Model-Based Methods in Quantitative MRI

Gopal Nataraj
Advisor: Dr. Jeffrey A. Fessler
Collaborators: Dr. Jon-Fredrik Nielsen, Dr. Daniel S. Weller

EE: Systems Qualifier Examination
Date Submitted: November 21, 2014

Contents

Abstract

In this work, we investigate the use of model-based image reconstruction (MBIR) for Quantitative Magnetic Resonance (QMR) imaging. In our context, MBIR refers to a broad class of algorithms that seek to recover images ("maps") of desired parameters, by minimizing a well-constructed cost function. Such cost functions typically involve (1) data fidelity term(s) based on the log-likelihood of (nonlinear) signal model(s), and (2) regularization term(s) that capture *a priori* assumptions, such as sparsity of the image gradients, to control noise.

In particular, we strive for fast and accurate quantification of spin-lattice T_1 and spin-spin T_2 relaxation parameter maps because of their importance for clinical MRI applications. Classical pulse sequences yield straightforward T_1 and T_2 estimates, but require undesirably lengthy acquisitions. By contrast, steady-state pulse sequences are considerably faster, but produce signals that are complicated functions of both desired and nuisance parameters. Conventional method-of-moments estimators often exhibit systematic errors because of approximations used to bypass nuisance parameter estimation. To improve mapping accuracy, yet retain fast scan times, we explore the use of MBIR on combinations of steady-state scans. Although we focus here on T_1 and T_2 estimation, MBIR can be applied to other parameter mapping problems as well.

This report is organized into four chapters. In Chapter One, we discuss T_1 estimation from Spoiled Gradient-Recalled Echo (SPGR) scans. In Chapter Two, we describe sequential T_1, then T_2 estimation from SPGR and Dual-Echo Steady-State (DESS) scans, respectively. In Chapter Three, we employ the Cramér-Rao Lower Bound (CRLB) to address the broader question of optimal scan selection for joint T_1, T_2 estimation, as a function of controllable scan parameters. In Chapter Four, we step away from QMR imaging to explore MBIR from accelerated, under-sampled acquisitions. We conclude with discussion of combining the ideas of the latter chapters as future work.

Chapter 1

Model-Based Estimation of T_1 Maps (MBET1) from SPGR Sequences

1.1 Introduction

Since its inception, Magnetic Resonance Imaging (MRI) has been widely accepted as powerful, non-invasive imaging modality that provides excellent soft-tissue contrast and avoids the use of ionizing radiation. Despite their advantages, MR images typically require professional interpretation because the received signals are non-physical, complicated functions of not only the underlying spin-lattice (T_1) and spin-spin (T_2) relaxation parameters that provide the desired contrast, but also numerous nuisance parameters governed by the scan protocol.

The field of Quantitive Magnetic Resonance (QMR) imaging arose to address the task of disentangling interpretable maps of physical parameters from raw data. Historically, researchers have approached this problem from two broad directions. One classical strategy has been to design the scan protocol such that the signal dynamics are dominated by the parameters of interest. While such techniques ensure easy recovery of desired parameters, they often require undesirably lengthy scan times. This has lead to the more recent strategy of using considerably faster scans to collect numerous datasets in as short a time as possible, leaving the burden of mapping as a post-processing step. The ideas behind clever scan design and careful post-processing are far from mutually exclusive, and can together yield methods for fast, accurate parameter mapping.

In this chapter, we are specifically interested in T_1 estimation. We avoid the lengthy scan times of classical pulse sequences such as Inversion Recovery (IR) and Saturation Recovery (SR) pulse sequences, and instead utilize a rapid series of Spoiled Gradient-Recalled Echo (SPGR) steady-state scans, repeated at variable flip angles. Conventional, log-linear least-squares T_1 estimation [1, 2] from this data exhibits systemic errors in low-signal regions. To improve T_1 mapping precision, we introduce Model-Based Estimation of T_1 Maps (MBET1) [3]. Though we focus here on SPGR data, the ideas behind MBET1 can be applied to any pulse sequence that provides high T_1 contrast.

In Section 2, we briefly derive the SPGR signal model from the Bloch equations, describe the short-comings of a conventional least-squares estimator, and introduce a penalized-likelihood estimator

that addresses these shortcomings. In Section 3, we discuss in detail an alternating minimization approach to solving this problem. In Section 4, we describe both synthetic and *in vivo* experiments. In Section 5, we briefly conclude our findings and describe how in later chapters, MBET1 will be taken as an intermediary step before T_2 estimation.

1.2 Problem Formulation

1.2.1 SPGR Signal Model

Beginning with the Bloch equations [4], we develop a model describing the steady-state signal received from an SPGR pulse sequence. In this analysis, we neglect B_1 inhomogeneity, and thus assume that all excitations are of the same flip angle. We will observe later that these assumptions are reasonable because the model mismatch they incur is dominated by that of incomplete spoiling. We orient the transverse axes such that RF excitation rotations occur about the i-axis.

Let $\vec{M}(0)$ denote the magnetization immediately following an RF pulse. Until the start of the next RF pulse, each voxel's magnetization $\vec{M}(t)$ evolves according to the Bloch equations:

$$\vec{M}(t) = \underbrace{\begin{bmatrix} E_2(t) & 0 & 0 \\ 0 & E_2(t) & 0 \\ 0 & 0 & E_1(t) \end{bmatrix} \begin{bmatrix} \cos\beta(t) & \sin\beta(t) & 0 \\ -\sin\beta(t) & \cos\beta(t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \vec{M}(0)}_{\mathbf{A}_\beta(t)} + \underbrace{M_0\left(1 - E_1(t)\right) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{\vec{b}(t)}, \quad (1.1)$$

where the longitudinal and transverse relaxation effects are given by $E_1(t) = e^{-t/T_1}$ and $E_2(t) = e^{-t/T_2}$, respectively, and the cumulative phase evolution due to all off-resonance effects (gradient fields, variations in susceptibility, B_0 inhomogeneity, chemical shift, etc.) are summarized as

$$\beta(t) = \int_0^t \Delta\omega(t') \, dt', \quad (1.2)$$

a phase rotation in the transverse plane. Note that off-resonance effects $\Delta\omega(t)$ are themselves usually time-varying effects, particularly with the use of gradients and spoiler fields.

Following readout but prior to RF excitation, residual transverse magnetization is removed through some form of spoiling [5]. An ideal spoiling operator can be represented as $\mathbf{S} := \mathrm{diag}\{0, 0, 1\}$.

After pulse time T, an RF pulse of duration $\tau \ll T_E$ and angle α is applied to complete one RF cycle. We assume that the pulse is sufficiently short that off-resonance effects, spin relaxation, and slice-selection gradients are negligible, as compared to the RF pulse. As mentioned before, we orient the transverse axes such that the rotation operator,

$$\mathbf{R}_\alpha := \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{bmatrix}, \quad (1.3)$$

occurs about the i-axis. Applying the spoiling operator followed by the RF pulse models the magnetization following one period, of duration $T_R := T + \tau$:

$$
\begin{aligned}
\vec{M}(T_R) &= \mathbf{R}_\alpha \mathbf{S} \vec{M}(T) \\
&= \mathbf{R}_\alpha \mathbf{S} \mathbf{A}_\beta(T) \vec{M}(0) + \mathbf{R}_\alpha \mathbf{S} \vec{b}(T) \,.
\end{aligned} \tag{1.4}
$$

By definition, the *steady-state* magnetization has equilibrated to a periodic pattern between RF pulses, *i.e.*, following each pulse, $\vec{M}^+ := \vec{M}((k+1)T_R) \approx \vec{M}(kT_R)$, as $k \to \infty$ is given as:

$$
\begin{aligned}
\vec{M}^+ &= \mathbf{R}_\alpha \mathbf{S} \mathbf{A}_\beta(T) \vec{M}^+ + \mathbf{R}_\alpha \mathbf{S} \vec{b}(T) \\
&= (\mathbf{I} - \mathbf{R}_\alpha \mathbf{S} \mathbf{A}_\beta(T))^{-1} \mathbf{R}_\alpha \mathbf{S} \vec{b}(T) \\
&= \frac{1}{1 - E_1(T)\cos\alpha} \begin{bmatrix} 0 \\ M_0 \sin\alpha(1 - E_1(T)) \\ M_0 \cos\alpha(1 - E_1(T)) \end{bmatrix} .
\end{aligned} \tag{1.5}
$$

Observe that with ideal spoiling, the steady-state magnetization does not depend on off-resonance effects $\beta(T)$. Furthermore, the steady-state magnetization immediately following each RF excitation is independent of T_2, thus simplifying T_1 estimation. We approximate $E_1(T) \approx E_1(T_R)$ and $E_2(T) \approx E_2(T_R)$ because a negligible amount of relaxation occurs over duration τ of the RF pulse. Then, the transverse magnetization immediately following the RF pulse $M_{xy}^+(kT_R)$, is given as

$$
M_{xy}^+(kT_R) := M_x^+ + i M_y^+ \approx \frac{i M_0 (\sin\alpha)(1 - E_1(T_R))}{1 - E_1(T_R)\cos\alpha} \,. \tag{1.6}
$$

Finally, the observed SPGR signal is collected at echo time T_E following RF excitation. We must therefore account for both T_2 relaxation, as well as phase accrual $\theta(t) \triangleq \int_0^t \Delta\omega(t')\,\mathrm{d}t' \approx t\Delta\omega_f$ of individual isochromats due to off-resonance effects. If the spins are assumed to dephase with broadening distribution $f(\Delta\omega_f) \sim \mathrm{Cauchy}(\Delta\bar{\omega}_f, R_2')$, the noiseless steady-state SPGR signal model (denoted S_{SPGR}) is computed by integrating over the many spins within each voxel:

$$
\begin{aligned}
S_{\mathrm{SPGR}} &:= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_0^{2\pi} M_{xy}^+(kT_R + T_E) f(\Delta\omega_f)\,\mathrm{d}\beta\,\mathrm{d}\Delta\omega_f \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_0^{2\pi} M_{xy}^+(kT_R) E_2(T_E) e^{i\theta(T_E)} f(\Delta\omega_f)\,\mathrm{d}\beta\,\mathrm{d}\Delta\omega_f \\
&\approx M_{xy}^+(kT_R) E_2(T_E) \int_{-\infty}^{\infty} e^{i\Delta\omega_f T_E} f(\Delta\omega_f)\,\mathrm{d}\Delta\omega_f \\
&= M_{xy}^+(kT_R) E_2(T_E) e^{i\Delta\bar{\omega}_f T_E - R_2' T_E} \\
&= M_0 e^{-T_E/T_2^*} e^{i(\Delta\bar{\omega}_f T_E + \frac{\pi}{2})} \frac{\sin\alpha(1 - E_1(T_R))}{1 - E_1(T_R)\cos\alpha} \,.
\end{aligned} \tag{1.7}
$$

In this work, we seek to accurately estimate T_1 maps from noise-corrupted SPGR data collected at $l = 1, \ldots, L$ flip angles, modeled as

$$
y_l = S_{\mathrm{SPGR}}(M_0, T_1, T_2^*; \alpha_l) + \epsilon \,, \tag{1.8}
$$

where $\epsilon \sim \mathbb{CN}(0, \sigma^2)$ is complex, white Gaussian noise [6, 7]. We observe however that nuisance parameters M_0 and T_2^* complicate this goal. Though other user-specified parameters (T_R, T_E, α_l) are known, these nuisance parameter are unknown, and thus must also be estimated.

3

1.2.2 Effects of Imperfect Spoiling
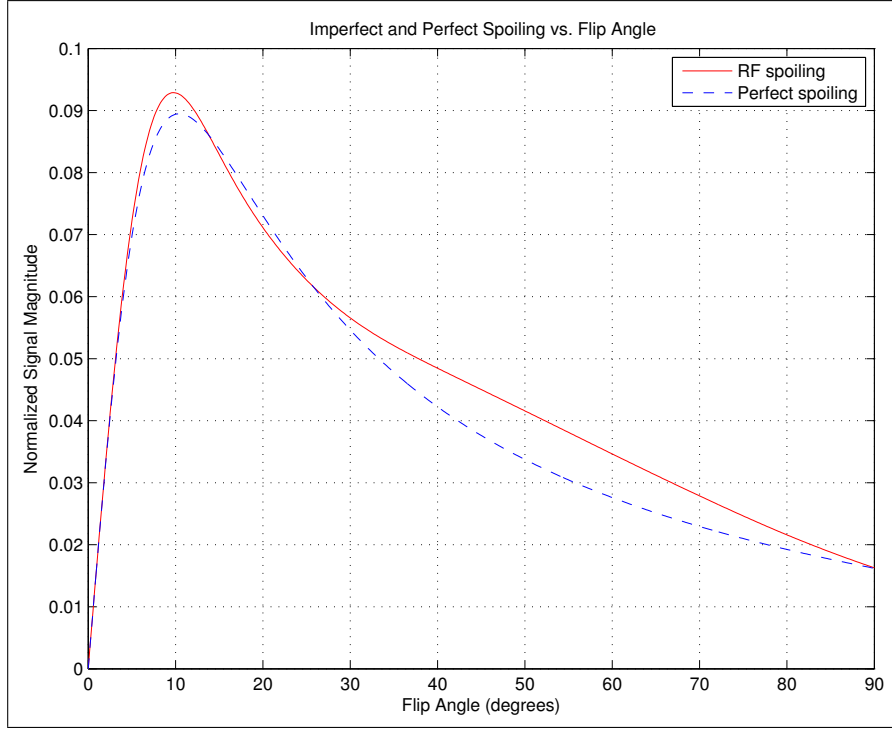
Figure 1.1: Bloch simulations of a single isochromat, with $T_1 = 600\text{ms}$, $T_2 = 100\text{ms}$, $T_R = 10\text{ms}$, $T_E = 2\text{ms}$. RF spoiling with $\psi = 117°$ linear phase cycling yields a signal close to that of perfect spoiling for low flip angles, but deviates from the ideal model for larger angles.

The simplicity of the matrix inversion in Equation (1.5) arose from our unrealistic assumption of perfect spoiling. In practice, it is quite difficult to completely remove all transverse magnetization before subsequent RF excitations. Zur proposed one reasonably effective method of spoiling by incorporating an additional transverse rotation $\phi(k) = k\psi + \psi_0$ into each RF excitation [5]. For certain well-selected linear phase increments ψ, selection of the (physically realizable) spoiling operator

$$\mathbf{S}(k) = \begin{bmatrix} \cos\phi(k) & \sin\phi(k) & 0 \\ -\sin\phi(k) & \cos\phi(k) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1.9}$$

causes the transverse phase to destructively interfere in steady state. Zur found that $\psi = 117°$ achieves good spoiling through this method for low α; this recommendation for RF spoiling is used throughout this chapter. It should be noted, however, that this spoiling mechanism is incomplete: we observe through Bloch simulation in Figure 1.1 that some model mismatch persists.[1]

---

[1] Simulation borrowed from `http://mrsrl.stanford.edu/~brian/bloch/`.

1.2.3 Conventional Method-of-Moments T_1 Estimation

Conventional approaches [1, 2] observe that if nuisance parameters M_0, T_2^* are collected into a single T_2^*-compensated proton density unknown,

$$M_0^* := iM_0 \exp\{-T_E/T_2^*\}\,, \tag{1.10}$$

then Equation (1.8) can be recast into a linear form:

$$\underbrace{\left[\frac{y_l}{\sin \alpha_l}\right]}_{\mathbf{y}} = \underbrace{\left[\frac{y_l}{\tan \alpha_l}\right]}_{\mathbf{x}} \underbrace{E_1(T_R)}_{m} + [\mathbf{1}]\underbrace{M_0^*(1 - E_1(T_R))}_{b} + \underbrace{[\epsilon_l']}_{\epsilon}\,, \tag{1.11}$$

where $\mathbf{y}$ and $\mathbf{x}$ are length-L data vectors, l again indexes $1, \ldots, L$ flip angles, $\mathbf{1}$ is a length-L vector of ones, and ϵ is a modified, length-L (flip angle dependent) noise vector. Then, slope m and intercept b can be estimated via linear regression from as few as two flip angles, on a per-voxel basis. If the noise vector ϵ is neglected, T_1 and M_0^* can subsequently be estimated as

$$\widehat{T}_1 = -T_R/\log \widehat{m}\,,$$
$$\widehat{M_0^*} = \widehat{b}/(1 - \widehat{m})\,. \tag{1.12}$$

This "log-linearized least squares" estimator yields noise-amplified T_1 maps, especially for low flip angles. Observe that if the image-domain data noise is distributed as $\epsilon \sim \mathbb{CN}(0, \sigma^2)$, then the linearization in Equation (1.11) will rescale the noise variance to be distributed as

$$\epsilon_l' \sim \mathbb{CN}\left(0, \left[\frac{1 - E_1(T_R)\cos \alpha_l}{\sin \alpha_l}\right]^2 \sigma^2\right)\,. \tag{1.13}$$

Particularly for small values of $E_1(T_R)$ (high T_R/T_1, which usually corresponds to regions of interest), this scaling factor is greater for small flip angles. To reduce noise scaling through linearization, larger flip angles are thus preferable. However, as seen from Figure 1.1, it is in fact preferable to scan at lower flip angles because of greater SNR, and closer signal model alignment between RF and ideal spoiling. These conflicting effects create tradeoffs in flip angle choice, which has caused researchers to expend significant effort in scan parameter selection [8, 9]. Such tradeoffs can be avoided if one instead directly approached the parameter estimation problem with a nonlinear signal model.

1.2.4 Penalized-Likelihood T_1 Estimation

We are interested in estimating spin-lattice relaxation parameter T_1 from steady-state incoherent MRI sequences. Towards this end, we must jointly estimate T_2^*-compensating proton density M_0^* as a nuisance parameter. In steady-state, the system model can be written as:

$$y_l = \frac{M_0^* \sin \alpha_l (1 - E_1)}{1 - E_1 \cos \alpha_l} + \epsilon =: f(M_0^*, E_1; \alpha_l) + \epsilon\,, \tag{1.14}$$

Algorithm 1 MBET1: Model-Based Estimation of T_1 from SPGR Data

---

Input: Data $\mathbf{y}_1, \ldots, \mathbf{y}_L$; Initial guesses $\mathbf{x}_1^{(0)}, \mathbf{x}_2^{(0)}$; Regularization parameters β_1, β_2
Output: Regularized parameter maps $\mathbf{M}_0^*, \mathbf{T}_1$
for $n = 1, \ldots, T$ **do**
    Update system matrices $\mathbf{A}_l^{(n)}$ with Equation (1.17)
    Update $\mathbf{D}_1^{(n)}$ with Equation (1.19)
    repeat
        Update $\mathbf{x}_1$ with Equation (1.18)
    until $\mathbf{x}_1$ convergence criterion satisfied
    Update $\mathbf{D}_2^{(n)}$ with Equation (1.28)
    repeat
        Update $\mathbf{x}_2$ with Equation (1.27)
    until $\mathbf{x}_2$ convergence criterion satisfied
end for
Embed $\mathbf{x}_1^{(T)}, \mathbf{x}_2^{(T)}$ into $\mathbf{M}_0^*, \mathbf{T}_1$

---

where y_l is the SPGR image data for the lth flip angle α_l; $E_1 := E_1(T_R) = e^{-T_R/T_1}$; $\epsilon \sim \mathbb{CN}\{0, \sigma^2\}$; and bold versions of the corresponding variable names indicate vectorized, full data sets. This problem can be investigated as a joint optimization over the cost function $\Psi(\mathbf{x}_1, \mathbf{x}_2)$:

$$\{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2\} = \underset{\{\mathbf{x}_1, \mathbf{x}_2 \in (0,1]\}}{\arg\min} \Psi(\mathbf{x}_1, \mathbf{x}_2), \text{ where}$$

$$\Psi(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{2} \sum_{l=1}^{L} ||\mathbf{y}_l - \mathbf{f}(\mathbf{x}_1, \mathbf{x}_2; \alpha_l)||_2^2 + \beta_1 R(\mathbf{x}_1) + \beta_2 R(\mathbf{x}_2); \tag{1.15}$$

$\mathbf{x}_1 := M_0^* \in \mathbb{C}^N$ and $\mathbf{x}_2 := E_1 \in \mathbb{R}^N$ are introduced to standardize notation; N is the number of voxels; $\mathbf{y}_l \in \mathbb{C}^N$; and β_1, β_2 are parameters used to control the influence of the (possibly non-quadratic) regularizer $R : \mathbb{C}^N \mapsto \mathbb{R}$. Note the slight recycling of notation by which we have extended the scalar function $f : \mathbb{C} \times \mathbb{R} \times \mathbb{R} \mapsto \mathbb{C}$, to a vectorized function $\mathbf{f} : \mathbb{C}^N \times \mathbb{R}^N \times \mathbb{R}^N \mapsto \mathbb{C}^N$.

The system model is linear in $\mathbf{x}_1$ but nonlinear in $\mathbf{x}_2$. We exploit this partially linear structure and derive a simple algorithm for solving the above problem by taking an alternating minimization approach, denoted MBET1. MBET1 is guaranteed to monotonically decrease the non-convex cost $\Psi(\mathbf{x}_1, \mathbf{x}_2)$ to a local minimum. In the following section, variable updates for the alternating minimization are presented in detail. A summary of MBET1 is provided in Algorithm 1.

1.3   Minimization Techniques

1.3.1   Updating the M_0^* Estimate

In the following, we will use i to index inner iterations, and $n = 1, \ldots, T$ to index outer iterations. To avoid unnecessarily cumbersome notation, we restart inner iteration indexing at $i = 1$ at the

beginning of each variable update. For the $\mathbf{x}_1$ update, we hold $\mathbf{x}_2^{(n)}$ fixed from the nth iteration and minimize a simplified, but equivalent objective function:

$$\hat{\mathbf{x}}_1^{(n+1)} = \underset{\{\mathbf{x}_1\}}{\arg\min} \ \Psi_1^{(n)}(\mathbf{x}_1), \ \text{where}$$

$$\Psi_1^{(n)}(\mathbf{x}_1) = \frac{1}{2}\sum_{l=1}^{L}||\mathbf{y}_l - \mathbf{A}_l^{(n)}\mathbf{x}_1||_2^2 + \beta_1 R(\mathbf{x}_1), \tag{1.16}$$

and the system matrix $\mathbf{A}_l^{(n)} \in \mathbb{C}^{N \times N}$ is diagonal, with entries $a_l^{(n)}$ given as

$$a_l^{(n)} = \frac{(\sin \alpha_l)(1 - x_2^{(n)})}{1 - x_2^{(n)}\cos \alpha_l}. \tag{1.17}$$

Unconstrained optimization of cost functions of the form (1.16) has been studied extensively. As an example method, a preconditioned gradient descent update suffices:

$$\hat{\mathbf{x}}_1^{(i+1)} = \hat{\mathbf{x}}_1^{(i)} - s_1[\mathbf{D}_1^{(n)}]^{-1}\left[\left(\sum_{l=1}^{L}(\mathbf{A}_l^{(n)})'(\mathbf{A}_l^{(n)}\hat{\mathbf{x}}_1^{(i)} - \mathbf{y}_l)\right) + \beta_1 \nabla R(\mathbf{x}_1^{(i)})\right], \ \text{where} \tag{1.18}$$

$$\mathbf{D}_1^{(n)} = \left(\sum_{l=1}^{L}(\mathbf{A}_l^{(n)})'\mathbf{A}_l^{(n)}\right) + \beta_1 \operatorname{maxeig}\{\nabla^2 R(\mathbf{x}_1^{(n)})\}\mathbf{I}_N, \tag{1.19}$$

and s_1 is an appropriately-selected step size, often selected using line search techniques. We repeat (1.18) until a convergence criterion is met, and assign the most recent update to $\hat{\mathbf{x}}_1^{(n+1)}$.

1.3.2 Updating the T_1 Estimate

For the $\mathbf{x}_2$ update, we hold $\mathbf{x}_1^{(n+1)}$ fixed at its most recent value and minimize the cost

$$\hat{\mathbf{x}}_2^{(n+1)} = \underset{\{\mathbf{x}_2 \in (0,1]\}}{\arg\min} \ \Psi_2^{(n)}(\mathbf{x}_2), \ \text{where}$$

$$\Psi_2^{(n)}(\mathbf{x}_2) = \frac{1}{2}\sum_{l=1}^{L}||\mathbf{y}_l - \mathbf{f}(\mathbf{x}_2; \mathbf{x}_1^{(n+1)}, \alpha_l)||_2^2 + \beta_2 R(\mathbf{x}_2). \tag{1.20}$$

A preconditioned gradient descent (PGD) method could be used to monotonically decrease this objective function if a preconditioner $\mathbf{P}^{(n)} = (\mathbf{D}_2^{(n)})^{-1}$ were selected such that $\mathbf{D}_2^{(n)}$ bounded the curvature of the cost function at each iteration:

$$\nabla^2\Psi_2^{(n)}(\mathbf{x}_2) \preceq \mathbf{D}_2^{(n)}, \ \forall \mathbf{x}_2. \tag{1.21}$$

Taking vector derivatives, the Hessian is given as

$$\nabla^2\Psi_2^{(n)}(\mathbf{x}_2) = \sum_{l=1}^{L}\nabla^2\Phi_{2,l}^{(n)}(\mathbf{x}_2) + \beta_2\nabla^2 R(\mathbf{x}_2), \ \text{where}$$

$$\nabla^2\Phi_{2,l}^{(n)}(\mathbf{x}_2) = \nabla^2\mathbf{f}(\mathbf{x}_2; \mathbf{x}_1^{(n+1)}, \alpha_l)\left[\mathbf{f}(\mathbf{x}_2; \mathbf{x}_1^{(n+1)}, \alpha_l) - \mathbf{y}_l\right] + \left[\nabla\mathbf{f}(\mathbf{x}_2; \mathbf{x}_1^{(n+1)}, \alpha_l)\right]^2. \tag{1.22}$$

7

Fortunately, $\nabla^2 \Phi_{2,l}^{(n)}(\mathbf{x}_2)$ is fully diagonal, and can therefore be considered on a voxel-by-voxel basis. Denoting $\ddot{\phi}_{2,l}^{(n)}(x_2) := \frac{\partial^2}{(\partial x_2)^2} \phi_{2,l}^{(n)}(x_2)$ as the jth entry on the diagonal of $\nabla^2 \Phi_{2,l}^{(n)}(\mathbf{x}_2)$,

$$\ddot{\phi}_{2,l}^{(n)}(x_2) = f''(x_2; x_1^{(n+1)}, \alpha_l)\left(f(x_2; x_1^{(n+1)}, \alpha_l) - y_l\right) + \left(f'(x_2; x_1^{(n+1)}, \alpha_l)\right)^2. \tag{1.23}$$

We next bound $\ddot{\phi}_{2,l}^{(n)}(x_2)$ to a value independent of x_2 by repeatedly exploiting the box constraint $0 < x_2 \le 1$ and assuming $0 < \alpha_l \le \frac{\pi}{2}$. For sake of readability, we omit flip angle subscripts (l) and iteration superscripts $(n+1)$ until the final step.

Recall that the scalar function (1.14) and its derivatives are given as

$$f(x_2; x_1, \alpha) := \frac{x_1 \sin\alpha(1 - x_2)}{1 - x_2\cos\alpha}$$

$$f'(x_2; x_1, \alpha) = \frac{-x_1 \sin\alpha(1 - \cos\alpha)}{(1 - x_2\cos\alpha)^2}$$

$$f''(x_2; x_1, \alpha) = \frac{-2x_1 \sin\alpha\cos\alpha(1 - \cos\alpha)}{(1 - x_2\cos\alpha)^3}. \tag{1.24}$$

Applying these derivatives, we bound $\ddot{\phi}_{2,l}^{(n)}(x_2)$ as follows:

$$
\begin{aligned}
\ddot{\phi}(x_2) &= f''(x_2)\Big[f(x_2) - y\Big] + \Big[f'(x_2)\Big]^2 \\
&= \frac{-2x_1^2 \sin^2\alpha\cos\alpha(1 - x_2)(1 - \cos\alpha) + x_1^2 \sin^2(\alpha)(1 - \cos\alpha)^2}{(1 - x_2\cos\alpha)^4} + \frac{2x_1 y \sin\alpha\cos\alpha(1 - \cos\alpha)}{(1 - x_2\cos\alpha)^3} \\
&\le \frac{-2x_1^2 \sin^2\alpha\cos\alpha(1 - x_2)(1 - \cos\alpha) + x_1^2 \sin^2(\alpha)(1 - \cos\alpha)^2}{(1 - \cos\alpha)^4} + \frac{2x_1 y \sin\alpha\cos\alpha(1 - \cos\alpha)}{(1 - \cos\alpha)^3} \\
&= \frac{-2x_1^2 \sin^2\alpha\cos\alpha(1 - x_2) + x_1^2 \sin^2(\alpha)(1 - \cos\alpha)}{(1 - \cos\alpha)^3} + \frac{2x_1 y \sin\alpha\cos\alpha}{(1 - \cos\alpha)^2} \\
&= \frac{x_1^2 \sin^2\alpha(1 - \cos\alpha - 2\cos\alpha(1 - x_2))}{(1 - \cos\alpha)^3} + \frac{x_1 y \sin 2\alpha}{(1 - \cos\alpha)^2} \\
&\le \frac{x_1^2 \sin^2\alpha(1 - \cos\alpha)}{(1 - \cos\alpha)^3} + \frac{x_1 y \sin 2\alpha}{(1 - \cos\alpha)^2} \\
&= \frac{(x_1^{(n+1)})^2 \sin^2\alpha_l + x_1^{(n+1)} y_l \sin 2\alpha_l}{(1 - \cos\alpha_l)^2} \quad := d_{2,l}^{(n)}. \tag{1.25}
\end{aligned}
$$

Vectorizing this inequality across all voxels, it is evident that

$$
\begin{aligned}
\nabla^2 \Phi_{2,l}^{(n)}(\mathbf{x}_2) &\le \operatorname{diag}\{d_{2,l}^{(n)}\} \\
\nabla^2 \Psi_2^{(n)}(\mathbf{x}_2) &\le \sum_{l=1}^{L} \operatorname{diag}\{d_{2,l}^{(n)}\} + \beta_2 \nabla^2 R(\mathbf{x}_2) \\
&\le \sum_{l=1}^{L} \operatorname{diag}\{d_{2,l}^{(n)}\} + \beta_2 \operatorname{maxeig}\{\nabla^2 R(\mathbf{x}_2^{(n)})\}\mathbf{I}_N \quad := \mathbf{D}_2^{(n)}. \tag{1.26}
\end{aligned}
$$

8

This diagonal choice of $\mathbf{D}_2^{(n)}$ allows for easy computation of the preconditioner, $\mathbf{P}^{(n)} = (\mathbf{D}_2^{(n)})^{-1}$. Setting step size $0 \leq s_2 < 2$ provides the following preconditioned gradient projection update:

$$\hat{\mathbf{x}}_2^{(i+1)} = \mathcal{P}_{(0,1]}\left\{\mathbf{x}_2^{(i)} - s_2(\mathbf{D}_2^{(n)})^{-1}\nabla\Psi_2^{(n)}(\mathbf{x}_2^{(i)})\right\}, \text{ where} \tag{1.27}$$

$$\mathbf{D}_2^{(n)} = \sum_{l=1}^{L}\mathrm{diag}\left\{d_{2,l}^{(n)}\right\} + \beta_2\,\mathrm{maxeig}\left\{\nabla^2 R(\hat{\mathbf{x}}_2^{(n)})\right\}\mathbf{I}_N, \text{ and} \tag{1.28}$$

$$\nabla\Psi_2^{(n)}(\mathbf{x}_2) = \sum_{l=1}^{L}\nabla\mathbf{f}(\mathbf{x}_2; \mathbf{x}_1^{(n+1)}, \alpha_l)\left[\mathbf{f}(\mathbf{x}_2; \mathbf{x}_1^{(n+1)}, \alpha_l) - \mathbf{y}_l\right] + \beta_2\nabla R(\mathbf{x}_2).$$

1.4 Experimentation

1.4.1 Results from Synthetic Data

We evaluate our method with data synthesized from the 95th slice of the BrainWeb digital phantom [10]. Using Equation (1.7), we combine ground truth M_0, T_1, and T_2^* maps to produce uncorrupted, 217×181 resolution SPGR data, at "ideal" flip angles $5°$ and $30°$ [8], $T_R = 20$ms, and $T_E = 5$ms. We corrupt the true data with complex, white Gaussian noise, to yield 40dB SNR raw data. We strive to recover precise T_1 maps from this raw data, with no prior knowledge of M_0 or T_2^*.

We first reconstruct the data through log-linearization and least-squares fitting [1, 2], and then use this estimate to initialize MBET1. We select R to be an edge-preserving regularizer,

$$R(\mathbf{x}) = \sum_{k=1}^{K}\psi_k\left([\mathbf{C}\mathbf{x}]_k\right), \tag{1.29}$$

where ψ_k is a hyperbolic potential function, $\mathbf{C}$ is a 2D spatial finite differencing matrix, and k indexes a total of K differencing operations. We choose regularization parameters $\beta_1 = 2^{-2}$ and $\beta_2 = 2^9$ heuristically, though automatic selection methods [11] can be used in practice. On a 3.5GHz desktop with 32GB RAM, with $T = 30$ outer iterations, reconstructions take 50s per slice.

Figures 1.2 and 1.3 compare MBET1 with conventional T_1 and M_0^* estimation, respectively, both before and after median filtering. Absolute difference images are paired with error measures in grey matter (GM) and white matter (WM) regions of interest. T_1 errors are reported in root mean-squared error (RMSE), while M_0^* errors are reported in normalized RMSE (NRMSE).

At typical low flip angles, the least-squares T_1 estimate is prone to noise amplification through linearization of the signal model. This is a major contributing factor to its relatively high RMSE. Median filtering reduces RMSE to some extent, but not as substantially as MBET1.

For completeness, nuisance parameter M_0^* maps are also shown. Here, the noise-amplification due to linearizing the signal model is offset by the manipulations in Eq. (1.12) used to recover $\widehat{M_0^*}$. Thus, both methods are expected to exhibit similar noise properties. Indeed, MBET1 does not achieve an appreciably different NRMSE than median filtering. Though both outperform unfiltered least-squares because both employ edge-preserving regularization, the MBET1 estimate in fact exhibits a

| (a) Least-Squares | (b) Median-filtered | (c) MBET1 | (d) Ground Truth |

| (e) RMSE = (136, 78)ms | (f) RMSE = (99, 53)ms | (g) RMSE = (63, 37)ms |
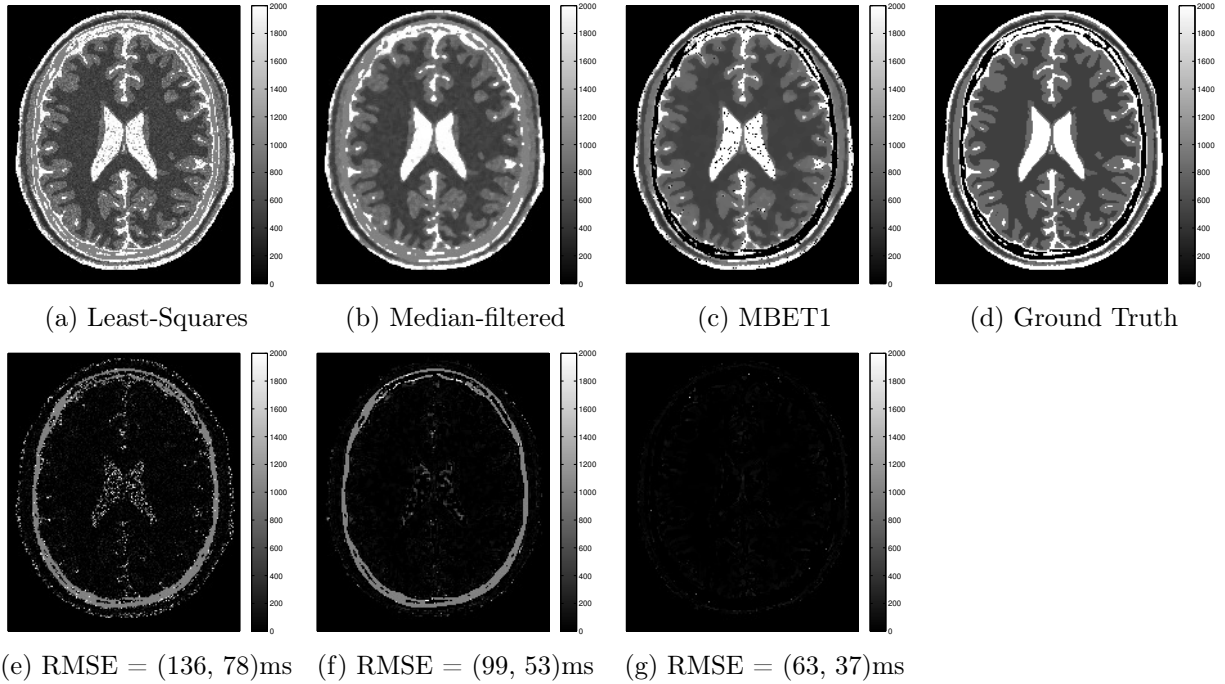
Figure 1.2: T_1 maps (a, b, c) and corresponding absolute-difference images (e, f, g), with respect to the ground truth (d). Linearized least-squares (a) and median-filtered (b) images reconstructed from 40dB SNR SPGR data at "optimal" flip angle combination $\alpha = 5, 30°$. MBET1 image (c) reconstructed using median-filtered initialization. Tissue-specific errors are spatially averaged over (GM, WM) voxels of interest, and are reported in RMSE.

slight increase in NRMSE in regions of interest, which is subsequently propagated as speckle noise in the T_1 reconstruction. Note, however, that we are primarily interested in T_1 mapping quality.

Tissue	Least-Squares	Median-Filtered	MBET1	Ground Truth
GM	853 ± 135	842 ± 99	827 ± 63	833 ± 0
WM	511 ± 77	522 ± 48	516 ± 29	500 ± 0

Table 1.1: T_1 estimates and standard deviations, spatially averaged over GM and WM regions of interest. Scan protocols match those in Figure 1.2. All values reported in (ms). MBET1 exhibits clear increases in precision over least-squares estimates, while maintaining similar levels of accuracy.

Table 1.1 quantitatively highlights that MBET1 achieves precision increases over least squares estimators, with or without median filtering, particularly in GM and WM regions of interest. These precision gains are due to (1) omission of noise amplification through linearization of the signal model, and (2) usage of edge-preserving regularization to encourage piecewise-smooth M_0^* and T_1 maps. RMSE reductions are sustained over wide range of SNR values (results omitted).

By contrast, we observe that MBET1 does not achieve considerably better accuracy, because MBET1 does not correct for any bias in the signal model that conventional methods neglect. However, by avoiding noise amplification through linearization of the signal model, we are free to collect data at lower flip angles, at which RF spoiling comes as close to perfect spoiling as possible. With

(a) Least-Squares   (b) Median-filtered   (c) MBET1   (d) Ground Truth

(e) NRMSE = (13.6, 13.8)%   (f) NRMSE = (6.1, 6.0)%   (g) NRMSE = (8.8, 8.0)%
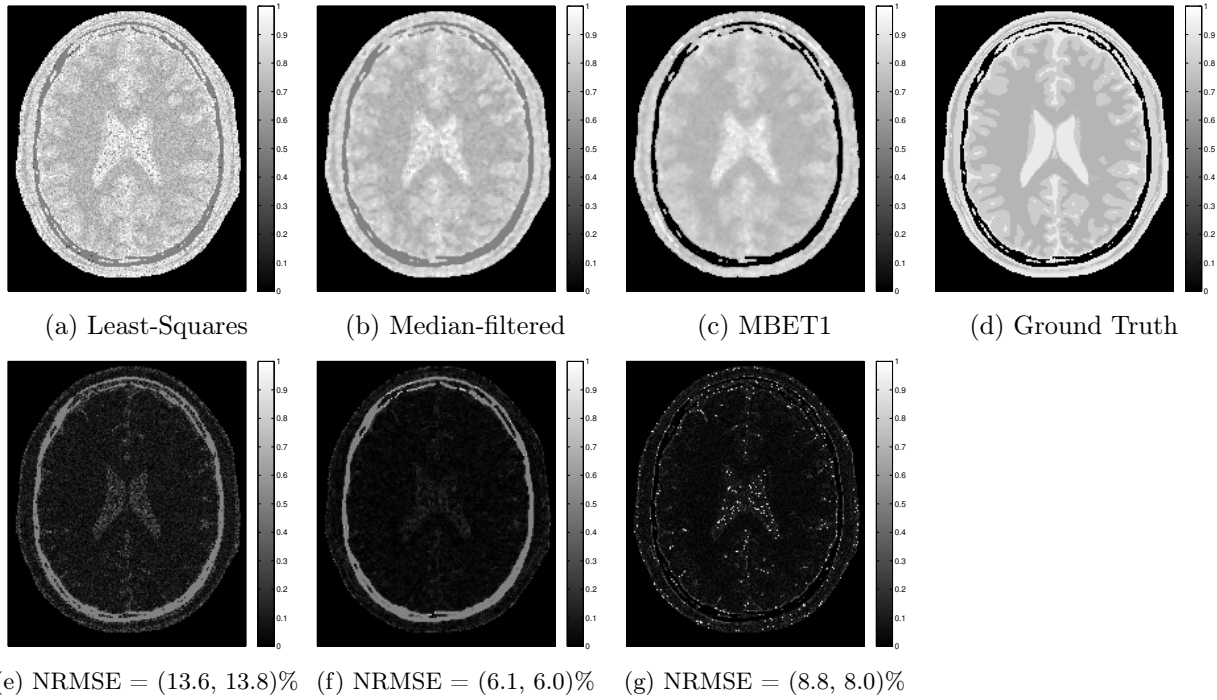
Figure 1.3: Nuisance parameter M_0^* maps (a, b, c) and corresponding absolute-difference images (e, f, g), with respect to ground truth (d). All images correspond to T_1 maps in Figure 1.2. Tissue-specific errors are spatially averaged over (GM, WM) voxels of interest, and are reported in NRMSE.

this choice, the image data would better match the ideal signal model, so one could expect to reconstruct more accurate T_1 maps. Scan selection for parameter estimation is a rich topic in its own right, and is a major subject of discussion in Chapter 3.

1.4.2   Results from *In Vivo* Data

(a) M_0^*, Least-Squares   (b) M_0^*, MBET1   (c) T_1, Least-Squares   (d) T_1, MBET1
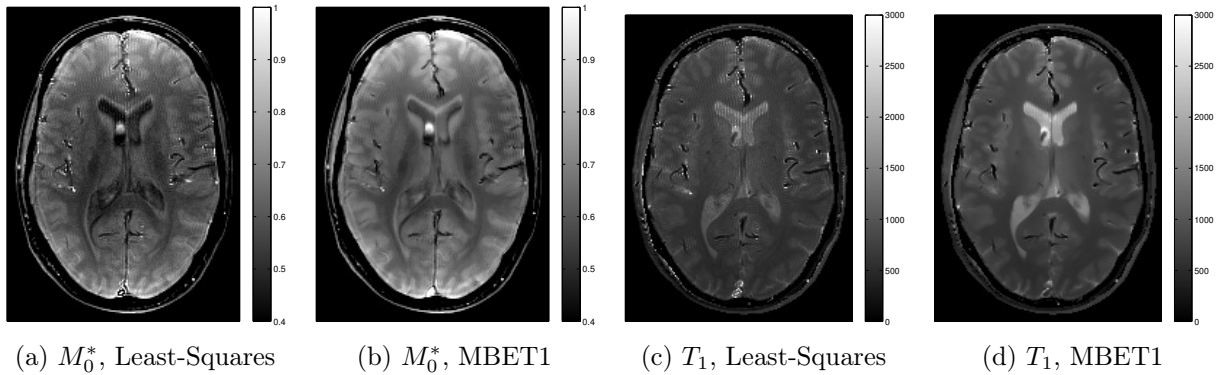
Figure 1.4: Least-squares and MBET1 reconstructions of SPGR data taken at $\alpha = 5, 10, 20, 30, 45$ degrees. Regularized estimation yields piecewise-smooth images. This translates to an increase in parameter mapping precision over conventional estimators.

Fig. 1.4 provides *in vivo* results from an 8-channel head coil in a 3T GE MRI scanner. We repeated an SPGR sequence at 5, 10, 20, 30, and 45 degree flip angles, with $T_R = 19.8$ms, $T_E = 5.1$ms. We acquired 10 axial slices at 5mm thickness, with a 24cm field of view (256×256 resolution). We reconstructed the data with both conventional and MBET1 methods.

We selected several WM and GM regions of interest for comparing parameter estimates. From the conventional method, T_1 was measured as 990 ± 105ms and 734 ± 54ms in grey and white matter, respectively. With the proposed method, T_1 was measured as 1181 ± 29ms and 802 ± 17ms in grey and white matter, respectively. These numbers highlight that the proposed method achieves higher precision *within* tissue types, while preserving high contrast *across* tissue types.

1.5   Conclusions and Future Work

We have presented MBET1, a model-based approach for T_1 relaxometry from SPGR sequences. As compared to conventional least-squares methods, we observe that MBET1 improves precision in T_1 estimation through omission of noise-amplifying linearization of the signal model and inclusion of edge-preserving regularization. We have shown that MBET1 improves T_1 mapping precision both for synthetic and *in vivo* data, at a wide range of noise thresholds and flip angle combinations.

Much of the simplicity in the formulation of MBET1 for T_1 estimation is derived from the simplicity of the ideal SPGR signal model, which manages to remove all T_2 dependence and inter-voxel off-resonance effects through spoiling. To our knowledge, a fast, steady-state pulse sequence that analogously isolates T_2 effects from nuisance parameters has not yet been developed. Indeed, certain steady-state signal models recently proposed for T_2 relaxometry tend to have complicated dependencies on numerous nuisance parameters, including M_0, T_1, and T_2^*. In the subsequent chapter, we find that MBET1 can serve as an intermediary step to facilitate T_2 estimation.

Chapter 2

Model-Based Estimation of T_2 Maps (MBET2) from DESS Sequences

2.1   Introduction

In this chapter, we take primary interest in estimating spin-spin relaxation parameter T_2, though we also obtain T_1 estimates through preliminary MBET1. Rapid, accurate T_2 quantification is of interest to the MRI community because it shows promise as a biomarker for subtle changes in pathology [12]. Indeed, T_2 may be useful for monitoring the progression of poorly-understood neurological and/or autoimmune disorders, such as Parkinson's disease, epilepsy, or multiple sclerosis.

Classical single-echo spin echo sequences yield simple methods for accurately estimating T_2, but require multiple scans at variable echo times, leading to undesirably long acquisitions. Multi-echo SE sequences [13] can eliminate the multiple-scan requirement, but are prone to systematic error, due to stimulated-echo contributions that can arise due to transmit field B_1 inhomogeneities [14].

Steady state free precession (SSFP) pulse sequences [15, 16] are, by contrast, much faster than spin-echo sequences due to a short repetition time T_R. SSFP techniques are particularly well-suited for relaxometry because the observed signals are highly sensitive to T_1 and T_2 variation. However, short T_R times also cause these signals to be complex functions of both desired and nuisance parameters, thereby complicating quantification. Furthermore, some such methods [8] still require scan repetition at multiple nutation angles, though individual scans are now considerably shorter.

Recently, the Dual-Echo Steady-State (DESS) sequence [17] was proposed as a promising SSFP imaging technique for T_2 estimation [18]. Since it contains two readouts per excitation, the DESS sequence can be used to reduce scan repetition requirements by collecting twice as much data per scan. As with other SSFP methods, the resulting signals are complex functions of T_1, T_2, and other parameters [19, 20]. To reduce unwanted dependencies, conventional approaches [14, 18] observe that in the large nutation-angle limit, the ratio between the second (SSFP-Echo, S^-) and first (SSFP-FID, S^+) signals is independent of T_1, yielding a simple method-of-moments T_2 estimator. Unfortunately, the low SNR at this limit requires usage of lower nutation angles in practice, thereby introducing bias into the T_2 estimate.

In this work, we mitigate this bias by approaching the T_2 estimation problem through a model-based framework, denoted MBET2 [21]. MBET2 also avoids noise amplification issues inherent to signal division. We first jointly estimate T_2^*-compensated proton density M_0^* and spin-lattice relaxation parameter T_1 maps through regularized reconstruction of DESPOT1 data from as few as two nutation angles [3, 22]. We then iteratively reconstruct T_2 maps using these nuisance parameter estimates and an initialization with the conventional DESS estimate, encouraging piecewise-smoothness through edge-preserving regularization. The resultant T_2 maps maintain significant improvements in both mapping accuracy and precision, over a wide range of nutation angles and repetition times.

Chapter 2 has a structure similar to Chapter 1, as it builds on many of the ideas presented previously. In Section 2, we derive in detail the DESS signal model from the Bloch equations. In Section 3, we describe the shortcomings of conventional T_2 estimation approaches, and introduce a penalized-likelihood optimization problem that addresses these shortcomings. In Section 4, we discuss a method for solving this problem. In Section 5, we describe both synthetic and *in vivo* experiments. In Section 6, we briefly conclude our findings and describe why joint T_1, T_2 estimation may be preferable to the sequential approach presented here.

2.2   Signal Model

Beginning with the Bloch equations [4], we develop a model describing the steady-state signal received from an DESS pulse sequence. Portions of this formulation are presented in earlier works [15, 19, 20], though some refinements given here are, to our knowledge, not explicitly published elsewhere. As in Chapter 1, we assume that all excitations are of the same flip angle, and orient our transverse axes such that rotations occur around the i-axis.

2.2.1   The Currently Popular Model

Let $\vec{M}(0)$ denote the magnetization immediately following an RF pulse. Until the start of the next RF pulse, each voxel's magnetization $\vec{M}(t)$ evolves according to the Bloch equations:

$$\vec{M}(t) = \begin{bmatrix} E_2(t) & 0 & 0 \\ 0 & E_2(t) & 0 \\ 0 & 0 & E_1(t) \end{bmatrix} \begin{bmatrix} \cos\beta(t) & \sin\beta(t) & 0 \\ -\sin\beta(t) & \cos\beta(t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \vec{M}(0) + M_0\left(1 - E_1(t)\right)\vec{k}, \qquad (2.1)$$

where the longitudinal and transverse relaxation parameters are given by $E_1(t) = e^{-t/T_1}$ and $E_2(t) = e^{-t/T_2}$, respectively, and the cumulative phase evolution due to all off-resonance effects (gradient fields, variations in susceptibility, B_0 inhomogeneity, chemical shift, etc.) are summarized as

$$\beta(t) = \int_0^t \Delta\omega(t')\,\mathrm{d}t', \qquad (2.2)$$

a phase rotation in the transverse plane. Note that off-resonance effects $\Delta\omega(t)$ are themselves usually time-varying effects, particularly with the use of gradients and spoiler fields.

Specifically, we are interested in the magnetization after the free precession interval, of duration T. If we define $\beta \triangleq \beta(T) = \int_0^T \Delta\omega(t')\,dt'$, and

$$\mathbf{R}_\beta \triangleq \begin{bmatrix} E_2(T) & 0 & 0 \\ 0 & E_2(T) & 0 \\ 0 & 0 & E_1(T) \end{bmatrix} \begin{bmatrix} \cos\beta & \sin\beta & 0 \\ -\sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} E_2(T)\cos\beta & E_2(T)\sin\beta & 0 \\ -E_2(T)\sin\beta & E_2(T)\cos\beta & 0 \\ 0 & 0 & E_1(T) \end{bmatrix},$$
(2.3)

then the magnetization at time T is given as $\vec{M}(T) = \mathbf{R}_\beta \vec{M}(0) + M_0(1 - E_1(T))\,\hat{k}$. (For those following [15], note the inconsistency in defining $\mathbf{R}_\beta$ in Eq. 3; it is corrected in subsequent steps).

To complete one RF cycle, we next apply an RF pulse of duration $\tau \ll T_2$ and flip angle α. We assume that the pulse is sufficiently short that off-resonance effects, spin relaxation, and slice-selection gradients are negligible in comparison to the RF pulse. As mentioned previously, we orient our transverse axes such that the pulse rotation $\mathbf{R}_\alpha$ occurs about the i-axis:

$$\mathbf{R}_\alpha \triangleq \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{bmatrix}.$$
(2.4)

Applying this pulse describes the magnetization following one period, of duration $T_R \triangleq T + \tau$:

$$\vec{M}(T_R) = \mathbf{R}_\alpha \vec{M}(T) = \mathbf{R}_\alpha \mathbf{R}_\beta \vec{M}(0) + M_0(1 - E_1(T))\mathbf{R}_\alpha\,\hat{k}.$$
(2.5)

In steady-state, we assume that the magnetization has equilibrated to a periodic pattern between RF pulses, meaning that following each RF pulse, $\vec{M}^+ \triangleq \vec{M}(kT_R) = \vec{M}(0)$, $\forall k \in \mathbb{Z}$ is given as:

$$\begin{aligned}
\vec{M}^+ &= \mathbf{R}_\alpha \mathbf{R}_\beta \vec{M}^+ + M_0(1 - E_1(T))\mathbf{R}_\alpha\,\hat{k} \\
&= M_0(1 - E_1(T))[\mathbf{R}_\alpha^{-1} - \mathbf{R}_\beta]^{-1}\,\hat{k} \\
&= Q_\beta \begin{bmatrix} E_2(T)\sin\alpha\sin\beta \\ \sin\alpha(1 - E_2(T)\cos\beta) \\ \cos\alpha(1 - E_2(T)\cos\beta) + E_2^2(T) - E_2(T)\cos\beta \end{bmatrix},
\end{aligned}$$
(2.6)
(2.7)

where pre-factor Q_β can be written (with some simplification) as [15]:

$$Q_\beta \triangleq \frac{M_0(1 - E_1(T))}{(1 - E_1(T)\cos\alpha)(1 - E_2(T)\cos\beta) - E_2(T)(E_1(T) - \cos\alpha)(E_2(T) - \cos\beta)}.$$
(2.8)

The steady-state magnetization immediately prior each RF excitation is similarly given as

$$\vec{M}^- = [\mathbf{R}_\alpha]^{-1}\vec{M}^+ = Q_\beta \begin{bmatrix} E_2(T)\sin\alpha\sin\beta \\ -\sin\alpha(E_2^2(T) - E_2(T)\cos\beta) \\ (1 - E_2(T)\cos\beta) + \cos\alpha(E_2^2(T) - E_2(T)\cos\beta) \end{bmatrix}.$$
(2.9)

To maintain the simplicity of this formulation, we henceforth approximate $E_1(T) \approx E_1(T_R)$ and $E_2(T) \approx E_2(T_R)$. These approximations are very reasonable because a negligible amount of relaxation occurs over the RF pulse. With this substitution, the transverse magnetization before and
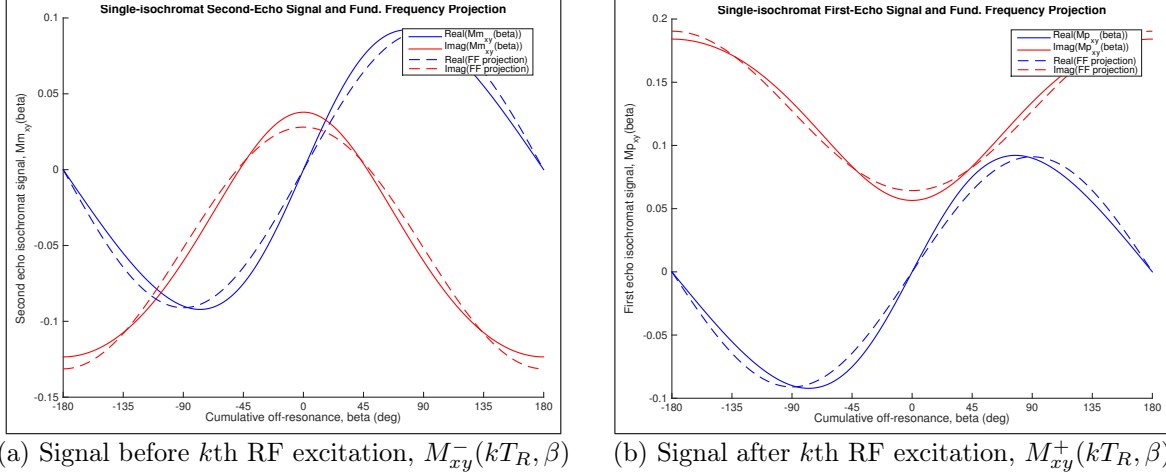
(a) Signal before kth RF excitation, $M_{xy}^-(kT_R, \beta)$ (b) Signal after kth RF excitation, $M_{xy}^+(kT_R, \beta)$

Figure 2.1: Single-isochromat DESS signals (solid lines), as functions of cumulative off-resonance effects β. Both signals are 2π periodic and contain most of their energy in the fundamental frequency, as shown by projections onto the subspace span$(1, \cos\beta, \sin\beta)$ (dashed lines). Results are displayed for $M_0 = 1$, $T_1 = 500$ms, $T_2 = 50$ms, $\alpha = 45°$, $T_R = 20$ms, $T_E = 5$ms.

after each RF pulse, denoted $M_{xy}^-(kT_R, \beta)$ and $M_{xy}^+(kT_R, \beta)$, respectively, can be written as

$$
\begin{aligned}
M_{xy}^-(kT_R, \beta) \triangleq M_x^- + iM_y^- &= Q_\beta \sin\alpha [E_2(T_R)\sin\beta - i(E_2^2(T_R) - E_2(T_R)\cos\beta)] \\
&= -iQ_\beta \sin\alpha [E_2^2(T_R) - E_2(T_R)e^{-i\beta}]; \quad\quad (2.10) \\
M_{xy}^+(kT_R, \beta) \triangleq M_x^+ + iM_y^+ &= Q_\beta \sin\alpha [E_2(T_R)\sin\beta + i(1 - E_2(T_R)\cos\beta)] \\
&= iQ_\beta \sin\alpha [1 - E_2(T_R)e^{i\beta}]. \quad\quad (2.11)
\end{aligned}
$$

These steady-state magnetization expressions are 2π-periodic functions of cumulative off-resonance phase β (Figure 2.1), which varies for different isochromats. However, if the DESS sequence includes a spoiler gradient, different spins within a given voxel will experience drastically different values of β. We can leverage this property by selecting spoiler areas that apply an integral number of spoiling cycles across each voxel, thereby creating a uniform distribution across $\beta \in [0, 2\pi)$. The signal intensities of interest, denoted $S^-(kT_R)$ and $S^+(kT_R)$ for before and after each RF excitation, respectively, will then be related to the *net* magnetization, found by integrating over β.

Assuming the imaging gradients are symmetric and balanced, just one complete spoiler phase cycle would theoretically suffice in creating a uniform distribution over β values. In practice however, B_0 field inhomogeneities can induce spurious through-voxel gradients that will modify the effective spoiler area and therefore create partial phase cycles, distorting the (nominally) uniform phase distribution. If ϕ denotes the true applied phase over a spoiler duration, we can separate $\phi =: 2\pi n + \tilde{\phi}$ to distinguish the effects of complete cycles n and partial phase $\tilde{\phi}$ on the bulk magnetization:

$$
\begin{aligned}
S^\pm(kT_R) &\triangleq \overbrace{\left[\frac{2\pi n}{\phi}\right] \frac{1}{2\pi n} \int_0^{2\pi n} M_{xy}^\pm(kT_R, \beta)\,\mathrm{d}\beta}^{\text{complete cycles}} + \overbrace{\frac{1}{\phi} \int_{2\pi n}^\phi M_{xy}^\pm(kT_R, \beta)\,\mathrm{d}\beta}^{\text{partial cycles}} \\
&= \left[\frac{2\pi n}{\phi}\right] \frac{1}{2\pi} \int_0^{2\pi} M_{xy}^\pm(kT_R, \beta)\,\mathrm{d}\beta + \frac{1}{\phi} \int_0^{\tilde{\phi}} M_{xy}^\pm(kT_R, \beta)\,\mathrm{d}\beta. \quad\quad (2.12)
\end{aligned}
$$

16

For fixed $\tilde{\phi}$, we observe that as nominal cycles n grows, $\frac{2\pi n}{\phi} \to 1$ and $\frac{1}{\phi} \to 0$, allowing the first term to increasingly dominate. From the perspective of dephasing completeness, it appears desirable to apply as many spoiling cycles as possible, given imaging time and gradient amplitude constraints. In practice, large gradient areas have been found to amplify diffusion effects in DESS [23], so appropriate selection of the number of spoiling cycles is essential to balancing the validity of subsequent full-cycle integrations with the validity of omitting diffusion effects from this analysis. In this work, we apply spoiler gradients along the slice-selective direction and acquire relatively thick slices to maximize through-voxel spoiler phase for a given gradient area. Nevertheless, the relationship between spoiler gradients and diffusion in DESS remains a topic of further research.

We continue this analysis assuming that a spoiler gradient imparts sufficient phase to create a uniform distribution of β over $[0, 2\pi)$ and neglect the resultant signal loss due to augmented diffusion effects. Following [19], we group terms involving β in Q_β with some new definitions:

$$
\begin{aligned}
Q_\beta &= \frac{M_0(1 - E_1(T_R))}{p - q\cos\beta}, \quad \text{where} \\
p &\triangleq (1 - E_1(T_R)\cos\alpha) - (E_1(T_R) - \cos\alpha)E_2^2(T_R) & (2.13) \\
q &\triangleq E_2(T_R)(1 - E_1(T_R))(1 + \cos\alpha). & (2.14)
\end{aligned}
$$

Physically, Q must be finite-valued for all possible values of p and q. To facilitate the subsequent integration, we can show that $p > q > 0$ because:

$$
\begin{aligned}
p - q &= [1 - E_2(T_R)][1 + E_1(T_R)E_2(T_R) - (E_1(T_R) + E_2(T_R))\cos\alpha] \\
&\geq [1 - E_2(T_R)][1 + E_1(T_R)E_2(T_R) - (E_1(T_R) + E_2(T_R))] \\
&= [1 - E_1(T_R)][1 - E_2(T_R)]^2 > 0. & (2.15)
\end{aligned}
$$

With these additional constraints, standard integration software is useful for bypassing some of the more tedious Taylor expansions necessary in earlier works [19, 20]. We first work on $S^+(kT_R)$:

$$
\begin{aligned}
S^+(kT_R) &= \frac{1}{2\pi}\int_0^{2\pi} M_{xy}^+(kT_R, \beta)\,\mathrm{d}\beta \\
&= \frac{iM_0\sin\alpha[1 - E_1(T_R)]}{2\pi}\int_0^{2\pi}\frac{1 - E_2(T_R)\cos\beta - iE_2(T_R)\sin\beta}{p - q\cos\beta}\,\mathrm{d}\beta \\
&= iM_0\sin\alpha[1 - E_1(T_R)]\left[\frac{1}{\sqrt{p^2 - q^2}} - \frac{E_2(T_R)}{q}\left(\frac{p}{\sqrt{p^2 - q^2}} - 1\right)\right] \\
&= iM_0\sin\alpha[1 - E_1(T_R)]\left[\frac{1}{\sqrt{p^2 - q^2}}\left(1 - \frac{E_2(T_R)p}{q}\right) + \frac{E_2(T_R)}{q}\right] \\
&= iM_0\frac{\sin\alpha}{1 + \cos\alpha}\left[1 - \frac{(E_1(T_R) - \cos\alpha)(1 - E_2^2(T_R))}{\sqrt{p^2 - q^2}}\right] \\
&= iM_0\tan\frac{\alpha}{2}\left[1 - \left(\frac{1 - E_2^2(T_R)}{\sqrt{p^2 - q^2}}\right)\left(E_1(T_R) - \cos\alpha\right)\right] & (2.16) \\
&= iM_0\tan\frac{\alpha}{2}\left[1 - \frac{1}{v_1}\sqrt{\frac{1 - E_2^2(T_R)}{1 - E_2^2(T_R)/v_1^2}}\right], \quad \text{where } v_1 \triangleq \frac{1 - E_1(T_R)\cos\alpha}{E_1(T_R) - \cos\alpha}. & (2.17)
\end{aligned}
$$

17

To arrive at (2.17) from (2.16), we borrow a clever substitution from [20]:

$$p^2 - q^2 = \left[1 - E_2^2(T_R)\right]\left[\left(1 - E_1(T_R)\cos\alpha\right)^2 + E_2^2(T_R)\left(E_1(T_R) - \cos\alpha\right)^2\right]. \tag{2.18}$$

A very similar integration over β yields the second signal's net magnetization:

$$
\begin{aligned}
S^-(kT_R) &= \frac{1}{2\pi}\int_0^{2\pi} M_{xy}^-(kT_R, \beta)\,\mathrm{d}\beta \\
&= \frac{-iM_0\sin\alpha[1 - E_1(T_R)]}{2\pi}\int_0^{2\pi}\frac{E_2^2(T_R) - E_2(T_R)\cos\beta - iE_2(T_R)\sin\beta}{p - q\cos\beta}\,\mathrm{d}\beta \\
&= -iM_0\sin\alpha[1 - E_1(T_R)]\left[\frac{1}{\sqrt{p^2 - q^2}}\left(E_2^2(T_R) - \frac{E_2(T_R)p}{q}\right) + \frac{E_2(T_R)}{q}\right] \\
&= -iM_0\tan\frac{\alpha}{2}\left[1 - \left(\frac{1 - E_2^2(T_R)}{\sqrt{p^2 - q^2}}\right)\left(1 - E_1(T_R)\cos\alpha\right)\right] \tag{2.19} \\
&= -iM_0\tan\frac{\alpha}{2}\left[1 - \sqrt{\frac{1 - E_2^2(T_R)}{1 - E_2^2(T_R)/v_1^2}}\right], \text{where } v_1 \triangleq \frac{1 - E_1(T_R)\cos\alpha}{E_1(T_R) - \cos\alpha}. \tag{2.20}
\end{aligned}
$$

Equations (2.16) and (2.19) are similar to the models given in [14], while (2.17) and (2.20) more clearly reveal the similarity between the steady-state signals, isolating all T_1 dependence into an auxiliary variable v_1, which can be estimated separately, for $\alpha \neq \arccos(E_1(T_R))$. Note that in transforming these expressions, care is taken in considering the sign of the radical term for all v_1. Formulations (2.17) and (2.20) are equivalent to (2.16) and (2.19), respectively, $\forall \alpha \neq \arccos(E_1(T_R))$.

2.2.2 Model Refinements

We can further improve the model for the received signal by accounting for the time evolution before and after RF excitations and readouts. Signal dynamics during these periods can be particularly significant in steady-state imaging because T_R is purposefully kept short, and thus T_E is very often on the order of T_R. We will assume that acquisitions are centered at times T_E and $T_R - T_E$.

S^+ Refinement

We first investigate the magnetization of a single isochromat at the time of the readout following RF excitation, $M_{xy}^+(kT_R + T_E, \beta)$, which is related to $M_{xy}^+(kT_R, \beta)$ by T_2 relaxation and additional phase accrual due to off-resonance effects:

$$M_{xy}^+(kT_R + T_E, \beta) = M_{xy}^+(kT_R, \beta)E_2(T_E)e^{i\theta(T_E)}, \tag{2.21}$$

where $\theta(t) \triangleq \int_0^t \Delta\omega(t')\,\mathrm{d}t'$ is now used to model off-resonance effects other than crusher gradients. For a given isochromat, we can reasonably assume that the effective off-resonance frequency, denoted $\Delta\omega_f$, is roughly constant for the short duration of T_E; then $\theta(t) \approx t\Delta\omega_f$.

18

We would once again like to describe the macroscopic effect of off-resonance by integrating over the many spins within each voxel. Since no crusher gradients are in effect during the period in which $\theta(t)$ accrues, we may not assume uniformly-distributed dephasing; instead, we must assume some other broadening distribution, denoted $f(\Delta\omega_f)$. Then the overall signal can be expressed as:

$$
\begin{aligned}
S^+(kT_R + T_E) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{0}^{2\pi} M_{xy}^+(kT_R + T_E, \beta) f(\Delta\omega_f) \, \mathrm{d}\beta \, \mathrm{d}\Delta\omega_f \\
&\approx \frac{E_2(T_E)}{2\pi} \int_{-\infty}^{\infty} \left[\int_{0}^{2\pi} M_{xy}^+(kT_R, \beta) \, \mathrm{d}\beta \right] e^{i\Delta\omega_f T_E} f(\Delta\omega_f) \, \mathrm{d}\Delta\omega_f. \quad (2.22)
\end{aligned}
$$

Note that phase accruals $\theta(t)$ from previous RF cycles have been crushed by corresponding dephasing gradients; thus, it is reasonable to assume that the inner integrand has no direct dependence on $\Delta\omega_f$, allowing separable integration. We model $f(\Delta\omega_f)$ to be distributed as $\text{Cauchy}(\Delta\bar\omega_f,\ R_2')$, with characteristic function $\Phi_{\Delta\omega_f}(t) = e^{i\Delta\bar\omega_f t - R_2'|t|}$, where $\Delta\bar\omega_f$ denotes the median off-resonance frequency and $R_2' \triangleq \frac{1}{T_2'}$ is the broadening bandwidth, both of which are voxel-dependent:

$$
\begin{aligned}
S^+(kT_R + T_E) &\approx \left[\frac{E_2(T_E)}{2\pi} \int_{0}^{2\pi} M_{xy}^+(kT_R, \beta) \, \mathrm{d}\beta \right] \left[\int_{-\infty}^{\infty} e^{i\Delta\omega_f T_E} f(\Delta\omega_f) \, \mathrm{d}\Delta\omega_f \right] \\
&= \left[E_2(T_E) S^+(kT_R) \right] \left[\Phi_{\Delta\omega_f}(T_E) \right] \\
&= S^+(kT_R) E_2(T_E) e^{i\Delta\bar\omega_f T_E - R_2' T_E} \\
&= M_0 E_2'(T_E) e^{i(\Delta\bar\omega_f T_E + \frac{\pi}{2})} E_2(T_E) \tan\frac{\alpha}{2} \left[1 - \frac{1}{v_1} \sqrt{\frac{1 - E_2^2(T_R)}{1 - E_2^2(T_R)/v_1^2}} \right]. \quad (2.23)
\end{aligned}
$$

S^- Refinement

A similar approach can be used to find the magnetization of a single isochromat during the readout prior to an RF excitation. It will behoove us to observe the signal of interest, $M_{xy}^-(kT_R - T_E, \beta)$, as it refocuses into $M_{xy}^-(kT_R, \beta)$ from phase accrual due to non-crusher, off-resonance effects:

$$
M_{xy}^-(kT_R, \beta) = M_{xy}^-(kT_R - T_E, \beta) E_2(T_E) e^{i\theta(T_E)}. \quad (2.24)
$$

Making identical assumptions on $\theta(T_E)$ as before, a similar computation yields $S^-(kT_R - T_E)$:

$$
\begin{aligned}
S^-(kT_R - T_E) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{0}^{2\pi} M_{xy}^-(kT_R - T_E, \beta) f(\Delta\omega_f) \, \mathrm{d}\beta \, \mathrm{d}\Delta\omega_f \\
&\approx \frac{E_2^{-1}(T_E)}{2\pi} \int_{-\infty}^{\infty} \left[\int_{0}^{2\pi} M_{xy}^-(kT_R, \beta) \, \mathrm{d}\beta \right] e^{i\Delta\omega_f(-T_E)} f(\Delta\omega_f) \, \mathrm{d}\Delta\omega_f \\
&= \left[E_2^{-1}(T_E) S^-(kT_R) \right] \left[\Phi_{\Delta\omega_f}(-T_E) \right] \\
&= M_0 E_2'(T_E) e^{-i(\Delta\bar\omega_f T_E + \frac{\pi}{2})} E_2^{-1}(T_E) \tan\frac{\alpha}{2} \left[1 - \sqrt{\frac{1 - E_2^2(T_R)}{1 - E_2^2(T_R)/v_1^2}} \right]. \quad (2.25)
\end{aligned}
$$

It is important to note that in steady-state, the dual echoes are *not* symmetrically dependent on $R_2^* \triangleq R_2 + R_2'$! Rather, echoes tend to refocus at the RF excitation and defocus afterwards, while T_2 relaxation always causes an exponential decrease in signal with time.

2.2.3 Formulation Used for this Work

The signal models presented in Equations (2.23) and (2.25) are intractable because of their $\Delta\bar{\omega}_f$ dependence. However, if we reasonably assume $\Delta\bar{\omega} \approx 0$ for short readouts, then

$$S^-(kT_R - T_E) \approx -iM_0 E_2'(T_E) E_2^{-1}(T_E) \tan\frac{\alpha}{2}\left[1 - \sqrt{\frac{1 - E_2^2(T_R)}{1 - E_2^2(T_R)/v_1^2}}\right] \tag{2.26}$$

$$S^+(kT_R + T_E) \approx iM_0 E_2'(T_E) E_2(T_E) \tan\frac{\alpha}{2}\left[1 - \frac{1}{v_1}\sqrt{\frac{1 - E_2^2(T_R)}{1 - E_2^2(T_R)/v_1^2}}\right]. \tag{2.27}$$

In this work, we seek to accurately estimate T_2 maps from DESS data modeled as Equations (2.26) and (2.27), but observe that nuisance parameters M_0, T_1, and R_2' complicate this goal. Note that in this estimation problem, T_1 is treated as a nuisance parameter, even though it may contain medically useful information. While other user-specified parameters (T_R, T_E, α) are known, nuisance parameters M_0, T_1, and R_2' are unknown, and thus must also be estimated.

2.2.4 Signal Behavior Under Limiting Conditions

We can use Figure 2.2 to visualize some of the limiting behavior of DESS signals, as well as the ratio between the two, as utilized in prior works [14]. We first observe that for small flip angles, the signals rapidly drop off to zero, as does the ratio: this is therefore a poor regime for taking data. For large flip angles ($\pm\pi$), the signals also decay to zero, but the signal ratio levels out to approximately $E_2^2(T_R - T_E)$, yielding an easy initial estimator of T_2. Because the signal reaches this limit rather quickly, we can obtain good initial approximations of T_2 even for $\alpha = \pm\frac{\pi}{2}$ via division. However, we wish to take data at angles with maximum SNR, which is achieved for "Ernst-like" flip angles (in this simulation, about $25°$). At these lower flip angles, the signal begins to depend on T_1 more significantly, and a joint estimation approach will be necessary to produce accurate T_2 maps.

The limiting behavior as T_R is varied is straightforward. For $T_R \ll T_1$, $E_1(T_R) \approx 1$, causing minimum signal in S^+ and maximum signal in S^-. In fact, since there is no relaxation as $T_R \to 0$, the signals are equal in the limit, and the ratio signal is a maximum of 1. In practice, this limit cannot exactly be achieved because $T_R \geq 2T_E \geq 2T_{acq}$, where T_{acq} is the time needed per readout. At the other extreme, for $T_R \gg T_1$, S^- energy unsurprisingly drops off to zero. In an effort to obtain reasonable SNR from each signal (and to minimize scan times), it is reasonable to set T_R small, though it should be noted that S^+ signal energy is sacrificed with this design choice.

We next observe limiting behavior as intrinsic parameters T_1 and T_2 vary over a region of interest. For $T_1 \gg T_R$ ($E_1(T_R) \approx 1$), the ratio of signals once again approaches approximately $E_2^2(T_R - T_E)$, but the signal energy of each individual signal unsurprisingly drops to zero, meaning that we will have very low SNR in such regions. For lower T_1 values ($T_R/T_1 > 0.15$), the first signal strengthens while the second signal weakens, causing their ratio to rapidly increase in its dependence on T_1. Maximum S^- signal energy is achieved at an intermediate value. In practice, T_R should never be selected such that the expected E_1 drops below this threshold, as (1) S^- SNR would decrease, (2) ratio dependence on T_1 would increase, and (3) scan times would increase.
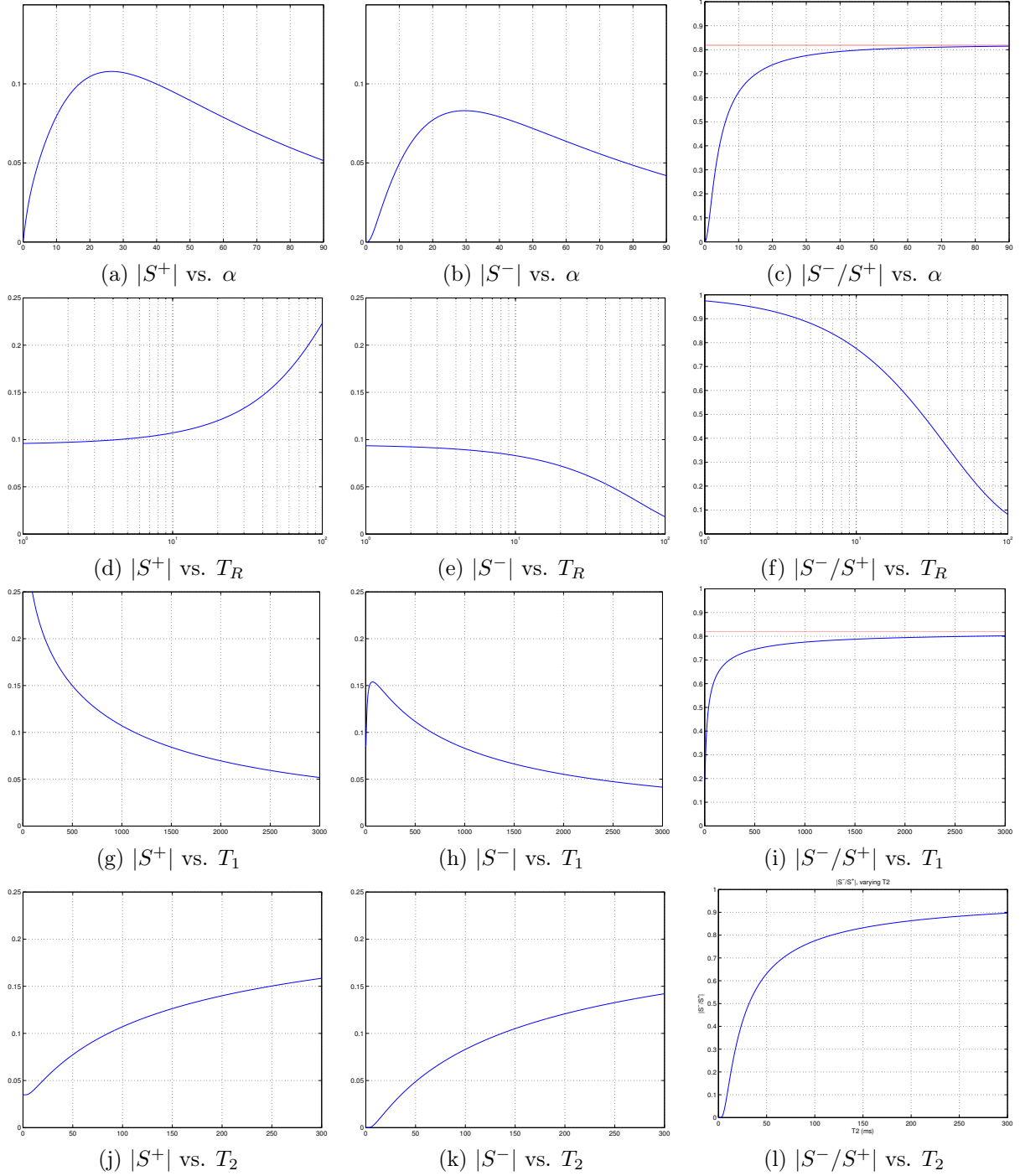
Figure 2.2: Theoretical behavior of noiseless, magnitude DESS signals (and their magnitude ratios) as functions of both adjustable and intrinsic parameters. $|S^+|$ (left), $|S^-|$ (center), and $|S^-/S^+|$ (right), are plotted as functions of α, T_R, T_1 and T_2 (top to bottom). Where relevant, other parameters are fixed reasonably, except $T_E = 0$ms for simplicity: $T_1 = 1000$ms, $T_2 = 100$ms, $T_R = 10$ms, and $\alpha = 30°$. Where shown, the red line highlights the deviation between the conventional method-of-moments approximation, $E_2^2(T_R - T_E)$, and the expected ratio.

Finally, signal dependencies on T_2 are straightforward: $|S^+|$, $|S^-|$, and $|S^+/S^-|$ all are maximized for $T_2 \gg T_R$, and monotonically decay to zero for shorter T_2 values. Again, this reiterates that T_R should be selected as small as possible to maximize signal energy.

2.3 Problem Formulation

2.3.1 Conventional Method-of-Moments T_2 Estimation

As observed above, conventional methods [8, 18] observe that for either the large nutation angle limit $\alpha \to \frac{\pi}{2}$ or the long spin-lattice relaxation limit $T_1 \to \infty$, the ratio between Equations (2.26) and (2.27) approaches a simple model, independent of all other unknowns:

$$\left| \frac{S^-(kT_R - T_E)}{S^+(kT_R + T_E)} \right| \to E_2^2(T_R - T_E) \,, \tag{2.28}$$

from which T_2 can be easily estimated. This method-of-moments estimator is prone to several error sources. First, signal division will amplify noise. Next, since large flip angles induce low-SNR signals, they are replaced with lower angles in practice. At typical flip angle choices, the signal ratio deviates from ((2.28)) through substantial T_1 dependence, inducing systematic error in the T_2 estimate. Even if DESS data is collected at multiple angles, corresponding total least-squares (TLS) estimates [24] will still suffer from this T_1-dependent bias. Accordingly, we exclusively consider a TLS-enhanced version of the method-of-moments estimator, henceforth referred to as the TLS estimator.

2.3.2 Penalized-Likelihood T_2 Estimation

We propose to mitigate T_1-related bias by first estimating unknown nuisance parameters from a fast, separate scan, before performing T_2 estimation. We observe that T_1 and

$$M_0^* := iM_0 E_2'(T_E)E_2(T_E) = iM_0 E_2^*(T_E) \tag{2.29}$$

can, for example, be initially estimated from fast Spoiled Gradient-Recalled Echo (SPGR) scans, repeated over as few as two flip angles [8]. We select T_E and T_R to be the same for all SPGR and DESS scans. We estimate M_0^* and T_1 nuisance parameter maps from the SPGR signal model [3],

$$S_{\mathrm{SPGR}} = \frac{M_0^* \sin \alpha (1 - E_1(T_R))}{1 - E_1(T_R) \cos \alpha} \,, \tag{2.30}$$

and take these estimates as fixed parameters during the subsequent T_2 estimation. We hereafter focus on estimating T_2 from Equations (2.26) and (2.27).

We estimate $\mathbf{x} \in \mathcal{X}$, with j indexing a total of N voxels and jth element $x_j := \exp\{-T_R/T_{2,j}\}$. We define noisy measurement vectors $\mathbf{y}_l^- := [y_{l,1}^-, \ldots, y_{l,N}^-]^{\mathrm{T}}$ and $\mathbf{y}_l^+ := [y_{l,1}^+, \ldots, y_{l,N}^+]^{\mathrm{T}}$, where l indexes a total of L nutation angles; $\epsilon_{l,j}^{\pm} \sim \mathbb{CN}(0, \sigma^2)$ is complex, white Gaussian noise [6, 7]; and

$$y_{l,j}^{\pm} := S^{\pm}(kT_R \pm T_E; T_{2,j}, \alpha_l) + \epsilon_{l,j}^{\pm} \,. \tag{2.31}$$

$\mathcal{X} := [E_{2,\min}, E_{2,\max}]^N \subset (0,1]^N$ is a solution space corresponding to a range $[T_{2,\min}, T_{2,\max}]$ over which we seek accurate estimation. An ML estimator for $\mathbf{x}$ is then given by:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x} \in \mathcal{X}} \Psi_{\text{ML}}(\mathbf{x}), \text{ where} \tag{2.32}$$

$$\Psi_{\text{ML}}(\mathbf{x}) = \frac{1}{2} \sum_{l=1}^{L} \left[||\mathbf{y}_l^- - \mathbf{f}_l^-(\mathbf{x})||^2 + ||\mathbf{y}_l^+ - \mathbf{f}_l^+(\mathbf{x})||^2 \right].$$

Observe that the vector functions $\mathbf{f}_l^{\pm}(\mathbf{x})$ are voxel-wise independent mappings from $\mathbb{R}^N \to \mathbb{C}^N$ that naturally extend the corresponding scalar functions $f_l^{\pm}(x) : \mathbb{R} \to \mathbb{C}$, defined as:

$$f_l^-(x) = -M_0^* \tan\frac{\alpha_l}{2} x^{-\frac{2T_E}{T_R}} \left[1 - \sqrt{\frac{1 - x^2}{1 - x^2/v_{1,l}^2}} \right]; \tag{2.33}$$

$$f_l^+(x) = M_0^* \tan\frac{\alpha_l}{2} \left[1 - \frac{1}{v_{1,l}} \sqrt{\frac{1 - x^2}{1 - x^2/v_{1,l}^2}} \right]. \tag{2.34}$$

In addition to the data, we assume prior knowledge that T_2 maps tend to be piecewise-smooth; that is, adjacent voxels are likely to contain similar tissue types, and are thus likely to have similar T_2 values. We thus extend the ML cost function to include an edge-preserving roughness penalty:

$$\Psi_{\text{MAP}}(\mathbf{x}) = \Psi_{\text{ML}}(\mathbf{x}) + \beta R(\mathbf{x}), \tag{2.35}$$

where $R(\mathbf{x}) := \sum_{k=1}^{K} \phi_k([\mathbf{C}\mathbf{x}]_k)$; $\mathbf{C}$ is a 2D or 3D spatial finite differencing matrix; ϕ_k is a convex, edge-preserving potential function; k indexes a total K differencing operations; and β is a regularization parameter that controls the level to which roughness in $\mathbf{x}$ is penalized. We then obtain a penalized-likelihood estimate of $\mathbf{x}$ by solving the (non-convex) constrained optimization problem:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x} \in \mathcal{X}} \Psi_{\text{MAP}}(\mathbf{x}). \tag{2.36}$$

2.4 Minimization Techniques

2.4.1 Preconditioned Gradient Projection Method (PGPM)

Since $\mathcal{X}$ is closed and convex, a preconditioned gradient projection method (PGPM) of the form

$$\mathbf{x}^{(n+1)} = \mathcal{P}_{\mathcal{X}}\left\{ \mathbf{x}^{(n)} - \mathbf{D}^{-1} \nabla \Psi_{\text{MAP}}(\mathbf{x}^{(n)}) \right\} \tag{2.37}$$

will monotonically decrease Ψ_{MAP} to a local minimum, as long as preconditioner $\mathbf{D}$ is selected such that $\forall \mathbf{x} \in \mathcal{X}$,

$$\nabla^2 \Psi_{\text{MAP}}(\mathbf{x}^{(n)}) = \nabla^2 \Psi_{\text{ML}}(\mathbf{x}^{(n)}) + \beta \nabla^2 R(\mathbf{x}^{(n)}) \preceq \mathbf{D}. \tag{2.38}$$

Here we have denoted $\mathcal{P}_{\mathcal{X}}\{\cdot\}$ as a projection operator onto $\mathcal{X}$. Observe that because functions $\mathbf{f}_l^{\pm}(\mathbf{x})$ are voxel-wise independent, the Hessian of the ML cost,

$$\nabla^2 \Psi_{\text{ML}}(\mathbf{x}) = \frac{1}{2} \sum_{l=1}^{L} \left[\nabla^2 ||\mathbf{y}_l^- - \mathbf{f}_l^-(\mathbf{x})||^2 + \nabla^2 ||\mathbf{y}_l^+ - \mathbf{f}_l^+(\mathbf{x})||^2 \right], \tag{2.39}$$
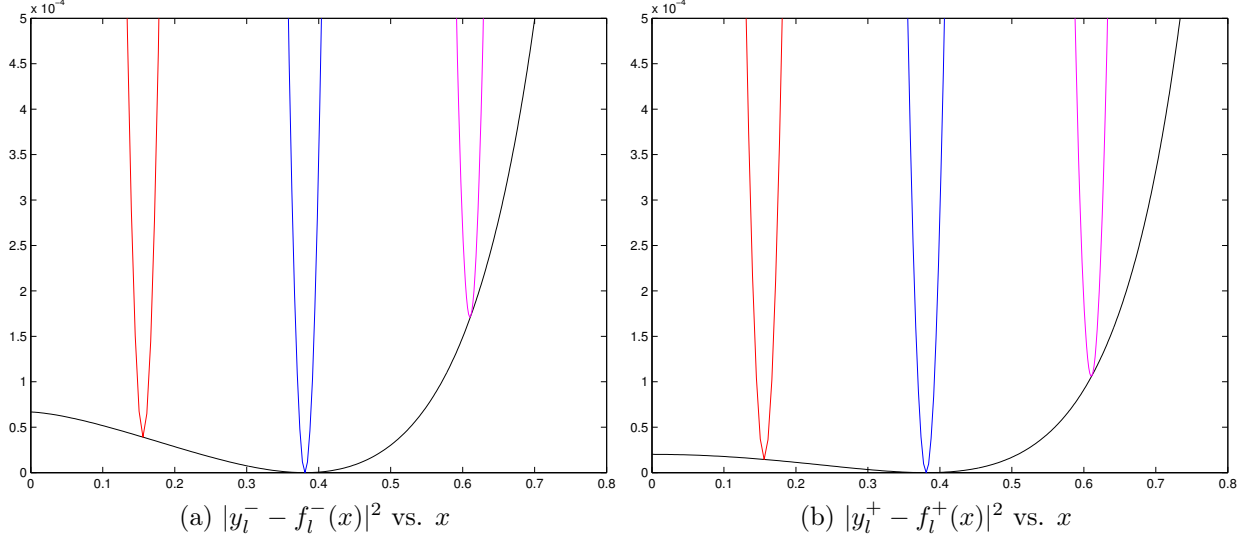
Figure 2.3: Comparisons of a single-voxel data-fit term (black) with quadratic majorizers (colored), whose curvatures are given by ((2.40)). High majorizer curvatures cause for slow PGPM convergence without acceleration. Figures created with $\alpha = \frac{\pi}{4}$, $T_R = 20$ms, $T_E = 5$ms, $T_1 = 1000$ms, and $\mathcal{X}$ corresponding to $T_2 = [5, 200]$ms.

is $N \times N$ diagonal. This allows us to first consider bounding the corresponding single-voxel data-fit curvatures $\ddot{\psi}_l^{\pm}(x)$. We then construct $\mathbf{D}$ via a trivial diagonal matrix extension.

We have shown that curvatures $\ddot{\psi}_l^{\pm}(x)$ are convex over $x \in \mathcal{X}, \forall \alpha \in [0, \frac{\pi}{2}]$ and $T_1 \in [0, \infty)$. We can therefore bound the curvatures by evaluating at the endpoints of the solution set. In practice, since $\lim_{x \to 1} \ddot{\psi}_l^{\pm}(x)$ can be unbounded, we select $T_{2,\max} < \infty$ (and thus $E_{2,\max} < 1$) to ensure that the endpoint curvatures are finite. With these restrictions, each of the data fidelity terms' curvatures can thereby be bounded as

$$d_l^{\pm} = \max_{x \in \{E_{2,\min}, E_{2,\max}\}} \left\{ \ddot{\psi}_l^{\pm}(x) \right\}. \tag{2.40}$$

After computing such bounds $d_{l,j}^{\pm}$ at each voxel $j = 1, \ldots, N$, the diagonal preconditioner

$$\mathbf{D} = \sum_{l=1}^{L} \operatorname{diag}_j \left\{ d_{l,j}^{+} + d_{l,j}^{-} \right\} + \beta \operatorname{maxeig} \left\{ \nabla^2 R \right\} \mathbf{I}_N \tag{2.41}$$

ensures that PGPM will monotonically decrease Ψ_{MAP} to a local minimum. The rate of this convergence can be increased at the expense of a tighter solution set $\mathcal{X}$.

2.4.2 Acceleration with Nesterov Momentum

We observe in Fig. 2.3 that even with a restrictive choice of $\mathcal{X}$, selection of $d_l^{\pm}$ via ((2.40)) yields a loose bound on the curvature. Quadratic majorizers thus have high curvatures, leading to slow convergence rates of majorize-minimize algorithms [25]. To accelerate convergence, we consider

(a) Total Least-Squares   (b) Median-filtered   (c) MBET2   (d) Ground Truth

(e) RMSE = (16.9, 12.3)ms (f) RMSE = (14.1, 10.2)ms (g) RMSE = (4.3, 2.6)ms
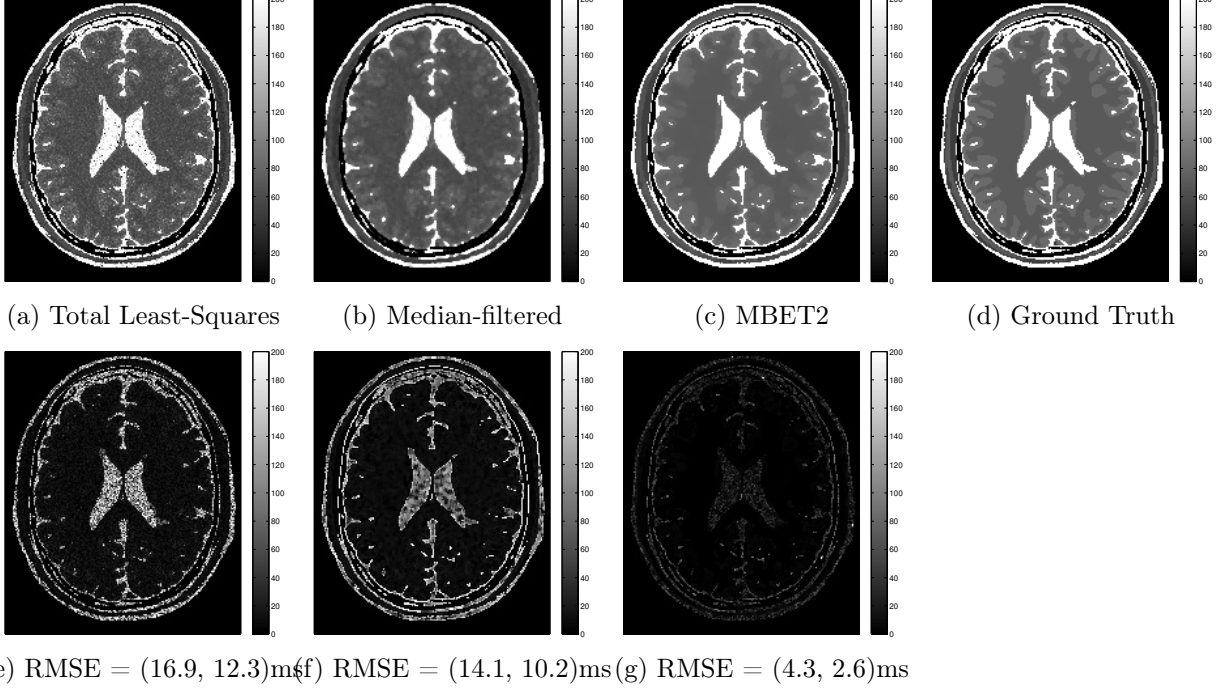
Figure 2.4: T_2 maps (a, b, c) and corresponding absolute-difference images (e, f, g), with respect to the ground truth (d). Total least squares (a) and median-filtered (b) images reconstructed from 40dB SNR DESS data at RMSE-minimizing flip angle combination $\alpha = 45, 65, 85°$. MBET2 image (c) reconstructed from 40dB SNR DESS data at $\alpha = 45°$, following MBET1 reconstruction of M_0^* and T_1 maps from 60dB SNR SPGR data at $\alpha = 5, 30°$. Tissue-specific errors are spatially averaged over (GM, WM) voxels of interest, and are reported in RMSE.

addition of Nesterov's momentum [26] to the PGPM algorithm. The original algorithm is extended in a direct manner to our case of a diagonal preconditioner [27], yielding a simple iteration:

$$t_{n+1} = \left(1 + \sqrt{1 + 4t_n^2}\right)/2$$

$$\mathbf{x}^{(n+1)} = \mathcal{P}_{\mathcal{X}}\left\{\mathbf{z}^{(n)} - \mathbf{D}^{-1}\nabla\Psi_{\text{MAP}}(\mathbf{z}^{(n)})\right\}$$

$$\mathbf{z}^{(n+1)} = \mathbf{x}^{(n+1)} + \frac{t_n - 1}{t_{n+1}}\left(\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\right), \tag{2.42}$$

where $\mathbf{z}^{(0)} = \mathbf{x}^{(0)}$ and $t_0 = 1$. This momentum extension to PGPM considerably accelerates convergence, with a marginal increase in storage requirement and computation.

2.5 Experimentation

2.5.1 Results from Synthetic Data

We combine true M_0, T_1, T_2, and T_2^* maps from the 95th slice of the BrainWeb digital phantom [10] to compute uncorrupted DESS data at 217×181 resolution, using Equations (2.26) and (2.27). We

corrupt the true data with complex, white Gaussian noise, to yield 40dB SNR raw data. We strive to recover accurate T_2 maps from this raw data, with no prior knowledge of M_0, T_1, or T_2^*.

To estimate the nuisance parameters, we simulate 60dB SPGR data using ((2.30)), at "ideal" [8] nutation angles $\alpha = 5, 30°$. From this data, we initially estimate M_0^* and T_1 maps using MBET1 [3]. These estimates are taken to be fixed parameters during subsequent T_2 estimation.

We compute a TLS T_2 estimate [18] from noisy DESS data generated at $\alpha = 45°$. We initialize MBET2 with a median-filtered TLS estimate and iterate until convergence criterion $||\mathbf{x}^{(n)} - \mathbf{x}^{(n-1)}|| < 10^{-6}||\mathbf{x}^{(n)}||$ is satisfied. We choose ϕ_k to be a hyperbolic potential function. We select $\beta = 2^7$ heuristically, though automatic selection methods [11] can be applied in practice. On a 3.5GHz desktop with 32GB RAM, with solution set $\mathcal{X}$ corresponding to a T_2 range of [5,1000]ms, reconstructions take roughly 35s per slice for MBET2 reconstruction of DESS data.

Fig. 2.4 compares MBET2 against conventional estimates, both before and after median filtering. To ensure a fair comparison in terms of scan time, the conventional estimates are computed from DESS data collected at three nutation angles. These angles, $\alpha = 45, 65, 85°$, are selected to minimize root mean-squared error (RMSE) in grey-matter (GM) and white-matter (WM) regions of interest.

As evidenced by absolute difference images with respect to the true T_2 map, the TLS estimate is prone to noise amplification due to signal division. Median-filtering reduces noise to some extent, but resulting T_2 maps remain prone to systematic error due to model-mismatch.

Tissue	TLS	Median-Filt.	Proposed	Truth
GM	75.1 ± 15.0	73.5 ± 9.6	82.3 ± 4.2	83.0 ± 0.0
WM	60.5 ± 8.0	60.6 ± 3.7	71.0 ± 2.5	70.0 ± 0.0

Table 2.1: T_2 estimates and standard deviations, spatially averaged over GM and WM regions of interest. Scan protocols match those in Fig. 2.4. All values reported in (ms). The proposed method exhibits increases in both accuracy and precision over TLS estimates.

Table 2.1 quantitatively highlights that MBET2 achieves improvements in both accuracy and precision over TLS estimates (with or without median filtering), particularly in GM and WM regions of interest. Accuracy improvements are due to reduced T_1 induced bias. Precision improvements are due to (1) omission of noise-amplifying signal division, and (2) usage of edge-preserving regularization to encourage piecewise-smooth T_2 estimates. Accuracy and precision increases combine to significantly reduce RMSE values, given in Fig. 2.4. Such improvements are sustained over a wide range of scan parameters and noise thresholds (results omitted).

2.5.2 Results from *In Vivo* Data

2.6 Conclusions and Future Work

We have presented MBET2, a novel, model-based approach for T_2 relaxometry from DESS sequences. Compared to conventional methods, we observe that the proposed method improves estimation *accuracy*, primarily due to a reduction of T_1-induced bias through preliminary M_0^*, T_1

joint estimation from SPGR data. We also observe improvements in estimation *precision* through omission of noise-amplifying signal division and utilization of edge-preserving regularization.

Residual MBET2 reconstruction errors are likely dominated by the propagation of M_0^*, T_1 estimation error. To bypass this error propagation, it is desirable to recast the T_1, T_2 estimation problem as a single joint optimization, as opposed to the sequential MBET1, followed by MBET2 framework presented here. Such problem reformulation invites the broader question of appropriate scan selection for effective simultaneous estimation. In the next chapter, we will explore different scan profiles and seek a fast pulse sequence that permits accurate and precise joint T_1, T_2 estimation.

Chapter 3

Optimized Scan Selection Methods for Joint T_1, T_2 Estimation (MBERP)

3.1 Introduction

In this chapter, we consider the problem of optimal scan selection for joint T_1, T_2 estimation. In general, MR scan selection is a challenging problem because there are many scan sequences to consider, and many scan parameters (for each candidate sequence) to control. Building upon the previous chapters, we consider different combinations of SPGR and DESS sequences only, but observe that even better scan combinations could be possible using other pulse sequences. Similar to prior works [8, 28], we restrict our optimization to flip angles and repetition rates, but again observe that expanding our search space to other scan parameters, *i.e.*, echo times, could permit better scan profiles. Fortunately, the work presented here is extendable to such larger search spaces.

In our context, optimality is loosely defined in terms of the worst-case variance in T_1 and T_2 estimates, over appropriate ranges of interest. We use the Cramér-Rao Lower Bound (CRLB) to analyze the minimum possible variance for unbiased, joint estimation of T_1, T_2 (as well as nuisance parameter M_0^*), for several reasonable combinations of SPGR and DESS scans. We observe that this analysis omits consideration of biasing regularization terms, but we anticipate that scan profiles selected from this analysis will also yield precise T_1, T_2 maps with regularization.

in Section 2, we briefly review the CRLB and apply it to a generalized M_0^*, T_1, T_2 mapping problem, for extraction of the minimum variance in T_1, T_2 unbiased estimates from an arbitrary scan profile. In Section 3, we describe a min-max optimization problem for scan selection, present associated heat maps, and decide on a specific scan profile for joint estimation. In Section 4, we describe a dictionary-based variable projection method for initialization, as well as an alternating minimization algorithm for regularized, joint T_1, T_2 estimation using the suggested scan profile. As it borrows ideas from MBET1 and MBET2, this method is labeled Model-Based Estimation of Relaxation Parameters (MBERP). In Section 5, we provide a performance comparison between sequential and joint approaches. In Section 6, we conclude our findings and suggest directions for future work.

3.2 CRLB Analysis in Quantitative MRI

3.2.1 A General Model for the CRLB

A broad class of pulse sequences useful for T_1, T_2 mapping produce per-voxel signals that can be described with the following general model:

$$y_m = f_m(\boldsymbol{\theta}; \alpha_m, T_{R,m}, T_{E,m}) + \epsilon_m \,, \tag{3.1}$$

where $\boldsymbol{\theta} := [M_0^*, T_1, T_2]^T \in \mathbb{C}^P$ collects the $P = 3$ unknown parameters; $\alpha_m, T_{R,m}, T_{E,m}$ collectively denote a particular choice of flip angle, repetition time, and echo time (respectively) for the mth dataset; and $f_m : \mathbb{C} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \mapsto \mathbb{C}$ is a (pulse-sequence dependent) function that models the noiseless signal value for the mth dataset. We further model $\epsilon_m \sim \mathbb{CN}(0, \sigma_m^2)$ to be complex, independent Gaussian noise [6, 7]. In this general framework, we allow datasets to exhibit different noise variances, though we later apply the CRLB with a constant-variance assumption.

A complete scan profile contains a total M datasets from a combination of pulse sequences. If $\mathbf{y} := [y_1, \ldots, y_m]^T \in \mathbb{C}^M$ and vector function $\mathbf{f} : \mathbb{C} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^M \times \mathbb{R}^M \times \mathbb{R}^M \mapsto \mathbb{C}^M$ naturally extends the scalar function f, then the log-likelihood function (still for a single voxel) is

$$\ln L(\boldsymbol{\theta}) = -\frac{1}{2} \|\mathbf{y} - \mathbf{f}(\boldsymbol{\theta}; \boldsymbol{\alpha}, \mathbf{T_R}, \mathbf{T_E})\|_{\boldsymbol{\Sigma}^{-1/2}}^2 \,, \tag{3.2}$$

where $\boldsymbol{\alpha} := [\alpha_1, \ldots, \alpha_M]^T$, $\mathbf{T_R} := [T_{R,1}, \ldots, T_{R,M}]^T$, and $\mathbf{T_E} := [T_{E,1}, \ldots, T_{E,M}]^T$ are vectors in $\mathbb{R}^M$ containing all M scan parameter choices. Here we have also introduced the covariance matrix $\boldsymbol{\Sigma} := \mathrm{diag}\{\sigma_1^2, \ldots, \sigma_M^2\}$, which is diagonal due to the independence assumption. Observing that

$$[\nabla \mathbf{f}(\boldsymbol{\theta})]^T := [\nabla \mathbf{f}(\boldsymbol{\theta}; \boldsymbol{\alpha}, \mathbf{T_R}, \mathbf{T_E})]^T = \begin{bmatrix} \partial/\partial M_0^* \\ \partial/\partial T_1 \\ \partial/\partial T_2 \end{bmatrix} [\mathbf{f}(\boldsymbol{\theta}; \boldsymbol{\alpha}, \mathbf{T_R}, \mathbf{T_E})]^T \tag{3.3}$$

is a $P \times M$ matrix of partial derivatives corresponding to a particular scan profile of M scans, we can express the $P \times P$ Fisher information matrix $\mathbf{F}(\boldsymbol{\theta})$ entirely independent of the data $\mathbf{y}$:

$$\begin{aligned}
\mathbf{F}(\boldsymbol{\theta}) &:= \mathbb{E}\left([\nabla \ln L(\boldsymbol{\theta})]^T [\nabla \ln L(\boldsymbol{\theta})]\right) \\
&= \mathbb{E}\left(\left[[\nabla \mathbf{f}(\boldsymbol{\theta})]^T \boldsymbol{\Sigma}^{-1}(\mathbf{y} - \mathbf{f}(\boldsymbol{\theta}))\right]\left[[\nabla \mathbf{f}(\boldsymbol{\theta})]^T \boldsymbol{\Sigma}^{-1}(\mathbf{y} - \mathbf{f}(\boldsymbol{\theta}))\right]^T\right) \\
&= [\nabla \mathbf{f}(\boldsymbol{\theta})]^T \boldsymbol{\Sigma}^{-1} \underbrace{\mathbb{E}\left((\mathbf{y} - \mathbf{f}(\boldsymbol{\theta}))(\mathbf{y} - \mathbf{f}(\boldsymbol{\theta}))^T\right)}_{\boldsymbol{\Sigma}} \boldsymbol{\Sigma}^{-1} [\nabla \mathbf{f}(\boldsymbol{\theta})] \tag{3.4} \\
&= [\nabla \mathbf{f}(\boldsymbol{\theta})]^T \boldsymbol{\Sigma}^{-1} [\nabla \mathbf{f}(\boldsymbol{\theta})] \,. \tag{3.5}
\end{aligned}$$

In this work, we assume that the noise statistics do not vary appreciably across datasets, $i.e.$, for a fixed variance σ^2, $\sigma_m^2 = \sigma^2$, $\forall m \in \{1, \ldots, M\}$. With this assumption, $\boldsymbol{\Sigma}$ simplifies to a scaled identity matrix, $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_M$, and the Fisher information matrix is simply $\mathbf{F}(\boldsymbol{\theta}) = \frac{1}{\sigma^2} [\nabla \mathbf{f}(\boldsymbol{\theta})]^T [\nabla \mathbf{f}(\boldsymbol{\theta})]$.

Finally, the matrix CRLB states that the covariance matrix associated with any unbiased estimator of $\boldsymbol{\theta}$ is bounded below by the inverse of the Fisher information matrix:

$$\mathrm{cov}(\boldsymbol{\theta}) \succeq \mathbf{F}^{-1}(\boldsymbol{\theta}) \,. \tag{3.6}$$

We can therefore bound the covariance matrix of an unbiased estimator for any choice of scan profile using (3.6). We next discuss desirable properties of $\mathbf{F}^{-1}(\boldsymbol{\theta})$, to help guide scan selection.

3.2.2 Heuristics for Scan Selection in QMRI

For QMRI, we are primarily interested in obtaining precise T_1, T_2 estimates; by contrast, the precision of M_0^* estimates, as well as the coupling effects associated with off-diagonal terms of $\mathbf{F}^{-1}(\boldsymbol{\theta})$, are of secondary importance. It is therefore reasonable to select scan profiles that yield small

$$\sigma_{T_1} := \sqrt{[\mathbf{F}^{-1}(\boldsymbol{\theta})]_{2,2}}, \quad \text{and} \tag{3.7}$$

$$\sigma_{T_2} := \sqrt{[\mathbf{F}^{-1}(\boldsymbol{\theta})]_{3,3}}. \tag{3.8}$$

A reasonable objective function to minimize over the (huge) parameter space would thus be

$$\Psi(\sigma_{T_1}, \sigma_{T_2}) = c\,\sigma_{T_1} + \sigma_{T_2}, \tag{3.9}$$

where $c \in [0,1]$ is a constant that controls the relative importance of T_1 versus T_2 estimation.

This objective does not account for the desirability of short scan times. To easily see this, note that repeating a scan τ times increases the Fisher matrix by a factor of τ, thereby reducing σ_{T_1}, σ_{T_2} and Ψ by a factor of $\sqrt{\tau}$. To enable fair comparison between scan profiles of differing duration, we must discourage long scan times. Funai [28] suggested the T_R-Compensated Deviation (TRCD) as a more useful measure of scan profile undesirability, with respect to the pth unknown parameter:

$$\tilde{\sigma}_p = \frac{\sigma_p}{\sigma}\sqrt{\|\mathbf{T_R}\|_1}, \tag{3.10}$$

for $p = T_1, T_2$. Note that, unlike the CRLB, the TRCD is unaffected by increases in scan repetition factor τ. A natural modification of (3.9) yields a more useful measure of scan undesirability:

$$\tilde{\Psi}(\sigma_{T_1}, \sigma_{T_2}) = \frac{1}{\sigma}\sqrt{\|\mathbf{T_R}\|_1}\,(c\,\sigma_{T_1} + \sigma_{T_2}). \tag{3.11}$$

Hereafter, we focus on optimizing $\tilde{\Psi}$ to find a desirable scan profile.

3.3 Optimal Scan Selection from SPGR and DESS Pulse Sequences

3.3.1 A Min-max Optimization Problem for Scan Selection

For a given scan profile, we seek a suitable set of controllable parameters that minimize the objective function $\tilde{\Psi}$ over a range of T_1, T_2 values typical of grey matter (GM) and white matter (WM) regions of interest. Mathematically, this can be expressed as a min-max optimization problem: we wish to *minimize* over the scan parameters $\boldsymbol{\alpha}, \mathbf{T_R}, \mathbf{T_E}$ the worst-case, *i.e., maximum* objective function viewed over the closed, convex parameter space of interest, $(T_1, T_2) \in \{\mathcal{T}_1 \times \mathcal{T}_2\}$:

$$\{\boldsymbol{\alpha}^*, \mathbf{T_R}^*, \mathbf{T_E}^*\} = \underset{\substack{\boldsymbol{\alpha} \in \mathcal{A}^M \\ \mathbf{T_R} \in \mathcal{T}_R^M \\ \mathbf{T_E} \in \mathcal{T}_E^M}}{\arg\min} \underset{\substack{T_1 \in \mathcal{T}_1 \\ T_2 \in \mathcal{T}_2}}{\max} \tilde{\Psi}\left(\sigma_{T_1}(\boldsymbol{\theta}; \boldsymbol{\alpha}, \mathbf{T_R}, \mathbf{T_E}), \sigma_{T_2}(\boldsymbol{\theta}; \boldsymbol{\alpha}, \mathbf{T_R}, \mathbf{T_E})\right), \tag{3.12}$$

where M_0^* is held fixed at unity. In practice, it is desirable for the scan profile to yield low-variance T_1, T_2 estimates, despite small perturbations of the selected scan parameters, $\{\boldsymbol{\alpha}^*, \mathbf{T_R^*}, \mathbf{T_E^*}\}$. If $\delta_\alpha, \delta_R, \delta_E \in [0, 1]$ denote tolerance levels associated with α, T_R, T_E, respectively, then

$$\tilde{\Psi}^{\text{worst}} := \max_{\substack{\boldsymbol{\alpha} \in [(1-\delta_\alpha)\boldsymbol{\alpha}^*, (1+\delta_\alpha)\boldsymbol{\alpha}^*] \\ \mathbf{T_R} \in [(1-\delta_R)\mathbf{T_R^*}, (1+\delta_R)\mathbf{T_R^*}] \\ \mathbf{T_E} \in [(1-\delta_E)\mathbf{T_E^*}, (1+\delta_E)\mathbf{T_E^*}]}} \max_{\substack{T_1 \in \mathcal{T}_1 \\ T_2 \in \mathcal{T}_2}} \tilde{\Psi}\left(\sigma_{T_1}(\boldsymbol{\theta}; \boldsymbol{\alpha}, \mathbf{T_R}, \mathbf{T_E}), \sigma_{T_2}(\boldsymbol{\theta}; \boldsymbol{\alpha}, \mathbf{T_R}, \mathbf{T_E})\right) \qquad (3.13)$$

is a useful measure of the scan profile's robustness. Hereafter, we make reasonable assumptions to approximate $\tilde{\Psi}^{\text{worst}}$ for several candidate scan profiles for joint T_1, T_2 estimation.

3.3.2 Scan Profile Choices

We perform a CRLB analysis on the following scan combinations, to assess their T_1, T_2 estimation performance: (1,1), (0,2), and (2,1), where the coordinate (S_s, S_d) denotes the number of (SPGR, DESS) scans in a profile. We motivate these particular profiles with the following observations:

- A minimum $M \geq P(= 3)$ datasets are necessary for well-conditioned estimation.

- Each DESS scan provides two datasets (2.26), (2.27). Thus, $M = S_s + 2S_d$.

- Practically, it is highly desirable to select as short a scan profile as possible.

- For T_2 estimation, we desire significant T_2 contrast across the datasets. The SPGR signal model (1.7) reveals little T_2 dependence, so at least 1 DESS dataset is desirable in the profile.

3.3.3 Simplifications and Assumptions

Because $\tilde{\Psi}$ is non-convex, gradient-based minimization techniques can only hope to arrive at local minima. To examine global optima, we perform an exhaustive search. In its general form, (3.12) describes an optimization over a $(3(S_s + S_d) + (P - 1))$-dimensional parameter space. To ease the computational burden, we impose two added constraints on the search space:

- We globally hold T_E fixed at 5ms. For short-T_R experiments, there is not much flexibility in T_E selection. Also, in (1.7) and (2.27), all T_E dependence is absorbed into nuisance parameter M_0^*. We thus hypothesize that T_E variation will not appreciably change the T_1 and T_2 CRLB.

- For a given scan profile, we hold T_R fixed across repetitions of SPGR or DESS sequences. For example, in the (2,1) profile, we perform a two-dimensional search across repetition times: one degree of freedom for the two SPGR scans, and the other for the single DESS scan.

The dimension d of the search space reduces to $(S_s + S_d) + (\mathbb{1}_{\{S_s > 0\}} + \mathbb{1}_{\{S_d > 0\}}) + (P - 1)$ with these constraints. For our choice of scan profiles, $S_s \leq 2, S_d \leq 2, P = 3$, so d never exceeds 7.

Several constants have yet to be specified for the following results:

- We set the T_1 vs. T_2 importance parameter to $c = 0.1$. Since T_1 values are $\sim$10 times greater than T_2 values, this corresponds to placing roughly equal importance on T_1, T_2 estimation.

- We select flip angle parameters to range from $\mathcal{A} = [5°, 85°]$.

- We choose repetition times to range from $\mathcal{T}_R = [8, 16]$ms for SPGR scans and $\mathcal{T}_R = [10, 20]$ms for DESS scans. The lower ends of these ranges reflect the physical limitations of MR hardware. We choose the upper ends to prevent scan times from being undesirably long.

- The unknown parameter space is taken to be $\mathcal{T}_1 = [400, 1000]$ms and $\mathcal{T}_2 = [60, 100]$ms, in correspondence with WM and GM voxels of interest.

- We select the controlled scan parameter tolerance thresholds to be $(\delta_\alpha, \delta_R) = (0.2, 0.2)$, corresponding to the reasonable assumption that both scan parameters can be specified to within 20% of their nominal value. Since T_E was held fixed, we do not need to specify δ_E.

Lastly, we highlight that the noise variance σ^2 need not be specified in this analysis, because the TRCD modifies the objective function $\tilde{\Psi}$ to yield costs (in ms) which are independent of σ.

3.3.4 CRLB Methods and Discussion

(a) $\log \tilde{\sigma}_{T_1}$ vs. $(\alpha_{\mathrm{SPGR}}, \alpha_{\mathrm{DESS}})$ (b) $\log \tilde{\sigma}_{T_2}$ vs. $(\alpha_{\mathrm{SPGR}}, \alpha_{\mathrm{DESS}})$ (c) $\log \tilde{\Psi}$ vs. $(\alpha_{\mathrm{SPGR}}, \alpha_{\mathrm{DESS}})$
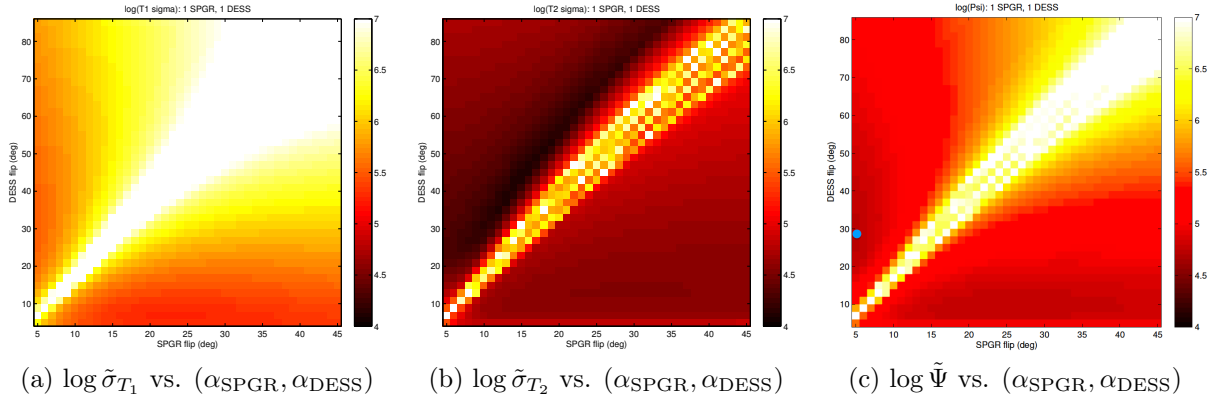
Figure 3.1: Plots of (a) $\log \tilde{\sigma}_{T_1}$, (b) $\log \tilde{\sigma}_{T_2}$, and (c) $\log \tilde{\Psi}$ versus flip angles, for the (1 SPGR, 1 DESS) scan profile. Each pixel represents the worst-case value observed over $\mathcal{T}_1 \times \mathcal{T}_2$. Other optimization dimensions $T_{R,\mathrm{SPGR}}$ and $T_{R,\mathrm{DESS}}$ are held fixed, such that the global minimum of $\tilde{\Psi}$ is displayed.

Figures 3.1, 3.2, and 3.3, summarize some of the most interesting behavior of TRCD values $\tilde{\sigma}_{T_1}, \tilde{\sigma}_{T_2}$ and objective function $\tilde{\Psi}$, with respect to the controllable parameter space. We find empirically that the effect of varying flip angles dominates the effect of varying T_R values. For this reason, the displayed heat-maps vary pairs of flip angles, but hold relevant T_R values fixed to their value at the global optimum. All heat-maps are presented on the same (base 10) logarithmic scale.

Figure 3.1 shows a projection of the results for the (1,1) scan profile, where $(T_{R,\mathrm{SPGR}}, T_{R,\mathrm{DESS}})$ are held fixed at the maximum tested values of $(16, 20)$ms, for a total scan time of 36ms per phase encode line. While $\tilde{\sigma}_{T_1}$ smoothly degrades as flip angles increase, $\tilde{\sigma}_{T_2}$ shows undesirably high sensitivity near its global minimum of $(18, 51)^\circ$. To avoid this sensitivity, $\tilde{\Psi}$ selects $(5, 29)^\circ$ as the optimal flip angle pair. However, as seen in Table 3.1, the corresponding $\tilde{\Psi} = 10^{4.74}$ and $\tilde{\Psi}^{\mathrm{worst}} = 10^{4.84}$ are considerably larger than corresponding values of other scan profiles. To explain this, we observe that while the (1,1) scan profile affords the fastest scan times, it only yields data at $M = 3$ contrasts. Unsurprisingly, fewer datasets correspond to higher estimation variance.

Scan	$T_{R,\mathrm{SPGR}}$	$T_{R,\mathrm{DESS}}$	α_{SPGR}	α_{DESS}	$\log\tilde\sigma_{T_1}$	$\log\tilde\sigma_{T_2}$	$\log\tilde\Psi$	$\log\tilde\sigma_{T_1}^{\mathrm{worst}}$	$\log\tilde\sigma_{T_2}^{\mathrm{worst}}$	$\log\tilde\Psi^{\mathrm{worst}}$
(1,1)	16	20	5°	29°	5.53	4.33	4.74	5.67	4.43	4.84
(2,1)	(16,16)	10	(5,22)°	47°	**5.32**	**4.15**	**4.54**	5.39	**4.19**	**4.59**
(0,2)	–	(20,20)	–	(11,37)°	5.34	4.21	4.58	**5.37**	4.26	4.61

Table 3.1: Optimized performance summary of different scan profiles. The first column denotes the number of (SPGR, DESS) scans in the profile. The next four columns describe optimized parameters, as computed through a min-max optimization similar to (3.12); the corresponding TRCD and objective function values are given in the following three columns. The last three columns summarize the robustness of scan selection, indicating worst-case performance degradation when parameters deviate up to 20% from their nominal optimal values. Repetition times and flip angles are reported in (ms) and (degrees), respectively, while $\log\tilde\sigma$ and $\log\tilde\Psi$ values are reported in (log ms).

Figure 3.2 shows three projections of the results for the (2,1) scan profile. Interestingly, the minimum $\tilde\Psi$ value occurs at the *maximum* tested $T_{R,\mathrm{SPGR}}$ value of 16ms and *minimum* tested $T_{R,\mathrm{DESS}}$ value of 10ms, for a total scan time of 42ms per phase encode line. Compared to the (1,1) profile, adding a second SPGR scan allows for lower estimation variance (particularly in T_2), for certain choices of $\boldsymbol{\alpha}$. However, certain σ_{T_2} local minima are sensitive to fluctuations in $\boldsymbol{\alpha}$. The objective function $\tilde\Psi$ balances the sometimes conflicting trends in σ_{T_1} versus σ_{T_2} variation, and finds the global optimum at $(5, 22, 47)°$. Table 3.1 affirms that $\tilde\Psi = 10^{4.54}$ is 34% lower than the corresponding value for the (1,1) scan profile, but degrades a substantial 11.7% to $\tilde\Psi^{\mathrm{worst}} = 10^{4.59}$.

Figure 3.3 shows a projection of the results for the (0,2) scan profile, where $T_{R,\mathrm{DESS}}$ is held fixed at the maximum tested value of 20ms, for a total scan time of 40ms per phase encode line. Minimizing values of $\tilde\sigma_{T_1}, \tilde\sigma_{T_2}$, and $\tilde\Psi$ are all observed along the diagonal, corresponding to two identical DESS flip angles. These values are discontinuities in an otherwise smooth heat-map; as they are not robust to fluctuations, they are ignored for parameter selection. Aside from the diagonal, the (0,2) scan exhibits smooth degradation in both $\tilde\sigma_{T_1}$ and $\tilde\sigma_{T_2}$. Because the TRCD maps also exhibit similar shapes, the $\tilde\Psi$-optimal $\boldsymbol{\alpha}^* = (11, 37)°$ is desirably close to the $\tilde\sigma_{T_1}$ and $\tilde\sigma_{T_2}$ local minima at $(11, 43)°$ and $(13, 33)°$, respectively. Table 3.1 notes that while the optimal $\tilde\Psi = 10^{4.58}$ is 8.1% higher than the corresponding value for the (2,1) profile, it gracefully degrades by only 6.7% to $\tilde\Psi^{\mathrm{worst}} = 10^{4.61}$. As a result, we observe a marginal 3.1% difference between the (0,2) and (2,1) worst-case costs.

For well-selected parameters, our analysis suggests that, in a $\tilde\Psi^{\mathrm{worst}}$-sense, collection of $M = 4$ datasets using either (2,1) or (0,2) SPGR and DESS scans will yield roughly equal unbiased estimation variance. We accordingly turn to qualitative reasoning for scan selection:

- The (0,2) scan profile yields local minima in $\tilde\sigma_{T_1}$ and $\tilde\sigma_{T_2}$ for similar sets of parameters, meaning that a broad range of c values will yield suitable $\boldsymbol{\alpha}^*$.

- The selected (0,2) scan profile is 2ms (4.8%) shorter than the best possible (2,1) scan profile, per phase encode. For a typically-sized $256 \times 256 \times 32$ problem, this yields a 16.38s saving.

- Most importantly, the (0,2) scan profile is most robust to performance degradation, over a broad range of flip angle and $T_{R,\mathrm{DESS}}$ choices.

We conclude that a scan profile consisting of **two DESS scans**, at $\boldsymbol{\alpha}^* = (11, 37)°$ and $T_R = 20$ms, will yield low variance for any unbiased estimator of T_1 and T_2 over broad ranges of interest. Furthermore, even if true scan parameters deviate from the nominal choices by 20%, the estimation

variance will remain low. Hereafter, we employ this scan profile for joint T_1, T_2 estimation. Though in practice, we will add regularization for biased estimation, we assume that our robustness considerations will still ensure that the chosen scan parameters are effective, if not strictly optimal. Future work could be to explore bounds for biased estimators [29].

(a) $\log \tilde{\sigma}_{T_1}$ vs. $(\alpha_{\mathrm{SPGR},1}, \alpha_{\mathrm{DESS}})$ (b) $\log \tilde{\sigma}_{T_2}$ vs. $(\alpha_{\mathrm{SPGR},1}, \alpha_{\mathrm{DESS}})$ (c) $\log \tilde{\Psi}$ vs. $(\alpha_{\mathrm{SPGR},1}, \alpha_{\mathrm{DESS}})$

(d) $\log \tilde{\sigma}_{T_1}$ vs. $(\alpha_{\mathrm{SPGR},2}, \alpha_{\mathrm{DESS}})$ (e) $\log \tilde{\sigma}_{T_2}$ vs. $(\alpha_{\mathrm{SPGR},2}, \alpha_{\mathrm{DESS}})$ (f) $\log \tilde{\Psi}$ vs. $(\alpha_{\mathrm{SPGR},2}, \alpha_{\mathrm{DESS}})$

(g) $\log \tilde{\sigma}_{T_1}$ vs. $(\alpha_{\mathrm{SPGR},1}, \alpha_{\mathrm{SPGR},2})$ (h) $\log \tilde{\sigma}_{T_2}$ vs. $(\alpha_{\mathrm{SPGR},1}, \alpha_{\mathrm{SPGR},2})$ (i) $\log \tilde{\Psi}$ vs. $(\alpha_{\mathrm{SPGR},1}, \alpha_{\mathrm{SPGR},2})$
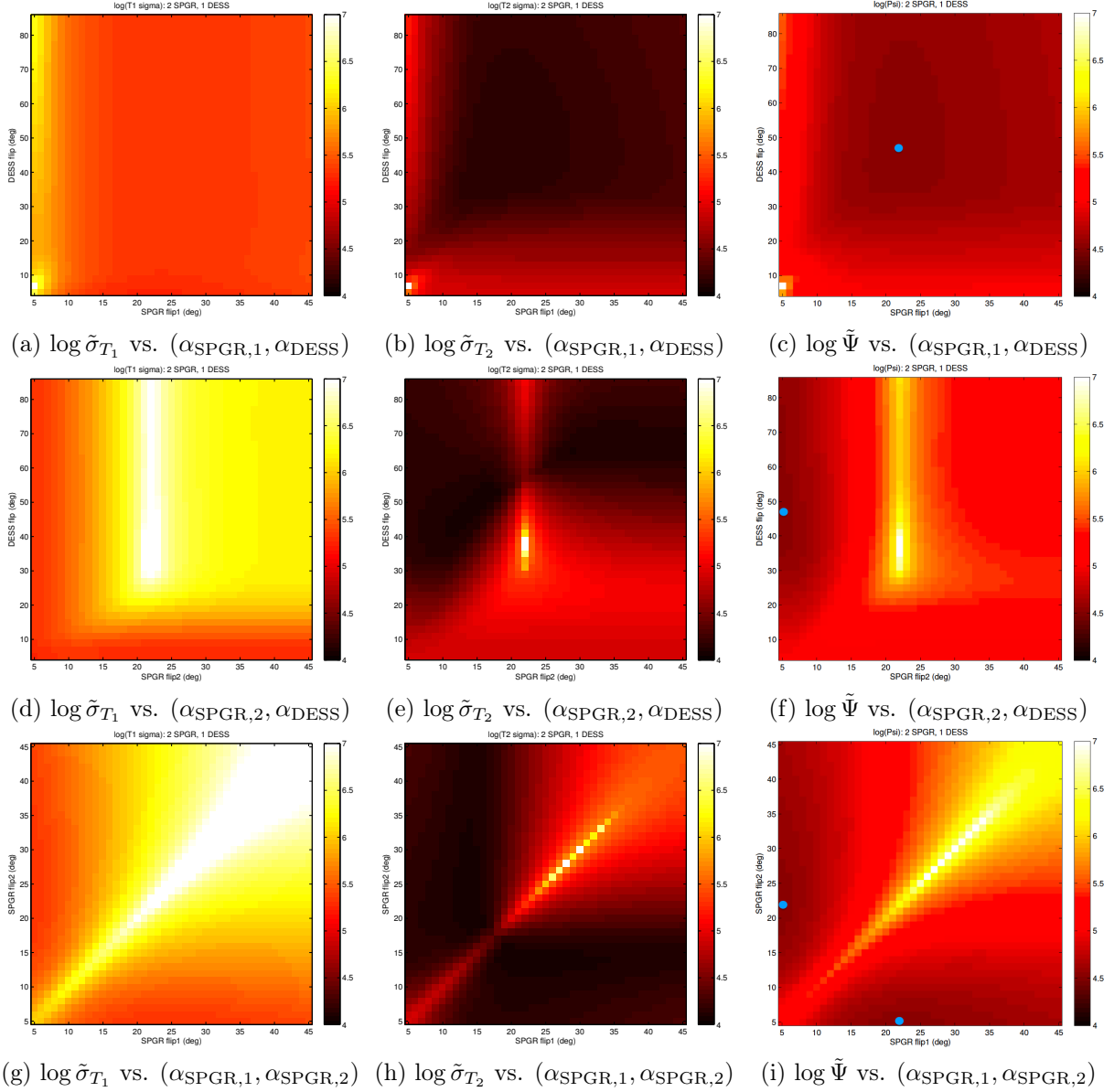
Figure 3.2: Plots of (left) $\log \tilde{\sigma}_{T_1}$, (center) $\log \tilde{\sigma}_{T_2}$, and (right) $\log \tilde{\Psi}$ versus combinations of flip angles, for the (2 SPGR, 1 DESS) scan profile. In each heat map, only two of the three flip angles are varied; by row, (top) $\alpha_{\mathrm{SPGR},\,2}$, (center) $\alpha_{\mathrm{SPGR},\,1}$, and (bottom) α_{DESS} are excluded. Each pixel represents the worst-case value observed over $\mathcal{T}_1 \times \mathcal{T}_2$. Other optimization dimensions $T_{R,\mathrm{SPGR}}$ and $T_{R,\mathrm{DESS}}$ are held fixed, such that the global minimum of $\tilde{\Psi}$ is displayed.
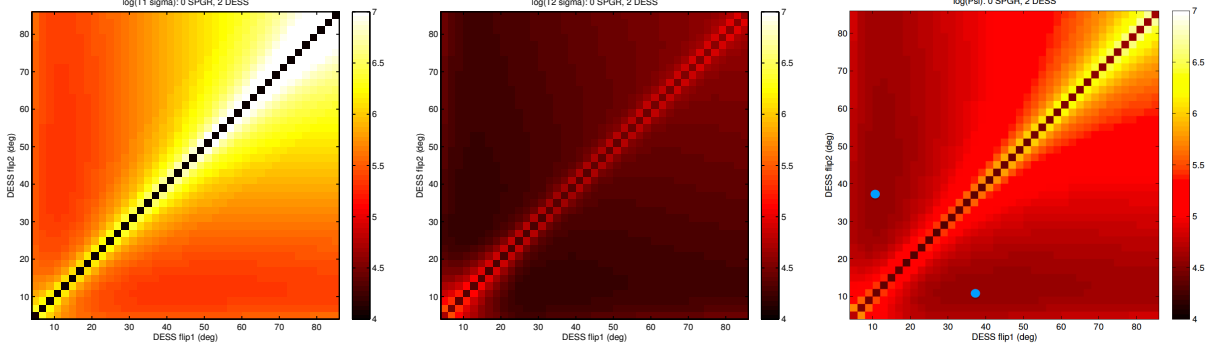
(a) $\log \tilde{\sigma}_{T_1}$ vs. $(\alpha_{\mathrm{DESS},1}, \alpha_{\mathrm{DESS},2})$ (b) $\log \tilde{\sigma}_{T_2}$ vs. $(\alpha_{\mathrm{DESS},1}, \alpha_{\mathrm{DESS},2})$ (c) $\log \tilde{\Psi}$ vs. $(\alpha_{\mathrm{DESS},1}, \alpha_{\mathrm{DESS},2})$

Figure 3.3: Plots of (a) $\log \tilde{\sigma}_{T_1}$, (b) $\log \tilde{\sigma}_{T_2}$, and (c) $\log \tilde{\Psi}$ versus flip angles, for the (0 SPGR, 2 DESS) scan profile. Each pixel represents the worst-case value observed over $\mathcal{T}_1 \times \mathcal{T}_2$. Other optimization dimension $T_{R,\mathrm{DESS}}$ is held fixed, such that the global minimum of $\tilde{\Psi}$ is displayed.

3.4 Joint Estimation of T_1, T_2 from Two DESS Sequences

Thus far, we have investigated the problem of scan design for optimized T_1, T_2 estimation. For the remainder of the chapter, we transition to the problem of jointly estimating T_1 and T_2 maps with high accuracy and precision, from the two-DESS scan profile we just designed.

3.4.1 Dictionary-based Initialization with the Variable Projection Method

Because data-fit terms associated with DESS sequences are non-convex (recall Figure 2.3), it is important to obtain reasonable initializations for iterative reconstruction. Though Welsch [18] proposed a (biased) T_2 estimator from DESS data, there is, to our knowledge, no literature describing a method for T_1 estimation from DESS. Indeed, even with two DESS datasets, it is difficult to massage the (noiseless) signal model into a form that isolates T_1 without unreasonable approximations.

We instead describe a fast method for dictionary-based T_1, T_2 initialization. This "variable-projection" method [30] exploits the partial linearity of the DESS signal model to factor out M_0^* effects and focus only on T_1, T_2 estimation. We later obtain an M_0^* initialization through least-squares. The resulting algorithm resembles a matched-filter, relying only on inner products with dictionary elements.

At each voxel, we wish to solve the following nonlinear least-squares (NLS) optimization problem:

$$\left\{ \widehat{M_0^*}, \widehat{T}_1, \widehat{T}_2 \right\} = \underset{M_0^*, T_1, T_2}{\arg\min} \frac{1}{2} \left\| \mathbf{y} - \mathbf{f}(M_0^*, T_1, T_2) \right\|_2^2 , \tag{3.14}$$

where $\mathbf{y} = [y_1^-, \ldots, y_L^-, y_1^+, \ldots, y_L^+]^T \in \mathbb{C}^{2L}$, with $y_l^\pm$ defined in (2.31); $\mathbf{f} = [f_1^-, \ldots, f_L^-, f_1^+, \ldots, f_L^+]^T :$ $\mathbb{C} \times \mathbb{R} \times \mathbb{R} \mapsto \mathbb{C}^{2L}$, with f_l^- and f_l^+ defined in (2.33) and (2.34), respectively; and l indexes a total L flip angles, as before. Note that $\mathbf{f}$ has been redefined from earlier in the chapter; for this case of a DESS-only sequence, we have $M = 2L$ datasets. Since $\mathbf{f}$ exhibits a linear dependence on M_0^*, we can decompose $\mathbf{f}(M_0^*, T_1, T_2) =: M_0^* \mathbf{A}(T_1, T_2)$, in which case the NLS problem can be rewritten as

Algorithm 2 Dictionary-Based T_1, T_2 Estimation using Variable Projection Method

---

Input: Data for all $j = 1, \ldots, N$ voxels, $\mathbf{Y} := [\mathbf{y}_1, \ldots, \mathbf{y}_N]$; Dictionary $\mathbf{D} := [\mathbf{d}_1, \ldots, \mathbf{d}_K]$

Output: Initializations $\left\{ \widehat{M}_{0,j}^*, \widehat{T}_{1,j}, \widehat{T}_{2,j} \right\}$, for $j = 1, \ldots, N$ voxels

for $k = 1, \ldots, K$ **do**

    Store $\mathbf{A}(T_1, T_2)_k \leftarrow \mathbf{d}_k$

    Compute $\tilde{\mathbf{Y}}_k = \left(\mathbf{A}(T_1, T_2)_k' \mathbf{A}(T_1, T_2)_k \right)^{-\frac{1}{2}} \mathbf{A}(T_1, T_2)_k' \mathbf{Y}$

    Store $\|\tilde{\mathbf{y}}_{k,j}\|_2$ for $j = 1, \ldots, N$

end for

for $j = 1, \ldots, N$ **do**

    $\widehat{k}_j = \arg\max_k \|\tilde{\mathbf{y}}_{k,j}\|_2$

    $\left\{ \widehat{T}_{1,j}, \widehat{T}_{2,j} \right\} \leftarrow \mathrm{LUT}(\widehat{k}_j)$, where LUT denotes a table look-up

    Compute $\widehat{M}_{0,j}^*(\widehat{T}_{1,j}, \widehat{T}_{2,j})$ using (3.16)

end for

$$\left\{ \widehat{M}_0^*, \widehat{T}_1, \widehat{T}_2 \right\} = \underset{M_0^*, T_1, T_2}{\arg\min} \frac{1}{2} \|\mathbf{y} - \mathbf{A}(T_1, T_2) M_0^*\|_2^2 \,, \tag{3.15}$$

Now, we can exploit the partially linear structure to simplify the joint estimation problem. For any fixed $\mathbf{T} := (T_1, T_2)$, the optimal $\widehat{M}_0^*$ can be expressed as the solution to a linear LS problem:

$$\widehat{M}_0^*(\mathbf{T}) = \left(\mathbf{A}(\mathbf{T})' \mathbf{A}(\mathbf{T}) \right)^{-1} \mathbf{A}(\mathbf{T})' \mathbf{y} \tag{3.16}$$

Substituting this LS solution back into (3.15) and simplifying, we obtain a simpler problem:

$$\begin{aligned}
\left\{ \widehat{T}_1, \widehat{T}_2 \right\} &= \underset{T_1, T_2}{\arg\min} \frac{1}{2} \left\| \mathbf{y} - \mathbf{A}(\mathbf{T}) \left(\mathbf{A}(\mathbf{T})' \mathbf{A}(\mathbf{T}) \right)^{-1} \mathbf{A}(\mathbf{T})' \mathbf{y} \right\|_2^2 \\
&= \underset{T_1, T_2}{\arg\min} \frac{1}{2} \left\| \underbrace{\left(\mathbf{I}_{2L} - \mathbf{A}(\mathbf{T}) \left(\mathbf{A}(\mathbf{T})' \mathbf{A}(\mathbf{T}) \right)^{-1} \mathbf{A}(\mathbf{T})' \right)}_{\mathbf{P}_{\mathfrak{N}(\mathbf{A}(\mathbf{T})')}} \mathbf{y} \right\|_2^2 \tag{3.17} \\
&= \underset{T_1, T_2}{\arg\min} \, \mathbf{y}' \mathbf{P}_{\mathfrak{N}(\mathbf{A}(\mathbf{T})')}' \mathbf{P}_{\mathfrak{N}(\mathbf{A}(\mathbf{T})')} \mathbf{y} \\
&= \underset{T_1, T_2}{\arg\min} \, \mathbf{y}' \left(\mathbf{I}_{2L} - \mathbf{A}(\mathbf{T}) \left(\mathbf{A}(\mathbf{T})' \mathbf{A}(\mathbf{T}) \right)^{-1} \mathbf{A}(\mathbf{T})' \right) \mathbf{y} \tag{3.18} \\
&= \underset{T_1, T_2}{\arg\max} \, \mathbf{y}' \mathbf{A}(\mathbf{T}) \left(\mathbf{A}(\mathbf{T})' \mathbf{A}(\mathbf{T}) \right)^{-1} \mathbf{A}(\mathbf{T})' \mathbf{y} \,, \tag{3.19}
\end{aligned}$$

where in (3.17), we introduce, for sake of clarity, the symmetric projection matrix $\mathbf{P}_{\mathfrak{N}(\mathbf{A}(\mathbf{T})')}$ onto the orthocomplement of the columnspace of $\mathbf{A}(\mathbf{T})$. In (3.18), we save some computation by exploiting the idempotency of projection operators. Note that because the $\mathbf{A}(\mathbf{T})' \mathbf{A}(\mathbf{T})$ is 1×1, the matrix inversion is trivial, and computation of such inner products as in (3.19) is extremely fast.

Algorithm 2 summarizes how we use the variable projection method for dictionary-based T_1, T_2 initialization. We first precompute and store a K-element dictionary $\mathbf{D} \in \mathbb{C}^{2L \times K}$ offline, whose kth column contains the $2L$ anticipated signals $\mathbf{A}(T_1, T_2)_k$ from the kth nominal $(T_1, T_2)_k$ pair. Then,

for every voxel, we compute (3.19) for each of the K dictionary elements, and select the (T_1, T_2) pair corresponding to the largest inner product. Lastly, M_0^* is initialized using (3.16).

Unlike previously-discussed initialization schemes, this dictionary-based estimator does not suffer from noise amplification due to manipulation of the signal model. In addition, as K is increased, the estimator can shrink its quantization bias (though not its bias due to estimator nonlinearity) to an arbitrarily small quantity. It is reasonable to ask why, if given a large enough dictionary, any further data processing is necessary. We list the most important reasons below:

1. This estimate does not account for (T_1, T_2) correlations between adjacent voxels; thus, edge-preserving regularization will likely improve estimation variance.

2. Off-the-shelf denoising methods like median-filtering are suboptimal. Useful quantization schemes of $\mathcal{T}_1 \times \mathcal{T}_2$ are nonlinear, causing for space-varying residual errors. As post-processing steps like median-filtering do not utilize the underlying data, they have no chance of reconciling these space-varying errors, causing for sub-optimality.

3. In practice, raw k-space data is under-sampled, complicating direct consideration of the image-domain data $\mathbf{y}$ used in this analysis. In this more realistic setting, $\mathbf{y}$ is typically of low-resolution, and iterative reconstruction is essential to extract fine details.

3.4.2 Problem Formulation for Regularized, Joint M_0^*, T_1, T_2 Estimation

We borrow much of the notation introduced in Chapter 2 to formulate the joint M_0^*, T_1, T_2 estimation problem. We wish to estimate $\{\mathbf{x}_{\mathrm{m}}, \mathbf{x}_1, \mathbf{x}_2\} \in \{\mathbb{C}^N \times \mathcal{X}_1 \times \mathcal{X}_2\}$, with j indexing a total of N voxels and jth elements defined as $x_{\mathrm{m},j} := M_{0,j}^*$, $x_{1,j} := \exp\{-T_R/T_{1,j}\}$, and $x_{2,j} := \exp\{-T_R/T_{2,j}\}$, respectively. As before, $\mathbf{y}_l^- := [y_{l,1}^-, \ldots, y_{l,N}^-]^{\mathrm{T}}$ and $\mathbf{y}_l^+ := [y_{l,1}^+, \ldots, y_{l,N}^+]^{\mathrm{T}}$ are noisy measurement vectors, where l indexes a total of L nutation angles; $\epsilon_{l,j}^\pm \sim \mathbb{CN}(0, \sigma^2)$ is complex, white Gaussian noise; and $y_{l,j}^\pm$ is defined in (2.31). For $t \in \{1, 2\}$, $\mathcal{X}_t := [E_{t,\min}, E_{t,\max}]^N \subset (0, 1]^N$ are solution spaces corresponding to ranges $[T_{t,\min}, T_{t,\max}]$ over which we seek accurate estimation. Recall that vector functions $\mathbf{f}_l^\pm(\mathbf{x}_{\mathrm{m}}, \mathbf{x}_1, \mathbf{x}_2) : \mathbb{C}^N \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{C}^N$ are voxel-wise extensions of the scalar functions $f_l^\pm(x_{\mathrm{m}}, x_1, x_2) : \mathbb{C} \times \mathbb{R} \times \mathbb{R} \to \mathbb{C}$, modified from (2.33) and (2.34) to be redefined as

$$f_l^-(x_{\mathrm{m}}, x_1, x_2) = -x_{\mathrm{m}} \tan\left(\frac{\alpha_l}{2}\right) x_2^{-\frac{2T_E}{T_R}} \left[1 - \sqrt{\frac{1 - x_2^2}{1 - x_2^2/v_{1,l}^2}}\right]; \tag{3.20}$$

$$f_l^+(x_{\mathrm{m}}, x_1, x_2) = x_{\mathrm{m}} \tan\left(\frac{\alpha_l}{2}\right) \left[1 - \frac{1}{v_{1,l}} \sqrt{\frac{1 - x_2^2}{1 - x_2^2/v_{1,l}^2}}\right], \tag{3.21}$$

where $v_{1,l} := \frac{1 - x_1 \cos\alpha_l}{x_1 - \cos\alpha_l}$. An ML joint estimator for $\{\mathbf{x}_{\mathrm{m}}, \mathbf{x}_1, \mathbf{x}_2\}$ is then given by minimizing

$$\Psi_{\mathrm{ML}}(\mathbf{x}_{\mathrm{m}}, \mathbf{x}_1, \mathbf{x}_2) = \frac{1}{2} \sum_{l=1}^{L} \left[\left\|\mathbf{y}_l^- - \mathbf{f}_l^-(\mathbf{x}_{\mathrm{m}}, \mathbf{x}_1, \mathbf{x}_2)\right\|_2^2 + \left\|\mathbf{y}_l^+ - \mathbf{f}_l^+(\mathbf{x}_{\mathrm{m}}, \mathbf{x}_1, \mathbf{x}_2)\right\|_2^2\right].$$

In addition to the data, we assume that M_0^*, T_1, T_2 maps tend to be piecewise-smooth; that is, adjacent voxels are likely to contain similar tissue types, and are thus likely to have similar $M_0^*, T_1,$

and T_2 values. We therefore extend the ML cost function into a reasonable MAP objective function through inclusion of edge-preserving roughness penalties:

$$\Psi_{\text{MAP}}(\mathbf{x}_{\text{m}}, \mathbf{x}_1, \mathbf{x}_2) = \Psi_{\text{ML}}(\mathbf{x}_{\text{m}}, \mathbf{x}_1, \mathbf{x}_2) + \beta_{\text{m}} R(\mathbf{x}_{\text{m}}) + \beta_1 R(\mathbf{x}_1) + \beta_2 R(\mathbf{x}_2), \quad (3.22)$$

where, as before, $R(\mathbf{x}) := \sum_{k=1}^{K} \phi_k([\mathbf{Cx}]_k)$; $\mathbf{C}$ is a 2D or 3D spatial finite differencing matrix; ϕ_k is a convex, edge-preserving potential function; and k indexes a total K differencing operations. Observe that, in general, different regularization parameters $\beta_{\text{m}}, \beta_1$, and β_2 are required to control the levels to which roughness in $\mathbf{x}_{\text{m}}, \mathbf{x}_1$, and $\mathbf{x}_2$ are penalized, respectively. We then obtain penalized-likelihood estimates of $\{\mathbf{x}_{\text{m}}, \mathbf{x}_1, \mathbf{x}_2\}$ by solving the (non-convex) constrained optimization problem:

$$\{\widehat{\mathbf{x}}_{\text{m}}, \widehat{\mathbf{x}}_1, \widehat{\mathbf{x}}_2\} = \underset{\mathbf{x}_{\text{m}} \in \mathbb{C}^N, \mathbf{x}_1 \in \mathcal{X}_1, \mathbf{x}_2 \in \mathcal{X}_2}{\arg \min} \Psi_{\text{MAP}}(\mathbf{x}_{\text{m}}, \mathbf{x}_1, \mathbf{x}_2). \quad (3.23)$$

3.4.3 Minimization Techniques

We observe that the system model (2.31) is linear in $\mathbf{x}_{\text{m}}$, but nonlinear in $\mathbf{x}_1$ and $\mathbf{x}_2$. We exploit this partially linear structure and describe below a simple iterative algorithm for solving (3.23) using an alternating minimization approach. This method, denoted MBERP, resembles MBET1 in that it updates one unknown at a time while holding other unknowns fixed. MBERP also utilizes MBET2 for one of its three inner minimization routines. MBERP tends to decrease the non-convex cost Ψ_{MAP} to a local minimum, though this tendency is not guaranteed because the $\mathbf{x}_1$ update presented here is non-monotonic. MBERP can modified easily to enjoy convergence guarantees by replacing this $\mathbf{x}_1$ update with convergent (but noticeably slower) first-order alternatives. Short descriptions of the variable updates are presented below. MBERP is summarized in Algorithm 3.

Updating the M_0^* Estimate

In the following, we use i to index inner iterations, and $n = 1, \ldots, T$ to index outer iterations. To avoid unnecessarily cumbersome notation, we restart inner iteration indexing at $i = 1$ at the beginning of each variable update. For the $\mathbf{x}_{\text{m}}$ update, we hold $\mathbf{x}_1^{(n)}$ and $\mathbf{x}_2^{(n)}$ fixed and minimize a simplified, but equivalent objective function:

$$\widehat{\mathbf{x}}_{\text{m}}^{(n+1)} = \underset{\mathbf{x}_{\text{m}} \in \mathbb{C}^N}{\arg \min} \Psi_{\text{m}}^{(n)}(\mathbf{x}_{\text{m}}), \quad \text{where}$$

$$\Psi_{\text{m}}^{(n)}(\mathbf{x}_{\text{m}}) := \frac{1}{2} \sum_{l=1}^{L} \left[\left\| \mathbf{y}_l^- - \mathbf{A}_l^- \mathbf{x}_{\text{m}} \right\|_2^2 + \left\| \mathbf{y}_l^+ - \mathbf{A}_l^+ \mathbf{x}_{\text{m}} \right\|_2^2 \right] + \beta_{\text{m}} R(\mathbf{x}_{\text{m}}), \quad (3.24)$$

and the system matrices $\mathbf{A}_l^{\pm}$ are $\mathbb{C}^{N \times N}$ diagonal, with entries

$$a_l^{\pm} = \frac{\partial}{\partial x_{\text{m}}} f_l^{\pm}(x_{\text{m}}, x_1^{(n)}, x_2^{(n)}). \quad (3.25)$$

Numerous methods exist to solve such unconstrained, linear optimization problems. We employ a (diagonally) preconditioned gradient descent (PGD) update, given by

$$\widehat{\mathbf{x}}_{\mathrm{m}}^{(i+1)} = \mathbf{x}_{\mathrm{m}}^{(i)} - s_{\mathrm{m}} \left[\mathbf{D}_{\mathrm{m}}^{(n)}\right]^{-1} \left[\left(\sum_{l=1}^{L} \begin{bmatrix} \mathbf{A}_l^- \\ \mathbf{A}_l^+ \end{bmatrix}' \left(\begin{bmatrix} \mathbf{A}_l^- \\ \mathbf{A}_l^+ \end{bmatrix} \mathbf{x}_{\mathrm{m}}^{(i)} - \begin{bmatrix} \mathbf{y}_l^- \\ \mathbf{y}_l^+ \end{bmatrix}\right)\right) + \beta_{\mathrm{m}} \nabla R(\mathbf{x}_{\mathrm{m}}^{(i)})\right], \text{ where } (3.26)$$

$$\mathbf{D}_{\mathrm{m}}^{(n)} = \left(\sum_{l=1}^{L} \begin{bmatrix} \mathbf{A}_l^- \\ \mathbf{A}_l^+ \end{bmatrix}' \begin{bmatrix} \mathbf{A}_l^- \\ \mathbf{A}_l^+ \end{bmatrix}\right) + \beta_1 \operatorname{maxeig}\left\{\nabla^2 R(\mathbf{x}_{\mathrm{m}}^{(n)})\right\} \mathbf{I}_N, \tag{3.27}$$

and s_{m} is an appropriately-selected step size, often selecting using line search techniques. We repeat (3.26) until a convergence criterion is met, and assign the most recent update to $\widehat{\mathbf{x}}_{\mathrm{m}}^{(n+1)}$.

Updating the T_1 Estimate

For the $\mathbf{x}_1$ update, we hold $\mathbf{x}_{\mathrm{m}}^{(n)}$ and $\mathbf{x}_2^{(n)}$ fixed and minimize the simplified objective function:

$$\widehat{\mathbf{x}}_1^{(n+1)} = \arg\min_{\mathbf{x}_1 \in \mathcal{X}_1} \Psi_1^{(n)}(\mathbf{x}_1), \text{ where}$$

$$\Psi_1^{(n)}(\mathbf{x}_1) := \frac{1}{2} \sum_{l=1}^{L} \left[\left\|\mathbf{y}_l^- - \mathbf{f}_l^-(\mathbf{x}_1; \mathbf{x}_{\mathrm{m}}^{(n)}, \mathbf{x}_2^{(n)})\right\|_2^2 + \left\|\mathbf{y}_l^+ - \mathbf{f}_l^+(\mathbf{x}_1; \mathbf{x}_{\mathrm{m}}^{(n)}, \mathbf{x}_2^{(n)})\right\|_2^2\right] + \beta_1 R(\mathbf{x}_1). \tag{3.28}$$

Because bounding this cost function's Hessian $\nabla^2 \Psi_1^{(n)}(\mathbf{x}_1)$ is difficult in general, we avoid methods that involve a preconditioner (such as PGD), but still seek fast updates. The Barzilai-Borwein gradient method (BBGM) [31] is an unconstrained first-order method that does not require preconditioning for fast (but non-monotonic) convergence to a local minimum [32]. We add a simple projection operation onto the closed, convex solution set $\mathcal{X}_1$ to obtain the following iteration:

$$\mathbf{x}_1^{(i+1)} = \mathcal{P}_{\mathcal{X}_1} \left\{\mathbf{x}_1^{(i)} - s_1^{(i)} \nabla \Psi_1^{(n)}(\mathbf{x}_1^{(i)})\right\}, \text{ where} \tag{3.29}$$

$$s_1^{(i)} = \frac{\left\|\mathbf{x}_1^{(i)} - \mathbf{x}_1^{(i-1)}\right\|_2^2}{\langle \mathbf{x}_1^{(i)} - \mathbf{x}_1^{(i-1)}, \nabla \Psi_1^{(n)}(\mathbf{x}_1^{(i)}) - \nabla \Psi_1^{(n)}(\mathbf{x}_1^{(i-1)})\rangle} \tag{3.30}$$

When applied to non-quadratic cost functions such as (3.28), this iteration alone has no convergence guarantees. Raydan [33] proposed the addition of a line search that retains reasonable speed by allowing a controlled amount of non-monotonic behavior. We employ these ideas at every inner iteration, by repeatedly shrinking $s_1^{(i)}$ by a factor $0 < \sigma_1 < 1$ until the condition

$$\Psi_1^{(n)}(\mathbf{x}_1^{(i+1)}) \le \max_{0 \le \lambda \le \Lambda} \Psi_1^{(n)}(\mathbf{x}_1^{(i-\lambda)}) + \gamma \langle \mathbf{x}_1^{(i+1)} - \mathbf{x}_1^{(i)}, \nabla \Psi_1^{(n)}(\mathbf{x}_1^{(i)})\rangle \tag{3.31}$$

is satisfied, where $\Lambda \in \mathbb{Q}^+$ controls how far back the line search examines cost function behavior, and $0 < \gamma \ll 1$ is a "slack" factor that controls how much the cost function can deviate from monotonicity. With the addition of this non-monotonic line search, iteration (3.30) is guaranteed to converge; once a convergence criterion is met, we assign the most recent update to $\widehat{\mathbf{x}}_1^{(n+1)}$.

Algorithm 3 MBERP: Model-Based Estimation of Relaxation Parameters T_1, T_2 from DESS Data

---

Input: Data $\mathbf{y}_1^\pm, \ldots, \mathbf{y}_L^\pm$; Initial guesses $\mathbf{x}_1^{(0)}, \mathbf{x}_2^{(0)}$; Regularization parameters $\beta_\mathrm{m}, \beta_1, \beta_2$
Input: BBGM constants $\sigma_1, \gamma, \Lambda$; Convergence criteria for $\mathbf{x}_\mathrm{m}, \mathbf{x}_1, \mathbf{x}_2$ inner minimizations
Output: Regularized parameter maps $\mathbf{M}_0^*, \mathbf{T}_1, \mathbf{T}_2$
Initialize $\mathbf{x}_\mathrm{m}^{(0)}$ by evaluating (3.16) at every voxel
for $n = 1, \ldots, T$ **do**
   Update system matrices $\mathbf{A}_l^\pm$ with Equation (3.25)
   Update $\mathbf{D}_\mathrm{m}^{(n)}$ with Equation (3.27)
   repeat
      Update $\mathbf{x}_\mathrm{m}$ with Equation (3.26)
   until $\mathbf{x}_\mathrm{m}$ convergence criterion satisfied
   repeat
      Update s_1 with Equation (3.30)
      repeat
         Evaluate candidate update $\mathbf{z}_1$ with Equation (3.29)
         Shrink step size: $s_1 \leftarrow \sigma_1 s_1$
      until Line search criterion (3.31) is satisfied, using $\mathbf{z}_1$
      Update $\mathbf{x}_1 \leftarrow \mathbf{z}_1$
   until $\mathbf{x}_1$ convergence criterion satisfied
   Update $\mathbf{D}_2^{(n)}$ with Equation (3.34)
   repeat
      Update $\mathbf{x}_2$ with Equation (3.33)
   until $\mathbf{x}_2$ convergence criterion satisfied
end for
Embed $\mathbf{x}_\mathrm{m}^{(T)}, \mathbf{x}_1^{(T)}, \mathbf{x}_2^{(T)}$ into $\mathbf{M}_0^*, \mathbf{T}_1, \mathbf{T}_2$

Updating the T_2 Estimate

For the $\mathbf{x}_2$ update, we hold $\mathbf{x}_\mathrm{m}^{(n)}$ and $\mathbf{x}_1^{(n)}$ fixed and minimize the simplified objective function:

$$\widehat{\mathbf{x}}_2^{(n+1)} = \arg\min_{\mathbf{x}_2 \in \mathcal{X}_2} \Psi_2^{(n)}(\mathbf{x}_2), \quad \text{where}$$

$$\Psi_2^{(n)}(\mathbf{x}_2) := \frac{1}{2} \sum_{l=1}^{L} \left[\left\| \mathbf{y}_l^- - \mathbf{f}_l^-(\mathbf{x}_2; \mathbf{x}_\mathrm{m}^{(n)}, \mathbf{x}_1^{(n)}) \right\|_2^2 + \left\| \mathbf{y}_l^+ - \mathbf{f}_l^+(\mathbf{x}_2; \mathbf{x}_\mathrm{m}^{(n)}, \mathbf{x}_1^{(n)}) \right\|_2^2 \right] + \beta_2 R(\mathbf{x}_2). \quad (3.32)$$

This cost function is identical to the MBET2 objective (2.35), so we may directly borrow the PGPM update used in MBET2 for the inner $\mathbf{x}_2$ iteration:

$$\mathbf{x}_2^{(i+1)} = \mathcal{P}_{\mathcal{X}_2} \left\{ \mathbf{x}_2^{(i)} - \left[\mathbf{D}_2^{(n)} \right]^{-1} \nabla \Psi_2^{(n)}(\mathbf{x}_2^{(i)}) \right\}, \quad \text{where} \quad (3.33)$$

$$\mathbf{D}_2^{(n)} := \sum_{l=1}^{L} \mathrm{diag}_j \left\{ d_{l,j}^+ + d_{l,j}^- \right\} + \beta_2 \, \mathrm{maxeig} \left\{ \nabla^2 R(\mathbf{x}_2^{(n)}) \right\} \mathbf{I}_N, \quad (3.34)$$

(a) $\left|\mathbf{y}_1^-\right|, \left|\mathbf{y}_2^-\right|$ (a) $\left|\mathbf{y}_1^+\right|, \left|\mathbf{y}_2^+\right|$
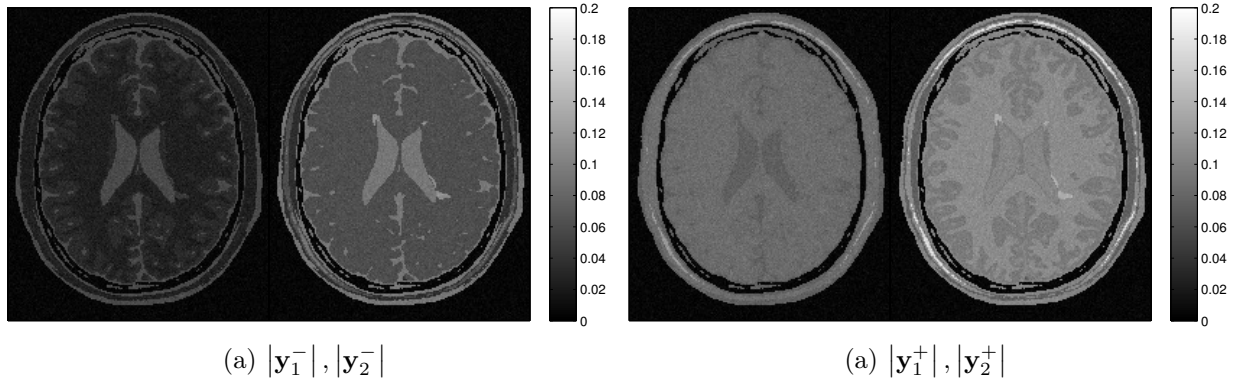
Figure 3.4: Image-domain raw data used for MBERP, with $L = 2$ optimized flip angles, $\boldsymbol{\alpha} = (11, 37)°$. Complex-valued raw data is displayed in magnitude. The relationship between raw data and underlying parameters of interest is clearly nonlinear, complicating estimation.

and single data-fit term curvature bounds $d_{l,j}^{\pm}$ are defined in (2.40), for the jth voxel. Since $\mathcal{X}_2$ is closed and convex, iteration (3.33) is guaranteed to monotonically converge to a local minimum of $\Psi_2^{(n)}(\mathbf{x}_2)$; once a convergence criterion is met, we assign the most recent update to $\widehat{\mathbf{x}}_2^{(n+1)}$.

Summary of MBERP

To reiterate, we cycle over each of these three variable updates while holding the other two fixed at the previous estimate, for a total T outer iterations. This simplifies the joint optimization problem into a sequence of single-variable optimization problems. As long as each of the inner minimization problems reduces the objective function (guaranteed for the M_0^* and T_2 updates, and empirically validated for the T_1 update), this algorithm, denoted MBERP (Alg. 3) will converge in cost.

3.5 Experimentation

3.5.1 Results from Synthetic Data

We combine true M_0, T_1, T_2, and T_2^* maps from the 95th slice of the BrainWeb digital phantom [10] to compute uncorrupted 217×181 DESS data at optimized flip angles $\boldsymbol{\alpha} = (11, 37)°$, using Equations (2.26) and (2.27). We select $T_E = 5$ms and $T_R = 20$ms, for a total scan time of 40ms per phase encode line. We corrupt the true data with unit-variance, complex white Gaussian noise, to yield roughly $(25, 38, 43, 47)$dB SNR raw data $(\mathbf{y}_1^-, \mathbf{y}_2^-, \mathbf{y}_1^+, \mathbf{y}_2^+)$, respectively (Fig. 3.4). We strive to jointly estimate T_1 and T_2 maps from this raw data, with no prior knowledge of M_0 or T_2^*.

As a comparison to assess performance gains strictly from joint (versus sequential) estimation, we separately combine the same true parameter maps to simulate two SPGR datasets and one DESS dataset at optimized flip angles $\boldsymbol{\alpha} = (5, 22, 47)°$, using Equations (2.30), (2.26), and (2.27). We initially selected $T_E = 5$ms and $(T_{R,\text{SPGR}}, T_{R,\text{DESS}}) = (16, 10)$ms, as suggested by our CRLB analysis; however, these repetition times yielded poor results, possibly due to biasing in the initialization,

(a) Sequential (MBET1/MBET2)   (b) Joint (MBERP)   (c) Ground Truth

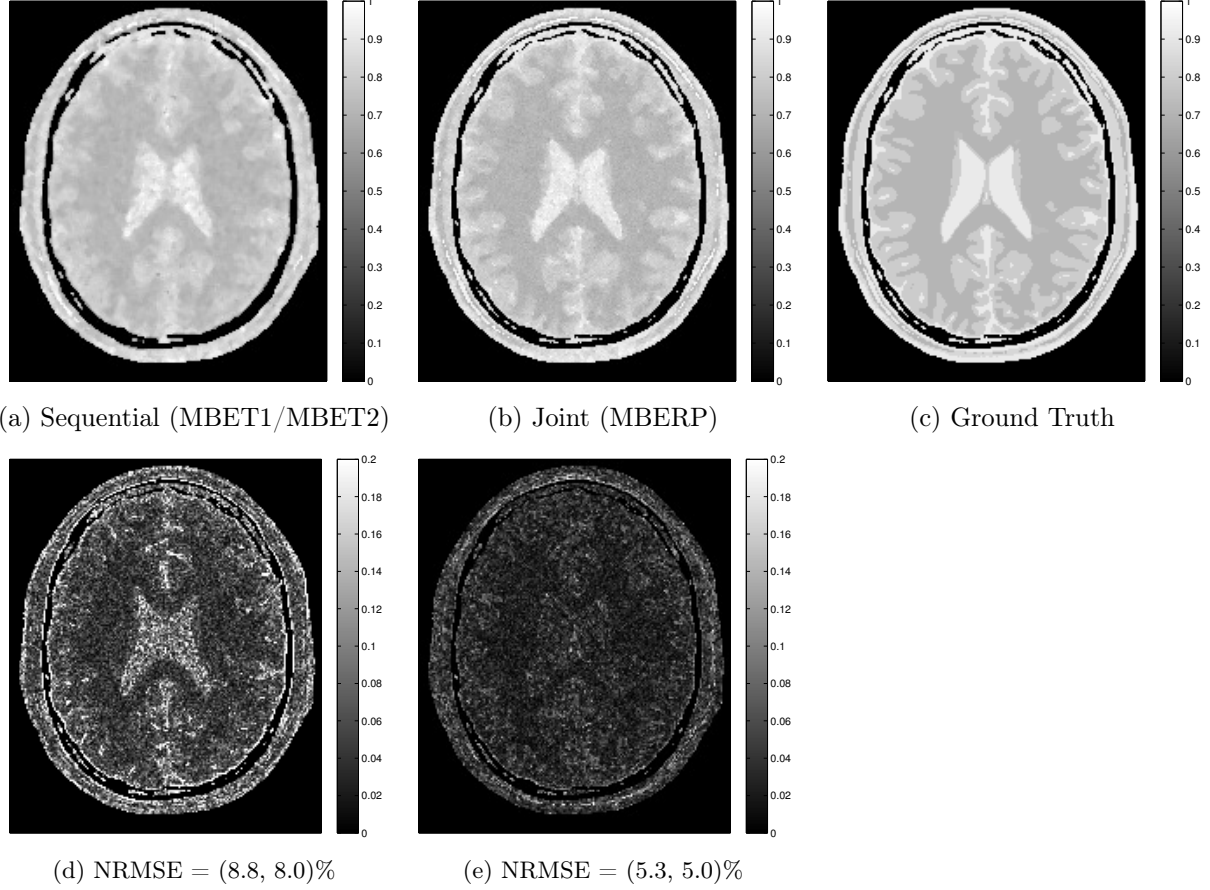(d) NRMSE = (8.8, 8.0)%   (e) NRMSE = (5.3, 5.0)%

Figure 3.5: Nuisance parameter M_0^* maps (a, b) and corresponding absolute-difference images (d, e), with respect to ground truth (c). Difference images are magnified $5\times$ to highlight differences. Tissue-specific errors are spatially averaged over (GM, WM) voxels, and are reported in NRMSE.

as discussed in Chapter 2. Because of this, we elect to keep scan profiles consistent in T_R and T_E, and choose to use $(T_{R,\mathrm{SPGR}}, T_{R,\mathrm{DESS}}) = (20, 20)$ms, for a total scan time of 60ms per phase encode line. We hold the noise level consistent by also corrupting the true SPGR and DESS data with unit-variance, complex white Gaussian noise, to yield roughly (36, 42)dB SNR raw SPGR datasets $(\mathbf{y}_1, \mathbf{y}_2)$, and (40, 47)dB SNR raw DESS datasets $(\mathbf{y}^-, \mathbf{y}^+)$, respectively. In this benchmark test, we strive to sequentially estimate accurate T_1 maps using MBET1, and then estimate accurate T_2 maps using MBET2, with no prior knowledge of M_0 or T_2^*.

We initialize MBERP using the variable projection method, with a dictionary of $K = 2500$ elements describing expected DESS signals at 50 T_1 and T_2 values logarithmically spaced between $[10^1, 10^4]$ and $[10^0, 10^3]$, respectively. We initialize MBET1 and MBET2 using median-filtered TLS estimates, as described in previous chapters. For both methods, we choose ϕ_k to be a hyperbolic potential function. For each method, we optimize regularization parameters to yield visually reasonable results: we choose $(\beta_{\mathrm{m}}, \beta_1, \beta_2) = (2^{-2}, 2^9, 2^8)$ for the sequential approach, and $(\beta_{\mathrm{m}}, \beta_1, \beta_2) = (2^3, 2^9, 2^7)$ for the joint approach. All inner minimizations are iterated until convergence criterion $||\mathbf{x}_t^{(i)} - \mathbf{x}_t^{(i-1)}|| < 10^{-6}||\mathbf{x}_t^{(i)}||$ is satisfied, where $t \in \{\mathrm{m}, 1, 2\}$. We select solution sets

(a) Sequential (MBET1/MBET2) (b) Joint (MBERP) (c) Ground Truth
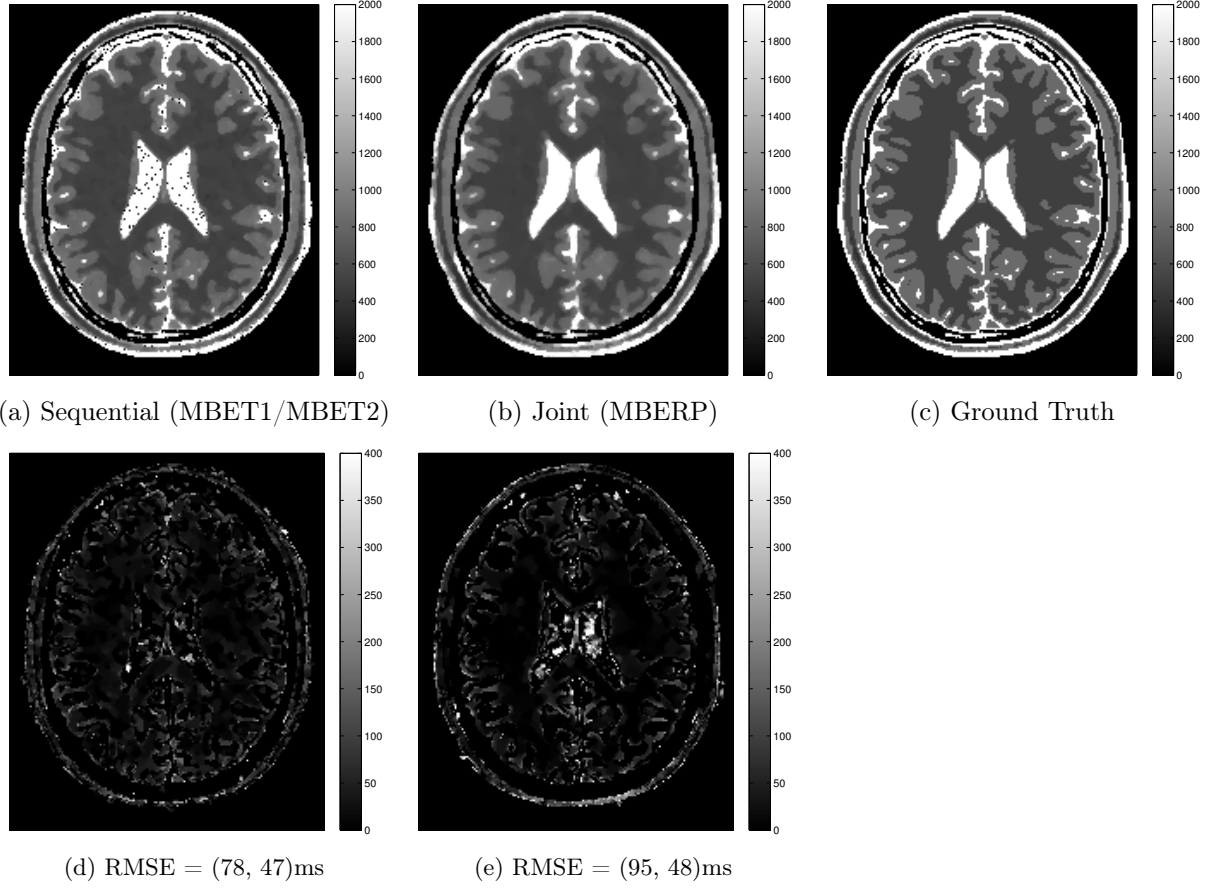
(d) RMSE = (78, 47)ms (e) RMSE = (95, 48)ms

Figure 3.6: Regularized T_1 maps (a, b) and corresponding absolute-difference images (d, e), with respect to ground truth (c). Difference images are magnified $5\times$ to highlight differences. Tissue-specific errors are spatially averaged over (GM, WM) voxels, and are reported in RMSE.

$\mathcal{X}_1$ and $\mathcal{X}_2$ to correspond with $T_1 \in [5, 5000]$ms and $T_2 \in [5, 500]$ms, respectively. On a 3.5GHz desktop with 32GB RAM, with $T = 30$ outer iterations, sequential (MBET1, then MBET2) and joint (MBERP) reconstructions take about 85s and 195s per slice, respectively.

Figures 3.5, 3.6, and 3.7 compare regularized M_0^*, T_1, and T_2 estimates using sequential (MBET1, then MBET2) vs. joint (MBERP) approaches. Difference images (with respect to the ground truth) are $5\times$ magnified as compared to the scalings of corresponding maps, to better distinguish quality of the estimates. Aside from speckle reduction in the cerebrospinal fluid (CSF), we visually observe a noticeable improvement in the distinction of grey-matter (GM) and white-matter (WM) regions of interest. MBERP recovers clearer, more accurate GM/WM boundaries than does MBET1/MBET2. This distinction is especially apparent when comparing the M_0^* and T_2 estimates.

We attribute MBERP's improvement to its effectively greater amount of accessible data at each inner minimization step. The joint approach utilizes all four datasets to individually update each of the three unknowns. By contrast, the sequential approach uses the two SPGR datasets to only estimate M_0^*, T_1, and two DESS datasets to only estimate T_2. With effectively twice as much data at each step, MBERP is able to better distinguish fluctuations in the data due to true intrinsic

43

(a) Sequential (MBET1/MBET2)   (b) Joint (MBERP)   (c) Ground Truth

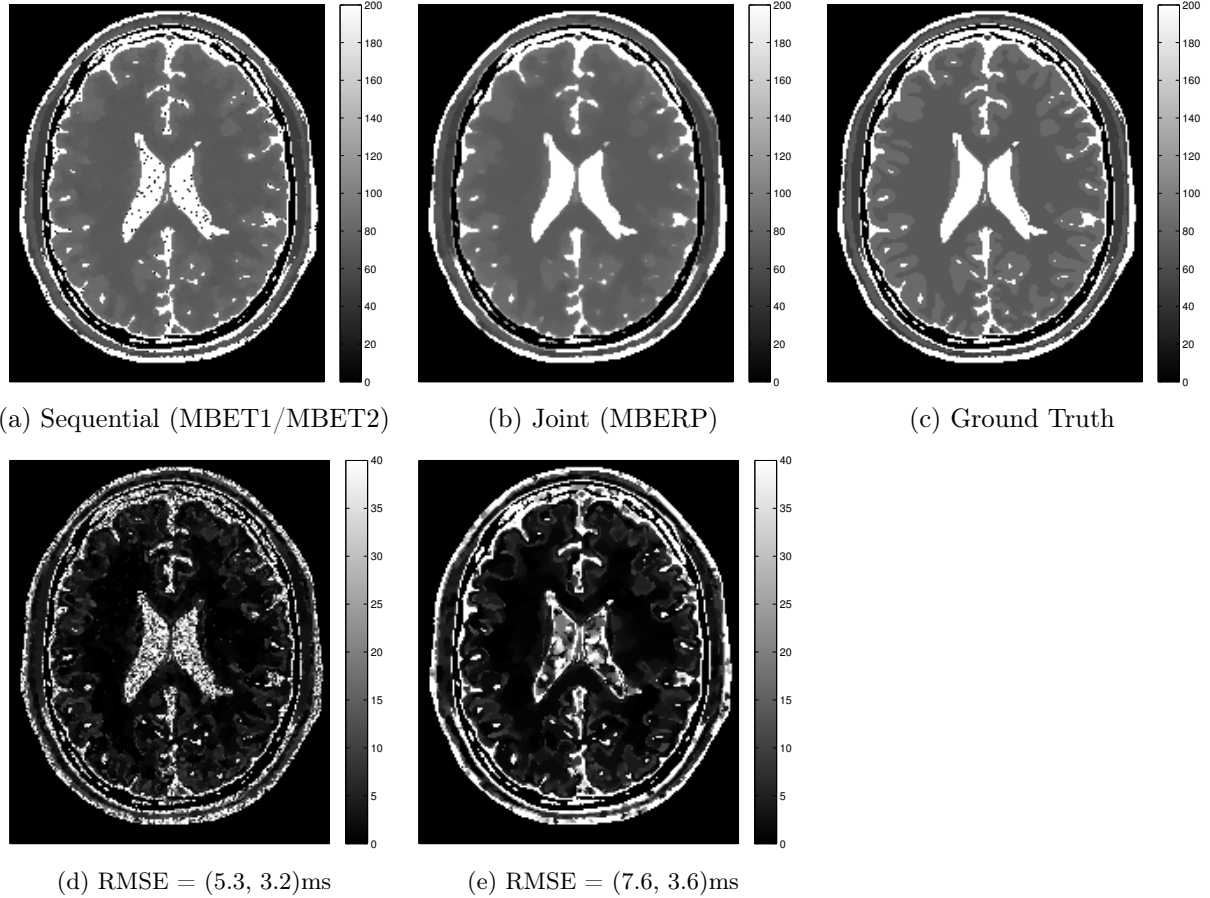(d) RMSE = (5.3, 3.2)ms   (e) RMSE = (7.6, 3.6)ms

Figure 3.7: Regularized T_2 maps (a, b) and corresponding absolute-difference images (d, e), with respect to ground truth (c). Difference images are magnified $5\times$ to highlight differences. Tissue-specific errors are spatially averaged over (GM, WM) voxels, and are reported in RMSE.

parameter variations, as opposed to noise. As a result, MBERP recovers tissue boundaries more accurately and reduces noise in the parameter estimates more effectively.

Table 3.2 provides quantitative comparison of the sequential and joint approaches in GM and WM regions of interest. As compared to method-of-moments and median-filtered estimates, both methods display substantial gains in accuracy and precision, due to a model-based formulation and edge-preserving regularization. Comparing MBERP to MBET1/MBET2 however, we find a mixture of advantages and disadvantages. Most notably, MBERP achieves greater T_1 and T_2 estimation *accuracy*, but suffers from lower estimation *precision*, across all GM voxels. Performance variations in WM voxels are insignificant. We again attribute this behavior to the imbalance in effective data availability for each of the T_1, T_2 estimation steps. With the joint approach, access to all four datasets affords greater reliance on data fidelity (yielding greater accuracy), and less reliance on regularization (yielding less precision). The opposite applies to the sequential approach. Returning to a broader scope, both MBERP and MBET1/MBET2 are accurate and precise methods for T_1, T_2 estimation, but we find MBERP to be preferable because of its robustness to scan variations, fewer scan parameters to tune, and faster (40ms vs. 60ms per phase encode, as tested) scan times.

Tissue	Sequential	Joint	Ground Truth
M_0^* GM	0.80 ± 0.07	0.79 ± 0.04	0.80 ± 0.00
M_0^* WM	0.73 ± 0.05	0.72 ± 0.03	0.71 ± 0.00
T_1 GM	820 ± 77	835 ± 94	833 ± 0
T_1 WM	528 ± 38	523 ± 42	500 ± 0
T_2 GM	82.6 ± 5.3	83.1 ± 7.6	83.0 ± 0.0
T_2 WM	71.7 ± 2.7	72.2 ± 2.9	70.0 ± 0.0

Table 3.2: Quantitative comparison of sequential (MBET1/MBET2) versus joint (MBERP) estimation approaches, in terms of M_0^*, T_1, and T_2 estimates and standard deviations, spatially averaged over GM and WM regions of interest. M_0^* values are normalized such that $M_0 = 1$; T_1, T_2 values are reported in (ms). The methods display no appreciable difference in estimation performance.

3.5.2 Results from *In Vivo* Data

3.6 Conclusions and Future Work

We have presented MBERP, a novel, model-based approach for simultaneous T_1, T_2 relaxometry, from as few two DESS sequences. We have also presented a CRLB-inspired method for optimal selection of parameters for joint estimation methods such as MBERP. Because of its more complete usage of data, we find that MBERP's joint-estimation approach better recovers boundaries between GM and WM regions of interest, as compared a sequential approach such as MBET1, followed by MBET2. MBERP preserves much of the accuracy and precision gains discovered by MBET1 and MBET2. At the expense of greater computation, MBERP permits collection of fewer scans, requires tuning fewer scan parameters, remains robust to scan variations, and enjoys faster scan times.

In this formulation of MBERP, we employ separate regularization terms to *independently* encourage M_0^*, T_1, T_2 estimates to be piecewise-smooth. This does not fully exploit the problem structure. We have the additional prior knowledge that different parameter maps are not only piecewise-smooth, but also tend to have tissue boundaries in similar locations. One method to reformulate the problem to capitalize on this added structure involves inclusion of a joint sparsity penalty [34, 35, 36]. Such exploitation of correlations not only *within*, but also *between* parameter maps might allow for more confident identification of tissue boundaries, and thus more accurate T_1, T_2 estimation.

In our work thus far, we have taken our data vectors to lie in the image domain, whereas in reality, the data is collected in the Fourier domain (commonly referred to as k-space). Our formulation thus implicitly assumes a fast, simple initial transformation from k-space to the image domain. This transformation is trivial when k-space is fully sampled along a Cartesian grid, but is more complicated [37] for other k-space trajectories. Furthermore, to accelerate MR scans, k-space is often under-sampled. As future work, it is desirable to accommodate such non-Cartesian and/or under-sampled k-space trajectories in our model-based framework for Quantitative MRI. In Chapter Four, we step away from QMRI and investigate such methods for reconstruction of accelerated MR acquisitions, with the eventual goal of accelerated Quantitative MRI in mind.

Finally, we have focused thus far on two particular types of fast MR pulse sequences (SPGR and DESS) for quantitative parameter mapping. As longer-term future work, it is of interest to apply

the model-based methods described here to scan profiles containing other fast scans, such as the recently-developed Triple-Echo Steady-State (TESS) pulse sequence [38].

Chapter 4

Reconstruction of Accelerated MRI Acquisitions which use Partial Fourier, Partial Parallel (PFPP) Imaging

4.1   Introduction

[1]Magnetic Resonance Imaging (MRI) is a popular medical imaging technique which has found applications in many different fields, including but not limited to medicine, psychology, neuroscience, etc. Although the use of MRI to image the human body has increased considerably in recent times, acquisition of MRI data takes a considerably long time, which has limited its application in certain areas such as imaging the cardiovascular or respiratory systems. Accelerating acquisitions while maintaining high spatial resolution would increase temporal resolution, an important improvement in certain applications like real-time imaging or dynamic contrast enhanced MRI. Such improvements may also help reduce motion artifacts caused due to physiological processes like cardiac and respiratory cycles. Most importantly, this would help reduce patient discomfort by reducing total scan times and subsequently help lower operating costs. In this regard, many accelerated acquisition techniques have been proposed in past studies and are still an ongoing topic of research.

MR images typically consist of many spatial frequencies. By convention, a spatial frequency is denoted by the letter k, and the set of all spatial frequencies is denoted k-space. Raw MRI data are samples in k-space; thus, an Inverse Fourier Transform (IFT) on this raw data recovers the desired image. Raw data is collected by systematically populating k-space in an order defined by a k-space trajectory. Previously proposed accelerated acquisition and reconstruction techniques like Partial Fourier or Parallel Imaging involve under-sampling k-space in some prescribed manner, and then reconstructing the under-sampled data using certain assumptions or additional information.

Partial Fourier (PF) reconstruction involves acquiring only a portion of k-space and then reconstructing the image using only the data sampled over this reduced area of k-space. In this technique,

---

[1]This work was completed as a group project for EECS 556 (Image Processing) at the University of Michigan. Other group members (and co-authors of this chapter) include Brandon Oselio and Yash Shah

a conjugate-symmetric assumption on k-space is used to reduce scan time, *i.e.*, we collect roughly one-half of k-space and populate the other half through conjugate symmetry. Typically, a speed up factor of approximately 2 can be obtained with PF reconstruction.

Parallel Imaging (PI) has been used to achieve better quality images by using multiple coil receivers during MR acquisition. Due to the geometry of these coils, they are more sensitive to specific parts of the object that are closer to them. This is defined by their sensitivity encoded maps (SENSE maps). Information from multiple coils together produce a higher quality reconstruction as compared to a single coil. Because of the increased data collection per point in k-space, PI allows for image reconstruction without aliasing, even if the data is under-sampled by skipping acquisition lines. Parallel imaging can be used either to reduce scan time, or improve signal to noise ratio (SNR), or a combination of both. Theoretically, parallel imaging can be employed to achieve an acceleration factor which is at most equal to the number of coils used.

In this chapter, we combine the Partial Fourier and Parallel imaging techniques, using a constrained reconstruction technique as described in [39]. Excessive dependence on the conjugate symmetry assumption may lead to phase artifacts; hence, we use a regularization scheme to vary the degree to which the PF assumption is enforced during reconstruction. This regularization parameter is used as a tradeoff between artifacts introduced in the PF technique versus the noise amplification suffered as a result of the PI technique. Such a reconstruction formulation is denoted Partial Fourier, Partially Parallel (PFPP) imaging. It is shown that PFPP is successful in obtaining higher quality images as compared to PF or PP alone. We therefore demonstrate that PFPP is a good choice for dramatically accelerating acquisitions, while retaining high SNR after reconstruction.

In Section 2, we describe the theory behind PF imaging and PP imaging, and describe how these techniques can be combined to obtain PFPP imaging. In Section 3, we briefly describe a technique for estimating receive coil sensitivities for PP imaging, formulate the PFPP optimization problem, describe an approach to solving this problem, and describe a novel extension. In Section 4, we present results for both the original optimization problem described in [39], as well as our extension. In Section 5, we briefly conclude our finding and discuss directions for future work.

4.2   Theory

4.2.1   Partial Fourier (PF) Imaging

Partially Fourier (PF) imaging [40] seeks to exploit the conjugate symmetry of k-space data corresponding to pure-real images by acquiring 50% of k-space and populating the remainder through simple conjugation. As shown in Figure 4.1, this amounts to an approximate two-fold acquisition acceleration if sampling is omitted along the phase-encode direction.

In practice, MR images are not entirely pure-real because of phase errors inherent to true MR acquisitions. As a result, reconstructions using pure PF scans (such as Figure 4.2(a)) are vulnerable to phase artifacts. One method to compensate for such phase errors is to collect slightly greater than 50% of k-space samples, thereby eliminating the reliance of conjugate symmetry for the lowest frequency scan lines in the phase-encoding direction, while retaining most of the two-fold acceleration. This modification increases resistance to phase errors (Figure 4.2(b)), as phase variation in
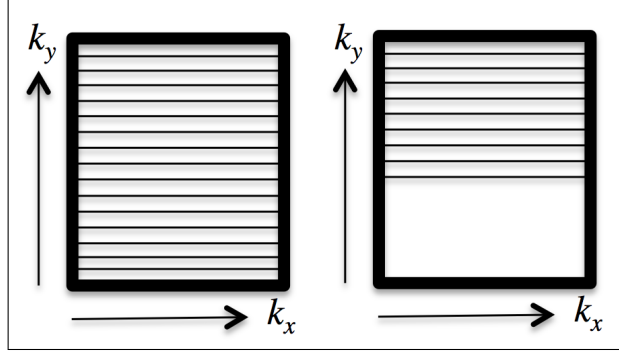
Figure 4.1: A (false) assumption of pure-real images allows for exploitation of conjugate symmetry of k-space in PF imaging. *Left*: A fully-sampled k-space sampling pattern. *Right*: Partially-sampled PF sampling pattern, which collects slightly greater than the upper 50% of k-space. As more data is omitted PF trajectories introduce greater phase artifacts for larger acceleration factors.
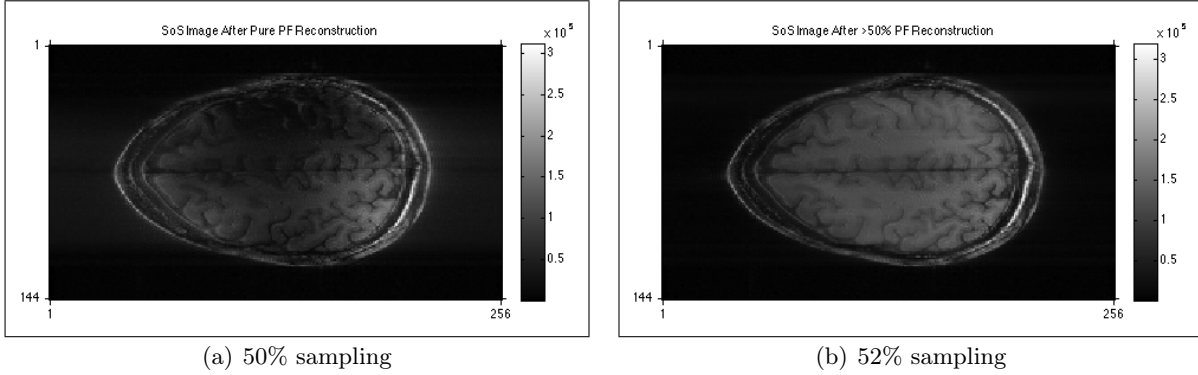

(a) 50% sampling

(b) 52% sampling

Figure 4.2: Reconstructed images using PF sampling scheme. 4.2(a): Exactly 50% of k-space data omitted, subjecting the reconstruction to phase errors at all frequencies. 4.2(b) 48% of k-space data omitted, protecting the reconstruction from low-frequency phase errors. Since most phase variation is slowly-varying in MR imaging, the reconstruction quality is dramatically improved.

MR images is typically low-frequency. However, phase artifacts remain in regions of rapid phase variation, such as in the skull boundary between brain matter and non-brain voxels. Increasing the number of scan lines improves these artifacts, but reduces acquisition acceleration.

4.2.2 Partial Parallel (PP) Imaging

In a conventional Cartesian MR acquisition, the k-space sampling pattern traverses a Cartesian grid, which is defined by a number of frequency-encode and phase-encode steps. Data acquisition in the frequency-encoding direction is typically very rapid (few milliseconds), and hence all the k_x samples along a single k_y line are acquired in a short time. However, a separate phase encode step is required to be applied to sample each value of k_y along the phase-encoding axis. Thus, k-space sampling using a prescribed number of phase-encoding steps is the bottleneck as far as acquisition time is concerned. This is the motivation behind using parallel imaging, in which acquisitions are
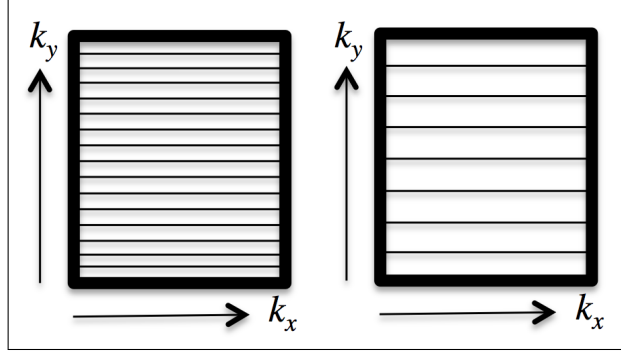
Figure 4.3: Representation of reduced k-space coverage, with PP imaging. *Left*: A fully-sampled k-space sampling pattern. *Right*: PP sampling pattern, with under-sampling factor $n_p = 2$ (leading to $2\times$ acquisition acceleration). To avoiding aliasing during reconstruction, a minimum of $n_c \geq 2$ coils must acquire data simultaneously for proper SENSE reconstruction.

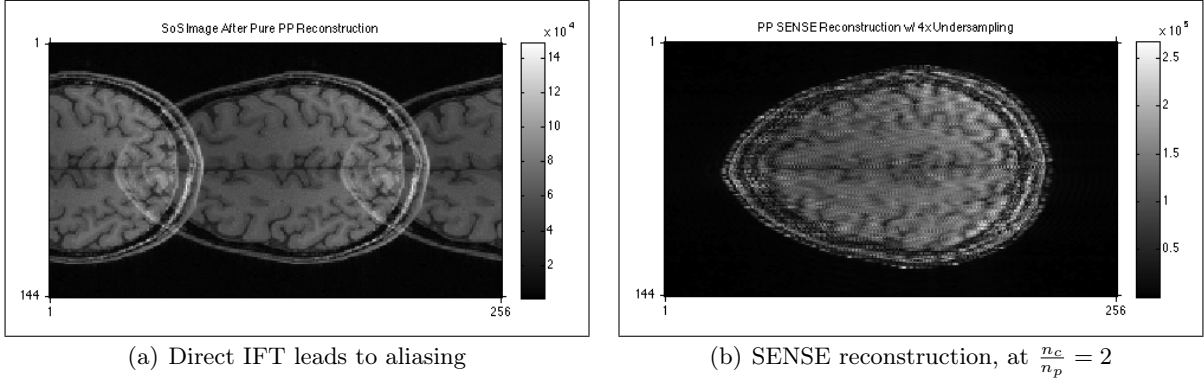accelerated by under-sampling k-space in the phase-encode direction, as shown in Figure 4.3.

(a) Direct IFT leads to aliasing

(b) SENSE reconstruction, at $\frac{n_c}{n_p} = 2$

Figure 4.4: Reconstructed images using a PP sampling scheme 4.4(a): Direct reconstruction through an IFT of sum-of-squares data collected along sampling pattern 4.3 causes aliasing along the phase-encoding direction. 4.4(b): Reconstructed image with $n_p = 4$ under-sampling, using $n_c = 8$ coil data. Aliasing is removed, but noise is amplified because no regularization is used.

As shown in Figure 4.4(a), direct reconstruction of PP data will cause aliasing in the image domain. To "undo" such aliasing, multiple coils are distributed around the object during scanning and simultaneously collect k-space data. A SENSE reconstruction [41, 42] is used to then solve a series of voxel-by-voxel least squares problems; ignoring regularization terms, this is simply

$$\hat{x} = \arg\min_x ||y - \mathbf{FSM}x||_2^2, \qquad (4.1)$$

where x and $\hat{x}$ are vectorized $NM \times 1$ representations of the true and estimated images, respectively; y is a vectorized $\frac{NMn_c}{n_p} \times 1$ representation of the k-space data from n_c coils, with an n_p under-sampling ratio; $\mathbf{F}$ is an $\frac{NMn_c}{n_p} \times NMn_c$ block-diagonal collection of n_c under-sampled DFT matrices; $\mathbf{S}$ is a $NMn_c \times NM$ stack of diagonal precomputed sensitivity maps; and $\mathbf{M}$ is an $NM \times NM$ precomputed diagonal mask, which simply zeroes out all non-brain voxels to prevent noise amplification.

The SENSE reconstruction problem must be overdetermined ($\frac{n_c}{n_p} \geq 1$) in order to guarantee a unique solution. In theory, this suggests that an arbitrary number of coils can be used to under-sample by arbitrarily high ratios. In practice, however, increased under-sampling factors cause noise amplification, as shown in Figure 4.4(b). Although some degree of regularization can be helpful in promoting smoother solutions, there is an inherent limit (as a function of B_0 field strength) [43] to the degree of under-sampling possible without sharply decreasing SNR.

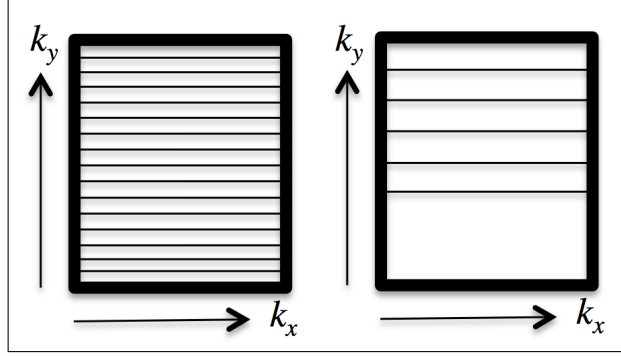### 4.2.3   Partial Fourier, Partial Parallel (PFPP) Imaging

Figure 4.5: Representation of reduced k-space coverage, with PFPP imaging. *Left*: A fully-sampled k-space sampling pattern. *Right*: PFPP sampling pattern, with under-sampling factor $n_p = 2$. Because of both PP and PF under-sampling effects, we now obtain nearly $4\times$ acquisition acceleration.

The idea of PFPP imaging is to extend the strengths of PP imaging by discouraging complex-valued solutions (especially those with rapid phase variation), allowing for omission of most phase-encode lines on one side of k-space and reliance on conjugate symmetry for reconstruction (Figure 4.5). By incorporating PF ideas into the PP framework, we may jointly introduce regularization into our objective function, while also increasing the effective under-sampling factor for each given n_p.

Defining the $\frac{NMn_c}{n_p} \times NM$ system matrix $\mathbf{B} = \mathbf{FSM}$ for simplicity, we may rewrite the standard SENSE equation, $y = \mathbf{B}x$, in a manner that decomposes all real and imaginary components:

$$\begin{bmatrix} [\mathrm{Re}]\{y\} \\ [\mathrm{Im}]\{y\} \end{bmatrix} = \begin{bmatrix} [\mathrm{Re}]\{\mathbf{B}\} & [\mathrm{Im}]\{-\mathbf{B}\} \\ [\mathrm{Im}]\{\mathbf{B}\} & [\mathrm{Re}]\{\mathbf{B}\} \end{bmatrix} \begin{bmatrix} [\mathrm{Re}]\{x\} \\ [\mathrm{Im}]\{x\} \end{bmatrix}. \tag{4.2}$$

We now introduce a regularization parameter λ into this (so far pure PP) formulation [39], which is used to control the PF proposition that solutions $\hat{x}$ should be pure real:

$$\begin{bmatrix} [\mathrm{Re}]\{y\} \\ [\mathrm{Im}]\{y\} \\ 0 \end{bmatrix} = \begin{bmatrix} [\mathrm{Re}]\{\mathbf{B}\} & [\mathrm{Im}]\{-\mathbf{B}\} \\ [\mathrm{Im}]\{\mathbf{B}\} & [\mathrm{Re}]\{\mathbf{B}\} \\ 0 & \lambda\mathbf{I} \end{bmatrix} \begin{bmatrix} [\mathrm{Re}]\{x\} \\ [\mathrm{Im}]\{x\} \end{bmatrix}, \tag{4.3}$$

where $\mathbf{I}$ is the identify matrix. This PFPP formulation is now a constrained least-squares problem, in which λ controls the level of enforcement against complex solutions. As λ tends to zero, Equation (4.3) reduces to Equation (4.2), and we return to a purely parallel imaging formulation, which

is prone to noise amplification at high acceleration factors. Conversely, as λ tends to infinity, the solution $\hat{x}$ is increasingly discouraged from having an imaginary component, meaning that it depends more heavily on the conjugate symmetry assumption of PF imaging, reducing noise but increasing phase artifacts. An intermediate solution (finite, positive λ) seeks to strike a balance between the undesirable consequences of pure PF or PP imaging, while simultaneously increasing the effective scan acceleration than with either method individually.

As in the PF discussion, this formulation expects the true object x to be pure real, when in fact this is generally not the case, due to phase errors accrued during scanning. In part, we seek to mitigate this issue by once again covering greater than 50% of k-space with scan lines, thereby accounting for low-frequency phase variations explicitly. However, the constrained least-squares formulation alone will continue to suppress these desired phase variations, as any other complex component. We compensate this issue by using an independent low-frequency pre-scan to independently encode the low-frequency phase information directly into the normalized sensitivity maps $\mathbf{S}$. During reconstruction, these "phase" maps are applied to every solution iterate x^n to help account for phase information loss. However, high-frequency phase variations are unlikely to be corrected by these low-resolution phase maps, preventing excessive increase of λ.

4.3 The PFPP Problem and Its Optimization

4.3.1 SENSE Map Estimation

As discussed previously, SENSE maps are jointly used in the PFPP framework to (1) compensate spatial weightings associated with signal drop-off due to the distribution of scanner coils, (2) differentiate coil-by-coil data for SENSE reconstruction, and (3) encode low-frequency phase variation due to scan imperfections, shielding it from complete suppression in the objective function.
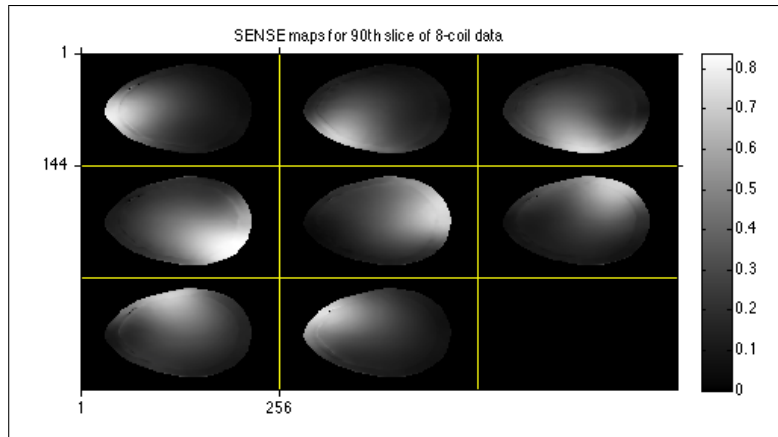

Figure 4.6: Normalized SENSE maps for the 90th slice of 8-coil data, computed through sum-of-squares estimation of an independent 64×64 fully-sampled, low-frequency pre-scan.

A simple technique is given for estimating normalized SENSE maps (Figure 4.6) for each coil, based on a fully-sampled, low-frequency pre-scan of k-space, performed prior to the PFPP scan:

1. The low-frequency pre-scans $K_i[k_x, k_y]$ (where the ith coil runs from 1 to n_c) for each coil are appropriately zero-padded, apodized to reduce ringing, and Inverse Fourier Transformed into the image domain, to produce a series of n_c blurry images, $k_i[n, m] = \mathcal{F}^{-1}\{K_i[k_x, k_y]\}$.

2. A reference SENSE map S_{ref} is computed as the voxel-by-voxel square root of the sum of squares (SoS) image across the n_c blurry images: $S_{ref}[n, m] = \sqrt{\sum_{i=1}^{N_{coils}} k_i[n, m]^2}$.

3. The non-brain voxels in each of the blurry images $k_i[n, m]$ are masked out through a heuristic thresholding operation, to avoid noise amplification before normalization.

4. The masked, blurry images are normalized (voxel-by-voxel) by the reference SENSE map S_{ref} to produced normalized SENSE maps for each coil, $S_i[n, m] = \frac{k_i[n,m]}{S_{ref}[n,m]}$.

4.3.2   Objective Function

Recalling that $\mathbf{B} = \mathbf{FSM}$, the least-squares solution of Equation (4.2) is equivalent to minimizing the objective function defined in Equation (4.1). A natural extension [39] insists then that the constrained least-squares solution of Equation (4.3) minimizes the following expanded objective:

$$\hat{x} = \arg\min_x ||y - \mathbf{FSM}x||_2^2 + \lambda^2 || \, [\mathrm{Im}] \, \{x\}||_2^2 \, . \tag{4.4}$$

In addition, [39] suggests to use a zeroth-order quadratic regularization penalty on the image as an energy minimization constraint, which modifies the objective function to

$$\hat{x} = \arg\min_x ||y - \mathbf{FSM}x||_2^2 + \lambda^2 || \, [\mathrm{Im}] \, \{x\}||_2^2 + \delta^2 ||x||_2^2 \, . \tag{4.5}$$

We minimize this objective using an iterative conjugate gradient (CG) method, described next.

4.3.3   Optimization with Conjugate Gradient (CG) Algorithm

Conjugate Gradient (CG) is an iterative algorithm [44] used to find solutions to the general matrix equation $Ax = b$. It is guaranteed to reach an optimal, unique solution $\hat{x}$ in n steps, where n is the length of x. Because CG approaches solutions while monotonically decreasing in error, it can also be used when running n iterations is infeasible, as is the case for typical vectorized images of dimension $NM \times 1$. The solution $\hat{x}$ is reconstructed from a series of conjugate directions with respect to the system matrix A; the weighting factors in this summation are selected through special inner products between the corresponding directions and the output b. By carefully choosing the directions in which to optimize, we can quickly approach the exact solution in a relatively few number of iterations, especially for large n. Originally implemented in [39], Algorithm 4 details an application of CG to the PFPP objective function, given in Equation (4.5).

Algorithm 4 Conjugate Gradient, as applied to PFPP objective, Equation (4.5).

$x = x_{initial}$
$d = x_{initial}$
$r_0 = (\mathbf{FSM})^H y$
repeat
  $q = \mathbf{M}^H \mathbf{S}^H \mathbf{F}^H \mathbf{FSM} d + i\lambda^2 [\text{Im}]\{d\} + \delta^2 d$
  if $(j == 1)$ **then**
    $r_1 = r_0 - q$
    $d = r_1$
  else
    $\alpha = \frac{r_j^H r_j}{[\text{Re}]\{d^H q\}}$
    $x = x + \alpha d$
    $r_{j+1} = r_j - \alpha q$
    $d = r_{j+1} + \frac{r_{j+1}^H r_{j+1}}{r_j^H r_j} d$
    $test = \frac{r_{j+1}^H r_{j+1}}{r_0^H r_0}$
  end if
  $j {+}{+}$
until $(j > count\_max)$ or $(test \leq tol)$

In keeping with the notation in Algorithm 4, x is the iterated image. d is the conjugated direction, and r_j is the jth residual. The algorithm will halt either when the normalized residual (with respect to the original) is small enough or when a maximum number of iteration cycles is reached. It should also be noted that it is impractical to implement this as written, in terms of matrices; the size of **FSM** is much too large. Fortunately, as shown in [42], these matrix operations can be implemented coil-by-coil using FFTs, with appropriate under-sampling procedures.

4.3.4 Extension: Reconstruction Through Sparsity Enforcement

Modifying the Objective Function

As an extension to [39], we explore advantages and disadvantages of an alternative objective function to Equation (4.5). While a quadratic regularization term such as $\delta^2 ||x||_2^2$ is analytically appealing due to its strict convexity, it is suboptimal in its method of promoting smoothness. Since zeroth-order regularization seeks to promote images with minimal energy, solutions will be encouraged to tend towards zero, which is undesirable because of the subsequent reduction in SNR.

Instead, we can better describe a "desirable" image as denoised and roughly piecewise-constant. One method to encourage such solutions is through enforcement of sparsity in some wavelet domain. Specifically, if $x = \mathbf{Q}\theta$ is the synthesis formulation of solution x from coefficients θ through a unitary wavelet transform $\mathbf{Q}$, then an alternative objective function such as

$$\hat{x} = \arg\min_x ||y - \mathbf{FSM}x||_2^2 + \lambda^2 ||[\text{Im}]\{x\}||_2^2 + \beta||\mathbf{Q}^H x||_1 \tag{4.6}$$

would better encourage piece-constant solutions. Note that we have introduced the ℓ_1 norm operator $||\mathbf{Q}^H x||_1$ as a convex (but not strictly convex) cost against pixel variation. Because of the complexity of solution x, it is difficult to minimize this objective function because of the nonlinearity associated with the ℓ_1 norm of a complex-valued function. A similar objective,

$$\hat{x} = \arg\min_x ||y - \mathbf{FSM}x||_2^2 + \lambda^2||\,[\mathrm{Im}]\,\{x\}||_2^2 + \beta_1||\mathbf{Q}^H\,[\mathrm{Re}]\,\{x\}||_1 + \beta_2||\mathbf{Q}^H\,[\mathrm{Im}]\,\{x\}||_1\,, \qquad (4.7)$$

alleviates this issue by helpfully separating the ℓ_1 norm penalty onto the real and imaginary components of $\hat{x}$ individually. This is in fact preferable to Equation (4.6), as it allows greater flexibility on piecewise-smoothness penalties of the real and imaginary components.

Because of the linearity of the unitary wavelet transform $\mathbf{Q}$, we can recast this minimization problem with a synthesis approach, estimating instead the transformed solution $\hat{\theta} = \mathbf{Q}^H \hat{x}$:

$$\hat{\theta} = \arg\min_\theta ||y - \mathbf{FSMQ}\theta||_2^2 + \lambda^2||\,[\mathrm{Im}]\,\{\theta\}||_2^2 + \beta_1||\,[\mathrm{Re}]\,\{\theta\}||_1 + \beta_2||\,[\mathrm{Im}]\,\{\theta\}||_1\,. \qquad (4.8)$$

In the ensuing discussion, it is of interest to minimize this non-quadratic objective through a series of simpler minimizations on the real and imaginary components of θ separately.

Objective Minimization via Optimization Transfer

If we denote the objective in Equation (4.8) as $\Phi(\theta)$, it suffices to sequentially minimize a series of majorant surrogate functions $\phi^i(\theta) \geq \Phi(\theta)$ as described in [45], each of which is strictly greater than the desired objective $\Phi(\theta)$ for all θ. The simplest such surrogate function takes the form

$$\phi^i(\theta) = \frac{\rho_{max}}{2}||z^{i-1} - \theta||_2^2 + \lambda^2||\,[\mathrm{Im}]\,\{\theta\}||_2^2 + \beta_1||\,[\mathrm{Re}]\,\{\theta\}||_1 + \beta_2||\,[\mathrm{Im}]\,\{\theta\}||_1\,, \qquad (4.9)$$

where z^{i-1} denotes the minimizer of the majorizer from the previous iteration and ρ_{max} is the Lipschitz constant corresponding to the system matrix $\mathbf{Q}^H\mathbf{M}^H\mathbf{S}^H\mathbf{F}^H\mathbf{FSMQ}$, *i.e.*,

$$\mathbf{Q}^H\mathbf{M}^H\mathbf{S}^H\mathbf{F}^H\mathbf{FSMQ} \preceq \rho_{max}\mathbf{I}\,. \qquad (4.10)$$

As ρ_{max} influences the step-size between iterations, we wish to minimize ρ_{max} as best as possible, while still bounding the system matrix, as in Equation (4.10). We begin by noting that the eigendecomposition of the product of under-sampled DFT matrices $\mathbf{F}^H\mathbf{F}$ is

$$\mathbf{F}^H\mathbf{F} = \mathbf{VDV}^H\,, \qquad (4.11)$$

where $\mathbf{V}$ is the fully sampled, unitary DFT matrix and $\mathbf{D}$ is the identity matrix, except with "0"s replacing "1"s along the diagonal on columns corresponding to the omitted columns of $\mathbf{F}$ after under-sampling $\mathbf{V}$. Consequently, $\mathbf{F}^H\mathbf{F} \preceq \mathbf{I}$, meaning that

$$\mathbf{Q}^H\mathbf{M}^H\mathbf{S}^H\mathbf{F}^H\mathbf{FSMQ} \preceq \mathbf{Q}^H\mathbf{M}^H\mathbf{S}^H\mathbf{SMQ}\,. \qquad (4.12)$$

By recalling the orthonormality of our inverse wavelet transform operator $\mathbf{Q}$, we observe that $\mathbf{Q}^H\mathbf{M}^H\mathbf{S}^H\mathbf{SMQ}$ and $\mathbf{M}^H\mathbf{S}^H\mathbf{SM}$ have identical eigenvalues; thus we can equivalently claim that

$$\mathbf{Q}^H\mathbf{M}^H\mathbf{S}^H\mathbf{F}^H\mathbf{FSMQ} \preceq \mathbf{M}^H\mathbf{S}^H\mathbf{SM}\,. \qquad (4.13)$$

Recalling the $NMn_c \times NM$ structure of $\mathbf{S}$, we observe that $\mathbf{S}^H\mathbf{S}$ is an $NM \times NM$ diagonal matrix of sum-of-squares; since M is another diagonal matrix of "1"s and "0"s,

$$\mathbf{M}^H\mathbf{S}^H\mathbf{S}\mathbf{M} \preceq \mathbf{S}^H\mathbf{S}. \tag{4.14}$$

Finally, $\mathbf{S}^H\mathbf{S}$ is trivially bounded by identity, scaled by its maximum diagonal entry:

$$\mathbf{S}^H\mathbf{S} \preceq \max\left[\sum_{i=1}^{n_c} S_i[n,m]^2\right]\mathbf{I}. \tag{4.15}$$

Combining Equations (4.10)-(4.15) through transitivity, we find a simple Lipschitz constant,

$$\rho_{max} = \max\left[\sum_{i=1}^{n_c} S_i[n,m]^2\right]. \tag{4.16}$$

If we are able to minimize each iteration's surrogate function $\phi^i(\theta)$ to produce a soft-thresholding solution, this constant ρ_{max} will provide insight in selecting step sizes.

Iterative Soft-Thresholding (IST) Solution to the Majorizing Surrogate, $\phi^i(\theta)$

Recalling Equation (4.8), expansion of the data-fit term of the ith surrogate function $\phi^i(\theta)$ in terms of real- and imaginary components of θ allows us to circumvent the use of complex calculus:

$$\phi^i(\theta_R, \theta_J) = \frac{\rho_{max}}{2}||z^{i-1} - \theta_R - j\theta_J||_2^2 + \lambda^2||\theta_J||_2^2 + \beta_1||\theta_R||_1 + \beta_2||\theta_J||_1, \tag{4.17}$$

where $\theta_R = [\mathrm{Re}]\{\theta\}$ and $\theta_J = [\mathrm{Im}]\{\theta\}$ are now explicitly written as pure-real variables over which to optimize over individually. This representation is simple to work with because of the explicit separation of real and imaginary components, in addition to voxel-by-voxel separability. A standard joint maximization over both θ_R and θ_J yields the following pure-real solutions:

$$\hat{\theta}_R^i = \mathrm{shrink}\{[\mathrm{Re}]\{z^{i-1}\}, \beta_1\} \tag{4.18}$$
$$\hat{\theta}_J^i = \left(\frac{1}{1+4\lambda^2}\right)\mathrm{shrink}\{[\mathrm{Im}]\{z^{i-1}\}, \beta_2\},$$

where the "shrink$\{m, p\}$" operation is defined as $\max\{|m| - p, 0\}\,\mathrm{sgn}\{m\}$. We reconstruct the ith complex estimate of the wavelet-transformed solution iterate, $\hat{\theta}^i$, through complex addition:

$$\hat{\theta}^i = \hat{\theta}_R^i + j\hat{\theta}_J^i. \tag{4.19}$$

After a series of sequential minimizations of this majorizer, we can monotonically approach the solution to the original objective, Equation (4.7), through a simple wavelet transform, $\hat{x} = Q\hat{\theta}$.

Algorithm 5 summarizes our novel application of an Iterative Soft-Thresholding (IST) technique [46] to the modified PFPP objective (Equation (4.5)), through a cycle of two separable shrinkage operations on the real and imaginary components of the wavelet-transformed image per iteration.

Algorithm 5 Iterative Soft Thresholding Algorithm (ISTA). A PFPP objective framework (Equation (4.5)) is minimized through a sequential series of surrogate objectives (Equation (4.8)).

$x^0 = S_{ref}$
$\rho_{max} = \max\left\{S_{ref}[n,m]^2\right\}$
$step = 1/\rho_{max}$
while ($k < count\_max$) and ($change > tol$) **do**
  $d = \mathbf{M}^H \mathbf{S}^H \mathbf{F}^H (\mathbf{FSM}x^k - y)$
  $r^k = x^k - 2 \times step \times d$
  $\theta^k = \mathbf{Q}^H r^k$
  $x_R^{k+1} = \mathbf{Q} \times \text{shrink}\{[\text{Re}]\{\theta^k\}, \beta_1\}$
  $x_J^{k+1} = (\frac{1}{1+4\lambda^2})\mathbf{Q} \times \text{shrink}\{[\text{Im}]\{\theta^k\}, \beta_2\}$
  $x^{k+1} = x_R^{k+1} + j x_J^{k+1}$
  $change = \frac{||x^{k+1}-x^k||_1}{||x^{k+1}||_1}$
  $k++$
end while

4.4 Experimentation

4.4.1 Data Acquisition and Analysis Metric

We base our results on a fully-3D, fully-sampled $256 \times 144 \times 128 \times 8$-coil brain data set.[2] The fully sampled data is first Inverse Fourier Transformed along the z-direction to produce slice-by-slice k-space data. A single slice is then extracted for 2D processing, although this work is easily extendable to 3D data through slice-by-slice repetition. We selected the 64th slice for all error comparisons, on the assumption of high SNR in the longitudinal center of the scan. Other images are created with the 90th slice because of increased visual perception of image artifacts.

To simulate the PFPP acquisition process, we retrospectively under-sampled the selected slice's raw data in the sampling pattern described in Figure 4.5. In addition, a 64×64 low-frequency pre-scan is extracted for initial SENSE map estimation. All other data is kept invisible to the algorithm. We compare our results to a "gold standard" image, created by taking the sum-of-squares (SoS) of images reconstructed from fully-sampled k-space data across all coils.

We use Normalized RMSE (NRMSE) as a metric of comparison. We note, however, that the NRMSE is not necessarily a good indicator of the ability of an image to portray relevant information to a clinician. It is, however, a convenient measure with which one can compare algorithms.

4.4.2 Main Results: PFPP with CG Reconstruction

Fig 4.7 displays results of PFPP reconstruction using a CG technique, for different levels of under-sampling, namely $n_p = 2, 4$, and 8, in comparison with the gold standard image. These under

---

[2]Courtesy Dr. Jeff Fessler and the fMRI Lab of University of Michigan, July 6, 2010

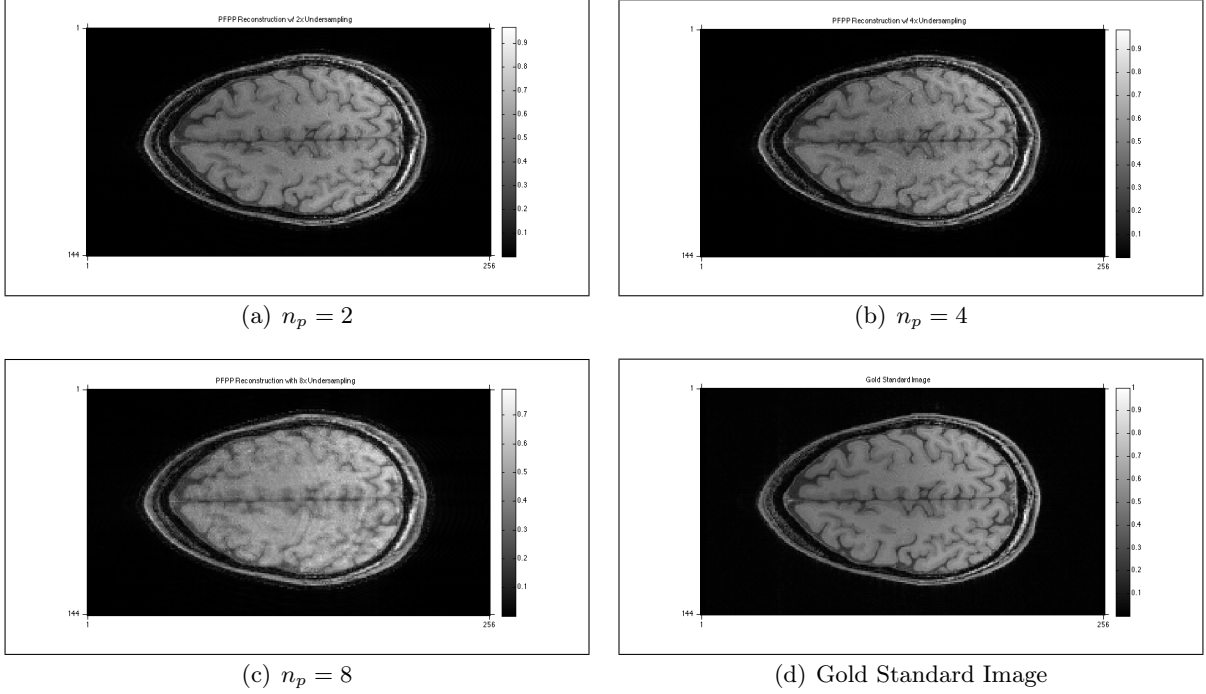|	
(a) $n_p = 2$	(b) $n_p = 4$
(c) $n_p = 8$	(d) Gold Standard Image

Figure 4.7: CG reconstructions of the 90th slice, minimizing objective ((4.5)). 4.7(a): $n_p = 2$, resulting in NRMSE = 0.1211. 4.7(b): $n_p = 4$, resulting in NRMSE = 0.1278. 4.7(c): $n_p = 8$, resulting in NRMSE = 0.1495 4.7(d): Gold Standard Image. The algorithm suppresses phase artifacts significantly better than a SENSE PP reconstruction, as in 4.2(b), and substantially reduces noise over pure PF, as in 4.4(b).

sampling-ratios equate to just under $4\times$, $8\times$, and $16\times$ acquisition acceleration factors, where imprecisions in these accelerations are due to a $\frac{9}{16}$ coverage of k-space, as described in Figure 4.2(b).

We ran these reconstructions with optimized (in an NRMSE sense) regularization parameters $\lambda = 0.0485$ and $\delta = 0.0222$, over normalized images. As expected, we observe a monotonic increase in NRMSE as n_p is increased. At under-sampling factors of $n_p = 2$ and $n_p = 4$, images are reconstructed with minimal noticeable aliasing. At $n_p = 8$, however, we approach an underdetermined problem (since there are only 8 coils of data), causing some aliasing to arise.

A parameter sweep (Figure 4.8) on λ and δ was performed at tolerance 10^{-8} in order to find the optimal values of $\lambda = 0.0485$ and $\delta = 0.0222$, as discussed previously. The average iteration counts to convergence was 34.175, and the average run time was 2.7367s. Note that despite the quadratic regularizer's sub-optimality with regards to "desirable" image qualities, the optimal error for this dataset is still achieved with a non-zero energy penalty. Although the lowest NRMSE is found with a non-zero δ, error increases quite quickly if the reconstruction is over-regularized with the quadratic energy penalty, *i.e.*, as δ increases. In practice, using a non-zero δ must be done with great caution; $\delta = 0$ is always a safe choice, and in this case is only marginally worse in terms of NRMSE.

We observe that the optimal λ is non-zero and finite, indicating that a mixture of PF and PP ideas produces images with less NRMSE than with either alone. Although our optimal parameters
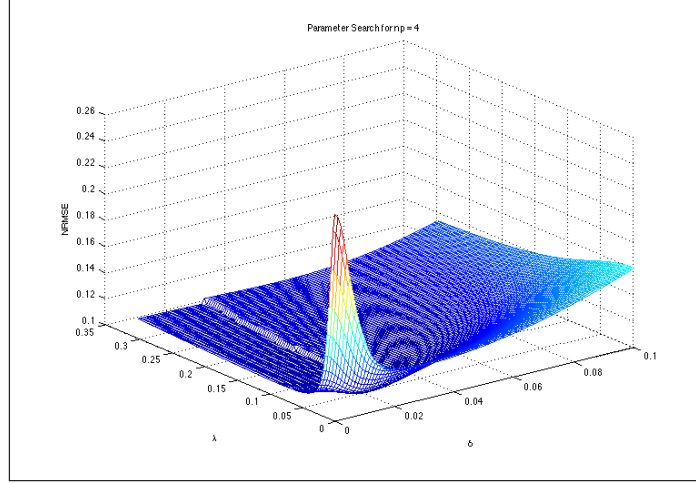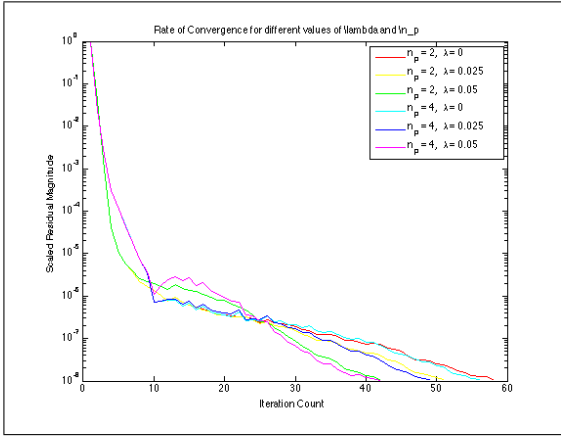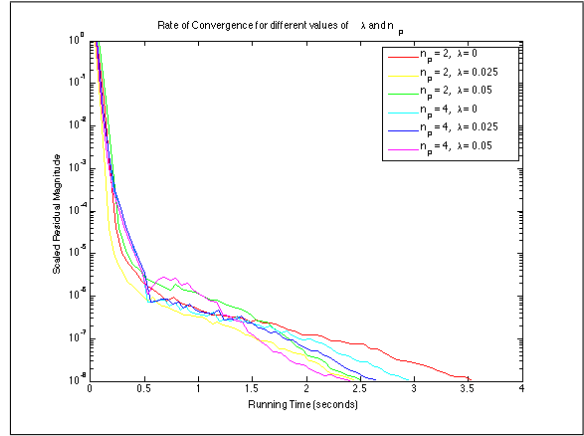
Figure 4.8: Parameter sweep over complex regularization constant λ and energy penalty constant δ. Optimal NRMSE of 10.39% is achieved at $\lambda = 0.0485$ and $\delta = 0.0222$ for $n_p = 4$ under-sampling.

unsurprisingly disagree with those of [39], this result is in support of PFPP imaging.

(a) Residual Magnitude vs. Iteration Count

(b) Residual Magnitude vs. Running Time

Figure 4.9: Convergence rates of CG algorithm over an array of under-sampling factors n_p, as a function of 4.9(a) number of iterations and 4.9(b) run time, respectively. The algorithm is monotonic during an initial phase of the optimization, and then loses monotonicity as solutions oscillate slightly.
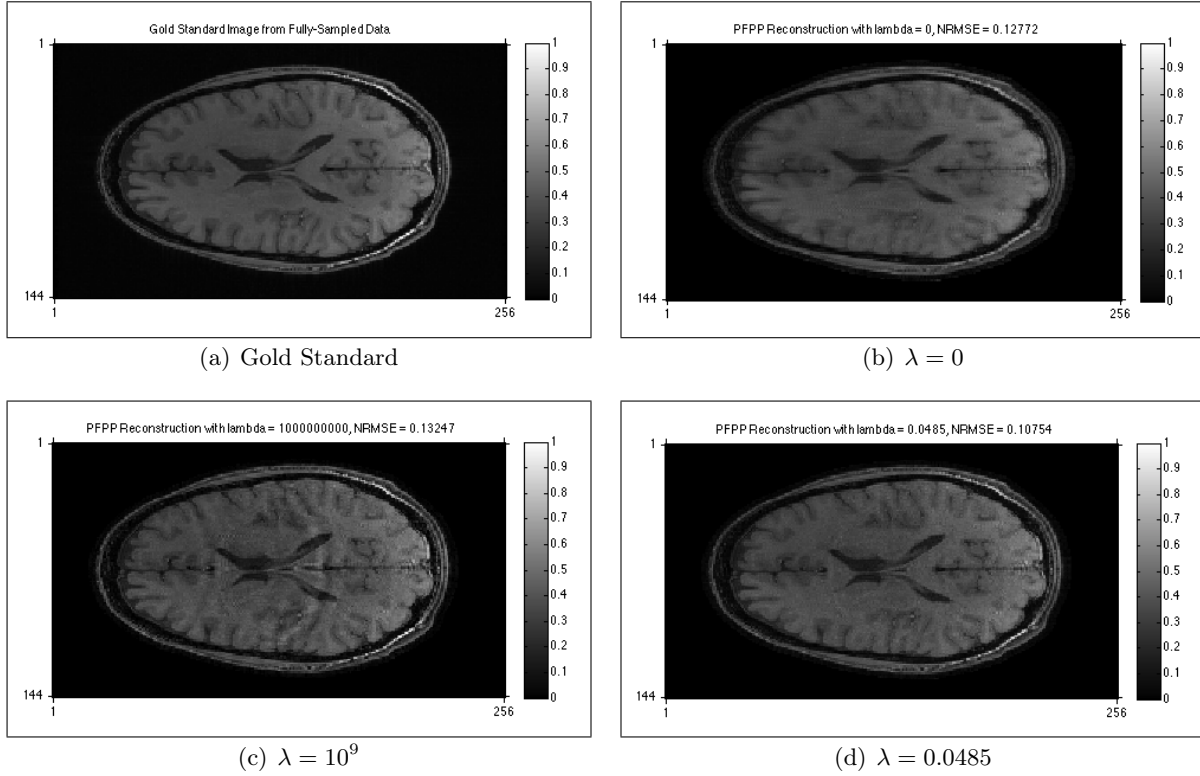
59

(a) Gold Standard

(b) $\lambda = 0$

(c) $\lambda = 10^9$

(d) $\lambda = 0.0485$

Figure 4.10: Asymptotic behavior of PFPP with respect to complex penalty λ, at $8\times$ acceleration. 4.10(a): Gold Standard Image, computed from fully-sampled k-space data. 4.10(b): With $\lambda = 0$, no imaginary constraint worsens conditioning of the problem, causing noise amplification and reduced contrast; 4.10(c): For large values of λ, over-suppression of imaginary components causes reconstruction errors in regions of rapid phase variation, making aliasing more apparent; 4.10(d) Intermediate values of λ strike a balance between these effects.

Figure 4.9 shows the scaled residual magnitude versus iteration count and run time, over an array of under-sampling factors n_p. The important aspect of this plot is that most of the residual magnitude is reduced monotonically in the first ten iterations, and then it decreases more slowly and slightly non-monotonically, perhaps due to the non-linearity of penalties on $\|[\text{Im}]\{x\}\|$.

It is interesting to explore image degradation at the asymptotic regimes of λ, i.e as λ goes to 0 and ∞ (Figure 4.10). Recall that when $\lambda = 0$, this reconstruction represents pure SENSE reconstruction, and as $\lambda \to \infty$, we approach a pure PF reconstruction (pure-real solution) [39].

Notice that for $\lambda = 0$, the reconstruction has reduced contrast due to noise amplification; this is to be expected, as we are not regularizing the imaginary component of the image. The improvement over Figure 4.4(b) is only due to the nonzero energy minimization constraint; image quality worsens in this regime dramatically as n_p is increased. As $\lambda \to \infty$, phase artifacts in the reconstructed image become very apparent, resulting in more visible aliasing, as is expected in pure-PF imaging. PFPP allows for a careful balance of the consequences of these two acceleration methods, allowing for reasonable noise suppression without excess phase artifacts introduced in the reconstruction.

60

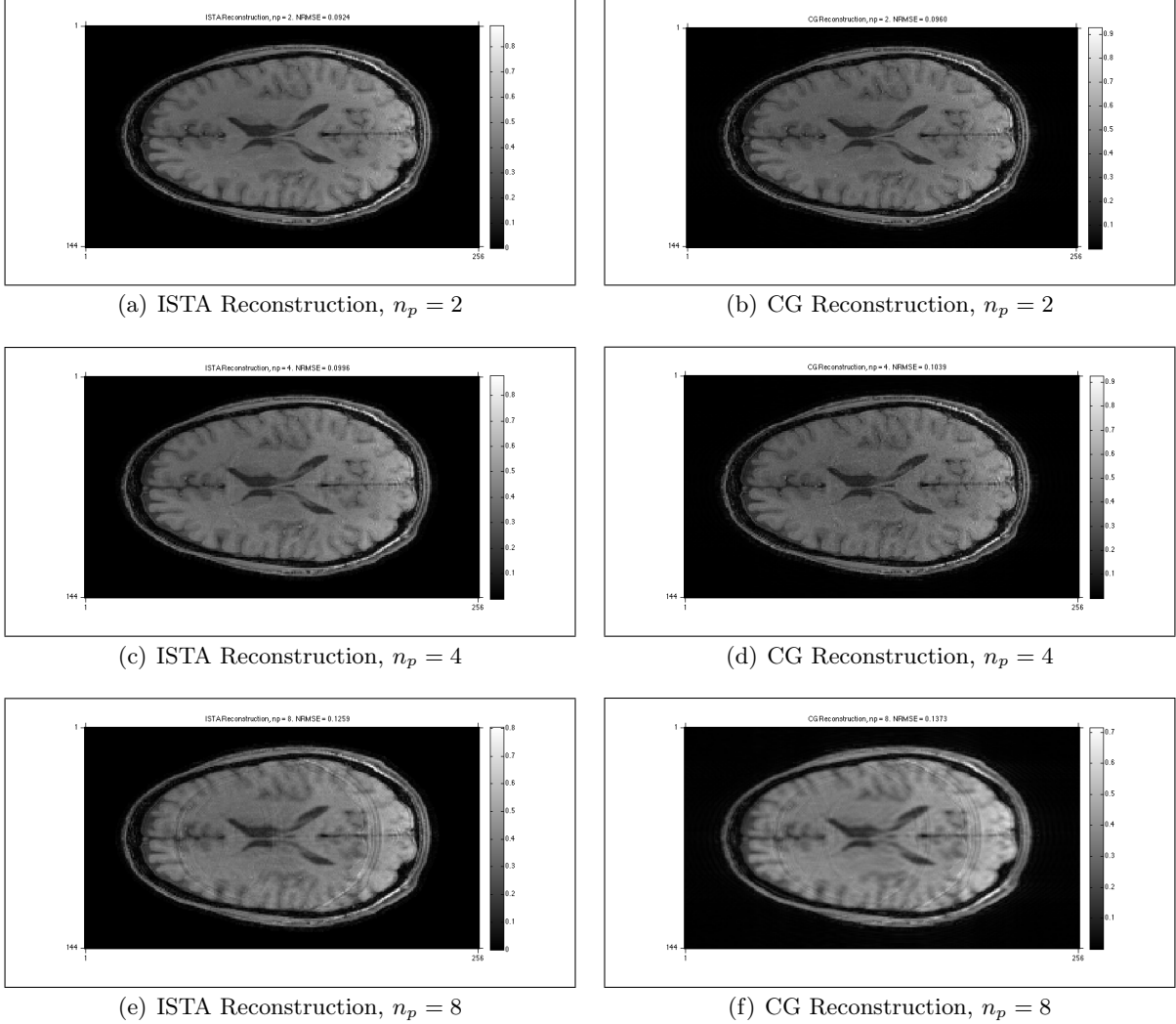### 4.4.3 Extension: Reconstruction Through Sparsity Enforcement

(a) ISTA Reconstruction, $n_p = 2$

(b) CG Reconstruction, $n_p = 2$

(c) ISTA Reconstruction, $n_p = 4$

(d) CG Reconstruction, $n_p = 4$

(e) ISTA Reconstruction, $n_p = 8$

(f) CG Reconstruction, $n_p = 8$

Figure 4.11: A comparison of the ISTA and CG methods. 4.11(a), 4.11(b): $n_p = 2$, resulting in NRMSE = 0.0924 for ISTA and NRMSE = 0.0960 for CG. 4.11(c), 4.11(d): $n_p = 4$, resulting in NRMSE = 0.0996 for ISTA and NRMSE = 0.1039 for CG. 4.11(e), 4.11(f): $n_p = 8$, resulting in NRMSE = 0.1259 for ISTA and NRMSE = 0.1373 for CG. At all levels of under-sampling, we achieve lower NRMSE using ISTA than CG; however, this is at the expense of increased computation time.

In Figure 4.11, we compare ISTA and CG reconstruction over several under-sampling ratios. Each reconstruction method is run using optimized parameters to compare best performances. Consistently, we observe a reduction in noise, through enforcement of piecewise-smoothness. Furthermore, image ranges are generally better constrained to the range of the gold standard because we no longer discourage total signal energy in our solution. These effects translate to consistently lower SNR across all values of n_p. Thus, our modification to the objective used in [39] consistently produces images which better represent the gold standard, in an NRMSE sense.
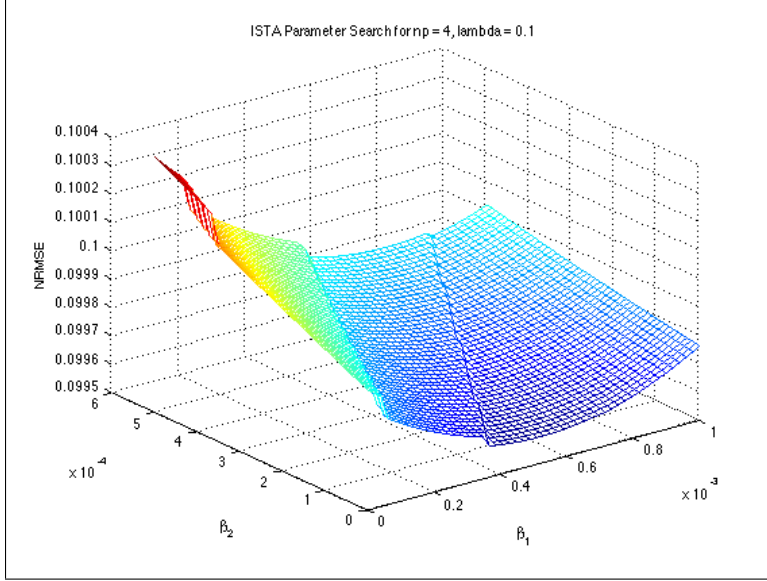
Figure 4.12: Parameter search at $n_p = 4$ under-sampling, over β_1 and β_2, holding optimal $\lambda = 0.10$ constant. NRMSE increases instantly as β_2 is increased from non-zero values due to over-regularization of the imaginary component of the solution. A finite and nonzero optimal $\beta_1 = 5.102 \times 10^{-4}$ indicates the effectiveness of ℓ_1 regularization on the real component of the image. A minimum NRMSE of 9.96% is achieved at optimal $(\lambda, \beta_1, \beta_2) = (0.10, 5.102 \times 10^{-4}, 0.000)$.

A parameter search (similar to the one shown in Figure 4.8) for the ISTA algorithm was also performed. For sake of simplicity, we hold $\lambda = 0.10$ constant at the optimal value, and vary β_1 and β_2. It is interesting to note that almost any increase in β_2 will result in a larger NRMSE. We hypothesize that this is because the combination of both quadratic regularization and ℓ_1 regularization results in pushing the imaginary component almost completely to 0. In a perfectly reconstructed image, however, we expect some phase contribution; because of this essential phase cancellation, we believe that the NRMSE begins increasing with β_2 immediately. However, β_1 assists in reducing NRMSE because the majority of image energy is contained in the real domain. As can be seen in Figure 4.11, the β parameters also help to smooth out the image somewhat, which is to be expected.

It should be noted that convergence times and iteration counts for our ISTA algorithm are noticeably less appealing than the CG formulation, as is often observed with basic ISTA implementations [46]. In essence, our extension trades convergence speed for noise reduction, piecewise smoothness and accuracy with respect to the gold standard. In practice, this could be an appealing trade in certain applications, as our methods allow for increased acceleration during acquisition (the more expensive step!), at the expense of reduced speed during reconstruction, for a given noise tolerance. However, as further discussed in the Extensions, it is desirable to explore alternative algorithms that produce acceleration increases at both the acquisition and reconstruction stages.

4.5 Conclusions and Future Work

Using the PFPP algorithm, there is a 2-5% decrease in NRMSE when compared to a simple SENSE reconstruction method (equivalent to $\lambda = 0$). Although the results of this study are not as strong as those in [39], we contend that the images used in this paper are far more realistic than those used as test subjects in [39]. Therefore, it is reasonable to imagine that the NRMSE reduction would not be as high as when the images to be reconstructed are not from real data.

The extension of ISTA to this reconstruction problem showed a further 1-3% decrease in NRMSE, and qualitatively the images look less noisy than their CG counterparts. In the end, ISTA illustrates a trade-off between time and quality; for the decrease in NRMSE we give up an average of about 30-40 iterations. If the computer that is running the reconstruction algorithm is fast and there are only a few reconstructions to be run, this is not a problem. However, as the number of reconstructions grows, the small gain in NRMSE might not be worth the extra time and computing cost. It should be noted that simple extensions to ISTA, i.e., [46] are available which could speed up the algorithm significantly. Also, in a low-SNR situation, the ISTA algorithm would be preferable, as it exudes greater robustness to noise in the image reconstruction process.

Overall, this PFPP reconstruction method is a promising step in reducing scan time, while preserving image integrity. Even with an acceleration factor of $8\times$, almost all of the detail from the gold standard is retained. The ability to reduce scan time by a significant factor is important, as decreasing scan time is associated with both decreased cost and patient discomfort. This reconstruction method succeeds in accelerating acquisitions, while still reconstructing acceptable images.

As future work, it is of interest to explore minimization of PFPP objectives similar to Equation (4.7) through optimization methods different from ISTA (or variations). Specifically, we are interested in applying variable-splitting methods explored in recent literature [47]. Such objectives have been examined by forming an Augmented Lagrangian (AL) through a transform of variables ($v = \mathbf{W}x$), and then attempting a joint minimization over both x and v using sequential updates:

$$\hat{x} = \arg\min_x \min_v ||y - \mathbf{FSM}x||_2^2 + \lambda^2||\mathrm{Im}\{x\}||_2^2 + \beta||v||_1 + \gamma||v - \mathbf{W}x||_2^2 \qquad (4.20)$$

Algorithms such as Alternating Directional Method of Multipliers (ADMM) can be used to perform this joint minimization. However, we will need to be careful when dealing with the additional $\mathrm{Im}\{x\}$ term. As with the ISTA minimization, it may behoove us to further separate the objective into alternating minimizations on $[\mathrm{Re}]\{x\}$ and $[\mathrm{Im}]\{x\}$ separately, and then stitch the estimated solution $\hat{x} = \hat{x_R} + \hat{x_J}$ back together after individual minimizations.

Bibliography

[1] 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, no. 1, pp. 231–5, Jan. 1977.

[2] J. Homer and M. S. Beevers, "Driven-equilibrium single-pulse observation of T1 relaxation. A reevaluation of a rapid "new" method for determining NMR spin-lattice relaxation times," *J. Mag. Res.*, vol. 63, no. 2, pp. 287–97, June 1985.

[3] G. Nataraj, J-F. Nielson, and J. A. Fessler, "Regularized, joint estimation of T1 and M0 maps," in *Proc. Intl. Soc. Mag. Res. Med.*, 2014, To appear.

[4] F. Bloch, "Nuclear induction," *Phys. Rev.*, vol. 70, no. 7-8, pp. 460–74, Oct. 1946.

[5] Y. Zur, M. L. Wood, and L. J. Neuringer, "Spoiling of transverse magnetization in steady-state sequences," *Mag. Res. Med.*, vol. 21, no. 2, pp. 251–63, Oct. 1991.

[6] A. Macovski, "Noise in MRI," *Mag. Res. Med.*, vol. 36, no. 3, pp. 494–7, Sept. 1996.

[7] T. Lei, "Statistics of MR signals: revisited," in *Proc. SPIE 6510 Medical Imaging 2007: Phys. Med. Im.*, 2007, p. 651052.

[8] 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, no. 3, pp. 515–26, Mar. 2003.

[9] 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, no. 1, pp. 194–9, Jan. 2004.

[10] C. A. Cocosco, V. Kollokian, R. K-S. Kwan, and A. C. Evans, "BrainWeb: Online interface to a 3D MRI simulated brain database," in *Proc. 3rd Intl. Conf. on Functional Mapping of the Human Brain*, 1997, p. S425, NeuroImage, vol. 5 (4, part2/4), May.

[11] S. Ramani, Z. Liu, J. Rosen, J-F. Nielsen, and J. A. Fessler, "Regularization parameter selection for nonlinear iterative image restoration and MRI reconstruction using GCV and SURE-based methods," *IEEE Trans. Im. Proc.*, vol. 21, no. 8, pp. 3659–72, Aug. 2012.

[12] 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, no. 4, pp. 805–24, Oct. 2012.

[13] 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, no. 5, pp. 1536–43, Nov. 2012.

[14] 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.*, 2014, In press.

[15] W. S. Hinshaw, "Image formation by nuclear magnetic resonance: The sensitive point method," *J. Appl. Phys.*, vol. 47, pp. 3709, 1976.

[16] 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.

[17] 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, no. 1, pp. 35–42, May 1988.

[18] G. H. Welsch, K. Scheffler, T. C. Mamisch, T. Hughes, S. Millington, M. Deimling, and S. Trattnig, "Rapid estimation of cartilage T2 based on double echo at steady state (DESS) with 3 Tesla," *Mag. Res. Med.*, vol. 62, no. 2, pp. 544–9, Aug. 2009.

[19] M. L. Gyngell, "The steady-state signals in short-repetition-time sequences," *J. Mag. Res.*, vol. 81, no. 3, pp. 474–83, Feb. 1989.

[20] W. Hänicke and H. U. Vogel, "An analytical solution for the SSFP signal in MRI," *Mag. Res. Med.*, vol. 49, no. 4, pp. 771–5, Apr. 2003.

[21] G. Nataraj, J-F. Nielson, 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*, 2014, Submitted.

[22] H. Wang and Y. Cao, "Spatially regularized T1 estimation from variable flip angles MRI," *Med. Phys.*, vol. 39, no. 7, pp. 4139–48, 2012.

[23] 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.

[24] G. H. Golub and C. F. Van Loan, "An analysis of the total least squares problem," *SIAM J. Numer. Anal.*, vol. 17, no. 6, pp. 883–93, Dec. 1980.

[25] H. Erdoğan and J. A. Fessler, "Monotonic algorithms for transmission tomography," *IEEE Trans. Med. Imag.*, vol. 18, no. 9, pp. 801–14, Sept. 1999.

[26] Y. Nesterov, "A method for unconstrained convex minimization problem with the rate of convergence $O(1/k^2)$," *Dokl. Akad. Nauk. USSR*, vol. 269, no. 3, pp. 543–7, 1983.

[27] D. Kim, S. Ramani, and J. A. Fessler, "Accelerating X-ray CT ordered subsets image reconstruction with Nesterov's first-order methods," in *Proc. Intl. Mtg. on Fully 3D Image Recon. in Rad. and Nuc. Med*, 2013, pp. 22–5.

[28] A. Funai and J. A. Fessler, "Cramer Rao bound analysis of joint B1/T1 mapping methods in MRI," in *Proc. IEEE Intl. Symp. Biomed. Imag.*, 2010, pp. 712–5.

[29] A. O. Hero, J. A. Fessler, and M. Usman, "Exploring estimator bias-variance tradeoffs using the uniform CR bound," *IEEE Trans. Sig. Proc.*, vol. 44, no. 8, pp. 2026–41, Aug. 1996.

[30] G. Golub and V. Pereyra, "Separable nonlinear least squares: the variable projection method and its applications," *Inverse Prob.*, vol. 19, no. 2, pp. R1–26, Apr. 2003.

[31] J. Barzilai and J. Borwein, "Two-point step size gradient methods," *IMA J. Numerical Analysis*, vol. 8, no. 1, pp. 141–8, 1988.

[32] M. Raydan, "On the Barzilai and Borwein choice of steplength for the gradient method," *IMA J. Numerical Analysis*, vol. 13, no. 3, pp. 321–6, 1993.

[33] M. Raydan, "The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem," *SIAM J. Optim.*, vol. 7, no. 1, pp. 26–33, 1997.

[34] M. F. Duarte, S. Sarvotham, D. Baron, M. B. Wakin, and R. G. Baraniuk, "Distributed compressed sensing of jointly sparse signals," in *Asilomar Conf. on Signals, Systems and Computers*, 2005, pp. 1537–41.

[35] D. Baron, M. B. Wakin, M. F. Duarte, S. Sarvotham, and R. G. Baraniuk, "Distributed compressed sensing," 2005.

[36] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Sig. Proc.*, vol. 53, no. 7, pp. 2477–88, July 2005.

[37] J. A. Fessler and B. P. Sutton, "Nonuniform fast Fourier transforms using min-max interpolation," *IEEE Trans. Sig. Proc.*, vol. 51, no. 2, pp. 560–74, Feb. 2003.

[38] R. Heule, C. Ganter, and O. Bieri, "Triple echo steady state (TESS) relaxometry," in *Proc. Intl. Soc. Mag. Res. Med.*, 2013, p. 467.

[39] M. Bydder and M. D. Robson, "Partial Fourier partially parallel imaging," *Mag. Res. Med.*, vol. 53, no. 6, pp. 1393–401, June 2005.

[40] K. F. King and L. Angelos, "SENSE with partial Fourier homodyne reconstruction," in *Proc. Intl. Soc. Mag. Res. Med.*, 2000, p. 153.

[41] D. C. Noll, D. G. Nishimura, and A. Macovski, "Homodyne detection in magnetic resonance imaging," *IEEE Trans. Med. Imag.*, vol. 10, no. 2, pp. 154–63, June 1991.

[42] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: sensitivity encoding for fast MRI," *Mag. Res. Med.*, vol. 42, no. 5, pp. 952–62, Nov. 1999.

[43] F. Wiesinger, P-F. V. . Moortele, G. Adriany, N. D. Zanche, K. Ugurbil, and K. P. Pruessmann, "Parallel imaging performance as a function of field strength - An experimental investigation using electrodynamic scaling," *Mag. Res. Med.*, vol. 52, no. 5, pp. 953–64, Nov. 2004.

[44] J. R. Shewchuk, "An introduction to the conjugate gradient method without the agonizing pain," 1994, http://atol.ucsd.edu/ pflatau/scatlib/conjugate.html.

[45] M. W. Jacobson and J. A. Fessler, "An expanded theoretical treatment of iteration-dependent majorize-minimize algorithms," *IEEE Trans. Im. Proc.*, vol. 16, no. 10, pp. 2411–22, Oct. 2007.

[46] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sci.*, vol. 2, no. 1, pp. 183–202, 2009.

[47] M. Le, S. Ramani, and J. A. Fessler, "An efficient variable splitting based algorithm for regularized SENSE reconstruction with support constraint," in *Proc. Intl. Soc. Mag. Res. Med.*, 2013, p. 2654.