Advances in magnetic resonance imaging using statistical signal processing

Kelvin Layton

Submitted in total fulfilment of the requirements of the degree of
Doctor of Philosophy

Department of Electrical & Electronic Engineering
THE UNIVERSITY OF MELBOURNE

Produced on archival quality paper

May, 2013

Abstract

OVER the last 30 years, magnetic resonance imaging (MRI) has revolutionised diagnostic radiology by producing anatomical images of remarkable quality. MRI is often considered the most flexible imaging technique compared to other imaging modalities and as such, the technology has helped answer fundamental questions about the structure and function of the body. Despite these advances, the core technology was developed at time when computer performance was limited, necessitating signal approximations and clever acquisition strategies. The dramatic increase in computer power available today means the full flexibility of MRI can be explored.

This thesis adopts a statistical signal processing framework. From this perspective, accurate models of the underlying signal and noise processes are crucial to extract the maximum information available in the measurements. This framework is applied to the advancement of two emerging MRI technologies: quantitative MRI and nonlinear spatial encoding.

Quantitative MRI aims to estimate the underlying parameters contributing to a magnetic resonance signal. Unlike traditional imaging, based on contrast alone, the estimation of physical parameters promises to enhance tissue classification, disease detection and pathology. The present work examines the estimation of transverse relaxation rates for two cases. Firstly, estimation in the presence of distortion due to finite sampling bandwidth is considered. Secondly, estimation of distributions of relaxation rates are considered to model voxels with multiple components. Bayesian techniques are developed that incorporate accurate signal models and result in state-of-the-art performance.

The recent advent of nonlinear encoding fields is testament to the flexibility inherent in MRI. These magnetic encoding fields vary nonlinearly over the field-of-view resulting in an image with spatially varying resolution. A new acquisition strategy that exploits this property is developed to produce images with improved resolution in a user-specified region of interest. This technique has many applications; for example, clinicians could acquire an image with high resolution detail of a brain tumour, beyond that achievable with traditional techniques. The use of nonlinear fields creates an additional spatial dependence on the image signal-to-noise ratio and a computationally efficient metric is derived to quantity this effect. Such performance metrics are required to design new acquisition schemes that take advantage of all possible degrees of freedom.

Declaration

This is to certify that

1. the thesis comprises only my original work towards the PhD,
2. due acknowledgement has been made in the text to all other material used,
3. the thesis is fewer than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices.

_____

Kelvin Layton

_____

Date

Publications

The work presented in this thesis has produced the following publications and conference presentations.

Journal papers

- Layton, K. J., Morelande, M., Wright, D., Farrell, P. M., Moran, B. and Johnson, L. A. "Modelling and estimation of multi-component T_2 distributions," *IEEE Transactions on Medical Imaging*, 2012, (*In press*)

- Layton, K. J., Gallichan, D., Testud, F., Cocosco, C. A., Welz, A. M., Barmet, C., Pruessmann, K. P., Hennig, J. and Zaitsev, M. "Single shot trajectory design for region-specific imaging using linear and nonlinear magnetic encoding fields," *Magnetic Resonance in Medicine*, 2012, (*Early view*)

- Layton, K. J., Morelande, M., Farrell, P. M., Moran, B. and Johnson, L. A. "Performance analysis for magnetic resonance imaging with nonlinear encoding fields," *IEEE Transactions on Medical Imaging*, 2012, 31(2), p. 391–404

Refereed conference papers

- Layton, K. J., Morelande, M., Johnson, L. A., Farrell, P. M. and Moran, B. "Improved quantification of MRI relaxation rates using Bayesian estimation," *Proceedings of the IEEE Conference on Acoustics, Speech, and Signal Processing*, 2010, p. 481–484

- Layton, K. J., Johnson, L. A., Farrell, P. M., Moran, B. and Morelande, M. "Estimation of relaxation time distributions in magnetic resonance imaging," *Proceedings of the IEEE Conference on Acoustics, Speech, and Signal Processing*, 2012, p. 697–700

- Layton, K. J., Morelande, M., Farrell, P. M., Moran, B. and Johnson, L. A. "Adapting magnetic resonance imaging performance using nonlinear encoding fields," *Proceedings of the 33rd IEEE Engineering in Medicine and Biology Conference (EMBC)*, 2011, p. 3740–3743

Conference presentations

- Layton, K. J., Morelande, M., Farrell, P. M., Moran, B. and Johnson, L. A. "An improved algorithm for the estimation of multi-component T_2 distributions," *Proceedings of the ISMRM 20th Scientific Meeting and Exhibition*, 2012
- Layton, K. J., Gallichan, D., Testud, F., Cocosco, C. A., Welz, A. M., Barmet, C., Pruessmann, K. P., Hennig, J. and Zaitsev, M. "Region-specific trajectory design for single-shot imaging using linear and nonlinear magnetic encoding fields," *Proceedings of the ISMRM 20th Scientific Meeting and Exhibition*, 2012
- Layton, K. J., Morelande, M., Farrell, P. M., Moran, B. and Johnson, L. A. "A performance measure for MRI with nonlinear encoding fields," *Proceedings of the ISMRM 19th Annual Meeting*, 2011
- Layton, K. J., Morelande, M., Farrell, P. M., Moran, B. and Johnson, L. A. "Magnetic Resonance Imaging Using Nonlinear Encoding Fields," *The 4th Australian Workshop on Mathematical and Computational Neuroscience*, 2010
- Layton, K. J., Morelande, M., Farrell, P. M., Moran, B. and Johnson, L. A. "Quantum Mechanical Model for Magnetic Resonance Imaging," *The 3rd Australian Workshop on Mathematical and Computational Neuroscience*, 2008

Acknowledgements

I would like to express my gratitude to the people who made this thesis possible.

Firstly, I owe many thanks to my supervisors: Mark Morelande, Bill Moran, Peter Farrell and Leigh Johnson for their endless patience, support and knowledge. Meetings with four enthusiastic supervisors made for many lively and enjoyable discussions.

I am very grateful for my friends and colleagues at the Department of Electrical and Electronic Engineering, the Florey Neuroscience Institutes and the Centre for Neural Engineering at The University of Melbourne: Michelle Chong, Tom Close, Michael Eager, Dean Freestone, Matthieu Gilson, Colin Hales, Amir Jafarian, Tania Kameneva, Rob Kerr, Sei Zhen Khong, Isabell Kiral-Kornek, Stephan Lau, Amanda Ng, Emily O'Brien, Elma O'Sullivan-Greene, Andre Peterson, Martin Spencer, Craig Savage, Evgeni Sergeev, Kyle Slater, Bahman Tahayori, Andrea Varsavsky and David Wright.

Thank you to NICTA for financial support and in particular Natasha Baxter, Domenic Santilli-Centofanti and Tracy Painter for help with conference travel and administration matters.

I was fortunate to spend two months with the MRI group in Freiburg, where I learnt a great deal. Thanks to Jürgen Hennig, Sebastian Littin, Hans Weber, Chris Cocosco, Cris Lovell-Smith and Benjamin Zahneisen. I'd especially like to thank Maxim Zaitsev for supervising me during my visit; and Frederik Testud and Daniel Gallichan for considerable and continued help with the nonlinear encoding work. Thank you to Julian Maclaren and Anna Welz for the loan of their bikes, which provided me many wonderful memories.

I am also greatly indebted to my friends. Members of Curtain Haus: Mitchell Lawrence, Bede Moore and Jill Pope for creating a wonderful retreat, particularly after days of little progress; Mark Ryan for his confidence and encouragement; Adam Baker for his grammatical and flair; the 'Fantastic Four'; and members of my social sporting teams, The Outsiders, Vege Bacon Burgers and UN United for providing a welcome distraction and much needed exercise.

I thank my loving parents, Rob and Karen, for the sacrifices they made to give me a sound educational start and for instilling a self-belief, without which, this thesis would never have begun. I am thankful for my sisters, Janey, Angie and Tiarni for their encouragement and perspective.

Finally, I am eternally grateful to my girlfriend, Nikki, whose unwavering love and support guided me through this often challenging time.

Contents

List of Figures

CHAPTER 1

Introduction

1.1 Introduction to MRI

MAGNETIC resonance imaging (MRI) has become an indispensable tool for both clinical use and fundamental research. The utility of MRI is due to a combination of factors. Firstly, MRI provides good contrast between different soft tissues, which makes it an ideal modality for imaging the human body. Secondly, it is very safe due to the use of non-ionizing radio frequency fields. Other major imaging modalities such as nuclear medicine and X-ray use ionizing radiation, which can be harmful if exposure is not limited. Another advantage of MRI lies in its flexibility. Techniques such as structural imaging, quantitative MRI, functional MRI, diffusion weighted imaging and MR angiography all provide complementary information about the structure and function of the body.

The major drawback of MRI is cost. Machines cost millions of dollars to purchase and maintain. Figure 1.1 illustrates the costly components of an MRI scanner. A large portion of the cost is due the superconducting magnetic that generates the main magnetic field. Adding other components such gradient fields, shim coils, and radio frequency systems, the cost increases further.

A further limitation of MRI is acquisition time. Depending on the application, a

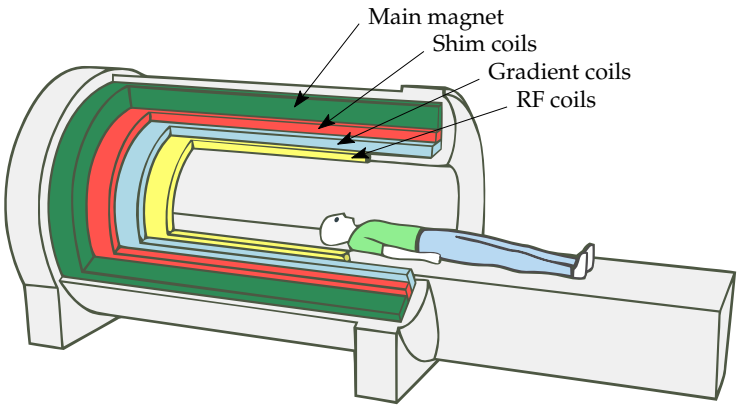


Figure 1.1: Schematic of an MRI scanner illustrating the main components. Image adapted from [1]

scan can last over an hour. This time is limited both by the hardware and physical constraints. The patient must lie still in a relatively confined space during the acquisition period, which can be unpleasant. Additionally, patient throughput is a priority for hospitals trying to maximise the benefit of a scanner. For these reasons, decreasing the acquisition time is a topic of ongoing research.

1.2 Motivation

The fundamentals of MRI were developed during a time when computational performance was often a limiting factor. Since the 1990s, a dramatic increase in computing power has been achieved simultaneously with a large reduction in computer cost. This new landscape paves the way for novel acquisition strategies and advanced processing algorithms in the field of MRI.

An MRI machine consists of many subsystems that interact in complex ways to produce the final output. Dependent on the application, several of these interactions are approximated to simplify the system description and the required computation. Whilst these approximations are often adequate to produce high quality results, there is still room for improvement. This thesis describes new MRI strategies with detailed physical models and advanced processing algorithms necessary to extract the desired information.

The increased computation, requisite for improved modelling, potentially permits two improvements. Firstly, the performance of current systems could be improved with a negligible increase in cost. Secondly, and perhaps more promisingly, the same level of performance might be achieved on cheaper systems. These systems could be cheaper for a number of reasons, including a reduced magnetic field strength, reduced field homogeneity and/or nonlinear gradient fields. Recent applications of these types of systems include portable MRI or ultra-low field MRI [2].

1.3 Overview of thesis

This thesis adopts a statistical signal processing framework. Through accurate models of the signal and noise processes, this framework provides tools for analysis and allows the maximum amount of information to be extracted from the measurements. The framework is applied to the emerging MRI technologies of quantitative MRI and nonlinear spatial encoding. The contributions to these research areas touch on a range of subcomponents of the MRI system and applications. Figure 1.2 displays a basic schematic for MRI highlighting the contributions

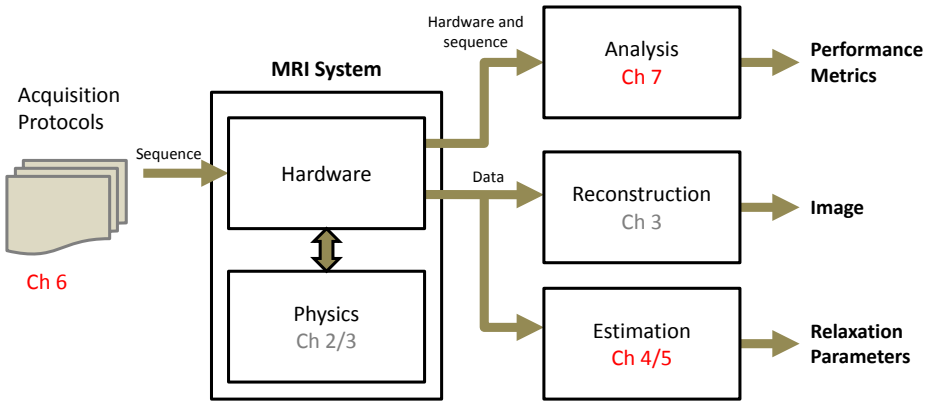


Figure 1.2: Overview of MRI processes and the relevant chapters discussing each component. Chapters coloured red indicate novel contributions.

made. The thesis is divided into three parts: background, quantitative MRI and nonlinear spatial encoding. Part I contains two chapters that establish the necessary background for the remainder of the thesis. Chapter 2 reviews the physics underpinning magnetic resonance from both a classical and quantum perspective. While not a strictly novel contribution, the chapter collates numerous sources in a manner not traditionally accessible in MRI literature. Boltzmann statistics are discussed as the mechanism to generate macroscopic magnetisation. The Bloch equation and the Schrödinger equation are presented and are used to derive the spin evolution during precession, excitation, and relaxation. The classical and quantum descriptions are presented alongside each other to highlight the similarities and provide a deeper understanding of the relevant physics. Chapter 3 presents the fundamentals of MRI including signal detection, echo generation, spatial encoding and image reconstruction. The concept of k-space is introduced and used to describe common acquisition sequences. Finally, some of the challenges specific to MRI data are discussed from a signal processing perspective.

Part II of this thesis includes two novel contributions to quantitative MRI. Chapter 4 examines the problem of relaxation time estimation in the presence of localised spatial blurring. The blurring is due to the presence of relaxation during data acquisition, described by a linear filter as in [3, 4]. This thesis extends the model in [4] to include contributions from both transverse relaxation and field inhomogeneity. The complete model is used to derive the statistical estimation bias for commonly used estimators. To overcome the issue of bias, the estimation problem is posed using the detailed signal model and a near optimal Bayesian estimator is developed. Chapter 5 discusses the problem of estimating a distribution

of relaxation times from a single decay curve, first examined in [5]. This difficult problem is analysed from a statistical viewpoint using the Cramér-Rao lower bound [6]. In light of the analysis, a novel algorithm based on the model proposed in [7, 8] is presented to estimate the main features of the distribution.

Part III contributes to the emerging field of nonlinear spatial encoding [9, 10, 11]. Chapter 6 develops a new acquisition strategy that improves the image resolution in a region of interest. The technique uses the notion of local k-space described in [12] to exploit the position dependent resolution inherent in nonlinear encoding fields. Chapter 7 provides a framework for image reconstruction with generalised spatial encoding using the theory of frames [13]. This framework is used to analyse the reconstruction problem for the existing schemes of SENSE [14], PatLoc [9] and O-Space [11] and a practical performance metric is developed based on the variance of the reconstructed pixels.

Part I

Background

CHAPTER 2

Physics of magnetic resonance

Contents

2.1 Introduction

SPIN is the fundamental property of nuclei that makes MR imaging possible. Understanding the physics describing the behaviour of spins is essential to accurately model the behaviour of the MR signal. Thus a detailed description of the physics describing systems of one or many spins is provided in this chapter. The behaviour is examined from both a classical and quantum perspective to give a deeper understanding of the topics covered. The chapter explores the basics of spin systems with focus on spin-1/2 systems relevant to MRI. An introduction

Table 2.1: Common notation used in Chapter 2

Symbol	Quantity	Units
	Physical constants	
K_B	Boltzmann constant, 1.3807×10^{-23}	J K^{-1}
$\hbar$	Reduced Planck's constant, 1.0546×10^{-34}	J s
γ	Gyromagnetic ratio, 2.6751×10^{8} (for ^{1}H)	rad/s/T
	Classical physics	
$\boldsymbol{b}$	Magnetic field	T
$\boldsymbol{m}$	Magnetisation	
T_1, T_2	Longitudinal and transverse relaxation times	s
$\hat{\boldsymbol{i}}$, $\hat{\boldsymbol{j}}$, $\hat{\boldsymbol{k}}$	Unit vectors along the x, y and z axes	
$p(\omega)$	Distribution of isochromats	
λ	Width of Lorentzian distribution	
	Quantum physics	
$\psi(t)$	Quantum state	
ρ	Density matrix	
$\lvert\alpha\rangle$, $\lvert\beta\rangle$	Eigenstates	
$\mathscr{H}$	Hamiltonian operator	
$\hat{A}$	Basis operators for Hamiltonian decomposition	
$F(t)$	Coefficients of Hamiltonian decomposition	
$J(\omega)$	Spectral density functions	
$\hat{I}_x$, $\hat{I}_y$, $\hat{I}_z$	Spin angular momentum operators	
$\hat{R}$	Rotation operator	
$\hat{\Gamma}$	Relaxation operator	

to concepts such as thermal equilibrium, precession, RF pulses and relaxation is provided.

2.1.1 Notation

Table 2.1 lists the important quantities and their associated notation used in this chapter. Other notation not listed in this table will be introduced as it is required.

2.2 Atomic nuclei and spin

Matter is made up of atoms which in turn are made up of electrons and *nuclei*. Nuclei consists of sub-atomic particles such as quarks and gluons. Nuclei have four physical properties that arise from the properties of these sub-atomic particles: mass, electric charge, magnetic moment and spin. The magnetic moment, $\boldsymbol{\mu}$,

is related to the spin, $\hat{\boldsymbol{S}}$, by a fundamental symmetry theorem[15, page 26]. That is,

$$\boldsymbol{\mu} = \gamma \hat{\boldsymbol{S}} \tag{2.1}$$

where γ is the *gyromagnetic ratio*. This means that depending on the sign of γ, the spin angular momentum is either aligned with the magnetic moment, or opposite it. The gyromagnetic ratio is positive for most magnetic nuclei (including ^{1}H) and negative for electrons and a few atomic nuclei.

2.3 Spin dynamics

In this section we describe both classical and quantum descriptions of a spin system. In both cases, the dynamics of the spins is governed by a differential equation: the Bloch equation for a classical description and the Schrödinger equation (or Liouville equation) for quantum states.

2.3.1 Classical description

Bulk magnetisation $\boldsymbol{m}$ is the sum of the individual proton's magnetic moments over a local volume V,

$$\boldsymbol{m}(t) = \frac{1}{V} \sum_i \boldsymbol{\mu}_i(t) \tag{2.2}$$

where $\boldsymbol{\mu}_i$ is the magnetic moment of the i^{th} proton.

In the presence of a magnetic field aligned along the z-axis, the motion of the bulk magnetisation is governed by the *Bloch equation* [16],

$$\frac{d}{dt}\boldsymbol{m}(t) = \gamma \boldsymbol{m}(t) \times \boldsymbol{b} - \frac{(m_z - M_0)\hat{\boldsymbol{k}}}{T_1} - \frac{m_x \hat{\boldsymbol{i}} + m_y \hat{\boldsymbol{j}}}{T_2} \tag{2.3}$$

where $\boldsymbol{m} = [m_x, m_y, m_z]^T$ is the vector of bulk magnetisation, $\boldsymbol{b}$ is the magnetic field, M_0 is the equilibrium magnetisation, T_1 is the spin-lattice (longitudinal) relaxation time and T_2 is the spin-spin (transverse) relaxation time. The notation $\boldsymbol{m} \times \boldsymbol{b}$ denotes the vector cross product.

The notation in (2.3) is different from much of the literature, where capital letters are used for the magnetisation and magnetic field. This deliberate choice creates consistency with the remainder of the thesis where capital letters refer to constants and bold capital letters refer to matrices.

2.3.1.1 Rotating Frame of Reference

A common convention in MRI is to define the main magnetic field along the z-direction such that

$$\boldsymbol{b} = B_0\hat{\boldsymbol{k}} \tag{2.4}$$

In this case, a useful tool for analysis is to transform the Bloch equation to a rotating frame of reference where the x-y plane rotates about the z axis at the Larmor frequency ω_0. The Bloch equation in the rotating frame of reference becomes

$$\frac{d}{dt}\boldsymbol{m}'(t) = \gamma\boldsymbol{m}'(t) \times \boldsymbol{b}_{\text{eff}} - \frac{(m_{z'} - M_0)\hat{\boldsymbol{k}}'}{T_1} - \frac{m_{x'}\hat{\boldsymbol{i}}' + m_{y'}\hat{\boldsymbol{j}}'}{T_2} \tag{2.5}$$

where $\boldsymbol{b}_{\text{eff}} = \boldsymbol{b}' + \boldsymbol{\omega}_{\text{rot}}/\gamma$ is the effective field in the rotating frame. The vector $\boldsymbol{\omega}_{\text{rot}}$ describes the speed and direction of the rotating frame and $\boldsymbol{b}'$ and $\boldsymbol{m}'$ are the magnetic field and magnetisation transformed to the rotating frame. By convention, $\boldsymbol{\omega}_{\text{rot}} = -\omega\hat{\boldsymbol{k}}$ so the reference frame rotates in the same direction as spin precession. Often throughout this thesis the prime notation is used to indicate the rotating frame of reference. To distinguish from the rotating frame of reference, we refer to the static or non-rotating frame as the 'laboratory' frame of reference.

2.3.2 Quantum description

The state of a single quantum particle is described by a complex-valued wavefunction, $\psi(x,t)$. In this thesis we consider stationary particles and drop the spatial variable, x from the notation. For a spin-1/2 particle, there are two eigenstates of angular momentum along the z-axis called *Zeeman eigenstates*,

$$|\alpha\rangle = \left|\frac{1}{2}, +\frac{1}{2}\right\rangle; \qquad |\beta\rangle = \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \tag{2.6}$$

where the notation, $|I, M\rangle$ specifies the eigenstate using two quantum numbers, I and M. These eigenstates are useful since an arbitrary state can be described as a linear combination (or *superposition*) of these basic states. The state $|\psi\rangle$ can be written as $|\psi\rangle = c_\alpha|\alpha\rangle + c_\beta|\beta\rangle$ or in vector form

$$|\psi\rangle = \begin{bmatrix} c_\alpha \\ c_\beta \end{bmatrix} \tag{2.7}$$

where the Zeeman eigenstates are written as

$$|\alpha\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \quad |\beta\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{2.8}$$

The dynamics of the spin state, $|\psi\rangle$, is given by the *time-dependent Schrödinger equation*. In natural units (scaled by $\hbar^{-1}$) the equation is

$$\frac{d}{dt}\psi(t) = -j\mathscr{H}\psi(t), \tag{2.9}$$

where $j = \sqrt{-1}$. This equation completely describes the evolution of the quantum states of the particle.

A magnetic resonance experiment is only capable of observing the average effect of all spins in an object. The average state of a spin ensemble is accurately described using the density matrix. Formally, the density matrix can be defined by considering the summation of N_s spins with states, $|\psi_i\rangle$, $i = 1, \ldots, N_s$ as follows [17],

$$\hat{\rho} = \frac{1}{N_s}\sum_{i=1}^{N_s} |\psi_i\rangle\langle\psi_i| \tag{2.10}$$

where $\langle\psi_i|$ denotes the transpose of $|\psi_i\rangle$. Each state, $|\psi_i\rangle$ can be defined as in (2.7) in terms of the Zeeman eigenstates, $|\alpha\rangle$ and $|\beta\rangle$ denoting the low and high energy states, respectively. Using this representation, the density matrix can be written as follows,

$$\hat{\rho} = \begin{bmatrix} \overline{c_\alpha c_\alpha^*} & \overline{c_\alpha c_\beta^*} \\ \overline{c_\beta c_\alpha^*} & \overline{c_\beta c_\beta^*} \end{bmatrix} = \begin{bmatrix} \rho_{\alpha\alpha} & \rho_{\alpha\beta} \\ \rho_{\beta\alpha} & \rho_{\beta\beta} \end{bmatrix} \tag{2.11}$$

The overbar indicates an average over the ensemble. Diagonal terms ($\rho_{\alpha\alpha}$ and $\rho_{\beta\beta}$) are called *populations* and off-diagonal terms ($\rho_{\alpha\beta}$ and $\rho_{\beta\alpha}$) are called *coherences*. They have the following relationships [17],

$$\rho_{\alpha\beta} = \rho_{\beta\alpha}^* \tag{2.12}$$

$$\rho_{\alpha\alpha} + \rho_{\beta\beta} = 1 \tag{2.13}$$

A physical interpretation of these quantities is useful to develop some intuition. The population difference represents the net polarisation along the external field direction. Thus if $\rho_{\alpha\alpha} - \rho_{\beta\beta} > 0$, the net polarisation is aligned with the external field. The presence of non-zero coherences indicates transverse spin magnetisa-

tion, that is a net polarisation perpendicular to the external field, B_0.

The system dynamics are governed by the time-dependent Schrödinger equation (2.9) from which the Liouville equation for the density operator is derived [18, page 369]

$$\frac{d}{dt}\rho(t) = -j[\mathscr{H}, \rho(t)] \tag{2.14}$$

where $[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$ is known as the commutator. Importantly, two operators $\hat{A}$ and $\hat{B}$ are said to commute if and only if $[\hat{A}, \hat{B}] = 0$.

2.3.2.1 Rotating Frame of Reference

Analogous to the classical rotating frame of reference, the Schrödinger equation can be transformed to rotating coordinate system [19]. This reference frame is often referred to as the *interaction frame*. The system is transformed using the operator, $\hat{R}_z$, representing a rotation around the z-axis,

$$\hat{R}_z(\phi) = \begin{bmatrix} \exp\{-j\frac{1}{2}\phi\} & 0 \\ 0 & \exp\{+j\frac{1}{2}\phi\} \end{bmatrix} \tag{2.15}$$

Using this operator, a single spin state is transformed to the rotating frame of reference according to

$$|\tilde{\psi}\rangle(t) = \hat{R}_z(\omega_{\text{ref}}t)|\psi\rangle(t) \tag{2.16}$$

where ω_{ref} is the frequency of rotation and $|\tilde{\psi}\rangle(t)$ is the state in the rotating frame of reference. The Hamiltonian in the rotating frame is transformed according to

$$\tilde{\mathscr{H}} = \hat{R}_z(-\omega_{\text{ref}}t)\mathscr{H}\hat{R}_z(\omega_{\text{ref}}t) - \omega_{\text{ref}}\hat{I}_z \tag{2.17}$$

Finally, the Schrödinger equation in the interaction frame is

$$\frac{d}{dt}\tilde{\psi}(t) = -j\tilde{\mathscr{H}}\tilde{\psi}(t) \tag{2.18}$$

For an ensemble of spins, the density operator in the rotating-frame, $\tilde{\rho}$, is defined as

$$\tilde{\rho} = \overline{|\tilde{\psi}\rangle\langle\tilde{\psi}|} \tag{2.19}$$

where $\tilde{\psi}$ is a single state in the rotating frame (defined in (2.16)) and the overbar denotes an ensemble average. Equivalently, the density operator can be transformed

directly using

$$\tilde{\rho}(t) = \hat{R}_z(\omega_{\text{ref}}t)\rho(t)\hat{R}_z(-\omega_{\text{ref}}t) \tag{2.20}$$

In this frame of reference the Liouville equation becomes

$$\frac{d}{dt}\tilde{\rho}(t) = -j[\mathscr{H}, \tilde{\rho}(t)]. \tag{2.21}$$

2.3.2.2 Observations

A key feature of quantum mechanics is that it only provides the probabilities of observing particular results. Each observation is associated with a Hermitian operator and the result of an experimental observation is an eigenvalue of that operator. Specifically, the probability of observing the eigenvalue λ_n for observable operator, $\hat{Q}$, is

$$\Pr(\lambda_n) = |\langle \xi_n | \psi \rangle|^2 \tag{2.22}$$

where $|\xi_n\rangle$ is the eigenvector associated with the eigenvalue λ_n. Note that if the state is initially an eigenvector the probability of observing the corresponding eigenvalue is 1.

Although the result of a single observation is undefined prior to the actual measurement, the expectation is well defined as

$$\langle \hat{Q} \rangle = \langle \psi | \hat{Q} | \psi \rangle \tag{2.23}$$

$$= \text{Tr}\{|\psi\rangle\langle\psi|\hat{Q}\} \tag{2.24}$$

where Tr denotes the matrix trace.

The motivation for using the density operator lies in its ability to describe macroscopic observations. The average outcome from measuring individual spins of an ensemble is the sum of the expectation values.

$$Q_{\text{obs}} = \text{Tr}\{(|\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2| + \cdots)\hat{Q}\} \tag{2.25}$$

Substituting the density operator definition in (2.10) gives

$$Q_{\text{obs}} = N_{\text{s}}\text{Tr}\{\hat{\rho}\hat{Q}\} \tag{2.26}$$

From this viewpoint, the contribution of each spin to the macroscopic observation is $\text{Tr}\{\hat{\rho}\hat{Q}\}$.

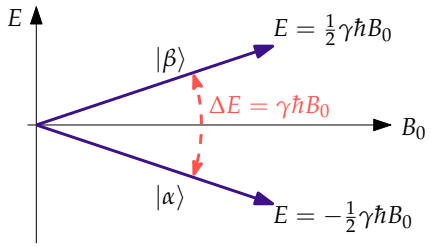


Figure 2.1: The energy levels of a spin-1/2 system in a magnetic field, illustrating low energy states, $|\alpha\rangle$, and high energy states, $|\beta\rangle$. The energy difference, ΔE, between states increases linearly with field strength, B_0.

The most useful measurement operators are the spin angular moment operators,

$$\hat{I}_x = \frac{1}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \quad \hat{I}_y = \frac{1}{2}\begin{bmatrix} 0 & -j \\ j & 0 \end{bmatrix}; \quad \hat{I}_z = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{2.27}$$

The notation of macroscopic observation can be used to relate the bulk magnetisation vector of the classical description, $\boldsymbol{m} = m_x\hat{\boldsymbol{i}} + m_y\hat{\boldsymbol{j}} + m_z\hat{\boldsymbol{k}}$, to the density matrix using

$$m_z = C\,\mathrm{Tr}\left(\rho\hat{I}_z\right) = \frac{1}{2}C(\rho_{\alpha\alpha} - \rho_{\beta\beta}) \tag{2.28a}$$

$$m_x = C\,\mathrm{Tr}\left(\rho\hat{I}_x\right) = C\,\mathrm{Re}(\rho_{\beta\alpha}) \tag{2.28b}$$

$$m_y = C\,\mathrm{Tr}\left(\rho\hat{I}_y\right) = C\,\mathrm{Im}(\rho_{\beta\alpha}) \tag{2.28c}$$

where the constant, C, is dependent on the thermal equilibrium and the number of spins, N_s. In Section 2.4 we will see how to define the constant in a meaningful way.

2.4 Thermal equilibrium

At thermal equilibrium, in the presence of an external magnetic field in the z-direction, the spins of a spin-1/2 system are observed in one of two possible orientations, spin-up or spin-down.

Spins in different orientations have different energy, a phenomena known as *Zeeman splitting*. The energy difference is $\Delta E = E_\beta - E_\alpha = \gamma\hbar B_0$. Figure 2.1 illustrates this phenomena and demonstrates that the energy difference is proportional to the field strength, B_0. Spins parallel to the B_0 field are in lower energy state than spins anti-parallel. This creates a population difference between the two spin

states that is related to the energy difference, ΔE, according to the Boltzmann distribution,

$$p(\alpha) = \frac{\exp(-E_\alpha / K_\mathrm{B} T)}{\exp(-E_\alpha / K_\mathrm{B} T) + \exp(-E_\beta / K_\mathrm{B} T)} \tag{2.29}$$

$$p(\beta) = \frac{\exp(-E_\beta / K_\mathrm{B} T)}{\exp(-E_\alpha / K_\mathrm{B} T) + \exp(-E_\beta / K_\mathrm{B} T)} \tag{2.30}$$

Since $\Delta E \ll K_\mathrm{B} T$ the exponentials can be approximated to first order by

$$\exp\left(-\frac{E}{K_\mathrm{B} T}\right) \approx 1 - \frac{E}{K_\mathrm{B} T} \tag{2.31}$$

Noting that $E_\alpha = -\frac{1}{2}\Delta E$ and $E_\beta = +\frac{1}{2}\Delta E$, the probabilities can be written as

$$p(\alpha) \approx \frac{1 + \mathbb{B}/2}{2}; \quad p(\beta) \approx \frac{1 - \mathbb{B}/2}{2} \tag{2.32}$$

where $\mathbb{B}$ denotes the *Boltzmann factor*, defined as

$$\mathbb{B} \triangleq \frac{\Delta E}{K_\mathrm{B} T} = \frac{\gamma \hbar B_0}{K_\mathrm{B} T}. \tag{2.33}$$

As an example we consider a typical MRI situation. An object is placed in a 3 T magnetic field at room temperature, $T = 300\,\mathrm{K}$. Substituting these values and the known physical constants into (2.33) results in probabilities such that approximately nine in one million protons contribute to the signal. Since the number of protons in a typical object is many orders of magnitude greater than one million, this tiny fraction is sufficient to provide high quality magnetic resonance data.

2.4.1 Semi-classical description

Thermal equilibrium can be described from a classical perspective by considering the number of up spins, $N_\uparrow$, and the number of down spins, $N_\downarrow$. This is given by the appropriate fraction of the total number of spins, N_s resulting in a population difference of

$$N_\uparrow - N_\downarrow = N_\mathrm{s} \left(p(\alpha) - p(\beta)\right) \approx N_\mathrm{s} \frac{\mathbb{B}}{2} \tag{2.34}$$

The spin orientation of an individual proton directly defines the direction of the magnetic moment vector according to according to the relation in (2.1). For spin-1/2 systems, the magnetic moment of the i^{th} proton is, $\boldsymbol{\mu}_i = \pm\frac{1}{2}\gamma\hbar\hat{\boldsymbol{k}}$, where the sign depends on the orientation of the proton's spin. The bulk magnetisation

is defined as the sum of magnetic moments (as in (2.2)) thus

$$\boldsymbol{m} = \frac{1}{2}(N_\uparrow - N_\downarrow)\gamma\hbar\hat{\boldsymbol{k}} \tag{2.35}$$

$$= \frac{\gamma^2\hbar^2 B_0 N_\mathrm{s}}{4K_\mathrm{B}T}\hat{\boldsymbol{k}} \tag{2.36}$$

It is useful to denote the magnetisation vector in thermal equilibrium as $\boldsymbol{m} = M_0\hat{\boldsymbol{k}}$ where we have defined the magnetisation strength as

$$M_0 \triangleq \frac{\gamma^2\hbar^2 B_0 N_\mathrm{s}}{4K_\mathrm{B}T} \tag{2.37}$$

2.4.2 Quantum description

Thermal equilibrium can be equivalently defined using the density matrix formulation. At thermal equilibrium the spin coherences are zero [17]. That is, $\rho_{\alpha\beta} = \rho_{\beta\alpha} = 0$ and the populations of energy states obey *Boltzmann distribution*. The high temperature approximation in (2.31) is adopted to give a density matrix at thermal equilibrium of

$$\hat{\rho}^0 = \begin{bmatrix} \frac{1}{2} + \frac{1}{4}\mathbb{B} & 0 \\ 0 & \frac{1}{2} - \frac{1}{4}\mathbb{B} \end{bmatrix} \tag{2.38}$$

These thermal equilibrium conditions provide an direct relationship between the bulk magnetisation vector of the classical description, $\boldsymbol{m}$ and the density matrix using the relationships in (2.28),

$$m_z = \gamma\hbar N_\mathrm{s}\frac{1}{2}(\rho_{\alpha\alpha} - \rho_{\beta\beta}) \tag{2.39a}$$

$$m_x = \gamma\hbar N_\mathrm{s}\,\mathrm{Re}(\rho_{\beta\alpha}) \tag{2.39b}$$

$$m_y = \gamma\hbar N_\mathrm{s}\,\mathrm{Im}(\rho_{\beta\alpha}) \tag{2.39c}$$

The normalisation has been chosen such that $\boldsymbol{m}(0) = M_0\hat{\boldsymbol{k}}$ as defined in (2.37).

2.5 Free precession

In the absence of an external magnetic field the spins are completely isotropic (all possible directions are equally represented). When a magnetic field is applied, the spin polarisation rotates around the field with a constant angle between the spin magnetic moment and the direction of the field. This is called precession. The

frequency of precession is given by the *Larmor frequency*, ω_0,

$$\omega_0 = \gamma B_0 \tag{2.40}$$

where B_0 is the magnetic field strength. This section examines this phenomena using classical and quantum descriptions.

2.5.1 Classical description

The presence of a static magnetic field causes the magnetisation to rotate around the direction of the field, a phenomenon known as precession. To see this, we solve the Bloch equation for a magnetic field aligned with the z-axis, $\boldsymbol{b} = B_0\hat{\boldsymbol{k}}$. We ignore relaxation effects by assuming $t \ll T_1, T_2$, which is realistic for short timescales. In this case,

$$\frac{d}{dt}\boldsymbol{m}(t) = \gamma \boldsymbol{m}(t) \times \boldsymbol{b}. \tag{2.41}$$

Given an arbitrary initial condition, $\boldsymbol{m}_0 = m_x(0)\hat{\boldsymbol{i}} + m_y(0)\hat{\boldsymbol{j}} + m_z(0)\hat{\boldsymbol{k}}$, the solution to (2.41) is

$$\boldsymbol{m}(t) = M_{xy}(0)\cos(\omega_0 t + \phi_0)\hat{\boldsymbol{i}} + M_{xy}(0)\sin(\omega_0 t + \phi_0)\hat{\boldsymbol{j}} + M_z(0)\hat{\boldsymbol{k}} \tag{2.42}$$

where $M_{xy}(0) = \sqrt{m_x^2(0) + m_y^2(0)}$, $\omega_0 = \gamma B_0$ and $\phi_0 = \arctan(m_y(0)/m_x(0))$. It is often instructive to consider the transverse magnetisation as a complex number, i.e.,

$$m_{xy}(t) = m_x(t) + jm_y(t) \tag{2.43}$$

With this notation the solution can be written as

$$m_{xy}(t) = M_{xy}(0)e^{-j(\omega_0 t + \phi_0)} \tag{2.44a}$$

$$m_z(t) = M_z(0) \tag{2.44b}$$

This clearly shows the bulk magnetisation precesses around the external field at the Larmor frequency $\omega_0 = \gamma B_0$.

Alternatively, we can use (2.5) to describe the magnetisation in a reference frame rotating at the Larmor frequency. We have $\boldsymbol{b} = B_0\hat{\boldsymbol{k}}$ and $\boldsymbol{\omega}_{\text{rot}} = -\gamma B_0\hat{\boldsymbol{k}}$ so $\boldsymbol{b}_{\text{eff}} = \boldsymbol{0}$. Neglecting relaxation leads to the condition,

$$\frac{d}{dt}\boldsymbol{m}'(t) = 0. \tag{2.45}$$

As expected, the bulk magnetisation is stationary in the rotating frame.

2.5.2 Quantum description

The Hamiltonian in the presence of a static magnetic field B_0 along the z-axis is $\mathscr{H} = \omega_0 \hat{I}_z$. The Schrödinger equation describing the state evolution is

$$\frac{d}{dt}|\psi\rangle(t) = -j\omega_0 \hat{I}_z |\psi\rangle(t) \tag{2.46}$$

The solution to this equation is

$$|\psi\rangle(t) = \exp\{-j\omega_0 t \hat{I}_z\}|\psi\rangle(0) \tag{2.47}$$

This can be recognised as a rotation about the z-axis and hence it can be described with the rotation operator in (2.15),

$$|\psi\rangle(t) = \hat{R}_z(\omega_0 t)|\psi\rangle(0) \tag{2.48}$$

This means that in the absence of an RF field, the Schrödinger equation implies that the spin rotates around the z-axis with frequency ω_0.

The Liouville equation can be solved to determine the evolution of the density matrix under a static magnetic field. Alternatively, we can use the evolution of a single state given in (2.48) to build the density matrix according to the definition in (2.10). In this case,

$$\rho(t) = \hat{R}(\omega_0 t)\rho(0)\hat{R}(-\omega_0 t) \tag{2.49}$$

Computing the individual elements of the matrix gives the following set of equations

$$\rho_{\alpha\alpha}(t) = \rho_{\alpha\alpha}(0) \tag{2.50a}$$
$$\rho_{\beta\beta}(t) = \rho_{\beta\beta}(0) \tag{2.50b}$$
$$\rho_{\alpha\beta}(t) = \exp\{-j\omega_0 t\}\rho_{\alpha\beta}(0) \tag{2.50c}$$
$$\rho_{\beta\alpha}(t) = \exp\{j\omega_0 t\}\rho_{\beta\alpha}(0) \tag{2.50d}$$

Ignoring relaxation, we see that the populations of the states remains unchanged and the coherences revolve around the complex plane at the resonant frequency ω_0.

2.6 RF pulse

2.6.1 Classical description

When a rotating magnetic field $B_1(t)$ is applied perpendicular to the B_0 field at the Larmor frequency of the spin system, the effect is to tilt the bulk magnetisation away from the z-axis. To see this consider a circularly polarised field,

$$\boldsymbol{b}_1(t) = B_1(t)\left\{\sin(\omega_{\text{rf}}t + \phi_p)\hat{\boldsymbol{i}} + \cos(\omega_{\text{rf}}t + \phi_p)\hat{\boldsymbol{j}}\right\} \tag{2.51}$$

where $B_1(t)$ is the amplitude envelope, ω_{rf} is the frequency of oscillation and ϕ_p is the initial phase angle.

To analyse the effect of such a pulse, we make the reasonable assumption that the RF pulse is short enough such that relaxation effects can be ignored for the duration of the pulse. The rotating frame Bloch equation in (2.5) reduces to

$$\frac{\partial}{\partial t}\boldsymbol{m}'(t) = \gamma \boldsymbol{m}'(t) \times \boldsymbol{b}_{\text{eff}} \tag{2.52}$$

where $\boldsymbol{b}_{\text{eff}} = \boldsymbol{b}' + \boldsymbol{\omega}_{\text{rot}}/\gamma$. Further, we assume the RF pulse has zero phase and constant envelope, $B_1(t) = B_1, 0 \leq t \leq T_{\text{p}}$, The RF pulse in (2.51) can be transformed to the rotating frame to give, $\boldsymbol{b}_1'(t) = B_1\hat{\boldsymbol{i}}$. When the resonance condition is satisfied (i.e. $\omega_{\text{rf}} = \omega_0$), and we adopt the Larmor rotating frame ($\boldsymbol{\omega}_{\text{rot}} = -\omega_0\hat{\boldsymbol{k}}$), the effective magnetic field is

$$\boldsymbol{b}_{\text{eff}} = B_0\hat{\boldsymbol{k}}' + B_1\hat{\boldsymbol{i}}' - \frac{\omega_0}{\gamma}\hat{\boldsymbol{k}} = B_1\hat{\boldsymbol{i}}' \tag{2.53}$$

The Bloch equation can be easily solved for an initial condition of $\boldsymbol{m}'(0) = M_0\hat{\boldsymbol{k}}'$ to give

$$\boldsymbol{m}'(t) = M_0\{\sin(\omega_1 t)\hat{\boldsymbol{j}}' + \cos(\omega_1 t)\hat{\boldsymbol{k}}'\} \tag{2.54}$$

where $\omega_1 = \gamma B_1$ is known as the Rabi frequency and represents the rotation speed of the bulk magnetisation away from the z-axis. This equation determines the motion of the bulk magnetisation for the duration of the pulse. At $t = T_{\text{p}}$ the bulk magnetisation has been rotated about the x'-axis through a flip angle $\theta = \omega_1 T_{\text{p}}$. For more complex envelopes the flip angle is given by

$$\theta = \int_0^{T_{\text{p}}} \gamma B_1(\tau)d\tau. \tag{2.55}$$

This equation is valid for on-resonance excitation with a duration sufficiently short compared to the relaxation times.

2.6.2 Quantum description

During an RF pulse the spin Hamiltonian is made up of the static Hamiltonian $\mathscr{H}_0$ and RF Hamiltonian, $\mathscr{H}_{\mathrm{rf}}(t)$. Ignoring any off-resonance effects, the complete Hamiltonian is given in [20] as

$$\mathscr{H}(t) = \mathscr{H}_0 + \mathscr{H}_{\mathrm{rf}} \tag{2.56}$$

$$= \omega_0 \hat{I}_z - \gamma B_1 \{\cos(\omega_{\mathrm{rf}} t + \phi_p)\hat{I}_x + \sin(\omega_{\mathrm{rf}} t + \phi_p)\hat{I}_y\} \tag{2.57}$$

We transform the system to a reference frame rotating at ω_{rf} about the z-axis. The transformed Hamiltonian is computed using (2.17). Simplifying with the "sandwich relationship", the final Hamiltonian can be expressed as

$$\tilde{\mathscr{H}} = \omega_{\mathrm{off}} \hat{I}_z + \omega_{\mathrm{nut}} \left(\hat{I}_x \cos\phi_p + \hat{I}_y \sin\phi_p\right) \tag{2.58}$$

where $\omega_{\mathrm{off}} = \omega_0 - \omega_{\mathrm{rf}}$ and $\omega_{\mathrm{nut}} = |\gamma B_1|$. Note the Hamiltonian in the rotating frame of reference is independent of time.

When the pulse is applied exactly on resonance ($\omega_{\mathrm{off}} = 0$), the solution to the Schrödinger equation is

$$|\tilde{\psi}\rangle(t) = \hat{R}_{\phi_p}(\omega_{\mathrm{nut}} t)|\tilde{\psi}\rangle(0). \tag{2.59}$$

This can be recognised as a rotation of the initial state using the general rotation operator, $\hat{R}_{\phi_p}(\theta) = \hat{R}_z(\phi_p)\hat{R}_x(\theta)\hat{R}_z(-\phi_p)$.

Now consider an on-resonance 'x-pulse', where $\phi_p = 0$. The rotating-frame spin Hamiltonian is $\tilde{\mathscr{H}} = \omega_{\mathrm{nut}}\hat{I}_x$ and the solution is a rotation about the x'-axis. For example, suppose the initial state is the eigenstate $|\alpha\rangle$; after time $\tau_{\mathrm{p}} = \pi/(2\omega_{\mathrm{nut}})$ the spin will be rotated through a flip angle of $\pi/2$ given by the equation

$$|\psi\rangle(\tau_{\mathrm{p}}) = \hat{R}_x(\pi/2)|\alpha\rangle \tag{2.60}$$

The evolution of the density matrix can be derived by averaging the effect of an RF pulse on individual states evolving according to (2.59). The evolution of the density operator is

$$\tilde{\rho}(t) = \hat{R}_{\phi_p}(\omega_{\mathrm{nut}} t)\tilde{\rho}(0)\hat{R}_{\omega_{\mathrm{nut}} t}(-\omega_{\mathrm{nut}} t) \tag{2.61}$$

After time τ_p the density matrix has been rotated by angle θ defined in (2.55). Notice that a π pulse inverts the population distribution while a $\pi/2$ pulse equalises the spin state populations and generates coherences.

2.7 Relaxation

There are two types of relaxation processes in a spin system caused by different physical mechanisms: spin-lattice relaxation and spin-spin relaxation.

Spin-lattice relaxation refers to the process by which the spin system exchanges energy with its external surroundings (or lattice). This is associated with transitions from high energy states to low energy states, which affects the population difference of the spin states and ultimately the longitudinal magnetisation. The time constant T_1 is associated with the time required for the spin system to reach thermal equilibrium. For these reasons spin-lattice relaxation is also known as T_1 or longitudinal relaxation.

Spin-spin relaxation, on the other hand, involves the spins exchanging energy among themselves. For example, one spin may cause a second spin to transition from high to low energy states while the first spin transitions from low to high energy. In this case, the population of states does not change and the longitudinal magnetisation will not be affected. However, transitions of this type result in a loss of coherence between spin states. This coherence loss manifests as a decrease in transverse magnetisation related to the time constant, T_2. This type of relaxation is thus referred to as T_2 or transverse relaxation.

This description illuminates the close relationship between the molecular structure of an object and relaxation. Images produced by MRI machines are strongly dependent on the relaxation properties and therefore the underlying structure. The difference between relaxation properties of normal and diseased tissues is often the basis for a diagnosis using MRI. Table 2.2 lists some typical values for T_1, T_2 and M_0 for various healthy tissue types at 1.5 T [21] although the values can vary significantly depending on the particular experiment (see e.g. [22]).

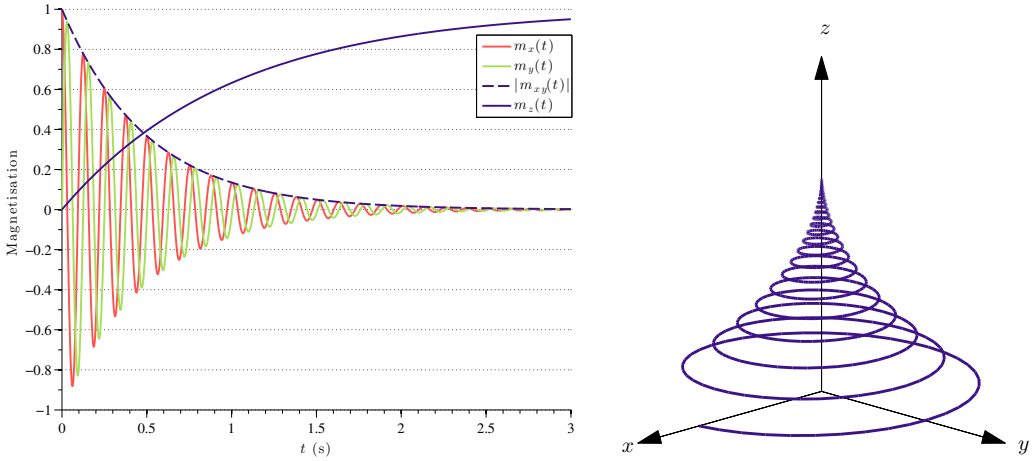
In the following section we examine these relaxation processes in more detail. This is important to gain a deeper understanding of the relaxation parameters we estimate in Chapters 4 and 5.

2.7.1 Classical description

Consider the evolution of the magnetisation after an RF pulse has been applied. The main B_0 magnetic field is still present along the z-direction. The Bloch equa-

Table 2.2: Typical values of proton density and relaxation time constants, T_1 and T_2 for biological tissue at 1.5 T. Sourced from [21].

Tissue Type	T_1 (ms)	T_2 (ms)	M_0
CSF	2400	160	1.0
White matter	780	90	
Gray Matter	900	100	
Muscle	870	45	
Liver	500	40	
Fat	270	80	0.9



(a) (b)

Figure 2.2: Relaxation viewed in the laboratory frame of reference illustrating (a) the individual components of the magnetisation vector during relaxation and (b) the trajectory of the tip of the magnetisation vector.

tion can be solved in the rotating frame using (2.5). Adopting the complex notation in (2.43) the solution is

$$m_{x'y'}(t) = m_{x'y'}(0)e^{-t/T_2} \tag{2.62a}$$

$$m_{z'}(t) = M_0(1 - e^{-t/T_1}) + m_{z'}(0)e^{-t/T_1} \tag{2.62b}$$

where $m_{x'y'}(0)$ and $m_{z'}(0)$ define the magnetisation immediately after the pulse. This solution can easily be interpreted as a decaying transverse magnetisation with time constant T_2, and a recovering longitudinal magnetisation with time constant T_1. Figure 2.2 illustrates the evolution of the bulk magnetisation for $T_1 = 1\,\text{s}$ and $T_2 = 500\,\text{ms}$. To emphasise the general behaviour, the system was simulated with a Larmor frequency many orders of magnitude below realistic values.

In addition to these two relaxation processes, spins in the object experience

different local fields due to inhomogeneity and/or applied gradients. The local field variation leads to spins with a distribution of different precession frequencies. The term *isochromat* is used to define a group of spins with the same precession frequency. In the laboratory frame of reference, the transverse magnetisation from a single isochromat (with Larmor frequency ω) is

$$m_{xy}(t,\omega) = m_{xy}(0)e^{-t/T_2}e^{j\omega t} \tag{2.63}$$

The final signal observed is the integration of the relative contributions of all such isochromats,

$$m_{xy}(t) = \int m_{xy}(t,\omega)p(\omega)d\omega \tag{2.64}$$

where $p(\cdot)$ is the distribution of isochromats.

Some intuition can be gained by considering a Lorentzian distribution of the Larmor frequencies,

$$p(\omega) = \frac{1}{\pi}\frac{\lambda}{\lambda^2 + (\omega - \omega_0)^2} \tag{2.65}$$

where ω_0 is the centre of the distribution and λ is half-width at half-maximum (HWHM) representing the distribution spread. It should be noted that, despite its prevalence, there is no physical reason to adopt a Lorentzian distribution. It is chosen for mathematical convenience. Nonetheless, it is useful to demonstrate some general properties of the free induction decay. In this case, the integral in (2.64) can be calculated in closed form using the Fourier transform, $\mathcal{F}$, as follows.

$$\begin{aligned} m_{xy}(t) &= \int m_{xy}(0)e^{-t/T_2}e^{j\omega t}p(\omega)d\omega && (2.66)\\ &= m_{xy}(0)e^{-t/T_2}\int e^{j\omega t}p(\omega)d\omega && (2.67)\\ &= m_{xy}(0)e^{-t/T_2}(\mathcal{F}p)(t) && (2.68)\\ &= m_{xy}(0)e^{-t/T_2}e^{-\lambda t}e^{-j\omega_0 t} && (2.69) \end{aligned}$$

This highlights the effect of the distribution on transverse relaxation. Specifically, the frequency distribution adds an additional relaxation component dependent on the spread of frequencies, λ. A very narrow distribution will result in an ideal T_2 decay, while a wide distribution will result in a faster decay. This can also be understood using the notion of dephasing discussed in Chapter 3. For the Lorentzian distribution, both relaxation processes can be aggregated into a single exponential

decay,

$$m_{xy}(t) = m_{xy}(0)e^{-t/T_2^*}e^{-j\omega_0 t} \tag{2.70}$$

where the time constant T_2^* consists of T_2 and $T_2' = 1/\lambda$ according to

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_2'} \tag{2.71}$$

2.7.2 Quantum description

At a quantum physics level, relaxation is mostly due to the following mechanisms: 1) dipole-dipole coupling, 2) chemical shift anisotropy, 3) spin-rotation interaction. The order listed is the usual order of importance where chemical shift becomes increasingly important at high field strengths and begins to compete with dipole-dipole coupling.

2.7.2.1 Internal Spin Interactions

The spin angular momentum is affected by the magnetic and electric fields of interacting particles in the sample. A single proton undergoes rapid movement and electromagnetic interactions with neighbouring particles (protons, electrons etc), each of which has its own magnetic moment. This causes tiny fluctuations in the magnetic field experienced by a given proton. The physical interaction mechanisms are described below:

Chemical Shift External magnetic field affects magnetism of electrons and they in turn affect the nuclear spin.

Dipole-Dipole Coupling Direct magnetic interactions of nuclear spins with each other.

J-Coupling Indirect magnetic interactions of nuclear spins, via interactions with electrons.

Spin-Rotation Interaction Nuclear spins interacting with magnetic fields generated by the rotational motion of the molecules.

Quadrupolar Coupling (for spins $> 1/2$) Electric interactions of nuclei with surrounding electric fields.

For spin-1/2 nuclei dipole-dipole and chemical shift interactions are the strongest [15]. All these interactions cause fluctuations in the local magnetic field and result in the relaxation effects described below.

The appropriate Hamiltonian describing these microscopic interactions can be used to derive the evolution of the spin system towards thermal equilibrium, i.e., relaxation.

2.7.2.2 Integral Form

The complete system Hamiltonian is composed as follows,

$$\mathscr{H}(t) = \mathscr{H}_0 + \mathscr{H}_1(t) \tag{2.72}$$

where $\mathscr{H}_0$ is the static Hamiltonian and $\mathscr{H}_1$ is the interaction Hamiltonian.

The system dynamics are governed by Liouville equation in (2.14). To simplify the calculations, we transform the system to the interaction frame using

$$\tilde{\rho}(t) = \exp\{+j\mathscr{H}_0 t\}\rho(t)\exp\{-j\mathscr{H}_0 t\} \tag{2.73}$$

$$\tilde{\mathscr{H}_1}(t) = \exp\{+j\mathscr{H}_0 t\}\mathscr{H}_1(t)\exp\{-j\mathscr{H}_0 t\} \tag{2.74}$$

The Liouville equation for the combined system in the interaction frame is

$$\frac{d}{dt}\tilde{\rho}(t) = -j[\tilde{\mathscr{H}_1}(t), \tilde{\rho}(t)] \tag{2.75}$$

Integrating this equation by successive approximations up to second order gives

$$\tilde{\rho}(t) \approx \tilde{\rho}(0) - j\int_0^t [\tilde{\mathscr{H}_1}(t'), \tilde{\rho}(0)]dt' - \int_0^t dt' \int_0^{t'} dt''[\tilde{\mathscr{H}_1}(t'), [\tilde{\mathscr{H}_1}(t''), \tilde{\rho}(0)]] \tag{2.76}$$

Differentiating (2.76) with respect to time gives the 'integrated to second order' approximation [23, page 276] and a change of variable ($\tau = t - t'$) gives

$$\frac{d}{dt}\tilde{\rho}(t) \approx -j[\tilde{\mathscr{H}_1}(t), \tilde{\rho}(0)] - \int_0^t d\tau[\tilde{\mathscr{H}_1}(t), [\tilde{\mathscr{H}_1}(t-\tau), \tilde{\rho}(0)]] \tag{2.77}$$

Since $\tilde{\mathscr{H}_1}(t)$ is a random operator, (2.77) indicates that $\tilde{\rho}(t)$ is also a random operator and the observable behaviour will be described by an average density operator $\bar{\tilde{\rho}}$ which is described by the above equation averaged over all the random Hamiltonians $\tilde{\mathscr{H}_1}(t)$. The average equation is

$$\frac{d}{dt}\overline{\tilde{\rho}(t)} \approx -j\overline{[\tilde{\mathscr{H}_1}(t), \tilde{\rho}(0)]} - \int_0^t d\tau\overline{[\tilde{\mathscr{H}_1}(t), [\tilde{\mathscr{H}_1}(t-\tau), \tilde{\rho}(0)]]} \tag{2.78}$$

Now the following assumptions are made:

A2.1 $\overline{\tilde{\mathscr{H}_1}(t)} = 0$. Otherwise, we can redefine $\mathscr{H}_0$ to include it.

A2.2 We can neglect the correlation between $\tilde{\mathscr{H}}_1(t)$ and $\tilde{\rho}(0)$ and average them separately.

A2.3 We can replace $\tilde{\rho}(0)$ by $\tilde{\rho}(t)$ on the right hand side of (2.78).

A2.4 We can extend the upper limit of the integral to $+\infty$.

A2.5 All unwritten higher-order terms can be neglected.

These assumptions, justified in [23, page 282], result in the following evolution equation,

$$\frac{d}{dt}\tilde{\rho}(t) = -\int_0^\infty d\tau \overline{[\tilde{\mathscr{H}}_1(t), [\tilde{\mathscr{H}}_1(t-\tau), \tilde{\rho}(t)]]}. \tag{2.79}$$

For the sake of brevity, we use the notation $\tilde{\rho}(t)$ to represent the density matrix averaged over the ensemble. In general, the density matrix in (2.79) evolves to zero, since the above theory does not model interactions with the lattice. In this case, an adjustment needs to be made for the semi-classical theory where $\tilde{\rho}(t)$ is replaced by $\tilde{\rho}(t) - \rho_0$ to ensure the system relaxes to the equilibrium state ρ_0.

2.7.2.3 Operator Form

We decompose the interaction Hamiltonian, $\mathscr{H}_1(t)$, as

$$\mathscr{H}_1(t) = \sum_q F^q(t)\hat{A}^q \tag{2.80}$$

where $F^q(t)$ are random functions of time representing classical stochastic forces independent of spin. A^q are operators acting on the variables of the system (the spins). Since $\hat{A}^q$ is not necessarily Hermitian and $\mathscr{H}_1(t)$ is required to be Hermitian, $\hat{A}^{-q} = \hat{A}^{q\dagger}$ and $F^{-q}(t) = F^{q*}(t)$ by convention [23]. Specific examples of Hamiltonians are described later.

Additionally, the operators $\hat{A}^q$ are decomposed into basis operators

$$\hat{A}^q = \sum_p \hat{A}^q_p \tag{2.81}$$

where the following relationship is satisfied,

$$[\mathscr{H}_0, \hat{A}^q_p] = \omega^q_p \hat{A}^q_p \tag{2.82}$$

Here the eigenoperators $\hat{A}^q_p$ correspond to transitions between different energy levels of the system, associated with a change in total magnetic quantum number

q. The eigenfrequencies ω_p^q correspond to the energy dissipated into the lattice by the spin transitions. [24]

The basis operators in (2.80) are transformed to the interaction frame, which gives them a time dependence.

$$\hat{A}^q(t) = \exp\{+j\mathscr{H}_0 t\}\hat{A}^q \exp\{-j\mathscr{H}_0 t\} \tag{2.83}$$

$$= \sum_p \hat{A}_p^q e^{j\omega_p^q t} \tag{2.84}$$

The proof can be found in [20, page 258]. Note that $\tilde{\mathscr{H}_1}(t)$ now contains two sources of time dependence: the fluctuating fields and the operators $\hat{A}^q(t)$ so the transformed Hamiltonian can be written as

$$\tilde{\mathscr{H}_1}(t) = \sum_q \sum_p F^q(t)\hat{A}_p^q e^{j\omega_p^q t} \tag{2.85}$$

Substituting (2.85) into (2.79) gives

$$\frac{d}{dt}\tilde{\rho}(t) = -\sum_{q,q'}\sum_{p,p'} e^{j(\omega_p^q+\omega_{p'}^{q'})t}[\hat{A}_{p'}^{q'},[\hat{A}_p^q,\tilde{\rho}(t)-\rho_0]]\int_0^\infty \overline{F^q(t)F^{q'}(t+\tau)}e^{j\omega_p^q\tau}d\tau \tag{2.86}$$

Note all the stochastic elements of the Hamiltonians are contained in the $F^q(t)$ functions and thus the ensemble average is taken over these.

With the assumption of stationarity, we define correlation functions,

$$g^{qq'}(\tau) = \overline{F^q(t)F^{q'}(t+\tau)} \tag{2.87}$$

and spectral density functions,

$$J^{qq'}(\omega) = \int_{-\infty}^{\infty} g^{qq'}(\tau)e^{-j\omega t}d\tau \tag{2.88}$$

The functions $F^q(t)$ and $F^{q'}(t)$ are assumed to be statistically independent such that the ensemble average vanishes unless $q' = -q$. Also noting that $\omega_{p'}^{-q} = -\omega_{p'}^q$, (2.86) becomes

$$\frac{d}{dt}\tilde{\rho}(t) = -\frac{1}{2}\sum_q\sum_{p,p'} e^{j(\omega_p^q-\omega_{p'}^q)t}[\hat{A}_{p'}^{-q},[\hat{A}_p^q,\tilde{\rho}(t)-\rho_0]]J^{qq}(\omega_p^q) \tag{2.89}$$

The factor of $\frac{1}{2}$ is introduced because of the Fourier transform limits. Also $g^{qq'}(\tau)$ is generally complex so the integral contains a complex part. This results in small energy shifts and can be included in the unperturbed Hamiltonian $\mathscr{H}_0$.

Terms in which $|\omega_p^q - \omega_{p'}^q| \gg 0$ are *nonsecular*, i.e., they do not affect the long-time behaviour of the system because the rapidly oscillating factors $e^{j(\omega_p^q - \omega_{p'}^q)t}$ average to zero faster than relaxation occurs. If none of the eigenfrequencies are degenerate (more than one eigenoperator associated with a single eigenfrequency), only *secular* terms in which $p = p'$ are non-zero. This leads to

$$\frac{d}{dt}\tilde{\rho}(t) = -\sum_q \sum_p J^{qq}(\omega_p^q)[\hat{A}_p^{-q}, [\hat{A}_p^q, \tilde{\rho}(t) - \rho_0]] \tag{2.90}$$

Converting (2.90) back to the laboratory frame yields the modified Liouville-von Neuman equation for relaxation,

$$\frac{d}{dt}\rho(t) = -j[\mathscr{H}_0, \rho(t)] - \hat{\Gamma}\left(\rho(t) - \rho_0\right) \tag{2.91}$$

where the relaxation operator is

$$\hat{\Gamma}(\rho) = \sum_q \sum_p J^{qq}(\omega_p^q)[\hat{A}_p^{-q}, [\hat{A}_p^q, \rho]] \tag{2.92}$$

2.7.2.4 Example

We now consider an example where the spins are coupled to a randomly fluctuating lattice. Spin-spin coupling is not considered but a similar analysis could be performed for Hamiltonians that model the interactions discussed in the beginning of this section. This example serves to demonstrate how macroscopic relaxation rates (T_1 and T_2) can be derived from the equations of motion with appropriate Hamiltonians. In our example a simple Hamiltonian is considered,

$$\mathscr{H}_1(t) = F^x(t)\hat{I}_x + F^y(t)\hat{I}_y + F^z(t)\hat{I}_z \tag{2.93}$$

which models random fluctuations of the angular moment operators. It is common to examine the relaxation of the density matrix by explicitly calculating the elements as in [25]; although, in this example, it is sufficient to following the procedure outlined in Section 2.7.2.3. Noting that F^x, F^y and F^z are statistically independent we obtain the relaxation operator,

$$\hat{\Gamma}(\rho) = J^{xx}(\omega_0)[\hat{I}_x, [\hat{I}_x, \rho]] + J^{yy}(\omega_0)[\hat{I}_y, [\hat{I}_y, \rho]] + J^{zz}(0)[\hat{I}_z, [\hat{I}_z, \rho]] \tag{2.94}$$

We seek the evolution of the expected value of $\hat{I}_x$ and $\hat{I}_y$ representing the transverse magnetisation components and $\hat{I}_z$ representing the longitudinal magnetisa-

tion. The differential equation for the expectation of operator, $\hat{Q}$, is

$$\frac{d}{dt}\langle\hat{Q}\rangle = \frac{d}{dt}\mathrm{Tr}\left(\hat{Q}\rho\right) = \mathrm{Tr}\left(\frac{d}{dt}\rho(t)\hat{Q}\right) \tag{2.95}$$

Inserting the Liouville equation with relaxation derived in (2.91) gives,

$$\frac{d}{dt}\langle\hat{Q}\rangle = \mathrm{Tr}\left\{\left(-j[\mathscr{H}_0,\rho] - \hat{\Gamma}(\rho)\right)\hat{Q}\right\} \tag{2.96}$$

We begin with the longitudinal magnetisation and compute $\langle\hat{I}_z\rangle$. In this case we use the fact that $\mathrm{Tr}([\mathscr{H}_0,\rho]I_z) = 0$ to simplify the general equation above to,

$$\frac{d}{dt}\langle\hat{I}_z\rangle = -\mathrm{Tr}\left(\hat{\Gamma}(\rho-\rho_0)\hat{I}_z\right) \tag{2.97}$$

$$\begin{aligned} &= -J^{xx}(\omega_0)\mathrm{Tr}\left([\hat{I}_x,[\hat{I}_x,\rho-\rho_0]]\hat{I}_z\right) - J^{yy}(\omega_0)\mathrm{Tr}\left([\hat{I}_y,[\hat{I}_y,\rho-\rho_0]]\hat{I}_z\right) \\ &\quad - J^{zz}(0)\mathrm{Tr}\left([\hat{I}_z,[\hat{I}_z,\rho-\rho_0]]\hat{I}_z\right) \end{aligned} \tag{2.98}$$

We let $\hat{D} = \rho - \rho_0$ and calculate the trace of each term above,

$$\mathrm{Tr}\left([\hat{I}_x,[\hat{I}_x,\hat{D}]]\hat{I}_z\right) = \mathrm{Tr}\left((\hat{I}_x^2\hat{D} + \hat{D}\hat{I}_x^2 - 2\hat{I}_x\hat{D}\hat{I}_x)\hat{I}_z\right) \tag{2.99}$$

$$= \frac{1}{2}\mathrm{Tr}\left(\hat{D}\hat{I}_z\right) - 2\mathrm{Tr}\left(\hat{D}\hat{I}_x\hat{I}_z\hat{I}_x\right) \tag{2.100}$$

$$= \frac{1}{2}\mathrm{Tr}\left(\hat{D}\hat{I}_z\right) + \frac{1}{2}\mathrm{Tr}\left(\hat{D}\hat{I}_z\right) \tag{2.101}$$

$$= \langle\hat{I}_z\rangle - \langle\hat{I}_z\rangle_0 \tag{2.102}$$

where $\langle\hat{I}_z\rangle_0$ is the expected value of the $\hat{I}_z$ operator at thermal equilibrium. The computations above have exploited the cyclic property of trace, the commutativity of $\hat{I}_x^2$ and $\hat{I}_z$ and the relation $\hat{I}_x\hat{I}_z\hat{I}_x = -\hat{I}_z/4$. Similarly for the second term we have

$$\mathrm{Tr}\left([\hat{I}_y,[\hat{I}_y,\rho-\rho_0]]\hat{I}_z\right) = \langle\hat{I}_z\rangle - \langle\hat{I}_z\rangle_0 \tag{2.103}$$

The third term is 0 which can be seen by expanding the commutator similarly to the first term above. Combining these results gives

$$\frac{d}{dt}\langle\hat{I}_z\rangle = -\left(J^{xx}(\omega_0) + J^{yy}(\omega_0)\right)\left(\langle\hat{I}_z\rangle - \langle\hat{I}_z\rangle_0\right) \tag{2.104}$$

This is related to the classical Bloch time constant, T_1, by

$$\frac{1}{T_1} = J^{xx}(\omega_0) + J^{yy}(\omega_0) \tag{2.105}$$

Next we calculate the transverse magnetisation given by $\langle \hat{I}_x \rangle$ and $\langle \hat{I}_y \rangle$. Unlike the longitudinal component $\hat{I}_x$ and $\hat{I}_y$ do not commute with $\mathscr{H}_0$ so the first term in (2.96) is not zero.

$$\begin{aligned} \mathrm{Tr}\left\{-j[\mathscr{H}_0, \rho]\hat{I}_x\right\} &= j\omega_0 \mathrm{Tr}\left([\hat{I}_z, \rho]\hat{I}_x\right) && (2.106)\\ &= j\omega_0 \mathrm{Tr}\left([\hat{I}_x, \hat{I}_z]\rho\right) && (2.107)\\ &= \omega_0 \mathrm{Tr}\left(\hat{I}_y \rho\right) && (2.108)\\ &= \omega_0 \langle \hat{I}_y \rangle && (2.109) \end{aligned}$$

The relaxation operator is composed of three terms which are evaluated below,

$$\begin{aligned} \mathrm{Tr}\left([\hat{I}_x, [\hat{I}_x, \rho - \rho_0]]\hat{I}_x\right) &= 0 && (2.110)\\ \mathrm{Tr}\left([\hat{I}_y, [\hat{I}_y, \rho - \rho_0]]\hat{I}_x\right) &= \langle \hat{I}_x \rangle && (2.111)\\ \mathrm{Tr}\left([\hat{I}_z, [\hat{I}_z, \rho - \rho_0]]\hat{I}_x\right) &= \langle \hat{I}_x \rangle && (2.112) \end{aligned}$$

where we assume $\langle \hat{I}_x \rangle_0 = 0$. Combining the results gives a differential equation for $\langle \hat{I}_x \rangle$,

$$\frac{d}{dt}\langle \hat{I}_x \rangle = \omega_0 \langle \hat{I}_y \rangle + \left(J^{yy}(\omega_0) + J^{zz}(0)\right) \langle \hat{I}_x \rangle \tag{2.113}$$

Similar computations for $\langle \hat{I}_y \rangle$ can be performed.

$$\frac{d}{dt}\langle \hat{I}_y \rangle = -\omega_0 \langle \hat{I}_x \rangle + \left(J^{xx}(\omega_0) + J^{zz}(0)\right) \langle \hat{I}_y \rangle \tag{2.114}$$

These coupled equations clearly represent precession and relaxation of the transverse magnetisation. In this example we assume $J^{xx}(\omega_0) = J^{yy}(\omega_0)$, which is reasonable for microscopic interactions. The relaxation is related to the classical Bloch time constant, T_2, by

$$\frac{1}{T_2} = J^{yy}(\omega_0) + J^{zz}(0) \tag{2.115}$$

The transverse relaxation time is made up of two terms which have different physical meanings. The term $J^{yy}(\omega_0)$ represents the *non-secular* contribution. This involves state transitions that are induced due to random fluctuations in the lattice. The frequency of this random fluctuation must match the energy difference between states, hence the $J(\omega_0)$ term. The second term $J^{zz}(0)$ represents the *secular* contribution. This does not involve transitions with the lattice (also called adiabatic relaxation). Instead, the random fluctuations in the z-direction, $F^z(t)$, superimpose on the static B_0 field to slightly alter the Larmor frequency of each

spin. This distribution of Larmor frequencies creates rapid dephasing. In MRI literature, the classical model is often modified by introducing different isochromats and integrating over a Lorentzian distribution to create the so-called T_2^*. In quantum mechanics, the spread is implicitly modelled by the stochastic nature of the interaction Hamiltonian $\mathscr{H}_1$.

As mentioned at the start of this section, the previous analysis could be performed for more complicated Hamiltonians to model actual relaxation effects such as dipole-dipole coupling and chemical shift anisotropy (CSA) relaxation. One would then expect physical parameters that define these Hamiltonians to appear in the various differential equations describing the behaviour of the ensemble. Such parameters may include the internuclear distance r, chemical shielding σ or J-coupling J. We would also expect terms such as $J^{zz}(2\omega_0)$ to appear, to model double-quantum transitions. For example, [15, page 537] states the transverse relaxation time-constant for dipole-dipole relaxation as

$$\frac{1}{T_2} = \left(\frac{\mu_0 \gamma^2 \hbar}{4\pi r^3}\right)^2 \{3J(0) + 5J(\omega_0) + 2J(2\omega_0)\} \tag{2.116}$$

We can interpret this to mean that T_2 is parameterised by four parameters: r, $J(0)$, $J(\omega_0)$, and $J(2\omega_0)$, which directly represent the physical parameters of the molecular environment.

This example demonstrates the underlying mechanisms of T_1 and in particular T_2 relaxation time, which is estimated in subsequent chapters of this thesis.

2.8 Summary

The physics behind spin systems and the generation of magnetisation is fundamental to MRI. This chapter has presented a detailed description using both classical and quantum mechanical theories. Phenomena such as thermal equilibrium, precession, control and relaxation have been described. In will be seen in the remaining chapters, that accurate models of the spin dynamics is essential for image reconstruction and advanced parameter estimation.

CHAPTER 3

Principles of magnetic resonance imaging

Contents

3.1 Introduction

A magnetic resonance scanner is composed of many components that work together to produce the final image. Figure 3.1 illustrates the main components of an MRI system. The scanner consists of a main magnet, spatial encoding gradients, an RF pulse generator to control the transmit coils and signal detection hardware including RF receive coils and analog-to-digital converters.

The main magnet is used to generate a large, static and homogeneous magnetic field. In the presence of this magnetic field, the signal behaves according the physical principles discussed in Chapter 2. The physics of spin systems is fundamental to the operation of the scanner.

All hardware components are controlled with an acquisition protocol known as a 'sequence'. The sequence contains a list of instructions and their precise timing; it can be considered the main input into the system.

In general only the transverse magnetisation is detectable so the first step in most MRI protocols is to excite the spin system to tip the magnetisation into the transverse plane. As described in the previous chapter, this can be achieved using an RF pulse tuned to the resonant frequency of the spin system. These pulses are controlled with RF hardware consisting of a pulse generator, amplifier and transmit coils.

Spatial information is encoded into the underlying signal using a set of magnetic fields known as gradients. These encoding gradients modulate the signal in a spatially dependent manner, which is later used to resolve the observed signal to form an image.

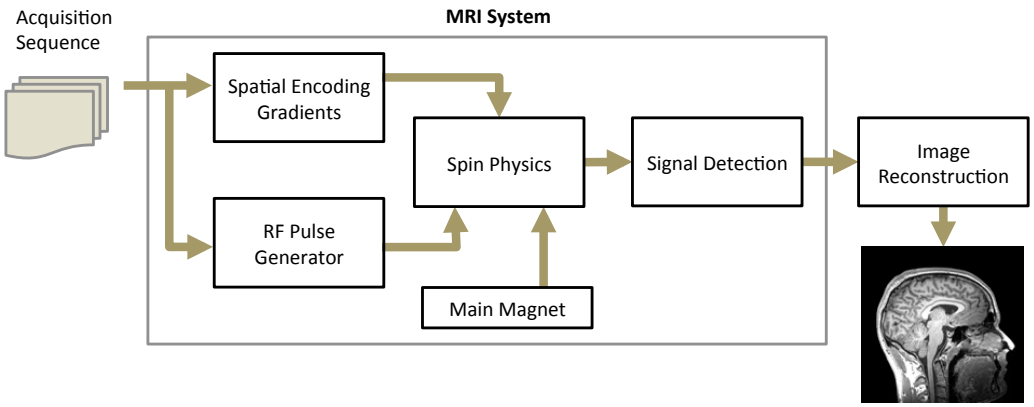


Figure 3.1: Overview of the main components in an MRI machine. The machine can be viewed as a system that inputs a list of precisely timed instructions (known as a 'sequence') and outputs raw data, which is used to produce the final image.

Table 3.1: Common notation used in Chapter 3

Symbol	Quantity	Units
Φ	Magnetic flux	Wb
$\boldsymbol{c}$	Coil sensitivity	
$\boldsymbol{m}$	Magnetisation vector	
V	Induced voltage	V
b	z-Component of the magnetic field	T
$\boldsymbol{g}$	Gradient vector	T/m
T_{p}	Duration of pulse	s
τ_{E}	Echo time	s
T_{acq}	Acquisition time (one readout)	s
T	Sampling period	s
SNR	Signal-to-noise ratio	
BW	Acquisition bandwidth	Hz

Signal detection involves analog-to-digital conversion of the voltage received on the RF coils. Receive RF coils are either volume coils or surface coils, depending on the application. Transmit and receive coils can be separate coils or combined into a single coil although there are two distinct roles: transmit coils excite the spins and rotate the magnetisation vector and receive coils detect the transverse magnetisation. Finally, the acquired signal is converted to an image, a process known as image reconstruction.

This chapter introduces the main components and basic principles of MRI. Additionally, the generation of echoes and the notion of k-space is introduced and used to describe common sequences. Finally, image reconstruction is considered from a general inverse problem perspective. The concepts outlined in this chapter are the foundation of the novel contributions in subsequent chapters of this thesis.

3.1.1 Notation

Table 3.1 lists the important symbols used in this chapter. Although the list is not exhaustive, it includes most quantities of interest.

3.2 Signal detection

Signal detection is the process of converting nuclear magnetisation into an electrical signal. The voltage induced on an RF receive coil, $V(t)$, is governed by Fara-

day's law of induction:

$$V(t) = -\frac{d\Phi_B(t)}{dt} \tag{3.1}$$

The magnetic flux through a coil is defined as

$$\Phi_B(t) = \int_{\text{obj}} \boldsymbol{m}(\boldsymbol{x}, t) \cdot \boldsymbol{c}(\boldsymbol{x}) d\boldsymbol{x} \tag{3.2}$$

where $\boldsymbol{m}$ is the magnetisation and $\boldsymbol{c}$ is the coil sensitivity. Combining (3.1) and (3.2) gives a received voltage of

$$V(t) = -\frac{\partial}{\partial t} \int_{\text{obj}} \boldsymbol{m}(\boldsymbol{x}, t) \cdot \boldsymbol{c}(\boldsymbol{x}) d\boldsymbol{x} \tag{3.3}$$

We adopt the common notation defined in (2.43) and let $m_{xy}(\boldsymbol{x}, t) = m_x(\boldsymbol{x}, t) + jm_y(\boldsymbol{x}, t)$. Since quadrature detection is used we can represent the coil sensitivity vector as

$$\boldsymbol{c}(\boldsymbol{x}) = c(\boldsymbol{x}) \begin{bmatrix} 1 \\ j \\ 0 \end{bmatrix} \tag{3.4}$$

The received voltage is demodulated to baseband to extract the underlying signal of interest, $s(t)$,

$$s(t) = V(t) e^{j\omega_0 t} \tag{3.5}$$

$$= -\left(\frac{\partial}{\partial t} \int_{\text{obj}} c(\boldsymbol{x}) m_{xy}(\boldsymbol{x}, t) d\boldsymbol{x} \right) e^{j\omega_0 t} \tag{3.6}$$

During detection, the transverse magnetisation, m_{xy}, is undergoing precession and relaxation as described in Chapter 2. The general form is

$$m_{xy}(\boldsymbol{x}, t) = M_0(\boldsymbol{x}) e^{-j(\omega_0 t + \phi_0)} e^{j\phi(\boldsymbol{x}, t)} e^{-t/T_2^*(\boldsymbol{x})}. \tag{3.7}$$

where M_0 is the initial magnetisation strength, ϕ_0 is the initial signal phase, $\phi(\boldsymbol{x}, t)$ is the spatial encoding phase and T_2^* is the free induction decay (FID) time. This leads to the detected signal,

$$s(t) = -\left(\frac{\partial}{\partial t} \int_{\text{obj}} c(\boldsymbol{x}) M_0(\boldsymbol{x}) e^{-j(\omega_0 t + \phi_0)} e^{j\phi(\boldsymbol{x}, t)} e^{-t/T_2^*(\boldsymbol{x})} d\boldsymbol{x} \right) e^{j\omega_0 t} \tag{3.8}$$

Simplifications of this general form are possible by making a series of assump-

tions. It is useful to invoke a different set of assumptions for different problems addressed in this thesis.

A3.1 The magnetisation precession is much faster than any time-dependent spatial encoding, inhomogeneity or relaxation processes. That is, we assume

$$\omega_0 \gg \frac{\partial \phi(\boldsymbol{x}, t)}{\partial t}, \quad \omega_0 \gg \frac{1}{T_2^*(\boldsymbol{x})}. \tag{3.9}$$

In this case,

$$s(t) \approx j\omega_0 \int_{\text{obj}} c(\boldsymbol{x}) M_0(\boldsymbol{x}) e^{j\phi_0} e^{j\phi(\boldsymbol{x},t)} e^{-t/T_2^*(\boldsymbol{x})} d\boldsymbol{x} \tag{3.10}$$

Notice the rotating ω_0 component is removed by the demodulation.

A3.2 The transverse magnetisation has zero phase initially. This corresponds to an initial bulk magnetisation on the x-axis.

A3.3 The ω_0 scaling factor can be considered constant across different experiments and hence we can arbitrarily set it to unity. Additionally the multiplication by j creates a global phase shift of π, which we ignore without loss of generality.

A3.4 Signal amplitude is constant during acquisition. In other words, the relaxation process occurs on a timescale much longer than the time of interest. During acquisition we have

$$e^{-t/T_2^*(\boldsymbol{x})} \approx e^{-\tau_{\text{E}}/T_2^*(\boldsymbol{x})} \tag{3.11}$$

Combining assumptions A3.1 to A3.4 the received signal can be written as

$$s(t) \approx \int_{\text{obj}} c(\boldsymbol{x}) M_0(\boldsymbol{x}) e^{-\tau_{\text{E}}/T_2^*(\boldsymbol{x})} e^{j\phi(\boldsymbol{x},t)} d\boldsymbol{x} \tag{3.12}$$

A3.5 The coil sensitivity is uniform and hence set to unity, i.e., $c(\boldsymbol{x}) = 1$. Neglecting receiver phase, this is a reasonable assumption for a birdcage coil [26]. Combining all previous assumptions, A3.1 to A3.5, yields,

$$s(t) \approx \int_{\text{obj}} m(\boldsymbol{x}) e^{j\phi(\boldsymbol{x},t)} d\boldsymbol{x} \tag{3.13}$$

where

$$m(\boldsymbol{x}) = M_0(\boldsymbol{x}) e^{-\tau_{\text{E}}/T_2^*(\boldsymbol{x})} \tag{3.14}$$

3.3 Signal echoes

Signal echoes are fundamental to some contrast mechanisms and form the basis of many MR imaging techniques. Due to the local field inhomogeneities or an applied gradient field the imaged object consists of a broad range of *isochromats*. That is, the spins are precessing at different frequencies centred about the resonance frequency. The distribution of frequencies mean that spins dephase very quickly relative to the intrinsic relaxation processes, resulting in rapid signal decay. It is possible to reverse the dephasing due to inhomogeneity to partially recover the signal, which is known as an 'echo'. An echo can be generated in two ways:

Gradient echo The distribution of frequencies is flipped by reversing the applied gradient field.

Spin echo The spins are rotated 180° using an RF pulse.

These echo generation mechanisms will be described in detail in the following sections.

3.3.1 Spin echoes and CPMG echoes

The generation of a spin echo was first demonstrated in [27] and since then, most MR imaging was performed by generating an echo some time after the initial excitation.

Figure 3.2 illustrates the generation of a spin echo. Figure 3.2a shows the bulk magnetisation at time $t = 0$, just after a 90° RF pulse. The spins begin to precess at slightly different speeds until τ seconds later when "faster" isochromats have a larger phase as shown in Figure 3.2b. At time τ a 180° pulse is applied to flip the vectors about the y-axis (Fig. 3.2c), causing the faster isochromats to lag the slower ones. After another τ seconds, the faster spins will have caught the slower ones and an echo is formed. The signal over the duration $\tau_E - T_{acq}/2 \leq t \leq \tau_E + T_{acq}/2$ is

$$s(t) = \int_{obj} M_0(\boldsymbol{x}) e^{-t/T_2(\boldsymbol{x})} e^{-|t-\tau_E|/T_2'(\boldsymbol{x})} e^{-j\boldsymbol{k}(t)\cdot\boldsymbol{x}} d\boldsymbol{x} \tag{3.15}$$

where T_2' captures the width of the off-resonance distribution as described in Chapter 2.

The process described by Figures 3.2 and 3.3 can be repeated by applying a train of 180° RF pulses spaced 2τ apart. In this case successive echoes will be generated at times, $t = 2\tau, 4\tau, \ldots, 2N_e\tau$. This sequence was first introduced by Carr and Purcell in [28] and modified by Meiboom and Gill in [29] and as such

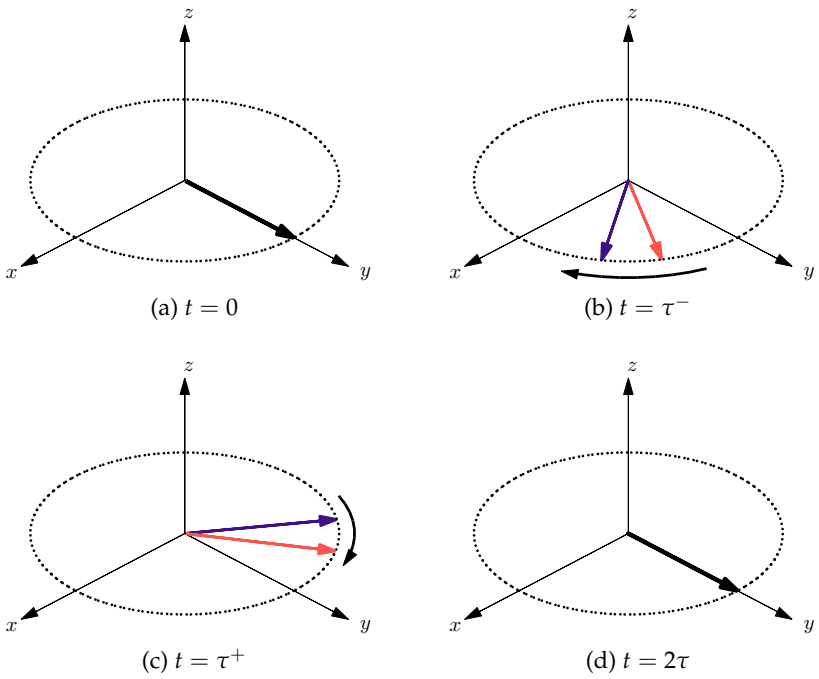


Figure 3.2: Formation of a spin echo illustrated by two vectors representing different isochromats. The 'fast' isochromat is indicated in blue and gains a positive phase at time $t = \tau^-$ relative to the 'slow' isochromat in red. After a 180° pulse at time $t = \tau^+$ the fast isochromat lags the slow isochromat. An echo is formed when the spins have the same phase again at $t = 2\tau$.

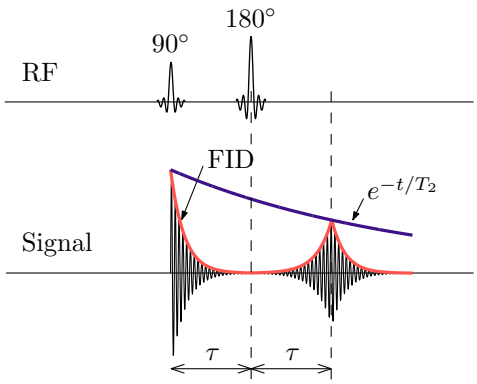


Figure 3.3: Timing diagram depicting the generation of a spin echo.

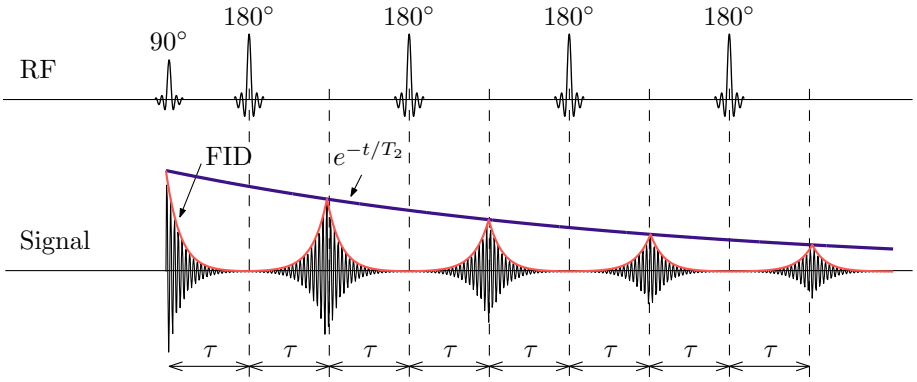


Figure 3.4: Timing diagram illustrating a CPMG sequence.

it is known as the CPMG sequence. Figure 3.4 demonstrates the timing of the RF pulses and subsequent echoes. The spin system undergoes spin-spin relaxation which represents a loss of coherence that cannot be recovered. As such the echo amplitudes are weighted by $e^{-2\tau n/T_2}$. This weighting makes the CPMG sequence very efficient for estimating the T_2 relaxation time from a set of echo amplitudes. The problem of T_2 estimation is explored in detail in Chapters 4 and 5.

3.3.2 Gradient echoes

Gradient fields are crucial for spatial encoding and they typically generate signal echoes as a consequence. During free induction decay (FID) the application of a gradient field over the sample creates a distribution of isochromats precessing at different frequencies. Analogous to the spin echo case, this distribution causes the spins to dephase and the overall signal decays much faster than the FID relaxation. This dephasing can be recovered by reversing the sign of the applied gradient field. The spins will be rephased after the reversed gradient has been applied for the same duration as the initial gradient, resulting in an echo. Figure 3.5 illustrates the generation of a gradient echo.

It is important to note that a gradient echo cannot reverse any dephasing due to underlying field inhomogeneity. As such, the echo amplitude will be weighted by the signal strength of the FID without gradients. The gradient echo occurs at time τ_E when the encoding phase $\phi = 0$. Under assumptions A3.1 to A3.5, the signal intensity during the echo is easily derived from (3.13) as

$$s(t) = \int_{\text{obj}} M_0(\boldsymbol{x}) e^{-\tau_E/T_2^*(\boldsymbol{x})} e^{j\phi(\boldsymbol{x},t)} d\boldsymbol{x}, \tag{3.16}$$

which highlights the dependance on T_2^*.

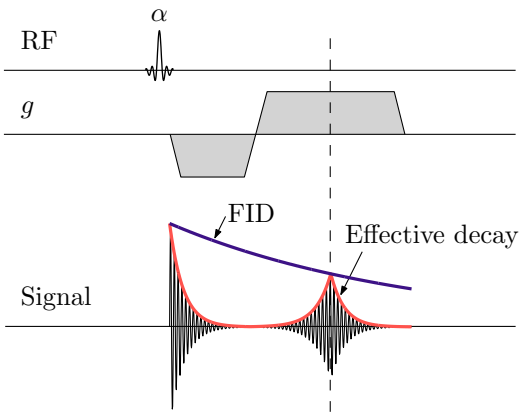


Figure 3.5: Timing diagram depicting the generation of a gradient echo.

3.4 Spatial encoding

Spatial encoding is the process of encoding the positions of different magnetisation vectors in an object. This encoding enables image reconstruction and is fundamental to all MRI techniques. The received signal was derived in Section 3.2 under assumptions A3.1 to A3.5. In the general case, spatial encoding is achieved by creating a phase distribution, ϕ, that is dependent on both time and space.

$$s(t) = \int_{\text{obj}} m(\boldsymbol{x}) e^{j\phi(\boldsymbol{x},t)} d\boldsymbol{x} \tag{3.17}$$

The phase is related to the encoding magnetic field by

$$\phi(\boldsymbol{x},t) = -\gamma \int_0^t b(\boldsymbol{x},\tau) d\tau. \tag{3.18}$$

The encoding field is the superposition of the field generated by each gradient coil. For three linear gradients shown in Figure 3.6 the total encoding field is

$$b(\boldsymbol{x},t) = g_x(t)x + g_y(t)y + g_z(t)z \tag{3.19}$$

For simplicity we let $\boldsymbol{g}(t) = \left[g_x(t), g_y(t), g_z(t)\right]^T$, thus the total field can be written as

$$b(\boldsymbol{x},t) = \boldsymbol{g}(t) \cdot \boldsymbol{x}, \tag{3.20}$$

where $\boldsymbol{a} \cdot \boldsymbol{b}$ denotes the inner product. Substituting into (3.18) the encoding phase

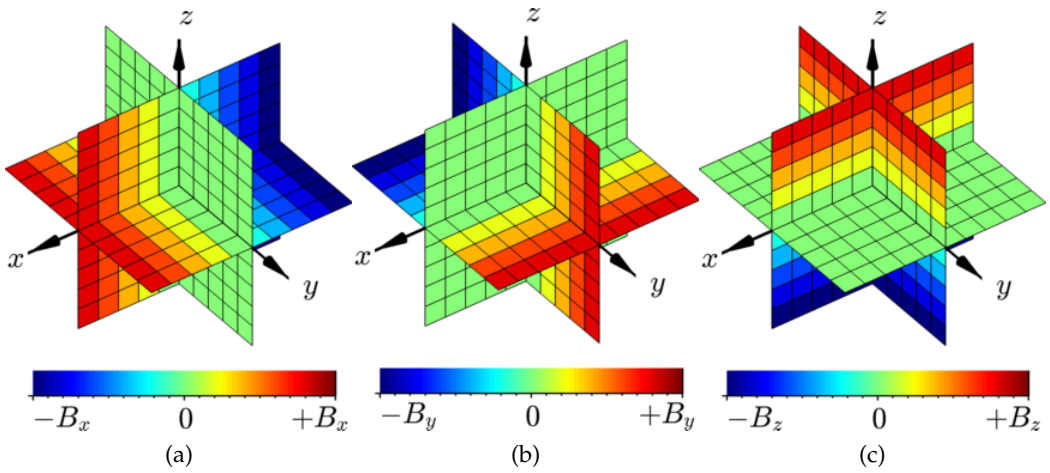


Figure 3.6: Linear encoding gradients for the three imaging dimensions illustrating the spatially-varying field offset imposed by each gradient. The maximum field of each gradient is given by $B = GL/2$ where G is the gradient strength (T/m) and L is the field of view (m).

is related to the gradient waveforms by

$$\phi(\boldsymbol{x}, t) = -\gamma \int_0^t \boldsymbol{g}(\tau) d\tau \cdot \boldsymbol{x}. \tag{3.21}$$

3.4.1 Slice selection

The first step in common two dimensional imaging sequences is to excite a thin slice of the object. This can be achieved by applying a gradient field and frequency selective RF pulse simultaneously. To select a slice in the x-y plane, the gradient field $b(z) = zG_z$ is first applied. The precession frequency, ω, of spins within the object becomes dependent on the z-coordinate,

$$\omega(z) = \omega_0 + \gamma G_z z \tag{3.22}$$

where ω_0 is the Larmor frequency due to the main magnetic field, G_z is the magnitude of the gradient change and z is the position along the z-axis.

After applying the gradient field the desired spins are excited using a bandlimited RF pulse oscillating at the required Larmor frequency. For example, an RF pulse resonating at $\omega(z_1)$, with a bandwidth of $\Delta\omega$ will excite a slice of the object between $z = z_1 - \Delta z/2$ and $z = z_1 + \Delta z/2$ where Δz is the slice thickness. Ideally, only spins in this slice will be rotated into the transverse plane and contribute to the received signal. In practice, however, the transverse magnetisation

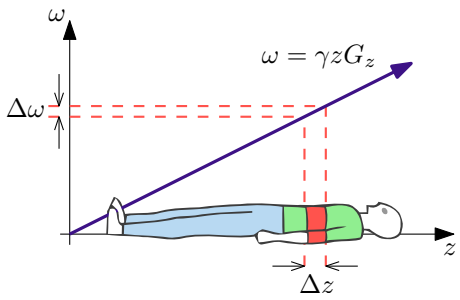


Figure 3.7: Slice selection diagram representing the spatially dependence frequency and an example slice.

follows a smooth slice profile and some "leakage" will occur, dependent on the RF pulse shape. Figure 3.7 illustrates the position dependent precession frequency and an example slice excited from a frequency selective RF pulse. Notice that the slice thickness is linearly related to the gradient strength (G_z) and bandwidth (Δw). This procedure can be repeated for different slices to encode the slice dimension. The remaining two dimensions can be encoded using a combination of frequency and phase encoding discussed below.

3.4.2 Frequency encoding

Frequency encoding is achieved by applying constant gradient over the entire time interval,

$$\boldsymbol{g}(t) = \boldsymbol{g}_{\mathrm{f}}, \quad t > 0 \tag{3.23}$$

The encoding phase in this case is

$$\phi(\boldsymbol{x}, t) = -\gamma \int_0^t \boldsymbol{g}(\tau) d\tau \cdot \boldsymbol{x} \tag{3.24}$$

$$= -\gamma t \boldsymbol{g}_{\mathrm{f}} \cdot \boldsymbol{x} \tag{3.25}$$

The received signal becomes

$$s(t) = \int_{\mathrm{obj}} m(\boldsymbol{x}) e^{-j\gamma t \boldsymbol{g}_{\mathrm{f}} \cdot \boldsymbol{x}} d\boldsymbol{x} \tag{3.26}$$

The signal can be interpreted as a 1D projection of the object, characterised by the gradient $\boldsymbol{g}_{\mathrm{f}}$. Intuitively, we can obtain different projections by altering the gradient and acquiring the signal again.

3.4.3 Phase encoding

Suppose a gradient field $\boldsymbol{g}_\mathrm{p}$ is briefly applied to the sample just after the RF pulse.

$$\boldsymbol{g}(t) = \begin{cases} \boldsymbol{g}_\mathrm{p}, & 0 \leq t \leq T_\mathrm{p}; \\ 0, & \text{otherwise.} \end{cases} \tag{3.27}$$

During the gradient duration ($0 \leq t \leq T_\mathrm{p}$) the received signal is frequency encoded as discussed above. For $t > T_\mathrm{p}$ the signal at every point returns to precession at ω_0 but has acquired a spatially dependent phase offset of

$$\phi(\boldsymbol{x}) = -\gamma T_\mathrm{p} \boldsymbol{g}_\mathrm{p} \cdot \boldsymbol{x}. \tag{3.28}$$

The acquired signal with phase encoding is

$$s(t) = \int_\mathrm{obj} m(\boldsymbol{x}) e^{-j\gamma T_\mathrm{p} \boldsymbol{g}_\mathrm{p} \cdot \boldsymbol{x}} d\boldsymbol{x}, \quad t > T_\mathrm{p}. \tag{3.29}$$

Notice the right hand side of (3.29) is independent of time but the magnetisation has been modulated with a phase characterised by the gradient, $\boldsymbol{g}_\mathrm{p}$, and the duration, T_p.

3.4.4 *k*-space

Frequency and phase encoding discussed above can be described in a common framework using the notion of *k*-space [30, 31]. The key point is to make the following substitution,

$$\boldsymbol{k}(t) = \bar{\gamma} \int_0^t \boldsymbol{g}(\tau) d\tau \tag{3.30}$$

where $\bar{\gamma} = \gamma / 2\pi$. The encoding phase in (3.21) becomes linear in both $\boldsymbol{k}$ and $\boldsymbol{x}$ with $\phi = -2\pi \boldsymbol{k} \cdot \boldsymbol{x}$ and the signal is written as

$$s(\boldsymbol{k}) = \int_\mathrm{obj} m(\boldsymbol{x}) e^{-j2\pi \boldsymbol{k} \cdot \boldsymbol{x}} d\boldsymbol{x} \tag{3.31}$$

This can be recognised as a multi-dimensional Fourier transform, where the spatial frequency domain is referred to as *k*-space. The definition in (3.30) highlights the important relationship between the applied gradient fields and the position in *k*-space. We can interpret frequency encoding discussed above as moving through *k*-space with a constant velocity whereas phase encoding can quickly move the *k*-space position a short distance. Importantly, different gradient waveforms lead to

different *trajectories* through k-space.

3.5 Basic sequences

In this section we examine some common sequences for two dimensional imaging. Prior to encoding, a gradient is applied in the z direction in combination with a frequency selective RF pulse to excite a slice of spins as described in Section 3.4.1. Thus spatial encoding is performed over the remaining two dimensions and k-space vectors are defined in $\mathbb{R}^2$. Sequences are defined by the timing of the gradient systems, RF pulses and the analog-to-digital converter (ADC).

3.5.1 Cartesian

The most common sampling strategies are based on a Cartesian trajectory. This type of trajectory starts with phase encoding using constant gradients for an initial period, T_p,

$$\boldsymbol{g}(t) = \begin{bmatrix} G_x \\ G_y \end{bmatrix}, \quad t < T_\mathrm{p} \tag{3.32}$$

Using (3.30) the k-space position after time T_p will be $\boldsymbol{k}_\mathrm{A} = \bar{\gamma} T_\mathrm{p} \left[G_x, G_x\right]^T$. To generate an echo a 180° pulse is applied, which inverts the phase of the signal. In terms of k-space, this is equivalent to the position being reflected about the origin so immediately after the pulse, $\boldsymbol{k}_\mathrm{B} = \bar{\gamma} T_\mathrm{p} \left[-G_x, -G_x\right]^T$. Frequency encoding is achieved by applying a constant g_x gradient during readout. This moves the trajectory along the k_x dimension of k-space as displayed in Figure 3.8a. During the readout period we have,

$$\boldsymbol{k}(t) = \bar{\gamma} \begin{bmatrix} (t - \tau_\mathrm{E}) G_x \\ G_y T_\mathrm{p} \end{bmatrix}, \quad |t - \tau_\mathrm{E}| < T_\mathrm{acq}/2 \tag{3.33}$$

The k_y gradient moment is set up during the initial phase encode and is constant during readout. The readout gradient, G_x, is constant for all excitations causing repetitive evolution of the k_x coordinate. The process is repeated for different values of G_y to define different horizontal lines across the k_x-k_y plane as shown in Figure 3.8b. Figure 3.9 displays a sequence diagram for the Cartesian trajectory, which illustrates the timing and function of the different subsystems.

Cartesian trajectories are popular since they are relatively insensitivity to field inhomogeneities and gradient nonlinearities [32]. Further, images are easy to re-

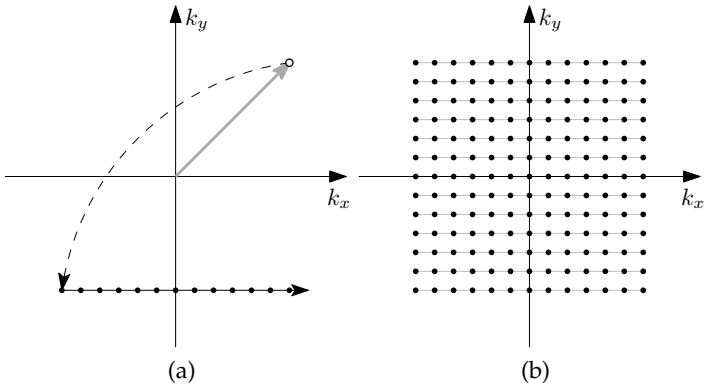


Figure 3.8: The *k*-space sampling pattern for a Cartesian spin echo trajectory illustrating (a) a single line consisting of a phase encode, refocusing pulse and readout; and (b) the complete sampling pattern from multiple excitations.

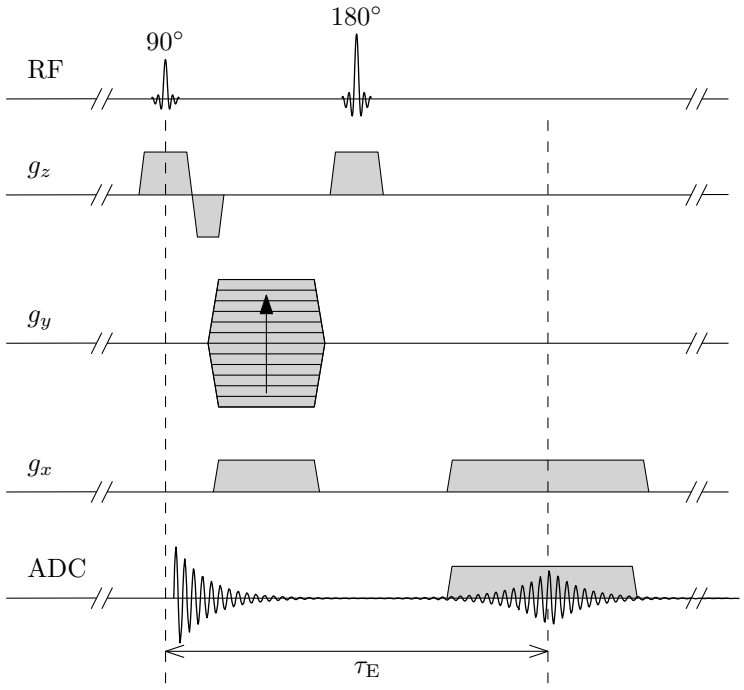


Figure 3.9: Sequence diagram for a Cartesian spin echo trajectory illustrating a single excitation, phase encode, refocusing pulse and readout.

construct using the discrete Fourier transform (DFT) as discussed in Section 3.6.

3.5.2 Radial

A simple radial trajectory begins much like the Cartesian trajectory described above. Specifically, a phase encoding gradient is applied such that the k-space position after time T_p will be $\boldsymbol{k}_\mathrm{A} = \bar{\gamma} T_\mathrm{p} \left[G_x, G_y\right]^T$ as described above. A refocusing pulse is applied to move k-space position to $\boldsymbol{k}_\mathrm{B} = -\boldsymbol{k}_\mathrm{A}$. Samples are collected when the constant gradients are switched on again in the period $\tau_\mathrm{E} - T_\mathrm{acq}/2 \leq t \leq \tau_\mathrm{E} + T_\mathrm{acq}/2$. The main difference from a Cartesian trajectory is that both gradients (g_x and g_y) are applied during the readout. The k-space trajectory during this period is

$$\boldsymbol{k}(t) = \boldsymbol{k}_\mathrm{B} + \bar{\gamma}(t - (\tau_\mathrm{E} - T_\mathrm{acq}/2)) \begin{bmatrix} G_x \\ G_y \end{bmatrix}, \quad |t - \tau_\mathrm{E}| < T_\mathrm{acq}/2 \tag{3.34}$$

The phase encoding defining point $\boldsymbol{k}_\mathrm{B}$ is only dependent on G_x, G_y and T_p, not time during the readout. This reveals the important property that phase encoding only affects the starting point of the k-space trajectory.

With careful selection of the phase encoding duration such that $T_\mathrm{p} = T_\mathrm{acq}/2$, we can reparameterise (3.34) as

$$\boldsymbol{k}(t) = C(t - \tau_\mathrm{E}) \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \quad |t - \tau_\mathrm{E}| < T_\mathrm{acq}/2 \tag{3.35}$$

where $\theta = \arctan(G_y/G_x)$ is the angle of the trajectory and $C = \bar{\gamma}\sqrt{G_x^2 + G_y^2}$ defines the trajectory speed. This process is illustrated in Figure 3.10a. Sufficient sampling of k-space is achieved by repeating the above process for different angles, θ, by defining appropriate values of G_x and G_y. The complete sampling pattern in shown in Figure 3.10b. Figure 3.11 displays a sequence diagram for the radial trajectory, which illustrates the timing of the gradient, RF and ADC systems.

Radial sampling in MRI is a direct analog to standard x-ray imaging and consequently it inherited a broad range of the prior knowledge concerning image reconstruction. Indeed the first MR images were produced using radial projections [33]. In general, however, the reconstruction algorithm is more complex than Fourier reconstruction of Cartesian samples. Another difficulty is that off-resonance effects due to field inhomogeneities create blurring along the radial direction whereas Fourier imaging is affected by shifts in the readout direction resulting in geometric distortion [32, 34, 35]. Despite these shortcomings, radial imaging remains use-

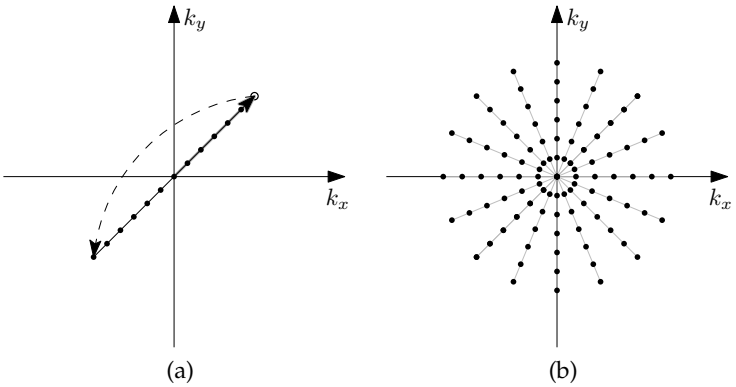


Figure 3.10: The *k*-space sampling pattern for a radial trajectory illustrating (a) a single spoke consisting of a phase encode, refocusing pulse and readout; and (b) the complete sampling pattern from multiple excitations.

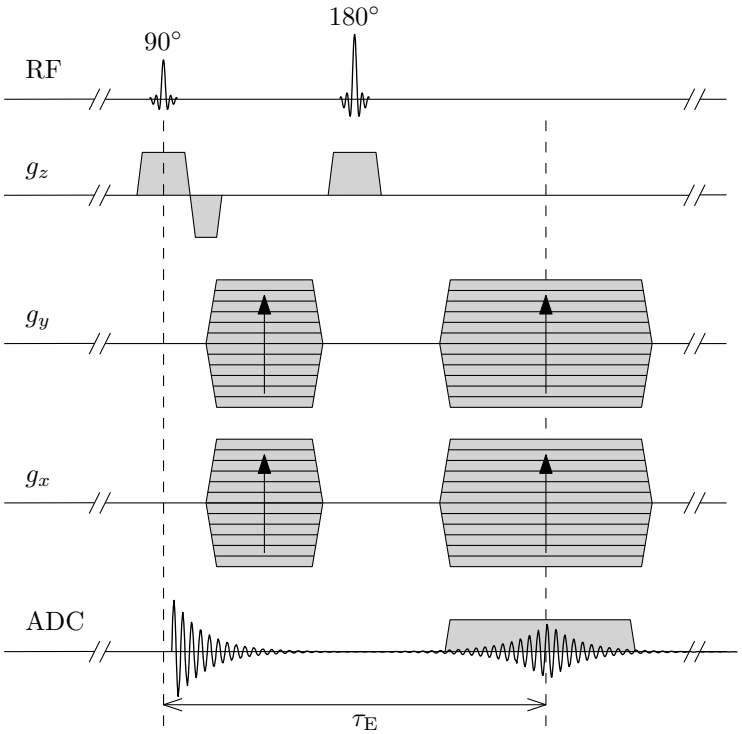


Figure 3.11: Sequence diagram for a radial trajectory illustrating a single excitation, phase encode, refocusing pulse and readout.

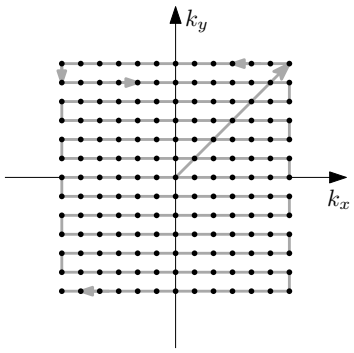


Figure 3.12: The *k*-space sampling pattern for a Cartesian EPI trajectory illustrating the initial phase encode and rapid trajectory through *k*-space from a single excitation.

ful for specialised applications. Since the readout can start immediately after the slice select gradient, this technique is useful to image tissue with short T_2 or T_2^*. Another advantage of radial imaging is the relative insensitivity to subject motion making it useful for dynamic imaging applications [36].

3.5.3 EPI

Echo planar imaging (EPI) was first introduced in [37] and later modified to produce Cartesian trajectories [38, 39]. Although the final sampling trajectory is similar to the Cartesian spin echo in Figure 3.8b, the underlying mechanism is different. Firstly, the trajectory is acquired with a single excitation and belongs to class of imaging techniques referred to as "single-shot". Secondly, the echoes are generated by switching the gradient polarity rather than a 180° refocusing pulse.

A simple EPI trajectory is shown in Figure 3.12, generated from the sequence diagram in Figure 3.13. The trajectory begins by exciting a slice of the object to rotate the magnetisation through a flip angle, α. As described above, prephasing gradients are applied to move the trajectory from the centre to the edge of the sampled *k*-space. Next, a series of echoes is generated by rapidly switching the polarity of the readout gradient. This moves the *k*-space trajectory back and forth along the horizontal direction. At the end of each line a brief phase encode gradient is applied to move the trajectory a small amount in the vertical direction in preparation for the next readout line. The final result is a Cartesian *k*-space sampling pattern acquired in under 50 ms, orders of magnitude faster than traditional spin echo techniques [39]. EPI imaging is insensitive to subject motion since the acquisition occurs on a shorter time scale than typical motion [40]. Disadvantages include an increased risk of peripheral nerve stimulation due to the rapidly switching gra-

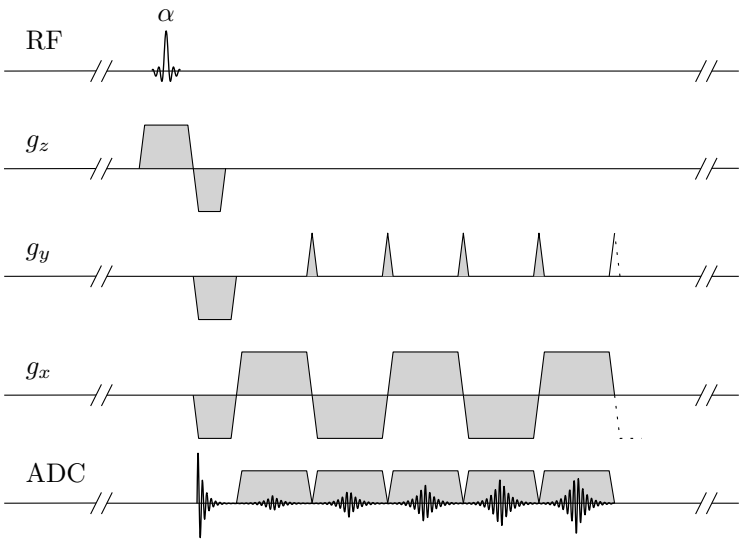


Figure 3.13: Sequence diagram for a Cartesian EPI trajectory illustrating the timing of the single excitation pulse followed by phase encoding and readout gradients.

dients [41, 42] and sensitivity to magnetic field inhomogeneity and the associated dephasing [43].

The EPI trajectory will be important in Chapter 6 and is the basis of our trajectory design for higher dimensional encoding schemes.

3.6 Image reconstruction

Image reconstruction is the process of generating an image from the measured MR signal. In the simplest case, the image reconstruction problem can be described as follows,

$$\begin{aligned} &\text{Find} \quad m(\boldsymbol{x}) \\ &\text{Given} \quad s(\boldsymbol{k}_i) = \int_{\text{obj}} m(\boldsymbol{x}) e^{-j2\pi \boldsymbol{k}_i \cdot \boldsymbol{x}} d\boldsymbol{x}, \quad \boldsymbol{k}_i \in \mathcal{K} \end{aligned} \tag{3.36}$$

The difficulty of this problem is revealed when we consider the preceding discussion regarding the inherent tradeoffs in MRI data.

The remainder of this section describes some commonly used reconstruction algorithms.

3.6.1 Direct Fourier

We consider the reconstruction of a 1D function to illustrate the method, although the theory is easily extended to 2D or 3D functions. Fourier reconstruction using

the DFT is suitable when samples are collected at a finite number of regularly spaced k-space locations [44]. That is,

$$\mathcal{K} = \{i\Delta k,\ -N/2 \leq i \leq N/2 - 1\} \tag{3.37}$$

Let $s[i] = s(i\Delta k)$ denote the measurements. A common reconstruction is one that satisfies the minimum-norm constraint. In this case the image is reconstructed using a truncated Fourier series,

$$\tilde{m}(\boldsymbol{x}) = \Delta k \sum_{i=-N/2}^{N/2-1} s[i] e^{j2\pi i \Delta k x}, \quad |x| < \frac{1}{\Delta k} \tag{3.38}$$

The reconstructed image function, $\tilde{m}(\cdot)$, is inherently bandlimited due to the truncation of its Fourier series. That is, $(\mathcal{F}\tilde{m})(k) = 0$ for $|k| > (N/2)\Delta k$. Nyquist sampling criterion states that this image function can be uniquely recovered from samples, $\tilde{m}[n] = \tilde{m}(n\Delta x)$, provided the pixel size, Δx, satisfies

$$\Delta x \leq \frac{1}{N\Delta k} \tag{3.39}$$

Sampling (3.38) at the Nyquist limit gives,

$$\tilde{m}[n] = \Delta k \sum_{i=-N/2}^{N/2-1} s[i] e^{j2\pi i n/N} \tag{3.40}$$

This can be recognised as the DFT of the measurements and as such this technique is called the direct DFT reconstruction.

3.6.2 Gridding

Gridding is used when the set $\mathcal{K}$ of k-space points are not regularly spaced in Cartesian coordinates. A simple approach in this case is to interpolate the measured points into a regular grid and then perform the direct DFT reconstruction described above. Let $\tilde{s}(\boldsymbol{k})$ be a sum of delta functions weighted by the measurements,

$$\tilde{s}(k) = \sum_{k_i \in \mathcal{K}} s(k_i)\delta(k - k_i) \tag{3.41}$$

These samples are convolved with an interpolation function, C, and sampled at regularly spaced points defined in the set $\mathcal{L}$,

$$\bar{s}[i] = (\tilde{s} * C)(k_i), \quad k_i \in \mathcal{L} \tag{3.42}$$

This sequence can now be reconstructed using a DFT.

3.6.3 Iterative

An alternative approach to image reconstruction is to formulate the problem as a matrix inversion. The signal equation is written as

$$\boldsymbol{s} = \boldsymbol{E}\boldsymbol{m} \tag{3.43}$$

where $\boldsymbol{s}$ is a vector of measurements, $\boldsymbol{E}$ is the encoding matrix and $\boldsymbol{m}$ is a vector of unknown magnetisations. This formulation allows us to specify an arbitrary encoding scheme, which may incorporate multiple receive coils, non-Cartesian trajectories or field inhomogeneity correction. Including these features into a Fourier reconstruction is not straightforward.

Reconstruction of a 256×256 image using fully sampled data from a 32 channel receive coil array has an encoding matrix with $\sim 10^{11}$ complex-valued elements. For double precision this would require approximately 2200 GB of memory. In addition to practical issues concerning memory, a direct inverse requires $\mathcal{O}(N^3)$ computations where $N = 256 \times 256$. Even with the dramatic increase in computing power of modern computers, a direct matrix inversion of (3.43) is infeasible for clinical applications. For these reasons we consider two iterative algorithms to solve (3.43). Both methods only require a single row of the encoding matrix at any given time and therefore avoid storing entire matrix in memory.

The first algorithm examined is the Kaczmarz iterative projection algorithm [45] otherwise known as the Algebraic Reconstruction Technique (ART) [46]. The algorithm is described in Algorithm 3.1. This method has been proven to converge to the minimum norm least squares solution for sufficiently small step sizes [47]. Further, it is relatively simple to include prior information into the reconstruction algorithm. In the case of Gaussian prior and noise distributions, the Kaczmarz algorithm converges to a regularised solution equivalent to the Bayesian maximum a posteriori (MAP) estimate [47].

Another iterative method is the conjugate gradient (CG) algorithm [48]. This algorithm was used in [49] for reconstruction of parallel imaging data with arbitrary trajectories. Instead of solving (3.43) directly the CG algorithm solves $\boldsymbol{E}'\boldsymbol{s} = \boldsymbol{E}'\boldsymbol{E}\boldsymbol{m}$.

Algorithm 3.1: Algebraic Reconstruction Technique

Input: $\boldsymbol{s}$ – signal, $\boldsymbol{E}$ – encoding matrix, N – number of measurements, P – number of iterations
Params: λ – step size
Output: $\hat{\boldsymbol{m}}$ – estimate

for $n \leftarrow 0, \ldots, P$ **do**
  for $i \leftarrow 1, \ldots, N$ **do**
    Assign the i^{th} row of $\boldsymbol{E}$ to $\boldsymbol{a}_i$
    $$\hat{\boldsymbol{m}}_{n+1} \leftarrow \hat{\boldsymbol{m}}_n + \lambda \frac{s_i - \langle \boldsymbol{a}_i, \hat{\boldsymbol{m}}_n \rangle}{\|\boldsymbol{a}_i\|^2} \boldsymbol{a}_i^*$$
Return estimate: $\hat{\boldsymbol{m}} \leftarrow \hat{\boldsymbol{m}}_P$

The algorithm is described in Algorithm 3.2. The algorithm only requires a single row of the encoding matrix at a time and has been proven to converge to the pseudo-inverse solution in a finite number of iterations [50].

Both the conjugate gradient algorithm and algebraic reconstruction technique are used extensively to reconstruct data resulting from the nonlinear encoding schemes examined in Chapters 7 and 6.

3.7 Properties of MRI signals

The reconstruction algorithms in Section 3.6 were developed without considering measurement noise. In reality, measurements are corrupted with thermal noise, which ultimately limits the imaging performance. Indeed the term 'reconstruction' should perhaps be replaced with 'estimation', since we can only extract information from the measurements up to the level of noise. This section explores some of the ramifications of the stochastic nature of the measurements.

3.7.1 Signal-to-noise ratio

The thermal noise variance, $\sigma^2_{\text{thermal}}$, in an imaging experiment is given by the following relationship, first observed by Johnson [51] and explained by Nyquist [52],

$$\sigma^2_{\text{thermal}} = 4K_{\text{B}}\, T_{\text{s}}\, R\, \text{BW} \tag{3.44}$$

where K_B is Boltzmann's constant, T_{s} is the system temperature, R is the resistance and BW is the acquisition bandwidth.

Algorithm 3.2: Conjugate Gradient Reconstruction

Input: $\boldsymbol{s}$ – signal, $\boldsymbol{E}$ – encoding matrix, P – number of iterations

Output: $\hat{\boldsymbol{m}}$ – estimate

$\hat{\boldsymbol{m}}_0 \leftarrow \boldsymbol{0}$

$\boldsymbol{p} \leftarrow \boldsymbol{E}\boldsymbol{s}$

$\boldsymbol{r}_0 \leftarrow \boldsymbol{E}\boldsymbol{s}$

for $n \leftarrow 0, \ldots, P$ **do**

$$\boldsymbol{q} \leftarrow \boldsymbol{E}'\boldsymbol{E}\boldsymbol{p}$$

$$\hat{\boldsymbol{m}}_{n+1} \leftarrow \hat{\boldsymbol{m}}_n + \frac{\boldsymbol{r}_n^* \boldsymbol{r}_n}{\boldsymbol{p}^* \boldsymbol{q}} \boldsymbol{p}$$

$$\boldsymbol{r}_{n+1} \leftarrow \boldsymbol{r}_n - \frac{\boldsymbol{r}_n^* \boldsymbol{r}_n}{\boldsymbol{p}^* \boldsymbol{q}} \boldsymbol{q}$$

$$\boldsymbol{p} \leftarrow \boldsymbol{r}_{n+1} + \frac{\boldsymbol{r}_{n+1}^* \boldsymbol{r}_{n+1}}{\boldsymbol{r}_n^* \boldsymbol{r}_n} \boldsymbol{p}$$

Return estimate: $\hat{\boldsymbol{m}} \leftarrow \hat{\boldsymbol{m}}_P$

The signal strength, s, is proportional to the square of the main magnetic field strength, i.e., $s \propto B_0^2$. However, it has been found that the resistance in (3.44) is frequency dependent, which results in the relation, $R \propto B_0$ [53]. Thus the signal-to-noise ratio (SNR), defined as $\mathsf{SNR} = s/\sigma_{\text{thermal}}$, has the following relationship to field strength [54],

$$\mathsf{SNR} \propto B_0 \tag{3.45}$$

The relationships in (3.44) and (3.45) mean the SNR for Fourier imaging is intrinsically linked to the acquisition parameters [53]. In particular,

$$\mathsf{SNR/voxel} \propto \frac{B_0 \, \Delta_x \Delta_y \Delta_z \sqrt{N_{\text{avg}}}}{\sqrt{\frac{\mathsf{BW}}{N_x N_y N_z}}} \tag{3.46}$$

where Δ_x, Δ_y and Δ_z define the voxel dimensions, N_x, N_y and N_z are the number of voxels in each dimension and N_{avg} is the number of signal averages. Since $\mathsf{BW} = 1/T_{\text{acq}}$ the SNR has the following relationship to the readout parameters,

$$\mathsf{SNR/voxel} \propto B_0 \, \Delta_x \sqrt{T_{\text{acq}}}. \tag{3.47}$$

The final SNR of an image is further influenced by physiological noise, systematic errors and reconstruction algorithm.

3.7.2 Signal processing challenges

The difficulty of many MRI problems is revealed when we consider the ideal solution from a users perspective. For both image reconstruction and parameter estimation problems, the user wants:

- fast acquisition times
- high resolution
- high SNR
- no image artifacts or bias

These are conflicting requirements for MRI data. For example, a fast acquisition time is often achieved by obtaining fewer measurements. On one hand, acquisition is limited by the signal relaxation, on other, gradient switching constraints are imposed due to hardware limitations and safety concerns regarding peripheral nerve stimulation. These constraints limit the amount of k-space that can be covered in a fixed time; conversely, reducing the imaging time is achieved with fewer k-space measurements.

Increasing the resolution of the reconstruction is equivalent to seeking a higher dimensional approximation of the underlying function. Processing a very high dimensional signal is difficult, even with the increase in modern computing power. When the number of measurements is limited, this reconstruction requires some form of prior knowledge for stable signal recovery in the presence of noise.

Prior knowledge is used to constrain the reconstruction, which is the basis of many state-of-the-art techniques, such as high resolution MR angiography [55] and dynamic cardiac imaging [56]. Although helpful, too much emphasis on prior knowledge leads to reconstruction artifacts or bias.

Tradeoffs are pervasive and cannot be avoided. Some typical tradeoffs and are summarised below:

$$\begin{array}{ccc} \uparrow \text{resolution} & \Leftrightarrow & \uparrow \text{signal dimension} \\ \downarrow \text{acquisition time} & \Leftrightarrow & \downarrow \text{number of measurements} \\ \uparrow \text{SNR} & \Leftrightarrow & \downarrow \text{resolution} \\ \downarrow \text{bias} & \Leftrightarrow & \downarrow \text{reconstruction constraints} \end{array}$$

With this in mind, the ideal requirements above translate to the following signal processing problem: *Reconstruct a very large dimensional function from a small number of noisy measurements without any prior knowledge about the function.* Clearly, no solution will be adequate for a realistic class of functions. We are forced find a compromise between requirements that is suitable for a particular application.

The difficulty of the problem necessitates the need for accurate signal models that capture the deterministic effects on the measurements. Only with precise models can the desired quantities be distinguished from random noise. This estimation-theoretic perspective is adopted for the remainder of this thesis and we will see that it leads to new techniques that push the boundaries of current MRI technology.

Part II

Quantitative MRI

CHAPTER 4

Estimation of relaxation rates in the presence of image distortion

Contents

4.1 Introduction

TRADITIONAL MRI generates image contrast dependent on a number of parameters, which may include the longitudinal relaxation time T_1, transverse relaxation time T_2, free induction decay time T_2^* or the initial magnetisation M_0. In quantitative MRI, the aim is to estimate the values of these parameters at every point in the object. The parameter values are useful to determine the underlying

structure of the object being imaged, which might not be evident using contrast alone.

In this chapter, we look at estimation of the spin-spin relaxation time, denoted T_2. Accurate estimation of T_2 values has been important for tissue classification, disease detection and pathology [57]. The underlying T_2 values of brain tissue are of particular interest to study white matter diseases such as multiple sclerosis [58]. Similar analyses of T_2^* values is used to quantify iron content within tissue to study many neurological conditions including Parkinson's disease and Huntington's disease (see [59] for a complete review).

Traditionally, a series of images is acquired after the initial excitation using a multi-echo spin-echo sequence suitable for T_2 estimation. A common spin-echo sequence for T_2 estimation was introduced by Carr and Purcell [28] and later modified by Meiboom and Gill [29]. The aptly named CPMG sequence is considered the gold standard for T_2 estimation.

The T_2 relaxation rate for each pixel is conventionally determined by a least squares fit of an exponential decay to the image intensities. A simple estimator such as this assumes that the intensity of each pixel depends only on the parameters at that pixel, and estimates its relaxation time separately [4]. However, the acquisition process used to generate MR images violates this assumption. Signal decay occurs while samples are being acquired, leading to spatial distortion (blurring) of the image across pixels [3]. This distortion leads to statistical bias of the traditional estimator that can be minimised by acquiring data samples over a small acquisition period. However, this comes at the expense of increased noise, creating a trade-off between SNR and bias. In this chapter, we investigate the statistical bias of the traditional estimator and propose a Bayesian estimation algorithm derived by including the distortion in the signal model.

The chapter is organised as follows: Section 4.2 analyses the image distortion due to signal decay. Section 4.3 presents a complete measurement model for relaxation time estimation incorporating signal distortion. Section 4.4 investigates the bias associated with naïve estimators. Section 4.5 presents the proposed estimation algorithm and Section 4.6 applies the algorithm to simulated data.

4.1.1 Notation

The important quantities and associated notation used in this chapter are list in Table 4.1. Additional notation will be introduced as required.

Table 4.1: Important quantities and notation used in Chapter 4

Symbol	Quantity
v, e	Noise
$f(\boldsymbol{k}, \boldsymbol{x})$	Relaxation terms
$h(\boldsymbol{x}, \boldsymbol{z})$	Spatial kernel
q, i	Excitation and time index (k-space)
n	Echo index
r, p	Pixel index
m	Initial magnetisation
$\zeta_1, \ldots, \zeta_G$	Posterior grid points
$w_1, \ldots, w_G$	Weights of gridded posterior
Δ_G	Grid spacing
$\boldsymbol{\Sigma}, \boldsymbol{\Lambda}$	Noise covariance and prior covariance

4.2 Image distortion due to relaxation

The acquired signal including relaxation (in the absence of measurement noise) was derived in Section 3.3 and can be written as

$$s(\boldsymbol{k}) = \int m(\boldsymbol{x}) f(\boldsymbol{k}, \boldsymbol{x}) e^{j\boldsymbol{k}\cdot\boldsymbol{x}} d\boldsymbol{x} \tag{4.1}$$

where $\boldsymbol{k}$ and $\boldsymbol{x}$ denote the k-space and position coordinates, respectively. The quantity $f(\boldsymbol{k}, \boldsymbol{x})$ represents the relaxation component and is a function of both position and k-space location. For a spin echo sequence it can be written as

$$f(\boldsymbol{k}(t), \boldsymbol{x}) = e^{-t/T_2(\boldsymbol{x})} e^{-|t-\tau_E|/T_2'(\boldsymbol{x})} \tag{4.2}$$

The dependence of the k-space trajectory on time has been made explicit by the notation $\boldsymbol{k}(t)$. As described in Chapter 2, it is useful to define the free induction decay time, T_2^*, to describe the aggregate effect of different relaxation processes,

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_2'} \tag{4.3}$$

The relevant decay curves and the measurement samples are illustrated in Figure 4.1. Notice that measurements are collected as the signal undergoes relaxation.

The image used for subsequent estimation is typically reconstructed using the discrete Fourier transform (DFT). This yields an image, $\tilde{m}(\cdot)$, which, in general, is not equivalent to the desired image. To analyse this in detail we first derive the Discrete Time Fourier Transform (DTFT) of the received signal and later sample

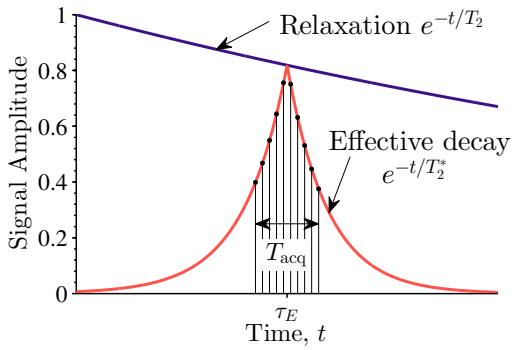


Figure 4.1: Signal amplitude of the bulk magnetisation during data acquisition for a spin echo sequence.

it to obtain the image reconstructed from the DFT. Note that 'time' in our case is actually a spatial variable.

$$\tilde{m}(\boldsymbol{z}) = \sum_i s(\boldsymbol{k}_i) e^{-j\boldsymbol{k}_i \cdot \boldsymbol{z}} \tag{4.4}$$

We include the expression for $s(\boldsymbol{k})$ in (4.1) to give,

$$\tilde{m}(\boldsymbol{z}) = \sum_i \int m(\boldsymbol{x}) f(\boldsymbol{k}_i, \boldsymbol{x}) e^{j\boldsymbol{k}_i \cdot \boldsymbol{x}} e^{-j\boldsymbol{k}_i \cdot \boldsymbol{z}} d\boldsymbol{x} \tag{4.5}$$

$$= \int m(\boldsymbol{x}) \sum_i f(\boldsymbol{k}, \boldsymbol{x}) e^{j\boldsymbol{k}_i \cdot (\boldsymbol{x} - \boldsymbol{z})} d\boldsymbol{x} \tag{4.6}$$

In the equations above $\boldsymbol{x}$ and $\boldsymbol{z}$ represent spatial variables while $\boldsymbol{k}_i$, $i = 1, \ldots, N$ are the spatial frequency or k-space samples.

4.2.1 Spatial filter interpretation

We can write (4.6) as a filter as follows,

$$\tilde{m}(\boldsymbol{z}) = \int m(\boldsymbol{x}) h(\boldsymbol{x}, \boldsymbol{z}) d\boldsymbol{x} \tag{4.7}$$

where the filter kernel $h(\boldsymbol{x}, \boldsymbol{z})$ is defined as

$$h(\boldsymbol{x}, \boldsymbol{z}) = \sum_i f(\boldsymbol{k}_i, \boldsymbol{x}) e^{j\boldsymbol{k}_i \cdot (\boldsymbol{x} - \boldsymbol{z})} \tag{4.8}$$

In this framework, the image is interpreted as a filtered version of the initial magnetisation $m(\boldsymbol{z})$ with the linear space-varying filter defined by $h(\boldsymbol{x}, \boldsymbol{z})$. An analysis of the kernel expression in (4.8) reveals the close links to the well-known

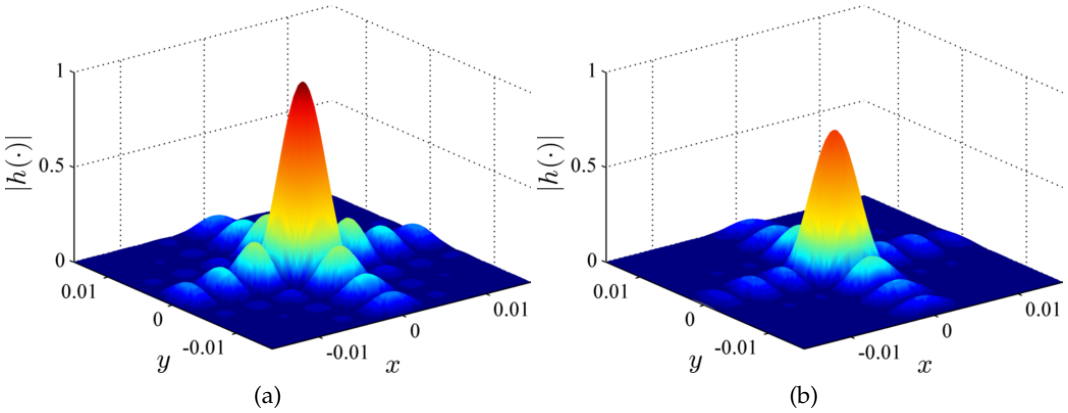


Figure 4.2: The 2D continuous kernel function for a single point with $T_2 = 100\,\text{ms}$ and $T_2' = 20\,\text{ms}$ for an acquisition time of (a) $T_{\text{acq}} = 1\,\text{ms}$ and (b) $T_{\text{acq}} = 25\,\text{ms}$. The x and y axes are in arbitrary units.

Dirichlet kernel common in Fourier theory. Appendix 4.A provides an introduction to the Dirichlet kernel and its properties. In Appendix 4.B we derive a closed form expression for the summation in (4.8), which greatly reduces the computation time in the proposed estimation algorithm. The general form is

$$h(\boldsymbol{x}, \boldsymbol{z}) = e^{-\tau_E / T_2(\boldsymbol{x})} \tilde{D}_N(u_y) \left[g_1(u_x) + g_2(u_x) \right] \tag{4.9}$$

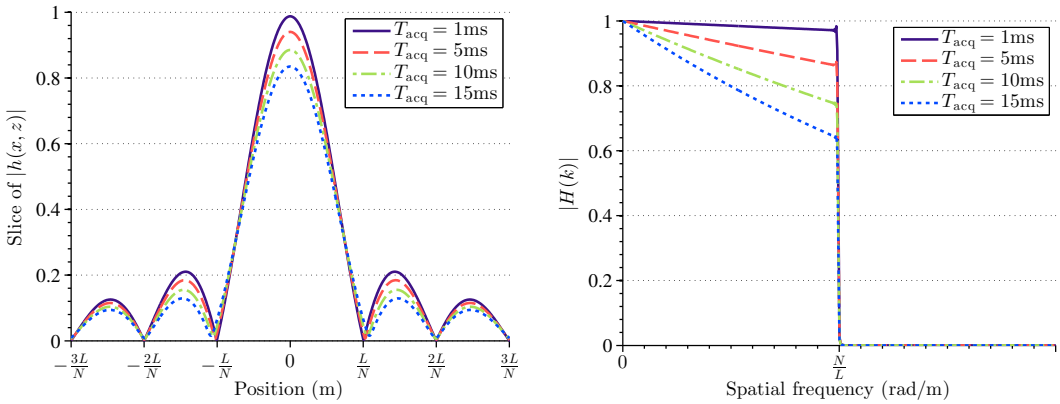
where u_x and u_y are the x and y coordinates of the difference, $\boldsymbol{x} - \boldsymbol{z}$, $\tilde{D}_N$ is a modified Dirichlet kernel and g_1 and g_2 are functions dependent on the readout parameters and relaxation parameters.

The closed-form kernel expression contains the multiplication of two decoupled components corresponding to each spatial dimension: a Dirichlet kernel in the phase encode direction, and a relaxation-dependent kernel in the readout direction.

Neglecting relaxation effects contained in f the kernel reduces to the Dirichlet kernel for both the readout and phase dimensions. The reconstructed function, $\tilde{m}$ is then the N^{th} degree Fourier series approximation to the underlying image, m.

Examples of the two dimensional kernel are shown in Figure 4.2. For short acquisition times, the kernel exhibits a strong peak and therefore minimal image distortion. Conversely, when the acquisition time is relatively long compared to the effective relaxation time, the kernel is attenuated and broader, which results in considerable image blurring.

Since the phase encoding does not depend on relaxation, the effect of the kernel on this dimension is shift invariant and independent of the relaxation parameters.



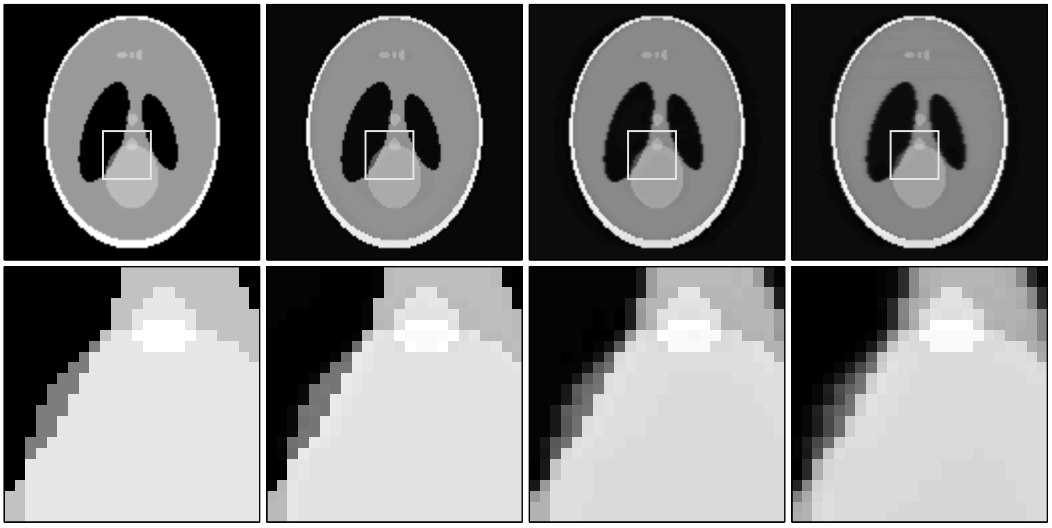
(a) (b)

Figure 4.3: Analysis of the filter kernel for different acquisition times displaying (a) the continuous filter kernel and (b) the corresponding frequency response. A slice along the readout direction is displayed for $T_2 = 100\,\text{ms}$ and $T_2' = 20\,\text{ms}$.

The filter in this direction is a Dirichlet kernel and its effect is well characterised by Fourier theory. Conversely the filter kernel along the readout direction is dependent on relaxation and creates an additional distortion. The purpose of this section is to investigate such a distortion. It is therefore constructive to focus our analysis on a 1D slice of the kernel along the readout direction.

We analyse the full kernel expression from a filtering perspective by calculating the frequency response of the windowed kernel at a location with $T_2 = 100\,\text{ms}$ and $T_2' = 20\,\text{ms}$ for different sampling rates. Figure 4.3 presents the spatial kernel and the corresponding frequency response for different acquisition times. The low pass nature of the filter is clear whereby the high frequency content is increasingly attenuated for larger acquisition times. In the spatial domain, attenuation of high frequencies corresponds to image blurring.

We investigate the low-pass characteristics of the filter by simulating the acquisition and reconstruction of a Shepp-Logan numerical phantom [60]. As the acquisition and sampling time decrease, $f(\boldsymbol{k}, \boldsymbol{x}) \to e^{-\tau_E / T_2(\boldsymbol{x})}$, and the image is a weighted $M_0(\cdot)$ as in (4.13). The effect of a non-zero sampling interval on reconstructed images is investigated in the Figure 4.4. To quantify the level of distortion, the mean square error (MSE) between the output image and the true image is plotted for different acquisition times as shown in Figure 4.5.



(a) Weighted image (b) $T_{\text{acq}} = 5\text{ms}$ (c) $T_{\text{acq}} = 15\text{ms}$ (d) $T_{\text{acq}} = 25\text{ms}$

Figure 4.4: Reconstructed images (top) and a region-of-interest (bottom) from simulated data using the linear filter model for increasing acquisition times. For long acquisition times considerable image blurring is present.

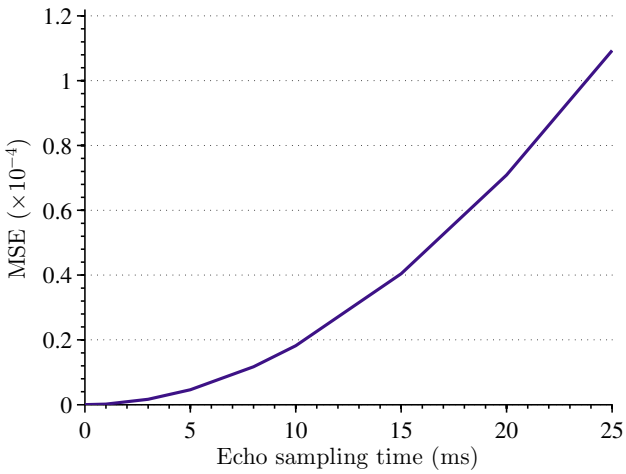


Figure 4.5: The mean squared error (MSE) of a reconstructed phantom as a function of acquisition time.

4.2.2 Filter kernel approximation

A simple approximation to the kernel function is obtained by introducing two assumptions:

A4.1 No relaxation occurs during the sampling of the observations, $s(\boldsymbol{k})$. This is valid for T_{acq} much smaller than the effective decay time. In this case $t \approx \tau_E$ and only the peak of the echo is sampled or in terms of the filter kernel, $f(\boldsymbol{k}(t), \boldsymbol{x}) \approx f(\boldsymbol{k}(\tau_E), \boldsymbol{x})$.

A4.2 The number of measurements is large such that the Dirichlet kernel approaches a delta function, $D_N(\boldsymbol{x}) \approx \delta(\boldsymbol{x})$.

For spin echo, these two assumptions result in a kernel

$$h(\boldsymbol{x}, \boldsymbol{z}) = e^{-\tau_E / T_2(\boldsymbol{z})} \delta(\boldsymbol{x} - \boldsymbol{z}) \tag{4.10}$$

Substituting this into the signal model gives,

$$\begin{aligned} \tilde{m}(\boldsymbol{z}) &= \int m(\boldsymbol{x}) h(\boldsymbol{x}, \boldsymbol{z}) d\boldsymbol{x} && (4.11) \\ &= \int m(\boldsymbol{x}) e^{-\tau_E / T_2(\boldsymbol{x})} \delta(\boldsymbol{x} - \boldsymbol{z}) d\boldsymbol{x} && (4.12) \\ &= m(\boldsymbol{z}) e^{-\tau_E / T_2(\boldsymbol{z})} && (4.13) \end{aligned}$$

Therefore, with this approximation we reconstruct a spatially weighted version of the initial magnetisation.

4.3 Measurement model for relaxation time estimation

A reasonable assumption for biological tissues (where $T_1 > T_2$) using acquisition sequences with long repetition times, is that the effect of T_1 relaxation is negligible. The unknown parameters depend on the specific imaging paradigm. For spin echo sequences the relevant parameters at the p^{th} object position are $\boldsymbol{\theta}_p = [T_2(\boldsymbol{x}_p), T_2'(\boldsymbol{x}_p), M_0(\boldsymbol{x}_p)]^T$ where $\boldsymbol{x}_p$ is a point on a $N_x \times N_y$ grid over the object. The vector of all important parameters is

$$\boldsymbol{\theta} = [\boldsymbol{\theta}_1^T, \ldots, \boldsymbol{\theta}_N^T]^T, \quad N = N_x \times N_y. \tag{4.14}$$

Estimation of the relaxation parameters is performed by first acquiring a sequence of M images at different echo times. As discussed in Chapter 3 the CPMG

sequence generates multiple echoes from a single excitation, which allows such data to be collected efficiently.

We wish to consider inversion of the signal model in (4.7). For this purpose it is useful to approximate the underlying signal by a sum of delta functions,

$$m(\boldsymbol{x}) \approx \sum_p m(\boldsymbol{x}_p)\delta(\boldsymbol{x} - \boldsymbol{x}_p) \tag{4.15}$$

We also recognise that the DFT used for image reconstruction produces discrete samples of the function, $\tilde{m}(\cdot)$, on the reconstruction grid, $z_1 \ldots, z_N$. The signal model now becomes

$$\tilde{m}(\boldsymbol{z}_r) = \sum_p m(\boldsymbol{x}_p)h(\boldsymbol{x}_p, \boldsymbol{z}_r) \tag{4.16}$$

The combination of the DFT and basis decomposition of the underlying signal leads to a sampling of the continuous kernel, $h(\cdot, \cdot)$, at the Fourier frequencies.

We denote the k-space observations corresponding to the n^{th} echo as $s_n(\boldsymbol{k}_i)$ where $\boldsymbol{k}_i$ is a grid point in 2-D k-space. The measurements can be expressed as,

$$s_n(\boldsymbol{k}_i) = \sum_{p=1}^{N} m(\boldsymbol{x}_p) f_{\boldsymbol{\theta}}^n(\boldsymbol{k}_i, \boldsymbol{x}_p) e^{j\boldsymbol{k}_i \cdot \boldsymbol{x}_p} + v_{n,i}. \tag{4.17}$$

The sequence $\{v_{n,i}\}$ is white complex Gaussian noise with variance σ^2 and $f_{\boldsymbol{\theta}}^n(\boldsymbol{k}, \boldsymbol{x})$ is the relaxation function, representing the acquisition and relaxation components.

As described in Section 4.2, these observations are collected in 2D frequency space and distortions occur in image space due to relaxation. We approximate the integral in (4.1) by a summation over the grid used for the parameter vector. Other modifications here are the superscript n denoting the image index and the notation $f_{\boldsymbol{\theta}}^n$ to acknowledge the dependence of the unknown parameters on the imaging kernel. The additive noise has also been explicitly included highlighting the stochastic nature of the measurements.

The relaxation function takes the form of (4.2) with the echo time now dependent on the image index, n. Specifically,

$$f_{\boldsymbol{\theta}}^n(\boldsymbol{k}_i, \boldsymbol{x}) = e^{-\tau_n / T_2(\boldsymbol{x})} e^{-|t_i - \tau_n| / T_2'(\boldsymbol{x})} \tag{4.18}$$

where τ_n is the echo time and t_i is the time after excitation. Figure 4.6 shows the acquisition parameters and signal amplitude for a spin echo sequence.

The standard method used to reconstruct the image, $\tilde{m}_n$, is the inverse discrete

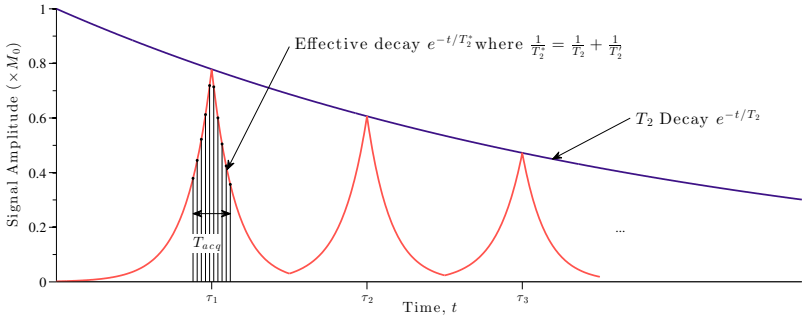


Figure 4.6: Signal amplitude of the bulk magnetisation during data acquisition for a spin echo sequence.

Fourier transform:

$$\tilde{m}_n(\boldsymbol{z}_r) = \sum_{p=1}^{N} m(\boldsymbol{x}_p) h_{\boldsymbol{\theta}}^n(\boldsymbol{x}_p, \boldsymbol{z}_r) + e_{n,r} \tag{4.19}$$

where $\boldsymbol{z}_r$ is a grid point in the output image, and the filter kernel, $h_{\boldsymbol{\theta}}^n(\boldsymbol{x}, \boldsymbol{z})$, is defined as

$$h_{\boldsymbol{\theta}}^n(\boldsymbol{x}, \boldsymbol{z}) = \sum_{i=1}^{N} f_{\boldsymbol{\theta}}^n(\boldsymbol{k}_i, \boldsymbol{x}) e^{j\boldsymbol{k}_i \cdot (\boldsymbol{x}-\boldsymbol{z})}. \tag{4.20}$$

The noise in image domain, $\{e_{n,r}\}$, is approximately white complex Gaussian for a sufficiently large N. A closed form for (4.20) has been found to simplify calculations. In this framework, the reconstructed image is interpreted as a filtered version of the initial magnetisation, $M_0(\boldsymbol{x})$, with the linear space-varying filter defined by the kernel, $h_{\boldsymbol{\theta}}^n(\boldsymbol{x}, \boldsymbol{z})$. This kernel models the spatial distortion in an image and provides an opportunity to reverse such distortion. Importantly, the kernel is dependent on both the acquisition parameters, $(T, T_{\text{acq}}, \tau_n)$, and the unknown relaxation parameters, $\boldsymbol{\theta}$. The properties of this kernel and its effect on relaxation time estimation are examined in detail in subsequent sections of this chapter.

Let $y_{n,r} = \tilde{m}_n(\boldsymbol{z}_r)$ and denote $\boldsymbol{y}_n = [y_{n,1}, \ldots, y_{n,N}]^T$. The observations from each echo image form a complete measurement vector, $\boldsymbol{y} = [\boldsymbol{y}_1^T, \ldots \boldsymbol{y}_M^T]^T$. The measurements are complex, $\boldsymbol{y} \in \mathbb{C}^{NM}$, and the parameters are positive and real, $\boldsymbol{\theta} \in [0, \infty)^{2N}$. The goal is to find an estimator for T_2^* and as such, M_0 is a nuisance parameter so we find an estimator $\hat{\boldsymbol{\theta}} : \mathbb{C}^{NM} \mapsto [0, \infty)^{2N}$.

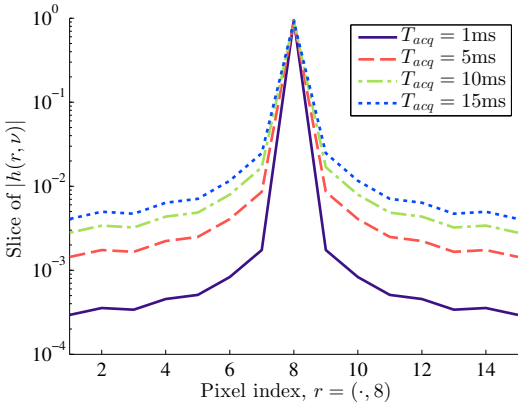


Figure 4.7: Magnitude slice through the centre of a two dimensional kernel, for the centre point in a 15x15 test object, for different T_{acq}.

4.4 Existing estimation method

The widespread approximation described in Section 4.2.2 is also used for the purposes of relaxation rate estimation. The filter kernel in (4.20) is approximated by,

$$h_{\boldsymbol{\theta}}^{n}(\boldsymbol{x},\boldsymbol{z}) \approx e^{-\tau_n/T_2(\boldsymbol{x})}\delta(\boldsymbol{x}-\boldsymbol{z}) \tag{4.21}$$

This reduces the measurement model in (4.19) to the well known forms [53],

$$\tilde{m}_n(\boldsymbol{z}_r) = m(\boldsymbol{z}_r)e^{-\tau_n/T_2(\boldsymbol{z}_r)} + e_r \tag{4.22}$$

These ubiquitous and somewhat naïve approximations assume that the signal amplitude is constant during data acquisition. In practice, relaxation and gradient dephasing occur during sampling, as shown in Figure 4.6, which compromises the Fourier encoding in (4.17), leading to distortion of the images. Figure 4.7 demonstrates the kernel arising from a simulated 15×15 object consisting of a constant T_2^* of 50 ms. Although the kernel approaches the 2-D weighted Kronecker delta function in (4.21) as $T_{\text{acq}} \to 0$, the non-zero sidelobes cause image blurring. That is, the signal from a given pixel is affected by the signal of neighbouring pixels. For a spin echo sequence, the distortion only occurs in the readout (x) direction, where relaxation occurs during sampling. Conversely, the phase (y) direction is encoded with separate excitations so no distortion occurs in this direction.

The approximation implicit in (4.22) is traditionally used for T_2 and T_2^* estimation. The advantage of this approximation is that it decouples the measurement at position $\boldsymbol{z}_r$ from the parameters of neighbouring locations, allowing $\boldsymbol{\theta}_r$ to be estimated independently using only M measurements, $\{\tilde{m}_1(\boldsymbol{z}_r), \tilde{m}_2(\boldsymbol{z}_r), \ldots, \tilde{m}_M(\boldsymbol{z}_r)\}$.

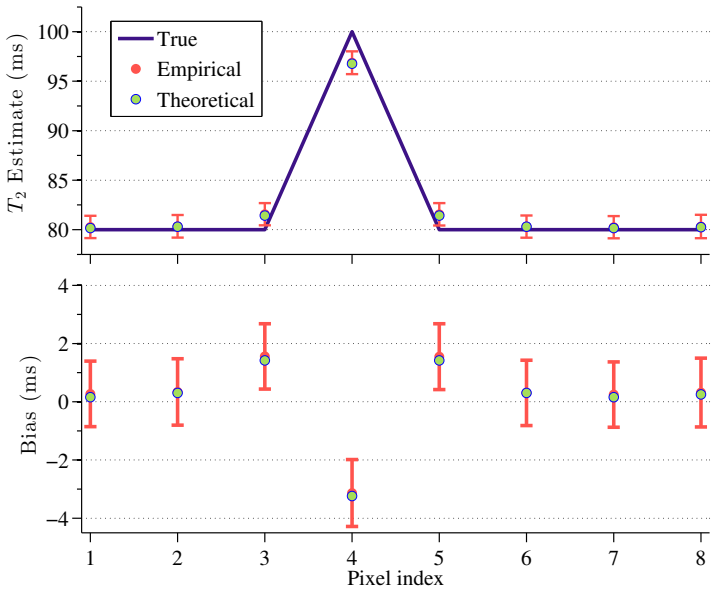


Figure 4.8: The theoretical and empirical estimation bias for a simulated 8 pixel object consisting of a T_2 spike of 100 ms in a background of $T_2 = 80$ ms.

Thus the estimation problem is reduced to N smaller problems of estimating two parameters from M complex measurements. Estimation using this simple model is performed using a least squares algorithm to fit a decaying exponential to the data. As will be shown in the following section, assuming a delta function kernel produces statistically biased estimates.

4.4.1 Analysis of estimation bias

In Appendix 4.C we perform a theoretical analysis of the least squares estimator using the assumed signal model in (4.22) when the actual data is generated from full signal model in (4.19). The analysis is based on a Taylor series expansion and gives a reasonably accurate indication of the expected bias for given experimental conditions. Figure 4.8 presents an example of the approximate bias and the true bias of the naive least squares estimator, simulated from 100 Monte Carlo trials of a simple object consisting of a T_2 spike of 100 ms in a background of 80 ms. A large bias is located around the sharp edge in the T_2 map. This is expected since the acquisition kernel blurs edges of an image, which affects the relaxation decay curve leading to increased estimation bias.

As noted previously, the measurement kernel is dependent on the acquisition parameters, and the unknown relaxation parameters, T_2, T_2' and M_0. Consequently the estimation bias is strongly dependent on these parameters. To quantify the bias

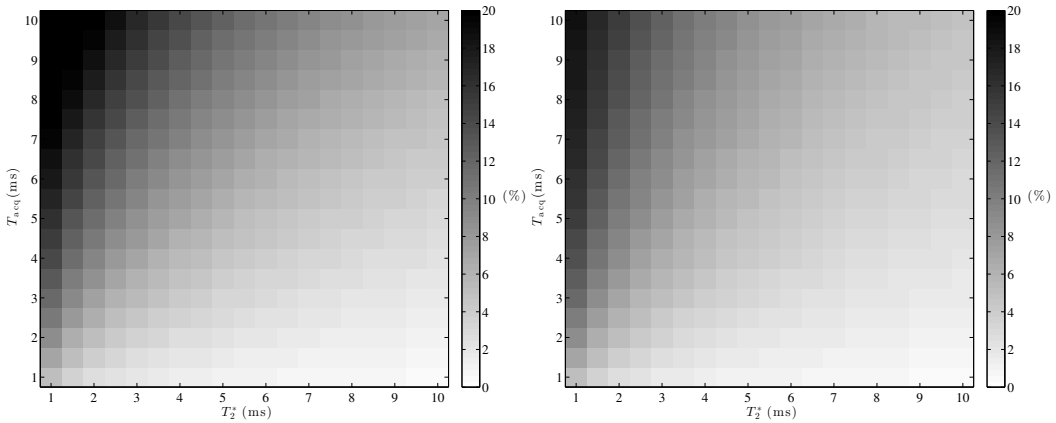


Figure 4.9: Estimation bias generated from (a) theoretical analysis and (b) 100 Monte Carlo trials as a function of the acquisition time T_{acq} and the effective relaxation time T_2^*. The simulated 8 pixel object consists of a T_2 spike of 100 ms in a background of $T_2 = 80$ ms.

over a range of situations, we repeat the experiment above for different pairs of T_{acq} and T_2^*. For each set of parameters we calculate the approximate bias and the empirical bias from 1000 Monte Carlo trials. Figure 4.9 shows the resultant bias as a function of T_{acq} and T_2^* and demonstrates the close match between our analytical bias and the true bias over a broad range of parameters. Note that some regions of the aggregate parameter space show near zero bias whereas other regions show significant bias. The regions with minimal bias are when $T_{\text{acq}} \ll T_2^*$, validating the assumption made in (4.22). However, it will be shown in the next section, that, improved estimates are achievable by acquiring samples with increased SNR and accounting for the image distortion.

4.5 Proposed estimation method

We propose estimation of relaxation times using the full measurement model in (4.19). In contrast to the pixel-independent estimation problems using (4.22), in this case, all the parameters, $\boldsymbol{\theta}$, must be estimated jointly. The joint estimation problem is substantially more difficult due to the large number of parameters and a simple non-linear least squares algorithm performs poorly and struggles to converge to the global minimum. Thus, we propose a Bayesian estimation algorithm.

4.5.1 Algorithm

In Bayesian estimation we aim to find the posterior probability density, $\pi(\boldsymbol{\theta}|\boldsymbol{y})$. With this density, the minimum mean-squared-error estimate, $\hat{\boldsymbol{\theta}}$, can be obtained by evaluating the conditional expectation, $\hat{\boldsymbol{\theta}} = \mathbb{E}_\pi[\boldsymbol{\theta}|\boldsymbol{y}]$. We cannot compute this expectation in closed form so we use Monte Carlo samples from $\pi(\boldsymbol{\theta}|\boldsymbol{y})$. The problem setup lends itself to a Gibbs sampler [61] for drawing the posterior samples. A Gibbs sampler is a Markov chain constructed by drawing from conditional distributions such that its invariant distribution is $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ [62]. The conditional distributions have the form $\pi(\theta_j|\boldsymbol{\theta}_{[\neg j]}, \boldsymbol{y})$ where $\boldsymbol{\theta}_{[\neg j]}$ are the elements in $\boldsymbol{\theta}$ excluding the j^{th} element. Bayes' rules states that

$$\pi(\theta_j|\boldsymbol{y}, \boldsymbol{\theta}_{[\neg j]}) \propto \ell(\boldsymbol{y}|\theta_j, \boldsymbol{\theta}_{[\neg j]})\pi_0(\theta_j|\boldsymbol{\theta}_{[\neg j]}). \tag{4.23}$$

We adopt a Gaussian prior for $\pi_0(\theta_j|\boldsymbol{\theta}_{[\neg j]})$ defined by $\pi_0(\boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{\theta}; \tilde{\boldsymbol{\theta}}, \boldsymbol{\Lambda})$, where $\tilde{\boldsymbol{\theta}}$ is the least squares estimate and $\boldsymbol{\Lambda}$ is a diagonal covariance matrix. In this work, we define $\boldsymbol{\Lambda}$ with variances of 50^2 ms for the relaxation times and 0.4^2 for the initial magnetisation, noting that the naive least squares estimator will provide a reasonable initial estimate. The signal model in (4.19) defines the expected signal as

$$\mu_{n,r}(\boldsymbol{\theta}) = \sum_{p=1}^{N} m(\boldsymbol{x}_p) h^n_{\boldsymbol{\theta}}(\boldsymbol{x}_p, \boldsymbol{z}_r). \tag{4.24}$$

For $\boldsymbol{\mu}(\boldsymbol{\theta}) = [\mu_{1,1}(\boldsymbol{\theta}), \mu_{1,2}(\boldsymbol{\theta}), \ldots, \mu_{M,N}(\boldsymbol{\theta})]^T$, the likelihood is

$$\ell(\boldsymbol{y}|\theta_j, \boldsymbol{\theta}_{[\neg j]}) = \mathcal{N}(\boldsymbol{y}; \boldsymbol{\mu}(\theta_j; \boldsymbol{\theta}_{[\neg j]}), \boldsymbol{\Sigma}) \tag{4.25}$$

We assume the measurement noise is white Gaussian and independent, thus $\boldsymbol{\Sigma} = \text{diag}(\sigma^2, \ldots, \sigma^2)$.

To draw samples, we approximate $\pi(\theta_j|\boldsymbol{y}, \boldsymbol{\theta}_{[\neg j]})$ by a piecewise uniform distribution [63], constructed by evaluating the posterior on an adaptive set of grid points, $\zeta_1, \ldots, \zeta_G$, to obtain weights $w_1, \ldots, w_G$. Each grid point, ζ_i, defines a uniform distribution on the interval $[\zeta_i - \Delta_G/2, \zeta_i + \Delta_G/2]$, denoted $\mathcal{U}_{[\zeta_i - \Delta_G/2, \zeta_i + \Delta_G/2]}$, creating a piece-wise continuous distribution across the grid. The grid is adaptive since it is redefined every iteration. The first point of the grid is chosen to be the current estimate and successive grid points are added in each direction until the calculated weights are close to zero. Furthermore, the grid spacing, Δ_G, is adjusted from one iteration to the next to keep the grid size close to a nominal value, denoted G_0.

The algorithm is run for P iterations and samples are collected after P_0 itera-

tions to allow the Markov chain to converge [64]. These samples, $\boldsymbol{\theta}^{(t)}, t = P_0, \ldots, P$, are used to calculate the point estimate, $\hat{\boldsymbol{\theta}} = \frac{1}{P-P_0} \sum_{t=P_0}^{P} \boldsymbol{\theta}^{(t)}$. The complete algorithm is listed in Algorithm 4.1.

Algorithm 4.1: Bayesian Estimation Algorithm

Input: $\boldsymbol{y}$ – observations
Output: $\hat{\boldsymbol{\theta}}$ – estimate
Params: $\boldsymbol{\Sigma}$ – noise covariance, $\boldsymbol{\Lambda}$ – prior covariance, P – number of samples, P_0 – Markov chain "burn in", G_0 – desired grid size, Δ_0 – initial grid spacing

Calculate least squares estimates $\tilde{\boldsymbol{\theta}}_1, \ldots, \tilde{\boldsymbol{\theta}}_N$.
$\boldsymbol{\theta} \leftarrow [\tilde{\boldsymbol{\theta}}_1, \ldots, \tilde{\boldsymbol{\theta}}_N]$
$\Delta_G \leftarrow \Delta_0$
for $t \leftarrow 1, \ldots, P$ **do**
  for $j \leftarrow 1, \ldots, 2N$ **do**
    Set $k \leftarrow 1, \quad \zeta_k \leftarrow \hat{\theta}_j$
    Set $w_{\max} = 0$
    repeat
      Compose: $\boldsymbol{\vartheta} \leftarrow [\zeta_k, \boldsymbol{\theta}_{[\neg j]}]$
      Calculate: $\boldsymbol{\mu}(\boldsymbol{\vartheta})$ from (4.24)
      Calculate: $w_k \leftarrow \mathcal{N}(\boldsymbol{y}; \boldsymbol{\mu}(\boldsymbol{\vartheta}), \boldsymbol{\Lambda}) \pi_0(\zeta_k | \boldsymbol{\theta}_{[\neg j]})$
      Assign: $w_{\max} \leftarrow w_k$ if $w_k > w_{\max}$
      Pick next grid point, ζ_{k+1}; $\quad k \leftarrow k+1$
    until $w_k < 0.01 w_{\max}$
    Normalise weights: $w_k \leftarrow \tilde{w}_k / \sum_i \tilde{w}_i$
    Sample $\theta_j^{(t)} \sim \sum_{k=1}^{g} w_k \mathcal{U}_{[\zeta_k - \Delta/2, \zeta_k + \Delta/2]}$
    Adjust grid spacing, $\Delta_G \leftarrow (1 + (G - G_0)/G_0)\Delta_G$
Return estimate: $\hat{\boldsymbol{\theta}} \leftarrow \frac{1}{P-P_0} \sum_{t=P_0}^{P} \boldsymbol{\theta}^{(t)}$

4.5.2 Properties of the Bayesian estimation algorithm

As explained in Section 3.7.1, the noise variance, σ^2, in an MRI experiment is proportional to the receiver bandwidth, $\mathsf{BW} = 1/T_{\text{acq}}$. The signal-to-noise ratio of the imaging system is defined as $\mathsf{SNR} = s/\sigma$ where s is the signal amplitude, which we assume to be constant. Thus

$$\mathsf{SNR} \propto \sqrt{T_{\text{acq}}}. \tag{4.26}$$

As Figure 4.9 illustrates, in many practical cases the distortion introduced by the finite sampling bandwidth (or non-zero acquisition time) is negligible. However, even in these cases, it is possible to obtain a smaller estimation error by se-

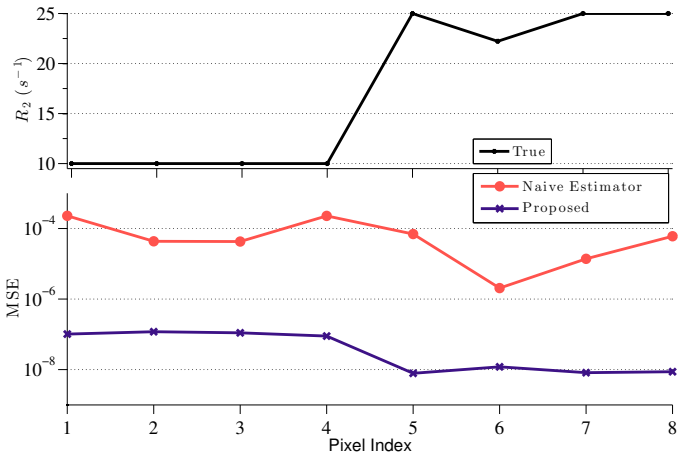


Figure 4.10: Comparison of the proposed T_2 estimation algorithm and the naïve least-squares approach showing the true $R_2 = 1/T_2$ profile and the mean square error for both approaches calculated from 100 noise realisations.

lecting a longer acquisition time and using the proposed estimator to account for the increased distortion. The increase in T_{acq} improves the SNR which decreases the estimation variance.

4.6 Simulations

The estimation algorithm developed in the previous sections is demonstrated using a CPMG sequence with multiple echoes. We simulate the spin echo sequence with echo times defined by $\tau_n = (4.5 + 3.5(n-1))$ ms for $n = 1, \ldots, M$. We use $M = 16$, and a relatively long $T_{\text{acq}} = 5.12$ ms, for a numerical object defined in Figure 4.10, with T_2 features typical of a human brain [65] and $M_0 = 1$ representing a constant proton density. The naïve estimator using non-linear least squares on the simple measurement model in (4.22) is compared with the proposed algorithm using the complete model in (4.19). The algorithm was run using $P = 1100$, $P_0 = 100$, $\mathbf{\Lambda} = \text{diag}(2.5 \times 10^{-5}, 10^{-2}, \ldots, 2.5 \times 10^{-5}, 10^{-2})$. The estimation was repeated for 100 realisations of noise.

The results are shown in Figure 4.10 where the rate, $R_2 = 1/T_2$ is displayed following the convention in MRI literature. The Bayesian algorithm has a smaller mean-square-error and successfully compensates for the distortion. Figure 4.11 demonstrates that the estimation error can always be decreased by increasing T_{acq} and reaping the benefits of improved SNR whilst counteracting the increase in distortion using the proposed estimator. Conversely, the performance of any algorithm that fails to account for the distortion, will degrade as T_{acq} increases.

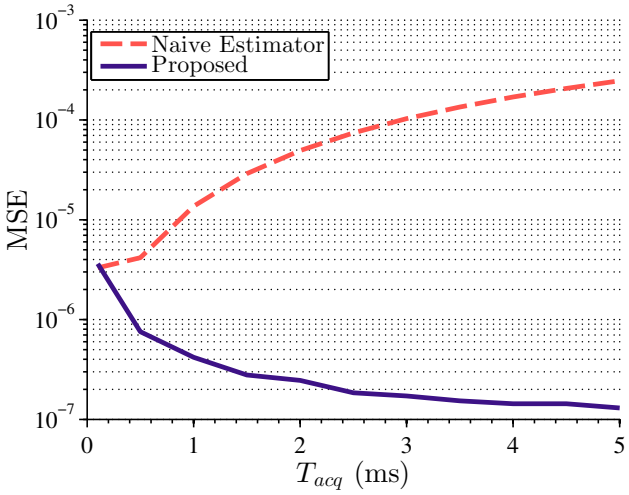


Figure 4.11: Mean square error, as T_{acq} increases, of the proposed algorithm and the naïve least-squared estimator. The simulation setup is otherwise the same as used for Figure 4.10.

An important implication of these results is revealed by considering the SNR dependencies discussed in Section 3.7.1. The SNR increases linearly with field strength and the MSE decreases with increasing SNR. Consequently, the same level of accuracy of current systems can be achieved with a much weaker field strength and an improved signal model.

Figure 4.12 displays the bias and variance of both estimators. This highlights that although the variance of the naive estimator deceases with increased SNR due to longer T_{acq}, the bias is the dominant source of error. Thus for the naive estimator, the potential improvement from increased SNR is outweighed by the increase in estimation bias.

4.7 Discussion and conclusion

In this chapter, the image distortion introduced by signal relaxation during sampling was investigated and modelled with a linear filter. The filter kernel was examined in detail, both analytically and through simulation. In the context of MRI relaxation rate estimation, it was revealed that this distortion leads to statistical bias when the commonly accepted signal model is used. Due to the acquisition noise properties, such estimators incur a tradeoff between bias and SNR. We have presented an algorithm suitable to estimate the transverse relaxation rate from spin echo sequences. The proposed estimator overcomes this tradeoff by accounting for the distortion in reconstructed images, resulting in improved estimates.

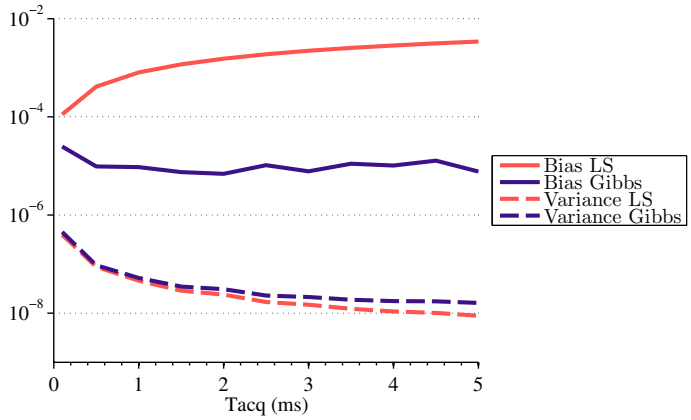


Figure 4.12: Bias and variance, as T_{acq} increases, of the proposed algorithm and the naive least-squared estimator. The simulation setup is otherwise the same as used for Figure 4.10.

An important issue not considered in this work is the constraints between the acquisition time and inter-echo spacing. From an estimation point of view we want the maximum number of echoes at the highest possible SNR. Unfortunately, the acquisition process creates a tradeoff between the number of echoes and the SNR. It may be better to acquire fewer echoes at higher SNR and use the proposed estimator to account for the distortion although further analysis is required to determine the optimal tradeoff.

Although our estimator provides a smaller estimation error than traditional estimators, it is relatively complex and very computationally demanding. For example, the results presented in this chapter took over 10 hours to produce using a typical desktop computer. Simplifications of the estimator and an optimised implementation are expected to greatly reduce the computation time, a topic for future work. For current clinical scanners the improvement in SNR may not be worth the extra complexity. It is expected that this technique will be most valuable for systems intrinsically limited in SNR, such as ultra low field or portable MRI systems [2] as discussed in Chapter 1. The technique could also be used for high-field systems to reduce the acquisition time (at the expense of SNR) with no loss in overall performance. The tradeoff between cost and SNR also suggest that similar estimation accuracy may be achieved on cheaper systems. For example the increased complexity may be justified if the same level of accuracy of a 3 T machine can be achieved on 1.5 T system, a cost saving in the order of a million dollars [66].

Appendices

4.A Introduction to the Dirichlet kernel

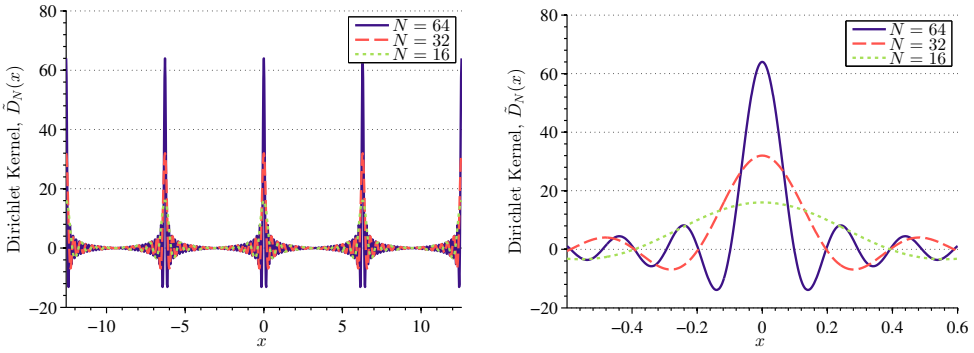
The Dirichlet kernel is defined as

$$D_N(x) := \sum_{k=-\frac{N-1}{2}}^{\frac{N-1}{2}} e^{jkx}. \tag{4.27}$$

which assumes symmetric sampling of k. That is, N is odd and k is summed from $-\frac{N-1}{2}$ to $\frac{N-1}{2}$. In reality, traditional use of the Fast Fourier Transform (FFT) favours an even number of samples. Furthermore, it is desirable to sample the point $k = 0$ and thus, as pointed out in [53, Page 240], MRI samples are typically collected from sampling from $-\frac{N}{2}$ to $\frac{N}{2} - 1$ for even N. In this case we have the relationship,

$$\tilde{D}_N(x) = D_{N-1}(x) + e^{-jN/2x} \tag{4.28}$$

This subtle discrepancy in the definition does not effect the general behaviour of the function; it merely adds a linear phase term, consistent our understanding of the Fourier sampling in MRI. Fig. 4.13 demonstrates the behaviour of $D_N(x)$ for increasing number of measurements. We see that the period is 2π regardless of the number of measurements and as $N \to \infty$, the Dirichlet kernel approaches a Dirac comb. For application in MRI, the period of the Dirichlet kernel is modified by scaling the argument by the sampling properties. In (7.41), the argument is scaled by $R_x \Delta_x W_x$ when $\boldsymbol{x} = (nW_x, mW_y)$.



(a) (b)

Figure 4.13: The Dirichlet kernel for three measurement sample sizes. (a) demonstrates the fixed period of the kernel while (b) indicates that the kernel approaches a Dirac comb for $N \to \infty$.

The main use of the Dirichlet kernel has been in Fourier series analysis. We see that, the convolution of $D_N(x)$ with a periodic function f with period 2π is the n^{th}-order Fourier series approximation to f where $n = (N-1)/2$. That is,

$$(D_N * f)(x) = \sum_{k=-\frac{N-1}{2}}^{\frac{N-1}{2}} \hat{f}(k) e^{jkx}. \tag{4.29}$$

where $\hat{f}(k)$ is the k^{th} Fourier coefficient. It follows naturally, that convergence properties of the Fourier series can be analysed by studying the properties of the Dirichlet kernel [67].

4.B Closed form expression for filter kernel

The time variable considered in this work evolves between excitations. As such, it is suitable to describe the evolution of the trajectory for a single readout line. In particular,

$$t_i = iT + \tau_E, \quad i = -K, \ldots, K-1 \tag{4.30}$$

The relaxation function, $f(\boldsymbol{k}(t), \boldsymbol{x})$ can be rewritten as

$$f(\boldsymbol{k}(t_i), \boldsymbol{x}) = e^{-(iT+\tau_E)/T_2(\boldsymbol{x})} e^{-|iT|/T_2'(\boldsymbol{x})} \tag{4.31}$$

The k-space sampling occurs according to

$$\boldsymbol{k}_{q,i} := \boldsymbol{k}_q(t_i) = (\Delta_x i, \Delta_y q), \quad i, q = -K, \ldots, K-1 \tag{4.32}$$

where i and q are the column and row indices of the 2D Fourier acquisition grid, respectively. The acquisition size is defined by $K = N/2$. For brevity we let $\boldsymbol{u} = \boldsymbol{x} - \boldsymbol{z}$ and analyse the expression below,

$$h(\boldsymbol{x}, \boldsymbol{u}) = \sum_{\boldsymbol{k}} f(\boldsymbol{k}, \boldsymbol{x}) e^{j\boldsymbol{k}\cdot\boldsymbol{u}} \tag{4.33}$$

$$= \sum_i \sum_q e^{-(iT+\tau_E)/T_2(\boldsymbol{x})} e^{-|iT|/T_2'(\boldsymbol{x})} e^{j(\boldsymbol{k}_{q,i}\cdot\boldsymbol{u})} \tag{4.34}$$

$$= e^{-\tau_E/T_2(\boldsymbol{x})} \sum_q e^{jq\Delta_y u_y} \sum_i e^{-iT/T_2(\boldsymbol{x})} e^{-|iT|/T_2'(\boldsymbol{x})} e^{ji\Delta_x u_x} \tag{4.35}$$

We first examine the summation over the phase-encode coordinate of k-space. We denote the summation as, $\tilde{D}_N$,

$$\tilde{D}_N(u_y) = \sum_{q=-K}^{K-1} e^{jq\Delta_y u_y} \tag{4.36}$$

$$= D_{K-1}(\Delta_y u_y) + e^{-jK\Delta_y u_y} \tag{4.37}$$

where D_{K-1} is the regular $(K-1)^{\text{th}}$ degree Dirichlet kernel described in Appendix 4.A.

Alternatively using a geometric progression we obtain,

$$\tilde{D}_N(u_y) = \sum_{q=-K}^{K-1} e^{jq\Delta_y u_y} \tag{4.38}$$

$$= e^{-jK\Delta_y u_y} \sum_{q=0}^{2K-1} e^{jq\Delta_y u_y} \tag{4.39}$$

$$= e^{-jK\Delta_y u_y} \frac{1 - e^{j2K\Delta_y u_y}}{1 - e^{j\Delta_y u_y}} \tag{4.40}$$

$$= \frac{e^{-jK\Delta_y u_y} - e^{jK\Delta_y u_y}}{1 - e^{j\Delta_y u_y}} \tag{4.41}$$

$$= \frac{-2j\sin(K\Delta_y u_y)}{1 - e^{j\Delta_y u_y}} \tag{4.42}$$

Multiplying numerator and denominator by $e^{-j\Delta_y u_y/2}$ yields

$$\tilde{D}_N(u_y) = \frac{-2j\sin(K\Delta_y u_y)}{1 - e^{j\Delta_y u_y}} \tag{4.43}$$

$$= \frac{-2j\sin(K\Delta_y u_y)e^{-j\Delta_y u_y/2}}{e^{-j\Delta_y u_y/2} - e^{j\Delta_y u_y/2}} \tag{4.44}$$

$$= \frac{-2j\sin(K\Delta_y u_y)e^{-j\Delta_y u_y/2}}{-2j\sin(\Delta_y u_y/2)} \tag{4.45}$$

$$= \frac{\sin(K\Delta_y u_y)e^{-j\Delta_y u_y/2}}{\sin(\Delta_y u_y/2)} \tag{4.46}$$

The expressions in (4.37) and (4.46) are equivalent.

The singularity at $u_y = 0$ is easily handled by l'Hôptial's rule [68] resulting in the full expression,

$$\tilde{D}_N(u_y) = \begin{cases} 2K & u_y = 0 \\ \frac{\sin(K\Delta_y u_y)e^{-j\Delta_y u_y/2}}{\sin(\Delta_y u_y/2)} & \text{otherwise} \end{cases} \tag{4.47}$$

Returning to the filter kernel expression, we now consider the summation over

the readout coordinate,

$$g(u_x) = \sum_{i=-K}^{K-1} e^{-iT/T_2(\boldsymbol{x})} e^{-|iT|/T_2'(\boldsymbol{x})} e^{ji\Delta_x u_x} \tag{4.48}$$

$$= \sum_{i=-K}^{-1} e^{-iT/T_2(\boldsymbol{x})} e^{iT/T_2'(\boldsymbol{x})} e^{ji\Delta_x u_x} + \sum_{i=0}^{K-1} e^{-iT/T_2(\boldsymbol{x})} e^{-iT/T_2'(\boldsymbol{x})} e^{ji\Delta_x u_x} \tag{4.49}$$

Now define

$$\frac{1}{T_2^*(\boldsymbol{x})} = \frac{1}{T_2(\boldsymbol{x})} + \frac{1}{T_2'(\boldsymbol{x})} \tag{4.50}$$

$$\frac{1}{T_2^-(\boldsymbol{x})} = \frac{1}{T_2(\boldsymbol{x})} - \frac{1}{T_2'(\boldsymbol{x})} \tag{4.51}$$

such that $g(u_x)$ can be written as

$$g(u_x) = \sum_{i=-K}^{-1} e^{-iT/T_2^-(\boldsymbol{x})} e^{ji\Delta_x u_x} + \sum_{i=0}^{K-1} e^{-iT/T_2^*(\boldsymbol{x})} e^{ji\Delta_x u_x} \tag{4.52}$$

Consider the first term, which we denote g_1,

$$g_1(u_x) = \sum_{i=-K}^{-1} e^{i(-T/T_2^-(\boldsymbol{x}) + j\Delta_x u_x)} \tag{4.53}$$

$$= e^{-K(-T/T_2^-(\boldsymbol{x}) + j\Delta_x u_x)} \sum_{i=0}^{K-1} e^{i(-T/T_2^-(\boldsymbol{x}) + j\Delta_x u_x)} \tag{4.54}$$

$$= e^{-K(-T/T_2^-(\boldsymbol{x}) + j\Delta_x u_x)} \frac{1 - e^{K(-T/T_2^-(\boldsymbol{x}) + j\Delta_x u_x)}}{1 - e^{(-T/T_2^-(\boldsymbol{x}) + j\Delta_x u_x)}} \tag{4.55}$$

$$= \frac{-1 + e^{-K(-T/T_2^-(\boldsymbol{x}) + j\Delta_x u_x)}}{1 - e^{(-T/T_2^-(\boldsymbol{x}) + j\Delta_x u_x)}} \tag{4.56}$$

Similarly for the second term, g_2, we get

$$g_2(u_x) = \sum_{i=0}^{K-1} e^{i(-T/T_2^*(\boldsymbol{x}) + j\Delta_x u_x)} \tag{4.57}$$

$$= \frac{1 - e^{K(-T/T_2^*(\boldsymbol{x}) + j\Delta_x u_x)}}{1 - e^{(-T/T_2^*(\boldsymbol{x}) + j\Delta_x u_x)}} \tag{4.58}$$

We can now write the expression for the filter kernel that does not involve any summations. That is,

$$h(\boldsymbol{x}, \boldsymbol{z}) = e^{-\tau_E/T_2(\boldsymbol{x})} \tilde{D}_N(u_y) \left[g_1(u_x) + g_2(u_x) \right] \tag{4.59}$$

where u_x and u_y are the x and y coordinates of the difference, $\boldsymbol{x} - \boldsymbol{z}$.

4.C Analysis of estimation bias

The least squares estimator minimises the following cost function

$$f(\boldsymbol{\theta}) = \sum_{i=1}^{M} (y_i - g_i(\boldsymbol{\theta}))^2 \tag{4.60}$$

where $g_i = me^{-t_i/T_2}$.

The derivative of the cost function at the least squares estimate will be zero. That is,

$$\boldsymbol{J}(\hat{\boldsymbol{\theta}}) = \nabla_{\boldsymbol{\theta}} f(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = \boldsymbol{0} \tag{4.61}$$

where $\nabla_{\boldsymbol{\theta}} = [\partial/\partial m, \partial/\partial T_2]^T$. Taking the first order Taylor series expansion around the true parameter $\boldsymbol{\theta}^\star$ yields

$$\boldsymbol{J}(\hat{\boldsymbol{\theta}}) = 0 \approx \boldsymbol{J}(\boldsymbol{\theta}^\star) + \boldsymbol{H}(\boldsymbol{\theta}^\star)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^\star) \tag{4.62}$$

where $\boldsymbol{H}(\tilde{\boldsymbol{\theta}}) = \nabla_{\boldsymbol{\theta}} \nabla_{\boldsymbol{\theta}}^T f(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}}$. The bias is then approximated by

$$\text{bias} = \mathbb{E}\left[\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^\star\right] \approx -\mathbb{E}\left[\boldsymbol{H}^{-1}(\boldsymbol{\theta}^\star)\boldsymbol{J}(\boldsymbol{\theta}^\star)\right] \tag{4.63}$$

In our case,

$$\boldsymbol{J}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial f}{\partial M} & \frac{\partial f}{\partial T} \end{bmatrix}^T \tag{4.64}$$

$$\boldsymbol{H}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial^2 f}{\partial M^2} & \frac{\partial^2 f}{\partial MT} \\ \frac{\partial^2 f}{\partial TM} & \frac{\partial^2 f}{\partial T^2} \end{bmatrix} \tag{4.65}$$

For the sake of brevity we let $f'_T = \frac{\partial f}{\partial T}$, $f'_M = \frac{\partial f}{\partial M}$, $f'_{TT} = \frac{\partial^2 f}{\partial T^2}$, $f'_{MM} = \frac{\partial f}{\partial M^2}$, $f'_{TM} = \frac{\partial f}{\partial TM}$. The inverse of the 2×2 matrix H can be computed explicitly as

$$\boldsymbol{H}^{-1}(\theta) = \frac{1}{f'_{TT}(\theta) f'_{MM}(\theta) - f'_{TM}(\theta) f'_{MT}(\theta)} \begin{bmatrix} f'_{TT}(\theta) & -f'_{TM}(\theta) \\ -f'_{MT}(\theta) & f'_{MM}(\theta) \end{bmatrix} \tag{4.66}$$

The relevant derivatives are

$$f'_V(\boldsymbol{\theta}) = -2 \sum_{i=1}^{M} (y_i - g_i(\boldsymbol{\theta})) g'_{i,V}(\boldsymbol{\theta}) \tag{4.67}$$

$$= -2\sum_{i=1}^{M}(\mu_t(\boldsymbol{\theta}^\star) - g_i(\boldsymbol{\theta}))g'_{i,V}(\boldsymbol{\theta}) - 2\sum_{i=1}^{M} v_i g'_{i,V}(\boldsymbol{\theta}) \tag{4.68}$$

and

$$f''_{VW}(\boldsymbol{\theta}) = -2\sum_{i=1}^{M}\left[(y_i - g_i(\boldsymbol{\theta}))g''_{i,VW}(\boldsymbol{\theta}) - g'_{i,W}(\boldsymbol{\theta})g'_{i,V}(\boldsymbol{\theta})\right] \tag{4.69}$$

$$= -2\sum_{i=1}^{M}\left[(\mu_t(\boldsymbol{\theta}^\star) - g_i(\boldsymbol{\theta}))g''_{i,VW}(\boldsymbol{\theta}) - g'_{i,W}(\boldsymbol{\theta})g'_{i,V}(\boldsymbol{\theta})\right] - 2\sum_{i=1}^{M} v_i g''_{i,VW}(\boldsymbol{\theta}) \tag{4.70}$$

where V and W can be different combinations of the variables T and M. The individual derivatives of $g(\boldsymbol{\theta})$ are

$$g_i(\boldsymbol{\theta}) = me^{-t_i/T_2} \tag{4.71}$$

$$g'_{i,T}(\boldsymbol{\theta}) = \frac{mt_i}{T_2^2}e^{-t_i/T_2} \tag{4.72}$$

$$g'_{i,M}(\boldsymbol{\theta}) = e^{-t_i/T_2} \tag{4.73}$$

The second derivatives are

$$g''_{i,TT}(\boldsymbol{\theta}) = \frac{-2mt_i}{T_2^3}e^{-t_i/T_2} + \frac{mt_i^2}{T_2^4}e^{-t_i/T_2} \tag{4.74}$$

$$g''_{i,MM}(\boldsymbol{\theta}) = 0 \tag{4.75}$$

$$g''_{i,TM}(\boldsymbol{\theta}) = g''_{i,MT}(\boldsymbol{\theta}) = \frac{t_i}{T_2^2}e^{-t_i/T_2} \tag{4.76}$$

The two elements of the matrix multiplication in (4.63) are obtained by

$$\text{bias}_M = \mathbb{E}\left[\frac{f'_M(\boldsymbol{\theta})f'_{TT}(\boldsymbol{\theta}) - f_T(\boldsymbol{\theta})f'_{TM}(\boldsymbol{\theta})}{f'_{TT}(\boldsymbol{\theta})f'_{MM}(\boldsymbol{\theta}) - f'_{TM}(\boldsymbol{\theta})f'_{MT}(\boldsymbol{\theta})}\right] \tag{4.77}$$

$$\text{bias}_T = \mathbb{E}\left[\frac{f_T(\boldsymbol{\theta})f'_{MM}(\boldsymbol{\theta}) - f_M(\boldsymbol{\theta})f'_{MT}(\boldsymbol{\theta})}{f'_{TT}(\boldsymbol{\theta})f'_{MM}(\boldsymbol{\theta}) - f'_{TM}(\boldsymbol{\theta})f'_{MT}(\boldsymbol{\theta})}\right] \tag{4.78}$$

Each f' factor above is a scalar that can be written in the form $\boldsymbol{\alpha}^T\boldsymbol{\beta} + \boldsymbol{\xi}\boldsymbol{v}$, where $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\xi}$ are deterministic and $\boldsymbol{v}$ is the stochastic noise. Multiplication of two such factors gives a quadratic form in $\boldsymbol{v}$. The desired bias is given by the expectation of the ratio of two such forms. The relevant expectation can be written as

$$\mathbb{E}\left[\frac{\eta}{\delta}\right] = \mathbb{E}\left[\frac{a + \boldsymbol{b}^T\boldsymbol{v} + \boldsymbol{v}^T A\boldsymbol{v}}{c + \boldsymbol{d}^T\boldsymbol{v} + \boldsymbol{v}^T B\boldsymbol{v}}\right] \tag{4.79}$$

To analyse the expression in (4.79), we consider the order of growth of each term. For stochastic terms, we use the order in probability, denoted $\mathcal{O}_p$. By definition a random variable X_n is said to be $\mathcal{O}_p(b_n)$ if, for each $\epsilon > 0$, there exists $N(\epsilon)$ and $B(\epsilon)$ such that, if $n \geq N(\epsilon)$,

$$\Pr(|X_n/b_n| < B(\epsilon)) \geq 1 - \epsilon. \tag{4.80}$$

We consider the stochastic term,

$$X_M = \boldsymbol{b}^T \boldsymbol{v} = \sum_{i=1}^{M} b_i v_i. \tag{4.81}$$

The mean of X_M is zero and since b_i is bounded for $i = 1, \ldots, M$ there is a k and $N(k)$ such that $\text{var}(X_M) \leq kM$ for $M \geq N(k)$. Chebyshev's inequality can be used to prove,

$$\Pr(|X_M| \leq a) \geq 1 - \text{var}(X_M)/a^2. \tag{4.82}$$

For $M \geq N(k)$,

$$\Pr(|X_M| \leq a) \geq 1 - kM/a^2 \tag{4.83}$$

$$\Pr(|X_M/\sqrt{M}| \leq \alpha) \geq 1 - k/\alpha^2 \tag{4.84}$$

$$\Pr(|X_M/\sqrt{M}| \leq \sqrt{k/\epsilon}) \geq 1 - \epsilon. \tag{4.85}$$

It follows that $\boldsymbol{b}^T \boldsymbol{v} = \mathcal{O}_p(\sqrt{M})$. An identical argument can be used to prove that $\boldsymbol{d}^T \boldsymbol{v} = \mathcal{O}_p(\sqrt{M})$.

We now consider the order of the quadratic term,

$$Y_M = \boldsymbol{v}^T A \boldsymbol{v} = \sum_{i,j=1}^{M} A_{i,j} v_i v_j. \tag{4.86}$$

Since the noise is white, we can show that there exists constants c and $N(c)$ such that $\mathbb{E}[Y_M] < cM$ for $M \geq N(c)$. Since Y_M is positive, we use Markov's inequality to prove that for $M \geq N(c)$,

$$\Pr(Y_M \leq a) \geq 1 - kM/a \tag{4.87}$$

$$\Pr(Y_M/M \leq \alpha) \geq 1 - k/\alpha \tag{4.88}$$

$$\Pr(Y_M/M \leq k/\epsilon) \geq 1 - \epsilon. \tag{4.89}$$

Therefore, $\boldsymbol{v}^T A \boldsymbol{v} = \mathcal{O}_p(M)$. Similarly for the other quadratic term, $\boldsymbol{v}^T B \boldsymbol{v} = \mathcal{O}_p(M)$.

The deterministic terms grow quadratically. That is,

$$a = \mathcal{O}(M^2) \tag{4.90}$$

$$c = \mathcal{O}(M^2) \tag{4.91}$$

To analyse the expectation in (4.79) we consider the stochastic parts as error terms. According to the above analysis, the numerator can be written as $\eta = a(1 + \mathcal{O}_p(M^{-1}))$ and the denominator can be written as $\delta = c(1 + \mathcal{O}_p(M^{-1}))$. The inverse of the denominator is analysed by a Taylor series expansion resulting in,

$$\delta^{-1} = \left(c(1 + \mathcal{O}_p(M^{-1}))\right)^{-1} \tag{4.92}$$

$$= \frac{1}{c}\left(1 + \mathcal{O}_p(M^{-1})\right). \tag{4.93}$$

The expectation in (4.79) can now be approximated as

$$\mathbb{E}\left[\frac{\eta}{\delta}\right] \approx \frac{a}{c} + \mathcal{O}(M^{-1}). \tag{4.94}$$

This approximation is used to compute the estimation bias for the initial magnetisation, M_0, and the relaxation time, T_2, when the least squares estimator is used with the simplified signal model.

CHAPTER 5

Estimation of multi-component relaxation rate distributions

Contents

5.1 Introduction

In an ideal environment, spin-spin relaxation can be modelled by a decaying exponential with a characteristic time constant, T_2. The previous chapter examined the accurate estimation of these time constants in the presence of distortion

due to finite sampling bandwidth. The signal model used there cannot always be applied in practice. Firstly, it was assumed that only a single exponential decay was present in each voxel. In reality, different spins within a single voxel can interact with vastly different molecular environments, leading to a signal decay that arises from a distribution of T_2 values [69]. Secondly, deviations in the refocusing flip angle due to B_1 field inhomogeneity, off-resonance effects or non-uniform slice profiles were ignored. These deviations lead to secondary and stimulated echoes which modulate the pure exponential decay and result in biased estimation of the underlying T_2 values [70]. In this chapter, T_2 estimation is considered for a signal model which incorporates both of these extensions.

Relaxation data for T_2 estimation is traditionally acquired using a multi-echo spin echo acquisition based on the CPMG sequence described in Section 3.3. This sequence relies on a train of 180° RF pulses to refocus the transverse magnetisation and produce an echo that is weighted by the T_2 components. In practice, the refocusing flip angle can deviate considerably from 180°. The echo amplitudes resulting from non-ideal refocusing pulses can be derived from the extended phase graph (EPG) algorithm [71, 72]. The EPG algorithm tracks the multiple coherence pathways of spins after consecutive periods that model precession, relaxation and refocusing. In this way, the problematic stimulated echoes can be accounted for using the improved signal model of the EPG algorithm.

Previous works to estimate T_2 distributions can be categorised into two different approaches, based on the underlying model for distribution. The first approach attempts to model a general continuous distribution. The method of Whittall and MacKay proposed in [5] is to approximate the distribution with a large number of delta functions at fixed locations and solve the resulting optimisation using the non-negative least squares (NNLS) algorithm. This technique has been used extensively for T_2 distribution estimation, e.g. [73, 74, 75, 76]. The recent work in [77] extended the NNLS algorithm to compensate for stimulated echoes arising from non-ideal flip angles.

The second approach taken in the literature is to assume the distribution consists of a small number of discrete components at unknown locations. A parametric model for a discrete distribution consisting of three pools was proposed and tested in [7] and [8], respectively. Estimation of the multiple discrete components was performed using a gradient based optimisation algorithm, equivalent to finding the maximum likelihood estimate (MLE) for additive Gaussian noise. However, in that simulation study, the algorithm was initialised to the true values. In general, the discrete model has proved unreliable when initialised away from ground truth [78]. Stimulated echo correction for non-ideal refocusing flip angles

has been applied in [79] for single component relaxation analysis. In that work a gradient based optimisation algorithm was used to jointly estimate the flip angle and T_2 value.

This chapter makes two novel contributions. Firstly, we propose a third approach for the modelling of multi-component T_2 distributions that is both parametric and continuous and use it calculate the expected accuracy of the estimated distributions. To this end, we demonstrate that it is not feasible to reliably estimate the width of the distribution, albeit we can expect some accuracy in the estimated locations of the distribution mass.

In light of this analysis, the second contribution of this work is to extend the discrete distribution model in [8] to include stimulated echo compensation using the EPG algorithm, analogous to [79] for the single component case. To overcome problems of the MLE associated with low SNR and algorithm initialisation, we develop a Bayesian algorithm to jointly estimate the parameters of the B_1 field and the weights and locations of the discrete T_2 components. A simplified version of this algorithm was proposed in [80] for the case of ideal 180° refocusing pulses.

The chapter is organised as follows. In Section 5.2 we present the general signal equation for T_2 distribution estimation. In Section 5.3 we review the existing approaches to T_2 estimation to date and show how all signal expressions derive from the same basic equation. We emphasise the inherent assumptions of each model. In Section 5.4 we propose a parametric continuous model and in Section 5.5 we use it to calculate bounds on the estimation performance and discuss the limited information available in the measurement. In Section 5.6 we present an improved estimation algorithm for the discrete model and apply it to simulated and experimental data. Finally, in Section 5.7, we discuss optimising the CPMG sequence for T_2 distribution estimation.

5.1.1 Notation

Table 5.1 lists the important symbols used in this chapter. The list is not exhaustive, and additional notation will be introduced as it is required.

5.2 Theory

The distribution of relaxation times is observed through the amplitudes of the acquired echo signals. In Chapter 4, the focus was on distortion due to finite sampling bandwidth, which resulted in pixel blurring modelled with a linear filter. In this chapter, we assume the readout acquisition times are sufficiently small to

Table 5.1: Common notation used in Chapter 5

Symbol	Quantity
$f(\cdot)$	Distribution of relaxation times
$g_n(\cdot)$	Relaxation function
M	Number of echoes
t_n	Echo time
N	Number of distribution components
w_i	Weight of the i^{th} component
τ_i	Relaxation time of the i^{th} component
L	NNLS grid size
$\boldsymbol{A}$	Linear measurement operator
$\boldsymbol{C}$	NNLS regularisation constraints
λ	Regularisation weight
α, β	Gamma distribution parameters
ν, ρ^2	Mean and variance of gamma distribution
$\boldsymbol{F}$	Fisher information matrix
$\boldsymbol{J}$	Jacobian matrix
$\gamma_1, \ldots, \gamma_P$	Correction weights
$\boldsymbol{\mu}_0, \boldsymbol{\Psi}_0$	Mean and covariance of prior
$\boldsymbol{K}$	Kalman gain
$\boldsymbol{\Sigma}$	Covariance of noise
χ^2	Misfit metric
α	Flip angle
F_k	Transverse coherence state
Z_k	Longitudinal coherence state
$\boldsymbol{x}$	Vector of coherence states
$\boldsymbol{P}$	Relaxation and precession operator
$\boldsymbol{T}$	Excitation operator
$\boldsymbol{E}$	Evolution operator
$\boldsymbol{c}$	Output selection vector

justify the delta function kernel as described in Section 4.2.2. In this case, the measurements from a pixel are independent and depend only on the properties of that pixel. The measured MRI signal for a single pixel at the n^{th} echo is described by the integration of the decaying signals from each contribution within the pixel. That is,

$$s_n = \int f(\tau) g_n(\tau) e^{j\phi} d\tau \tag{5.1}$$

where f is the unknown distribution of relaxation times, g_n is a function describing the signal amplitude for a given relaxation rate and ϕ is the signal phase. Notice the complex nature of the signal, $s_n \in \mathbb{C}$.

The measurements are noisy samples of the signal described by (5.1),

$$y_n = s_n + v_n, \quad n = 1, \ldots, M \tag{5.2}$$

where v_n is complex Gaussian additive noise and M is the number of echoes. This noise model may not be suitable for advanced reconstruction algorithms such as those used in undersampled parallel imaging. We can derive all existing signal models and T_2 estimation algorithms from this general form in (5.1) and (5.2), with different assumptions placed on f, g_n and v_n.

Traditionally the signal from a single T_2 component, measured with a spin echo sequence, is described by a decaying exponential. That is,

$$g_n(\tau) = e^{-t_n/\tau} \tag{5.3}$$

where t_n is the echo time and τ is the T_2 time constant.

We also consider the effect of B_1 inhomogeneity on the signal model. The repeated application of RF pulses at flip angles other than the ideal 180° causes secondary and stimulated echoes that contribute to the observed signal strength. These echoes can be tracked using the extended phase graph (EPG) algorithm [71, 72]. In this case, the signal decay function, g_n, includes additional parameters for spin-lattice relaxation, T_1, and the actual flip angle, α. The distribution of different phase states at the n^{th} echo are stored in a vector, $\boldsymbol{x}_n$. The EPG algorithm describes the evolution of this vector via a linear recursive relation, from which the echo amplitudes can be extracted. The algorithm, described in detail in Appendix 5.B, can be summarised as,

$$\boldsymbol{x}_n = \boldsymbol{E}(\tau, T_1, \alpha)\boldsymbol{x}_{n-1} \tag{5.4a}$$

$$g_n = \boldsymbol{c}^T \boldsymbol{x}_n. \tag{5.4b}$$

where the vectors $\boldsymbol{c}$ and $\boldsymbol{x}_n$ and the matrix $\boldsymbol{E}$ are defined in Appendix 5.B and the superscript ^{T} denotes matrix transpose. Non-uniform slice profiles (such as those in [79]) can be considered under the same formalism by defining the appropriate function g_n.

The next section describes different approaches to distribution estimation resulting from a chosen form of g_n together with a model for the distribution, f.

5.3 Existing approaches

5.3.1 Pseudo-continuous model

The first approach considered is to represent a continuous distribution by values on a grid over the parameter space. In this case, a large but known sequence of relaxation times is defined, $\hat{\tau}_1, \ldots, \hat{\tau}_L$, covering a physically plausible range of times. This approximation and subsequent non-negative least squares (NNLS) algorithm was proposed in [5]. It has been applied extensively for the case of ideal flip angles by defining g_n as shown in (5.3), e.g. [73, 74, 75, 76]. Mathematically, the distribution is made up of L delta functions,

$$f(\tau) = \sum_{i=1}^{L} w_i \delta(\tau - \hat{\tau}_i). \tag{5.5}$$

This pseudo-continuous model is considered to be non-parametric due to the minimal assumptions about the distribution shape.

An extension was proposed in [77] to account for stimulated echoes arising from non-ideal flip angles. In this case, g_n is defined by the EPG algorithm in (5.4).

All these works take the magnitude of the measurements and implicitly model the phase, ϕ, as zero. The Rician noise can be approximated as Gaussian for sufficiently high SNR [81]. Alternatively, an unknown offset can be included in the signal equation in an attempt to correct for any discrepancy [76].

The signal equation for the pseudo-continuous model is obtained by combining (5.1), (5.2), (5.4) and (5.5) to give

$$y_n = \sum_{i=1}^{L} w_i g_n(\hat{\tau}_i, \hat{T}_1, \hat{\alpha}) + v_n. \tag{5.6}$$

Given values of $\hat{T}_1$ and $\hat{\alpha}$, the EPG algorithm gives the expected signal contribution at a grid point $\hat{\tau}_i$. A vector of the measurements, $\boldsymbol{y} = [y_1, \ldots, y_M]^T$ and weights, $\boldsymbol{w} = [w_1, \ldots, w_L]^T$ can be constructed and (5.6) can be rewritten as the

matrix equation,

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{w} + \boldsymbol{v} \tag{5.7}$$

where $\boldsymbol{A}$ is an $M \times L$ matrix with elements, $A_{n,i} = g_n(\hat{\tau}_i, \hat{T}_1, \hat{\alpha})$. The problem of estimating the distribution is reduced to estimating the vector of weights, $\boldsymbol{w}$. The number of measurements is often considerably smaller than the number of weights to estimate ($M < L$) and the system is underdetermined. Consequently, a least squares estimator that solves (5.7) will have poor noise performance. The solution proposed in [5] is to regularise the optimisation and solve,

$$\begin{aligned} &\text{minimise} \quad \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{w}\|_2^2 + \lambda \|\boldsymbol{C}\boldsymbol{w}\|_2^2 \\ &\text{subject to} \quad \boldsymbol{w} > 0 \end{aligned} \tag{5.8}$$

where $\boldsymbol{C}$ contains additional constraints (such as smoothness), weighted by the regularisation parameter, λ. This problem can be solved efficiently using the non-negative least squares (NNLS) algorithm and often the term NNLS is used to refer to the discrete grid model in (5.5) together with the optimisation in (5.8). In this chapter, the regularisation parameter, λ, is selected to give a χ^2 misfit approximately 1% greater than the minimum misfit (when $\lambda = 0$) as in [82].

It is not practical to estimate T_1 directly from the echo data since the signal is only weakly dependent on the T_1 value. For example, a typical two component decay curve changes by less than 1% for T_1 values above 100 ms. Thus one can confidently fix $\hat{T}_1 = 1\,\text{s}$ (as in [77]) or $\hat{T}_1 = \infty$ (as in [79]) with minimal effect on the estimation results. All results in this chapter assume $\hat{T}_1 = 1\,\text{s}$.

The flip angle $\hat{\alpha}$ is estimated in a separate optimisation stage, which minimises the sum of squared errors between the measured and predicted signals, as in [77].

5.3.2 Discrete model with unknown locations

The approach proposed in [7] assumes the distribution is made up of a small number of discrete components. This assumption is motivated by the nature of biological tissue, which can consist of myelinated water, inter/extra-cellular water and cerebrospinal fluid (CSF) each with a different relaxation time. Consequently, the distribution is modelled by a small number of weighted delta functions with unknown weights and locations,

$$f(\tau) = \sum_{i=1}^{N} w_i \delta(\tau - \tau_i). \tag{5.9}$$

Two signal equations can be derived depending on whether or not the assumed decay function g_n includes stimulated echoes. The case of ideal 180° pulses yields an exponential decay function for each T_2 value. The corresponding integral in (5.1) becomes

$$y_n = e^{j\phi} \sum_{i=1}^{N} w_i e^{-t_n/\tau_i} + v_n. \tag{5.10}$$

To include stimulated echo compensation, the EPG algorithm is employed. The signal can still be described by a weighted sum of the decay functions but in this case the flip angle must be estimated along with the weights, locations and phase.

$$y_n = e^{j\phi} \sum_{i=1}^{N} w_i g_n(\tau_i, \hat{T}_1, \alpha) + v_n \tag{5.11}$$

This equation is fundamentally different from the non-parametric form in (5.6), since the times τ_i are unknown and must be estimated along with the weights. In this case, the estimation problem is nonlinear, unlike the linear problem in (5.6). Accurate estimation with this nonlinear signal model is only viable when the distribution has a small number of modes, i.e., $N \ll M$. Although the equation in (5.11) is a natural extension of [8] and [79], it has not yet been formally proposed in the literature for T_2 distribution estimation.

Typical solutions to the estimation problems above attempt to find the maximum likelihood estimate (MLE) [6]. Magnitude data is commonly used, which is modelled by assigning $\phi = 0$, although the idea can be extended to the signal equations in (5.10) and (5.11). For Gaussian noise with known variance, maximising the likelihood is equivalent to the following optimisation problem,

$$\arg\min_{\boldsymbol{\theta}} \sum_{n=1}^{M} \left(y_n - \sum_{i=1}^{N} w_i g_n(\boldsymbol{\theta}) \right)^2 \tag{5.12}$$

where $\boldsymbol{\theta}$ is a vector of unknown parameters, depending on the chosen decay function. We will see that even for a simple signal equation, a naive MLE optimisation algorithm provides unreliable results. Instead, we develop a Bayesian algorithm that produces near-optimal estimates.

Table 5.2 summarises the different signal equations presented in this chapter. It demonstrates that all signal equations are derived from the integral equation in (5.1), with different assumptions on the distribution, f and the decay function, g_n.

5.4 A novel parametric and continuous model

Neither of the models presented in the previous section can represent a realistic distribution in a parsimonious way. The discrete grid (NNLS) can model an arbitrary continuous distribution given a sufficiently large number of grid points, however a grid does not represent smoothness in a parsimonious way. Although smoothness can be enforced using regularisation, many of the grid points will become superfluous. On the other hand, the discrete model with a small number of unknown locations is a simple parametric model but it cannot model the continuous nature of realistic distributions. In this section we present a third model for T_2 distributions that is both parametric and continuous. We then analyse the estimation performance of the various models.

The alternative model considered in this chapter is the estimation of a mixture of parametric distributions. We assume the distribution consists of a small number of modes, each with an inverse-gamma distribution. The inverse-gamma distribution is particularly well-suited to our problem since: it can approximate a wide range of distributions by appropriate selection of the parameters; it has positive support, which is suitable for relaxation times; and it leads to a tractable integration in (5.1), important for analysis and the development of efficient estimation algorithms [6]. The inverse-gamma mixture is given by

$$f(\tau) = \sum_{i=1}^{N} w_i \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \tau^{-\alpha_i - 1} e^{-\beta_i/\tau}. \tag{5.13}$$

The three parameters, w_i, α_i, β_i, characterise the weight, location and scale of the i^{th} mode, respectively. Figure 5.1 illustrates the shape of the inverse-gamma distribution for different values of α and β. Importantly, both parameters are required to uniquely specify the centre and width of the distribution.

We assume the ideal exponential decay function for g_n given in (5.3), and together with the inverse-gamma model, this leads to a closed-form signal model. Substituting the parametric form (5.13) into the signal model (5.1) gives,

$$y_n = e^{j\phi} \sum_{i=1}^{N} w_i \left(\frac{\beta_i}{t_n + \beta_i} \right)^{\alpha_i} + v_n. \tag{5.14}$$

See Appendix 5.A for the proof of (5.14). Although the ideal exponential decay is used for our analysis, this result can be extended to the EPG algorithm.

Analogous to the discrete model, imposing the inverse-gamma mixture on the distribution shape reduces the estimation problem to finding $3N$ parameters. As

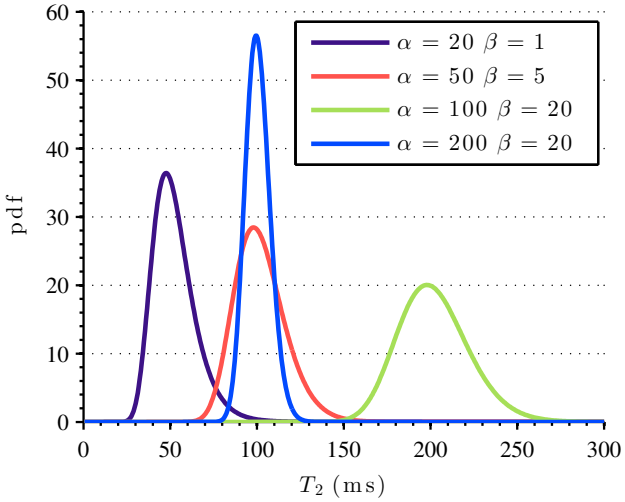


Figure 5.1: Examples of inverse-gamma distributions for different values of the location parameter, α, and scale parameter, β.

Table 5.2: List of signal models and corresponding assumptions

Name	Distribution Model, f	g_n	s_n	Ref.
NNLS_α	grid	EPG	$\mathbb{R}$	[77]
$\text{Discrete}_{180^\circ}$	discrete	exp	$\mathbb{R}$	[7, 8]
Discrete_α	discrete	EPG	$\mathbb{R}/\mathbb{C}$	[79]
Gamma	inv gamma mixture	exp	$\mathbb{C}$	

such, this model can be seen as a parametric alternative to NNLS. It is instructive to consider an alternative parameterisation of each mode in the mixture, specified by the mean and variance of the inverse-gamma distribution, given by

$$\nu_j = \frac{\beta_i}{\alpha_i - 1}; \qquad \rho_j^2 = \frac{\beta_i^2}{(\alpha_i - 1)^2(\alpha_i - 2)}. \tag{5.15}$$

This parametrisation is useful to analyse the fundamental ability to estimate the location and spread of the relaxation times.

5.5 Estimation of continuous distributions

5.5.1 CRLB analysis

The minimum variance of an unbiased estimator is given by the Cramér-Rao lower bound (CRLB) [6]. This bound is often used as a benchmark to assess suboptimal estimation algorithms. In this section, we use the bound more directly to anal-

yse the intrinsic uncertainty in the estimation problems for different distribution models. The CRLB reveals the theoretical limit on variance of estimated model parameters for a given experimental setup. Specifically the following bound applies to the covariance of an unbiased estimator,

$$\operatorname{cov} \hat{\boldsymbol{\theta}} \geq \boldsymbol{F}^{-1}(\boldsymbol{\theta}) \tag{5.16}$$

where $\hat{\boldsymbol{\theta}}$ denotes an estimator and $\boldsymbol{F}$ is the Fisher information matrix (FIM). The FIM has elements given by

$$F_{i,j}(\boldsymbol{\theta}) = \mathbb{E}\left[\left(\frac{\partial}{\partial\theta_i}\log\ell(\boldsymbol{\theta}|\boldsymbol{y})\right)\left(\frac{\partial}{\partial\theta_j}\log\ell(\boldsymbol{\theta}|\boldsymbol{y})\right)\right]. \tag{5.17}$$

where ℓ is the likelihood function for the given the signal model.

In this work we calculate the CRLB for three different models: the inverse-gamma model in (5.14) with unknown location and width parameters; a constrained inverse gamma model with unknown location and known width; and the discrete distribution model in (5.10). These models are representative of the different approaches to distribution estimation.

The inverse-gamma mixture is ideal to consider the estimation performance in terms of the location and width of the unknown distribution components, since it is the only parametric model to incorporate width. To this end, we use the relationships in (5.15) to transform the FIM for this model (initially in terms of α_i and β_i) according to

$$\boldsymbol{F}(\nu,\rho) = \boldsymbol{J}^T\boldsymbol{F}(\alpha,\beta)\boldsymbol{J} \tag{5.18}$$

where $\boldsymbol{J}$ is the Jacobian matrix of the mapping associated with (5.15).

We evaluate the CRLB for an inverse gamma mixture distribution consisting of two modes defined using (5.15): a slow mode with location $\nu_1 = 100\,\text{ms}$ and width $\rho_1 = 10\,\text{ms}$ and a fast mode with location $\nu_2 = 20\,\text{ms}$ and width $\rho_2 = 10\,\text{ms}$. The CRLB for the discrete model signal in (5.10) was calculated for two modes with equivalent locations, $\tau_1 = 100\,\text{ms}$ and $\tau_2 = 20\,\text{ms}$. Both models had weights, $w_1 = 0.7$ and $w_2 = 0.3$ for the slow and fast modes, respectively. These values were chosen to emulate a voxel with a dominant intra/extra cellular component and a small myelin water component. The simulated MRI sequence is a multi-echo sequence consisting of $M = 32$ spin echoes with echo times spaced 12 ms apart, i.e. $t_n = (12n)\,\text{ms}, n = 1,\ldots,32$.

Fig. 5.2 displays the CRLB for the parameters of the first mode of the multicom-

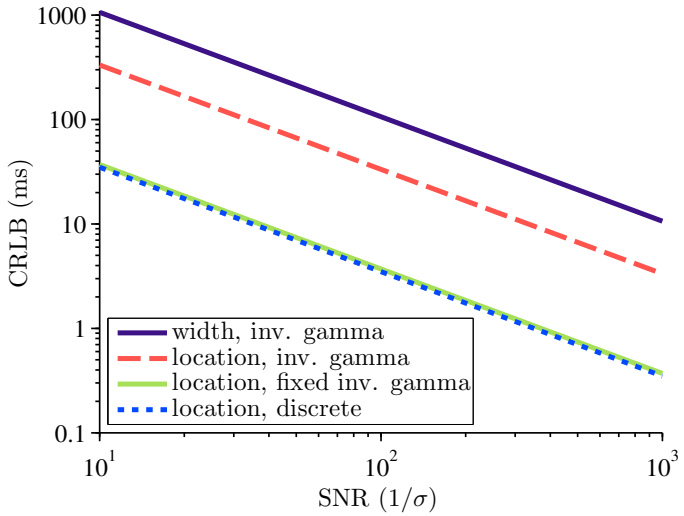


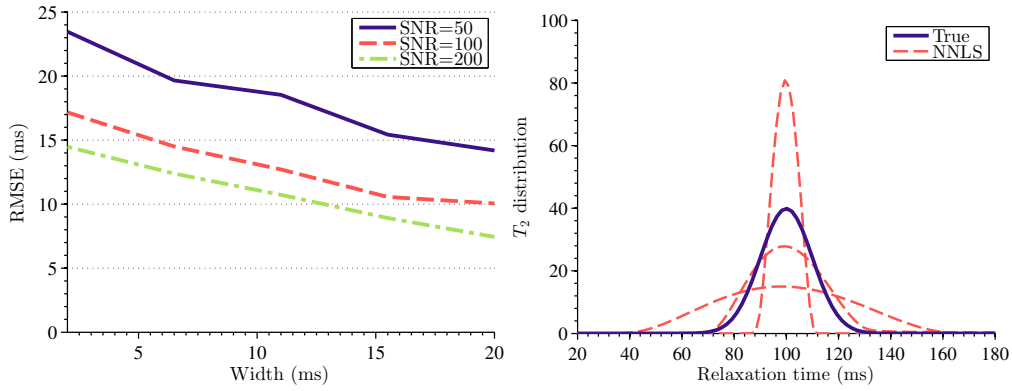
Figure 5.2: The estimation bounds for different parametric distributions. The CRLB for the slow mode parameters is displayed for the width (blue solid line) and location (red dashed line) of the inverse gamma model. For comparison, the CRLB of the location parameter of a fixed-width inverse gamma model (green solid line) and discrete model (blue dotted line) are displayed.

ponent distribution using different models. The plot demonstrates that the width parameter of the inverse gamma model is exceedingly difficult to estimate. For an SNR of 100, the width of the slow mode can only be estimated with a standard deviation of $\sim$100 ms, five times greater than the true width parameter. These results also highlight that estimating the location is much harder when the width is unknown.

To achieve useful estimates of the weight, location and width parameters of the distribution, the SNR would need to be $>$1000, beyond that achievable in a clinical setting. Alternatively we would need to collect in the order of 10^4 echoes, which is completely impractical. Similar plots for the second mode or different parameter values yield the same conclusions: it is not possible to reliably estimate the width of the relaxation time distribution. The next section provides empirical evidence that this result is not specific to an inverse-gamma distribution and holds for the pseudo-continuous model used in NNLS.

5.5.2 NNLS and distribution width

The inability of the available data to provide information about distribution width is a fundamental property which also affects NNLS. To investigate this further, a signal was simulated from a single component Gaussian distribution with mean



(a) (b)

Figure 5.3: Demonstration of the ability of NNLS to estimate distribution width of a single component. The (a) root mean squared error (RMSE) of the width estimates extracted from NNLS distributions for SNR of 50, 100 and 200, and (b) the true (solid line) and three example distributions (dashed line) estimated from different noise realisations for SNR of 100.

$T_2 = 100$ ms and five different distribution widths evenly spaced from 2 ms to 20 ms. For each signal, independent noise was added and NNLS was performed. The distribution width was extracted by fitting a Gaussian to the generated pseudo-continuous NNLS distributions. This was repeated for 1000 independent noise trials to calculate the mean squared error. Fig. 5.3a displays the root mean squared error (RMSE) of the width estimates, which demonstrate the NNLS distributions do not accurately represent the true distribution width. Three example NNLS distributions are shown in Fig. 5.3b for a width of 10 ms and SNR of 100. These results were produced for a single component distribution and even weaker results are obtained for multicomponent distributions. Although estimation performance improves considerably as the SNR increases, the width estimates are not sufficiently reliable for typical data.

The reason for such poor estimation performance is that a large number of distributions will produce very similar measurements. For reasonable noise levels, the differences are indistinguishable and we cannot discern the correct model width. For example, Fig. 5.4 displays the signal on a log-scale from two-component gamma distributions with locations, $\tau_1 = 20$ ms, $\tau_2 = 100$ ms and width parameters of 10 ms, 30 ms and the limiting case of a discrete distribution. This figure highlights the small difference between signals generated from vastly different width parameters.

Since the measurements do not contain sufficient information to estimate the distribution width, we must constrain the width estimates or adopt a discrete dis-

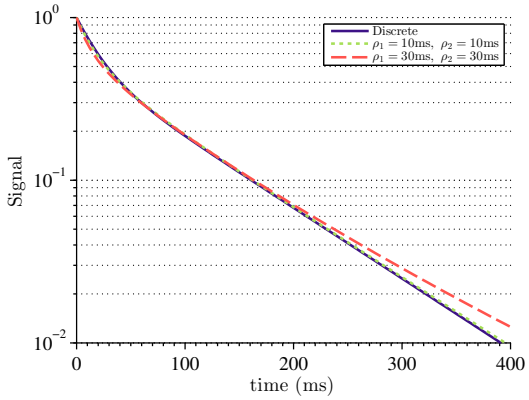


Figure 5.4: The noiseless signal (log-scale) generated from a two-component discrete distribution and a gamma distribution with widths of 10 ms and 30 ms. There is minimal difference between the discrete distribution (solid) and a continuous distribution with width of 10 ms (dotted). A notable difference is present when the width increases threefold to 30 ms (dashed).

tribution model that doesn't model it. We adopt the latter approach since the CRLB results in Figure 5.2 suggest adverse performance of location estimation when the width is unknown. We will see that the discrete distribution is suitable for estimating the main contributions to the signal. Existing estimation techniques using the discrete distribution model are based on the MLE and have not demonstrated satisfactory performance. This motivated the development of a Bayesian algorithm.

5.6 Proposed estimation algorithm for the discrete model

The multicomponent signal model in (5.11) does not lend itself well to a simple estimation algorithm despite the seeming simplicity of the model. The nonlinear relationship between the parameters and the signal creates a poorly behaved MLE cost function, with local minima and large regions in parameter space where the cost function is essentially flat. These features are problematic for a naive gradient-based optimisation algorithm.

We adopt a Bayesian framework that leads to a numerically robust algorithm and allows us to incorporate prior information about the biological tissue. The cornerstone equation is Bayes' rule,

$$\pi(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{\ell(\boldsymbol{\theta}|\boldsymbol{y})\pi_0(\boldsymbol{\theta})}{\pi(\boldsymbol{y})} \tag{5.19}$$

where $\boldsymbol{y}$ is a vector of measurements, $\boldsymbol{\theta}$ is a vector of unknown parameters, $\ell(\cdot)$

is the likelihood and $\pi_0(\cdot)$ is the prior distribution. The discrete complex model has two parameters for each distribution component in addition to a phase and flip angle. Although we ultimately require estimates in terms of the relaxation time constants, τ, our implementation estimates the relaxation rates, $r = 1/\tau$. This is particularly useful to avoid the large signal discontinuity about $\tau = 0$, which can be problematic when estimating small time constants from noisy data. Consequently, the unknown parameters are

$$\boldsymbol{\theta} = [w_1, r_1, \ldots, w_N, r_N, \phi, \alpha]^T. \tag{5.20}$$

It is not possible to compute an exact, finite-dimensional representation of the posterior (5.19) so approximations must be sought. Finding an accurate but low-dimensional approximation is complicated by the relatively wide prior, π_0, and narrow likelihood, ℓ. Progressive correction is used to overcome this difficulty [83]. Although we do not use a Monte Carlo approximation, we apply the same principle of 'flattening' the likelihood and iteratively correcting our estimate of the posterior. This approach is very computationally efficient compared to Monte Carlo techniques and demonstrates improved accuracy compared to simple linearisation schemes. We define a schedule of P corrections, $\gamma_1, \ldots, \gamma_P$, with the intermediate posterior at the j^{th} correction step given by

$$\pi_j(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{\ell^{\gamma_j}(\boldsymbol{\theta})\pi_{j-1}(\boldsymbol{\theta}|\boldsymbol{y})}{\eta_j} \tag{5.21}$$

where η_j is a normalising constant. When $\sum_j \gamma_j = 1$, the final posterior π_P is the required one defined in (5.19).

For small γ_j, it is much easier to accurately approximate (5.21) than (5.19). In particular, a Gaussian approximation, obtained by linearising the likelihood, can be used. This approximation would not be accurate using (5.19) directly. The approximate likelihood is obtained by linearising the nonlinear function that describes the EPG signal. Recall the signal in (5.11) is,

$$s_n(\boldsymbol{\theta}) = e^{j\phi} \sum_{i=1}^{N} w_i g_n(\tau_i, \hat{T}_1, \alpha). \tag{5.22}$$

Importantly, we use the raw complex-valued data to jointly estimate the phase and relaxation times. This has the advantage of the noise remaining Gaussian instead of Rician as is the case when magnitude data is used [81], provided there is minimal phase artefacts from physiological noise and reconstruction. The Jacobian

for the signal in (5.22) is

$$\boldsymbol{J}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial s_1}{\partial w_1} & \frac{\partial s_1}{\partial r_1} & \dots & \frac{\partial s_1}{\partial w_N} & \frac{\partial s_1}{\partial r_N} & \frac{\partial s_1}{\partial \phi} & \frac{\partial s_1}{\partial \alpha} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial s_M}{\partial w_1} & \frac{\partial s_M}{\partial r_1} & \dots & \frac{\partial s_M}{\partial w_N} & \frac{\partial s_M}{\partial r_N} & \frac{\partial s_M}{\partial \phi} & \frac{\partial s_M}{\partial \alpha} \end{bmatrix} \tag{5.23}$$

The partial derivatives can be calculated efficiently using a set of recursive relationships set out in Appendix 5.C.

We adopt a Gaussian prior with mean, $\boldsymbol{\mu}_0$, and covariance $\boldsymbol{\Psi}_0$. Further, π_{j-1} is assumed to be Gaussian with mean $\boldsymbol{\mu}_{j-1}$ and covariance matrix $\boldsymbol{\Psi}_{j-1}$. The approximate likelihood is

$$\hat{\ell}^{\gamma_j}(\boldsymbol{\theta}) = \mathcal{N}\left(\boldsymbol{y}; \boldsymbol{s}(\boldsymbol{\mu}_{j-1}) + \boldsymbol{J}(\boldsymbol{\mu}_{j-1})(\boldsymbol{\theta} - \boldsymbol{\mu}_{j-1}), \boldsymbol{\Sigma}/\gamma_j\right) \tag{5.24}$$

where $\boldsymbol{J}$ is the Jacobian in (5.23) and $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes a multivariate Gaussian PDF with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. In this case, the prior and likelihood are linear and Gaussian at each step. This is the same situation as a Kalman filter and thus the posterior approximation is a Gaussian with mean and variance given by the Kalman filter update equations,

$$\boldsymbol{\mu}_j = \boldsymbol{\mu}_{j-1} + \boldsymbol{K}_j(\boldsymbol{y} - \boldsymbol{s}(\boldsymbol{\mu}_{j-1})), \tag{5.25a}$$

$$\boldsymbol{\Psi}_j = (\boldsymbol{I} - \boldsymbol{K}_j\boldsymbol{J})\boldsymbol{\Psi}_{j-1} \tag{5.25b}$$

where $\boldsymbol{I}$ is the identity matrix and $\boldsymbol{K}_j$ is the gain at the j^{th} iteration defined as

$$\boldsymbol{K}_j = \boldsymbol{\Psi}_{j-1}\boldsymbol{J}'(\boldsymbol{J}\boldsymbol{\Psi}_{j-1}\boldsymbol{J}' + \boldsymbol{\Sigma}/\gamma_j)^{-1}. \tag{5.26}$$

Initially, the linearisation will be inaccurate but the likelihood will be wide due do severe flattening, so the approximation of π_j remains accurate. As the algorithm progresses, the approximate posterior approaches the true posterior and the linearisation becomes more accurate. This process is very similar to simulated annealing [84] and we will see that the resulting algorithm is very robust to local minima, particularly at low SNR. The final algorithm is described in Algorithm 5.1. The algorithm is computationally efficient as matrix inversions can exploit the diagonal nature of the measurement covariance.

Algorithm 5.1: Estimation Algorithm

Input: Data, $\boldsymbol{y}$; Covariance, $\boldsymbol{\Sigma} = \sigma^2 I$; Prior, $\boldsymbol{\mu}_0$, $\boldsymbol{\Psi}_0$; Correction schedule, $\gamma_1, \ldots, \gamma_P$

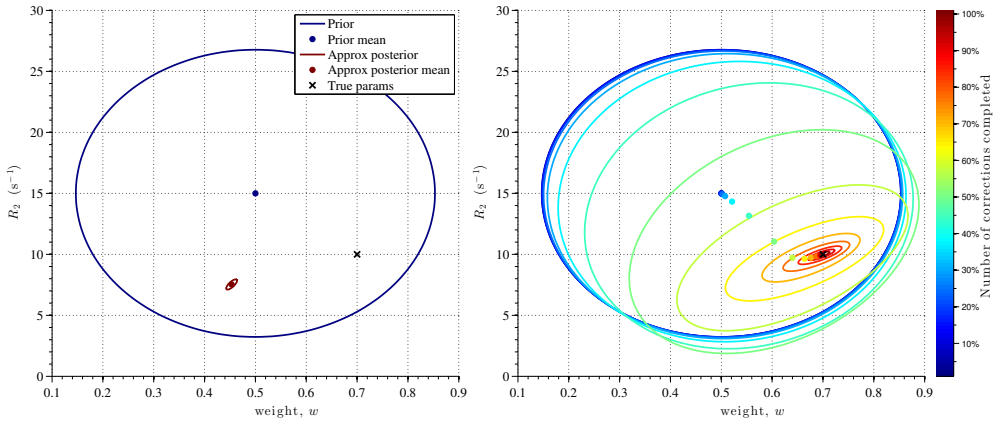
Output: Estimate, $\boldsymbol{\mu}_P$

for $j \leftarrow 1 \ldots P$ **do**
- Calculate Jacobian, $\boldsymbol{J} \leftarrow \boldsymbol{J}(\boldsymbol{\mu}_{j-1})$
- Calculate gain, $\boldsymbol{K}_j \leftarrow \boldsymbol{\Psi}_{j-1}\boldsymbol{J}'(\boldsymbol{J}\boldsymbol{\Psi}_{j-1}\boldsymbol{J}' + \boldsymbol{\Sigma}/\gamma_j)^{-1}$
- Update mean, $\boldsymbol{\mu}_j \leftarrow \boldsymbol{\mu}_{j-1} + \boldsymbol{K}_j(\boldsymbol{y} - \boldsymbol{s}(\boldsymbol{\mu}_{j-1}))$
- Update covariance, $\boldsymbol{\Psi}_j \leftarrow (\boldsymbol{I} - \boldsymbol{K}_j\boldsymbol{J})\boldsymbol{\Psi}_{j-1}$

5.6.1 Demonstration of algorithm

The operation of this algorithm is demonstrated with an example consisting a single discrete component with weight, $w = 0.7$, and relaxation time, $\tau = 100$ ms. The aim is to estimate these two values from 32 echo measurements spaced 12 ms apart. The prior distribution is relatively wide reflecting minimal knowledge about the true parameter values. We test two correction schedules: firstly we perform a single correction ($\gamma_1 = 1$, $P = 1$) equivalent to Bayesian estimation using the standard Extended Kalman Filter equations. The second schedule consists of $P = 30$ corrections, logarithmically spaced between 10^{-6} and 1, i.e., $\gamma_i = 10^{6i/29-6}, i = 0, \ldots, 29$ and normalised such that $\sum_j \gamma_j = 1$. The logarithmic spacing ensures that small corrections are made initially, when the linearised approximation can be poor, and progressively larger corrections are performed as the estimates becomes more accurate.

Figure 5.5a displays error ellipses for the prior and the approximate posterior derived using the linearised likelihood and a single correction step. The linearisation occurs about the prior mean, which, in this case, is too far from the true parameters to provide an accurate likelihood approximation. The resulting posterior approximation is poor and not representative of the true parameters. Figure 5.5b illustrates the progressive corrections from the prior (at 0%) to the posterior (at 100%). In this case, the flattening of the likelihood results in small updates to the posterior approximation indicated by the evolution of the error ellipse. The algorithm mitigates the difficulties of using a wide prior with a narrow likelihood function and the final posterior approximation is very close to the true parameters.



(a) (b)

Figure 5.5: Error ellipses for the prior and posterior distributions demonstrating the proposed algorithm using (a) a single correction and (b) a schedule of 30 corrections. The true parameters are marked with a $\times$ and the distribution mean is indicated by a $\bullet$. The colours reflects the number of corrections performed where 0% gives the prior and 100% gives the posterior approximation.

5.6.2 Simulations

In this work we define a Gaussian prior for two components with mean

$$\boldsymbol{\mu}_0 = [0.7 \;\; 0.1^{-1} \;\; 0.3 \;\; 0.02^{-1} \;\; 0 \;\; 2\pi/3]^T, \tag{5.27}$$

and covariance

$$\boldsymbol{\Psi}_0 = \begin{bmatrix} 1^2 & & & & & & \\ & 100^2 & & & & & \\ & & 0.2^2 & & & & \\ & & & 100^2 & & & \\ & & & & 35^2 & & \\ & & & & & 35^2 & \end{bmatrix}. \tag{5.28}$$

Recall the parameter vector in (5.20) is composed of weights, relaxation rates, a phase angle and a flip angle. The prior is relatively wide such that the estimation bias is negligible. A similar prior can be created for three or more components. The correction schedule is the same as above and consists of 30 corrections, logarithmically spaced between 10^{-6} and 1 and normalised such that $\sum_j \gamma_j = 1$.

To examine the accuracy of different estimation algorithms for varying levels of SNR, a decay curve was generated from a multicomponent distribution consist-

ing of two discrete modes: a slow component with weight 0.7 and T_2 of 100 ms, and a fast component with weight 0.3 and T_2 of 20 ms. These general features are typical of white matter tissue in the cortex. A flip angle of 160° was simulated to represent a non-ideal B_1 field. The signal was calculated from the extended phase graph (EPG) algorithm to model secondary and stimulated echoes as described by (5.11) and (5.4). Simulations consisted of 2000 independent trials where i.i.d. complex noise was added. The phase angle in (5.11) was randomly selected each trial from a uniform distribution between $-\pi$ and π to demonstrate that the algorithm is insensitive to phase. The relaxation modes were estimated using three approaches: a discrete grid model with NNLS including stimulated echo compensation, as in [77]; a discrete model with a gradient based MLE optimisation similar to [79] extended to multiple components; and the discrete model using the proposed Bayesian algorithm. The iterative MLE algorithm was initialised at the prior mean in (5.27). To compare the continuous NNLS distributions, component locations were extracted using the geometric mean between 0 ms and 40 ms for the fast component and 40 ms and 200 ms for the slow component [75]. The MSE was calculated and the process was repeated for 30 SNR values logarithmically spaced between 10 and 1000. Average execution times for the three algorithms were calculated from all trials.

The ability to estimate a wide range of flip angles was examined by another simulation of 2000 trials with an SNR of 300 (as in [77]) for 30 flip angles evenly spaced between 60° and 180°. The MSE was calculated at each flip angle for the three algorithms. Fig. 5.6 displays the performance of the NNLS algorithm, a gradient based MLE algorithm and the proposed Bayesian algorithm. This plot demonstrates that the MLE optimisation algorithm achieves the same performance as the Bayesian algorithm for sufficiently high SNR. This is expected for a sufficiently wide prior, when the MLE and MAP estimates converge. However, the MLE algorithm fails to produce reliable estimates at low SNR values, due to an optimisation cost function that is poorly behaved (it is relatively flat with local minima). NNLS exhibits suboptimal performance, although it achieves reasonable results across a range of SNR values.

Both the proposed algorithm and NNLS are reasonably robust to flip angle variations. Fig. 5.7 illustrates that the error in the estimated flip angle increases with the true flip angle. This can be understood by noting that the derivative of the signal tends to zero as the flip angle approaches 180°. The 8-point NNLS algorithm as implemented in [77] exhibits elevated error around 80° and 160°. This is due to the algorithm testing only 8 flip angles and using spline interpolation to extract the minimum sum of squares error. A smooth curve using 64 test angles displays

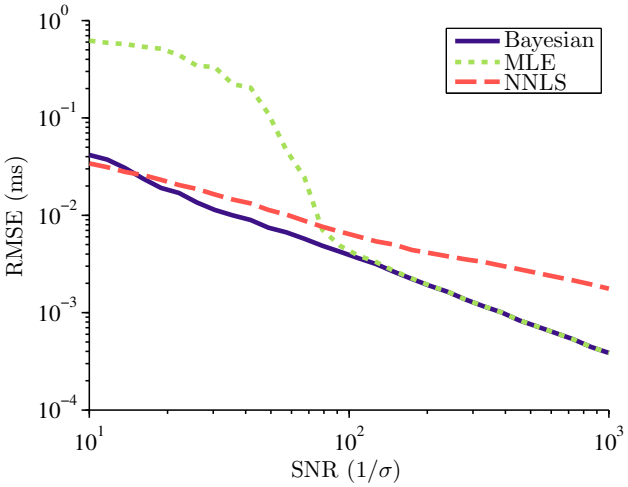


Figure 5.6: The root mean squared error (RMSE) of the estimated T_2 components calculated at different SNR values for the Bayesian algorithm (solid line), the gradient based MLE (dotted line) and NNLS (dashed line).

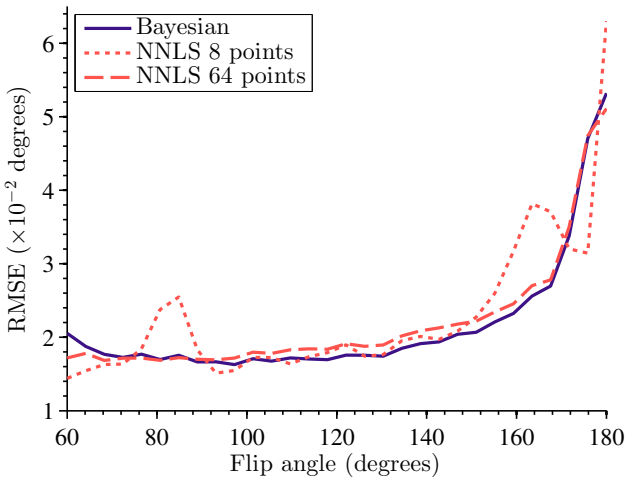


Figure 5.7: The RMSE of flip angles estimated from multiecho data with an SNR of 300. The error is plotted against different values of the true flip angle for the Bayesian algorithm (solid line) and the NNLS algorithm with 8 (dotted line) and 64 (dashed line) interpolation points.

only slightly degraded performance compared to the proposed algorithm. This modest gain in performance, however, increases the computation time of NNLS by an order of magnitude.

The Bayesian algorithm proposed above had an average execution time of 0.022 s per voxel compared to NNLS with 8-point stimulated echo correction, which averaged 0.17 s per voxel. This represents an almost eight-fold improvement.

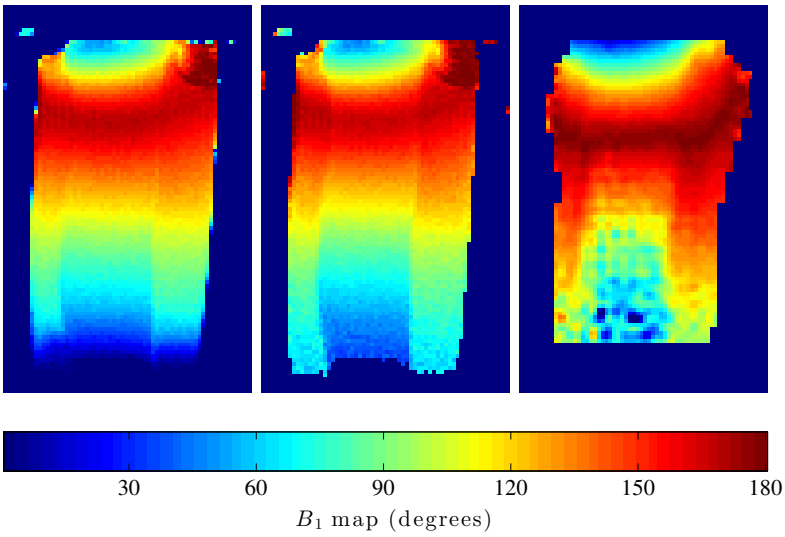
5.6.3 Optic nerve experiments

All experiments were performed on a 4.7 T Bruker BioSpec small bore MRI scanner fitted with a high performance gradient set. A cryogenically cooled surface coil was used to improve SNR. A multi-echo CPMG sequence with 24 echoes was run with a first echo time of 12 ms and an echo spacing of 12 ms. The slice thickness was 1 mm for all experiments.

A sample was constructed consisting of agar gel with a sheep optic nerve fixed parallel to the transverse plane. The TR $=$ 2500 ms, FOV $=$ 6.4 mm $\times$ 12.8 mm with a matrix size of 64 $\times$ 128. B_1 field mapping was performed by processing two spin echo images acquired with different excitation flip angles as in [85]. Data was processed with the proposed algorithm using both the exponential model, which assumes ideal 180° flip angles; and the EPG model, which jointly estimates the T_2 components, weights and B_1 map. The algorithms were run for both a single component distribution and a two component distribution.

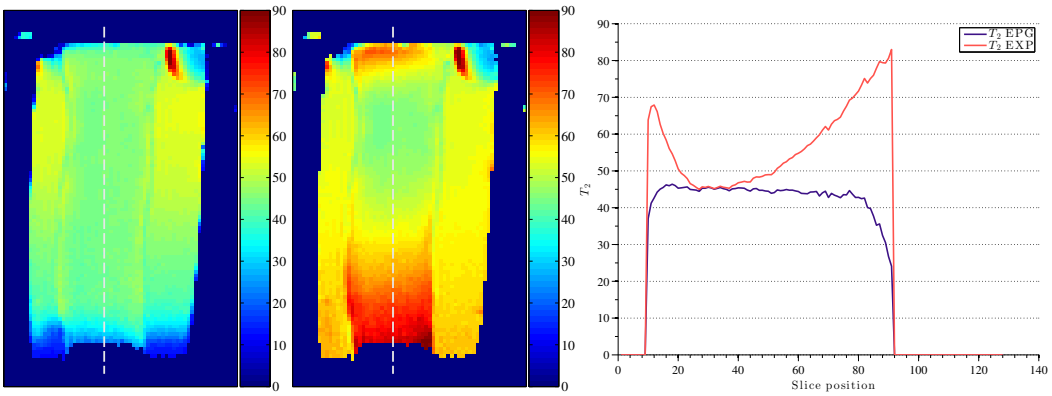
The B_1 maps for the optic nerve sample are displayed in Fig. 5.8. The B_1 map generated from the proposed estimation algorithm is consistent with that generated from a low resolution B_1 field mapping protocol. The estimated maps are very similar for a single component distribution and a two component distribution. The nature of the surface coil creates a dramatic flip angle inhomogeneity throughout the sample, with only a small region exhibiting a flip angle close to the prescribed 180°. Note that regions above the 180° arc experience a flip angle greater than 180° but are reflected about 180° as this is equivalent in the signal model.

Fig. 5.9 presents the single component T_2 maps generated with different signal models. Fig. 5.9a displays the single T_2 value estimated using the EPG algorithm, which models stimulated echoes resulting from nonideal flip angles. The estimated T_2 value is similar along the length of the optic nerve, despite the large flip angle variation demonstrated in Fig. 5.8. Conversely, Fig. 5.9b shows the results of an exponential decay model that neglects flip angle inhomogeneity. The T_2 maps in this case are overestimated by an increasing amount as the flip angle deviates from the ideal 180°. Figure 5.9c displays the T_2 profiles along a line through



(a)

Figure 5.8: The B_1 map (left) estimated using a single component model, (middle) estimated using a multi-component model and (right) measured from separate spin echo images.



(a) (b) (c)

Figure 5.9: The single component T_2 maps (ms) of an optic nerve sample estimated from (a) the extended phase graph (EPG) model and (b) the exponential decay model (EXP). Profiles along the dashed lines of the T_2 estimates are shown in (c).

the optic nerve. This demonstrates the close match between estimates when the flip angle is close to 180° and the overestimation as the flip angle decreases. This result is consistent with those previously reported in [77, 79].

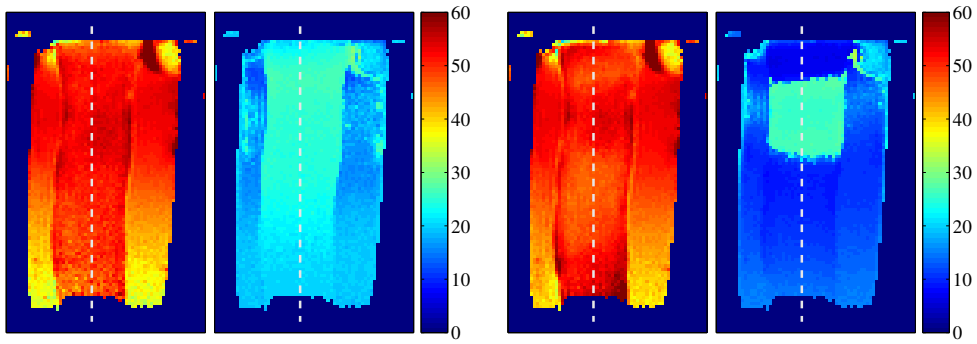
Fig. 5.10 presents the multicomponent T_2 maps generated with different signal models. Fig. 5.10a displays the two components estimated using the EPG algorithm. The estimated T_2 values are similar along the length of the optic nerve, despite the large flip angle variation. Conversely, Fig. 5.10b displays the T_2 components from an exponential decay model that neglects flip angle inhomogeneity. The T_2 maps in this case exhibit a large variation along the nerve, with the fast component fitted in an attempt fix the disparity of the first few echo amplitudes resulting from stimulated echoes. T_2 profiles along the nerve from both algorithms is displayed in Figure 5.10c. As expected the estimates are similar in regions corresponding to flip angles close to 180°. In other regions, however, the model mismatch causes large variation dependent more on the flip angle than the underlying structure.

5.6.4 Mouse brain experiments

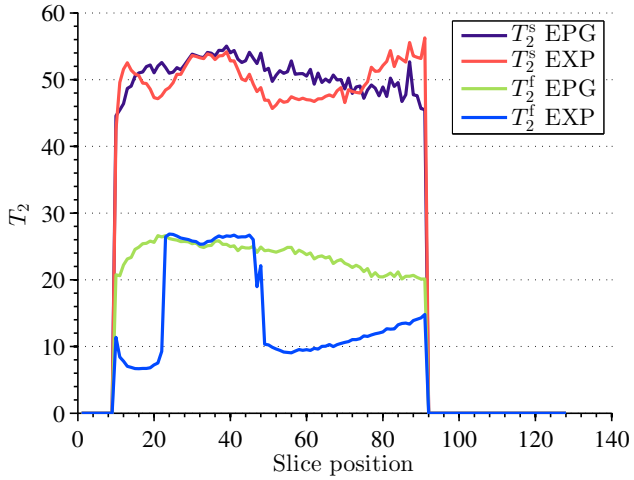
An ex-vivo mouse brain was scanned with the same 24 echo sequence with TR $=$ 2500 ms, FOV $=$ 15 mm $\times$ 15 mm and a matrix size of 192 $\times$ 192 for an in-plane resolution of 78.1 µm. A single mid-axial slice of 1 mm was acquired. This data was processed with the proposed algorithm using the EPG model and the NNLS modification in [77]. Both algorithms attempt to estimate multicomponent distributions and account for stimulated echoes by estimating both the T_2 distribution and the B_1 map. As described above, T_2 values were extracted from the NNLS distribution using the geometric mean to obtain location estimates of the fast and slow components as in [75].

To further investigate the reliability of the gradient based MLE algorithm, complex noise was added to the mouse data to yield an SNR of 200. This data was then processed with the proposed algorithm and a gradient based MLE algorithm. Fig. 5.11a displays boxplots of the T_2 value of the first component for the two algorithms. The plots depict the median, first and third quartiles, and values considered outliers. The estimates from the MLE optimisation and the Bayesian algorithm are similar, as shown by similar medians and first and third quartiles. However, approximately 6% of the MLE estimates can be considered outliers, many of which are far away from the true value. These outliers manifest as erroneous speckled values in the T_2 map, as illustrated in Fig. 5.11b.

Estimation results from the mouse brain data are displayed in Fig. 5.12. Estimated

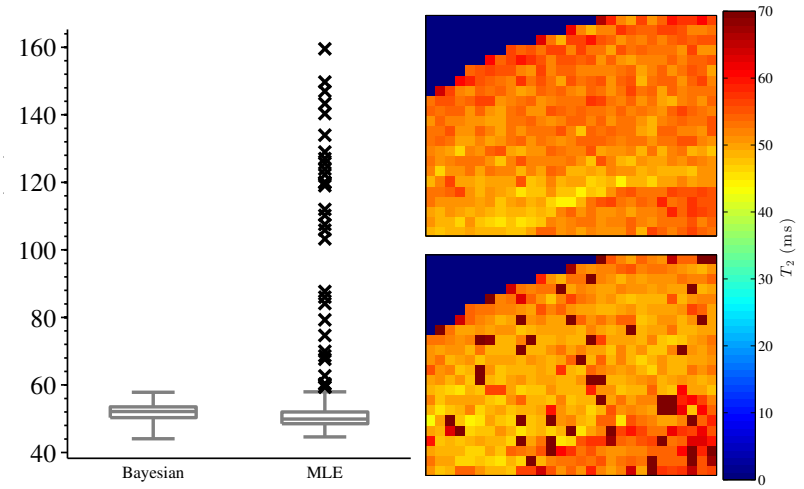


(a) (b)



(c)

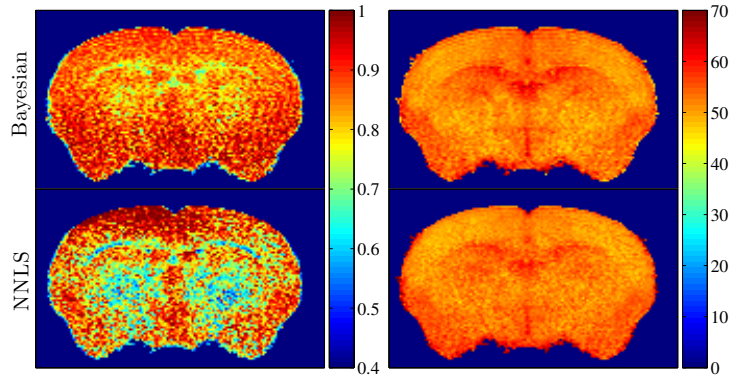
Figure 5.10: The multicomponent T_2 maps (ms) of an optic nerve sample estimated from (a) the extended phase graph (EPG) model and (b) the exponential decay model (EXP). The left and right sides display the slow and fast components, respectively. Profiles along the dashed lines of the T_2 estimates for the different components are shown in (c).



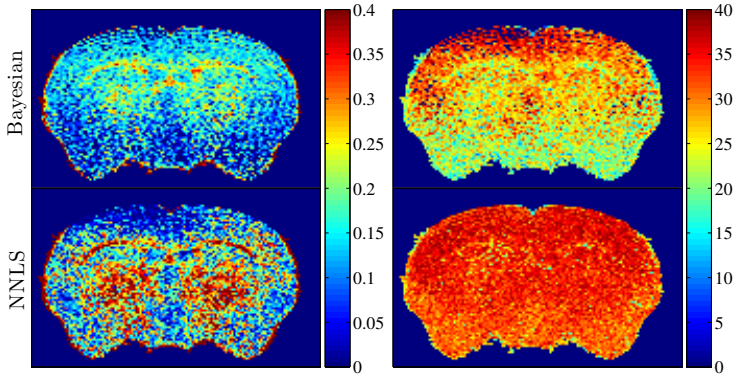
(a) (b)

Figure 5.11: Analysis of outliers from a gradient based optimisation algorithm. (a) displays a summary of the different estimates in the region of interest for the Bayesian algorithm and the gradient based MLE algorithm. (b) displays the region of interest of the mouse brain estimates for the (top) Bayesian and (bottom) MLE algorithms.

values from the proposed algorithm are shown in the top image of each figure part, while the bottom image displays values extracted from the NNLS distribution. The weights and times of the slow component (calculated using the geometric mean between 40 ms and 200 ms) are displayed in Figs. 5.12a and 5.12b, respectively. The Bayesian and NNLS algorithms produce similar T_2 maps for this component. Figs. 5.12c and 5.12d display the fast component and demonstrate an essential difference between the two algorithms. The maps produced by the proposed algorithm contain anatomical features such as a decreased T_2 below the ventricles, consistent with the presence of white matter. Fast components above 40 ms are not displayed. NNLS fast T_2 maps were calculated between 0 ms and 40 ms and exhibit a relatively constant T_2 value. NNLS appears to represent the white matter by increasing the weighting of the relatively constant T_2. This is due to the regularisation of the NNLS algorithm which decreases the estimation variance at the expense of increased bias. On the other hand, the proposed algorithm produces a fast component with both an increased weight and a decreased time constant in the white matter region. It is worthwhile to mention that both algorithms often produce fast component weights close to zero as shown in Fig. 5.12c. In this case, the T_2 estimates of the corresponding voxels are unlikely to be reliable since they have minimal contribution to the fitted signal.



(a) w_1 (b) τ_1



(c) w_2 (d) τ_2

Figure 5.12: The weights (w_1, w_2) and mean T_2 (τ_1, τ_2) of the two distribution components estimated from experimental data of a mouse brain. Each image contains results from the proposed estimation algorithm (top) and NNLS (bottom).

5.6.5 Model selection

The algorithm proposed above assumes the number of distribution components is known. Fitting too few or too many components will result in degraded performance and this could be considered the main disadvantage of the discrete distribution model. For example, the weights of the fast T_2 component were estimated close to zero in the agar gel region of the optic nerve sample. This is due to the fact that the relaxation curve of the gel is described accurately by a single decaying component. It would be preferable to only fit a single component in this case, as the second component is simply fit to the noise and has minimal contribution to the signal. There are many model selection techniques that choose the appropriate model order automatically based on a statistical metric, see e.g. [86, 87]. A simple measure known as the Bayesian Information Criteria (BIC) [88] is defined as

$$\mathrm{BIC}(\mathcal{M}) = -2\ln(\ell_{\mathcal{M}}) + k_{\mathcal{M}}\ln(M) \tag{5.29}$$

where $\ell_{\mathcal{M}}$ is maximum value of the likelihood function for model $\mathcal{M}$, $k_{\mathcal{M}}$ is the number of parameters to fit and M is the number of measurements.

Although this criteria is based on the maximum likelihood estimate (MLE) of each candidate model, we use the proposed Bayesian algorithm. Under certain non-restrictive conditions, the Bayesian estimate asymptotically converges to the MLE [89]. This justifies the use of the Bayesian estimate in place of the MLE, and the reliability it provides. We evaluate the BIC at each pixel in the optic nerve data for two models: a single component model and a two-component model, both with stimulated echo correction using the EPG algorithm. In general, the model with smallest BIC is preferred as it provides a good tradeoff between model complexity and model fit. The BIC difference between the models gives a quantitative measure of support for the two models.

Figure 5.13 displays the BIC difference between the one and two component models. In the agar gel region of the sample, there is minimal difference between the models suggesting that the data is accurately represented by a single component. Conversely, the optic nerve displays a large negative difference between the one component and two component model. This provides strong support for a two component model. As the signal strength decreases further away from the surface coil, the strength of support for a two component model decreases as the criteria cannot discern between the model and noise. Nonetheless, these results suggest that automatic model selection techniques such as the BIC may prove useful for the estimation of multi-component T_2 distributions.

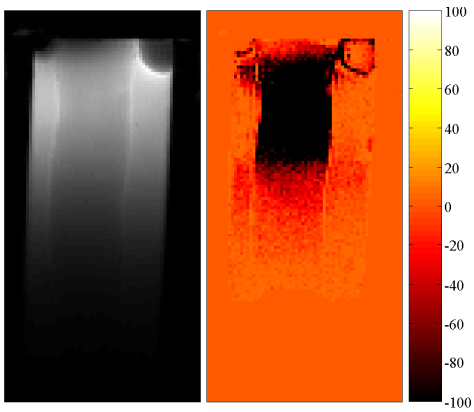


Figure 5.13: (left) Magnitude image of the optic nerve sample acquired with a surface coil and (right) the difference between Bayesian Information Criteria metrics between the two component model and the one component model. Large negative values indicate strong evidence for the two component model.

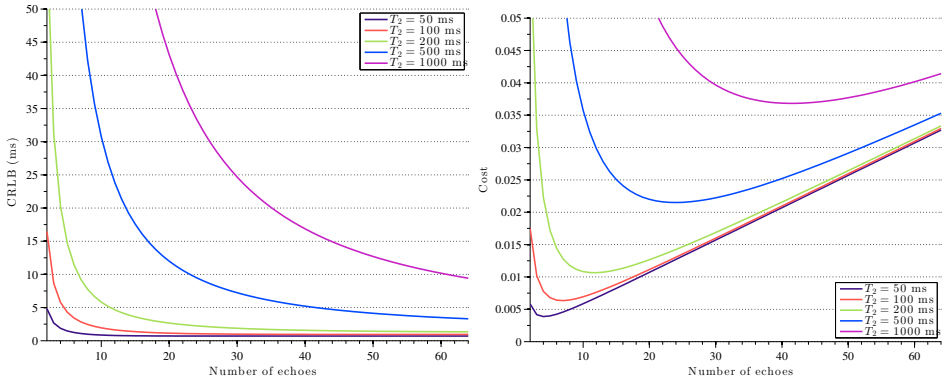
5.7 Optimal experiment design

This section addresses an often overlooked question: how to optimise the acquisition sequence for T_2 estimation. Previous work in [90] considered the joint estimation of T_1, T_2 and M_0 and optimised sequence parameters to give the minimum estimation variance. The study was limited to a single component per pixel. In this work, we consider sequence optimisation in the context of multicomponent T_2 distribution estimation using a CPMG sequence.

The CPMG sequence parameters we can control are the first echo time, t_1, echo spacing, Δt, and number of echoes, M. Unfortunately, the optimal selection of these parameters depends on the unknown T_2 components we wish to estimate. The Cramér-Rao lower bound (CRLB) was calculated in Section 5.5.1 to analyse the estimation variance for multicomponent T_2 distributions. In this section, we employ the CRLB to select the sequence parameters that give the lowest possible variance.

The CRLB is defined by the inverse of the Fisher information matrix (FIM). In general, the inverse FIM can be written as a function of the unknown T_2 components, $\boldsymbol{\theta}$, and the experimental design parameters, $\boldsymbol{\vartheta}$, denoted $\boldsymbol{F}^{-1}(\boldsymbol{\theta};\boldsymbol{\vartheta})$. Various optimisation problems can be defined to find the optimal sequence parameters, $\tilde{\boldsymbol{\vartheta}}$. For example, we may wish to minimise the worst case sum of variances over a range of T_2 components. That is,

$$\tilde{\boldsymbol{\vartheta}} = \arg\min_{\boldsymbol{\vartheta}} \left(\max_{\boldsymbol{\theta}\in\Theta} \operatorname{Tr}\left\{ \boldsymbol{F}^{-1}(\boldsymbol{\theta};\boldsymbol{\vartheta}) \right\} \right) \tag{5.30}$$



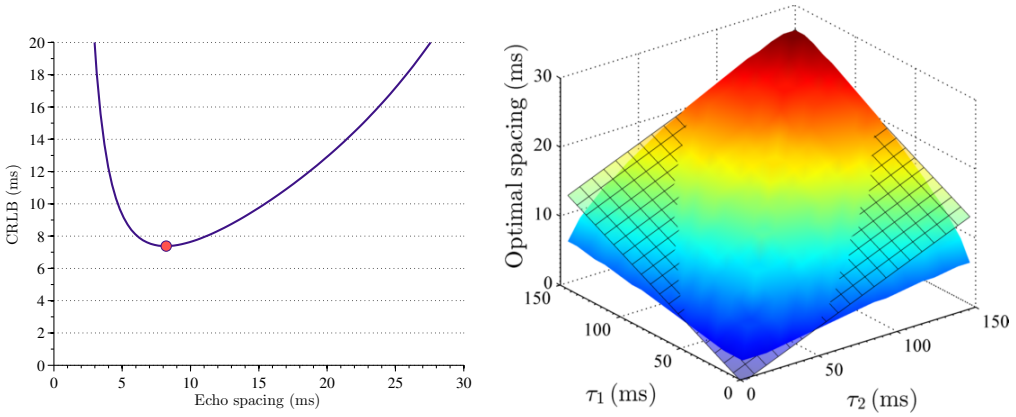
(a) (b)

Figure 5.14: (a) The CRLB displaying a monotonic performance improvement for increasing number of echoes and (b) a cost function consisting of CRLB and a linear penalty term for number of echoes. The coloured curves represent different T_2 values.

where Θ is the set of T_2 components that we expect to encounter and Tr denotes the matrix trace (equivalent to the sum of variances). For the present study, we consider a simpler problem and aim to determine the optimal sequence for a given T_2 distribution. This problem is equivalent to (5.30) when Θ contains a single element.

The first question we consider is how many echoes to acquire. For this problem we fix the first echo time and inter-echo spacing at 10 ms and optimise over the number of echoes. We consider a single component distribution with data acquired with an SNR of 100. Figure 5.14a displays the CRLB as a function of the number of echoes. The notion that acquiring echoes in the "noise floor" can degrade T_2 estimation is only applicable when magnitude data is used. The performance of the full model using complex measurements is monotonic in the number of echoes. Although the estimation performance never degrades with extra echoes, the increase in performance tends to zero as the signal becomes noise. Figure 5.14b illustrates a cost function consisting of the CRLB and a linear penalty term for the number of echoes. This penalty term could model practical factors such as an increase in the RF power deposition for additional echoes. In this case, the optimal number of echoes is finite and increases as the T_2 value increases. For example, with our chosen penalty, 10 echoes is optimal to estimate a single component with $T_2 = 200$ ms while 22 echoes may be preferred to estimate a component with $T_2 = 500$ ms. Results will vary strongly depending on the choice of penalty so care must be taken in selecting an appropriate penalty for a given application.

The second sequence parameter we optimise is the echo spacing. We assume



(a) (b)

Figure 5.15: (a) The CRLB for different echo spacings for a two component distribution with $\tau_1 = 20\,\text{ms}$ and $\tau_2 = 100\,\text{ms}$. The optimal spacing is marked with a circle. (b) displays the optimal echo spacing as a function of the two components. A linear fit to this function is displayed as a grid overlay.

the number of echoes is fixed at 32, and we aim to determine the optimal echo spacing for different T_2 distributions. In this case, we adopt a brute force approach and calculate the optimal spacing for all pairs of T_2 components. Figure 5.15a illustrates this optimisation for a single pair representing a two component distribution with $\tau_1 = 20\,\text{ms}$ and $\tau_2 = 100\,\text{ms}$. The optimal echo spacing in this case must trade the estimation accuracy between fast and slow components. The minimum estimation variance is obtained when the echo spacing is $\sim$8 ms. We the repeat this optimisation for all practical values of the fast and slow components of a multi-component distribution. Figure 5.15b depicts the optimal echo spacing as a function of the two distribution components. This function is reasonably well approximated by a plane shown as a grid overlay. The least squares fitting of this plane to the optimal function gives the following approximate relationship,

$$(\Delta t)_{\text{opt}} \approx 0.1(\tau_1 + \tau_2) \tag{5.31}$$

This relationship is useful to determine the optimal echo spacing (given a total of 32 echoes) for accurate estimation of two given T_2 components. A similar analysis could be performed for non-uniform echo spacing although this increases size of the parameter space to optimise.

This section has explored two examples of optimal experiment design using the Cramér-Rao lower bound as a cost function. This demonstrates that experiment design should be considered jointly with the problem of T_2 distribution estimation.

5.8 Discussion and conclusion

The estimation of T_2 distributions from CPMG data is a challenging problem. In this chapter, we have developed a parametric and continuous model for multicomponent T_2 distribution analysis and used it to show that it is infeasible to accurately estimate the width of a continuous distribution. Instead we have considered a discrete distribution together with the extended phase graph (EPG) algorithm to account for nonideal flip angles and the resulting stimulated echoes. We have demonstrated that a simple gradient-based MLE algorithm is unsuitable in this case and developed a Bayesian alternative that exhibits near-optimal performance even at very low SNR. The proposed algorithm jointly estimates the weights and locations of the T_2 components as well as the signal phase and refocusing flip angle.

It is commonly thought that the regularisation in NNLS produces 'more realistic' distributions, primarily due to the smooth and continuous nature of the resulting curves [77]. In this work, however, we have shown that one cannot reliably estimate the distribution width, thus a regularised distribution, although continuous, is not necessarily closer to the true distribution than a discrete counterpart.

The distributions resulting from the NNLS algorithm have been most useful to date by extracting summary statistics, such as the mean intra/extra-cellular peak and the myelin water fraction [75]. These measures integrate the distribution and consequently are only weakly dependent on the estimated width. Our analysis justifies the use of these statistics, as these features can be reasonably well estimated, despite minimal knowledge of the distribution width. However, we have shown that regularisation can bias the distribution shape of the fast components, leading to a lack of anatomical information in the T_2 maps.

One drawback of the discrete multicomponent T_2 model is the need to know *a priori* the number of components. For example, the weights of the fast T_2 component were estimated close to zero in the agar gel region of the optic nerve sample (results not shown). This is due to the fact that the relaxation curve of the gel is described accurately by a single decaying component. It would be preferable to only fit a single component in this case, as the second component is simply fit to the noise and has minimal contribution to the signal. Many model order selection techniques exist, such as the Akaike or Bayesian information criteria, that choose an appropriate model order based on a statistical metric, see e.g. [87]. These criteria are applicable to MRI signal decay data although this is beyond the scope this thesis.

In the presence of non-ideal flip angles, the estimated T_2 components for a

discrete distribution with an exponential decay function exhibit more complicated behaviour than the single component counterpart. For a single component, the T_2 value was increasingly overestimated as the flip angle deviated from 180° as reported in [79]. In the case when a two component distribution was used, the second component is estimated to compensate for the stimulated echoes. This results in a large variation of the second component and relatively small variation of the first component.

The advantage of the discrete distribution model is that the associated estimation problem is inherently better conditioned than the NNLS pseudo-continuous grid model. This reduces the need for regularisation and potentially minimises the estimation bias. The minimal regularisation is apparent in our Bayesian approach since the prior distribution for the unknown parameters is relatively wide. Additionally, the proposed algorithm requires substantially less computation compared to both gradient-based optimisation and the NNLS algorithm.

The EPG signal model together with the proposed estimation algorithm makes quantitative multicomponent T_2 analysis viable in the presence of large B_1 inhomogeneity, such as those produced from a transceive surface coil.

Appendices

5.A Signal model for inverse-gamma mixture

The measured signal from a distribution consisting of an inverse-gamma mixture is derived from the general measurement equation,

$$y_n = e^{j\phi} \int_0^\infty f(\tau) e^{-t/\tau} d\tau \tag{5.32}$$

Substituting the inverse gamma mixture gives,

$$y_n = e^{j\phi} \int_0^\infty \sum_{i=1}^N w_i \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \tau^{-\alpha_i - 1} e^{-\beta_i/\tau} e^{-t/\tau} d\tau \tag{5.33}$$

$$= e^{j\phi} \sum_{i=1}^N w_i \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \int_0^\infty \tau^{-\alpha_i - 1} e^{-(t+\beta_i)/\tau} d\tau \tag{5.34}$$

Let $u = (t + \beta_i)/\tau$. Taking derivatives gives,

$$d\tau = -\frac{(t+\beta_i)}{u^2} du \tag{5.35}$$

Changing the variable of integration, and noting the new limits of $-\infty$ to 0, gives,

$$y_n = e^{j\phi} \sum_{i=1}^N w_i \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \int_{-\infty}^0 \left(\frac{t+\beta_i}{u}\right)^{-\alpha_i - 1} e^{-u} \frac{-(t+\beta_i)}{u^2} du \tag{5.36}$$

$$= e^{j\phi} \sum_{i=1}^N w_i \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} (t+\beta_i)^{-\alpha_i - 1} (t+\beta_i) \int_0^\infty u^{\alpha_i - 1} e^{-u} du \tag{5.37}$$

$$= e^{j\phi} \sum_{i=1}^N w_i \frac{\beta_i^{\alpha_i} \Gamma(\alpha_i)}{(t+\beta_i)^{\alpha_i} \Gamma(\alpha_i)} \tag{5.38}$$

$$= e^{j\phi} \sum_{i=1}^N w_i \left(\frac{\beta_i}{t+\beta_i}\right)^{\alpha_i} \tag{5.39}$$

5.B The extended phase graph algorithm

The extended phase graph algorithm is a recursive algorithm that tracks the evolution of the phase coherence pathways [72]. It describes the spin system at a given time by a distribution over different phase states. There are three types of states, $F_k^{(n)}$ (dephasing), $F_{-k}^{(n)}$ (rephasing) and $Z_k^{(n)}$ (longitudinal) for each coherence level, k, and echo number, n. After the initial excitation pulse, all the magnetisation is assumed to be coherent and on the transverse plane so $F_0^{(0)} = 1$ and

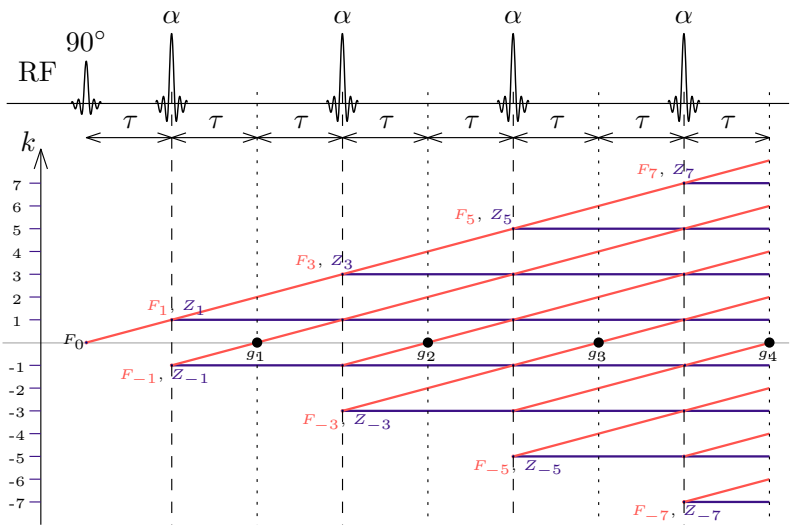


Figure 5.16: Graphical illustration of the extended phase graph (EPG) algorithm for the CPMG sequence. An RF pulse mixes coherence states into (blue) longitudinal and (red) transverse components. Echoes occur when the coherence level of the transverse magnetisation is zero.

$F_k^{(0)} = F_{-k}^{(0)} = Z_k^{(0)} = 0$ for $k = 1, \ldots, M$. The evolution around subsequent refocusing pulses is modelled by three periods of relaxation/precession, excitation, followed by relaxation/precession again. Figure 5.16 illustrates the basic principles of the algorithm. Precession causes the initial transverse magnetisation, F_0, to dephase, which is modelled with a transition to F_1. The first refocusing pulse mixes this magnetisation into four components: transverse states, F_1 and F_{-1}, depicted in red and longitudinal states, Z_1, $Z_{-1} = -Z_1$, depicted in blue. During the next period, the F_{-1} component is rephased to form the first echo, g_1. The other components precess and relax during this period until the next refocusing pulse, which then mixes the components to determine the amplitude of the second echo. The process is repeated for all RF pulses to obtain a sequence of echo amplitudes.

Formally, the relaxation and precession periods are modelled by the recursive relation,

$$F_k^{(n)\prime} = F_{k-1}^{(n)} e^{-\tau/T_2} \tag{5.40a}$$

$$Z_k^{(n)\prime} = Z_k^{(n)} e^{-\tau/T_1} \tag{5.40b}$$

where the prime notation indicates the new state. We define a vector containing all the states at the n^{th} echo as $\boldsymbol{x}_n = [F_0^{(n)}, F_1^{(n)}, F_{-1}^{(n)}, Z_1^{(n)}, \ldots, F_M^{(n)}, F_{-M}^{(n)}, Z_M^{(n)}]^T$. The relation in (5.40) can be described by a single matrix, $\boldsymbol{P}$, consisting of relaxation

and coherence level shifts,

$$\boldsymbol{P} = \begin{bmatrix} e^{-\tau/T_2} & & & \\ & \boldsymbol{R}_0 & & \\ & & \boldsymbol{R}_0 & \\ & & & \ddots \end{bmatrix} \boldsymbol{S} \tag{5.41}$$

where $\boldsymbol{R}_0 = \text{diag}\left(e^{-\tau/T_2}, e^{-\tau/T_2}, e^{-\tau/T_1}\right)$ and $\boldsymbol{S}$ is a permutation of the identity matrix that shifts the transverse states up a coherence level.

An RF pulse of flip angle α mixes states of equal coherence level described by,

$$\begin{bmatrix} F_k^{(n)\prime} \\ F_{-k}^{(n)\prime} \\ Z_k^{(n)\prime} \end{bmatrix} = \underbrace{\begin{bmatrix} \cos^2(\alpha/2) & \sin^2(\alpha/2) & \sin\alpha \\ \sin^2(\alpha/2) & \cos^2(\alpha/2) & -\sin\alpha \\ -\frac{1}{2}\sin\alpha & \frac{1}{2}\sin\alpha & \cos\alpha \end{bmatrix}}_{\boldsymbol{T}_0} \begin{bmatrix} F_k^{(n)} \\ F_{-k}^{(n)} \\ Z_k^{(n)} \end{bmatrix} \tag{5.42}$$

The mixing in (5.42) can also be applied to all states simultaneously using a block diagonal matrix, $\boldsymbol{T}$,

$$\boldsymbol{T} = \begin{bmatrix} 1 & & & \\ & \boldsymbol{T}_0 & & \\ & & \boldsymbol{T}_0 & \\ & & & \ddots \end{bmatrix} \tag{5.43}$$

where $\boldsymbol{T}_0$ is the matrix defined in (5.42).

Combining these operators the evolution of the entire state vector between echoes is described as

$$\boldsymbol{x}_n = \underbrace{\boldsymbol{P}\boldsymbol{T}\boldsymbol{P}}_{\boldsymbol{E}} \boldsymbol{x}_{n-1}. \tag{5.44}$$

The echo amplitude is given by $F_0^{(n)}$ at each time, n. That is,

$$g_n = \boldsymbol{c}^T \boldsymbol{x}_n \tag{5.45}$$

where $\boldsymbol{c} = [1, 0, \ldots, 0]^T$.

5.C Partial derivatives

The derivative of the echo amplitude with respect to the unknown parameters is useful for a number of applications including CRLB analysis, gradient descent optimisation and approximations of the likelihood. As such, an efficient means to compute the partial derivatives is included here.

The recursion in (5.44) is linear in the state vector thus the partial derivatives can also be calculated in a recursive manner. First notice that both $\boldsymbol{x}_n$ and $\boldsymbol{E}$ depend on the parameter of interest, θ, so the product rule must be applied,

$$\frac{\partial \boldsymbol{x}_n}{\partial \theta} = \frac{\partial \boldsymbol{E}}{\partial \theta}\boldsymbol{x}_{n-1} + \boldsymbol{E}\frac{\partial \boldsymbol{x}_{n-1}}{\partial \theta}. \tag{5.46}$$

Next we use the decomposition $\boldsymbol{P} = \boldsymbol{P}(r_1, r_2)$ and $\boldsymbol{T} = \boldsymbol{T}(\alpha)$. The partial derivative of the matrix elements in $\boldsymbol{E} = \boldsymbol{P}\boldsymbol{T}\boldsymbol{P}$ are

$$\frac{\partial \boldsymbol{E}}{\partial \alpha} = \boldsymbol{P}\frac{\partial \boldsymbol{T}}{\partial \alpha}\boldsymbol{P}, \tag{5.47}$$

$$\frac{\partial \boldsymbol{E}}{\partial r_2} = \frac{\partial \boldsymbol{P}}{\partial r_2}\boldsymbol{T}\boldsymbol{P} + \boldsymbol{P}\boldsymbol{T}\frac{\partial \boldsymbol{P}}{\partial r_2}. \tag{5.48}$$

For convenience we define additional intermediate states $\boldsymbol{x}_n^{(r)}$, $\boldsymbol{x}_n^{(\alpha)}$, $\boldsymbol{z}_i^{(r)}$ and $\boldsymbol{z}_i^{(\alpha)}$ to calculate the different terms in (5.46)–(5.48). We also let $\boldsymbol{P}' = \frac{\partial \boldsymbol{P}}{\partial r_2}$ and $\boldsymbol{T}' = \frac{\partial \boldsymbol{T}}{\partial \alpha}$. The recursive algorithm is shown in Algorithm 5.2. Note that the operators $\boldsymbol{P}$, $\boldsymbol{P}'$, $\boldsymbol{T}$ and $\boldsymbol{T}'$ can be executed very efficiently using the inherent structure described in (5.40) and (5.42). The derivatives of the EPG signal are simply the first element of the derivatives of the state vector, $\partial \boldsymbol{x}_n / \partial r_2$ and $\partial \boldsymbol{x}_n / \partial \alpha$.

Algorithm 5.2: Recursive algorithm for EPG derivatives

Input: Relaxation parameters, $r_1 = 1/T_1$, $r_2 = 1/T_2$; Flip angle, α
Output: Partial derivatives, $\frac{\partial \boldsymbol{x}_n}{\partial r_2}$ and $\frac{\partial \boldsymbol{x}_n}{\partial \alpha}$ for $n = 1, \ldots, M$

Calculate matrices, $\boldsymbol{P}$ and $\boldsymbol{T}$ using (5.41) and (5.43)
Calculate derivatives, $\boldsymbol{P}' = \frac{\partial \boldsymbol{P}}{\partial r_2}$ and $\boldsymbol{T}' = \frac{\partial \boldsymbol{T}}{\partial \alpha}$
for $n \leftarrow 1 \ldots M$ **do**
    // Calculate first term of product rule
    $\boldsymbol{z}_n^{(r)} \leftarrow \boldsymbol{P}\boldsymbol{T}\boldsymbol{P}'\boldsymbol{x}_{n-1}$
    $\boldsymbol{z}_n^{(\alpha)} \leftarrow \boldsymbol{P}\boldsymbol{T}'\boldsymbol{P}\boldsymbol{x}_{n-1}$
    // Calculate second term of product rule
    $\boldsymbol{x}_n^{(r)} \leftarrow \boldsymbol{P}\boldsymbol{T}\boldsymbol{P}\frac{\partial \boldsymbol{x}_{n-1}}{\partial r_2}$
    $\boldsymbol{x}_n^{(\alpha)} \leftarrow \boldsymbol{P}\boldsymbol{T}\boldsymbol{P}\frac{\partial \boldsymbol{x}_{n-1}}{\partial \alpha}$
    // Combine terms for the derivatives
    $\frac{\partial \boldsymbol{x}_n}{\partial r_2} \leftarrow 2\boldsymbol{z}_n^{(r)} + \boldsymbol{x}_n^{(r)}$
    $\frac{\partial \boldsymbol{x}_n}{\partial \alpha} \leftarrow \boldsymbol{z}_n^{(\alpha)} + \boldsymbol{x}_n^{(\alpha)}$
    // Calculate state evolution
    $\boldsymbol{x}_n \leftarrow \boldsymbol{P}\boldsymbol{T}\boldsymbol{P}\boldsymbol{x}_{n-1}$

Part III

Nonlinear Spatial Encoding

CHAPTER 6

Region-specific trajectory design using nonlinear encoding fields

Contents

6.1 Introduction

SPATIAL encoding in MRI is traditionally achieved using magnetic fields with linear gradients, applied to each of the three spatial dimensions. The linear gradient fields in these schemes are typically setup to acquire samples of the object

in the Fourier domain, or *k*-space, as described in Chapter 3. Image reconstruction for this approach is straightforward and based on the inverse discrete Fourier transform (DFT). This supports one's intuition regarding the relationship between sampling parameters and image properties such as SNR and resolution. Importantly, the trade-off between SNR and resolution can be characterised precisely.

Recently, magnetic fields that vary nonlinearly across the field-of-view have been proposed to improve on the linear gradient approaches. The concept of PatLoc (Parallel Imaging Technique using Localized Gradients) uses higher order non-bijective encoding fields with the potential to overcome safety limits associated with peripheral nerve stimulation [9]. Similar to conventional SENSE imaging, the encoding inherent in the receiver coils is crucial to resolving the ambiguity associated with non-bijectivity. For a Cartesian PatLoc acquisition, the use of ambiguous encoding fields together with multiple RF coils provides an attractive alternative for acquisition acceleration compared to traditional undersampling in parallel imaging [10]. In both cases reduced imaging time is achieved since the non-bijective encoding leads to a reduced field-of-view, which requires fewer samples for the same resolution. An important feature of a Cartesian PatLoc trajectory is the spatially varying resolution. Since the quadrupolar fields have a large gradient at the periphery of the field-of-view, this region has improved resolution. Conversely, the quadrupolar fields provide no spatial encoding at the centre of the field-of-view (where the gradient is zero) and the resolution is severely degraded, resulting in pronounced blurring.

O-Space imaging [11] and Null Space Imaging [91] use nonlinear fields with the aim to provide spatial encoding that complements the spatial encoding provided by multiple RF coil sensitivities. O-Space imaging uses a quadratic encoding field in conjunction with traditional linear fields. The linear fields translate the quadratic encoding function after each excitation. Consequently, projections are acquired along rings evolving outward from each centre point [11]. It is purported that the improved encoding along the radial direction complements the inherent encoding provided by the receive coils along the azimuthal direction, allowing for much faster imaging times. Null Space Imaging searches for the optimal encoding shapes using a singular value decomposition of the coil sensitivity matrix. In general the optimal encoding shapes cannot be realised with the available hardware, so a least square fit is performed to determine the closest achievable encoding shapes to use for projection imaging [91]. This scheme promises to accelerate image acquisition even further.

4D-RIO is a new PatLoc encoding strategy that uses linear and nonlinear fields simultaneously [12]. Two out-of-phase radial trajectories are implemented on the

linear and quadrupolar pairs such that the linear gradient is maximum when the nonlinear gradient is minimum and vise versa. The reconstructed images exhibit relatively uniform resolution compared to a Cartesian PatLoc trajectory without linear fields.

This chapter develops a novel acquisition strategy that improves the resolution in a region of interest. The technique uses nonlinear encoding fields; however, in contrast to existing techniques, this is the first method to design the acquisition around the desired resolution. The trajectory design is posed as an optimisation problem, which can be solved numerically, to produce a customised trajectory based on the specified region of interest.

The chapter is organised as follows. Section 6.2 presents the theory of spatial encoding using multiple encoding fields with arbitrary shapes. The theory can be considered a generalisation of that presented in Section 3.4. Section 6.3 examines the existing nonlinear encoding strategies of Cartesian PatLoc, O-Space, and 4D-RIO. This serves to highlight some basic properties of nonlinear spatial encoding. Section 6.4 develops a novel acquisition strategy that uses nonlinear encoding fields to enhance the resolution in a user-specified region and Section 6.5 verifies the technique through simulation and experiments.

6.1.1 Notation

Table 6.1 lists the important symbols used in this chapter. Although the list is not exhaustive, it includes most quantities of interest. Notice that the individual elements of some vector quantities can have different units depending if they describe a linear or nonlinear encoding field.

6.2 Theory

We assume the availability of N_c receive coils. Neglecting relaxation effects, the magnetic resonance signal for the l^{th} RF coil, s_l, was derived in Section 3.2,

$$s_l(t) = \int m(\boldsymbol{x}) c_l(\boldsymbol{x}) e^{j\phi(\boldsymbol{x},t)} d\boldsymbol{x} \tag{6.1}$$

where $m(\boldsymbol{x})$ is the magnetisation at position $\boldsymbol{x}$, c_l is the sensitivity of coil l and ϕ is the spatially dependent encoding phase. The encoding phase is given by

$$\phi(\boldsymbol{x},t) = -\gamma \int_0^t b(\boldsymbol{x},\tau) d\tau, \tag{6.2}$$

Table 6.1: Common notation used in Chapter 6

Symbol	Quantity	Units	
		Linear	Quadratic
m	Magnetisation (image)		
c_l	Sensitivity of the l^{th} coil		
ϕ	Encoding phase	rad	rad
$\boldsymbol{g}$	Gradients	T/m	T/m^2
$\boldsymbol{\kappa}$	Gradient moments	rad/m	rad/m^2
$\boldsymbol{\psi}$	Field shapes	m	m^2
$\boldsymbol{k}$	Local k-space	rad/m	rad/m
$\tilde{\boldsymbol{k}}$	Concatenated k-space trajectories	rad/m	rad/m
$\tilde{\boldsymbol{k}}_{\text{tar}}$	Concatenated target k-space trajectories	rad/m	rad/m
∇	Vector differential operator		
$\boldsymbol{J}$	Transpose of the Jacobian of $\boldsymbol{\psi}$		m
$\boldsymbol{A}$	Local k-space operator		
$\boldsymbol{B}, \boldsymbol{C}$	Optimisation constraints		
$\boldsymbol{S}$	Selection matrix		
$\boldsymbol{U}$	Up-sampling matrix		
f	Holomorphic function describing fields		

where γ is the gyromagnetic ratio, $b(\boldsymbol{x}, t)$ is the encoding magnetic field, and $t = 0$ denotes the time the excitation pulse is applied. The encoding field is generated by the contribution of different gradient channels. The case of three linear gradients described in Section 3.4 can be generalised to N_g encoding channels,

$$b(\boldsymbol{x}, t) = \sum_{i=1}^{N_g} g_i(t)\psi_i(\boldsymbol{x}) \tag{6.3}$$

where $g_i(t)$ is the gradient waveform proportional to the current applied through the coil and $\psi_i(\boldsymbol{x})$ is the shape of the encoding field at position $\boldsymbol{x}$. At a hardware level, the imaging system is controlled by applying a time-varying gradient. In general, the shape of each field given by ψ_i is fixed and determined by the coil geometry. The gradients and field shapes for all channels can be written as vectors $\boldsymbol{g}(t) = [g_1(t), \ldots, g_{N_g}(t)]^T$ and $\boldsymbol{\psi}(\boldsymbol{x}) = [\psi_1(\boldsymbol{x}), \ldots, \psi_{N_g}(\boldsymbol{x})]^T$. In this notation, the encoding fields is

$$b(\boldsymbol{x}, t) = \boldsymbol{g}(t) \cdot \boldsymbol{\psi}(\boldsymbol{x}), \tag{6.4}$$

where $\boldsymbol{a} \cdot \boldsymbol{b}$ denotes the inner product of two vectors. Notice for the linear case $\boldsymbol{g}(t) = [g_x(t), g_y(t), g_z(t)]^T$, $\boldsymbol{\psi}(\boldsymbol{x}) = [x, y, z]^T$ and (6.4) reduces to the linear form in

(3.20). The encoding phase is calculated by integration,

$$\phi(\boldsymbol{x}, t) = -\gamma \int_0^t \boldsymbol{g}(\tau) d\tau \cdot \boldsymbol{\psi}(\boldsymbol{x}) \tag{6.5}$$

Analogous to the formulation of k-space in Section 3.4.4, we define the gradient moments as

$$\boldsymbol{\kappa}(t) = \gamma \int_0^t \boldsymbol{g}(\tau) d\tau \tag{6.6}$$

In this case, the encoding phase can be decomposed as [10],

$$\phi(\boldsymbol{x}, t) = -\boldsymbol{\kappa}(t) \cdot \boldsymbol{\psi}(\boldsymbol{x}) \tag{6.7}$$

The vector $\boldsymbol{\kappa}(t)$ has a dimension equal to the number of encoding gradients, N_g. Thus in the general case, we refer to $\boldsymbol{\kappa}$ as a multidimensional trajectory. The gradient moments correspond to traditional k-space for the special case of linear imaging ($\boldsymbol{\psi}(\boldsymbol{x}) = \boldsymbol{x}$). For more general encoding schemes the concept of k-space can be generalised to the notion of "local k-space".

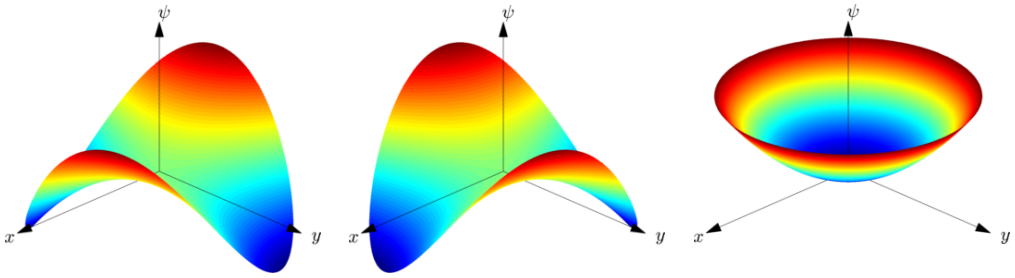
6.2.1 Local k-space

It has been shown in [12, 92] that the concept of local k-space can be defined in a way that is consistent with the traditional notions of k-space for linear imaging. The local k-space at position $\boldsymbol{x}$ is defined as

$$\boldsymbol{k}(\boldsymbol{x}, t) := -\nabla \phi(\boldsymbol{x}, t). \tag{6.8}$$

From this perspective, k-space is given by the spatial derivative of the accumulated phase, which, in the general case, can vary over the image space. Combining (6.7) and (6.8), the local k-vector field can be written as a linear transformation of the gradient moments, $\boldsymbol{k}(\boldsymbol{x}, t) = \boldsymbol{J}(\boldsymbol{x})\boldsymbol{\kappa}(t)$. In this notation, $\boldsymbol{J}$ is defined as the transpose of the Jacobian of $\boldsymbol{\psi}$. For slice-selective imaging where $\boldsymbol{x} = (x, y)$ defines a position in two dimensions, $\boldsymbol{J}(\cdot)$ is a tensor field defining a $2 \times N_g$ matrix at each point. The concept can be easily generalised to three dimensional imaging.

In the special case of linear encoding, the derivative is spatially independent and $\boldsymbol{k}(\boldsymbol{x}, t)$ reduces to the gradient moments, $\boldsymbol{\kappa}(t)$. This is explained by considering the encoding function $\boldsymbol{\psi}(\boldsymbol{x}) = [x, y]^T$ which gives $\boldsymbol{J}(\boldsymbol{x}) = \boldsymbol{I}_{2\times 2}$ (with $\boldsymbol{I}$ the identity matrix) and $\boldsymbol{k}(t) = \boldsymbol{\kappa}(t)$. This highlights the equivalence of the gradient moments and the conventional k-space trajectory for linear encoding. For nonlinear encod-



(a) (b) (c)

Figure 6.1: Example field shapes used for nonlinear spatial encoding. (a) and (b) depict two orthogonal quadrupolar fields and (c) illustrates a quadratic rotationally symmetric field.

ing the derivative of the phase is spatially dependent and the gradient moments alone are insufficient to represent the k-space trajectory.

We denote the concatenation of gradient moments as $\bar{\boldsymbol{\kappa}} = [\boldsymbol{\kappa}^T(t_1), \ldots, \boldsymbol{\kappa}^T(t_n)]^T$ and the corresponding k-space points as $\bar{\boldsymbol{k}}(\boldsymbol{x}) = [\boldsymbol{k}^T(\boldsymbol{x}, t_1), \ldots, \boldsymbol{k}^T(\boldsymbol{x}, t_n)]^T$. The entire k-space trajectory at a location $\boldsymbol{x}$ can be written simply as

$$\bar{\boldsymbol{k}}(\boldsymbol{x}) = \boldsymbol{A}(\boldsymbol{x})\bar{\boldsymbol{\kappa}} \tag{6.9}$$

where $\boldsymbol{A}(\boldsymbol{x}) = \operatorname{diag}(\boldsymbol{J}(\boldsymbol{x}), \ldots, \boldsymbol{J}(\boldsymbol{x}))$. This linear matrix form is well suited to automated trajectory design, as we will see Section 6.4.

6.3 Existing nonlinear encoding techniques

The first application of nonlinear encoding considered curvilinear multipolar fields as a replacement for traditional linear fields [9, 10]. Combinations of traditional linear gradients together with these curvilinear multipolar fields was considered in [12]. Linear fields together with quadratic rotation-symmetric fields was proposed in [11]. The analysis presented in the previous section is applicable to all these schemes, where each encoding field is used to specify the analytical form of an element in $\boldsymbol{\psi}$.

Figure 6.1 illustrates the encoding shapes used for nonlinear spatial encoding fields to date. Figures 6.1a and 6.1b depicts two quadrupolar fields in the shape of a hyperbolic paraboloid or "saddle". The saddle shapes are rotated 45° with respect to each other making them orthogonal. Figure 6.1c depicts a circular paraboloid or "bowl" shaped field.

6.3.1 Cartesian PatLoc

Perhaps the simplest application of nonlinear spatial encoding is known as Cartesian PatLoc. In this technique, the linear phase encoding and readout gradients of a simple Cartesian trajectory are replaced with the quadrupolar fields in Figures 6.1a and 6.1b. In the ideal case, these fields can be described by scaling the spherical harmonics. Specifically,

$$\psi_a(\boldsymbol{x}) = x^2 - y^2; \qquad \psi_b(\boldsymbol{x}) = 2xy \tag{6.10}$$

The vector of field shapes is

$$\boldsymbol{\psi}(\boldsymbol{x}) = \begin{bmatrix} x^2 - y^2 \\ 2xy \end{bmatrix} \tag{6.11}$$

Since the gradients of the fields are orthogonal at all points ($\nabla\psi_a \cdot \nabla\psi_b = 0$), the two fields can be used for frequency and phase encoding, analogous to encoding with linear fields. Specifically, the trajectory, $\boldsymbol{\kappa}(t)$, is given by the Cartesian sequence diagram described in Section 3.5.1. A detailed analysis of this imaging strategy will be provided in Chapter 7. In the present discussion we look at some of the basic properties of this scheme.

Simulated data was generated from a checkerboard phantom using a Cartesian trajectory with 64 phase encode and 64 readout points. Reconstruction was performed using the conjugate gradient algorithm described in Section 3.6. Two schemes were simulated: a traditional setup using two linear gradients and the PatLoc setup with orthogonal quadrupolar fields replacing the phase encode and readout gradients. To visualise the spatial dependence of the encoding, the local k-space trajectory was calculated at points on a 5×5 grid over the field of view.

Figure 6.2 displays the local k-space plots alongside the reconstruction results. Figure 6.2a displays the local k-space trajectory when linear encoding gradients are used. In this case, the local k-space reduces to traditional notion of k-space and the square sampling pattern characteristic of a Cartesian trajectory is shown everywhere. Figure 6.2b displays the corresponding reconstruction demonstrating uniform resolution. Figure 6.2c depicts the local k-space for a Cartesian trajectory using the quadrupolar fields in (6.10). The spatial dependence of the local k-space is evident with a singularity in the center where the gradient of the fields is zero and extended coverage at the periphery where the gradient is larger than the linear fields. Figure 6.2d display the corresponding image reconstruction which demonstrates spatially varying resolution. In the center, the resolution is poor with pro-

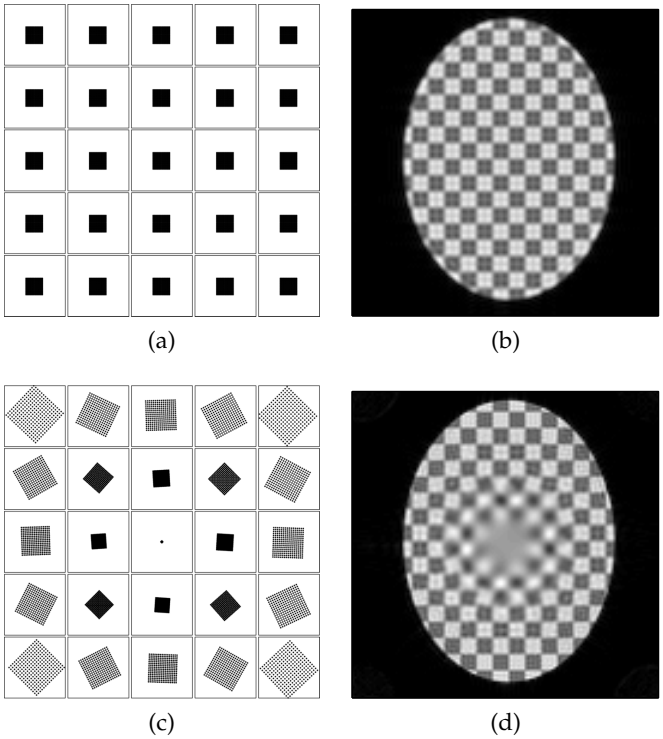


Figure 6.2: Example of a Cartesian trajectory with nonlinear encoding fields. The local *k*-space is depicted in (a) for linear fields and (c) for quadrupolar PatLoc fields. Reconstructions of a simulated checkerboard phantom are shown in (b) and (d).

nounced blurring and in the periphery the resolution is improved compared to linear encoding. Importantly, this demonstrates that the behaviour of the local k-space trajectories is indicative of the spatially varying resolution characteristic of nonlinear encoding fields. This relationship is exploited in Section 6.4 to design novel encoding strategies with improved resolution in a region of interest.

In addition to the spatially varying resolution, another important feature of nonlinear fields is that the encoding is non-bijective. That is, two points on opposite sides of the image will be encoded with identical phase. This ambiguity is resolved using multiple receive coils since each coil modulates the signal by its complex-valued sensitivity. The measurements from each coil provide a different view of the object, which can be used to successfully reconstruct the image. This is similar to undersampling with linear gradients where aliasing is resolved using multiple receive coils [14]. In Chapter 7 we will analyse this problem in more detail.

6.3.2 O-Space

O-Space is a technique that combines linear fields with a quadratic "bowl shaped" field and a constant B_0 field offset. The quadratic field as shown in Figure 6.1c is described by

$$\psi_c(\boldsymbol{x}) = x^2 + y^2 \tag{6.12}$$

The constant field can be represented with $\psi_d(\boldsymbol{x}) = 1$. Including the linear gradients, the entire vector of field shapes for O-space imaging is

$$\boldsymbol{\psi}(\boldsymbol{x}) = \begin{bmatrix} x \\ y \\ x^2 + y^2 \\ 1 \end{bmatrix} \tag{6.13}$$

In addition to this particular combination of fields, O-space is specified by a unique trajectory with gradients given by

$$g_x(t) = 2Gr_x \tag{6.14}$$
$$g_y(t) = 2Gr_y \tag{6.15}$$
$$g_c(t) = -G \tag{6.16}$$
$$g_d(t) = -G(r_x^2 + r_y^2) \tag{6.17}$$

where r_x and r_y parameterise the projection[1]. Different projections are acquired by altering the values of r_x and r_y. The entire acquisition can be specified by a set of vectors, $\boldsymbol{r}_q = (r_x^{(q)}, r_y^{(q)})$ for $q = 1, \ldots, N_e$. In this case, the encoding field in (6.4) becomes

$$b(\boldsymbol{x}, t) = G\|\boldsymbol{x} - \boldsymbol{r}_q\|^2 \tag{6.18}$$

It is evident that the role of the two linear gradients and the B_0 offset is to shift the centre of the quadratic field. With this field shape, projections of the object are obtained along rings expanding outward from the centre point, $\boldsymbol{r}_q$. Indeed, the term 'O-space' was motivated by the circular isocontours of the encoding field. By translating the centre point each projection, all regions obtain some spatial encoding and the central singularity in Cartesian PatLoc is avoided.

6.3.3 4D-RIO

Another technique, termed 4D-RIO, simultaneously employs two linear and two quadrupolar fields for spatial encoding. The fields are specified by the vector,

$$\boldsymbol{\psi}(\boldsymbol{x}) = \begin{bmatrix} x \\ y \\ x^2 - y^2 \\ 2xy \end{bmatrix} \tag{6.19}$$

The trajectory for this technique is based on two standard radial trajectories used to drive the linear and quadratic coils. The trajectories are out-of-phase such that the gradient waveforms for the nonlinear channels are maximum at times when the linear gradients are minimum and vice versa.

6.4 Region specific trajectory design

As demonstrated above, an important consequence of nonlinear encoding fields is that the image resolution is spatially varying. The combination of linear and nonlinear encoding fields was examined in [12] and it was demonstrated that different trajectories produce different regions of improved and degraded resolution. A simple forward analysis was presented based on the concept of local k-space, where the resolution can be predicted for a given trajectory. We propose the opposite; we design a multidimensional trajectory given the required resolution. In

[1] This is similar to angle parameterising a radial trajectory as discussed in Section 3.5.2.

this way, it is possible to design trajectories with improved resolution in a region of interest.

We define an optimisation problem based on the local *k*-space trajectory at different points within the region of interest. Numerical optimisation of gradient waveforms has been previously proposed for linear encoding in a range of works, e.g. [93] and references therein. In our case, the addition of nonlinear encoding means that no solution exists that implements our target trajectory at all locations within the desired region. Instead, we seek to minimise the difference between the local *k*-space trajectories and the target trajectory. To increase the clarity of the proposed technique, our target trajectory is based on a simple EPI scheme with no ramp-sampling, although more complex trajectories are possible.

To date, nonlinear spatial encoding has only been applied to projection imaging with multiple excitation pulses. In addition to trajectory design, we demonstrate the first application of nonlinear encoding fields for single-shot imaging. Single-shot imaging is particularly well suited to multidimensional trajectories, where the additional encoding gradient channels can overcome the limitations associated with fast switching.

6.4.1 Trajectory optimisation

Similar to the global *k*-space trajectory for linear encoding, the sampling trajectory of local *k*-space contains valuable information on properties of the reconstructed image. In this work, we focus on resolution, as we aim to improve the resolution in a localised region. The global image resolution for linear encoding is determined by the width of the main lobe of the point spread function, which is inversely proportional to the extent of the traditional *k*-space trajectory. Likewise, for nonlinear encoding, the resolution at a given position is determined by the extent of the local *k*-space trajectory at that position.

Based on *k*-space intuition, a trajectory that covers twice the maximum spatial frequency will produce images with a two-fold improvement in resolution. Additionally, the trajectory should cover *k*-space with sufficient uniformity to ensure the reconstruction is reasonably well conditioned. The assumption here is that any undersampling resulting from an increased distance between *k*-space points can be resolved using multiple RF coils and the information from the associated sensitivity maps.

This has close links to traditional parallel imaging where, for a fixed number of measurements, *k*-space is undersampled so the maximum spatial frequency is greater than that achievable by satisfying Nyquist limit. It is well known that this

undersampling leads to aliasing which can be resolved using information from the coil sensitivity profiles [14, 94]. In our case, a similar situation exists except that we aim to control the region where undersampling occurs.

We use the notion of local k-space to define an optimisation problem for trajectory design. Specifically, we define a target k-space trajectory for points inside the region of interest and try to find the gradient moments to achieve this. For a fixed readout time, conventional single-shot trajectories (such as EPI or spiral) using only linear gradients cover a maximum spatial frequency that is limited by gradient performance and safety considerations. As such, good candidates for the target trajectory are single-shot trajectories that cover an increased spatial frequency. The aim is that additional nonlinear encoding channels can be used to extend the coverage of local k-space (at the expense of undersampling) to improve the resolution in the desired region.

It is impractical to calculate the local k-space trajectory at every pixel in the region of interest. Instead we calculate the trajectory at a set of control points within the region. Since the encoding fields are smooth, the local k-space trajectory will vary smoothly over the imaging region. Thus it is sufficient to specify the desired trajectory at a small number of points and find the gradient waveforms that approximately achieve this while satisfying the physical constraints of the scanner hardware. It will be demonstrated in this work that the addition of nonlinear encoding fields allows the calculation of gradient waveforms that steer the areas of improved resolution to the region of interest.

To formulate an optimisation problem we define the target k-space trajectory, $\tilde{\boldsymbol{k}}_{\text{tar}}$, as the concatenation of the target trajectories at each control point. Denoting the set of control points by $\{\boldsymbol{z}_i\}_{i=1,\ldots,p}$, the concatenation of the local trajectories at each point is represented by the single matrix equation,

$$\tilde{\boldsymbol{k}} = \tilde{\boldsymbol{A}}\bar{\boldsymbol{\kappa}} \tag{6.20}$$

where

$$\tilde{\boldsymbol{k}} = \begin{bmatrix} \bar{\boldsymbol{k}}(\boldsymbol{z}_1) \\ \vdots \\ \bar{\boldsymbol{k}}(\boldsymbol{z}_p) \end{bmatrix}, \quad \tilde{\boldsymbol{A}} = \begin{bmatrix} \boldsymbol{A}(\boldsymbol{z}_1) \\ \vdots \\ \boldsymbol{A}(\boldsymbol{z}_p) \end{bmatrix}. \tag{6.21}$$

Ideally, we wish to find a set of gradient moments, $\bar{\boldsymbol{\kappa}}$, such that $\tilde{\boldsymbol{k}}_{\text{tar}} = \tilde{\boldsymbol{A}}\bar{\boldsymbol{\kappa}}$. In the general case, this is not always realisable due to the limited shapes of the encoding fields in $\boldsymbol{\psi}$ and the constraints imposed by safety considerations and hardware performance. Instead we seek to minimise the difference between the

desired trajectory and the actual trajectory, $\|\tilde{\boldsymbol{k}}_{\text{tar}} - \tilde{\boldsymbol{A}}\bar{\boldsymbol{\kappa}}\|$, measured with a suitable norm.

The physical constraints fall under two categories: equality constraints and inequality constraints. The equality constraints concern the initial and final states of the gradient system. Specifically, the gradient waveforms must be zero at the start and end of the sequence. It is also assumed that the system is in a resting state prior to imaging, i.e. $g(t) = 0$ for $t < 0$, which means the initial gradient moment is also zero. The equality constraints are summarised as

$$\boldsymbol{g}(t_1) = \boldsymbol{0}, \tag{6.22a}$$
$$\boldsymbol{g}(t_n) = \boldsymbol{0}, \tag{6.22b}$$
$$\boldsymbol{\kappa}(t_1) = \boldsymbol{0}. \tag{6.22c}$$

The inequality constraints deal with safety and hardware limitations concerning the maximum gradient amplitude and the gradient slew rate. Constraints of this form already exist in clinical scanners for linear gradients and the same constraints should be applied to additional nonlinear gradients. The gradient amplitude is limited by $\boldsymbol{g}_{\max}$ and the slew rate by $\boldsymbol{s}_{\max}$. These constraints can be written as

$$|\boldsymbol{g}(t)| \leq \boldsymbol{g}_{\max}\,, \tag{6.23a}$$
$$\left|\frac{\partial \boldsymbol{g}(t)}{\partial t}\right| \leq \boldsymbol{s}_{\max}\,. \tag{6.23b}$$

The constraints in (6.22) and (6.23) are simple linear constraints and using finite difference approximations for the time derivative, we can write them as $\boldsymbol{B}\bar{\boldsymbol{\kappa}} = \boldsymbol{0}$ and $\boldsymbol{C}\bar{\boldsymbol{\kappa}} \leq \boldsymbol{b}$, respectively. In this work, we only constrain the switching rates of individual channels. In future work we plan to adapt the framework to consider the effect of all channels simultaneously.

We can now define the optimisation as

$$\begin{aligned} &\text{minimise} && \|\tilde{\boldsymbol{k}}_{\text{tar}} - \tilde{\boldsymbol{A}}\bar{\boldsymbol{\kappa}}\| \\ &\text{subject to} && \boldsymbol{B}\bar{\boldsymbol{\kappa}} = \boldsymbol{0} \\ &&& \boldsymbol{C}\bar{\boldsymbol{\kappa}} \leq \boldsymbol{b}. \end{aligned} \tag{6.24}$$

Once the achievable gradient moments are found, the optimal gradient waveforms are computed using finite differences to generate a waveform suitable for implementation on the scanner.

Although the optimisation problem in (6.24) forms the basis of our trajectory design it is not directly applicable for our hardware. The scanner used in this study

(see "experimental setup" below) requires a gradient waveform sampled every 10 µs, known as the gradient raster time. However the ADC can acquire samples faster. In our phantom experiments, we set the ADC sampling time to 2.5 µs, so four data samples are acquired for each gradient raster. Consequently, four local k-space points are also defined every gradient sample. To apply the optimisation, the gradient waveform (or gradient moment) must be related to the local k-space, which is achieved using an up-sampling matrix, $\boldsymbol{U}$. We define $\boldsymbol{U}$ based on linear interpolation. For example, an up-sampling matrix by a factor of two is,

$$\boldsymbol{U} = \begin{bmatrix} 0.5 & & \\ 1 & & \\ 0.5 & 0.5 & \\ & 1 & \\ & 0.5 & \ddots \end{bmatrix} \otimes \boldsymbol{I}_q \tag{6.25}$$

where $\otimes$ is the Kronecker product and $\boldsymbol{I}_q$ is the $q \times q$ identity matrix representing the q encoding channels that require up-sampling.

Another consideration for EPI-based trajectories is the end-of-line transition time. This extra time after each acquisition line allows the gradient system to change polarity. Our goal is to design a gradient waveform, $\boldsymbol{g}(t)$, for the entire trajectory, ensuring the constraints are satisfied at each time point. However, our objective is to match only the appropriate portion of the trajectory to the target k-space. We solve these differing requirements by seeking a complete trajectory using a cost function dependent only on the 'central' portion. This is implemented using a rectangular selection matrix, $\boldsymbol{S}$, which reduces the entire trajectory to the central part relevant for optimisation. Specifically,

$$\boldsymbol{S} = \begin{bmatrix} 0 & \boldsymbol{I}_{\text{row}} & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & \boldsymbol{I}_{\text{row}} & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & \boldsymbol{I}_{\text{row}} & \\ & & & & & & \ddots \end{bmatrix} \otimes \boldsymbol{I}_q \tag{6.26}$$

where $\boldsymbol{I}_{\text{row}}$ is an identity matrix selecting a single row of the trajectory, the first column of zeros discards the prephase and subsequent zero columns discard the end-of-line transitions. Figure 6.3 illustrates the various waveform vectors and the matrices relating them.

Inclusion of the up-sampling and selection matrices defined in (6.25) and (6.26),

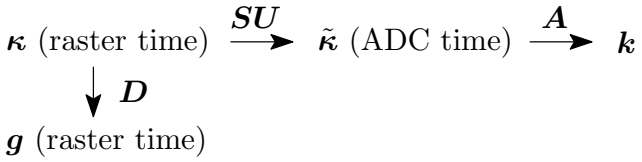


Figure 6.3: The waveform vectors used for trajectory design and the transformation matrices relating each vector. The gradient waveforms, $\boldsymbol{g}$, are computed from finite differences of the gradient moments, $\boldsymbol{\kappa}$, via the matrix $\boldsymbol{D}$. The matrix $\boldsymbol{U}$ up-samples the gradient moments to the ADC rate and the matrix $\boldsymbol{S}$ extracts the central portion of the trajectory. The local k-space, $\boldsymbol{k}$, is computed using $\boldsymbol{A}$, which contains the spatial derivatives of the encoding fields.

respectively, allows us to specify the complete optimisation problem.

$$\begin{aligned} &\text{minimise} && \|\tilde{\boldsymbol{k}}_{\text{tar}} - \tilde{\boldsymbol{A}}\boldsymbol{S}\boldsymbol{U}\bar{\boldsymbol{\kappa}}\| \\ &\text{subject to} && \boldsymbol{B}\bar{\boldsymbol{\kappa}} = \boldsymbol{0} \\ & && \boldsymbol{C}\bar{\boldsymbol{\kappa}} \leq \boldsymbol{b}. \end{aligned} \tag{6.27}$$

where $\tilde{\boldsymbol{A}}$ is defined in (6.21), and $\boldsymbol{B}$ and $\boldsymbol{C}$ implement the constraints in (6.22) and (6.23), respectively.

The optimisation problem in (6.27) is equivalent to optimising the square of the norm. In this work we use the $\mathcal{L}_2$ norm since the minimisation can then be written as a quadratic program which can be solved using computationally efficient algorithms [95].

Once the optimal gradient moments are found, the corresponding gradient waveform is computed using finite differences and a constant scale factor specified by the matrix $\boldsymbol{D}$. A major advantage of solving for the optimal gradient moments instead of the gradient waveforms directly is that the corresponding transformation matrix is extremely sparse, which greatly reduces the computation time. The optimisation problem above was solved using the MATLAB algorithm `quadprog`, which implements an interior point method capable of handling sparse matrices.

6.4.2 Encoding fields and target regions

We consider three different quadratic fields for additional spatial encoding. Two quadrupolar encoding fields, used in PatLoc imaging to date, are specified using a holomorphic function to define two orthogonal hyperbolic paraboloids (or "saddle" shapes) that closely approximate the fields created from the available hardware [96]. Specifically,

$$\psi_a(\boldsymbol{x}) = \operatorname{Re} f(\boldsymbol{x}), \quad \psi_b(\boldsymbol{x}) = \operatorname{Im} f(\boldsymbol{x}) \tag{6.28}$$

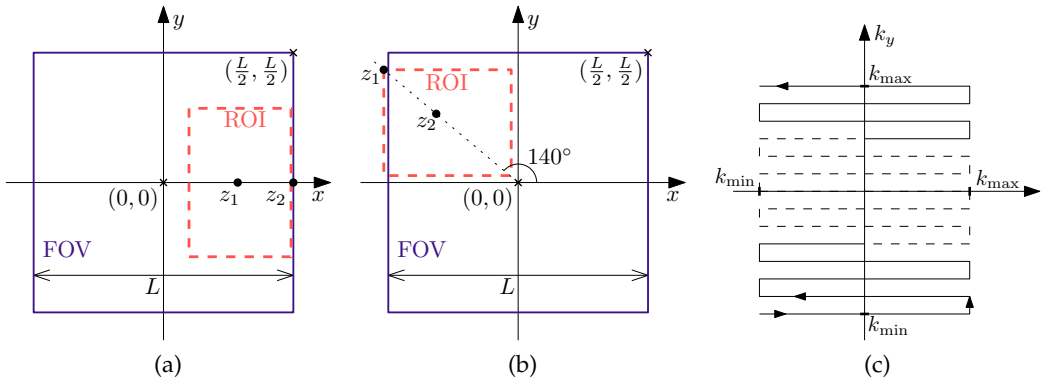


Figure 6.4: Field-of-view (FOV), region of interest (ROI) and control points (z_1, z_2) for (a) the 3-component trajectory and (b) the 4-component trajectory. The local k-space trajectory at the control points is optimised with respect to the target trajectory in (c).

where $f(\boldsymbol{x}) = (x + iy)^2 e^{-i\pi/4}$. The $e^{-i\pi/4}$ factor represents a 22.5° rotation from the main axis, consistent with the location of the physical coils. The third quadratic field is a circular paraboloid (or "bowl" shaped), currently used in O-space imaging [11], represented by the function

$$\psi_c(\boldsymbol{x}) = x^2 + y^2. \tag{6.29}$$

Although our site does not have hardware to generate this nonlinear encoding field, other sites have reported on its performance, and therefore this field is studied to demonstrate the generality and broader applicability of the present approach.

With these analytical forms the multidimensional function, $\boldsymbol{\psi}$, can by specified for any combination of linear and nonlinear fields. Importantly, all physically-realisable fields satisfy Laplace's equation thus the spatial derivatives are well-defined and the matrix $\boldsymbol{J}$ can be specified. We can then select the control points and target trajectory and solve the optimisation problem to generate the appropriate gradient waveforms. In this work we consider two regimes illustrated in Fig. 6.4:

3-component: Linear & Circular Paraboloid, Right-Sided Enhancement In this setup we use a combination of two linear gradients and a circular paraboloid field to enhance the resolution on the right side of the image. As illustrated in Fig. 6.4a, the region of interest is defined using the control points,

$$z_1 = (L/4, 0), \quad z_2 = (L/2, 0) \tag{6.30}$$

where L is the field-of-view in both the x and y directions. The encoding function and its spatial derivative are

$$\boldsymbol{\psi}(\boldsymbol{x}) = \begin{bmatrix} x & y & x^2+y^2 \end{bmatrix}^T, \quad \boldsymbol{J}(\boldsymbol{x}) = \begin{bmatrix} 1 & 0 & 2x \\ 0 & 1 & 2y \end{bmatrix}. \tag{6.31}$$

4-component: Linear & Hyperbolic Paraboloid, Top-Left Enhancement In this regime we use the two linear gradients and the two PatLoc quadrupolar fields to target the top-left region of the image. The control points for this regime are shown in Fig. 6.4b. Specifically,

$$z_1 = ((L/2)\cos 140^\circ, (L/2)\sin 140^\circ), \quad z_2 = ((L/4)\cos 140^\circ, (L/4)\sin 140^\circ). \tag{6.32}$$

A 140° line of focus is selected to demonstrate the flexibility of the target region, beyond the regions naturally selected by the field geometry. The encoding function and spatial derivative in this case are

$$\boldsymbol{\psi}(\boldsymbol{x}) = \begin{bmatrix} x & y & \operatorname{Re} f(\boldsymbol{x}) & \operatorname{Im} f(\boldsymbol{x}) \end{bmatrix}^T, \tag{6.33}$$

$$\boldsymbol{J}(\boldsymbol{x}) = \begin{bmatrix} 1 & 0 & \operatorname{Re} \frac{\partial f}{\partial x}(\boldsymbol{x}) & \operatorname{Im} \frac{\partial f}{\partial x}(\boldsymbol{x}) \\ 0 & 1 & \operatorname{Re} \frac{\partial f}{\partial y}(\boldsymbol{y}) & \operatorname{Im} \frac{\partial f}{\partial y}(\boldsymbol{x}) \end{bmatrix} \tag{6.34}$$

where $\frac{\partial f}{\partial x}(\boldsymbol{x}) = 2xe^{-i\pi/4}$ and $\frac{\partial f}{\partial y}(\boldsymbol{x}) = 2ye^{-i\pi/4}$.

In both regimes the target trajectory is a Cartesian EPI illustrated in Fig. 6.4c. For a fixed acquisition time, the maximum k-space position of a standard linear-gradient EPI, $k_{\max}^{\text{lin}}$, is limited by the gradient and slew constraints. To improve the resolution, we define the target trajectory to cover twice the maximum spatial frequency of the linear-only EPI. That is, $k_{\max} = 2k_{\max}^{\text{lin}}$. This was chosen to obtain approximately a two-fold improvement in resolution, based on k-space intuition. We consider slice-selective imaging, where the linear z-gradient is used to excite a slice and the remaining fields are used to encode the 2D image. Thus the trajectory names '3-component' and '4-component' refer to the number of encoding channels used to encode two spatial dimensions.

Each regime above specified the region of interest using two control points. A single control point creates a simpler optimisation problem although this results in resolution with a larger variation across the region of interest. On the other hand, it is possible to specify more than two control points; however, it was found that this did not greatly change the resulting trajectories, due to the fact that the encoding

fields vary relatively slowly across space.

In both cases the target trajectory consisted of 64 lines with 64 samples each line. For safety considerations the maximum gradient, $g_{\max}$, for the linear channels was set to a conservative 22 mT/m. This is the minimum strength required to implement a standard EPI in the given time. The maximum gradient for the nonlinear channels was set to 115 mT/m^2, which was conservatively chosen to be 15% below the limits of our current hardware. The maximum slew rate, $s_{\max}$, was set to 170 T/m/s for the linear channels and 772 T/m^2/s for the nonlinear channels, which leads to equivalent dB/dt at the edge of the field-of-view. To meet the slew rate constraints, extra time must be allocated to allow the gradients to safely change polarity after each line. We refer to this time as the 'end-of-line transition'. Additionally, a prephase time is needed to initially move the trajectory from the centre of k-space to the corner of the acquisition grid. In these experiments a prephase time of 640 µs and end-of-line transition time of 360 µs was allocated. These relatively long transition times were chosen to prevent ramp-sampling, which simplifies the presentation of our method. The total duration of the single-shot trajectories was 44.16 ms.

6.4.3 Simulations

All simulations were conducted using MATLAB (The Mathworks, Natick, MA). Given the trajectory, $\bar{\boldsymbol{\kappa}}_{\text{opt}}$, resulting from the optimisation, the corresponding local k-space at the point $\boldsymbol{x}_i$ was computed using

$$\bar{\boldsymbol{k}}(\boldsymbol{x}_i) = \boldsymbol{A}(\boldsymbol{x}_i)\boldsymbol{S}\boldsymbol{U}\bar{\boldsymbol{\kappa}}_{\text{opt}}. \tag{6.35}$$

To visualise the local k-space over the entire field-of-view, a 5×5 grid of locations was defined and the local k-space of each location is plotted on an equivalent grid.

The local resolution of each trajectory was determined by calculating the width of the point spread function (PSF). The PSF was approximated by back-projecting measurements simulated from a single source pixel (similar to calculating a row of $\boldsymbol{E}'\boldsymbol{E}$ where $\boldsymbol{E}$ is the encoding matrix). The accuracy of this approach was verified at eight points around the field of view by calculating the full PSF using conjugate gradient reconstruction. Good quantitative agreement was found between our approximation and the full PSF. The width of the 2D PSF was quantified in the x and y directions by integrating the absolute value of the PSF along a horizontal or vertical line centred at the source pixel. Since the height of the PSF is independent of pixel location (neglecting coil sensitivities), the sum provides a measure of width. The maximum width over both directions was chosen as a measure of

resolution. This was repeated for each pixel to build a quantitative map of the resolution.

Simulated measurements were created from the designed trajectories using a discrete approximation of the signal model in (6.1). The coil sensitivities of eight RF coils were simulated in the magnetostatic approximation of the Biot-Savart law for single coil loops arranged concentrically around the field-of-view. The encoding phase, ϕ, was calculated using the optimal trajectory and the encoding function $\boldsymbol{\psi}$ defined analytically as the appropriate combination of linear fields (x, y) and quadratic fields (ψ_a, ψ_b, ψ_c given by (6.28) and (6.29)).

The magnetisation, m, was defined by a numerical phantom, which consisted of a checkerboard pattern multiplied by a circular mask and filtered using a Hamming window to remove the very high spatial frequencies. This phantom was generated at a significantly higher resolution (512×512) compared to the final reconstructed images (128×128), to increase the accuracy of the discrete approximation used to simulate the integral in (6.1). Finally, independent complex noise was added with zero mean and a standard deviation equal to 5% of the mean signal intensity, identical to previous studies [12]. Images were reconstructed using the iterative conjugate-gradient method described in Section 3.6.

6.4.4 Experiments

In vivo and phantom experiments were conducted on a 3 T clinical imaging system (MAGNETOM, Trio Tim, Siemens Healthcare, Erlangen, Germany) fitted with a custom-built gradient insert-coil. The insert-coil generates two fields which approximate a hyperbolic paraboloid where one field is rotated 45° with respect to the other to create two orthogonal encoding fields. These encoding fields have been previously described in [97].

The coils for each quadrupolar field can be driven with currents up to 80 A using additional high-performance gradient amplifiers. The scanner architecture was modified to control both the linear gradient coils and the PatLoc gradient coils, which allows us to simultaneously and independently drive the five encoding gradients. A Siemens head coil, used in the PET-MR scanner, was fitted inside the gradient insert and contained a single RF coil used for excitation and an eight channel coil array for receiving.

This hardware cannot generate a circular paraboloid shaped field so only the 4-component regime is tested experimentally. Thus after the slice selection gradient, the remaining two linear and two quadrupolar fields are driven simultaneously to encode the two dimensional slice. To achieve this, an existing MRI sequence for

single-shot three-dimensional encoding was modified for our purposes. Importantly, the sequence allows us to specify arbitrary gradient waveforms for each of the five encoding channels.

The phantom used in experiments is a 190 mm diameter plastic cylinder containing approximately 130 thin Plexiglas tubes (with diameter ∼4 mm) parallel to the longitudinal axis of the cylinder. The cylinder and tubes were filled with water doped with nickel sulfate and sodium chloride to give suitable contrast.

All acquisitions had a field-of-view of 220 mm and images were reconstructed to a 128×128 grid using the conjugate gradient algorithm described in Section 3.6. The trajectories had an effective TE $\approx$ 21.8 ms, which is half the readout duration. A single ADC event was used throughout the single-shot readout, all samples of which were included in the image reconstruction. The data were sampled with an ADC dwell time of 2.5 µs and 5 µs for the phantom and in vivo experiments, respectively. The ADC sampling rate must be high enough to capture the largest possible phase change between consecutive readout points. If the sampling rate is too low, the local frequency may be outside the acquisition bandwidth and the signal from these regions will be attenuated by the scanner's anti-aliasing filter. The slice thickness was 2 mm for phantom experiments and 3 mm for in vivo experiments. Thin slices were chosen to minimise signal dropouts due to through-plane dephasing [98].

6.4.5 Calibration

As the single-shot trajectories become shorter and the hardware limits of the gradient coils are approached, the actual trajectory can deviate significantly from the desired trajectory [99]. These deviations originate from many sources including eddy currents, thermal drifts, hardware delays and concomitant fields. Without adequate correction, these inaccuracies can cause blurring, geometric distortions or ghosting in reconstructed images [100]. The problems have been partially addressed by characterising the eddy-currents [101], measuring the k-space trajectory [99] and by correcting for distortions due to concomitant fields [102]. In recent years, the magnetic field has been monitored by measuring the signal from field probes [103, 104, 105, 106, 107, 108] and fitting the signal phase to a model of the field evolution. This paradigm was used to successfully reconstruct images from e.g. spiral trajectories [104, 106] or echo-planar imaging [106] in the presence of undesirable field perturbations. In [108], the field evolution was modelled by real solid harmonics up to third order to account for higher order field perturbations in diffusion imaging. In this work where a custom-built nonlinear gradient coil

insert is used, the problem of gradient field deviations is even more pronounced. In order to calibrate the simultaneously applied gradient waveforms from the linear and the nonlinear gradient coils, a field camera with 16 field probes is used [105, 108, 109].

Reconstruction of the experimental data requires accurate estimates of the RF sensitivity maps $c_l(\cdot)$, trajectory $\boldsymbol{\kappa}(\cdot)$ and gradient shapes $\boldsymbol{\psi}(\cdot)$. In this work we measure the effective trajectories using a 'field camera' constructed with 16 ^{1}H field probes [105, 106, 108]. The field probes were placed approximately equidistant around the surface of a sphere with a radius of 9 cm. The probes were excited using a separate transmit chain, similar to [110], consisting of a signal generator (N5181A, Agilent, Santa Clara, CA, USA), a power amplifier (75A400, Amplifier Research, Souderton, PA, USA) and a power splitter (MITEQ-ESTONIA, Estonia). The separate transmit chain was controlled via the optical trigger from the scanner. The inserted RF coil and the eight channel head coil array were kept in the gradient coil and connected to a coil test-bench for detuning and grounding during the trajectory calibration measurements. The field probes were excited with a rectangular RF pulse of 5 µs duration and the resulting signal was sampled with an ADC dwell time of 2.5 µs. This setup is based on work in [105, 108, 110] and has been previously used in [109] to characterise the PatLoc gradients.

The encoding phase evolution was modelled with 16 real valued spherical harmonics up to 3$^{\text{rd}}$ order. The 16 basis functions define the vector of field shapes $\boldsymbol{\psi}$. The 16 corresponding coefficients were fit to the probe measurements using a standard least squares algorithm [106] to define the trajectory $\boldsymbol{\kappa}$ (referred to as the 'measured trajectory'). In this way we obtain a very accurate estimate of the actual encoding phase at each time point. With our current hardware setup we measured the trajectory in a separate scan.

For comparison purposes, we also reconstructed images using the 'nominal trajectory' and encoding field shapes obtained using data from a custom 8 echo GRE sequence (128×128, 5 mm slice thickness, 220 mm field-of-view, 4 ms echo spacing) used previously in [111]. The field mapping was performed on the imaged object immediately prior to acquisition. The nominal trajectory was obtained by adding a 0.6 µs delay to the linear channels and a 5.3 µs delay to the nonlinear channels of the optimised trajectory. The time delays were estimated empirically in order to reconstruct the best possible image. For both the measured and nominal reconstructions, the RF sensitivities were estimated using an adaptive method based on a stochastic formulation of the matched filter, as in [112], with data from the first echo of the field mapping sequence.

6.4.6 Safety considerations

In vivo experiments were performed on a healthy volunteer after suitable IRB approval was obtained. Since this was the first application of single-shot imaging using linear and nonlinear fields simultaneously, it was important to address potential safety issues. The safety of the gradient-insert coil for traditional PatLoc imaging has been described in [113]. The implementation of single-shot imaging brings an increased risk of peripheral nerve stimulation (PNS) due to the rapidly switching linear and nonlinear gradients. To mitigate the risk the gradient waveforms were initially implemented at 40% of the desired strength, which reduces the slew rates by the same factor [114]. This ensured the maximum magnitude of dB/dt was below 20 T/s throughout the region covering the subject's body. The amplitudes were then increased in 10% increments, each time confirming with the volunteer that no PNS was experienced.

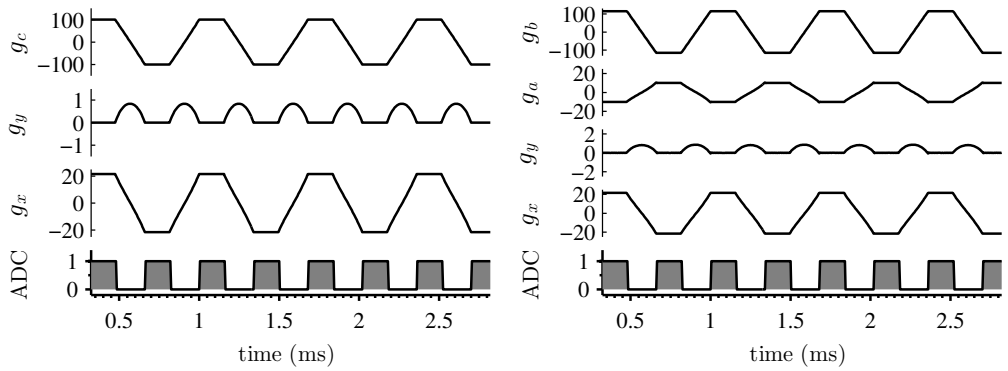
A further safety consideration was the elevated acoustic noise generated from the custom gradient-insert. Previous experiments in [113, 114] reported a sound pressure level (SPL) above regulatory guidelines for a gradient switching frequency between 480–650 Hz. A Fourier analysis of the optimised trajectories in this chapter indicated a dominant frequency of ∼750 Hz, well above the problematic frequencies. After measuring the SPL of the sequence with a calibrated microphone, the risk of auditory damage was deemed negligible with the use of suitable earplugs.

6.5 Results

6.5.1 Trajectory optimisation and simulations

The gradient waveforms for the 3-component and 4-component encoding schemes were calculated from the optimisation procedure described above. Figure 6.5 displays the 3-component and 4-component gradient waveforms necessary to enhance the desired regions of the image. Since the target trajectory is an undersampled EPI trajectory (Fig. 6.4c), the waveforms are similar to a conventional EPI sequence. Most noticeable are the constant gradients during readout lines and blips of g_y between lines.

The local k-space trajectories for the linear, 3-component and 4-component encoding regimes are plotted in Fig. 6.6, together with the simulation results. The local k-space trajectories in Figs. 6.6d and 6.6g appear 'skewed' and are not exactly equal to the target trajectory in the desired regions due to the gradient constraints and nonlinear encoding fields. However, the optimised local k-space trajectories retain the desired features: the samples are regularly spaced and have increased



(a) (b)

Figure 6.5: The gradient waveforms using (a) 3-component encoding with two linear and one quadratic and (b) 4-component encoding using two linear and two quadrupolar fields. The linear waveforms (g_x, g_y) are in units of mT/m whereas the nonlinear waveforms (g_a, g_b, g_c) are in units of mT/m^2 reflecting the quadratic nature of the corresponding fields over the spatial dimension. The plots display the first few milliseconds after the pre-phasing.

coverage. The fine details of the checkerboard phantom are completely blurred out for conventional linear encoding as shown in Fig. 6.6c. The blurring occurs because the hardware and safety constraints imposed for single-shot imaging limit the k-space extent achievable in a fixed acquisition time. The addition of extra nonlinear encoding channels allows us to increase the boundaries of the local k-space trajectory in the regions we are interested in, effectively increasing the resolution in those parts of the image. This is demonstrated in the simulation results of Figs. 6.6f and 6.6i where the checkerboard details can be shown in the desired regions, i.e., the right side for the 3-component encoding and top-left corner for the 4-component encoding scheme.

Some intuition about the gradient waveforms resulting from the optimisation can be gained by considering the local k-space equation. For example, the k-space for the 3-component trajectory is given by,

$$\boldsymbol{k}(\boldsymbol{x},t) = \begin{bmatrix} k_x(\boldsymbol{x},t) \\ k_y(\boldsymbol{x},t) \end{bmatrix} = \begin{bmatrix} \kappa_x(t) + \kappa_c(t)x \\ \kappa_y(t) + \kappa_c(t)y \end{bmatrix}. \tag{6.36}$$

During a readout line we must move the x-coordinate of the local k-space trajectory from $-k_{\max}$ to $k_{\max}$ in a relatively short time. Once the linear channel, $\kappa_x(t)$, is at maximum strength the additional term of $\kappa_c(t)x$ is crucial. The right side of the image is represented by $x > 0$, thus to enhance the k-space coverage in this re-

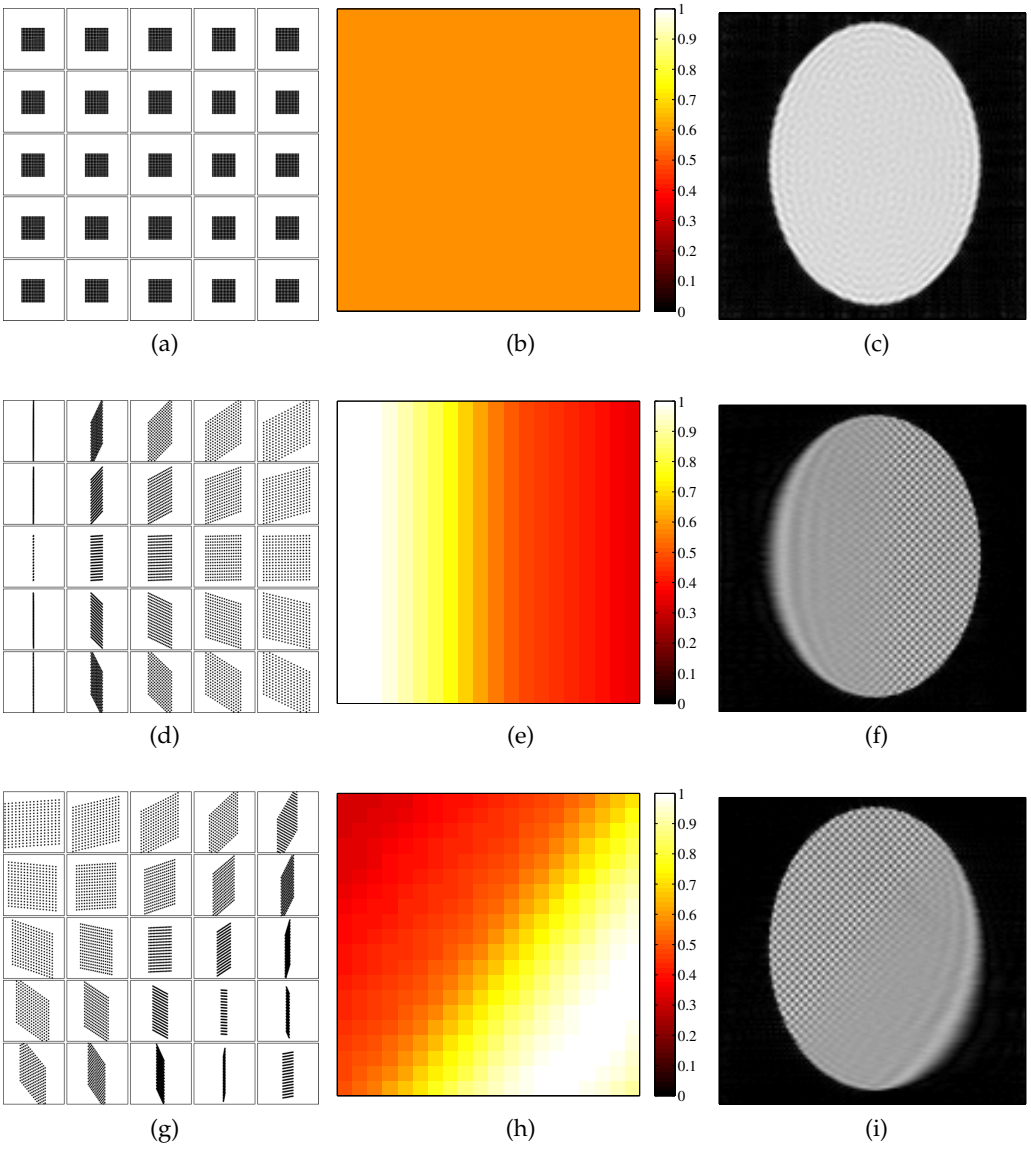


Figure 6.6: The (left) local *k*-space, (middle) estimated PSF width and (right) simulated reconstructions of a numerical checkerboard phantom for different gradient waveforms. The linear, 3-component and 4-component encoding schemes are presented in the top, middle and bottom rows, respectively.

gion we require positive values of $\kappa_c(t)$ during a left-to-right readout and negative values during a right-to-left readout line. This explains why the corresponding gradient waveform $g_c(t)$ follows the waveform of the x-coordinate, $g_x(t)$. This analysis also suggests that to enhance the left side of the image, we can simply flip the quadratic channel waveform by applying $-g_c(t)$ instead.

More subtle behaviour is observed in the y-coordinate, $k_y(\boldsymbol{x}, t)$. In this case, we are interested in the transition between readout lines and the appropriate gradient waveforms to increment the trajectory by a small amount. The gradient waveform between the end of readout line, τ_1 and the beginning of the next, τ_2 is symmetric about the time axis, so the corresponding gradient moment after this period is zero,

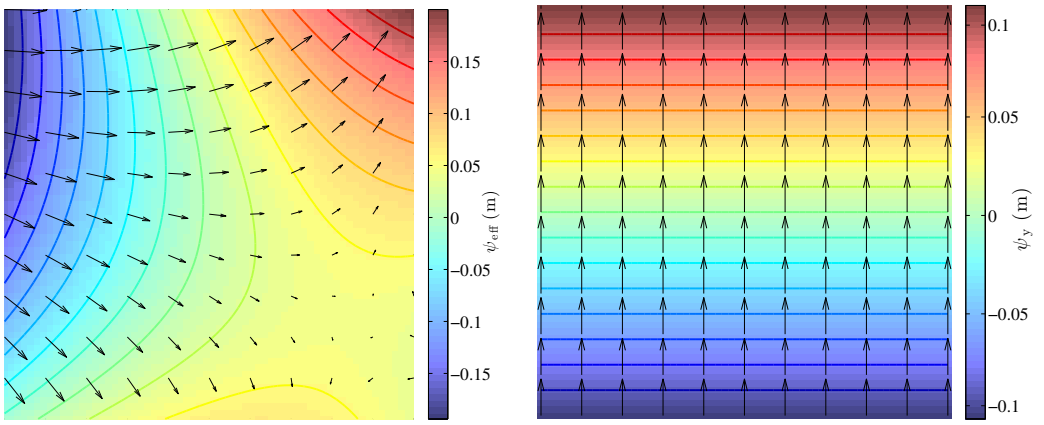
$$\kappa_c(t) = \int_{\tau_1}^{\tau_2} g_c(\tau) d\tau = 0. \tag{6.37}$$

Thus $k_y(\boldsymbol{x}, t) = \kappa_y(t)$ is only dependent on the y-gradient, which is not limited by amplitude or switching constraints for our simple target trajectory. Similarly, the local k-space extent in the y direction is not a function of position, unlike the x direction described above.

Another interpretation of the results is available by noting that during the readout the gradient waveforms of the additional channels are proportional to the linear readout gradient. This means the 'readout' encoding fields (all fields except ψ_y) can be replaced by an effective field, which is a linear combination of the encoding fields. To elucidate this point we consider the 4-component gradient waveforms in Fig. 6.5b. During the readout periods the gradient moments are related by constants λ and η such that $\kappa_a(t) = \lambda \kappa_x(t)$ and $\kappa_b(t) = \eta \kappa_x(t)$. According to (6.7), the phase can be written as

$$\phi(\boldsymbol{x}, t) = \kappa_x(t) \underbrace{\left(\psi_x(\boldsymbol{x}) + \lambda \psi_a(\boldsymbol{x}) + \eta \psi_b(\boldsymbol{x})\right)}_{\psi_{\text{eff}}(\boldsymbol{x})} + \kappa_y(t) \psi_y(\boldsymbol{x}). \tag{6.38}$$

We see that, in this case, the 4-component encoding is equivalent to traditional 2D encoding with a different readout field given by the effective encoding field $\psi_{\text{eff}}(\boldsymbol{x})$. Figure 6.7 illustrates the two effective encoding fields of the 4-component trajectory in Fig. 6.5b. The effective readout field in Fig. 6.7a has maximum gradient in the region of interest demonstrated by a narrow distance of neighbouring contours and length of the arrows representing the spatial gradient. The gradient is also maximally orthogonal to the phase encoding gradient in Fig. 6.7b in the region of interest.



(a) (b)

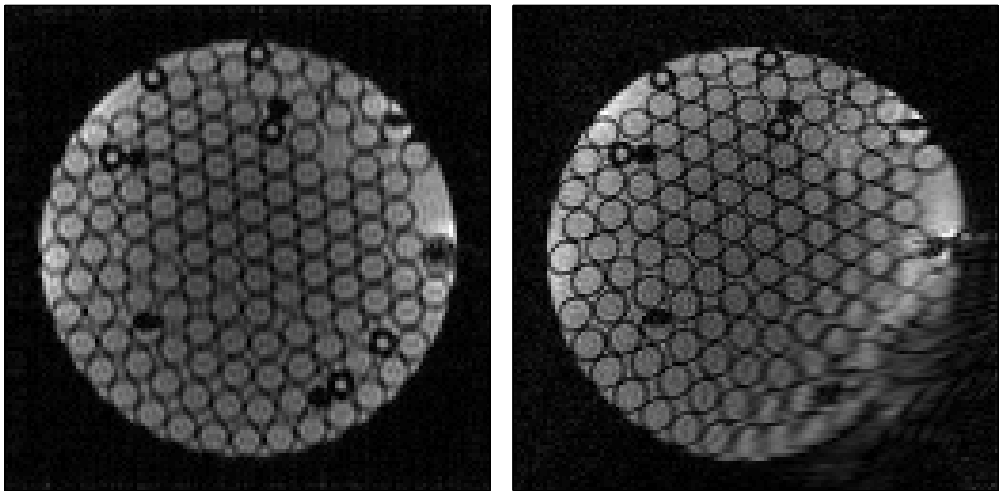
Figure 6.7: The two 'effective' encoding fields of the 4-component trajectory (measured in meters). (a) the readout encoding field exhibiting a spatially varying gradient and (b) the phase encoding field equivalent to traditional linear imaging. The coloured lines are the contours of the field and the arrows represent the spatial gradient.

This analysis gives an explanation of why the waveforms in Fig. 6.5 produce the results in Fig. 6.6, yet it is not intended as a design tool. Instead, the waveforms are obtained from the optimisation algorithm. The optimisation finds the gradients such that the local *k*-space trajectories are as close as possible to the target trajectory, given the gradient constraints and the available encoding fields.

6.5.2 Experiments

Figure 6.8 presents the reconstructed images from an EPI trajectory using two linear encoding fields and the optimised 4-component single-shot trajectory using two linear and two quadrupolar fields. The trajectory was measured directly using a field camera as described above. The spatially varying resolution characteristic of nonlinear encoding is evident in Fig. 6.8b with improved resolution in the top-left corner at the expense of pronounced blurring in other regions of the image. There are some small geometric distortions present in Fig. 6.8b; however, preliminarily investigations suggest these may be resolvable using an improved basis for fitting the field camera data. Figures 6.8c and 6.8d show the target region for the linear and 4-component reconstructions, respectively. This demonstrates the desired resolution improvement attainable using both linear and nonlinear encoding fields and an optimally designed trajectory.

During the in vivo acquisitions, no PNS was reported by the subject from either



(a) (b)

(c) (d)

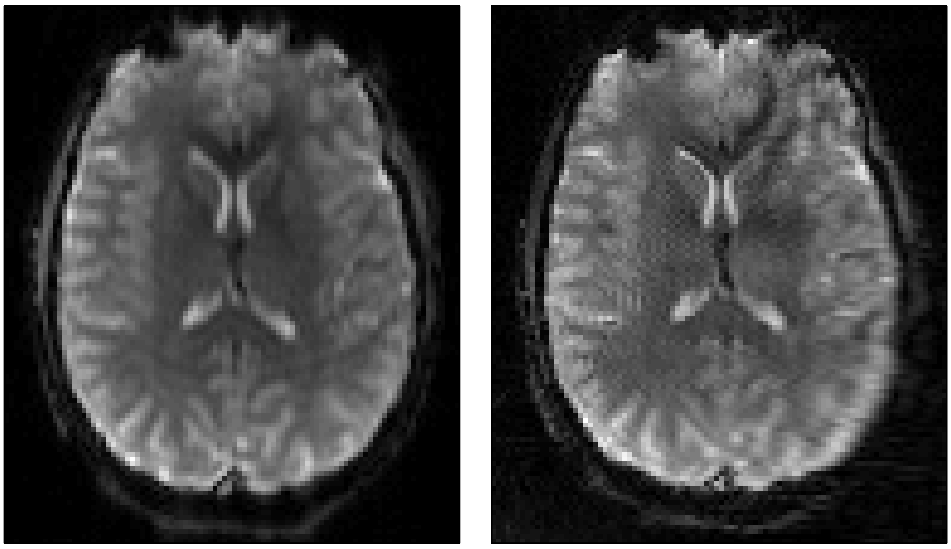
Figure 6.8: Reconstructions of a phantom experiment from (a) the linear EPI trajectory and (b) the optimised 4-component single-shot trajectory. The 4-component trajectory has been designed to provide improved resolution in the top-left region of the image. The target region is magnified in (c) and (d).

the linear-only or 4-component trajectories. Figure 6.9 displays the in vivo images, acquired using linear EPI and the optimised 4-component trajectory. The reconstructed image from the 4-component trajectory has a spatially varying resolution as predicted by the simulation results. As with the phantom images, the resolution is visibly enhanced in the desired region (Fig. 6.9d) compared to a standard EPI using only linear gradients (Fig. 6.9c). All the in vivo and phantom images have little geometric distortion or ghosting due to accurate knowledge of the trajectories, which were measured using the field monitoring device. However, there is a small checkerboard-like artifact present in the in-vivo image. This may be due to small residual errors in the calibration data. The images acquired using the 4-component trajectory also exhibit an elevated noise level within the target region, which reflects the usual trade-off between resolution and SNR.

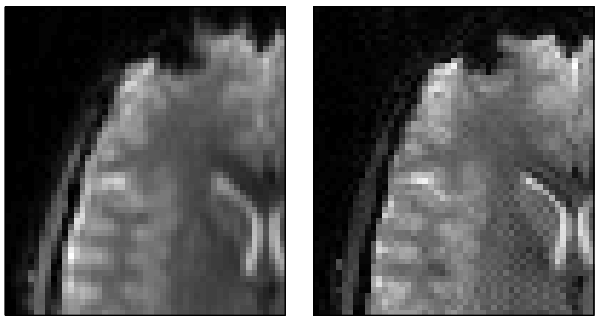
Figure 6.10 demonstrates the importance of using the field monitoring device to measure the trajectory. Figure 6.10a displays the 16 spherical harmonic coefficients used to define the encoding phase during the first 30 ms of the optimised single-shot trajectory. For a comparable display, the coefficients are scaled to represent the accumulated phase from each component at the edge of the field-of-view. As expected, the dominant components are the first order components x and y (κ_1 and κ_2, respectively) describing the linear gradients and the second order components xy and $x^2 - y^2$ (κ_4 and κ_8, respectively) representing the two hyperbolic paraboloid encoding fields. To illustrate an example of the differences between the nominal 4-component and the measured 16-component trajectories, Fig. 6.10b shows the initial rows of the resulting local k-space trajectories calculated at the first optimisation control point, $(\frac{L}{2}\cos 140^\circ, \frac{L}{2}\sin 140^\circ)$. The higher-order terms of the 16-component trajectory lead to a significant difference between the measured and nominal local k-space trajectories. In particular, the drift in the vertical direction is visible after only 8 rows of the trajectory, and is significantly worse by the end of the readout. Figure 6.10c displays a reconstructed image from the in vivo data using the nominal 4-component trajectory. The difference between the nominal and actual trajectories leads to significant artifacts and geometric distortion compared to Fig. 6.9b, which was reconstructed using the measured trajectories.

6.6 Discussion

We have developed a method for automated design of multidimensional single-shot trajectories that improves the imaging resolution in a region of interest. The use of additional encoding fields allowed us to overcome physical constraints concerning the rapid switching necessary for single-shot imaging.

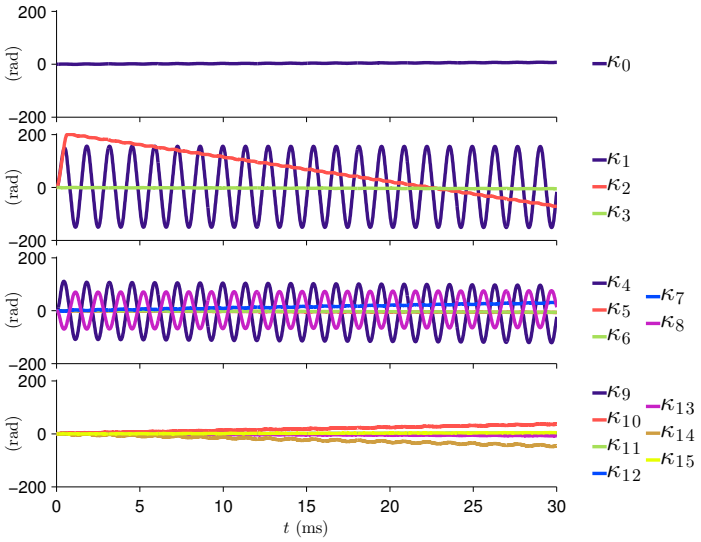


(a) (b)

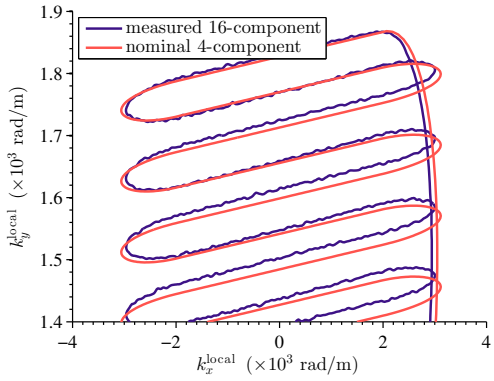


(c) (d)

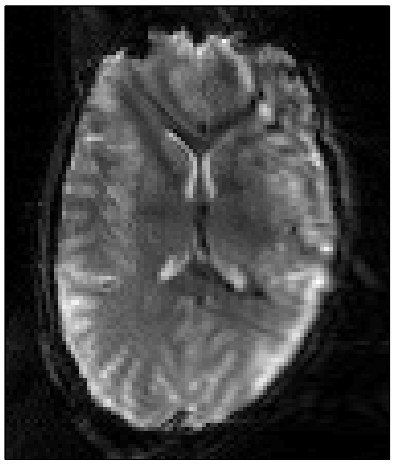
Figure 6.9: Reconstructions of a healthy brain from (a) the linear EPI trajectory and (b) the optimised 4-component single-shot trajectory. The region of interest is magnified in (c) and (d).



(a)



(b)



(c)

Figure 6.10: A comparison of the nominal and measured trajectories. (a) the 16 measured spherical harmonic coefficients, converted to the phase contributions (in radians) at a radius of $L/2$ from the centre; (b) the first 8 rows of the local k-space trajectory at the first optimisation control point, z_1, derived using the nominal 4-component trajectory and the measured 16-component trajectory; (c) the reconstruction of the in vivo data using the nominal 4-component trajectory, for comparison with the measured trajectory reconstruction in Fig. 6.9b.

The reconstructed images in this chapter have similarities to those presented in previous PatLoc literature. Images from a traditional Cartesian PatLoc trajectory using two orthogonal quadrupolar fields exhibit improved resolution at the periphery at the expense of pronounced blurring in the central region, where the local encoding gradient tends to zero [10]. The "Dual Cartesian" trajectory presented in [12] shifts the blurred region by simultaneously applying linear and nonlinear gradients. As required, the images presented in this chapter also exhibit regions with improved and degraded resolution; however, their locations are chosen prior to trajectory design. It is important to note that the region of improved resolution can be steered to other areas of the field-of-view, beyond what was demonstrated above. In general, improvement can be expected for any region towards the periphery, where these particular nonlinear fields are known to provide resolution enhancement.

The target trajectory used in this study was a simple EPI trajectory with a generous end-of-line transition time. This produced simple waveforms and allowed an analytical analysis of the results. However, the optimisation technique could also be applied to an EPI trajectory with ramp-sampling, allowing for a much shorter end-of-line transition [40]. In this case, the optimisation will trade off between increasing local k-space extent and compensating for the non-uniform sampling under the ramp. Additionally, the phase encode blips are often slew-rate limited and thus, analogous to the readout, the optimisation would include gradient moments from the nonlinear channels during the blip. This is a topic of follow-up studies.

A 16-channel field camera was used to measure the actual encoding trajectory, including higher order terms, in order to account for hardware imperfections, concomitant fields and eddy currents. A similar 16-channel field camera based on the same field sensors was used in [108] to minimise geometric distortion in diffusion-weighted images. In this work, where fast-switching trajectories drive both linear and nonlinear gradients, the field camera was necessary to prevent severe artifacts in the reconstructed images.

The design framework, based on the definition of local k-space, is applicable to any set of encoding fields. For example, the method can be applied when the analytical form of the encoding field shapes is not known or difficult to characterise, such as [115]. In this general case, it is sufficient to find a suitable set of basis functions that can accurately approximate the field shapes. Once this basis is chosen, the spatial derivatives can be computed and the appropriate optimisation problem can be defined.

The general trajectory design method proposed in this chapter is also applica-

ble to imaging with multiple excitations. The simplest modification would be to alter the slew rate constraints such that they only apply during each readout period. Since there is sufficient time between echoes, we can build up the necessary gradient moment prior to the next readout. Another technique is to parameterise the k-space trajectory to restrict the possible gradient moments, e.g. by the angle defining a spoke in a radial trajectory. This can be included in the optimisation so that we solve for the set of parameters instead of the trajectory directly.

An important problem is the likelihood of peripheral nerve stimulation (PNS) of the designed trajectories. In principle, a global resolution improvement is possible using a linear gradient coil insert and each channel could be switched independently to overcome the hardware constraints, as demonstrated in [116]. However, large rapid field changes would be present across the entire volume, increasing the risk of PNS [117]. On the other hand, an optimised nonlinear field trajectory would only apply large field changes over local regions, which may reduce the risk of nerve stimulation. Although we can prove that the time derivative of the B_z field is reduced, a complete investigation of this proposition would require a computational model for PNS risk including the influence of other switched field components (B_x, B_y) on the entire body [42]. The optimisation framework could be adapted to include such a model though this is beyond the scope of this thesis. Analysing this risk, however, will become increasingly important as the maximum gradient and slew rates are increased beyond the conservative values used in this work.

The spatially varying resolution characteristic of nonlinear encoding fields leads to an additional spatial dependence on the SNR and a trade-off exists between the two quantities. Indeed, the reconstructed images exhibited an increased noise level in regions with improved local resolution. It may be desirable in future work to include a noise metric, e.g. [118], into the optimisation algorithm such that a trajectory could be designed satisfying both resolution and SNR requirements, although the computational requirements may become prohibitive in this case.

The inherent resolution/SNR trade-off suggests the current framework is most useful for acquisition schemes limited primarily by resolution, such as single-shot imaging. As such, this technique is readily applicable to functional MRI studies concerning a particular anatomical region of the brain.

6.7 Conclusion

In this chapter, we have developed an automated procedure to create single-shot multidimensional trajectories that improve the resolution in a region of interest.

The use of linear and nonlinear encoding fields creates an additional degree of freedom that is used to steer the local resolution improvement. We have validated our technique in simulations as well as phantom and in vivo experiments for a hardware configuration consisting of two linear and two quadrupolar fields. A custom-built field camera was used to measure the actual trajectories, which was necessary to minimise artifacts in the reconstructed images.

CHAPTER 7

Noise performance for imaging with nonlinear encoding fields

Contents

7.1 Introduction

NONLINEAR spatial encoding has the potential to improve current imaging technologies by removing the linearity constraint on the gradient coils. These novel schemes require a characterisation of the imaging performance using quantitative metrics for resolution and noise. The previous chapter focused on the spatially varying resolution inherent in encoding with nonlinear fields. In this chapter,

we examine the noise properties of images reconstructed from nonlinear spatial encoding schemes.

In principle, analysis of the linear reconstruction employed by nonlinear encoding schemes is straightforward and the pixel covariance matrix is easily defined. The difficulty lies in computing this covariance for the large matrices associated with practical image resolutions. Fourier imaging and its variations possess a structure that can be exploited to simplify the computation. For example, the noise performance for traditional parallel imaging techniques such as SENSE [14] and GRAPPA [94] has been well quantified, using the familiar g-factor [14, 119]. In [10], it was shown that PatLoc is a generalisation of SENSE for nonlinear fields and thus inherits the analytical results concerning the SNR and g-factor, with a correction for the nonlinear field. However, the computation is intractable for arbitrary encoding schemes, such as O-space, 4D-RIO, Null Space Imaging, and the optimised schemes of Chapter 6. This motivates the development of an approximate metric. As the tractability of SENSE and PatLoc is due to the inherent structure of the encoding schemes, we categorise these schemes as "block-structured encoding schemes". Although the term PatLoc can refer to class of imaging schemes (as in [12]), in this chapter we use the term exclusively for the PatLoc scheme with a Cartesian trajectory described in Section 6.3.1.

In this chapter, we present a natural unification of block-structured and arbitrary encoding schemes by considering frames in the object space. Frame theory has primarily been used to analyse the windowed Fourier transform and wavelets. See [13, 120] for an introduction to this field. In the context of MRI, frames have been used to derive an SNR expression for schemes that use RF pulse encoding techniques in conjunction with linear encoding gradients [121]. Additionally, the "generalised point spread function" in [122] is described by a frame operator, although this was not explicitly mentioned in the paper, using it to assess different RF excitation schemes.

In this work, we apply the theory of frames to develop an analysis framework applicable to any MRI encoding scheme. Imaging performance is often measured by the image SNR, which ultimately depends on the variance of the reconstructed pixels. We derive an expression for the variance, which includes the previous results for SENSE and PatLoc while offering a novel understanding. Furthermore, we provide a viable means to analyse the performance of general encoding schemes which we demonstrate using O-space as a case study.

The chapter is organised as follows. In Section 7.2 we review frame theory in the MRI context. In Section 7.3 we derive an expression for the pixel variance and apply it to the block-structured encoding schemes of SENSE and PatLoc to gain

a novel perspective on these techniques. In Section 7.4 we use our framework to derive an approximation for the variance of the reconstructed pixels and apply it to O-space imaging. Finally, in Section 7.5 we demonstrate the wide applicability of our metric and present a quantitative comparison of PatLoc and O-space.

7.1.1 Notation

Table 7.1 lists the important symbols used in this chapter. Although the list is not exhaustive, it includes the majority of quantities. Additional notation will be introduced as it is required.

Table 7.1: Common notation used in Chapter 7

Symbol	Quantity
	Measurements
m	Magnetisation
$y_{l,q,i}$	Measurements indexed by coil, echo and time sample
N_c, N_e, N_t	Number of coils, excitations, time samples
c_l	Coil sensitivity
$\phi_{q,i}$	Encoding phase distribution
$\boldsymbol{\psi}$	Vector of field shapes
$\beta_{l,q,i}$	Encoding function
$b(\boldsymbol{x}, t)$	Encoding magnetic field
	Frame theory
f	Function to reconstruct
$\Phi_j, \tilde{\Phi}_j$	Frame elements, dual frame elements
A, B	Frame bounds
T, $\boldsymbol{T}$	Analysis operator (continuous and discrete)
T^*, $\boldsymbol{T}'$	Synthesis operator (continuous and discrete)
S, $\boldsymbol{S}$	Frame operator (continuous and discrete)
F, $\boldsymbol{F}$	Reconstruction operator (continuous and discrete)
χ, ξ, ζ	Pixel bases
	Analysis
D_N	Dirichlet kernel
$\boldsymbol{\Gamma}$	Weight matrix to whiten noise
$\boldsymbol{\Sigma}$	Measurement covariance
$\tilde{d}$	Volumetric correction factor
$\boldsymbol{D}$	Diagonal matrix of correction factors
$\boldsymbol{C}$	Coil sensitivity matrix
$\boldsymbol{\mathcal{F}}$	Fourier transform
$\varkappa$, $\tilde{\varkappa}$	Coil kernels
$h(\boldsymbol{u}, \boldsymbol{u}')$	Block structure kernel
$\hat{X}_n$	Variance approximation

7.2 Image reconstruction using frame theory

As described in Section 6.2, the general form of an acquired MRI signal is an integral transform of the underlying magnetisation. Ignoring relaxation effects the signal equation is

$$s_l(t) = \int m(\boldsymbol{x}) c_l(\boldsymbol{x}) e^{-j\phi(\boldsymbol{x},t)} d\boldsymbol{x}. \tag{7.1}$$

where c_l is the coil sensitivity of the l^{th} coil, ϕ is the accumulated phase distribution and m is unknown magnetisation.

In general, the signal is acquired over a number of RF excitations, each of which resets the phase distribution. Additionally, measurements are collected at discrete time points, denoted t_i. We introduce the notation $\phi_{q,i}(\boldsymbol{x})$ which represents the phase distribution for the q^{th} excitation and i^{th} time point. The discrete measurements, denoted $y_{l,q,i}$, can be viewed as inner products on the space of functions,

$$y_{l,q,i} = \langle m, \beta_{l,q,i} \rangle, \tag{7.2}$$

where there is an index for each coil, l, echo, q, and time, i. The encoding functions, $\beta_{l,q,i}$ are given by

$$\beta_{l,q,i}(\boldsymbol{x}) = c_l(\boldsymbol{x}) e^{j\phi_{q,i}(\boldsymbol{x})}. \tag{7.3}$$

The reconstruction problem is to estimate the object, $m(\cdot)$, from a set of projections $\{y_{l,q,i}\}$. From this perspective, the reconstruction problem lends itself to a frame theoretic formulation. Frame theory provides a performance analysis and optimal reconstruction algorithm for a function observed through a sequence of projections, exactly the situation constructed in MRI. This formulation underscores the idea that the MRI reconstruction problem exists within a broader mathematical context. Such abstraction may lead to further advances in the field of image reconstruction.

7.2.1 Review of frame theory

We proceed with a brief introduction to frame theory followed by a matrix formulation that is readily applied to the discretised objects used for MRI reconstruction. For a more rigourous and detailed account consult, for example, [13].

Frame theory is concerned with functions, f, belonging to a Hilbert space, $\mathcal{V}$. In the context of MRI, f is the unknown magnetisation and $\mathcal{V}$ is the space of phys-

ically plausible magnetisation distributions.

By definition, a set of functions, $\{\Phi_j\}_{j\in\mathcal{I}}$, in a space, $\mathcal{V}$, defined over a countable set, $\mathcal{I}$, is a frame if there exists constants, $0 < A \leq B < \infty$ such that for all $f \in \mathcal{V}$,

$$A\|f\|^2 \leq \sum_{j\in\mathcal{I}} |\langle f, \Phi_j\rangle|^2 \leq B\|f\|^2 \tag{7.4}$$

where A and B are the lower and upper frame bounds, respectively. We refer to the functions $\{\Phi_j\}_{j\in\mathcal{I}}$ as frame elements. The frame bounds reflect the robustness of the frame elements to perturbations, e.g. additive random noise on the projections. If $A = B = 1$ and $\|\Phi_j\| = 1, \forall j \in \mathcal{I}$, then $\{\Phi_j\}_{j\in\mathcal{I}}$ is an orthonormal basis. In this sense, the frame elements can be thought of as a generalised basis. The set $\mathcal{I}$ can be infinite for theoretical analysis although it is typically finite for practical applications.

In most applications, we obtain measurements (or frame coefficients) of an unknown function via the analysis operator, $T : \mathcal{V} \to \mathbb{C}^M$ where M is the number of frame elements. The operator is defined as

$$Tf = \{\langle f, \Phi_j\rangle\}_{j\in\mathcal{I}}. \tag{7.5}$$

Similarly, we define the synthesis operator $T^* : \mathbb{C}^M \to \mathcal{V}$ as

$$T^*\{y_j\}_{j\in\mathcal{I}} = \sum_{j\in\mathcal{I}} y_j\Phi_j(\boldsymbol{x}), \tag{7.6}$$

which transforms the sequence of coefficients, $\{y_j\}_{j\in\mathcal{I}}$, into a function in $\mathcal{V}$. Notice that T^* is the adjoint of T. The frame operator $S : \mathcal{V} \to \mathcal{V}$ is the composition, $S = T^* \circ T$ or

$$(Sf)(\boldsymbol{x}) = \sum_{j\in\mathcal{I}} \langle f, \Phi_j\rangle\Phi_j. \tag{7.7}$$

If the functions, $\{\Phi_j\}_{j\in\mathcal{I}}$, form a frame by satisfying the frame condition in (7.4), we can reconstruct any function from a sequence of noiseless transform coefficients as

$$f = \sum_{j\in\mathcal{I}} \langle f, \Phi_j\rangle\tilde{\Phi}_j = \sum_{j\in\mathcal{I}} \langle f, \tilde{\Phi}_j\rangle\Phi_j, \tag{7.8}$$

where $\{\tilde{\Phi}_j\}_{j\in\mathcal{I}}$ is the dual frame with the functions given by

$$\tilde{\Phi}_j = S^{-1}\Phi_j. \tag{7.9}$$

The relationship in (7.8) is exact and valid as long as (7.4) is satisfied. Note the duality in (7.8) in the sense that either Φ_j or $\tilde{\Phi}_j$ can be used for decomposition with the other used for reconstruction. Adopting the convention of measuring with the frame elements, $\{\Phi_j\}_{j\in\mathcal{I}}$, and reconstructing with the dual, we can use (7.8) to define the reconstruction operator, $F : \mathbb{C}^M \to \mathcal{V}$ as the composition,

$$F = S^{-1} \circ T^*. \tag{7.10}$$

In MRI the unknown function is the object's magnetisation, i.e. $f = m$ and the operator in (7.10) would be used to reconstruct the magnetisation, m, from a sequence of measurements, $\{y_{l,q,i}\}$. The theory informs us that if the frame elements satisfy the frame condition in (7.4) then reconstruction is possible using the dual frame. Hence, it is desirable to find a frame in the two or three dimensional object space using encoding functions realisable in the magnetic resonance setting. For MRI it is useful to consider frame elements $\Phi_{l,q,i}$ indexed by coil, echo, and time sample. The results above are readily applicable to these frame elements by considering indices from the set of tuples, $\mathcal{I} = \{(l,q,i) \in [1,\ldots,N_c] \times [1,\ldots,N_e] \times [1,\ldots,N_t]\}$, where N_c, N_e and N_t denote the number of coils, echoes and time samples, respectively.

In practice, we can only acquire a finite number of measurements so reconstruction of an arbitrary continuous function is impossible. Consequently, we are forced to restrict the space of functions by discretising the unknown functions into pixels. For this purpose we select a pixel basis, $\{\chi_n(\boldsymbol{x})\}_{n=1}^N$, so the object can be represented by a linear combination of the functions, $\chi_n(\boldsymbol{x})$. For example, images can be represented using a rectangular basis where $\chi_n(\boldsymbol{x}) = \mathrm{rect}_W(\boldsymbol{x} - \boldsymbol{x}_n)$ for pixel width W. Mathematically, the basis is used to define a subspace, $\mathcal{U}$, representing the space of allowable functions,

$$\mathcal{U} = \left\{ f \in \mathcal{V} : f(\boldsymbol{x}) = \sum_{n=1}^{N} f_n \chi_n(\boldsymbol{x}) \right\}. \tag{7.11}$$

In this space, the problem of reconstructing an arbitrary object is transformed to the problem of estimating a set of coefficients, $\{f_n\}$.

Since a function $f \in \mathcal{U}$ is entirely described by the coefficients, $f_1,\ldots,f_N$, the subspace $\mathcal{U}$ is isomorphic to $\mathbb{C}^N$. Therefore the analysis operator, T, acting on functions in $\mathcal{U}$ can be considered as a mapping from $\mathbb{C}^N$ to $\mathbb{C}^M$ and thus admits a matrix representation, $\boldsymbol{T}$. The corresponding matrix has elements given by

$$T_{(l,q,i),n} = \langle \chi_n, \Phi_{l,q,i} \rangle. \tag{7.12}$$

The synthesis operator in this subspace, $T^* : \mathbb{C}^M \to \mathcal{U}$, also has a matrix representation,

$$T^*_{n,(l,q,i)} = \langle \Phi_{l,q,i}, \chi_n \rangle. \tag{7.13}$$

This can be represented as $\boldsymbol{T}'$, where $\boldsymbol{T}'$ is the conjugate transpose of the $\boldsymbol{T}$ defined by (7.12).

The frame operator in the discrete subspace is $S : \mathcal{U} \to \mathcal{U}$, defined by the composition $S = T^* \circ T$. Equivalently this operator is represented as a matrix, called the frame matrix, defined by the equation,

$$\boldsymbol{S} = \boldsymbol{T}'\boldsymbol{T} \tag{7.14}$$

A common example is when $\boldsymbol{T}$ is a full-rank matrix with orthonormal columns (such as a DFT matrix). In this case, the inverse is given by $\boldsymbol{T}'$ and the frame matrix $\boldsymbol{S}$ is equal to the identity matrix.

The reconstruction operator using the dual frame in (7.10) simplifies to the Moore-Penrose pseudoinverse of the discrete analysis operator, $\boldsymbol{T}$,

$$\boldsymbol{F} = \boldsymbol{S}^{-1}\boldsymbol{T}' = (\boldsymbol{T}'\boldsymbol{T})^{-1}\boldsymbol{T}'. \tag{7.15}$$

When we select the frame elements to be the encoding functions used in MRI, $\Phi_{l,q,i} = \beta_{l,q,i}$, the analysis matrix in (7.12) becomes the standard encoding matrix [14],

$$E_{(l,q,i),n} = \langle \chi_n, \beta_{l,q,i} \rangle \tag{7.16}$$

In this case the measurements $y_{l,q,i}$ correspond to the frame coefficients.

An interesting question, given little attention in the literature, is how to choose the pixel basis, $\{\chi_n\}_{n=1}^N$. The pixel basis must satisfy two conflicting requirements; firstly, the basis should lead to simple computation of the frame matrix and its inverse, and secondly the basis should be able to accurately approximate the true functions. The Dirac delta distribution is ubiquitous since it satisfies the requirement of simple computation. A rectangular basis potentially better represents the functions. Appendix 7.A compares the two choices and shows that the difference in the final reconstructions for Fourier encoding is $\mathcal{O}(N^{-5/2})$ where N is the number of pixels. This confirms the common notion that Dirac distributions sufficiently approximate the functions if the pixel grid is fine enough.

7.3 Analysis of block-structured encoding schemes

The impact of stochastic noise in the measurements must be considered for practical applications of the presented theory. In the frame literature, reconstruction using the dual frame has been shown to give the minimum mean-square error (MMSE) estimate, with global performance given by the lower frame bound [120]. However, exact quantification in terms of the variance of each pixel is desirable for MRI applications. In this section we relate the pixel-wise variance to the frame operator and apply the theory to the block-structured encoding schemes of SENSE and PatLoc.

7.3.1 Reconstruction variance

The signal model in (7.2) is modified to include additive noise,

$$y_{l,q,i} = \langle m, \beta_{l,q,i} \rangle + v_{l,q,i}. \tag{7.17}$$

A discretised object can be represented by the vector, $\boldsymbol{f} = [f_1, \ldots, f_N]^T$, where the elements are $f_n = \langle \chi_n, m \rangle$. We construct the vectors, $\boldsymbol{y} = [y_{1,1,1}, \ldots, y_{N_c,N_e,N_t}]^T$ and $\boldsymbol{v} = [v_{1,1,1}, \ldots, v_{N_c,N_e,N_t}]^T$ so (7.17) can be written as the matrix equation,

$$\boldsymbol{y} = \boldsymbol{E}\boldsymbol{f} + \boldsymbol{v}. \tag{7.18}$$

The noise is complex Gaussian distributed, $\boldsymbol{v} \sim \mathcal{N}(0, \boldsymbol{\Sigma})$ for a general covariance matrix $\boldsymbol{\Sigma}$. The size of $\boldsymbol{\Sigma}$ is $N_c N_e N_t \times N_c N_e N_t$.

We desire a frame operator that adequately models the covariance so we consider the general case where the frame elements are a linear combination of encoding functions,

$$\Phi_{l,q,i} = \sum_{l',q',i'} \Gamma_{(l,q,i),(l',q',i')} \beta_{l',q',i'}. \tag{7.19}$$

This leads to frame coefficients that are a linear combination of the MRI measurements. Let $\boldsymbol{\Gamma}$ be a matrix of weights with elements $\Gamma_{(l,q,i),(l',q',i')}$; the coefficients can be defined by the matrix equation, $\boldsymbol{\eta} = \boldsymbol{\Gamma}\boldsymbol{y}$, with

$$\boldsymbol{\eta} \sim \mathcal{N}(\boldsymbol{\Gamma}\boldsymbol{E}\boldsymbol{f}, \boldsymbol{\Gamma}\boldsymbol{\Sigma}\boldsymbol{\Gamma}'). \tag{7.20}$$

The analysis operator is $\boldsymbol{T} = \boldsymbol{\Gamma}\boldsymbol{E}$ and recall the frame matrix is $\boldsymbol{S} = \boldsymbol{T}'\boldsymbol{T}$. The reconstructed function coefficients are given by $\hat{\boldsymbol{f}} = \boldsymbol{F}\boldsymbol{\eta}$, where $\boldsymbol{F}$ is the reconstruc-

tion operator in (7.15). The covariance, $\boldsymbol{X}$, of the reconstructed coefficients is

$$\boldsymbol{X} = \boldsymbol{F}\boldsymbol{\Gamma}\boldsymbol{\Sigma}\boldsymbol{\Gamma}'\boldsymbol{F}' \tag{7.21}$$

$$= (\boldsymbol{T}'\boldsymbol{T})^{-1}\boldsymbol{T}'\boldsymbol{\Gamma}\boldsymbol{\Sigma}\boldsymbol{\Gamma}'\boldsymbol{T}(\boldsymbol{T}'\boldsymbol{T})^{-1} \tag{7.22}$$

Since $\boldsymbol{\Sigma}$ is a covariance matrix, it is conjugate symmetric and thus the following decomposition is well defined, $\boldsymbol{\Sigma}^{-1} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}^{-1/2}$. The main result of this section is that selecting the set of weights to be

$$\Gamma_{(l,q,i),(l',q',i')} = [\boldsymbol{\Sigma}^{-1/2}]_{(l,q,i),(l',q',i')} \tag{7.23}$$

yields

$$\boldsymbol{X} = \boldsymbol{S}^{-1} \tag{7.24}$$

This equation reveals the close link between the frame elements, frame operator, and the resulting covariance. Importantly, the relationship is applicable for any noise covariance, by the appropriate selection of the weighting matrix.

When the condition in (7.23) is satisfied, the frame matrix can be written in terms of the discretised encoding functions as

$$\boldsymbol{S} = \boldsymbol{E}'\boldsymbol{\Sigma}^{-1}\boldsymbol{E} \tag{7.25}$$

In terms of the matrix elements we have,

$$S_{n,m} = \sum_{l,q,i}\sum_{l',q',i'} \langle \beta_{l,q,i}, \chi_n \rangle \Sigma^{-1}_{(l,q,i),(l',q',i')} \langle \chi_m, \beta_{l',q',i'} \rangle. \tag{7.26}$$

This process of transforming the measurements is the same as 'noise whitening' in linear estimation theory [6].

For white noise, $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}$, and the pixel variance is a scalar multiple of that obtained by adopting an identity weighting matrix. In this case, the frame elements are simply the MRI encoding functions. For correlated noise, the selection of transformation weights, ensures the frame operator captures the properties of the reconstruction covariance. This is useful to analyse the noise performance of existing parallel imaging schemes, where noise is correlated between receive channels [14].

In the remainder of this section we apply our analysis to the existing imaging schemes of SENSE and PatLoc. The analysis of SENSE confirms the well-known results concerning SNR and coil sensitivities. Our analysis of PatLoc demonstrates

its relationship to SENSE as described in [10]. For both techniques we present the existing analysis reported in the literature and our analysis using frame theory. The intention here is to provide a novel perspective on the workings of these schemes.

7.3.2 SENSE imaging

SENSE imaging is a type of Fourier imaging and thus there is a linear relationship between the phase and position, $\phi_{q,i}(\boldsymbol{x}) = -2\pi \boldsymbol{k}_{q,i} \cdot \boldsymbol{x}$ as described in Section 3.4.4. The main feature of SENSE is that the encoding functions include the sensitivities from multiple coils,

$$\beta_{l,q,i}(\boldsymbol{x}) = c_l(\boldsymbol{x}) e^{j2\pi \boldsymbol{k}_{q,i} \cdot \boldsymbol{x}}. \tag{7.27}$$

The quantity $\boldsymbol{k}_{q,i}$ defines the position in spatial frequency space, k-space, as described in Section 3.4.4. Measurements are collected for a set of k-space coordinates (the trajectory), parameterised by time sample i and excitation number q. The gradients are changed after each excitation, defining the direction of the line through k-space, which is then traversed as time evolves.

For faster acquisition, one can undersample by integer factors R_x and R_y in the x and y directions, respectively. This reduces the number of measurements by a factor of $R = R_x R_y$. Practically, only undersampling in the phase direction is useful to reduce imaging time since the sampling time in the readout direction is negligible. Nonetheless, our analysis is kept general to cover the case of 3D Fourier imaging with two phase encode directions [123]. We collect $N_k = N_t N_e$ measurements per coil with $N_t = N_x/R_x$, $N_e = N_y/R_y$. Given a field-of-view of $L_x \times L_y$, the maximum frequency, $k_x^{\max} = N_x/L_x$, $k_y^{\max} = N_y/L_y$ is unchanged by the undersampling, indicating the fundamental resolution is not affected. The measurements are collected at frequency points, $\boldsymbol{k}_{q,i} = (i\tilde{\Delta}_x, q\tilde{\Delta}_y)$ where $\tilde{\Delta}_x = R_x\Delta_x = R_x/L_x$ and likewise $\tilde{\Delta}_y = R_y\Delta_y$.

7.3.2.1 Existing analysis of SENSE

The standard SENSE analysis originally presented in [14] proceeds by adopting the delta distribution as the pixel basis and analysing the corresponding matrix equation for the measurement vector. Identical to (7.18) we have,

$$\boldsymbol{y} = \boldsymbol{E}\boldsymbol{f} + \boldsymbol{v} \tag{7.28}$$

where $\boldsymbol{v} \sim \mathcal{N}(0, \boldsymbol{\Sigma})$ and $E_{(l,q,i),n} = c_l(\boldsymbol{x}_n)e^{-j2\pi \boldsymbol{k}_{q,i} \cdot \boldsymbol{x}_n}$. The optimal reconstruction matrix was proved to be [14]

$$\boldsymbol{F} = (\boldsymbol{E}'\boldsymbol{\Sigma}^{-1}\boldsymbol{E})^{-1}\boldsymbol{E}'\boldsymbol{\Sigma}^{-1}. \tag{7.29}$$

It is observed that for a given pixel centred at $\boldsymbol{x}_p = (x_p, y_p)$, the group of aliased pixels due to undersampling is given by the set,

$$\begin{aligned} \mathcal{P}_p = \big\{ &(x_p + nL_x/R_x,\ y_p + mL_y/R_y) : \\ &n \in [0, \ldots, R_x - 1], m \in [0, \ldots, R_y - 1] \big\} \end{aligned} \tag{7.30}$$

The problem of reconstruction is reduced to resolving this small set of pixels. We denote the $R \times 1$ vector of coefficients for the aliased pixels as $\boldsymbol{f}^{(p)}$. After performing an inverse DFT, this group of aliased pixels are related by

$$\boldsymbol{\eta} = \boldsymbol{C}^{(p)}\boldsymbol{f}^{(p)} \tag{7.31}$$

where $\boldsymbol{C}^{(p)}$ is an $N_c \times R$ matrix with elements $\boldsymbol{C}^{(p)}_{ln} = c_l(\boldsymbol{x}_n)$. The folding matrix, $\boldsymbol{C}^{(p)}$, highlights the crucial role of the coil sensitivities in successfully estimating the function coefficients. This structure is also revealed in the calculation of the reconstruction covariance matrix,

$$\boldsymbol{X} = \boldsymbol{F}\boldsymbol{\Sigma}\boldsymbol{F}' \tag{7.32}$$

$$= \frac{1}{N_t N_e}(\boldsymbol{E}'\boldsymbol{\Sigma}^{-1}\boldsymbol{E})^{-1} \tag{7.33}$$

Exploiting the block structure of $\boldsymbol{E}$ and $\boldsymbol{\Sigma}$ [14], the noise matrix can be defined for each group of aliased pixels,

$$\boldsymbol{X}^{(p)} = \frac{1}{N_t N_e}(\boldsymbol{C}^{(p)'}{\boldsymbol{\Sigma}^{(p)}}^{-1}\boldsymbol{C}^{(p)})^{-1} \tag{7.34}$$

which involves the inversion of a small $R \times R$ matrix. The well-known g-factor is derived by considering the ratio of SNRs for the reduced and full acquisitions. A pixel is now indexed as (r, p) indicating the pixel and aliased group, respectively. For notational simplicity, the superscript (p) is dropped in the following g-factor equations. For full acquisition the covariance is denoted $\bar{\boldsymbol{X}}$ and $N_t = N_x$, $N_e = N_y$.

Assuming that the signal strength is constant,

$$\frac{\mathrm{SNR}_{r,p}^{\mathrm{full}}}{\mathrm{SNR}_{r,p}^{\mathrm{red}}} = \frac{\sqrt{\boldsymbol{X}_{r,r}}}{\sqrt{\bar{\boldsymbol{X}}_{r,r}}} = \sqrt{R}\frac{\sqrt{[(\boldsymbol{C}'\boldsymbol{\Sigma}^{-1}\boldsymbol{C})^{-1}]_{r,r}}}{\sqrt{[(\bar{\boldsymbol{C}}'\boldsymbol{\Sigma}^{-1}\bar{\boldsymbol{C}})^{-1}]_{r,r}}} \tag{7.35}$$

The case of full Fourier encoding yields a $\bar{\boldsymbol{C}}$ that is an $N_c \times 1$ column vector. In this situation there is no aliasing so our blocks contain only a single pixel and $\bar{\boldsymbol{C}}$ is identical to the n^{th} column of the $\boldsymbol{C}$ matrix. Thus,

$$\frac{1}{\hat{X}_{r,r}} = [\boldsymbol{C}'\boldsymbol{\Sigma}^{-1}\boldsymbol{C}]_{r,r} \tag{7.36}$$

Finally, the ratio of SNRs is expressed as

$$\frac{\mathrm{SNR}_{r,p}^{\mathrm{full}}}{\mathrm{SNR}_{r,p}^{\mathrm{red}}} = \sqrt{R}\sqrt{[\boldsymbol{C}'\boldsymbol{\Sigma}^{-1}\boldsymbol{C}]_{r,r}[(\boldsymbol{C}'\boldsymbol{\Sigma}^{-1}\boldsymbol{C})^{-1}]_{r,r}} \tag{7.37}$$

From this, the g-factor is defined as

$$g_{r,p} = \sqrt{[\boldsymbol{C}'\boldsymbol{\Sigma}^{-1}\boldsymbol{C}]_{r,r}[(\boldsymbol{C}'\boldsymbol{\Sigma}^{-1}\boldsymbol{C})^{-1}]_{r,r}} \tag{7.38}$$

Recall $\boldsymbol{C}$ and $\boldsymbol{\Sigma}$ are indexed by the aliased block, p, and the pair (r, p) is needed to index a pixel.

7.3.2.2 Frame analysis of SENSE

The g-factor analysis above emerges naturally from the frame matrix, which emphasises the source of the block structure. Central to this analysis is the periodic sinc, also known as the Dirichlet kernel, $D_N(x)$, defined as

$$D_N(x) := \sum_{k=-(N-1)/2}^{(N-1)/2} e^{jkx}. \tag{7.39}$$

This definition can be obtained by taking the Discrete Time Fourier Transform (DTFT) of a rectangular window with length N. See Appendix 4.A for a brief introduction to the Dirichlet kernel and its properties. In this work, we will see that the block structure in SENSE is revealed by evaluating the Dirichlet kernel.

The pixel covariance is calculated from the frame matrix as explained in Section 7.3.1. As such, we select frame elements as a linear combination of the SENSE encoding functions according to the 'noise whitening' condition in (7.23). In this case, the frame matrix elements are given by (7.26). Additionally, the noise in a

parallel MRI experiment is correlated between channels, not between frequency samples or echoes, thus $\Sigma^{-1}_{(l,q,i),(l',q',i')} = \delta_{q,q'}\delta_{i,i'}\Sigma^{-1}_{l,l'}$. This reduces the matrix elements to

$$S_{n,m} = \sum_{l,l'} c^*_l(\boldsymbol{x}_n)\Sigma^{-1}_{l,l'}c_{l'}(\boldsymbol{x}_m)\sum_{q,i} e^{j2\pi \boldsymbol{k}_{q,i}\cdot(\boldsymbol{x}_n-\boldsymbol{x}_m)} \tag{7.40}$$

$$\begin{aligned} &= \varkappa(\boldsymbol{x}_n,\boldsymbol{x}_m)D_{N_t}(2\pi R_x\Delta_x(\boldsymbol{x}_n-\boldsymbol{x}_m)\cdot\hat{\boldsymbol{\imath}})\times \\ &\qquad D_{N_e}(2\pi R_y\Delta_y(\boldsymbol{x}_n-\boldsymbol{x}_m)\cdot\hat{\boldsymbol{\jmath}}) \end{aligned} \tag{7.41}$$

where $\hat{\boldsymbol{\imath}} = (1,0)$, $\hat{\boldsymbol{\jmath}} = (0,1)$ and $\varkappa$ is the noise-adjusted coil kernel given by

$$\varkappa(\boldsymbol{x}_n,\boldsymbol{x}_m) = \sum_{l,l'} c^*_l(\boldsymbol{x}_n)\Sigma^{-1}_{l,l'}c_{l'}(\boldsymbol{x}_m). \tag{7.42}$$

This coil kernel arises from the use of multiple receiver coils with spatially varying sensitivities.

In our application, the undersampling means that the Dirichlet kernel in the x and y directions will have a period of L_x/R_x and L_y/R_y, respectively, so R points will not be resolved by Fourier encoding alone. The additional encoding comes from the coil kernel, $\varkappa(\boldsymbol{x},\boldsymbol{x}')$, which resolves the ambiguity.

A key to SENSE imaging is the careful selection of the pixel size, $W_x \times W_y$, and sampling spacing, Δ_x, Δ_y such that $\Delta_x W_x = 1/N_t$ and $\Delta_y W_y = 1/N_e$. In this case, the Dirichlet kernels evaluate to

$$D_{N_t}(2\pi R_x\Delta_x(\boldsymbol{x}_n-\boldsymbol{x}_m)\cdot\hat{\boldsymbol{\imath}}) = \begin{cases} N_t, & \boldsymbol{x}_n,\boldsymbol{x}_m \in \mathcal{P}^{(x)}_p; \\ 0, & \text{otherwise.} \end{cases} \tag{7.43}$$

and

$$D_{N_e}(2\pi R_y\Delta_y(\boldsymbol{x}_n-\boldsymbol{x}_m)\cdot\hat{\boldsymbol{\jmath}}) = \begin{cases} N_e, & \boldsymbol{x}_n,\boldsymbol{x}_m \in \mathcal{P}^{(y)}_p; \\ 0, & \text{otherwise.} \end{cases} \tag{7.44}$$

The sets $\mathcal{P}^{(x)}_p$ and $\mathcal{P}^{(y)}_p$ are all points aliased onto the point $\boldsymbol{x}_p$ in the x and y directions, respectively, such that $\mathcal{P}_p = \mathcal{P}^{(x)}_p \cap \mathcal{P}^{(y)}_p$. Thus,

$$S_{n,m} = \begin{cases} N_tN_e\varkappa(\boldsymbol{x}_n,\boldsymbol{x}_m), & \boldsymbol{x}_n,\boldsymbol{x}_m \in \mathcal{P}_p; \\ 0, & \text{otherwise.} \end{cases} \tag{7.45}$$

Furthermore, if the basis, $\{\chi_n\}_{n=1}^N$, is ordered according to the groups in $\mathcal{P}_p$ then $\boldsymbol{S}$

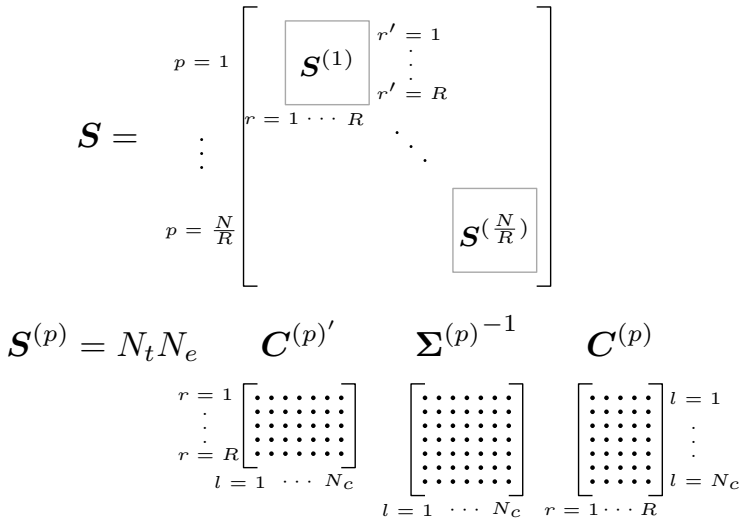


Figure 7.1: The block diagonal structure of the frame matrix, $\boldsymbol{S}$. Each block represents a group of aliased pixels and depends on the number of measurements, the coil sensitivities and the measurement noise covariance.

will have a block diagonal structure with blocks given by,

$$\boldsymbol{S}^{(p)} = N_t N_e \boldsymbol{C}^{(p)'} \boldsymbol{\Sigma}^{(p)^{-1}} \boldsymbol{C}^{(p)}. \tag{7.46}$$

The structure of the frame matrix given by (7.46) is illustrated in Fig. 7.1.

Inversion of the block diagonal frame operator is achieved by inversion of the individual blocks. Thus,

$$\boldsymbol{X}^{(p)} = \frac{1}{N_t N_e} (\boldsymbol{C}^{(p)'} \boldsymbol{\Sigma}^{(p)^{-1}} \boldsymbol{C}^{(p)})^{-1}, \tag{7.47}$$

identical to (7.34). The expression for g-factor follows as in (7.38).

This perspective emphasises the role of the Dirichlet kernel in Fourier encoding. The kernel explains the source of aliasing at its most fundamental level and provides an intuition into the structure of the system. Evaluation of the kernels reveals the block structure exactly as we expect with our prior understanding of aliasing. In both the existing analysis and our frame-based analysis, the structure of the system is exploited to derive the g-factor.

7.3.3 PatLoc imaging

Understanding of the PatLoc imaging technique described in Section 6.3.1 can also benefit from analysis in terms of the frame matrix. In this section we again show both the previously published analysis in [10] and our frame-based approach. Pat-

Loc employs nonlinear non-bijective fields, and similar to SENSE imaging, the additional encoding provided by the receive coils is crucial to resolving the non-bijectivity.

The phase function for general imaging schemes is described by (6.7). Introducing indices for the q^{th} excitation and i^{th} time point the accumulated phase can be written as $\phi_{q,i}(\boldsymbol{x}) = \boldsymbol{\kappa}_{q,i} \cdot \boldsymbol{\psi}(\boldsymbol{x})$, where $\boldsymbol{\psi}$ is the nonlinear and non-bijective mapping associated with the spatial encoding magnetic fields. The corresponding encoding functions are

$$\beta_{l,q,i}(\boldsymbol{x}) = c_l(\boldsymbol{x}) e^{j\boldsymbol{\kappa}_{q,i} \cdot \boldsymbol{\psi}(\boldsymbol{x})}. \tag{7.48}$$

7.3.3.1 Existing analysis of PatLoc

The existing PatLoc analysis in [10] showed that PatLoc can be thought of as an extension of SENSE for nonlinear fields. Consequently, the noise analysis can be performed identically to the SENSE analysis in [14] discussed earlier and similar results are reported concerning the SNR and g-factor. This is elaborated below.

The PatLoc analysis involves the selection of a pixel basis, $\{\chi_n\}$, for the discretisation of the encoding operator according to (7.16).

$$E_{(l,q,i),n} = \int_{\mathcal{V}} \chi_n(\boldsymbol{x}) \beta^*_{l,q,i}(\boldsymbol{x}) d\boldsymbol{x} \tag{7.49}$$

$$= \int_{\mathcal{V}} \chi_n(\boldsymbol{x}) c_l(\boldsymbol{x}) e^{-j\boldsymbol{\kappa}_{q,i} \cdot \boldsymbol{\psi}(\boldsymbol{x})} d\boldsymbol{x} \tag{7.50}$$

The non-bijective nature of $\boldsymbol{\psi}$ is handled by splitting the imaging region into R bijective regions $\mathcal{V}_1, \ldots, \mathcal{V}_R$ such that $\mathcal{V} = \bigcup_{r=1}^{R} \mathcal{V}_r$. Let $\boldsymbol{\psi}^k(\boldsymbol{x}) = \boldsymbol{\psi}|_{\mathcal{V}_k}$, i.e. the function restricted to the subregion $\mathcal{V}_k$. A basis element, $\chi_{r,p}$, is now indexed as a pixel p within bijective region, r. This allows for a change of variable in the integration $\boldsymbol{u} = \boldsymbol{\psi}(\boldsymbol{x})$ and a new space, known as the encoding space, $\mathcal{W} = \psi(\mathcal{V}_k)$, is naturally defined. Specifically,

$$E_{(l,q,i),(r,p)} = \sum_{k=1}^{R} \int_{\mathcal{V}_k} \chi_{r,p}(\boldsymbol{x}) c_l(\boldsymbol{x}) e^{-j\boldsymbol{\kappa}_{q,i} \cdot \boldsymbol{\psi}(\boldsymbol{x})} d\boldsymbol{x} \tag{7.51}$$

$$= \sum_{k=1}^{R} \int_{\mathcal{W}} \tilde{\chi}^k_{r,p}(\boldsymbol{u}) \tilde{c}^k_l(\boldsymbol{u}) \tilde{d}^k(\boldsymbol{u}) e^{-j\boldsymbol{\kappa}_{q,i} \cdot \boldsymbol{u}} d\boldsymbol{u} \tag{7.52}$$

where the transformed functions are

$$\tilde{c}^k_l(\boldsymbol{u}) = c_l((\boldsymbol{\psi}^k)^{-1}(\boldsymbol{u})), \tag{7.53}$$

$$\tilde{\chi}^k_{r,p}(\boldsymbol{u}) = \chi_{r,p}((\boldsymbol{\psi}^k)^{-1}(\boldsymbol{u})), \tag{7.54}$$

and the volumetric correction factor is

$$\tilde{d}^k(\boldsymbol{u}) = |\det(D\boldsymbol{\psi})(\boldsymbol{z})|^{-1}_{\boldsymbol{z}=(\boldsymbol{\psi}^k)^{-1}(\boldsymbol{u})}. \tag{7.55}$$

Notice that the correction factor, $\tilde{d}^k$, is the inverse of the Jacobian determinant for the mapping to the curvilinear coordinate system.

With this change of variable, the basis $\chi_{r,p}(\boldsymbol{x}) = \delta(\boldsymbol{\psi}^r(\boldsymbol{x}) - \boldsymbol{u}_p)$ is adopted so

$$\tilde{\chi}^k_{r,p}(\boldsymbol{u}) = \delta(\boldsymbol{\psi}^r((\boldsymbol{\psi}^k)^{-1}(\boldsymbol{u})) - \boldsymbol{u}_p) \tag{7.56}$$

$$= \delta_{k,s}\delta(\boldsymbol{u} - \boldsymbol{u}_p) \tag{7.57}$$

which simplifies the analysis matrix to

$$E_{(l,q,i),(r,p)} = \tilde{c}^r_l(\boldsymbol{u}_p)\tilde{d}^r(\boldsymbol{u}_p)e^{-j\boldsymbol{\kappa}_{q,i}\cdot\boldsymbol{u}_p}. \tag{7.58}$$

Further structure can be imposed on the analysis operator by selecting the pixel grid, $\{\boldsymbol{u}_p\}$, so the encoding functions form a Discrete Fourier Transform. That is,

$$\boldsymbol{E} = \tilde{\boldsymbol{\mathcal{F}}}\tilde{\boldsymbol{C}}\tilde{\boldsymbol{D}}, \tag{7.59}$$

where $\tilde{\boldsymbol{\mathcal{F}}} = \mathrm{diag}(\boldsymbol{\mathcal{F}}, \ldots, \boldsymbol{\mathcal{F}})$ contains N_c separate DFT blocks, $\tilde{\boldsymbol{C}}$ contains the coil sensitivities and $\tilde{\boldsymbol{D}}$ is a diagonal matrix containing information on the field structure, $\tilde{d}^k(\boldsymbol{u}_p)$.

The relationship to SENSE imaging is established by transforming the unknown coefficients to "encoding space" using $\boldsymbol{m} = \tilde{\boldsymbol{D}}\boldsymbol{f}$. Denoting $\bar{\boldsymbol{E}} = \tilde{\boldsymbol{\mathcal{F}}}\tilde{\boldsymbol{C}}$, the measurement model is written as,

$$\boldsymbol{y} = \bar{\boldsymbol{E}}\boldsymbol{m} + \boldsymbol{v} \tag{7.60}$$

which is identical to the SENSE model in (7.28). Hence, the optimal reconstruction of $\boldsymbol{m}$ is obtained by applying the SENSE reconstruction matrix in (7.29) (with $\boldsymbol{E}$ replaced by $\bar{\boldsymbol{E}}$). Following this, the inverse transformation, $\boldsymbol{f} = \tilde{\boldsymbol{D}}^{-1}\boldsymbol{m}$, is applied to give the final coefficients. The covariance matrix for $\boldsymbol{f}$ inherits the block-diagonal structure from SENSE where the blocks are given by

$$\boldsymbol{X}^{(p)} = \frac{1}{N_t N_e}\boldsymbol{D}^{-1}(\boldsymbol{C}'\boldsymbol{\Sigma}^{-1}\boldsymbol{C})^{-1}\boldsymbol{D}^{-1}. \tag{7.61}$$

Considering the diagonal elements of a block, we obtain an SNR expression of

$$\frac{\mathsf{SNR}_{r,p}^{\text{PatLoc}}}{\mathsf{SNR}_{r,p}^{\text{linear}}} = \frac{\tilde{d}^r(\boldsymbol{u}_p)}{\sqrt{R}\, g_{r,p}} \tag{7.62}$$

where $g_{r,p}$ is the standard g-factor calculated for groups of ambiguous pixels.

It should also be noted that the 'volumetric correction factor' presented in this analysis is very similar to the 'density compensation factor' in the field of nonuniform k-space sampling e.g. [124]. Both cases are concerned with compensating for nonuniform samples, the former involves non-uniform samples in the image domain whereas the latter is concerned with nonuniform samples in the spatial frequency domain.

The analysis above emphasises the role of the weights, $\tilde{d}^r(\boldsymbol{u}_p)$, and corresponding matrix, $\boldsymbol{D}$, in calculating the reconstruction variance. The SENSE-like nature of PatLoc is revealed in the expressions for covariance and SNR where $\tilde{d}^r(\boldsymbol{u}_p)$ captures the non-linearity and appropriately modifies the existing SENSE equations.

7.3.3.2 Frame analysis of PatLoc

We derive the PatLoc SNR results using the frame matrix with emphasis on the selection of the discretisation basis. We will see that a normalised basis is more appropriate to reveal the inherent block structure of PatLoc. To achieve this we view the problem of image reconstruction from a signal processing perspective, where the role of the basis is explicit. Specifically, a basis $\{\chi_n\}$ is selected for discretisation and reconstruction; however, a different basis $\zeta_n(\boldsymbol{x}) = \text{rect}_{W_n}(\boldsymbol{x} - \boldsymbol{x}_n)$ is used to display the discretised function. Notice in the general case the pixel width, W_n, can vary between pixels.

Our aim is to display the function $\tilde{m}(\boldsymbol{x})$, a piecewise constant approximation to the true function, decomposed using $\{\zeta_n(\boldsymbol{x})\}$ as,

$$m(\boldsymbol{x}) \approx \tilde{m}(\boldsymbol{x}) = \sum_n m_n \operatorname*{rect}_{W_n}(\boldsymbol{x} - \boldsymbol{x}_n). \tag{7.63}$$

The reconstruction basis, $\chi_n(\boldsymbol{x})$, is used to decompose the function as

$$m(\boldsymbol{x}) \approx f(\boldsymbol{x}) = \sum_n f_n \chi_n(\boldsymbol{x}). \tag{7.64}$$

This basis reduces the measurement equation to a linear matrix equation in (7.18), which can be solved numerically for the basis coefficients, $\{f_1, \ldots, f_N\}$.

In general, there is a mismatch between the reconstructed coefficients, $\{f_n\}$,

and the desired pixel values, $\{m_n\}$. To compensate for this mismatch we consider the area of the n^{th} basis function, defined as

$$a_n := \int \chi_n(\boldsymbol{x}) d\boldsymbol{x}. \tag{7.65}$$

Since the coefficient, f_n, represents the weight of the n^{th} pixel basis, the quantity $f_n a_n$ is the weight of a pixel basis scaled to unit area. Finally, to convert from a weighted pixel with unit area, to the function height, m_n, we divide by the pixel width. Thus,

$$m_n = \frac{a_n}{W_n} f_n. \tag{7.66}$$

We now apply this result to PatLoc reconstruction. Contrary to the existing analysis, we use a normalised basis, which is standard convention in signal processing literature. Normalised in this context means the integral of a basis element is unity. The chosen basis naturally separates the SENSE-like structure from the nonlinear effects.

As described above we can partition the object space into its bijective regions and define a reconstruction basis indexed by region, r, and pixel position, p. A normalised, non-uniform basis that elicits a desirable structure in the frame matrix is

$$\xi_{r,p}(\boldsymbol{x}) = \frac{1}{\tilde{d}^r(\boldsymbol{u}_p)} \delta(\boldsymbol{\psi}^r(\boldsymbol{x}) - \boldsymbol{u}_p). \tag{7.67}$$

In this case, the area of the basis functions is unity i.e. $a_{r,p} = 1$. We apply this basis to the calculation of the frame matrix, $\boldsymbol{S}$, and note that the covariance matrix for parallel imaging has the structure, $\Sigma^{-1}_{(l,q,i),(l',q',i')} = \delta_{q,q'}\delta_{i,i'}\Sigma^{-1}_{l,l'}$, as described earlier. The frame matrix in (7.26) simplifies to

$$S_{(r,p),(r',p')} = \tilde{\varkappa}^{(r,r')}(\boldsymbol{u}_{p'}, \boldsymbol{u}_p) h(\boldsymbol{u}_{p'}, \boldsymbol{u}_p). \tag{7.68}$$

The coil kernel is

$$\tilde{\varkappa}^{(r,r')}(\boldsymbol{u}_{p'}, \boldsymbol{u}_p) = \sum_{l,l'} \tilde{c}^{s'}_l(\boldsymbol{u}_{p'}) \Sigma^{-1}_{l,l'} \tilde{c}^{r}_{l'}(\boldsymbol{u}_p) \tag{7.69}$$

and

$$h(\boldsymbol{u}, \boldsymbol{u}') = D_{N_t}(\Delta_x(\boldsymbol{u} - \boldsymbol{u}') \cdot \hat{\boldsymbol{\imath}}) D_{N_e}(\Delta_y(\boldsymbol{u} - \boldsymbol{u}') \cdot \hat{\boldsymbol{\jmath}}), \tag{7.70}$$

where $D_N(\boldsymbol{x})$ is the Dirichlet kernel defined in (7.39).

The chosen pixel basis suggests a natural ordering of the pixels so $\boldsymbol{S}$ is block-diagonal. The matrix is indexed by p, the block of aliased pixels and r, the pixel index within a block. The samples are acquired according to the pixel grid, $\boldsymbol{u}_p$, so the Fourier encoding is exact in the encoding space and

$$h(\boldsymbol{u}_{p'}, \boldsymbol{u}_p) = N_t N_e \delta_{p,p'}. \tag{7.71}$$

Thus a block of $\boldsymbol{S}$ is defined when $p = p'$,

$$S^{(p)}_{(r,r')} := S_{(r,p),(r',p)} = N_t N_e \tilde{\varkappa}^{(r,r')}(\boldsymbol{u}_p, \boldsymbol{u}_p). \tag{7.72}$$

Let $\bar{\boldsymbol{x}}_{r,p} = (\boldsymbol{\psi}^r)^{-1}(\boldsymbol{u}_p)$ denote the non-uniform pixel positions. The elements of the p^{th} block are

$$S^{(p)}_{(r,r')} = N_t N_e \sum_{l,l'} c^*_{l'}(\bar{\boldsymbol{x}}_{r,p}) \Sigma^{-1}_{l,l'} c_l(\bar{\boldsymbol{x}}_{r',p}) \tag{7.73}$$

In matrix notation, $\boldsymbol{S}^{(p)} = (N_t N_e) \boldsymbol{C}' \boldsymbol{\Sigma}^{-1} \boldsymbol{C}$. This frame matrix is identical to the SENSE equivalent in (7.46).

The theory derived earlier shows that noise performance can be directly calculated from the frame matrix,

$$\boldsymbol{X}^{(p)} = {\boldsymbol{S}^{(p)}}^{-1} \tag{7.74}$$

$$= \frac{1}{N_t N_e} (\boldsymbol{C}' \boldsymbol{\Sigma}^{-1} \boldsymbol{C})^{-1} \tag{7.75}$$

This covariance is identical to SENSE and contains no nonlinear adjustment, as expected since PatLoc is a generalisation of SENSE.

The nonlinear effects enter into the system by relating the basis coefficients, $f_{r,p}$, with the pixel coefficients, $m_{r,p}$. Recall the pixel area is $a_{r,p} = 1$ and the pixel width is $W_{r,p} = \tilde{d}^r(\boldsymbol{u}_p) W^u$ where W^u is the pixel width in encoding space. Without loss of generality, we set $W^u = 1$ and by virtue of (7.66),

$$m_{r,p} = \frac{f_{r,p}}{\tilde{d}^r(\boldsymbol{u}_p)}. \tag{7.76}$$

As a result, the covariance of the pixel values, $\tilde{X}^{(p)}$, is related to the covariance of the basis coefficients, $X^{(p)}$, by

$$\tilde{\boldsymbol{X}}^{(p)} = \boldsymbol{D}^{-1} \boldsymbol{X}^{(p)} \boldsymbol{D}^{-1} \tag{7.77}$$

$$= \frac{1}{N_t N_e} \boldsymbol{D}^{-1} (\boldsymbol{C}' \boldsymbol{\Sigma}^{-1} \boldsymbol{C})^{-1} \boldsymbol{D}^{-1} \tag{7.78}$$

which is identical to (7.61) and the resulting SNR ratio is identical to (7.62).

Here we gain a different perspective on the PatLoc reconstruction problem. On one hand, we can adopt the unnormalised basis, which results in a variance expression that directly includes the nonlinearity. Alternatively, we can use a normalised basis that provides the desired matrix structure with the nonlinearity entering the equations in an intuitive way: a transformation between the reconstruction basis and the display basis. The transformation given in (7.76) also reveals the trade-off between resolution and SNR. Specifically, the variance is scaled by the inverse square of the pixel size (given by $\tilde{d}^r$).

It should also be noted that (7.61) or (7.78) is the variance of the nonuniform pixel values and usually these are interpolated to a regular grid prior to display. This interpolation modifies the variance of the final pixels and a rigourous noise analysis should include the effect of this interpolation. Although this discrepancy can be minimised by careful selection of the interpolation grid, this point has not been highlighted in the PatLoc literature.

7.4 Analysis of arbitrary encoding schemes

As demonstrated above, there is an intrinsic relationship between resolution and SNR. This relationship becomes increasingly difficult to characterise for arbitrary nonlinear encoding schemes. Although a detailed investigation of the trade-off between SNR and resolution is beyond the scope of this thesis, it is worthwhile to note that the frame matrix can also be used to estimate the local resolution. In Chapter 6 we approximated the point spread function by extracting rows of $\boldsymbol{E}'\boldsymbol{E}$, equivalent to the frame matrix when the noise covariance is the identity matrix. We adopted a similar approach in [125] to define simple metric related to the width of the point spread function. In this chapter, we restrict our attention to calculating the variance from the frame matrix.

The analyses in the previous sections exploited inherent structure in the encoding schemes to simplify the calculation of SNR, which depends on the pixel covariance. This may not be possible for arbitrary encoding schemes so we seek an alternative metric. Consequently we require a performance metric that: 1) reflects the variance of the reconstructed pixels, 2) is valid for arbitrary encoding fields, and 3) is computationally efficient in terms of time and memory.

7.4.1 Approximate reconstruction variance

The pixel covariance is exactly specified by the inverse of the frame matrix, as stated in (7.24). The pixel variance is obtained by extracting the diagonal elements of the covariance matrix. This is not a feasible performance metric since the inverse is difficult to compute in the general case. In this section, we propose a performance metric using a series expansion of the matrix inverse. The resulting metric is a second-order approximation to the variance, which is simple to compute.

To this end, we exploit the fact that, for most practical imaging schemes, the frame matrix is approximately diagonal. For example, in standard Fourier imaging the matrix is exactly diagonal and the only distortion is due to the truncation effects associated with the projection operator. The approximate diagonal nature of the frame matrix, $\boldsymbol{S}$, suggests a decomposition of the matrix into its diagonal part, $\boldsymbol{\Lambda}$, and off-diagonal part, $\boldsymbol{Z}$, such that $\boldsymbol{Z} = \boldsymbol{S} - \boldsymbol{\Lambda}$. We take the Taylor series of $\boldsymbol{S}^{-1}$ about the matrix $\boldsymbol{S} = \boldsymbol{\Lambda}$ to obtain the expansion

$$\boldsymbol{S}^{-1} = \sum_{n=0}^{\infty} (-\boldsymbol{\Lambda}^{-1}\boldsymbol{Z})^n \boldsymbol{\Lambda}^{-1}. \tag{7.79}$$

If $\|\boldsymbol{\Lambda}^{-1}\boldsymbol{Z}\| < 1$ then the series will converge, where $\|\cdot\|$ denotes some matrix norm, e.g. the induced L_2 norm. The zeroth, first, and second order approximations are

$$\boldsymbol{S}_0^{-1} = \boldsymbol{\Lambda}^{-1} \tag{7.80}$$

$$\boldsymbol{S}_1^{-1} = \boldsymbol{\Lambda}^{-1} - \boldsymbol{\Lambda}^{-1}\boldsymbol{Z}\boldsymbol{\Lambda}^{-1} \tag{7.81}$$

$$\boldsymbol{S}_2^{-1} = \boldsymbol{\Lambda}^{-1} - \boldsymbol{\Lambda}^{-1}\boldsymbol{Z}\boldsymbol{\Lambda}^{-1} + \boldsymbol{\Lambda}^{-1}\boldsymbol{Z}\boldsymbol{\Lambda}^{-1}\boldsymbol{Z}\boldsymbol{\Lambda}^{-1} \tag{7.82}$$

By definition, $\boldsymbol{\Lambda}$ is diagonal so $\boldsymbol{\Lambda}^{-1}$ is easily computed by inverting the non-zero diagonal elements. This leads to approximations that can be efficiently computed. If we consider the diagonal elements of $\boldsymbol{S}^{-1}$ then we notice that

$$\operatorname{diag}(\boldsymbol{\Lambda}^{-1}\boldsymbol{Z}\boldsymbol{\Lambda}^{-1}) = \boldsymbol{0} \tag{7.83}$$

so we need to use the second order approximation in (7.82) for useful results.

The computation is further reduced since the diagonal components can be calculated individually as follows. Let $z_{n,m}$ denote the elements of $\boldsymbol{Z}$ and $\boldsymbol{z}_m$ denote the m^{th} column of $\boldsymbol{Z}$. The diagonal elements of $\boldsymbol{\Lambda}^{-1}$ are $\lambda_n, n = 1, \ldots, N$, which only depend on the coil sensitivities and noise covariance. The final performance

metric, $\hat{X}_n$, is obtained by extracting the diagonal elements of $\boldsymbol{S}_2^{-1}$,

$$\hat{X}_n := [\boldsymbol{S}_2^{-1}]_{nn} \tag{7.84}$$

$$= \lambda_n + \lambda_n^2 \boldsymbol{z}_n' \boldsymbol{\Lambda}^{-1} \boldsymbol{z}_n \tag{7.85}$$

$$= \lambda_n \left(1 + \lambda_n \sum_{m=1}^{N} z_{n,m}^2 \lambda_m \right) \tag{7.86}$$

which can be calculated row by row, in $\mathcal{O}(N)$, without the need to calculate or store the large matrix, $\boldsymbol{S}$. Since there are N rows to calculate the overall cost is $\mathcal{O}(N^2)$.

When $\|\boldsymbol{\Lambda}^{-1}\boldsymbol{Z}\| > 1$ the series will not converge so the approximation is no longer suitable. This can happen for schemes with a wide point spread function or very few echoes, causing a reduction in the diagonal dominance of the frame matrix. In this case, we can take the Taylor series of $\boldsymbol{S}^{-1}$ about a different matrix, $\boldsymbol{\Omega}$, such that the condition for convergence is satisfied, i.e., $\|\boldsymbol{\Omega}^{-1}\boldsymbol{Y}\| < 1$ with $\boldsymbol{Y} = \boldsymbol{S} - \boldsymbol{\Omega}$. To retain the computational efficiency of the expansion, we must choose an $\boldsymbol{\Omega}$ that is easily inverted. For example, a tridiagonal, block-diagonal or general sparse matrix has this property. This generalisation will be considered in future work.

The metric defined by (7.86) satisfies all of our requirements; it is relatively easy to compute, applicable to general encoding schemes and reflects the variance of individual pixels.

We now apply our frame analysis to O-space [11], a recent imaging technique whose performance has not been rigourously analysed like SENSE or PatLoc. In Section 7.5 we apply the approximate variance metric to an O-space imaging scheme to produce a high resolution variance map. In this case, direct inversion of the frame matrix is impractical and no alternative analysis exists.

7.4.2 O-Space imaging

O-Space imaging was described in Section 6.3.2. The technique uses a nonlinear encoding field to achieve rapid imaging times [11]. The encoding field for this scheme has a quadratic shape with its centre translated after every echo. The encoding field for O-space is defined in (6.18) as

$$b_q(\boldsymbol{x}, t) = b_q(\boldsymbol{x}) = G\|\boldsymbol{x} - \boldsymbol{r}_q\|^2 \tag{7.87}$$

where G is the field strength and $\boldsymbol{r}_q$ is an echo-dependent centre point. The encoding field is constant during readout thus the phase functions are $\phi_{q,i}(\boldsymbol{x}) =$

$-\gamma G\|\boldsymbol{x}-\boldsymbol{r}_q\|^2 t_i$. The corresponding encoding functions are

$$\beta_{l,q,i}(\boldsymbol{x}) = c_l(\boldsymbol{x}) e^{j\gamma G\|\boldsymbol{x}-\boldsymbol{r}_q\|^2 t_i}. \tag{7.88}$$

This is fundamentally different from PatLoc since a Jacobian (such as that in (7.55)) cannot be defined for O-space imaging, since it is not a mapping between two (Cartesian) coordinates and two (curvilinear) coordinates. This precludes an analysis similar to that of the PatLoc paper [10].

The global performance of O-space was previously examined for a numerical phantom to determine the optimal placement of the centre points [126]. In that work, full reconstruction was performed and the reconstruction error was calculated. The analysis did not, however, capture the local properties of the reconstruction error. Furthermore, it is only valid for the single phantom considered. The frame-oriented approach developed in this chapter overcomes these limitations.

7.4.2.1 Frame analysis of O-Space

We use the frame matrix to analyse the O-space imaging technique. The frame matrix elements are given by (7.26) with a measurement covariance of the form,

$$\Sigma^{-1}_{(l,q,i),(l',q',i')} = \delta_{q,q'}\delta_{i,i'}\Sigma^{-1}_{l,l'} \tag{7.89}$$

Furthermore, the result of Appendix 7.A justifies the use of the ideal delta functions, $\chi_i(\boldsymbol{x}) = \delta(\boldsymbol{x}-\boldsymbol{x}_i)$, thus the elements of the frame matrix are

$$S_{n,m} = \sum_{l,l'} c^*_{l'}(\boldsymbol{x}_n)\Sigma^{-1}_{l,l'} c_l(\boldsymbol{x}_m) \sum_{q,i} e^{j\gamma(b_q(\boldsymbol{x}_n)-b_q(\boldsymbol{x}_m))t_i} \tag{7.90}$$

$$= \varkappa(\boldsymbol{x}_n,\boldsymbol{x}_m) \sum_q D_{N_t}(\gamma\tau_s(b_q(\boldsymbol{x}_n)-b_q(\boldsymbol{x}_m))) \tag{7.91}$$

where $\varkappa(\boldsymbol{x}_n,\boldsymbol{x}_m)$ is the coil kernel defined in (7.42) and b_q is the O-space encoding field defined in (7.87). We assume regular sampling during readout with sampling period, τ_s, such that $t_i = i\tau_s$. This allows us to remove the temporal summation in (7.90), resulting in the simplified expression (7.91). This simplification can be applied whenever the encoding field is constant during readout and the time samples are regularly spaced. It could therefore be used to simplify the frame matrix for other nonlinear schemes such as 4D-RIO [12] or Null Space Imaging [91].

For O-space, the points where the encoding Dirichlet kernel is maximum are

given by $b_q(\boldsymbol{x}_n) - b_q(\boldsymbol{x}_m) = 0$ which implies that

$$\|\boldsymbol{x}_n - \boldsymbol{r}_m\|^2 = \|\boldsymbol{x}_p - \boldsymbol{r}_q\|^2. \tag{7.92}$$

This is the equation for a circle, centred at $\boldsymbol{r}_q$ with radius $\|\boldsymbol{x}_n - \boldsymbol{r}_m\|$. Thus the n^{th} row of the frame matrix, $\boldsymbol{S}$, is approximately given by the sum of all circles passing through the point $\boldsymbol{x}_n$, with centres $\boldsymbol{r}_q$, weighted by the coil kernel, $\varkappa(\boldsymbol{x}_n, \boldsymbol{x}_m)$. More precisely, the circles will have ringing defined by the Dirichlet kernel in (7.91), which depends on the readout sampling parameters, τ_s and N_t. Furthermore, for non-symmetric sampling the circles will be complex-valued. This complicated geometry means the structure of the frame matrix $\boldsymbol{S}$ is difficult to exploit when calculating the reconstruction covariance, leading to an impractical matrix inversion.

An interesting alternative is to exploit to the fact that the frame matrix contains many small elements, particularly in the sidelobes of the circles defined above. Thus a sparse approximation can be obtained by thresholding the frame matrix, which could then be inverted using numerically efficient algorithms that take advantage of the sparsity. This is equivalent to [127], where the encoding matrix was replaced by a sparse approximation, by transforming it using a Discrete Fourier Transform and then applying a threshold.

Since our goal is to quantify the reconstruction performance as accurately as possible, we use the series expansion in (7.86) of the frame matrix given by (7.91) to approximate the pixel variance. The results of Section 7.5 demonstrate the efficacy of this approximation through simulation examples.

7.5 Simulations

The power of the proposed performance metric is demonstrated using O-space imaging as a case study. As detailed above, O-space exhibits a structure that is difficult to handle unlike the schemes of SENSE and PatLoc. In the latter cases, precise and computationally efficient expressions for the variance of the reconstructed pixels can be derived as explained in Section 7.3. No such expressions can be derived for O-space imaging and the advantage of our performance metric becomes apparent. In particular, the metric allows for the first quantitative comparison of the local noise performance of O-space and PatLoc.

7.5.1 Methods

Three simulations were conducted to demonstrate the utility of the metric developed in Section 7.4. Firstly, the accuracy of the metric was verified for a low-

resolution O-space imaging regime. The setup was comparable to the original implementation of O-space, which utilises the shim coil to provide a quadratic encoding field [11]. As such, the field gradient was limited to $600\,\mathrm{Hz/cm^2}$, which is the maximum field produced by the Z2 shim coil of a 4.7 T Bruker small bore scanner. The field-of-view was set to 4 cm. Eight coils were placed concentrically around the field-of-view with sensitivities simulated by the magnetostatic limit of the Biot-Savart equation. The acquisition consisted of 32 echoes, each generated with a unique encoding field. The encoding fields for O-space are defined by their centre points, which were arranged in concentric rings with radii 0.8 cm and 1.6 cm as well as a single centre point at the origin. The arrangement is similar to [11]. The sampling parameters were $T_{\mathrm{acq}} = 50\,\mathrm{ms}$, $N_t = 64$. The final image size was 32×32. The low resolution image means the true variance of the individual pixels can be calculated by matrix inversion of the frame matrix as in (7.24). This was compared to the metric in (7.86) and the percentage error was calculated to evaluate the accuracy of the second order approximation. The metric was further validated by calculating the empirical variance of each reconstructed pixel over 500 simulated acquisitions. Each acquisition was simulated with O-space encoding of a uniform disk phantom with different realisations of Gaussian noise added.

The second simulation compared the local performance of O-space and PatLoc imaging schemes using the theory developed in Sections 7.3 and 7.4. The final image for both schemes was 256×256. Two PatLoc imaging schemes were examined that consisted of orthogonal quadrupolar encoding fields with undersampling by factors of 2 and 4 in the phase-encoding direction. This produces net acceleration factors of 4 and 8 compared to traditional Fourier imaging. The sampling parameters were chosen according to Nyquist criteria for combined PatLoc-SENSE acquisition [10]. The non-rectangular field-of-view was chosen to entirely encompass the object. Two time-equivalent O-space schemes were defined consisting of 64 and 32 echoes representing "acceleration factors" of 4 and 8, respectively. Technically speaking, the acceleration factor for an O-space scheme is difficult to define since any loss of resolution is not included. Nonetheless, we use the term to define the number of echoes. The number of time samples was $N_t = 512$. The coil locations and all other parameters were identical to the low-resolution protocol above. The variance of the reconstructed pixels for the PatLoc schemes were calculated exactly by inverting small matrices as in (7.61). The O-space schemes do not possess an exploitable structure so matrix inversion is impractical; instead we use the metric in (7.86) to approximate the variance.

Thirdly, nonlinear encoding was simulated using the PatLoc and O-space schemes with acquisition parameters defined above. These nonlinear encoding schemes

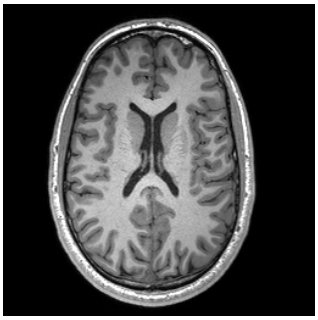


Figure 7.2: The numerical phantom used for simulations, taken from a T_1 weighted acquisition of a healthy brain.

were simulated using the numerical phantom in Fig. 7.2, which was obtained by imaging a healthy cortex using a standard sequence with linear gradients. In these simulations, the image intensity of the phantom is used to represent the object's magnetisation. Reconstructed images for O-space and PatLoc were generated from this simulated data to provide concrete examples of our performance analysis. The PatLoc image was reconstructed according to the Cartesian PatLoc reconstruction algorithm described in [10]. The O-space image was reconstructed using the Algebraic Reconstruction Technique (or Kaczmarz algorithm [45]) as in the original O-space implementation [11]. The images were reconstructed using data simulated from the brain phantom.

7.5.2 Results

Figure 7.3 displays performance maps of the low-resolution O-space imaging regime. The true variance is displayed in Fig. 7.3a and the second order approximation metric in Fig. 7.3b. The absolute difference for each pixel between the two metrics was calculated and the mean of this difference was $\sim$2%, which demonstrates the close match between the proposed metric and the true variance. The empirical variance of 500 reconstructions with independent noise realisations is presented in Fig. 7.3c, which shows a variance map consistent with the theoretical analysis. Figure 7.4 plots line profiles through the centre of the performance maps for the low-resolution O-Space scheme in Fig. 7.3. Good quantitate agreement between the variance measures is demonstrated, which reflects the usefulness of the proposed approximate variance metric.

Fig. 7.5 displays the performance of PatLoc and O-space imaging schemes for different acceleration factors. The PatLoc performance for four-fold and eight-fold acceleration are shown in Figures 7.5a and 7.5b, respectively. The O-space performance for equivalent acceleration factors is illustrated in Fig. 7.5c and 7.5d. The

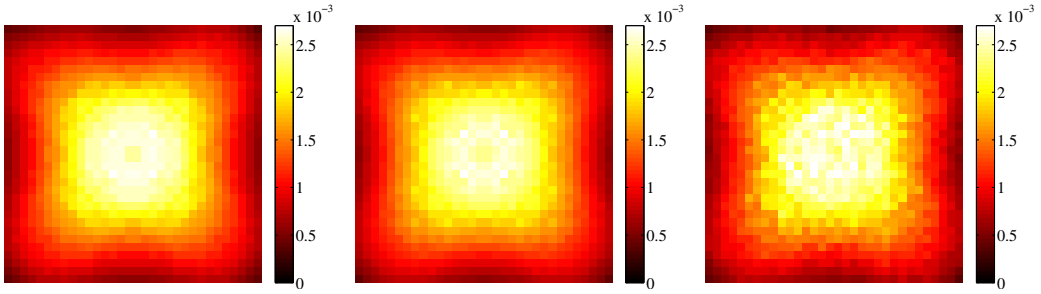


Figure 7.3: The (a) actual, (b) approximate and (c) empirical variance of a low-resolution O-space imaging scheme.

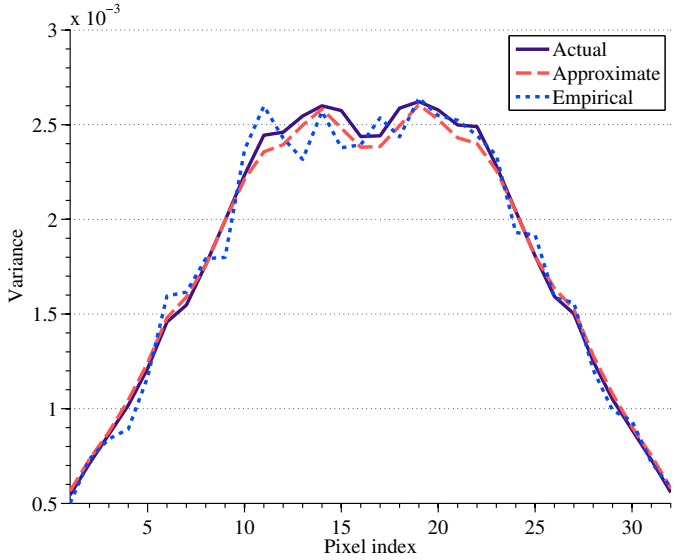


Figure 7.4: Profiles along vertical lines through the centre of the performance maps in Fig. 7.3. The actual (solid line), approximate (dashed line) and empirical (dotted line) variance demonstrate good agreement.

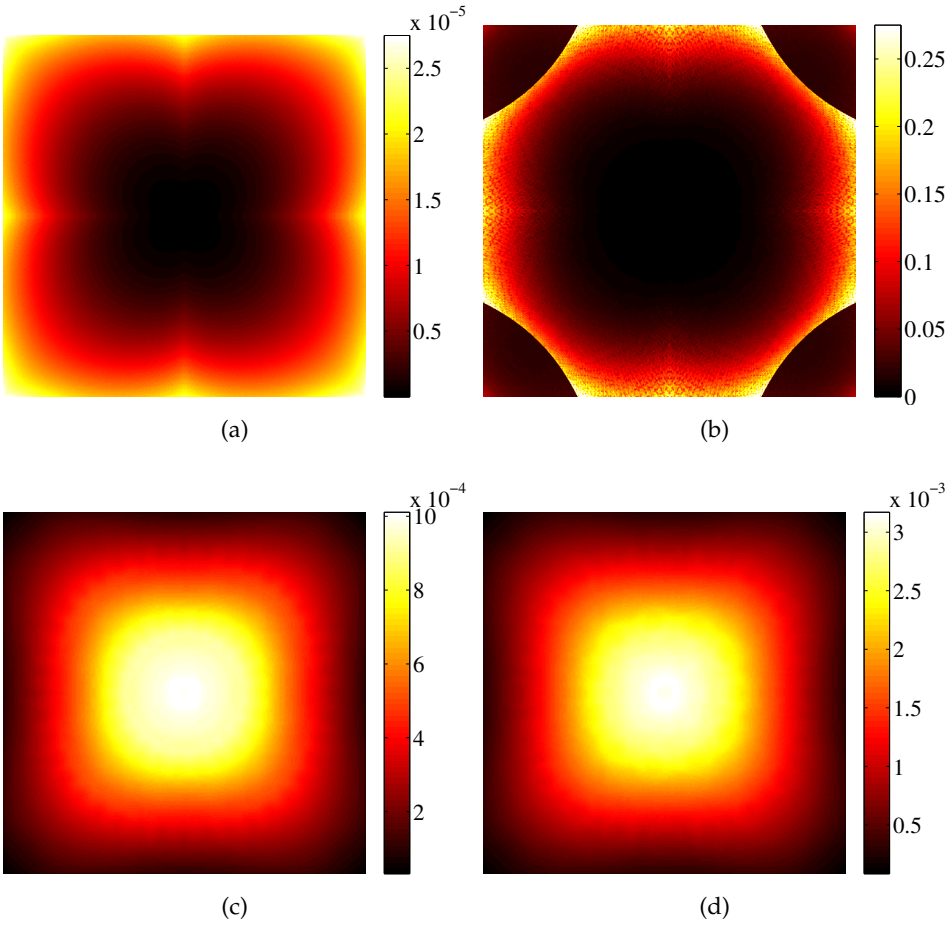


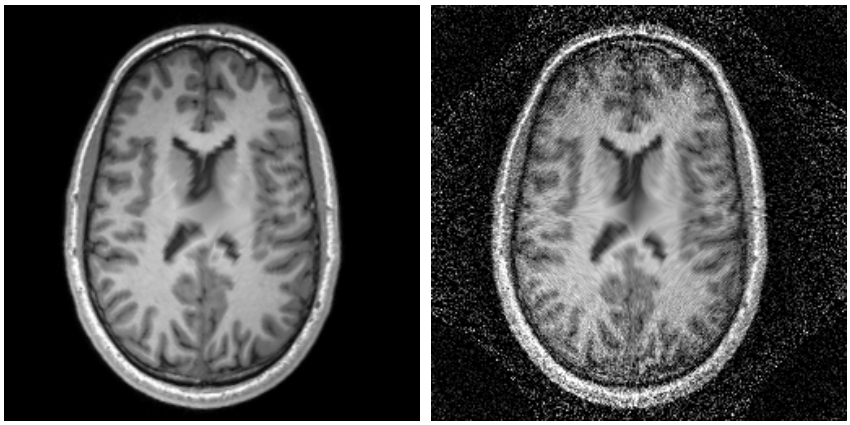
Figure 7.5: A high resolution variance map for PatLoc and O-space imaging schemes. The PatLoc maps in (a) and (b) were calculated from the variance obtained in Eq. (7.61) for accelerations factors of 4 and 8, respectively. The O-space maps in (c) and (d) we calculated using our performance metric in Eq. (7.86) for equivalent acceleration factors. Notice the different scales in each sub-figure.

metric displayed is an approximation to the pixel variance as defined in (7.86). Notice the different scales of each performance map. PatLoc imaging for modest acceleration exhibits low reconstruction variance in the central region of the image. This corresponds to the region where the resolution is poor indicated by a large effective pixel size. The usual trade-off exists between resolution and SNR. An acceleration factor of 8 leads to a large increase of the pixel variance, indicating the transition from an overdetermined system to a square system of equations. In this case the inversion in (7.75) becomes ill-conditioned, resulting in pixels with large variance.The O-space scheme produces images with increased variance towards the centre of the image, reflecting the loss of encoding provided by the coil profiles. The performance of O-space imaging exhibits a gradual decline for increasing acceleration factors. Although this was demonstrated though simulated reconstructions in [11], our metric quantifies the noise performance, based on a theoretical analysis. Importantly, the metric is valid irrespective of the object being imaged.

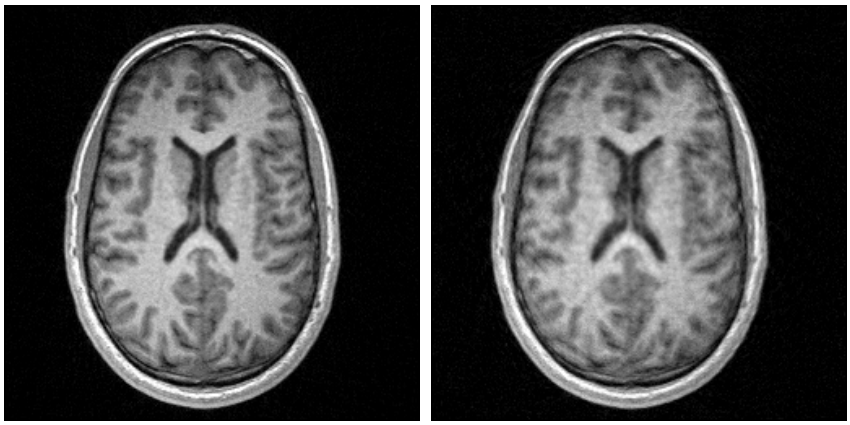
Fig. 7.6 presents the reconstruction results from PatLoc and O-space imaging regimes with acceleration factors of $R = 4$ and $R = 8$. The PatLoc images in Fig. 7.6a and 7.6b exhibit an obvious loss of resolution towards the centre. The large effective pixel size in the centre blurs out both the object features and the noise. Consequently, these regions have low variance as predicted by the metric maps in Fig. 7.5. On the other hand, the O-space images in Fig. 7.6c and 7.6d have noise spread across the image, with more uniform resolution. It should be noted, however, that the resolution of O-space at the periphery is lower than PatLoc, and it degrades significantly with increasing acceleration factor. This degradation of resolution in O-space imaging is traded for only a modest increase in reconstruction variance from $R = 4$ to $R = 8$. The PatLoc examples maintain a fixed (but spatially-varying) resolution for increasing acceleration, which results in a large increase in pixel variance at $R = 8$. These results highlight that the developed metric is useful to predict the noise performance of nonlinear imaging schemes. However, noise performance alone is not sufficient to completely characterise the imaging performance and the spatially-varying resolution must also be considered.

7.6 Discussion

In addition to deriving a suitable performance metric, the frame oriented approach provides insight into a particular encoding scheme. We have demonstrated that the frame operator for PatLoc elucidates the fundamental source of aliasing, de-



(a) (b)



(c) (d)

Figure 7.6: Images reconstructed from simulated data with nonlinear encoding. PatLoc images (a) and (b) are for acceleration factors of 4 and 8, respectively. O-Space reconstructions (c) and (d) are for equivalent accelerations of 4 and 8, respectively.

fined by the Dirichlet kernel. PatLoc imposes a structure on the frame operator by the careful selection of a nonuniformly spaced pixel basis. The structure allows for an exact performance analysis and a simple reconstruction algorithm, which are significant advantages of this scheme. The difficulty lies in displaying the nonuniform pixels and often the pixel values are interpolated to a regular grid. This interpolation reflects a loss in the fundamental resolution of the final pixels displayed. Regions with low variance may have large bias due to interpolation, which results in significant image blurring. On the other hand, O-space exhibits a relatively uniform variance map, but suffers from a loss of resolution for increasing acceleration. Thus using variance as the sole metric will fail to capture the true imaging performance. In other words, the variance is necessary but not sufficient to quantify the performance of nonlinear encoding schemes. A complete characterisation of the performance of nonlinear encoding schemes must jointly examine the SNR and the spatially-dependent resolution. It is foreseeable that a hybrid of variance and resolution metrics may prove the most useful for assessment. This is the direction of our current work.

The metric presented in this chapter does not capture the tradeoff between variance and resolution, and will be of limited use for encoding schemes with large variation in resolution. These schemes require an appropriate pixel basis to improve the conditioning of the reconstruction problem. This basis is naturally defined for Cartesian PatLoc and is based on the Jacobian of the nonlinear mapping associated with the encoding fields. For arbitrary schemes such as those in Chapter 6, no such Jacobian can be defined. In this case, a direct reconstruction of a uniform pixel basis will amplify noise in regions with minimal encoding, since the inherent pixel resolution is not compensated for. In reality, the reconstruction is regularised and the proposed metric no longer valid. However, for schemes such as O-space, which have a relatively constant resolution, the metric is an effective measure of the reconstruction variance.

The development of nonlinear imaging schemes has revealed the need to consider both the encoding fields and the coil sensitivities jointly in designing new state-of-the-art imaging techniques. In additional to this encoding, spatial encoding can be provided by tailored RF excitation profiles which excite a different subset of spins every echo [122]. In principle, the framework presented in this chapter is readily applicable to the performance analysis of schemes that utilise a combination of nonlinear fields, multiple receive coils and tailored RF excitation.

7.7 Conclusion

In this chapter, we have developed a framework to analyse the noise performance of imaging schemes with general nonlinear encoding fields. Each scheme is uniquely defined by the spatial encoding functions which have a contribution from the accumulated phase function and the coil sensitivity. The encoding functions for a given scheme define an associated frame operator, from which performance metrics are derived. Highly structured schemes such as SENSE and PatLoc exhibit a structure in the frame operator that can be exploited to provide the closed form performance metrics reported in the literature. For arbitrary schemes such as O-space, the frame operator can be used to define a tractable and intuitive noise metric based on the covariance of the reconstructed pixels. We demonstrated the utility of this computationally efficient performance metric through simulation examples.

The continued advancement of MRI relies on the ability to quantitatively assess new techniques as they are created. The development of suitable performance metrics, such as those presented in this chapter, will help guide designers to find optimised imaging schemes that utilise all possible degrees of freedom.

Appendices

7.A Basis selection

The finite dimensional representation of the frame operator, described in Section 7.2, was developed using a pixel basis. We now turn our attention to the precise definition of this basis. The basis should be capable of accurately representing the true function while allowing efficient computation of the encoding matrix in (7.16). Two common bases are the rectangular basis $\zeta_p(\boldsymbol{x}) = \frac{1}{W}\,\mathrm{rect}_W(\boldsymbol{x} - \boldsymbol{x}_p)$ and the Dirac delta basis $\chi_p(\boldsymbol{x}) = \delta(\boldsymbol{x} - \boldsymbol{x}_p)$. The rectangular basis is more realistic while the delta basis is better for computation. In this section we calculate the difference between these two bases for simple linear encoding. We quantify the difference in the reconstructed pixels between the delta basis and rectangular basis.

This investigation uses a one dimensional object with linear encoding to examine the differences between the bases although the analysis can be extended to higher dimensions and more complicated encoding schemes. The phase function we use for demonstration is

$$\phi_i(x) = k_i x \tag{7.93}$$

where k_i is selected so the encoding functions, $\{\Phi_{q,i}\}$, are the Fourier basis. Additionally, an ideal volume coil is assumed such that $c_l(\boldsymbol{x}) = 1$. The measurements are given by

$$y_i = \int m(x) e^{-jk_i x} dx \tag{7.94}$$

This equation is discretised with both χ_p and ζ_p. This is equivalent to making one of the following approximations,

$$m(x) \approx f(x) = \sum_j f_j \chi_j(x) = \sum_j f_j \delta(x - x_j) \tag{7.95}$$

$$m(x) \approx g(x) = \sum_j g_j \zeta_j(x) = \sum_j g_j \tfrac{1}{W} \operatorname*{rect}_W(x - x_j) \tag{7.96}$$

We denote the vector containing the coefficients of the delta basis approximation as $\boldsymbol{f} = [f_1, \ldots, f_N]'$. Likewise, the coefficient vector for the rectangular basis is $\boldsymbol{g} = [g_1, \ldots, g_N]'$ and the measurement vector is $\boldsymbol{y} = [y_1, \ldots, y_{N_t}]'$. Applying the approximations above the measurement equation can be written as two different

matrix equations, depending on our choice of basis.

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{E}\boldsymbol{f} \tag{7.97}$$

$$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{E}\boldsymbol{g} \tag{7.98}$$

where $E_{i,n} = e^{-jk_i x_n}$ is the Fourier matrix and the elements of the diagonal weighting matrices, $\boldsymbol{H}$ and $\boldsymbol{G}$, are calculated by

$$\alpha_n = \int \chi(x) e^{-jk_n x} dx \tag{7.99}$$

which is the Fourier transform of the basis shape. This results in

$$\boldsymbol{H} = \text{diag}(t_1, \ldots, t_{N_t}); \qquad t_i = 1,\ i = 1, \ldots, N_t \tag{7.100}$$

$$\boldsymbol{G} = \text{diag}(\tau_1, \ldots, \tau_{N_t}); \qquad \tau_i = \text{sinc}(Wk_i),\ i = 1, \ldots, N_t \tag{7.101}$$

We assume that both systems of equations can be inverted and note that $\boldsymbol{E}^{-1} = \frac{1}{n}\boldsymbol{E}'$. Given the measurement vector, $\boldsymbol{y}$, the reconstructed coefficients for both systems are

$$\boldsymbol{f} = \boldsymbol{E}^{-1}\boldsymbol{H}^{-1}\boldsymbol{y} \tag{7.102}$$

$$\boldsymbol{g} = \boldsymbol{E}^{-1}\boldsymbol{G}^{-1}\boldsymbol{y} \tag{7.103}$$

These coefficients are displayed as image pixels in the reconstructed image; thus we are interested in the difference between both sets of coefficients. We calculate $\|\boldsymbol{f} - \boldsymbol{g}\|$ and use the fact that, $\|\boldsymbol{E}^{-1}\boldsymbol{y}\| = \frac{1}{\sqrt{n}}\|\boldsymbol{y}\|$,

$$\boldsymbol{f} - \boldsymbol{g} = \boldsymbol{E}^{-1}(\boldsymbol{H}^{-1} - \boldsymbol{G}^{-1})\boldsymbol{y} \tag{7.104}$$

$$\|\boldsymbol{f} - \boldsymbol{g}\| = \frac{1}{\sqrt{n}}\|(\boldsymbol{H}^{-1} - \boldsymbol{G}^{-1})\boldsymbol{y}\| \tag{7.105}$$

We consider $\boldsymbol{y}$ with elements defined by the continuous integral in (7.94) and assume that $k_i = (i - N_t/2)\Delta_k$ where the frequency spacing, Δ_k, is fixed (determined from the reciprocal of the FOV). Furthermore, we assume that the Fourier transform of $m(x)$ has support in the interval $[-\Omega/2, \Omega/2]$. Denote $K = \frac{\Omega}{\Delta_k}$ as the number of non-zero measurements. Thus,

$$\|(\boldsymbol{H}^{-1} - \boldsymbol{G}^{-1})\boldsymbol{y}\|_2^2 = \|\mathbf{Y}_K(\boldsymbol{H}^{-1} - \boldsymbol{G}^{-1})\boldsymbol{y}\|_2^2 \tag{7.106}$$

$$\leq \|\mathbf{Y}_K(\boldsymbol{H}^{-1} - \boldsymbol{G}^{-1})\|_2^2 \, \|\boldsymbol{y}\|_2^2 \tag{7.107}$$

where $\mathbf{Y}_K : \mathbb{C}^{N\times N} \to \mathbb{C}^{N\times K}$ is a truncation operator that selects K columns of an $N \times N$ matrix. We now consider $\|\mathbf{Y}_K(\boldsymbol{H}^{-1} - \boldsymbol{G}^{-1})\|_2^2$ using the relation $\|\boldsymbol{C}\|_2 = \sqrt{\lambda_{\max}(\boldsymbol{C}'\boldsymbol{C})}$ where $\lambda_{\max}(\boldsymbol{A})$ denotes the maximum eigenvalue of $\boldsymbol{A}$. We compute

$$\|\mathbf{Y}_K(\boldsymbol{H}^{-1} - \boldsymbol{G}^{-1})\|_2 = \sqrt{\max_{i\in[1,\ldots,K]} \left(t_i^{-1} - \tau_i^{-1}\right)^2} \tag{7.108}$$

$$= \max_{i\in[1,\ldots,K]} \left|t_i^{-1} - \tau_i^{-1}\right| \tag{7.109}$$

$$= \max_{i\in[1,\ldots,K]} \left|1 - \frac{1}{\operatorname{sinc}(Wk_i)}\right| \tag{7.110}$$

Recall for Fourier sampling $\Delta_k W = \frac{1}{N}$ so

$$\|\mathbf{Y}_K(\boldsymbol{H}^{-1} - \boldsymbol{G}^{-1})\|_2 = \max_{i\in[1,\ldots,K]} \left|1 - \frac{1}{\operatorname{sinc}(i/N)}\right| \tag{7.111}$$

Since $K \leq N$, the largest difference will occur at $i = K$ so

$$\|\mathbf{Y}_K(\boldsymbol{H}^{-1} - \boldsymbol{G}^{-1})\|_2 = \frac{1}{\operatorname{sinc}(K/N)} - 1 \tag{7.112}$$

To analyse this we let $x = \frac{\pi K}{N}$ and expand $\sin(x)$ in a Taylor series,

$$\|\mathbf{Y}_K(\boldsymbol{H}^{-1} - \boldsymbol{G}^{-1})\|_2 = \frac{x}{\sin x} - 1 \tag{7.113}$$

$$= \frac{x - \sin x}{\sin x} \tag{7.114}$$

$$= \frac{x - (x - \frac{x^3}{3!} + \mathcal{O}(x^5))}{(x - \frac{x^3}{3!} + \mathcal{O}(x^5))} \tag{7.115}$$

$$= \frac{\frac{x^3}{3!} + \mathcal{O}(x^5)}{(x - \frac{x^3}{3!} + \mathcal{O}(x^5))} \tag{7.116}$$

$$= \mathcal{O}(x^2) \tag{7.117}$$

Thus,

$$\|\boldsymbol{f} - \boldsymbol{g}\|_2 \leq \frac{1}{\sqrt{N}} \|\mathbf{Y}_K(\boldsymbol{H}^{-1} - \boldsymbol{G}^{-1})\|_2 \|\boldsymbol{y}\|_2 \tag{7.118}$$

$$\frac{\|\boldsymbol{f} - \boldsymbol{g}\|_2}{\|\boldsymbol{y}\|_2} = \mathcal{O}\left(N^{-\frac{1}{2}}\right) \mathcal{O}\left(N^{-2}\right) \tag{7.119}$$

$$= \mathcal{O}\left(N^{-\frac{5}{2}}\right) \tag{7.120}$$

This analysis demonstrates that the difference between images reconstructed using

a delta basis and rectangular basis decays rapidly for increasing N. In other words, the difference between the delta basis and rectangular basis is negligible for a finely spaced pixel grid. Although this is well accepted in the research community, a rigourous analysis such as above has to the best of our knowledge, not appeared in the literature.

CHAPTER 8

Conclusion

THE fundamental aim of MRI is to non-invasively determine the detailed structure and function of a human body. This information is used for the diagnosis and treatment of patients as well as fundamental neuroscience and pathology. To satisfy all these applications, ideally MRI would be able to classify tissue, detect and localise diseases, and map different functions of the brain.

This ambitious goal is challenging for a multitude of reasons. The structure of the body is made up of an enormous number of atoms. MRI is only sensitive to atoms with a particular spin. Further, these spins are not directly observable; measurements only represent the average behaviour of the relevant spins. To make matters worse, the number of measurements is severely limited due to physical constraints, hardware constraints and safety constraints. Reasonably, one can expect that limited information will be attainable from the measurements of such a system.

The saving grace of MRI is also its main limitation – the ability to improve one property of the system at the expense of another. These tradeoffs are either inherent in the physical principles or imposed by the engineer.

The overarching conclusion of this thesis is the following. *In order to maximise the performance of MRI for a specific application, one must jointly consider accurate signal models, algorithm design and acquisition strategies.*

Although this statement is intuitive, the results in this thesis highlight several areas where performance improvements can be made by following this principle.

8.1 Summary of original contributions

Chapter 4 examined the tradeoff in Fourier reconstruction between SNR (related to sampling bandwidth) and image distortion. In the context of the estimation of transverse relaxation times, this translates to a tradeoff between bias and variance. A detailed analysis of the estimation problem revealed that the bias is often substantial when a simplified signal model is assumed. To overcome this, the estimation problem was reposed using an accurate signal model, which incorporates the image distortion. A Bayesian algorithm was developed using the improved signal

model to overcome the bias/variance tradeoff. Improved results were demonstrated at the expense of increased algorithm complexity and computation time.

Chapter 5 analysed the problem of estimating a distribution of relaxation times, which is more representative of the underlying structure than the assumption of a unique discrete component. However, it was shown that the accurate estimation of some properties of this distribution is infeasible due to the spin averaging. It was concluded from the results that although the location of the distribution components can be accurately estimated, the distribution width cannot be determined from typical measurements. Current state-of-the art techniques employ regularisation to improve the estimation results although this can unfavourably bias the distribution shape. This highlights the important point that one should not attempt to extract more information than is available in the measurements. The alternative proposed in this thesis is to estimate the locations of the distribution mass, assuming a width of zero. This assumption is justified since the width has minimal effect on the measurements; indeed this is the same reason the width cannot be estimated. The resulting estimation problem was cast in a Bayesian framework, from which a computationally efficient and reliable algorithm was developed. The algorithm exhibits near-optimal performance and improves on the currently available techniques.

Chapter 6 delved into the emerging field of nonlinear spatial encoding. A unique property of these techniques is that, unlike linear encoding, the fundamental resolution is spatially varying. This create both challenges and opportunities. A novel acquisition strategy was developed to improve the image resolution in a user-specified region. The technique uses nonlinear encoding fields together with the design of optimal gradient waveforms. In this way the spatially varying resolution inherent in nonlinear fields is exploited to improve the spatial encoding in the region of interest. The result is an image with enhanced resolution in the desired region at the expense of pronounced blurring in other regions. This direct tradeoff is useful for applications that are only interested in a particular region. This is often the case for brain imaging as practitioners may be focused on the region surrounding a brain tumour, or on the function of sub-cortical structures such as the hippocampus. Another application is cardiac imaging where the acquisition could be designed to optimally encode the heart region.

Chapter 7 revisited the basic image reconstruction problem applicable for encoding with nonlinear fields. The image reconstruction problem was recast using frame theory to provide a novel perspective on the noise properties of existing techniques. The source of aliasing and noise amplification for these techniques was elucidated using the Dirichlet kernel. Additionally, a computationally effi-

cient and intuitive noise metric was developed, based on an approximation of the reconstructed pixel covariance. This novel metric is applicable to general encoding schemes that use any combination of linear and nonlinear fields.

The work in this thesis underscores the importance of rigourously analysing the signals obtained from MRI experiments. The accuracy of estimated biological parameters was examined in Chapters 4 and 5. In both chapters, an initial analysis lead to an improved signal model and estimation algorithm. The spatially varying resolution inherent in nonlinear encoding fields was investigated in Chapter 6, which lead to novel acquisition strategies. Finally, the noise variance associated with image reconstruction of general encoding schemes was explored in Chapter 7, which we anticipate will be useful to evaluate emerging new techniques.

8.2 Future work

Although impressive performance is achievable with modern MRI technology, there are many opportunities for further enhancements. This thesis has demonstrated several specific improvements that can be made to current state-of-the-art techniques. Additionally, this work has confirmed that detailed signal analysis can lead to new insights and the development of novel techniques. This paradigm can also serve as a roadmap for future development.

8.2.1 Statistical estimation

The statistical estimation algorithms developed in this thesis are most useful for situations with low SNR. In this case, improved signal models and computationally demanding algorithms are required to extract the maximum amount of information available from the noisy measurements. Future research could apply similar algorithms to the recent applications of portable MRI and ultra-low field imaging, since these operate at inherently low SNR. The increased computation and complexity is mitigated by the dramatic increasing in computing performance in recent years. This, coupled with a reduction in hardware cost, makes the application of advanced estimation algorithms feasible.

Further research can be conducted into the estimation algorithms developed in this thesis. Improvements to the computational performance of the Gibbs sampling algorithm of Chapter 4 is required to make this technique feasible in a clinical setting. For example, the spatial variation of T_2 could be modelled with a Markov random field to capture the local similarities of biological tissue. This prior information would improve the sampling performance of the estimation algorithm.

The progressive correction technique in Chapter 5 would benefit from a theoretical analysis and proof that the posterior density approximation approaches the optimal density. In principle, diffusion weighting could also be added to the CPMG sequence and incorporated into the signal model, in order to distinguish the exchange component of the T_2 model.

8.2.2 Nonlinear encoding

Nonlinear spatial encoding is still a relatively new idea and future research can take a number of directions. Generalised encoding with multiple gradient channels incorporating linear and nonlinear fields adds degrees of freedom to the design of acquisition strategies. How to best utilise the extra degrees of freedom for a given application remains an open question. The optimisation in Chapter 6 is based on the single criteria of local k-space trajectories. Other factors such as noise performance, receive coil geometries, and acquisition speed may be incorporated into gradient waveform design.

A detailed investigation into peripheral nerve stimulation is required to verify the hypothesis that nonlinear encoding can reduce the risk of nerve stimulation. Although it is challenging to incorporate all factors that lead to peripheral nerve stimulation, such a study will be necessary prior to the clinical adoption of nonlinear encoding.

Chapter 7 analyses the SNR of general nonlinear encoding schemes. The performance metric developed, although useful, is not perfect. The covariance approximation relies on a diagonally dominant frame matrix, which neglects the structure of the off-diagonal terms. An ideal noise analysis would consider all the nuances of the encoding scheme, without having to compute the entire covariance matrix.

In addition to noise performance, further research is required into the resolution of nonlinear encoding schemes. Currently, resolution metrics are based on the point spread function which depends on the reconstruction choice. A metric that depends only on the encoding scheme would be useful to determine the fundamental resolvability of neighbouring features in the presence of noise. Analogous to the Cramér-Rao lower bound, an ideal resolution metric would provide a benchmark for reconstruction algorithms, not depend on them.

A complete characterisation of the imaging performance of nonlinear encoding schemes requires a joint analysis of both noise and resolution. A metric that combines both quantities would prove useful. The mean squared error naturally combines the quantities of bias and variance into a single performance metric and

is often used to benchmark different techniques. The prevalence of this error metric suggests that researchers working on nonlinear encoding could benefit from an analogous metric that combines SNR and resolution.

8.2.3 Signal modelling

There are signal dependencies other than those addressed in this thesis that are often not modelled. These include magnetic susceptibility, diffusion, blood flow and partial volume effects to name just a few. Simple assumptions about these processes causes a mismatch between the assumed signal model and the actual measurements. This is often referred to as 'systematic error' in MRI literature or 'model uncertainty' in signal processing literature. The mismatch is problematic for image reconstruction and leads to non-convergent algorithms or ill-defined artifacts. This problem could be alleviated by incorporating these effects into the appropriate signal model and reconstruction algorithm.

Further research is needed to quantify the effect of model uncertainty on image reconstruction and parameter estimation. As the boundaries of performance are pushed and more information is extracted from noisy measurements, clinicians will need measures of certainty in order to completely trust the images and parameter maps that are produce from an MRI scanner.

Bibliography

[1] M. Poole, *Improved equipment and techniques for dynamic shimming in high field MRI*. PhD thesis, The University of Nottingham, 2007. (Cited on page 1.)

[2] A. Macovski and S. Conolly, "Novel approaches to low-cost MRI," *Magnetic Resonance in Medicine*, vol. 30, no. 2, pp. 221–230, 1993. (Cited on pages 2 and 76.)

[3] J. P. Mugler and J. R. Brookeman, "The optimum data sampling period for maximum signal-to-noise ratio in MR imaging," *Reviews of magnetic resonance in medicine*, vol. 3, no. 1, pp. 1–51, 1988. (Cited on pages 3 and 60.)

[4] E. M. Haacke, "The effects of finite sampling in spin-echo or field-echo magnetic resonance imaging," *Magnetic Resonance in Medicine*, vol. 4, no. 5, pp. 407–421, 1987. (Cited on pages 3 and 60.)

[5] K. P. Whittall and A. L. MacKay, "Quantitative interpretation of NMR relaxation data," *Journal of Magnetic Resonance (1969)*, vol. 84, no. 1, pp. 134–152, 1989. (Cited on pages 4, 86, 90 and 91.)

[6] S. M. Kay, *Fundamentals of statistical signal processing: Estimation theory*. Prentice-Hall, 1993. (Cited on pages 4, 92, 93, 94 and 167.)

[7] J. L. Lancaster, T. Andrews, L. J. Hardies, S. Dodd, and P. T. Fox, "Three-pool model of white matter," *Journal of Magnetic Resonance Imaging*, vol. 17, no. 1, pp. 1–10, 2003. (Cited on pages 4, 86, 91 and 94.)

[8] T. Andrews, J. L. Lancaster, S. J. Dodd, C. Contreras-Sesvold, and P. T. Fox, "Testing the three-pool white matter model adapted for use with T2 relaxometry," *Magnetic Resonance in Medicine*, vol. 54, no. 2, pp. 449–454, 2005. (Cited on pages 4, 86, 87, 92 and 94.)

[9] J. Hennig, A. Welz, G. Schultz, J. Korvink, Z. Liu, O. Speck, and M. Zaitsev, "Parallel imaging in non-bijective, curvilinear magnetic field gradients: A concept study," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 21, pp. 5–14, 2008. (Cited on pages 4, 126 and 130.)

[10] G. Schultz, P. Ullmann, H. Lehr, A. Welz, J. Hennig, and M. Zaitsev, "Reconstruction of MRI data encoded with arbitrarily shaped, curvilinear, nonbijective magnetic fields," *Magnetic Resonance in Medicine*, vol. 64, no. 5, pp. 1390–1403, 2010. (Cited on pages 4, 126, 129, 130, 155, 160, 168, 172, 173, 181, 183 and 184.)

[11] J. P. Stockmann, P. A. Ciris, G. Galiana, L. Tam, and R. T. Constable, "O-space imaging: Highly efficient parallel imaging using second-order nonlinear fields as encoding gradients with no phase encoding," *Magnetic Resonance in Medicine*, vol. 64, no. 2, pp. 447–456, 2010. (Cited on pages 4, 126, 130, 140, 180, 183, 184 and 187.)

[12] D. Gallichan, C. A. Cocosco, A. Dewdney, G. Schultz, A. Welz, J. Hennig, and M. Zaitsev, "Simultaneously driven linear and nonlinear spatial encoding fields in MRI," *Magnetic Resonance in Medicine*, vol. 65, no. 3, pp. 702–714, 2011. (Cited on pages 4, 126, 129, 130, 134, 143, 155, 160 and 181.)

[13] O. Christensen, *Frames and bases: An introductory course*. Birkhäuser, 2008. (Cited on pages 4, 160 and 162.)

[14] K. Pruessmann, M. Weiger, M. Scheidegger, and P. Boesiger, "SENSE: sensitivity encoding for fast MRI," *Magnetic Resonance in Medicine*, vol. 42, no. 5, pp. 952–962, 1999. (Cited on pages 4, 133, 136, 160, 165, 167, 168, 169 and 173.)

[15] M. H. Levitt, *Spin dynamics: basics of nuclear magnetic resonance*. John Wiley and Sons, 2001. (Cited on pages 9, 24 and 31.)

[16] F. Bloch, "Nuclear induction," *Physics Review*, vol. 70, pp. 460–473, 1946. (Cited on page 9.)

[17] K. Blum, *Density matrix theory and applications*. Plenum Press, second ed., 1996. (Cited on pages 11 and 16.)

[18] E. Merzbacher, *Quantum mechanics*. John Wiley and Sons, third ed., 2004. (Cited on page 12.)

[19] I. I. Rabi, N. F. Ramsey, and J. Schwinger, "Use of rotating coordinates in magnetic resonance problems," *Rev. Mod. Phys.*, vol. 26, pp. 167–171, Apr 1954. (Cited on page 12.)

[20] J. Cavanagh, W. J. Fairbrother, A. G. Palmer, III, and N. J. Skelton, *Protein NMR spectroscopy, principles and practice*. Academic Press, 1996. (Cited on pages 20 and 27.)

[21] J. T. Bushberg, J. A. Seibert, E. M. Leidholdt, and J. M. Boone, *The essential physics of medical imaging*. Lippincott Williams and Wilkins, second ed., 2002. (Cited on pages 21 and 22.)

[22] P. A. Bottomley, T. H. Foster, R. E. Argersinger, and L. M. Pfeifer, "A review of normal tissue hydrogen NMR relaxation times and relaxation mechanisms from 1–100 MHz: dependence on tissue type, NMR frequency, temperature, species, excision, and age," *Medical Physics*, vol. 11, pp. 425–448, 1984. (Cited on page 21.)

[23] A. Abragam, *The principles of nuclear magnetism*. London: Oxford University Press, 1961. (Cited on pages 25 and 26.)

[24] P. Luginbühl and K. Wüthrich, "Semi-classical nuclear spin relaxation theory revisited for use with biological macromolecules," *Progress in Nuclear Magnetic Resonance Spectroscopy*, vol. 40, no. 3, pp. 199–247, 2002. (Cited on page 27.)

[25] A. G. Redfield, "On the theory of relaxation processes," *IBM J. Res. Dev*, vol. 1, pp. 19–31, 1957. (Cited on page 28.)

[26] M. Alecci, C. M. Collins, M. B. Smith, and P. Jezzard, "Radio frequency magnetic field mapping of a 3 Tesla birdcage coil: Experimental and theoretical dependence on sample properties," *Magnetic Resonance in Medicine*, vol. 46, no. 2, pp. 379–385, 2001. (Cited on page 37.)

[27] E. L. Hahn, "Spin echoes," *Physical Review*, vol. 80, pp. 580–594, 1950. (Cited on page 38.)

[28] H. Y. Carr and E. M. Purcell, "Effects of diffusion on free precession in nuclear magnetic resonance experiments," *Phys. Rev.*, vol. 94, pp. 630–638, May 1954. (Cited on pages 38 and 60.)

[29] S. Meiboom and D. Gill, "Modified spin-echo method for measuring nuclear relaxation times," *Review of Scientific Instruments*, vol. 29, no. 8, pp. 688–691, 1958. (Cited on pages 38 and 60.)

[30] S. Ljunggren, "A simple graphical representation of Fourier-based imaging methods," *Journal of Magnetic Resonance*, vol. 54, pp. 338–343, 1983. (Cited on page 44.)

[31] D. B. Twieg, "The k-trajectory formulation of the NMR imaging process with applications in analysis and synthesis of imaging methods," *Medical Physics*, vol. 10, no. 5, pp. 610–621, 1983. Times Cited: 154. (Cited on page 44.)

[32] M. L. Lauzon and B. K. Rutt, "Effects of polar sampling in k-space," *Magnetic Resonance in Medicine*, vol. 36, no. 6, pp. 940–949, 1996. (Cited on pages 45 and 47.)

[33] P. Lauterbur, "Image formation by induced local interactions: examples employing nuclear magnetic resonance," *Nature*, vol. 242, no. 5394, pp. 190–191, 1973. (Cited on page 47.)

[34] M. O'Donnell and W. A. Edelstein, "NMR imaging in the presence of magnetic field inhomogeneities and gradient field nonlinearities," *Medical Physics*, vol. 12, pp. 20–26, 1985. (Cited on page 47.)

[35] J. R. Reichenbach, R. Venkatesan, D. A. Yablonskiy, M. R. Thompson, S. Lai, and E. M. Haacke, "Theory and application of static field inhomogeneity effects in gradient-echo imaging," *Journal of Magnetic Resonance Imaging*, vol. 7, no. 2, pp. 266–279, 1997. (Cited on page 47.)

[36] K. Scheffler and J. Hennig, "Reduced circular field-of-view imaging," *Magnetic Resonance in Medicine*, vol. 40, no. 3, pp. 474–480, 1998. (Cited on page 49.)

[37] P. Mansfield, A. A. Maudsley, and T. Bains, "Fast scan proton density imaging by NMR," *Journal of Physics E: Scientific Instruments*, vol. 9, no. 4, pp. 271–278, 1976. (Cited on page 49.)

[38] I. L. Pykett and R. R. Rzedzian, "Instant images of the body by magnetic resonance," *Magnetic Resonance in Medicine*, vol. 5, no. 6, pp. 563–571, 1987. (Cited on page 49.)

[39] M. S. Cohen and R. M. Weisskoff, "Ultra-fast imaging," *Magnetic Resonance Imaging*, vol. 9, no. 1, pp. 1–37, 1991. (Cited on page 49.)

[40] C. P. Davis, G. C. Mckinnon, J. F. Debatin, and G. K. Schulthess, "Ultra-high-speed MR imaging," *European Radiology*, vol. 6, pp. 297–311, 1996. (Cited on pages 49 and 155.)

[41] J. A. Den Boer, J. D. Bourland, J. A. Nyenhuis, C. L. Ham, J. M. Engels, F. X. Hebrank, G. Frese, and D. J. Schaefer, "Comparison of the threshold for peripheral nerve stimulation during gradient switching in whole body MR systems," *Journal of Magnetic Resonance Imaging*, vol. 15, no. 5, pp. 520–525, 2002. (Cited on page 50.)

[42] P. P. M. So, M. A. Stuchly, and J. A. Nyenhuis, "Peripheral nerve stimulation by gradient switching fields in magnetic resonance imaging," *Biomedical Engineering, IEEE Transactions on*, vol. 51, no. 11, pp. 1907–1914, 2004. (Cited on pages 50 and 156.)

[43] H. Zeng and R. T. Constable, "Image distortion correction in EPI: Comparison of field mapping with point spread function mapping," *Magnetic Resonance in Medicine*, vol. 48, no. 1, pp. 137–146, 2002. (Cited on page 50.)

[44] Z. P. Liang and P. C. Lauterbur, *Principles of magnetic resonance imaging: a signal processing perspective*. IEEE Press, 2000. (Cited on page 51.)

[45] S. Kaczmarz, "Angenäherte auflösung von systemen linearer gleichungen," *Bull. Int. Acad. Polon. Sci. A*, vol. 35, pp. 355–357, 1937. (Cited on pages 52 and 184.)

[46] R. Gordon, R. Bender, and G. T. Herman, "Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and X-ray photography," *Journal of Theoretical Biology*, vol. 29, no. 3, pp. 471–481, 1970. (Cited on page 52.)

[47] G. T. Herman, *Fundamentals of computerized tomography: Image reconstruction from projections*. Springer, second ed., 2009. (Cited on page 52.)

[48] M. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, pp. 409–436, 1952. (Cited on page 52.)

[49] K. Pruessmann, M. Weiger, P. Börnert, and P. Boesiger, "Advances in sensitivity encoding with arbitrary k-space trajectories," *Magnetic Resonance in Medicine*, vol. 46, no. 4, pp. 638–651, 2001. (Cited on page 52.)

[50] Luenberger, *Introduction to linear and nonlinear programming*. Kluwer Academic Publishers, second ed., 2003. (Cited on page 53.)

[51] J. Johnson, "Thermal agitation of electricity in conductors," *Physical Review*, vol. 32, no. 1, pp. 97–109, 1928. (Cited on page 53.)

[52] H. Nyquist, "Thermal agitation of electric charge in conductors," *Physical Review*, vol. 32, no. 1, pp. 110–113, 1928. (Cited on page 53.)

[53] E. M. Haacke, R. W. Brown, M. R. Thompson, and R. Venkatesan, *Magnetic resonance imaging: Physical principles and sequence design*. John Wiley and Sons, 1999. (Cited on pages 54, 69 and 77.)

[54] W. A. Edelstein, G. H. Glover, C. J. Hardy, and R. W. Redington, "The intrinsic signal-to-noise ratio in NMR imaging," *Magnetic Resonance in Medicine*, vol. 3, no. 4, pp. 604–618, 1986. (Cited on page 54.)

[55] C. R. Haider, H. H. Hu, N. G. Campeau, J. Huston, and S. J. Riederer, "3D high temporal and spatial resolution contrast-enhanced MR angiography of the whole brain," *Magnetic Resonance in Medicine*, vol. 60, no. 3, pp. 749–760, 2008. (Cited on page 55.)

[56] J. Tsao, P. Boesiger, and K. P. Pruessmann, "k-t BLAST and k-t SENSE: Dynamic MRI with high frame rate exploiting spatiotemporal correlations," *Magnetic Resonance in Medicine*, vol. 50, no. 5, pp. 1031–1042, 2003. (Cited on page 55.)

[57] I. R. Levesque and G. B. Pike, "Characterizing healthy and diseased white matter using quantitative magnetization transfer and multicomponent T2 relaxometry: A unified view via a four-pool model," *Magnetic Resonance in Medicine*, vol. 62, no. 6, pp. 1487–1496, 2009. (Cited on page 60.)

[58] A. L. MacKay, I. M. Vavasour, A. Rauscher, S. H. Kolind, B. MÃd'dler, G. R. W. Moore, A. L. Traboulsee, D. K. B. Li, and C. Laule, "MR relaxation in multiple sclerosis," *Neuroimaging Clinics of North America*, vol. 19, no. 1, pp. 1–26, 2009. (Cited on page 60.)

[59] G. Bartzokis, T. Tishler, P. Lu, P. Villablanca, L. Altshuler, M. Carter, D. Huang, N. Edwards, and J. Mintz, "Brain ferritin iron may influence age- and gender-related risks of neurodegeneration," *Neurobiology of Aging*, vol. 28, no. 3, pp. 414–423, 2007. (Cited on page 60.)

[60] L. A. Shepp and B. F. Logan, "The fourier reconstruction of a head section," *Nuclear Science, IEEE Transactions on*, vol. 21, no. 3, pp. 21–43, 1974. (Cited on page 64.)

[61] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, no. 6, pp. 721–741, 1984. (Cited on page 72.)

[62] A. Smith and G. Roberts, "Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 55, no. 1, pp. 3–23, 1993. (Cited on page 72.)

[63] C. Ritter and M. Tanner, "Facilitating the Gibbs sampler: the Gibbs stopper and the griddy-Gibbs sampler," *J. Amer. Statistical Assoc.*, vol. 87, no. 419, pp. 861–868, 1992. (Cited on page 72.)

[64] J. S. Lui, *Monte Carlo Strategies in Scientific Computing*. Springer, 2008. (Cited on page 73.)

[65] K. Uludağ, B. Müller-Bierl, and K. Uğurbil, "An integrative model for neuronal activity-induced signal changes for gradient and spin echo functional imaging," *NeuroImage*, vol. 48, pp. 150–165, 2009. (Cited on page 74.)

[66] D. Price, I. Delakis, C. Renaud, and R. Dickinson, "MRI scanners: a buyer's guide." (Cited on page 76.)

[67] P. Bloomfield, *Fourier analysis of time series: an introduction*. John Wiley & Sons, 2000. (Cited on page 78.)

[68] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Dover, 1964. (Cited on page 79.)

[69] A. MacKay, C. Laule, I. Vavasour, T. Bjarnason, S. Kolind, and B. Mädler, "Insights into brain microstructure from the T2 distributions," *Magnetic Resonance Imaging*, vol. 24, no. 4, pp. 515–525, 2006. (Cited on page 86.)

[70] C. Poon and R. Henkelman, "Practical T2 quantitation for clinical applications," *Journal of Magnetic Resonance Imaging*, vol. 2, no. 5, pp. 541–553, 1992. (Cited on page 86.)

[71] J. Hennig, "Multiecho imaging sequences with low refocusing flip angles," *J Magn Reson*, vol. 78, no. 3, pp. 397–407, 1988. (Cited on pages 86 and 89.)

[72] J. Hennig, "Echoes–how to generate, recognize, use or avoid them in MR-imaging sequences. Part I: Fundamental and not so fundamental properties of spin echoes," *Concepts in Magnetic Resonance*, vol. 3, no. 3, pp. 125–143, 1991. (Cited on pages 86, 89 and 117.)

[73] G. J. Stanisz, E. E. Odrobina, J. Pun, M. Escaravage, S. J. Graham, M. J. Bronskill, and R. M. Henkelman, "T1, T2 relaxation and magnetization transfer in tissue at 3T," *Magnetic Resonance in Medicine*, vol. 54, no. 3, pp. 507–512, 2005. (Cited on pages 86 and 90.)

[74] 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*, vol. 31, no. 6, pp. 673–677, 1994. (Cited on pages 86 and 90.)

[75] K. Whittall, A. Mackay, D. Graeb, R. Nugent, D. Li, and D. Paty, "In vivo measurement of T2 distributions and water contents in normal human brain," *Magnetic Resonance in Medicine*, vol. 37, no. 1, pp. 34–43, 1997. (Cited on pages 86, 90, 103, 107 and 115.)

[76] S. Zheng and Y. Xia, "On the measurement of multi-component T2 relaxation in cartilage by MR spectroscopy and imaging," *Magnetic Resonance Imaging*, vol. 28, no. 4, pp. 537–545, 2010. (Cited on pages 86 and 90.)

[77] T. Prasloski, B. Mädler, Q.-S. Xiang, A. MacKay, and C. Jones, "Applications of stimulated echo correction to multicomponent T2 analysis," *Magnetic Resonance in Medicine*, 2011. (in press). (Cited on pages 86, 90, 91, 94, 103, 107 and 115.)

[78] K. J. Layton, M. Morelande, P. M. Farrell, B. Moran, and L. A. Johnston, "An improved algorithm for the estimation of multi-component T2 distributions," in *Proceedings of the ISMRM 20th Scientific Meeting and Exhibition*, (Melbourne), p. 2395, 2012. (Cited on page 86.)

[79] R. M. Lebel and A. H. Wilman, "Transverse relaxometry with stimulated echo compensation," *Magnetic Resonance in Medicine*, vol. 64, no. 4, pp. 1005–1014, 2010. (Cited on pages 87, 90, 91, 92, 94, 103, 107 and 116.)

[80] K. J. Layton, L. A. Johnston, P. M. Farrell, B. Moran, and M. Morelande, "Estimation of relaxation time distributions in magnetic resonance imaging," in *Proceedings of the IEEE Conference on Acoustics, Speech, and Signal Processing*, pp. 697–700, 2012. (Cited on page 87.)

[81] J. Sijbers and A. J. Den Dekker, "Maximum likelihood estimation of signal amplitude and noise variance from MR data," *Magnetic Resonance in Medicine*, vol. 51, no. 3, pp. 586–594, 2004. (Cited on pages 90 and 99.)

[82] S. J. Graham, P. L. Stanchev, and M. J. Bronskill, "Criteria for analysis of multicomponent tissue T2 relaxation data," *Magnetic Resonance in Medicine*, vol. 35, no. 3, pp. 370–378, 1996. (Cited on page 91.)

[83] N. Oudjane and C. Musso, "Progressive correction for regularized particle filters," in *Proceedings of the Third International Conference on Information Fusion*, vol. 2, pp. 10–17, 2002. (Cited on page 99.)

[84] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983. (Cited on page 100.)

[85] J. Wang, M. Qiu, Q. X. Yang, M. B. Smith, and R. T. Constable, "Measurement and correction of transmitter and receiver induced nonuniformities in vivo," *Magnetic Resonance in Medicine*, vol. 53, no. 2, pp. 408–417, 2005. (Cited on page 105.)

[86] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974. (Cited on page 111.)

[87] K. P. Burnham and D. R. Anderson, *Model selection and multimodel inference: a practical information-theoretic approach*. Springer, 2nd ed., 2002. (Cited on pages 111 and 115.)

[88] G. Schwarz, "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, no. 2, pp. 461–464, 1978. (Cited on page 111.)

[89] M. T. Chao, "The asymptotic behavior of Bayes' estimators," *The Annals of Mathematical Statistics*, vol. 41, no. 2, pp. 601–608, 1970. (Cited on page 111.)

[90] H. Iwaoka, T. Hirata, and H. Matsuura, "Optimal pulse sequences for magnetic resonance imaging-computing accurate T1, T2, and proton density images," *IEEE Transactions on Medical Imaging*, vol. 6, pp. 360–369, dec. 1987. (Cited on page 112.)

[91] L. Tam, J. P. Stockmann, and R. T. Constable, "Null space imaging: a novel gradient encoding strategy for highly efficient parallel imaging," in *Proceedings of the ISMRM 18th Annual Meeting*, (Stockholm), p. 2868, 2010. (Cited on pages 126 and 181.)

[92] D. Noll, "Rapid MR image acquisition in the presence of background gradients," in *Biomedical Imaging, 2002. Proceedings. 2002 IEEE International Symposium on*, pp. 725–728, 2002. (Cited on page 129.)

[93] B. A. Hargreaves, D. G. Nishimura, and S. M. Conolly, "Time-optimal multidimensional gradient waveform design for rapid imaging," *Magnetic Resonance in Medicine*, vol. 51, no. 1, pp. 81–92, 2004. (Cited on page 135.)

[94] M. Griswold, P. Jakob, R. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase, "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magnetic Resonance in Medicine*, vol. 47, no. 6, pp. 1202–1210, 2002. (Cited on pages 136 and 160.)

[95] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2009. (Cited on page 139.)

[96] G. Schultz, A. Welz, J. Hennig, and M. Zaitsev, "Generalized two-dimensional orthogonal spatial encoding fields," in *Proceedings of the ISMRM 16th Annual Meeting*, p. 2992, 2008. (Cited on page 139.)

[97] A. M. Welz, M. Zaitsev, F. Jia, Z. Liu, J. Korvink, H. Schmidt, H. Lehr, H. Post, A. Dewney, and J. Hennig, "Development of a non-shielded PatLoc gradient insert for human head imaging," in *Proceedings of the ISMRM 17th Scientific Meeting and Exhibition*, p. 3073, 2009. (Cited on page 143.)

[98] O. Speck, J. Stadler, and M. Zaitsev, "High resolution single-shot EPI at 7T," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 21, pp. 73–86, 2008. (Cited on page 144.)

[99] J. H. Duyn, Y. Yang, J. A. Frank, and J. W. van der Veen, "Simple correction method for k-space trajectory deviations in MRI," *Journal of Magnetic Resonance*, vol. 132, no. 1, pp. 150–153, 1998. (Cited on page 144.)

[100] D. Norris and J. Hutchison, "Concomitant magnetic field gradients and their effects on imaging at low magnetic field strengths," *Magnetic Resonance Imaging*, vol. 8, no. 1, pp. 33–37, 1990. (Cited on page 144.)

[101] P. Jezzard, A. S. Barnett, and C. Pierpaoli, "Characterization of and correction for eddy current artifacts in echo planar diffusion imaging," *Magnetic Resonance in Medicine*, vol. 39, no. 5, pp. 801–812, 1998. (Cited on page 144.)

[102] K. F. King, A. Ganin, X. J. Zhou, and M. A. Bernstein, "Concomitant gradient field effects in spiral scans," *Magnetic Resonance in Medicine*, vol. 41, no. 1, pp. 103–112, 1999. (Cited on page 144.)

[103] N. D. Zanche, C. Barmet, J. A. Nordmeyer-Massner, and K. P. Pruessmann, "NMR probes for measuring magnetic fields and field dynamics in MR systems," *Magn Reson Med*, vol. 60, pp. 176–186, Jul 2008. (Cited on page 144.)

[104] P. Sipilä, D. Lange, S. Lechner, W. Löw, P. Gross, M. Baller, G. Wachutka, and F. Wiesinger, "Robust, susceptibility-matched NMR probes for compensation of magnetic field imperfections in magnetic resonance imaging (MRI)," *Sensors and Actuators A: Physical*, vol. 145–146, pp. 139–146, 2008. (Cited on page 144.)

[105] C. Barmet, B. J. Wilm, M. Pavan, and K. P. Pruessmann, "A third-order field camera with microsecond resolution for MR system diagnostics," in *Proceedings ISMRM, 17th Scientific Meeting and Exhibition*, (Honolulu), p. 781, 2009. (Cited on pages 144 and 145.)

[106] C. Barmet, N. D. Zanche, and K. P. Pruessmann, "Spatiotemporal magnetic field monitoring for MR," *Magnetic Resonance in Medicine*, vol. 60, no. 1, pp. 187–197, 2008. (Cited on pages 144 and 145.)

[107] H. Han, R. P. MacGregor, and B. J. Balcom, "Pure phase encode magnetic field gradient monitor," *Journal of Magnetic Resonance*, vol. 201, pp. 212–217, Dec 2009. (Cited on page 144.)

[108] B. J. Wilm, C. Barmet, M. Pavan, and K. P. Pruessmann, "Higher order reconstruction for MRI in the presence of spatiotemporal field perturbations," *Magnetic Resonance in Medicine*, vol. 65, no. 6, pp. 1690–1701, 2011. (Cited on pages 144, 145 and 155.)

[109] F. Testud, D. Gallichan, C. Barmet, J. Vannesjö, A. M. Welz, C. A. Cocosco, K. PrÃijssmann, J. Hennig, and M. Zaitsev, "Characterization of PatLoc gradient with a field camera," in *Proceedings of the ISMRM 20th Scientific Meeting and Exhibition*, (Melbourne), p. 2598, 2012. (Cited on page 145.)

[110] C. Barmet, N. De Zanche, B. J. Wilm, and K. P. Pruessmann, "A transmit/receive system for magnetic field monitoring of in vivo MRI," *Magnetic Resonance in Medicine*, vol. 62, no. 1, pp. 269–276, 2009. (Cited on page 145.)

[111] D. Gallichan, C. Cocosco, G. Schultz, H. Weber, A. Welz, J. Hennig, and M. Zaitsev, "Practical considerations for in vivo MRI with higher dimensional spatial encoding," *Magnetic Resonance Materials in Physics, Biology and Medicine*, 2012. (early view). (Cited on page 145.)

[112] D. O. Walsh, A. F. Gmitro, and M. W. Marcellin, "Adaptive reconstruction of phased array MR imagery," *Magnetic Resonance in Medicine*, vol. 43, no. 5, pp. 682–690, 2000. (Cited on page 145.)

[113] C. Cocosco, A. Dewdney, P. Dietz, M. Semmler, A. Welz, D. Gallichan, H. Weber, G. Schultz, J. Hennig, and M. Zaitsev, "Safety considerations for a PatLoc gradient insert coil for human head imaging," in *Proceedings of the ISMRM 18th Scientific Meeting and Exhibition*, (Stockholm), p. 3946, 2010. (Cited on page 146.)

[114] C. A. Cocosco, D. Gallichan, A. J. Dewdney, G. Schultz, A. M. Welz, W. R. Witschey, H. Weber, J. Hennig, and M. Zaitsev, "First in-vivo results with a PatLoc gradient insert coil for human head imaging," in *Proceedings of the ISMRM 19th Scientific Meeting and Exhibition*, (Montreal), p. 714, 2011. (Cited on page 146.)

[115] S. Littin, A. M. Welz, D. Gallichan, G. Schultz, C. A. Cocosco, J. Hennig, W. DeBoer, and M. Zaitsev, "Planar gradient system for imaging with non-linear gradients," in *Proceedings of the ISMRM 19th Annual Meeting*, (Montreal), p. 1837, 2011. (Cited on page 155.)

[116] D. L. Parker, K. C. Goodrich, J. R. Hadley, S.-E. Kim, S. M. Moon, B. A. Chronik, U. Fontius, and F. Schmitt, "Magnetic resonance imaging with composite (dual) gradients," *Concepts in Magnetic Resonance Part B: Magnetic Resonance Engineering*, vol. 35B, no. 2, pp. 89–97, 2009. (Cited on page 156.)

[117] D. J. Schaefer, J. D. Bourland, and J. A. Nyenhuis, "Review of patient safety in time-varying gradient fields," *Journal of Magnetic Resonance Imaging*, vol. 12, no. 1, pp. 20–29, 2000. (Cited on page 156.)

[118] K. J. Layton, M. Morelande, P. M. Farrell, B. Moran, and L. A. Johnston, "Performance analysis for magnetic resonance imaging with nonlinear encoding fields," *IEEE Transactions on Medical Imaging*, vol. 31, pp. 391–404, feb. 2012. (Cited on page 156.)

[119] F. A. Breuer, S. A. Kannengiesser, M. Blaimer, N. Seiberlich, P. M. Jakob, and M. A. Griswold, "General formulation for quantitative g-factor calculation in GRAPPA reconstructions," *Magnetic Resonance in Medicine*, vol. 62, no. 3, pp. 739–746, 2009. (Cited on page 160.)

[120] S. Mallat, *A wavelet tour of signal processing*. Academic Press, 1999. (Cited on pages 160 and 166.)

[121] Z. Xu and A. Chan, "Encoding with frames in MRI and analysis of the signal-to-noise ratio," *IEEE Transactions on Medical Imaging*, vol. 21, no. 4, pp. 332–342, 2002. (Cited on page 160.)

[122] L. Panych, "Theoretical comparison of Fourier and wavelet encoding in magnetic resonance imaging," *IEEE Transactions on Medical Imaging*, vol. 15, no. 2, pp. 141–153, 1996. (Cited on pages 160 and 189.)

[123] M. Weiger, K. Pruessmann, and P. Boesiger, "2D SENSE for faster 3D MRI," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 14, pp. 10–19, 2002. (Cited on page 168.)

[124] R. D. Hoge, R. K. Kwan, and G. B. Pike, "Density compensation functions for spiral MRI," *Magnetic Resonance in Medicine*, vol. 38, no. 1, pp. 117–128, 1997. (Cited on page 175.)

[125] K. Layton, M. Morelande, P. Farrell, B. Moran, and L. Johnston, "A performance measure for MRI with nonlinear encoding fields," in *Proceedings of the ISMRM 19th Annual Meeting*, (Montreal), p. 482, 2011. (Cited on page 178.)

[126] P. A. Ciris, J. P. Stockmann, L. Tam, and R. T. Constable, "O-space imaging: tailoring encoding gradients to coil profiles for highly accelerated imaging," in *Proceedings of the ISMRM 17th Annual Meeting*, (Honolulu), p. 4556, 2009. (Cited on page 181.)

[127] G. Schultz, D. Gallichan, M. Reisert, M. Zaitsev, and J. Hennig, "Fast image reconstruction for generalized projection imaging," in *Proceedings of the ISMRM 19th Annual Meeting*, (Montreal), p. 2868, 2011. (Cited on page 182.)