

Magnetic Resonance Imaging: Myelin Water Imaging and Model-Based Image Reconstruction

by

Steven T. Whitaker

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical and Computer Engineering)
in the University of Michigan
2022

Doctoral Committee:

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

Steven T. Whitaker
stwhit@umich.edu
ORCID iD: 0000-0003-1170-7653

© Steven T. Whitaker 2022

DEDICATION

*To Nicala, my best friend
and the mother of our beautiful children.*

ACKNOWLEDGMENTS

Earning a PhD requires many resources, not the least of which is time. I spent a lot of time in this endeavor, of course, but the time I spent was made more productive and pleasant because of the time others invested in me. I cannot hope to mention everyone by name, but thank you to all who have spent some of their time to bless my life in any way over the past five years.

I thank my co-advisors, Jeffrey A. Fessler and Jon-Fredrik Nielsen, for their constant support. Probably the biggest reason why I chose to come to Michigan was because I wanted to work with Jeff. From day one, Jeff has always done an excellent job at pointing me to resources and giving suggestions while also encouraging autonomy and giving me time to work through things at my own pace. I especially thank him for his support as I went through some difficult setbacks. Jeff also introduced me to the Julia programming language, which quickly became my favorite language to work with. For his part, Jon has similarly been a great mentor, and I was glad to have him be one of my co-advisors. I especially thank him for the time he spent getting me data for my work and helping me understand many of the practical considerations one needs to keep in mind when scanning. Both Jeff and Jon gave valuable feedback on my work, and it was a pleasure having the perspective both of a signal processing person as well as a physicist. I should mention also that both of my co-advisors did an excellent job adjusting to the COVID-19 pandemic. Thank you, Jeff and Jon, for your guidance, even in uncertain times.

I thank my committee members Douglas Noll, Clayton Scott, and Nicole Seiberlich. The perspective they brought to my work was enlightening and invaluable, and the time they spent on my committee has helped make this dissertation a better document.

When I started my PhD, I had the privilege of working with then-senior PhD student Gopal Nataraj and then-undergraduate student Mingjie Gao. I thank them for the time they spent helping me get up to speed with many of the methods I ended up using in this dissertation, particularly the PERK estimator and scan parameter optimization using the Cramér-Rao Lower Bound.

I thank Scott Swanson for the time he has taken to discuss tissue modeling and various considerations for multi-compartment, exchanging tissue models. Much of the knowledge he imparted is reflected in the tissue model I used in Chapter 4 in this dissertation.

I thank my fellow labmates for our shared time together. In particular, I thank Cameron Blocker for being my BYU connection and letting me know that Jeff was a great advisor; Caroline Crockett for going through the PhD program at the same time as me; David Hong for sharing his enthusiasm for Julia; and Naveen Murthy for working on a similar project so we could bounce ideas off each other.

I thank my parents, siblings, and their spouses, as well as my in-laws, for their support for as long as I've known them (which goes back further than five years). I am especially grateful for their words of encouragement and their prayers on my behalf.

I thank my children Lynelle, William, and Michael for bringing joy to my life and giving me opportunities to become more patient and kind. Each of them is unique, and it is exciting to see them grow and learn and interact with each other and others around them.

Most of all, I thank my dear wife Nicala for her love, encouragement, patience, and kindness for the six years we have been married. She is constantly focused on helping each of our children (and me!) become the best we can be. Nicala, you truly are the light of my life, and there is no other person I would rather spend eternity with, let alone complete a PhD. Thank you for your constant support and prayers for me. This dissertation is dedicated to you.

TABLE OF CONTENTS

DEDICATION	ii
ACKNOWLEDGMENTS	iii
LIST OF FIGURES	viii
LIST OF TABLES	xvii
ABSTRACT	xx
Chapter 1: Introduction	1
Chapter 2: Background	3
2.1 MRI Physics	3
2.1.1 Bloch Equation	5
2.1.2 Bloch-McConnell Equation	7
2.1.3 MRI Scans	9
2.2 Small-Tip Fast Recovery (STFR)	9
2.3 Parameter Estimation via Regression with Kernels (PERK)	11
2.4 Myelin Water Imaging (MWI)	13
2.4.1 Multi-Echo Spin Echo (MESE)	13
2.4.2 Tissue Models for MWI	15
Chapter 3: Myelin Water Fraction Estimation Using Small-Tip Fast Recovery MRI	19
3.1 Introduction	19
3.2 Background	21
3.2.1 STFR	21
3.2.2 MESE	23
3.2.3 Scan Design Using the Cramér-Rao Lower Bound	24
3.2.4 Parameter Estimation via Regression with Kernels (PERK)	25
3.3 Methods	26

3.3.1	Scan Design	26
3.3.2	MWF Estimation	27
3.4	Results	30
3.4.1	Scan Design	30
3.4.2	MWF Estimation	31
3.5	Discussion	34
3.6	Conclusion	40
3.7	Supporting Information	41
3.7.1	Estimator RMSE for White and Gray Matter Tissue Values	41
3.7.2	Estimator Performance with Model Mismatch	41
3.7.3	Estimator Bias	44

Chapter 4: End-to-End MRI Scan Design Approach for Improved Myelin

	Water Imaging	49
4.1	Introduction	49
4.2	Theory	51
4.3	Methods	52
4.3.1	Scan Designs	54
4.3.2	MWF Estimation	56
4.4	Results	58
4.4.1	Scan Designs	58
4.4.2	MWF Estimation	62
4.5	Discussion	62
4.6	Conclusion	68
4.7	Supporting Information	69
4.7.1	Choosing PERK Training Distributions for $T_{1,f}$ and $T_{1,s}$	69
4.7.2	Comparison of Initial and Optimized Scan Designs	71
4.7.3	<i>In Vivo</i> MWF Map Using Computed Noise	72
4.7.4	Noise Cross-Validation	72
4.8	Appendix	77
4.8.1	De-biasing	77
4.8.2	Other BrainWeb Parameter Estimates	78
4.8.3	Other <i>In Vivo</i> Parameter Estimates	81

Chapter 5: Intravoxel B0 Corrected Image Reconstruction with RF Prephasing

5.1	Introduction	87
-----	------------------------	----

5.2	Background	88
5.3	Theory	90
5.4	Methods	91
5.5	Results	94
5.6	Discussion	98
5.7	Conclusion	101
Chapter 6: Future Work		102
6.1	Model Intravoxel Dephasing in MWI	102
6.2	Jointly Reconstruct Under-sampled STFR Scans	102
6.3	Replace PERK with Deep Learning	104
6.4	Explore Bias-Variance Trade-Off in End-to-End Scan Design	104
6.5	Optimize Magnetic Resonance Fingerprinting Scans	105
6.6	Investigate CSF Partial Volume Effects	106
Appendices		108
Bibliography		113

LIST OF FIGURES

- 2.1 Image of myelin from an electron microscope. The myelin sheath (shaded in blue) consists of layers of myelin (myelin bilayers). Water trapped within the myelin bilayers is called myelin water, and all other water (both within the axon—the area enclosed by the myelin sheath—and without the axon) is called non-myelin water. (Original image from <http://www.teachmentortexts.com/2013/01/be-agent-of-change.html>, accessed September 18, 2020 at 9:57 AM EDT.) 14
- 2.2 Tissue models for MWI. The two compartments of the two-compartment model (a) are myelin water and non-myelin water, each with its own T_1 , T_2 , compartment fraction, and (for myelin water) off-resonance frequency. Exchange between myelin water and non-myelin water can also be modeled (not shown). The three-compartment model (b) adds a macromolecule compartment, with its own T_1 , T_2 , and compartment fraction, as well as exchange from myelin water to macromolecules. The subscript “f” is for “fast” or “fast-relaxing” and is used for myelin water because it experiences faster relaxation than non-myelin water; conversely, non-myelin water uses the subscript “s” for “slow” or “slow-relaxing”. The subscript “m” stands for “macromolecule”. 16
- 3.1 Workflow of the proposed methods. We first optimized a set of STFR scan parameters by minimizing a Cramér-Rao Lower Bound, then acquired data using those scans, as well as Bloch-Siegert (BS) scans. Two of the STFR scans were equivalent to SPGR scans, so were used to estimate $\Delta\omega$, and the BS scans were used to estimate κ . These parameters were treated as known values in the MWF estimation step. We then generated noisy training data using an STFR signal model. Finally, we passed the training data, acquired STFR images, and known parameters to PERK to estimate MWF voxel-by-voxel. 22

3.2	<p>The two optimized scans (see Table 3.2) were evaluated to explore how including the additional myelin water off-resonance $\Delta\omega_f$ in the design process affects performance. Design A (where $\Delta\omega_f$ was included in the optimization) has a better expected CRLB for every value of $\Delta\omega_f$ within the 5–35 Hz range over which design A was optimized. Design B (where $\Delta\omega_f$ was ignored) understandably has a better expected CRLB for $\Delta\omega_f = 0$. For each value of $\Delta\omega_f$ investigated in this plot, the expected CRLB was computed with that value of $\Delta\omega_f$ held constant but unknown. Typical values of $\Delta\omega_f$ in white matter are 5–35 Hz [1]. Although these CRLB values predict estimator standard deviation (for an unbiased estimator) on the order of 100% MWF or more, we are not restricted by these large values because we are using a Bayesian estimator (see Figure 3.11 for details).</p>	32
3.3	<p>RMSE of MWF estimates from simulated test data for various ways of accounting for the additional myelin water off-resonance $\Delta\omega_f$. The diamond-markered green curve was generated using scan design B (where $\Delta\omega_f$ was ignored) and using PERK training data where $\Delta\omega_f = 0$. In other words, $\Delta\omega_f$ was not considered in any aspect, neither in the scan design nor when training. The square-markered red curve was also generated using scan design B, but the training data included a range of $\Delta\omega_f$ values. The circle-markered blue curve was generated using design A and a range of values of $\Delta\omega_f$. The latter two methods look almost identical, but both have much better RMSE than the first method as $\Delta\omega_f$ increases.</p>	33
3.4	<p><i>Right:</i> MWF maps from five methods using simulated test data for a three-compartment tissue model with exchange. Table 3.3 reports numerical results. The proposed STFR3-PERK estimates are closer to the true MWF value for white matter tissue values than are the MESE-NNLS estimates. <i>Left:</i> Bulk off-resonance $\Delta\omega$ and flip angle scaling κ maps used in this simulation. . . .</p>	35
3.5	<p>MWF maps from five methods using simulated test data for a nine-compartment tissue model with exchange. These results are essentially the same as when using the three-compartment exchanging model (see Figure 3.4). Thus, even though STFR3-PERK was trained with a three-compartment exchanging model, it still produced good MWF estimates from signal generated using a nine-compartment exchanging model.</p>	36

3.6	MWF maps from <i>in vivo</i> MESE data and STFR data using scan design A. Table 3.4 shows numerical results for several manually selected regions of interest. The MESE-NNLS MWF map appears noisier than those shown in other works. This is likely due to the lower SNR of our data due to differences in voxel size. To match the STFR resolution, we acquired MESE with 1.1 mm isotropic voxels, whereas often MESE data is collected with slice thickness of 5 mm and 1.6 mm or greater in the phase encode direction. Remarkably, MESE-PERK is much less noisy than MESE-NNLS. This is likely due to PERK being a Bayesian estimator that discourages estimates that are far from the mean MWF training value.	37
3.7	RMSE of MWF estimates for white matter and gray matter simulated test data. Scan design A has better RMSE in white matter for values of $\Delta\omega_f$ we expect to see in white matter. This better RMSE in white matter is at the cost of worse RMSE in gray matter. Note that the values of $T_{1,f}$ and $T_{1,s}$ for gray matter were outside of the range of values used for the scan designs and for training our estimator.	42
3.8	MWF maps from five methods using simulated test data for a four-compartment tissue model with exchange. The four compartments considered were myelin water, axonal water (i.e., water in myelinated axons), all other water, and macromolecules. The results are similar to those using the three-compartment model with exchange. Table 3.5 shows numerical results. . . .	43
3.9	MWF maps from five methods using simulated test data for a three-compartment tissue model without exchange. Without exchange, the three-compartment model becomes essentially a two-compartment model because the T_2 of the macromolecular pool is so small. Thus it makes sense that STFR2-PERK performs well. Surprisingly, MESE-PERK still produces good MWF estimates, even though it is trained with the three-compartment exchanging model (like STFR3-PERK). This could be because the T_R of the MESE scan is long compared to the residence times governing exchange. Furthermore, it is possible that if the training ranges for the residence times were adjusted appropriately (increased) then STFR3-PERK would also do well. Table 3.6 shows numerical results.	45

3.10	<p><i>In vivo</i> images for two SPGR and nine STFR scans using scan design A. Each image is the square root sum of squares combination of the individual coil data. STFR produces contrast similar to balanced SSFP, including a similar off-resonance profile that induces the characteristic banding artifact of balanced SSFP. Different points of this profile are sampled as the phase ϕ of the STFR tip-up excitation varies. The nine STFR images are sorted by increasing ϕ, so this off-resonance profile is easily visualized. In the lower right is the field map estimated from the two SPGR scans.</p>	46
3.11	<p>Comparison of biased and unbiased CRLBs for white matter tissue values using the two-compartment non-exchanging model. The biased CRLB is much lower than the unbiased CRLB, suggesting that bias is the reason why our STFR-based MWF estimation results in estimates with low variance. However, our proposed method still shows sensitivity to changes in MWF (see Figure 3.12).</p>	47
3.12	<p>Expected MWF estimates from the proposed STFR3-PERK MWF estimation technique for fixed white matter tissue values from a three-compartment exchanging model. The proposed method is (mildly) biased, yet it is still very sensitive to changes in true MWF value. Furthermore, bias decreases as SNR increases. (An unbiased estimator would have estimates along the line of identity, i.e., along the dashed line.)</p>	48
4.1	<p>Diagram of the end-to-end scan design process, in which scan parameters \mathbf{P} are optimized subject to constraints \mathcal{P} to minimize the end-to-end cost function (4.3). Validation and training sets of tissue parameters are sampled before and held fixed during the iterative optimization. The end-to-end cost function is computed in several steps. First, the current candidate scan design \mathbf{P} is used to simulate noisy data, using a signal model \mathbf{f}, from both the validation and training tissue parameter sets. The training set of tissue parameters is then paired with the associated noisy signals and used to train the estimator $\hat{\mathbf{x}}(\cdot)$. Next, the noisy signals from the validation set of tissue parameters are given to the estimator to estimate the underlying tissue parameters. Finally, these estimates are compared to the true tissue parameters to compute a (weighted) MSE. The scan design process seeks to minimize this MSE.</p>	53

4.2	Optimized RMSE of MWF estimates from end-to-end scan designs with varying initializations and number of STFR scans. The 10 points in the plot for each given number of scans correspond to the scan designs optimized starting from 10 different initializations. From these results, we decided that using 18 scans was a good trade-off between scan time and performance.	59
4.3	Comparison of 10 CRLB scan designs to 10 end-to-end scan designs, each with 18 STFR scans. We computed the square root of the expected CRLB of MWF estimates (left subplot) and the RMSE of MWF estimates (right subplot) for each of the 20 scan designs. Optimizing scan parameters to minimize the RMSE results in scan designs that have lower RMSEs than scan designs optimized to minimize the expected CRLB. (The one CRLB design that has a RMSE on par with the end-to-end designs, indicated by the white triangle marker, was initialized with an end-to-end design and did not change much, i.e., was stuck near a poor local minimum, during the optimization process.)	60
4.4	Scan parameters of the best end-to-end scan design and the best CRLB scan design, sorted by increasing tip-down angle. Each design consists of 18 STFR scans, each with its own tip-down angle, tip-up angle, and tip-up phase (that were optimized). Each scan additionally had fixed $T_{\text{free}} = 8$ ms, $T_{\text{g}} = 2.8$ ms, and $T_{\text{E}} = 4$ ms. Interestingly, the end-to-end design has much lower flip angles, in general, than the CRLB design.	61
4.5	Histograms of estimated MWF versus true MWF obtained when using the end-to-end scan design and when using the CRLB scan design for the experiment where we drew the test data randomly from the distributions in Figure 4.1. Though only MWF is shown as varying, we varied all tissue parameters. The table shows the ensemble RMSE, average absolute bias, and average standard deviation of the MWF estimates. The MWF estimates obtained using the end-to-end scan design were more accurate, less biased, and less noisy than those obtained using the CRLB design.	63

4.6	Estimated MWF maps obtained when using the end-to-end scan design and when using the CRLB scan design. The true MWF varied from 0.1 to 0.2 in WM and from 0.02 to 0.05 in GM from image bottom to top. The results using the end-to-end design were more accurate, less biased, and less noisy than those obtained using the CRLB design. The second row shows the error maps for reference. The table shows the RMSE, mean, and standard deviation (SD) of the MWF estimates across all WM and GM voxels. For reference, the true MWF map has a mean value of 0.151 for WM and 0.035 for GM. . . .	64
4.7	Estimated MWF maps obtained from <i>in vivo</i> data acquired using the end-to-end scan design and the CRLB scan design. The MWF map from the end-to-end design was less noisy and appeared to have more reasonable values and decent WM/GM contrast. The table shows the mean and standard deviation (SD) of the MWF estimates across a pooled region of interest (ROI) consisting of WM voxels (see Figure 4.8).	65
4.8	One STFR scan with WM ROIs shown from which we computed the mean and standard deviation of the MWF estimates from the <i>in vivo</i> data. . . .	66
4.9	Estimated MWF maps obtained from <i>in vivo</i> data acquired using the end-to-end scan design and the CRLB scan design, with PERK trained with a noise standard deviation that was computed from background voxels of the reconstructed images. The MWF map from the end-to-end design had higher values than what one might expect in the brain, but was less noisy than the MWF map from the CRLB design. The table shows the mean and standard deviation (SD) of the MWF estimates across a pooled region of interest (ROI) consisting of WM voxels (see Figure 4.8).	74
4.10	MSE of fitted STFR signals versus PERK training noise standard deviation for the modified BrainWeb experiment. The noise standard deviation added to the BrainWeb data was $2^{-9.6}$ (denoted by the vertical line). The cross-validated noise standard deviation performed well in this case.	75
4.11	MSE of fitted STFR signals versus PERK training noise standard deviation for the <i>in vivo</i> data. The training noise standard deviation we manually chose based on visual preference (used for Figure 4.7) was $2^{-10.8}$ (denoted by the vertical line). The cross-validated noise standard deviation did not perform well in this case.	76

4.12	Trend lines used for de-biasing MWF estimates. We estimated MWF using PERK for 1,000 sets of non-scan parameters and averaged over 100 noise realizations. Here we show 2D histograms of the average estimates versus the true MWF values. We then fit a trend line to the scatter plot underlying each histogram. This trend line was used in de-biasing MWF estimates; when an estimate was given by PERK, we would look up the corresponding true MWF value predicted by the trend line and return that MWF value as the de-biased estimate.	77
4.13	M_0 BrainWeb estimates. The true M_0 values are shown before scaling by 3.	78
4.14	$T_{1,f}$ BrainWeb estimates.	78
4.15	$T_{1,s}$ BrainWeb estimates.	79
4.16	$T_{2,f}$ BrainWeb estimates.	79
4.17	$T_{2,s}$ BrainWeb estimates.	79
4.18	$\tau_{f \rightarrow s}$ BrainWeb estimates.	80
4.19	$\Delta\omega_f$ BrainWeb estimates.	80
4.20	$\Delta\omega$ BrainWeb estimates.	80
4.21	κ BrainWeb estimates.	81
4.22	M_0 <i>in vivo</i> estimates.	81
4.23	$T_{1,f}$ <i>in vivo</i> estimates.	82
4.24	$T_{1,s}$ <i>in vivo</i> estimates.	82
4.25	$T_{2,f}$ <i>in vivo</i> estimates.	83
4.26	$T_{2,s}$ <i>in vivo</i> estimates.	83
4.27	$\tau_{f \rightarrow s}$ <i>in vivo</i> estimates.	84
4.28	$\Delta\omega_f$ <i>in vivo</i> estimates.	84
4.29	$\Delta\omega$ <i>in vivo</i> estimates alongside a reference B0 map acquired from two T_E -offset SPGR scans. While some tissue structure remains visible in the $\Delta\omega$ estimates from the end-to-end design, overall there is good visual agreement between the estimated B0 map and the reference one.	85
4.30	κ <i>in vivo</i> estimates.	86
5.1	BrainWeb image weighted by the four individual coil sensitivities used in the simulation experiments. Only the middle slice of the simulated object is shown. Each coil “sees” only part of the object, but the information from each coil is used jointly during reconstruction to determine the true underlying object.	92

5.2	Off-resonance map and its spatial gradients used in the simulation experiments. Only the values corresponding to the middle slice are shown. The B0 gradients in the x and y directions range from -3.6 to 3.6 Hz/voxel. For $T_E = 21.6$ ms, this corresponds to less than $1/12$ cycle/voxel at most. In the z-direction, the peak of -61 Hz/voxel leads to more than one cycle/voxel.	93
5.3	True and reconstructed images for k-space data simulated with and without ideal RF prephasing. We compared three reconstruction approaches: square root sum-of-squares (SSoS) reconstruction and model-based reconstruction with and without modeling intravoxel B0 and RF prephasing effects. The SSoS reconstructed images were scaled for display. Without RF prephasing, the reconstructed images had severe signal loss due to intravoxel dephasing. This signal loss was mitigated, but not entirely removed due to (near) complete signal loss that occurs at and near the zero-crossings of the sinc function in (5.4), when reconstructing with a model that incorporated such intravoxel dephasing. Ideal RF prephasing by itself essentially eliminated all signal loss due to intravoxel dephasing, regardless of reconstruction method.	96
5.4	True and reconstructed images for k-space data simulated with and without ideal RF prephasing using a single-shot 3D EPI readout. The reconstruction approaches were the same as in Figure 5.3. The SSoS reconstructed images were scaled for display. Overall, these reconstructed images are of worse quality than those from the stack-of-EPI data, but they do show similar trends. The most notable difference, however, is that ideal RF prephasing did not recover the true image without also being combined with model-based reconstruction using a model that includes intravoxel B0 and RF prephasing effects.	97
5.5	Reconstructed images for k-space data simulated with varying amounts of RF prephasing. We reconstructed these images using a model that included intravoxel B0 and RF prephasing effects. As expected, the reconstructed image quality improved as the rephasing time increased from 0 (no prephasing) to T_E (ideal prephasing). The residual signal loss artifact was still severe for rephasing times $t \in \{0, 5\}$ ms. Starting at a rephasing time of 10 ms, however, the reconstructed image quality improved significantly.	98

5.6	Reconstructed images for k-space data simulated with varying amounts of RF prephasing. We reconstructed these images using a model that did not include intravoxel B0 and RF prephasing effects. The image quality of these reconstructed images degrades much less gracefully as the rephasing time decreases compared to Figure 5.5, which has images reconstructed using a model that included intravoxel B0 and RF prephasing effects.	99
A.1	Steady state STFR signal acquired at $T_E = 0.5$ ms for a three-compartment, exchanging system using instantaneous and non-instantaneous RF excitation pulses versus bulk off-resonance frequency $\Delta\omega$. The signals are essentially the same in both cases.	111

LIST OF TABLES

3.1	Unknown and known parameters used in scan design and in simulation. Values were chosen to match literature values for white matter [1–4]. The columns “Design A” and “Design B” contain the tissue distributions used for computing the expectation in the CRLB scan design cost function (3.6). We used $M_0 = 1$ for the scan designs because it only scales the STFR signal. The columns “WM” and “GM” specify the tissue parameters used for simulating white matter and gray matter tissues, respectively. For such simulations, we hold the tissue parameters fixed across different voxels, but vary $\Delta\omega$ and κ . The last column contains the tissue parameter distributions from which we sampled sets of tissue parameters for training PERK. The line below κ separates parameters used in both the two-compartment and three-compartment models (above) from those used only in the three-compartment exchanging model (below). WM: white matter. GM: gray matter.	28
3.2	Optimized STFR scan parameters. All values have units of degrees. For design A, the additional myelin water off-resonance $\Delta\omega_f$ was taken to be unknown and distributed uniformly from 5 to 35 Hz. For design B, $\Delta\omega_f$ was ignored (i.e., taken to be known and equal to 0). Both design A and design B also include two fixed SPGR scans with $\alpha = 5^\circ$	31
3.3	Comparison of various methods of MWF estimation. The reported time refers to the entire estimation, combining the time to estimate MWF in white matter voxels and gray matter voxels; it also includes training time for the methods that use PERK. The best value in each column is highlighted. See Figure 3.4 for a visual comparison of these methods.	34
3.4	<i>Left</i> : White matter (WM) and gray matter (GM) regions of interest (ROIs). The underlying image is from a standard MP-RAGE acquisition, acquired in the same scan session and registered to the other scans. <i>Right</i> : Sample means \pm standard deviations of MWF estimates for four WM ROIs and one GM ROI. Figure 3.6 shows corresponding MWF maps.	35
3.5	Numerical results for Figure 3.8.	42
3.6	Numerical results for Figure 3.9.	44

4.1	Tissue parameter distributions used in scan designs and in simulation results, unless otherwise noted in the text. Parameters were uniformly distributed between the given minimum and maximum values. PERK training also used these same distributions, except for a slightly expanded M_0 distribution ($[0.6, 1.2]$ instead of the values in the table) to account for the uncertainty in <i>in vivo</i> M_0 values due to unknown scanner gain. Note that we chose the range for $\Delta\omega$ to cover one full period of the magnitude STFR signal model (which is periodic in $\Delta\omega$ with period $1/T_{\text{free}}$, where we fixed $T_{\text{free}} = 8$ ms for the scan design process).	54
4.2	Comparison of RMSEs of MWF estimates from simulated data of optimized 18-scan end-to-end designs with combinations of bulk off-resonance $\Delta\omega$ and flip angle scaling κ being assumed unknown and known parameters. The RMSE is unchanged regardless of whether a separate B0 map ($\Delta\omega$) is acquired, whereas acquiring a separate B1+ map (κ) improved the RMSE by about 13%. In this work, despite the predicted increase in MWF estimation error, we decided to forego acquiring separate B0 and B1+ maps for ease of practical <i>in vivo</i> implementation and to lessen the overall scan time.	58
4.3	Computed myelin water and non-myelin water T_1 values ($T_{1,f}$ and $T_{1,s}$, respectively) determined by combining the water and macromolecule pools associated with myelin and non-myelin, respectively, using (4.5). Values for R_m , R_w , $k_{m \rightarrow w}$, and $k_{w \rightarrow m}$ were determined using values in Table 1 in [5]. We also computed T_1 values $T_{1,f}^a$ and $T_{1,s}^a$ with an adjusted R_m to account for field strength, in which case R_m was scaled by $\frac{4.7 \text{ T}}{3.0 \text{ T}}$. Units for all table values are s^{-1} except for $T_{1,f/s}^{(a)}$, which have units of ms. Based on this range of values for $T_{1,f}$ and $T_{1,s}$, we trained PERK using $T_{1,f}$ values uniformly distributed from 200 to 500 ms and $T_{1,s}$ values uniformly distributed from 1500 to 2500 ms.	71
4.4	CRLB (a) and RMSE (b) of MWF estimates computed using the optimized 18-scan CRLB and end-to-end scan designs as well as the scan designs used for initialization. The end-to-end design used to initialize a CRLB scan design was initialized with the hand-crafted initialization. (The optimized RMSE of the end-to-end design initialized with the hand-crafted design is different than the initial RMSE for the CRLB design initialized with the end-to-end design due to slightly different PERK hyperparameter values.) The displayed percentages indicate the change relative to the initial value. Despite resulting in drastically worse CRLBs compared to the initial values in nearly all cases, the end-to-end designs significantly improved the RMSEs in all cases.	73

5.1	NRMSEs for each of the reconstructed images in (a) Figure 5.3 and (b) Figure 5.4. (a) Even though the NRMSE for the image reconstructed with a model including intravoxel B0 effects from data without RF prephasing was only 2.3%, residual artifacts are still clearly visible in the reconstructed image. On the other hand, both model-based reconstructed images from data with ideal RF prephasing had excellent NRMSEs, regardless of whether or not intravoxel B0 and RF prephasing effects were modeled. (b) The NRMSEs from the single-shot 3D EPI data were worse than those from the stack-of-EPI data (a), at least for the model-based reconstructed images. In this case, reconstructing the data with ideal RF prephasing requires using a model that includes intravoxel B0 and prephasing effects for good performance.	95
5.2	NRMSEs for each of the reconstructed images in Figures 5.5 and 5.6. As expected, the NRMSE improved as the amount of RF prephasing increased from no prephasing to ideal prephasing. The NRMSEs for the images reconstructed using a model that included intravoxel B0 and RF prephasing effects were much better overall than those of the images reconstructed without those intravoxel effects.	99
A.1	Steady state single-compartment magnetization using instantaneous and non-instantaneous RF excitation pulses. A macromolecule magnetization (especially the longitudinal component M_z) is much more sensitive to the duration of the excitation pulse than is a white matter magnetization. iRF: instantaneous excitation. niRF: non-instantaneous excitation.	111

ABSTRACT

Magnetic resonance imaging (MRI) is a useful imaging modality governed by complicated physics. One important application of MRI is myelin water imaging (MWI), in which myelin water fraction (MWF), a measure of myelin content, is estimated. As a measure of myelin content, MWF can be used to track the onset and progression of demyelinating diseases such as multiple sclerosis. The traditional method used for MWI uses a multi-echo spin echo (MESE) scan that is prohibitively slow and has not been widely adopted clinically. One aim of this dissertation is to introduce new methods for MWI that are faster than the traditional method and that will allow for higher quality MWF maps. This dissertation introduces using the small-tip fast recovery (STFR) MRI scan, which is faster than the traditional scan, for MWI. In addition, this dissertation develops a new method for optimizing scan parameters for improved MWI by minimizing the estimation error of the estimator to be used. A set of STFR scans are optimized to be informative about MWF to enable accurate estimation of MWF. MWF is estimated using parameter estimation via regression with kernels (PERK), a recently developed learning-based method that can learn nonlinear functions with practical training time. PERK is trained with data simulated using a tissue model that incorporates parameters that are neglected in the traditional method for MWI, namely chemical exchange and off-resonance frequency differences. STFR-based *in vivo* MWF estimates comparable to MESE-based MWF estimates are obtained in about 1/5 the scan time. In simulation, the normalized root mean squared error (NRMSE) of MWF estimates is reduced from 42% to 14%.

In addition to advances in MWI, another aim of this dissertation is to improve the quality of MRI images by correcting for physics-induced image imperfections. As implied by its name, MRI uses a large magnet for imaging. Spatial variations in the ideally uniform magnetic field due to magnetic susceptibility differences of different tissues can result in image artifacts, or imperfections in the reconstructed image. One such image artifact is signal loss, which occurs when the magnetic field varies too quickly across space, causing signals at different spatial locations to be out of phase with each other. Approaches exist that can correct for this effect to an extent, but not in very extreme cases. Therefore, this dissertation introduces a novel acquisition and reconstruction approach to tackle this problem. In particular, data is acquired with prephasing so the signals at different spatial locations start out of phase but then rephase as the signal is acquired. Then an image

is reconstructed from the data using a novel model-based reconstruction technique that accounts for both the spatial variation of the magnetic field as well as the proposed prephasing acquisition scheme. In simulation, the NRMSE of a reconstructed image is reduced from 2.3% when reconstructed without prephasing to 0.2% when reconstructed using the proposed prephasing-based approach.

CHAPTER 1

Introduction

Magnetic resonance imaging (MRI) is a popular medical imaging modality. It offers a non-invasive view of the imaged object with excellent soft-tissue contrast without the need for ionizing radiation. Furthermore, MRI is flexible, having the ability to obtain information about many different physical processes by adjusting acquisition parameters.

One challenge in MRI is obtaining information about tissue microstructure. Typically in MRI, the resolution of an image (i.e., the size of an image *voxel*, or 3D pixel) is on the order of 1 mm^3 . Such resolution suffices for many applications (e.g., tumor detection), but for phenomena that occur on a smaller scale (e.g., demyelination), more advanced techniques are needed. Typically, microstructural information is obtained from MRI by acquiring multiple MRI scans and fitting the data to some model that accounts for multiple sub-voxel compartments.

One application of MRI to obtaining information about tissue microstructure is myelin water imaging (MWI). MWI images myelin water, or the water trapped between the lipid bilayers of the myelin sheath, and provides information about the integrity of myelin structure. It does so by computing the myelin water fraction (MWF), the fraction of MRI signal that comes from myelin water. MWI is useful for tracking myelination and demyelination, and can be used to monitor related diseases such as multiple sclerosis. Though MWI was first demonstrated *in vivo* in humans in 1994 [6], and has been shown to correlate well with histological measures of myelin content [7], it is still not a clinically accepted method for determining myelin content due to the excessive scan times needed for conventional MWI methods. Furthermore, conventional MWI methods neglect to model some aspects of MRI physics, namely chemical exchange and differences in off-resonance frequencies between myelin water and non-myelin water.

One aim of this dissertation is to address the speed and modeling issues of conventional MWI methods. Specifically, I propose to optimize and use a faster MRI scan, small-tip fast recovery (STFR) [8], in a novel way for MWI. I estimate MWF from multiple STFR scans (optimized to be informative about MWF) using parameter estimation via regression with

kernels (PERK) [9], a recently developed learning-based method that can learn nonlinear functions with minimal training time. I train PERK with simulated data using a tissue model that models physics more faithfully by incorporating exchange and off-resonance frequency differences.

Another challenge in MRI is correcting for physics-induced image imperfections, or artifacts. The signal obtained in MRI is governed by many physical phenomena. Not accounting for the physics when reconstructing an image from MRI data makes reconstruction simpler and faster, but the result is a poorer quality image. Therefore, another aim of this dissertation is to improve the quality of MRI images by correcting for physics-induced artifacts by incorporating the physics into a model used in model-based reconstruction.

One physics-induced image artifact is signal loss due to spatial variations in the main B0 magnetic field. When the magnetic field varies too quickly across space, particularly within a single voxel, signals at different spatial locations within the voxel have different phases; thus the overall voxel signal is smaller than if everything was in phase. Approaches exist that can correct for this effect to an extent, but not in very extreme cases.

Therefore, in this dissertation I introduce a novel acquisition and reconstruction approach in which I acquire data with *prephasing* so the signals at different spatial locations within a voxel start out of phase but then *rephase* as the total signal is acquired. Then I reconstruct an image from the data using a novel model-based reconstruction technique that models both the intravoxel B0 variation as well as the proposed prephasing acquisition scheme.

The rest of the dissertation (including my contributions, which are contained in Chapters 3–5 and Appendix A and in the provided citations) is organized as follows:

- Chapter 2 explains relevant background material.
- Chapter 3 introduces the use of STFR and PERK for MWI. Its content comes from published journal paper [10] that extends conference abstracts [11, 12].
- Chapter 4 introduces an end-to-end approach to optimizing scan parameters for MWI. Its content comes from in-preparation journal paper [13] that extends conference abstract [14].
- Chapter 5 introduces a novel model-based reconstruction approach for correcting for intravoxel B0 variation using a prephasing acquisition. Its content is an extension of conference abstract [15].
- Chapter 6 suggests several possible areas of future work.
- Appendix A contains a brief investigation into the effects of magnetization transfer (MT) on PERK MWF estimates from STFR scans.

CHAPTER 2

Background

This chapter discusses necessary background information. I will start with a discussion of relevant MRI physics and describe the origin of MRI signals. I will continue with a discussion about the STFR scan and introduce a model describing the MRI signal that arises from a STFR scan. Next, I will describe PERK, a supervised learning algorithm for parameter estimation. Finally, I will describe MWI, including a definition of MWF and the current reference method for estimating it. I will describe background related to model-based reconstruction and intravoxel B0 effects in the relevant chapter (Chapter 5).

2.1 MRI Physics

Fundamentally, MRI is a quantum mechanical problem involving the spin property of individual nuclei. However, a classical description often suffices because for a 1 mm^3 volume (a typical voxel size in MRI) there are *a lot* of nuclei.

A nucleus with nonzero spin rotates around its own axis, creating a magnetic field (due to the nonzero charge of nuclei). In a classical description, this magnetic field is represented by a vector called the magnetic moment. Normally, each nucleus' magnetic moment is randomly aligned, resulting in no net macroscopic magnetization. In the presence of an applied magnetic field \mathbf{B}_0 , however, each magnetic moment assumes one of two or more discrete energy states (the number of energy states depends on the type of nucleus, e.g., ^1H or ^{31}P nuclei) that dictate the orientation of the magnetic moment with respect to the external magnetic field [17]. Lower energy states are preferred, and thus the vector sum of a large number of magnetic moments results in a nonzero net magnetization vector aligned with the external magnetic field. This net magnetization vector is denoted by $\vec{\mathbf{M}} = [M_x, M_y, M_z]^\top \in \mathbb{R}^3$, and, since by convention the direction of the external magnetic field is along the z-axis, $\vec{\mathbf{M}} = \vec{\mathbf{M}}_0 \triangleq [0, 0, M_0]^\top$ in thermal equilibrium, where the equilibrium magnetization M_0 is proportional both to the number of nuclei whose magnetic moments make up $\vec{\mathbf{M}}$ and to the strength of the external magnetic field [17].

The net magnetization will stay in thermal equilibrium until acted upon, or *excited*, by another applied magnetic field, \mathbf{B}_1 . This magnetic field is applied perpendicularly to \mathbf{B}_0 (in the *transverse* plane) and must rotate at the *resonance* frequency of the nuclei (also called the *Larmor* frequency) to push $\vec{\mathbf{M}}$ out of thermal equilibrium. The Larmor frequency is given by $\omega_0 \triangleq 2\pi\gamma B_0$, where $B_0 \triangleq \|\mathbf{B}_0\|_2$ and γ is a nucleus-dependent physical constant called the gyromagnetic ratio. Because each type of nucleus has a different resonance frequency, \mathbf{B}_1 can excite one type of nucleus while leaving other nuclei undisturbed. Typically, MRI targets ^1H nuclei having a gyromagnetic ratio of $\gamma = 42.58 \text{ MHz/T}$, resulting in a resonance frequency of about 128 MHz for a typical $B_0 = 3 \text{ T}$ magnet. Because \mathbf{B}_1 is typically applied for a short time and oscillates in the radio-frequency range, \mathbf{B}_1 is frequently called an RF (excitation) pulse [17].

Following excitation, $\vec{\mathbf{M}}$ will experience *free precession* (rotation of the transverse component $M_{xy} \triangleq M_x + iM_y$, where $i \triangleq \sqrt{-1}$, at the Larmor frequency) and *relaxation* (decay of M_{xy} and recovery of the *longitudinal* component M_z) as it returns to thermal equilibrium. Precession causes time-varying magnetic flux that by Faraday's law can induce a voltage in a receiver coil tuned to the Larmor frequency [17]. Thus, the signal received in MRI measures M_{xy} . Specifically, the signal received from a voxel centered at \mathbf{r} is $s(\mathbf{r}, t) = \int_{\mathcal{V}(\mathbf{r})} M_{xy}(\mathbf{r}', t) d\mathbf{r}'$, where $\mathcal{V}(\mathbf{r})$ represents the spatial extent of the voxel.

An MRI experiment typically consists of alternating periods of excitation (to generate observable signal) and free precession and relaxation (during which the signal is recorded). The time between successive excitation pulses is called the *repetition time* (T_R), and the time from the center of the RF pulse to the center of signal recording is called the *echo time* (T_E).¹ This process of alternating between excitation and signal acquisition is repeated until enough data has been collected, and then the MRI image can be reconstructed.

¹The technically correct definition of T_E actually depends on the so-called *readout trajectory*, which is related to signal localization. This dissertation, however, does not focus on signal localization, so the given definition (which is correct for traditional *Cartesian readouts*) suffices.

2.1.1 Bloch Equation

The dynamics of the net magnetization $\vec{M}(\mathbf{r}, t)$ for nuclei at position \mathbf{r} are described by the Bloch equation²:

$$\frac{\partial}{\partial t} \vec{M}(\mathbf{r}, t) = \begin{bmatrix} -\frac{1}{T_2(\mathbf{r})} & \Delta\omega(\mathbf{r}) & -2\pi\gamma B_{1,y}(\mathbf{r}, t) \\ -\Delta\omega(\mathbf{r}) & -\frac{1}{T_2(\mathbf{r})} & 2\pi\gamma B_{1,x}(\mathbf{r}, t) \\ 2\pi\gamma B_{1,y}(\mathbf{r}, t) & -2\pi\gamma B_{1,x}(\mathbf{r}, t) & -\frac{1}{T_1(\mathbf{r})} \end{bmatrix} \vec{M}(\mathbf{r}, t) + \begin{bmatrix} 0 \\ 0 \\ \frac{M_0(\mathbf{r})}{T_1(\mathbf{r})} \end{bmatrix}. \quad (2.1)$$

$B_{1,x}$ and $B_{1,y}$ are the x and y components, respectively, of the spatially- and time-varying RF excitation pulse \mathbf{B}_1 ; T_1 is the spin-lattice time constant governing the rate of recovery of M_z to thermal equilibrium (M_0); T_2 is the spin-spin time constant governing the rate of decay of M_{xy} to 0; and $\Delta\omega$ is the off-resonance frequency, representing deviations in resonance frequency from the Larmor frequency due to spatial variations in the main magnetic field \mathbf{B}_0 and chemical shift (deviations in the magnetic field experienced by a nucleus due to its particular chemical environment). As written, (2.1) has no analytical solution; however, if one assumes the RF pulse is short relative to the effects of free precession and relaxation (a common assumption in MRI, where a typical RF pulse lasts about 1 ms, compared to T_2 values from about 20–100 ms and T_1 values from about 300–3000 ms) then one can solve the Bloch equation for the separate cases of excitation and free precession and relaxation.

Excitation

For excitation, i.e., neglecting the effects of free precession and relaxation (and assuming \mathbf{B}_1 is aligned with the x-axis³), the Bloch equation (2.1) becomes

$$\frac{\partial}{\partial t} \vec{M}(\mathbf{r}, t) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 2\pi\gamma B_{1,x}(\mathbf{r}, t) \\ 0 & -2\pi\gamma B_{1,x}(\mathbf{r}, t) & 0 \end{bmatrix} \vec{M}(\mathbf{r}, t). \quad (2.2)$$

²Technically, (2.1) presents the dynamics of \vec{M} in a rotating frame of reference that rotates at the Larmor frequency. Using the rotating frame simplifies notation, but otherwise is essentially equivalent. In the stationary frame, $\Delta\omega(\mathbf{r})$ is replaced by $\omega_0 + \Delta\omega(\mathbf{r})$. Additionally, if $B_{1,x}(\mathbf{r}, t)$ and $B_{1,y}(\mathbf{r}, t)$ are viewed as the real and imaginary parts, respectively, of $B_1(\mathbf{r}, t) \in \mathbb{C}$, then in the stationary frame they are replaced by the real and imaginary parts, respectively, of $B_1(\mathbf{r}, t)e^{-i\omega_0 t}$. Thus, only the rotational frequency of \mathbf{B}_1 is affected. These changes in the Bloch equation (2.1) do not affect the dynamics of M_z , and the only difference between M_{xy} in the rotating frame and the stationary frame is its rotational frequency (i.e., a factor of $e^{-i\omega_0 t}$).

³Excitation about an arbitrary axis in the transverse plane is possible by appropriately setting $B_{1,x}$ and $B_{1,y}$. This case can be viewed as first rotating the reference frame such that the arbitrary axis becomes the x-axis in the new frame of reference, applying the solution to (2.2) (since, intuitively, changing the reference frame does not change the dynamics of the system), and then rotating back to the original reference frame.

Its solution is simply rotation of $\vec{\mathbf{M}}$ about the x-axis by angle (called the *flip angle*) $\alpha(\mathbf{r}) \triangleq 2\pi\gamma \int_{t_0}^{t_0+\Delta t} B_{1,x}(\mathbf{r}, t) dt$, where t_0 is the time at which excitation begins and Δt is the duration of the excitation pulse (again, assumed to be small). (Often it is convenient to represent the flip angle as $\alpha(\mathbf{r}) = \kappa(\mathbf{r})\alpha$, where α is a constant prescribed or nominal flip angle and $\kappa(\mathbf{r})$ captures spatially varying imperfections in \mathbf{B}_1 .) Described as a matrix equation, the solution to (2.2) is

$$\vec{\mathbf{M}}(\mathbf{r}, t_0 + \Delta t) = \mathbf{R}_x(\mathbf{r})\vec{\mathbf{M}}(\mathbf{r}, t_0), \quad (2.3)$$

where

$$\mathbf{R}_x(\mathbf{r}) \triangleq \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha(\mathbf{r}) & \sin \alpha(\mathbf{r}) \\ 0 & -\sin \alpha(\mathbf{r}) & \cos \alpha(\mathbf{r}) \end{bmatrix}$$

describes rotation of the magnetization about the x-axis.

Free Precession and Relaxation

For free precession and relaxation, $\mathbf{B}_1 = \mathbf{0}_3$, so the Bloch equation (2.1) becomes

$$\frac{\partial}{\partial t} \vec{\mathbf{M}}(\mathbf{r}, t) = \begin{bmatrix} -\frac{1}{T_2(\mathbf{r})} & \Delta\omega(\mathbf{r}) & 0 \\ -\Delta\omega(\mathbf{r}) & -\frac{1}{T_2(\mathbf{r})} & 0 \\ 0 & 0 & -\frac{1}{T_1(\mathbf{r})} \end{bmatrix} \vec{\mathbf{M}}(\mathbf{r}, t) + \begin{bmatrix} 0 \\ 0 \\ \frac{M_0(\mathbf{r})}{T_1(\mathbf{r})} \end{bmatrix}. \quad (2.4)$$

Its solution for $t > t_0$ is

$$\vec{\mathbf{M}}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r}, t - t_0)\mathbf{F}(\mathbf{r}, t - t_0)\vec{\mathbf{M}}(\mathbf{r}, t_0) + (\mathbf{I}_3 - \mathbf{E}(\mathbf{r}, t - t_0)\mathbf{F}(\mathbf{r}, t - t_0))\vec{\mathbf{M}}_0(\mathbf{r}), \quad (2.5)$$

where $\vec{\mathbf{M}}_0(\mathbf{r}) \triangleq [0, 0, M_0(\mathbf{r})]^\top$ is the magnetization vector in thermal equilibrium,

$$\mathbf{E}(\mathbf{r}, t) \triangleq \begin{bmatrix} e^{-t/T_2(\mathbf{r})} & 0 & 0 \\ 0 & e^{-t/T_2(\mathbf{r})} & 0 \\ 0 & 0 & e^{-t/T_1(\mathbf{r})} \end{bmatrix}$$

describes relaxation of the magnetization, and

$$\mathbf{F}(\mathbf{r}, t) \triangleq \begin{bmatrix} \cos(\Delta\omega(\mathbf{r})t) & \sin(\Delta\omega(\mathbf{r})t) & 0 \\ -\sin(\Delta\omega(\mathbf{r})t) & \cos(\Delta\omega(\mathbf{r})t) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

describes precession of the transverse magnetization about the z-axis.

2.1.2 Bloch-McConnell Equation

The Bloch equation (2.1) uses a single-compartment model, assuming that the imaged object is homogeneous at position \mathbf{r} and thus can be described by a single T_1 and T_2 . However, a more realistic model is a multi-compartment model, where at position \mathbf{r} there exist multiple chemical environments, giving rise to a net magnetization vector that is a composite of the magnetization vectors originating from each of the chemical environments. Additionally, nuclei may move from one compartment to another through chemical exchange. The dynamics of the magnetization vectors for each compartment in a multi-compartment model with C compartments are described by the Bloch-McConnell [18] equation⁴:

$$\frac{\partial}{\partial t} \vec{\mathbf{M}}_{\text{MC}}(\mathbf{r}, t) = \begin{bmatrix} \mathbf{A}_{c_1}(\mathbf{r}, t) & \mathbf{E}_{c_2 \rightarrow c_1}(\mathbf{r}) & \cdots & \mathbf{E}_{c_C \rightarrow c_1}(\mathbf{r}) \\ \mathbf{E}_{c_1 \rightarrow c_2}(\mathbf{r}) & \mathbf{A}_{c_2}(\mathbf{r}, t) & \cdots & \mathbf{E}_{c_C \rightarrow c_2}(\mathbf{r}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{E}_{c_1 \rightarrow c_C}(\mathbf{r}) & \mathbf{E}_{c_2 \rightarrow c_C} & \cdots & \mathbf{A}_{c_C}(\mathbf{r}, t) \end{bmatrix} \vec{\mathbf{M}}_{\text{MC}}(\mathbf{r}, t) + \begin{bmatrix} \vec{\mathbf{M}}_{0,c_1}(\mathbf{r})/T_{1,c_1}(\mathbf{r}) \\ \vdots \\ \vec{\mathbf{M}}_{0,c_C}(\mathbf{r})/T_{1,c_C}(\mathbf{r}) \end{bmatrix}, \quad (2.6)$$

where

$$\mathbf{A}_{c_i}(\mathbf{r}, t) \triangleq \begin{bmatrix} -\frac{1}{T_{2,c_i}(\mathbf{r})} - \sum_{j \neq i} \frac{1}{\tau_{c_i \rightarrow c_j}(\mathbf{r})} & \Delta\omega_{c_i}(\mathbf{r}) & -2\pi\gamma B_{1,y}(\mathbf{r}, t) \\ -\Delta\omega_{c_i}(\mathbf{r}) & -\frac{1}{T_{2,c_i}(\mathbf{r})} - \sum_{j \neq i} \frac{1}{\tau_{c_i \rightarrow c_j}(\mathbf{r})} & 2\pi\gamma B_{1,x}(\mathbf{r}, t) \\ 2\pi\gamma B_{1,y}(\mathbf{r}, t) & -2\pi\gamma B_{1,x}(\mathbf{r}, t) & -\frac{1}{T_{1,c_i}(\mathbf{r})} - \sum_{j \neq i} \frac{1}{\tau_{c_i \rightarrow c_j}(\mathbf{r})} \end{bmatrix}$$

for $i = 1, \dots, C$,

$$\mathbf{E}_{c_i \rightarrow c_j}(\mathbf{r}) \triangleq \text{diag} \left(\left[\frac{1}{\tau_{c_i \rightarrow c_j}(\mathbf{r})}, \frac{1}{\tau_{c_i \rightarrow c_j}(\mathbf{r})}, \frac{1}{\tau_{c_i \rightarrow c_j}(\mathbf{r})} \right]^\top \right)$$

for $i, j = 1, \dots, C$, $\text{diag}(\cdot)$ creates a diagonal matrix with its vector input along the main diagonal, T_{1,c_i} is the T_1 for compartment c_i , T_{2,c_i} is the T_2 for compartment c_i , $\Delta\omega_{c_i}$ is the off-resonance frequency for compartment c_i , $\tau_{c_i \rightarrow c_j}$ is the residence time (inverse exchange rate) for exchange from compartment c_i to compartment c_j , $\vec{\mathbf{M}}_{0,c_i}(\mathbf{r}) \triangleq [0, 0, f_{c_i}(\mathbf{r})M_0(\mathbf{r})]^\top$ is the thermal equilibrium magnetization vector for compartment c_i , f_{c_i} is the fraction of nuclei at

⁴Like the Bloch equation (2.1), the Bloch-McConnell equation (2.6) is presented in the rotating frame of reference.

position \mathbf{r} that correspond to compartment c_i ($\sum_{i=1}^C f_{c_i} = 1$), and $\vec{\mathbf{M}}_{\text{MC}} \triangleq [\vec{\mathbf{M}}_{c_1}^\top \cdots \vec{\mathbf{M}}_{c_C}^\top]^\top$ collects the magnetization vectors for each compartment. The overall net magnetization vector is $\vec{\mathbf{M}} = \sum_{i=1}^C \vec{\mathbf{M}}_{c_i}$. Note that when the number of compartments is $C = 1$ the Bloch-McConnell equation (2.6) is the same as the Bloch equation (2.1). For analyzing the Bloch-McConnell equation, it is again convenient to split into two cases: one for excitation, and another for free precession, relaxation, and exchange.

Excitation

For excitation, i.e., neglecting the effects of free precession, relaxation, and exchange (and again assuming \mathbf{B}_1 is aligned with the x-axis), the Bloch-McConnell equation (2.6) becomes

$$\frac{\partial}{\partial t} \vec{\mathbf{M}}_{\text{MC}}(\mathbf{r}, t) = \left(\mathbf{I}_C \otimes \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 2\pi\gamma B_{1,x}(\mathbf{r}, t) \\ 0 & -2\pi\gamma B_{1,x}(\mathbf{r}, t) & 0 \end{bmatrix} \right) \vec{\mathbf{M}}_{\text{MC}}(\mathbf{r}, t), \quad (2.7)$$

where \otimes denotes the Kronecker product. In this case, the dynamics of each compartment are decoupled, and so the solution to (2.7) is the same as the Bloch equation solution (2.3) applied to each compartment individually, i.e.,

$$\begin{aligned} \vec{\mathbf{M}}_{\text{MC}}(\mathbf{r}, t_0 + \Delta t) &= (\mathbf{I}_C \otimes \mathbf{R}_x(\mathbf{r})) \vec{\mathbf{M}}_{\text{MC}}(\mathbf{r}, t_0) \\ &= \begin{bmatrix} \mathbf{R}_x(\mathbf{r}) \vec{\mathbf{M}}_{c_1} \\ \vdots \\ \mathbf{R}_x(\mathbf{r}) \vec{\mathbf{M}}_{c_C} \end{bmatrix}. \end{aligned} \quad (2.8)$$

Free Precession, Relaxation, and Exchange

In the absence of \mathbf{B}_1 , the Bloch-McConnell equation (2.6) becomes

$$\frac{\partial}{\partial t} \vec{\mathbf{M}}_{\text{MC}}(\mathbf{r}, t) = \mathbf{A}(\mathbf{r}) \vec{\mathbf{M}}_{\text{MC}}(\mathbf{r}, t) + \mathbf{c}(\mathbf{r}), \quad (2.9)$$

where

$$\mathbf{A}(\mathbf{r}) \triangleq \begin{bmatrix} \tilde{\mathbf{A}}_{c_1}(\mathbf{r}) & \mathbf{E}_{c_2 \rightarrow c_1}(\mathbf{r}) & \cdots & \mathbf{E}_{c_C \rightarrow c_1}(\mathbf{r}) \\ \mathbf{E}_{c_1 \rightarrow c_2}(\mathbf{r}) & \tilde{\mathbf{A}}_{c_2}(\mathbf{r}) & \cdots & \mathbf{E}_{c_C \rightarrow c_2}(\mathbf{r}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{E}_{c_1 \rightarrow c_C}(\mathbf{r}) & \mathbf{E}_{c_2 \rightarrow c_C} & \cdots & \tilde{\mathbf{A}}_{c_C}(\mathbf{r}) \end{bmatrix}, \quad \mathbf{c}(\mathbf{r}) \triangleq \begin{bmatrix} \vec{\mathbf{M}}_{0,c_1}(\mathbf{r})/T_{1,c_1}(\mathbf{r}) \\ \vdots \\ \vec{\mathbf{M}}_{0,c_C}(\mathbf{r})/T_{1,c_C}(\mathbf{r}) \end{bmatrix},$$

and

$$\tilde{\mathbf{A}}_{c_i}(\mathbf{r}) \triangleq \begin{bmatrix} -\frac{1}{T_{2,c_i}(\mathbf{r})} - \sum_{j \neq i} \frac{1}{\tau_{c_i \rightarrow c_j}(\mathbf{r})} & \Delta\omega_{c_i}(\mathbf{r}) & 0 \\ -\Delta\omega_{c_i}(\mathbf{r}) & -\frac{1}{T_{2,c_i}(\mathbf{r})} - \sum_{j \neq i} \frac{1}{\tau_{c_i \rightarrow c_j}(\mathbf{r})} & 0 \\ 0 & 0 & -\frac{1}{T_{1,c_i}(\mathbf{r})} - \sum_{j \neq i} \frac{1}{\tau_{c_i \rightarrow c_j}(\mathbf{r})} \end{bmatrix}$$

for $i = 1, \dots, C$. The solution to (2.9) for $t > t_0$ is

$$\vec{\mathbf{M}}_{\text{MC}}(\mathbf{r}, t) = e^{(t-t_0)\mathbf{A}(\mathbf{r})} \vec{\mathbf{M}}_{\text{MC}}(\mathbf{r}, t_0) + (\mathbf{I}_{3C} - e^{(t-t_0)\mathbf{A}(\mathbf{r})}) \vec{\mathbf{M}}_{0,\text{MC}}(\mathbf{r}), \quad (2.10)$$

where $\vec{\mathbf{M}}_{0,\text{MC}} \triangleq [\vec{\mathbf{M}}_{0,c_1}^\top \cdots \vec{\mathbf{M}}_{0,c_C}^\top]^\top$, and $e^{\mathbf{X}} \triangleq \sum_{k=0}^{\infty} \frac{1}{k!} \mathbf{X}^k$ is the matrix exponential function. In general, there is no simpler form that avoids the matrix exponential. However, if there is no exchange between a compartment and any of the others then that compartment's dynamics are decoupled from the other compartments' dynamics, and thus can be solved independently using the solution to the Bloch equation (2.5).

As stated previously, MRI measures transverse magnetization M_{xy} . For the multi-compartment case, the measured signal is a composite of the transverse magnetizations for each individual compartment, i.e., $M_{xy} = \sum_{i=1}^C M_{xy,c_i}$. In the absence of exchange, each compartment's dynamics evolve independently (and thus can ignore compartment fraction). Then the measured signal is a weighted sum of each compartment's transverse magnetization, i.e., $M_{xy} = \sum_{i=1}^C f_{c_i} M_{xy,c_i}$.

2.1.3 MRI Scans

The Bloch and Bloch-McConnell equations (2.1) and (2.6) can be used to model MRI scans. Essentially, an MRI scan is a way to manipulate magnetization vectors to obtain a desired signal (which of course depends on tissue properties such as T_1 and T_2). Different scans differentiate themselves by changing the dynamics magnetization vectors experience, accomplished by changing the T_R , T_E , flip angle(s), number of RF pulses per T_R , and phase of the RF pulse(s), among other things. The large number of user-controlled parameters results in a variety of unique MRI scans.

2.2 Small-Tip Fast Recovery (STFR)

One type of MRI scan is small-tip fast recovery (STFR) [8]. STFR includes two RF pulses per T_R , providing greater flexibility in controlling magnetization dynamics (and thus the

observed signal) than traditional MRI scans that include only one RF pulse.

One repetition of STFR begins with an initial tip-down excitation with flip angle α . Then there is time T_{free} during which free precession occurs, after which there is a tip-up excitation (“fast recovery”) where magnetization is rotated up towards the +z-axis with flip angle β and phase ϕ . Finally, there is gradient spoiling⁵ for time T_g . For STFR, $T_R = T_{\text{free}} + T_g$.

STFR is a steady-state scan, meaning the signal is acquired after many repetitions such that the net magnetization experiences the same dynamics from repetition to repetition. An expression for the steady-state magnetization can be obtained by setting up a recurrence relation. Note that the solutions to the Bloch and Bloch-McConnell equations (2.1) and (2.6) for the simplified cases of excitation and free precession, relaxation, and exchange all have the form $\vec{M}^+ = \mathbf{A}\vec{M}^- + \mathbf{c}$, where \vec{M}^- is the magnetization before the dynamics occur and \vec{M}^+ is the magnetization after (and \mathbf{c} is the zero vector in the case of excitation); thus \mathbf{A} and \mathbf{c} describe the dynamics that the magnetization experiences. Let \vec{M}_1 be the magnetization just after the tip-down pulse. This magnetization then experiences free precession, relaxation, and exchange for time T_{free} (with dynamics \mathbf{A}_{free} and \mathbf{c}_{free}), followed by the tip-up pulse (with dynamics $\mathbf{A}_{\text{tip-up}}$), gradient spoiling, and finally the tip-down pulse (with dynamics $\mathbf{A}_{\text{tip-down}}$). For simplicity, the gradient spoiling can be modeled as free precession, relaxation, and exchange for time T_g (with dynamics $\mathbf{A}_{\text{spoil}}$ and $\mathbf{c}_{\text{spoil}}$) followed by (assumed) ideal spoiling in which the x- and y-components of the magnetization (for each compartment) are set to 0 (modeled by multiplication with a matrix \mathbf{S}). Written mathematically,

$$\begin{aligned}\vec{M}_2 &= \mathbf{A}_{\text{free}}\vec{M}_1 + \mathbf{c}_{\text{free}} \\ \vec{M}_3 &= \mathbf{A}_{\text{tip-up}}\vec{M}_2 \\ \vec{M}_4 &= \mathbf{A}_{\text{spoil}}\vec{M}_3 + \mathbf{c}_{\text{spoil}} \\ \vec{M}_5 &= \mathbf{S}\vec{M}_4 \\ \vec{M}_6 &= \mathbf{A}_{\text{tip-down}}\vec{M}_5.\end{aligned}$$

In steady-state, $\vec{M}_{\text{ss}} \triangleq \vec{M}_1 = \vec{M}_6$. Thus,

$$\vec{M}_{\text{ss}} = (\mathbf{I}_{3C} - \mathbf{A}_{\text{tip-down}}\mathbf{S}\mathbf{A}_{\text{spoil}}\mathbf{A}_{\text{tip-up}}\mathbf{A}_{\text{free}})^{-1} (\mathbf{A}_{\text{tip-down}}\mathbf{S}(\mathbf{A}_{\text{spoil}}\mathbf{A}_{\text{tip-up}}\mathbf{c}_{\text{free}} + \mathbf{c}_{\text{spoil}})), \quad (2.11)$$

where \vec{M}_{ss} is the steady-state magnetization immediately following the tip-down pulse. The

⁵Gradient spoiling can be viewed as free precession with purposely induced off-resonance with the intent to zero out the signal within each voxel.

steady-state magnetization at the echo time is then given by

$$\vec{M}_{\text{ss,TE}} = \mathbf{A}_{\text{TE}} \vec{M}_{\text{ss}} + \mathbf{c}_{\text{TE}}, \quad (2.12)$$

where \mathbf{A}_{TE} and \mathbf{c}_{TE} describe the dynamics of free precession, relaxation, and exchange for time T_{E} .

For a single compartment, (2.12) has an analytical expression; the steady-state signal obtained at a given spatial location from an STFR scan at echo time $T_{\text{E}} \in (0, T_{\text{free}})$ (typically, $T_{\text{E}} = T_{\text{free}}/2$ for Cartesian scans) is given by [19]

$$s(M_0, T_1, T_2, \Delta\omega, \kappa, T_{\text{free}}, T_{\text{g}}, \alpha, \beta, \phi, T_{\text{E}}) = \frac{M_0 \sin(\kappa\alpha) \left[e^{-T_{\text{g}}/T_1} (1 - e^{-T_{\text{free}}/T_1}) \cos(\kappa\beta) + (1 - e^{-T_{\text{g}}/T_1}) \right] e^{-T_{\text{E}}/T_2} e^{-i\Delta\omega T_{\text{E}}}}{1 - e^{-T_{\text{g}}/T_1} e^{-T_{\text{free}}/T_2} \sin(\kappa\alpha) \sin(\kappa\beta) \cos(\Delta\omega \cdot T_{\text{free}} - \phi) - e^{-T_{\text{g}}/T_1} e^{-T_{\text{free}}/T_1} \cos(\kappa\alpha) \cos(\kappa\beta)}. \quad (2.13)$$

For multiple compartments with exchange, there is no analytical expression for the steady-state signal due to the matrix exponential in (2.10). For any compartment c_i that does not exchange with the other compartments, however, its steady-state signal can be computed using (2.13) (using $f_{c_i} M_0$ instead of M_0) and then added to the signal generated from the other compartments that do exchange with each other (computed using the Bloch-McConnell equation (2.6)).

The STFR signal depends on both known, user-controlled scan parameters and unknown tissue parameters. If multiple STFR scans are acquired with different scan parameters, then the unknown tissue parameters can be estimated using a variety of different estimation techniques. (Figure 3.10 shows examples of such STFR scans.)

2.3 Parameter Estimation via Regression with Kernels (PERK)

One estimation technique is parameter estimation via regression with kernels (PERK) [9], a supervised learning method for estimating unknown parameters \mathbf{x} given data observations \mathbf{y} and known parameters $\boldsymbol{\nu}$ (which are typically estimated separately and then treated as known values in PERK). In the context of MRI, $\mathbf{y} \in \mathbb{C}^D$ collects the acquired signal for a single voxel across D scans, and \mathbf{x} and $\boldsymbol{\nu}$ collect the unknown and known parameters (e.g., T_1 , $\Delta\omega$, etc.) for the corresponding voxel. Training data is obtained by simulating data observations \mathbf{y}_i (e.g., by using the Bloch-McConnell equation (2.6) or STFR signal model (2.13) for each

of the D scans) for various values of unknown and known parameters \mathbf{x}_i and $\boldsymbol{\nu}_i$; these T training data points are collected as $(\mathbf{q}_1, \mathbf{x}_1), \dots, (\mathbf{q}_T, \mathbf{x}_T)$, where $\mathbf{q}_i \triangleq [|\mathbf{y}_i|^\top \boldsymbol{\nu}_i^\top]^\top \in \mathbb{R}^Q$ and $|\cdot|$ denotes element-wise complex modulus. Each \mathbf{q}_i is a training feature vector with Q features, and the test feature vector \mathbf{q} is formed analogously for each voxel of acquired data. PERK computes estimates via regularized linear regression (ridge regression) after first transforming the feature vectors \mathbf{q} (for both training and testing) via some user-defined feature map (which is never directly used but is indirectly specified through the choice of kernel function). The PERK estimator is [9]

$$\hat{\mathbf{x}}(\mathbf{q}) = \frac{1}{T} \mathbf{X} \mathbf{1}_T + \mathbf{X} \mathbf{M} (\mathbf{M} \mathbf{K} \mathbf{M} + T \rho \mathbf{I}_T)^{-1} \mathbf{k}(\mathbf{q}), \quad (2.14)$$

where $\mathbf{X} \triangleq [\mathbf{x}_1 \cdots \mathbf{x}_T]$ denotes all of the unknown parameters in the training data, $\mathbf{1}_T \in \mathbb{R}^T$ is a vector of all ones, $\mathbf{M} \triangleq \mathbf{I}_T - \frac{1}{T} \mathbf{1}_T \mathbf{1}_T^\top$ is a de-meaning operator, ρ is a regularization parameter, the Gram matrix $\mathbf{K} \in \mathbb{R}^{T \times T}$ has entries $K_{i,j} \triangleq k(\mathbf{q}_i, \mathbf{q}_j)$, where $\mathbf{k}(\mathbf{q}) = [k(\mathbf{q}, \mathbf{q}_1), \dots, k(\mathbf{q}, \mathbf{q}_T)]^\top - \frac{1}{T} \mathbf{K} \mathbf{1}_T$, and $k(\mathbf{q}, \mathbf{q}')$ is the user-specified kernel function. A common choice of kernel function is the Gaussian kernel

$$k(\mathbf{q}, \mathbf{q}') = \exp\left(-\frac{1}{2} \|\boldsymbol{\Lambda}^{-1}(\mathbf{q} - \mathbf{q}')\|_2^2\right), \quad (2.15)$$

where $\boldsymbol{\Lambda}$ is a positive definite weighting matrix. PERK with a Gaussian kernel corresponds to first transforming the feature vectors \mathbf{q} via a nonlinear feature map into infinite-dimensional features, and then applying ridge regression on the transformed features. This lifting of features to a higher-dimensional space improves the ability to capture the nonlinear dependence of the signal on the unknown parameters to be estimated.

When the number of training examples T is large, (2.14) requires inverting a large matrix. To reduce computation, one can approximate (2.14) using random Fourier features [9, 20]. Specifically, features $\mathbf{q} \in \mathbb{R}^Q$ can be approximated with

$$\mathbf{z}(\mathbf{q}) = \sqrt{\frac{2}{H}} [z(\mathbf{v}_1, s_1; \mathbf{q}), \dots, z(\mathbf{v}_H, s_H; \mathbf{q})]^\top,$$

where H is the kernel approximation order. It is an approximation in the sense that, by the strong law of large numbers,

$$\lim_{H \rightarrow \infty} \mathbf{z}(\mathbf{q})^\top \mathbf{z}(\mathbf{q}') \xrightarrow{a.s.} \mathbb{E}_{\mathbf{v}, s} [z(\mathbf{v}, s; \mathbf{q}) z(\mathbf{v}, s; \mathbf{q}')] = k(\mathbf{q}, \mathbf{q}')$$

for translation-invariant kernels, an appropriate function z , and appropriate probability dis-

tributions for \mathbf{v} and s . For the (translation-invariant) Gaussian kernel (2.15),

$$z(\mathbf{v}, s; \mathbf{q}) = \cos(2\pi(\mathbf{v}^\top \mathbf{q} + s)),$$

$\mathbf{v}_1, \dots, \mathbf{v}_H$ are drawn independently from $p_{\mathbf{v}} \sim \mathcal{N}(\mathbf{0}_Q, (2\pi\mathbf{\Lambda})^{-2})$ (note that this $\mathbf{\Lambda}$ is the same as in (2.15)), and s_1, \dots, s_H are drawn independently from $p_s \sim \text{unif}(0, 1)$. With this approximation, the PERK estimator (2.14) becomes [9]

$$\hat{\mathbf{x}}(\mathbf{q}) = \frac{1}{T} \mathbf{X} \mathbf{1}_T + \frac{1}{T} \mathbf{X} \mathbf{M} \mathbf{Z}^\top \left(\frac{1}{T} \mathbf{Z} \mathbf{M} \mathbf{Z}^\top + \rho \mathbf{I}_H \right)^{-1} \left(\mathbf{z}(\mathbf{q}) - \frac{1}{T} \mathbf{Z} \mathbf{1}_T \right), \quad (2.16)$$

where $\mathbf{Z} \triangleq [z(\mathbf{q}_1) \cdots z(\mathbf{q}_T)]$ is such that $\mathbf{Z}^\top \mathbf{Z}$ is a low-rank approximation to \mathbf{K} .

2.4 Myelin Water Imaging (MWI)

One parameter that we wish to estimate is myelin water fraction (MWF). Myelin plays an important role in the transmission of electrical signals throughout the brain, and myelin loss is associated with decreased cognitive function. As a result, being able to measure myelin content is of interest. The structure of myelin is smaller than typical MRI voxel sizes, so direct imaging is not possible. However, one can account for multiple compartments within a voxel to create quantitative maps of myelin content. Specifically, myelin water imaging (MWI) involves creating an image of MWF, the fraction of MRI signal from a given voxel that comes from ^1H nuclei in myelin water (the water trapped between lipid bilayers within the myelin sheath, as illustrated in Figure 2.1). The differences in the tissue parameter values of myelin water and non-myelin water—in particular their very different T_2 times (~ 20 ms for myelin water versus ~ 80 ms for non-myelin water)—result in different signal characteristics, enabling separation of the two compartments.

2.4.1 Multi-Echo Spin Echo (MESE)

The standard method for myelin water imaging uses the multi-echo spin echo (MESE) MRI scan. A MESE scan samples the signal at multiple time points after excitation, giving information about the T_2 decay characteristics of the signal. This information is useful for separating myelin water from non-myelin water based on T_2 values.

One repetition of MESE [21] that uses a Carr-Purcell-Meiboom-Gill (CPMG) train consists of an initial excitation with flip angle $\alpha_{\text{ex}} = 90^\circ$ about the x-axis followed by a sequence of N_{ref} refocusing excitations with flip angle $\alpha_{\text{ref}} = 180^\circ$ about the y-axis. The signal is

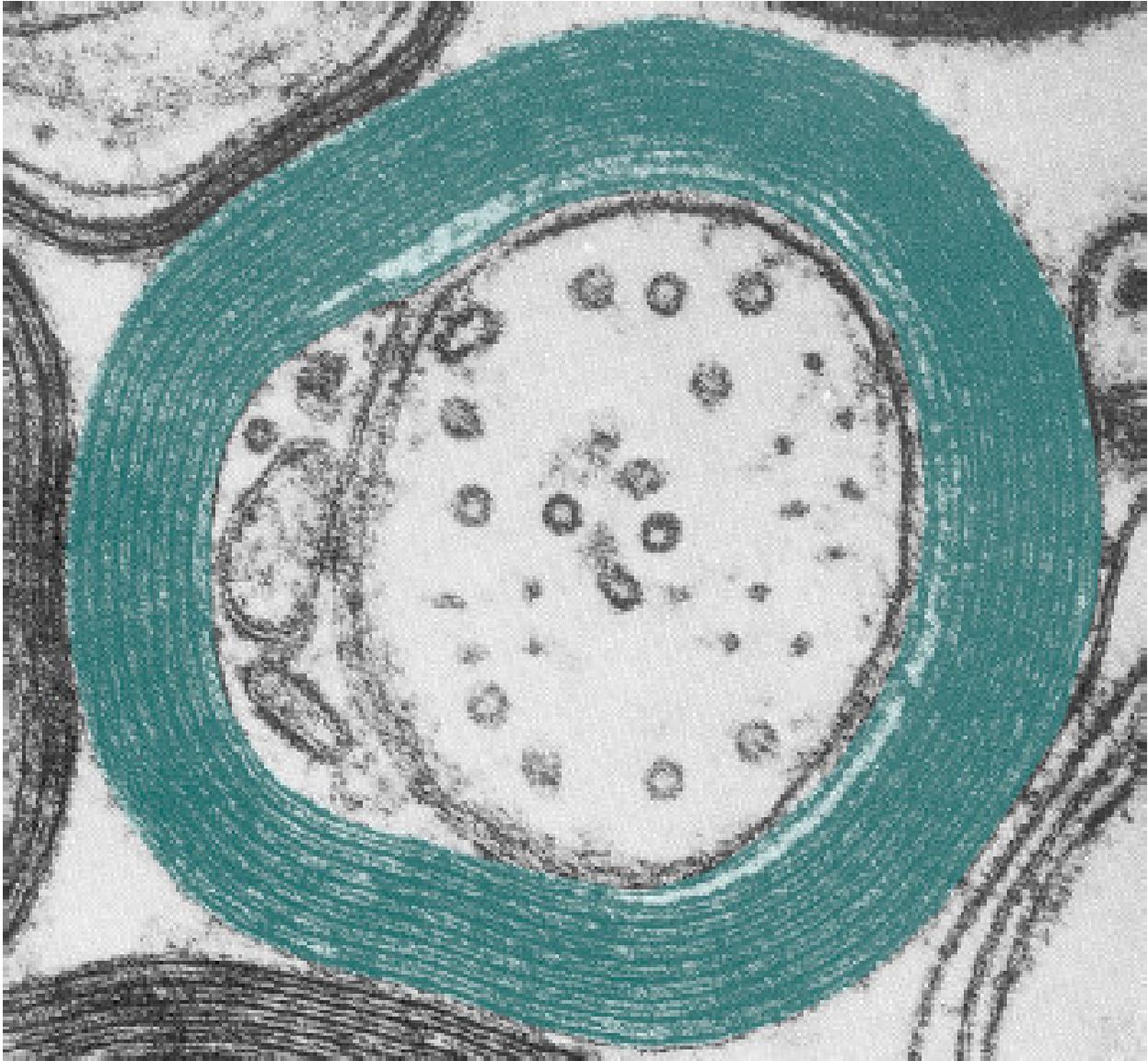


Figure 2.1: Image of myelin from an electron microscope. The myelin sheath (shaded in blue) consists of layers of myelin (myelin bilayers). Water trapped within the myelin bilayers is called myelin water, and all other water (both within the axon—the area enclosed by the myelin sheath—and without the axon) is called non-myelin water. (Original image from <http://www.teachmentortexts.com/2013/01/be-agent-of-change.html>, accessed September 18, 2020 at 9:57 AM EDT.)

sampled at times $T_E, 2T_E, \dots, N_{\text{ref}}T_E$ after the initial excitation, resulting in N_{ref} images in one MESE scan.

MWF is estimated from the MESE images using the nonnegative least squares (NNLS) algorithm, which solves

$$\hat{\mathbf{x}}(\mathbf{y}) = \arg \min_{\mathbf{x} \geq \mathbf{0}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \beta \|\mathbf{x}\|_2^2, \quad (2.17)$$

where $\mathbf{x} \geq \mathbf{0}$ constrains each element of \mathbf{x} to be nonnegative, and $\beta \geq 0$ is a regularization parameter. In the context of myelin water imaging with MESE, $\mathbf{y} \in \mathbb{R}^{N_{\text{ref}}}$ represents the acquired magnitude MESE signals (one after each refocusing excitation) for a given voxel, $\mathbf{A} \in \mathbb{R}^{N_{\text{ref}} \times C}$ is a dictionary where each column represents the N_{ref} simulated magnitude MESE signals for a particular choice of T_2 , and $\mathbf{x} \in \mathbb{R}^C$ represents how much of the different T_2 compartments are present. MWF is calculated by summing up the elements of $\hat{\mathbf{x}}$ that correspond to T_2 values within a certain range corresponding to myelin water (typically 10–40 ms) divided by $\sum_{c=1}^C \hat{x}_c$.

One benefit of using NNLS for myelin water imaging is that a large number of compartments C can be modeled at one time by specifying C different T_2 values when constructing the signal dictionary \mathbf{A} . On the other hand, this method ignores exchange between any of the compartments, as the signals arising from different compartments are just linearly combined to model the acquired signals \mathbf{y} . Another drawback is that the dictionary entries vary only with T_2 (i.e., other tissue parameters are held constant), though there are generalizations that overcome this issue (at the cost of increased computation) [22].

2.4.2 Tissue Models for MWI

When using NNLS, one assumes a certain tissue model, one wherein the tissue is characterized by C different compartments that differ only in T_2 values and that do not exchange with each other. There are other tissue models that can be utilized as well. This dissertation considers three other models: a two-compartment model with and without exchange and a three-compartment model with exchange. These models are described next and are summarized in Figure 2.2.

Two-Compartment Tissue Model with and without Exchange

The two-compartment tissue model without exchange consists of two compartments that represent myelin water and non-myelin water and that differ in T_1 , T_2 , and off-resonance frequency, as shown in Figure 2.2(a). The T_1 and T_2 values of the myelin water compartment

(a) Two-compartment model without exchange. (b) Three-compartment model with exchange.

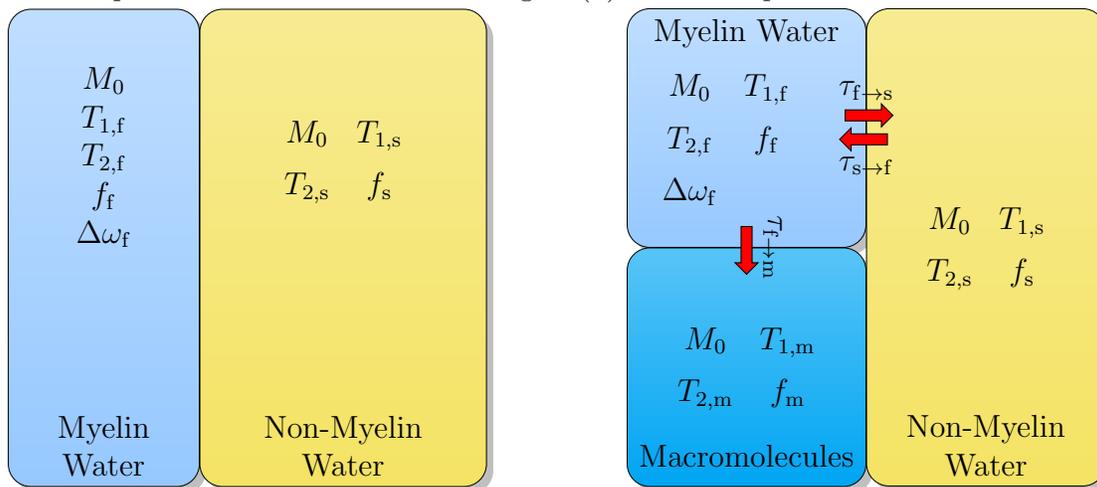


Figure 2.2: Tissue models for MWI. The two compartments of the two-compartment model (a) are myelin water and non-myelin water, each with its own T_1 , T_2 , compartment fraction, and (for myelin water) off-resonance frequency. Exchange between myelin water and non-myelin water can also be modeled (not shown). The three-compartment model (b) adds a macromolecule compartment, with its own T_1 , T_2 , and compartment fraction, as well as exchange from myelin water to macromolecules. The subscript “f” is for “fast” or “fast-relaxing” and is used for myelin water because it experiences faster relaxation than non-myelin water; conversely, non-myelin water uses the subscript “s” for “slow” or “slow-relaxing”. The subscript “m” stands for “macromolecule”.

are denoted as $T_{1,f}$ and $T_{2,f}$ (the subscript “f” refers to “fast” or “fast-relaxing” and arises because the T_1 and T_2 values of the myelin water compartment are smaller than those of the non-myelin water compartment, leading to faster relaxation rates). Similarly, the T_1 and T_2 values of the non-myelin water compartment are denoted as $T_{1,s}$ and $T_{2,s}$ (“s” for “slow” or “slow-relaxing”). The two compartments share the same M_0 , but differ in compartment fraction (f_f for myelin water, f_s for non-myelin water). In this sense, M_0 represents the total equilibrium magnetization in the voxel, while M_0f_f and M_0f_s represent the effective M_0 for each compartment individually. MWF is represented by f_f , and $f_f + f_s = 1$ since they represent fractions of the entire observed signal for one voxel. Off-resonance frequency $\Delta\omega$ (also called the bulk off-resonance frequency) is not tissue specific, and thus is not included in the tissue model; however, other sources of off-resonance effects can be tissue specific, leading to different off-resonance frequencies for tissues that experience such effects. This additional off-resonance frequency is denoted by $\Delta\omega_f$, and plays a role only in the myelin water compartment [1]. (To be clear, the total off-resonance frequency experienced by myelin water is $\Delta\omega + \Delta\omega_f$, and that experienced by non-myelin water is just $\Delta\omega$.) The myelin water off-resonance shift $\Delta\omega_f$ comes from magnetic susceptibility differences between myelin and the surrounding water and varies with the orientation of the myelinated axons with respect to \mathbf{B}_0 [23].

The two-compartment model can also incorporate exchange, characterized by the residence time or inverse rate of exchange. Specifically, magnetization exchanges from the myelin water compartment to the non-myelin water compartment (with residence time $\tau_{f \rightarrow s}$) and vice versa (with residence time $\tau_{s \rightarrow f}$). This model also assumes the two compartments are in chemical equilibrium, meaning $f_f\tau_{s \rightarrow f} = f_s\tau_{f \rightarrow s}$.

Three-Compartment Tissue Model with Exchange

The three-compartment tissue model with exchange consists of three compartments that represent myelin water, non-myelin water, and macromolecules and that differ in T_1 , T_2 , and off-resonance frequency, as shown in Figure 2.2(b). This model starts with the two-compartment exchanging model and adds another compartment for magnetization in macromolecules. The associated T_1 , T_2 , and compartment fraction are $T_{1,m}$, $T_{2,m}$, and f_m (“m” for “macromolecule”). As in the two-compartment model, MWF is represented by f_f , $f_f + f_s + f_m = 1$, and only the myelin water compartment experiences additional off-resonance. Unlike the two-compartment exchanging model, the three-compartment model incorporates exchange from the myelin water compartment to the macromolecule compartment (with residence time $\tau_{f \rightarrow m}$) [3].

In this dissertation, I used these different tissue models for MWI. In particular, in Chap-

ter 3 I used the three-compartment model with exchange. I chose to use this model because not only does it model myelin water and non-myelin water, but it also accounts for macromolecules that, despite having very short T_2 times, nevertheless influence the acquired signal through exchange [3]. Later, in Chapter 4, I opted to use the two-compartment model with exchange to have a model with fewer parameters and because exchange between water and macromolecules manifests as a change in water T_1 values (see Section 4.8). In both cases, however, I sought to use a model that incorporated relevant MR physics to enable accurate MWF estimates.

CHAPTER 3

Myelin Water Fraction Estimation Using Small-Tip Fast Recovery MRI

This chapter presents published journal paper [10] that extends conference abstracts [11] and [12]. Sections 3.1–3.6 are taken from the main published paper [10], and Section 3.7 is taken from the supporting information for the paper.

3.1 Introduction

Quantitative magnetic resonance imaging (QMRI) is the application of MRI to estimate parameters of interest. One QMRI application of growing interest is myelin water imaging, where one seeks quantitative maps of myelin water fraction (MWF) [2, 24]. The MWF is the proportion of MRI signal in a given voxel that originates from water bound within the myelin sheath. MWF maps are desirable for tracking progression of demyelinating diseases [24], e.g., multiple sclerosis [6].

The most widely accepted myelin water imaging techniques use the multi-echo spin echo (MESE) MRI scan (or variants) [6, 22, 25]. MESE is the standard for clinical MWF imaging to which alternative MWF mapping techniques are typically compared. However, MESE traditionally suffers from long scan times, impeding its routine clinical use. On the other hand, a combined gradient and spin echo (GRASE) MRI scan, a variant of MESE, has been shown to enable whole-brain MWF maps in under 8 minutes [26].

An alternative to MESE-based myelin water imaging is multicomponent driven equilibrium single pulse observation of T_1 and T_2 (mcDESPOT) that uses faster, steady-state MRI scans [27] that can acquire whole-brain MWF maps in 7 minutes [26]. Despite evidence showing that this method produces reproducible MWF maps (thus enabling longitudinal studies), there are concerns about overestimating the true MWF [28, 29] and its precision [30]. Other steady-state methods have also been explored for MWF estimation, such as multi-echo gradient echo (GRE) [31–33] and dual-echo steady-state (DESS). [34–36].

To our knowledge, most of these myelin water imaging techniques ignore potential differences in the effective magnetic field experienced by myelin-bound water compared to water outside of myelin (an exception being [33]). However, it has been shown that in cerebral white matter, myelin-bound water does in fact experience a different effective magnetic field [1].

In preliminary work [11], we showed that modeling the additional off-resonance experienced by myelin water reduces the Cramér-Rao Lower Bound (CRLB) of estimates of MWF using small-tip fast recovery (STFR) MRI [8]. We showed that the STFR sequence is sensitive to the frequency differences, suggesting that the difference in off-resonance between myelin and non-myelin water is a potentially useful contrast mechanism containing information that can help estimate MWF [11]. Simulations using optimized STFR scan parameters led to MWF estimates with lower errors when there was a fixed, non-zero (but unknown) difference in off-resonance, compared to when there was no (still unknown) frequency difference. To our knowledge, this work was the first to specifically design scans for myelin water imaging that exploit frequency differences. Because the actual frequency difference is unknown and might vary between voxels or disease conditions, the proposed approach treats the difference as an unknown parameter that is estimated alongside other unknown parameters like the T_1 and T_2 values of the various tissue compartments.

One limitation of our previous work was its tissue model. In [11] we assumed a two-compartment, non-exchanging model for simplicity in computing the STFR signal. However, neglecting exchange can lead to biases in MWF estimates [37]. Therefore, the method proposed in this chapter uses a three-compartment model. The three compartments are myelin water, non-myelin water, and a macromolecular pool; myelin water and non-myelin water are in exchange, while myelin water exchanges with the macromolecular pool [3].

We previously estimated MWF from optimized STFR scans using parameter estimation via regression with kernels (PERK), a recently developed learning-based technique for parameter estimation in MRI that uses kernel ridge regression at its core [9,11]. One alternative method for MWF estimation is nonlinear least squares, which requires iterative methods for solving and can get stuck in a local minimum. Another alternative is dictionary search, which requires evaluating the STFR signal model on a discretized grid of the signal model parameters, which quickly becomes impractical as the number of parameters increases due to the exponential scaling of the dictionary size and compute time. Yet another alternative is to use a neural network. While neural networks can lead to good parameter estimates, they require a lot of training data and long training time. In contrast, PERK trains quickly and avoids the other problems associated with nonlinear least squares and dictionary search. Therefore, this work again uses PERK.

This chapter substantially builds upon our previous work. First, we re-optimize the STFR scan parameters to model variations of bulk off-resonance and to account for two spoiled gradient-recalled echo (SPGR) scans that are used for separate bulk off-resonance estimation. Next, we compare STFR-based MWF estimates to MESE-based MWF estimates in simulation. In particular, we estimate MWF from the optimized STFR scans with PERK [9] using a three-compartment tissue model with exchange. Finally, we compare our proposed STFR-based MWF estimation method to MESE-based MWF estimation *in vivo*. Figure 3.1 illustrates the proposed approach.

The organization of this chapter is as follows. Section 3.2 provides background information on the scans used in this work (STFR and MESE), the scan design process, and PERK. Section 3.3 outlines our experiments, both for the STFR scan design and for MWF estimation in simulation and *in vivo*. Section 3.4 reports the experimental results. Section 3.5 discusses our results. Section 3.6 gives concluding remarks.

3.2 Background

3.2.1 STFR

One repetition of STFR [8] begins with an initial tip-down excitation with flip angle α . Then there is time T_{free} during which free precession occurs, after which there is a tip-up excitation (“fast recovery”) where magnetization is rotated up towards the +z-axis with flip angle β and phase ϕ . Finally there is gradient spoiling for time T_g . For a single compartment, the signal obtained at a given spatial location from a STFR scan is given by [19]

$$s_1(M_0, T_1, T_2, \Delta\omega, \kappa, T_{\text{free}}, T_g, \alpha, \beta, \phi) = \frac{M_0 \sin(\kappa\alpha) \left[e^{-T_g/T_1} (1 - e^{-T_{\text{free}}/T_1}) \cos(\kappa\beta) + (1 - e^{-T_g/T_1}) \right] e^{-T_{\text{free}}/(2T_2)} e^{-i\Delta\omega T_{\text{free}}/2}}{1 - e^{-T_g/T_1} e^{-T_{\text{free}}/T_2} \sin(\kappa\alpha) \sin(\kappa\beta) \cos(\Delta\omega \cdot T_{\text{free}} - \phi) - e^{-T_g/T_1} e^{-T_{\text{free}}/T_1} \cos(\kappa\alpha) \cos(\kappa\beta)}, \quad (3.1)$$

where M_0 is the equilibrium magnetization, T_1 and T_2 are the spin-lattice and spin-spin time constants, respectively, $\Delta\omega$ is the off-resonance frequency, and κ is a flip angle scaling constant (to account for differences between the prescribed and actual flip angles). Note that approximating flip angle error as a scale factor is accurate for the small flip angles used in this work, but typically inaccurate at larger flip angles. STFR with $\beta = 0$ is the same as SPGR with $T_R = T_{\text{free}} + T_g$.

For myelin water imaging, more than one compartment must be modeled. In a two-compartment model, one compartment consists of spins within myelin (myelin water), and

Scan Design

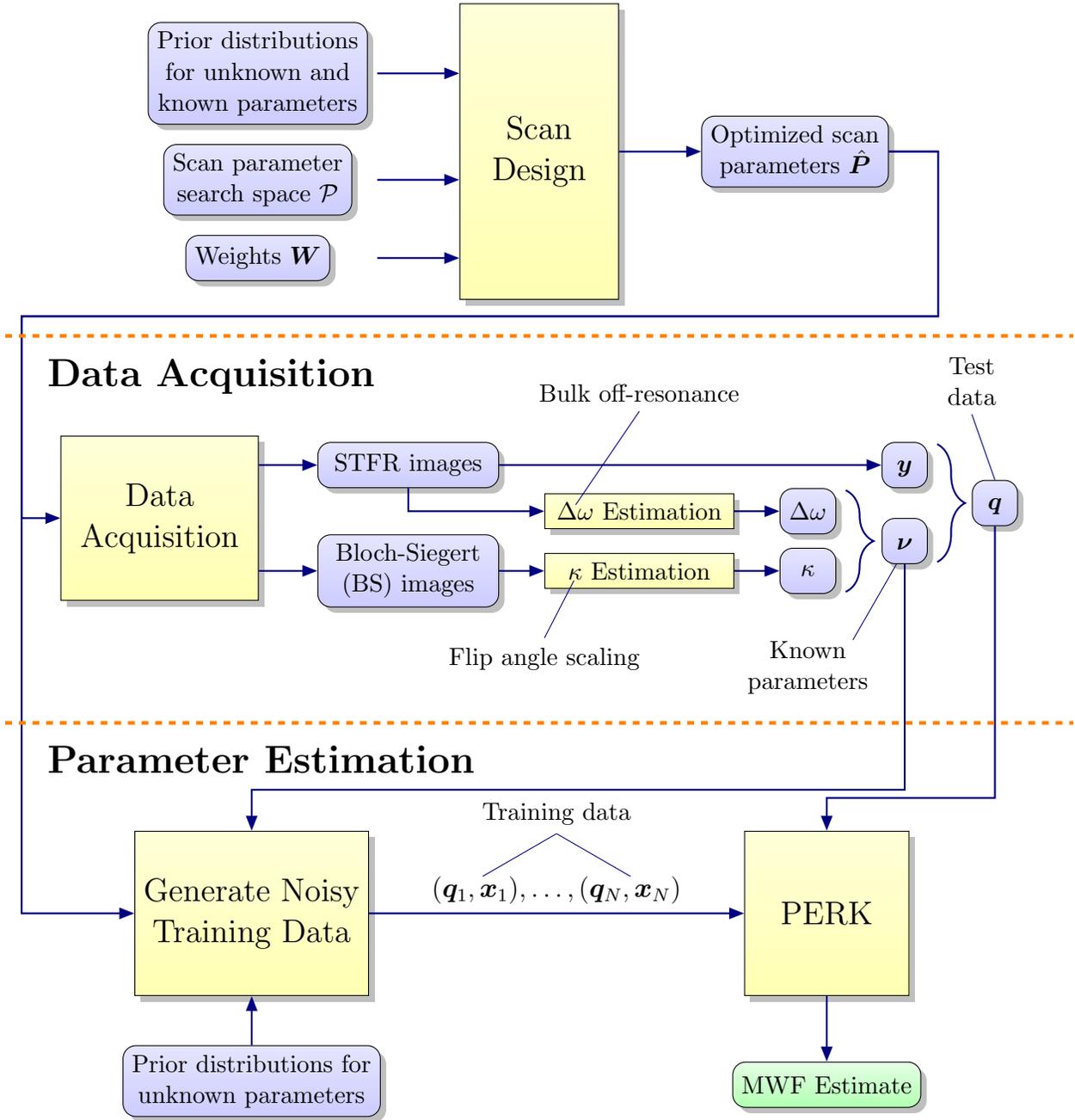


Figure 3.1: Workflow of the proposed methods. We first optimized a set of STFR scan parameters by minimizing a Cramér-Rao Lower Bound, then acquired data using those scans, as well as Bloch-Siebert (BS) scans. Two of the STFR scans were equivalent to SPGR scans, so were used to estimate $\Delta\omega$, and the BS scans were used to estimate κ . These parameters were treated as known values in the MWF estimation step. We then generated noisy training data using an STFR signal model. Finally, we passed the training data, acquired STFR images, and known parameters to PERK to estimate MWF voxel-by-voxel.

the other compartment consists of other spins (non-myelin water). If one neglects exchange, then the STFR signal at a given spatial location is the weighted sum of the single-compartment STFR signals of the individual compartments:

$$\begin{aligned}
s_2(M_0, f_f, T_{1,f}, T_{1,s}, T_{2,f}, T_{2,s}, \Delta\omega_f, \Delta\omega, \kappa, T_{\text{free}}, T_g, \alpha, \beta, \phi) = \\
f_f \cdot s_1(M_0, T_{1,f}, T_{2,f}, \Delta\omega + \Delta\omega_f, \kappa, T_{\text{free}}, T_g, \alpha, \beta, \phi) + \\
(1 - f_f) \cdot s_1(M_0, T_{1,s}, T_{2,s}, \Delta\omega, \kappa, T_{\text{free}}, T_g, \alpha, \beta, \phi), \tag{3.2}
\end{aligned}$$

where the weight f_f is the myelin water fraction (MWF), $T_{1,f}$ and $T_{2,f}$ are the T_1 and T_2 time constants for the fast-relaxing, myelin water compartment, $T_{1,s}$ and $T_{2,s}$ are the T_1 and T_2 time constants for the slow-relaxing, non-myelin water compartment, and $\Delta\omega_f$ is the additional off-resonance that is experienced only by myelin water [1].

Although (3.2) has a convenient analytical expression, a more accurate tissue model for cerebral white matter consists of three compartments (non-myelin water, myelin water, and a macromolecule water pool) with exchange between the non-myelin and myelin water compartments and from the myelin water compartment to the macromolecule compartment [3]. In this case, the STFR signal is also a function of the macromolecule compartment volume fraction f_m , the macromolecule compartment $T_{1,m}$ and $T_{2,m}$, the residence time for exchange from myelin water to non-myelin water $\tau_{f \rightarrow s}$, and the residence time for exchange from myelin water to the macromolecule compartment $\tau_{f \rightarrow m}$, in addition to the previously mentioned parameters. We assume the myelin water and non-myelin water compartments are in chemical equilibrium, which means that $f_f \tau_{s \rightarrow f} = (1 - f_f - f_m) \tau_{f \rightarrow s}$, and we assume there is no other exchange, i.e., $\tau_{m \rightarrow f} = \tau_{s \rightarrow m} = \tau_{m \rightarrow s} = \infty$ [3]. Because of exchange, the STFR signal no longer has an analytical expression and must be computed using the Bloch-McConnell equation [18].

3.2.2 MESE

One repetition of MESE [21] consists of an initial excitation with flip angle α_{ex} (typically 90°) followed by a sequence of N_{ref} refocusing excitations with flip angle α_{ref} (typically 180°). The signal is sampled at times $T_E, 2T_E, \dots, N_{\text{ref}}T_E$ after the initial excitation, resulting in N_{ref} images in one MESE scan. The repetition time T_R is typically chosen to be long enough so that the net magnetization of the spins is in equilibrium prior to each repetition. Thus, the MESE signal is a function of α_{ex} , α_{ref} , T_E , and T_R , as well as the same tissue parameters as the STFR signal; but if T_R is sufficiently long there is little dependence on T_1 (of any compartment). Additionally, for myelin water imaging using MESE, normally the acquired

MESE signal is modeled as a weighted sum of MESE signals from individual compartments, i.e., ignoring exchange between compartments. When ignoring exchange, we computed the MESE signal using the extended phase graph (EPG) method [38]. When accounting for exchange, we used Bloch-McConnell simulation.

3.2.3 Scan Design Using the Cramér-Rao Lower Bound

MR image data for a single voxel in a single scan is often modeled as

$$y = f(\mathbf{x}, \boldsymbol{\nu}, \mathbf{p}) + \epsilon, \quad (3.3)$$

where $f(\mathbf{x}, \boldsymbol{\nu}, \mathbf{p}) \in \mathbb{C}$ is the MR signal that is a function of unknown parameters \mathbf{x} , known parameters $\boldsymbol{\nu}$, and scan parameters \mathbf{p} ; and $\epsilon \sim \mathcal{CN}(0, \sigma^2)$ is additive complex Gaussian noise. When there are D scans then the data for a single voxel across each scan is collected into a vector:

$$\mathbf{y} = \mathbf{f}(\mathbf{x}, \boldsymbol{\nu}, \mathbf{P}) + \boldsymbol{\epsilon}, \quad (3.4)$$

where $\mathbf{y} \in \mathbb{C}^D$, $\mathbf{f}(\mathbf{x}, \boldsymbol{\nu}, \mathbf{P}) = [f_1(\mathbf{x}, \boldsymbol{\nu}, \mathbf{p}_1), \dots, f_D(\mathbf{x}, \boldsymbol{\nu}, \mathbf{p}_D)]^T$, f_d is the signal given by the d th scan for $d = 1, \dots, D$, $\mathbf{P} = (\mathbf{p}_1, \dots, \mathbf{p}_D)$ denotes the collection of all scan parameters, and the noise vector is $\boldsymbol{\epsilon} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma})$. We assume that each scan has noise independent of the other scans, and we assume that each scan has the same noise variance σ^2 ; thus $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_D$, where \mathbf{I}_D is the $D \times D$ identity matrix.

For simplicity in computing the Fisher information matrix (see below), we further assume that the MR signal model f is real-valued. We also take the magnitude of the received signal y , resulting in a Rician distributed signal [39]; however, we assume sufficiently high SNR so that this magnitude signal is approximately normally distributed with mean $f(\mathbf{x}, \boldsymbol{\nu}, \mathbf{p})$ and variance σ^2 .

Under these assumptions, the Fisher information matrix for the magnitude of signal model (3.4) is [40]

$$\mathbf{I}(\mathbf{x}, \boldsymbol{\nu}, \mathbf{P}) = \frac{1}{\sigma^2} (\nabla_{\mathbf{x}} \mathbf{f}(\mathbf{x}, \boldsymbol{\nu}, \mathbf{P}))^\top (\nabla_{\mathbf{x}} \mathbf{f}(\mathbf{x}, \boldsymbol{\nu}, \mathbf{P})), \quad (3.5)$$

where $\nabla_{\mathbf{x}}$ denotes a row gradient with respect to the unknown parameters \mathbf{x} . The inverse Fisher information matrix gives the Cramér-Rao Lower Bound (CRLB) for unbiased estimators [41]. In particular, the variance of an unbiased estimator for the i th unknown parameter x_i has a lower bound given by the i th diagonal element of the inverse Fisher information matrix, i.e., $\text{var}(\hat{x}_i) \geq [(\mathbf{I}(\mathbf{x}, \boldsymbol{\nu}, \mathbf{P}))^{-1}]_{i,i}$. This bound on the precision of unbiased estimators is useful for optimizing experimental designs. The CRLB has been used to optimize

MR sequence parameters for a variety of pulse sequences and applications, e.g., [9, 42–44]. In this work, we optimize scan parameters of a set of D STFR scans for MWF estimation by minimizing an expected weighted sum of the CRLB for each unknown parameter [45]:

$$\hat{\mathbf{P}} = \arg \min_{\mathbf{P} \in \mathcal{P}} \mathbb{E}_{\mathbf{x}, \boldsymbol{\nu}}[\text{trace}(\mathbf{W}(\mathbf{I}(\mathbf{x}, \boldsymbol{\nu}, \mathbf{P}))^{-1})], \quad (3.6)$$

where \mathcal{P} denotes the scan parameter search space, $\mathbb{E}_{\mathbf{x}, \boldsymbol{\nu}}$ denotes an expectation over \mathbf{x} and $\boldsymbol{\nu}$, and \mathbf{W} is a diagonal weighting matrix used to indicate the relative importance of precisely estimating the different unknown parameters.

3.2.4 Parameter Estimation via Regression with Kernels (PERK)

This section describes the PERK method we use to estimate MWF from STFR scans. Suppose a set of scan parameters \mathbf{P} is given, typically the $\hat{\mathbf{P}}$ from the scan design process (3.6). We seek to estimate unknown parameters \mathbf{x} after acquiring data using the D scans corresponding to these scan parameters. We generate training data by simulating data \mathbf{y}_n via (3.4) with appropriate signal models \mathbf{f} for various values of unknown and known parameters \mathbf{x}_n and $\boldsymbol{\nu}_n$; these N training data points are collected as $(\mathbf{q}_1, \mathbf{x}_1), \dots, (\mathbf{q}_N, \mathbf{x}_N)$, where $\mathbf{q}_n = [|\mathbf{y}_n|^T, \boldsymbol{\nu}_n^T]^T$ and $|\cdot|$ denotes element-wise complex modulus. After scanning (with the scan parameters \mathbf{P}), we have test data \mathbf{q} for each voxel (where $\boldsymbol{\nu}$ collects separately estimated parameters, such as B1+ maps, that are treated as known values), and we want to estimate \mathbf{x} . PERK computes estimates via regularized linear regression (ridge regression), after first transforming the feature vectors \mathbf{q} (for both training and testing) via some user-defined feature map (which is never directly used but is indirectly specified through the choice of kernel function). The PERK estimator is [9]

$$\hat{\mathbf{x}}(\mathbf{q}) = \frac{1}{N} \mathbf{X} \mathbf{1}_N + \mathbf{X} \mathbf{M} (\mathbf{M} \mathbf{K} \mathbf{M} + N \rho \mathbf{I}_N)^{-1} \mathbf{k}(\mathbf{q}), \quad (3.7)$$

where $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N]$ denotes all of the training data, $\mathbf{1}_N \in \mathbb{R}^N$ is a vector of all ones, $\mathbf{M} = \mathbf{I}_N - \frac{1}{N} \mathbf{1}_N \mathbf{1}_N^T$ is a de-meaning operator, ρ is a regularization parameter, the Gram matrix $\mathbf{K} \in \mathbb{R}^{N \times N}$ has entries $K_{i,j} = k(\mathbf{q}_i, \mathbf{q}_j)$, where $\mathbf{k}(\mathbf{q}) = [k(\mathbf{q}, \mathbf{q}_1), \dots, k(\mathbf{q}, \mathbf{q}_N)]^T - \frac{1}{N} \mathbf{K} \mathbf{1}_N$, and $k(\mathbf{q}, \mathbf{q}')$ is the user-specified kernel function. In this work, we used the Gaussian kernel

$$k(\mathbf{q}, \mathbf{q}') \triangleq \exp\left(-\frac{1}{2} \|\boldsymbol{\Lambda}^{-1}(\mathbf{q} - \mathbf{q}')\|_2^2\right), \quad (3.8)$$

where $\boldsymbol{\Lambda}$ is a positive definite weighting matrix. PERK with a Gaussian kernel corresponds to first transforming the feature vectors \mathbf{q} via a nonlinear feature map into infinite-dimensional

features, and then applying ridge regression on the transformed features. This lifting of features to a higher-dimensional space improves the ability to capture the nonlinear dependence of the signal on the unknown parameters we wish to estimate.

To reduce storage and computational needs, we approximated (3.7) using random Fourier features [9, 20].

3.3 Methods

This section describes the experiments performed in this work. We first explain the scan design process for optimizing a set of STFR scans for MWF estimation. We then explain simulated MWF estimation experiments that compare our proposed method to MESE-based MWF estimation. Finally, we explain an experiment to test our proposed method *in vivo*. The code for reproducing the methods and results in this chapter is available at <https://github.com/StevenWhitaker/STFR-MWF>. The raw data is available at <https://doi.org/10.7302/nw6e-1d66>.

3.3.1 Scan Design

For the STFR scan design, we computed the CRLB using the two-compartment non-exchanging signal model (3.2). (We used the non-exchanging model as opposed to the three-compartment exchanging model for computational reasons.) We chose the weighting matrix \mathbf{W} to place full weight on the CRLB for f_t (i.e., the diagonal entries of \mathbf{W} were all 0 except for a 1 in the location corresponding to f_t). We took the flip angle scaling κ and bulk off-resonance $\Delta\omega$ to be known, i.e., part of $\boldsymbol{\nu}$, and we optimized a set of $D = 11$ STFR scans. Two of these scans were SPGR scans with fixed scan parameters and an echo time shift. We included these scans to enable the option of estimating $\Delta\omega$ using conventional techniques and then treat $\Delta\omega$ as known for further parameter estimation.

We fixed $T_g = 2.8$ ms across all 11 scans. We fixed $\alpha = 5^\circ$ and $T_{\text{free}} = 10.3$ ms for the two SPGR scans. We chose a small flip angle to limit contrast between white matter and gray matter in the SPGR scans (and thus in the estimated field map). The T_E of each STFR scan and the first SPGR scan was 4 ms. The echo time shift between the two SPGR scans was 2.3 ms, chosen such that fat would be out of phase between the two SPGR scans at 3 T. For the remaining nine STFR scans, we fixed $T_{\text{free}} = 8$ ms and we constrained $\alpha \in [1, 15]^\circ$, $\beta \in [0, 15]^\circ$, and $\phi \in [-180, 180]^\circ$. We chose the fixed timing parameters to be small but still physically realizable, and we chose the flip angle constraints to ensure the resulting scans would be well below specific absorption rate (SAR) constraints.

The expectation in (3.6) requires choices for the probability distributions of the unknown and known parameters. Table 3.1 shows the distributions we used. To explore the effect that the additional myelin water off-resonance $\Delta\omega_f$ has on the CRLB of f_f , we performed one scan design (design A) where we took $\Delta\omega_f$ to be unknown, and another (design B) where we ignored $\Delta\omega_f$ (i.e., we assumed it was known and equal to 0). To solve the optimization in (3.6), we used the NLOpt package [46–49] in the Julia programming language [50].

3.3.2 MWF Estimation

For MWF estimation, we compared several estimation protocols. The proposed method, which we call STFR3-PERK, uses PERK to estimate MWF from the optimized STFR scans, with training data generated using the three-compartment exchanging model. This method assumes bulk off-resonance $\Delta\omega$ and flip angle scaling κ are known (unless otherwise noted). Another method, STFR2-PERK, is the same as STFR3-PERK, except training data is generated using the two-compartment non-exchanging model. Again, $\Delta\omega$ and κ are assumed known. The reference method, MESE-NNLS, uses regularized NNLS to estimate MWF from a MESE scan. Following [22], we fit 40 different T_2 components spaced logarithmically from 15 ms to 2000 ms, and computed MWF as the proportion of signal coming from components with $T_2 \leq 40$ ms to the total signal. This method does not assume knowledge of $\Delta\omega$ or κ , but jointly estimates κ . A fourth method, MESE-PERK, estimates MWF from a MESE scan using PERK, with training data generated using the three-compartment exchanging model. This method was included to determine whether performance differences were due to the estimation method (i.e., NNLS versus PERK), or due to the scans (i.e., MESE versus STFR). MESE-PERK does not assume knowledge of $\Delta\omega$ or κ . Finally, because the proposed STFR3-PERK assumes $\Delta\omega$ and κ are known, whereas the reference MESE-NNLS does not, we compared a fifth method, STFR3-PERK-JE, that is the same as STFR3-PERK except $\Delta\omega$ and κ are assumed unknown. Table 3.1 shows the training ranges for the methods that use PERK.

The methods that use PERK require specifying the regularization parameter ρ and the positive definite matrix $\mathbf{\Lambda}$ in the Gaussian kernel. For the Gaussian kernel, to eliminate dependence on scale we set $\mathbf{\Lambda} = \lambda \text{diag}(\mathbf{m}_{|q|})$, where λ is a regularization parameter and $\mathbf{m}_{|q|}$ denotes the sample average across all voxels of the magnitude test data $|q|$, where \mathbf{q} collects the magnitude STFR signals $|\mathbf{y}|$ and the known parameters $\boldsymbol{\nu}$ (see Section 3.2.4). We chose $\rho = 2^{-60}$ and $\lambda = 2^{3.5}$ for the regularization parameters, which we tuned using a holdout process described in [9, Section S.II].

Table 3.1: Unknown and known parameters used in scan design and in simulation. Values were chosen to match literature values for white matter [1–4]. The columns “Design A” and “Design B” contain the tissue distributions used for computing the expectation in the CRLB scan design cost function (3.6). We used $M_0 = 1$ for the scan designs because it only scales the STFR signal. The columns “WM” and “GM” specify the tissue parameters used for simulating white matter and gray matter tissues, respectively. For such simulations, we hold the tissue parameters fixed across different voxels, but vary $\Delta\omega$ and κ . The last column contains the tissue parameter distributions from which we sampled sets of tissue parameters for training PERK. The line below κ separates parameters used in both the two-compartment and three-compartment models (above) from those used only in the three-compartment exchanging model (below). WM: white matter. GM: gray matter.

Parameter	Design A	Design B	WM	GM	PERK Training Ranges
M_0	1	1	0.77	0.86	$\text{unif}(0, M_0^{\max})^a$
f_f	$\text{unif}(0.03, 0.31)$	$\text{unif}(0.03, 0.31)$	0.15	0.03	$\text{unif}(0.03, 0.31)$
$T_{1,f}$ (ms)	$\mathcal{N}(400, 80^2)$	$\mathcal{N}(400, 80^2)$	400	500	$\text{unif}(320, 480)$
$T_{1,s}$ (ms)	$\mathcal{N}(1000, 200^2)$	$\mathcal{N}(1000, 200^2)$	832	1331	$\text{unif}(800, 1200)$
$T_{2,f}$ (ms)	$\mathcal{N}(20, 4^2)$	$\mathcal{N}(20, 4^2)$	20	20	$\text{unif}(16, 24)$
$T_{2,s}$ (ms)	$\mathcal{N}(80, 16^2)$	$\mathcal{N}(80, 16^2)$	80	80	$\text{unif}(64, 96)$
$\Delta\omega_f$ (Hz)	$\text{unif}(5, 35)$	0*	15	5	$\text{unif}(0, 35)$
$\Delta\omega$ (Hz)	$\text{unif}(-50, 50)^*$	$\text{unif}(-50, 50)^*$	Varies	Varies	$\text{unif}(-50, 50)^b$
κ	$\text{unif}(0.8, 1.2)^*$	$\text{unif}(0.8, 1.2)^*$	Varies	Varies	$\text{unif}(0.8, 1.2)^b$
f_m	N/A	N/A	0.1	0.03	$\text{unif}(0.03, 0.31)$
$T_{1,m}$ (ms)	N/A	N/A	1000	1000	$\text{unif}(800, 3000)$
$T_{2,m}$ (ms)	N/A	N/A	0.02	0.02	$\text{unif}(0.01, 0.1)$
$\tau_{f \rightarrow s}$ (ms)	N/A	N/A	100	20	$\text{unif}(80, 150)$
$\tau_{f \rightarrow m}$ (ms)	N/A	N/A	50	10	$\text{unif}(40, 75)$

N/A – Not applicable (scan designs only used the two-compartment model).

$\text{unif}(a, b)$ – Uniform distribution on the interval $[a, b]$.

$\mathcal{N}(\mu, \sigma^2)$ – Normal distribution with mean μ and variance σ^2 .

*Known parameter.

^a M_0^{\max} given by the maximum signal value from data divided by the mean signal value from the signal model with $M_0 = 1$.

^bUnless the parameter is known at test time, in which case the training range covers the range of values in the known map.

Numerical Simulation

We compared the two optimized sets of STFR scans to validate the scan design process. We simulated test data using the two-compartment non-exchanging STFR signal model (3.2) using the range of tissue parameters over which the scans were optimized (to match the scan design assumptions), and we estimated MWF using STFR2-PERK. We measured the root mean squared error (RMSE) of the MWF estimates versus the additional myelin water off-resonance $\Delta\omega_f$ for three cases: first, using design B and training data that ignored $\Delta\omega_f$ (i.e., all training points had $\Delta\omega_f = 0$); second, using design B and training data that accounted for $\Delta\omega_f$; and third, using design A and training data that accounted for $\Delta\omega_f$. Section 3.7.1 describes another experiment where test data was generated using the two-compartment model with fixed white matter and gray matter tissue values (see Table 3.1).

Next, we investigated the effects of exchange and compared STFR-based MWF estimates to MESE-based estimates. We simulated STFR scans using design A and a MESE scan using $\alpha_{\text{ex}} = 90^\circ$, $\alpha_{\text{ref}} = 180^\circ$, $N_{\text{ref}} = 32$, $T_E = 10$ ms, and $T_R = 1200$ ms. We simulated test data using the three-compartment model with exchange and tissue parameters corresponding to white matter and gray matter (see Table 3.1). Additionally, we chose bulk off-resonance $\Delta\omega$ values to vary from -30 to 30 Hz and κ values to vary from 0.8 to 1.2 . We compared MWF estimates from each of the aforementioned methods (STFR3-PERK, STFR2-PERK, MESE-NNLS, MESE-PERK, and STFR3-PERK-JE). For these simulations we added complex Gaussian noise corresponding to a SNR in white matter ranging from 7–28 across the STFR scans and from 2–122 across the 32 MESE echoes (to match the SNR of the *in vivo* data), where SNR was calculated by dividing the white matter signal mean by the noise standard deviation.

The proposed method (STFR3-PERK) uses a model that matches the model used to generate the test data in the previous experiment. To investigate the effects of model mismatch, we repeated the previous experiment using a nine-compartment tissue model with exchange for the test data. The nine compartments were created by splitting each of the three compartments in the three-compartment model into three sub-compartments. For example, the myelin water compartment with fraction f_f and relaxation time $T_{2,f}$ was split into three compartments with fractions $0.5f_f$, $0.25f_f$, and $0.25f_f$ and relaxation times $T_{2,f}$, $0.8T_{2,f}$, and $1.2T_{2,f}$. Section 3.7.2 repeats this experiment for a four-compartment model with exchange and a three-compartment model without exchange.

In Vivo Experiments

Under an IRB-approved protocol, we scanned a healthy volunteer to compare the proposed STFR-based MWF estimation to MESE-based MWF estimation. We used 3D acquisitions for both the STFR and MESE scans to avoid slice profile effects. The STFR scans used design A, and the RF pulses had time-bandwidth product of 8 and duration of 1 ms; the two SPGR scans took 58 s and the nine STFR scans took 3 min 36 s for a total scan time of 4 min 34 s. We also acquired a pair of Bloch-Siegert (BS) scans for separate estimation of κ [51]; the excitation RF pulse of these scans had time-bandwidth product of 8 and duration of 1 ms, and used ± 4 kHz off-resonant Fermi pulses between excitation and readout. The total duration of the BS scans was 2 min 40 s. Therefore, our overall STFR-based MWF estimation scan protocol lasted 7 min 14 s. For the MESE scan, we used the same scan parameters as in simulation (described above); the initial excitation RF pulse had time-bandwidth product of 6, duration of 3 ms, and slab thickness of 0.9 cm, and each refocusing pulse had time-bandwidth product of 2, duration of 2 ms, and slab thickness of 2.1 cm. Each refocusing pulse was also flanked with crusher gradients, each of which imparted 14 cycles of phase across the imaging volume. The total duration of the MESE scan was 36 min 11 s. For all scans, we acquired a $22 \times 22 \times 0.99$ cm³ field of view (FOV) with matrix size $200 \times 200 \times 9$ (except the BS scans used matrix size $200 \times 50 \times 9$). We implemented the protocol in TOPPE [52].

We used a GE DiscoveryTM MR750 3.0T scanner with a 32-channel Nova Medical[®] head coil. We used conventional inverse FFT reconstruction followed by square-root of sum-of-squares coil combination to make the magnitude images used for MWF estimation. We estimated the SNR in the white matter brain regions (pooling the four white matter regions of interest (ROIs) in Table 3.4 for each scan/echo) to vary from 8–17 across the STFR scans and from 6–73 for across MESE echoes.

We analyzed the center slice of the acquired data. We estimated MWF using STFR3-PERK, STFR2-PERK, MESE-NNLS, and MESE-PERK. In this case, for STFR3-PERK and STFR2-PERK we took bulk off-resonance $\Delta\omega$ to be unknown (but still assumed κ to be known).

3.4 Results

3.4.1 Scan Design

Table 3.2 reports the two optimized scan design parameters. For design A, the additional myelin water off-resonance $\Delta\omega_f$ was taken to be unknown and distributed uniformly from 5

Table 3.2: Optimized STFR scan parameters. All values have units of degrees. For design A, the additional myelin water off-resonance $\Delta\omega_f$ was taken to be unknown and distributed uniformly from 5 to 35 Hz. For design B, $\Delta\omega_f$ was ignored (i.e., taken to be known and equal to 0). Both design A and design B also include two fixed SPGR scans with $\alpha = 5^\circ$.

	Scan #	1	2	3	4	5	6	7	8	9
Design A	α	15.0	15.0	15.0	15.0	15.0	15.0	15.0	11.4	15.0
	β	15.0	15.0	11.6	15.0	13.3	15.0	14.9	0.3	14.4
	ϕ	-139.3	-108.1	-66.0	-28.0	25.9	64.4	104.1	146.3	173.0
Design B	α	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0	15.0
	β	15.0	14.5	14.9	14.8	14.8	14.9	0.0	14.5	15.0
	ϕ	-139.3	-113.3	-63.7	-14.3	14.3	63.7	83.2	113.3	139.3

to 35 Hz. For design B, $\Delta\omega_f$ was ignored (i.e., taken to be known and equal to 0).

Figure 3.2 compares the expected CRLB of the standard deviation of MWF of these two scan designs versus $\Delta\omega_f$, where at each data point $\Delta\omega_f$ is fixed (unlike the other parameters that vary according to the distributions in Table 3.1) but still unknown, i.e., is contained in \mathbf{x} (see Section 3.2.3). For these CRLB calculations, we used a noise standard deviation that corresponds to SNR ranging from 9–15 in white matter across the STFR scans to match the SNR of the 1.1 mm isotropic *in vivo* data. Figure 3.2 shows that modeling $\Delta\omega_f$ improves the precision of the optimized scan design, and that MWF becomes easier to estimate as $\Delta\omega_f$ increases.

3.4.2 MWF Estimation

Numerical Simulation

We computed the RMSE of MWF estimates for test data generated using (3.2) with different values of $\Delta\omega_f$ and a range of tissue parameters. For design A, we estimated MWF using training data that was generated with a range of $\Delta\omega_f$ values. For design B, in one experiment we estimated MWF using training data that was generated with a range of $\Delta\omega_f$, and in another experiment the training data included only $\Delta\omega_f = 0$. Figure 3.3 shows the results. The mean MWF value in the test data was 0.17, so the minimum RMSE of 0.045 corresponds to about 26% relative error. Figure 3.7 reports an analogous experiment using fixed white matter and gray matter tissue values.

Furthermore, we investigated the effects of exchange on MWF estimates. We simulated STFR scans (using design A) and a MESE scan, and we used the three-compartment tissue model with exchange using tissue parameters corresponding to white matter and gray matter.

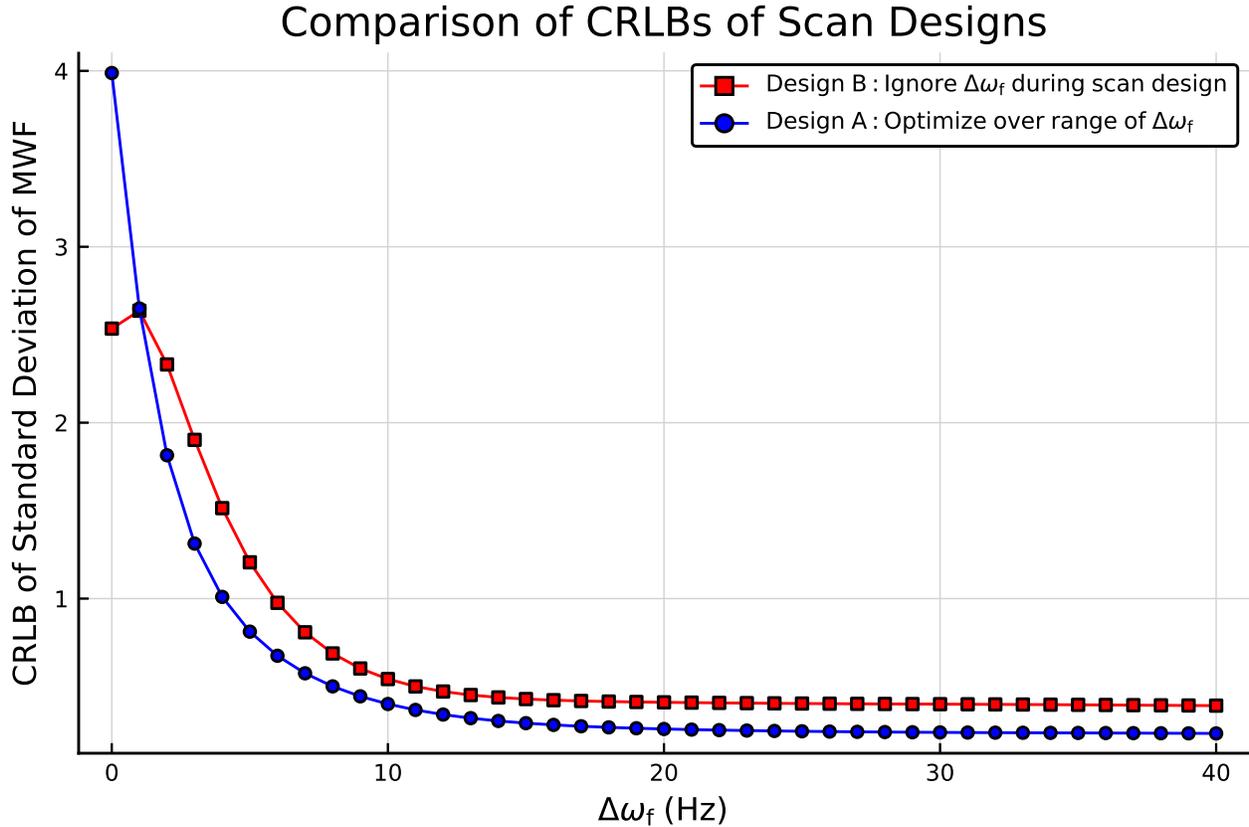


Figure 3.2: The two optimized scans (see Table 3.2) were evaluated to explore how including the additional myelin water off-resonance $\Delta\omega_f$ in the design process affects performance. Design A (where $\Delta\omega_f$ was included in the optimization) has a better expected CRLB for every value of $\Delta\omega_f$ within the 5–35 Hz range over which design A was optimized. Design B (where $\Delta\omega_f$ was ignored) understandably has a better expected CRLB for $\Delta\omega_f = 0$. For each value of $\Delta\omega_f$ investigated in this plot, the expected CRLB was computed with that value of $\Delta\omega_f$ held constant but unknown. Typical values of $\Delta\omega_f$ in white matter are 5–35 Hz [1]. Although these CRLB values predict estimator standard deviation (for an unbiased estimator) on the order of 100% MWF or more, we are not restricted by these large values because we are using a Bayesian estimator (see Figure 3.11 for details).

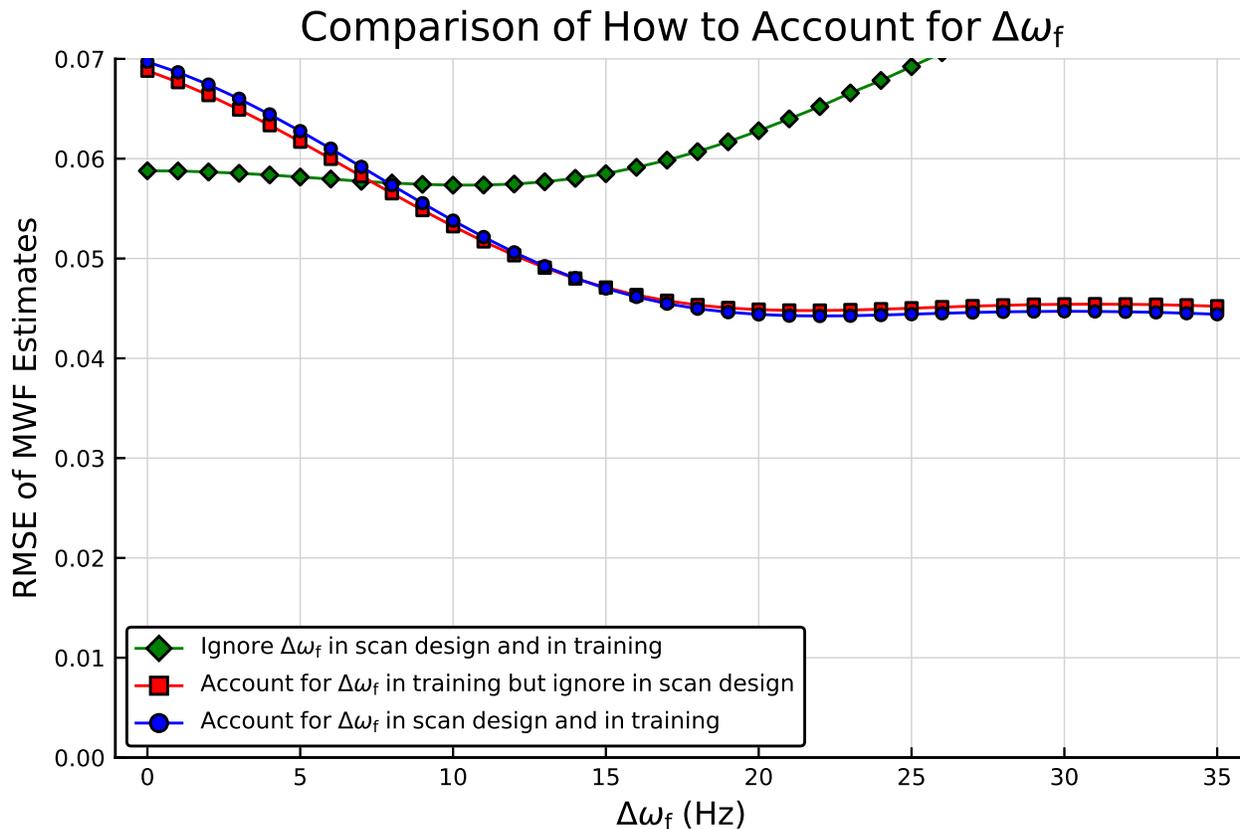


Figure 3.3: RMSE of MWF estimates from simulated test data for various ways of accounting for the additional myelin water off-resonance $\Delta\omega_f$. The diamond-marked green curve was generated using scan design B (where $\Delta\omega_f$ was ignored) and using PERK training data where $\Delta\omega_f = 0$. In other words, $\Delta\omega_f$ was not considered in any aspect, neither in the scan design nor when training. The square-marked red curve was also generated using scan design B, but the training data included a range of $\Delta\omega_f$ values. The circle-marked blue curve was generated using design A and a range of values of $\Delta\omega_f$. The latter two methods look almost identical, but both have much better RMSE than the first method as $\Delta\omega_f$ increases.

Table 3.3: Comparison of various methods of MWF estimation. The reported time refers to the entire estimation, combining the time to estimate MWF in white matter voxels and gray matter voxels; it also includes training time for the methods that use PERK. The best value in each column is highlighted. See Figure 3.4 for a visual comparison of these methods.

	White Matter (MWF = 0.15)			Gray Matter (MWF = 0.03)			Time (s)
	RMSE	Mean	SD	RMSE	Mean	SD	
STFR2-PERK	0.215	0.349	0.082	0.185	0.209	0.047	21.9
STFR3-PERK	0.021	0.158	0.020	0.046	0.074	0.015	43.1
STFR3-PERK-JE	0.026	0.145	0.026	0.044	0.069	0.021	41.3
MESE-NNLS	0.063	0.092	0.025	0.029	0.001	0.003	1602.4
MESE-PERK	0.029	0.134	0.025	0.026	0.026	0.025	142.7

We estimated MWF using STFR2-PERK, STFR3-PERK, MESE-NNLS, MESE-PERK, and STFR3-PERK-JE. Table 3.3 shows the RMSEs, means, and standard deviations of the MWF estimates. Figure 3.4 shows the ground truth map and a visual comparison of the estimated MWF maps. Figure 3.5 shows the results of this experiment when using a nine-compartment exchanging model. Figures 3.8 and 3.9 show results when using a four-compartment exchanging model and a three-compartment non-exchanging model, respectively. The anatomy for the simulated data used in these experiments came from BrainWeb [53].

In Vivo Experiments

We scanned a healthy volunteer using scan design A. Figure 3.10 shows images of the two SPGR and nine STFR scans of the subject. In the same scan session, we also scanned the volunteer with a MESE scan. Figure 3.6 shows MWF maps that were computed from the STFR and MESE scans. In this case, we made the STFR-based MWF estimates without using a separately estimated $\Delta\omega$ map because the MWF estimates made with the separately estimated $\Delta\omega$ map exhibited spatial variation that mimicked the field map spatial variations, which we do not expect in MWF maps (i.e., we expect myelin content to be independent of $\Delta\omega$). Table 3.4 shows numerical results for the *in vivo* data for several ROIs.

3.5 Discussion

Almost all of the optimized flip angles α and β for both scan designs A and B are equal to (or are very close to) the upper constraint, and there is a wide spread of tip-up phases (see Table 3.2). This seems to suggest that most of the information needed for estimating

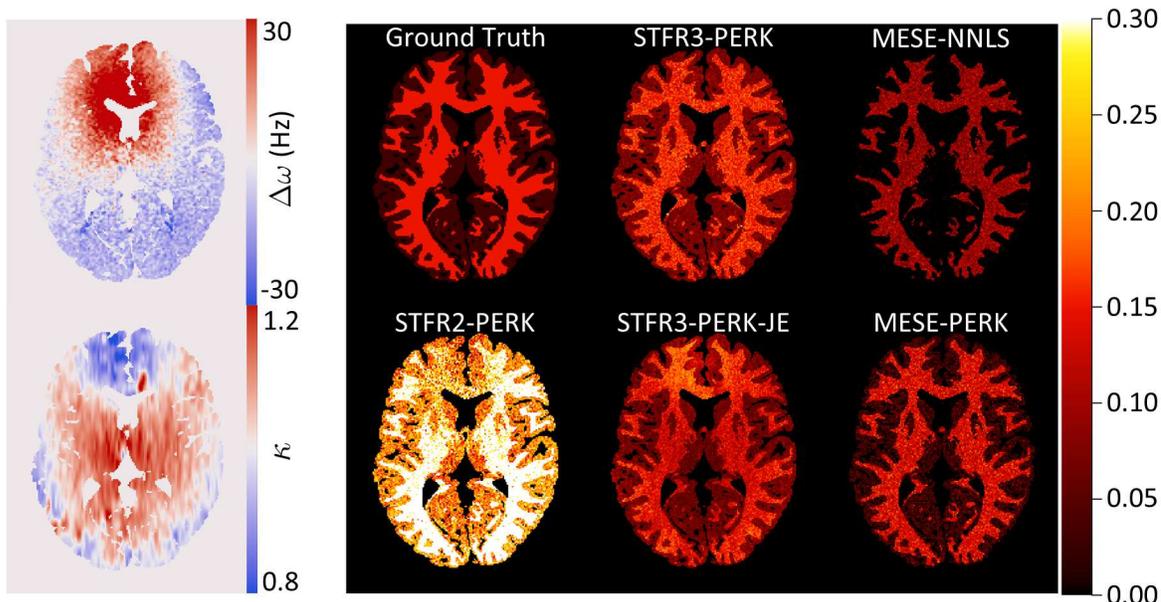
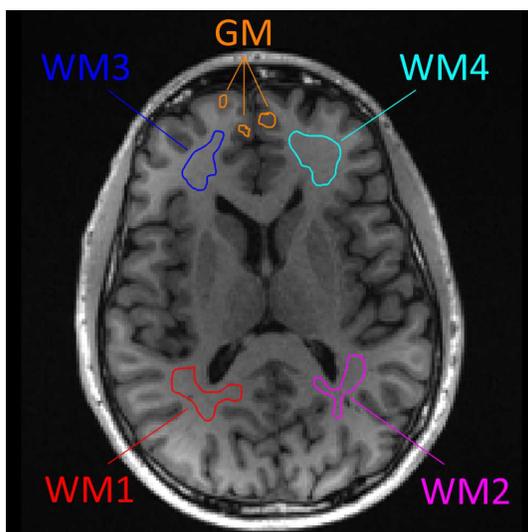


Figure 3.4: *Right*: MWF maps from five methods using simulated test data for a three-compartment tissue model with exchange. Table 3.3 reports numerical results. The proposed STFR3-PERK estimates are closer to the true MWF value for white matter tissue values than are the MESE-NNLS estimates. *Left*: Bulk off-resonance $\Delta\omega$ and flip angle scaling κ maps used in this simulation.

Table 3.4: *Left*: White matter (WM) and gray matter (GM) regions of interest (ROIs). The underlying image is from a standard MP-RAGE acquisition, acquired in the same scan session and registered to the other scans. *Right*: Sample means \pm standard deviations of MWF estimates for four WM ROIs and one GM ROI. Figure 3.6 shows corresponding MWF maps.



ROI	STFR2-PERK	STFR3-PERK
WM1	0.175 ± 0.021	0.116 ± 0.029
WM2	0.175 ± 0.009	0.117 ± 0.011
WM3	0.206 ± 0.010	0.133 ± 0.010
WM4	0.195 ± 0.008	0.138 ± 0.010
GM	0.187 ± 0.034	0.110 ± 0.029
ROI	MESE-NNLS	MESE-PERK
WM1	0.096 ± 0.042	0.105 ± 0.030
WM2	0.089 ± 0.046	0.097 ± 0.023
WM3	0.108 ± 0.036	0.133 ± 0.014
WM4	0.121 ± 0.039	0.141 ± 0.014
GM	0.034 ± 0.035	0.085 ± 0.034

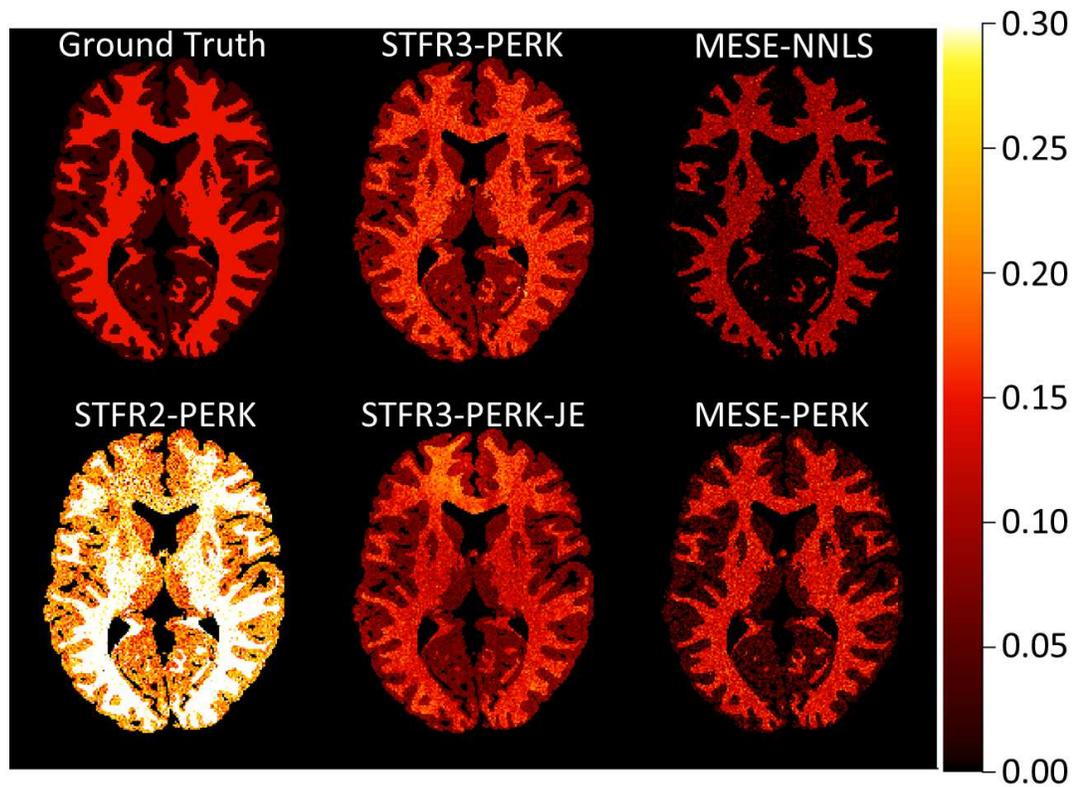


Figure 3.5: MWF maps from five methods using simulated test data for a nine-compartment tissue model with exchange. These results are essentially the same as when using the three-compartment exchanging model (see Figure 3.4). Thus, even though STFR3-PERK was trained with a three-compartment exchanging model, it still produced good MWF estimates from signal generated using a nine-compartment exchanging model.

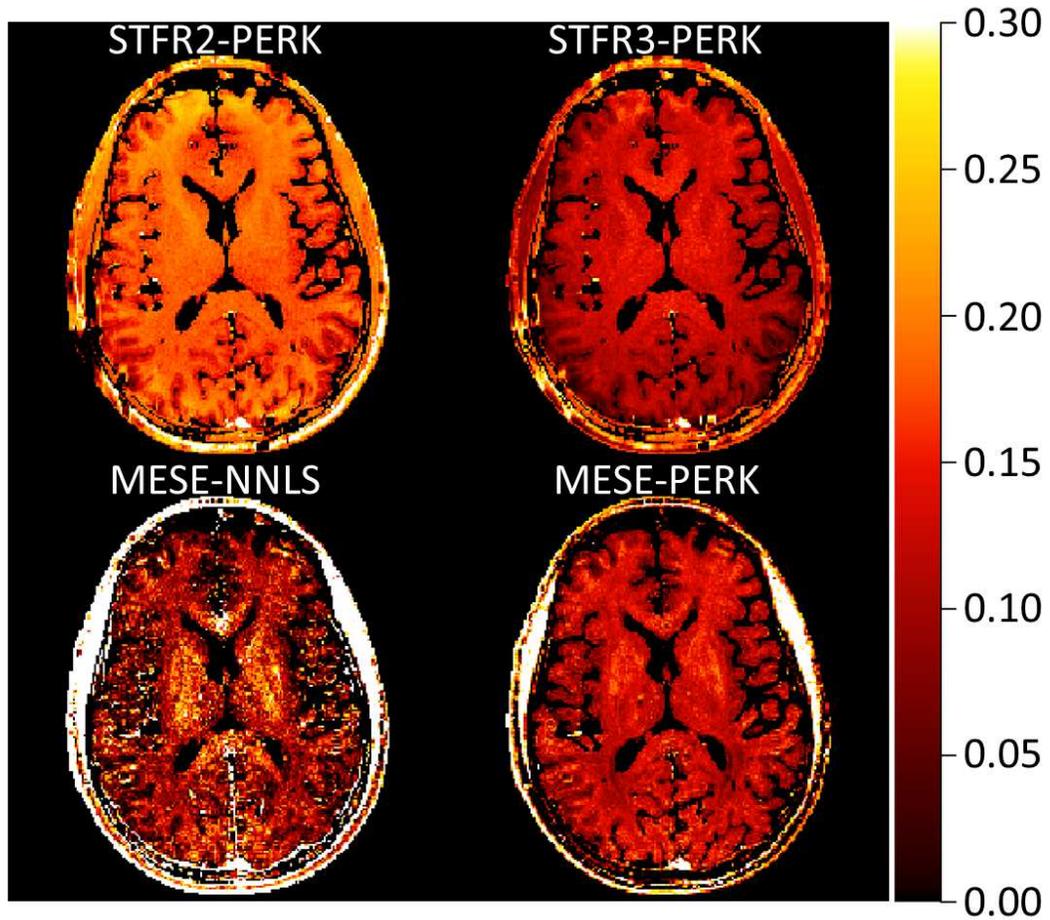


Figure 3.6: MWF maps from *in vivo* MESE data and STFR data using scan design A. Table 3.4 shows numerical results for several manually selected regions of interest. The MESE-NNLS MWF map appears noisier than those shown in other works. This is likely due to the lower SNR of our data due to differences in voxel size. To match the STFR resolution, we acquired MESE with 1.1 mm isotropic voxels, whereas often MESE data is collected with slice thickness of 5 mm and 1.6 mm or greater in the phase encode direction. Remarkably, MESE-PERK is much less noisy than MESE-NNLS. This is likely due to PERK being a Bayesian estimator that discourages estimates that are far from the mean MWF training value.

MWF lies in the phase accrual that occurs between the tip-down and tip-up excitations, so the flip angles should be chosen to maximize SNR. Interestingly, however, an unreported experiment showed that a scan design with flip angles set to 15° and with an even spread of tip-up phases ϕ resulted in CRLBs that were many orders of magnitude worse than the optimized scans. This result emphasizes the importance of the scan design process in choosing scan parameters, because these optimized parameters are robust across a range of $\Delta\omega_f$ values (see Figure 3.2). We also looked at optimized scan parameters when fixing T_{free} to 6 ms. We found that the tip-up phases still covered a spread of values, but the range of phases was slightly smaller, which makes sense because a smaller T_{free} leads to less off-resonance precession.

The expected CRLB for scan design A is better than that of design B when compared across many values of the additional myelin water $\Delta\omega_f$ (see Figure 3.2), as expected because the optimization of design B ignored the presence of $\Delta\omega_f$. Figure 3.2 also illustrates the impact that $\Delta\omega_f$ has on estimates of MWF; MWF becomes harder to estimate as $\Delta\omega_f$ approaches 0. These findings appear to be at variance with the findings in [54], where in multi-GRE MWF estimation modeling $\Delta\omega_f$ led to worse estimates at 3T. However, there is likely more information about $\Delta\omega_f$ in the STFR scans because of the optimized tip-up phases, which could explain why modeling $\Delta\omega_f$ in this work improved MWF estimation.

Simulated test data showed that scan design A and scan design B gave similarly good estimates of MWF across many values of $\Delta\omega_f$, at least for a range of tissue parameters (see Figure 3.3). At first glance, one may be surprised that design A performed noticeably better than design B with respect to the expected CRLB, and yet the two designs had similar RMSE values. One may also be surprised that the RMSE values were relatively small (about 25% of the mean MWF value) even though the expected CRLB predicted errors of 100% or more. However, PERK is a Bayesian estimator; thus, the unbiased CRLB does not necessarily predict the precision of MWF estimates computed by PERK. We minimized the unbiased CRLB during scan design because of its simplicity, but recognize that other objective functions for scan optimization may be better suited for MWF estimates from PERK or other Bayesian estimation methods. We investigated the effect of bias in Section S3 of the Supporting Information.

Simulated test data also shows that STFR-based myelin water imaging compares well to MESE-based approaches. Compared to the conventional MESE-NNLS, STFR3-PERK gives more accurate results in simulated white matter voxels, in addition to reducing estimation time by more than an order of magnitude (see Table 3.3). This result is interesting because the simulated MESE echoes generally had much higher SNR than the STFR scans. Combining MESE with PERK improves upon the NNLS results. However, the MESE scan is longer

than the combined time of all the STFR and BS scans. Furthermore, MWF estimation using MESE-PERK takes longer than STFR3-PERK because when simulating the MESE signal one must simulate a collection of spins to account for stimulated echoes, which is not necessary for STFR. This simulated data also shows that ignoring exchange when estimating MWF with STFR scans results in drastically overestimated MWF values (see Figure 3.4), so it is essential to generate training data that accounts for exchange. These same results hold even when the test data was generated using a nine-compartment exchanging model (see Figure 3.5). See Section S2 of the Supporting Information for results using a four-compartment exchanging model and a three-compartment non-exchanging model.

The *in vivo* MWF estimates in Figure 3.6 further emphasize the importance of modeling exchange. The MWF map given by STFR2-PERK has higher MWF values than the map given by STFR3-PERK. The three-compartment model led to maps that better agreed with the MWF maps estimated from MESE data. Table 3.4 indicates that in all white matter ROIs the STFR3-PERK estimates are within one standard deviation of the mean MESE-NNLS estimates. In gray matter it is different; however, this difference could be due to how the STFR training data were simulated, as typical values for gray matter $T_{1,f}$ and $T_{1,s}$ are slightly outside of the range of values generated for training. Figure 3.6 and Table 3.4 also demonstrate the effect that the estimation method has on MWF estimates: the MESE-PERK estimates have decreased standard deviation compared to the MESE-NNLS estimates. Furthermore, Table 3.4 demonstrates that STFR-based MWF estimates have lower standard deviation than MESE-based MWF estimates, despite the MESE scans being $5\times$ longer than the STFR scans.

For the *in vivo* data, we did not use a separately acquired bulk off-resonance $\Delta\omega$ map as a known parameter for our proposed STFR-based MWF estimation technique, even though the scan design and simulations assumed that $\Delta\omega$ was known. When we attempted to use the separately acquired $\Delta\omega$ map for the *in vivo* data, the MWF estimates appeared to be more biased in regions with high $\Delta\omega$ values. Further work is needed to investigate this behavior.

For the *in vivo* data we acquired a 9 mm slab in about 7 minutes total scan time with 1.1 mm isotropic resolution. Whole brain coverage would require 4 times as much data (with 2 mm slices), so our proposed approach would take about 28 minutes, which is longer than the 8 minutes achieved by the GRASE method. However, in [26] the authors under-sampled the GRASE data by a factor of 4, whereas we acquired fully sampled data. By under-sampling by the same factor the proposed STFR approach would achieve whole-brain coverage in about 7 minutes.

3.6 Conclusion

This work optimized a set of STFR scans that can be used to estimate MWF. We found that estimates of MWF are more precise for larger values of the frequency difference $\Delta\omega_f$ between myelin water and non-myelin water. Fortunately, in white matter reported values of $\Delta\omega_f$ that are far enough away from 0 to aid estimation of MWF [1]. We also found that modeling exchange (i.e., using a more accurate tissue model) greatly impacts the MWF estimates from STFR scans. When modeling exchange, STFR with PERK yields MWF estimates that are comparable to MESE-based MWF estimates.

This is the first work to compare STFR-based MWF estimation to MESE-based MWF estimation. Additionally, to our knowledge, this is the first work to generate MWF estimates from a MESE scan using PERK. While this estimation method was not the main point of this chapter, it illustrates another potential method for MWF estimation.

This study was a proof-of-concept study to see if STFR could be applied to estimating MWF. As such, only a single healthy volunteer was scanned. While the initial comparison of STFR to MESE is promising, future studies should compare the two methods across multiple volunteers. Additionally, our proposed MWF estimation method should be validated in pathology to verify that it can detect, e.g., multiple sclerosis lesions. Such verification is especially important because the proposed method assumes a fixed number of tissue compartments, which may or may not inhibit its sensitivity to anomalies.

There are several ways in which the scan design process could be further explored. Our choice to optimize nine STFR scans for scan design was somewhat arbitrary, so one could explore different numbers of scans to see how the CRLB is affected for a given scan time budget. Additionally, one could change the space of scan parameters over which to optimize; especially interesting would be to increase the upper bound on α and β to see if the optimized scans would have a greater variety of flip angles. Another route to explore is to adjust the weighting matrix \mathbf{W} to optimize STFR scans for estimating other parameters in addition to or instead of MWF. In particular, since the results here suggest that the STFR scans are sensitive to the effects of exchange, it could be interesting to optimize STFR scan design for quantifying exchange parameters. Additionally, future work could explore what parameters to include as known parameters versus unknown parameters, in both the scan design and in PERK.

Finally, to reduce the scan time of the STFR scans, either to allow for more scans or to reduce scan time, one could under-sample the MRI k-space data. The image reconstruction would then be under-determined, thus requiring some sort of regularized reconstruction. Methods that jointly reconstruct all D scans at once would be a natural approach, e.g.,

[55, 56].

3.7 Supporting Information

This supporting information presents additional results and discussion for experiments not included in the main body of the chapter.

3.7.1 Estimator RMSE for White and Gray Matter Tissue Values

We compared MWF estimates from scan designs A and B. We simulated test data using the two-compartment non-exchanging STFR signal model using tissue values typical of white matter and gray matter (see Table 3.1), and we estimated MWF using STFR2-PERK. We plotted RMSE of MWF estimates from both scan designs versus the additional myelin water off-resonance $\Delta\omega_f$. Figure 3.7 shows the results.

Figure 3.7 indicates that scan design A gives better MWF estimates in white matter over values of $\Delta\omega_f$ we expect to see, but scan design B performs better in gray matter. However, the values of $T_{1,f}$ and $T_{1,s}$ for gray matter are (slightly) outside of the range of values used for the scan designs and for training our estimator. When quantifying MWF in gray matter is of interest, one probably should use a wider range of values for scan design and training.

3.7.2 Estimator Performance with Model Mismatch

We compared MWF estimates from STFR2-PERK, STFR3-PERK, MESE-NNLS, MESE-PERK, and STFR3-PERK-JE for different ground truth models. First, we generated test data for white matter and gray matter tissue values using a four-compartment exchanging model. The four compartments were myelin water, axonal water (i.e., water in myelinated axons), all other water, and macromolecules. Myelin water was in exchange with the macromolecular pool, myelin water and axonal water exchanged with each other, and myelin water and all other water exchanged with each other. Figure 3.8 shows the results, and Table 3.5 reports numerical values. STFR3-PERK still provides good MWF estimates despite the model mismatch between the test data and the training data.

We then generated test data for white matter and gray matter tissue values using a three-compartment non-exchanging model. The three compartments were the same as in the three-compartment exchanging model that STFR3-PERK was trained with, except no exchange occurred (i.e., the exchange rates were set to 0). Figure 3.9 shows the results, and Table 3.6 reports numerical values. Without exchange, the three-compartment model becomes essentially a two-compartment model because the T_2 of the macromolecular pool

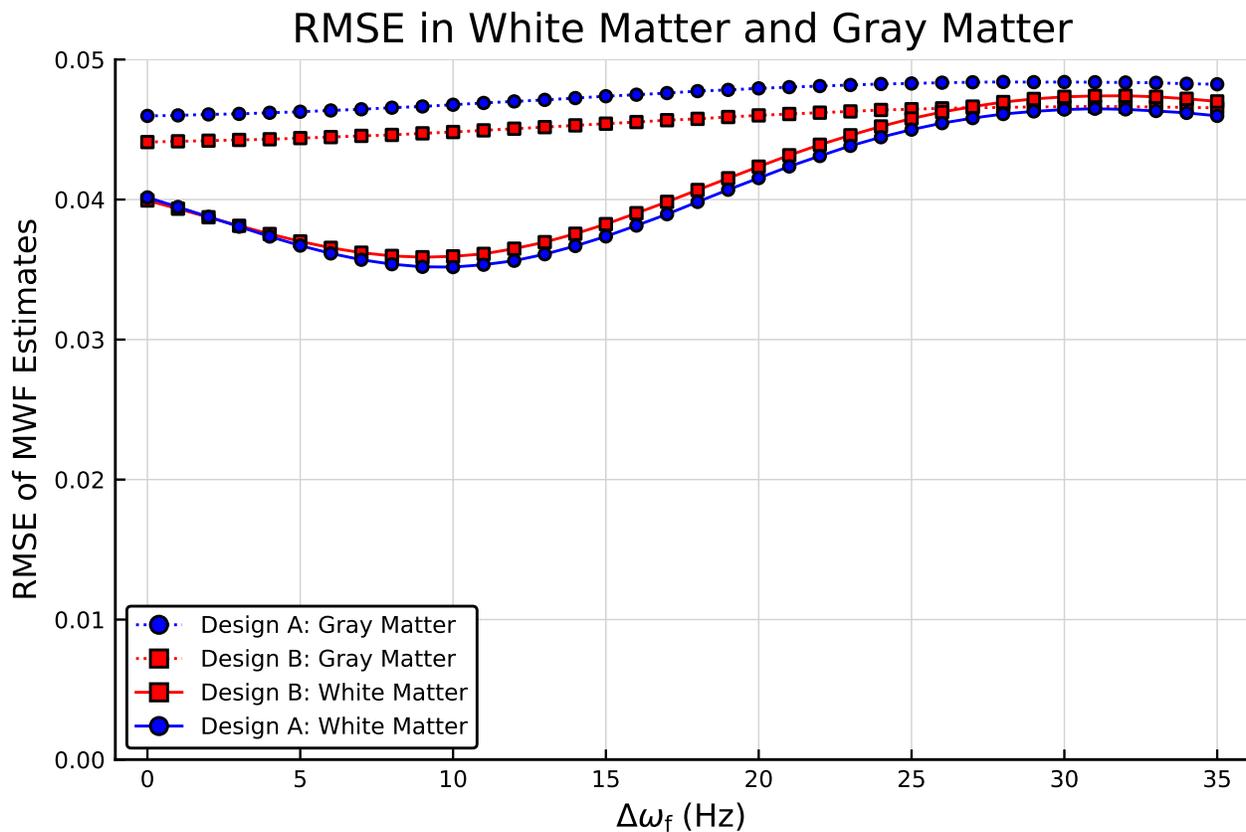


Figure 3.7: RMSE of MWF estimates for white matter and gray matter simulated test data. Scan design A has better RMSE in white matter for values of $\Delta\omega_f$ we expect to see in white matter. This better RMSE in white matter is at the cost of worse RMSE in gray matter. Note that the values of $T_{1,f}$ and $T_{1,s}$ for gray matter were outside of the range of values used for the scan designs and for training our estimator.

Table 3.5: Numerical results for Figure 3.8.

	White Matter (MWF = 0.15)			Gray Matter (MWF = 0.03)			Time (s)
	RMSE	Mean	SD	RMSE	Mean	SD	
STFR2-PERK	0.170	0.308	0.062	0.112	0.133	0.044	14.7
STFR3-PERK	0.028	0.130	0.020	0.028	0.052	0.017	42.1
STFR3-PERK-JE	0.040	0.120	0.026	0.028	0.046	0.022	42.2
MESE-NNLS	0.071	0.084	0.024	0.029	0.001	0.004	1623.6
MESE-PERK	0.033	0.127	0.023	0.056	-0.005	0.043	167.3

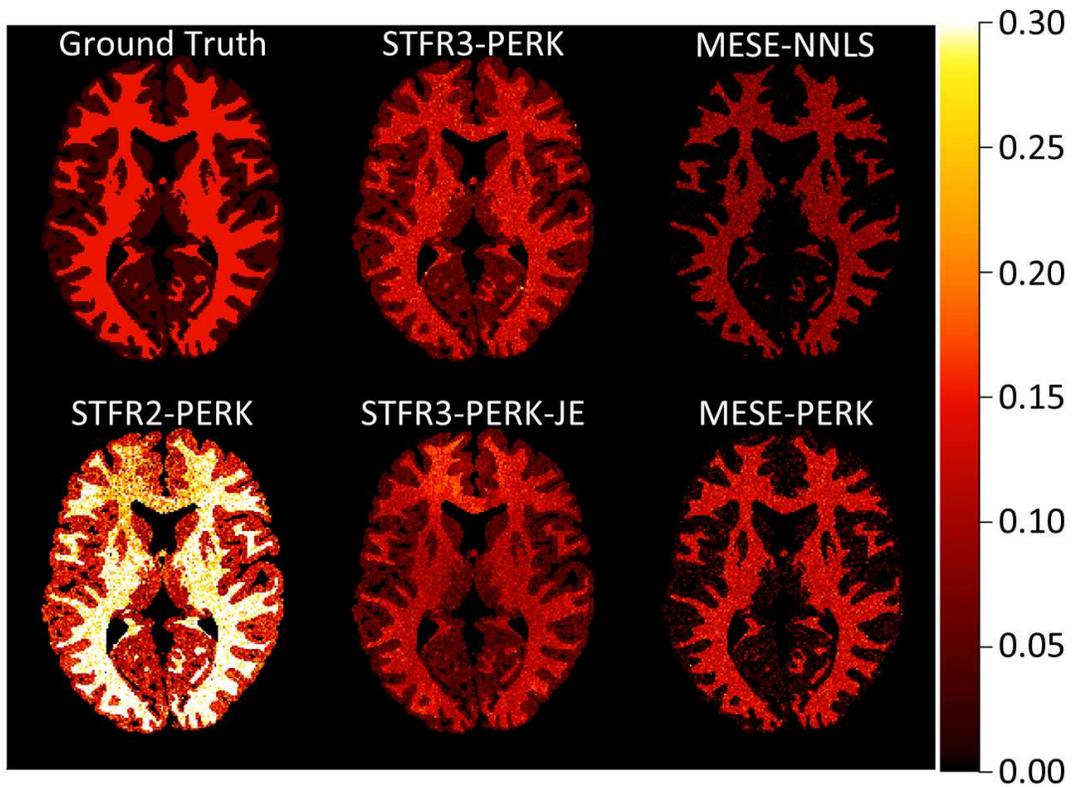


Figure 3.8: MWF maps from five methods using simulated test data for a four-compartment tissue model with exchange. The four compartments considered were myelin water, axonal water (i.e., water in myelinated axons), all other water, and macromolecules. The results are similar to those using the three-compartment model with exchange. Table 3.5 shows numerical results.

Table 3.6: Numerical results for Figure 3.9.

	White Matter (MWF = 0.15)			Gray Matter (MWF = 0.03)			Time (s)
	RMSE	Mean	SD	RMSE	Mean	SD	
STFR2-PERK	0.048	0.181	0.037	0.047	0.045	0.044	14.8
STFR3-PERK	0.097	0.055	0.020	0.051	0.058	0.043	41.9
STFR3-PERK-JE	0.092	0.061	0.024	0.047	0.045	0.045	41.9
MESE-NNLS	0.031	0.148	0.031	0.027	0.007	0.013	1606.2
MESE-PERK	0.038	0.178	0.025	0.046	0.066	0.029	142.1

is so small. Thus it makes sense that STFR2-PERK gives good MWF estimates. The overestimation of MWF could be because the macromolecular pool has a nonzero f_m , but since it contributes no signal the estimator assumes that the smaller signal is due to a larger MWF. MESE-NNLS does better without exchange, though it still underestimates gray matter MWF, while STFR3-PERK does poorly. It is possible, though, that if the training ranges for the residence times were adjusted appropriately (increased, to allow for less exchange) then STFR3-PERK would also do well, although doing so might cause greater estimator bias. It is somewhat surprising that MESE-PERK still gives good MWF estimates, despite being trained with the three-compartment exchanging model. This could be because the T_R of the MESE scan is long compared to the residence times (more than $10\times$ longer).

3.7.3 Estimator Bias

To assess the effect of MWF estimator bias, we computed the biased CRLB [30] of scan design A for fixed white matter tissue values (for the two-compartment non-exchanging model). (This is unlike what we did in Figure 3.2, where we calculated an expected CRLB over distributions of the parameters.) The biased CRLB was computed by multiplying the inverse Fisher information matrix (3.5) by the partial derivatives with respect to the unknown parameters of the expected parameter estimates given by PERK [30]. The biased CRLB indeed was smaller than the unbiased CRLB (see Figure 3.11), suggesting that estimator bias is why our estimates had low variance. We investigated the bias of our STFR3-PERK estimator for test data using the three-compartment exchanging model with fixed white matter tissue values. We found that even with (mild) estimator bias, our proposed MWF estimation technique is still sensitive to changes in MWF (see Figure 3.12). Furthermore, our estimator bias decreases as SNR increases (e.g., by using larger voxels). Thus, while the proposed method is biased, it still shows promise for detecting changes in MWF.

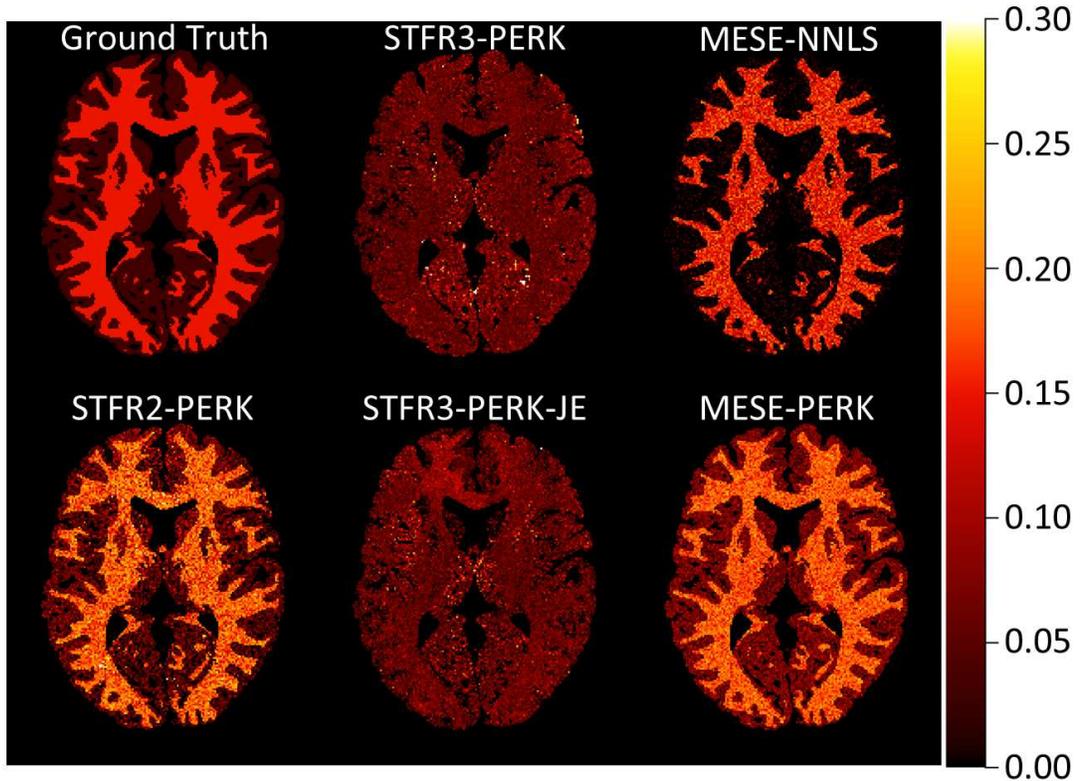


Figure 3.9: MWF maps from five methods using simulated test data for a three-compartment tissue model without exchange. Without exchange, the three-compartment model becomes essentially a two-compartment model because the T_2 of the macromolecular pool is so small. Thus it makes sense that STFR2-PERK performs well. Surprisingly, MESE-PERK still produces good MWF estimates, even though it is trained with the three-compartment exchanging model (like STFR3-PERK). This could be because the T_R of the MESE scan is long compared to the residence times governing exchange. Furthermore, it is possible that if the training ranges for the residence times were adjusted appropriately (increased) then STFR3-PERK would also do well. Table 3.6 shows numerical results.

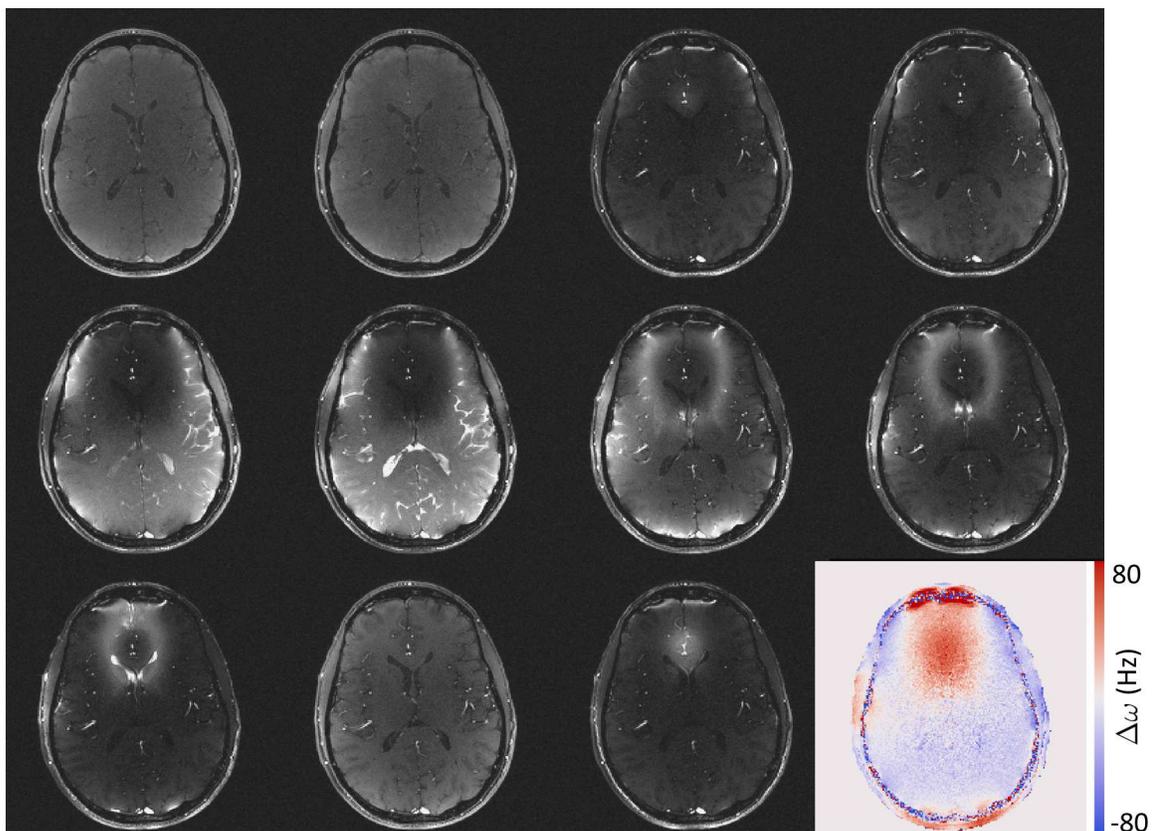


Figure 3.10: *In vivo* images for two SPGR and nine STFR scans using scan design A. Each image is the square root sum of squares combination of the individual coil data. STFR produces contrast similar to balanced SSFP, including a similar off-resonance profile that induces the characteristic banding artifact of balanced SSFP. Different points of this profile are sampled as the phase ϕ of the STFR tip-up excitation varies. The nine STFR images are sorted by increasing ϕ , so this off-resonance profile is easily visualized. In the lower right is the field map estimated from the two SPGR scans.

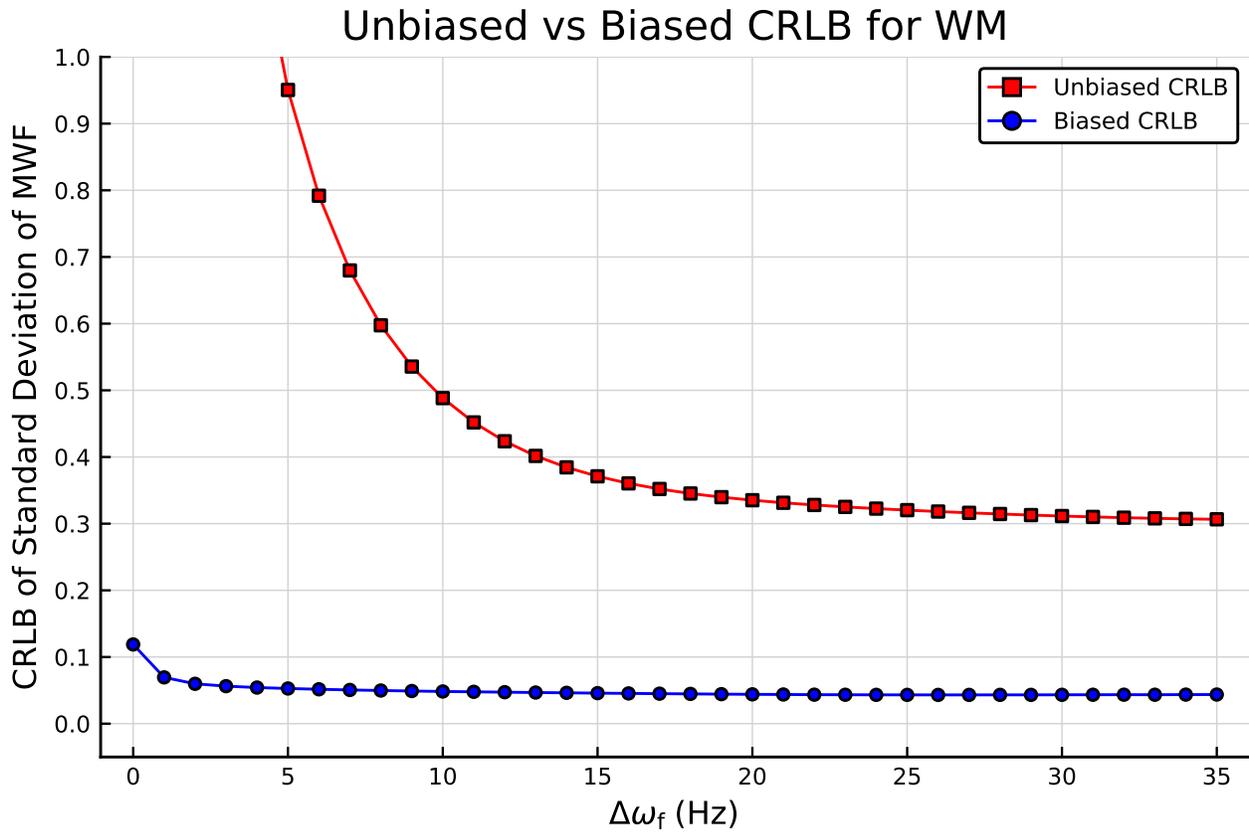


Figure 3.11: Comparison of biased and unbiased CRLBs for white matter tissue values using the two-compartment non-exchanging model. The biased CRLB is much lower than the unbiased CRLB, suggesting that bias is the reason why our STFR-based MWF estimation results in estimates with low variance. However, our proposed method still shows sensitivity to changes in MWF (see Figure 3.12).

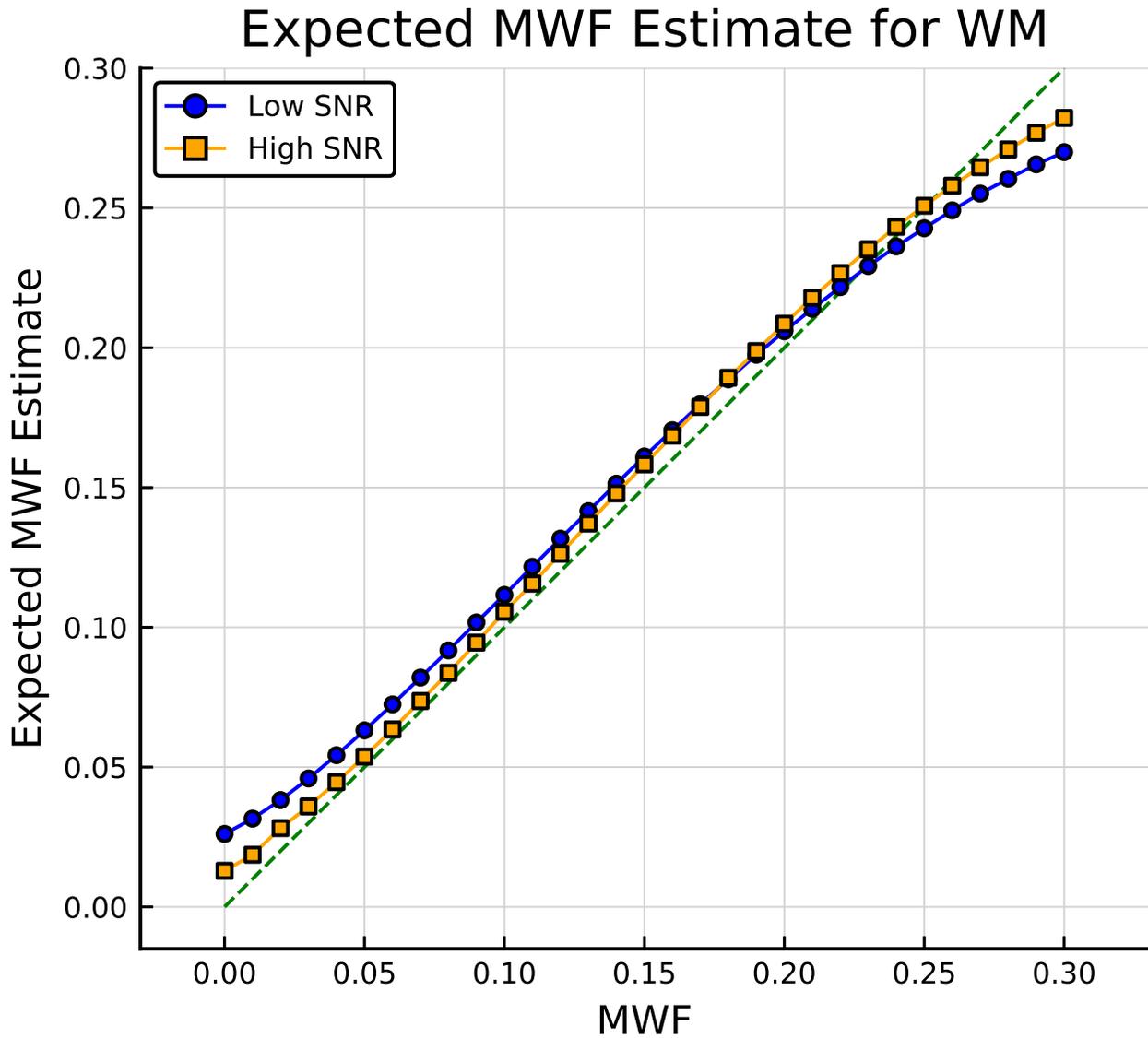


Figure 3.12: Expected MWF estimates from the proposed STFR3-PERK MWF estimation technique for fixed white matter tissue values from a three-compartment exchanging model. The proposed method is (mildly) biased, yet it is still very sensitive to changes in true MWF value. Furthermore, bias decreases as SNR increases. (An unbiased estimator would have estimates along the line of identity, i.e., along the dashed line.)

CHAPTER 4

End-to-End MRI Scan Design Approach for Improved Myelin Water Imaging

This chapter presents in-preparation journal paper [13] that extends conference abstract [14]. Sections 4.1–4.6 are taken from the main paper [13], and Section 4.7 is taken from the supporting information of the paper. Further results not found in the paper are also presented in Section 4.8.

4.1 Introduction

Myelin water imaging (MWI) consists of making quantitative maps of myelin water fraction (MWF) [2], the proportion of the total magnetic resonance imaging (MRI) signal within a voxel that arises from myelin water, or water trapped within the myelin sheath. MWF maps have potential clinical utility in that they can allow tracking the progression of demyelinating diseases such as multiple sclerosis [6].

The most widely accepted MWI technique uses the multi-echo spin echo (MESE) scan [6]. A faster variant of this approach uses a combined gradient and spin echo (GRASE) scan [25]. Steady-state scans also have been used for MWI, including combinations of balanced steady-state free precession (bSSFP) and spoiled gradient-recalled echo (SPGR) scans [27], multi-echo gradient echo scans (mGRE) [31–33], dual-echo steady-state (DESS) scans [34–36], and small-tip fast recovery (STFR) scans [10–12, 14]. With the exception of the MWI approaches using DESS and STFR, these approaches use manually chosen scan parameters, which may limit their overall effectiveness.

In MRI, there are a number of scan parameters that are user-controlled and that affect the acquired signal. In situations where many different scans are run, e.g., in MWI or in quantitative MRI (QMRI) in general, one would like to choose scan parameters that produce images that are as informative as possible. Because of complicated nonlinear relationships between scan parameters and the acquired images, optimizing the scan parameters is not

always straightforward. One way to automate the design of scan parameters is first to define an objective function that provides some measure of how much information a set of images produced by given scan parameters has, and then optimize the objective function to find the scan design that gives the best objective function value.

The Cramér-Rao Lower Bound (CRLB), a lower bound on the variance of parameter estimates from an unbiased estimator, often is used for optimizing scan parameters in QMRI applications. For example, the CRLB has been used for optimizing scan parameters for estimating relaxation times [9, 40, 42, 44, 57], for B1 mapping [43], for MWI [10–12, 34–36], for arterial spin labeling [58], in magnetic resonance fingerprinting (MRF) [57, 58], and when estimating parameters using machine learning [9–12, 34–36, 58]. However, the CRLB characterizes only *unbiased* estimators; therefore, scan parameters optimized using the CRLB may be suboptimal when paired with a biased estimator (such as regularized least squares methods and machine learning models).

Other methods also exist for designing scan parameters. For example, in MRF one can maximize the discrimination between tissue types present in the pattern matching dictionary [59, 60].

An alternative scan design approach is to use an end-to-end approach, where one incorporates an estimator into the scan design process and directly optimizes the quality of parameter estimates from that estimator rather than the CRLB or any other metric that is not specific to any one estimator. Such an approach has been used in, e.g., coded aperture design, in which a neural network was trained while jointly optimizing the set of coded apertures [61]. End-to-end approaches also have been used in MRF to jointly train a neural network to estimate relaxation times while optimizing the MRF scan parameters [62] and to design a MRF scan to minimize the error of relaxation time maps estimated using pattern matching [63].

In this work, we introduce an end-to-end scan design approach for QMRI and apply it to MWI. Specifically, we optimize scan parameters for a set of STFR scans [8] by minimizing the mean squared error (MSE) of MWF estimates obtained using parameter estimation via regression with kernels (PERK) [9]. We compare this proposed approach to optimizing STFR scan parameters by minimizing the CRLB of MWF estimates and show improved MWF estimates using the proposed approach. This work expands upon the results of a recent conference abstract [14].

The organization of this chapter is as follows. Section 4.2 details the end-to-end scan design approach. Section 4.3 describes our experiments. Section 4.4 reports the experimental results. Section 4.5 provides some discussion of the results and the proposed scan design approach. Section 4.6 gives concluding remarks.

4.2 Theory

Suppose, for a given voxel, there are L tissue parameters (e.g., relaxation times, field map¹, etc.) to estimate that describe the MRI signal for the voxel. Further suppose we want to optimize P scan parameters (e.g., flip angles, repetition times, etc.). Let $\mathbf{f} : \mathbb{R}^L \times \mathbb{R}^P \rightarrow \mathbb{C}^D$ be a signal model that takes tissue and scan parameters \mathbf{x} and \mathbf{P} as input and outputs the signal produced for a single voxel by a set of D scans (or a time series of length D in the MRF setting). Note that \mathbf{f} may also depend on other parameters (scan parameters or not) that are not estimated/optimized, but we assume for simplicity of notation that these parameters are included in \mathbf{f} . (Such parameters could include scan parameters that are held fixed or tissue parameters that are assumed to be known, e.g., off-resonance determined via a separately acquired field map, in which case \mathbf{f} would differ from voxel to voxel according to the spatial variation of the known parameter.) Thus \mathbf{P} contains all the scan parameters to be optimized for each of the D scans (or for the MRF schedule).

The scan design process involves solving

$$\arg \min_{\mathbf{P} \in \mathcal{P}} \Psi(\mathbf{P}), \quad (4.1)$$

where Ψ is a cost function to minimize and \mathcal{P} is the set of possible scan designs. (\mathcal{P} includes scans that are physically realizable, e.g., scans with positive timing parameters, and that satisfy user-defined constraints such as flip angle limits.) For a CRLB-based scan design, Ψ computes, e.g., an average CRLB over a distribution of tissue parameters one expects to see when scanning [10, 40]. For an end-to-end scan design, Ψ will involve parameter estimates given by a particular choice of estimator $\hat{\mathbf{x}}(\cdot)$ that estimates the tissue parameters given the output of a signal model \mathbf{f} .

In this work, we use MSE for the cost function. Given a signal model \mathbf{f} and N sets of tissue parameters $\{\mathbf{x}_n\}_{n=1}^N$, the cost function has the form

$$\Psi(\mathbf{P}) = \frac{1}{N} \sum_{n=1}^N \|\mathbf{W}(\hat{\mathbf{x}}(\mathbf{f}_\epsilon(\mathbf{x}_n, \mathbf{P})) - \mathbf{x}_n)\|_2^2, \quad (4.2)$$

where \mathbf{W} is a diagonal weighting matrix used to place emphasis on the parameters of interest, and $\mathbf{f}_\epsilon(\mathbf{x}_n, \mathbf{P}) = \mathbf{f}(\mathbf{x}_n, \mathbf{P}) + \epsilon_n$ is the signal model with noise added to simulate scanner

¹In this work, we use the term “tissue parameters” to mean all parameters that influence the acquired MRI signal that are not user-controlled scan parameters. We include non-tissue-specific parameters such as B0 and B1+ inhomogeneities as well. An alternative label could be “non-scan parameters”, but that seems less clear (though more precise). Because the main “tissue parameter” of interest in this work is MWF, which is an actual tissue parameter, we trade precision of writing for clarity of intent.

noise. With this cost function, we optimize scan parameters specifically to improve the error of our chosen estimator, thus tailoring the scan parameters to the estimator’s needs.

Typically, the tissue parameters \mathbf{x}_n will be drawn randomly from user-chosen distributions. If $\hat{\mathbf{x}}(\cdot)$ is a supervised machine learning method, then T training data pairs $\{(\tilde{\mathbf{x}}_t, \mathbf{f}_\varepsilon(\tilde{\mathbf{x}}_t, \mathbf{P}))\}_{t=1}^T$ also are required, in which case the training tissue parameters $\tilde{\mathbf{x}}_t$ also typically will be drawn randomly from user-chosen distributions (usually the same as used for the \mathbf{x}_n draws). We randomly draw these parameters once and then hold them constant throughout the scan design process.

Another consideration for supervised machine learning estimators is that the training data pairs $(\tilde{\mathbf{x}}_t, \mathbf{f}_\varepsilon(\tilde{\mathbf{x}}_t, \mathbf{P}))$ depend on \mathbf{P} . As a result, one cannot train $\hat{\mathbf{x}}(\cdot)$ once and use it throughout the scan design process; rather, $\hat{\mathbf{x}}(\cdot)$ must be retrained whenever \mathbf{P} is updated. In this case (or in any case where the estimator depends on \mathbf{P} , such as when generating a dictionary for pattern matching in MRF), the MSE cost function more appropriately would be written

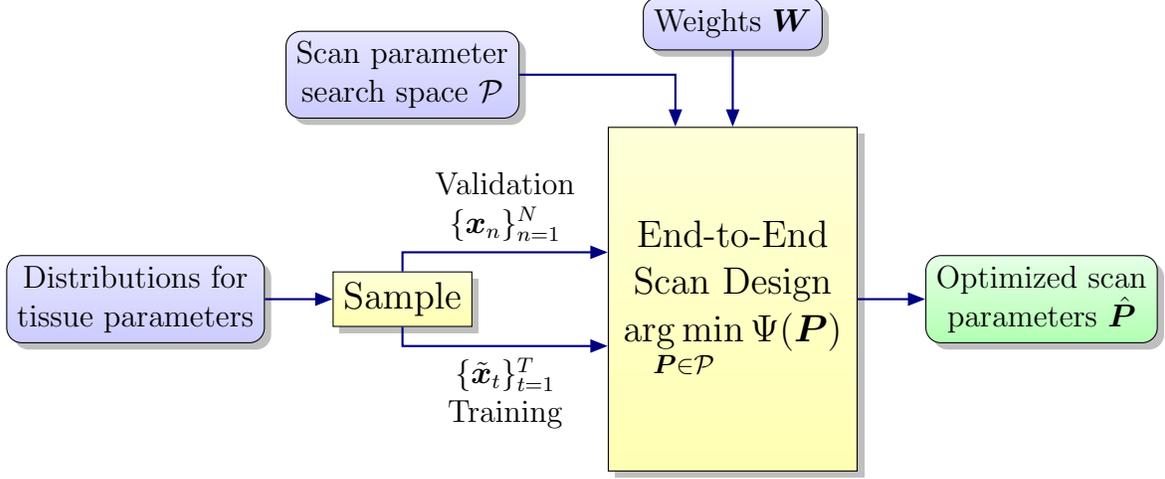
$$\Psi(\mathbf{P}) = \frac{1}{N} \sum_{n=1}^N \|\mathbf{W}(\bar{\mathbf{x}}(\mathbf{f}_\varepsilon(\mathbf{x}_n, \mathbf{P}), \mathbf{P}) - \mathbf{x}_n)\|_2^2, \quad (4.3)$$

where $\bar{\mathbf{x}}(\cdot)$ first trains $\hat{\mathbf{x}}(\cdot)$ using \mathbf{P} and then applies $\hat{\mathbf{x}}(\cdot)$ to estimate tissue parameters. Figure 4.1 shows a block diagram illustrating the end-to-end scan design approach using (4.3).

4.3 Methods

In this work, we used a tissue model with two compartments: myelin water and non-myelin water. In each voxel we modeled $L = 10$ tissue parameters: the overall M_0 , MWF f_f , T_1 for each compartment ($T_{1,f}$ and $T_{1,s}$), T_2 for each compartment ($T_{2,f}$ and $T_{2,s}$), residence time $\tau_{f \rightarrow s}$ of water exchange from myelin water to non-myelin water, bulk off-resonance frequency $\Delta\omega$, additional off-resonance frequency offset $\Delta\omega_f$ experienced by myelin water [1], and flip angle scaling κ . (The subscript “f” refers to the fast-relaxing compartment, or myelin water, while the subscript “s” refers to the slow-relaxing compartment, or non-myelin water.) We assumed the two water pools were in chemical equilibrium, and thus the residence time of exchange from non-myelin water to myelin water was determined from f_f and $\tau_{f \rightarrow s}$ [10]. We evaluated the STFR signal model with this two-compartment exchanging tissue model using the Bloch-McConnell equation [18] implemented in the Julia [50] package BlochSim.jl [64]. Table 4.1 lists the parameter distributions used for generating $N = 1,000$ validation and $T = 20,000$ training sets of tissue parameters for the scan designs and for training PERK. These parameter distributions were chosen to cover a range of tissue parameter values seen in

End-to-End Scan Design



End-to-End Cost Function $\Psi(\mathbf{P})$

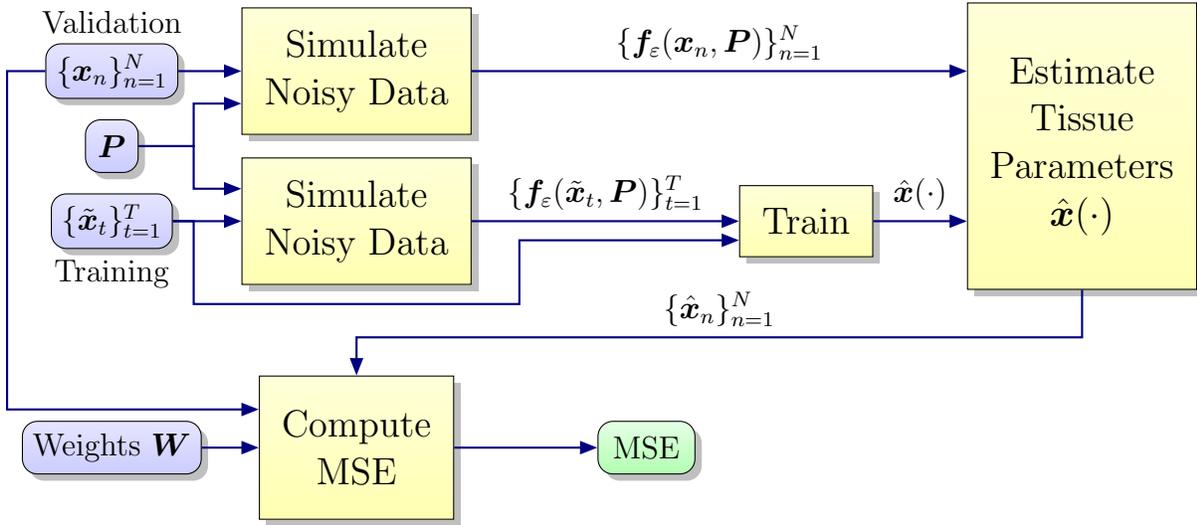


Figure 4.1: Diagram of the end-to-end scan design process, in which scan parameters \mathbf{P} are optimized subject to constraints \mathcal{P} to minimize the end-to-end cost function (4.3). Validation and training sets of tissue parameters are sampled before and held fixed during the iterative optimization. The end-to-end cost function is computed in several steps. First, the current candidate scan design \mathbf{P} is used to simulate noisy data, using a signal model \mathbf{f} , from both the validation and training tissue parameter sets. The training set of tissue parameters is then paired with the associated noisy signals and used to train the estimator $\hat{\mathbf{x}}(\cdot)$. Next, the noisy signals from the validation set of tissue parameters are given to the estimator to estimate the underlying tissue parameters. Finally, these estimates are compared to the true tissue parameters to compute a (weighted) MSE. The scan design process seeks to minimize this MSE.

Table 4.1: Tissue parameter distributions used in scan designs and in simulation results, unless otherwise noted in the text. Parameters were uniformly distributed between the given minimum and maximum values. PERK training also used these same distributions, except for a slightly expanded M_0 distribution ($[0.6, 1.2]$ instead of the values in the table) to account for the uncertainty in *in vivo* M_0 values due to unknown scanner gain. Note that we chose the range for $\Delta\omega$ to cover one full period of the magnitude STFR signal model (which is periodic in $\Delta\omega$ with period $1/T_{\text{free}}$, where we fixed $T_{\text{free}} = 8$ ms for the scan design process).

Parameter	Minimum	Maximum
M_0	0.7	1
f_f	0.02	0.31
$T_{1,f}$ (ms)	200	500
$T_{1,s}$ (ms)	1500	2500
$T_{2,f}$ (ms)	10	25
$T_{2,s}$ (ms)	60	100
$\tau_{f \rightarrow s}$ (ms)	10	150
$\Delta\omega_f$ (Hz)	0	10
$\Delta\omega$ (Hz)	-62.5	62.5
κ	0.8	1.2

cerebral white matter (WM) and gray matter (GM) [2–5, 23]. (See Section 4.7 and Table 4.3 for a discussion of using the results in [5] to determine values to use for $T_{1,f}$ and $T_{1,s}$.)

4.3.1 Scan Designs

We sought to optimize the scan parameters of D STFR scans for improved MWF estimation (i.e., we chose \mathbf{W} in (4.3) to place emphasis only on MWF, thus computing the MSE of MWF estimates). We chose to use STFR scans because STFR is a fast, steady-state scan that also incorporates an additional tip-up RF pulse, giving additional degrees of freedom (relative to typical steady-state scans that include just one RF pulse per repetition) for the scan design process to explore and potentially exploit for improved MWF estimation. Furthermore, in addition to relaxation times, STFR is also sensitive to $\Delta\omega_f$ [10], which we modeled in this work. (We note, however, that the end-to-end scan design approach proposed in this work is general and can be applied using other scans or even a combination of different scans.) For simplicity, we fixed all timing parameters with $T_{\text{free}} = 8$ ms (time between the excitation, or tip-down, pulse and the tip-up pulse), $T_g = 2.8$ ms (time after the tip-up pulse during which gradient spoiling occurs), and $T_E = 4$ ms. We optimized the tip-down angle $\alpha \in [1, 75]^\circ$, the tip-up angle $\beta \in [0, 75]^\circ$, and the phase of the tip-up pulse $\phi \in [-180, 180]^\circ$ for each STFR scan. Thus we optimized a total of $P = 3D$ scan parameters.

We used the BFGS optimizer in the Optim.jl [65] package to optimize the scan parameters. We computed gradients of the cost function and for computing the CRLB using the Zygote.jl [66] and ForwardDiff.jl [67] automatic differentiation packages. We ran the scan designs until reaching convergence to a local minimum (when the ℓ_2 -norm of the gradient of the cost function was less than 10^{-8} , the default Optim.jl tolerance) or after running for 16 days. We wrote the code for running the scan designs (and for the experiments below) in the Julia programming language [50]. The code is publicly available at <https://github.com/StevenWhitaker/EndtoEndScanDesign>.

In addition to optimizing scan parameters, we also selected the two hyperparameters used in PERK: the Tikhonov regularization parameter ρ and the Gaussian kernel length scale λ . We used cross-validation to select the hyperparameters. Specifically, we randomly selected 20,000 sets of training points and 1,000 sets of validation points using the distributions in Table 4.1. Then for each $(\rho, \lambda) \in \{2^{-35}, 2^{-34}, \dots, 2^0\} \times \{2^{-5}, 2^{-4.9}, \dots, 2^5\}$, we trained PERK and computed the MSE of MWF estimates from the validation data and then chose the ρ and λ that resulted in the least MSE. We did cross-validation before optimizing scan parameters, i.e., using the initial scan design, held ρ and λ fixed during the scan design, and then did cross-validation again with the optimized scan parameters.

In previous work [10], $\Delta\omega$ and κ were taken to be known parameters, simplifying the computation of the CRLB during the scan design process at the cost of needing to acquire separate B0 and B1+ maps when scanning. Here, we first performed four end-to-end scan designs of 18 STFR scans, varying whether $\Delta\omega$ and κ were known or unknown parameters to determine the impact knowledge of those parameters has on the optimized root mean squared error (RMSE) of MWF estimates, thus determining whether acquiring separate B0 and B1+ maps is worth the additional scan time needed. For each of the four scan designs, we initialized 16 of the STFR scans with α linearly spaced in the range $[15, 30]^\circ$, $\beta = c \cdot \alpha$ with c linearly spaced in the range $[0.9, 1]$, and ϕ linearly spaced in the range $(-180, 180]^\circ$. The remaining two STFR scans were initialized with $T_{\text{free}} = 10.3$ ms, $T_g = 2.8$ ms, $\alpha = 5^\circ$, $\beta = 0$, $\phi = 0$, and two different echo times: $T_{E,1}/T_{E,2} = 4/6.3$ ms. These two scans are equivalent to a pair of T_E -offset spoiled gradient recalled echo (SPGR) scans that can be used for separate $\Delta\omega$ estimation. Thus, for the two designs where $\Delta\omega$ was taken to be known these two SPGR scans were held fixed, whereas for the two designs where $\Delta\omega$ was taken to be unknown the α , β , and ϕ of these two STFR scans were optimized as for the other 16 STFR scans. Note that in the case where $\Delta\omega$ was taken to be known we included the fixed SPGR scans with the STFR scans during the scan design so the scan design could optimize the STFR scans while accounting for any (potential) additional knowledge provided by the SPGR scans beyond just a B0 map. We did not include Bloch-Siegert scans (used for separate κ estimation) with

the STFR scans because neither type of scan is a special case of the other. After these four scan designs, we took $\Delta\omega$ and κ to be unknown for the remaining scan designs.

We then investigated how the number of optimized STFR scans impacts the RMSE of MWF estimates. We varied the number of scans D from 11 to 20 and we ran 10 scan designs with different initializations for each, for a total of 100 scan designs (separate from the four scan designs from the previous paragraph). We used multiple starting points because the end-to-end scan design cost function is non-convex, and thus there are possibly many global and local minima, and the choice of initialization influences which local minimum is found. For each D , one initialization was similar to the initialization described above, excluding the two SPGR scans. For the other initializations, we uniformly randomly chose the STFR scan parameters for each scan such that $\alpha \in [5, 70]^\circ$, $\beta \in [0, \alpha]$, and $\phi \in [-180, 180]^\circ$.

We next ran 10 CRLB scan designs, in which we optimized the STFR scan parameters for $D = 18$ scans to minimize the mean CRLB of MWF estimates, as in [10]. We computed this mean by averaging the CRLB for 20,000 sets of tissue parameters, randomly selected according to the distributions in Table 4.1. We initialized these CRLB designs in the same way as the 18-scan end-to-end designs, except we replaced one of the random initializations with an optimized end-to-end scan design (so that CRLB design started where the end-to-end design stopped).

4.3.2 MWF Estimation

We compared the MWF estimates of the 18-scan end-to-end design with the lowest MWF estimation error to those of the 18-scan CRLB design with the lowest CRLB of MWF estimates. We computed the MWF estimates using PERK, as in the end-to-end scan design. Unlike in the scan design, in the following experiments we also de-biased the MWF estimates. (De-biasing in the end-to-end scan design was avoided for computational reasons.) To de-bias, we first randomly generated 1,000 sets of tissue parameters according to the distributions in Table 4.1. We then used the scan design to simulate the 18 STFR signals for each set of tissue parameters, repeated 100 times, each time with a different noise realization. Thus there were 100,000 sets of 18 noisy STFR signals each. Each set of STFR signals was given to PERK from which to estimate MWF, resulting in 100,000 MWF estimates total, or 100 MWF estimates for each true MWF value. For each true MWF value, we averaged the associated 100 MWF estimates to obtain the average PERK MWF estimate for that true MWF value. We then fit a line to the 1,000 average MWF estimates to obtain an estimate of the average PERK MWF estimate for any given value of MWF. Then, whenever PERK produced a MWF estimate from test data, we took the estimate and returned the

corresponding estimate of the true MWF value from the line fit above.

We first compared the MWF estimates of the two scan designs in simulation. As was done for the de-biasing process, we randomly generated 1,000 sets of tissue parameters, added 100 noise realizations for each, and then estimated MWF using PERK, resulting in 100,000 MWF estimates for each scan design. We then compared the de-biased estimates to the true MWF values.

We next compared the end-to-end design to the CRLB design using a modified BrainWeb digital phantom [53]. We included WM and GM tissues with the following fixed tissue parameters for WM/GM: $M_0 = 0.77/0.86$, $T_{1,f} = 250/250$ ms, $T_{1,s} = 2000/2000$ ms, $T_{2,f} = 15/15$ ms, $T_{2,s} = 80/80$ ms, $\tau_{f \rightarrow s} = 100/20$ ms, and $\Delta\omega_f = 7/2$ Hz. The flip angle scaling κ was fixed to 1 throughout the digital phantom. We varied f_f from 0.1/0.02 to 0.2/0.05 from image bottom to top for WM/GM, and we varied $\Delta\omega$ from -62.5 to 62.5 Hz from image left to right. We used each scan design to simulate noisy signals for each voxel in the phantom. To mimic differences in scale between acquired data and simulated training data, we scaled the simulated test data by a factor of 3 (that was unknown by the training data). We then added noise to achieve a SNR comparable to that of the *in vivo* data we acquired. After simulating the noisy test data, we then normalized each voxel by the mean STFR signal for that voxel across the 18 scans of the scan design. For training, we tried to add noise to result in a similar training SNR as the test SNR by taking the noise standard deviation that was used for the test data and scaling it by the ratio of the mean test STFR signal (across all voxels and scans) to the mean STFR signal generated from 10,000 randomly selected sets of tissue parameters from the distributions in Table 4.1. We also normalized the noisy training data by the mean STFR signal across STFR scans. Because we ran cross-validation previously without signal normalization, we ran cross-validation again for both scan designs as before, but this time scaling the validation signals by 3 and finding an appropriate training noise standard deviation as just described. We also altered the de-biasing approach to take into account signal normalization. We compared the de-biased MWF estimates to the true MWF map.

Under an IRB-approved protocol, we scanned a healthy volunteer with both the end-to-end scan design and the CRLB scan design. We implemented the protocol in TOPPE [52]. For the implemented STFR scans, $T_{\text{free}} = 8.2$ ms, $T_g = 4.2$ ms, and $T_E = 4.3$ ms. We used 3D acquisitions to avoid slice profile effects. The RF pulses had time-bandwidth product of 4 and duration of 1.5 ms. Each STFR scan took 19.9 s for a total of 5 min 57 s for each scan design. We acquired a $24 \times 24 \times 1.07$ cm³ field of view (FOV) with matrix size $180 \times 180 \times 8$. We used a GE UHP 3T scanner with a 32-channel Nova Medical[®] head coil. We ensured the scanner gain was the same across all 36 STFR scans. We used conventional inverse

Table 4.2: Comparison of RMSEs of MWF estimates from simulated data of optimized 18-scan end-to-end designs with combinations of bulk off-resonance $\Delta\omega$ and flip angle scaling κ being assumed unknown and known parameters. The RMSE is unchanged regardless of whether a separate B0 map ($\Delta\omega$) is acquired, whereas acquiring a separate B1+ map (κ) improved the RMSE by about 13%. In this work, despite the predicted increase in MWF estimation error, we decided to forego acquiring separate B0 and B1+ maps for ease of practical *in vivo* implementation and to lessen the overall scan time.

	$\Delta\omega$ Unknown	$\Delta\omega$ Known
κ Unknown	0.046	0.046
κ Known	0.041	0.041

FFT reconstruction followed by square-root sum-of-squares coil combination to make the magnitude images used for MWF estimation. We analyzed the fifth slice of the end-to-end data and the sixth slice of the CRLB data (there was slight motion between the scans, so we chose the CRLB data slice that visually looked closest to the chosen end-to-end data slice). For PERK, we used the updated hyperparameters we used in the BrainWeb experiment. For the training data noise, we used a noise standard deviation that was approximately half that found by taking the standard deviation of background noise voxels and scaling it as we did in the BrainWeb experiment. To choose the noise standard deviation, we manually inspected the MWF maps resulting from training PERK with different noise standard deviations and chose the one that seemed to give the most visually reasonable MWF map for the end-to-end design. (In the Supporting Information we also estimated MWF training PERK without manually tuning the training noise standard deviation.) We compared the de-biased MWF estimates obtained from both scan designs.

4.4 Results

4.4.1 Scan Designs

All four end-to-end scan designs that we ran to investigate the effects of taking $\Delta\omega$ and κ to be known or unknown ran to local convergence. Table 4.2 shows that the MWF RMSE of these scan designs was virtually unaffected by whether $\Delta\omega$ was known or unknown, whereas the RMSE increased by about 13% (from 0.041 to 0.046) when κ was taken to be unknown compared to when it was known.

All 100 end-to-end scan designs that we ran to investigate how the number of scans influences the RMSE of MWF estimates ran to local convergence. When varying the number of STFR scans in the end-to-end scan design, we found that the best MWF RMSE in each

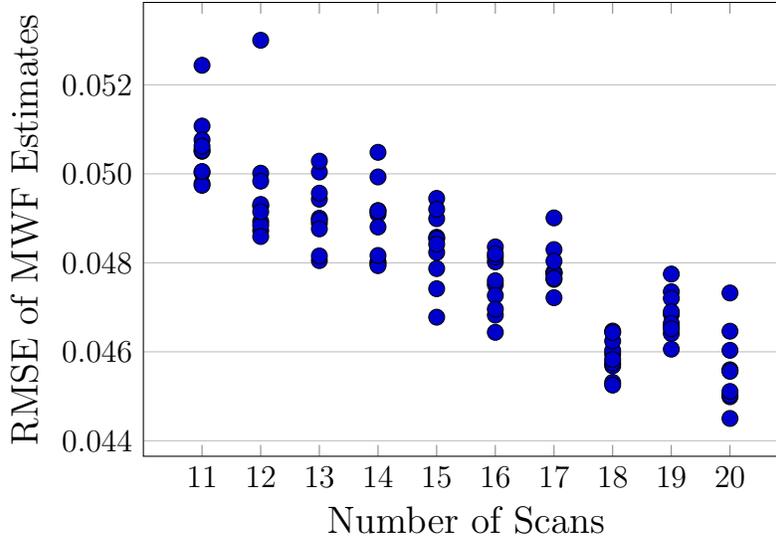


Figure 4.2: Optimized RMSE of MWF estimates from end-to-end scan designs with varying initializations and number of STFR scans. The 10 points in the plot for each given number of scans correspond to the scan designs optimized starting from 10 different initializations. From these results, we decided that using 18 scans was a good trade-off between scan time and performance.

case monotonically decreased (with two exceptions), as seen in Figure 4.2. From these results, we decided that using 18 scans was a good trade-off between scan time and performance.

None of the CRLB scan designs ran to local convergence, though in all cases the optimized designs yielded lower expected CRLBs than their respective initializations. Figure 4.3 compares the 10 18-scan end-to-end designs to the 10 18-scan CRLB designs. The expected CRLBs of 8 of the CRLB designs are understandably much lower than those of the end-to-end designs. Surprisingly, one of the CRLB designs has a much higher expected CRLB than do all the end-to-end designs; this is likely due to a poor initialization and possibly due to not having enough time to reach a local minimum. One CRLB design was initialized with an end-to-end design; the corresponding optimized scan design, though having a (very slightly) lower expected CRLB, remained virtually the same as the initialization, which explains why in both plots of Figure 4.3 there is one CRLB design that matches up with one of the end-to-end designs (the one with which the CRLB design was initialized). On the other hand, all of the end-to-end designs have much lower RMSE of MWF estimates than all of the CRLB designs (with one exception just discussed).

Figure 4.4 shows the scan parameters of the best end-to-end scan design and the best CRLB scan design that were used in the remaining experiments.

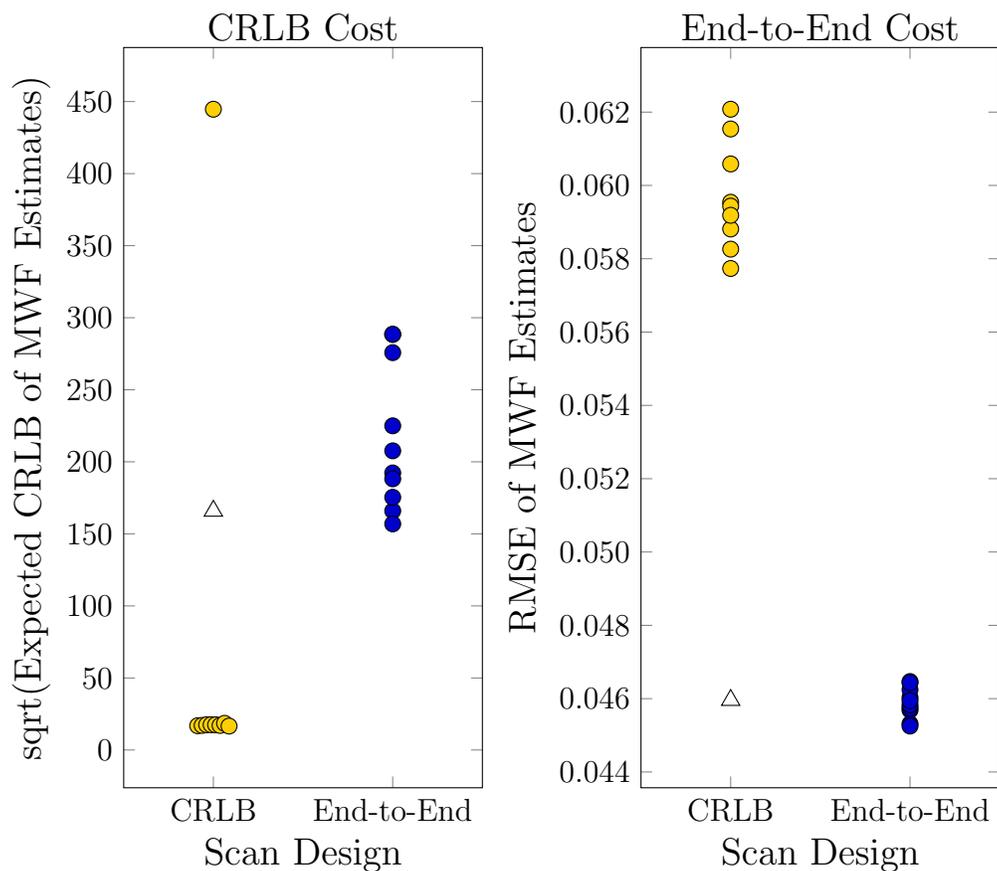


Figure 4.3: Comparison of 10 CRLB scan designs to 10 end-to-end scan designs, each with 18 STFR scans. We computed the square root of the expected CRLB of MWF estimates (left subplot) and the RMSE of MWF estimates (right subplot) for each of the 20 scan designs. Optimizing scan parameters to minimize the RMSE results in scan designs that have lower RMSEs than scan designs optimized to minimize the expected CRLB. (The one CRLB design that has a RMSE on par with the end-to-end designs, indicated by the white triangle marker, was initialized with an end-to-end design and did not change much, i.e., was stuck near a poor local minimum, during the optimization process.)

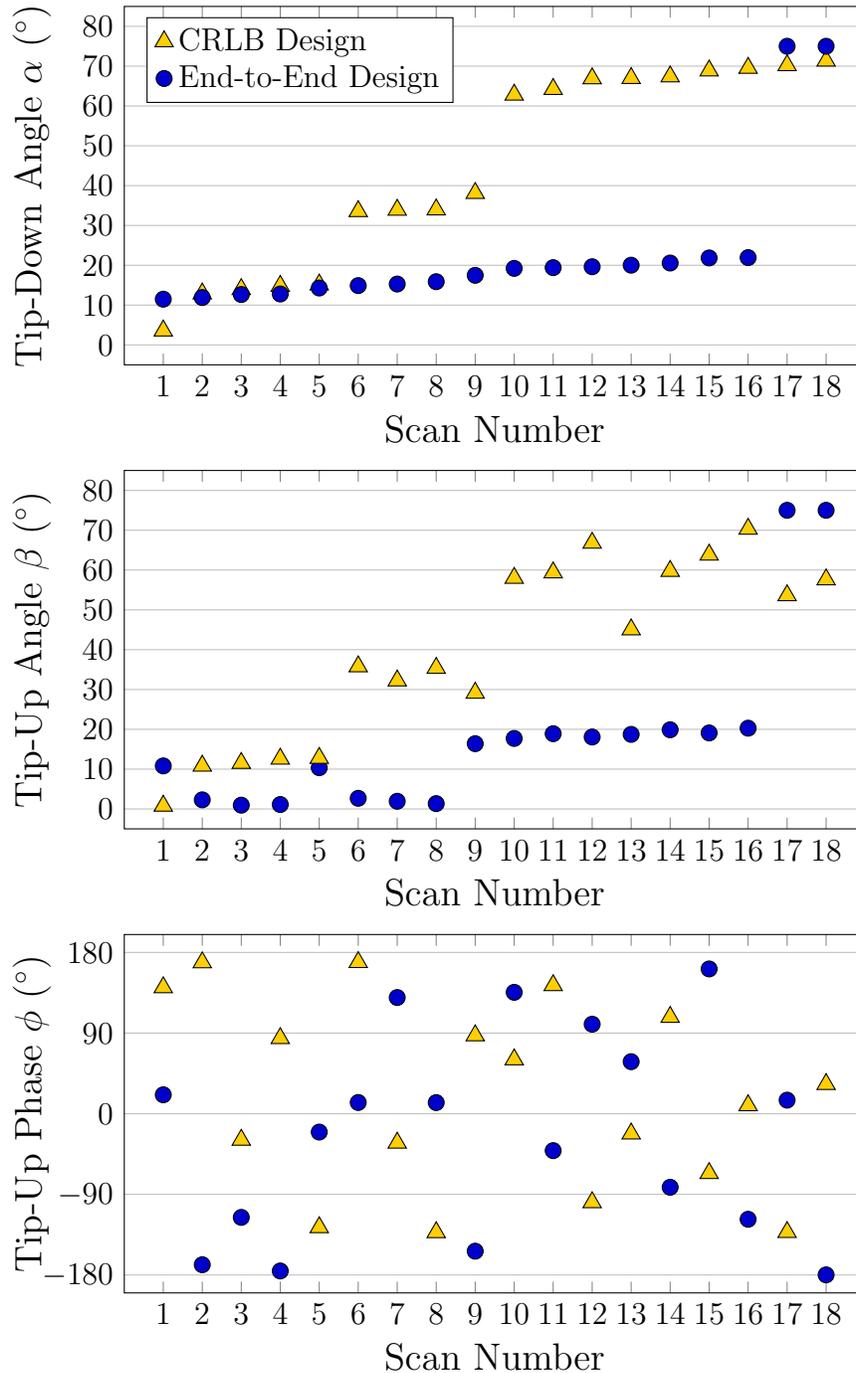


Figure 4.4: Scan parameters of the best end-to-end scan design and the best CRLB scan design, sorted by increasing tip-down angle. Each design consists of 18 STFR scans, each with its own tip-down angle, tip-up angle, and tip-up phase (that were optimized). Each scan additionally had fixed $T_{\text{free}} = 8$ ms, $T_g = 2.8$ ms, and $T_E = 4$ ms. Interestingly, the end-to-end design has much lower flip angles, in general, than the CRLB design.

4.4.2 MWF Estimation

The de-biased MWF estimates obtained using the end-to-end scan design were more accurate, less biased, and less noisy than those obtained using the CRLB design. The end-to-end/CRLB designs' estimates had a RMSE of 0.060/0.080, an average absolute bias of 0.040/0.043, and an average standard deviation of 0.035/0.060. Figure 4.5 shows a 2D histogram of the estimates versus the true MWF values for each scan design.

Figure 4.6 shows the results of the modified BrainWeb experiment. As in the previous experiment, the de-biased MWF estimates obtained using the end-to-end scan design were more accurate, less biased, and less noisy than those obtained using the CRLB design.

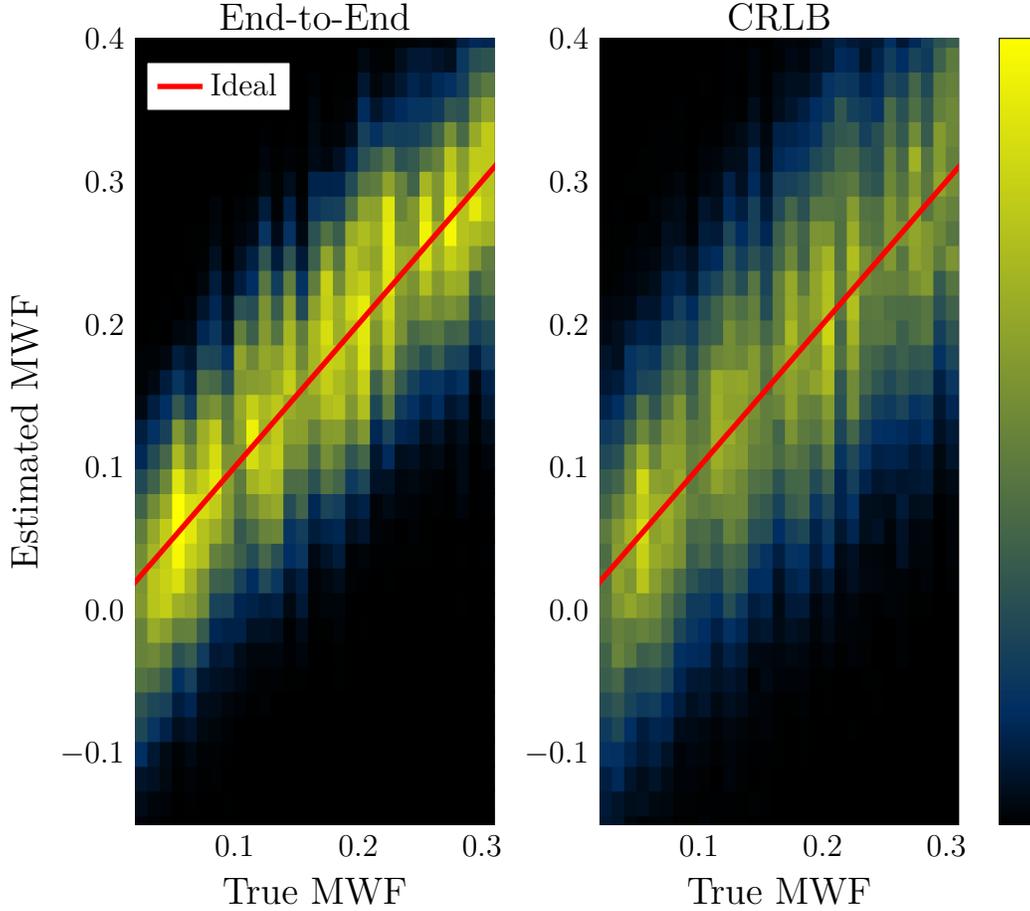
Figure 4.7 shows the de-biased MWF maps estimated from the *in vivo* data. Figure 4.8 shows one of the STFR scans with WM regions of interest (ROIs) shown from which we computed the mean and standard deviation of the MWF estimates. The MWF map from the end-to-end design was less noisy and appeared to have more reasonable values and decent WM/GM contrast.

4.5 Discussion

Even though treating the flip-angle scaling factor κ as a known parameter improved the MWF RMSE of end-to-end scan designs (see Table 4.2), we decided to take it as an unknown parameter to simplify our method and to lessen the overall scan time by avoiding a separate B1+ acquisition (e.g., a pair of Bloch-Siegert scans). Future work could explore this trade-off more in-depth.

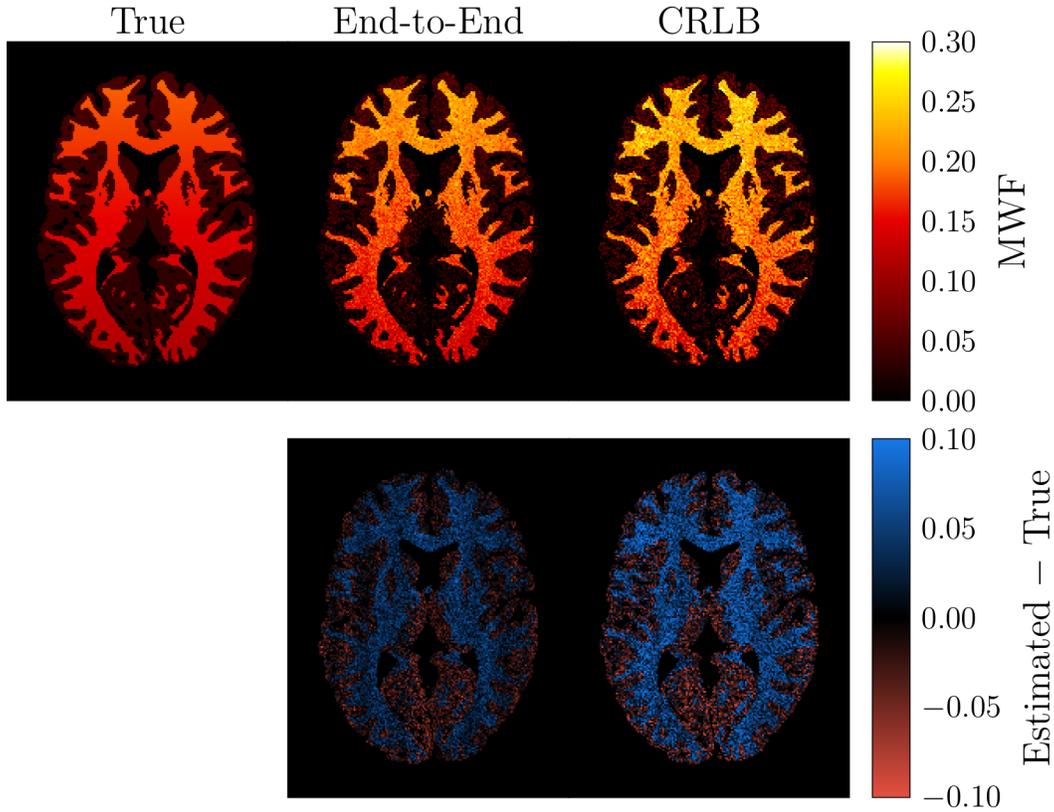
Although Figure 4.2 shows many points (representing the optimized RMSE resulting from different initializations) for each value of D (number of scans), the most relevant RMSE in terms of minimizing the end-to-end cost is the lowest value; it is the value closest to the true global minimum. Intuitively, this value should monotonically decrease as D increases, and for the most part we observe this behavior. Strangely, however, the minimum RMSE for $D = 17$ scans is greater than that for $D = 16$ (and again for $D = 19$ versus $D = 18$). Further investigation is needed to determine why this is the case. One should keep in mind, however, that the end-to-end scan design optimization problem is non-convex, and thus Figure 4.2 does not show the global minimum RMSE for each value of D ; this global minimum, if it could be determined, very likely would decrease monotonically with increasing D as expected.

We note for reference that the 18-scan end-to-end scan designs typically took about 7–8 days to run to convergence on a 20-core, 2.8 GHz CPU with 64 GB memory running Ubuntu 18.04.6. However, we used the default tolerances for determining convergence, which may



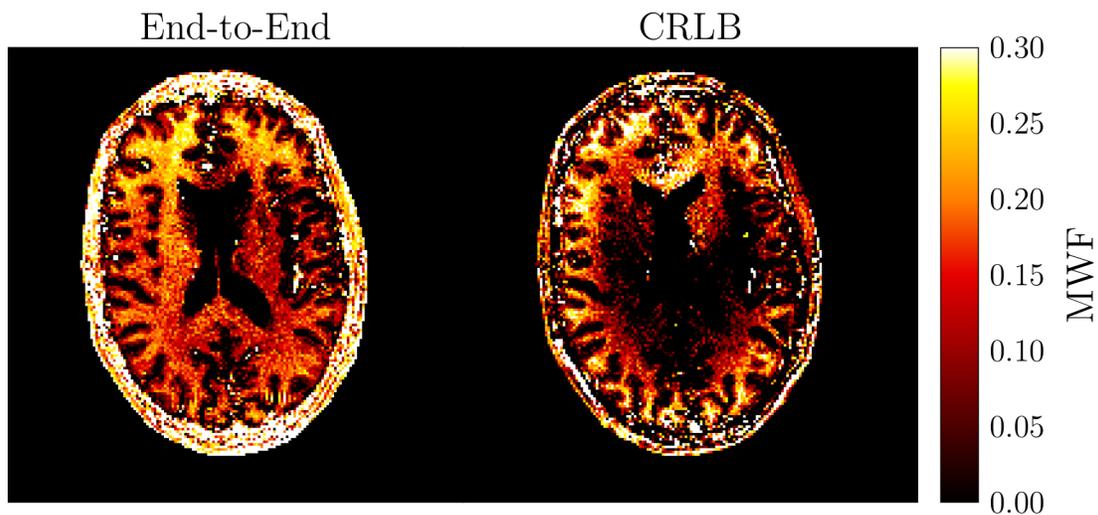
	RMSE	Average Bias	Average SD
End-to-End	0.060	0.040	0.035
CRLB	0.080	0.043	0.060

Figure 4.5: Histograms of estimated MWF versus true MWF obtained when using the end-to-end scan design and when using the CRLB scan design for the experiment where we drew the test data randomly from the distributions in Figure 4.1. Though only MWF is shown as varying, we varied all tissue parameters. The table shows the ensemble RMSE, average absolute bias, and average standard deviation of the MWF estimates. The MWF estimates obtained using the end-to-end scan design were more accurate, less biased, and less noisy than those obtained using the CRLB design.



	RMSE	WM Mean	SD	RMSE	GM Mean	SD
End-to-End	0.035	0.180	0.031	0.024	0.024	0.023
CRLB	0.057	0.200	0.038	0.035	0.023	0.034

Figure 4.6: Estimated MWF maps obtained when using the end-to-end scan design and when using the CRLB scan design. The true MWF varied from 0.1 to 0.2 in WM and from 0.02 to 0.05 in GM from image bottom to top. The results using the end-to-end design were more accurate, less biased, and less noisy than those obtained using the CRLB design. The second row shows the error maps for reference. The table shows the RMSE, mean, and standard deviation (SD) of the MWF estimates across all WM and GM voxels. For reference, the true MWF map has a mean value of 0.151 for WM and 0.035 for GM.



	Mean	SD
End-to-End	0.178	0.059
CRLB	0.109	0.089

Figure 4.7: Estimated MWF maps obtained from *in vivo* data acquired using the end-to-end scan design and the CRLB scan design. The MWF map from the end-to-end design was less noisy and appeared to have more reasonable values and decent WM/GM contrast. The table shows the mean and standard deviation (SD) of the MWF estimates across a pooled region of interest (ROI) consisting of WM voxels (see Figure 4.8).

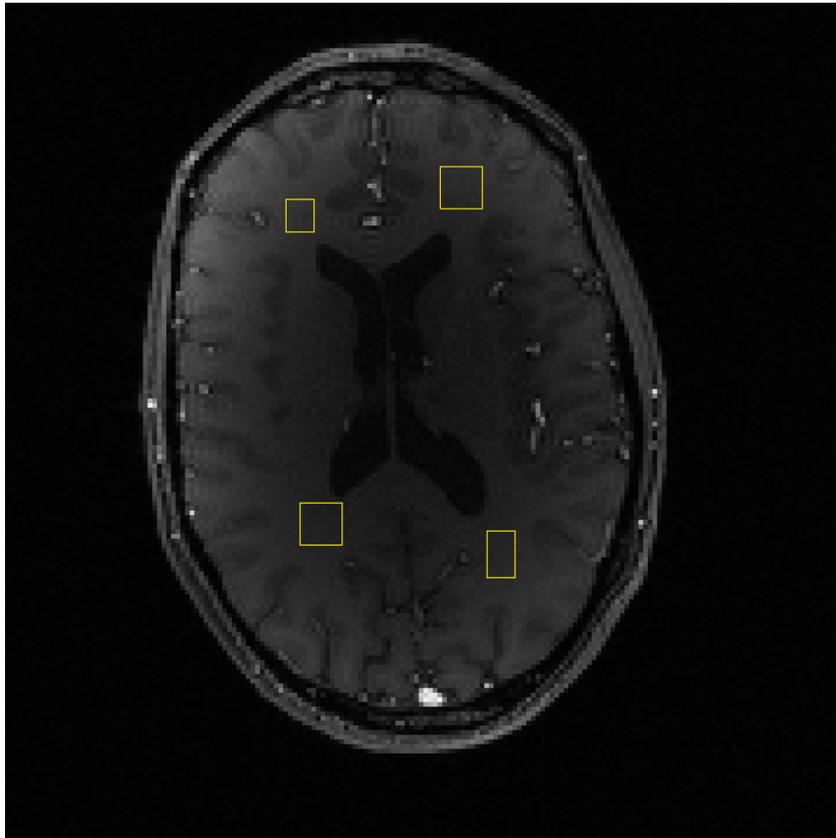


Figure 4.8: One STFR scan with WM ROIs shown from which we computed the mean and standard deviation of the MWF estimates from the *in vivo* data.

have been too stringent; relaxing the tolerances may have led to quicker convergence without noticeably impacting the resultant scan designs. It is possible the CRLB designs also would have converged to local minima (instead of timing out at 16 days) had we used less stringent stopping criteria. However, we find it unlikely (particularly in light of Figure 4.3) that having the CRLB scan designs converge would significantly impact the overall findings of this work.

The results in Figure 4.3 underscore that having a low CRLB does not imply having a low RMSE, at least under the conditions explored in this work. Therefore, if the goal is to have low estimation error then an end-to-end scan design approach is preferable. Future work could extend the analysis here to consider other applications and/or estimators to determine whether the result seen here generalizes.

The simulation results (see Figures 4.5 and 4.6) further illustrate the improved performance of the end-to-end scan design compared to the CRLB design. The MWF estimates resulting from the end-to-end design outperform those resulting from the CRLB design in terms of RMSE, bias, and standard deviation; this holds when considering an ensemble of tissue parameters (Figure 4.5) as well as when considering WM- and GM-like parameters (Figure 4.6).

The *in vivo* MWF maps in Figure 4.7 also demonstrate the potential for improved performance of the end-to-end scan design. The end-to-end design resulted in a less noisy MWF map, and overall it appeared to have more reasonable MWF values. However, we had to manually adjust the noise standard deviation to use for PERK training. Further work is needed to determine a more principled way to choose an appropriate noise standard deviation to use for training PERK to estimate MWF from *in vivo* data.

We also tried using the noise standard deviation computed from background voxels and then scaled as we did for the BrainWeb experiment (this scaled noise standard deviation was about two times that used in Figure 4.7). Figure 4.9 shows the resulting MWF maps. In this case, the MWF maps for both scan designs had higher values than shown in Figure 4.7. However, the end-to-end design still resulted in cleaner MWF estimates. Additionally, the CRLB design's MWF map had spatial variations that appeared to mimic coil sensitivity patterns (i.e., diminished values in the center of the brain, and increasing towards the edges of the brain), whereas the end-to-end design's MWF map did not. Having said that, it is difficult to conclude much about the accuracy of the *in vivo* MWF maps without a ground truth. Work towards developing a well-characterized phantom including multiple exchanging water compartments would be very valuable in this regard.

As illustrated in the previous paragraph, the way PERK is trained can significantly impact the resulting MWF estimates. We have found in unreported experiments that the distributions used for generating tissue parameters for training PERK, as well as the PERK

hyperparameters ρ and λ , can also influence the resulting MWF estimates. Future work could more thoroughly characterize how the way PERK is trained impacts parameter estimates.

Because the end-to-end scan design approach incorporates a specific estimator, in principle one could optimize estimator hyperparameters jointly with scan parameters. Doing so would increase the computational cost of each iteration of optimization but could also result in a better local minimizer, leading to better parameter estimates. PERK has two hyperparameters (ρ and λ), but we chose not to optimize them jointly because we found in unreported experiments that jointly optimizing the hyperparameters only marginally improved the MSE of MWF estimates compared to selecting the hyperparameters via cross-validation after optimizing the scan parameters. Other estimators with more tunable parameters (such as neural networks) potentially could benefit from such joint optimization.

Because the training data depends on \mathbf{P} , PERK must be retrained each time \mathbf{P} is updated during the scan design process. Training PERK is non-iterative and so retraining PERK is not computationally prohibitive. However, using a deep neural network as the estimator could be infeasible due to the computationally expensive training required. Thus, for neural networks, it may be preferable to optimize scan parameters to minimize training error as in [62] rather than optimizing scan parameters to minimize generalization error as proposed in this work.

The end-to-end scan design approach we proposed in this work is similar to the approach in [63] in that scan parameters are optimized to minimize a validation loss (as opposed to a training loss). One major difference, however, is that the loss in [63] is computed over three specific sets of tissue parameters, whereas in this work we compute the loss over a large ensemble of tissue parameters. We are able to do so because we consider collections of individual voxels, whereas an entire image is considered in [63] to enable the modeling of under-sampling artifacts, which are severe in the MRF approach considered in that work. Future work could apply the end-to-end scan design proposed here to MRF and compare this approach to that in [63].

4.6 Conclusion

In this work, we proposed an end-to-end scan design approach and applied it to MWI using STFR scans. We compared an end-to-end scan design to a CRLB design in simulation and *in vivo*. In simulation, we found that the end-to-end design resulted in more accurate, less biased, and less noisy MWF estimates than the CRLB design, both for an ensemble of tissue parameters and for WM- and GM-like parameters. In the *in vivo* case, we likewise saw that the MWF map from the end-to-end design was less noisy and appeared qualitatively more

reasonable than that from the CRLB design. However, there was no gold standard method to which to compare our MWF maps, so it is difficult to make a definitive conclusion about whether the end-to-end scan design worked better than the CRLB design.

The aim of this work was to introduce the end-to-end scan design approach. As such, only a single healthy volunteer was scanned. Further studies are needed to use the proposed method across multiple volunteers, both with and without pathology, to validate the efficacy of the proposed approach.

There are several ways in which the end-to-end design approach could be further explored. For example, one could consider a different cost to minimize rather than MSE. One approach could be to decompose MSE as a sum of variance and bias squared terms and then introduce a tuning parameter to control how much emphasis the optimization places on minimizing bias versus variance. Or, rather than having a tuning parameter, one could minimize the bias of parameter estimates subject to a constraint on the maximum allowed variance in the estimates. Alternatively, one could replace MSE with another metric such as the mean absolute error of parameter estimates.

Future work could also use the end-to-end design approach itself in different ways. In this work, we only explored one estimator (PERK) and one application (MWI with STFR scans). One could use the end-to-end scan design approach to optimize the scan parameters of a different type (or a collection of different types) of scans for estimating one (or more) different parameters using a different estimator.

Finally, in this work we used fully sampled data. Thus another avenue for future work is to acquire under-sampled data and see how parameter estimates look when estimated from the reconstructed under-sampled images even though the end-to-end design did not account for under-sampling effects. Regularized reconstruction techniques would be needed in this case. Methods that reconstruct all D images jointly (such as [55, 56]) would be a natural approach.

4.7 Supporting Information

4.7.1 Choosing PERK Training Distributions for $T_{1,f}$ and $T_{1,s}$

To choose training distributions for $T_{1,f}$ (myelin water T_1) and $T_{1,s}$ (non-myelin water T_1) to use with PERK (and to inform the range of those values to use for the simulated evaluation data during the end-to-end scan design), we relied on the work of Manning et al. [5]. Manning et al. developed a four-pool model for describing white matter (WM) longitudinal relaxation. The pools included both water and macromolecules for both myelin and non-

myelin. Relative pool sizes, T_1 values for each pool, and exchange times between the pools were estimated by fitting the model for several tissue samples.

To simplify the model from four to two pools to use in our work, we combined the water and macromolecule pools associated with myelin and non-myelin by determining an apparent T_1 for myelin water and non-myelin water, resulting in one myelin water pool and one non-myelin water pool with T_1 values that reflect contributions from both water and macromolecules. We determined the combined T_1 values by computing the eigenvalues of the Bloch-McConnell equation for longitudinal relaxation [18]:

$$\begin{bmatrix} \frac{dM_{z,m}}{dt} \\ \frac{dM_{z,w}}{dt} \end{bmatrix} = \begin{bmatrix} -R_m - k_{m \rightarrow w} & k_{w \rightarrow m} \\ k_{m \rightarrow w} & -R_w - k_{w \rightarrow m} \end{bmatrix} \begin{bmatrix} M_{z,m} \\ M_{z,w} \end{bmatrix} + \begin{bmatrix} M_{0,m}R_m \\ M_{0,w}R_w \end{bmatrix}, \quad (4.4)$$

where the subscripts “m” and “w” respectively stand for “macromolecules” and “water”, $M_{z,m/w}$ is the longitudinal magnetization, $R_{m/w}$ is the longitudinal relaxation rate, $k_{m/w \rightarrow w/m}$ is the exchange rate from one pool to the other, and $M_{0,m/w}$ is the equilibrium magnetization. The eigenvalues² are given by [68]

$$\lambda_{f/s} = \frac{1}{2}(R_m + R_w + k_{m \rightarrow w} + k_{w \rightarrow m}) \pm \frac{1}{2}\sqrt{(R_m - R_w + k_{m \rightarrow w} - k_{w \rightarrow m})^2 + 4k_{m \rightarrow w}k_{w \rightarrow m}}. \quad (4.5)$$

The combined T_1 is then given by $1/\lambda_s$. As there are two pairs of water and macromolecule pools to combine, we used (4.5) twice: once to obtain $T_{1,f}$ by combining myelin water and myelin macromolecules, and again to obtain $T_{1,s}$ by combining non-myelin water and non-myelin macromolecules.

Manning et al. estimated model parameters for three WM tissue samples. We computed $T_{1,f}$ and $T_{1,s}$ using those parameters for each tissue sample. Values for R_m , R_w , $k_{m \rightarrow w}$, and $k_{w \rightarrow m}$ were determined using values in Table 1 in [5]. For the myelin water pool, $R_m = 1/T_{1,M}$, $R_w = 1/T_{1,MW}$, $k_{m \rightarrow w} = k_{w \rightarrow m}M_{MW}(\infty)/M_M(\infty)$, and $k_{w \rightarrow m} = M_M(\infty)/(T_{cr,M/MW}(M_{MW}(\infty) + 1))$. Similar expressions hold for the non-myelin water pool. We also investigated accounting for the B0 dependence of R_m assuming an inversely proportional relationship [68], in which case R_m was scaled by $\frac{4.7 \text{ T (field strength in [5])}}{3.0 \text{ T (our field strength)}}$. Table 4.3 shows the resulting $T_{1,f}$ and $T_{1,s}$ values. Based on this range of values, we trained PERK using $T_{1,f}$ values uniformly distributed from 200 to 500 ms and $T_{1,s}$ values uniformly distributed from 1500 to 2500 ms.

²Technically, (4.5) gives the negated eigenvalues (to result in positive values). The original eigenvalues are negative because they describe exponential decay.

Table 4.3: Computed myelin water and non-myelin water T_1 values ($T_{1,f}$ and $T_{1,s}$, respectively) determined by combining the water and macromolecule pools associated with myelin and non-myelin, respectively, using (4.5). Values for R_m , R_w , $k_{m \rightarrow w}$, and $k_{w \rightarrow m}$ were determined using values in Table 1 in [5]. We also computed T_1 values $T_{1,f}^a$ and $T_{1,s}^a$ with an adjusted R_m to account for field strength, in which case R_m was scaled by $\frac{4.7 \text{ T}}{3.0 \text{ T}}$. Units for all table values are s^{-1} except for $T_{1,f/s}^{(a)}$, which have units of ms. Based on this range of values for $T_{1,f}$ and $T_{1,s}$, we trained PERK using $T_{1,f}$ values uniformly distributed from 200 to 500 ms and $T_{1,s}$ values uniformly distributed from 1500 to 2500 ms.

Myelin		R_m	R_w	$k_{m \rightarrow w}$	$k_{w \rightarrow m}$	λ_s	$T_{1,f}$	λ_s^a	$T_{1,f}^a$
	Sample WM-fr		4.3	0.33	12	21	2.7	370	4.1
Sample WM-sp1		5.3	0.42	9.9	17	3.3	310	4.8	210
Sample WM-sp2		6.7	0.45	11	15	3.6	280	5.2	190
Non-Myelin		R_m	R_w	$k_{m \rightarrow w}$	$k_{w \rightarrow m}$	λ_s	$T_{1,s}$	λ_s^a	$T_{1,f}^a$
	Sample WM-fr		1.6	0.33	10	1.8	0.50	2000	0.61
Sample WM-sp1		1.0	0.42	13	1.3	0.47	2100	0.51	1900
Sample WM-sp2		0.40	0.45	12	1.3	0.44	2300	0.46	2200

We also determined the overall apparent T_1 by applying (4.5) again, this time with the myelin water pool (subscript “m”) and the non-myelin water pool (subscript “w”). In this case, we let $R_{m/w} = 1/T_{1,f/s}$, $k_{m \rightarrow w} \in \{13, 21, 18\} \text{ s}^{-1}$ (each entry corresponding to one of the tissue samples in Table 4.3 and computed from values in Table 1 of [5]), and $k_{w \rightarrow m} = f_f k_{m \rightarrow w} / (1 - f_f)$, where f_f is the myelin water fraction (MWF). For f_f , we used values of 0.12, 0.078, and 0.10 (derived from [5] for each tissue sample) as well as 0.15 and 0.20. The range of overall apparent T_1 values was about 990 to 1500 ms, which is in line with the largest eigenvalues (i.e., the overall apparent T_1 computed using a four-pool version of (4.5)) reported in Table 2 in [5]. When accounting for field strength, the range of apparent T_1 values was about 790 to 1300 ms, which is consistent with [68] (see λ_s in Table 2 therein).

4.7.2 Comparison of Initial and Optimized Scan Designs

Here we compare the CRLB and RMSE of MWF estimates of the optimized 18-scan CRLB and end-to-end scan designs to those of their corresponding initializations. Table 4.4 shows the comparison, revealing several interesting points. First, the CRLB designs resulted in significantly improved CRLBs than the initial values (the one exception being the CRLB design that was initialized with an end-to-end design); despite this, the RMSEs were much more mildly affected. Second, the end-to-end designs resulted in significantly improved RMSEs than the initial values despite resulting in drastically worse CRLBs in nearly all

cases. Third, the hand-crafted initialization (described in Section 4.3.1) had a CRLB about 40 times worse than the average random initialization, and yet it had the best RMSE. Finally, despite having the best initial RMSE, the hand-crafted initialization actually resulted in the *worst* end-to-end design. The best end-to-end design had a RMSE 13% better than the hand-crafted design.

Table 4.4 (a) reports CRLB values that are much higher than a nominal MWF value of 0.15. However, the results in this table are expected CRLB values, where the expectation is taken over quite a large range of parameter values. Computing the CRLB using literature WM tissue parameters lowers the CRLB of the best CRLB design from 16.6 to 2.8. While the CRLB still is huge compared to a nominal MWF value, Figure 3.11 (though computed with a tissue model without exchange) indicates that bias is likely responsible for PERK still obtaining MWF estimates with low variance.

4.7.3 *In Vivo* MWF Map Using Computed Noise

Here we present MWF maps estimated from *in vivo* data acquired using the end-to-end scan design and the CRLB scan design. We followed the same PERK training procedure as described in the main text, but instead of manually adjusting the noise standard deviation to use for the training data, we computed the standard deviation of background noise voxels from the acquired data and scaled it by the ratio of the mean acquired STFR signal (across all voxels and scans, pooling the data for both scan designs) to the mean STFR signal generated from 10,000 randomly selected sets of tissue parameters from the distributions in Table 4.1 using all 36 scans from both scan designs. We used this scaled standard deviation as the training noise standard deviation for training PERK for both scan designs. Figure 4.9 shows the resulting maps. In this case, the MWF maps for both scan designs had higher values than shown in Figure 4.7. However, the end-to-end design still resulted in cleaner MWF estimates. Additionally, the CRLB design’s MWF map had spatial variations that appeared to mimic coil sensitivity patterns (i.e., diminished values in the center of the brain, and increasing towards the edges of the brain), whereas the end-to-end design’s MWF map did not.

4.7.4 Noise Cross-Validation

Figures 4.7 and 4.9 show the MWF estimates resulting from using two different noise standard deviations for training PERK, and the resulting maps are quite different. We thus investigated a way to choose the noise standard deviation via cross-validation. We swept different values of noise standard deviation in the set $\{2^{-15}, 2^{-14.9}, \dots, 2^{-5}\}$. For each value,

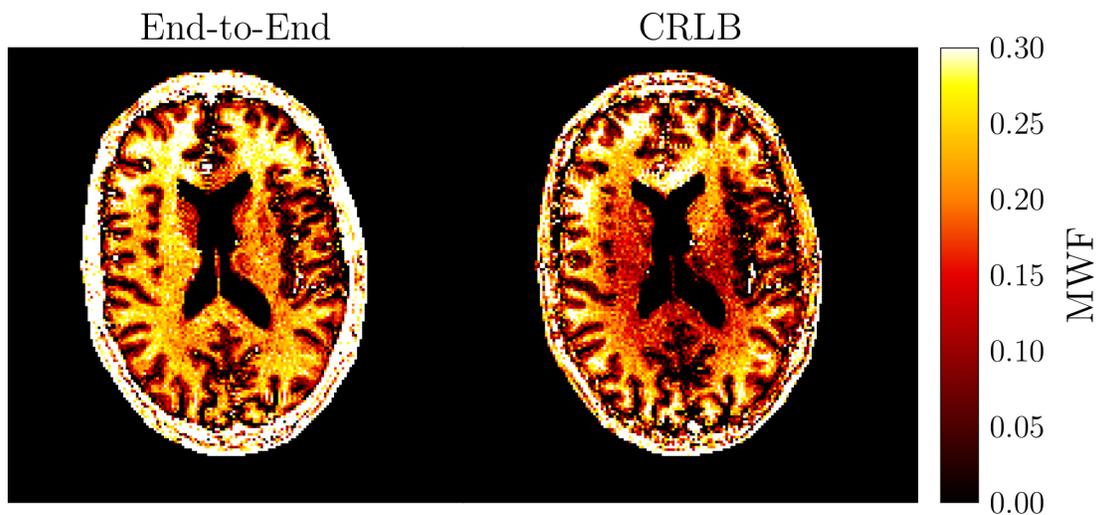
Table 4.4: CRLB (a) and RMSE (b) of MWF estimates computed using the optimized 18-scan CRLB and end-to-end scan designs as well as the scan designs used for initialization. The end-to-end design used to initialize a CRLB scan design was initialized with the hand-crafted initialization. (The optimized RMSE of the end-to-end design initialized with the hand-crafted design is different than the initial RMSE for the CRLB design initialized with the end-to-end design due to slightly different PERK hyperparameter values.) The displayed percentages indicate the change relative to the initial value. Despite resulting in drastically worse CRLBs compared to the initial values in nearly all cases, the end-to-end designs significantly improved the RMSEs in all cases.

(a) sqrt(Expected CRLB of MWF Estimates)

Initialization	Initial	CRLB Design	End-to-End Design
Hand-Crafted	4710	445 (−91%)	166 (−96%)
Random 1	102	16.8 (−83%)	276 (+171%)
Random 2	105	17.1 (−84%)	208 (+97%)
Random 3	118	17.5 (−85%)	157 (+33%)
Random 4	72.6	N/A	192 (+165%)
Random 5	96.6	17.5 (−82%)	288 (+198%)
Random 6	121	17.5 (−86%)	188 (+55%)
Random 7	98.9	17.1 (−83%)	225 (+127%)
Random 8	134	18.5 (−86%)	289 (+116%)
Random 9	79.3	16.6 (−79%)	175 (+121%)
End-to-End	166	166 (−0%)	N/A

(b) RMSE of MWF Estimates

Initialization	Initial	CRLB Design	End-to-End Design
Hand-Crafted	0.0519	0.0595 (+15%)	0.0465 (−10%)
Random 1	0.0599	0.0577 (−3.5%)	0.0457 (−24%)
Random 2	0.0604	0.0583 (−3.5%)	0.0453 (−25%)
Random 3	0.0597	0.0621 (+3.9%)	0.0460 (−23%)
Random 4	0.0576	N/A	0.0462 (−20%)
Random 5	0.0605	0.0606 (+0.1%)	0.0457 (−24%)
Random 6	0.0612	0.0588 (−4.0%)	0.0458 (−25%)
Random 7	0.0571	0.0594 (+4.2%)	0.0453 (−21%)
Random 8	0.0596	0.0615 (+3.3%)	0.0464 (−22%)
Random 9	0.0593	0.0592 (−0.3%)	0.0459 (−23%)
End-to-End	0.0462	0.0460 (−0.6%)	N/A



	Mean	SD
End-to-End	0.249	0.045
CRLB	0.201	0.056

Figure 4.9: Estimated MWF maps obtained from *in vivo* data acquired using the end-to-end scan design and the CRLB scan design, with PERK trained with a noise standard deviation that was computed from background voxels of the reconstructed images. The MWF map from the end-to-end design had higher values than what one might expect in the brain, but was less noisy than the MWF map from the CRLB design. The table shows the mean and standard deviation (SD) of the MWF estimates across a pooled region of interest (ROI) consisting of WM voxels (see Figure 4.8).

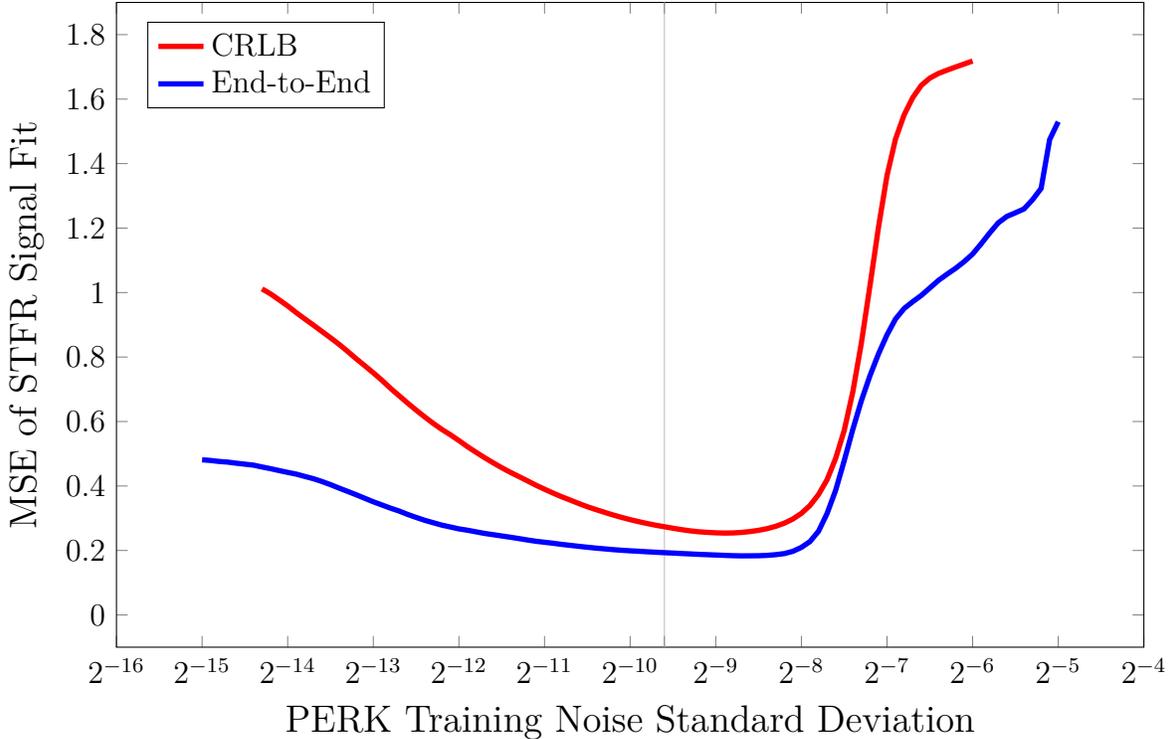


Figure 4.10: MSE of fitted STFR signals versus PERK training noise standard deviation for the modified BrainWeb experiment. The noise standard deviation added to the BrainWeb data was $2^{-9.6}$ (denoted by the vertical line). The cross-validated noise standard deviation performed well in this case.

we trained PERK as usual and then estimated tissue parameters from several voxels in WM ROIs. After clamping the parameter estimates to physically reasonable values, we used them to simulate the STFR signals for each voxel, normalizing the signals on a per-voxel basis as done previously. We then computed the error between the predicted STFR signals and the actual STFR signals for each voxel. We used the noise standard deviation resulting in the lowest error to train PERK for estimating MWF over the whole object.

For the modified BrainWeb experiment, we found that this cross-validation process chose a training noise standard deviation that resulted in MWF estimates that were just as good as those in Figure 4.6. Figure 4.10 shows the signal fitting MSE versus training noise standard deviation.

For the *in vivo* data, however, the cross-validation chose a training noise standard deviation that did not result in (visually) good MWF estimates; the resulting maps were closer to those in Figure 4.9 (and actually had a bit higher MWF values) than to those in Figure 4.7. Figure 4.11 shows the *in vivo* signal fitting MSE versus training noise standard deviation.

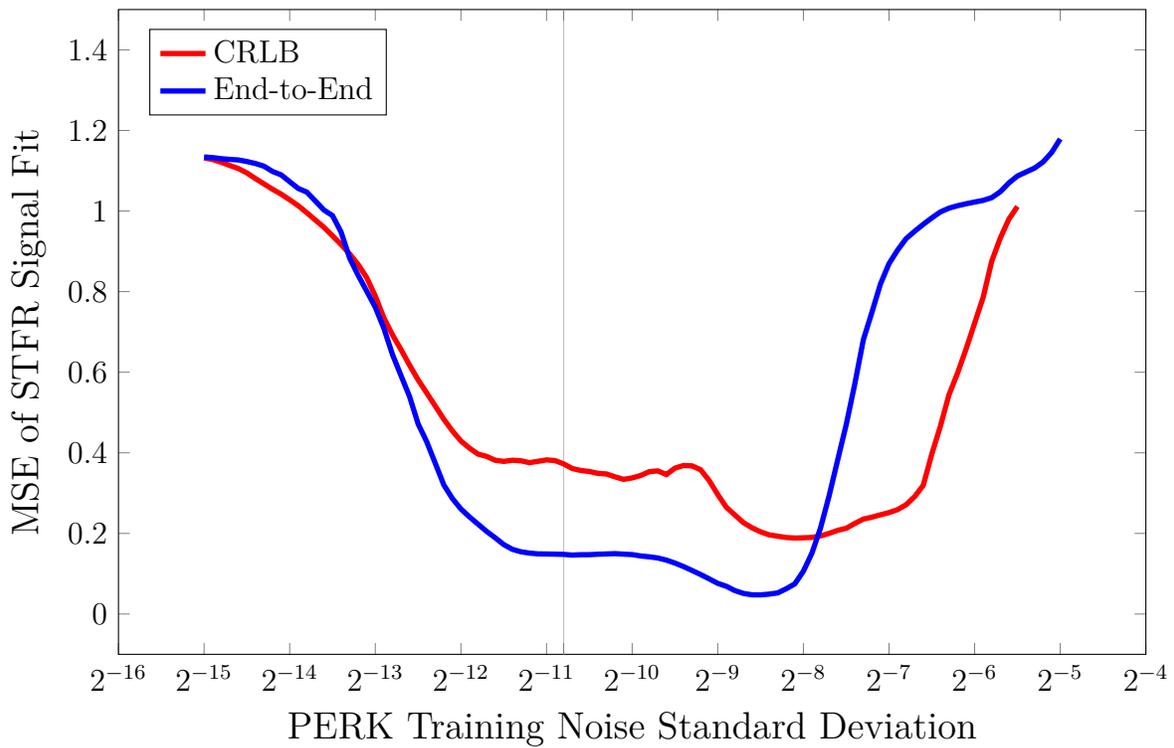


Figure 4.11: MSE of fitted STFR signals versus PERK training noise standard deviation for the *in vivo* data. The training noise standard deviation we manually chose based on visual preference (used for Figure 4.7) was $2^{-10.8}$ (denoted by the vertical line). The cross-validated noise standard deviation did not perform well in this case.

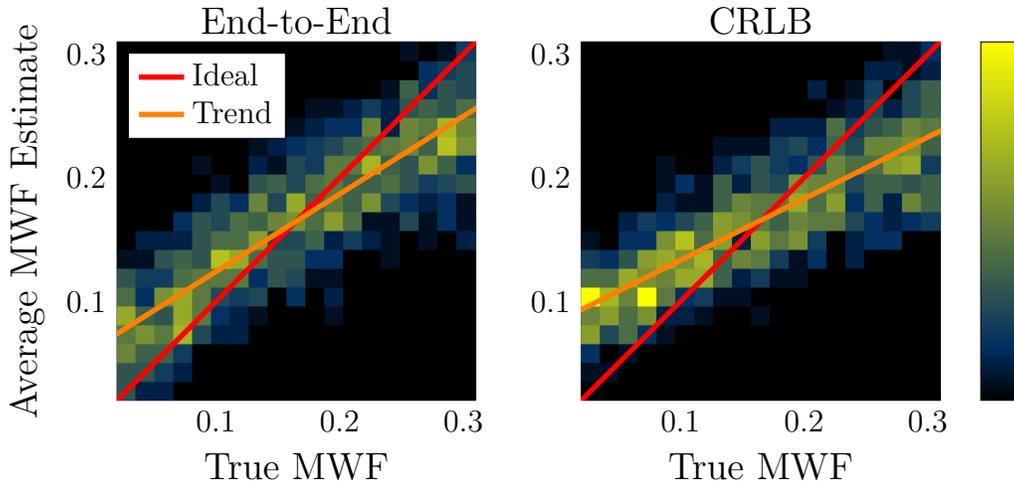


Figure 4.12: Trend lines used for de-biasing MWF estimates. We estimated MWF using PERK for 1,000 sets of non-scan parameters and averaged over 100 noise realizations. Here we show 2D histograms of the average estimates versus the true MWF values. We then fit a trend line to the scatter plot underlying each histogram. This trend line was used in de-biasing MWF estimates; when an estimate was given by PERK, we would look up the corresponding true MWF value predicted by the trend line and return that MWF value as the de-biased estimate.

4.8 Appendix

This appendix presents additional results not included in the main body of the chapter.

4.8.1 De-biasing

Figure 4.12 shows a 2D histogram of the MWF estimates and fitted trend line used for de-biasing the MWF estimates in the chapter for both the end-to-end scan design and the CRLB design. To de-bias, when an estimate was given by PERK, we would look up the corresponding true MWF value predicted by the trend line and return that MWF value as the de-biased estimate. From the trend lines, we observe that the end-to-end design appears to be more informative of MWF than the CRLB design because its trend line is closer to the ideal case. (A scan design with no information about MWF would result in a horizontal trend line, i.e., would regress to the mean training MWF value.)

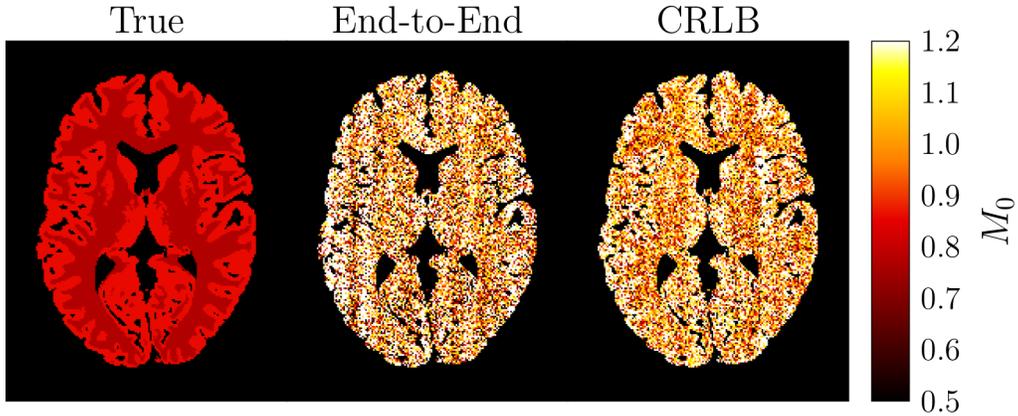


Figure 4.13: M_0 BrainWeb estimates. The true M_0 values are shown before scaling by 3.

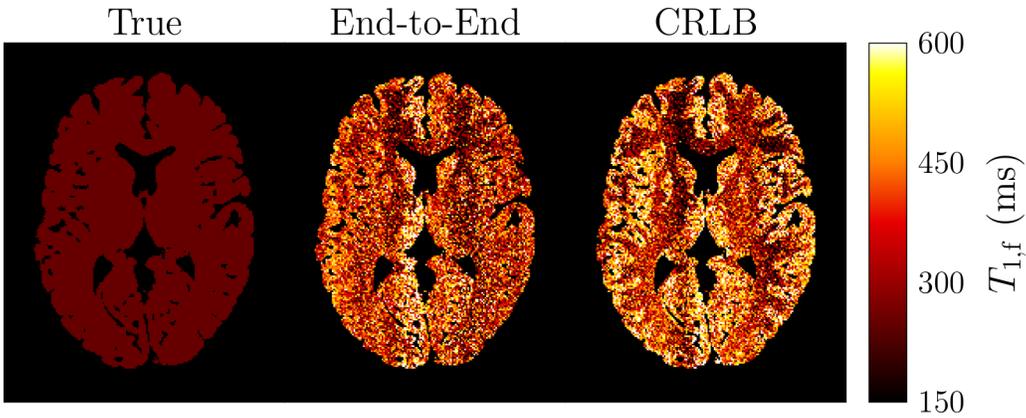


Figure 4.14: $T_{1,f}$ BrainWeb estimates.

4.8.2 Other BrainWeb Parameter Estimates

Here we show the other parameter estimates obtained with PERK for the modified BrainWeb digital phantom experiment for both the end-to-end scan design and the CRLB design. Note that the scan designs were optimized for just MWF, so we do not have any expectation as to how good the other parameter estimates should be. Figure 4.13 shows M_0 , Figure 4.14 shows $T_{1,f}$, Figure 4.15 shows $T_{1,s}$, Figure 4.16 shows $T_{2,f}$, Figure 4.17 shows $T_{2,s}$, Figure 4.18 shows $\tau_{f \rightarrow s}$, Figure 4.19 shows $\Delta\omega_f$, Figure 4.20 shows $\Delta\omega$, and Figure 4.21 shows κ .

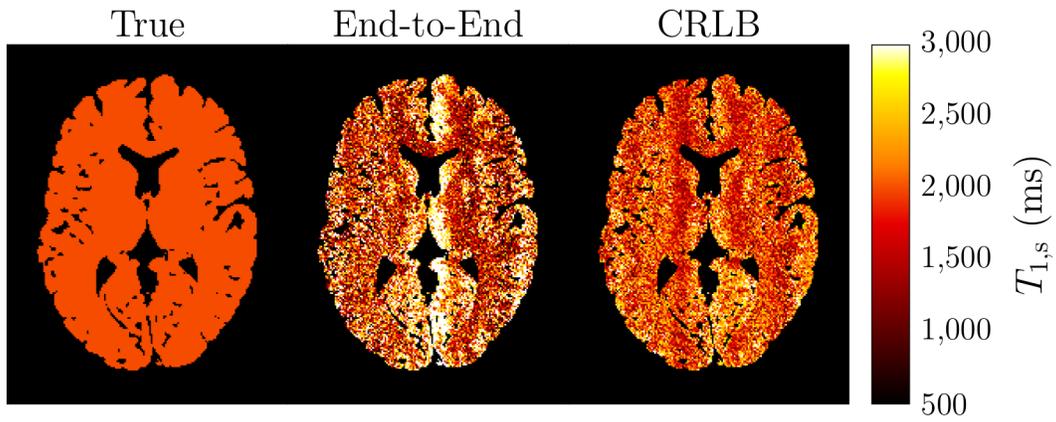


Figure 4.15: $T_{1,s}$ BrainWeb estimates.

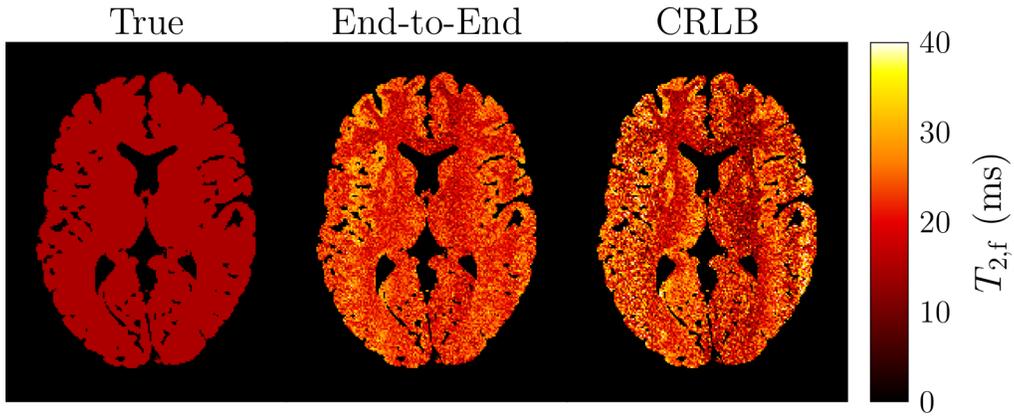


Figure 4.16: $T_{2,f}$ BrainWeb estimates.

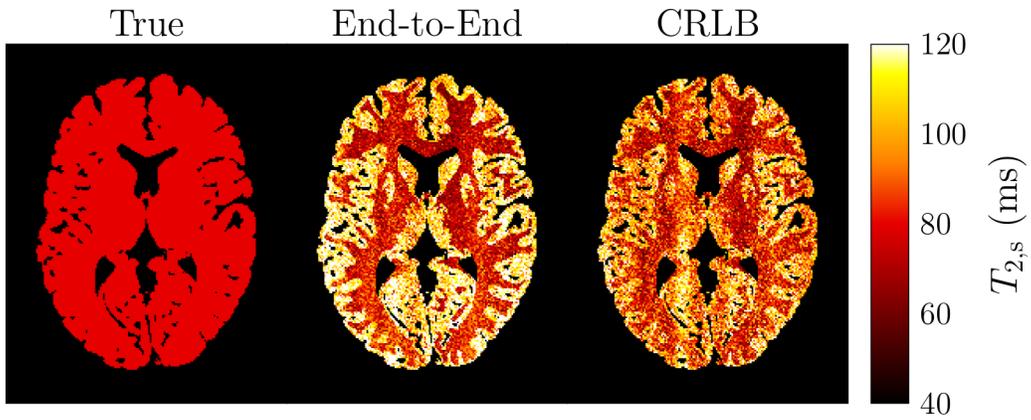


Figure 4.17: $T_{2,s}$ BrainWeb estimates.

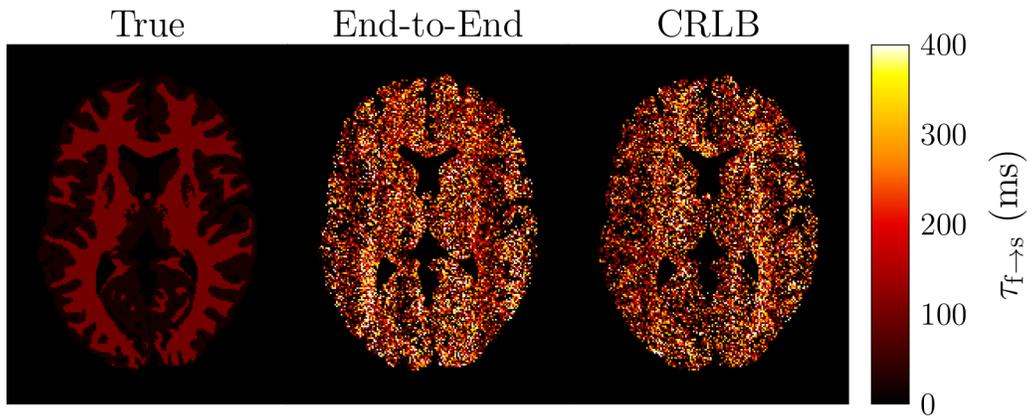


Figure 4.18: $\tau_{f \rightarrow s}$ BrainWeb estimates.

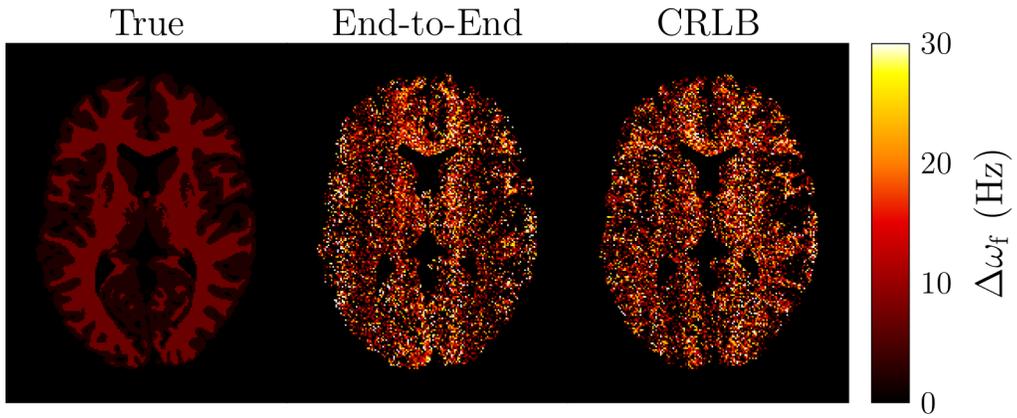


Figure 4.19: $\Delta\omega_f$ BrainWeb estimates.

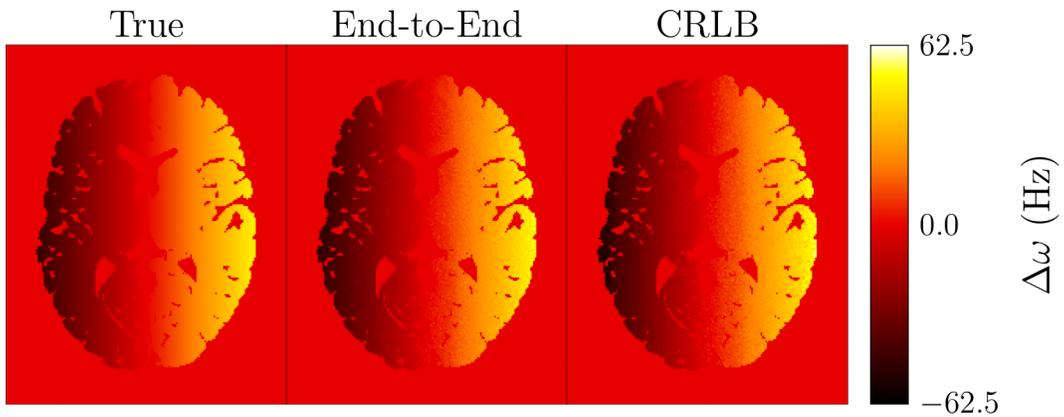


Figure 4.20: $\Delta\omega$ BrainWeb estimates.

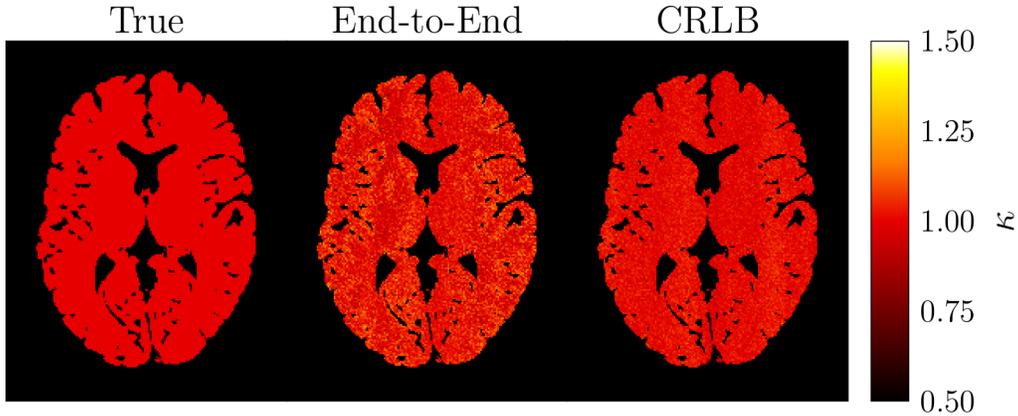


Figure 4.21: κ BrainWeb estimates.

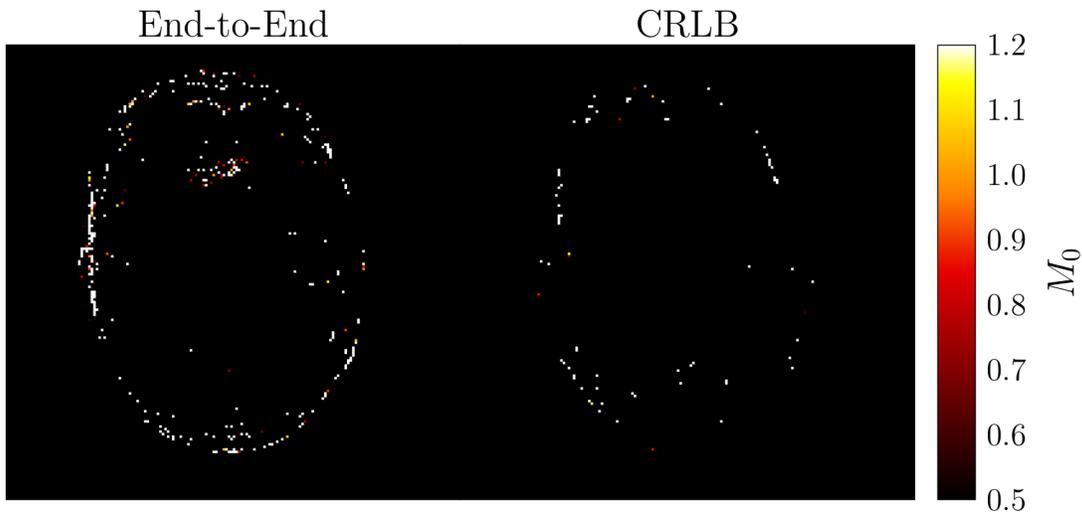


Figure 4.22: M_0 *in vivo* estimates.

4.8.3 Other *In Vivo* Parameter Estimates

Here we show the other parameter estimates obtained with PERK from the *in vivo* data for both the end-to-end scan design and the CRLB design. These estimates correspond to Figure 4.7 (i.e., we trained PERK with a manually chosen noise level). Figure 4.22 shows M_0 , Figure 4.23 shows $T_{1,f}$, Figure 4.24 shows $T_{1,s}$, Figure 4.25 shows $T_{2,f}$, Figure 4.26 shows $T_{2,s}$, Figure 4.27 shows $\tau_{f \rightarrow s}$, Figure 4.28 shows $\Delta\omega_f$, Figure 4.29 shows $\Delta\omega$, and Figure 4.30 shows κ . Of these parameter estimates, only $\Delta\omega$ appears reasonable, as verified with the independently measured field map. All other parameter estimates are too high or low compared to values typically reported in literature. We again note that the scan designs were optimized for just MWF. Despite the poor parameter estimates for most of the other parameters, the MWF estimates from the end-to-end design still seem reasonable.

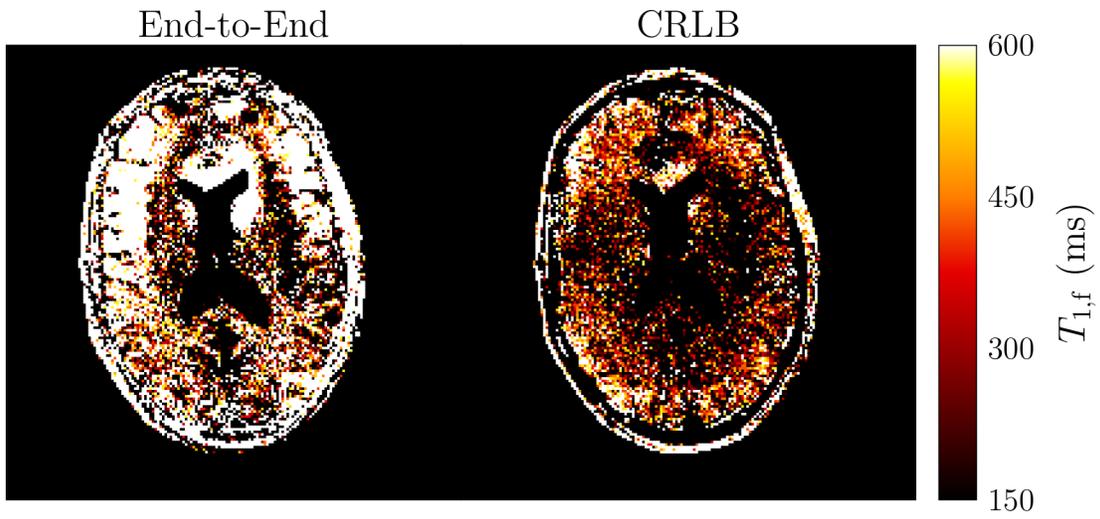


Figure 4.23: $T_{1,f}$ *in vivo* estimates.

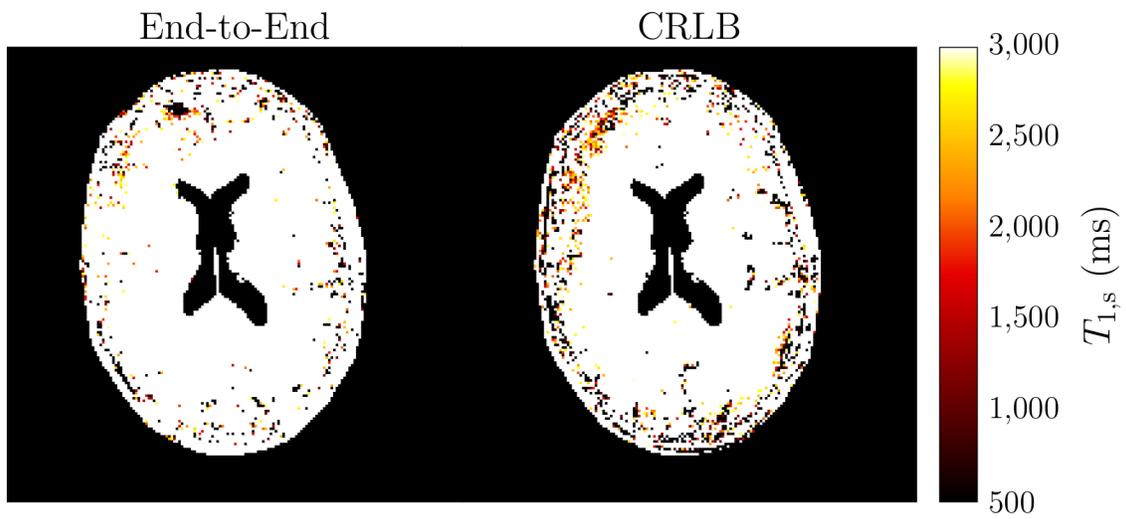


Figure 4.24: $T_{1,s}$ *in vivo* estimates.

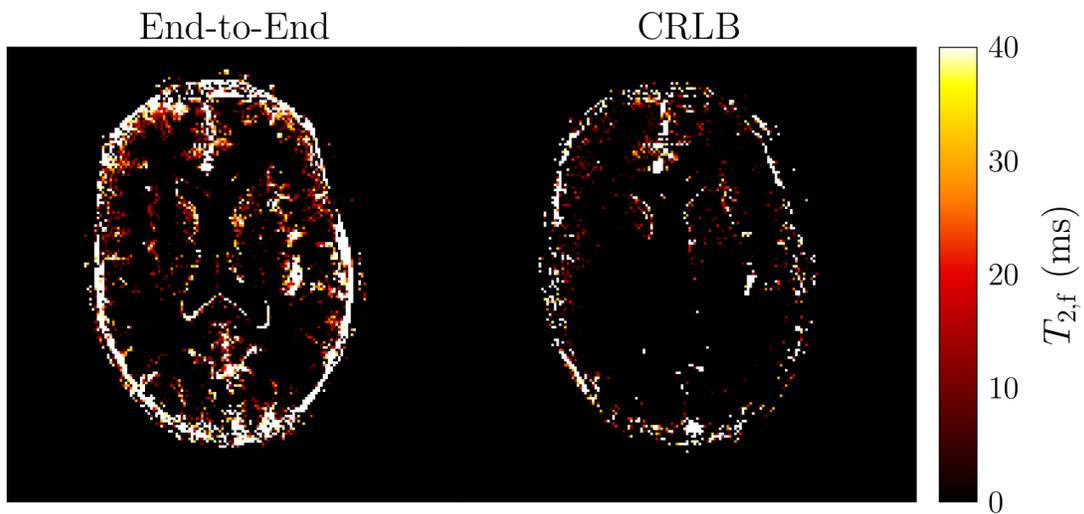


Figure 4.25: $T_{2,f}$ *in vivo* estimates.

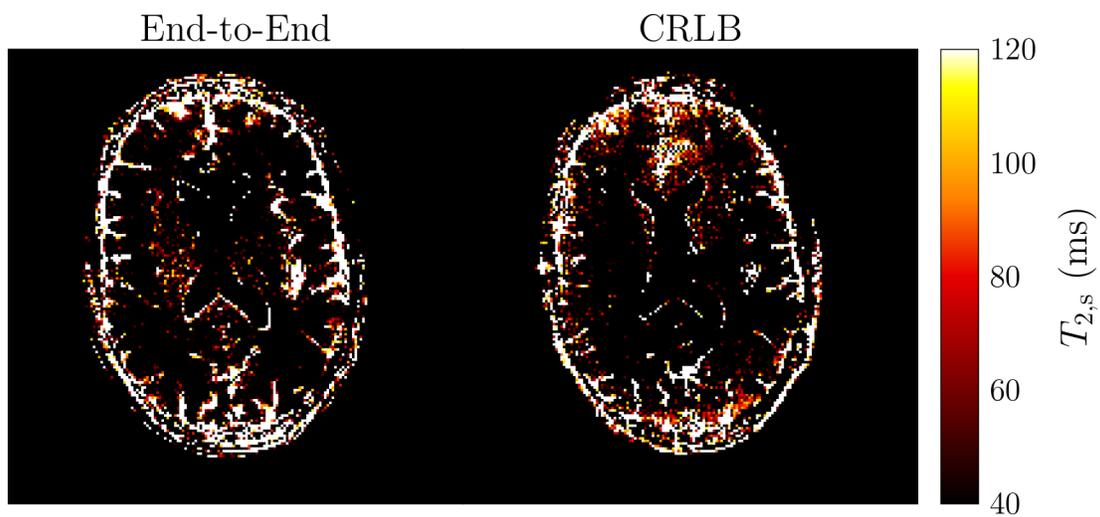


Figure 4.26: $T_{2,s}$ *in vivo* estimates.

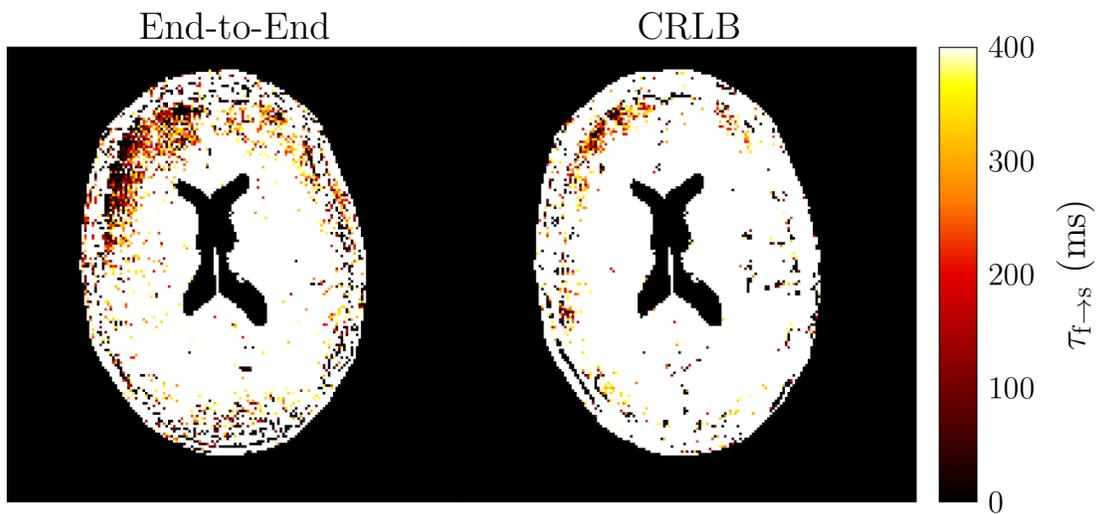


Figure 4.27: $\tau_{f \rightarrow s}$ *in vivo* estimates.

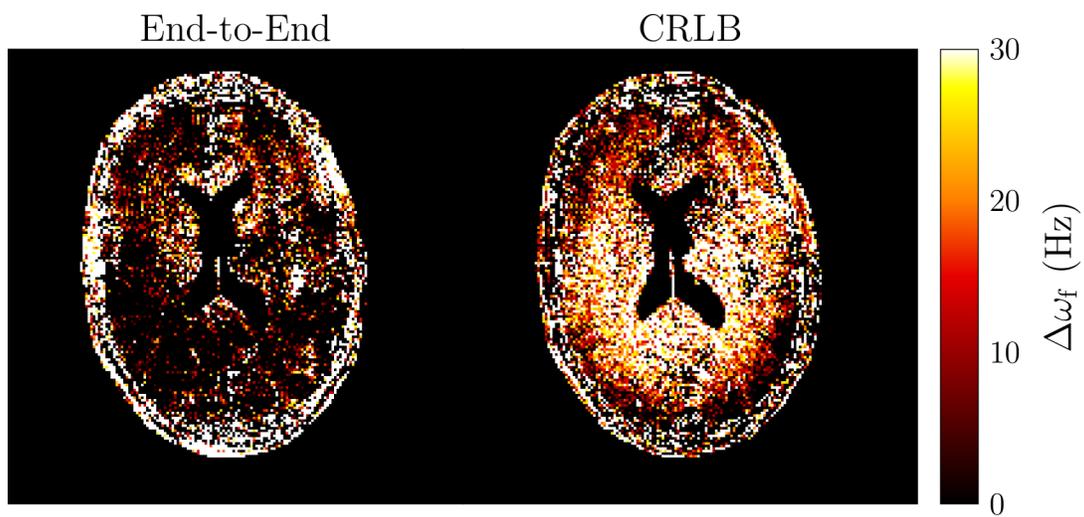


Figure 4.28: $\Delta\omega_f$ *in vivo* estimates.

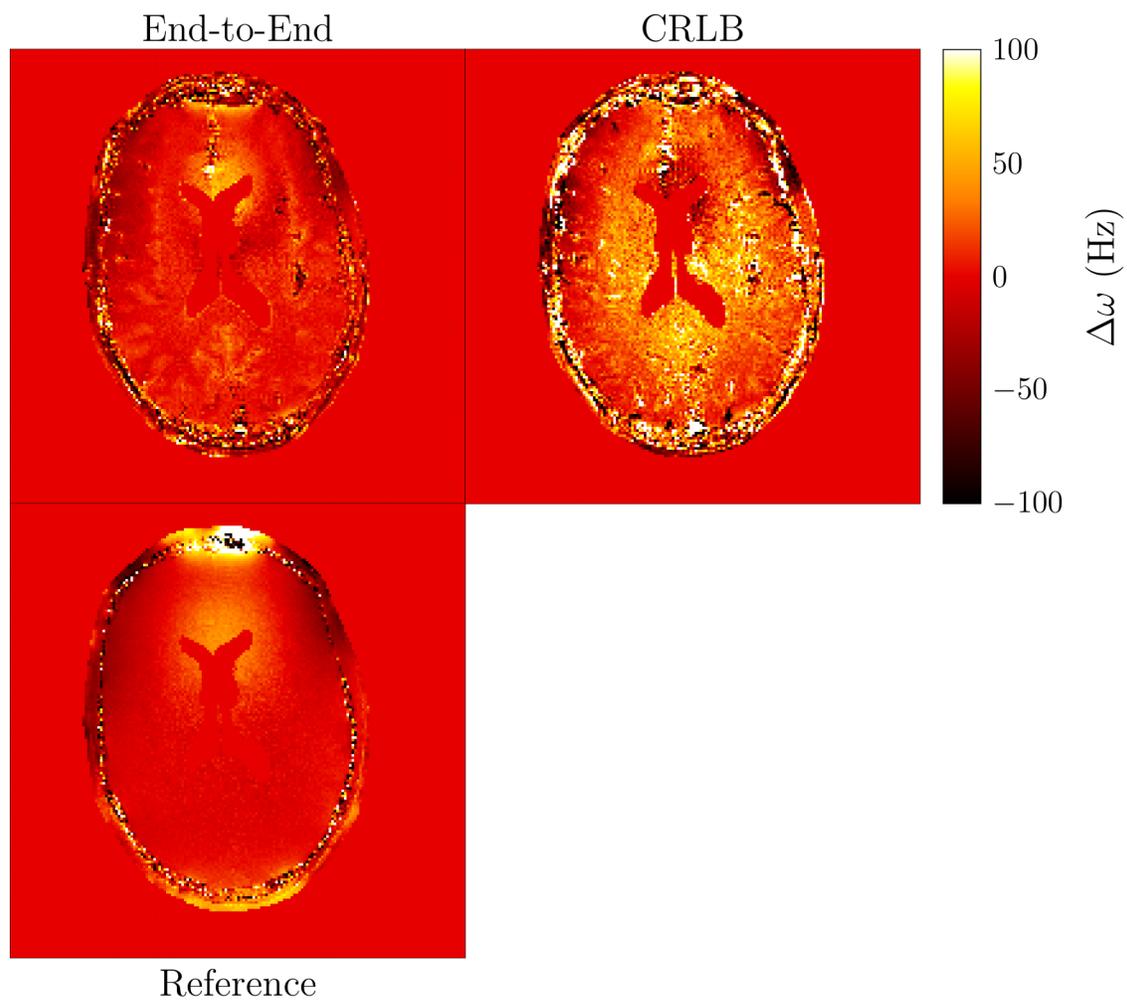


Figure 4.29: $\Delta\omega$ *in vivo* estimates alongside a reference B0 map acquired from two T_E -offset SPGR scans. While some tissue structure remains visible in the $\Delta\omega$ estimates from the end-to-end design, overall there is good visual agreement between the estimated B0 map and the reference one.

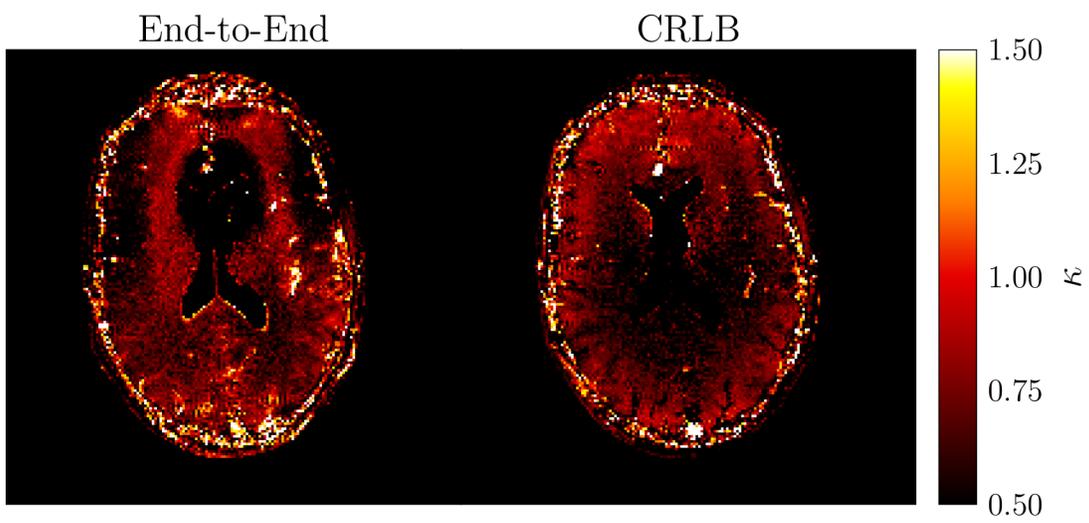


Figure 4.30: κ *in vivo* estimates.

CHAPTER 5

Intravoxel B0 Corrected Image Reconstruction with RF Prephasing

This chapter presents preliminary work extending conference abstract [15].

5.1 Introduction

Various MRI methods require acquiring signal with either a long readout and/or a long TE (e.g., fMRI, multi-gradient echo sequences, etc.), allowing a long time for off-resonance precession to occur. In areas of relatively constant B0 field inhomogeneity, this precession impacts only the phase of the image. However, in voxels having large B0 gradients, spins within a single voxel can dephase with respect to each other, inducing signal loss in the reconstructed image.

One approach to reduce the effect of this intravoxel dephasing is to sample higher frequency components in k-space [69], thus decreasing the size of voxels and limiting the intravoxel variation of off-resonance. A drawback of such an approach, however, is that longer scan times are needed. Another drawback of smaller voxels is lower SNR because there are fewer spins contributing signal in each voxel.

Another approach is to use model-based reconstruction using a model that accounts for intravoxel B0 effects [70–72]. Reconstructing images in this way can help alleviate artifacts from B0-induced signal loss. However, in areas of severe signal loss such as near the sinuses, such reconstruction can fail to fully recover the image due to near complete signal loss.

Therefore, in this work we propose to combine model-based reconstruction with the use of a *prephasing* RF pulse so that the intravoxel spread of spins is approximately zero at the echo time, mitigating the associated signal loss. We incorporate the effects of this prephasing into a model-based reconstruction method and demonstrate improved reconstructed images.

The organization of this chapter is as follows. Section 5.2 provides some background necessary for understanding model-based image reconstruction. Section 5.3 explains the

model-based image reconstruction technique proposed in this chapter. Section 5.4 describes our experiments. Section 5.5 reports the experimental results. Section 5.6 provides some discussion of the results. Section 5.7 gives concluding remarks and discusses ideas for future work.

5.2 Background

The data d_c acquired by the c th RF receive coil at time t in a MRI scan can be modeled as [72]

$$d_c(t) = \int_{-\infty}^{\infty} s_c(\mathbf{r})\rho(\mathbf{r})e^{-i2\pi\Delta f(\mathbf{r})t}e^{-i2\pi\mathbf{k}(t)\cdot\mathbf{r}} d\mathbf{r}, \quad (5.1)$$

where s_c is the spatially varying receive coil sensitivity map for the c th coil, ρ is the imaged object, $\Delta f(\mathbf{r}) = \frac{\Delta\omega(\mathbf{r})}{2\pi}$ (see (2.1)) is the off-resonance map that accounts for spatially varying differences in spins' resonance frequencies due to deviations in the main \mathbf{B}_0 field, the integration is over 3D spatial location $\mathbf{r} = (x, y, z)$, and $\mathbf{k}(t) = (k_x(t), k_y(t), k_z(t))$ is the 3D k-space location that represents systematic spatial variations in \mathbf{B}_0 introduced by separate *gradient* coils used for spatial localization of signal. These gradient coils are controlled by the user and are used to determine what points \mathbf{k} to sample. As seen from the Fourier transform structure of the last exponential in (5.1), these points are points in the spatial frequency domain (also called k-space) of the imaged object. The set of sampled k-space points and the time at which they are sampled comprise a readout, or k-space, trajectory.

One common readout trajectory is the echo planar imaging (EPI) readout. The EPI readout is a Cartesian sampling scheme in which, after RF excitation, k_z is fixed but a matrix of k_x and k_y samples are acquired in rapid succession. This readout can be repeated for as many k_z samples as desired, resulting in a stack-of-EPI readout. The EPI readout also generalizes directly to 3D, in which case k_z is not fixed and a 3D array of data is acquired. This is called a single-shot 3D EPI readout.

Introducing a voxel basis function $\psi(\cdot)$ (representing the shape of a voxel, or the distribution of signal within a voxel) and discretizing in space allows one to replace $s_c(\mathbf{r})\rho(\mathbf{r})$ with $\sum_{n=1}^N s_{c,n}x_n\psi(\mathbf{r} - \mathbf{r}_n)$, where N is the number of voxels, $s_{c,n}$ is the coil sensitivity for the n th voxel and c th coil, x_n is the image intensity for the n th voxel, and \mathbf{r}_n is the spatial position of the center of the n th voxel. Combining this discretization with (5.1) yields

$$d_c(t) = \sum_{n=1}^N s_{c,n}x_n \int_{-\infty}^{\infty} \psi(\mathbf{r} - \mathbf{r}_n)e^{-i2\pi\Delta f(\mathbf{r})t}e^{-i2\pi\mathbf{k}(t)\cdot\mathbf{r}} d\mathbf{r}. \quad (5.2)$$

Note that typically ψ is taken to be a Dirac impulse, resulting in

$$d_c(t) = \sum_{n=1}^N s_{c,n} x_n e^{-i2\pi\Delta f(\mathbf{r}_n)t} e^{-i2\pi\mathbf{k}(t)\cdot\mathbf{r}_n}, \quad (5.3)$$

which, if one ignores the off-resonance term and samples in time are taken such that k-space samples lie on a suitable Cartesian grid, permits the use of the Fast Fourier Transform (FFT) for image reconstruction.

To model intravoxel B0 effects, ψ cannot be a Dirac impulse (i.e., it must have some spatial extent). One choice is to use a rectangular basis. Further assuming off-resonance varies linearly within each voxel and discretizing in time yields [72]

$$\begin{aligned} d_c(t_m) &= \sum_{n=1}^N s_{c,n} x_n \int_{-\infty}^{\infty} \text{rect}_3\left(\frac{\mathbf{r} - \mathbf{r}_n}{\Delta}\right) e^{-i2\pi(f_n + \mathbf{g}_n \cdot (\mathbf{r} - \mathbf{r}_n))t_m} e^{-i2\pi\mathbf{k}_m \cdot \mathbf{r}} d\mathbf{r} \\ &= \sum_{n=1}^N s_{c,n} x_n e^{-i2\pi f_n t_m} \int_{-\infty}^{\infty} \text{rect}_3\left(\frac{\mathbf{r} - \mathbf{r}_n}{\Delta}\right) e^{-i2\pi\mathbf{g}_n \cdot (\mathbf{r} - \mathbf{r}_n)t_m} e^{-i2\pi\mathbf{k}_m \cdot \mathbf{r}} d\mathbf{r} \\ &= \sum_{n=1}^N s_{c,n} x_n e^{-i2\pi f_n t_m} \text{sinc}_3((\mathbf{k}_m + \mathbf{g}_n t_m) \odot \Delta) e^{-i2\pi\mathbf{k}_m \cdot \mathbf{r}_n}, \end{aligned} \quad (5.4)$$

where t_m is the k-space sample time, f_n is the off-resonance frequency at the center of the n th voxel, sinc_3 is the 3D sinc function (i.e., $\text{sinc}_3(x, y, z) = \text{sinc}(x) \text{sinc}(y) \text{sinc}(z)$), rect_3 is the 3D rectangle function (defined similarly), \mathbf{k}_m is the k-space location for the m th sample time, \mathbf{g}_n is the spatial gradient of the B0 field map at the n th voxel, Δ is the voxel size, and \odot is element-wise (Hadamard) product. Signal loss due to the intravoxel dephasing of spins is modeled via the sinc_3 term (which includes dephasing due to variations in off-resonance frequency as well as due to the gradient coils used for spatial localization, i.e., k-space location).

Collecting M time sample points into data vector \mathbf{y}_c , one can rewrite (5.4) in matrix-vector form:

$$\mathbf{y}_c = \mathbf{A}_c \mathbf{x}, \quad (5.5)$$

where $\mathbf{x} = [x_1, \dots, x_N]^\top$ and \mathbf{A}_c is the system model, which, for (5.4), is given by

$$\mathbf{A}_c = (\mathbf{W} \odot \mathbf{F}) \mathbf{S}_c, \quad (5.6)$$

where $\mathbf{S}_c = \text{diag}([s_{c,1}, \dots, s_{c,N}]^\top)$, \mathbf{F} is the Fourier operator, and the elements of \mathbf{W} are given by [72, 73] $W_{mn} = e^{-i2\pi f_n t_m} \text{sinc}_3((\mathbf{k}_m + \mathbf{g}_n t_m) \odot \Delta)$. In this case, one cannot use the FFT to compute $\mathbf{A}_c \mathbf{x}$, even when assuming $M = N$ and the k-space samples lie on a Cartesian

grid (in which case \mathbf{F} is the 3D Discrete Fourier Transform (DFT) matrix).

To improve computational efficiency, one can use a rank- L approximation of \mathbf{W} and then rewrite (5.5) as [72]

$$\begin{aligned} \mathbf{y}_c &= \left(\left(\sum_{l=1}^L \mathbf{u}_l \mathbf{v}_l' \right) \odot \mathbf{F} \right) \mathbf{S}_c \mathbf{x} \\ &= \sum_{l=1}^L \mathbf{u}_l \odot (\mathbf{F}(\mathbf{v}_l^* \odot \mathbf{S}_c \mathbf{x})), \end{aligned} \quad (5.7)$$

where \mathbf{u}_l and \mathbf{v}_l are the columns of \mathbf{U} and \mathbf{V} , respectively, such that $\mathbf{W} \approx \mathbf{U}\mathbf{V}'$, $(\cdot)'$ denotes conjugate transpose, and $(\cdot)^*$ denotes element-wise conjugation. Now, instead of performing one matrix-vector product to compute $\mathbf{A}_c \mathbf{x}$, L FFTs can be used, greatly reducing the computational burden for $L \ll N$ ($L \approx 10$ usually suffices in practice).

5.3 Theory

In this work, we propose a model-based image reconstruction technique that incorporates the effects of both intravoxel dephasing and RF prephasing. Typically, RF excitation pulses are chosen not to impart any phase on the imaged object. However, one can adjust the phase of the excitation pulse, thus adjusting the phase of the image. Furthermore, this imparted phase can vary spatially. Purposely imparting phase on an object using the RF excitation pulse is called RF prephasing.

The effects of RF prephasing can be incorporated into (5.1):

$$d_c(t) = \int_{-\infty}^{\infty} s_c(\mathbf{r}) \rho(\mathbf{r}) e^{-i2\pi\phi(\mathbf{r})} e^{-i2\pi\Delta f(\mathbf{r})t} e^{-i2\pi\mathbf{k}(t)\cdot\mathbf{r}} d\mathbf{r}, \quad (5.8)$$

where ϕ is the spatially varying phase (in units of cycles) imparted by the RF pulse. After discretizing using the rectangular basis, assuming ϕ varies linearly within each voxel, and rewriting in matrix-vector form, the data vector again is given by $\mathbf{y}_c = \mathbf{A}_c \mathbf{x}$. This time, however, the system matrix \mathbf{A}_c includes the effects of RF prephasing. In particular,

$$\mathbf{A}_c = (\mathbf{W} \odot \mathbf{F}) \mathbf{\Phi} \mathbf{S}_c, \quad (5.9)$$

where $\mathbf{\Phi} = \text{diag}([e^{-i2\pi\phi(\mathbf{r}_1)}, \dots, e^{-i2\pi\phi(\mathbf{r}_N)}]^\top)$ is a diagonal matrix containing the bulk phase imparted to the center of each voxel by the RF prephasing pulse and now the elements of

\mathbf{W} are

$$W_{mn} = e^{-i2\pi f_n t_m} \text{sinc}_3((\mathbf{k}_m + \mathbf{g}_n t_m + \boldsymbol{\gamma}_n) \odot \boldsymbol{\Delta}), \quad (5.10)$$

where $\boldsymbol{\gamma}_n = \nabla\phi(\mathbf{r})|_{\mathbf{r}=\mathbf{r}_n}$ is the spatial gradient of the RF prephasing phase map at the n th voxel. (Note that under the assumption that ϕ varies linearly within each voxel, $\phi(\mathbf{r}) = \sum_{n=1}^N (\phi_n + \boldsymbol{\gamma}_n \cdot (\mathbf{r} - \mathbf{r}_n)) \text{rect}_3(\frac{\mathbf{r}-\mathbf{r}_n}{\boldsymbol{\Delta}})$, where ϕ_n is the phase imparted by the RF prephasing pulse to the center of the n th voxel.)

Note that $\boldsymbol{\Phi}$ and $\boldsymbol{\gamma}_n$ are the key new aspects that incorporate the effects of RF prephasing into the system model. An ideal RF prephasing pulse will impart phase such that $f_n T_E + \phi_n = 0$ and an intravoxel spread of phase such that $\mathbf{g}_n T_E + \boldsymbol{\gamma}_n = \mathbf{0}_3 \forall n$, i.e., all spins throughout the object will be in phase at the echo time.

We use the system matrix \mathbf{A}_c in (5.9) in a regularized model-based reconstruction framework for multi-channel data. In particular, we solve

$$\arg \min_{\mathbf{x}} \sum_{c=1}^C \frac{1}{2} \|\mathbf{A}_c \mathbf{x} - \mathbf{y}_c\|_2^2 + \lambda R(\mathbf{x}), \quad (5.11)$$

where \mathbf{y}_c is the acquired k-space data from coil c , C is the total number of coils, R is a regularizer, and λ is a tuning parameter. We note that (5.11) assumes we know all the terms that make up the system matrix \mathbf{A}_c . In particular, coil sensitivity maps (\mathbf{S}_c) and a B0 field map (f_n values) must be separately acquired. The RF prephasing pulse is then designed using the B0 field map and the echo time to determine the ideal amount of prephasing to impart. (Thus the RF prephasing pulse is designed in an online way and then used to scan.) Because the RF pulse is known, the RF prephasing map ($\boldsymbol{\Phi}$) is known. The spatial gradients of the field map (\mathbf{g}_n values) and of the RF prephasing map ($\boldsymbol{\gamma}_n$ values) are computed via finite differences.

5.4 Methods

For the experiments in this work, we reconstructed raw MRI k-space data using (5.11). We used the ℓ_2 -norm of 3D finite differences in the three principal (x, y, and z) directions for the regularizer R with $\lambda = 0.1$. We solved the minimization problem using 300 iterations of conjugate gradient (`ncg` in the Julia [50] package `MIRT.jl` [74]), which was enough to converge to the global minimum cost function value in all cases. We used a rank-10 approximation of \mathbf{W} . To ease the computational burden of computing the singular vectors of \mathbf{W} , we approximately computed them by randomly sampling 1,000 rows and columns of \mathbf{W} , following what was done in [72].

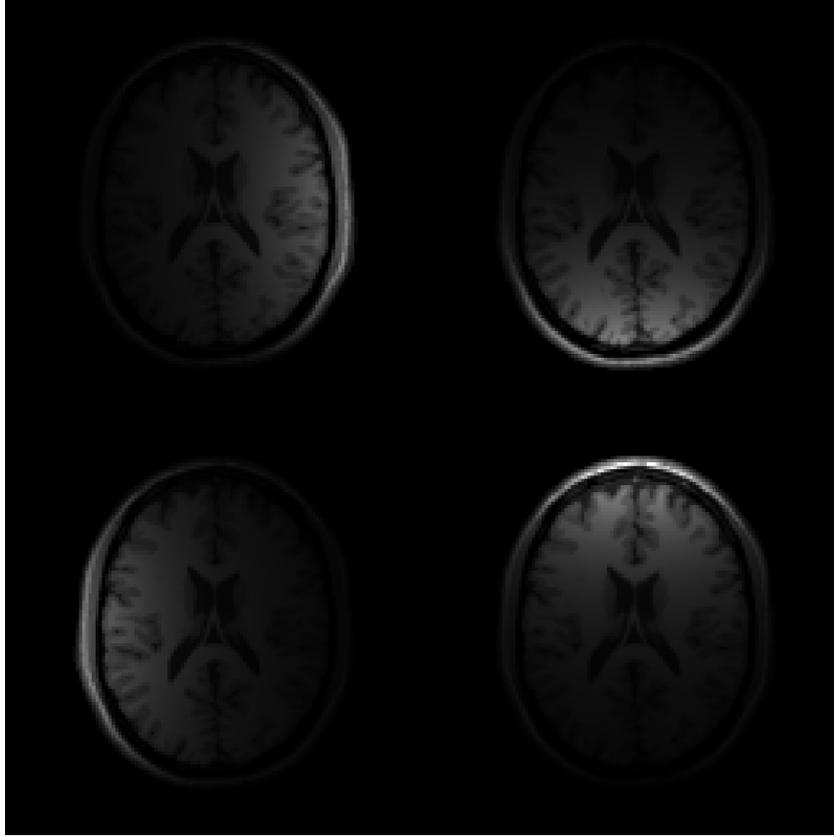


Figure 5.1: BrainWeb image weighted by the four individual coil sensitivities used in the simulation experiments. Only the middle slice of the simulated object is shown. Each coil “sees” only part of the object, but the information from each coil is used jointly during reconstruction to determine the true underlying object.

For the simulation experiments, we used a digital BrainWeb phantom [53] with matrix size $128 \times 128 \times 3$. We used four synthetic coil sensitivity maps (shown in Figure 5.1). We generated an off-resonance map with values 0–85 Hz and spatial gradients with values -3.6 – 3.6 Hz/voxel in the x direction, -3.6 – 3.6 Hz/voxel in the y direction, and -61 – 0 Hz/voxel in the z direction (see Figure 5.2). In other words, each voxel had a linear intravoxel spread of off-resonance frequencies in each of the x, y, and z directions. For example, if a voxel had a spatial gradient of 3 Hz/voxel in the x direction, then spins within the voxel on the right border of the voxel would precess with an off-resonance frequency 3 Hz greater than spins on the left border of the voxel. The spatial gradients were computed using central finite differences for all non-border voxels and left or right finite differences otherwise. We generated noisy k-space data by simulating a Cartesian stack-of-EPI readout, unless otherwise noted. Each readout occurred from 3 to 40 ms with $T_E = 21.6$ ms.

In the first simulation experiment, we compared different reconstruction methods for data

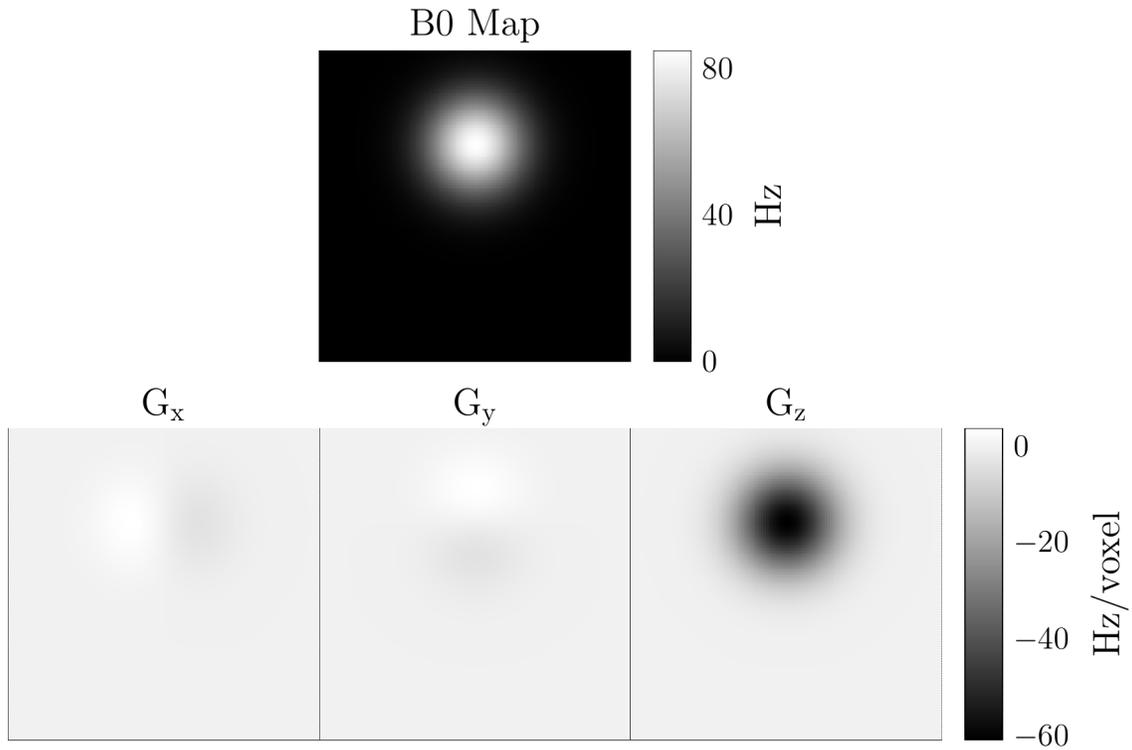


Figure 5.2: Off-resonance map and its spatial gradients used in the simulation experiments. Only the values corresponding to the middle slice are shown. The B0 gradients in the x and y directions range from -3.6 to 3.6 Hz/voxel. For $T_E = 21.6$ ms, this corresponds to less than $1/12$ cycle/voxel at most. In the z-direction, the peak of -61 Hz/voxel leads to more than one cycle/voxel.

simulated both with and without RF prephasing. The data simulated with RF prephasing used ideal prephasing. We compared three reconstruction techniques for each data set:

1. *Square root sum-of-squares (SSoS)*: We reconstructed the data for each coil individually using an inverse FFT and then combined the individual coil images with a square root sum-of-squares combination. Note that this method uses a model that ignores coil sensitivity effects, bulk B0 and prephasing effects, and intravoxel effects, i.e., $\mathbf{A}_c = \mathbf{F}$. This method also has no regularization.
2. *Model-based reconstruction without intravoxel effects*: We reconstructed the data using system matrix (5.9) but with $\mathbf{g}_n = \boldsymbol{\gamma}_n = \mathbf{0}_3 \forall n$. Thus we ignored intravoxel dephasing due to off-resonance and RF prephasing, but we modeled coil sensitivity effects, bulk B0 and prephasing effects, and intravoxel dephasing due to k-space location.
3. *Model-based reconstruction with intravoxel effects*: We reconstructed the data using the proposed system model (5.9) (which is the same as (5.6) when RF prephasing is not used during data acquisition), thus accounting for both intravoxel B0 effects and (when applicable) intravoxel RF prephasing effects.

We also repeated this experiment with a single-shot 3D EPI readout. We kept the readout timing the same (from 3 to 40 ms, effectively traversing k-space three times more quickly than in the initial experiment), but the echo time changed slightly to $T_E = 21.5$ ms.

In the second simulation experiment, we investigated what happens when RF prephasing is not ideal. We simulated data with RF prephasing varying from none to ideal. In other words, instead of having $f_n T_E + \phi_n = 0$ and $\mathbf{g}_n T_E + \boldsymbol{\gamma}_n = \mathbf{0}_3 \forall n$, we had RF prephasing such that $f_n t + \phi_n = 0$ and $\mathbf{g}_n t + \boldsymbol{\gamma}_n = \mathbf{0}_3 \forall n$ for rephasing times $t \in \{0, 5, 10, 12, 15, T_E\}$ ms. We reconstructed the data using model-based reconstruction both with and without intravoxel effects (methods 2 and 3 above). For the cases where we reconstructed the data accounting for intravoxel effects, we assumed we exactly knew the amount of non-ideal RF prephasing that was done.

5.5 Results

Figure 5.3 shows the ground truth image and the results of the different reconstruction methods for the k-space data simulated with and without RF prephasing. Only the middle slice is shown. Table 5.1(a) shows the corresponding normalized root mean squared error (NRMSE) for each reconstructed image, where $\text{NRMSE} = \|\hat{\mathbf{x}} - \mathbf{x}\|_2^2 / \|\mathbf{x}\|_2^2$, where \mathbf{x} is the true image and $\hat{\mathbf{x}}$ is the reconstructed image. Without RF prephasing, there was severe signal

Table 5.1: NRMSEs for each of the reconstructed images in (a) Figure 5.3 and (b) Figure 5.4. (a) Even though the NRMSE for the image reconstructed with a model including intravoxel B0 effects from data without RF prephasing was only 2.3%, residual artifacts are still clearly visible in the reconstructed image. On the other hand, both model-based reconstructed images from data with ideal RF prephasing had excellent NRMSEs, regardless of whether or not intravoxel B0 and RF prephasing effects were modeled. (b) The NRMSEs from the single-shot 3D EPI data were worse than those from the stack-of-EPI data (a), at least for the model-based reconstructed images. In this case, reconstructing the data with ideal RF prephasing requires using a model that includes intravoxel B0 and prephasing effects for good performance.

(a) Stack-of-EPI readout		
	No RF Prephasing	Ideal RF Prephasing
SSoS reconstruction	72.5%	67.2%
Reconstruction without intravoxel effects	21.5%	0.1%
Reconstruction with intravoxel effects	2.3%	0.2%
(b) Single-shot 3D EPI readout		
	No RF Prephasing	Ideal RF Prephasing
SSoS reconstruction	72.0%	66.8%
Reconstruction without intravoxel effects	34.8%	117.8%
Reconstruction with intravoxel effects	5.1%	0.7%

loss in the area of the image with large B0 gradients. Even when taking intravoxel B0 effects into account during image reconstruction, there was still residual artifact. The reconstructed images from the data with ideal RF prephasing were all free from signal loss artifacts, though the SSoS-reconstructed image still had incorrect image scaling due to unmodeled coil sensitivities. The two images reconstructed using model-based reconstruction were close to the true image and were very similar despite one not modeling intravoxel B0 and RF prephasing effects.

Figure 5.4 is the same as Figure 5.3 but with the data acquired using a simulated single-shot 3D EPI readout. Table 5.1(b) shows the corresponding NRMSEs. Overall, these reconstructed images are of worse quality than those from the stack-of-EPI data, but they do show similar trends. The most notable difference, however, is that ideal RF prephasing did not recover the true image without also being combined with model-based reconstruction using a model that included intravoxel B0 and RF prephasing effects.

Figure 5.5 shows the images reconstructed with a model that included intravoxel effects from data with different amounts of RF prephasing. Again, only the middle slice is shown. Table 5.2 shows the corresponding NRMSEs. As expected, as the amount of RF prephas-

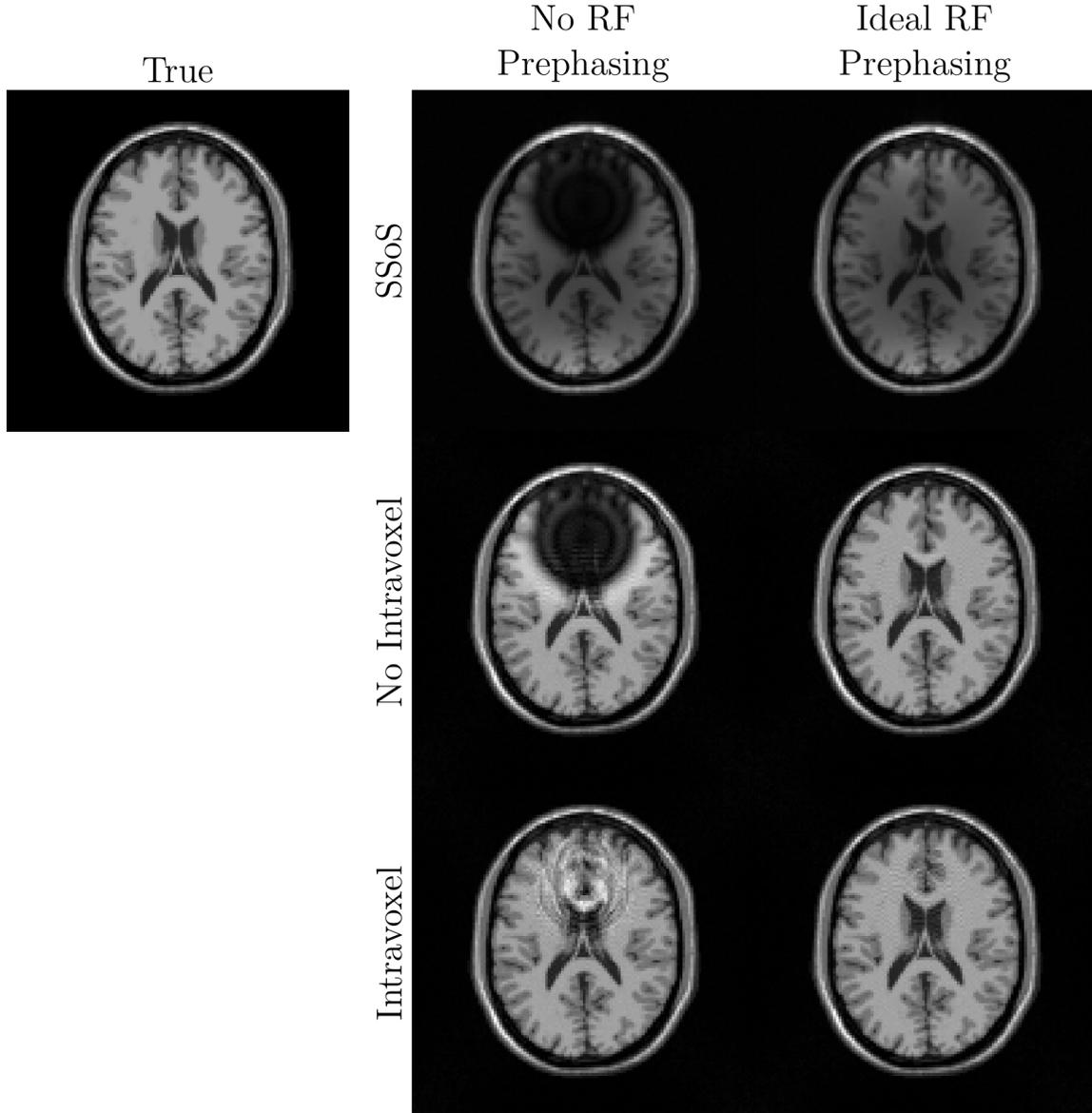


Figure 5.3: True and reconstructed images for k-space data simulated with and without ideal RF prephasing. We compared three reconstruction approaches: square root sum-of-squares (SSoS) reconstruction and model-based reconstruction with and without modeling intravoxel B0 and RF prephasing effects. The SSoS reconstructed images were scaled for display. Without RF prephasing, the reconstructed images had severe signal loss due to intravoxel dephasing. This signal loss was mitigated, but not entirely removed due to (near) complete signal loss that occurs at and near the zero-crossings of the sinc function in (5.4), when reconstructing with a model that incorporated such intravoxel dephasing. Ideal RF prephasing by itself essentially eliminated all signal loss due to intravoxel dephasing, regardless of reconstruction method.

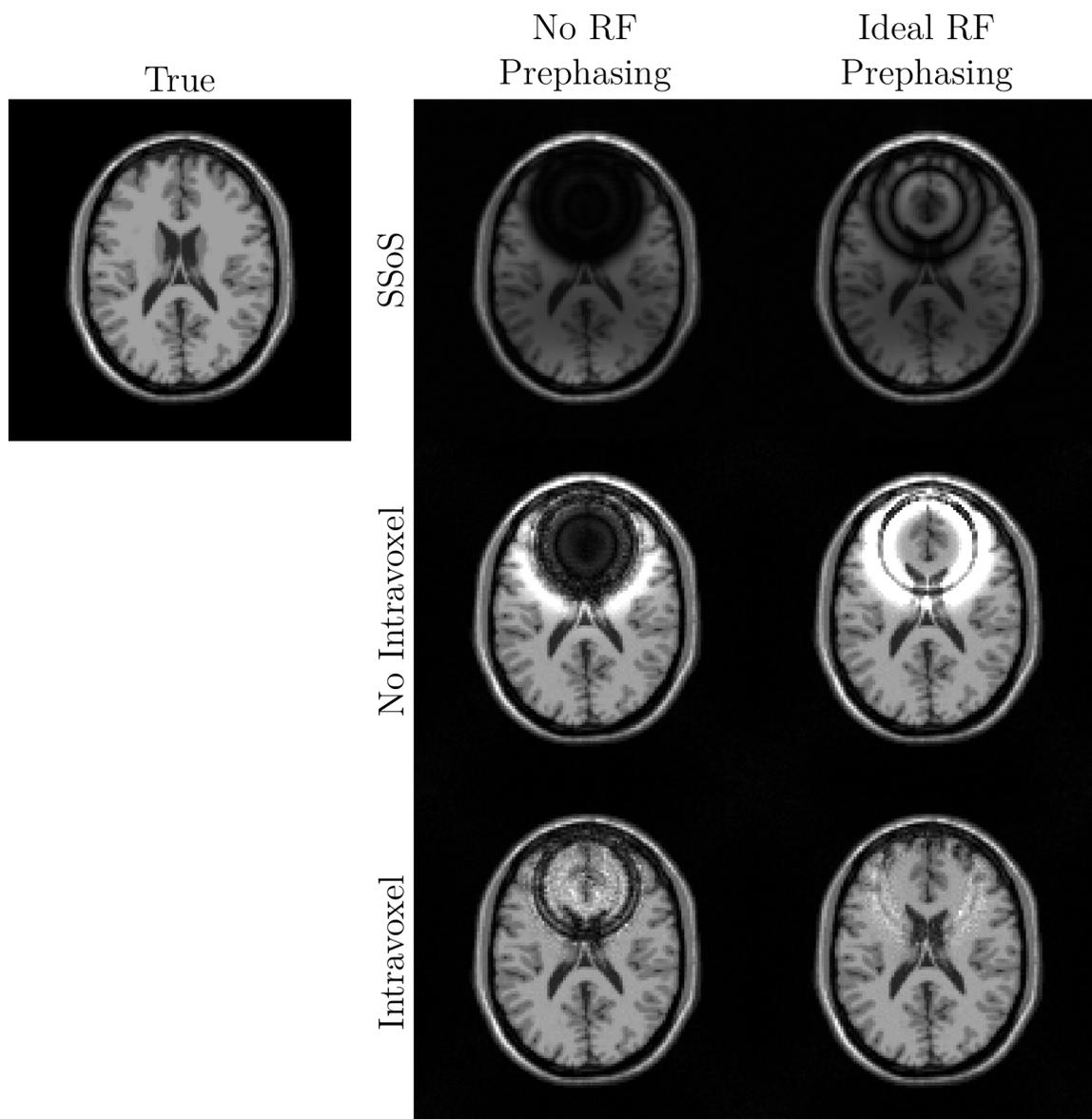


Figure 5.4: True and reconstructed images for k-space data simulated with and without ideal RF prephasing using a single-shot 3D EPI readout. The reconstruction approaches were the same as in Figure 5.3. The SSoS reconstructed images were scaled for display. Overall, these reconstructed images are of worse quality than those from the stack-of-EPI data, but they do show similar trends. The most notable difference, however, is that ideal RF prephasing did not recover the true image without also being combined with model-based reconstruction using a model that includes intravoxel B0 and RF prephasing effects.

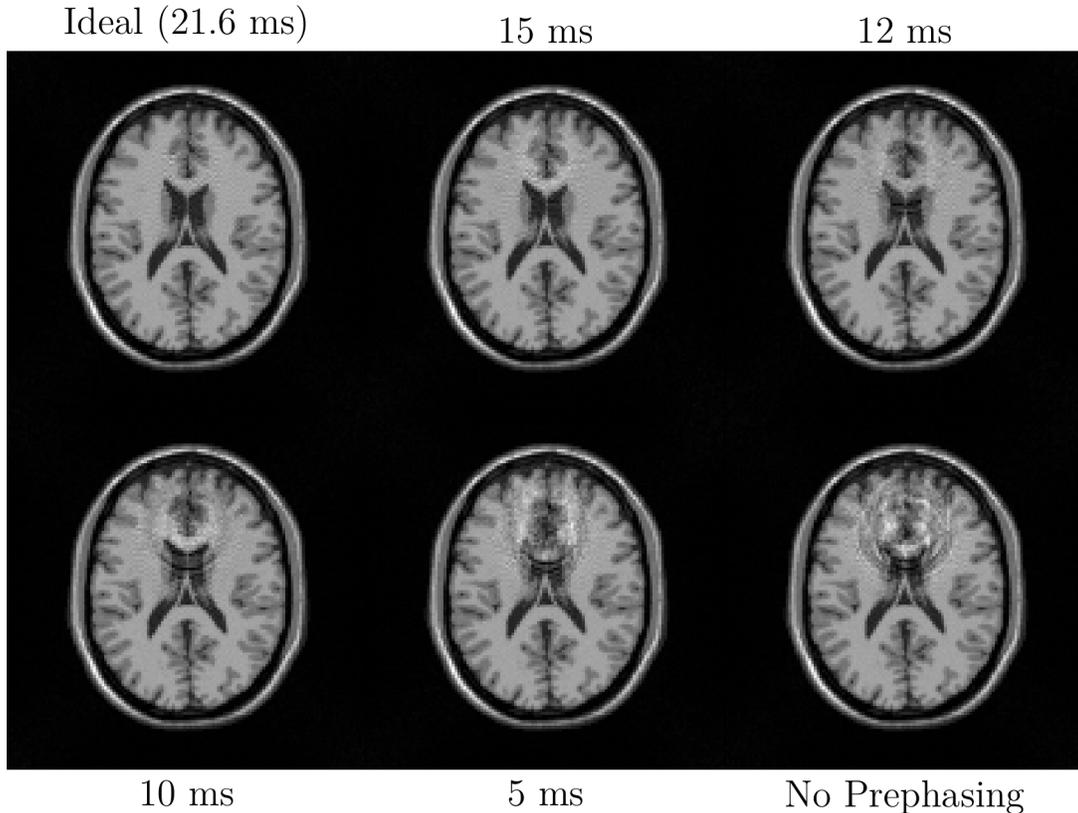


Figure 5.5: Reconstructed images for k-space data simulated with varying amounts of RF prephasing. We reconstructed these images using a model that included intravoxel B0 and RF prephasing effects. As expected, the reconstructed image quality improved as the rephasing time increased from 0 (no prephasing) to T_E (ideal prephasing). The residual signal loss artifact was still severe for rephasing times $t \in \{0, 5\}$ ms. Starting at a rephasing time of 10 ms, however, the reconstructed image quality improved significantly.

ing increased from none to ideal, the quality of the reconstructed images improved. The residual signal loss artifact was still severe for rephasing times $t \in \{0, 5\}$ ms. Starting at a rephasing time of 10 ms, however, the reconstructed image quality improved significantly. When reconstructing the images without modeling intravoxel B0 and prephasing effects, a similar trend was seen but with more severe signal loss as the rephasing time decreased (see Figure 5.6 and Table 5.2).

5.6 Discussion

In the simulated experiments of this work, we used an artificially generated off-resonance map with severe through-voxel gradients to illustrate the benefits of RF prephasing and model-based reconstruction. We note that, in practice, the intravoxel B0 gradients, expressed in

Table 5.2: NRMSEs for each of the reconstructed images in Figures 5.5 and 5.6. As expected, the NRMSE improved as the amount of RF prephasing increased from no prephasing to ideal prephasing. The NRMSEs for the images reconstructed using a model that included intravoxel B0 and RF prephasing effects were much better overall than those of the images reconstructed without those intravoxel effects.

		Model Intravoxel Effects?	
		Yes	No
Rephasing Time	0 ms	2.0%	21.5%
	5 ms	1.7%	16.1%
	10 ms	0.7%	5.5%
	12 ms	0.5%	2.2%
	15 ms	0.3%	1.2%
	21.6 ms (T_E)	0.2%	0.2%

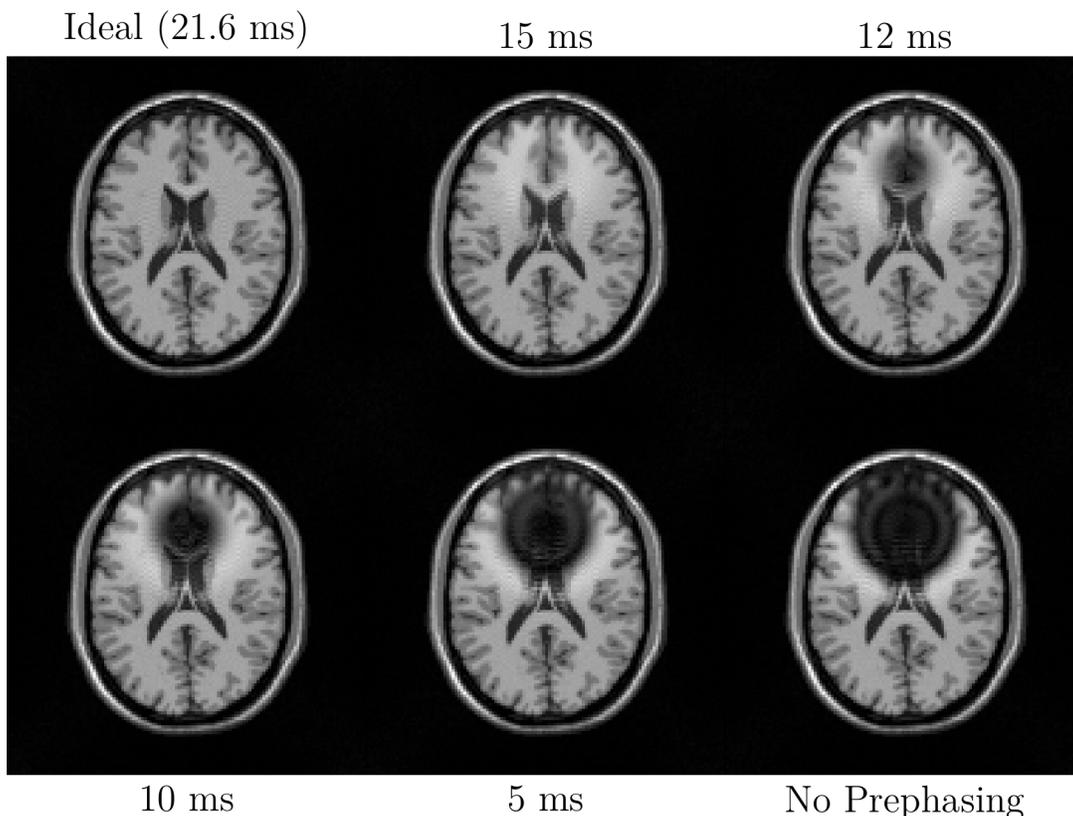


Figure 5.6: Reconstructed images for k-space data simulated with varying amounts of RF prephasing. We reconstructed these images using a model that did not include intravoxel B0 and RF prephasing effects. The image quality of these reconstructed images degrades much less gracefully as the rephasing time decreases compared to Figure 5.5, which has images reconstructed using a model that included intravoxel B0 and RF prephasing effects.

units of Hz/mm, are similar in all three principal directions, so the off-resonance map we simulated in this work is not necessarily realistic. Therefore, future work could extend this analysis to evaluate what effect RF prephasing and model-based reconstruction have for a range of typical values for the spatial gradients of off-resonance maps encountered in practice, particularly when scanning near the sinuses.

Figures 5.3 and 5.4 both illustrate the effect ideal RF prephasing has in the presence of large B0 gradients. Even when taking intravoxel B0 effects into account during image reconstruction, there was still residual artifact when reconstructing data simulated without RF prephasing, indicating that model-based reconstruction is not entirely sufficient on its own to correct for the signal loss due to intravoxel B0 effects. Combining the model-based reconstruction with ideal RF prephasing resulted in greatly improved reconstructed image quality.

Figure 5.3 seems to indicate that modeling intravoxel B0 and RF prephasing effects is unimportant when ideal RF prephasing is used; the reconstructed image had less than 1% error regardless of whether or not those intravoxel effects were modeled. However, Figure 5.4 demonstrates that other factors (in this case, the readout trajectory) can influence the relative impact of reconstructing data with ideal RF prephasing using a model including intravoxel B0 and prephasing effects versus using a model without those effects. Future work could further investigate different readout trajectories to see in what circumstances using a model including intravoxel effects is necessary for good reconstruction quality. Additionally, other factors (e.g., under-sampling) could be investigated to see if they also influence the relative impact of using such intravoxel modeling.

Even with the stack-of-EPI acquisition, Figures 5.5 and 5.6 show that modeling intravoxel B0 and RF prephasing effects becomes more important when the RF prephasing is not ideal. Using such modeling during image reconstruction results in a more graceful decline in reconstructed image quality as the rephasing time decreases (becomes less ideal) than when reconstructing using a model without those intravoxel effects.

Figure 5.5 also indicates that ideal RF prephasing is not necessary to overcome the effects of severe through-voxel B0 gradients; the NRMSE of the reconstructed image was still below 1% even for a rephasing time of 10 ms, which is less than half the ideal rephasing time ($T_E = 21.6$ ms). This result bodes well for RF prephasing in practice, where achieving ideal RF prephasing may at times be impractical.

Additionally, Figure 5.6 gives some sense of how robust the proposed reconstruction approach is to inaccuracies in the spatial gradients of the B0 and prephasing maps. In this work, we assumed perfect knowledge of the B0 and prephasing maps and their spatial gradients. In practice, however, these estimates will be imperfect, and thus the reconstruction will

assume some amount of dephasing exists that is more or less than the amount of dephasing actually present in the data. In Figure 5.6, the case with a rephasing time of $t = 15$ ms assumed there was no dephasing at the echo time (i.e., $\mathbf{g}_n T_E + \boldsymbol{\gamma}_n = \mathbf{0}_3$), while in reality the amount of dephasing was $\mathbf{g}_n(T_E - t)$ (i.e., there was some dephasing that occurred from the rephasing time t to the echo time T_E). Despite this model mismatch, which can also be viewed as an underestimation of \mathbf{g}_n by about 30% or an overestimation of $\boldsymbol{\gamma}_n$ by about 44%, the reconstructed image has high visual quality and low NRMSE (1.2%). Thus the proposed reconstruction approach appears to be robust to errors in the spatial gradients on the order of 30%. Of course, the robustness of the proposed approach to such inaccuracies likely also depends on the readout trajectory. Future work could do a more thorough characterization of such robustness for different readout trajectories.

In this work, we used a rectangular basis function for representing the shape of voxels. However, it may be that other basis functions lead to better reconstructed images. In particular, a sinc basis may be a better choice due to the nature of the Fourier sampling. Future work could explore different voxel basis functions in more detail and analyze the impact they have on image reconstruction, both with and without RF prephasing.

5.7 Conclusion

In this work, we proposed incorporating RF prephasing into model-based reconstruction that models intravoxel B0 and prephasing effects. We compared this approach in simulation to model-based reconstruction with intravoxel B0 effects but without RF prephasing. The proposed approach resulted in a reconstructed image that more closely matched the true underlying image. We also found that while RF prephasing can be sufficient on its own to eliminate the signal loss artifact due to intravoxel dephasing, combining RF prephasing with model-based reconstruction that incorporates intravoxel B0 and prephasing effects provides better reconstructed images for a wider range of situations (e.g., for a different readout trajectory, or when RF prephasing is not ideal, as explored in this work).

The aim of this work was to introduce the proposed reconstruction method. So far, only simulations have been done. These results need to be verified with actual data. For this, it will be necessary to develop algorithms for designing RF prephasing pulses. Future work could investigate doing so with pulses based on spectral [75] and/or spatial [76, 77] excitation profiles.

Furthermore, the proposed reconstruction method can be used, e.g., as part of a fMRI study where long echo times are needed to observe the BOLD effect. Therefore, future work could investigate the effect of the proposed method on fMRI activation maps.

CHAPTER 6

Future Work

This chapter offers ideas for future work that are more broadly applicable than for just one of the previous chapters. The discussion and conclusion sections of Chapters 3, 4, and 5 contain ideas for future work that are more specific to the corresponding chapters.

6.1 Model Intravoxel Dephasing in MWI

When using STFR and PERK for MWI as in Chapters 3 and 4, we used a tissue model that ignored so-called T_2^* effects, i.e., variations in off-resonance frequencies due to interactions between spins at a microscopic level. These interactions cause a spread of off-resonance frequencies within a voxel, which in turn causes magnetization vectors within a voxel to dephase with respect to each other, causing signal loss. This phenomenon is similar to intravoxel B0 effects as discussed in Chapter 5, but is different in that it is tissue-specific. Another difference is that the intravoxel spread of off-resonance values due to T_2^* effects is not well modeled by a uniform distribution of off-resonance values, but rather a Lorentzian distribution. T_2^* effects influence the STFR signal [19]; therefore, an area of future work is to model T_2^* effects in the STFR signal model used to train PERK for MWI and investigate how PERK MWF estimates are affected. Additionally, intravoxel B0 variations could also be modeled, and their effect on PERK MWF estimates studied.

6.2 Jointly Reconstruct Under-sampled STFR Scans

The *in vivo* results presented in Chapters 3 and 4 used fully sampled STFR scans. The scan time can be reduced by under-sampling, at the cost of a more complicated image reconstruction. Normally, under-sampled MRI images are reconstructed one at a time by solving the following optimization problem:

$$\hat{\mathbf{x}}_d = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{A}_d \mathbf{x} - \mathbf{y}_d\|_2^2 + \beta R(\mathbf{x}), \quad (6.1)$$

where $\hat{\mathbf{x}}_d$ represents the reconstructed image for the d th scan, \mathbf{A}_d is the system matrix that converts the true image \mathbf{x}_d into the observed data \mathbf{y}_d , R is a regularizer that encourages, e.g., sparsity and/or low-rankness in some domain [78], and $\beta > 0$ is a regularization parameter. Note that, ignoring system imperfections and coil sensitivity profiles, and for the single-coil case, \mathbf{A}_d is given by taking the DFT matrix and removing the rows corresponding to the unsampled data points for the d th scan. This formulation (6.1), however, does not take into account image similarities across all scans.

One way to set up the reconstruction to use information across all scans is to jointly reconstruct all images together, i.e.,

$$(\hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_D) = \arg \min_{\mathbf{x}_1, \dots, \mathbf{x}_D} \sum_{d=1}^D \frac{1}{2} \|\mathbf{A}_d \mathbf{x}_d - \mathbf{y}_d\|_2^2 + \beta R(\mathbf{x}_1, \dots, \mathbf{x}_D), \quad (6.2)$$

where now the regularizer R is a function of all scans together. For MWI with STFR, multiple STFR scans are acquired of the same anatomy, so choosing R to encourage common structure across all scans, such as group sparsity [79, 80] or low-rankness [80], should improve the reconstructions. Also, for each spatial location of the imaged object, the D STFR values lie on a manifold whose dimension is the number of free parameters in the STFR signal model, so another choice for R encourages the STFR signals for each voxel to lie close to that manifold. In this case,

$$R(\mathbf{x}_1, \dots, \mathbf{x}_D) = \sum_{n=1}^N \min_{\boldsymbol{\theta}_n} \left\| \left[x_1^{(n)}, \dots, x_D^{(n)} \right]^\top - \mathbf{s}(\boldsymbol{\theta}_n) \right\|_2^2, \quad (6.3)$$

where $x_d^{(n)}$ represents the n th voxel of the d th scan and \mathbf{s} is the STFR signal model as a function of the unknown parameters $\boldsymbol{\theta}_n$.

A direction for future work is to investigate different methods for reconstructing the under-sampled STFR scans. In particular, one idea is to develop a method for group sparsity using either a mixed ℓ_2 - ℓ_0 -norm or an ℓ_0 -norm on the singular values of the matrix $\mathbf{X} \triangleq [\mathbf{x}_1 \cdots \mathbf{x}_D]$, using an exact reformulation of the ℓ_0 -norm to make the problem tractable [81, 82], and compare it to the other aforementioned methods.

In addition to developing a method for reconstructing the under-sampled STFR scans, one could also investigate the trade-off between sampling in space versus along the contrast dimension. For a given amount of scan time, one can try to under-sample each STFR scan as much as possible to maximize the number of STFR scans that can be acquired, increasing sampling along the contrast dimension at the expense of less spatial sampling. (This approach is akin to magnetic resonance fingerprinting [83].) On the other extreme, one

can fully sample each STFR scan, minimizing the number of contrasts that can be acquired. Future work could investigate the trade-off between the two extremes to determine how to balance spatial sampling and contrast sampling.

6.3 Replace PERK with Deep Learning

In this dissertation, I used PERK for MWF estimation. Increasingly, however, deep neural networks are being used to achieve state-of-the-art performance in various MRI applications, including QMRI. One attractive aspect of using a deep neural network for QMRI is the ability to incorporate a signal model into network training. One way to do this is to train a network to estimate tissue parameters from the acquired images but optimize the network parameters to minimize the error between the acquired images and the images obtained by passing the estimated tissue parameters through a signal model [84–86]. An avenue for future work could be to investigate different network architectures (e.g., fully connected, CNN, U-Net) and jointly train the network and optimize the STFR scan parameters in an end-to-end fashion similar to the one introduced in Chapter 4.

6.4 Explore Bias-Variance Trade-Off in End-to-End Scan Design

In the end-to-end scan design (Chapter 4), I minimized the MSE of MWF estimates from PERK. MSE is equal to the sum of the variance and squared bias of the estimates, so minimizing the MSE places equal emphasis on both aspects. However, it may be useful to place unequal emphasis on the two. For example, since humans are good at reading through noise, it may be desirable to place more emphasis on minimizing the bias of estimates. One way to achieve this is to introduce a weighting hyperparameter $\gamma \in (0, 2)$ and then minimize

$$\mathbb{E}_{\mathbf{x}}[\gamma \|\text{Bias}(\hat{\mathbf{x}})\|_2^2 + (2 - \gamma) \text{var}(\hat{\mathbf{x}})], \quad (6.4)$$

where the bias and variance are the conditional bias and variance of parameter estimates given a particular true underlying parameter. Then, as in Chapter 4, the scan design would minimize an expected cost, averaged over distributions of tissue parameters. In this way, choosing a higher value of γ places more emphasis on minimizing bias, while choosing a lower value of γ places more emphasis on minimizing variance. Note that choosing $\gamma = 1$ is equivalent to minimizing the MSE. On one hand, adding γ introduces another hyperparameter that must be tuned, but on the other hand, adding γ allows finer control over the

true objective of interest. Furthermore, adding γ does not introduce any complexity to the actual minimization (i.e., minimizing the modified objective function is just as easy as minimizing MSE, in the sense that no constraints are added). Note also that using this approach would likely make the de-biasing procedure described in Chapter 4 unnecessary, as one could increase γ to decrease the bias of parameter estimates.

Besides adding γ , another way to place unequal emphasis on bias and variance is to minimize one with a constraint on the other. For example, one may desire to minimize bias subject to some constraint on the variance, resulting in MWF estimates that are as unbiased as possible given some allowable amount of noise. While adding a constraint makes the minimization problem harder, doing so can help ensure that MWF estimates have an acceptable level of noise or bias, which is not guaranteed in the previous formulation.

6.5 Optimize Magnetic Resonance Fingerprinting Scans

Magnetic resonance fingerprinting (MRF) is a successful, relatively new approach to QMRI [83]. Rather than using the same scan parameters (flip angles, repetition times, etc.) for each repetition of excitation and signal acquisition in a scan, in MRF the scan parameters are allowed to vary from repetition to repetition. The result is a signal time course that can then be fit to a model to estimate tissue parameters.

One idea for future work is to investigate optimizing MRF scan parameters using the end-to-end scan design approach from Chapter 4. To adapt the experiments in that chapter to MRF, instead of having a signal model to generate 18 STFR signals for a given set of tissue parameters, one would use a signal model to generate a MRF time course of, e.g., 1,000 time points. Furthermore, if not using PERK for parameter estimation, one would replace PERK training with generating the dictionary used for MRF pattern matching or with training a learning-based estimator (e.g., a neural network) as applicable.

There are a few things to consider when applying the end-to-end scan design approach to MRF. The first is that many MRF scans contain on the order of 1,000 scan parameters (many more than the 54 STFR scan parameters I optimized in Chapter 4). Thus the computational burden of computing gradients may be an issue, particularly when using a signal model that involves the Bloch-McConnell equation (2.6) (i.e., when modeling exchange between two or more compartments). One approach to lowering the number of parameters to optimize is to vary the scan parameters according to some parameterized function and then optimize that function's parameters. Then only a few parameters will need to be optimized, at the

expense of adding constraints on the scan parameters.

Another consideration (that is applicable in general but is especially pertinent to MRF) is under-sampling. MRF data often are severely under-sampled, resulting in reconstructed images with severe artifacts. However, the end-to-end scan design ignores the effects of under-sampling. If using a dictionary and pattern matching for parameter estimation, ignoring under-sampling effects in the scan design may be acceptable due to the incoherence of such artifacts with respect to the dictionary time courses [83]. On the other hand, if using a learned estimator for parameter estimation, under-sampled time courses should be used for training, and so should not be ignored in the scan design. To incorporate under-sampling effects into the end-to-end scan design approach, instead of drawing sets of tissue parameters from specified distributions (effectively simulating signals on a voxel-by-voxel basis), one would need to simulate full images with realistic anatomical features (e.g., using a BrainWeb digital phantom [53]), both for the validation data and for the training data. (Such an end-to-end approach would be much closer to the work in [63].) One would need to investigate how many realistic simulated images would be needed, particularly for training; for example, could one image have a wide range of tissue parameter values, or would the resulting under-sampling artifacts be too unrealistic?

Finally, the end-to-end scan design approach is amenable to optimizing continuous variables. Discrete variables, such as how many inversion pulses a MRF scan should have and when to include them, would need to be selected in some other way.

6.6 Investigate CSF Partial Volume Effects

One physical effect I ignored in this dissertation is partial volume effects, in which two (or more) *macroscopic* environments (such as WM and GM) exist within a single voxel. (This is in contrast to multiple *microscopic* environments, such as myelin water and non-myelin water, existing within a single voxel, which I did consider in this dissertation.) Partial volume effects with cerebrospinal fluid (CSF) are of particular interest; for example, voxels near the cortical surface of the brain will have a mix of GM and CSF. These effects also become more pronounced as the brain ages and atrophies. Since PERK assumes each voxel comprises just one of either WM or GM or CSF, partial volume effects lead to model mismatch. One avenue of future work is to add a CSF compartment to the tissue model and to investigate the effect this has on MWF estimates and, in particular, WM/GM contrast *in vivo*. For simplicity, one could assume the CSF compartment does not exchange with the other compartments. This assumption is reasonable for partial volume effects that are at the interface between brain tissue and CSF in, e.g., the ventricles, where there is a natural barrier that prevents

exchange. However, this assumption breaks down in pathological conditions such as edema, where brain tissue and CSF are in direct contact [87]. First, one could investigate the model mismatch by simulating test data using the new tissue model but training PERK using the model that does not include CSF. If significant differences in the MWF estimates are found, the new tissue model model could be used to train PERK for estimating MWF from *in vivo* data to see whether WM/GM contrast improves.

Appendices

APPENDIX A

Effects of Magnetization Transfer on STFR-Based Myelin Water Imaging

Chapter 3 estimated MWF by simulating the STFR signal for training PERK. The signals were simulated assuming instantaneous RF excitation pulses. However, real excitation pulses have finite duration (about 1–2 ms). For magnetization with relaxation times much longer than the excitation pulse duration, the difference between instantaneous and non-instantaneous pulses is negligible, but for other magnetization, such as is found in macromolecules with T_2 values on the order of 10 μ s, the difference can lead to significant changes in the simulated signal. Though magnetization with such short T_2 is not typically MR-visible, it impacts the observed signal through exchange with other tissue compartments, a process called magnetization transfer (MT). This appendix presents work investigating how MT affects the STFR signal and how that, in turn, affects MWF estimates.

A.1 Introduction

It was shown in [88] that MT effects cause the steady state bSSFP signal to be much lower than predicted by conventional theory. STFR produces a similar signal to bSSFP, so we investigated to what extent MT effects influence the steady state STFR signal and, in turn, MWF estimates from STFR scans.

A.2 Methods

A.2.1 Modeling MT Effects

Brain tissues were modeled using a three-compartment model with exchange, as in Chapter 3. When simulating the STFR signal for training PERK in that chapter, however, instantaneous RF excitation pulses were assumed, i.e., the magnetization simply was rotated by

the prescribed flip angle without undergoing any relaxation or free-precession. To assess the effects of model mismatch, we accounted for MT effects by simulating sinc-shaped RF excitation pulses with 1 ms duration and time-bandwidth product of 8. Each RF pulse was played out in 401 time points, each of which was treated as an instantaneous RF pulse, and between each hard pulse, the magnetization experienced relaxation and free-precession.

A.2.2 Experiments

First, we looked at how a single-compartment macromolecule magnetization with $M_0 = 1$, $T_1 = 1000$ ms, and $T_2 = 50 \mu\text{s}$ was affected by the non-instantaneous RF excitation pulse. We also looked at a single-compartment white matter magnetization with $M_0 = 1$, $T_1 = 833$ ms, and $T_2 = 80$ ms for comparison. Next, we looked at the magnetization of a three-compartment, exchanging system with $M_0 = 1$, $f_f = 0.15$, $f_m = 0.1$, $T_{1,f} = 400$ ms, $T_{1,s} = 1000$ ms, $T_{1,m} = 1000$ ms, $T_{2,f} = 20$ ms, $T_{2,s} = 80$ ms, $T_{2,m} = 50 \mu\text{s}$, $\tau_{f \rightarrow s} = 100$ ms, $\tau_{f \rightarrow m} = 50$ ms, $\Delta\omega_f = 15$ Hz, $\kappa = 1$, and $\Delta\omega$ ranging from -100 to 100 Hz. In both cases, we looked at the magnetization arising from an STFR scan in steady state with $T_{\text{free}} = 8$ ms, $T_g = 2.8$ ms, $\alpha = 15^\circ$, $\beta = 15^\circ$, $\phi = 0$, and T_E at either 0 ms (immediately after the instantaneous pulse), 0.5 ms (immediately after the 1 ms pulse), or 4 ms (the conventional echo time).

After looking at the STFR steady state signals for a single STFR scan, we then investigated how PERK MWF estimates were influenced by non-instantaneous RF excitation pulses. We generated test data by simulating STFR scans with 1 ms sinc excitation pulses, and we compared training PERK with instantaneous excitation pulses to training PERK with non-instantaneous pulses. The STFR scans were simulated using scan design A from Table 3.2, with $T_{\text{free}} = 8$ ms, $T_g = 2.8$ ms, and $T_E = 4$ ms. Both test and training data were simulated from a range of tissue parameters given in Table 3.1.

A.3 Results

A single-compartment macromolecule magnetization is greatly affected by the duration of the excitation pulse (especially the longitudinal component M_z), unlike a single-compartment white matter magnetization, as seen in Table A.1. For a three-compartment, exchanging system, however, the acquired signal in steady state is essentially the same for both instantaneous and non-instantaneous excitation pulses, as shown in Figure A.1. The results in Figure A.1 are for $T_E = 0.5$ ms, i.e., immediately after the 1 ms sinc excitation pulse.

Table A.1: Steady state single-compartment magnetization using instantaneous and non-instantaneous RF excitation pulses. A macromolecule magnetization (especially the longitudinal component M_z) is much more sensitive to the duration of the excitation pulse than is a white matter magnetization. iRF: instantaneous excitation. niRF: non-instantaneous excitation.

		$T_E = 0$	$T_E = 0.5$ ms		$T_E = 4$ ms	
		iRF	niRF	iRF	niRF	iRF
Macromolecule	M_{xy}	0.035	-0.001	0.000	0.000	0.000
	M_z	0.131	0.197	0.132	0.200	0.135
White Matter	M_{xy}	0.175	0.173	0.174	0.166	0.166
	M_z	0.652	0.652	0.652	0.653	0.654

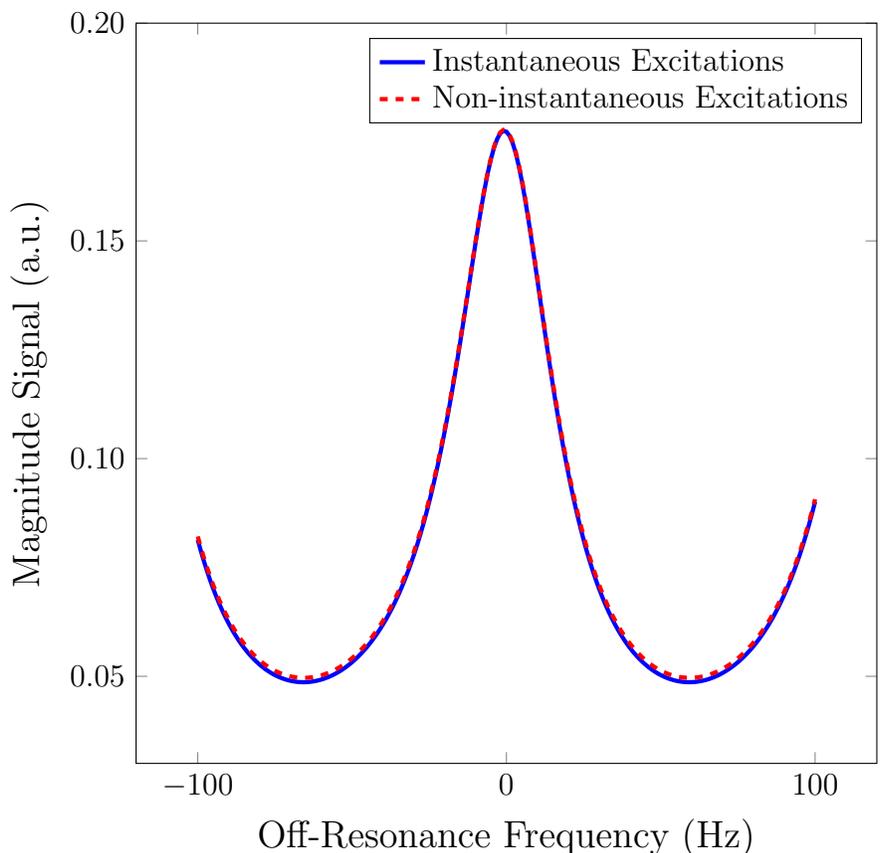


Figure A.1: Steady state STFR signal acquired at $T_E = 0.5$ ms for a three-compartment, exchanging system using instantaneous and non-instantaneous RF excitation pulses versus bulk off-resonance frequency $\Delta\omega$. The signals are essentially the same in both cases.

The RMSE of PERK MWF estimates was essentially equal when training with instantaneous excitation pulses as it was when training with non-instantaneous pulses (0.0206 versus 0.0199).

A.4 Conclusion

We found that while the steady state of a single macromolecule magnetization changes dramatically depending on whether the RF excitation pulses are instantaneous or not, the steady state STFR signal from a three-compartment, exchanging system does not change much at all. We also found that MT effects did not affect PERK MWF estimates very much, which is in line with the finding that the steady state STFR signal did not change much. In sum, it appears that for STFR-based MWF estimation using PERK modeling MT effects is not important.

Bibliography

- [1] K. L. Miller, S. M. Smith, and P. Jezzard. Asymmetries of the balanced SSFP profile. Part II: White matter. Mag. Res. Med., 63(2):396–406, February 2010. (Cited on [ix](#), [xvii](#), [17](#), [20](#), [23](#), [28](#), [32](#), [40](#), [52](#).)
- [2] E. Alonso-Ortiz, I. R. Levesque, and G. B. Pike. MRI-based myelin water imaging: A technical review. Mag. Res. Med., 73(1):70–81, January 2015. (Cited on [xvii](#), [19](#), [28](#), [49](#), [54](#).)
- [3] G. J. Stanisz, A. Kecojevic, M. J. Bronskill, and R. M. Henkelman. Characterizing white matter with magnetization transfer and T2. Mag. Res. Med., 42(6):1128–36, December 1999. (Cited on [xvii](#), [17](#), [18](#), [20](#), [23](#), [28](#), [54](#).)
- [4] S. C. L. Deoni. Correction of main and transmit magnetic field (B0 and B1) inhomogeneity effects in multicomponent-driven equilibrium single-pulse observation of T1 and T2. Mag. Res. Med., 65(4):1021–35, April 2011. (Cited on [xvii](#), [28](#), [54](#).)
- [5] A. P. Manning, A. L. MacKay, and C. A. Michal. Understanding aqueous and non-aqueous proton T1 relaxation in brain. J. Mag. Res., 323:106909, February 2021. (Cited on [xviii](#), [54](#), [69](#), [70](#), [71](#).)
- [6] A. Mackay, K. Whittall, J. Adler, D. Li, D. Paty, and D. Graeb. In vivo visualization of myelin water in brain by magnetic resonance. Mag. Res. Med., 31(6):673–7, June 1994. (Cited on [1](#), [19](#), [49](#).)
- [7] C. Laule, E. Leung, D. K. B. Li, A. L. Traboulsee, D. W. Paty, A. L. MacKay, and G. R. W. Moore. Myelin water imaging in multiple sclerosis: quantitative correlations with histopathology. Multiple Sclerosis J., 12(6):747–53, November 2006. (Cited on [1](#).)
- [8] J-F. Nielsen, D. Yoon, and D. C. Noll. Small-tip fast recovery imaging using non-slice-selective tailored tip-up pulses and radiofrequency-spoiling. Mag. Res. Med., 69(3):657–66, March 2013. (Cited on [1](#), [9](#), [20](#), [21](#), [50](#).)
- [9] G. Nataraj, J-F. Nielsen, C. D. Scott, and J. A. Fessler. Dictionary-free MRI PERK: Parameter estimation via regression with kernels. IEEE Trans. Med. Imag., 37(9):2103–14, September 2018. (Cited on [2](#), [11](#), [12](#), [13](#), [20](#), [21](#), [25](#), [26](#), [27](#), [50](#).)

- [10] S. T. Whitaker, G. Nataraj, J-F. Nielsen, and J. A. Fessler. Myelin water fraction estimation using small-tip fast recovery MRI. Mag. Res. Med., 84(4):1977–90, October 2020. (Cited on [2](#), [19](#), [49](#), [50](#), [51](#), [52](#), [54](#), [55](#), [56](#).)
- [11] S. T. Whitaker, G. Nataraj, M. Gao, J-F. Nielsen, and J. A. Fessler. Myelin water fraction estimation using small-tip fast recovery MRI. In Proc. Intl. Soc. Mag. Res. Med., page 4403, 2019. (Cited on [2](#), [19](#), [20](#), [49](#), [50](#).)
- [12] S. T. Whitaker, G. Nataraj, J-F. Nielsen, and J. A. Fessler. Myelin water imaging using STFR with exchange. In Proc. Intl. Soc. Mag. Res. Med., page 1020, 2020. (Cited on [2](#), [19](#), [49](#), [50](#).)
- [13] S. T. Whitaker, J-F. Nielsen, and J. A. Fessler. End-to-end MRI scan design approach for improved myelin water imaging. In preparation. (Cited on [2](#), [49](#).)
- [14] S. T. Whitaker, J-F. Nielsen, and J. A. Fessler. End-to-end scan parameter optimization for improved myelin water imaging. In Proc. Intl. Soc. Mag. Res. Med., page 1273, 2022. (Cited on [2](#), [49](#), [50](#).)
- [15] S. T. Whitaker, J-F. Nielsen, and J. A. Fessler. Intravoxel B0 corrected image reconstruction with RF prephasing. In Proc. Intl. Soc. Mag. Res. Med., page 2340, 2022. (Cited on [2](#), [87](#).)
- [16] N. Murthy, J-F. Nielsen, S. T. Whitaker, M. W. Haskell, S. D. Swanson, N. Seiberlich, and J. A. Fessler. Quantifying myelin water exchange using optimized bSSFP sequences. In Proc. Intl. Soc. Mag. Res. Med., page 2068, 2022. (Cited on [2](#).)
- [17] Z-P. Liang and P. C. Lauterber. Principles of magnetic resonance imaging. IEEE, New York, 2000. (Cited on [3](#), [4](#).)
- [18] H. M. McConnell. Reaction rates by nuclear magnetic resonance. J. of Chemical Phys., 28(3):430–31, March 1958. (Cited on [7](#), [23](#), [52](#), [70](#).)
- [19] H. Sun, J. A. Fessler, D. C. Noll, and J-F. Nielsen. Steady-state functional MRI using spoiled small-tip fast recovery (STFR) imaging. Mag. Res. Med., 73(2):536–43, February 2015. (Cited on [11](#), [21](#), [102](#).)
- [20] A. Rahimi and B. Recht. Random features for large-scale kernel machines. In NIPS, pages 1177–84, 2007. (Cited on [12](#), [26](#).)
- [21] H. Y. Carr and E. M. Purcell. Effects of diffusion on free precession in nuclear magnetic resonance experiments. Phys. Rev., 94(3):630–8, May 1954. (Cited on [13](#), [23](#).)
- [22] T. Prasloski, B. Mädler, Q-S. Xiang, A. MacKay, and C. Jones. Applications of stimulated echo correction to multicomponent T2 analysis. Mag. Res. Med., 67(6):1803–14, June 2012. (Cited on [15](#), [19](#), [27](#).)
- [23] W. C. Chen, S. Foxley, and K. L. Miller. Detecting microstructural properties of white matter based on compartmentalization of magnetic susceptibility. NeuroImage, 70:1–9, April 2013. (Cited on [17](#), [54](#).)

- [24] K. L. West, N. D. Kelm, R. P. Carson, D. F. Gochberg, K. C. Ess, and M. D. Does. Myelin volume fraction imaging with MRI. NeuroImage, 182:511–21, November 2018. (Cited on 19.)
- [25] T. Prasloski, A. Rauscher, A. L. MacKay, M. Hodgson, I. M. Vavasour, C. Laule, and B. Mädler. Rapid whole cerebrum myelin water imaging using a 3D GRASE sequence. NeuroImage, 63(1):533–9, October 2012. (Cited on 19, 49.)
- [26] J. Zhang, I. Vavasour, S. Kolind, B. Baumeister, A. Rauscher, and A. L. MacKay. Advanced myelin water imaging techniques for rapid data acquisition and long T2 component measurements. In Proc. Intl. Soc. Mag. Res. Med., page 824, 2015. (Cited on 19, 39.)
- [27] S. C. L. Deoni, B. K. Rutt, T. Arun, C. Pierpaoli, and D. K. Jones. Gleaning multi-component T1 and T2 information from steady-state imaging data. Mag. Res. Med., 60(6):1372–87, December 2008. (Cited on 19, 49.)
- [28] J. Zhang, S. H. Kolind, C. Laule, and A. L. MacKay. Comparison of myelin water fraction from multiecho T2 decay curve and steady-state methods. Mag. Res. Med., 73(1):223–32, January 2015. (Cited on 19.)
- [29] D. J. West, R. P. A. G. Teixeira, T. C. Wood, J. V. Hajnal, J-D. Tournier, and S. J. Malik. Inherent and unpredictable bias in multi-component DESPOT myelin water fraction estimation. NeuroImage, 195(1):78–88, July 2019. (Cited on 19.)
- [30] C. L. Lankford and M. D. Does. On the inherent precision of mcDESPOT. Mag. Res. Med., 69(1):127–36, January 2013. (Cited on 19, 44.)
- [31] D. Hwang, D-H. Kim, and Y. P. Du. In vivo multi-slice mapping of myelin water content using T2* decay. NeuroImage, 52(1):198–204, August 2010. (Cited on 19, 49.)
- [32] C. Lenz, M. Klarhöfer, and K. Scheffler. Feasibility of in vivo myelin water imaging using 3d multigradient-echo pulse sequences. Mag. Res. Med., 68(2):523–8, August 2012. (Cited on 19, 49.)
- [33] Y. Nam, J. Lee, D. Hwang, and D-H. Kim. Improved estimation of myelin water fraction using complex model fitting. NeuroImage, 116:214–221, March 2015. (Cited on 19, 20, 49.)
- [34] G. Nataraj, J-F. Nielsen, and J. A. Fessler. Myelin water fraction estimation from optimized steady-state sequences using kernel ridge regression. In Proc. Intl. Soc. Mag. Res. Med., page 5076, 2017. (Cited on 19, 49, 50.)
- [35] G. Nataraj, J-F. Nielsen, M. Gao, and J. A. Fessler. Fast, precise myelin water quantification using DESS MRI and kernel learning. Technical report, Univ. of Michigan, 2018. <http://arxiv.org/abs/1809.08908>. (Cited on 19, 49, 50.)

- [36] G. Nataraj, M. Gao, J-F. Nielsen, and J. A. Fessler. Kernel regression for fast myelin water imaging. In ISMRM Workshop on Machine Learning Part 2, page 65, 2018. (Cited on 19, 49, 50.)
- [37] K. D. Harkins, A. N. Dula, and M. D. Does. Effect of intercompartmental water exchange on the apparent myelin water fraction in multiexponential T2 measurements of rat spinal cord. Mag. Res. Med., 67(3):793–800, March 2012. (Cited on 20.)
- [38] J. Hennig. Multiecho imaging sequences with low refocusing flip angles. J. Mag. Res., 88(3):397–407, July 1988. (Cited on 24.)
- [39] G. McGibney and M. R. Smith. An unbiased signal-to-noise ratio measure for magnetic resonance images. Med. Phys., 20(4):1077–8, July 1993. (Cited on 24.)
- [40] G. Nataraj, J-F. Nielsen, and J. A. Fessler. Optimizing MR scan design for model-based T1, T2 estimation from steady-state sequences. IEEE Trans. Med. Imag., 36(2):467–77, February 2017. (Cited on 24, 50, 51.)
- [41] S. M. Kay. Fundamentals of statistical signal processing: Estimation theory. Prentice-Hall, New York, 1993. (Cited on 24.)
- [42] J. A. Jones, P. Hodgkinson, A. L. Barker, and P. J. Hore. Optimal sampling strategies for the measurement of spin-spin relaxation times. J. Mag. Res. B, 113(1):25–34, October 1996. (Cited on 25, 50.)
- [43] T. Voigt, K. Nehrke, O. Doessel, and U. Katscher. T1 corrected B1 mapping using multi-TR gradient echo sequences. Mag. Res. Med., 64(3):725–33, 2010. (Cited on 25, 50.)
- [44] J. Asslander, R. Lattanzi, D. K. Sodickson, and M. A. Cloos. Optimized quantification of spin relaxation times in the hybrid state. Mag. Res. Med., 82(4):1385–1397, 2019. (Cited on 25, 50.)
- [45] G. Nataraj. Advances in quantitative MRI: acquisition, estimation, and application. PhD thesis, Univ. of Michigan, Ann Arbor, MI, 48109-2122, Ann Arbor, MI, 2018. (Cited on 25.)
- [46] S. G. Johnson. The NLOpt nonlinear-optimization package. <https://github.com/JuliaOpt/NLOpt.jl>. (Cited on 27.)
- [47] A. H. G. Rinnooy Kan and G. T. Timmer. Stochastic global optimization methods part I: Clustering methods. Mathematical Programming, 39(1):27–56, September 1987. (Cited on 27.)
- [48] A. H. G. Rinnooy Kan and G. T. Timmer. Stochastic global optimization methods part II: Multi level methods. Mathematical Programming, 39(1):57–78, September 1987. (Cited on 27.)

- [49] D. Kraft. Algorithm 733: TOMP-Fortran modules for optimal control calculations. ACM Trans. Math. Softw., 20(3):262–81, September 1994. (Cited on 27.)
- [50] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah. Julia: A fresh approach to numerical computing. SIAM Review, 59(1):65–98, 2017. (Cited on 27, 52, 55, 91.)
- [51] L. I. Sacolick, F. Wiesinger, I. Hancu, and M. W. Vogel. B1 mapping by Bloch-Siegert shift. Mag. Res. Med., 63(5):1315–22, May 2010. (Cited on 30.)
- [52] J-F. Nielsen and D. C. Noll. TOPPE: A framework for rapid prototyping of MR pulse sequences. Mag. Res. Med., 79(6):3128–34, June 2018. (Cited on 30, 57.)
- [53] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans. Design and construction of a realistic digital brain phantom. IEEE Trans. Med. Imag., 17(3):463–8, June 1998. (Cited on 34, 57, 92, 106.)
- [54] Eva Alonso-Ortiz, Ives R Levesque, and G Bruce Pike. Impact of magnetic susceptibility anisotropy at 3 T and 7 T on T2*-based myelin water fraction imaging. NeuroImage, 182:370–378, September 2018. (Cited on 38.)
- [55] S. Mandava, M. B. Keerthivasan, Z. Li, D. R. Martin, M. I. Altbach, and A. Bilgin. Accelerated MR parameter mapping with a union of local subspaces constraint. Mag. Res. Med., 80(6):2744–58, December 2018. (Cited on 41, 69.)
- [56] X. Wang, V. Roeloffs, J. Klosowski, Z. Tan, D. Voit, M. Uecker, and J. Frahm. Model-based T1 mapping with sparsity constraints using single-shot inversion-recovery radial FLASH. Mag. Res. Med., 79(2):730–40, February 2018. (Cited on 41, 69.)
- [57] B. Zhao, J. P. Haldar, C. Liao, D. Ma, Y. Jiang, M. A. Griswold, K. Setsompop, and L. L. Wald. Optimal experiment design for magnetic resonance fingerprinting: Cramér-Rao bound meets spin dynamics. IEEE Trans. Med. Imag., 38(3):844–61, March 2019. (Cited on 50.)
- [58] A. Lahiri, J. A. Fessler, and L. Hernandez-Garcia. Optimizing MRF-ASL scan design for precise quantification of brain hemodynamics using neural network regression. Mag. Res. Med., 83(6):1979–91, June 2020. (Cited on 50.)
- [59] O. Cohen, M. Sarracanie, B. D. Armstrong, J. L. Ackerman, and M. S. Rosen. Magnetic resonance fingerprinting trajectory optimization. In Proc. Intl. Soc. Mag. Res. Med., page 27, 2014. (Cited on 50.)
- [60] O. Cohen and M. S. Rosen. Algorithm comparison for schedule optimization in MR fingerprinting. Mag. Res. Im., 41:15–21, September 2017. (Cited on 50.)
- [61] J. Bacca, T. Gelvez, and H. Arguello. Deep coded aperture design: an end-to-end approach for computational imaging tasks, May 2021. <https://arxiv.org/abs/2105.03390>. (Cited on 50.)

- [62] B. Zhu, J. Liu, N. Koonjoo, B. R. Rosen, and M. S. Rosen. AUTOMated pulse SE-Quence generation (AUTOSEQ) and neural network decoding for fast quantitative MR parameter measurement using continuous and simultaneous RF transmit and receive. In Proc. Intl. Soc. Mag. Res. Med., page 1090, 2019. (Cited on 50, 68.)
- [63] S. P. Jordan, S. Hu, I. Rozada, D. F. McGivney, R. Boyacıoğlu, D. C. Jacob, S. Huang, M. Beverland, H. G. Katzgraber, M. Troyer, M. A. Griswold, and D. Ma. Automated design of pulse sequences for magnetic resonance fingerprinting using physics-inspired optimization. Proc. Natl. Acad. Sci., 118(40):1–8, October 2021. (Cited on 50, 68, 106.)
- [64] S. T. Whitaker. BlochSim.jl. <https://github.com/StevenWhitaker/BlochSim.jl>. Version: v0.5.1. (Cited on 52.)
- [65] P. K. Mogensen and A. N. Riseth. Optim: A mathematical optimization package for Julia. Journal of Open Source Software, 3(24):615, 2018. (Cited on 55.)
- [66] M. Innes. Zygote.jl. <https://github.com/FluxML/Zygote.jl>. Version: v0.6.36. (Cited on 55.)
- [67] J. Revels, M. Lubin, and T. Papamarkou. Forward-mode automatic differentiation in Julia, 2016. <https://arxiv.org/abs/1607.07892>. (Cited on 55.)
- [68] Y. Wang, P. van Gelderen, J. A. de Zwart, and J. H. Duyn. B₀-field dependence of MRI T₁ relaxation in human brain. NeuroImage, 213:116700, June 2020. (Cited on 70, 71.)
- [69] G. H. Glover. 3D z-shim method for reduction of susceptibility effects in BOLD fMRI. Mag. Res. Med., 42(3):290–9, September 1999. (Cited on 87.)
- [70] B. P. Sutton, D. C. Noll, and J. A. Fessler. Compensating for within-voxel susceptibility gradients in BOLD fMRI. In Proc. Intl. Soc. Mag. Res. Med., page 349, 2004. (Cited on 87.)
- [71] Y. Zhuo and B. P. Sutton. Iterative image reconstruction model including susceptibility gradients combined with Z-shimming gradients in fMRI. In embc, pages 5721–4, 2009. (Cited on 87.)
- [72] F. Lam and B. P. Sutton. Intravoxel B₀ inhomogeneity corrected reconstruction using a low-rank encoding operator. Mag. Res. Med., 84(2):885–94, August 2020. (Cited on 87, 88, 89, 90, 91.)
- [73] J. A. Fessler and D. C. Noll. Model-based MR image reconstruction with compensation for through-plane field inhomogeneity. In Proc. IEEE Intl. Symp. Biomed. Imag., pages 920–3, 2007. Invited paper. (Cited on 89.)
- [74] J. A. Fessler. MIRT.jl. <https://github.com/JeffFessler/MIRT.jl>. Version: v0.15.0. (Cited on 91.)
- [75] J. Asslander, S. J. Glaser, and Juergen Hennig. Spin echoes in the regime of weak dephasing. Mag. Res. Med., 75(1):150–60, January 2016. (Cited on 101.)

- [76] T. Luo, D. C. Noll, J. A. Fessler, and J-F. Nielsen. Joint design of RF and gradient waveforms via auto-differentiation for 3D tailored excitation in MRI. IEEE Trans. Med. Imag., 2021. (Cited on [101](#).)
- [77] C. Yip, D. Yoon, V. Olafsson, S. Lee, W. A. Grissom, J. A. Fessler, and D. C. Noll. Spectral-spatial pulse design for through-plane phase precompensatory slice selection in T2*-weighted functional MRI. Mag. Res. Med., 61(5):1137–47, May 2009. (Cited on [101](#).)
- [78] M. Lustig, D. Donoho, and J. M. Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. Mag. Res. Med., 58(6):1182–95, December 2007. (Cited on [103](#).)
- [79] A. Majumdar and R. K. Ward. Accelerating multi-echo T2 weighted MR imaging: Analysis prior group-sparse optimization. J. Mag. Res., 210(1):90–7, May 2011. (Cited on [103](#).)
- [80] B. Zhao, W. Lu, T. K. Hitchens, F. Lam, C. Ho, and Z-P. Liang. Accelerated MR parameter mapping with low-rank and sparsity constraints. Mag. Res. Med., 74(2):489–98, August 2015. (Cited on [103](#).)
- [81] E. Soubies, L. Blanc-Feraud, and G. Aubert. A continuous exact ℓ_0 penalty (CEL0) for least squares regularized problem. SIAM J. Imaging Sci., 8(3):1607–39, 2015. (Cited on [103](#).)
- [82] A. Bechensteen, L. Blanc-Feraud, and G. Aubert. New l2-l0 algorithm for single-molecule localization microscopy. Biomed. Opt. Express, 11(2):1153–74, February 2020. (Cited on [103](#).)
- [83] D. Ma, V. Gulani, N. Seiberlich, K. Liu, J. L. Sunshine, J. L. Duerk, and M. A. Griswold. Magnetic resonance fingerprinting. Nature, 495:187–93, March 2013. (Cited on [103](#), [105](#), [106](#).)
- [84] F. Liu, L. Feng, and R. Kijowski. MANTIS: Model-Augmented Neural neTwork with Incoherent k-space Sampling for efficient MR parameter mapping. Mag. Res. Med., 82(1), July 2019. (Cited on [104](#).)
- [85] B. Kang, B. Kim, M. Schar, H. Park, and H. Y. Heo. Unsupervised deep learning-based magnetization transfer contrast (MTC) MR fingerprinting and CEST MRI. In Proc. Intl. Soc. Mag. Res. Med., page 510, 2020. (Cited on [104](#).)
- [86] B. Kang, B. Kim, M. Schär, H. Park, and H. Heo. Unsupervised learning for magnetization transfer contrast MR fingerprinting: Application to CEST and nuclear Overhauser enhancement imaging. Mag. Res. Med., 85(4):2040–54, April 2021. (Cited on [104](#).)
- [87] S. C. L. Deoni, L. Matthews, and S. H. Kolind. One component? Two components? Three? The effect of including a nonexchanging “free” water component in multicomponent driven equilibrium single pulse observation of T1 and T2. Mag. Res. Med., 70(1):147–54, July 2013. (Cited on [107](#).)

- [88] O. Bieri and K. Scheffler. On the origin of apparent low tissue signals in balanced SSFP. Mag. Res. Med., 56(5):1067–74, November 2006. (Cited on [109](#).)