

8-2019

## An Information Theory Model for Optimizing Quantitative Magnetic Resonance Imaging Acquisitions

Drew Mitchell

Follow this and additional works at: [https://digitalcommons.library.tmc.edu/utgsbs\\_dissertations](https://digitalcommons.library.tmc.edu/utgsbs_dissertations)



Part of the [Applied Mathematics Commons](#), and the [Medicine and Health Sciences Commons](#)

---

### Recommended Citation

Mitchell, Drew, "An Information Theory Model for Optimizing Quantitative Magnetic Resonance Imaging Acquisitions" (2019). *The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access)*. 926.

[https://digitalcommons.library.tmc.edu/utgsbs\\_dissertations/926](https://digitalcommons.library.tmc.edu/utgsbs_dissertations/926)

This Dissertation (PhD) is brought to you for free and open access by the The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences at DigitalCommons@TMC. It has been accepted for inclusion in The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access) by an authorized administrator of DigitalCommons@TMC. For more information, please contact [digitalcommons@library.tmc.edu](mailto:digitalcommons@library.tmc.edu).

AN INFORMATION THEORY MODEL FOR OPTIMIZING  
QUANTITATIVE MAGNETIC RESONANCE IMAGING ACQUISITIONS

by

Drew Palmer Mitchell, B.Sc.

APPROVED:

---

David Thomas Alfonso Fuentes, Ph.D., Advisor

---

Ken-Pin Hwang, Ph.D.

---

James Bankson, Ph.D.

---

Erik N.K. Cressman, M.D., Ph.D.

---

James Long, Ph.D.

APPROVED:

---

Dean, The University of Texas

MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences

AN INFORMATION THEORY MODEL FOR OPTIMIZING  
QUANTITATIVE MAGNETIC RESONANCE IMAGING ACQUISITIONS

A  
DISSERTATION

Presented to the Faculty of  
The University of Texas  
Health Science Center at Houston  
and  
The University of Texas  
MD Anderson Cancer Center  
Graduate School of Biomedical Sciences  
in Partial Fulfillment  
of the Requirements  
for the Degree of  
DOCTOR OF PHILOSOPHY

by  
Drew Palmer Mitchell, B.Sc.  
Houston, Texas

August, 2019

©Copyright by Drew Palmer Mitchell, 2019.  
All rights reserved.

*To my dad. Your memory reverberates throughout these pages and all else I do. You are dearly missed.*

# Acknowledgments

I thank my advisor, David Fuentes, for his mentorship throughout this challenging and rewarding process. I thank the members of my advisory committee for helping me refine my work and for their patience while I learned how to communicate it. I thank Ken Hwang in particular for his time and invaluable assistance collecting data for this project. I thank my medical physics peers for their camaraderie over the years of courses, rotations, and research.

I thank my wife, Julika Kaplan, for her constant love and encouragement. She is an inspiration and an unrelenting force of good. I thank my mother, Melanie Mitchell, who has always been first in supporting me through every difficulty and every success in my life. I thank my father, Randall Mitchell. I aspire to be as genuine and kind as he always was. I thank my both of my parents for enabling me to freely pursue my academic interests throughout my life and for encouraging my curiosity always. I thank my brothers, Matthew and Kyle Mitchell. They are true friends. I could not have done this without my family. They have my eternal love and gratitude.

AN INFORMATION THEORY MODEL FOR OPTIMIZING QUANTITATIVE MAGNETIC  
RESONANCE IMAGING ACQUISITIONS

Drew Palmer Mitchell, B.Sc.

Advisory Professor: David Thomas Alfonso Fuentes, Ph.D.

Quantitative magnetic resonance imaging (qMRI) is a powerful group of imaging techniques with a growing number of clinical applications, including synthetic image generation in post-processing, automatic segmentation, and diagnosis of disease from quantitative parameter values. Currently, acquisition parameter selection is performed empirically for quantitative MRI. Tuning parameters for different scan times, tissues, and resolutions requires some measure of trial and error. There is an opportunity to quantitatively optimize these acquisition parameters in order to maximize image quality and the reliability of the previously mentioned methods which follow image acquisition.

The objective of this work is to introduce and evaluate a quantitative method for selecting parameters that minimize image variability. An information theory framework was developed for this purpose and applied to a 3D-quantification using an interleaved Look-Locker acquisition sequence with T2 preparation pulse (3D-QALAS) signal model for synthetic MRI. In this framework, mutual information is used to measure the information gained by a measurement as a function of acquisition parameters, quantifying the information content of the acquisition parameters and allowing informed parameter selection.

The information theory framework was tested on synthetic data generated from a representative mathematical phantom, measurements acquired on a qMRI multiparametric imaging standard phantom, and *in vivo* measurements in a human brain. The application of this information theory framework resulted in successful parameter optimization with respect to mutual information. Both the phantom and *in vivo* measurements showed that higher mutual information calculated by the model correlated with smaller standard deviation in the reconstructed parametric maps.

With this framework, optimal acquisition parameters can be selected to improve image quality, image repeatability, or scan time. This method could reduce the time and labor necessary to achieve images of the desired quality. Making an informed acquisition parameter selection reduces uncertainty in the imaging output and optimizes information gain within the bounds of clinical constraints.

# Contents

<b>Acknowledgments</b>	<b>v</b>
<b>Abstract</b>	<b>vi</b>
<b>Table of Contents</b>	<b>vii</b>
<b>List of Figures</b>	<b>x</b>
<b>List of Tables</b>	<b>xvi</b>
<b>Abbreviations and symbol names</b>	<b>xvii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Hypothesis . . . . .	2
1.2 Specific Aims . . . . .	3
1.3 Dissertation Organization . . . . .	5
<b>2 Quantitative MRI</b>	<b>6</b>
2.1 Introduction . . . . .	6
2.2 Relaxometry . . . . .	6
2.2.1 QRAPMASTER . . . . .	7
2.2.2 3D-QALAS . . . . .	8
2.3 Post-Processing Utility . . . . .	8
2.3.1 Tissue Segmentation . . . . .	9
2.3.2 Synthetic Image Synthesis . . . . .	9
2.3.3 Magnetic Resonance Spectroscopy . . . . .	10
2.3.4 Physical Models . . . . .	10
2.4 Clinical Applications . . . . .	11
2.4.1 Brain Metastases . . . . .	11
2.4.2 Multiple Sclerosis . . . . .	12
2.4.3 Other Diagnostic Utilities . . . . .	12
2.5 Performance of Quantitative Imaging . . . . .	13
2.6 Application of Information Theory . . . . .	15
<b>3 Information Theory</b>	<b>16</b>
3.1 Introduction . . . . .	16
3.2 Mutual Information . . . . .	17
3.3 Information Theory Metrics . . . . .	22
3.3.1 Mutual Information . . . . .	22



3.3.2	Joint Mutual Information . . . . .	24
3.3.3	Conditional Mutual Information . . . . .	25
3.3.4	Conditional Mutual Information Relationships . . . . .	26
3.4	Numeric Methods . . . . .	26
<b>4</b>	<b>Development and Verification of an Information Model</b>	<b>30</b>
4.1	Introduction . . . . .	30
4.2	Theory . . . . .	31
4.2.1	General Problem Statement . . . . .	31
4.2.2	Gauss-Hermite Quadrature . . . . .	34
4.2.3	Linear Approximation . . . . .	35
4.2.4	Nonlinear Computation . . . . .	37
4.3	Methods . . . . .	39
4.3.1	Verification of Information Theory Framework . . . . .	39
4.3.2	Justification of Model Assumptions for 3D-QALAS . . . . .	43
4.4	Results . . . . .	44
4.4.1	Linear Approximation . . . . .	44
4.4.2	Nonlinear Model . . . . .	47
4.4.3	Justification of Model Assumptions for 3D-QALAS . . . . .	50
4.5	Discussion . . . . .	55
<b>5</b>	<b>Information Quantification of Acquisition Parameters for 3D-QALAS in Phantom</b>	<b>59</b>
5.1	Introduction . . . . .	59
5.2	Theory . . . . .	60
5.2.1	3D-QALAS Signal Model . . . . .	60
5.2.2	3D-QALAS Operator . . . . .	62
5.2.3	Mutual Information . . . . .	63
5.3	Methods . . . . .	64
5.3.1	Synthetic Phantom Definition . . . . .	65
5.3.2	Phantom Acquisitions . . . . .	69
5.4	Results . . . . .	71
5.4.1	Synthetic Model . . . . .	71
5.4.2	ISMIRM NIST Phantom . . . . .	74
5.5	Discussion . . . . .	77
<b>6</b>	<b>Conditional Mutual Information Generalization and In Vivo Model Validation</b>	<b>80</b>
6.1	Introduction . . . . .	80
6.2	Theory . . . . .	81
6.2.1	Conditional Mutual Information Relationships . . . . .	81
6.2.2	Mutual Information for Jointly Gaussian Measurements . . . . .	83
6.2.3	Mutual Information for Independent Subsampling Masks . . . . .	85
6.3	Methods . . . . .	88
6.3.1	Synthetic Model . . . . .	89
6.3.2	Phantom Acquisitions . . . . .	89
6.3.3	In Vivo Acquisitions . . . . .	93
6.4	Results . . . . .	95

6.4.1	Synthetic Model . . . . .	95
6.4.2	Phantom Acquisitions . . . . .	95
6.4.3	In Vivo Acquisitions . . . . .	100
6.5	Discussion . . . . .	104
<b>7</b>	<b>Conclusion</b>	<b>108</b>
<b>A</b>	<b>Important Terms</b>	<b>113</b>
<b>B</b>	<b>Mathematical Symbols</b>	<b>115</b>
<b>C</b>	<b>Mathematical Derivations</b>	<b>118</b>
C.1	Chapter 3 Derivations . . . . .	118
C.1.1	Mutual Information . . . . .	118
C.1.2	Joint Mutual Information . . . . .	119
C.1.3	Conditional Mutual Information . . . . .	120
C.1.4	Conditional Mutual Information Relationships . . . . .	120
C.2	Chapter 4 and 5 Derivations . . . . .	122
C.2.1	Linear Approximation . . . . .	122
C.2.2	Nonlinear Computation . . . . .	125
C.3	Chapter 6 Derivations . . . . .	129
C.3.1	Conditional Mutual Information for Jointly Gaussian Measurements	129
C.3.2	Joint Mutual Information for Independent Subsampling Masks . .	134
C.3.3	Conditional Mutual Information for Independent Subsampling Masks	135
C.3.4	Form of the Conditional Probability Distribution . . . . .	137
	<b>Bibliography</b>	<b>142</b>
	<b>Vita</b>	<b>172</b>

# List of Figures

3.1	Fair coin flip example. Two coins are flipped sequentially, both with equal probability of landing on either side. The first coin provides no information about the second, so mutual information between the events is zero. . . . .	18
3.2	Unfair coin flip example. Two coins are flipped sequentially, both with 75% chance of landing on one side. The favored side is not known beforehand. The first coin flip now provides some information about the second, so mutual information between the two events is greater than zero. . . . .	18
3.3	Representation of a suboptimal 3D-QALAS acquisition. (A) Longitudinal magnetization in a 3D-QALAS acquisition with suboptimal acquisition parameters. Each of the five measurements is shown as a red x. The machine noise that corrupts the measurements is represented as a set of error bars on each measurement. (B) Venn diagram showing the relationships between entropies resulting from the information theory modeling of the acquisition depicted in (A). In Bayesian terms, the entropy of the prior distribution of quantitative parameters of the image subject is illustrated by the red outline. The entropy of the posterior distribution of quantitative parameters of the image subject is illustrated by the red filled portion of the circle. This entropy is reduced increasing mutual information (the purple filled portion of the circle). (C) Synthetic reconstruction of the T1 map (right) from the previous suboptimal acquisition in (A). The probability density functions (left) show the spread of reconstructed M0, T1, and T2 values in gray matter for the synthetic reconstruction. . . . .	21
3.4	Representation of an optimal 3D-QALAS acquisition. (A) Longitudinal magnetization in a 3D-QALAS acquisition with optimal acquisition parameters. Each of the five measurements is shown as a red x. The machine noise that corrupts the measurements is represented as a set of error bars on each measurement. (B) Venn diagram showing the relationships between entropies resulting from the information theory modeling of the acquisition depicted in (A). In Bayesian terms, the entropy of the prior distribution of quantitative parameters of the image subject is illustrated by the red outline. The entropy of the posterior distribution of quantitative parameters of the image subject is illustrated by the red filled portion of the circle. This entropy is reduced increasing mutual information (the purple filled portion of the circle). (C) Synthetic reconstruction of the T1 map (right) from the previous optimal acquisition in (A). The probability density functions (left) show the spread of reconstructed M0, T1, and T2 values in gray matter for the synthetic reconstruction. The dashed lines show the probability distribution functions resulting from the suboptimal acquisition in Figure 3.3. . . . .	23

4.1	(A) Plot of the well-known relationship between the Ernst angle and the ratio of TR and T1 for a spoiled GRE sequence. (B) Plot of the signal intensity in a spoiled GRE sequence as a function of flip angle for characteristic values of two brain tissues. T1 for white matter was set to 1000 ms. T1 for cerebrospinal fluid was set to 4000 ms. TR for the acquisition was selected to be 500 ms. . . . .	40
4.2	Flow chart for the mutual information-based optimization of flip angle in a spoiled GRE pulse sequence. . . . .	41
4.3	(A) Comparison between SNR calculated from a spoiled GRE signal model and mutual information calculated from an information theory model of the spoiled GRE sequence. Calculations are performed for a single scenario where T1 is 1000 ms and TR is 500 ms. The signal model SNR is plotted against the left ordinate, and the information model mutual information is plotted against the right ordinate. The maxima of both are located at approximately 53°. (B) Correlation between SNR and mutual information. SNR and mutual information are positively correlated. . . . .	45
4.4	(A) Theoretical Ernst angle calculated over a range of T1 and TR values. Ernst angle values are calculated from the deterministic equation which maximizes spoiled GRE signal intensity for the flip angle. (B) Optimal flip angle selected by linear approximation of information gain from spoiled GRE acquisitions over a range of T1 and TR values. Information gain is quantified by mutual information between measurement and tissue property uncertainties. Greater information gain corresponds to greater measurement SNR. . . . .	45
4.5	Relative error between analytical Ernst angle and flip angle optimization using information theory. . . . .	46
4.6	The regression line between 6561 measurement pairs of the analytical Ernst angle and the linear information-optimized flip angle in a spoiled GRE pulse sequence over a range of T1 and TR values. Each measurement pair is plotted with a red dot. The regression line is shown in black. Fit parameters and goodness of fit metrics are displayed in the upper left corner. . . . .	47
4.7	The Bland Altman plot of 6561 measurement pairs of the analytical Ernst angle and the linear information-optimized flip angle in a spoiled GRE pulse sequence over a range of T1 and TR values. Each measurement pair is plotted with a red dot. The mean measurement difference, $d$ , is shown as a black line. The bounds of the 95% confidence interval assuming normally distributed measurement differences are shown as black dash-dot lines. . . . .	48
4.8	Histogram of differences between measurement pairs of the analytical Ernst angle and the linear information-optimized flip angle in a spoiled GRE pulse sequence over a range of T1 and TR values. . . . .	48
4.9	(A) Theoretical Ernst angle calculated over a range of T1 and TR values. (B) Optimal flip angle selected by nonlinear model of information gain from spoiled GRE acquisitions over a range of T1 and TR values. . . . .	49

- 4.10 The regression line between 6561 measurement pairs of the analytical Ernst angle and the nonlinear information-optimized flip angle in a spoiled GRE pulse sequence over a range of T1 and TR values. Each measurement pair is plotted with a red dot. The regression line is shown in black. Fit parameters and goodness of fit metrics are displayed in the upper left corner. . . . . 50
- 4.11 The Bland Altman plot of 6561 measurement pairs of the analytical Ernst angle and the nonlinear information-optimized flip angle in a spoiled GRE pulse sequence over a range of T1 and TR values. Each measurement pair is plotted with a red dot. The mean measurement difference,  $d$ , is shown as a black line. The bounds of the 95% confidence interval are shown as black dash-dot lines. . . . . 51
- 4.12 Assumed normal prior distributions of M0, T1, and T2 for gray matter, white matter, and cerebrospinal fluid. Ellipsoidal nested sampling is used to sample the posterior distributions resulting from these prior distributions. Diagonal entries on the correlation matrix show distributions of each combination of parametric map value and tissue type. Non-diagonal lower-triangular entries show a two-dimension scatter plot of the joint distribution between the parameters labeled on the same row and column. Non-diagonal upper-triangular entries show a two-dimensional histogram of the joint distribution between the parameters labeled on the same row and column. . . . . 52
- 4.13 Posterior distributions of M0, T1, and T2 for gray matter, white matter, and cerebrospinal fluid resulting from ellipsoidal nested sampling and normal prior distributions shown in Figure 4.12. Diagonal entries on the correlation matrix show distributions of each combination of parametric map value and tissue type. Non-diagonal lower-triangular entries show a two-dimension scatter plot of the joint distribution between the parameters labeled on the same row and column. Non-diagonal upper-triangular entries show a two-dimensional histogram of the joint distribution between the parameters labeled on the same row and column. . . . . 52
- 4.14 Assumed uniform prior distributions of M0, T1, and T2 for gray matter, white matter, and cerebrospinal fluid. Ellipsoidal nested sampling is used to sample the posterior distributions resulting from these prior distributions. Diagonal entries on the correlation matrix show distributions of each combination of parametric map value and tissue type. Non-diagonal lower-triangular entries show a two-dimension scatter plot of the joint distribution between the parameters labeled on the same row and column. Non-diagonal upper-triangular entries show a two-dimensional histogram of the joint distribution between the parameters labeled on the same row and column. . . . . 54
- 4.15 Posterior distributions of M0, T1, and T2 for gray matter, white matter, and cerebrospinal fluid resulting from ellipsoidal nested sampling and uniform prior distributions shown in Figure 4.14. Diagonal entries on the correlation matrix show distributions of each combination of parametric map value and tissue type. Non-diagonal lower-triangular entries show a two-dimension scatter plot of the joint distribution between the parameters labeled on the same row and column. Non-diagonal upper-triangular entries show a two-dimensional histogram of the joint distribution between the parameters labeled on the same row and column. . . . . 54

4.16	Correlation matrix for M0, T1, and T2 values measured from in vivo data in gray matter, white matter, and cerebrospinal fluid. Diagonal entries show distributions of each combination of parametric map value and tissue type. Non-diagonal lower-triangular entries show a two-dimension scatter plot of the joint distribution between the parameters labeled on the same row and column. Non-diagonal upper-triangular entries show a two-dimensional histogram of the joint distribution between the parameters labeled on the same row and column. . . . .	56
5.1	Diagram of QALAS acquisition sequence. The longitudinal magnetization is shown in red as a function of time. The five acquisition and delay times are labeled at the bottom. . . . .	62
5.2	Flowchart for the mutual information-based optimization of acquisition parameters in a synthetic model. . . . .	65
5.3	Flowchart for the application of the synthetic model to real-world measurements in the System Standard Model 130 phantom is shown on the right. . . . .	66
5.4	A two-dimensional axial slice of the tissue labels used to define the synthetic phantom. Tissue label 1 is gray matter, 2 is white matter, and 3 is cerebrospinal fluid. . . . .	68
5.5	Measurements of the ISMRM NIST phantom with each of the 42 phantom elements segmented. The three layers of elements (PD, T1, and T2) are labeled. There are 14 elements in each layer. . . . .	70
5.6	Example T1 parametric map reconstructed from measurements of the ISMRM NIST phantom. . . . .	70
5.7	Evolution of acquisition parameters over the course of an optimization (A), and the optimization objective function, mutual information (B). . . . .	72
5.8	The red histogram shows the final distribution resulting from a measurement at the optimal location in parameter space. This is the location where mutual information and, consequently, the information gain from a measurement are greatest. The blue histogram shows the final distribution from a measurement at the initial point in parameter space. . . . .	73
5.9	We simulate a reconstruction using a measurement with the initial, non-optimized acquisition parameters. The absolute error between these parametric maps and the ground truth maps is shown on the top row. We simulate a second reconstruction using a measurement with the optimized acquisition parameters. The absolute error between these parametric maps and the ground truth is on the bottom row. . . . .	74
5.10	Distribution of reconstructed M0, T1, and T2 values in the 14 T2 elements of the System Standard Model 130 phantom for the second set of acquisition parameters. The first row shows the distribution of the absolute M0, T1, and T2 values. The second row shows the M0, T1, and T2 values normalized by the element mean. The red bars represents the median M0 value. The blue boxes are the middle two quartiles, and the black lines are the outer two quartiles. A red plus denotes an outlier. . . . .	75
5.11	Comparison of T1 mapping performance between a variable flip angle T1 mapping sequence and 3D-QALAS on the PD elements of the ISMRM NIST phantom. T1 means and standard deviations for these elements are not available in the phantom manual. . . . .	75

5.12	Comparison of T1 mapping performance between a variable flip angle T1 mapping sequence and 3D-QALAS on the T2 elements of the ISMRM NIST phantom. True T1 values use the T1 means and standard deviations listed for these elements in the phantom manual. . . . .	76
5.13	Comparison of T1 mapping performance between a variable flip angle T1 mapping sequence and 3D-QALAS on the T1 elements of the ISMRM NIST phantom. True T1 values use the T1 means and standard deviations listed for these elements in the phantom manual. . . . .	76
5.14	Reconstructed M0, T1, and T2 standard deviation in each of the 36 analyzed phantom elements for both scans as a function of mutual information. For reconstructed M0, T1, and T2 values, standard deviation is negatively correlated with mutual information across all data sets. . . . .	77
6.1	High-resolution tissue label maps of ten synthetic brain models. . . . .	89
6.2	Low-resolution tissue label maps of ten synthetic brain models. . . . .	90
6.3	(A) Example longitudinal magnetization of ISMRM NIST phantom from high-resolution acquisition. (B) Longitudinal magnetization of same view from ISMRM NIST phantom from low-resolution acquisition. . . . .	90
6.4	(A) Optimization history of acquisition parameters. Five delay times are shown over the course of 14 optimization iterations. (B) Optimization history of mutual information over the course of 14 optimization iterations. Mutual information is the objective function to be maximized during optimization. . . . .	94
6.5	(A) Conditional mutual information between simulated high- and low-resolution acquisitions in the ten synthetic brain models (Figures 6.1 and 6.2) as a function of variable delay times. Delay times are fixed equal to one another. Positive CMI predicts an improvement in image quality, whereas negative CMI predicts a decrease in image quality. Optimum delay times are selected by maximizing CMI. Optimum values are marked by a red x for each synthetic brain model. (B) CMI between simulated high- and low-resolution acquisitions as a function of delay times, zoomed around optimum values. . . . .	96
6.6	Inter-scan differences in reconstructed M0 standard deviations as a function of conditional mutual information between low-resolution pre-scan and high-resolution scan. (A) Low-resolution pre-scan at scan 1 acquisition parameters and high-resolution scan at scan 2 acquisition parameters (Table 6.1). (B) Low-resolution pre-scan at scan 2 acquisition parameters and high-resolution scan at scan 1 acquisition parameters (Table 6.1). . . . .	97
6.7	Inter-scan differences in reconstructed T1 standard deviations as a function of conditional mutual information between low-resolution pre-scan and high-resolution scan. (A) Low-resolution pre-scan at scan 1 acquisition parameters and high-resolution scan at scan 2 acquisition parameters (Table 6.1). (B) Low-resolution pre-scan at scan 2 acquisition parameters and high-resolution scan at scan 1 acquisition parameters (Table 6.1). . . . .	98
6.8	Inter-scan differences in reconstructed T2 standard deviations as a function of conditional mutual information between low-resolution pre-scan and high-resolution scan. (A) Low-resolution pre-scan at scan 1 acquisition parameters and high-resolution scan at scan 2 acquisition parameters (Table 6.1). (B) Low-resolution pre-scan at scan 2 acquisition parameters and high-resolution scan at scan 1 acquisition parameters (Table 6.1). . . . .	99

6.9	Inter-element differences in reconstructed M0 (A), T1 (B), and T2 (C) standard deviations as a function of inter-element conditional mutual information. Positive CMI predicts decrease in reconstructed parametric map value standard deviation between different ISMRM NIST phantom elements in the same scan. Negative CMI predicts increase in standard deviation. . . . .	101
6.10	Standard deviations of simulated posterior distributions. Clinical default is obtained by sampling with clinical default acquisition parameters (scan 1). Clinical optimized is obtained by sampling with implemented optimal acquisition parameters (scan 2). Theoretical optimal is obtained by sampling with calculated optimal acquisition parameters that could not be feasible tested <i>in vivo</i> . Theoretical minimum is obtained by sampling with no biological uncertainty. . . . .	102
6.11	(A) Reconstructed parametric maps from acquisition with default clinical parameters (scan 1). (B) Reconstructed parametric maps from acquisition with optimized parameters (scan 2). . . . .	103
6.12	Change in standard deviation in several tissues between two <i>in vivo</i> scans. Negative change in standard deviation denotes an improvement in reconstruction going from the default clinical parameters to the optimized parameters. Analysis was performed in frontal lobe white matter (Fr. WM) and gray matter (Fr. GM), parietal lobe white matter (Pa. WM) and gray matter (Pa. GM), temporal lobe white matter (Tm. WM) and gray matter (Tm. GM), occipital lobe white matter (Oc. WM) and gray matter (Oc. GM), and cerebrospinal fluid (CSF). Change in standard deviation is shown separately for M0, T1, and T2 maps. . . . .	104



# List of Tables

4.1	Optimization parameters used in calculating information-optimized flip angle for both linear and nonlinear approaches. . . . .	43
5.1	Prior distributions for tissue properties (M0, T1, and T2) in three tissues (gray matter, white matter, and cerebrospinal fluid). . . . .	69
5.2	Acquisition parameter values for two scans performed on ISMRM NIST system phantom. <sup>a</sup> TD1 is the time between the T1 sensitizing pulse and the first acquisition, TD2 is the time between the end of the first acquisition and the start of the second acquisition, and so on. . . . .	69
6.1	Acquisition parameter values for two scans performed on ISMRM NIST system phantom. <sup>a</sup> TD1 is the time between the T1 sensitizing pulse and the first acquisition, TD2 is the time between the end of the first acquisition and the start of the second acquisition, and so on. . . . .	91
6.2	ISMRM NIST phantom proton density sphere properties. . . . .	91
6.3	ISMRM NIST phantom T1 contrast sphere properties at 3.0 T. . . . .	92
6.4	ISMRM NIST phantom T2 contrast sphere properties at 3.0 T. . . . .	92
6.5	Acquisition parameter values for <i>in vivo</i> 3D-QALAS acquisitions. <sup>a</sup> TD1 is the time between the T1 sensitizing pulse and the first acquisition, TD2 is the time between the end of the first acquisition and the start of the second acquisition, and so on. . . . .	93
6.6	The p-values from two-sample F-tests for equal variances between distributions from the default and optimized acquisitions. Entries outlined in green are significant at the $p < 0.05$ level. Entries outlined in red are nonsignificant at the $p < 0.05$ level. . . . .	103

# Abbreviations and symbol names

<b>2D</b>	Two-Dimensional
<b>3D</b>	Three-Dimensional
<b>BART</b>	Berkeley Advanced Reconstruction Toolbox
<b>CSF</b>	Cerebrospinal Fluid
<b>CMI</b>	Conditional Mutual Information
<b>FOV</b>	Field of View
<b>GM</b>	Gray Matter
<b>GRE</b>	Gradient Recalled Echo
<b>ISMRM</b>	International Society of Magnetic Resonance in Medicine
<b>MI</b>	Mutual Information
<b>MR</b>	Magnetic Resonance
<b>MRI</b>	Magnetic Resonance Imaging
<b>NIST</b>	National Institute of Standards and Technology
<b>PD</b>	Proton Density
<b>PDF</b>	Probability Density Function
<b>QALAS</b>	Quantification using an interleaved Look-Locker Acquisition Sequence with T2 preparation pulse
<b>qMRI</b>	Quantitative Magnetic Resonance Imaging

---

<b>QRAPMASTER</b>	<b>Q</b> uantification of <b>R</b> elaxation times <b>A</b> nd <b>P</b> roton density by <b>M</b> ultiecho <b>A</b> cquisition of a <b>S</b> aturation-recovery using <b>T</b> urbo spin- <b>E</b> cho <b>R</b> eadout
<b>R1</b>	Longitudinal relaxation rate
<b>R2</b>	Transverse relaxation rate
<b>R2*</b>	Effective transverse relaxation rate
<b>RMSE</b>	<b>R</b> oot- <b>M</b> ean- <b>S</b> quare <b>E</b> rror
<b>SSE</b>	<b>S</b> um of <b>S</b> quared <b>E</b> rrors
<b>T</b>	<b>T</b> esla
<b>T1</b>	Longitudinal relaxation time
<b>T1-W</b>	<b>T1-Weighted</b>
<b>T2</b>	Transverse relaxation time
<b>T2*</b>	Effective transverse relaxation time
<b>T2-W</b>	<b>T2-Weighted</b>
<b>TE</b>	<b>E</b> cho <b>T</b> ime
<b>TI</b>	<b>I</b> nversion <b>T</b> ime
<b>TR</b>	<b>R</b> epetition <b>T</b> ime
<b>VFA</b>	<b>V</b> ariable <b>F</b> lip <b>A</b> ngle
<b>WM</b>	<b>W</b> hite <b>M</b> atter

# Chapter 1

## Introduction

Quantitative MRI is a single MRI quantification scan which measures physical properties. [1–5] Synthetic MRI methods allow conventional contrast images to be synthesized in post-processing using any combination of repetition time (TR), echo time (TE), and inversion time (TI). The ability to synthesize images reduces the scan time required to produce multiple series of different contrast weightings. Furthermore, the quantitative images produced by synthetic MRI provide additional diagnostic utility. Current acquisition parameters—the quantities which determine the acquisition sequence, such as flip angle, TR, TE, and delay times—are selected based on a combination of simple models and experience. As an alternative, this work proposes a quantitative framework using mutual information to evaluate the information content [6] of quantitative MRI acquisitions and to guide optimization of acquisition parameter selection for multi-parameter mapping.

In conventional MRI, diagnosis relies on comparison of relative image intensities, since absolute image intensity cannot by itself be used for this purpose. This usually necessitates several different contrast scans. Quantitative MRI addresses these shortcomings by allowing a pixel-wise examination of pathology to determine absolute deviation

from normal values.[1] It allows direct material identification by simultaneous quantification of multiple parameters, *e.g.* cancer cells from T1, T2, and self-diffusion tensor changes. However, many existing methods require clinically unacceptable scan times, and current fast methods have a narrow range of accuracy or employ fitting algorithms which require high SNR to obtain adequate estimates.

Information theory can address these drawbacks. This work aims to obtain a quantitative understanding of the information content of the acquisition parameters. Making an informed acquisition parameter selection instead of an empirical selection reduces the uncertainty in model output or optimizes information gain for fixed time. Measurements of an event with less entropy are by definition more repeatable. In other words, for given clinical constraints, acquisition parameters can be selected which maximize synthetic MRI repeatability. This gives the greatest consistency to post-processed synthetic MR images, automatic segmentation results, and diagnoses made from quantitative images.

Toward this goal, a quantitative framework is developed using mutual information to evaluate acquisition parameter selection for the optimization of multi-parameter mapping. An important feature of this framework is that it is application agnostic. It should be applicable to most clinical modeling problems and could provide a quantitative understanding of model parameter information content in clinical image reconstruction, treatment planning, and other applications.

## 1.1 Hypothesis

**The hypothesis of this research is that quantitative parameter optimization via information theory significantly reduces the variance of reconstructed**

**parametric map values in quantitative MRI.** The hypothesis will be tested through completion of the following three specific aims:

## 1.2 Specific Aims

**Specific Aim 1: Design and validate information theory framework to quantify information gain of quantitative MRI acquisitions with variable acquisition parameters.**

This information theory approach models uncertainties in image subject properties and measurement accuracy as probability distributions. These uncertainties propagate through the model of measurement acquisition, allowing quantification of uncertainty reduction as a function of controllable measurement parameters. This is ideally applied to signal models which are nonlinear operators with no analytical solution to validate the method, but it must first be validated on a model with a known analytical solution.

The working hypothesis in this aim is that mutual information can quantify the information gain of an MRI acquisition and predict the optimal acquisition parameters needed to maximize SNR of the measurement. To investigate this hypothesis, a spoiled GRE sequence is modeled to compare information-optimized flip angle to the analytically known Ernst angle, where SNR is maximum.

**Specific Aim 2: Develop a computational methodology to quantify information content of 3D-QALAS acquisitions relative to a representative synthetic brain model and validate in phantom.**

To demonstrate feasibility of information optimization in 3D-QALAS, a synthetic brain model is employed to represent the state of information in the system before image

acquisition. Information gained through image acquisition with 3D-QALAS is computed by Gauss-Hermite quadrature. Current 3D-QALAS reconstruction relies on empirical parameter selection during acquisition, but these parameters can be quantitatively optimized by maximizing modeled information gain.

The working hypothesis of this aim is that information-optimized acquisition parameters for a 3D-QALAS signal model correlate to smaller variances in reconstructed parametric maps. To test this hypothesis, reconstruction accuracy will be compared between optimal and sub-optimal acquisitions of phantoms with known T1 and T2 values. Theoretical reconstruction uncertainty reduction will be quantified for information-guided acquisition parameters.

**Specific Aim 3: Develop a computational methodology to optimize conditional information gain of 3D-QALAS acquisitions relative to low-information tuning acquisitions of the image subject and validate both in phantom and *in vivo*.**

A more flexible and accurate implementation of this information optimization must select a more representative model of the state of information in the system before image acquisition. A low-resolution tuning acquisition provides such a model. The conditional form of mutual information is computed to allow optimization of subsequent measurements with independent acquisition parameters. Conditional mutual information optimization will be used to locate the most informative points in parameter space. The information theory framework will quantify the information content of full-resolution 3D-QALAS acquisitions with updated acquisition parameters.

The working hypothesis of this aim is that information-optimized acquisition parameters for a 3D-QALAS signal model result in reduction of reconstructed parametric

map variance. To test this hypothesis, the variability of reconstructed parametric maps as a function of acquisition parameters will be predicted by conditional mutual information in a phantom with known T1 and T2 values. Additionally, an information-optimized acquisition of a human brain will be compared to an acquisition with default clinical parameters.

### 1.3 Dissertation Organization

The organization of this dissertation is as follows. Chapter 2 covers background information about synthetic MRI, the role of quantitative MRI in diagnostic imaging, and the importance of reproducibility testing for quantitative imaging modalities. Chapter 3 reviews relevant concepts from information theory. Chapters 4, 5, and 6 detail the completion of Specific Aims 1, 2, and 3, respectively. Chapter 7 draws conclusions from this work. Appendix A defines important terms used throughout this work. Appendix B defines mathematical symbols and conventions used consistently across chapters. Appendix C gives mathematical derivations omitted from the main chapters.



## Chapter 2

# Quantitative MRI

### 2.1 Introduction

Quantitative magnetic resonance imaging refers to the parametric mapping of meaningful physical or chemical properties, such as proton density, relaxation times, and B1 inhomogeneity.[1–5, 7] This differs from the qualitative images of conventional MRI, the image values of which possess no inherent meaning except relative to one another. When combined with other techniques like image synthesis and automatic segmentation, it is sometimes referred to as SyMRI.[8] Quantitative MRI is poised to bring numerous improvements to patient care through quantitative disease diagnoses, more robust automatic segmentation, and generation of synthetic images in post-processing.[1]

### 2.2 Relaxometry

The principles of quantifying relaxation times have been employed for decades.[9] Dozens of studies and methods have been performed to achieve rapid quantification of proton

density (PD), R1, R2, or R2\* individually.[10–22] Several recent methods have simultaneously quantified PD, R1, and R2, including quantification of relaxation times and proton density by multiecho acquisition of a saturation-recovery using turbo spin-echo readout (QRAPMASTER),[1, 23] magnetic resonance fingerprinting,[24] and inversion-recovery true fast imaging with steady state precession (IR TrueFISP).[25]

The most significant hurdles to widespread clinical implementation have been excessive scan time and lack of clinical experience with directly interpreting absolute PD, T1, and T2 maps. Previous multi-parameter mapping approaches have required 20–25 minutes to perform with partial parallel imaging.[26, 27] Substantial progress in recent years has allowed the absolute quantification of T1, T2, proton density, and B1 inhomogeneity in 5 minutes.[1, 28, 29] However, fast methods must employ fitting algorithms that require a high signal-to-noise ratio to produce acceptable results and are accurate only within a limited range of inputs.[30]

### 2.2.1 QRAPMASTER

QRAPMASTER is a widely used quantitative MRI pulse sequence that functions by repeating two phases: a slice-selective saturation pulse followed by spoiling acting on one slice and a slice-selective turbo spin-echo acquisition of a second, different slice. Variable delays between the execution of each phase on one specific slice render a matrix of measurements that is used to retrieve R1 and R2 relaxation times through a least squares fit. Local B1 field estimation is also possible due to the use of the saturation pulse (R1 curve position depends on B1). R1, R2, and B1 are then used to estimate unsaturated magnetization M0.[1] Simultaneous quantification of these parameters results in parametric maps which are automatically coregistered with one another, which is a distinct advantage over earlier methods that required separate scans for each parameter. The

QRAPMASTER reconstruction assumes monoexponential decay for the computation of R1 and R2, but multiexponential decay may cause errors near CSF interfaces.[31] Nevertheless, the method meets accuracy and reproducibility requirements for clinical use.[23]

### 2.2.2 3D-QALAS

A pulse signal of particular interest and used as a signal model in this work is 3D-QALAS (3D-Quantification using an interleaved Look-Locker acquisition sequence with T2 preparation pulse).[32] 3D-QALAS is a novel technique based on a multi-acquisition 3D gradient echo sequence. The sequence takes place in a T2 sensitization phase and a T1 sensitization phase. During the T2 sensitization phase, a series of pulses encode T2 relaxation on the  $M_z$  axis, and a gradient echo acquisition is performed. In the T1 sensitization phase, an inversion pulse is applied, and then four gradient echo acquisitions are performed as the longitudinal magnetization relaxes. M0, T1, and T2 parametric maps are fitted to these measurements from the five gradient echo acquisitions. The original application of 3D-QALAS is rapid cardiac mapping, but it may be adapted for 3D multi-parameter quantification in the brain with different timing constraints.

## 2.3 Post-Processing Utility

In addition to the diagnostic benefits, synthetic 3D MRI could fundamentally alter MRI workflow.[29] Ideally, only a single quantification scan is required, replacing survey and reference scans, as well as various contrast images at different orientations, potentially reducing MRI scanning and planning times. As with CT, images with the preferred contrast and orientation would be generated in post-processing.

### 2.3.1 Tissue Segmentation

Absolute parametric maps have also led to improvements in automatic tissue segmentation. Accurate segmentations for volume calculations are important in several diseases, including multiple sclerosis,[33] Alzheimer disease,[34] and vascular dementia.[35] Quantitative MR images have been found to make automatic segmentation results more robust.[36, 37] There is no need to normalize signal intensities as with segmentation done from conventional MR images,[38–40] and furthermore, absolute parametric maps are not affected by several machine and pulse sequence imperfections.[1] Segmentations performed with multiple contrasts[39] can model intra-voxel tissue type mixtures, which mitigates partial volume errors.[41] An automatic segmentation method is currently in use that uses quantitative outputs from the QRAPMASTER sequence[42] and has been found to have good repeatability.[43] It uses absolute parametric maps to develop lookup tables of tissue values corresponding to mixtures of brain tissues.[42] Addition of artificial noise gave relaxation times and proton density representative distributions in order to define confidence intervals.[44] Separate lookup tables are needed for 1.5T and 3.0T scanners, but segmentations from the two methods showed good agreement.[45–47]

### 2.3.2 Synthetic Image Synthesis

PD, T1, and T2 parametric maps are unlikely to be used clinically as the sole diagnostic tool for the time being simply because there is little experience making diagnoses from these images.[1] However, synthetic image synthesis can mimic the contrast weightings that radiologists see routinely, and studies have shown good agreement between diagnoses made on synthetic and conventional MRI.[48, 49] A wide range of contrast weightings can be achieved, as well as inversion recovery and double inversion recovery

images.[50–52] Contrasts can be customized to specific diseases[51] or to highly variable tissue properties, such as developing brains in pediatric imaging.[48] Synthetic MRI can allow shorter scan times by replacing multiple conventional acquisitions selected to provide multiple contrast weighting of diagnostic interest. This can be particularly useful when imaging time is limited or when patients require shorter scan times.[53] Rapid generation of synthetic images in post-processing is now possible and has greatly improved the clinical feasibility of this approach.[49, 54]

### 2.3.3 Magnetic Resonance Spectroscopy

Absolute metabolite concentrations cannot be calculated from conventional magnetic resonance spectroscopy. Instead, the ratios of metabolite concentrations relative to creatine are usually produced.[55] Quantification of spectroscopy results is valuable, because it reduces variations in the relative ratios as a function of age and other factors such as multiple sclerosis,[56, 57] as well as obviating the need for external reference phantoms.[58, 59] It is possible to calibrate magnetic resonance spectroscopy results using quantitative MRI by exploiting correlations of various metabolites with relaxation parameters and water concentration.[60]

### 2.3.4 Physical Models

Absolute-valued PD, T1, and T2 from quantitative MRI can be used to construct more complex models of physical tissue properties. Of particular importance are models of myelin content and edema, which are useful in the evaluation of multiple sclerosis. Several methods exist to construct such models,[48, 61, 62] and efforts have been made to include myelin modeling in the SyMRI workflow.[63] Quantitative comparison of brain

anatomy is also made possible through spatial normalization of parametric maps into a standardized stereotactic space.[64] This allows development of a repository of healthy brain data for which anatomical differences are eliminated. Such an approach could improve automated diagnosis of disease.

## 2.4 Clinical Applications

Magnetic resonance imaging methods that simultaneously quantify multiple parameters enable new diagnostic approaches—for example, the identification of cancer cells on the basis of T1, T2, and self-diffusion tensor changes.[24] Normal brain tissues possess relatively small ranges of T1, T2, and PD values,[65, 66] but pathological tissues deviate significantly from normal values.[1] Dozens of diagnostic applications have been investigated across a wide range of diseases.

### 2.4.1 Brain Metastases

Quantitative MRI has been used previously in the evaluation of brain metastases.[67] T1-weighted and T1-weighted inversion recovery images have historically had the greatest diagnostic utility for this purpose.[68] Hagiwara[50] et al. have compared diagnoses of brain metastases between synthetic and conventional MR images and found that more lesions were detected with synthetic T1-weighted inversion recovery than with its conventionally acquired counterpart.

### 2.4.2 Multiple Sclerosis

Multiple sclerosis is perhaps one of the diseases most amenable to analysis by quantitative MRI. Detection and evaluation of multiple sclerosis via quantitative imaging has already been achieved in several studies.[30, 69, 69–72] Treatment effectiveness is often evaluated by detecting new or progressing focal lesions.[73] Compared to conventional methods, synthetic MRI has been found to enable detection of a greater number of multiple sclerosis plaques.[37, 74–76] Two contrasts that are known to be effective for locating multiple sclerosis plaques—phase-sensitive inversion recovery and double inversion recovery images—are often not acquired due to time constraints. However, they can be synthetically generated in post-processing from absolute parametric maps.[51] Analyses of brain tissue segmentations have found that brain parenchymal fraction is an important biomarker in multiple sclerosis.[77–79] It is a predictor of patient outcome, including disability and relapsing-remitting disease,[80] and thus accurate and reproducible volumetric analysis is critical.

### 2.4.3 Other Diagnostic Utilities

The quantification of tissue relaxation times and proton density has been used in the detection and staging of many other diseases: epilepsy,[81] Parkinson disease,[82] Alzheimer disease,[29] meningitis,[83–85] thalassemia,[86, 87] Sturge-Weber syndrome,[50, 52, 88, 89] and idiopathic normal pressure hydrocephalus.[90–92] It has also been utilized in the general characterization of lesions[93, 94] and trauma,[95] for monitoring response to radiation therapy[96], for assessment of atherosclerotic plaques,[97] and for rapid quantification of metabolites in spectroscopic imaging.[98] Additionally, postmortem imaging is most commonly performed with CT,[99] but whole-body quantitative MRI[100] offers

better soft tissue discrimination.[101] It is necessary to perform temperature corrections so the tissue properties of dead tissues approximate those of tissues in living patients before synthetic images are generated.[102–105]

## 2.5 Performance of Quantitative Imaging

Synthetic MRI is an important technology with promising diagnostic applications, and it is imperative to quantitatively address of images produced by this method. This work addresses variability in 3D-QALAS, in particular. In general, the framework developed herein may be used to predict repeatability or reproducibility as a function of manually selected acquisition parameters for a wide range of imaging modalities and acquisition schemes.

Variability is a key concern in all quantitative imaging modalities. It is central to the mission statement of the Quantitative Imaging Biomarkers Alliance[106, 107], which is to "improve the value and practicality of quantitative imaging biomarkers by reducing variability across devices, patients, and time." [108] The metrology papers[108–112] put forth by QIBA outline terminology and methods for standardizing and optimizing performance evaluations of quantitative imaging biomarkers (QIBs).

Two accepted metrics of QIB performance are repeatability and reproducibility. Repeatability concerns measurement precision under identical testing conditions, whereas reproducibility concerns measurement precision under varying testing conditions, *e.g.* scanner, location, operator, or institution.[113]

The standardized metrics of repeatability and reproducibility suggested by Raunig[111] et al. include the repeatability coefficient (RC) and the reproducibility coefficient (RDC).



These coefficients are defined as the least significant difference between two measurements in a repeatability or reproducibility experiment, respectively. Alternatively, they can be understood as the width of the 95% confidence interval on the measurement distribution in these experiments. The repeatability coefficient (RC) is given by

$$\text{RC} = 1.96\sqrt{2s_w^2} = 2.77s_w, \quad (2.1)$$

where  $s_w^2$  is an estimate of  $\sigma_w^2$ , the within-subject variance.[111] The reproducibility coefficient (RDC) is given by

$$\text{RDC} = 2.77\sqrt{\sigma_\delta^2 + \sigma_{\gamma\delta}^2 + \sigma_\epsilon^2}, \quad (2.2)$$

where  $\sigma_\delta^2$  is the site variance,  $\sigma_{\gamma\delta}^2$  is the case by site variance, and  $\sigma_\epsilon^2$  is the variance between replicates within site and case.[111]

The guiding principles of QIBA have been implemented in variability studies across many quantitative imaging modalities, including dynamic contrast-enhanced (DCE) MRI,[114, 115] functional MRI (fMRI),[116, 117] MR elastography,[118] computed tomography (CT),[119–129] fluorodeoxyglucose 18F positron emission tomography (FDG PET),[130, 131] PET/CT,[132–137] and ultrasound elastography.[138, 139] Additional studies have demonstrated that variability in imaging protocols impacts reported progression-free survival results[140] and used Monte Carlo simulation to determine the threshold between measurement error and real change in the patient.[141]

## 2.6 Application of Information Theory

Acquisition parameters—the quantities that determine the acquisition sequence, such as excitation angle, TR, TE, and delay times—are currently selected on the basis of a combination of simple models and experience. As an alternative, this work develops a quantitative framework using mutual information to evaluate the information content [6] of quantitative MRI measurements and to guide acquisition parameter selection for the optimization of multi-parameter mapping.

The information theory framework addresses the accuracy requirements of fast methods and seeks to improve reproducibility. Development of this framework involved derivation of a numerical method to compute mutual information of a complex MRI pulse sequence model as well as an optimization method for parameter selection.

In this work, a quantitative understanding is obtained of the effect of information-theory-optimized acquisition parameters on the accuracy and reproducibility of quantitative tissue measurements. It was found that making an informed acquisition parameter selection instead of an empirical selection reduces uncertainty in model output and optimizes information gain for fixed time. These improvements are relevant to the clinic: for given clinical constraints, acquisition parameters can be selected that maximize synthetic MRI reproducibility, increasing the confidence with which one can make diagnoses based on quantitative images. Finally, the information theory framework is application agnostic: it is expected to be applicable to most clinical modeling problems as well as to clinical image reconstruction and treatment planning.

## Chapter 3

# Information Theory

### 3.1 Introduction

Information is often understood as the resolution of uncertainty.[\[142\]](#) The information entropy of an event is a measure of how much information it contains. An event that is nearly certain to occur possesses very little information, so greater information gain results from events with greater information entropy. These quantities have long been used in myriad signal processing and communications applications. Unsurprisingly, information theory intersects with imaging science and imaging physics as well, and its tools find a natural place in many imaging modeling problems. Mutual information[\[143\]](#) is the optimization metric used in conjunction with stochastic modeling[\[144\]](#) of the magnetic resonance imaging acquisitions in this work.

## 3.2 Mutual Information

The reduction in variance from a measurement is quantified using an information theory concept called mutual information. Mutual information is a quantity which describes the amount of information one random variable contains about another random variable. More intuitively, it is the reduction in one random variable's entropy due to knowledge of another variable. It is helpful to visualize mutual information as a Venn diagram of entropies for simple distributions (Figures 3.1 and 3.2). Generally, for some total entropy of a joint probability function of measurements and parametric maps, the entropy of the parametric map distribution is reduced by a known measurement to a smaller entropy. This reduction is mutual information.

Mutual information can be understood intuitively by applying it to a simple example, such as flipping two coins. Consider two fair coins flipped one after another (Figure 3.1). It is assumed that the second coin is equally likely to land on either side. Flipping the first coin gives no information about the outcome of the second coin flip. Thus, mutual information between the two is zero.

Now, consider two unfair coins flipped one after another (Figure 3.2). These coins land on one side 75% of the time. It is not known beforehand which side they favor, so it can only be assumed that both outcomes are equally likely. Because landing on one side is actually more probable, the first coin flip gives some information about the second. So, mutual information is greater than zero.

In imaging, acquiring a measurement is like flipping the first coin. The goal is to learn as much as possible about the distribution of image values, or analogously the distribution the second coin flip results, despite having an imperfect measurement due to machine noise and other factors. Acquisition parameters should be selected that are akin

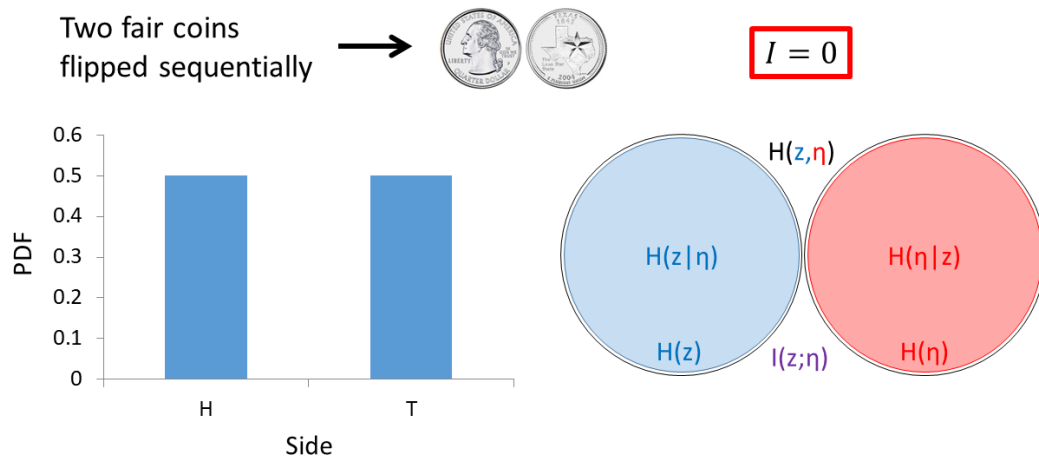


FIGURE 3.1: Fair coin flip example. Two coins are flipped sequentially, both with equal probability of landing on either side. The first coin provides no information about the second, so mutual information between the events is zero.

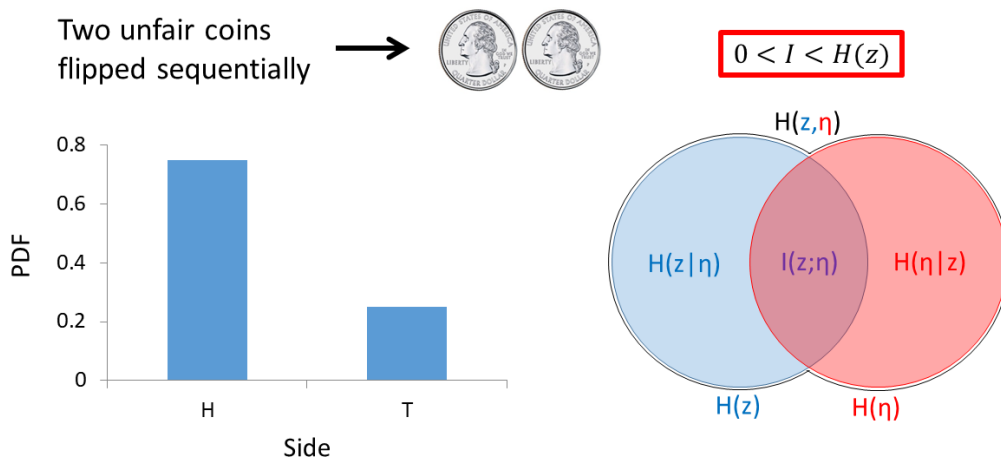


FIGURE 3.2: Unfair coin flip example. Two coins are flipped sequentially, both with 75% chance of landing on one side. The favored side is not known beforehand. The first coin flip now provides some information about the second, so mutual information between the two events is greater than zero.

to flipping the most unfair coin possible. Recall that information can be defined as the resolution of uncertainty.[142] Information known about an event reduces its uncertainty. In imaging applications, mutual information measures how much the measured signal reduces entropy of reconstructed image parameters, which correlates with reduction of the variance of reconstructed image parameters.

For an uninformative set of acquisition parameters, mutual information is small, and the variance of reconstructed parametric maps is large. Figure 3.3 illustrates a scenario in which a suboptimal 3D-QALAS acquisition is performed. Figure 3.3a shows a representation of the longitudinal magnetization in a 3D-QALAS acquisition with suboptimal acquisition parameters. Each of the five measurements is shown as a red x. The machine noise that corrupts the measurements is represented as a set of error bars on each measurement. The second, third, fourth, and fifth measurements are used to fit the longitudinal relaxation curve and determine T1. The fact that the third, fourth, and fifth measurements occur after most of the longitudinal relaxation has already occurred make this acquisition spacing highly suboptimal. The Venn diagram in Figure 3.3b shows the relationships between entropies resulting from the information theory modeling of the acquisition depicted in Figure 3.3a. In Bayesian terms, the entropy of the prior distribution of quantitative parameters of the image subject is illustrated by the red outline. These are the assumed distributions of the quantitative parameters before any imaging information is acquired. The entropy of the posterior distribution of quantitative parameters of the image subject is illustrated by the red filled portion of the circle. These are the information-model-predicted distributions of the quantitative parameters after imaging information is acquired. This entropy is reduced increasing mutual information (the purple filled portion of the circle). Thus, mutual information acts as a measure of acquisition optimality by its relationship to predicted entropy of the quantitative

parameters of the image subject. Figure 3.3c shows a synthetic reconstruction of the T1 map (right) from the previous suboptimal acquisition in Figure 3.3a. The image is visibly noisy. It is important to note that this is the result of relatively high uncertainty in the reconstruction of the quantitative parameters of the image subject and not a result of increased electronic or machine noise. The probability density functions (left) show the spread of reconstructed M0, T1, and T2 values in gray matter for the synthetic reconstruction. The spread of these values can be quantified by standard deviation (or higher order moments for more general distributions). The entropy of these distributions for one acquisition relative to another is predicted by the mutual information of the information theory model for that acquisition relative to the other.

Alternatively, the greatest mutual information exists for the set of acquisition parameters for which a measurement produces M0, T1, and T2 maps with the least entropy. Mutual information can be optimized over this acquisition parameter space to find the optimal set of parameters. Figure 3.4 illustrates a scenario in which an optimal 3D-QALAS acquisition is performed. Figure 3.4a shows a representation of the longitudinal magnetization in a 3D-QALAS acquisition with optimal acquisition parameters. Each of the five measurements is shown as a red x. The machine noise that corrupts the measurements is represented as a set of error bars on each measurement. The second, third, fourth, and fifth measurements are used to fit the longitudinal relaxation curve and determine T1. The fact that the second, third, fourth, and fifth measurements are well-spaced across the relaxation curve make this acquisition spacing more optimal than the suboptimal acquisition example. The Venn diagram in Figure 3.4b shows the relationships between entropies resulting from the information theory modeling of the acquisition depicted in Figure 3.4a. In Bayesian terms, the entropy of the prior distribution of quantitative parameters of the image subject is illustrated by the red outline.

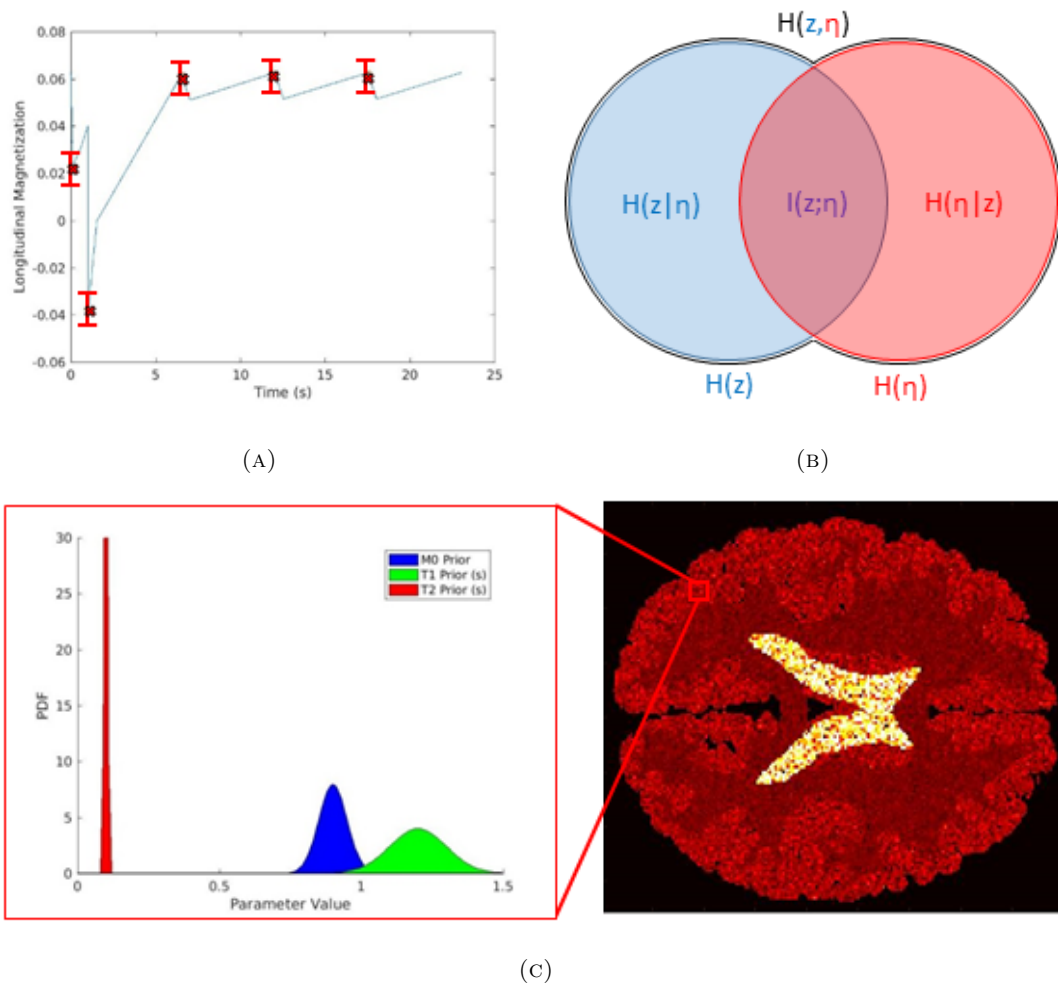


FIGURE 3.3: Representation of a suboptimal 3D-QALAS acquisition. (A) Longitudinal magnetization in a 3D-QALAS acquisition with suboptimal acquisition parameters. Each of the five measurements is shown as a red x. The machine noise that corrupts the measurements is represented as a set of error bars on each measurement. (B) Venn diagram showing the relationships between entropies resulting from the information theory modeling of the acquisition depicted in (A). In Bayesian terms, the entropy of the prior distribution of quantitative parameters of the image subject is illustrated by the red outline. The entropy of the posterior distribution of quantitative parameters of the image subject is illustrated by the red filled portion of the circle. This entropy is reduced increasing mutual information (the purple filled portion of the circle). (C) Synthetic reconstruction of the T1 map (right) from the previous suboptimal acquisition in (A). The probability density functions (left) show the spread of reconstructed M0, T1, and T2 values in gray matter for the synthetic reconstruction.



These are the assumed distributions of the quantitative parameters before any imaging information is acquired. The entropy of the posterior distribution of quantitative parameters of the image subject is illustrated by the red filled portion of the circle. These are the information-model-predicted distributions of the quantitative parameters after imaging information is acquired. This entropy is reduced increasing mutual information (the purple filled portion of the circle). Thus, mutual information acts as a measure of acquisition optimality by its relationship to predicted entropy of the quantitative parameters of the image subject. Compared to the previous figure, mutual information is greater, and the entropy of the posterior distribution is lesser. Figure 3.4c shows a synthetic reconstruction of the T1 map (right) from the previous suboptimal acquisition in Figure 3.4a. The image is visibly less noisy than the corresponding image in the previous figure. It is important to note that this is the result of relatively low uncertainty in the reconstruction of the quantitative parameters of the image subject and not a result of decreased electronic or machine noise. The probability density functions (left) show the spread of reconstructed M0, T1, and T2 values in gray matter for the synthetic reconstruction. The spread of these values can be quantified by standard deviation (or higher order moments for more general distributions). The entropy of these distributions for one acquisition relative to another is predicted by the mutual information of the information theory model for that acquisition relative to the other.

### 3.3 Information Theory Metrics

#### 3.3.1 Mutual Information

Mutual information can be understood as the expected value of the Kullback-Leibler divergence between two distributions[143]. For two discrete random variables,  $X$  and

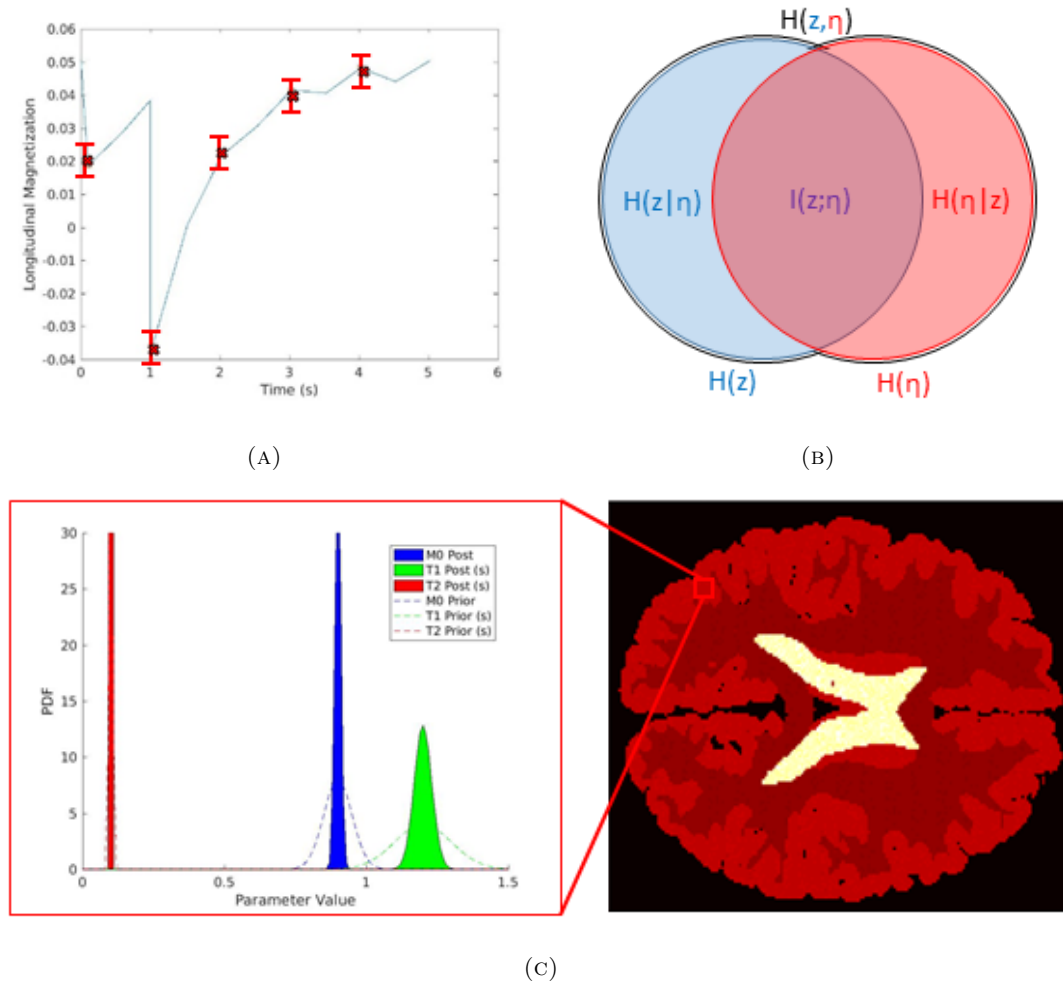


FIGURE 3.4: Representation of an optimal 3D-QALAS acquisition. (A) Longitudinal magnetization in a 3D-QALAS acquisition with optimal acquisition parameters. Each of the five measurements is shown as a red x. The machine noise that corrupts the measurements is represented as a set of error bars on each measurement. (B) Venn diagram showing the relationships between entropies resulting from the information theory modeling of the acquisition depicted in (A). In Bayesian terms, the entropy of the prior distribution of quantitative parameters of the image subject is illustrated by the red outline. The entropy of the posterior distribution of quantitative parameters of the image subject is illustrated by the red filled portion of the circle. This entropy is reduced increasing mutual information (the purple filled portion of the circle). (C) Synthetic reconstruction of the T1 map (right) from the previous optimal acquisition in (A). The probability density functions (left) show the spread of reconstructed M0, T1, and T2 values in gray matter for the synthetic reconstruction. The dashed lines show the probability distribution functions resulting from the suboptimal acquisition in Figure 3.3.

$Y$ , with joint probability distribution  $p(x, y)$  and marginal probability distributions  $p(x)$  and  $p(y)$ , mutual information between  $X$  and  $Y$  is defined

$$I(X; Y) = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log \left( \frac{p(x, y)}{p(x)p(y)} \right). \quad (3.1)$$

Similarly, if  $X$  and  $Y$  are continuous random variables, mutual information between  $X$  and  $Y$  is

$$I(X; Y) = \int_y \int_x p(x, y) \log \left( \frac{p(x, y)}{p(x)p(y)} \right) dx dy. \quad (3.2)$$

Mutual information may also be defined in terms of the difference of entropies:

$$I(X; Y) = H(Y) - H(Y|X) = H(X) - H(X|Y). \quad (3.3)$$

This definition is the most closely related to the qualitative understanding of mutual information as the reduction of uncertainty in one random variable given knowledge of another random variable. The derivation for this definition is shown in Section C.1.1 in Appendix C.

### 3.3.2 Joint Mutual Information

It is also necessary to define the joint mutual information between one random variable,  $X$ , and a pair of random variables,  $Y$  and  $Z$ . For discrete random variables, the joint mutual information is

$$I(X; Y, Z) = \sum_{z \in Z} \sum_{y \in Y} \sum_{x \in X} p(x, y, z) \log \left( \frac{p(x, y, z)}{p(x)p(y, z)} \right). \quad (3.4)$$

For continuous random variables, the joint mutual information is

$$I(X; Y, Z) = \int_z \int_y \int_x p(x, y, z) \log \left( \frac{p(x, y, z)}{p(x)p(y, z)} \right) dx dy dz. \quad (3.5)$$

A similar definition in terms of the difference of entropies is again possible:

$$I(X; Y, Z) = H(Y, Z) - H(Y, Z|X) = H(X) - H(X|Y, Z). \quad (3.6)$$

The derivation for this definition is shown in Section C.1.2 in Appendix C.

### 3.3.3 Conditional Mutual Information

Finally, it is necessary to define the conditional mutual information between random variables  $X$  and  $Y$  given random variable  $Z$ . For discrete random variables, the conditional mutual information is

$$I(X; Y|Z) = \sum_{y \in Y} \sum_{x \in X} p(x, y|z) \log \left( \frac{p(x, y|z)}{p(x|z)p(y|z)} \right). \quad (3.7)$$

For continuous random variables, the conditional mutual information is

$$I(X; Y|Z) = \int_y \int_x p(x, y|z) \log \left( \frac{p(x, y|z)}{p(x|z)p(y|z)} \right) dx dy. \quad (3.8)$$

A similar definition in terms of the difference of entropies is again possible:

$$I(X; Y|Z) = H(Y|Z) - H(Y|X, Z) = H(X|Z) - H(X|Y, Z). \quad (3.9)$$

The derivation for this definition is shown in Section C.1.3 in Appendix C.

### 3.3.4 Conditional Mutual Information Relationships

Several generalized conditional mutual information relationships can be derived for a new measurement,  $z$ , given  $N$  previous measurements,  $d_1, \dots, d_N$ .

$$I(\eta; z|d) = H(\eta|d) - H(\eta|z, d) \quad (3.10)$$

$$I(\eta; z|d) = H(\eta) - I(\eta; d_1) - \sum_{i=2}^N I(\eta; d_i|d_{i-1}, \dots, d_1) - H(\eta; z, d) \quad (3.11)$$

$$I(\eta; z|d) = I(\eta; z, d) - I(\eta; d_1) - \sum_{i=2}^N I(\eta; d_i|d_{i-1}, \dots, d_1) \quad (3.12)$$

$$I(\eta; z|d) = I(\eta; z, d) - I(\eta; d) \quad (3.13)$$

These relationships are derived in full in Section C.1.4 in Appendix C.

## 3.4 Numeric Methods

The information theory framework uses mutual information as an objective function. Computation of mutual information is an expensive high-dimensional integration problem requiring careful consideration of available numeric methods. Numerical integration of multivariate functions has been accomplished by several methods which are largely independent of the problem dimension, each of which is well-suited to certain function classes.[145] These methods include Monte Carlo and Quasi-Monte Carlo[146], lattice rules[147], adaptive subdivision[148, 149], and neural network approximation[150].

For mutual information calculation in particular, several integration and sampling methods have been employed. For special cases and simplified parameter estimation problems, conjugate priors [151], binary models[152], and linear Gaussian models[153]

have been used to compute evidence analytically. Numerical quadrature [154] has been used to approximate the posterior and evidence in order to calculate mutual information. Sequential Monte Carlo algorithms have been used to approximate the posterior and evidence for parameter estimation problems[155]. Importance sampling has been used in a computationally expensive method to determine Kullback-Leibler distance between prior and posterior distributions [156]. The same Kullback-Leibler distance has also been calculated by polynomial chaos approximation[157, 158] and nested Monte Carlo integration[159, 160].

Calculation of posterior variance can be accomplished through filters for parameter and state estimation. This has already been accomplished using the Kalman filter[6]. Several additional filters could be applied to this problem[161], including linear Gaussian filters (modified Kalman filter[162]), nonlinear Gaussian filters (extended Kalman filter[163], unscented Kalman filter[164], Gauss-Hermite filter[165, 166], quadrature Kalman filter[165], cubature Kalman filter[167]), adaptive filtering (adaptive filter with gain adaptation [168, 169], multiple models adaptive filtering [170–172]), and robust filtering (robust Kalman filter[173], variable structure filtering,  $H_\infty$  filtering[174]).

The dimensionality of problems can be mitigated by Smolyak’s construction (also called discrete blending method, Boolean method, or sparse grid method) for special function classes. Cubature can be performed by combinations of tensor products of one-dimensional quadrature formulas. This allows the number of function evaluations to be a logarithmic function of the problem dimensionality[145]. Different one-dimensional basis rules have been used to accelerate numerical integration, including the midpoint rule, the rectangle rule [175], the trapezoidal rule, the Clenshaw-Curtis rule [176, 177], and Gauss rules[145, 178]. The Newton-Cotes formulas use equidistant abscissas. Weights are

obtained from integrating Lagrange polynomials. For large numbers of points, Newton-Cotes formulas become numerically unstable. Because of this, the usual application is iterated versions of low degree formulas[175]. Clenshaw-Curtis formulas use non-equidistant abscissas obtained from the Chebyshev polynomials and are numerically more stable. If extreme points are used, the quadrature formulas are nested[179]. For Gauss formulas, the abscissas are determined from Legendre polynomials and the weights from integrating associated Lagrange polynomials. They are usually not nested, but they have the maximum possible polynomial degree of exactness[145]. Patterson[180] iterated Kronrod's Gauss quadrature extension formula recursively to retain maximal degree of exactness. The result was a sequence of nested quadrature formulas. It is not clear whether Patterson extensions exist for large numbers of cubature points.

Approximation of posterior distributions can be accelerated by combining variational inference and Markov Chain Monte Carlo[181]. Multi-index Monte Carlo (MIMC) is an extension of multilevel Monte Carlo (MLMC) and a stochastic version of a sparse combination technique[182]. Multifidelity methods have been applied to several problems in uncertainty propagation and optimization. Low-fidelity methods include simplified models (natural problem hierarchies, early-stopping criteria, coarse-grid approximations), projection-based models (proper orthogonal decomposition (POD), reduced basis method, Krylov subspace methods), and data-fit models (interpolation/regression, kriging, support vector machines)[183]. Multifidelity model management uses adaption, fusion, or filtering to either combine information from low- and high-fidelity models or reduce the number of high-fidelity model computations required. Efficient global optimization (EGO) is a multifidelity method used in optimization problems. The low-fidelity model is a kriging method. Kriging approximates a high-fidelity model as a stochastic process  $f_{hi} \approx b + \epsilon(z)$ .  $b \in \mathbb{R}$  is either the mean of the stochastic process or some

---

function depending on the input.  $\epsilon(z)$  is a normally distributed stochastic process.  $\epsilon$  simulates uncertainty caused by limited samples from the high-fidelity model[184]. Efficient global optimization employs surrogate-based optimization and multifidelity trust-region methods[185–190].



## Chapter 4

# Development and Verification of an Information Model

### 4.1 Introduction

In this chapter, an information theory framework for quantifying information content of imaging acquisitions is designed and introduced. This approach is verified on a spoiled gradient echo sequence, which is well-understood. The flip angle which maximizes signal intensity, called the Ernst angle, is known analytically. Thus, flip angle is a prime candidate for optimization by maximizing information gain of the acquisition. The information-optimized flip angle is compared to the analytically calculated Ernst angle, and their agreement is analyzed.

## 4.2 Theory

This section will describe in general terms the framework used to model an image subject and an image acquisition method. It will then demonstrate two methods of computing mutual information between modeled biological uncertainties in the image subject and measurement uncertainties in the acquisition method as a function of acquisition parameters.

For brevity and clarity, throughout this work the function  $N(\mathbf{x}|\mathbf{m}, \Sigma)$  will be used to represent a multivariate normal probability density function of  $N$ -dimensional random vector  $\mathbf{X}$  with mean  $\mathbf{m}$  and covariance matrix  $\Sigma$ .

$$N(\mathbf{x}|\mathbf{m}, \Sigma) = \frac{\exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{m})^T \Sigma^{-1}(\mathbf{x} - \mathbf{m})\right)}{\sqrt{(2\pi)^N |\Sigma|}} \quad (4.1)$$

### 4.2.1 General Problem Statement

Consider image subject  $\mathbf{x}$  on image domain  $\Omega$ :

$$\mathbf{x} \in \Omega. \quad (4.2)$$

The image subject on domain  $\Omega$  is defined by  $N$  mutually disjoint tissue labels, each of which occupy their own subdomain,  $\Omega_n$ .

$$\bigcup_{n=1}^N \Omega_n = \Omega \quad (4.3)$$

$$\Omega_n \cap \Omega_m = \emptyset \quad \text{for } n \neq m \quad (4.4)$$

Each tissue label has associated with it some multivariate distribution of physical tissue properties,  $\boldsymbol{\eta}$ . Thus, the distribution of physical tissue properties in the image subject is

$$\boldsymbol{\eta}(\mathbf{x}) = \sum_{n=1}^N \boldsymbol{\eta}_n U_n(\mathbf{x}), \quad (4.5)$$

where  $U_n$  is the indicator function for tissue label  $n$ :

$$U_n(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in \Omega_n \\ 0 & \text{if } \mathbf{x} \notin \Omega_n \end{cases}. \quad (4.6)$$

Physical tissue properties are modeled based on the signal model selection for image acquisition. For example, in a quantitative MRI acquisition  $\boldsymbol{\eta}(\mathbf{x}) = [M_0(\mathbf{x}), T_1(\mathbf{x}), T_2(\mathbf{x})]$ . These physical properties may be selected to simulate independent tissues of interest (white matter, gray matter, CSF, etc.) in each subdomain  $\Omega_n$ . The physical property distributions are normally distributed about literature values  $\mathbf{m}_\eta$  with covariance matrix  $\Sigma_\eta$  to simulate biological uncertainty in tissue properties.

$$p(\boldsymbol{\eta}) = N(\boldsymbol{\eta} | \mathbf{m}_\eta, \Sigma_\eta) \quad (4.7)$$

$$\mathbf{m}_\eta = \begin{bmatrix} m_{\eta_1} \\ \vdots \\ m_{\eta_N} \end{bmatrix} \quad (4.8)$$

$$\Sigma_{\eta} = \begin{bmatrix} \sigma_{\eta_1}^2 & 0 & \dots & 0 & 0 \\ 0 & \sigma_{\eta_2}^2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \sigma_{\eta_{N-1}}^2 & 0 \\ 0 & 0 & \dots & 0 & \sigma_{\eta_N}^2 \end{bmatrix} \quad (4.9)$$

These distributions are meant to encompass a range of feasible values that could be measured in any human brain. They may be thought of as the distributions of parameter values over a large population of human brains. Any individual brain likely has narrower distributions of these parameter values, *i.e.* smaller “biological uncertainty.” If patient-specific estimates of these distributions were known beforehand, they could serve as assumed prior distributions. Because this is not possible, it is most appropriate to model prior distributions after the range of values expected from a large population. 3D-QALAS measurement data is used to justify the assumed form of the prior distribution later in this chapter. It should be noted that the distribution of parameter values retrieved by measurement can never be less than that resulting from the combined effects of the true patient-specific biological uncertainty and the measurement uncertainty.

Image measurements  $\mathbf{z}$  are generated by operating on the modeled physical properties of the image subject with the signal model  $\mathcal{G}$  at some set of acquisition parameters  $\boldsymbol{\mu}$ .

$$\mathbf{z} = \mathcal{G}_{\boldsymbol{\mu}}(\boldsymbol{\eta}(\mathbf{x})) + \nu = g_{\boldsymbol{\mu}} + \nu \quad (4.10)$$

The modeled measurement  $g_{\boldsymbol{\mu}}$  is corrupted by simulated measurement noise  $\nu$ . Zero-mean, normally distributed measurement noise is sufficient for many applications.

$$p(\nu) = N(\nu | \mathbf{0}, \Sigma_{\nu}) \quad (4.11)$$

$$\Sigma_\nu = \begin{bmatrix} \sigma_{\nu_1}^2 & 0 & \dots & 0 & 0 \\ 0 & \sigma_{\nu_2}^2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \sigma_{\nu_{N-1}}^2 & 0 \\ 0 & 0 & \dots & 0 & \sigma_{\nu_N}^2 \end{bmatrix} \quad (4.12)$$

### 4.2.2 Gauss-Hermite Quadrature

The multi-dimensional integration required by mutual information is computed using Gauss-Hermite quadrature. The dimensionality of this application is on the cusp of what is computationally feasible for quadrature, but a deterministic method allowed development of intuition for the quantities involved. More complex applications will require stochastic methods, such as Markov chain Monte Carlo, to be computationally feasible.

Gauss-Hermite quadrature approximates integrals of the form  $\int_{-\infty}^{\infty} e^{-x^2} f(x) dx$ :

$$\int_{-\infty}^{\infty} e^{-x^2} f(x) dx \approx \sum_{i=1}^N \omega_i f(x_i). \quad (4.13)$$

For  $n$  quadrature points, the function  $f$  is evaluated at  $x_i$ , the roots of the physicists' Hermite polynomial  $H_n$ . The associated Gauss-Hermite weights are defined

$$\omega_i = \frac{2^{n-1} n! \sqrt{\pi}}{n^2 [H_{n-1}(x_i)]^2}. \quad (4.14)$$

To numerically compute mutual information of high-dimensional uncertainties, it is necessary to perform high-dimensional integration. Because uncertainties are often modeled as normal distributions, it will be necessary to approximate the integration

of the product of a multivariate normal probability density function and an arbitrary function. If the change of variables

$$\mathbf{x} = \frac{1}{\sqrt{2}}\Sigma^{-1}(\boldsymbol{\xi} - \boldsymbol{\mu}) \Leftrightarrow \boldsymbol{\xi} = \sqrt{2}\Sigma\mathbf{x} + \boldsymbol{\mu} \quad (4.15)$$

is made, the form of integrals approximated by Gauss-Hermite quadrature can be written

$$\int_{-\infty}^{\infty} ((2\pi)^2|\Sigma|)^{-1/2} \exp\left(-\frac{1}{2}(\boldsymbol{\xi} - \boldsymbol{\mu})^T \Sigma^{-1}(\boldsymbol{\xi} - \boldsymbol{\mu})\right) h(\boldsymbol{\xi}) d\boldsymbol{\xi}. \quad (4.16)$$

In particular, it can be shown that integrating the product of some function  $h(\boldsymbol{\xi})$  and an  $N$ -dimensional multivariate normal distribution with independent variables results in the Gauss-Hermite quadrature approximation

$$\int f(\boldsymbol{\xi}|\boldsymbol{\mu}, \Sigma)h(\boldsymbol{\xi})d\boldsymbol{\xi} \approx \pi^{-N/2} \sum_{i_1=1}^N \omega_{i_1} \dots \sum_{i_N=1}^N \omega_{i_N} h\left(\sqrt{2}\Sigma\mathbf{x} + \boldsymbol{\mu}\right). \quad (4.17)$$

This approximation allows the computation of mutual information for models in which uncertainties can be modeled as multivariate normally distributed random variables.

### 4.2.3 Linear Approximation

Computational expense can be saved by making the (often poor) assumption that the signal model operator  $\mathcal{G}$  is a linear operator. If this is true, the distribution  $p(\mathbf{z})$  may be treated as normal.

$$I(\eta; z) = H(z) - H(z|\eta) = \frac{1}{2} \ln((2\pi e)^2 \cdot |\Sigma_z|) - \frac{1}{2} \ln((2\pi e)^2 \cdot |\Sigma_\nu|) \quad (4.18)$$

Conditional entropy  $H(z|\eta)$  is trivial to compute from the problem definition:

$$|\Sigma_\nu| = \prod_{i=1}^D \sigma_{\nu,i}^2. \quad (4.19)$$

The full derivation of the covariance matrix of the evidence  $\Sigma_z$  is left for Section C.2.1 in Appendix C, but the quadrature approximation of  $\Sigma_z$  is

$$\begin{aligned} \Sigma_z = & \sum_{q_1=1}^{Q_1} \omega_{q_1} \dots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \dots \sum_{k_N=1}^{K_N} \omega_{k_N} \left( \sqrt{2}\Sigma_\nu \mathbf{x}_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right) \left( \sqrt{2}\Sigma_\nu \mathbf{x}_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right)^T \\ & - \sum_{q_1=1}^{Q_1} \omega_{q_1} \dots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \dots \sum_{k_N=1}^{K_N} \omega_{k_N} \left( \sqrt{2}\Sigma_\nu \mathbf{x}_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right) \\ & \cdot \left[ \sum_{q_1=1}^{Q_1} \omega_{q_1} \dots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \dots \sum_{k_N=1}^{K_N} \omega_{k_N} \left( \sqrt{2}\Sigma_\nu \mathbf{x}_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right) \right]^T. \end{aligned} \quad (4.20)$$

Approximation of mutual information is straightforward once the covariance matrix  $\Sigma_z$  is known.

$$I(\eta; z) = H(z) - H(z|\eta) = \frac{1}{2} \ln \left( (2\pi e)^D \cdot |\Sigma_z| \right) - \frac{1}{2} \ln \left( (2\pi e)^D \prod_{i=1}^D \sigma_{\nu,i}^2 \right) \quad (4.21)$$

For a scalar-valued measurement  $z$ , Equation 4.21 simplifies to the following:

$$\begin{aligned} I(\eta; z) = & \frac{1}{2} \ln \left\{ 2\pi e \left[ \sum_{q_1=1}^{Q_1} \omega_{q_1} \dots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \left( \sqrt{2}\Sigma_\nu \mathbf{x}_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right)^2 \right. \right. \\ & \left. \left. - \left( \sum_{q_1=1}^{Q_1} \omega_{q_1} \dots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \left( \sqrt{2}\Sigma_\nu \mathbf{x}_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right) \right)^2 \right] \right\} \\ & - \frac{1}{2} \ln (2\pi e \sigma_\nu^2). \end{aligned} \quad (4.22)$$

#### 4.2.4 Nonlinear Computation

The signal model  $\mathbf{z}$  is a nonlinear function of parameters  $\boldsymbol{\eta}$ .

$$\mathbf{z} = \mathcal{G}_{\boldsymbol{\mu}}(\boldsymbol{\eta}(\mathbf{x})) + \nu \quad (4.23)$$

The conditional probability  $p(\mathbf{z}|\boldsymbol{\eta})$  is defined as a multivariate normal distribution  $N(\boldsymbol{\xi}|\boldsymbol{\mu}, \Sigma)$ , a function of  $\boldsymbol{\xi}$  with mean  $\boldsymbol{\mu}$  and covariance  $\Sigma$ , in the problem statement.

$$p(\mathbf{z}|\boldsymbol{\eta}) = N(\mathbf{z}|\mathcal{G}_{\boldsymbol{\mu}}(\boldsymbol{\eta}(\mathbf{x})), \Sigma_{\nu}) \quad (4.24)$$

Similarly, the probability  $p(\boldsymbol{\eta})$  is defined as another multivariate normal distribution in the problem statement.

$$p(\boldsymbol{\eta}) = N(\boldsymbol{\eta}|\mathbf{m}_{\boldsymbol{\eta}}, \Sigma_{\boldsymbol{\eta}}) \quad (4.25)$$

It is most straightforward to begin with mutual information defined as the difference in entropies:

$$I(\boldsymbol{\eta}; \mathbf{z}) = H(\mathbf{z}) - H(\mathbf{z}|\boldsymbol{\eta}) \quad (4.26)$$

The entropy of a multivariate normal distribution is  $H(N) = \frac{1}{2} \ln((2\pi e)^D \det(\Sigma))$ , so the term  $H(\mathbf{z}|\boldsymbol{\eta})$  can be immediately written

$$H(\mathbf{z}|\boldsymbol{\eta}) = \frac{1}{2} \ln \left( (2\pi e)^D \prod_{i=1}^D \sigma_{\nu,i}^2 \right). \quad (4.27)$$

The term  $H(\mathbf{z})$ , however, must be approximated through multiple applications of Gauss-Hermite quadrature. This approximation is possible because  $H(\mathbf{z})$  may be represented



in terms of distributions defined to be normal in the problem statement:

$$\begin{aligned}
H(\mathbf{z}) &= \int p(\mathbf{z}) \ln(p(\mathbf{z})) d\mathbf{z} \\
&= \int \left( \int p(\mathbf{z}|\boldsymbol{\eta}) p(\boldsymbol{\eta}) d\boldsymbol{\eta} \right) \cdot \ln \left( \int p(\mathbf{z}|\boldsymbol{\eta}) p(\boldsymbol{\eta}) d\boldsymbol{\eta} \right) d\mathbf{z} \\
&= \int \left( \int N(\mathbf{z}|\mathcal{G}_{\boldsymbol{\mu}}(\boldsymbol{\eta}), \Sigma_{\nu}) N(\boldsymbol{\eta}|\mathbf{m}_{\boldsymbol{\eta}}, \Sigma_{\boldsymbol{\eta}}) d\boldsymbol{\eta} \right) \\
&\quad \cdot \ln \left( \int N(\mathbf{z}|\mathcal{G}_{\boldsymbol{\mu}}(\boldsymbol{\eta}), \Sigma_{\nu}) N(\boldsymbol{\eta}|\mathbf{m}_{\boldsymbol{\eta}}, \Sigma_{\boldsymbol{\eta}}) d\boldsymbol{\eta} \right) d\mathbf{z}.
\end{aligned} \tag{4.28}$$

The full derivation of the evidence entropy  $H(\mathbf{z})$  is left for Section C.2.2 in Appendix

C. The final result of quadrature approximation of  $H(\mathbf{z})$  is

$$\begin{aligned}
H(\mathbf{z}) &= \pi^{-N/2-P/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \cdots \sum_{k_P=1}^{K_P} \omega_{k_P} \\
&\quad \cdot \ln \left( \pi^{-N/2} \sum_{s_1=1}^{S_1} \omega_{s_1} \cdots \sum_{s_N=1}^{S_N} \omega_{s_N} \cdot N \left( \left( \sqrt{2} \Sigma_{\nu} \mathbf{x}_k + \mathcal{G}_{\boldsymbol{\mu}}(\boldsymbol{\eta}_q) \right) \middle| \mathcal{G}_{\boldsymbol{\mu}}(\boldsymbol{\eta}_s), \Sigma_{\nu} \right) \right).
\end{aligned} \tag{4.29}$$

Thus, the computation of mutual information takes the following form:

$$\begin{aligned}
I(\boldsymbol{\eta}; \mathbf{z}) &= H(\mathbf{z}) - H(\mathbf{z}|\boldsymbol{\eta}) \\
&= \pi^{-N/2-P/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \cdots \sum_{k_P=1}^{K_P} \omega_{k_P} \\
&\quad \cdot \ln \left( \pi^{-N/2} \sum_{s_1=1}^{S_1} \omega_{s_1} \cdots \sum_{s_N=1}^{S_N} \omega_{s_N} \cdot N \left( \left( \sqrt{2} \Sigma_{\nu} \mathbf{x}_k + \mathcal{G}_{\boldsymbol{\mu}}(\boldsymbol{\eta}_q) \right) \middle| \mathcal{G}_{\boldsymbol{\mu}}(\boldsymbol{\eta}_s), \Sigma_{\nu} \right) \right) \\
&\quad - \frac{1}{2} \ln \left( (2\pi e)^D \prod_{i=1}^D \sigma_{\nu,i}^2 \right).
\end{aligned} \tag{4.30}$$

For a scalar-valued measurement  $z$ , Equation 4.30 reduces to

$$\begin{aligned}
I(\boldsymbol{\eta}; \mathbf{z}) &= H(\mathbf{z}) - H(\mathbf{z}|\boldsymbol{\eta}) \\
&= \pi^{-N/2-1/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \\
&\quad \cdot \ln \left( \pi^{-N/2} \sum_{s_1=1}^{S_1} \omega_{s_1} \cdots \sum_{s_N=1}^{S_N} \omega_{s_N} \cdot N \left( \left( \sqrt{2}\sigma_\nu x_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right) \middle| \mathcal{G}_\mu(\boldsymbol{\eta}_s), \sigma_\nu \right) \right) \\
&\quad - \frac{1}{2} \ln (2\pi e \sigma_\nu^2).
\end{aligned} \tag{4.31}$$

## 4.3 Methods

### 4.3.1 Verification of Information Theory Framework

Linear and nonlinear methods for calculating mutual information are developed in Section 4.2. Mutual information is used as the objective function in an acquisition parameter optimization for a spoiled GRE pulse sequence. This sequence was selected to enable comparison to an analytical optimum value and analysis of error in the information optimization results.

The spoiled GRE pulse sequence is described mathematically by

$$g = M_0 \frac{\sin \alpha (1 - e^{-TR/T_1})}{(1 - (\cos \alpha) e^{-TR/T_1})} e^{-TE/T_2^*}, \tag{4.32}$$

where  $M_0$  is the fully relaxed longitudinal magnetization,  $\alpha$  is the flip angle,  $TR$  is repetition time,  $T_1$  is longitudinal relaxation time,  $TE$  is echo time,  $T_2^*$  is effective transverse relaxation time. The flip angle at which signal intensity is maximized is called the Ernst angle.[191] This can be determined analytically by setting  $\frac{dg}{d\alpha} = 0$  from

Equation 4.32. The result is

$$\alpha_E = \arccos(e^{-TR/T1}). \quad (4.33)$$

Figure 4.1a shows the well-known relationship between the Ernst angle and the ratio between TR and T1. Figure 4.1b shows signal intensity as a function of flip angle for two different T1 relaxation times representative of white matter and cerebrospinal fluid. The Ernst angles are located at the maxima of the two functions and can be seen to differ for different tissue T1 values. Ernst angle can be calculated iteratively by maximizing signal intensity, or equivalently if assuming constant noise, by maximizing SNR. Because mutual information measures uncertainty reduction in the measured signal, it is inversely correlated with noise and positively correlated with SNR. Thus, mutual information may be used as the objective function in a flip angle optimization for a spoiled GRE sequence, and the result can be expected to agree with the analytical Ernst angle.

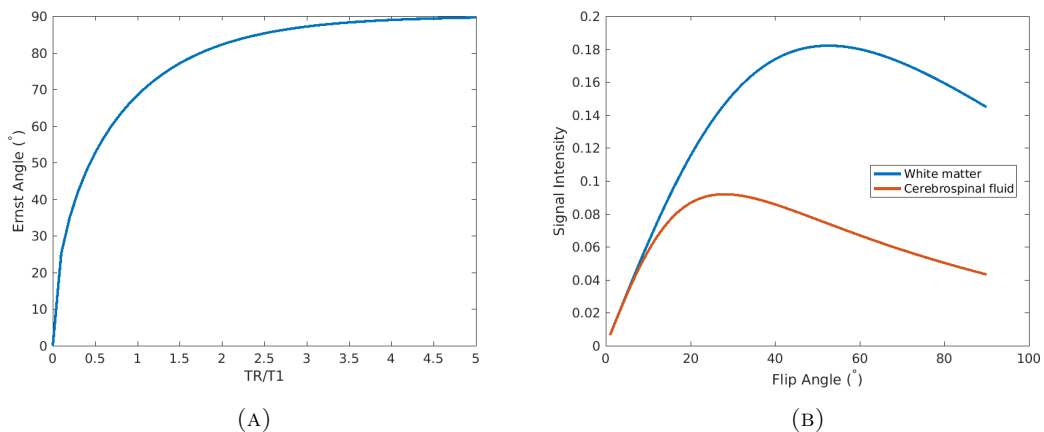


FIGURE 4.1: (A) Plot of the well-known relationship between the Ernst angle and the ratio of TR and T1 for a spoiled GRE sequence. (B) Plot of the signal intensity in a spoiled GRE sequence as a function of flip angle for characteristic values of two brain tissues. T1 for white matter was set to 1000 ms. T1 for cerebrospinal fluid was set to 4000 ms. TR for the acquisition was selected to be 500 ms.

Figure 4.2 shows a flowchart illustrating the process by which information-optimized flip angles are obtained. The image subject is simplified to a single tissue label consisting

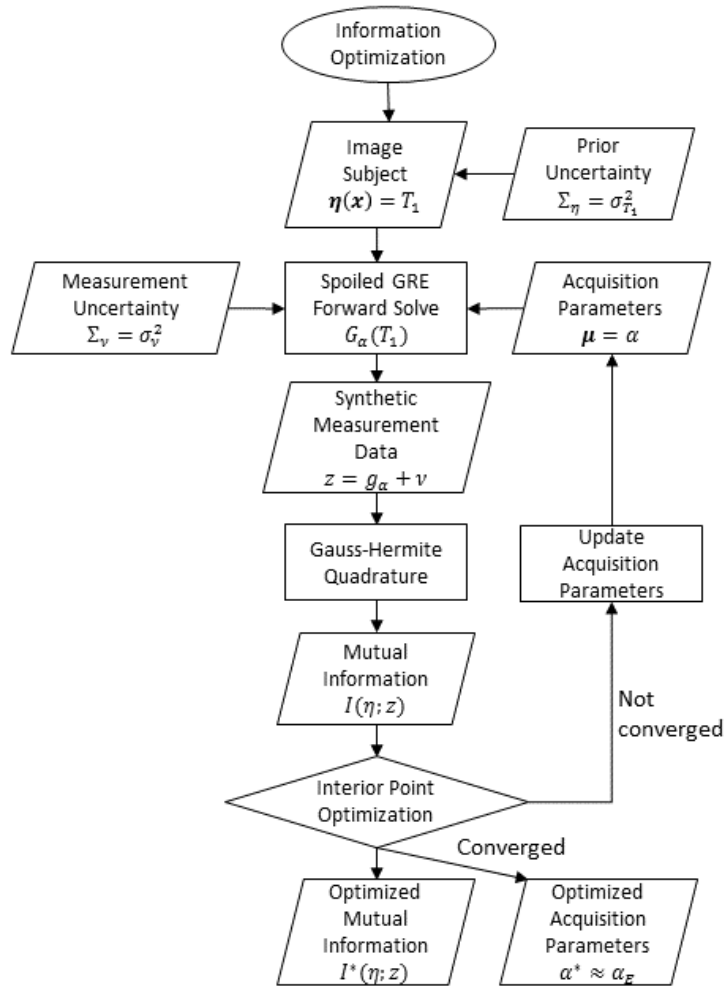


FIGURE 4.2: Flow chart for the mutual information-based optimization of flip angle in a spoiled GRE pulse sequence.

of a single voxel. The only relevant material properties for this acquisition model are the longitudinal relaxation time  $T_1$  and the effective transverse relaxation time  $T_2^*$ . Thus, the general vector of material properties  $\boldsymbol{\eta}$  is reduced to a two-dimension vector

$$\boldsymbol{\eta}(\boldsymbol{x}) = \begin{bmatrix} T_1 \\ T_2^* \end{bmatrix}. \quad (4.34)$$

The prior uncertainty is generally described by the covariance matrix of the physical properties  $\boldsymbol{\eta}$  in all tissue labels.  $T_1$  and  $T_2^*$  are assumed to be uncorrelated, so the

covariance matrix is

$$\Sigma_{\eta} = \begin{bmatrix} \sigma_{T_1}^2 & 0 \\ 0 & \sigma_{T_2^*}^2 \end{bmatrix}, \quad (4.35)$$

where  $\sigma_{T_1}^2$  is the T1 variance and  $\sigma_{T_2^*}^2$  is the  $T_2^*$  variance of the image subject. Signal intensity is a function of three acquisition parameters:

$$\boldsymbol{\mu} = \begin{bmatrix} \alpha \\ TR \\ TE \end{bmatrix}. \quad (4.36)$$

Flip angle  $\alpha$  is optimized for various fixed values of TR and TE. The signal model is given previously in Equation 4.32. Because signal intensity is a real-valued scalar, the measurement uncertainty is given as

$$\Sigma_{\nu} = \sigma_{\nu}^2, \quad (4.37)$$

where  $\sigma_{\nu}^2$  is the variance of the simulated signal intensity due to modeled acquisition error. The optimization itself is written

$$\begin{aligned} \alpha^*(T_1, TR) &= \underset{\alpha}{\operatorname{argmin}} I(\boldsymbol{\eta}; z(\boldsymbol{\mu})) \\ &= \underset{\alpha}{\operatorname{argmin}} I(T_1, T_2; z(\alpha, TR, TE)) \\ &\approx \alpha_E(T_1, TR), \end{aligned} \quad (4.38)$$

where  $\alpha^*$  is the information-optimized flip angle. The optimization was performed using an interior point method with parameter values listed in Table 4.1. Mutual information is calculated using Equation 4.21 for the nonlinear approximation or Equation 4.30 for the full nonlinear solution.

<b>Optimization Parameter</b>	<b>Value</b>
$m_{T_2^*}$ (ms)	100
$\sigma_{T_1}$ (ms)	100
$\sigma_{T_2^*}$ (ms)	10
$M_0$	1
<b>TE</b> (ms)	100
$\sigma_\nu$	0.05
<b>Termination tolerance</b>	$1.0 \times 10^{-20}$
<b>Optimality tolerance</b>	$1.0 \times 10^{-20}$
<b>Finite difference step size</b>	$2.0 \times 10^{-26}$

TABLE 4.1: Optimization parameters used in calculating information-optimized flip angle for both linear and nonlinear approaches.

Information-optimized flip angle and analytical Ernst angle are calculated for 81 T1 relaxation times evenly spaced between 600 ms and 1400 ms, inclusive, and 81 TR values evenly spaced between 100 ms and 500 ms, inclusive. This results in 6561 points of comparison with the known solution for each optimization method. The relative error is plotted for the linear optimization for visual inspection of trends. The trends and biases of optimization errors are analyzed more thoroughly through linear regression and Bland Altman analysis.[192] The Bland Altman plot is recommended in QIBA metrology by Raunig et al.[111] in order to show trends in variability of measurements over the measuring interval.

### 4.3.2 Justification of Model Assumptions for 3D-QALAS

The assumption of independence between the nine variables of the prior distribution (M0, T1, and T2 for gray matter, white matter, and CSF) is tested with simulated and measured data. Given the multivariate normal prior distribution described previously, the posterior distribution was sampled using ellipsoidal nested sampling.[193, 194] The Pearson correlation coefficient is used to evaluate the strength of the correlation between parameters:

$$\rho_{i,j} = \frac{\text{cov}(\eta_i, \eta_j)}{\sigma_{\eta_i} \sigma_{\eta_j}}. \quad (4.39)$$

Correlations between parameters are compared between simulated posterior distributions and real measurements from a 3D-QALAS acquisition.

## 4.4 Results

The information theory framework was successfully verified using a spoiled GRE signal model. Simulated measurements made with the assumptions about the prior distribution are shown to agree well with actual measured data.

### 4.4.1 Linear Approximation

To visualize the relationship between SNR and mutual information as functions of a single acquisition parameter for a spoiled GRE sequence, a global search of these quantities was performed over flip angles between  $0^\circ$  and  $90^\circ$ . Figure 4.3a shows these results for a single scenario with fixed T1 and TR at 1000 ms and 500 ms, respectively. The analytical solution which maximizes SNR is  $52.66^\circ$ , given by Equation 4.33. Maximum mutual information was achieved at  $53^\circ$ . Precision was limited by search grid resolution. The correlation between SNR and mutual information is shown in Figure 4.3b. SNR and mutual information are positively correlated.

Figure 4.4 shows a comparison of Ernst angle (Figure 4.4a) and information-optimized flip angle (Figure 4.4b) for T1 values between 600 ms and 1400 ms and TR values between 100 ms and 500 ms. The Ernst angle maximizes spoiled GRE signal intensity for given T1 and TR, and it is calculated analytically from Equation 4.33. Information-optimized flip angles are computed by solving the optimization problem

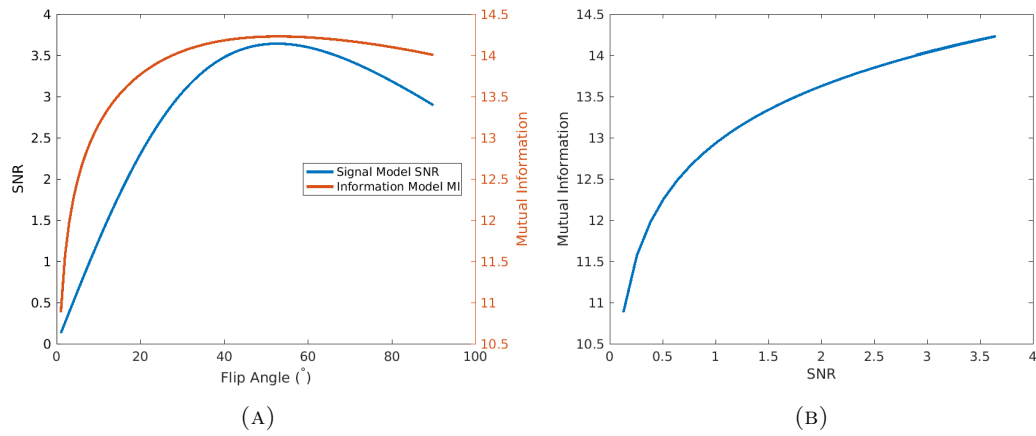


FIGURE 4.3: (A) Comparison between SNR calculated from a spoiled GRE signal model and mutual information calculated from an information theory model of the spoiled GRE sequence. Calculations are performed for a single scenario where T1 is 1000 ms and TR is 500 ms. The signal model SNR is plotted against the left ordinate, and the information model mutual information is plotted against the right ordinate. The maxima of both are located at approximately 53°. (B) Correlation between SNR and mutual information. SNR and mutual information are positively correlated.

in Equation 4.38 and calculating mutual information as in Equation 4.21. Information gain is quantified by mutual information between measurement and tissue property uncertainties. Greater information gain corresponds to greater measurement SNR.

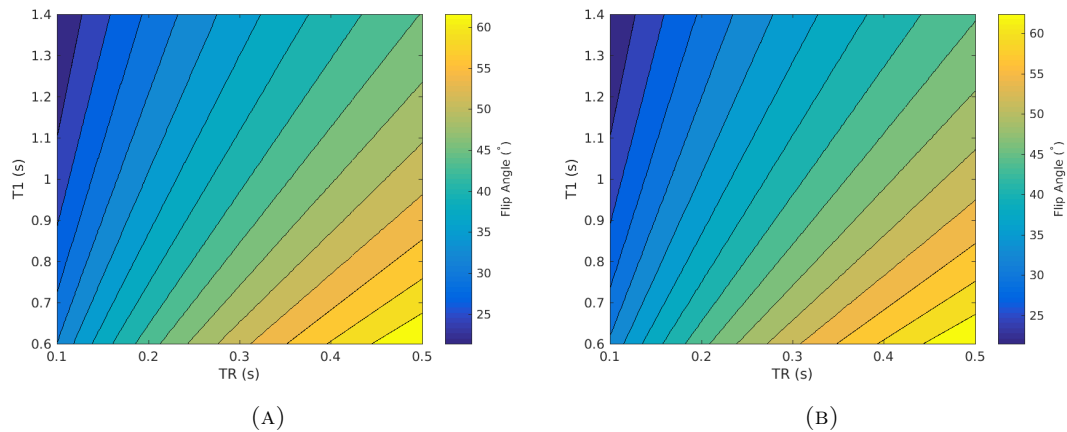


FIGURE 4.4: (A) Theoretical Ernst angle calculated over a range of T1 and TR values. Ernst angle values are calculated from the deterministic equation which maximizes spoiled GRE signal intensity for the flip angle. (B) Optimal flip angle selected by linear approximation of information gain from spoiled GRE acquisitions over a range of T1 and TR values. Information gain is quantified by mutual information between measurement and tissue property uncertainties. Greater information gain corresponds to greater measurement SNR.

Figure 4.5 shows the relative error between the analytical Ernst angle (Figure



4.4a) and the linear information-optimized flip angle (Figure 4.4b). Mean relative error is approximately 0.79%, and maximum relative error is approximately 2.12%. Relative error displays a clear trend and is inversely correlated to both T1 and TR.

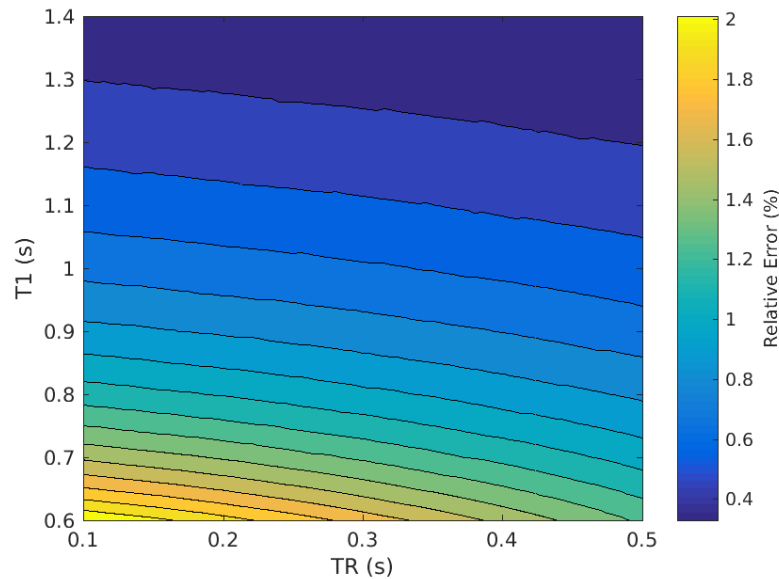


FIGURE 4.5: Relative error between analytical Ernst angle and flip angle optimization using information theory.

Figure 4.6 shows the regression line between the analytical Ernst angle and the linear information-optimized flip angle in a spoiled GRE pulse sequence over the same range of T1 and TR values. Each plotted point corresponds to the analytical Ernst angle (abscissa) and the information-optimized flip angle (ordinate) at identical T1 and TR. The regression line slope and intercept are 1.0143 and 0.2538, respectively. The adjusted coefficient of determination ( $r^2$ ) is 0.9997. The sum of squared errors (SSE) and root-mean-square error (RMSE) are 162.67 and 0.1575, respectively.

Figure 4.7 shows a Bland Altman plot of the analytical Ernst angle and the non-linear information-optimized flip angle measurement pairs plotted in Figures 4.4 and 4.6. The mean measurement difference,  $d$ , is  $0.3426^\circ$ . The bounds of the 95% confidence interval (defined by  $d \pm 1.96s$ , where  $s$  is the standard deviation of measurement

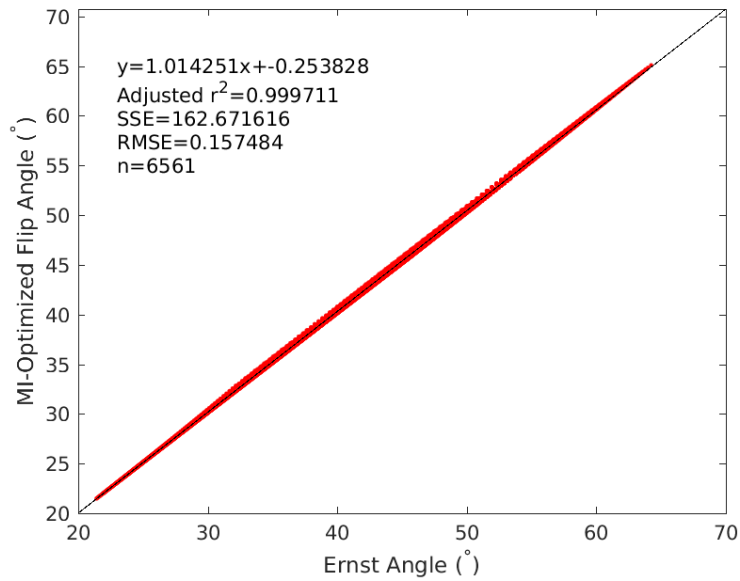


FIGURE 4.6: The regression line between 6561 measurement pairs of the analytical Ernst angle and the linear information-optimized flip angle in a spoiled GRE pulse sequence over a range of T1 and TR values. Each measurement pair is plotted with a red dot. The regression line is shown in black. Fit parameters and goodness of fit metrics are displayed in the upper left corner.

differences) are  $-0.0577^\circ$  and  $0.7429^\circ$ . If these measurement differences were normally distributed, 95% of all information-optimized flip angles obtained with the nonlinear MI model will differ from the analytical Ernst angle by between  $-0.0577^\circ$  and  $0.7429^\circ$ .

Figure 4.8 shows a histogram of the distribution of measurement differences (ordinate of Figure 4.7). The distribution is highly non-Gaussian and has a long tail in the direction of larger negative flip angle differences.

#### 4.4.2 Nonlinear Model

Figure 4.9 shows a comparison of Ernst angle (Figure 4.9a) and information-optimized flip angle (Figure 4.9b) for T1 values between 600 ms and 1400 ms and TR values between 100 ms and 500 ms. The Ernst angle maximizes spoiled GRE signal intensity for given T1 and TR, and it is calculated analytically from Equation 4.33. Information-optimized flip angles are computed by solving the optimization problem in Equation 4.38

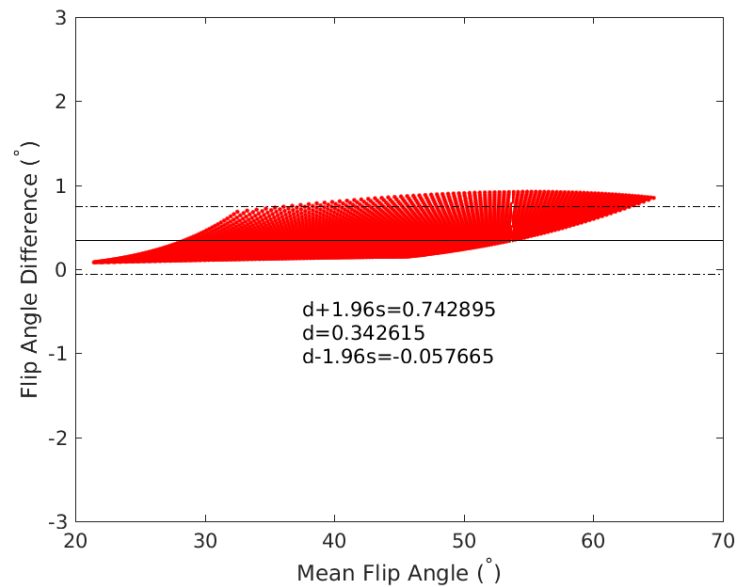


FIGURE 4.7: The Bland Altman plot of 6561 measurement pairs of the analytical Ernst angle and the linear information-optimized flip angle in a spoiled GRE pulse sequence over a range of T1 and TR values. Each measurement pair is plotted with a red dot. The mean measurement difference,  $d$ , is shown as a black line. The bounds of the 95% confidence interval assuming normally distributed measurement differences are shown as black dash-dot lines.

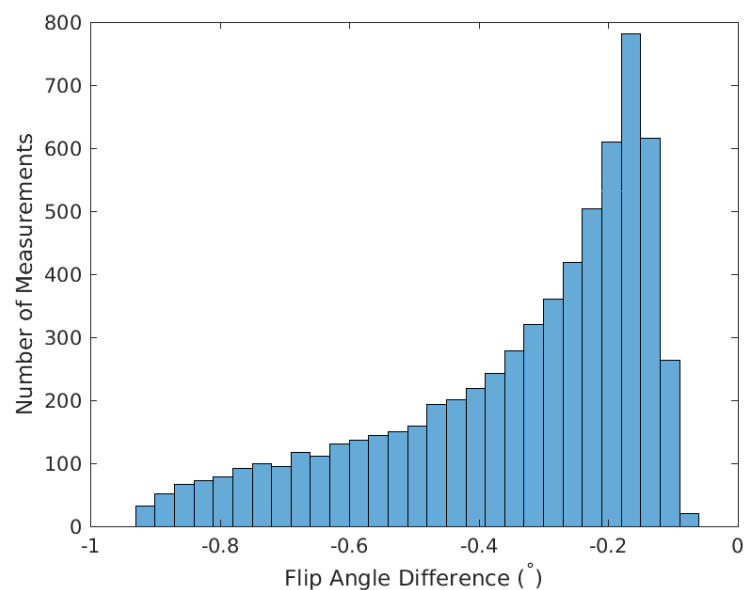


FIGURE 4.8: Histogram of differences between measurement pairs of the analytical Ernst angle and the linear information-optimized flip angle in a spoiled GRE pulse sequence over a range of T1 and TR values.

and calculating mutual information as in Equation 4.30. Information gain is quantified by mutual information between measurement and tissue property uncertainties. Greater information gain corresponds to greater measurement SNR. The maximum relative error is approximately 0.005%.

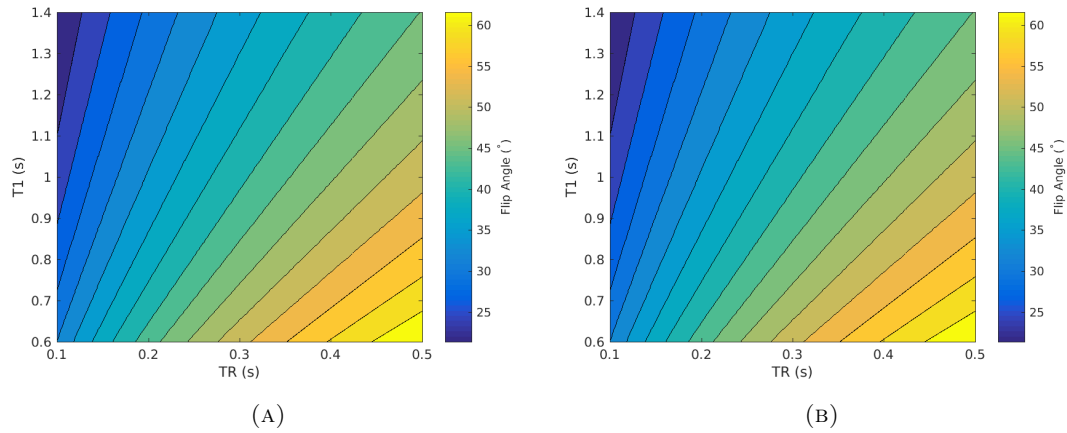


FIGURE 4.9: (A) Theoretical Ernst angle calculated over a range of T1 and TR values. (B) Optimal flip angle selected by nonlinear model of information gain from spoiled GRE acquisitions over a range of T1 and TR values.

Figure 4.10 shows the regression line between the analytical Ernst angle and the nonlinear information-optimized flip angle in a spoiled GRE pulse sequence over the same range of T1 and TR values. Each plotted point corresponds to the analytical Ernst angle (abscissa) and the information-optimized flip angle (ordinate) at identical T1 and TR. The regression line slope and intercept are 0.9999 and 0.0009, respectively. The adjusted coefficient of determination ( $r^2$ ) is 1.0000. The sum of squared errors (SSE) and root-mean-square error (RMSE) are 0.0015 and 0.0005, respectively.

Figure 4.11 shows a Bland Altman plot of the analytical Ernst angle and the nonlinear information-optimized flip angle measurement pairs plotted in Figures 4.9 and 4.10. The mean measurement difference,  $d$ , is  $-0.0018^\circ$ . The bounds of the 95% confidence interval (defined by  $d \pm 1.96s$ , where  $s$  is the standard deviation of measurement

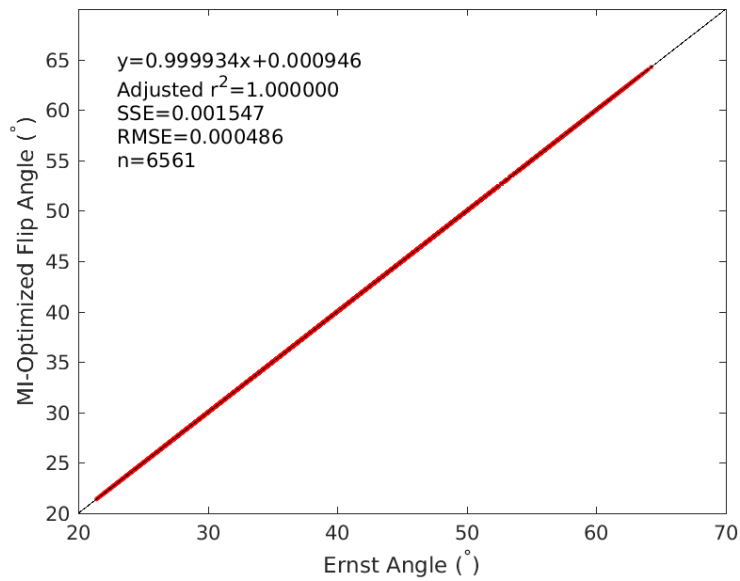


FIGURE 4.10: The regression line between 6561 measurement pairs of the analytical Ernst angle and the nonlinear information-optimized flip angle in a spoiled GRE pulse sequence over a range of T1 and TR values. Each measurement pair is plotted with a red dot. The regression line is shown in black. Fit parameters and goodness of fit metrics are displayed in the upper left corner.

differences) are  $-0.0033^\circ$  and  $-0.0003^\circ$ . Thus, 95% of all information-optimized flip angles obtained with the nonlinear MI model will differ from the analytical Ernst angle by between  $-0.0033^\circ$  and  $-0.0003^\circ$ .

#### 4.4.3 Justification of Model Assumptions for 3D-QALAS

The multivariate normal prior distribution is shown in Figure 4.12. The sampled posterior distribution is shown in Figure 4.13. The QALAS signal model introduces some covariance between the parameters. The Pearson correlation coefficient is used to evaluate the strength of the correlation between parameters. The matrix of Pearson correlation coefficients is shown below for the simulated posterior distribution resulting from the multivariate normal prior distribution.

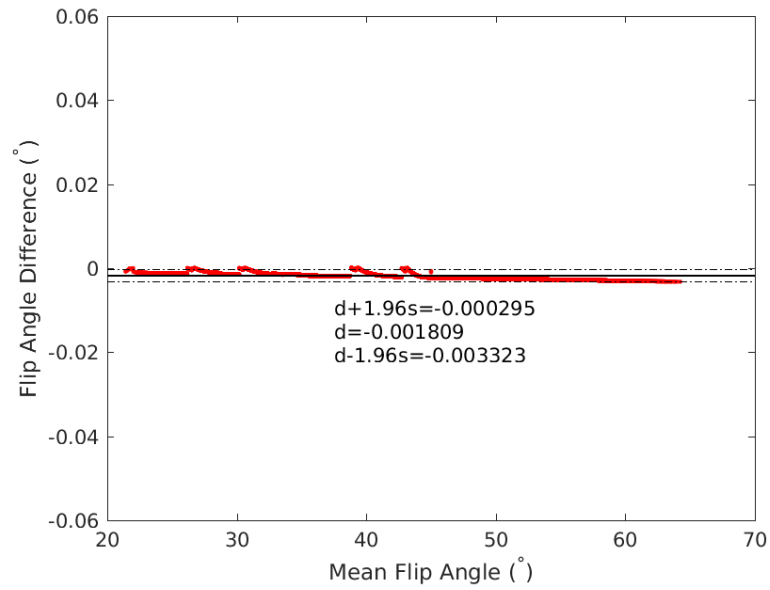


FIGURE 4.11: The Bland Altman plot of 6561 measurement pairs of the analytical Ernst angle and the nonlinear information-optimized flip angle in a spoiled GRE pulse sequence over a range of T1 and TR values. Each measurement pair is plotted with a red dot. The mean measurement difference,  $d$ , is shown as a black line. The bounds of the 95% confidence interval are shown as black dash-dot lines.

$\rho_{\text{normal}} =$

$$\begin{bmatrix} 1.000 & 0.918 & -0.353 & 0.001 & 0.005 & 0.003 & 0.030 & 0.018 & -0.052 \\ 0.918 & 1.000 & -0.239 & -0.003 & -0.004 & -0.007 & 0.004 & -0.006 & -0.044 \\ -0.353 & -0.239 & 1.000 & -0.050 & -0.035 & 0.016 & -0.011 & 0.004 & 0.020 \\ 0.001 & -0.003 & -0.050 & 1.000 & 0.904 & -0.227 & 0.050 & 0.056 & 0.031 \\ 0.005 & -0.004 & -0.035 & 0.904 & 1.000 & -0.166 & 0.054 & 0.057 & 0.018 \\ 0.003 & -0.007 & 0.016 & -0.227 & -0.166 & 1.000 & -0.008 & 0.001 & -0.016 \\ 0.030 & 0.004 & -0.011 & 0.050 & 0.054 & -0.008 & 1.000 & 0.940 & -0.213 \\ 0.018 & -0.006 & 0.004 & 0.056 & 0.057 & 0.001 & 0.940 & 1.000 & -0.132 \\ -0.052 & -0.044 & 0.020 & 0.031 & 0.018 & -0.016 & -0.213 & -0.132 & 1.000 \end{bmatrix}$$

(4.40)

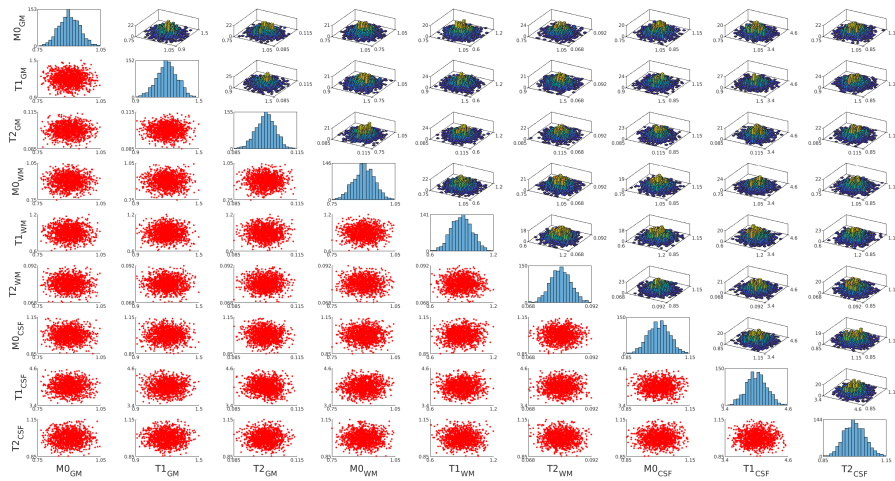


FIGURE 4.12: Assumed normal prior distributions of M0, T1, and T2 for gray matter, white matter, and cerebrospinal fluid. Ellipsoidal nested sampling is used to sample the posterior distributions resulting from these prior distributions. Diagonal entries on the correlation matrix show distributions of each combination of parametric map value and tissue type. Non-diagonal lower-triangular entries show a two-dimension scatter plot of the joint distribution between the parameters labeled on the same row and column. Non-diagonal upper-triangular entries show a two-dimensional histogram of the joint distribution between the parameters labeled on the same row and column.

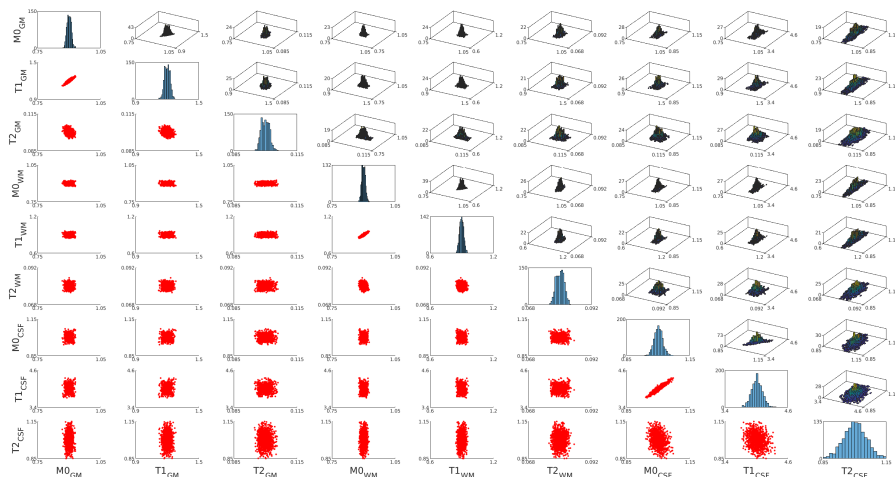


FIGURE 4.13: Posterior distributions of M0, T1, and T2 for gray matter, white matter, and cerebrospinal fluid resulting from ellipsoidal nested sampling and normal prior distributions shown in Figure 4.12. Diagonal entries on the correlation matrix show distributions of each combination of parametric map value and tissue type. Non-diagonal lower-triangular entries show a two-dimension scatter plot of the joint distribution between the parameters labeled on the same row and column. Non-diagonal upper-triangular entries show a two-dimensional histogram of the joint distribution between the parameters labeled on the same row and column.

In particular, there are strong correlations between M0 and T1 values for identical tissues. There are no significant correlations between other parameters and tissues. This process was repeated for independent uniform prior distributions, where  $\rho(\eta_i) \sim U(-2\sigma_{\eta_i}, 2\sigma_{\eta_i})$ . The uniform prior distributions are shown in Figure 4.14. The sampled posterior distribution is shown in Figure 4.15. Again, the QALAS signal model introduces some covariance between the parameters. The matrix of Pearson correlation coefficients is shown below for the multivariate uniform prior distribution.

$$\rho_{\text{uniform}} = \begin{bmatrix} 1.000 & 0.927 & -0.370 & 0.043 & 0.031 & -0.094 & 0.023 & 0.018 & -0.019 \\ 0.927 & 1.000 & -0.303 & 0.050 & 0.044 & -0.104 & 0.038 & 0.029 & -0.021 \\ -0.370 & -0.303 & 1.000 & -0.031 & -0.020 & 0.015 & 0.034 & 0.047 & 0.024 \\ 0.043 & 0.050 & -0.031 & 1.000 & 0.908 & -0.238 & 0.019 & 0.035 & -0.016 \\ 0.031 & 0.044 & -0.020 & 0.908 & 1.000 & -0.159 & -0.008 & 0.012 & -0.013 \\ -0.094 & -0.104 & 0.015 & -0.238 & -0.159 & 1.000 & -0.002 & -0.003 & 0.029 \\ 0.023 & 0.038 & 0.034 & 0.019 & -0.008 & -0.002 & 1.000 & 0.963 & -0.321 \\ 0.018 & 0.029 & 0.047 & 0.035 & 0.012 & -0.003 & 0.963 & 1.000 & -0.254 \\ -0.019 & -0.021 & 0.024 & -0.016 & -0.013 & 0.029 & -0.321 & -0.254 & 1.000 \end{bmatrix} \quad (4.41)$$

Strong correlations exist between M0 and T1 values for identical tissues. No other significant correlations exist between other parameters and tissues. Furthermore, the sampled posterior distribution is nearly identical to that resulting from the multivariate normal prior distribution.



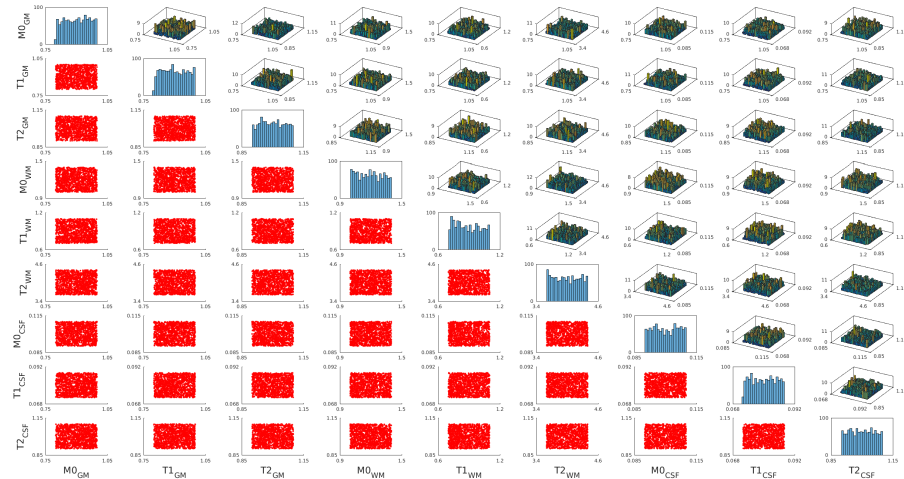


FIGURE 4.14: Assumed uniform prior distributions of M0, T1, and T2 for gray matter, white matter, and cerebrospinal fluid. Ellipsoidal nested sampling is used to sample the posterior distributions resulting from these prior distributions. Diagonal entries on the correlation matrix show distributions of each combination of parametric map value and tissue type. Non-diagonal lower-triangular entries show a two-dimension scatter plot of the joint distribution between the parameters labeled on the same row and column. Non-diagonal upper-triangular entries show a two-dimensional histogram of the joint distribution between the parameters labeled on the same row and column.

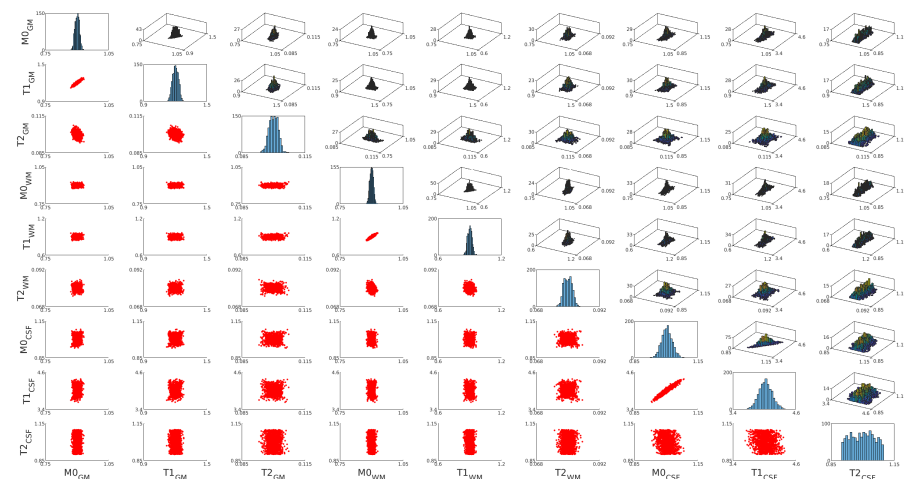


FIGURE 4.15: Posterior distributions of M0, T1, and T2 for gray matter, white matter, and cerebrospinal fluid resulting from ellipsoidal nested sampling and uniform prior distributions shown in Figure 4.14. Diagonal entries on the correlation matrix show distributions of each combination of parametric map value and tissue type. Non-diagonal lower-triangular entries show a two-dimension scatter plot of the joint distribution between the parameters labeled on the same row and column. Non-diagonal upper-triangular entries show a two-dimensional histogram of the joint distribution between the parameters labeled on the same row and column.

Finally, the distribution of parametric map values taken from in vivo data is compared to the simulated posterior distributions. Figure 4.16 shows the distribution of measured parameter values. The matrix of Pearson correlation coefficients is used to evaluate correlation between parameters and is shown below for in vivo 3D-QALAS measurements:

$$\rho_{\text{measured}} = \begin{bmatrix} 1.000 & 0.722 & -0.073 & -0.016 & -0.012 & 0.023 & -0.031 & -0.043 & 0.027 \\ 0.722 & 1.000 & 0.031 & 0.023 & 0.042 & 0.004 & -0.001 & -0.056 & 0.068 \\ -0.073 & 0.031 & 1.000 & -0.026 & -0.021 & -0.095 & 0.092 & 0.055 & -0.029 \\ -0.016 & 0.023 & -0.026 & 1.000 & 0.970 & -0.121 & -0.041 & 0.002 & -0.009 \\ -0.012 & 0.042 & -0.021 & 0.970 & 1.000 & -0.084 & -0.057 & -0.006 & -0.010 \\ 0.023 & 0.004 & -0.095 & -0.121 & -0.084 & 1.000 & -0.001 & 0.048 & -0.001 \\ -0.031 & -0.001 & 0.092 & -0.041 & -0.057 & -0.001 & 1.000 & 0.777 & -0.029 \\ -0.043 & -0.056 & 0.055 & 0.002 & -0.006 & 0.048 & 0.777 & 1.000 & -0.057 \\ 0.027 & 0.068 & -0.029 & -0.009 & -0.010 & -0.001 & -0.029 & -0.057 & 1.000 \end{bmatrix} \quad (4.42)$$

Significant correlation exists only between M0 and T1 for identical tissues. This agrees very well with correlations observed in the simulated posterior distributions.

## 4.5 Discussion

In this aim, a general information theory framework for the optimization of acquisition parameters in MR pulse sequences was developed. The objective function for this

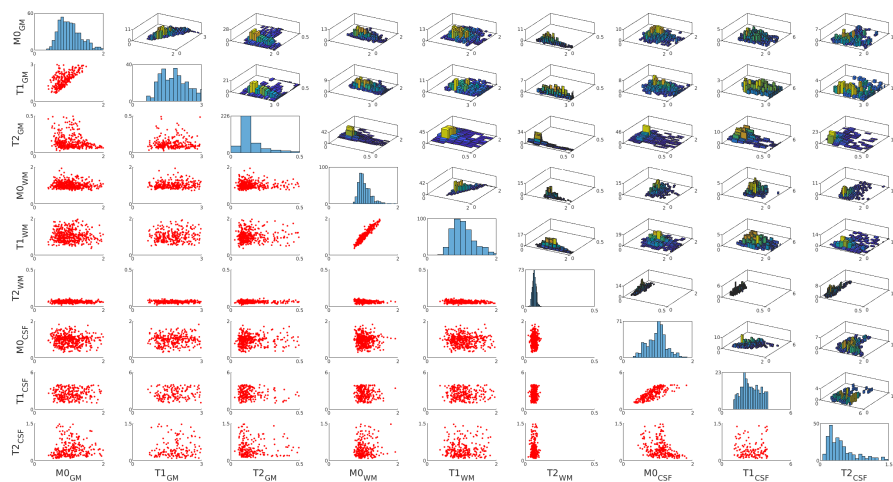


FIGURE 4.16: Correlation matrix for M0, T1, and T2 values measured from in vivo data in gray matter, white matter, and cerebrospinal fluid. Diagonal entries show distributions of each combination of parametric map value and tissue type. Non-diagonal lower-triangular entries show a two-dimension scatter plot of the joint distribution between the parameters labeled on the same row and column. Non-diagonal upper-triangular entries show a two-dimensional histogram of the joint distribution between the parameters labeled on the same row and column.

optimization was mutual information between subject T1 uncertainty and acquisition uncertainty. Both a linear approximation and a full nonlinear computation of mutual information were verified on a simulated spoiled GRE pulse sequence in which the optimized acquisition parameter was flip angle. Selection of a relatively simple signal model enabled optimization results from the information theory framework to be directly compared to a known analytical optimum value. Model assumptions were justified for 3D-QALAS by comparison to actual 3D-QALAS measurements.

Figure 4.3a illustrates suitability of mutual information for maximizing information gain. Mutual information correlates with SNR, suggesting that information content of an image acquisition, as measured by mutual information, is a predictor of image quality. This is potentially invaluable as a tool for optimization of image quality before imaging occurs in scenarios where scanner time is limited or opportunities for rescanning do not exist.

The information-optimized flip angles relying on the linear approximation to mutual information show fairly good agreement with the known Ernst angle (Figure 4.4). The maximum relative error is slightly greater than 2% (Figure 4.5). However, the relative error shows concerning bias and appears to be a function of T1 and TR. The regression line and Bland Altman plot (Figures 4.6 and 4.7) confirm the bias of linear method. Mean optimization error is  $0.3426^\circ$ , which is non-negligible. Furthermore, optimization error is not normally distributed (Figure 4.8).

The nonlinear mutual information calculation is more computationally expensive, but it allows the model evidence to take any general non-Gaussian form. This results in the elimination of bias from the information optimization. Figure 4.9 shows the nonlinear solution also agrees reasonably well with analytical values. The regression line between information optimized flip angles and Ernst angles in Figure 4.10 shows an improvement of agreement with the analytical solution compared to that of the linear approximation. This is reflected in the regression line parameters and goodness of fit metrics. The adjusted  $r^2$  value was used to attempt to compensate for the large number of points fit by the regression line. Even so, the difference in adjusted  $r^2$  between the linear ( $r^2 = 0.9997$ ) and nonlinear ( $r^2 = 1.000$ ) methods is quite small. SSE and RMSE both decrease multiple orders of magnitude—a more appreciable difference between the two results. SSE decreases from 162.7 to 0.0015, and RMSE decreases from 0.1575 to 0.0005.

The Bland Altman plot for the nonlinear method in Figure 4.11 shows near elimination of bias, a significant reduction in error, and elimination of non-Gaussian behavior in error. These results suggest that the more computationally intensive approach is necessary to capture nonlinearities in even the simplest signal models. It is likely that a

linear approximation to mutual information would introduce error and bias in analyses of more complex signal models as well.

Strong correlations can be seen between M0 and T1 values for identical tissues in Figures 4.13 and 4.15. No other significant correlations exist between other parameters and tissues. Furthermore, the sampled posterior distribution is nearly identical to that resulting from the multivariate normal prior distribution. This suggests that there is at least some degree of robustness to the selection of prior distribution. For real 3D-QALAS data, significant correlation exists only between M0 and T1 for identical tissues, as seen in Figure 4.16. This agrees very well with correlations observed in both of the simulated posterior distributions. These results suggest that correlation is introduced by the 3D-QALAS signal model (in particular, the curve fitting during which M0 and T1 are simultaneously determined) and justifies using prior distributions with independent variables.

## Chapter 5

# Information Quantification of Acquisition Parameters for 3D-QALAS in Phantom

### 5.1 Introduction

In this chapter, the feasibility of the information theory framework is tested on 3D-QALAS acquisitions, a signal model with significantly increased complexity. A computational methodology is developed to quantify information content of 3D-QALAS acquisitions relative to a representative synthetic brain model and validated in phantom measurements. The hypothesis that information-optimized acquisition parameters for a 3D-QALAS signal model correlate to smaller variances in reconstructed parametric maps is tested.

## 5.2 Theory

### 5.2.1 3D-QALAS Signal Model

The signal model of interest is 3D-QALAS (3D-Quantification using an interleaved Look-Locker acquisition sequence with T2 preparation pulse).[32] 3D-QALAS is described briefly in Section 2.2.2. The pulse sequence can be generally divided into a T2-sensitization phase and a T1-sensitization phase. The longitudinal magnetization over the two phases is modeled mathematically in the next section (Section 5.2.2).  $M_0$ , T1, and T2 parametric maps are fitted to measurements from five gradient echo acquisitions over the two phases. The original application of 3D-QALAS is rapid cardiac mapping, but an adaptation for 3D multi-parameter quantification in the brain, which has different timing constraints, is used in this work.

The QALAS forward solve,  $\mathcal{Q}_\mu$ , describes the set of  $N_{\text{acq}}$  measurements,  $\mathbf{M}$ , retrieved at a combination of acquisition parameters,  $\boldsymbol{\mu}$ , for a subject with properties  $\boldsymbol{\eta} = [M_0(\mathbf{x}), T_1(\mathbf{x}), T_2(\mathbf{x})]$  in image space  $\mathbf{x} \in \Omega$ , where  $\Omega$  is the synthetic phantom domain:

$$\mathbf{M}(\boldsymbol{\mu}, \mathbf{x}) = \mathcal{Q}_\mu(M_0(\mathbf{x}), T_1(\mathbf{x}), T_2(\mathbf{x})) = \mathcal{Q}_\mu(\boldsymbol{\eta}(\mathbf{x})). \quad (5.1)$$

Acquisition parameters can generally include flip angle, TR, TE,  $\text{TE}_{\text{T2prep}}$ , TI, number of acquisitions, acquisition times, and delay times. The QALAS inverse solve describes the fitting of the parametric map values,  $\boldsymbol{\eta}$ , such that the norm of the difference between the measurements,  $\mathbf{M}_{\text{meas}}$ , and the modeled signal,  $\mathcal{Q}(\boldsymbol{\eta})$ , is minimized:

$$\boldsymbol{\eta}_{\text{meas}} = \underset{\boldsymbol{\eta}}{\text{argmin}} (\|\mathbf{M}_{\text{meas}} - \mathcal{Q}(\boldsymbol{\eta})\|_2). \quad (5.2)$$

The QALAS signal model used in this work,  $\mathbf{z}$ , is the Fourier transform of the QALAS forward solve, subsampled by subsampling mask  $S$ , which may generally be a function of subsampling parameters  $\boldsymbol{\theta}$ , and subjected to normally distributed noise  $\nu$ :

$$\begin{aligned} \mathbf{z}(\boldsymbol{\mu}, \boldsymbol{\theta}, \mathbf{k}) &= S_{\boldsymbol{\theta}} \odot \int_{\Omega} \mathbf{M}(\boldsymbol{\mu}, \mathbf{x}) e^{-2\pi i \mathbf{k} \cdot \mathbf{x}} d\mathbf{x} + \nu(\mathbf{k}) = S_{\boldsymbol{\theta}} \odot \mathcal{FQ}_{\boldsymbol{\mu}}(\boldsymbol{\eta}) + \nu(\mathbf{k}) \\ \nu &\sim \mathcal{N}(\mathbf{0}, \Sigma_{\nu}) \\ \Sigma_{\nu} &= \begin{bmatrix} \sigma_{\nu,r}^2 & 0 \\ 0 & \sigma_{\nu,i}^2 \end{bmatrix}. \end{aligned} \tag{5.3}$$

The model is defined such that the real and imaginary components of the noise are independent and identically distributed ( $\sigma_{\nu,r}^2 = \sigma_{\nu,i}^2$ ).

The probability of observing a measurement  $\mathbf{z}$  made in tissue with properties  $\boldsymbol{\eta}$ , then, is normally distributed about the model evaluation  $S \odot \mathcal{Q}$  with covariance  $\Sigma_{\nu}$ :

$$p(\mathbf{z}|\boldsymbol{\eta}) \sim \mathcal{N}(S_{\boldsymbol{\theta}} \odot \mathcal{FQ}_{\boldsymbol{\mu}}(\boldsymbol{\eta}), \Sigma_{\nu}). \tag{5.4}$$

Because the noise components are independent and identically distributed, the multivariate distribution  $p(\mathbf{z}|\boldsymbol{\eta})$  is equal to the product of the univariate distributions  $p(z_r|\boldsymbol{\eta})$  and  $p(z_i|\boldsymbol{\eta})$ :

$$\begin{aligned} p(\mathbf{z}|\boldsymbol{\eta}) &\sim \mathcal{N}(S_{\boldsymbol{\theta}} \odot \mathcal{FQ}_{\boldsymbol{\mu}}(\boldsymbol{\eta}), \Sigma_{\nu}) \\ &= \mathcal{N}_{z_r}(S_{\boldsymbol{\theta}} \odot \Re(\mathcal{FQ}_{\boldsymbol{\mu}}(\boldsymbol{\eta})), \sigma_{\nu_r}) \cdot \mathcal{N}_{z_i}(S_{\boldsymbol{\theta}} \odot \Im(\mathcal{FQ}_{\boldsymbol{\mu}}(\boldsymbol{\eta})), \sigma_{\nu_i}) \\ &\sim p(z_r|\boldsymbol{\eta})p(z_i|\boldsymbol{\eta}). \end{aligned} \tag{5.5}$$



### 5.2.2 3D-QALAS Operator

The QALAS forward solve operator,  $Q$ , mentioned previously is defined by a series of recursive equations describing the longitudinal magnetization during various points in the acquisition. During delay, the longitudinal magnetization relaxes according to the following equation:

$$M_{n+1} = M_0 - (M_0 - M_n) \cdot e^{-\Delta t/T_1}. \quad (5.6)$$

During signal acquisition the longitudinal magnetization relaxes toward  $M_0^*$  at rate  $T_1^*$  according to the following two equations:

$$\frac{T_1^*}{T_1} = \frac{M_0^*}{M_0} = \frac{1 - e^{-T_R/T_1}}{1 - \cos(\alpha) \cdot e^{-T_R/T_1}} \equiv S \quad (5.7)$$

$$M_{n+1} = M_0^* - (M_0^* - M_n) \cdot e^{-\Delta t/T_1^*} = SM_0 - (SM_0 - M_n) \cdot e^{-\Delta t/ST_1}. \quad (5.8)$$

Figure 5.1 shows the full QALAS acquisition sequence.

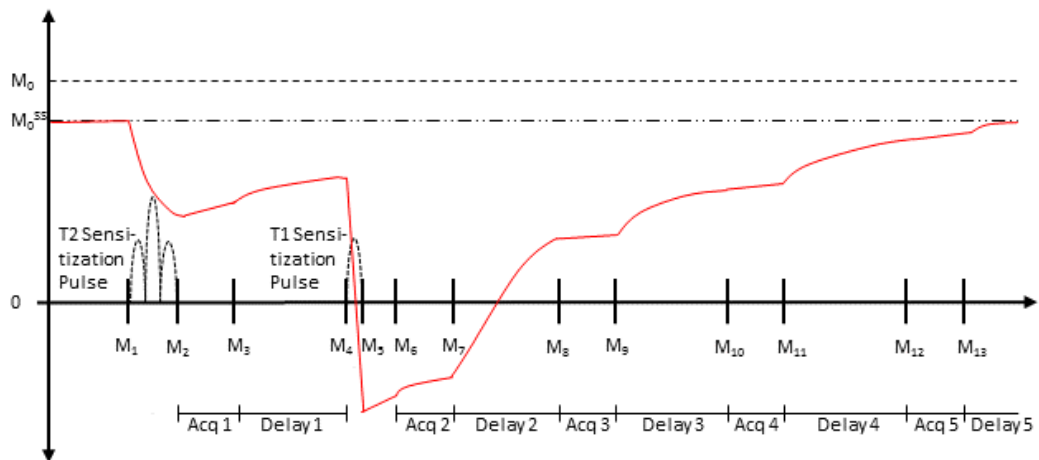


FIGURE 5.1: Diagram of QALAS acquisition sequence. The longitudinal magnetization is shown in red as a function of time. The five acquisition and delay times are labeled at the bottom.

### 5.2.3 Mutual Information

In general, measurement errors, noise, and other perturbations usually make it impossible to know inputs to systems exactly. For this reason, inputs can often be modeled as random variables, and the propagation of this uncertainty also makes the output of the system a random variable. For this model, the five 3D-QALAS acquisitions are the random variables inputted into the system. Their distributions are selected to simulate machine noise that would cause uncertainty in the measurements. This uncertainty is propagated through the model, resulting in uncertainty in the M0, T1, and T2 maps, which are the system outputs. The uncertainty of the system outputs is inversely related to the reproducibility of the parametric maps. The goal is to minimize the uncertainty of these system outputs by taking a measurement with the optimal parameters.

The information gain is quantified from a measurement using mutual information. Mutual information is a quantity that describes the amount of information one random variable contains about another random variable. Alternatively, it is the reduction in one random variable's uncertainty owing to knowledge of another variable, i.e. the difference between the marginal entropy and the conditional entropy:

$$I(\mathbf{z}; \boldsymbol{\eta}) = H(\mathbf{z}) - H(\mathbf{z}|\boldsymbol{\eta}). \quad (5.9)$$

In this framework, mutual information measures the reduction in uncertainty of the parametric maps resulting from the 3D-QALAS signal model owing to knowledge of the measured signal.

For some total entropy of a joint probability function of measurements,  $\mathbf{z}$ , and parametric maps,  $\boldsymbol{\eta}$ , the entropy of  $\boldsymbol{\eta}$ ,  $H(\boldsymbol{\eta})$ , is reduced by a known measurement to the

entropy of  $\boldsymbol{\eta}$  given  $\boldsymbol{z}$ . The measurement of this reduction is mutual information. The optimization problem to find the best set of acquisition parameters,  $\boldsymbol{\mu}^*$ , is written

$$\boldsymbol{\mu}^* = \underset{\boldsymbol{\mu}}{\operatorname{argmax}} I(\boldsymbol{z}(\boldsymbol{\mu}); \boldsymbol{\eta}) = \underset{\boldsymbol{\mu}}{\operatorname{argmax}} (H(\boldsymbol{z}(\boldsymbol{\mu})) - H(\boldsymbol{z}(\boldsymbol{\mu})|\boldsymbol{\eta})). \quad (5.10)$$

In this equation,  $H(\boldsymbol{z}|\boldsymbol{\eta})$  is known and is a function only of the normally distributed model machine noise:

$$H(\boldsymbol{z}|\boldsymbol{\eta}) = \frac{1}{2} \ln ((2\pi e)^2 \cdot |\Sigma_{\boldsymbol{\eta}}|) = \frac{1}{2} \ln ((2\pi e)^2 \sigma_{\nu,r}^2 \sigma_{\nu,i}^2). \quad (5.11)$$

The difficulty is in calculating the entropy  $H(\boldsymbol{z})$  of the non-Gaussian distribution  $p(\boldsymbol{z})$ , which is a function of a nonlinear signal model. A detailed description of the mutual information computation is in Appendix C.

Mutual information for a set of acquisition parameters corresponds to this uncertainty reduction for a signal measurement at this set of acquisition parameters. The greatest mutual information exists for the set of acquisition parameters for which a measurement produces M0, T1, and T2 maps with the least uncertainty. Thus, mutual information can be optimized over this acquisition parameter space to find the optimal set of parameters.

### 5.3 Methods

Experiments were performed on both synthetic phantom data and the ISMRM NIST qMRI Multiparametric Imaging Standard phantom (QalibreMD, Boulder, CO).[195] The synthetic phantom data were used to optimize the selected acquisition parameters (five independent delay times) with respect to mutual information. Multiple measurements of

the ISMRM NIST system phantom with different combinations of the selected acquisition parameters were used to validate the synthetic phantom mutual information model by measuring the correlation between mutual information and reconstruction variance. These processes are detailed below and shown schematically in Figures 5.2 and 5.3.

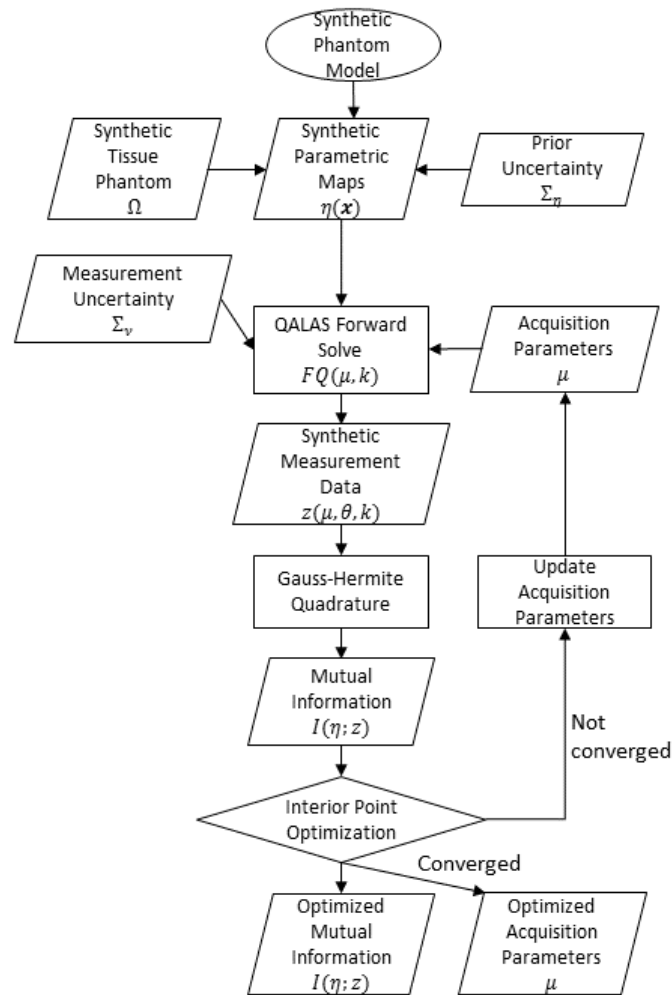


FIGURE 5.2: Flowchart for the mutual information-based optimization of acquisition parameters in a synthetic model.

### 5.3.1 Synthetic Phantom Definition

To test the information theory framework, measurements were simulated on a mathematical phantom segmented into gray matter, white matter, and cerebrospinal fluid

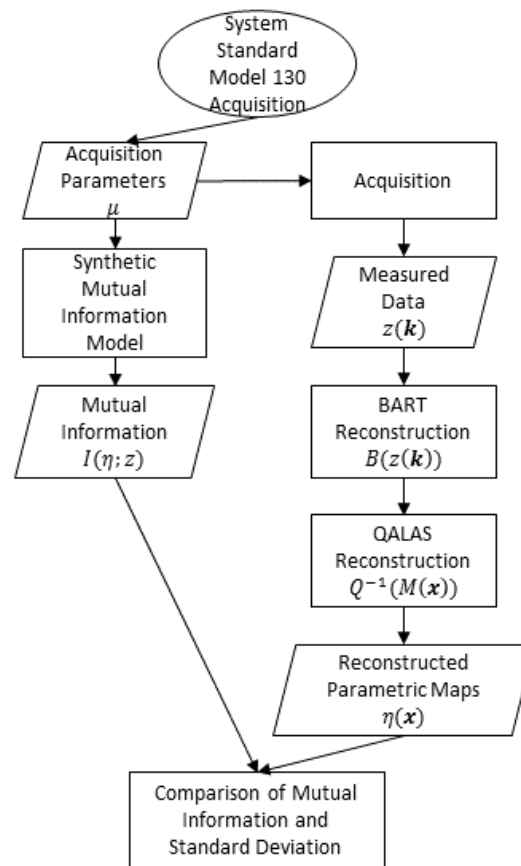


FIGURE 5.3: Flowchart for the application of the synthetic model to real-world measurements in the System Standard Model 130 phantom is shown on the right.

(CSF). A two-dimensional cross-section of this synthetic phantom is shown in Figure 5.4. Gold standard measurements were synthesized using literature values for proton densities and relaxation times, as well as acquisition parameter values from the current evaluation point in acquisition parameter space. To support the assumption that this model is a representative population average, mutual information sensitivity to geometry was approximated by dilating and eroding tissue boundaries.

The synthetic phantom on domain  $\Omega$  is defined by  $N$  mutually disjoint tissue labels:

$$\begin{aligned}
 \boldsymbol{\eta}(\mathbf{x}) &= \sum_{n=1}^N \boldsymbol{\eta}_n U(\mathbf{x} - \Omega_n) \\
 \bigcup_{n=1}^N \Omega_n &= \Omega \\
 \Omega_n \cap \Omega_m &= \emptyset \quad \text{for } n \neq m \\
 U(\mathbf{x} - \Omega_n) &= \begin{cases} 1, & \mathbf{x} \in \Omega_n \\ 0, & \text{otherwise} \end{cases}
 \end{aligned} \tag{5.12}$$

Physical tissue properties,  $\boldsymbol{\eta}(\mathbf{x}) = [M_0(\mathbf{x}), T_1(\mathbf{x}), T_2(\mathbf{x})]$ , for tissue  $n$  (white matter, gray matter, CSF, etc.) are normally distributed about literature values  $\mathbf{m}_{\eta,n}$  with covariance matrix  $\Sigma_{\eta,n}$ :

$$\begin{aligned}
 p_n(\boldsymbol{\eta}) &\sim \mathcal{N}(\mathbf{m}_{\eta,n}, \Sigma_{\eta,n}) \\
 \mathbf{m}_{\eta,n} &= \begin{bmatrix} m_{\eta_1} \\ \vdots \\ m_{\eta_N} \end{bmatrix} \\
 \Sigma_{\eta,n} &= \begin{bmatrix} \sigma_{\eta_1}^2 & 0 & \dots & 0 & 0 \\ 0 & \sigma_{\eta_2}^2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \sigma_{\eta_{N-1}}^2 & 0 \\ 0 & 0 & \dots & 0 & \sigma_{\eta_N}^2 \end{bmatrix}
 \end{aligned} \tag{5.13}$$

The acquisition parameters explored were flip angle and four independent delay times between the T1 sensitization phase acquisitions during which the longitudinal magnetization relaxes. The dimensionality of the acquisition parameter space can be increased by including TR, TE of the T2 preparation pulse, acquisition length, and

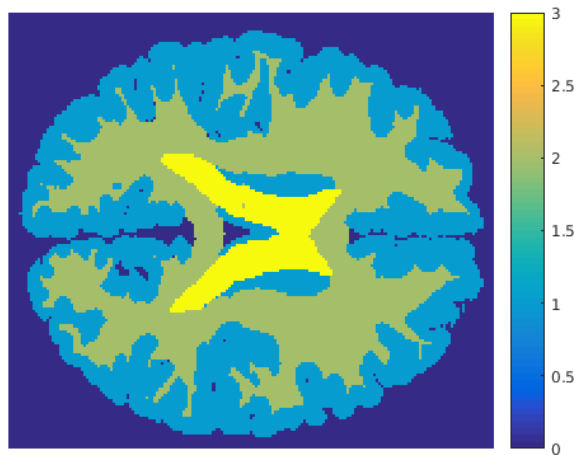


FIGURE 5.4: A two-dimensional axial slice of the tissue labels used to define the synthetic phantom. Tissue label 1 is gray matter, 2 is white matter, and 3 is cerebrospinal fluid.

various parameters defining the subsampling strategy, but for the sake of computation time, these parameters were held constant. Optimization was performed over this acquisition parameter space using an interior point method with mutual information as the objective function to be maximized. For a full derivation of the mutual information calculation used, see Appendix C.

To validate the optimization, a clustered ellipsoidal nested sampling method was used to sample the distributions of the parametric map values after simulating a measurement. Greater uncertainty reduction should result in narrower distributions. The prior distributions for T1 and T2 in gray matter, white matter, and CSF are shown in Table 5.1. Several acquisition parameters are fixed to constant values ( $TR = 2.6$  ms,  $TI = 30$  ms, T2 preparation pulse echo time ( $TE_{T2\text{prep}}$ ) = 100 ms, acquisition time ( $T_{\text{acq}}$ ) = 338 ms, number of gradient echo acquisitions ( $N_{\text{acq}} = 5$ ).

Parameter		GM	WM	CSF
T1 (ms)	Mean	1400	1000	4000
	Std. dev.	100	100	100
T2 (ms)	Mean	100	75	600
	Std. dev.	5	5	30
M0	Mean	0.8	0.8	1.0
	Std. dev.	0.1	0.1	0.1

TABLE 5.1: Prior distributions for tissue properties (M0, T1, and T2) in three tissues (gray matter, white matter, and cerebrospinal fluid).

### 5.3.2 Phantom Acquisitions

To validate the mutual information model, it was applied to actual measurements of the ISMRM NIST system phantom (Figure 5.5). This phantom has 14 T1 elements and 14 T2 elements with well-characterized relaxation times. The phantom also has 14 proton density elements with known material properties. Two scans were performed with different sets of acquisition parameters using a 3.0T MR 750W MRI scanner (GE Healthcare, Waukesha, WI). Reconstruction was performed using geOrchestra and the Berkeley Advanced Reconstruction Toolbox (BART, The University of California, Berkeley, Berkeley, CA), with L2 regularization (Figure 5.6).[196] The acquisition parameters are listed in Table 5.2.

Parameter	Scan 1	Scan 2
Flip angle ( $^{\circ}$ )	4	4
TR (ms)	6.6	6.6
TI (ms)	100	100
TE <sub>T2prep</sub> (ms)	100	100
T <sub>acq</sub> (ms)	674.8	674.8
N <sub>acq</sub>	5	5
TD1 <sup>a</sup> (ms)	230.6	500
TD2 <sup>a</sup> (ms)	325.2	142.6
TD3 <sup>a</sup> (ms)	325.2	142.6
TD4 <sup>a</sup> (ms)	325.2	142.6
TD5 <sup>a</sup> (ms)	221.6	500.0

TABLE 5.2: Acquisition parameter values for two scans performed on ISMRM NIST system phantom. <sup>a</sup>TD1 is the time between the T1 sensitizing pulse and the first acquisition, TD2 is the time between the end of the first acquisition and the start of the second acquisition, and so on.



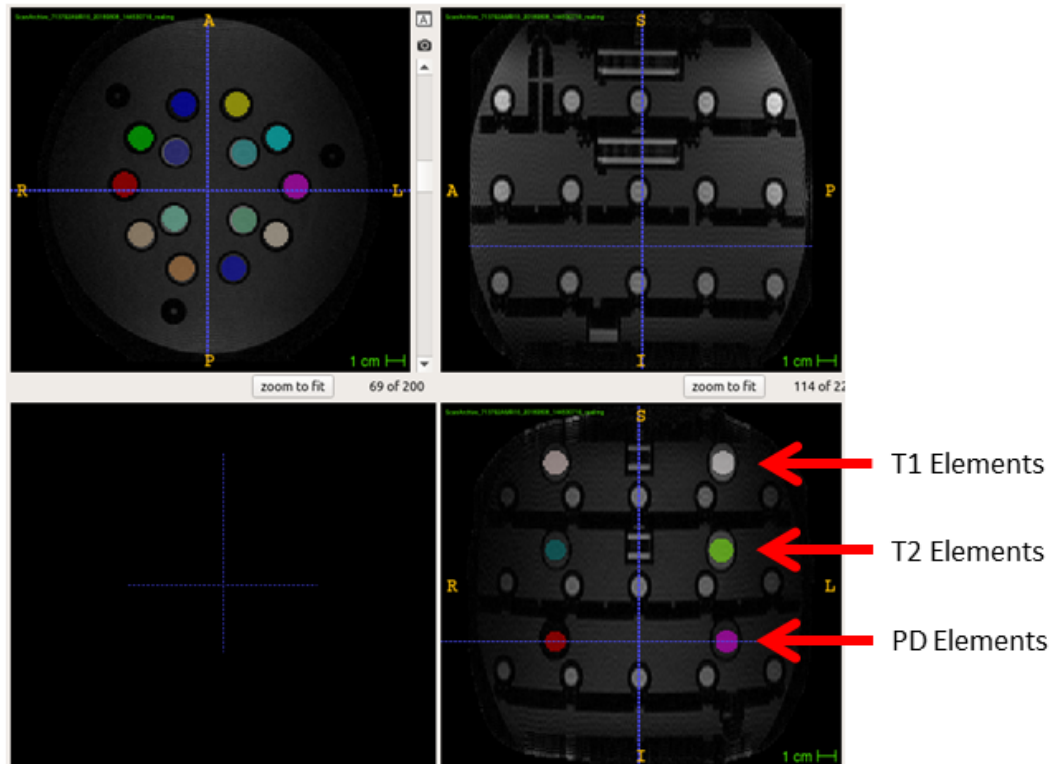


FIGURE 5.5: Measurements of the ISMRM NIST phantom with each of the 42 phantom elements segmented. The three layers of elements (PD, T1, and T2) are labeled. There are 14 elements in each layer.

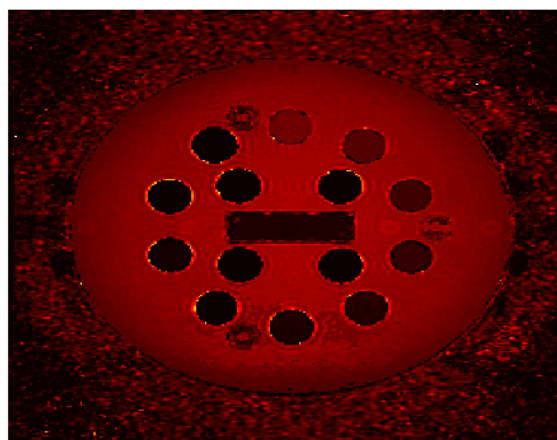


FIGURE 5.6: Example T1 parametric map reconstructed from measurements of the ISMRM NIST phantom.

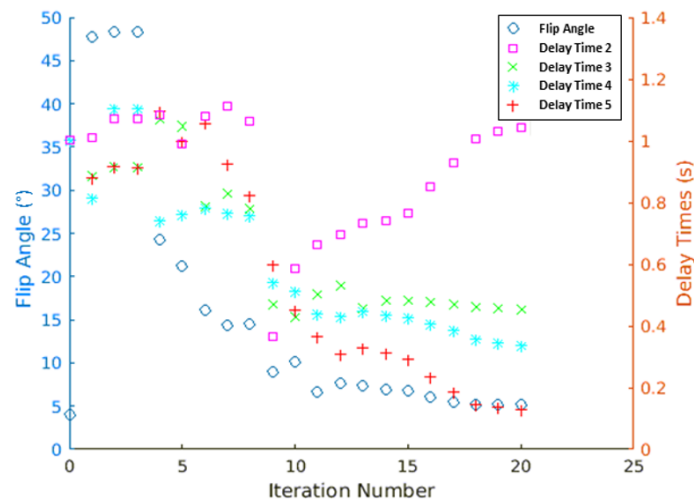
## 5.4 Results

### 5.4.1 Synthetic Model

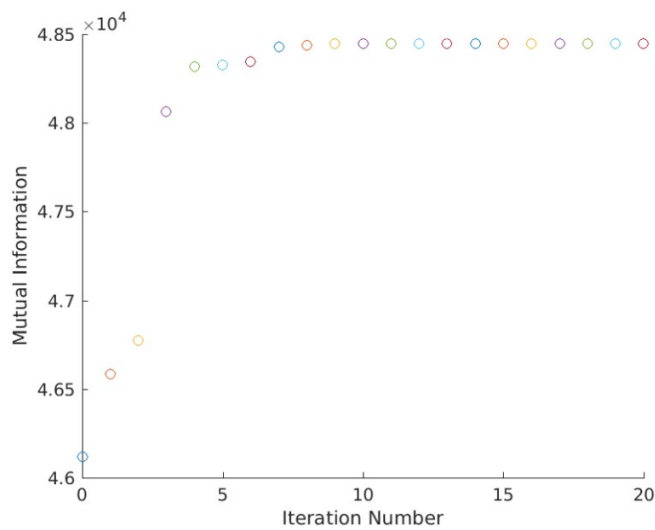
For a fixed set of acquisition parameters, dilating and eroding the tissue boundaries by one voxel resulted in relative changes in mutual information of less than 1%. For fixed geometry and a variable flip angle, mutual information decreased by almost 7% and increased by more than 2% for flip angles between  $2^\circ$  and  $6^\circ$ . For fixed geometry and variable delay times, mutual information increased by almost 2% and decreased by more than 2% for fixed delay times between 100 ms and 500 ms. This exercise demonstrates that the mutual information computed using the synthetic phantom model was much more sensitive to the acquisition parameters being optimized than the geometry of the tissue labels, which suggests that the synthetic phantom is adequately representative of a population average.

The calculation of mutual information in the synthetic phantom was used to optimize acquisition parameter selection. The acquisition parameter optimization history is displayed in Figure 5.7. The optimal parameters were approximately the following: flip angle =  $5^\circ$ , delay 2 = 1050 ms, delay 3 = 450 ms, delay 4 = 330 ms, and delay 5 = 125 ms. Delay 2 is the time between the T1 sensitizing pulse and the second acquisition, and the other delays are the times between the start of the corresponding acquisition and the end of the previous acquisition.

Ellipsoidal nested sampling was also used to generate and compare two posterior distributions from reconstructions of two measurements—one with acquisition parameters from a non-optimal point in parameter space and one with acquisition parameters



(A)



(B)

FIGURE 5.7: Evolution of acquisition parameters over the course of an optimization (A), and the optimization objective function, mutual information (B).

from the optimal location in parameter space, where mutual information and, consequently, the information gain from a measurement are greatest. As expected, narrower distributions were observed, which resulted from greater uncertainty reduction in all three parametric maps for all three tissues (Figure 5.8).

To better visualize the impact of the result of reducing the parametric map uncertainties, a reconstruction was simulated using a measurement with the initial, non-optimized acquisition parameters and the absolute error was calculated between these

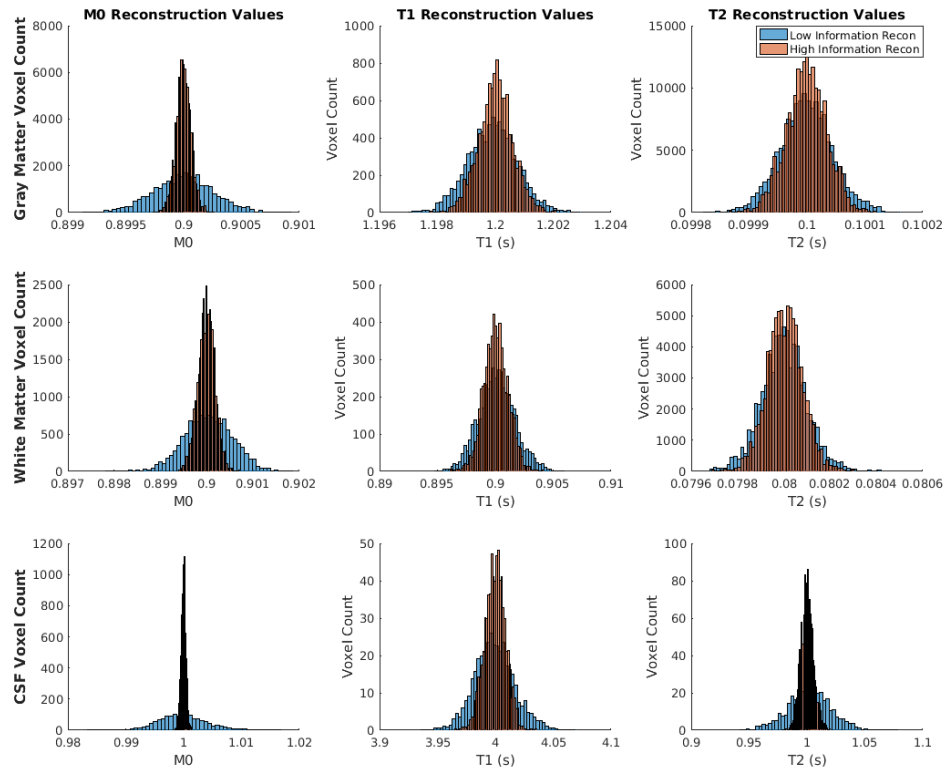


FIGURE 5.8: The red histogram shows the final distribution resulting from a measurement at the optimal location in parameter space. This is the location where mutual information and, consequently, the information gain from a measurement are greatest. The blue histogram shows the final distribution from a measurement at the initial point in parameter space.

parametric maps and the ground truth maps (Figure 5.9, top). A second reconstruction was then simulated using a measurement with the optimized acquisition parameters. The absolute error between these parametric maps and the ground truth was much smaller than that for the first reconstruction (Figure 5.9, bottom). This finding illustrates how the reproducibility of these images is improved by optimal selection of acquisition parameters.

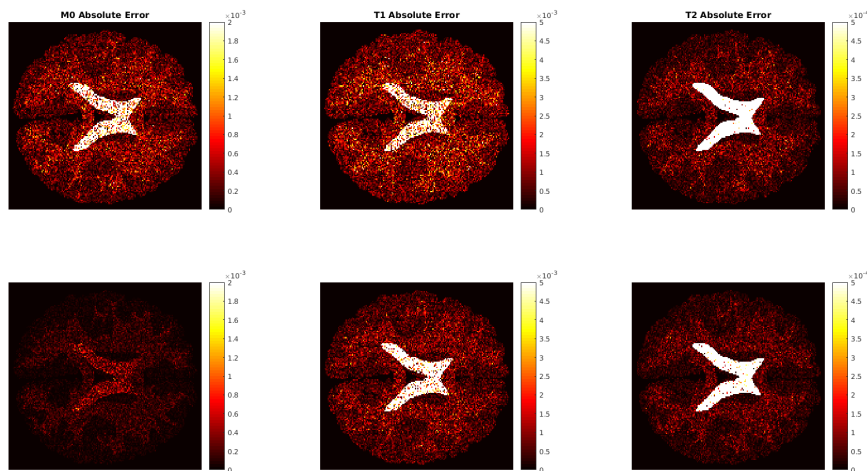


FIGURE 5.9: We simulate a reconstruction using a measurement with the initial, non-optimized acquisition parameters. The absolute error between these parametric maps and the ground truth maps is shown on the top row. We simulate a second reconstruction using a measurement with the optimized acquisition parameters. The absolute error between these parametric maps and the ground truth is on the bottom row.

#### 5.4.2 ISMRM NIST Phantom

Two scans were performed on the ISMRM NIST system phantom with different sets of acquisition parameters. Using the model derived from the synthetic phantom with parameter values listed in Table 5.1, the mutual information of the first scan was calculated to be  $7.29 \times 10^5$ . The mutual information of the second scan was  $7.49 \times 10^5$ . The M0, T1, and T2 parametric maps were reconstructed for the phantom's 14 T1 elements. Figure 5.10 shows the distributions of the M0, T1, and T2 values, reconstructed from the first scan, for the PD, T1, and T2 elements. Omitting T1 and T2 elements 1 and 2 owing to poor reconstruction, nine of 12 T1 elements displayed less variable reconstructions for the higher mutual information scan, and nine of 12 T2 elements displayed less variable reconstructions for the higher mutual information scan. Figures 5.11, 5.12, and 5.13 show a comparison of reconstructed T1 values and 95% confidence intervals between 3D-QALAS and variable flip angle (VFA) T1 mapping in the ISMRM NIST phantom

PD, T1, and T2 elements, respectively. True T1 values and standard deviations are taken from the ISMRM NIST phantom manual.

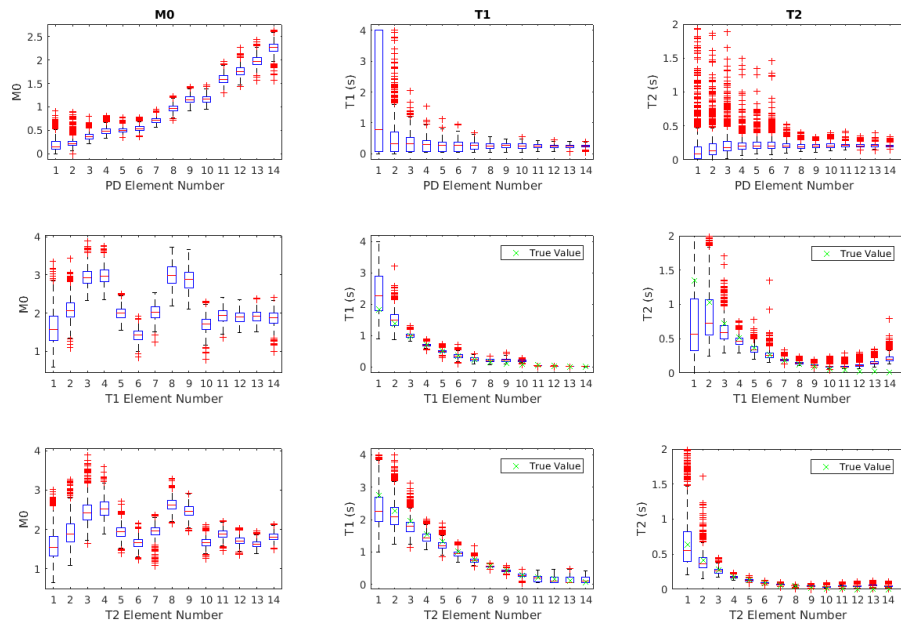


FIGURE 5.10: Distribution of reconstructed M0, T1, and T2 values in the 14 T2 elements of the System Standard Model 130 phantom for the second set of acquisition parameters. The first row shows the distribution of the absolute M0, T1, and T2 values. The second row shows the M0, T1, and T2 values normalized by the element mean. The red bars represents the median M0 value. The blue boxes are the middle two quartiles, and the black lines are the outer two quartiles. A red plus denotes an outlier.

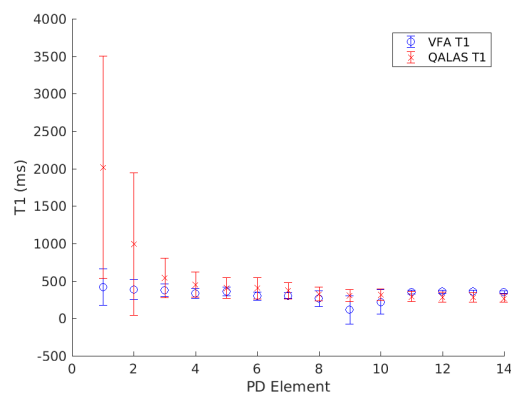


FIGURE 5.11: Comparison of T1 mapping performance between a variable flip angle T1 mapping sequence and 3D-QALAS on the PD elements of the ISMRM NIST phantom. T1 means and standard deviations for these elements are not available in the phantom manual.

Adjusting the T1 and T2 relaxation times of the synthetic phantom to match those

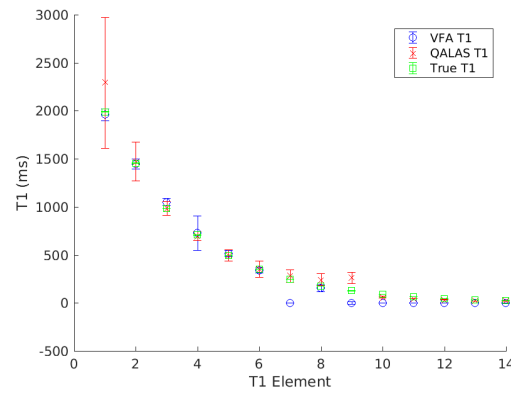


FIGURE 5.12: Comparison of T1 mapping performance between a variable flip angle T1 mapping sequence and 3D-QALAS on the T2 elements of the ISMRM NIST phantom. True T1 values use the T1 means and standard deviations listed for these elements in the phantom manual.

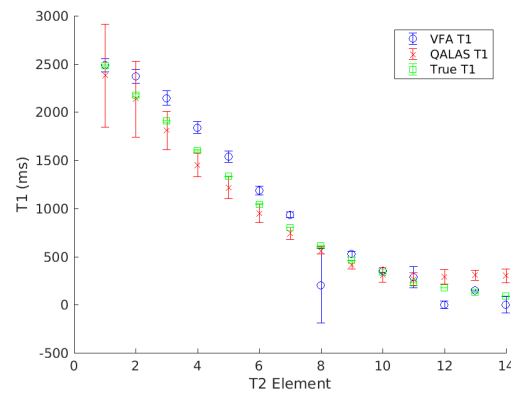


FIGURE 5.13: Comparison of T1 mapping performance between a variable flip angle T1 mapping sequence and 3D-QALAS on the T1 elements of the ISMRM NIST phantom. True T1 values use the T1 means and standard deviations listed for these elements in the phantom manual.

of the various ISMRM NIST system phantom elements allowed the calculation of an independent mutual information value for each phantom element and each set of acquisition parameters. Figure 5.14 shows the reconstructed M0, T1, and T2 standard deviations as a function of the independent mutual information calculated for 12 T1 elements and 12 T2 elements for the two scans. Relative variance was negatively correlated with mutual information, as expected. This result suggests that higher mutual information calculated by this model corresponds to lower standard deviation in the reconstructed image.

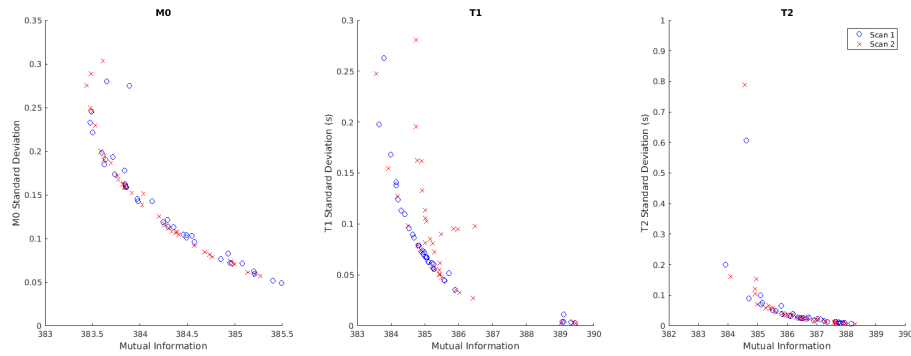


FIGURE 5.14: Reconstructed M0, T1, and T2 standard deviation in each of the 36 analyzed phantom elements for both scans as a function of mutual information. For reconstructed M0, T1, and T2 values, standard deviation is negatively correlated with mutual information across all data sets.

## 5.5 Discussion

The synthetic phantom results show that this mathematical model can minimize QALAS reconstruction uncertainty by optimizing mutual information as a function of acquisition parameters. The optimization is stable in a parameter space of flip angle and four delay times (Figure 5.7). The optimized parameter values differ substantially from the initial values, which are currently used for QALAS acquisitions. The high-information and low-information acquisitions also show substantially different distributions (Figure 5.8). The distributions of parametric map values resulting from a point in acquisition parameter space with greater mutual information are narrower than the distributions resulting from a point with less mutual information. Reduced reconstruction uncertainty corresponds directly to reduced variability (Figure 5.9). This method for quantitatively optimizing reproducibility could have a significant impact on the reliability of quantitative MRI as diagnostic tool.

The ISMRM NIST system phantom measurements demonstrate that the mathematical model tested on the synthetic phantom is applicable to real-world measurements.



The mutual information model can quantify the potential of various combinations of acquisition parameters to reduce QALAS reconstruction uncertainty in phantom measurements. The 3D-QALAS model reconstructed parametric maps from the two acquisitions reasonably accurately (Figure 5.10), with the exception of proton density elements 1 and 2, T1 elements 1 and 2, and T2 elements 1 and 2. These elements were excluded from further analysis. The distribution width indicates how well suited the scan parameters are to reconstructing tissues with the properties of a given element—i.e., optimal scan parameters for an element result in narrower distribution of reconstructed values for that element. The elements of the ISMRM NIST system phantom cover a wide range of M0, T1, and T2 values, including several outside the normal range of human brain tissue. Using literature M0, T1, and T2 values for human brain tissue in the construction of the synthetic phantom may not be representative for ISMRM NIST system phantom elements with properties significantly different from those of human brain tissue. Furthermore, it is expected that different acquisition timings are optimal for different tissue properties, so comparing the performances of two scans across widely differing phantom elements may not be informative. For this reason, mutual information was calculated independently for each element by adjusting the model tissue properties to match the phantom element properties. Mutual information is then a function of the scan parameters and the phantom element properties. Scan parameters that are well suited for reconstructing parametric maps for tissue properties similar to a particular element will then result in greater mutual information. This model showed success when applied to real data (Figure 5.14). The reconstruction performance, measured by standard deviation of the voxel values within the phantom element, was clearly negatively correlated with the mutual information modeled for the element. This result suggests that as long as the model is representative of the subject, mutual information is an accurate predictor

of reconstruction performance.

## Chapter 6

# Conditional Mutual Information

## Generalization and In Vivo

## Model Validation

### 6.1 Introduction

In this chapter, the information theory framework is validated on *in vivo* data from a human brain. It is also generalized to include conditioning on previous measurements. This allows additional flexibility for application to real-world problems. The feasibility of the conditional information theory framework is tested on 3D-QALAS acquisitions conditioned on low-resolution pre-scan acquisitions with independent acquisition parameters. A computational methodology is developed to quantify information content of 3D-QALAS acquisitions relative to a representative synthetic brain model and validated

in phantom measurements. The hypothesis that information-optimized acquisition parameters for a 3D-QALAS signal model result in reduction in reconstructed parametric map variance is tested.

## 6.2 Theory

### 6.2.1 Conditional Mutual Information Relationships

The low-resolution pre-scan measurement  $\mathbf{d}$  is acquired with subsampling parameters  $\boldsymbol{\theta}_1$  and acquisition parameters  $\boldsymbol{\mu}_1$ . The QALAS operator  $\mathcal{Q}_\mu$  is a nonlinear operator representing the solution of the QALAS forward problem with pulse sequence parameters  $\boldsymbol{\mu}$ .

$$\mathbf{d} = S_{\boldsymbol{\theta}_1}(\mathbf{k}) \odot (\mathcal{F}\mathcal{Q}_{\boldsymbol{\mu}_1}(\boldsymbol{\eta}(\mathbf{x})) + \boldsymbol{\nu}(\mathbf{k})) = S_{\boldsymbol{\theta}_1}(\mathbf{k}) \odot (\mathcal{G}_{\boldsymbol{\mu}_1}(\boldsymbol{\eta})) + \boldsymbol{\nu}(\mathbf{k}) \quad (6.1)$$

The optimized scan  $\mathbf{z}$  is acquired with subsampling parameters  $\boldsymbol{\theta}_2$  and acquisition parameters  $\boldsymbol{\mu}_2$ .

$$\mathbf{z} = S_{\boldsymbol{\theta}_2}(\mathbf{k}) \odot (\mathcal{F}\mathcal{Q}_{\boldsymbol{\mu}_2}(\boldsymbol{\eta}(\mathbf{x})) + \boldsymbol{\nu}(\mathbf{k})) = S_{\boldsymbol{\theta}_2}(\mathbf{k}) \odot (\mathcal{G}_{\boldsymbol{\mu}_2}(\boldsymbol{\eta})) + \boldsymbol{\nu}(\mathbf{k}) \quad (6.2)$$

Conditional probability of the low-resolution pre-scan measurement  $p(\mathbf{d}|\boldsymbol{\eta})$  is normally distributed about the signal model evaluation with subsampling parameters  $\boldsymbol{\theta}_1$ , acquisition parameters  $\boldsymbol{\mu}_1$ , and measurement variance  $\Sigma_{\nu_1}$ .

$$p(\mathbf{d}|\boldsymbol{\eta}) = N(\mathbf{d} | S_{\boldsymbol{\theta}_1}(\mathbf{k}) \odot \mathcal{G}_{\boldsymbol{\mu}_1}(\boldsymbol{\eta}(\mathbf{x})), \Sigma_{\nu_1}) \quad (6.3)$$

Similarly, conditional probability of the high-resolution full scan  $p(\mathbf{z}|\boldsymbol{\eta})$  is normally

distributed about the signal model evaluation with subsampling parameters  $\boldsymbol{\theta}_2$ , acquisition parameters  $\boldsymbol{\mu}_2$ , and measurement variance  $\Sigma_{\nu_2}$ .

$$p(\mathbf{z}|\boldsymbol{\eta}) = N\left(\mathbf{z} \mid S_{\boldsymbol{\theta}_2}(\mathbf{k}) \odot \mathcal{G}_{\boldsymbol{\mu}_2}(\boldsymbol{\eta}(\mathbf{x})), \Sigma_{\nu_2}\right) \quad (6.4)$$

Because measurements  $\mathbf{z}$  and  $\mathbf{d}$  are independent, the joint conditional probability distribution  $p(\mathbf{z}, \mathbf{d}|\boldsymbol{\eta})$  is the multivariate normal distribution

$$p(\mathbf{z}, \mathbf{d}|\boldsymbol{\eta}) = N\left(\begin{bmatrix} \mathbf{z} \\ \mathbf{d} \end{bmatrix} \mid \begin{bmatrix} S_{\boldsymbol{\theta}_2}(\mathbf{k}) \odot \mathcal{G}_{\boldsymbol{\mu}_2}(\boldsymbol{\eta}(\mathbf{x})) \\ S_{\boldsymbol{\theta}_1}(\mathbf{k}) \odot \mathcal{G}_{\boldsymbol{\mu}_1}(\boldsymbol{\eta}(\mathbf{x})) \end{bmatrix}, \begin{bmatrix} \Sigma_{\nu_2} & 0 \\ 0 & \Sigma_{\nu_1} \end{bmatrix}\right). \quad (6.5)$$

Recall the generalized mutual information relationships for a new measurement,  $\mathbf{z}$ , given  $N$  previous measurements,  $\mathbf{d}_1, \dots, \mathbf{d}_N$ :

$$I(\boldsymbol{\eta}; \mathbf{z}|\mathbf{d}) = H(\boldsymbol{\eta}|\mathbf{d}) - H(\boldsymbol{\eta}|\mathbf{z}, \mathbf{d}) \quad (6.6)$$

$$I(\boldsymbol{\eta}; \mathbf{z}|\mathbf{d}) = H(\boldsymbol{\eta}) - I(\boldsymbol{\eta}; \mathbf{d}_1) - \sum_{i=2}^N I(\boldsymbol{\eta}; \mathbf{d}_i|\mathbf{d}_{i-1}, \dots, \mathbf{d}_1) - H(\boldsymbol{\eta}; \mathbf{z}, \mathbf{d}) \quad (6.7)$$

$$I(\boldsymbol{\eta}; \mathbf{z}|\mathbf{d}) = I(\boldsymbol{\eta}; \mathbf{z}, \mathbf{d}) - I(\boldsymbol{\eta}; \mathbf{d}_1) - \sum_{i=2}^N I(\boldsymbol{\eta}; \mathbf{d}_i|\mathbf{d}_{i-1}, \dots, \mathbf{d}_1) \quad (6.8)$$

$$I(\boldsymbol{\eta}; \mathbf{z}|\mathbf{d}) = I(\boldsymbol{\eta}; \mathbf{z}, \mathbf{d}) - I(\boldsymbol{\eta}; \mathbf{d}) \quad (6.9)$$

If the measurement parameters are explicitly written out, conditional mutual information takes the following form:

$$\begin{aligned} & I(\boldsymbol{\eta}; \mathbf{z}(\boldsymbol{\mu}_{N+1}, \boldsymbol{\theta}_{N+1})|\mathbf{d}_1(\boldsymbol{\mu}_1, \boldsymbol{\theta}_1), \dots, \mathbf{d}_N(\boldsymbol{\mu}_N, \boldsymbol{\theta}_N)) \\ &= I(\boldsymbol{\eta}; \mathbf{z}(\boldsymbol{\mu}_{N+1}, \boldsymbol{\theta}_{N+1}), \mathbf{d}_1(\boldsymbol{\mu}_1, \boldsymbol{\theta}_1), \dots, \mathbf{d}_N(\boldsymbol{\mu}_N, \boldsymbol{\theta}_N)) - I(\boldsymbol{\eta}; \mathbf{d}_1(\boldsymbol{\mu}_1, \boldsymbol{\theta}_1)) \quad (6.10) \\ & \quad - \sum_{i=2}^N I(\boldsymbol{\eta}; \mathbf{d}_i(\boldsymbol{\mu}_i, \boldsymbol{\theta}_i)|\mathbf{d}_{i-1}(\boldsymbol{\mu}_{i-1}, \boldsymbol{\theta}_{i-1}), \dots, \mathbf{d}_1(\boldsymbol{\mu}_1, \boldsymbol{\theta}_1)). \end{aligned}$$

For the case where  $\mathbf{d}$  is one low-resolution pre-scan with sub-optimal parameters, conditional mutual information simplifies to

$$\begin{aligned}
I(\boldsymbol{\eta}; \mathbf{z}(\boldsymbol{\mu}_2, \boldsymbol{\theta}_2) | \mathbf{d}_1(\boldsymbol{\mu}_1, \boldsymbol{\theta}_1)) \\
&= I(\boldsymbol{\eta}; \mathbf{z}(\boldsymbol{\mu}_2, \boldsymbol{\theta}_2), \mathbf{d}_1(\boldsymbol{\mu}_1, \boldsymbol{\theta}_1)) - I(\boldsymbol{\eta}; \mathbf{d}_1(\boldsymbol{\mu}_1, \boldsymbol{\theta}_1)) \\
&= H(\mathbf{z}, \mathbf{d}) - H(\mathbf{z}, \mathbf{d} | \boldsymbol{\eta}) - (H(\mathbf{d}) - H(\mathbf{d} | \boldsymbol{\eta})).
\end{aligned} \tag{6.11}$$

Because the conditional probability functions  $p(\mathbf{d} | \boldsymbol{\eta})$  and  $p(\mathbf{z}, \mathbf{d} | \boldsymbol{\eta})$  are defined to be normally distributed, their respective entropies,  $H(\mathbf{d} | \boldsymbol{\eta})$  and  $H(\mathbf{z}, \mathbf{d} | \boldsymbol{\eta})$ , are known analytically.

$$H(\mathbf{d} | \boldsymbol{\eta}) = \frac{1}{2} \ln \left( (2\pi e)^2 \sigma_\nu^{2N_{\text{acq},d}} \right) \tag{6.12}$$

$$H(\mathbf{z}, \mathbf{d} | \boldsymbol{\eta}) = \frac{1}{2} \ln \left( (2\pi e)^2 \sigma_\nu^{2(N_{\text{acq},d} + N_{\text{acq},z})} \right) \tag{6.13}$$

### 6.2.2 Mutual Information for Jointly Gaussian Measurements

For measurements which are jointly Gaussian, conditional mutual information can be approximated by Gauss-Hermite quadrature. This derivation is given in Section C.3.1 in Appendix C, but the general forms of  $H(\mathbf{d})$ ,  $H(\mathbf{z}, \mathbf{d})$ , and  $I(\boldsymbol{\eta}; \mathbf{z} | \mathbf{d})$  used in the computation of conditional mutual information are shown below.

Conditional mutual information  $I(\boldsymbol{\eta}; \mathbf{z}|\mathbf{d})$  requires approximation of entropies  $H(\mathbf{d})$  and  $H(\mathbf{z}, \mathbf{d})$ . The approximation of  $H(\mathbf{d})$  can be computed by

$$\begin{aligned}
H(\mathbf{d}) &\approx \pi^{-N/2-P} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_{1,r}=1}^{K_{1,r}} \omega_{k_{1,r}} \sum_{k_{1,i}=1}^{K_{1,i}} \omega_{k_{1,i}} \cdots \sum_{k_{P,r}=1}^{K_{P,r}} \omega_{k_{P,r}} \sum_{k_{P,i}=1}^{K_{P,i}} \omega_{k_{P,i}} \\
&\cdot \ln \left[ \pi^{-N/2} \sum_{s_1=1}^{S_1} \omega_{s_1} \cdots \sum_{s_N=1}^{S_N} \omega_{s_N} \right. \\
&\quad \cdot \prod_{p=1}^P p(\sqrt{2}\Sigma_\nu \mathbf{x}_{k_{p,r}} + S_{\theta_1} \odot \mathcal{FQ}_{\mu_1}(\sqrt{2}\Sigma_\eta \mathbf{x}_q + \mathbf{m}_\eta) | \sqrt{2}\Sigma_\eta \mathbf{x}_s + \mathbf{m}_\eta) \\
&\quad \left. \cdot p(\sqrt{2}\Sigma_\nu \mathbf{x}_{k_{p,i}} + S_{\theta_1} \odot \mathcal{FQ}_{\mu_1}(\sqrt{2}\Sigma_\eta \mathbf{x}_q + \mathbf{m}_\eta) | \sqrt{2}\Sigma_\eta \mathbf{x}_s + \mathbf{m}_\eta) \right].
\end{aligned} \tag{6.14}$$

The approximation of  $H(\mathbf{z}, \mathbf{d})$  can be computed by

$$\begin{aligned}
H(\mathbf{z}, \mathbf{d}) &= \int \int p(\mathbf{z}, \mathbf{d}) \ln(p(\mathbf{z}, \mathbf{d})) dz d\mathbf{d} \\
&\approx \pi^{-N/2-P} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \cdot \sum_{j_{1,r}=1}^{J_{1,r}} \omega_{j_{1,r}} \sum_{j_{1,i}=1}^{J_{1,i}} \omega_{j_{1,i}} \cdots \sum_{j_{P,r}=1}^{J_{P,r}} \omega_{j_{P,r}} \sum_{j_{P,i}=1}^{J_{P,i}} \omega_{j_{P,i}} \\
&\cdot \sum_{k_{1,r}=1}^{K_{1,r}} \omega_{k_{1,r}} \sum_{k_{1,i}=1}^{K_{1,i}} \omega_{k_{1,i}} \cdots \sum_{k_{P,r}=1}^{K_{P,r}} \omega_{k_{P,r}} \sum_{k_{P,i}=1}^{K_{P,i}} \omega_{k_{P,i}} \cdot \ln \left\{ \pi^{-N/2} \sum_{s_1=1}^{S_1} \omega_{s_1} \cdots \sum_{s_N=1}^{S_N} \omega_{s_N} \right. \\
&\cdot \prod_{p=1}^P p \left( \sqrt{2}\Sigma_\nu \begin{bmatrix} \mathbf{x}_{j_{p,r}} \\ \mathbf{x}_{k_{p,r}} \end{bmatrix} + \begin{bmatrix} S_{\theta_2} \odot \mathcal{FQ}_{\mu_2}(\sqrt{2}\Sigma_\eta \mathbf{x}_q + \mathbf{m}_\eta) \\ S_{\theta_1} \odot \mathcal{FQ}_{\mu_1}(\sqrt{2}\Sigma_\eta \mathbf{x}_q + \mathbf{m}_\eta) \end{bmatrix} \middle| \sqrt{2}\Sigma_\eta \mathbf{x}_s + \mathbf{m}_\eta \right) \\
&\cdot \left. p \left( \sqrt{2}\Sigma_\nu \begin{bmatrix} \mathbf{x}_{j_{p,i}} \\ \mathbf{x}_{k_{p,i}} \end{bmatrix} + \begin{bmatrix} S_{\theta_2} \odot \mathcal{FQ}_{\mu_2}(\sqrt{2}\Sigma_\eta \mathbf{x}_q + \mathbf{m}_\eta) \\ S_{\theta_1} \odot \mathcal{FQ}_{\mu_1}(\sqrt{2}\Sigma_\eta \mathbf{x}_q + \mathbf{m}_\eta) \end{bmatrix} \middle| \sqrt{2}\Sigma_\eta \mathbf{x}_s + \mathbf{m}_\eta \right) \right\}.
\end{aligned} \tag{6.15}$$

After computing  $H(\mathbf{d})$  and  $H(\mathbf{z}, \mathbf{d})$ , conditional mutual information is straightforward to determine:

$$\begin{aligned}
I(\boldsymbol{\eta}; \mathbf{z}|\mathbf{d}) &= H(\boldsymbol{\eta}|\mathbf{d}) - H(\boldsymbol{\eta}|\mathbf{z}, \mathbf{d}) = H(\mathbf{z}, \mathbf{d}) - H(\mathbf{z}, \mathbf{d}|\boldsymbol{\eta}) - (H(\mathbf{d}) - H(\mathbf{d}|\boldsymbol{\eta})) \\
&= H(\mathbf{z}, \mathbf{d}) - H(\mathbf{d}) + \frac{1}{2} \ln \left( (2\pi e)^2 \sigma_\nu^{2N_{\text{acq},d}} \right) - \frac{1}{2} \ln \left( (2\pi e)^2 \sigma_\nu^{2(N_{\text{acq},d} + N_{\text{acq},z})} \right) \\
&= H(\mathbf{z}, \mathbf{d}) - H(\mathbf{d}) + \frac{1}{2} \ln(\sigma_\nu^{-2N_{\text{acq},z}}).
\end{aligned} \tag{6.16}$$

### 6.2.3 Mutual Information for Independent Subsampling Masks

When optimizing  $k$ -space subsampling approaches, the inclusion of new subsampling locations conditional upon existing subsampled measurements can be represented by independent subsampling masks in this model. Computation of conditional mutual information is simplified significantly for this special case.

For FFT reconstruction, the FFT operator is not a function of the subsampling mask like the BART reconstruction operator. Let  $g_\mu$  be a fully-sampled, noiseless set of measurements resulting from  $\mathcal{FQ}_\mu(\boldsymbol{\eta}(\mathbf{x}))$ .

$$\mathbf{z}(\boldsymbol{\mu}, \boldsymbol{\theta}, \mathbf{k}) = S_{\boldsymbol{\theta}} \odot (\mathcal{FQ}_\mu(\boldsymbol{\eta}(\mathbf{x})) + \nu(\mathbf{k})) = S(\boldsymbol{\theta}, \mathbf{k}) \odot (g_\mu + \nu(\mathbf{k})) \tag{6.17}$$

Subsequent measurements,  $\mathbf{z}_1$  and  $\mathbf{z}_2$ , with separate acquisitions,  $\mathcal{G}_{\mu_1}$  and  $\mathcal{G}_{\mu_2}$  are defined as follows:

$$\begin{aligned}
\mathbf{z}_1 &= S_{\boldsymbol{\theta}_1} \odot (\mathcal{FQ}_{\mu_1}(\boldsymbol{\eta}(\mathbf{x})) + \nu_1(\mathbf{k})) = S_{\boldsymbol{\theta}_1} \odot (\mathcal{G}_{\mu_1}(\boldsymbol{\eta}) + \nu_1(\mathbf{k})) \\
\mathbf{z}_2 &= S_{\boldsymbol{\theta}_2} \odot (\mathcal{FQ}_{\mu_2}(\boldsymbol{\eta}(\mathbf{x})) + \nu_2(\mathbf{k})) = S_{\boldsymbol{\theta}_2} \odot (\mathcal{G}_{\mu_2}(\boldsymbol{\eta}) + \nu_2(\mathbf{k})).
\end{aligned} \tag{6.18}$$



Now, subsequent subsampling realizations,  $\xi_1$  and  $\xi_2$ , of the same fully-sampled measurement,  $g_\mu$ , are defined as follows:

$$\begin{aligned}\xi_1|g_\mu &= S_{\theta_1} \odot (\mathcal{F}Q_\mu(\boldsymbol{\eta}(\mathbf{x})) + \nu(\mathbf{k})) = S_{\theta_1} \odot (g_\mu + \nu(\mathbf{k})) \\ \xi_2|g_\mu &= S_{\theta_2} \odot (\mathcal{F}Q_\mu(\boldsymbol{\eta}(\mathbf{x})) + \nu(\mathbf{k})) = S_{\theta_2} \odot (g_\mu + \nu(\mathbf{k}))\end{aligned}\tag{6.19}$$

The subsampled  $\xi_1$  and  $\xi_2$  are independent for the same measurement  $g_\mu$ . However,  $\mathbf{z}_1$  and  $\mathbf{z}_2$  are not independent.

The joint mutual information of these subsampling masks is equal to the sum of the independent mutual information terms.

$$I(\boldsymbol{\eta}; \xi_1, \xi_2) = I(\boldsymbol{\eta}; \xi_1) + I(\boldsymbol{\eta}; \xi_2)\tag{6.20}$$

The derivation of this relationship is shown in Section C.3.2 in Appendix C. It is straightforward to show through induction that

$$I(\boldsymbol{\eta}; \boldsymbol{\xi}) = \sum_{i=1}^N I(\boldsymbol{\eta}; \xi_i).\tag{6.21}$$

An expected result of the independence of subsampling masks is that conditional mutual information reduces to mutual information:

$$\begin{aligned}I(\boldsymbol{\eta}; \xi_2|\xi_1) &= I(\boldsymbol{\eta}; \xi_2, \xi_1) - I(\boldsymbol{\eta}; \xi_1) \\ &= I(\boldsymbol{\eta}; \xi_2) + I(\boldsymbol{\eta}; \xi_1) - I(\boldsymbol{\eta}; \xi_1) \\ &= I(\boldsymbol{\eta}; \xi_2).\end{aligned}\tag{6.22}$$

Thus, the general result for conditional mutual information of independent subsampling masks is

$$I(\boldsymbol{\eta}; \boldsymbol{\xi}_M | \boldsymbol{\xi}_N) = \sum_{i \in M} I(\boldsymbol{\eta}; \boldsymbol{\xi}_i) \quad \text{for } M \cap N = \emptyset. \quad (6.23)$$

To compute conditional mutual information from Equation 6.23, the difference of entropies is used.

$$I(\boldsymbol{\eta}; \boldsymbol{\xi}_M | \boldsymbol{\xi}_N) = \sum_{i \in M} I(\boldsymbol{\eta}; \boldsymbol{\xi}_i) = \sum_{i \in M} (H(\boldsymbol{\xi}_i) - H(\boldsymbol{\xi}_i | \boldsymbol{\eta})) \quad (6.24)$$

As previously, the probability density functions  $p(\boldsymbol{\xi}_1 | \boldsymbol{\eta})$ ,  $p(\boldsymbol{\xi}_2 | \boldsymbol{\eta})$ , and  $p(\boldsymbol{\eta})$  are assumed to be normal.

$$\begin{aligned} p(\boldsymbol{\xi}_1 | \boldsymbol{\eta}) &= N(\boldsymbol{\xi}_1 | S_{\boldsymbol{\theta}_1}(\mathbf{k}) \odot g_{\boldsymbol{\mu}}(\mathbf{k}), \Sigma_{\nu}) \\ p(\boldsymbol{\xi}_2 | \boldsymbol{\eta}) &= N(\boldsymbol{\xi}_2 | S_{\boldsymbol{\theta}_2}(\mathbf{k}) \odot g_{\boldsymbol{\mu}}(\mathbf{k}), \Sigma_{\nu}) \\ p(\boldsymbol{\eta}) &= N(\boldsymbol{\eta} | \mathbf{m}_{\boldsymbol{\eta}}, \Sigma_{\boldsymbol{\eta}}) \end{aligned} \quad (6.25)$$

The conditional entropy term is again known from the problem statement.

$$H(\boldsymbol{\xi}_i | \boldsymbol{\eta}) = \frac{1}{2} \ln \left( (2\pi e)^2 \sigma_{\nu}^{2N_{acq,i}} \right) \quad (6.26)$$

$H(\xi_i)$  can be approximated via Gauss-Hermite quadrature.

$$\begin{aligned}
H(\xi) \approx & \pi^{-N/2-P} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_{1,r}=1}^{K_{1,r}} \omega_{k_{1,r}} \sum_{k_{1,i}=1}^{K_{1,i}} \omega_{k_{1,i}} \cdots \sum_{k_{P,r}=1}^{K_{P,r}} \omega_{k_{P,r}} \sum_{k_{P,i}=1}^{K_{P,i}} \omega_{k_{P,i}} \\
& \cdot \ln \left[ \pi^{-N/2} \sum_{s_1=1}^{S_1} \omega_{s_1} \cdots \sum_{s_N=1}^{S_N} \omega_{s_N} \right. \\
& \quad \cdot \prod_{p=1}^P p(\sqrt{2}\Sigma_\nu \mathbf{x}_{k_{p,r}} + S \odot \mathcal{Q}(\sqrt{2}\Sigma_\eta \mathbf{x}_q + \mathbf{m}_\eta) | \sqrt{2}\Sigma_\eta \mathbf{x}_s + \mathbf{m}_\eta) \\
& \quad \left. \cdot p(\sqrt{2}\Sigma_\nu \mathbf{x}_{k_{p,i}} + S \odot \mathcal{Q}(\sqrt{2}\Sigma_\eta \mathbf{x}_q + \mathbf{m}_\eta) | \sqrt{2}\Sigma_\eta \mathbf{x}_s + \mathbf{m}_\eta) \right]
\end{aligned} \tag{6.27}$$

The derivation of this relationship is shown in Section C.3.3 in Appendix C. The complete form of conditional mutual information for independent subsampling masks is

$$I(\boldsymbol{\eta}; \xi_M | \xi_N) = \sum_{i \in M} I(\boldsymbol{\eta}; \xi_i) = \sum_{i \in M} \left( H(\xi_i) - \frac{1}{2} \ln \left( (2\pi e)^2 \sigma_\nu^{2N_{\text{acq},i}} \right) \right). \tag{6.28}$$

### 6.3 Methods

Proof-of-concept experiments were again performed on both synthetic phantom data and the ISMRM NIST qMRI Multiparametric Imaging Standard phantom (QalibreMD, Boulder, CO).[195] The synthetic phantom data were used to demonstrate the range of optimized acquisition parameter values compared to the static empirical parameter selection currently implemented clinically. Multiple measurements of the ISMRM NIST system phantom at high and low resolutions and with different combinations of the selected acquisition parameters were used to validate the conditional mutual information model. The correlation between mutual information and reconstruction variance was measured.

### 6.3.1 Synthetic Model

To demonstrate potential clinical usefulness of real-time parameter tuning, measurements are simulated on a cohort of ten synthetic brain phantoms (Figure 6.1). The tissue label boundaries are eroded and dilated to create anatomically distinct synthetic patients. Additionally, the patient-specific biological uncertainty was randomly assigned by sampling the population distribution of tissue properties (Table 5.1).

Conditional mutual information was calculated conditioned on synthesized low-resolution measurements (using tissue labels in Figure 6.2) made with all delay times set to 500 ms, which are the current clinically implemented values of these parameters. The optimized values of these delay times are then compared to the current clinical value.

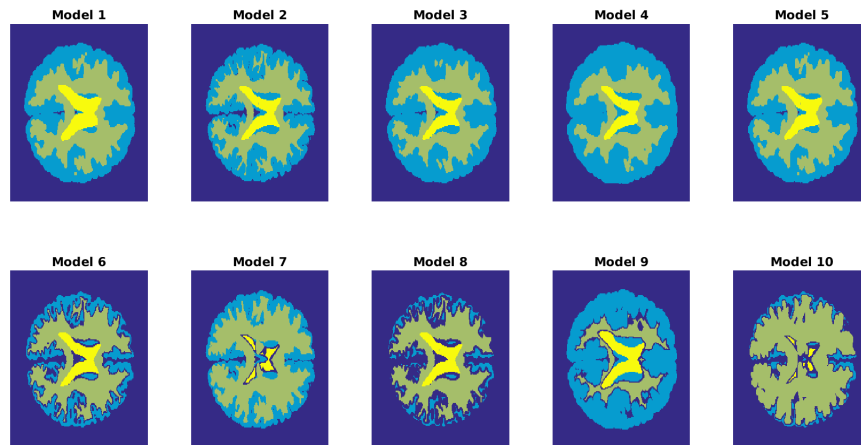


FIGURE 6.1: High-resolution tissue label maps of ten synthetic brain models.

### 6.3.2 Phantom Acquisitions

To validate the mutual information model, we applied it to actual measurements of the ISMRM NIST system phantom (Figure 5.5). Multiple measurements were acquired at high (Figure 6.3a) and low resolutions (Figure 6.3b) and with different combinations of the selected acquisition parameters using a 3.0T MR 750W MRI scanner (GE Healthcare,

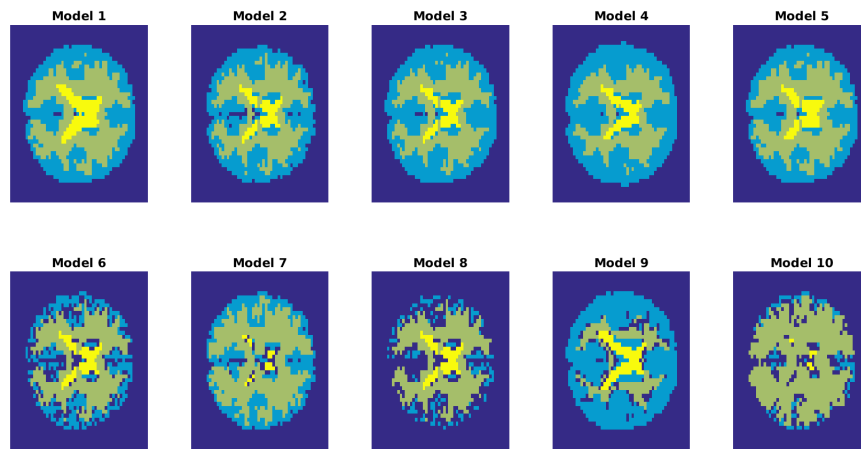


FIGURE 6.2: Low-resolution tissue label maps of ten synthetic brain models.

Waukesha, WI). The acquisition parameters are shown in Table 6.1, and the physical properties of the 42 ISMRM NIST phantom elements are shown in Tables 6.2, 6.3, and 6.4. Reconstruction was performed using geOrchestra and the Berkeley Advanced Reconstruction Toolbox (BART, The University of California, Berkeley, Berkeley, CA), with L2 regularization (Figure 5.6).[196] Correlation between mutual information and reconstruction variance was measured.

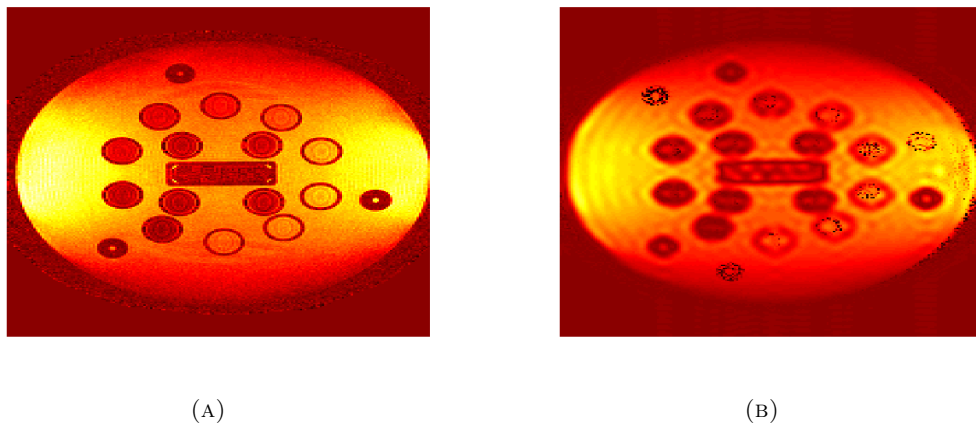


FIGURE 6.3: (A) Example longitudinal magnetization of ISMRM NIST phantom from high-resolution acquisition. (B) Longitudinal magnetization of same view from ISMRM NIST phantom from low-resolution acquisition.

The ISMRM NIST phantom contains 42 elements with differing material properties. This provides 42 data points of well-characterized physical properties. Conditional

Parameter	Scan 1	Scan 2
<b>Flip angle</b> ( $^{\circ}$ )	4	4
<b>TR</b> (ms)	6.6	6.6
<b>TI</b> (ms)	100	100
<b>TE<sub>T2prep</sub></b> (ms)	100	100
<b>T<sub>acq</sub></b> (ms)	674.8	674.8
<b>N<sub>acq</sub></b>	5	5
<b>TD1<sup>a</sup></b> (ms)	230.6	500
<b>TD2<sup>a</sup></b> (ms)	325.2	142.6
<b>TD3<sup>a</sup></b> (ms)	325.2	142.6
<b>TD4<sup>a</sup></b> (ms)	325.2	142.6
<b>TD5<sup>a</sup></b> (ms)	221.6	500.0

TABLE 6.1: Acquisition parameter values for two scans performed on ISMRM NIST system phantom. <sup>a</sup>TD1 is the time between the T1 sensitizing pulse and the first acquisition, TD2 is the time between the end of the first acquisition and the start of the second acquisition, and so on.

Contrast ID	Concentration (% Water)	H2O mass (g)	D2O mass (g)
<b>PD-1</b>	5	0.746	15.758
<b>PD-2</b>	10	1.491	14.932
<b>PD-3</b>	15	2.238	14.106
<b>PD-4</b>	20	2.238	13.279
<b>PD-5</b>	25	3.731	12.451
<b>PD-6</b>	30	4.478	11.624
<b>PD-7</b>	35	5.226	10.796
<b>PD-8</b>	40	5.974	9.968
<b>PD-9</b>	50	7.470	8.310
<b>PD-10</b>	60	8.968	6.651
<b>PD-11</b>	70	10.468	4.990
<b>PD-12</b>	80	11.968	3.328
<b>PD-13</b>	90	13.470	1.668
<b>PD-14</b>	100	14.973	0.000

TABLE 6.2: ISMRM NIST phantom proton density sphere properties.

mutual information was calculated for the physical properties of each phantom element between a low-resolution acquisition at one set of acquisition parameters and a theoretical high-resolution acquisition at the other set of acquisition parameters. In other words, one data point would result from the calculation of conditional mutual information using the physical properties of element T1-4, conditional on the low-resolution acquisition with Scan 1 parameters, and assuming a theoretical high-resolution acquisition with Scan 2 parameters. Then, the difference in standard deviations of reconstructed M0, T1, and

Sample Name	T1 (ms)	T1 Standard Deviation (ms)	T2 (ms)	T2 Standard Deviation (ms)
T1-1	1989	1.0	1465	1.0
T1-2	1454	2.5	1076	1.8
T1-3	984.1	0.33	717.9	1.12
T1-4	706	1.5	510.1	1.36
T1-5	496.7	0.41	359.6	0.22
T1-6	351.5	0.91	255.5	0.07
T1-7	247.13	0.086	180.8	0.04
T1-8	175.3	0.11	127.3	0.14
T1-9	125.9	0.33	90.3	0.14
T1-10	89.0	0.17	64.3	0.05
T1-11	62.7	0.13	45.7	0.12
T1-12	44.53	0.090	31.86	0.02
T1-13	30.84	0.016	22.38	0.02
T1-14	21.719	0.0054	15.83	0.03

TABLE 6.3: ISMRM NIST phantom T1 contrast sphere properties at 3.0 T.

Sample Name	T1 (ms)	T1 Standard Deviation (ms)	T2 (ms)	T2 Standard Deviation (ms)
T2-1	2480	10.8	581.3	0.39
T2-2	2173	14.7	403.5	0.55
T2-3	1907	10.3	278.1	0.28
T2-4	1604	7.2	190.94	0.011
T2-5	1332	0.8	133.27	0.073
T2-6	1044	3.2	96.89	0.049
T2-7	801.7	1.70	64.07	0.034
T2-8	608.6	1.03	46.42	0.014
T2-9	458.4	0.33	31.97	0.083
T2-10	336.5	0.18	22.56	0.012
T2-11	224.2	0.09	15.813	0.0061
T2-12	176.6	0.09	11.237	0.0057
T2-13	126.9	0.03	7.911	0.0037
T2-14	90.9	0.05	5.592	0.0055

TABLE 6.4: ISMRM NIST phantom T2 contrast sphere properties at 3.0 T.

T2 values in element T1-4 between low-resolution Scan 1 and high-resolution Scan 2 can be used to evaluate the success of the performance prediction made by conditional mutual information. By using all phantom elements, a correlation between the standard deviation differences and conditional mutual information is used to evaluate model prediction performance.

In order to more strongly corroborate the trend seen in image quality prediction

with the available data, predictions are also made between NIST phantom elements for constant scan parameters. This is equivalent to changing the image subject properties between scans instead of the acquisition parameters. While this is obviously not a realistic scenario, it provides further validation of the predictive capabilities of the conditional mutual information model.

### 6.3.3 In Vivo Acquisitions

*In vivo* 3D-QALAS measurements were acquired on one volunteer patient. Two scans were performed, one with default clinical acquisition parameters and one with an optimized set of acquisition parameters. Optimized parameters were determined similarly to the optimization performed in Section 5.4. The five delay times were optimized as shown in Figure 6.4. Because it was not feasible to implement the exact delay times selected via optimization at the time of acquisition, a set of delay times were chosen that were acceptably close in order to demonstrate the differences in reconstruction. These three sets of acquisition parameters are displayed in Table 6.5.

Parameter	Default Acquisition	Optimized Acquisition	Theoretical Optimum
Flip angle ( $^{\circ}$ )	4	4	4
TR (ms)	6.684	6.684	6.684
TI (ms)	100	100	100
TE <sub>T2prep</sub> (ms)	100	100	100
T <sub>acq</sub> (ms)	871	871	871
N <sub>acq</sub>	5	5	5
TD1 <sup>a</sup> (ms)	23.4	500.0	625.5
TD2 <sup>a</sup> (ms)	117.4	150.0	71.3
TD3 <sup>a</sup> (ms)	117.4	0.5	0.0
TD4 <sup>a</sup> (ms)	117.4	0.5	198.0
TD5 <sup>a</sup> (ms)	0.0	500.0	1608.6

TABLE 6.5: Acquisition parameter values for *in vivo* 3D-QALAS acquisitions. <sup>a</sup>TD1 is the time between the T1 sensitizing pulse and the first acquisition, TD2 is the time between the end of the first acquisition and the start of the second acquisition, and so on.



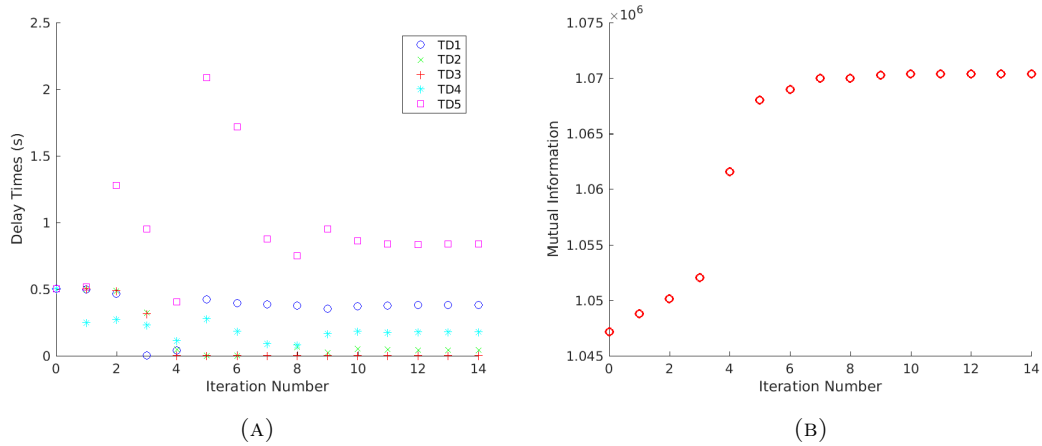


FIGURE 6.4: (A) Optimization history of acquisition parameters. Five delay times are shown over the course of 14 optimization iterations. (B) Optimization history of mutual information over the course of 14 optimization iterations. Mutual information is the objective function to be maximized during optimization.

Performance of these acquisition parameters is prospectively evaluated by sampling the posterior distributions resulting from each set of parameters using ellipsoidal nested sampling.[193, 194] Standard deviation of the posterior distributions was used to predict relative performance between the scans. Mutual information was calculated using Equation 4.30 for scans acquired with current clinical parameters, optimized acquisition parameters, and theoretical optimum parameters.

Two scans were performed with acquisition parameters listed in Table 6.1 using a 3.0T MR 750W MRI scanner (GE Healthcare, Waukesha, WI). Reconstruction was performed using geOrchestra, the Berkeley Advanced Reconstruction Toolbox (BART, The University of California, Berkeley, Berkeley, CA), and a fitting algorithm for the 3D-QALAS signal model developed for use in this work.

ROIs were drawn manually in gray matter and white matter in the frontal, parietal, temporal, and occipital lobes, as well as cerebrospinal fluid in the ventricles. ROI selection was performed manually in order to exclude interfaces between tissue types. This minimizes the effects of partial voluming, for which the model did not account. ROIs were analyzed in the four lobes of the cerebral cortex in both hemispheres in order

to minimize sensitivity to ROI selection. Summary statistics were calculated for the parameter value distributions in gray and white matter in each of the four lobes, as well as CSF. The change in standard deviation between the reconstructions of the default and optimized scans was evaluated. Two-sample F-tests for equal variances were used to determine the significance of the change in standard deviation.

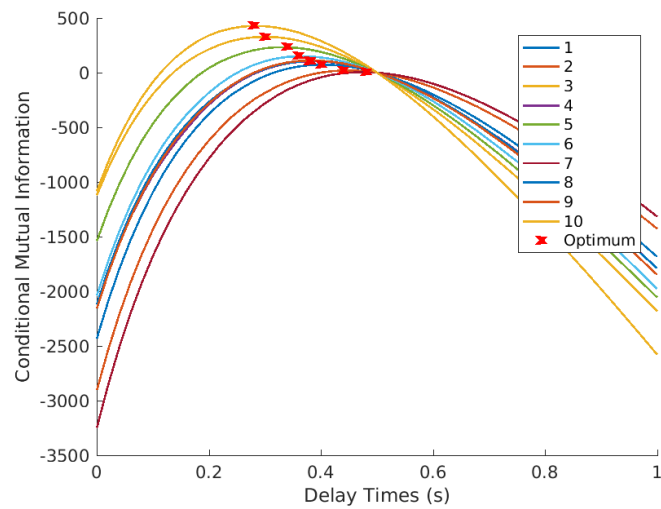
## 6.4 Results

### 6.4.1 Synthetic Model

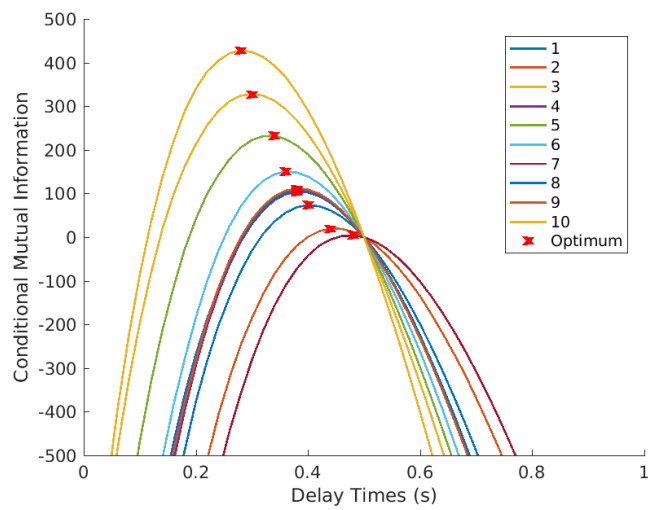
Figure 6.5 shows the wide range of optimal parameter values that arise from realistic perturbations of the synthetic brain model. Optimal delay time varied from 280 ms to 480 ms. The currently implemented clinical value is 500 ms. Thus, patient-specific anatomical and tissue property perturbations resulted in an optimized delay time up to 44% less than the clinical value. The range of optimized parameter values indicates that this method may result in significant information gain in clinical settings.

### 6.4.2 Phantom Acquisitions

Figure 6.6 shows the change in M0 standard deviation in NIST phantom elements between a low-resolution pre-scan at one set of acquisition parameters and a high-resolution scan at a second set of acquisition parameters as a function of conditional mutual information. It can be seen that the change in M0 standard deviation is negatively correlated with conditional mutual information. For conditioning on the low-resolution acquisition at Scan 1 parameters, the adjusted  $r^2$  from a robust linear regression is 0.3436. The root-mean-squared error is 0.0108. For conditioning on the low-resolution acquisition at



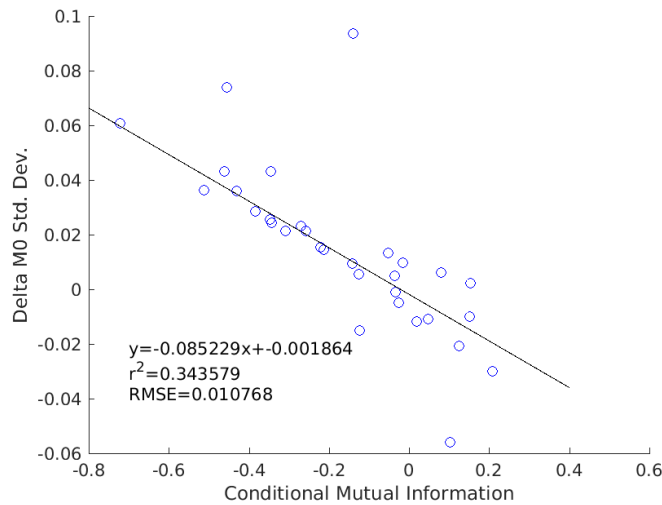
(A)



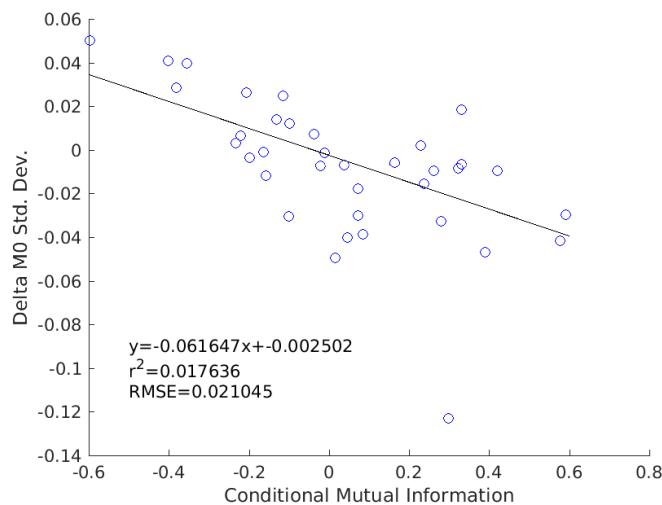
(B)

FIGURE 6.5: (A) Conditional mutual information between simulated high- and low-resolution acquisitions in the ten synthetic brain models (Figures 6.1 and 6.2) as a function of variable delay times. Delay times are fixed equal to one another. Positive CMI predicts an improvement in image quality, whereas negative CMI predicts a decrease in image quality. Optimum delay times are selected by maximizing CMI. Optimum values are marked by a red x for each synthetic brain model. (B) CMI between simulated high- and low-resolution acquisitions as a function of delay times, zoomed around optimum values.

Scan 2 parameters, the adjusted  $r^2$  from a robust linear regression is 0.0176, and the root-mean-squared error is 0.0210.



(A)

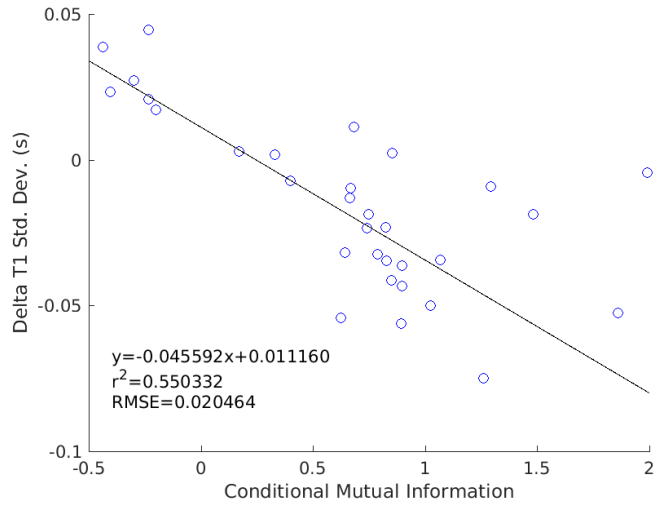


(B)

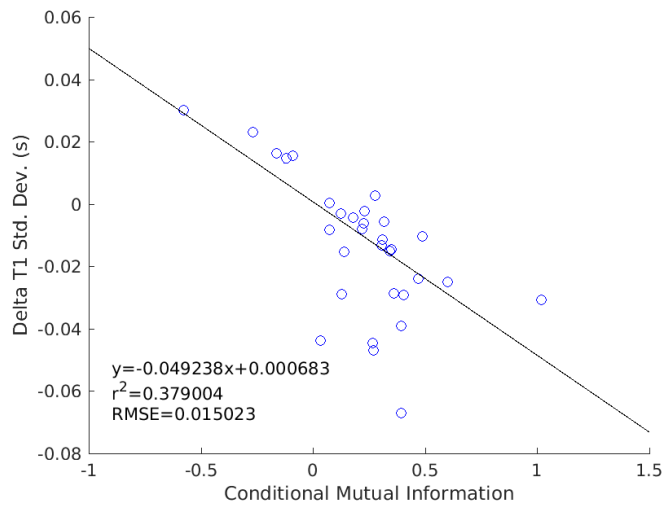
FIGURE 6.6: Inter-scan differences in reconstructed M0 standard deviations as a function of conditional mutual information between low-resolution pre-scan and high-resolution scan. (A) Low-resolution pre-scan at scan 1 acquisition parameters and high-resolution scan at scan 2 acquisition parameters (Table 6.1). (B) Low-resolution pre-scan at scan 2 acquisition parameters and high-resolution scan at scan 1 acquisition parameters (Table 6.1).

Similar trends are seen in T1 reconstruction in Figure 6.7. For conditioning on the low-resolution acquisition at Scan 1 parameters, the adjusted  $r^2$  from a robust linear regression is 0.5503. The root-mean-squared error is 0.0205. For conditioning on the

low-resolution acquisition at Scan 2 parameters, the adjusted  $r^2$  from a robust linear regression is 0.3790, and the root-mean-squared error is 0.0150.



(A)

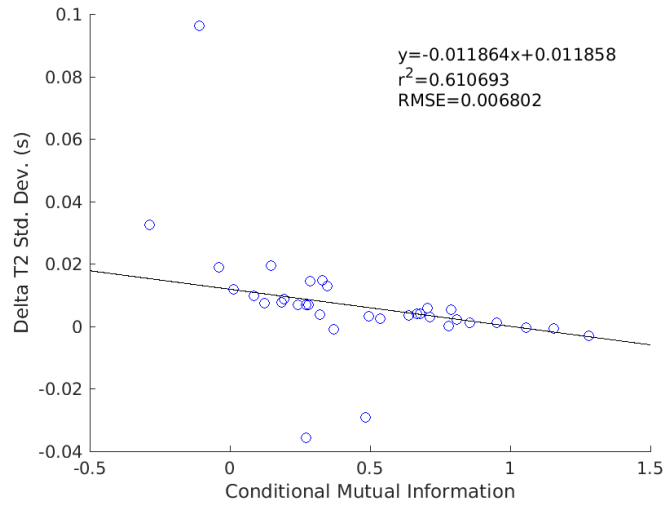


(B)

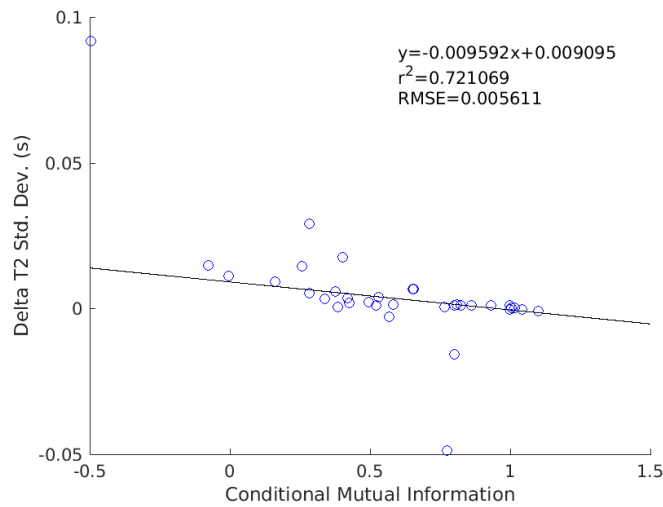
FIGURE 6.7: Inter-scan differences in reconstructed T1 standard deviations as a function of conditional mutual information between low-resolution pre-scan and high-resolution scan. (A) Low-resolution pre-scan at scan 1 acquisition parameters and high-resolution scan at scan 2 acquisition parameters (Table 6.1). (B) Low-resolution pre-scan at scan 2 acquisition parameters and high-resolution scan at scan 1 acquisition parameters (Table 6.1).

Once more, similar trends are seen in T2 reconstructions in Figure 6.8. For conditioning on the low-resolution acquisition at Scan 1 parameters, the adjusted  $r^2$  from a

robust linear regression is 0.6107. The root-mean-squared error is 0.0068. For conditioning on the low-resolution acquisition at Scan 2 parameters, the adjusted  $r^2$  from a robust linear regression is 0.7211, and the root-mean-squared error is 0.0056.



(A)



(B)

FIGURE 6.8: Inter-scan differences in reconstructed T2 standard deviations as a function of conditional mutual information between low-resolution pre-scan and high-resolution scan. (A) Low-resolution pre-scan at scan 1 acquisition parameters and high-resolution scan at scan 2 acquisition parameters (Table 6.1). (B) Low-resolution pre-scan at scan 2 acquisition parameters and high-resolution scan at scan 1 acquisition parameters (Table 6.1).

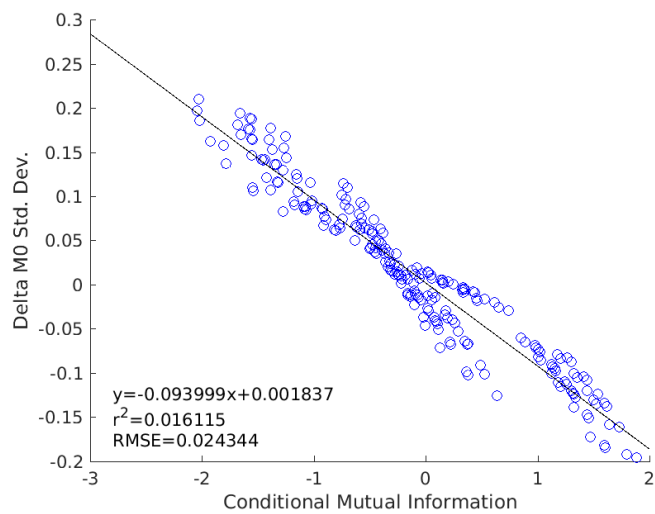
Finally, Figure 6.9 shows the same trends for inter-element calculations with unchanged acquisition parameters. A robust linear regression of M0 predictions resulted

in an adjusted  $r^2$  of 0.0161 and RMSE of 0.0243. A similar fit of T1 predictions yielded an adjusted  $r^2$  of 0.0261 and RMSE of 0.0342. Another fit of T2 predictions produced an adjusted  $r^2$  of 0.0665 and RMSE of 0.2445.

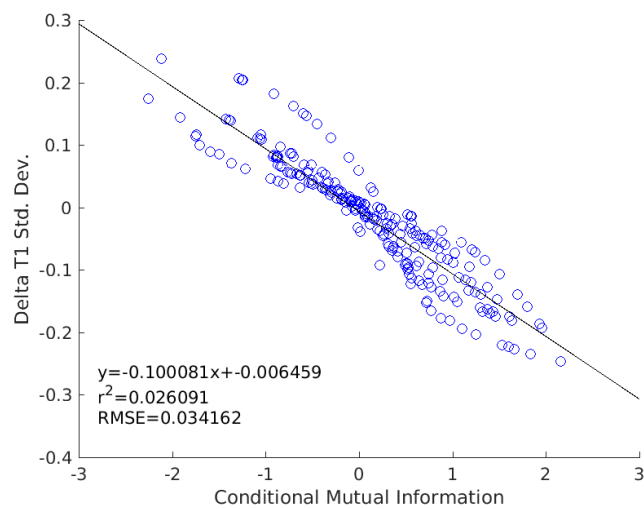
### 6.4.3 In Vivo Acquisitions

Three acquisitions were simulated with scan parameters listed in Table 6.5. For scan 1, mutual information was 1090210, for scan 2 it was 1108215, and for the theoretical optimum parameters it was 1112305. The standard deviations of the simulated posterior distributions from these acquisitions are shown in Figure 6.10. Additionally, the posterior distribution resulting from no uncertainty in the prior distribution is shown in Figure 6.10 and labeled as theoretical minimum. This is the smallest possible standard deviation of the posterior distribution, and all uncertainty results from uncertainty in the measurement itself. This serves as an approximation to the minimum detectable change in standard deviation from optimization, *i.e.* the smallest improvement in reconstruction that is differentiable from variance introduced by machine noise. It should be noted that the sampled posterior distributions are meant to predict relative changes in standard deviation, rather than accurately predict the magnitude of the standard deviations of measured distributions.

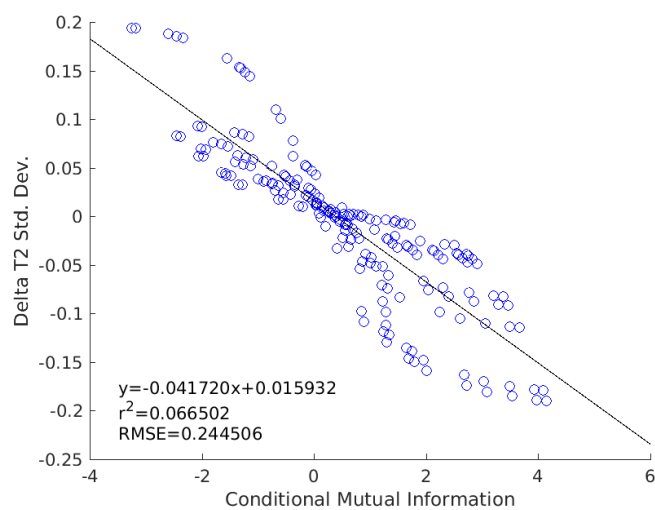
Reconstructed parametric maps from an acquisition with default clinical acquisition parameters are shown in Figure 6.11a, and reconstructed parametric maps from an acquisition with optimized acquisition parameters are shown in Figure 6.11b. The analysis of parameter value distributions from these two acquisitions is shown in Figure 6.12. The difference in standard deviations in M0, T1, and T2 distributions between the reconstructions from the two acquisitions is shown, where negative values correspond



(A)



(B)



(C)

FIGURE 6.9: Inter-element differences in reconstructed M0 (A), T1 (B), and T2 (C) standard deviations as a function of inter-element conditional mutual information. Positive CMI predicts decrease in reconstructed parametric map value standard deviation between different ISMRM NIST phantom elements in the same scan. Negative CMI predicts increase in standard deviation.



to an improvement in performance for an acquisition with optimized parameters. Distributions were analyzed from gray and white matter in the frontal, parietal, temporal, and occipital lobes, as well as cerebrospinal fluid in the ventricles. Distributions from nearly all ROIs showed improvement, with the exception of T2 in occipital lobe white matter. Table 6.6 shows p-values from two-sample F-tests for equal variances between distributions from the default and optimized acquisitions. Six out of nine changes in standard deviation were statistically significant in M0 maps. All changes in standard deviation were statistically significant in T1 maps. Four out of nine changes in standard deviation were statistically significant in T2 maps.

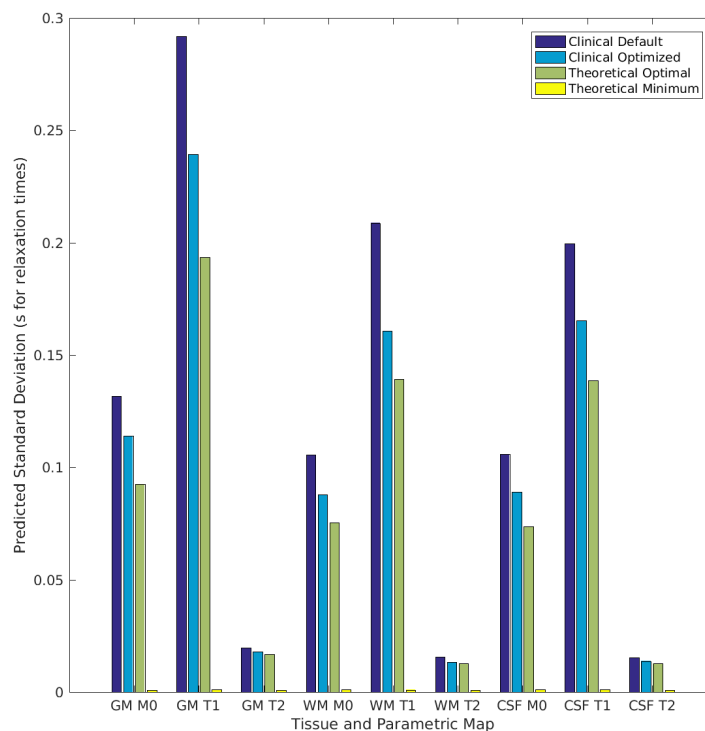


FIGURE 6.10: Standard deviations of simulated posterior distributions. Clinical default is obtained by sampling with clinical default acquisition parameters (scan 1). Clinical optimized is obtained by sampling with implemented optimal acquisition parameters (scan 2). Theoretical optimal is obtained by sampling with calculated optimal acquisition parameters that could not be feasible tested *in vivo*. Theoretical minimum is obtained by sampling with no biological uncertainty.

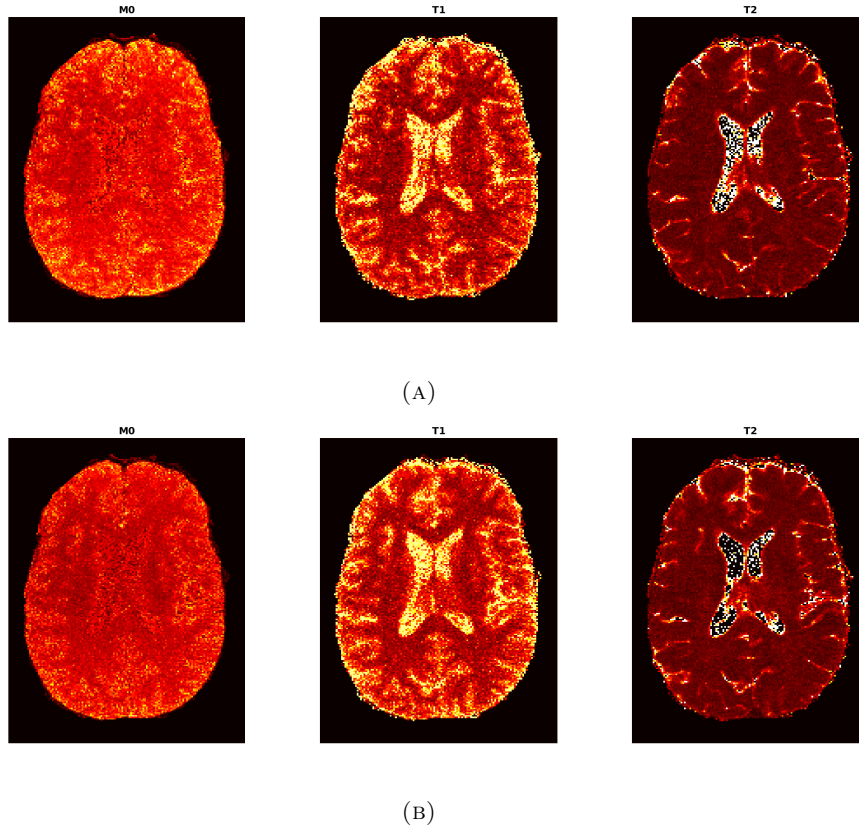


FIGURE 6.11: (A) Reconstructed parametric maps from acquisition with default clinical parameters (scan 1). (B) Reconstructed parametric maps from acquisition with optimized parameters (scan 2).

	M0	T1	T2
Fr. WM	0.0150	0.0000	0.1547
Fr. GM	0.0014	0.0000	0.0000
Pa. WM	0.0000	0.0000	0.9792
Pa. GM	0.0000	0.0005	0.0000
Tm. WM	0.0000	0.0000	0.0000
Tm. GM	0.4567	0.0029	0.6346
Oc. WM	0.5353	0.0476	0.8956
Oc. GM	0.0002	0.0000	0.5565
CSF	0.5559	0.0000	0.0007

TABLE 6.6: The  $p$ -values from two-sample F-tests for equal variances between distributions from the default and optimized acquisitions. Entries outlined in green are significant at the  $p < 0.05$  level. Entries outlined in red are nonsignificant at the  $p < 0.05$  level.

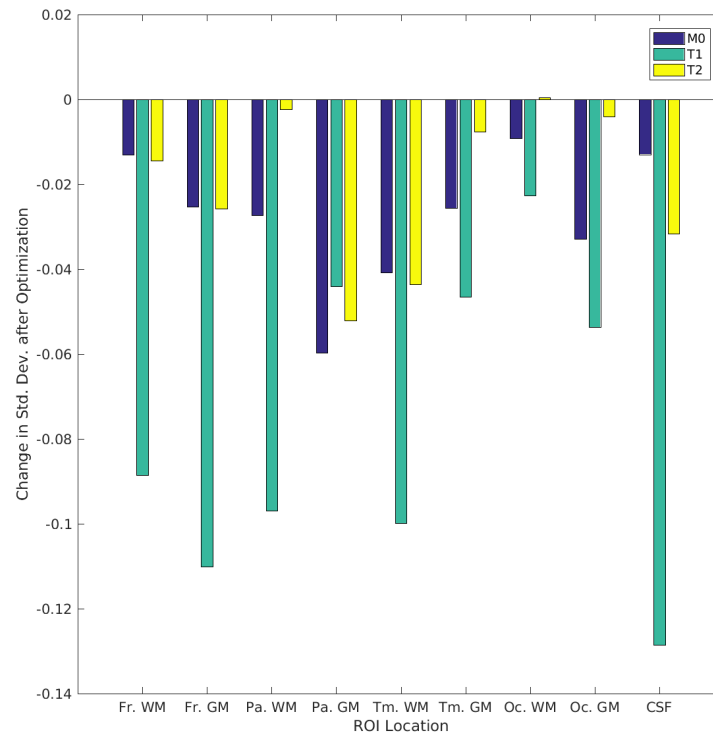


FIGURE 6.12: Change in standard deviation in several tissues between two *in vivo* scans. Negative change in standard deviation denotes an improvement in reconstruction going from the default clinical parameters to the optimized parameters. Analysis was performed in frontal lobe white matter (Fr. WM) and gray matter (Fr. GM), parietal lobe white matter (Pa. WM) and gray matter (Pa. GM), temporal lobe white matter (Tm. WM) and gray matter (Tm. GM), occipital lobe white matter (Oc. WM) and gray matter (Oc. GM), and cerebrospinal fluid (CSF). Change in standard deviation is shown separately for M0, T1, and T2 maps.

## 6.5 Discussion

Figure 6.5 shows the potential value of information optimization. Ten realistic scenarios produced optimum delay times ranging from 280 ms to 480 ms. Current empirical delay times are selected to be 500 ms. Additionally, this does not account for further improvements through independent delay time optimization.

It should be noted that the low-resolution simulation is performed with acquisition spacing of 500 ms. Conditional mutual information is larger for scenarios in which the magnitude of the difference between the optimal delay time and 500 ms is larger.

This behavior is expected, because a highly sub-optimal first acquisition leaves more information unknown before for a second, optimized acquisition. Conversely, scenarios in which the optimal acquisition spacing is near 500 ms show very nearly zero conditional mutual information between the first and second acquisitions. Finally, for delay times outside the advisable range, conditional mutual information becomes negative, denoting a loss of information for extremely poor parameter selection.

Figure 6.6 shows a clear trend in the difference in standard deviation for reconstructed  $M_0$  values as a function of conditional mutual information between a low-resolution scan at scan 1 parameters and a high-resolution scan at scan 2 parameters (Figure 6.6a) and between a low-resolution scan at scan 2 parameters and a high-resolution scan at scan 1 parameters (Figure 6.6b). Positive conditional mutual information indicates an improvement in the information content of the acquisition, whereas negative conditional mutual information indicates a worsening. The fact that the linear regression intercept is approximately zero is evidence that the information framework accurately captures this behavior. As expected, the change in standard deviation is negatively correlated with conditional mutual information. The significance of this trend is that positive conditional mutual information predicts a decrease in reconstruction variance, and negative mutual information predicts an increase. These results demonstrate that conditional mutual information is a feasible metric for use in image quality prediction given a quickly acquired pre-scan.

Figures 6.7 and 6.8 show the same results for standard deviation in reconstructed T1 and T2 values, respectively. These parametric maps displayed the same trends with respect to conditional mutual information. The correlation is stronger for the T1 and T2 results.

Generally, these figures indicate that conditional mutual information is predictive of the change in image variability. Reconstruction uncertainty is negatively correlated with conditional mutual information. Positive conditional mutual information corresponds to decreased reconstruction standard deviation, while negative CMI corresponds to increased reconstruction standard deviation.

In order to more strongly corroborate the trend seen in image quality prediction with the available data, predictions are made between NIST phantom elements. Figure 6.9 shows the additional data analysis used to confirm the relationship observed in the previous figures. Here, conditional mutual information is calculated between the physical properties of two different phantom elements while the scan parameters are unchanged. This is equivalent to changing the image subject properties between scans instead of the acquisition parameters. While this is obviously not a scenario encountered in the clinic, it still provides further validation of the predictive capabilities of the conditional mutual information model. A similar trend can be seen in this data, where an increase in conditional mutual information predicts better performance, *i.e.* a decrease in standard deviation of reconstructed values, and vice versa.

For the predicted *in vivo* performance, mutual information correlated with the standard deviation of simulated posterior distributions in the expected manner. The default clinical parameters resulted in the smallest mutual information among the three configurations simulated. In turn, the standard deviations of the posterior distributions of M0, T1, and T2 values in gray matter, white matter, and cerebrospinal fluid for an acquisition with default clinical parameters were the greatest among the three configurations. The theoretically optimal acquisition parameters reached through interior point optimization of mutual information of course resulted in the greatest mutual information. The standard deviations of the posterior distributions for an acquisition

with the optimum acquisition parameters were the smallest. The practical optimized acquisition parameters that were employed in the acquisition of *in vivo* data resulted in mutual information greater than an acquisition with default parameters, but less than that of the theoretically optimal acquisition. As expected, the standard deviations of the simulated posterior distributions from this acquisition fall between the default clinical acquisition and the theoretically optimal acquisition. Furthermore, the magnitude of the theoretical minimum standard deviation resulting solely from measurement uncertainty is more than an order of magnitude smaller than the differences in posterior distribution standard deviations between default and optimized acquisitions. This suggests performance improvement through optimization is distinguishable from the variance introduced through machine noise.

Information optimization of acquisition parameters resulted in decreases in standard deviation of reconstructed M0, T1, and T2 distributions in all but one case. Nineteen out of 26 of these improvements were statistically significant. The one increase in standard deviation was not statistically significant. These results confirm the hypothesis of this work.

## Chapter 7

# Conclusion

The information theory analysis described in this work allows quantitative guidance of synthetic MRI acquisition parameters across multiple applications. The results represent an advanced quantitative understanding of the relationship between acquisition parameters and reproducibility of parametric maps, which in quantitative imaging is especially critical for reliable diagnosis. The major contribution of this research is the quantitative assessment of the impact of acquisition parameter selection on reconstructed parametric map uncertainty, which is currently understood only empirically. The tool described in this work guides the optimized selection of acquisition parameters for a given application, and it may be useful in the evaluation of repeatability and reproducibility in clinical settings.

This information theory approach becomes increasingly valuable for increasingly complex acquisition signal models. For the simplest possible signal models, optimal parameters are either known analytically or readily intuited. Complex signal models, such as 3D-QALAS, benefit greatly from quantitative optimization, but there is no existing analytical optimum against which to validate the information optimization. For

this reason, in specific aim 1 the information optimization framework was tested and validated on a spoiled GRE acquisition signal model. The optimal flip angle is known analytically and defined by the Ernst angle equation. A nonlinear mutual information optimization of the flip angle showed near perfect agreement with the analytical solution, giving confidence to the application of this approach to more complex signal models. In specific aim 2, the model was applied to acquisitions of an ISMRM NIST phantom. The results showed that acquisitions with greater mutual information resulted in reconstructed parametric maps with smaller standard deviations. In specific aim 3, the model was extended to consider conditional mutual information and applied to *in vivo* data. It was demonstrated that information theory optimization of acquisition parameters led to a statistically significant improvement in the standard deviations of reconstructed parametric maps in nearly all of the sampled regions of the *in vivo* images. This confirms the hypothesis defined in Chapter 1.

Conditional mutual information has potential applications in corrective updates to acquisitions in real time. Real time updates could range in complexity from updating  $k$ -space locations of new measurements in undersampled acquisitions to altering pulse sequence parameters mid-scan to maximize information acquired within clinical constraints. Another important feature of this framework is that it is application agnostic. It should be applicable to most clinical modeling problems and could provide a quantitative understanding of model parameter information content in clinical image reconstruction, treatment planning, and other applications.

Another approach to quantitative imaging, called magnetic resonance fingerprinting, boasts high accuracy and robustness to parameter variance.[24, 197, 198] In fingerprinting, reproducibility is a function of the dictionary and randomization of parameters,



rather than fixed parameter selection. It may be possible to apply an information theory approach to quantifying magnetic resonance fingerprinting accuracy as a function of dictionary selection, or to inform randomization of parameters based on information content.

Compressed sensing is also based on information theory principles, but it is not mutually exclusive with this approach. Compressed sensing allows measurement of a small number of signal samples without previous knowledge of the signal or image, whereas the technique described in this work provides a method to quantitatively optimize acquisition parameters given knowledge of the signal or imaging location.[6, 199–201] Furthermore, compressed sensing can be included in the signal model, so optimal acquisition parameters can be obtained for such acquisitions, and compressed sensing parameters can even be included in the optimization space.

Model complexity is currently a limiting factor in this work. These results have demonstrated the importance of a representative model, but in cases where modeling a subject requires high-dimensional parameter spaces, the current method struggles with computation time. Using Gauss-Hermite quadrature becomes prohibitively expensive as the problem dimensionality increases. In future work, Markov chain Monte Carlo integration will be implemented to calculate mutual information for higher-dimensional problems.

For the development of new protocols, this method could determine a theoretical optimal set of parameters for image quality faster than an empirical approach comprising the iteration of scans with different sets of parameters. Currently, some amount of trial and error is unavoidable when tuning parameters for different scan times, tissues, and resolutions. Depending on the availability of machines, volunteers, and patients, this

process can even span days or weeks. A quantitative parameter optimization method reduces the logistical burden of developing new protocols by minimizing scanner time and labor. A promising example application would be determining the set of optimized parameters for a target range of T1 or T2 values. With a short enough computation time, an information theory analysis could be run *ad hoc* at a scanner and acquisition parameters could be tailored to any given situation. A brief computation time also enables the alteration of existing protocols through informed adjustments that ensure consistent image quality. For example, acquisition parameters change when coverage or field of view is adjusted, but there is not always enough scan time to optimize parameters again empirically. Furthermore, evaluation of image quality during parameter tuning is almost always visual and subjective. Mutual information is a useful metric for quantitative evaluation of measurement quality and possibly for standardization of parameter tuning processes.

This mutual information framework could also prove valuable in model-based reconstruction for several other modalities. The framework is applicable to any other method with an acceptable model, including most magnetic resonance pulse sequences and reconstructions. The physics model will determine the prior and posterior statistics, and the mutual information framework can be used to similarly determine the most informative data to be used in reconstruction.

Overall, these results show that more informative acquisition parameters, which are quantified by mutual information, can be selected to reduce output uncertainty. In the case of 3D-QALAS, selecting a more informative combination of acquisition delay times in the sequence can reduce M0, T1, and T2 parametric map uncertainty. Thus, an information theory analysis of the construction of parametric maps in synthetic MRI using

mutual information shows promise for quantifying and guiding selection of acquisition parameters in synthetic MRI.

# Appendix A

## Important Terms

Herein is a list of important terms and their definitions used throughout this work.

**Coefficient of repeatability:** The least significant difference between two measurements in a repeatability experiment.

**Coefficient of reproducibility:** The least significant difference between two measurements in a reproducibility experiment.

**Differential entropy:** A measure of average self-information of a random variable.

**Information:** The resolution of uncertainty. Information known about an event reduces uncertainty and entropy.

**Information entropy:** A measure of how much information an event contains. An event that is nearly certain to occur possesses very little information, so greater information gain results from events with greater information entropy.

**Moment:** The  $n^{\text{th}}$  moment is the expectation value of the  $n^{\text{th}}$  power of the difference between  $x$  and the mean.

**Mutual information:** The amount of information one random variable contains about another random variable. Equivalently, it is the reduction in one random variable's entropy due to knowledge of another variable. Alternatively, it is also the expected value of the Kullback-Leibler divergence between the probability distributions of the two random variables.

**Noise:** The random variation in image intensity that is often the result of electronic noise.

**Population variance:** The expectation value of the squared difference between a random variable and the population mean. It is also the second central moment.

**Repeatability:** Repeatability concerns measurement precision under identical testing conditions.

**Reproducibility:** Reproducibility concerns measurement precision under varying testing conditions, *e.g.* scanner, location, operator, or institution.

**Sample variance:** The estimation of population variance from a sample of finite size.

**Signal-to-noise ratio (SNR):** The ratio of average signal intensity in the foreground to background noise.

**Standard deviation:** The square root of the variance.

**Uncertainty:** The absence of information or knowledge of a random variable. It is quantified by entropy.

## Appendix B

# Mathematical Symbols

Herein is a list of symbols used in derivations and model descriptions throughout this work.

$\alpha$	Flip angle or excitation angle.
$\eta$	General tissue properties of the image subject. In the case of quantitative MRI in brain, this is usually a vector of M0, T1, and T2 values, <i>i.e.</i> $\eta(\mathbf{x}) = [M_0(\mathbf{x}), T_1(\mathbf{x}), T_2(\mathbf{x})]$ .
$\mathcal{F}$	Discrete Fourier transform.
$\mathcal{G}$	Signal model operator. In this work, it is often shown as $\mathcal{G}_{\boldsymbol{\mu}}(\eta)$ , which denotes the signal model with acquisition parameters, $\boldsymbol{\mu}$ , operating on physical tissue properties of the image subject, $\eta$ , to create signal values $\mathbf{z}$ .

$\boldsymbol{\mu}$	General acquisition parameters. These are used in the optimization of acquisition parameters. For example, if the signal model is a spoiled GRE sequence, one subsampling parameter may be flip angle. If the signal model is 3D-QALAS, $\boldsymbol{\mu}$ may be a vector of five delay times which partly define acquisition spacing.
$\Omega$	Image subject domain. The $n^{\text{th}}$ tissue label of the image subject is described by $\boldsymbol{x} \in \Omega_n$ .
$\sigma$	Standard deviation.
$\Sigma$	Covariance matrix.
$\boldsymbol{\theta}$	General subsampling parameters. These are used if optimizing the $k$ -space subsampling trajectory. For example, if variable density Poisson disc subsampling is used, one subsampling parameter may be acceleration of variable spacing in the $x$ -direction, and another may be the same in the $y$ -direction.
$I(X; Y)$	Mutual information between random variable $X$ and random variable $Y$ .
$N(\boldsymbol{x} \boldsymbol{m}, \Sigma)$	Multivariate normal probability density function of $\boldsymbol{x}$ with mean $\boldsymbol{m}$ and covariance $\Sigma$ .

---

$R_1$	Longitudinal relaxation rate. This is the inverse of $T_1$ .
$R_2$	Transverse relaxation rate. This is the inverse of $T_2$ .
$R_2^*$	Effective transverse relaxation rate. This is the inverse of $T_2^*$ .
$T_1$	Longitudinal relaxation time.
$T_2$	Transverse relaxation time.
$T_2^*$	Effective transverse relaxation time.
$\mathbf{x}$	Position in image space.
$z$	Signal value. This is complex valued in $k$ -space for 3D-QALAS.



# Appendix C

## Mathematical Derivations

This appendix contains mathematical derivations omitted from the main body for brevity and clarity. Derivations are organized into appropriate sections and referenced in the main text where relevant.

### C.1 Chapter 3 Derivations

#### C.1.1 Mutual Information

Derivation of relationship to entropy:

$$\begin{aligned} I(X; Y) &= \int_y \int_x p(x, y) \log \left( \frac{p(x, y)}{p(x)p(y)} \right) dx dy \\ &= \int_y \int_x p(x, y) \log \frac{p(x, y)}{p(x)} dx dy - \int_y \int_x p(x, y) \log p(y) dx dy \\ &= \int_y \int_x p(x, y) \log p(y|x) dx dy - \int_y p(y) \log p(y) dy \\ &= H(Y) - H(Y|X) \end{aligned} \tag{C.1}$$

Symmetric case:

$$\begin{aligned}
I(X; Y) &= \int_y \int_x p(x, y) \log \left( \frac{p(x, y)}{p(x)p(y)} \right) dx dy \\
&= \int_y \int_x p(x, y) \log \frac{p(x, y)}{p(y)} dx dy - \int_y \int_x p(x, y) \log p(x) dx dy \\
&= \int_y \int_x p(x, y) \log p(x|y) dx dy - \int_x p(x) \log p(x) dx \\
&= H(X) - H(X|Y)
\end{aligned} \tag{C.2}$$

### C.1.2 Joint Mutual Information

Derivation of relationship to entropy:

$$\begin{aligned}
I(X; Y, Z) &= \int_z \int_y \int_x p(x, y, z) \log \left( \frac{p(x, y, z)}{p(x)p(y, z)} \right) dx dy dz \\
&= \int_z \int_y \int_x p(x, y, z) \log \frac{p(x, y, z)}{p(x)} dx dy dz - \int_z \int_y \int_x p(x, y, z) \log p(y, z) dx dy dz \\
&= \int_z \int_y \int_x p(x, y, z) \log p(x|y, z) dx dy dz - \int_z \int_y p(y, z) \log p(y, z) dy dz \\
&= H(Y, Z) - H(Y, Z|X)
\end{aligned} \tag{C.3}$$

Symmetric case:

$$\begin{aligned}
I(X; Y, Z) &= \int_z \int_y \int_x p(x, y, z) \log \left( \frac{p(x, y, z)}{p(x)p(y, z)} \right) dx dy dz \\
&= \int_z \int_y \int_x p(x, y, z) \log \frac{p(x, y, z)}{p(y, z)} dx dy dz - \int_z \int_y \int_x p(x, y, z) \log p(x) dx dy dz \\
&= \int_z \int_y \int_x p(x, y, z) \log p(x|y, z) dx dy dz - \int_x p(x) \log p(x) dx \\
&= H(X) - H(X|Y, Z)
\end{aligned} \tag{C.4}$$

### C.1.3 Conditional Mutual Information

Derivation of relationship to entropy:

$$\begin{aligned}
I(X; Y|Z) &= \int_y \int_x p(x, y|z) \log \left( \frac{p(x, y|z)}{p(x|z)p(y|z)} \right) dx dy \\
&= \int_y \int_x p(x, y|z) \log \frac{p(x, y|z)}{p(x|z)} dx dy - \int_y \int_x p(x, y|z) \log p(y|z) dx dy \\
&= \int_y \int_x p(x, y|z) \log p(y|x, z) dx dy - \int_y p(y|z) \log p(y|z) dy \\
&= H(Y|Z) - H(Y|X, Z)
\end{aligned} \tag{C.5}$$

Symmetric case:

$$\begin{aligned}
I(X; Y|Z) &= \int_y \int_x p(x, y|z) \log \left( \frac{p(x, y|z)}{p(x|z)p(y|z)} \right) dx dy \\
&= \int_y \int_x p(x, y|z) \log \frac{p(x, y|z)}{p(y|z)} dx dy - \int_y \int_x p(x, y|z) \log p(x|z) dx dy \\
&= \int_y \int_x p(x, y|z) \log p(x|y, z) dx dy - \int_x p(x|z) \log p(x|z) dx \\
&= H(X|Z) - H(X|Y, Z)
\end{aligned} \tag{C.6}$$

### C.1.4 Conditional Mutual Information Relationships

Conditional mutual information relationships derived for three measurements,  $\mathbf{z}_1$ ,  $\mathbf{z}_2$ , and  $\mathbf{z}_3$ , and parametric maps  $\boldsymbol{\eta} = [M0, T1, T2]$ :

$$\begin{aligned}
I(\boldsymbol{\eta}; \mathbf{z}_1) &= H(\boldsymbol{\eta}) - H(\boldsymbol{\eta}|\mathbf{z}_1) \\
I(\boldsymbol{\eta}; \mathbf{z}_2|\mathbf{z}_1) &= H(\boldsymbol{\eta}|\mathbf{z}_1) - H(\boldsymbol{\eta}|\mathbf{z}_2, \mathbf{z}_1) \\
I(\boldsymbol{\eta}; \mathbf{z}_3|\mathbf{z}_2, \mathbf{z}_1) &= H(\boldsymbol{\eta}|\mathbf{z}_2, \mathbf{z}_1) - H(\boldsymbol{\eta}|\mathbf{z}_3, \mathbf{z}_2, \mathbf{z}_1)
\end{aligned} \tag{C.7}$$

$$\begin{aligned}
I(\boldsymbol{\eta}; \mathbf{z}_3 | \mathbf{z}_2, \mathbf{z}_1) &= H(\boldsymbol{\eta} | \mathbf{z}_2, \mathbf{z}_1) - H(\boldsymbol{\eta} | \mathbf{z}_3, \mathbf{z}_2, \mathbf{z}_1) \\
&= H(\boldsymbol{\eta} | \mathbf{z}_1) - I(\boldsymbol{\eta}; \mathbf{z}_2 | \mathbf{z}_1) - H(\boldsymbol{\eta} | \mathbf{z}_3, \mathbf{z}_2, \mathbf{z}_1) \\
&= H(\boldsymbol{\eta}) - I(\boldsymbol{\eta}; \mathbf{z}_1) - I(\boldsymbol{\eta}; \mathbf{z}_2 | \mathbf{z}_1) - H(\boldsymbol{\eta} | \mathbf{z}_3, \mathbf{z}_2, \mathbf{z}_1)
\end{aligned} \tag{C.8}$$

$$\begin{aligned}
I(\boldsymbol{\eta}; \mathbf{z}_3 | \mathbf{z}_2, \mathbf{z}_1) &= I(\boldsymbol{\eta}; \mathbf{z}_3, \mathbf{z}_2, \mathbf{z}_1) - I(\boldsymbol{\eta}; \mathbf{z}_1) - I(\boldsymbol{\eta}; \mathbf{z}_2 | \mathbf{z}_1) \\
&= I(\boldsymbol{\eta}; \mathbf{z}_3, \mathbf{z}_2, \mathbf{z}_1) - I(\boldsymbol{\eta}; \mathbf{z}_2) - I(\boldsymbol{\eta}; \mathbf{z}_1 | \mathbf{z}_2) \\
&= I(\boldsymbol{\eta}; \mathbf{z}_3, \mathbf{z}_2, \mathbf{z}_1) - I(\boldsymbol{\eta}; \mathbf{z}_2, \mathbf{z}_1)
\end{aligned} \tag{C.9}$$

Generalized mutual information relationships for a new measurement,  $\mathbf{z}$ , given  $N$  previous measurements,  $\mathbf{d}_1, \dots, \mathbf{d}_N$ :

$$I(\boldsymbol{\eta}; \mathbf{z} | \mathbf{d}) = H(\boldsymbol{\eta} | \mathbf{d}) - H(\boldsymbol{\eta} | \mathbf{z}, \mathbf{d}) \tag{C.10}$$

$$I(\boldsymbol{\eta}; \mathbf{z} | \mathbf{d}) = H(\boldsymbol{\eta}) - I(\boldsymbol{\eta}; \mathbf{d}_1) - \sum_{i=2}^N I(\boldsymbol{\eta}; \mathbf{d}_i | \mathbf{d}_{i-1}, \dots, \mathbf{d}_1) - H(\boldsymbol{\eta}; \mathbf{z}, \mathbf{d}) \tag{C.11}$$

$$I(\boldsymbol{\eta}; \mathbf{z} | \mathbf{d}) = I(\boldsymbol{\eta}; \mathbf{z}, \mathbf{d}) - I(\boldsymbol{\eta}; \mathbf{d}_1) - \sum_{i=2}^N I(\boldsymbol{\eta}; \mathbf{d}_i | \mathbf{d}_{i-1}, \dots, \mathbf{d}_1) \tag{C.12}$$

$$I(\boldsymbol{\eta}; \mathbf{z} | \mathbf{d}) = I(\boldsymbol{\eta}; \mathbf{z}, \mathbf{d}) - I(\boldsymbol{\eta}; \mathbf{d}) \tag{C.13}$$

## C.2 Chapter 4 and 5 Derivations

### C.2.1 Linear Approximation

The distribution  $p(\mathbf{z})$  is assumed to be normal.

$$I(\boldsymbol{\eta}; \mathbf{z}) = H(\mathbf{z}) - H(\mathbf{z}|\boldsymbol{\eta}) = \frac{1}{2} \ln \left( (2\pi e)^2 \cdot |\Sigma_{\mathbf{z}}| \right) - \frac{1}{2} \ln \left( (2\pi e)^2 \cdot |\Sigma_{\nu}| \right) \quad (\text{C.14})$$

Recall conditional entropy  $H(\mathbf{z}|\boldsymbol{\eta})$  from the problem definition:

$$|\Sigma_{\nu}| = \prod_{i=1}^D \sigma_{\nu,i}^2. \quad (\text{C.15})$$

The covariance matrix of the evidence is defined as

$$\Sigma_{\mathbf{z}} = \text{cov}[\mathbf{z}, \mathbf{z}] = E[(\mathbf{z} - m_{\mathbf{z}})(\mathbf{z} - m_{\mathbf{z}})^T] = E[\mathbf{z}\mathbf{z}^T] - m_{\mathbf{z}}m_{\mathbf{z}}^T, \quad (\text{C.16})$$

or, in matrix form,

$$\Sigma_{\mathbf{z}} = \begin{bmatrix} E[z_1 z_1] - m_{z_1} m_{z_1} & E[z_1 z_2] - m_{z_1} m_{z_2} & \dots & E[z_1 z_n] - m_{z_1} m_{z_n} \\ E[z_2 z_1] - m_{z_2} m_{z_1} & E[z_2 z_2] - m_{z_2} m_{z_2} & \dots & E[z_2 z_n] - m_{z_2} m_{z_n} \\ \vdots & \vdots & \ddots & \vdots \\ E[z_n z_1] - m_{z_n} m_{z_1} & E[z_n z_2] - m_{z_n} m_{z_2} & \dots & E[z_n z_n] - m_{z_n} m_{z_n} \end{bmatrix}. \quad (\text{C.17})$$

The mean of  $\mathbf{z}$  can be computed with Gauss-Hermite quadrature.

$$\begin{aligned}
m_{\mathbf{z}} &= E[\mathbf{z}] = \int \mathbf{z} p(\mathbf{z}) d\mathbf{z} \\
&= \int \mathbf{z} \left( \int p(\mathbf{z}|\boldsymbol{\eta}) p(\boldsymbol{\eta}) d\boldsymbol{\eta} \right) d\mathbf{z} \\
&= \int \mathbf{z} \left( \int p(\mathbf{z}|\boldsymbol{\eta}) N(\boldsymbol{\eta}|\mathbf{m}_{\boldsymbol{\eta}}, \Sigma_{\boldsymbol{\eta}}) d\boldsymbol{\eta} \right) d\mathbf{z} \\
&= \int \mathbf{z} \left( \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} p(\mathbf{z}|\boldsymbol{\eta}_q) \right) d\mathbf{z} \\
&= \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \int \mathbf{z} p(\mathbf{z}|\boldsymbol{\eta}_q) d\mathbf{z} \\
&= \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \int \mathbf{z} N(\mathbf{z}|\mathcal{G}_{\boldsymbol{\mu}}(\boldsymbol{\eta}_q), \Sigma_{\nu}) \\
&= \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \cdots \sum_{k_N=1}^{K_N} \omega_{k_N} \mathbf{z}_{q,k} \\
&= \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \cdots \sum_{k_N=1}^{K_N} \omega_{k_N} \left( \sqrt{2}\Sigma_{\nu} \mathbf{x}_k + \mathcal{G}_{\boldsymbol{\mu}}(\boldsymbol{\eta}_q) \right)
\end{aligned} \tag{C.18}$$

$$\boldsymbol{\eta}_q = [\eta_{q_1}, \dots, \eta_{q_N}] = \sqrt{2}\Sigma_{\boldsymbol{\eta}} \mathbf{x}_q + \mathbf{m}_{\boldsymbol{\eta}} \tag{C.19}$$

$$\mathbf{z}_{q,k} = \sqrt{2}\Sigma_{\nu} \mathbf{x}_k + \mathcal{G}_{\boldsymbol{\mu}}(\boldsymbol{\eta}_q) \tag{C.20}$$

Likewise for the term  $E[\mathbf{z}\mathbf{z}^T]$ :

$$\begin{aligned}
E[\mathbf{z}\mathbf{z}^T] &= \int \mathbf{z}\mathbf{z}^T p(\mathbf{z}) d\mathbf{z} \\
&= \int \mathbf{z}\mathbf{z}^T \left( \int p(\mathbf{z}|\boldsymbol{\eta}) p(\boldsymbol{\eta}) d\boldsymbol{\eta} \right) d\mathbf{z} \\
&= \int \mathbf{z}\mathbf{z}^T \left( \int p(\mathbf{z}|\boldsymbol{\eta}) N(\boldsymbol{\eta}|\mathbf{m}_\eta, \Sigma_\eta) d\boldsymbol{\eta} \right) d\mathbf{z} \\
&= \int \mathbf{z}\mathbf{z}^T \left( \sum_{q_1=1}^{Q_1} \omega_{q_1} \dots \sum_{q_N=1}^{Q_N} \omega_{q_N} p(\mathbf{z}|\boldsymbol{\eta}_q) \right) d\mathbf{z} \\
&= \sum_{q_1=1}^{Q_1} \omega_{q_1} \dots \sum_{q_N=1}^{Q_N} \omega_{q_N} \int \mathbf{z}\mathbf{z}^T p(\mathbf{z}|\boldsymbol{\eta}_q) d\mathbf{z} \\
&= \sum_{q_1=1}^{Q_1} \omega_{q_1} \dots \sum_{q_N=1}^{Q_N} \omega_{q_N} \int \mathbf{z}\mathbf{z}^T N(\mathbf{z}|\mathcal{G}_\mu(\boldsymbol{\eta}_q), \Sigma_\nu) \\
&= \sum_{q_1=1}^{Q_1} \omega_{q_1} \dots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \dots \sum_{k_N=1}^{K_N} \omega_{k_N} \mathbf{z}_{q,k} \mathbf{z}_{q,k}^T \\
&= \sum_{q_1=1}^{Q_1} \omega_{q_1} \dots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \dots \sum_{k_N=1}^{K_N} \omega_{k_N} \\
&\quad \cdot \left( \sqrt{2}\Sigma_\nu \mathbf{x}_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right) \left( \sqrt{2}\Sigma_\nu \mathbf{x}_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right)^T.
\end{aligned} \tag{C.21}$$

Thus, the quadrature approximation of the covariance matrix of the evidence  $\Sigma_z$  is

$$\begin{aligned}
\Sigma_z &= \sum_{q_1=1}^{Q_1} \omega_{q_1} \dots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \dots \sum_{k_N=1}^{K_N} \omega_{k_N} \left( \sqrt{2}\Sigma_\nu \mathbf{x}_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right) \left( \sqrt{2}\Sigma_\nu \mathbf{x}_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right)^T \\
&\quad - \sum_{q_1=1}^{Q_1} \omega_{q_1} \dots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \dots \sum_{k_N=1}^{K_N} \omega_{k_N} \left( \sqrt{2}\Sigma_\nu \mathbf{x}_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right) \\
&\quad \cdot \left[ \sum_{q_1=1}^{Q_1} \omega_{q_1} \dots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \dots \sum_{k_N=1}^{K_N} \omega_{k_N} \left( \sqrt{2}\Sigma_\nu \mathbf{x}_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right) \right]^T.
\end{aligned} \tag{C.22}$$

Approximation of mutual information is straightforward once the covariance matrix  $\Sigma_z$  is known.

$$I(\boldsymbol{\eta}; \mathbf{z}) = H(\mathbf{z}) - H(\mathbf{z}|\boldsymbol{\eta}) = \frac{1}{2} \ln \left( (2\pi e)^D \cdot |\Sigma_z| \right) - \frac{1}{2} \ln \left( (2\pi e)^D \prod_{i=1}^D \sigma_{\nu,i}^2 \right) \quad (\text{C.23})$$

For a scalar-valued measurement  $z$ , Equations C.18, C.21, and C.23 simplify to the following:

$$m_z = \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \left( \sqrt{2}\Sigma_\nu x_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right) \quad (\text{C.24})$$

$$E[z^2] = \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \left( \sqrt{2}\Sigma_\nu x_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right)^2 \quad (\text{C.25})$$

$$I(\boldsymbol{\eta}; z) = \frac{1}{2} \ln \left\{ 2\pi e \left[ \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \left( \sqrt{2}\Sigma_\nu x_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right)^2 - \left( \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \left( \sqrt{2}\Sigma_\nu x_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right) \right)^2 \right] \right\} - \frac{1}{2} \ln (2\pi e \sigma_\nu^2) \quad (\text{C.26})$$

## C.2.2 Nonlinear Computation

The signal model  $\mathbf{z}$  is a nonlinear function of parameters  $\boldsymbol{\eta}$ .

$$\mathbf{z} = \mathcal{G}_\mu(\boldsymbol{\eta}(\mathbf{x})) + \nu \quad (\text{C.27})$$

The conditional probability  $p(\mathbf{z}|\boldsymbol{\eta})$  is defined as a multivariate normal distribution  $N(\boldsymbol{\xi}|\boldsymbol{\mu}, \Sigma)$ , a function of  $\boldsymbol{\xi}$  with mean  $\boldsymbol{\mu}$  and covariance  $\Sigma$ , in the problem statement.

$$p(\mathbf{z}|\boldsymbol{\eta}) = N(\mathbf{z}|\mathcal{G}_\mu(\boldsymbol{\eta}(\mathbf{x})), \Sigma_\nu) \quad (\text{C.28})$$



Similarly, the probability  $p(\boldsymbol{\eta})$  is defined as another multivariate normal distribution in the problem statement.

$$p(\boldsymbol{\eta}) = N(\boldsymbol{\eta} | \mathbf{m}_\eta, \Sigma_\eta) \quad (\text{C.29})$$

It is most straightforward to begin with mutual information defined as the difference in entropies:

$$I(\boldsymbol{\eta}; \mathbf{z}) = H(\mathbf{z}) - H(\mathbf{z} | \boldsymbol{\eta}) \quad (\text{C.30})$$

The entropy of a multivariate normal distribution is  $H(N) = \frac{1}{2} \ln((2\pi e)^D \det(\Sigma))$ , so the term  $H(\mathbf{z} | \boldsymbol{\eta})$  can be immediately written

$$H(\mathbf{z} | \boldsymbol{\eta}) = \frac{1}{2} \ln \left( (2\pi e)^D \prod_{i=1}^D \sigma_{\nu,i}^2 \right). \quad (\text{C.31})$$

The term  $H(\mathbf{z})$ , however, must be approximated through multiple applications of Gauss-Hermite quadrature. This approximation is possible because  $H(\mathbf{z})$  may be represented in terms of distributions defined to be normal in the problem statement:

$$\begin{aligned} H(\mathbf{z}) &= \int p(\mathbf{z}) \ln(p(\mathbf{z})) d\mathbf{z} \\ &= \int \left( \int p(\mathbf{z} | \boldsymbol{\eta}) p(\boldsymbol{\eta}) d\boldsymbol{\eta} \right) \cdot \ln \left( \int p(\mathbf{z} | \boldsymbol{\eta}) p(\boldsymbol{\eta}) d\boldsymbol{\eta} \right) d\mathbf{z} \\ &= \int \left( \int N(\mathbf{z} | \mathcal{G}_\mu(\boldsymbol{\eta}), \Sigma_\nu) N(\boldsymbol{\eta} | \mathbf{m}_\eta, \Sigma_\eta) d\boldsymbol{\eta} \right) \\ &\quad \cdot \ln \left( \int N(\mathbf{z} | \mathcal{G}_\mu(\boldsymbol{\eta}), \Sigma_\nu) N(\boldsymbol{\eta} | \mathbf{m}_\eta, \Sigma_\eta) d\boldsymbol{\eta} \right) d\mathbf{z}. \end{aligned} \quad (\text{C.32})$$

The first quadrature iteration is performed by making the substitution  $\boldsymbol{\eta} = \sqrt{2}\Sigma_\eta\mathbf{x} + \mathbf{m}_\eta$ .

$$\begin{aligned}
H(\mathbf{z}) &= \int p(\mathbf{z}) \ln(p(\mathbf{z})) d\mathbf{z} \\
&= \int \left( \int p(\mathbf{z}|\boldsymbol{\eta}) p(\boldsymbol{\eta}) d\boldsymbol{\eta} \right) \ln(p(\mathbf{z})) d\mathbf{z} \\
&= \int \left( \int p(\mathbf{z}|\boldsymbol{\eta}) N(\boldsymbol{\eta}|\mathbf{m}_\eta, \Sigma_\eta) d\boldsymbol{\eta} \right) \cdot \ln(p(\mathbf{z})) d\mathbf{z} \\
&\approx \int \left[ \pi^{-N/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \dots \sum_{q_N=1}^{Q_N} \omega_{q_N} p(\mathbf{z}|\boldsymbol{\eta}_q) \right] \cdot \ln(p(\mathbf{z})) d\mathbf{z}
\end{aligned} \tag{C.33}$$

The resulting quadrature points are defined  $\boldsymbol{\eta}_q = [\eta_{q_1}, \dots, \eta_{q_N}] = \sqrt{2}\Sigma_\eta\mathbf{x}_q + \mathbf{m}_\eta$ , where  $\mathbf{x}_q = [x_{q_1}, \dots, x_{q_N}]$  is the combination of roots of the physicists' Hermite polynomial  $H_n$  corresponding to the indices  $q_1$  through  $q_N$ . The second quadrature iteration is performed almost identically and approximates the integral inside the natural logarithm by making the substitution  $\boldsymbol{\eta} = \sqrt{2}\Sigma_\eta\mathbf{x} + \mathbf{m}_\eta$ .

$$\begin{aligned}
H(\mathbf{z}) &= \int \left[ \pi^{-N/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \dots \sum_{q_N=1}^{Q_N} \omega_{q_N} p(\mathbf{z}|\boldsymbol{\eta}_q) \right] \cdot \ln(p(\mathbf{z})) d\mathbf{z} \\
&= \int \left[ \pi^{-N/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \dots \sum_{q_N=1}^{Q_N} \omega_{q_N} p(\mathbf{z}|\boldsymbol{\eta}_q) \right] \cdot \ln \left( \int p(\mathbf{z}|\boldsymbol{\eta}) p(\boldsymbol{\eta}) d\boldsymbol{\eta} \right) d\mathbf{z} \\
&= \int \left[ \pi^{-N/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \dots \sum_{q_N=1}^{Q_N} \omega_{q_N} p(\mathbf{z}|\boldsymbol{\eta}_q) \right] \cdot \ln \left( \int p(\mathbf{z}|\boldsymbol{\eta}) N(\boldsymbol{\eta}|\mathbf{m}_\eta, \Sigma_\eta) d\boldsymbol{\eta} \right) d\mathbf{z} \\
&= \int \left[ \pi^{-N/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \dots \sum_{q_N=1}^{Q_N} \omega_{q_N} p(\mathbf{z}|\boldsymbol{\eta}_q) \right] \\
&\quad \cdot \ln \left( \pi^{-N/2} \sum_{s_1=1}^{S_1} \omega_{s_1} \dots \sum_{s_N=1}^{S_N} \omega_{s_N} \cdot p(\mathbf{z}|\boldsymbol{\eta}_s) \right) d\mathbf{z}
\end{aligned} \tag{C.34}$$

The resulting quadrature points are defined  $\boldsymbol{\eta}_s = [\eta_{s_1}, \dots, \eta_{s_N}] = \sqrt{2}\Sigma_\eta \mathbf{x}_s + \mathbf{m}_\eta$ . The third quadrature iteration is performed by making the substitution  $\mathbf{z} = \sqrt{2}\Sigma_\nu \mathbf{x} + \mathcal{G}_\mu(\boldsymbol{\eta})$ .

$$\begin{aligned}
H(\mathbf{z}) &= \int \left[ \pi^{-N/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} p(\mathbf{z}|\boldsymbol{\eta}_q) \right] \\
&\quad \cdot \ln \left( \pi^{-N/2} \sum_{s_1=1}^{S_1} \omega_{s_1} \cdots \sum_{s_N=1}^{S_N} \omega_{s_N} \cdot p(\mathbf{z}|\boldsymbol{\eta}_s) \right) d\mathbf{z} \\
&= \pi^{-N/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \\
&\quad \cdot \int N(\mathbf{z}|\mathcal{G}_\mu(\boldsymbol{\eta}_q), \Sigma_\nu) \cdot \ln \left( \pi^{-N/2} \sum_{s_1=1}^{S_1} \omega_{s_1} \cdots \sum_{s_N=1}^{S_N} \omega_{s_N} \cdot N(\mathbf{z}|\mathcal{G}_\mu(\boldsymbol{\eta}_s), \Sigma_\nu) \right) d\mathbf{z} \\
&= \pi^{-N/2-P/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \cdots \sum_{k_P=1}^{K_P} \omega_{k_P} \\
&\quad \cdot \ln \left( \pi^{-N/2} \sum_{s_1=1}^{S_1} \omega_{s_1} \cdots \sum_{s_N=1}^{S_N} \omega_{s_N} \cdot N \left( \left( \sqrt{2}\Sigma_\nu \mathbf{x}_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right) \middle| \mathcal{G}_\mu(\boldsymbol{\eta}_s), \Sigma_\nu \right) \right)
\end{aligned} \tag{C.35}$$

Thus, the computation of mutual information takes the following form:

$$\begin{aligned}
I(\boldsymbol{\eta}; \mathbf{z}) &= H(\mathbf{z}) - H(\mathbf{z}|\boldsymbol{\eta}) \\
&= \pi^{-N/2-P/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \cdots \sum_{k_P=1}^{K_P} \omega_{k_P} \\
&\quad \cdot \ln \left( \pi^{-N/2} \sum_{s_1=1}^{S_1} \omega_{s_1} \cdots \sum_{s_N=1}^{S_N} \omega_{s_N} \cdot N \left( \left( \sqrt{2}\Sigma_\nu \mathbf{x}_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right) \middle| \mathcal{G}_\mu(\boldsymbol{\eta}_s), \Sigma_\nu \right) \right) \\
&\quad - \frac{1}{2} \ln \left( (2\pi e)^D \prod_{i=1}^D \sigma_{\nu,i}^2 \right).
\end{aligned} \tag{C.36}$$

For a scalar-valued measurement  $z$ , Equation C.36 reduces to

$$\begin{aligned}
I(\boldsymbol{\eta}; z) &= H(z) - H(z|\boldsymbol{\eta}) \\
&= \pi^{-N/2-1/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_1=1}^{K_1} \omega_{k_1} \\
&\quad \cdot \ln \left( \pi^{-N/2} \sum_{s_1=1}^{S_1} \omega_{s_1} \cdots \sum_{s_N=1}^{S_N} \omega_{s_N} \cdot N \left( \left( \sqrt{2} \sigma_\nu x_k + \mathcal{G}_\mu(\boldsymbol{\eta}_q) \right) \middle| \mathcal{G}_\mu(\boldsymbol{\eta}_s), \sigma_\nu \right) \right) \\
&\quad - \frac{1}{2} \ln (2\pi e \sigma_\nu^2).
\end{aligned} \tag{C.37}$$

### C.3 Chapter 6 Derivations

#### C.3.1 Conditional Mutual Information for Jointly Gaussian Measurements

The signal model  $\mathbf{z}$  is a nonlinear function of parameters  $\boldsymbol{\eta}$ .

$$\mathbf{z}(\boldsymbol{\eta}, \boldsymbol{\theta}, \mathbf{k}) = S(\boldsymbol{\theta}, \mathbf{k}) \odot \mathcal{FQ}(\boldsymbol{\eta}, \mathbf{k}) + \nu(\mathbf{k}) \tag{C.38}$$

The conditional probability  $p(\mathbf{z}|\boldsymbol{\eta})$  is defined as a multivariate normal distribution  $N(\boldsymbol{\xi}|\boldsymbol{\mu}, \Sigma)$ , a function of  $\boldsymbol{\xi}$  with mean  $\boldsymbol{\mu}$  and covariance  $\Sigma$ , in the problem statement.

$$p(\mathbf{z}(\boldsymbol{\eta}, \boldsymbol{\theta}, \mathbf{k})|\boldsymbol{\eta}(\mathbf{x})) = N(\mathbf{z}(\boldsymbol{\eta}, \boldsymbol{\theta}, \mathbf{k})|S(\boldsymbol{\theta}, \mathbf{k}) \odot \mathcal{FQ}(\boldsymbol{\eta}, \mathbf{k}), \Sigma_\nu) \sim N(S(\boldsymbol{\theta}, \mathbf{k}) \odot \mathcal{FQ}(\boldsymbol{\eta}, \mathbf{k}), \Sigma_\nu) \tag{C.39}$$

Similarly, the probability  $p(\boldsymbol{\eta})$  is defined as another multivariate normal distribution in the problem statement.

$$p(\boldsymbol{\eta}(\mathbf{x})) = N(\boldsymbol{\eta}(\mathbf{x})|\mathbf{m}_\eta, \Sigma_\eta) \sim N(\mathbf{m}_\eta, \Sigma_\eta) \tag{C.40}$$

The derivation begins with mutual information defined as the difference in entropies:

$$I(\boldsymbol{\eta}; \mathbf{z}) = H(\mathbf{z}) - H(\mathbf{z}|\boldsymbol{\eta}) = H(\mathbf{z}) - \frac{1}{2} \ln \left( (2\pi e)^{N_{\text{meas}}} \sigma_{\nu,r}^{2N_{\text{meas}}} \sigma_{\nu,i}^{2N_{\text{meas}}} \right). \quad (\text{C.41})$$

The entropy of a multivariate normal distribution is  $H = \frac{1}{2} \ln \left( (2\pi e)^N \det(\Sigma) \right)$ , so the term  $H(\mathbf{z}|\boldsymbol{\eta})$  can be immediately written

$$H(\mathbf{z}|\boldsymbol{\eta}) = \frac{1}{2} \ln \left( (2\pi e)^{N_{\text{meas}}} \sigma_{\nu,r}^{2N_{\text{meas}}} \sigma_{\nu,i}^{2N_{\text{meas}}} \right). \quad (\text{C.42})$$

The term  $H(\mathbf{z})$ , however, must be approximated through multiple applications of Gauss-Hermite quadrature. This approximation is possible because  $H(\mathbf{z})$  may be represented in terms of distributions defined to be normal in the problem statement:

$$\begin{aligned} H(\mathbf{z}) &= \int p(\mathbf{z}) \ln(p(\mathbf{z})) d\mathbf{z} \\ &= \int \left( \int p(\mathbf{z}|\boldsymbol{\eta}) p(\boldsymbol{\eta}) d\boldsymbol{\eta} \right) \cdot \ln \left( \int p(\mathbf{z}|\boldsymbol{\eta}) p(\boldsymbol{\eta}) d\boldsymbol{\eta} \right) d\mathbf{z} \\ &= \int \left( \int f(\mathbf{z}|S \odot \mathcal{FQ}(\boldsymbol{\eta}), \Sigma_\nu) f(\boldsymbol{\eta}|\mathbf{m}_\eta, \Sigma_\eta) d\boldsymbol{\eta} \right) \\ &\quad \cdot \ln \left( \int f(\mathbf{z}|S \odot \mathcal{FQ}(\boldsymbol{\eta}), \Sigma_\nu) f(\boldsymbol{\eta}|\mathbf{m}_\eta, \Sigma_\eta) d\boldsymbol{\eta} \right) d\mathbf{z}. \end{aligned} \quad (\text{C.43})$$

The first quadrature iteration is performed by making the substitution  $\boldsymbol{\eta} = \sqrt{2}\Sigma_{\boldsymbol{\eta}}\boldsymbol{x} + \mathbf{m}_{\boldsymbol{\eta}}$ .

$$\begin{aligned}
H(\mathbf{d}) &= \int p(\mathbf{d}) \ln(p(\mathbf{d})) d\mathbf{d} = \int \left( \int p(\mathbf{d}|\boldsymbol{\eta}) p(\boldsymbol{\eta}) d\boldsymbol{\eta} \right) \ln(p(\mathbf{d})) d\mathbf{d} \\
&= \int \left( \int p(\mathbf{d}|\boldsymbol{\eta}) f(\boldsymbol{\eta}|\mathbf{m}_{\boldsymbol{\eta}}, \Sigma_{\boldsymbol{\eta}}) d\boldsymbol{\eta} \right) \cdot \ln(p(\mathbf{d})) d\mathbf{d} \\
&\approx \int \left[ \pi^{-N/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} p(\mathbf{d}|\boldsymbol{\eta}_q) \right] \cdot \ln(p(\mathbf{d})) d\mathbf{d} \\
\boldsymbol{\eta}_q &= \sqrt{2}\Sigma_{\boldsymbol{\eta}}\mathbf{x}_q + \mathbf{m}_{\boldsymbol{\eta}}
\end{aligned} \tag{C.44}$$

The resulting quadrature points are defined  $\boldsymbol{\eta}_q = [\eta_{q_1}, \dots, \eta_{q_N}] = \sqrt{2}\Sigma_{\boldsymbol{\eta}}\mathbf{x}_q + \mathbf{m}_{\boldsymbol{\eta}}$ , where  $\mathbf{x}_q = [x_{q_1}, \dots, x_{q_N}]$  is the combination of roots of the physicists' Hermite polynomial  $H_n$  corresponding to the indices  $q_1$  through  $q_N$ . The second quadrature iteration is performed similarly by making the substitution  $\mathbf{z}_q = \sqrt{2}\Sigma_{\nu}\mathbf{x} + S \odot \mathcal{FQ}(\boldsymbol{\eta}_q)$ .

$$\begin{aligned}
H(d) &= \pi^{-N/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \int p(\mathbf{d}|\boldsymbol{\eta}_q) \cdot \ln(p(\mathbf{d})) d\mathbf{d} \\
&= \pi^{-N/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \int f(\mathbf{d}|S \odot \mathcal{Q}(\boldsymbol{\eta}_q), \Sigma_{\nu}) \cdot \ln(p(\mathbf{d})) d\mathbf{d} \\
&\approx \pi^{-N/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \\
&\quad \cdot \left[ \pi^{-P} \sum_{k_{1,r}=1}^{K_{1,r}} \omega_{k_{1,r}} \sum_{k_{1,i}=1}^{K_{1,i}} \omega_{k_{1,i}} \cdots \sum_{k_{P,r}=1}^{K_{P,r}} \omega_{k_{P,r}} \sum_{k_{P,i}=1}^{K_{P,i}} \omega_{k_{P,i}} \cdot \ln(p(\mathbf{d}_k)) \right] \\
\mathbf{d}_k &= \sqrt{2}\Sigma_{\nu}\mathbf{x}_k + S \odot \mathcal{Q}(\boldsymbol{\eta}_q)
\end{aligned} \tag{C.45}$$

The resulting quadrature points are now  $\mathbf{z}_{q,k} = \sqrt{2}\Sigma_{\nu}\mathbf{x}_k + S \odot \mathcal{FQ}(\boldsymbol{\eta}_q)$ . The second term is evaluated for the quadrature point  $\boldsymbol{\eta}_q$  before a second combination of roots of  $H_n$  (assuming  $n$  quadrature points are also used for this step) are computed for indices

$k_{1,r}$  through  $k_{P,r}$  and  $k_{1,i}$  through  $k_{P,i}$ . The third quadrature iteration approximates the integral inside the natural logarithm by making the substitution  $\boldsymbol{\eta} = \sqrt{2}\Sigma_\eta\mathbf{x} + \mathbf{m}_\eta$  again.

$$\begin{aligned}
H(\mathbf{d}) &= \pi^{-N/2-P} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_{1,r}=1}^{K_{1,r}} \omega_{k_{1,r}} \sum_{k_{1,i}=1}^{K_{1,i}} \omega_{k_{1,i}} \cdots \sum_{k_{P,r}=1}^{K_{P,r}} \omega_{k_{P,r}} \sum_{k_{P,i}=1}^{K_{P,i}} \omega_{k_{P,i}} \\
&\quad \cdot \ln \left[ \int p(\mathbf{d}_k|\boldsymbol{\eta})p(\boldsymbol{\eta})d\boldsymbol{\eta} \right] \\
&= \pi^{-N/2-P} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_{1,r}=1}^{K_{1,r}} \omega_{k_{1,r}} \sum_{k_{1,i}=1}^{K_{1,i}} \omega_{k_{1,i}} \cdots \sum_{k_{P,r}=1}^{K_{P,r}} \omega_{k_{P,r}} \sum_{k_{P,i}=1}^{K_{P,i}} \omega_{k_{P,i}} \\
&\quad \cdot \ln \left[ \int p(\mathbf{d}_k|\boldsymbol{\eta})f(\boldsymbol{\eta}|\mathbf{m}_\eta, \Sigma_\eta)d\boldsymbol{\eta} \right] \\
&\approx \pi^{-N/2-P} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_{1,r}=1}^{K_{1,r}} \omega_{k_{1,r}} \sum_{k_{1,i}=1}^{K_{1,i}} \omega_{k_{1,i}} \cdots \sum_{k_{P,r}=1}^{K_{P,r}} \omega_{k_{P,r}} \sum_{k_{P,i}=1}^{K_{P,i}} \omega_{k_{P,i}} \\
&\quad \cdot \ln \left[ \pi^{-N/2} \sum_{s_1=1}^{S_1} \omega_{s_1} \cdots \sum_{s_N=1}^{S_N} \omega_{s_N} p(\mathbf{d}_k|\boldsymbol{\eta}_s) \right] \\
&= \pi^{-N/2-P} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_{1,r}=1}^{K_{1,r}} \omega_{k_{1,r}} \sum_{k_{1,i}=1}^{K_{1,i}} \omega_{k_{1,i}} \cdots \sum_{k_{P,r}=1}^{K_{P,r}} \omega_{k_{P,r}} \sum_{k_{P,i}=1}^{K_{P,i}} \omega_{k_{P,i}} \\
&\quad \cdot \ln \left[ \pi^{-N/2} \sum_{s_1=1}^{S_1} \omega_{s_1} \cdots \sum_{s_N=1}^{S_N} \omega_{s_N} \prod_{p=1}^P p(\mathbf{d}_{k_{p,r}}|\boldsymbol{\eta}_s)p(\mathbf{d}_{k_{p,i}}|\boldsymbol{\eta}_s) \right] \\
\boldsymbol{\eta}_s &= \sqrt{2}\Sigma_\eta\mathbf{x}_s + \mathbf{m}_\eta
\end{aligned} \tag{C.46}$$

The third set of quadrature points are  $\boldsymbol{\eta}_s = \sqrt{2}\Sigma_\eta\mathbf{x}_s + \mathbf{m}_\eta$ . The probability distribution  $p(\mathbf{z}|\boldsymbol{\eta})$  is normal and can be evaluated at each quadrature point.

$$\begin{aligned}
p(\mathbf{z}_{q,k_{p,r}}|\boldsymbol{\eta}_s) &= N(\mathbf{z}_{q,k_{p,r}}|S(\boldsymbol{\theta}) \odot \mathcal{FQ}(\boldsymbol{\eta}_s), \Sigma_\nu) \\
&= N\left(\sqrt{2}\Sigma_\nu\mathbf{x}_{k_{p,r}} + \Re\left(S(\boldsymbol{\theta}) \odot \mathcal{FQ}\left(\sqrt{2}\Sigma_\eta\mathbf{x}_q + \mathbf{m}_\eta\right)\right)\right. \\
&\quad \left. \middle| S(\boldsymbol{\theta}) \odot \mathcal{FQ}\left(\sqrt{2}\Sigma_\eta\mathbf{x}_s + \mathbf{m}_\eta\right), \Sigma_\nu\right)
\end{aligned} \tag{C.47}$$

$$\begin{aligned}
p(\mathbf{z}_{q,k_{p,i}}|\boldsymbol{\eta}_s) &= N(\mathbf{z}_{q,k_{p,i}}|S(\boldsymbol{\theta}) \odot \mathcal{FQ}(\boldsymbol{\eta}_s), \Sigma_\nu) \\
&= N\left(\sqrt{2}\Sigma_\nu \mathbf{x}_{k_{p,i}} + \mathfrak{S}\left(S(\boldsymbol{\theta}) \odot \mathcal{FQ}\left(\sqrt{2}\Sigma_\eta \mathbf{x}_q + \mathbf{m}_\eta\right)\right)\right. \\
&\quad \left. |S(\boldsymbol{\theta}) \odot \mathcal{FQ}\left(\sqrt{2}\Sigma_\eta \mathbf{x}_s + \mathbf{m}_\eta\right), \Sigma_\nu\right)
\end{aligned} \tag{C.48}$$

The entropy approximation is then

$$\begin{aligned}
H(\mathbf{d}) &\approx \pi^{-N/2-P} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_{1,r}=1}^{K_{1,r}} \omega_{k_{1,r}} \sum_{k_{1,i}=1}^{K_{1,i}} \omega_{k_{1,i}} \cdots \sum_{k_{P,r}=1}^{K_{P,r}} \omega_{k_{P,r}} \sum_{k_{P,i}=1}^{K_{P,i}} \omega_{k_{P,i}} \\
&\cdot \ln \left[ \pi^{-N/2} \sum_{s_1=1}^{S_1} \omega_{s_1} \cdots \sum_{s_N=1}^{S_N} \omega_{s_N} \right. \\
&\quad \cdot \prod_{p=1}^P p(\sqrt{2}\Sigma_\nu \mathbf{x}_{k_{p,r}} + S_{\boldsymbol{\theta}_1} \odot \mathcal{FQ}_{\boldsymbol{\mu}_1}(\sqrt{2}\Sigma_\eta \mathbf{x}_q + \mathbf{m}_\eta) | \sqrt{2}\Sigma_\eta \mathbf{x}_s + \mathbf{m}_\eta) \\
&\quad \left. \cdot p(\sqrt{2}\Sigma_\nu \mathbf{x}_{k_{p,i}} + S_{\boldsymbol{\theta}_1} \odot \mathcal{FQ}_{\boldsymbol{\mu}_1}(\sqrt{2}\Sigma_\eta \mathbf{x}_q + \mathbf{m}_\eta) | \sqrt{2}\Sigma_\eta \mathbf{x}_s + \mathbf{m}_\eta) \right].
\end{aligned} \tag{C.49}$$



Entropy of the joint distribution is approximated by the same process.

$$\begin{aligned}
H(\mathbf{z}, \mathbf{d}) &= \int \int p(\mathbf{z}, \mathbf{d}) \ln(p(\mathbf{z}, \mathbf{d})) \, d\mathbf{z} d\mathbf{d} \\
&\approx \pi^{-N/2-P} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \cdot \sum_{j_{1,r}=1}^{J_{1,r}} \omega_{j_{1,r}} \sum_{j_{1,i}=1}^{J_{1,i}} \omega_{j_{1,i}} \cdots \sum_{j_{P,r}=1}^{J_{P,r}} \omega_{j_{P,r}} \sum_{j_{P,i}=1}^{J_{P,i}} \omega_{j_{P,i}} \\
&\cdot \sum_{k_{1,r}=1}^{K_{1,r}} \omega_{k_{1,r}} \sum_{k_{1,i}=1}^{K_{1,i}} \omega_{k_{1,i}} \cdots \sum_{k_{P,r}=1}^{K_{P,r}} \omega_{k_{P,r}} \sum_{k_{P,i}=1}^{K_{P,i}} \omega_{k_{P,i}} \cdot \ln \left\{ \pi^{-N/2} \sum_{s_1=1}^{S_1} \omega_{s_1} \cdots \sum_{s_N=1}^{S_N} \omega_{s_N} \right. \\
&\cdot \prod_{p=1}^P p \left( \sqrt{2}\Sigma_\nu \begin{bmatrix} \mathbf{x}_{j_{p,r}} \\ \mathbf{x}_{k_{p,r}} \end{bmatrix} + \begin{bmatrix} S_{\theta_2} \odot \mathcal{FQ}_{\mu_2}(\sqrt{2}\Sigma_\eta \mathbf{x}_q + \mathbf{m}_\eta) \\ S_{\theta_1} \odot \mathcal{FQ}_{\mu_1}(\sqrt{2}\Sigma_\eta \mathbf{x}_q + \mathbf{m}_\eta) \end{bmatrix} \middle| \sqrt{2}\Sigma_\eta \mathbf{x}_s + \mathbf{m}_\eta \right) \\
&\cdot p \left( \sqrt{2}\Sigma_\nu \begin{bmatrix} \mathbf{x}_{j_{p,i}} \\ \mathbf{x}_{k_{p,i}} \end{bmatrix} + \begin{bmatrix} S_{\theta_2} \odot \mathcal{FQ}_{\mu_2}(\sqrt{2}\Sigma_\eta \mathbf{x}_q + \mathbf{m}_\eta) \\ S_{\theta_1} \odot \mathcal{FQ}_{\mu_1}(\sqrt{2}\Sigma_\eta \mathbf{x}_q + \mathbf{m}_\eta) \end{bmatrix} \middle| \sqrt{2}\Sigma_\eta \mathbf{x}_s + \mathbf{m}_\eta \right) \left. \right\} \\
\end{aligned} \tag{C.50}$$

Thus, the computation of mutual information takes the following form:

$$\begin{aligned}
I(\boldsymbol{\eta}; \mathbf{z}|\mathbf{d}) &= H(\boldsymbol{\eta}|\mathbf{d}) - H(\boldsymbol{\eta}|\mathbf{z}, \mathbf{d}) = H(\mathbf{z}, \mathbf{d}) - H(\mathbf{z}, \mathbf{d}|\boldsymbol{\eta}) - (H(\mathbf{d}) - H(\mathbf{d}|\boldsymbol{\eta})) \\
&= H(\mathbf{z}, \mathbf{d}) - H(\mathbf{d}) + \frac{1}{2} \ln \left( (2\pi e)^2 \sigma_\nu^{2N_{acq,d}} \right) - \frac{1}{2} \ln \left( (2\pi e)^2 \sigma_\nu^{2(N_{acq,d} + N_{acq,z})} \right) \\
&= H(\mathbf{z}, \mathbf{d}) - H(\mathbf{d}) + \frac{1}{2} \ln(\sigma_\nu^{-2N_{acq,z}}). \\
\end{aligned} \tag{C.51}$$

### C.3.2 Joint Mutual Information for Independent Subsampling Masks

For measurements with independent subsampling masks,  $\boldsymbol{\xi}_1$  and  $\boldsymbol{\xi}_2$ , it is shown below that the mutual information between the image subject parameters and the joint distribution of measurements is equal to the sum of the mutual information between the

image subject parameters and each of the marginal measurement distributions.

$$\begin{aligned}
I(\boldsymbol{\eta}; \boldsymbol{\xi}_1, \boldsymbol{\xi}_2) &= \int_{\boldsymbol{\eta}} \int_{\boldsymbol{\xi}_1} \int_{\boldsymbol{\xi}_2} p(\boldsymbol{\eta}, \boldsymbol{\xi}_1, \boldsymbol{\xi}_2) \log \left( \frac{p(\boldsymbol{\eta}, \boldsymbol{\xi}_1, \boldsymbol{\xi}_2)}{p(\boldsymbol{\eta})p(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)} \right) d\boldsymbol{\eta} d\boldsymbol{\xi}_1 d\boldsymbol{\xi}_2 \\
&= \int_{\boldsymbol{\eta}} \int_{\boldsymbol{\xi}_1} \int_{\boldsymbol{\xi}_2} p(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 | \boldsymbol{\eta}) p(\boldsymbol{\eta}) \log \left( \frac{p(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 | \boldsymbol{\eta})}{p(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)} \right) d\boldsymbol{\eta} d\boldsymbol{\xi}_1 d\boldsymbol{\xi}_2 \\
&= \int_{\boldsymbol{\eta}} \int_{\boldsymbol{\xi}_1} \int_{\boldsymbol{\xi}_2} p(\boldsymbol{\xi}_1 | \boldsymbol{\eta}) p(\boldsymbol{\xi}_2 | \boldsymbol{\eta}) p(\boldsymbol{\eta}) \log \left( \frac{p(\boldsymbol{\xi}_1 | \boldsymbol{\eta}) p(\boldsymbol{\xi}_2 | \boldsymbol{\eta})}{p(\boldsymbol{\xi}_1) p(\boldsymbol{\xi}_2)} \right) d\boldsymbol{\eta} d\boldsymbol{\xi}_1 d\boldsymbol{\xi}_2 \\
&= \int_{\boldsymbol{\eta}} \int_{\boldsymbol{\xi}_1} \int_{\boldsymbol{\xi}_2} p(\boldsymbol{\xi}_1 | \boldsymbol{\eta}) p(\boldsymbol{\xi}_2 | \boldsymbol{\eta}) p(\boldsymbol{\eta}) \log \left( \frac{p(\boldsymbol{\xi}_1 | \boldsymbol{\eta})}{p(\boldsymbol{\xi}_1)} \right) d\boldsymbol{\eta} d\boldsymbol{\xi}_1 d\boldsymbol{\xi}_2 \\
&\quad + \int_{\boldsymbol{\eta}} \int_{\boldsymbol{\xi}_1} \int_{\boldsymbol{\xi}_2} p(\boldsymbol{\xi}_1 | \boldsymbol{\eta}) p(\boldsymbol{\xi}_2 | \boldsymbol{\eta}) p(\boldsymbol{\eta}) \log \left( \frac{p(\boldsymbol{\xi}_2 | \boldsymbol{\eta})}{p(\boldsymbol{\xi}_2)} \right) d\boldsymbol{\eta} d\boldsymbol{\xi}_1 d\boldsymbol{\xi}_2 \\
&= \int_{\boldsymbol{\eta}} \int_{\boldsymbol{\xi}_1} p(\boldsymbol{\eta}, \boldsymbol{\xi}_1) \log \left( \frac{p(\boldsymbol{\eta}, \boldsymbol{\xi}_1)}{p(\boldsymbol{\eta})p(\boldsymbol{\xi}_1)} \right) \left( \int_{\boldsymbol{\xi}_2} p(\boldsymbol{\xi}_2 | \boldsymbol{\eta}) d\boldsymbol{\xi}_2 \right) d\boldsymbol{\eta} d\boldsymbol{\xi}_1 \\
&\quad + \int_{\boldsymbol{\eta}} \int_{\boldsymbol{\xi}_2} p(\boldsymbol{\eta}, \boldsymbol{\xi}_2) \log \left( \frac{p(\boldsymbol{\eta}, \boldsymbol{\xi}_2)}{p(\boldsymbol{\eta})p(\boldsymbol{\xi}_2)} \right) \left( \int_{\boldsymbol{\xi}_1} p(\boldsymbol{\xi}_1 | \boldsymbol{\eta}) d\boldsymbol{\xi}_1 \right) d\boldsymbol{\eta} d\boldsymbol{\xi}_2 \\
&= \int_{\boldsymbol{\eta}} \int_{\boldsymbol{\xi}_1} p(\boldsymbol{\eta}, \boldsymbol{\xi}_1) \log \left( \frac{p(\boldsymbol{\eta}, \boldsymbol{\xi}_1)}{p(\boldsymbol{\eta})p(\boldsymbol{\xi}_1)} \right) d\boldsymbol{\eta} d\boldsymbol{\xi}_1 \\
&\quad + \int_{\boldsymbol{\eta}} \int_{\boldsymbol{\xi}_2} p(\boldsymbol{\eta}, \boldsymbol{\xi}_2) \log \left( \frac{p(\boldsymbol{\eta}, \boldsymbol{\xi}_2)}{p(\boldsymbol{\eta})p(\boldsymbol{\xi}_2)} \right) d\boldsymbol{\eta} d\boldsymbol{\xi}_2 \\
&= I(\boldsymbol{\eta}; \boldsymbol{\xi}_1) + I(\boldsymbol{\eta}; \boldsymbol{\xi}_2)
\end{aligned} \tag{C.52}$$

### C.3.3 Conditional Mutual Information for Independent Subsampling Masks

The entropy of the distribution of measurements with independent subsampling masks,  $H(\boldsymbol{\xi})$ , is approximated via Gauss-Hermite quadrature as demonstrated previously. The

first quadrature approximation is made as follows:

$$\begin{aligned}
H(\boldsymbol{\xi}) &= \int p(\boldsymbol{\xi}) \ln(p(\boldsymbol{\xi})) d\boldsymbol{\xi} = \int \left( \int p(\boldsymbol{\xi}|\boldsymbol{\eta}) p(\boldsymbol{\eta}) d\boldsymbol{\eta} \right) \ln(p(\boldsymbol{\xi})) d\boldsymbol{\xi} \\
&= \int \left( \int p(\boldsymbol{\xi}|\boldsymbol{\eta}) f(\boldsymbol{\eta}|m_\eta, \Sigma_\eta) d\boldsymbol{\eta} \right) \cdot \ln(p(\boldsymbol{\xi})) d\boldsymbol{\xi} \\
&\approx \int \left[ \pi^{-N/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} p(\boldsymbol{\xi}|\boldsymbol{\eta}_q) \right] \cdot \ln(p(\boldsymbol{\xi})) d\boldsymbol{\xi} \\
\eta_q &= \sqrt{2}\Sigma_\eta \mathbf{x}_q + \mathbf{m}_\eta
\end{aligned} \tag{C.53}$$

The second quadrature approximation is made as follows:

$$\begin{aligned}
H(\boldsymbol{\xi}) &= \pi^{-N/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \int p(\boldsymbol{\xi}|\boldsymbol{\eta}_q) \cdot \ln(p(\boldsymbol{\xi})) d\boldsymbol{\xi} \\
&= \pi^{-N/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \int f(\boldsymbol{\xi}|S \odot \mathcal{Q}(\boldsymbol{\eta}_q), \Sigma_\nu) \cdot \ln(p(\boldsymbol{\xi})) d\boldsymbol{\xi} \\
&\approx \pi^{-N/2} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \\
&\quad \cdot \left[ \pi^{-P} \sum_{k_{1,r}=1}^{K_{1,r}} \omega_{k_{1,r}} \sum_{k_{1,i}=1}^{K_{1,i}} \omega_{k_{1,i}} \cdots \sum_{k_{P,r}=1}^{K_{P,r}} \omega_{k_{P,r}} \sum_{k_{P,i}=1}^{K_{P,i}} \omega_{k_{P,i}} \cdot \ln(p(\boldsymbol{\xi}_k)) \right] \\
\boldsymbol{\xi}_k &= \sqrt{2}\Sigma_\nu \mathbf{x}_k + S \odot \mathcal{Q}(\boldsymbol{\eta}_q).
\end{aligned} \tag{C.54}$$

The final quadrature approximation is made as follows:

$$\begin{aligned}
H(\boldsymbol{\xi}) &= \pi^{-N/2-P} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_{1,r}=1}^{K_{1,r}} \omega_{k_{1,r}} \sum_{k_{1,i}=1}^{K_{1,i}} \omega_{k_{1,i}} \cdots \sum_{k_{P,r}=1}^{K_{P,r}} \omega_{k_{P,r}} \sum_{k_{P,i}=1}^{K_{P,i}} \omega_{k_{P,i}} \\
&\quad \cdot \ln \left[ \int p(\boldsymbol{\xi}_k | \boldsymbol{\eta}) p(\boldsymbol{\eta}) d\boldsymbol{\eta} \right] \\
&= \pi^{-N/2-P} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_{1,r}=1}^{K_{1,r}} \omega_{k_{1,r}} \sum_{k_{1,i}=1}^{K_{1,i}} \omega_{k_{1,i}} \cdots \sum_{k_{P,r}=1}^{K_{P,r}} \omega_{k_{P,r}} \sum_{k_{P,i}=1}^{K_{P,i}} \omega_{k_{P,i}} \\
&\quad \cdot \ln \left[ \int p(\boldsymbol{\xi}_k | \boldsymbol{\eta}) f(\boldsymbol{\eta} | \mathbf{m}_\eta, \Sigma_\eta) d\boldsymbol{\eta} \right] \\
&\approx \pi^{-N/2-P} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_{1,r}=1}^{K_{1,r}} \omega_{k_{1,r}} \sum_{k_{1,i}=1}^{K_{1,i}} \omega_{k_{1,i}} \cdots \sum_{k_{P,r}=1}^{K_{P,r}} \omega_{k_{P,r}} \sum_{k_{P,i}=1}^{K_{P,i}} \omega_{k_{P,i}} \\
&\quad \cdot \ln \left[ \pi^{-N/2} \sum_{s_1=1}^{S_1} \omega_{s_1} \cdots \sum_{s_N=1}^{S_N} \omega_{s_N} p(\boldsymbol{\xi}_k | \boldsymbol{\eta}_s) \right] \\
&= \pi^{-N/2-P} \sum_{q_1=1}^{Q_1} \omega_{q_1} \cdots \sum_{q_N=1}^{Q_N} \omega_{q_N} \sum_{k_{1,r}=1}^{K_{1,r}} \omega_{k_{1,r}} \sum_{k_{1,i}=1}^{K_{1,i}} \omega_{k_{1,i}} \cdots \sum_{k_{P,r}=1}^{K_{P,r}} \omega_{k_{P,r}} \sum_{k_{P,i}=1}^{K_{P,i}} \omega_{k_{P,i}} \\
&\quad \cdot \ln \left[ \pi^{-N/2} \sum_{s_1=1}^{S_1} \omega_{s_1} \cdots \sum_{s_N=1}^{S_N} \omega_{s_N} \prod_{p=1}^P p(\boldsymbol{\xi}_{k_{p,r}} | \boldsymbol{\eta}_s) p(\boldsymbol{\xi}_{k_{p,i}} | \boldsymbol{\eta}_s) \right]
\end{aligned}$$

$$\boldsymbol{\eta}_s = \sqrt{2} \Sigma_\eta \mathbf{x}_s + \mathbf{m}_\eta.$$

(C.55)

### C.3.4 Form of the Conditional Probability Distribution

The form of the conditional probability distribution  $p(\mathbf{z} | \boldsymbol{\eta}, \mathbf{d})$  for a second measurement  $\mathbf{z}$ , conditional on a first measurement  $\mathbf{d}$  and the parametric maps  $\boldsymbol{\eta}$  is of interest. The probability distribution of a set of measurements  $\mathbf{z}$  with acquisition parameters  $\boldsymbol{\mu}$  of an exactly known set of parametric maps  $\boldsymbol{\eta}(\mathbf{x})$  have previously been defined to be normally distributed about the forward model solution  $\mathcal{FQ}_\mu(\boldsymbol{\eta}(\mathbf{x}))$  with covariance matrix  $\Sigma_\nu$  representing the uncertainty from machine noise. Making the reasonable assumption that measurements are independent from one another with consistent uncertainty for

each measurement, the covariance matrix becomes diagonal,  $\Sigma_\nu = \sigma_\nu^2 I$ . It also follows that a second set of measurements  $\mathbf{z}$  is independent from the first set of measurements  $\mathbf{d}$ . Therefore, the distributions  $p(\mathbf{z}|\boldsymbol{\eta})$ ,  $p(\mathbf{d}|\boldsymbol{\eta})$ , and  $p(\mathbf{z}, \mathbf{d}|\boldsymbol{\eta})$  are given as follows:

$$p(\mathbf{z}|\boldsymbol{\eta}) = \prod_{i=1}^n p(z_i|\boldsymbol{\eta}) \sim \mathcal{N} \left( \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}, \begin{bmatrix} \sigma_{\eta_1}^2 & 0 & \dots & 0 & 0 \\ 0 & \sigma_{\eta_2}^2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \sigma_{\eta_{N-1}}^2 & 0 \\ 0 & 0 & \dots & 0 & \sigma_{\eta_N}^2 \end{bmatrix} \right) \quad (\text{C.56})$$

$$p(\mathbf{d}|\boldsymbol{\eta}) = \prod_{j=1}^m p(d_j|\boldsymbol{\eta}) \sim \mathcal{N} \left( \begin{bmatrix} d_1 \\ \vdots \\ d_m \end{bmatrix}, \begin{bmatrix} \sigma_{\eta_1}^2 & 0 & \dots & 0 & 0 \\ 0 & \sigma_{\eta_2}^2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \sigma_{\eta_{N-1}}^2 & 0 \\ 0 & 0 & \dots & 0 & \sigma_{\eta_N}^2 \end{bmatrix} \right) \quad (\text{C.57})$$

$$p(\mathbf{z}, \mathbf{d}|\boldsymbol{\eta}) = \prod_{i=1}^n p(z_i|\boldsymbol{\eta}) \prod_{j=1}^m p(d_j|\boldsymbol{\eta}) \sim \mathcal{N} \left( \begin{bmatrix} z_1 \\ \vdots \\ z_n \\ d_1 \\ \vdots \\ d_m \end{bmatrix}, \begin{bmatrix} \sigma_{\eta_1}^2 & 0 & \dots & 0 & 0 \\ 0 & \sigma_{\eta_2}^2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \sigma_{\eta_{N-1}}^2 & 0 \\ 0 & 0 & \dots & 0 & \sigma_{\eta_N}^2 \end{bmatrix} \right) \quad (\text{C.58})$$

The first set of measurements can also be written

$$\mathbf{d} = S_{\theta_1} \odot (\mathcal{F}\mathcal{Q}_{\mu_1}(\boldsymbol{\eta}(\mathbf{x})) + \nu(\mathbf{k})) = S_{\theta_1} \odot (\mathcal{G}_{\mu_1}(\boldsymbol{\eta}(\mathbf{x})) + \nu(\mathbf{k})). \quad (\text{C.59})$$

If the second set of measurements is conditional on the first,  $\mathbf{z}|\mathbf{d}$ , then in this model the subsampling parameters ( $\boldsymbol{\theta}_1 \rightarrow \boldsymbol{\theta}_2$ ) and the acquisition parameters ( $\boldsymbol{\mu}_1 \rightarrow \boldsymbol{\mu}_2$ ) can be changed as a result of acquiring the first set of measurements. The given parametric maps,  $\boldsymbol{\eta}(\mathbf{x})$ , and the uncertainty from machine noise,  $\nu(\mathbf{k})$ , remain the same. Therefore, the second set of measurements takes the form

$$\mathbf{z}|\mathbf{d} = S_{\boldsymbol{\theta}_2} \odot (\mathcal{F}\mathcal{Q}_{\boldsymbol{\mu}_2}(\boldsymbol{\eta}(\mathbf{x})) + \nu(\mathbf{k})) = S_{\boldsymbol{\theta}_2} \odot (\mathcal{G}_{\boldsymbol{\mu}_2}(\boldsymbol{\eta}(\mathbf{x})) + \nu(\mathbf{k})). \quad (\text{C.60})$$

where  $\boldsymbol{\theta}_2$  and  $\boldsymbol{\mu}_2$  are conditional on  $\mathbf{d}$ . This is still a set of normally distributed, independent measurements, and the mean is only conditional on  $\mathbf{d}$  insofar as  $\mathbf{d}$  affects the selection of model parameters for the second measurement. So,  $p(\mathbf{z}|\boldsymbol{\eta}, \mathbf{d})$  is defined as

$$p(\mathbf{z}|\boldsymbol{\eta}, \mathbf{d}) = \prod_{i=1}^n p(z_i|\boldsymbol{\eta}, \mathbf{d}) \sim \mathcal{N} \left( \begin{array}{c} \left[ \begin{array}{c} z_1(\boldsymbol{\theta}_2, \boldsymbol{\mu}_2|\mathbf{d}) \\ \vdots \\ z_n(\boldsymbol{\theta}_2, \boldsymbol{\mu}_2|\mathbf{d}) \end{array} \right] \\ \left[ \begin{array}{ccccc} \sigma_{\eta_1}^2 & 0 & \dots & 0 & 0 \\ 0 & \sigma_{\eta_2}^2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \sigma_{\eta_{N-1}}^2 & 0 \\ 0 & 0 & \dots & 0 & \sigma_{\eta_N}^2 \end{array} \right] \end{array} \right). \quad (\text{C.61})$$

For the purposes of maximizing conditional mutual information to select an optimum  $\mathbf{z}$  given an already acquired  $\mathbf{d}$ , it makes sense to think of  $\mathbf{z}$  and  $\mathbf{d}$  as independent (which, absent selection criteria for parameters of a second measurement that depend on the first, they are), because the model parameters are explored during optimization. The optimization problem to select optimum model parameters  $\boldsymbol{\theta}_2^*$  and  $\boldsymbol{\mu}_2^*$  for the second measurement  $\mathbf{z}$  is of the form

$$[\boldsymbol{\theta}_2^*, \boldsymbol{\mu}_2^*] = \underset{\boldsymbol{\theta}, \boldsymbol{\mu}}{\operatorname{argmax}} I(\boldsymbol{\eta}; \mathbf{z}|\mathbf{d}) = \underset{\boldsymbol{\theta}, \boldsymbol{\mu}}{\operatorname{argmax}} (I(\boldsymbol{\eta}; \mathbf{z}(\boldsymbol{\theta}, \boldsymbol{\mu}), \mathbf{d}) - I(\boldsymbol{\eta}, \mathbf{d})). \quad (\text{C.62})$$

As for the form of the conditional probability  $p(\mathbf{z}|\mathbf{d})$ , it cannot be computed analytically, but it can be written in terms of known probability distributions  $p(\mathbf{z}, \mathbf{d}|\boldsymbol{\eta})$ ,  $p(\mathbf{d}|\boldsymbol{\eta})$ , and  $p(\boldsymbol{\eta})$  with normal distribution  $N$ :

$$\begin{aligned} p(\mathbf{z}|\mathbf{d}) &= \frac{p(\mathbf{z}, \mathbf{d})}{p(\mathbf{d})} = \frac{\int p(\mathbf{z}, \mathbf{d}|\boldsymbol{\eta})p(\boldsymbol{\eta})d\boldsymbol{\eta}}{\int p(\mathbf{d}|\boldsymbol{\eta})p(\boldsymbol{\eta})d\boldsymbol{\eta}} \\ &= \frac{\int N(S_{\boldsymbol{\theta}_1} \odot \mathcal{G}_{\boldsymbol{\mu}_1}(\boldsymbol{\eta}), \Sigma_\nu)N(S_{\boldsymbol{\theta}_2} \odot \mathcal{G}_{\boldsymbol{\mu}_2}(\boldsymbol{\eta}), \Sigma_\nu)N(\mathbf{m}_\eta, \Sigma_\eta)d\boldsymbol{\eta}}{\int N(S_{\boldsymbol{\theta}_2} \odot \mathcal{G}_{\boldsymbol{\mu}_2}(\boldsymbol{\eta}), \Sigma_\nu)N(\mathbf{m}_\eta, \Sigma_\eta)d\boldsymbol{\eta}}. \end{aligned} \quad (\text{C.63})$$

This can be computed using Gauss-Hermite quadrature as in previous sections.

Conditional mutual information can be computed through this probability distribution in the following way:

$$\begin{aligned} I(\boldsymbol{\eta}; \mathbf{z}|\mathbf{d}) &= H(\mathbf{z}|\mathbf{d}) - H(\mathbf{z}|\boldsymbol{\eta}, \mathbf{d}) \\ &= H(\mathbf{z}|\mathbf{d}) - H(\boldsymbol{\eta}, \mathbf{z}, \mathbf{d}) + H(\boldsymbol{\eta}, \mathbf{d}) \\ &= H(\mathbf{z}|\mathbf{d}) - H(\mathbf{z}, \mathbf{d}|\boldsymbol{\eta}) - H(\boldsymbol{\eta}) + H(\mathbf{d}|\boldsymbol{\eta}) + H(\boldsymbol{\eta}) \\ &= H(\mathbf{z}|\mathbf{d}) + H(\mathbf{d}|\boldsymbol{\eta}) - H(\mathbf{z}, \mathbf{d}|\boldsymbol{\eta}). \end{aligned} \quad (\text{C.64})$$

Then, known entropies can be substituted.

$$\begin{aligned} &H(\mathbf{z}|\mathbf{d}) + H(\mathbf{d}|\boldsymbol{\eta}) - H(\mathbf{z}, \mathbf{d}|\boldsymbol{\eta}) \\ &= - \int \int p(\mathbf{z}, \mathbf{d}) \ln(p(\mathbf{z}|\mathbf{d})) dz d\mathbf{d} + \frac{1}{2} \ln \left( (2\pi e)^2 \sigma_\nu^{2N_{\text{acq},d}} \right) \\ &\quad - \frac{1}{2} \ln \left( (2\pi e)^2 \sigma_\nu^{2(N_{\text{acq},d} + N_{\text{acq},z})} \right) \\ &= - \int \int (p(\mathbf{z}, \mathbf{d}|\boldsymbol{\eta})p(\boldsymbol{\eta})d\boldsymbol{\eta}) \ln(p(\mathbf{z}|\mathbf{d})) dz d\mathbf{d} + \frac{1}{2} \ln(\sigma_\nu^{-2N_{\text{acq},z}}) \end{aligned} \quad (\text{C.65})$$

Finally, in terms of known normal distributions  $N$ , conditional mutual information is

$$\begin{aligned}
I(\boldsymbol{\eta}; \mathbf{z}|\mathbf{d}) = & - \int \int \int N(\mathbf{S}_{\boldsymbol{\theta}_1} \odot \mathcal{G}_{\boldsymbol{\mu}_1}(\boldsymbol{\eta}), \Sigma_\nu) N(\mathbf{m}_\eta, \Sigma_\eta) d\boldsymbol{\eta} \\
& \cdot \ln \left( \frac{\int N(\mathbf{S}_{\boldsymbol{\theta}_1} \odot \mathcal{G}_{\boldsymbol{\mu}_1}(\boldsymbol{\eta}), \Sigma_\nu) N(\mathbf{S}_{\boldsymbol{\theta}_2} \odot \mathcal{G}_{\boldsymbol{\mu}_2}(\boldsymbol{\eta}), \Sigma_\nu) N(\mathbf{m}_\eta, \Sigma_\eta) d\boldsymbol{\eta}}{\int N(\mathbf{S}_{\boldsymbol{\theta}_2} \odot \mathcal{G}_{\boldsymbol{\mu}_2}(\boldsymbol{\eta}), \Sigma_\nu) N(\mathbf{m}_\eta, \Sigma_\eta) d\boldsymbol{\eta}} \right) dz d\mathbf{d} \\
& + \frac{1}{2} \ln(\sigma_\nu^{-2N_{\text{acq},z}}).
\end{aligned}$$

(C.66)



# Bibliography

- [1] J. B M Warntjes, O. Dahlqvist Leinhard, J. West, and P. Lundberg. Rapid magnetic resonance quantification on the brain: Optimization for clinical usage. *Magnetic Resonance in Medicine*, 60(2):320–329, 2008.
- [2] S J Riederer, S A Suddarth, S A Bobman, J N Lee, H Z Wang, and J R MacFall. Automated MR image synthesis: feasibility studies. *Radiology*, 153(1):203–6, 1984.
- [3] D A Ortendahl, N Hylton, L Kaufman, J C Watts, L E Crooks, C M Mills, and D D Stark. Analytical tools for magnetic resonance imaging. *Radiology*, 153(2):479–88, 1984.
- [4] Hideto Iwaoka, Takaaki Hirata, and Hiroyuki Matsuyara. Optimal Pulse Sequences for Magnetic Resonance Imaging—Computing Accurate T1, T2, and Proton Density Images. *IEEE Transactions on Medical Imaging*, 6(4):360–369, 1987.
- [5] D A Feinberg, C M Mills, J P Posin, D A Ortendahl, N M Hylton, L E Crooks, J C Watts, L Kaufman, M Arakawa, and J C Hoenninger. Multiple spin-echo magnetic resonance imaging. *Radiology*, 155(2):437–442, 1985.
- [6] R. Madankan, W. Stefan, S. J. Fahrenholtz, C. J. Maclellan, J. D. Hazle, R. J. Stafford, J. S. Weinberg, G. Rao, and D. Fuentes. Accelerated magnetic resonance

- thermometry in the presence of uncertainties. *Physics in Medicine and Biology*, 62(1):214–245, 2017.
- [7] Carlo Pierpaoli. Quantitative Brain MRI. *Top Magn Reson Imaging*, 21(2):3–4, 2010.
- [8] Akifumi Hagiwara, Marcel Warntjes, Masaaki Hori, Christina Andica, Misaki Nakazawa, Kanako Kunishima Kumamaru, Osamu Abe, and Shigeki Aoki. SyMRI of the brain. *Investigative Radiology*, 52(10):647–657, 2017.
- [9] L. E. Drain. A direct method of measuring nuclear spin-lattice relaxation times. *Proceedings of the Physical Society. Section A*, 62(5):301–306, 1949.
- [10] Kenneth P. Whittall, Alex L. MacKay, Douglas A. Graeb, Robert A. Nugent, David K B Li, and Donald W. Paty. In vivo measurement of T2 distributions and water contents in normal human brain. *Magnetic Resonance in Medicine*, 37(1):34–43, 1997.
- [11] Elizabeth Henderson, Graeme McKinnon, Ting Yim Lee, and Brian K. Rutt. A fast 3D look-locker method for volumetric T1mapping. *Magnetic Resonance Imaging*, 17(8):1163–1171, 1999.
- [12] Charles A. McKenzie, Zuoqun Chen, Dick J. Drost, and Frank S. Prato. Fast acquisition of quantitative T2 maps. *Magnetic Resonance in Medicine*, 41(1):208–212, 1999.
- [13] Stuart Clare and Peter Jezzard. Rapid T1 mapping using multislice echo planar imaging. *Magnetic Resonance in Medicine*, 45(4):630–634, 2001.

- [14] Cynthia F. Maier, Steve G. Tan, Hari Hariharan, and Hollis G. Potter. T2 quantitation of articular cartilage at 1.5 T. *Journal of Magnetic Resonance Imaging*, 17(3):358–364, 2003.
- [15] Mark Westwood, Lisa J. Anderson, David N. Firmin, Peter D. Gatehouse, Clare C. Charrier, Beatrix Wonke, and Dudley J. Pennell. A single breath-hold multiecho T2\* cardiovascular magnetic resonance technique for diagnosis of myocardial iron overload. *Journal of Magnetic Resonance Imaging*, 18(1):33–39, 2003.
- [16] Hannes Dahnke and T. Schaeffter. Limits of detection of SPIO at 3.0 T using T2\* relaxometry. *Magnetic Resonance in Medicine*, 53(5):1202–1206, 2005.
- [17] Ralf Deichmann. Fast high-resolution T1 mapping of the human brain. *Magnetic Resonance in Medicine*, 54(1):20–27, 2005.
- [18] Sean C.L. Deoni, Terry M. Peters, and Brian K. Rutt. High-resolution T1 and T2 mapping of the brain in a clinically acceptable time with DESPOT1 and DESPOT2. *Magnetic Resonance in Medicine*, 53(1):237–241, 2005.
- [19] David C. Zhu and Richard D. Penn. Full-brain T1 mapping through inversion recovery fast spin echo imaging with time-efficient slice ordering. *Magnetic Resonance in Medicine*, 54(3):725–731, 2005.
- [20] H. Neeb, K. Zilles, and N. J. Shah. A new method for fast quantitative mapping of absolute water content in vivo. *NeuroImage*, 31(3):1156–1168, 2006.
- [21] Alessia Pepe, Vincenzo Positano, Maria Filomena Santarelli, Fortunato Sorrentino, Eliana Cracolici, Daniele De Marchi, Aurelio Maggio, Massimo Midiri, Luigi Landini, and Massimo Lombardi. Multislice multiecho T2\* cardiovascular magnetic

- resonance for detection of the heterogeneous distribution of myocardial iron overload. *Journal of Magnetic Resonance Imaging*, 23(5):662–668, 2006.
- [22] Rajesh Kumar, Sean Delshad, Mary A. Woo, Paul M. MacEy, and Ronald M. Harper. Age-related regional brain T2-relaxation changes in healthy adults. *Journal of Magnetic Resonance Imaging*, 35(2):300–308, 2012.
- [23] Wolfgang Krauss, Martin Gunnarsson, Torbjörn Andersson, and Per Thunberg. Accuracy and reproducibility of a quantitative magnetic resonance imaging method for concurrent measurements of tissue relaxation times and proton density. *Magnetic Resonance Imaging*, 33(5):584–591, 2015.
- [24] Dan Ma, Vikas Gulani, Nicole Seiberlich, Kecheng Liu, Jeffrey L. Sunshine, Jeffrey L. Duerk, and Mark A. Griswold. Magnetic resonance fingerprinting. *Nature*, 495(7440):187–192, 2013.
- [25] Philipp Ehses, Nicole Seiberlich, Dan Ma, Felix A. Breuer, Peter M. Jakob, Mark A. Griswold, and Vikas Gulani. IR TrueFISP with a golden-ratio-based radial readout: Fast quantification of T1, T2, and proton density. *Magnetic Resonance in Medicine*, 69(1):71–81, 2013.
- [26] Sampada Bhawe, Sajan Goud Lingala, Casey P. Johnson, Vincent A. Magnotta, and Mathews Jacob. Accelerated whole-brain multi-parameter mapping using blind compressed sensing. *Magnetic Resonance in Medicine*, 75(3):1175–1186, 2016.
- [27] Nikolaus Weiskopf, John Suckling, Guy Williams, Marta M. Correia M., Becky Inkster, Roger Tait, Cinly Ooi, Edward T. Bullmore T., and Antoine Lutti. Quantitative multi-parameter mapping of R1, PD\*, MT, and R2\* at 3T: A multi-center validation. *Frontiers in Neuroscience*, 7(7 JUN):1–11, 2013.

- [28] Antoine Lutti, Joerg Stadler, Oliver Josephs, Christian Windischberger, Oliver Speck, Johannes Bernarding, Chloe Hutton, and Nikolaus Weiskopf. Robust and fast whole brain mapping of the RF transmit field B1 at 7T. *PLoS ONE*, 7(3):1–7, 2012.
- [29] J. B.M. Warntjes, O. Dahlqvist, and P. Lundberg. Novel method for rapid, simultaneous T1, T2\*, and proton density quantification. *Magnetic Resonance in Medicine*, 57(3):528–537, 2007.
- [30] Ewa E. Odrobina, T. Y J Lam, Teresa Pun, Rajiv Midha, and Greg J. Stanisz. MR properties of excised neural tissue following experimentally induced demyelination. *NMR in Biomedicine*, 18(5):277–284, 2005.
- [31] Christian Labadie, Jing Huei Lee, William D. Rooney, Silvia Jarchow, Monique Aubert-Frécon, Charles S. Springer, and Harald E. Möller. Myelin water mapping by spatially regularized longitudinal relaxographic imaging at high magnetic fields. *Magnetic Resonance in Medicine*, 71(1):375–387, 2014.
- [32] S Kvernby, M J Warntjes, H Haraldsson, C J Carlhall, J Engvall, and T Ebbers. Simultaneous three-dimensional myocardial T1 and T2 mapping in one breath hold with 3D-QALAS. *J Cardiovasc Magn Reson*, 16:102, 2014.
- [33] Àlex Rovira and Adelaida León. MR in the diagnosis and monitoring of multiple sclerosis: An overview. *European Journal of Radiology*, 67(3):409–414, 2008.
- [34] Andrew Simmons, Eric Westman, Sebastian Muehlboeck, Patrizia Mecocci, Bruno Vellas, Magda Tsolaki, Iwona Kloszewska, Lars Olof Wahlund, Hilka Soininen, Simon Lovestone, Alan Evans, and Christian Spenger. MRI measures of alzheimer’s disease and the addneuromed study. *Annals of the New York Academy of Sciences*, 1180(0):47–55, 2009.

- [35] Geon Ha Kim, Jae Hong Lee, Sang Won Seo, Jeong Hun Kim, Joon Kyung Seong, Byoung Seok Ye, Hanna Cho, Young Noh, Hee Jin Kim, Cindy W. Yoon, Seung Jun Oh, Jae Seung Kim, Yearn Seong Choe, Kyung Han Lee, Sung Tae Kim, Jung Won Hwang, Jee Hyang Jeong, and Duk L. Na. Hippocampal volume and shape in pure subcortical vascular dementia. *Neurobiology of Aging*, 36(1):485–491, 2015.
- [36] K. Ambarki, T. Lindqvist, A. Wahlin, E. Petterson, M.J.B. Warntjes, R. Birgander, J. Malm, and A. Eklund. Evaluation of Automatic Measurement of the Intracranial Volume Based on Quantitative MR Imaging. *American Journal of Neuroradiology*, 33(10):1951–1956, 2012.
- [37] T. Granberg, Y. Forslin, L.E. Nordin, F. Hashim, S. Shams, M. Kristoffersen-Wiberg, J. Berglund, S. Fredrikson, P. Aspelin, C. Cananau, and M. Uppman. Clinical Feasibility of Synthetic MRI in Multiple Sclerosis: A Diagnostic and Volumetric Validation Study. *American Journal of Neuroradiology*, 37(6):1023–1029, 2016.
- [38] David W. Shattuck, Stephanie R. Sandor-Leahy, Kirt A. Schaper, David A. Rotenberg, and Richard M. Leahy. Magnetic resonance image tissue classification using a partial volume model. *NeuroImage*, 13(5):856–876, 2001.
- [39] Anders H. Andersen, Zhiming Zhang, Malcolm J. Avison, and Don M. Gash. Automated segmentation of multispectral brain MR images. *Journal of Neuroscience Methods*, 122(1):13–23, 2002.
- [40] Artem Mikheev, Gregory Nevsky, Siddharth Govindan, Robert Grossman, and Henry Rusinek. Fully Automatic Segmentation of the Brain From T1-Weighted. *Journal of magnetic resonance imaging : JMRI*, 27(6), 2008.

- [41] H.S. Choi, D.R. Haynor, and Y. Kim. Partial volume tissue classification of multi-channel magnetic resonance images-a mixel model. *IEEE Transactions on Medical Imaging*, 10(3):395–407, 1991.
- [42] J. West, J. B.M. Warntjes, and P. Lundberg. Novel whole brain segmentation and volume estimation using quantitative MRI. *European Radiology*, 22(5):998–1007, 2012.
- [43] J. B.M. Warntjes, A. Tisell, A. M. Landtblom, and P. Lundberg. Effects of gadolinium contrast agent administration on automatic brain tissue classification of patients with multiple sclerosis. *American Journal of Neuroradiology*, 35(7):1330–1336, 2014.
- [44] Lynn M. Fletcher, John B. Barsotti, and Joseph P. Hornak. A multispectral analysis of brain tissues. *Magnetic Resonance in Medicine*, 29(5):623–630, 1993.
- [45] Adolf Pfefferbaum, Torsten Rohlfing, Margaret J. Rosenbloom, and Edith V. Sullivan. Combining atlas-based parcellation of regional brain data acquired across scanners at 1.5T and 3.0T field strengths. *NeuroImage*, 60(2):940–951, 2012.
- [46] Janne West, Ida Blystad, Maria Engström, Jan B.M. Warntjes, and Peter Lundberg. Application of Quantitative MRI for Brain Tissue Segmentation at 1.5 T and 3.0 T Field Strengths. *PLoS ONE*, 8(9):1–12, 2013.
- [47] Aline Hocq, Nicolas Brouette, Sven Saussez, Michel Luhmer, Pierre Gillis, and Yves Gossuin. Variable-field relaxometry of iron-containing human tissues: A preliminary study. *Contrast Media and Molecular Imaging*, 4(4):157–164, 2009.
- [48] Hollie West, James L. Leach, Blaise V. Jones, Marguerite Care, Rupa Radhakrishnan, Arnold C. Merrow, Enrique Alvarado, and Suraj D. Serai. Clinical validation

- of synthetic brain MRI in children: initial experience. *Neuroradiology*, 59(1):43–50, 2017.
- [49] I. Blystad, J. B.M. Warntjes, O. Smedby, A. M. Landtblom, P. Lundberg, and E. M. Larsson. Synthetic MRI of the brain in a clinical setting. *Acta Radiologica*, 53(10):1158–1163, 2012.
- [50] Akifumi Hagiwara, Masaaki Hori, Michimasa Suzuki, Christina Andica, Misaki Nakazawa, Kouhei Tsuruta, Nao Takano, Shuji Sato, Nozomi Hamasaki, Mariko Yoshida, Kanako Kunishima Kumamaru, Kuni Ohtomo, and Shigeki Aoki. Contrast-enhanced synthetic MRI for the detection of brain metastases. *Acta Radiologica Open*, 5(2):205846011562675, 2016.
- [51] X A Hagiwara, X M Hori, X K Yokoyama, X M Y Takemura, X C Andica, X T Tabata, X K Kamagata, and X M Suzuki. Synthetic MRI in the Detection of Multiple Sclerosis Plaques. pages 1–7, 2017.
- [52] Christina Andica, Akifumi Hagiwara, Misaki Nakazawa, Kanako K Kumamaru, Masaaki Hori, Mitsuru Ikeno, Toshiaki Shimizu, and Shigeki Aoki. Synthetic MR Imaging in the Diagnosis of Bacterial Meningitis. *Magnetic Resonance in Medical Sciences*, 16(2):91–92, 2017.
- [53] Aaron M. Betts, James L. Leach, Blaise V. Jones, Bin Zhang, and Suraj Serai. Brain imaging with synthetic MR in children: clinical quality assessment. *Neuroradiology*, 58(10):1017–1026, 2016.
- [54] Ranjan Maitra and John J. Riddles. Synthetic magnetic resonance imaging revisited. *IEEE Transactions on medical imaging*, 29(3):895–902, 2010.



- [55] Stephen W. Provencher. Estimation of metabolite concentrations from localized in vivo proton NMR spectra. *Magnetic Resonance in Medicine*, 30(6):672–679, 1993.
- [56] Anders Tisell, Olof Dahlqvist Leinhard, Jan Bertus Marcel Warntjes, Anne Aalto, Örjan Smedby, Anne Marie Landtblom, and Peter Lundberg. Increased Concentrations of Glutamate and Glutamine in Normal-Appearing White Matter of Patients with Multiple Sclerosis and Normal MR Imaging Brain Scans. *PLoS ONE*, 8(4):1–8, 2013.
- [57] Zografos Caramanos, Sridar Narayanan, and Douglas L. Arnold. 1H-MRS quantification of tNA and tCr in patients with multiple sclerosis: A meta-analytic review. *Brain*, 128(11):2483–2506, 2005.
- [58] J. Hennig, H. Pfister, T. Ernst, and D. Ott. Direct Absolute Quantification of Metabolites in the Human Brain with. *Brain*, 5(August 1991):193–199, 1992.
- [59] S. F. Keevil, B. Barbiroli, J. C.W. Brooks, E. B. Cady, R. Canese, P. Carlier, D. J. Collins, P. Gilligan, G. Gobbi, J. Hennig, H. Kügel, M. O. Leach, D. Metzler, V. Mlynárik, E. Moser, M. C. Newbold, G. S. Payne, P. Ring, J. N. Roberts, I. J. Rowland, T. Thiel, I. Tkác, S. Topp, H. J. Wittsack, M. Wylezinska, P. Zaniol, O. Henriksen, and F. Podo. Absolute metabolite quantification by in vivo NMR spectroscopy: II. A multicentre trial of protocols for in vivo localised proton studies of human brain. *Magnetic Resonance Imaging*, 16(9):1093–1106, 1998.
- [60] A. Tisell, O. Dahlqvist Leinhard, J. B.M. Warntjes, and P. Lundberg. Procedure for quantitative 1H magnetic resonance spectroscopy and tissue characterization of human brain tissue based on the use of quantitative magnetic resonance imaging. *Magnetic Resonance in Medicine*, 70(4):905–915, 2013.

- [61] Eva Alonso-Ortiz, Ives R. Levesque, and G. Bruce Pike. MRI-based myelin water imaging: A technical review. *Magnetic Resonance in Medicine*, 73(1):70–81, 2015.
- [62] David Glasser, Matthew; Van Essen. NIH Public Access Mapping Human Cortical Areas in vivo Based on Myelin Content as Revealed by T1- and T2-weighted MRI. *J Neurosci*, 31(32):11597–11616, 2012.
- [63] Marcel Warntjes, Maria Engström, Anders Tisell, and Peter Lundberg. Modeling the presence of myelin and edema in the brain based on multi-parametric quantitative MRI. *Frontiers in Neurology*, 7(FEB):1–15, 2016.
- [64] Jan B M Warntjes, Maria Engström, Anders Tisell, and Peter Lundberg. Brain Characterization Using Normalized Quantitative Magnetic Resonance Imaging. *PLoS ONE*, 8(8), 2013.
- [65] Bruno Alfano, Arturo Brunetti, Eugenio M. Covelli, Mario Quarantelli, Maria Rosaria Panico, Andrea Ciarmiello, and Marco Salvatore. Unsupervised, automated segmentation of the normal brain using a multispectral relaxometric magnetic resonance approach. *Magnetic Resonance in Medicine*, 37(1):84–93, 1997.
- [66] Robert K. Breger, Felix W. Wehrli, H. Cecil Charles, James R. MacFall, and Victor M. Haughton. Reproducibility of relaxation and spin-density parameters in phantoms and the human brain measured by MR imaging at 1.5T. *Magnetic Resonance in Medicine*, 3(5):649–662, 1986.
- [67] Joonmi Oh, Soonmee Cha, Ashley H. Aiken, Eric T. Han, Jason C. Crane, Jeffrey A. Stainsby, Graham A. Wright, William P. Dillon, and Sarah J. Nelson.

- Quantitative apparent diffusion coefficients and T2 relaxation times in characterizing contrast enhancing brain tumors and regions of peritumoral edema. *Journal of Magnetic Resonance Imaging*, 21(6):701–708, 2005.
- [68] Y. F. Qian, C. L. Yu, C. Zhang, and Yong Qiang Yu. MR T1-weighted inversion recovery imaging in detecting brain metastases: Could it replace T1-weighted spin-echo imaging? *American Journal of Neuroradiology*, 29(4):701–704, 2008.
- [69] Khader M Hasan, Indika S Walimuni, Humaira Abid, Sushmita Datta, Jerry S Wolinsky, and Ponnada A Narayana. Human brain atlas-based multimodal MRI analysis of volumetry, diffusimetry, relaxometry and lesion distribution in multiple sclerosis patients and healthy adult controls: implications for understanding the pathogenesis of multiple sclerosis and consolidat. *Journal of the neurological sciences*, 313(1-2):99–109, 2012.
- [70] Mohit Neema, James Stankiewicz, Ashish Arora, Venkata S.R. Dandamudi, Courtney E. Batt, Zachary D. Guss, Ali Al-Sabbagh, and Rohit Bakshi. T1- and T2-based MRI measures of diffuse gray matter and white matter damage in patients with multiple sclerosis. *Journal of Neuroimaging*, 17(SUPPL. 1):16–21, 2007.
- [71] H. B.W. Larsson, J. Frederiksen, J. Petersen, A. Nordenbo, I. Zeeberg, O. Henriksen, and J. Olesen. Assessment of demyelination, edema, and gliosis by in vivo determination of T1 and T2 in the brain of patients with acute attack of multiple sclerosis. *Magnetic Resonance in Medicine*, 11(3):337–348, 1989.
- [72] H. B.W. Larsson, J. Frederiksen, L. Kjær, O. Henriksen, and J. Olesen. In vivo determination of T1 and T2 in the brain of patients with severe but stable multiple sclerosis. *Magnetic Resonance in Medicine*, 7(1):43–55, 1988.

- [73] L. Bonzano, L. Roccatagliata, G. L. Mancardi, and Maria Pia Sormani. Gadolinium-enhancing or active T2 magnetic resonance imaging lesions in multiple sclerosis clinical trials? *Multiple Sclerosis*, 15(9):1043–1047, 2009.
- [74] Flavia Nelson, A. H. Poonawalla, P. Hou, F. Huang, J. S. Wolinsky, and P. A. Narayana. Improved identification of intracortical lesions in multiple sclerosis with phase-sensitive inversion recovery in combination with fast double inversion recovery MR imaging. *American Journal of Neuroradiology*, 28(9):1645–1649, 2007.
- [75] Barry J. Bedell and Ponnada A. Narayana. Implementation and evaluation of a new pulse sequence for rapid acquisition of double inversion recovery images for simultaneous suppression of white matter and CSF. *Journal of Magnetic Resonance Imaging*, 8(3):544–547, 1998.
- [76] Ping Hou, Khader M. Hasan, Clark W. Sitton, Jerry S. Wolinsky, and Ponnada A. Narayana. Phase-sensitive T1 inversion recovery imaging: A time-efficient interleaved technique for improved tissue contrast in neuroimaging. *American Journal of Neuroradiology*, 26(6):1432–1438, 2005.
- [77] Maria Engström, Jan B.M. Warntjes, Anders Tisell, Anne Marie Landtblom, and Peter Lundberg. Multi-parametric representation of voxel-based quantitative magnetic resonance imaging. *PLoS ONE*, 9(11), 2014.
- [78] R. A. Rudick, E. Fisher, J.-C. Lee, J. Simon, and L. Jacobs. Use of the brain parenchymal fraction to measure whole brain atrophy in relapsing-remitting MS. *Neurology*, 53(8):1698–1698, 1999.
- [79] Jan Kassubek, Hayrettin Tumani, Daniel Ecker, Anja Kurt, Albert C. Ludolph, and Freimut D. Juengling. Age-related brain parenchymal fraction is significantly

- decreased in young multiple sclerosis patients: A quantitative MRI study. *NeuroReport*, 14(3):427–430, 2003.
- [80] Hugo Vrenken, Dirk L Knol, L Noor van Dijk, Vincenzo Dattola, Bas Jasperse, Ronald A van Schijndel, Chris H Polman, Jonas A Castelijns, Frederik Barkhof, Petra J Pouwels, H Vrenken, J J Geurts, D L Knol, L N van Dijk, V Dattola, B Jasperse, R A van Schijndel, C H Polman, J A Castelijns, F Barkhof, P J Pouwels, and Jeroen J Geurts. Whole-Brain T1 Mapping in Multiple Sclerosis: Global Changes of Normal-appearing Gray and White Matter. *Radiology*, 240(3):1–10, 2006.
- [81] Andrea Bernasconi, Neda Bernasconi, Zografos Caramanos, David C. Reutens, Frederick Andermann, François Dubeau, Donatella Tampieri, Bruce G. Pike, and Douglas L. Arnold. T2 relaxometry can lateralize mesial temporal lobe epilepsy in patients with normal MRI. *NeuroImage*, 2000.
- [82] J Vymazal, a Righini, R a Brooks, M Canesi, C Mariani, M Leonardi, and G Pezzoli. T1 and T2 in the brain of healthy subjects, patients with Parkinson disease, and patients with multiple system atrophy: relation to iron content. *Radiology*, 211(2):489–495, 1999.
- [83] Saman Hazany, John L. Go, and Meng Law. Magnetic resonance imaging of infectious meningitis and ventriculitis in adults. *Topics in Magnetic Resonance Imaging*, 23(5):315–325, 2014.
- [84] Alessandra Splendiani, Edoardo Puglielli, Rosanna De Amicis, Stefano Necozone, Carlo Masciocchi, and Massimo Gallucci. Contrast-enhanced FLAIR in the early diagnosis of infectious meningitis. *Neuroradiology*, 47(8):591–598, 2005.

- [85] Christina Andica, Akifumi Hagiwara, Misaki Nakazawa, Kouhei Tsuruta, Nao Takano, Masaaki Hori, Hiroharu Suzuki, Hidenori Sugano, Hajime Arai, and Shigeki Aoki. The Advantage of Synthetic MRI for the Visualization of Early White Matter Change in an Infant with Sturge-Weber Syndrome. *Magnetic Resonance in Medical Sciences*, 15(4):347–348, 2016.
- [86] Efthymia Alexopoulou, Fotini Stripeli, Panagiotis Baras, Ioannis Seimenis, Antonis Kattamis, Vasilis Ladis, Efstathios Efstathopoulos, Elias N. Brountzos, Alexis D. Kelekis, and Nikolaos L. Kelekis. R2 Relaxometry With MRI for the Quantification of Tissue Iron Overload in Beta-Thalassemic Patients. *Magnetic Resonance Imaging*, 23:163–170, 2006.
- [87] L. J. Anderson, S. Holden, B. Davis, E. Prescott, C. C. Charrier, N. H. Bunce, D. N. Firmin, B. Wonke, J. Porter, J. M. Walker, and D. J. Pennell. Cardiovascular T2-star ( $T2^*$ ) magnetic resonance for the early diagnosis of myocardial iron overload. *European Heart Journal*, 22(23):2171–2179, 2001.
- [88] Kristin A. Thomas-Sohl, Dale F. Vaslow, and Bernard L. Maria. Sturge-Weber syndrome: A review. *Pediatric Neurology*, 30(5):303–310, 2004.
- [89] C. Di Rocco and G. Tamburrini. Sturge-Weber syndrome. *Child's Nervous System*, 22:909–921, 2006.
- [90] Evans. Encephalographic Abnormality. *Arch Neurol Psychiatry*, 47:931–937, 1942.
- [91] Ullrich Meier, S. Paris, A. Gräwe, D. Stockheim, A. Hajdukova, and S. Mutze. Is there a correlation between operative results and change in ventricular volume after shunt placement? A study of 60 cases of idiopathic normal-pressure hydrocephalus. *Neuroradiology*, 45(6):377–380, 2003.

- [92] J. Virhammar, M. Warntjes, K. Laurell, and E. M. Larsson. Quantitative MRI for rapid and user-independent monitoring of intracranial CSF volume in hydrocephalus. *American Journal of Neuroradiology*, 37(5):797–801, 2016.
- [93] Janne West, Anne Aalto, Anders Tisell, Olof Dahlqvist Leinhard, Anne Marie Landtblom, Örjan Smedby, and Peter Lundberg. Normal appearing and diffusely abnormal white matter in patients with multiple sclerosis assessed with quantitative MR. *PLoS ONE*, 9(4):1–9, 2014.
- [94] A. Hagiwara, M. Hori, K. Yokoyama, M. Y. Takemura, C. Andica, K. K. Kumamaru, M. Nakazawa, N. Takano, H. Kawasaki, S. Sato, N. Hamasaki, A. Kunitatsu, and S. Aoki. Utility of a multiparametric quantitative MRI model that assesses myelin and edema for evaluating plaques, periplaque white matter, and normal-appearing white matter in patients with multiple sclerosis: A feasibility study. *American Journal of Neuroradiology*, 38(2):237–242, 2017.
- [95] Augusto Elias Mamere, L. A.L. Saraiva, A. L.M. Matos, A. A.O. Carneiro, and A. C. Santos. Evaluation of delayed neuronal and axonal damage secondary to moderate and severe traumatic brain injury using quantitative MR imaging techniques. *American Journal of Neuroradiology*, 30(5):947–952, 2009.
- [96] C. Wagner-Manslau, P. Lukas, M. Herzog, R. Kau, and K. Beckers. MRI and proton-NMR relaxation times in diagnosis and therapeutic monitoring of squamous cell carcinoma. *European Radiology*, 4(4):314–323, 1994.
- [97] Chun Yuan and William S. Kerwin. MRI of atherosclerosis. *Journal of Magnetic Resonance Imaging*, 19(6):710–719, 2004.
- [98] Charles Gasparovic, Tao Song, Deidre Devier, H. Jeremy Bockholt, Arvind Caprihan, Paul G. Mullins, Stefan Posse, Rex E. Jung, and Leslie A. Morrison. Use of

- tissue water as a concentration reference for proton spectroscopic imaging. *Magnetic Resonance in Medicine*, 55(6):1219–1226, 2006.
- [99] Klaus Poulsen and Jørn Simonsen. Computed tomography as routine in connection with medico-legal autopsies. *Forensic Science International*, 171(2-3):190–197, 2007.
- [100] Christian Jackowski, Marcel J.B. Warntjes, Johan Kihlberg, Johan Berge, Michael J. Thali, and Anders Persson. Quantitative MRI in Isotropic Spatial Resolution for Forensic Soft Tissue Documentation. Why and How? *Journal of Forensic Sciences*, 56(1):208–215, 2011.
- [101] Seiji Shiotani, Mototsugu Kohno, Noriyoshi Ohashi, Kentaro Yamazaki, Hidetsugu Nakayama, Ko Watanabe, Yuji Oyake, and Yuji Itai. Non-traumatic postmortem computed tomographic (PMCT) findings of the lung. *Forensic Science International*, 139(1):39–48, 2004.
- [102] Wolf Dieter Zech, Anna Lena Hottinger, Nicole Schwendener, Frederick Schuster, Anders Persson, Marcel J. Warntjes, and Christian Jackowski. Post-mortem 1.5T MR quantification of regular anatomical brain structures. *International Journal of Legal Medicine*, 130(4):1071–1080, 2016.
- [103] Wolf Dieter Zech, Nicole Schwendener, Anders Persson, Marcel J. Warntjes, and Christian Jackowski. Temperature dependence of postmortem MR quantification for soft tissue discrimination. *European Radiology*, 25(8):2381–2389, 2015.
- [104] Wolf Dieter Zech, Nicole Schwendener, Anders Persson, Marcel J. Warntjes, Fabiano Riva, Frederick Schuster, and Christian Jackowski. Postmortem quantitative 1.5-T MRI for the differentiation and characterization of serous fluids, blood, CSF,



- and putrefied CSF. *International Journal of Legal Medicine*, 129(5):1127–1136, 2015.
- [105] Nicole Schwendener, Christian Jackowski, Anders Persson, Marcel J. Warntjes, Frederick Schuster, Fabiano Riva, and Wolf Dieter Zech. Detection and differentiation of early acute and following age stages of myocardial infarction with quantitative post-mortem cardiac 1.5 T MR. *Forensic Science International*, 270:248–254, 2017.
- [106] Andrew J. Buckler, Linda Bresolin, N. Reed Dunnick, Daniel C. Sullivan, and For the Group. A Collaborative Enterprise for Multi-Stakeholder Participation in the Advancement of Quantitative Imaging. *Radiology*, 258(3):906–914, 2011.
- [107] Andrew J. Buckler, Linda Bresolin, N. Reed Dunnick, and Daniel C. Sullivan. Quantitative Imaging Test Approval and Biomarker Qualification : Interrelated but Distinct Activities. *Radiology*, 259(3):875–884, 2011.
- [108] Larry G Kessler, Huiman X Barnhart, Andrew J Buckler, Kingshuk Roy Choudhury, Marina V Kondratovich, Alicia Toledano, Alexander R Guimaraes, Ross Filice, Zheng Zhang, and Daniel C Sullivan. The emerging science of quantitative imaging biomarkers terminology and definitions for scientific studies and regulatory submissions. *Statistical Methods in Medical Research*, 24(1):9–26, 2015.
- [109] Erich P Huang, Xiao-Feng Wang, Kingshuk Roy Choudhury, Lisa M McShane, Mithat Gönen, Jingjing Ye, Andrew J Buckler, Paul E Kinahan, Anthony P Reeves, Edward F Jackson, Alexander R Guimaraes, and Gudrun Zahlmann. Meta-analysis of the technical performance of an imaging procedure: Guidelines and statistical methodology. *Statistical Methods in Medical Research*, 24(1):141–174, 2015.

- [110] Nancy A Obuchowski, Huiman X Barnhart, Andrew J Buckler, Gene Pennello, Xiao-Feng Wang, Jayashree Kalpathy-Cramer, Hyun J (Grace) Kim, and Anthony P Reeves. Statistical issues in the comparison of quantitative imaging biomarker algorithms using pulmonary nodule volume as an example. *Statistical Methods in Medical Research*, 24(1):107–140, 2015.
- [111] David L Raunig, Lisa M McShane, Gene Pennello, Constantine Gatsonis, Paul L Carson, James T Voyvodic, Richard L Wahl, Brenda F Kurland, Adam J Schwarz, Mithat Gönen, Gudrun Zahlmann, Marina V Kondratovich, Kevin O'Donnell, Nicholas Petrick, Patricia E Cole, Brian Garra, and Daniel C Sullivan. Quantitative imaging biomarkers: A review of statistical methods for technical performance assessment. *Statistical Methods in Medical Research*, 24(1):27–67, 2015.
- [112] Daniel C. Sullivan, Nancy A. Obuchowski, Larry G. Kessler, David L. Raunig, Constantine Gatsonis, Erich P. Huang, Marina Kondratovich, Lisa M. McShane, Anthony P. Reeves, Daniel P. Barboriak, Alexander R. Guimaraes, and Richard L. Wahl. Metrology Standards for Quantitative Imaging Biomarkers. *Radiology*, 277(3):813–825, 2015.
- [113] Nancy a Obuchowski, Anthony P Reeves, Erich P Huang, Xiao-Feng Wang, Andrew J Buckler, Hyun J Grace Kim, Huiman X Barnhart, Edward F Jackson, Maryellen L Giger, Gene Pennello, Alicia Y Toledano, Jayashree Kalpathy-Cramer, Tatiyana V Apanasovich, Paul E Kinahan, Kyle J Myers, Dmitry B Goldgof, Daniel P Barboriak, Robert J Gillies, Lawrence H Schwartz, and Daniel C Sullivan. Quantitative imaging biomarkers: A review of statistical methods for computer algorithm comparisons. *Statistical methods in medical research*, 0(0):1–39, jun 2014.

- [114] Michael J. Paldino and Daniel P. Barboriak. Fundamentals of Quantitative Dynamic Contrast-Enhanced MR Imaging. *Magnetic Resonance Imaging Clinics of North America*, 17(2):277–289, 2009.
- [115] Wei Huang, Xin Li, Yiyi Chen, Xia Li, Ming-Ching Chang, Matthew J Oborski, Dariya I Malyarenko, Mark Muzi, Guido H Jajamovich, Andriy Fedorov, Alina Tudorica, Sandeep N Gupta, Charles M Laymon, Kenneth I Marro, Hadrien A Dyvorne, James V Miller, Daniel P Barbodiak, Thomas L Chenevert, Thomas E Yankeelov, James M Mountz, Paul E Kinahan, Ron Kikinis, Bachir Taouli, Fiona Fennessy, and Jayashree Kalpathy-Cramer. Variations of dynamic contrast-enhanced magnetic resonance imaging in evaluation of breast cancer therapy response: a multicenter data analysis challenge. *Translational oncology*, 7(1):153–66, 2014.
- [116] James T. Voyvodic. Reproducibility of single-subject fMRI language mapping with AMPLE normalization. *Journal of Magnetic Resonance Imaging*, 36(3):569–580, 2012.
- [117] Domenico Zacà, Jorge Jovicich, Sreenivasan R. Nadar, James T. Voyvodic, and Jay J. Pillai. Cerebrovascular reactivity mapping in patients with low grade gliomas undergoing presurgical sensorimotor mapping with BOLD fMRI. 40(2):383–390, 2014.
- [118] Kang Wang, Paul Manning, Nikolaus Szeverenyi, Tanya Wolfson, Gavin Hamilton, Michael S. Middleton, Florin Vaida, Meng Yin, Kevin Glaser, Richard L. Ehman, and Claude B. Sirlin. Repeatability and reproducibility of 2D and 3D hepatic MR elastography with rigid and flexible drivers at end-expiration and end-inspiration in healthy volunteers. *Abdominal Radiology*, 42(12):2843–2854, 2017.

- [119] P. D. Mozley, L. H. Schwartz, C. Bendtsen, B. Zhao, N. Petrick, and A. J. Buckler. Change in lung tumor volume as a biomarker of treatment response: A critical review of the evidence. *Annals of Oncology*, 21(9):1751–1755, 2010.
- [120] A. J. Buckler, L. H. Schwartz, N. Petrick, M. McNitt-Gray, B. Zhao, C. Fenimore, A. P. Reeves, P. D. Mozley, and R. S. Avila. Data sets for the qualification of volumetric CT as a quantitative imaging biomarker in lung cancer. *Optics Express*, 18(14):15267, 2010.
- [121] Andrew J. Buckler, P. David Mozley, Lawrence Schwartz, Nicholas Petrick, Michael McNitt-Gray, Charles Fenimore, Kevin O’Donnell, Wendy Hayes, Hyun J. Kim, Laurence Clarke, and Daniel Sullivan. Volumetric CT in Lung Cancer. *Academic Radiology*, 17(1):107–115, 2009.
- [122] Nicholas Petrick, Hyun J. Grace Kim, David Clunie, Kristin Borradaile, Robert Ford, Rongping Zeng, Marios A. Gavrielides, Michael F. McNitt-Gray, Z. Q. John Lu, Charles Fenimore, Binsheng Zhao, and Andrew J. Buckler. Comparison of 1D, 2D, and 3D Nodule Sizing Methods by Radiologists for Spherical and Complex Nodules on Thoracic CT Phantom Images. *Academic Radiology*, 21(1):30–40, 2014.
- [123] Baiyu Chen, Huiman Barnhart, Samuel Richard, James Colsher, Maxwell Amurao, and Ehsan Samei. Quantitative CT: Technique dependence of volume estimation on pulmonary nodules. *Physics in Medicine and Biology*, 57(5):1335–1348, 2012.
- [124] Marios A. Gavrielides, Qin Li, Rongping Zeng, Kyle J. Myers, Berkman Sahiner, and Nicholas Petrick. Minimum detectable change in lung nodule volume in a phantom ct study. *Academic Radiology*, 20(11):1364–1370, 2013.

- [125] Baiyu Chen, Huiman Barnhart, Samuel Richard, Marthony Robins, James Colsher, and Ehsan Samei. Volumetric quantification of lung nodules in CT with iterative reconstruction (ASiR and MBIR). *Medical Physics*, 40(11):111902, 2013.
- [126] A. Rodriguez, F. N. Ranallo, P. F. Judy, D. S. Gierada, and S. B. Fain. CT reconstruction techniques for improved accuracy of lung CT airway measurement. *Medical Physics*, 41(11), 2014.
- [127] James L. Mulshine, David S. Gierada, Samuel G. Armato, Rick S. Avila, David F. Yankelevitz, Ella A. Kazerooni, Michael F. McNitt-Gray, Andrew J. Buckler, and Daniel C. Sullivan. Role of the Quantitative Imaging Biomarker Alliance in optimizing CT for the evaluation of lung cancer screen-detected nodules. *Journal of the American College of Radiology*, 12(4):390–395, 2015.
- [128] Maria Athelougou, Hyun J. Kim, Alden Dima, Nancy Obuchowski, Adele Peskin, Marios A. Gavrielides, Nicholas Petrick, Ganesh Saiprasad, Dirk Colditz Colditz, Hubert Beaumont, Estanislao Oubel, Yongqiang Tan, Binsheng Zhao, Jan Martin Kuhnigk, Jan Hendrik Moltz, Guillaume Orioux, Robert J. Gillies, Yuhua Gu, Ninad Mantri, Gregory Goldmacher, Luduan Zhang, Emilio Vega, Michael Bloom, Rudresh Jarecha, Grzegorz Soza, Christian Tietjen, Tomoyuki Takeguchi, Hitoshi Yamagata, Sam Peterson, Osama Masoud, and Andrew J. Buckler. Algorithm Variability in the Estimation of Lung Nodule Volume From Phantom CT Scans: Results of the QIBA 3A Public Challenge. *Academic Radiology*, 23(8):940–952, 2016.
- [129] Andrew J. Buckler, Jovanna Danagoulian, Kjell Johnson, Adele Peskin, Marios A. Gavrielides, Nicholas Petrick, Nancy A. Obuchowski, Hubert Beaumont, Lubomir

- Hadjiiski, Rudresh Jarecha, Jan-Martin Kuhnigk, Ninad Mantri, Michael McNitt-Gray, Jan Hendrik Moltz, Gergely Nyiri, Sam Peterson, Pierre Tervé, Christian Tietjen, Etienne von Lavante, Xiaonan Ma, Samantha St. Pierre, and Maria Athelougou. Inter-method Performance Study of Tumor Volumetry Assessment on Computed Tomography Test-retest Data. *Academic Radiology*, 22(11):1393–1408, 2015.
- [130] Elske Quak, Pierre Yves Le Roux, Michael S. Hofman, Philippe Robin, David Bourhis, Jason Callahan, David Binns, Cédric Desmots, Pierre Yves Salaun, Rodney J. Hicks, and Nicolas Aide. Harmonizing FDG PET quantification while maintaining optimal lesion detection: prospective multicentre validation in 517 oncology patients. *European Journal of Nuclear Medicine and Molecular Imaging*, 42(13):2072–2082, 2015.
- [131] Vadim S. Koshkin, Vanessa Bolejack, Lawrence H. Schwartz, Richard L. Wahl, Rashmi Chugh, Denise K. Reinke, Binsheng Zhao, H. O. Joo, Shreyaskumar R. Patel, Scott M. Schuetze, and Laurence H. Baker. Assessment of imaging modalities and response metrics in Ewing sarcoma: Correlation with survival. *Journal of Clinical Oncology*, 34(30):3680–3685, 2016.
- [132] Richard Frank. Quantitative Imaging Biomarkers Alliance FDG–PET/CT Working Group Report. *Molecular Imaging and Biology*, 10(6):305–305, 2008.
- [133] A. J. Buckler and R. Boellaard. Standardization of Quantitative Imaging: The Time Is Right, and  $^{18}\text{F}$ -FDG PET/CT Is a Good Place to Start. *Journal of Nuclear Medicine*, 52(2):171–172, 2011.

- [134] Larry A. Pierce, Brian F. Elston, David A. Clunie, Dennis Nelson, and Paul E. Kinahan. A Digital Reference Object to Analyze Calculation Accuracy of PET Standardized Uptake Value. *Radiology*, 277(2):538–545, 2015.
- [135] Syahir Mansor, Elisabeth Pfaehler, Dennis Heijtel, Martin A. Lodge, Ronald Boellaard, and Maqsood Yaqub. Impact of PET/CT system, reconstruction protocol, data analysis method, and repositioning on PET/CT precision: An experimental evaluation using an oncology and brain phantom: An. *Medical Physics*, 44(12):6413–6424, 2017.
- [136] Michael M. Graham, Richard L. Wahl, John M. Hoffman, Jeffrey T. Yap, John J. Sunderland, Ronald Boellaard, Eric S. Perlman, Paul E. Kinahan, Paul E. Christian, Otto S. Hoekstra, and Gary S. Dorfman. Summary of the UPICT Protocol for 18F-FDG PET/CT Imaging in Oncology Clinical Trials. 56(6):955–961, 2015.
- [137] Joo Hyun O, Heather Jacene, Brandon Lubner, Hao Wang, Minh-Huy Huynh, Jeffrey P. Leal, and Richard L. Wahl. Quantitation of Cancer Treatment Response by 18 F-FDG PET/CT: Multicenter Assessment of Measurement Variability . *Journal of Nuclear Medicine*, 58(9):1429–1434, 2017.
- [138] Mark L. Palmeri, Bo Qiang, Shigao Chen, and Matthew W. Urban. Guidelines for finite-element modeling of acoustic radiation force-induced shear wave propagation in tissue-mimicking media. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 64(1):78–92, 2017.
- [139] Yufeng Deng, Ned C. Rouze, Mark L. Palmeri, and Kathryn R. Nightingale. On System Dependent Sources of Uncertainty and Bias in Ultrasonic Quantitative Shear Wave Imaging. 63(3):381–393, 2016.

- [140] Daniel Carl Sullivan, Lawrence H Schwartz, and Binsheng Zhao. The Imaging Viewpoint: How Imaging Affects Determination of Progression-free Survival. *19(10):2621–2628*, 2013.
- [141] Nancy A. Obuchowski. Interpreting Change in Quantitative Imaging Biomarkers. *Academic Radiology*, 25(3):372–379, 2018.
- [142] C. E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3):379–423, 1948.
- [143] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. 2005.
- [144] P. S. Maybeck and George M. Siouris. Stochastic Models, Estimation, and Control, Volume I, 1980.
- [145] Thomas Gerstner and Michael Griebel. Numerical integration using sparse grids. *Numerical Algorithms*, 18:209–232, 1998.
- [146] H. Niederreiter. *Random Number Generation and Monte Carlo Methods*. 1992.
- [147] I H Sloan, S Joe, and S.L.M.S. Joe. *Lattice Methods for Multiple Integration*. Oxford science publications. Clarendon Press, 1994.
- [148] Alan Genz. *A Package for Testing Multiple Integration Subroutines*, pages 337–340. Springer Netherlands, Dordrecht, 1987.
- [149] Paul van Dooren and Luc de Ridder. An adaptive algorithm for numerical integration over an n-dimensional cube. *Journal of Computational and Applied Mathematics*, 2(3):207–217, 1976.
- [150] Andrew R. Barron. Approximation and Estimation Bounds for Artificial Neural Networks. *Machine Learning*, 14(1):115–133, 1994.



- 
- [151] David M. Borth. A Total Entropy Criterion for the Dual Problem of Model Discrimination and Parameter Estimation. *Royal Statistical Society*, 1975.
- [152] Janne V. Kujala and Tuomas J. Lukka. Bayesian adaptive estimation: The next dimension. *Journal of Mathematical Psychology*, 50(4):369–389, 2006.
- [153] Jeremy Lewi, Robert Butera, and Liam Paninski. Sequential optimal design of neurophysiology experiments. *Neural Computation*, 21(3):619–687, 2009.
- [154] DR Daniel R Cavagnaro, JI Jay I Myung, Mark a Pitt, and Janne V JV Kujala. Adaptive design optimization: A mutual information-based approach to model discrimination in cognitive science. *Neural computation*, 22(1956):1–15, 2010.
- [155] Christopher C. Drovandi, James M. McGree, and Anthony N. Pettitt. A sequential Monte Carlo algorithm to incorporate model uncertainty in Bayesian sequential design. *Journal of Computational and Graphical Statistics*, 23(1):3–24, 2014.
- [156] Elizabeth G. Ryan, Christopher C. Drovandi, M. Helen Thompson, and Anthony N. Pettitt. Towards Bayesian experimental design for nonlinear models that require a large number of sampling times. *Computational Statistics and Data Analysis*, 70:45–60, 2014.
- [157] Xun Huan and Youssef M. Marzouk. Simulation-based optimal Bayesian experimental design for nonlinear systems. *Journal of Computational Physics*, 232(1):288–317, 2013.
- [158] Xun Huan and Youssef M Marzouk. Gradient-Based Stochastic Optimization Methods in Bayesian Experimental Design. *International Journal for Uncertainty Quantification*, 4(6):479–510, 2014.

- 
- [159] Kenneth J Ryan. Estimating Expected Information Gains for Experimental Designs With Application to the Random Fatigue-Limit Model. *Journal of Computational and Graphical Statistics*, 12(3):585–603, 2003.
- [160] Elizabeth G. Ryan, Christopher C. Drovandi, James M. McGree, and Anthony N. Pettitt. A Review of Modern Computational Algorithms for Bayesian Optimal Design. *International Statistical Review*, 84(1):128–154, 2016.
- [161] H. H. Afshari, S. A. Gadsden, and S. Habibi. Gaussian filters for parameter and state estimation: A general review of theory and recent trends. *Signal Processing*, 135(December 2016):218–238, 2017.
- [162] C. L. Thornton and G. J. Bierman. Givens transformation techniques for Kalman filtering. *Acta Astronautica*, 4(7-8):847–863, 1977.
- [163] Dan Simon and Optimal State Estimation. *Optimal State Estimation*, volume 2. 2006.
- [164] S.J. Julier and J.K. Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3):401–422, 2004.
- [165] I. Arasaratnam, S. Haykin, and R.J. Elliott. Discrete-Time Nonlinear Filtering Algorithms Using Gauss–Hermite Quadrature. *Proceedings of the IEEE*, 95(5):953–977, 2007.
- [166] Kazufumi Ito and Kaiqi Xiong. Gaussian filters for nonlinear filtering problems. *IEEE Transactions on Automatic Control*, 45(5):910–927, 2000.
- [167] Simon Haykin and Ienkaran Arasaratnam. Cubature Kalman Filters. *Ieee Transactions on Automatic Control*, 54(6):1254–1269, 2009.

- [168] Xiaojun Yang, Keyi Xing, Kunlin Shi, and Quan Pan. Joint state and parameter estimation in particle filtering and stochastic optimization. *Journal of Control Theory and Applications*, 6(2):215–220, 2008.
- [169] Afshin Rahimi, Krishna Dev Kumar, and Hekmat Alighanbari. Enhanced adaptive unscented Kalman filter for reaction wheels. *IEEE Transactions on Aerospace and Electronic Systems*, 51(2):1568–1575, 2015.
- [170] D. Magill. Optimal adaptive estimation of sampled stochastic processes. *IEEE Transactions on Automatic Control*, 10(4):434–439, 1965.
- [171] D. G. Lainiotis. Optimal adaptive estimation: Structure and parameter adaption. *IEEE Transactions on Automatic Control*, AC-16(2):160–170, 1971.
- [172] D G Lainiotis. Partitioning: A unifying framework for adaptive systems, I: Estimation. *Proceedings of the IEEE*, 64(8):1126–1143, 1976.
- [173] C. Masreliez and R. Martin. Robust bayesian estimation for the linear model and robustifying the Kalman filter. *IEEE Transactions on Automatic Control*, 22(3):361–371, 1977.
- [174] G Zames. Feedback and optimal sensitivity: model reference transformation, multiplicative semiforms and approximate inverse. *IEEE Transactions on Automatic Control*, 26(2):301–320, 1981.
- [175] S. H. Paskov. *Average Case Complexity of Multivariate Integration for Smooth Functions*, 1991.
- [176] Erich Novak and Klaus Ritter. High dimensional integration of smooth functions over cubes. *Numerische Mathematik*, 75(1):79–97, 1996.

- [177] Erich Novak and Klaus Ritter. The curse of dimension and a universal method for numerical integration. *International Series of Numerical Mathematics*, pages 177–188, 1997.
- [178] Erich Novak and Klaus Ritter. Simple cubature formulas with high polynomial exactness. *Constructive Approximation*, 15(4):499–522, 1999.
- [179] Charles W Clenshaw and Alan R Curtis. A method for numerical integration on an automatic computer. *Numerische Mathematik*, 2(1):197–205, 1960.
- [180] T. N. L. Patterson. The optimum addition of points to quadrature formulae. *Mathematics of Computation*, 22(104):847–847, 1968.
- [181] Tim Salimans, Diederik P. Kingma, and Max Welling. Markov Chain Monte Carlo and Variational Inference: Bridging the Gap. 2014.
- [182] Abdul Lateef Haji-Ali, Fabio Nobile, and Raúl Tempone. *Multi-index Monte Carlo: when sparsity meets sampling*, volume 132. Springer Berlin Heidelberg, 2016.
- [183] Benjamin Peherstorfer, Karen Willcox, and Max Gunzburger. Survey of multi-fidelity methods in uncertainty propagation, inference, and optimization. *ACDL Technical Report TR16-1*, pages 1–57, 2016.
- [184] Jay D. Martin and Timothy W. Simpson. Use of Kriging Models to Approximate Deterministic Computer Models. *AIAA Journal*, 43(4):853–863, 2005.
- [185] Donald R. Jones, Matthias Schonlau, and William J. Welch. Efficient Global Optimization of Expensive Black-Box Functions. *Journal of Global Optimization*, 13(4):455–492, 1998.
- [186] Donald Jones. A Taxonomy of Global Optimization Methods Based on Response Surfaces. *Journal of Global Optimization*, 21(4):345–383, 2001.

- [187] J Nocedal and S Wright. *Numerical optimization*. 2006.
- [188] Martin A. Grepl, Yvon Maday, Ngoc C. Nguyen, and Anthony T. Patera. Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 41(3):575–605, 2007.
- [189] Alexander I J Forrester. *Engineering Design via Surrogate Modelling: A Practical Guide - Constructing a Surrogate*. 2008.
- [190] Anirban Chaudhuri and Raphael T Haftka. Efficient Global Optimization with Adaptive Target Setting. *AIAA Journal*, 52(7):1573–1578, 2014.
- [191] R. R. Ernst and W. A. Anderson. Application of Fourier Transform Spectroscopy to Magnetic Resonance. *The Review of Scientific Instruments*, 37(1):93–102, 1966.
- [192] Giavarina D. Understanding Bland Altman analysis Lessons in biostatistics. *Biochemia Medica*, 25(2):141–51, 2015.
- [193] F. Feroz and M. P. Hobson. Multimodal nested sampling: An efficient and robust alternative to Markov Chain Monte Carlo methods for astronomical data analyses. *Monthly Notices of the Royal Astronomical Society*, 384(2):449–463, 2008.
- [194] F. Feroz, M. P. Hobson, and M. Bridges. MultiNest: An efficient and robust Bayesian inference tool for cosmology and particle physics. *Monthly Notices of the Royal Astronomical Society*, 398(4):1601–1614, 2009.
- [195] Kathryn E. Keenan, Maureen Ainslie, Alex J. Barker, Michael A. Boss, Kim M. Cecil, Cecil Charles, Thomas L. Chenevert, Larry Clarke, Jeffrey L. Evelhoch, Paul Finn, Daniel Gembris, Jeffrey L. Gunter, Derek L.G. Hill, Clifford R. Jack, Edward F. Jackson, Guoying Liu, Stephen E. Russek, Samir D. Sharma, Michael

- Steckner, Karl F. Stupic, Joshua D. Trzasko, Chun Yuan, and Jie Zheng. Quantitative magnetic resonance imaging phantoms: A review and the need for a system phantom. *Magnetic Resonance in Medicine*, 79(1):48–61, 2018.
- [196] Martin Uecker, Jonathan I Tamir, Frank Ong, and Michael Lustig. BART Toolbox for Computational Magnetic Resonance Imaging. <https://Mrirecon.Github.io/Bart/>.
- [197] Leo Breiman. Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). *Statistical Science*, 2001.
- [198] European Society of Radiology (ESR). Magnetic Resonance Fingerprinting - a promising new approach to obtain standardized imaging biomarkers from MRI. *Insights into Imaging*, 6(2):163–165, 2015.
- [199] Emmanuel J. Candes and Terence Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE Transactions on Information Theory*, 2006.
- [200] Emmanuel J. Candès, Justin Romberg, and Terence Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 2006.
- [201] Michael Lustig, David L. Donoho, Juan M. Santos, and John M. Pauly. Compressed sensing MRI. *Signal Processing Magazine, IEEE*, 2008.

# Vita

Drew Palmer Mitchell was born in Houston, Texas to Randall Glenn Mitchell and Melanie Ann Mitchell. His two younger brothers are Matthew Alexander Mitchell and Kyle Glenn Mitchell. After graduating as the valedictorian of Friendswood High School in 2007, Drew attended the Eastman School of Music in Rochester, New York to pursue a degree in music performance. Drew transferred to Texas A&M University in College Station, Texas in 2009 to pursue degrees in physics and nuclear engineering. In 2013, he received the degree of Bachelor of Science, *summa cum laude*, with a major in Physics and the degree of Bachelor of Science, *summa cum laude*, with a major in Nuclear Engineering. In 2013, he entered The University of Texas Graduate School of Biomedical Sciences at Houston and began a doctor of philosophy degree in medical physics. In 2018, Drew married Julika Ayla Kaplan.

Permanent address as of June 25, 2019:

8333 Braesmain Drive, Apt. 2303

Houston, Texas 77025