

Advances in Quantitative MRI: Acquisition, Estimation, and Applications

by

Gopal Nataraj

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical Engineering and Computer Science)
in the University of Michigan
2016

Doctoral Committee:

Professor Jeffrey A. Fessler, Co-Chair
Assistant Research Scientist Jon-Fredrik Nielsen, Co-Chair
Professor Douglas C. Noll
Associate Professor Clayton Scott
Associate Research Scientist Scott Swanson

©Gopal Nataraj

2016

TABLE OF CONTENTS

List of Figures	iii
List of Tables	v
List of Appendices	vii
List of Abbreviations	viii
Abstract	ix
Chapter	
1 Introduction	1
1.1 Thesis Overview	2
1.2 Thesis Organization	2
2 Background	4
2.1 Relevant MR Physics	4
2.1.1 Bloch Equations	4
2.1.2 Steady-State Sequences	8
2.2 Optimization in QMRI	15
2.2.1 Iterative Local Optimization with Constraints	15
2.2.2 Partially Linear Models and the Variable Projection Method	16
3 MRI Parameter Estimation from Likelihood Models	18
3.1 Introduction	18
3.2 Likelihood-Based Estimation in QMRI	18
3.2.1 The QMRI Scan Profile	18
3.2.2 Latent Object Parameter Estimation	20
3.3 Experimentation	23
3.4 Summary	23
4 Optimizing MR Scan Design for Model-Based Relaxometry	24
4.1 Introduction	24
4.2 A CRB-Inspired Scan Selection Method	26
4.2.1 The CRB and its Relevance to QMRI	26
4.2.2 Min-max Optimization Problem for Scan Design	27
4.3 Optimizing SS Sequences for Relaxometry in the Brain	28

4.3.1	Scan Design Details	28
4.3.2	Scan Profile Comparisons	30
4.4	Experimental Validation and Results	32
4.4.1	Numerical Simulations	33
4.4.2	Phantom Experiments	36
4.4.3	<i>In Vivo</i> Experiments	43
4.5	Discussion and Future Work	46
4.6	Conclusion	47
5	MRI Parameter Estimation via Kernel Regression	49
5.1	Introduction	49
5.2	A Function Optimization Problem & Kernel Solution	51
5.3	Implementation Considerations	53
5.3.1	A Kernel Approximation	53
5.3.2	Model Selection	55
5.4	Experimentation	58
5.5	Summary and Future Work	61
6	Myelin Water Fraction Estimation from Steady-State Sequences	63
7	Future Work	64
	Appendices	65
	Bibliography	67

LIST OF FIGURES

4.1	Worst-case standard deviations $\tilde{\sigma}_{T_1}^t$ (top), $\tilde{\sigma}_{T_2}^t$ (middle), and cost $\tilde{\Psi}^t$ (bottom), versus pairs of nominal flip angles, holding other scan parameters fixed at selected profile $\hat{\mathbf{P}}$. Subfigures (a)-(i), (j)-(l), and (m)-(o) correspond to scan profiles containing $(S_{\text{SPGR}}, S_{\text{DESS}}) = (2, 1), (1, 1), \text{ and } (0, 2)$ SPGR and DESS scans, respectively. Selected scan parameters (starred) are within $\delta = 1\%$ of global minimizers and retain as much estimator precision as possible over a wide range of latent object parameters. All axes range from 5 to 90 degrees, in 5-degree increments. Colorbar ranges are $[0, 100]$, $[0, 10]$, and $[0, 20]$ milliseconds for rows of $\tilde{\sigma}_{T_1}^t$, $\tilde{\sigma}_{T_2}^t$, and $\tilde{\Psi}^t$ subfigures, respectively. The optimized $(0, 2)$ profile appears most robust to transmit field spatial variation.	31
4.2	Histograms of T_1 and T_2 estimates from noisy independent measurements of a <i>single</i> nominal WM or GM value. In each plot, two normal distributions are overlaid, each with latent means T_1 and T_2 . In (a)-(b) and (c)-(d), the solid green curve is $\mathcal{N}(T_1, (\tilde{\sigma}_{T_1}^t)^2)$ and $\mathcal{N}(T_2, (\tilde{\sigma}_{T_2}^t)^2)$, respectively. In (a)-(d), the dashed maroon curves have variances computed from the Fisher information at <i>a priori</i> unknown T_1, T_2 values in WM or GM. These plots correspond to an optimized $(0, 2)$ scan profile; analogous plots for other profiles are visually similar. At realistic noise levels, parameter estimates distribute with minimal bias and near-Gaussian shape. Thus, the CRB reliably approximates \hat{T}_1^{ML} and \hat{T}_2^{ML} errors.	35
4.3	Colorized T_1 and T_2 ML and RL estimates from an HPD [®] quantitative phantom. Columns correspond to scan profiles consisting of (2 SPGR, 1 DESS), (1 SPGR, 1 DESS), (0 SPGR, 2 DESS), and (4 IR, 4 SE) acquisitions. Rows distinguish T_1 and T_2 ML and RL estimators. Fig. 4.4 provides identical grayscale images that enumerate vials. Colorbar ranges are in milliseconds.	38
4.4	Grayscale T_1 and T_2 ML and RL estimates from an HPD [®] quantitative phantom. Columns correspond to scan profiles consisting of (2 SPGR, 1 DESS), (1 SPGR, 1 DESS), (0 SPGR, 2 DESS), and (4 IR, 4 SE) acquisitions. Rows distinguish T_1 and T_2 ML and RL estimators. Vials are enumerated and color-coded to correspond with data points in Fig. 4.5. Fig. 4.3 provides identical colorized images. Colorbar ranges are in milliseconds.	39

4.5	Phantom within-ROI sample statistics of T_1 and T_2 ML and RL estimates from optimized SPGR/DESS and reference IR/SE scan profiles, versus NIST NMR measurements [1]. Markers and error bars indicate ROI sample means and ROI sample standard deviations within the 14 labeled and color-coded vials in Fig. 4.4. Tight \mathbb{X}^t and broad \mathbb{X}^b latent parameter ranges are highlighted in orange and yellow, respectively. Table 4.3 replicates sample statistics within Vials 5-8. Our MR measurements are at 293K and NIST NMR measurements are at 293.00K. Within the designed parameter ranges, estimates from different acquisitions are in reasonable agreement with NIST measurements.	40
4.6	Colorized T_1 and T_2 ML and RL estimates from the brain of a healthy volunteer. Columns correspond to profiles consisting of (2 SPGR, 1 DESS), (1 SPGR, 1 DESS), (0 SPGR, 2 DESS), and (4 IR, 4 SE) acquisitions. Rows distinguish T_1 and T_2 ML and RL estimators. Table 4.5 presents corresponding WM/GM within-ROI sample statistics. Colorbar ranges are in milliseconds.	44
5.1	True f_F (<i>left</i>) and estimated \hat{f}_F fast-relaxing compartmental fraction maps, in simulation. Maximum-likelihood estimation via variable projection method and grid search (<i>center</i>) is accurate but is computationally expensive. In contrast, kernel ridge regression (<i>right</i>) is very fast and achieves comparable precision, at the expense of slightly increased bias (<i>cf.</i> Table 5.1). Voxels outside WM/GM regions are masked out in post-processing for visual clarity.	60

LIST OF TABLES

4.1	Performance summary of different scan profiles, optimized by solving (4.9) subject to scan time constraint $T_{\max} = 41.9\text{ms}$. The first row defines each profile. The next four rows describe $\hat{\mathbf{P}}$. The latter three pairs of rows show how worst-case values degrade from tight to broad ranges. Flip angles are in degrees; all other values are in milliseconds.	32
4.2	Sample means \pm sample standard deviations of T_1 and T_2 ML estimates in WM and GM ROIs of simulated data, compared across different optimized scan profiles. Sample means exhibit insignificant bias, and sample standard deviations are consistent with worst-case standard deviations $\tilde{\sigma}_{T_1}^t$ and $\tilde{\sigma}_{T_2}^t$ reported in Table 4.1. All values are reported in milliseconds.	34
4.3	Phantom within-ROI sample means \pm sample standard deviations of T_1 and T_2 estimates from optimized SPGR/DESS and reference IR/SE scan profiles, versus NIST NMR measurements (<i>cf.</i> slide 22 of e-poster corresponding to [1]). For sake of brevity, sample statistics corresponding only to phantom vials within (or nearly within) tight design range \mathbb{X}^t (color-coded orange in Fig. 4.4) are reported. Fig. 4.5 plots sample statistics for all vials. ‘V#’ abbreviates vial numbers. All values are reported in milliseconds.	41
4.4	Phantom pooled sample standard deviations \pm pooled standard errors of sample standard deviations, from optimized SPGR/DESS scan profiles. Each entry is a measure of uncertainty of a typical voxel’s T_1 or T_2 ML estimate, estimated over 10 repeated acquisitions. For sake of brevity, sample statistics corresponding only to phantom vials within (or nearly within) tight design range \mathbb{X}^t (color-coded orange in Fig. 4.4) are reported. ‘V#’ abbreviates vial numbers. All values are reported in milliseconds.	42
4.5	<i>Left:</i> WM/GM ROIs, overlaid on a representative anatomical (coil-combined, IR) image. Separate WM ROIs are distinguished by anterior-right (AR), anterior-left (AL), posterior-right (PR), and posterior-left (PL) directions. Four small anterior (A) cortical GM polygons are pooled into a single ROI. <i>Right:</i> Within-ROI sample means \pm within-ROI sample standard deviations of T_1 and T_2 ML and RL estimates from the brain of a healthy volunteer (Fig. 4.6 presents corresponding images). Sample statistics are computed within ROIs indicated in the anatomical image. All values are reported in milliseconds.	45

5.1	Sample means \pm sample standard deviations of fast-relaxing compartmental fraction estimates \hat{f}_F , computed over simulated WM- and GM-like voxels. Each sample statistic is rounded off to the highest place value of its (unreported) standard error, which is computed via formulas in [2].	61
-----	--	----

LIST OF APPENDICES

A Coil Data Combination from Multiple Datasets	65
B DESS in the Presence of Diffusion	66

LIST OF ABBREVIATIONS

ABSTRACT

**Advances in Quantitative MRI:
Acquisition, Estimation, and Applications**

by

Gopal Nataraj

Co-Chairs: Jeffrey A. Fessler and Jon-Fredrik Nielsen

todo

CHAPTER 1

Introduction

{c,intro}

Magnetic resonance imaging (MRI) is a non-invasive tool that has earned widespread clinical adoption due (among other factors) to its potential for excellent soft tissue contrast, its avoidance of ionizing radiation, and its flexibility to characterize a diversity of physical phenomena. Despite its numerous advantages, MRI requires highly specialized hardware, ongoing liquid-helium cooling of its superconducting main magnet, and comparably long scan times. For these reasons, MRI is expensive relative to other medical imaging modalities. Accordingly, one broad initiative recently advocated by the MR community is to increase the *value* of MRI examinations.

Two reasonable measures of an MRI acquisition’s value are its *sensitivity* to a given disorder and its *specificity* in distinguishing it from others. The field of *quantitative MRI* (QMRI) seeks to use MRI data to build MR *biomarkers*, or measurable tissue properties that can increase the sensitivity and specificity of MRI for specific disorders of interest.

QMRI has potential to be more informative than conventional MRI. Conventional MRI is *qualitative*: it produces images comprised of *voxels* (*i.e.*, three-dimensional pixels) that are informative only relative to each other, not individually. Conventional MRI voxels are qualitative because they localize the MR signal, which is typically a complex function of not only biomarkers but also two types of confounds: *nuisance markers* that characterize undesired signal sources and/or MRI system imperfections; and *acquisition parameters* that characterize the MRI system’s tunable “knobs”. QMRI seeks to remove confound influence by instead imaging the biomarkers directly. Each QMR image voxel is thus a measurement of a given biomarker at a specific location. QMRI can therefore provide localized biomarker measurements (*e.g.*, myelin water content) related to a specific physiological process (*e.g.*, demyelination) that can, through longitudinal study, be used to monitor the onset and progression of disease (*e.g.*, multiple sclerosis).

QMRI poses several challenges beyond those of conventional MRI that currently limit its feasibility for routine clinical use. For example, accurate biomarker quantification tra-

ditionally requires multiple MR scans and thus long scan times. Furthermore, it has previously been unclear how to tune acquisition parameters of these multiple scans to ensure that biomarkers can be quantified precisely. Finally, MR biomarker quantification is a challenging estimation problem for which efficient algorithms have previously been unavailable. Addressing these challenges is essential for widespread clinical adoption of QMRI.

1.1 Thesis Overview

{s,intro,over}

This thesis seeks to address the above challenges by building an automated workflow for QMRI. We borrow tools from optimization, statistics, and machine learning to develop fast workflows for quantifying biomarkers that characterize specific physiological processes. We apply this framework to challenging QMRI problems of clinical interest. Our goal is to introduce fast, automated tools that will increase the clinical value of QMRI.

Our solutions to two distinct subproblems in QMRI constitute two stages of our proposed QMRI workflow. Questions in *acquisition design* (Chapters 4, 6) ask how to assemble fast collections of scans that yield data rich in information about physical processes of interest. Questions in *parameter estimation* (Chapters 3, 5) ask how to quickly and reliably quantify biomarkers associated with these relevant physical processes. The overall workflow seeks to first design fast and informative scans based on the application, and to then accurately and precisely estimate clinically relevant biomarkers.

1.2 Thesis Organization

{s,intro,org}

The main body of this thesis is organized as follows:

- Chapter 2 reviews relevant background material on MRI and optimization.
- Chapter 3 discusses methods for MRI parameter estimation from likelihood models and applies these methods for model-based MR relaxometry, (*i.e.*, estimation of relaxation parameters T_1, T_2), of interest for many neurological applications. It derives some content (especially regarding applications) from conference papers [3, 4].
- Chapter 4 introduces a minimax optimization approach to aid robust and application-specific MR scan selection and optimization for precise latent parameter estimation. It optimizes several practical acquisitions and uses the likelihood-based estimation techniques introduced in Chapter 3 to assess the utility of scan optimization through

simulations, phantom studies, and *in vivo* experiments. It derives content mainly from journal paper [5] and conference paper [6].

- Chapter 5 describes MRI parameter estimation using kernel ridge regression. It derives content from conference paper [7].
- Chapter 6 introduces a multi-compartmental model for relevant MR pulse sequences and proposes a new acquisition useful for myelin water fraction estimation, of interest in white matter disorders. It applies kernel-based MR parameter estimation to estimate myelin water fraction, in simulations and *in vivo* experiments. It derives some content from conference paper [8].
- Chapter 7 summarizes several items of possible future work (on both short- and long-term timescales) and presents a timeline for completion of this thesis.

The appendices are organized as follows:

- Appendix A proposes an algorithm for combining multiple MRI datasets (as is necessary for many parameter estimation problems), when each dataset is acquired using multiple receiver coils.
- Appendix B presents an analysis of model mismatch due to the presence of diffusion, shows that neglecting diffusive effects during T_2 estimation can cause significant bias, and suggests acquisition modifications for mitigating this bias.

CHAPTER 2

Background

{c,bkgrd}

This chapter focuses only on background information pertinent to multiple subsequent chapters. We present further topic-specific information at the beginnings of corresponding chapters. Section 2.1 places emphasis on reviewing necessary MR fundamentals, and Section 2.2 proceeds to a shorter discussion regarding optimization as it pertains to QMRI.

2.1 Relevant MR Physics

{s,bkgrd,mri}

This section begins with the fundamental Bloch equations and derives the signal models associated with two MR pulse sequences used extensively in this thesis. Our coverage of MRI is far from comprehensive, and omits fundamental but tangential topics such as signal localization. We refer the interested reader to books such as [9, 10, 11].

2.1.1 Bloch Equations

{ss,bkgrd,mri,bloch}

The Bloch equations [12] describe the macroscopic magnetization dynamics of *spin*, or (loosely) atomic nuclei with nonzero angular momentum and thus nonzero magnetic moment, *e.g.* ^1H . If the dominant source of magnetic flux arises (as is typical in MRI) from a main magnetic field that is oriented along the z -axis, the equations read

{eq:bloch-mxy}
$$\frac{\partial}{\partial t} m_{xy}(\mathbf{r}, t) = i\gamma(m_z(\mathbf{r}, t)b_{xy}(\mathbf{r}, t) - m_{xy}(\mathbf{r}, t)b_z(\mathbf{r}, t)) - \frac{m_{xy}(\mathbf{r}, t)}{T_2(\mathbf{r})}; \quad (2.1)$$

{eq:bloch-mz}
$$\frac{\partial}{\partial t} m_z(\mathbf{r}, t) = \gamma(m_x(\mathbf{r}, t)b_y(\mathbf{r}, t) - m_y(\mathbf{r}, t)b_x(\mathbf{r}, t)) - \frac{m_z(\mathbf{r}, t) - m_0(\mathbf{r})}{T_1(\mathbf{r})}. \quad (2.2)$$

Here, $m_{xy}(\mathbf{r}, t) := m_x(\mathbf{r}, t) + im_y(\mathbf{r}, t) \in \mathbb{C}$ and $m_z(\mathbf{r}, t) \in \mathbb{R}$ are the transverse and longitudinal components of the magnetization vector at position $\mathbf{r} := [x, y, z]^T \in \mathbb{R}^3$ and time $t \geq 0$; $b_{xy}(\mathbf{r}, t) := b_x(\mathbf{r}, t) + ib_y(\mathbf{r}, t) \in \mathbb{C}$ and $b_z(\mathbf{r}, t) \in \mathbb{R}$ are the transverse and longitudinal components (in an inertial reference frame) of the applied magnetic field; $T_1(\mathbf{r})$

and $T_2(\mathbf{r})$ are spin-lattice and spin-spin relaxation time constants; $m_0(\mathbf{r})$ is the equilibrium magnetization and is proportional to the density of spins per unit volume as well as the main field strength; γ is the gyromagnetic ratio; and $i := \sqrt{-1}$. As written, (2.1)-(2.2) specifically model the temporal dynamics of a single spin *isochromat*, or collection of macroscopically similar spins; later chapters consider second-order effects such as multiple (possibly interacting) isochromat compartments (Chapter 6) and diffusion (Appendix B).

only model dominant temporal dynamics; later chapters consider second-order effects such as multiple magnetization compartments (Chapter 6) and diffusion (Appendix B).

It is often convenient to study Bloch dynamics in a non-inertial reference frame rotating clockwise about the z -axis at Larmor frequency $\omega_0 := \gamma B_0$, where $B_0 \hat{k}$ is the (nearly uniform) main magnetic field. In these coordinates, the apparent transverse magnetic field $b'_{xy}(\mathbf{r}, t) = b'_x(\mathbf{r}, t) + ib'_y(\mathbf{r}, t) := b_{xy}(\mathbf{r}, t)e^{i\omega_0 t}$ transforms only in phase, but the apparent longitudinal magnetic field $b'_z(\mathbf{r}, t) := b_z(\mathbf{r}, t) - B_0$ is greatly reduced in magnitude. The magnetization components transform more simply as $m'_{xy}(\mathbf{r}, t) = m'_x(\mathbf{r}, t) + im'_y(\mathbf{r}, t) := m_{xy}(\mathbf{r}, t)e^{i\omega_0 t}$ and $m'_z(\mathbf{r}, t) := m_z(\mathbf{r}, t)$. Remarkably, inserting these coordinate transformations into (2.1)-(2.2) does not change the form of the dynamical equations:

$$\{\text{eq:bloch-mxyp}\} \quad \frac{\partial}{\partial t} m'_{xy}(\mathbf{r}, t) = i\gamma(m'_z(\mathbf{r}, t)b'_{xy}(\mathbf{r}, t) - m'_{xy}(\mathbf{r}, t)b'_z(\mathbf{r}, t)) - \frac{m'_{xy}(\mathbf{r}, t)}{T_2(\mathbf{r})}; \quad (2.3)$$

$$\{\text{eq:bloch-mzp}\} \quad \frac{\partial}{\partial t} m'_z(\mathbf{r}, t) = \gamma(m'_x(\mathbf{r}, t)b'_y(\mathbf{r}, t) - m'_y(\mathbf{r}, t)b'_x(\mathbf{r}, t)) - \frac{m'_z(\mathbf{r}, t) - m_0(\mathbf{r})}{T_1(\mathbf{r})}. \quad (2.4)$$

It thus suffices to consider how perturbations $\mathbf{b}'(\mathbf{r}, t)$ to main field $B_0 \hat{k}$ influence rotating-frame magnetization $\mathbf{m}'(\mathbf{r}, t)$ via Eqs. (2.3)-(2.4). The inertial-frame magnetization $\mathbf{m}(\mathbf{r}, t)$ is then easily constructed via $m_{xy}(\mathbf{r}, t) = m'_{xy}(\mathbf{r}, t)e^{-i\omega_0 t}$ and $m_z(\mathbf{r}, t) = m'_z(\mathbf{r}, t)$.

It is challenging to explicitly solve Eqs. (2.3)-(2.4) for arbitrary field perturbations $\mathbf{b}'(\mathbf{r}, t)$. We discuss relevant special cases in the following.

2.1.1.1 Non-Selective Excitation

Here, we derive solutions to Eqs. (2.3)-(2.4) in the case of short, spatially non-selective excitations. We take the following common assumptions:

- We assume negligible spatial variation in the main magnetic field, so $b'_z(\mathbf{r}, t) \approx 0$.
- We assume the transverse field separates in position and time; oscillates at the Larmor frequency (commonly in the radiofrequency (RF) range); and aligns at initial time $t \leftarrow t_0$ with the x -axis. Together, these assumptions restrict the so-called RF

excitation to take form $b'_{xy}(\mathbf{r}, t) \approx s^t(\mathbf{r})b'_{1,x}(t)\hat{i} + 0\hat{j}$, where $s^t(\mathbf{r}) \in \mathbb{R}$ is the RF transmit coil spatial variation and $b'_{1,x}(t) \in \mathbb{R}$ is the RF excitation envelope.

- We assume that the duration T_P of RF excitation (often $T_P \sim 1\text{ms}$) is much shorter than relaxation time constants (typically $T_1 \sim 1000\text{ms}$ and $T_2 \sim 50\text{ms}$ in brain tissue) and thus neglect relaxation effects during excitation.

Under these assumptions, Eqs. (2.3)-(2.4) reduce to the linear system

$$\frac{\partial}{\partial t} \begin{bmatrix} m'_x(\mathbf{r}, t) \\ m'_y(\mathbf{r}, t) \\ m'_z(\mathbf{r}, t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \gamma s^t(\mathbf{r})b'_{1,x}(t) \\ 0 & -\gamma s^t(\mathbf{r})b'_{1,x}(t) & 0 \end{bmatrix} \begin{bmatrix} m'_x(\mathbf{r}, t) \\ m'_y(\mathbf{r}, t) \\ m'_z(\mathbf{r}, t) \end{bmatrix}. \quad (2.5)$$

Eq. (2.5) admits the simple solution (for $t \geq t_0$)

$$\begin{bmatrix} m'_x(\mathbf{r}, t) \\ m'_y(\mathbf{r}, t) \\ m'_z(\mathbf{r}, t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha(\mathbf{r}, t; t_0)) & \sin(\alpha(\mathbf{r}, t; t_0)) \\ 0 & -\sin(\alpha(\mathbf{r}, t; t_0)) & \cos(\alpha(\mathbf{r}, t; t_0)) \end{bmatrix} \begin{bmatrix} m'_x(\mathbf{r}, t_0) \\ m'_y(\mathbf{r}, t_0) \\ m'_z(\mathbf{r}, t_0) \end{bmatrix}, \quad (2.6)$$

where $\mathbf{m}'(\mathbf{r}, t_0) := [m'_x(\mathbf{r}, t_0), m'_y(\mathbf{r}, t_0), m'_z(\mathbf{r}, t_0)]^\top$ is the initial magnetization and

$$\alpha(\mathbf{r}, t; t_0) := \gamma s^t(\mathbf{r}) \int_{t_0}^t b'_{1,x}(\tau) d\tau \quad (2.7)$$

is the nutation (or “flip”) angle at time t . Eq. (2.6) reveals that on-resonance RF excitation causes the magnetization vector to rotate clockwise about an axis parallel to the direction of excitation. The nutation angle accumulated over an RF pulse of duration T_P is often decomposed as $\alpha(\mathbf{r}, t_0 + T_P; t_0) =: \alpha_0 s^t(\mathbf{r})$, where α_0 is a prescribed nominal flip angle.

For deriving signal models in later sections, it is convenient and intuitive to define matrix operators that summarize relevant dynamics. Here, we rewrite Eq. (2.6) as

$$\mathbf{m}'(\mathbf{r}, t) = \mathbf{R}_{x'}(\alpha(\mathbf{r}, t; t_0))\mathbf{m}'(\mathbf{r}, t_0), \quad (2.8)$$

where $\mathbf{R}_{x'}(\alpha(\mathbf{r}, t; t_0))$ denotes a clockwise rotation of angle $\alpha(\mathbf{r}, t; t_0)$ about the x' -axis.

2.1.1.2 Free Precession and Relaxation

Next, we derive solutions to the rotating-frame Bloch equations when no RF excitation is present, *i.e.* $b'_{xy}(\mathbf{r}, t) \approx 0$. In this case, Eqs. (2.3)-(2.4) decouple, yielding separate

dynamical equations for the transverse and longitudinal magnetization components:

$$\{\text{eq:bloch-free-mxyp}\} \quad \frac{\partial}{\partial t} m'_{xy}(\mathbf{r}, t) = -i\gamma m'_{xy}(\mathbf{r}, t) b'_z(\mathbf{r}, t) - \frac{m'_{xy}(\mathbf{r}, t)}{T_2(\mathbf{r})}; \quad (2.9)$$

$$\{\text{eq:bloch-free-mzp}\} \quad \frac{\partial}{\partial t} m'_z(\mathbf{r}, t) = -\frac{m'_z(\mathbf{r}, t) - m_0(\mathbf{r})}{T_1(\mathbf{r})}. \quad (2.10)$$

Eqs. (2.9)-(2.10) admit simple solutions with no further assumptions:

$$\{\text{eq:mxy-fp}\} \quad m'_{xy}(\mathbf{r}, t) = m'_{xy}(\mathbf{r}, t_0) e^{-(t-t_0)/T_2(\mathbf{r})} e^{-i\phi'(\mathbf{r}, t; t_0)}; \quad (2.11)$$

$$\{\text{eq:mz-fp}\} \quad m'_z(\mathbf{r}, t) = m'_z(\mathbf{r}, t_0) e^{-(t-t_0)/T_1(\mathbf{r})} + m_0(\mathbf{r}) (1 - e^{-(t-t_0)/T_1(\mathbf{r})}), \quad (2.12)$$

where $m'_{xy}(\mathbf{r}, t_0)$ and $m'_z(\mathbf{r}, t_0)$ are the initial magnetization components and

$$\{\text{eq:ph-def}\} \quad \phi'(\mathbf{r}, t; t_0) := \gamma \int_{t_0}^t b'_z(\mathbf{r}, \tau) d\tau \quad (2.13)$$

denotes the phase accumulation due to main field inhomogeneity (often called off-resonance effects). Eq. (2.11) reveals that without RF excitations, the transverse magnetization $m'_{xy}(\mathbf{r}, t)$ relaxes to zero exponentially fast with time constant $T_2(\mathbf{r})$, while accruing phase due to off-resonance effects. Eq. (2.12) similarly reveals that without RF excitations, longitudinal magnetization $m'_z(\mathbf{r}, t)$ recovers to $m_0(\mathbf{r})$ exponentially fast with time constant $T_1(\mathbf{r})$.

As in Section 2.1.1.2, we rewrite Eqs. (2.11)-(2.12) for $t \geq t_0$ using matrix operators:

$$\{\text{eq:mtx-pr}\} \quad \mathbf{m}'(\mathbf{r}, t) = \mathbf{R}_{z'}(\phi'(\mathbf{r}, t; t_0)) \mathbf{E}(\mathbf{r}, t; t_0) \mathbf{m}'(\mathbf{r}, t_0) + \mathbf{m}_0(\mathbf{r}, t; t_0) \quad (2.14)$$

where $\mathbf{m}_0(\mathbf{r}, t; t_0) := m_0(\mathbf{r}) (1 - e^{-(t-t_0)/T_1(\mathbf{r})}) \hat{k}$;

$$\{\text{eq:op-rotz}\} \quad \mathbf{R}_{z'}(\phi'(\mathbf{r}, t; t_0)) := \begin{bmatrix} \cos(\phi'(\mathbf{r}, t; t_0)) & \sin(\phi'(\mathbf{r}, t; t_0)) & 0 \\ -\sin(\phi'(\mathbf{r}, t; t_0)) & \cos(\phi'(\mathbf{r}, t; t_0)) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.15)$$

denotes a clockwise rotation of angle $\phi'(\mathbf{r}, t; t_0)$ about the z' -axis; and

$$\{\text{eq:op-relax}\} \quad \mathbf{E}(\mathbf{r}, t; t_0) := \begin{bmatrix} e^{-(t-t_0)/T_2(\mathbf{r})} & 0 & 0 \\ 0 & e^{-(t-t_0)/T_2(\mathbf{r})} & 0 \\ 0 & 0 & e^{-(t-t_0)/T_1(\mathbf{r})} \end{bmatrix} \quad (2.16)$$

is an exponential relaxation operator. Section 2.1.2 (and later chapters) use matrix dynamical representations (2.8) and (2.14) to succinctly describe pulse sequence signal models.

{ss,bkgrd,mri,ss}

2.1.2 Steady-State Sequences

MRI experiments typically involve repeated cycles of (pulsed) RF excitation; signal localization (not discussed here); and transverse T_2 relaxation and free precession, alongside (relatively slow) longitudinal T_1 recovery. We can build models of the received MR signal by considering the magnetization dynamics induced by specific pulse sequences.

Classical pulse sequences use relatively long cycle repetition times T_R to ensure near-complete T_1 recovery of the magnetization vector back to equilibrium state $m_0(\mathbf{r})\hat{k}$ prior to the start of each RF cycle. For such long- T_R sequences, it suffices to approximate the magnetization as fully recovered (*i.e.*, $\mathbf{m}'(\mathbf{r}, t_0 + rT_R) \approx m_0(\mathbf{r})\hat{k}, \forall r \in \{0, 1, 2, \dots\}$) just prior to each RF excitation. This approximation yields a sequence of initial conditions and allows computation of the magnetization at corresponding times of data acquisition via direct application of Bloch dynamics (2.8) and (2.14). Resulting signal models are typically simple expressions of relaxation parameters $T_1(\mathbf{r})$ and $T_2(\mathbf{r})$; however, model accuracy often depends strongly on the long- T_R assumption, which requires long acquisitions.

Steady-state (SS) sequences [13] utilize short T_R , and can thus achieve much faster scan times. Due to short repetition times, SS sequences achieve only partial T_1 recovery in between RF excitations; thus, their magnetization responses do not obey the simple classical initial conditions (for the second RF cycle onwards). Although their transient magnetization dynamics can be complicated, SS sequences produce (under certain assumptions [14]) long-time magnetization responses that eventually¹ achieve a steady-state condition:

{eq:ss-cond}

$$\lim_{t_0 \rightarrow \infty} \mathbf{m}'(\mathbf{r}, t_0 + rT_R) = \mathbf{m}'(\mathbf{r}, t_0), \quad (2.17)$$

where repetition count $r \in \{1, 2, \dots\}$ for fixed RF excitations and off-resonance induced phase increments (as is assumed in the following). Subsections 2.1.2.1 and 2.1.2.2 use SS condition (2.17) and Bloch equation matrix operators introduced in (2.8) and (2.14) to derive long-time signal models for Spoiled Gradient-Recalled Echo (SPGR) and Dual-Echo Steady-State (DESS), two SS pulse sequences useful for quantitative MRI.

2.1.2.1 Spoiled Gradient-Recalled Echo (SPGR) Sequence

SPGR [16] is a fast pulse sequence that repeats cycles of fixed RF excitation (such that $b'_{1,x}(t + rT_R) = b'_{1,x}(t), \forall t \in [t_0, t_0 + T_P], r \in \{1, 2, \dots\}$); data acquisition; relaxation and recovery; and residual transverse magnetization “spoiling” (discussed later). Here we

¹The progression to steady state takes on the order of $5T_2/T_R$ RF cycles [14], typically a small but not insignificant period during which data acquisition is often foregone. This transition can (in some cases) be accelerated by prepending SS sequences with tailored “magnetization-catalyzing” modules [15].

{sss,bkgrd,mri,ss,spgr}

develop a simple and popular steady-state SPGR signal model.

Let $\mathbf{m}'(\mathbf{r}, t_0)$ denote the magnetization at an initial time t_0 selected well into the steady-state and just prior to excitation. The SPGR sequence first applies an RF excitation, which rotates the initial magnetization as per (2.8):

$$\{\text{eq:spgr-ex}\} \quad \mathbf{m}'(\mathbf{r}, t_0 + T_P) = \mathbf{R}_{x'}(\alpha(\mathbf{r}, t_0 + T_P; t_0))\mathbf{m}'(\mathbf{r}, t_0). \quad (2.18)$$

The excited magnetization then precesses and relaxes as per (2.14) until data acquisition, defined to occur at “echo time” $T_E \in [T_P, T_R]$ after the (midpoint of) RF excitation:

$$\{\text{eq:spgr-daq}\} \quad \begin{aligned} \mathbf{m}'\left(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E\right) &= \mathbf{R}_{z'}\left(\phi'\left(\mathbf{r}, \frac{T_P}{2} + T_E; T_P\right)\right)\mathbf{E}\left(\mathbf{r}, \frac{T_P}{2} + T_E; T_P\right)\mathbf{m}'(\mathbf{r}, t_0 + T_P) \\ &+ \mathbf{m}_0\left(\mathbf{r}, \frac{T_P}{2} + T_E; T_P\right). \end{aligned} \quad (2.19)$$

Following signal reception, the remaining transverse magnetization is spoiled² while the longitudinal component is unaffected. We model an ideal spoiling operation as

$$\{\text{eq:spgr-spoil}\} \quad \mathbf{S}\mathbf{m}'\left(\mathbf{r}, \frac{T_P}{2} + T_E\right), \text{ where } \mathbf{S} := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (2.20)$$

After spoiling, the longitudinal magnetization (partially) recovers until $t \leftarrow t_0 + T_R$:

$$\{\text{eq:spgr-pr}\} \quad \begin{aligned} \mathbf{m}'(\mathbf{r}, t_0 + T_R) &= \mathbf{R}_{z'}\left(\phi'\left(\mathbf{r}, T_R; \frac{T_P}{2} + T_E\right)\right)\mathbf{E}\left(\mathbf{r}, T_R; \frac{T_P}{2} + T_E\right)\mathbf{S}\mathbf{m}'\left(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E\right) \\ &+ \mathbf{m}_0\left(\mathbf{r}, T_R; \frac{T_P}{2} + T_E\right). \end{aligned} \quad (2.21)$$

In steady-state, one cycle of excitation, acquisition, spoiling, and recovery returns the magnetization back to its initial state. We enforce this through the steady-state condition

$$\{\text{eq:spgr-ss}\} \quad \mathbf{m}'(\mathbf{r}, t_0 + T_P) = \mathbf{R}_{x'}(\alpha(\mathbf{r}, t_0 + T_P; t_0))\mathbf{m}'(\mathbf{r}, t_0 + T_R) \quad (2.22)$$

²Transverse signal spoiling is often (nearly) achieved in practice using strong induced field inhomogeneities (which cause rapid transverse signal dephasing) in tandem with RF excitations that additionally impart nonlinear (often quadratically increasing) transverse magnetization phase [16]. Though the nonlinear RF phase used in so-called “RF-spoiling” prevents any one spin from reaching a true steady-state, the signal integrated over a typically-sized voxel achieves SS-like behavior [17].

which yields an algebraic system of equations. When it exists, the solution is

$$\mathbf{m}'(\mathbf{r}, t_0 + T_P) = \frac{1}{1 - e^{-(T_R - T_P)/T_1(\mathbf{r})} \cos(\alpha(\mathbf{r}))} \begin{bmatrix} 0 \\ m_0(\mathbf{r}) \sin(\alpha(\mathbf{r})) (1 - e^{-(T_R - T_P)/T_1(\mathbf{r})}) \\ m_0(\mathbf{r}) \cos(\alpha(\mathbf{r})) (1 - e^{-(T_R - T_P)/T_1(\mathbf{r})}) \end{bmatrix}, \quad (2.23)$$

{eq:spgr-bmmp-t0}

where $\alpha(\mathbf{r}) := \alpha(\mathbf{r}, t_0 + T_P; t_0)$ is a slight abuse of notation. Remarkably, the SPGR steady-state magnetization immediately after excitation is approximately independent of both off-resonance effects and $T_2(\mathbf{r})$. Researchers more often cite the expression

$$\begin{aligned} m'_{xy}(\mathbf{r}, t_0 + T_P) &= m'_x(\mathbf{r}, t_0 + T_P) + im'_y(\mathbf{r}, t_0 + T_P) \\ &= \frac{im_0(\mathbf{r}) \sin(\alpha(\mathbf{r})) (1 - e^{-T_R/T_1(\mathbf{r})})}{1 - e^{-T_R/T_1(\mathbf{r})} \cos(\alpha(\mathbf{r}))} \end{aligned} \quad (2.24)$$

{eq:spgr-mxyp-t0}

for the complex transverse magnetization as it modifies (2.23) to include a simple first-order correction for unaccounted T_1 recovery during the RF pulse. Substituting (2.24) into (2.19) yields an expression for the transverse magnetization at the echo time:

$$\begin{aligned} m'_{xy}\left(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E\right) &= m'_{xy}(\mathbf{r}, t_0 + T_P) e^{-(T_E - T_P/2)/T_2(\mathbf{r})} e^{-i\phi'(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E; t_0 + T_P)} \\ &\approx m'_{xy}(\mathbf{r}, t_0 + T_P) e^{-T_E/T_2(\mathbf{r})} e^{-i\phi'(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E; t_0 + \frac{T_P}{2})}, \end{aligned} \quad (2.25)$$

{eq:spgr-mxyp-te}

where the approximation again keeps in line with literature expressions.

The received signal is approximately proportional to the integrated transverse magnetization over a volume \mathbb{V} . To derive expressions, we take a few more usual assumptions:

- We assume that the signal is localized to a scale over which there is off-resonance phase variation, but minimal variation of $m_0(\mathbf{r})$, $T_1(\mathbf{r})$, $T_2(\mathbf{r})$, and $\alpha(\mathbf{r})$. This assumption is reasonable³ when describing the signal arising from a typical voxel.
- We assume that (free-precession) off-resonance phase grows linearly with time, *i.e.* $\phi'(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E; t_0 + \frac{T_P}{2}) \approx \omega'(\mathbf{r})T_E$. We further assume that off-resonance frequency $\omega'(\mathbf{r})$ is distributed over the localized voxel as $p_{\omega'} \leftarrow \text{Cauchy}(\bar{\omega}', R'_2)$, where $\bar{\omega}'(\mathbf{r})$ is the median off-resonance frequency and $R'_2(\mathbf{r})$ is the broadening bandwidth.

With these additional assumptions, the received steady-state SPGR (noiseless) signal model

³Model mismatch due to within-voxel spatial variation of relaxation parameters can be significant, especially for large voxels. Chapter 6 studies so-called partial volume effects and uses them for QMRI.

for a typically sized voxel centered at position \mathbf{r} is (to within constants):

$$\{eq:spgr-int\} \quad s_S\left(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E\right) \propto \int_{\mathbb{V}(\mathbf{r})} m'_{xy}\left(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E\right) d^3 \mathbf{r} \quad (2.26)$$

$$\begin{aligned} &\approx m_{xy}(\mathbf{r}, t_0 + T_P) e^{-T_E/T_2(\mathbf{r})} \int_{\mathbb{R}} e^{-i\omega' T_E} p_{\omega'}(\omega') d\omega' \\ &= m_{xy}(\mathbf{r}, t_0 + T_P) e^{-T_E/T_2(\mathbf{r})} e^{-R'_2(\mathbf{r}) T_E - i\bar{\omega}'(\mathbf{r}) T_E} \\ \{eq:spgr-model\} &= \frac{im_0(\mathbf{r}) \sin(\alpha(\mathbf{r})) (1 - e^{-T_R/T_1(\mathbf{r})})}{1 - e^{-T_R/T_1(\mathbf{r})} \cos(\alpha(\mathbf{r}))} e^{-T_E/T_2^*(\mathbf{r})} e^{-i\bar{\omega}'(\mathbf{r}) T_E}, \quad (2.27) \end{aligned}$$

where $T_2^*(\mathbf{r}) := \left(\frac{1}{T_2} + R'_2\right)^{-1}$ is a modified spin-spin relaxation time that accounts for additional transverse magnetization decay due to off-resonance effects.

2.1.2.2 Dual-Echo Steady-State (DESS) Sequence

DESS [18, 19] is a fast pulse sequence that interlaces fixed, constant-phase RF excitations with fixed dephasing “gradients” (*i.e.*, induced main field inhomogeneities that vary nearly linearly with space) to produce two distinct signals per RF excitation. Here we develop simple steady-state DESS signal models.

As in Subsection 2.1.2.1, let $\mathbf{m}'(\mathbf{r}, t_0)$ denote the magnetization at an initial time t_0 selected well into the steady-state and just prior to excitation. The DESS sequence first applies a fixed RF rotation $\alpha(\mathbf{r}) := \alpha(\mathbf{r}, t_0 + rT_R + T_P; t_0 + rT_R), \forall r \in \{0, 1, 2, \dots\}$:

$$\{eq:dess-ex\} \quad \mathbf{m}'(\mathbf{r}, t_0 + T_P) = \mathbf{R}_{x'}(\alpha(\mathbf{r})) \mathbf{m}'(\mathbf{r}, t_0). \quad (2.28)$$

The excited transverse magnetization contributes to a first acquired signal; dephases (but does not spoil completely) due to gradient dephasing⁴ and contributes again to a second (smaller, but nonzero) acquired signal. Since (with proper selection) dephasing gradients mainly contribute to off-resonance phase accrual, the net effect after data acquisition and gradient dephasing is well described simply by precession and relaxation:

$$\{eq:dess-pr\} \quad \mathbf{m}'(\mathbf{r}, t_0 + T_R) = \mathbf{R}_{z'}(\phi'(\mathbf{r})) \mathbf{E}(\mathbf{r}, T_R; T_P) \mathbf{m}'(\mathbf{r}, t_0 + T_P) + \mathbf{m}_0(\mathbf{r}, T_R; T_P), \quad (2.29)$$

where the abbreviation $\phi'(\mathbf{r}) := \phi'(\mathbf{r}, t_0 + (r+1)T_R; t_0 + rT_R + T_P), \forall r \in \{0, 1, 2, \dots\}$ implies fixed phase accrual (due to gradient dephasing, field inhomogeneity, and other

⁴It is worth distinguishing gradient dephasing (commonly but somewhat misleadingly referred to as gradient spoiling) from RF spoiling. Gradient dephasing (used in DESS) primarily affects magnetization phase and is modeled simply as precession. RF spoiling (used in SPGR) combines gradient dephasing with nonlinear RF phase cycling and suppresses magnetization magnitude in steady-state.

unaccounted effects) over each repetition cycle.

In steady-state, one cycle of excitation, first acquisition, gradient spoiling, second acquisition, and (partial) recovery returns the magnetization back to its initial state. We enforce this through the steady-state condition

$$\{\text{eq:dess-ss}\} \quad \mathbf{m}'(\mathbf{r}, t_0) = \mathbf{m}'(\mathbf{r}, t_0 + T_R) \quad (2.30)$$

which yields an algebraic system of equations. When it exists, the solution gives the steady-state magnetization just prior to RF excitation:

$$\{\text{eq:dess-bmmp-t0}\} \quad \mathbf{m}'(\mathbf{r}, t_0) = \begin{bmatrix} E_2(\mathbf{r}, T_F) \sin \alpha(\mathbf{r}) \sin \phi'(\mathbf{r}) \\ -E_2(\mathbf{r}, T_F) \sin \alpha(\mathbf{r}) (E_2(\mathbf{r}, T_F) - \cos \phi'(\mathbf{r})) \\ 1 - E_2(\mathbf{r}, T_F) \cos \phi'(\mathbf{r}) + E_2(\mathbf{r}, T_F) \cos \alpha(\mathbf{r}) (E_2(\mathbf{r}, T_F) - \cos \phi'(\mathbf{r})) \end{bmatrix} q(\mathbf{r}, T_F), \quad (2.31)$$

where $T_F := T_R - T_P$ is the free precession interval; $E_1(\mathbf{r}, t) := e^{-t/T_1(\mathbf{r})}$ and $E_2(\mathbf{r}, t) := e^{-t/T_2(\mathbf{r})}$ are relaxation operators; and $q(\mathbf{r}, t) :=$

$$\frac{m_0(\mathbf{r})(1 - E_1(\mathbf{r}, t))}{(1 - E_1(\mathbf{r}, t) \cos \alpha(\mathbf{r}))(1 - E_2(\mathbf{r}, t) \cos \phi'(\mathbf{r})) - E_2(\mathbf{r}, t)(E_1(\mathbf{r}, t) - \cos \alpha(\mathbf{r}))(E_2(\mathbf{r}, t) - \cos \phi'(\mathbf{r}))}.$$

Substituting (2.31) into (2.28) produces a similar expression for the steady-state magnetization immediately following RF excitation:

$$\{\text{eq:dess-bmmp-tp}\} \quad \mathbf{m}'(\mathbf{r}, t_0 + T_P) = \begin{bmatrix} E_2(\mathbf{r}, T_F) \sin \alpha(\mathbf{r}) \sin \phi'(\mathbf{r}) \\ \sin \alpha(\mathbf{r})(1 - E_2(\mathbf{r}, T_F) \cos \phi'(\mathbf{r})) \\ \cos \alpha(\mathbf{r})(1 - E_2(\mathbf{r}, T_F) \cos \phi'(\mathbf{r})) + E_2(\mathbf{r}, T_F)(E_2(\mathbf{r}, T_F) - \cos \phi'(\mathbf{r})) \end{bmatrix} q(\mathbf{r}, T_F). \quad (2.32)$$

The transverse magnetizations before and after RF excitation are then

$$\{\text{eq:dess-mxyp-t0}\} \quad m'_{xy}(\mathbf{r}, t_0) = -i \sin \alpha(\mathbf{r}) E_2(\mathbf{r}, T_R) (E_2(\mathbf{r}, T_R) - e^{-i\phi'(\mathbf{r})}) q(\mathbf{r}, T_R); \quad (2.33)$$

$$\{\text{eq:dess-mxyp-tp}\} \quad m'_{xy}(\mathbf{r}, t_0 + T_P) = +i \sin \alpha(\mathbf{r}) (1 - E_2(\mathbf{r}, T_R) e^{i\phi'(\mathbf{r})}) q(\mathbf{r}, T_R), \quad (2.34)$$

where (2.33)-(2.34) include simple first-order corrections for yet-unaccounted relaxation and recovery during excitation. Frequently, the DESS signals are acquired at symmetric echo times T_E before and after the center of each RF pulse. Substituting (2.34) into (2.9)

gives the magnetization at the data acquisition time after RF excitation:

$$m'_{xy}\left(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E\right) = m'_{xy}(\mathbf{r}, t_0 + T_P) e^{-(T_E - T_P/2)/T_2(\mathbf{r})} e^{-i\phi'(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E; t_0 + T_P)}$$

$$\approx m'_{xy}(\mathbf{r}, t_0 + T_P) e^{-T_E/T_2(\mathbf{r})} e^{-i\phi'(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E; t_0 + \frac{T_P}{2})} \quad (2.35)$$

{eq:dess-mxyp-te1-ph}
{eq:dess-mxyp-te1}

$$\approx m'_{xy}(\mathbf{r}, t_0 + T_P) e^{-T_E/T_2(\mathbf{r})} e^{-i\omega'(\mathbf{r})T_E}, \quad (2.36)$$

where in (2.35) we again approximately correct for relaxation during excitation and in (2.36) we assume linear off-resonance phase accrual during free precession. To compute the magnetization at the acquisition time before excitation, we consider the free precession and relaxation that occurs between⁵ signal reception and excitation:

$$m'_{xy}(\mathbf{r}, t_0) = m'_{xy}\left(\mathbf{r}, t_0 - \left(T_E - \frac{T_P}{2}\right)\right) e^{-(T_E - T_P/2)/T_2(\mathbf{r})} e^{-i\phi'(\mathbf{r}, t_0; t_0 - (T_E - \frac{T_P}{2}))}. \quad (2.37)$$

Rearranging (2.37) and applying approximations similar to those of (2.35)-(2.36),

$$m'_{xy}\left(\mathbf{r}, t_0 + \frac{T_P}{2} - T_E\right) = m'_{xy}(\mathbf{r}, t_0) e^{+(T_E - T_P/2)/T_2(\mathbf{r})} e^{+i\phi'(\mathbf{r}, t_0; t_0 - (T_E - \frac{T_P}{2}))}$$

$$\approx m'_{xy}(\mathbf{r}, t_0) e^{+T_E/T_2(\mathbf{r})} e^{+i\phi'(\mathbf{r}, t_0 + \frac{T_P}{2}; t_0 + \frac{T_P}{2} - T_E)} \quad (2.38)$$

{eq:dess-mxyp-te2-ph}
{eq:dess-mxyp-te2}

$$\approx m'_{xy}(\mathbf{r}, t_0) e^{+T_E/T_2(\mathbf{r})} e^{+i\omega'(\mathbf{r})T_E}. \quad (2.39)$$

The received signal is approximately proportional to the integrated transverse magnetization over a volume \mathbb{V} . To derive expressions, we retake assumptions used in Subsection 2.1.2.1 and append an additional assumption on the full-repetition phase accrual $\phi'(\mathbf{r})$:

- We assume that the signal is localized to a scale over which there is off-resonance phase variation, but minimal variation of $m_0(\mathbf{r})$, $T_1(\mathbf{r})$, $T_2(\mathbf{r})$, and $\alpha(\mathbf{r})$. This assumption is reasonable⁶ when describing the signal arising from a typical voxel.
- We assume that free precession off resonance frequency $\omega'(\mathbf{r})$ is distributed over the localized voxel as $p_{\omega'} \leftarrow \text{Cauchy}(\bar{\omega}', R'_2)$, where $\bar{\omega}'(\mathbf{r})$ is the median off-resonance frequency and $R'_2(\mathbf{r})$ is the broadening bandwidth.
- We assume that the dephasing gradient imparts an integral number n_{cyc} of across-

⁵Observe that we do not attempt to express the magnetization prior to (the next) RF excitation by simply operating on the magnetization after (the current) RF excitation with further precession and relaxation. The reason is due to the intermediate dephasing gradient, which causes phase accrual in excess of off-resonance effects and thus forbids an approximation akin to (2.36).

⁶Model mismatch due to within-voxel spatial variation of relaxation parameters can be significant, especially for large voxels. Chapter 6 studies so-called partial volume effects and uses them for QMRI.

voxel phase cycles⁷ such that full-repetition phase accrual $\phi'(\mathbf{r})$ is distributed essentially uniformly as $p_{\phi'} \leftarrow \text{uniform}(0, 2\pi n_{\text{cyc}}), n_{\text{cyc}} \in \{1, 2, 3, \dots\}$.

With these assumptions, the received steady-state DESS (noiseless) signal models for a typically sized voxel centered at position \mathbf{r} are (to within constants):

$$\begin{aligned}
 \text{\{eq:dess-def-int\}} \quad s_D\left(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E\right) &\propto \int_{\mathbb{V}(\mathbf{r})} m'_{xy}\left(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E\right) d^3 \mathbf{r} \\
 &\approx \int_{\mathbb{R}} \int_{\mathbb{R}} m'_{xy}\left(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E\right) p_{\phi'}(\phi') p_{\omega'}(\omega') d\phi' d\omega' \\
 &\approx e^{-T_E/T_2(\mathbf{r})} \int_{\mathbb{R}} m'_{xy}(\mathbf{r}, t_0 + T_P) p_{\phi'}(\phi') d\phi' \int_{\mathbb{R}} e^{-i\omega' T_E} p_{\omega'}(\omega') d\omega' \\
 &= +im_0(\mathbf{r}) E_2(\mathbf{r}, T_E) e^{-(R'_2(\mathbf{r}) - i\bar{\omega}'(\mathbf{r})) T_E} \tan \frac{\alpha(\mathbf{r})}{2} \left[1 - \frac{\eta(\mathbf{r}, T_R)}{\xi(\mathbf{r}, T_R)}\right]; \\
 \text{\{eq:dess-def-model\}} & \tag{2.41}
 \end{aligned}$$

$$\begin{aligned}
 \text{\{eq:dess-ref-int\}} \quad s_D\left(\mathbf{r}, t_0 + \frac{T_P}{2} - T_E\right) &\propto \int_{\mathbb{V}(\mathbf{r})} m'_{xy}\left(\mathbf{r}, t_0 + \frac{T_P}{2} - T_E\right) d^3 \mathbf{r} \\
 &\approx \int_{\mathbb{R}} \int_{\mathbb{R}} m'_{xy}\left(\mathbf{r}, t_0 + \frac{T_P}{2} - T_E\right) p_{\phi'}(\phi') p_{\omega'}(\omega') d\phi' d\omega' \\
 &\approx e^{+T_E/T_2(\mathbf{r})} \int_{\mathbb{R}} m'_{xy}(\mathbf{r}, t_0) p_{\phi'}(\phi') d\phi' \int_{\mathbb{R}} e^{+i\omega' T_E} p_{\omega'}(\omega') d\omega' \\
 &= -im_0(\mathbf{r}) E_2^{-1}(\mathbf{r}, T_E) e^{-(R'_2(\mathbf{r}) + i\bar{\omega}'(\mathbf{r})) T_E} \tan \frac{\alpha(\mathbf{r})}{2} [1 - \eta(\mathbf{r}, T_R)], \\
 \text{\{eq:dess-ref-model\}} & \tag{2.43}
 \end{aligned}$$

where (2.41) and (2.43) introduce intermediate variables

$$\begin{aligned}
 \eta(\mathbf{r}, t) &:= \sqrt{\frac{1 - E_2^2(\mathbf{r}, t)}{1 - E_2^2(\mathbf{r}, t)/\xi^2(\mathbf{r}, t)}}; \\
 \xi(\mathbf{r}, t) &:= \frac{1 - E_1(\mathbf{r}, t) \cos \alpha(\mathbf{r})}{E_1(\mathbf{r}, t) - \cos \alpha(\mathbf{r})}.
 \end{aligned}$$

In steady-state, the DESS signal is typically greatest immediately following excitation and defocuses with rate $\frac{1}{T_2} + R'_2$ until what we hereafter denote the *defocusing* echo time.

⁷In theory, it suffices to design dephasing gradients to impart as few as one complete cycle of net phase variation across a voxel. In practice, field inhomogeneities will induce spurious through-voxel field gradients that modify the effective dephasing gradient moment and thereby create partial phase cycles that distort the nominally uniform phase distribution. To reduce model mismatch due to such “partial spoiling” effects, dephasing gradients are usually designed to nominally impart multiple complete cycles of across-voxel phase variation. However, larger dephasing gradients cause greater DESS model mismatch due to diffusive signal loss. Appendix B studies diffusion in DESS and discusses regimes of dephasing gradient moments which balance partial-spoiling versus diffusive sources of model mismatch.

After a low-signal period between RF pulses, the DESS signal then refocuses with rate $\frac{1}{T_2} - R'_2$ from what we hereafter denote the *refocusing* echo time until just prior the next excitation. Fortuitously, the defocusing (2.41) and refocusing (2.43) DESS signal models have significantly different dependence on relaxation parameters (especially T_2) and thus together are quite useful for relaxation parameter estimation.

2.2 Optimization in QMRI

{s,bkgrd,opt}

This section overviews how optimization methods are leveraged in a substantial portion of this thesis to solve practical QMRI problems. For such problems, the central idea is to construct a suitable scalar cost function Ψ of some design variables \mathbf{x} , whose output $\Psi(\mathbf{x}) \in \mathbb{R}$ is designed to provide a measure of the undesirability of \mathbf{x} . We then employ tailored optimization algorithms to find an \mathbf{x} that minimizes Ψ over a set \mathbb{X} , written as

{eq:opt-global}

$$\mathbf{x}^* \in \left\{ \arg \min_{\mathbf{x} \in \mathbb{X}} \Psi(\mathbf{x}) \right\}. \quad (2.44)$$

In either optimization-based parameter estimation (Chapter 3) or acquisition design (Chapter 4), we have reason to design Ψ to depend on corresponding design variables \mathbf{x} through MR signal models. Because these models are often (strongly) nonlinear functions of design variables, corresponding cost functions are usually non-convex in \mathbf{x} (though the search space \mathbb{X} is almost always assumed convex in this thesis). Thus, most QMRI problems in the form of (2.44) are non-convex optimization problems.

In general, solving (2.44) is more challenging when Ψ is non-convex in \mathbf{x} than otherwise, due in part to the possible presence of local extrema and/or saddle points. In the following, we discuss two strategies used in this thesis to cope with non-convex optimization. Subsection 2.2.1 relaxes (2.44) to instead seek a local minimizer via iterative methods. Subsection 2.2.2 restricts attention to signal models that are linear in a portion of \mathbf{x} and discusses a specific problem for which (2.44) simplifies for such partially linear structures.

2.2.1 Iterative Local Optimization with Constraints

{ss,bkgrd,opt,loc}

This subsection overviews a method for finding a local minimizer $\hat{\mathbf{x}}$ of possibly non-convex cost function Ψ over convex constraint set \mathbb{X} . Such $\hat{\mathbf{x}} \in \mathbb{X}$ must satisfy for some $\delta > 0$

{eq:opt-local}

$$\Psi(\hat{\mathbf{x}}) \leq \Psi(\mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{X} : \|\hat{\mathbf{x}} - \mathbf{x}\|_2 < \delta. \quad (2.45)$$

Observe that a global optimizer \mathbf{x}^* satisfies (2.45) for arbitrarily large δ ; thus, any global minimizer is a local minimizer (but the converse is not necessarily true unless Ψ is convex).

As even locally optimal minimizers are often challenging to compute analytically, many algorithms construct $\hat{\mathbf{x}}$ by iteratively updating an initial guess $\mathbf{x}^{(0)}$ until some convergence criterion is satisfied. For a differentiable cost and convex constraints, the gradient projection method [20] is one such iterative algorithm and repeats the following simple update:

$$\{\text{eq:gpm}\} \quad \mathbf{x}^{(i)} \leftarrow \mathbf{P}_{\mathbb{X}}(\mathbf{x}^{(i-1)} - \mathbf{\Pi} \nabla_{\mathbf{x}} \Psi(\mathbf{x}^{(i-1)})), \quad (2.46)$$

where $\mathbf{P}_{\mathbb{X}}$ denotes projection onto \mathbb{X} ; $\nabla_{\mathbf{x}}$ denotes row gradient with respect to \mathbf{x} ; and $\mathbf{\Pi}$ is a diagonal preconditioning matrix that permits elements of \mathbf{x} to take scale-informed step sizes along the negative gradient direction.

If Ψ is convex and sufficiently smooth, iterates produced via (2.46) converge to a limit point [21] that is a constrained global minimum (for appropriately selected $\mathbf{\Pi}$). If instead Ψ is non-convex (but \mathbb{X} is still convex), statements regarding convergence⁸ to a particular constrained local minimizer require additional (strong) assumptions regarding initialization and in general are still much weaker than in the convex case.

Since non-convex cost functions can have many local extrema (whose associated costs can vary dramatically), the utility of a locally optimal solution depends strongly on initialization quality. Accordingly, this thesis uses iterative local optimization for non-convex QMRI problems where a reasonable initialization is available and global optimization (to within quantization error) via exhaustive grid search is intractable.

2.2.2 Partially Linear Models and the Variable Projection Method

(Constrained, weighted) nonlinear least-squares is a specific non-convex optimization problem that is useful for many parameter estimation problems:

$$\{\text{eq:nonlin-ls}\} \quad \mathbf{x}^* \in \left\{ \arg \min_{\mathbf{x} \in \mathbb{X}} \|\mathbf{y} - \mathbf{f}(\mathbf{x})\|_{\mathbf{W}}^2 \right\}, \quad (2.47)$$

where $\mathbf{f} : \mathbb{X} \mapsto \mathbb{C}^D$ is a nonlinear forward model that (barring noise) relates parameters $\mathbf{x} \in \mathbb{X} \subseteq \mathbb{C}^L$ to data $\mathbf{y} \in \mathbb{C}^D$; weighted 2-norm $\|\boldsymbol{\iota}\|_{\mathbf{W}} := \sqrt{\boldsymbol{\iota}^H \mathbf{W} \boldsymbol{\iota}}$ for a symmetric, positive-semidefinite weighting matrix $\mathbf{W} \in \mathbb{R}^{D \times D}$ and arbitrary vector $\boldsymbol{\iota} \in \mathbb{C}^D$; and $(\cdot)^H$

⁸For example, it suffices to assume that $\mathbf{x}^{(0)}$ lies in the *attraction basin* $\mathbb{B}_{\tilde{\mathbf{x}}}$ of a given unconstrained local minimum $\tilde{\mathbf{x}}$, where attraction basin is defined here as the largest convex set containing $\tilde{\mathbf{x}}$ over which Ψ is convex. If $\mathbb{B}_{\tilde{\mathbf{x}}} \cap \mathbb{X}$ is nonempty and step sizes within $\mathbf{\Pi}$ are small enough to contain iterates within $\mathbb{B}_{\tilde{\mathbf{x}}}$, then iterates converge to the limit point $\mathbf{P}_{\mathbb{X}}(\tilde{\mathbf{x}})$.

denotes conjugate transpose. The variable projection method [22] reduces the complexity of (2.47) when the forward model takes the partially linear structure $\mathbf{f}(\mathbf{x}) \equiv \mathbf{A}(\mathbf{x}_N)\mathbf{x}_L$ and the feasible set takes the partially unconstrained form $\mathbb{X} \equiv \mathbb{C}^{L_L} \times \mathbb{X}_N$, where $\mathbf{x}_L \in \mathbb{C}^{L_L}$; $\mathbf{x}_N \in \mathbb{X}_N$; and $\mathbf{A} : \mathbb{X}_N \mapsto \mathbb{C}^{D \times L_L}$ is a matrix function. These restrictions on (2.47) define a so-called separable least-squares problem:

$$\{\text{eq:sep-ls}\} \quad (\mathbf{x}_L^*, \mathbf{x}_N^*) \in \left\{ \arg \min_{\substack{\mathbf{x}_L \in \mathbb{C}^{L_L} \\ \mathbf{x}_N \in \mathbb{X}_N}} \|\mathbf{y} - \mathbf{A}(\mathbf{x}_N)\mathbf{x}_L\|_{\mathbf{W}}^2 \right\}. \quad (2.48)$$

The variable projection method simplifies (2.48) by exploiting the partially linear structure of \mathbf{f} to explicitly express the optimal \mathbf{x}_L^* as a function of any fixed $\mathbf{x}_N \in \mathbb{X}_N$:

$$\begin{aligned} \mathbf{x}_L^*(\mathbf{x}_N) &= \arg \min_{\mathbf{x}_L \in \mathbb{C}^{L_L}} \|\mathbf{y} - \mathbf{A}(\mathbf{x}_N)\mathbf{x}_L\|_{\mathbf{W}}^2 \\ &= (\mathbf{W}^{1/2} \mathbf{A}(\mathbf{x}_N))^\dagger \mathbf{W}^{1/2} \mathbf{y} \end{aligned} \quad (2.49)$$

$$\begin{aligned} &= (\mathbf{A}^H(\mathbf{x}_N) \mathbf{W} \mathbf{A}(\mathbf{x}_N))^{-1} \mathbf{A}^H(\mathbf{x}_N) \mathbf{W} \mathbf{y}, \end{aligned} \quad (2.50)$$

$\{\text{eq:sep-ls-lin}\}$
 $\{\text{eq:sep-ls-fullrnk}\}$

where $(\cdot)^\dagger$ denotes pseudoinverse; $\mathbf{W}^{1/2}$ denotes principle (matrix) square root; and (2.50) holds if the matrix inversion within exists. Substituting (2.50) into (2.48) yields a new non-convex optimization problem that contains L_L fewer unknowns than before:

$$\begin{aligned} \mathbf{x}_N^* &\in \left\{ \arg \min_{\mathbf{x}_N \in \mathbb{X}_N} \left\| \mathbf{y} - \mathbf{A}(\mathbf{x}_N) (\mathbf{A}^H(\mathbf{x}_N) \mathbf{W} \mathbf{A}(\mathbf{x}_N))^{-1} \mathbf{A}^H(\mathbf{x}_N) \mathbf{W} \mathbf{y} \right\|_{\mathbf{W}}^2 \right\} \\ &\equiv \left\{ \arg \max_{\mathbf{x}_N \in \mathbb{X}_N} \mathbf{y}^H \mathbf{W} \mathbf{A}(\mathbf{x}_N) (\mathbf{A}^H(\mathbf{x}_N) \mathbf{W} \mathbf{A}(\mathbf{x}_N))^{-1} \mathbf{A}^H(\mathbf{x}_N) \mathbf{W} \mathbf{y} \right\}, \end{aligned} \quad (2.51)$$

$\{\text{eq:sep-ls-nonlin}\}$

where the equivalence leading to (2.51) omits terms independent of \mathbf{x}_N .

In low-dimensional QMRI applications (*e.g.*, those discussed in Chapter 3), reduced problem (2.51) may be tractable via exhaustive grid search, in which case a global optimum $(\mathbf{x}_L^*(\mathbf{x}_N^*), \mathbf{x}_N^*)$ is achievable to within quantization error. However, larger estimation problems involving more nonlinear parameters might still be tractable only via iterative optimization (see Subsection 2.2.1) towards a local solution $(\hat{\mathbf{x}}_L(\hat{\mathbf{x}}_N), \hat{\mathbf{x}}_N)$. For such higher-dimensional applications, Chapters 5-6 introduce novel methods that tackle problems similar to (2.47), while circumventing initialization-dependent local optimization.

CHAPTER 3

MRI Parameter Estimation from Likelihood Models

{c,relax}

3.1 Introduction

{s,relax,intro}

todo: connect with ss,bkgrd,opt,vpm; more motivation; see jfn 2017-01-18

This chapter introduces methods for QMRI parameter estimation from statistical likelihood models and applies these methods to simple problems in T_1 and T_2 relaxometry, which are of interest for monitoring the progression of various disorders [23]. Section 3.2 introduces the notion of a QMRI scan profile, describes a signal model for parameter estimation, formulates several likelihood-based estimators using this model, and discuss practical implementation issues. Section 3.3 demonstrates the utility of likelihood-based parameter estimation over conventional methods through simulation, phantom, and *in vivo* experiments. Section 3.4 provides brief concluding remarks and suggests future directions.

3.2 Likelihood-Based Estimation in QMRI

{s,relax,meth}

3.2.1 The QMRI Scan Profile

{ss,relax,meth,prof}

After image reconstruction, many MRI pulse sequences useful for parameter estimation produce at each voxel centered at position \mathbf{r} a set of noisy voxel values $\{y_1(\mathbf{r}), \dots, y_D(\mathbf{r})\}$, each of which can be described with the following general model:

{eq:relax,mod-scalar}

$$y_d(\mathbf{r}) = s_d(\mathbf{x}(\mathbf{r}); \boldsymbol{\nu}(\mathbf{r}), \mathbf{p}_d) + \epsilon_d(\mathbf{r}), \quad (3.1)$$

where $d \in \{1, \dots, D\}$. Here, $\mathbf{x}(\mathbf{r}) \in \mathbb{C}^L$ collects L *latent* object parameters at \mathbf{r} ; $\boldsymbol{\nu}(\mathbf{r}) \in \mathbb{C}^K$ collects K *known* object parameters at \mathbf{r} ; $s_d : \mathbb{C}^L \times \mathbb{C}^K \times \mathbb{R}^A \mapsto \mathbb{C}$ is a (pulse-sequence dependent) function that models the noiseless signal obtained from the d th dataset

using *acquisition* parameter $\mathbf{p}_d \in \mathbb{R}^A$; and $\epsilon_d \sim \mathbb{CN}(0, \sigma_d^2)$ is assumed for simplicity¹ to be (circularly-symmetric) complex Gaussian noise [26, 27] with zero mean and variance σ_d^2 . Colon positions in signal model (3.1) and similar expressions throughout this thesis distinguish unknown and known parameters. Concrete examples follow shortly.

For accurate, well-conditioned QMRI parameter estimation, it is typically necessary to acquire a collection of datasets, which we refer to hereafter as a *scan profile*. A scan profile consists of D datasets from up to D pulse sequences (some sequences yield more than one dataset, *e.g.* DESS). Let $\mathbf{y}(\mathbf{r}) := [y_1(\mathbf{r}), \dots, y_D(\mathbf{r})]^\top \in \mathbb{C}^D$ collect noisy voxel values centered at \mathbf{r} from a given scan profile. Then the vector signal model

{eq:relax,mod-vec}

$$\mathbf{y}(\mathbf{r}) = \mathbf{s}(\mathbf{x}(\mathbf{r}); \boldsymbol{\nu}(\mathbf{r}), \mathbf{P}) + \boldsymbol{\epsilon}(\mathbf{r}) \quad (3.2)$$

helps define the noiseless signal $\mathbf{s} := [s_1, \dots, s_D]^\top : \mathbb{C}^L \times \mathbb{C}^K \times \mathbb{R}^{A \times D} \mapsto \mathbb{C}^D$ and acquisition parameter $\mathbf{P} := [\mathbf{p}_1, \dots, \mathbf{p}_D] \in \mathbb{R}^{A \times D}$ associated with that scan profile. Here, noise $\boldsymbol{\epsilon}(\mathbf{r}) := [\epsilon_1(\mathbf{r}), \dots, \epsilon_D(\mathbf{r})]^\top \in \mathbb{C}^D$ typically has diagonal covariance structure $\boldsymbol{\Sigma} := \text{diag}([\sigma_1, \dots, \sigma_D]^\top)$ due to independence across datasets, where $\text{diag}(\cdot)$ assigns its argument to the diagonal entries of an otherwise zero (square) matrix.

The following subsections describe two concrete scan profiles whose signals can be modeled via (3.2) and that we study through experiments later in this chapter.

3.2.1.1 Example: An SPGR Scan Profile for T_1 estimation

We first consider the problem of $T_1(\mathbf{r})$ estimation at \mathbf{r} from as few SPGR scans as possible, given a prior estimate of transmit field variation $s^t(\mathbf{r})$ (see (2.7)). Examining SPGR model (2.27) makes clear that by fixing echo time T_E across scans, SPGR signal dependence is reduced to just two spatially varying latent parameters: desired parameter $T_1(\mathbf{r}) \in \mathbb{R}$ and nuisance parameter $c_1(\mathbf{r}) := im_0(\mathbf{r})e^{-T_E/T_2^*(\mathbf{r})}e^{-i\bar{\omega}'(\mathbf{r})T_E} \in \mathbb{C}$. We assign $\mathbf{x} \leftarrow [T_1, c_1]^\top$ and $\boldsymbol{\nu} \leftarrow s^t$ for $L \leftarrow 2$ latent and $K \leftarrow 1$ known parameters, respectively.

With T_E fixed, prescribed flip angles α_0 and repetition times T_R are the only remaining $A \leftarrow 2$ acquisition parameters available to choose that appear explicitly in (2.27). Thus, an SPGR scan profile useful for T_1 estimation must vary $\mathbf{p}_d \leftarrow [\alpha_0, T_R]^\top \forall d \in \{1, \dots, D\}$ over S_{SPGR} scan repetitions to produce $D \geq L \leftarrow 2$ datasets for well-conditioned estimation.

¹Though the noise distribution of \mathbf{k} -space raw data is usually well-modeled as complex white Gaussian, the noise distribution of the d th reconstructed image y_d depends both on the acquisition and reconstruction. If single receive channel \mathbf{k} -space data is fully-sampled on a Cartesian grid, each dataset y_d is recoverable via separate Fourier transform, and is thus complex Gaussian and independent across datasets. However if \mathbf{k} -space data is multi-channel, undersampled, and/or Cartesian, it may be preferable that y_d be estimated by more sophisticated techniques, *e.g.* [24, 25]. In such cases, reconstructed image noise is unlikely to be Gaussian-distributed.

3.2.1.2 Example: A DESS Scan Profile for T_2 estimation

We next consider the problem of $T_2(\mathbf{r})$ estimation at \mathbf{r} from as few DESS scans as possible. Examining DESS models (2.41) and (2.43) makes clear that even with fixed T_E over possibly several acquisitions, there is signal dependence on five distinct object parameters: $s^t(\mathbf{r}) \in \mathbb{R}$, $T_1(\mathbf{r}) \in \mathbb{R}$, $\bar{\omega}'(\mathbf{r}) \in \mathbb{R}$, $c_2(\mathbf{r}) := m_0(\mathbf{r})e^{-T_E/T_2^*(\mathbf{r})} \in \mathbb{C}$, and $T_2(\mathbf{r}) \in \mathbb{R}$. In this chapter, we take $s^t(\mathbf{r}) \in \mathbb{R}$ and $T_1(\mathbf{r}) \in \mathbb{R}$ as known for simplicity. To avoid (separate or joint) $\bar{\omega}'(\mathbf{r})$ estimation, we choose to use magnitude DESS data, at the expense of slight model mismatch² due to Rician noise. These choices assign $\boldsymbol{\nu} \leftarrow [s^t, T_1]^\top$ as $K \leftarrow 2$ known parameters and leave $L \leftarrow 2$ latent parameters $\mathbf{x} \leftarrow [T_2, c_2]^\top$ to be estimated.

With T_E again fixed, $\mathbf{p}_d \leftarrow [\alpha_0, T_R]^\top \forall d \in \{1, \dots, D\}$ collects the remaining $A \leftarrow 2$ tunable scan parameters that appear explicitly in (2.41) and (2.43). As in Example 3.2.1.1, $D \geq L \leftarrow 2$ datasets are necessary for well-conditioned estimation. Unlike before however, a minimum $D \leftarrow 2$ datasets need not require scan repetition, since S_{DESS} DESS scan repetitions produce $D \leftarrow 2S_{\text{DESS}}$ datasets.

3.2.2 Latent Object Parameter Estimation

3.2.2.1 Signal Model and Problem Statement

A scan profile's reconstructed images can be modeled to discretize the bulk MR signal into V localized voxels centered at positions $\mathbf{r}_1, \dots, \mathbf{r}_V$:

$$\mathbf{Y} = \mathbf{S}(\mathbf{X}; \mathbf{N}, \mathbf{P}) + \mathbf{E}. \quad (3.3)$$

Here, signal model $\mathbf{S} : \mathbb{C}^{L \times V} \times \mathbb{C}^{K \times V} \times \mathbb{R}^{A \times D} \mapsto \mathbb{C}^{D \times V}$ is a matrix function that maps latent $\mathbf{X} := [\mathbf{x}(\mathbf{r}_1), \dots, \mathbf{x}(\mathbf{r}_V)] \in \mathbb{C}^{L \times V}$ and known $\mathbf{N} := [\boldsymbol{\nu}(\mathbf{r}_1), \dots, \boldsymbol{\nu}(\mathbf{r}_V)] \in \mathbb{C}^{L \times V}$ parameter images (with fixed acquisition parameter \mathbf{P}) to reconstructed image data $\mathbf{Y} := [\mathbf{y}(\mathbf{r}_1), \dots, \mathbf{y}(\mathbf{r}_V)] \in \mathbb{C}^{D \times V}$, save for noise image $\mathbf{E} := [\boldsymbol{\epsilon}(\mathbf{r}_1), \dots, \boldsymbol{\epsilon}(\mathbf{r}_V)] \in \mathbb{C}^{D \times V}$. The goal in QMRI parameter estimation is to estimate latent parameter images \mathbf{X} from MR image data \mathbf{Y} , for a fixed scan profile defined by \mathbf{S} and \mathbf{P} and given (separately acquired, estimated, and here assumed) known parameter images \mathbf{N} .

²The assumption of complex Gaussian noise in noisy MRI images implies that corresponding magnitude MRI images are Rician-distributed. However, the statistical estimators we will develop in Subsection 3.2.2 are based on Gaussian data. Fortunately, this source of model mismatch is negligible (less than 1%) for signal-to-noise ratio (SNR) in excess of 10 [28], and the acquisitions we examine here are capable of producing SNR in tissue of at minimum 100 and usually more.

3.2.2.2 Maximum Likelihood Methods

In maximum likelihood (ML) estimation, one seeks to find model parameters that maximize the likelihood of observing output data. We apply ML estimation to QMRI by first constructing a *likelihood function* that describes the probability of observing image data \mathbf{Y} given latent parameters \mathbf{X} . We then formulate ML latent parameter estimate $\hat{\mathbf{X}}_{\text{ML}}(\mathbf{Y}; \mathbf{N}, \mathbf{P})$ by finding an \mathbf{X} that maximizes this likelihood function.

We first construct the likelihood function for the v th voxel's data $\mathbf{y}(\mathbf{r}_v)$ and latent parameter $\mathbf{x}(\mathbf{r}_v)$. For complex Gaussian noise, the likelihood function is

$$\mathcal{L}(\mathbf{x}(\mathbf{r}_v)) \propto \exp \left(-\|\mathbf{y}(\mathbf{r}_v) - \mathbf{s}(\mathbf{x}(\mathbf{r}_v); \boldsymbol{\nu}(\mathbf{r}_v), \mathbf{P})\|_{\boldsymbol{\Sigma}^{-1}}^2 \right), \quad (3.4)$$

where (3.4) omits constants that are independent of $\mathbf{x}(\mathbf{r}_v)$ and are therefore irrelevant. Assuming noise independence across image voxels, we can next build a simple and practical likelihood function of the full image data as

$$\mathcal{L}(\mathbf{X}) = \prod_{v=1}^V \mathcal{L}(\mathbf{x}(\mathbf{r}_v)). \quad (3.5)$$

We form an ML parameter estimate by finding \mathbf{X} that maximizes this likelihood function:

$$\begin{aligned} \hat{\mathbf{X}}_{\text{ML}}(\mathbf{Y}; \mathbf{N}, \mathbf{P}) &\in \left\{ \arg \max_{\mathbf{X} \in \mathbb{X}^V} \mathcal{L}(\mathbf{X}) \right\} \\ &\equiv \left\{ \arg \min_{\mathbf{X} \in \mathbb{X}^V} -\log \mathcal{L}(\mathbf{X}) \right\} \end{aligned} \quad (3.6)$$

$$\begin{aligned} &= \left\{ \arg \min_{\mathbf{X} \in \mathbb{X}^V} \sum_{v=1}^V \|\mathbf{y}(\mathbf{r}_v) - \mathbf{s}(\mathbf{x}(\mathbf{r}_v); \boldsymbol{\nu}(\mathbf{r}_v), \mathbf{P})\|_{\boldsymbol{\Sigma}^{-1}}^2 \right\} \\ &= \left\{ \arg \min_{\mathbf{X} \in \mathbb{X}^V} \left\| \boldsymbol{\Sigma}^{-1/2} (\mathbf{Y} - \mathbf{S}(\mathbf{X}; \mathbf{N}, \mathbf{P})) \right\|_{\text{F}}^2 \right\}, \end{aligned} \quad (3.7)$$

where \mathbb{X} is a (typically convex) latent parameter search space; the set equivalence in (3.6) uses the monotonicity of the log function; and $\|\cdot\|_{\text{F}}$ denotes the Frobenius matrix norm.

Typically, QMR image model \mathbf{S} is nonlinear in \mathbf{X} and so ML estimation problem (3.7) involves non-convex optimization, which is challenging in general (see Section 2.2). Two properties of (3.7) guide our solution strategies. First, (3.7) is separable across voxels, so problem non-convexity is addressable on a voxel-by-voxel basis. Second, MR signal models are usually partially linear, in which case we may employ the variable projection method (described in Section 2.2.2) to further reduce problem complexity. For applications studied in this chapter, these properties allow for (3.7) to be solved via simple grid search.

3.2.2.3 Regularized Likelihood Methods

In regularized likelihood (RL) estimation, we modify ML estimation problem (3.6) to include additional information in the form of *regularization*:

$$\widehat{\mathbf{X}}_{\text{RL}}(\mathbf{Y}; \mathbf{N}, \mathbf{P}) \in \left\{ \arg \min_{\mathbf{X} \in \mathbb{X}^V} -\log L(\mathbf{X}) + R(\mathbf{X}) \right\}. \quad (3.8)$$

Here, we have freedom to design regularizer $R : \mathbb{C}^{L \times V} \mapsto \mathbb{R}$ to encourage desirable structure in estimates of \mathbf{X} . We observe that it is usually reasonable to assume that each latent object parameter map is *piecewise smooth* as a function of space: that is, each parameter is likely to vary smoothly in space, except for sharp discontinuities at tissue boundaries. To encourage piecewise-smoothness in parameter estimates, we use the regularizer

$$R(\mathbf{X}) := \sum_{l=1}^L \beta_l \sum_{j=1}^J \phi_l \left([\mathbf{J}\mathbf{X}^\top]_{jl} \right), \text{ where} \quad (3.9)$$

$$\phi_l(\cdot) := \gamma_l^2 \left(\sqrt{1 + |\cdot/\gamma_l|^2} - 1 \right) \quad (3.10)$$

is a differentiable approximation of the absolute value function; $\mathbf{J} \in \mathbb{R}^{J \times V}$ evaluates J (multi-dimensional) finite-differencing operations; $[\cdot]_{jl}$ extracts the (j, l) th matrix element; and β_l is a regularization parameter that controls the relative importance of smoothing the l th latent object parameter image. Conceptually, this regularizer penalizes inconsistencies in adjacent latent parameter image voxels, but with a severity that depends on the degree of inconsistency. “Small” voxel-to-voxel differences are likely due to image data noise within a single tissue type and are penalized near-quadratically, while “large” differences are likely due to tissue boundaries and are penalized near-linearly. Useful notions of small versus large differences are governed by shape parameters $\gamma_l \forall l \in \{1, \dots, L\}$, and vary for different latent parameter maps based on their units and relative scale.

In general, QMRI image signal model \mathbf{S} is nonlinear in \mathbf{X} and so RL estimation problem (3.8) requires non-convex optimization. Unlike in ML estimation, (3.8) is not separable across voxels due to regularization, precluding global optimization (via grid search or other methods). We instead take the corresponding ML estimate as initialization and solve (3.8) via iterative constrained local optimization (detailed in Section 2.2.1).

3.3 Experimentation

{s,relax,exp}

3.4 Summary

{s,relax,summ}

CHAPTER 4

Optimizing MR Scan Design for Model-Based Relaxometry

{c,scn-dsgn}

4.1 Introduction

{s,scn-dsgn,intro}

Fast, accurate *relaxometry*, or quantification of spin-lattice and spin-spin relaxation parameters T_1 and T_2 has been of longstanding interest in MRI. Many researchers have suggested that T_1, T_2 “maps” (*i.e.*, estimated parameter images) may serve as biomarkers for monitoring the progression of various disorders [23]. Neurological applications include: lesion classification in multiple sclerosis [29]; tumor characterization [30, 31]; and symptom onset prediction in stroke [32, 33]. In addition, T_1, T_2 have shown promise for detecting hip and knee cartilage degeneration [34, 35] and for assessing cardiac dysfunction due to iron overload [36] or edema [37]. Motivated by this broad interest in T_1, T_2 mapping, this chapter describes a systematic method to guide QMRI scan design.

Classical pulse sequences such as inversion/saturation recovery (IR/SR) or (single) spin echo (SE) yield relatively simple methods for T_1 or T_2 estimation, respectively; however, these methods require several scans, each with long repetition time T_R , leading to undesirably long acquisitions. Numerous modifications such as the Look-Locker method [38], multi-SE trains [39], or fast k-space trajectories [40, 41, 42] have been proposed to accelerate T_1 [43, 44, 45, 46] and T_2 [47, 48, 49, 50] relaxometry with these classical sequences. These techniques are more sensitive to model non-idealities [51, 52, 53], and are still speed-limited by the long T_R required for (near)-complete T_1 recovery.

Steady-state (SS) pulse sequences [13, 14] permit short T_R , and are thus inherently much faster than classical counterparts. SS techniques are well-suited for relaxometry because the signals produced are highly sensitive to T_1 and T_2 variation. However, short T_R times also cause SS signals to be complex functions of both desired and undesired (*nuisance*) parameters, complicating quantification. Furthermore, some such methods [54, 55] still require scan repetition, though individual scans are now considerably shorter. Despite

these difficulties, the potential for rapid scanning with high T_1 , T_2 sensitivity has motivated numerous SS relaxometry studies [56, 54, 55, 57, 58, 59, 60, 61, 62, 63].

The dual-echo steady-state (DESS) sequence [19] was recently proposed as a promising SS imaging technique for T_2 estimation [60]. Because it produces two distinct signals per excitation, the DESS sequence can reduce scan repetition requirements by recording twice as much data per scan. As with other SS methods, the resulting signals [64, 65] are complicated functions of T_1 , T_2 , and other parameters (see Section 2.1.2.2 for derivations). Prior works have isolated T_2 dependencies using either algebraic manipulations of the first- and second-echo signals [60, 61] or separate scans to first estimate nuisance parameters [4]. Although DESS concurrently encodes rich T_1 and T_2 information, these methods have shied away from using DESS for T_1 estimation, either through bias-inducing approximations, or noise-propagating sequential estimation, respectively.

Whether it be with DESS, other sequences, or even combinations thereof, it is generally unclear how to best assemble a *scan profile* (*i.e.*, a collection of scans) for a fixed amount of scan time. Furthermore, for a given scan profile, it is typically not obvious how to best select acquisition parameters (*e.g.*, flip angles, repetition times, etc.) for relaxometry. In this and subsequent chapters, the term *scan design* refers to the related problems of scan profile selection and acquisition parameter optimization.

Historically, scan design for relaxometry has mainly been explored using figures of merit related to estimator precision. In particular, several studies have used the Cramér-Rao Bound (CRB), a statistical tool that bounds the minimum variance of an unbiased estimator. Earlier works have used the CRB and variations to select inversion times for recovery experiments [66, 67], flip angles for spoiled gradient-recalled echo (SPGR) sequences [68], and echo times for SE experiments [69]. More recent studies have considered additional scan design challenges, including scan time constraints [70], multiple latent parameters [71], multiple scan parameter types [72], and latent parameter spatial variation [73, 74].

The aforementioned studies consider scan parameter optimization for profiles consisting of *only one* pulse sequence. In contrast, this chapter introduces a general framework for robust, application-specific scan design for parameter estimation from *combinations* of pulse sequences. The framework first finds multiple sets of scan parameters that achieve precise estimation within a tight, *application-specific* range of object parameters (*e.g.*, T_1 , T_2 , etc.). The framework then chooses the one scan parameter set most *robust* to estimator precision degradation over a broader range of object parameters. As a detailed example, we optimize three combinations of SPGR and DESS sequences for T_1 , T_2 mapping. For a fixed total scan time, we find that well-chosen DESS scans alone can be used to estimate both T_1 and T_2 with precision and robustness comparable to combinations of SPGR

and DESS. This example illustrates that, with careful scan profile design, well-established pulse sequences can find use in new estimation problems.

This chapter is organized as follows. Section 4.2 describes a CRB-inspired min-max optimization problem for robust, application-specific scan optimization. Section 4.3 optimizes three practical DESS/SPGR combinations to show that, even in the presence of radiofrequency (RF) field inhomogeneity, DESS is a promising option for T_1, T_2 relaxometry. Section 4.4 describes simulation, phantom, and *in vivo* experiments and discusses corresponding results. Section 4.5 discusses practical challenges and suggests future directions. Section 4.6 summarizes key contributions.

4.2 A CRB-Inspired Scan Selection Method

{s,scn-dsgn,crb}

4.2.1 The CRB and its Relevance to QMRI

Recall from Section 3.2.1 that after image reconstruction, we can model the single-voxel MR image domain data associated with a particular scan profile as

$$\mathbf{y} = \mathbf{s}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) + \boldsymbol{\epsilon}, \quad (4.1)$$

where signal model $\mathbf{s} := [s_1, \dots, s_D]^T : \mathbb{C}^L \times \mathbb{C}^K \times \mathbb{R}^{A \times D} \mapsto \mathbb{C}^D$ relates latent $\mathbf{x} \in \mathbb{C}^L$, known $\boldsymbol{\nu} \in \mathbb{C}^K$, and acquisition $\mathbf{P} \in \mathbb{R}^{A \times D}$ parameters to noisy scan profile image data $\mathbf{y} \in \mathbb{C}^D$, barring noise $\boldsymbol{\epsilon} \in \mathbb{C}^D$. Assuming (as in Section 3.2.1) complex Gaussian noise $\boldsymbol{\epsilon} \sim \mathbb{CN}(\mathbf{0}, \boldsymbol{\Sigma})$, the likelihood function (3.4) is (to within constants independent of \mathbf{x})

$$L(\mathbf{x}|\mathbf{y}) \propto \exp \left(-\|\mathbf{y} - \mathbf{s}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P})\|_{\boldsymbol{\Sigma}^{-1}}^2 \right). \quad (4.2)$$

Under suitable regularity conditions¹, the Fisher information matrix $\mathbf{F}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) \in \mathbb{C}^{L \times L}$ [75] characterizes the imprecision of unbiased estimates of \mathbf{x} from \mathbf{y} , given $\boldsymbol{\nu}$ and \mathbf{P} :

$$\begin{aligned} \mathbf{F}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) &:= \mathbb{E}_{\mathbf{y}} \left((\nabla_{\mathbf{x}} \log L(\mathbf{x}|\mathbf{y}))^H \nabla_{\mathbf{x}} \log L(\mathbf{x}|\mathbf{y}) \right) \\ &= (\nabla_{\mathbf{x}} \mathbf{s}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}))^H \boldsymbol{\Sigma}^{-1} \nabla_{\mathbf{x}} \mathbf{s}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}), \end{aligned} \quad (4.3)$$

¹In particular, \mathbf{s} must be analytic in complex components of \mathbf{x} .

where $E_y(\cdot)$ denotes element-wise expectation with respect to y . In particular, the matrix CRB [76] ensures that any unbiased estimator $\hat{\mathbf{x}}$ satisfies

$$\text{cov}(\hat{\mathbf{x}}; \boldsymbol{\nu}, \mathbf{P}) \succeq \mathbf{F}^{-1}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}), \quad (4.4)$$

where for arbitrary, equally-sized \mathbf{C}_1 and \mathbf{C}_2 , matrix inequality $\mathbf{C}_1 \succeq \mathbf{C}_2$ means $\mathbf{C}_1 - \mathbf{C}_2$ is positive semi-definite. In the following, we design an optimization problem based on the CRB to guide QMRI scan design for relaxometry.

4.2.2 Min-max Optimization Problem for Scan Design

Following [77], we focus on minimizing a weighted average of the variances in each of the L latent object parameter estimates. A reasonable objective function for overall estimator precision is therefore given by

$$\Psi(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) = \text{tr}(\mathbf{W}\mathbf{F}^{-1}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P})\mathbf{W}^T), \quad (4.5)$$

where $\mathbf{W} \in \mathbb{R}^{L \times L}$ is a diagonal, application-specific matrix of weights, preselected to control the relative importance of precisely estimating the L latent object parameters. For scan design, we would like to minimize (4.5) with respect to scan parameters \mathbf{P} .

The CRB depends not only on \mathbf{P} but also on the spatially varying object parameters \mathbf{x} and $\boldsymbol{\nu}$. Thus, one cannot perform scan design by “simply” minimizing Ψ with respect to scan parameters \mathbf{P} . Instead, we pose a *min-max* optimization problem for scan design: we seek candidate scan parameters $\check{\mathbf{P}}$ over a search space \mathbb{P} that *minimize* the worst-case (*i.e.*, *maximum*) cost $\tilde{\Psi}^t$, as viewed over “tight” object parameter ranges \mathbb{X}^t and \mathbb{N}^t :

$$\check{\mathbf{P}} \in \left\{ \arg \min_{\mathbf{P} \in \mathbb{P}} \tilde{\Psi}^t(\mathbf{P}) \right\}, \text{ where} \quad (4.6)$$

$$\tilde{\Psi}^t(\mathbf{P}) := \max_{\substack{\mathbf{x} \in \mathbb{X}^t \\ \boldsymbol{\nu} \in \mathbb{N}^t}} \Psi(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}). \quad (4.7)$$

Here, we select *latent* parameter set \mathbb{X}^t based on the application and *known* parameter set \mathbb{N}^t based on the spatial variation typically observed in the known parameters $\boldsymbol{\nu}$. Min-max approach (4.9) should ensure good estimation precision over a range of parameter values.

Since Ψ is in general non-convex with respect to \mathbf{P} , it may have multiple global minimizers as well as other scan parameters that are nearly global minimizers. To improve robustness to object parameter variations, we form an expanded set of candidate scan parameters by also including scan parameters that yield costs to within a tolerance $\delta \ll 1$

of the optimum. Mathematically, we define this expanded set of candidate scan parameter combinations (for a given scan profile) as

$$\{\text{eq:scn-dsgn,set}\} \quad \check{\mathbb{P}} := \left\{ \mathbf{P} : \tilde{\Psi}^t(\mathbf{P}) - \tilde{\Psi}^t(\check{\mathbf{P}}) \leq \delta \tilde{\Psi}^t(\check{\mathbf{P}}) \right\}. \quad (4.8)$$

To select amongst these candidate scan parameters, we employ a robustness criterion: we select the single scan parameter $\hat{\mathbf{P}}$ that degrades the least when the worst-case cost is viewed over widened object parameter sets $\mathbb{X}^b \supseteq \mathbb{X}^t$ and $\mathbb{N}^b \supseteq \mathbb{N}^t$:

$$\{\text{eq:scn-dsgn,P-star}\} \quad \hat{\mathbf{P}} = \arg \min_{\mathbf{P} \in \check{\mathbb{P}}} \tilde{\Psi}^b(\mathbf{P}), \text{ where} \quad (4.9)$$

$$\{\text{eq:scn-dsgn,cost-broad}\} \quad \tilde{\Psi}^b(\mathbf{P}) := \max_{\substack{\mathbf{x} \in \mathbb{X}^b \\ \boldsymbol{\nu} \in \mathbb{N}^b}} \Psi(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}). \quad (4.10)$$

To compare different scan profiles, we select corresponding search spaces \mathbb{P} to satisfy acquisition constraints (*e.g.*, total scan time), but otherwise hold optimization parameters \mathbf{W} , δ , \mathbb{X}^t , \mathbb{X}^b , \mathbb{N}^t , \mathbb{N}^b fixed. Since Ψ is data-independent, we can solve (4.6) and (4.9) offline for each scan profile. The result of each profile’s min-max optimization process (4.9) is a corresponding optimized scan parameter matrix $\hat{\mathbf{P}}$ that is suitable for the range of latent \mathbf{x} and known $\boldsymbol{\nu}$ object parameters specified in \mathbb{X}^t and \mathbb{N}^t , and is robust to variations in those parameters over broader sets \mathbb{X}^b and \mathbb{N}^b , respectively.

4.3 Optimizing SS Sequences for Relaxometry in the Brain

$\{\text{s,scn-dsgn,opt}\}$

This section applies the methods of Section 4.2.2 to the problem of scan design for joint T_1, T_2 estimation from combinations of SS sequences. Section 4.3.1 details how we use optimization problems (4.6) and (4.9) to tailor three SPGR and DESS scan combinations for precise T_1, T_2 estimation in white matter (WM) and grey matter (GM) regions of the brain. Section 4.3.2 compares the predicted performance of the three optimized scan profiles.

4.3.1 Scan Design Details

$\{\text{ss,scn-dsgn,opt,design}\}$

There are numerous candidate scan profiles involving DESS and/or other pulse sequences that may be useful for fast, accurate T_1, T_2 mapping. In this chapter, we consider combinations of magnitude SPGR and DESS scans for estimating the $L \leftarrow 3$ latent parameters T_1, T_2 , and proportionality constant c_2 (defined in Example 3.2.1.2), given knowledge of transmit field inhomogeneity s^t as $K \leftarrow 1$ known parameter. With proper RF phase cycling and gradient spoiling, the SPGR signal s_S (as expressed in (2.27)) contains no explicit T_2

dependence. SPGR's reduced dependence on spatially varying unknowns is reason for its use in T_1 mapping [56, 55, 57] and subsequent T_2 mapping from other sequences [54, 4]. In a similar spirit, we examine scan profiles containing SPGR over other SS sequences because we predict that the SPGR sequence's T_2 -independence may help estimators disentangle T_2 from other unknown sources of DESS signal contrast.

As respectively discussed in Examples 3.2.1.1-3.2.1.2, each SPGR and DESS scan leaves $\mathbf{p} \leftarrow [\alpha_0, T_R]^\top$ as $A \leftarrow 2$ acquisition parameters available to optimize. A given scan profile consisting of S_{SPGR} SPGR and S_{DESS} DESS scans yields $D \leftarrow S_{\text{SPGR}} + 2S_{\text{DESS}}$ datasets. We optimize such a scan profile by solving (4.9) over a dimension- $AD \leftarrow 2(S_{\text{SPGR}} + 2S_{\text{DESS}})$ space of scan parameters.

We select constraints on search space \mathbb{P} based on hardware limitations and desired scan profile properties. Since each pair of DESS signals must share the same \mathbf{p} , the search space \mathbb{P} is reduced to $\mathbb{A}_{0,\text{SPGR}}^{S_{\text{SPGR}}} \times \mathbb{A}_{0,\text{DESS}}^{S_{\text{DESS}}} \times \mathbb{T}_{R,\text{SPGR}}^{S_{\text{SPGR}}} \times \mathbb{T}_{R,\text{DESS}}^{S_{\text{DESS}}}$ (superscripts denote Cartesian powers). We assign flip angle ranges $\mathbb{A}_{0,\text{SPGR}} \leftarrow [5, 90]^\circ$ and $\mathbb{A}_{0,\text{DESS}} \leftarrow [5, 90]^\circ$ to restrict RF energy deposition. We set feasible T_R solution sets $\mathbb{T}_{R,\text{SPGR}} \leftarrow [12.2, +\infty)\text{ms}$ and $\mathbb{T}_{R,\text{DESS}} \leftarrow [17.5, +\infty)\text{ms}$ based on pulse sequence designs that control for other scan parameters. These control parameters are described in further detail in Section 4.4, and are held fixed in all subsequent SPGR and DESS experiments. To equitably compare optima from different scan profiles, we require

$$\mathbf{T}_R := [T_{R,1}, \dots, T_{R,S_{\text{SPGR}}}, T_{R,S_{\text{SPGR}}+1}, \dots, T_{R,S_{\text{SPGR}}+S_{\text{DESS}}}]^\top$$

to satisfy a total time constraint, $\|\mathbf{T}_R\|_1 \leq T_{\text{max}}$. For a scan profile consisting of S_{SPGR} SPGR and S_{DESS} DESS scans, these constraints collectively reduce the search space dimension from AD to $2(S_{\text{SPGR}} + S_{\text{DESS}}) - 1$.

Prior works have considered T_1 or T_2 estimation from as few as 2 SPGR [68, 54] or 1 DESS [60] scan(s), respectively. We likewise elect to optimize the $(S_{\text{SPGR}}, S_{\text{DESS}}) \leftarrow (2, 1)$ scan profile as a benchmark. We choose $T_{\text{max}} \leftarrow 2(12.2) + 1(17.5) = 41.9\text{ms}$ and select other scan profiles capable of meeting this time constraint. Requiring that candidate profiles contain $S_{\text{DESS}} \geq 1$ DESS scans for T_2 contrast and satisfy $D \geq L (= 3)$ for well-conditioned estimation, we note that $(1, 1)$ and $(0, 2)$ are the only other eligible profiles.

In the ensuing experiments, we focus on precise T_1, T_2 estimation in the brain. Noting that $T_1 \sim 10T_2$, we choose $\mathbf{W} \leftarrow \text{diag}(0.1, 1, 0)$ to place approximately equal importance on precise T_1 versus T_2 estimation and zero weight on proportionality constant c_2 estimation (obviating the need for complex differentiation in (4.3)). Since Ψ then depends on c_2 through only a scale factor, it suffices to fix $c_2 \leftarrow 1$ and design the la-

tent object parameter range as $\mathbb{X}^t \leftarrow \mathbb{T}_1^t \times \mathbb{T}_2^t \times 1$. Here, $\mathbb{T}_1^t \leftarrow [800, 1400]\text{ms}$ and $\mathbb{T}_2^t \leftarrow [50, 120]\text{ms}$ correspond with WM and GM regions of interest (ROIs) at 3T [78, 79]. We take $\mathbb{N}^t \leftarrow [0.9, 1.1]$ to account for 10% transmit field spatial variation. Broadened ranges $\mathbb{X}^b \leftarrow [400, 2000]\text{ms} \times [40, 200]\text{ms} \times 1$ and $\mathbb{N}^b \leftarrow [0.5, 2]$ are constructed to encourage solutions robust to a realistically wide range of object parameters. We assume constant noise variance $\sigma_1^2 = \dots = \sigma_D^2 := \sigma^2$, where $\sigma^2 \leftarrow 1.49 \times 10^{-7}$ is selected to reflect measurements from normalized phantom datasets (*cf.* Section 4.4.2.1 for acquisition details). Lastly, we set $\delta \leftarrow 0.01$ to select a robust scan parameter $\hat{\mathbf{P}}$ with associated worst-case cost $\tilde{\Psi}^t(\hat{\mathbf{P}})$ within 1% of global optimum $\tilde{\Psi}^t(\hat{\mathbf{P}})$.

4.3.2 Scan Profile Comparisons

We solve (4.6) and (4.9) via grid search to allow illustration of $\tilde{\Psi}^t(\mathbf{P})$ as well as worst-case T_1, T_2 standard deviations $\tilde{\sigma}_{T_1}^t(\mathbf{P})$ and $\tilde{\sigma}_{T_2}^t(\mathbf{P})$, each defined as

$$\tilde{\sigma}_{T_1}^t(\mathbf{P}) := \max_{\substack{\mathbf{x} \in \mathbb{X}^t \\ \boldsymbol{\nu} \in \mathbb{N}^t}} \sigma_{T_1}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}); \quad (4.11)$$

$$\tilde{\sigma}_{T_2}^t(\mathbf{P}) := \max_{\substack{\mathbf{x} \in \mathbb{X}^t \\ \boldsymbol{\nu} \in \mathbb{N}^t}} \sigma_{T_2}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}), \quad (4.12)$$

where $\sigma_{T_1}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P})$ and $\sigma_{T_2}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P})$ are corresponding diagonal elements of inverse Fisher matrix $\mathbf{F}^{-1}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P})$. Grid searches for the (2, 1), (1, 1), and (0, 2) profiles each took about 4, 43, and 28 minutes, respectively. All experiments described hereafter were carried out using MATLAB[®] R2013a on a 3.5GHz desktop with 32GB RAM.

Table 4.1 compares optimized scan parameters for profiles consisting of (2, 1), (1, 1), and (0, 2) SPGR and DESS scans, respectively. In addition to $\tilde{\sigma}_{T_1}^t(\hat{\mathbf{P}})$ and $\tilde{\sigma}_{T_2}^t(\hat{\mathbf{P}})$, Table 4.1 presents analogous worst-case standard deviations $\tilde{\sigma}_{T_1}^b(\hat{\mathbf{P}})$ and $\tilde{\sigma}_{T_2}^b(\hat{\mathbf{P}})$ over $\mathbb{X}^b \times \mathbb{N}^b$ to show how each estimator degrades over the broadened object parameter range. When viewed over tight range $\mathbb{X}^t \times \mathbb{N}^t$, the (0, 2) profile provides a 11.5% reduction in worst-case cost over the other choices. Extending to broadened range $\mathbb{X}^b \times \mathbb{N}^b$, this reduction grows dramatically to 31.4%. We thus observe that while the different optimized profiles afford similar estimator precision over a narrow range of interest, the (0, 2) profile may be preferable due to its robustness to a wide range of object parameters.

Fig. 4.1 displays heat maps of worst-case latent parameter standard deviations $\tilde{\sigma}_{T_1}^t, \tilde{\sigma}_{T_2}^t$ and worst-case cost $\tilde{\Psi}^t$ as pairs of flip angles are varied away from the optimized scan design $\hat{\mathbf{P}}$. Boxes group subfigures corresponding to the same scan profile. Viewing the bottom row of subfigures, it is evident that $\tilde{\Psi}^t(\hat{\mathbf{P}})$ takes similar values for the different scan

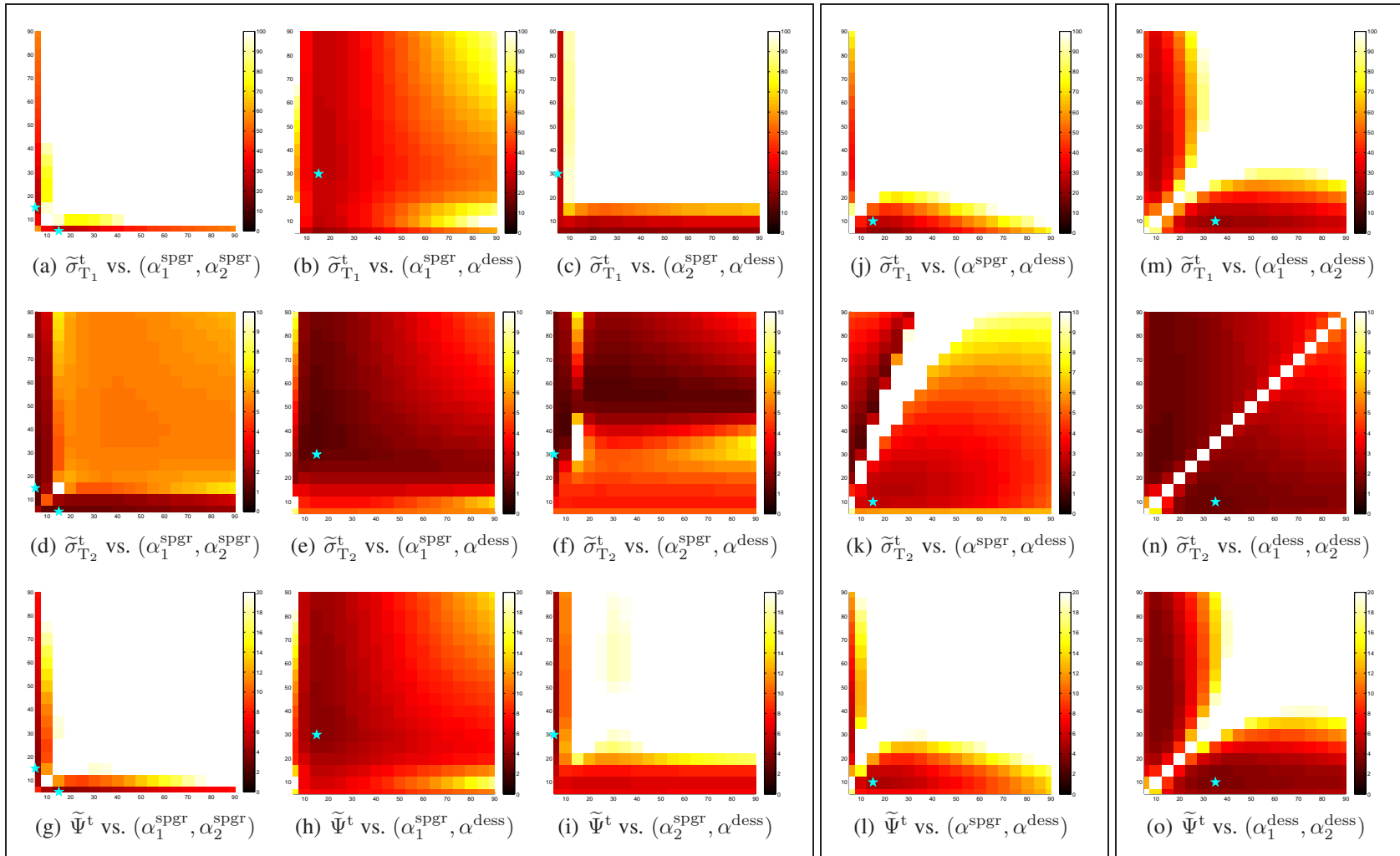


Figure 4.1: Worst-case standard deviations $\tilde{\sigma}_{T_1}^t$ (top), $\tilde{\sigma}_{T_2}^t$ (middle), and cost $\tilde{\Psi}^t$ (bottom), versus pairs of nominal flip angles, holding other scan parameters fixed at selected profile $\hat{\mathbf{P}}$. Subfigures (a)-(i), (j)-(l), and (m)-(o) correspond to scan profiles containing $(\mathcal{S}_{\text{SPGR}}, \mathcal{S}_{\text{DESS}}) = (2, 1)$, $(1, 1)$, and $(0, 2)$ SPGR and DESS scans, respectively. Selected scan parameters (starred) are within $\delta = 1\%$ of global minimizers and retain as much estimator precision as possible over a wide range of latent object parameters. All axes range from 5 to 90 degrees, in 5-degree increments. Colorbar ranges are $[0, 100]$, $[0, 10]$, and $[0, 20]$ milliseconds for rows of $\tilde{\sigma}_{T_1}^t$, $\tilde{\sigma}_{T_2}^t$, and $\tilde{\Psi}^t$ subfigures, respectively. The optimized $(0, 2)$ profile appears most robust to transmit field spatial variation.

Scan	(2, 1)	(1, 1)	(0, 2)
$\hat{\alpha}_0^{\text{SPGR}}$	(15,5) $^\circ$	15 $^\circ$	–
$\hat{\alpha}_0^{\text{DESS}}$	30 $^\circ$	10 $^\circ$	(35,10) $^\circ$
\hat{T}_R^{SPGR}	(12.2, 12.2)	13.9	–
\hat{T}_R^{DESS}	17.5	28.0	(24.4, 17.5)
$\tilde{\sigma}_{T_1}^t(\hat{\mathbf{P}})$	28	27	21
$\tilde{\sigma}_{T_1}^b(\hat{\mathbf{P}})$	154	169	113
$\tilde{\sigma}_{T_2}^t(\hat{\mathbf{P}})$	1.3	2.8	1.5
$\tilde{\sigma}_{T_2}^b(\hat{\mathbf{P}})$	9.1	8.8	6.0
$\tilde{\Psi}^t(\hat{\mathbf{P}})$	4.0	4.9	3.5
$\tilde{\Psi}^b(\hat{\mathbf{P}})$	17.7	17.9	12.2

Table 4.1: Performance summary of different scan profiles, optimized by solving (4.9) subject to scan time constraint $T_{\max} = 41.9\text{ms}$. The first row defines each profile. The next four rows describe $\hat{\mathbf{P}}$. The latter three pairs of rows show how worst-case values degrade from tight to broad ranges. Flip angles are in degrees; all other values are in milliseconds.

{table:profile}

profiles. However, it is apparent that the $(S_{\text{SPGR}}, S_{\text{DESS}}) = (0, 2)$ profile is substantially more robust to transmit field variation than other tested profiles (namely, (2, 1) and (1, 1)). Optimized worst-case cost over broadened latent parameter ranges $\tilde{\Psi}^b(\hat{\mathbf{P}})$ captures this by expanding the range of possible flip angles from $\mathbb{N}^t = [0.9, 1.1]$ to $\mathbb{N}^b = [0.5, 2]$ to account for factor-of-two spatial variation in relative flip angle. As a result, we find that the properties of “broad” search criterion $\tilde{\Psi}^b$ provide a stronger reason to select the (0, 2) scan for joint T_1, T_2 estimation in the brain than the properties of “tight” search criterion $\tilde{\Psi}^t$.

As the DESS sequence has already found success for T_2 mapping from even one scan [60], it is reassuring but unsurprising that our analysis finds two DESS scans to yield the most precise T_2 estimates. More interestingly, our methods suggest that, with a minimum $S_{\text{DESS}} = 2$ scans, DESS can be used to simultaneously estimate T_1 as well. In fact, for certain choices of parameter ranges, a second DESS scan is predicted to afford \hat{T}_1 precision comparable to two SPGR scans.

4.4 Experimental Validation and Results

{s,scn-dsgn,exp}

To test our approach to optimized scan design (described in Section 4.2.2), we estimate \mathbf{T}_1 and \mathbf{T}_2 maps (using maximum likelihood (ML) and regularized likelihood (RL) methods detailed in Section 3.2.2) from datasets collected using the scan profiles optimized in Section 4.3. In Section 4.4.1, we study estimator statistics from simulated data. In Sec-

tions 4.4.2-4.4.3, we progress to phantom and *in vivo* datasets to evaluate scan profile performance under increasingly complex settings. For the latter experiments, we use reference parameter maps from classical (long) pulse sequences, in lieu of ground truth maps.

4.4.1 Numerical Simulations

We select T_1 and T_2 WM and GM values based on previously reported measurements at 3T [78, 79] and extrapolate other nuisance latent object parameters m_0 and T_2^* from measurements at 1.5T [80]. We assign these parameter values to the discrete anatomy of the BrainWeb digital phantom [81, 80] to create ground truth $\mathbf{M}_0, \mathbf{T}_1, \mathbf{T}_2, \mathbf{T}_2^* \in \mathbb{R}^V$ maps. We then choose acquisition parameters based on Table 4.1 (with fixed $T_E = 4.67\text{ms}$) and apply models (2.27) and (2.41)-(2.43) to the 81st slices of these true maps to compute noiseless 217×181 SPGR and DESS image-domain data, respectively.

For each scan profile, we corrupt the corresponding (complex) noiseless dataset \mathbf{S} with additive complex Gaussian noise, whose variance $\sigma^2 \leftarrow 1.49 \times 10^{-7}$ is set to match CRB calculations. This yields realistically noisy datasets \mathbf{Y} ranging from 105-122 signal-to-noise ratio (SNR), where SNR is defined here as

$$\text{SNR}(\mathbf{S}, \mathbf{Y}) := \frac{\|\mathbf{S}\|_F}{\|\mathbf{Y} - \mathbf{S}\|_F}. \quad (4.13)$$

We use each profile’s noisy magnitude dataset $|\mathbf{Y}|$ to compute estimates $\hat{\mathbf{T}}_1$ and $\hat{\mathbf{T}}_2$. We then evaluate estimator bias and variance from latent ground truth \mathbf{T}_1 and \mathbf{T}_2 maps.

In these simulations, we intentionally neglect to model a number of physically realistic effects because their inclusion would complicate study of estimator statistics. First and foremost, we assume knowledge of a uniform transmit field, to avoid confounding s^t and T_1, T_2 estimation errors. For a similar reason, spatial variation in the sensitivity of a single receive coil is also not considered. We omit modeling partial volume effects to ensure deterministic knowledge of WM and GM ROIs. We will explore the influence of these (and other) nuisance effects on scan design in later subsections and chapters.

To isolate bias due to estimator nonlinearity from regularization bias, we solve ML problem (3.7) only, and do not proceed to solve RL problem (3.8). This permits consideration of T_1, T_2 estimation from each of the 7733 WM or 9384 GM data points as voxel-wise independent realizations of the same estimation problem. To minimize quantization bias, we optimize (3.7) using a finely spaced dictionary of signal vectors from 1000 T_1 and T_2 values logarithmically spaced between $[10^2, 10^{3.5}]$ and $[10^1, 10^{2.5}]$, respectively. Using 10^6 dictionary elements, solving (3.7) took less than 7 minutes for each tested scan design $\hat{\mathbf{P}}$.

Scan	(2, 1)	(1, 1)	(0, 2)	Truth
WM \hat{T}_1^{ML}	830 ± 17	830 ± 15	830 ± 14	832
GM \hat{T}_1^{ML}	$1330 \pm 30.$	1330 ± 24	1330 ± 24	1331
WM \hat{T}_2^{ML}	$80. \pm 1.0$	$80. \pm 2.1$	79.6 ± 0.94	79.6
GM \hat{T}_2^{ML}	$110. \pm 1.4$	$110. \pm 3.0$	$110. \pm 1.6$	110

Table 4.2: Sample means \pm sample standard deviations of \mathbf{T}_1 and \mathbf{T}_2 ML estimates in WM and GM ROIs of simulated data, compared across different optimized scan profiles. Sample means exhibit insignificant bias, and sample standard deviations are consistent with worst-case standard deviations $\tilde{\sigma}_{T_1}^t$ and $\tilde{\sigma}_{T_2}^t$ reported in Table 4.1. All values are reported in milliseconds.

{table:numerical}

Table 4.2 verifies² that, despite model nonlinearity and Rician noise, estimation bias in WM- and GM-like voxels is negligible. Sample standard deviations are consistent with $\tilde{\sigma}_{T_1}^t$ and $\tilde{\sigma}_{T_2}^t$ (cf. Table 4.1). We observe that the (1, 1) and (0, 2) profiles afford high \hat{T}_1^{ML} precision, while the (2, 1) and (0, 2) scans afford high \hat{T}_2^{ML} precision. In agreement with the predictions of $\tilde{\Psi}^t$ and $\tilde{\Psi}^b$, these simulation studies suggest that at these SNR levels, an optimized profile containing 2 DESS scans can permit \mathbf{T}_1 and \mathbf{T}_2 estimation precision in WM and GM comparable to optimized profiles containing SPGR/DESS combinations.

Fig. 4.2 histograms (voxel-wise independent) ML estimates \hat{T}_1^{ML} and \hat{T}_2^{ML} from the (0, 2) scan profile. Each histogram is over a WM or GM ROI, within which all voxels are assigned the same single-component true T_1 and T_2 nominal value, listed in Table 4.2.

Overlaid in dashed maroon are normal distributions with latent means T_1 and T_2 and variances computed from the Fisher matrix at T_1, T_2 values in WM or GM. It is apparent that despite finite SNR and Rician noise, \hat{T}_1^{ML} and \hat{T}_2^{ML} exhibit negligible bias and near-Gaussian shape, suggesting locally linear behavior of the DESS signal model in T_1 and T_2 (\hat{T}_1^{ML} and \hat{T}_2^{ML} distributions from other profiles are similar).

The subfigures of Fig. 4.2 superimpose in solid green a second set of normal distributions, with the same means T_1 and T_2 as before, but worst-case standard deviations $\tilde{\sigma}_{T_1}^t$ and $\tilde{\sigma}_{T_2}^t$. The separations between these distribution pairs visually depict how estimator variances specific to WM or GM T_1 and T_2 values differ from worst-case variances. Using the fixed latent object parameters to optimize scan profiles can tailor scans for precise estimation in *either* WM *or* GM. In contrast, the proposed min-max formulation finds scan parameters that ensure precise estimation in *both* WM *and* GM.

²Each sample statistic presented in this chapter is rounded off to the highest place value of its corresponding uncertainty measure. For simplicity, each uncertainty measure is itself endowed one extra significant figure. Decimal points indicate the significance of trailing zeros.

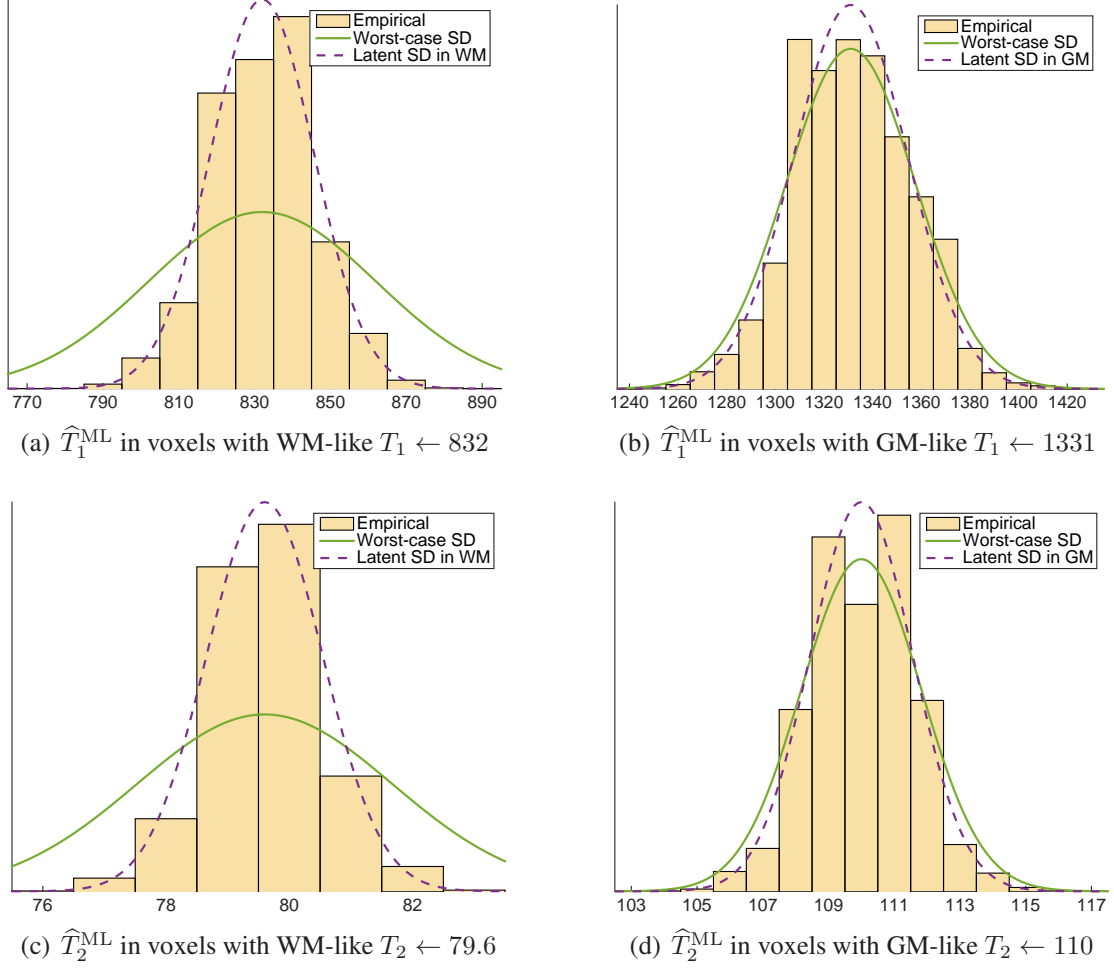


Figure 4.2: Histograms of T_1 and T_2 estimates from noisy independent measurements of a *single* nominal WM or GM value. In each plot, two normal distributions are overlaid, each with latent means T_1 and T_2 . In (a)-(b) and (c)-(d), the solid green curve is $\mathcal{N}(T_1, (\tilde{\sigma}_{T_1}^t)^2)$ and $\mathcal{N}(T_2, (\tilde{\sigma}_{T_2}^t)^2)$, respectively. In (a)-(d), the dashed maroon curves have variances computed from the Fisher information at *a priori* unknown T_1, T_2 values in WM or GM. These plots correspond to an optimized $(0, 2)$ scan profile; analogous plots for other profiles are visually similar. At realistic noise levels, parameter estimates distribute with minimal bias and near-Gaussian shape. Thus, the CRB reliably approximates \hat{T}_1^{ML} and \hat{T}_2^{ML} errors.

{fig:normality}

4.4.2 Phantom Experiments

This subsection describes two experiments. In the first experiment, we compare SPGR/DESS scan profiles described in Table 4.1 (as well as a reference profile consisting of IR and SE scans) against nuclear magnetic resonance (NMR) measurements from the National Institute for Standards and Technology (NIST) [1]. These measurements provide information about *ROI sample means* and *ROI sample standard deviations* (Fig. 4.5), which we define as first- and second-order statistics computed across voxels within an ROI. In the second experiment, we repeat the SPGR/DESS scan profiles 10 times and compute *sample standard deviation maps* across repetitions. Taking ROI sample means of these maps gives *pooled sample standard deviations* (Table 4.4), which indicate relative scan profile precision.

4.4.2.1 Within-ROI Statistics

We acquire combinations of (2, 1), (1, 1), and (0, 2) SPGR and DESS coronal scans of a High Precision Devices® MR system phantom T_2 array. For each scan profile, we prescribe the optimized flip angles $\hat{\alpha}_0$ and repetition times \hat{T}_R listed in Table 4.1, and hold all other scan parameters fixed. We achieve the desired nominal flip angles by scaling a 20mm slab-selective Shinnar-Le Roux excitation [82], of duration 1.28ms and time-bandwidth product 4. For each DESS (SPGR) scan, we apply 2 (10) spoiling phase cycles over a 5mm slice thickness. We acquire all steady-state phantom and *in vivo* datasets with a $256 \times 256 \times 8$ matrix over a $240 \times 240 \times 30$ mm³ field of view (FOV). Using a 31.25kHz readout bandwidth, we acquire all data at minimum $T_E \leftarrow 4.67$ ms before or after RF excitations. To avoid slice-profile effects, we sample k-space over a 3D Cartesian grid. After Fourier transform of the raw datasets, only one of the excited image slices is used for subsequent parameter mapping. Including time to reach steady-state, each steady-state scan profile requires 1m37s scan time.

To validate a reference scan profile for use in *in vivo* experiments, we also collect 4 IR and 4 SE scans. For (phase-sensitive, SE) IR, we hold $(T_R, T_E) \leftarrow (1400, 14)$ ms fixed and vary (adiabatic) inversion time $T_I \in \{50, 150, 450, 1350\}$ ms across scans. For SE, we similarly hold $T_R \leftarrow 1000$ ms fixed and vary echo time $T_E \in \{10, 30, 60, 150\}$ ms across scans. We prescribe these scan parameters to acquire 256×256 datasets over the same $240 \times 240 \times 5$ mm³ slice processed from the SPGR/DESS datasets. Each IR and SE scan requires 5m58s and 4m16s, for a total 40m58s scan time.

We additionally collect a pair of Bloch-Siegert shifted 3D SPGR scans for separate transmit field estimation [83]. We insert a 9ms Fermi pulse at ± 8 kHz off-resonance into an SPGR sequence immediately following on-resonant excitation. We estimate regularized

transmit field maps [84] from the resulting pair of datasets. We normalize this transmit field map estimate by the 0.075G peak Fermi pulse amplitude to estimate transmit coil spatial variation map s_t . After calibration via separate measurements, we take s_t as known. For consistency, we account for flip angle variation when estimating T_1 and T_2 from both candidate (SPGR/DESS) and reference (IR/SE) scan profiles. With a repetition time of 21.7ms, this transmit field mapping acquisition requires 1m40s total scan time.

We acquire all phantom datasets using a GE DiscoveryTM MR750 3.0T scanner with an 8-channel receive head array. We separately normalize and combine coil data from each scan profile using a natural extension of [85] to the case of multiple datasets. For each optimized SPGR/DESS scan profile \hat{P} , we pre-cluster known parameter maps N into 10 clusters using k -means++ [86] and use each of the 10 cluster means to compute a corresponding dictionary of signal vectors from 300 T_1 and T_2 values logarithmically spaced between $[10^{1.5}, 10^{3.5}]$ and $[10^{0.5}, 10^3]$, respectively. We then iterate over clusters and use each dictionary in conjunction with corresponding coil-combined magnitude image data to produce ML parameter estimates $\hat{X}_{ML}(N, \hat{P})$. We subsequently solve RL problem (3.8) with initialization $\hat{X}_{ML}(N, \hat{P})$ to obtain regularized estimates $\hat{X}_{RL}(N, \hat{P})$ for each \hat{P} . We design regularizer (3.9) to encourage RL parameter estimates from different scan profiles to exhibit similar levels of smoothness. Letting $l \in \{1, 2, 3\}$ enumerate latent object parameters T_1 , T_2 , and the proportionality constant, we choose mild regularization parameters $(\beta_1, \beta_2, \beta_3) := D \times (2^{-21}, 2^{-23}, 2^{-26})$ to scale with the number of datasets and fix shape parameters $(\gamma_1, \gamma_2, \gamma_3) := (2^5 \text{ ms}, 2^2 \text{ ms}, 2^{-2})$ to values on the order of anticipated standard deviations. We iteratively update X until convergence criterion

$$\{\text{eq:conv,crit}\} \quad \|X^{(i)} - X^{(i-1)}\|_F < 10^{-7} \|X^{(i)}\|_F \quad (4.14)$$

is satisfied. For all steady-state profiles tested, ML initializations and RL reconstructions of phantom datasets require less than 3m30s and 9s, respectively.

We next describe sequential³ T_1 , then T_2 estimation from IR and SE reference scans. We first jointly coil-combine all 8-channel IR and SE phantom datasets to produce complex images. We next estimate T_1 along with a (nuisance parameter) inversion efficiency map via (3.7) and (3.8) from the 4 complex coil-combined IR images. By using the same flip angle scaling map s_t as is used for SPGR/DESS profiles, we estimate T_1 using a signal model similar to one proposed in [88], which accounts for imperfect excitation/refocusing

³ We initially attempted to circumvent sequential T_1 , then T_2 estimation by instead jointly estimating M_0 , T_1 , T_2 , and inversion efficiency from the IR and SE datasets together. Even using magnitude data and signal models, this resulted in heavily biased parameter maps, possibly due to the dependence of adiabatic inversion efficiency on relaxation parameters [87].

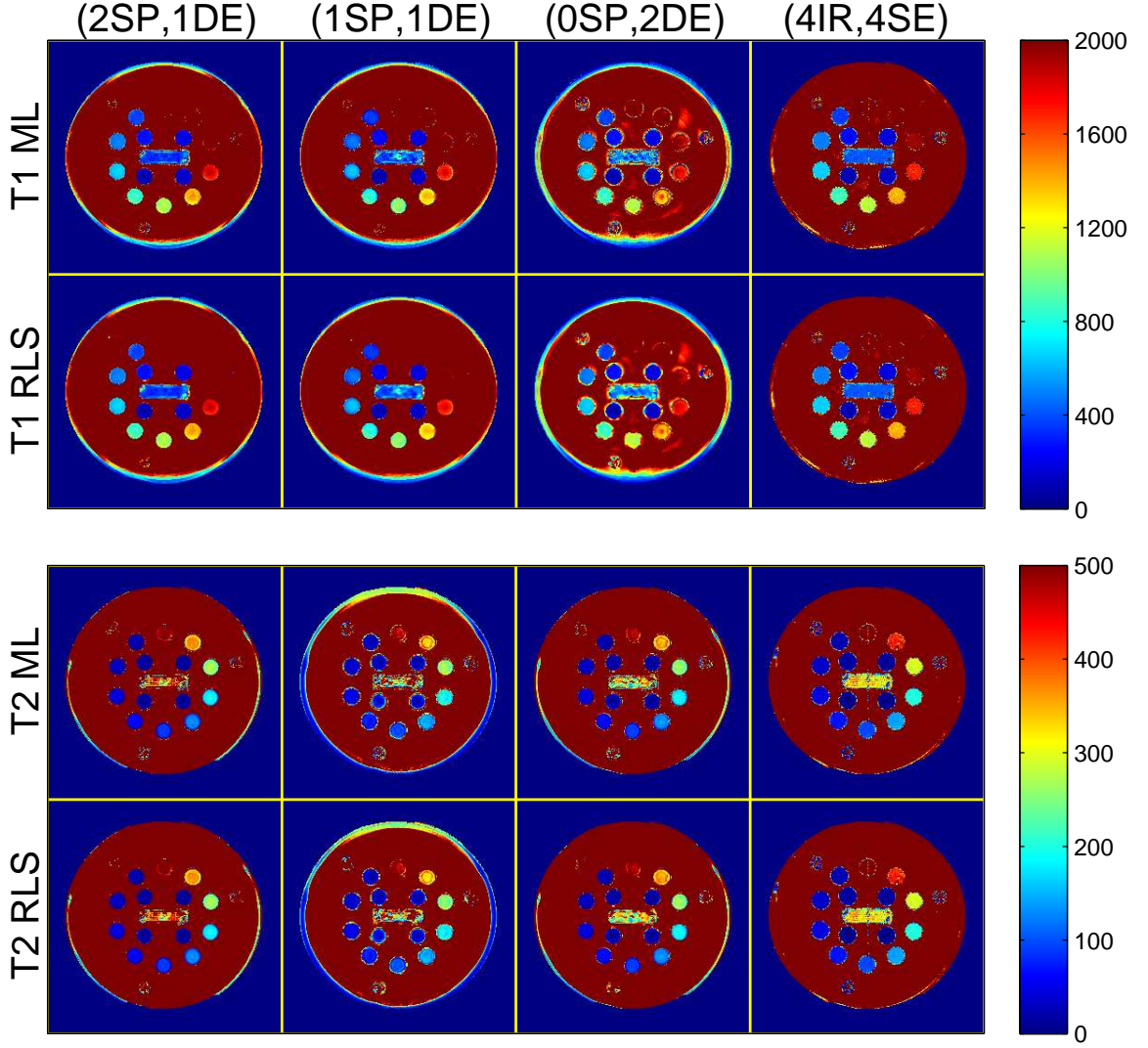


Figure 4.3: Colorized T_1 and T_2 ML and RL estimates from an HPD[®] quantitative phantom. Columns correspond to scan profiles consisting of (2 SPGR, 1 DESS), (1 SPGR, 1 DESS), (0 SPGR, 2 DESS), and (4 IR, 4 SE) acquisitions. Rows distinguish T_1 and T_2 ML and RL estimators. Fig. 4.4 provides identical grayscale images that enumerate vials. Colorbar ranges are in milliseconds.

{fig:hpj,et}

and imperfect inversion. We then take both T_1 and s_t as known and estimate T_2 along with nuisance parameter M_0 (accounting for imperfect excitation/refocusing and incomplete recovery) via (3.7) and (3.8) from the 4 complex coil-combined SE images. We hold all other reconstruction details identical to those of SPGR/DESS reconstructions.

Figs. 4.3-4.4 compare in color and grayscale phantom T_1 and T_2 ML and RL estimates from optimized scan profiles. Vials are enumerated in Fig. 4.5 in descending T_1 and T_2 order. Vials corresponding to tight \mathbb{X}^t and broad \mathbb{X}^b parameter ranges are highlighted with

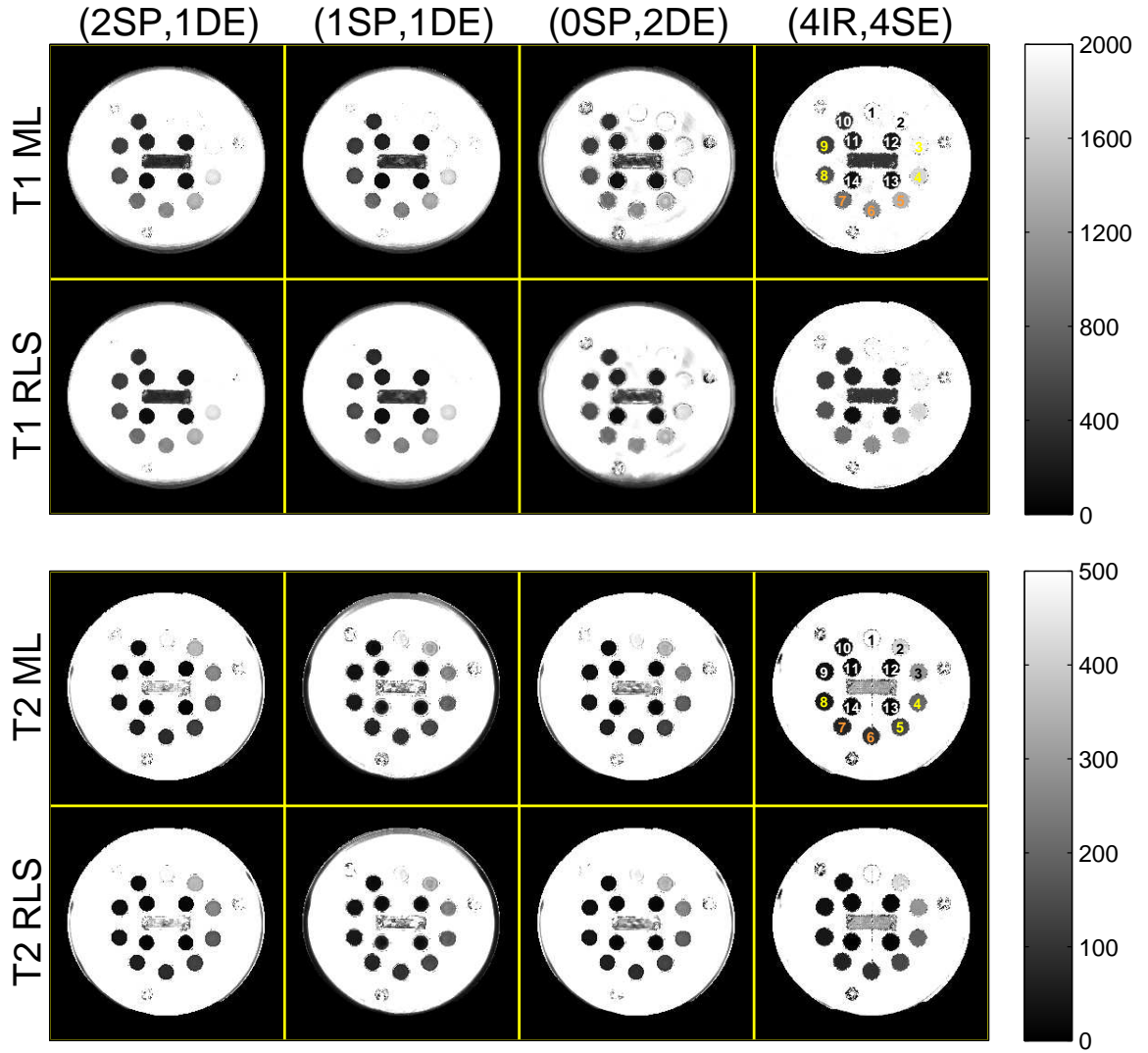


Figure 4.4: Grayscale T_1 and T_2 ML and RL estimates from an HPD[®] quantitative phantom. Columns correspond to scan profiles consisting of (2 SPGR, 1 DESS), (1 SPGR, 1 DESS), (0 SPGR, 2 DESS), and (4 IR, 4 SE) acquisitions. Rows distinguish T_1 and T_2 ML and RL estimators. Vials are enumerated and color-coded to correspond with data points in Fig. 4.5. Fig. 4.3 provides identical colorized images. Colorbar ranges are in milliseconds.

orange and yellow labels, respectively. Within these vials of interest, parameter maps from different scans appear visually similar.

In higher- T_1 vials (and the surrounding water), more bias is apparent in \hat{T}_1 ML and RL estimates from the (0, 2) scan profile than from the (2, 1) and (1, 1) scan profiles. With the signal models used in this study, the images suggest that scan profiles consisting of at least one SPGR scan may offer increased protection against T_1 estimation bias.

Fig. 4.5 plots sample means and sample standard deviations computed within circular

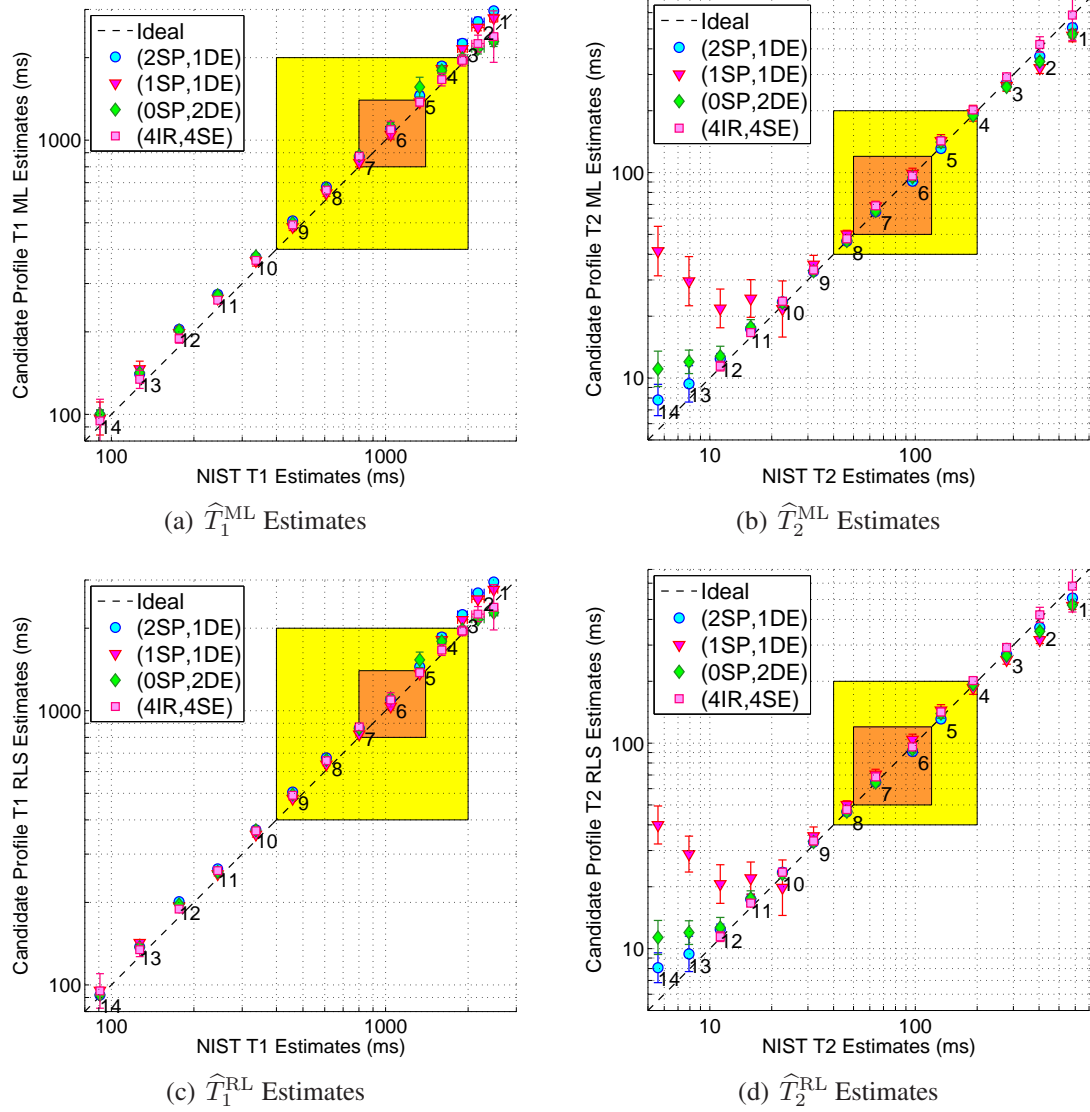


Figure 4.5: Phantom within-ROI sample statistics of T_1 and T_2 ML and RL estimates from optimized SPGR/DESS and reference IR/SE scan profiles, versus NIST NMR measurements [1]. Markers and error bars indicate ROI sample means and ROI sample standard deviations within the 14 labeled and color-coded vials in Fig. 4.4. Tight \mathbb{X}^t and broad \mathbb{X}^b latent parameter ranges are highlighted in orange and yellow, respectively. Table 4.3 replicates sample statistics within Vials 5-8. Our MR measurements are at 293K and NIST NMR measurements are at 293.00K. Within the designed parameter ranges, estimates from different acquisitions are in reasonable agreement with NIST measurements.

{fig:hpdl,ml-rls}

	(2SP,1DE)	(1SP,1DE)	(0SP,2DE)	(4IR,4SE)	NIST NMR
V5 \hat{T}_1^{ML}	1450 \pm 50.	1380 \pm 41	1600 \pm 130	1380 \pm 44	1332 \pm 0.8
V5 \hat{T}_1^{RL}	1450 \pm 26	1370 \pm 16	1540 \pm 98	1380 \pm 37	
V6 \hat{T}_1^{ML}	1100 \pm 30.	1050 \pm 39	1120 \pm 39	1100 \pm 74	1044 \pm 3.2
V6 \hat{T}_1^{RL}	1100 \pm 15	1040 \pm 14	1110 \pm 16	1100 \pm 64	
V7 \hat{T}_1^{ML}	870 \pm 22	830 \pm 29	880 \pm 29	870 \pm 25	801.7 \pm 1.70
V7 \hat{T}_1^{RL}	865 \pm 7.1	820 \pm 11	860 \pm 18	870 \pm 21	
V8 \hat{T}_1^{ML}	680 \pm 12	640 \pm 18	670 \pm 12	658 \pm 8.8	608.6 \pm 1.03
V8 \hat{T}_1^{RL}	674 \pm 7.6	637 \pm 7.4	662 \pm 6.6	658 \pm 7.1	
V5 \hat{T}_2^{ML}	131 \pm 5.5	140 \pm 10.	141 \pm 8.4	143 \pm 4.9	133.27 \pm 0.073
V5 \hat{T}_2^{RL}	131 \pm 5.2	145 \pm 9.1	139 \pm 7.1	142 \pm 4.8	
V6 \hat{T}_2^{ML}	91 \pm 3.5	99 \pm 6.0	95 \pm 4.2	96 \pm 2.7	96.89 \pm 0.049
V6 \hat{T}_2^{RL}	91 \pm 3.4	104 \pm 6.2	93 \pm 3.7	96 \pm 2.6	
V7 \hat{T}_2^{ML}	64 \pm 2.2	69 \pm 3.9	65 \pm 2.1	69 \pm 1.2	64.07 \pm 0.034
V7 \hat{T}_2^{RL}	65 \pm 2.1	71 \pm 4.3	64 \pm 1.9	69 \pm 1.2	
V8 \hat{T}_2^{ML}	46 \pm 1.5	50. \pm 2.3	46 \pm 1.1	47.6 \pm 0.87	46.42 \pm 0.014
V8 \hat{T}_2^{RL}	46 \pm 1.5	50. \pm 2.3	46 \pm 1.0	47.5 \pm 0.85	

Table 4.3: Phantom within-ROI sample means \pm sample standard deviations of T_1 and T_2 estimates from optimized SPGR/DESS and reference IR/SE scan profiles, versus NIST NMR measurements (*cf.* slide 22 of e-poster corresponding to [1]). For sake of brevity, sample statistics corresponding only to phantom vials within (or nearly within) tight design range \mathbb{X}^t (color-coded orange in Fig. 4.4) are reported. Fig. 4.5 plots sample statistics for all vials. ‘V#’ abbreviates vial numbers. All values are reported in milliseconds.

ROIs of phantom T_1 and T_2 ML and RL estimates. The highlighted orange and yellow parameter spaces correspond to design ranges \mathbb{X}^t and \mathbb{X}^b . T_1 estimates from both the candidate (2, 1), (1, 1), and (0, 2) (SPGR, DESS) and reference (4, 4) (IR, SE) profiles are in reasonable agreement with NIST estimates [1] across the vial range. T_2 estimates from all profiles are also in good agreement with NIST for vials within \mathbb{X}^b . SPGR/DESS profiles likely underestimate large T_2 values ($\geq 200\text{ms}$) due to greater influence of diffusion in DESS [89, 90, 91], (studied further in Appendix B). SPGR/DESS profiles possibly overestimate and the IR/SE profile likely underestimates short ($\leq 30\text{ms}$) and very short ($\leq 15\text{ms}$) T_2 values, respectively, due to poorly conditioned estimation. Table 4.3 replicates sample statistics in Fig. 4.5 for vials 5-8. Compared to ML initializations, (weakly) regularized estimates reduce error bars without introducing substantial additional bias.

	(2SP,1DE)	(1SP,1DE)	(0SP,2DE)
V5 $\hat{\sigma}_{\hat{T}_1^{\text{ML}}}$	50 ± 12	$40 \pm 10.$	39 ± 9.4
V6 $\hat{\sigma}_{\hat{T}_1^{\text{ML}}}$	70 ± 18	60 ± 15	70 ± 16
V7 $\hat{\sigma}_{\hat{T}_1^{\text{ML}}}$	60 ± 13	50 ± 13	50 ± 13
V8 $\hat{\sigma}_{\hat{T}_1^{\text{ML}}}$	23 ± 5.4	$20. \pm 4.7$	18 ± 4.3
V5 $\hat{\sigma}_{\hat{T}_2^{\text{ML}}}$	2.6 ± 0.63	6 ± 1.4	3.5 ± 0.84
V6 $\hat{\sigma}_{\hat{T}_2^{\text{ML}}}$	1.9 ± 0.46	5 ± 1.1	2.3 ± 0.54
V7 $\hat{\sigma}_{\hat{T}_2^{\text{ML}}}$	1.4 ± 0.34	3.4 ± 0.80	1.5 ± 0.35
V8 $\hat{\sigma}_{\hat{T}_2^{\text{ML}}}$	1.1 ± 0.26	3.5 ± 0.84	1.4 ± 0.33

Table 4.4: Phantom pooled sample standard deviations \pm pooled standard errors of sample standard deviations, from optimized SPGR/DESS scan profiles. Each entry is a measure of uncertainty of a typical voxel’s T_1 or T_2 ML estimate, estimated over 10 repeated acquisitions. For sake of brevity, sample statistics corresponding only to phantom vials within (or nearly within) tight design range \mathbb{X}^t (color-coded orange in Fig. 4.4) are reported. ‘V#’ abbreviates vial numbers. All values are reported in milliseconds.

4.4.2.2 Across-Repetition Statistics

In a second study, we repeat the (2, 1), (1, 1), and (0, 2) scan profiles 10 times each and separately compute T_1 and T_2 ML estimates for each repetition of each scan profile. We then estimate the standard deviation across repetitions on a per-voxel basis, to produce sample standard deviation maps for each profile. Each ROI voxel of the sample standard deviation map is a better estimate of the *population standard deviation* (which the CRB characterizes) than the ROI sample standard deviation from a single repetition, because the latter estimate is contaminated with slight spatial variation of voxel population means (due to imaging non-idealities such as Gibbs ringing due to k-space truncation).

Table 4.4 reports pooled sample standard deviations and pooled standard errors of the sample standard deviations (computed via expressions in [2]) for phantom vials within (or nearly within) tight design range \mathbb{X}^t (marked orange in Fig. 4.4). Due to error propagation from coil combination and s_t estimation, pooled ML sample standard deviations cannot be compared *in magnitude* to worst-case predicted standard deviations (Table 4.1); however, *trends* of empirical and theoretical standard deviations are overall similar. In particular, the optimized (0, 2) DESS-only scan profile affords T_1 ML estimation precision (in vials whose T_1, T_2 is similar to that of WM/GM) comparable to optimized (2, 1) and (1, 1) mixed (SPGR, DESS) profiles. Also in agreement with predictions, the optimized (2, 1) and (0, 2) profiles afford greater T_2 ML estimation precision than the optimized (1, 1) profile.

4.4.3 In Vivo Experiments

In a single long study of a healthy volunteer, we acquire the same optimized scan profiles containing (2, 1), (1, 1), and (0, 2) SPGR/DESS scans (*cf.* Table 4.1), as well as the reference profile containing (4, 4) IR/SE scans. We obtain axial slices from a 32-channel Nova Medical® receive head array. To address bulk motion between acquisitions and to compare within-ROI statistics, we rigidly register⁴ each coil-combined image to an IR image prior to parameter mapping. All acquisition and reconstruction details are otherwise the same as in phantom experiments (*cf.* Section 4.4.2.1). For all SS scan profiles tested, ML and RL reconstructions of brain datasets require less than 3m30s and 9s, respectively.

Fig. 4.6 compares brain T_1 and T_2 ML and RL estimates from optimized scan profiles. Though in-plane motion is largely compensated via registration, through-plane motion and non-bulk motion likely persist, and will influence ROI statistics. Due to motion (and scan duration) considerations, we examine within-ROI statistics from a single repetition as in Section 4.4.2.1, and do not attempt across-repetition statistics as in Section 4.4.2.2.

Visually, \hat{T}_1 maps from steady-state profiles exhibit similar levels of contrast in WM/GM regions well away from cerebrospinal fluid (CSF) as that seen in the reference \hat{T}_1 estimate. Since we did not optimize any scan profiles for estimation in high- T_1 regions, it is expected that greater differences may emerge in voxels containing or nearby CSF. In particular, T_1 is significantly underestimated within and near CSF by the (0, 2) DESS-only profile. This suggests that with the signal models used in this work, including at least one SPGR scan in an optimized profile may offer greater protection against estimation bias in high- T_1 regions.

Table 4.5 summarizes within-ROI sample means and sample standard deviations computed⁵ over four separate WM ROIs containing 96, 69, 224, and 148 voxels and one pooled cortical GM ROI containing 156 voxels. Within-ROI \hat{T}_1 sample standard deviations are comparable across SS profiles. In agreement with Table 4.1, T_2 estimates from the optimized (1, 1) scan profile exhibit higher within-ROI sample variation than corresponding (2, 1) and (0, 2) \hat{T}_2 maps. Compared to ML counterparts, RL estimates generally reduce within-ROI sample variation and do not significantly change within-ROI sample means.

In most cases, \hat{T}_1 within-ROI sample means from optimized SPGR/DESS scan profiles

⁴For each coil-combined dataset, we compute a separate 2D rigid transformation (with respect to the $T_1 = 50\text{ms}$ IR dataset) via the MATLAB® function `imregtform` and then apply the transformation via `imwarp`. We choose to use rigid transformations instead of affine distortions to avoid scaling; however in doing so we sacrifice compensating for small through-plane rotations. We do not find registration to substantially change subsequently estimated relaxation maps; however, this extra step substantially improves alignment of (especially cortical GM) ROIs in T_1 and T_2 estimates from different scan profiles.

⁵We have taken effort to select ROIs that reflect expected anatomy in all coil-combined and registered images, including adjacent slices in images from 3D acquisitions. However, we acknowledge the possibility of some contamination across tissue boundaries, especially WM and/or CSF contamination into cortical GM.

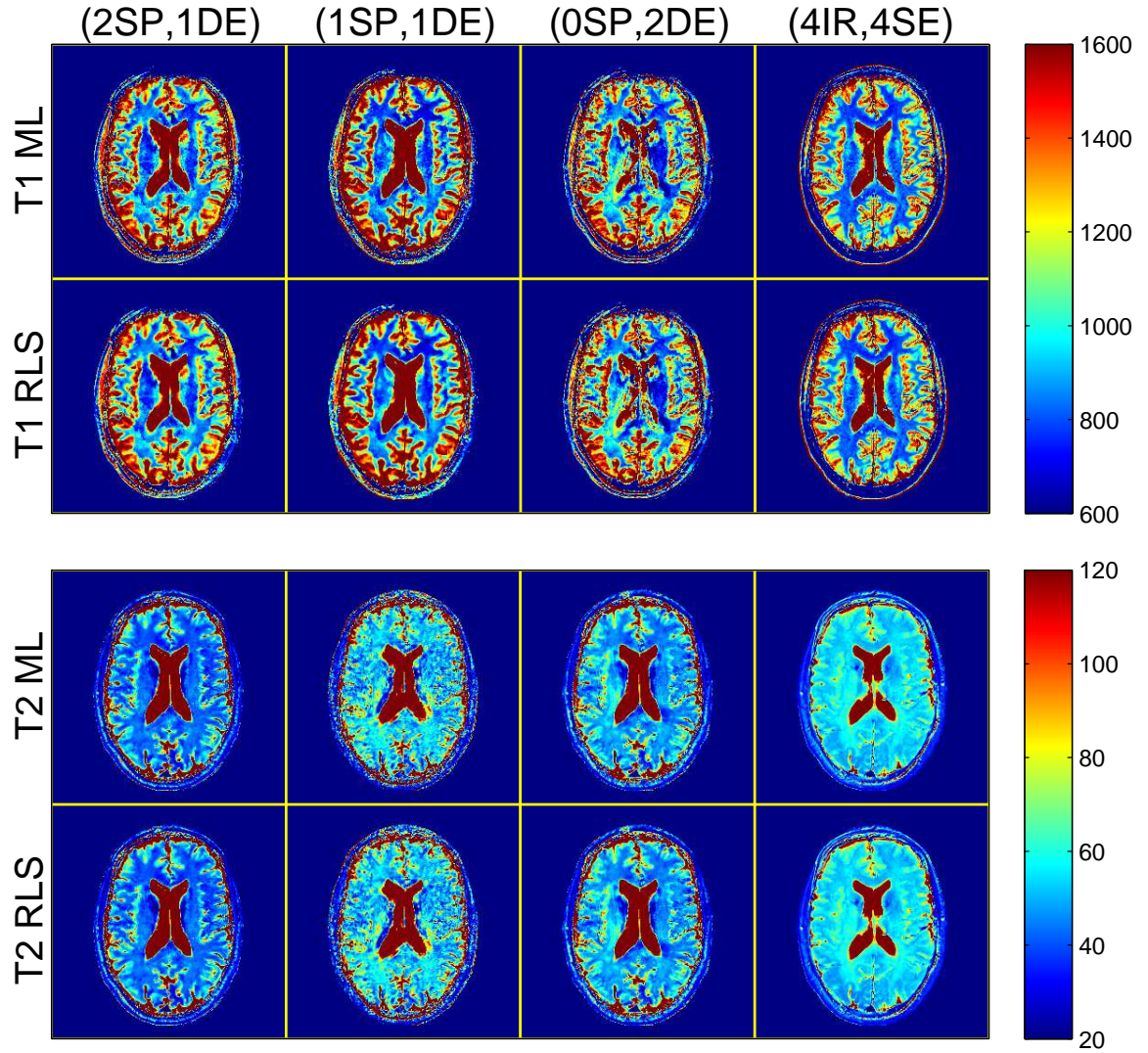
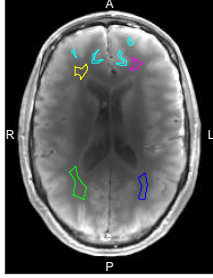


Figure 4.6: Colorized T_1 and T_2 ML and RL estimates from the brain of a healthy volunteer. Columns correspond to profiles consisting of (2 SPGR, 1 DESS), (1 SPGR, 1 DESS), (0 SPGR, 2 DESS), and (4 IR, 4 SE) acquisitions. Rows distinguish T_1 and T_2 ML and RL estimators. Table 4.5 presents corresponding WM/GM within-ROI sample statistics. Colorbar ranges are in milliseconds.

{fig:brain,jet}

{fig:brain.roi.gray}



	ROI	(2SP,1DE)	(1SP,1DE)	(0SP,2DE)	(4IR,4SE)
\hat{T}_1^{ML}	AR WM	840 ± 32	770 ± 31	840 ± 43	780 ± 22
	AL WM	740 ± 61	660 ± 45	740 ± 55	760 ± 24
	PR WM	890 ± 88	860 ± 72	960 ± 84	810 ± 26
	PL WM	860 ± 70.	850 ± 61	880 ± 79	820 ± 37
	A GM	1200 ± 210	1200 ± 230	1300 ± 230	1300 ± 180
\hat{T}_1^{RL}	AR WM	840 ± 24	770 ± 20.	840 ± 43	780 ± 20.
	AL WM	740 ± 51	670 ± 37	740 ± 54	760 ± 23
	PR WM	890 ± 79	860 ± 61	960 ± 82	810 ± 24
	PL WM	870 ± 62	850 ± 50.	880 ± 78	820 ± 35
	A GM	1200 ± 200	1200 ± 220	1300 ± 230	1300 ± 180
\hat{T}_2^{ML}	AR WM	40. ± 1.3	54 ± 3.8	46 ± 1.5	55 ± 1.9
	AL WM	40. ± 1.7	50. ± 4.5	44 ± 1.7	53 ± 1.8
	PR WM	43 ± 2.7	60. ± 6.9	51 ± 3.6	59 ± 2.1
	PL WM	43 ± 1.8	57 ± 4.9	49 ± 2.5	57 ± 1.8
	A GM	50 ± 12	60 ± 15	60 ± 11	59 ± 6.0
\hat{T}_2^{RL}	AR WM	40. ± 1.3	54 ± 3.4	46 ± 1.5	55 ± 1.9
	AL WM	40. ± 1.7	50. ± 4.4	43 ± 1.7	53 ± 1.8
	PR WM	43 ± 2.8	60. ± 6.7	51 ± 3.7	58 ± 2.3
	PL WM	43 ± 1.7	57 ± 4.7	49 ± 2.5	57 ± 1.8
	A GM	50 ± 12	60 ± 15	60 ± 11	59 ± 6.4

{table:brain}

Table 4.5: *Left*: WM/GM ROIs, overlaid on a representative anatomical (coil-combined, IR) image. Separate WM ROIs are distinguished by anterior-right (AR), anterior-left (AL), posterior-right (PR), and posterior-left (PL) directions. Four small anterior (A) cortical GM polygons are pooled into a single ROI. *Right*: Within-ROI sample means ± within-ROI sample standard deviations of T_1 and T_2 ML and RL estimates from the brain of a healthy volunteer (Fig. 4.6 presents corresponding images). Sample statistics are computed within ROIs indicated in the anatomical image. All values are reported in milliseconds.

do not deviate substantially from each other or from reference IR/SE measurements. Two notable exceptions are \hat{T}_1^{ML} in anterior left and posterior right WM from (1, 1) and (0, 2) profiles: these estimates are significantly lower and higher than analogous estimates from other profiles, respectively. Results thus suggest that the optimized (2, 1) scan profile yields WM \hat{T}_1^{ML} estimates that are more consistently similar to IR WM \hat{T}_1^{ML} estimates than other optimized SPGR/DESS profiles.

Systematic differences in \hat{T}_2 sample means are evident across scan profiles, particularly within WM ROIs. Curiously, the (1, 1) profile agrees most consistently (in WM/GM \hat{T}_2^{ML} within-ROI sample mean) with reference estimates, though with relatively high sample variation. The (2, 1) and (0, 2) SPGR/DESS profiles produce consistently lower WM \hat{T}_2^{ML} than the reference IR/SE profile, though the (0, 2) profile is in reasonable agreement with other steady-state estimates [92]. These discrepancies may due to differences in sensitivity

to multi-compartmental relaxation [93]. Specifically, different signal models with different scan parameter choices might be more or less sensitive to the model mismatch incurred by neglecting to distinguish the multiple T_2 components within each voxel. Chapter 6 studies multi-compartmental relaxation in much greater detail.

4.5 Discussion and Future Work

{s,scn-dsgn,disc}

Phantom experiments show that optimized scan profiles consisting of (2, 1), (1, 1), and (0, 2) (SPGR, DESS) scans yield accurate WM/GM T_1, T_2 estimates, and that empirical precision trends across profiles agree reasonably with CRB-based predictions. However, *in vivo* experiments reveal that even with scan optimization, it may be challenging to achieve clinically viable levels of precision from the aforementioned SS profiles, at least at 3T. At the expense of greater scan time, it is of course possible that optimized profiles containing greater numbers of SPGR, DESS, and/or other SS scans can provide clinically acceptable precision levels. For these and other more complicated scan profiles, estimator dependence on scan parameters becomes even less intuitive, increasing the need for scan design.

The proposed scan design framework addresses spatial variation in object parameters through a min-max design criterion. The min-max criterion guarantees an upper bound on a weighted sum of variances and assumes no prior knowledge of distributions. However, in general it is non-differentiable in \mathbf{P} , precluding gradient-based optimization. Furthermore, it is conservative by nature, and often selects scan parameters based on corner cases of the object parameter space. To reduce the influence of corner cases, it may be desirable to instead construct a cost function related to the coefficient of variation as in [69, 67, 70, 71], perhaps by setting parameter weights $\mathbf{W}^{-1} \leftarrow \text{diag}(\mathbf{x})$ for $\mathbf{x} \neq 0$ in (4.5).

As a less conservative alternative to min-max design, other recent works [73, 74] have addressed object parameter spatial variation by instead constructing cost functions related to the Bayesian CRB [94], which characterizes the expected precision with respect to a prior distribution on object parameters. Bayesian cost functions are usually differentiable and can also, with appropriate priors, penalize object parameter coefficients of variation instead of variances, as in [73]. However, prior distributions are generally unknown, and may need to be estimated from data, as in [74].

Careful calibration of flip angle scaling s_t is essential for accurate T_1, T_2 estimation from SPGR/DESS scan profiles. In this work, we estimate s_t from *separate* acquisitions and adjust nominal flip angles prior to reconstruction, but acknowledge that non-idealities in those separate acquisitions may themselves cause resultant transmit field estimation errors to propagate into our T_1, T_2 estimates. To reduce error propagation, it may be de-

sirable to instead design scan profiles to permit *joint* estimation of \mathbf{s}_t , in addition to other latent object parameters. Unfortunately, we find that optimizing the (2, 1) or (0, 2) profile to allow for four-parameter $\mathbf{x}(\mathbf{r}) \leftarrow [T_1(\mathbf{r}), T_2(\mathbf{r}), c_2(\mathbf{r}), s^t(\mathbf{r})]^T$ estimation results in unacceptably high amplification of the worst-case T_1 standard deviation. (Incidentally however, precise T_2 ML and RL estimation alone from the (2, 1) or (0, 2) profile is possible [4].) It remains an open scan design question as to whether time spent collecting Bloch-Siegrert data for separate \mathbf{s}_t mapping could instead be better spent collecting additional SPGR, DESS, and/or other data for joint estimation.

By working with closed-form signal expressions, we neglect to model several higher-order effects. However, it is apparent that the nonlinear estimation procedures required for many mapping problems can amplify the influence of these secondary effects, often inducing substantial bias. Since the CRB (as described) applies only to unbiased estimators, it is thus desirable to use signal models that are as complete as possible for CRB-based scan design. In theory, scan optimization approach (4.9) is even compatible with acquisitions where a closed-form model relating data to latent and scan parameters is unknown, as in [49, 95]. In practice, difficulties arise in efficient computation of signal gradients required in (4.3), which may demand more specialized techniques, as in [96]. Designing scan profiles involving such complex signal models would likely necessitate optimization techniques more involved than the simple grid searches used in this work.

4.6 Conclusion

{s,scn-dsgn,conc}

This chapter has introduced a CRB-inspired min-max optimization approach to guide MR scan design for precise parameter estimation. As a detailed example, we have optimized combinations of fast SPGR and DESS scans for T_1, T_2 relaxometry in WM and GM regions of the human brain at 3T. Numerical simulations show that at typical noise levels and with accurate flip angle prior knowledge, WM- and GM-like T_1, T_2 ML estimates from optimized scans are nearly unbiased, and so worst-case CRB predictions yield reliable bounds on ROI sample variances. Phantom accuracy experiments show that optimized combinations of (2, 1), (1, 1), or (0, 2) (SPGR, DESS) scans are in excellent agreement with NIST and IR/SE measurements over the designed latent object parameter range of interest. Phantom precision experiments show that these SPGR/DESS combinations exhibit trends in pooled sample standard deviations that reasonably reflect CRB predictions.

In vivo experiments suggest that with optimization, the (0, 2) profile can yield comparable \hat{T}_1, \hat{T}_2 precision to the more conventional (2, 1) [4] scan profile in well-isolated WM/GM ROIs; however, the (0, 2) T_1 estimates are unreliable within and near the CSF

and do not agree with IR measurements in WM as consistently as the (2, 1) profile. This and other disagreements across profiles *in vivo* may be attributable to differences in signal model sensitivities to neglected higher-order effects. Nevertheless, the example application studied in this chapter illustrates that scan optimization can enable new parameter mapping techniques from established pulse sequences.

CHAPTER 5

MRI Parameter Estimation via Kernel Regression

{c,krr}

5.1 Introduction

{s,krr,intro}

In MRI *parameter estimation*, one seeks to quantify biomarker “maps” (*i.e.*, parameter images) from data. Because MRI acquisitions are tunably sensitive to many physical processes (*e.g.*, relaxation [12], diffusion [97], and chemical exchange [98]), MRI parameter estimation is important in many QMRI applications (*e.g.*, relaxometry [99], diffusivity tensor imaging [100], and multi-compartmental imaging [93]). Motivated by such widespread applications, this chapter describes a method for fast MRI parameter estimation.

Chapter 3 applied a common parameter estimation strategy to QMRI that involves minimization of an objective function related to the likelihood function (and possibly regularization terms). Because MR signal models are typically nonlinear functions of latent object parameters, such likelihood-based estimation requires non-convex optimization in general. To seek global optima, several of our [3, 4, 5] and others’ [101, 95, 49, 102] works approach estimation via exhaustive grid search, which requires either storing or computing on-the-fly a “dictionary” of scan profile signal vectors. These works estimate a small (2-3) number of latent parameters, and so grid search is practical. However, for even moderately sized problems, the required number of dictionary elements renders grid search undesirable or even intractable, unless one assumes artificially restrictive latent parameter constraints.

There are numerous QMRI applications that could benefit from an alternative MRI parameter estimation method that scales well with the number of latent parameters. For example, vector (*e.g.*, flow [103]) and tensor (*e.g.*, diffusivity [100] or conductivity [104]) field mapping techniques require estimation of at minimum 4 and 7 latent parameters per voxel, respectively. Phase-based longitudinal [105] or transverse [106, 83] field mapping could avoid noise-amplifying algebraic manipulations on reconstructed image data that are

conventionally used to reduce signal dependencies on nuisance latent parameters. Compartmental fraction mapping [93] from steady-state pulse sequences requires estimation of at least 7 [107] and as many as 10 [108] latent parameters per voxel. In these and other applications, greater estimation accuracy requires more complete signal models that involve more latent parameters, which only increases the need for scalable estimation methods.

The fundamental challenge of scalable MRI parameter estimation stems from MR signal model nonlinearity: standard linear estimators would be scalable but also inaccurate. One natural solution strategy involves preprocessing the image data such that the transformed data is at least approximately linear in the latent parameters. Without further insight however, such transformation could dramatically increase problem dimensionality, hindering scalability. Fortunately, a celebrated result in approximation theory [109] showed that simple transformations involving so-called *reproducing kernel* functions [110] could accurately represent solutions to a relevant class of nonlinear problems while scaling the associated linear problem’s dimensionality in a mild, controlled manner. These kernel methods later found popularity in machine learning (initially for classification [111] and quickly thereafter for other applications, *e.g.* regression [112]) because they provided simple, scalable nonlinear extensions to otherwise linear algorithms.

This chapter introduces a fast, scalable method for nonlinear MRI parameter estimation via kernel ridge regression (KRR). We observe that for voxel-wise separable MRI parameter estimation problems, one can rapidly simulate many instances of latent parameter inputs and signal outputs from the nonlinear signal model. We take such input-output pairs as simulated *training points* and propose to then *learn* (using an appropriate kernel function) a nonlinear *regression function* (*i.e.*, non-iterative estimator) from the training points. The proposed KRR-based estimator scales considerably better with the number of estimated latent parameters than previously-discussed likelihood-based estimators.

The remainder of this chapter is organized as follows. Section 5.2 reviews the general signal model for an MR scan profile, constructs an appropriate functional optimization problem, and efficiently solves this problem using kernels. Section 5.3 further reduces computational requirements via a kernel approximation and provides guidelines for model selection. Section 5.4 applies KRR-based estimation to quantify six parameters arising from models describing the steady-state magnetization dynamics of two water compartments, a challenging application of interest in myelin water fraction imaging (discussed in Chapter 6). Section 5.5 discusses possible extensions and provides concluding remarks.

5.2 A Function Optimization Problem & Kernel Solution

{s,krr, meth}

Recall from Section 3.2.1 that after image reconstruction, many MRI acquisitions produce at each voxel position a sequence of noisy voxel values $\mathbf{y} \in \mathbb{C}^D$, modeled as

{eq:krr,model}

$$\mathbf{y} = \mathbf{s}(\mathbf{x}; \boldsymbol{\nu}) + \boldsymbol{\epsilon}, \quad (5.1)$$

where $\mathbf{x} \in \mathbb{R}^L$ denotes L *latent* object parameters (*e.g.*, relaxation time constants); $\boldsymbol{\nu} \in \mathbb{R}^K$ denotes K *known* object parameters (*e.g.*, separately acquired and estimated field maps); $\mathbf{s} : \mathbb{R}^L \times \mathbb{R}^K \mapsto \mathbb{C}^D$ models the noiseless signals that arise from D datasets; and $\boldsymbol{\epsilon} \in \mathbb{C}^D$ is complex Gaussian noise, assumed to be distributed as $\mathcal{CN}(\mathbf{0}_D, \boldsymbol{\Sigma})$. Unlike signal models in earlier chapters (*cf.* (4.1)), Equation 5.1 omits for simplicity explicit model dependence on acquisition parameters, as these are fixed during parameter estimation.

Here we seek to estimate on a per-voxel basis each latent parameter \mathbf{x} from corresponding data sequence \mathbf{y} and known parameter $\boldsymbol{\nu}$. To develop (or train) a non-iterative estimator $\hat{\mathbf{x}}$, we simulate many instances of input-output relation (5.1) and use kernels to develop a nonlinear inverse relation. We sample $\mathbb{R}^L \times \mathbb{R}^K \times \mathbb{C}^D$ and evaluate (5.1) N times to produce sets of training inputs $\{(\mathbf{x}_1, \boldsymbol{\nu}_1, \boldsymbol{\epsilon}_1), \dots, (\mathbf{x}_N, \boldsymbol{\nu}_N, \boldsymbol{\epsilon}_N)\}$ and data sequences $\{\mathbf{y}_1, \dots, \mathbf{y}_N\}$. We seek a function $\hat{\mathbf{h}} : \mathbb{R}^Q \mapsto \mathbb{R}^L$ for $Q := 2D + K$ and an offset $\hat{\mathbf{b}} \in \mathbb{R}^L$ that together map each pure-real regressor $\mathbf{q}_n := [\text{Re}(\mathbf{y}_n)^\top, \text{Im}(\mathbf{y}_n)^\top, \boldsymbol{\nu}_n^\top]^\top$, $n \in \{1, \dots, N\}$, to an estimate $\hat{\mathbf{x}}(\mathbf{q}_n) := \hat{\mathbf{h}}(\mathbf{q}_n) + \hat{\mathbf{b}}$ that is “close” to corresponding regressand \mathbf{x}_n :

{eq:krr,prob}

$$(\hat{\mathbf{h}}, \hat{\mathbf{b}}) \in \left\{ \arg \min_{\substack{\mathbf{h} \in \mathbb{H}^L \\ \mathbf{b} \in \mathbb{R}^L}} \Psi(\mathbf{h}, \mathbf{b}; (\mathbf{x}_1, \mathbf{q}_1), \dots, (\mathbf{x}_N, \mathbf{q}_N)) \right\}, \text{ where} \quad (5.2)$$

{eq:krr,cost}

$$\Psi(\dots) = \sum_{l=1}^L \Psi_l(h_l, b_l; (x_{l,1}, \mathbf{q}_1), \dots, (x_{l,N}, \mathbf{q}_N)); \quad (5.3)$$

{eq:krr,cost-1}

$$\Psi_l(\dots) = \frac{1}{N} \sum_{n=1}^N (h_l(\mathbf{q}_n) + b_l - x_{l,n})^2 + \rho_l \|h_l\|_{\mathbb{H}}^2. \quad (5.4)$$

Here, each $h_l : \mathbb{R}^Q \mapsto \mathbb{R}$ is a scalar function that maps to the l th component of the output of \mathbf{h} ; each $b_l, x_{l,n} \in \mathbb{R}$ are scalar components of \mathbf{b}, \mathbf{x} ; \mathbb{H} is a (presently unspecified) Hilbert function space, whose norm $\|\cdot\|_{\mathbb{H}}$ is induced by inner product $\langle \cdot, \cdot \rangle_{\mathbb{H}} : \mathbb{H} \times \mathbb{H} \mapsto \mathbb{R}$; each ρ_l controls for regularity in h_l ; and $(\cdot)^\top$ denotes vector transpose.

Since (5.3) is separable in the components of \mathbf{h} , it suffices to consider optimizing each (h_l, b_l) by separately minimizing (5.4) for each $l \in \{1, \dots, L\}$. Nevertheless, minimizing (5.4) is a challenging functional optimization problem without further restriction on

the Hilbert space \mathbb{H} . Following many other works involving kernels, we proceed by hereafter restricting \mathbb{H} to be a *reproducing kernel* Hilbert space (RKHS) [110]. Remarkably, a generalization [113] of the Representer Theorem [109], restated as is relevant here for completeness, then reduces (5.4) to a finite-dimensional optimization problem:

Theorem 1 (Generalized Representer, [113]) *Define $k : \mathbb{R}^Q \times \mathbb{R}^Q \mapsto \mathbb{R}$ to be the (symmetric positive definite) kernel function associated with RKHS \mathbb{H} , such that reproducing property $h_l(\mathbf{q}) = \langle h_l, k(\cdot, \mathbf{q}) \rangle_{\mathbb{H}}$ holds for all $h_l \in \mathbb{H}$ and $\mathbf{q} \in \mathbb{R}^Q$. Then any minimizer (\hat{h}_l, \hat{b}_l) of (5.4) over $\mathbb{H} \times \mathbb{R}$ admits a representation of \hat{h}_l of the form*

{thm:krr,rep}

$$\hat{h}_l(\cdot) \equiv \sum_{n=1}^N a_{l,n} k(\cdot, \mathbf{q}_n), \quad (5.5)$$

{eq:krr,rep}

where each $a_{l,n} \in \mathbb{R}$ for $n \in \{1, \dots, N\}$.

Theorem 1 ensures that any solution to

$$(\hat{\mathbf{a}}_l, \hat{b}_l) \in \left\{ \arg \min_{\substack{\mathbf{a}_l \in \mathbb{R}^N \\ b_l \in \mathbb{R}}} \frac{1}{N} \sum_{n=1}^N \left(\sum_{n'=1}^N a_{l,n'} k(\mathbf{q}_n, \mathbf{q}_{n'}) + b_l - x_{l,n} \right)^2 + \rho_l \left\| \sum_{n'=1}^N a_{l,n'} k(\cdot, \mathbf{q}_{n'}) \right\|_{\mathbb{H}}^2 \right\} \quad (5.6)$$

{eq:krr,cvx}

corresponds via (5.5) to a minimizer of (5.4) over $\mathbb{H} \times \mathbb{R}$, where $\mathbf{a}_l := [a_{l,1}, \dots, a_{l,N}]^T$. Fortunately, a solution of (5.6) exists uniquely for $\rho_l > 0$ and can be expressed as

{eq:krr,a-hat}

$$\hat{\mathbf{a}}_l = (\mathbf{M}\mathbf{K} + N\rho_l \mathbf{I}_N)^{-1} \mathbf{M}\mathbf{x}_l; \quad (5.7)$$

{eq:krr,b-hat}

$$\hat{b}_l = \frac{1}{N} \mathbf{1}_N^T (\mathbf{x}_l - \mathbf{K}\hat{\mathbf{a}}_l), \quad (5.8)$$

where $\mathbf{K} \in \mathbb{R}^{N \times N}$ is the Gram matrix consisting of entries $k(\mathbf{q}_n, \mathbf{q}_{n'})$ for $n, n' \in \{1, \dots, N\}$; $\mathbf{M} := \mathbf{I}_N - \frac{1}{N} \mathbf{1}_N \mathbf{1}_N^T$ is a de-meaning operator; $\mathbf{x}_l := [x_{l,1}, \dots, x_{l,N}]^T$; $\mathbf{I}_N \in \mathbb{R}^{N \times N}$ is the identity matrix; and $\mathbf{1}_N \in \mathbb{R}^N$ is a vector of ones. In the special case where each $\rho_l \leftarrow \rho$ for fixed $\rho > 0$, corresponding scalar estimators $\left\{ \hat{x}_l(\cdot) := \hat{h}_l(\cdot) + \hat{b}_l \right\}_1^L$ that arise from plugging (5.7) into (5.5) can be concisely concatenated as

{eq:krr,x-hat}

$$\hat{\mathbf{x}}(\cdot) \leftarrow \mathbf{X} \left(\frac{1}{N} \mathbf{1}_N + \mathbf{M}(\mathbf{K}\mathbf{M} + N\rho \mathbf{I}_N)^{-1} \mathbf{k}(\cdot) \right), \quad (5.9)$$

where $\mathbf{k}(\cdot) := [k(\cdot, \mathbf{q}_1), \dots, k(\cdot, \mathbf{q}_N)]^T - \frac{1}{N} \mathbf{K} \mathbf{1}_N : \mathbb{R}^Q \mapsto \mathbb{R}^N$ is a kernel embedding operator and $\mathbf{X} := [\mathbf{x}_1, \dots, \mathbf{x}_N] = [\mathbf{x}_1, \dots, \mathbf{x}_L]^T \in \mathbb{R}^{L \times N}$ collects the regressands.

For $\rho > 0$, estimator (5.9) minimizes (5.3) over $(\mathbb{H} \times \mathbb{R})^L$. However, the utility of (5.9) depends on the choice of kernel k , which induces a choice on the RKHS \mathbb{H} and thus the function space $(\mathbb{H} \times \mathbb{R})^L$ over which (5.2) optimizes. For example, if k was selected as the canonical dot product $k(\mathbf{q}, \mathbf{q}') \leftarrow \langle \mathbf{q}, \mathbf{q}' \rangle_{\mathbb{R}^Q} := \mathbf{q}^\top \mathbf{q}'$ (for which RKHS $\mathbb{H} \leftarrow \mathbb{R}^Q$), then (5.9) would reduce to affine ridge regression [114], which is optimal over $(\mathbb{R}^Q \times \mathbb{R})^L$ but is unlikely to be useful when signal model \mathbf{s} is nonlinear in \mathbf{x} .

Since we expect a useful estimate $\hat{\mathbf{x}}(\mathbf{q})$ to depend nonlinearly (but smoothly) on \mathbf{q} , we instead use a (symmetric, positive definite) k that is likewise nonlinear in its arguments and thus corresponds to a RKHS richer than \mathbb{R}^Q . Specifically, we use a Gaussian kernel

$$\{eq:krr,kern\} \quad k(\mathbf{q}, \mathbf{q}') \leftarrow \exp \left(-\frac{1}{2} \|\mathbf{q} - \mathbf{q}'\|_{\Lambda^{-2}}^2 \right), \quad (5.10)$$

where symmetric, positive definite $\Lambda \in \mathbb{R}^{Q \times Q}$ controls the length scales in \mathbf{q} over which the estimator $\hat{\mathbf{x}}$ smooths. We focus on the Gaussian kernel over other candidate kernels because of its popularity in the machine learning community for representing smooth functions.

Interestingly, the RKHS associated with Gaussian kernel (5.10) is infinite-dimensional. Thus, Gaussian KRR can be interpreted as first “lifting” via a nonlinear *feature map* $\mathbf{z} : \mathbb{R}^Q \mapsto \mathbb{H}$ each \mathbf{q} into an infinite-dimensional *feature* $\mathbf{z}(\mathbf{q}) \in \mathbb{H}$, and then performing affine ridge regression on the features via inner products of the form $k(\mathbf{q}, \mathbf{q}') \leftarrow \langle \mathbf{z}(\mathbf{q}), \mathbf{z}(\mathbf{q}') \rangle_{\mathbb{H}}$. From this perspective, the challenges of nonlinear estimation via likelihood models are avoided because we *select* (through the choice of kernel) characteristics of the nonlinear dependence that we wish for our regression function to model and need only *estimate* via (5.6) the linear dependence of $\hat{\mathbf{x}}$ on the corresponding features.

5.3 Implementation Considerations

$\{s,krr,pract\}$

This section focuses on important practical implementation issues. Subsection 5.3.1 discusses a scalable and conceptually intuitive approximation of KRR estimator (5.9). Subsection 5.3.2 describes guidelines for nearly-automatic model selection.

5.3.1 A Kernel Approximation

$\{ss,krr,pract,apprx\}$

In practical problems with even moderately large ambient dimension Q , the necessarily large number of training samples N complicates storage of (dense) Gram matrix \mathbf{K} . Using a kernel approximation can mitigate storage issues. Here we choose to sample Random Fourier Features (RFF) [115], a recent method for approximating shift-invariants kernels

having form $k(\mathbf{q}, \mathbf{q}') \equiv k(\mathbf{q} - \mathbf{q}')$. This subsection reviews the main result of [115] for the purpose of constructing an intuitive and computationally efficient approximation of (5.9).

The strategy of [115] is to construct independent probability distributions p_v and p_s associated with random $\mathbf{v} \in \mathbb{R}^Q$ and random $s \in \mathbb{R}$ as well as a random function (that is parameterized by \mathbf{q}) $\tilde{z}(\cdot, \cdot; \mathbf{q}) : \mathbb{R}^Q \times \mathbb{R} \times \mathbb{R}^Q \mapsto \mathbb{R}$, such that

$$\{eq:krr,exp\} \quad \mathbb{E}_{v,s}(\tilde{z}(\mathbf{v}, s; \mathbf{q})\tilde{z}(\mathbf{v}, s; \mathbf{q}')) = k(\mathbf{q} - \mathbf{q}'), \quad (5.11)$$

where $\mathbb{E}_{v,s}(\cdot)$ denotes expectation with respect to $p_v p_s$. If such a construction exists, one can build approximate feature maps $\tilde{\mathbf{z}}_Z$ by concatenating evaluations of $\tilde{z}_Z := \sqrt{2/Z}\tilde{z}$ on Z samples $\{(\mathbf{v}_1, s_1), \dots, \mathbf{v}_Z, s_Z\}$ of (\mathbf{v}, s) (drawn jointly albeit independently), to produce approximate feature vectors

$$\{eq:krr,feat\} \quad \tilde{\mathbf{z}}_Z(\mathbf{q}) := [\tilde{z}_Z(\mathbf{v}_1, s_1; \mathbf{q}), \dots, \tilde{z}_Z(\mathbf{v}_Z, s_Z; \mathbf{q})]^\top \quad (5.12)$$

for any \mathbf{q} . Then by the strong law of large numbers,

$$\{eq:krr,lln\} \quad \lim_{Z \rightarrow \infty} \langle \tilde{\mathbf{z}}_Z(\mathbf{q}), \tilde{\mathbf{z}}_Z(\mathbf{q}') \rangle_{\mathbb{R}^Z} \xrightarrow{a.s.} k(\mathbf{q}, \mathbf{q}') \quad \forall \mathbf{q}, \mathbf{q}', \quad (5.13)$$

which, in conjunction with strong performance guarantees for finite Z [115, 116], justifies the interpretation of $\tilde{\mathbf{z}}_Z$ as an approximate (and now finite-dimensional) feature map.

We use the Fourier construction of [115] that assigns $\tilde{z}(\mathbf{v}, s; \mathbf{q}) \leftarrow \cos(2\pi(\mathbf{v}^\top \mathbf{q} + s))$. If $s \sim \text{unif}(0, 1)$, then $\mathbb{E}_{v,s}(\tilde{z}(\mathbf{v}, s; \mathbf{q})\tilde{z}(\mathbf{v}, s; \mathbf{q}'))$ simplifies to

$$\{eq:krr,ft\} \quad \int_{\mathbb{R}^Q} \cos(2\pi \mathbf{v}^\top (\mathbf{q} - \mathbf{q}')) p_v(\mathbf{v}) \, d\mathbf{v}. \quad (5.14)$$

For symmetric positive definite k , (5.14) exists [117] and is the Fourier transform of p_v . Thus, for Gaussian kernel (5.10), choosing $\mathbf{v} \sim \mathcal{N}(\mathbf{0}_Q, (2\pi\Lambda)^{-2})$ satisfies (5.11).

Subsequent sampling of \mathbf{v}, s and construction of $\tilde{\mathbf{Z}}_Z := [\tilde{\mathbf{z}}_Z(\mathbf{q}_1), \dots, \tilde{\mathbf{z}}_Z(\mathbf{q}_N)] \in \mathbb{R}^{Z \times N}$ via (5.12) gives for $Z \ll N$ a low-rank approximation $\tilde{\mathbf{Z}}_Z^\top \tilde{\mathbf{Z}}_Z$ of Gram matrix \mathbf{K} . Substituting this approximation into (5.9) and applying the matrix inversion lemma [118] yields

$$\{eq:krr,x-apx\} \quad \hat{\mathbf{x}}(\cdot) \leftarrow \mathbf{m}_x + \mathbf{C}_{x\tilde{\mathbf{z}}}(\mathbf{C}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}} + \rho \mathbf{I}_Z)^{-1}(\tilde{\mathbf{z}}_Z(\cdot) - \mathbf{m}_{\tilde{\mathbf{z}}}), \quad (5.15)$$

where $\mathbf{m}_x := \frac{1}{N} \mathbf{X} \mathbf{1}_N$ and $\mathbf{m}_{\tilde{\mathbf{z}}} := \frac{1}{N} \tilde{\mathbf{Z}}_Z \mathbf{1}_N$ are sample mean vectors; and $\mathbf{C}_{x\tilde{\mathbf{z}}} := \frac{1}{N} \mathbf{X} \mathbf{M} \tilde{\mathbf{Z}}_Z^\top$ and $\mathbf{C}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}} := \frac{1}{N} \tilde{\mathbf{Z}}_Z \mathbf{M} \tilde{\mathbf{Z}}_Z^\top$ are sample covariance matrices. Estimator (5.15) is a regularized variation of the linear minimum mean-squared error estimator on the features, and illus-

trates that Gaussian KRR via estimator (5.9) is asymptotically (in Z) equivalent to affine ridge regression after nonlinear, high-dimensional feature mapping.

{ss,krr,pract,mod}

5.3.2 Model Selection

This subsection proposes guidelines for data-driven model selection through a mixture of intuitive arguments and referenced analytical results. Our goal here is to leave for uninformed manual tuning as few model parameters as possible.

5.3.2.1 Choosing Sampling Distributions

{sss,krr,pract,mod,dist}

To sample training points, one must select prior distributions on object parameters \mathbf{x} , $\boldsymbol{\nu}$ and noise vector ϵ . If offline training is required, all of these distributions of course must either be assumed or empirically estimated from previous experiments. Here we focus on online training, where one trains alongside data processing.

Latent object parameter distributions must account for model physics (to focus samples on useful parameter ranges) while also leaving generously heavy distribution tails (to enable precise estimation over a broad parameter range). Known object parameter distributions can be estimated directly from known $\boldsymbol{\nu}$ via density estimation methods, *e.g.* [119]. Noise covariance is well-estimated from low-signal regions of test data.

5.3.2.2 Choosing Regularization Parameters

{sss,krr,pract,mod,rho}

Regularization parameter selection is well guided by considering KRR from an alternative Bayesian formulation. In this perspective, the unknown regression function $\mathbf{h}(\cdot) + \mathbf{b}$ arises randomly from a Gaussian process prior distribution with mean function zero and covariance function given by the kernel. The observed regressand model $\mathbf{x}(\mathbf{q}) = \mathbf{h}(\mathbf{q}) + \mathbf{b} + \epsilon_{\mathbf{x}}$ includes additive *latent parameter noise* $\epsilon_{\mathbf{x}}$ because noise is *anticipated* in testing data and is thus intentionally *simulated* in training data. If this latent parameter noise is assumed to be distributed as $\mathcal{N}\left(\mathbf{0}_L, N \text{diag}\left([\rho_1, \dots, \rho_L]^\top\right)\right)$, then the mean function of the regression function's posterior distribution (after observation of training points) happens to equal the KRR solution (5.7)-(5.8) for each $l \in \{1, \dots, L\}$ (see [120, Ch. 2] for derivations). In a Bayesian perspective then, $N\rho_1, \dots, N\rho_L$ should reflect latent parameter noise covariance.

In MRI parameter estimation, we can more easily measure noise covariance in image data Σ than in latent parameters. We relate these two covariances in the following.

A first-order Taylor expansion of the signal model around $\mathbf{E}_{\mathbf{x}}(\mathbf{x})$ is

{eq:krr,model-approx}

$$\mathbf{y} \approx \mathbf{E}_{\mathbf{y}}(\mathbf{y}) + [\nabla_{\mathbf{x}} \mathbf{s}(\mathbf{E}_{\mathbf{x}}(\mathbf{x}), \boldsymbol{\nu})](\mathbf{x} - \mathbf{E}_{\mathbf{x}}(\mathbf{x})). \quad (5.16)$$

Substituting (5.16) into a covariance definition gives

$$\begin{aligned} \Sigma &= \mathbb{E}_{\mathbf{y}} \left((\mathbf{y} - \mathbb{E}_{\mathbf{y}}(\mathbf{y}))(\mathbf{y} - \mathbb{E}_{\mathbf{y}}(\mathbf{y}))^H \right) \\ &\approx [\nabla_{\mathbf{x}} \mathbf{s}(\mathbb{E}_{\mathbf{x}}(\mathbf{x}), \boldsymbol{\nu})] \text{cov}(\mathbf{x}) [\nabla_{\mathbf{x}} \mathbf{s}(\mathbb{E}_{\mathbf{x}}(\mathbf{x}), \boldsymbol{\nu})]^H. \end{aligned} \quad (5.17)$$

Taking the Moore-Penrose pseudoinverse (denoted by $(\cdot)^\dagger$) of each side yields

$$\begin{aligned} &[\nabla_{\mathbf{x}} \mathbf{s}(\mathbb{E}_{\mathbf{x}}(\mathbf{x}), \boldsymbol{\nu})]^H (\Sigma)^{-1} [\nabla_{\mathbf{x}} \mathbf{s}(\mathbb{E}_{\mathbf{x}}(\mathbf{x}), \boldsymbol{\nu})] \approx \\ &[\nabla_{\mathbf{x}} \mathbf{s}(\mathbb{E}_{\mathbf{x}}(\mathbf{x}), \boldsymbol{\nu})]^H \left([\nabla_{\mathbf{x}} \mathbf{s}(\mathbb{E}_{\mathbf{x}}(\mathbf{x}), \boldsymbol{\nu})]^H \right)^\dagger (\text{cov}(\mathbf{x}))^{-1} ([\nabla_{\mathbf{x}} \mathbf{s}(\mathbb{E}_{\mathbf{x}}(\mathbf{x}), \boldsymbol{\nu})])^\dagger [\nabla_{\mathbf{x}} \mathbf{s}(\mathbb{E}_{\mathbf{x}}(\mathbf{x}), \boldsymbol{\nu})]. \end{aligned} \quad (5.18)$$

If the Jacobian matrix $[\nabla_{\mathbf{x}} \mathbf{s}(\mathbb{E}_{\mathbf{x}}(\mathbf{x}), \boldsymbol{\nu})]$ is of full column rank (and is thus left-invertible), Eq. (5.18) simplifies after matrix inversion to

$$\text{cov}(\mathbf{x}; \boldsymbol{\nu}) \approx \left([\nabla_{\mathbf{x}} \mathbf{s}(\mathbb{E}_{\mathbf{x}}(\mathbf{x}), \boldsymbol{\nu})]^H (\Sigma)^{-1} [\nabla_{\mathbf{x}} \mathbf{s}(\mathbb{E}_{\mathbf{x}}(\mathbf{x}), \boldsymbol{\nu})] \right)^{-1}, \quad (5.19)$$

an expression reminiscent of the Cramér-Rao Bound (*cf.* Eq. (4.3)-(4.4)).

Expression (5.19) approximates latent parameter noise covariance but is not yet useful because it remains a function of known parameter $\boldsymbol{\nu}$, which varies spatially. As a simple solution, one could set regularization parameters ρ_1, \dots, ρ_L using diagonal elements of

$$\left(N [\nabla_{\mathbf{x}} \mathbf{s}(\mathbb{E}_{\mathbf{x}}(\mathbf{x}), \bar{\boldsymbol{\nu}})]^H (\Sigma)^{-1} [\nabla_{\mathbf{x}} \mathbf{s}(\mathbb{E}_{\mathbf{x}}(\mathbf{x}), \bar{\boldsymbol{\nu}})] \right)^{-1}, \quad (5.20)$$

where $\mathbb{E}_{\mathbf{x}}(\mathbf{x})$ is taken with respect to the distributions used to sample training points and $\bar{\boldsymbol{\nu}} := \frac{1}{V} \sum_{v=1}^V \boldsymbol{\nu}(\mathbf{r}_v)$ is the sample mean of the known object parameters.

Regularization parameter selection via (5.20) is extremely fast but requires signal model gradients that might be undesirable to compute or altogether unavailable. As an alternative, one could compute the sample covariance matrix $\mathbf{C}_{\mathbf{xx}} := \frac{1}{N} \mathbf{X} \mathbf{M} \mathbf{X}^T$ from simulated training data points and set ρ_1, \dots, ρ_L as diagonal elements of $\frac{1}{N} \mathbf{C}_{\mathbf{xx}}$.

5.3.2.3 Choosing Smoothing Length Scale

With online training, it is intuitive to set the Gaussian kernel's smoothing length scale based on test data sample means $\bar{\mathbf{y}} := \frac{1}{V} \sum_{v=1}^V \mathbf{y}(\mathbf{r}_v)$ and known parameter sample means $\bar{\boldsymbol{\nu}}$:

$$\Lambda \leftarrow \lambda_0 \text{diag} \left(\left[\text{Re}(\bar{\mathbf{y}})^T, \text{Im}(\bar{\mathbf{y}})^T, \bar{\boldsymbol{\nu}}^T \right]^T \right). \quad (5.21)$$

Here, $\lambda_0 > 0$ remains unspecified as a unitless scalar free parameter that balances training point consistency and estimator regularity.

5.3.2.4 Choosing Kernel Approximation Order

Selection of kernel approximation order Z can be guided through performance analysis. To this end, we apply a bound on the expected maximal approximation error of RFF [115], studied in Proposition 4 of [116] and restated in our notation below for convenience. We examine this bound over others because it requires intuitive selection of only one error criterion; tighter bounds are available, but may not be as simple to use.

Theorem 2 ([116]) *Let \mathbb{Q} be a compact subset of \mathbb{R}^Q with diameter $d_{\mathbb{Q}} := \sup_{\mathbf{q}, \mathbf{q}' \in \mathbb{Q}} \|\mathbf{q} - \mathbf{q}'\|_2$. Define $k : \mathbb{Q} \times \mathbb{Q} \mapsto \mathbb{R}$ to be a shift-invariant symmetric positive definite kernel, such that $k(\mathbf{q}, \mathbf{q}') \equiv k(\mathbf{q} - \mathbf{q}')$ for each $\mathbf{q}, \mathbf{q}' \in \mathbb{Q}$. Normalize $k(\mathbf{0}_Q) = 1$ and ensure $\nabla^2 k(\mathbf{0}_Q)$ exists. Further suppose k is v -Lipschitz over $\mathbb{Q}_{\Delta} := \{\mathbf{q} - \mathbf{q}' | \mathbf{q}, \mathbf{q}' \in \mathbb{Q}\}$. Construct approximate feature vectors $\tilde{\mathbf{z}}_Z$ (defined in (5.12)) from Z independently drawn samples of $(\boldsymbol{\nu}, s)$, where $\boldsymbol{\nu}$ is distributed with a probability density defined by the Fourier transform of k and $s \sim \text{unif}(0, 1)$. Take $r := \mathbb{E}_{\boldsymbol{\nu}}(\max(\|\boldsymbol{\nu}_1\|_2, \dots, \|\boldsymbol{\nu}_Z\|_2))$. Assume \mathbb{Q} and Z are not so small as to prevent $\tilde{\mathbf{z}}_Z(\mathbf{q})^T \tilde{\mathbf{z}}_Z(\mathbf{q}') - k(\mathbf{q}, \mathbf{q}')$ from crossing zero, as is typical. Then*

$$\mathbb{E} \left(\sup_{\mathbf{q}, \mathbf{q}' \in \mathbb{Q}} \left| \tilde{\mathbf{z}}_Z(\mathbf{q})^T \tilde{\mathbf{z}}_Z(\mathbf{q}') - k(\mathbf{q}, \mathbf{q}') \right| \right) \leq \frac{48c_{\mathbb{Q}}d_{\mathbb{Q}}\sqrt{Q}}{\sqrt{Z}}(v + r), \quad (5.22)$$

where $c_{\mathbb{Q}} \in [0.803, 1.542]$ achieves its lower/upper bound for \mathbb{Q} a ball/sphere.

For concreteness, we apply Theorem 2 explicitly to Gaussian kernel (5.10). It suffices to consider \mathbb{Q} a ball with diameter $d_{\mathbb{Q}} \leftarrow 2 \max(\|\mathbf{q}(\mathbf{r}_1)\|_2, \dots, \|\mathbf{q}(\mathbf{r}_V)\|_2)$. Then $c_{\mathbb{Q}} \leftarrow 0.803$. The Gaussian kernel admits a global Lipschitz constant as

$$v := \sup_{\Delta \in \mathbb{Q}_{\Delta}} \|\nabla_{\Delta} k\|_2 \leq \sup_{\Delta \in \mathbb{R}^Q} \|\nabla_{\Delta} k\|_2 = \sqrt{\frac{Q}{e \text{tr}(\boldsymbol{\Lambda}^2)}}. \quad (5.23)$$

The best lower bound we have presently for r is

$$r \leq \frac{1}{2\pi \min \text{eig}(\boldsymbol{\Lambda})} \left(\sqrt{2 \log Z} + \sqrt{Q} \right), \quad (5.24)$$

where $\min \text{eig}(\cdot)$ extracts the minimum eigenvalue. Our overall bound is then

$$\begin{aligned} & \mathbb{E} \left(\sup_{\mathbf{q}, \mathbf{q}' \in \mathbb{Q}} \left| \tilde{\mathbf{z}}_Z(\mathbf{q})^\top \tilde{\mathbf{z}}_Z(\mathbf{q}') - k(\mathbf{q}, \mathbf{q}') \right| \right) \\ & \leq \frac{96(0.803)(\max_v(\|\mathbf{q}(\mathbf{r}_v)\|_2))\sqrt{Q}}{\sqrt{Z}} \left(\sqrt{\frac{Q}{e \text{tr}(\mathbf{\Lambda}^2)}} + \frac{\sqrt{2 \log Z} + \sqrt{Q}}{2\pi \min \text{eig}(\mathbf{\Lambda})} \right) \end{aligned} \quad (5.25)$$

$$\leq \frac{96(0.803)(\max_v(\|\mathbf{q}(\mathbf{r}_v)\|_2))\sqrt{Q}}{\min \text{eig}(\mathbf{\Lambda}) \sqrt{Z}} \left(\frac{1}{\sqrt{e}} + \frac{\sqrt{2 \log Z} + \sqrt{Q}}{2\pi} \right). \quad (5.26)$$

Error bound (5.25) is conservative due to large constants, but at least ensures asymptotically that scaling Z quadratically with Q will maintain a given expected maximal approximation error tolerance. Error bound (5.26) is looser, but more clearly shows that selecting $\mathbf{\Lambda}$ via (5.21) makes the bound somewhat invariant to the scale of test data.

5.4 Experimentation

{s,krr,exp}

As proof of concept, we apply kernel-based estimation to quantify parameters describing the magnetization dynamics of multiple water compartments, a challenging application of clinical interest, *e.g.*, for myelin water imaging (application details in Chapter 6). Specifically, we use a simple model of two non-exchanging compartments and seek to estimate the associated latent parameters; more complex models would only add parameters and thereby increase the need for an alternative to grid search.

Such challenging estimation problems typically require multiple data acquisitions and thus long scans. To reduce scan times of classical methods [93], fast steady-state pulse sequences were recently proposed for two-compartment parameter estimation [107, 122]. We similarly take interest in steady-state sequences, but modify acquisition details to address possible concerns of insufficient estimation precision [123].

We begin with two-compartment models of the signals arising from Spoiled Gradient-Recalled Echo (SPGR) [16] and Dual-Echo Steady-State (DESS) [19] pulse sequences. We make appropriate assumptions to reduce model dependencies to seven free parameters per voxel: flip angle spatial variation (due to transmit field inhomogeneity) s^t ; fast-relaxing compartmental fraction f_F ; (spin-lattice, spin-spin) relaxation time constants for the fast-relaxing $(T_{1,F}, T_{2,F})$ and slow-relaxing $(T_{1,S}, T_{2,S})$ compartments; and a complex proportionality constant¹ m_0 . We assume prior knowledge of $\nu \leftarrow s^t$ (which in practice can be estimated from separate fast acquisitions, *e.g.* [83]) and collect the remaining $L \leftarrow 6$ latent

¹We collect off-resonance effects in m_0 by approximating broadening distributions to be constant across compartments, as in prior works [107, 122]. We acknowledge this could lead to some bias in practice.

parameters as $\mathbf{x} \leftarrow [f_F, T_{1,F}, T_{2,F}, T_{1,S}, T_{2,S}, m_0]^\top$.

In light of clinical need, we focus on estimating fast-relaxing fraction f_F and tailor our simulation accordingly. We optimize (by methods similar to those described in Section 4.2) the flip angles and repetition times of four SPGR and three DESS scans² for precise estimation of f_F in white matter (WM) and grey matter (GM) regions of the human brain. We consider the other five latent parameters to be nuisance parameters and thus do not evaluate the performance of their estimators in the following.

We simulate data to arise from two non-exchanging water pools with nominal fast $(T_{1,F}, T_{2,F}) \leftarrow (500, 20)$ ms and slow $(T_{1,S}, T_{2,S}) \leftarrow (1000, 80)$ ms relaxation time constants selected from prior measurements [93, 122]. We assign fast-compartment fractions $f_F \leftarrow 0.15$ in WM and $f_F \leftarrow 0.03$ in GM and constrain corresponding slow-compartment fractions as $1 - f_F$. We prescribe these parameter values to the anatomy of the BrainWeb digital phantom [81] to produce ground truth parameter maps. Using optimized acquisition parameters and allowing s^t to model $\pm 20\%$ flip angle variation, we apply two-compartment SPGR and DESS models to the 81st axial slice of the true parameter maps. We corrupt these (complex) noiseless signals with additive complex Gaussian noise whose covariance $\Sigma \leftarrow (1.49 \times 10^{-7})\mathbf{I}_{D \leftarrow 10}$ reflects measurements from normalized datasets [5]. This yields realistically noisy 217×181 (image-domain) datasets ranging from 24.3-48.8dB SNR in WM and 26.4-49.5dB SNR in GM.

Because the first and second DESS signals depend differently on phase accrual due to off-resonance effects [5], off-resonance related phase (unlike signal loss) cannot be collected into m_0 . To avoid (separate or joint) estimation of an off-resonance field map, we elect to estimate parameters using magnitude SPGR and DESS image data and account for consequently Rician-distributed noise during training.

To sample training points, we assume prior distributions on latent object parameters \mathbf{x} (that distinguish the two compartments but are otherwise conservative) and directly measure known parameter ν and noise ϵ distributions from test data. We take f_F to be uniformly distributed on $[-0.1, 0.4]$ and $T_{1,F}, T_{2,F}, T_{1,S}, T_{2,S}$ to be log-uniformly distributed on $[50, 700]$ ms, $[5, 50]$ ms, $[700, 2000]$ ms, $[50, 300]$ ms, respectively. We match the scaling of test data in training by taking m_0 to be uniformly distributed on $[2.22 \times 10^{-16}, u]$, where u is set as $5 \times$ the maximum value of test data. We estimate the distribution of ν via kernel density estimation [119]. We assume noise covariance Σ of form $\sigma^2\mathbf{I}_{10}$ and compute noise variance estimate $\hat{\sigma}^2$ from Rayleigh-distributed noise regions of magnitude test data, using estimators described in [124].

²Since SPGR (DESS) yields one (two) signal(s) per excitation, four SPGR and three DESS scans produce a total $D \leftarrow 10$ datasets.

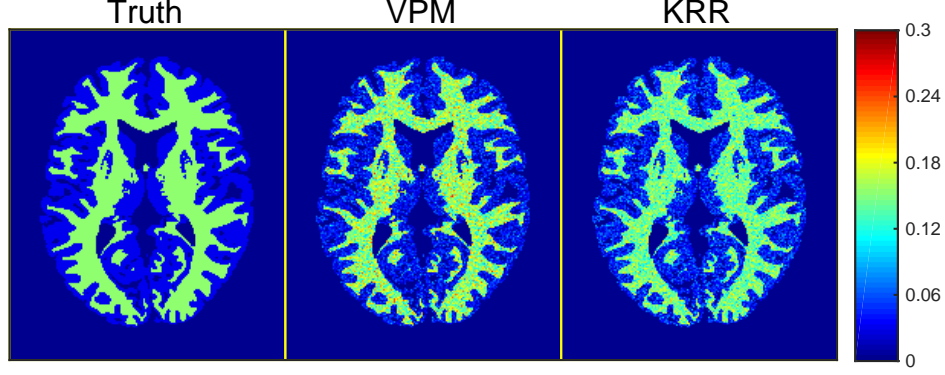


Figure 5.1: True f_F (left) and estimated \hat{f}_F fast-relaxing compartmental fraction maps, in simulation. Maximum-likelihood estimation via variable projection method and grid search (center) is accurate but is computationally expensive. In contrast, kernel ridge regression (right) is very fast and achieves comparable precision, at the expense of slightly increased bias (cf. Table 5.1). Voxels outside WM/GM regions are masked out in post-processing for visual clarity.

We sample $N \leftarrow 10^6$ training inputs from these distributions and use two-compartment SPGR/DESS signal models to evaluate corresponding (noisy, magnitude) responses. We set smoothing length scale Λ via (5.21), with $\lambda_0 \leftarrow 1$. We sample $(\mathbf{v}, s) \mathcal{Z} \leftarrow 10^3$ times to construct approximate feature mapping $\tilde{\mathbf{z}}_Z$. We apply $\tilde{\mathbf{z}}_Z$ to training data and compute sample means \mathbf{m}_x , $\mathbf{m}_{\tilde{\mathbf{z}}}$ and sample covariance matrices $\mathbf{C}_{x\tilde{\mathbf{z}}}$, $\mathbf{C}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}}$, and \mathbf{C}_{xx} . Lastly, we set ρ_1, \dots, ρ_L as diagonal entries of $\frac{1}{N}\mathbf{C}_{xx}$ and evaluate (5.15) for each latent parameter using test datasets and s^t on a per-voxel basis. On a 3.5GHz desktop computer with 32GB RAM running MATLAB[®] R2013a, training and estimating each took less than 40s and 2s.

Fig. 5.1 compares KRR estimates of fast-relaxing compartmental fraction f_F against not only ground truth maps but also maximum-likelihood estimates achieved via the “variable projection” method (VPM) [22] and grid search. As presented, the VPM estimate utilizes a dictionary of nearly 8×10^6 signal vectors computed using finely spaced samples on an unrealistically narrow feasible region consisting of a hypercube with boundaries set as $[-0.1, 0.4]$ in f_F and $\pm 20\%$ away from the truth in other dimensions. Using equal computational resources, estimation via VPM took nearly 4h.

Table 5.1 reports f_F sample statistics computed over 7810 WM-like and 9162 GM-like voxels. Overall, KRR and VPM achieve comparable estimation performance. In WM, KRR attains precision slightly higher than and accuracy similar to VPM. In GM, KRR attains precision comparable to and accuracy slightly lower than VPM. KRR attains root mean squared errors lower than VPM in WM (0.0229 versus 0.0295) and comparable to VPM in GM (0.0254 versus 0.0235).

	Truth	VPM	KRR
WM \hat{f}_F	0.15	0.1538 ± 0.0292	0.1440 ± 0.0221
GM \hat{f}_F	0.03	0.0336 ± 0.0232	0.0407 ± 0.0231

Table 5.1: Sample means \pm sample standard deviations of fast-relaxing compartmental fraction estimates \hat{f}_F , computed over simulated WM- and GM-like voxels. Each sample statistic is rounded off to the highest place value of its (unreported) standard error, which is computed via formulas in [2].

{tab:ff}

5.5 Summary and Future Work

{s,krr,summ}

This chapter is introducing a fast and computationally efficient method for MRI parameter estimation from nonlinear models via KRR. As proof of concept, we have thus far applied KRR to quantify in simulation fast-relaxing compartmental fraction f_F maps (along with several nuisance parameters) using two-compartment signal models of realistic SPGR and DESS acquisitions. In simulated \hat{f}_F WM/GM regions, KRR achieved comparable estimation performance as VPM-accelerated grid search, with dramatically reduced computation. Due to its generality, KRR could potentially accelerate MRI parameter estimation in many other applications, particularly those involving multiple latent parameters and/or cumbersome if not altogether unavailable signal models (as in, *e.g.*, MR fingerprinting [95]).

We foresee several avenues for ongoing research. Below, we discuss items of future work specific to the refinement of KRR for general MRI parameter estimation problems. Section ?? will discuss future work items specific to myelin water imaging.

Performance Analysis We assert (with some support from one simulation study) that scalability with the number of latent parameters L is the major computational advantage of KRR over conventional grid search. It is natural then to ask for the rate of this improvement. Specifically, we ask how slowly the number of training points N may scale with L (and regressor length Q) while still ensuring small test error.

We see two strategies to address this question. The first approach is through mathematical arguments. To this end, kernel ridge regression is one of several so-called *structural risk minimization* problems in statistical machine learning, for which well-established analytical results are available. While these results may aid in proving that our method has sub-exponential complexity in L (and thus outperforms grid search), the rates we find may be conservative. To construct a practical performance criterion, we may need to employ a second (likely more cumbersome) approach involving empirical parameter selection.

Exploiting Partially Linear Structure to Incorporate Scale Invariance Because there is ambiguity in the scale of MR data, it is desirable to construct a regression function that is unaffected by changes in data scale between training and testing. The present formulation addresses scaling ambiguity by setting the sampling distribution of proportionality constant m_0 based on test data, thereby matching simulated training data scale to test data scale. This strategy requires retraining between acquisitions, which may be undesirable in practice.

Since m_0 is the only latent parameter that is affected by data scale, it is natural to seek problem structure in m_0 to build scale invariance into KRR-based MRI parameter estimation. Here we surmise that appropriately leveraging the partial linearity of most MR signal models in m_0 will allow us to separate estimating scale-variant m_0 (via standard linear estimators) from estimating scale-invariant nonlinear parameters (via kernel regression).

Joint Image Reconstruction and Parameter Estimation We have thus far considered parameter estimation to be separate from image reconstruction, which affords fast data processing but may leave room for improved estimation, especially when raw data is undersampled. Alternatively, one could employ a procedure for joint image reconstruction and parameter estimation; one such algorithm iterates the following variable updates:

$$\{\text{eq:krr,joint-recon}\} \quad \mathbf{Y}^{(t)} \leftarrow \arg \min_{\mathbf{Y}} \|\mathbf{D} - \mathbf{Y}\mathbf{A}\|_{\text{F}}^2 + \eta^{(t)} \|\mathbf{S}^{(t-1)} - \mathbf{Y}\|_{\text{F}}^2 \quad (5.27)$$

$$\{\text{eq:krr,joint-lift}\} \quad \tilde{\mathbf{Z}}_Z^{(t)} \leftarrow \left[\tilde{\mathbf{z}}_Z(\mathbf{y}_1^{(t)}, \boldsymbol{\nu}_1), \dots, \tilde{\mathbf{z}}_Z(\mathbf{y}_V^{(t)}, \boldsymbol{\nu}_V) \right] \quad (5.28)$$

$$\{\text{eq:krr,joint-est}\} \quad \mathbf{X}^{(t)} \leftarrow \mathbf{m}_x \mathbf{1}_V^{\text{T}} + \mathbf{C}_{x\tilde{z}} (\mathbf{C}_{\tilde{z}\tilde{z}} + \rho \mathbf{I}_N)^{-1} \left(\tilde{\mathbf{Z}}_Z^{(t)} - \mathbf{m}_{\tilde{z}} \mathbf{1}_V^{\text{T}} \right) \quad (5.29)$$

$$\{\text{eq:krr,joint-model}\} \quad \mathbf{S}^{(t)} \leftarrow \left[\mathbf{s}(\mathbf{x}_1^{(t)}, \boldsymbol{\nu}_1), \dots, \mathbf{s}(\mathbf{x}_V^{(t)}, \boldsymbol{\nu}_V) \right] \quad (5.30)$$

where \mathbf{A} is a linear sensing operator that models how raw data \mathbf{D} relates to image data \mathbf{Y} ; and $(\cdot)^{(t)}$ denotes the t th iterate. The general concept involves cycling image reconstruction (5.27), feature mapping (5.28), parameter estimation via KRR (5.29), and signal modeling (5.30). Setting $\eta^{(0)} \leftarrow 0$ provides a natural initialization, and increasing $\eta^{(t)}$ for subsequent iterations encourages image iterates $\mathbf{Y}^{(t)}$ to not only maintain raw data fidelity but also add training point consistency, which might provide a regularization-like effect.

Two points of inquiry may guide refinement of the above algorithm (5.27)-(5.30). First, it is desirable to reformulate (5.27)-(5.30) as steps towards the solution of a well-defined optimization problem. This will likely provide a route towards convergence analysis and may also give intuition in setting $\eta^{(t)}$ or other algorithmic parameters. Second, it is desirable to explicitly incorporate the signal model's partially linear structure in variable updates, as this may help incorporate scale invariance in the KRR step (5.29).

CHAPTER 6

Myelin Water Fraction Estimation from Steady-State Sequences

{c,mwf}

todo: discuss negative f_F range

CHAPTER 7

Future Work

{c,future}

APPENDIX A

Coil Data Combination from Multiple Datasets

{a,cc-multi}

APPENDIX B

DESS in the Presence of Diffusion

{a, dess-diff}

BIBLIOGRAPHY

- [1] K. E. Keenan, K. F. Stupic, M. A. Boss, S. E. Russek, T. L. Chenevert, P. V. Prasad, W. E. Reddick, K. M. Cecil, J. Zheng, P. Hu, and E. F. Jackson, "Multi-site, multi-vendor comparison of T1 measurement using ISMRM/NIST system phantom," in *Proc. Intl. Soc. Mag. Res. Med.*, p. 3290, 2016.
- [2] S. Ahn and J. A. Fessler, "Standard errors of mean, variance, and standard deviation estimators," Tech. Rep. 413, Comm. and Sign. Proc. Lab., Dept. of EECS, Univ. of Michigan, Ann Arbor, MI, 48109-2122, July 2003.
- [3] G. Nataraj, J.-F. Nielsen, and J. A. Fessler, "Regularized, joint estimation of T1 and M0 maps," in *Proc. Intl. Soc. Mag. Res. Med.*, p. 3128, 2014.
- [4] G. Nataraj, J.-F. Nielsen, and J. A. Fessler, "Model-based estimation of T2 maps with dual-echo steady-state MR imaging," in *Proc. IEEE Intl. Conf. on Image Processing*, pp. 1877–81, 2014.
- [5] G. Nataraj, J.-F. Nielsen, and J. A. Fessler, "Optimizing MR scan design for model-based T1, T2 estimation from steady-state sequences," *IEEE Trans. Med. Imag.*, 2017. To appear.
- [6] G. Nataraj, J.-F. Nielsen, and J. A. Fessler, "A min-max CRLB optimization approach to scan selection for relaxometry," in *Proc. Intl. Soc. Mag. Res. Med.*, p. 1672, 2015.
- [7] G. Nataraj, J.-F. Nielsen, and J. A. Fessler, "Dictionary-free MRI parameter estimation via kernel ridge regression," in *Proc. IEEE Intl. Symp. Biomed. Imag.*, 2017. To appear.
- [8] G. Nataraj, J.-F. Nielsen, and J. A. Fessler, "Myelin water fraction estimation from optimized steady-state sequences using kernel ridge regression," in *Proc. Intl. Soc. Mag. Res. Med.*, 2017. Submitted.
- [9] A. Macovski, *Medical imaging systems*. New Jersey: Prentice-Hall, 1983.
- [10] E. M. Haacke, R. W. Brown, M. R. Thompson, and R. Venkatesan, *Magnetic resonance imaging: Physical principles and sequence design*. New York: Wiley, 1999.
- [11] D. G. Nishimura, "Principles of magnetic resonance imaging," 1996. Unpublished textbook.

- [12] F. Bloch, "Nuclear induction," *Phys. Rev.*, vol. 70, pp. 460–74, Oct. 1946.
- [13] W. S. Hinshaw, "Image formation by nuclear magnetic resonance: The sensitive point method," *J. Appl. Phys.*, vol. 47, p. 3709, Aug. 1976.
- [14] K. Scheffler, "A pictorial description of steady-states in rapid magnetic resonance imaging," *Concepts in Magnetic Resonance*, vol. 11, no. 5, pp. 291–304, 1999.
- [15] B. A. Hargreaves, S. S. Vasanawala, J. M. Pauly, and D. G. Nishimura, "Characterization and reduction of the transient response in steady-state MR imaging," *Mag. Res. Med.*, vol. 46, pp. 149–58, July 2001.
- [16] Y. Zur, M. L. Wood, and L. J. Neuringer, "Spoiling of transverse magnetization in steady-state sequences," *Mag. Res. Med.*, vol. 21, pp. 251–63, Oct. 1991.
- [17] V. Denolin, C. Azizieh, and T. Metens, "New insights into the mechanisms of signal formation in rf-spoiled gradient echo sequences," *Mag. Res. Med.*, vol. 54, pp. 937–54, Oct. 2005.
- [18] T. Redpath and R. A. Jones, "Fade-a new fast imaging sequence," *Mag. Res. Med.*, vol. 6, pp. 224–34, Feb. 1988.
- [19] H. Bruder, H. Fischer, R. Graumann, and M. Deimling, "A new steady-state imaging sequence for simultaneous acquisition of two MR images with clearly different contrasts," *Mag. Res. Med.*, vol. 7, pp. 35–42, May 1988.
- [20] J. B. Rosen, "The gradient projection method for nonlinear programming, Part I: Linear constraints," *SIAM J. Appl. Math.*, vol. 8, no. 1, pp. 181–217, 1960.
- [21] C. Byrne, "A unified treatment of some iterative algorithms in signal processing and image reconstruction," *Inverse Prob.*, vol. 20, pp. 103–20, Feb. 2004.
- [22] G. Golub and V. Pereyra, "Separable nonlinear least squares: the variable projection method and its applications," *Inverse Prob.*, vol. 19, pp. R1–26, Apr. 2003.
- [23] H.-L. M. Cheng, N. Stikov, N. R. Ghugre, and G. A. Wright, "Practical medical applications of quantitative MR relaxometry," *J. Mag. Res. Im.*, vol. 36, pp. 805–24, Oct. 2012.
- [24] J. A. Fessler and B. P. Sutton, "Nonuniform fast Fourier transforms using min-max interpolation," *IEEE Trans. Sig. Proc.*, vol. 51, pp. 560–74, Feb. 2003.
- [25] M. J. Muckley, D. C. Noll, and J. A. Fessler, "Fast parallel MR image reconstruction via B1-based, adaptive restart, iterative soft thresholding algorithms (BARISTA)," *IEEE Trans. Med. Imag.*, vol. 34, pp. 578–88, Feb. 2015.
- [26] A. Macovski, "Noise in MRI," *Mag. Res. Med.*, vol. 36, pp. 494–7, Sept. 1996.
- [27] T. Lei, "Statistics of MR signals: revisited," in *Proc. SPIE 6510 Medical Imaging 2007: Phys. Med. Im.*, p. 651052, 2007.

- [28] H. Gudbjartsson and S. Patz, "The Rician distribution of noisy MRI data," *Mag. Res. Med.*, vol. 34, pp. 910–4, Dec. 1995.
- [29] H. B. W. Larsson, J. Frederiksen, L. Kjaer, O. Henriksen, and J. Olesen, "In vivo determination of T1 and T2 in the brain of patients with severe but stable multiple sclerosis," *Mag. Res. Med.*, vol. 7, pp. 43–55, May 1988.
- [30] T. Kurki, N. Lundbom, M. Komu, and M. Kormano, "Tissue characterization of inter cranial tumors by magnetization transfer and spin-lattice relaxation parameters in vivo," *J. Mag. Res. Im.*, vol. 6, pp. 573–9, Aug. 1996.
- [31] E. Englund, A. Brun, Z. Gyorffy-Wagner, E. Larsson, and B. Persson, "Relaxation times in relation to grade of malignancy and tissue necrosis in astrocytic gliomas," *Mag. Res. Im.*, vol. 4, no. 5, pp. 425–9, 1986.
- [32] S. Siemonsen, K. Mouridsen, B. Holst, T. Ries, J. Finsterbusch, G. Thomalia, L. Ostergaard, and J. Fiehler, "Quantitative T2 values predict time from symptom onset in acute stroke patients," *Stroke*, vol. 40, pp. 1612–6, May 2009.
- [33] L. D. DeWitt, J. P. Kistler, D. C. Miller, E. P. Richardson, and F. S. Buonanno, "NMR-neuropathologic correlation in stroke," *Stroke*, vol. 18, no. 2, pp. 342–51, 1987.
- [34] S. J. Matzat, J. V. Tiel, G. E. Gold, and E. H. G. Oei, "Quantitative MRI techniques of cartilage composition," *Quant. Imaging Med. Surg.*, vol. 3, pp. 162–74, June 2013.
- [35] T. J. Mosher and B. J. Dardzinski, "Cartilage MRI T2 relaxation time mapping: overview and applications," *Semin. Musculoskelet. Radiol.*, vol. 8, no. 4, pp. 355–68, 2004.
- [36] H. Guo, W.-Y. Au, J. S. Cheung, D. Kim, J. H. Jensen, P.-L. Khong, Q. Chan, K. C. Chan, C. Tosti, H. Tang, T. R. Brown, W. W. M. Lam, S.-Y. Ha, G. M. Brittenham, and E. X. Wu, "Myocardial T2 quantification in patients with iron overload at 3 Tesla," *J. Mag. Res. Im.*, vol. 30, pp. 394–400, Aug. 2009.
- [37] S. Giri, Y. C. Chung, A. Merchant, G. Mihai, S. Rajagopalan, S. V. Raman, and O. P. Simonetti, "T2 quantification for improved detection of myocardial edema," *Cardiovasc. Magn. Reson.*, vol. 11, no. 1, pp. 56–68, 2009.
- [38] D. C. Look and D. R. Locker, "Time saving in measurement of NMR and EPR relaxation times," *Rev Sci Instrum.*, vol. 41, pp. 250–1, Feb. 1970.
- [39] H. Y. Carr and E. M. Purcell, "Effects of diffusion on free precession in nuclear magnetic resonance experiments," *Phys. Rev.*, vol. 94, pp. 630–8, May 1954.
- [40] M. K. Stehling, R. Turner, and P. Mansfield, "Echo-planar imaging: magnetic resonance imaging in a fraction of a second," *Science*, vol. 254, pp. 43–50, Oct. 1991.

- [41] C. B. Ahn, J. H. Kim, and Z. H. Cho, "High-speed spiral-scan echo planar NMR imaging - I," *IEEE Trans. Med. Imag.*, vol. 5, pp. 2–7, Mar. 1986.
- [42] C. H. Meyer, B. S. Hu, D. G. Nishimura, and A. Macovski, "Fast spiral coronary artery imaging," *Mag. Res. Med.*, vol. 28, pp. 202–13, Dec. 1992.
- [43] I. Kay and R. M. Henkelman, "Practical Implementation and Optimization of One-shot T1 imaging," *Mag. Res. Med.*, vol. 22, pp. 414–24, Dec. 1991.
- [44] P. A. Gowland and M. O. Leach, "Fast and accurate measurements of T1 using a multi-readout single inversion-recovery sequence," *Mag. Res. Med.*, vol. 26, pp. 79–88, July 1992.
- [45] D. R. Messroghli, A. Radjenovic, S. Kozerke, D. M. Higgins, M. U. Sivananthan, and J. P. Ridgway, "Modified Look-Locker inversion recovery (MOLLI) for high-resolution T_1 mapping of the heart," *Mag. Res. Med.*, vol. 52, pp. 141–6, July 2004.
- [46] M. K. Stehling, R. J. Ordidge, R. Coxon, and P. Mansfield, "Inversion-recovery Echo-planar imaging (IR-EPI) at 0.5T," *Mag. Res. Med.*, vol. 13, pp. 514–7, Mar. 1990.
- [47] J.-M. Bonny, M. Zanca, J.-Y. Boire, and A. Veyre, "T2 maximum likelihood estimation from multiple spin-echo magnitude images," *Mag. Res. Med.*, vol. 36, pp. 287–93, Aug. 1996.
- [48] D. Kumar, T. D. Nguyen, S. A. Gauthier, and A. Raj, "Bayesian algorithm using spatial priors for multiexponential T2 relaxometry from multiecho spin echo MRI," *Mag. Res. Med.*, vol. 68, pp. 1536–43, Nov. 2012.
- [49] N. Ben-Eliezer, D. K. Sodickson, and K. T. Block, "Rapid and accurate T2 mapping from multi-spin-echo data using Bloch-simulation-based reconstruction," *Mag. Res. Med.*, vol. 73, pp. 809–17, Feb. 2015.
- [50] T. D. Nguyen, C. Wisnieff, M. A. Cooper, D. Kumar, A. Raj, P. Spincemaille, Y. Wang, T. Vartanian, and S. A. Gauthier, "T2prep three-dimensional spiral imaging with efficient whole brain coverage for myelin water quantification at 1.5 tesla," *Mag. Res. Med.*, vol. 67, pp. 614–21, Mar. 2012.
- [51] S. Majumdar, S. C. Orphanoudakis, A. Gmitro, M. O'Donnell, and J. C. Gore, "Error in the measurements of T2 using multiple-echo MRI techniques: 1. Effect of radiofrequency pulse imperfections," *Mag. Res. Med.*, vol. 3, pp. 397–417, June 1986.
- [52] S. Majumdar, S. C. Orphanoudakis, A. Gmitro, M. O'Donnell, and J. C. Gore, "Error in the measurements of T2 using multiple-echo MRI techniques: 2. Effects of static field inhomogeneity," *Mag. Res. Med.*, vol. 3, pp. 562–74, Aug. 1986.

- [53] F. Farzaneh, S. J. Riederer, and N. J. Pelc, "Analysis of T2 limitations and off-resonance effects on spatial resolution and artifacts in echo-planar imaging," *Mag. Res. Med.*, vol. 14, pp. 123–39, Apr. 1990.
- [54] S. C. L. Deoni, B. K. Rutt, and T. M. Peters, "Rapid combined T1 and T2 mapping using gradient recalled acquisition in the steady state," *Mag. Res. Med.*, vol. 49, pp. 515–26, Mar. 2003.
- [55] L.-C. Chang, C. G. Koay, P. J. Basser, and C. Pierpaoli, "Linear least-squares method for unbiased estimation of T1 from SPGR signals," *Mag. Res. Med.*, vol. 60, pp. 496–501, Aug. 2008.
- [56] E. K. Fram, R. J. Herfkens, G. A. Johnson, G. H. Glover, J. P. Kaaris, A. Shimakawa, T. G. Perkins, and N. J. Pelc, "Rapid calculation of T1 using variable flip angle gradient refocused imaging," *Mag. Res. Im.*, vol. 5, no. 3, pp. 201–8, 1987.
- [57] H. Wang and Y. Cao, "Spatially regularized T1 estimation from variable flip angles MRI," *Med. Phys.*, vol. 39, pp. 4139–48, July 2012.
- [58] S. C. L. Deoni, H. A. Ward, T. M. Peters, and B. K. Rutt, "Rapid T_2 estimation with phase-cycled variable nutation steady-state free precession," *Mag. Res. Med.*, vol. 52, pp. 435–9, Aug. 2004.
- [59] S. C. L. Deoni, "Transverse relaxation time (T2) mapping in the brain with off-resonance correction using phase-cycled steady-state free precession imaging," *J. Mag. Res. Im.*, vol. 30, pp. 411–7, Aug. 2009.
- [60] G. H. Welsch, K. Scheffler, T. C. Mamisch, T. Hughes, S. Millington, M. Deimling, and S. Trattnig, "Rapid estimation of cartilage T2 based on double echo at steady state (DESS) with 3 Tesla," *Mag. Res. Med.*, vol. 62, pp. 544–9, Aug. 2009.
- [61] R. Heule, C. Ganter, and O. Bieri, "Rapid estimation of cartilage T2 with reduced T1 sensitivity using double echo steady state imaging," *Mag. Res. Med.*, vol. 71, pp. 1137–43, Mar. 2014.
- [62] T. Stöcker, F. Keil, K. Vahedipour, D. Brenner, E. Pracht, and N. J. Shah, "MR parameter quantification with magnetization-prepared double echo steady-state (MP-DESS)," *Mag. Res. Med.*, vol. 72, pp. 103–11, July 2014.
- [63] R. Heule, C. Ganter, and O. Bieri, "Triple echo steady-state (TESS) relaxometry," *Mag. Res. Med.*, vol. 71, pp. 230–7, Jan. 2014.
- [64] M. L. Gyngell, "The steady-state signals in short-repetition-time sequences," *J. Mag. Res.*, vol. 81, pp. 474–83, Feb. 1989.
- [65] W. Hänicke and H. U. Vogel, "An analytical solution for the SSFP signal in MRI," *Mag. Res. Med.*, vol. 49, pp. 771–5, Apr. 2003.

- [66] G. H. Weiss, R. K. Gupta, J. A. Ferretti, and E. D. Becker, "The choice of optimal parameters for measurement of spin-lattice relaxation times. I. Mathematical formulation," *J. Mag. Res.*, vol. 37, pp. 369–79, Feb. 1980.
- [67] Y. Zhang, H. N. Yeung, M. O'Donnell, and P. L. Carson, "Determination of sample time for T1 measurement," *J. Mag. Res. Im.*, vol. 8, pp. 675–81, May 1998.
- [68] H. Z. Wang, S. J. Riederer, and J. N. Lee, "Optimizing the precision in T1 relaxation estimation using limited flip angles," *Mag. Res. Med.*, vol. 5, pp. 399–416, Nov. 1987.
- [69] J. A. Jones, P. Hodgkinson, A. L. Barker, and P. J. Hore, "Optimal sampling strategies for the measurement of spin-spin relaxation times," *J. Mag. Res. B*, vol. 113, pp. 25–34, Oct. 1996.
- [70] J. Imran, François. Langevin, and Hervé. Saint-Jalmes, "Two-point method for T1 estimation with optimized gradient-echo sequence," *Mag. Res. Im.*, vol. 17, pp. 1347–56, Nov. 1999.
- [71] S. C. L. Deoni, T. M. Peters, and B. K. Rutt, "Determination of optimal angles for variable nutation proton magnetic spin-lattice, T_1 , and spin-spin, T_2 , relaxation times measurement," *Mag. Res. Med.*, vol. 51, pp. 194–9, Jan. 2004.
- [72] L. Fleyscher, R. Fleyscher, S. Liu, W. Zaaraoui, and O. Gonen, "Optimizing the precision-per-unit-time of quantitative MR metrics: Examples for T_1 , T_2 , and DTI," *Mag. Res. Med.*, vol. 57, pp. 380–7, Feb. 2007.
- [73] M. Akçakaya, S. Weingärtner, Sébastien. Roujol, and R. Nezafat, "On the selection of sampling points for myocardial T1 mapping," *Mag. Res. Med.*, vol. 73, pp. 1741–53, May 2015.
- [74] C. M. Lewis, S. A. Hurley, M. E. Meyerand, and C. G. Koay, "Data-driven optimized flip angle selection for T1 estimation from spoiled gradient echo acquisitions," *Mag. Res. Med.*, vol. 76, pp. 792–802, Sept. 2016. To appear.
- [75] R. A. Fisher, "Theory of statistical estimation," *Proc. Cambridge Philosophical Society*, vol. 22, pp. 700–25, July 1925.
- [76] H. Cramér, *Mathematical methods of statistics*. Princeton: Princeton Univ. Press, 1946.
- [77] H. Chernoff, "Locally optimal designs for estimating parameters," *Ann. Math. Stat.*, vol. 24, pp. 586–602, Dec. 1953.
- [78] J. P. Wansapura, S. K. Holland, R. S. Dunn, and W. S. Ball, "NMR relaxation times in the human brain at 3.0 Tesla," *J. Mag. Res.*, vol. 9, pp. 531–8, Apr. 1999.
- [79] G. J. Stanisz, E. E. Odobina, J. Pun, M. Escaravage, S. J. Graham, M. J. Bronskill, and R. M. Henkelman, " T_1 , T_2 relaxation and magnetization transfer in tissue at 3T," *Mag. Res. Med.*, vol. 54, pp. 507–12, Sept. 2005.

- [80] R. K.-S. Kwan, A. C. Evans, and G. B. Pike, "MRI simulation-based evaluation of image-processing and classification methods," *IEEE Trans. Med. Imag.*, vol. 18, pp. 1085–97, Nov. 1999.
- [81] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans, "Design and construction of a realistic digital brain phantom," *IEEE Trans. Med. Imag.*, vol. 17, pp. 463–8, June 1998.
- [82] J. Pauly, P. Le Roux, D. Nishimura, and A. Macovski, "Parameter relations for the Shinnar-Le Roux selective excitation pulse design algorithm," *IEEE Trans. Med. Imag.*, vol. 10, pp. 53–65, Mar. 1991.
- [83] L. I. Sacolick, F. Wiesinger, I. Hancu, and M. W. Vogel, "B1 mapping by Bloch-Siebert shift," *Mag. Res. Med.*, vol. 63, pp. 1315–22, May 2010.
- [84] H. Sun, W. A. Grissom, and J. A. Fessler, "Regularized estimation of Bloch-Siebert B1+ Maps in MRI," in *Proc. IEEE Intl. Conf. on Image Processing*, pp. 3646–50, 2014.
- [85] L. Ying and J. Sheng, "Joint image reconstruction and sensitivity estimation in SENSE (JSENSE)," *Mag. Res. Med.*, vol. 57, pp. 1196–1202, June 2007.
- [86] D. Arthur and S. Vassilvitskii, "K-means++: The advantages of careful seeding," in *Proc. 18th Annual ACM-SIAM Symp. Disc. Alg. (SODA)*, pp. 1027–35, 2007.
- [87] L. R. Frank, E. C. Wong, and R. B. Buxton, "Slice profile effects in adiabatic inversion: Application to multislice perfusion imaging," *Mag. Res. Med.*, vol. 38, pp. 558–64, Oct. 1997.
- [88] J. K. Barral, E. Gudmundson, N. Stikov, M. Etezadi-Amoli, P. Stoica, and D. G. Nishimura, "A robust methodology for in vivo T1 mapping," *Mag. Res. Med.*, vol. 64, pp. 1057–67, Oct. 2010.
- [89] C. E. Carney, S. T. S. Wong, and S. Patz, "Analytical solution and verification of diffusion effect in SSFP," *Mag. Res. Med.*, vol. 19, pp. 240–6, June 1991.
- [90] E. X. Wu and R. B. Buxton, "Effect of diffusion on the steady-state magnetization with pulsed field gradients," *J. Mag. Res.*, vol. 90, pp. 243–53, Nov. 1990.
- [91] R. Kaiser, E. Bartholdi, and R. R. Ernst, "Diffusion and field-gradient effects in NMR Fourier spectroscopy," *J. Chem. Phys.*, vol. 60, pp. 2966–79, Apr. 1974.
- [92] R. Heule, P. Bär, C. Mirkes, K. Scheffler, S. Trattnig, and O. Bieri, "Triple-echo steady-state T2 relaxometry of the human brain at high to ultra-high fields," *NMR in Biomedicine*, vol. 27, pp. 1037–45, Sept. 2014.
- [93] A. Mackay, K. Whittall, J. Adler, D. Li, D. Paty, and D. Graeb, "In vivo visualization of myelin water in brain by magnetic resonance," *Mag. Res. Med.*, vol. 31, pp. 673–7, June 1994.

- [94] R. D. Gill and B. Y. Levit, "Applications of the van Trees inequality: A Bayesian Cramér-rao bound," *Bernoulli*, vol. 1, no. 1/2, pp. 59–79, 1995.
- [95] D. Ma, V. Gulani, N. Seiberlich, K. Liu, J. L. Sunshine, J. L. Duerk, and M. A. Griswold, "Magnetic resonance fingerprinting," *Nature*, vol. 495, pp. 187–93, Mar. 2013.
- [96] B. Zhao, J. Haldar, K. Setsompop, and L. L. Wald, "Optimal experiment design for magnetic resonance fingerprinting," in *embc*, 2016.
- [97] H. C. Torrey, "Bloch equations with diffusion terms," *Phys. Rev.*, vol. 104, pp. 563–5, 1956.
- [98] H. M. McConnell, "Reaction rates by nuclear magnetic resonance," *J. of Chemical Phys.*, vol. 28, pp. 430–31, Mar. 1958.
- [99] N. Bloembergen, E. M. Purcell, and R. V. Pound, "Relaxation effects in nuclear magnetic resonance absorption," *Phys. Rev.*, vol. 73, pp. 679–712, Apr. 1948.
- [100] D. L. Bihan, J.-F. Mangin, C. Poupon, C. A. Clark, S. Pappata, N. Molko, and H. Chabriet, "Diffusion tensor imaging: Concepts and applications," *J. Mag. Res. Im.*, vol. 13, pp. 534–546, Apr. 2001.
- [101] E. Staroswiecki, K. L. Granlund, M. T. Alley, G. E. Gold, and B. A. Hargreaves, "Simultaneous estimation of T2 and apparent diffusion coefficient in human articular cartilage in vivo with a modified three-dimensional double echo steady state (DESS) sequence at 3 T," *Mag. Res. Med.*, vol. 67, no. 4, pp. 1086–96, 2012.
- [102] B. Zhao, K. Setsompop, H. Ye, S. Cauley, and L. L. Wald, "Maximum likelihood reconstruction for magnetic resonance fingerprinting," *IEEE Trans. Med. Imag.*, vol. 35, pp. 1812–23, Aug. 2016.
- [103] D. A. Feinberg, L. E. Crooks, P. Sheldon, J. H. III, J. Watts, and M. Arakawa, "Magnetic resonance imaging the velocity vector components of fluid flow," *Mag. Res. Med.*, vol. 2, pp. 555–66, Dec. 1985.
- [104] D. S. Tuch, V. J. Wedeen, A. M. Dale, J. S. George, and J. W. Belliveau, "Conductivity tensor mapping of the human brain using diffusion tensor MRI," *Proc. Natl. Acad. Sci.*, vol. 98, pp. 11697–701, Sept. 2001.
- [105] K. Sekihara, S. Matsui, and H. Kohno, "NMR imaging for magnets with large nonuniformities," *IEEE Trans. Med. Imag.*, vol. 4, pp. 193–9, Dec. 1985.
- [106] G. R. Morrell, "A phase-sensitive method of flip angle mapping," *Mag. Res. Med.*, vol. 60, pp. 889–94, Oct. 2008.
- [107] S. C. L. Deoni, B. K. Rutt, T. Arun, C. Pierpaoli, and D. K. Jones, "Gleaning multi-component T1 and T2 information from steady-state imaging data," *Mag. Res. Med.*, vol. 60, pp. 1372–87, Dec. 2008.

- [108] S. C. L. Deoni, L. Matthews, and S. H. Kolind, “One component? Two components? Three? The effect of including a nonexchanging ”free” water component in multi-component driven equilibrium single pulse observation of T1 and T2,” *Mag. Res. Med.*, vol. 70, pp. 147–54, July 2013.
- [109] G. S. Kimeldorf and G. A. Wahba, “A correspondence between bayesian estimation on stochastic processes and smoothing by splines,” *Ann. Math. Statist.*, vol. 41, no. 2, pp. 495–502, 1970.
- [110] N. Aronszajn, “Theory of reproducing kernels,” *tams*, vol. 68, pp. 337–404, May 1950.
- [111] C. Cortes and V. Vapnik, “Support-vector networks,” *Mach. Learn.*, vol. 20, pp. 273–97, Sept. 1995.
- [112] C. Saunders, A. Gammerman, and V. Vovk, “Ridge regression learning algorithm in dual variables,” in *In Proceedings of the 15th International Conference on Machine Learning*, pp. 515–21, Morgan Kaufmann, 1998.
- [113] B. Schölkopf, R. Herbrich, and A. J. Smola, “A generalized representer theorem,” in *Proc. Computational Learning Theory (COLT)*, pp. 416–426, 2001. LNCS 2111.
- [114] A. E. Hoerl and R. W. Kennard, “Ridge regression: biased estimation for nonorthogonal problems,” *Technometrics*, vol. 12, pp. 55–67, Feb. 1970.
- [115] A. Rahimi and B. Recht, “Random features for large-scale kernel machines,” in *NIPS*, 2007.
- [116] D. J. Sutherland and J. Schneider, “On the error of random fourier features,” in *In Proceedings of the 31st International Conference on Uncertainty in Artificial Intelligence*, 2015.
- [117] Z. Wu, “Generalized Bochner’s theorem for radial function,” *Approximation Theory and its Applications*, vol. 13, no. 3, pp. 47–57, 1997.
- [118] M. A. Woodbury, “Inverting modified matrices,” 1950. Tech. Report 42, Stat. Res. Group, Princeton Univ.
- [119] E. Parzen, “On estimation of a probability density function and mode,” *Ann. Math. Stat.*, vol. 33, pp. 1065–76, Sept. 1962.
- [120] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005.
- [121] E. Alonso-Ortiz, I. R. Levesque, and G. B. Pike, “MRI-based myelin water imaging: A technical review,” *Mag. Res. Med.*, vol. 73, pp. 70–81, Jan. 2015.
- [122] S. C. L. Deoni, “Correction of main and transmit magnetic field (B0 and B1) inhomogeneity effects in multicomponent-driven equilibrium single-pulse observation of T1 and T2,” *Mag. Res. Med.*, vol. 65, pp. 1021–35, Apr. 2011.

- [123] C. L. Lankford and M. D. Does, “On the inherent precision of mcDESPOT,” *Mag. Res. Med.*, vol. 69, pp. 127–36, Jan. 2013.
- [124] M. M. Siddiqui, “Statistical inference for Rayleigh distributions,” *RADIO SCIENCE Journal of Research NBS/USNC-URSI*, vol. 68D, pp. 1005–10, Sept. 1964.