NEW METHODS IN QUANTIFICATION AND RF PULSE OPTIMISATION FOR MAGNETIC RESONANCE SPECTROSCOPY

MAX STEWART CHANDLER

A thesis submitted in partial fulfilment of the requirement
for the degree of Doctor of Philosophy to the
School of Computer Science and Informatics
Cardiff University, Wales

December, 2019

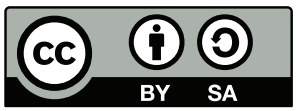
Max Chandler: *New Methods in Quantification and RF Pulse Optimisation for Magnetic Resonance Spectroscopy*
A thesis submitted in partial fulfilment of the requirement
for the degree of Doctor of Philosophy to the
School of Computer Science and Informatics
Cardiff University, Wales

For my parents,
who taught me critical thinking.

Jennifer Hogg,
for your incredible patience and support.

And Dillan Mistry;
I guess this means I won the bet.

LICENSE



SUMMARY

Magnetic Resonance Spectroscopy (MRS) is a powerful medical diagnostic and research tool that enables us to identify metabolite concentrations in a region of interest (ROI) *in-vivo*. This non-invasive diagnostic technique provides a large amount of information about a certain region in the body, such as the brain or spinal cord, with no impact on patient wellbeing. MRS is readily available in many clinical units across the UK with an MRI machine and no additional hardware is needed.

MRS has a number of challenges, including the requirement of a much higher level of magnetic field calibration compared to MRI, and detecting and analysing a substantially weaker signal per metabolite. To complicate the matter, there is a broad range of metabolites found *in-vivo* with overlapping proton spectra, obscuring signals and making spectral analysis very challenging.

The primary focus of this thesis is to explore methods to aid quantification of metabolites by exploring two ends of the issue, focusing specifically on GABA, NAA, Creatine quantification, of interest to a range of neuroscience studies. Firstly, the focus is on the analysis of the acquired spectral data utilizing the MEGA-PRESS pulse sequence, specifically aimed at GABA. Comprehensively benchmarking the current state-of-the-art spectral quantification methods with experimental data from phantoms of known composition lays the foundation for devising an improved quantification technique. This novel quantification method utilises a convolutional neural network for MEGA-PRESS spectra and can outperform the state-of-the-art.

Secondly, an optimisation method to find RF pulses that create specific excitations in the metabolites is devised, leading to spectra that are simpler to analyse. Such customisation of the spectra allows the removal of overlapping or obscuring features, creating chemically selective spectral acquisition methods. Moreover, the RF pulses are optimised over a range of scanner uncertainties to improve robustness. Simulations demonstrate that this approach can separate GABA, NAA and Creatine as well as Glutamine and Glutamate at 3 Tesla.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor Frank Langbein for his continued support and guidance. His ability to deconstruct and critically examine ideas along with his passion for discussion has been an incredible eye-opening learning experience. Without his patience and insightful discussions, the journey of working on this PhD would not have happened. Thank you for encouraging me to apply for doctoral funding all those years ago.

Secondly, Sophie Shermer for her relentless passion for science and research. Thank you for all of the lengthy discussions, and invaluable practical and theoretical knowledge in quantum control, magnetic resonance and phantom design. Chris Jenkins, thank you for sharing the experience of the long hours discussing ideas, debugging code and analysing data over the last few years.

Cardiff University and EPSRC Doctoral Training Grant for funding; without I would not have been able to afford to study. Finally, thank you to the staff at Cardiff University for their teaching, support and feedback over the last seven years.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LISTINGS

LIST OF ALGORITHMS

LIST OF ACRONYMS

ADC	apparent diffusion coefficient
AMARES	advanced method for accurate, robust and efficient spectral fitting
ANN	artificial neural network
AQSES	automated quantification of short echo time MRS signals
BFGS	Broyden–Fletcher–Goldfarb–Shanno algorithm
BN	batch normalisation
BW	bandwidth
CHESS	chemical shift-selective
Cho	Choline
CI	confidence interval
CLI	command line interface
CNN	convolutional neural network
conv	convolution
Cr	Creatine
CT	computed tomography
DAC	digital-to-analogue converter
DICOM	digital imaging and communications in medicine
DO	dropout
DSS	sodium trimethylsilylpropanesulfonate
FC	fully connected
FID	free induction decay
FID-A	FID appliance
FWHM	full-width half-maximum
GABA	γ-aminobutyric acid
GAMMA	general approach to magnetic resonance mathematical analysis
GANNET	GABA-MRS analysis tool
Gln	Glutamine
Glu	Glutamate
Glx	Glutamine (Gln) and Glutamate (Glu) combined
GPU	graphics processing unit
GSH	Glutathione
GUI	graphical user interface

HERMES	Hadamard encoding and reconstruction of MEGA-edited spectroscopy
HLSVD	Hankel Lanczos squares SVD
HWHM	half-width half-maximum
jMRUI	Java-based MR user interface
Lac	Lactate
L-BFGS	limited-memory BFGS
L-BFGS-B	L-BFGS with bound-constraints
LCModel	linear combination of models
LReLU	leaky ReLU
MAE	mean absolute error
MAPE	mean absolute percentage error
MEGA-PRESS	Mesher-Garwood PRESS
MLW	multi-line width
MR	magnetic resonance
MRI	magnetic resonance imaging
MRI/S	magnetic resonance imaging and spectroscopy
MRS	magnetic resonance spectroscopy
MRSNet	MRS network
MSE	mean squared error
NAA	N-Acetylaspartic acid
NLLS	non-linear least squares
NMR	nuclear magnetic resonance
NN	neural network
PCA	principal component analysis
PET	positron emission tomography
PPM	parts per million
PRESS	point resolved spectroscopy
PyGamma	Python-GAMMA
QUEST	quantification based on quantum estimation
R^2	coefficient of determination
ReLU	rectified linear unit
RF	radio frequency
ROI	region of interest
RWA	rotating wave approximation
SAR	specific absorption rate
SE	spin-echo
SNR	signal-to-noise ratio
SPECT	single-photon emission CT
STEAM	stimulated echo acquisition mode

SVD	singular values decomposition
TARQUIN	totally automatic robust quantitation in NMR
T_E	echo time
T_R	repetition time
VeSPA	versatile simulation, pulses and analysis
WIP	work-in-progress
WS	water suppression
WSF	WS factor

1

INTRODUCTION

Magnetic resonance imaging (MRI) is a non-invasive imaging modality that has revolutionised medical imaging. MRI is based on the principle of nuclear magnetic resonance (NMR) [6, 84], utilising strong static magnetic fields combined with field gradients for spatial localisation and radio frequency (RF) pulses for control. MRI scanners are widely available, with MRI scans accounting for 3.39M of 41.3M (12.2%) medical imaging scans performed by the NHS in the UK in 2018 [104]. Although magnetic resonance imaging (MRI) scans are still considerably more costly than alternatives such as computed tomography (CT) scans, unlike alternatives such as X-rays, positron emission tomography (PET) and single-photon emission CT (SPECT), MRI does not use ionising radiation and offers superior soft-tissue contrast and a variety of acquisition modes including diffusion and perfusion imaging, chemical shift imaging and spectroscopy.

Magnetic resonance (MR) is an excellent platform for research and development ranging from hardware development, such as transmit and receive coils to improve signal-to-noise ratio (SNR) [48, 107] to the development and optimisation of new pulse sequences. The latter generally does not require any hardware modifications and new pulses sequences therefore can easily be deployed on existing MRI equipment.

The focus of this thesis is on proton magnetic resonance spectroscopy (MRS) [89, 154] and the quantification of metabolites present in a spectroscopy voxel by matching basis sets of known metabolite signals. This provides knowledge of the chemical composition of biological tissue and insight into biophysical processes and disorders. MRS has enabled significant advances in detecting, classifying and modelling malignant tissue in the brain [2, 40, 171], prostate [72, 121, 161] and breast [36, 50, 79]. In modern psychology and neuroscience, MRS provides insights into normal brain function, as well as abnormalities associated with conditions such as anxiety, schizophrenia, depression and other mood disorders [25, 29, 150, 152, 158]. A key biomarker for these studies is γ-aminobutyric acid (GABA), the primary inhibitory neurotransmitter [88].

Although MRI has been clinically available for many years, MRS is not yet in widespread clinical use. As metabolite signal of interest are

often several orders of magnitude weaker $\times 10,000$ than the water signal, obtaining good SNR is challenging. Furthermore, to achieve spectral resolution required to distinguish metabolites requires extremely homogeneous magnetic field over the entire spectroscopy volume (voxel) and precise calibration of RF control fields. These requirements for MRS create a new set of significant challenges for clinical MR, such as improving scanner calibration, pulse sequence design for low signal metabolites and improvements to signal acquisition and spectral analysis (quantification). Consequently, any improvements in these areas of MRS will have a knock-on effect on other MR modalities, such as calibration improvements potentially leading to sharper MR images, or pulse sequence optimisation leading to faster or lower energy imaging sequences.

1.1 PROBLEM DEFINITION

This thesis presents a range of new techniques for MRS to improve current practices in the quantification of metabolites with a focus on metabolites of interest in neuroscience including γ-aminobutyric acid (GABA), N-Acetylaspartic acid (NAA), Creatine (Cr), Glutamate (Glu) and Glutamine (Gln).

It can be split into two parts: signal acquisition and signal analysis. The signal acquisition includes all the necessary steps for acquiring a spectrum — from the design and implementation of the pulse sequence to scanner calibration and prepossessing of raw data on by the scanner. Signal analysis is focused on decomposition of the signals from the scanner into their constituent metabolic contributions and is usually done off-line using independent software. These parts are directly linked as the signal acquisition determines the analysis. The key objectives explored in-depth are

- **Improved quantification of spectra** obtained using existing pulses sequences such as Mesher-Garwood PRESS (MEGA-PRESS) based on decomposition by Convolutional Neural Networks (convolutional neural network (CNN)).
- **Development of novel pulse sequences** designed to enhance signals of interest and suppress unwanted signals via chemically selective excitation pulses derived using quantum control techniques.

To assess the performance of the new tools a review of current state of state-of-the-art quantification techniques and benchmarking results for a range of calibrated phantom spectra will be presented.

1.1.1 *Magnetic Resonance Spectroscopy: Challenges in Signal Acquisition*

Simplified, MR works by using a strong homogeneous magnetic field, B_0, to align spins, typically ^1H and utilises RF fields for control. Spins are excited by the RF field, and will eventually relax back into alignment with the B_0 field, emitting a signal. For spectroscopy, each metabolites' signal is a consequence of its chemical structure, essentially forming an electromagnetic fingerprint , which is the basis for MRS. These signals, converted to the frequency domain to form spectra, can later be matched to a basis set of known signals, indicating the chemical composition of a region of interest (ROI).

MRS requires an extremely homogeneous magnetic field as spatial or temporal fluctuations in the B_0 field cause frequency drift and line broadening, resulting in misidentification of spectral features and obscuring smaller features. The challenge of performing MRS *in-vivo* is that biological systems are complex and noisy. This combined with the presence of many metabolites in often low concentrations makes it generally difficult to quantify more than a handful due to poor signal-to-noise ratios and spectral overlap. To improve SNR multiple repeats are necessary for noise filtering, outlier rejection and averaging of the signal, which makes MRS time-consuming.

MRS pulse sequences are primarily designed around broadband excitation, exciting all metabolites simultaneously. In theory, this should provide maximal information but in practice is not the case. Due to the high spectral overlap between metabolite peaks, more prevalent metabolites, especially those with high concentrations typically obscure peaks of lower concentration metabolites, making accurate quantification of these metabolites not possible. To address this issue pulse sequences have been developed that utilise frequency-selective editing, such as MEGA-PRESS [91, 92] or Hadamard encoding and reconstruction of MEGA-edited spectroscopy (HERMES) [125].

These editing pulse sequences are commonly used to aid the quantification of the primary inhibitory neurotransmitter GABA. Due to the importance of this metabolite and the difficulty quantifying it *in-vivo* there is large focus in the MRS community on improving its quantification [31, 95, 101, 103, 106]. In Chapter 5, a method is presented to customise spectral fingerprints of different metabolites to make them more distinguishable.

1.1.2 *Spectral Quantification: A complex Inverse Problem*

Spectral quantification is reverse engineering a spectrum to obtain absolute or relative concentrations of metabolites that have contributed to it. As explained above, quantification of *in-vivo* spectra is a complex task due to low SNR, a high level of spectral overlap and a range of non-linear noise sources distorting the spectrum.

Several software packages are available to perform quantification, usually focused on handling the task as an inverse-problem by fitting a "basis-set" of known signals to the spectrum. The issue with this approach is that basis set selection is complicated and typically separate basis sets are required for different magnetic field strengths, pulse sequence, pulse sequence timings and scanner hardware. To further complicate the matter, basis set creation is complex and time-consuming. Basis sets derived from experimental spectra suffer from issues with spectral quality. Computationally derived basis rely on accurate replication of pulse sequences and accurate models for metabolites for simulation. Finally, creating numerical models to match and correct or remove noise sources and distortion is complex. This leads to software that can be complicated to use, requiring specialist knowledge, with multiple options available to the user for adjusting fit parameters. It is also an obstacle to widespread clinical use, which calls for programs that do not require lengthy specialist training and enable clinicians to quickly obtain relevant, reliable data for diagnosis without having to tune many parameters. Ideally, quantification should be automated to speed up the process and remove any bias or uncertainty introduced by manual adjustments of parameters.

The current state of quantification in MRS is difficult to assess due to a lack of benchmarking of existing methods. Using *in-vivo* data for verification of quantification methods is problematic due to the absence of ground truth data, while simulated spectra are generally not sufficiently representative of experimental data (see Fig. 4.6). Reviews of different tools therefore often use linear combination of models (LCModel) as a benchmark, accepting the results it generates as the ground truth, with analysis of how well the reviewed method can match this data. Methodologically, this has substantial flaws as will be discussed in Chapter 3.

1.1.3 *Pulse Sequence Optimisation: A novel approach*

Pulse sequence optimisation is rarely found in MRS, where optimisations are typically focused on protocol enhancements for current pulse sequences. These protocol enhancements have significant benefits to the clinician or patients, where a reduction in total scan time or reduced total energy has a direct impact on patient wellbeing. This, however, is not the focus of this thesis. Instead, it focuses on a novel method for the design of RF control pulse shapes found through numerical optimisation.

Classical broadband pulses are based around rotations, designed to act universally on all spins rotating them by a fixed angle, e.g. 90° or 180°. Alternatively, pulses that selectively excite specific frequency ranges have been implemented and combined with universal pulses in the previously mentioned MEGA-PRESS [91, 92] and HERMES [125] sequences to improve chemical selectivity. However, the design of pulse sequences to detect low-signal metabolites such as GABA is challenging, the process of designing and testing these pulse sequences can be lengthy [1] and there is considerable room for improvement.

This thesis focuses instead on the design of pulses that selectively excite certain modes in specific metabolites. We show later in Chapter 5 we are able to custom define how the spectra for each metabolite should behave. This is done by selecting individual protons on a per–metabolite basis that should be in an excited state, effectively silencing protons that are not of interest and removing their signal from the resulting spectrum. This 'target-state' is then used as an optimisation target, with the aim of producing an RF pulse that is able to place multiple metabolites in their target-states concurrently. This sort of pulse sequence optimisation focused on clinical MR is not typically seen, where the focus is usually on NMR instruments where the level of robustness needed is substantially less. This is the second strength of this method, with experimentally validated ranges of instabilities integrated directly into the optimisation, putting these theoretical pulses within reach of experimental validation.

1 For example, the well known MEGA-PRESS pulse sequence is still considered to be work-in-progress (WIP) since its conception in 1988.

1.2 PUBLISHED MATERIAL

Journal Papers

P1 C. Jenkins et al. 'Seeking Ground Truth for GABA Quantification by Edited Magnetic Resonance Spectroscopy:Comparative Analysis of TARQUIN, LCModel, JMRUI and GANNET'. In: Mc (Sept. 2019), pp. 1–23. arXiv: 1909.02163. URL: http://arxiv.org/abs/1909.02163 — under revision

P2 M. Chandler et al. 'MRSNet: Metabolite Quantification from Edited Magnetic Resonance Spectra With Convolutional Neural Networks'. In: Mc (Sept. 2019), pp. 1–12. arXiv: 1909.03836. URL: http://arxiv.org/abs/1909.03836 — under revision

P3 Max Chandler et al. 'Robust Quantum Optimal Control for Unique Metabolic Spectra in Magnetic Resonance Spectroscopy'. 2019 — in preparation

Conference Posters

P5 Max Chandler et al. 'Quantum control for magnetic resonance spectroscopy'. In: *Proceedings of the All Wales Medical conference*. 2017

P6 Max Chandler et al. 'Advanced detection and quantification of biomarkers in magnetic resonance spectroscopy'. In: *Proceedings of the EPSRC UK Image-Guided Therapies Network+ Launch*. 2016

P7 Chris Jenkins et al. 'Quantification of edited magnetic resonance spectroscopy: a comparative phantom based study of analysis methods'. In: *Proceedings of ISMRM Annual Meeting*. 2019

1.3 CONTRIBUTIONS

C1 Contribution to benchmarking of existing tools for GABA quantification based on MEGA-PRESS spectra at 3 T.

C2 A new quantification method utilising a CNN for MEGA-PRESS spectra: MRSNet.

C3 Application of quantum control to design optimal pulses for selective excitation of metabolites.

C4 New technique to optimise pulses stable with respect to a range of B_0 and B_1 instabilities.

Contribution 1 (C1) is related to Chapter 3 and publication 1 (P1).
While this work was mainly completed by collaborators from Swansea University I contributed to the analysis of the benchmarking datasets, in particular, the initial planning and research for the project, aided in the creation of two phantom datasets and performing the LCModel analysis of the spectra. The novelty and importance of this work lies in providing benchmarking data for the most popular MRS quantification tools using experimental datasets with associated ground truth data. By publicly releasing the experimental benchmark datasets of spectra, it allows and encourages others to use it to analyse the performance of new or updated methods.

C2 is related to Chapter 4 and publication 2 (P2).
This work aims to improve quantification of MEGA-PRESS spectra with reference to the benchmark results from C1 by exploring the utility and performance of CNNs. A large range of options are explored, including how to present the data to the networks so they can learn effectively. It ultimately leads to a CNN that outperforms LCModel in terms of speed, simplicity and overall quantification accuracy.

C3 and C4 are directly related to Chapter 5 and P3-P7.
This chapter cover the largest body of work for this thesis with the aim to produce RF pulses to excite metabolites into specific states utilising control methods used for quantum spin-networks. C3 is the initial investigation of applying these quantum control techniques to metabolites and observing the resulting spectrum. The main focus of this contribution is to explore the limits of the control and what is realisable. This contribution shows that it is possible to produce controls that can manipulate multiple metabolites concurrently, enabling a chemically-selective editing of the spectra.

C4 focuses on improving the techniques from C3 to make them theoretically robust and experimentally realisable. The optimised controls generated by C3 are sensitive to noise, and due to the nature of MR the controls would not be effective in this environment. This contribution focuses on producing controls that are robust with respect to a range of expected instability from a clinical MR machine. By incorporating a range of uncertainties into the optimisation, it is possible to produce controls that are robust with respect to the two dominant sources of instability: magnetic field instability and control accuracy.

Additionally, as part of the work completed for this thesis a range of code and data sources have been created:

- Max Chandler and Frank Langbein. *MR spectral quantification using convolutional neural networks.* 2018
- S.M. Shermer; C. Jenkins; M. Chandler; F. C. Langbein. *Magnetic resonance spectroscopy data for GABA quantification using MEGAPRESS pulse sequence*. 2019. DOI: 10.21227/ak1d-3s20. URL: http://dx.doi.org/10.21227/ak1d-3s20
- Max Chandler. *A scheduler for automated execution of code written in bash.* 2015. URL: https://github.com/MaxChandler/scheduler
- Max Chandler, Frank Langbein and Sophie Shermer. *A framework for optimising MRS RF Pulses for metabolite target states*. 2017. URL: https://qyber.black/MRIS/control

All are publicly available, except for the QControl framework. It is a proprietary optimisation framework written in MATLAB for the optimisation of the MR control pulses. It is currently closed source but access is available upon request.

1.4 THESIS STRUCTURE

This thesis is split into six chapters, including this introduction, with three distinct contribution chapters: Benchmarking Quantification (C1), MRSNet (C2) and Pulse Sequence optimisation (C3 & 4). Each of these chapters will have an in-depth evaluation, with a summary evaluation of the thesis found in the epilogue.

Chapter 2 introduces the background and major themes of this thesis, starting with a brief introduction to the fundamentals of MR, the underlying physics and the hardware found in MR systems. An introduction to simulating metabolite spectra, the mathematical formalism and the current challenges in MRS simulation are outlined, followed by a brief review of existing quantification methods considered in the benchmarking exercise in Chapter 3. The notions of control optimisation in MR including optimisation currently used in MRS and high-field NMR systems are briefly summarised as a preamble to Chapter 5 on pulse sequence optimisation.

Chapter 3 is focused on benchmarking the current state-of-the-art for quantification of MRS MEGA-PRESS spectra for GABA quantification at

3 T. This chapter examines the reliability of a range of quantification tools when presented with experimental data that has associated ground truth data, something that has not been done before on such scale. There is an in-depth explanation of the tools and methods used to create the range of phantoms for experimental validation, along with an explanation of the fit routines used by each tool. This is followed by an analysis of the performance of these tools.

Chapter 4 focuses on the development of a novel quantification method utilising a convolutional neural network (CNN), leveraging the knowledge gained from Chapter 3 as a benchmark. This chapter examines how to present spectra to a CNN and how this affects the performance of the network. MEGA-PRESS presents a unique opportunity as multiple spectra (edit-on, edit-off and difference) are available. Additionally, there is a review of the performance effects of using basis sets from different sources that are used to generate spectra to train the network. This chapter shows that it is possible to improve upon current quantification methods in terms of speed, accuracy and simplicity with a CNN approach.

Chapter 5 introduces the theoretical validation of a novel method for producing MRS RF pulses. These pulses are designed by a numerical optimisation approach, based on techniques used for state transfer in quantum processes. The goal of this method is to produce chemically selective RF pulses that alter the resulting spectra, removing obscuring signals. The targets are selected to excite specific protons or quantum states for selected metabolites only, effectively allowing complete control of the spectrum (if a suitable and realistic pulse can be found). This chapter covers the challenges of this method, such as picking appropriate targets, realistic time and energy constraints, and the integration of scanner instabilities into the target function to produce robust pulses. These robust pulses are the first step toward experimental validation of this method. This chapter finishes with a comprehensive evaluation of the utility of these pulses, and how well they are expected to be able to theoretically perform with respect to a large number of scanner uncertainties.

Chapter 6 evaluates and discusses the contributions made and presents overall conclusions regarding the work presented on the current state of quantification, the newly proposed MRSNet method, and pulse sequence optimisation for chemically selective spectroscopy. The chapter concludes with a discussion of future work.

2

BACKGROUND

This chapter provides the background necessary for this thesis in magnetic resonance spectroscopy (MRS), quantification, neural networks, and optimal control of quantum systems. Section 2.1 provides a background to MR, the underlying physics, hardware, scanner control with pulse sequences, and calibration to mitigate the challenging environment. Section 2.2 dives deeper into the theory of MR and links with quantum spin-networks which provides a primer to the later sections on control optimisation. Next, Section 2.3 explores the aims and challenges, followed by an in-depth literature-review of current methods, ultimately highlighting the need for improvement. Followed by this is a brief background in neural networks as a background for Chapter 4 in Section 2.4. Finally, Section 2.5 covers optimal control, applications to quantum systems and general optimisation found in MR.

2.1 PRINCIPLES OF MAGNETIC RESONANCE SPECTROSCOPY

Magnetic resonance (MR), as the name suggests, utilises a strong magnetic field in combination with RF pulses for control to produce a signal from biological material. MR is commonly known for imaging, with clinical MR machines commonly referred to as MRI scanners. Magnetic resonance spectroscopy (MRS) [46, 73] is another modality of MR that requires no additional hardware than is necessary for standard MRI, making it readily accessible in all clinical units and hospitals. MRS aims to gain an understanding of the metabolic composition of a defined region of interest (ROI). The spatial isolation is achieved through the use of gradients that encode spacial information. This signal typically originates form a quantum mechanical property called spin. Protons (^{1}H) are typically used for signal acquisition due to their naturally high abundance *in-vivo*, but other nuclei can be used for spectroscopy such as ^{13}C or ^{31}P. However, unlike MRI, MRS actively tries to suppress the largest source of the proton signal, water, to measure the substantially weaker metabolite signal. The resulting signal is Fourier transformed to create a spectrum, that can later be decomposed to gain an understanding of the metabolic content. MRS is effectively NMR performed on living

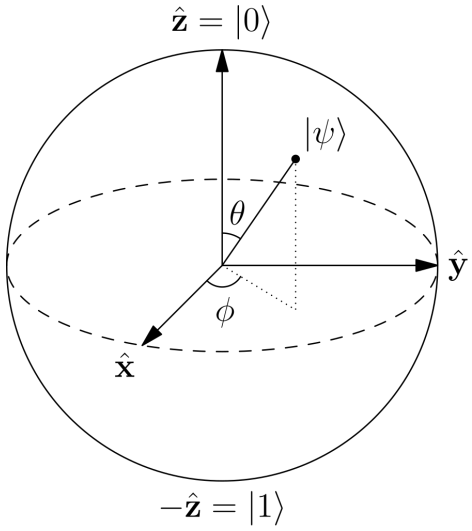


Figure 2.1: Bloch Sphere. Image courtesy of Glosser.ca (CC BY-SA 3.0)

tissue, presenting a new set of challenges due to the higher temperatures, constantly moving liquids with spacial localisation and calibration issues across a much larger scan area.

2.1.1 *Spin*

Spin [4, 138] is the quantum mechanical property that enables MR. Spin relates to the angular momentum of elementary particles, with the large static magnetic field and electromagnetic signals to controls being able to alter the magnetic moment, creating a signal. Simplified, spin (n) is the angular momentum quantum number, of the form $\frac{n}{2}$, where n must be a non-negative integer. It is a fundamental property of particles, with the spin number indicating how the spin can be quantised. For MRS, any spin that can have its angular momentum quantised can be used (i.e. any non-zero spin), enabling interaction and measurement within a magnetic field. This thesis focuses specifically on the spin-$\frac{1}{2}$ hydrogen-1 (^{1}H) nuclei used in standard, clinical MRI, due to the high availability *in-vivo*.

In order to more easily visualise spin states in the following section, the Bloch sphere is introduced. The Bloch sphere shown in Figure 2.1 is a visualisation tool that provides a framework to project a quantum state of a single spin, or an ensemble onto a 3D sphere.

The Larmor frequency ω of a spin is determined by the gyromagnetic ratio of the spin and the local chemical environment, which alters the resonance frequency. The resonance frequency of a spin is directly

proportional to the strength of the main static external magnetic field (B_0).

$$\omega = -\gamma B_0 \tag{2.1}$$

The gyromagnetic ratio γ is 42.58 MHz per Tesla [23] for ^1H; for a typical 3 T clinical scanner, the operating frequency is near 127.74 MHz. This constant gyromagnetic ratio conveniently allows for spectra to be rescaled to be in units of parts per million (PPM) (δ), enabling a comparison of spectra between different strength magnetic fields.

$$\delta = \frac{f - \omega}{\omega} \times 10^6. \tag{2.2}$$

State distribution in spin systems is determined by temperature and the strength of the magnetic field as described by the Boltzmann distribution in Equation 2.3.

$$\frac{z+}{z-} = \frac{e^{-\Delta E}}{kT} \tag{2.3}$$

Where $z+$ and $z-$ represent the populations of spins in the upper and lower energy states, ΔE is the energy difference between the states, T is the temperature in Kelvin and k is the Boltzmann constant.

To visualise this, the reader should imagine a set of spins tumbling in space, which are subject to a static magnetic field. The rate of tumbling is determined by the temperature, with spins either parallel or anti-parallel with the magnetic field. These two states have nearly identical populations due to the small energy difference between the spin states. The thermal tumbling provides enough energy to 'knock' spins between the higher and lower excitation states. Creating a small differential in the number of spins either parallel or anti-parallel with the magnetic field, with the favour being toward the lower energy state, in alignment with the field. When a MR signal is emitted after excitation, the bulk of the spins' signals cancel each other signal out as they are in opposing states. This leaves the remaining signal that originates from the small additional percentage of spins previously in the lower energy state.

The aforementioned metaphor effectively describes a free induction decay (FID) pulse sequence. An FID sequence is the most simplistic pulse sequence, comprising of a single excitation RF pulse, followed by an immediate readout. In the case of a group of single protons, a B_1 field is applied, tipping alignment away from the main static field (B_0) typically to a prescribed 'flip-angle' of 90°. The B_1 field is turned off when the desired tip angle is achieved. With the protons magnetic moment now aligned with the transverse field, a sine-wave can be read

by a receiver coil in the transverse plane as the protons oscillate at their Larmor frequency.

These spins eventually relax back into alignment with the magnetic field due to relaxation; the natural process of entropy in MR. Relaxation happens continuously, with the following two primary sources of energy dissipation.

Spin-lattice relaxation (T_1)

Spin-lattice relaxation is the longitudinal relaxation of spins. T_1 is the decay constant for the time taken to recover the z component of the magnetism. T_1 defines the time taken for spins to re-align with the static field, after being excited and rotated into the transverse plane.

Spin-spin relaxation (T_2)

Spin-spin relaxation is responsible for the loss of phase coherence of spins in the transverse plane. Spin-spin relaxation is caused by random processes, such as interactions with nearby spins, local magnetic field disturbances and molecular tumbling effects. T_2 is the decay constant for spin-spin relaxation and is generally faster than T_1. T_2 is dependent on the composition of the local environment, with different times found *in-vivo* depending on the tissue type, e.g. for the brain, liver or prostate [38].

Molecules

So far, the discussion has been based on a free single proton in isolation which is not realistic. In practice, spins are predominately bound in stable chemical structures *in-vivo*, with the focus of this thesis specifically on on Metabolites. Metabolites are any substance that is necessary for metabolic function, e.g. for growth, signalling or stimulation and inhibition. The chemical structure of these compounds directly affects the resonances of the spins in the following ways.

Chemical shift

Chemical shift is the resonance frequency of a spin, relative to the static magnetic field B_0 frequency. The spectrometer frequency is directly proportional to the strength of the magnetic field, which is later discussed in Section 2.1.2. The chemical shift is determined by electron distribution, which itself is determined by the chemical structure of the molecule. This electron distributions partly shield the spin from the full effect of the

external magnetic field, altering their resonance frequency, providing an offset from the spectrometer frequency.

Chemical shift is a key property enabling identification of molecules with MR spectroscopy, where the frequency offset from the spectrometer is dependent on the chemical structure of the molecule.

J-coupling

J-coupling (indirect dipole-dipole coupling) is the hyperfine interaction between two spins in the same metabolite, due to local interactions occurring between electrons and nuclei. J-couplings are enabled via the chemical bonds of a molecule, and they enable the transition of energy between spins via these structures. Generally, the strength of these bonds is based on proximity, with nuclei separated by fewer bonds experiencing a stronger coupling effect. J-couplings are invariant with respect to scanner strength; this property means that as the scanner frequency increases the distance between the split peaks decreases. Additionally, J-couplings can be positive or negative; they are symmetric in strength between two spins.

J-couplings are responsible for line-splitting seen in MR spectroscopy, the multiplicity of which is based on the number of protons in the coupled group. More coupled protons make for a more complex spectra increasing the amount of line-splitting seen. This line-splitting provides a great deal of additional information as it varies the spectra on a per metabolite basis, compared to chemical shift alone which would simply produce singlet peaks. J-couplings combined with chemical shift essentially gives each molecule an electromagnetic fingerprint based on its chemical structure.

Conveniently line-splitting can only occur with nuclei that have an integer spin value, meaning that the naturally abundant Carbon ^{12}C and Oxygen ^{16}O isotopes do not cause line-splitting, simplifying the spectra and simulation. A comprehensive introduction into line splitting is unnecessary for this thesis, but an excellent background is available from the Hans J. Reich at the University of Wisconsin [119].

2.1.2 *MR Hardware*

The primary component in an MR machine is the large static magnetic field. This is usually generated with a superconducting magnet, cooled with liquid helium and nitrogen. This magnet is typically in the $1.5-3$ Tesla range for a clinical scanner with $7\,T$ for more advanced clinical

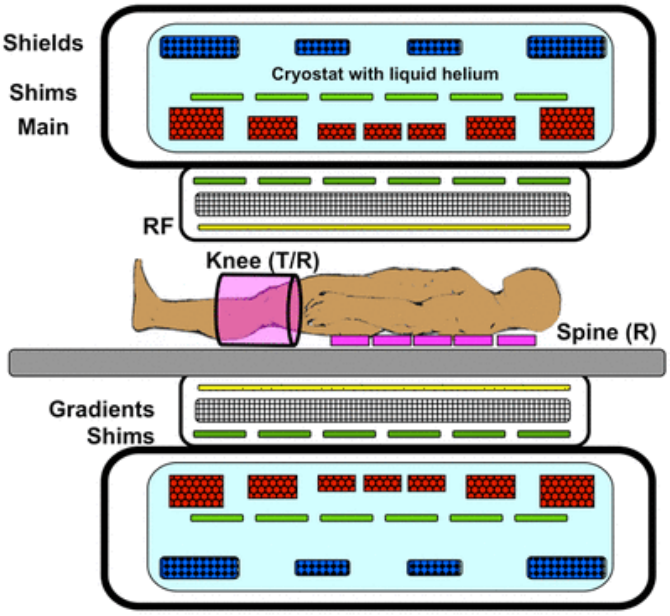


Figure 2.2: The layers of a typical MR machines shim, gradient and RF coils. Image Courtesy of Allen D. Elster, MRIquestions.com [34].

and research units. The focus of a clinical MR scanner is to provide a large homogeneous magnetic field across a large scanner bore (<60 cm), requiring a large magnet to generate a sufficient field.

For pre-clinical MR machines, it is not uncommon to find 9.4 T-15.2 T [141] strength magnets; however, these have small scanner bores (<10 cm) and are typically used for studies of small animals such as mice and rats. NMR systems have magnets that are significantly stronger (15-25 T) with very small bore diameters (<5 cm), where the focus is shifted to studying the chemical structure of molecules, rather than spacial localisation and imaging. This allows NMR machines to be comparatively more compact, all while achieving a much stronger and homogeneous field across the scan area. Fundamentally, all of these machines utilise the same principle but have significantly different focus.

As the strength of the principal magnetic field increases, it benefits MRS by increasing SNR, and spectral and spatial resolution, leading to a cleaner spectrum. However, at 7 T and above, the operating frequency for ^{1}H (298 MHz) begins to reach the microwave band. As such, more attention must be placed on limiting tissue heating from RF energy exposure and reducing specific absorption rate (SAR), in addition to reduced T_1 relaxation times that must be compensated for.

An MR machine has many layers, as shown by figure 2.2. First, inside the magnet bore are the shim coils, used to adjust the B_0 field, reducing

any inhomogeneities via "shimming". Next, there is a selection of gradient coils, used for spatial localisation of signals, used heavily in MRI to construct the images. These gradient coils are usually omitted from NMR machines as they are not required as NMR typically excites the whole volume inside the scanner bore. For the best SNR RF coils are placed last, closest to the patient or specimen. For clinical MR machines, they typically come with a selection of additional RF coils, such as spine coils located in the patient table, or independently moveable coils for specific applications such as head or knee coils. It is often more beneficial to use these as these provide a much better SNR compared to the much larger body coil due to the proximity.

2.1.3 *Control (pulse sequences)*

Pulse sequences are a programmed set of instructions for the RF transmit and receive, and the gradients required for the operation of an MR scanner. Each pulse sequence has a specific desired outcome, such as imaging or spectroscopy. Pulses sequences additionally contain the waveform to be transmitted by each coil at which time, along with the desired amplitude and duration of the pulse. Whether these are achieved depends on the characteristics of the waveform generator, amplifiers and coil properties. Gradients are an excellent example of this, where a large amount of energy must be ramped up and down through the coils. Near instantaneous switching of this much energy is not possible, creating a rise time before the maximum amplitude is achieved. This factor is summarised by the 'slew-rate' metric, which describes how quickly a particular model of a scanner can ramp up to the maximum gradient energy. With this in mind, and limits to the amount of energy a patient can absorb, characterised by the SAR, pulse sequences must be designed to be energy conservative and realistic where possible.

As an example MRS pulse sequence is the spin-echo (SE) pulse sequence. First, a 90° excitation pulse is applied to tip the magnetisation into the transverse plane, away from alignment with the Z-axis. Spins begin to precess in the transverse plane, some faster than others. This excitation pulse is followed by a 180° pulse at TE/2, to rephase the spins that are precessing at different rates due to T_2 relaxation. If timed correctly, this leads to the maximal signal at the readout window TE. This method of applying additional pulses for rephasing is widely employed

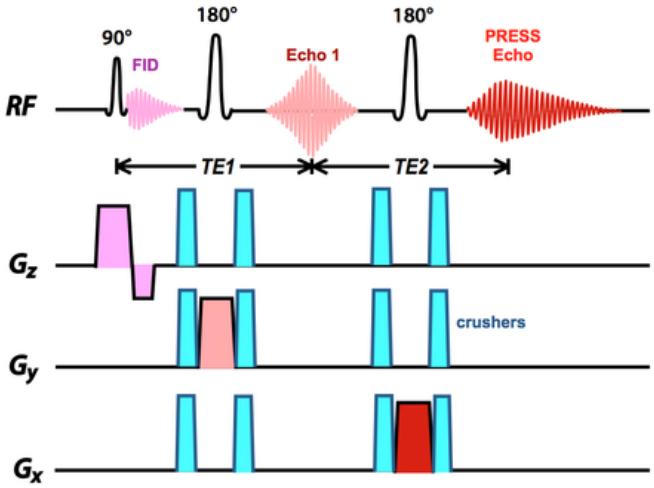


Figure 2.3: PRESS pulse sequence diagram. Note the use of gradients used for spacial localisation. Image Courtesy of Allen D. Elster, MRIquestions.com [34].

in other pulse sequences to negate the effects of phase decoherence in the transverse plane.

To solve the spatial localisation problem gradients are employed in MRS to encode spatial information by providing out-of-volume suppression and "spoiling" between RF excitation pulses to reduce any steady-state magnetization in the transverse plane. Gradient localisation schemes commonly used in MRI do not apply to MRS as they alter the resonant frequency of the protons, which would destroy the spectra. Instead, RF pulses combined with gradients in the x, y, z directions isolate the ROI. Protons only in the intersection of each plane will experience all of the RF pulses, creating the readout with the others dephased producing no signal.

Spatial localisation schemes such as point resolved spectroscopy (PRESS) [7], and stimulated echo acquisition mode (STEAM) [92] utilise these gradient pulses to enable broadband excitation with spatial localisation. PRESS, shown in Figure 2.3 utilises one 90° followed by two 180° pulses, whereas STEAM uses three 90° pulses. PRESS is the preferred scheme, typically producing twice the signal as STEAM. However, with PRESS, the pulse sequence is limited in how fast the readout window can be, reducing its utility at 7T due to the aforementioned T_2 relaxation issues. STEAM, on the other hand, can read out substantially faster than PRESS, enabling the detection of short T_2 metabolites, and has the added benefit of a much lower SARs than PRESS.

There are more complex pulse sequences available in MRS, such as the edited pulse sequence Mesher-Garwood PRESS (MEGA-PRESS) which uses frequency-selective editing pulses to suppress or enhance spectral features. MEGA-PRESS is the main focus of Chapters 3 and 4 of this thesis for its ability to resolve the low signal metabolite GABA.

MEGA-PRESS is based off the PRESS pulse sequence, with the same structure of one 90° followed by two 180° pulses for excitation and refocusing. MEGA-PRESS has two acquisitions, where the first acquisition is essentially a PRESS sequence, acquiring an "edit off" spectrum. The second acquisition includes two frequency-selective editing pulses placed symmetrically around the final 180° refocusing pulse for the "edit on" spectra.

The frequency selective pulses are utilised to excite spins that are coupled to target spins to induce an excitation elsewhere in the spectra. For the common application of GABA detection by removal of the overlapping Cr peak, a frequency selective editing pulse is applied at 1.9 ppm to excite the coupled GABA spins near 3.1 ppm. As the Cr peak is unaffected by this editing pulse, the spectra produced for the both the edit off and edit on acquisitions is the same, allowing it to be subtracted to form the difference spectra (edit on-edit off). Mullins *et al.* [101] produced an excellent review of the current practice and methods used for GABA quantification with MEGA-PRESS.

More recently, the HERMES pulse sequence [125] has been developed. This pulse sequence is an adaptation of the MEGA-PRESS pulse sequence in an effort to unify the pulse sequence across vendors, as the current implementation of MEGA-PRESS varies from vendor to vendor.

2.1.4 *Robustness, Environment and Calibration*

Some of the largest challenges facing clinical MRS are due to the environment and performing accurate calibration to mitigate the impact of this. Calibrating the magnetic field across such a large volume, with the added problems caused by physical movement of the patient and molecular tumbling *in-vivo* makes MRS very challenging. As spectroscopy is essentially the measurement of offsets from the B_0 field, any local perturbation seen by spins in this field can cause major problems, ultimately drifting the location of spins relative to B_0, migrating their location in the spectrum. This effect causes the characteristic Lorentzian, Gaussian and Voigt lineshape [86] in MRS, where the distribution of protons at a

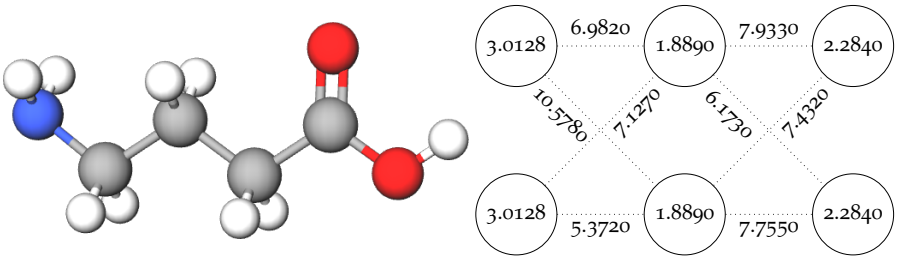
given resonance mimics the distribution of the B_0 field across the ROI. To reduce this effect, shimming is required for each new ROI *in-vivo*, with each area having its own magnetic characteristics to be compensated for. Automated shimming methods are typically a simple optimisation routine, using a SE pulse sequence to measure the linewidth of the water peak in the ROI, automatically adjusting the current to the shim coils, attempting to reduce this linewidth. It is not uncommon to find a linewidth of $>$12 Hz at 3 T *in-vivo*, where a range of $<$8 Hz is suggested for acceptable spectra quality at 1.5 T [27]. It is also common to find that the automatic shim routines perform sub-optimally, and it is usually possible to bring the linewidth down to around 4 Hz with manual adjustments.

Additional calibration for the tip-angle is performed before scanning, which aims to optimise the voltage of the RF transmitter to achieve the maximum signal from the ROI. Effectively, this is an optimisation for the Ernst angle, which describes precisely this effect. Similarly to before, the calibration is performed by iteratively searching RF transmission voltages with a SE sequence to maximise readout signal. Finally, adjustments to the pulse sequence timing parameters T_E (echo time; the time between RF pulse and echo readout) and T_R (repetition time; the time between corresponding points in a repeating series of pulses and echoes) also have a significant impact on the final spectra. These two timing values are heavily dependent on the goal, such as suppression of macromolecule signals with a longer T_E time.

There are additional components of MRS control that are necessary to acquire high-quality spectra, with one key area being water suppression. chemical shift-selective (CHESS) [49] pulses along with gradient dephasing is typically used before a pulse sequence, to suppress the dominant water signal found around 4.7 ppm. Without this suppression, due to the intensity of the water peak, the base of the peak can cover metabolite any surrounding metabolite signal up to $\pm$1.5ppm either side.

2.2 SIMULATION

Simulation of MRS spectra is theoretically simple, however in practice (see the later Figure 4.6) accurate simulation of the MR environment and pulse sequences is incredibly challenging.This section introduces the mathematics of closed spin systems, followed by open system dynamics as the role of the environment in MR is substantial.



(a) GABA chemical structure

(b) GABA J-coupling network

Figure 2.4: Chemical structure of GABA alongside the corresponding Chemical shift (ppm) and J-coupling (Hz) network. Values from Govindaraju *et al.* [45].

Simulation of metabolites is one of the core themes of this thesis, where it is used in two contexts; creation of basis sets for use as a source of truth in quantification, and optimisation of pulse sequences. In Chapters 3 and 4 a range of simulators are used to create basis sets for quantification. Later in Chapter 5 a custom Matlab framework [15] is used to optimise RF pulses for simulated metabolites.

Simulation of MR is performed in the rotating-frame, simplifying the simulation by removing the rapidly oscillating term that originates from the B_0 field. The underlying density-matrix simulation method for molecules have been well studied, the reader is directed to the following texts for a more comprehensive background [43, 55, 84, 122].

2.2.1 *Metabolites as Spin-Networks*

As previously mentioned molecules are a collection of atoms, bound into a chemical structure. We consider protons in the metabolite molecules. These protons form a network of spin-$\frac{1}{2}$ particles, as determined by their chemical shift and J-couplings as shown in figure 2.4.

Spin-networks are not only found in molecules but are an essential theoretical research tool for many areas of quantum systems in quantum information transfer with spin-networks [74], to new device research and manufacturing for quantum technologies [21]. Conveniently, the simulation of quantum spin-networks and metabolites utilise the same underlying mathematics, enabling the use of quantum control techniques developed for such networks to optimise pulse sequences for MR.

2.2.2 *Closed System Dynamics*

A closed system is one that isolated, with no interaction with an external environment. In MR, such a system only experiences unitary evolution under its dynamics and from the influence of two control fields (σ_x & σ_y). These control fields represent the interaction the system experiences from the two primary RF fields.

The Schrödinger equation describes the dynamics of a closed system,

$$\imath\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle \tag{2.4}$$

$\hat{H}$ is the total Hamiltonian, which is the sum of the system Hamiltonian and the contributions from the Hamiltonians which define our controls, and $|\Psi\rangle$ describing the wave function of the system. $\hbar$ is the reduced Planck constant $\hbar = \frac{h}{2\pi}$, where h is the Planck constant.

The system Hamiltonian H_0 is comprised of the chemical shifts and the J-couplings between nuclear spins, for N spins in the molecule.

$$H_0 = \sum_{n=1}^{N} \epsilon_n \sigma_z^n + \sum_{l=1}^{N}\sum_{k=1}^{N} J_{l,k}\left(\sigma_x^l\sigma_x^k + \sigma_y^l\sigma_y^k + \sigma_z^l\sigma_z^k\right) \tag{2.5}$$

σ_x^n, σ_y^n, σ_z^n are the extended Pauli spin operators acting on spin n. These are N-fold tensor products whose nth factor is one of the Pauli matrices respectively, and all other factors are the 2×2 identity matrix I.

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -\imath \\ \imath & 0 \end{pmatrix}, \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

The chemical shifts ϵ_n reflect the proton resonance frequencies relative to the static magnetic field B_0. The J-couplings $J_{l,k}$ are the strength of the interaction between two given spins via the hyperfine structure of the molecule. Contributions from short and long-range dipole coupling terms to the system Hamiltonian are ignored as their contribution becomes averaged in isotropic liquids due to their chaotic motion.

Controls represent two orthogonal x and y RF coils found in all MR systems with the corresponding control Hamiltonians H_x, H_y, which are formed by the tensor product of the σ_x^l and σ_y^l Pauli operators acting on all spins l uniformly. The total Hamiltonian of the system becomes

$$\hat{H}(t) = H_0 + \sum_{j=x,y} H_j u_j(t) \tag{2.6}$$

Time is split into K uniform time steps of length Δt, from time 0 to a prescribed final time T. Each control signal is a piecewise constant

signal that is given by two vectors $u_x = (u_{(x,0)}, \ldots, u_{(x,K)})$ and $u_y = (u_{(y,0)}, \ldots, u_{(y,K)})$, providing a control amplitude for each time step. The modulation of these amplitudes provides the ability to steer the system from some initial state towards the desired target state. This provides a realistic control simulation, analogous to the digital-to-analogue converter (DAC) found in MR systems for control generation, which outputs a piecewise constant amplitude-modulated signal.

System simulation in this thesis is performed on the density matrix ρ for easy translation to a dissipative simulation. The density matrix is formed as the outer product of the wave function $|\Psi\rangle$.

$$\rho = |\Psi\rangle \langle\Psi| = \sum_j p_j \left|\Psi_j\right\rangle \left\langle\Psi_j\right| \tag{2.7}$$

Where p_j is the probability that the system is in state $\left|\Psi_j\right\rangle$ and the $\left|\Psi_j\right\rangle$ form the (computational) basis of the Hilbert space of all wavefunctions of the system. All density matrices are normalised such that $\mathrm{Tr}(\rho) = 1$. The Schrödinger Equation (2.4) describing the dynamics of the system becomes the Liouville-von Neumann Equation

$$\imath\hbar\dot{\rho}(t) = [H(t), \rho(t)] \tag{2.8}$$

where $[A, B] = AB - BA$ is the Lie bracket.

Solving the Liouville-von Neumann equation for a piecewise constant control is directly derived from the Schrödinger equation solution, employing matrix exponentials. Hence, the dynamics of the system from an initial state, described by the density matrix ρ_0, is then given by

$$\rho(T) = U(T)\rho_0 U(T)^\dagger \tag{2.9}$$

where $U(T)$ is the propagator from time 0 to T. Under piecewise constant evolution it is computed as

$$\begin{aligned} U(T) &= U_K U_{K-1} \ldots U_2 U_1 \text{ where} \\ U_t &= exp\{-i/\hbar \Delta t \hat{H}(t)\}, t = 1, \ldots, K. \end{aligned} \tag{2.10}$$

The readout is calculated by propagating the system under its dynamics. At each time step, a signal is measured by projecting the density matrix with the lowering operator in Equation 2.11.

$$\sigma_- = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}. \tag{2.11}$$

Finally, a simple decay model is imposed onto the readout signal to simulate decoherence,

$$S(t) = S(t)e^{(-t\Delta t)/\delta} \tag{2.12}$$

where δ is the prescribed half-width half-maximum (HWHM) linewidth of a Lorentzian in Hz.

2.2.3 *Open System Dynamics*

To simulate dissipation, the mechanics are extended by Markvoain dynamics to introduce the influence of the environment. This is necessary to correctly simulate how the optimised pulses will perform in the presence of T_1 and T_2 relaxation.

We define the Lindbladian, describing the effect of the environment,

$$\mathcal{L} = \Sigma_n \left(\gamma_1 \mathcal{D} \left(\sigma_-^{(n)} \right) + \gamma_2 \mathcal{D} \left(\sigma_z^{(n)} \right) \right) \tag{2.13}$$

where γ_1, and γ_2 are the inverse of T_1, and T_2 respectively, with σ_- being the lowering operator previously defined in Equation 2.11. $\mathcal{D}$ is a super-operator acting on some operator ν

$$\mathcal{D}(\nu) = \nu\rho\nu^\dagger - \frac{1}{2}\left(\nu^\dagger\nu\rho + \rho\nu^\dagger\nu \right). \tag{2.14}$$

With the Liouville-von Neumann Equation, this leads to a Lindblad master equation

$$\imath\hbar\dot{\rho}(t) = [H(t), \rho(t)] + \mathcal{L}(t). \tag{2.15}$$

Readout is calculated in the same way as non-dissipative simulation by repeatability sampling with the lowering operator.

2.2.4 *Challenges*

Simulation of quantum systems is inherently challenging as the memory requirement for the state vector grows exponentially with system size (2^N). Naturally, the exponential growth of the matrix size causes a close to exponential growth in time requirement due to the computational cost of calculating a matrix exponential required for simulation.

Figure 2.5 displays the memory requirement for the matrices for simulation along with the average runtime for the inbuilt matrix exponential function in Matlab (R2019a) performed on randomly generated complex matrices. The sharp increase in runtime growth when compared to memory requirement is due to the larger matrices being pushed to slower memory types, moving from CPU cache down the memory-speed hierarchy to RAM; for much larger systems, this data would ultimately overflow RAM and end up in swap on disk, crippling performance.

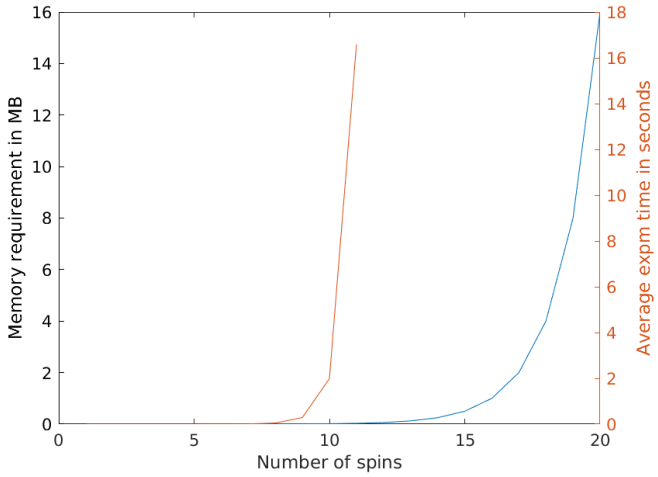


Figure 2.5: Simulation memory and time requirement. "expm" is the built in Matlab (R2019a) matrix exponential function.

For a simplified optimisation with a pulse sequence has 1,000 steps, requiring 1,000 matrix exponentials, each optimisation iteration for a seven spin system would take 45 seconds, eight spins would take 171 and nine 328. It is clear to see that this problem quickly grows out of hand. However, conveniently the majority of metabolites that are of interest in for this thesis are six spin systems, with the most significant being NAA with seven.

MR simulation

Simulation of metabolites in the MR environment is particularly challenging, mimicking the broad range of potential noise sources, control distortions and magnetic field instabilities. Unfortunately, even running an open system simulation does not capture the wide range of potential distortions, such as a poorly calibrated tip-angle leading to an under or over-rotation of spins, changing the characteristics of the pulse sequences and affecting the results. A large range of other effects, such as RF coil ring-down, non-linear amplifiers favouring specific power or frequency ranges and even the ramp up and down times of the RF and gradients are unreasonable to characterise and simulate in the time frame of this thesis. In general, a highly accurate MR simulator is complex, as hardware uncertainties cannot be easily considered, or further extrapolated to other MR machines, as they may each have their own peculiarities.

To further complicate the matter, simulation parameters for metabolites are uncertain [70] and are still widely debated. These simulation parameters designate the chemical shift of the protons and the J-coupling

structure. This is later shown to be a large issue in Chapter 4, where a range of GABA models (Govindaraju *et al.* [45], Kaiser *et al.* [60] and Near *et al.* [103]) all fail to match experimentally collected GABA spectra shown in Figure 4.6.

In addition, simulation of MR pulse sequences is complex, wherein some cases (such as the MEGA-PRESS pulse sequence), the implementation differs from vendor to vendor. Furthermore, there are many possible ways to achieve rotations in MR simulation, with either unrealistic instantaneous rotations or through the use of shaped pulses. This choice of pulse implementation not accurately recreate what is seen experimentally by the spins, due to the range of potential distortions mentioned at the start of this sub-section.

As mentioned in Section 2.2.4, simulation of MR spectra is computationally expensive. This imposes a limit to the number of spins a system can have before it can no longer be simulated in a reasonable time on standard computers, with current hardware, this limit tends to be around $10-12$ spins. Hogben *et al.* has developed 'Spinach' [54], a Matlab based simulation library for spin dynamics that can simulate upto 40 spin systems. This is achieved through the removal of unpopulated states and long-range couplings that have a minimal effect on the system dynamics. Additional methods such as Spin-Scenario [20] and the Fokker–Planck formalism [32, 71] have been recently displayed to aid in speeding up spatially encoded MR simulations. Techniques used by these methods to speed up MR simulation have not been employed in this thesis, for a range of reasons. Primarily, as speed is currently not a concern for the methods used as the largest metabolite (NAA) has 7 spins, but it is on the roadmap for future work in Chapter 5. Secondly, by restricting the number of state-spaces that are reachable to speed up computation would theoretically constrain the optimisation of controls by restricting the number of states that could be reached. In practice, these states may never be used or even be attainable in a restricted MRS simulation, but this would require significant investigation that has been constrained to future work.

2.3 SPECTRAL QUANTIFICATION

Quantification aims to gain an understanding of the ROI by quantifying the levels of individual metabolic contribution to the spectrum. Quantification is inherently an inverse problem, where the metabolite signal

is linearly proportional to the quantity. As such, it is typically tackled by fitting a basis set of normalised pure individual metabolite spectra to decompose the input spectra plus additional terms for baseline correction and noise sources.

However, there are many ways to tackle this problem, focusing on time or frequency domain analysis, fitting signals to a basis set, singular values decomposition (SVD), principal component analysis (PCA), machine learning, or peak integration, to name a few methods. Quantification can be reported as a relative value by using a common metabolic marker that is considered to have a stable quantity, such as NAA, Choline (Cho) Cr or as an absolute value by using the water signal as the reference.

Due to the complexity of MRS, quantification is a challenging problem; with a very low SNR, 10,000 times weaker than water, spectra that frequently overlap and obscure features, with a large variety of noise sources, and with calibration and environmental factors additionally distort the spectrum. Furthermore, additional sources of distortion can come from spectral processing methods, such as the coil-combine algorithm, removal of the residual water signal, baseline correction or phase correction when incorrectly applied, often requiring an experienced operator to manually correct. To further complicate the matter, quantification methods that utilise a basis set require a separate basis for each pulse sequence, timing of the sequence (e.g. the T_E for PRESS), scanner B_0 strength, and even vendor pulse sequence implementation as these can vary [125].

An ideal quantification method should aim to be quick and automated, enabling a clinician to quickly receive results, reducing workload and negating the need for specialist training to use the software. Similarly, it should also be verifiably accurate and precise across a broad range of spectra, removing doubt in the accuracy of the results. However, current quantification methods can be complex as a consequence of the difficulty of the task, requiring experienced users to achieve accurate results. Many methods have an extensive range of options to alter the quantification method, altering the results, and potentially introducing a source of bias.

Methods that utilise basis sets as prior information have additional hurdles to overcome when collecting the basis sets. These basis sets are difficult to acquire; experimental basis sets are time-consuming to create and suffer from the same environmental issues in their spectra, whereas simulated basis sets are comparatively quick to simulate, they suffer from the aforementioned simulation challenges in Section 2.2.4. Methods that employ basis sets usually come bundled with a simulator to generate them.

Due to the range of presented challenges, quantification can be an extremely time-consuming task, where time must be spent acquiring or simulating basis sets, and adjusting fit, processing algorithms and parameters. In the following section, there will be a brief background that describes commonly used software packages available for quantification and simulation, with a summary provided in Table 2.1.

2.3.1 *Metabolites of Interest*

A metabolite is a molecule that is involved with metabolism. They are usually associated with a function, such as signalling, growth or fuel. The role of metabolites *in-vivo* is complex and interlinked with multiple processes. This complexity makes understanding a metabolite's function and its overall role in the human body a significant challenge. The study of metabolites attempts to untangle these complex processes, how they interact with each other, and how metabolite-related disorder can have knock-on effects. This thesis focuses on five metabolites that are of particular interest in MRS community for their links with metabolic function and the insight they can provide in other areas. All metabolite models in this section have their chemical shift (ppm) and J-coupling (Hz) values from Govindaraju *et al.* [45].

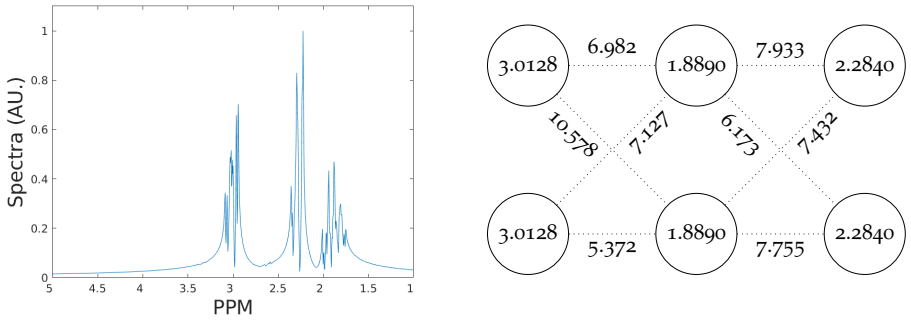


Figure 2.6: γ-aminobutyric acid (GABA) spectra (left), chemical shift (ppm) and J-coupling (Hz) network (right).

γ-*aminobutyric acid (GABA)*

GABA (Figure 2.6) is the primary inhibitory neurotransmitter [88]. Its quantification is especially difficult as it is present in low concentration *in-vivo* and its characteristic MRS features overlap with those of much more abundant metabolites, such as NAA and Cr, obscuring its signature in MR spectra. This had lead to the development of new spectroscopic

techniques, such as edited spectroscopy, that attempt to selectively edit out certain features to make others observable. GABA is of particular interest in psychology due to the wide-reaching effects when there is a disorder in GABA levels. GABA levels have been correlated with many disorders, such as anxiety [82], depression [126], schizophrenia [87], obsessive-compulsive disorder [132], sleep disorder and post-traumatic stress disorder [93], and social function in autism [67].

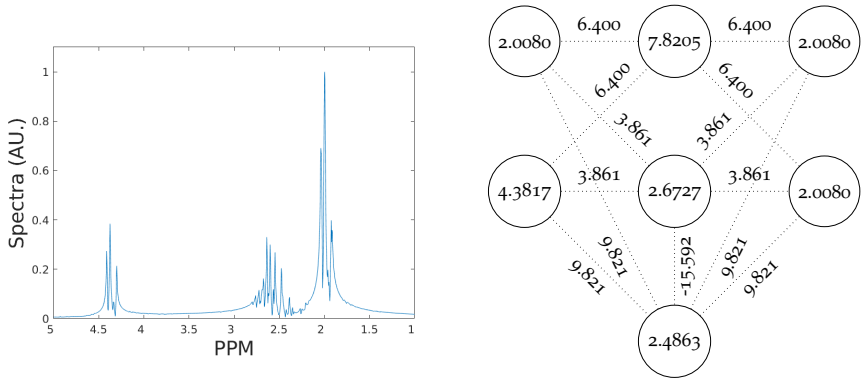


Figure 2.7: N-Acetylaspartic acid (NAA) spectra (left), chemical shift (ppm) and J-coupling (Hz) network (right).

N-Acetylaspartic acid (NAA)

NAA (Figure 2.7) is the second most prevalent metabolite in the brain. It has a clear, distinct peak around 2 ppm that dominates the spectrum and is often used as a reference for relative quantification. Because of this, any improvements seen in the accuracy of NAA quantification has direct effects on the quantification of all other metabolites. NAA is typically associated with the repair and building of neurons in the brain, and as such reduced levels are seen in patients after traumatic brain injury [97], patients with early stages of Parkinson's disease [105] and Alzheimer's disease [44]. Similarly, increased levels of NAA have links with improved neuronal viability in old age, with increased concentrations found as a consequence of aerobic exercise [35]. In cancer research, increased NAA in cancer has been linked with a reduced survival rate [170]. For psychological disorders, decreased NAA has been linked with schizophrenia, schizoaffective disorder and bipolar disorder [61, 69].

Creatine (Cr)

Cr (Figure 2.8) facilitates recycling the energy currency of cells and is taken up by tissues with high energy demands such as muscles and the

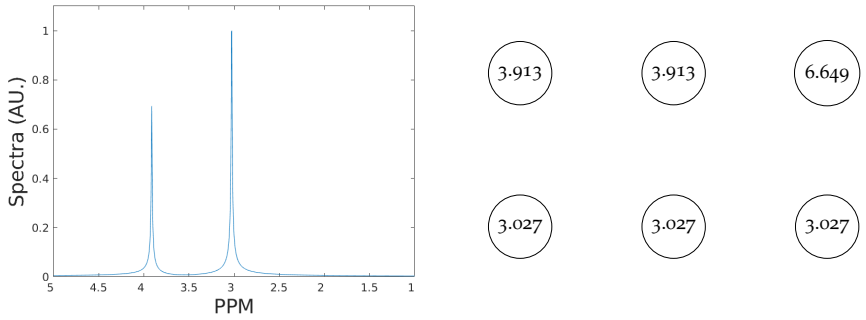


Figure 2.8: Creatine spectra (left), chemical shift (ppm) and lack of J-couplings.

brain. Cr is unusual when compared to the other metabolites that have been selected, as it has no couplings creating straightforward spectra without any split peaks caused by J-coupling. Cr is often used as an internal reference compound, much like NAA due to the well-resolved peaks and high temperature and pH stability.

Due to the importance of the role of Cr, reduced concentrations are usually a symptom of a much larger problem as Cr is predominately synthesised by the body. The effect of this deficiency is incredibly broad and as such as a parent term for disorders stemming from it: Creatine deficiency syndromes (CDS). CDS has links with a large range of disorders, such as epilepsy, intellectual disability, muscle disorders and gastrointestinal problems among others [80, 85].

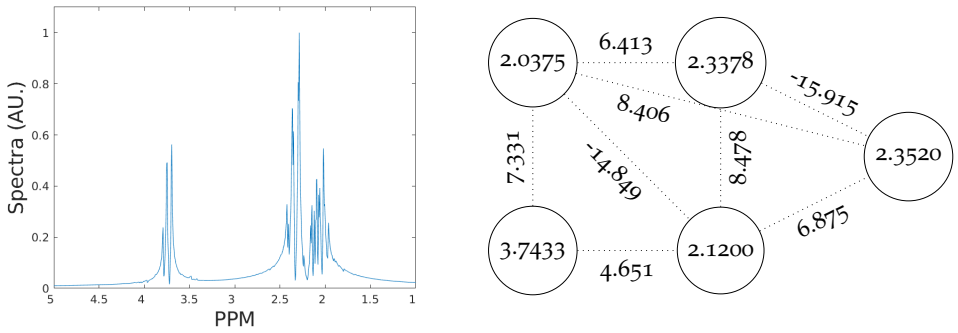


Figure 2.9: Glutamate spectra (left), chemical shift (ppm) and J-coupling (Hz) network (right).

Glutamate (Glu)

Glu (Figure 2.9) is the most abundant neurotransmitter which primarily serves an excitatory function, and due to the prevalence, it is linked to a wide range of biochemical processes. The excitatory process in excess has been linked to cell death through a process called excitotoxicity [173]. Glu

receptors are linked to cancer growth, particularly melanoma opening new treatment pathways to reduce tumour volume [143, 149]. The lack of Glu specific studies is covered below under Glx.

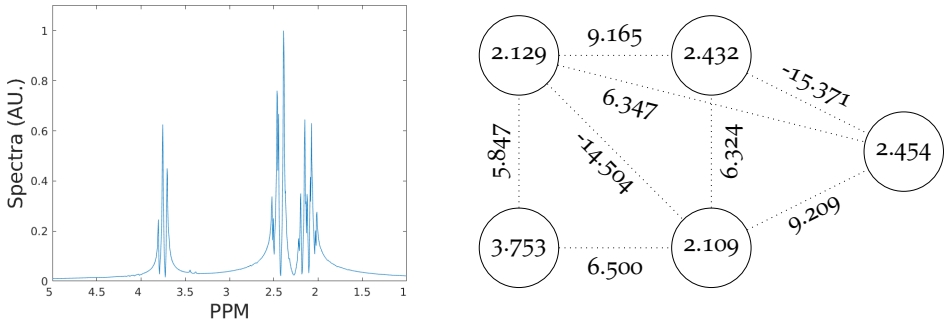


Figure 2.10: Glutamine spectra (left), chemical shift (ppm) and J-coupling (Hz) network (right).

Glutamine (Gln)

Gln (Figure 2.10) is the most abundant metabolite in human blood. Gln has an incredibly similar structure to Glu, and is a chemical precursor to Glu. Like Glu, Gln is also linked to a large range of metabolic processes. As before, the lack of Gln specific studies is covered below under Glx.

Glutamine (Gln) and Glutamate (Glu) combined (Glx)

Glx is not single a metabolite, but Glutamate (Glu) and Glutamine (Gln) combined. Glx is frequently used in the MRS community to describe the combined signal from both of the metabolites. This stems from the inability to accurately separate them due to their highly similar chemical structure and spectrum, particularly at 3 T and lower. Due to this, studies performed at 3 T and lower that individually quantify Glu and Gln are hard to find and are problematic with regards to their accuracy [90].

However, Glx is resolvable and is commonly studied with an excellent review by Ramadan *et al.*Ramadan2013a. Increased Glx has been linked with bipolar disorder [42], restless legs syndrome and sleep disruptions [96], and severe symptoms of attention-deficit hyperactivity disorder in adults [85]. Decreased Glx and NAA linked with autism spectrum disorder [153] and major depressive disorder [81].

In order to display the difficulty of quantification for this limited set of metabolites, Figure 2.11 displays a simulated spectrum. This spectrum has no noise and has a very reasonable FWHM linewidth of 2 Hz; concentration values are taken from Govindaraju *et al.* [45] for the human brain.

(a) Individual metabolite spectrum (b) Combined spectra

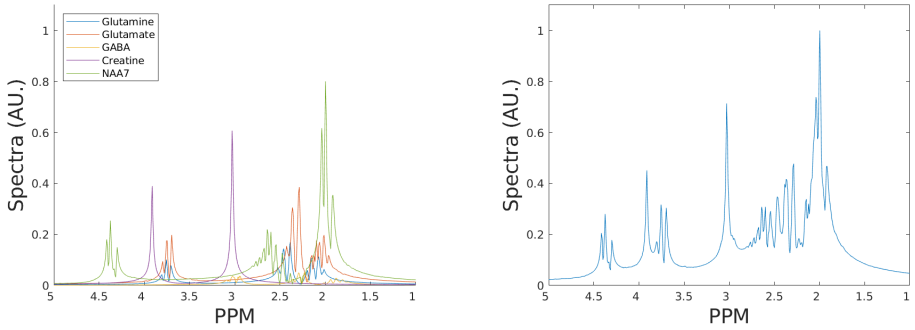


Figure 2.11: Simulated combined FID spectra, 3 T, 2 Hz full-width half-maximum (FWHM), no added noise. Concentration values from Govindaraju *et al.* [45], note how relatively low the GABA signal is.

This FID spectra is very idealistic for MRS and is only intended to display the overlap between this small subset of metabolites. In reality, these spectra are substantially more complex, with many more metabolites found *in-vivo*, with external nose, and a substantially broader FWHM.

These metabolites have been chosen due to their importance in addition to the difficulty to quantify them. They encompass the current challenges in MRS quantification well with two distinct challenges of accurate GABA quantification in the presence of Cr and NAA, and secondly the separation of Glx at 3 T. These two targets are later used in Chapter 5 as they are excellent for showcasing the method used to optimise the RF pulses.

It is clear to see the importance of studying these five metabolites due to the wide-reaching impact any accuracy improvements to quantification would bring. However, there are many more metabolites in the human body, each of which will have a particular contribution to metabolism. Methods explored later in this thesis are not restricted to these five metabolites, but they are utilised to display the utility of these methods to current relevant research.

2.3.2 *Basis Sets*

Basis sets are sets of metabolite signals that are used by methods that attempt to quantify spectra as an inverse problem, by fitting the known set of signals to quantify the spectrum. Selection or creation of the right basis set is itself an incredibly complex task, a basis set that poorly represents experimental data will ultimately lead to a poor fit. To further complicate the matter, separate basis sets are required for each combination of pulse

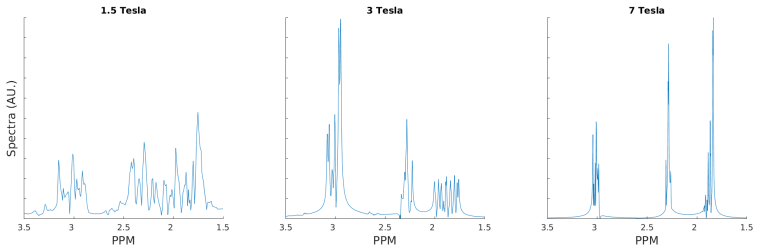


Figure 2.12: Effect of static field strength B_0 for simulated GABA PRESS spectra. Absolute spectra shown.

sequence, B_0 field strength, pulse sequence timing (T_E, T_R) and in some instances scanner manufacturer as each of these factors alter the resulting spectrum in a variety of ways.

Basic simulators for basis sets often assume the molecule is prepared in a well-defined initial state such as the ground state. They approximate the pulse sequence by a series of instantaneous unitary operations acting on the state at certain times, followed by readout of the MR signal, to simulate the full MRS pulse sequence. More sophisticated simulators perform time-resolved calculations for finite-duration pulses and include the effects of relaxation or field inhomogeneities. However, the choices of pulse shapes and pulse timings in vendor-specific implementations of the MEGA-PRESS sequence differ [95] and are often not known precisely. Therefore, quite often educated guesses must be made. Similarly, other parameters, such as relaxation parameters or field inhomogeneities, are typically not known precisely. Moreover, even metabolite models, including chemical shift and J-coupling parameters, still have uncertainties [70]. A major source of these parameters for common metabolites in MRS is the landmark paper by Govindaraju *et al.* [45], but alternative models have been suggested, e.g., for GABA by Near *et al.* [103] and Kaiser *et al.* [60].

Acquisition of an experimental basis requires the creation of a range of calibrated phantoms, with the correct pH and temperature and additionally added chemical markers, such as sodium trimethylsilylpropanesulfonate (DSS) to provide a reference signal. Phantom creation is typically done for each metabolite individually to get a pure signal and later scaled against the additional reference signal. For efficiency, a large range of pulse sequences and timing variants should be scanned at the same time, as phantoms typically have a short shelf life and are usually discarded after use, requiring expensive large blocks of uninterrupted scanner time. Experimentally collected basis sets often suffer from the same general MR issues, where instabilities in the scanner, calibration issues and noise

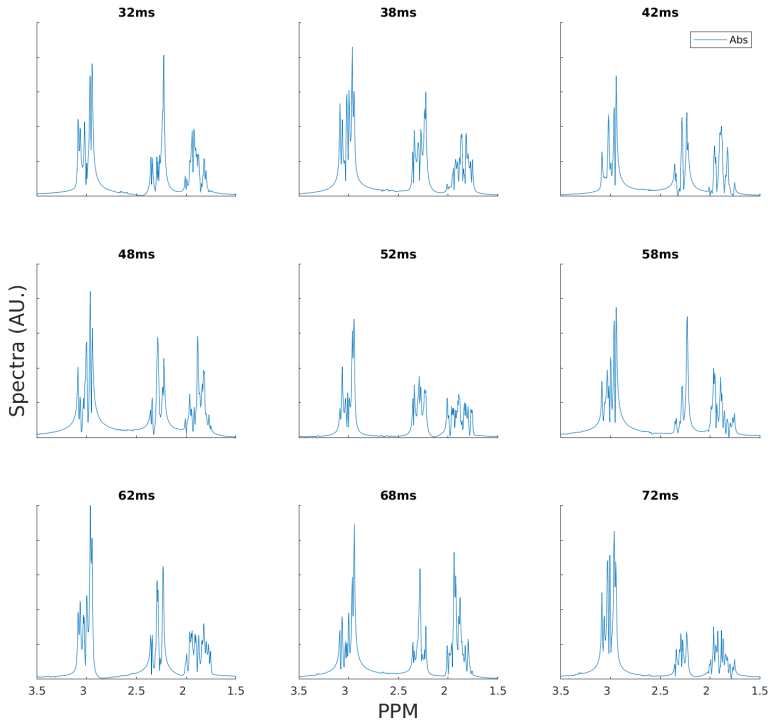


Figure 2.13: Effect of varying echo time (T_E) times for simulated GABA PRESS spectra. Absolute spectra shown.

can distort the resulting spectrum, ultimately later influencing the quality of the fit.

Simulation is the alternative to experimental basis set creation and is typically preferred due to the substantially shorter time needed, lower cost, and it removes the noise issues found experimentally. Simulation enables alteration of parameters and is typically quick (> 60 seconds) to simulate for metabolites of up-to seven spins. However, mimicking the MR environment and pulse sequences is not a trivial task, and is further complicated with the aforementioned issues of the accuracy of metabolic model parameters for simulation [70].

2.3.3 *State-of-the-art Quantification Methods*

This section covers a selection of the most popular quantification software. Table 2.1 shows a summary of the quantification tools mentioned along with their approach to solving the issue. The most popular method is to utilise a basis set, combined with a variant of the non-linear least squares (NLLS) algorithm to solve the inverse problem. However, basis sets and their creation presents additional challenges, where simulation can be inaccurate and experimental basis sets suffer from the same general issues as MRS.

Table 2.1: Summary of currently available quantification methods and software.

	Basis set	**Peak integration**	**Machine Learning**
Frequency Domain	LCModel TARQUIN AQSES† VeSPA	GANNET	MRSNet
Time Domain	QUEST†		

† Denotes an analysis method that is part of the Java-based MR user interface (jMRUI) software package.

Linear combination of models (LCModel) [114, 115]

LCModel is widely considered to be the gold standard of spectral quantification. It uses the Marquardt modification of the Gauss-Newton NLLS algorithm to decompose spectra in the frequency domain and spline fits for removal of macro-molecule baseline signals. It also uses additional

terms with normal distribution priors to correct for each metabolites' T_2 relaxation and frequency drift.

LCModel is commercial, Linux-based and costs $13,200 per non-commercial license (as of September 2019) [148]. LCModel is a black-box approach, with few options to modify the process available to the user, with the general focus on simplicity and robustness through automation. Basis sets are provided by email from the creator or suggested external sources, as LCModel is only able to simulate macro-molecule contributions, not complete basis sets.

Totally automatic robust quantitation in NMR (TARQUIN) [120, 166]

TARQUIN is an MRS analysis tool designed for short echo sequences, written in C++ under the GPL license and cross-platform on Windows, Mac and Linux. It aims to solve quantification in the time domain using the Lawson-Hanson non-negative NLLS [76] algorithm. Automated phase and frequency correction is performed, along with the removal of the water peak using HSVD (Hankel Singular Value Decomposition) [5, 108]. Much like LCModel, TARQUIN focuses on automation, with both a graphical user interface (GUI) and command line interface (CLI) for batch processing of spectra. TARQUIN can simulate and export basis sets, with an inbuilt density-matrix simulator.

Java-based MR user interface (jMRUI) [147]

Java-based MR user interface (jMRUI) is a vast magnetic resonance imaging and spectroscopy (MRI/S) toolbox with a range of quantification, simulation and processing methods available. jMRUI is free to use for non-commercial applications and is available on Windows and Linux. jMRUI has a large range of available algorithms, including the NLLS based advanced method for accurate, robust and efficient spectral fitting (AMARES) and automated quantification of short echo time MRS signals (AQSES), and the SVD based HLSVD (Hankel Lanczos Squares Singular Values Decomposition), HTLS (Hankel Lanczos Total Least Squares), LPSVD (Linear Prediction Singular Value Decomposition) and quantification based on quantum estimation (QUEST).

AQSES [41, 112] and QUEST [47] are the two most predominantly used methods from the jMRUI toolkit. AQSES is a method designed for short-time echo sequences, uses a macromolecular baseline and automatically removes any residual water signal. QUEST is a time-domain

fitting method, utilising the simulators bundled with jMRUI for basis set creation.

Versatile simulation, pulses and analysis (VeSPA) [140]

Versatile simulation, pulses and analysis (VeSPA) is a complete software suite for simulation of spectra, pulse sequences, creation of basis sets and spectral analysis. It is open-source, written in Python 2 and can be run on Windows, Mac and Linux. VeSPA quantifies spectra in the frequency domain, using the NLLS algorithm. The spectral quantification is complex with a large range of options available to the user. Simulation of basis sets uses the PyGamma library, a Python wrapper on the C++ NMR simulator GAMMA [116]. The internal basis sets are stored as sets of peaks with frequency, magnitude and phase components that can be used to re-simulate the spectra with different lineshape parameters and models including Gaussian, Lorentzian and Voigt, this allows for an incredibly flexible fit model.

GABA-MRS analysis tool (GANNET) [31]

GANNET is an open-source tool, dedicated to automated GABA quantification for edited spectroscopy, e.g. the MEGA-PRESS or HERMES pulse sequences. It is written in MATLAB, so requires a software license to MATLAB and a set of additional toolboxes to run; however, it is open-source, and it can be run on Windows, Mac and Linux. GANNET fits peaks in the frequency domain using the NLLS algorithm. It fits a Gaussian peak for the 3 ppm GABA signal and a Lorentzian for the 3 ppm Cr. Currently, it can only run with a single target per run from: GABA, Glx, 'GABAGlx', Glutathione (GSH) and Lactate (Lac) for the MEGA-PRESS pulse sequence. It cannot quantify other pulse sequences, such as PRESS or STEAM. GANNET incorporates a range of pre-processing tools that automate ingress of raw time-domain data, performing channel combination, frequency and phase corrections and outlier rejection.

2.3.4 *Review of the state-of-the-art*

The motivation for this section comes from the clear disconnect in current methods, and the need to improve on current standards. Studying the performance of these tools is intrinsically difficult, as *in-vivo* data has no associated ground truth, so the majority of studies measure performance comparatively between tools. There are excellent reviews of these tools

available here [47]. The following section is a review of literature that has benchmarked current methods, motivating the work reported in Chapter 3.

Comparative studies

There has been a limited amount of work comparing the performance of these methods. [106] studied the reproducibility of quantification of GABA with MEGA-PRESS at 3 T for LCModel, jMRUI and Matlab (GANNET) using 28 healthy volunteers. Results show that LCModel is the most consistent at reproducibility with Glx= 6%, GABA= 7%, GANNET came second with Glx= 9%, GABA= 12% and jMRUI (AMARES) came in last with Glx= 18%, GABA= 9%. This work is obviously missing ground truth data as it has been done *in-vivo*. It also does not list the actual concentrations reported by the tools to check if they report similar values. This result can show how precise these tools are, but not how accurate.

TARQUIN was compared to LCModel across a range of *in-vivo* and simulated PRESS 1.5 T and 3 T data sources. It was reported that it is "acceptable for most purposes" and "in agreement with the popular frequency domain fitting program LCModel" [166]. However, upon looking more closely at the data contained in the paper, they note a $< 20\%$ error for Cr, TNAA, TCho and Ins with a $< 40\%$ error for Glx in 95% in agreement between LCModel and TARQUIN for *in-vivo* data. This *in-vivo* comparison with LCModel has no ground truth data along with it, the reported metabolite concentrations in mM are compared to reference values for healthy males in the same age range. There is mention of a Monte-Carlo spectra simulator used to verify both LCModel and TARQUIN. However, only one fit is shown as a figure that appears to be fairly accurate.

This study reports that TARQUIN performs in the same region as LCModel on a range of *in-vivo* data, but it fails to assess the accuracy of both of these methods, which could have been investigated more thoroughly with their Monte-Carlo simulator.

In the literature, GANNET is cited to have "approximately equal" performance to LCModel, jMRUI and TARQUIN [101], when discussing the estimated GABA concentration for MEGA-PRESS spectra. This study focuses on the quantitative analysis

Despite the success and considerable improvement of quantification methods, it has been shown that there are significant errors in quan-

tification between the different methods [62, 94, 129], in addition to disagreement with phantoms of known composition [57]. These studies highlight the need for improvement, for these methods to be reliable enough for general clinical use.

The studies presented in this section have highlighted the challenges in benchmarking quantification methods. There is no defined method for benchmarking these tools, and as a consequence, it is hard to understand the relative performance of these methods. While on paper they all perform well, the majority of them compare their performance to LCModel in simulation or with *in-vivo* data, which is either not representative of the real spectra found, or lacking ground truth data. This shows the need for more reliable methods for quantification of MR spectra for MRS to become a reliable tool for clinical use, especially for difficult-to-quantify metabolites such as GABA.

2.3.5 *Machine Learning for Quantification*

The use of machine learning for MRS quantification has only started to emerge in the last few years. Initial work was done by Das *et al.* [26], focused on the use of random forests trained on 1 million simulated PRESS (3 T, $T_E = 35$ ms, $BW = 2500$ Hz) spectra, and tested on 287 experimental spectra. This work showed that quantification could be performed quickly, with minimal interaction from the user, and can account for lipid and a macro-molecular baseline. Results from this study are unclear, other than the proposed solution is 'comparable' to LCModel.

This was followed by Hatami *et al.* [51], which was the first use of a CNN for MRS quantification. They proposed a framework for simulating and generating large numbers of PRESS (3 T, $T_E = 30$ ms, $BW = 4000$ Hz) spectra for training the CNN. This paper was compared to the QUEST quantification method from jMRUI and their own random forest implementation of the work done by Das *et al.* [26]. They were able to show that a CNN can outperform the random forest and QUEST quantification methods.

Finally, recently Lee *et al.* [78] trained a CNN on 40,000 PRESS (3 T, $T_E = 30$ ms, $BW = 2000$ Hz) simulated spectra to quantify a range of simulated and *in-vivo* data from five healthy volunteers. They utilised zeroth-order phase correction from jMRUI for their experimental data, along with the Hankel Lanczos squares SVD (HLSVD) filter for removing the residual water signal. They highlighted the performance benefits of

utilising CNNs for quantification and the impact that it would have on a clinical application. This study further confirmed that a well-trained CNN has the ability to outperform LCModel.

In NMR a method utilising Bayesian inference has been developed for the analysis of time-domain signals [165]. This method leveraged prior knowledge with associated probability distributions and estimates 'nuisance parameters' such as phase-shifts, decay constants and noise parameters, outperforming Fourier based NMR methods. Due to much higher spectral resolution and the focus on molecular structure elucidation, the issues and applications of high-field NMR are different, and the methods have not been applied to MRS quantification *in-vivo*.

2.4 ARTIFICIAL NEURAL NETWORKS

This section is a brief introduction into neural networks necessary for Chapter 4, where a convolutional neural network (CNN) is introduced for spectral quantification. This section starts at the foundations of modern neural network (NN)s, with artificial neural network (ANN)s. ANNs are graph-like like structures that are intended to mimic the human brain. These networks operate by taking a "sensory" input and returning some output with the aim of solving a specific problem. This output is usually tied to a classification or regression problem, depending on the desired use-case of the network, such as guessing the genus of a flower based on a photograph.

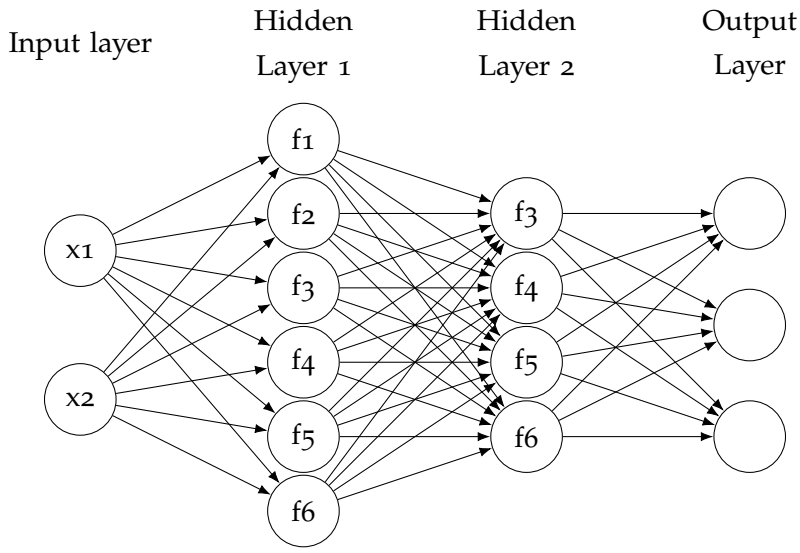


Figure 2.14: An example artificial neural network (ANN) with two hidden layers, two inputs and three outputs.

Figure 2.14 shows a simple ANN structure, with two inputs [x1, x2], two hidden layers and three outputs. All nodes are fully connected to every other node in the next layer, with data flowing from left to right. In this simplistic model, the values seen by each node is the sum of its inputs. Activation of nodes is typically based on a threshold value, with the output edge having a weight scaling the strength of the signal, determining any modulation of the value between nodes. These internal values are usually randomly initialised and optimised through several rounds of training.

A 'deep' neural network typically refers to a ANN that has multiple hidden layers in its architecture. The term is loosely defined but is used to generally describe modern network architectures, such as ResNET [77], which has a 152 layer implementation. Deep networks are able to perform much more complex tasks than previously possible. This is in part due to improvements in computing power and leveraging modern graphics processing unit (GPU)s to reduce training time greatly, along with new layer types and architectures.

There is a large range of other types of neural networks, such as recurrent neural networks (RNN), generative adversarial networks (GAN) and radial basis function networks (RBF) among many more.

2.4.1 *Network Architecture*

The architecture of a NN should be tailored to suit the specific problem it is trying to solve. It is important to recognise what information needs to be preserved in the data that is being passed through the network. NNs have evolved a long way from the early fully-connected ANN architectures, with an ever-increasing range of layers available for network design. This range of NN layers are sandwiched together into a near limitless configuration, limited by computational power. These layers do not necessarily feed directly into each other, there can be forks and joins, with multiple outputs for different trained tasks, creating complex dataflows. Designing modern NNs can be challenging, typically requiring a large number of iterations on the architecture before a 'good' solution is found.

2.4.2 *Network Layers*

A full in-depth discussion into NN types and layer types is out of the scope of this thesis, but there is an excellent resource by Gu *et al.* [172] that covers Recent Advances in NNs including layer types. This subsection will briefly go over a few key components that are used later in Chapter 4.

Fully connected (FC)

Fully connected layers are where there is a connection between every node in the current layer and the next. These have more recently been described as 'classification layers' in CNNs but were historically the only component in early ANNs. The use of fully connected layers has generally been phased out of modern CNNs as it has been shown that a fully connected layer can be represented by multiple, much faster and memory-efficient convolutional layers [144].

Convoloutional Layers

Convolutional layers are arguably one of the most significant developments for NNs in recent years. Convolutional layers lead to the creation of convolutional neural network (CNN)s, which recently have dominated current NN research. Convolutional layers operate by providing each neuron with a receptive field of the input, such as a (3×3) grid of a photo rather than a single pixel. This allows the node to obtain a broader context than has previously been possible.

Pooling

Pooling layers can be thought of analogous to compression, where a range of inputs is averaged, or max pooled to form the output. Similarly to convolutional layers, pooling layers also have a receptive field such as (2×2) or (5×5). Pooling layers are typically used to reduce the amount of data in the network, speeding up computation, and aiding in noise reduction by reducing unnecessary data, while retaining the necessary information for the desired task.

Activation Layers

Activation function layers are used to alter the data in the network by applying a non-linear function. There is a broad range of functions

that are typically used, with rectified linear unit (ReLU) [102] being one of the most popular functions. When choosing activation functions, it is important to consider the effect these have on the gradient back-propagation, ReLU, for example, is not differentiable when the input is $\geqslant 0$. This stops gradient back-propagation in its tracks leading to no learning when values are $\geqslant 0$ inside networks, leading to the creation of leaky ReLU (LReLU) [169] addressing this problem.

2.4.3 *Hyperparameters*

Hyperparameters broadly define any global variables that affect the network, they are variables such as the learning rate (how much the network will alter weights in the nodes) and the batch size (how many training examples are fed to the network at once for training). Each of these values is usually optimised in parallel (either manually or automatically) when testing different network architectures, as each network will perform differently based on these parameters. Poorly tuned hyperparameters will cause a network to never correctly learn and could mimic a network architecture that is not fit for purpose. Optimisation of these parameters can be lengthy depending on the size of the input data, requiring a new network to be trained each time.

2.4.4 *Loss Function*

A loss function is required to train a neural network; it is a mathematical description of how well the network performs a particular task. Typically, the output of a neural network is a vector with each value corresponding to an output node, and the loss function is used to compare this output with a vector that contains the ground truth. Choice of loss functions is highly dependent on the problem at hand, such as using cross-entropy for classification problems or mean squared error (MSE) for regression problems. Similarly to the activation functions, the choice of the loss function directly impacts the gradients seen in backpropagation and should be carefully considered.

2.4.5 *Optimisation*

Training a NN in the context of this thesis is focused on gradient back-propagation from supervised learning. Supervised learning requires a

training set with the associated ground truth data. Simplified, back-propagation works by showing a network an example input, getting the current solution and adjusting the internal values of the network based on the difference with the ground truth. For back-propagation to work successfully, suitable activation functions need to be chosen in the network that has exact numerical gradients supplied.

Backpropagation struggles with deep networks with exploding and vanishing gradients, where any error in the back-propagation destroys the accuracy of the gradients; in extreme cases, this will overflow to NaN. This leads to models having issues gaining traction and never accurately learning. There are many methods to combat this issue, such as architectural changes, weight regularisation and gradient clipping, but these are out of the scope of this thesis.

2.4.6 *Network Dimensionality*

Chapter 4 primarily deals with 2D neural networks, but it is worth briefly mentioning that there are other network architecture dimensions available. The dimensionality of the network refers to the organisation of the data and how this is handled internally. The dimensionality of the networks is entirely dependent on the type of input data and the goal for the network. For example, it may make sense to split a 2D image into multiple layers for individual colours if the task performed by the network needs to be particularly sensitive to colour. Similarly, if the task is detecting movement in the video, it might make sense to have the input be black and white single frames that are stacked into a 3D rectangle.

2.5 OPTIMAL CONTROL AND OPTIMISATION IN MRS

This section introduces the notion of optimal control in the general case as a primer to Chapter 5. It is also contextualised with a literature review of current optimisation found in NMR and MRI/S. We find that control and optimisation in the MR context has a broad range of meanings. For clinical MR it is typically related to protocol enhancements, and in NMR it is optimal control. Finally, optimal control is presented in relation to quantum systems, with examples of what can be achieved based off of current literature.

2.5.1 *Optimisation*

In this thesis, optimal control is concerned with driving a system (metabolite) from an initial state to the desired state at a prescribed time. To achieve this, metabolites are simulated in a MR environment, with the optimisation adjusting the amplitudes for the RF used to control the rotation of the spins. This provides the ability to steer the system towards the desired state. The desired state corresponds to a complex excitation state, for a pre-determined spectrum that the metabolite will exhibit. This methods essentially tailor the resulting spectrum. A much more in-depth explanation of this method can be found later, in Chapter 5.

Optimisation in MR is a broad topic, ranging from new clinical MRI pulse sequences, energy reduction techniques to RF coil design [110, 151] and advanced decoupling sequences in NMR. To simplify, this section, it is broadly split into clinical and non-clinical MR, due to the vastly different requirements for development and implementation in these areas. Clinical MR optimisations are typically focused on protocol enhancements and modifications that provide new features, such as the development of MEGA-PRESS, which is an alteration of the PRESS pulse sequence but enables spectral editing. Generally, the difficulty of the clinical MR environment makes it difficult to optimise controls for, with comparatively low magnet strengths, poor B_0 homogeneity, and scanner energy limits must all be taken into consideration when designing pulse sequences.

On the other hand, non-clinical MR provides a better platform for experimentation with control techniques, when compared to the clinical MR environment due to the increased level of control and general robustness due to the substantially stronger, more homogeneous magnet with a larger range of coil and gradient channels available that can reproduce controls much more accurately. Additionally, these machines are designed with higher energy limits in mind, as they do not need to be safe *in-vivo*. MR controls can be 'squeezed' to produce the same results in a much shorter time, at the cost of higher energy, but with the benefit of less time for the environment to disturb the control and spin states. Unfortunately, due to the high level of control achievable in NMR, these optimisation results typically do not translate back to clinical MR, as the lack of stability degrades the accuracy of the controls.

Clinical MR optimisation

There has been a diverse range of optimisations in clinical MR in recent years, ranging from RF coil design to improve B_0 homogeneity [168], open RF coil design to improve musculoskeletal MR imaging [39] and design of asymmetric gradient coils for improved performance [151]. In protocol enhancements and optimisation, work is generally focused on a specific use case, such as optimisation of MRI protocols for detection of lesions in multiple sclerosis [156] or the visibility of the appendix in pregnant women with suspected appendicitis [130]. For general imaging, recent work has been focused on optimising the RF pulses to improve contrast and SNR [118, 159], spatially selective RF [162] and optimisation of RF and slice-selective gradients together [124].

There is minimal research that focuses on optimisation for MRS due to the aforementioned issues in Section 2.5.1, as this is typically done in the NMR context. However, advances have been made in speeding up 3D spin-echo sequences [100] and the JPRESS pulse sequence [10], and dynamic nuclear polarisation for ^{13}C [163]. For pulse sequence optimisation, work has been done suppressing NAA for Glu and Gln detection by optimising a PRESS pulse sequence [3].

Work presented later in Chapter 5 appears to be the first attempt at optimising pulse sequences for metabolite states focused on a clinical MR environment.

Non-clinical MR

NMR has a plethora of optimisation based research, covering a broad range of topics. For RF detectors, work has been done improving the spectral linewidth with in-situ detectors [37]. Design and optimisation of broadband excitation pulses in NMR has been very successful, focusing on improving performance and reducing the total energy of the pulse [28, 65, 68, 134–136], selective decoupling [33] and band-selective excitation [137]. More recently, work has been done investigating double-quantum spectroscopy to improve ultrafast 2D NMR for analytical chemistry [123] and optimisation of the localisation routines for NMR field cameras, that are used for mapping and calibrating the B_0 field [19].

There is an excellent summary of techniques for NMR simulation and optimisation in [160]. This summary provides an excellent background into general spin control, and is framed in the context of utilising NMR for quantum computing by creating quantum gates through optimal control techniques. While this work is similar to later work in Chapter 5,

it is focused around spin manipulation in a highly controlled environment for single spins. The work presented later is focused on concurrently controlling multiple metabolites, each of which have many coupled spins in a complex and unstable environment.

2.5.2 *Optimisation of Quantum Processes*

Control of quantum processes is an ideal application of optimal control theory, steering a dynamical system towards a desired state. Typically these applications are developed on simulations of quantum processes, as large-scale reliable quantum systems are incredibly hard to produce experimentally, with many hurdles to overcome in manufacturing processes and material research to become viable. Despite this, there are numerous applications of quantum control that are applicable and realisable now, such as MR. MR is already a quantum control problem as it capitalises on the quantum mechanical property of spin. Application of quantum control techniques to MR is typically overlooked due to the increased difficulty due to the large role of the environment. Later, in Chapter 5 we display a method that optimises controls with knowledge of instabilities in the environment to great effect.

Outside of clinical MR, optimal control of other quantum systems has had numerous successes with control of spin systems, from optimal information routing in spintronic networks with the use of static bias controls [74], creation of high precision sensors [109] and, importantly, has been shown to work well within the rotating wave approximation (RWA) approximation [63]. An excellent review that benchmarks control and optimisation methods applied to a range of open and closed quantum systems, with varying targets is provided by [83]. Additionally, there is an excellent review by Wershnik *et al.* [164] that covers quantum control utilising lasers and is contextualised with examples in quantum chemistry.

In the context of this thesis quantum control techniques are utilised in Chapter 5 to optimise RF pulses for metabolites. This provides the ability to steer the system with the knowledge of the underlying system dynamics towards a desired target state. The result of this is the ability to prescribe a state that would produce desirable spectra that would be easier to analyse.

There is a range of optimal control methods that have been developed for quantum systems, from the gradient free Nelder-Mead with CRAB [117],

the first-order GRAPE [65] method, second-order Broyden–Fletcher–Goldfarb–Shanno algorithm (BFGS) adaptation of GRAPE [28], to hybrid methods such as GROUP [142] which combines GRAPE and CRAB.

Of these methods, there are a few schools of thought on control update schemes, where control amplitudes are adjusted either sequentially as with the Krotov-type methods, in parallel as done with GRAPE, or hybrid update schemes which update subset of control amplitudes [83]. Generally, of the three methods, a parallel update scheme, combined with limited-memory BFGS (L-BFGS) is usually preferred as it is able to outperform the other methods [83].

These optimisation methods have been explored in the NMR context [53, 64, 65, 160], but they have not been studied in the clinical MR context likely due to the difficulty of the environment.

2.6 SUMMARY

MR is a rich and diverse field, with a wide variety of potential avenues for improvement, with spectroscopy being one key area of many. The applications of MRS are wide-reaching, impacting many areas of research and aiding in medical diagnosis. This chapter has provided background to the major themes of this thesis, providing a sufficient background with relevant links to key studies where applicable. It has introduced a wide variety of topics including the physics of MR, metabolites and the simulation of metabolites, quantification, NNs and finally control and optimisation of MR and quantum systems.

While there has been considerable progress in the accuracy of quantification methods, there is still plenty of room for improvement. It is clear that there is no generally accepted standard to benchmark current and new methods. This is problematic, as there is no clear indication for current researchers which tool to use, or if their results will be accurate and reliable. Additionally, due to the large range of quantification options available in some programs, it is incredibly easy to introduce bias accidentally, ultimately skewing the results of the study. The benchmarking issue is addressed in Chapter 3, which ultimately shows there should be a more concerted effort on the whole to validate the performance of quantification methods.

NNs are an incredibly powerful tool with a rich, ever-growing research community to back them up. They are reshaping the landscape of research with new, successful applications being found daily. MRS spectral

quantification could be one such task, enabling methods that are computationally fast, practically simple and require no specialist knowledge to use, only to train. This application is later investigated in Chapter 4, building on the earlier benchmarking work in Chapter 3.

Finally, quantum control has been successful in a variety of fields, and it is an excellent fit for MR. At the core of it, MR is a quantum control problem, but the design of new controls typically uses conventional methods, such as broadband excitation pulses. Utilising quantum control methods to design pulse sequences could bring a new era for MR, much like how the application of fractal mathematics has substantially changed antenna design. This idea is explored in Chapter 5, where these techniques are utilised to create RF control fields for specific MRS targets.

3

BENCHMARKING QUANTIFICATION OF MEGA-PRESS SPECTRA

This chapter takes a deep dive into benchmarking quantification methods, specifically focused on quantifying GABA from spectra that has been acquired with the MEGA-PRESS pulse sequence. We aim to answer the question of "How well do current quantification methods perform?" and investigate "How can we benchmark current methods reliably?". These two questions are essential for improving quantification. The current state of the performance of quantification methods is difficult to navigate, with a distinct lack of studies that comprehensively benchmark multiple tools with experimental data that has associated ground truth data.

Benchmarking in this chapter is done with a set of experimental datasets, collected from calibrated phantoms of known composition. Typically quantification methods are benchmarked using simulated spectra, or have their results compared to LCModel under the assumption that it is a good source of truth. The novelty in this chapter is the fact that there has been no previous efforts to benchmark a broad range of quantification tools using an experimental dataset that has associated ground truth data. This work has a strong secondary goal; to encourage other researchers to use our dataset by publicly releasing it, with the hope others will create and release their own analysis and phantom datasets.

3.1 INTRODUCTION

Quantification is a necessary counterpart to MRS data acquisition, decoding the complex spectra into information about metabolite levels in the ROI. In a perfect world, quantification would be a simple inverse-problem, as the data from any MR scanner would perfectly match simulated signals. However, as introduced in Section 2.3, the reality is that quantification is incredibly complex, with a vast range of sources distorting the spectra. In addition, collection or generation of ground truth basis sets through experimental or simulated methods is similarly difficult, with the former suffering from the same noise issues and the latter struggling to accurately reproduce the experimentally observed spectra.

This proposes a difficult challenge, where designing a universal quantification method is simply not possible due to the wide variety in spectra, scanners and pulse sequences; even within the same pulse sequence for one metabolite (see Figure 2.13). As such a wide range of tools has been developed, with the most popular methods utilising a basis set of known signals to decompose the spectrum. This practice isolates the difficulty of sourcing, or defining the ground truth to base the quantification on, and instead focuses on minimising the distortion and noise found in the spectrum. Other methods, such as peak fitting are typically highly specialised, focusing on a handful of metabolites and pulse sequences, as generalisation is too complex.

As part of the quantification process, it is important to correct issues with the experimental spectra before attempting quantification. Many of these can be expected, such as frequency drifts or incorrectly phased data, due to the way the scanner combines signals from multiple coils; with the best quantification methods being able to handle these distortions successfully. It should also be noted that any improvement seen downstream that would improve the data quality, such as scanner calibration or pulse sequence design, makes the job of quantification easier and, as such, should be a priority.

The ultimate goal of quantification is accurate, precise and timely measurements that are immune to operator bias and easy to achieve universally. This would provide a reliable platform to base medical diagnosis and research on, without the need for a specialist to analyse the incoming spectra. For this to happen, there needs to be a good understanding of the current state of quantification, which methods are performing the best in which areas and what needs to be improved. This requires a comprehensive testing of methods, with high quality test sets covering a wide variety of realistic spectra. This is naturally extremely difficult, with collecting ground truth data associated with *in-vivo* scans is not possible, and developing accurate phantoms time consuming, costly and requires specialist knowledge.

The primary goal of this chapter is to benchmark the current state of quantification for GABA with the MEGA-PRESS pulse sequence, acquired at 3 T. This is done by creating a range of water and gel based phantoms with a handful of metabolites, with the GABA concentration varied across multiple scans. This is followed by quantifying the spectra collected experimentally on a range of popular quantification programs. The benchmarking performed in this chapter is used as motivation for

the work in Chapter 4, where a new quantification method is presented and benchmarked against the results presented here.

We additionally release the experimental datasets, with the aim of encouraging other researchers to utilise these sets for their own benchmarking, and ideally contribute their own. An ideal outcome would be to eventually have a large and varied dataset of spectra with ground truth, collected at multiple sites, on multiple scanners, with a large range of metabolites to be used in analysing new tools. This is common practice in other areas where creation and collection of high-quality datasets is extremely time consuming, such as machine learning where it is used to train and evaluate new models; it would be extremely advantageous to do the same for spectroscopy.

This chapter starts with the methods used to prepare the phantoms, acquire the data from the scanner and how the quantification is performed across all the chosen methods. This is followed by the presentation of results, where there is a discussion on the quality of the spectra produced and the results from the quantification of the spectra with each of the methods. Next is the evaluation and discussion, which covers both an analysis of the phantoms and the quantification, which is broken down on a per-method basis. From here, there is a short section on recommendations for the future of benchmarking quantification, and how the methods presented here could be improved. Finally, there is a summary followed by a conclusion, wrapping up the findings in this chapter.

3.2 METHOD

The work in this chapter is broadly split into two distinct sections: signal acquisition and quantification. Signal acquisition encompasses the entire work flow from creating the phantoms, calibrating the scanner, scanning and transporting the data off of the scanner. Quantification covers all of the considerations that had to be made to attempt a reasonable and fair comparison of the methods.

3.2.1 *Phantom Preparation*

The key focus when preparing phantoms is accuracy, spending time double checking measurements and calculations is essential as the quantities of metabolite powder used is minute. The metabolites and concentrations used were pre-determined, and made in a controlled environment.

(a) Water phantom (b) Gel phantom

Figure 3.1: Water and Gel phantoms. The phantoms are approximately 5.8 cm in diameter for the gel, and 8.2 cm for the water.

A simplified view of the process is: Firstly, the calculation and measurement for the metabolite powder for a volume of distilled water, mixing the solution, followed by pH calibration and temperature measurement, and finally scanning. This process typically takes an hour or two for a water based phantom, for gel phantoms there is additional steps for heating the solution to dissolve the gelling agent, followed by a cooling period which is typically overnight. Once phantoms are prepared, it is important to scan them quickly, as without preservatives, they tend to decompose after a week or two. Although adding preservatives seems advantageous to keep hold of the phantoms, we chose to abstain in order to keep the produced spectra clear of unknown peaks.

A range of phantoms was prepared and scanned over a six month period, three water and a range of gel-based phantoms. Spherical phantom containers were chosen to reduce magnetic susceptibility effects, that can be scanned in any orientation, reduce the size of the air-phantom interface and most importantly to minimise the spectral linewidth.

Phantom Containers

For good magnetic homogeneity, minimised voids and reduced air contact, containers for phantoms should be spherical [**Bluemer2016**, 113]. Using a spherical phantom provides a shape is not only highly regular, but it also enables easy analysis of any distortion seen in preliminary MRI field maps. Additionally, by removing complex shapes it minimises the possibility of any voids or pockets of air in the phantom, that could cause susceptibility affects.

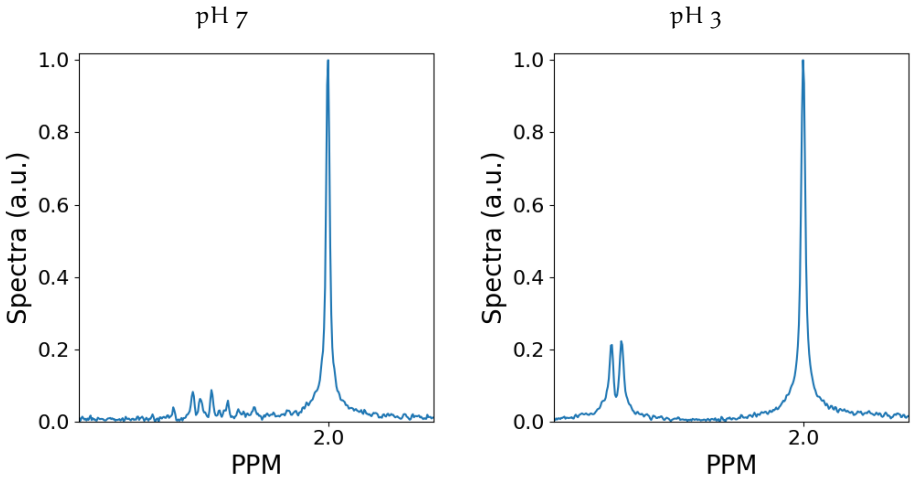


Figure 3.2: The effect of low pH on the NAA spectrum. Edit off acquisition of the MEGA-PRESS pulse sequence.

For our phantoms, shown in Figure 3.1, we used a 290 ml glass, round bottom boiling flask for the water based phantoms, and hard acrylic moulds for the gel based phantoms.

pH *Calibration*

pH calibration is essential to the creation of accurate MRS spectra, where improperly balanced solutions cause peak-drifts [113], such as the 2.3 ppm NAA peaks as shown in Figure 3.2. This is important as the pH of the phantom should match what is seen in the ROI the phantom is mimicking *in-vivo*. For the phantoms used in this work, as they are mimicking a subset of metabolite found in the human brain, the pH should be as close to 7.2 as possible [11].

pH adjustments were made using a 36% hydrochloric acid (CAS-7647-01-0) solution and a 3.99% sodium hydroxide solution (CAS-1310-73-2) to achieve a pH of 7.2 ± 0.2.

Water Phantom

Water phantoms are comparatively quick to create, compared to the gel phantoms. However, it is still a time consuming procedure where care must be taken to ensure the correct measures are used as the quantities of chemicals involved can be incredibly minute, especially for the ad-hoc pH balancing.

For the water phantoms, we chose to use 290 ml round bottomed boiling flasks to contain the solution. These provide excellent B_0 homo-

geneity, with a large homogeneous region in the centre for acquisition of high SNR spectra. De-ionised water is used as the solvent, with the metabolites added in powdered form. Weights for each metabolite are calculated for the correct mMol/ml, and dissolved into the solution with a glass stir rod until there is no visible precipitate. In order to save time, a base solution is made that contains static concentrations. GABA is then added in small quantities to the flask to increase the GABA concentration.

A 350 ml metabolite solution is made, with constant levels of [Cr & NAA] for E1 & E2 and additionally [Glu & Gln] for E3 & E4. 290 ml was placed into the round bottom flask, with an additional portion reserved for the concentrated GABA solution.

The water phantoms were altered during scanning sessions by increasing the GABA concentrations. These variations in the GABA concentration are achieved by removing 1 ml of phantom solution, and replacing it with a highly concentrated metabolite solution, equivalent to 1 mMol/ml. Between each scanning session pH was recorded, and adjusted if necessary using the previously mentioned method. This method introduces a potential methodical issue, as 0.004% of the previous GABA solution is removed on each iteration, which is later corrected for in the phantom concentration values.

Gel Phantom

Gel phantoms are designed to be tissue mimicking, simulating an environment that is much closer to one that is found *in-vivo* [52, 111]. They are an excellent middle ground between water phantoms and biological material, that retain the ability to customise the metabolite composition, however, they are more time consuming to create.

Preparation of gel based phantoms require that the metabolite solution is heated to around $90 - 100°$ before the gel crystals can be successfully dissolved. Once heated, the solution can then be poured into the phantom moulds and left to set. Although these phantoms can be re-heated to be liquefied again, this additional heating process makes gel phantoms substantially less mailable than their water counterparts. Because of this, a range of phantoms must be created at the same time, rather than one water based one that can be adjusted in-between scans. Overall, this makes gel phantoms more costly due to the higher amount of metabolite powder needed along with the additional materials for moulds, gelling agent and extra time needed to heat and pour the solution.

For our gel phantoms, we prepared 800 ml of the metabolite solution containing constant levels of [NAA, Cr, Glu & Gln], split into eight 100 ml portions, and had a range of GABA added as shown in Table 3.2. Each solution was pH calibrated and 1 g of agar was added as gelling agent. The mixtures were then heated to 90 °C to 100 °C while being stirred until the agar had fully dissolved.

For our phantoms moulds, we used clear acrylic spherical Christmas decorations, with a small hole (5 mm) made to allow for filling made with a soldering iron on low temperature. The solution was added to the phantom moulds with a syringe, and allowed to cool overnight, before adding silicone to seal the hole.

3.2.2 *Phantom Composition*

Our phantoms can be grouped into two sets; [GABA, NAA, Cr] and [GABA, NAA, Cr, Glu, Gln]. These sets of metabolites were chosen to specifically benchmark the accuracy of GABA quantification. The former, ([GABA, NAA, Cr]) provides the smallest reasonable set of metabolites for benchmarking MEGA-PRESS. The latter set included fixed quantities of Glx to see if this has any effect on the quality of the fit.

Table 3.1: Benchmark phantom composition of constant metabolite concentrations in mM (mmol/l).

Series	Medium	pH	#	NAA	Cr	Glu	Gln
E1	Water	7.2 ± 0.2	13	15.0	0/8.0	0.0	0.0
E2	Water	3.0 ± 0.2	15	15.0	8.0	0.0	0.0
E3	Water	7.2 ± 0.2	15	15.0	8.0	12.0	3.0
E4	Gel	7.2 ± 0.2	8	15.0	8.0	12.0	3.0

Tables 3.1 and 3.2 show the medium used, the pH and the metabolite concentrations in mM (mmol/l). In general GABA was the only metabolite where the quantity was varied over the course of the scans, with the rest of the metabolites having a fixed value, with the exception of E1, where the first scan only had NAA, the second NAA and Cr.

The phantom concentrations were selected to mimic the ranges of the selected metabolites in a normal human adult brain Govindaraju *et al.* [45], with the exception of GABA. GABA typically has a concentration

Table 3.2: Benchmark phantom GABA concentration in mM (mmol/l).

	Scan number														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
E1	0	0	0.5	1	1.5	2	2.5	3	4	6	8	10	11.6		
E2	0	0.5	1	1.5	2	2.5	3	4	5	6	7	8	9	10	11.8
E3	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
E4	0	1	2	3	4	6	8	10							

of $1.3 - 1.9$mmol, where in our phantoms it extends all the way up-to 14mmol providing ample opportunity to correctly quantify the signal.

Phantoms E1, E3 and E4 were all pH adjusted, no pH calibration was performed on E2; the pH was found to be 3.0 ± 0.2.

3.2.3 *Scanning Conditions and Settings*

Phantoms were scanned on a 3 T Siemens MAGNETOM Skyra located at Swansea University, UK which has an operating B_0 field of around 127 MHz (3 T). The scanner room and phantom temperature was controlled at $20 \pm 0.6°$C by the air conditioning. Phantom temperature was measured between each scanning run to check for any heating affects, of which the variation was negligible.

The WIP MEGA-PRESS 859D [1] pulse sequence ($T_E = 68$ ms, $BW = 1250$ Hz) was used with CHESS water suppression pulses, over a $20 \times 20 \times 20$mm^3 voxel with spectra averaged over 160 acquisitions. Editing pulses were applied at 1.9 ppm during the off and 7.4 ppm during the on acquisition to remove the obscuring Cr resonance from the GABA signal at 3 ppm.

Phantoms were placed at isocentre in the scanner bore for the best B_0 homogeneity. The signal was acquired from the spine coils located in the patient table, due to them exhibiting the highest SNR of all the available coils due to their proximity to the phantoms. Shim and flip-angle calibration was first done using the automatic calibration routines on the scanner, and further manually adjusted to improve the strength of the apparent diffusion coefficient (ADC) and to minimise the linewidth of the spectra.

1 Work in Progress; the product is currently under development and is not for sale in the US and in other countries. Its future availability cannot be ensured.

3.2.4 *Quantification Procedures*

We aimed to perform quantification in the simplest way possible, without the addition of custom code or scripts where possible, in an attempt to reduce operator bias. For all programs, only the difference spectrum (diff) was used for quantification (edit on - edit off), with no additional input of a water unsuppressed spectra, to focus on relative quantification. All tools were provided with the unedited Siemens ".IMA" digital imaging and communications in medicine (DICOM) in order for the input data to be compatible and consistent across all tools.

However, special care had to be taken due to the comparatively low temperature of the phantoms, compared to *in-vivo*. The resonance frequency of the water peak is directly dependant on temperature, and as such this leads to the location of it shifting. Many quantification methods use the water peak as a frequency reference; for some quantification methods other reference peaks had to be used, or the frequency axis manually corrected.

While using these tools, the creators were not contacted in regards to optimising fitting for MEGA-PRESS, to remove any source of potential bias. The aim was to use these tools as any other user would, starting from a place without specialist knowledge of their usage, but utilising the recommended settings in the user guides.

Totally automatic robust quantitation in NMR (TARQUIN)

TARQUIN version 4.3.11 was used with the inbuilt MEGA-PRESS basis set for quantification. For frequency referencing, the NAA peak was used. It is worth noting that TARQUIN has a comparatively different approach to MEGA-PRESS quantification to other basis set methods. It attempts to adjust the signal and phase to correct the negative NAA peak found, and treat quantification more like a PRESS spectrum.

Linear combination of models (LCModel)

LCModel version 6.3-1L was used, with the MEGA-PRESS basis sets were sourced from Dr. Dydak's lab at Purdue University, USA [30] under the recommendation from the creator of LCModel, Stephen Provencher. This is the standard process for acquiring LCModel basis sets; emailing Stephen Provencher, as indicated in the user manual [113]. The 3 T Siemens difference basis set with the Kaiser couplings [60] was used for quantification. The settings shown in Listing 3.1 were used to enable

MEGA-PRESS quantification, as recommended from the LCModel user guide.

Listing 3.1: Recommended LCModel MEGA-PRESS quantification settings [113]

```
SPTYPE = 'mega-press-3'
PPMST = 4.2
PPMEND = 1.95
```

Java-based MR user interface (jMRUI)

jMRUI version 6.0 beta was used, with both, the time domain quantification based on quantum estimation (QUEST) and automated quantification of short echo time MRS signals (AQSES) methods assessed. The basis set was simulated using jMRUIs inbuilt simulator: NMR-SCOPE-B [146]. The MEGA-PRESS protocol for simulation was created matching the MEGA-PRESS implementation utilised experimentally. Manual adjustments to frequency and phase calibration along with apodization on the input spectra was performed, before being passed onto the quantification methods. The fit amplitudes were used as the reported units.

GABA-MRS analysis tool (GANNET)

GANNET version 3.0 was used and run with MATLAB (R2018a), with adjustments made to the the pre-initialisation script to compensate for the room temperature phantoms, by adjusting the frequency offset. GANNET requires no basis sets as it is a peak fitting method. As GANNET reports the area of the peaks, rather than the expected concentration, the area of GABA and NAA is multiplied by $\frac{3}{2}$ to account for the proton weightings, and to obtain the GABA-to-NAA ratio.

LWFit

LWFit is a proprietary fitting method developed as part of this study by our collaborators S. Shermer and C. Jenkins [59]. The method is simplistic, and is an attempt to see how well a basic numerical integration routine would perform when compared to the more complex available methods. It uses a numerical peak integration method to fit peaks over a predetermined frequency range per metabolite. As before with GANNET, the ratios are achieved by multiplying the GABA and NAA reported ratios by $\frac{3}{2}$ to account for the proton ratios to the respective peaks.

3.3 RESULTS

The results are presented in two sections: firstly, the quality of the observed spectra, followed by the presentation of the results from the quantification methods.

3.3.1 *Spectra Quality*

Tables 3.3 and 3.4 display the average linewidths of the water and NAA peaks from a Lorentzian fit for all series. Additionally, the WS factor (WSF) and SNR for the NAA peak is shown. Overall, the linewidths for the solution phantoms is excellent in the $2-4$Hz range, suggesting an accurate B_0 shim was achieved producing high quality spectra. For the gel phantoms, the linewidths are expectedly worse as they mimic the environment found *in-vivo*, with the water peak around 12 Hz and the NAA around 4 Hz, however this should be more than acceptable for spectroscopy at 3 T.

Table 3.3: Linewidths (Hz) of the H_2O (mean±std) peak with water suppression (WS) On and Off for the Edit Off spectra, as well as mean WSF.

Series	WS Off	WS On	WSF
E1	2.40 ± 0.67	2.10 ± 0.54	1840
E2	3.04 ± 1.44	2.82 ± 1.03	1873
E3	2.98 ± 0.90	4.16 ± 2.02	665
E4	11.16 ± 3.19	12.42 ± 9.01	478

Table 3.4: Linewidths (Hz) of the NAA peak (mean±std) for Edit Off and difference spectra and the mean SNR of NAA peak.

Series	NAA Edit Off	NAA Diff.	NAA SNR
E1	1.19 ± 0.31	1.19 ± 0.31	209.6
E2	1.84 ± 0.33	1.85 ± 0.33	223.0
E3	1.97 ± 0.25	1.98 ± 0.25	222.0
E4	4.20 ± 1.48	4.25 ± 1.55	109.7

E1

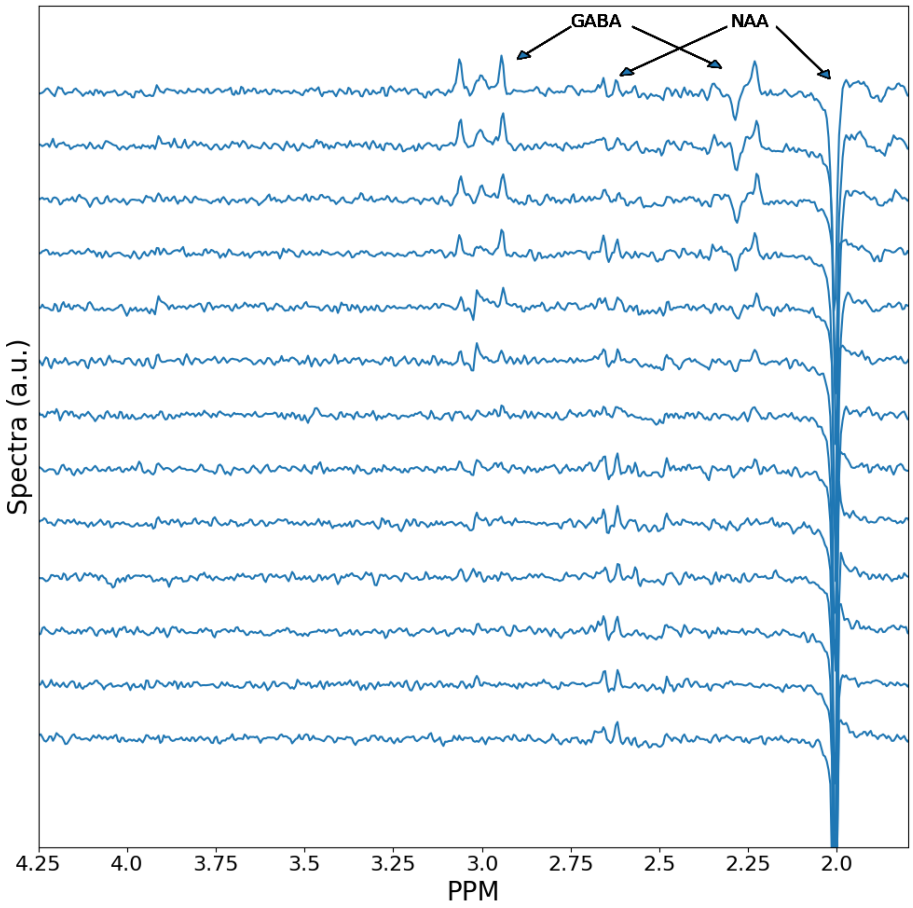


Figure 3.3: E1 real difference spectrum. GABA concentration increasing from bottom to top.

Figure 3.3 shows a waterfall plot of the real difference spectrum for E1. The spectra are well resolved, with clear distinct features indicative of the lowest linewidth of all the series as shown in Table 3.4. The editing efficiency is also excellent, with no visible residual Cr signal in the difference spectrum, clearly displaying the GABA triplet around 3 ppm.

E2

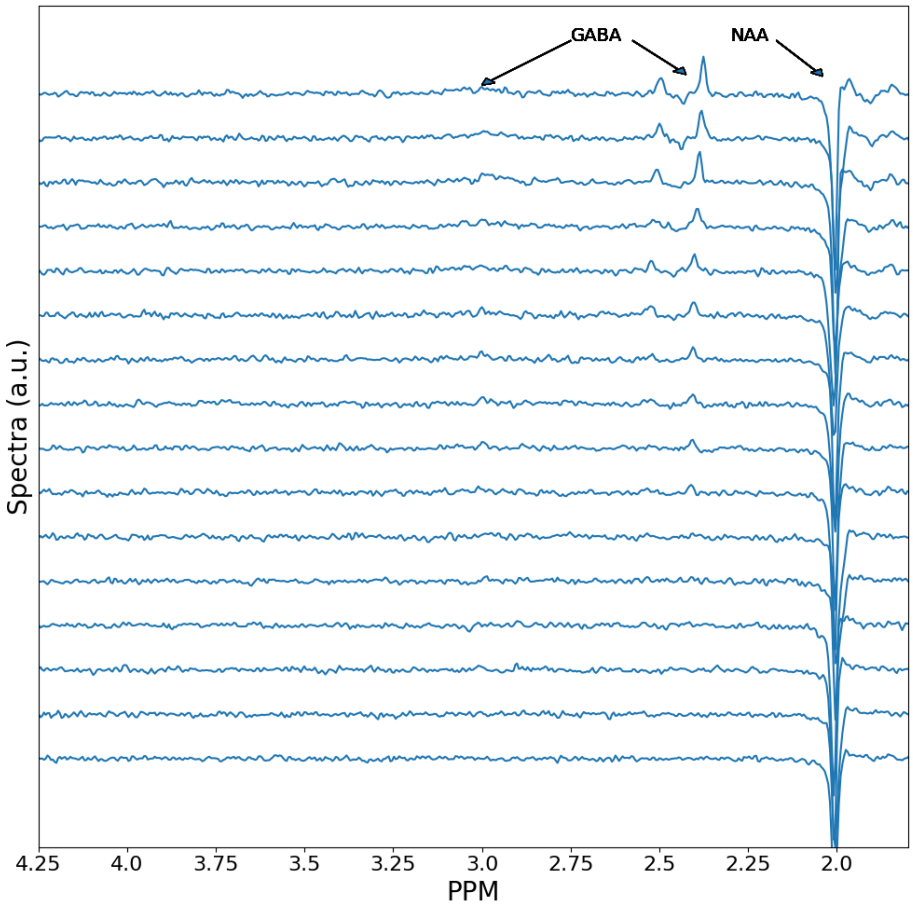


Figure 3.4: E2 real difference spectrum. GABA concentration increasing from bottom to top.

Figure 3.4 shows a waterfall plot of the real difference spectrum for E2, the low pH phantom. As before with E1, the spectra are of good quality, with well resolved sharp peaks. The linewidth of the NAA peak as shown in Table 3.4 is still below 2 Hz, which is excellent. However, the effect of the low pH is clear to see, with the NAA peaks drifting from around 2.3 ppm all the way up-to 2.8 ppm. This also appears to have a large effect on the GABA peaks seen around 3 ppm, where in E1 they are visible, in E2 are now a small 'mound' and have lost their distinctive shape despite having very similar concentrations in both phantoms.

E3

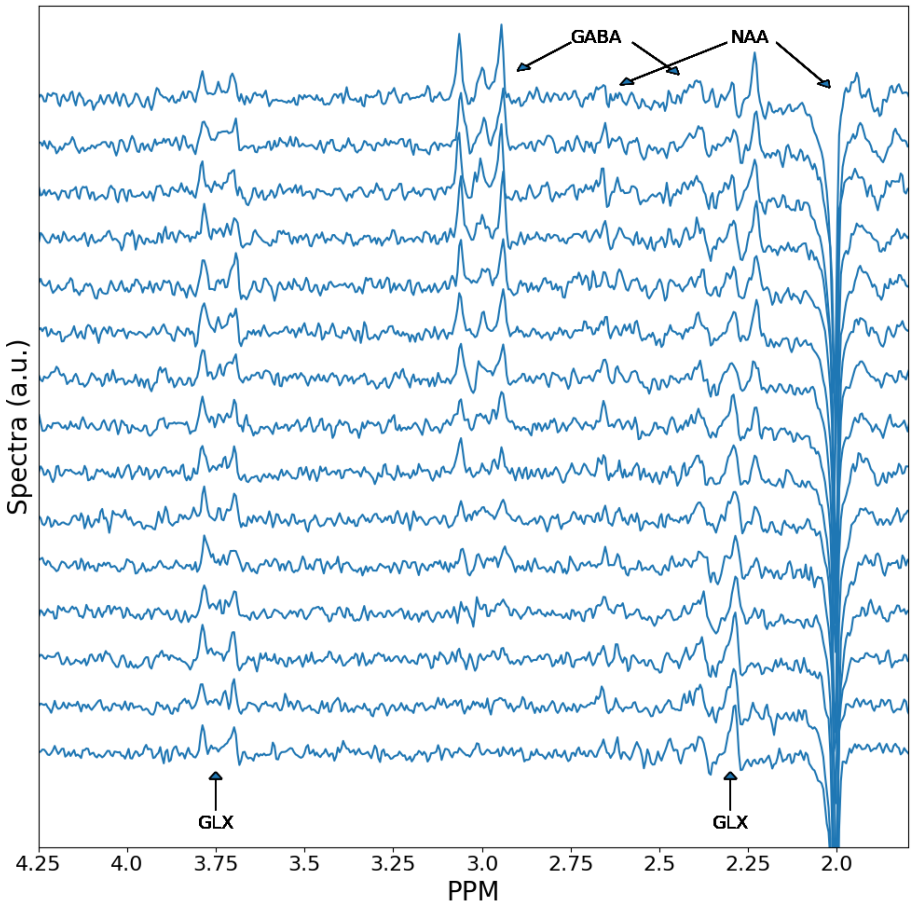


Figure 3.5: E3 real difference spectrum. GABA concentration increasing from bottom to top.

The E3 phantom introduces Glu and Gln to the composition, complicating the spectrum as shown in Figure 3.5. As before, the spectra are well resolved, still maintaining the linewidth of NAA less than 2 Hz. However, this water phantom has the lowest WSF of the three, suggesting that the CHESS routine was not fully effective. Despite this, it does not appear to have an adverse affect on the quality of the spectra. The spectra from E3 are more complex than E1 and E2 due to the additional peaks from Glu and Gln. The GABA peaks around 3 ppm in the difference spectrum are well resolved and visible with an additional contribution from Glx around 3.8 ppm.

E4

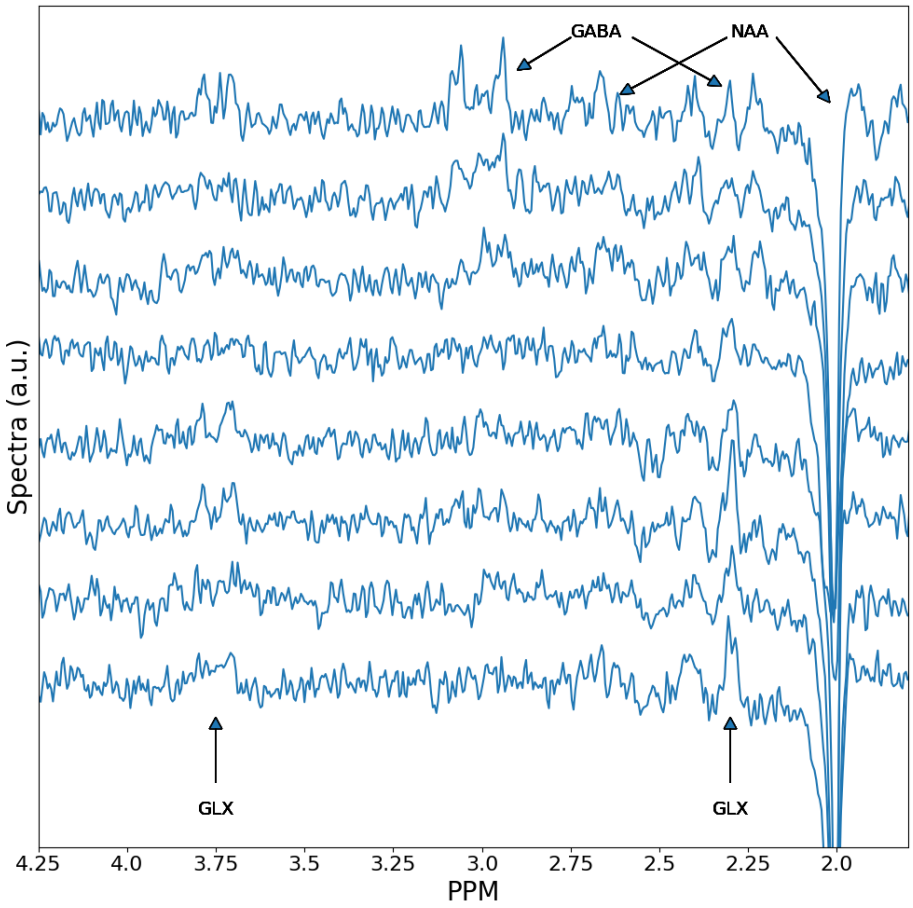


Figure 3.6: E4 real difference spectrum. GABA concentration increasing from bottom to top.

Our final set of spectra are from the E4 Gel phantom as shown in Figure 3.6. The peaks in the spectra are visibly wider, with the linewidth of the NAA peak being two times wider than the water based E1-E3 phantoms. In addition, this phantom has the lowest WSF as shown in Table 3.3. However, these two effects are to be expected by the tissue mimicking gels. The NAA and Cr peaks are still well resolved, with the GABA triplet visible at 3 ppm and the Glx contribution around 3.8 ppm.

3.3.2 *Quantification Performance*

As relative quantification is being performed, a reference metabolite is needed to produce a ratio. Cr and NAA are common reference metabolites. However, as the Cr signal is edited out in the difference spectrum

for MEGA-PRESS, we use NAA as the reference. As such, performance is assessed on the GABA-to-NAA ratio, comparing the reported values to the known values. Results are analysed by comparing the slope and intercept of the real-to-predicted values reported over a 95% confidence interval (CI), along with the coefficient of determination (R^2). The ideal quantification method to have an intercept at 0, a slope of 1 and an R^2 value of 1.

Results from all quantification methods are summarised in Table 3.5. The following paragraphs contain figures plotting the slope and intercept of each quantification method, for each phantom.

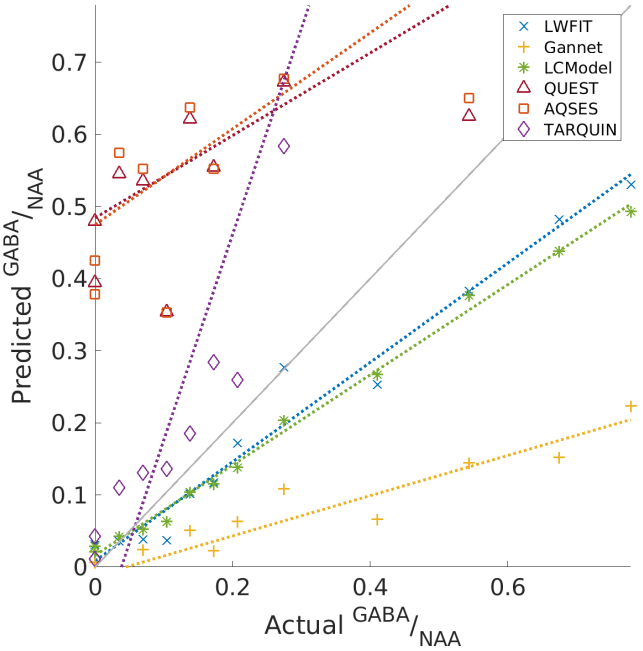


Figure 3.7: GABA/NAA ratio for E1 displayed as the predicted-to-actual ratio. Dashed lines are linear fits per quantification method. The light grey line represents an ideal quantification method, with a 1:1 ratio.

E1

Figure 3.7 shows the GABA-to-NAA ratio for the E1 series. Visually, LCModel and LWFIT perform the best, with the closest slope and intercept to the ideal quantification method, with well grouped linear estimations, leading to their high R^2 values in Table 3.5. GANNET similarly produces a reasonably linear set of predictions with a good intercept near 0, however the slope of 0.27 displays how significantly it is underestimating the ratio. TARQUIN, AQSES and QUEST all over-predict the

ratio, with TARQUIN producing a relatively linear set of estimations (R^2 of 0.95) and an intercept near zero of -0.11. However, it grossly over estimates the slope. Both of the jMRUI based methods (AQSES and QUEST) perform the worst of all the methods by detecting GABA where there is none present with an intercept near 0.5, in addition to having a poor scaling with a slope near 0.6 and poor linearity with an R^2 of 0.68.

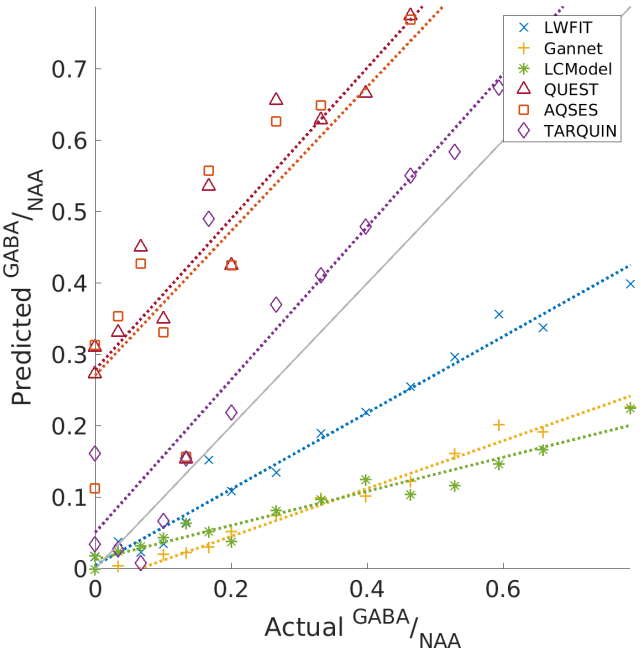


Figure 3.8: GABA/NAA ratio for E2 displayed as the predicted-to-actual ratio. Dashed lines are linear fits per quantification method. The light grey line represents an ideal quantification method, with a 1:1 ratio.

E2

Figure 3.8 displays the quantification results for the E2 phantom, which is the low pH water phantom. Overall, all quantification methods make more stable estimations, with a mean R^2 value of 0.90 compared to 0.85 for E1. In general, similar trends continue, with LCModel, GANNET and LWFIT continuing to underestimate the ratio, and with TARQUIN, AQSES and QUEST over-estimating. TARQUIN provides the best fit for this dataset, with a good slope of 1.07, intercept of 0.05 but with a comparatively average R^2 of 0.9. LCModel performs significantly worse than on the E1 phantom. However, this is to be expected as the basis sets provided no longer match the experimental spectra. LWFit overall performs reasonably, which is to be expected as it does not use a basis set

but peak integration, where at the same time GANNET performs poorly, perhaps trying to fit a Lorentzian peak to the uncharacteristic signal from GABA as seen in Figure 3.4.

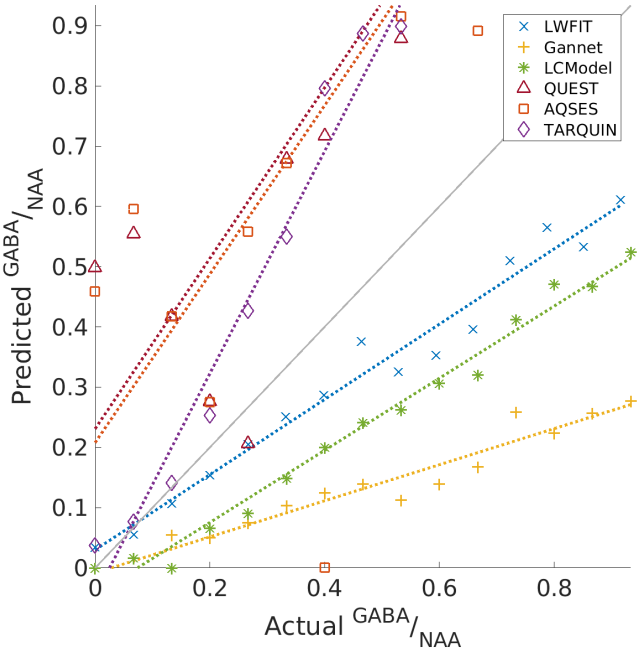


Figure 3.9: GABA/NAA ratio for E3 displayed as the predicted-to-actual ratio. Dashed lines are linear fits per quantification method. The light grey line represents an ideal quantification method, with a 1:1 ratio.

E3

Figure 3.9 shows the quantification results for the water based E3 phantom, that unlike E1 and E2 now contains Glu and Gln. Overall, quantification performance is relatively poor, where the best methods are LCModel and LWFIT with the best slope and intercept values, but they still underestimate the ratio by around 30%. TARQUIN, AQSES and QUEST continue to over-estimate the ratio, with all having a slope greater than 1.2. The jMRUI based methods continue to detect GABA when it is not present in the phantom with an intercept of around 0.2 with all other methods having intercepts near 0.0. LCModel, GANNET and LWFIT are the three methods that have similar results to E1, suggesting the inclusion of Glx has little effect on their ability to quantify NAA and GABA. In contrast, TARQUIN, QUEST and AQSES have large differences in their slope and intercept values, but similar NAA variation, alluding to difficulties quantifying GABA when Glx is present.

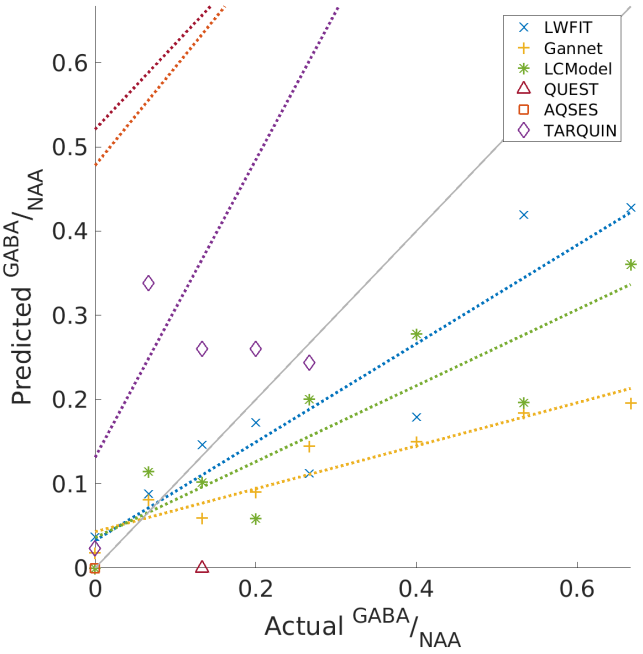


Figure 3.10: GABA/NAA ratio for E4 displayed as the predicted-to-actual ratio. Dashed lines are linear fits per quantification method. The light grey line represents an ideal quantification method, with a 1:1 ratio.

E4

Finally, Figure 3.10 displays the quantification performance on the gel E4 phantom. Overall quantification is the worst across the board compared to the other phantoms, with the lowest average R^2 of all three phantoms of 0.65, suggesting the difficulty of quantification is greater. The same quantification trends continue, with LCModel, LWFIT and GANNET under-estimating the ratio and TARQUIN, AQSES, QUEST over-estimating. Overall LWFIT performs the best of all methods, with an intercept of 0.03. However, with a slope of 0.58 it underestimates the ratio by 42%. AQSES and QUEST have few visible points inside the plot area, and again with intercepts of around 0.5, detecting GABA when it is not present. TARQUIN has the worse R^2 seen for any result across all phantoms, with a range of scans decreasing in their predictions as seen between $0.2 - 0.4$ in the figure, it would appear that TARQUIN struggles quantifying with the higher linewidth.

3.4 EVALUATION AND DISCUSSION

Overall, we have highlighted the importance of benchmarking quantification methods with experimental data. All of methods tested, at best, are concerningly far from the ground truth and, at worst, will incorrectly report high levels of GABA when none is present. In addition, we have shown that it is possible to create calibrated phantoms relatively easily to achieve high quality spectra.

However, as the spectra were collected from one scanner, with one pulse sequence and we only focused on the GABA-to-NAA ratio, it is hard to generalise. In theory this should apply across the board, but there may be environmental factors or user error altering the quality of the spectra from the scanner. It would be excellent to see this extended into a multi-site study, with different scanner manufacturers, magnet strengths and a range of pulse sequences for a range of metabolites, but this is clearly beyond the scope of this work.

The following two sections will explore in greater depth into the analysis of the phantoms and the quality of the quantification.

3.4.1 *Phantoms*

Overall we were able to get a range of 51 MEGA-PRESS spectra with varying concentrations of GABA. These phantoms were time consuming to create and scan, typically requiring at least a day or two for the water phantoms and three for the gel series. However, as shown in Tables 3.4 and 3.3 a high level of calibration was achieved, leading to low NAA peak widths less than 2 Hz for the water phantoms and 4.2 Hz for the gel phantom. The spectra, as shown earlier in Section 3.3.1, are high quality, with easily visible metabolite features.

The E2 phantom spectra, as shown in Figure 3.4, highlights the importance of correct pH balancing when creating phantoms. This was shown in the results, as the majority of the quantification methods struggled to quantify the E2 spectra when compared to the E1 phantom which has a very similar chemical composition, but a neutral pH.

The gel phantom E4 produced expectedly worse spectra, mimicking the environment found *in-vivo*. Despite this, the linewidth achieved on the NAA peak is within reasonable values for 3 T *in-vivo* spectroscopy [167]. It is interesting to note that all quantification methods struggled with

this, suggesting that more research is needed into benchmarking these tools with tissue mimicking gels.

3.4.2 *Quantification*

The results highlight some interesting shortcomings for the selected quantification methods for MEGA-PRESS spectra. Generally, for all phantoms LWFIT performed the best, with the closest average gradient and intercepts. However, it was specifically designed in conjunction with this study. Of the external tools, LCModel is the best performing on all phantoms, but it consistently under-estimates the ratio by around 50%. Overall, all methods performed worse on the issue mimicking E4 phantom. This result is worrying and highlights the need for more testing of these quantification methods, as they are designed for *in-vivo* quantification.

TARQUIN

TARQUIN consistently overestimates the GABA-to-NAA ratio, with the exception of the E2 phantom, for all others the mean predicted slope is 2.15. In addition, TARQUIN struggles to consistently estimate the constant NAA concentration. The large variation for E1 may be due to an unexplained rescaling of metabolite amplitudes between scans 10 and 11. The % variation of NAA for scans 1 – 10 is 6.09%. There is no simple explanation for the large variation reported by TARQUIN for E4. For the water based phantoms (E1-E3), TARQUIN is relatively consistent in the linearity of the predictions as indicated by the high R^2 values. However, for E4 this drops substantially, struggling with consistency for the gel based phantoms.

LCModel

LCModel was the best performing external tool, consistently suggesting a low intercept value and typically having the highest R^2 value per phantom. Of the external tools, it was consistently the closest to the ideal quantification method. However, it consistently under-estimates the GABA-to-NAA ratio, and expectedly performs poorly on the low pH phantom due to the basis not matching the spectra. Worryingly, in line with the other methods, the performance drops for the gel phantom, underestimating the ratio by 55%. Overall, this suggests that LCModel is

a good tool for highly calibrated spectra, but it begins to struggle with distorted spectra shapes as seen in E4.

jMRUI

Both AQSES and QUEST from jMRUI were consistently the worst of all the quantification methods for consistently having an intercept that is out of an acceptable range, and with the exception of E4 consistently making the least consistent predictions as suggested by the R^2 value. This is worrying, as both methods report a quantity ofGABA when there is none. Both methods appear to be relatively consistent with each other, with their slopes and intercepts for each phantom typically having close values. AQSES and QUEST are the only methods in this set that share the same basis set simulated from jMRUI. In addition, both methods were supplied spectra that had a range of manual pre-processing steps performed by a user for filtering and frequency and phase calibration from jMRUIs interface. Both of these could potentially be sources of methodological bias, where a more experienced user may be able to optimise each step of the process to improve the fit. This specifically highlights the difficulty and complexity of quantification, and where possible user interaction and the number of user-made choices should be minimised.

LWFIT

Overall, LWFIT was the best performing method, outperforming LCModel in terms of gradient accuracy marginally in most cases, and significantly for the E2 phantom by 30%. Overall, it reports an intercept value close to 0, consistently achieves a high R^2 value and had the lowest mean NAA variation of 2.68%. This is rather surprising, as it is a simple peak integration method. However, despite the successes of the method, like LCModel and GANNET, it consistently under-estimates the GABA-to-NAA ratio by around 40%. This result may be unsurprising, as the method was developed in conjunction with this study and so presents a significant source of methodological bias. However, with this in mind, it fails to perform significantly better than the second place LCModel. As this method is peak integration directly measuring the observed area, it would be difficult to apply to more complex spectra with spectral overlap without careful consideration of peak locations and fitting methods. As such, it is not a generalisable method, but it appears to work well for edited spectra where the overlap issue is removed.

GANNET

GANNET was the only external non-basis set method, and unfortunately it performed poorly at estimating the GABA-to-NAA ratio. It consistently underestimated the slope, where in the best case it estimated it to be 0.33 for E2 and 0.25 in the worst for E4. However, like many other methods it was able to consistently estimate an intercept near 0 along with having a low overall variation in the NAA estimation. This consistent underestimation of GABA may be due to the Gaussian model used to fit the GABA peaks, where visually our well resolved peaks are more Lorentzian. Nevertheless, GANNET is able to produce a reasonably consistent set of estimations, as shown by the high R^2 values.

3.4.3 *Methodological Issues*

This work was an initial study, where our group was investigating the need for benchmarking. While this has shown some rather surprising results, there are a few issues that should be addressed in future work before the results should be considered generalisable.

Firstly, this work was conducted at a single site and only assess the performance of the GABA-to-NAA ratio. This makes this work difficult to generalise to different strength scanners and even to other scanner manufacturers, as MEGA-PRESS is known to have vendor specific implementations [101]. In addition, although it may be unlikely, any issues with the spectra produced at Swansea would be difficult to detect without having data from another site to compare to. In future, it this study should be expanded to include multiple scanners at different sites, from a range of vendors and with a broader range of pulse sequences. In addition, a range of gel phantoms could be made to have specific T1 and T2 relaxation properties, it would be interesting to explore a range of phantoms that mimic specific areas *in-vivo*, and their relaxation times. Additionally, the use of a vendor made spectroscopy specific phantom would be interesting to explore, as this would provide us with a high quality reference spectra. This would also make a multi-site study substantially easier, assuming that other locations also have the same phantom available. This were not explored as we did not have one available, and the cost to purchase them is substantial.

Our phantoms were designed to provide a "'best-case"' scenario, where only a minimal number of metabolites were used to simplify the quantification problem. It is important to recognise that the environment *in-vivo*

is much more complex, with a diverse range of metabolites. As such, it is important to continue this work by expanding the range of metabolites in the phantoms, such as Choline, Glutathione, Phosphocreatine and Myo-inositol. Introducing more metabolites would make the spectra more complex and the quantification more challenging, but would bring the benchmarking spectra closer to what is seen *in-vivo*. By collecting a larger range of spectra, it would allow the results to be more generalisable.

At the initial conception of this study, there were two groups of thought as to how to approach the quantification. One was to use the same basis set for all of the basis set methods, with the second being to use the recommended basis set or simulator for each method. During the initial research, the former method raised an interesting set of issues, where most quantification methods are unable to export and read each other's basis sets. From our set only LCModel and TARQUIN are able to share the same basis set by being able to read LCModels proprietary ".BASIS" files. This has left us with the latter method of using the recommended settings for each quantification method. As the performance of the basis set methods is directly reliant on the quality of the basis set, it is extremely difficult to separate this effect in our analysis. In addition, it appears that this may even be the case for the jMRUI results, as the two independent methods consistently have similar results, potentially due to their shared basis. In order to correctly assess the underlying performance of the basis methods, there would need to be a standardised format for saving and loading of basis sets used by all tools. Currently the LCModel ".BASIS" and ".RAW" appear to be the most transferable, but with a very poorly documented standard, and it is based on an old Fortran file format that is not particularly human readable. Ideally, a new format should be specified and used by all programs, perhaps in a easily human and machine readable mark-up language, such as YAML or JSON.

For most of the basis set quantification methods, the Govindaraju *et al.* [45] GABA model is used. However, the LCModel basis set used the Kaiser *et al.* [60] model for GABA. As before with the different basis sets, by changing the metabolite model this adds another uncontrolled variable for this study, making it difficult to understand a cause and effect with the results. In theory, the success of LCModel could potentially be due to this metabolite model. In the future, the metabolite models should be standardised across all quantification methods.

Overall the results suggest that peak integration methods are more robust than basis set methods for edited MRS, perhaps a consequence of the simplified spectral landscape, eliminating the requirement for

basis set fitting. This is in agreement with the observations of Mosconi *et al.* [99], who also reported the robustness of these methods *in-vitro* but note the poor performance of integration methods *in-vivo*. However, due to the absence of a ground truth, the accuracy of quantification is difficult to establish for *in-vivo* data.

3.5 SUMMARY

This chapter has presented work for benchmarking a range of quantification tools using experimentally collected MEGA-PRESS spectra at 3 T. Overall, we have shown that there is a great deal of need to have a more rigorous method to benchmark these quantification methods, in order to improve them. We have also shown the procedures used to create a range of water and gel based spectroscopy phantoms for this method are not complex, and are within reach of many other researchers.

The quantification methods investigated present a worrying picture for the current accuracy of GABA quantification with MEGA-PRESS at 3 T, with the best external quantification method (LCModel) consistently underestimating the GABA-to-NAA ratio by 38%-55% for the pH neutral phantoms. For the other methods, AQSES, QUEST and TARQUIN were very far off the mark, and GANNET performing similarly to LCModel but typically underestimating by at least an additional 20%.

The systematic errors observed require further study to elucidate the precise nature of the variation and this work highlights the need for standardisation of existing methods and the development of new approaches to quantification of MRS data.

The methodological dependence of the quantification results observed also suggests that care must be taken when comparing results across studies, where analysis pipelines may differ and standardisation is desirable where possible. Furthermore, the methodological dependence is not restricted to instantaneous measurements of GABA, but is in fact found to propagate into reported changes in concentration, with tools presenting a range of GABA-to-NAA gradients. This finding is of great relevance to clinical studies, with our results suggesting that reported differences between normative and disease state GABA measurements will also be influenced by the choice of analysis method.

Finally, we have made our spectra datasets publicly available [75], allowing others to use and expand on this dataset for benchmarking new or updated methods. This is the first step in generalised improvements

in MR quantification, as there is a clear need for experimental validation of methods.

Table 3.5: Quantification results for GABA-to-NAA ratios on E1 - E4 phantoms. Gradient (Grad.), intercept (Int.) are displayed over a 95% confidence interval (CI) in addition to coefficient of determination (R^2) and percentage of NAA variation.

Series	Tool	Grad. (95% CI)	Int. (95% CI)	R^2	NAA %
E1	TARQUIN	2.85 (2.41,3.30)	-0.11 (-0.27 ,0.05)	0.95	38.2
	AQSES	0.66 (0.36,0.96)	0.47 (0.37 ,0.58)	0.68	3.86
	QUEST	0.57 (0.29,0.84)	0.48 (0.39 ,0.58)	0.66	4.03
	LCModel	0.62 (0.59,0.65)	0.01 (0.01 ,0.03)	0.99	2.76
	GANNET	0.27 (0.20,0.34)	-0.01 (-0.09 ,0.01)	0.86	3.39
	LWFIT	0.68 (0.61,0.77)	0.009 (−0.02 ,0.04)	0.95	2.61
E2	TARQUIN	1.07 (0.86,1.27)	0.05 (-0.03 ,0.19)	0.90	2.87
	AQSES	1.01 (0.73,1.29)	0.27 (0.16 ,0.38)	0.81	4.95
	QUEST	1.06 (0.72,1.39)	0.27 (0.15 ,0.41)	0.77	5.52
	LCModel	0.23 (0.20,0.27)	0.013 (0.01 ,0.03)	0.95	4.33
	GANNET	0.33 (0.30,0.36)	-0.021 (-0.04 ,-0.01)	0.98	2.85
	LWFIT	0.53 (0.48,0.59)	−0.005 (−0.02 ,0.03)	0.96	4.22
E3	TARQUIN	1.85 (1.72,1.97)	-0.04 (-0.12 ,0.03)	0.99	5.55
	AQSES	1.40 (0.86,1.93)	0.20 (-0.09 ,0.50)	0.71	3.59
	QUEST	1.41 (1.04,1.79)	0.23 (0.03 ,0.43)	0.84	3.31
	LCModel	0.59 (0.55,0.64)	-0.04 (-0.07 ,0.02)	0.98	3.77
	GANNET	0.30 (0.25,0.34)	-0.01 (-0.03 ,0.02)	0.95	3.31
	LWFIT	0.62 (0.56,0.69)	0.03 (−0.00 ,0.06)	0.97	1.59
E4	TARQUIN	1.77 (-0.54,4.08)	0.13 (-0.70 ,0.96)	0.37	42.1
	AQSES	1.17 (0.15,2.18)	0.47 (0.12 ,0.84)	0.57	5.85
	QUEST	1.02 (0.18,2.22)	0.52 (-0.09 ,0.95)	0.42	6.11
	LCModel	0.45 (0.21,0.68)	0.04 (-0.05 ,0.12)	0.79	7.83
	GANNET	0.25 (0.16,0.34)	0.04 (0.01 ,0.07)	0.90	2.82
	LWFIT	0.58 (0.35,0.82)	−0.03 (−0.05 ,0.12)	0.86	2.29

4

QUANTIFICATION OF MEGA-PRESS SPECTRA WITH A CONVOLUTIONAL NEURAL NETWORK

In this chapter, we explore a novel use of convolutional neural networks as a method of quantification for 3 T MEGA-PRESS spectra. We explore in depth how to present the spectra to the network, as MEGA-PRESS produces multiple spectra due to the three acquisitions: Edit Off, Edit On and Difference. In addition, the spectra can be presented as different combinations of real, imaginary or absolute numerical values to the networks. Every combination of acquisitions and data-types is later tested, and shown that this ultimately has a large impact on the performance of the network. For benchmarking the quality of the networks, this chapter uses the data from Chapter 3. Using the benchmark results and spectra dataset, we show that convolutional neural networks are able to outperform the state-of-the-art on the supplied experimental spectra.

MEGA-PRESS is the focus of this chapter as, similarly to Chapter 3, it provides an opportunity to quantify the relatively low signal intensity metabolite GABA. As previously introduced in Section 2.3.1, GABA is of high research interest and given the poor performance of current quantification methods shown in Chapter 3, there is a significant need to improve on quantification accuracy of GABA with MEGA-PRESS.

Using a machine learning approach, such as CNNs, has extremely promising prospects for MRS quantification. Once a network is trained and validated, quantification is quick and requires no specialist knowledge or adjustment of fitting parameters. CNNs are excellent at being able to automatically extract and learn features from input data. This shifts the focus when designing a quantification method from numerical modelling of the spectra and noise to network architecture and input data design. This chapter investigates in detail the range of choices in presenting MEGA-PRESS spectra to a CNN along with a selection of network architectures.

The major result of this chapter is a CNN architecture: 'MRSNet'. MRSNet is designed to solve the multi-class regression problem that is spectral quantification. It accepts a selection of spectra as the input, and returns the relative quantities of five metabolites: [NAA, Cr, GABA, Glu and Gln]. The following sections describe the process of investigating a

range of choices that shaped the network design. Section 4.1 outlines the network architectures tested, along the range of input data shapes, data sources for training and testing and finally the experimental datasets used for validation. Section 4.2 steps through a range of experiments that feed sequentially into each-other, affecting experimental design decisions along the way in order to limit the search space. This section covers a range of network architectures, spectra representation into the network, basis set selection and a comparison with the state of the art. This specific sequence of experiments was chosen as a logical progression of network improvements, starting with the fundamentals and iterating on the design. Finally, there is an evaluation and summary of the work in this chapter. MRSNet outperforms LCModel in the majority of cases of the specific benchmark spectra dataset introduced in Chapter 3, where LCModel is the best performing publicly available method in that chapter.

4.1 METHOD

MRS network (MRSNet) is a 2D CNN. Input data to the network is a $N \times 2048$ matrix, where each row contains one component of a single spectrum. The maximum number of input rows is $N = 9$ for the real and imaginary part and magnitude of up to three spectra (edit-off, edit-on and difference). Spectra are zero-padded and trimmed to obtain a length of 2048 in the range from 4.5 ppm to 1.5 ppm, i.e. each column is a single frequency bin 0.0014 ppm wide, providing ample resolution for small features. The majority of convolutions are $1 \times M$, spanning horizontally across the frequency bins to align with the separate rows of the input spectra. This network structure is loosely based on the Visual Geometry Group's (VGG) CNN architecture [131] and experimentally derived. Multiple layer 1D and 2D networks and single-layer 3D networks were explored by varying the arrangement of acquisitions and data type representations. However, they generally have worse performance than the suggested network structure in this chapter and, hence, are not discussed further here.

The architecture of the network is designed to train identification of spectral features in individual rows (component of a single spectrum) early on, to subsequently be combined further down the network. These reduction layers are implemented in the middle of the network and repeated to combine the rows until the output tensor has one row. Using convolutions spanning multiple rows of the input early on lead to very

poor results, whereas combining the convoluted rows near the middle of the network provided the best results.

The complete network architecture is shown in Table 4.1 for the 'small' variant without pooling. For every variant of the network, the layer class *reduction1* is repeated until the output is a tensor with one row. For input spectra with nine rows, the layer is repeated four times, whereas for an input of a tensor with one row the layer has a convolutional kernel sized 1×3 and is used only once. This is where the information contained in the different rows (spectra) of the input tensor are combined. A variety of substitutions are made to test different 'medium' and 'large' networks and pooling variants as shown in Tables 4.2 and 4.3. The final layer uses the softmax activation function,

$$S(y)_i = \frac{e^{y_i}}{\sum_j^K e^{y_j}}, \tag{4.1}$$

to rescale the output to be a ratio that sums to 1.

The network is trained using the ADAM [66] weight update with an experimentally derived learning rate of 10^{-4} and suggested beta values of $\beta_1 = 0.9$ and $\beta_2 = 0.999$ respectively. Mean squared error (MSE) is used as the loss function over mean absolute percentage error (MAPE) or mean absolute error (MAE) due to the generally higher overall performance seen experimentally. The training was performed on an NVIDIA Titan X, using Python 2.7.15, Keras 2.2.4 [22] with Tensorflow 1.9.0 [1] and CUDA 9.0.176.

4.1.1 *Data*

Spectral datasets for quantifying mixtures of NAA, Gln, Glu, GABA and Cr are used due to their importance for MRS applications as introduced in Section 2.3.1. These metabolites are in principle detectable by MEGA-PRESS. The networks are trained and validated using simulated data generated by combining individual metabolite spectra from basis sets. A basis set contains the normalised characteristic signals in the frequency domain for each metabolite, which are combined to create mixed spectra. The goal is to have the CNN distinguish and identify these spectral features on a per-metabolite basis.

We use simulated datasets for training and validation, with an experimental dataset for benchmarking. Simulated spectral basis sets were deemed more efficient than deriving basis sets experimentally from calibrated phantoms. Using experimental data to generate basis sets is also

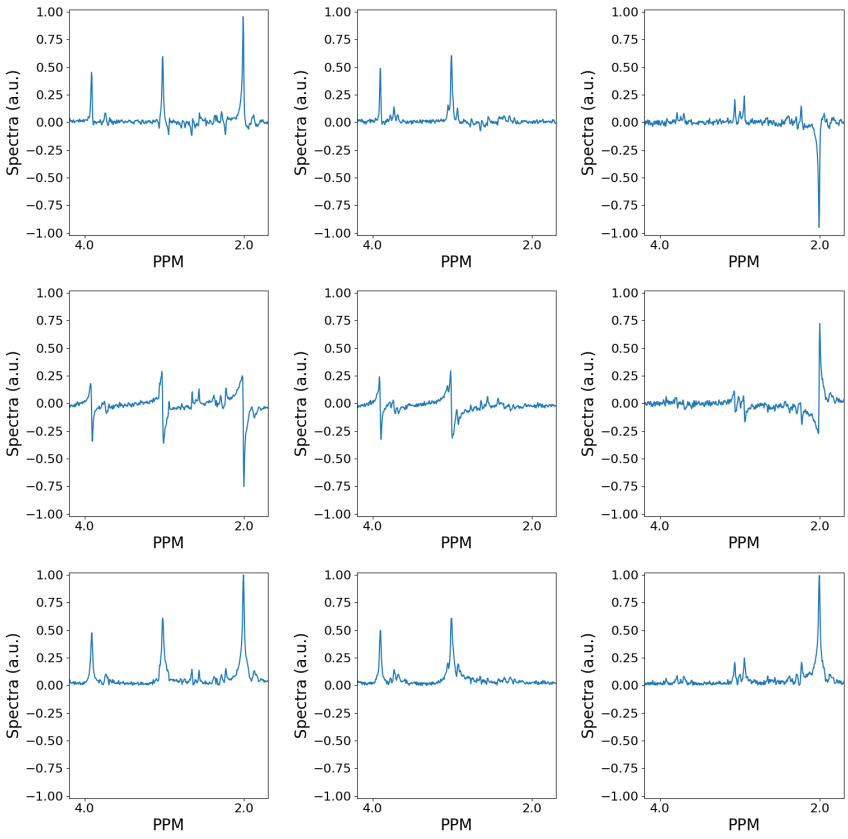


Figure 4.1: Acquisitions of MEGA-PRESS, from left to right: edit off, edit on and difference. data-types of the Fourier transform from top to bottom: real, imaginary and absolute.

problematic due to the difficulty of obtaining high-quality spectra [113]. To assess the performance of the networks, experimental data from carefully calibrated phantoms are used as a benchmark.

Selection of basis sets is as important as the selection of the quantification method, as it supplies the ground truth for the expected metabolite signals. In Section 2.3.2 we introduced the difficulties of basis set simulation. Hence, we use and test different basis sets from state-of-the-art simulation and quantification software: FID appliance (FID-A) [133], Python-GAMMA (PyGamma) [116] and LCModel [114]. It is important to mention, that each of these the three simulated bases utilise different GABA models: LCModel uses the Kaiser *et al.* [60] model, VeSPA uses the Govindaraju *et al.* [45] model, and finally FID-A uses the Near *et al.* [103] model.

Table 4.1: MRSNet architecture. 'Small' network structure without pooling. Ellipsis indicate repeated layers of the same class. Keys: convolution (conv), batch normalisation (BN), dropout (DO), fully connected (FC), rectified linear unit (ReLU).

Layer class	**Description**
input	$[1-9]\times2048$ spectra
conv1	$256\times1\times7$ conv, 1×2 stride, ReLU, BN, DO= 0.4
conv2	$256\times1\times5$ conv, 1×2 stride, ReLU, DO= 0.4
reduction1	$256\times[1-3]\times3$ conv, ReLU, DO= 0.25
reduction1	...
conv3	$256\times1\times3$ conv, padding=same, ReLU, DO= 0.25
reduction2	$256\times1\times3$ conv, 1×3 stride, ReLU, DO= 0.25
conv3	...
reduction2	...
conv4	$512\times1\times3$ conv, padding=same, ReLU, DO= 0.25
reduction3	$512\times1\times3$ conv, 1×3 stride, ReLU, DO= 0.25
conv4	...
reduction3	...
dense1	1024 FC
output	5 FC, Softmax activation

LCModel

Basis sets for LCModel are acquired by emailing the creator (Stephen Provincer) directly. For MEGA-PRESS spectra users are directed to the Perdue basis sets [30]. These basis sets are density matrix simulations of the vendor specific implementation of MEGA-PRESS, as it varies between Siemens, Phillips and GE [125]. For MRSNet we use the Siemens basis set with Kaiser *et al.* [60] couplings to match the basis set used for LCModel quantification performed in Chapter 3. As this data source is a basis set, and not a simulator unlike the other two basis set sources, the linewidth is fixed at 1 Hz.

Custom Python routines were written to read in the LCModel basis sets as they use a proprietary file format that is poorly documented in the user manual [113].

Table 4.2: Network layer class substitutions for 'medium' and 'large' networks, differences are in bold.

Network	Layer class	Description
Medium	conv1	256×1×**9** conv, 1×2 stride,
		ReLU, BN, DO= 0.4
	conv2	256×1×**7** conv, 1×2 stride,
		ReLU, DO= 0.4
	reduction1	256×[1 − 3]×**5** conv, ReLU, DO= 0.25
Large	conv1	256×1×**16** conv, 1×2 stride,
		ReLU, BN, DO= 0.4
	conv2	256×1×**8** conv, 1×2 stride,
		ReLU, DO= 0.4
	reduction1	256×[1 − 3]×**7** conv, ReLU, DO= 0.25

Table 4.3: Network substitutions for pooling variants where strided convolutions are replaced by a MaxPool layer following a convolution layer.

Network	Layer class	Description
All (pooling)	reduction2	256×1×3 conv, ReLU, DO= 0.25
		1×3 MaxPool
	reduction3	512×1×3 conv, ReLU, DO= 0.25
		1×3 MaxPool

VeSPA (PyGamma)

VeSPA is a Python based MRS toolbox and uses PyGamma as the underlying simulator, and in order to be able to programatically and automatically create basis sets, we directly call PyGamma with the pulse sequence code and metabolite models from VeSPA. The MEGA-PRESS pulse sequence code was kindly shared by Brian Soher, one of the creators of VeSPA. The PyGamma simulation code bundled in the MRSNet repository [14] is able to simulate FID, STEAM, PRESS and MEGA-PRESS pulse sequences. Spectra are simulated with a linewidth of 1 Hz, with a later multi-line width (MLW) training dataset containing 0.75 Hz, 1 Hz and 1.25 Hz spectra. These 'MLW' basis sets are created to later investigate the effect of training the network on a range of line widths, as in an

experimental setting, line width varies depending on the quality of the B_0 shimming over the ROI.

FID-A

FID-A is a Matlab MRS simulation and processing tool. MEGA-PRESS simulation is performed using their example MEGA-PRESS shaped editing routine, which has been modified to save spectra to disk and later loaded by Python. As with the PyGamma basis set, spectra are simulated with a linewidth of 1 Hz, with a later MLW training dataset containing 0.75 Hz, 1 Hz and 1.25 Hz spectra.

Experimental Dataset

In order to evaluate network performance, experimental datasets [75] created as part of Chapter 3 are used. These MEGA-PRESS ($T_E = 68$ ms and repetition time (T_R)$= 2000$ ms) spectra were acquired on a MAGNETOM Skyra 3 T (Siemens Healthcare GmbH, Erlangen, Germany) system at Swansea University using the Siemens WIP 859D.[1].

Four datasets for phantoms of known composition, E1, E3 & E4a and E4b, were used. E1, E3 and E4a were acquired with an acquisition bandwidth of 1250 Hz, while E4b was acquired at 2000 Hz (for the same phantoms as E4a). The published experimental dataset also contains a non-pH calibrated dataset E2, which was excluded as it was deemed not representative of the *in-vivo* environment. E4c and E4d, also contained in the dataset, are repeat runs of E4a and E4b and we obtained similar results for them, so they are not further discussed here.

4.1.2 *Spectra Pre-Processing*

Before the spectra are fed to the network, they are B_0 corrected w.r.t. the 2 ppm NAA singlet peak due to NAA's low sensitivity to temperature of 0.01 ppm/°C [24]. For the experimental spectra, they are filtered using a first-order Butterworth filter to reduce noise. No phase correction done for simplicity, as it is complex and can often require human interaction [8]. Experimental and simulated spectra are mean-centred and normalised so that the largest peak has amplitude ± 1 across acquisitions with $x' = x/\max(|x|)$. As different acquisition bandwidths are found across the basis sets and experimental datasets, these are dealt with by zero-filling

1 Work in Progress; the product is currently under development and is not for sale in the US and in other countries. Its future availability cannot be ensured.

the time domain signal to achieve a spectral resolution of 2048 points in the selected ppm range.

4.1.3 *Dataset Generation*

Training and validation datasets are generated by taking linear combinations of individual metabolite signals from a basis set. Each metabolite has a scaling factor in $[0, 1]$, corresponding to a relative concentration. The scaling factors are sampled using a low discrepancy (quasi-Monte Carlo) Sobol sequence [139], which provides good uniform coverage of possible states with a low number of data points. Time-domain noise is added from a normal distribution ($\mu = 0$ and σ randomly chosen in the range $[0, 0.25]$) to 50% of the dataset to improve simulation accuracy and network robustness. This noise model was chosen as it closely resembles what is seen in experimental spectra, by characterising the noise profile from spectral areas that do not contain a metabolic signal for $4,160$ phantom spectra.

4.1.4 *Performance Evaluation*

Performance of the networks is evaluated by the error ϵ, which is calculated as the mean of the absolute differences of the actual $a_{i,j}$ and the predicted $p_{i,j}$ relative concentration for every label (metabolite) $j = 1, \ldots, L$, for every prediction (spectrum) $i = 1, \ldots, N$:

$$\epsilon = \frac{1}{NL} \sum_{i=1}^{N} \sum_{j=1}^{L} |a_{i,j} - p_{i,j}|. \tag{4.2}$$

The standard deviation σ is calculated in the usual way,

$$\sigma^2 = \sum_{i=1}^{N} \sum_{j=1}^{L} (p_{i,j} - \epsilon)^2 / (NL) \tag{4.3}$$

This provides a good indication of overall network performance but is insensitive to low concentration metabolites. To counter this, an in-depth regression and MAPE analysis is later performed for the best network on a per-metabolite basis in Section 4.2.4.

4.1.5 *Experiments*

A range of network structures, spectra, representations and basis sets are investigated to find a favourable combination. Network structure

and batch sizes are investigated first, followed by the effect of different representations of the spectra and the choice of basis set. The LCModel basis set is used for the first two experiments as the analysis program has been shown to perform well in Chapter 3, suggesting a good fit with experimental data. Training and validation are performed using simulated spectra as described in Section 4.1.3. Initial experiments in Sections 4.2.1 and 4.2.2 are trained for 100 epochs, with 4,000 training, 1,000 validation samples. For the final experiment in Section 4.2.3, more compute time is dedicated to refining the networks over 200 epochs with 5,000 training samples. A light early stopping criteria is utilised throughout to prevent over-fitting, with a minimum loss decrease of 10^{-12}, a patience (number of epochs without improvement) of 15 with 'restore best weights' enabled. Network performance is benchmarked using the previously mentioned experimental datasets in Section 4.1.1.

4.2 RESULTS

The following section presents results from a range of experiments exploring network architecture, input data formatting, basis set choices and comparison with state-of-the-art quantification methods.

4.2.1 *Network Structure Investigation*

An initial investigation is performed to explore the effect of mini-batch sizes, convolutional kernel widths and pooling vs. stridden convolutions on performance with the benchmark datasets. Three variants of the network are tested (small, medium and large) over a range of mini-batch sizes (64, 32, 16) for two major variants of each network, either using stridden convolutions or max-pooling. These two variants are explored as convolutions with strides have been shown to be advantageous over max-pooling methods [144]. Networks are trained using all three spectra (edit-off, edit-on and difference) and all three data representations (real, imaginary and magnitude/absolute value) stacked to form a 9×2048 input tensor for each sample.

The results in Tables 4.4, 4.5 and 4.6 show the general trend that using convolutions with strides obtains a substantially higher performance over max-pooling. Networks with smaller convolutional kernels are substantially quicker than their larger counterparts as they are less computationally expensive to calculate. The 'small' network, using stridden

Table 4.4: Performance of MRSNet trained with LCModel basis set, all three acquisitions, all three data types using convolutions with strides.

Btch.	Sze.	Trn.	Vld.	E1	E3	E4a	E4b
16	S	0.00933	0.01113	0.12954	0.08898	0.07845	0.10344
		σ0.008	σ0.011	σ0.099	σ0.075	σ0.056	σ0.080
16	M	0.00858	0.01090	0.14440	0.08013	0.08157	0.11970
		σ0.008	σ0.011	σ0.126	σ0.064	σ0.055	σ0.074
16	L	0.00769	0.00968	0.12680	0.07382	0.08439	0.12502
		σ0.007	σ0.011	σ0.106	σ0.051	σ0.057	σ0.068
32	S	0.01096	0.01293	0.11980	0.08003	0.06619	0.13130
		σ0.010	σ0.013	σ0.099	σ0.063	σ0.053	σ0.093
32	M	0.00948	0.01194	0.14928	0.08752	0.09360	0.12649
		σ0.009	σ0.012	σ0.127	σ0.070	σ0.059	σ0.076
32	L	0.01169	0.01409	0.17127	0.06996	0.06271	0.10637
		σ0.010	σ0.013	σ0.134	σ0.043	σ0.046	σ0.081
64	S	0.01150	0.01274	0.11599	0.07378	0.07168	0.12210
		σ0.010	σ0.012	σ0.092	σ0.060	σ0.048	σ0.090
64	M	0.00961	0.01135	0.12914	0.09217	0.11139	0.13296
		σ0.008	σ0.011	σ0.111	σ0.067	σ0.049	σ0.074
64	L	0.01348	0.01567	0.11031	0.11308	0.10906	0.08609
		σ0.011	σ0.014	σ0.097	σ0.080	σ0.093	σ0.085

convolutions and a mini-batch size of 64 achieves a good balance between performance and training time on the benchmark datasets, making it the chosen network architecture for the following experiments.

4.2.2 *Data-Type and Channels*

Using MEGA-PRESS data provides a unique opportunity to explore how training networks on different combinations of acquired spectra for one scan (edit-off, edit-on and difference) affects performance. Additionally, three data types have been chosen to represent the spectrum by taking the real, imaginary or magnitude component of the frequency domain signal. These components, shown previously in Figure 4.1 are explored by using the same generated dataset for training and validation but considering different combinations of spectra and data types.

Table 4.5: Performance of network trained with LCModel basis set, all three acquisitions, and all three data types using max pooling.

Btch.	Sze.	Trn.	Vld.	E1	E3	E4a	E4b
16	S	0.02319	0.02537	0.23780	0.12593	0.14338	0.16654
		σ0.020	σ0.022	σ0.183	σ0.086	σ0.100	σ0.093
16	M	0.02155	0.02419	0.25608	0.12456	0.14460	0.18350
		σ0.018	σ0.022	σ0.203	σ0.095	σ0.097	σ0.108
16	L	0.02955	0.03267	0.23548	0.10779	0.11319	0.15390
		σ0.024	σ0.027	σ0.203	σ0.085	σ0.088	σ0.100
32	S	0.02550	0.02848	0.24714	0.14428	0.15845	0.16844
		σ0.021	σ0.024	σ0.194	σ0.107	σ0.115	σ0.097
32	M	0.02825	0.03044	0.26010	0.11713	0.11767	0.16157
		σ0.022	σ0.025	σ0.203	σ0.084	σ0.089	σ0.086
32	L	0.02463	0.02652	0.23353	0.11028	0.10509	0.11382
		σ0.019	σ0.022	σ0.153	σ0.061	σ0.068	σ0.074
64	S	0.02502	0.02701	0.25574	0.13671	0.14247	0.14566
		σ0.021	σ0.024	σ0.175	σ0.103	σ0.109	σ0.089
64	M	0.03168	0.03358	0.24458	0.12839	0.15139	0.14814
		σ0.025	σ0.027	σ0.197	σ0.095	σ0.105	σ0.097
64	L	0.04081	0.04236	0.24200	0.11145	0.10594	0.15611
		σ0.028	σ0.030	σ0.196	σ0.087	σ0.082	σ0.097

Results in Tables 4.7, 4.8 and 4.9 show a clear performance advantage of utilising the magnitude representation for all acquisition types. This is expected, as there is a level of phase uncertainty from the experimental signal leading to a potential disagreement with the real and phase spectra from the benchmark set when compared to the basis set. However, using the magnitude spectra does increase the linewidth when compared to only the real spectrum; which is typically preferred for quantification for this reason. Despite this, the networks prefer the more predictable magnitude spectra, at the cost of dealing with the increased linewidth.

The best combination of acquisitions is the edit-off and difference, as these provide the maximal amount of information with no repetition of data. The reduced performance for the difference only acquisition can be attributed to the Creatine (Cr) spectra, wherein the LCModel basis set that was used to train the network assumes perfect editing, where

Table 4.6: Timing performance of MRSNet trained with LCModel basis set, all three acquisitions (edit off, edit on and diff), all three data (real, imaginary and magnitude) types using convolutions with strides.

Batchsize	Network	Seconds per Epoch
16	Small	51
16	Medium	90
16	Large	112
32	Small	50
32	Medium	84
32	Large	100
64	Small	48
64	Medium	81
64	Large	96

the normal and inverted spectrum in the edit-on and edit-off acquisition match amplitude. As such, there is no residual Cr signal in the difference spectrum (which is not always the case in practice), as it has been edited out leading to difficulties in quantification. Networks perform worse when redundant data is supplied, for example, in the case of using the edit-off, on and difference magnitude spectra. This could be because the network is provided with too many degrees of freedom, increasing the training difficulty. For the following experiments, we choose networks that use the magnitude edit-off and difference spectra as inputs.

4.2.3 *Choice of Basis Set*

We compare the performance of three basis sets from LCModel, PyGamma and FID-A as described in Section 4.1.1 as differences in the basis sets may have a significant impact on quantification performance [56, 57]. Networks are trained and tested using the magnitude of the edit-off and difference spectrum with an increased number of training samples 5,000, 1,000 validation and test spectra, and with an increased number of 200 epochs for training.

Datasets were generated using the same metabolite concentration and noise values across the three basis sets to create a like-for-like comparison. Results in Table 4.10 suggest that for overall quantification, the single

linewidth (1 Hz) FID-A basis performs the best for E1, E4a and E4b, while the LCModel basis outperforms the others for the E3 dataset. The PyGamma basis performs substantially worse overall. This suggests that overall the FID-A basis set is the best choice for training the network for general quantification.

We also investigate the effect of training the networks on basis sets generated with multi-line width (MLW). In practice, the linewidth of the experimental spectra is variable, depending on a multitude of environmental factors as described in Section 2.1.4. Spectra are simulated using basis sets with line-widths of 0.75 Hz, 1 Hz and 1.25 Hz for the PyGamma and FID-A basis sets. For LCModel this was not possible as there is only one fixed linewidth basis set and we do not have access to the underlying generator. The network is trained using the same values for metabolite concentrations and noise as the single linewidth experiment. The generated spectra are split evenly between the defined line-widths. Table 4.10 shows that the overall accuracy of the networks trained with PyGamma improves but is marginally worse for FID-A. The use of simulators over fixed basis sets could be advantageous here as training a network on a range of line-widths should allow it to generalise to broader line-widths, typically seen in practice. However, in this instance, it appears that the linewidth of 1 Hz for FID-A closely matches the experimental spectra and using multiple line-widths has provided no improvement.

4.2.4 *Comparison with state-of-the-art*

Unlike previous sections where performance was measured against all five metabolites, the performance in this section is evaluated on reduced sets of metabolites (see Table 4.11). The sets are based on which metabolites are present in the phantom and an intersection of metabolites the selected programs can report.

Cr is omitted from the analysis as it is not reported by LCModel for MEGA-PRESS difference analysis. Glu and Gln values are combined and reported as Glx as they are generally considered unresolvable at 3 T due to their similar chemical structure and resulting spectrum. For the reduced metabolite sets, reported concentration values are re-scaled to $c'_r = c_r / \sum_{r \in R} c_r$ where R is the index set of the reduced metabolite set. Absolute quantification is not performed, as this requires a reference which is not possible for our water-suppressed spectra.

Quantification software settings, results and procedures are available in Section 3.2.4. The error measures have been changed to focus on individual metabolite error for this article rather than the NAA/GABA ratio.

An additional quantification method has been included in this chapter, compared to Chapter 3: VeSPA. It has been added as we are using a basis set generated by their simulator, and the work required once a basis set is simulated is not substantial. Comparing the performance of the quantification performance with VeSPA and MRSNet trained with the VeSPA basis will indicate the underlying quality of the basis set. VeSPA quantification was performed using the Voigt line shape, with automatic B_0 correction, phase shift and basic baseline removal using wavelet filtering. The basis set used for the VeSPA quantification is the the same simulated MEGA-PRESS basis used for training MRSNet as described in Section 4.1.1.

The results in Tables 4.12 and 4.13 show that overall MRSNet is more accurate and precise than other methods for quantification using MEGA-PRESS. In contrast to the results in Table 4.10, training the network with the LCModel basis set provides the best performance in this instance. This performance increase is due to the reduced set of metabolites for analysis. However, when evaluating MRSNet performance in the general case, across all metabolites, the FID-A basis set outperforms the LCModel basis on average.

4.2.4.1 *Regression Analysis*

The following subsections present graphed regression analysis for the E1,E3,E4a and E4b phantoms comparing the best MRSNet architecture ('small', stridden convolutions) trained with the LCModel basis set, compared with LCModel. R^2, Slope (sl), intercept (int), p-value (p), and standard error (se) values are displayed on each graph in the legend. The ideal quantification method would have a slope of 1 and an intercept at 0, as represented by the faint grey diagonal line.

These following results show that the performance of MRSNet is comparable to LCModel, except for E4b, where both struggle. The E4b dataset is considerably noisier than the others and the performance may be due to LCModel's pre-processing steps. Analysis of all series shows that only E4b suffers from inversion of the edit-on or off the spectrum, resulting in a poor difference spectrum, caused by the MR scanner software (also see [56]).

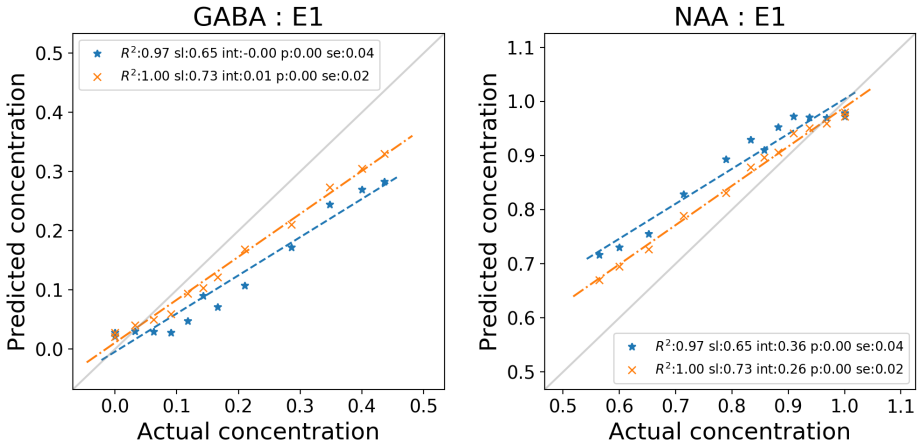


Figure 4.2: Regression analysis for MRSNet basis set (blue, $\star$) compared with LCModel analysis (orange, $\times$) for GABA (top row) and NAA (bottom row) for E1.

Figure 4.2 shows the regression analysis for the best network architecture trained using the LCModel basis set, compared with the fit from LCModel for both GABA and NAA fit to the E1 phantom. Overall performance is comparable between the two methods, with LCModel outperforming MRSNet with an overall more accurate slope, intercept and more consistent predictions based on the R^2 values for both GABA and NAA. Interestingly, both methods correctly predict an intercept near 0 for GABA, but they both substantially over-predict the intercept for NAA.

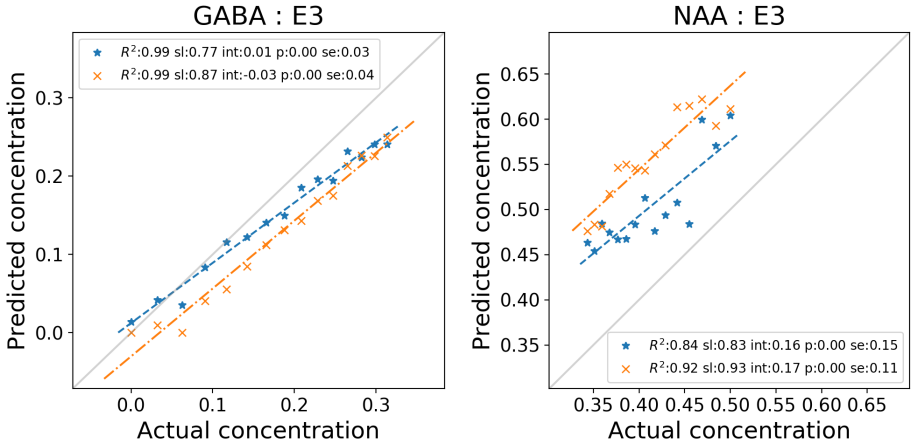


Figure 4.3: Regression analysis for MRSNet basis set (blue, $\star$) compared with LCModel analysis (orange, $\times$) for GABA (top row) and NAA (bottom row) for E3.

Figure 4.3 shows the regression analysis for the E3 phantom, comparing MRSNet with LCModel. LCModel outperforms MRSNet in terms of slope

accuracy, but the intercepts are worse, leading to a set of predictions that are further away from the ideal quantification method. Overall, MRSNet is closer to the ground truth for E3. Both methods perform well in terms of R^2 for GABA (0.99), however the R^2 values for NAA are lower with 0.84 for LCModel 0.92 for MRSNet; this is a stark contrast to E1 R^2 performance for both methods.

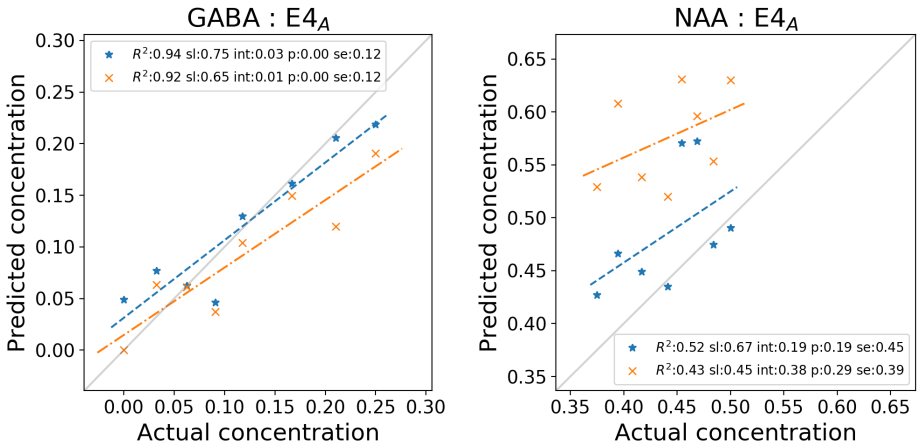


Figure 4.4: Regression analysis for MRSNet basis set (blue, $\star$) compared with LCModel analysis (orange, $\times$) for GABA (top row) and NAA (bottom row) for E4a.

Figure 4.4 shows the regression analysis for MRSNet compared with LCModel for E4a. As with E3, MRSNet is overall closer to the ideal quantification method represented by the grey line, in addition to having a more accurate slope for GABA and NAA, and a more consistent set of predictions indicated by the higher R^2 value. This is an interesting result, as the E4 phantom is the tissue-mimicking Gel and LCModel is designed to quantify *in-vivo* data. The E4 data has a substantially broader linewidth (4 Hz) than the datasets used to train MRSNet (1 Hz), suggesting that it can reasonably generalise broader line widths. It is important to recognise that both methods struggled with accurate quantification of NAA, but not GABA which is another surprising trend as the NAA signal is not hard to identify as it dominates in the spectra.

Finally, Figure 4.5 shows the regression analysis for the 2000 Hz bandwidth (BW) E4 phantom. Here it is clear that MRSNet struggles to quantify both metabolites with low R^2 values, a slope near 0.1 for NAA and a negative slope for GABA. This is likely due to the much higher linewidth for the E4b series which is around (8 Hz) for the NAA peak. LCModel outperforms MRSNet, with a comparatively reasonable set of predictions for GABA and NAA. However, as with E3 and E4a LCModel continues

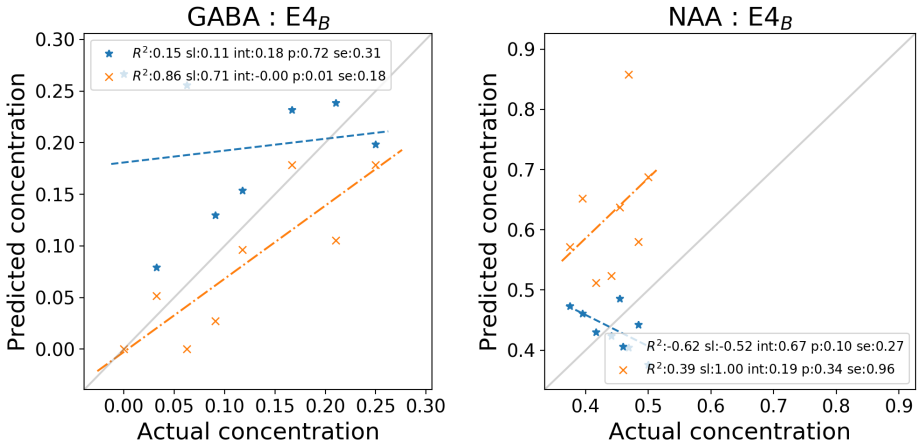


Figure 4.5: Regression analysis for MRSNet basis set (blue, $\star$) compared with LCModel analysis (orange, $\times$) for GABA (top row) and NAA (bottom row) for E4b.

to struggle to accurately quantify NAA, with a intercept of 0.19 and R^2 of 0.39.

4.2.4.2 *MAPE analysis*

Mean absolute percentage error (MAPE) is used to further understand and analyse the performance of the network on an individual metabolite basis for GABA and NAA across all benchmark datasets. This is done for both LCModel and MRSNet trained with the LCModel basis due the performance advantage shown in Table 4.14. A similar trend continues to the regression analysis, where MRSNet outperforms LCModel on average, except for GABA quantification which bounces between LCModel and MRSNet trained with the FID-A 'MLW' basis, suggesting that the GABA model from FID-A may be a better fit to experimental data.

The large error seen in Tables 4.12 and 4.13 can be attributed to the performance of NAA quantification, as it has the highest concentration in all benchmark phantoms. Any improvement seen in NAA quantification has the largest impact on the error in Eq. (4.2) and is reflected in the overall performance in Tables 4.12 and 4.13. While NAA is not typically the target of edited MRS, it can be utilised as an internal reference compound. So any improvement in the accuracy of NAA would indirectly improve quantification for all other metabolites.

Table 4.7: Network performance for different combinations of acquired spectra (edit-off, edit-on, difference) and data types (real: R, imaginary: I, magnitude: M) using the 'small' network with strided convolutions and the LCModel basis for training and validation dataset generation.

AQ	DAT	TRN	VLD	E1	E3	E4a	E4b
O	R	0.01386	0.01521	0.13292	0.10046	0.08394	0.10424
		σ0.011	σ0.014	σ0.089	σ0.080	σ0.065	σ0.084
O	I	0.01265	0.01494	0.13159	0.07989	0.09348	0.11512
		σ0.012	σ0.014	σ0.092	σ0.054	σ0.067	σ0.070
O	M	0.01565	0.01949	0.06863	0.07800	0.06419	0.07662
		σ0.016	σ0.024	σ0.054	σ0.063	σ0.050	σ0.054
O	RI	0.01203	0.01400	0.09873	0.10943	0.09235	0.08462
		σ0.011	σ0.013	σ0.064	σ0.075	σ0.067	σ0.072
O	RIM	0.01105	0.01366	0.10024	0.09750	0.08600	0.09015
		σ0.010	σ0.013	σ0.062	σ0.050	σ0.059	σ0.078
N	R	0.01453	0.01695	0.22904	0.13183	0.11652	0.17211
		σ0.013	σ0.017	σ0.192	σ0.124	σ0.100	σ0.077
N	I	0.01256	0.01570	0.22129	0.13696	0.12930	0.19410
		σ0.012	σ0.017	σ0.191	σ0.133	σ0.119	σ0.112
N	M	0.01944	0.02567	0.21743	0.15283	0.15292	0.16059
		σ0.019	σ0.030	σ0.200	σ0.100	σ0.098	σ0.095
N	RI	0.01369	0.01648	0.23515	0.13795	0.13242	0.19135
		σ0.012	σ0.018	σ0.182	σ0.130	σ0.114	σ0.101
N	RIM	0.01468	0.01851	0.23231	0.17241	0.18429	0.19757
		σ0.014	σ0.019	σ0.194	σ0.135	σ0.101	σ0.095
D	R	0.04300	0.04888	0.08484	0.09208	0.09231	0.10738
		σ0.041	σ0.047	σ0.073	σ0.059	σ0.071	σ0.081
D	I	0.03993	0.04649	0.06341	0.11511	0.11253	0.10329
		σ0.040	σ0.046	σ0.066	σ0.066	σ0.070	σ0.092
D	M	0.04142	0.05125	0.04967	0.08806	0.08591	0.10017
		σ0.041	σ0.050	σ0.063	σ0.069	σ0.066	σ0.074
D	RI	0.03963	0.04671	0.07428	0.07755	0.07148	0.08196
		σ0.040	σ0.047	σ0.066	σ0.047	σ0.052	σ0.063
D	RIM	0.03918	0.04647	0.07753	0.06767	0.08112	0.09246
		σ0.040	σ0.047	σ0.079	σ0.046	σ0.057	σ0.061

Table 4.8: Network performance for different combinations of acquired spectra (edit-off, edit-on, difference) and data types (real: R, imaginary: I, magnitude: M) using the 'small' network with strided convolutions and the LCModel basis for training and validation dataset generation.

AQ	DAT	TRN	VLD	E1	E3	E4a	E4b
ON	R	0.01396	0.01626	0.20498	0.18993	0.16938	0.14621
		σ0.011	σ0.014	σ0.159	σ0.146	σ0.135	σ0.082
ON	I	0.01425	0.01640	0.23126	0.16219	0.15234	0.19090
		σ0.012	σ0.014	σ0.170	σ0.128	σ0.120	σ0.095
ON	M	0.01415	0.01642	0.14484	0.11244	0.12549	0.13234
		σ0.012	σ0.016	σ0.118	σ0.086	σ0.078	σ0.070
ON	RI	0.01434	0.01636	0.18232	0.16390	0.15403	0.15630
		σ0.013	σ0.015	σ0.135	σ0.093	σ0.095	σ0.072
ON	RIM	0.01618	0.01801	0.15784	0.17001	0.12872	0.12797
		σ0.013	σ0.016	σ0.132	σ0.130	σ0.105	σ0.080
OD	R	0.01361	0.01440	0.06866	0.08917	0.07371	0.09721
		σ0.011	σ0.013	σ0.053	σ0.051	σ0.051	σ0.076
OD	I	0.01203	0.01358	0.07983	0.04921	0.05918	0.10686
		σ0.011	σ0.013	σ0.053	σ0.028	σ0.042	σ0.081
OD	M	0.01549	0.01759	0.05956	0.04519	0.04288	0.06034
		σ0.013	σ0.017	σ0.042	σ0.031	σ0.034	σ0.054
OD	RI	0.01292	0.01461	0.11167	0.05455	0.04936	0.10697
		σ0.011	σ0.013	σ0.090	σ0.043	σ0.037	σ0.082
OD	RIM	0.01344	0.01509	0.07955	0.11050	0.09290	0.10666
		σ0.012	σ0.015	σ0.057	σ0.074	σ0.059	σ0.085
ND	R	0.01704	0.01819	0.20824	0.14069	0.14328	0.17063
		σ0.014	σ0.016	σ0.177	σ0.106	σ0.103	σ0.080
ND	I	0.01439	0.01552	0.19468	0.13248	0.13502	0.17786
		σ0.012	σ0.014	σ0.161	σ0.087	σ0.084	σ0.098
ND	M	0.01888	0.02156	0.15265	0.10383	0.12364	0.12842
		σ0.018	σ0.022	σ0.131	σ0.078	σ0.076	σ0.056
ND	RI	0.01673	0.01840	0.22594	0.15966	0.15320	0.19460
		σ0.014	σ0.016	σ0.191	σ0.118	σ0.106	σ0.086
ND	RIM	0.01957	0.02295	0.23033	0.16489	0.17125	0.17671
		σ0.019	σ0.022	σ0.184	σ0.111	σ0.101	σ0.091

Table 4.9: Network performance for different combinations of acquired spectra (edit-off, edit-on, difference) and data types (real: R, imaginary: I, magnitude: M) using the 'small' network with strided convolutions and the LCModel basis for training and validation dataset generation.

AQ	DAT	TRN	VLD	E1	E3	E4a	E4b
OND	R	0.01326 σ0.011	0.01486 σ0.013	0.13909 σ0.086	0.09442 σ0.067	0.08806 σ0.058	0.10422 σ0.083
OND	I	0.01251 σ0.010	0.01393 σ0.012	0.16848 σ0.112	0.08873 σ0.081	0.09467 σ0.081	0.14555 σ0.085
OND	M	0.01336 σ0.011	0.01517 σ0.014	0.07679 σ0.055	0.06491 σ0.053	0.06009 σ0.040	0.06971 σ0.046
OND	RI	0.01624 σ0.013	0.01710 σ0.015	0.15457 σ0.107	0.13011 σ0.067	0.14073 σ0.074	0.15130 σ0.086
OND	RIM	0.01418 σ0.012	0.01612 σ0.014	0.14309 σ0.104	0.08350 σ0.073	0.09529 σ0.070	0.11978 σ0.081

Table 4.10: A comparison of basis set influence on MRSNet performance, using the 'small' network with stridden convolutions. Networks are trained using the magnitude of the edit-off and difference acquisition. MLW denotes that the dataset is comprised of multiple linewidth spectra.

Basis	TRN	VLD	E1	E3	E4a	E4b
LCModel	0.01206 σ0.009	0.01547 σ0.017	0.07936 σ0.061	0.04553 σ0.035	0.04991 σ0.038	0.06702 σ0.046
FID-A	0.01120 σ0.009	0.01374 σ0.013	0.04744 σ0.035	0.05323 σ0.036	0.04670 σ0.033	0.05960 σ0.049
FID-A MLW	0.00995 σ0.008	0.01308 σ0.013	0.04773 σ0.035	0.05963 σ0.041	0.05407 σ0.040	0.06302 σ0.051
PyGamma	0.01337 σ0.012	0.01688 σ0.017	0.09081 σ0.071	0.09496 σ0.057	0.09396 σ0.064	0.10269 σ0.066
PyGamma MLW	0.01133 σ0.009	0.03436 σ0.037	0.08465 σ0.061	0.07521 σ0.058	0.06994 σ0.062	0.08147 σ0.064

Table 4.11: Metabolites used to compare MRSNet to the state-of-the-art.

	E1	**E3**	**E4a**	**E4b**
Metabolite set	[NAA, GABA]	[NAA, GABA, Glx]	[NAA, GABA, Glx]	[NAA, GABA, Glx]

LCModel, VeSPA and jMRUI report individual Glu and Gln values which have been combined into Glx. GANNET and TARQUIN report the combined Glx.

Table 4.12: State-of-the-art quantification performance.

Analysis program	**E1**	**E3**	**E4a**	**E4b**
VeSPA	0.1192 σ0.096	0.1199 σ0.081	0.1552 σ0.105	0.2016 σ0.122
TARQUIN	0.1125 σ0.093	0.0708 σ0.049	0.0705 σ0.070	0.1341 σ0.090
LCModel	0.0462 σ0.030	0.0956 σ0.041	0.0919 σ0.054	0.1266 σ0.095
AQSES †	0.2081 σ0.092	0.1432 σ0.088	0.1865 σ0.117	-
QUEST †	0.2044 σ0.098	0.1402 σ0.080	0.1958 σ0.144	-
GANNET	0.1405 σ0.103	0.3092 σ0.158	0.3012 σ0.179	0.5334 σ0.603

† Quantification was not performed for the E4b dataset.

Table 4.13: MRSNet quantification performance.

MRSNet (basis source)	**E1**	**E3**	**E4a**	**E4b**
LCModel	0.0318	0.0321	0.0402	0.0720
	σ0.014	σ0.021	σ0.022	σ0.050
FID-A	0.0746	0.0617	0.0402	0.0735
	σ0.044	σ0.034	σ0.033	σ0.060
FID-A MLW	0.0636	0.0714	0.0722	0.0864
	σ0.032	σ0.044	σ0.041	σ0.054
PyGamma	0.0408	0.0842	0.0674	0.0994
	σ0.021	σ0.057	σ0.041	σ0.063
PyGamma MLW	0.0501	0.0708	0.0679	0.0704
	σ0.022	σ0.044	σ0.044	σ0.055

Table 4.14: Individual metabolite MAPE for state-of-the-art quantification programs compared to the best network ('small' with stridden convolutions). Multiple networks have been trained with different basis sets, including multi-line width (MLW) bases.

	E1		**E3**			**E4a**			**E4b**		
	NAA	GABA	NAA	GABA	Glx	NAA	GABA	Glx	NAA	GABA	Glx
Program											
TARQUIN	16.26%	69.98%	10.64%	47.63%	18.62%	13.04%	158.19%	28.55%	7.93%	100.93%	27.17%
VeSPA	16.96%	66.64%	43.52%	49.96%	23.00%	48.17%	82.69%	43.79%	54.42%	217.30%	55.48%
LCModel	6.53%	24.62%	35.35%	34.68%	21.82%	31.08%	35.20%	24.61%	42.49%	39.08%	33.62%
jMRUI (AQSES)	24.12%	221.87%	18.65%	133.39%	40.34%	18.37%	312.78%	57.14%		-	
jMRUI (QUEST)	23.47%	215.59%	12.90%	134.90%	47.68%	21.50%	332.47%	57.12%		-	
GANNET	20.13%	76.41%	114.12%	57.56%	87.59%	102.48%	44.97%	90.06%	99.92%	139.99%	135.18%
MRSNet (trained with)											
LCModel	3.89%	25.29%	4.59%	27.93%	10.53%	5.15%	52.59%	13.30%	11.22%	104.94%	17.08%
FID-A	10.37%	42.56%	22.63%	18.20%	15.02%	11.60%	30.97%	10.42%	13.00%	85.83%	16.41%
FID-A MLW	8.71%	36.90%	25.87%	11.45%	20.08%	22.85%	29.17%	19.70%	20.05%	84.95%	21.46%
PyGamma	5.18%	30.33%	14.13%	57.37%	29.63%	5.71%	123.52%	19.85%	13.35%	160.74%	22.14%
PyGamma MLW	6.14%	44.89%	13.15%	53.56%	23.65%	8.61%	83.27%	22.63%	7.00%	87.71%	22.52%

4.2.5 *Basis Comparison*

The variety of performance shown in the basis comparison experiment and subsequent MAPE analysis (see Tables 4.10 and 4.14) can be explained by the variance of spectra generated from three data sources. As mentioned in Section 2.3.2, dataset and simulator creators have a large range of options available to them for simulation parameters, along with a range of metabolite models, resulting in a large range of possible spectra.

Fig. 4.6 compares the difference spectra from PyGamma, FID-A and LCModel with an additional in-house experimentally acquired pure GABA phantom spectrum using MEGA-PRESS. From visual inspection, it is clear to see that none of the simulated data sources aligns perfectly with the experimental spectrum. This may be due to a multitude of reasons, including each simulator using a different GABA model: PyGamma uses Govindaraju Govindaraju *et al.* [45], FID-A uses Near *et al.* [103] and LCModel uses Kaiser *et al.* [60]. All simulators use Govindaraju *et al.* [45] values for the other metabolite models. A different GABA model is often chosen due to the well-known issue that the values in [45] are rough approximations of the true values [70]. In addition to simulation parameters, there is a multitude of experimental factors that can alter the resulting spectrum, further complicating the issue of matching simulation to experimental results. In an ideal scenario, networks should be trained on a large range of experimental data to cover the potential variation of spectra. However, in practice, this is a non-trivial task, with phantom creation being difficult and time-consuming in addition to needing a large amount of data to cover the potential experimental variations of the spectra.

4.3 EVALUATION AND DISCUSSION

MRSNet is quick to train on GPUs (18 minutes on an NVIDIA Titan X GPU), with a relatively low number of samples (5,000). Once trained, it requires no special hardware and it takes on average 24 ± 2ms to process a single spectrum with the network, using a dual-core 2.7 GHz i7-7500u CPU, with the network occupying 95MB disk space. Furthermore, it requires no interaction from a user with specialist knowledge for potentially complicated processing steps, such as phase-correction.

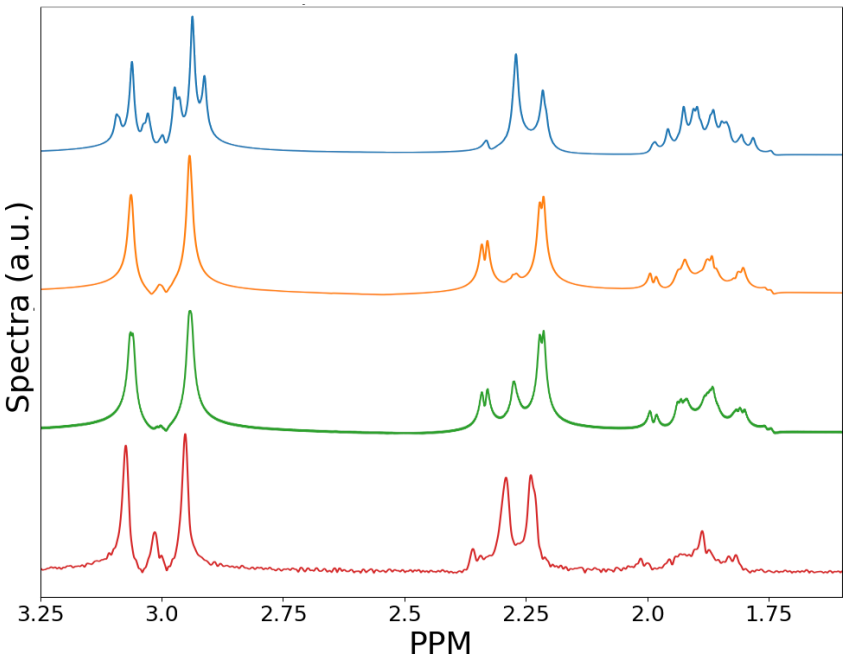


Figure 4.6: Example of spectra differences for the magnitude MEGA-PRESS difference GABA spectra from multiple sources, from top to bottom: PyGamma, FID-A, LCModel, experimental spectrum.

The data type and channel experiments in Section 4.2.2 is not generally applicable to all basis sets in Section 4.2.3, as this study was only performed using the LCModel basis set. As shown, other basis sets have different spectra and may have a more accurate representation of real or phase data, but it is expected that the magnitude spectra will remain the best performing due to the issue of uncertainty in phase reconstruction from the scanner.

Networks in this thesis are only trained and tested with one timing variant of a single pulse sequence (MEGA-PRESS at 3 T with $T_E = 68$ ms, $T_R = 2000$ ms) with a benchmark dataset collected from one scanner, and for a single frequency window (4.5 ppm to 1.5 ppm). This network should generalise to different MEGA-PRESS spectra, but it is unlikely that it will work as accurately with different strength B_0 fields, scanner manufacturers, pulse sequences or pulse sequence timings without further training.

The performance improvements that MRSNet has over current quantification methods as shown in Table 4.14 may be due to the fact that it uses two spectra as an input (edit off and difference), compared to the single difference spectra used by the quantification programs. This additional spectrum naturally provides more information to the networks. This is a benefit of utilising CNNs, a network could be trained to quantify a much

larger number of spectra at once, such as stacking multiple acquisitions that vary the T_E time. This variation of T_E alters the spectra considerably *in-vivo*, where individual protons in metabolites relax at different rates, shifting the intensity of spectral peaks, thus providing more information to the networks. This multi T_E implementation would of-course require retraining the network to learn and accept the new input data format and would be an interesting avenue to explore for future work.

One major disadvantage of the MRSNet architecture is that it has been designed under the assumption that there will always be at least one of the metabolites it has been trained to detect present. The softmax normalisation function on the final layer of the network normalises all the outputs to have a sum of one, which restricts the network to always report a ratio and to never report that there are no metabolites present. This is not necessarily a concern for an *in-vivo* application, as there would never be a spectra that does not contain NAA, GABA, Cr, Glu or Gln. This is more a consideration for when testing the networks, to ensure that there is at least one metabolite present in the phantoms or generated datasets. This issue is a consequence of designing a normalisation scheme for both the input spectra and output labels to the network. A route for future work would be to explore different scaling and normalisation schemes for the input and output data, such as using linear or sigmoid based outputs.

Finally, *in-vivo* data has not been used to evaluate this method due to the lack of ground truth data, which would, of course, be very difficult to obtain. Additionally, we have chosen to explore a limited number of metabolic signals, while there is a much larger range of spectra and macromolecule signals that are obtainable with MRS *in-vivo*. This is something to be explored as future work, expanding the range of metabolites that the network is able to detect.

4.4 SUMMARY

The contribution in this chapter is novel for a range of reasons: it is the first work that quantifies a multi-acquisition protocol such as MEGA-PRESS. In addition, there is a great deal of depth into how these multi-acquisitions can be represented to the neural network. This has shown that the edit-off and difference absolute spectra provide the best results when optimising fitting for the benchmark datasets. Additionally, networks are trained on a range of simulated data sources

(PyGamma, LCModel, FID-A) to compare performance, and are finally benchmarked against current state of the art methods, such as LCModel, jMRUI (AQSES and QUEST), TARQUIN, GANNET and VeSPA. This has shown that in the general case, training the network should be done with the FID-A basis, however, when looking at a reduced set of metabolites that excludes Cr, the LCModel basis outperforms the alternatives. This highlights the non-trivial task of basis set selection and the need for more accurate characterisation.

This chapter has demonstrated that the overall accuracy and precision of metabolite quantification in MRS is improved by a convolutional neural network by comparing its performance to current state-of-the-art methods. We have found that a 2D CNN, using stridden convolutions, and utilising 'small' 1×N convolutions along the frequency axis of the input spectra is the best performing network architecture.

Finally, all of the code for collection and simulation of basis sets along with training and testing of the network have been made publicly available here [14].

5

PULSE SEQUENCE OPTIMISATION

This chapter focuses on a different approach to improving quantification. In previous chapters the approach was focused on post-data-collection methods improving MRS through quantification tools. This chapter investigates improving the data coming from the scanner with a novel method of pulse sequence design that enables a chemically selective pulse-sequence. These control pulses are optimised through the use of quantum control techniques applied to metabolites within a simulated MRS environment.

The goal of this chapter is to determine if there are suitable controls that can be designed and be successful in a substantially more difficult environment than typically seen in quantum optimal control. This is a novel application, where the range of inhomogeneity in MRS are large and vast which propose a significant challenge. As such a method has been developed that allows the optimisation of controllers that are robust with respect to a range of B_0 and B_1 instabilities.

5.1 INTRODUCTION

We propose a method to find excitation pulses that ensure the spectra for different metabolites are sufficiently distinct to significantly simplify the analysis of the resulting combined spectra. For each metabolite of interest, a target quantum state is pre-selected that ensures the resulting spectra do not overlap. Optimal control techniques for quantum systems are used to find RF pulses that simultaneously achieve the desired state for multiple metabolites and with that a distinct spectrum for each metabolite. Moreover, the pulse sequences are optimised such that they are robust to a range of environmental instabilities and typically MRI scanner uncertainties.

Two metabolite sets are used to demonstrate the method in simulation: (i) Gln and Glu; (ii) NAA, Cr and GABA. Both of these targets are challenging due to significant overlap and large differences in concentrations. These metabolites are of particular clinical and research interest, as previously introduced in Section 2.3.1.

Computational results indicate that suitable pulses that are robust to B_0 and B_1 instabilities can be found. They are theoretically realisable on current 3 T MR scanners, enabling the separation of GABA, NAA, Cr as well as Gln, Glu in the combined spectra with a single RF pulse for each target. Additionally, we show that these pulses can be optimised within realistic energy and time constraints ($\leqslant 100\,ms$), and remain robust to scanner instabilities such as inhomogeneity of the main magnetic field (B_0 $\pm 5\,Hz$) and uncertainties in the RF control fields ($B_1 \pm 2.5\%$).

In this section, we introduce the underlying physics of the simulation, our algorithm for optimising the RF pulse, the target functional and target state selection. Our algorithm finds an RF control pulse to simultaneously place multiple metabolites into states such that their spectra are easily distinguishable. It is based on an optimisation approach from optimal control of quantum systems [83] and requires models of the hydrogen nuclear spin dynamics. We first describe the spin dynamics models, how simulation is performed, and outline our control algorithm. Finally, the algorithm is adjusted to optimise RF pulses within realistic ranges of expected uncertainties in the system and realistic scanner limits.

5.2 METHOD

A simple pulse sequence is optimised that is composed of a single RF pulse followed by an immediate readout. Control amplitudes, duration and number of steps are initially randomly initialised and optimised with the gradient-based, L-BFGS with bound-constraints (L-BFGS-B) [9, 98, 174], as this has been shown to provide excellent performance [83, 157]. Control amplitudes are updated concurrently to reduce the number of computationally expensive function calls.

Simulation for this method has been introduced and covered in Section 2.2 of the background of this thesis. Metabolites are simulated as closed-systems to reduce the computational complexity. Simulation and optimisation code is implemented in a MATLAB framework developed for the optimisation of MR RF pulses; QControl [15]. Computation is divided among several machines, with each machine responsible for one optimisation run (one set of metabolites and pulse sequence parameters). No additional parallelisation was done, as these were found to be detrimental in testing, leading to an increased runtime and memory requirement when using MATLAB due to the overhead of Matlab's `parpool` parallelisation method.

Optimisation of RF pulses is done in two phases. Initially, pulses are found for a 'simple' target, which is a set of systems at resonance with a set of randomly generated values for duration, the number of time steps and initial RF pulse. The second stage of optimisation performs the much more computationally expensive and harder robust optimisation, where the starting point is one of the best pulses from the initial phase with an error $\geqslant 0.05$. This is done to reduce the total overall time spent optimising pulses, by finding the best case initial values to optimise for. There is an assumed relationship here; pulses that can drive the set of metabolites into the desired state for the simpler non-robust case would be a good starting point for the robust optimisation.

Target state selection and controllability

Target states represent the orientation of each of the spins in the metabolite. By individually selecting spins to be placed in the transverse plane, it is possible to manipulate the resulting spectrum, by leaving the other spins in the ground state, in alignment with the field.

For simplicity, target states for a group of metabolites are chosen manually by selecting specific resonances per metabolite that create non-overlapping spectra with a maximal distance between peaks. These are selected by inspecting the chemical shift values of all protons of the metabolites in the set and choosing a combination that has the maximum distance between metabolite peaks. Several similar targets are chosen per metabolite set later in Section 5.3 to maximise the possibility of finding a pulse sequence that achieves one of the targets. Target state selection is currently done manually for simplicity, however, this could be automated by inspecting the chemical shifts of each metabolite and selecting combinations that maximise range and uniformity.

Additional resonances are added to the target state if they are within $\pm 0.1\,\mathrm{ppm}$ of the main chosen resonance. This is primarily done to enable robust optimisation w.r.t. B_0 further down the line; as the B_0 inhomogeneity increases, the frequency selective components of the control pulses become less accurate, affecting neighbouring spins. As we are simulating a 3 T scanner, $\pm 0.1\,\mathrm{ppm}$ is roughly equal to $\pm 12.6\,\mathrm{Hz}$ so gives us room for drift in B_0 where later we optimise pulses for a $\pm 5\,\mathrm{Hz}$ B_0 range. Secondly, protons that have similar chemical shifts in a metabolite are typically strongly coupled, potentially making the target harder, as it would need to precisely excite only one of the spins.

These selected resonances are translated to a target density matrix and placed into a superposition with the ground state. These targets are equivalent to a 90° rotation around the x-axis to bring target spins into the transverse plane, ready for readout. Spins that are not selected for excitation are left in the ground state, in alignment with the B_0 field.

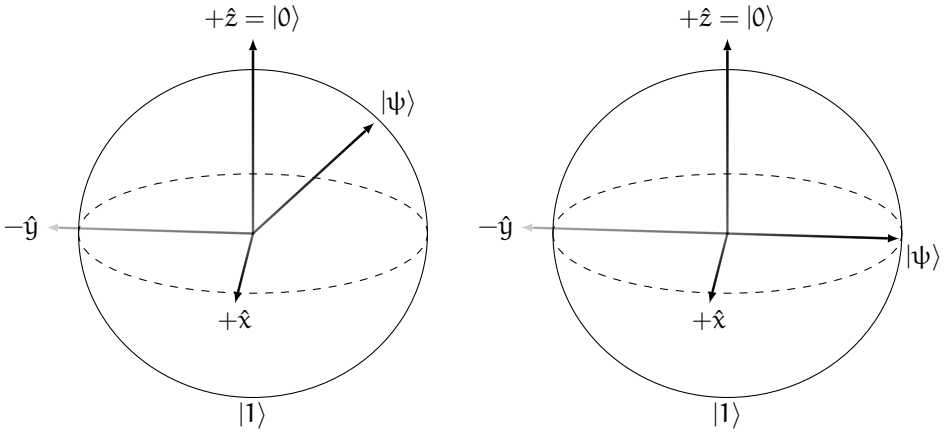
Target states that select more than one spin naturally create more complex targets, as a larger number of excitation subspaces need to be utilised to achieve a full 90° rotation. An excitation subspace is a logical-grouping of the state probabilities, in this instance states that represent a single excitation are grouped into the first excitation subspace, e.g. for a two spin system: $|01\rangle$ & $|10\rangle$. For the second excitation subspace, this would be $|10\rangle$ for a two spin system, and so on.

For example, for a three spin system, where two spins have been selected as the target the states in superposition would be:

$$|\psi_T\rangle = |000\rangle + |100\rangle + |010\rangle + |110\rangle\,. \tag{5.1}$$

The target state vectors would be expanded to the density matrix as in Equation 5.2.

$$\rho = \sum_j p_j \left|\Psi_j\right\rangle \left\langle\Psi_j\right|, \tag{5.2}$$



(a) $|000\rangle + |110\rangle$ (b) $|000\rangle + |100\rangle + |010\rangle + |110\rangle$

Figure 5.1: Projection of multiple excitation target states on the Bloch Sphere.

Figure 5.1 shows a projection of the effect of using only the second excitation subspace if there are two spins in the targets state. To achieve a full 90° rotation into the transverse plane, the number of excitation subspaces utilised must match the number of excited spins.

Using multiple excitation subspaces makes the targets harder to optimise, as transitions between subspaces are not possible from the system

dynamics alone; energy must be driven between the subspaces sequentially by the control fields $u_x(t)H_x$, $u_y(t)H_y$, or lowered down with the dissipative lowering operator σ^-. As we are optimising over a closed-system simulation, only the RF control fields can facilitate energy transfer between the subspaces.

Simple, one spin targets only require one subspace transition and balancing the energy between two total subspaces (ground and first). Where for two spins, this becomes two subspace transitions and balancing the energy between three subspaces (ground, first and second), and so on. This balancing act becomes harder once optimisation is performed for multiple metabolites concurrently, each with their targets, potentially exciting a different number of subspaces. This is ultimately where the difficulty lies in the optimisation.

It is worth mentioning, that there is a multitude of other targets that could be optimised for. In this instance, the goal is to still receive a signal from each of the metabolites in the set. It could be easily changed to only receive a signal from one metabolite by setting the target states for the remaining molecules in the simulated set to remain in the ground state.

Finally, it is important to touch on the motion of controllability of quantum systems. As we are working with a set number of pre-determined controls $[H_x, H_y]$, that have hard limits on the maximum amplitude they can impose, in addition to having restrictions on how short each Δt can be due to a scanner transmit limitation of 2.5×10^{-8}ms, we expect that targets states will not be fully obtainable. Besides, there will be some metabolites that do not satisfy requirements for complete controllability as outlined in [127, 128]. However, in this instance, we do not need to have full control, and it will be possible to reach a near-optimal solution that will produce the desired unique spectra, given enough time and with the appropriate targets. The constraints are ultimately the largest issue in regards to the controllability of these systems, with the effect of this displayed as the Pareto front visible in the optimisation landscape Figures 5.10 and 5.11 for error over the duration.

Target functions

Two target functions are defined, one for a simple optimisation, where the set of metabolites is simulated without any B_0 or B_1 instabilities, with the second optimising the best pulses from the first phase over a range of B_0 and B_1 instabilities. This is done to find good candidate solutions to the much harder and computationally expensive robust optimisation,

this eliminates settings and pulses early on that would perform poorly as robust controllers.

Both of the target functions aim to minimise the error when transitioning from an initial state ρ_0 to a pre-determined target state ρ_{tar} at time T:

$$f(u, T) = \mathrm{Re}\left(\mathrm{Tr}\left\{\rho_{tar}\rho(T)^\dagger\right\}\right) \tag{5.3}$$

Where $\rho(T) = U_T\rho_0U_T^\dagger$ and U_T is computed according to Eq. (2.10) and ρ_0 is the initial state of the system.

As the eigendecomposition is readily available from the calculation of the propagators, it is quick to compute the gradients using the spectral method. The numerical gradient is used as it performs substantially better than the approximate gradient for L-BFGS [83].

As introduced in Section 2.2, the piecewise constant controls form a matrix $u = [u_{j,k}]$ where j is the control signal index and k the time interval index in $[0, K]$. The derivative for the gradient with respect to a single control amplitude $u_{j,k}$ is as follows:

$$\begin{aligned}\frac{\partial f(u,T)}{\partial u_{j,k}} = &\mathrm{Re}\left(\mathrm{Tr}\left\{\rho_{tar}U_K\cdots\left(\frac{\partial U_k}{\partial u_{j,k}}\right)\cdots U_0\rho_0U_0^\dagger\cdots U_K^\dagger\right\}\right.\\ &\left.+\,\mathrm{Tr}\left\{\rho_{tar}U_K\cdots U_0\rho_0U_0^\dagger\cdots\left(\frac{\partial U_k^\dagger}{\partial u_{j,k}}\right)\cdots U_K^\dagger\right\}\right)\end{aligned} \tag{5.4}$$

with $\frac{\partial U_k}{\partial u_{j,k}}$ as follows:

$$\begin{aligned}&\left\langle\lambda_l\left|\frac{\partial U}{\partial u_{j,k}}\lambda_m\right.\right\rangle =\\ &\begin{cases}-i\Delta t\left\langle\lambda_l\right|H_j\left|\lambda_m\right\rangle e^{-i\Delta t\lambda_l} & \text{if } \lambda_l = \lambda_m\\ -i\Delta t\left\langle\lambda_l\right|H_j\left|\lambda_m\right\rangle\frac{e^{-i\Delta t\lambda_l}-e^{-i\Delta t\lambda_m}}{-i\Delta t(\lambda_l-\lambda_m)} & \text{if } \lambda_l \neq \lambda_m\end{cases}.\end{aligned} \tag{5.5}$$

The derivative is given element-wise in the orthonormal eigenbasis $\{|\lambda_i\rangle\}$ to the real eigenvalues $\{\lambda_i\}$.

Full details are out of the scope of this thesis, but an extensive explanation of this target function can be found in [83]. As optimisation is done on a set of metabolites, the target function is run individually per metabolite, with the error and gradient values being averaged over the set.

Algorithm 1 Pseudo-code of robust error function.

```
function ErrorFunction(ctrl, metabolites, targets, B0 offsets, B1 offsets)
    err, grad, n ← [], [], 0
    for all (M, T) in (metabolites, targets) do
        err(n), grad(n) ← eval(ctrl, M, T, 0, 0)
        n++
        for B0 in B0 offsets do
            err(n), grad(n) ← eval(ctrl, M, T, B0, 0)
            n++
        end for
        for B1 in B1 offsets do
            err(n), grad(n) ← eval(ctrl, M, T, 0, B1)
            n++
        end for
    end for
return mean(err), mean(grad)
end function
```

To produce pulses that are robust to environmental factors, a range of B_0 and B_1 instabilities are integrated into the target function shown in Algorithm 1. Both are introduced by simulating an additional set of metabolites with a given level of B_0 or B_1 distortion in a cross from resonance, where one arm is responsible for simulating a $\pm$ B_0 range, with the other simulating the $\pm B_1$ range. A cross has been chosen due to the large increase in computational time, simulating an entire grid would be computationally prohibitive as displayed in Table 5.5. Error and gradient are calculated in the usual way and is averaged out over all metabolites. Here, we use a B_0 range of ± 5 Hz at 1 Hz intervals, with a B_1 range of $\pm 5\%$ at 1% intervals, which requires 21 sets of metabolites to be simulated.

B_0 is included by simulating metabolite with an additional σ_z offset to the system Hamiltonian. The larger the optimised B_0 range the more chance of proton resonances overlapping with others, leading to frequency selective components of the pulse driving additional spins, creating an extremely difficult control problem. B_1 instabilities are more difficult to characterise, where potential noise sources, distortions and environmental effects are numerous. To cover a large range of sources of distortion, B_1 instability is integrated by modulating the global pulse amplitude and simulating the effect on the metabolites.

Optimisation for the robust target uses the best non-robust pulse sequences as the starting point for the harder target, this has been done to reduce the total time needed to compute these pulses. The assumption is that if we cannot produce successful controllers for the simplified problem, we can rule out using those initial conditions for the more complex, robust optimisation. Additionally, when attempting to optimise robust pulses from an initial random or zero pulse, it often fails within a few iterations as it fails to converge due to flat gradients, by seeding the optimisation we avoid this issue, but we may be converging to a local-minima.

Pulses are chosen as the starting point for optimisation if they have not been used for a robust optimisation before, and if they have a final error of $\leqslant 0.05$, effectively filtering out the poorly performing controllers before running the much more expensive optimisation. This allows us to save time computing pulses for settings that would never work effectively, such as those that are too short in duration.

Constraints

Scanner limits are integrated into the generation and optimisation of the RF pulse, based off the 3 T Siemens Magnetrom located at Swansea University, UK. RF amplitude generation and optimisation are capped at $800\,\mathrm{Rad/s}$, in line with experimentally validated values, this limit is imposed by running the box constrained optimiser L-BFGS-B. Generation of the minimum RF transmit time is limited to $25\,\mathrm{ns}$, in-line with the scanner limitation. In addition, we are aiming to achieve a total RF duration under $100\,\mathrm{ms}$, as the scanner does not allow single pulses longer than $100\,\mathrm{ms}$. There are schemes to enable longer pulses, by stacking them back-to-back, but the ideal candidate pulses will be under $100\,\mathrm{ms}$ in duration.

Optimisation conditions

All metabolites start in the ground state, aligned with the B_0 field. Simulation Larmor frequency is 2.89 T (123.32 MHz), in line with the B_0 strength of the 3 T scanner at Swansea University, UK. All metabolite models use the values from the landmark paper by Govindaraju *et al.* [45], however, issues of model accuracies are discussed later on in 5.4.4. The pulse sequence being optimised is an RF pulse followed by immediate readout; the duration and number of time steps of the pulse are randomly chosen within the following ranges: $\Delta t[1 \times 10^{-5} : 5 \times 10^{-5} : 2 \times 10^{-1}]$ time steps

: [10 : 10 : 2000]. With the RF being randomly generated and optimised within the scanner limits of $\pm 800\,\mathrm{rad/s}$.

5.3 RESULTS

We explore the effectiveness of this method with two metabolite sets that are particularly relevant to current research; Glx : [Glu, Gln] and GABA: [Cr, NAA, GABA].

5.3.1 *Glutamine and Glutamate*

Glutamine-Glutamate as a target proposes a significant challenge, as they share a very similar chemical structure. As a consequence they produce heavily overlapping spectra as shown in Figure 5.2, both are tightly coupled producing complex multiplet structures between 2 ppm and 3.8 ppm. This is a substantial problem when scanning at 3 T or below, they are usually jointly combined to Glx as the peaks are incredibly difficult to resolve accurately.

Multiple targets have been selected, as shown in Table 5.1. These targets provide maximal distance between peaks, removing the overlap between the two metabolites.

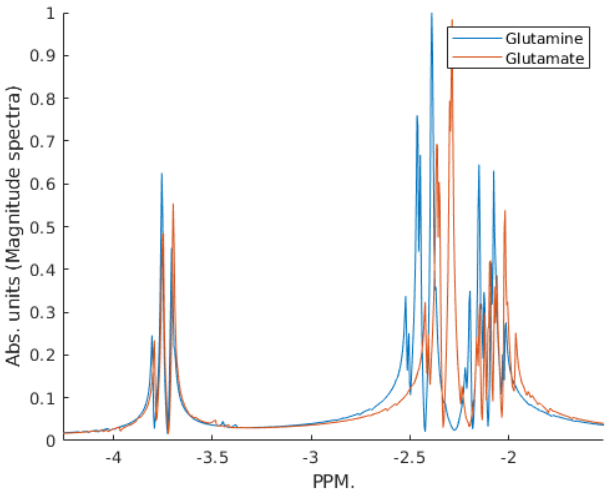


Figure 5.2: Combined Glx spectrum from a simulated FID pulse sequence.

Figure 5.4 displays the best non-robust spectra achieved for each of the targets. All of the spectra meet the objective, they have no overlapping peaks, and are individually resolvable.

Table 5.2 shows the best-achieved error values for the Glx targets for the non-robust and robust optimisation. Overall the results show that pulse sequences can be achieved for this difficult target, and can reach high fidelities. As expected the error for the robust optimisation is increased but is still acceptable.

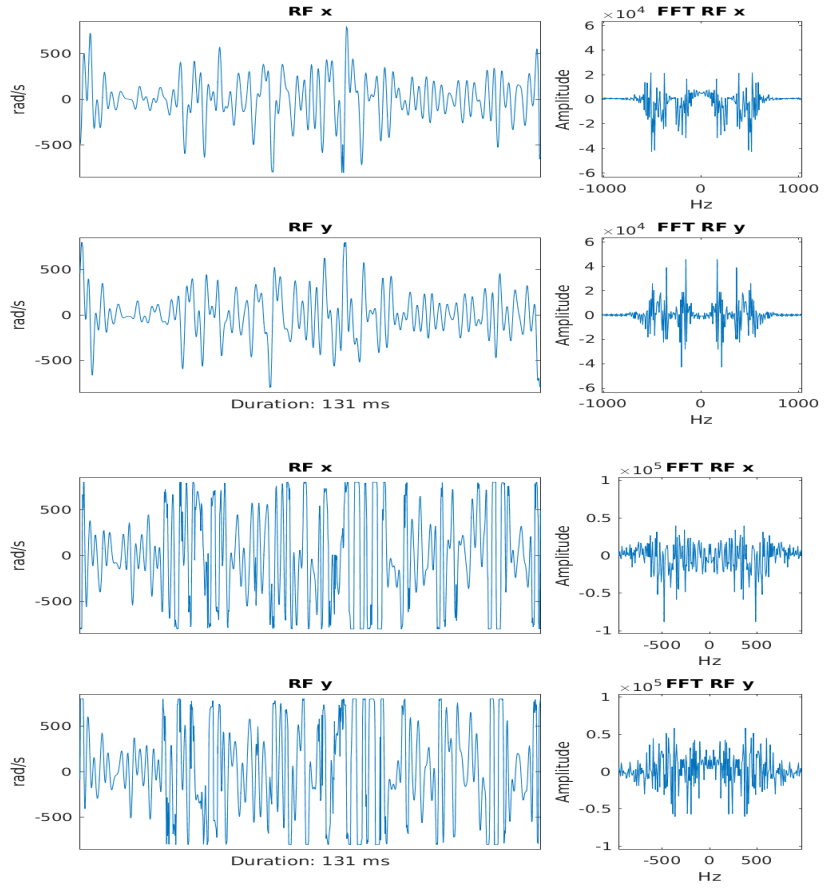


Figure 5.3: Best robust pulse for the Glx 1 (bottom), with the non-robust pulse used as the starting point for the optimisation (top). Duration:131 ms, $\Delta t = 1.3 \times 10^{-4}$ms, 1000 time steps.

Figure 5.3 shows the best pulse optimised for the robust Glx 1 pulse, and the non-robust pulse used as the starting point for the robust optimisation. It is clear to see how complex these pulses are, compared to the traditional pulses that are typically much smoother, such as a Gaussian sync pulse. The non-robust case as is relatively smooth, with only a few points hitting the maximum amplitude. In contrast, the pulse for the robust target is substantially more aggressive, frequently reaching near the maximum allowed amplitude. This is an indication of how much harder the robust targets are to reach, and how much more control is required for a high fidelity result.

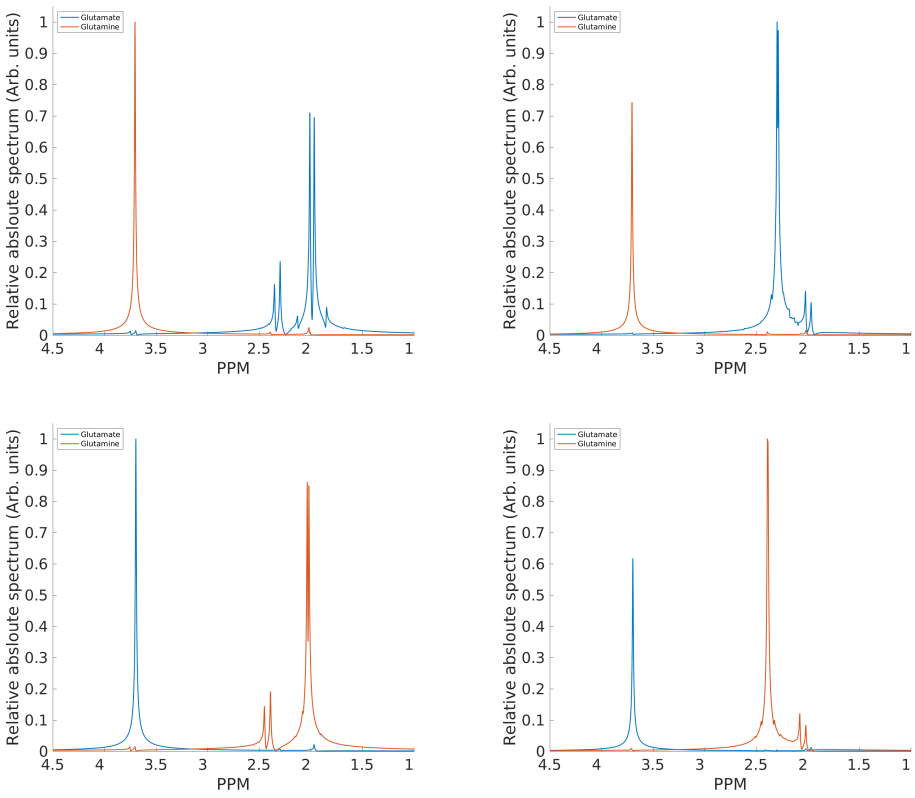


Figure 5.4: The best achieved spectra for the non-robust Glx targets. Top row: Glx 1, Glx2 Bottom row: Glx 3 Glx 4.

Table 5.1: Target resonances for Glx, peak values are in ppm. Peaks that fall within ±0.1 ppm range are grouped into the same target.

	Glx Target			
	1	2	3	4
State				
Gln	$\|01100\rangle$	$\|00011\rangle$	$\|10000\rangle$	$\|10000\rangle$
Glu	$\|10000\rangle$	$\|10000\rangle$	$\|01100\rangle$	$\|00011\rangle$
Resonances (ppm)				
Gln	2.10 2.12	2.43 2.45	3.75	3.75
Glu	3.74	3.74	2.03 2.12	2.33 2.35

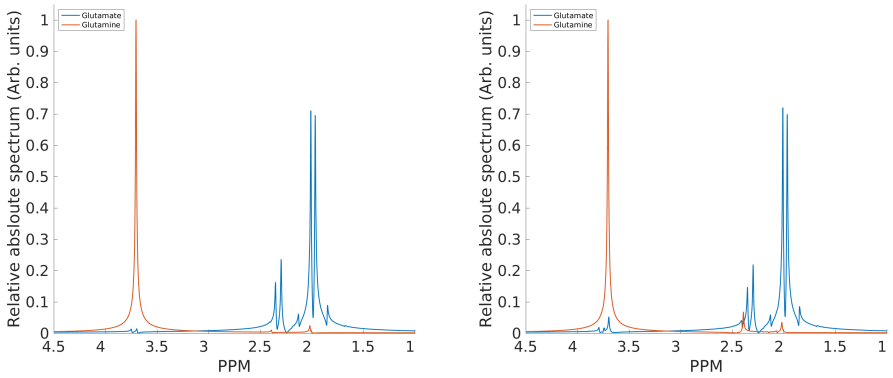


Figure 5.5: A comparison of the best spectra for Glx Target 1 for the non-robust optimisation on the left and the robust optimisation on the right.

The target presented here is a significant challenge, due to the strength of the couplings and the similarity of the structure of the metabolites. These factors combined with the optimised B_0 drift of ±5 Hz which is equivalent to ±0.04 ppm at 123.23 MHz. This means that some of the protons are not individually addressable with frequency-selective pulses as each resonance from Glutamine will overlap with at least one from Glutamate, and vice versa.

However, despite this, Figure 5.5 shows a side by side comparison of the best spectra achieved for Glx 1 in the non-robust and robust case. The results clearly show the similarities between the spectra, with minimal distortion between them. This is further backed up by the best performance errors shown in Table 5.2, where the difference in fidelity for the non-robust and robust pulses is 3.4×10^{-3}, however, this translates to a minimal difference in the spectra.

Table 5.2: Performance of Glx targets. "# Runs" refers to the number of completed optimisation runs.

		Best overall		**Best under** 100 ms	
Glx Target	# Runs	Error	Duration	Error	Duration
Non-robust					
1	345	5.3×10^{-5}	186 ms	5.3×10^{-3}	96 ms
2	356	1.7×10^{-3}	176 ms	1.8×10^{-2}	96 ms
3	344	9.2×10^{-5}	193 ms	3.9×10^{-3}	96 ms
4	361	2.3×10^{-3}	171 ms	1.9×10^{-2}	96 ms
Robust					
1	177	1.7×10^{-2}	131 ms	4.9×10^{-2}	94 ms
2	139	3.2×10^{-2}	151 ms	6.0×10^{-2}	99 ms
3	169	2.8×10^{-2}	175 ms	6.6×10^{-2}	91 ms
4	161	2.2×10^{-2}	191 ms	6.0×10^{-2}	98 ms

Due to the high fidelity achieved of the Glx 1 target, it will be chosen going forward as the main target for analysis of the robustness optimisation.

5.3.2 *GABA, Creatine and NAA*

GABA has six observable spins arranged into three pairs of magnetically equivalent nuclei. There are three resulting resonances — two triplets at 2.28 ppm and 3.01 ppm, and a quintet at 1.89 ppm. Detection of GABA

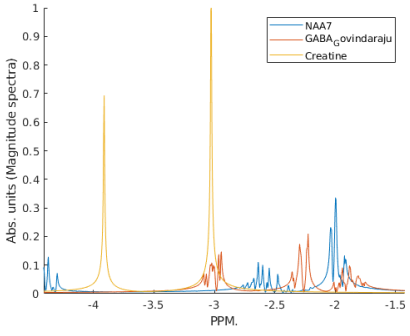


Figure 5.6: Combined GABA, Cr and NAA spectrum from a simulated FID pulse sequence.

using MRS is complicated by the presence of other, more plentiful molecules that produce interfering signals. For instance, NAA, Cr, and Glx all produce peaks in the same range as the GABA resonances as shown in Figure 5.6. It is possible to acquire *in-vivo* GABA spectra using J-difference editing techniques such as MEGA-PRESS, which removes the obscuring Cr signal at 3 ppm. However such techniques are susceptible to signal instabilities such as frequency drift [155].

Multiple targets have been selected, to produce the largest peak distance between the metabolites, as shown in Table 5.3. The high-frequency protons (6.9 ppm Cr and 7.4 ppm NAA) are ignored, as they are not conventionally used in MRS.

Table 5.4 displays the error achieved for optimisation of the GABA targets in the robust and non-robust optimisation with Figure 5.8 displaying the best spectra achieved for the non-robust case. Similarly to the **Glx** targets, the results are excellent, with all targets achieving a higher fidelity when compared to the Glx targets. Additionally, the resulting spectra are distinguishable. GABA target 3 has been chosen going forward for the robustness analysis, again due to the high overall fidelity achieved.

Figure 5.7 displays the best robust optimised pulse sequence for the GABA 3 target and the non-robust pulse used as the starting point for the optimisation. The non-robust pulse at the top of the figure is very calm, utilising only a small range of the possible amplitudes available. In contrast to the non-robust Glx pulse, this could be an indication of how this is overall an easier target, due to the very different structure of the molecules involved. The robust pulse at the bottom of the figure is much more aggressive, with fast switches to higher amplitudes. However, this still reaches lower energy limits than the Glx pulse, further hinting that the GABA target might be easier overall.

The robustness optimisation for the GABA pulse has also been successful, with a decrease in the fidelity of 1.5×10^{-3} that has a minimal impact on the resulting spectra as shown in Figure 5.9. There are additional features that can be seen from GABA that overlap with the Cr peak at 3 ppm, but these are incredibly small and would be typically be lost below the noise floor in experimental conditions.

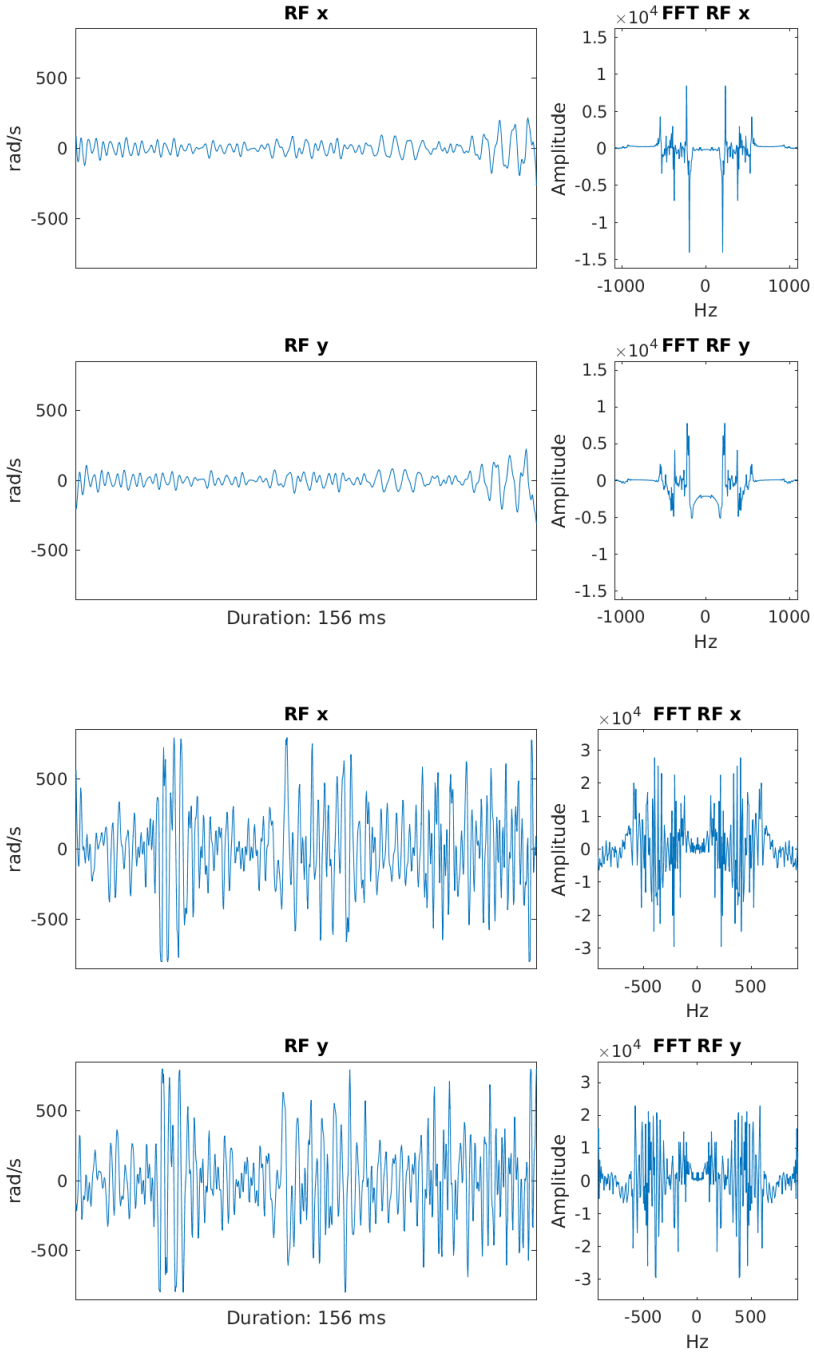


Figure 5.7: Best pulse sequence for the GABA 3 target, non-robust pulse top, robust-pulse bottom. Duration:156 ms, $\Delta t = 1.95 \times 10^{-4}$ms, 800 time steps.

Table 5.3: GABA target vectors and corresponding resonances. Peaks that fall within ±0.1 ppm range are grouped into the same target.

	GABA Target		
	1	2	3
State			
Cr	$\|00111\rangle$	$\|11000\rangle$	$\|11000\rangle$
NAA	$\|0111000\rangle$	$\|0111000\rangle$	$\|00001000\rangle$
GABA	$\|110000\rangle$	$\|000011\rangle$	$\|001100\rangle$
Resonances (ppm)			
Cr	3.91	3.02	3.02
NAA	2.00	2.00	4.38
GABA	3.01	2.28	1.88

Table 5.4: Performance of GABA targets. "# Runs" refers to the number of completed optimisation runs.

		Best overall		**Best under** 100 ms	
GABA Target	# Runs	Error	Duration	Error	Duration
Non-robust					
1	437	7.4×10^{-6}	198 ms	7.3×10^{-4}	99 ms
2	491	1.8×10^{-5}	195 ms	1.4×10^{-3}	96 ms
3	438	7.3×10^{-6}	193 ms	1.4×10^{-4}	98 ms
Robust					
1	161	8.3×10^{-3}	136 ms	4.0×10^{-2}	99 ms
2	201	1.4×10^{-2}	172 ms	1.4×10^{-2}	91 ms
3	166	7.9×10^{-3}	156 ms	2.3×10^{-2}	98 ms

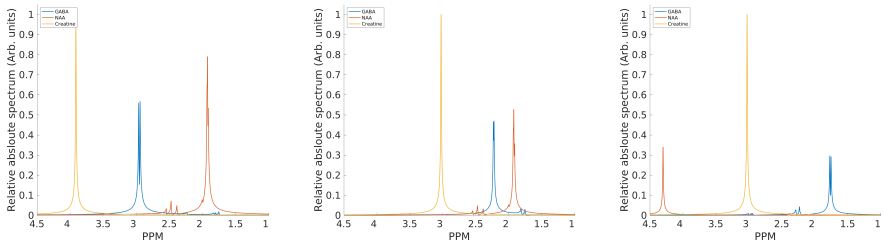


Figure 5.8: Best spectra found for the three GABA targets in the non-robust case. From left to right: Target GABA 1, GABA 2, GABA 3.

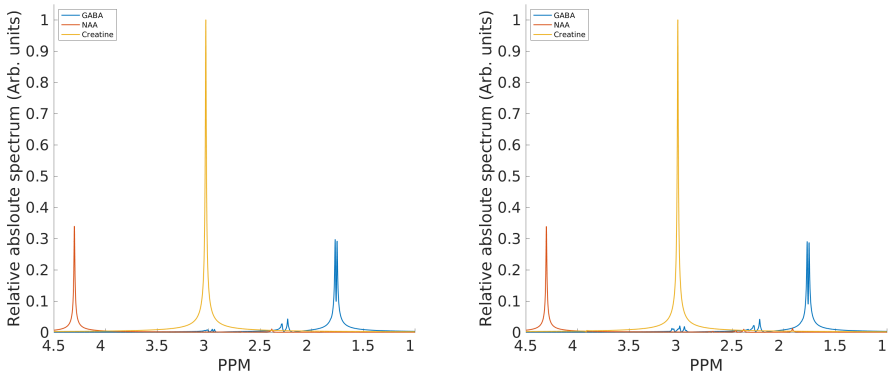


Figure 5.9: GABA 3 target spectra for the best non-robust pulse (left) and best robust pulse (right).

5.4 EVALUATION

In the following section, we review the effectiveness of the best pulses over a range of instabilities, metabolite models, and dissipation, the optimisation landscape and the importance of individual J-couplings.

5.4.1 *Optimisation Landscape*

Pulse performance is assessed on three key indicators: RF duration, RF Δt and 'aggression'. Aggression is a rudimentary measure of how quickly a pulse needs to switch large amplitudes. We define the *aggression metric* as:

$$A = \frac{1}{M(K-1)} \sum_{c=1}^{M} \sum_{t=1}^{K-1} |P(c,t) - P(c,t+1)|, \tag{5.6}$$

where P are the the pulse channels, in this case $P = [u_x, u_y]$ with M being the number of pulse channels. This metric aims to produce a normalised value across all pulses to determine how fast and hard the pulse switches between time steps. The assumption is that more "aggressive" pulses will be harder to implement, pushing the scanner hardware close to operating limits, increasing the amount of distortion.

Figure 5.10 displays all of the previously mentioned properties for both the non-robust Glx 1 and GABA 3 targets. There are a few surprising characteristics of this problem, firstly the optimisation is well behaved with regards to the error vs. duration graph found on the right-hand side. The gradual slope, with minimal outliers, suggests that these targets are relatively easy to achieve but we have hit a quantum speed limit, visible as the Pareto front. The speed limit is due to how fast we can force

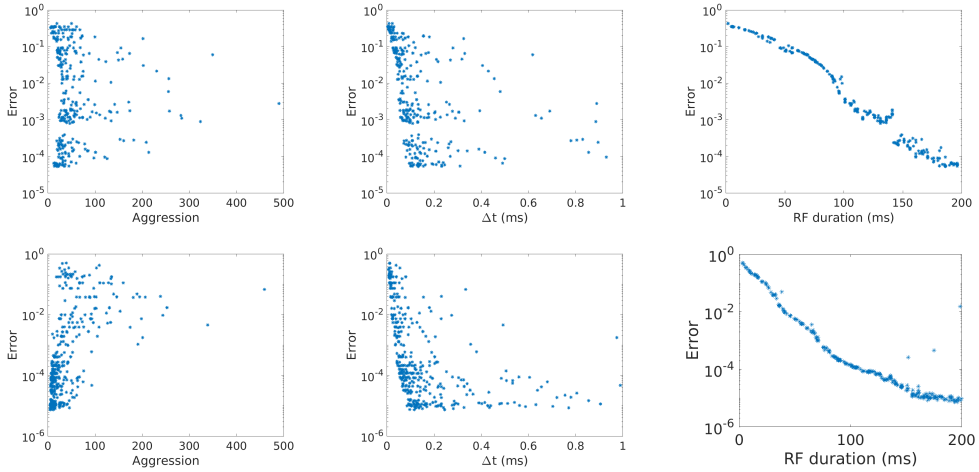


Figure 5.10: Optimisation landscape of Glx target 1 (top) and GABA target 3 (bottom). All graphs plot log error on the y axis, from left to right x axis values: Aggression vs Error, time step vs Error, and Duration vs Error.

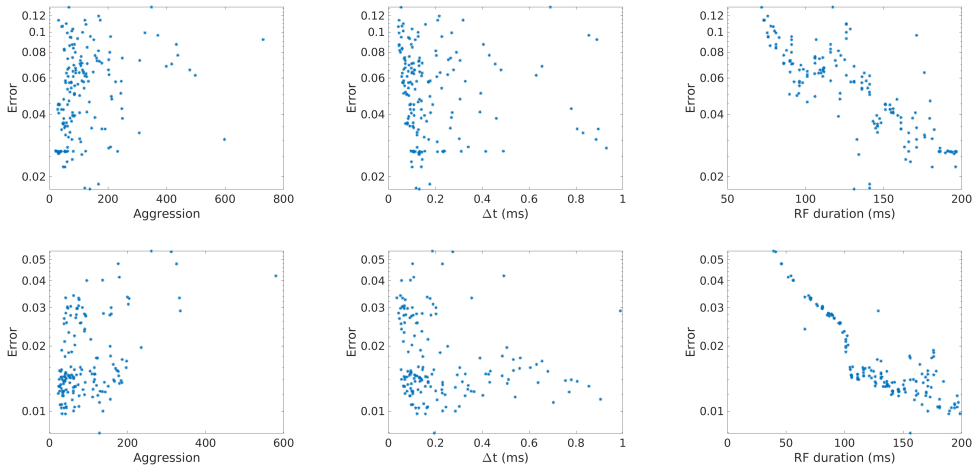


Figure 5.11: Optimisation landscape of robust Glx target 1 (top) and GABA target 3 (bottom). All graphs plot log error on the y axis, from left to right x axis values: Aggression vs Error, time step vs Error, and Duration vs Error.

Table 5.5: Mean runtime metrics of each of the targets, runtime shown is the mean number of hours.

	Mean Runtime	
Target	Non-Robust	Robust
Glx		
1	1.18σ0.86	13.02σ9.96
2	1.15σ0.86	12.10σ8.59
3	1.19σ0.87	11.99σ9.59
4	1.26σ0.89	12.37σ9.13
GABA		
1	6.45σ4.51	56.49σ48.25
2	6.78σ4.49	64.88σ48.55
3	6.43σ4.28	54.68σ41.00

these metabolites into the desired states, which is constrained by the minimum time step duration of the pulse and the energy limits imposed in Section 5.2.

Secondly, when assessing 'aggression', both the GABA and Glx targets favour smooth pulses. This is a good indicator that these will be realisable on physical hardware. Finally, there appears to be a negative effect of having too fine-grained control over the size of the time step, as error increases as this decreases. This could be because as the time step decreases, the gradient does too; effectively inhibiting the optimiser from successfully finding minima.

Figure 5.11 shows the same selection of metrics for the robust optimisation of the Glx 1 and GABA 3 targets. The results, in general, are more scattered, but the general trends from before continue. The error vs. duration is relatively well behaved, pulses favour low aggression and too fine a time step leads to an increase in error. The results are in general noisier, suggesting that the problem is harder and it is more difficult to find a similar Pareto front, to do so would require more optimisation runs with more variation in the initial conditions.

Table 5.5 shows the average runtime and error statistics from the optimisations. The simulation and optimisation was run in MATLAB (2018b) independently across 42 machines (Intel i7-4790 CPU, 16GB RAM). Machines pick optimisation jobs from a custom in-house scheduler at ran-

Table 5.6: Mean error metrics of each of the targets with standard deviation.

	Mean Error	
Target	Non-Robust	Robust
Glx		
1	$5.93\times10^{-2}\sigma9.93\times10^{-2}$	$5.78\times10^{-2}\sigma2.90\times10^{-2}$
2	$7.84\times10^{-2}\sigma9.25\times10^{-2}$	$5.85\times10^{-2}\sigma1.94\times10^{-2}$
3	$6.68\times10^{-2}\sigma1.03\times10^{-1}$	$7.34\times10^{-2}\sigma3.22\times10^{-2}$
4	$6.12\times10^{-2}\sigma7.63\times10^{-2}$	$5.35\times10^{-2}\sigma2.41\times10^{-2}$
GABA		
1	$3.39\times10^{-2}\sigma8.70\times10^{-2}$	$3.26\times10^{-2}\sigma1.58\times10^{-2}$
2	$4.32\times10^{-2}\sigma9.93\times10^{-2}$	$2.70\times10^{-2}\sigma1.24\times10^{-2}$
3	$2.52\times10^{-2}\sigma6.93\times10^{-2}$	$2.01\times10^{-2}\sigma1.40\times10^{-2}$

dom. There are no parallelism techniques employed as running sequential code was found to be faster experimentally.

In general, the Glx targets are substantially faster to compute, this is due to the reduced number of metabolites (2 compared to Glxs 3), and fewer spins per metabolite ($[2\times5]$ vs $[2\times6, 1\times7]$ for GABA).

As the density matrix of each system is 2^N, where N is the number of spins. Scaling of the problem between the non-robust and robust optimisation is relatively linear. The robust optimisation simulates 21 times the number of systems for each iteration but is run for 500 total iterations, rather than 1000, the average timings for the robust are around 13.5 times the non-robust times. This is a problem that would lend itself well to heavy parallelisation, however, due to time and memory constraints, this was not implemented in this instance.

5.4.2 B_0 & B_1 *Robustness*

To assess the performance achieved through our B_0 B_1 optimisation scheme outlined in Section 5.2, pulse sequences are simulated over a regular grid of B_0 and B_1 offsets. The range of these grids extends far beyond acceptable limits for spectroscopy (B_0 ±50 Hz and B_1 $\pm25\%$), but are useful for highlighting the expected performance over the range. The values in these graphs are the mean error for the simulated systems at the given offset, anything over 0.1 is considered to be unacceptable, as this

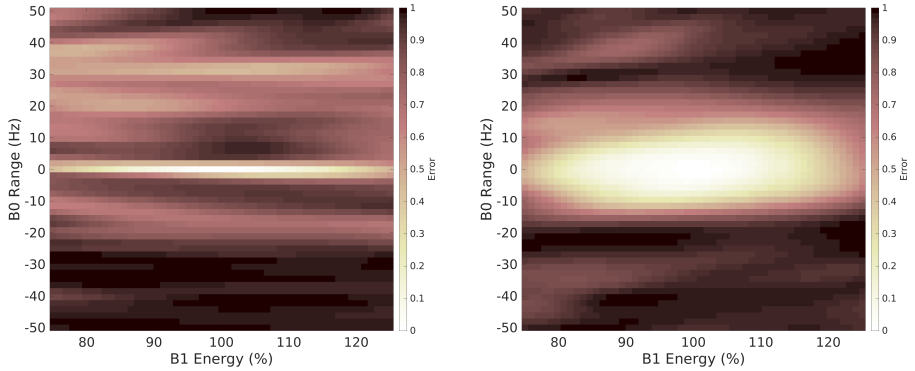


Figure 5.12: B_0, B_1 landscape for Glx target 1. Left, simple pulse sequence optimised at resonance, right pulse sequence optimised over $B_0 \pm 5\,\text{Hz}$ and $B_1 \pm 5\%$ ranges.

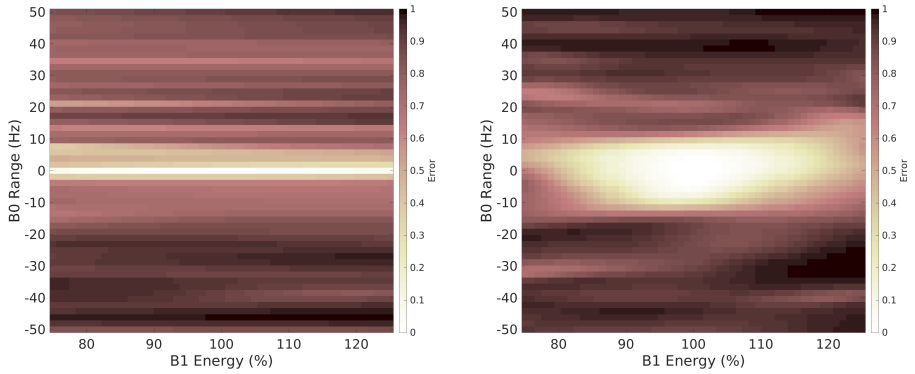


Figure 5.13: B_0, B_1 landscape of target GABA target 3. Simple target on the left, robust target on the right

is typically where the spectra begin to visually deviate from the target substantially.

Figure 5.12 shows the overall performance impact that can be achieved by optimising over a range of instabilities for Glx. Surprisingly as shown on the left, a large variance in total B_1 energy has a limited impact on performance for pulses that are not optimised for it. However, these same non-robust are very unstable with respect to any B_0 drift. The robust optimisation changes this landscape considerably, creating a much larger acceptable region of B_0 instability while maintaining much of the original B_1 range. From this analysis, it would appear that a $\pm 5\,\text{Hz}$ B_0 and up-to a $\pm 15\%$ B_1 range would be acceptable in an experimental setting.

Similar trends continue for GABA target 3 as shown in Figure 5.13. The non-robust pulse produces a solution that is relatively stable for B_1 instabilities but falls over quickly when B_0 varies. The robust pulse has a much larger effective area over the B_0 and B_1 range, where it appears

that a $\pm 12.5\%$ B_1 and $\pm 7\%$ B_0 would work in practice. However, the B_0 range is not symmetric around the resonance frequency but shifted to favour a positive B_0 offset by around 2 Hz.

5.4.3 *Controls and the Hyperfine Structure*

Evaluating the importance of the J-couplings for the controls is a complex task, where a full investigation would requite a continuous evaluation across all couplings at each time step of the system evolution. To simplify this, we have chosen to evaluate the contribution of each coupling removing them from the structure of the molecule, re-simulating the control and evaluating the impact on performance. Although this method is not exhaustive, and will not show if there is a chain of couplings that are used in unison, it will highlight if there are couplings that are incredibly important for a particular pulse sequence and target.

Glx

Figure 5.14 displays the error difference found when removing each of the J-couplings for Glutamine and Glutamate, for both the robust and non-robust pulses.

The J-couplings play a large role in the non-robust pulse, for Glutamate, the pulse sequence is heavily dependant on the majority of the couplings, suggesting that the optimised pulse is utilising the dynamics of the system to achieve the desired target state. For Glutamine, the effect is similar but predominantly utilises the couplings between the protons linking the spins in the 2.1 ppm to 2.4 ppm range.

For the robust pulse, the effect of removing the J-couplings is much less substantial but the removal of the majority of the couplings would lead to an increase in error and spectra that would be unsuitable. Glutamate is dependant on multiple couplings to remain successful, whereas Glutamine is only heavily dependant on one coupling between 2.432 ppm and 2.129 ppm to achieve adequate control. This suggests that these controls are not simply frequency selective, but are utilising the hyperfine structure in both the non-robust and robust case.

GABA

Figure 5.15 shows the effect of the removal of the J-couplings and the resulting performance of the pulse for the GABA 1 target. In contrast to the Glx J-couplings, the dynamics of the systems are very different. In

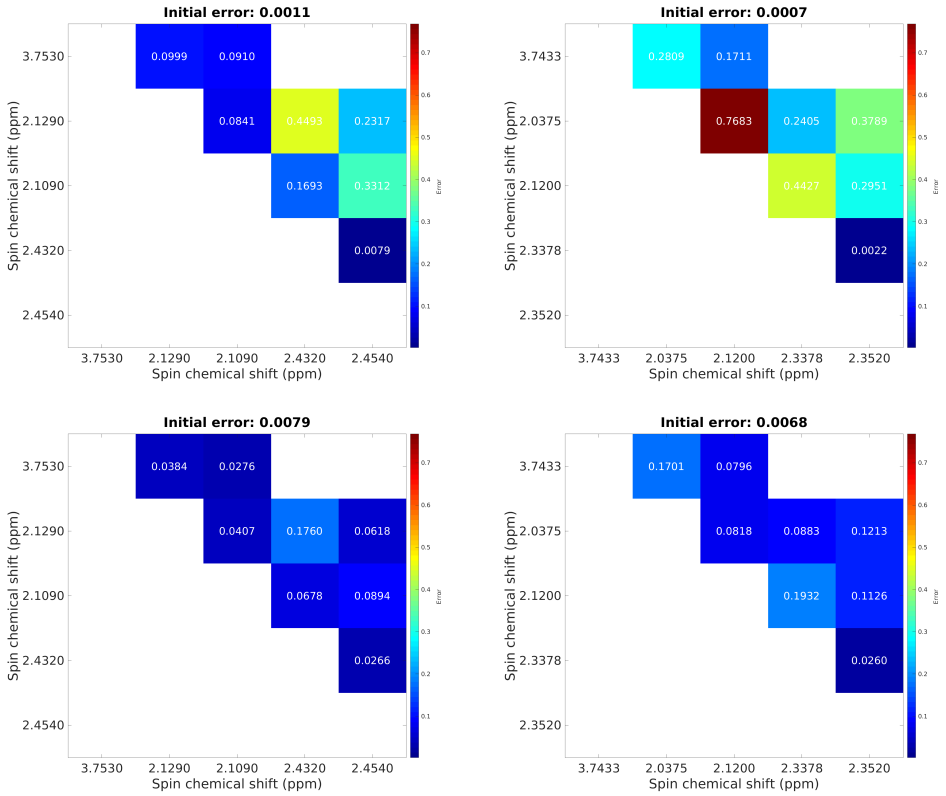


Figure 5.14: Effect of removing J-couplings for the best Glx 1 pulse. Glutamate left, Glutamine Right, non-robust pulse top and robust pulse bottom.

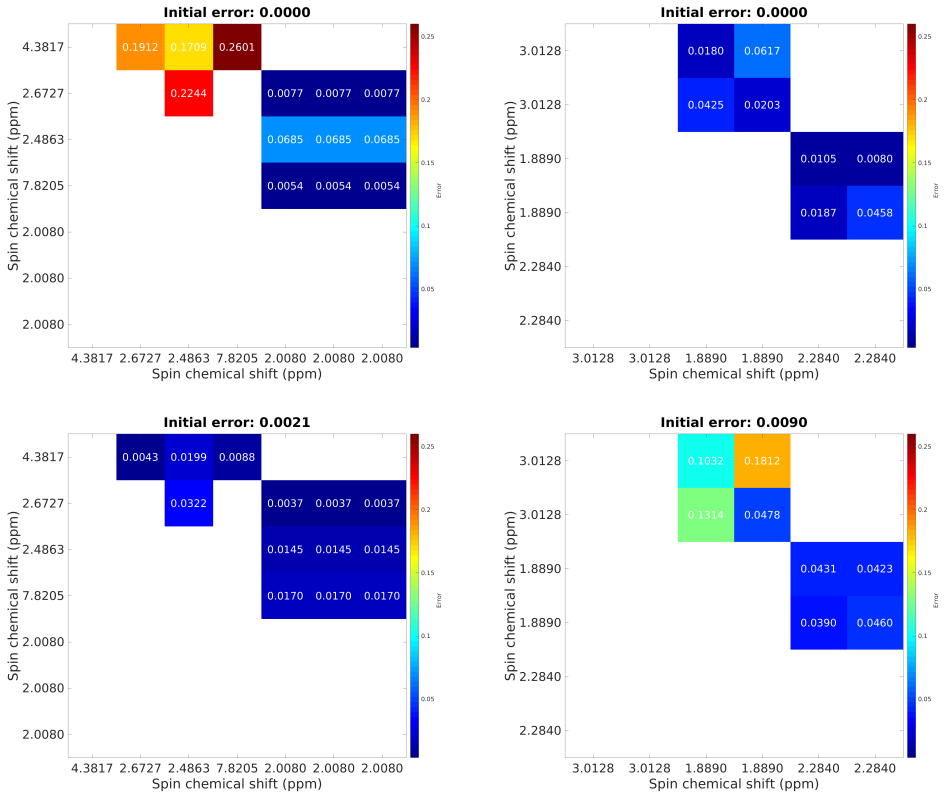
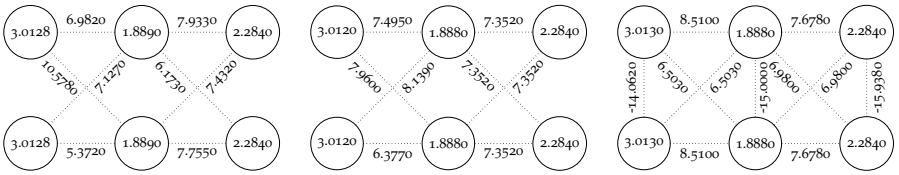


Figure 5.15: Effect of removing J-couplings for the best GABA 3 pulse. GABA left, NAA Right, non-robust pulse top and robust pulse bottom. Creatine has been omitted as it has no couplings.



(a) Govindaraju *et al.* [45] (b) Kaiser *et al.* [60] (c) Near *et al.* [103]

Figure 5.16: Different GABA models, nodes have chemical shift values in ppm, edges are J-coupling strength in Hz. Note the additional couplings in the Near model.

the non-robust case, the error difference for removing any of the GABA couplings is not all that substantial, which is surprising as it is a tightly coupled molecule. For NAA, the pulse is heavily dependant on the four couplings in the upper left, where the pulse surprisingly utilises the coupling to the high-frequency 7.8 ppm spin.

In contrast, in the robust case, the GABA pulse is much more heavily dependent on the J-couplings and NAA would remain largely unaffected by the removal of a large majority.

This represents a large change in the way the pulse sequences are achieving their targets in the robust case, where NAA appears to be generally ignoring the hyperfine structure of the molecule and the GABA control component utilising it.

5.4.4 *Metabolite Models*

Metabolite models are a disputed area in MRS [70], with GABA alone having three models from Govindaraju, Kaiser *et al.* [60] and Near *et al.* [103], displayed as weighted graphs in Figure 5.16. As optimisation is performed over a ± 5 Hz offset, any minor variations in the resonance frequencies of the protons would be insignificant, as this target effectively allows these to drift up to ± 0.04 ppm. The major consideration when testing these metabolite models is the changes in the strength of the J-couplings and the structure itself.

All three of the GABA models have very similar spectra for the robust optimised GABA target. This is likely due to the reduced dependency on the J-coupling structure as displayed in Figure 5.15. Additionally, Table 5.7 shows the error differences between the metabolite models for all three targets and confirms the high performance of the different models. This result is surprising for the Near model, as it is substantially different from the others, with a total of 11 couplings compared to 8 used

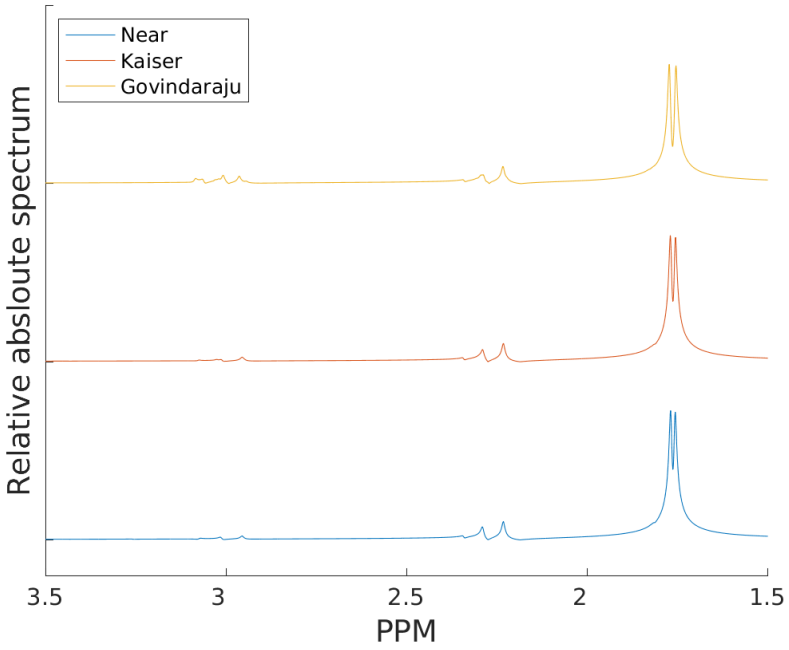


Figure 5.17: GABA 3 target robust pulse simulated with multiple GABA models. From top to bottom: Near *et al.* [103], Kaiser *et al.* [60] and Govindaraju *et al.* [45].

in the Govindaraju and Kaiser models. It also highlights the difficulty in defining accurate models, there seems to be a number of possible 'correct' solutions that produce similar spectra.

5.4.5 *Open system simulation*

So far, we have only discussed the spectra in for the open system simulation, for each of the best robust pulses, we have simulated the dissipative

Table 5.7: Individual metabolite error for different GABA models for the best GABA robust pulses. The Govindaraju *et al.* [45] model is used for the optimisation.

	Model source		
GABA Target	Govindaraju	Kaiser	Near
1	8.673×10^{-3}	3.715×10^{-3}	3.658×10^{-3}
2	1.435×10^{-2}	3.530×10^{-3}	3.743×10^{-3}
3	8.999×10^{-3}	3.668×10^{-3}	4.478×10^{-3}

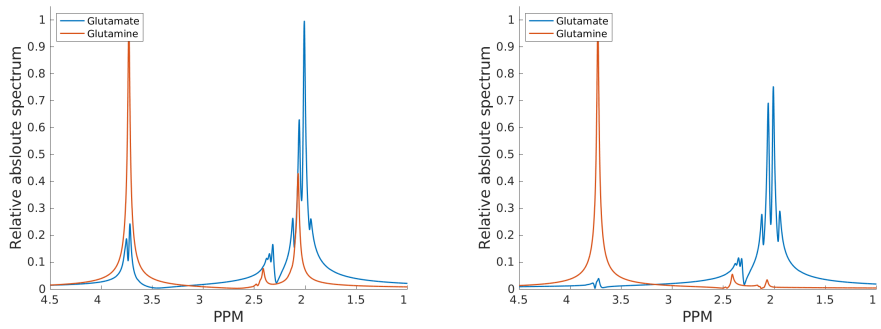


Figure 5.18: Open system simulation of two robust pulses for the Glx 1 target, best pulse under 100 ms on the left, best overall pulse on the right

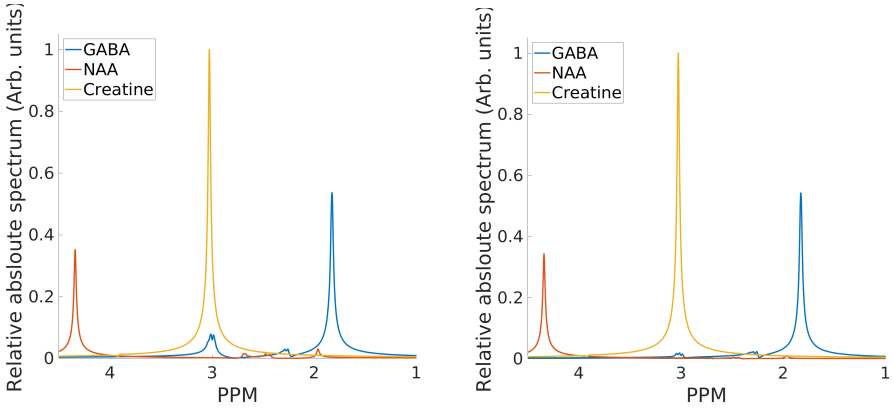


Figure 5.19: Open system simulation of two robust pulses for the GABA 3 target, best pulse under 100 ms on the left, best overall pulse on the right (156 ms).

spectra. The values used in this paper are based on the average relaxation times in white matter at 3 T of $T_1 = 69$ ms and $T_2 = 1084$ ms [145].

Figure 5.18 shows the effect of a complete dissipative simulation of the robust pulses. The higher fidelity, longer robust pulses fares much better, maintaining the target spectra through dissipation. In contrast, the shorter pulse falls apart, with elements of each metabolites overlapping substantially. This result suggests that longer pulses will be necessary to achieve the molecule states that will produce the spectra in an experimental setting.

Similarly, Figure 5.19 shows the effect of a complete dissipative simulation of the GABA 3 robust pulse. Following a similar storyline to before, the more accurate, longer control has more stability when robustness is considered. This result suggests that longer pulses will be necessary to achieve the molecule states that will produce the spectra in an experimental setting.

5.5 EVALUATION AND DISCUSSION

Overall, we have shown that optimisation of MR pulse sequences with quantum control techniques can create extremely desirable characteristics in the resulting spectrum. The optimised RF pulses for all seven targets produce distinct spectra for each of the metabolites. Additionally, it is possible to further optimise the non-robust pulses for each of the targets to be robust to a range of control and system instabilities. Furthermore, all of these pulse sequences have been found with a realistic constraint on the maximum energy and duration of each pulse.

Optimisations were only done to a subset of metabolites, there is no indication of how well this method will work with a larger set of metabolites. Additionally, no simulations were done to see how the pulse sequences distort the spectra of other metabolites. This would be key to accurate quantification, as the current state of the art methods utilise inverse problem solving, fitting known signals to the input spectrum. This would require the generation of a new basis set for quantification with each pulse sequence, which would be straight forward. Hence existing tools could be used with minor modification.

Optimisation of larger ranges of inhomogeneity would require more complex targets, as they would require more spin resonances to be absorbed into the target, as when the B_0 range increases, protons are no longer individually addressable. This would severely limit the number of possible targets for more metabolites. We have proposed a range (± 7 Hz) that is reasonable for spectroscopy with current hardware. If over time the calibration accuracy of scanner hardware w.r.t. B_0 should increase, the B_0 range used for optimisation could decrease.

Current MRS pulse sequences utilise a spacial localisation scheme, we have provided no such method. In addition, the use of refocusing pulses is not appropriate for this application as it destroys the target states. One possible way around this is that it may be possible to optimise the final RF pulse in a PRESS or STEAM pulse sequence to perform the selective excitation. Other solutions would be to perform outer volume suppression and/or 3D phase encoding.

There is a hard limit on the size of metabolites that would be feasible to simulate due to the computational and memory requirements scaling exponentially with the number of spins. It is possible to split the metabolite graphs into sub-graphs if there is no interaction between the spins, such as splitting Cr into 6 individual spin systems, as there are no J-couplings. However, this method requires careful consideration of

how operators interact with these systems. Hogben *et al.* [54] have implemented a method in their Matlab based simulation library 'Spinach', that enables simulation of systems with 40+ spins. This is achieved by removing long-range couplings that have a negligible effect and reducing the state space of the system, under reasonable assumptions on the limits of MR based experiments. This, in theory, could be used to great success, however as we are performing a non-typical MR experiment, removing states that are used by the optimal controls would be detrimental to their ability to perform. This would have to be carefully checked and validated before restricting the range of system states that could be used as control paths for the optimised RF pulses.

Robust optimisation is currently very slow. This problem would lend itself to massively parallel computing extremely well, bringing down the overall runtime substantially by simulating each 'set' of metabolites on a separate machine. This could bring the time for the robust optimisation in line with the non-robust if there were enough free machines to satisfy one metabolite set for each system.

There are many directions for future work, most importantly would be experimental validation of this method. Once this has been achieved, a larger range of metabolites and targets should be tested, including a review of the effect of the pulses on the large range of metabolites found *in-vivo*. Following this, exploring a range of scanner strengths and spatial localisation schemes to move the pulse sequence further towards a generally available method. Once this technique has been experimentally validated, it would then be imperative to benchmark the performance of quantification when compared with other pulse sequences. Finally, much more could be done on the computational side by improving the overall performance by utilising parallel computing methods. A greater speed-up could be achieved through the use of GPUs to calculate the system states, as long as the metabolite models can fit into GPU memory.

5.6 SUMMARY

This chapter has displayed that it is theoretically possible to optimise RF pulses for specific metabolite states concurrently, enabling a level of customisation to the resulting spectra that has previously not been seen. Also, high-quality RF pulses can found that are robust over a realistic range of B_0 and B_1 instabilities and within energy and time restrictions.

This chapter has also presented a method of target state selection that places spins of interest into the transverse plane and leaves other spins in alignment with the static field. This allows for the spectra to only be composed of the spins of interest, however, this target state selection is arbitrary. In theory, any target state could be selected and optimised for, such as only exciting one metabolite out of a handful of brain metabolites. This is likely not realisable given the energy and time constraints, but it is important to understand the flexibility of this method.

Overall, this chapter has shown that quantum control techniques can successfully be applied in simulations to the unstable MR environment. The optimisation can be leveraged to transfer energy to desired states to produce predetermined spectra that have desirable characteristics. The resulting spectrum provides the ability to uniquely resolve metabolites that have significantly overlapping spectral features with popular pulse sequences, such as PRESS or STEAM. We have also shown that by integrating a large range of realistic and expected instabilities with regards to system parameters and control, we can produce robust pulse sequences to mitigate the impact of this and making them theoretically implementable.

6

EPILOGUE

The theme of the research presented in this thesis is to further develop MRS methods for quantification and RF pulse design. We summarise the main findings in Section 6.1, discuss future work in Section 6.2 and conclude with Section 6.3.

6.1 THESIS SUMMARY

The four contributions of this work laid out in Section 1.3 are covered by Chapter 3 (C1), Chapter 4 (C2), and Chapter 5 (C3,C4) respectively and are summarised below.

6.1.1 *Benchmarking Quantification*

We started with benchmarking current quantification methods, presented in Chapter 3. The focus of this chapter is on GABA quantification with MEGA-PRESS.

A range of experimental phantoms were collected by collaborators at Swansea University and the spectra analysed with state-of-the-art quantification methods (LCModel, GANNET, TARQUIN and jMRUI). It was overall found that LCModel was the best performing tool, but the general accuracy was poor with the GABA-to-NAA ratio by underestimated by $38\% - 55\%$ for the pH neutral phantoms. This is despite LCModel being considered to be the gold standard of quantification methods.

This work has highlighted the need for more rigorous benchmarking of quantification methods. However, this itself is a complex task due to the difficulty of acquiring spectra with the associated ground truth, and the complexity of the quantification software adding not only a significant time requirement, but also the potential for bias due to different users. In addition, when benchmarking basis set methods, additional time must be spent simulating or sourcing basis sets, which ultimately have a large impact on the quality of the fit. In an ideal world, there would be a transferable basis set file format to enable a like-for-like comparison between quantification methods, but currently this simply is not the case.

This leads to difficulty evaluating the fitting method in isolation, as a different basis set end up being used for different fitting methods and the two aspects cannot be easily separated.

This work has lead to a range of recommendations for the future of benchmarking quantification. Attempts should be made to create a benchmarking standard, to empower researchers to test, develop and improve on quantification methods. In order to do this, additional phantoms should be made and scanned at more locations with a larger range of pulse sequences and their T_E timing variants. In addition, efforts should be made to unify a basis set format to enable saving and loading between quantification methods, as the simulation of basis sets should not be coupled to the process of quantification.

6.1.2 *MRSNet*

The second objective was to investigate the use of machine learning for MRS quantification as discussed in Chapter 4. While research into machine learning application to MRS quantification is not new [26, 78], this work is the first exploration of using machine learning for edited spectroscopy specifically with MEGA-PRESS for the quantification of GABA.

This chapter explores a large range of possible ways to represent the spectra to a neural network, as, unlike some other pulse sequences, MEGA-PRESS has two acquisitions (edit off, edit on) and their difference. This creates a unique opportunity to study the effect of varying combinations of the acquisitions as training data to the network, combined with altering the data representation (real, imaginary and/or absolute values). Furthermore, a range of neural network architecture variants and parameters were explored to find the best performing network.

Work presented in this chapter has shown that a CNN approach has comparable performance with LCModel, which is the best performing quantification method from Chapter 3. Specifically, we found that using the absolute value edit off and difference spectra provides the best performance. On the specific experimental dataset we were even able to demonstrate improved performance compared to LCModel, but the limited nature of the study, as already explained, means that this may not hold in general. Nevertheless, the approach is highly promising and, in particular, it is quite simple to adapt this to a range of spectra acquired with pulse sequences similar to MEGA-PRESS, assuming suitable train-

ing data can be simulated and of course it would have to be validated again.

6.1.3 *Pulse Sequence Optimisation*

Finally, we devised of a novel method of RF pulse optimisation towards chemically selective MRS in Chapter 5.

Excitation of spins in MRS is typically done with broadband excitation pulses, exciting all metabolites in the ROI to generate the spectra. While in many cases this is advantageous, for the detection of low-signal metabolites this is a major issue, especially where spectra are obscured by more prevalent metabolites, such as GABA. While there are spectral editing sequences, such as MEGA-PRESS [92] and HERMES [123] to address this issue, they require careful consideration, especially due to the multiple acquisitions involved.

A method was developed where individual spins per metabolite can be chosen for excitation. The selected spins are translated into target states for a group of metabolites, with two groups used as a case study: [Glu & Gln] and [GABA, NAA & Cr]. These targets are then set via a target functional for an optimisation method over the RF pulse. It is quite complex to decide whether such a pulse exists in theory, so instead high-fidelity pulses are sought numerically. If a RF pulse realising the targets to high fidelity has been found, it enables customising the resulting spectra from the selected metabolites, removing unwanted spectral features that overlap with signals of interest. This significantly simplifies spectral analysis and quantification.

In addition, the presented optimisation approach allows us to consider a range of scanner uncertainties to also optimise for the robustness of the pulses. This integration of scanner instabilities into the optimisation is imperative for these pulses to be experimentally validated, as without they are extremely sensitive to changes to the B_0 and B_1 fields. The method presented showed that it is possible to optimise pulses for the aforementioned metabolite sets over a $\pm 5\,\mathrm{Hz}$ B_0 range and a B_1 range of $\pm 5\%$ of total pulse energy.

The work has shown that it is theoretically possible to create optimised RF pulses that enable chemically selective excitation at 3 T. In addition, these optimised pulses are realistic with respect to scanner constraints and a range of B_0 and B_1 instabilities, as shown by simulations.

6.2 FUTURE WORK

Contributions made towards this thesis can be used as a foundation to a range of future work to enhance MRS.

6.2.1 *Benchmarking Quantification*

Firstly, one of the main motivations for Chapter 3 was a recognition of the lack of open experimental datasets with ground truth data. This is something that is commonly done in MRI and machine learning communities, and should be extended to include MRS.

In order to facilitate this, an open source standard for the storage of basis set data should be created, as there is currently no universally accepted format between quantification software. LCModels proprietary file format is the closest to being universally adopted, but it is poorly documented and as such it is difficult to develop robust methods to save and load data in this format. A mark up language, with a high level of compatibility across programming languages, such as YAML or JSON could be used.

A wider collaboration on the creation of experimental basis sets along with a unified resource to publicly upload and release them should be created, such as a website. By creating a unified resource, it would greatly simplify the process of sourcing these datasets. This website could act as a repository for spectra, include guidance in create and contribute their own phantom spectra, and allow users to upload their created spectra. The overall aim would be to expand the dataset, adding additional pulse sequences, scanners and magnet strengths and to make this publicly available. In addition, this website could automate the process of benchmarking and optimising quantification methods and offer recommendations for the currently best performing settings.

6.2.2 *MRSNet*

From the work completed in Chapter 4, there are a range of directions for future work. The study in this thesis focused on the frequency domain for a single pulse sequence at one T_E, T_R timing for five metabolites. Different representations of the input data could be investigated, such as using the complex time-domain signal, multiple T_E acquisitions, multiple short-time Fourier transforms or reduced frequency domain data such as peak

locations, amplitudes and phases. Networks could be trained and tested for a larger number of metabolites and a range of pulse sequences and timings to see how well they can learn to generalise. Importantly, such work must continue to be linked to experimental work with calibrated phantoms to try to ensure accuracy and reliability *in-vivo*.

In addition, if the work for a centralised repository of benchmark data was created, this could greatly benefit the development of quantification methods. If enough data was available, it would be trivial to augment experimental datasets for training and validation of new neural networks.

Of course the MRSNet approach could also be adapted to the spectra acquired with the optimised RF pulses. For this to be studied further, however, these would first have to be realised experimentally to better understand any practical issues that may be involved. Given that these spectra are expected to be much simpler, a simpler fitting or machine learning approach, such as support vector machines or random forests, may be sufficient.

6.2.3 *Pulse Sequence Optimisation*

Finally, Chapter 5 has shown that it is in theory possible to create pulses for specific metabolite excitation states. The next logical step from this work is to experimentally realise these pulses and to test their effectiveness. This should be closely followed by comparing the quantification performance with current pulse sequences, by generating basis sets to be used by current state-of-the-art methods. Experimental realisation, however, is not a trivial task, as it requires careful calibration and characterisation to ensure the custom RF pulses are realised with sufficient high accuracy.

Once the method has been validated, there is a large range of potential research avenues to pursue. Firstly, the method presented only looks at a handful of metabolites. As this is not the case *in-vivo*, optimising pulse sequences to be effective on a larger range of metabolites would be interesting. By doing this, it would begin to highlight the limits of this method where controlling a large range of metabolites concurrently is no longer possible. In order for this to be achieved, a range of performance improvements would be necessary to maximise the number of metabolites that can be taken into account. This could be achieved by utilising GPUs for calculation of metabolite states and utilising methods employed in Spinach [54] to enable the simulation of larger spin systems.

Moreover, there is no spatial localisation scheme provided for the RF pulses, and development of a method is needed, after the approach has been in principle validated experimentally. In development of a spatial localisation scheme, the optimisation routine could be extended to additionally control the gradients for this to be achieved. Alternatively, existing spatial localisation schemes may be adapted, but care has to be taken that these do not counter the effects of the custom RF pulse.

6.3 FINAL REMARKS

This thesis has covered a range of MRS topics in quantification and optimisation of pulse sequences. Work that has been completed for this thesis has shown improvements over current methods in quantification and metabolite separation with pulse sequences. While the work does not represent a complete solution that is ready for implementation in a clinical environment, it is a significant step forward on the foundations to lead towards this broader goal.

BIBLIOGRAPHY

[1] Martín Abadi et al. 'TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems'. In: *Netw. Comput. Neural Syst.* 16.2-3 (2016), pp. 121–138. ISSN: 0954898X. DOI: 10.1080/09548980500300507. arXiv: 1603.04467. URL: http://arxiv.org/abs/1603.04467.

[2] Abhishek Aggarwal et al. 'Role of multivoxel intermediate TE 2D CSI MR spectroscopy and 2D echoplanar diffusion imaging in grading of primary glial brain tumours'. In: *J. Clin. Diagnostic Res.* 11.6 (2017), TC05–TC08. ISSN: 0973709X. DOI: 10.7860/JCDR/2017/24982.9984.

[3] Li An et al. 'Detection of glutamate, glutamine, and glutathione by radiofrequency suppression and echo time optimization at 7 Tesla'. In: *Magn. Reson. Med.* 73.2 (2015), pp. 451–458. ISSN: 15222594. DOI: 10.1002/mrm.25150.

[4] E Raymond Andrew. 'Nuclear Magnetic'. In: *NMR In Physiology and Biomedicine* (2013), p. 1.

[5] H Barkhuijsen, R de Beer and D van Ormondt. 'Improved algorithm for noniterative time-domain model fitting to exponentially damped magnetic resonance signals'. In: *Journal of Magnetic Resonance (1969)* 73.3 (1987), pp. 553–557. ISSN: 0022-2364. DOI: https://doi.org/10.1016/0022-2364(87)90023-0. URL: http://www.sciencedirect.com/science/article/pii/0022236487900230.

[6] John H F Bothwell et al. 'An introduction to biological nuclear magnetic resonance spectroscopy'. In: 44 (2011), pp. 493–510. DOI: 10.1111/j.1469-185X.2010.00157.x.

[7] Paul A. Bottomley. 'Spatial Localization in NMR Spectroscopy in Vivo'. In: *Ann. N. Y. Acad. Sci.* 508.1 Physiological (Nov. 1987), pp. 333–348. ISSN: 0077-8923. DOI: 10.1111/j.1749-6632.1987.tb32915.x. URL: http://doi.wiley.com/10.1111/j.1749-6632.1987.tb32915.x.

[8] Hans de Brouwer. 'Evaluation of algorithms for automated phase correction of NMR spectra'. In: *J. Magn. Reson.* 201.2 (Dec. 2009), pp. 230–238. ISSN: 10907807. DOI: 10.1016/j.jmr.2009.09.017.

[9] Richard H. Byrd et al. 'A Limited Memory Algorithm for Bound Constrained Optimization'. In: *SIAM J. Sci. Comput.* 16.5 (Sept. 1995), pp. 1190–1208. ISSN: 1064-8275. DOI: 10.1137/0916069. arXiv: arXiv:1011.1669v3. URL: http://epubs.siam.org/doi/10.1137/0916069.

[10] Dominic Carlin et al. 'Short-acquisition-time JPRESS and its application to paediatric brain tumours'. In: *Magn. Reson. Mater. Physics, Biol. Med.* 32.2 (2019), pp. 247–258. ISSN: 13528661. DOI: 10.1007/s10334-018-0716-6. URL: https://doi.org/10.1007/s10334-018-0716-6.

[11] Joseph R. Casey, Sergio Grinstein and John Orlowski. 'Sensors and regulators of intracellular pH'. In: *Nat. Rev. Mol. Cell Biol.* 11.1 (2010), pp. 50–61. ISSN: 14710072. DOI: 10.1038/nrm2820.

[12] M. Chandler et al. 'MRSNet: Metabolite Quantification from Edited Magnetic Resonance Spectra With Convolutional Neural Networks'. In: Mc (Sept. 2019), pp. 1–12. arXiv: 1909.03836. URL: http://arxiv.org/abs/1909.03836.

[13] Max Chandler. *A scheduler for automated execution of code written in bash.* 2015. URL: https://github.com/MaxChandler/scheduler.

[14] Max Chandler and Frank Langbein. *MR spectral quantification using convolutional neural networks.* 2018.

[15] Max Chandler, Frank Langbein and Sophie Shermer. *A framework for optimising MRS RF Pulses for metabolite target states*. 2017. URL: https://qyber.black/MRIS/control.

[16] Max Chandler et al. 'Advanced detection and quantification of biomarkers in magnetic resonance spectroscopy'. In: *Proceedings of the EPSRC UK Image-Guided Therapies Network+ Launch*. 2016.

[17] Max Chandler et al. 'Quantum control for magnetic resonance spectroscopy'. In: *Proceedings of the All Wales Medical conference*. 2017.

[18] Max Chandler et al. 'Robust Quantum Optimal Control for Unique Metabolic Spectra in Magnetic Resonance Spectroscopy'. 2019.

[19] Paul Chang et al. 'Constrained optimization for position calibration of an NMR field camera'. In: *Magn. Reson. Med.* 80.1 (2018), pp. 380–390. ISSN: 15222594. DOI: 10.1002/mrm.27010.

[20] Yan Chang et al. 'Spin-Scenario: A flexible scripting environment for realistic MR simulations'. In: *J. Magn. Reson.* 301 (2019), pp. 1–9. ISSN: 10960856. DOI: 10.1016/j.jmr.2019.01.016. URL: https://doi.org/10.1016/j.jmr.2019.01.016.

[21] Lilian Childress and Ronald Hanson. 'Diamond NV centers for quantum computing and quantum networks'. In: *MRS Bull.* 38.2 (2013), pp. 134–138. ISSN: 08837694. DOI: 10.1557/mrs.2013.20.

[22] Francois and others Chollet. *Keras*. 2015. URL: https://keras.io.

[23] *CODATA Value: proton gyromagnetic ratio*. URL: https://physics.nist.gov/cgi-bin/cuu/Value?gammap (visited on 17/12/2019).

[24] Daniel Coman et al. 'Brain temperature and pH measured by 1 H chemical shift imaging of a thulium agent'. In: *NMR Biomed.* 22.2 (Feb. 2009), pp. 229–239. ISSN: 09523480. DOI: 10.1002/nbm.1312. URL: http://doi.wiley.com/10.1002/nbm.1312.

[25] M. I. Cordero et al. 'Effects of paternal and peripubertal stress on aggression, anxiety, and metabolic alterations in the lateral septum'. In: *Eur. Neuropsychopharmacol.* 26.2 (2016), pp. 357–367. ISSN: 18737862. DOI: 10.1016/j.euroneuro.2015.11.017. URL: http://dx.doi.org/10.1016/j.euroneuro.2015.11.017.

[26] Dhritiman Das et al. 'Quantification of Metabolites in Magnetic Resonance Spectroscopic Imaging Using Machine Learning'. In: *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. Vol. 10435 LNCS. 2017, pp. 462–470. ISBN: 9783319661780. DOI: 10.1007/978-3-319-66179-7_53. arXiv: arXiv:1805.10201v1. URL: http://link.springer.com/10.1007/978-3-319-66179-7%7B%5C_%7D53.

[27] F David et al. 'RF Coil Technology for Small-Animal MRI'. In: *NMR Biomed.* 20.3 (2007), pp. 304–325. DOI: 10.1002/nbm. URL: http://www.dotynmr.com/PDF/SA%7B%5C_%7DRFCoils2b.pdf.

[28] P. De Fouquieres et al. 'Second order gradient ascent pulse engineering'. In: *J. Magn. Reson.* 212.2 (2011), pp. 412–417. ISSN: 10907807. DOI: 10.1016/j.jmr.2011.07.023. arXiv: 1102.4096. URL: http://dx.doi.org/10.1016/j.jmr.2011.07.023.

[29] Stefano Delli Pizzi et al. 'GABA content within the ventromedial prefrontal cortex is related to trait anxiety'. In: *Soc. Cogn. Affect. Neurosci.* 11.5 (2016), pp. 758–766. ISSN: 17495024. DOI: 10.1093/scan/nsv155.

[30] Dr. Jim Murdoch. *LCModel Basis Sets*. URL: http://purcell.healthsciences.purdue.edu/mrslab/basis%7B%5C_%7Dsets.html (visited on 13/11/2018).

[31] R A E Edden et al. 'Gannet: a batch-processing tool for the quantitative analysis of GABA-edited MRS spectra'. In: *J Magn Reson Imag* 4.164 (2011), pp. 1445–1452. ISSN: 15378276.

[32] Luke J Edwards et al. 'Grid-free powder averages : on the applications of the Fokker-Planck equation to solid state NMR'. In: ().

[33] Bénédicte Elena, Gaël de Paëpe and Lyndon Emsley. 'Direct spectral optimisation of proton–proton homonuclear dipolar decoupling in solid-state NMR'. In: *Chem. Phys. Lett.* 398.4-6 (Nov. 2004), pp. 532–538. ISSN: 00092614. DOI: 10.1016/j.cplett.2004.09.122. URL: https://linkinghub.elsevier.com/retrieve/pii/S0009261404015295.

[34] ELSTER LLC. *Coil Overview*. URL: http://mriquestions.com/many-kinds-of-coils.html (visited on 18/10/2019).

[35] Kirk I. Erickson et al. 'Beyond vascularization: Aerobic fitness is associated with N-acetylaspartate and working memory'. In: *Brain Behav.* 2.1 (2012), pp. 32–41. ISSN: 21623279. DOI: 10.1002/brb3.30.

[36] Reza Fardanesh et al. 'Proton MR spectroscopy in the breast: Technical innovations and clinical applications'. In: *J. Magn. Reson. Imaging* (2019). ISSN: 15222586. DOI: 10.1002/jmri.26700.

[37] Graeme Finch, Ali Yilmaz and Marcel Utz. 'An optimised detector for in-situ high-resolution NMR in microfluidic devices'. In: *J. Magn. Reson.* 262 (2016), pp. 73–80. ISSN: 10960856. DOI: 10.1016/j.jmr.2015.11.011. URL: http://dx.doi.org/10.1016/j.jmr.2015.11.011.

[38] J. Frahm et al. 'Localized proton NMR spectroscopy in different regions of the human brain in vivo. Relaxation times and concentrations of cerebral metabolites'. In: *Magn. Reson. Med.* 11.1 (1989), pp. 47–63. ISSN: 15222594. DOI: 10.1002/mrm.1910110105.

[39] Fei Gao et al. 'A target field design of open multi-purpose RF coil for musculoskeletal MR imaging at 3 T'. In: *Magn. Reson. Imaging* 34.8 (2016), pp. 1064–1070. ISSN: 18735894. DOI: 10.1016/j.mri.2016.04.018. URL: http://dx.doi.org/10.1016/j.mri.2016.04.018.

[40] Roberto García Figueiras et al. 'Proton magnetic resonance spectroscopy in oncology: the fingerprints of cancer?' In: *Diagnostic Interv. Radiol.* 22.1 (Dec. 2015), pp. 75–89. ISSN: 13053825. DOI: 10.5152/dir.2015.15009. URL: http://www.dirjournal.org/sayilar/80/buyuk/75-89.pdf.

[41] Maria Isabel Osorio Garcia et al. 'Quantification of in vivo magnetic resonance spectroscopy signals with baseline and lineshape corrections'. In: *2010 IEEE Int. Conf. Imaging Syst. Tech. IST 2010 - Proc.* (2010), pp. 349–352. DOI: 10.1109/IST.2010.5548503.

[42] Alexandre Duarte Gigante et al. 'Brain glutamate levels measured by magnetic resonance spectroscopy in patients with bipolar disorder: A meta-analysis'. In: *Bipolar Disord.* 14.5 (2012), pp. 478–487. ISSN: 13985647. DOI: 10.1111/j.1399-5618.2012.01033.x.

[43] Steffen J. Glaser et al. 'Training Schrödinger's cat: Quantum optimal control: Strategic report on current status, visions and goals for research in Europe'. In: *Eur. Phys. J. D* 69.12 (2015), pp. 279–279. ISSN: 14346079. DOI: 10.1140/epjd/e2015-60464-1. arXiv: 1508.00442. URL: http://download.springer.com/static/pdf/212/art%7B%5C%%7D253A10.1140%7B%5C%%7D252Fepjd%7B%5C%%7D252Fe2015-60464-1.pdf?originUrl=http%7B%5C%%7D253A%7B%5C%%7D252F%7B%5C%%7D252Flink.springer.com%7B%5C%%7D252Farticle%7B%5C%%7D252F10.1140%7B%5C%%7D252Fepjd%7B%5C%%7D252Fe2015-60464-1%7B%5C&%7Dtoken2=exp=1490190658%7B~%7Dacl=%7B%5C%%7D252Fstatic%7B%5C%%7D252Fpdf%7B%5C%%7D252F212%7B%5C%%7D252Fart%7B%5C%%7D2525.

[44] Lidia Glodzik et al. 'Global N-acetylaspartate in normal subjects, mild cognitive impairment and Alzheimer's disease patients'. In: *J. Alzheimer's Dis.* 43.3 (2015), pp. 939–947. ISSN: 18758908. DOI: 10.3233/JAD-140609.

[45] Varanavasi Govindaraju, Karl Young and Andrew A. Maudsley. 'Proton NMR chemical shifts and coupling constants for brain metabolites'. In: *NMR Biomed.* 13.3 (May 2000), pp. 129–153. ISSN: 09523480. DOI: 10.1002/1099-1492(200005)13:3<129::AID-

NBM619>3.0.CO;2-V. URL: http://doi.wiley.com/10.1002/1099-1492(200005)13:3%7B%5C%%7D3C129::AID-NBM619%7B%5C%%7D3E3.0.CO;2-V.

[46] Marinette van der Graaf. 'In vivo magnetic resonance spectroscopy: basic methodology and clinical applications'. In: *Eur. Biophys. J.* 39.4 (Mar. 2010), pp. 527–540. ISSN: 0175-7571. DOI: 10.1007/s00249-009-0517-y. URL: http://link.springer.com/10.1007/s00249-009-0517-y.

[47] Danielle Graveron-Demilly. 'Quantification in magnetic resonance spectroscopy based on semi-parametric approaches'. In: *Magn. Reson. Mater. Physics, Biol. Med.* 27.2 (2014), pp. 113–130. ISSN: 09685243. DOI: 10.1007/s10334-013-0393-4.

[48] Bernhard Gruber et al. 'RF coils: A practical guide for nonphysicists'. In: *J. Magn. Reson. Imaging* 48.3 (2018), pp. 590–604. ISSN: 15222586. DOI: 10.1002/jmri.26187.

[49] A. Haase et al. '1 H NMR chemical shift selective (CHESS) imaging'. In: *Phys. Med. Biol.* 30.4 (Apr. 1985), pp. 341–344. ISSN: 0031-9155. DOI: 10.1088/0031-9155/30/4/008.

[50] Ihab S. Haddadin et al. 'Metabolite quantification and high-field MRS in breast cancer'. In: *NMR Biomed.* 22.1 (Jan. 2009), pp. 65–76. ISSN: 09523480. DOI: 10.1002/nbm.1217. URL: http://doi.wiley.com/10.1002/nbm.1217.

[51] Nima Hatami, Michaël Sdika and Hélène Ratiney. 'Magnetic Resonance Spectroscopy Quantification using Deep Learning'. In: *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* 11070 LNCS (June 2018), pp. 467–475. ISSN: 16113349. DOI: 10.1007/978-3-030-00928-1_53. arXiv: 1806.07237. URL: http://arxiv.org/abs/1806.07237.

[52] Kengo Hattori et al. 'Development of MRI phantom equivalent to human tissues for 3.0-T MRI'. In: *Medical Physics* 40.3 (2013), p. 032303. DOI: 10.1118/1.4790023. eprint: https://aapm.onlinelibrary.wiley.com/doi/pdf/10.1118/1.4790023. URL: https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.4790023.

[53] Rahel Heule et al. 'Local quantum control of Heisenberg spin chains'. In: *Phys. Rev. A* 82.5 (Nov. 2010), p. 52333. ISSN: 1050-2947. DOI: 10.1103/PhysRevA.82.052333. arXiv: arXiv:1007.2572v2. URL: http://link.aps.org/doi/10.1103/PhysRevA.82.052333.

[54] H. J. Hogben et al. 'Spinach - A software library for simulation of spin dynamics in large spin systems'. In: *J. Magn. Reson.* 208.2 (2011), pp. 179–194. ISSN: 10907807. DOI: 10.1016/j.jmr.2010.11.008. URL: http://dx.doi.org/10.1016/j.jmr.2010.11.008.

[55] Matthew James. 'Quantum Control Theory'. In: *arXiv* (June 2014), pp. 1–35. ISSN: {<}null{>}. arXiv: 1406.5260. URL: http://arxiv.org/abs/1406.5260.

[56] C. Jenkins et al. 'Seeking Ground Truth for GABA Quantification by Edited Magnetic Resonance Spectroscopy:Comparative Analysis of TARQUIN, LCModel, JMRUI and GANNET'. In: Mc (Sept. 2019), pp. 1–23. arXiv: 1909.02163. URL: http://arxiv.org/abs/1909.02163.

[57] Christopher Jenkins et al. 'Quantification of edited magnetic resonance spectroscopy : a comparative phantom based study of analysis methods'. In: *Int. Soc. Magn. Reson. Med.* Vol. 62. 2018, pp. 2–5.

[58] Chris Jenkins et al. 'Quantification of edited magnetic resonance spectroscopy: a comparative phantom based study of analysis methods'. In: *Proceedings of ISMRM Annual Meeting*. 2019.

[59] Chris Jenkins et al. 'Quantification of edited magnetic resonance spectroscopy: a comparative phantom based study of analysis methods'. In: *Proc. ISMRM Annu. Meet.* 2019.

[60] L. G. Kaiser et al. 'A detailed analysis of localized J-difference GABA editing: theoretical and experimental study at 4 T'. In: *NMR Biomed.* 21.1 (Jan. 2008), pp. 22–32. ISSN: 09523480. DOI: 10.1002/nbm.1150. URL: http://doi.wiley.com/10.1002/nbm.1150.

[61] Demet Kalayci et al. 'A proton magnetic resonance spectroscopy study in schizoaffective disorder: Comparison of bipolar disorder and schziophrenia'. In: *Prog. Neuro-Psychopharmacology Biol. Psychiatry* 37.1 (2012), pp. 176–181. ISSN: 02785846. DOI: 10.1016/j.pnpbp.2012.01.010.

[62] Martin Kanowski et al. 'Quantitation of simulated short echo time1H human brain spectra by LCModel and AMARES'. In: *Magn. Reson. Med.* 51.5 (May 2004), pp. 904–912. ISSN: 0740-3194. DOI: 10.1002/mrm.20063. URL: http://doi.wiley.com/10.1002/mrm.20063.

[63] Maximilian Keck et al. 'Quantum optimal control within the rotating-wave approximation'. In: *Phys. Rev. A - At. Mol. Opt. Phys.* 92.3 (Sept. 2015), pp. 1–7. ISSN: 10941622. DOI: 10.1103/PhysRevA.92.033402. arXiv: arXiv:1502.07739v1. URL: http://link.aps.org/doi/10.1103/PhysRevA.92.033402.

[64] Navin Khaneja et al. 'Optimal control of spin dynamics in the presence of relaxation'. In: *J. Magn. Reson.* 162.2 (June 2003), pp. 311–319. ISSN: 10907807. DOI: 10.1016/S1090-7807(03)00003-X. arXiv: 0208050 [quant-ph]. URL: http://linkinghub.elsevier.com/retrieve/pii/S109078070300003X.

[65] Navin Khaneja et al. 'Optimal control of coupled spin dynamics: Design of NMR pulse sequences by gradient ascent algorithms'. In: *J. Magn. Reson.* 172.2 (Feb. 2005), pp. 296–305. ISSN: 10907807. DOI: 10.1016/j.jmr.2004.11.004.

[66] Diederik P. Kingma and Jimmy Ba. 'Adam: A Method for Stochastic Optimization'. In: *Int. Conf. Learn. Represent.* (Dec. 2015), pp. 58–62. ISSN: 09252312. DOI: 10.1063/1.4902458. arXiv: 1412.6980. URL: http://arxiv.org/abs/1412.6980.

[67] Melissa Kirkovski et al. 'Short communication: Sex-linked differences in gamma-aminobutyric acid (GABA) are related to social functioning in autism spectrum disorder'. In: *Psychiatry Res. - Neuroimaging* 274.April 2017 (2018), pp. 19–22. ISSN: 18727506. DOI: 10.1016/j.pscychresns.2018.02.004. URL: https://doi.org/10.1016/j.pscychresns.2018.02.004.

[68] Kyryl Kobzar et al. 'Exploring the limits of broadband 90° and 180° universal rotation pulses'. In: *J. Magn. Reson.* 225 (2012), pp. 142–160. ISSN: 10907807. DOI: 10.1016/j.jmr.2012.09.013.

[69] Nina Vanessa Kraguljac et al. 'Neurometabolites in schizophrenia and bipolar disorder - A systematic review and meta-analysis'. In: *Psychiatry Res. - Neuroimaging* 203.2-3 (2012), pp. 111–125. ISSN: 09254927. DOI: 10.1016/j.pscychresns.2012.02.003. URL: http://dx.doi.org/10.1016/j.pscychresns.2012.02.003.

[70] Roland Kreis and Christine Sandra Bolliger. 'The need for updates of spin system parameters, illustrated for the case of γ-aminobutyric acid'. In: *NMR Biomed.* 25.12 (Dec. 2012), pp. 1401–1403. ISSN: 09523480. DOI: 10.1002/nbm.2810. URL: http://doi.wiley.com/10.1002/nbm.2810.

[71] Ilya Kuprov. 'Fokker-Planck formalism in magnetic resonance simulations'. In: *J. Magn. Reson.* 270 (2016), pp. 124–135. ISSN: 10960856. DOI: 10.1016/j.jmr.2016.07.005. URL: http://dx.doi.org/10.1016/j.jmr.2016.07.005.

[72] John Kurhanewicz and Daniel B. Vigneron. 'Advances in MR Spectroscopy of the Prostate'. In: *Magn. Reson. Imaging Clin. N. Am.* 16.4 (2008), pp. 697–710. ISSN: 10649689. DOI: 10.1016/j.mric.2008.07.005.

[73] Joseph B Lambert, Eugene P Mazzola and Clark D Ridge. *Nuclear magnetic resonance spectroscopy: an introduction to principles, applications, and experimental methods*. Wiley, 2019.

[74] Frank C. Langbein, Sophie Schirmer and Edmond Jonckheere. 'Time optimal information transfer in spintronics networks'. In: *Proc. IEEE Conf. Decis. Control* 54rd IEEE (2015), pp. 6454–6459. ISSN: 07431546. DOI: 10.1109/CDC.2015.7403236. arXiv: arXiv:1508.00928v1.

[75] S.M. Shermer; C. Jenkins; M. Chandler; F. C. Langbein. *Magnetic resonance spectroscopy data for GABA quantification using MEGAPRESS pulse sequence*. 2019. DOI: 10.21227/ak1d-3s20. URL: http://dx.doi.org/10.21227/ak1d-3s20.

[76] C.L. Lawson and R.J. Hanson. *Solving Least Squares Problems*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 1995. ISBN: 9780898713565. URL: https://books.google.co.uk/books?id=ROw4hU85nz8C.

[77] Dongwook Lee et al. 'Deep residual learning for accelerated MRI using magnitude and phase networks'. In: *IEEE Trans. Biomed. Eng.* 65.9 (2018), pp. 1985–1995. ISSN: 15582531. DOI: 10.1109/TBME.2018.2821699. arXiv: arXiv:1804.00432v1.

[78] Hyeong Hun Lee and Hyeonjin Kim. 'Intact metabolite spectrum mining by deep learning in proton magnetic resonance spectroscopy of the brain'. In: *Magn. Reson. Med.* 82.November 2018 (Mar. 2019), pp. 33–48. ISSN: 07403194. DOI: 10.1002/mrm.27727. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.27727.

[79] Kin Men Leong, Peter Lau and Saadallah Ramadan. 'Utilisation of MR spectroscopy and diffusion weighted imaging in predicting and monitoring of breast cancer response to chemotherapy'.

In: *J. Med. Imaging Radiat. Oncol.* 59.3 (2015), pp. 268–277. ISSN: 17549485. DOI: 10.1111/1754-9485.12310.

[80] Nicola Longo et al. 'Disorders of creatine transport and metabolism'. In: *Am. J. Med. Genet. Part C Semin. Med. Genet.* 157.1 (2011), pp. 72–78. ISSN: 15524868. DOI: 10.1002/ajmg.c.30292.

[81] J. J. Luykx et al. 'Region and state specific glutamate downregulation in major depressive disorder: A meta-analysis of 1H-MRS findings'. In: *Neurosci. Biobehav. Rev.* 36.1 (2012), pp. 198–205. ISSN: 01497634. DOI: 10.1016/j.neubiorev.2011.05.014. URL: http://dx.doi.org/10.1016/j.neubiorev.2011.05.014.

[82] R. Bruce Lydiard. 'The role of GABA in anxiety disorders'. In: *J. Clin. Psychiatry* 64.SUPPL. 3 (2003), pp. 21–27. ISSN: 01606689.

[83] S. Machnes et al. 'Comparing, optimizing, and benchmarking quantum-control algorithms in a unifying programming framework'. In: *Phys. Rev. A - At. Mol. Opt. Phys.* 84.2 (Aug. 2011), pp. 1–23. ISSN: 10502947. DOI: 10.1103/PhysRevA.84.022305. arXiv: 1011.4874. URL: http://arxiv.org/abs/1011.4874%20http://link.aps.org/doi/10.1103/PhysRevA.84.022305.

[84] Malcolm H. Levitt. 'Spin dynamics: Basics of Nuclear Magnetic Resonance, Second Edition'. In: *Med. Phys.* 37.1 (Dec. 2009), pp. 406–407. ISSN: 00942405. DOI: 10.1118/1.3273534. URL: http://doi.wiley.com/10.1002/nbm.1356%20http://doi.wiley.com/10.1118/1.3273534.

[85] S. Maltezos et al. 'Glutamate/glutamine and neuronal integrity in adults with ADHD: A proton MRS study'. In: *Transl. Psychiatry* 4.3 (2014). ISSN: 21583188. DOI: 10.1038/tp.2014.11.

[86] Ian Marshall et al. 'Use of Voigt lineshape for quantification of in vivo 1H spectra'. In: *Magn. Reson. Med.* 37.5 (1997), pp. 651–657. ISSN: 07403194. DOI: 10.1002/mrm.1910370504.

[87] Anouk Marsman et al. 'GABA and glutamate in schizophrenia: A 7 T 1H-MRS study'. In: *NeuroImage Clin.* 6 (2014), pp. 398–407. ISSN: 22131582. DOI: 10.1016/j.nicl.2014.10.005. URL: http://dx.doi.org/10.1016/j.nicl.2014.10.005.

[88] D. A. McCormick. 'GABA as an inhibitory neurotransmitter in human cerebral cortex'. In: *J. Neurophysiol.* 62.5 (Nov. 1989), pp. 1018–1027. ISSN: 0022-3077. DOI: 10.1152/jn.1989.62.5.1018. URL: http://www.physiology.org/doi/10.1152/jn.1989.62.5.1018.

[89] Donald A. McQuarrie. 'Proton magnetic resonance spectroscopy'. In: *J. Chem. Educ.* 65.5 (2009), p. 426. ISSN: 0021-9584. DOI: 10.1021/ed065p426. URL: http://dx.doi.org/10.1016/j.nic.2012.10.002.

[90] Kate Merritt, Philip McGuire and Alice Egerton. 'Relationship between glutamate dysfunction and symptoms and cognitive function in psychosis'. In: *Front. Psychiatry* 4.NOV (2013), pp. 1–8. ISSN: 16640640. DOI: 10.3389/fpsyt.2013.00151.

[91] M. Mescher et al. 'Solvent suppression using selective echo dephasing'. In: *J. Magn. Reson. - Ser. A* 123.2 (1996), pp. 226–229. ISSN: 10641858. DOI: 10.1006/jmra.1996.0242.

[92] M. Mescher et al. 'Simultaneous in vivo spectral editing and water suppression'. In: *NMR Biomed.* 11.6 (Oct. 1998), pp. 266–272. ISSN: 0952-3480. DOI: 10.1002/(SICI)1099-1492(199810)11:6<266::AID-NBM530>3.0.CO;2-J.

[93] Dieter J. Meyerhoff et al. 'Cortical Gamma-Aminobutyric Acid and Glutamate in Posttraumatic Stress Disorder and Their Relationships to Self-Reported Sleep Quality'. In: *Sleep* 37.5 (2014), pp. 893–900. ISSN: 0161-8105. DOI: 10.5665/sleep.3654.

[94] Šárka Mierisová and Mika Ala-Korpela. 'MR spectroscopy quantitation: A review of frequency domain methods'. In: *NMR Biomed.* 14.4 (2001), pp. 247–259. ISSN: 09523480. DOI: 10.1002/nbm.697.

[95] Mark Mikkelsen et al. 'Big GABA: Edited MR spectroscopy at 24 research sites'. In: *Neuroimage* 159.4 (Oct. 2017), pp. 32–45. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2017.07.021. arXiv: 15334406. URL: http://www.ncbi.nlm.nih.gov/pubmed/24655651.

[96] J Gordon Millichap. 'Thalamic Glutamate/Glutamine in Restless Legs Syndrome'. In: *Pediatr. Neurol. Briefs* 27.7 (2013), p. 55. ISSN: 1043-3155. DOI: 10.15844/pedneurbriefs-27-7-10.

[97] John R. Moffett et al. 'N-Acetylaspartate reductions in brain injury: Impact on post-injury neuroenergetics, lipid synthesis, and protein acetylation'. In: *Front. Neuroenergetics* 5.DEC (2013), pp. 1–19. ISSN: 16626427. DOI: 10.3389/fnene.2013.00011.

[98] José Luis Morales and Jorge Nocedal. 'Remark on "algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound constrained optimization"'. In: *ACM Trans. Math. Softw.* 38.1 (Nov. 2011), pp. 1–

4. ISSN: 00983500. DOI: 10.1145/2049662.2049669. URL: http://dl.acm.org/citation.cfm?doid=2049662.2049669.

[99] E. Mosconi et al. 'Different quantification algorithms may lead to different results: A comparison using proton MRS lipid signals'. In: *NMR Biomed.* 27.4 (2014), pp. 431–443. ISSN: 10991492. DOI: 10.1002/nbm.3079.

[100] John P. Mugler. 'Optimized three-dimensional fast-spin-echo MRI'. In: *J. Magn. Reson. Imaging* 39.4 (2014), pp. 745–767. ISSN: 15222586. DOI: 10.1002/jmri.24542.

[101] Paul G Mullins et al. 'Current practice in the use of MEGA-PRESS spectroscopy for the detection of GABA Cardiff Symposium on MRS of GABA'. In: (). DOI: 10.1016/j.neuroimage.2012.12.004. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3825742/pdf/nihms433653.pdf.

[102] Vinod Nair and Geoffrey E. Hinton. 'Rectified linear units improve restricted boltzmann machines'. In: *ICML*. Oct. 2010.

[103] Jamie Near et al. 'J -difference editing of gamma-aminobutyric acid (GABA): Simulated and experimental multiplet patterns'. In: *Magn. Reson. Med.* 70.5 (Nov. 2013), pp. 1183–1191. ISSN: 07403194. DOI: 10.1002/mrm.24572. URL: http://doi.wiley.com/10.1002/mrm.24572.

[104] NHS England Analytical Services. 'Diagnostic Imaging Dataset Annual Statistical Release 2013/14'. In: (2014), pp. 1–18. URL: https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2014/06/Annual-Statistical-Release-2013-14-DID-pdf-1118KB.pdf.

[105] Kun Nie et al. 'Marked N-acetylaspartate and choline metabolite changes in Parkinson's disease patients with mild cognitive impairment'. In: *Park. Relat. Disord.* 19.3 (2013), pp. 329–334. ISSN: 13538020. DOI: 10.1016/j.parkreldis.2012.11.012. URL: http://dx.doi.org/10.1016/j.parkreldis.2012.11.012.

[106] Ruth L. O'Gorman et al. 'In vivo detection of GABA and glutamate with MEGA-PRESS: Reproducibility and gender effects'. In: *J. Magn. Reson. Imaging* 33.5 (2011), pp. 1262–1267. ISSN: 10531807. DOI: 10.1002/jmri.22520.

[107] Michael A. Ohliger and Daniel K. Sodickson. 'An introduction to coil array design for parallel MRI'. In: *NMR Biomed.* 19.3 (2006), pp. 300–315. ISSN: 09523480. DOI: 10.1002/nbm.1046.

[108] WWF Pijnappel et al. 'SVD-based quantification of magnetic resonance signals'. In: *Journal of Magnetic Resonance (1969)* 97.1 (1992), pp. 122–134.

[109] F. Poggiali, P. Cappellaro and N. Fabbri. 'Optimal Control for One-Qubit Quantum Sensing'. In: *Phys. Rev. X* 8.2 (2018), p. 21059. ISSN: 21603308. DOI: 10.1103/PhysRevX.8.021059. URL: https://doi.org/10.1103/PhysRevX.8.021059.

[110] Michael S. Poole and N. Jon Shah. 'Convex optimisation of gradient and shim coil winding patterns'. In: *J. Magn. Reson.* 244 (2014), pp. 36–45. ISSN: 10960856. DOI: 10.1016/j.jmr.2014.04.015.

[111] Ziyafer Gizem Portakal et al. 'Design and Characterisation of Tissue-Mimicking Gel Phantoms for Diffusion Kurtosis Imaging'. In: (Aug. 2016). arXiv: 1608.08542. URL: http://arxiv.org/abs/1608.08542.

[112] Jean-Baptiste Poullet et al. 'An automated quantitation of short echo time MRS spectra in an open source software environment: AQSES'. In: *NMR Biomed.* 20.5 (Aug. 2007), pp. 493–504. ISSN: 09523480. DOI: 10.1002/nbm.1112. URL: http://www.dotynmr.com/PDF/SA%7B%5C_%7DRFCoils2b.pdf%20http://doi.wiley.com/10.1002/nbm.1112.

[113] Stephen W Provencher. 'LCModel & LCMgui user's manual'. In: (2019).

[114] Stephen W. Provencher. 'Estimation of metabolite concentrations from localized in vivo proton NMR spectra.' In: *Magn. Reson. Med.* 30.6 (Dec. 1993), pp. 672–9. ISSN: 0740-3194. DOI: 10.1002/mrm.1910300604. URL: http://www.ncbi.nlm.nih.gov/pubmed/8139448.

[115] Stephen W. Provencher. 'Automatic quantitation of localized in vivo 1H spectra with LCModel'. In: *NMR Biomed.* 14.4 (2001), pp. 260–264. ISSN: 09523480. DOI: 10.1002/nbm.698.

[116] *PyGamma – The GAMMA Library*. URL: https://scion.duhs.duke.edu/vespa/gamma/wiki/PyGamma (visited on 02/05/2019).

[117] N. Rach et al. 'Dressing the chopped-random-basis optimization: A bandwidth-limited access to the trap-free landscape'. In: *Phys. Rev. A* 92.6 (June 2015). ISSN: 24699934. DOI: 10.1103/PhysRevA.92.062343. arXiv: 1506.04601. URL: http://arxiv.org/abs/1506.04601.

[118] Eric Van Reeth et al. 'A simplified framework to optimize MRI contrast preparation'. In: *Magn. Reson. Med.* 81.1 (2019), pp. 424–438. ISSN: 15222594. DOI: 10.1002/mrm.27417.

[119] Hans J. Reich. '5.3 Spin-Spin Splitting: J-Coupling'. In: *Struct. Determ. Using NMR Note*. 2017.

[120] Greg Reynolds et al. 'An algorithm for the automated quantitation of metabolites in in vitro NMR signals'. In: *Magn. Reson. Med.* 56.6 (2006), pp. 1211–1219. ISSN: 07403194. DOI: 10.1002/mrm.21081.

[121] S. F. Riches et al. 'Multivariate modelling of prostate cancer combining magnetic resonance derived T2, diffusion, dynamic contrast-enhanced and spectroscopic parameters'. In: *Eur. Radiol.* 25.5 (2015), pp. 1247–1256. ISSN: 14321084. DOI: 10.1007/s00330-014-3479-0.

[122] Ángel Rivas and Susana F. Huelga. *Open Quantum Systems. An Introduction*. SpringerBriefs in Physics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 97–147. ISBN: 9783642233531. DOI: 10.1007/978-3-642-23354-8. arXiv: 1104.5242. URL: http://arxiv.org/abs/1104.5242%7B%5C%%7D0Ahttp://dx.doi.org/10.1007/978-3-642-23354-8.

[123] Laetitia Rouger et al. 'Ultrafast double-quantum NMR spectroscopy with optimized sensitivity for the analysis of mixtures'. In: *Analyst* 141.5 (2016), pp. 1686–1692. ISSN: 13645528. DOI: 10.1039/c6an00089d.

[124] Armin Rund et al. 'Magnetic Resonance RF Pulse Design by Optimal Control with Physical Constraints'. In: *IEEE Trans. Med. Imaging* 37.2 (2018), pp. 461–472. ISSN: 1558254X. DOI: 10.1109/TMI.2017.2758391.

[125] Muhammad G. Saleh et al. 'Multi-vendor standardized sequence for edited magnetic resonance spectroscopy'. In: *Neuroimage* 189.January (2019), pp. 425–431. ISSN: 10959572. DOI: 10.1016/j.neuroimage.2019.01.056.

[126] Gerard Sanacora et al. 'Reduced cortical γ-aminobutyric acid levels in depressed patients determined by proton magnetic resonance spectroscopy'. In: *Arch. Gen. Psychiatry* 56.11 (1999), pp. 1043–1047. ISSN: 0003990X. DOI: 10.1001/archpsyc.56.11.1043.

[127] S. G. Schirmer, H. Fu and A. I. Solomon. 'Complete controllability of quantum systems'. In: *arXiv Prepr.* 34.May (May 2000), pp. 1–8. ISSN: 1050-2947. DOI: 10.1103/PhysRevA.63.063410. arXiv: 0010031 [quant-ph]. URL: http://arxiv.org/abs/quant-ph/0010031%7B%5C%%7D0Ahttp://dx.doi.org/10.1103/PhysRevA.63.063410.

[128] S. G. Schirmer, I. C. H. Pullen and A. I. Solomon. 'Controllability of Quantum Systems'. In: (2003), p. 6. ISSN: 14746670. DOI: 10.1016/S1474-6670(17)38905-X. arXiv: 0302121 [quant-ph]. URL: http://arxiv.org/abs/quant-ph/0302121.

[129] Joseph Scott et al. 'A comparison of two post-processing analysis methods to quantify cerebral metabolites measured via proton magnetic resonance spectroscopy in HIV disease'. In: *Br. J. Radiol.* 89.1060 (Apr. 2016), p. 20150979. ISSN: 0007-1285. DOI: 10.1259/bjr.20150979. URL: http://www.birpublications.org/doi/10.1259/bjr.20150979.

[130] Ilah Shin et al. 'Optimisation of the MR protocol in pregnant women with suspected acute appendicitis'. In: *Eur. Radiol.* 28.2 (2018), pp. 514–521. ISSN: 14321084. DOI: 10.1007/s00330-017-5038-y.

[131] Karen Simonyan and Andrew Zisserman. 'Very Deep Convolutional Networks for Large-Scale Image Recognition'. In: *Int. Conf. Learn. Represent.* Sept. 2015. arXiv: 1409.1556. URL: http://arxiv.org/abs/1409.1556.

[132] Helen B. Simpson et al. 'Investigation of cortical glutamate-glutamine and γ-aminobutyric acid in obsessive-compulsive disorder by proton magnetic resonance spectroscopy'. In: *Neuropsychopharmacology* 37.12 (2012), pp. 2684–2692. ISSN: 0893133X. DOI: 10.1038/npp.2012.132.

[133] Robin Simpson et al. 'Advanced processing and simulation of MRS data using the FID appliance (FID-A)—An open source, MATLAB-based toolkit'. In: *Magn. Reson. Med.* 77.1 (2017), pp. 23–33. ISSN: 15222594. DOI: 10.1002/mrm.26091.

[134] Thomas E. Skinner et al. 'Application of optimal control theory to the design of broadband excitation pulses for high-resolution NMR'. In: *J. Magn. Reson.* 163.1 (July 2003), pp. 8–15. ISSN: 10907807. DOI: 10.1016/S1090-7807(03)00153-8. URL: http://linkinghub.elsevier.com/retrieve/pii/S1090780703001538.

[135] Thomas E. Skinner et al. 'Reducing the duration of broadband excitation pulses using optimal control with limited RF amplitude'. In: *J. Magn. Reson.* 167.1 (2004), pp. 68–74. ISSN: 10907807. DOI: 10.1016/j.jmr.2003.12.001.

[136] Thomas E. Skinner et al. 'Optimal control design of constant amplitude phase-modulated pulses: Application to calibration-free broadband excitation'. In: *J. Magn. Reson.* 179.2 (2006), pp. 241–249. ISSN: 10907807. DOI: 10.1016/j.jmr.2005.12.010.

[137] Thomas E. Skinner et al. 'Optimal control design of band-selective excitation pulses that accommodate relaxation and RF inhomogeneity'. In: *J. Magn. Reson.* 217 (Nov. 2012), pp. 53–60. ISSN: 10907807. DOI: 10.1016/j.jmr.2012.02.007. arXiv: 1111.6650. URL: http://arxiv.org/abs/1111.6650.

[138] Charles P Slichter. *Principles of magnetic resonance*. Vol. 1. Springer Science & Business Media, 2013.

[139] I.M Sobol'. 'On the distribution of points in a cube and the approximate evaluation of integrals'. In: *USSR Comput. Math. Math. Phys.* 7.4 (Jan. 1967), pp. 86–112. ISSN: 00415553. DOI: 10.1016/0041-5553(67)90144-9. URL: https://linkinghub.elsevier.com/retrieve/pii/0041555367901449.

[140] Bj Soher et al. 'VeSPA: Integrated applications for RF pulse design, spectral simulation and MRS data analysis'. In: *Proc. 19th Annu. Meet. ISMRM* 19.19 (2011), p. 1410. URL: http://scion.duhs.duke.edu/vespa/%7B%5C%%7D0Ahttp://www.swig.org..

[141] Unparalleled M R I Solutions. *Preclinical MRI Systems for Mice up to 15 . 2 Tesla*. Tech. rep. 2018.

[142] J. J.W.H. Sørensen et al. 'Quantum optimal control in a chopped basis: Applications in control of Bose-Einstein condensates'. In: *Phys. Rev. A* 98.2 (2018). ISSN: 24699934. DOI: 10.1103/PhysRevA.98.022119. arXiv: 1802.07509.

[143] Cecilia L. Speyer et al. 'Metabotropic glutamate receptor-1: A potential therapeutic target for the treatment of breast cancer'. In: *Breast Cancer Res. Treat.* 132.2 (2012), pp. 565–573. ISSN: 01676806. DOI: 10.1007/s10549-011-1624-x.

[144] Jost Tobias Springenberg et al. 'Striving for Simplicity: The All Convolutional Net'. In: *Int. Conf. Learn. Represent.* (Dec. 2015), pp. 1–14. ISSN: 02548704. DOI: 10.1163/_q3_SIM_00374. arXiv: 1412.6806. URL: http://arxiv.org/abs/1412.6806.

[145] Greg J. Stanisz et al. 'T1, T2 relaxation and magnetization transfer in tissue at 3T'. In: *Magn. Reson. Med.* 54.3 (2005), pp. 507–512. ISSN: 07403194. DOI: 10.1002/mrm.20605.

[146] Zenon Starčuk et al. 'Simulation of coupled-spin systems in the steady-state free-precession acquisition mode for fast magnetic resonance (MR) spectroscopic imaging'. In: *Measurement Science and Technology* 20.10 (Sept. 2009), p. 104033. DOI: 10.1088/0957-0233/20/10/104033. URL: https://doi.org/10.1088%2F0957-0233%2F20%2F10%2F104033.

[147] D Stefan et al. 'Quantitation of magnetic resonance spectroscopy signals: the jMRUI software package'. In: *Meas. Sci. Technol.* 20.10 (Oct. 2009), p. 104035. ISSN: 0957-0233. DOI: 10.1088/0957-0233/20/10/104035.

[148] Stephen Provencher. *LCModel License*. URL: http://s-provencher.com/lcm-license.shtml (visited on 27/09/2019).

[149] Andrzej Stepulak et al. 'Glutamate and its receptors in cancer'. In: *J. Neural Transm.* 121.8 (2014), pp. 933–944. ISSN: 14351463. DOI: 10.1007/s00702-014-1182-6.

[150] Alina Strasser et al. 'Nucleus accumbens neurochemistry in human anxiety: A 7 T 1 H-MRS study'. In: *Eur. Neuropsychopharmacol.* 29.3 (2019), pp. 365–375. ISSN: 18737862. DOI: 10.1016/j.euroneuro.2018.12.015. URL: https://doi.org/10.1016/j.euroneuro.2018.12.015.

[151] Fangfang Tang et al. 'An improved asymmetric gradient coil design for high-resolution MRI head imaging'. In: *Phys. Med. Biol.* 61.24 (2016), pp. 8875–8889. ISSN: 13616560. DOI: 10.1088/1361-6560/61/24/8875.

[152] Reggie Taylor et al. 'Neurometabolic abnormalities in schizophrenia and depression observed with magnetic resonance spectroscopy at 7 T'. In: *BJPsych Open* 3.01 (2017), pp. 6–11. DOI: 10.1192/bjpo.bp.116.003756.

[153] L. Tebartz Van Elst et al. 'Disturbed cingulate glutamate metabolism in adults with high-functioning autism spectrum disorder: Evidence in support of the excitatory/inhibitory imbalance hypothesis'. In: *Mol. Psychiatry* 19.12 (2014), pp. 1314–1325. ISSN: 14765578. DOI: 10.1038/mp.2014.62.

[154] Joshua M. Tognarelli et al. 'Magnetic Resonance Spectroscopy: Principles and Techniques: Lessons for Clinicians'. In: *J. Clin. Exp. Hepatol.* 5.4 (2015), pp. 320–328. ISSN: 22133453. DOI: 10.1016/j.jceh.2015.10.006. URL: http://dx.doi.org/10.1016/j.jceh.2015.10.006.

[155] Shang Yueh Tsai et al. 'Effects of Frequency Drift on the Quantification of Gamma-Aminobutyric Acid Using MEGA-PRESS'. In: *Sci. Rep.* 6.September 2015 (2016), pp. 1–8. ISSN: 20452322. DOI: 10.1038/srep24564.

[156] N. Tubridy et al. 'Optimisation of unenhanced MRI for detection of lesions in multiple sclerosis: A comparison of five pulse sequences with variable slice thickness'. In: *Neuroradiology* 40.5 (1998), pp. 293–297. ISSN: 00283940. DOI: 10.1007/s002340050587.

[157] Thomas Schulte-herbr Uggen. 'Gradient Flows for Optimization in Quantum Information and Quantum Dynamics :' in: *Rev. Math. Phys.* 22.6 (Feb. 2010), pp. 597–667. ISSN: 0129-055X. DOI: 10.1142/S0129055X10004053. arXiv: 0802.4195. URL: http://www.worldscientific.com/doi/abs/10.1142/S0129055X10004053.

[158] Anna S. Urrila et al. 'Frontal Cortex Myo-Inositol Is Associated with Sleep and Depression in Adolescents: A Proton Magnetic Resonance Spectroscopy Study'. In: *Neuropsychobiology* 75.1 (2017), pp. 21–31. ISSN: 0302-282X. DOI: 10.1159/000478861. URL: https://www.karger.com/Article/FullText/478861.

[159] Eric Van Reeth et al. 'Optimal control theory for applications in Magnetic Resonance Imaging'. In: *Pacific J. Math. Ind.* 9.1 (2017), pp. 1–19. ISSN: 2198-4115. DOI: 10.1186/s40736-017-0034-3. arXiv: 1709.02751.

[160] L. M.K. Vandersypen and I. L. Chuang. 'NMR techniques for quantum control and computation'. In: *Rev. Mod. Phys.* 76.4 (Jan. 2004), pp. 1037–1069. ISSN: 00346861. DOI: 10.1103/RevModPhys.76.1037. arXiv: 0404064 [quant-ph]. URL: http://link.aps.org/doi/10.1103/RevModPhys.76.1037.

[161] Sadhna Verma et al. 'Prostate MRI and 3D MR spectroscopy: How we do it'. In: *Am. J. Roentgenol.* 194.6 (2010), pp. 1414–1426. ISSN: 0361803X. DOI: 10.2214/AJR.10.4312.

[162] Mads S. Vinding et al. 'Fast numerical design of spatial-selective rf pulses in MRI using Krotov and quasi-Newton based optimal

control methods'. In: *J. Chem. Phys.* 137.5 (2012). ISSN: 00219606. DOI: 10.1063/1.4739755.

[163] Mads S. Vinding et al. 'Dynamic nuclear polarization and optimal control spatial-selective 13C MRI and MRS'. In: *J. Magn. Reson.* 227 (Feb. 2013), pp. 57–61. ISSN: 10907807. DOI: 10.1016/j.jmr.2012.12.002. URL: http://dx.doi.org/10.1016/j.jmr.2012.12.002.

[164] J Werschnik and E K U Gross. 'Quantum Optimal Control Theory'. In: (July 2007). arXiv: 0707.1883. URL: http://arxiv.org/abs/0707.1883.

[165] Andrew Gordon Wilson et al. 'Bayesian Inference for NMR Spectroscopy with Applications to Chemical Quantification'. In: 1992 (Feb. 2014), pp. 1–26. arXiv: 1402.3580. URL: http://arxiv.org/abs/1402.3580.

[166] Martin Wilson et al. 'A constrained least-squares approach to the automated quantitation of in vivo ^{1}H magnetic resonance spectroscopy data.' In: *Magn. Reson. Med.* 65.1 (Jan. 2011), pp. 1–12. ISSN: 1522-2594. DOI: 10.1002/mrm.22579. URL: http://doi.wiley.com/10.1002/mrm.22579.

[167] Martin Wilson et al. 'Methodological consensus on clinical proton MRS of the brain: Review and recommendations'. In: *Magn. Reson. Med.* 82.2 (2019), pp. 527–550. ISSN: 15222594. DOI: 10.1002/mrm.27742.

[168] Eric C. Wong, A. Jesmanowicz and James S. Hyde. 'Coil optimization for MRI by conjugate gradient descent'. In: *Magn. Reson. Med.* 21.1 (1991), pp. 39–48. ISSN: 15222594. DOI: 10.1002/mrm.1910210107.

[169] Bing Xu et al. 'Empirical Evaluation of Rectified Activations in Convolutional Network'. In: (2015). arXiv: 1505.00853. URL: http://arxiv.org/abs/1505.00853.

[170] Behrouz Zand et al. 'Role of Increased n-acetylaspartate Levels in Cancer'. In: *J. Natl. Cancer Inst.* 108.6 (2016), pp. 1–11. ISSN: 14602105. DOI: 10.1093/jnci/djv426.

[171] Niloufar Zarinabad et al. 'Application of pattern recognition techniques for classification of pediatric brain tumors by in vivo 3T 1H-MR spectroscopy—A multi-center study'. In: *Magn. Reson. Med.* 79.4 (2018), pp. 2359–2366. ISSN: 15222594. DOI: 10.1002/mrm.26837.

[172] Qianru Zhang et al. 'Recent advances in convolutional neural network acceleration'. In: *Neurocomputing* 323 (2019), pp. 37–51. ISSN: 18728286. DOI: 10.1016/j.neucom.2018.09.038. arXiv: 1807.08596.

[173] Y. Zhou and N. C. Danbolt. 'Glutamate as a neurotransmitter in the healthy brain'. In: *J. Neural Transm.* 121.8 (2014), pp. 799–817. ISSN: 14351463. DOI: 10.1007/s00702-014-1180-8.

[174] Ciyou Zhu et al. 'Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization'. In: *ACM Trans. Math. Softw.* 23.4 (Dec. 1997), pp. 550–560. ISSN: 00983500. DOI: 10.1145/279232.279236. URL: http://portal.acm.org/citation.cfm?doid=279232.279236.

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede and Ivo Pletikosić. The style was inspired by Robert Bringhurst's seminal book on typography *"The Elements of Typographic Style"*. `classicthesis` is available for both LaTeX and LyX:

https://bitbucket.org/amiede/classicthesis/

Final Version as of 3rd May 2020 (classicthesis v4.6).