapter to handle center-frequency shifting.

There are five real-valued unknown parameters in (4.1) and (4.2) that can be estimated: Ω, $\text{Re}\{KM_0\}$, $\text{Im}\{KM_0\}$, T_1, and T_2. The parameters assumed to be known are: $\Delta\Omega_n$, α, TE, and TR; however, as will be shown next, α does not have to be known for reconstructing band-free images, but only to get explicit estimates of KM_0 and T_1.

We introduce the following variables:

$$\begin{aligned}
S_0 &= KMe^{-TE/T_2}, \quad \theta = \Omega TR, \\
\Delta\theta_n &= \Delta\Omega_n TR, \qquad \theta_n = \theta + \Delta\theta_n.
\end{aligned} \qquad (4.4)$$

This enables us to rewrite the voxelwise signal model in (4.1) as

$$S_n = S_0 e^{i\theta TE/TR}\frac{1 - ae^{-i\theta_n}}{1 - b\cos\theta_n} + v_n = g_n(\mathbf{u}) + v_n, \qquad (4.5)$$

where $g_n(\mathbf{u})$ is the noise-free data model of the n:th phase-cycled image, and the vector of new unknown model parameters is denoted by $\mathbf{u} = [\theta,\ \text{Re}\{S_0\},\ \text{Im}\{S_0\},\ a,\ b]^{\text{T}}$. Acquiring images with different phase increments $\Delta\theta_n$ allows us to estimate the unknown model parameters of (4.5). Using these parameters, we can reconstruct a band-free image from the model. It can be noted that α does not explicitly occur in (4.5), and hence, no prior information regarding the flip angle is needed when using (4.5) for band reduction.

Even though the phase increments can be arbitrary, using four images ($N = 4$) with the phase increments $\Delta\boldsymbol{\theta} = [0,\ \pi/2,\ \pi,\ 3\pi/2]^{\text{T}}$ is common practice and will therefore be considered here as well. It is possible to optimize the phase increments in some sense, especially if some prior information is available. However, a preliminary study [15] showed that the gain from doing so is small, assuming no prior knowledge, and therefore, this option will not pursued further here.

4.2.2 Derivation of the signal model

The model in (4.1) can be derived from the Bloch equations (2.3). The following derivation was inspired by [45], but generalized to account for an arbitrary echo time TE, and to accommodate the use of phase cycling.

Assume that we have a right-handed coordinate system with the z-axis along the static magnetic field B_0. The magnetization vector ($\mathbf{M}$) immediately before the kth RF pulse is related to that after by a simple rotation of $-\alpha$. Assuming that the pulse is along the x-axis, we have

$$\mathbf{M}(kTR^+) = \mathbf{R}_x(-\alpha)\mathbf{M}(kTR^-), \tag{4.6}$$

where the matrix

$$\mathbf{R}_x(-\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{bmatrix}, \tag{4.7}$$

describes a rotation around the x-axis. Furthermore, we have defined $\mathbf{M}(t^\pm) = \lim_{\tau \to t^\pm} \mathbf{M}(\tau)$. With linear phase cycling, the phase of the pulse will be incremented by the same phase for each repetition. Introducing a coordinate system which aligns with the RF phase at each excitation, we can still express the flips as rotations around the x-axis. During each TR, free precession causes the magnetization to rotate about the z-axis, and relaxation causes an exponential recovery towards thermal equilibrium (M_0). When using a discretely rotating frame that rotates with the linear phase increment, an additional rotation about the z-axis (equal and opposite the phase increment) is introduced. Using the same type of rotation matrix defined in (4.6), but now properly modified to give a rotation around the z-axis, we get the following expression:

$$\mathbf{M}((k+1)TR^-) = \\ \mathbf{R}_z(\Omega TR)\mathbf{R}_z(\Delta\Omega TR)\mathbf{D}(TR)\mathbf{M}(kTR^+) + (1 - E_1)\mathbf{M_0}, \quad (4.8)$$

where $\mathbf{D}(t) = \mathrm{diag}\left([e^{-t/T_2},\ e^{-t/T_2},\ e^{-t/T_1}]\right)$ is a damping matrix, $E_1 = e^{-TR/T_1}$, $\mathbf{M_0} = [0,\ 0,\ M_0]^\mathrm{T}$ is the equilibrium magnetization directed along the z-axis, $\Omega = 2\pi\gamma\Delta B_0$ corresponds to the off-resonance frequency in radians per second, and $\Delta\Omega TR$ is the user-controlled phase increment of the RF pulse.

Substituting (4.8) into (4.6) and using the fact that $\mathbf{M}((k+1)TR^+) = \mathbf{M}(kTR^+)$ at steady-state, the resulting system of equations can be solved to obtain $\mathbf{M}(TR^+)$, where the arbitrary integer k has been omit-

ted to simplify the notation. The solution is

$$\mathbf{M}(TR^+) = (\mathbf{I} - \mathbf{R}_x(-\alpha)\mathbf{R}_z(\Omega TR)\mathbf{R}_z(\Delta\Omega TR)\mathbf{D}(TR))^{-1} \times$$
$$(1 - E_1)\mathbf{R}_x(-\alpha)\mathbf{M_0}. \quad (4.9)$$

At the echo time TE, the free precession has rotated the magnetization in the transverse plane by an angle $2\pi\gamma B_0 TE = \Omega TE$ under decay from the initial $\mathbf{M}(TR^+)$. This can be expressed as

$$\mathbf{M}(TE) = \mathbf{R}_z(\Omega TE)\mathbf{D}(TE)\mathbf{M}(TR^+). \quad (4.10)$$

Note that there is no accumulation of phase due to the rotation of the frame, since the frame rotates in discrete steps just before each excitation.

We are only interested in the transverse component of (4.10). Expressing it as a complex number $S_n = M_x(TE) + iM_y(TE)$ and simplifying, we get

$$S_n = M e^{-\frac{TE}{T_2}} e^{i\Omega TE} \frac{1 - ae^{-i(\Omega+\Delta\Omega_n)TR}}{1 - b\cos\left[(\Omega + \Delta\Omega_n)TR\right]}, \quad (4.11)$$

where we have defined

$$M = iM_0 \frac{(1 - E_1)\sin\alpha}{1 - E_1\cos\alpha - (E_1 - \cos\alpha)E_2^2},$$
$$a = E_2, \quad (4.12)$$
$$b = \frac{E_2(1 - E_1)(1 + \cos\alpha)}{1 - E_1\cos\alpha - (E_1 - \cos\alpha)E_2^2},$$

and $E_2 = e^{-TR/T_2}$. Including a coil sensitivity K and a multiplicative constant, we arrive at the expression in (4.1).

4.2.3 The Cramér-Rao bound

For the model in (4.5) the Jacobian vector in the FIM of (3.14) is given by

$$\frac{\partial g_n(\mathbf{u})}{\partial \mathbf{u}} =$$

$$e^{i\theta\frac{TE}{TR}}\frac{1 - ae^{-i\theta_n}}{1 - b\cos(\theta_n)}\begin{bmatrix} S_0\left(i\frac{ae^{-i\theta_n}}{1-ae^{-i\theta_n}} - \frac{b\sin(\theta_n)}{1-b\cos(\theta_n)} + i\frac{TE}{TR}\right) \\ 1 \\ i \\ S_0\frac{e^{-i\theta_n}}{1-ae^{-i\theta_n}} \\ S_0\frac{\cos(\theta_n)}{(1-b\cos(\theta_n))} \end{bmatrix}^{\mathrm{T}}. \quad (4.13)$$

The CRB matrix $\mathbf{C}_{\text{CRB}}$ is then obtained through (3.12).

Additionally, we derived the CRB with respect to $\mathbf{u}_o = [\theta, \text{Re}\{KM_0\}, \text{Im}\{KM_0\}, T_1, T_2]^T$, using MATLAB's Symbolic Math Toolbox. This was done to draw conclusions about the initial estimation problem based on the model in (4.1).

4.2.4 The LORE-GN algorithm

To estimate the unknown parameters and remove the off-resonance artifacts, we propose a two-step algorithm. The first step is named *Linearization for Off-Resonance Estimation* (LORE). The second step is a Gauss-Newton (GN) nonlinear search, hence we name the full algorithm LORE-GN. In the first step we rewrite the model in (4.5) so that it becomes linear in the unknown parameters, by making use of an over-parameterization. This enables the application of ordinary linear least squares (LS), which is both fast and robust. However, the resulting estimates will be biased in general. In the following step, the final estimates are obtained using GN, initialized with the LORE estimates. This removes any bias and makes the estimates NLS optimal, which under the assumption of identically distributed Gaussian noise is the maximum likelihood (ML) estimate.

From the estimates of $S_0 \in \mathbb{C}$ and $a, b, \theta \in \mathbb{R}$ it is possible to recover the original parameters $\mathbf{u}_o$ by successively inverting the equations for a and b in (4.2), and substituting the results. We have

$$
\begin{aligned}
\hat{E}_1 &= \frac{\hat{a}(1 + \cos\alpha - \hat{a}\hat{b}\cos\alpha) - \hat{b}}{\hat{a}(1 + \cos\alpha - \hat{a}\hat{b}) - \hat{b}\cos\alpha}, \\
\hat{E}_2 &= \hat{a},
\end{aligned}
\tag{4.14}
$$

which in turn can be used to compute the estimates:

$$
\begin{aligned}
\hat{T}_1 &= -TR/\log(\hat{E}_1), \\
\hat{T}_2 &= -TR/\log(\hat{E}_2), \\
K\hat{M}_0 &= \hat{S}_0 \frac{1 - \hat{E}_1\cos\alpha - (\hat{E}_1 - \cos\alpha)\hat{E}_2^2}{ie^{-TE/\hat{T}_2}(1 - \hat{E}_1)\sin\alpha}.
\end{aligned}
\tag{4.15}
$$

From (4.15) it is clear that T_2 can be estimated without any knowledge of the flip angle α. However, to obtain KM_0 and T_1, an estimate of α is needed. Assuming a constant flip angle will introduce errors since the effective flip can vary significantly over the image. For example, the T_1 dependence on α can be quite sensitive, and a small error in the flip angle can lead to significant changes in the estimated T_1. Several techniques for estimating the flip angle (B_1 mapping) are available in

the literature, see for example [36, 137, 103]. Furthermore, since M_0 is ideally real-valued, the phase of K can also be obtained, giving a partial separation of the two variables.

Step 1: Parameter estimation using LORE

For the LORE algorithm, we introduce the following complex-valued parameters:

$$\eta = S_0 e^{i\theta TE/TR},$$
$$\beta = S_0 a e^{i\theta(TE/TR-1)}, \qquad\qquad (4.16)$$
$$\zeta = b e^{i\theta}.$$

Note the slight over-parameterization with six real-valued parameters as opposed to five in (4.5). This enables us to rewrite the noise-free part of (4.5) as

$$S_n = \frac{\eta - \beta e^{-i\Delta\theta_n}}{1 - \mathrm{Re}\{\zeta e^{i\Delta\theta_n}\}}. \qquad\qquad (4.17)$$

If center-frequency shifting is used instead of phase cycling, we can use the identity $S_n = \tilde{S}_n e^{-i\Delta\theta_n TE/TR}$, which follows from (4.1), (4.3), and (4.4). Multiplying the samples $\tilde{S}_n$ with the known factor $e^{-i\Delta\theta_n TE/TR}$ enables direct use of (4.17) and the following algorithm, in conjunction with frequency-shifted data.

To simplify the notation we introduce the subscripts r and i to denote the real and imaginary part, respectively. Multiplying both sides by the denominator we can now express (4.17) in linear form:

$$S_n \left[1 - \zeta_r \cos\Delta\theta_n + \zeta_i \sin\Delta\theta_n\right] = \eta - \beta e^{-i\Delta\theta_n}. \qquad (4.18)$$

Furthermore, the noise is amplified by at most a factor of two from this operation, since it can be shown that $0 \le b \le 1$, which, in turn, implies that $|1 - \zeta_r \cos\Delta\theta_n + \zeta_i \sin\Delta\theta_n| \le 2$. Moving the unknown variables to the right hand side and gathering the real and imaginary parts of S_n separately in a vector $\mathbf{y}_n = [S_{r,n}\ S_{i,n}]^{\mathrm{T}}$, we can write (4.18) in matrix form:

$$\mathbf{y}_n = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -\cos(\Delta\theta_n) & \sin(\Delta\theta_n) \\ -\sin(\Delta\theta_n) & -\cos(\Delta\theta_n) \\ S_{r,n}\cos(\Delta\theta_n) & S_{i,n}\cos(\Delta\theta_n) \\ -S_{r,n}\sin(\Delta\theta_n) & -S_{i,n}\sin(\Delta\theta_n) \end{bmatrix}^{\mathrm{T}}}_{\mathbf{A}_n} \underbrace{\begin{bmatrix} \eta_r \\ \eta_i \\ \beta_r \\ \beta_i \\ \zeta_r \\ \zeta_i \end{bmatrix}}_{\mathbf{x}}. \qquad (4.19)$$

By stacking all measurements in a vector $\mathbf{y} = [\mathbf{y}_1 \cdots \mathbf{y}_N]^{\mathrm{T}}$, and constructing the matrix $\mathbf{A} = \left[\mathbf{A}_1^{\mathrm{T}} \cdots \mathbf{A}_N^{\mathrm{T}}\right]^{\mathrm{T}}$, where N is the number of

phase-cycled images, we obtain a model of the form $\mathbf{y} = \mathbf{A}\mathbf{x}$. Using (3.6), the LS estimate of $\mathbf{x}$ is readily found as:

$$\hat{\mathbf{x}} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{y}. \tag{4.20}$$

The estimates of the sought parameters in (4.5) can then be obtained as

$$\begin{aligned}
\hat{\theta} &= -\arg\left(\hat{\beta}/\hat{\eta}\right), \\
\hat{a} &= |\hat{\beta}/\hat{\eta}|, \\
\hat{b} &= |\hat{\zeta}|, \\
\hat{S}_0 &= \hat{\eta}e^{-i\hat{\theta}TE/TR},
\end{aligned} \tag{4.21}$$

where $\arg(\cdot)$ denotes the phase of a complex number. The information in ζ regarding the off-resonance is not used since ζ can be small in magnitude, leading to unreliable estimates of θ. While LORE can provide accurate estimates, they are sub-optimal in the NLS sense, as the noise enters the regressor matrix $\mathbf{A}$ through the measured data S_n. To tackle this, the LORE estimates can be used as an initial guess for the next step.

Step 2: Fine tuning using Gauss-Newton
We propose to use a Gauss-Newton iterative method to truly minimize the NLS and further improve the results. GN is chosen since it is simple, computationally efficient, and has fast convergence [114]. However, this minimization method is unconstrained, so any physical constraints on the parameters cannot be taken into account. Given a good initial estimate, here provided by LORE, GN converges to the correct global optimum with high probability. This is what distinguishes LORE-GN from other general nonlinear methods.

The NLS criterion is given by

$$L(\mathbf{u}) = \sum_{n=1}^{N} |S_n - g_n(\mathbf{u})|^2. \tag{4.22}$$

Letting $\mathbf{r}$ denote the residual vector, according to

$$\mathbf{r} = \begin{bmatrix} \mathrm{Re}\{\mathbf{S}\} \\ \mathrm{Im}\{\mathbf{S}\} \end{bmatrix} - \begin{bmatrix} \mathrm{Re}\{\mathbf{g}(\mathbf{u})\} \\ \mathrm{Im}\{\mathbf{g}(\mathbf{u})\} \end{bmatrix}, \tag{4.23}$$

the update formula for GN with the search direction $\mathbf{p}_k$ is

$$\mathbf{u}_{k+1} = \mathbf{u}_k + c\mathbf{p}_k = \mathbf{u}_k + c(\mathbf{J}_k^{\mathrm{T}}\mathbf{J}_k)^{-1}\mathbf{J}_k^{\mathrm{T}}\mathbf{r}_k, \tag{4.24}$$

where $\mathbf{J}_k = \mathbf{J}(\mathbf{u})\big|_{\mathbf{u}=\mathbf{u}_k}$ is the Jacobian matrix evaluated at the current point in the parameter space $\mathbf{u}_k$. In the same manner, we derived the GN algorithm with respect to the original model parameters $\mathbf{u}_{\mathrm{o}} = [\theta, \ \mathrm{Re}\{KM_0\}, \ \mathrm{Im}\{KM_0\}, \ T_1, \ T_2]^{\mathrm{T}}$.

The step length c is chosen by back-tracking so that the *Armijo condition* is fulfilled, that is, $c = 2^{-m}$, where m is the smallest non-negative integer that fulfills

$$L(\mathbf{u}_{k+1}) \leq L(\mathbf{u}_k) - \mu c \mathbf{r}_k^{\mathsf{T}} \mathbf{J}_k (\mathbf{J}_k^{\mathsf{T}} \mathbf{J}_k)^{-1} \mathbf{J}_k^{\mathsf{T}} \mathbf{r}_k, \qquad (4.25)$$

and $\mu \in [0,\ 1]$ is a constant [90]. The second term of the right-hand-side of (4.25) is proportional to both the step length and the directional derivative of the criterion along the search direction, and it is used to enforce a sufficient decrease in the criterion. Similarly, if the function changes rapidly along the search direction, the step will be made smaller. A stopping condition based on the norm of the gradient $\|\mathbf{J}_k^{\mathsf{T}} \mathbf{r}_k\|$ was used. In the following, μ was set to 0.5 and the stopping condition to $\|\mathbf{J}_k^{\mathsf{T}} \mathbf{r}_k\| < 10^{-8}$.

The obtained estimates can then be used to reconstruct band-free images with bSSFP contrast by using the model in (4.5), setting $\theta = 0$, and letting $\Delta\theta$ be any constant value. When considering the explicit parameter estimates, however, phase wrapping and nonphysical optima can cause ambiguities. These problems are treated in the next section.

The MATLAB code for the algorithm is available for general use at: `https://github.com/AAAArcus/LORE-GN`

4.2.5 Post-processing

By analyzing the model, a few interesting properties can be seen. The following equation holds:

$$S_0 e^{i\theta TE/TR} \frac{1 - a e^{-i\theta_n}}{1 - b\cos\theta_n} =$$
$$S_0 e^{\mp i\pi TE/TR} e^{i(\theta \pm \pi)TE/TR} \frac{1 + a e^{-i(\theta_n \pm \pi)}}{1 + b\cos(\theta_n \pm \pi)}. \quad (4.26)$$

This means that for a set of optimal parameters a, b, θ, and S_0, the NLS criterion will have another global optimum at $\tilde{a} = -a$, $\tilde{b} = -b$, $\tilde{\theta} = \theta \pm \pi$, and $\tilde{S}_0 = S_0 e^{\mp i\pi TE/TR}$. It can be shown that a and b are positive, and hence, we can remove the resulting non-physical minima. It should be noted that LORE provides $a, b \geq 0$ by design, and does not suffer from this ambiguity; however, GN is unconstrained and could potentially estimate negative a and b. In practice, if a minimization algorithm is initialized close to a feasible optimum, this is unlikely to

occur. Furthermore, we have

$$
S_0 e^{i(\theta+2\pi k)TE/TR} \frac{1 - ae^{-i(\theta_n+2\pi k)}}{1 - b\cos(\theta_n + 2\pi k)} =
$$
$$
S_0 e^{i2\pi kTE/TR} e^{i\theta TE/TR} \frac{1 - ae^{-i\theta_n}}{1 - b\cos\theta_n}, \quad \forall k \in \mathbb{Z}, \quad (4.27)
$$

that is, a shift of θ by $2\pi k$ is equivalent to a phase shift of S_0 by $2\pi kTE/TR$. The estimate of θ is confined in the interval $(-\pi, \ \pi]$ (wrapped phase), meaning that if the true θ is outside of our estimation interval, we will obtain the wrong phase of S_0. The number of possible S_0 estimates is the smallest integer k for which kTE/TR is an integer, which could be a large number. In practice, however, the phase will only wrap a few times, since it is proportional to the deviation of the static magnetic field B_0, and the number of solutions will be further limited by this fact. It is important to realize that the magnitude signal is not affected by (4.26) and (4.27), and hence, the post-processing step is not needed for band reduction. The problems only arise when estimating the absolute off-resonance and a complex-valued S_0, in which case phase unwrapping is needed to get consistent estimates. By assuming that θ is close to zero in the center of the image, which can be obtained through proper shimming, we can unwrap the estimated phase to obtain k in (4.27) in each voxel, and then compensate our S_0 estimates according to

$$
\hat{S}_{0\mathrm{u}} = \hat{S}_0 e^{-i2\pi kTE/TR}. \tag{4.28}
$$

Given that proper shimming has indeed been ensured, we obtain the true estimate of S_0. Phase unwrapping in two or three dimensions is a common problem in MRI, and there are several methods available in the literature, see for example [32] for a review. Here, a MATLAB implementation of the quality-guided 2D phase-unwrapping algorithm was used [49], see Section 3.2.2 for more details.

The two correction steps, that is (4.26) and phase unwrapping together with (4.28), constitute the post-processing step, which can be used to avoid some local minima of the criterion function in (4.22).

4.3 Methods

4.3.1 Simulations and the CRB

Simulations were performed with the following parameters:

$$
\begin{array}{llll}
T_1 = 675 \text{ ms}, & T_2 = 75 \text{ ms}, & KM_0 = 1 \\
\alpha = 30°, & TR = 5 \text{ ms}, & TE = 2.5 \text{ ms}.
\end{array} \tag{4.29}
$$

These parameters were chosen as a representative case targeting brain white matter at 1.5 T, and they were the basis of all simulations unless stated otherwise.

The simulated data was generated by adding i.i.d. circular complex-Gaussian noise of appropriate variance σ^2 to the model $g_n(\mathbf{u})$ in (4.5). The variance was chosen to achieve a certain SNR as defined by

$$\text{SNR} = \frac{\sum_{n=1}^{N} |g_n(\mathbf{u})|}{N\sigma}, \tag{4.30}$$

which is the common definition of SNR in the MRI community [89].

The root mean square error (rMSE) of the parameter estimates is defined as

$$\text{rMSE}(\hat{z}) = \sqrt{\frac{1}{M} \sum_{m=1}^{M} |\hat{z}_m - z|^2}, \tag{4.31}$$

where $\hat{z}_m$ is the parameter estimate in simulation m, z is the true parameter value, and M is the number of simulations.

The simulations and calculations were performed in MATLAB on an HP desktop computer with a 2.8 GHz Intel Core i7 860 quad-core processor and 16GB RAM. All computation times were measured when running a single thread.

Monte Carlo simulations provided the rMSE of the parameter estimates. The performance of the proposed algorithm was compared to 1) the LM algorithm suggested by Santini and Scheffler [104], 2) LM with our suggested post-processing step (LMpost), and 3) the optimum performance given by the CRB. The standard MATLAB LM implementation in the function "lsqnonlin" was utilized. The estimates obtained with LORE were also included in the comparison to illustrate the accuracy of this linear approximation algorithm, and hence, the accuracy of the initial estimates used by GN. The cross-solution proposed in [136] is not included here, as it does not estimate the model parameters.

To illustrate the CRB of T_1 and T_2, the minimum SNR needed to obtain a 5% relative standard deviation (RSD) in the estimates was calculated. The RSD was defined as the CRB standard deviation of the parameter estimate, relative to the true parameter value, that is for parameter j in the CRB matrix $\mathbf{C}_{\text{CRB}}$, $\text{RSD}_j = \sqrt{[\mathbf{C}_{\text{CRB}}]_{jj}}/u_j$. Numerical simulations were performed for true parameter values in the ranges $T_1 = 100 - 3000$ ms and $T_2 = 5 - 200$ ms.

4.3.2 Phantom and in-vivo data

A phantom and an in-vivo brain dataset was acquired using a 1.5 T scanner (GE Healthcare, Milwaukee, WI). Each dataset consisted of

four complex-valued 3D bSSFP images with linear phase increments of
0, $\pi/2$, π, and $3\pi/2$. The scan parameters were: FOV = $24 \times 24 \times 16$
cm^3, matrix size = $128 \times 128 \times 32$, $TR = 5$ ms, $TE = 2.5$ ms, $\alpha = 30°$.
The phantom dataset was included since the banding artifacts remain-
ing after the processing are easily visualized when the ideal intensity
is uniform. For demonstration purposes, in order to induce more sig-
nificant banding artifacts in the phantom and the in-vivo datasets, the
automatic shimming was disabled at 1.5 T. The estimated average SNR
was 170 and 33 for the phantom and in vivo data, respectively.

Similarly, an in-vivo brain dataset was acquired using a 7 T system
(GE Healthcare). The scan parameters were as follows: FOV = $20 \times
20 \times 16$ cm^3, matrix size = $200 \times 200 \times 160$, $TR = 10$ ms, $TE = 5$
ms, $\alpha = 10°$. High-order shimming was used to achieve best-possible
field homogeneity. The estimated average SNR was 11. This dataset was
included to show that the proposed method can be applied to higher field
strengths, where banding artifacts are typically more significant. Also,
the low flip angles used at high field strengths due to specific absorption
rate (SAR) constraints give a non-flat passband in the bSSFP profile,
which is problematic for many of the competing approaches. The longer
TR for the 7 T dataset is motivated by the application in [139], where
phase-cycled bSSFP at 7 T is used for high resolution imaging of the
hippocampus.

Before running the LORE-GN algorithm, the data was masked to
remove the background, thereby reducing the computation time. This
was done by thresholding the sum-of-squares image and masking voxels
with intensity below a certain percentage of the maximum value, in this
case 15% for 1.5 T, and 6% for 7 T. The resulting number of computed
voxels was 9467, 22712, and 22844, for the 1.5 T phantom, 1.5 T in vivo,
and 7 T in vivo data, respectively.

Here we reconstructed the images at $\Delta\theta = \pi/2$ to be able to compare
directly with one of the original phase-cycled images. This corresponds
to the image closest to maximum SNR reconstruction for both white and
gray matter. Since the computed reconstructions, as well as the the
collected phase-cycled images, are complex-valued, the corresponding
magnitudes were used when displaying the images.

4.4 Results

4.4.1 Simulations and the CRB

The results of the Monte Carlo simulations, based on 10000 noise real-
izations, are shown in Fig. 4.1. As can be seen, the LORE-GN algorithm
is efficient when estimating the parameters of (4.1) given an SNR above
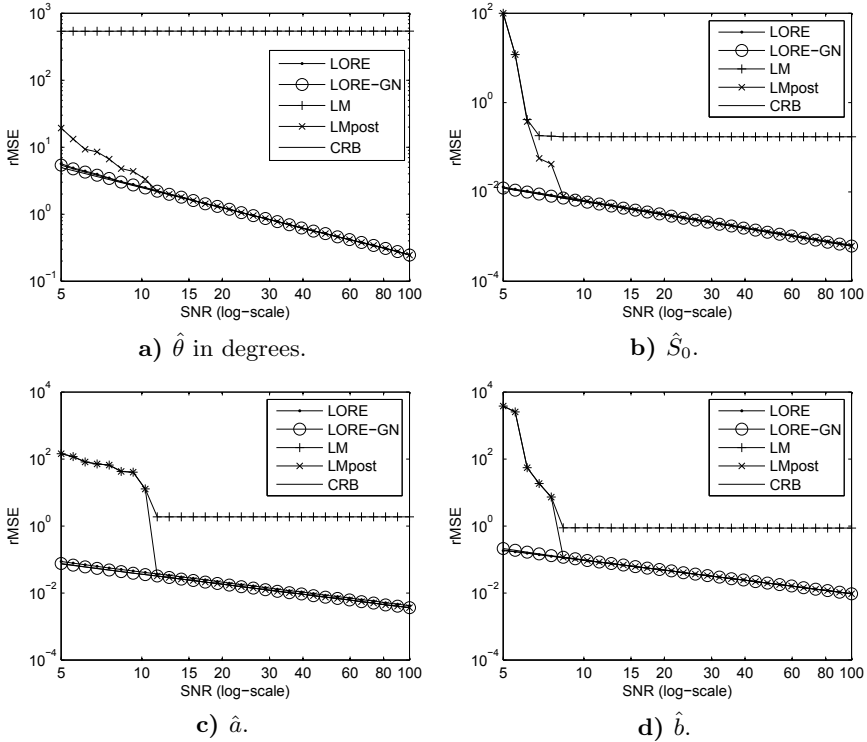5, as it achieves the CRB; and furthermore, LORE alone has compara-

Figure 4.1: rMSE vs. SNR for the parameter estimates of a) $\hat{\theta}$ (degrees), b) $\hat{S}_0$, c) $\hat{a}$, and d) $\hat{b}$, using the different methods, and the associated CRB. The results are based on 10000 Monte Carlo simulations. The true values were $\theta = 0$, $S_0 = 0.1207i$, $a = 0.9355$, and $b = 0.4356$.

ble performance in this case. SNRs below 5 have been excluded since the variance is bound to be high, which would generally not result in useful estimates.

Indirect estimates of T_1, T_2 and KM_0 can be obtained from (4.15), and the corresponding performance is shown in Fig. 4.2. In these figures, outliers have been removed at the lower SNR values. This was done by omitting estimates with an absolute distance larger than 20 CRB standard deviations from the true value. The reason for removing outliers is that single large values, caused by a noise realization leading to singularity, can have a great impact on the rMSE values, but these cases are easily detected and removed. The estimates in Fig. 4.2 have a rather high variance in general, and an SNR above 50 is needed to achieve the CRB for T_1 and T_2.

Figures 4.3a and 4.3b show how the SNR needed for accurate T_1 and T_2 estimation varies for different true values of T_1 and T_2. It can be seen
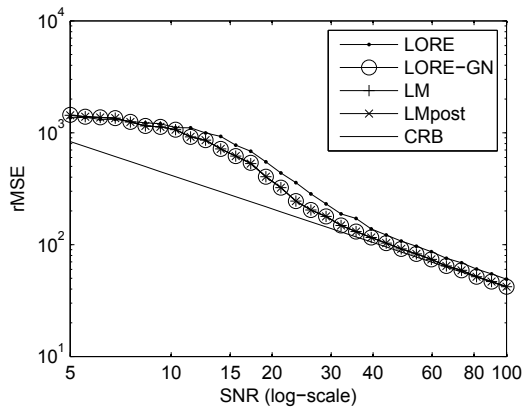
a) $\hat{T}_1$.
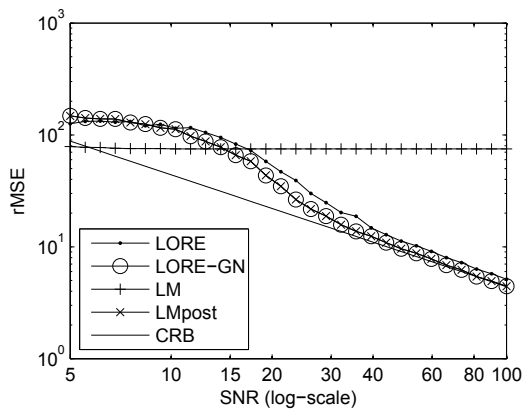
b) $\hat{T}_2$.
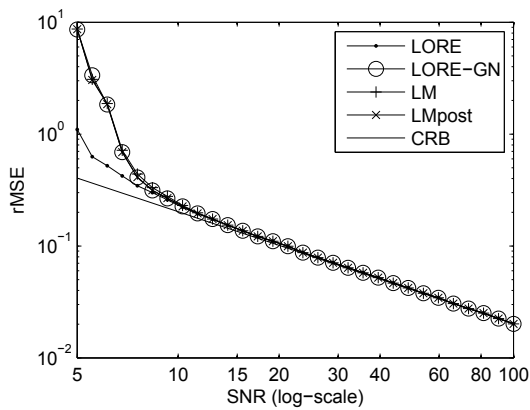
c) $K\hat{M}_0$.

Figure 4.2: rMSE vs. SNR for the indirect parameter estimates of a) T_1, b) T_2, and c) KM_0, using the different methods, and the associated CRB. The results are based on 10000 Monte Carlo simulations. The true values were $T_1 = 675$ ms, $T_2 = 75$ ms, and $KM_0 = 1$.

that the minimum required SNR is generally quite high. For example, an SNR of roughly 72 is needed to estimate the relaxation parameters of white matter and gray matter in the brain with 5% RSD.

4.4.2 Phantom example

Four phase-cycled images from a central slice of a 1.5 T 3D phantom dataset, all showing some degree of banding, are shown in Fig. 4.4a. The LORE-GN reconstructed image at $\Delta\theta = \pi/2$ is shown in Fig. 4.4b, together with the sum-of-squares, maximum-intensity, and cross-solution images in Figs. 4.4c, 4.4d, and 4.4e, respectively. The proposed method results in an image showing no bands, while in the sum-of-squares and maximum-intensity images, some artifacts still remain. The cross-solution gives a uniform image, but has an intensity corresponding to the parameter S_0. The off-resonance frequency estimated by LORE-GN is shown in Fig. 4.4f. For comparison, the off-resonance was also estimated using two gradient-echo images with $TE = 4$ and 5 ms, respectively, and computing the field map from the phase difference of these images as is described in [9]. The result is shown in Fig. 4.4g. As can be seen, LORE-GN provides a smooth low noise estimate that corresponds well with the gradient echo based technique.

4.4.3 In-vivo examples

The phase-cycled images from a central slice of the 1.5 T 3D in-vivo dataset are shown in Fig. 4.5a. LORE-GN was applied to estimate the model parameters. The reconstructed image at $\Delta\theta = \pi/2$ is shown in Fig. 4.5b, together with the sum-of-squares, maximum-intensity and cross-solution images in Figs. 4.5c, 4.5d and 4.5e, respectively. As expected, banding artifacts are more subtle than in the phantom experiments. By scaling and subtracting the sum-of-squares, maximum-intensity and cross-solution images from the LORE-GN estimate, the differences can be visualized more clearly. These images are shown in Figs. 4.5f, 4.5g and 4.5h, respectively. Sum-of-squares shows some nonuniformity compared to the proposed estimate, while the maximum-intensity and cross-solution approaches give a similar level of uniformity. Using any constant initialization of LM seemed to give estimation errors in some parts of the image, leading to defects in the reconstruction. An example of this is shown in Fig. 4.6, where the upper right corner of the image has been zoomed in to visualize the problem more clearly. The image was reconstructed at $\Delta\theta = \pi/2$, similarly to Fig. 4.5b. In Fig. 4.7, additional LORE-GN reconstructions at different $\Delta\theta$ for the
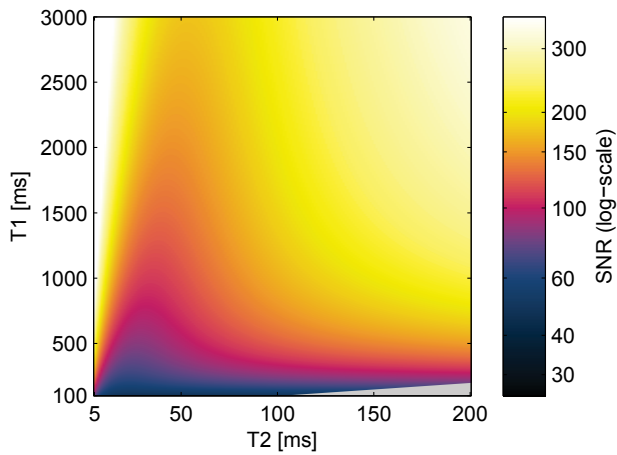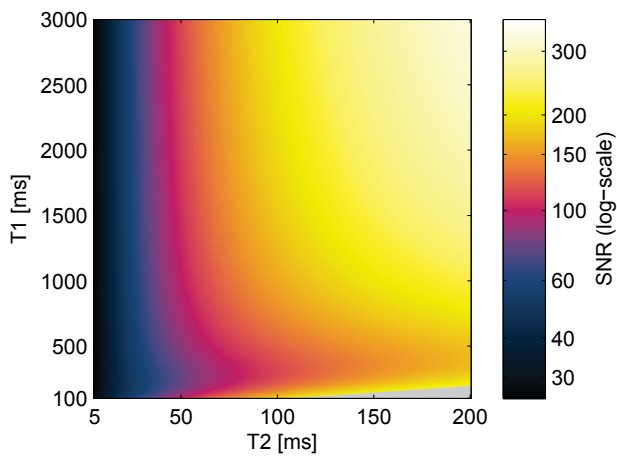
a) Estimating T_1.

b) Estimating T_2.

Figure 4.3: SNR needed to achieve a standard deviation equal to 5% of the true value of the parameters a) T_1, and b) T_2. The values for $T_2 > T_1$ are excluded (gray) and the SNR range is saturated at a maximum of 400.

a) Phase-cycled images.

b) LORE-GN. **c)** Sum-of-squares. **d)** Maximum-intensity. **e)** Cross-solution.

f) $\hat{f}_{OR}$. **g)** Reference f_{OR}.

Figure 4.4: a) 1.5 T phantom dataset of four images with phase increments 0, $\pi/2$, π, $3\pi/2$, b) the reconstructed image with $\Delta\theta = \pi/2$, c) the sum-of-squares image, d) the maximum-intensity image, e) the cross-solution estimate, f) the estimated off-resonance frequency, and g) the reference off-resonance frequency. LORE-GN and the cross-solution show a uniform intensity, while bands still remain in the sum-of-squares and maximum-intensity images. The estimated average SNR of the data was 170. The automatic shimming was disabled to induce more significant banding artifacts. Each image in b)–e) is displayed with the same scale as the original images in a), and all images have been cropped prior to display.

Table 4.1: Average MATLAB run times in seconds for the different methods and the 1.5 T datasets.

	Dataset	
	Phantom	In vivo
LORE	1.5	3.6
LORE-GN	16	40
LM	122	245
LMpost	123	247

1.5 T in-vivo data of 4.5a are shown. As can be seen, the SNR and contrast vary depending on the reconstruction.

Four phase-cycled images from a central slice of the 7 T 3D in-vivo dataset are shown in Fig. 4.8a. Here LORE was used alone to esti-mate the parameters, due to the relatively low SNR. The reason for this will be further explained in the discussion. In Fig. 4.8b the recon-structed image at $\Delta\theta = \pi/2$ is shown, together with the sum-of-squares, maximum-intensity and cross-solution images in Figs. 4.8c, 4.8d and 4.8e, respectively. No banding can be seen with LORE at this low flip angle, while sum-of-squares and maximum-intensity do not fully sup-press the bands, as indicated by the arrows. The cross-solution method has problems due to the low SNR, and the resulting image is severely degraded.

4.4.4 Run times

The run times for the phantom and the in-vivo data are shown in Table 4.1. The current implementation of LORE-GN is approximately 7 times faster than using LM with a fixed initialization. It can also be seen that LORE accounts for less than 10% of the LORE-GN run time, providing a speedup factor of approximately 80 compared to LM. The algebraic post-processing used in LMpost has an insignificant impact on the run time.

4.5 Discussion

The fast linear algorithm LORE is the main contribution in this chapter. In many cases, LORE provides accurate estimates on its own, but it can also be used to initialize a nonlinear algorithm minimizing the NLS criterion. By adding post-processing steps we can also separate between the several optima that are inherent to the model, to obtain consistent parameter estimates. Using these estimates, we can reconstruct band-free images through the model in (4.5). As mentioned in 4.2.4, the

a) Phase-cycled images.
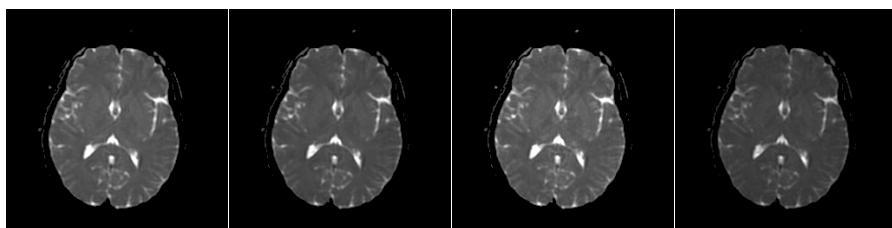
b) LORE-GN. **c)** Sum-of-squares. **d)** Maximum-intensity. **e)** Cross-solution.
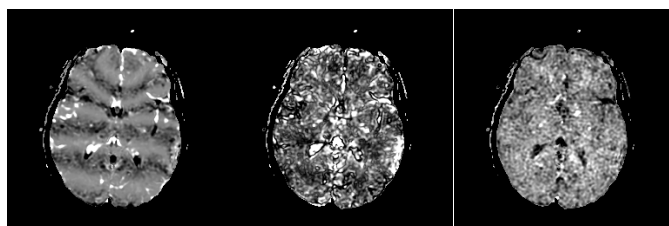
f) Difference image **g)** Difference image **h)** Difference image
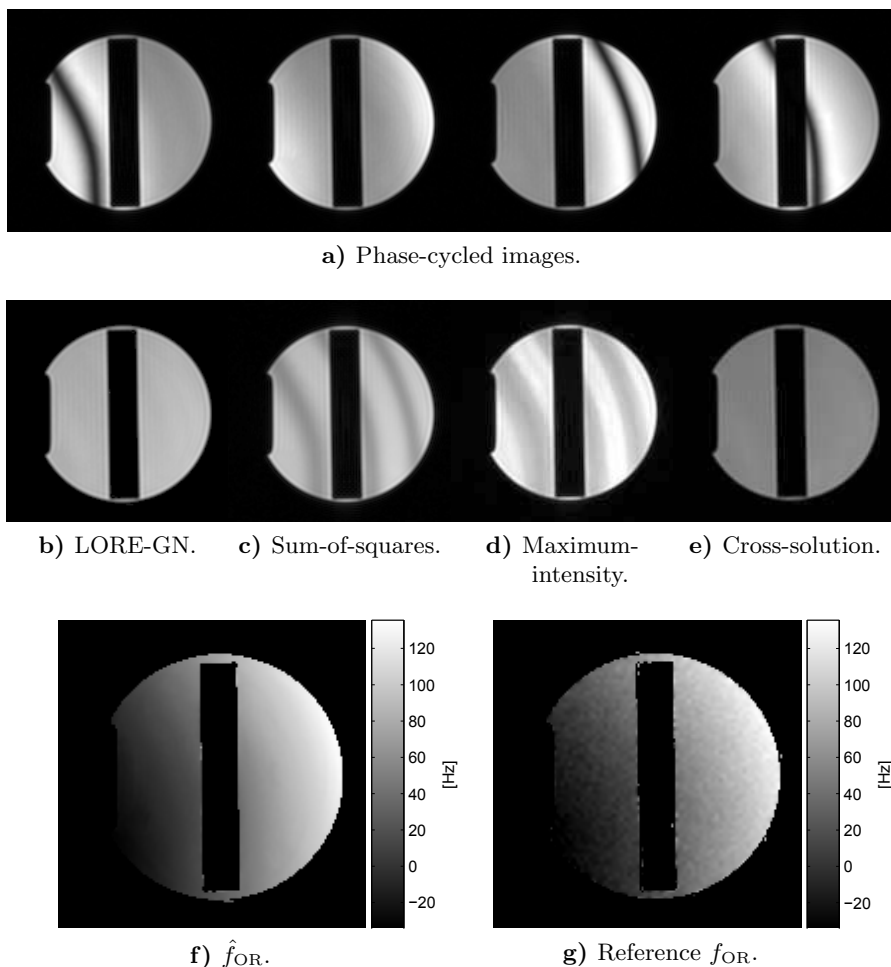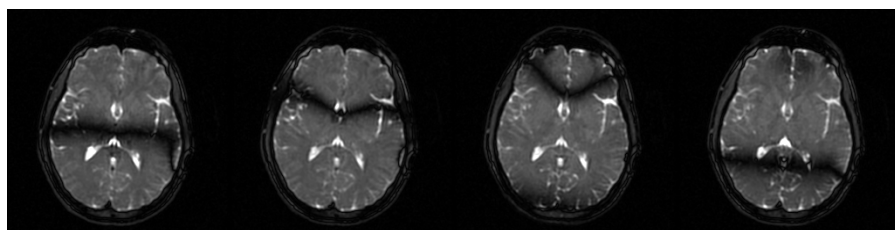c)−b) d)−b) e)−b)

Figure 4.5: a) 1.5 T in-vivo brain dataset of four images with phase increments 0, $\pi/2$, π, $3\pi/2$; b) the reconstructed image with $\Delta\theta = \pi/2$; c) the sum-of-squares image; d) the maximum-intensity image; e) the estimate obtained with the cross-solution. The relative difference between the proposed estimate and: f) sum-of-squares, g) maximum-intensity, and h) cross-solution; where the difference in average intensity has been removed, shows that some bands remain in the sum-of-squares image. The estimated average SNR was 33. The automatic shimming was disabled to induce more significant banding artifacts. Each image in b)–e) is displayed with the same scale as the original images in a).
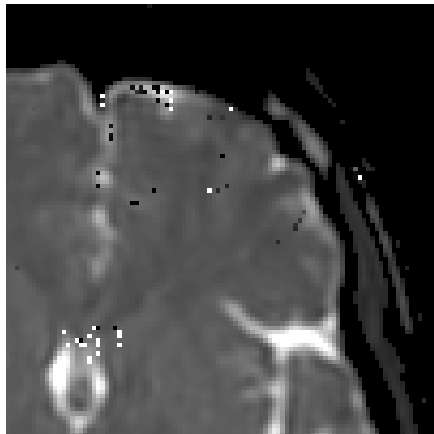
Figure 4.6: A zoomed-in example of the $\Delta\theta = \pi/2$ reconstruction from LM with constant initialization when applied to the 1.5 T in-vivo data of Fig. 4.5a. The reconstruction has defects due to convergence issues of the LM algorithm.

reconstructed data is the ML estimate of the true signal, under the Gaussian assumption, and is hence minimally distorted by noise.

4.5.1 Simulations and the CRB

The case shown in Fig. 4.1 was chosen to illustrate when LM fails due to wrong initialization of θ. This leads LM into an ambiguous optimum, which is the reason for the poor performance where the rMSE is high and remains constant when increasing the SNR. The suggested post-processing mostly corrects for this, however, noise minima can occur at low SNR, making the initialization increasingly important. This can be seen in Fig. 4.1 as an increased rMSE for all parameters when using LMpost compared to LORE-GN at low SNR.

LORE and LORE-GN show robustness at low SNR while the other methods sometimes give outliers. This is, for example, seen in Fig. 4.1d where the LORE rMSE is even slightly below the CRB at low SNR. This is possible due to the biased estimates provided by LORE. However, the problem gets quite sensitive at low SNR when only four phase-cycled images are used, which can lead to outliers. For LORE-GN, this is mainly a problem when estimating T_1, T_2 and KM_0, which is why outliers were removed when generating Fig. 4.2. It can be seen in Fig. 4.2 that the rMSEs at low SNR are typically higher for LORE-GN than LORE. This can be explained, as GN involves the inversion of a matrix, and for high noise levels, this matrix could potentially be close-to-singular. Therefore, it is advisable to use LORE alone in these cases, or to apply a gradient-based method instead of GN, which does

$$\Delta\theta = 0 \qquad\qquad \Delta\theta = \pi/4$$

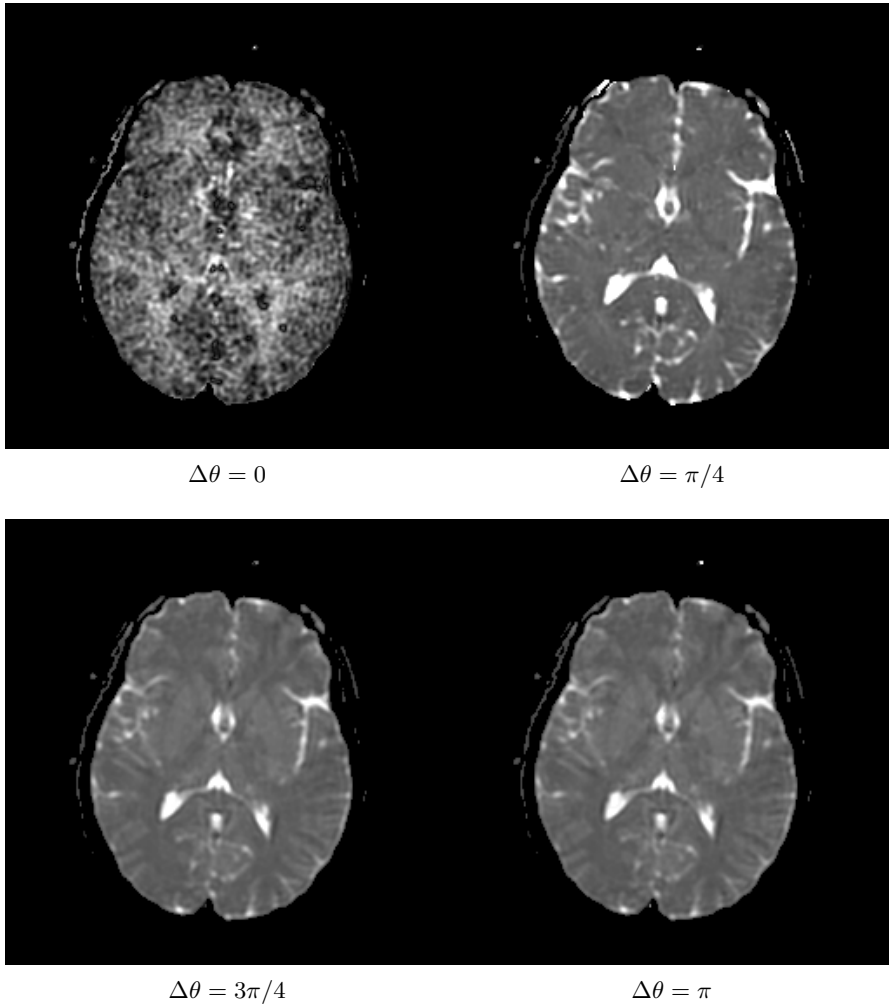$$\Delta\theta = 3\pi/4 \qquad\qquad \Delta\theta = \pi$$

Figure 4.7: Reconstructions with different $\Delta\theta$ using the LORE-GN estimates obtained from the 1.5 T in-vivo dataset of Fig. 4.5a. Each image is displayed in a different scale to make the median intensity comparable. As can be seen, the contrast and SNR of the reconstruction depend on $\Delta\theta$.

a) Phase-cycled images.

b) LORE. **c)** Sum-of-squares. **d)** Maximum-intensity. **e)** Cross-solution.

f) LORE (zoomed-in). **g)** Sum-of-squares (zoomed-in). **h)** Maximum-intensity (zoomed-in). **i)** Cross-solution (zoomed-in).
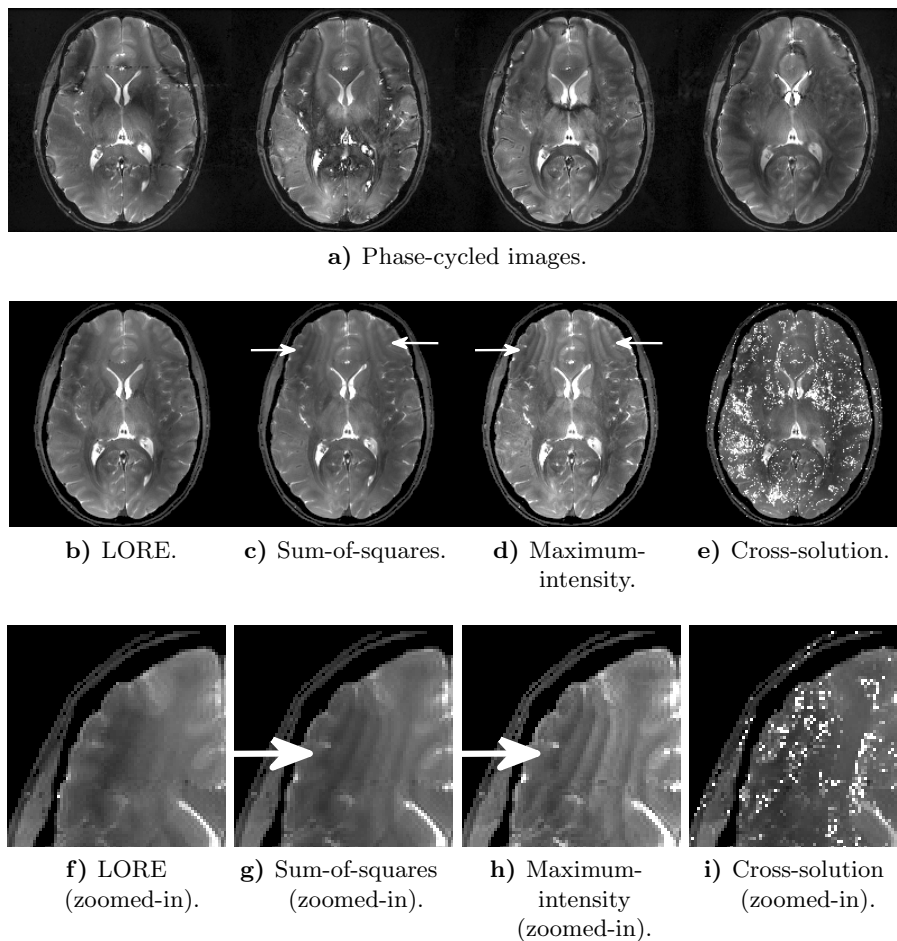
Figure 4.8: a) 7 T in-vivo brain dataset of four images with phase increments 0, $\pi/2$, π, $3\pi/2$; b) the LORE reconstructed image with $\Delta\theta = \pi/2$; c) the sum-of-squares image; d) the maximum-intensity image; e) the estimate obtained with the cross-solution. f)–i) shows zoomed-in versions of b)–e), focusing on the top left part of the image. The proposed method shows no banding artifacts while the sum-of-squares and maximum-intensity have some remaining bands as indicated by the arrows. The cross-solution fails due to the low SNR, which was estimated to be 11. A high-order shim was used to achieve the best-possible field homogeneity. Each image in b)–i) is displayed with the same scale as the original images in a).

not involve any matrix inversion. However, LORE-GN is advantageous at higher SNR due to its ML formulation.

The difference in signal power and the conditioning of the model are the two major contributors to the variations in Figs. 4.3a and 4.3b for different values of T_1 and T_2. As can be seen, the SNR required to accurately estimate T_1 and T_2 from a set of four phase-cycled bSSFP images is rather high. This means that using purely phase-cycled bSSFP to simultaneously estimate T_1 and T_2 is not a viable approach. It is therefore of little use to derive a method that is efficient at lower SNR. Figure 4.3b indicates that it is easier to estimate T_2 than T_1 in the region of short T_2 values. However, achieving high SNR is harder for short T_2 species. In practice the noise variance will be fixed and the signal magnitude will vary over the image, in this case the decrease in SNR will contribute to an increased CRB at short T_2. It should be noted that the difficulty to estimate T_1 and T_2 explicitly does not affect the quality of the image reconstruction, since these estimates are not used in (4.5).

4.5.2 Phantom example

For the phantom data, the proposed method, Fig. 4.4b, has a clear advantage compared to the sum-of-squares and maximum-intensity images, as shown in Figs. 4.4c and 4.4d. The cross-solution, Fig. 4.4e, provides a uniform intensity similar to the proposed approach. However, since the cross-solution assumes $TE = 0$, it is not strictly valid in this case, and therefore, it provides a biased estimate with a different contrast.

It can also be noted that LORE-GN provides an accurate estimate of the off-resonance, Fig. 4.4f, when comparing to the reference field map in Fig. 4.4g. Field map estimation is an application in its own, and several methods are described in the literature, see for example [46, 75]. Since LORE provides an efficient estimate of the off-resonance, and a rather low variance even at an SNR of 10, it could be a useful method for B_0 field mapping.

4.5.3 In-vivo examples

The main difference when using highly structured in-vivo images, as opposed to the phantom dataset, is the partial volume effects at tissue borders. As can be seen in Fig. 4.5b, the reconstructed image using the LORE-GN estimates seem to have a level of detail similar to the sum-of-squares and maximum-intensity images of Figs. 4.5c and 4.5d, respectively. The reconstructed image shows no bands as opposed to the

sum-of-squares. The banding is further highlighted by the difference image in Fig. 4.5f. However, no clear advantage in terms of banding artifacts can be seen when compared to maximum-intensity and cross-solution, as can be seen in Figs. 4.5g and 4.5h, respectively. Since the contrast, for example between white and gray matter, varies for different phase increments, the maximum-intensity will not provide the largest possible contrast. With the proposed approach, several images with different phase increments $\Delta\theta$ can be reconstructed. There is a trade-off between contrast and SNR in the resulting images; however, by choosing the optimal $\Delta\theta$, it is possible to obtain both a higher contrast and SNR compared to maximum-intensity, while getting similar or superior band reduction. The reconstructions in Fig. 4.7 show that significantly different contrasts and SNRs can be obtained, and in the end, the SNR defines the extent to which these images can be used. For example, $\Delta\theta = 0$ gives a low SNR, while theoretically, this corresponds to the image with the maximum contrast between gray and white matter.

As can be seen in Fig. 4.6, LM with fixed initialization does not converge properly, leading to defects in the reconstructed image. These defects are due to local minima caused by noise, and adding the post-processing does not correct for this. Using LORE, however, solves the problem, which underlines the importance of a proper initialization.

The 7 T dataset shows that the method is applicable at higher field strengths and low flip angles, as the resulting LORE reconstruction in Fig. 4.8b gives superior band reduction compared to the sum-of-squares and maximum-intensity, shown in Figs. 4.8c and 4.8d, respectively. The low flip angle and the resulting narrow pass band of the bSSFP profile is the main reason for the incomplete band reduction provided by the sum-of-squares and maximum-intensity. Furthermore, this low SNR case gives an example where using LORE alone is favorable, as GN can suffer from stability problems. As can be seen, the LORE reconstruction does not suffer from the low SNR artifacts present in the cross-solution estimate shown in Fig. 4.8e.

4.5.4 Run times

LORE-GN is fast due to its simplicity and the few GN iterations needed to converge when the initial value is close to the optimum, as was seen in Table 4.1. The initial value is provided by LORE, which due to its linear formulation is more computationally efficient than the nonlinear counterparts. The run times are approximately a factor of two longer for all algorithms when applied to in-vivo data, compared to the phantom. This can be explained by the difference in SNR. The exact computation time will also depend on the implementation, and optimizations in this

respect are possible. For example a C implementation would significantly reduce the computation time compared to running the code in a MATLAB environment. Furthermore, because of the voxelwise computations, all algorithms can be parallelized on multi-core computers to significantly decrease total run time.

4.5.5 Limitations

One limitation of the current method is that magnetization transfer (MT) is not taken into account. Determining the impact of MT is beyond the scope of this discussion. The effects could be modeled, but adding MT parameters will increase the complexity of the model [51]; and it is unlikely that LORE could be generalized to this case, since it is specific to the model presented here. Furthermore, the number of images would have to be increased, leading to a longer acquisition time.

Imperfections in the slice profile can be the source of poor estimation, especially for 2D datasets. The flip angle is here assumed to be a constant in each voxel; however, in practice the flip angle may vary, and if the constant approximation is poor this can make the model invalid. Here we have only used 3D acquisitions in which the slice profile is approximately constant for the slice of interest. Finally, the model does not take partial volume effects into account, but no problems were observed when applying LORE-GN to the in-vivo images.

4.6 Conclusion

We have successfully minimized off-resonance effects in bSSFP. The LORE-GN algorithm is designed to be general, and can be applied to any phase-cycled bSSFP datasets with three or more images, regardless of TE and TR. For band removal and off-resonance estimation, the flip angle does not have to be known, and the method provides uniform reconstructed images even at low flip angles, where other techniques often fail. The fast linear estimator LORE is user-parameter free, which in turn makes the method simple to use and robust, and it provides rather accurate estimates. Adding a nonlinear optimization step, we can efficiently minimize the NLS and provide reconstructed images with optimal SNR in the ML sense, under the assumption of Gaussian noise. We have also demonstrated that it is inherently difficult to explicitly estimate T_1 and T_2 from pure phase-cycled bSSFP, as the obtained variance is bound to be high at common SNR.

Chapter 5

Multi-component T_2 relaxometry and myelin-water imaging

5.1 Introduction

Relaxometry provides quantitative tissue characterization and is useful in many branches of MRI [126, 87]. Commonly, a single exponential decay is assumed [17, 112, 26], but the data analyzed often consist of a continuous spectrum of decays with several significant components [135]. In cases with high SNR and several images, it is possible to estimate more than one decaying component, using a model that consists of a sum of damped exponentials. This model is commonly used in multi-component T_2 relaxometry based on spin echo data, but there are also other methods based on, for example, steady-state data [40, 78, 39, 140]. There are several practical applications that utilize the multi-component T_2 parameter estimates, such as: quantification of myelin in the brain by the myelin water fraction (MWF) [27, 81, 88, 85, 91, 131, 47, 84], with potential future use in Multiple-Sclerosis treatment; and characterization of breast tissue [56], cartilage [100, 141], and skeletal muscle [102]. Moreover, multi-component exponential fitting is important for T_2 relaxometry in general, as the single-exponential model can lead to significant bias of the T_2 estimates when the underlying data originate from several decaying components [135].

In the brain, a varying number of exponential components is expected, depending on the type of tissue. Typically, two to four significant components are present, and need to be estimated to compute the MWF. To facilitate the estimation, a multi-echo spin-echo sequence is used to sample the T_2 decay curve in time.

Several parameter estimation techniques for the sum-of-exponentials model are available in the literature, cf. [69]. The problem has been

shown to be difficult [29, 57], as the components are typically strongly correlated, and therefore, methods with high sensitivity are needed for accurate estimation. For this purpose, the parametric methods offer superior performance compared to their nonparametric counterparts.

The parametric methods are often based on minimizing the least squares (LS) criterion, due to its well established statistical properties. However, for the exponential estimation problem, the resulting nonlinear minimization is typically difficult due to local minima and ill conditioning [53, 109, 120]. Considering the maximum likelihood estimator may further complicate the optimization, making the global minimization of the criterion function intractable. There are three problems to be tackled: 1) a good initialization is needed for the minimization; 2) given the initialization, the minimization method needs to converge to an acceptable solution, without getting stuck in local minima; and 3) the number of exponential components in the data has to be estimated.

There are several methods for obtaining an initial guess of the parameters available; for example, the total-least-squares-based Prony algorithm [77], used in [25], or the subspace-based method used in [61]. However, these methods have varying performance, and can sometimes provide poor initial values. Furthermore, even with reasonable initial values, many minimization approaches can converge to suboptimal stationary points due to the ill-conditioning of the problem. Finally, many methods for order selection rely on user choices, for example, by setting an input parameter. This makes them more difficult to use in practice, and can also cause systematic errors.

The commonly used non-negative least squares (NNLS) spectrum approach [133, 81, 85, 91, 47, 102], circumvents the numerical problems of 1) and 2) above by gridding the nonlinear parameters. This results in a constrained linear LS problem for estimating the amplitude spectrum. Furthermore, the issue in 3) can be partially avoided by determining the order from the number of peaks in the spectrum. However, NNLS requires interaction with the user to set the grid and the regularization parameter controlling the smoothness of the spectrum. The parameter estimator EASI-SM (Exponential Analysis via System Identification using Steiglitz-McBride), recently studied in [120], is user-parameter free, and avoids spurious local minima and the numerical problems mentioned above by reparameterizing the model. Furthermore, by adding an information-based order selection method [123], the number of components is automatically estimated.

In this work, we analyze the performance of NNLS compared to EASI-SM; derive a performance measure based on the minimum achievable variance of the parameter estimates, as given by the CRB; and show that the NNLS is less efficient than EASI-SM when applied to simulated brain data. As an experimental example, the MWF is estimated by applying

the two algorithms to a multi-echo spin-echo dataset with 32 in-vivo brain images to obtain T_2 estimates. The MWF then is obtained by computing the amount of the myelin water component ($T_2 \approx 20-30$ ms) in relation to the total amount of all components, including, for example, intra- and extra-cellular water ($T_2 \approx 80 - 100$ ms), and cerebrospinal fluid ($T_2 > 1000$ ms) [81, 84, 134].

5.2 Theory

5.2.1 Signal model

Due to local tissue variations, the observed signal from a single voxel, denoted $\bar{s}(t)$, can be modeled as a continuous distribution of T_2 relaxation times:

$$\bar{s}(t) = \int_0^\infty p(T_2)e^{-t/T_2} \, \mathrm{d}T_2, \tag{5.1}$$

where the distribution $p(T_2)$ describes the amplitude corresponding to each T_2 value. Typically, this distribution captures a peak widening compared to an idealistic discrete set of T_2 values. The widening is due to local variations in the tissue under study, and should not be confused with an uncertainty in T_2.

To enable a parametric approach, $p(T_2)$ in (5.1) will have to be parameterized. However, even for a low dimensional parameterization of the T_2 distribution, the estimation problem becomes ill conditioned due to the fact that the data is rather insensitive to the width of the peak. This is illustrated in Fig. 5.1, where a discrete exponential was used to approximate data generated by the continuous model in (5.1), for a single peak around $T_2 = 80$ ms. As can be seen, the relative model error over time (the magnitude of the error between the continuous and discrete model data, divided by the continuous model data) is smaller than the relative noise standard deviation (standard deviation divided by the continuous model data) even at a relatively high SNR of 200 (see the SNR definition in (5.3)). This suggests that a discrete component model [2, 25, 42, 102], will often provide sufficient accuracy. Therefore, we propose to model the intensity over time in a single voxel as

$$s(t_n) = \sum_{m=1}^{M} c_m e^{-t_n/T_{2m}} + w(t_n)$$
$$= g(\boldsymbol{\theta}, t_n) + w(t_n), \quad n = 0, \ldots, N-1, \tag{5.2}$$

where M is the number of discrete exponential components, c_m and T_{2m} are the amplitude and relaxation time of component m, respectively, $w(t_n)$ is the noise, t_n is the n:th sampling time instant, and N is the

total number of samples. Furthermore, we introduce the noise-free signal model $g(\boldsymbol{\theta}, t_n)$, and the corresponding vector of model parameters, $\boldsymbol{\theta} = [c_1, \ldots, c_M, \ T_{21}, \ldots, T_{2M}]^{\mathrm{T}}$.

In the MRI application, the model in (5.2) is typically the result of taking the magnitude of a complex-valued signal, therefore the distribution of $w(t_n)$ is not strictly Gaussian, but Rice distributed. At high SNR, the Gaussian noise assumption holds reasonably well, and enforcing it typically only leads to a small bias in the resulting estimates. However, if the tail of the decay is heavily sampled or the SNR is low, the Rician noise can lead to bias when using LS, and could even cause false detection of slow components [12]. When complex-valued datasets are available, the phase correction methods presented in [12] and [19] could be used to generate Gaussian-distributed real-valued decay data, which would alleviate these two problems; and in the case when only magnitude data is available, one alternative is to estimate a baseline in the data to reduce the bias and the occurrence of spurious components [133]. Therefore, the noise in (5.2), $\{w(t_n)\}_{n=0}^{N-1}$, is here modeled as independent and Gaussian distributed with zero mean and variance σ^2.

Throughout this chapter, the following SNR definition will be used:

$$\mathrm{SNR} = \frac{\sum_{n=0}^{N-1} |g(\boldsymbol{\theta}, t_n)|}{N\sigma}, \tag{5.3}$$

that is, the mean intensity of the signal over the noise standard deviation. This is the standard SNR definition for MR images, however, here it is used in the time domain. Note that this SNR has a significant dependence on T_2, as opposed to using only the first image of the dataset.

Assuming uniform sampling with sampling interval Δt, we can reparameterize the model as

$$s_n = \sum_{m=1}^{M} c_m \lambda_m^n + w_n, \qquad n = 0, \ldots, N-1, \tag{5.4}$$

where we have defined $\lambda_m = e^{-\Delta t / T_{2m}}$, $s_n = s(t_n)$, and $w_n = w(t_n)$. If $t_0 \neq 0$, we can define $c_m' = c_m e^{-t_0/T_{2m}}$, but in the following the prime is omitted for ease of notation. By stacking the samples into a vector $\mathbf{s} = [s_0, \ldots, s_{N-1}]^{\mathrm{T}}$, and defining the Vandermonde matrix

$$\mathbf{V} = \begin{bmatrix} 1 & \cdots & 1 \\ \lambda_1 & \cdots & \lambda_M \\ \vdots & & \vdots \\ \lambda_1^{N-1} & \cdots & \lambda_M^{N-1} \end{bmatrix}, \tag{5.5}$$

we obtain the model in matrix form:

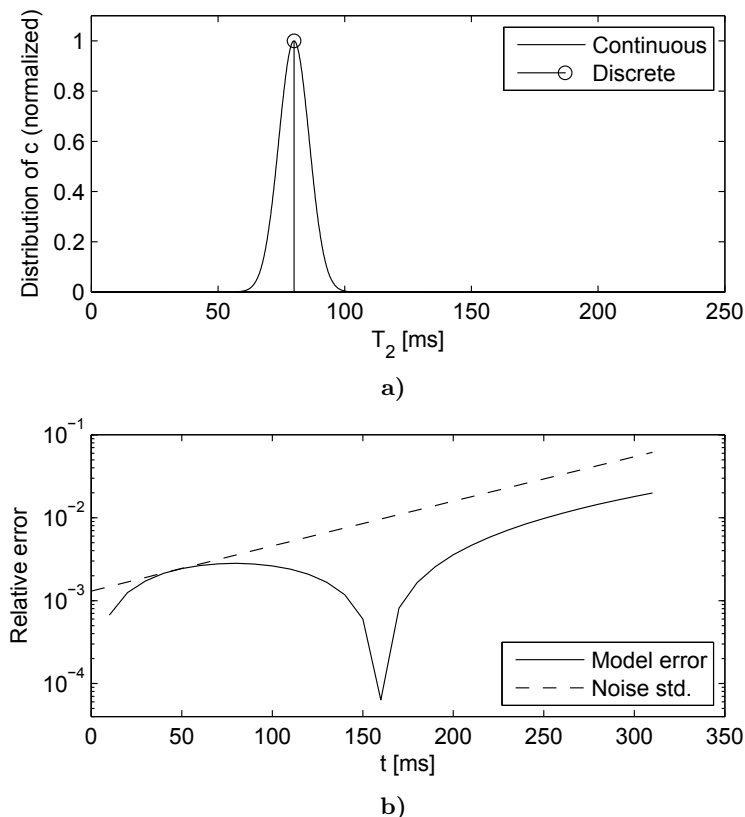$$\mathbf{s} = \mathbf{V}\mathbf{c} + \mathbf{w}. \tag{5.6}$$

Figure 5.1: a) A T_2 distribution with one peak centered at 80 ms, and the corresponding discrete component (normalized), and b) the relative model error versus time when approximating the continuous data generated by (5.1) using a discrete exponential at $T_2 = 80$ ms generated by (5.2), together with the relative noise standard deviation at SNR = 200.

5.2.2 Cramér-Rao Bound

For the discrete component model in (5.2), at any given sampling time t_n, we have the Jacobian vector

$$\frac{\partial g(\boldsymbol{\theta}, t_n)}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \left\{ \frac{\partial g(\boldsymbol{\theta}, t_n)}{\partial c_m} \right\}_{m=1}^M \\ \left\{ \frac{\partial g(\boldsymbol{\theta}, t_n)}{\partial T_{2m}} \right\}_{m=1}^M \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} \left\{ e^{-t_n/T_{2m}} \right\}_{m=1}^M \\ \left\{ c_m \frac{t_n}{T_{2m}^2} e^{-t_n/T_{2m}} \right\}_{m=1}^M \end{bmatrix}^{\mathrm{T}}. \quad (5.7)$$

Substituting (5.7) into (3.16), and using (3.12), gives the CRB matrix $\mathbf{C}_{\mathrm{CRB}}$. The CRB is used in the next section to define a benchmark performance measure at various SNR.

5.2.3 Estimation algorithms

The estimation problem is usually formulated as a LS fitting, although other formulations are possible [69, 112, 26, 61]. By using (5.4) we get

$$\underset{\{c_m, \lambda_m\}_{m=1}^{M}}{\text{minimize}} \sum_{n=0}^{N-1} \left(s_n - \sum_{m=1}^{M} c_m \lambda_m^n \right)^2, \tag{5.8}$$

which is a highly nonlinear problem. Under the assumption of Gaussian noise, (5.8) gives the maximum likelihood estimates of $\{c_m, \lambda_m\}_{m=1}^{M}$; however, solving the problem by finding the global minimum is rather difficult in general.

EASI-SM

This method is based on the Steiglitz-McBride (SM) algorithm [118, 124], originally suggested for estimating the parameters of linear systems. It was introduced for the purpose of estimating the parameters of sum-of-exponential models in [120]. It approximates the LS solution, while avoiding some of the drawbacks associated with minimizing nonlinear functions, like getting stuck in spurious local minima. A more detailed analysis of the convergence properties of SM is given in, for example, [124].

The data can be viewed as the impulse response of a discrete-time linear system. Let the polynomials

$$A(q^{-1}) = 1 + a_1 q^{-1} + \ldots + a_M q^{-M}, \tag{5.9}$$
$$B(q^{-1}) = b_0 + b_1 q^{-1} + \ldots + b_{M-1} q^{-M+1}, \tag{5.10}$$

be defined through

$$\frac{B(q^{-1})}{A(q^{-1})} = \sum_{m=1}^{M} \frac{c_m}{1 - \lambda_m q^{-1}}, \tag{5.11}$$

where q^{-1} is the unit delay operator, that is, $q^{-1} s_n = s_{n-1}$. By using this polynomial parameterization, we can rewrite the problem in (5.8) as

$$\underset{\{a_m\}_{m=1}^{M}, \{b_m\}_{m=0}^{M-1}}{\text{minimize}} \sum_{n=0}^{N-1} \left(s_n - \frac{B(q^{-1})}{A(q^{-1})} u_n \right)^2, \tag{5.12}$$

where u_n, is a unit impulse, that is, $u_n = 1$ for $n = 0$ and zero otherwise ($n = 1, \ldots, N-1$), and the quotient of $B(q^{-1})$ and $A(q^{-1})$ represents a linear dynamical system, or a filter. The SM algorithm for approximately solving the problem in (5.12) can be described as follows. Let $\hat{A}(q^{-1})$ be an estimate of $A(q^{-1})$ (to start we set $\hat{A}(q^{-1}) = 1$); compute

---

Algorithm 5.1: MATLAB implementation of EASI-SM

1: **Inputs:** 3D image matrix: 'im'
       model order: 'M'
       sampling interval: 'dt'
       number of iterations: 'maxIter'
2: **for each** voxel **do**
3:    Extract time signal vector 's' from 3D matrix 'im'
4:    [b, a] = `stmcb`(s, M-1, M, maxIter);
5:    [c_hat, lambda_hat] = `residue`(b, a);
6:    T2_hat=−dt./log(lambda_hat);
7: **end for**

---

new estimates of $A(q^{-1})$ and $B(q^{-1})$ by solving the linear LS problem:

$$
\underset{\{a_m\}_{m=1}^{M}, \{b_m\}_{m=0}^{M-1}}{\text{minimize}} \sum_{n=0}^{N-1} \left[A(q^{-1}) \left[\frac{1}{\hat{A}(q^{-1})} \right] s_n \right.
$$

$$
\left. -B(q^{-1}) \left[\frac{1}{\hat{A}(q^{-1})} \right] u_n \right]^2. \quad (5.13)
$$

Next, use the so-obtained updated estimate of $A(q^{-1})$ as $\hat{A}(q^{-1})$ and recompute the solution to (5.13). Iterate until some predefined stopping criterion is met. Estimates $\{\hat{\lambda}_m\}_{m=1}^{M}$, and hence $\{\hat{T}_{2m}\}_{m=1}^{M}$, can then be obtained by computing the roots of the polynomial $\hat{A}(q^{-1})$, while $\{\hat{c}_m\}_{m=1}^{M}$ are given by the residues:

$$
\hat{c}_m = (1 - \hat{\lambda}_m q^{-1}) \frac{\hat{B}(q^{-1})}{\hat{A}(q^{-1})} \bigg|_{q=\hat{\lambda}_m}. \quad (5.14)
$$

It should be noted that there is no guarantee that the estimates $\{\hat{\lambda}_m\}_{m=1}^{M}$ satisfy $0 < \hat{\lambda}_m < 1$, or that they are even real-valued. However, practically, this is not a problem at sufficiently high SNR and a properly chosen model order, as there is no frequency component in the real-valued data.

As is shown in Algorithm 5.1, the method is easily implemented in MATLAB using the existing functions `stmcb()` and `residue()`. The MATLAB code for the algorithms presented below is available at: `https://github.com/AAAArcus/multiT2`

EASI-SM Order selection
For the in-vivo data, the number of components varies on a voxel basis, and their true number is unknown. Therefore, a method of automatic

order selection is needed for EASI-SM. Two methods have been considered, one based on the Bayesian Information Criterion (BIC), and a heuristic approach trying to enforce the physical constraints on the parameters by selecting the appropriate order.

Using BIC, the LS criterion is modified to include a penalty on the model order M [123, 122]:

$$\underset{M,\,\{c_m, \lambda_m\}_{m=1}^M}{\text{minimize}} \; N \log \left(\sum_{n=0}^{N-1} \left(s_n - \sum_{m=1}^M c_m \lambda_m^n \right)^2 \right) + 2M \log(N). \quad (5.15)$$

This problem is approximately solved by EASI-SM for different values of M up to some maximum order $M_{\max}$, and the solution providing the minimum criterion value is chosen. In this way, only components that are statistically motivated by the data are included, which promotes parsimony and suppresses spurious estimates due to the noise.

The heuristic method starts at order $M_{\max}$, and reduces the order until all parameters satisfy their physical bounds, that is, $c_m > 0$ and $T_{2m} > 0$ for all $m = 1, \dots, M$. This procedure can be intuitively motivated, but it does not possess the statistical properties of BIC. We call this method the Feasibility-based Order Selection (FOS).

MWF estimation via EASI-SM

The MWF is computed as

$$\text{MWF} = \sum_{\substack{m,\ \text{such that:} \\ T_{2\,\min} < T_{2m} < T_{2\,\max}}} c_m \Bigg/ \sum_m c_m \;, \quad (5.16)$$

where $[T_{2\,\min},\ T_{2\,\max}]$ is the interval containing the myelin water component. Therefore, for the application of T_2 estimation considered here, estimating the amplitude of the myelin water component is of main importance. The sum of all amplitudes, occurring in the denominator of (5.16), is significantly easier to estimate, as it does not depend on the T_2 values of the corresponding components.

Since the problem is ill conditioned, different models yield similar fits to the data. In particular, when a slowly decaying component is present in a dataset with few samples, it is often possible to accurately model the signal using a combination of faster components. This may cause problems for EASI-SM if the slow component is not captured adequately, as this can lead to significant bias in the faster decays, such as the myelin water component. Moreover, as previously mentioned, using magnitude data can in some cases cause detection of spurious slowly decaying components, due to the offset in the tail of the decay introduced by the Rician noise. This offset can be reduced by including

a unknown baseline constant in the estimation. This type of approach was discussed in [133] for NNLS, and in fact, the NNLS implementation in this chapter is implicitly using a baseline component in the form of a very slowly decaying exponential included in the T_2 grid.

EASI-SM can be modified to explicitly account for a constant baseline in the data, corresponding to a slow decay on the current time scale. The LS criterion in (5.12) can then be rewritten as

$$\underset{\{a_m\}_{m=1}^{M-1}, \{b_m\}_{m=0}^{M-2}}{\text{minimize}} \sum_{n=0}^{N-1} \left(s_n - k - \frac{B(q^{-1})}{A(q^{-1})} u_n \right)^2, \qquad (5.17)$$

where the total model order M now also includes the constant k. For a fixed k, an approximation of the LS estimate can be obtained by SM, as before; while for fixed $A(q^{-1})$ and $B(q^{-1})$, an estimate of k is readily obtained as

$$\hat{k} = \frac{1}{N} \sum_{n=0}^{N-1} \left(s_n - \frac{B(q^{-1})}{A(q^{-1})} u_n \right), \qquad (5.18)$$

that is, the mean of the residuals without the constant term. The modified algorithm can be summarized as follows: given an initial guess for k (e.g. $k = 0$), run SM to update $A(q^{-1})$ and $B(q^{-1})$, then update k using (5.18), and iterate these two steps until convergence. By using this modification, it is possible to significantly reduce the bias of the myelin water component in cases with low SNR and few samples.

Non-Negative Least Squares (NNLS)

This approach was proposed in [133], and is based on gridding the non-linear parameter (T_2) and solving for the amplitudes using LS with a non-negativity constraint. Thus NNLS implicitly assumes that there is an exponential component at every point of the grid, with an unknown, possibly zero, magnitude. The resulting, typically underdetermined, problem can be formulated as

$$\begin{aligned} \underset{\tilde{\mathbf{c}}}{\text{minimize}} \quad & \|\mathbf{s} - \tilde{\mathbf{V}}\tilde{\mathbf{c}}\|^2 \\ \text{subject to} \quad & \tilde{\mathbf{c}} \geq \mathbf{0} \end{aligned}, \qquad (5.19)$$

where $\tilde{\mathbf{V}} \in \mathbb{R}^{N \times P}$ is a Vandermonde matrix similar to $\mathbf{V}$ in (5.6), and $\tilde{\mathbf{c}} \in \mathbb{R}^{P \times 1}$ is the vector of corresponding amplitudes, and P is the number of grid points. Due to the structure of the problem and the positivity constraint, solving (5.19) produces a sparse vector of amplitudes, with most elements equal to zero, that can be viewed as a spectrum, or distribution, of the damping time constants. However, due to the noise and the fact that the problem is ill-conditioned, a single exponential component in the data can be accurately fitted by several closely spaced

exponentials with nonzero amplitudes. To avoid this issue, and reduce the variance of the estimates by improving the conditioning of the problem, a regularization term penalizing the first-order differences of the vector $\tilde{\mathbf{c}}$ is commonly used, leading to a smoothing of the spectrum. The resulting problem can be written as

$$
\begin{aligned}
\underset{\tilde{\mathbf{c}}}{\text{minimize}} \quad & \|\mathbf{s} - \tilde{\mathbf{V}}\tilde{\mathbf{c}}\|^2 + \rho\|\mathbf{R}\tilde{\mathbf{c}}\|^2 \\
\text{subject to} \quad & \tilde{\mathbf{c}} \geq \mathbf{0}
\end{aligned}
\quad , \tag{5.20}
$$

where ρ is a regularization parameter that needs to be set by the user, and

$$
\mathbf{R} = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & -1 \end{bmatrix}_{P-1 \times P} . \tag{5.21}
$$

To be able to compare the NNLS estimates with the true parameter values in the simulations, point estimates $\{\hat{c}_m, \hat{T}_{2m}\}_{m=1}^M$ are needed. By determining the T_2 values of the peaks in the spectrum, and re-estimating the corresponding amplitudes $\{c_m\}_{m=1}^M$ by LS, estimates of $\{c_m, T_{2m}\}_{m=1}^M$ can be obtained. When the number of peaks in the spectrum exceeds M, the subset of M peaks giving the best fit to the data is chosen as the final estimates. We will refer to this procedure as subset selection. Note that, in simulation, this subset selection method might give an advantage to NNLS, as it will enforce the correct model order.

It should be noted that the discrete unregularized spectrum given by the solution of (5.19), is more easily compared to discrete estimates provided by EASI-SM; however, the central values of the smooth peaks in the regularized NNLS are typically provides better estimates. Furthermore, at typical SNR, regularized NNLS provides no accurate information regarding the amplitude distribution apart from the peak locations (central values), as the continuous distribution of amplitudes is solely due to the regularization, which in turn is controlled by a user parameter.

5.2.4 Evaluating the parameter estimates

Due to the fact that the ordering of the parameters is arbitrary, there is an ambiguity in the model that complicates the simulations and MSE evaluations. If a component known to exist in the data fails to be estimated, which effectively leads to a lower order M, the estimated parameters can be matched to wrong true values, giving an increased MSE. To avoid this problem, a component-wise metric based on the

CRB is defined by the fraction of estimates $\{\hat{c}_m, \hat{T}_{2m}\}$ falling within an uncertainty region R_m in the $\{c, T\}$-plane, for different noise realizations. More specifically, we use the detection rate D_r as a performance measure, defined as the fraction of realizations where M components are estimated within their corresponding uncertainty regions. Similarly, we can define D_r for detecting $\tilde{M} < M$ components, to illustrate each method's ability for partial detection. The uncertainty regions are defined using the CRB as follows:

$$R_m = \left\{ \hat{c}_m, \hat{T}_{2m} : |c_m - \hat{c}_m| < 3\sigma_{c_m} \wedge |T_{2m} - \hat{T}_{2m}| < 3\sigma_{T_{2m}} \right\}, \quad (5.22)$$

where the standard deviations $\sigma_{T_{2m}}$ and σ_{c_m} are obtained from the corresponding diagonal elements of the matrix $\mathbf{C}_{\mathrm{CRB}}$. However, at low SNR the CRB can be rather large, causing the above rectangles to overlap. This overlap indicates a fundamental resolution problem, yet it would give a fictitiously high D_r at low SNR. To prevent this from happening, we restrict the size of the regions to be maximally $\pm 60\%$ of the true parameter values, giving a fixed rectangle size at low SNR.

The detection rate D_r measures the statistical performance of the estimators by comparing them to fundamental variance limits: at lower SNR it shows the fraction of estimates that are practically acceptable; while at high SNR, it shows the fraction that matches the statistical benchmark given by the CRB. This performance measure is based on the discrete component model of (5.2), but can also be applied to data generated by the continuous model of (5.1) to give an indication of performance even in this under-modeled case. However, since the assumptions made in the derivation of the CRB are not strictly valid in the latter case, statistical efficiency cannot be assessed.

5.3 Methods

5.3.1 Simulation

Monte Carlo (MC) simulations with different noise realizations were performed to compute D_r for the two methods. Data containing three exponential components was generated, using both the continuous and the discrete component model, given by (5.1) and (5.2), respectively. The parameters used for the discrete model were

$$\begin{aligned} T_2 &= [20, \ 80, \ 200] \ \text{ms}, \\ c &= [0.4, \ 1, \ 0.2], \end{aligned} \quad (5.23)$$

inspired by previous research [81, 131]. For the continuous model, the c values were normally distributed around the above T_2 values with variances 9, 36, and 4 ms, respectively. The signal was then normalized

to the same total energy as for the discrete case, i.e. the area of each continuous peak corresponds to the height of the discrete component. The used continuous and discrete T_2 distributions are shown in Fig. 5.2, where the continuous distribution has been normalized to unit amplitude to enable clear visualization. Complex-valued Gaussian noise with variance $2\sigma^2$ (i.e. variance σ^2 for both the real- and imaginary parts) was added to achieve the appropriate SNR, as defined by (5.3). Then, the magnitude was computed to generate Rice-distributed data, aimed at mimicking the in-vivo images. The variance of the magnitude data will then be approximately σ^2 at high SNR.

In this three-component example, $N = 48$ samples with sampling interval $\Delta t = 10$ ms, were generated within the chosen SNR range: $40 - 300$. This range was chosen to enable reliable detection of all three components, while including practical SNR values. The SNR of the in-vivo data was estimated to be 70, and therefore, SNR $= 40$ was deemed to be well within the practical range. Reliable estimation at even lower SNR values requires the use of a lower model order. For the EASI-SM order selection methods, $M_{\max} = 4$ was used as reasonable upper limit of the model order for the given range of SNRs. For NNLS, 500 grid points uniformly spaced on a logarithmic scale from 4 - 5000 ms, were used. By using this nonuniform grid, the number of grid points can be reduced compared to a uniform grid, without sacrificing accuracy, as the variance of $\hat{T}_2$ naturally increases for larger T_2 values. The use of 500 grid points is motivated by statistical analysis, where it is important that the grid is dense compared to the expected spread in the estimates for the investigated range of SNRs. Using a sparser grid caused significant performance loss, in terms of D_r, at higher SNR; however, in situations with low SNR, fewer grid points can be considered with only a minor performance loss. No negative consequences were observed from using this dense grid, however, increasing the number of grid points further can eventually lead to numerical problems.

5.3.2 Data Acquisition

The in-vivo data was collected at the McConnell Brain Imaging Centre of McGill University, using a 1.5T Sonata scanner (Siemens Healthcare, Erlangen, Germany) with a single-channel head coil. A single-slice multi-echo spin-echo sequence with nonselective composite refocusing pulses [84] and a crusher gradient scheme [96] was used. Acquisition parameters: 32 echoes, first echo time $TE = 10$ ms, echo spacing $\Delta TE = 10$ ms, $TR = 3000$ ms, FOV $= 22 \times 22$ cm, slice thickness $= 5$ mm, matrix $= 256 \times 128$, NSA $= 1$. The subject was a 24-year-old healthy female volunteer, scanned with approval from the Ethics Com-
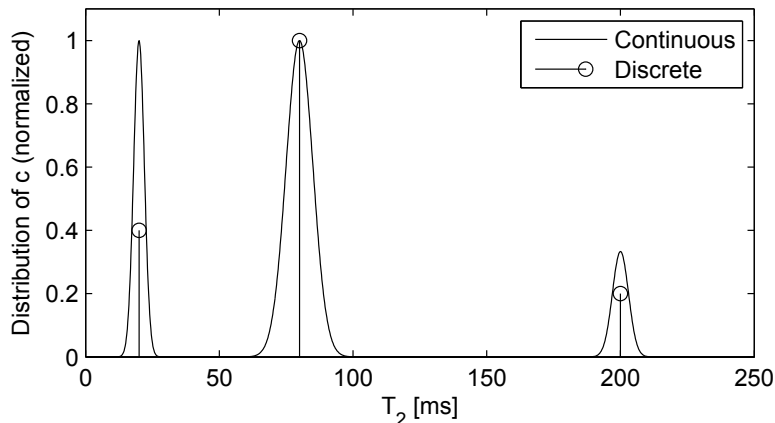
Figure 5.2: Normalized distributions of the amplitudes c over T_2, for both the discrete and continuous models used in the simulations. The area of the continuous peaks correspond to the height of the discrete stems.

mittee of the Montreal Neurological Institute. The total scan time was 26 minutes.

5.4 Results

5.4.1 Simulation

To illustrate the performance measure D_r associated with the estimates produced by the EASI-SM and NNLS algorithms, 200 MC simulations were performed at SNR = 150 and 70. Here, $\rho = 0.05$ was used as described in the next paragraph. The obtained estimates, plotted in the $\{c, T\}$-plane, are shown in Fig. 5.3, together with the true values, and the uncertainty regions. For EASI-SM, most estimates fall within the corresponding uncertainty regions at both SNRs, which corresponds to a high D_r, while the NNLS estimates are generally more spread out. The number of components for EASI-SM was automatically detected by FOS; while for NNLS the three most significant peaks (the true value) in the NNLS spectrum were chosen by the subset-selection method described in 5.2.3, to generate the point estimates.

To find a suitable value for ρ in the regularized NNLS, MC simulations were performed at different SNRs and for a range of ρ values. For each value of ρ considered, D_r was computed, and the results for data generated by the discrete and continuous model, at SNR = 100, are shown in Fig. 5.4. The corresponding detection rates for one and two components were also included, and plotted in an accumulated area plot. In this way, both the proportions of each type of detection, as well as the

probability of detecting, for example, *at least* two components, is shown. As can be seen, the variation in D_r with respect to ρ is similar for both models. For a wide range of SNRs (not shown), setting $\rho = 0.05$ gave an increased performance for detecting all components compared to no regularization; therefore, this value was used in the D_r simulations. For reference, this value of ρ gave 1% average increase in the χ^2 error of the NNLS fitting, compared to using no regularization. Note that choosing ρ to optimize D_r is not possible in practice, and therefore, this approach provides a rather unfair comparison to EASI-SM, in the sense that it favors NNLS.

The D_r obtained from 5000 MC simulations using, both the discrete and continuous data models, is shown in Fig. 5.5. As can be seen, EASI-SM can estimate two or three components in 95% of the realizations for an SNR above 70, while to detect all three components in 95% of the cases, an SNR around 200 is needed. This holds for both the discrete and continuous data. NNLS, on the other hand, estimates all three components only in approximately 60% of the realizations at SNR = 200. At the lowest SNR of 40, the performance of the two algorithms is similar, but the 3-component detection rate of EASI-SM increases more rapidly than for NNLS as the SNR improves.

To illustrate the performance of EASI-SM for the application of MWF estimation, 32 images with a constant MWF were simulated at an SNR of 70, similar to the in-vivo dataset used in the next section. The data was generated by the continuous model of (5.1) using the parameters in (5.23). The results obtained using EASI-SM and NNLS are shown in Fig. 5.6. For this short and relatively low SNR dataset, the baseline estimation approach described around (5.17) was used with EASI-SM. This baseline can capture both very slowly decaying components, whose relaxation time is hard to estimate with high accuracy, and the bias introduced by the Rician noise in the magnitude data. For the regularized NNLS, ρ was chosen to give a 1% increase of the χ^2 fitting term in the criterion, as is commonly done in vivo [2]. The corresponding bias, standard deviation and MSE of the estimated MWFs are listed in Table 5.1. As can be seen, the EASI-SM method has a slightly higher bias in this example, but the lower standard deviation gives a smaller MSE. Furthermore, the order selection methods, FOS and BIC, estimated the model order to be three in 93% and 94% of the cases, respectively; in the remaining cases the order was estimated to be four (not shown here).

5.4.2 In-vivo

Images from the in-vivo dataset at echo times 10, 50, 100, and 200 ms, are shown in Fig. 5.7. As can be seen, there is a flow artifact causing a
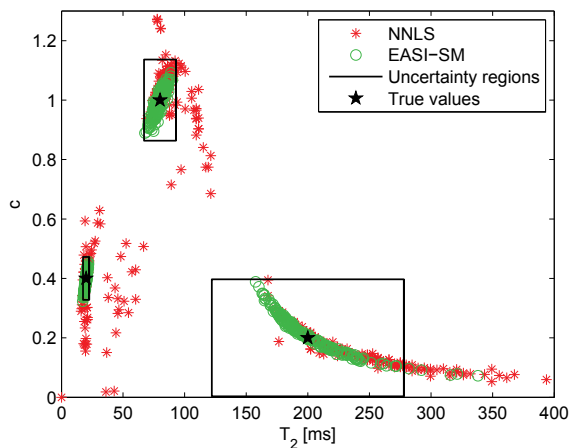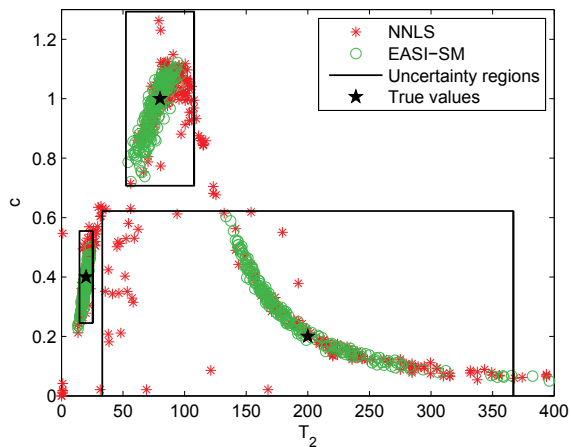
Figure 5.3: Estimates obtained with NNLS ($\rho = 0.05$) and EASI-SM in 200 Monte Carlo simulations at, a) SNR = 150, and b) SNR = 70, along with the corresponding uncertainty regions given by three CRB standard deviations, illustrating the metric used to determine the performance of the investigated algorithms. The true parameter values are indicated by the three stars.

Table 5.1: Bias, standard deviation, and root mean square error of the estimated MWF for the methods in Fig. 5.6.

Method	Bias	Std	rMSE
EASI-SM BIC	0.284	0.026	0.042
EASI-SM FOS	0.283	0.025	0.042
NNLS ($\rho = 0.15$)	0.274	0.044	0.050
NNLS ($\rho = 0$)	0.280	0.077	0.083

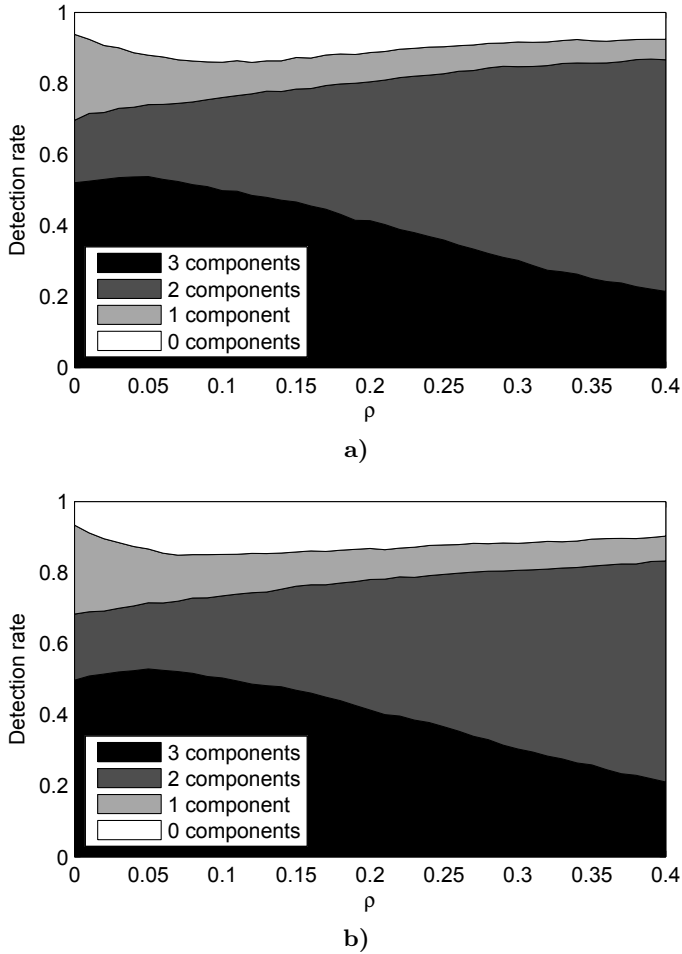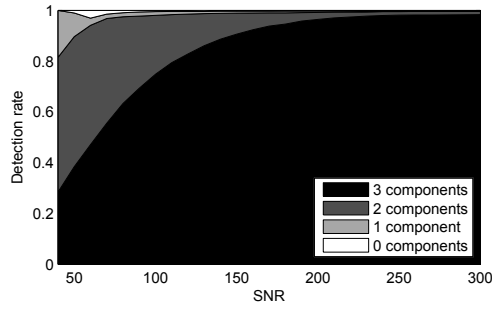**Figure 5.4:** NNLS detection rate at SNR = 100 based on 5000 MC simulations, versus the regularization parameter ρ for a) discrete components, and b) continuously distributed components. The corresponding detection rates for one and two components are also shown accumulatively, to indicate the probability of detecting, for example, *at least* two components.
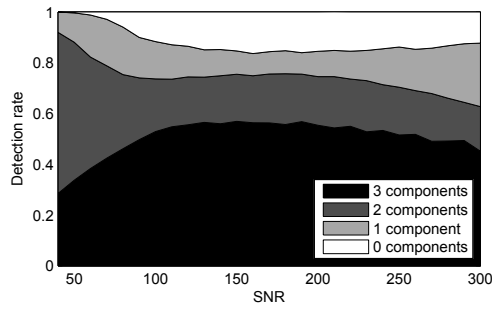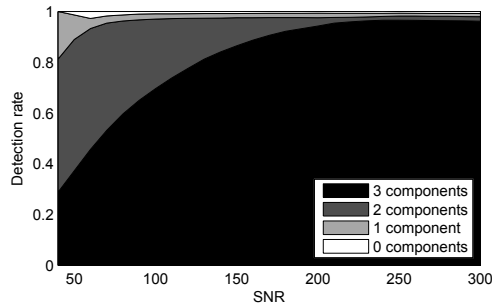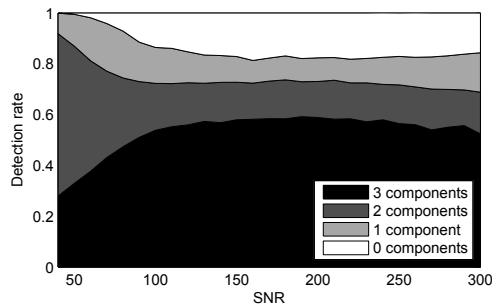
Figure 5.5: Detection rate D_r of the exponential components versus SNR, for a) EASI-SM, and b) NNLS ($\rho = 0.05$), using the discrete model; and the corresponding D_r for the continuous model in c) and d), respectively.
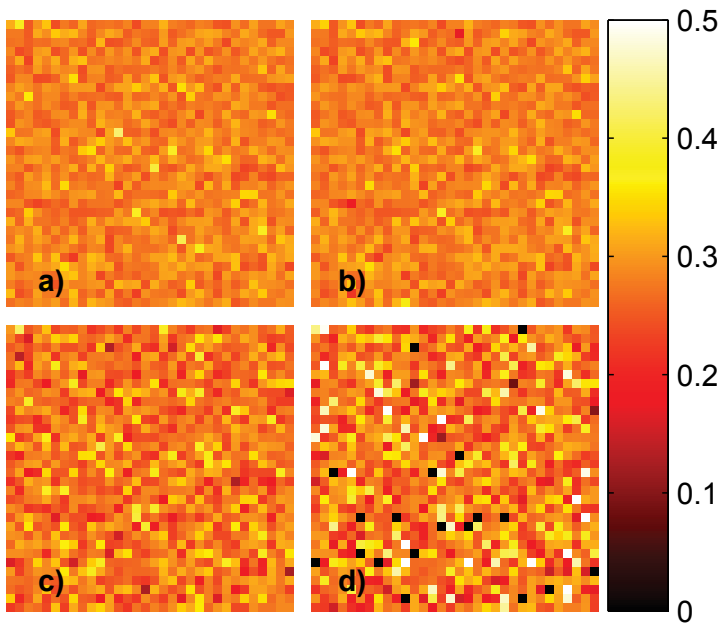
Figure 5.6: Estimated MWF for simulated data using: a) the proposed EASI-SM method with BIC order selection, b) EASI-SM with feasibility-based order selection (FOS), c) NNLS method with regularization ($\rho = 0.15$), and d) NNLS without regularization ($\rho = 0$). The EASI-SM method estimated a baseline in the data. The true value of the MWF was 0.25, and SNR = 70.

horizontal line of deviating values in the center of the image; however, this defect is not essential for the analysis of the methods, and can therefore be disregarded. The average SNR of the dataset was estimated to be approximately 70, by computing the ratio of signal energy to the estimated Gaussian noise standard deviation in the signal-leakage-free background voxels. To reduce the computation time, the low signal background voxels outside the brain were masked prior to applying the algorithms, by excluding background voxels below 20% of the maximum intensity voxel, and the skull.

Using the EASI-SM estimates, the MWF was computed using (5.16). For the NNLS, the estimated spectrum was integrated over the T_2 interval of interest, as in previous works; however, using the corresponding point estimates gave similar results (not shown here). The MWF maps estimated by EASI-SM, using both BIC and FOS for order selection, and by NNLS with $\rho = 0.1$ and $\rho = 0$ (no regularization), are shown in Fig. 5.8. Again, ρ was chosen to give a 1% increase of the χ^2 fitting term in the criterion. As can be seen, EASI-SM generally yields more spatially concentrated MWF estimates compared to NNLS. The difference between using EASI-SM with BIC or FOS is minor. The MWF estimates from NNLS using $\rho = 0$ where also filtered spatially, using a 3×3 Gaussian filter with $\sigma = 0.5$, and compared to the estimates obtained with NNLS using $\rho = 0.1$. The results are shown in Fig. 5.9, and indicate that the regularized method provide MWF estimates that are visually similar to a spatially smoothed version of the unregularized estimates.

To illustrate the T_2 estimation performance in vivo, T_2 maps were generated showing the relaxation time corresponding to the component with the highest estimated amplitude, in each voxel. The results obtained with EASI-SM using BIC and FOS, and NNLS with and without regularization, are shown in Fig. 5.11. As can be seen, the overall T_2 values obtained with EASI-SM and NNLS are similar, but the noise levels in the maps are different. The average T_2 values, and the corresponding standard deviations, in the two 10×10 voxel regions indicated in Fig. 5.11 were approximately 74(3.7) ms and 92(3.5) ms for EASI-SM with BIC, and 74(7.2) ms and 94(6.3) ms for regularized NNLS. As a reference, the single component T_2 estimates, as well as the estimates obtained using a single exponential together with a constant baseline, were computed, and are shown in Fig. 5.12. As can be seen, using a single component leads to significantly longer T_2 relaxation times in some parts of the image, while the estimates obtained when including a baseline are closer to the multi-component estimates of Fig. 5.11.

The estimated EASI-SM model order for each voxel using BIC and FOS is shown in Fig. 5.10, along with the histograms that indicate the relative frequency of each order. As can be seen, regions where the model
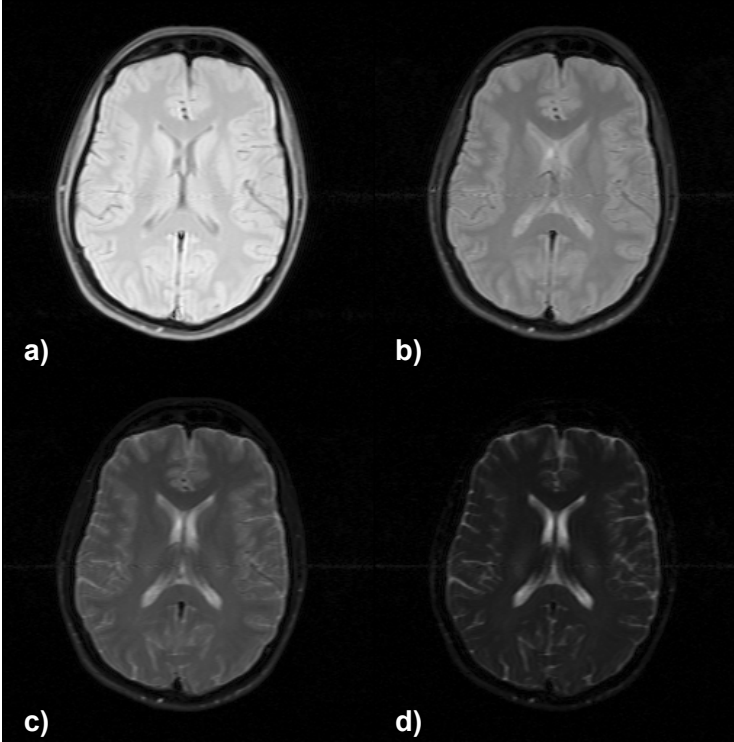
Figure 5.7: Magnitude brain data used for MWF estimation, with an echo spacing of $\Delta t = 10$ ms. Echo time: a) 10 ms, b) 50 ms, c) 100 ms, and d) 200 ms.

order was estimated to be three corresponds well to parts of the brain with a high estimated MWF. Furthermore, BIC typically estimates a higher order than FOS, but the extra components are relatively small, and as shown, the impact on the MWF is minor.

5.5 Discussion

5.5.1 Simulation

For NNLS, the maximum detection rate of all three components occurred when ρ was small. Using a larger ρ causes problems for detection as it introduces bias, and in the worst case, can even make well separated peaks in the spectrum merge into one. The ρ parameter was introduced in [133] to reduce the occurrence of multiple closely spaced spurious components, as well as the variance of the estimates. However, for explicit parameter estimation, the gain from regularization is relatively small compared to setting $\rho = 0$ and choosing a subset of the peaks in the spectrum based on the model fit. Furthermore, when using
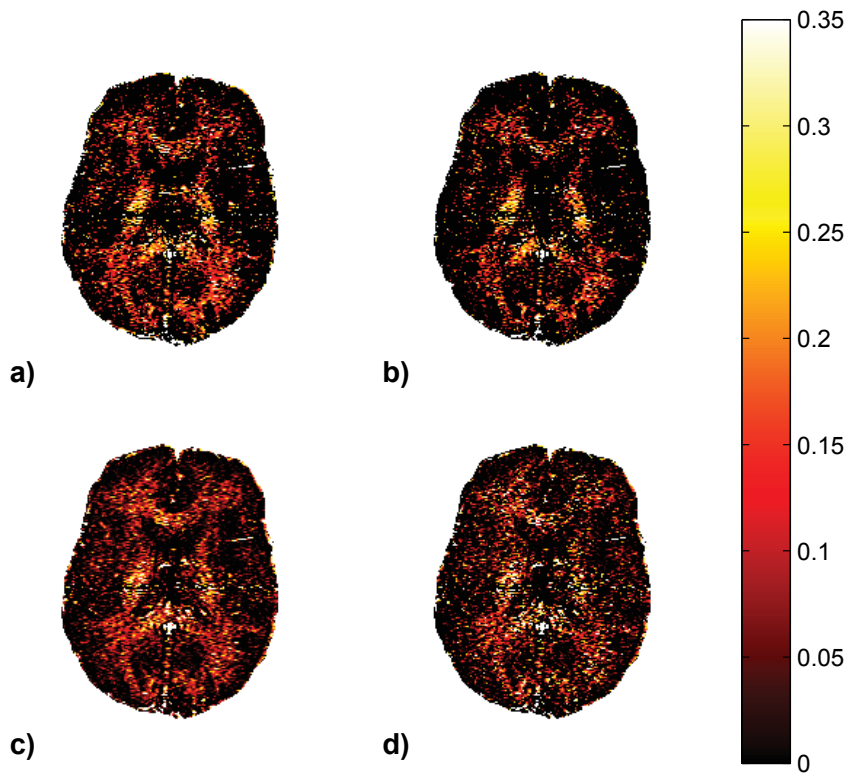
Figure 5.8: Estimated MWF for the in-vivo brain dataset using: a) the proposed EASI-SM method with BIC order selection, b) EASI-SM with feasibility-based order selection (FOS), c) NNLS method with regularization ($\rho = 0.1$), and d) NNLS without regularization ($\rho = 0$). The EASI-SM method estimated a baseline in the data.
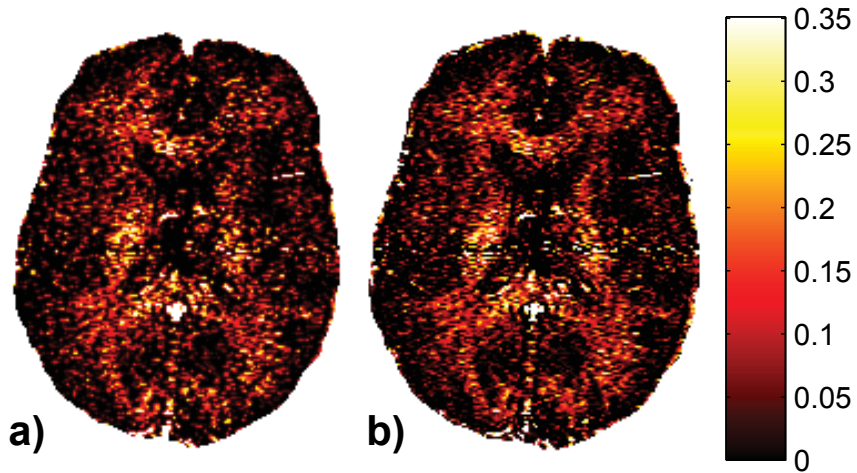
Figure 5.9: a) Gaussian filtered version of the MWF estimates obtained using NNLS without regularization (see Fig. 5.8c), compared to the regularized NNLS method ($\rho = 0.1$) from Fig. 5.8d. The images have been cropped from clarity.
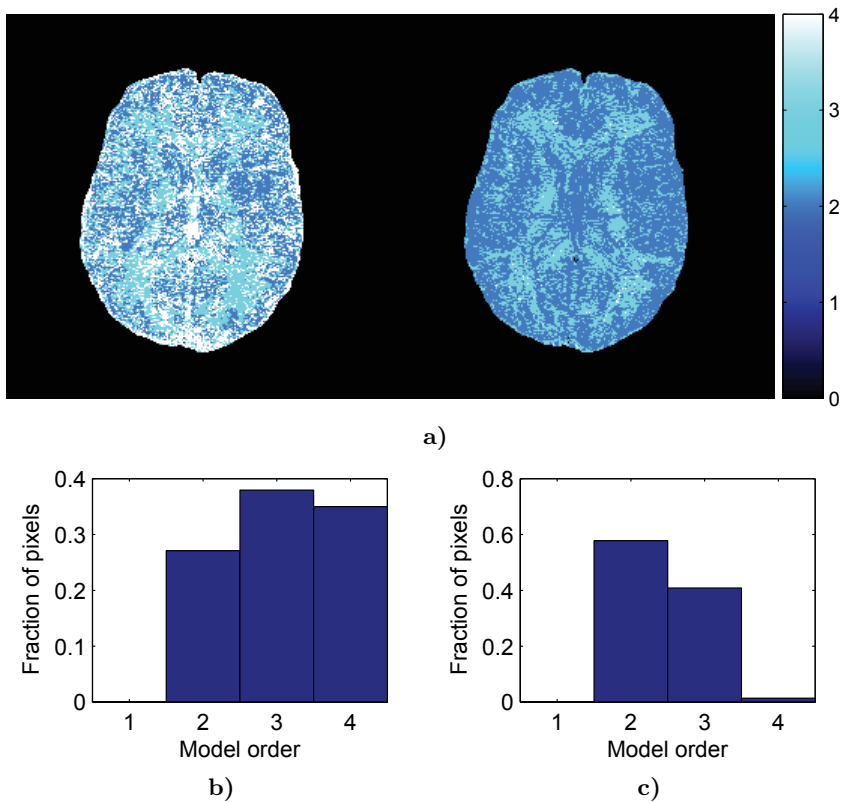
Figure 5.10: a) The estimated order for each voxel of the in-vivo data using EASI-SM with BIC (left), and FOS (right), together with the corresponding relative frequency of each order for b) BIC, and c) FOS.
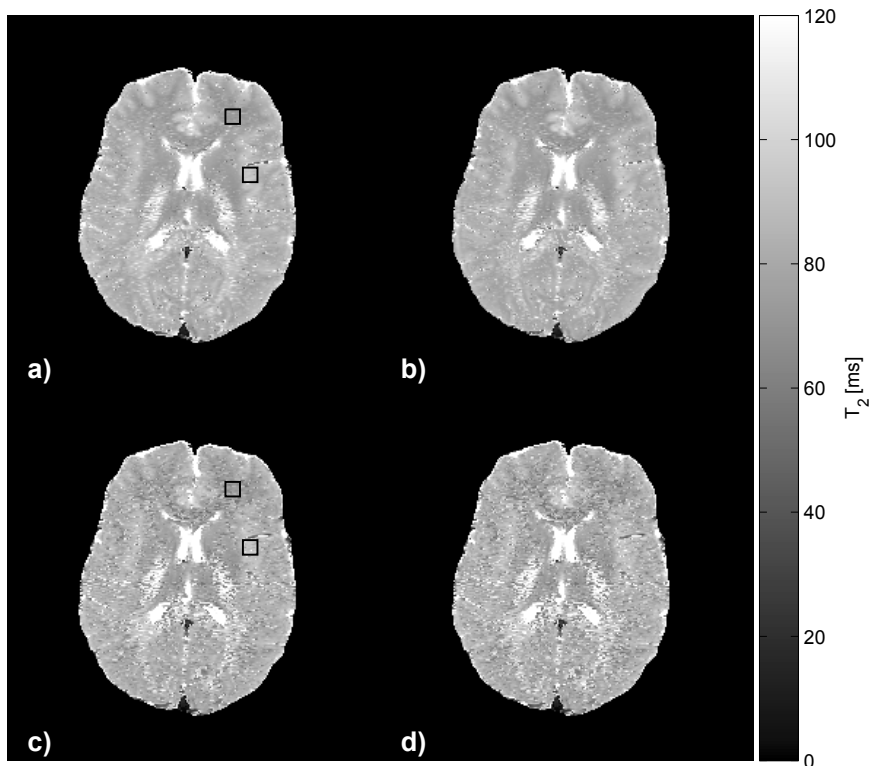
Figure 5.11: Estimated in-vivo T_2 maps corresponding to the most significant decaying component in each voxel, using: a) EASI-SM with BIC order selection, b) EASI-SM with FOS, c) NNLS with regularization ($\rho = 0.1$), and d) NNLS without regularization ($\rho = 0$). The average T_2 values and the corresponding standard deviations in the indicated regions were: a) 74.0(3.7) ms and 92.4(3.5) ms, and c) 73.7(7.2) ms and 93.0(6.3) ms. The EASI-SM method estimated a baseline in the data.
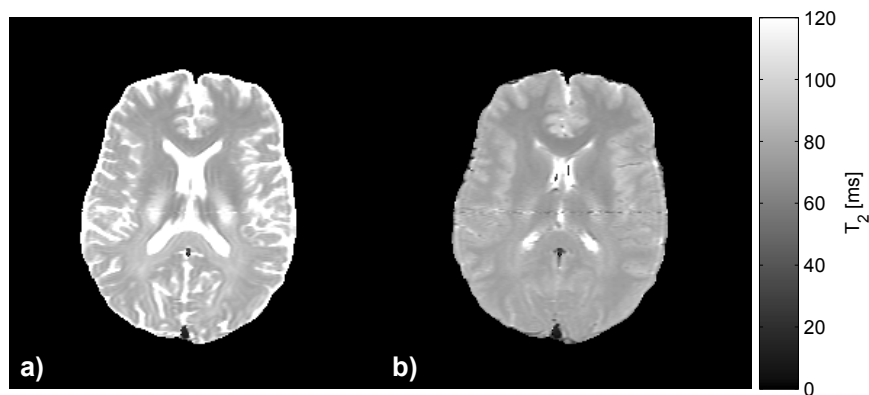
Figure 5.12: Estimated in-vivo T_2 maps using: a) a single exponential component, and b) a single exponential component together with a baseline constant.

regularization, the problem of choosing a suitable ρ for a given dataset remains to be properly solved.

The NNLS has suboptimal performance when compared the CRB, as was indicated by the performance measure D_r in Fig. 5.5. EASI-SM, on the other hand, has a detection rate that improves monotonically as the SNR increases, and even at low SNR it provides a slightly higher detection rate compared to NNLS. Moreover, for the discrete component data, EASI-SM achieves the CRB for all components at high SNR, which is implicit by the fact that D_r for all three components approaches its theoretical maximum. NNLS is not able to efficiently estimate the parameters at any of the investigated SNRs, and at higher SNR the performance has virtually saturated. The slight decrease in D_r that NNLS experiences at high SNR is both due to the use of a relatively large ρ for this SNR, and the T_2 grid being too sparse compared to the variance of the estimates. A denser grid and less regularization could partially cancel this effect (assuming that the need for making these selections could somehow be determined). Furthermore, NNLS has a rather high D_r for two components at low SNR. This can be explained by the constraints being active. More precisely, since $\hat{T}_2$ is fixed on a limited grid, and the amplitudes are positive by construction, there is a limited amount of freedom in the estimates. This, in combination with the large uncertainty regions not taking the positivity of the parameters into account, can cause a high D_r at low SNR for a constrained method, such as NNLS. As SNR increases, however, the 2-component D_r goes down as the uncertainty regions shrink faster than the speed at which the NNLS estimates improve.

The model error associated with using a discrete model to estimate the continuous data leads to some bias, but this has only a small impact on the detection rate, as was shown in Fig. 5.5. This indicates that the discrete model, and the associated methods, are useful even in this case.

The estimate of the order, given by the number of peaks in the NNLS spectrum, can be useful as no separate order selection is needed. However, it can also be a drawback when more peaks than what is physically motivated are found, since there is no reliable way of deciding which ones should be retained. The BIC order selection sometimes includes infeasible estimates, for example, complex-valued T_2 values, which lead to oscillations in the model. The corresponding amplitudes are, however, small in general, as the oscillations are due to fitting the noise. Because of this, EASI-SM with BIC generally achieves better fitting compared to FOS, but the extra parameter estimates are typically of no practical interest. FOS eliminates complex-valued estimates, and therefore usually results in a lower estimated order, as was shown in Fig. 5.10. It is possible to combine these methods and choose the best *feasible* model based on BIC.

For the simulated MWF example in Fig. 5.6, EASI-SM showed superior performance. However, in simulations it is sometimes possible to choose the NNLS regularization parameter, using knowledge of the true signal, to obtain improved results. The problem is how to make this decision in a practical setting, where the true signal parameters are unknown. Moreover, some simulated examples indicated that using a seemingly large value of ρ lead to heavily distorted estimates of the amplitude distribution compared to the true one, but nevertheless resulted in good estimates when computing the MWF using (5.16). It is, however, unclear why NNLS would work in such cases, and exploring it further is beyond the scope of this discussion. However, it should be noted that small perturbations of the true parameters can lead to significant changes in both the minimum MSE and the corresponding optimal ρ.

5.5.2 In-vivo

The 32 image in-vivo dataset is not of high enough SNR to reliably estimate three exponential components, as was done in the simulations. However, as was discussed in 5.2.3, accurate estimation of all components is not necessary to obtain good estimates of the MWF. The regularized NNLS gives a smoother MWF compared to the point estimate-based NNLS approach ($\rho = 0$), and the variance of the estimates is significantly reduced for this relatively low SNR dataset. This can be expected, since the estimated spectrum itself is smoother. For example, a component with $T_2 = 60$ ms can leak energy into the [15, 50] ms range solely due to the regularization used when estimating the spectrum. This can give a false increase of the MWF, as the true value $T_2 = 60$ ms would not normally pertain to myelin water. In turn, this leads to more voxels with small MWFs. It was also noted that the regularized NNLS provided MWF estimates similar to a *spatially* smoothed version of the MWF map obtained by unregularized NNLS, which could be seen from Fig. 5.9. This could indicate a problem with regularized NNLS, as a loss of detail in the final image, similar to a Gaussian smoothing, is typically not desired.

Generally speaking, the EASI-SM with baseline estimation provides similar estimates to NNLS. However, the MWF map estimated by EASI-SM, as displayed in Fig. 5.8, appears to show a more concentrated distribution of myelin in the brain, and a lower noise floor. This is shown as a large proportion of the EASI-SM estimates being identically equal to zero, while the NNLS approach has many scattered estimates throughout the brain with an MWF close to zero. A potential reason for this could be that NNLS is actually detecting myelin water in the gray mat-

ter. However, these low concentrations are difficult to detect given the current image SNR, and since there is no clear structure in the NNLS gray matter estimates, these non-zero values could also be caused by the noise. When comparing to previous research in [91] and [81], similar experimental setups lead to maximum MWFs around 0.10 and 0.20, respectively; the latter value agrees well with the estimates presented herein.

As was seen in Fig. 5.11, the noise levels in the EASI-SM-based T_2 maps are lower than in the NNLS counterparts, leading to a more clear image of the gray and white matter. Furthermore, the computed standard deviation of the T_2 estimates in the two indicated square regions, reflecting the noise level, is almost a factor of two higher for NNLS compared to EASI-SM. This mirrors the T_2 estimation results obtained in simulation, listed in Table 5.1. Overall, the T_2 maps presented herein are significantly different from the one obtained using a single exponential fit, supporting the claim that single-component T_2 relaxometry can lead to significant bias in the resulting estimates, in vivo. The in-vivo T_2 estimation performance of EASI-SM reflects the results obtained in simulation, which indicates that method is practically applicable, and that the simulations made are relevant to evaluate the algorithms.

5.6 Conclusion

The user-parameter free EASI-SM algorithm generally showed superior parameter estimation performance in simulations, when compared to NNLS. At low SNR, however, both methods have similar troubles to estimate all three components in the simulated example, which is explained by the intrinsic difficulty of the problem, and reflected by a high CRB. For the in-vivo application of estimating the MWF in brain tissue, the images resulting from EASI-SM with baseline estimation are similar to those obtained with the regularized NNLS; however, EASI-SM gave a more concentrated distribution of myelin water, and an apparently lower noise level. The lower noise level was also observed in the EASI-SM in-vivo T_2 maps. We believe that using two independent methods and obtaining similar results can increase the confidence in the MWF estimates, and provide a useful cross-check.

Chapter 6

Edge-preserving denoising of T_2 estimates

6.1 Introduction

Quantitative MRI has brought new ways of imaging based on tissue specific physical quantities, such as mapping the longitudinal and transverse relaxation times. One standard approach to estimate T_2 is to acquire several images at different echo times using the spin echo sequence; and then fit a decaying exponential to the magnitude data in each voxel individually [30]. The magnitude images are Rice distributed meaning that the LS estimate will be suboptimal. A few suggestions of how to solve this problem are available in the literature [111, 130], where the authors apply maximum likelihood (ML) methods taking the Rice distribution into account. To further denoise the images or the resulting estimates, techniques like total variation (TV) regularization can be used [74]. The idea is to penalize the total variation in the image, usually quantified by the L_1 norm of some first order difference measure. The resulting estimates tend to be piecewise constant, which is often a good approximation. However, more gradual changes can also be present in the images, which can lead to artifacts. The authors in [74] present a total *generalized* variation regularization that can be used for image denoising while suppressing these so called staircase artifacts. Solving the optimization problem resulting from a TV based approach can, however, be nonlinear and relatively time consuming, and the authors of [74] propose an implementation on a graphics processing unit (GPU). Given that the algorithm can be parallelized efficiently, a GPU implementation gives a significant decrease in computation time; however, this comes at the cost of usability and ease of implementation. Furthermore, the TV-based methods usually come with user parameters that might have to

be chosen iteratively, meaning that the optimization would have to be performed several times.

In this chapter we treat the T_2 estimation problem when only two magnitude images are available, in which case the noise can be a major problem. Applying advanced statistical methods to account for the Rice distributed noise is not likely to be fruitful when only two samples are available in each voxel, especially as the ML properties only hold asymptotically. Moreover, at common signal-to-noise ratios (SNR) the Rice distribution will be accurately approximated by a Gaussian. We focus on the problem of variance reduction of the T_2 estimates. The idea is to improve the estimates by using inter-voxel information in the images, while preserving the contrast between tissues. We propose two methods: 1) a fast local LS method which is easy to implement, and 2) a more general TV based method that can be cast as an LP. Due to the efficiency of solving LPs, this approach is computationally more efficient than the standard TV. Both methods provide an intuitive way to choose the user parameters. We then compare the performance and computation time of the proposed methods to the voxelwise counterpart.

6.2 Theory

6.2.1 Signal model

Here, we solve the estimation problem using magnitude data, which is typical for spin echo T_2 estimation. The resulting data samples are Rice distributed with a PDF given by (3.20). The data at an arbitrary voxel can be modeled as an observation of a Rice distributed stochastic variable parameterized by ρ, T_2, t_n, and σ:

$$s_{\mathrm{R}}(t_n) = S_{\mathrm{R}}(\rho, T_2, \sigma, t_n) \sim \mathrm{Rice}(f(\rho, T_2, t_n), \sigma), \qquad (6.1)$$

where according to (3.19), $f(\rho, T_2, t_n) = \rho e^{-t_n/T_2}$ is a model for the magnitude signal, and $t_n < t_{n+1}, \forall n$ are the echo times of the images.

The SNR is defined as $\mathrm{SNR} = (s_1 + s_2)/2\sigma$, which is the definition commonly used in MRI. The reason for avoiding the root mean square SNR definition here is because it will give less weight to the second sample of the damped signal. At an SNR larger than 20 dB, the Rice distribution is accurately approximated by a Gaussian, as was shown in Fig. 3.2, meaning that least squares (LS) will be close to maximum likelihood. We can then approximately model the data as

$$s(t_n) = \rho e^{-t_n/T_2} + v(t_n), \qquad (6.2)$$

where $v(t_n)$ is i.i.d Gaussian noise. Due to the high complexity associated with using a Rician distribution compared to the accuracy gained,

we propose to use the Gaussian approximation model of (6.2) in all voxels containing signal.

By assuming that there is no noise, the transverse relaxation time T_2 can be estimated from (6.2) by simple algebra. Given that ρ and T_2 are positive and real-valued, we have

$$\lambda \triangleq \frac{s(t_2)}{s(t_1)} = e^{(t_1 - t_2)/T_2}, \tag{6.3}$$

$$\hat{T}_2 = (t_1 - t_2)/\ln(\lambda). \tag{6.4}$$

By applying (6.4) for each voxel p, we can estimate T_2 in an entire image. We will refer to this method as the voxelwise approach. The variable ρ in (6.2) represents the initial magnetization which is proportional to the proton density. Using the found $\hat{T}_2$, an estimate of ρ can be obtained by LS.

The estimate in (6.4) is actually the NLS estimate, obtained by solving the problem

$$\operatorname*{minimize}_{\rho, T_2} \sum_{n=1}^{2} \left(s_n - \rho e^{-t_n/T_2} \right)^2, \tag{6.5}$$

where we have denoted $s_n = s(t_n)$ for ease of notation. This can be seen by inserting the LS estimate of ρ into (6.5) to obtain

$$\operatorname*{minimize}_{T_2} \sum_{n=1}^{2} \left| s_n - \frac{\sum_{j=1}^{2} s_j e^{-\frac{t_j}{T_2}}}{\sum_{j=1}^{2} e^{-\frac{2t_j}{T_2}}} e^{-\frac{t_n}{T_2}} \right|^2, \tag{6.6}$$

which can be factored into

$$\operatorname*{minimize}_{T_2} \left| \frac{e^{-\frac{t_1+t_2}{T_2}}}{e^{-\frac{2t_1}{T_2}} + e^{-\frac{2t_2}{T_2}}} \right|^2 \left(1 + \frac{1}{\lambda^2} \right) |s_2 - \lambda s_1|^2, \tag{6.7}$$

where $\lambda \geq 0$ was defined in (6.3). Since $s_1 \geq 0$ and $s_2 \geq 0$, there is always a λ such that $|s_2 - \lambda s_1|^2 = 0$, hence we can state the equivalent problem

$$\operatorname*{minimize}_{\lambda} (s_2 - \lambda s_1)^2, \tag{6.8}$$

which clearly has the solution in (6.4). The expression in (6.8) can also be obtained by construction of the annihilating filter $G(z) = 1 - \lambda q^{-1}$, where q^{-1} is the backward time shift operator, which gives the output $y = s_2 - \lambda s_1 + v_2 - \lambda v_1$. Furthermore, the annihilating filter shows that by using the formulation in (6.8), the noise will be amplified by a factor $|1 - \lambda q^{-1}| \leq 2$, since $0 \leq \lambda \leq 1$. This will be of interest in Section 6.3, when generalizing this problem to cover several voxels.

6.2.2 Noise variance estimation

To estimate the standard deviation of the Gaussian noise σ from the signal free background of the images, the Rayleigh distribution is used, which is a special case of the Rician distribution in (3.20) when the magnitude of the signal component η is equal to zero. The maximum likelihood estimate of σ is [111]

$$\hat{\sigma} = \sqrt{\frac{1}{4N_b} \sum_{p \in R_b} (s_{1p}^2 + s_{2p}^2)} \,, \tag{6.9}$$

where R_b is the set of all background voxels and N_b is its cardinality. Similarly, we will denote the cardinality of any set R_x as N_x. Note that the corresponding voxels in both images are used, since σ is assumed to be constant. In practice the voxels containing signal are separated from the noise only background by thresholding the first image and keeping voxels above a fixed percentage of the maximum intensity. This procedure defines the set R_b in (6.9), as well as the set of signal voxels R_s used in the proposed algorithms.

6.2.3 The Cramér-Rao bound

Using the Rician probability distribution defined in (3.20), we can write the likelihood function of the samples s_1 and s_2 as

$$\mathcal{L}(s_1, s_2, \rho, T_2) = \prod_{n=1}^{2} p_{\mathrm{R}}(s_n | f(\rho, T_2, t_n), \sigma). \tag{6.10}$$

Substituting (6.10) into (3.11), we obtain the FIM, which is a relatively complicated expression. As opposed to the Gaussian CRB, evaluating the Rician CRB is numerically unstable, and as far as the authors know, an analytical expression is not available. The difference between the Gaussian and Rician CRBs for the T_2 parameter is illustrated in Fig. 6.1. The Rician CRB is initially higher than the Gaussian counterpart, but as the SNR increases they approach each other. Because of this, the Gaussian CRB will be used in the following as a lower bound and a high SNR approximation, of the true CRB.

6.3 Method

By using the definition in (6.3), we can express the problem in (6.8) in matrix form by stacking the data from each image in column vectors $\mathbf{s}_n$:

$$\underset{\boldsymbol{\lambda}}{\text{minimize}} \; \|\mathbf{s}_2 - \mathrm{diag}(\mathbf{s}_1)\boldsymbol{\lambda}\|_2^2, \tag{6.11}$$
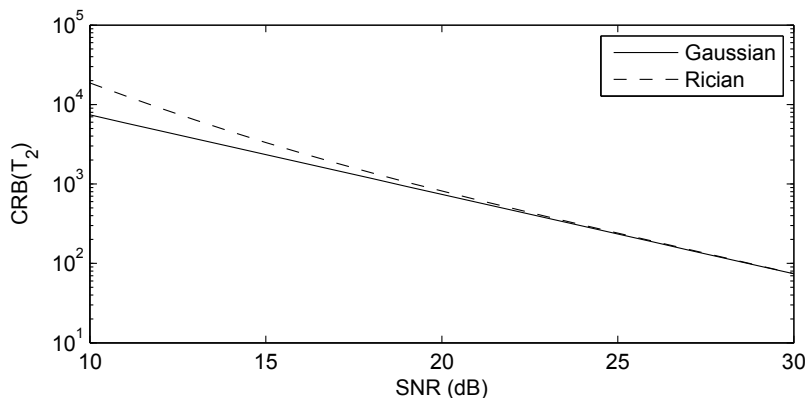
Figure 6.1: Difference between Gaussian and Rician CRB for the special case $\rho = 1$, $T_2 = 70$ ms, $\mathbf{t} = [21, \ 100]^{\mathrm{T}}$ ms

where $\text{diag}(\mathbf{v})$ is a square matrix with the elements of $\mathbf{v}$ along its diagonal.

The problem in (6.11) is redundant in its current form, since there is no coupling between the voxels, meaning that we can solve the problem for each voxel using (6.4). However, if a smoothness constraint is imposed on the estimates, the voxels will be coupled and an optimization formulation becomes vital. In the following, two different coupling approaches and the corresponding estimation algorithms are presented. The goal is to preserve the contrast between different types of tissue, while reducing the MSE of the estimates.

6.3.1 Local Least Squares approach

A simple method to reduce the variance is to assume that locally over a small region around the voxel of interest, the value of T_2 is constant. By using this assumption the estimate of λ in (6.8) can be obtained by linear LS over this region. Here we will avoid smoothing across tissue borders by detecting sudden large changes in the voxelwise T_2 estimates obtained from (6.4). Based on this information we adjust the neighborhood region R_p used in LS (see below). The algorithm then estimates λ, and hence T_2, assigns this estimate to the center voxel only, and then moves to the next voxel. We call this method the localLS.

We can formulate the problem in each region R_p, identified by the center voxel p, as

$$\underset{\lambda_p}{\text{minimize}} \ \sum_{j \in R_p} (s_{2j} - \lambda_p s_{1j})^2, \qquad \forall p \in R_s, \tag{6.12}$$
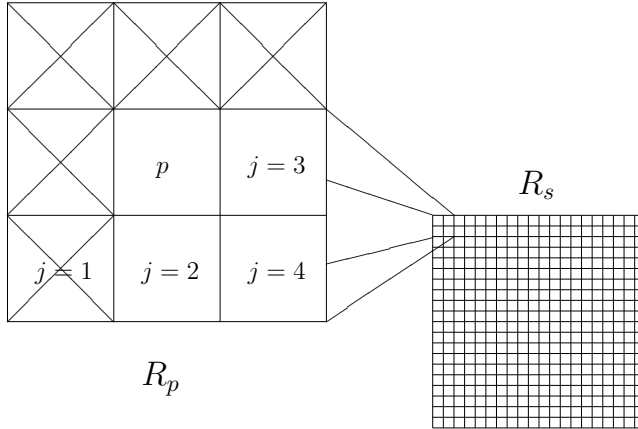
Figure 6.2: Example of the smoothing region R_p at a border voxel p of the signal-containing voxels R_s. The voxels marked $j = 1 \ldots 4$ are the set of unique coupling voxels that are neighbors to p, while the crossed out voxels are outside R_s and will not be used.

where R_s is the set of all voxels containing signal. Denoting the vector of neighborhood data points by $\mathbf{s}_{1p}$ and $\mathbf{s}_{2p}$ we get

$$\hat{\lambda}_p = \frac{\mathbf{s}_{1p}^{\mathrm{T}}\mathbf{s}_{2p}}{\mathbf{s}_{1p}^{\mathrm{T}}\mathbf{s}_{1p}}, \qquad \hat{T}_{2p} = -\frac{t_2 - t_1}{\ln(\hat{\lambda}_p)}. \qquad (6.13)$$

The region R_p is initially defined as the intersection of R_s and the 8 voxels surrounding voxel p; however, to eliminate constraints occurring more than once, only four neighbors are needed. The coupling scheme used, as well as the intersection with R_s, is illustrated in Fig. 6.2. The voxelwise estimates are then computed and checked for feasibility, that is $0 < T_2 < \infty$, and any infeasible voxels are removed from R_p. If the center voxel is infeasible, it is initially set to the median value of the feasible surrounding estimates. The outlier removal procedure then further limits the set R_p. This is done by computing the CRB of T_2 using the center voxel as the true value and the estimated noise standard deviation from (6.9). The CRB gives a lower bound on the parameter standard deviation, which is used as an estimate of the variation in the current area. Any voxel T_2 estimate further away than $k_{\mathrm{LS}}\sigma$ from the center T_2 estimate is considered an outlier. The threshold k_{LS} is a user parameter that will be discussed further in Section 6.4. It is straightforward to choose a different size of the surrounding neighborhood than the 3x3 voxels used here, but a preliminary study showed no clear gain in doing so, and therefore it is here kept constant and not discussed further.

The problem stated in (6.12) is not equivalent to the NLS as in the voxelwise case (6.4). This is due to the re-parameterization using λ

instead of T_2. By expressing the NLS for the region R_p in vector form, where $\mathbf{r}(T_{2p}) = [e^{-t_1/T_{2p}} \; e^{-t_2/T_{2p}}]^{\mathrm{T}}$ and $\tilde{\mathbf{s}}_j = [s_{1j} \; s_{2j}]^{\mathrm{T}}$, we get

$$\underset{\rho_j, T_{2p}}{\text{minimize}} \; \sum_j \|\tilde{\mathbf{s}}_j - \rho_j \mathbf{r}(T_{2p})\|_2^2. \tag{6.14}$$

By LS we have that for each j

$$\rho_j(T_{2p}) = \frac{\mathbf{r}^{\mathrm{T}}(T_{2p})\tilde{\mathbf{s}}_j}{\mathbf{r}^{\mathrm{T}}(T_{2p})\mathbf{r}(T_{2p})}. \tag{6.15}$$

Substituting this into (6.14) we can rewrite the NLS criterion as

$$\underset{T_{2p}}{\text{minimize}} \; \sum_{j \in R_p} \tilde{\mathbf{s}}_j^{\mathrm{T}} \left(I - \frac{\mathbf{r}(T_{2p})\mathbf{r}^{\mathrm{T}}(T_{2p})}{\mathbf{r}^{\mathrm{T}}(T_{2p})\mathbf{r}(T_{2p})} \right) \tilde{\mathbf{s}}_j, \tag{6.16}$$

or by introducing

$$A(T_{2p}) = I - \frac{\mathbf{r}(T_{2p})\mathbf{r}^{\mathrm{T}}(T_{2p})}{\mathbf{r}^{\mathrm{T}}(T_{2p})\mathbf{r}(T_{2p})}, \tag{6.17}$$

$$S = \begin{bmatrix} s_{11} & s_{12} & \dots & s_{1N_p} \\ s_{21} & s_{22} & \dots & s_{2N_p} \end{bmatrix}, \tag{6.18}$$

we get

$$\underset{T_{2p}}{\text{minimize}} \; \text{tr}\left(S^{\mathrm{T}} A(T_{2p}) S \right). \tag{6.19}$$

By using the estimates obtained form (6.13) as an initial guess, a nonlinear minimization method can be applied to solve the problem in (6.14) and obtain the NLS estimate of T_2 in the region R_p. The method obtained by adding this NLS tuning step is abbreviated as the localNLS. For implementation, A can be precomputed on a given grid. Also, since the trace is invariant under cyclic permutations, that is, $\text{tr}(S^{\mathrm{T}}A(T_{2p})S) = \text{tr}\left(A(T_{2p})SS^{\mathrm{T}} \right)$, where the later is only a 2x2 matrix, it is possible to reduce number of flops, and the computation time.

6.3.2 L$_1$ Total Variation approach

In this approach we suggest imposing a global smoothing criterion such that each estimate of T_2 is *close* to the neighboring estimates. The problem is typically formulated as a quadratic fitting with a total variation constraint based on the L$_1$ norm. However, this leads to a quadratic program, which can be quite demanding to solve for the problem size at hand. Therefore we propose to use a L$_1$-norm fitting, leading to an LP that can be efficiently solved. Furthermore, we formulate the problem as a minimization of the total variation, subject to a bound on the L1

norm of the error. Using this formulation, the right hand side of the constraint can be estimated by applying the Cauchy-Schwarz inequality to the error vector $\boldsymbol{\epsilon}$:

$$\|\boldsymbol{\epsilon}\|_1 = \mathbf{1}^{\mathrm{T}}|\boldsymbol{\epsilon}| \le \|\boldsymbol{\epsilon}\|_2\|\mathbf{1}\|_2 = \hat{\sigma}N_s, \tag{6.20}$$

where the last equality holds if the errors have zero mean. The problem can then be formulated as

$$
\begin{aligned}
\underset{\boldsymbol{\lambda}}{\text{minimize}} \quad & \sum_{p \in R_s}\sum_{j \in R_p} |\lambda_p - \lambda_j| \\
\text{subject to} \quad & \|\mathbf{s}_2 - \mathrm{diag}(\mathbf{s}_1)\boldsymbol{\lambda}\|_1 \le k_{\mathrm{TV}}\hat{\sigma}N_s \\
& \mathbf{0} < \boldsymbol{\lambda} < \mathbf{1}
\end{aligned} \tag{6.21}
$$

where $\mathbf{0}$ and $\mathbf{1}$ are the column vectors of appropriate length containing zeros and ones, respectively, and $\boldsymbol{\lambda}$ is the vector of all $\lambda_p \in R_s$. The proportionality constant k_{TV} is a user parameter that will be discussed further in Section 6.4. The use of the L$_1$ norm allows for piecewise smoothness, and does not penalize sudden jumps in the estimate as severely as the L$_2$ norm. We also require that T_2 is positive and finite, corresponding to $\mathbf{0} < \boldsymbol{\lambda} < \mathbf{1}$. It is possible to set stricter bounds on T_2 using prior information, but this option will not be pursued here. By introducing two new variables $\mathbf{w}$ and $\mathbf{z}$, we can write the problem in (6.21) as

$$
\begin{aligned}
\underset{\boldsymbol{\lambda},\mathbf{z},\mathbf{w}}{\text{minimize}} \quad & \mathbf{1}^{\mathrm{T}}z \\
\text{subject to} \quad & -\mathbf{z} \le \left(\mathrm{bdiag}\left(\{\mathbf{1}_{N_p}\}_{p=1}^{P}\right) - B\right)\boldsymbol{\lambda} \le \mathbf{z} \\
& -\mathbf{w} \le \mathbf{s}_2 - \mathrm{diag}(\mathbf{s}_1)\boldsymbol{\lambda} \le \mathbf{w} \\
& \mathbf{1}^{\mathrm{T}}\mathbf{w} \le k_{\mathrm{TV}}\sigma N_s \\
& \mathbf{0} < \boldsymbol{\lambda} < \mathbf{1}
\end{aligned} \tag{6.22}
$$

where the B matrix specifies the unique neighbors of each voxel, and $\mathrm{bdiag}\left(\{\mathbf{1}_{N_p}\}_{p=1}^{P}\right)$ is the block diagonal matrix with P blocks, and the vector $\mathbf{1}_{N_p}$ of length N_p in block p. The problem in (6.22) is an LP [28], and solving it for $\boldsymbol{\lambda}$ gives the variance reduced estimates of T_2. We call this method L1TV.

6.4 Results

6.4.1 Simulations

Empirically, it can be shown that the optimal echo times for estimating T_2, in terms of the CRB, can be accurately described by the equation

$t_2 = 1.1T_2 + t_1$. This relies on the Gaussian assumption, and hence, only holds approximately for high SNR. In practice, the true T_2 is not known and will vary over the image, but the tissue of interest can still provide an approximate expected T_2. In terms of actual CRB values, choosing t_1 as low as possible is favorable, which can be expected since the signal decays in time.

A simulated dataset with Rician noise was created to be able to quantitatively compare the performance of the proposed methods to the voxelwise approach. The echo times were $t_1 = 21$ ms and $t_2 = 100$ ms, chosen to mimic the in-vivo data in Section 6.4.2. A range of T_2 values from 50 ms to 300 ms, and several types of T_2 variations: smooth gradients, large and small abrupt changes, fine detail and larger constant areas, were included in the dataset; while ρ was kept constant, as it only affects the SNR. The two raw data images are shown in Fig. 6.3. The total mean square error (MSE) was computed for different values of the user parameters k_{LS} and k_{TV}, and is shown in Fig. 6.4. As can be seen the MSE is minimized for $k_{\mathrm{LS}} > 2.5$ in the localLS method, and L1TV has a clear minimum around $k_{\mathrm{TV}} = 0.8$. The results were similar for different SNRs, although a small variation could be seen. As a compromise between visual appearance, bias, and MSE, default values of $k_{\mathrm{LS}} = 2$ and $k_{\mathrm{TV}} = 0.5$ were set for the localLS and L1TV respectively. It can also be seen from Fig. 6.4 that the localNLS has similar performance to localLS. Visually the results where indistinguishable and the localNLS will therefore not be studied in the following.

The T_2 estimates obtained from localLS and L1TV using the found values of k_{LS} and k_{TV}, are shown in Fig. 6.5. The localLS and L1TV have similar performance in terms of MSE, and show a clear reduction in the variance. However, the L1TV has a slight problem resolving the thin bright lines on the dark background compared to localLS. The error histograms for each method are shown in Fig. 6.6. As can be seen, the errors are generally smaller for L1TV than for localLS, even though the total MSE is slightly higher. This can be explained by a few larger errors that are not shown in the histogram, which have a significant impact on the MSE.

6.4.2 In-vivo data

Two spin echo images of a brain at echo times $t_1 = 21$ ms and $t_2 = 100$ ms where acquired using a 1.5 T scanner. The average image SNR was 31 dB, meaning that the Gaussian assumption should be valid. By employing the user parameters suggested from the simulations, the T_2 estimates were calculated using the proposed methods, localLS and L1TV, and compared to the voxelwise approach. The results are shown
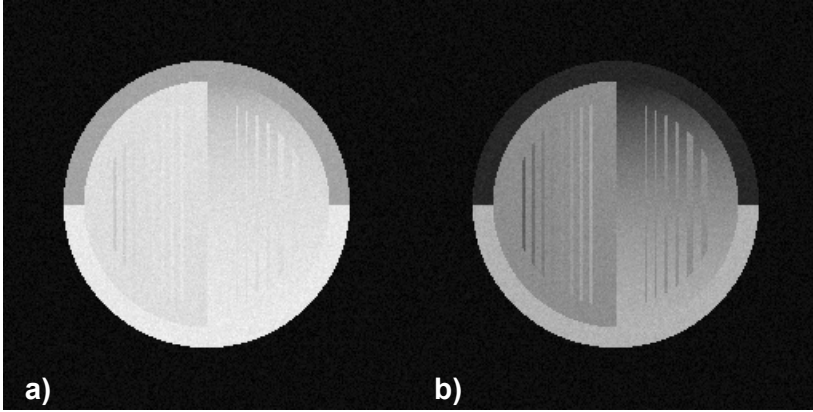
Figure 6.3: The two simulated raw data Rice distributed magnitude images, a) $t_1 = 21$ ms, b) $t_2 = 100$ ms. Average SNR = 30 dB.
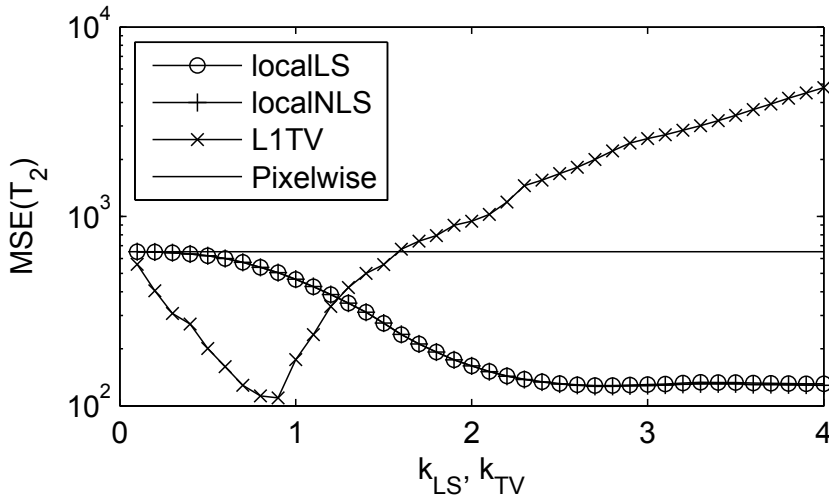
Figure 6.4: The average MSE of the T_2 estimates for the simulated data versus the parameters k_{LS} and k_{TV} that controls the smoothness enforced.

in Fig. 6.7. The images are truncated at 200 ms to make the results in the target (lower) range of T_2 values more clear. Again, both the localLS and L1TV show a clear reduction in the variance without compromising contrast. Visually there is no clear gain in using the L1TV compared to the simpler localLS approach; however, looking at the zoomed-in images in Fig. 6.8, some minor differences can be seen. Only minor staircase artifacts can be seen in the L1TV estimates for the chosen degree of smoothing. By examining the absolute difference between the estimates obtained by the proposed methods and voxelwise approach, shown in Fig. 6.9, it can be seen that the localLS has a more uniform noise reduction, indicated by the low amount of structure in the difference image.
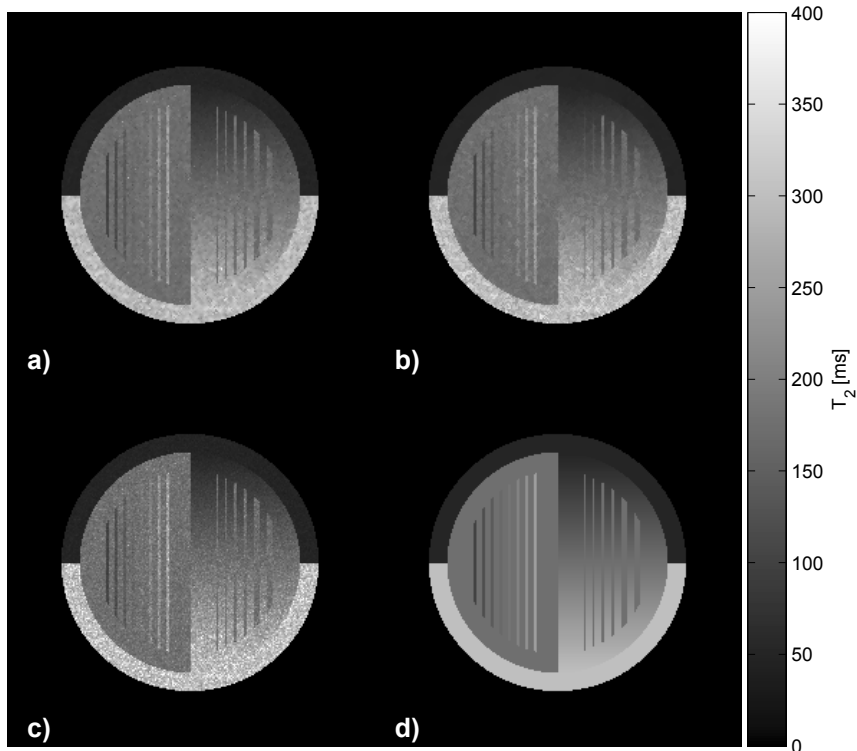
Figure 6.5: Estimates of T_2 obtained by: a) localLS (MSE = 172), b) L1TV (MSE = 209), and c) voxelwise (MSE = 652), together with d) the true values.

The relative difference of the LS estimates compared to the NLS (not shown here) were less than 1% in over 99 % of the voxels. This indicates that the LS gives a good approximation of the NLS, and the extra computational burden involved in obtaining the NLS estimates is not motivated.

The MOSEK linear program solver was run through MATLAB on a Intel Core i7 860 system at 2.93 GHz and with 16 GB of memory. The average run times for the different methods when applied to the brain data (25800 voxels) were 2.3 s for the localLS and 27.5 s for L1TV, while the voxelwise approach is instant.

6.5 Conclusions

We have proposed two methods that can be used to reduce the variance of the T_2 estimates obtained from two spin echo images, without compromising the resolution at tissue boundaries. Both localLS and L1TV include a way of choosing the user parameters, k_{LS} and k_{TV}, re-
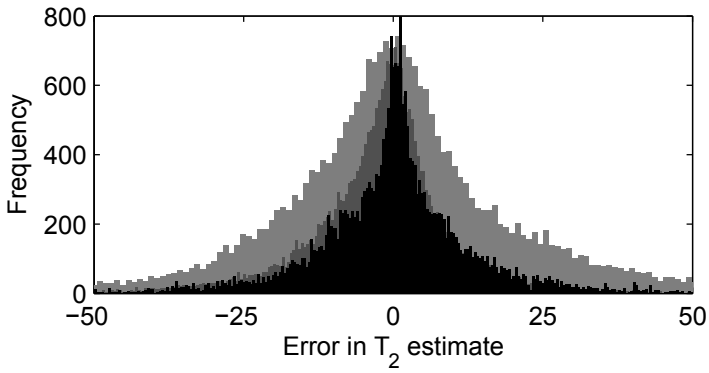
Figure 6.6: Overlay histograms for the errors in the T_2 estimates from Fig. 6.5 using: L1TV (solid), localLS (medium), and voxelwise approach (light).
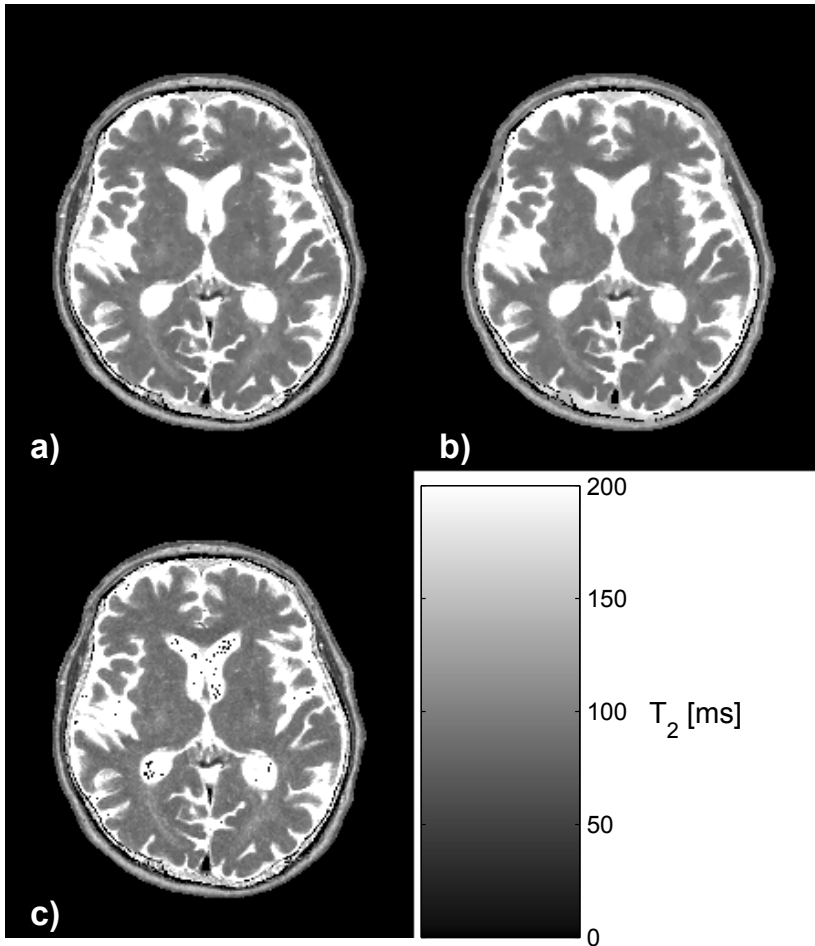
Figure 6.7: Estimates of T_2 obtained by: a) localLS, b) L1TV, and c) voxel-wise approach.
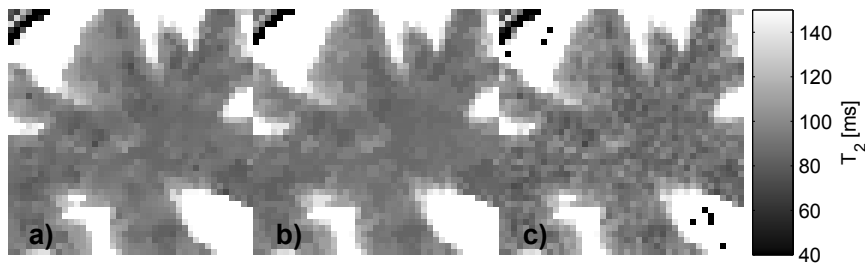
Figure 6.8: Zoomed in version of Fig. 6.7 for a) localLS, b) L1TV, and c) voxelwise approach, showing more detail in the white matter T_2 estimates and the tissue contrast.
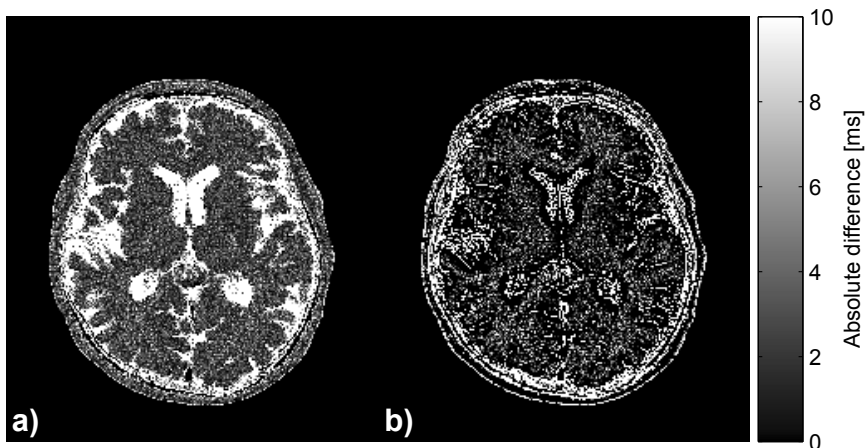
Figure 6.9: Absolute difference between the T_2 estimates obtained by the voxelwise approach and a) localLS, b) L1TV.

spectively, and automatically adapt to the local image conditions. The L1TV approach decides what is considered to be an outlier based on the whole image, not just the center voxel versus the surrounding as in LS. Furthermore, L1TV is set in a more general optimization framework. The localLS algorithm, on the other hand, is easy to use, computationally efficient, and generally gives similar or superior performance compared to L1TV. It is also easy to implement and memory efficient enough to be generally applicable for T_2 estimation on any standard computer.

Chapter 7

Temporal phase correction

7.1 Introduction

The magnitude images commonly used for T_2 relaxometry are Rice distributed [67], which means that the LS parameter estimates will be suboptimal. A few suggestions on how to solve this sub-optimality problem are available in the literature [111, 130], where the authors apply ML methods, taking the Rice distribution into account; however, these approaches are nonlinear and rather involved, which complicates the implementation and can lead to convergence problems. To make the problem more tractable, many algorithms for T_2 estimation are based on minimizing the LS criterion; however, it has been suggested that LS-based approaches can lead to tissue mischaracterization caused by the Rician noise [12].

In this chapter, we present two methods that, based on complex-valued data, compute Gaussian-distributed real-valued images, which can be used in LS-based T_2 estimation without inflicting any bias. By splitting the problem into two parts: one of phase correction to obtain an estimate of the underlying magnitude decay, and another of estimating the relaxation components, the highly nonlinear Rician estimation problem is avoided. Moreover, this enables increased accuracy when using the non-negative least squares (NNLS) algorithm [84, 80], or the method called Exponential Analysis via System Identification based on Steiglitz-McBride (EASI-SM), for multi-component T_2 estimation [120]. It should be noted that the phase correction methods presented herein can be used for other problems than T_2 relaxometry, as only the phase of the data is modeled; however, this option will not be considered here.

A method for temporal phase correction (TPC) has been proposed in [12], where the phase of the signal in an arbitrary voxel is fitted by two fourth-order polynomials using weighted LS. The estimated phase

is then removed from the data, effectively projecting the signal onto the real axis. However, the TPC method does not make the best use of the data, and can be both simplified and generalized. Moreover, no simulations were performed in [12] to show the statistical performance of TPC. The main drawbacks of TPC from a signal processing perspective are: 1) the data is split into two parts to counter any even/odd-echo oscillations, which effectively reduces the number of samples used for estimation; 2) a suboptimal weighting based on the echo number is applied when solving the LS problem, which can be problematic, particularly for nonuniform echo spacing; and 3) estimating the parameters of high-order polynomials, such as the two fourth-order polynomials used in TPC, is rather sensitive to noise when only a few samples are available, which can lead to spurious oscillations in the phase. Typically, a simpler description of the phase is sufficient.

Another approach, that was applied in [4], estimates the signal decay and the noise properties from the magnitude images, and uses this information to transform the magnitude data into a Gaussian signal that can be used for T_2 estimation. This method is shown to improve the T_2 estimates when only the magnitude images are available.

The aim is to statistically improve upon the method in [12], generalize it to any linearly parameterized phase function, and extend it to multi-coil data. The first method proposed makes use of all the available samples to fit a single function to the phase of the data by weighted linear LS. The weights are chosen based on the magnitude of the measured signal, which directly relates to the variance of the phase noise. Using these weights in the LS problem corresponds to an approximation of the BLUE [114]. Moreover, the number of parameters can be chosen based on the observed variations in the data, or other prior information, avoiding the potential over fitting associated with always using a fourth-order polynomial. The alternating phase phenomenon described in [12], which causes the phase to change sign from echo to echo, can be absorbed into the known parts of the equations. Doing so simplifies the algorithm, and makes it more robust compared to splitting the data into even and odd echoes, and performing separate phase estimation for each of these datasets. The resulting method is called Weighted Linear Phase Estimation (WELPE).

Furthermore, we present a ML estimator of the true magnitude decay. This results in a nonlinear problem; however, when the phase changes linearly with time, the ML estimator can be efficiently implemented using the FFT. This case is of particular interest, as linear phase variations were observed in the in-vivo data used both in this chapter and in [12].

In the following, we compare the phase-correction performance of WELPE, ML, and TPC using simulated data; and illustrate the difference in accuracy when estimating multiple T_2 components using phase

corrected data and magnitude data. Furthermore, WELPE is applied to a multi-echo spin-echo dataset comprising 32 in-vivo brain images, to evaluate the practical feasibility of the algorithm. As the focus of this chapter is to compare phase correction methods, and not approaches to estimate the model parameters, only EASI-SM will be used for parameter estimation. This simplifies the bias comparison, as EASI-SM estimates discrete components rather than a continuous spectrum of T_2 values as, for example, NNLS does.

The chapter is structured as follows: In the next section the signal and noise models are presented. Section 7.3 contains details on the two algorithms proposed, while Section 7.4 describes the simulation and data acquisition procedure. The estimation results for both simulated and in-vivo data are shown in Section 7.5, and the discussion of the results follows in Section 7.6. Finally, some concluding remarks are found in Section 7.7.

7.2 Theory

7.2.1 Signal model

For a multi-echo spin-echo acquisition, the received time domain signal from coil j, in an arbitrary voxel, can be modeled as a sum of complex exponentials:

$$\tilde{s}_j(t_n) = k_j e^{iP(t_n)} \sum_{m=1}^{M} c_m e^{-t_n/T_{2m}} + \tilde{\epsilon}_j(t_n) \qquad (7.1)$$

$$= \tilde{g}_j(t_n) + \tilde{\epsilon}_j(t_n), \qquad n = 1, \ldots, N,$$

where $k_j \in \mathbb{C}$ is the coil sensitivity, M is the number of exponential components, $c_m \in \mathbb{R}_+$ and $T_{2m} \in \mathbb{R}_+$ are the amplitude and relaxation time of component m, respectively, $P(t_n)$ is a function describing the phase change over time, $\tilde{\epsilon}_j(t_n)$ is zero-mean i.i.d. complex-Gaussian noise with variance $2\sigma^2$ (i.e. σ^2 in both the real- and imaginary part), t_n is the sampling time relative to excitation, and N is the number of samples (echoes). It is important to note that the factor $e^{iP(t_n)}$ is common to all damped exponentials in a voxel, and all coils, as this phase is mainly attributed to field inhomogeneity and drift [67]. It is also possible to model the signal using a continuous distribution of T_2 values [76], rather than the discrete sum of components in (7.1). The phase-correction algorithms presented herein can be used with either of these models; however, to keep the description concise, the continuous model will not be considered here.

As mentioned, the magnitude data commonly used in T_2 relaxometry is Rice distributed, and can de described by the PDF given in (3.20).

However, by estimating the phase function $P(t)$ in (7.1), including the constant phase contribution from the coil, and removing it from the data, the real part of the resulting signal can be modeled as

$$s_j(t_n) = |k_j| \sum_{m=1}^{M} c_m e^{-t_n/T_{2m}} + \epsilon_j(t_n) \qquad (7.2)$$
$$= g_j(t_n) + \epsilon_j(t_n),$$

where now, $\epsilon_j(t_n)$ is a Gaussian distributed noise with variance σ^2. Unlike the magnitude data, $s_j(t_n)$ is suitable for use in any LS-based method, like NNLS or EASI-SM, without causing bias problems. In the next section two methods for estimating $g_j(t_n)$ from complex-valued data are suggested. These estimated samples will be Gaussian distributed, and can hence be described by the model in (7.2).

Here, $P(t_n)$ is assumed to be linearly parameterized, that is

$$P(t_n) = p_0 + \sum_{q=1}^{Q} p_q \psi_q(t_n), \qquad (7.3)$$

where $\{\psi_q(t_n)\}_{q=1}^{Q}$ are a set of predefined basis functions, Q is the order of the basis expansion, and p_0 captures any constant phase offset in the data. A wide class of phase variations are described by this parameterization, which should enable general use. For example, $\psi_q(t_n) = t_n^q$ gives the set of polynomials of order Q, as used in [12] with $Q = 4$. In most cases, using a polynomial basis is sufficient, for example when no other assumption than smoothness of the phase can be made. Therefore, this is the general recommendation; however, given some prior information on the expected phase variation, other basis functions could be useful.

The SNR is defined as

$$\text{SNR} = \frac{\sum_{n=1}^{N} g(t_n)}{N\sigma}, \qquad (7.4)$$

that is, the mean intensity of the signal over the noise standard deviation. This is the common SNR definition for MR images; however, here it is used across the echoes.

7.3 Methods

7.3.1 WELPE

The unwrapped angle of the model in (7.1) is given by

$$a_j(t_n) = \arg_u(\tilde{s}_j(t_n)) = P(t_n) + \phi_j + v_j(t_n), \qquad (7.5)$$

where $\arg_{\mathrm{u}}(\cdot)$ is the argument of a complex number, unwrapped across time, ϕ_j is the phase of the coil sensitivity k_j, and $v_j(t_n)$ is uncorrelated noise with time dependent variance. The unwrapping is performed by the simple heuristic method in 3.2.2. Note that ϕ_j can be absorbed into p_0, giving the coil dependent parameter $p_{0j} = p_0 + \phi_j$. The phase noise, $v_j(t_n)$, in (7.5) is random and cannot be unwrapped, therefore, $v_j(t_n)$ will be wrapped-normal distributed rather than normal distributed [86].

Constructing the BLUE estimator of $\{p_q\}_{q=0}^Q$ from (7.5) only requires knowledge of the variance of $v_j(t_n)$ for each t_n, independent of the distribution [114]; however, BLUE assumes that the noise is zero mean, which does not hold for the wrapped-normal distribution at low SNR. It is possible to make use of circular statistics to avoid this problem, cf. [86], but for practical purposes it will be sufficient to use the variance of $v_j(t_n)$ to obtain an approximation of the BLUE.

Using (7.3), the model in (7.5) can be rewritten as

$$a_j(t_n) = \begin{bmatrix} 1 & \psi_1(t_n) & \cdots & \psi_Q(t_n) \end{bmatrix} \begin{bmatrix} p_{0j} \\ p_1 \\ \vdots \\ p_Q \end{bmatrix} + v_j(t_n). \qquad (7.6)$$

As only one parameter depends on the coil index j, there is a benefit in using the data from all coils to simultaneously estimate the model parameters. Using N_c coils, this reduces the number of unknowns to $(N_c + Q)$, compared to $(N_c + QN_c)$ when solving the problem for each coil separately; a significant decrease, particularly when $Q > 1$.

By stacking the time samples and the basis functions into vectors, we can define $\mathbf{a}_j = [a_j(t_1), \ldots, a_j(t_N)]^{\mathrm{T}}$, $\mathbf{v}_j = [v_j(t_1), \ldots, v_j(t_N)]^{\mathrm{T}}$, and $\Psi_q = [\psi_q(t_1), \ldots, \psi_q(t_N)]^{\mathrm{T}}$, to obtain the multi-coil matrix model:

$$\mathbf{a} = \begin{bmatrix} \mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} & \Psi_1 & \cdots & \Psi_Q \\ \mathbf{0} & \mathbf{1} & \ddots & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} & \vdots & & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{1} & \Psi_1 & \cdots & \Psi_Q \end{bmatrix} \begin{bmatrix} p_{01} \\ \vdots \\ p_{0N_c} \\ p_1 \\ \vdots \\ p_Q \end{bmatrix} + \mathbf{v}$$

$$= \mathbf{R}\boldsymbol{\theta} + \mathbf{v}, \qquad (7.7)$$

where $\mathbf{a} = [\mathbf{a}_1^{\mathrm{T}}, \ldots, \mathbf{a}_{N_c}^{\mathrm{T}}]^{\mathrm{T}}$, $\mathbf{v} = [\mathbf{v}_1^{\mathrm{T}}, \ldots, \mathbf{v}_{N_c}^{\mathrm{T}}]^{\mathrm{T}}$, $\mathbf{R} \in \mathbb{R}^{NN_c \times (N_c+Q)}$, and $\mathbf{0}, \mathbf{1} \in \mathbb{R}^{N \times 1}$ are column vectors with all elements equal to zero or one, respectively. By construction, the corresponding estimation problem is linear in $\boldsymbol{\theta}$, and by introducing weights proportional to the inverse of the variance of each sample, we can obtain the BLUE estimate of $\boldsymbol{\theta}$ [114].

A full derivation of the variance of the samples, and hence the weights, is mathematically rather involved and not particularly useful in practice; however, to provide some guidance for how to choose the weights, we rewrite (7.1) as

$$\tilde{s} = e^{i(P+\phi)} \left(g + |\epsilon| e^{i\beta} \right), \tag{7.8}$$

where β is a uniformly distributed random phase, and the dependence on j and t_n has been dropped for notational convenience. As the phase $P + \phi$ is to be estimated, the disturbing phase is given by $\arg(g + |\epsilon| e^{i\beta})$. It can be shown that, for high SNR, this phase disturbance can be approximated as $v \approx \mathrm{Re}\{\epsilon\} / g$, and hence, the variance becomes $\mathrm{var}(v) \approx \sigma^2/g^2$ [128]. Therefore, the BLUE weights at high SNR can be approximated as

$$w^*(t_n) = g^2(t_n), \tag{7.9}$$

where the constant σ^2 has been omitted as it has no effect on the BLUE. However, as $g(t_n)$ is not known, we propose to use the magnitude of the noisy data from each coil j to approximate the quantity in (7.9), that is

$$w_j(t_n) = |\tilde{s}_j(t_n)|^2. \tag{7.10}$$

For details on the accuracy of this approximation, particularly at lower SNR, see Section 7.5.

The weighted LS solution can now be written as

$$\hat{\boldsymbol{\theta}} = (\mathbf{R}^{\mathrm{T}} \mathbf{W} \mathbf{R})^{-1} \mathbf{R}^{\mathrm{T}} \mathbf{W} \mathbf{a}, \tag{7.11}$$

and the weight matrix $\mathbf{W}$ is given by

$$\mathbf{W} = \mathrm{diag}\big([\mathbf{w}_1^{\mathrm{T}}, \ \ldots, \ \mathbf{w}_{N_c}^{\mathrm{T}}]\big), \tag{7.12}$$

where $\mathbf{w}_j = [w_j(t_1), \ \ldots, \ w_j(t_N)]^{\mathrm{T}}$, and $\mathrm{diag}(\mathbf{x})$ is a diagonal matrix with the elements of the vector $\mathbf{x}$ along the diagonal.

If the data is collected in a manner that yields alternating phases, as was the case in [12], this is easily detected and compensated for by conjugating every other sample of the data prior to running the estimation method. This procedure does not alter the noise properties, and is in fact equivalent to including the alternating sign in the model.

Next, the estimated phase function is removed from the measured data, giving an estimate of the true signal magnitude:

$$\hat{\mathbf{g}}_j = \mathrm{Re}\Big\{ \mathrm{diag}(\tilde{\mathbf{s}}_j) e^{-i\mathbf{R}_j \hat{\boldsymbol{\theta}}_j} \Big\}, \quad j = 1, \ldots, N_c, \tag{7.13}$$

where $\mathbf{R}_j \hat{\boldsymbol{\theta}}_j$ is the estimated phase function corresponding to coil element j, $\tilde{\mathbf{s}}_j = [\tilde{s}_j(t_1), \ \ldots, \ \tilde{s}_j(t_N)]^{\mathrm{T}}$, and $\mathbf{g}_j = [g_j(t_1), \ \ldots, \ g_j(t_N)]^{\mathrm{T}}$.

The phases of the low-SNR samples cannot be successfully unwrapped as these are practically noise. This unwrapping problem occurs when

---

Algorithm 7.1: Summary of WELPE for phase correction

1: **Inputs:**   Complex-valued data: $\{\tilde{s}(t_n)\}_{n=1}^{N}$
Sampling times: $\{t_n\}_{n=1}^{N}$
model order: $Q \geq 1$
Basis functions: $\{\psi_q(t_n)\}_{q=1}^{Q}$

2: Construct the matrix in (7.7)
3: **for each** voxel **do**
4:    Compute the unwrapped angle of the data (7.5)
5:    Construct the matrix in (7.12)
6:    Compute the WELPE parameter estimates (7.11)
7:    Project data to the real axis (7.13)
8: **end for**

---

the noise intensity is of the same order of magnitude as the signal. In WELPE, this corresponds to a very small weight in the criterion, meaning that the impact on estimation is typically negligible. Moreover, unwrapping the phase noise does not affect the noise properties, and therefore, phase unwrapping can generally be applied to the full dataset without effecting the outcome of the estimation procedure.

The order of the basis expansion, Q, and the basis functions, $\{\psi_q\}_{q=1}^{Q}$, can be chosen based on the data, and as mentioned, the set of polynomials of first or second order is often sufficient. It is also possible to choose Q automatically on a voxel basis by utilizing an order selection method, for example, the Bayesian Information Criterion [123]. This is, however, beyond the scope of our discussion.

The fitting performance of WELPE depends on the SNR, the model order, and the choice of basis functions. For low-order polynomials the phase function can be accurately estimated at rather low SNR values. In particular, WELPE typically works at a much lower SNR than what is required for the application of multi-component T_2 estimation [2, 57].

The estimation and phase correction procedure is summarized in Algorithm 7.1.

7.3.2 Maximum likelihood estimator

When only one coil is used, it is possible to derive a special form of the ML estimator of $g(t_n)$. This method can also be used when multiple coil images have been combined into one image using a method that preserves the data model. For example, taking the mean of the complex-valued images across the coils, or combining the data corresponding to the maximum intensity voxels across the coils into one image, preserves

the model in (7.1), while the sum-of-squares approach does not have this property.

To derive the ML estimator, the true signal magnitude, $g(t_n) \in \mathbb{R}_+$ is seen as a time varying amplitude, that is, no model of $g(t_n)$ is assumed. By rewriting (7.1) as

$$\tilde{s}(t_n) = g(t_n)e^{iP(t_n)} + \tilde{\epsilon}(t_n), \tag{7.14}$$

where the coil index j has been dropped, and the coil-sensitivity parameters have been absorbed into $g(t_n)$ and $P(t)$, the NLS estimates of $g(t_n)$ and $P(t_n)$ can be obtained by minimizing the criterion function

$$
\begin{aligned}
L &= \sum_{n=1}^{N} \left| \tilde{s}(t_n) - g(t_n)e^{iP(t_n)} \right|^2 \\
&= \sum_{n=1}^{N} \left\{ |\tilde{s}(t_n)|^2 + \left[g(t_n) - \mathrm{Re}\left\{ e^{-iP(t_n)}\tilde{s}(t_n) \right\} \right]^2 \right. \\
&\qquad\qquad\qquad \left. - \left[\mathrm{Re}\left\{ e^{-iP(t_n)}\tilde{s}(t_n) \right\} \right]^2 \right\},
\end{aligned} \tag{7.15}
$$

where the second equality follows from expanding and completing the square. Estimates obtained by globally minimizing (7.15) are ML under the assumption of Gaussian noise. Given $\{\hat{p}_q\}_{q=0}^{Q}$ and the corresponding function estimate $\hat{P}(t_n)$, the estimate of $g(t_n)$ is immediately found to be

$$\hat{g}(t_n) = \mathrm{Re}\left\{ e^{-i\hat{P}(t_n)}\tilde{s}(t_n) \right\}. \tag{7.16}$$

By substituting (7.16) into (7.15), it can be seen that the NLS estimate of $\{p_q\}_{q=0}^{Q}$ can be obtained by maximizing the following function:

$$\tilde{L} = 2\sum_{n=1}^{N} \left[\mathrm{Re}\left\{ e^{-iP(t_n)}\tilde{s}(t_n) \right\} \right]^2. \tag{7.17}$$

Using the fact that for any complex number z:

$$[\mathrm{Re}\{z\}]^2 = \frac{1}{2}[|z|^2 + \mathrm{Re}\{z^2\}], \tag{7.18}$$

we can rewrite (7.17) as

$$
\begin{aligned}
\tilde{L} &= \sum_{n=1}^{N} \left[|\tilde{s}(t_n)|^2 + \mathrm{Re}\left\{ e^{-2iP(t_n)}\tilde{s}^2(t_n) \right\} \right] \\
&= \mathrm{const.} + |h(p_1, \ldots, p_Q)| \cos\left(\arg(h(p_1, \ldots, p_Q)) - 2p_0 \right), \tag{7.19}
\end{aligned}
$$

where we have defined

$$h(p_1, \ldots, p_Q) = \sum_{n=1}^{N} \tilde{s}^2(t_n) \exp\left(-2i \sum_{q=1}^{Q} p_q \Psi_q(t_n)\right). \qquad (7.20)$$

It now follows directly that the maximum with respect to p_0 is given by

$$\hat{p}_0 = \frac{1}{2} \arg(h(\hat{p}_1, \ldots, \hat{p}_Q)), \qquad (7.21)$$

where the estimates of $\{p_q\}_{q=1}^{Q}$ are

$$\{\hat{p}_q\}_{q=1}^{Q} = \operatorname*{argmax}_{\{p_q\}_{q=1}^{Q}} |h(p_1, \ldots, p_Q)|. \qquad (7.22)$$

By computing $\{\hat{p}_q\}_{q=1}^{Q}$ from (7.22), and back-substituting into (7.21) to obtain p_0, (7.16) can be used to obtain the ML estimate of $g(t_n)$. Note that this approach does not require phase unwrapping, nor is there a need to use a potentially sub-optimal weighting in the criterion.

For $Q = 1$, finding a solution to (7.22) is a one-dimensional problem, and the maximization can be performed by brute force search on a grid of potential p_1 values. Moreover, when the phase function $P(t_n)$ can be modeled as a linear function of t_n, that is

$$P(t_n) = p_0 + p_1 t_n, \qquad (7.23)$$

the expression in (7.20) is similar to a discrete Fourier transform, and for uniform echo spacing T_s, (7.22) can be efficiently implemented using the FFT:

$$\hat{p}_1 = \frac{1}{2T_s} \operatorname*{argmax}_{f \in [0,\ 2\pi]} \left| \mathrm{FFT}_K\{\tilde{s}^2(t_n)\} \right|, \qquad (7.24)$$

where f corresponds to the normalized frequency in the transform, and K is the number of evaluated frequencies. Typically, $K \approx 10N$ (rounded to the closest power of two) provides sufficient accuracy in practical applications. For nonuniform echo times, the corresponding nonuniform FFT can be used, cf. [44]; alternatively, it is possible to precompute a nonuniform Fourier matrix based on grid of p_1 values, and use this matrix at each voxel.

For higher orders of the phase function, that is $Q \geq 2$, the estimates in (7.22) can be obtained by nonlinear maximization, given proper initialization; however, in such a case we can only expect convergence to a local maximum of $|h(p_1, \ldots, p_Q)|$, and hence, the obtained estimates may be suboptimal.

MATLAB implementations of both WELPE and the ML algorithm are available at:
https://github.com/AAAArcus/PhaseCorrection.

7.4 Simulation and Data Acquisition

As the phase variation observed in the in-vivo data was approximately linear over time (see Section 7.5), the simulations will focus on this scenario. Inspired by [12], Monte Carlo (MC) simulations were performed at SNR = 70, using datasets with 48 samples and linear alternating phase, where the first 32 echoes were spaced by 10 ms, and the following 16 by 50 ms. However, to enable reliable estimation of more than two T_2 components at SNR = 70, 128 echoes, uniformly spaced by 10 ms, were used in the T_2 estimation example. The data was generated using the model in (7.1), and complex-valued Gaussian white noise of variance $2\sigma^2$ was added to achieve the appropriate SNR, as defined by (7.4). To enable direct comparison with TPC, only a single coil was simulated, and the coil sensitivity k was set to unity (without loss of generality). Furthermore, to obtain a high accuracy of the sample means and variances, 10000 MC simulations were performed in all examples.

The chosen parameter set used to generate the data in the simulation examples was

$$T_2 = [20, \ 80, \ 200] \text{ ms},$$
$$c = [0.4, \ 1, \ 0.1]. \tag{7.25}$$

Single slice, multi-echo spin-echo in-vivo brain data was collected at Uppsala University Hospital using a 1.5 T Philips scanner equipped with a 8 element SENSE headcoil. For an echo spacing of 10 ms and $TR = 2200$, 32 images with echo times ranging from 10 ms to 320 ms, were acquired. No averaging was used. An imaging matrix of 240×188 voxels was used to obtain a resolution of $1 \times 1 \times 5$ mm. The total scan time was 7 minutes and 11 seconds.

7.5 Results

7.5.1 Simulation

To illustrate the accuracy of the proposed WELPE weights, given by (7.10), the variance of the phase noise was empirically approximated by the means of MC simulation. Noisy datasets were generated using the model parameters in (7.25), and the variance of the phase at each sampling instance was computed. The corresponding weights are shown in Fig. 7.1, along with the high SNR approximation, $g^2(t_n)$, and the proposed square magnitude of a noisy samples given by (7.10). The $1/n$ weights used in TPC [12] were also included for comparison. As can be seen, the empirical weights are relatively well approximated by the proposed weights in this example (note the logarithmic scale). Moreover, (7.10) provides a better approximation to the empirical weights than

Table 7.1: Root mean square errors (rMSE) of the $\{T_{2m}\}_{m=1}^{3}$ estimates using magnitude, WELPE and TPC data, computed from 10000 MC simulations. The simulated data consisted of 128 echoes with uniform spacing of 10 ms, and SNR = 70.

Method	rMSE [ms]		
	T_{21}	T_{22}	T_{23}
Magnitude	1.1	4.9	87.0
WELPE	0.5	1.6	13.3
TPC	1.2	3.6	16.4

$g^2(t_n)$ at later samples, that is, at lower SNR. TPC, on the other hand, gives a relatively low weight to the initial, most informative, samples, and too high weight to the noisy samples at the end of the decay.

An example of a simulated dataset with linear alternating phase is shown in Fig. 7.2, in terms of the magnitude and phase of the noisy data. The generated data closely resembles the measurements from a white matter voxel shown in [12]. As can be seen, the phase information is very noisy when the signal magnitude is low, even though the true phase is linear and alternating.

To show the statistical improvement associated with using WELPE or ML, compared to TPC, when estimating the true magnitude decay, MC simulations were performed. The average of the estimated phase-corrected signals are shown in Fig. 7.3. As can be seen, both WELPE and ML are on the average close to the true noise-free magnitude decay, while TPC gives a signal that is more similar to the magnitude of the noisy data.

To illustrate the multi-component T_2 estimation performance based on WELPE, TPC, and magnitude data, MC simulations were performed using the parameters in (7.25). The root mean square errors (rMSE) of the T_2 estimates, obtained using EASI-SM on the different datasets, are listed in Table 7.1. As can be seen, the WELPE phase-corrected data provides the highest parameter estimation accuracy, followed by TPC, and the magnitude data. The estimates from the first 200 MC simulations, plotted in the $\{c, T\}$-plane, are shown in Fig. 7.4, together with the true parameter values. As can be seen, using the magnitude data causes a bias in the estimates, particularly for the slowest T_2 component. TPC reduces this bias, but the estimates are slightly more spread out compared to when using the data corrected by WELPE. Moreover, TPC leads to a few outliers in the parameter estimates, shown as squares clearly separated from the main cluster of estimates. ML provided similar results to WELPE, but was omitted from Fig. 7.4 for clarity.
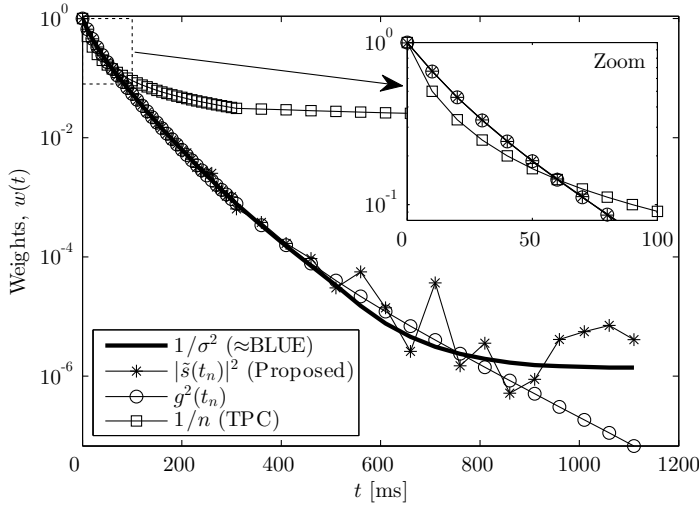
Figure 7.1: An example of the LS weights given by the inverse of the variance of the phase (approximate BLUE weights), the proposed weights obtained directly from the data, and the TPC weights given by the inverse of the echo number.

7.5.2 In-vivo

To reduce the computation time, the collected in-vivo images were initially masked by excluding voxels below 20% of the maximum intensity voxel. This masked region corresponds to the low signal background of the images, mainly located outside of the skull. The magnitude and phase of the 80 ms echo from the collected 32 echo dataset is shown in Fig. 7.5, together with the estimated magnitude and phase using WELPE, and the corresponding difference images ($|\tilde{s}| - \hat{g}$ and $\arg(\tilde{s}) - \hat{P}$). As can be seen, the magnitude image and the WELPE corrected data are close to one another. Moreover, the phase is accurately modeled as no major structure remains in the difference image, and a clear denoising effect is observed in the estimated phase image. An example of the WELPE fitting of the phase for a single voxel is shown in Fig. 7.6. The phase initially varies approximately linearly, while as the signal magnitude decays, the phase noise grows larger, leading to a more random behavior. Because of the weighting in the criterion, WELPE is able to fit to the initial linear trend, while effectively disregarding the heavily distorted samples from the later echoes.

In Fig. 7.7, a histogram of the imaginary part of the phase corrected image from Fig. 7.5 is shown. As can be seen, the samples are approximately Gaussian distributed with zero mean and a relatively small standard deviation.
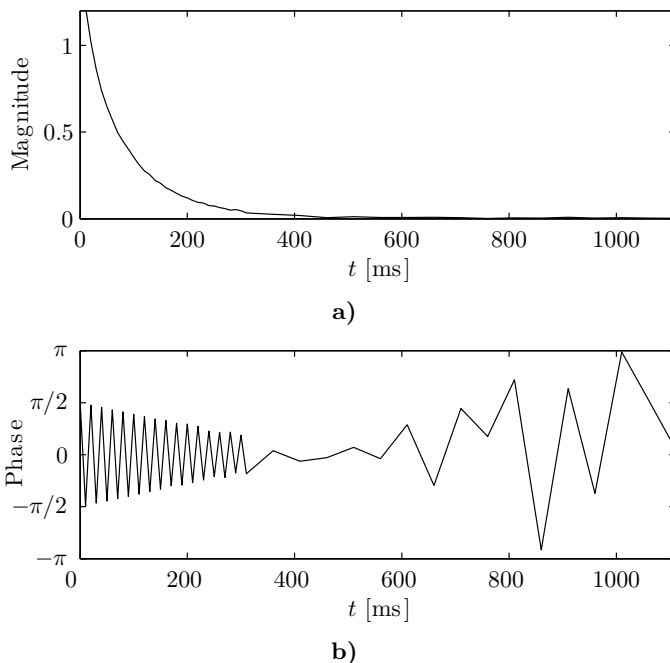
Figure 7.2: a) Magnitude and b) phase [rad] over time, for a simulated noisy dataset where the true phase is linear and alternating.

7.5.3 Computational details

The computational burden of the considered methods varies depending on the settings and the implementation; however, we can still give an indication of the run times to be expected. The computation times for all methods when performing 10000 MC simulations on a single processor thread using an Intel i7 860 at 2.93 GHz, are shown in Table 7.2. The simulated data consisted of 128 echoes with uniform spacing of 10 ms, similar to the example in Table 7.1. A uniform sampling was used to be able to show the computational performance of the ML estimator based on the FFT implementation. WELPE is the fastest method, running approximately twice as fast as TPC. ML has a slightly longer runtime than WELPE for $K = 1024$, but is still more computationally efficient than TPC in this case.

7.6 Discussion

7.6.1 Simulation

The proposed weights used in the WELPE criterion are data adaptive, provide a good approximation of the empirical BLUE weights, are easily

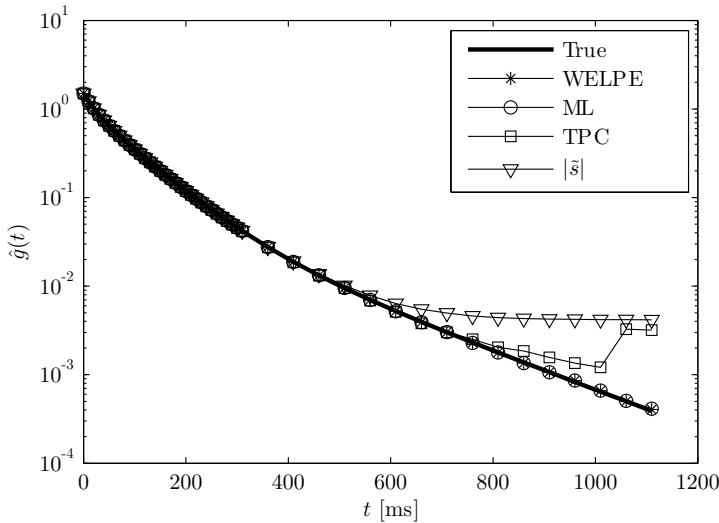**Figure 7.3:** The average estimated magnitude decay $\hat{g}(t_n)$ from 10000 MC simulations for WELPE, ML, TPC, and the magnitude data, compared to the true decay $g(t_n)$.

Table 7.2: MATLAB run times in seconds for phase correcting 10000 datasets using WELPE, ML, and TPC, at different SNRs.

	SNR		
Method	50	70	150
WELPE	1.2	1.2	1.2
ML ($K = 1024$)	1.9	1.9	1.8
TPC	2.8	2.7	2.7

computed, and lead to superior statistical performance compared to the $1/n$ weights used in TPC.

The simulated data with linear and alternating phase displayed in Fig. 7.2 showed that a linear phase model, rather than a fourth-order polynomial, would have been more suitable to model the phase of the data in [12]. This motivates both WELPE, which has the option to choose the number of parameters used to model the phase based on the data, and the ML method, which provides a simple FFT-based solution for the case of linear phase variations.

As was seen in Fig. 7.3, the TPC data is on the average more similar to the magnitude data, in this example with non-uniformly spaced echoes. This is partially due to the suboptimal weights used, which puts too high weight to the noisy samples at the end of the decay. The data adaptive weights used by WELPE, on the other hand, provide an
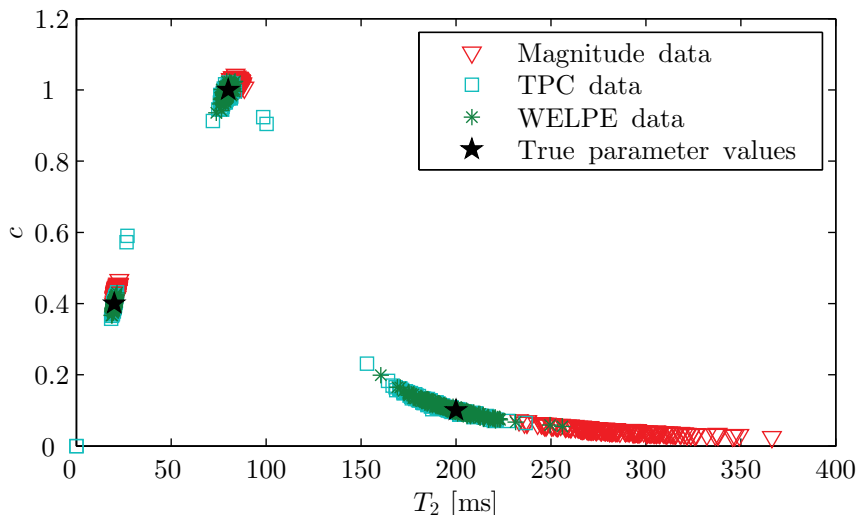
Figure 7.4: Estimates of c and T_2 from 200 MC simulations, obtained by applying EASI-SM to phase-corrected data generated by WELPE and TPC, as well as to the magnitude of the noisy data. The true parameter values are indicated by the stars. The estimates correspond to the first 200 realizations of the simulations in Table 7.1.

accurate estimate of the true signal magnitude. The jump discontinuity shown in the last two samples of the TPC estimate in Fig. 7.3 is due to robustness issues in the phase fitting. The two fourth-order polynomials used in TPC usually fits the initial linear phase, where the SNR is high, quite well. However, as a consequence of the high order model, several extreme points of the polynomials are placed in the low SNR region, essentially fitting to the noise. This can cause rapid variations in the estimated phase function, and in turn, can lead to an increase in the bias. Thresholding the data and removing samples below a certain SNR from the fitting procedure, as was suggested in [12], eliminates the discontinuity of TPC in Fig. 7.3, but overall makes the estimated decay curve more similar to the magnitude data (not shown here).

As was shown in Fig. 7.4, using the magnitude of the noisy data and an LS-based approach can lead to a significant bias in the T_2 estimates, and therefore, mischaracterization of the tissue. This can be expected, as the Rician noise will effectively raise the tail of the decay curve, which is interpreted as a slower decay. By applying WELPE to correct the phase, however, the data becomes Gaussian distributed, which eliminates the bias problem. The TPC approach also improves the T_2 estimation significantly, but results in a higher rMSE than WELPE, and occasionally causes outliers in the estimates. The outliers are partially due to the robustness issues in the polynomial fitting mentioned above,
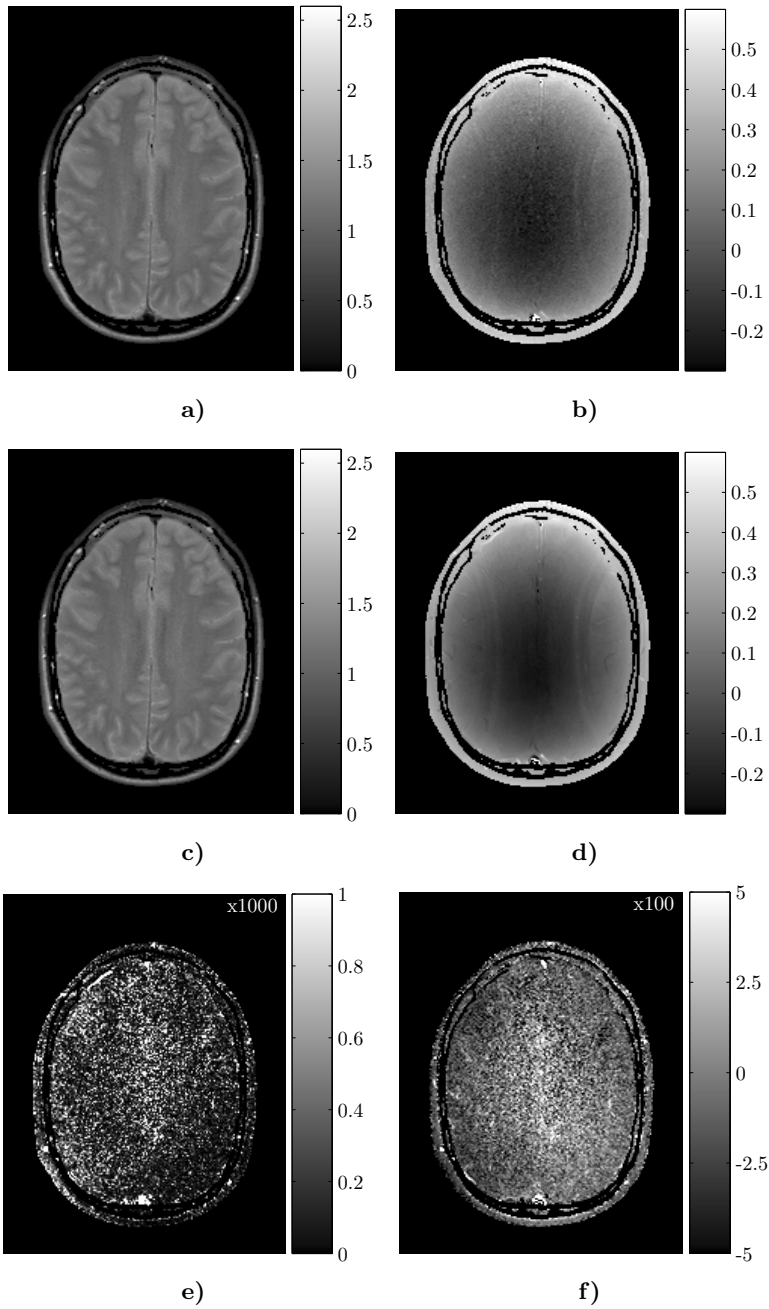
Figure 7.5: a) Magnitude, and b) phase [rad] at echo time 80 ms of the single slice in-vivo dataset; and the estimated c) magnitude, and d) phase, provided by WELPE, together with the error in e) magnitude (amplified by a factor of 1000), and f) phase (amplified by a factor of 100).
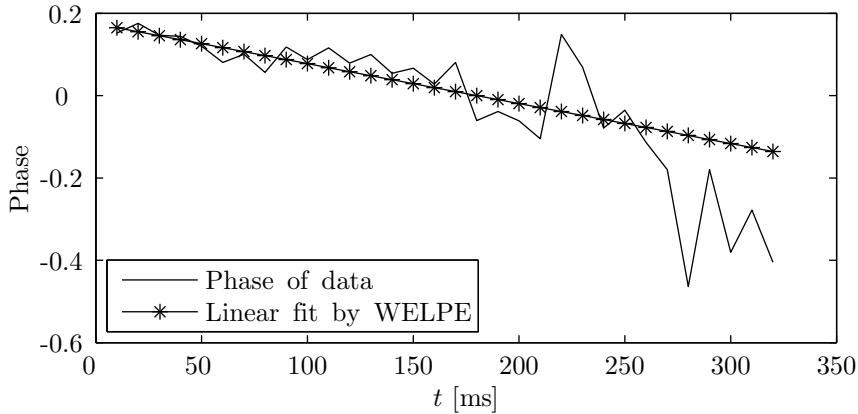
Figure 7.6: Phase [rad] over time for a single voxel of the in-vivo dataset together with the linear fit provided by WELPE.
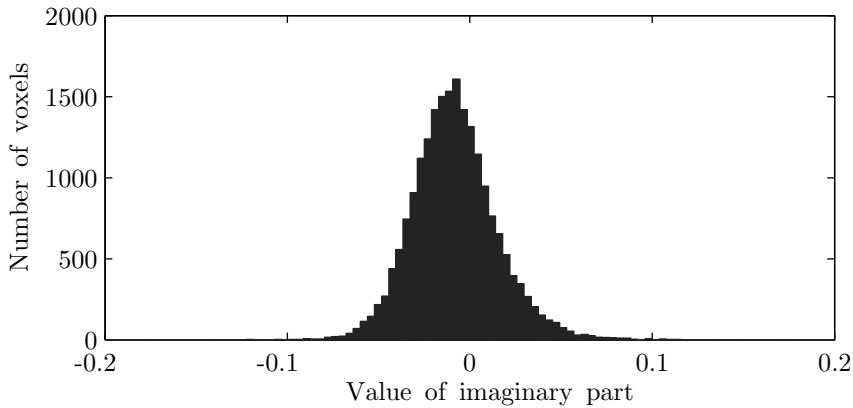
Figure 7.7: Distribution of imaginary part of the phase corrected image at echo time 80 ms.

which for some noise realizations can lead to poor estimates of the true phase function.

7.6.2 In-vivo

For the collected 32 echo in-vivo dataset, WELPE is able to accurately model the phase in the whole image, and provides an estimate of the true magnitude decay that is close to the magnitude of the noisy data for high SNR, as was shown in Fig. 7.5. Figure 7.6 showed that the linear fit to the phase variation in time was a relatively accurate approximation, and that it would be hard to motivate the fitting of a higher order phase function in this example.

The histogram in Fig. 7.7, showed that the imaginary part of the data after phase correction is small and approximately Gaussian distributed. This would be expected in the case of successful phase correction, and indicates that WELPE is able to project a large proportion of the signal energy to the real axis, leaving mainly noise in the imaginary part.

7.6.3 Computational details

Since TPC requires two systems of equations to be solved and the results recombined, the computation times for WELPE are typically lower. The FFT implementation of the ML estimator is also fast for $K \approx 10N$, which provides reasonable accuracy. Note that it is possible to trade off computation time and accuracy by setting the number of evaluation points, K, in the FFT.

7.7 Conclusion

Two methods for phase correction have been presented, and through simulations, the algorithms have been shown to be useful for avoiding bias in multi-component T_2 estimation by accurately estimating the true magnitude decay. WELPE is statistically sound and is easy to implement; moreover it works with multi-coil data, general sampling schemes, and a wide range of phase functions. The ML estimator is optimal, and does not require phase unwrapping or weighting of the criterion. Moreover, it can be rather efficiently implemented for linear phase variations in time; however, in the general case, finding the ML estimates is computationally more intensive than using WELPE.

Chapter 8

Sequence design for excitation

8.1 Introduction

Sequence, or waveform, design aims to generate sequences with specific desired properties, such as a certain spectral content, or good correlation properties. There is a wide range of applications, for example, in communications, active sensing, and MRI [52, 65, 62]. Typically, the signal to be designed is characterized by means of an optimization problem. Solving the problem globally can be difficult when the criterion is non-convex; however, in some cases a local minimization algorithm is sufficient to find a good solution. Indeed, different local optima correspond to possible candidates for a signal, and since the problem is usually solved offline, it is possible to generate several signals and choose the best among these based on the criterion.

In this chapter, we derive a cyclic algorithm that locally solves a class of sequence design problems where a constraint on the magnitude of the designed complex-valued signal is enforced. This formulation has applications in MRI [59], but could also apply to other fields where low-cost amplifiers are used. Typically, such amplifiers are single stage and are not equipped with feedback control [35]. This can cause nonlinear distortion of the signal when there are rapid variations in the magnitude [94]. By penalizing such variations in the design, the resulting sequences can be amplified and transmitted with higher fidelity.

8.2 Problem formulation

In general terms, the criterion to be minimized can be formulated as

$$f(\mathbf{x}) = \|\mathbf{d} - \mathbf{A}\mathbf{x}\|_{\mathbf{W}}^2 + \lambda\|\mathbf{R}|\mathbf{x}\|\|^2, \tag{8.1}$$

where $\mathbf{d} \in \mathbb{C}^M$ is the desired signal, $\mathbf{x} \in \mathbb{C}^N$ is the signal to be designed, $\mathbf{A} \in \mathbb{C}^{M \times N}$ and $\mathbf{R} \in \mathbb{C}^{P \times N}$ are arbitrary linear transformation matrices, and $\mathbf{W} \in \mathbb{C}^{M \times M}$ is a positive semidefinite weighting matrix. The regularization term contains a magnitude vector, which makes this function non-convex in general. The minimization of (8.1) with respect to $\mathbf{x}$ can be done in several ways; however, for large problems it is necessary to find an efficient method with low computational complexity. The algorithm used in [59] is similar to the heuristic Iterative Quadratic Maximum Likelihood (IQML) algorithm and is not guaranteed to converge, nor is it a true minimization algorithm for the criterion [121]. However, IQML does typically converge to a vector fairly close to a minimizer of the stated criterion.

For the criterion in (8.1) the IQML algorithm can be described as follows. The vector $\mathbf{x}$ can be elementwise partitioned into its magnitude and phase as

$$x_k = |x_k|e^{i\phi_k}, \quad k = 1 \ldots N. \tag{8.2}$$

By stacking the phases $\{\phi_k\}_{k=1}^N$ into a vector $\boldsymbol{\phi}$, we can form a criterion function:

$$g(\mathbf{x}, \boldsymbol{\phi}) = \|\mathbf{d} - \mathbf{A}\mathbf{x}\|_{\mathbf{W}}^2 + \lambda \|\mathbf{R}\text{diag}(e^{-i\boldsymbol{\phi}})\mathbf{x}\|^2, \tag{8.3}$$

where we have defined $e^{-i\boldsymbol{\phi}} = [e^{-i\phi_1}, \ldots, e^{-i\phi_N}]^{\mathrm{T}}$ for notational convenience, and $\text{diag}(e^{-i\boldsymbol{\phi}})$ is a square matrix with the elements of $e^{-i\boldsymbol{\phi}}$ along its diagonal. Under the constraint that $\boldsymbol{\phi} = \text{arg}(\mathbf{x})$, where the argument is taken elementwise, we have $g(\mathbf{x}, \boldsymbol{\phi}) = f(\mathbf{x})$; however, by relaxing this constraint and keeping $\boldsymbol{\phi}$ fixed, the minimization of (8.3) with respect to $\mathbf{x}$ becomes quadratic. After solving for $\mathbf{x}$, the phases can be updated as $\boldsymbol{\phi} = \text{arg}(\mathbf{x})$. These two steps are then iterated until some predefined stopping condition is satisfied.

Since IQML is not a minimizer of (8.1), it does not get stuck in local minima in the same way as a true minimization algorithm does. This property, together with the observation that IQML often converges rather rapidly, makes IQML a potential candidate for initialization of the local minimization algorithm described in Section 8.3.1. However, when there is no optimal vector $\mathbf{x}_{\text{opt}}$ such that $\mathbf{A}\mathbf{x}_{\text{opt}}$ is close enough to $\mathbf{d}$, IQML tends to have poor performance, and in the worst case, might not converge. An example of this type of behavior is shown in Section 8.5.

In [59], IQML is initialized by $\boldsymbol{\phi} = \mathbf{0}$, meaning that the first optimization step consists of solving the following least-squares problem:

$$\underset{\mathbf{x}}{\text{minimize}} \ \|\mathbf{d} - \mathbf{A}\mathbf{x}\|_{\mathbf{W}}^2 + \lambda \|\mathbf{R}\mathbf{x}\|^2. \tag{8.4}$$

The solution to the complex-valued smoothing problem above provides a reasonably good initialization for the non-convex problem in (8.1).

8.3 Magnitude-Constrained Cyclic Optimization (MACO)

8.3.1 Description of the Algorithm

Using (8.2), and defining $z_k = |x_k| \geq 0$, we can re-write the problem of minimizing (8.1) as

$$\underset{\mathbf{z}, \boldsymbol{\phi}}{\text{minimize}} \quad \left\| \mathbf{d} - \sum_{k=1}^{N} \mathbf{a}_k z_k e^{i\phi_k} \right\|_{\mathbf{W}}^2 + \lambda \|\mathbf{R}\mathbf{z}\|^2 \quad , \tag{8.5}$$
$$\text{subject to} \quad \mathbf{z} \geq \mathbf{0}$$

where $\mathbf{a}_k$ is the k:th column of $\mathbf{A}$. Assuming $\mathbf{z}$ and $\{\phi_k\}_{k \neq p}$ are given, let

$$\mathbf{d}_p = \mathbf{d} - \sum_{\substack{k=1 \\ k \neq p}}^{N} \mathbf{a}_k z_k e^{i\phi_k}, \tag{8.6}$$

and observe that

$$\|\mathbf{d}_p - \mathbf{a}_p z_p e^{i\phi_p}\|_{\mathbf{W}}^2 + \lambda\|\mathbf{R}\mathbf{z}\|^2 =$$
$$= \|\mathbf{d}_p\|_{\mathbf{W}}^2 + z_p^2 \|\mathbf{a}_p\|_{\mathbf{W}}^2 + \lambda\|\mathbf{R}\mathbf{z}\|^2 - 2\,\mathrm{Re}\{z_p e^{-i\phi_p} \mathbf{a}_p^* \mathbf{W} \mathbf{d}_p\}$$
$$= -2 z_p |\mathbf{a}_p^* \mathbf{W} \mathbf{d}_p| \cos(\arg(\mathbf{a}_p^* \mathbf{W} \mathbf{d}_p) - \phi_p) + C, \tag{8.7}$$

where the constant term C is independent of ϕ_p. Then it follows that the ϕ_p that minimizes the criterion in (8.5) is

$$\hat{\phi}_p = \arg(\mathbf{a}_p^* \mathbf{W} \mathbf{d}_p), \tag{8.8}$$

for each p. By cycling through the entire $\boldsymbol{\phi}$ vector we obtain an updated estimate, $\hat{\boldsymbol{\phi}}$, for the next iteration.

Once the phase vector is updated, we have to solve the minimization problem in (8.5) with respect to $\mathbf{z} \geq \mathbf{0}$ for a fixed $\boldsymbol{\phi} = \hat{\boldsymbol{\phi}}$. Similarly to the approach above for $\{\phi_k\}$, it is possible to determine the $\{z_k\}$ one-by-one. First, we rewrite the criterion in (8.5) as

$$\left\| \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{A}\,\mathrm{diag}(e^{i\boldsymbol{\phi}}) \\ -\sqrt{\lambda}\mathbf{R} \end{bmatrix} \mathbf{z} \right\|_{\tilde{\mathbf{W}}}^2 \triangleq \|\mathbf{c} - \mathbf{B}\mathbf{z}\|_{\tilde{\mathbf{W}}}^2 , \tag{8.9}$$

where

$$\tilde{\mathbf{W}} = \begin{bmatrix} \mathbf{W} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_P \end{bmatrix}, \tag{8.10}$$

and $\mathbf{I}_P$ is the identity matrix of size $P \times P$. If we assume that $\boldsymbol{\phi}$ and $\{z_k\}_{k \neq p}$ are given, and define

$$\mathbf{c}_p = \mathbf{c} - \sum_{\substack{k=1 \\ k \neq p}}^{N} \mathbf{b}_k \hat{z}_k, \tag{8.11}$$

Algorithm 8.1: MACO Sequence Design

1: **Input: A, R, d, W,** λ, initial guess of **z**
2: **repeat**
3:   **Step 1:**
4:   **for all** p **do**
5:     Compute $\mathbf{d}_p$ using (8.6)
6:     Compute $\hat{\phi}_p$ using (8.8)
7:   **end for**
8:   **Step 2:**
9:   **for all** p **do**
10:     Compute $\mathbf{c}_p$ using (8.11)
11:     Compute $\hat{z}_p$ using (8.13)
12:   **end for**
13: **until** convergence
14: **Output:** Compute $\hat{\mathbf{x}}$ from $\hat{\mathbf{z}}$ and $\hat{\phi}$ using (8.2)

---

we can write (8.9) as

$$\|\mathbf{c}_p - \mathbf{b}_p z_p\|^2_{\tilde{\mathbf{W}}} = \|\mathbf{c}_p\|_{\tilde{\mathbf{W}}} + z_p^2 \|\mathbf{b}_p\|^2_{\tilde{\mathbf{W}}} - 2 z_p \operatorname{Re}\left\{\mathbf{b}_p^* \tilde{\mathbf{W}} \mathbf{c}_p\right\}$$

$$= \text{const.} + \|\mathbf{b}_p\|^2_{\tilde{\mathbf{W}}} \left[z_p - \frac{\operatorname{Re}\left\{\mathbf{b}_p^* \tilde{\mathbf{W}} \mathbf{c}_p\right\}}{\|\mathbf{b}_p\|^2_{\tilde{\mathbf{W}}}} \right]^2, \qquad (8.12)$$

where $\mathbf{b}_p$ is the p:th column of $\mathbf{B}$, and the constant term is independent of z_p. The minimizer $\hat{z}_p \geq 0$ of (8.12) has the following simple expression:

$$\hat{z}_p = \begin{cases} \dfrac{\operatorname{Re}\left\{\mathbf{b}_p^* \tilde{\mathbf{W}} \mathbf{c}_p\right\}}{\|\mathbf{b}_p\|^2_{\tilde{\mathbf{W}}}} & \text{if } \operatorname{Re}\left\{\mathbf{b}_p^* \tilde{\mathbf{W}} \mathbf{c}_p\right\} > 0 \\ \\ 0 & \text{otherwise} \end{cases} . \qquad (8.13)$$

This can be used to update $\hat{\mathbf{z}}$, element by element, in the same manner as for $\hat{\phi}$. By iterating the two steps, (8.8) and (8.13), the criterion function in (8.5) will decrease monotonically, as each step minimizes a part of the criterion. Since the criterion is bounded from below, it follows that the algorithm will converge to a local minimum. The proposed MACO algorithm is summarized in Algorithm 8.1.

The MACO algorithm can be initialized in several ways. A good guess is, typically, provided by solving (8.4). The other option considered here is to initialize the algorithm by IQML, given that it has converged properly. By using the estimate obtained from IQML as initialization,

MACO is guaranteed to perform at least as well, while taking advantage of IQMLs potential ability to avoid some local minima.

It should be noted that the problem of minimizing (8.9) with respect to $\mathbf{z} \geq \mathbf{0}$ is a linearly constrained quadratic program (LCQP), which can typically be solved rather efficiently by, for example, interior-point methods [28]. We denote the corresponding method MACO-LCQP. However, for large dimensions it might be favorable to determine the $\{z_k\}$ one-by-one, as was done above.

The computations in (8.6) and (8.11) can be performed recursively to reduce the computational burden. We have

$$\mathbf{d}_p = \mathbf{d} - \sum_{\substack{k=1 \\ k \neq p}}^{N} \mathbf{a}_k e^{i\phi_k} z_k = \mathbf{d} - \mathbf{A}\mathbf{x} + \mathbf{a}_p x_p, \qquad (8.14)$$

where x_p is the current estimate. After obtaining the updated estimate $\hat{x}_p$ we can express the next residual as

$$\mathbf{d}_{p+1} = \mathbf{d}_p - \mathbf{a}_p \hat{x}_p + \mathbf{a}_{p+1} x_{p+1}. \qquad (8.15)$$

Because of this, $\mathbf{d} - \mathbf{A}\mathbf{x}$ only has to be computed once; although, to prevent accumulating numerical errors in the recursion, a full re-computation of the residual can be done at each step of Algorithm 8.1. A similar recursion holds for $\mathbf{c}_p$.

8.3.2 Note on convergence

One interesting property is that the optimal phases $\phi^\star$ are independent of $\mathbf{z} \geq \mathbf{0}$ when $\mathbf{A}^*\mathbf{W}\mathbf{A}$ is diagonal. This can be shown by inserting the polar form of $\mathbf{x}$ from (8.2) into (8.1) and expanding. We then obtain the following criterion to be minimized with respect to ϕ:

$$\tilde{f}(\mathbf{z}, \phi) = [e^{-i\phi}]^{\mathrm{T}} \mathbf{Z}^* \mathbf{A}^* \mathbf{W}\mathbf{A}\mathbf{Z} e^{i\phi} - 2\operatorname{Re}\{\mathbf{d}^*\mathbf{W}\mathbf{A}\mathbf{Z} e^{i\phi}\}, \qquad (8.16)$$

where $\mathbf{Z} = \operatorname{diag}(\mathbf{z})$, and the terms that are constant with respect to ϕ have been omitted. If $\mathbf{A}^*\mathbf{W}\mathbf{A}$ is diagonal, the ϕ:s in the first term cancel, and reminiscent of (8.7), the remaining term can be written as

$$-2\sum_{k=1}^{N} z_k \cos\left(\arg(\mathbf{a}_k^*\mathbf{W}\mathbf{d}) - \phi_k\right), \qquad (8.17)$$

which decouples for all k. Therefore, the minimizer of (8.1) with respect to ϕ, for a diagonal $\mathbf{A}^*\mathbf{W}\mathbf{A}$, is given by

$$\phi^\star = \arg(\mathbf{A}^*\mathbf{W}\mathbf{d}). \qquad (8.18)$$

Moreover, using (8.8) to solve for $\{\phi_k\}_{k=1}^{N}$ converges in one step, and given that the remaining problem for $\mathbf{z}$ is solved by an LCQP, the optimum is reached instantly.

8.4 Application to MRI

In MRI, the problem is to design sequences used to excite, or tip, the magnetic field vector in a certain region of a subject. Typically, such excitation pulses have rapidly varying magnitudes [71, 142]. As mentioned, the low-cost amplifiers commonly used in parallel MRI can distort these signals, leading to artifacts in the resulting images. Therefore, the problem consists of finding a signal with smooth magnitude, while trying to maintain the desired excitation pattern. The multi-coil problem can be stated as follows [59]:

$$\operatorname*{argmin}_{\{\mathbf{x}_j\}_{j=1}^{N_c}} \left\| \mathbf{d} - \sum_{j=1}^{N_c} \operatorname{diag}(\mathbf{s}_j)\tilde{\mathbf{A}}\mathbf{x}_j \right\|_{\mathbf{W}}^2 + \lambda \left\| \sum_{j=1}^{N_c} \tilde{\mathbf{R}}|\mathbf{x}_j| \right\|^2, \qquad (8.19)$$

where N_c is the number of parallel transmit channels in the coil array, $\mathbf{s}_j \in \mathbb{C}^M$ is the vectorized spatial sensitivity of coil j, and $\mathbf{x}_j$ is the corresponding complex-valued signal to be designed. By stacking the N_c signal vectors in one vector $\mathbf{x} = [\mathbf{x}_1^{\mathrm{T}} \ \cdots \ \mathbf{x}_{N_c}^{\mathrm{T}}]^{\mathrm{T}}$, and defining the matrices $\mathbf{A} = [\operatorname{diag}(\mathbf{s}_1)\tilde{\mathbf{A}} \ \cdots \ \operatorname{diag}(\mathbf{s}_{N_c})\tilde{\mathbf{A}}]$, and $\mathbf{R} = \mathbf{I}_{N_c} \otimes \tilde{\mathbf{R}}$, where $\otimes$ is the Kronecker product, we get the problem in the same form as (8.1). The desired signal $\mathbf{d}$ is in this case a vectorized multi-dimensional excitation pattern in space. The matrix $\mathbf{A}$ corresponds to a Fourier-type matrix that captures the, possibly nonuniform, sampling trajectory in k-space over time. The regularization matrix $\mathbf{R}$ can, for example, be determined by using a linear approximation of the filtering occurring in the amplifier, and computing the expected distortion filter. However, this requires knowledge of, or direct measurements from, the amplifier used. The distortion of the amplifiers used in [59] was shown to be fairly accurately modeled by a first-order difference filter, that is

$$\mathbf{R} = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & -1 \end{bmatrix}_{N-1 \times N}, \qquad (8.20)$$

which is the approximation we will consider here. For a more detailed explanation of how one can achieve a multi-dimensional excitation pattern in space from one or several scalar time series, see for example [138].

8.5 Numerical examples

8.5.1 Example 1: A simple design

Let $\mathbf{W}$ and $\mathbf{A}$ in (8.1) be identity matrices, then the optimal phases can be obtained in closed form as $\phi = \arg(\mathbf{d})$, by making use of (8.18). The resulting LCQP for $\mathbf{z}$ can be written as follows

$$
\begin{aligned}
&\underset{\mathbf{z}}{\text{minimize}} \quad \||\mathbf{d}| - \mathbf{z}\|^2 + \lambda \|\mathbf{Rz}\|^2 \\
&\text{subject to} \quad \mathbf{z} \geq \mathbf{0}
\end{aligned}
\tag{8.21}
$$

Since the globally optimal solution of (8.21) can be computed, this special case can be used as benchmark to compare the IQML and MACO algorithms. In this example, $N = 100$ and $\lambda = 1$ was used. Each element in $\mathbf{d}$ and the initialization $\mathbf{x}_0$, was generated from uniform distributions for both the phase (between 0 and 2π) and the magnitude (between 0 and 1), and the elements of $\mathbf{R} \in \mathbb{R}^{100 \times 100}$ were drawn from a zero-mean Gaussian distribution with unit variance.

Monte Carlo simulations were performed by generating 1000 random initializations, and using these to start each algorithm. The problem parameters, $\mathbf{d}$ and $\mathbf{R}$ were kept fixed in all simulations. The resulting mean criterion as a function of the iteration number is shown in Fig. 8.1, together with the spread in terms of two standard deviations. As can be seen, the proposed method converges to the optimal solution in less than 20 iterations for all initializations, while IQML does not converge at all. Even the initialization given by (8.4) resulted in a similar behavior (not shown here). This indicates that IQML will have poor performance in some cases, which is a partial motivation for the local minimization algorithm presented herein.

8.5.2 Example 2: An MRI design

To make this example easy to follow, we will consider the problem with a fully sampled rectangular grid in k-space, no weighting, and a single transmitter coil. Furthermore, $\lambda = 10$, and the desired 2D excitation pattern, $\mathbf{D} \in \mathbb{R}^{32 \times 32}$, is a 10×10 square passband with unit magnitude, centered in space, as shown in Fig. 8.3a. In 2D, the problem can be formulated as

$$
\underset{\mathbf{X}}{\arg\min} \left\|\mathbf{D} - \mathbf{FXF}^{\mathrm{T}}\right\|_{\mathrm{F}}^2 + \lambda \left\|\mathbf{R}|\text{vec}(\mathbf{X})|\right\|^2,
\tag{8.22}
$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm, $\mathbf{F} \in \mathbb{C}^{32 \times 32}$ is a inverse discrete Fourier transform matrix, and $\text{vec}(\cdot)$ is the columnwise vectorizing operator. By letting $\mathbf{x} = \text{vec}(\mathbf{X}) \in \mathbb{C}^{1024}$, $\mathbf{d} = \text{vec}(\mathbf{D}) \in \mathbb{R}^{1024}$, and $\mathbf{A} = \mathbf{F} \otimes \mathbf{F} \in \mathbb{C}^{1024 \times 1024}$, we can re-write the problem in the same form
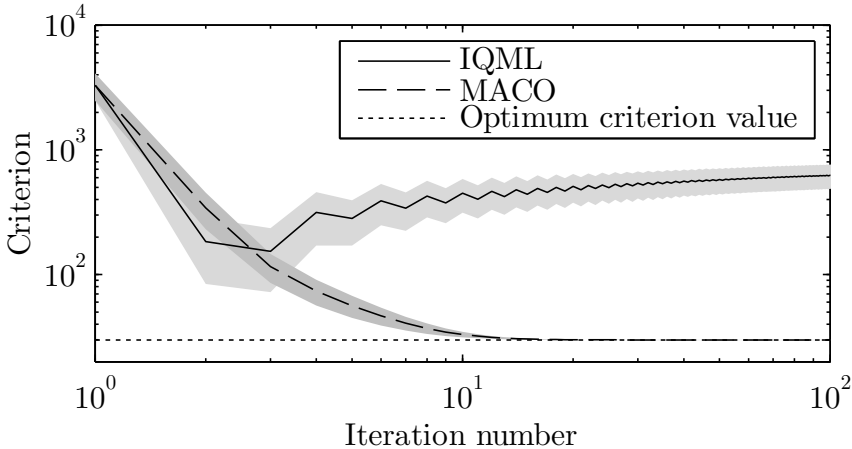
Figure 8.1: The mean criterion for IQML and MACO versus the number of iterations, when applied to the simple design problem of (8.21) using 1000 random initializations. The light and dark gray fields show the spread of the criterion ($\pm 2\sigma$) for IQML and MACO, respectively.

as (8.19). Here, $\mathbf{A}^*\mathbf{A}$ becomes diagonal, and as was shown in Section 8.3.2, the optimal phase vector $\phi^\star$ will therefore be independent of $\mathbf{z}$. As a consequence, MACO-LCQP will reach the optimum in one iteration. Here, the LCQP was solved by MOSEK. Again, it should be noted that solving an LCQP might become intractable for large problems, in which case the elementwise update approach is preferable.

MACO, MACO-LCQP, and IQML were used to find the solution to the problem in (8.22). The initial guess for all algorithms was obtained by solving the least-squares problem in (8.4). The convergence in terms of the criterion function versus the number of iterations is shown in Fig. 8.2. The magnitudes of the excitation patterns obtained after 30 iterations are shown in Fig. 8.3b–d. The resulting stopband and passband ripples, together with the sub-criteria for the fit (first term of (8.1)) and the magnitude smoothing (second term of (8.1)), are listed in Table 8.1. The stopband and passband ripples were defined as the maximum magnitude deviation from the desired excitation pattern in the respective areas.

In this example, the regularization is easier to handle than in the first example, and IQML converges. MACO-LCQP converges in one iteration, as expected, while the standard MACO has a slightly slower convergence rate. The time until convergence, with a tolerance of 10^{-6}, was 1041 s, 40 s, and 4 s, for IQML, MACO, and MACO-LCQP, respectively. At iteration 30, MACO closely approximates the MACO-LCQP solution, while IQML provides a smoother estimate with both lower fit and higher ripple values.
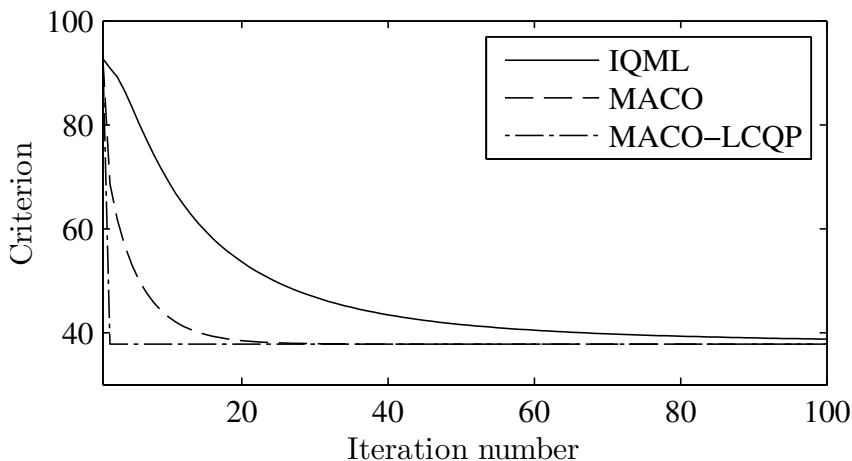
Figure 8.2: Comparison of the criterion for IQML, MACO, and MACO-LCQP versus the number of iterations, when applied to the MRI example.

Table 8.1: Ripples and sub-criteria at iteration 30 for the different methods when applied to the MRI example.

	MACO	MACO-LCQP	IQML
Passband ripple	0.56	0.53	0.60
Stopband ripple	0.56	0.57	0.75
Fit-term	26.9	26.7	35.9
Smoothness-term	1.10	1.11	1.04

For smaller values of λ, that is, less smoothness imposed, IQML might outperform MACO for a given initialization as it does not get stuck in local minima. However, IQML would typically be used to initialize MACO in these cases, and therefore an improvement can still be expected.

8.6 Conclusion

We have derived a simple algorithm with low computational complexity, for solving LS problems with magnitude constraints. The proposed MACO algorithm does not suffer from the potential convergence problems of IQML, and can further improve the results from IQML by truly minimizing the design criterion. The algorithm is useful for designing RF pulse excitation sequences in parallel MRI, which can be transmitted without compromising signal fidelity in the amplifier stage.
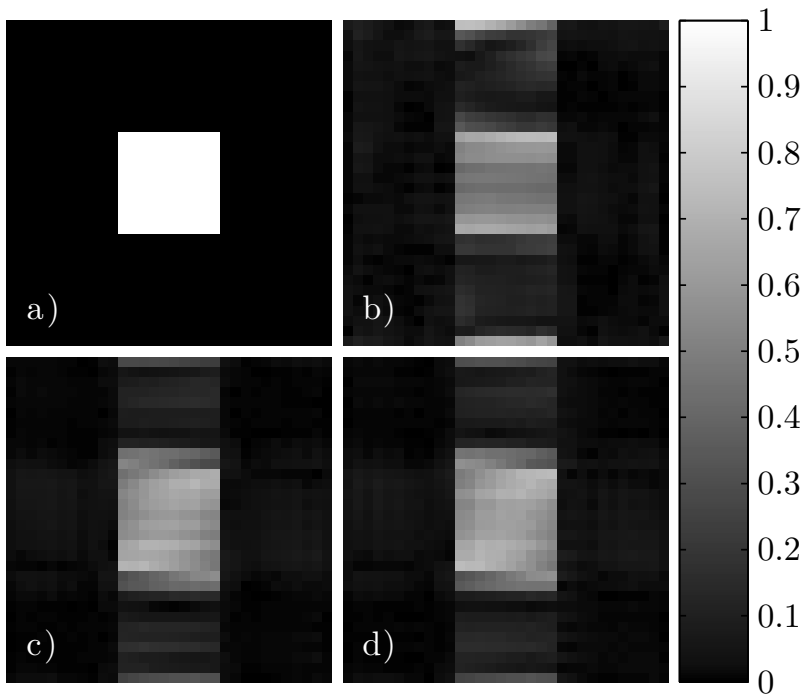
Figure 8.3: a) The desired excitation pattern for the MRI example. Excitation patterns corresponding to the sequences designed by: b) IQML, c) MACO, and d) MACO-LCQP, obtained after 30 iterations.

Chapter 9

Magnetic resonance thermometry

9.1 Introduction

The proton resonance frequency (PRF) generally depends on the local molecular bonds, and for the hydrogen found in water, this frequency varies with temperature. By detecting the resonance frequency of hydrogen in water, the relative temperature distribution can be determined by a linear relation [82]. However, the absolute water resonance frequency cannot be estimated directly, due to an unknown static magnetic field inhomogeneity. Here, fat-bound hydrogen is used as a reference, as its frequency is largely independent of the temperature, enabling noninvasive absolute temperature mapping in three dimensions, through MRI.

Separating the water and fat contributions in an image can improve diagnosis, and is useful to quantify the amount of fatty tissue and estimate its distribution. In fat-water separation, the amplitudes of the resonances are of main interest, given that the frequency separation is known [99]. In the present application we are interested in estimating the frequency separation of the fat and water signals, which further complicates the problem. The PRF-based fat referenced approach is one out of several methods of absolute MR thermometry, cf. [101, 58].

MR thermometry can be used to monitor tissue temperature, which for example can be used to guide thermal therapy [37, 125]. Another application is to detect the activation of brown adipose tissue, which in turn is related to metabolism [50]. Studying this activation could be important when developing future treatments for obesity.

The fat signal contains several resonances depending on the internal bonds in the fat molecules, but the methylene peak is the most significant. The water-bound protons, on the other hand, correspond to a single resonance. Figure 9.1 shows an example of the theoretical signal spectrum from a simulated tissue containing equal proportions
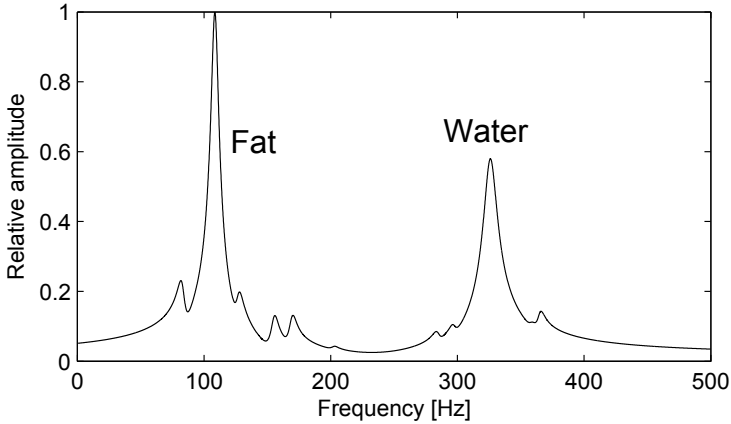
Figure 9.1: Example of the theoretical spectrum of a received signal from one voxel, including a frequency offset caused by field inhomogeneity. The multiple smaller peaks originate from fat.

of fat and water. By modeling this signal, nonlinear estimation methods can be used to find the parameters, and hence the temperature. A fat-referenced parametric-modeling approach has been previously attempted with different models, both in the time and frequency-domain [82, 117].

The contribution of this chapter is to: i) extend the model in [82] to include multiple fat resonances, ii) analyze the Cramér-Rao bound with respect to the experimental setup, particularly in fat tissue, and iii) discuss identifiability and parameter estimation. We illustrate certain problems and limitations, both from a physical and a theoretical point of view. The main focus is temperature mapping in fat (adipose) tissue, which to our knowledge has not been explicitly treated in the literature before. Previous analyses have been performed for cases that do not reflect the typical fat/water distribution found in fat tissue [115]. The goal is to answer fundamental questions regarding feasibility and experiment design, which is essential before attempting to develop more efficient algorithms for estimation.

9.2 Signal model

The time-domain model for the signal in each voxel is given by

$$s(t) = \left(\rho_{\mathrm{w}} e^{-\nu_{\mathrm{w}} t} e^{i\omega_{\mathrm{w}} t} + \rho_{\mathrm{f}} e^{-\nu_{\mathrm{f}} t} \sum_{r=1}^{R} \alpha_r e^{i\omega_r t} \right) e^{i\omega_0 t} + v(t)$$

$$= f(\boldsymbol{\theta}, t) + v(t), \tag{9.1}$$

where $\rho_{\mathrm{w}}, \rho_{\mathrm{f}}$ are the complex-valued water and fat amplitudes; $\nu_{\mathrm{w}}, \nu_{\mathrm{f}}$ are the water and fat damping factors, corresponding to $1/T_2^*$; ω_{w} is the water resonance frequency; $\{\omega_r, \alpha_r\}_{r=1}^R$ is the set of fat resonance frequencies, and the corresponding relative amplitudes; and ω_0 is the frequency shift caused by the magnetic field inhomogeneity ΔB_0. Furthermore, (9.1) defines the noise-free signal model $f(\boldsymbol{\theta}, t)$, where $\boldsymbol{\theta} = [\mathrm{Re}\{\rho_w\}, \mathrm{Im}\{\rho_w\}, \nu_{\mathrm{w}}, \omega_{\mathrm{w}}, \mathrm{Re}\{\rho_f\}, \mathrm{Im}\{\rho_f\}, \nu_{\mathrm{f}}, \omega_0]^{\mathrm{T}}$ is the vector of unknown real-valued model parameters, and $v(t)$ is i.i.d. complex-Gaussian noise. The resonance frequencies, and hence the spectrum, depend linearly on the applied static magnetic field B_0 according to (2.1). Ideally, the fat and water components should be in phase, but due to system imperfections their phases are modeled independently through $\rho_{\mathrm{w}}, \rho_{\mathrm{f}}$.

The parameters $\{\omega_r, \alpha_r\}_{r=1}^R$ are considered known, giving a predefined fat profile that can be moved spectrally and scaled. The amplitudes α_r are normalized to achieve $\sum_r \alpha_r = 1$. To simplify the notation we introduce the known complex-valued function $F(t)$ describing the fat resonance profile:

$$F(t) = \sum_{r=1}^R \alpha_r e^{i\omega_r t}. \tag{9.2}$$

The absolute water resonance frequency can be mapped to temperature by a linear mapping:

$$T = a\omega_{\mathrm{w}} + b, \tag{9.3}$$

where a and b are assumed known. Using the calibration in [82] at $B_0 = 1.5$ T, the constants $a = -0.244°\mathrm{C}/(\mathrm{rad}/\mathrm{s})$ and $b = 499.7°\mathrm{C}$ are obtained. It should be noted that this calibration implicitly depends on the static field strength B_0; in fact $a \propto 1/B_0$. This means that increasing the field strength will improve the conditioning of the temperature calculation, something which will be discussed further in Section 9.6.1.

The measured signal modeled by (9.1) consists of damped complex-valued exponentials, and therefore, the instantaneous signal to noise ratio (SNR) will decrease over time. This makes it difficult to define SNR in a reasonable way. In the following we use the SNR definition

$$\mathrm{SNR} = \frac{\sum\limits_{n=1}^N |f(\boldsymbol{\theta}, t_n)|^2}{N\sigma^2}, \tag{9.4}$$

where σ is the noise standard deviation and N is the number of samples. This definition represents the total amount of information in the data, but it will also be strongly dependent on the damping factors $\nu_{\mathrm{w}}, \nu_{\mathrm{f}}$. This means that simulations at constant SNR can be hard to interpret

when changing the parameters. In practice, the noise level is constant, and in the following we want to compare effects of changing model parameters that also influence the average signal power. Therefore, a given value of the SNR will correspond a fixed noise level, based on a set of predefined parameters, unless stated otherwise. For measured data, the noise variance can be estimated from the image background, where no tissue is present. Furthermore, the signal power can be estimated using the power of the noisy data and the noise power.

Other definitions of SNR could also be considered, given that they provide a straightforward method for estimating the SNR from data. For example, defining a separate SNR for the water and fat signals could be useful in simulation, since high enough signals from both resonances are needed for accurate estimation, as discussed in the next section. However, to achieve the same goal, we will in the following refer to the total SNR and the corresponding proportions of water and fat.

9.3 Practical considerations

Estimation of ω_{w} and ω_0 requires that both the water and fat resonances are nonzero. In the case with no water present, ω_0 can be estimated but no information regarding ω_{w} is available. In the case of no fat present, ω_{w} and ω_0 cannot be separated since they occur additively, hence we cannot get an absolute measure of ω_{w}, which is the sought quantity.

In practice, most tissue types contain a large proportion of either water or fat. For example, fat tissue typically contains about 3-5% water (in terms of the resonance amplitude), and normal muscle tissue contains virtually no fat at all. Because of this, fat referenced MR thermometry cannot be used in all tissue types. There are, nonetheless, tissue types where both water and fat are present to greater extent, such as bone marrow [117], but henceforth we will focus on fat tissue, which represents a difficult case with low water content.

For the problem at hand, we can define the performance in terms of standard deviations of the parameter estimates. To accurately detect natural variations in the body temperature a standard deviation of at most $0.1°C$ is likely to be needed, while for guiding thermal therapy, a standard deviation of $1°C$ might be sufficient.

Extending the scan time generally enables a higher SNR, as discussed in 2.2.5; however, the maximum total scan time will be limited in practice. Here we assume that there is a maximum scan time allowed, but this time will vary depending on the application. Furthermore, the MRI scanner used can limit the choice of sampling intervals, and will define the static magnetic field strength B_0.

9.4 The Cramér-Rao Bound

Under the model assumption of Gaussian noise, the FIM is conveniently given by (3.14), which only requires the Jacobian with respect to the model parameter vector $\boldsymbol{\theta}$, to be computed. Performing the differentiation gives the Jacobian vector

$$\frac{\partial f(\boldsymbol{\theta}, t)}{\partial \boldsymbol{\theta}} = \begin{bmatrix} e^{(-\nu_{\mathrm{w}}+i(\omega_{\mathrm{w}}+\omega_0))t} \\ ie^{(-\nu_{\mathrm{w}}+i(\omega_{\mathrm{w}}+\omega_0))t} \\ -\rho_{\mathrm{w}}te^{(-\nu_{\mathrm{w}}+i(\omega_{\mathrm{w}}+\omega_0))t} \\ i\rho_{\mathrm{w}}te^{(-\nu_{\mathrm{w}}+i(\omega_{\mathrm{w}}+\omega_0))t} \\ F(t)e^{(-\nu_{\mathrm{f}}+i\omega_0)t} \\ iF(t)e^{(-\nu_{\mathrm{f}}+i\omega_0)t} \\ -\rho_{\mathrm{f}}tF(t)e^{(-\nu_{\mathrm{f}}+i\omega_0)t} \\ ite^{i\omega_0 t}(\rho_{\mathrm{w}}e^{(-\nu_{\mathrm{w}}+i\omega_{\mathrm{w}})t} + \rho_{\mathrm{f}}F(t)e^{-\nu_{\mathrm{f}}t}) \end{bmatrix}. \tag{9.5}$$

The CRB matrix $\mathbf{C}_{\mathrm{CRB}}$ is then obtained through (3.12). The analytical expression of the CRB matrix is hard to analyze, and we therefore resort to numerical analysis. For MR thermometry, the variance of the temperature estimate is of most importance, and henceforth, the CRB will refer to the bound on $\hat{T}$.

9.5 Experimental setup

As a representative case for fat tissue the following true parameters were chosen for $B_0 = 1.5$ T (unless stated otherwise):

$$\begin{aligned} \omega_{\mathrm{w}} &= 1887 \text{ rad/s}, & \omega_0 &= 241 \text{ rad/s}, \\ \rho_{\mathrm{w}} &= 0.05e^{i\pi/4}, & \rho_{\mathrm{f}} &= 0.95e^{2i\pi/9}, \\ \nu_{\mathrm{w}} &= 35, & \nu_{\mathrm{f}} &= 20. \end{aligned} \tag{9.6}$$

The frequencies are given relative to the demodulation performed at the receiver. The numerical values of the initial phases are not essential for the analysis, and have been chosen arbitrarily. Note, however, that there is a small phase difference between the water and fat components, which represents a deviation from the ideal scenario. The assumed experimental setup was: SNR = 25 dB, $B_0 = 1.5$ T and $N = 32$ samples at $t_k = t_0 + k\Delta t, \ k = 0, \ldots, N-1$, where the sampling interval is $\Delta t = 3.5$ ms and $t_0 = 2.4$ ms. The parameters of the fat profile, defined by (9.2), are given in Table 9.1, and the corresponding theoretical spectrum of the signal, specified by (9.6), is shown in Fig. 9.2. As can be seen, the water resonance is relatively small, and is comparable in size to the less significant fat resonances.
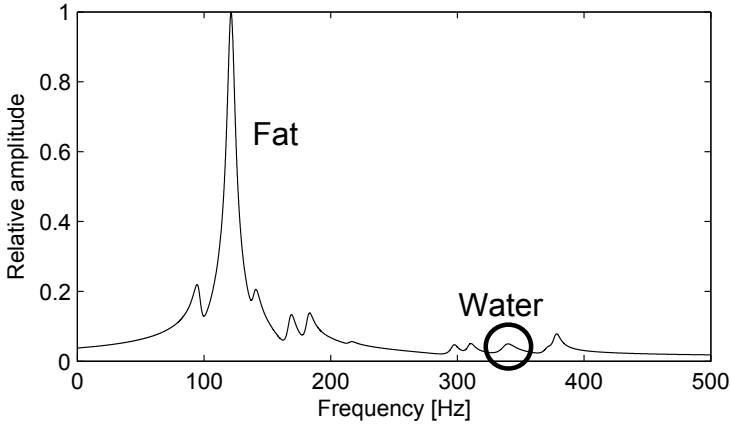
Figure 9.2: The theoretical spectrum of the received signal from fat-tissue, based on the simulated parameters in (9.6).

Table 9.1: Parameters of the fat profile defined by (9.2).

r	α_r	ω_r (rad/s)
1	0.0834	361.3
2	0.6581	521.9
3	0.0556	638.4
4	0.0592	815.0
5	0.0556	903.4
6	0.0060	1112.1
7	0.0186	1626.1
8	0.0186	1706.4
9	0.0093	2091.8
10	0.0356	2131.9

9.6 Results and discussion

9.6.1 Simulation

A plot of the CRB of the temperature estimates, for different relative water and fat contents and constant noise variance, is shown in Fig. 9.3. As can be seen, the lower bound approaches infinity as either the water or fat content goes to zero. This is expected since both components are needed for absolute temperature estimation. For approximately equal amounts of fat and water, the CRB is minimal, however, this fat/water distribution is uncommon in practice. The minimum CRB is slightly shifted towards higher water content due to the more rapid damping of this component, as indicated in (9.6). For fat tissue containing 5% water and 95% fat, the minimum standard deviation is significantly increased, making the corresponding estimates less reliable.
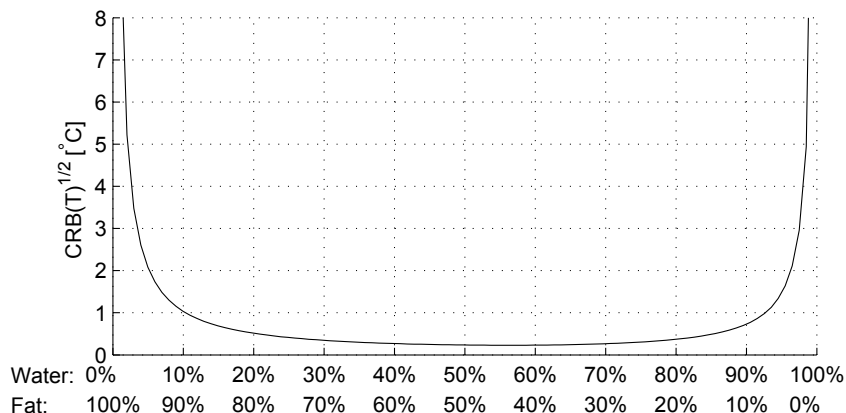
Figure 9.3: The CRB of the temperature (standard deviation) for different proportions of water and fat ($|\rho_{\mathrm{w}}|, |\rho_{\mathrm{f}}|$) at a fixed noise variance (at $|\rho_{\mathrm{w}}| = 0.05$ and $|\rho_{\mathrm{f}}| = 0.95$, SNR = 25 dB). The other parameters were kept fixed according to the representative case.

The CRB is only weakly dependent on the true values of the frequencies $\omega_{\mathrm{w}}, \omega_0$ (not shown here); however, as shown in Fig. 9.4, the CRBs dependence on the two damping factors is more significant. As can be seen, the water damping factor ν_{w} has a large influence on the obtainable performance in this fat-tissue example, while the dependence on ν_{f} is weaker. This can be expected, since a higher damping of the water signal will effectively reduce the information in an already weak signal. Since the fat signal is large, the effect of more rapid damping is relatively small. The chosen representative case, given by (9.6), is just one out of many possible scenarios, with vastly different CRB of the temperature estimates; however, as can be seen from Fig. 9.4, the problem is difficult for a wide range of ν_{w} and ν_{f}, and not only for the chosen default parameters. If the field inhomogeneity ΔB_0 can be kept low by tuning the hardware, the damping of the water component ν_{w} will be slow, effectively improving the SNR for the small water component. However, even in the best case shown in Fig. 9.4, that is, $\nu_{\mathrm{w}} = 0.5$ and $\nu_{\mathrm{f}} = 10$, the CRB is 0.32 °C, which is relatively high for detecting natural temperature variations in the body. Furthermore, this scenario is not practical, as it would correspond to the decay rates for free water and fat in a perfectly homogeneous field. The main bottleneck for the CRB is the low amplitude of the water component, and this cannot be changed by improving the hardware.

The CRB of the temperature estimate versus SNR, using a 1.5 T, 3 T, and 7 T (B_0) scanner, is shown in Fig. 9.5. Increasing the magnetic field scales the spectrum of the signal and gives a larger spectral separation between the resonances. However, this only mildly influences the
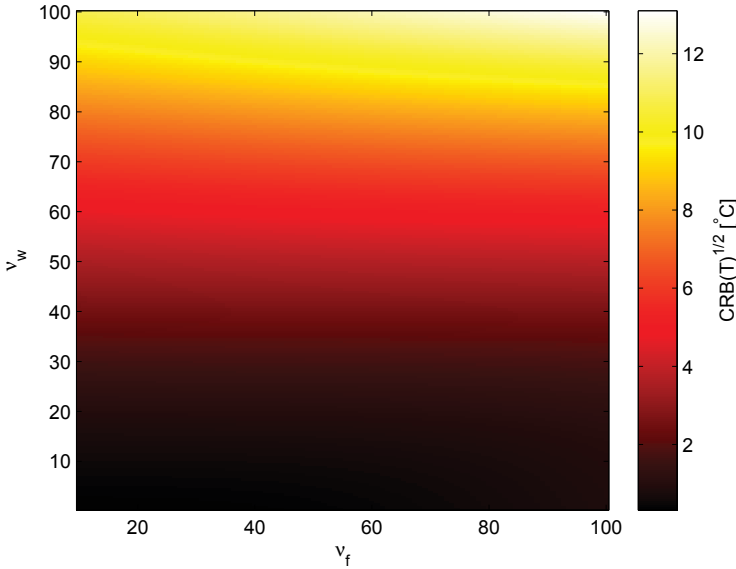
Figure 9.4: The CRB of the temperature (standard deviation) for different damping factors ν_w and ν_f at a fixed noise variance corresponding to SNR = 25 dB using the parameters in (9.6).

CRB of ω_w. Increasing the frequency of the received signal, however, leads to a higher frequency sensitivity, as a given frequency error will have a larger impact on the model fit. Moreover, increasing B_0 also improves the conditioning of the temperature transformation given by (9.3), as calibration coefficients depend on the field strength. This significantly lowers the CRB of $\hat{T}$ for a given SNR, but in practice, the SNR would also improve by increasing B_0, providing an additional boost in estimation performance.

For a uniform sampling interval Δt, the corresponding optimal number of samples can be studied with respect to the CRB. In general, optimizing the sampling times with respect to the CRB requires knowledge of the true system. However, some overall properties are common to all choices of model parameters. Figure 9.6 shows how the bound depends on the sampling interval for the parameter set given in (9.6), assuming a fixed number of samples ($N = 32$). As can be seen, the bound significantly increases for some choices of Δt. This can be explained by aliasing causing signal cancellation. In fact, the increased bound occurs when $2\pi/\Delta t$ is approximately equal to the frequency spacing between the water resonance and the largest fat resonance (at 1.5 T). In the case of a single fat resonance, this particular choice of sampling interval would cause the two aliased resonance frequencies to coincide, indepen-
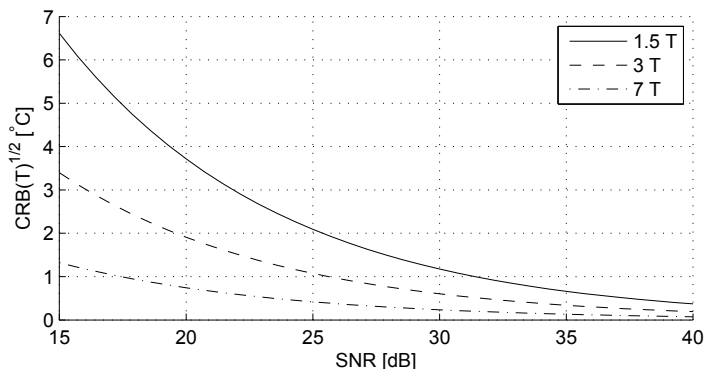
Figure 9.5: CRB of the temperature (standard deviation) for different static magnetic field strengths B_0, versus SNR.

dent of ω_0 and the other model parameters. It should be noted that although the variance increases in this case, the separation of the peaks is still possible, as there are several fat peaks as well as different damping factors of the water and fat components. In practice, the frequency spacing is approximately known, so the corresponding range of sampling intervals can be avoided.

Given a desired range of sampling intervals, the best sampling strategy for lowering the CRB within a predefined time frame can be examined; decreasing Δt and getting more samples or decreasing the number of samples and applying averaging to increase the SNR. The results are of course data dependent, but can still illustrate the usefulness of such optimizations. The CRB as a function of both Δt and N with compensation for the averaging that can be applied if the acquisition time is reduced, is shown in Fig. 9.7. The zeroed-out region (black) does not comply with the time limit of original sequence, and is therefore not of interest. Averaging cannot be done over non-integer numbers of experiments, which accounts for the banded structure in Fig. 9.7. It can be seen that maintaining the initial sampling interval $\Delta t = 3.5$ ms and averaging over two acquisitions of 16 samples is preferable, compared to using 32 samples in one acquisition. This can be explained by the fact that the signal samples later in the sequence carry less information due to the damping. Figure 9.7 also shows that using a shorter sampling interval, if possible in practice, can lower the CRB. However, a minimal sampling interval is not desired in general, since the total acquisition time would be too short to capture the relevant frequencies, which again increases the CRB. It should be noted that the gain from optimizing the sampling scheme subject to a fixed total scan time is relatively small, but any means of improving the estimates are essential to make fat tissue temperature mapping practically feasible.
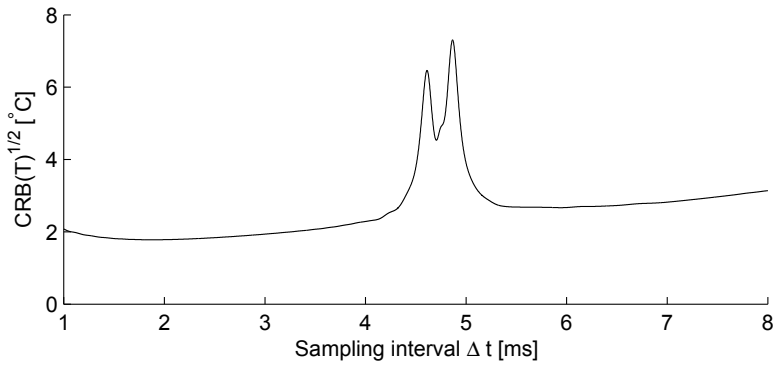
Figure 9.6: The CRB of the temperature (standard deviation) versus the sampling interval Δt for $N = 32$ samples, SNR = 25 dB.
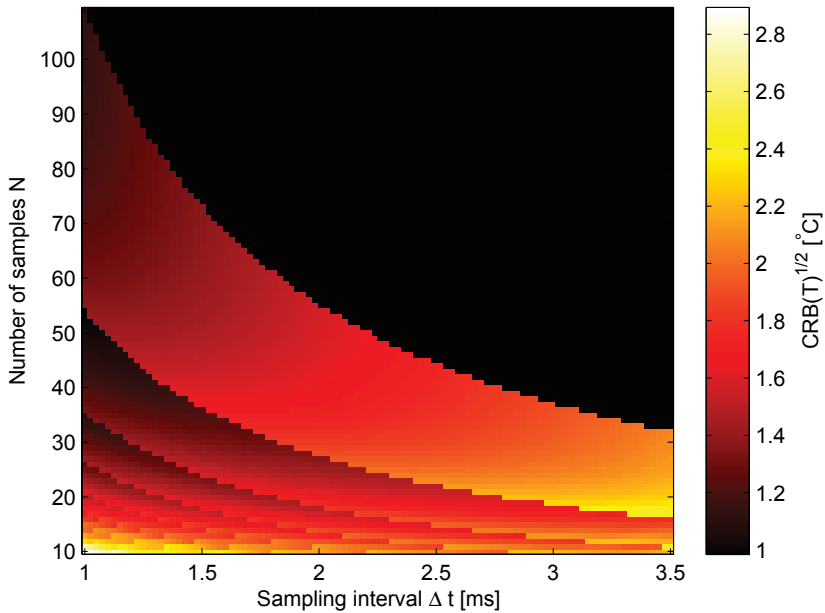

Figure 9.7: The CRB of the temperature (standard deviation) versus Δt and N with compensation for the averaging that can be done when the measurement time is reduced.
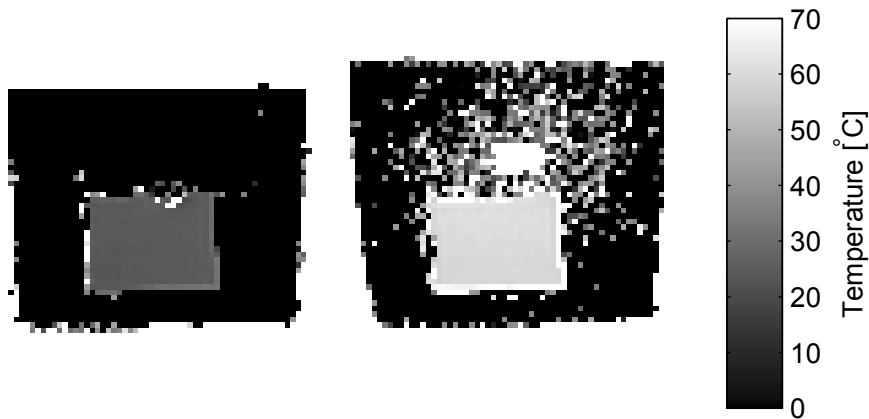
Figure 9.8: Temperature estimates for a phantom image consisting of two cartons of whipping cream at different temperatures, each lowered in a box of water. The averages were 24.5(0.63)°C (left), and 60.1(0.93)°C (right), while the true temperature was measured to be 21°C (left), and 55°C (right). The noise in the water region of the image should be disregarded, as this region is not identifiable.

9.6.2 Phantom data

To show the validity and practical use of fat referenced absolute MR thermometry, the temperatures in a collected phantom image were estimated. The NLS criterion was minimized voxelwise, by a properly initialized Gauss-Newton algorithm. The phantom contained two cartons of whipping cream (40% fat weight) at different temperatures (left: 21°C, right: 55°C), each lowered in a box of water. The data consisted of $N = 32$ samples with a sampling interval of $\Delta t = 3.0$ ms, acquired with a 1.5 T scanner. The estimated average SNR was 26.4 dB. The temperature map is shown in Fig. 9.8. Using the calibration from [82] the corresponding estimates were 24.5°C and 60.1°C when averaged over the interior of the two cartons, respectively. The corresponding standard deviations were 0.63°C and 0.93°C which is higher than what is expected from the CRB, assuming similar decay constants as in the simulations; but the obtained variance is significantly lower than what is expected in fat tissue. The bias is likely due to imperfect calibration. It should be noted that the water surrounding the cream phantom does not enable unique identification of the parameters, which causes the partially noisy appearance of Fig. 9.8. The smooth sections of the background are due to imposed constraints on the estimates w_{w}.

9.7 Conclusions

There are fundamental limitations that make fat-referenced PRF-based temperature estimation difficult, the main reason being that it requires a detectable signal from both fat and water. In fat tissue, methods for obtaining high SNR are needed to enable the use of a 1.5 T scanner, due to the low water content. Applying higher field strengths can significantly improve the estimation performance, given that the homogeneity of the static B_0 field is not compromised which could otherwise increase the damping. In general, the problem of temperature estimation in fat tissue is sensitive to the assumptions made. An accurate model with multiple fat peaks is needed to separate the small water component from the fat, but this model also relies on calibration. Application specific optimization of the sampling scheme is possible but provides a relatively small gain. It is, however, important to avoid specific choices of the sampling interval to prevent signal cancellation due to aliasing. Intelligent estimation algorithms that utilizes as much of the available prior information as possible can help to provide high precision estimates of the absolute temperature, but the potential model mismatch can induce bias, which in turn can be a problem in some applications.

References

[1] C. Ahn and Z. Cho. Analysis of the eddy-current induced artifacts and the temporal compensation in nuclear magnetic resonance imaging. *IEEE Transactions on Medical Imaging*, 10(1):47–52, 1991.

[2] E. Alonso-Ortiz, I. R. Levesque, and G. B. Pike. MRI-based myelin water imaging: A technical review. *Magnetic Resonance in Medicine*, 73(1):70–81, 2015.

[3] P. Babu and P. Stoica. Connection between SPICE and square-root LASSO for sparse parameter estimation. *Signal Processing*, 95:10–14, 2014.

[4] R. Bai, C. G. Koay, E. Hutchinson, and P. J. Basser. A framework for accurate determination of the T_2 distribution from multiple echo magnitude MRI images. *Journal of Magnetic Resonance*, 244:53 – 63, 2014.

[5] N. K. Bangerter, B. A. Hargreaves, S. S. Vasanawala, J. M. Pauly, G. E. Gold, and D. G. Nishimura. Analysis of multiple-acquisition SSFP. *Magnetic Resonance in Medicine*, 51(5):1038–47, 2004.

[6] J. K. Barral, E. Gudmundson, N. Stikov, M. Etezadi-Amoli, P. Stoica, and D. G. Nishimura. A robust methodology for in vivo T_1 mapping. *Magnetic Resonance in Medicine*, 64(4):1057–1067, 2010.

[7] P. J. Basser and D. K. Jones. Diffusion-tensor MRI: theory, experimental design and data analysis – a technical review. *NMR in Biomedicine*, 15(7-8):456–467, 2002.

[8] J. Berglund, L. Johansson, H. Ahlström, and J. Kullberg. Three-point Dixon method enables whole-body water and fat imaging of obese subjects. *Magnetic Resonance in Medicine*, 63(6):1659–1668, 2010.

[9] M. Bernstein, K. King, and X. Zhou. *Handbook of MRI Pulse Sequences*. Academic Press. Elsevier Academic Press, Burlington, MA, USA, 2004.

[10] O. Bieri, K. Scheffler, G. H. Welsch, S. Trattnig, T. C. Mamisch, and C. Ganter. Quantitative mapping of T_2 using partial spoiling. *Magnetic Resonance in Medicine*, 66(2):410–418, 2011.

[11] J. Bioucas-Dias, M. Figueiredo, and J. Oliveira. Total variation-based image deconvolution: a majorization-minimization approach. In *Proc. 31st IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-2006)*, volume 2, pages 861–864, May 2006.

[12] T. A. Bjarnason, C. Laule, J. Bluman, and P. Kozlowski. Temporal phase correction of multiple echo T_2 magnetic resonance images. *Journal of Magnetic Resonance*, 231:22–31, 2013.

[13] M. Björk, J. Berglund, J. Kullberg, and P. Stoica. Signal modeling and the Cramér-Rao Bound for absolute magnetic resonance thermometry in fat tissue. In *Proc. 45th Asilomar Conference on Signals, Systems, and Computers*, pages 80–84, Pacific Grove, CA, USA, 2011.

[14] M. Björk, E. Gudmundson, J. K. Barral, and P. Stoica. Signal processing algorithms for removing banding artifacts in MRI. In *Proc. 19th European Signal Processing Conference (EUSIPCO-2011)*, pages 1000–1004, Barcelona, Spain, 2011.

[15] M. Björk, R. R. Ingle, J. K. Barral, E. Gudmundson, D. G. Nishimura, and P. Stoica. Optimality of equally-spaced phase increments for banding removal in bSSFP. In *Proc. 20th Annual Meeting of ISMRM*, page 3380, Melbourne, Australia, 2012.

[16] M. Björk, R. R. Ingle, E. Gudmundson, P. Stoica, D. G. Nishimura, and J. K. Barral. Parameter estimation approach to banding artifact reduction in balanced steady-state free precession. *Magnetic Resonance in Medicine*, 72(3):880–892, 2014.

[17] M. Björk and P. Stoica. Fast denoising techniques for transverse relaxation time estimation in MRI. In *Proc. 21st European Signal Processing Conference (EUSIPCO-2013)*, pages 1–5, Marrakech, Morocco, 2013.

[18] M. Björk and P. Stoica. Magnitude-constrained sequence design with application in MRI. In *Proc. 39th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-2014)*, pages 4943–4947, Florence, Italy, 2014.

[19] M. Björk and P. Stoica. New approach to phase correction in multi-echo T_2 relaxometry. *Journal of Magnetic Resonance*, 249:100–107, 2014.

[20] M. Björk, D. Zachariah, J. Kullberg, and P. Stoica. A multicomponent T_2 relaxometry algorithm for myelin water imaging of the brain. *Magnetic Resonance in Medicine*, 2015. DOI: 10.1002/mrm.25583.

[21] M. Blaimer, F. Breuer, M. Mueller, R. M. Heidemann, M. A. Griswold, and P. M. Jakob. SMASH, SENSE, PILS, GRAPPA: how to choose the optimal method. *Topics in Magnetic Resonance Imaging*, 15(4):223–236, 2004.

[22] K. T. Block, M. Uecker, and J. Frahm. Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint. *Magnetic Resonance in Medicine*, 57(6):1086–1098, 2007.

[23] P. Bloomfield and W. Steiger. Least absolute deviations curve-fitting. *SIAM Journal on Scientific and Statistical Computing*, 1(2):290–301, 1980.

[24] J.-M. Bonny. Methods and applications of quantitative MRI. *Annual Reports on NMR Spectroscopy*, 56:213–229, 2005.

[25] J.-M. Bonny, O. Boespflug-Tanguly, M. Zanca, and J.-P. Renou. Multi-exponential analysis of magnitude MR images using a

quantitative multispectral edge-preserving filter. *Journal of Magnetic Resonance*, 161(1):25–34, 2003.

[26] J.-M. Bonny, M. Zanca, J.-Y. Boire, and A. Veyre. T_2 maximum likelihood estimation from multiple spin-echo magnitude images. *Magnetic Resonance in Medicine*, 36(2):287–293, 1996.

[27] M. R. Borich, A. L. Mackay, I. M. Vavasour, A. Rauscher, and L. A. Boyd. Evaluation of white matter myelin water fraction in chronic stroke. *Neuroimage Clin*, 2:569–580, 2013.

[28] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004.

[29] G. L. Bretthorst. How accurately can parameters from exponential models be estimated? A Bayesian view. *Concepts in Magnetic Resonance Part A*, 27A(2):73–83, 2005.

[30] H. Y. Carr and E. M. Purcell. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Physical Review*, 94:630–638, 1954.

[31] T. Chan, J. Shen, and H.-M. Zhou. Total variation wavelet inpainting. *Journal of Mathematical Imaging and Vision*, 25(1):107–125, 2006.

[32] S. Chavez, Q.-S. Xiang, and L. An. Understanding phase maps in MRI: a new cutline phase unwrapping method. *IEEE Transactions on Medical Imaging*, 21(8):966–977, 2002.

[33] J. Y. Cheng, M. T. Alley, C. H. Cunningham, S. S. Vasanawala, J. M. Pauly, and M. Lustig. Nonrigid motion correction in 3D using autofocusing with localized linear translations. *Magnetic Resonance in Medicine*, 68(6):1785–1797, 2012.

[34] M. S. Cohen. Parametric analysis of fMRI data using linear systems methods. *NeuroImage*, 6(2):93–103, 1997.

[35] S. C. Cripps. *RF Power Amplifiers for Wireless Communications*. Artech House, Inc., Norwood, MA, USA, 2nd edition, 2006.

[36] C. H. Cunningham, J. M. Pauly, and K. S. Nayak. Saturated double-angle method for rapid B1+ mapping. *Magnetic Resonance in Medicine*, 55(6):1326–1333, 2006.

[37] B. Denis de Senneville, B. Quesson, and C. T. W. Moonen. Magnetic resonance temperature imaging. *International Journal of Hyperthermia*, 21(6):515–531, 2005.

[38] S. C. Deoni. Transverse relaxation time (T_2) mapping in the brain with off-resonance correction using phase-cycled steady-state free precession imaging. *Journal of Magnetic Resonance Imaging*, 30(2):411–417, 2009.

[39] S. C. Deoni and S. H. Kolind. Investigating the stability of mcDESPOT myelin water fraction values derived using a stochastic region contraction approach. *Magnetic Resonance in Medicine*, 73(1):161–169, 2015.

[40] S. C. Deoni, B. K. Rutt, T. Arun, C. Pierpaoli, and D. K. Jones. Gleaning multicomponent T_1 and T_2 information from steady-state imaging data. *Magnetic Resonance in Medicine*, 60(6):1372–1387, 2008.

[41] S. C. Deoni, B. K. Rutt, and T. M. Peters. Rapid combined T_1 and T_2 mapping using gradient recalled acquisition in the steady state. *Magnetic Resonance in Medicine*, 49(3):515–526, 2003.

[42] A. N. Dula, D. F. Gochberg, and M. D. Does. Optimal echo spacing for multi-echo imaging measurements of bi-exponential T_2 relaxation. *Journal of Magnetic Resonance*, 196(2):149–156, 2009.

[43] C. L. Dumoulin, S. P. Souza, M. F. Walker, and W. Wagle. Three-dimensional phase contrast angiography. *Magnetic Resonance in Medicine*, 9(1):139–149, 1989.

[44] J. A. Fessler and B. P. Sutton. Nonuniform fast Fourier transforms using min-max interpolation. *IEEE Transactions on Signal Processing*, 51(2):560–574, 2003.

[45] R. Freeman and H. D. W. Hill. Phase and intensity anomalies in Fourier transform NMR. *Journal of Magnetic Resonance*, 4:366–383, 1971.

[46] A. Funai, J. Fessler, D. Yeo, V. Olafsson, and D. Noll. Regularized field map estimation in MRI. *IEEE Transactions on Medical Imaging*, 27(10):1484–1494, 2008.

[47] P. J. Gareau, B. K. Rutt, S. J. Karlik, and J. R. Mitchell. Magnetization transfer and multicomponent T_2 relaxation measurements with histopathologic correlation in an experimental model of MS. *Journal of Magnetic Resonance Imaging*, 11(6):586–595, 2000.

[48] C. F. G. C. Geraldes and S. Laurent. Classification and basic properties of contrast agents for magnetic resonance imaging. *Contrast Media & Molecular Imaging*, 4(1):1–23, 2009.

[49] D. C. Ghiglia and M. D. Pritt. *Two-Dimensional Phase Unwrapping: Theory, Algorithms and Software*. Wiley, New York, NY, USA, 1998.

[50] A. Gifford, T. F. Towse, M. J. Avison, and E. B. Welch. Temperature mapping in human brown adipose tissue using fat-water MRI with explicit fitting of water peak location. In *Proc. 22nd Annual Meeting of ISMRM*, page 2354, 2014.

[51] M. Gloor, K. Scheffler, and O. Bieri. Quantitative magnetization transfer imaging using balanced SSFP. *Magnetic Resonance in Medicine*, 60(3):691–700, 2008.

[52] S. W. Golomb and G. Gong. *Signal Design for Good Correlation: for Wireless Communication, Cryptography, and Radar*. Cambridge University Press, Cambridge, UK, 2005.

[53] G. Golub and V. Pereyra. The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate. *SIAM Journal on Numerical Analysis*, 10(2):413–432, 1973.

[54] G. H. Golub, M. Heath, and G. Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21(2):215–223, 1979.

[55] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 3rd edition, 2006.

[56] S. J. Graham and M. J. Bronskill. MR measurement of relative water content and multicomponent T_2 relaxation in human breast. *Magnetic Resonance in Medicine*, 35(5):706–715, 1996.

[57] S. J. Graham, P. L. Stanchev, and M. J. Bronskill. Criteria for analysis of multicomponent tissue T_2 relaxation data. *Magnetic Resonance in Medicine*, 35(3):370–378, 1996.

[58] W. Grissom, K. Pauly, M. Lustig, V. Rieke, J. Pauly, and N. McDannold. Regularized referenceless temperature estimation in PRF-shift MR thermometry. In *Proc. IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI '09)*, pages 1235–1238, 2009.

[59] W. A. Grissom, A. B. Kerr, P. Stang, G. C. Scott, and J. M. Pauly. Minimum envelope roughness pulse design for reduced amplifier distortion in parallel excitation. *Magnetic Resonance in Medicine*, 64(5):1432–1439, 2010.

[60] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magnetic Resonance in Medicine*, 47(6):1202–1210, 2002.

[61] P. D. Groen and B. D. Moor. The fit of a sum of exponentials to noisy data. *Journal of Computational and Applied Mathematics*, 20(0):175–187, 1987.

[62] E. M. Haacke, R. W. Brown, M. R. Thompson, and R. Venkatesan. *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. Wiley-Liss, New York, NY, USA, 1999.

[63] R. Hashemi, W. Bradley, and C. Lisanti. *MRI: The Basics*. Lippincott Williams & Wilkins, Philadelphia, PA, USA, 3rd edition, 2010.

[64] T. Hastie, R. Tibshirani, J. Friedman, T. Hastie, J. Friedman, and R. Tibshirani. *The Elements of Statistical Learning*. Springer, New York, NY, USA, 2nd edition, 2009.

[65] H. He, J. Li, and P. Stoica. *Waveform Design for Active Sensing Systems: A Computational Approach*. Cambridge University Press, Cambridge, UK, 2012.

[66] K. He, J. Sun, and X. Tang. Guided image filtering. In *Computer Vision - ECCV 2010*, volume 6311 of *Lecture Notes in Computer Science*, pages 1–14. Springer Berlin Heidelberg, 2010.

[67] R. M. Henkelman. Measurement of signal intensities in the presence of noise in MR images. *Medical Physics*, 12(2):232–233, 1985.

[68] R. M. Henkelman, G. J. Stanisz, and S. J. Graham. Magnetization transfer in MRI: a review. *NMR in Biomedicine*, 14(2):57–64, 2001.

[69] A. Istratov and O. F. Vyvenko. Exponential analysis in physical phenomena. *Review of Scientific Instruments*, 70(2):1233–1257, 1999.

[70] K. Itoh. Analysis of the phase unwrapping algorithm. *Applied Optics*, 21(14):2470–2470, 1982.

[71] U. Katscher, P. Börnert, C. Leussler, and J. S. van den Brink. Transmit SENSE. *Magnetic Resonance in Medicine*, 49(1):144–150, 2003.

[72] S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*, volume 1. Prentice Hall, Upper Saddle River, NJ, USA, 1993.

[73] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky. An interior-point method for large-scale L1-regularized least squares. *IEEE Journal of Selected Topics in Signal Processing*, 1(4):606–617, 2007.

[74] F. Knoll, K. Bredies, T. Pock, and R. Stollberger. Second order total generalized variation (TGV) for MRI. *Magnetic Resonance in Medicine*, 65(2):480–491, 2011.

[75] T. Knopp, H. Eggers, H. Dahnke, J. Prestin, and J. Senegas. Iterative off-resonance and signal decay estimation and correction for multi-echo MRI. *IEEE Transactions on Medical Imaging*, 28(3):394–404, 2009.

[76] R. M. Kroeker and R. M. Henkelman. Analysis of biological NMR relaxation data with continuous distributions of relaxation times. *Journal of Magnetic Resonance*, 69(2):218 – 235, 1986.

[77] R. Kumaresan and D. Tufts. Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 30(6):833–840, 1982.

[78] C. L. Lankford and M. D. Does. On the inherent precision of mcDESPOT. *Magnetic Resonance in Medicine*, 69(1):127–136, 2013.

[79] M. L. Lauzon and R. Frayne. Analytical characterization of RF phase-cycled balanced steady-state free precession. *Concepts in Magnetic Resonance Part A*, 34A(3):133–143, 2009.

[80] C. L. Lawson and R. J. Hanson. *Solving Least Squares Problems*. Prentice Hall, Englewood Cliffs, NJ, USA, 1974.

[81] I. R. Levesque, P. S. Giacomini, S. Narayanan, L. T. Ribeiro, J. G. Sled, D. L. Arnold, and G. B. Pike. Quantitative magnetization transfer and myelin water imaging of the evolution of acute multiple sclerosis lesions. *Magnetic Resonance in Medicine*, 63(3):633–640, 2010.

[82] C. Li, X. Pan, K. Ying, Q. Zhang, J. An, D. Weng, W. Qin, and K. Li. An internal reference model-based PRF temperature mapping method with Cramér-Rao lower bound noise performance analysis. *Magnetic Resonance in Medicine*, 62(5):1251–1260, 2009.

[83] M. Lustig, D. Donoho, and J. M. Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, 58(6):1182–1195, 2007.

[84] A. Mackay, K. Whittall, J. Adler, D. Li, D. Paty, and D. Graeb. In vivo visualization of myelin water in brain by magnetic resonance. *Magnetic Resonance in Medicine*, 31(6):673–677, 1994.

[85] B. Mädler, S. A. Drabycz, S. H. Kolind, K. P. Whittall, and A. L. MacKay. Is diffusion anisotropy an accurate monitor of myelination? Correlation of multicomponent T_2 relaxation and diffusion tensor anisotropy in human brain. *Magnetic Resonance Imaging*, 26(7):874 – 888, 2008.

[86] K. V. Mardia and P. E. Jupp. *Directional Statistics*, volume 494. John Wiley & Sons, Chichester, UK, 2009.

[87] H.-L. Margaret Cheng, N. Stikov, N. R. Ghugre, and G. A. Wright. Practical medical applications of quantitative MR relaxometry. *Journal of Magnetic Resonance Imaging*, 36(4):805–824, 2012.

[88] C. R. McCreary, T. A. Bjarnason, V. Skihar, J. R. Mitchell, V. W. Yong, and J. F. Dunn. Multiexponential T_2 and magnetization transfer MRI of demyelination and remyelination in murine spinal cord. *Neuroimage*, 45(4):1173 – 1182, 2009.

[89] D. Nishimura. *Principles of Magnetic Resonance Imaging*. Stanford University, Stanford, CA, USA, 1996.

[90] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, NewYork, NY, USA, 1999.

[91] J. Oh, E. T. Han, D. Pelletier, and S. J. Nelson. Measurement of in vivo multi-component T_2 relaxation times for brain tissue using multi-slice T_2 prep at 1.5 and 3 T. *Magnetic Resonance Imaging*, 24(1):33–43, 2006.

[92] M. Ott, M. Blaimer, P. Ehses, P. M. Jakob, and F. Breuer. Phase sensitive PC-bSSFP: simultaneous quantification of T_1, T_2 and spin density M_0. In *Proc. 20th Annual Meeting of ISMRM*, page 2387, Melbourne, Australia, 2012.

[93] J. Pauly, P. Le Roux, D. Nishimura, and A. Macovski. Parameter relations for the Shinnar-Le Roux selective excitation pulse design algorithm. *IEEE Transactions on Medical Imaging*, 10(1):53–65, 1991.

[94] J. C. Pedro and S. A. Maas. A comparative overview of microwave and wireless power-amplifier behavioral modeling approaches. *IEEE Transactions on Microwave Theory and Techniques*, 53(4):1150–1163, 2005.

[95] B. Picinbono. Second-order complex random vectors and normal distributions. *IEEE Transactions on Signal Processing*, 44(10):2637–2640, 1996.

[96] C. S. Poon and R. M. Henkelman. Practical T_2 quantitation for clinical applications. *Journal of Magnetic Resonance Imaging*, 2(5):541–553, 1992.

[97] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger. SENSE: Sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine*, 42(5):952–962, 1999.

[98] S. Ramani and J. Fessler. Parallel MR image reconstruction using augmented Lagrangian methods. *IEEE Transactions on Medical Imaging*, 30(3):694–706, March 2011.

[99] S. B. Reeder, A. R. Pineda, Z. Wen, A. Shimakawa, H. Yu, J. H. Brittain, G. E. Gold, C. H. Beaulieu, and N. J. Pelc. Iterative decomposition of water and fat with echo asymmetry and least-squares estimation (IDEAL): Application with fast spin-echo imaging. *Magnetic Resonance in Medicine*, 54(3):636–644, 2005.

[100] D. A. Reiter, P.-C. Lin, K. W. Fishbein, and R. G. Spencer. Multicomponent T_2 relaxation analysis in cartilage. *Magnetic Resonance in Medicine*, 61(4):803–809, 2009.

[101] V. Rieke and K. Butts Pauly. MR thermometry. *Journal of Magnetic Resonance Imaging*, 27(2):376–390, 2008.

[102] G. Saab, R. T. Thompson, and G. D. Marsh. Multicomponent T_2 relaxation of in vivo skeletal muscle. *Magnetic Resonance in Medicine*, 42(1):150–157, 1999.

[103] L. I. Sacolick, F. Wiesinger, I. Hancu, and M. W. Vogel. B1 mapping by Bloch-Siegert shift. *Magnetic Resonance in Medicine*, 63(5):1315–1322, 2010.

[104] F. Santini and K. Scheffler. Reconstruction and frequency mapping with phase-cycled bSSFP. In *Proc. 18th Annual Meeting of ISMRM*, page 3089, Stockholm, Sweden, 2010.

[105] L. Scharf and C. Demeure. *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis.* Addison-Wesley series in electrical and computer engineering: Digital Signal Processing. Addison-Wesley Publishing Company, Reading, MA, USA, 1991.

[106] K. Scheffler and J. Hennig. T_1 quantification with inversion recovery TrueFISP. *Magnetic Resonance in Medicine*, 45(4):720–723, 2001.

[107] V. D. Schepkin, F. C. Bejarano, T. Morgan, S. Gower-Winter, M. Ozambela, and C. W. Levenson. In vivo magnetic resonance imaging of sodium and diffusion in rat glioma at 21.1 T. *Magnetic Resonance in Medicine*, 67(4):1159–1166, 2012.

[108] P. Schmitt, M. A. Griswold, P. M. Jakob, M. Kotas, V. Gulani, M. Flentje, and A. Haase. Inversion recovery TrueFISP: Quantification of T_1, T_2, and spin density. *Magnetic Resonance in Medicine*, 51(4):661–667, 2004.

[109] G. Seber and C. Wild. *Nonlinear Regression.* Wiley Series in Probability and Statistics. Wiley, Hoboken, NJ, USA, 2003.

[110] J. Sijbers, A. den Dekker, P. Scheunders, and D. Van Dyck. Maximum-likelihood estimation of Rician distribution parameters. *IEEE Transactions on Medical Imaging*, 17(3):357–361, 1998.

[111] J. Sijbers, A. J. den Dekker, E. Raman, and D. Van Dyck. Parameter estimation from magnitude MR images. *International Journal of Imaging Systems and Technology*, 10(2):109–114, 1999.

[112] J. Sijbers, A. J. den Dekker, M. Verhoye, E. Raman, and D. Van Dyck. Optimal estimation of T_2 maps from magnitude MR images. *Proc. of SPIE Medical Imaging*, 3338:384–390, 1998.

[113] J. G. Sled and G. B. Pike. Correction for B1 and B0 variations in quantitative T_2 measurements using MRI. *Magnetic Resonance in Medicine*, 43(4):589–593, 2000.

[114] T. Söderström and P. Stoica. *System Identification.* Prentice Hall International, Cambridge, UK, 1989.

[115] B. J. Soher, C. Wyatt, S. B. Reeder, and J. R. MacFall. Noninvasive temperature mapping with MRI using chemical shift water-fat separation. *Magnetic Resonance in Medicine*, 63(5):1238–1246, 2010.

[116] W. M. Spees, N. Buhl, P. Sun, J. J. Ackerman, J. J. Neil, and J. R. Garbow. Quantification and compensation of eddy-current-induced magnetic-field gradients. *Journal of Magnetic Resonance*, 212(1):116 – 123, 2011.

[117] S. M. Sprinkhuizen, C. J. G. Bakker, and L. W. Bartels. Absolute MR thermometry using time-domain analysis of multi-gradient-echo magnitude images. *Magnetic Resonance in Medicine*, 64(1):239–248, 2010.

[118] K. Steiglitz and L. McBride. A technique for the identification of linear systems. *IEEE Transactions on Automatic Control*, 10(4):461–464, 1965.

[119] V. Stevenson, G. Parker, G. Barker, K. Birnie, P. Tofts, D. Miller, and A. Thompson. Variations in T1 and T2 relaxation times of normal appearing white matter and lesions in multiple sclerosis. *Journal of the neurological sciences*, 178(2):81–87, 2000.

[120] P. Stoica and P. Babu. Parameter estimation of exponential signals: A system identification approach. *Digital Signal Processing*, 23(5):1565 – 1577, 2013.

[121] P. Stoica, J. Li, and T. Söderström. On the inconsistency of IQML. *Signal Processing*, 56(2):185–190, 1997.

[122] P. Stoica and R. Moses. *Spectral Analysis of Signals.* Prentice Hall, Upper Saddle River, NJ, USA, 2005.

[123] P. Stoica and Y. Selen. Model-order selection: a review of information criterion rules. *IEEE Signal Processing Magazine*, 21(4):36–47, 2004.

[124] P. Stoica and T. Söderström. The Steiglitz-McBride identification algorithm revisited—convergence analysis and accuracy aspects. *IEEE Transactions on Automatic Control*, 26(3):712–717, 1981.

[125] B. A. Taylor, K.-P. Hwang, A. M. Elliott, A. Shetty, J. D. Hazle, and R. J. Stafford. Dynamic chemical shift imaging for image-guided thermal therapy: Analysis of feasibility and potential. *Medical Physics*, 35(2):793–803, 2008.

[126] P. Tofts. *Quantitative MRI of the Brain: Measuring Changes Caused by Disease.* Wiley, Chichester, UK, 2003.

[127] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Proc. 6th International Conference on Computer Vision*, pages 839–846, Jan 1998.

[128] S. Tretter. Estimating the frequency of a noisy sinusoid by linear regression. *IEEE Transactions on Information Theory*, 31(6):832–835, 1985.

[129] H. L. Van Trees. *Detection, Estimation, and Modulation Theory, Optimum Array Processing.* John Wiley & Sons, New York, NY, USA, 2004.

[130] S. Walker-Samuel, M. Orton, L. D. McPhail, J. K. R. Boult, G. Box, S. A. Eccles, and S. P. Robinson. Bayesian estimation of changes in transverse relaxation rates. *Magnetic Resonance in Medicine*, 64(3):914–921, 2010.

[131] S. Webb, C. A. Munro, R. Midha, and G. J. Stanisz. Is multicomponent T_2 a good measure of myelin content in peripheral nerve? *Magnetic Resonance in Medicine*, 49(4):638–645, 2003.

[132] D. Weishaupt, V. Köchli, and B. Marincek. *How does MRI work? An Introduction to the Physics and Function of Magnetic Resonance Imaging.* Springer, Heidelberg, Germany, 2nd edition, 2006.

[133] K. P. Whittall and A. L. MacKay. Quantitative interpretation of NMR relaxation data. *Journal of Magnetic Resonance*, 84(1):134 – 152, 1989.

[134] K. P. Whittall, A. L. Mackay, D. A. Graeb, R. A. Nugent, D. K. B. Li, and D. W. Paty. In vivo measurement of T_2 distributions and water

contents in normal human brain. *Magnetic Resonance in Medicine*, 37(1):34–43, 1997.

[135] K. P. Whittall, A. L. MacKay, and D. K. Li. Are mono-exponential fits to a few echoes sufficient to determine T_2 relaxation for in vivo human brain? *Magnetic Resonance in Medicine*, 41(6):1255–1257, 1999.

[136] Q.-S. Xiang and M. N. Hoff. Simple cross-solution for banding artifact removal in bSSFP imaging. In *Proc. 18th Annual Meeting of ISMRM*, page 74, Stockholm, Sweden, 2010.

[137] V. L. Yarnykh. Actual flip-angle imaging in the pulsed steady state: A method for rapid three-dimensional mapping of the transmitted radiofrequency field. *Magnetic Resonance in Medicine*, 57(1):192–200, 2007.

[138] C. Y. Yip, J. A. Fessler, and D. C. Noll. Iterative RF pulse design for multidimensional, small-tip-angle selective excitation. *Magnetic Resonance in Medicine*, 54(4):908–917, 2005.

[139] M. Zeineh, M. Parekh, J. Su, and B. Rutt. Ultra-high resolution 0.4 mm isotropic structural imaging of the human hippocampus in vivo utilizing phase-cycled bSSFP at 7.0 t. In *Proc. American Society for Neuroradiology 51st Annual Meeting*, page 31, San Diego, CA, USA, 2013. American Society for Neuroradiology.

[140] J. Zhang, S. H. Kolind, C. Laule, and A. L. MacKay. Comparison of myelin water fraction from multiecho T_2 decay curve and steady-state methods. *Magnetic Resonance in Medicine*, 73(1):223–232, 2015.

[141] S. Zheng and Y. Xia. Multi-components of T_2 relaxation in ex vivo cartilage and tendon. *Journal of Magnetic Resonance*, 198(2):188 – 196, 2009.

[142] Y. Zhu. Parallel excitation with an array of transmit coils. *Magnetic Resonance in Medicine*, 51(4):775–784, 2004.

[143] Y. Zur, S. Stokar, and P. Bendel. An analysis of fast imaging sequences with steady-state transverse magnetization refocusing. *Magnetic Resonance in Medicine*, 6(2):175–193, 1988.

Sammanfattning på svenska

Nedan följer en kort sammanfattning av avhandlingen, som på svenska får titeln *"Bidrag till signalbehandling för magnetresonanstomografi"*.

Magnetresonanstomografi (MRT) kan användas för att avbilda mjuka vävnader utan joniserande strålning, och är ett viktigt verktyg för medicinsk diagnos. Utöver anatomi kan man också fånga meta/-bol/-ism och diffusion av olika molekyler, samt mäta temperatur i tre dimensioner och indikera aktivering av hjärnan.

Signalbehandling är en centralt många av de steg som behövs för att skapa en MR bild, och används i allt från design av excitation och kodning av spatiell information, till att rekonstruera en bild från uppmätta data och utföra bildbehandling. Dessutom kan man genom avancerad signalbehandling förbättra bilderna ytterligare, ta bort artifakter, och till och med ta fram information bortom bilderna i sig. Detta är det huvudsakliga ämnet i denna avhandling. Inom kvantitativ MRT är målet vanligen att skatta fysikaliska parametrar baserat på ett flertal insamlade bilder. De resulterande optimeringsproblemen är ofta olinjära vilket kräver smarta och specialdesignade algoritmer för att undvika suboptimala lösningar och lokala minimum. I denna avhandling presenteras flera algoritmer av detta slag som löser olika parameterskattningsproblem, antingen för att skatta fysikaliska storheter, eller för att minimera artifakter och hantera brus i bilderna. Med hjälp av dessa skattningar kan man karaktärisera vävnad på ett bättre sätt, vilket leder till bättre förutsättningar för diagnos. Dessutom behandlas ett designproblem, med målet att finna excitationssekvenser som minimerar artifakter när man använder förstärkare med begränsad kvalité. Detta leder i sin tur till bättre bilder för diagnos och kan även minska skattningsfelet vid kvantitativ MRT.

Först ges en introduktion till fysiken bakom MRT och hur en bild skapas samt en genom/-gång av några användbara begrepp inom signalbehandling. Resterande kapitel behandlar mer specifika signalbe-

handlingsproblem inom MRT, som att ta bort bandartifakter, skatta tidskonstanter för exponentiellt avtagande, kompensera för fasfel, designa excitation och mäta temperatur. En sammanfattning av respektive problem följer nedan.

Signalförsluts i form av bandartifakter kan skymma detaljer i MR bilder och försvåra diagnos. Dessa artifakter är ett vanligt problem för den annars effektiva bSSFP-sekvensen. I kapitel 4 presenteras en snabb tvåstegsalgoritm för att 1) estimera de okända parametrarna i bSSFP-signalmodellen baserat på flera insamlade bilder med olika fasinkrement, och 2) rekonstruera ett bandfritt resultat. Det första steget, kallat LORE (Linearization of off-resonance estimation), löser det olinjära problemet approximativt genom en robust linjär metod. I det andra steget används Gauss-Newton, med LORE-estimaten som initialisering, för att minimera det ursprungliga kriteriet. Hela algoritmen kallas LORE-GN. Genom att härleda Cramér-Rao-gränsen (CRB) visas att LORE-GN är statistiskt effektiv; och vidare att det är teoretiskt svårt att simultant skatta T_1 och T_2 från bSSFP-data, genom att CRB är högt vid normala signal-brus-förhållanden. Med hjälp av simulerat, fantom, och in-vivo data, så påvisas LORE-GNs bandreduktionsegenskaper jämfört med andra vanliga metoder, så som "summan av kvadraterna". LORE-GN minimerar effektivt bandartifakter i bSSFP där andra metoder misslyckas, till exempel vid hög fältstyrka.

Modeller baserade på en summa av dämpade exponentialer förekommer i många tillämpningar, och specifikt vid skattning av multipla T_2-komponenter. Problemet med att uppskatta relaxationsparametrar och motsvarande amplituder är känt för att vara svårt, speciellt när antalet komponenter ökar. I kapitel 5 jämförs en nyligen presenterad parameterskattningsalgoritm, kallad EASI-SM, med den positiva minstakvadratmetoden för att skatta T_2-spektrum (NNLS), som vanligen används inom MR-fältet. Prestandan hos de två algoritmerna utvärderas via simulering med hjälp av Cramér-Rao-gränsen. Dessutom appliceras algoritmerna på in-vivo hjärndata med 32 spinneko-bilder, för att uppskatta myelin-vattenfraktionen och den mest signifikanta T_2-komponenten. I simulering ger EASI-SM bättre parameterskattningsprestanda jämfört med NNLS, och resulterar dessutom i en lägre varians av T_2-estimaten in vivo. EASI-SM är ett effektivt och parameterfritt alternativ till NNLS, och ger ett nytt sätt att uppskatta variationer av myelin i hjärnan.

Att skatta den transversella relaxationstiden, T_2, från magnituden av flera spinn-ekobilder är ett vanligt problem inom kvantitativ MRT. Standardmetoden är att använda pixelvisa estimat, dock kan dessa vara ganska brusiga då endast två bilder finns att tillgå. Genom att använda information från närliggande pixlar är det möjligt att minska variansen i skattningarna, men detta försämrar vanligtvis detaljrikedomen i bilden,

särskilt vid vävnadsgränser. I kapitel 6, presenteras två snabba metoder för att minska variansen av T_2-estimaten utan att påverka kontrasten mellan vävnader. Metoderna använder en suboptimal formulering av problemet då en optimal lösning skulle vara för tidskrävande i praktiken. Den första metoden är en enkel lokal minstakvadratbaserad lösning, medan den andra bygger på global optimering över hela bilden, med begränsningar på totala variationen i de resulterande estimaten. Båda dessa metoder utvärderas med hjälp av simulerat- och in-vivo-data. Resultaten visar att variansen av estimaten för de föreslagna T_2-skattningsmetoderna är mindre än för standardmetoden, och att kontrasten bevaras.

Bruset i de magnitudbilder som vanligen används vid skattning av den transversella relaxationstiden är Rice-fördelat, vilket kan leda till ett betydande systematiskt fel när minstakvadratmetorder används. Ett sätt att undvika dessa problem är att skatta ett reellt och Gaussiskt fördelat dataset från komplexa data, istället än att använda magnituden. I kapitel 7 presenteras två algoritmer för faskorrigering som kan användas för att generera reella data som lämpar sig för LS-baserade parameterskattningsmetoder. Den första kallas WELPE, och beräknar ett fasestimat genom en viktad linjär metod. WELPE ger en förbättring jämfört med en tidigare publicerad algoritm, samtidigt som den förenklar skattningsförfarandet och kan hantera data från flera mottagarspolar. I varje bildelement anpassar algoritmen en linjärt parametriserad funktion till fasen av flera spinn-ekobilder, och utifrån den skattade fasen projiceras datat på den reella axeln. Den andra algoritmen är en maximum likelihood-metod som estimerar den sanna avklingande signalmagnituden, och som kan implementeras effektivt när fasvariationen är linjär i tiden. Prestandan hos algoritmerna demonstreras via Monte Carlo-simuleringar, genom att jämföra noggrannheten hos de uppskattade tidskonstanterna. Vidare visas att de föreslagna algoritmerna möjliggör mer exakta T_2-skattningar, genom att de minskar det systematiska felet vid skattning av flera T_2-komponenter jämfört mot när man använder magnituddata. WELPE tillämpas också på ett dataset med 32 ekon av en in-vivo-hjärna för att visa att algoritmen också fungerar i ett praktiskt scenario.

Vid användning av lågkostnadsförstärkare med begränsad kvalitet kan signaldistorsion kan vara ett problem vid excitationen, vilket i sin tur kan resultera i bildartefakter. I kapitel 8 presenteras en algoritm för att designa excitation då man har begränsningar på signalens magnitud. Sådana sekvenser kan till exempel användas för att uppnå önskat excitationsmönster i frekvensdomän samtidigt som man utnyttjar flera spolar med budgetförstärkare. Det resulterande icke-konvexa optimeringskriteriet minimeras lokalt med hjälp av en cyklisk algoritm som består av två enkla algebraiska delsteg. Eftersom den föreslagna algoritmen min-

imerar det faktiska kriteriet så är de erhållna sekvenserna garanterat bättre än de skattningar som erhålls från en tidigare publicerad algoritm, som bygger på den heuristiska principen av iterativ kvadratisk maximum likelihood. Prestandan för den föreslagna algoritmen illustreras i två numeriska exempel, där den också jämförs med den tidigare föreslagna metoden.

Acta Universitatis Upsaliensis
Uppsala Dissertations from the Faculty of Science
Editor: The Dean of the Faculty of Science

1–11: 1970–1975

12. *Lars Thofelt*: Studies on leaf temperature recorded by direct measurement and by thermography. 1975.
13. *Monica Henricsson*: Nutritional studies on Chara globularis Thuill., Chara zeylanica Willd., and Chara haitensis Turpin. 1976.
14. *Göran Kloow*: Studies on Regenerated Cellulose by the Fluorescence Depolarization Technique. 1976.
15. *Carl-Magnus Backman*: A High Pressure Study of the Photolytic Decomposition of Azoethane and Propionyl Peroxide. 1976.
16. *Lennart Källströmer*: The significance of biotin and certain monosaccharides for the growth of Aspergillus niger on rhamnose medium at elevated temperature. 1977.
17. *Staffan Renlund*: Identification of Oxytocin and Vasopressin in the Bovine Adenohypophysis. 1978.
18. *Bengt Finnström*: Effects of pH, Ionic Strength and Light Intensity on the Flash Photolysis of L-tryptophan. 1978.
19. *Thomas C. Amu*: Diffusion in Dilute Solutions: An Experimental Study with Special Reference to the Effect of Size and Shape of Solute and Solvent Molecules. 1978.
20. *Lars Tegnér*: A Flash Photolysis Study of the Thermal Cis-Trans Isomerization of Some Aromatic Schiff Bases in Solution. 1979.
21. *Stig Tormod*: A High-Speed Stopped Flow Laser Light Scattering Apparatus and its Application in a Study of Conformational Changes in Bovine Serum Albumin. 1985.
22. *Björn Varnestig*: Coulomb Excitation of Rotational Nuclei. 1987.
23. *Frans Lettenström*: A study of nuclear effects in deep inelastic muon scattering. 1988.
24. *Göran Ericsson*: Production of Heavy Hypernuclei in Antiproton Annihilation. Study of their decay in the fission channel. 1988.
25. *Fang Peng*: The Geopotential: Modelling Techniques and Physical Implications with Case Studies in the South and East China Sea and Fennoscandia. 1989.
26. *Md. Anowar Hossain*: Seismic Refraction Studies in the Baltic Shield along the Fennolora Profile. 1989.
27. *Lars Erik Svensson*: Coulomb Excitation of Vibrational Nuclei. 1989.
28. *Bengt Carlsson*: Digital differentiating filters and model based fault detection. 1989.
29. *Alexander Edgar Kavka*: Coulomb Excitation. Analytical Methods and Experimental Results on even Selenium Nuclei. 1989.
30. *Christopher Juhlin*: Seismic Attenuation, Shear Wave Anisotropy and Some Aspects of Fracturing in the Crystalline Rock of the Siljan Ring Area, Central Sweden. 1990.

31. *Torbjörn Wigren*: Recursive Identification Based on the Nonlinear Wiener Model. 1990.
32. *Kjell Janson*: Experimental investigations of the proton and deuteron structure functions. 1991.
33. *Suzanne W. Harris*: Positive Muons in Crystalline and Amorphous Solids. 1991.
34. *Jan Blomgren*: Experimental Studies of Giant Resonances in Medium-Weight Spherical Nuclei. 1991.
35. *Jonas Lindgren*: Waveform Inversion of Seismic Reflection Data through Local Optimisation Methods. 1992.
36. *Liqi Fang*: Dynamic Light Scattering from Polymer Gels and Semidilute Solutions. 1992.
37. *Raymond Munier*: Segmentation, Fragmentation and Jostling of the Baltic Shield with Time. 1993.

Prior to January 1994, the series was called *Uppsala Dissertations from the Faculty of Science*.

Acta Universitatis Upsaliensis
Uppsala Dissertations from the Faculty of Science and Technology
Editor: The Dean of the Faculty of Science

1–14: 1994–1997. 15–21: 1998–1999. 22–35: 2000–2001. 36–51: 2002–2003.
52. *Erik Larsson*: Identification of Stochastic Continuous-time Systems. Algorithms, Irregular Sampling and Cramér-Rao Bounds. 2004.
53. *Per Åhgren*: On System Identification and Acoustic Echo Cancellation. 2004.
54. *Felix Wehrmann*: On Modelling Nonlinear Variation in Discrete Appearances of Objects. 2004.
55. *Peter S. Hammerstein*: Stochastic Resonance and Noise-Assisted Signal Transfer. On Coupling-Effects of Stochastic Resonators and Spectral Optimization of Fluctuations in Random Network Switches. 2004.
56. *Esteban Damián Avendaño Soto*: Electrochromism in Nickel-based Oxides. Coloration Mechanisms and Optimization of Sputter-deposited Thin Films. 2004.
57. *Jenny Öhman Persson*: The Obvious & The Essential. Interpreting Software Development & Organizational Change. 2004.
58. *Chariklia Rouki*: Experimental Studies of the Synthesis and the Survival Probability of Transactinides. 2004.
59. *Emad Abd-Elrady*: Nonlinear Approaches to Periodic Signal Modeling. 2005.
60. *Marcus Nilsson*: Regular Model Checking. 2005.
61. *Pritha Mahata*: Model Checking Parameterized Timed Systems. 2005.
62. *Anders Berglund*: Learning computer systems in a distributed project course: The what, why, how and where. 2005.
63. *Barbara Piechocinska*: Physics from Wholeness. Dynamical Totality as a Conceptual Foundation for Physical Theories. 2005.
64. *Pär Samuelsson*: Control of Nitrogen Removal in Activated Sludge Processes. 2005.

65. *Mats Ekman*: Modeling and Control of Bilinear Systems. Application to the Activated Sludge Process. 2005.
66. *Milena Ivanova*: Scalable Scientific Stream Query Processing. 2005.
67. *Zoran Radović´*: Software Techniques for Distributed Shared Memory. 2005.
68. *Richard Abrahamsson*: Estimation Problems in Array Signal Processing, System Identification, and Radar Imagery. 2006.
69. *Fredrik Robelius*: Giant Oil Fields – The Highway to Oil. Giant Oil Fields and their Importance for Future Oil Production. 2007.
70. *Anna Davour*: Search for low mass WIMPs with the AMANDA neutrino telescope. 2007.
71. *Magnus Ågren*: Set Constraints for Local Search. 2007.
72. *Ahmed Rezine*: Parameterized Systems: Generalizing and Simplifying Automatic Verification. 2008.
73. *Linda Brus*: Nonlinear Identification and Control with Solar Energy Applications. 2008.
74. *Peter Nauclér*: Estimation and Control of Resonant Systems with Stochastic Disturbances. 2008.
75. *Johan Petrini*: Querying RDF Schema Views of Relational Databases. 2008.
76. *Noomene Ben Henda*: Infinite-state Stochastic and Parameterized Systems. 2008.
77. *Samson Keleta*: Double Pion Production in dd→απα Reaction. 2008.
78. *Mei Hong*: Analysis of Some Methods for Identifying Dynamic Errors-invariables Systems. 2008.
79. *Robin Strand*: Distance Functions and Image Processing on Point-Lattices With Focus on the 3D Face-and Body-centered Cubic Grids. 2008.
80. *Ruslan Fomkin*: Optimization and Execution of Complex Scientific Queries. 2009.
81. *John Airey*: Science, Language and Literacy. Case Studies of Learning in Swedish University Physics. 2009.
82. *Arvid Pohl*: Search for Subrelativistic Particles with the AMANDA Neutrino Telescope. 2009.
83. *Anna Danielsson*: Doing Physics – Doing Gender. An Exploration of Physics Students' Identity Constitution in the Context of Laboratory Work. 2009.
84. *Karin Schönning*: Meson Production in pd Collisions. 2009.
85. *Henrik Petrén*: η Meson Production in Proton-Proton Collisions at Excess Energies of 40 and 72 MeV. 2009.
86. *Jan Henry Nyström*: Analysing Fault Tolerance for ERLANG Applications. 2009.
87. *John Håkansson*: Design and Verification of Component Based Real-Time Systems. 2009.
88. *Sophie Grape*: Studies of PWO Crystals and Simulations of the $\bar{p}p \to \bar{\Lambda}\Lambda, \bar{\Lambda}\Sigma^0$ Reactions for the PANDA Experiment. 2009.
90. *Agnes Rensfelt*. Viscoelastic Materials. Identification and Experiment Design. 2010.
91. *Erik Gudmundson*. Signal Processing for Spectroscopic Applications. 2010.
92. *Björn Halvarsson*. Interaction Analysis in Multivariable Control Systems. Applications to Bioreactors for Nitrogen Removal. 2010.
93. *Jesper Bengtson*. Formalising process calculi. 2010.
94. *Magnus Johansson*. Psi-calculi: a Framework for Mobile Process Calculi. Cook your own correct process calculus – just add data and logic. 2010.
95. *Karin Rathsman*. Modeling of Electron Cooling. Theory, Data and Applications. 2010.

96. *Liselott Dominicus van den Bussche*. Getting the Picture of University Physics. 2010.