

# **Advances in Quantitative MRI: Acquisition, Estimation, and Application**

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  
2017

## **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

---

2017

# TABLE OF CONTENTS

<b>List of Figures</b> . . . . .	<b>iii</b>
<b>List of Tables</b> . . . . .	<b>v</b>
<b>List of Appendices</b> . . . . .	<b>vii</b>
<b>Abstract</b> . . . . .	<b>viii</b>
<b>Chapter</b>	
<b>1 Introduction</b> . . . . .	<b>1</b>
1.1 Thesis Overview . . . . .	2
1.2 Thesis Proposal Organization . . . . .	2
<b>2 Background</b> . . . . .	<b>4</b>
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
<b>3 MRI Parameter Estimation from Likelihood Models</b> . . . . .	<b>18</b>
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
3.3.1 $T_1$ estimation from two SPGR scans . . . . .	23
3.3.2 $T_2$ estimation from one DESS scan . . . . .	25
3.4 Summary . . . . .	25
<b>4 Optimizing MR Scan Design for Model-Based Relaxometry</b> . . . . .	<b>26</b>
4.1 Introduction . . . . .	26
4.2 A CRB-Inspired Scan Selection Method . . . . .	28
4.2.1 The CRB and its Relevance to QMRI . . . . .	28
4.2.2 Min-max Optimization Problem for Scan Design . . . . .	29
4.3 Optimizing SS Sequences for Relaxometry in the Brain . . . . .	30

4.3.1	Scan Design Details . . . . .	31
4.3.2	Scan Profile Comparisons . . . . .	32
4.4	Experimental Validation and Results . . . . .	35
4.4.1	Numerical Simulations . . . . .	35
4.4.2	Phantom Experiments . . . . .	38
4.4.3	<i>In Vivo</i> Experiments . . . . .	44
4.5	Discussion and Future Work . . . . .	48
4.6	Conclusion . . . . .	49
<b>5</b>	<b>Dictionary-Free MRI Parameter Estimation via Regression with Kernels (PERK)</b>	<b>51</b>
5.1	Introduction . . . . .	51
5.2	A Function Optimization Problem & Kernel Solution . . . . .	53
5.3	Implementation Considerations . . . . .	55
5.3.1	A Kernel Approximation . . . . .	55
5.3.2	Model Selection . . . . .	57
5.4	Experimentation . . . . .	60
5.5	Summary and Future Work . . . . .	63
<b>6</b>	<b>Fast Myelin Water Fraction Imaging via Scan Design and PERK</b>	<b>66</b>
6.1	Introduction . . . . .	66
6.2	Multi-Compartmental Models for SS Sequences . . . . .	68
6.2.1	A Two-Compartment SPGR Signal Model . . . . .	68
6.2.2	A Two-Compartment DESS Signal Model . . . . .	74
6.3	A Fast Acquisition for Precise MWF Estimation . . . . .	80
6.3.1	Scalable Acquisition Design . . . . .	80
6.3.2	SPGR/DESS Scan Design Implementation Details . . . . .	82
6.4	Experimentation . . . . .	85
6.5	Summary and Future Work . . . . .	88
<b>7</b>	<b>Future Work</b>	<b>90</b>
	<b>Appendices</b>	<b>92</b>
	<b>Bibliography</b>	<b>94</b>

## LIST OF FIGURES

3.1	<p><math>T_1</math> MOM, ML, and RL estimates and corresponding error images, from two simulated SPGR scans. Magnitude error images are <math>10\times</math> magnified. Voxels not assigned WM- or GM-like relaxation times are masked out in post-processing for display. Table 3.1 presents corresponding sample statistics. . . . .</p>	24
4.1	<p>Worst-case standard deviations <math>\tilde{\sigma}_{T_1}^t</math> (top), <math>\tilde{\sigma}_{T_2}^t</math> (middle), and cost <math>\tilde{\Psi}^t</math> (bottom), versus pairs of nominal flip angles, holding other scan parameters fixed at selected profile <math>\hat{\mathbf{P}}</math>. Subfigures (a)-(i), (j)-(l), and (m)-(o) correspond to scan profiles containing <math>(S_{\text{SPGR}}, S_{\text{DESS}}) = (2, 1), (1, 1), \text{ and } (0, 2)</math> SPGR and DESS scans, respectively. Selected scan parameters (starred) are within <math>\delta = 1\%</math> 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 <math>[0, 100]</math>, <math>[0, 10]</math>, and <math>[0, 20]</math> milliseconds for rows of <math>\tilde{\sigma}_{T_1}^t</math>, <math>\tilde{\sigma}_{T_2}^t</math>, and <math>\tilde{\Psi}^t</math> subfigures, respectively. The optimized <math>(0, 2)</math> profile appears most robust to transmit field spatial variation. . . . .</p>	34
4.2	<p>Histograms of <math>T_1</math> and <math>T_2</math> 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 <math>T_1</math> and <math>T_2</math>. In (a)-(b) and (c)-(d), the solid green curve is <math>\mathcal{N}(T_1, (\tilde{\sigma}_{T_1}^t)^2)</math> and <math>\mathcal{N}(T_2, (\tilde{\sigma}_{T_2}^t)^2)</math>, respectively. In (a)-(d), the dashed maroon curves have variances computed from the Fisher information at <i>a priori</i> unknown <math>T_1, T_2</math> values in WM or GM. These plots correspond to an optimized <math>(0, 2)</math> 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 <math>\hat{T}_1^{\text{ML}}</math> and <math>\hat{T}_2^{\text{ML}}</math> errors. . . . .</p>	37
4.3	<p>Colorized <math>T_1</math> and <math>T_2</math> ML and RL estimates from an HPD<sup>®</sup> 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 <math>T_1</math> and <math>T_2</math> ML and RL estimators. Fig. 4.4 provides identical grayscale images that enumerate vials. Colorbar ranges are in milliseconds. . . . .</p>	41
4.4	<p>Grayscale <math>T_1</math> and <math>T_2</math> ML and RL estimates from an HPD<sup>®</sup> 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 <math>T_1</math> and <math>T_2</math> 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. . . . .</p>	42

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. . . . .	43
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. . . . .	45
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. . . . .	63
6.1	Block diagram of greedy scan profile construction. We set an original acquisition consisting of three SPGR and three DESS scans whose flip angles and repetition times are respectively initialized randomly and minimally. We optimize the original acquisition subject to several constraints: $\alpha \in [1, 40]^\circ$ and $T_R \geq 11.8\text{ms}$ for SPGR; $\alpha \in [1, 60]^\circ$ and $T_R \geq 17.5\text{ms}$ for DESS; and $\sum_{d=1}^D T_{R,d} \leq 263.7\text{ms}$ (selected to allow 9 SPGR and 9 DESS scans each at minimal $T_R$ , comparable to [2]). We then seek to improve the acquisition by appending another SPGR or DESS scan and repeating (constrained) optimization. This process iterates until additional scans no longer improve estimation precision. . . . .	84
6.2	MWF estimates, using: our optimized SPGR/DESS scan profile and ( <i>top left</i> ) KRR $f_F$ estimation, followed by ( <i>top right</i> ) iterative ML estimation; ( <i>bottom left</i> ) GRASE [3], a state-of-the-art acceleration to the slow gold-standard MESE acquisition originally used in [4]; and ( <i>bottom right</i> ) mcDESPOT [2], a fast steady-state acquisition that lacks sufficient estimation precision [5]. All MWF estimates are from healthy volunteers, but those from GRASE/mcDESPOT are from a separate study in a different subject and are reprinted from [6], a recent comparison study. The proposed precision-optimized SPGR/DESS scan profile is as fast as mcDESPOT and yields MWF images comparable to those of GRASE. . . . .	87

## LIST OF TABLES

3.1	Sample means $\pm$ sample standard deviations of MOM, ML, and RL $T_1$ , $c_1$ estimates from two simulated SPGR datasets, computed over 3001 WM-like and 1151 GM-like voxels. Each sample statistic is rounded off to the highest place value of its (unreported) standard error, computed via formulas in [7]. $T_1$ values are in milliseconds. $c_1$ values are unitless. Fig. 3.1 presents corresponding images. . . . .	25
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. . . . .	33
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. . . . .	36
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. . . . .	44
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. . . . .	46

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. . . . .	47
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 [7]. . . . .	63
6.1	SPGR/DESS flip angles and repetition times that comprise $\hat{\mathbf{P}}$ , a scan parameter matrix designed under total repetition time budget $\sum_{d=1}^D T_{R,d} \leq 263.7\text{ms}$ for precise $f_F$ estimation in WM via the sequence of optimization problems described in Figure 6.1. For our noise variance measurements, this acquisition is expected to yield 28.5% relative standard deviation in asymptotically unbiased $f_F$ estimates from two-compartment SPGR (6.22) and DESS (6.41)-(6.42) signal models. Remarkably, this optimized scan profile involving (substantial) $T_R$ variation achieves better $f_F$ estimation precision than a similarly optimized scan profile with fixed minimum- $T_R$ acquisitions, even though the latter can utilize many more scans for the given time budget. . . . .	84
6.2	<i>Left:</i> WM/GM ROIs, overlaid on a representative anatomical (coil-combined SPGR) image. Four WM ROIs and four GM ROIs are each pooled into a single WM and a single GM ROI, over which sample statistics are computed. <i>Right:</i> Within-ROI sample means $\pm$ within-ROI sample standard deviations of $f_F$ estimates, using KRR only as well as KRR with ML refinement (Fig. 6.2 presents corresponding images). . . . .	86



## **LIST OF APPENDICES**

<b>A Coil Data Combination from Multiple Datasets . . . . .</b>	<b>92</b>
<b>B DESS in the Presence of Diffusion . . . . .</b>	<b>93</b>

# **ABSTRACT**

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

**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 absence 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 directly localize the MR signal, a typically 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 algorithms 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, ??) 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 Proposal Organization

{s,intro,org}

The main body of this thesis proposal 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 to 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 [8, 9].
- 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 [10], which substantially extends conference paper [11].

- Chapter ?? describes scalable MRI parameter estimation using kernel ridge regression. It derives some content from conference paper [12].
- Chapter 6 studies multi-compartmental models 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 preliminary *in vivo* experiments. It derives some content from conference paper [13].
- 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.

In the thesis, the appendices will be organized as follows:

- Appendix A will propose 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 will present an analysis of model mismatch due to the presence of diffusion, will show that neglecting diffusive effects during  $T_2$  estimation can cause significant bias, and will suggest 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 [14, 15, 16].

### 2.1.1 Bloch Equations

{ss,bkgrd,mri,bloch}

The Bloch equations [17] describe the macroscopic magnetization dynamics of *spin*, or (loosely) atomic nuclei with nonzero angular momentum and thus nonzero magnetic moment, *e.g.*  $^1\text{H}$ . 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;  $\gamma$  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).

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  $b'_{xy}(\mathbf{r}, t) \approx \kappa^t(\mathbf{r})b'_{1,x}(t)\hat{i} + 0\hat{j}$ , where  $\kappa^t(\mathbf{r}) \in \mathbb{R}$  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

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

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  $\mathbf{m}'(\mathbf{r}, t_0) := [m'_x(\mathbf{r}, t_0), m'_y(\mathbf{r}, t_0), m'_z(\mathbf{r}, t_0)]^T$  is the initial magnetization and

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

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  $\alpha(\mathbf{r}, t_0 + T_P; t_0) =: \alpha_0 \kappa^t(\mathbf{r})$ , where  $\alpha_0$  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 [18] 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 [19]) long-time magnetization responses that eventually<sup>1</sup> 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 [21] 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

---

<sup>1</sup>The progression to steady state takes on the order of  $5T_2/T_R$  RF cycles [19], 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 [20].

{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”  $T_E \in [\frac{T_P}{2}, T_R]$  after the (midpoint of) RF excitation:

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

The remaining transverse magnetization is spoiled<sup>2</sup> following signal reception<sup>3</sup> while the longitudinal component is unaffected. We model an ideal spoiling operation as

$$\{\text{eq:spgr-spoil}\} \quad \mathbf{S}\mathbf{m}'\left(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E\right), \text{ where } \mathbf{S} := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (2.20)$$

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

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

---

<sup>2</sup>Transverse signal spoiling is often (nearly) achieved in practice using *gradient fields*, or strong induced field inhomogeneities that cause rapid transverse signal dephasing, in tandem with RF excitations that additionally impart nonlinear (often quadratically increasing) transverse magnetization phase [21]. 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 [22].

<sup>3</sup>MRI signal reception also uses gradient fields to spatially localize signal; thus signal reception would in general impart additional magnetization phase. However, we assume hereafter that MR signal is received using “balanced” imaging gradients that impart zero net phase per repetition cycle; under this restriction, balanced imaging gradients have a negligible net affect on the magnetization vector.

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 reasonable<sup>4</sup> 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  $p_{\omega'} := \text{Cauchy}(\bar{\omega}', R'_2)$ , 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

---

<sup>4</sup>Model 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):

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

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

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 [23, 24] 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 dephasing<sup>5</sup> 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

---

<sup>5</sup>It 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

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

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

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 between<sup>6</sup> 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 reasonable<sup>7</sup> 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  $n_{\text{cyc}}$  of across-

---

<sup>6</sup>Observe 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).

<sup>7</sup>Model 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 cycles<sup>8</sup> such that full-repetition phase accrual  $\phi'(\mathbf{r})$  is distributed essentially uniformly as  $\mathbf{p}_{\phi'} \leftarrow \text{unif}(0, 2\pi n_{\text{cyc}}), n_{\text{cyc}} \in \{1, 2, 3, \dots\}$ .

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):

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

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

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.

---

<sup>8</sup>In 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  $\Psi$  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  $\Psi$  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  $\Psi$  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  $\Psi$  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  $\Psi$  over convex constraint set  $\mathbb{X}$ . Such  $\hat{\mathbf{x}} \in \mathbb{X}$  must satisfy for some local neighborhood's radius  $\delta > 0$

{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  $\delta$ ; thus, any global minimizer is a local minimizer (but the converse is not necessarily true unless  $\Psi$  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 [25] is one such iterative algorithm and repeats the following simple update:

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

where  $\mathbf{P}_{\mathbb{X}}$  denotes projection onto  $\mathbb{X}$ ;  $\nabla_{\mathbf{x}} := \left[ \frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_L} \right]$  denotes row gradient with respect to length- $L$  vector  $\mathbf{x}$ ; 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  $\Psi$  is convex and sufficiently smooth, iterates produced via (2.46) converge to a limit point [26] that is a constrained global minimum (for appropriately selected  $\mathbf{\Pi}$ ). If instead  $\Psi$  is non-convex (but  $\mathbb{X}$  is still convex), statements regarding convergence<sup>9</sup> 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:

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

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  $(\cdot)^H$

<sup>9</sup>For 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  $\Psi$  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 [27] 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  $\mathbf{x}_N \in \mathbb{X}_N$ :

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

$$\{\text{eq:sep-ls-fullrnk}\} \quad = (\mathbf{A}^H(\mathbf{x}_N) \mathbf{W} \mathbf{A}(\mathbf{x}_N))^{-1} \mathbf{A}^H(\mathbf{x}_N) \mathbf{W} \mathbf{y}, \quad (2.50)$$

where  $(\cdot)^\dagger$  denotes pseudoinverse;  $\mathbf{W}^{1/2}$  denotes principal (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:

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

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 only be tractable via iterative optimization (see Subsection 2.2.1) towards a local solution  $(\hat{\mathbf{x}}_L(\hat{\mathbf{x}}_N), \hat{\mathbf{x}}_N)$ .

## CHAPTER 3

# MRI Parameter Estimation from Likelihood Models

{c,relax}

### 3.1 Introduction

{s,relax,intro}

This brief chapter describes methods for QMRI parameter estimation from statistical likelihood models. The main purpose of this chapter is to serve as a bridge between the background information reviewed in Chapter 2 and more novel ideas introduced in later chapters. As such, we place emphasis here on development of notation and terminology over thorough validation. As instructional examples, we demonstrate likelihood-based parameter estimation on simple problems involving estimation of relaxation parameters  $T_1$  and  $T_2$ , applications that Chapter 4 motivates and studies in much greater detail.

The remainder of this chapter is organized as follows. 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 discusses practical implementation issues. Section 3.3 demonstrates the utility of likelihood-based parameter estimation over conventional estimators through comparisons in simulation. Section 3.4 provides concluding remarks and motivates the future work items studied in subsequent chapters.

### 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 simplicity<sup>1</sup> to be (circularly-symmetric) complex Gaussian noise [30, 31] with zero mean and variance  $\sigma_d^2$ . Semicolon 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

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

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  $\kappa^t(\mathbf{r})$  (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  $\boldsymbol{\nu} \leftarrow \kappa^t$  for  $L \leftarrow 2$  latent and  $K \leftarrow 1$  known parameters, respectively.

With  $T_E$  fixed, prescribed flip angles  $\alpha_0$  and repetition times  $T_R$  are the only remaining

---

<sup>1</sup>Though 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.* [28, 29]. In such cases, reconstructed image noise is unlikely to be Gaussian-distributed.

$A \leftarrow 2$  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  $S_{\text{SPGR}}$  scan repetitions to produce  $D \geq L \leftarrow 2$  datasets for well-conditioned estimation.

### 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:  $\kappa^t(\mathbf{r}) \in \mathbb{R}$ ,  $T_1(\mathbf{r}) \in \mathbb{R}$ ,  $\bar{\omega}'(\mathbf{r}) \in \mathbb{R}$ ,  $c_2(\mathbf{r}) := m_0(\mathbf{r})e^{-T_E/T_2^*(\mathbf{r})} \in \mathbb{C}$ , and  $T_2(\mathbf{r}) \in \mathbb{R}$ . In this chapter, we take  $\kappa^t(\mathbf{r}) \in \mathbb{R}$  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 mismatch<sup>2</sup> due to Rician noise. These choices assign  $\boldsymbol{\nu} \leftarrow [\kappa^t, T_1]^\top$  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 ??,  $D \geq L \leftarrow 2$  datasets are necessary for well-conditioned estimation. Unlike before however, a minimum  $D \leftarrow 2$  datasets need not require scan repetition, since  $S_{\text{DESS}}$  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  $\mathbf{N} := [\boldsymbol{\nu}(\mathbf{r}_1), \dots, \boldsymbol{\nu}(\mathbf{r}_V)] \in \mathbb{C}^{K \times V}$  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

<sup>2</sup>The 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 [32], and the acquisitions we examine here are capable of producing SNR in tissue of at minimum 100 and usually more.

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  $\mathbf{N}$ .

### 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 estimator  $\hat{\mathbf{X}}_{\text{ML}}$  by finding an estimate  $\hat{\mathbf{X}}_{\text{ML}}(\mathbf{Y}; \mathbf{N}, \mathbf{P})$  of  $\mathbf{X}$  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*:

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

Here, we have freedom to design regularizer  $\mathbf{R} : \mathbb{C}^{L \times V} \mapsto \mathbb{R}$  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

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

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

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  $\beta_l$  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}

This section demonstrates likelihood-based estimation through two experiments in simulation that correspond to the simple problems defined in Section 3.2.1. Subsection 3.3.1 continues Example 3.2.1.1 in studying  $T_1$  estimation from two SPGR scans. Subsection 3.3.2 continues Example 3.2.1.2 in studying  $T_2$  estimation from one DESS scan.

#### 3.3.1 $T_1$ estimation from two SPGR scans

{ss,relax,exp,t1}

We selected  $T_1$  and  $T_2$  WM and GM values based on previously reported measurements at 3T [33, 34] and extrapolated other nuisance latent object parameters  $m_0$  and  $T_2^*$  from measurements at 1.5T [35]. For simplicity, we assumed no flip angle variation  $\kappa^t \leftarrow 1$  and no phase accrual due to off-resonance effects  $\bar{\omega}' \leftarrow 0$ . We assigned these parameter values to the 81st slice of the BrainWeb digital phantom [36, 35] to create ground truth  $\mathbf{M}_0, \mathbf{T}_1, \mathbf{T}_2, \mathbf{T}_2^* \in \mathbb{R}^V$  maps. We simulated  $217 \times 181$  noiseless single-coil SPGR image datasets, varying nominal flip angles  $\alpha_0 \leftarrow 5, 30^\circ$  and fixing repetition times  $T_R \leftarrow 12.2\text{ms}$  and echo times  $T_E \leftarrow 4.67\text{ms}$  across  $S_{\text{SPGR}} \leftarrow 2$  scans. We corrupted noiseless datasets with additive complex Gaussian noise to yield  $D \leftarrow 2$  noisy complex datasets with signal-to-noise ratio (SNR) ranging from 105-122, where SNR is defined here as

{eq:relax,snr}

$$\text{SNR}(\mathbf{S}, \mathbf{Y}) := \frac{\|\mathbf{S}\|_{\text{F}}}{\|\mathbf{Y} - \mathbf{S}\|_{\text{F}}}. \quad (3.11)$$

We estimated latent parameter maps  $\mathbf{T}_1, \mathbf{C}_1$  using a conventional method-of-moments (MOM) estimator [37], the ML estimator (3.7), and the RL estimator (3.8). The MOM estimator applies linear regression voxel-by-voxel to an appropriately transformed version of the noiseless magnitude SPGR signal model that is linear in  $T_1, c_1$ ; see *e.g.* [37, 38] for details. We next describe our implementations of ML and RL estimation in turn.

The ML estimator applies the variable projection method (VPM; described in Subsection 2.2.2) to separate nonlinear  $\mathbf{T}_1$  estimation from linear  $\mathbf{C}_1$  estimation. Specifically, the algorithm first estimates  $\mathbf{T}_1$  voxel-by-voxel via an exhaustive grid search (over 1000  $T_1$  values logarithmically spaced between  $(10^{1.5}, 10^{3.5})\text{ms}$ ) for a maximizer of the separated least squares cost (2.51). It then estimates  $\mathbf{C}_1$  via per-voxel linear regression (2.50).

The RL estimator applies a preconditioned variant of the classical gradient projection method (GPM; described in Subsection 2.2.1) to iteratively descend towards a local optimizer of the RL cost defined in (3.8). We designed the preconditioner as the inverse of a positive definite diagonal majorizer of the RL cost function’s Hessian matrix, updated for the first five iterations and fixed thereafter. We employed a diagonal preconditioner to

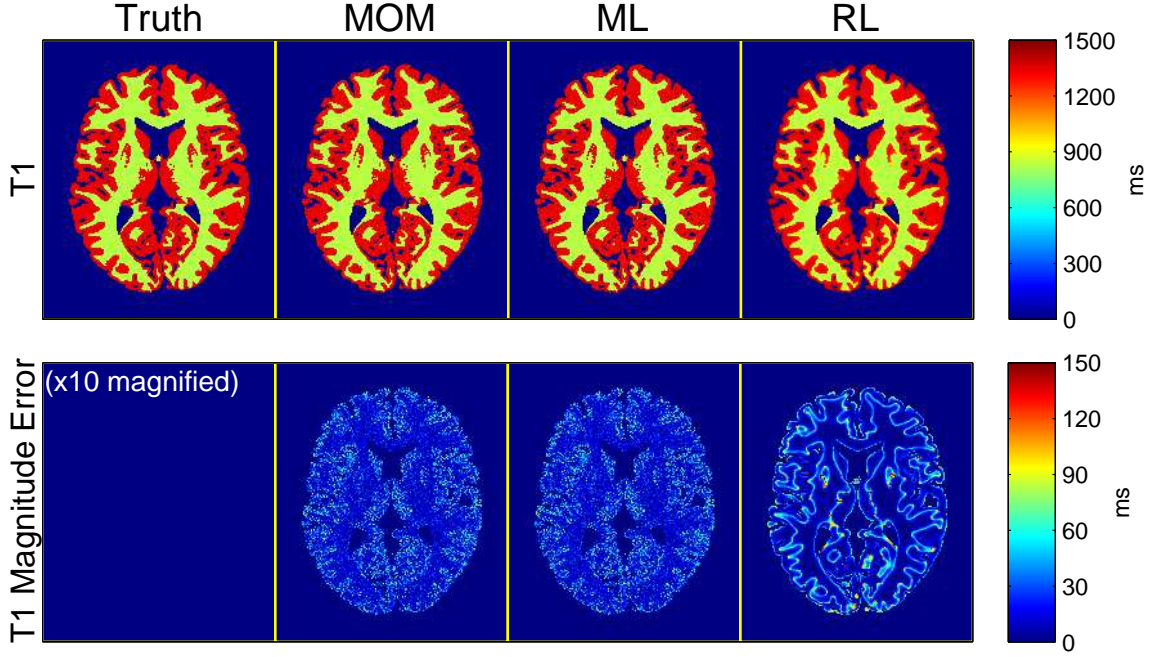


Figure 3.1:  $T_1$  MOM, ML, and RL estimates and corresponding error images, from two simulated SPGR scans. Magnitude error images are  $10\times$  magnified. Voxels not assigned WM- or GM-like relaxation times are masked out in post-processing for display. Table 3.1 presents corresponding sample statistics.

{fig:relax,sim,t1}

retain the linear convergence guarantees of GPM [39] yet approach the practical performance of other unprojected second-order methods (*e.g.*, Newton’s method). We employed a simple step-halving line search at each iteration to ensure monotone local convergence in cost. We initialized GPM with the ML estimates. We selected regularization parameters as described in Subsection 4.4.2.1. We used the Michigan image reconstruction toolbox [40] to construct the regularizer and rapidly evaluate its gradient and Hessian. We used the MATLAB<sup>®</sup> symbolic toolbox to generate analytical expressions for the gradient and Hessian of the SPGR signal model. At each iteration, we used these gradient and Hessian expressions to compute a preconditioned descent direction, update the iterate (possibly after backtracking to ensure descent), and project each voxel’s  $T_1$  iterate to within  $[10, 3000]$ ms. We continued iterations until the convergence criterion

{eq:relax,conv-crit}

$$\|\Omega^{-1}(\mathbf{X}^{(i)} - \mathbf{X}^{(i-1)})\|_F < 10^{-7} \|\Omega^{-1}(\mathbf{X}^{(i)})\|_F \quad (3.12)$$

was satisfied, where  $(\cdot)^{(i)}$  denotes the  $i$ th iterate,  $\Omega := \text{diag}(\text{med}(\mathbf{X}^{(0)}))$  is a weighting matrix, and  $\text{med}(\cdot)$  takes the median across the columns of its argument.

Fig. 3.1 compares MOM, ML, and RL  $T_1$  estimates alongside  $10\times$  magnified absolute

	Truth	MOM	ML	RL
WM $T_1$	832	$832.7 \pm 15.6$	$832.7 \pm 15.6$	$834.00 \pm 2.77$
GM $T_1$	1331	$1332 \pm 34.9$	$1332 \pm 34.9$	$1332.2 \pm 6.3$
WM $c_1$	0.77	$0.7266 \pm 0.00744$	$0.7314 \pm 0.00749$	$0.73184 \pm 0.00475$
GM $c_1$	0.86	$0.8245 \pm 0.0108$	$0.8301 \pm 0.0109$	$0.8287 \pm 0.0059$

Table 3.1: Sample means  $\pm$  sample standard deviations of MOM, ML, and RL  $T_1, c_1$  estimates from two simulated SPGR datasets, computed over 3001 WM-like and 1151 GM-like voxels. Each sample statistic is rounded off to the highest place value of its (unreported) standard error, computed via formulas in [7].  $T_1$  values are in milliseconds.  $c_1$  values are unitless. Fig. 3.1 presents corresponding images.

difference images with respect to the ground truth. Overall, all three estimators produce reasonable  $T_1$  maps. The MOM and ML  $T_1$  estimates are visually quite similar. The RL  $T_1$  estimate exhibits less variability well within WM and GM regions of interest (ROIs), but incurs systematic bias along edges and provides reduced spatial resolution.

Table 3.1 presents  $T_1, c_1$  samples statistics within WM-like and GM-like ROIs selected to contain voxels that are well away from tissue interfaces. In both WM and GM, MOM and ML  $T_1, C_1$  estimates are comparable. RL estimates consistently exhibit lower variability but the RL  $T_1$  estimate exhibits greater bias in WM. All RL bias values would be significantly greater if ROIs instead contained voxels at tissue interfaces.

### 3.3.2 $T_2$ estimation from one DESS scan

## 3.4 Summary

In the completed thesis, this section will demonstrate likelihood-based estimation in two simulation experiments that correspond to the problems defined in Examples ??-??. The results will involve minor implementation improvements to similar results presented in early conference papers [8] and [9], respectively. The results will motivate our use in Chapter 4 of likelihood-based estimation over more conventional methods.

In the completed thesis, this section will call into question the somewhat arbitrary choices of acquisition parameters used in the previous examples and ask whether these acquisition parameters can themselves be designed through optimization. This question will serve as motivation for and as transition into Chapter 4.

## 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 [41]. Neurological applications include: lesion classification in multiple sclerosis [42]; tumor characterization [43, 44]; and symptom onset prediction in stroke [45, 46]. In addition,  $T_1, T_2$  have shown promise for detecting hip and knee cartilage degeneration [47, 48] and for assessing cardiac dysfunction due to iron overload [49] or edema [50]. 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 [51], multi-SE trains [52], or fast k-space trajectories [53, 54, 55] have been proposed to accelerate  $T_1$  [56, 57, 58, 59] and  $T_2$  [60, 61, 62, 63] relaxometry with these classical sequences. These techniques are more sensitive to model non-idealities [64, 65, 66], and are still speed-limited by the long  $T_R$  required for (near)-complete  $T_1$  recovery.

Steady-state (SS) pulse sequences [18, 19] 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 [38, 67] 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 [68, 38, 67, 69, 70, 71, 72, 73, 74, 75].

The dual-echo steady-state (DESS) sequence [24] was recently proposed as a promising SS imaging technique for  $T_2$  estimation [72]. 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 [76, 77] 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 [72, 73] or separate scans to first estimate nuisance parameters [9]. 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 [78, 79], flip angles for spoiled gradient-recalled echo (SPGR) sequences [80], and echo times for SE experiments [81]. More recent studies have considered additional scan design challenges, including scan time constraints [82], multiple latent parameters [83], multiple scan parameter types [84], and latent parameter spatial variation [85, 86, 87].

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  $\mathbf{x}$ )

{eq:scn-dsgn,lf-vec}

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

Under suitable regularity conditions<sup>1</sup>, the Fisher information matrix  $\mathbf{F}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) \in \mathbb{C}^{L \times L}$  [88] characterizes the imprecision of unbiased estimates of  $\mathbf{x}$  from  $\mathbf{y}$ , given  $\boldsymbol{\nu}$  and  $\mathbf{P}$ :

{eq:scn-dsgn,fisher}

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

---

<sup>1</sup>In 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 [89] ensures that any unbiased<sup>2</sup> 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 [90], 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  $\Psi$  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

---

<sup>2</sup>Provided that signal model  $s(\cdot)$  is injective and continuously differentiable in  $\mathbf{x}$ , maximum-likelihood (ML) estimates of  $\mathbf{x}$  (discussed in Subsection 3.2.2.2) are asymptotically consistent and efficient, and thus asymptotically achieve the CRB. For Gaussian noise models, increasing sample size is statistically equivalent to increasing signal-to-noise ratio. Thus, in regions where the data provides sufficiently high SNR (and is thus approximately Gaussian-distributed even in magnitude [32]), ML estimates will exhibit minimal estimation bias, and the CRB can be used to reliably predict ML estimation error. Table 4.2 in Subsection 4.4.1 empirically explores the validity of this high-SNR assumption, through simulations at realistic noise levels.



approach (4.9) should ensure good estimation precision over a range of parameter values.

Since  $\Psi$  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

$$\mathbb{P} := \left\{ \mathbf{P} : \tilde{\Psi}^t(\mathbf{P}) - \tilde{\Psi}^t(\hat{\mathbf{P}}) \leq \delta \tilde{\Psi}^t(\hat{\mathbf{P}}) \right\}. \quad (4.8)$$

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  $\mathbb{N}^b \supseteq \mathbb{N}^t$ :

$$\hat{\mathbf{P}} = \arg \min_{\mathbf{P} \in \mathbb{P}} \tilde{\Psi}^b(\mathbf{P}), \text{ where} \quad (4.9)$$

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

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  $\mathbf{W}$ ,  $\delta$ ,  $\mathbb{X}^t$ ,  $\mathbb{X}^b$ ,  $\mathbb{N}^t$ ,  $\mathbb{N}^b$  fixed. Since  $\Psi$  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

{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

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 ??), given knowledge of transmit field inhomogeneity  $\kappa^t$  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 [68, 67, 69] and subsequent  $T_2$  mapping from other sequences [38, 9]. 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 ??-??, each SPGR and DESS scan leaves  $\mathbf{p} \leftarrow [\alpha_0, T_R]^T$  as  $A \leftarrow 2$  acquisition parameters available to optimize. A given scan profile consisting of  $S_{\text{SPGR}}$  SPGR and  $S_{\text{DESS}}$  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  $\mathbb{A}_{0,\text{SPGR}}^{S_{\text{SPGR}}} \times \mathbb{A}_{0,\text{DESS}}^{S_{\text{DESS}}} \times \mathbb{T}_{\text{R},\text{SPGR}}^{S_{\text{SPGR}}} \times \mathbb{T}_{\text{R},\text{DESS}}^{S_{\text{DESS}}}$  (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  $\mathbb{T}_{\text{R},\text{SPGR}} \leftarrow [12.2, +\infty)\text{ms}$  and  $\mathbb{T}_{\text{R},\text{DESS}} \leftarrow [17.5, +\infty)\text{ms}$  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

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

to satisfy a total time constraint,  $\|\mathbf{T}_R\|_1 \leq T_{\text{max}}$ . For a scan profile consisting of  $S_{\text{SPGR}}$  SPGR and  $S_{\text{DESS}}$  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 [80, 38] or 1 DESS [72] 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  $\Psi$  then depends on  $c_2$  through only a scale factor, it suffices to fix  $c_2 \leftarrow 1$  and design the latent 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 [33, 34]. 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  $\tilde{\Psi}^t(\check{\mathbf{P}})$ .

### 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<sup>®</sup> 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

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) $^\circ$
$\hat{T}_R^{\text{spgr}}$	(12.2, 12.2)	13.9	–
$\hat{T}_R^{\text{dess}}$	17.5	28.0	(24.4, 17.5)
$\tilde{\sigma}_{T_1}^t(\hat{\mathbf{P}})$	28	27	<b>21</b>
$\tilde{\sigma}_{T_1}^b(\hat{\mathbf{P}})$	154	169	<b>113</b>
$\tilde{\sigma}_{T_2}^t(\hat{\mathbf{P}})$	<b>1.3</b>	2.8	1.5
$\tilde{\sigma}_{T_2}^b(\hat{\mathbf{P}})$	9.1	8.8	<b>6.0</b>
$\tilde{\Psi}^t(\hat{\mathbf{P}})$	4.0	4.9	<b>3.5</b>
$\tilde{\Psi}^b(\hat{\mathbf{P}})$	17.7	17.9	<b>12.2</b>

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}

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 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 [72], 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.

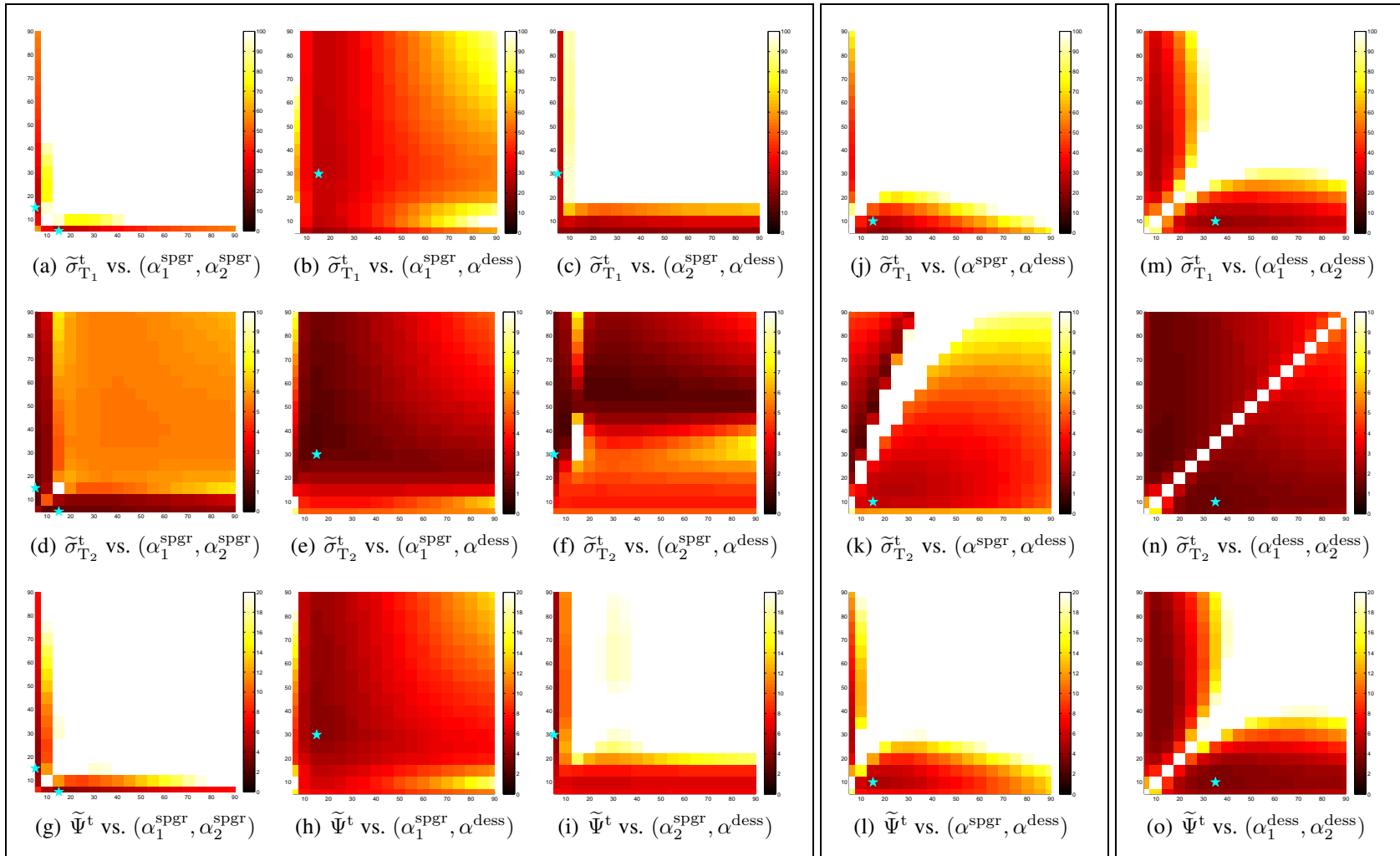


Figure 4.1: Worst-case standard deviations  $\tilde{\sigma}_{T_1}^t$  (top),  $\tilde{\sigma}_{T_2}^t$  (middle), and cost  $\tilde{\Psi}^t$  (bottom), versus pairs of nominal flip angles, holding other scan parameters fixed at selected profile  $\hat{\mathbf{P}}$ . Subfigures (a)-(i), (j)-(l), and (m)-(o) correspond to scan profiles containing  $(S_{\text{SPGR}}, S_{\text{DESS}}) = (2, 1), (1, 1), \text{ and } (0, 2)$  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.

## 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  $T_1$  and  $T_2$  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 Sections 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

{ss,scn-dsgn,exp,sim}

We select  $T_1$  and  $T_2$  WM and GM values based on previously reported measurements at 3T [33, 34] and extrapolate other nuisance latent object parameters  $m_0$  and  $T_2^*$  from measurements at 1.5T [35]. We assign these parameter values to the discrete anatomy of the BrainWeb digital phantom [36, 35] 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  $217 \times 181$  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

{eq:scn-dsgn,snr}

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

We use each profile’s noisy magnitude dataset  $|\mathbf{Y}|$  to compute estimates  $\hat{T}_1$  and  $\hat{T}_2$ . 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  $\kappa^t$  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 considera-

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

Table 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.

{table:numerical}

tion 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  $\hat{\mathbf{P}}$ .

Table 4.2 verifies<sup>3</sup> 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  $\hat{T}_1^{\text{ML}}$  precision, while the (2, 1) and (0, 2) scans afford high  $\hat{T}_2^{\text{ML}}$  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  $T_1$  and  $T_2$  estimation precision in WM and GM comparable to optimized profiles containing SPGR/DESS combinations.

Fig. 4.2 histograms (voxel-wise independent) ML estimates  $\hat{T}_1^{\text{ML}}$  and  $\hat{T}_2^{\text{ML}}$  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,  $\hat{T}_1^{\text{ML}}$  and  $\hat{T}_2^{\text{ML}}$  exhibit negligible bias and near-Gaussian shape, suggesting locally linear behavior of the DESS signal model in  $T_1$  and  $T_2$  ( $\hat{T}_1^{\text{ML}}$  and  $\hat{T}_2^{\text{ML}}$  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

<sup>3</sup>Each 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.

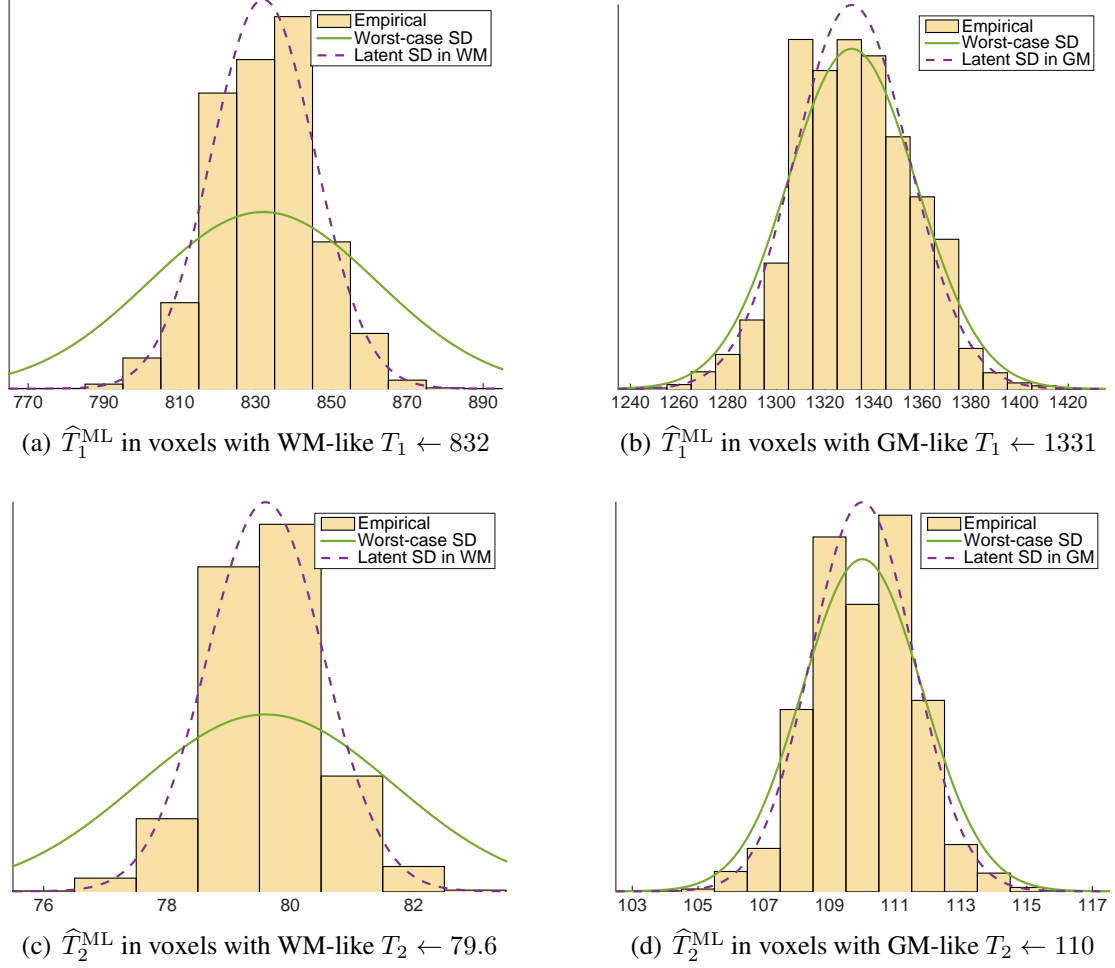


Figure 4.2: Histograms of  $T_1$  and  $T_2$  estimates from noisy independent measurements of a *single* nominal WM or GM value. In each plot, two normal distributions are overlaid, each with latent means  $T_1$  and  $T_2$ . In (a)-(b) and (c)-(d), the solid green curve is  $\mathcal{N}(T_1, (\tilde{\sigma}_{T_1}^t)^2)$  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  $\hat{T}_1^{\text{ML}}$  and  $\hat{T}_2^{\text{ML}}$  errors.

{fig:normality}

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.

## 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<sup>®</sup> 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 [91], 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  $240 \times 240 \times 40$  mm<sup>3</sup> 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  $256 \times 256$  datasets over the same



$240 \times 240 \times 5 \text{ mm}^3$  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-Siebert shifted 3D SPGR scans for separate transmit field estimation [92]. We insert a 9ms Fermi pulse at  $\pm 8\text{kHz}$  off-resonance into an SPGR sequence immediately following on-resonant excitation. We estimate regularized transmit field maps [93] 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 Discovery<sup>TM</sup> 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 [94] 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++ [95] 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

{eq:conv,crit}

$$\|X^{(i)} - X^{(i-1)}\|_F < 10^{-7} \|X^{(i)}\|_F \quad (4.14)$$

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 sequential<sup>4</sup>  $T_1$ , then  $T_2$  estimation from IR and SE reference scans.

---

<sup>4</sup> 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 [96].

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 [97], which accounts for imperfect excitation/refocusing 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 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 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 [98, 99, 100], (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.

#### 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

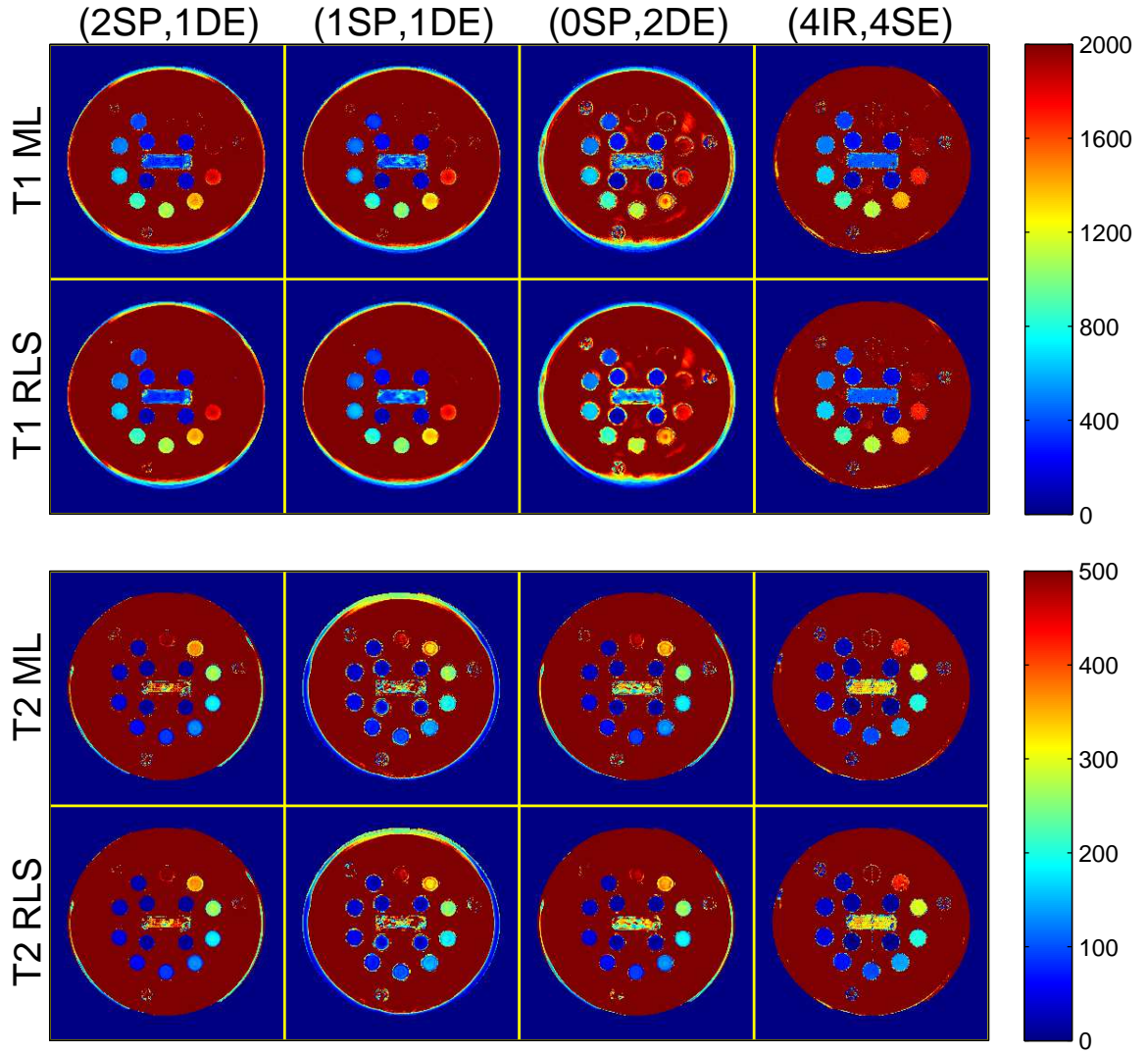


Figure 4.3: Colorized  $T_1$  and  $T_2$  ML and RL estimates from an HPD<sup>®</sup> quantitative phantom. Columns correspond to scan profiles consisting of (2 SPGR, 1 DESS), (1 SPGR, 1 DESS), (0 SPGR, 2 DESS), and (4 IR, 4 SE) acquisitions. Rows distinguish  $T_1$  and  $T_2$  ML and RL estimators.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

{fig:hpj, jet}

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 [7]) for phantom vials within (or nearly within) tight design range  $\mathbb{X}^t$  (marked orange in Fig. 4.4). Due to error propagation

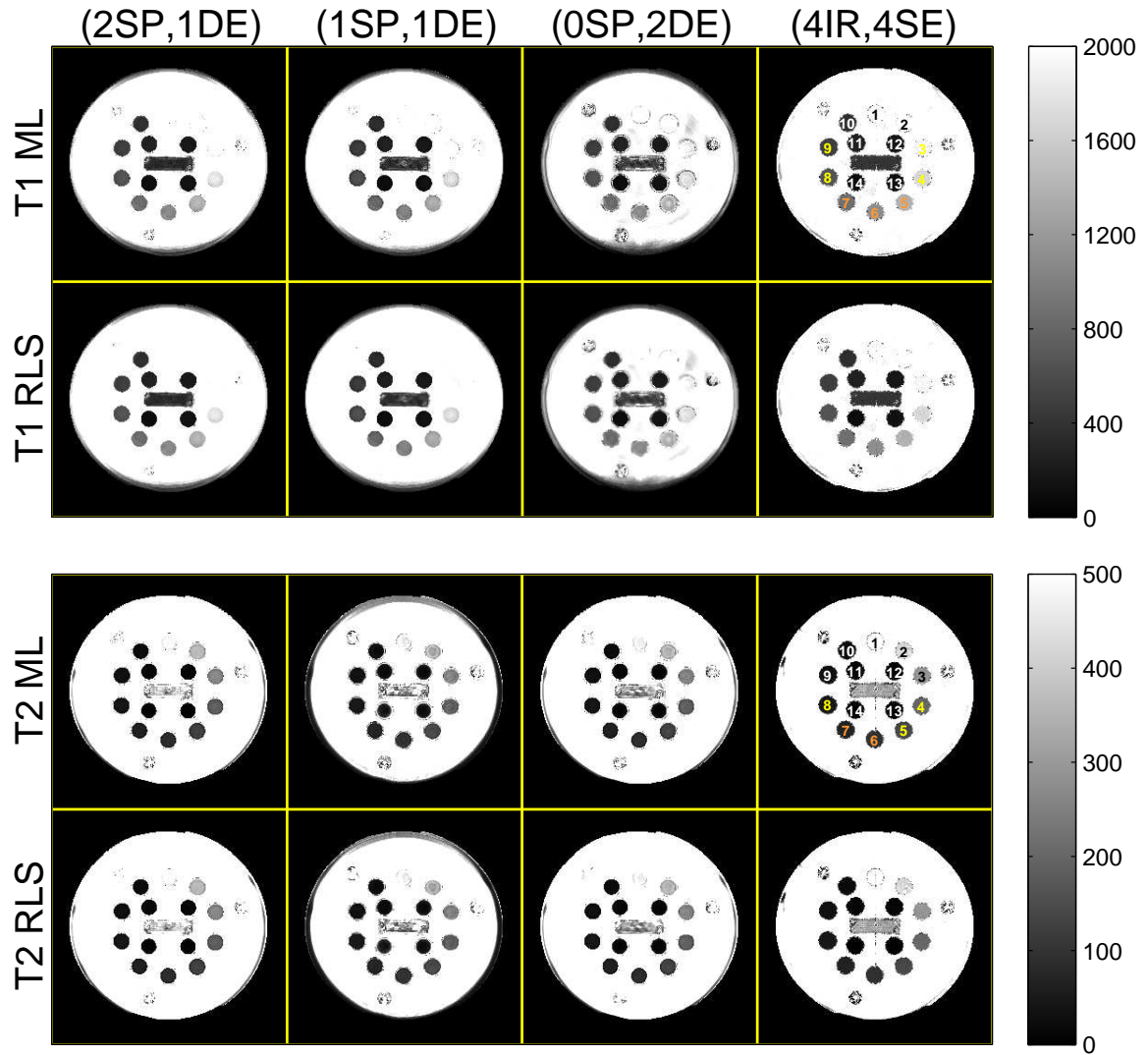


Figure 4.4: Grayscale  $T_1$  and  $T_2$  ML and RL estimates from an HPD<sup>®</sup> quantitative phantom. Columns correspond to scan profiles consisting of (2 SPGR, 1 DESS), (1 SPGR, 1 DESS), (0 SPGR, 2 DESS), and (4 IR, 4 SE) acquisitions. Rows distinguish  $T_1$  and  $T_2$  ML and RL estimators. 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.

{fig:hpd,gray}

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.

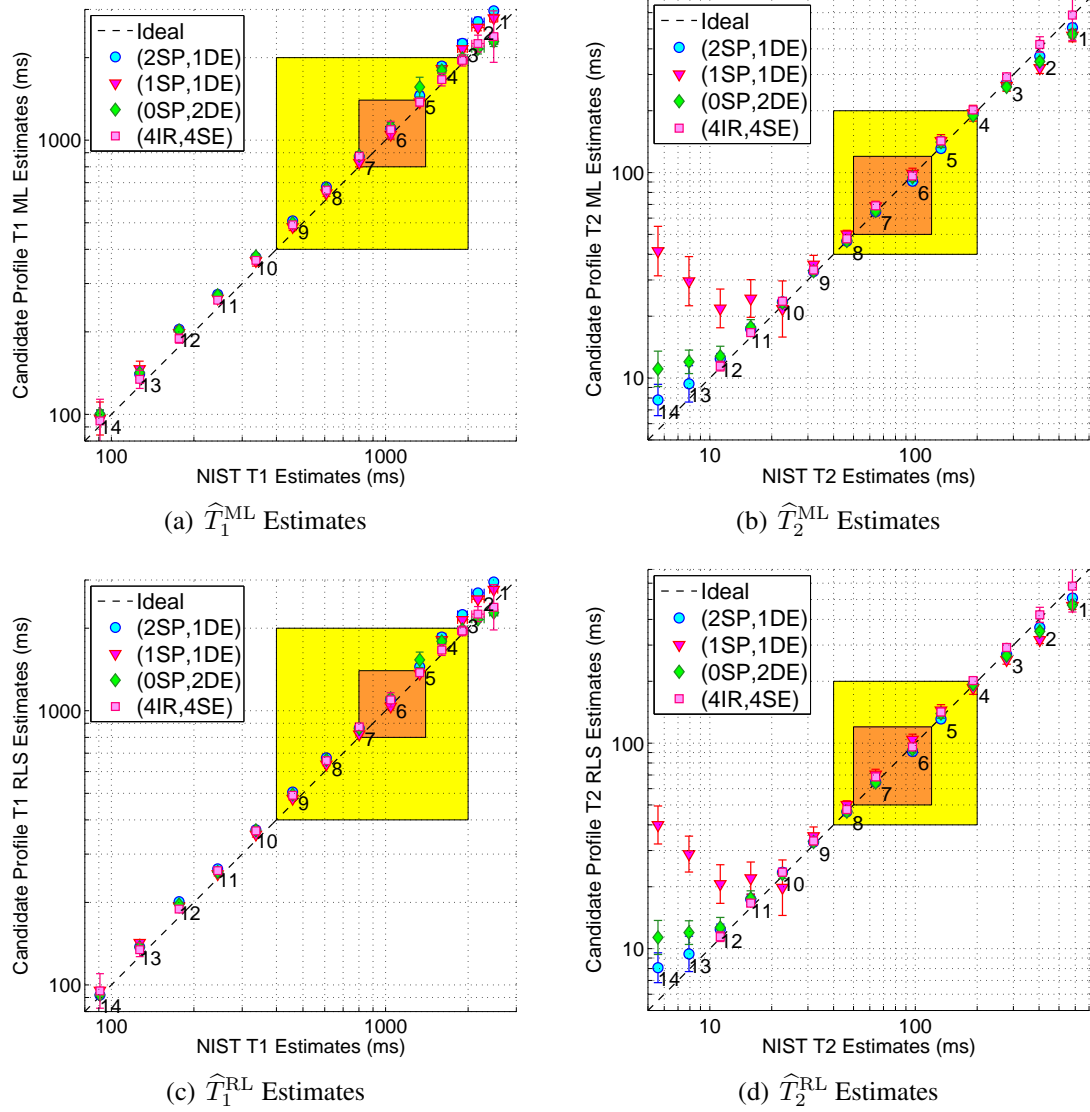


Figure 4.5: Phantom within-ROI sample statistics of  $T_1$  and  $T_2$  ML and RL estimates from optimized SPGR/DESS and reference IR/SE scan profiles, versus NIST NMR measurements [1]. Markers and error bars indicate ROI sample means and ROI sample standard deviations within the 14 labeled and color-coded 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:hpdp,ml-rls}



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

#### 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<sup>®</sup> receive head array. To address bulk motion between acquisitions and to compare within-ROI statistics, we rigidly register<sup>5</sup> 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.

<sup>5</sup>For 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<sup>®</sup> 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.

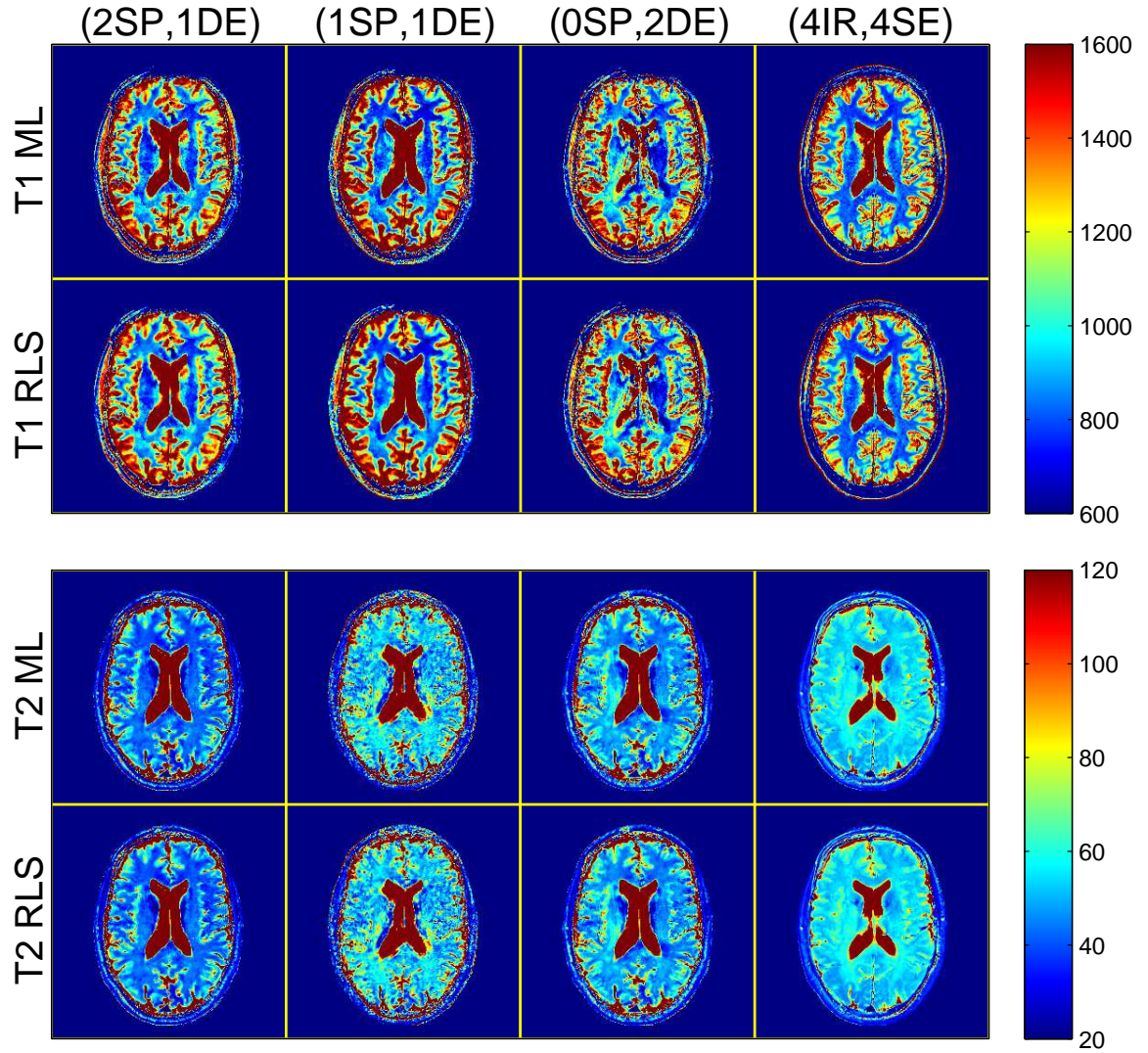


Figure 4.6: Colorized  $T_1$  and  $T_2$  ML and RL estimates from the brain of a healthy volunteer. Columns correspond to profiles consisting of (2 SPGR, 1 DESS), (1 SPGR, 1 DESS), (0 SPGR, 2 DESS), and (4 IR, 4 SE) acquisitions. Rows distinguish  $T_1$  and  $T_2$  ML and RL estimators. Table 4.5 presents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

{fig:brain,jet}

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

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.

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.

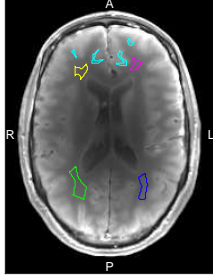
Visually,  $\hat{T}_1$  maps from steady-state profiles exhibit similar levels of contrast in WM/GM regions well away from cerebrospinal fluid (CSF) as that seen in the reference  $\hat{T}_1$  estimate. Since we did not optimize any scan profiles for estimation in high- $T_1$  regions, it is expected 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 computed<sup>6</sup> 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.

<sup>6</sup>We 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.



{fig:brain.roi.gray}



	ROI	(2SP,1DE)	(1SP,1DE)	(0SP,2DE)	(4IR,4SE)
$\hat{T}_1^{\text{ML}}$	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
$\hat{T}_1^{\text{RL}}$	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
$\hat{T}_2^{\text{ML}}$	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
$\hat{T}_2^{\text{RL}}$	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.

In most cases,  $\hat{T}_1$  within-ROI sample means from optimized SPGR/DESS scan profiles do not deviate substantially from each other or from reference IR/SE measurements. Two notable exceptions are  $\hat{T}_1^{\text{ML}}$  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  $\hat{T}_1^{\text{ML}}$  estimates that are more consistently similar to IR WM  $\hat{T}_1^{\text{ML}}$  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  $\hat{T}_2^{\text{ML}}$  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  $\hat{T}_2^{\text{ML}}$  than the reference IR/SE profile, though the (0, 2) profile is in reasonable agreement with

other steady-state estimates [101]. These discrepancies may be due to differences in sensitivity to multi-compartmental relaxation [4]. 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 [81, 79, 82, 83], 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 [85, 86] have addressed object parameter spatial variation by instead constructing cost functions related to the Bayesian CRB [102], 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 [85]. However, prior distributions are generally unknown, and may need to be estimated from data, as in [86].

Careful calibration of flip angle scaling  $\kappa^t$  is essential for accurate  $T_1, T_2$  estimation from SPGR/DESS scan profiles. In this work, we estimate  $\kappa^t$  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 er-

rors to propagate into our  $T_1, T_2$  estimates. To reduce error propagation, it may be desirable to instead design scan profiles to permit *joint* estimation of  $\kappa^t$ , in addition to other latent object parameters. Unfortunately, we find that optimizing the (2, 1) or (0, 2) profile to allow for four-parameter  $\mathbf{x}(\mathbf{r}) \leftarrow [T_1(\mathbf{r}), T_2(\mathbf{r}), c_2(\mathbf{r}), \kappa^t(\mathbf{r})]^T$  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 [9].) It remains an open scan design question as to whether time spent collecting Bloch-Siegert data for separate  $\kappa^t$  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 [62, 103]. In practice, difficulties arise in efficient computation of signal gradients required in (4.3), which may demand more specialized techniques, as in [104]. 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) [9] 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

# Dictionary-Free MRI Parameter Estimation via Regression with Kernels (PERK)

{c,perk}

### 5.1 Introduction

{s,krr,intro}

In MRI *parameter estimation*, one seeks to quantify biomarker “maps” (*i.e.*, parameter images) from data. Because MR acquisitions are tunably sensitive to many physical processes (*e.g.*, relaxation [17], diffusion [105], and chemical exchange [106]), MRI parameter estimation is important in many QMRI applications (*e.g.*, relaxometry [107], diffusion tensor imaging [108], and multi-compartmental imaging [4]). 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 [8, 9, 10] and others’ [109, 103, 62, 110] works approach estimation through exhaustive grid search that 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 [111]) and tensor (*e.g.*, diffusivity [108] or conductivity [112]) field mapping techniques require estimation of at minimum 4 and 7 latent parameters per voxel, respectively. Phase-based longitudinal [113] or transverse [114, 92] 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 [4] from steady-state pulse sequences requires estimation of at least 7 [115] and as many as 10 [116] latent parameters per voxel. In these and other applications, greater estimation accuracy requires more complete signal models that involve more latent parameters, increasing 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 [117] showed that simple transformations involving so-called *reproducing kernel* functions [118] 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 [119] and quickly thereafter for other applications, *e.g.* regression [120]) 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  $\rho_l$  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) [118]. Remarkably, a generalization [121] of the Representer Theorem [117], restated as is relevant here for completeness, then reduces (5.4) to a finite-dimensional optimization problem:

**Theorem 1 (Generalized Representer, [121])** *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*

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

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

Theorem 1 ensures that any solution to the  $(N + 1)$ -dimensional optimization problem

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

corresponds via (5.5) to a minimizer of (5.4) over  $\mathbb{H} \times \mathbb{R}$ , where  $\mathbf{a}_l := [a_{l,1}, \dots, a_{l,N}]^\top$ . Because the objective function within (5.6) is a strictly convex quadratic function in  $(\mathbf{a}_l, b_l)$  for  $\rho_l > 0$ , a solution to (5.6) exists uniquely and is expressed as

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

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

where  $\mathbf{K} \in \mathbb{R}^{N \times N}$  is the Gram matrix consisting of entries  $k(\mathbf{q}_n, \mathbf{q}_{n'})$  for  $n, n' \in \{1, \dots, N\}$ ;  $\mathbf{M} := \mathbf{I}_N - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^\top$  is a de-meaning operator;  $\mathbf{x}_l := [x_{l,1}, \dots, x_{l,N}]^\top$ ;  $\mathbf{I}_N \in \mathbb{R}^{N \times N}$  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

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

where  $\mathbf{k}(\cdot) := [k(\cdot, \mathbf{q}_1), \dots, k(\cdot, \mathbf{q}_N)]^\top - \frac{1}{N}\mathbf{K}\mathbf{1}_N : \mathbb{R}^Q \mapsto \mathbb{R}^N$  is a kernel embedding operator and  $\mathbf{X} := [\mathbf{x}_1, \dots, \mathbf{x}_N] = [\mathbf{x}_1, \dots, \mathbf{x}_L]^\top \in \mathbb{R}^{L \times N}$  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  $k$  that 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 [122], which is optimal over  $(\mathbb{R}^Q \times \mathbb{R})^L$  but is unlikely to be useful when signal model  $\mathbf{s}(\cdot)$  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) kernel  $k$  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 matrix  $\Lambda \in \mathbb{R}^{Q \times Q}$  controls the length scales in  $\mathbf{q}$  over which the estimator  $\hat{\mathbf{x}}$  smooths. We use the Gaussian kernel over other candidates 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) [123], a recent method for approximating shift-invariant kernels

having form  $k(\mathbf{q}, \mathbf{q}') \equiv k(\mathbf{q} - \mathbf{q}')$ . This subsection reviews the main result of [123] for the purpose of constructing an intuitive and computationally efficient approximation of (5.9).

The strategy of [123] 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

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

where  $E_{v,s}(\cdot)$  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  $Z$  [123, 124], justifies the interpretation of  $\tilde{\mathbf{z}}_Z$  as an approximate (and now finite-dimensional) feature map.

We use the Fourier construction of [123] 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  $E_{v,s}(\tilde{z}(\mathbf{v}, s; \mathbf{q})\tilde{z}(\mathbf{v}, s; \mathbf{q}'))$  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 [125] 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 [126] 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 automatically tune as many model parameters as possible, thereby leaving as few parameters as possible to manual selection. In this spirit, we focus on “online” model selection, where one chooses model parameters for training the estimator  $\hat{\mathbf{x}}(\cdot)$  *after* observation of (unlabeled) real test data. This approach is unconventional in traditional machine learning applications because training often takes considerably longer than testing. However, many voxel-wise separable QMRI parameter estimation problems are comparatively low-dimensional and so KRR-based estimators can be quickly trained using only a moderate number of simulated training examples; in fact, applying the estimator (especially on full-volume high-resolution imaging data) can take time comparable to or even greater than training. For these reasons, online model selection is practical for KRR-based QMRI parameter estimation.

{sss,krr,pract,mod,dist}

#### 5.3.2.1 Choosing Sampling Distributions

To sample training points, we must select prior distributions on object parameters  $\mathbf{x}$ ,  $\boldsymbol{\nu}$  and noise vector  $\epsilon$ . Because we do not possess many separate prior estimates of latent object parameters, we must assume a prior distribution on  $\mathbf{x}$ , and do so balancing prior knowledge of the model physics (to focus samples on useful parameter ranges) against generously heavy distribution tails (to enable precise estimation over broad parameter ranges). However, we *can* automatically estimate distributions on  $\boldsymbol{\nu}$  and  $\epsilon$  with online model selection. We estimate known object parameter distributions directly from known  $\boldsymbol{\nu}$  via kernel density estimation [127]. We estimate the noise covariance from low-signal regions of real image data (and could alternatively have used readouts with no RF excitation).

{sss,krr,pract,mod,rho}

#### 5.3.2.2 Choosing Regularization Parameters

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 variability*  $\epsilon_{\mathbf{x}}$  because variation is *anticipated* in parameter estimates and is thus intentionally *simulated* (implicitly through additive measurement

noise) during training. If this latent parameter variability is assumed to be distributed as  $\mathcal{N}(\mathbf{0}_L, N \text{diag}([\rho_1, \dots, \rho_L]^\top))$ , 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 [128, Ch. 2] for derivations). In a Bayesian perspective then,  $N\rho_1, \dots, N\rho_L$  should reflect latent parameter variability.

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

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

$$\{\text{eq:krr,model-approx}\} \quad \mathbf{s}(\mathbf{x}, \boldsymbol{\nu}) \approx \mathbf{E}_x(\mathbf{x}) + [\nabla_x \mathbf{s}(\mathbf{E}_x(\mathbf{x}); \boldsymbol{\nu})](\mathbf{x} - \mathbf{E}_x(\mathbf{x})). \quad (5.16)$$

Substituting (5.16) into a standard covariance identity gives

$$\{\text{eq:krr,cov-y}\} \quad \text{cov}(\mathbf{y}) = \text{cov}(\mathbf{s}(\mathbf{x}; \boldsymbol{\nu})) + \text{cov}(\boldsymbol{\epsilon}) \quad (5.17)$$

$$\{\text{eq:krr,cov-y,approx}\} \quad \approx [\nabla_x \mathbf{s}(\mathbf{E}_x(\mathbf{x}); \boldsymbol{\nu})] \text{cov}(\mathbf{x}) [\nabla_x \mathbf{s}(\mathbf{E}_x(\mathbf{x}); \boldsymbol{\nu})]^\text{H} + \boldsymbol{\Sigma}. \quad (5.18)$$

where in (5.17) we have assumed that latent parameters  $\mathbf{x}$  and noise  $\boldsymbol{\epsilon}$  are uncorrelated. Subtracting  $\boldsymbol{\Sigma}$  from each side, taking the Moore-Penrose pseudoinverse (denoted by  $(\cdot)^\dagger$ ), and pre- and post-multiplying by  $[\nabla_x \mathbf{s}(\mathbf{E}_x(\mathbf{x}); \boldsymbol{\nu})]^\text{H}$  and  $[\nabla_x \mathbf{s}(\mathbf{E}_x(\mathbf{x}); \boldsymbol{\nu})]$  together yield

$$\{\text{eq:krr,messy}\} \quad \begin{aligned} & [\nabla_x \mathbf{s}(\mathbf{E}_x(\mathbf{x}); \boldsymbol{\nu})]^\text{H} (\text{cov}(\mathbf{y}) - \boldsymbol{\Sigma})^\dagger [\nabla_x \mathbf{s}(\mathbf{E}_x(\mathbf{x}); \boldsymbol{\nu})] \approx \\ & [\nabla_x \mathbf{s}(\mathbf{E}_x(\mathbf{x}); \boldsymbol{\nu})]^\text{H} \left( [\nabla_x \mathbf{s}(\mathbf{E}_x(\mathbf{x}); \boldsymbol{\nu})]^\text{H} \right)^\dagger (\text{cov}(\mathbf{x}))^{-1} ([\nabla_x \mathbf{s}(\mathbf{E}_x(\mathbf{x}); \boldsymbol{\nu})])^\dagger [\nabla_x \mathbf{s}(\mathbf{E}_x(\mathbf{x}); \boldsymbol{\nu})]. \end{aligned} \quad (5.19)$$

If the Jacobian matrix  $[\nabla_x \mathbf{s}(\mathbf{E}_x(\mathbf{x}); \boldsymbol{\nu})]$  is of full column rank (and is thus left-invertible), Eq. (5.19) simplifies after matrix inversion to

$$\{\text{eq:krr,cov-x}\} \quad \text{cov}(\mathbf{x}) \approx \left( [\nabla_x \mathbf{s}(\mathbf{E}_x(\mathbf{x}); \boldsymbol{\nu})]^\text{H} (\text{cov}(\mathbf{y}) - \boldsymbol{\Sigma})^\dagger [\nabla_x \mathbf{s}(\mathbf{E}_x(\mathbf{x}); \boldsymbol{\nu})] \right)^{-1}, \quad (5.20)$$

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

Expression (5.20) approximates latent parameter noise covariance but is not yet useful because it remains a function of  $\mathbf{y}$  and  $\boldsymbol{\nu}$ , both of which vary spatially. As a simple solution, one could set regularization parameters  $\rho_1, \dots, \rho_L$  using diagonal elements of

$$\{\text{eq:krr,cov-x,approx}\} \quad \left( N [\nabla_x \mathbf{s}(\mathbf{E}_x(\mathbf{x}); \bar{\boldsymbol{\nu}})]^\text{H} (\text{cov}(\bar{\mathbf{y}}) - \boldsymbol{\Sigma})^\dagger [\nabla_x \mathbf{s}(\mathbf{E}_x(\mathbf{x}); \bar{\boldsymbol{\nu}})] \right)^{-1}, \quad (5.21)$$

where  $\bar{\mathbf{y}} := \frac{1}{V} \sum_{v=1}^V \mathbf{y}(\mathbf{r}_v)$  and  $\bar{\boldsymbol{\nu}} := \frac{1}{V} \sum_{v=1}^V \boldsymbol{\nu}(\mathbf{r}_v)$  are sample means of the unlabeled real test data and the known object parameters, respectively.

Regularization parameter selection via (5.21) 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  $\mathbf{C}_{\mathbf{x}\mathbf{x}} := \frac{1}{N}\mathbf{X}\mathbf{M}\mathbf{X}^\top$  from simulated training data points and set  $\rho_1, \dots, \rho_L$  as diagonal elements of  $\frac{1}{N}\mathbf{C}_{\mathbf{x}\mathbf{x}}$ . In the ensuing preliminary experiment, we set regularization parameters using the sample covariance matrix.

### 5.3.2.3 Choosing Smoothing Length Scale

With online training, it is intuitive to set the Gaussian kernel's smoothing length scale  $\Lambda$  based on test data sample means  $\bar{\mathbf{y}}$  and known parameter sample means  $\bar{\boldsymbol{\nu}}$ :

$$\Lambda \leftarrow \lambda_0 \text{diag} \left( \left[ \text{Re}(\bar{\mathbf{y}})^\top, \text{Im}(\bar{\mathbf{y}})^\top, \bar{\boldsymbol{\nu}}^\top \right]^\top \right). \quad (5.22)$$

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 [123], studied in Proposition 4 of [124] 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 be more challenging to use.

**Theorem 2 ([124])** *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  $\tilde{\mathbf{z}}_Z(\mathbf{q})^\top \tilde{\mathbf{z}}_Z(\mathbf{q}') - k(\mathbf{q}, \mathbf{q}')$  from crossing zero, as is typical. Then*

$$\mathbb{E} \left( \sup_{\mathbf{q}, \mathbf{q}' \in \mathbb{Q}} \left| \tilde{\mathbf{z}}_Z(\mathbf{q})^\top \tilde{\mathbf{z}}_Z(\mathbf{q}') - k(\mathbf{q}, \mathbf{q}') \right| \right) \leq \frac{48c_{\mathbb{Q}}d_{\mathbb{Q}}\sqrt{Q}}{\sqrt{Z}}(v + r), \quad (5.23)$$

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

$$\{eq:krr,lipschitz\} \quad v := \sup_{\Delta \in \mathbb{Q}_\Delta} \|\nabla_\Delta k\|_2 \leq \sup_{\Delta \in \mathbb{R}^Q} \|\nabla_\Delta k\|_2 = \sqrt{\frac{Q}{e \operatorname{tr}(\mathbf{\Lambda}^2)}}. \quad (5.24)$$

The best lower bound we have presently for  $r$  is

$$\{eq:krr,max-nu\} \quad r \leq \frac{1}{2\pi \min \operatorname{eig}(\mathbf{\Lambda})} \left( \sqrt{2 \log Z} + \sqrt{Q} \right), \quad (5.25)$$

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

$$\begin{aligned} \{eq:approx-tight\} \quad & \mathbb{E} \left( \sup_{\mathbf{q}, \mathbf{q}' \in \mathbb{Q}} \left| \tilde{\mathbf{z}}_Z(\mathbf{q})^\top \tilde{\mathbf{z}}_Z(\mathbf{q}') - k(\mathbf{q}, \mathbf{q}') \right| \right) \\ & \leq \frac{96(0.803)(\max_v(\|\mathbf{q}(\mathbf{r}_v)\|_2))\sqrt{Q}}{\sqrt{Z}} \left( \sqrt{\frac{Q}{e \operatorname{tr}(\mathbf{\Lambda}^2)}} + \frac{\sqrt{2 \log Z} + \sqrt{Q}}{2\pi \min \operatorname{eig}(\mathbf{\Lambda})} \right) \end{aligned} \quad (5.26)$$

$$\{eq:approx-loose\} \quad \leq \frac{96(0.803)(\max_v(\|\mathbf{q}(\mathbf{r}_v)\|_2))\sqrt{Q}}{\min \operatorname{eig}(\mathbf{\Lambda}) \sqrt{Z}} \left( \frac{1}{\sqrt{e}} + \frac{\sqrt{2 \log Z} + \sqrt{Q}}{2\pi} \right). \quad (5.27)$$

After observing unlabeled real test data and selecting  $\mathbf{\Lambda}$  (perhaps automatically via (5.22)), all quantities except  $Z$  on the right-hand side of inequality (5.26) are known. Thus, one could intuitively set  $Z$  by choosing a expected maximal approximation error threshold and finding the minimum  $Z$  for which the right-hand side of (5.26) is less than the threshold.

In practice, we find that this procedure yields conservatively large choices for  $Z$ , and so we presently instead tune  $Z$  manually to yield low empirical approximation error. Nevertheless, we have gained from the above analysis two points of intuition that may aid in guiding manual tuning. First, we observe that scaling  $Z$  quadratically with  $Q$  will asymptotically maintain a given expected maximal approximation error tolerance. Second, we see (more clearly via loosened bound (5.27)) that selecting  $\mathbf{\Lambda}$  via (5.22) makes the approximation error largely invariant to the scale of test data; thus  $Z$  can be reasonably selected predominantly based on regressor dimension  $Q$  (as opposed to other model parameters).

## 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 (six) 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 [4], fast steady-state pulse sequences were recently proposed for two-compartment parameter estimation [115, 2]. We similarly investigate steady-state sequences, but modify acquisition details to address possible concerns of insufficient estimation precision [5].

We begin with two-compartment models of the signals arising from Spoiled Gradient-Recalled Echo (SPGR) [21] and Dual-Echo Steady-State (DESS) [24] 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)  $\kappa^t$ ; 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 constant<sup>1</sup>  $c_3$  (defined explicitly in Section 6.4). We assume prior knowledge of  $\nu \leftarrow \kappa^t$  (which in practice can be estimated from separate fast acquisitions, *e.g.* [92]) and collect the remaining  $L \leftarrow 6$  latent parameters as  $\mathbf{x} \leftarrow [f_F, T_{1,F}, T_{2,F}, T_{1,S}, T_{2,S}, c_3]^T$ .

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 scans<sup>2</sup> 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 arising from two non-exchanging water compartments with nominal fast ( $T_{1,F}, T_{2,F} \leftarrow (500, 20)\text{ms}$ ) and slow ( $T_{1,S}, T_{2,S} \leftarrow (1000, 80)\text{ms}$ ) relaxation time constants selected from prior measurements [4, 2]. 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 [36] to produce realistic ground truth parameter maps. Using optimized acquisition parameters and allowing  $\kappa^t$  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 [10]. This yields realistically noisy  $217 \times 181$  (image-domain) datasets ranging

<sup>1</sup>We collect off-resonance effects in  $c_3$  by approximating broadening distributions to be constant across compartments, as in [115, 2]. We acknowledge this simplification could lead to some bias in practice.

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



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 [10], off-resonance related phase (unlike signal loss) cannot be collected into  $c_3$ . 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  $\nu$  and noise  $\epsilon$  distributions from test data. We take fast-compartment fraction  $f_F$  to be uniformly distributed on  $[-0.1, 0.4]$ , and note that we sample negative values<sup>3</sup> with nonzero probability to ensure sampling bias does not contribute significantly to  $f_F$  estimation bias (especially in low- $f_F$  regions). We take  $T_{1,F}$ ,  $T_{2,F}$ ,  $T_{1,S}$ ,  $T_{2,S}$  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  $c_3$  to be uniformly distributed on  $[2.2 \times 10^{-16}, u]$ , where  $u$  is set as  $5 \times$  the maximum value of test data. We estimate the distribution of  $\nu$  via kernel density estimation [127]. We assume noise covariance  $\Sigma$  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 [129].

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  $\Lambda$  via (5.22), with  $\lambda_0 \leftarrow 1$ . We sample  $(\mathbf{v}, s) \ Z \leftarrow 10^3$  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  $\rho_1, \dots, \rho_L$  as diagonal entries of  $\frac{1}{N} \mathbf{C}_{xx}$  and evaluate (5.15) for each latent parameter using test datasets and  $\kappa^t$  on a per-voxel basis. On a 3.5GHz desktop computer with 32GB RAM running MATLAB<sup>®</sup> R2013a, training and estimating each took 35s and 5s.

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) [27] and grid search. As presented, the VPM estimate utilizes a dictionary of nearly  $8 \times 10^6$  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

---

<sup>3</sup>Observe that our two-compartment signal models are linear in  $f_F$  and are therefore well-defined for zero or even negative  $f_F$  values.



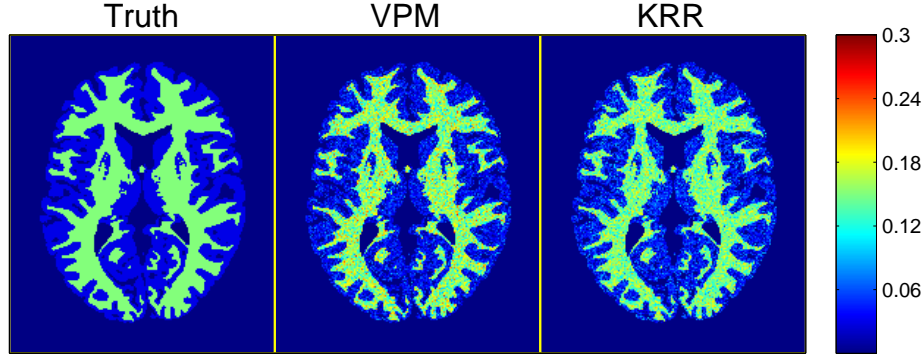


Figure 5.1: True  $f_F$  (left) and estimated  $\hat{f}_F$  fast-relaxing compartmental fraction maps, in simulation. Maximum-likelihood estimation via variable projection method and grid search (center) is accurate but is computationally expensive. In contrast, kernel ridge regression (right) is very 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.

	Truth	VPM	KRR
WM $\hat{f}_F$	0.15	$0.1540 \pm 0.0298$	$0.1481 \pm 0.0237$
GM $\hat{f}_F$	0.03	$0.0333 \pm 0.0230$	$0.0398 \pm 0.0241$

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 [7].

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 (0.0238 versus 0.0300) and comparable to VPM in GM (0.0260 versus 0.0232).

## 5.5 Summary and Future Work

This chapter introduced 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 [103]).

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 6.5 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 [119], 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  $c_3$  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  $c_3$  is the only latent parameter that is affected by data scale, it is natural to seek problem structure in  $c_3$  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  $c_3$  will allow us to separate estimating scale-variant  $c_3$  (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:

$$\{\text{eq:krr,joint-recon}\} \quad \mathbf{Y}^{(t)} \leftarrow \arg \min_{\mathbf{Y}} \|\mathbf{D} - \mathbf{YA}\|_{\text{F}}^2 + \eta^{(t)} \|\mathbf{S}^{(t-1)} - \mathbf{Y}\|_{\text{F}}^2 \quad (5.28)$$

$$\{\text{eq:krr,joint-lift}\} \quad \tilde{\mathbf{Z}}_Z^{(t)} \leftarrow \left[ \tilde{\mathbf{z}}_Z(\mathbf{y}_1^{(t)}, \boldsymbol{\nu}_1), \dots, \tilde{\mathbf{z}}_Z(\mathbf{y}_V^{(t)}, \boldsymbol{\nu}_V) \right] \quad (5.29)$$

$$\{\text{eq:krr,joint-est}\} \quad \mathbf{X}^{(t)} \leftarrow \mathbf{m}_x \mathbf{1}_V^{\text{T}} + \mathbf{C}_{x\tilde{\mathbf{z}}} (\mathbf{C}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}} + \rho \mathbf{I}_N)^{-1} (\tilde{\mathbf{Z}}_Z^{(t)} - \mathbf{m}_{\tilde{\mathbf{z}}} \mathbf{1}_V^{\text{T}}) \quad (5.30)$$

$$\{\text{eq:krr,joint-model}\} \quad \mathbf{S}^{(t)} \leftarrow \left[ \mathbf{s}(\mathbf{x}_1^{(t)}, \boldsymbol{\nu}_1), \dots, \mathbf{s}(\mathbf{x}_V^{(t)}, \boldsymbol{\nu}_V) \right] \quad (5.31)$$

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 (5.28), feature mapping (5.29), parameter estimation via KRR (5.30), and signal modeling (5.31). 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 (5.28)-(5.31). First, it is desirable to reformulate (5.28)-(5.31) 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 (5.30).

## CHAPTER 6

# Fast Myelin Water Fraction Imaging via Scan Design and PERK

{c,mwf}

### 6.1 Introduction

{s,mwf,intro}

This chapter applies techniques (developed in earlier chapters) for QMRI acquisition design and parameter estimation to design new methods for imaging a specific MR biomarker of clinical interest. In particular, we study a biomarker for *myelin content* in the human brain.

Myelin is a lipid-rich material that forms an insulating sheath encasing neuronal axons predominantly in white matter (WM) regions of the human brain [130]. Demyelination (*i.e.*, myelin loss) is central to the development of several neurodegenerative disorders such as multiple sclerosis (MS) [131]. Non-invasive myelin quantification in WM is thus desirable for monitoring the onset and progression of neurodegenerative disease.

MR relaxation time constants (especially spin-spin time constant  $T_2$ ) depend on the macromolecular environment surrounding excited water molecules. In nervous tissue, these environments vary spatially on scales much smaller than the millimeter-scale resolutions used in typical MR imaging experiments. Thus, there is significant variation of relaxation times within a typical imaging voxel containing nervous tissue.

Many researchers have attempted to characterize tissue microstructure by estimating the distribution of MR relaxation time constants and associating certain ranges of time constants with particular “compartments” or “pools” of water molecules that exist in similar macromolecular environments. *In vitro* NMR studies of nervous animal tissue ascribed a fast-relaxing<sup>1</sup> water compartment with  $T_2 \sim 10\text{-}40\text{ms}$  initially to general protein and phospholipid structures [132] and later more specifically to water trapped between the phospholipid bilayers of myelin [133, 134]. Shortly thereafter, the first MR images of so-called

---

<sup>1</sup>The fast-relaxing compartment bears its name with reference to the portion of the  $T_2$  distribution in water (from about 10ms to at least 1000ms) that is typically observable in MRI.

*myelin water fraction* (MWF), defined as the proportion of MR signal arising from the fast-relaxing water compartment relative to total MR signal, were demonstrated *in vivo* in the human brain [4]. More recently, MWF has been shown to correlate well with histological measurements of myelin content in animal models of nerve injury [135] and demyelination [136]. In humans, MWF has also been measured to be markedly lower in “normally appearing” WM of MS patients versus controls [137], and to correlate strongly with post-mortem histological measurements of myelin content in MS patients [138]. Thus, there is reasonably strong evidence that MWF (as defined in [4]) is a specific and non-invasive biomarker for myelin content in WM.

All of the aforementioned studies estimate MWF images from a multi-echo spin echo (MESE) MRI pulse sequence [52] with long repetition time  $T_R \geq 2s$  to ensure sufficient recovery of the longitudinal magnetization in nervous tissue. Whole-brain MWF imaging using such long- $T_R$  MESE acquisitions at a typical imaging resolution would require hours of scan time and thus may not be clinically feasible. As a more practical alternative, scan profiles consisting of short- $T_R$  steady-state (SS) sequences were proposed for whole-brain MWF imaging in about 30m scan time [115]. Despite recent further refinements [2, 116], MWF images from SS pulse sequences have thus far been shown to be incomparable with MWF images from MESE pulse sequences [6], likely due in part to insufficient parameter estimation precision [5].

This chapter introduces a rapid SS MRI scan profile for precise MWF imaging. We apply QMRI acquisition design (developed in Chapter 4) to optimize the flip angles and repetition times of combinations of spoiled gradient-recalled echo (SPGR) [21] and dual-echo steady-state (DESS) [23, 24] sequences for precise MWF estimation. We rapidly estimate MWF and other nuisance latent object parameters using QMRI parameter estimation via kernel ridge regression (KRR) (developed in Chapter ??). We obtain proof-of-concept MWF maps *in vivo* that are comparable to results reported in MESE literature.

The remainder of this chapter is organized as follows. Section 6.2 reviews and develops simple two-compartment signal models for SPGR and DESS pulse sequences, respectively. Section 6.3 designs a new SPGR/DESS scan profile for precisely estimating MWF parameter images from SPGR/DESS image data. Section 6.4 applies the proposed acquisition *in vivo* and qualitatively compares resultant MWF images to state-of-the-art MESE results. Section 6.5 summarizes contributions thus far and discusses future work.

## 6.2 Multi-Compartmental Models for SS Sequences

{s,mwf,model}

This section develops multi-compartmental signal models for the SPGR and DESS pulse sequences. Subsection 6.2.1 reviews and extends a concise Bloch-matrix derivation [139] of an SPGR signal model that accounts for exchange between multiple compartments. Subsection 6.2.2 applies the Bloch-matrix representation to derive analogous (but previously unpublished) multi-compartmental DESS signal models. Though the derivations below focus for simplicity on only two exchanging compartments, the Bloch-matrix formulation generalizes readily to greater numbers of interacting compartments.

### 6.2.1 A Two-Compartment SPGR Signal Model

{ss,mwf,model,spgr}

The McConnell equations [106] extend the Bloch equations [17] to account for physical exchange<sup>2</sup> between two or more intra-voxel compartments. Here, we consider the interaction of a fast-relaxing water compartment (characterized by comparatively short spin-lattice  $T_{1,F}$  and spin-spin  $T_{2,F}$  relaxation times) with a slow-relaxing water compartment (characterized by longer relaxation times  $T_{1,S}, T_{2,S}$ ). In primed coordinates rotating clockwise about the longitudinal  $z$ -axis at the Larmor frequency, the dynamics of corresponding fast-relaxing and slow-relaxing compartmental magnetization vectors  $\mathbf{m}'_F := [m'_{x,F}, m'_{y,F}, m'_{z,F}]^T$  and  $\mathbf{m}'_S := [m'_{x,S}, m'_{y,S}, m'_{z,S}]^T$  are coupled via first-order exchange rates  $r_{F \rightarrow S}$  (from fast to slow compartment) and  $r_{S \rightarrow F}$  (vice-versa). During periods when no RF excitation is present, these magnetization dynamics decouple in transverse and longitudinal components; the two-compartment transverse equations extend (2.9) to read

$$\frac{\partial}{\partial t} m'_{xy,F}(\mathbf{r}, t) = -i\gamma m'_{xy,F}(\mathbf{r}, t) \omega'_F(\mathbf{r}) - \frac{m'_{xy,F}(\mathbf{r}, t)}{T_{2,F}(\mathbf{r})} - r_{F \rightarrow S}(\mathbf{r}) m'_{xy,F}(\mathbf{r}, t) + r_{S \rightarrow F}(\mathbf{r}) m'_{xy,S}(\mathbf{r}, t); \quad (6.1)$$

{eq:mwf,mxy-f}

$$\frac{\partial}{\partial t} m'_{xy,S}(\mathbf{r}, t) = -i\gamma m'_{xy,S}(\mathbf{r}, t) \omega'_S(\mathbf{r}) - \frac{m'_{xy,S}(\mathbf{r}, t)}{T_{2,S}(\mathbf{r})} - r_{S \rightarrow F}(\mathbf{r}) m'_{xy,S}(\mathbf{r}, t) + r_{F \rightarrow S}(\mathbf{r}) m'_{xy,F}(\mathbf{r}, t), \quad (6.2)$$

{eq:mwf,mxy-s}

---

<sup>2</sup>The word “exchange” is often a source of confusion because it is used in different contexts to refer to a variety of transport phenomena. As originally described in [106], *chemical exchange* specifically refers to the rapid reversible transfer of a nucleus between two or more molecular environments (*e.g.*, hydrogen bonding in water or proton exchange between water and a macromolecule). More recently, exchange has also been loosely used to describe several other processes that can be characterized with similar physical equations but where explicit nuclear transfer does not occur. One such process that we denote for clarity as *physical exchange* involves the transfer of intact molecules between different compartments (*e.g.*, water molecules across a membrane). In the ensuing derivations, we refer specifically to physical exchange but note that the McConnell equations describe chemical exchange as well.

where  $m'_{xy,F}(\mathbf{r}, t) := m'_{x,F}(\mathbf{r}, t) + im'_{y,F}(\mathbf{r}, t)$  and  $m'_{xy,S}(\mathbf{r}, t) := m'_{x,S}(\mathbf{r}, t) + im'_{y,S}(\mathbf{r}, t)$  are complex compartmental representations of the transverse magnetization at position  $\mathbf{r}$  and time  $t$ ;  $\omega'_F(\mathbf{r}) \in \mathbb{R}$  and  $\omega'_S(\mathbf{r}) \in \mathbb{R}$  allow for compartment-specific but time-invariant off-resonance effects; and  $\gamma \in \mathbb{R}$  is the gyromagnetic ratio. During periods when no excitation is present, analogous two-compartment longitudinal equations extend (2.10) to read

$$\frac{\partial}{\partial t} m'_{z,F}(\mathbf{r}, t) = -\frac{m'_{z,F}(\mathbf{r}, t) - f_F(\mathbf{r})m_0(\mathbf{r})}{T_{1,F}(\mathbf{r})} - r_{F \rightarrow S}(\mathbf{r})m'_{z,F}(\mathbf{r}, t) + r_{S \rightarrow F}(\mathbf{r})m'_{z,S}(\mathbf{r}, t); \quad (6.3)$$

$$\frac{\partial}{\partial t} m'_{z,S}(\mathbf{r}, t) = -\frac{m'_{z,S}(\mathbf{r}, t) - f_S(\mathbf{r})m_0(\mathbf{r})}{T_{1,S}(\mathbf{r})} - r_{S \rightarrow F}(\mathbf{r})m'_{z,S}(\mathbf{r}, t) + r_{F \rightarrow S}(\mathbf{r})m'_{z,F}(\mathbf{r}, t), \quad (6.4)$$

where  $f_F(\mathbf{r}) \in [0, 1]$  and  $f_S(\mathbf{r}) \in [0, 1]$  denote fast- and slow-relaxing compartmental fractions and  $m_0(\mathbf{r})$  denotes the total<sup>3</sup> equilibrium magnetization.

Equations (6.1)-(6.4) comprise a first-order affine dynamical system and can be equivalently and concisely written in matrix form as

$$\frac{\partial}{\partial t} \mathbf{m}'(\mathbf{r}, t) = \mathbf{A}(\mathbf{r})\mathbf{m}'(\mathbf{r}, t) + \mathbf{c}(\mathbf{r}), \quad (6.5)$$

where  $\mathbf{m}'(\mathbf{r}, t) := \text{vec}\left([m'_F(\mathbf{r}, t), m'_S(\mathbf{r}, t)]^T\right) \in \mathbb{R}^6$  collects the compartmental magnetization vectors and  $\text{vec}(\cdot)$  denotes vectorization. Here, system matrix  $\mathbf{A}(\mathbf{r}) \in \mathbb{R}^{6 \times 6}$  admits a block diagonal form<sup>4</sup>  $\mathbf{A}(\mathbf{r}) \equiv [\mathbf{A}_{xy}(\mathbf{r}), \mathbf{0}_{4 \times 2}; \mathbf{0}_{2 \times 4}, \mathbf{A}_z(\mathbf{r})]$ , with  $\mathbf{A}_{xy}(\mathbf{r}) :=$

$$\begin{bmatrix} -\frac{1}{T_{2,F}(\mathbf{r})} - r_{F \rightarrow S}(\mathbf{r}) & r_{S \rightarrow F}(\mathbf{r}) & \omega'_F(\mathbf{r}) & 0 \\ r_{F \rightarrow S}(\mathbf{r}) & -\frac{1}{T_{2,S}(\mathbf{r})} - r_{S \rightarrow F}(\mathbf{r}) & 0 & \omega'_S(\mathbf{r}) \\ -\omega'_F(\mathbf{r}) & 0 & -\frac{1}{T_{2,F}(\mathbf{r})} - r_{F \rightarrow S}(\mathbf{r}) & r_{S \rightarrow F}(\mathbf{r}) \\ 0 & -\omega'_S(\mathbf{r}) & r_{F \rightarrow S}(\mathbf{r}) & -\frac{1}{T_{2,S}(\mathbf{r})} - r_{S \rightarrow F}(\mathbf{r}) \end{bmatrix} \quad (6.6)$$

<sup>3</sup>The numerous parameters introduced here are often interdependent. In regions with only two water compartments present,  $f_F(\mathbf{r}) + f_S(\mathbf{r}) = 1$ . In regions where only two water pools exhibit exchange and that exchange is in chemical equilibrium,  $f_F(\mathbf{r})r_{F \rightarrow S}(\mathbf{r}) = f_S(\mathbf{r})r_{S \rightarrow F}(\mathbf{r})$ .

<sup>4</sup>For conciseness we introduce here the MATLAB<sup>®</sup>-like shorthand

$$[\mathbf{M}_{1,1}, \mathbf{M}_{1,2}; \mathbf{M}_{2,1}, \mathbf{M}_{2,2}] \equiv \begin{bmatrix} \mathbf{M}_{1,1} & \mathbf{M}_{1,2} \\ \mathbf{M}_{2,1} & \mathbf{M}_{2,2} \end{bmatrix}$$

for all appropriately sized submatrices  $\mathbf{M}_{1,1}, \mathbf{M}_{1,2}, \mathbf{M}_{2,1}, \mathbf{M}_{2,2}$ .

collecting transverse dynamics and

$$\mathbf{A}_z(\mathbf{r}) := \begin{bmatrix} -\frac{1}{T_{1,F}(\mathbf{r})} - r_{F \rightarrow S}(\mathbf{r}) & r_{S \rightarrow F}(\mathbf{r}) \\ r_{F \rightarrow S}(\mathbf{r}) & -\frac{1}{T_{1,S}(\mathbf{r})} - r_{S \rightarrow F}(\mathbf{r}) \end{bmatrix} \quad (6.7)$$

collecting longitudinal dynamics; and  $\mathbf{c}(\mathbf{r}) := \left[0, 0, 0, 0, \frac{f_F(\mathbf{r})m_0(\mathbf{r})}{T_{1,F}(\mathbf{r})}, \frac{f_S(\mathbf{r})m_0(\mathbf{r})}{T_{1,S}(\mathbf{r})}\right]^T$ . If  $\mathbf{A}(\mathbf{r})$  is invertible, a matrix exponential solution to (6.5) exists and reads

$$\mathbf{m}'(\mathbf{r}, t) = e^{(t-t_0)\mathbf{A}(\mathbf{r})}\mathbf{m}'(\mathbf{r}, t_0) + (e^{(t-t_0)\mathbf{A}(\mathbf{r})} - \mathbf{I}_6)(\mathbf{A}(\mathbf{r}))^{-1}\mathbf{c}(\mathbf{r}) \quad \forall t \geq t_0, \quad (6.8)$$

where  $\mathbf{m}'(\mathbf{r}, t_0)$  is the magnetization at an initial time  $t_0$  and  $\mathbf{I}_6 \in \mathbb{R}^{6 \times 6}$  is an identity matrix. In chemical equilibrium (described in Footnote 3), a direct calculation involving compartmental equilibrium magnetization  $\mathbf{m}_0(\mathbf{r}) := [0, 0, 0, 0, f_F(\mathbf{r})m_0(\mathbf{r}), f_S(\mathbf{r})m_0(\mathbf{r})]^T$  reveals that  $\mathbf{A}(\mathbf{r})\mathbf{m}_0(\mathbf{r}) = -\mathbf{c}(\mathbf{r})$ , in which case (6.8) simplifies to

$$\mathbf{m}'(\mathbf{r}, t) = e^{(t-t_0)\mathbf{A}(\mathbf{r})}\mathbf{m}'(\mathbf{r}, t_0) + (\mathbf{I}_6 - e^{(t-t_0)\mathbf{A}(\mathbf{r})})\mathbf{m}_0(\mathbf{r}) \quad \forall t \geq t_0, \quad (6.9)$$

a form that bears much resemblance to the matrix operator solution (2.14) to the single-compartment Bloch equations in the presence of only free precession and relaxation.

To develop a two-compartment SPGR signal model, we modify the derivation of the single-compartment SPGR model (presented in Subsection 2.1.2.1) to account for inter-compartmental exchange (via (6.9)) during intervals of free precession and relaxation. As before, 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 RF excitation. The SPGR sequence first applies an RF excitation with pulse duration  $T_P$ . If we neglect relaxation, off-resonance, and exchange during excitation (which is reasonable for sufficiently short  $T_P$ ) and further assume that each compartment experiences this excitation with equal transmit field sensitivity, we may model RF excitation as a rotation by a compartment-wise constant (but spatially varying) nutation angle  $\alpha(\mathbf{r}, t_0 + T_P; t_0)$  (defined in (2.7)), abbreviated  $\alpha(\mathbf{r})$  hereafter. For clockwise rotations about the  $x'$ -axis, such near-instantaneous rotation may be represented as

$$\mathbf{m}'(\mathbf{r}, t_0 + T_P) = \mathbf{R}_{x'}^{\otimes 2}(\alpha(\mathbf{r}))\mathbf{m}'(\mathbf{r}, t_0), \quad (6.10)$$

where  $\mathbf{R}_{x'}^{\otimes 2}(\alpha(\mathbf{r})) := \mathbf{R}_{x'}(\alpha(\mathbf{r})) \otimes \mathbf{I}_2 \in \mathbb{R}^{6 \times 6}$  is a compartment-wise rotation operation;  $\mathbf{R}_{x'}(\alpha(\mathbf{r}))$  is defined in (2.6); and  $\otimes$  denotes the Kronecker product.

The compartments exchange while their magnetization vectors precess and relax as per



(6.8) until data acquisition at echo time  $T_E \in [\frac{T_P}{2}, T_R]$  after the midpoint of RF excitation:

$$\mathbf{m}'\left(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E\right) = e^{(T_E - \frac{T_P}{2})\mathbf{A}(\mathbf{r})}\mathbf{m}'(\mathbf{r}, t_0 + T_P) + \left(\mathbf{I}_6 - e^{(T_E - \frac{T_P}{2})\mathbf{A}(\mathbf{r})}\right)\mathbf{m}_0(\mathbf{r}). \quad (6.11)$$

Following signal reception (which we assume as in Footnote 3 of Chapter 2 to negligibly influence magnetization), the SPGR sequence spoils the remaining transverse magnetization in both compartments while leaving unaffected the longitudinal magnetization components in each compartment. We model ideal spoiling in both compartments as

$$\mathbf{S}^{\otimes 2}\mathbf{m}'\left(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E\right), \quad (6.12)$$

where  $\mathbf{S}^{\otimes 2} := \mathbf{S} \otimes \mathbf{I}_2 \in \mathbb{R}^{6 \times 6}$  is a compartment-wise spoiling operator and  $\mathbf{S} \in \mathbb{R}^{3 \times 3}$  is the diagonal binary matrix defined in (2.20). After spoiling, compartments continue to exchange while their longitudinal magnetization components partially recover until  $t \leftarrow t_0 + T_R$ , completing one repetition cycle:

$$\begin{aligned} \mathbf{m}'(\mathbf{r}, t_0 + T_R) &= e^{(T_R - (\frac{T_P}{2} + T_E))\mathbf{A}(\mathbf{r})}\mathbf{S}^{\otimes 2}\mathbf{m}'\left(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E\right) \\ &+ \left(\mathbf{I}_6 - e^{(T_R - (\frac{T_P}{2} + T_E))\mathbf{A}(\mathbf{r})}\right)\mathbf{m}_0(\mathbf{r}). \end{aligned} \quad (6.13)$$

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

$$\mathbf{m}'(\mathbf{r}, t_0 + T_P) = \mathbf{R}_{x'}^{\otimes 2}(\alpha(\mathbf{r}))\mathbf{m}'(\mathbf{r}, t_0 + T_R), \quad (6.14)$$

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

$$\mathbf{m}'(\mathbf{r}, t_0 + T_P) = (\mathbf{I}_6 - \mathbf{R}_{x'}^{\otimes 2}(\alpha(\mathbf{r}))\mathbf{S}^{\otimes 2}e^{(T_R - T_P)\mathbf{A}(\mathbf{r})})^{-1}\mathbf{R}_{x'}^{\otimes 2}(\alpha(\mathbf{r})) \quad (6.15)$$

$$\begin{aligned} &\times \left(\mathbf{I}_6 - \mathbf{S}^{\otimes 2}e^{(T_R - T_P)\mathbf{A}(\mathbf{r})} - e^{(T_R - (\frac{T_P}{2} + T_E))\mathbf{A}(\mathbf{r})}(\mathbf{I}_6 - \mathbf{S}^{\otimes 2})\right)\mathbf{m}_0(\mathbf{r}) \\ &= (\mathbf{I}_6 - \mathbf{R}_{x'}^{\otimes 2}(\alpha(\mathbf{r}))\mathbf{S}^{\otimes 2}e^{(T_R - T_P)\mathbf{A}(\mathbf{r})})^{-1}\mathbf{R}_{x'}^{\otimes 2}(\alpha(\mathbf{r})) \quad (6.16) \\ &\times (\mathbf{I}_6 - \mathbf{S}^{\otimes 2}e^{(T_R - T_P)\mathbf{A}(\mathbf{r})})\mathbf{m}_0(\mathbf{r}), \end{aligned}$$

where (6.15) is due to straightforward matrix operations and (6.16) makes use of the ideal spoiling property  $(\mathbf{I}_6 - \mathbf{S}^{\otimes 2})\mathbf{m}_0(\mathbf{r}) = \mathbf{0}_6 \forall \mathbf{r}$ . Substituting (6.16) into (6.11) yields an ex-

pression for the compartmental magnetization at the echo time.

The matrix exponentials in (6.16) would be cumbersome to expand in general but fortunately always arise prepended by spoiling matrix  $\mathbf{S}^{\otimes 2}$ . The combined form reduces to

$$\mathbf{S}^{\otimes 2} e^{(T_R - T_P)\mathbf{A}(\mathbf{r})} \equiv [\mathbf{0}_{4 \times 4}, \mathbf{0}_{4 \times 2}; \mathbf{0}_{2 \times 4}, e^{(T_R - T_P)\mathbf{A}_z(\mathbf{r})}] \quad (6.17)$$

and admits an explicit representation for (6.16) that we find via standard solvers<sup>5</sup>. Note that if RF pulses are assumed instantaneous (*i.e.*,  $T_P \leftarrow 0$ ) the longitudinal components of (6.16) are equivalent to [139, Eq. 34]. Another instructive (and much simpler) special case neglects exchange (*i.e.*,  $r_{F \rightarrow S} \leftarrow 0$  and  $r_{S \rightarrow F} \leftarrow 0$ ), in which case (6.16) reduces to

$$\mathbf{m}'(\mathbf{r}, t_0 + T_P) = \begin{bmatrix} 0 \\ 0 \\ \frac{f_F(\mathbf{r}) \sin(\alpha(\mathbf{r})) (1 - e^{-(T_R - T_P)/T_{1,F}(\mathbf{r})})}{1 - e^{-(T_R - T_P)/T_{1,F}(\mathbf{r})} \cos(\alpha(\mathbf{r}))} \\ \frac{f_S(\mathbf{r}) \sin(\alpha(\mathbf{r})) (1 - e^{-(T_R - T_P)/T_{1,S}(\mathbf{r})})}{1 - e^{-(T_R - T_P)/T_{1,S}(\mathbf{r})} \cos(\alpha(\mathbf{r}))} \\ \frac{f_F(\mathbf{r}) \cos(\alpha(\mathbf{r})) (1 - e^{-(T_R - T_P)/T_{1,F}(\mathbf{r})})}{1 - e^{-(T_R - T_P)/T_{1,F}(\mathbf{r})} \cos(\alpha(\mathbf{r}))} \\ \frac{f_S(\mathbf{r}) \cos(\alpha(\mathbf{r})) (1 - e^{-(T_R - T_P)/T_{1,S}(\mathbf{r})})}{1 - e^{-(T_R - T_P)/T_{1,S}(\mathbf{r})} \cos(\alpha(\mathbf{r}))} \end{bmatrix} m_0(\mathbf{r}), \quad (6.18)$$

a two (non-exchanging) compartment extension to the SPGR steady-state solution (2.23).

The received signal is approximately proportional to the integrated transverse magnetization arising from both compartments. To derive expressions, we take further assumptions, the first two of which are two-compartment extensions of assumptions used in §2.1.2.1:

1. We assume that the signal is localized to a scale over which there is within-voxel variation of each compartment's off-resonance frequency, but minimal intra-voxel variation of other space-varying parameters  $m_0, f_F, f_S, T_{1,F}, T_{1,S}, T_{2,F}, T_{2,S}, \alpha, r_{F \rightarrow S}, r_{S \rightarrow F}$ . This assumption effectively fixes all compartmental properties other than  $\omega'_F$  and  $\omega'_S$  over the volume  $\mathbb{V}$  of a sufficiently small voxel.
2. We assume that the fast-relaxing and slow-relaxing compartments' off-resonance frequencies are independently distributed within a localized voxel with marginal distributions  $p_{\omega'_F} := \text{Cauchy}(\bar{\omega}'_F, R'_{2,F})$  and  $p_{\omega'_S} := \text{Cauchy}(\bar{\omega}'_S, R'_{2,S})$ , where  $\bar{\omega}'_F, \bar{\omega}'_S$  are

<sup>5</sup>Since we require many signal evaluations for acquisition design and parameter estimation, it is well worthwhile from a computational standpoint to symbolically solve algebraic systems of equations whenever possible. However, for complicated equation systems such as (6.16), the manipulations required are tedious and error-prone. In MATLAB<sup>®</sup>, we used the Symbolic Toolbox and the function `matlabFunction` to symbolically solve (6.16) and automatically generate memory-friendly function handles.

{item:spgr,freq}

median off-resonance frequencies and  $R'_{2,F}$ ,  $R'_{2,S}$  are broadening bandwidths.

3. We assume for short echo time  $T_E$  that minimal exchange occurs between excitation and signal reception. This strong assumption is only used to facilitate first-order expansion of the matrix exponential in (6.20), thereby separating broadening integrals by compartment. Without this assumption, (6.20) remains valid.

{item:spgr,exchg0}

With these assumptions, the noiseless two-compartment steady-state SPGR signal model for a voxel centered at position  $\mathbf{r}$  and with volume defined by  $\mathbb{V}(\mathbf{r})$  is (to within constants):

{eq:mwf,spgr,int}

$$s_S\left(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E\right) \propto \int_{\mathbb{V}(\mathbf{r})} [1, 1, i, i, 0, 0] \mathbf{m}'\left(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E\right) d^3 \mathbf{r} \quad (6.19)$$

$$\approx \left( \int_{\mathbb{R}^2} [1, 1, i, i, 0, 0] e^{\left(T_E - \frac{T_P}{2}\right) \mathbf{A}(\mathbf{r})} \mathbf{p}_{\omega'_F}(\omega'_F) \mathbf{p}_{\omega'_S}(\omega'_S) d\omega'_F d\omega'_S \right) \mathbf{m}'(\mathbf{r}, t_0 + T_P) \quad (6.20)$$

{eq:mwf,spgr,int-freq}

{eq:mwf,spgr,int-exchg0}

$$\approx \begin{bmatrix} e^{-\left(1/T_{2,F}(\mathbf{r}) + R'_{2,F}(\mathbf{r}) + i\bar{\omega}'_F(\mathbf{r})\right)\left(T_E - \frac{T_P}{2}\right)} \\ e^{-\left(1/T_{2,S}(\mathbf{r}) + R'_{2,S}(\mathbf{r}) + i\bar{\omega}'_S(\mathbf{r})\right)\left(T_E - \frac{T_P}{2}\right)} \\ ie^{-\left(1/T_{2,F}(\mathbf{r}) + R'_{2,F}(\mathbf{r}) + i\bar{\omega}'_F(\mathbf{r})\right)\left(T_E - \frac{T_P}{2}\right)} \\ ie^{-\left(1/T_{2,S}(\mathbf{r}) + R'_{2,S}(\mathbf{r}) + i\bar{\omega}'_S(\mathbf{r})\right)\left(T_E - \frac{T_P}{2}\right)} \\ 0 \\ 0 \end{bmatrix}^T \mathbf{m}'(\mathbf{r}, t_0 + T_P) \quad (6.21)$$

{eq:mwf,spgr,mod-exchg0}

$$= m'_{xy,F}(\mathbf{r}, t_0 + T_P) e^{-\left(1/T_{2,F}(\mathbf{r}) + R'_{2,F}(\mathbf{r}) + i\bar{\omega}'_F(\mathbf{r})\right)\left(T_E - \frac{T_P}{2}\right)} + m'_{xy,S}(\mathbf{r}, t_0 + T_P) e^{-\left(1/T_{2,S}(\mathbf{r}) + R'_{2,S}(\mathbf{r}) + i\bar{\omega}'_S(\mathbf{r})\right)\left(T_E - \frac{T_P}{2}\right)}. \quad (6.22)$$

Observe that (6.20) uses Assumption 1 to leave  $\mathbf{m}'(\mathbf{r}, t_0 + T_P)$  outside of broadening integrals defined by Assumption 2 and (6.21) uses Assumption 3 to expand the matrix exponential and evaluate broadening integrals component-wise. Eq. (6.22) naturally extends the single-compartment SPGR model (2.27) and shows clearly that compartmental signals simply add if exchange between excitation and signal reception is neglected.

Though the manipulations leading to (6.22) require strong assumptions, they serve to minimally demonstrate the nontrivial two-compartment SPGR signal dependence on off-resonance distributions. Unlike in the single-compartment case, the signal decay and dephasing terms due to off-resonance effects  $e^{-(R'_{2,F} + i\bar{\omega}'_F)T_E}$  and  $e^{-(R'_{2,S} + i\bar{\omega}'_S)T_E}$  may *not* be considered to simply modify spin density  $m_0$  (in fact, to first order they modify compartmental fractions  $f_F$  and  $f_S$  of usual interest). Since off-resonance distributions do often differ significantly in cerebral tissue [140, 141], accurate *in vivo* parameter estimation from

this model requires either joint estimation of compartmental distributions or acquisition with echo times much shorter than compartmental broadening timescale differences.

## 6.2.2 A Two-Compartment DESS Signal Model

To develop two-compartment DESS signal models, we first adapt the McConnell equation solutions of Subsection 6.2.1 to describe compartmental magnetization evolution in the presence of time-dependent field inhomogeneities (such as the unbalanced dephasing gradient used in DESS). We then modify corresponding derivations of single-compartment DESS models (presented in Subsection 2.1.2.2), to account for inter-compartmental exchange during intervals of free precession and relaxation.

As discussed in Subsection 2.1.2.2, the DESS pulse sequence interlaces fixed RF excitations with fixed dephasing gradients to produce two distinct signals per excitation. Because these dephasing gradients create time-dependence in compartmental off-resonance effects, DESS signal dynamics cannot in general be described by matrix exponential solution (6.9). Instead, the dynamics (during periods without excitation) arise as a solution to

$$\frac{\partial}{\partial t} \mathbf{m}'(\mathbf{r}, t) = \mathbf{A}(\mathbf{r}, t) \mathbf{m}'(\mathbf{r}, t) + \mathbf{c}(\mathbf{r}), \quad (6.23)$$

where  $\mathbf{A}(\mathbf{r}, t) \equiv [\mathbf{A}_{xy}(\mathbf{r}, t), \mathbf{0}_{4 \times 2}; \mathbf{0}_{2 \times 4}, \mathbf{A}_z(\mathbf{r})]$  contains now time-dependent compartmental off-resonance effects  $\omega'_F(\mathbf{r}, t), \omega'_S(\mathbf{r}, t)$  that appear only in submatrix  $\mathbf{A}_{xy}(\mathbf{r}, t) :=$

$$\begin{bmatrix} -\frac{1}{T_{2,F}(\mathbf{r})} - r_{F \rightarrow S}(\mathbf{r}) & r_{S \rightarrow F}(\mathbf{r}) & \omega'_F(\mathbf{r}, t) & 0 \\ r_{F \rightarrow S}(\mathbf{r}) & -\frac{1}{T_{2,S}(\mathbf{r})} - r_{S \rightarrow F}(\mathbf{r}) & 0 & \omega'_S(\mathbf{r}, t) \\ -\omega'_F(\mathbf{r}, t) & 0 & -\frac{1}{T_{2,F}(\mathbf{r})} - r_{F \rightarrow S}(\mathbf{r}) & r_{S \rightarrow F}(\mathbf{r}) \\ 0 & -\omega'_S(\mathbf{r}, t) & r_{F \rightarrow S}(\mathbf{r}) & -\frac{1}{T_{2,S}(\mathbf{r})} - r_{S \rightarrow F}(\mathbf{r}) \end{bmatrix}. \quad (6.24)$$

If  $\mathbf{A}(\mathbf{r}, t)$  is invertible, an exponential series solution to (6.23) exists and could be expressed exactly using the Magnus expansion [142]. Here, we utilize a first-order Magnus expansion

that in chemical equilibrium yields a simple approximate<sup>6</sup> solution for all  $t \geq t_0$ :

$$\mathbf{m}'(\mathbf{r}, t) \approx e^{\int_{t_0}^t \mathbf{A}(\mathbf{r}, t') dt'} \mathbf{m}'(\mathbf{r}, t_0) + \left( e^{\int_{t_0}^t \mathbf{A}(\mathbf{r}, t') dt'} - \mathbf{I}_6 \right) \mathbf{m}_0(\mathbf{r}). \quad (6.27)$$

Higher-order expansions would better capture compartmental magnetization interactions due to off-resonance effects and are not considered hereafter for simplicity.

We next use (6.27) to develop a two-compartment DESS signal model. As before, let  $\mathbf{m}'(\mathbf{r}, t_0)$  denote the two-compartment magnetization at an initial time  $t_0$  selected well into the steady-state and just prior to RF excitation. The DESS sequence first applies a fixed RF excitation, which we again assume is of sufficiently short duration  $T_P$  as to permit neglect of within-pulse relaxation, off-resonance, and exchange effects. Further assuming that each compartment experiences excitation with equal transmit field sensitivity, we model RF excitation as a simple rotation operation by angle  $\alpha(\mathbf{r})$ :

$$\mathbf{m}'(\mathbf{r}, t_0 + T_P) = \mathbf{R}_{x'}^{\otimes 2}(\alpha(\mathbf{r})) \mathbf{m}'(\mathbf{r}, t_0). \quad (6.28)$$

The transverse components of  $\mathbf{m}'(\mathbf{r}, t_0 + T_P)$  contribute to a first acquired signal; dephase (but do not spoil completely) due to gradient dephasing, and contribute again to a second (smaller, but nonzero) acquired signal. We assume that the dephasing gradient is of sufficiently small gradient area so as to contribute mainly to compartmental off-resonance phase accrual and negligibly to compartmental magnetization attenuation due to self-diffusion. Then, compartmental magnetization evolution during data acquisition and gradient dephasing and until repetition time  $T_R$  is reasonably described by (6.27):

$$\mathbf{m}'(\mathbf{r}, t_0 + T_R) \approx e^{\int_{t_0+T_P}^{T_R} \mathbf{A}(\mathbf{r}, t') dt'} \mathbf{m}'(\mathbf{r}, t_0 + T_P) + \left( e^{\int_{t_0+T_P}^{T_R} \mathbf{A}(\mathbf{r}, t') dt'} - \mathbf{I}_6 \right) \mathbf{m}_0(\mathbf{r}). \quad (6.29)$$

Conveniently, approximation (6.29) depends on compartmental off-resonance effects en-

---

<sup>6</sup>Approximation (6.27) is exact if commutation relation

$$\mathbf{A}(\mathbf{r}, t_1) \mathbf{A}(\mathbf{r}, t_2) = \mathbf{A}(\mathbf{r}, t_2) \mathbf{A}(\mathbf{r}, t_1)$$

holds for all  $t_1, t_2 \geq t_0$ , which in turn holds pointwise if and only if

$$r_{F \rightarrow S}(\mathbf{r})(\omega'_F(\mathbf{r}, t_1) - \omega'_S(\mathbf{r}, t_1) - (\omega'_F(\mathbf{r}, t_2) - \omega'_S(\mathbf{r}, t_2))) = 0; \quad (6.25)$$

$$r_{S \rightarrow F}(\mathbf{r})(\omega'_F(\mathbf{r}, t_1) - \omega'_S(\mathbf{r}, t_1) - (\omega'_F(\mathbf{r}, t_2) - \omega'_S(\mathbf{r}, t_2))) = 0. \quad (6.26)$$

Conditions (6.25)-(6.26) hold exactly in the special cases of a time-independent difference in compartmental off-resonance frequencies (*i.e.*,  $\omega'_F(\mathbf{r}, t_1) - \omega'_S(\mathbf{r}, t_1) = \omega'_F(\mathbf{r}, t_2) - \omega'_S(\mathbf{r}, t_2) \forall \mathbf{r}, \forall t_1, t_2 \geq t_0$ ) or no exchange (*i.e.*,  $r_{F \rightarrow S}(\mathbf{r}) \leftarrow 0$  and  $r_{S \rightarrow F}(\mathbf{r}) \leftarrow 0 \forall \mathbf{r}$ ).

tirely through compartmental phase functions  $\phi'_F(\mathbf{r}) := \int_{t_0+T_P}^{T_R} \omega'_F(\mathbf{r}, t') \, dt'$  and  $\phi'_S(\mathbf{r}) := \int_{t_0+T_P}^{T_R} \omega'_S(\mathbf{r}, t') \, dt'$ , which will later aid intra-voxel integration.

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

$$\{\text{eq:mwf,dess-ss}\} \quad \mathbf{m}'(\mathbf{r}, t_0) = \mathbf{m}'(\mathbf{r}, t_0 + T_R) \quad (6.30)$$

which yields an algebraic system of equations. The solution (if it exists) gives the steady-state compartmental magnetization just prior to RF excitation:

$$\{\text{eq:mwf,dess-m-t0}\} \quad \mathbf{m}'(\mathbf{r}, t_0) = \left( \mathbf{I}_6 - e^{\int_{t_0+T_P}^{T_R} \mathbf{A}(\mathbf{r}, t') \, dt'} \mathbf{R}_{x'}^{\otimes 2}(\alpha(\mathbf{r})) \right)^{-1} \left( e^{\int_{t_0+T_P}^{T_R} \mathbf{A}(\mathbf{r}, t') \, dt'} - \mathbf{I}_6 \right) \mathbf{m}_0(\mathbf{r}). \quad (6.31)$$

Substituting (6.31) into (6.28) would yield a similar expression for the steady-state compartmental magnetization immediately following RF excitation. Equivalently, one may substitute (6.30) into (6.28) before solving for  $\mathbf{m}'(\mathbf{r}, t_0 + T_P)$  directly, which gives

$$\{\text{eq:mwf,dess-m-tp}\} \quad \begin{aligned} \mathbf{m}'(\mathbf{r}, t_0 + T_P) &= \left( \mathbf{I}_6 - \mathbf{R}_{x'}^{\otimes 2}(\alpha(\mathbf{r})) e^{\int_{t_0+T_P}^{T_R} \mathbf{A}(\mathbf{r}, t') \, dt'} \right)^{-1} \\ &\times \mathbf{R}_{x'}^{\otimes 2}(\alpha(\mathbf{r})) \left( e^{\int_{t_0+T_P}^{T_R} \mathbf{A}(\mathbf{r}, t') \, dt'} - \mathbf{I}_6 \right) \mathbf{m}_0(\mathbf{r}). \end{aligned} \quad (6.32)$$

Unlike the SPGR two-compartment magnetization (6.16), analogous DESS expressions (6.31)-(6.32) depend strongly on transverse submatrix  $\int_{t_0}^{T_P} \mathbf{A}_{xy}(\mathbf{r}, t') \, dt'$ . Consequently, it remains challenging (even with computer solvers) to expand corresponding matrix exponentials and thereby find explicit representations of (6.31)-(6.32) in general.

Frequently, the DESS signals are acquired at symmetric echo times  $T_E$  before and after the center of each RF pulse. Since no gradient dephasing occurs during intervals between refocusing and defocusing echoes that contain excitations, we can reasonably assume linear off-resonance phase accrual over these intervals and therefore use (6.9) (instead of (6.27)) to evolve the magnetization accordingly. Substituting (6.32) into (6.9) gives the magnetization at the data acquisition time after RF excitation:

$$\{\text{eq:mwf,dess-m-te-def}\} \quad \mathbf{m}'\left(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E\right) = e^{(T_E - \frac{T_P}{2})\mathbf{A}(\mathbf{r})} \mathbf{m}'(\mathbf{r}, t_0 + T_P) + \left( \mathbf{I}_6 - e^{(T_E - \frac{T_P}{2})\mathbf{A}(\mathbf{r})} \right) \mathbf{m}_0(\mathbf{r}). \quad (6.33)$$

To compute the magnetization at the acquisition time before excitation, we consider the free

precession, relaxation, and exchange that occurs between signal reception and excitation:

$$\mathbf{m}'(\mathbf{r}, t_0) = e^{(T_E - \frac{T_P}{2})\mathbf{A}(\mathbf{r})} \mathbf{m}'\left(\mathbf{r}, t_0 + \frac{T_P}{2} - T_E\right) + \left(\mathbf{I}_6 - e^{(T_E - \frac{T_P}{2})\mathbf{A}(\mathbf{r})}\right) \mathbf{m}_0(\mathbf{r}). \quad (6.34)$$

Inserting (6.31) into (6.34) and rearranging gives an expression for  $\mathbf{m}'(\mathbf{r}, t_0 + \frac{T_P}{2} - T_E)$ .

The received signal is approximately proportional to the integrated transverse magnetization arising from both compartments. To derive expressions, we take assumptions very similar to those used in Subsection 6.2.1 and append additional distributional assumptions on the compartmental phase accrual functions  $\phi'_F(\mathbf{r})$  and  $\phi'_S(\mathbf{r})$ :

1. We assume that the signal is localized to a scale over which there is within-voxel variation of compartmental off-resonance effects, but minimal intra-voxel variation of other space-varying parameters  $m_0, f_F, f_S, T_{1,F}, T_{1,S}, T_{2,F}, T_{2,S}, \alpha, r_{F \rightarrow S}, r_{S \rightarrow F}$ . This assumption effectively fixes all compartmental properties other than  $\phi'_F, \phi'_S, \omega'_F$ , and  $\omega'_S$  over the volume  $\mathbb{V}$  of a sufficiently small voxel.
2. We assume that the dephasing gradient imparts a sufficiently large integral number  $n_{\text{cyc}}$  of intra-voxel phase cycles such that full-repetition compartmental phase accruals  $\phi'_F$  and  $\phi'_S$  are distributed essentially uniformly as  $\mathbf{p}_{\phi'_F} \leftarrow \text{unif}(0, 2\pi n_{\text{cyc}})$  and  $\mathbf{p}_{\phi'_S} \leftarrow \text{unif}(0, 2\pi n_{\text{cyc}})$ , where  $n_{\text{cyc}} \in \{1, 2, 3, \dots\}$ .
3. We again assume that compartmental off-resonance phase accrues linearly between each excitation and its adjacent data acquisition periods, and that off-resonance frequencies are independently distributed within a localized voxel with marginal distributions  $\mathbf{p}_{\omega'_F} := \text{Cauchy}(\bar{\omega}'_F, R'_{2,F})$  and  $\mathbf{p}_{\omega'_S} := \text{Cauchy}(\bar{\omega}'_S, R'_{2,S})$ , where  $\bar{\omega}'_F, \bar{\omega}'_S$  are median off-resonance frequencies and  $R'_{2,F}, R'_{2,S}$  are broadening bandwidths.
4. We assume for short echo times  $T_E$  that negligible exchange occurs between each excitation and its adjacent data acquisition periods. This assumption facilitates expansion of the matrix exponentials explicitly visible in (6.33)-(6.34) and separates off-resonance frequency broadening integrals by compartment.

With these assumptions, the noiseless two-compartment steady-state DESS signal models

for a voxel centered at position  $\mathbf{r}$  and with volume defined by  $\mathbb{V}(\mathbf{r})$  are (to within constants):

$$\{eq:mwf, dess-def, int\} \quad s_D\left(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E\right) \propto \int_{\mathbb{V}(\mathbf{r})} [1, 1, i, i, 0, 0] \mathbf{m}'\left(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E\right) d^3 \mathbf{r} \quad (6.35)$$

$$\{eq:mwf, dess-def, int-freq\} \quad \approx \left( \int_{\mathbb{R}^2} [1, 1, i, i, 0, 0] e^{\left(T_E - \frac{T_P}{2}\right) \mathbf{A}(\mathbf{r})} \mathbf{p}_{\omega'_F}(\omega'_F) \mathbf{p}_{\omega'_S}(\omega'_S) d\omega'_F d\omega'_S \right) \times \int_{\mathbb{R}^2} \mathbf{m}'(\mathbf{r}, t_0 + T_P) \mathbf{p}_{\phi'_F}(\phi'_F) \mathbf{p}_{\phi'_S}(\phi'_S) d\phi'_S d\phi'_F \quad (6.36)$$

$$\{eq:mwf, dess-def, int-exchg0\} \quad = \begin{bmatrix} e^{-\left(1/T_{2,F}(\mathbf{r}) + R'_{2,F}(\mathbf{r}) + i\bar{\omega}'_F(\mathbf{r})\right)\left(T_E - \frac{T_P}{2}\right)} \\ e^{-\left(1/T_{2,S}(\mathbf{r}) + R'_{2,S}(\mathbf{r}) + i\bar{\omega}'_S(\mathbf{r})\right)\left(T_E - \frac{T_P}{2}\right)} \\ ie^{-\left(1/T_{2,F}(\mathbf{r}) + R'_{2,F}(\mathbf{r}) + i\bar{\omega}'_F(\mathbf{r})\right)\left(T_E - \frac{T_P}{2}\right)} \\ ie^{-\left(1/T_{2,S}(\mathbf{r}) + R'_{2,S}(\mathbf{r}) + i\bar{\omega}'_S(\mathbf{r})\right)\left(T_E - \frac{T_P}{2}\right)} \\ 0 \\ 0 \end{bmatrix}^T \times \int_{\mathbb{R}^2} \mathbf{m}'(\mathbf{r}, t_0 + T_P) \mathbf{p}_{\phi'_F}(\phi'_F) \mathbf{p}_{\phi'_S}(\phi'_S) d\phi'_S d\phi'_F; \quad (6.37)$$

$$\{eq:mwf, dess-ref, int\} \quad s_D\left(\mathbf{r}, t_0 + \frac{T_P}{2} - T_E\right) \propto \int_{\mathbb{V}(\mathbf{r})} [1, 1, i, i, 0, 0] \mathbf{m}'\left(\mathbf{r}, t_0 + \frac{T_P}{2} - T_E\right) d^3 \mathbf{r} \quad (6.38)$$

$$\{eq:mwf, dess-ref, int-freq\} \quad \approx \left( \int_{\mathbb{R}^2} [1, 1, i, i, 0, 0] e^{-\left(T_E - \frac{T_P}{2}\right) \mathbf{A}(\mathbf{r})} \mathbf{p}_{\omega'_F}(\omega'_F) \mathbf{p}_{\omega'_S}(\omega'_S) d\omega'_F d\omega'_S \right) \times \int_{\mathbb{R}^2} \mathbf{m}'(\mathbf{r}, t_0) \mathbf{p}_{\phi'_F}(\phi'_F) \mathbf{p}_{\phi'_S}(\phi'_S) d\phi'_S d\phi'_F \quad (6.39)$$

$$\{eq:mwf, dess-ref, int-exchg0\} \quad = \begin{bmatrix} e^{+\left(1/T_{2,F}(\mathbf{r}) - R'_{2,F}(\mathbf{r}) + i\bar{\omega}'_F(\mathbf{r})\right)\left(T_E - \frac{T_P}{2}\right)} \\ e^{+\left(1/T_{2,S}(\mathbf{r}) - R'_{2,S}(\mathbf{r}) + i\bar{\omega}'_S(\mathbf{r})\right)\left(T_E - \frac{T_P}{2}\right)} \\ ie^{+\left(1/T_{2,F}(\mathbf{r}) - R'_{2,F}(\mathbf{r}) + i\bar{\omega}'_F(\mathbf{r})\right)\left(T_E - \frac{T_P}{2}\right)} \\ ie^{+\left(1/T_{2,S}(\mathbf{r}) - R'_{2,S}(\mathbf{r}) + i\bar{\omega}'_S(\mathbf{r})\right)\left(T_E - \frac{T_P}{2}\right)} \\ 0 \\ 0 \end{bmatrix}^T \times \int_{\mathbb{R}^2} \mathbf{m}'(\mathbf{r}, t_0) \mathbf{p}_{\phi'_F}(\phi'_F) \mathbf{p}_{\phi'_S}(\phi'_S) d\phi'_S d\phi'_F. \quad (6.40)$$

Observe that (6.36) and (6.39) use Assumption 1 to separate the phase and frequency broadening integrals respectively defined in Assumptions 2 and 3. Similar to SPGR calculations in Subsection 6.2.1, (6.37) and (6.40) use Assumption 4 to evaluate the frequency broadening integral compartment-wise. However, the phase broadening integrals in (6.37) and (6.40) do not separate. Thus, even with the simple first-order Magnus expansion taken in



(6.27), we find in DESS that exchange causes magnetization compartments to interact not only via exchange rates but also through differences across compartments in off-resonance phase accrual. In other words, additional assumptions on the relationship between compartmental off-resonance distributions (*e.g.*, that they are the same) would affect the DESS models *even if* the signal were received immediately following excitation (at  $T_E \leftarrow \frac{T_P}{2}$ ).

In the special case where exchange is altogether neglected (which is a stronger assumption than Assumption 4, especially for longer  $T_R$ ), the phase broadening integrals in (6.37) and (6.40) separate across voxels and admit the closed-form expressions

$$\text{f,dess-def,mod-exchg0} \quad s_D\left(\mathbf{r}, t_0 + \frac{T_P}{2} + T_E\right) \propto +im_0(\mathbf{r}) \tan \frac{\alpha(\mathbf{r})}{2} \quad (6.41)$$

$$\times \left( f_F(\mathbf{r}) \left( 1 - \frac{\eta_F(\mathbf{r}, T_R - T_P)}{\xi_F(\mathbf{r}, T_R - T_P)} \right) e^{-(1/T_{2,F}(\mathbf{r}) + R'_{2,F}(\mathbf{r}) + i\bar{\omega}'_F(\mathbf{r}))(T_E - \frac{T_P}{2})} \right. \\ \left. + f_S(\mathbf{r}) \left( 1 - \frac{\eta_S(\mathbf{r}, T_R - T_P)}{\xi_S(\mathbf{r}, T_R - T_P)} \right) e^{-(1/T_{2,S}(\mathbf{r}) + R'_{2,S}(\mathbf{r}) + i\bar{\omega}'_S(\mathbf{r}))(T_E - \frac{T_P}{2})} \right);$$

$$\text{f,dess-ref,mod-exchg0} \quad s_D\left(\mathbf{r}, t_0 + \frac{T_P}{2} - T_E\right) \propto -im_0(\mathbf{r}) \tan \frac{\alpha(\mathbf{r})}{2} \quad (6.42)$$

$$\times \left( f_F(\mathbf{r}) (1 - \eta_F(\mathbf{r}, T_R - T_P)) e^{+(1/T_{2,F}(\mathbf{r}) - R'_{2,F}(\mathbf{r}) + i\bar{\omega}'_F(\mathbf{r}))(T_E - \frac{T_P}{2})} \right. \\ \left. + f_S(\mathbf{r}) (1 - \eta_S(\mathbf{r}, T_R - T_P)) e^{+(1/T_{2,S}(\mathbf{r}) - R'_{2,S}(\mathbf{r}) + i\bar{\omega}'_S(\mathbf{r}))(T_E - \frac{T_P}{2})} \right),$$

where  $\eta_F$ ,  $\eta_S$ ,  $\xi_F$ , and  $\xi_S$  are intermediate variables defined as

$$\eta_F(\mathbf{r}, t) := \sqrt{\frac{1 - (\exp(-t/T_{2,F}(\mathbf{r})))^2}{1 - (\exp(-t/T_{2,F}(\mathbf{r}))/\xi_F(\mathbf{r}, t))^2}}; \\ \xi_F(\mathbf{r}, t) := \frac{1 - \exp(-t/T_{1,F}(\mathbf{r})) \cos \alpha(\mathbf{r})}{\exp(-t/T_{1,F}(\mathbf{r})) - \cos \alpha(\mathbf{r})}; \\ \eta_S(\mathbf{r}, t) := \sqrt{\frac{1 - (\exp(-t/T_{2,S}(\mathbf{r})))^2}{1 - (\exp(-t/T_{2,S}(\mathbf{r}))/\xi_S(\mathbf{r}, t))^2}}; \\ \xi_S(\mathbf{r}, t) := \frac{1 - \exp(-t/T_{1,S}(\mathbf{r})) \cos \alpha(\mathbf{r})}{\exp(-t/T_{1,S}(\mathbf{r})) - \cos \alpha(\mathbf{r})}.$$

Eqs. (6.41)-(6.42) naturally extend single-compartment models (2.41) and (2.43), and elucidate the intuitive result that compartmental signals simply add if exchange is neglected.

## 6.3 A Fast Acquisition for Precise MWF Estimation

{s,mwf,acq}

This section develops a new scan profile consisting of fast steady-state pulse sequences for precise MWF estimation. Subsection 6.3.1 first motivates the need for more scalable scan design in MWF imaging and then adapts our method for QMRI scan design (introduced for a simpler problem in Chapter 4) appropriately. Subsection 6.3.2 details how we use scalable scan design (with the two-compartment signal models developed in Section 6.2) to design a fast SPGR/DESS acquisition for precise MWF imaging.

### 6.3.1 Scalable Acquisition Design

{ss,mwf,acq,design}

Recall from Subsection 4.2.1 that the inverse of the Fisher information  $\mathbf{F}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) \in \mathbb{C}^{L \times L}$  lower-bounds the covariance of unbiased estimates of  $L$  latent object parameters  $\mathbf{x} \in \mathbb{C}^L$ , given  $K$  known object parameters  $\boldsymbol{\nu} \in \mathbb{C}^K$  and  $A$  tunable acquisition parameters for each of  $D$  datasets  $\mathbf{P} \in \mathbb{R}^{A \times D}$ . As before, we continue to focus on minimizing a weighted average of the latent parameter variances and thus study the objective function

{eq:mwf,cost}

$$\Psi(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) := \text{tr}(\mathbf{W}\mathbf{F}^{-1}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P})\mathbf{W}^T), \quad (6.43)$$

where  $\mathbf{W}$  is a diagonal weighting matrix and  $\text{tr}(\cdot)$  denotes the matrix trace operation. For scan design, we seek to minimize  $\Psi$  with respect to acquisition parameters  $\mathbf{P}$ .

In Subsection 4.2.2, we addressed the dependance of  $\Psi$  on space-varying object parameters  $\mathbf{x}$  and  $\boldsymbol{\nu}$  through a min-max optimization problem. The associated “worst-case” design criterion requires only weak assumptions on object parameter distributions but is non-differentiable in  $\mathbf{P}$ . For the relatively simple application described in Section 4.3, the min-max criterion was studied through exhaustive search and so non-differentiability did not matter. However, MWF imaging requires estimation of several more latent parameters and thus necessitates scan parameter selection for a greater number of datasets and thus over a larger search space. Since exhaustive search over this higher-dimensional search space is prohibitively expensive computationally, we study here an alternate design criterion that is differentiable in  $\mathbf{P}$  and is thus amenable to gradient-based local optimization. Specifically, we seek an acquisition parameter  $\check{\mathbf{P}}$  that minimizes the *expected* weighted average of latent parameter variances over a search space  $\mathbb{P}$ :

{eq:mwf,P-hat}

$$\check{\mathbf{P}} \in \left\{ \arg \min_{\mathbf{P} \in \mathbb{P}} \bar{\Psi}(\mathbf{P}) \right\}, \text{ where} \quad (6.44)$$

{eq:mwf,expcost}

$$\bar{\Psi}(\mathbf{P}) := \mathbb{E}_{\mathbf{x}, \boldsymbol{\nu}}(\Psi(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P})) \quad (6.45)$$

and  $\mathbb{E}_{\mathbf{x},\boldsymbol{\nu}}(\cdot)$  denotes joint expectation with respect to prior joint distribution  $\mathbf{p}_{\mathbf{x},\boldsymbol{\nu}}$  on  $\mathbf{x}, \boldsymbol{\nu}$ .

Unlike min-max cost (4.7), expected cost (6.45) is often differentiable in  $\mathbf{P}$ . We next construct the gradient matrix  $\nabla_{\mathbf{P}}\bar{\Psi}(\mathbf{P}) \in \mathbb{R}^{A \times D}$  and provide sufficient conditions for when this gradient matrix exists. Our simple strategy involves first constructing  $\nabla_{\mathbf{P}}\Psi(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P})$  element-wise for fixed  $\mathbf{x}, \boldsymbol{\nu}$  and then relating  $\nabla_{\mathbf{P}}\bar{\Psi}(\mathbf{P})$  to  $\nabla_{\mathbf{P}}\Psi(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P})$ . Let  $\frac{\partial}{\partial p_{a,d}}$  be the  $(a, d)$ th element of matrix operator  $\nabla_{\mathbf{P}}$ . By standard matrix derivative identities, we have

$$\begin{aligned} \frac{\partial}{\partial p_{a,d}}\Psi(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) &= \frac{\partial}{\partial p_{a,d}} \text{tr}(\mathbf{W}\mathbf{F}^{-1}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P})\mathbf{W}^{\top}) \\ &= -\text{tr}\left(\mathbf{W}\mathbf{F}^{-1}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P})\frac{\partial}{\partial p_{a,d}}(\mathbf{F}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}))\mathbf{F}^{-1}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P})\mathbf{W}^{\top}\right). \end{aligned} \quad (6.46)$$

For image data corrupted by additive complex Gaussian noise with zero mean and covariance  $\boldsymbol{\Sigma}$ , the Fisher information (repeated from (4.3) for clarity) is given by

$$\mathbf{F}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) = (\nabla_{\mathbf{x}}\mathbf{s}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}))^{\text{H}}\boldsymbol{\Sigma}^{-1}\nabla_{\mathbf{x}}\mathbf{s}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}), \quad (6.47)$$

where  $\mathbf{s} := [s_1, \dots, s_D]^{\top}$  is the noiseless signal model. Furthermore, if elements of each measurement vector are assumed independent (as is typical),  $\boldsymbol{\Sigma}$  takes the diagonal structure  $\boldsymbol{\Sigma} \leftarrow \text{diag}(\sigma_1^2, \dots, \sigma_D^2)$  and

$$\begin{aligned} \frac{\partial}{\partial p_{a,d}}(\mathbf{F}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P})) &= \frac{\partial}{\partial p_{a,d}} \sum_{d'=1}^D \frac{1}{\sigma_{d'}^2} (\nabla_{\mathbf{x}} s_{d'}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{p}_{d'}))^{\text{H}} \nabla_{\mathbf{x}} s_{d'}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{p}_{d'}) \\ &= \frac{1}{\sigma_d^2} \frac{\partial}{\partial p_{a,d}} \left( (\nabla_{\mathbf{x}} s_d(\mathbf{x}; \boldsymbol{\nu}, \mathbf{p}_d))^{\text{H}} \nabla_{\mathbf{x}} s_d(\mathbf{x}; \boldsymbol{\nu}, \mathbf{p}_d) \right). \end{aligned} \quad (6.48)$$

Substituting (6.47)-(6.48) into (6.46) gives expressions in terms of signal model derivatives for each element of  $\nabla_{\mathbf{P}}\Psi(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P})$ . These expressions are well-defined if  $\mathbf{F}$  is invertible and if mixed partial derivative matrices  $\nabla_{\mathbf{p}_1}(\nabla_{\mathbf{x}} s_1)^{\top}, \dots, \nabla_{\mathbf{p}_D}(\nabla_{\mathbf{x}} s_D)^{\top} \in \mathbb{C}^{L \times A}$  exist and are continuous in  $\mathbf{x}, \mathbf{P}$ . Further assuming that  $\nabla_{\mathbf{P}}\Psi(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P})$  remains bounded for all  $\mathbf{x}, \boldsymbol{\nu}$ ,

$$\begin{aligned} \nabla_{\mathbf{P}}\bar{\Psi}(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P}) &= \nabla_{\mathbf{P}}\mathbb{E}_{\mathbf{x},\boldsymbol{\nu}}(\Psi(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P})) \\ &= \mathbb{E}_{\mathbf{x},\boldsymbol{\nu}}(\nabla_{\mathbf{P}}\Psi(\mathbf{x}; \boldsymbol{\nu}, \mathbf{P})), \end{aligned} \quad (6.49)$$

which provides an expression for the gradient of the expected cost, as desired.

If (6.49) exists, one could solve (6.44) iteratively for convex search space  $\mathbb{P}$  via updates

$$\mathbf{P}^{(i)} \leftarrow \mathbf{P}_{\mathbb{P}}(\mathbf{P}^{(i-1)} - \nabla_{\mathbf{P}}\bar{\Psi}(\mathbf{P}^{(i-1)})), \quad (6.50)$$

where  $P_{\mathbb{P}}(\cdot)$  denotes projection onto  $\mathbb{P}$  and  $i$  indexes iteration. Since  $\bar{\Psi}$  is non-convex in  $\mathbf{P}$  in general, such iterations achieve only locally optimal convergence in cost (further discussed in Subsection 2.2.1) and the locally convergent minimizer depends on initialization.

### 6.3.2 SPGR/DESS Scan Design Implementation Details

{ss,mwf,acq,detail}

This subsection applies scalable scan design problem (6.44) to optimize fast scan profiles consisting of SPGR and DESS scans for precise MWF estimation. Intuitively, we study SPGR/DESS scan profiles over previously studied steady-state scan combinations [115, 2, 116] because SPGR/DESS signals are relatively insensitive to off-resonance related non-idealities, which as demonstrated in Section 6.2 can be difficult to model accurately in multi-compartmental systems and in fact here (as in [2, 116]) are largely neglected.

For the preliminary study discussed in the remainder of this chapter, we specifically assume (as in Section 5.4) that signal arises from two non-exchanging water compartments (*i.e.*,  $r_{F \rightarrow S} \leftarrow 0$ ,  $r_{S \rightarrow F} \leftarrow 0$ , and  $f_F + f_S = 1$ ) with identical broadening distributions (*i.e.*,  $R'_{2,F} \equiv R'_{2,S}$  and  $\bar{\omega}'_F \equiv \bar{\omega}'_S$ ). These simplifications provide closed-form expressions for the SPGR (6.22) and DESS (6.41)-(6.42) signal models as well as their gradients. To reduce SPGR/DESS signal dependence on off-resonance effects, we use magnitude signal models for scan design and highlight that Rician distributed noise in corresponding magnitude image data is well-approximated as Gaussian for sufficiently large SNR [32]. We further fix  $T_P$ ,  $T_E$  across scans and thereby reduce model dependencies to seven free object parameters per voxel:  $f_F$ ,  $T_{1,F}$ ,  $T_{2,F}$ ,  $T_{1,S}$ ,  $T_{2,S}$ ,  $\kappa^t$ , and  $c_3 := m_0 e^{-R'_{2,F} T_E} \equiv m_0 e^{-R'_{2,S} T_E}$ ; and two acquisition parameters per dataset:  $\mathbf{p}_d \leftarrow [\alpha_0, T_R]^T, \forall d \in \{1, \dots, D\}$ . As in Chapter 4, we assume prior knowledge of transmit field sensitivity  $\boldsymbol{\nu} \leftarrow \kappa^t$  (which in practice can be estimated from separate fast acquisitions, *e.g.* [92]) and collect the remaining  $L \leftarrow 6$  latent parameters as  $\mathbf{x} \leftarrow [f_F, T_{1,F}, T_{2,F}, T_{1,S}, T_{2,S}, c_3]$ .

We take fast-relaxing compartmental fraction  $f_F$  to be a proxy for MWF. To design a MWF imaging acquisition, we therefore tailor our design criterion to encourage precise  $f_F$  estimation. We consider the other five latent parameters to be nuisance parameters and accordingly set weighting matrix  $\mathbf{W} \leftarrow \text{diag}((E_{\mathbf{x},\boldsymbol{\nu}}(f_F))^{-1}, \mathbf{0}_5)$ , thereby ignoring nuisance parameter estimation precision. Here, fast-fraction variance weight  $(E_{\mathbf{x},\boldsymbol{\nu}}(f_F))^{-1}$  assigns interpretable meaning to  $\sqrt{\bar{\Psi}(\mathbf{P})}$  (equivalently,  $\bar{\Psi}(\mathbf{P})$ ) as a unitless measure of the expected relative standard deviation (equivalently, expected coefficient of variation) afforded by  $\mathbf{P}$  in asymptotically unbiased estimates of  $f_F$ .

We approximate expectations of form  $E_{\mathbf{x},\boldsymbol{\nu}}(\cdot)$  by taking empirical averages using samples of  $\mathbf{x}, \boldsymbol{\nu}$  drawn from the prior distribution defined as follows. We assume for simplicity

that all components of  $\mathbf{x}, \boldsymbol{\nu}$  are independent. We model  $f_F \sim \text{unif}(0.03, 0.21)$  to conservatively contain state-of-the-art MESE MWF measurements in WM [6]. We assume  $T_{1,F}$ ,  $T_{2,F}$ ,  $T_{1,S}$ , and  $T_{2,S}$  are Gaussian distributed, with means 400ms, 20ms, 1000ms, and 80ms selected from literature measurements [4, 2] and standard deviations that are 20% of corresponding means. Since  $\mathbf{W}$  places zero weight on estimating  $c_3$ , it suffices to fix  $c_3 \leftarrow 1$  and to assign noise variance  $\boldsymbol{\Sigma} \leftarrow (1.49 \times 10^{-7})\mathbf{I}_{10}$  based on separate measurements in unit-normalized image data (*cf.* Section 4.4.2.1 for acquisition details). Lastly, we model  $\kappa^t \sim \text{unif}(0.9, 1.1)$  to account for 10% flip angle spatial variation.

We constrain our search space  $\mathbb{P}$  to reflect hardware, safety, and model-accuracy limitations and to avoid undesirably long acquisitions. To control RF energy deposition, we restrict DESS flip angles to range between  $1^\circ$  and  $60^\circ$ . We further restrict SPGR flip angles to be between  $1^\circ$  and  $40^\circ$  to avoid excessive model mismatch due to partial spoiling effects [21]. To comply with other fixed pulse sequence timing requirements, we require DESS and SPGR repetition times to be no less than 17.5ms and 11.8ms respectively. We of course insist each pair of DESS defocusing- and refocusing-echo datasets be assigned the same flip angle and repetition time. Lastly, we impose a total scan time constraint  $\sum_{d=1}^D T_{R,d} \leq 263.7\text{ms}$  that permits equal data collection at minimal  $T_R$  (given our acquisition timing restrictions) as in a state-of-the-art fast steady-state acquisition [2]. These constraints together define a convex search space over which  $\bar{\Psi}$  can be locally optimized.

To this point, we have only considered scan design via acquisition parameter optimization for a fixed combination of SPGR and DESS scans. In single-component  $T_1, T_2$  relaxometry (studied in Section 4.3), we were able to restrict our total scan time constraint such that only three such scan profiles were feasible, each of which we studied explicitly. However in MWF imaging, more datasets are required for well-conditioned estimation of a larger number of latent parameters. To permit scan profiles with many datasets, we must optimize with respect to a larger time constraint that allows combinatorially many more scan profiles. As an alternative to exhaustively comparing all scan profiles (that too, only via local optima), we employ a greedy method for scan profile construction, diagrammed in Figure 6.1. We first optimize an initial scan profile consisting of only 3 SPGR and 3 DESS scans. We then seek to improve the acquisition by appending another SPGR or DESS scan to the previously optimized scan profile and repeating constrained optimization. If we find that at least one appended scan profile reduces the objective while meeting all constraints, we update the previous scan profile with the appended scan profile that better improved the objective and repeat the process. Otherwise, we return the previously optimized scan profile. Observe that while this approach is far from optimal, we can easily assess solution quality via expected  $f_F$  relative standard deviation  $\sqrt{\bar{\Psi}(\hat{\mathbf{P}})}$ .

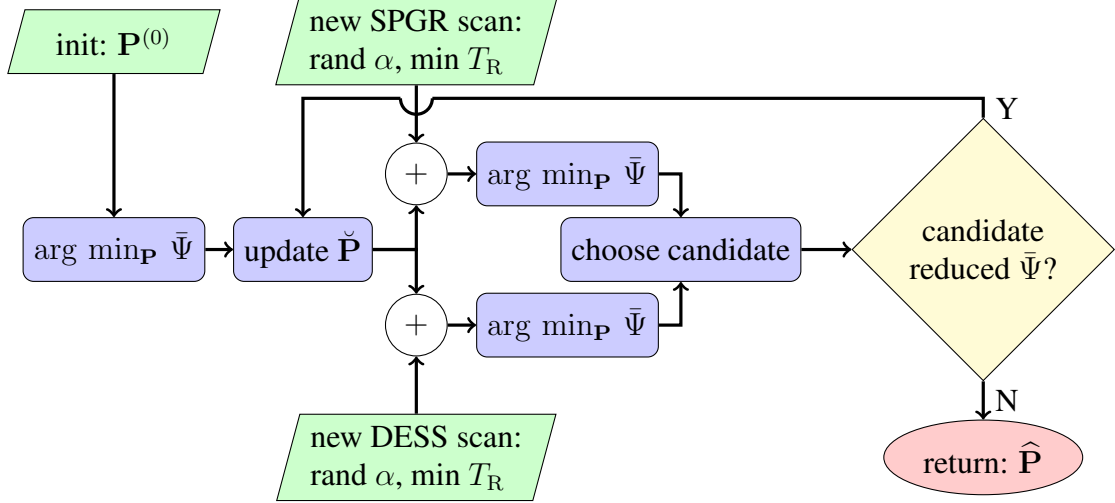


Figure 6.1: Block diagram of greedy scan profile construction. We set an original acquisition consisting of three SPGR and three DESS scans whose flip angles and repetition times are respectively initialized randomly and minimally. We optimize the original acquisition subject to several constraints:  $\alpha \in [1, 40]^\circ$  and  $T_R \geq 11.8\text{ms}$  for SPGR;  $\alpha \in [1, 60]^\circ$  and  $T_R \geq 17.5\text{ms}$  for DESS; and  $\sum_{d=1}^D T_{R,d} \leq 263.7\text{ms}$  (selected to allow 9 SPGR and 9 DESS scans each at minimal  $T_R$ , comparable to [2]). We then seek to improve the acquisition by appending another SPGR or DESS scan and repeating (constrained) optimization. This process iterates until additional scans no longer improve estimation precision.

{fig:mwf,schem}

	Optimized flip angles (deg)	Optimized repetition times (ms)
SPGR	$[38.1, 12.9, 9.2, 33.5]^\top$	$[50.2, 32.4, 16.4, 11.8]^\top$
DESS	$[32.0, 40.3, 52.9]^\top$	$[17.5, 98.0, 37.6]^\top$

Table 6.1: SPGR/DESS flip angles and repetition times that comprise  $\hat{\mathbf{P}}$ , a scan parameter matrix designed under total repetition time budget  $\sum_{d=1}^D T_{R,d} \leq 263.7\text{ms}$  for precise  $f_F$  estimation in WM via the sequence of optimization problems described in Figure 6.1. For our noise variance measurements, this acquisition is expected to yield 28.5% relative standard deviation in asymptotically unbiased  $f_F$  estimates from two-compartment SPGR (6.22) and DESS (6.41)-(6.42) signal models. Remarkably, this optimized scan profile involving (substantial)  $T_R$  variation achieves better  $f_F$  estimation precision than a similarly optimized scan profile with fixed minimum- $T_R$  acquisitions, even though the latter can utilize many more scans for the given time budget.

{tab:mwf,acq}

Table 6.1 summarizes the optimized scan parameter  $\hat{\mathbf{P}}$  that arises from one representative realization of our scan profile construction procedure. Using `fmincon` in MATLAB<sup>®</sup> R2013a for each inner-step optimization (with the “active set” option and with an objective function convergence tolerance of  $10^{-7}$ ), this realization took 64s run time on a 3.5GHz desktop computer with 32GB RAM. We find that  $\sqrt{\bar{\Psi}(\hat{\mathbf{P}})} = 0.285$ , meaning that at a realistic noise level, the SPGR/DESS acquisition defined by  $\hat{\mathbf{P}}$  is expected to yield 28.5%

relative standard deviation in asymptotically unbiased  $f_F$  estimates from SPGR (6.22) and DESS (6.41)-(6.42) non-exchanging two-compartment signal models.

Interestingly, we observe empirically that the above scan profile construction procedure tends to append far fewer additional scans than is possible under our time budget: here, it halts at a small scan profile consisting of only 4 SPGR and 3 DESS scans. We investigated this phenomenon further by repeating the above procedure under the same time constraint, but now initializing with a profile consisting of 9 SPGR and 9 DESS scans whose repetition times are implicitly restricted to be minimal. Over many random flip angle initializations, we found after optimization that this expanded minimum- $T_R$  scan profile could achieve an expected  $f_F$  relative standard deviation no better than 0.387, suggesting that  $T_R$  variation across scans is important for building scan profiles that permit precise  $f_F$  estimation.

## 6.4 Experimentation

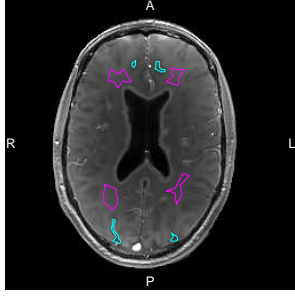
{s,mwf,exp}

As a second step toward validation (the first being simulation results presented in Section 5.4), we estimate  $f_F$  maps from *in vivo* datasets acquired using the precision-optimized scan profile developed in Section 6.3. For preliminary discussion, we qualitatively compare our  $f_F$  maps with other state-of-the-art MWF proxies described in recent literature.

In a single continuous study of a healthy volunteer, we collected a scan profile consisting of 4 SPGR and 3 DESS datasets whose flip angles and repetition times (listed in Table 6.1) are optimized to permit precise  $f_F$  estimation. We obtained 3D axial slices using a GE Discovery™ MR750 3.0T scanner with a 32-channel Nova Medical® receive head array. All other acquisition details were identical to those described in Section 4.4.2.1. Most importantly, we acquired all data at fixed  $T_E \leftarrow 4.67\text{ms}$  before and after RF excitations, which is minimal given other pulse sequence timing requirements and thus minimizes model mismatch due to compartmental off-resonance effects (characterized in Section 6.2). Also as in Section 4.4.2.1, we separately estimated transmit field sensitivity  $\kappa^t$  from a Bloch-Siegert acquisition [92] and took  $\kappa^t$  as known for  $f_F$  estimation. For  $256 \times 256 \times 8$  fully-sampled 3D Cartesian acquisitions (each over a  $240 \times 240 \times 40\text{mm}^3$  field of view), SPGR/DESS and Bloch-Siegert acquisitions respectively took totals of 10m8s and 1m40s.

After Fourier transform of the raw coil SPGR/DESS data, we normalized and combined one central slice of each coil image using an extension of [94] for the case of multiple datasets, detailed in Appendix A. We then applied KRR parameter estimation exactly as in simulations (discussed in Section 5.4) to estimate  $f_F$  and five other nuisance parameter maps from magnitude coil-combined SPGR/DESS image data. To refine results, we lastly





ROI (# voxels)	KRR	KRR, then ML
WM $\hat{f}_F$ (725)	$0.130 \pm 0.074$	$0.134 \pm 0.070$
GM $\hat{f}_F$ (176)	$0.063 \pm 0.106$	$0.063 \pm 0.098$

Table 6.2: *Left*: WM/GM ROIs, overlaid on a representative anatomical (coil-combined SPGR) image. Four WM ROIs and four GM ROIs are each pooled into a single WM and a single GM ROI, over which sample statistics are computed. *Right*: Within-ROI sample means  $\pm$  within-ROI sample standard deviations of  $f_F$  estimates, using KRR only as well as KRR with ML refinement (Fig. 6.2 presents corresponding images).

solved ML parameter estimation problem (3.7) via iterative local optimization<sup>7</sup>, using KRR parameter estimates as initialization. KRR training (with simulated data), KRR testing (with real data), and iterative refinement respectively took 35.2s, 5.4s, and 27.5s.

Figure 6.2 compares resulting  $f_F$  estimates using KRR only versus KRR followed by iterative ML estimation. ML iterations produce visible improvement in cerebrospinal fluid but leave  $f_F$  estimates in WM and GM ROIs largely unchanged.

For qualitative comparison, Figure 6.2 also reprints from a comparison study [6] MWF proxy estimates from two other acquisitions: GRASE [3], a state-of-the-art acceleration to the slow gold-standard MESE acquisition [4]; and mcDESPOT [2], a faster steady-state acquisition that lacks sufficient estimation precision [5]. These images are from a separate study of a different healthy volunteer, and utilize different algorithms to estimate different MWF proxies from multi-compartmental signal models of different acquisitions. Acknowledging the lack of controls, we cautiously observe that the proposed optimized SPGR/DESS scan profile is as fast as the mcDESPOT acquisition (controlling for scanner timing limitations) and yields MWF maps at least comparable in WM to those of GRASE.

Table 6.2 summarizes within-ROI samples means and sample standard deviations computed over manually-selected ROIs containing 725 WM and 176 GM voxels. Compared to its KRR initialization, local ML optimization does not appreciably change WM/GM  $\hat{f}_F$  sample means and mildly reduces sample standard deviations. Our WM  $\hat{f}_F$  sample means agree well with literature MESE measurements (*e.g.*, see [6, Fig. 3]). Our GM  $\hat{f}_F$  sample means are not statistically meaningful because of high within-ROI sample variation, likely

<sup>7</sup>We ran preconditioned gradient projection method iterations (2.46) with the same constraint set used to sample training points for KRR and a diagonal preconditioning matrix defined as the inverse diagonal entries of the cost function’s Hessian matrix. We used a step-halved backtracking line search to ensure monotone descent in cost over iterations. We iterated until convergence criterion (4.14) was met with tolerance  $10^{-8}$ .



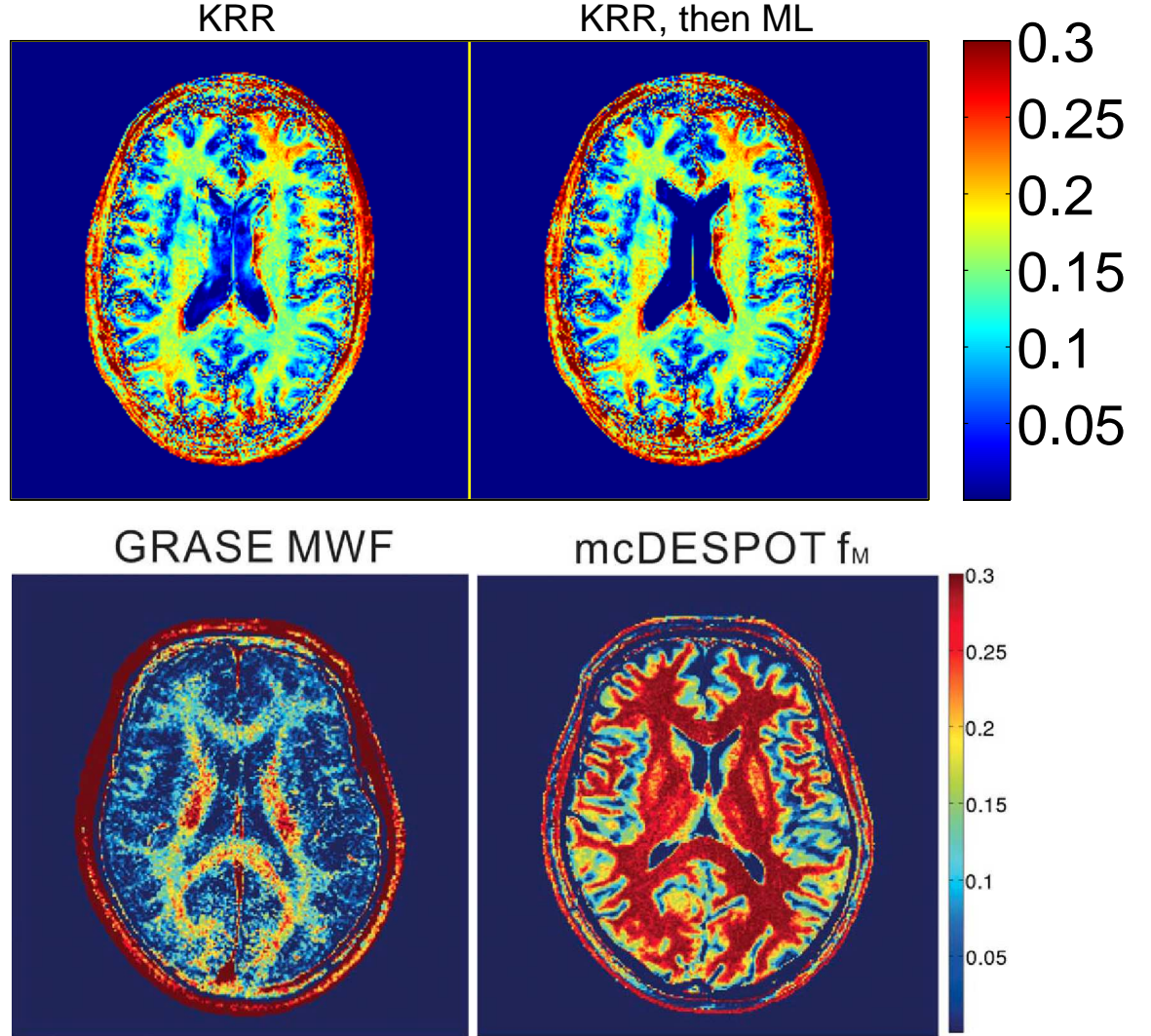


Figure 6.2: MWF estimates, using: our optimized SPGR/DESS scan profile and (*top left*) KRR  $f_F$  estimation, followed by (*top right*) iterative ML estimation; (*bottom left*) GRASE [3], a state-of-the-art acceleration to the slow gold-standard MESE acquisition originally used in [4]; and (*bottom right*) mcDESPOT [2], a fast steady-state acquisition that lacks sufficient estimation precision [5]. All MWF estimates are from healthy volunteers, but those from GRASE/mcDESPOT are from a separate study in a different subject and are reprinted from [6], a recent comparison study. The proposed precision-optimized SPGR/DESS scan profile is as fast as mcDESPOT and yields MWF images comparable to those of GRASE.

due to the common challenge of poor GM ROI selection. Nevertheless, it warrants mention that MWF in healthy GM was reported in [6, Fig. 3] to be about 0.05 using GRASE.

## 6.5 Summary and Future Work

{s,mwf,summ}

This chapter describes initial work on developing a fast steady-state acquisition for precise MWF imaging. We have thus far developed a two-compartment signal model for the DESS pulse sequence; used a simple non-exchanging version of this model (along with a similar SPGR model from literature) to design a fast scan profile for precisely estimating fast-relaxing compartmental fraction  $f_F$  as a proxy for MWF; and applied this precision-optimized acquisition *in vivo* to produce to our knowledge the first MWF images from combinations of fast SPGR/DESS scans. Though preliminary results appear qualitatively comparable in healthy WM to state-of-the-art literature MESE results, thorough validation is required and will constitute a major portion of future work, discussed in the following.

**Comparisons with MESE MWF Estimates** We would like to claim that fast-relaxing fraction  $f_F$  estimates from precision-optimized SPGR/DESS scan profiles are accurate measurements of MWF as originally defined in [4]. To build support for this hypothesis, we will need to systematically compare our  $f_F$  estimates with MWF estimates from a MESE acquisition. Such comparison is challenging because the methods differ not only in acquisitions but also in modeling assumptions<sup>8</sup>, initialization (KRR estimate (5.9) versus the zero vector), objective functions (ML cost (3.5) versus a regularized nonnegative least-squares (NNLS) cost [143, Eq. 8]), and estimation algorithms (gradient projection method [25] versus least-distance programming (LDP) [144, §23.4]).

We plan to approach validation as follows. First, we will compare *in vivo* MWF estimates from a fixed signal model of a fixed MESE acquisition using KRR, KRR-initialized ML, KRR-initialized regularized NNLS, and zero-initialized regularized NNLS<sup>9</sup>. Here, all three iterative algorithms can be solved via gradient projection method for consistency. These results will provide direct insight as to how KRR and KRR-initialized ML estimation perform for *in vivo* MWF estimation from a classical acquisition and signal model.

Using KRR as above for a convex MWF estimation problem is important for systematic validation but is overkill in practice. For a more interesting usage, we will next consider

---

<sup>8</sup>MESE MWF estimation assumes  $T_1$  is known and fixed over many compartments. As presented, SPGR/DESS  $f_F$  estimation assumes unknown variable  $T_1$  across two compartments. Both models neglect exchange and multi-compartmental off-resonance effects, and assume known flip angle spatial variation.

<sup>9</sup>Both KRR- and zero-initialized NNLS should yield near-equivalent solutions because the regularized NNLS problem is convex, assuming as in [143, 4] that all latent parameters other than MWF are known.

joint estimation of MWF and transmit field variation  $\kappa^t$  (still assuming knowledge of other latent parameters) from the same MESE dataset. NNLS is inapplicable for this non-convex problem; however, we can still quickly validate KRR estimates via voxel-wise exhaustive grid search due to low problem dimensionality. If promising, these results will support the claim that KRR could be useful for fast *in vivo* MWF estimation from nonlinear models.

To compare SPGR/DESS and MESE acquisitions, we can modify SPGR/DESS model assumptions to those taken in [3] for the MESE model (listed in Footnote 8) and thereby formulate an analogous  $T_2$  distribution estimation problem, for which linear algorithms such as regularized NNLS and LDP are applicable. Our task will then be to design a fast SPGR/DESS acquisition under these modified model assumptions that permits precise voxel-wise estimation of the  $T_2$  distribution. The resulting estimated  $T_2$  distributions will be directly comparable to  $T_2$  distributions from MESE acquisitions and could be used to form SPGR/DESS MWF maps that are directly comparable with MESE MWF maps.

Upon isolated validation of acquisitions, initializations, and estimation algorithms, we can explore the effects of different modeling assumptions using the full flexibility of optimized scan design and KRR estimation. As a first experiment, we could try to optimize under the variable- $T_1$  two-compartment model a MESE acquisition similar to the SPGR/DESS acquisition presented in Table 6.1 and compare resulting SPGR/DESS and MESE  $\hat{f}_F$  maps. If  $\hat{f}_F$  from these models do not measure MWF with sufficient specificity, we could explore similar comparisons using more complicated models involving exchange, three (or more) compartments, and/or compartmental off-resonance effects. All of these model comparisons will require highly nonlinear estimation problems involving many latent parameters, for which KRR-initialized ML estimation is well-suited.

**Comparisons with Different Myelin Biomarkers** We are also interested in how  $f_F$  estimates (or other MWF proxies) correlate with other myelin biomarkers. One contending noninvasive marker arises from MR pulse sequences sensitized to the inhomogeneous magnetization transfer (ihMT) effect [145], which has recently been argued to be specific to the large membrane lipids that comprise much of myelin [146, 147]. Outside MRI, invasive measurements from histology have been used to study myelin (as in *e.g.*, [135, 136]) and could serve as a gold-standard *in situ* marker. Multi-compartmental and ihHT MRI markers can be successively compared through phantom, healthy volunteer, and patient studies. Both MRI and histological markers can be compared through phantom and *in situ* studies. All of these studies are highly amenable to clinical collaborations.

## CHAPTER 7

### Future Work

{c,future}

This chapter organizes the numerous avenues of future work itemized in Sections 5.5 and 6.5 into a timeline for completion of this dissertation. For the short term, we place emphasis on future work items that are likely to lead to journal publications that expand on Chapters ??-6. Other opportunities are realistically listed as longer-term research goals.

2017-04	<p>prepare for and travel to conferences</p> <ul style="list-style-type: none"> <li>• partially validate KRR estimates from SPGR/DESS acquisition given in 6.1 via exhaustive grid search</li> <li>• present [12] at ISBI</li> <li>• present [13] at ISMRM annual meeting</li> </ul>
2017-05	propose dissertation
2017-05	<p>experimentally validate KRR estimation</p> <ul style="list-style-type: none"> <li>• implement and test MESE acquisition</li> <li>• research MESE forward model; implement NNLS estimation</li> <li>• replicate <i>in vivo</i> MESE MWF maps in [6], using KRR versus gold-standard NNLS</li> <li>• repeat above while jointly estimating transmit field sensitivity, using KRR versus exhaustive grid search</li> </ul>
2017-08	<p>prepare journal paper on scalable MRI parameter estimation via KRR</p> <ul style="list-style-type: none"> <li>• analyze performance through theory and simulations</li> <li>• seek to exploit partially linear structure</li> <li>• present simulation and experimental results, first showing that KRR can replicate classical MWF maps <i>in vivo</i>, and then showing that KRR has flexibility to handle more complicated nonlinear models of faster (precision-optimized) SPGR/DESS acquisitions</li> </ul>
2017-10	<p>experimentally validate <math>f_F</math> estimates from SPGR/DESS acquisition</p> <ul style="list-style-type: none"> <li>• design SPGR/DESS acquisition under MESE model assumptions</li> <li>• compare <i>in vivo</i> MWF estimates from optimized SPGR/DESS vs. classical MESE acquisitions</li> <li>• repeat scan design/comparison using increasingly complex model assumptions until MWF estimates from SPGR/DESS and MESE agree reasonably</li> </ul>
2018-01	prepare journal paper on MWF imaging from SPGR/DESS acquisitions
2018-03	defend dissertation
	<p>longer-term research directions</p> <ul style="list-style-type: none"> <li>• combine image reconstruction and KRR parameter estimation</li> <li>• correlate our MWF estimates with other myelin biomarkers</li> <li>• apply KRR to other MRI parameter estimation problems</li> </ul>

## APPENDIX A

### Coil Data Combination from Multiple Datasets

{a,cc-multi}

In the completed thesis, this appendix will describe an unpublished algorithm for combining multiple MRI datasets when each dataset is acquired using multiple receiver coils with fixed coil geometry. Such sequences of coil datasets arise naturally in many quantitative MRI applications, including the ones studied in Chapters 4-???. The algorithm extends [94] (which considers joint image reconstruction and coil sensitivity estimation in the case of a single dataset) to exploit the fixed coil sensitivity across several datasets, thereby improving estimation conditionality as compared to separate coil combination across datasets.

## APPENDIX B

### DESS in the Presence of Diffusion

{a,deSS-diff}

In the completed thesis, this appendix will present an unpublished analysis of model mismatch in single-compartment DESS signal models (2.41) and (2.43) in the presence of diffusion. We first develop models that describe the single-compartment DESS signal in the presence of diffusion. We then use these models to show through phantom experiments that neglecting diffusive effects during  $T_2$  estimation (as in Chapters 3-4) induces significant estimation bias. We conclude with recommendations for MR acquisition settings (specifically, for dephasing gradient area) that reduce diffusion-related estimation bias without excessively imparting other bias. These recommendations were used to guide DESS acquisition design for all relevant experiments considered in the main body of this thesis.

## 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. 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.
- [3] T. Prasloski, A. Rauscher, A. L. MacKay, M. Hodgson, I. M. Vavasour, C. Laule, and B. Mädler, "Rapid whole cerebrum myelin water imaging using a 3D GRASE sequence," *NeuroImage*, vol. 63, pp. 533–9, Oct. 2012.
- [4] 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.
- [5] C. L. Lankford and M. D. Does, "On the inherent precision of mcDESPOT," *Mag. Res. Med.*, vol. 69, pp. 127–36, Jan. 2013.
- [6] J. Zhang, S. H. Kolind, C. Laule, and A. L. MacKay, "Comparison of myelin water fraction from multiecho T2 decay curve and steady-state methods," *Mag. Res. Med.*, vol. 73, pp. 223–32, Jan. 2015.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] 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.*, vol. 36, pp. 467–77, Feb. 2017.



- [11] 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.
- [12] 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.*, pp. 5–9, 2017.
- [13] 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.*, p. 5076, 2017.
- [14] A. Macovski, *Medical imaging systems*. New Jersey: Prentice-Hall, 1983.
- [15] E. M. Haacke, R. W. Brown, M. R. Thompson, and R. Venkatesan, *Magnetic resonance imaging: Physical principles and sequence design*. New York: Wiley, 1999.
- [16] D. G. Nishimura, "Principles of magnetic resonance imaging," 1996. Unpublished textbook.
- [17] F. Bloch, "Nuclear induction," *Phys. Rev.*, vol. 70, pp. 460–74, Oct. 1946.
- [18] W. S. Hinshaw, "Image formation by nuclear magnetic resonance: The sensitive point method," *J. Appl. Phys.*, vol. 47, p. 3709, Aug. 1976.
- [19] 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.
- [20] 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.
- [21] 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.
- [22] 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.
- [23] T. W. Redpath and R. A. Jones, "FADE-A new fast imaging sequence," *Mag. Res. Med.*, vol. 6, pp. 224–34, Feb. 1988.
- [24] 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.
- [25] 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.

- [26] C. Byrne, “A unified treatment of some iterative algorithms in signal processing and image reconstruction,” *Inverse Prob.*, vol. 20, pp. 103–20, Feb. 2004.
- [27] 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.
- [28] 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.
- [29] 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.
- [30] A. Macovski, “Noise in MRI,” *Mag. Res. Med.*, vol. 36, pp. 494–7, Sept. 1996.
- [31] T. Lei, “Statistics of MR signals: revisited,” in *Proc. SPIE 6510 Medical Imaging: Phys. Med. Im.*, p. 651052, 2007.
- [32] H. Gudbjartsson and S. Patz, “The Rician distribution of noisy MRI data,” *Mag. Res. Med.*, vol. 34, pp. 910–4, Dec. 1995.
- [33] 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.
- [34] 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.
- [35] 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.
- [36] 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.
- [37] R. K. Gupta, “A new look at the method of variable nutation angle for the measurement of spin-lattice relaxation times using Fourier transform NMR,” *J. Mag. Res.*, vol. 25, pp. 231–5, Jan. 1977.
- [38] 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.
- [39] D. P. Bertsekas, “Projected Newton methods for optimization problems with simple constraints,” *SIAM J. Cont. Opt.*, vol. 20, no. 2, pp. 221–46, 1982.
- [40] J. A. Fessler, “Michigan image reconstruction toolbox (MIRT) for Matlab,” 2016. Available from [web.eecs.umich.edu/~fessler](http://web.eecs.umich.edu/~fessler).

- [41] 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.
- [42] 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.
- [43] 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.
- [44] 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.
- [45] 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.
- [46] 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.
- [47] 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.
- [48] 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.
- [49] 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.
- [50] 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.
- [51] 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.
- [52] 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.
- [53] 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.

- [54] 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.
- [55] 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.
- [56] 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.
- [57] 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.
- [58] 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.
- [59] 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.
- [60] 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.
- [61] 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.
- [62] 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.
- [63] 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.
- [64] 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.
- [65] 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.

- [66] 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.
- [67] 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.
- [68] 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.
- [69] H. Wang and Y. Cao, "Spatially regularized T1 estimation from variable flip angles MRI," *Med. Phys.*, vol. 39, pp. 4139–48, July 2012.
- [70] 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.
- [71] 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.
- [72] G. H. Welsch, K. Scheffler, T. C. Mamsch, 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.
- [73] 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.
- [74] 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.
- [75] R. Heule, C. Ganter, and O. Bieri, "Triple echo steady-state (TESS) relaxometry," *Mag. Res. Med.*, vol. 71, pp. 230–7, Jan. 2014.
- [76] M. L. Gyngell, "The steady-state signals in short-repetition-time sequences," *J. Mag. Res.*, vol. 81, pp. 474–83, Feb. 1989.
- [77] 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.
- [78] 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.

- [79] 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.
- [80] 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.
- [81] 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.
- [82] 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.
- [83] 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.
- [84] L. Fleysheer, R. Fleysheer, 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.
- [85] 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.
- [86] 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.
- [87] Y. Liu, J. R. Buck, and V. N. Ikonomidou, "Generalized min-max bound-based MRI pulse sequence design framework for wide-range T1 relaxometry: A case study on the tissue specific imaging sequence," *PLoS One*, vol. 12, pp. 1–20, 02 2017.
- [88] R. A. Fisher, "Theory of statistical estimation," *Proc. Cambridge Philosophical Society*, vol. 22, pp. 700–25, July 1925.
- [89] H. Cramér, *Mathematical methods of statistics*. Princeton: Princeton Univ. Press, 1946.
- [90] H. Chernoff, "Locally optimal designs for estimating parameters," *Ann. Math. Stat.*, vol. 24, pp. 586–602, Dec. 1953.
- [91] 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.

- [92] 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.
- [93] 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.
- [94] L. Ying and J. Sheng, “Joint image reconstruction and sensitivity estimation in SENSE (JSENSE),” *Mag. Res. Med.*, vol. 57, pp. 1196–1202, June 2007.
- [95] 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.
- [96] 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.
- [97] 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.
- [98] 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.
- [99] 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.
- [100] 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.
- [101] 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.
- [102] 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.
- [103] 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.
- [104] B. Zhao, J. Haldar, K. Setsompop, and L. L. Wald, “Optimal experiment design for magnetic resonance fingerprinting,” in *embc*, pp. 453–6, 2016.
- [105] H. C. Torrey, “Bloch equations with diffusion terms,” *Phys. Rev.*, vol. 104, pp. 563–5, 1956.
- [106] H. M. McConnell, “Reaction rates by nuclear magnetic resonance,” *J. of Chemical Phys.*, vol. 28, pp. 430–31, Mar. 1958.

- [107] 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.
- [108] D. L. Bihan, J.-F. Mangin, C. Poupon, C. A. Clark, S. Pappata, N. Molko, and H. Chabriat, "Diffusion tensor imaging: Concepts and applications," *J. Mag. Res. Im.*, vol. 13, pp. 534–546, Apr. 2001.
- [109] 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.
- [110] 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.
- [111] 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.
- [112] 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.
- [113] 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.
- [114] G. R. Morrell, "A phase-sensitive method of flip angle mapping," *Mag. Res. Med.*, vol. 60, pp. 889–94, Oct. 2008.
- [115] 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.
- [116] 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.
- [117] G. S. Kimeldorf and G. A. Wahba, "A correspondence between Bayesian estimation on stochastic processes and smoothing by splines," *Ann. Math. Stat.*, vol. 41, pp. 495–502, Apr. 1970.
- [118] N. Aronszajn, "Theory of reproducing kernels," *Trans. Amer. Math. Soc.*, vol. 68, pp. 337–404, May 1950.
- [119] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, pp. 273–97, Sept. 1995.



- [120] C. Saunders, A. Gammerman, and V. Vovk, "Ridge regression learning algorithm in dual variables," in *Proc. Intl. Conf. Mach. Learn.*, pp. 515–21, 1998.
- [121] B. Schölkopf, R. Herbrich, and A. J. Smola, "A generalized representer theorem," in *Proc. Computational Learning Theory (COLT)*, pp. 416–426, 2001. LNCS 2111.
- [122] A. E. Hoerl and R. W. Kennard, "Ridge regression: biased estimation for nonorthogonal problems," *Technometrics*, vol. 12, pp. 55–67, Feb. 1970.
- [123] A. Rahimi and B. Recht, "Random features for large-scale kernel machines," in *NIPS*, 2007.
- [124] D. J. Sutherland and J. Schneider, "On the error of random Fourier features," in *Proc. Intl. Conf. on Uncertainty in AI*, 2015.
- [125] Z. Wu, "Generalized Bochner's theorem for radial function," *Approximation Theory and its Applications*, vol. 13, no. 3, pp. 47–57, 1997.
- [126] M. A. Woodbury, "Inverting modified matrices," 1950. Tech. Report 42, Stat. Res. Group, Princeton Univ.
- [127] E. Parzen, "On estimation of a probability density function and mode," *Ann. Math. Stat.*, vol. 33, pp. 1065–76, Sept. 1962.
- [128] C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning (adaptive computation and machine learning)*. MIT Press, 2005.
- [129] M. M. Siddiqui, "Statistical inference for Rayleigh distributions," *RADIO SCIENCE Journal of Research NBS/USNC-URSI*, vol. 68D, pp. 1005–10, Sept. 1964.
- [130] P. Morell, *Myelin*. Springer, 1984.
- [131] M. M. Goldenberg, "Multiple sclerosis review," *Pharmacy and Therapeutics*, vol. 37, pp. 175–84, Mar. 2012.
- [132] V. Vasilescu, E. Katona, V. Simplaceanu, and D. Demco, "Water compartments in the myelinated nerve. III. Pulsed NMR results," *Experientia*, vol. 34, pp. 1443–4, Nov. 1978.
- [133] R. S. Menon and P. S. Allen, "Application of continuous relaxation time distributions to the fitting of data from model systems and excised tissue," *Mag. Res. Med.*, vol. 20, pp. 214–27, Aug. 1991.
- [134] W. A. Stewart, A. L. Mackay, K. P. Whittall, G. R. W. Moore, and D. W. Paty, "Spin-spin relaxation in experimental allergic encephalomyelitis. Analysis of CPMG data using a non-linear least-squares method and linear inverse theory," *Mag. Res. Med.*, vol. 29, pp. 767–75, June 1993.

- [135] P. J. Gareau, B. K. Rutt, S. J. Karlik, and J. R. Mitchell, "Magnetization transfer and multicomponent T2 relaxation measurements with histopathologic correlation in an experimental model of MS," *J. Mag. Res. Im.*, vol. 11, pp. 586–95, June 2000.
- [136] S. Webb, C. A. Munro, R. Midha, and G. J. Stanisz, "Is multicomponent T2 a good measure of myelin content in peripheral nerve?," *Mag. Res. Med.*, vol. 49, pp. 628–45, Apr. 2003.
- [137] C. Laule, I. M. Vavasour, G. R. W. Moore, J. Oger, D. K. B. Li, D. W. Paty, and A. L. MacKay, "Water content and myelin water fraction in multiple sclerosis," *J. Neurol.*, vol. 251, pp. 284–93, Mar. 2004.
- [138] C. Laule, E. Leung, D. K. B. Li, A. L. Traboulsee, D. W. Paty, A. L. MacKay, and G. R. W. Moore, "Myelin water imaging in multiple sclerosis: quantitative correlations with histopathology," *Multiple Sclerosis*, vol. 12, pp. 747–53, Nov. 2006.
- [139] R. G. Spencer and K. W. Fishbein, "Measurement of spin-lattice relaxation times and concentrations in systems with chemical exchange using the one-pulse sequence: breakdown of the Ernst model for partial saturation in nuclear magnetic resonance spectroscopy," *J. Mag. Res.*, vol. 142, pp. 120–35, Jan. 2000.
- [140] K. L. Miller, "Asymmetries of the balanced SSFP profile. Part I: Theory and observation," *Mag. Res. Med.*, vol. 63, pp. 385–95, Feb. 2010.
- [141] K. L. Miller, S. M. Smith, and P. Jezzard, "Asymmetries of the balanced SSFP profile. Part II: White matter," *Mag. Res. Med.*, vol. 63, pp. 396–406, Feb. 2010.
- [142] W. Magnus, "On the exponential solution of differential equations for a linear operator," *Comm. Pure Appl. Math.*, vol. 7, pp. 649–73, Nov. 1954.
- [143] K. P. Whittall and A. L. MacKay, "Quantitative interpretation of NMR relaxation data," *J. Mag. Res.*, vol. 84, pp. 134–52, Aug. 1989.
- [144] C. L. Lawson and R. J. Hanson, *Solving least squares problems*. Prentice-Hall, 1974.
- [145] G. Varma, G. Duhamel, C. de Bazelaire, and D. C. Alsop, "Magnetization transfer from inhomogeneously broadened lines: A potential marker for myelin," *Mag. Res. Med.*, vol. 73, pp. 614–22, Feb. 2015.
- [146] G. Varma, O. M. Girard, V. H. Prevost, G. Duhamel, and D. C. Alsop, "Interpretation of magnetization transfer from inhomogeneously broadened lines (ihMT) in tissues as a dipolar effect within motion restricted molecules," *J. Mag. Res.*, vol. 260, pp. 67–76, Nov. 2015.
- [147] S. D. Swanson, D. I. Malyarenko, M. L. Fabiilli, R. C. Welsh, J.-F. Nielsen, and A. Srinivasan, "Molecular, dynamic, and structural origin of inhomogeneous magnetization transfer in lipid membranes," *Mag. Res. Med.*, vol. 77, pp. 1318–28, Mar. 2017.