The Analysis & Automatic Classification of Nuclear Magnetic Resonance Signals

Catherine A. Ojo

Doctor of Philosophy

Institute for Adaptive and Neural Computation
School of Informatics
University of Edinburgh

2009

Abstract

The human brain consists of a myriad of chemical compounds critical to its functioning. A group of these compounds, collectively known as metabolites, have been a research interest for years because the pathogenesis of neurodegenerative diseases, a tumours classification, the effectiveness of a drug, etc., can be investigated via variations in brain metabolite concentration levels. Nuclear Magnetic Resonance Spectroscopy (NMRS) enables investigators to conduct non-invasive in vivo studies of metabolites in the human brain and the rest of the body. However a number of problems have hindered the usage of NMRS as a clinical diagnostic tool. One is the non-uniqueness of the most widely used analysis methods, i.e. as the parameters and/or prior knowledge data of an analysis method are changed, the results also change. A second problem is the lack of a method that can automatically classify the signal components estimated via signal decomposition based signal analysis methods. Additionally, some of the most widely used analysis methods, by virtue of their algorithms, intrinsically assume the nature of NMRS signals, e.g. stationary, linear, Lorentzian, etc. Hence, this thesis explores a new analysis approach, based on a theoretical and practical understanding of NMRS, that (a) avoids making assumptions about the nature of experimentally acquired NMRS signals, (b) relies on a unique decomposition analysis method, and (c) automatically classifies the estimated peaks of an analysis. Unique decomposition analysis was conducted via the rarely used *unique* and *non-linear* signal decomposition method − the Fast Padé Transform (FPT). The FPT is compared with the main decomposition based NMRS analysis methods via a detailed mathematical analysis, and a comparative analysis. Automatic classification was conducted via a novel classification method, which is introduced herein, and which is based on quantum mechanical predictions of metabolite NMRS behaviour.

Acknowledgements

My thanks to my supervisors Professors McLauglin and Marshall, I know their supervision would stand me in good stead. My thanks to Professor Williams and Dr. Lymer for the helpful discussions. My thanks to Trevor Capenter, James Withers, Dr. Lymer, and the radiographers Elaine Sandeman, Iona Hamilton, and Jenny Boyd-Ellison (SFC Brain Imaging Research Centre, The University of Edinburgh) for their help in relation to the experiments. My thanks to Neuroinformatics DTC for its support, and to Penicuik Library, where this thesis was written, for providing a conducive working space.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Catherine A. Ojo*)

Table of Contents

Nomenclature

$(a\ \ b)$	Range a to b, excluding a and b
$<<$	Much Less Than
$[a\ \ b]$	Range a to b, including a and b
$\cap$	Intersect
$\cdot$	dot product
$\dagger$	Pseudo-inverse $A^{\dagger} = (A^T A)^{-1} A$
δ	Chemical Shift
$\exists$	There Exists
γ	Gyromagnetic Ratio
$\Im$	Imaginary Numbers
$\mathcal{H}$	Hamiltonian
$\mathcal{H}$	Hankel Matrix
$\mathcal{N}$	Natural Numbers
μ	Nuclear Magnetic Moment
ν	Lamor Frequency
ϕ_k	The k^{th} mono-component's phase
$\Re$	Real Numbers
$\Rightarrow$	Implies, which/that Implies
τ	Sampling Time
$\varnothing$	Empty Set
^{1}H	Hydrogen
^{31}P	Phosphorus
a_k	The k^{th} mono-component's amplitude

d_k	The k^{th} mono-component's decay rate
f_k	The k^{th} mono-component's frequency
g_k	The k^{th} mono-component's damping factor correction term
H	Hermitian Conjugate
P_α	Probability of false alarm
P_δ	Probability of detection
T_E	Echo Time
T_R	Repetition Time
B	External Magnetic Field
I	Angular Momentum
m	Magnetic Quantum Number
Tr	The *trace* of a matrix, i.e. the sum of its diagonal elements.

N.B. In general, a matrix is denoted by an upper case letter, a vector is denoted by a bold lower case letter (rarely, by an upper case letter), a quote from a book or journal paper is placed between quotation marks " ", and italics is used to emphasize a point. Table, figure, and equation labels with a letter prefix, e.g. *Table D.2* or *Fig. F.6*, are located in the appendix chapter of the same letter.

Acronyms & Abbreviations

Measurement Units

dB	Decibels
Hz	Hertz

Metabolites & Compounds

Ace	acetate
Bet	betaine
Ch	choline
Cr	creatine
Eth	ethanol
EA	ethanolamine
GABA	γ-amino-butyric acid
Glyc	glycerol
Glye	glycine
GTT	glutathione
GPC	glycero-phosphoryl-choline
LAl	l alanine
LAsp	l aspartic
LGlc	l glutamic
LGle	l glutamine
LLac	l lactate
LPA	l phenylalanine
LSer	l serine

LThr l threonine

LTryp l tryptophan

LTyr l tyrosine

LVal l valine

MSM methyl sulphonyl methane

MI myoinositol

NAA n-acetyl-l-aspartic acid

NAAG n-acetyl-l-aspartic glutamic acid

PCr phosphocreatine

PCh phosphorylcholine

Py pyruvic

SI scylloinositol

Suc succinic

Tau taurine

Miscellaneous Abbreviations

a.u. arbitrary units

d.p. decimal places

s.c. sinus cardinal

w.r.t. with respect to

p.p.m. parts per million

i.i.d independent and identically distributed

Miscellaneous Acronyms

AP Arithmetic Progression

ADC Analogue-to-Digital Converter

AMARES Advanced Method for Accurate, Robust and Efficient Spectral

AQSES Automated Quantitation of Short Echo Time MRS Spectra

CDF Cumulative Distribution Function

FID Free Induction Decay

FIR Finite Impulse Response

FPR False Positive Rate

FPT Fast Padé Transform

GLRT Generalized Likelihood Ratio Test

HLSVD Hankel-Lanczos Singular Value Decomposition

HLTLS Hankel-Lanczos Total Least Squares

HSVD Hankel Singular Value Decomposition

HTLS Hankel Total Least Squares

IRL Implicitly Restarted Lanczos

ISIS Image-Selected *In Vivo* Spectroscopy

LC Linear Combination

NMR Nuclear Magnetic Resonance

NMRS Nuclear Magnetic Resonance Spectroscopy

NSA Number of Signal Averages

MRI Magnetic Resonance Imaging

PA Padé Approximant

PDF Probability Density Function

PRESS Point Resolved Spectroscopy

PRO Partial Re-orthogonalization

QM Quantum Mechanics

QMP Quantum Mechanically Predicted

QR Orthogonal Triangular Decomposition

QUEST Quantum Estimation

ROC Receiver Operating Characteristics

ROI Region of Interest

RF Radio Frequency

SNR Signal-to-Noise Ratio

SSE Sum of the Squared Errors

STEAM Stimulated Echo Acquisition Mode

SVD Singular Value Decomposition

SVS Single Voxel Spectroscopy

TMS Tetra-Methyl-Silane

TPR True Positive Rate

VARPRO Variable Projection

VEST Volume Excitation using Stimulated echoes

WGN White Gaussian Noise

Chapter 1

Introduction

Nuclear Magnetic Resonance Spectroscopy (NMRS) enables non-invasive in vivo studies of metabolites in the human brain, and the rest of the body, by exploiting the magnetic properties of a metabolite's nuclei. In the early days of human NMRS, ^{31}P spectroscopy experiments, i.e. experiments that exploit the nuclear magnetic properties of phosphorus atoms, were the most common form of in vivo spectroscopy experiments conducted [5, 61]. This was partly due to the ease with which the experiments could be conducted, and partly due to scientists' interest in the role of phosphorus in energy metabolism [5, 61]. However, the analysis of ^{31}P spectra is problematic due to the [relatively] low sensitivity of ^{31}P. The sensitivity of a nucleus [type] is a measure of the energy required to move the nucleus from one energy level to another; the greater the energy required to achieve this goal, the greater the energy released as the nucleus returns to its original energy level, hence the greater its sensitivity [1].

Presently, the most common form of human spectroscopy studies are ^{1}H (Hydrogen) NMRS studies, due to the relatively higher sensitivity of ^{1}H (refer to [1], page 5, for a comparison of the sensitivity of different nuclei). This greater sensitivity is due to a combination of the high natural abundance of ^{1}H and its high gyromagnetic ratio.[1] The prevalence of ^{1}H NMRS studies has also been aided by the advent of Magnetic Resonance Imaging (MRI), which relies on the ^{1}H nucleus; the instruments are designed "primarily for proton MRI" [61] (in this context, and in the field of NMRS generally, proton implies ^{1}H). Hence this thesis focuses on ^{1}H NMRS.

1.1 Key Problems

In spite of the sensitivity advantage, a number of problems have hindered the usage of ^{1}H NMRS as a clinical diagnostic tool. At the signal analysis level, one of the most important of these problems is the subjective and non-unique character of the most widely used NMRS signal analysis methods [8]. In this text

Definition 1 (Subjective) *A method is subjective if its parameters, or prior knowledge data set, are set by the user.*[2]

[1] The gyromagnetic ratio of a particle is the ratio of its magnetic moment to its angular momentum [1].

[2] In NMRS prior knowledge refers to the prior information we have about the spectrum of a metabolite –

and

Definition 2 (Non-unique) *A method is non-unique if the results of its analysis change (or do not converge) as the parameters, or prior knowledge data set, are changed.*

Each of the most widely used NMRS analysis methods is [both] subjective and non-unique with respect to (w.r.t.) the above definitions. Hence, it can be argued that the quality of the results obtained from these methods is as dependent on the user as it is on the method. Clearly, this is not ideal, there are times when an element of subjectivity cannot be avoided, e.g. spectrum shifting in NMRS which ensures that the spectral frequencies are shifted to their equivalent standard frequency values (refer to *Fig. 1.1a* & *1.1b*). However wherever subjectivity cannot be avoided, its effect should be simple and straightforward; the spectrum shifting example illustrated in *Fig. 1.1* is quite simple and straightforward.

The definition of non-uniqueness, *Definition 2*, suggests that sometimes prior knowledge data is not effectively exploited by NMRS analysis methods. This leads to a second problem − the way prior knowledge is used by *interactive* NMRS signal analysis methods, i.e. methods that use prior knowledge. Each of the main interactive methods use prior knowledge data via an iterative fitting algorithm, hence the methods are susceptible to **(1)** under-fitting − due to "model mismatch errors" [40], and **(2)** over-fitting − due to "model estimation errors" [40], both leading to non-unique behaviour. Explicitly, a model mismatch error occurs when the assumed functional form of an event cannot describe the data. For example, the most widely used functional forms used by *model based interactive NMRS analysis methods*, for describing NMRS signal behaviour, are variants of the form

$$x_n = \sum_{k=1}^{K} a_k e^{i\phi_k} e^{-d_k(1-g_k+g_k t_n)t_n} e^{i2\pi f_k t_n}, \qquad x_n \in X, \quad n = 1, \ldots, N \tag{1.1.1}$$

whereby x_n is the n^{th} data point of the *noise free* $N \times 1$ time series signal X. The parameters a_k, ϕ_k, d_k, g_k, and f_k are the respective amplitude, phase, decay constant, decay correction term, and frequency of the k^{th} component of an NMRS multi-component signal. K is the total number of assumed spectral frequencies present at a source, and it is based on the assumed number of spectral frequencies per metabolite assumed present at a source. If a selected variant of the functional form of *Eq. 1.1.1* does not adequately describe the data because it makes inadequate assumptions about the nature of the signal being analysed, its solutions will be quite varied, i.e. it will exhibit non-unique behaviour.

e.g. its spectral frequencies, the amplitude pattern of its peaks, etc. − assumed to be present at a source.

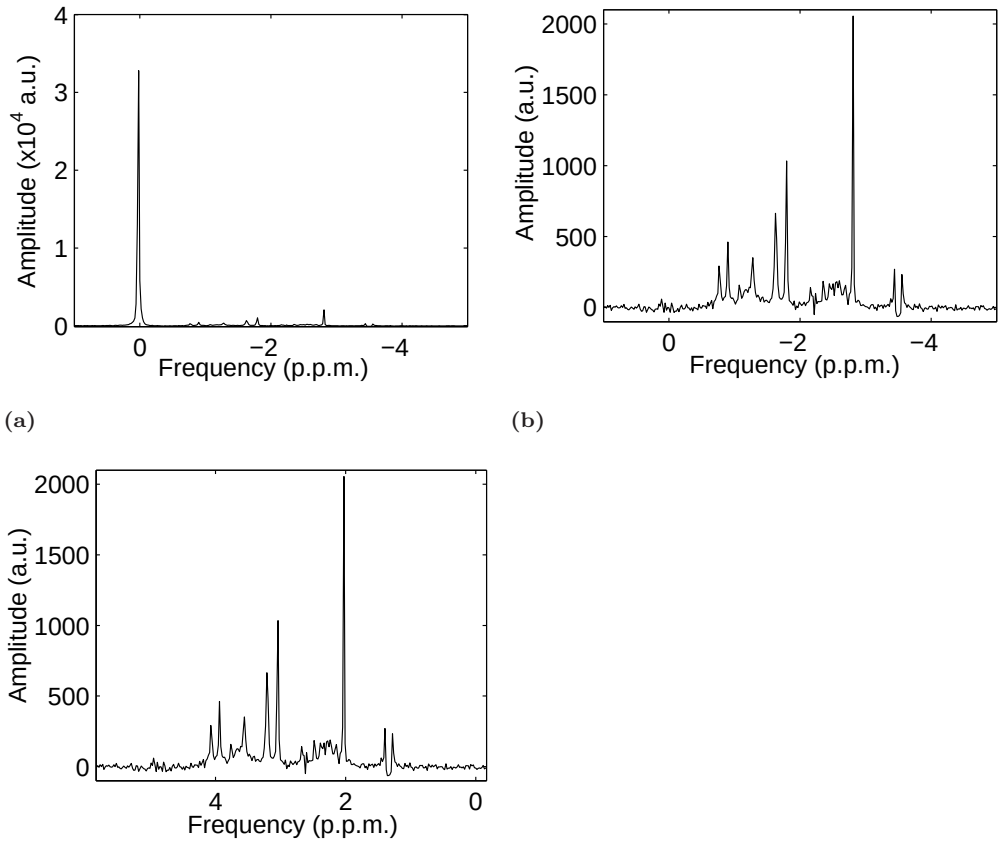


Figure 1.1: The Fourier spectrum of an NMRS signal acquired from a phantom (*Table C.1, page 147*) in a 1.5 tesla NMR scanner; the acquired signal is a time series signal of length 2048. **(a)** depicts the Fourier spectrum before the large water peak at 0 p.p.m. was removed, and **(b)** is the spectrum after the water peak was removed. The peaks of most metabolites appear within the negative scale. However, it is standard practice to discuss the ^{1}H NMRS spectral frequencies as positive scale frequencies, in relation to the reference compound TMS [1, 60]. The standard values are obtained by shifting the spectrum of an acquired signal by the resonance frequency of water's ^{1}H nuclei; after shifting TMS appears at 0 p.p.m. The resonance frequency of water's ^{1}H nuclei, which depends linearly on temperature within a specific range [28], has to be determined by the user using any of a number of methods, hence a degree of subjectivity. In this example the evaluated resonance frequency of water's ^{1}H nuclei − evaluated by calculating the gap between the water and main NAA peak, based on [28] − was 4.835 p.p.m., thus **(b)** was shifted accordingly, leading to **(c)**. The water peak of this signal has been filtered out using the HLSVD Filter of jMRUI [75], a NMRS signal analysis package. **Parts Per Million (p.p.m.)** is the standard unit used for frequency in the field of NMRS [60]. The equivalent frequency in Hertz is given by (frequency (p.p.m.) × Scanner Strength (MHzT^{-1}))/10^6 (this relationship is defined on page 16, by *Eq. 2.3.5*). For example, the scanner used to acquire the above signal has a strength of 63.87 MHzT^{-1}, therefore the frequency in Hertz at 2 p.p.m. is 127.74 Hz.

In the case of *basis set reliant interactive NMRS analysis methods*, whose general functional form is[3]

$$X = \sum_{j=1}^{J} q_j \mathrm{Met}_j \tag{1.1.2}$$

J : The number of metabolites assumed to be present at a source

q_j : Metabolite quantity parameter.

the model mismatch errors will be due to Met_j, which is the predicted or experimentally deduced time series signal or frequency (Fourier) signal of a metabolite, w.r.t. a specific experiment type. In this case non-uniqueness will be an issue if the metabolite signal prediction methods are inadequate descriptors of metabolite behaviour. (N.B. J is the number of metabolites assumed to be present at a source *not* the number of spectral frequencies. Whereas K, of *Eq. 1.1.1*, is the total number of assumed spectral frequencies, based on the assumed number of spectral frequencies per metabolite assumed present at a source.)

On the other hand model estimation errors − i.e. inaccurate estimated parameter values − occur when an inappropriate number of parameters, and/or flawed starting values are used. In both model based and basis set interactive analysis methods, the number of parameters is proportional to the number of metabolites that are assumed to be present in an acquired signal. The starting values of these parameters are normally based on the prior knowledge data of the metabolites assumed present. However the identity of the metabolites present in a region of the human body or brain, and the number of metabolites present, is rarely known. Additionally the aim of the iterative fitting algorithms, which minimize the objective/cost/error functions of the form

$$\|Y - X\|^2 \tag{1.1.3}$$

Y : noisy NMRS multi-component signal

X : a method's assumed functional form or forms

is to ensure that *Eq. 1.1.3* is *minimal* not *minimal and physically plausible*, i.e. its objective is to find a solution that *fits* the assumptions made. Consequently, as the number of parameters and metabolite choices are changed, the results of the interactive analysis methods change, i.e. they exhibit non-unique behaviour. Unfortunately none of the widely used interactive methods includes a systematic means of excluding an incorrect metabolite choice, or of estimating the appropriate number of parameters. An implicit consequence of model estimation errors is the under-estimation or over-estimation of metabolite quantities; problematic w.r.t. clinical diagnostics. All in all, the prior knowledge usage advantage of interactive NMRS analysis methods is tempered by the way the prior knowledge is used.

The most widely used non-interactive methods are also non-unique because their results change as a user's assumptions w.r.t an NMRS signal change. The non-interactive methods

[3] A basis set is a set of metabolite spectra whereby the spectrum of each metabolite has been deduced experimentally or quantum mechanically under the same conditions that would be used for scanning a prospective patient or subject.

decompose an NMRS signal into a noise space, and a *noise free* signal space whose components have the assumed functional form

$$x_n = \sum_{k=1}^{K} a_k e^{i\phi_k} e^{-d_k t_n} e^{i2\pi f_k t_n}, \qquad x_n \in X, \quad n = 1, \ldots, N \tag{1.1.4}$$

i.e. a variant of *Eq. 1.1.1* with $g_k = 0$. In this case the constant K is the number of components assumed present in an NMRS signal. As the value of K is changed the estimated amplitude a_k, phase ϕ_k, decay constant d_k, and frequency f_k parameter values also change. In relation to the common non-interactive methods, it is important to note that if we assume that an NMRS signal has $K = J$ or $K = L$ components, $J \neq L$, the resulting set of estimated parameter values may not overlap, or may only overlap in a few cases, i.e.

$$Set\{a_j, \phi_j, d_j, f_j\} \cap Set\{a_l, \phi_l, d_l, f_l\} = \varnothing \tag{1.1.5}$$

or

$$Set\{a_j, \phi_j, d_j, f_j\} \cap Set\{a_l, \phi_l, d_l, f_l\} < min(J, L) \tag{1.1.6}$$

This occurs because the aim of the common non-interactive NMRS analysis methods is to ensure that

$$x_n + e_n = y_n \tag{1.1.7}$$

y_n : the nth point of the noisy $N \times 1$ NMRS signal

e_n : the nth point of the $N \times 1$ noise signal

i.e. to ensure that the sum of the resulting components and the noise residue, of the noise space, is always equal to the decomposed signal, rather than extracting the physically plausible components. Thus as K is changed, the parameter values are evaluated to ensure that *Eq. 1.1.7* holds. This issue is discussed in more mathematical detail in *Section 3.1.1*, page 38, of the *Non-Interactive Methods* review section.

An additional issue, which mainly applies to the common non-interactive methods, is the metabolite associations of the estimated components. Presently, a user has to *manually* decide which metabolite an estimated component is associated with because there are no automatic methods available for classifying the estimated components or peaks of an analysis.

These three NMRS signal analysis issues − i.e. (1) subjectiveness & non-uniqueness, (2) prior knowledge usage, and (3) the lack of an automatic classification method for the estimated components of non-interactive methods − are the focus of this thesis.

1.2 Project Objective

In order to address the outlined problems, the aim of this thesis was the development of a new analysis approach that[4]

[4] An *analysis approach* refers to a combination of data processing methods (e.g. water filtering, explicit de-noising), a signal analysis method (e.g. the black-box decomposition method HSVD), and statistical assessments (e.g. the probability that an estimated mono-component is a genuine signal) used in NMRS studies.

1. relies on a unique, objective, and non-linear, signal analysis method that does not assume the nature of the signal being analysed.

2. exploits prior knowledge, but not via a fitting method.

3. includes an automatic classification method if the signal analysis method is a decomposition based method.

1.3 Project Achievements

Point 1, of the project objective section above, was addressed via the Fast Padé Transform (FPT) decomposition method. It is a unique and non-linear decomposition analysis method which does not assume the nature of the signal being analysed. The FPT was recently introduced to the field of NMRS signal analysis by Belkić [7, 18, 19]. Although a user-friendly Matlab®, and C++, program is referred to in [7,13], neither program is available to the public, therefore Matlab® code was developed for the FPT work reported in this thesis.

Additionally, because the peaks associated with a metabolite must be known in order to evaluate the absolute quantity of a metabolite [53], the estimated components resulting from FPT analysis were automatically classified via a new classification method; this addresses **Point 3**. The new classification method uses quantum mechanically predicted spectral frequencies of metabolites to classify the estimated components, i.e. **Point 2** (prior knowledge exploitation) was achieved via the new classification method. The classification method can be used to classify the peaks estimated via any NMRS signal analysis method whose results are in the form of estimated components.

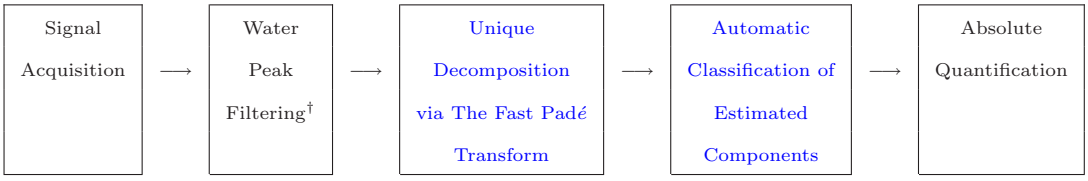


Figure 1.2: Analysis approach framework. The project's core achievement is the development and/or implementation of the modules printed in blue. †Water peak filtering of the residual water peak of an NMRS signal.

Hence, within the analysis approach framework, *Fig. 1.2*, the implementation of the FPT, and the development and implementation of a new classification method, are the core achievements of this thesis. Another main achievement is the detailed mathematical review, and assessment, of the FPT and of the widely used non-interactive NMRS analysis methods, which makes it easier to identify and understand a method's algorithmic problem points. It is the first time that a detailed mathematical review, of the type conducted herein, has been done. The next section introduces the thesis structure.

1.4 Project Overview

The outlined objectives of this thesis were addressed via a number of steps. ***Chapter 2*** focuses on the theoretical and practical aspects of nuclear magnetic resonance spectroscopy (NMRS). The chapter is divided into two parts. The first part starts with a discussion of the physical and chemical properties of metabolites that are exploited by NMRS experiments. This is followed by a discussion of the theoretical models that have been used to describe the behaviour of metabolites during NMRS experiments. It concludes with a discussion of the most widely accepted model of metabolite NMRS behaviour, and how it is used to predict metabolite NMRS behaviour; the latter in preparation for the classification method. The second part of *Chapter 2* discusses the nature of the signal acquired in practice. As in many fields, the nature of the signal acquired in practice differs slightly from theoretical expectations. Hence, the discussion is focused on the factors that affect the nature of an acquired signal, e.g. thermal noise, and the empirical effect of these factors, e.g. poor Signal-to-Noise Ratio (SNR). The overall aim of *Chapter 2* is to understand the range of factors that ought to be taken into consideration during signal analysis.

Chapter 3 is focused on a novel detailed review of the most widely used NMRS analysis methods, bearing in mind the observations of *Chapter 2*. The first part of the chapter discusses the strengths and weaknesses of the interactive and non-interactive NMRS analysis methods. The non-interactive methods discussion includes a detailed mathematical analysis of the methods, which helps to pinpoint the areas of potential algorithm improvement. The second part of the chapter is focused on a comparative analysis of the non-interactive methods.

Based on a general understanding of the theoretical and practical aspects of NMRS as discussed in *Chapter 2*, and the strengths and weaknesses of the analysis methods discussed in *Chapter 3*, ***Chapter 4*** introduces a new automatic classification method. Subsequently, the receiver operating characteristics of the new classification method are assessed via noisy artificial signals.

Chapter 5 is focused on the quantitative performance analysis of the classification method introduced in *Chapter 4*. The classification method is tested on phantom data, of known chemical components, and on human data, of unknown components.

Chapter 6 is the concluding chapter, it summarizes the findings and contributions of the project, and suggests future work.

Chapter 2

Theoretical & Practical Aspects of NMRS

The most commonly used models in NMRS data analysis and quantification are variants of the model

$$y_n = x_n + e_n = \sum_{k=1}^{K} \left(a_k e^{i\phi_k} e^{-d_k(1-g_k+g_k t_n)t_n} e^{i2\pi f_k t_n} \right) + e_n \tag{2.0.1}$$

$$y_n \in Y, \quad x_n \in X, \quad n = 1, \ldots, N$$

i.e. each metabolite is assumed to behave like a decaying sinusoid. y_n is the n^{th} point of the noisy $N \times 1$ NMRS signal Y, whilst x_n is the n^{th} point of the *noise free* (in theory) $N \times 1$ signal X; $n = 1, 2, ..., N$, and N is the number of data points acquired from a voxel.[1] The signal acquired in NMRS experiments is a complex time series signal, and its length is always 2^u, $u \in \mathcal{N}$ (natural numbers), [50,60]; all the experimentally acquired and simulated time series signals discussed in this text are of length $2^{11} = 2048$, unless stated otherwise. The t_n terms are [generally] uniformly spaced sampling time points; $t_n = n\tau + t_0$ whereby τ is the sampling time and t_0 is the receiver dead time (the time delay before signal acquisition starts). e_n is the noise term. The subscript k, $k = 1, 2, ..., K$, refers to the k^{th} mono-component of the acquired multi-component signal; some decomposition methods do not resolve two mono-components of the same frequency but different decaying rates, whether or not this resolution occurs sometimes depends on parameter settings and sometimes on the number of available data points (This is discussed in full in *Chapter 3*). K is the maximum number of peaks extracted from the data. a_k is the amplitude and it is proportional to a compound's concentration, and the number of ^{1}H nuclei in the compound (in the case of ^{1}H NMRS). ϕ_k is the phase, which is due to RF electronics imperfections [49,60]; RF pulses tend not to interact with the magnetization vectors of a nucleus as theoretically expected. d_k is the k^{th} component signal's decay rate, it is the inverse of the transverse relaxation time constant T_2^* [26, 60]. T_2^* is affected by two magnetic field types − local magnetic fields arising from inter-molecular and intra-molecular interactions and external magnetic field inhomogeneity (as discussed in the *Spin Relaxation* section, *page*

[1] A voxel's typical size is approximately 8×10^{-6} m^3.

28). g_k is the damping factor's correction term, this value controls the form of an assumed line-shape (as discussed in the *Line-Shapes* section, *page 30*). f_k is the frequency (Hertz) of the k^{th} mono-component signal. The aim of the first part of the chapter is to try and understand the theoretical and analytical roots of the above model's, *Eq. 2.0.1*, functional form in order to later

- understand how, if at all, the most widely used analysis approaches account for or reflect the nature of the signal being analysed (w.r.t. both intrinsic and extrinsic factors).
- develop/implement an informed NMRS signal analysis approach .
- develop an informed automatic classification method.

The second part of the chapter considers the practical aspects of NMRS, especially because the NMRS signal acquired in practice differs slightly from what is expected in theory.

Part I: Theoretical Aspects of NMRS

2.1 The Fundamental Property

The fundamental property required of a particle in an NMR experiment is the nuclear spin property [1]. A nucleus, of a particle, that exhibits this property has intrinsic angular momentum I, and consequently nuclear magnetic moment μ (HzT^{-1}) properties. μ is directly proportional to I, $\mu \propto I$, and explicitly [1]

$$\mu = I\gamma \,(\mathrm{HzT}^{-1}) \tag{2.1.1}$$

The nuclear magnetic moment μ is a torque i.e. it describes a turning effect, which − in this case − is due to I. I is unitless, it is actually a number that highlights whether or not a particle has intrinsic angular momentum. Hence I is also known as the spin quantum number [61]; $I = u/2$, $u \in 0, 1, \ldots$, i.e. u is a member of the *whole number* [1]. γ is the gyromagnetic ratio (MHzT^{-1}). Each nucleus type has a different gyromagnetic ratio, e.g. the gyromagnetic ratios of ^{1}H, ^{13}C & ^{31}P are 42.58, 10.71 & 17.25 MHzT^{-1} [61] respectively; unless stated otherwise, this thesis' discussions are in terms of the ^{1}H nuclei. In the absence of an external magnetic field, nuclei rotate randomly about their axes. However, in the presence of an external magnetic field B, measured in tesla T, the allowed orientations of a nucleus depends on I. There exists ($\exists$)

$$2I + 1 \tag{2.1.2}$$

allowed orientations, and each has a different magnetic quantum number m . The magnetic quantum numbers of a nucleus are given by the Arithmetic Progression (AP)

$$m = -I, -I + 1, ..., +I \tag{2.1.3}$$

with common difference 1 [1]. For a ^{1}H nucleus $I = \frac{1}{2}$, thus it has two allowed orientations in the presence of B. The orientations correspond to parallel or anti-parallel alignment to

the magnetic flux direction; nuclei in this state are in an equilibrium state [60]. The ratio of parallel to anti-parallel nuclei depends on the magnetic field strength. Slightly more nuclei align parallel to the flux, rather than anti-parallel, because this requires less energy; magnetic resonance imaging (MRI) depends on the detection of these slight differences [36, 37]. The energies are discussed next.

2.2 The Energy Levels

The energy of interaction between an applied magnetic field B and a nucleus, is

$$E = -\mu B = -m\gamma B\,(\text{Hz}) \tag{2.2.1}$$

N.B. Energy is measured in Hertz in the field of NMRS.

and because each nucleus type has more than one magnetic quantum number m (*Expression 2.1.3*), each has a group of energies, known as Energy Levels, associated with it. The number of energy levels per nucleus, or group of nuclei of the same type, is

$$(2I+1)^{\text{Number of Nuclei}} \tag{2.2.2}$$

In the case of ^1H spectroscopy, where $I = \frac{1}{2}$, a single nucleus has two energy levels

$$E_\alpha = -\frac{1}{2}\gamma B \qquad \& \qquad E_\beta = +\frac{1}{2}\gamma B \tag{2.2.3}$$

E_α is the lower energy level (*Table 2.1a*); $\gamma = 42.58$ MHzT^{-1}, and $B \geq 0$ always. In ^{1}H spectroscopy the $m = +\frac{1}{2}$ state is labelled α, whilst the $m = -\frac{1}{2}$ state is labelled β, and they are called the *spin up* and *spin down* states respectively. The energies correspond to the energies at which measurements can be made, <u>not</u> the energies at which nuclei can exist. This is a consequence of a quantum mechanics theory which states that the measured energy from a particle will always equal one of its energy levels or rather measurements can only be taken at energy levels [60]. NMR spectroscopy measures the difference between energy levels not the energy at a level. This energy difference, known as transition energy, is the energy released when particles try to re-attain their equilibrium position (due to magnetic field B), after being bombarded with a radio frequency pulse sequence [1, 60]. In 1D NMRS, which is the focus of this thesis, the only measurable/observable transitions (between energy levels) are those whereby [60]

$$|\Delta M| = 1 \tag{2.2.4}$$

M is the sum of the magnetic quantum numbers at a level. The diagrams of *Tables 2.1a & 2.1b*, *page 14*, are the energy level diagrams for a single and two coupled ^{1}H nuclei respectively; the corresponding M and energies are detailed in *Tables 2.1c & 2.1d* respectively. For a single $\frac{1}{2}$-spin nucleus (a nucleus whose $I = \frac{1}{2}$), the transitions $\alpha \rightarrow \beta$ & $\beta \rightarrow \alpha$, which are the absorption and dispersion processes respectively, are the transitions observed in 1D NMR. Our main interest is absorption. For a single ^{1}H nucleus the transition energy is

$$\begin{aligned} \Delta E &= (-m_\beta \gamma B) - (-m_\alpha \gamma B) \\ &= (\frac{1}{2}\gamma B) - (-\frac{1}{2}\gamma B) \\ &= \gamma B(\text{Hz}) \end{aligned} \tag{2.2.5}$$

For $B > 0$ and $\gamma \neq 0$, $\Delta E \neq 0$ if and only if $I \neq 0$; this is not an issue in the case of ^{1}H NMRS because $I = \frac{1}{2}$. More critically

Theorem 1 *To detect ΔE, an applied energy, due to radio frequency pulses in the case of NMRS, must equal $-\Delta E$ i.e.*

$$\nu = -\Delta E = -\gamma B \,(\text{Hz}) \tag{2.2.6}$$

ν (Hz) is the Lamor (or resonant) Frequency due to B . In different chemical environments the Lamor frequency of the same nucleus type varies from compound to compound due to nuclear shielding [1], which is caused by electrons circulating a nucleus and creating a magnetic field opposed to an applied field. However, in numerous cases, the circulating electrons create a magnetic field aligned to the applied field, thereby increasing the magnetic field at the nucleus [50]. In either case the field at a nucleus is not equal to the applied field, instead it is

$$(1 + \delta)B \tag{2.2.7}$$

whereby δB is the nuclear shielding; the term δ is known as the **Chemical Shift** and it is either positive or negative, i.e. the field due to the circulating electrons is either aligned to or opposed to the applied field respectively. Hence *Eq. 2.2.6* becomes

$$\nu = -\gamma(1 + \delta)B \tag{2.2.8}$$

Hence

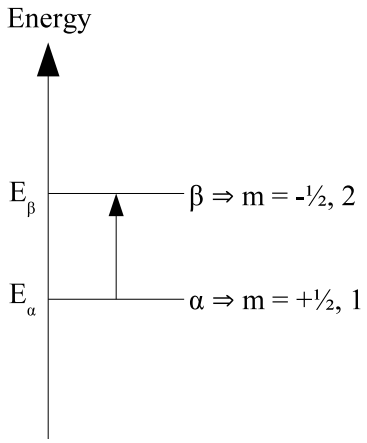
Corollary 1 *(Of Theorem 1) To detect the transition energy ΔE of a spin in NMRS, the applied energy due to radio frequency pulse must be*

$$\nu = -\gamma(1 + \delta)B(\text{Hz}) \tag{2.2.9}$$

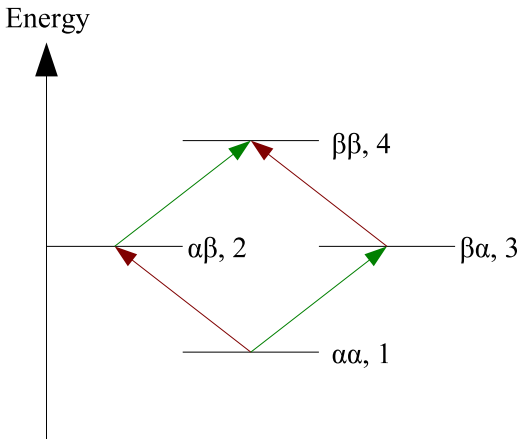
It is the chemical shift property that enables the detection of a variety of compounds, via the same nucleus type, in NMRS experiments. The key element of an NMRS experiment that exploits the δ property is the Radio Frequency (RF) pulse. The RF pulse is a radiating RF field, i.e. an energy emitting field [60]. If an RF pulse, of appropriate strength/energy, is applied to equilibrium state nuclei (induced by B) it displaces them from their equilibrium states, by transferring energy to the nuclei. The energy moves the nuclei from one energy level to another, as previously noted. When the RF pulse is switched off the nuclei start precessing, i.e. start releasing the absorbed energy, in an attempt to regain their B induced equilibrium state. However, due to their different chemical shifts δ the nuclei precess at different rates, which means they produce energy signals of varying frequency. The signal recorded in NMRS experiments, known as the Free Induction Decay (FID), is the sum of these energy signals, i.e. it is a multi-component signal. It is important to note here that the energy signals that make up the FID of NMRS experiments are due to $-\gamma\delta B$ (Eq. 2.2.9) only, because a rotating frame of reference, rotating at $-\gamma B$ Hz, is used in NMRS experiments. Also of import is the

fact that the RF field strength used in NMRS experiments must be broad enough to affect the relevant range of ^{1}H resonance frequencies present in the relevant chemicals in a Region of Interest (ROI) [60].

A few models have been developed to simplify, and aid our understanding of, the concepts that underlie nuclei behaviour in the presence of B or a RF pulse. These are discussed next.



(a)



(b)

Level	Spin States	M	Energy
1	α	$+\frac{1}{2}$	$+\frac{1}{2}\nu_1$
2	β	$-\frac{1}{2}$	$-\frac{1}{2}\nu_1$

(c)

Level	Spin States	M	Energy
1	$\alpha\alpha$	1	$+\frac{1}{2}\nu_1 + \frac{1}{2}\nu_2 + \frac{1}{4}J_{12}$
2	$\alpha\beta$	0	$+\frac{1}{2}\epsilon - \frac{1}{4}J_{12}$
3	$\beta\alpha$	0	$-\frac{1}{2}\epsilon - \frac{1}{4}J_{12}$
4	$\beta\beta$	-1	$-\frac{1}{2}\nu_1 - \frac{1}{2}\nu_2 + \frac{1}{4}J_{12}$

(d)

Table 2.1: The energy levels of ^{1}H nuclei. The diagrams and tables illustrate and detail the energy levels of a single ^{1}H nucleus (a), and two coupled ^{1}H nuclei (b). Only the observable 1D NMRS transition lines of a nucleus are displayed in (a) & (b); in (b) green for nucleus 1, and red for nucleus 2. In **(d)** the energy levels are for two strongly coupled spins, i.e. the difference between the frequencies ν_1 & ν_2 are comparable; $\epsilon = \left((\nu_1 - \nu_2)^2 + J_{12}^2\right)^{1/2}$, whereby J_{12} is the coupling constant. The limiting case $|J_{12}| << |\nu_1 - \nu_2|$, which describes weak coupling, leads to $\epsilon \approx \left((\nu_1 - \nu_2)^2\right)^{1/2} = \nu_1 - \nu_2$.

2.3 Models of Spin Behaviour during NMRS Experiments

At a simple level, the effect of a radio frequency pulse on an equilibrium state nucleus (with angular momentum I) can be understood via the Vector Model [49, 60]. A simple NMRS experiment is illustrated in *Fig. 2.1.* Let M_0 be the magnetization vector due to the effect of

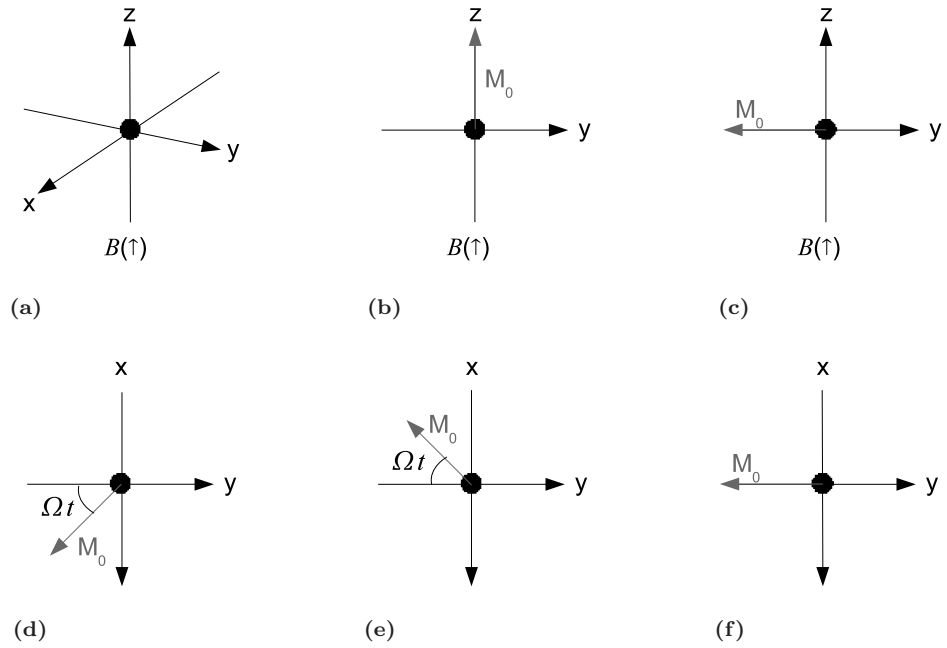


Figure 2.1: The vector model and the spin echo pulse sequence. The spin echo pulse sequence [47] is $90^\circ_x \xrightarrow{t} 180^\circ_y \xrightarrow{t}$ Acquisition. **(a)** Frame of reference. **(b)** A ^{1}H nucleus in an equilibrium state due to a field B applied along the z axis. M_0 is the resulting magnetization vector. The field at the nucleus is given by Expression 2.2.7. **(c)** Next a 90° RF pulse is applied about the x-axis (Notation 90°_x). The RF pulse is − should be − resonant with the Lamor frequency of the nucleus. **(d)** Ω is the Lamor frequency offset ($rad\,s^{-1}$) in the rotating frame. Thus the angular distance travelled by M_0 after a time t is Ωt. **(e)** The application of another RF pulse; 180°_y. **(f)** After t seconds of free precession

the external magnetic field B on a single spin (nucleus). After a 90° RF pulse applied about the x-axis (*Fig. 2.1c*) is switched off, M_0 precesses about the z-axis (*Fig. 2.1d*), i.e. in the x-y plane, as it starts the process of re-attaining its equilibrium position w.r.t. the field B.[2] The transverse (x & y) components of M_0 give the two aforementioned measures of transition energy: absorption and dispersion. If a signal is acquired at the *Fig. 2.1d* stage then the x and y components of the signal are

$$M_x(t) = M_0 \sin(\Omega t) \tag{2.3.1}$$

$$M_y(t) = -M_0 \cos(\Omega t) \tag{2.3.2}$$

[2]NMRS signals are recorded within the first moments of precession during which a magnetization vector precesses in the transverse plane only. Eventually the vector will *rise off* the transverse plane and nutate towards the z-axis.

and in the case of several uncoupled/non-interacting ^{1}H nuclei

$$M_x = M_{0,1}\sin(\Omega_1 t) + M_{0,2}\sin(\Omega_2 t) + ... + M_{0,p}\sin(\Omega_i t) \quad (2.3.3)$$

$$M_y = -M_{0,1}\cos(\Omega_1 t) - M_{0,2}\cos(\Omega_2 t) - ... - M_{0,p}\cos(\Omega_i t) \quad (2.3.4)$$

In this latter case the bulk magnetization vector $M_0 = M_{0,1} + M_{0,2} + M_{0,3} + ... + M_{0,P}$, and the Fourier transform of the signal leads to peaks at frequencies Ω_i which is converted to parts per million (p.p.m.) by the expression

$$\frac{\Omega/2\pi}{\gamma B} \times 10^6 \quad (2.3.5)$$

whereby γB is known as the transmitter frequency of the scanner. Signals decay over time, and it is assumed that they decay exponentially [1, 60] i.e.

$$\frac{dM_l}{dt} = -dM_l(t) \Rightarrow M_l(t) = \text{constant}\, e^{-dt} \quad (2.3.6)$$

whereby l is x or y, and d is the decay constant.

The vector model is useful for understanding the underlying aspects of NMRS experiments. In practice, however, spins interact with each other and their surroundings in a manner which is much more complicated than the vector model suggests. This much more complicated behaviour of nuclei within a magnetic field can be understood via the Quantum Mechanics (QM) model of NMRS.

Critical to the understanding of NMR via quantum mechanics is the Wave Function Ψ, which describes a physical system. Another important QM object is the Operator $\mathcal{L}$; an operator is a mathematical object that acts on a function, leading to a new function. In QM theory an operator represents something that can be measured e.g. energy. However, to acquire this information the operator must interact with a relevant wave function. Regardless of how $\mathcal{L}$ & Ψ interact, the level at which a measurement is taken is not known. This is dealt with by evaluating the *expectation value of the operator* $\langle\mathcal{L}\rangle$. [49, 60]

In NMR the observable energy operator is the Hamiltonian Operator $\mathcal{H}$ [49,60]. For a single nucleus with angular momentum I, and an applied magnetic field B

$$\mathcal{H} = -\gamma\delta BI(\text{Hz}) \quad (2.3.7)$$

is the chemical shift energy of the nucleus, i.e. the energy of interaction between a single nucleus, of chemical shift δ, and a magnetic field B. Single or uncoupled (non-interacting) spins only have chemical shift energy. On the other hand coupled (interacting) spins have coupling energy in addition to chemical shift energy, i.e. the total energy at a level is [1]

$$\mathcal{H} = \textit{Chemical Shift Energy + Coupling Energy} \quad (2.3.8)$$

The coupling energy is the interaction energy between any two nuclei, and it is proportional to the splitting/coupling constant J_{uw}(Hz) between a pair of nuclei u & w, and to their respective angular momenta I_u & I_w i.e.

$$\textit{Coupling Energy} = I_u I_w J_{uw} \quad (2.3.9)$$

The J_{uw}, which is due to chemical bond effects, leads to peaks that are split about a spin's resonant frequency, and the split peaks of a spin are always a constant distance J_{uw}(Hz) apart; the distance is independent of magnetic field strength. [1, 60] According to the QM theory of NMR, the behaviour of a single $\frac{1}{2}$-spin nucleus (a nucleus whose $I = \frac{1}{2}$) during free precession, when NMRS signals are recorded, can be understood via the time dependent Schrödinger equation [49]

$$\frac{d}{dt}\Psi = i\mathcal{H}\Psi \Rightarrow \Psi = e^{i\mathcal{H}t} \tag{2.3.10}$$

i.e. the quantum mechanics model of NMRS behaviour expects the system Ψ (i.e. the NMRS system) to behave exponentially. However, and as illustrated in the vector model example of *Fig. 2.1*, the free precession periods of an NMRS pulse sequence are preceded by an RF pulse application. The QM expression describing the evolution of a system w.r.t. their Hamiltonian and pulse sequence is given by the Liouville von Neumann Equation [49], which is a function of the Schrödinger equation. For example, the behaviour of the same single spin of *Eq. 2.3.10*, of angular momentum I, and w.r.t. the pulse sequence of *Fig. 2.1*, i.e.

$$90^\circ_x \xrightarrow{t} 180^\circ_y \xrightarrow{t} \text{Acquisition} \tag{2.3.11}$$

is, by the Liouville von Neumann Equation,

$$\rho(t) = e^{-i\mathcal{H}t}e^{-i\pi I_y}e^{-i\mathcal{H}t}e^{-i\frac{\pi}{2}I_x}\rho(0)e^{+i\frac{\pi}{2}I_x}e^{+i\mathcal{H}t}e^{+i\pi I_y}e^{+i\mathcal{H}t} \tag{2.3.12}$$

$$\rho(0) : \rho(0) = I_z, \text{because the field } B \text{ directed along the } z \text{ axis in } \textit{Fig. 2.1}$$

And in all cases the y and x components of the free induction decay are [49]

$$\begin{aligned} M_y &= \mathrm{Tr}\left[F_y\rho(t)\right] \\ M_x &= \mathrm{Tr}\left[F_x\rho(t)\right] \end{aligned} \tag{2.3.13}$$

Tr is the *trace* of a matrix, i.e. the sum of its diagonal elements. F_l, $l = x, y,$ or z, is the l component of the *angular momentum value or matrix* F of one or more spins [49]. In this example we are dealing with a single spin, therefore in *Eq. 2.3.13* $F_y \equiv I_y$ and $F_x \equiv I_x$. Both M_y and M_x reduce to exponential expressions, this is proved in detail in Chapters $7-10$ of [49]. Using this same approach, evolution expressions can also be deduced for coupled and uncoupled nuclei [49,60]. For example, the Liouville von Neumann expression for two strongly coupled ^{1}H nuclei, w.r.t. the pulse sequence of *Eq. 2.3.11*, is

$$\rho(t) = e^{-i\mathcal{H}t}e^{-i\pi F_y}e^{-i\mathcal{H}t}e^{-i\frac{\pi}{2}F_x}\rho(0)e^{+i\frac{\pi}{2}F_x}e^{+i\mathcal{H}t}e^{+i\pi F_y}e^{+i\mathcal{H}t} \tag{2.3.14}$$

$$\rho(0) : \rho(0) = F_z$$

In this example F is the angular momentum matrix of two interacting spins, it is evaluated by the convolution of the individual momenta of a system [49], e.g. of a metabolite; refer to [23,49] for a detailed step-by-step guide to calculating the convolution of two or more momenta. The texts by Hore et al. (Chapters $7-10$) [49], and Keeler (Chapters 3,4, & 6) [60], illustrate how complex the expressions become as the number of coupled spins increase, and as the pulse sequence becomes more complex. The important point is that − whether we are dealing with

coupled or uncoupled spins the expected underlying behaviour of the nuclei remains the same i.e. the transition energy measurements are in terms of complex exponentials. The data required for modelling, quantum mechanically, the NMRS behaviour of nuclei are

- chemical shifts
- coupling constant data
- pulse sequence parameters

The quantum theory of NMRS is considered a well-defined theory of NMRS behaviour [1, 49]. Thus quantum theory based prediction models − e.g. the product operator model [45, 60, 99], and the density matrix model [60] − have been developed in order to predict the spectral frequencies of metabolites w.r.t. a pulse sequence type. The density matrix model is basically the Liouville von Neumann Equation. The product operator model is a simplification, by linearization, of the density matrix model [60, 99]. The linearization of a complex Liouville von Neumann expression can be quite challenging, thus the product operator model could mainly handle weakly coupled systems initially [99] − the product operator model of strongly coupled systems was considered quite complicated and cumbersome − but the work of Graveron-Demilly et al. [45] extended the model to strongly coupled systems (within certain size limits).

The Liouville von Neumann Equation gives an insight into why the functional form used by almost all the analysis methods are variations of the form in *Eq. 2.0.1* i.e.

$$a_k e^{i\phi_k} e^{-d_k(1-g_k+g_k t_n)t_n} e^{i2\pi f_k t_n} \tag{2.3.15}$$

Thus the functional forms of the analysis models used in NMRS data analysis have roots in our quantum mechanical understanding of NMR behaviour. However, because the signals acquired experimentally are affected by a variety of factors, the models encountered in NMRS analysis are extended forms of the model encountered in NMR theory. Before discussing the characteristics of a signal acquired experimentally, spectra prediction based on the QM theory of NMRS is briefly discussed.

2.4 QM Based NMRS Predictions for Prior Knowledge Purposes

The model of an event or process can be used to develop data analysis algorithms and/or prediction algorithms. The preceding section introduced the QM based models of NMRS behaviour. These models have been used to develop simulation software for predicting the NMRS spectra of metabolites because such data can be used as prior knowledge information in NMRS signal analysis; as discussed in the introduction and as illustrated in [29, 88, 89]. This section discusses the software predictions, and the prior knowledge potential of the predictions.

The most common prediction tools used in human NMRS studies are NMR SCOPE [45] and GAMMA [97], and both are QM based NMRS simulation software; a few other simula-

tion software packages are listed at www.spincore.com/nmrinfo/software_s.html. The GAMMA software is not consistently supported, hence this section is focused on NMR SCOPE.[3]

The NMR SCOPE program is part of the jMRUI [74, 75, 105] NMRS data analysis package. It can simulate the behaviour of a maximum of 8 weakly coupled or 6 strongly coupled spins. This constraint can be overcome by simulating the behaviour of non-interacting moieties separately (this approach was adopted for the spectra simulation of the metabolites glutathione, glycero-phosphoryl-choline (GPC), and n-acetyl-l-aspartic glutamic acid (NAAG), whose chemical structures and couplings are detailed in *Tables F.1*, *F.2*, *F.3*, and *F.4*). To simulate a metabolite's NMRS behaviour NMR SCOPE needs the chemical shifts, coupling constants, and pulse sequence parameters data. Presently, the chemical shifts and coupling constants data are obtained empirically. Consequently, different studies do report slightly different values (e.g. compare the on-line databases of [104, 115, 116] and the records of [43]), and it is difficult to assess the accuracy of the reported values. The chemical shifts and coupling constants values used in this text are detailed in *Tables F.1* & *F.2* and *Tables F.3* & *F.4*. Mathematical theories have been developed for these constants e.g. Shoolery's rule [1] for a specific class of chemical shifts, and the Karplus equation for coupling constants [1, 54–57], but these have not been exploited in human NMR spectroscopy data analysis.[4]

NMR SCOPE predicts the multi-component time series signal of a metabolite. Thus our prior knowledge possibilities are the

- metabolite's spectrum.
- mono-component parameters, i.e. the amplitude, frequency, decay constant and phase of the metabolite's mono-components.

However the data must be used with care because there are some differences between what some of the above parameters represent in theory, and what they represent in practice. For example, the time series signal of the spectrum displayed in *Fig. 1.1* was acquired in a 1.5 tesla General Electric Signa Scanner (Western General Hospital, Edinburgh) from a General Electric Head Phantom containing 6 metabolites: choline (3 mM), creatine (10 mM), l-glutamate (12.5 mM), lactate (5 mM), n-acetly-l-aspartic acid (12.5 mM) & myoinositol (7.5 mM) (refer to *Table C.1* for more details about the phantom). The signal was acquired from a single voxel of size 20 mm × 20 mm × 15 mm using the Point Resolved Spectroscopy (PRESS) [21, 22] pulse sequence

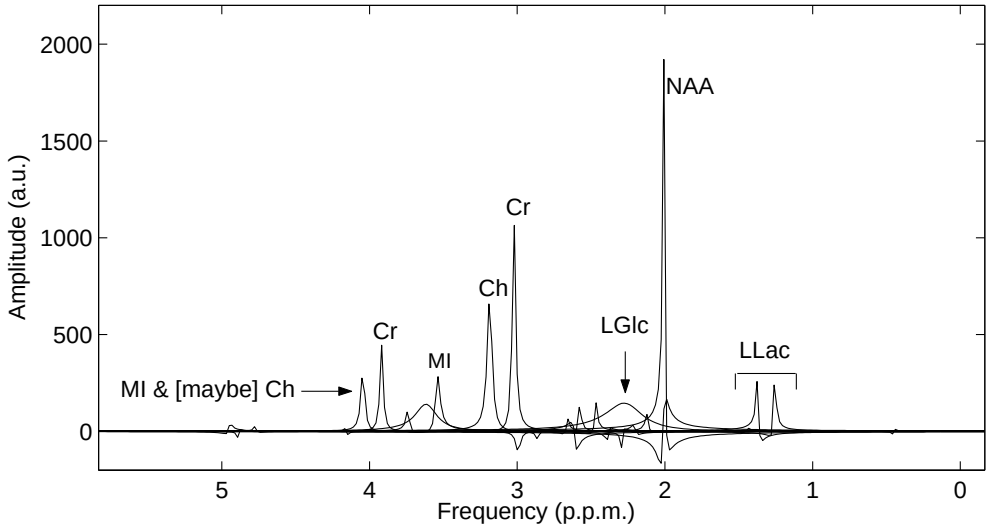
$$\underbrace{90^\circ_x}_{1.8} \xrightarrow{5.6} \underbrace{180^\circ_y}_{5.2} \xrightarrow{12.3} \underbrace{180^\circ_z}_{5.2} \xrightarrow{4.9} \text{Acquisition} \tag{2.4.1}$$

$T_R = 1500$ ms, NSA $= 128$, Red text: Time (ms)

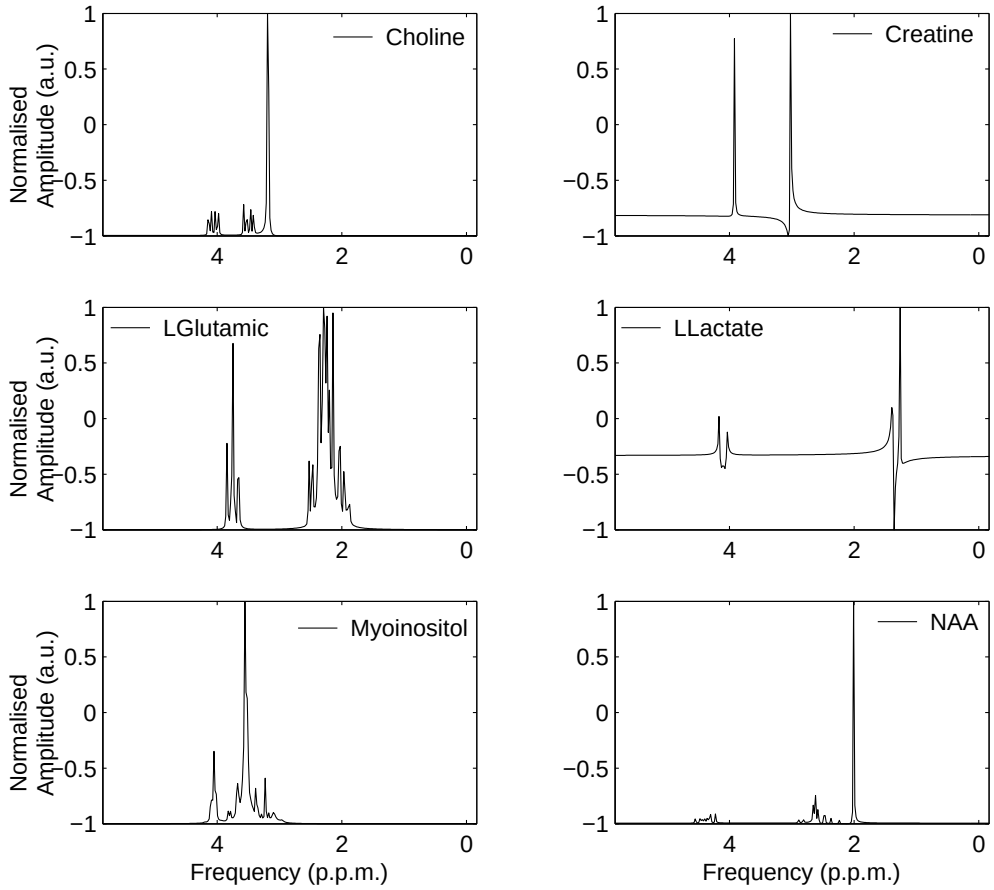
i.e. the echo time $T_E = 1.8 + 5.6 + 5.2 + 12.3 + 5.2 + 4.9 = 35$ ms. NMR SCOPE was used to simulate the spectra of the same set of metabolites also using the pulse sequence of *Eq. 2.4.1*; all the NMR Scanner and NMR SCOPE parameters used in this text are outlined in *Tables D.1* &

[3] The web pages were last updated in 2003, however a graphical user interface was developed for GAMMA by Soher et al. [98] in 2007; the team is not connected to the original command line program team of [97].

[4] In a 2007 paper by Allouche et al. [4], the authors use quantum chemistry theorems to evaluate the Hamiltonian parameters – i.e. δ & J – of the brain metabolite GABA.



(a)



(b)

Figure 2.2: Experiment based estimated components and comparable QMP spectra. **(a)** The estimated peaks of the signal (*Fig. 1.1*). **(b)** The Quantum Mechanically Predicted (QMP) spectra. The QMP spectra have not been phase corrected. This is intentional − the phase of a signal is inconsequential to our main interest − the frequencies and amplitudes of the time series signal. Hence, what we are comparing here is the spectral frequencies, not the line-shapes. Choline (Ch), Creatine (Cr), L Glutamic (LGlc), L Lactate (LLac), Myoinositol (MI), N-Acetyl-L-Aspartic Acid (NAA). Amplitude normalization method: *Eq. B.0.2, page 145*

	Ch	Cr		LGlc					LLac		MI	NAA			
f	3.186	3.023	3.919	2.289	2.119	2.386	3.737	2.013	1.367	1.259	3.542	2.005	2.666	2.472	2.577
f_p	3.183	3.027	3.920	2.287	2.133	2.369	3.744	2.035	1.373	1.265	3.544	2.008	2.649	2.475	2.573
σ	*0.002*	0.002	0.001	0.001	0.007	0.009	0.004	0.011	0.003	0.003	0.001	0.002	0.009	0.002	0.002
a	2.179	2.300	0.899	0.107	0.188	0.109	0.247	0.963	0.455	0.472	0.906	2.607	0.120	0.273	0.242
a_p	8.559	2.555	1.723	1.455	0.571	0.461	0.332	0.302	1.078	1.062	2.643	2.996	0.428	0.339	0.339
d	-7.646	-5.052	-4.943	-2.593	-5.021	-3.984	-4.823	-6.874	-3.007	-3.146	-7.658	-3.011	-3.219	-3.875	-3.501
d_p	-2.000	-2.000	-2.000	-10.522	-40.188	-4.979	-2.490	-3.675	-2.000	-2.000	-13.000	-2.000	-2.000	-2.001	-2.000
d^n		0.505	0.495	0.111	0.216	0.171	0.207	0.295	0.489	0.511		0.221	0.237	0.285	0.257
d_p^n		0.500	0.500	0.170	0.650	0.080	0.040	0.059	0.500	0.500		0.250	0.250	0.250	0.250
$\sigma^\dagger$		0.003	0.003	0.029	0.217	0.045	0.083	0.118	0.006	0.006		0.014	0.007	0.017	0.004
ϕ	0.069	-0.046	0.002	-2.543	0.429	-2.033	0.397	-1.547	0.718	-0.562	-0.265	0.383	-0.740	-0.457	-0.306
ϕ_p	0.690	0.376	-0.073	1.870	2.605	2.235	-0.188	2.265	1.767	0.296	0.086	2.135	1.968	1.015	1.338
$\lambda^{\dagger\dagger}$	1	1	2	1	2	3	6	7	1	2	1	1	3	4	4

Table 2.2: The estimated parameter values of an acquired spectrum, and of QMP metabolite spectra. f, a, d, and ϕ are the estimated frequency, amplitude, decay constant, and phase values respectively. Whilst f_p, a_p, d_p, and ϕ_p are the [*comparable*] respective QMP values. † The standard deviation between the normalized values d^n (of d) and d_p^n (of d_p). In this case, the normalized values of a vector $\boldsymbol{u}$ is given by $\boldsymbol{u}/\sum_{i=1} u_i$. †† The importance of a predicted spectral frequency f_p per metabolite, and w.r.t. its predicted amplitude a_p; 1 denotes most important, i.e. largest, amplitude. Ch: Choline, Cr: Creatine, LGlc: L Glutamic Acid, LLac: L Lactate, MI: Myoinositol, NAA: N-Acetyl-L-Aspartic Acid

D.2 respectively. In *Fig. 2.2a* the estimated component peaks of the acquired signal, w.r.t. the non-interactive analysis method HSVD [80], are displayed, and the QMP metabolite spectra are displayed below (*Fig. 2.2b*). The peaks of *Fig. 2.2a* were manually labelled using values recorded or summarized in literature, e.g. [5,43,115], or on-line databases, e.g. [104,116]. None of the QMP spectra of this text is phase corrected because this thesis' focus is the amplitude and frequency of a time series signal, not the line-shape aesthetics of the frequency spectra, which only have a bearing on certain types of analysis methods (as discussed in the previous chapter).

For quantitative comparison purposes, the *comparable* component parameter values of the acquired signal and the QMP metabolite signals are detailed in *Table 2.2*; the comparison was conducted manually, and the QMP metabolite time series signals were decomposed via FPT decomposition [8].[5]

A few remarks can be made about the *Table 2.2* & *Fig. 2.2*. First, a straightforward comparison can be made between the estimated f and predicted f_p spectral frequencies. Second, because the estimated amplitudes a of an experimentally acquired signal are proportional to concentration and the number of ^{1}H nuclei, whereas the predicted amplitudes a_p of a simulation are proportional to the number ^{1}H nuclei only, a straightforward comparison cannot be made between the amplitudes. However, the predicted amplitudes evaluated by the QM based models are proportional to the probability that the spectral frequency will exist w.r.t. the specified pulse sequence and scanner magnet strength [49, 60]. Hence, it is quite probable that the observable spectral frequencies per metabolite will be the frequencies with the highest QMP amplitudes. The λ values of *Table 2.2* seem to support this suggestion. The λ values per metabolite highlight the importance of a QMP spectral frequency f_p w.r.t. its corresponding a_p; $1 \Rightarrow$ highest, hence most important, amplitude. As an example, choline's full range of λ values, w.r.t. its simulated QMP spectrum in *Fig. 2.2*, is displayed in *Fig. 2.3*. The graph (*Fig. 2.3*) shows that choline's most important spectral frequency is approximately 3.183 p.p.m.

It is important to note that the spectral frequency with the highest a_p does not necessarily have the highest estimated amplitude a, as the results table (*Table 2.2*) suggests. This behaviour is probably due to the decay values of a metabolite's nuclei. If the decay values of a metabolite's nuclei are equal then their decay behaviour will be similar − the nucleus with the highest amplitude at time zero will still have the highest amplitude at a future time t, etc. However if the decay values are sometimes unequal, the type of behaviour illustrated in *Fig. 2.4* is possible. Unequal ^{1}H nuclei decay rates may account for the varying effects of pulse sequence timings on a metabolite's spectral pattern [3, 32]. It is difficult to compare estimated and predicted decay values in cases whereby the nuclei of a metabolite have different decay rates − the mathematical relationship is not clear. Hence decay values might be difficult prior knowledge candidates. Another set of values that will be difficult to compare are the phase values. The phase values of acquired signals are mainly due to RF electronics imperfections [49, 60], and as

[5] The FPT [8] was used to decompose all the simulated spectra used in this text because there are no subjective parameter decisions to be made. The estimable number of parameters per time series signal depends on the FPT itself. The FPT is discussed in detail in Chapter 4.

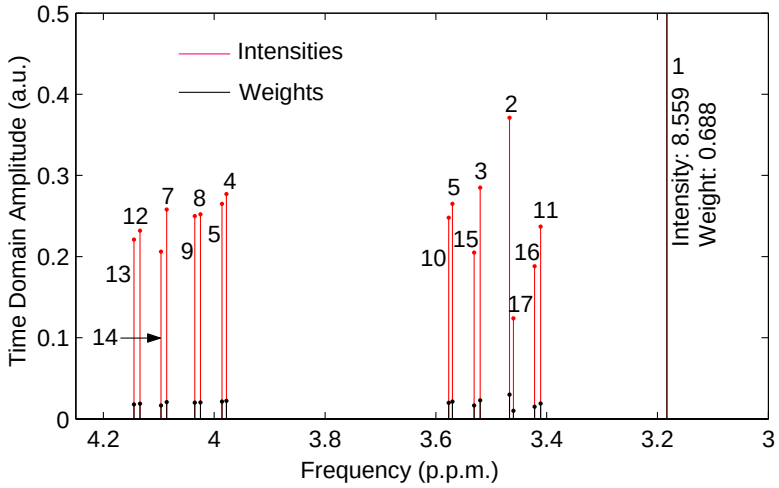


Figure 2.3: The QMP spectral frequencies of the choline spectrum of *Fig. 2.2b*. The spectral frequencies, and time domain amplitudes, were obtained by decomposing the QMP choline time series signal.

a consequence the RF pulses tend not to interact with the magnetization vectors as predicted by the QM theory of NMRS behaviour. Whereas the phase values of predicted signals are probably due to theoretical pulse sequence rotation effects on evolution.

Hence our best prior knowledge candidate, especially for the classification aim of this project, is QMP spectral frequency. Whilst the probabilistic property of the QMP amplitudes can be used as one of the probabilistic measures in a classification model.

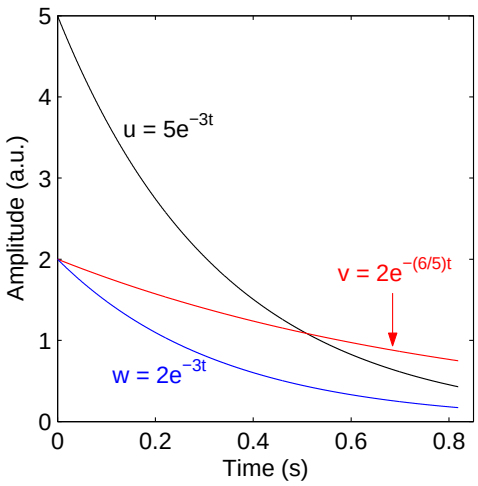


Figure 2.4: Three decaying signals. At time zero the amplitudes of signals u & v are 5 & 2 respectively, but the decay rate of u (3) is much higher than that of v (6/5). Therefore at some point in time $v > u$ is possible. In fact $v > u$ when $2e^{-(6/5)t} > 5e^{-3t}$, i.e. when $t > \frac{5}{9}\ln\left(\frac{5}{2}\right) \approx 0.5s$, as illustrated by the graph. Whereas $w \ngeq u \quad \forall \quad t$ because the converse implies that $2e^{-3t} \geq 5e^{-3t}$, i.e. $1 > \frac{5}{2}$, which is false.

2.5 Summary (Part I)

The first part of this chapter has summarized the theory and models that underpin NMRS. The quantum mechanical theory of NMRS is considered to be the most well-defined of the postulated theories [1, 49]. Consequently, most of the developed NMRS signal analysis algorithms have a functional form based on NMR quantum theory's exponential model. Additionally, the variety of simulation software developed for predicting the NMRS spectra of a chemical compound is based on the quantum theory model of NMRS. The prior knowledge potential of QMP data was also discussed, and QMP spectral frequencies were considered to be the best, especially for peak classification purposes. Next we consider the practical aspects of NMRS.

Part II: Practical Aspects of NMRS

As indicated earlier, the NMRS signal acquired in practice differs slightly from what is expected in theory. This is due to a number of factors including hardware and experimental imperfections, molecular properties and the limitations of the method (^{1}H NMRS) itself. The following sections discuss these limiting factors, and hence their effect on a signal's characteristics.

2.6 Hardware, Experimental, and Molecular Factors

2.6.1 Hardware Factors

A rather surprising number of problems are associated with hardware faults or effects (refer to [66] for a detailed discussion).

1. *Noise* is perhaps the main problem in NMRS signal analysis, and a major noise source is the coil used to detect the NMRS signal [60]. Electronic equipment (signal amplifiers, computers, etc.) and other peripheral devices are also noise sources [60,112]. Noise affects the SNR of a signal.

2. *Inhomogeneous magnetic fields* due to faulty or poorly shimmed magnets leads to poor spectra resolution, i.e. peaks with poor line-shapes and line-widths [66]. Shimming refers to "the process of optimizing the uniformity of the static magnetic field" [61] over a region.

3. *Eddy currents* are generated when certain equipment are switched on and off rather quickly during an experiment. Eddy currents affect the workings of crusher gradients − gradients used to crush signals from outside a ROI − resulting in shifted signals, which are signified by asymmetric line-shapes with negative feet. [66].

4. *RF leakage* describes the leakage of RF signals into the area where the energy signals of a ROI are recorded. If these signals are recorded along with the other energy signals that make up the FID, we end up with a pseudo-signal that might have the same spectral frequency as a genuine chemical compound [66].

Additionally hardware imperfections such as *RF coil instability*, *field inhomogeneity*, *faulty signal amplifier*, *defective Analogue-to-Digital Converter (ADC)* and *RF coil thermal noise* tend to affect an acquired signal non-linearly [52, 66].

2.6.2 Experimental Factors

The main experimental problems are due to spatial localization, and body movements or small movements within organs [66,112]. One of the most critical issues that would have a bearing on NMR spectroscopy's future in the clinical sphere is ***spatial localization*** i.e. the capacity to acquire a signal from a specific ROI. Poor spatial localization leads to signal contamination by signals outwith the ROI [32]. It also degrades the magnetic field homogeneity; consequently both the water suppression and spectral resolution are poorer [32]. There are three broad classes of localization procedures **(1)** surface coil procedures, **(2)** Single Voxel Spectroscopy (SVS) procedures, and **(3)** magnetic resonance spectroscopy imaging (MRSI) procedures. The surface coil methods are not widely used due to their poor spatial localization [5, 61]; the methods involved varying the position of a surface RF coil, in relation to the area of interest, and using simple pulse sequences.

The SVS methods use a narrow band of RF pulses and static [magnetic] field gradients to localize a ROI [61]. Before the implementation of a pulse sequence, three mutually orthogonal slice selective pulses are applied to a ROI, and a cuboid (usually) ROI is formed at the

intersection point. The aim of the various SVS pulse sequences is to acquire a signal from the ROI only, whilst eliminating signal contaminants from without. Each SVS method achieves this to varying degrees as summarized by Keevil, 2006 [61], and Barker & Lin, 2006 [5]. The most widely used ^{1}H NMRS SVS methods are Point Resolved Spectroscopy (PRESS) [21, 22], and Stimulated Echo Acquisition Mode (STEAM) [38], which was first developed by Granot [44] as Volume Excitation using Stimulated echoes (VEST) [44,61]; Image-Selected *In Vivo* Spectroscopy (ISIS) [78] is the most common ^{31}P NMRS SVS method [61]. The focus of this thesis is SVS, which does have some critical limitations and problems, as follows

1. Presently, a localization procedure that can acquire a signal from any shape/ROI as defined (by the scanner parameters), located in any region of the brain, whilst ensuring that information is neither lost from the region nor gained from without, does <u>not</u> exist [5, 61, 66]. This should be borne in mind when comparing studies conducted in different scanners.

2. Whether or not the ROI from which a signal is acquired coincides with the ROI being targetted is questionable [61], and should not be taken for granted.

3. Presently, the spatial heterogeneity of spectral patterns [5] cannot be determined, this feature is critical to some diagnostic decisions.

4. In SVS the ROI is usually much smaller than the surrounding area. If the surrounding area has a higher concentration of a nucleus of interest, the acquired signal will be susceptible to 'outer volume signal bleed' [61, 66], i.e. the acquired signal might include signals from outside the ROI. Extraneous signal suppression techniques are not considered sufficiently robust [61].

5. The extent to which an acquired signal reflects the chemistry of an area depends on the localization procedure used; each procedure type leads to varying degrees of signal loss or contamination [53, 66].

The last point refers to the properties/parameters of a localization procedure. The parameters have a bearing on the acquirable metabolite quantities, not just area localization. For example, the repetition time of a pulse sequence, if too short, may affect the acquirable metabolite quantities, and hence the estimable metabolite quantities [53]. In fact points 1,2,4, & 5 will affect the acquirable, and hence estimable, quantities of metabolites, and because our knowledge of these problems is limited, their effect is outside the control of an analysis method. In fact it is quite probable that only people familiar with expected metabolite quantities per region will notice a problem. On the other hand poor localization, as already noted, leads to poor signal resolution, which might also affect the estimable quantities of metabolites, but mainly if an inappropriate NMRS analysis method is used. For example, iterative fitting analysis methods are considered inappropriate for poor resolution data due to their reliance on line-widths [66]. A feature of spatial localization procedures that affects a spectrum's line-shapes is ***unbalanced gradient crusher pulses*** [5, 66]. Crusher gradients are used to "prevent refocusing of unwanted echoes or FIDs" [66] within the ROI and/or for eliminating signals from outside the ROI [5]. Unbalanced crusher gradients are usually signified by peaks with negative

feet [66]. Depending on the NMRS analysis method used, estimated metabolite quantities might be affected [66].

The last group of methods, the MRSI methods, address some of the problems of the prior two methods, and offers some interesting opportunities. The MRSI methods, which are quite different from the afore-discussed methods [5, 61], are multi-voxel techniques providing concurrent spatial and spectral information.

Another problem encountered in NMRS is ***physiological movements*** [53, 66], e.g. movements such as the repetitive or pulsatile motions associated with cardiac and respiratory activity. Physiological movements affect the spectral resolution of an acquired signal [66]; affected spectra are characterized by increased line-widths, frequency shifting, reduced peak area, poorer water peak suppression [66]. Another movement issue is ***subject movements***, such as head movements, such movements can lead to a misdiagnosis if the metabolic profile of the [mistakenly] scanned region is different from that of the targetted region [66]. Finally, placing a subject in an NMR scanner affects/spoils the ***shimming*** quality of the scanner's shim [30], and the poorer the shimming quality the poorer the spectral resolution of an acquired NMRS signal [66].

The 2004 paper by Kreis [66] discusses the experimental issues in detail; [5] and [32] are also good sources.

2.6.3 Biological or Molecular Factors

2.6.3.1 Water.

In ^{1}H NMRS the water signal is approximately 10^3 to 10^4 times the signal of metabolites [5, 100], hence water peak suppression, and eventual removal, is an important part of ^{1}H NMRS experiments and analysis. The original, and full, water peak is suppressed during NMRS experiments. The water peak left in the acquired NMRS signal is called the residual water peak. Poor residual water peak filtering leads to erroneous metabolite quantity estimates, because the *tail* remnants of a poorly filtered water peak are inadvertently included in the peak areas of other signals. For most analysis methods the water peak has to be filtered out prior to analysis. There are a few filtering algorithms available [69, 100]; the HLSVD filtering algorithm of jMRUI [74, 75, 105], the NMR data analysis package, was used for all the experimentally acquired signals reported in this thesis.

2.6.3.2 Macromolecules & Lipids.

Macromolecules, and sometimes lipids, are – relatively – high molecular weight compounds [53], and their decay rates in ^{1}H NMRS are high relative to those of metabolites [91]. Due to the characteristics of the Fourier Transform, the energies of the fast decaying components are spread out in a Fourier spectrum. This, in part, accounts for the large baselines that tend to characterize the FT spectra of brain NMRS signals, and it highlights a disadvantage of using Fourier based NMRS analysis methods. Researchers have tried to use a variety of methods to suppress/eliminate the baseline effects of macromolecules and lipids, e.g. truncation

of the first few points of a signal [88, 89], wavelets [93, 94], experimental methods [67, 92], baseline correction [30], etc. Baseline correction is a frequency domain, post-processing method. It ensures that the baseline of an estimated spectrum is approximately at zero. It involves deducting an average from each frequency domain value

$$F(\omega) = F(\omega) - Average \tag{2.6.1}$$

whereby ω is the frequency ($rad\,s^{-1}$), and $F(\omega)$ is the corresponding amplitude in the frequency domain. There are a few options for evaluating the average, e.g. the mean of the first n points, the mean of the first and last $\frac{n}{2}$ points, or (and most commonly) a non-linear/polynomial average. The values deducted tend to be quite small and relatively peripheral, however this method should be used carefully to avoid loss of information.

2.6.3.3 Spin Relaxation.

As already discussed in *Section 2.2*, after the application of a RF pulse, each nucleus tries to re-attain its B induced equilibrium state. This process is generally known as spin relaxation. Spin relaxation time constants are widely exploited in NMRS experiments, and there are two different types: longitudinal and transverse relaxation. Referring back to *Fig. 2.1*, as the transverse components, M_x & M_y, of the magnetization vector decreases, its longitudinal component M_z increases i.e. it progressively realigns with the +z axis in its quest to re-attain its B induced equilibrium state. This latter process is known as longitudinal relaxation. Nuclei achieve longitudinal relaxation by loosing their excess energy − the excess energy is due to the RF pulse − to their surroundings. The longitudinal relaxation time [constant] T_1 depends on the strength of the applied magnetic field and the internal motion of particles. Prior to scientists theoretical understandings of nuclear spin behaviour during NMRS experiments, it was assumed that $M_z \rightarrow M_0$ exponentially

$$\frac{dM_z}{dt} = \frac{1}{T_1}(M_0 - M_z) \tag{2.6.2}$$

in a static magnetic field applied along the +z axis. The hypothesis is based on Bloch theory [26]. Present knowledge of NMRS spin behaviour of $\frac{1}{2}$-spin nuclei suggests that this is an accurate model of longitudinal relaxation in most cases [26].

Initially after applying a RF pulse, the nuclear spins precess in phase in the transverse plane (Section 2.3). However, due to local magnetic field variations and external magnetic field inhomogeneity, the individual magnetization vectors start precessing at different rates, this leads to a progressive loss of phase coherence. Consequently, the vector sum M_{xy} of the transverse components of the individual magnetization vectors decreases; instead of equivalent vector values different negative and positive values are being summed. The local magnetic field variations are due to intra-molecular and inter-molecular interactions. It involves the exchange of energy between spins, rather than energy dissipation, and this is the natural transverse relaxation process. Its de-phasing time [constant] T_2 is more or less independent of the strength of B [112]. The second factor affecting M_{xy} is external magnetic field inhomogeneity, which is due to hardware (magnet) flaws. This inhomogeneity also affects the phase coherence of the individual magnetization vectors. Let T_{B_0} be the de-phasing time [constant] due to B

inhomogeneity then the relaxation time constant due to both de-phasing time constants is [26]

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_{B_0}} \tag{2.6.3}$$

A measured FID signal represents a decreasing/decaying M_{xy}, and the decay rate d of each component of the FID is the inverse of the de-phasing time constant. Naturally the desired de-phasing time constant is $T_2^* = T_2$, such that the evaluated decay rates are independent of field inhomogeneity effects. However this would not be achieved, especially for small molecules, if the magnet shimming (a common physical method for eliminating magnetic field inhomogeneity) and/or the inhomogeneity processing method (e.g. QUALITY [31]: a time domain deconvolution method that removes inhomogeneity effects by converting all line-shapes to Lorentzian line-shapes) is poor.[6] Whether a molecule's transverse relaxation process is dominated by local magnetic field or B inhomogeneity effects, or neither, seems to depend on the size of the molecule [26]. According to Claridge [26], B inhomogeneity effects dominate the transverse magnetization decay rate of small molecules. Thus, effective elimination of B inhomogeneity effects is critical to the quantification of small metabolite compounds. An observable effect of B inhomogeneities is line-shape broadening [60]. This line-shape broadening is almost always modelled as an exponentially decaying effect even though the peculiar line-shapes that result from field inhomogeneities suggests this is a flawed assumption [48, 60].

2.6.4 Section Summary

Three main factors affect the acquirable and/or estimable metabolite signals: hardware, experimental, and biological or molecular factors. They mainly affect the SNR, spectral resolution, and line-shapes of a signal. The next section discusses these characteristics.

2.7 Signal Characteristics

In the previous section a number of factors that affect the nature of an acquired NMRS signal were discussed. It was observed that they affect three key signal characteristics: SNR, spectral resolution, and line-shapes.

2.7.1 SNR & Spectral Resolution

The time domain SNR is defined as [66]

$$20 \log_{10} \left(\frac{Amp.}{\sigma_{res.}} \right) \tag{2.7.1}$$

$Amp.$: the amplitude of a signal at time 0

σ_{res} : the standard deviation of the noise residue

[6] In order to address magnetic field inhomogeneity effects in MRSI Khalidov et al., 2006 [63], suggested a constrained reconstruction technique. This method has been used before in NMRS for spectral localization purposes [51]. A good background description of the method can be found in [46].

This is a standard NMRS definition [66]. The spectral resolution is a measure of the extent to which individual peaks can be differentiated. Hence spectral resolution is a function of line-width because as the line-widths of a spectrum's peaks increase, the more the peaks overlap, thereby reducing the spectral resolution. The quality of both the SNR and spectral resolution can be *improved* via a range of *controllable* experiment related parameters: voxel size, receiver bandwidth, number of acquisitions, scan parameters (echo time, repetition time, flip angle), magnetic field strength, shimming, and coil options [5, 30, 60, 66, 112].

The SNR of a signal increases as the voxel size is increased, i.e. SNR is directly proportional to voxel size, however voxel size is inversely proportional to spectral resolution. The latter happens because the degree of magnetic field inhomogeneity within a voxel increases as the voxel size is increased, and as discussed in the previous section poor magnetic field homogeneity is characterized by poor spectral resolution. An option is to increase the number scans per voxel or to increase the repetition time (thus increasing the scanning time), however both are expensive in relation to time and costs. The receiver bandwidth refers to the range of frequencies acquired during frequency encoding, and it is inversely proportional to SNR, i.e. as the receiver bandwidth is reduced, the SNR increases. SNR also improves with increasing field strength B. Unfortunately, as B increases the line-widths of a peak also increase leading to decreased peak heights [5], and consequently data is lost to the baseline area. The SNR also improves with increasing metabolite concentration.

An analytical solution that is sometimes used for enhancing the SNR or spectral resolution of a signal is apodisation. Apodisation refers to the mathematical method for dealing with sharp discontinuities. In MRS apodisation is used for sensitivity (SNR) and spectral resolution enhancement [30]. In each case the time series signal is multiplied by a weighting function. In the case of SNR enhancement, the time series signal is multiplied by a decaying exponential function, $e^{-(constant)t}$, to reduce the contribution of the virtually constant noise [30, 60]. However if an inappropriate decaying weighting function is used, the improved SNR will be at the expense of spectral resolution. To address this, the matched filter method [30, 60] is usually used for evaluating the optimal SNR weighting function . For resolution enhancement the Lorentz-Gauss apodisation method is used. The time series signal is multiplied by a composite weighting function, which consists of a rising and decaying exponential; $e^{\rho t}e^{-\sigma^2 t^2}$, where $\rho \geq 0$ & $\sigma \geq 0$ [30, 60].

All in all, for optimal SNR and spectral resolution the voxel size has to be chosen carefully, and any field inhomogeneities must be dealt with effectively [5, 112]. Additionally, analytical techniques can be used to enhance the SNR or spectral resolution of a signal.

2.7.2 Line-Shapes

NMRS analysis methods that specify line-shape forms usually assume that the NMR spectra line-shapes are either Lorentzian, Gaussian or Voigt line-shapes [48, 109]. Mathematically, if $g_k = 0$ in *Eq. 2.0.1* the line-shape is Lorentzian, if $g_k = 1$ the line-shape is Gaussian, and if $g_k \in (0 \quad 1)$ the line-shape is a Voigt line-shape [48]; refer to *Fig. 2.5* for an illustrated example. Theoretically, the Gaussian line-shape is due to the Doppler effect, i.e. the broadening of spectral lines due to the temperature dependent random thermal motion of particles. The

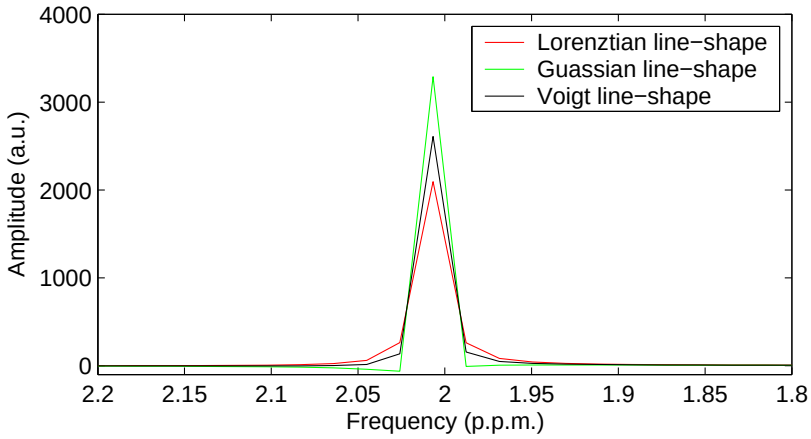


Figure 2.5: An illustration of the line-shape types. Each of the line-shapes is a Fourier transform of a time series of amplitude $a_k = 2.681$, sampling time $\tau = 0.0004$ s, decay rate $d_k = 2.855$ Hz, frequency $f_k = 2.008$ p.p.m. (128.25096 Hz w.r.t a scanner of strength 63.87 MHzT^{-1}, and by the conversion equation *Eq. 2.3.5*), and length $N = 2048$ points. The latter implies that each time point $t_n = n\tau$ whereby $n = 1, 2, \ldots, N$. Each time series was evaluated using *Eq. 2.0.1*; without noise. For the Lorentzian line-shape $g_k = 0$, for the Gaussian line-shape $g_k = 1$, and for the illustrated Voigt line-shape $g_k = 0.5$.

Lorentzian line-shape is due to colliding particles, which leads to pressure broadening of spectral lines [64]. A Voigt line results if a spectral line is broadened by both the Gaussian and Lorentzian [48] line broadening systems, i.e. by independent superimposition of both broadening systems. A factor that affects the line-shape form, as previously discussed, is magnetic field inhomogeneity, and the resulting line-shape form is not clear. The effects of inhomogeneity can be improved via better shimming [66] or via a range of analytical methods [82], e.g. the aforementioned QUALITY (*page 29*) that converts all line-shapes to Lorentzian line-shapes [31]. The same range of analytical methods can be used to counter the effect of eddy currents on line-shapes; eddy current effects are signified by asymmetric line-shapes [66]. In general, line-shape assumptions by an NMRS analysis method might lead to quantification errors due to model mismatching (*page 1.1*). Perhaps the best way to avoid line-shape related errors is to use an analysis method that does not assume the form of a component's line-shape.

2.7.3 Section Summary

The negative effect of the range of factors discussed in the previous section, *Section 2.6*, on a signal's characteristics can be reduced by a number of experimental techniques or analytical methods. Considering the fact that the methods mainly reduce rather than eliminate these negative effects, an NMRS signal analysis method must take these issues into account. *Table 2.3* summarizes the main features addressed, and how they are addressed.

Feature	**Contributing Factors**	**Analytical & Experimental Solutions**
Poor SNR	Small voxel size Low magnetic field scanner Hardware induced noise	Apodisation & Matched Filtering [30, 60], Increased scanning time [66, 112], Larger voxel size [112], Reduced receiver bandwidth [112], A scanner of higher magnetic field strength [5]
Poor Spectra resolution, Large Line-width	Poor spatial localization Inhomogeneous magnetic field Large voxel size	Lorentz-Gauss apodisation [30, 60], Smaller voxel size [112]
Unknown or uncharacteristic Line-shapes	Molecular activity Inhomogeneous magnetic field Unbalanced crusher gradients Eddy currents	Re-modelling of line-shapes into a known or analytical form [31, 82] An analysis method that does not assume the line-shape form
Large Baseline	Macromolecules Lipids	Baseline correction [30], Truncation (of the first few points) [88, 89]
Water†		SVD Based Filters [69], Finite Impulse Response Filter [100], Experimental water suppression [5]
Non-linearities	RF coil instability Field inhomogeneity Faulty signal amplifier Defective ADC RF coil thermal noise	A signal analysis method that does not assume the linearity or non-linearity of a signal

Table 2.3: Potential problems with acquired NMRS signals. † The residual water peak filtering that is conducted prior to NMRS signal analysis is usually in addition to experimental water suppression. The residual water peak of all the experimentally acquired signals in this text were filtered out using the Singular Value Decomposition (SVD) based Hankel-Lanczos Singular Value Decomposition (HLSVD) method [69], of jMRUI [74,75]. [Other filter options are the Finite Impulse Response (FIR) Filter [100], implemented in the AQSES [81] NMRS analysis application, and the three other SVD based methods discussed in [69]: Orthogonal Triangular Decomposition (QR), Partial Re-orthogonalization (PRO), and Implicitly Restarted Lanczos (IRL).]

2.8 Summary (Part II)

The second part of this chapter has summarized a range of hardware, experimental, and molecular or biological factors that affect the main signal characteristics − SNR, spectral resolution (line-width), line-shape, baseline − of a signal. The relationships between the main signal characteristics and a range of experiment variables, which can improve the signal characteristics, were also discussed. The next chapter reviews the main NMRS analysis methods in light of what has been discussed thus far.

Chapter 3

A Review of Analysis Methods

The previous chapters have discussed the theoretical and experimental aspects of NMRS, and the reasons why the experimentally acquired signals might differ from theoretical predictions. Part of the difference can be attributed to hardware factors, to experimental factors, and to biological or molecular factors. The effect of these factors on an acquired signal can sometimes be observed via the SNR, spectral resolution, line-shape and baseline characteristics of a signal. Therefore, these characteristics should be considered when developing an analysis method. This chapter discusses the most widely used analysis methods in relation to how each addresses any of the previous chapters' problems; it also discusses an additional method, the Fast Padé Transform [8], which is rarely used (almost all the journal papers concerned with the application of the FPT to NMRS signal analysis have been published by Belkić & Belkić, e.g. [7, 13, 17–19]; the only exceptions are the 2006 and 2003 papers by Williamson et al. [113, 114]). The chapter consists of two sections. The first section discusses the methods, which have been broken down into two broad groups: interactive and non-interactive. Within either group the analysis methods can be grouped further (*Table 3.1*), e.g. frequency or time domain, fitting or non-fitting, parametric fitting or non-parametric fitting, etc. Instead of further groupings the characteristics of a method are highlighted as the discussion proceeds. An implicit aim of the first section is to identify areas of strength and weakness in order to identify areas of possible algorithm improvement, or to identify important considerations for any new analysis method. The second section is a comparative analysis of the discussed non-interactive methods. The aim of the second section is to compare the optimal set-ups of the non-interactive methods in order to understand how, why, and the extent to which they fail as noise levels increase, and to assess how consistent their analysis is w.r.t. noisy signals of similar noise level.

Definitions of terms used

Definition 3 (Basis Set) *A basis set is a set of metabolite spectra whereby the spectrum of each metabolite has been deduced experimentally or quantum mechanically under the same conditions that would be used for scanning a prospective patient or subject.*

Methods	NL FF	Interactive	PK Type	Fitting Method	Iterative FM	Fitting Algorithm	NL FM	Parametric FM	Fit Form	Analysis Domain	References
HSVD	✓	✗		✓	✗	SVD	✗	✓	Peaks	T	[80]
HLSVD	✓	✗		✓	✗	Lanczos SVD	✗	✓	Peaks	T	[69, 80]
HTLS	✓	✗		✓	✗	SVD	✗	✓	Peaks	T	[107]
HLTLS	✓	✗		✓	✗	Lanczos SVD	✗	✓	Peaks	T	[107] & [108]
AMARES	✓	✓	Peak†	✓	✓	NL2SOL	✓	✓	Peaks	T	[110]
VARPRO	✓	✓	Peak†	✓	✓	VARPRO & LM	✓	✓	Peaks	T	[106]
AQSES	*	✓	Basis Set	✓	✓	AV, NL2SOL	✓	✓	Spectra	T	[81]
LC Model	*	✓	Basis Set	✓	✓	LM	✓	✓	Spectra	F	[84]
QUEST	*	✓	Basis Set	✓	✓	VARPRO & LM	✓	✓	Spectra	F	[88, 89]

Table 3.1: The most common NMRS signal analysis methods. *The functional forms of AQSES, LC Model, and QUEST depend on the spectra in their basis set. †The prior knowledge of AMARES and VARPRO is per mono-component, i.e. per peak consisting of a single decaying sinusoid only. **General Abbreviations & Acronyms** → NL: Non-Linear, FF: Functional Form, PK: Prior Knowledge, MP: Mono-component Parameters, FM: Fitting Method, T: Time, F: Frequency. **Fitting Algorithms** → SVD: Singular Value Decomposition [42], VARPRO: Variable Projection [41, 58], AV: Adapted VARPRO [96], LM: Levenberg Marquardt [72], NL2SOL: Non-Linear Least Squares [34].

Definition 4 (Mono-component) *In this text a mono-component, of a time series signal, is a signal formed by a single set of signal parameters, i.e. frequency, amplitude, decay, and phase parameters.*

Definition 5 (Mono-component Parameters) *The mono-component parameters are a mono-component's frequency, amplitude, decay, and phase parameters, and also the parameters that describe the relationship between any number of mono-components.*

Definition 6 (Component) *Sometimes, a decomposition method may not be able to resolve the overlapping mono-components – i.e. mono-components of the same frequency but different decay rates – of a multi-component signal. Hence some of the decomposition method's estimated mono-components may not [strictly] be mono-components. In order to highlight such possibilities, the term component rather than mono-component is used to refer to the estimated mono-components of relevant decomposition methods.*

Definition 7 (Multi-component Signal) *In this text, a multi-component signal is a time series signal consisting of two or more mono-component time series signals.*

3.1 Non-Interactive Methods

Non-interactive methods are methods which cannot use prior knowledge. Most of the methods are not entirely non-interactive. For example, within the jMRUI [74,75] application a user may change four default parameter values of the SVD based methods.[1] The four parameters affect the structure of the vectors or matrices that are directly, or indirectly, used to estimate the mono-component parameters of an NMRS signal. In this text, the four parameters are termed *structural parameters*, in order to differentiate them from the mono-component parameters, and the parameters are discussed in context as the SVD methods discussion proceeds.

This section focuses on the Singular Value Decomposition (SVD) based analysis methods [69, 80, 107, 108] and the rarely used Fast Padé Transform (FPT) [8]. Both are discussed in detail because: **(1)** there are various forms of the FPT and therefore it is important to outline the FPT approach used; **(2)** problem points can be identified clearly; and **(3)** a core objective of this project is the automatic classification of estimated components.

3.1.1 Singular Value Decomposition Methods

Several SVD methods have been developed for NMRS data analysis [82,109]. The most widely used of these are Hankel Singular Value Decomposition (HSVD) [80], Hankel-Lanczos Singular Value Decomposition (HLSVD) [69,80], Hankel Total Least Squares (HTLS) [107], and Hankel-Lanczos Total Least Squares (HLTLS) [107] & [108]. These methods use the same model function

$$y_n = x_n + e_n = \sum_{k=1}^{K} a_k e^{i\phi_k} e^{-d_k t_n} e^{i2\pi f_k t_n} + e_n \tag{3.1.1}$$

[1] At present there is no systematic method for determining the values of the four parameters.

i.e. they all assume that the mono-components of an acquired NMRS signal have Lorentzian line-shapes. As previously defined within *Chapter 2*, *page 10*, y_n is the n^{th} data point of the noisy $N \times 1$ NMRS signal Y, x_n is the n^{th} data point of the *noise free* (in theory) $N \times 1$ time series signal X, and parameters a_k, ϕ_k, d_k, and f_k are the amplitude, phase, decay constant, and frequency of the k^{th} component of an NMRS multi-component signal. e_n is the noise term. The analysis steps are as follows

Step 1

Build a *user defined* $L \times M$ Hankel matrix $\mathcal{H}$, such that $N = L + M - 1$, $L > K$, and $M > K$. In the HSVD paper by Pijnappel et al., 1992 [80], the authors quote the range $0.5 \leq \frac{L}{M} \leq 2$ as the optimal range for the row L and column M ratio. The Hankel matrix is of the form

$$\mathcal{H}_{L\times M} = \begin{pmatrix} y_1 & y_2 & \cdots & y_M \\ y_2 & y_3 & & y_{M+1} \\ \vdots & \vdots & & \vdots \\ y_L & y_{L+1} & \cdots & y_{L+M-1} \end{pmatrix} \tag{3.1.2}$$

L, the number of rows of the Hankel matrix, is one of the four structural parameters of the SVD based methods. The default value of L in jMRUI [74, 75] is $N/2$, i.e. $M = N - N/2 + 1$. This is a sound choice. As previously noted, in NMRS the number of acquired points is always 2^u, $u \in \mathcal{N}$ (natural numbers), i.e. even [50, 60]. Therefore by the constraint $N = L + M - 1$, if N is always even, then $L + M$ is always odd, and hence the difference between L and M is at least 1. The latter implies that we are closest to a square matrix when $|L - M| = 1$. This occurs when either L or M is $N/2$, which means the other is $N/2 + 1$. Therefore the maximum possible rank of $\mathcal{H}$ is $\min(N/2, N/2 + 1) = N/2$, which is the same as the default value of L in jMRUI [74, 75].[2] Hence jMRUI's default value for L is an optimal setting w.r.t. obtaining the maximum number of linearly independent vectors of $\mathcal{H}$.

Step 2

Evaluate the SVD of $\mathcal{H}$

$$\mathcal{H}_{L\times M} = U_{L\times L}\Sigma_{L\times M}V^H_{M\times M} \tag{3.1.3}$$

The $L \times L$ matrix U is the matrix of normalized eigenvectors of $\mathcal{H}\mathcal{H}^T$, whilst the $M \times M$ matrix V is the matrix of normalized eigenvectors of $\mathcal{H}^T\mathcal{H}$; $U^HU = UU^H = I, V^HV = VV^H = I$, superscript $H \Rightarrow$ Hermitian conjugate , $I \Rightarrow$ identity matrix. Σ is the diagonal matrix of the square roots of the eigenvalues of $\mathcal{H}^T\mathcal{H}$, in decreasing order. The HSVD & HTLS methods evaluate the SVD via the normal/general SVD algorithm [77]. However the normal SVD algorithm is computationally expensive [69]. The HLSVD & HLTLS methods address this issue by using a Lanczos based algorithm [69] to evaluate the SVD. Lanczos algorithms evaluate the eigenvalues and eigenvectors of a matrix by taking advantage of the matrix's

[2] The rank of a matrix is "the maximum number of linearly independent row or column vectors of a matrix" Section 6.4 & 6.6 of [68].

structure [69]; such structure exploitation usually gives Lanczos based algorithms a *computation time* advantage. The only difference between HSVD & HLSVD, and between HTLS & HLTLS, is this SVD evaluation step.

Step 3

Truncate the matrices U, Σ, and V w.r.t. a *user defined* value of K ($K \in \mathcal{N}$)

$$\mathcal{H}_{L\times M} = U_{L\times L}\Sigma_{L\times M}V^H_{M\times M} \rightarrow \mathcal{H}_{L\times M} = U_{L\times K}\Sigma_{K\times K}V^H_{K\times M} \tag{3.1.4}$$

Here the user assumes that there are K mono-components or components in the acquired multi-component signal. K is also one of the four structural parameters. The default jMRUI [74, 75] setting is $K = 15$, but the maximum number of mono-components/components determinable per method is 50. Hence $K = 50$ will be considered the optimal K setting w.r.t. jMRUI's SVD based analysis.

Step 4

Let U_K denote the truncated U matrix $U_{L\times K}$, let $\underline{U}_K$ denote the partial U_K matrix whose last row has been deleted, and let $\overline{U}_K$ denote the partial U_K matrix whose first row has been deleted.

HSVD/HLSVD: Evaluate the eigenvalues of

$$\underline{U}^{\dagger}_K\overline{U}_K \qquad \left[\underline{U}^{\dagger}_K = \left(\underline{U}^T_K\underline{U}_K\right)^{-1}\underline{U}_K\right] \tag{3.1.5}$$

whereby † denotes pseudo-inverse.

HTLS/HLTLS: Evaluate the eigenvalues of

$$-\mathcal{B}_{12}\mathcal{B}^{-1}_{22} \tag{3.1.6}$$

whereby

$$\left[\underline{U}_K\overline{U}_K\right]_{L-1\times 2K} = \mathcal{A}_{L-1\times L-1}\mathcal{D}\mathcal{B}^H_{2K\times 2K}, \qquad \mathcal{B} = \begin{pmatrix} \mathcal{B}_{11} & \mathcal{B}_{12} \\ \mathcal{B}_{21} & \mathcal{B}_{22} \end{pmatrix}$$

i.e. *Expression 3.1.6* is determined via the Singular Value Decomposition (SVD) of $\underline{U}_K\overline{U}_K$.

Step 5

Let $\hat{\boldsymbol{z}}$ be the $K \times 1$ vector of estimated eigenvalues of either *Eq. 3.1.5 or 3.1.6*. The decay constant and frequency values are related to $\hat{\boldsymbol{z}}$ by the expression (refer to [80, 107] for more details)

$$\hat{z}_k = e^{(-d_k+i2\pi f_k)\tau} \tag{3.1.7}$$

$$\Longrightarrow$$

$$\hat{z}_k = e^{-d_k\tau}\left(\cos(2\pi f_k\tau) + i\sin(2\pi f_k\tau)\right) \tag{3.1.8}$$

$\therefore$

$$\hat{f}_k = \frac{1}{2\pi\tau} \arctan\left(\frac{\Im(\hat{z}_k)}{\Re(\hat{z}_k)}\right) \tag{3.1.9}$$

and

$$\hat{d}_k = -\frac{1}{2\tau} \ln\left(\Re(\hat{z}_k)^2 + \Im(\hat{z}_k)^2\right) \tag{3.1.10}$$

$\hat{f}_k$ and $\hat{d}_k$ imply estimated frequency and decay constant values respectively, and τ is the sampling time.

Step 6

Express the model function, *Eq. 3.1.1*, in terms of the estimated values i.e.

$$y_n = \hat{x}_n + e_n = \sum_{k=1}^{K} a_k e^{i\phi_k} e^{-\hat{d}_k t_n} e^{i2\pi\hat{f}_k t_n} + e_n \tag{3.1.11}$$

Let $c_k = a_k e^{i\phi_k}$; similarly $\boldsymbol{c} = \boldsymbol{a} \cdot e^{i\boldsymbol{\phi}}$. Estimate each c_k by minimizing the cost function

$$\text{SSE} = \sum_{n=0}^{N-1} \left| x_n - \left(\sum_{k=1}^{K} c_k e^{-\hat{d}_k t_n + i2\pi\hat{f}_k t_n} \right) \right|^2 \tag{3.1.12}$$

i.e. by minimizing the Sum of the Squared Errors (SSE); the SSE is discussed in the *Minimization Algorithms* section (*page 141*) of *Appendix A*. In matrix form, Eq. 3.1.12 becomes

$$\text{SSE} = \|X - \Gamma\mathbf{c}\| \tag{3.1.13}$$

whereby

$$\Gamma = e^{T\left(-\hat{D}+2\pi\hat{F}\right)}$$

T is the $N \times 1$ vector of time points, i.e. $T = \{t_n\}, n = 1, \ldots, N$. $\hat{D}$ is the $K \times 1$ vector of estimated decay constants, i.e. $\hat{D}_{K\times 1} = \{\hat{d}_k\}$. $\hat{F}$ is the $K \times 1$ vector of estimated frequencies, i.e. $\hat{F}_{K\times 1} = \{\hat{f}_k\}$.

The cost function is minimized by setting SSE = 0, i.e.

$$\hat{\mathbf{c}} = \Gamma^{\dagger} X \tag{3.1.14}$$

Step 7

Evaluate the a_k and ϕ_k estimates from

$$\hat{c}_k = a_k e^{i\phi_k} = a_k(\cos(\phi_k) + i\sin(\phi_k)) \tag{3.1.15}$$

$\Longrightarrow$

$$\hat{\phi}_k = \arctan\left(\frac{\Im(\hat{c}_k)}{\Re(\hat{c}_k)}\right) \tag{3.1.16}$$

$$\hat{a}_k = \left(\Im(\hat{c}_k)^2 + \Re(\hat{c}_k)^2\right)^{1/2} \tag{3.1.17}$$

Comments

The steps involved in estimating the parameters a_k, ϕ_k, d_k, and f_k, per method, give an insight why each method is subjective and non-unique. Foremost, the structure of the Hankel matrix $\mathcal{H}$ of *Step 2* depends on the L or M value specified by the user; hence the subjectivity of the methods. As the value of L or M is varied, the structure of $\mathcal{H}$ also varies, subsequently leading to different estimated parameter values per specific value K; hence the non-uniqueness of the methods. Secondly, the value K is also specified by the user; contributing to the subjectivity of the methods, and the estimated parameter values change as K is changed.

It is assumed that the SVD and truncation, *Steps 2 & 3*, will relegate the noise components of a signal to the truncated end of the SVD matrices, i.e. separate signal and noise subspaces. This is problematic because the quality of the signal subspace varies with the Hankel matrix, i.e. it is quite easy to obtain an erroneous signal subspace. Pijnappel et al., 1992 [80], quote the range $0.5 \leq \frac{L}{M} \leq 2$ as the optimal range for the row and column values; $N = L + M - 1$ and $L \ \& \ M > K$. Secondly, this theory only holds if the SNR of the signal is not too low [69].

Because the estimated parameter values change with the defined Hankel matrix and K, the decompositions are non-unique. Basically, a signal can be decomposed into any number of plausible mono-components or components, and there is no systematic way of selecting the most probable decomposition. Another issue of import is the fact that the user has to manually discern the metabolite associations of each estimated mono-component/component.

Thus far two of the four structural parameters − the number of rows of the Hankel matrix L (*Step 1*) and the assumed number of components K (*Step 2*) − have been mentioned. The other two are (1) the number of data points $N_p \leq N$, and (2) a time series shift structural parameter τ_Δ known as number of truncation points (unfortunately). A user may opt to use only the first N_p points of a noisy time series signal Y of length N − this reduces the analysis time, but it is ill-advised because if the NMRS signal has not fully decayed, cutting-off the last few points affects the signal analysis [30, 60]. Hence the optimal setting of N_p is $N_p = N$.

The time series shift structural parameter is quite interesting. It allows a user to shift a time series back in time by τ_Δ points, perform the required analysis on the new signal, and then re-adjust (shift back) accordingly. The aim is to reduce the background or baseline during the analysis. However in contrast to the truncation pre-processing method of [88, 89], which permanently eliminates part of the baseline, the baseline absent during the analysis appears as part of the residue (jMRUI Application Notes [74, 75]). Note that the number of points does not change, i.e. the signal is extended by the same number of points τ_Δ, whereas in the case of the truncation pre-processing method the number of points is reduced [88, 89]. The optimal setting of τ_Δ is 0 because in this case the whole, original, signal is available for analysis.

3.1.2 FPT

The Fast Padé Transform (FPT) models a signal as the quotient of two polynomials. There are quite a number of ways of implementing the FPT [7–10, 12, 13, 18]. In the case of NMRS signals, the most appropriate FPT is the para-diagonal FPT because its parametric functional form, which is introduced towards the end of the FPT discussion (*Eq. 3.1.38 & 3.1.39*), is equivalent

to a *finite Green function* description of a system's quantum mechanical behaviour [16], also introduced later.[3] Hence, by opting for the para-diagonal FPT we are implicitly assuming that the NMRS behaviour of metabolites is quantum mechanical. Mathematically, the para-diagonal FPT is the FPT set-up whereby the degree of the denominator polynomial is greater than that of the numerator polynomial by exactly 1. Herein, the para-diagonal FPT, which is used for all the FPT based estimates within this text, is summarized and discussed; it includes a few additional steps, and the purpose of each is clearly discussed.

From the underlying theory of the FPT [9,10], the frequency spectrum of a time series signal is defined (is modelled) as the convergent quotient of two polynomials

$$\sum_{n=0}^{N_p-1} y_n z^{-n} \approx \frac{P_L^{\pm}(z^{\pm 1})}{Q_M^{\pm}(z^{\pm 1})} \qquad z = e^{i\tau\omega}, \quad \Im(\omega) > 0, \quad N_p \leq N \tag{3.1.18}$$

whereby

$$FPT^{-}: \frac{P_L^{-}(z^{-1})}{Q_M^{-}(z^{-1})} = \frac{\sum_{l=0}^{L} p_l^{-} z^{-l}}{\sum_{m=0}^{M} q_m^{-} z^{-m}} \quad \& \quad FPT^{+}: \frac{P_L^{+}(z^{+1})}{Q_M^{+}(z^{+1})} = \frac{\sum_{l=1}^{L} p_l^{+} z^{+l}}{\sum_{m=0}^{M} q_m^{+} z^{+m}} \tag{3.1.19}$$

y_n : is the n^{th} point of the noisy $N \times 1$ NMRS signal Y

N_p : partial length, i.e. $y_n \in \left(y_0, \ldots, y_{N_p-1}\right)^T$, $N_p \leq N$

and the quotients are convergent in the regions

$$\mathrm{FPT}^{-}: |z| > 1 \qquad \& \qquad \mathrm{FPT}^{+}: |z| < 1 \tag{3.1.20}$$

The FPT^{-} and FPT^{+} are *complementary* forms of the FPT [9]. $p_l^{\pm}$ and $q_m^{\pm}$ are their polynomial coefficients; the exponents are used to differentiate between the FPT forms. As initially noted, the para-diagonal FPT [9, 11] is used throughout this text, which means $L \equiv M - 1$. Either part of the times series signal, the first N_p points, or the whole signal, i.e. $N_p = N$ points, can be used in FPT analysis; the context will become clear as we work through the section. The parameters of a mono-component are estimated via the poles of $\mathrm{FPT}^{\pm}$, however the coefficients $q_m^{\pm}$ of the denominator polynomials are unknown, therefore they must be determined first.

Step 1: Determining the polynomial coefficients

In the case of FPT^{+}, the coefficients p_l^{+} & q_m^{+} are determined by cross-multiplying *Eq. 3.1.18*, leading to

$$\left(\sum_{m=0}^{M} q_m^{+} z^{+m}\right)\left(\sum_{n=0}^{N_p-1} y_n z^{-n}\right) = \sum_{l=1}^{M-1} p_l^{+} z^{+l} \tag{3.1.21}$$

[3] The Green function description of a system's quantum mechanical behaviour is *equivalent* to Schrödinger's description [10, 14], which − in conjunction with the Liouville von Neumann Equation − was used to describe NMRS metabolite behaviour in *Chapter 2* (*Theoretical Aspects of NMRS*, *page 2*). *Equivalent* means their mathematical expressions give the same result, although they are functions of different variables and parameters.

and then resolving the system of equations resulting from the expansion of *Eq. 3.1.21*, i.e. [10]

$$\begin{pmatrix} y_1 & y_2 & y_3 & \cdots & y_M \\ y_2 & y_3 & y_4 & \cdots & y_{M+1} \\ \vdots & \vdots & \vdots & & \vdots \\ y_{J+1} & y_{J+2} & y_{J+3} & \cdots & y_{J+M} \end{pmatrix} \begin{pmatrix} q_1^+ \\ q_2^+ \\ \vdots \\ q_M^+ \end{pmatrix} = - \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_J \end{pmatrix} \tag{3.1.22}$$

and

$$\begin{pmatrix} p_1^+ \\ p_2^+ \\ \vdots \\ p_{M-1}^+ \end{pmatrix} = \begin{pmatrix} y_0 & y_1 & y_2 & \cdots & y_{M-2} \\ 0 & y_0 & y_1 & \cdots & y_{M-3} \\ 0 & 0 & y_0 & \cdots & y_{M-4} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \ldots & y_0 \end{pmatrix} \begin{pmatrix} q_1^+ \\ q_2^+ \\ \vdots \\ q_{M-1}^+ \end{pmatrix} \tag{3.1.23}$$

Similarly, the coefficients p_l^- & q_m^- are determined by resolving the system of equations

$$\begin{pmatrix} y_M & y_{M-1} & y_{M-2} & \cdots & y_1 \\ y_{M+1} & y_M & y_{M-1} & \cdots & y_2 \\ \vdots & \vdots & \vdots & & \vdots \\ y_{M+J-1} & y_{M+J-2} & y_{M+J-3} & \cdots & y_J \end{pmatrix} \begin{pmatrix} q_1^- \\ q_2^- \\ \vdots \\ q_M^- \end{pmatrix} = - \begin{pmatrix} y_{M+1} \\ y_{M+2} \\ \vdots \\ y_{M+J} \end{pmatrix} \tag{3.1.24}$$

and

$$\begin{pmatrix} p_0^- \\ p_1^- \\ \vdots \\ p_{M-1}^- \end{pmatrix} = \begin{pmatrix} y_0 & 0 & 0 & \ldots & 0 \\ y_1 & y_0 & 0 & \ldots & 0 \\ y_2 & y_1 & y_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_M & y_{M-1} & \cdots & \cdots & y_0 \end{pmatrix} \begin{pmatrix} 1 \\ q_1^- \\ \vdots \\ q_{M-1}^- \end{pmatrix} \tag{3.1.25}$$

In both cases $J = N_p - 1 - M$, and the standard settings $p_0^+ = 0$, $q_0^+ = 1$, and $q_0^- = 1$ [10] were used for all calculations. It was stated earlier that the mono-component parameters are determined via the poles of FPT$^{\pm}$. By *Eq. 3.1.19*, each FPT form has a maximum of M poles since each has a denominator polynomial of degree M. And in order to extract all the poles that correspond to the mono-components embedded in a time series signal, in addition to the noise related poles, the conditions [9, 14]

$$FPT^- : N_p \geq 2M, \quad FPT^+ : N_p > 2M \tag{3.1.26}$$

must be satisfied. *Eq. 3.1.26* implies that the utilized length of the time series signal must be at least twice the number of poles. The number of deducible poles, w.r.t. the optimal pole value that can be converged to per pole, is constrained by the signal length of a well-sampled signal. By the Nyquist sampling theorem a well-sampled signal is a signal whose sampling time τ is such that [83]

$$\tau \leq \frac{1}{2 \times f_{max}} \tag{3.1.27}$$

f_{max} : maximum frequency component of a continuous signal

The smaller the sampling time, the smaller the information lost due to sampling, and the more similar the sampled and continuous signals are.[4] Referring back to the conditions of *Eq. 3.1.26* and optimal pole values, the optimal FPT$^-$ setting is $M = N/2$ [9], i.e. for a given signal the optimal pole values occur when M is half the full length of the signal. In the case of FPT$^+$ the optimal setting is $M = N/2 - 1$. All in all, in order to obtain the optimal results per signal the maximum settings of M are

$$\text{FPT}^- : M = \frac{N}{2}, \quad \text{FPT}^+ : M = \frac{N}{2} - 1 \quad \Longrightarrow N_p = N \tag{3.1.28}$$

These are the settings used for all the FPT analysis of this thesis. When will such settings fail? The FPT$^-$, for example, has at most $N/2$ poles w.r.t. its optimal setting. Each pole is either associated with a genuine peak or a noise peak. As the SNR of an acquired signal decreases, the greater the number of noise peaks, and hence the greater the number of poles associated with noise peaks. Neither the genuine nor noise associated poles are determined first − more genuine poles are found as the degree (i.e. M) of the denominator polynomial is increased to the optimal maximum $N/2$. After a genuine pole has been found, its estimated value converges to its *correct value* as the degree M is increased. By extension, because the parameters of a mono-component are estimated via its pole, the parameter values also converge to their *correct value*, as illustrated by the examples of *Fig. 3.1*. If the estimated values of a genuine pole converge to the pole's *correct value* when M is much smaller than $N/2$, the estimated values thereafter will be approximately equal.[5] On the other hand, if the estimated values of a genuine pole are still converging when $M = N/2$ then the resulting mono-component parameter values may or may not be quite erroneous, depending on the convergence curve. Altogether, the ability to detect a genuine pole, and the rate of convergence of its value, depends on a signal's SNR, and improves as the length of a well-sampled signal increases. Signal length is increased either by

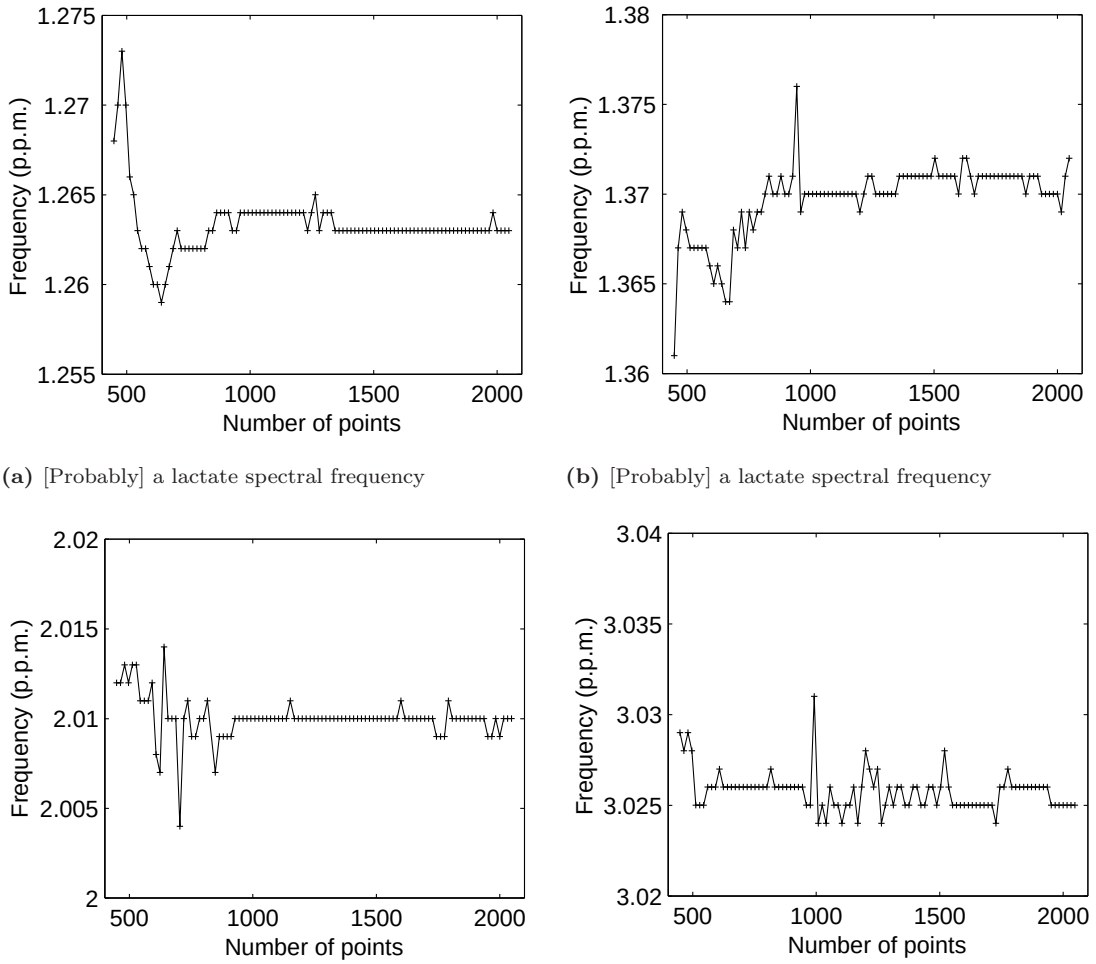
- reducing the sampling time, thereby acquiring more data points within a time period, and hence retaining more continuous-signal information, or
- increasing the length of the acquired signal without changing the sampling time, which will only be useful if a shorter signal means acquiring a signal that does not fully decay.

Hence the settings might fail, or rather the accuracy of the estimated mono-component parameter values might be poor, if there are not enough points to determine most of the genuine poles within a reasonable level of accuracy.

Considering the last point, lets return to the examples of *Fig. 3.1* & *3.2*. The graphs illustrate the convergence behaviour of 4 estimated spectral frequencies, and their corresponding estimated amplitudes, as the number of points N_p used for FPT$^-$ analysis is increased from $448 \rightarrow N = 2048$, using a step length of 16. The estimates were extracted from *Fig. 1.1*'s time series, which was experimentally acquired from the phantom detailed in *Table C.1*. The experiment's parameter values − $T_E = 35$ ms (*Table D.1*), NSA = 128, voxel size 20 mm ×

[4] The maximum frequency component of all the acquired NMRS signals of this text was 1250Hz, and the sampling time per experiment was 0.0004s, i.e. the condition of *Eq. 3.1.27* was always satisfied.

[5] The behaviour of convergent series is discussed in detail in *Chapters 4* and *14* of [68].



(a) [Probably] a lactate spectral frequency

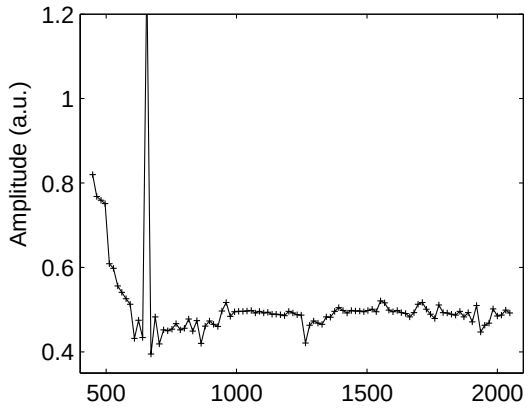
(b) [Probably] a lactate spectral frequency

(c) [Probably] an NAA spectral frequency

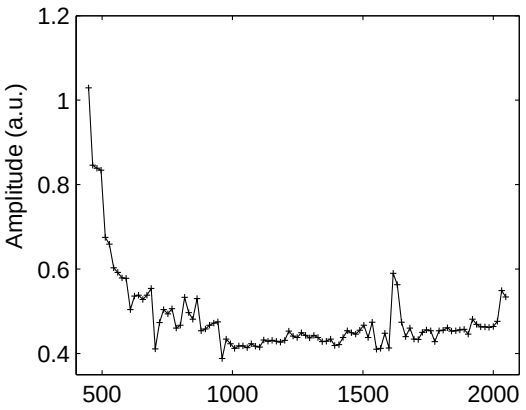
(d) [Probably] a creatine spectral frequency

Figure 3.1: Examples of convergence behaviour: Frequency estimates. These graphs illustrate the convergence behaviour of four estimated spectral frequencies as the number of points N_p used for FPT$^-$ analysis is increased from 448 $\rightarrow$ N, where N is the length of the NMRS time series signal, using a step length of 16. The estimates were extracted from *Fig. 1.1*'s time series, which was acquired from the phantom, of known contents, detailed in *Table C.1*, hence the captions. The acquisition parameters used are $T_E = 35$ ms (*Table D.1*), NSA = 128, and voxel size 20 mm $\times$ 20 mm $\times$ 15 mm. The length of the acquired signal $N = 2048$.

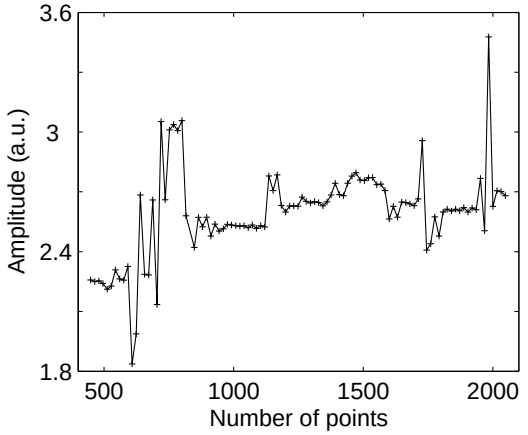
20 mm $\times$ 15 mm, and the number of data points acquired $N = 2048$ − are standard values, i.e. they are the same as the values used for some human based studies [5]. The graphs, *Fig. 3.1* & *3.2*, suggest that the convergence rate of the estimated amplitude values is slower than that of the estimated spectral frequency values. Additionally, only the amplitude of *Fig. 3.1a* (and perhaps *Fig. 3.1b*) seems to have converged. Altogether, considering the experiment's parameter values and the fact that $N > 2048$ is rare in NMRS studies, *it is quite probable that the amplitude estimates of the FPT, depending on an acquired signal's SNR and length (w.r.t. sampling time and how well the FID decays), may [generally or sometimes] have a lower level of accuracy than the frequency estimates.* This point also applies to the decay constant estimates,



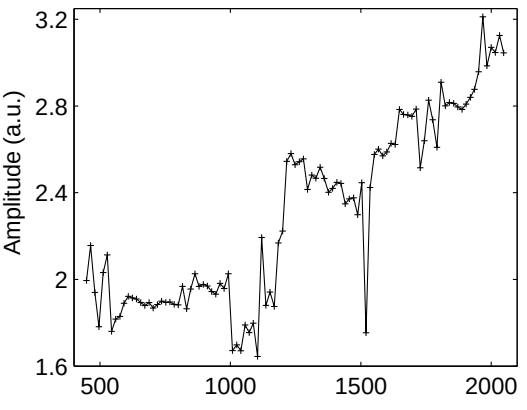
(a) Amplitude estimates w.r.t. *Fig. 3.1a*



(b) Amplitude estimates w.r.t. *Fig. 3.1b*



(c) Amplitude estimates w.r.t. *Fig. 3.1c*



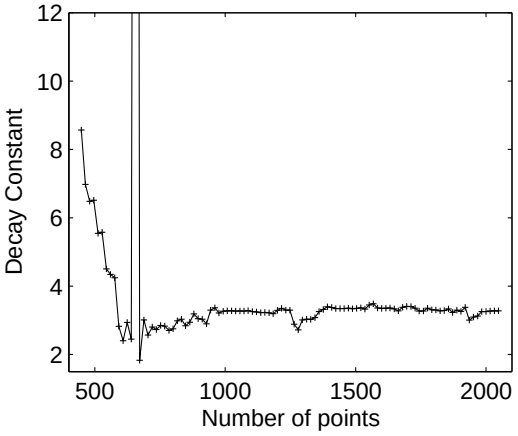
(d) Amplitude estimates w.r.t. *Fig. 3.1d*

Figure 3.2: Examples of convergence behaviour: Amplitude estimates. These graphs illustrate the convergence behaviour of the estimated amplitudes associated with the four estimated spectral frequencies of *Fig. 3.1*.

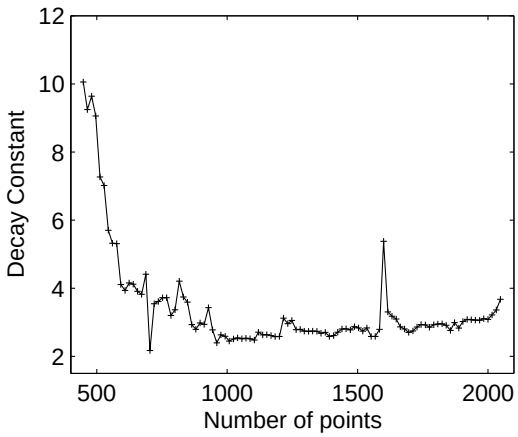
as *Fig. 3.3* illustrates, however the convergence rate of the decay estimates seem to be higher than those of the amplitude estimates. A final point concerns the non-smoothness of the illustrated convergence curves. The non-smoothness is due to noise and numerical/rounding errors; this also applies to the large spikes of *Fig. 3.2a* & *Fig. 3.3a*.

Step 2: The valid poles of FPT$^{\pm}$

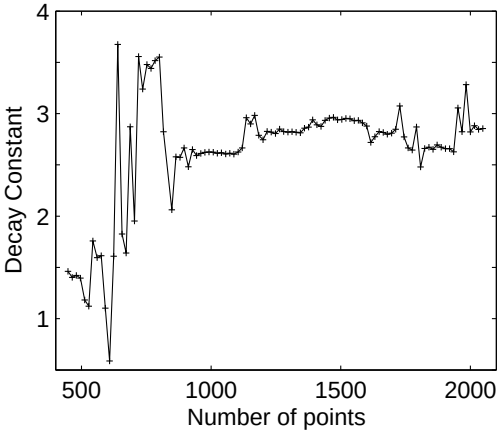
After substituting for $p_l^{\pm}$ & $q_m^{\pm}$ in FPT$^{\pm}$ (*Eq. 3.1.19*), the roots of the numerator and denominator polynomials are evaluated; in this text the roots were evaluated via the *root function* of Matlab®. A key property, especially in the case of noisy signals, is the presence of noise related pole-zero pairs, known as Froissart Doublets [9], which cancel each other out. Unfortunately, due to rounding errors, pole-zero cancellation is not a straightforward case of comparing the numerator and denominator roots. For this project, this problem was addressed via the inner product rule (*Chapter 8* [68]). Let a be a complex zero, and let b be a complex pole − their



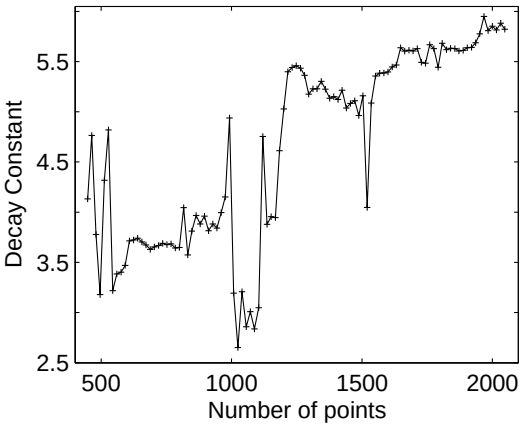
(a) Decay constant estimates w.r.t. *Fig. 3.1a*



(b) Decay constant estimates w.r.t. *Fig. 3.1b*



(c) Decay constant estimates w.r.t. *Fig. 3.1c*



(d) Decay constant estimates w.r.t. *Fig. 3.1d*

Figure 3.3: Examples of convergence behaviour: Decay constant estimates. These graphs illustrate the convergence behaviour of the estimated decay constant associated with the four estimated spectral frequencies of *Fig. 3.1.*

pairing is considered a Froissart Doublet if

$$\theta = \arccos\left[\frac{(a_r \quad a_i) \bullet (b_r \quad b_i)}{|a|\,|b|}\right] < 0.000005, \quad \text{whereby } \bullet \text{ denotes dot product} \tag{3.1.29}$$

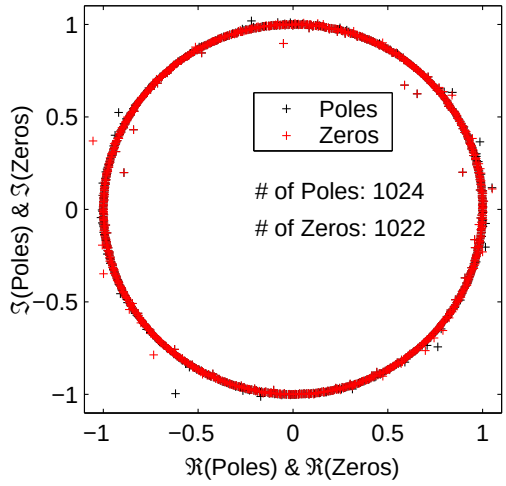
i.e. the angle between the geometric lines is approximately zero, correct to 4 decimal places (d.p.); $a_r \equiv \Re(a)$, $a_i \equiv \Im(a)$, $b_r \equiv \Re(b)$, and $b_i \equiv \Im(b)$. A second property, in addition to Froissart Doublets, used for extracting genuine peak poles is the convergence region condition of FPT$^{\pm}$ [9]. In the case of FPT^{-}, and w.r.t a complex pole b, b is accepted if $|b| > 1$, whereas in the case of FPT^{+} b is accepted if $|b| < 1$, i.e.

$$\text{FPT}^{-}: \quad \text{if } |\text{pole}| > 1, \text{accept pole} \tag{3.1.30}$$

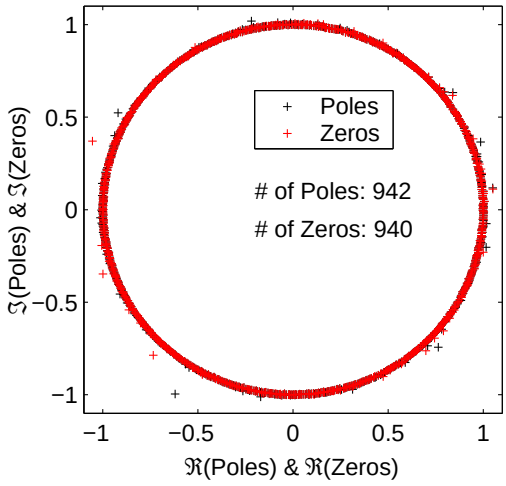
$$\text{FPT}^{+}: \quad \text{if } |\text{pole}| < 1, \text{accept pole} \tag{3.1.31}$$

Also, the pole of either FPT forms is accepted if and only if

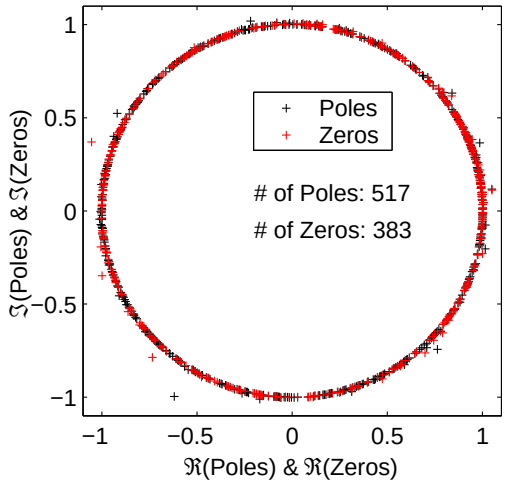
$$\Im\left\{\mp\frac{i}{\tau}\ln(b^{\pm})\right\} > 0 \tag{3.1.32}$$



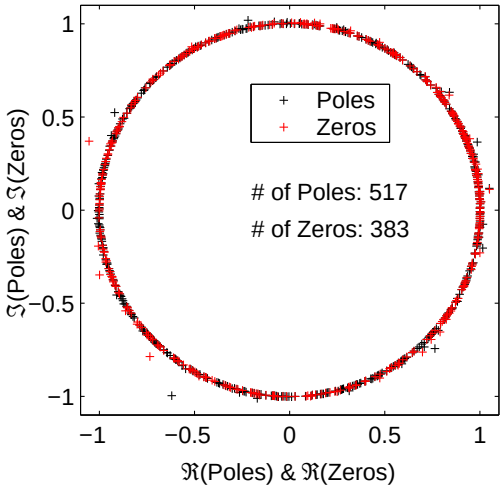
(a) All poles & zeros



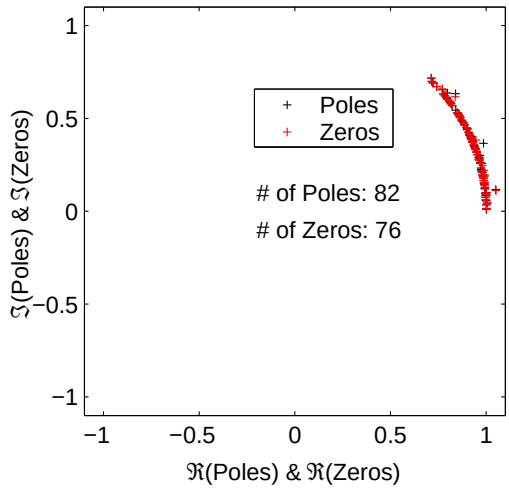
(b) After doublet cancellation



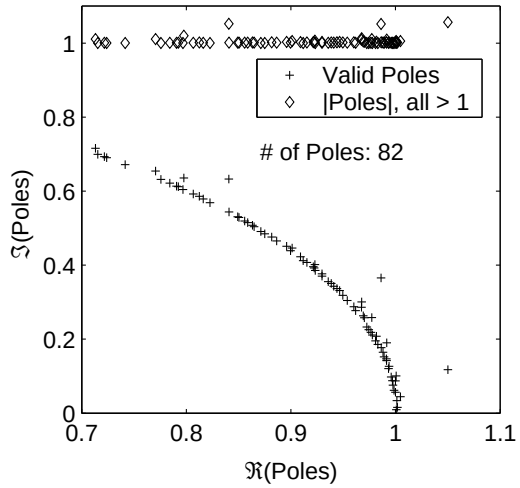
(c) $|\text{Poles}| > 1$, $|\text{Zeros}| > 1$



(d) $\Im(\omega) > 0$ (*Eq. 3.1.32*)



(e) Poles & Zeros within -5p.p.m. & 0p.p.m., cf. *Fig. 1.1a*



(f) **ZOOMING IN ON THE POLES OF (e).** Their moduli is also plotted.

Figure 3.4: Poles and zeros: an FPT$^-$ example

This is the same as the condition of *Eq. 3.1.18*, i.e. $\omega^{\pm} = \mp(i/\tau)\ln(pole^{\pm})$. In *Fig. 3.4* the signal of *Fig. 1.1a* is used to illustrate the sequential effect of a set of FPT^{-} conditions on the original set of poles and zeros deduced via FPT^{-}. *Fig. 3.4b* illustrates the effect of doublet cancellation on the original set of poles and zeros (*Fig. 3.4a*). Next, the condition of *Eq. 3.1.30* was applied to the poles and zeros of *Fig. 3.4b* (i.e. the results of doublet cancellation), the result is displayed in *Fig. 3.4c*. The condition of *Eq. 3.1.32* is then applied to the poles and zeros of *Fig. 3.4c*, the results are displayed in *Fig. 3.4d*. The next condition is discussed in the next step.

Step 3: The mono-component parameters

Let $z_k^{\pm}$ refer to the k^{th} valid pole of FPT$^{\pm}$. The ω term of *Eq. 3.1.18* is defined as

$$\omega_k^{\pm} = \mp\frac{i}{\tau}\ln(z_k^{\pm}) \tag{3.1.33}$$

and its real and imaginary parts are the frequency and decay constant respectively, i.e.

$$f_k^{\pm} = \Re(\omega_k^{\pm}), \qquad d_k^{\pm} = \Im(\omega_k^{\pm}) \tag{3.1.34}$$

Basically, the condition of *Eq. 3.1.18* (or *Eq. 3.1.32*) ensures that the decay constants of the valid poles (and zeros) are greater than zero. The spectral frequencies f_k can be used to impose an additional condition w.r.t. NMRS signal analysis − a frequency range of interest condition [15]. Most *in vivo* or *in vitro* NMRS studies concentrate on spectral frequencies that occur upfield w.r.t. water, i.e. within the range $[-5 \quad 0]$ p.p.m. $\Rightarrow [0.1 \quad 4.9]$ p.p.m., because the downfield spectral frequencies are usually not detected. For example, all the detected spectral frequencies of *Fig. 1.1a* occur before the 0p.p.m. mark, the approximate position of the removed water peak; the water peak normally appears around 0p.p.m. However it is important to note that the detection of downfield spectral frequencies is possible if an appropriate mixture of "short echo time, appropriate water suppression methods, and high magnetic field strengths" [5] is used. In the case of the NMRS experiment and simulation types conducted for this thesis, detailed in *Tables D.1 & D.2*, only the upfield frequencies ($[-5 \quad 0]$ p.p.m., i.e. $[0.1 \quad 4.9]$ p.p.m.) are detectable, therefore the condition

$$f_k^{\pm} \in [-5 \quad 0] \text{ p.p.m.} \tag{3.1.35}$$

was applied to the poles (and zeros). Hence, continuing with the example of *Step 2*, the condition of *Eq. 3.1.35* was applied to the poles and zeros of *Fig. 3.4d*, and the results are displayed in *Fig. 3.4e*. *Fig. 3.4f* zooms in on the poles of *Fig. 3.4e*, since we are mainly interested in the poles, as indicated by *Eq. 3.1.33 & 3.1.34*. Within the same graph the corresponding |pole| of each accepted pole is plotted.

The amplitude and phase values of the estimated mono-components are given by [8, 9, 13]

$$a_{k,j_k}^{\pm} = \left|\zeta_{k,j_k}^{\pm}\right|, \qquad \phi_{k,j_k}^{\pm} = Arg\left(\zeta_{k,j_k}^{\pm}\right), \qquad j_k = 1, 2, 3, \ldots \tag{3.1.36}$$

whereby

$$\zeta_{k,j_k}^{\pm} = P_L^{\pm}(z_k^{\pm 1}) \Big/ \left(\frac{d}{dz^{\pm 1}}\right)^{j_k} Q_M^{\pm}(z_k^{\pm}) \tag{3.1.37}$$

The premise of *Eq. 3.1.37*, which is called the Cauchy residue [9,13], is discussed in [8,9,13]. The denominator of *Eq. 3.1.37* is the j_k^{th} derivative of the denominator of *Eq. 3.1.18*. The subscript j_k refers to the multiplicity number of a pole. The multiplicity or order of a pole is the number of times the same pole value occurs. In FPT analysis $max(j_k) > 1$ implies overlapping mono-components, i.e. mono-components of the same frequency but different decay rates [7,9]. Thus the FPT has an important feature of interest for NMRS signal analysis − the ability to resolve overlapping peaks. Another feature of interest is the parametric form of the FPT, which is an indirect function of the mono-component parameters. Earlier, it was noted that the parametric form of the para-diagonal FPT is equivalent to the finite Green function description of quantum mechanical behaviour [16]; refer back to *page 42*. The non-parametric forms of the FPT$^{\pm}$, described by *Eq. 3.1.19*, and the parametric forms of the FPT$^{\pm}$, are related as follows

$$FPT^{-}: \frac{P_L^{-}(z^{-1})}{Q_M^{-}(z^{-1})} = \frac{\sum_{l=0}^{L} p_l^{-} z^{-l}}{\sum_{m=0}^{M} q_m^{-} z^{-m}} = \sum_{k=1}^{K} \sum_{j_k=1}^{J_k} \frac{\zeta_{k,j_k}^{-}}{\left(z^{-1} - z_k^{-1}\right)^{j_k}} \tag{3.1.38}$$

$$FPT^{+}: \frac{P_L^{+}(z^{+1})}{Q_M^{+}(z^{+1})} = \frac{\sum_{l=1}^{L} p_l^{+} z^{+l}}{\sum_{m=0}^{M} q_m^{+} z^{+m}} = \sum_{k=1}^{K} \sum_{j_k=1}^{J_k} \frac{\zeta_{k,j_k}^{+}}{\left(z^{+1} - z_k^{+1}\right)^{j_k}} \tag{3.1.39}$$

whereby

$\zeta_{k,j_k}^{\pm}$ is defined by *Eq. 3.1.37*

$z_k^{\pm}$ are the valid poles of FPT$^{\pm}$ (*page 50*)

The expressions on the right-hand-side of *Eq. 3.1.38* & *3.1.39* are the parametric para-diagonal FPT descriptions of a time series signal. Both expressions are also the finite (or truncated) Green function descriptions of a system's quantum mechanical behaviour [14].

Step 4: Internal Cross-Validation FPT^{+} & FPT^{-}

The FPT^{-} & FPT^{+} are complementary forms whose aim is the same − in this case determining the mono-components of a signal, and their parameter values. Therefore only the mono-components that exist in both FPT forms are considered genuine [13], i.e. mono-components such that

$$\omega_k^{+} \approx \omega_k^{-}, \text{i.e. } f_k^{+} \approx f_k^{-} \quad \& \quad d_k^{+} \approx d_k^{-} \tag{3.1.40}$$

and consequently $a_k^{+} \approx a_k^{-}$. Hence, referring back to the example of *Fig. 3.4*. The FPT^{-} poles of *Fig. 3.4f* were compared with the FPT^{+} poles of the same level. This cross-validation, which must be a one-to-one mapping of FPT^{-} & FPT^{+} poles, resulted in fewer, and an equivalent, number of poles per FPT form (*Fig. 3.5*); prior to cross-validation FPT^{-} had 82 poles left (*Fig. 3.4f*), whilst FPT^{+} had 83 (not shown), afterwards each had 70 poles left. In this text the cross-validation was conducted via the frequency $f_k^{\pm}$ and decay $d_k^{\pm}$ values of a pole; *Eq. 3.1.33* and *Eq. 3.1.34* outlined how to evaluate the frequency and decay values of a pole. For each f_k^{-}

value, the closest f_k^+ value is found from a set of f_k^+ values. The *dsearchn function* of Matlab® was used for this purpose. Subsequently, if more than one f_k^- value is closest to the same f_k^+ value then their corresponding $d_k^{\pm}$ values were also compared. In this latter case it is the pairing with the best/smallest combined frequency & decay value deviation, that is accepted. Hence the one-to-one mapping of FPT^- and FPT^+ poles was evaluated by determining the closest frequency pairs or the closest frequency & decay pairs.

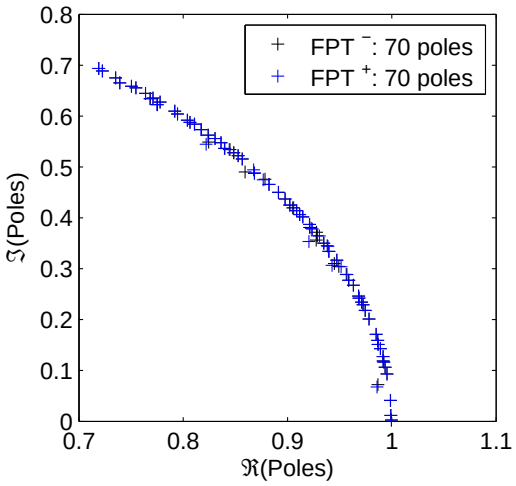


Figure 3.5: Internal cross-validation of the FPT

Comments

The FPT^- converges much more quickly than the FPT^+ [6]. Hence, noting that the estimated f_k, d_k, a_k, and ϕ_k values converge to their *real* values as $N_p \to \infty$, the FPT^- values will be much more accurate. This is especially true in the case of noisy signals because more data points are needed to attain convergence, as discussed previously. $N_p = N$ is used for all FPT analysis in this text.

The FPT is not a widely used method, and the reason why is not clear. It is quite possible that the absence of a computer application, and the fact that it is mathematically involved, has discouraged potential users. The brief review of this method by Poullet et al. [82] also suggests a lack of interest in exploring the strengths and weaknesses of the method. The paper by Poullet et al. [82] was published in 2008, but refers to only two papers published during the early stages of applying FPT to NMRS: [13] (method development), and [113] (application). Subsequent advances, e.g. the 2006 journal papers [9–11, 14], were not referred to. Consequently some of the paper's assertions are wrong. For example, and contrary to the paper's claim, the FPT can resolve overlapping mono-components, as noted above. However if two signals from different biochemical sources have the same frequency and decay values then it is highly improbable that any analysis method will be able to resolve this. The best that can be done is to assess whether such a mono-component can actually be associated with a single biochemical.

If we consider some of the potential problems listed in *Table 2.3*, two advantages of the FPT are its line-shape independence (i.e. it can extract line-shapes of any known or unknown form

from a multi-component signal, therefore peculiar line-shapes do not need to be re-modelled), and its non-linearity.

3.1.3 Section Summary

This section has discussed a set of non-interactive NMRS signal analysis methods − the widely used Singular Value Decomposition based methods and the Fast Padé Transform. The probable problem points, which may lead to poor or sub-optimal results, where highlighted. The problems are due to one or more of the following: an inappropriate analysis structure imposed by a user via structural parameters, an insufficient number of well-sampled data points, rounding errors, poor SNR. The number of well-sampled data points can be increased either by **(1)** reducing the sampling time, thereby acquiring more data points within a time period, and hence reducing information loss due to sampling, or **(2)** increasing the length of the acquired signal without changing the sampling time, which will only be useful if a shorter signal means acquiring a signal that does not fully decay. The optimal settings, in theory, that can lead to the best set of results for a given well-sampled signal, of certain length, were also discussed. The optimal settings of the discussed non-interactive methods are summarized in *Table 3.2*. Next the main interactive methods are briefly discussed.

	N_p	M	K	τ_Δ	L
SVD Based Methods & jMRUI	N	−	50	0	$\frac{N}{2}$
FPT$^-$	N	$\frac{N}{2}$	−	−	−
FPT$^+$	N	$\frac{N}{2}-1$	−	−	−

Table 3.2: Optimal settings of the non-interactive methods. N_p: The number of points of a $N \times 1$ signal used for an analysis. M: The degree of the denominator polynomials. K: Number of mono-components or components. τ_Δ: Time series shift parameter. L: Number of Hankel matrix rows.

3.2 Interactive Methods

The interactive methods are methods which use prior knowledge. In some cases prior knowledge usage is optional − e.g. Variable Projection method [106], Advanced Method for Accurate, Robust and Efficient Spectral fitting [110] − in some cases it is compulsory − e.g. Quantum Estimation [88, 89], Automated Quantitation of Short Echo Time MRS Spectra [81], Linear Combination Model [84]. These five interactive methods are the most widely used, and each is briefly discussed below.

3.2.1 VARPRO

The Variable Projection method [106] is a non-linear time domain parametric fitting method. Its model function is

$$y_n = x_n + e_n = \left(\sum_{k=1}^{K} a_k e^{i\phi_k} e^{-d_k(1-g+gt_n)t_n + i2\pi f_k t_n} \right) + e_n \tag{3.2.1}$$

i.e. a variant of *Eq. 2.0.1*, where all the terms have been defined (*page 9*). The damping correction g, which controls the form of an assumed line-shape, is set a priori. The same g value is assigned to all prospective components. Briefly, if $g = 0$ the line-shapes are Lorentzian, if $g = 1$ the line-shapes are Gaussian, and if $g \in (0\,1)$ the line-shapes are Voigt line-shapes [48]. The VARPRO method uses the Variable Projection method [41, 58] to modify the normal cost function, hence its name. The normal cost function is

$$C(\mathbf{a}, \mathbf{d}, \mathbf{f}, \phi, t) = \sum_{n=0}^{N-1} |y_n - x_n|^2 = \|Y - \Psi_{VP} L\|^2 \tag{3.2.2}$$

such that Ψ_{VP} is a matrix of terms $e^{-d_k(1-g-gt_n)t_n + i2\pi f_k t_n}$ and $L = \left(a_1 e^{i\phi_1} \ldots a_K e^{i\phi_K}\right)^T$; this means Ψ_{VP} is a function of decay constant, frequency and time, i.e. $\Psi_{VP}(\mathbf{d}, \mathbf{f}, t)$, whilst L is a function of amplitude and phase, i.e. $L(\mathbf{a}, \phi)$. The aim is to minimize *Eq. 3.2.2*, and the VARPRO method starts off by setting

$$\|Y - \Psi_{VP} L\|^2 = 0 \tag{3.2.3}$$

which leads to

$$Y = \Psi_{VP} L \;\Rightarrow\; L = \Psi_{VP}^{\dagger} Y \tag{3.2.4}$$

i.e. the initial damping factor and frequency values are used to estimate $L(\mathbf{a}, \phi)$

$$\hat{L}(\mathbf{a}, \phi) = \Psi_{VP}^{\dagger} Y \tag{3.2.5}$$

Hence, the initial or prior amplitude and phase information is not required in the VARPRO method. By substituting for L in *Eq. 3.2.2*, the **modified cost function** is

$$C(\mathbf{d}, \mathbf{f}, t) = \left\| Y - \Psi_{VP} \Psi_{VP}^{\dagger} Y \right\|^2 \tag{3.2.6}$$

and this function, known as the variable projection functional, **is the function that is minimized**. It is minimized, w.r.t. decay constant and frequency, using a variation of the Levenberg-Marquardt (LM) algorithm [72, 79]. A peculiar feature of the method is the minimization of *Eq. 3.2.6* with respect to $\sqrt{d_k}$ to ensure that $d_k \geq 0 \,\forall\, k$. The structure of the Levenberg-Marquardt algorithm allows for limited prior knowledge usage; the Levenberg-Marquardt algorithm is discussed briefly within the *Minimization Algorithms* section of *Appendix A* (*page 143*). There are a few problems with the VARPRO method: all peaks are constrained to the same line-shape form, the prior knowledge options are limited, and the minimization algorithm performs best when handling linear or approximately linear problems [76].

A user *may* use prior knowledge via the parameters a_k, ϕ_k, d_k, g_k, and f_k; prior knowledge usage is not compulsory. It is possible to link the parameter values, e.g. a set of frequency

parameters associated with the same metabolite may be defined in terms of the differences between them. Non-uniqueness is also a feature of this method − it is possible to obtain a range of plausible results per signal. The results depend on the number of mono-components that are assumed, by the user, to be embedded in the multi-component signal. The method then tries to *fit* the result to the user's assumptions. Hence VARPRO is susceptible to under-estimation or over-estimation of mono-components. If prior knowledge is not used the user has to manually discern the metabolite associations of the estimated components. However, if prior knowledge is used, the user should know the metabolite associations of the estimated peaks for which prior knowledge has been provided.

3.2.2 AMARES

The Advanced Method for Accurate, Robust and Efficient Spectral fitting [110] is a non-linear time domain parametric fitting method. Its model function is

$$y_n = x_n + e_n = \left(\sum_{k=1}^{K} a_k e^{i\phi_k} e^{-d_k(1-g_k+g_k t_n)t_n + i2\pi f_k t_n} \right) + e_n \tag{3.2.7}$$

which is similar to the VARPRO model function. However in contrast to VARPRO each peak can have a different line-shape form, hence g_k rather than g. **It is the cost function**

$$C(\mathbf{a}, \mathbf{d}, \mathbf{f}, \phi, t) = \sum_{n=0}^{N-1} |y_n - x_n|^2 = \|Y - \Psi_{AM} L\|^2 \tag{3.2.8}$$

Ψ_{AM} : matrix of terms $e^{-d_k(1-g_k-g_k t_n)t_n + i2\pi f_k t_n}$

$L = \left(a_1 e^{i\phi_1} \ldots a_K e^{i\phi_K}\right)^T$

that is minimized, not a transformed version of it, as is the case with the VARPRO method of the previous section (*Eq. 3.2.6*). The cost function is an explicit function of all the parameters being estimated. An adaptive non-linear least squares algorithm, NL2SOL [33–35], is used to minimize *Eq. 3.2.8*. The NL2SOL algorithm is discussed briefly within the *Minimization Algorithms* section of *Appendix A* (*page 144*). An advantage of the NL2SOL algorithm is its flexibility; parameters can be constrained easily, hence, prior knowledge can be utilized extensively in the AMARES method. The review by Nazareth [76] suggests that it is a much more robust minimization algorithm for non-linear problems, compared to the Levenberg-Marquardt algorithm.

Similar to VARPRO, a user *may* use prior knowledge via the parameters a_k, ϕ_k, d_k, g_k, and f_k; prior knowledge usage is not compulsory. The parameter values may also be linked, e.g. an amplitude parameter value might be defined as a specified fraction of another. Non-uniqueness is also a feature of this method because a range of plausible results are plausible per signal. AMARES also tries to *fit* the result to a user's assumptions, i.e. the number of components that are assumed, by the user, to be embedded in the multi-component signal. Therefore it is also susceptible to under-estimation or over-estimation of mono-components. The prior knowledge conclusions of VARPRO also apply to AMARES.

3.2.3 QUEST

Quantum Estimation [88, 89] is a non-linear time domain semi-parametric method. Its functional form is

$$x_n = e^{i\phi_0} \sum_{k=1}^{K} \left(a_k \hat{x}_n^{(k)} e^{(\Delta\alpha_k + i\Delta\omega_k)t_n + i\Delta\phi_k} \right) \tag{3.2.9}$$

K is the number of individual peaks. Each $\hat{x}^{(k)}$ is a simulated (using QMP or literature values) or in vitro sample of a metabolite − together they form a metabolite basis set. Each evaluated parameter is the difference between the estimated parameters of corresponding simulated and NMR signal peaks. One of the dangers of this method is loss of information due to truncation of the first few points of the signal. The authors did this in order to reduce the effect of lipid and macromolecule base line; both tend to decay within the first few milliseconds of the acquisition period. To address the information loss in relation to other peaks, the authors back-extrapolate. Another issue is the quality of a set of results, which depends on the quality of the predicted metabolite basis set. Additionally, the method is subjective and non-unique because the metabolite contents of the basis set depend on the user's assumptions, and the results change as the contents of the basis set change.

The cost function minimization is via Levenberg-Marquardt [72].

3.2.4 AQSES

Automated Quantitation of Short Echo Time MRS Spectra [81] is perhaps one of the most complex NMR analysis methods available. Its functional form is

$$x_n = \sum_{k=1}^{K} \left(a_k e^{i\phi_k} e^{-d_k(1-g_k+g_k t_n)t_n + i2\pi f_k t_n} e^{i\epsilon_k t_n^2} \eta_k(t) \right) + b_n + w_n \tag{3.2.10}$$

Whereby η_k is the metabolite basis set, $e^{i\epsilon_k t_n^2}$ the eddy current term, b_n the baseline term and w_n the residual water component. The metabolite basis set has to be simulated or acquired in vitro. In contrast to the methods discussed so far, AQSES includes a water filter, it is a Finite Impulse Response Water Filter [100]; before using the other analysis methods the residual water signal is usually filtered out using a separate tool e.g. HLSVD & HLSVD Pro [69], which are two of the most popular water filters in NMRS signal analysis.

The baseline term $b(t)$ refers to the macromolecule and lipid components of the signal, and these are modelled non-parametrically using a basis of splines [95, 96]. Afterwards to ensure that the modelled baseline is concurrent with the signal being fit a regularized non-linear least squares cost function is minimized i.e. instead of minimizing $\sum_{n=0}^{N-1} |y_n - x_n|^2$ the problem is [96]

$$minimize \ \frac{1}{N} \sum_{n=0}^{N-1} |y_n - x_n|^2 + \lambda^2 \|Dc\|^2 \tag{3.2.11}$$

$\mathbf{D}$ measures the frequency of the baseline in the FD, the components of $\mathbf{c}$, $\mathbf{c} = c_1, ..., c_r$ where $r < N$, are the coefficients of the spline functions, and λ is a fixed regularization/penalty parameter. All in all $\lambda^2 \|Dc\|^2$ ascertains a degree of concurrency between $b(t)$ and the signal.

Expression 3.2.11 is a regularized non-linear LS parametric fitting problem and before it is minimized it is transformed into a variable projection functional, using an adapted variable projection method [96] based on the work of [41, 58]. The variable projection functional is minimized using *an extension of the Levenberg-Marquardt algorithm*, basically a NL2SOL algorithm [34]. The AQSES method is subjective and non-unique because the metabolite contents of the basis set depend on the user's assumptions, and the results change as the contents of the basis set change. Additionally, the order in which the metabolites of the basis set are introduced to the method matter − the results change as the order is changed.

3.2.5 LC Model

The Linear Combination (LC) Model [84] is a frequency domain semi-parametric analysis method. The algorithm is not outlined in detail in [84], perhaps because the plan, all along, was to implement it in commercially available products only. In fact the LC Model is only implemented in commercially available products, to which the project had no access. This section's review will have some shortcomings due to the limited information available.

The LC Model analyses *in vivo* NMRS signals using a basis set of metabolites acquired *in vitro*; it does not use quantum mechanically simulated spectra. Within a specified region of a frequency spectrum, e.g. [0 5]p.p.m., each spectral frequency point f_k of a grid of K_G spectral frequency points, i.e. $k = 1, \ldots, K_G$, is modelled as [84]

$$x(f_k) = e^{-i(\phi_0 + f_k \phi_1)} \left(\sum_{j=1}^{N_B} \beta_j B(f_k) + \sum_{l=1}^{N_M} C_l \sum_{n=-N_S}^{N_S} S_n M_l(\boldsymbol{f}; \gamma_l, \epsilon_l) \right) \tag{3.2.12}$$

subject to the constraints

$$C_l \geq 0 \quad \gamma_l \geq 0 \quad \sum_{n=-N_S}^{N_S} S_n = 1 \tag{3.2.13}$$

Each $x(f_k)$ model's the form a Fourier spectrum - which is the sum of the Fourier spectra of N_M metabolite time series signals acquired *in vitro* - at frequency f_k. The parameters ϕ_0 and ϕ_1 are the zero-order and first-order phases respectively. The zero-order phase is the frequency independent phase, whilst the first-order phase is an additional phase term that depends on the frequency and/or other parameters at time zero [27, 109]. B_j denotes cubic B-splines, which model the baseline of the frequency spectrum; β_j are the B-spline coefficients, and N_B is the number of splines. The concentration of each metabolite in the basis set is denoted by C_l, and N_M is the number of metabolites in the basis set. The S_n terms are the line-shape coefficients, which model the line-shape effects of field inhomogeneities, eddy currents, etc (*Table 2.3*). M_l is the *in vitro* metabolite basis set. Its explicit function is a function of **(1)** a line-broadening parameter γ_l, which accounts for relaxation time variations, and **(2)** a shift parameter ϵ_l, which accounts for small variations between a metabolite's basis set spectrum and a prospective *in vivo* fit. This suggests that M_l model's the decay rate given that the decay rate is a function of relaxation time [26, 60].

In line with the afore discussed interactive fitting methods, the required estimates of the LC Model are obtained via the minimization of a cost function. The cost function minimized

in LC Model analysis is the regularized cost function

$$\frac{1}{\sigma}\sum_{k=1}^{K_G}\left(y(f_k)-x(f_k)\right)^2+\|p_S R_S S\|+\|p_{\boldsymbol{\beta}} R_{\boldsymbol{\beta}}\boldsymbol{\beta}\| \tag{3.2.14}$$

$$+ \quad \text{prior normal probability distributions of } \gamma_l \ \&\ \epsilon_l$$

S : vector of line-shape coefficients S_n

$\boldsymbol{\beta}$: vector of B-spline coefficients β_j

Each $y(f_k)$ is a point, at frequency f_k, of the Fourier spectrum of an acquired NMRS time series signal. Hence *Eq. 3.2.14* compares the Fourier spectrum of an acquired signal with the Fourier spectrum that is the sum of the Fourier spectra of the metabolites assumed to be present in the acquired signal. The matrices R_S and $R_{\boldsymbol{\beta}}$ are the regularization matrices of the line-shapes and baselines respectively, these matrices are discussed in detail in [85, 86]. These regularization matrices, in combination with their regularization parameters p_S (for R_S) and $p_{\boldsymbol{\beta}}$ (for $R_{\boldsymbol{\beta}}$), constrain/regularize the forms of the line-shapes and baselines. The prior normal probability distribution of γ_l and ϵ_l are summarized in [84]. The complexity of LC Model's analysis is perhaps comparable with that of AQSES (previous discussion). The preparation required for using the LC Model can be quite time consuming and expensive. A user must first build a metabolite basis set using signals acquired from an *in vitro* solution of metabolites that are assumed to be present at a source. Additionally, each metabolite signal must be acquired under the same conditions − pulse sequence type (e.g. PRESS), pulse sequence parameters (e.g. $T_E = 35$, NSA $= 128$), ROI size, etc − as that of a prospective subject. Even if the time and expense issues are not a problem, the subjectivity and non-uniqueness of the method [8, 32] is, in spite of claims to the contrary [84, 87].

3.2.6 Section Summary

Amongst the interactive methods the main advantage of QUEST, AQSES, and LC Model over VARPRO, and AMARES is that they are spectrum rather than peak fitting methods, thus the links between the components of a metabolite do not need to be set by the user. Another advantage of the QUEST, AQSES, and LC Model methods is that they automatically associate each estimated spectrum with a metabolite. In the case of VARPRO and AMARES this only happens if prior knowledge is used. The main disadvantage of all the interactive methods is their subjectivity and non-uniqueness. Unfortunately, none of the methods includes a systematic way of determining the probable number of components of a signal, which could then inform the choice of metabolites for mono-component parameters or basis set purposes. The introductory chapter of [8] discusses, in detail, the non-uniqueness and subjectivity issue w.r.t. the LC Model; it also discusses other strengths and weaknesses of the method in detail. Most of the observations discussed therein also apply to all interactive methods.

3.3 A Comparative Analysis of Non-Interactive Methods

This section has two objectives:

1. To assess the **consistency** of an analysis' decompositions w.r.t. different noise signals of the same strength, i.e. reproducibility assessment, and

2. To assess the effect of increasing noise on the **decomposition accuracy** of an analysis.

per non-interactive method's optimal setting (*Table 3.2*). To this end, an artificial time series signal X_{anf}, of length $N = 2048$, was created and two computer experiments were conducted. The artificial signal's mono-component parameter values are detailed in *Table C.2*, and the resulting Fourier spectrum is also displayed within *Table C.2*. All the analysis detailed herein are time domain analysis, i.e. each analysed and/or discussed signal is a time series signal. The length of each time series signal, unless stated otherwise, is 2048.

3.3.1 The Experiments

1. The Reproducibility Assessment. Four different time series noise signals of the same strength were added to the artificial *noise free* signal X_{anf} in order to assess how consistent, or otherwise, a method's analysis is. The noise signals were created using the jMRUI [74, 75] application's simulation module, and the noise signals were added to X_{anf} using the same simulation module. The noise signals of jMRUI are created by

1. Creating a time series signal consisting of random values in the range [-0.5 0.5], and of the same length as a noise free time series signal to be corrupted.

2. Multiplying the signal created above by a value equal to a percentage of the largest time domain mono-component amplitude, of a multi-component time series signal. For example, if a multi-component time series signal consists of three mono-component time series signals of time domain amplitudes 2, 4, and 6. Then the signal created in 1. will be multiplied by a value that is a percentage of 6.

In the case of the 4 noise signals of similar strength used for this experiment, the maximum amplitude of each noise signal was 5% of X_{anf}'s largest mono-component time domain amplitude. The SNR of the resulting 4 noisy signals is given by

$$20 \log_{10} \left(\frac{\max(X_{an})}{\sigma_{noise}} \right) \tag{3.3.1}$$

X_{an} : Artificial noisy time series signal

σ_{noise} : The standard deviation of the added noise

Hence the approximate SNR values of the artificial noisy time series signals are

39.205, 39.287, 39.350, and 39.292 dB

As expected, the SNR values are approximately equal since the noise signals are of similar strength.

2. The Effect of Noise on Decomposition Accuracy. In this case, four different noise signals of increasing strength were added to X_{anf} in order to assess the extent to which increasing noise levels affect decomposition accuracy. As before, the noise signals were added to X_{anf} via the jMRUI [74, 75] application's simulation module. The maximum amplitude of each noise signal was 5%, 10%, 15%, and 20% of X_{anf}'s largest mono-component amplitude. The approximate SNR values of the resulting artificial noisy time series signals are

39.285, 33.383, 29.500, and 26.154 dB

for cases 5%, 10%, 15%, and 20% respectively. The spectra of two of the these artificial noisy signals are displayed in *Fig. 3.6*.

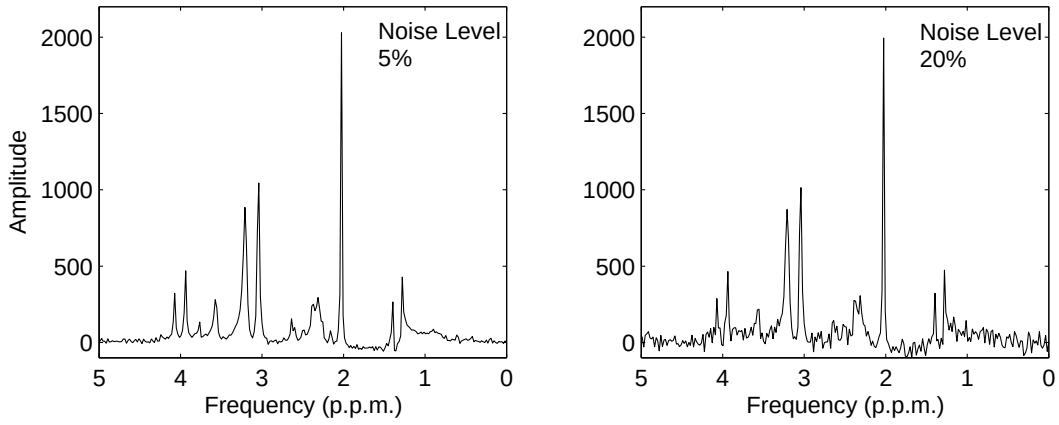


Figure 3.6: Examples of artificial noisy signals. Refer to *Table C.2* for a graph of the original noise free signal, and details of its mono-component parameters.

3.3.2 Analysis Measures

Number of Partnered Estimated Components All the estimated components within the frequency ROI, [0 4.28] p.p.m., were compared with the original mono-components of *Table C.2*, in order to find out which components each method was able to estimate per noisy signal. In each case the number of partnered estimates was recorded. Similar to the FPT's cross validation procedure (*page 51*) the comparison was conducted via the frequency and decay values. For each estimated frequency value, the closest original frequency value is found from the set of *Table C.2*. The *dsearchn function* of Matlab® was used for this purpose. If the standard deviation between any pairings is greater than 6, such a pairing is rejected; the value 6 is arbitrary, but it ensured that very few pairings were rejected. Subsequently, if more than one estimated frequency is *closest* to the same original frequency value then their corresponding decay values were also compared. In this latter case it is the pairing with the best/smallest combined frequency & decay value deviation that is accepted.

Percentage Error Per partnered estimated component, the percentage error of the estimated parameter values was evaluated. The errors are summarized via box plots, which provide median

and margin of error, or error boundary, information as defined in *Fig. 3.7b*. For comparison purposes the error measure used for the box plots is

$$E = \log_{10}\left(10^{-4} + \%\text{Error}\right) \tag{3.3.2}$$

The value 10^{-4} ensures that $\log_{10}(0)$ is avoided; when the percentage error is zero $E = -4$. The error grid of *Fig. 3.7a* demonstrates the link between a percentage error value and a box plots' values.

3.3.3 Results: Reproducibility Assessment

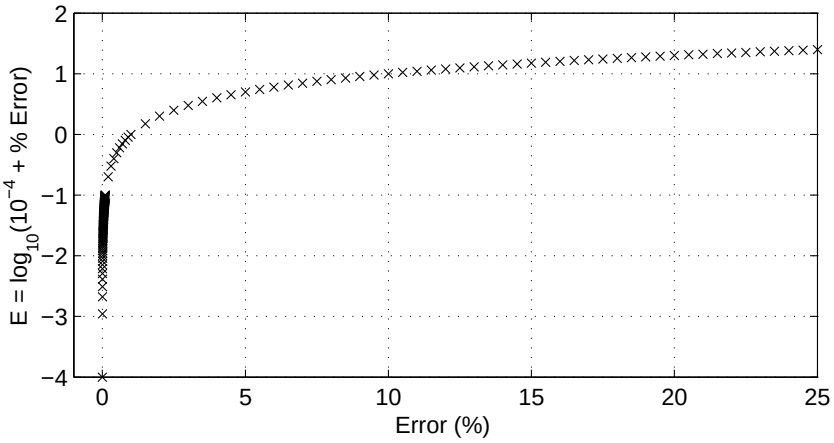
Number of Estimated Components. The number of components that each method was able to estimate per signal is detailed in *Table 3.3*. In most cases the FPT estimated the highest number of components because it seems more sensitive to fast decaying signals; the specific components identified per method have been recorded in *Tables E.1* & *E.2* (*Appendix E*).

	1	2	3	4	Total
HSVD	18(21,50)	17(21,50)	17(19,50)	18(22,50)	70
HLSVD	17(19,36)	15(20,32)	17(19,31)	18(21,35)	68
HTLS	17(21,50)	17(20,50)	17(19,50)	18(21,50)	69
HLTLS	17(19,36)	17(19,32)	17(19,31)	18(20,35)	69
FPT†	20(52,M)	20(59,M)	18(57,M)	20(57,M)	78

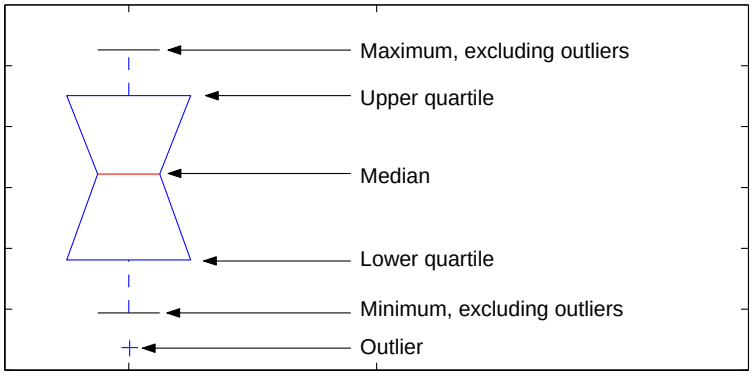
Table 3.3: Partnered estimates. Format: Number of Partnered Estimates (Number of Estimates Within ROI, Number of Estimated Components). $^{\dagger}M = N/2$ or $N/2-1$ for the FPT (*Table 3.2*). The HLSVD and HLTLS methods were unable to estimate the maximum number of components in any case.

E & % Error. The box plots, and descriptive statistics, of *Fig. 3.8* summarize the spread of the errors E (*Eq. 3.3.2*) w.r.t. the estimated spectral frequency, amplitude, and decay parameter values. For this analysis each method's error values from all four signals have been pooled together; the total number of values per box plot is printed above or below the box plot, and is equal to the tabulated total (*Table 3.3*).

In general the frequency estimates are the most accurate *Fig. 3.8a* & *3.8b*; the median percentage error (red line) per method is less than 0.1%. On closer inspection, the frequency estimates of the FPT are the most accurate as indicated by the highly skewed box plot; the frequency-error graphs of *Fig. 3.10* illustrate this point. However, because most of the extra components estimated by the FPT are fast decaying components, and because the accuracy level of these estimates is relatively lower than those of slow decaying components − as the arrow indicated examples of *Fig. 3.10a* & *3.10b* suggest − the FPT's median [frequency estimates] error is relatively high. The frequency estimates of the SVD based methods are quite similar, as the similarity between their error box plots, and their corresponding descriptive statistics, illustrates.



(a) The error grid.



(b) The information provided by the box plots.

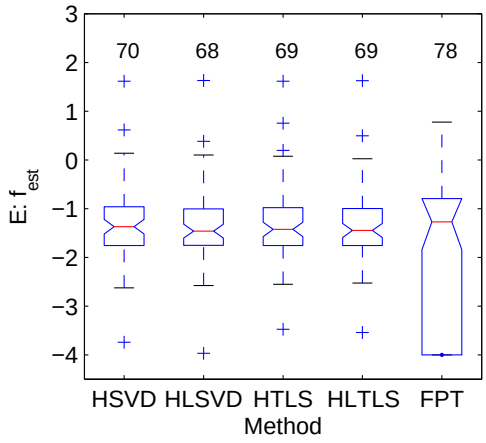
Figure 3.7: The error grid, and the format of the box plots used to summarize the errors evaluated per analysis method (refer to *Fig. 3.8*). For a given set of error values, the box plots summarize the error values as defined in **(b)**.

The outliers and the data points, w.r.t. the errors of the frequency estimates, greater than the 75^{th} percentile are probably noise rather than genuine peaks. However, assuming that these points represent

- poorly estimated components in the case of the SVD methods, and
- components whose parameter values are still converging in the case of FPT analysis.

then the upper error boundaries of the SVD frequency estimates are quite high, [41.22 42.453]%, but that of the FPT frequency estimates, 5.955%, is perhaps acceptable. Most of the data points lie within the inter-quartile range. The inter-quartile range boundaries of the percentage errors are approximately [0.017 0.109]% for the SVD based methods, and [0 0.16]% for the FPT. Both are reasonable ranges.

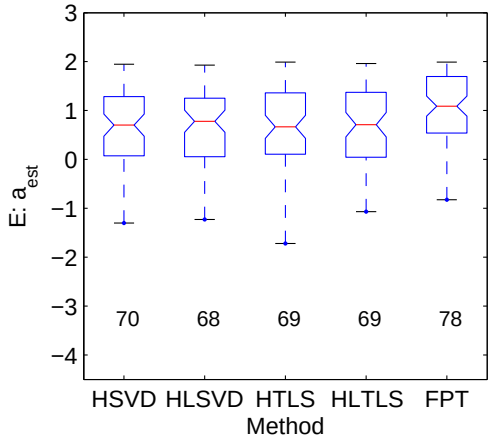
The amplitude estimates of the SVD based methods are the most accurate estimates. Even though the FPT amplitude estimates of fast decaying components are again less accurate than those of slow decaying components, as *Fig. 3.9* highlights, *Fig. 3.9* also illustrates that the SVD



(a) Noise Level: 5%

	HSVD	HLSVD	HTLS	HLTLS	FPT
n	70	68	69	69	78
p_{75}	0.109	0.099	0.105	0.101	0.160
p_{25}	0.017	0.017	0.017	0.017	0
iqr	0.092	0.081	0.087	0.083	0.160
med	0.043	0.034	0.037	0.036	0.053
max	41.314	42.453	41.22	42.338	5.955
min	0.000	0.000	0.000	0.000	0

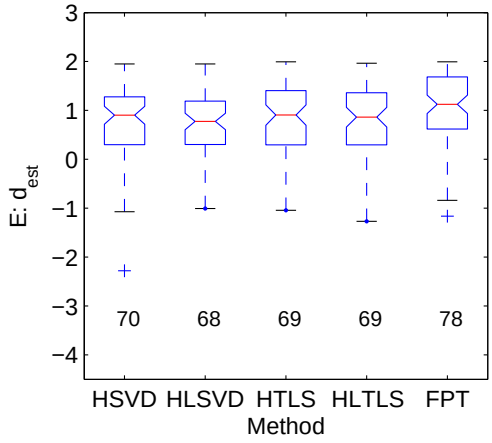
(b) Desc. statistics of **(a)**'s corresponding % Errors



(c) Noise Level: 5%

	HSVD	HLSVD	HTLS	HLTLS	FPT
n	70	68	69	69	78
p_{75}	19.348	17.87	23.002	23.536	49.552
p_{25}	1.189	1.134	1.281	1.105	3.448
iqr	18.159	16.737	21.721	22.43	46.104
med	5.085	6.037	4.650	5.149	12.354
max	88.676	84.597	98.112	91.558	97.8
min	0.050	0.059	0.019	0.085	0.149

(d) Desc. statistics of **(c)**'s corresponding % Errors



(e) Noise Level: 5%

	HSVD	HLSVD	HTLS	HLTLS	FPT
n	70	68	69	69	78
p_{75}	19.018	15.517	25.499	22.999	48.341
p_{25}	2.000	2.010	1.983	1.972	4.158
iqr	17.017	13.508	23.516	21.027	44.183
med	7.986	5.995	8.041	7.280	13.302
max	89.545	89.151	98.41	92.315	98.432
min	0.005	0.098	0.090	0.054	0.069

(f) Desc. statistics of **(e)**'s corresponding % Errors

Figure 3.8: Reproducibility Assessment. n: Number of estimates, p_{75}: 75^{th} percentile, 25^{th} percentile, iqr: inter-quartile range, med: median, max: maximum, min: minimum, Desc.: Descriptive, f_{est}: Frequency Estimates, a_{est}: Amplitude Estimates, d_{est}: Decay Constant Estimates.

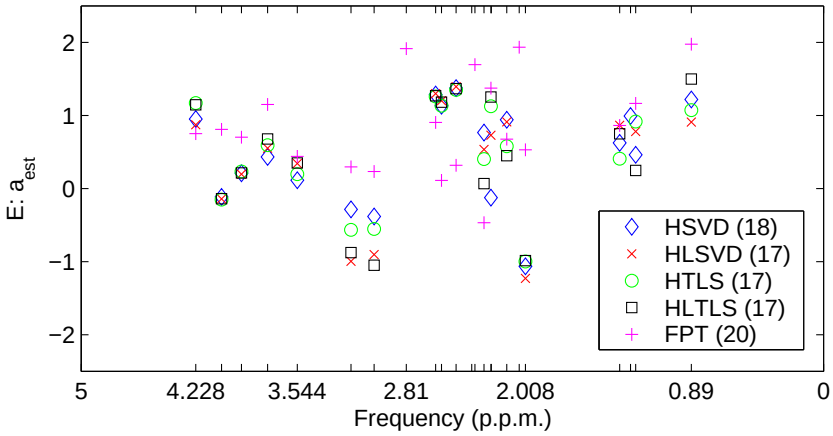


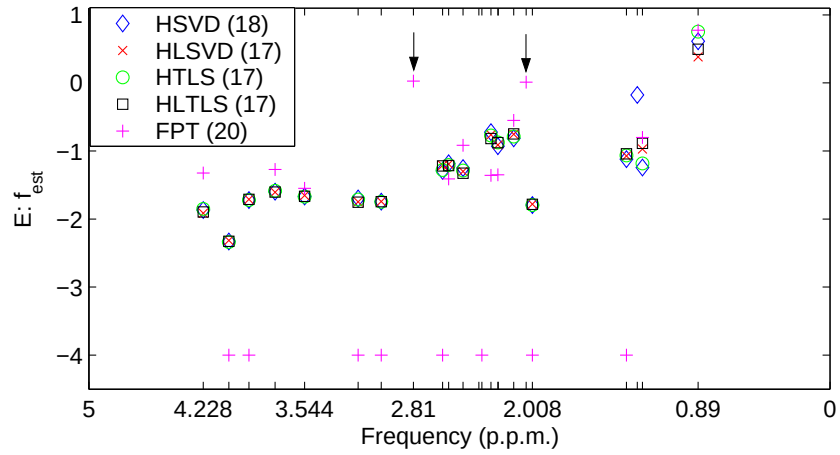
Figure 3.9: An example of estimated amplitude errors per method. The errors E per method are w.r.t. the components that each method was able to estimate.

estimates are usually more accurate. The FPT's less accurate estimates is probably due to the important convergence observation made within *Step 1* of the FPT (*page 43*), i.e. that

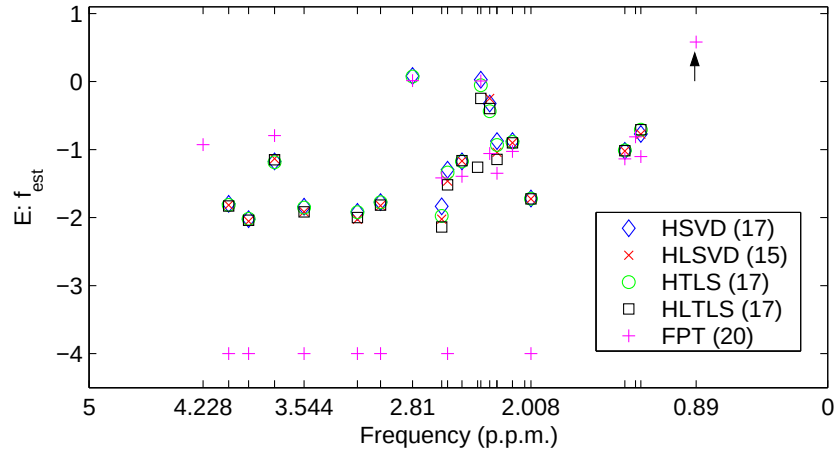
> *The convergence rate of the amplitude estimates is slower than that of the frequency estimates*

Therefore, for the amplitude estimates to have the same level of accuracy as the frequency estimates, relatively more well-sampled data points are required for the amplitude estimates. Especially if the signal is noisy. This point also applies to the decay constant estimates, as highlighted within *Step 1* of the FPT (*page 43*). Although, considering the fact that the convergence rates of the FPT's decay constant estimates seem to be higher than those of the amplitude estimates, much more accurate decay constant estimates, than those of *Fig. 3.8e* & *3.8f*, were expected. The box plots of *Fig. 3.8e* indicate that the SVD decay constant estimates are more accurate than those of the FPT.

In general the amplitude and decay constant estimates are not as accurate as the frequency estimates.



(a)



(b)

Figure 3.10: These graphs summarize the difference between an estimated spectral frequency value and the original, per signal analysis method, and w.r.t. two of the artificial noisy signals. The errors are w.r.t. the spectral frequencies that each method was able to estimate. E: errors (*Eq. 3.3.2*).

3.3.4 Results: The Effect of Noise on the Decomposition Accuracy

Number of Estimated Components. The number of accepted and partnered estimated components is summarized in *Table. 3.4*. The number of partnered estimated components per SVD analysis method decreased as the noise level increased. This was expected. As the noise level increased it became increasingly difficult for the SVD based methods to **(1)** detect fast decaying signals, and **(2)** detect mono-components characterized by low amplitudes and relatively high decay rates. The behaviour of the SVD based methods is in line with observations from other studies [109]: **(a)** the accuracy level is excellent at high SNR, **(b)** at low SNR the SVD based methods find it difficult to estimate or detect the components of a signal, and **(c)** the estimates obtained from low SNR signals usually lack physical meaning. The specific components identified per method have been recorded in *Tables E.3* & *E.4*. The FPT numbers are rather interesting, a similar number of components were estimated per noise level. The results of *Fig. 3.11* & *Table 3.5* suggest it is the accuracy of the FPT's detections, rather than the ability to detect, that is most affected by increasing noise levels. If the ability to detect a mono-component depends on the number of well-sampled data points, and the SNR, of a signal then the current observations suggests that there are more than enough points for detecting the high SNR and/or slow decaying mono-components.

	5%	10%	15%	20%
HSVD	17(20, 50)	14(18, 50)	11(17, 50)	9(18, 50)
HLSVD	18(21, 38)	14(18, 40)	11(16, 41)	8(18, 49)
HTLS	18(19, 50)	16(17, 50)	14(17, 50)	10(15, 50)
HLTLS	18(19, 38)	15(17, 40)	14(16, 41)	10(16, 49)
FPT†	20(65, M)	19(58, M)	19(59, M)	19(56, M)

Table 3.4: Partnered estimates. Format: Number of Partnered Estimates (Number of Estimates Within ROI, Number of Estimated Components). $^{\dagger}M = N/2$ or $N/2-1$ for the FPT (*Table 3.2*). The HLSVD and HLTLS methods were unable to estimate the maximum number of components in any case.

E & % Error. Considering the number of estimates being compared per method, comparing the SVD based estimates with the FPT estimates is not straightforward. In order to give a clearer picture of accuracy levels, the graphs of E *versus* spectral frequency (instead of box plots) have been drawn per noise level N_L (*Fig. 3.11*). The corresponding descriptive statistics of the actual percentage errors, i.e. $10^{E} - 10^{-4}$, are detailed within *Table 3.5*. Once again, the frequency estimates are the most accurate estimates.

Generally, the accuracy of the spectral frequency estimates, per method, decreases with increasing noise level. It is important to note that the accuracy of the FPT frequency estimates decreases at a much faster rate than those of the SVD (*Fig. 3.11a*), albeit the percentage error ranges are still good (*Table 3.5*), and there are much fewer SVD estimates. The observed decreasing accuracy of the FPT frequency estimates, with decreasing SNR, is probably due to an insufficient number of well-sampled data points as suggested within *Step 1* of the FPT,

page 43, i.e. *the accuracy of the genuine poles of the FPT, via which the mono-component parameter values are estimated, will decrease with decreasing signal SNR if there are not enough well-sampled signal data points to ensure the convergence of each genuine pole*

The fact that the accuracy of most, rather than a select few, of the frequency estimates decreases with increasing noise level, supports this suggestion.

There is no discernible error trend amongst the estimated amplitudes of any method. This does not necessarily imply that amplitude estimates are similar regardless of noise level. The lack of an observable trend might be due to the way the amplitudes are estimated and/or an insufficient number of well-sampled signal data points. In the case of the FPT, for example, it has already been noted that the mono-component parameter values converge to their genuine values, and the rate of convergence of the amplitude parameters is the slowest amongst the 3 main parameter types. Additionally, the noisier a signal is, the more the number of well-sampled data points required for accurate parameter estimates. Bearing these points in mind, and assuming that convergence behaviour is similar faraway from a convergence point, regardless of noise level, but differs with noise level closer to a convergence point − then the $N_L : 5\%$ graph of *Fig. 3.11a* suggests that the FPT amplitude estimates are far from their convergence points, considering their accuracy levels, and hence the number of well-sampled points is insufficient.[6] Given that the artificial noisy signals are of the same length, if the 5% level estimates are far from their convergence points, it is quite probable that the estimates at higher noise levels are also far from their convergence points, hence the similar accuracy levels across the varying noise levels.

For perhaps the same reasons as above, there is no discernible trend amongst the estimated decay constants of any method.

[6] This is perhaps comparable to the behaviour of the Newton minimization algorithm (*Appendix A.2*), which is rather haphazard faraway from the minimum point, but smooth and fast closer to the minimum point [40,68].

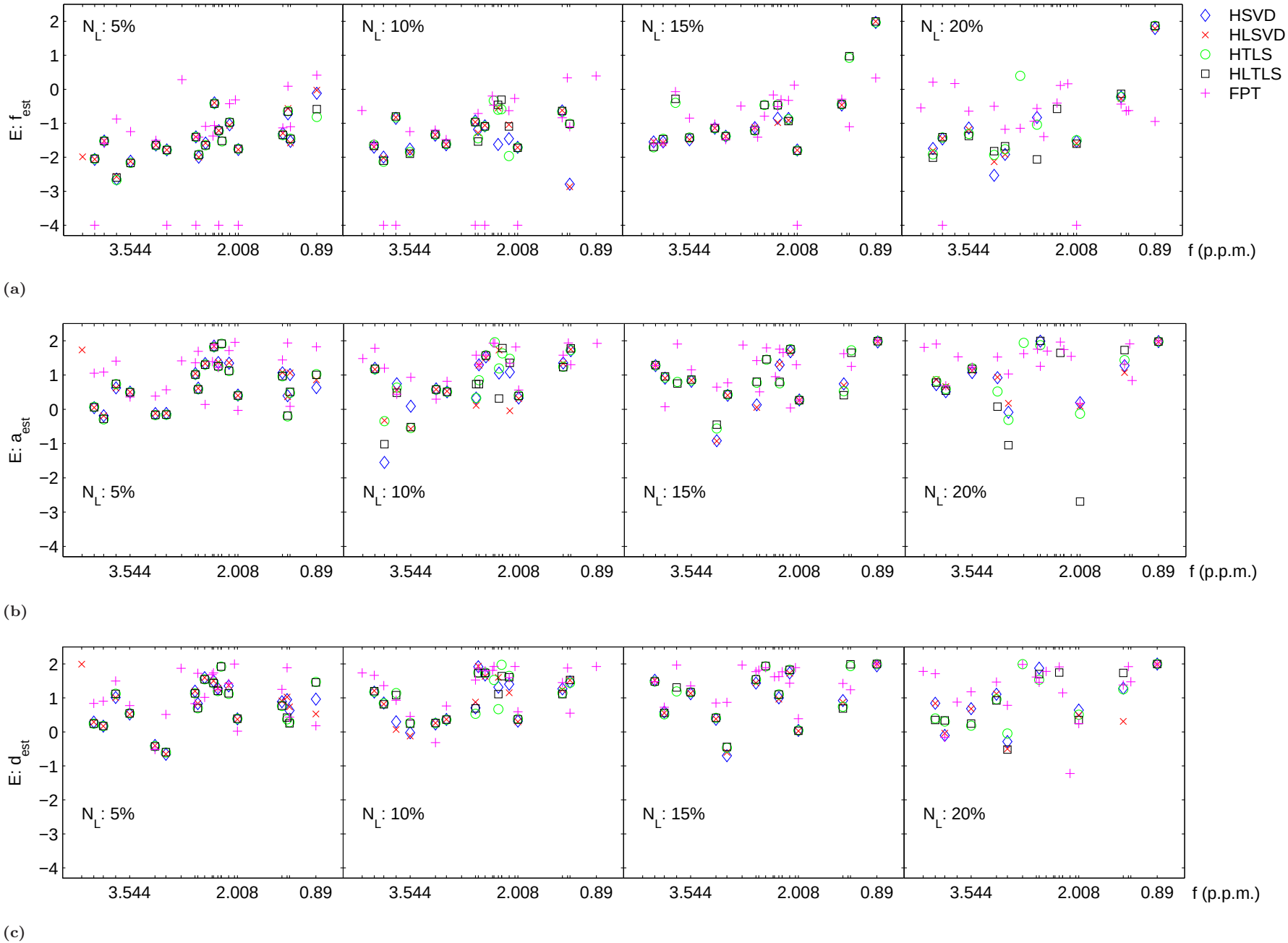


Figure 3.11: Accuracy and increasing noise levels. N_L: Noise level.

		5%					10%					15%					20%				
		n	p_{75}	p_{25}	med	max	n	p_{75}	p_{25}	med	max	n	p_{75}	p_{25}	med	max	n	p_{75}	p_{25}	med	max
f_{est}	HSVD	17	0.068	0.015	0.029	0.758	14	0.084	0.020	0.029	0.231	11	0.140	0.029	0.071	92.174	9	0.260	0.017	0.034	61.088
	HLSVD	18	0.061	0.010	0.026	0.939	14	0.108	0.019	0.049	0.282	11	0.117	0.030	0.072	98.749	8	0.299	0.014	0.030	61.352
	HTLS	18	0.060	0.016	0.030	0.380	16	0.184	0.021	0.063	0.464	14	0.359	0.038	0.101	92.101	10	0.594	0.017	0.043	73.386
	HLTLS	18	0.061	0.016	0.030	0.380	15	0.145	0.022	0.080	0.491	14	0.360	0.038	0.094	98.751	10	0.264	0.015	0.032	73.741
	FPT	20	0.254	0.013	0.065	2.584	19	0.315	0.006	0.079	2.472	19	0.505	0.048	0.162	2.135	19	0.386	0.081	0.237	1.628
a_{est}	HSVD	17	14.104	2.155	4.211	69.918	14	19.868	2.134	8.474	51.296	11	19.467	2.052	6.578	92.435	9	37.801	2.827	8.370	97.330
	HLSVD	18	21.680	2.323	5.712	70.178	14	21.834	1.309	3.762	55.810	11	20.405	2.009	7.056	96.751	8	12.566	3.059	8.044	93.927
	HTLS	18	13.068	1.172	4.650	81.785	16	33.805	2.892	10.836	91.080	14	27.897	3.371	6.636	98.135	10	74.405	3.304	11.243	95.128
	HLTLS	18	13.287	1.147	4.649	82.722	15	21.293	2.595	5.433	61.082	14	28.662	2.727	6.866	96.839	10	53.786	1.192	10.535	98.945
	FPT	20	50.429	3.067	23.666	90.000	19	64.621	10.372	30.189	86.496	19	54.053	4.817	19.032	98.500	19	62.327	17.230	35.036	97.600
d_{est}	HSVD	17	16.550	2.303	7.687	39.153	14	25.121	2.072	10.966	82.225	11	31.492	2.717	9.982	87.523	9	33.270	3.489	6.938	98.546
	HLSVD	18	19.221	2.338	7.480	98.191	14	33.057	1.979	10.843	85.463	11	31.495	2.703	9.181	99.395	8	9.922	1.472	3.937	97.558
	HTLS	18	15.946	1.933	5.353	83.114	16	39.020	2.904	13.765	93.997	14	67.076	3.275	15.054	97.701	10	34.318	2.024	6.140	99.539
	HLTLS	18	16.008	1.778	5.454	83.030	15	38.610	2.971	12.868	54.165	14	66.640	3.602	17.458	99.602	10	54.092	2.180	5.311	99.118
	FPT	20	50.550	4.600	14.074	98.738	19	71.670	6.452	39.397	84.088	19	74.274	18.598	41.748	98.656	19	60.270	9.165	29.160	99.112

Table 3.5: The descriptive statistics of the % Errors. n: Number of estimates, p_{75}: 75^{th} percentile, 25^{th} percentile, med: median, max: maximum, f_{est}: Frequency Estimates, a_{est}: Amplitude Estimates, d_{est}: Decay Constant Estimates.

3.4 Conclusions

This chapter has summarized the most widely used NMRS signal analysis methods, and the FPT. Within the non-interactive methods sections, problem areas were highlighted within each method. A comparative analysis of the non-interactive methods highlighted the advantages and disadvantages of each method.

Due to the sometimes slow rate of convergence of the FPT's estimated parameter values, the accuracy level of the FPT's estimates is sometimes low. Consequently, the SVD based methods tend to have much more accurate estimates at high SNR, and sometimes at low SNR, however there are much fewer SVD estimates at low SNR. The latter point highlights the fact that SVD based methods find it difficult to detect the mono-components of a signal at low SNR [69]. The lower accuracy of the FPT due to slower convergence, w.r.t. noisy signals, can be countered by increasing the number of well-sampled signal data points either by

- reducing the sampling time, thereby acquiring more data points within a time period, and hence retaining more continuous-signal information, or
- increasing the length of the acquired signal without changing the sampling time, which will only be useful if a shorter signal means acquiring a signal that does not fully decay.

The slow convergence of the FPT estimates tempers the other advantages of the method. The FPT can detect low concentration components when the noise level is high, albeit at a lower level of accuracy. The FPT analysis method is independent of line-shapes, i.e. it can extract line-shapes of any known or unknown form from a multi-component signal, therefore peculiar line-shapes do not need to be re-modelled. The FPT can extract fast decaying components that dominate baselines. The FPT is a non-linear analysis method.

A core objective of this thesis was the development of an automatic classification method that can automatically classify the decompositions of the type discussed within the non-interactive methods, and comparative analysis, sections. A new automatic classification method is introduced next.

Chapter 4

A Novel Classification Method

The objective of this thesis, as outlined in the introductory chapter, is the development of a new analysis approach that

1. relies on a unique, objective, and non-linear, signal analysis method that does not assume the nature of the signal being analysed;
2. exploits prior knowledge, but not via a fitting method;
3. includes an automatic classification method if the signal analysis method estimates the components of an NMRS signal.

Only the Fast Padé Transform (previous chapter), which estimates the mono-components of an NMRS signal, meets all the conditions of **Point 1**, and hence a classifier is required to identify the metabolite associations of its estimates. This chapter introduces a prior knowledge based classification method for this purpose, i.e. this chapter addresses **Points 2** & **3**.

In light of what has been discussed thus far, the nature of the time series multi-component signal acquired from a ^{1}H NMRS experiment generally depends on (i) noise [66], (ii) temperature effects [30], (iii) pulse sequence effects [32,102], (iv) 180° refocusing pulse effects [102], (v) magnet strength [5], (vi) magnetic field inhomogeneities [5, 61, 66], (vii) metabolite concentration [53, 66], and (viii) a metabolite's coupling constants, chemical shifts, and hence coupling type. These effects affect either the number of deducible mono-components or the estimable value of a mono-component's parameters (e.g. amplitude, phase, decay constant or frequency). The estimated parameters are also affected by an analysis method's baseline sensitivity, resolution power, and signal-to-noise (SNR) sensitivity [66]. In general, the functional forms of these effects, on a metabolite's estimated mono-component parameters, are either linear, non-linear, or unknown [3,32,66]. Therefore in order to classify the estimated mono-components of an NMRS signal, the classification feature has to be selected carefully. Analytically, the best understood mono-component parameter is the spectral frequency parameter, as *Section 2.4* illustrates. The factors affecting the spectral frequency parameter value − temperature [28], pH-value [5], and coupling constant, chemical shift & magnet strength [1, 49, 60] − have been studied in detail. For example, the afore-discussed QM model of the NMRS behaviour of metabolites [49, 60], which predicts the spectral frequencies of a metabolite, explicitly models

the combined effect of coupling constants, magnet strength, and chemical shifts on a metabolite's spectral frequency values (*Section 2.4*). Secondly, the observed linear effect of temperature on spectral frequency values [28] can be used to adjust for temperature effects on spectral frequencies. Finally, the known pH dependency of certain metabolites [5,43,65], albeit of unknown functional form, may be used to inform decisions w.r.t. unexpected spectral frequencies. On the other hand, the relationship between the observed amplitudes of a metabolite's estimated mono-components and two key affecting factors − concentration [53], pulse sequence parameters [3, 32] − is neither clearly defined nor understood. Consequently, using a combination of estimated spectral frequencies *and* their corresponding estimated amplitudes, as the classification features, is not currently feasible. Therefore the most appropriate feature for classifying the estimated components of an NMRS signal is the estimated spectral frequency. Hence this chapter introduces a classification method that classifies the estimated components of an NMRS signal by comparing the estimated spectral frequencies with the QMP spectral frequencies of 32 known human brain metabolites, and other compounds. The method exploits the distinct spectral frequency combinations per metabolite as predicted by the QM model of metabolite NMRS behaviour. Additionally, because the QMP intensities of a metabolite are probabilistic [49, 60], the QMP intensities values are exploited for probabilistic purposes.

In general, bearing in mind the selected model feature (spectral frequency) and the factors which affect it, the following were considered critical to the classification method ***(a)*** *a supervised classification approach in order to exploit prior knowledge i.e. spectral frequency patterns,* ***(b)*** *invariance to temperature affects,* ***(c)*** *not too susceptible to break in cases of missing data due to low SNR, noise, substantial baseline, etc,* ***(d)*** *probabilistic predictions to aid assignations and rejections during the decision phase, and to aid novelty detection (e.g. the detection of new biochemicals), and* ***(e)*** *the ability to assign mono-components of the same spectral frequency value to the appropriate metabolite.* The classification method introduced herein takes these critical issues into account. It consists of a conditional probability model, discussed next, and a decision algorithm, which is discussed thereafter. Afterwards a statistical test commonly used in detection theory for differentiating between genuine and noise signals is discussed; the purpose of this test is to filter out some of the noise peaks amongst the estimated mono-components prior to classification. Finally the classification method's strengths and weaknesses are explored via an example that uses artificial noisy signals. How the classification method, and the statistic test, fit into an overall analysis approach is summarized in *Box 1* (*page 86*) The classification method is assessed using real data in the following chapter.

4.1 The Conditional Probability Model

A simple expression of the problem in question is

> What is $p(3.185 \text{ p.p.m.} \in \text{Choline} \mid \text{frequency } 3.185 \text{ p.p.m., and other frequencies})$?

Consider a set of N estimated spectral frequencies

$$\mathbf{f} = (f_1, f_2, \ldots, f_n, \ldots, f_N)^T, \quad n = 1, \ldots, N \tag{4.1.1}$$

each probably associated with one, or more, of the metabolites C_k, $k \in 1, \ldots, K$. A standard multi-class probabilistic modelling choice, in statistics, for the above problem type is the *normalized exponential function* [20][1]

$$p(\{f_n\} \in C_k|\mathbf{f}, \mathbf{w}_k) = \frac{\exp(\mathbf{w}_k^T \boldsymbol{\phi}_n)}{\sum_j \exp(\mathbf{w}_j^T \boldsymbol{\phi}_n)} \tag{4.1.2}$$

$$k = 1, \ldots, K, \qquad n = 1, \ldots, N$$

Eq. 4.1.2 evaluates the probability that each f_n of $\boldsymbol{f}$, denoted $\{f_n\}$, is associated with a metabolite C_k. Because we have N estimated frequencies, *Eq. 4.1.2* leads to a $N \times 1$ vector of probabilities per metabolite C_k.

The term $\mathbf{w}_k$ is the weight parameter vector associated with C_k. The weights are usually estimated by defining a distribution over the weights, and subsequently estimating the weights via a Bayesian, or other, probabilistic modelling approach [20]. Each defined distribution, which will be per metabolite in this case, should characterize the properties of a metabolite, e.g. the time series features of the metabolites time series signal.

Each $\boldsymbol{\phi}_n$ - explicitly $\boldsymbol{\phi}_n = \boldsymbol{\phi}(f_n)$, i.e. it is a function of f_n - is a basis function, each represents a relationship between an estimated frequency f_n and each estimated frequency of $\mathbf{f}$, i.e. each $\boldsymbol{\phi}_n$ is also a $N \times 1$ vector. The basis functions are members of a $N \times N$ Kernel matrix Φ. Explicitly

$$\begin{aligned} \Phi &= Kernel(\mathbf{f}, \mathbf{f}) \\ &= [\phi_1 \ \boldsymbol{\phi}_2 \ \ldots \ \boldsymbol{\phi}_n \ \ldots \ \boldsymbol{\phi}_N], \qquad n = 1, 2, \ldots, N \end{aligned} \tag{4.1.3}$$

Kernel functions, that evaluate kernel matrices, are designed to exploit intrinsic data features [90]. For the current problem, we require a kernel function which can exploit the distances between the frequency points of a metabolite's spectrum, and which is invariant to frequency shifts (e.g. due to temperature effects). Hence the Gaussian kernel function

$$\boldsymbol{\phi}_n = \exp\left(-\frac{\|\mathbf{f} - f_n\|^2}{\sigma^2}\right) \tag{4.1.4}$$

which is translation invariant [90], is an appropriate kernel function for the problem in question. The denominator σ is the kernel width; the smaller the value the lower the flexibility of the model, $\sigma = 1$ was used throughout this text. If the QMP spectral frequencies are to be used as prior knowledge *Eq. 4.1.4* should be adjusted to reflect this. Hence, let

- $\mathbf{f}$ be the $N \times 1$ vector of estimated frequencies resulting from an NMRS time series signal analysis, i.e. $\mathbf{f} = (f_1, f_2, \ldots, f_n, \ldots, f_N)^T$, $n = 1, \ldots, N$.

- $\boldsymbol{\psi}_k$ be the $R \times 1$ vector of QMP spectral frequencies (p.p.m.) w.r.t metabolite C_k, i.e. $\boldsymbol{\psi}_k = (\psi_{1k}, \psi_{2k}, \ldots, \psi_{rk}, \ldots, \psi_{Rk})^T$, $r = 1, \ldots, R$.

[1] N.B. N now refers to the number of spectral frequencies, or corresponding amplitudes, etc., estimated by a time series decomposition method, hence the number of estimated components.

Then in order to use the QMP spectral frequencies $\boldsymbol{\psi}_k$, of metabolite C_k, as prior knowledge the kernel matrix $\Phi_k = Kernel(\mathbf{f}, \boldsymbol{\psi}_k)$ was evaluated instead. Explicitly

$$\Phi_k = \left[\exp\left(-\frac{\|\mathbf{f} - \psi_{1k}\|^2}{\sigma^2}\right), \ldots, \exp\left(-\frac{\|\mathbf{f} - \psi_{Rk}\|^2}{\sigma^2}\right)\right] \tag{4.1.5}$$

remembering that $\psi_{1k}, \ldots, \psi_{Rk} \in \boldsymbol{\psi}_k$

i.e. *Eq. 4.1.5* is a $N \times R$ kernel matrix formed between the N estimated frequencies of $\boldsymbol{f}$ and the R QMP frequencies of $\boldsymbol{\psi}_k$ (of metabolite C_k); the subscripts k are used to identify the metabolite. Hence, instead of *Eq. 4.1.4*, each column basis vector is of the form

$$\exp\left(-\frac{\|\mathbf{f} - \psi_{rk}\|^2}{\sigma^2}\right) \tag{4.1.6}$$

$r = 1, 2, \ldots, R$

i.e. the basis functions of the kernel matrix are now functions of both the estimated frequencies and the QMP frequencies, not estimated frequencies only. For later use, $\boldsymbol{\phi}_{nk}$ is used to denote the $1 \times R$ n^{th} row basis vector of kernel matrix Φ_k. The R QMP spectral frequencies of a metabolite are deduced from the QMP time series signal of the metabolite. The QMP time series signal of each metabolite is predicted using NMR SCOPE [74, 75], as discussed in *Section 2.4* (*QM based NMRS Predictions for Prior Knowledge Purposes*, *page 18*), and the predictions are w.r.t. strong coupling, which implicitly accounts for weak coupling. In summary

- The QMP times series signal per metabolite was predicted using the same parameters used to acquire an experimentally acquired signal whose estimated frequencies are to be classified; this includes predicting a QMP time series signal of the same length as that of the experimentally acquired signal.

- Then each time series signal is decomposed using the optimal settings of the Fast Padé Transform (FPT), as discussed in detail in the FPT section (*Section 3.1.2*, *page 42*) of *Chapter 3*. The QMP time series signals are *noise free*, therefore all of FPT's estimated parameters per metabolite were used as QMP data. As detailed in the FPT section (*Section 3.1.2*, *page 42*), the FPT estimates the frequency, amplitude, decay, and phase values of the mono-component time series signals of a multi-component time series signal.

Referring back to *Eq. 4.1.2*, it was previously noted that the weights are usually estimated by defining a distribution over the weights, and subsequently estimating the weights via a Bayesian, or other, probabilistic modelling approach [20]. Regardless of the approach, the aim is to minimize misclassification via the weights i.e. the weights are more or less discriminants. In the proposed model the weights are a function of a probabilistic feature of the QMP spectral frequencies, as detailed next.

Let

- $\mathbf{w}_k$ be the $R \times 1$ weight vector corresponding to the $R \times 1$ QMP frequencies $\boldsymbol{\psi}_k$ of metabolite C_k. Explicitly, $\boldsymbol{w}_k = (w_{1k}, w_{2k}, \ldots, w_{rk}, \ldots, w_{Rk})^T$, $r = 1, \ldots, R$.

In the proposed model, each weight will be a measure of a QMP spectral frequency's importance within a metabolite, as discussed in the *QM Based NMRS Predictions for Prior Knowledge Purposes* section (*Section 2.4*). This means each weight vector will be a function of a QMP spectral frequency's theoretical, and probabilistic [49, 60], intensity (w.r.t. a pulse sequence type). For example, the theoretical time domain intensities of choline's QMP spectral frequencies at approximately 1.5 tesla are plotted in *Fig.* 2.3, and the plot of weights is superimposed. The graph shows that the QMP frequency at $3.183 p.p.m.$ has the highest intensity, hence its weight, normalized intensity, is the highest. Explicitly, the weight vector of metabolite C_k is defined as

$$\mathbf{w}_k = \frac{1}{\sum_{r=1}^{r=R} a_{rk}} \left(a_{1k}, a_{2k}, \ldots, a_{rk}, \ldots, a_{Rk}\right)^T \tag{4.1.7}$$

whereby $\left(a_{1k}, a_{2k}, \ldots, a_{rk}, \ldots, a_{Rk}\right)^T$ is the $R \times 1$ vector of QMP amplitudes of a metabolite C_k. The R amplitudes correspond to the R QMP frequencies $\boldsymbol{\psi}_k$ of the same metabolite. These defined weights (*Eq. 4.1.7*) are a measure of importance, not observed intensities i.e. whilst the most important frequency of a metabolite, w.r.t. weight and a specific magnet strength, will be present in an acquired signal, it is not necessarily the frequency with the highest amplitude, as discussed in the *QM Based NMRS Predictions for Prior Knowledge Purposes* section (*Section 2.4*). And as discussed in *Section 2.4*, this behaviour is probably due to the differing decay rates of the ^{1}H nuclei of a metabolite, as suggested by the varying effects of pulse sequence timings on a metabolite's spectral pattern [3, 32]. Hence the weights are not a sufficient discriminating factor, and consequently the proposed conditional probability model includes two extra discriminating terms, which will be discussed following the model's introduction.

Hence, let

- $\Gamma_k = Kernel(\boldsymbol{\psi}_k, \boldsymbol{\psi}_k)$ be the $R \times R$ kernel matrix of metabolite C_k, i.e. the kernel matrix formed by the QMP frequencies of C_k.
- $\boldsymbol{\gamma}_{rk}$ be the $1 \times R$ row basis vector of Γ_k.
- $\boldsymbol{\nu}_k$ be the $N \times 1$ vector of standard deviations between each f_n and the QMP frequency $\psi_{rk} \in \boldsymbol{\psi}_k$ of metabolite C_k, that each f_n is closest to, **and**
- $\boldsymbol{\eta}_k$ be the corresponding $N \times 1$ vector of weights associated with each ψ_{rk}, of a metabolite, that each f_n is closest to.

then

$$p\left(\{f_n\} \in C_k | \Phi_k, \mathbf{w}_k\right) = \frac{\boldsymbol{\eta}_k \cdot \boldsymbol{\vartheta}_k \cdot \exp\left[(\Phi_k \mathbf{w}_k) \cdot \boldsymbol{\vartheta}_k\right]}{\sum_{j=1}^{J} \boldsymbol{\eta}_j \cdot \boldsymbol{\vartheta}_j \cdot \exp\left[(\Phi_j \mathbf{w}_j) \cdot \boldsymbol{\vartheta}_j\right]} \quad \forall\ k,\ \ J = K \tag{4.1.8}$$

whereby $\boldsymbol{\vartheta}_k = \ln\left(1 + \frac{1}{1^{-10} + \boldsymbol{\nu}_k}\right)$, similarly for $\boldsymbol{\vartheta}_j$

and $\{f_n\} \Rightarrow$ Each f_n

is the proposed conditional probabilities model for evaluating the probability that each estimated frequency f_n is associated with each metabolite C_k. Each probability of the vector $p\left(\{f_n\} \in C_k|\Phi_k, \mathbf{w}_k\right)$ is a Bernoulli probability w.r.t. metabolite C_k, i.e. $p(f_n \in C_k|\Phi_k, \mathbf{w}_k) = 1 - p(f_n \notin C_k|\Phi_k, \mathbf{w}_k)$. Therefore

$$\left[p\left(\{f_n\} \in C_1|\Phi_1, \mathbf{w}_1\right), \ldots, p\left(\{f_n\} \in C_K|\Phi_K, \mathbf{w}_K\right)\right] \tag{4.1.9}$$

is the $N \times K$ matrix of Bernoulli probabilities; there are N estimated frequencies, and K metabolites. Each row's probabilities are associated with an estimated frequency f_n and

$$\sum_{k=1}^{K} p\left(f_n \in C_k|\Phi_k, \mathbf{w}_k\right) = 1 \tag{4.1.10}$$

The terms

$$\boldsymbol{\vartheta}_k \tag{4.1.11}$$

$$\boldsymbol{\eta}_k.\boldsymbol{\vartheta}_k \tag{4.1.12}$$

are the extra discriminating terms. As the standard deviation ν_{nk} between ϕ_{nk} & γ_{ck} increases, *Eq. 4.1.11* ensures that the corresponding conditional probability $p(f_n \in C_k|\ldots)$ decreases. Secondly, because it is quite probable that more than one estimated frequency will be in the neighbourhood of the same QMP spectral frequency in $\boldsymbol{\psi}_k$, *Eq. 4.1.12* discriminates amongst them via the standard deviation measure ν_{nk}; estimated frequencies f_n with the same closest neighbour $\psi_{ck} \in \boldsymbol{\psi}_k$ will have the same η_{nk} value, hence $\boldsymbol{\eta}_k \cdot \boldsymbol{\vartheta}_k$ is a discriminating term w.r.t. the denominator ν_{nk}. The behaviour of each term is illustrated in *Fig. 4.1*. Thus, *Step*

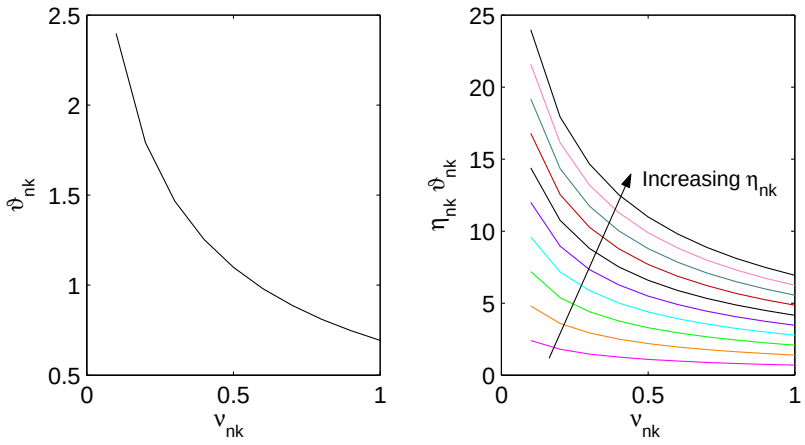


Figure 4.1: The discriminants $\vartheta_{nk} = \ln\left[1 + 1/\left(1^{-10} + \nu_{nk}\right)\right]$, and $\eta_{nk}\vartheta_{nk}$. As $\vartheta_{nk} \to 0$, and as $\eta_{nk}\vartheta_{nk} \to 0$, the conditional probability $p\left(f_n \in C_k|\Phi_k, \mathbf{w}_k\right) \to 0$.

1 of the proposed classification system evaluates the Bernoulli probability that an estimated frequency f_n is associated with a metabolite C_k. Next, a decision has to be made w.r.t. the probability results. There are two main options: either (i) assign each f_n to the metabolite with the highest probability value, or (ii) systematically evaluate the results, e.g. via a defined set of rules, before assigning an estimated frequency to one or more of the metabolites with the

highest probabilities. Because a pair of metabolites may have mono-components with the same spectral frequency, option (i) will be sub-optimal, therefore option (ii) was chosen. Hence, a decision algorithm was developed to systematically evaluate the probability results. To aid the decision algorithm discussion, a new matrix Θ is defined w.r.t. *Eq. 4.1.9*. Explicitly *Eq. 4.1.9* $\Rightarrow$

$$\begin{array}{c} \\ f_1 \\ \vdots \\ f_N \end{array} \begin{array}{c} \begin{array}{ccc} k=1 & \dots & k=K \end{array} \\ \begin{pmatrix} p(f_1 \in C_1|\Phi_1, \mathbf{w}_1) & \dots & p(f_1 \in C_K|\Phi_K, \mathbf{w}_K) \\ \vdots & \ddots & \vdots \\ p(f_N \in C_1|\Phi_1, \mathbf{w}_1) & \dots & p(f_N \in C_K|\Phi_K, \mathbf{w}_K) \end{pmatrix} \end{array} \tag{4.1.13}$$

and hence Θ is defined as

$$\begin{array}{c} \\ f_1 \\ \vdots \\ f_N \end{array} \begin{array}{c} \begin{array}{ccc} p\left(f_n \in C_k|\Phi_k, \mathbf{w}_k\right) \overset{decreasing}{\rightarrow} & & \end{array} \\ \begin{pmatrix} \max\limits_{k\in 1,\dots,K} p(f_1 \in C_k|\dots) & \dots & \min\limits_{k\in 1,\dots,K} p(f_1 \in C_k|\dots) \\ \vdots & \ddots & \vdots \\ \max\limits_{k\in 1,\dots,K} p(f_N \in C_k|\dots) & \dots & \min\limits_{k\in 1,\dots,K} p(f_N \in C_k|\dots) \end{pmatrix} \end{array} = \Theta \tag{4.1.14}$$

i.e. the probabilities of each f_n are in descending order. Let $\boldsymbol{\theta}_i$ be the i^{th} column of Θ, and let $\theta_{ni} \in \boldsymbol{\theta}_i$. Let ε_{ni} refer to the metabolite associated with the i^{th} column of the n^{th} row of Θ, i.e.

$$\begin{array}{c} \\ f_1 \\ \vdots \\ f_N \end{array} \begin{array}{c} \begin{array}{cccc} & & & \end{array} \\ \begin{pmatrix} \varepsilon_{11} & \dots & \varepsilon_{1i} & \dots \\ \vdots & \ddots & & \vdots \\ \varepsilon_{N1} & \dots & \varepsilon_{Ni} & \dots \end{pmatrix} \end{array} = \mathcal{E} \tag{4.1.15}$$

is the $N \times K$ matrix of metabolites parallel to Θ (*Eq. 4.1.14*). Hence, the decision algorithm is introduced.

4.2 The Decision Algorithm

The decision algorithm is based on the idea that the spectral frequencies of a metabolite are of varying importance, and the more important a frequency is, the more probable its presence in an experimentally acquired signal (as considered in *Section 2.4, QM Based NMRS Predictions for Prior Knowledge Purposes, page 18*). As previously discussed, w.r.t. weights $\mathbf{w}_k$, **the importance of a spectral frequency is quantified in terms of its evolution dependent QMP intensity** − the higher the intensity the more important the frequency. For example, referring back to choline's spectral frequencies in *Fig. 2.3* (*page 23*), the 3.183 p.p.m. frequency is considered the most important, and the most likely to be observed in a mixture that includes choline, because the 3.183 spectral frequency has the highest QMP probabilistic intensity. The numbers within *Fig. 2.3*'s plot signify the importance of Choline's spectral frequencies at approximately 1.5 tesla, and $T_E = 35$ ms. The number 1 is allocated to the frequency with the highest amplitude (probabilistic intensity) − henceforth termed *core frequency* − and the

other [secondary] frequencies are numbered accordingly. Hence, each probability of matrix Θ (*Eq. 4.1.14*) has an associated number which signifies the importance of f_n's closest neighbour in $\boldsymbol{\psi}_k \; \forall k$. These numbers are hence termed *indicants*.

Every metabolite has a core frequency, and this leads to the fundamental assumption of the decision algorithm. It is assumed that *in the absence of a metabolite's core frequency it is highly improbable that the metabolite is present in an acquired NMRS signal*. Because each time a metabolite is estimated, or deduced, or considered present in a mixture, the core frequency is also present, as illustrated in *Table 2.2* of the *QM Based NMRS Predictions for Prior Knowledge Purposes* section (*page 21*). Hence, let Λ be Θ's parallel matrix of indicants

$$\begin{matrix} f_1 \\ \vdots \\ f_N \end{matrix} \begin{pmatrix} \lambda_{11} & \dots & \lambda_{1i} & \dots \\ \vdots & \ddots & \vdots & \\ \lambda_{N1} & \dots & \lambda_{Ni} & \dots \end{pmatrix} = \Lambda \tag{4.2.1}$$

$$i \in 1, \dots, K$$

Each λ_{ni} is a record of the importance of the QMP frequency, of metabolite ε_{ni}, that an f_n is closest to. The subscripts in each case, λ_{ni} & ε_{ni}, mean that each term is associated with the i^{th} column of the n^{th} row of Θ. For example, $\lambda_{12} = 3$ is associated with probability $\theta_{12} \in \Theta$, and the importance of the QMP frequency of metabolite ε_{12}, to which f_n is closest, to is 3. Hence, let (Θ, $\mathcal{E}$, Λ) denote the $N \times K \times 3$ matrix of decreasing probabilities Θ (*Eq. 4.1.14*), per f_n, and the corresponding metabolite associations $\mathcal{E}$ (*Eq. 4.1.15*) and indicants Λ (*Fig. 4.2*).

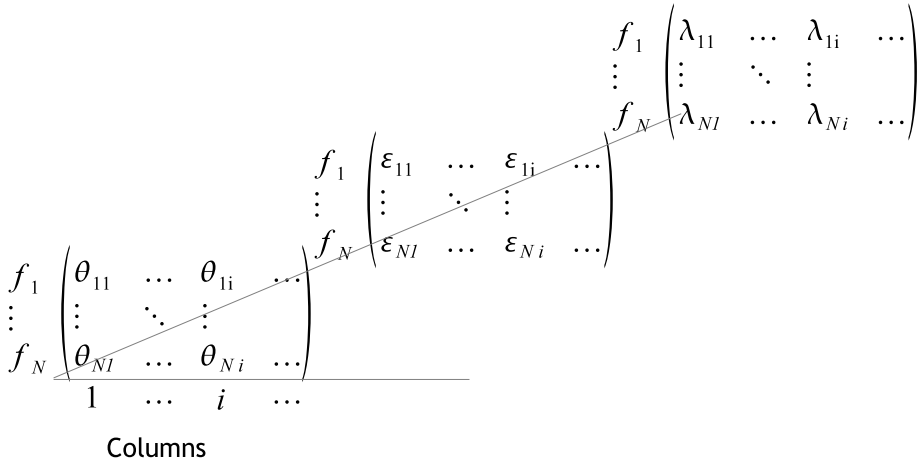


Figure 4.2: The $N \times K \times 3$ matrix, (Θ, $\mathcal{E}$, Λ), of decreasing probabilities Θ (*Eq. 4.1.14*), per f_n,and the corresponding metabolite associations $\mathcal{E}$ (*Eq. 4.1.15*) and indicants Λ (Eq. 4.2.1). In this text each column of this matrix is of size $N \times 1 \times 3$.

Then an estimated component with estimated spectral frequency f_n will be assigned to metabolite ε_{ni}, of a set of metabolites C_k, with regard to the following 4 steps:

Step 1: Extracting Potential $f_n \in C_k$

- The decision algorithm searches the first three columns of $(\Theta, \mathcal{E}, \Lambda)$ for any f_n associated with C_k **and** whose $\lambda_{ni} = 1$. The latter implies that f_n is associated with the core QMP spectral frequency of C_k. **If** such f_n do not exist, all analysis w.r.t. C_k ceases, else
- The decision algorithm then searches the first six columns of $(\Theta, \mathcal{E}, \Lambda)$ for any f_n associated with C_k.

Step 2: d_n Limits

The correlation time τ_c of a molecule − the time it takes a molecule to rotate through one radian [26] − is sometimes used to estimate the de-phasing time constant T_2 (discussed on *page 28*) of the molecule [24, 26, 117]. A widely used estimate of τ_c is [24, 26]

$$\tau_c = \kappa \mathcal{F}_w 10^{-12} \tag{4.2.2}$$

$\mathcal{F}_w$ is the formula weight of the molecule, and κ is a constant proportional to the size of the molecule in question [24]. In the case of the metabolites C_k of the classification method, which are small molecules, $\kappa \approx 10^{-4}$.[2] The relationship between T_2 and τ_c is given by [117]

$$\frac{1}{T_2} = \frac{(-\nu_f)^2 \tau_c}{2 + (-\nu_f \tau_c)^2} \tag{4.2.3}$$

$$\text{whereby } -\nu_f = \gamma(1+f)B \tag{4.2.4}$$

Eq. 4.2.4 is analogous to *Eq. 2.2.8* (*page 12*) but instead of the chemical shift δ we have the maximum spectral frequency of a metabolite. For the problem in question $f = max(\boldsymbol{\psi}_k)$, i.e. the maximum QMP spectral frequency of metabolite C_k. Noting that the decay constant $d = (T_2)^{-1}$, a f_n associated with C_k remains associated if and only if

$$d_n < \frac{(-\nu_f)^2 \tau_c}{2 + (-\nu_f \tau_c)^2} \tag{4.2.5}$$

d_n : the estimated decay value associated with f_n

Afterwards, if there are no more f_n associated with the core QMP spectral frequency of C_k all analysis w.r.t. C_k ceases, else

Step 3: 1-to-1 Mapping

The remaining f_n associated with C_k are compared with the QMP spectral frequencies $\boldsymbol{\psi}_k$ of C_k to ensure that each f_n is associated with a different QMP spectral frequency of $\boldsymbol{\psi}_k$. If two or more f_n values are associated the same QMP spectral frequency the *closest* is accepted.

Steps 1, 2 & ***3*** constitute the first part of the decision algorithm. Once all possible metabolite association have been determined, the decision algorithm then analyses frequencies associated with more than 1 metabolite, i.e.

[2] Generally, a small molecule is a molecule that weighs < 1500Da [116].

Step 4: f_n & Multiple Associations

Multiple associations are resolved, or left unchanged, w.r.t. the following rules

- If a f_n is associated with the core QMP spectral frequency of one metabolite, but with the secondary frequency of another, f_n is assigned to the former metabolite.
- If a f_n is associated with a secondary QMP spectral frequency of two or more metabolites, then it is assigned to the metabolite with the lowest λ_{ni} value.
- If the core QMP spectral frequencies of a pair of metabolites are associated with the same f_n, both results are presented.

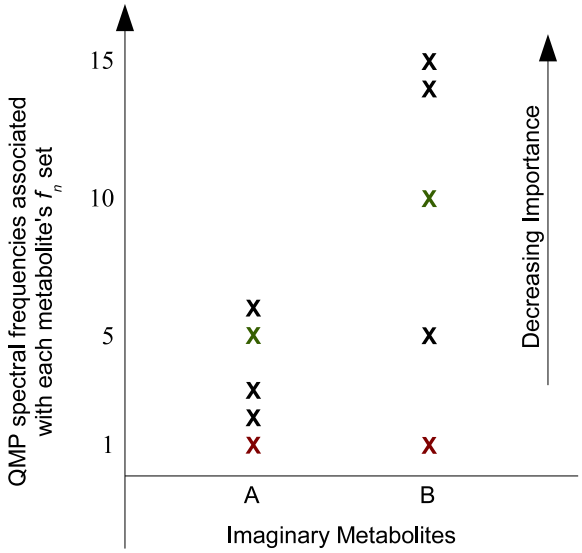


Figure 4.3: Pattern of associations and decision making. Here we have two imaginary metabolites A & B. The graph highlights the importance of the set of f_n associated with each metabolite, based on the importance of each f_n's counterpart QMP spectral frequency. The coloured marks denote cases whereby the same f_n has been assigned to both metabolites. In this case the same f_n has been associated with the core QMP spectral frequency of both metabolites. If we consider the overall pattern of associations Metabolite A is the most plausible option because all the f_n associated with it are also associated with the *top* QMP spectral frequencies of Metabolite A.

The last point requires further explanation. Consider, for example, the QMP core spectral frequencies of creatine and phosphocreatine, 3.027 p.p.m. & 3.029 p.p.m. respectively. Because the frequencies are quite close, there will be cases whereby the same f_n will be associated with the core QMP frequency of both. A user can decide on the most plausible association either by **(1)** comparing the conditional probability values that associate a f_n with the core QMP spectral frequencies, or **(2)** considering the pattern of associations of all the f_n associated with each metabolite (refer to the illustrated example of *Fig. 4.3*). It should be noted again that the FPT, whose estimated components are used herein, extracts overlapping components from a multi-component signal. Hence, if both creatine and phosphocreatine are present in a mixture the FPT will provide separate estimated components to account for each metabolite.

Metabolites & Other Compounds

Common	Other
Acetate (Ace)	Ethanol^{e} (Eth)
Betaine (Bet)	Ethanolamine (EA)
Choline (Ch)	Glycerol (Glyc)
Creatine (Cr)	Glycine^{p} (Glye)
γ-Amino-Butyric Acid (GABA)	Glutathione (GTT)
Glycero-Phosphoryl-Choline (GPC)	L Alanine^{p} (LAl)
L Aspartic (LAsp)	L Phenylalanine^{p} (LPA)
L Glutamic (LGlc) Acid	L Serine (LSer)
L Glutamine (LGle)	LThreonine (LThr)
L Lactate (LLac)	LTryptophan (LTryp)
Myoinositol (MI)	LTyrosine (LTyr)
N-Acetyl-l-Aspartic Acid (NAA)	LValine (LVal)
N-Acetyl-l-Aspartic Glutamic (NAAG) Acid	Methyl Sulphonyl Methane^{e} (MSM)
Phosphocreatine (PCr)	Pyruvic^{p} (Py)
Phosphorylcholine (PCh)	Succinic^{p} (Suc)
Scylloinositol (SI)	
Taurine (Tau)	

Table 4.1: The 32 metabolites/compounds of the classification method whose QMP prior knowledge is used by the classification method. Here they have been divided into commonly detected metabolites and other metabolites and compounds associated with human brain ^{1}H NMRS [5,43,116]. ^{e}Exogenous compounds [5]. ^{p}Compounds usually detected under pathological or abnormal conditions [5]. Lactate is much easier to detect under pathological or abnormal conditions [5,43]; it is sometimes difficult to detect in a healthy human brain because its normal concentration is low [43].

All 4 steps of the decision algorithm have now been outlined. Two important points. First, the classification method uses the QMP data of the 32 metabolites, and compounds listed in *Table 4.1*.[3] And due to the degree of overlap amongst their QMP spectral frequencies – e.g. *Tables F.1, F.2, F.3* & *F.4* – sometimes the conditional probabilities associated with a f_n might be generally low. Therefore specifying a conditional probabilities lower limit is tricky. In this text an f_n metabolite association whereby $p(f_n \in C_k|\ldots) < 0.001$, i.e. whereby the conditional probability is 0 correct to 3d.p., is ignored. Secondly, the results are divided into *commonly detected metabolites* and *other metabolites and compounds* (*Table 4.1*) associated with human brain ^{1}H NMRS, according to the discussions of [5,43], and of the on-line database [116]. The *other metabolites and compounds* consist of metabolites that are rarely detected or difficult to detect, compounds introduced orally (e.g. ethanol via alcohol consumption), and compounds "which may be detected under pathological or other abnormal conditions" [5]. The main results are those related to the commonly detected metabolites.

[3]The chemical shifts and coupling constants used to predict the quantum mechanical data of each metabolite/compound is detailed in *Tables F.1* & *F.2* and *Tables F.3* & *F.4*.

All aspects of the classification method − the conditional probability model and the decision algorithm − have now been discussed, and how they fit into an overall analysis approach is outlined in *Box 1* (*page 86*). Prior to introducing the example, a probability measure used quite widely in signal detection theory is introduced.

4.3 The Generalized Likelihood Ratio Test

Using the optimal settings of the FPT can sometimes produce a relatively (relative to the number of genuine components) large number of estimated components within a spectrum's ROI; as *Tables 3.8* & *3.11* of the *Comparative Analysis of Non-Interactive Methods* section suggests (*Section 3.3*). An option for dealing with a large number of estimated components is to filter out probable noise signals prior to introducing a set of estimated components to a classifier. In signal detection theory it is common to assess whether a detected or estimated signal is a genuine signal [59]. For this text this is achieved via the Generalized Likelihood Ratio Test (GLRT), a *composite hypothesis test statistic* for discriminating between noise and genuine signals [59].[4] In addition to accepting or rejecting a signal w.r.t. a defined hypothesis, the GLRT evaluates the probability of detection of an estimated signal. This is an invaluable measure because it is independent of a classification method but can be used in conjunction with a classification method's results for assessing the plausibility of the classifications. For example, consider the metabolite NAA. In human brain ^{1}H NMRS, the concentration of NAA is [almost always] greater than that of any other metabolite. Therefore we expect the probability of detection of NAA's main peak to be quite high. If however an estimated component classified as NAA's main peak has a low probability of detection − and assuming the component has not been misclassified − then either

1. the subject's NAA level is actually low, or
2. the acquired signal has been corrupted by an experimental error/problem, consequently the classified peak is actually a noise peak that shares some of NAA's characteristics.

In each case the probability of detection acts as a warning signal in relation to a disease, a flawed experiment, or a flawed analysis. Hence, prior to classifying each estimated frequency f_n of $\boldsymbol{f}$ (i.e. as defined on page 75), the estimated time series mono-component signal, denoted $\hat{\boldsymbol{x}}_n$, that each f_n is associated with, will be assessed by the GLRT as follows.

Consider the hypotheses

$$\begin{aligned} &\mathcal{H}_0 : \hat{\mathbf{x}}_n \text{ is a noise signal} \\ &\mathcal{H}_1 : \hat{\mathbf{x}}_n \text{ is a genuine mono-component time series signal of a metabolite} \end{aligned} \tag{4.3.1}$$

by the GLRT [59] $\hat{\boldsymbol{x}}_n$ is accepted under $\mathcal{H}_1$ if

$$max\, I_n > -\frac{\sigma^2_{res}}{\mathcal{C}} \ln(P_\alpha) = \epsilon \tag{4.3.2}$$

[4] A composite hypothesis test statistic is a hypothesis test statistic whose assumed distribution Probability Density Function (PDF), w.r.t. to an hypothesis, has at least one estimated parameter [59].

whereby I_n is the periodogram of $\hat{\boldsymbol{x}}_n$. σ_{res} is the standard deviation of the residue that remains after a signal analysis method has extracted all the *possible* mono-component time series signals from a multi-component NMRS time series signal. $\mathcal{C}$ is the length of $\hat{\boldsymbol{x}}_n$, and P_α is the pre-set probability of false alarm P_α . P_α (also known as Significance Level) is the probability of committing a Type I Error, i.e., and w.r.t. the defined hypotheses, the probability of accepting a signal under $\mathcal{H}_1$, albeit $\mathcal{H}_0$ is true [59]. Hence a small P_α value should ensure that only a minimal number of noise signals are accepted as genuine signals. However if a signal is accepted under $\mathcal{H}_0$ albeit $\mathcal{H}_1$ is true, denoted $P(\mathcal{H}_0;\, \mathcal{H}_1)$, this is a Type II Error [59]. The aim of any decision problem is to minimize $\mathrm{P}(\mathcal{H}_0;\, \mathcal{H}_1)$, or maximize $P_\delta = 1 - P(\mathcal{H}_0; \mathcal{H}_1)$, where P_δ is the probability of detection . In relation to the GLRT, P_δ is defined as [59]

$$P_\delta = p\{I_n > \epsilon; \mathcal{H}_1\} \tag{4.3.3}$$

$$= Q_{\chi_2'^2(\lambda)}\left(2\ln\frac{\mathcal{C}/2-1}{P_\alpha}\right) \tag{4.3.4}$$

whereby

$$\lambda = 10\log_{10}\frac{max(I_n)^2\mathcal{C}}{2\sigma_{res}^2} \tag{4.3.5}$$

$Q_{\chi_2'^2(\lambda)}$ is the Cumulative Distribution Function (CDF) of the non-central chi-squared PDF $\chi_2'^2(\lambda)$ [59]. Throughout this text $P_\alpha = 0.01$. The GLRT is summarized, w.r.t. NMRS signals, in *Appendix A*; a general, and more detailed, discussion of GLRT can be found in [59].

AN OUTLINE OF THE ANALYSIS APPROACH W.R.T. THE CLASSIFICATION METHOD, AND GLRT, DISCUSSED IN THIS CHAPTER

1. An NMRS multi-component time series signal is acquired.
2. Its residual water peak is removed (refer to the example of *Fig. 1.1*, *page 3*), leaving a multi-component time series signal Y, of length $\mathcal{C}$.
3. Y is decomposed via the Fast Padé Transform (FPT) (FPT, *page 42*). Leading to N estimated time series mono-component signals $\hat{x}_n$, each also of length $\mathcal{C}$. Each $\hat{x}_n$ has an estimated frequency f_n, estimated amplitude a_n, and estimated decay value d_n.
4. Each $\hat{x}_n$ is assessed by the Generalized Likelihood Ratio Test (GLRT) (GLRT, *page 84*) in order to (a) determine whether it should be rejected as a noise signal, or passed on for classification, and (b) calculate the probability of detection P_δ of each $\hat{x}_n$.
5. Each remaining $\hat{x}_n$ is classified via its corresponding f_n by the classification method, i.e.
   - the conditional probability, *Eq. 4.1.8*, that an f_n is associated with a metabolite (*The Conditional Probability Model*, *page 74*) is calculated.
   - the calculated probabilities are used by the decision algorithm (*The Decision Algorithm*, *page 79*), which decides the most probable metabolite that an f_n is associated with, and hence the most probable metabolite that $\hat{x}_n$ is associated with.

Box 1: A summary of the pre-classification and classification steps w.r.t. classifying each estimated frequency $f_n \in \boldsymbol{f}$. Classifying a f_n is the same as classifying the corresponding estimated time series signal $\hat{x}_n$.

4.4 An Illustrative Example using Synthetic Data

The aim of this exercise is to illustrate the classification method's ability to correctly classify a set of estimated mono-component time series signals, via their estimated frequencies, as the number of noise components amongst them increases. To this end, this exercise uses the

same set of noisy artificial time series signals, each of length 2048, used for the *Decomposition Accuracy Assessment* in the *Comparative Analysis of Non-Interactive Methods* section (*Section 3.3*). That assessment assessed 4 noisy time series signals consisting of the same artificial *noise free* signal X_{anf} (*Table C.2*), but different levels of noise. The noise levels were 5, 10, 15, & 20% of the highest X_{anf} mono-component amplitude; refer to page 59 of the *Comparative Analysis of Non-Interactive Methods* section for details of how these signals were formed. As detailed in *Table C.2*, 18 of the mono-components are associated with 6 metabolites, whilst the remaining 6 mono-components are associated with unknown macromolecules and lipids [91]. Hence the classification method will be tested w.r.t. the 6 metabolites.

The True Positive Rate (TPR), i.e. Sensitivity, and the False Positive Rate (FPR), i.e. 1 - Specificity, measures are used to assess the classification method's performance. The TPR and FPR are defined as

$$\text{TPR} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}} \tag{4.4.1}$$

$$\text{FPR} = \frac{\text{False Positives (FP)}}{\text{False Positives (FP)} + \text{True Negatives (TN)}} \tag{4.4.2}$$

and the TP, FP, FN, and TN terms are defined by *Table 4.2*. As discussed in the previous section each estimated mono-component is first assessed via the GLRT, and only the estimated components accepted as signals by the GLRT are analysed by the classification method. In order to assess the *decision value* of the GLRT's P_δ values, the TPR and FPR values are evaluated per P_δ threshold, denoted P_δ^τ. Explicitly, P_δ^τ is increased from 0 to 0.9 $[0 : 0.01 : 0.9]$, and the estimated components whose P_δ values are below the current P_δ^τ are automatically categorized as False Negative or True Negative. If the number of false negatives is high whilst P_δ^τ is low (perhaps less than 0.5) then the classified estimated components with low P_δ are probably as important as the estimates with high P_δ values. On the other hand, if the number of false negatives is low whilst P_δ^τ is low, then the classified estimated components with low P_δ values might not be genuine components, therefore their status (genuine or spurious) should be considered carefully. Thus the decision value rationale of P_δ. The actual behaviour of the classification method is observed when $P_\delta^\tau = 0$. The accuracy of the estimated component's parameter values, per artificial noise signal, has already been assessed in *Section 3.3.4*. The focus herein is classification.

		ACTUAL	
		Genuine	Noise
CLASSIFICATION	Genuine	TP	FP
METHOD	Noise	FN	TN

Table 4.2: Receiver Operating Characteristics (ROC) Parameters. The estimated components whose P_δ values are below a specified threshold P_δ^τ will either fall into the FN or TN category

	The Noisy Artificial Signals					
	5%	10%	15%	20%	10%	
Number of Peaks					Ext.	
Within ROI	65	58	59	56	107	cf. *Fig. 3.11b*
Accepted by GLRT	48	30	32	36	53	$\equiv \sum (TP + FN + FP + TN)$
Falsely Rej. by GLRT†	0	1	0	0	0	

Table 4.3: The classification method and artificial noisy signals. †The estimated parameter values of the probably genuine signal rejected by the GLRT are $f = 4.218$, $a = 0.037$ & $d = 0.771$, whereas the original values are $f = 4.228$, $a = 0.053$ & $d = 1.688$. Rej.: Rejected, Ext.: Extended.

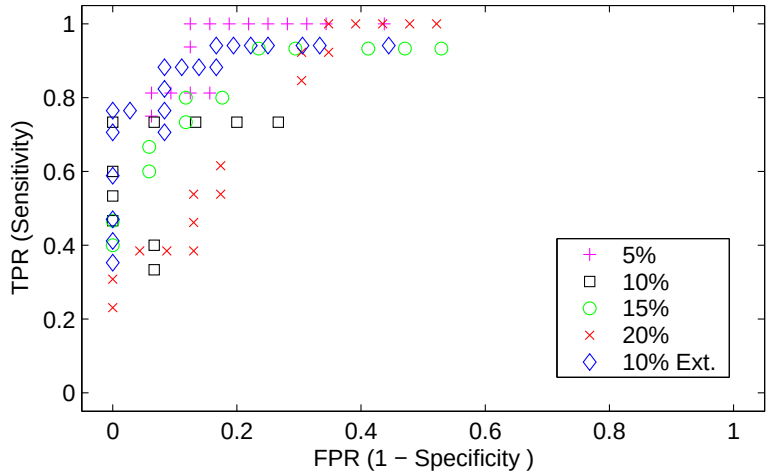
4.4.1 Observations

N.B. An additional noisy signal, 10% Ext., has been included in this analysis. The signal is twice the length of the original four signals, and it is an extended version of the 10% signal. The purpose of this extended signal is introduced, and discussed, as the discussion proceeds.

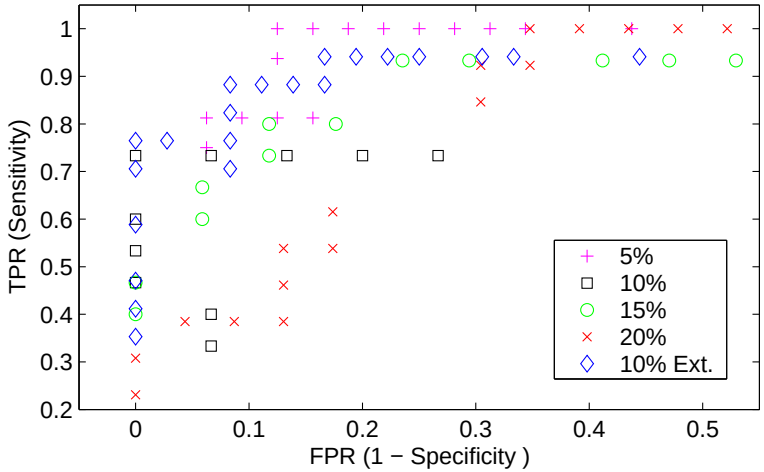
GLRT. The P_α (probability of false alarm or significance level) of the GLRT is quite high, $P_\alpha = 0.01$, therefore the GLRT rarely rejects a genuine peak, as *Table 4.3* illustrates. Of course a high P_α means that the classification method would receive a high number of noise peaks relative to the number of genuine signals. However the aim of using the GLRT is to reduce the number of noise signals not eliminate them because the classification method should reject spurious classification candidates.

The ROC. As preferred, the focal point of the ROC points is the top left corner of the ROC graph (*Fig. 4.4a*). The graph of *Fig. 4.4b* zooms in on the focal point, and highlights the areas where each of the noisy signals' ROC points are clustered. There are 91 ROC points per noisy signal − the P_δ range, i.e. $0 : 0.01 : 0.9$, consists of 91 points − however these points are not obvious in the graphs of *Fig. 4.4a* & *4.4b* because the points overlap. The three dimensional graph of *Fig. 4.4c* highlights the hidden points. Initially ignoring the 10% Ext. classifications, the decreasing order of classification performance w.r.t. the noisy signals is $X_{an}^{5\%}$, $X_{an}^{15\%}$, $X_{an}^{20\%}$, and $X_{an}^{10\%}$; $X_{an}^{n\%}$ denotes an artificial noisy signal of noise level $n\%$. The performance w.r.t. $X_{an}^{10\%}$ was not expected. The results of *Table 4.5* indicate why the performance w.r.t. $X_{an}^{10\%}$ is relatively poor − there are fewer TP classifications, and more FN classifications, w.r.t. $X_{an}^{10\%}$ than there are for $X_{an}^{15\%}$ & $X_{an}^{20\%}$.

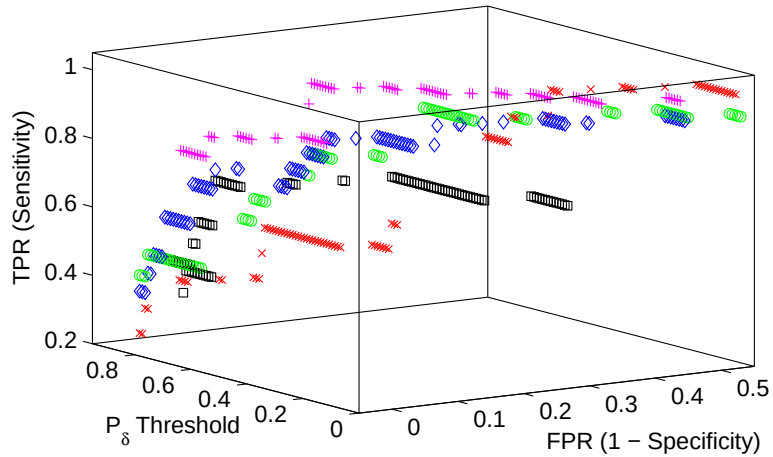
Normally we expect the TP values to decrease as the noise level of a system increases, mainly because we expect more noise components, and less accurate components, to be delivered to the classifier. However the higher number of FN w.r.t. $X_{an}^{10\%}$ suggests that the FPT mono-component estimates of this signal might be poorer than those extracted from the noisier signals. The most obvious reason for this anomaly is noise complexity, i.e. *it is quite possible that noise complexity might affect the detection and/or convergence of the FPT's mono-component poles.* To test this hypothesis, a longer version of the artificial noisy signal $X_{an}^{10\%}$, i.e. 10% Ext. of length $2 \times \text{length}(X_{an}^{10\%})$, was decomposed via the FPT, and its decompositions were also



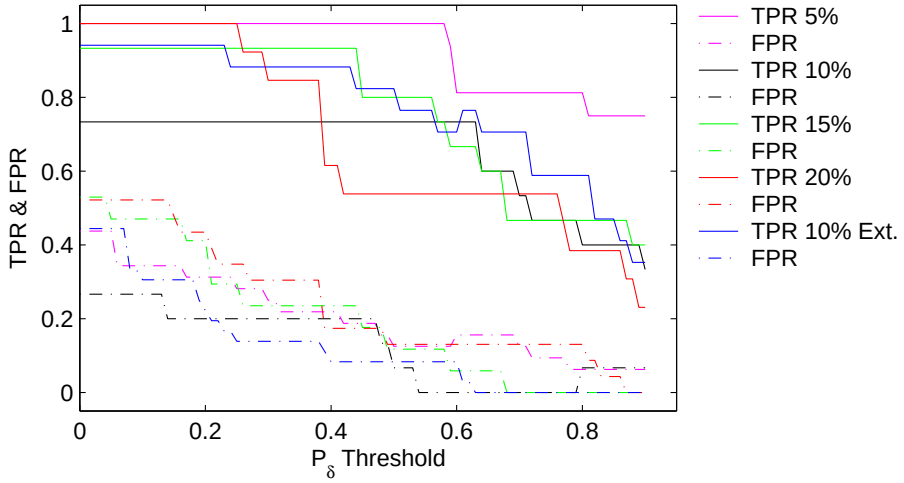
(a)



(b)

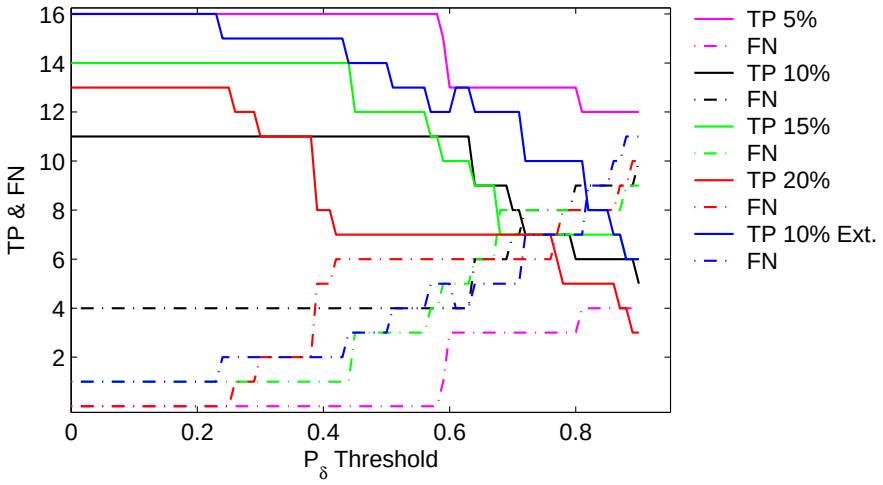


(c)

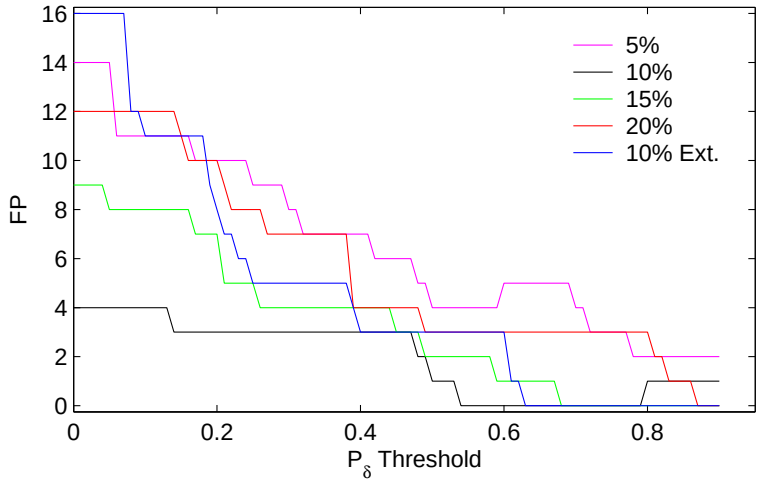


(d)

Figure 4.4: The ROC graphs of the classification method



(a)



(b)

Figure 4.5: The ROC graphs of the classification method

passed on to the classifier. The results w.r.t. the extended signal, 10% Ext., seem to support the hypothesis. The new signal has a higher number of true positive classifications than the original $X_{an}^{10\%}$. (N.B. The first half of the extended signal is exactly the same as the whole original signal $X_{an}^{10\%}$)

The graph of *Fig. 4.4d* gives a clear picture of the effect of P_δ^τ on the TPR, and the FPR. Both rates decrease as P_δ^τ is increased, i.e. the classifier's sensitivity (TPR) decreases, whilst its specificity (1 - FPR) increases. This is reasonable because as P_δ^τ is increased the number of estimated components, especially the noise associated components, passed on to the classifier decreases. In general, the TPR decreases at a much slower pace than the FPR. This suggests that genuine components tend to have high P_δ values. Hence, P_δ might be a useful decision tool when deciding to accept or reject the classification method's results. Whether or not we should automatically, and arbitrarily, consider classifications whose P_δ values are above a certain threshold is a different issue. For example, consider the results of *Table 4.5*, which compares the classifications at $P_\delta^\tau = 0$ & $P_\delta^\tau = 0.3$, w.r.t. each artificial noisy signal. The results suggests that using $P_\delta^\tau > 0$ might be beneficial, if P_δ^τ is chosen wisely. When $P_\delta^\tau = 0$, the number of classified noise peaks, i.e. FP, per experiment is sometimes quite high. However when $P_\delta^\tau = 0.3$ the maximum FP value is 8, the FN values differ from those at $P_\delta^\tau = 0$ by a maximum of 2, and the TP values differ from those at $P_\delta^\tau = 0$ by a maximum of 2 (*Table 4.5*). Thus $P_\delta^\tau = 0.3$ is a reasonable threshold, but it is subjective. Hence it is probably better to provide the user with all the classifications, and P_δ values, in prioritized form, and let the user make the final decision.

The graphs of *Fig. 4.5* highlight the trend of the TP, FN, and FP values, as P_δ^τ is increased. The trend is as expected, i.e.

- the number of FP, i.e. the number of classified noise components, decreases.
- the number of true positives, i.e. the number of correctly classified components, decreases because as P_δ^τ is increased, more genuine peaks are left out, and consequently
- the number of false negatives, i.e. the number of unclassified genuine components, increases.

Generally, the false positives consist of noise components that are similar to **(1)** the QMP secondary spectral frequencies of genuine components, or **(2)** a metabolite's QMP core spectral frequency, hence any noise components similar to the metabolite's QMP secondary spectral frequencies will also be classified. There is no clear way to deal with these two situations, other than by exploiting P_δ^τ and/or by considering patterns of association, as discussed within the *Decision Algorithm* section (*Section 4.2*). For example, and for the case $P_\delta^\tau = 0$, consider the subset of false positives detailed in *Table 4.4*. All the estimated spectral frequencies of *Table 4.4* were classified as NAAG components. This is an improbable result, not mainly because the λ values of the classified components are rather far apart (refer to the *Decision Algorithm* section, *Section 4.2*), but because the P_δ of the core frequency, 2.060 p.p.m., is extremely low.[5]

[5] λ denotes the importance of a spectral frequency (or mono-component), as discussed in *Section 4.2*, it is a cardinal number that signifies how critical a spectral frequency is to a metabolite's spectrum. 1 is allocated to the most important spectral frequency; importance decreases with increasing λ value.

f_{est} (p.p.m.)	a_{est} (a.u.)	d_{est}	p	λ	P_δ
1.529	0.150	11.516	0.002	32	0.785
2.060	0.015	0.352	0.950	1	0.054
2.238	0.354	6.839	0.358	2	0.925
2.521	0.059	4.453	0.022	12	0.491
2.722	0.054	0.054	0.169	9	0.457

Table 4.4: A subset of the FP components at $P_\delta^\tau = 0$. All the components were classified as NAAG components. f_{est}: estimated spectral frequency, a_{est}: estimated amplitude, p: $p(f_{est} \in C_k| \ldots)$

It should be noted that the classification of the estimated mono-component whose spectral frequency is 2.238 p.p.m., is a borderline FP/FN classification *because* the spectral frequency estimate is exactly equal to the spectral frequency of the 16^{th} mono-component of the artificial signal (*Table C.2*), but the amplitude and decay constant values differ by 125.478% and 57.581% respectively. Hence the estimated mono-component is either **(1)** a noise peak similar to the 16^{th} mono-component of the artificial signal, or **(2)** a genuine estimate of the 16^{th} mono-component of the artificial signal, whose amplitude and decay constant parameter values are converging extremely slowly.

Altogether the results illustrate that the classification method's performance, unsurprisingly, does depend on the quality of the data it receives.

		$P_\delta^\tau = 0$					$P_\delta^\tau = 0.3$				
		5%	10%	15%	20%	10% Ext.	5%	10%	15%	20%	10% Ext.
LLac	1.265	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	1.373	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
NAA	2.008	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	2.613	✓	✓	✓		✓	✓	✓	✓		✓
	2.475	✓	FN	✓	✓	✓	✓	FN	✓	FN	✓
	2.573	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	4.228					FN					FN
Cr	3.027	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	3.920	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Ch	3.183	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
MI	3.544	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	4.055	✓	✓	✓	✓	✓	✓	✓	✓	FN	✓
LGlc	2.287	✓		✓	✓	✓	✓		✓	✓	✓
	2.133	✓	FN	✓	✓	✓	✓	FN	✓	✓	✓
	2.369	✓	FN				✓	FN			
	2.238			FN	✓	✓			FN	✓	FN
	3.744	✓	✓	✓		✓	✓	✓	✓		✓
	2.348	✓	FN			✓	✓	FN			✓
	TP	16	11	14	13	16	16	11	14	11	15
	FN	0	4	1	0	1	0	4	1	2	2
	FP	14	4	9	12	16	8	3	4	7	5
	TN	18	11	8	11	20	24	12	13	16	31
	$\sum$	48	30	32	36	53	48	30	32	36	53

Table 4.5: A comparison of classifications: $P_\delta^\tau = 0$, and $P_\delta^\tau = 0.3$. The table highlights the peaks that were correctly, and incorrectly, classified when the $P_\delta^\tau = 0$ & 0.3. The false positives are noise peaks that were classified, whilst the false negatives are unclassified or misclassified genuine peaks. Compare the TP, FP, and FN values with the graphed values of *Fig. 4.5*

4.5 Conclusions

This chapter introduced a new classification method for classifying the estimated components of an NMRS signal. The classification method consists of a conditional probability model, and a decision algorithm, which both use QMP metabolite data to classify estimated mono-components via each mono-component's estimated spectral frequencies only. Additionally, the generalized likelihood ratio test (GLRT), which is used in detection theory for differentiating between genuine and noise components [59], was discussed as an optional tool for reducing the number of noise components passed on to the classifier.

The classification method's performance was assessed via a set of ROC graphs that were evaluated using an artificial data set. The ROC graphs (*Fig. 4.4* & *4.5*) illustrated that the classification method can classify estimated components quite well. This suggests that it is possible to classify an estimated mono-component using its estimated frequency only, and that QMP metabolite data is a good prior knowledge candidate. The GLRT rarely rejects genuine components, as *Table 4.3* highlights, therefore it might be beneficial to use the classification method in conjunction with the GLRT. Using a probability of detection P_δ threshold might also be beneficial, however there is a greater risk of losing genuine components (*Fig. 4.4c*, *Fig. 4.4d*, *Fig. 4.5*, and *Fig. 4.5*).

An important cautionary note − the classification method will classify a noise peak that is quite similar to a genuine peak, and which passes all the tests of the decision algorithm. However such a classification, akin to the example of *Table 4.4* (page 93), would probably be characterized by low P_δ values, and/or disparate λ values.

The next chapter tests the classification method on experimentally acquired phantom and volunteer data.

Chapter 5

Quantitative Analysis of the Classification Method

In the previous chapter a new automatic classification method for classifying the estimated mono-components of an NMRS signal was introduced. In this chapter the classification method is tested using data acquired from a 1.5 tesla General Electric Signa Scanner (Western General Hospital, Edinburgh). The first section of this chapter tests the classification method on NMRS signals acquired from a phantom, whilst the second section tests it on NMRS signals acquired from the brain of human volunteers. Each experiment, and the classification results, are detailed clearly within each experiment. Prior to discussing the experiments and results, the analysis steps and measures are discussed below.

5.1 Definitions, Analysis Steps, and Analysis Measures

5.1.1 Definitions of terms used

Definition 8 (Core Conditional Probability p_{core}) *The core conditional probability p_{core} is the conditional probability, as defined by Eq. 4.1.8, that an estimated frequency is associated with the core QMP spectral frequency of a metabolite.*

Definition 9 (Importance λ) *The importance λ of a spectral frequency (or mono-component), as discussed in Section 4.2, is a cardinal number that signifies how critical a spectral frequency is to a metabolite's spectrum.* 1 *is allocated to the most important spectral frequency; importance decreases with increasing λ value.*

5.1.2 Analysis Steps

Each experimentally acquired signal discussed in this chapter had a residual water peak (i.e. the water peak left-over after experimental water suppression), therefore the analysis steps are

1. Residual water peak filtering/removal using the HLSVD method [69], of jMRUI [74, 75].
2. FPT decomposition using the optimal settings of *Table 3.2*.

3. Generalized Likelihood Ratio Test (*Section 4.3*) to **(a)** filter out noise peaks, and **(b)** evaluate the P_δ (probability of detection) of each estimated mono-component. [N.B. A P_δ threshold was not used.]

4. Classification of estimated mono-components.

5.1.3 Analysis Measures

Except in the case of a phantom, we are never certain of the chemical composition, and quantities, at a source, hence the number of applicable statistical performance measures is limited. Even in the case of a phantom our options are limited − there is still no standard method for relating chemical concentrations to estimated amplitudes [53]. Altogether, there are three implicit or explicit measures used by the analysis:

1. **The Probability of Detection (P_δ)** The P_δ values provided by the GLRT (*Section 4.3, page 84*), which can be used to assess whether or not the classifications of the classification method are acceptable or plausible.

2. **The Importance of the Classified Estimated Components per Metabolite** This enables us to judge the plausibility of the classifier's classifications per metabolite, as illustrated by the example of *Fig. 4.3* (*page 82*).

3. **The Number of Misclassifications** Misclassifications are either classified noise peaks, or misclassified genuine peaks. In instances whereby we know the chemical composition of a signal's source, e.g. the phantom, it is possible to manually check for classification errors because the spectral frequencies of many metabolites have been recorded in some journal papers, e.g. [5, 43, 71], or can be predicted quantum mechanically. Within each experiment cum results section, w.r.t. the phantom, the results tables are such that **misclassifications are printed in red**.

As discussed in the previous chapter, the classification model's conditional probabilities are used by the decision algorithm for assigning estimated mono-components to metabolites . However using these probability values to assess whether or not a classification should be accepted is tricky because metabolite spectral frequencies are highly overlapping (as *Fig. F.1*, F.2, F.3, and F.4 illustrate), which sometimes leads to a spread of low conditional probabilities. Therefore the three measures above will be the main analysis measures, and the conditional probabilities will be discussed in some cases.

In cases of overlap − cases whereby an estimated spectral frequency is classified as the core spectral frequency of more than 1 metabolite − only the *best* of the overlaps is tabulated in the main results tables. The *best* option is chosen with regard to either the core conditional probabilities p_{core}, or the pattern of associations (*Step 4* of the *Decision Algorithm* section, *page 82*). Explicitly, either the metabolite with the highest p_{core}, or the metabolite with the best set of λ values is selected.

5.2 Analysis of Phantom Data

To test the classification method, a set of experiments were conducted in a 1.5 tesla General Electric Signa Scanner (Western General Hospital, Edinburgh). These experiments are detailed below. The 4 experiments of this section were conducted using the aforementioned GE Phantom Head (*Table C.1*), which contains 6 metabolites: choline (3 mM), creatine (10 mM), l-glutamate (12.5 mM), lactate (5 mM), n-acetly-l-aspartic acid (12.5 mM) & myoinositol (7.5 mM). In each experiment SVS signals, each a time series signal of length 2048 points, were acquired using a PRESS (Point Resolved Spectroscopy) pulse sequence of repetition time $T_R = 1500$ ms, and the acquisition sampling time $\tau = 0.0004$ s per experiment. The echo time T_E, voxel size, and the Number of Signal Averages (NSA) values, used per experiment, are detailed per experiment.

5.2.1 Experiment and Results: Noise Level Analysis

In order to assess the effect of noise on the classification method's performance, the phantom head was scanned 5 times using different NSA values: 32, 64, 128, 256, 512. The experiments were conducted consecutively; on the same day, and within minutes of each other. Generally, as the NSA is increased, the SNR of a signal increases [66]. Therefore it is expected that at low NSA not all the metabolites will be identified, but as the NSA increases **(a)** all the metabolites in the phantom will be identified, and **(b)** the identifications, w.r.t. phantom metabolites, will *stabilize* as the NSA is increased. The latter means that once a metabolite present in the phantom has been identified, it must be identified in signals of higher NSA thereafter. The other scanning parameters were $T_E = 35$ ms, voxel size 20 mm $\times$ 20 mm $\times$ 15 mm; refer to *Table D.1* for a detailed breakdown. The Fourier spectrum of each acquired, and water filtered, signal is displayed in *Fig. 5.3*.

Observations

Identified Phantom Metabolites. The main results of this experiment are detailed in *Tables 5.2, 5.3*, & *5.4*. At first, i.e. NSA = 32, only 5 of the phantom's 6 metabolites were correctly identified. Afterwards all the metabolites were identified per acquired signal (*Tables 5.2, 5.3*, & *5.4*), i.e. the identifications w.r.t. phantom metabolites stabilize.

Misclassifications. As NSA is increased, the SNR of a signal improves [66], thus fewer non-phantom metabolites were expected with increasing NSA. However the NSA = 256 & NSA = 512 tables include 3 non-phantom metabolites each, whilst the others include only 1 each. Whether this is an anomaly is not quite clear. L-aspartic acid (LAsp), one of the misclassifications, recurs quite frequently amongst the phantom's classification results; amongst the results of this experiment and of other experiments. This suggests that the spectral frequencies associated with it are probably genuine, albeit l-aspartic acid is not included in the phantom. L-aspartic acid is actually a moiety of NAA, one of the phantom's metabolites; as their 2-dimensional shapes, *Fig. F.5*, illustrate. Consequently, it is not unreasonable to expect a slight overlap between their QMP spectral frequencies. However the QMP spectral frequencies of NAA and l-aspartic acid hardly overlap (*Fig. 5.1*). Hence it is quite possible that the NAA

Experiment	C_k	λ	p_{core}	C_k	λ	p_{core}	C_k	λ	p_{core}
NSA = 32	Cr	1	0.755	PCh	1, 5	0.983			
	PCr	1	0.256	GPC	1, 14	0.009			
NSA = 64	Cr	1, 2	0.373						
	PCr	1	0.138						
NSA = 128	Cr	1, 2	0.390						
	PCr	1	0.201						
NSA = 256	Cr	1, 2	0.427						
	PCr	1	0.171						
NSA = 512	Cr	1, 2	0.425	LGlc	1, 2, 3	0.251	LAsp	1, 2, 3, 10, 12	0.360
	PCr	1, 2	0.153	GABA	1, 2	0.671	GPC	1, 7, 10, 12, 18, 20	0.006

Table 5.1: Overlapping Pairs: NSA experiments. In each case the *best* of each pair is tabulated in *Tables 5.2*, *5.3*, and *5.4*. **Variables** $\longrightarrow C_k$: Metabolite, λ: Importance (*Definition 9*, *page 97*). p_{core}: Core conditional probability (*Definition 8*, *page 97*).

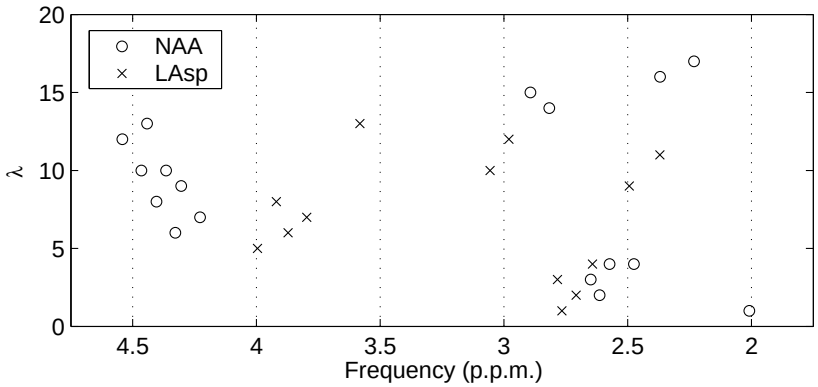


Figure 5.1: The QMP spectral frequencies of NAA & L Aspartic Acid

chemical shifts, and coupling constants, garnered from literature are inaccurate or incomplete. The NAAG classifications of NSA = 512 also recurs in later experiments. The issue in this case is much more straight forward − the 1^{st} & 2^{nd} most important QMP spectral frequencies of NAAG overlap the 7^{th} & 4^{th} most important QMP spectral frequencies of l-glutamic acid (LGlc), respectively. In fact the QMP spectral frequencies of these two metabolites are highly overlapping between 1.85 p.p.m. and 2.60 p.p.m. (*Fig. 5.2*). Another misclassification of interest, with regards to NSA = 32, is the phosphorylcholine (PCh) misclassification. In a few cases, when the same estimated frequency is not associated with the core QMP spectral frequency of choline (Ch) and phosphorylcholine (PCh), a frequency of approximately 3.20 p.p.m. is assigned to phosphorylcholine (*Tables 5.2*, *5.6*, and *5.9*), albeit it is not present in the phantom. The recurrence of this frequency, and its high P_δ value in each case, suggest that the corresponding estimated mono-component is probably a genuine metabolite component. The core QMP spectral frequency of phosphorylcholine is 3.208 p.p.m., hence the classification is

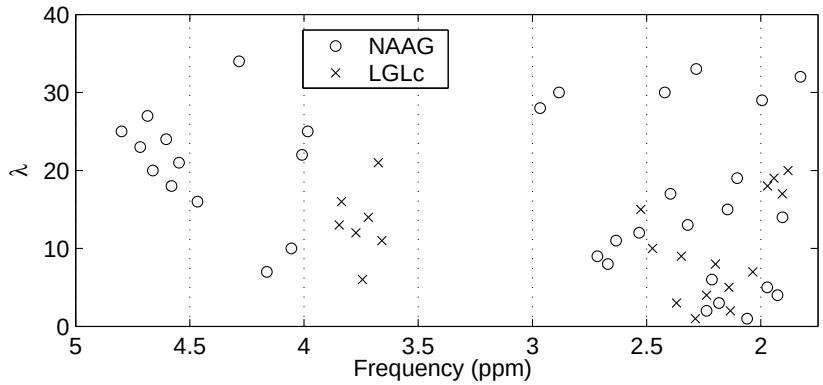


Figure 5.2: The QMP spectral frequencies of NAAG & L Glutamic Acid

plausible, although wrong. At present the source of this component is not clear, but the most probable source is myoinositol. At present there are various sets of chemical shifts & coupling constants for myoinositol [2, 25, 43, 104, 116], illustrating how difficult it is to determine the physico-chemical constants of this compound. Using different combinations of the constants of the papers [2, 25, 43, 104, 116] does sometimes lead to peaks around the 3.20 p.p.m. mark. For the myoinositol QMP spectral frequencies of this text, the constants of [43] were used.

Altogether, it seems a number of the misclassifications are due to overlapping instances or inaccurate data. There are fewer misclassifications w.r.t. the NSA = 32, 64, and 128 experiments probably because there are fewer estimated components, especially genuine estimated components, to assess. The misclassifications that are not associated with l-aspartic acid, NAAG, or phosphorylcholine, and some of those associated with them, are probably noise peaks considering their low P_δ values.

Probability of Detection P_δ. The core P_δ values of the classifications associated with non-phantom metabolites are usually lower than those of the phantom metabolites. If we exclude l-aspartic acid and NAAG classifications, most of the other classifications have low P_δ values. The odd result is the PCh classification w.r.t. NSA = 32.

Overlapping Instances. The overlapping instances of this experiment are detailed in *Table 5.1*. In the case of NSA = 32 there are two overlapping cases. First, creatine (Cr) and phosphocreatine (PCr). This is not a surprising case because the spectra of these two metabolites are quite similar [43]. The QMP spectral frequencies are 3.027 p.p.m.(0.597) & 3.92 p.p.m.(0.403) for creatine, and 3.029 p.p.m.(0.599) & 3.93 p.p.m.(0.401) for phosphocreatine; the bracketed numbers are the normalized probabilistic intensities. Creatine is detailed in *Table 5.2* because it has the highest core conditional probability p_{core}. The second NSA = 32 overlapping case is that of phosphorylcholine and GPC. In fact neither metabolite is present; the classifier has probably classified a noise peak. Phosphorylcholine is displayed because its λ values are closer together. The trickiest overlapping case occurs when NSA = 512 − case glutamate (LGlc) and GABA. The p_{core} of GABA is greater than that of glutamate, but glutamate was selected be-

cause there are more estimates associated with it, and their λ values are as close together, and as important, as those of GABA. In general, only metabolites that have overlapping structures (*Fig. F.5*, *F.6*, and *F.7*) and/or overlapping spectral frequencies (*Fig. F.1*, *F.2*, *F.3*, and *F.4*), overlap.

The Estimated Spectral Frequencies & Amplitudes. The graphs of *Fig. 5.4a* & *b* summarize how similar, or dissimilar, the estimated spectral frequencies and amplitudes, of all 5 experiments, are. The spectral frequency estimates, *Fig. 5.4a*, are quite similar. The amplitude estimates, *Fig. 5.4b*, are slightly varied, which is in line with the observations of the comparative analysis section (*The Effect of Noise on Decomposition Accuracy*, *Section 3.3.4*), i.e. the rate of convergence of the FPT's amplitude estimates is lower than that of the frequency estimates, and more so as the SNR decreases.

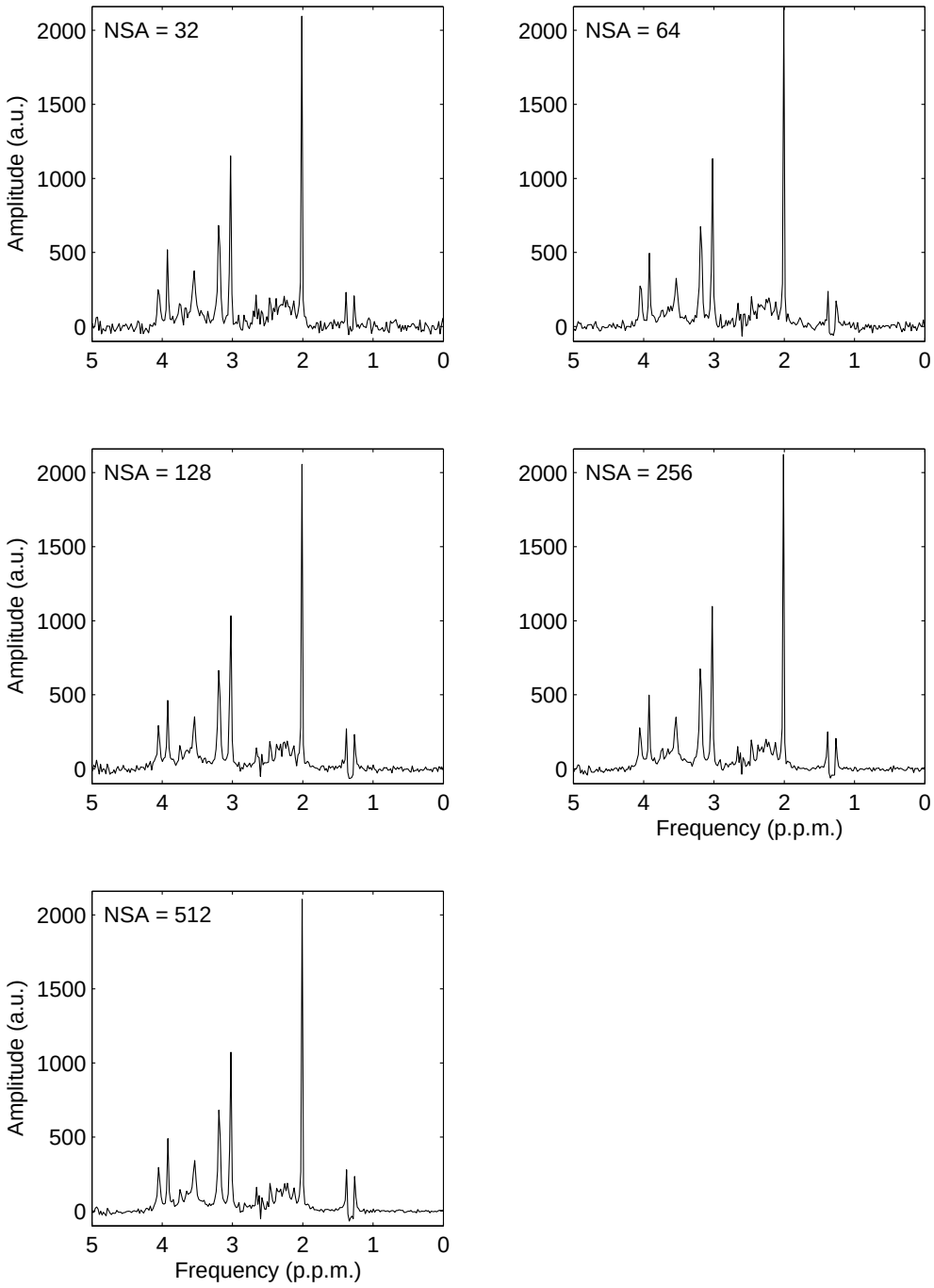
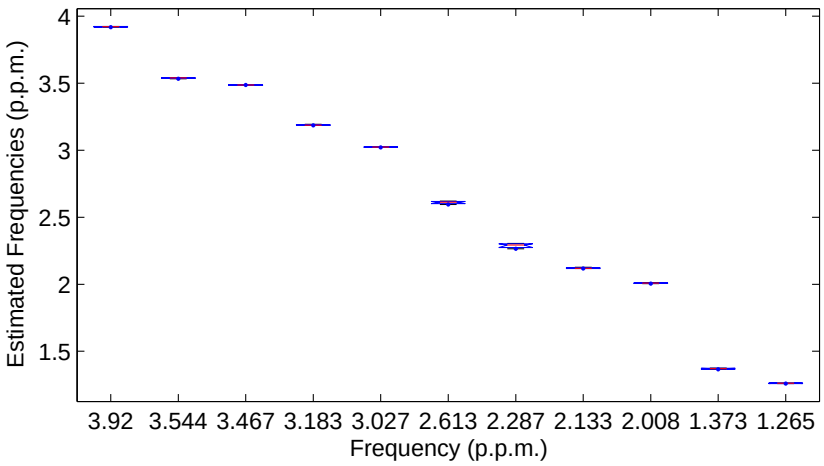
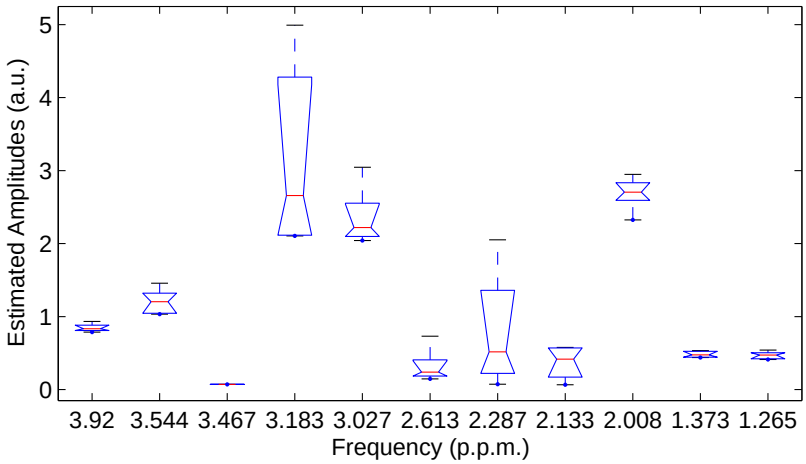


Figure 5.3: The Fourier spectra of the five, water filtered, NMRS signals acquired for the noise level analysis experiment. The classifications per NMRS signal are detailed in the tables that follow (*Tables 5.2*, *5.3*, and *5.4*).



(a) Comparing the FPT estimated frequencies with the original values.



(b) Comparing the FPT estimated amplitudes with the original values.

Figure 5.4: Noise Level Analysis. The difference, using the standard deviation measure, between the estimated parameter values, of a classified estimated component, and the original values.

NSA = 32					NSA = 64				
C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ	C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ
Ch	3.193	0.755	0.994	1	**Ace**	1.985	0.048	0.158	1
	3.971	0.001	0.429	4	Ch	3.186	0.308	0.991	1
Cr	3.026	0.353	0.979	1		3.406	0.041	0.401	11
LLac	1.375	0.921	0.809	1	Cr	3.022	0.373	0.988	1
	1.263	0.901	0.836	2		3.918	0.052	0.945	2
MI	3.537	0.212	0.941	1	LGlc	2.304	0.040	0.926	1
	4.053	0.109	0.851	4		2.119	0.141	0.290	2
	2.963	0.155	0.771	24		2.379	0.007	0.623	3
NAA	2.012	0.890	0.983	1		2.237	0.340	0.917	4
	2.595	0.012	0.618	2	LLac	1.365	0.814	0.875	1
	2.634	0.017	0.547	3		1.259	0.885	0.868	2
	2.542	0.013	0.668	4		4.670	0.116	0.278	3
					MI	3.533	0.544	0.960	1
PCh	3.207	0.983	0.968	1		4.046	0.059	0.946	4
	4.259	0.031	0.157	5		3.110	0.002	0.550	7
						3.224	0.087	0.396	13
						2.904	0.011	0.498	24
					NAA	2.005	0.676	0.989	1
						2.603	0.028	0.935	2
						2.647	0.063	0.550	3
						2.575	0.459	0.968	4
						2.823	0.035	0.131	14

Table 5.2: The noise level analysis classification results: cases NSA = 32, and NSA = 64. **Variables** $\longrightarrow$ C_k: Metabolite, f_n: estimated spectral frequency, $p(f_n \in C_k \vert \ldots)$: the conditional probability (*Eq. 4.1.8*), P_δ: probability of detection, λ: importance (*Definition 9*, *page 97*).

NSA = 128					NSA = 256				
C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ	C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ
Ch	3.186	0.308	0.998	1	**Ace**	1.955	0.095	0.146	1
	3.732	0.115	0.815	10	Ch	3.185	0.974	0.996	1
Cr	3.025	0.390	0.996	1		3.486	0.001	0.582	2
	3.923	0.004	0.974	2	Cr	3.021	0.427	0.996	1
LAsp	2.741	0.512	0.830	1		3.920	0.022	0.979	2
	2.704	0.775	0.718	2	**LAsp**	2.758	0.695	0.695	1
	2.823	0.184	0.652	3		2.816	0.050	0.640	3
	2.913	0.002	0.576	12	LGlc	2.300	0.029	0.936	1
LGlc	2.284	0.726	0.993	1		2.118	0.008	0.967	2
	2.127	0.646	0.858	2		2.374	0.050	0.493	3
	2.374	0.050	0.410	3		2.250	0.181	0.892	4
	2.231	0.596	0.735	4	LLac	1.367	0.849	0.950	1
	1.976	0.045	0.854	18		1.259	0.885	0.945	2
	1.688	0.022	0.098	20		4.625	0.094	0.571	3
LLac	1.372	0.916	0.940	1	MI	3.539	0.481	0.992	1
	1.263	0.901	0.933	2		4.047	0.329	0.967	4
	4.743	0.114	0.518	3		3.644	0.018	0.983	5
MI	3.541	0.075	0.978	1		4.093	0.919	0.843	10
	4.049	0.335	0.952	4		3.054	0.014	0.793	11
	3.660	0.716	0.967	5		3.847	0.488	0.619	16
	3.376	0.070	0.088	12		2.899	0.044	0.532	24
NAA	2.010	0.959	0.995	1	NAA	2.006	0.848	0.998	1
	2.612	0.261	0.872	2		2.617	0.025	0.857	2
	2.663	0.006	0.741	3		2.655	0.083	0.896	3
	2.579	0.015	0.731	4		2.580	0.027	0.876	4
						2.222	0.002	0.995	17
					Tau	3.432	0.321	0.555	1

Table 5.3: The noise level analysis classification results (continued): cases NSA = 128, and NSA = 256. **Variables** $\longrightarrow$ C_k: Metabolite, f_n: estimated spectral frequency, $p(f_n \in C_k \vert \ldots)$: the conditional probability (*Eq. 4.1.8*), P_δ: probability of detection, λ: importance (*Definition 9*, *page 97*)

NSA = 512					NSA = 512, continued				
C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ	C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ
Ace	1.890	0.944	0.280	1	NAA	2.009	0.953	0.999	1
Ch	3.189	0.886	0.998	1		2.620	0.609	0.857	2
	4.074	0.001	0.956	7		2.660	0.003	0.914	3
	3.401	0.157	0.159	11		2.586	0.479	0.933	4
	4.158	0.001	0.492	13	**NAAG**	2.038	0.295	0.953	1
Cr	3.023	0.425	0.998	1		2.231	0.130	0.594	2
	3.922	0.009	0.989	2		1.941	0.614	0.836	4
LAsp	2.752	0.360	0.952	1		1.970	0.923	0.972	5
	2.733	0.037	0.717	2		2.216	0.811	0.996	6
	2.827	0.196	0.850	3		2.527	0.071	0.954	12
	3.048	0.074	0.921	10					
	2.900	0.004	0.559	12					
LGlc	2.264	0.252	0.702	1					
	2.118	0.008	0.978	2					
	2.376	0.039	0.854	3					
LLac	1.370	0.891	0.976	1					
	1.262	0.898	0.972	2					
	4.735	0.136	0.761	3					
MI	3.541	0.075	0.994	1					
	4.050	0.065	0.983	4					
	3.665	0.132	0.904	5					
	4.016	0.002	0.195	6					

Table 5.4: The noise level analysis classification results (continued): case NSA = 512. **Variables** $\longrightarrow$ C_k: Metabolite, f_n: estimated spectral frequency, $p(f_n \in C_k \vert \ldots)$: the conditional probability (*Eq. 4.1.8*), P_δ: probability of detection, λ: importance (*Definition 9*, *page 97*)

5.2.2 Experiment and Results: Reproducibility Assessment

In order to assess how consistent the classification method's classifications are, 4 SVS signals were acquired on different days using the same scanning parameters. The main scanning parameters are $T_E = 35$ ms, voxel size 20 mm × 20 mm × 15 mm, and NSA = 128; refer to *Table D.1* for a detailed breakdown.

Observations

Experiment	C_k	λ	p_{core}	C_k	λ	p_{core}	C_k	λ	p_{core}
a	Cr	1,2	0.299	Ch	1, 13	0.361	LGlc	1,2	0.738
	PCr	1	0.376	PCh	1	0.387	GABA	1,19,20,22	0.243
b	Cr	1,2	0.307	PCh	1,4	0.968	LGlc	1,5,10	0.670
	PCr	1	0.275	GPC	1,8,13	0.006	GABA	1,8,14,20	0.358
c	PCr	1,2	0.331	Ch	1	0.755	LGlc	1,2,3,4,18,20	0.852
	Cr	1	0.326	PCh	1,3	0.130	GABA	1,3,10	0.092
d	Cr	1,2	0.387						
	PCr		0.201						

Table 5.5: Overlapping Pairs: Reproducibility assessment experiments. In each case the best of each pair is tabulated in *Tables 5.6 & 5.7*. **Variables** $\longrightarrow C_k$: Metabolite, λ: Importance (*Definition 9, page 97*). p_{core}: Core conditional probability (*Definition 8, page 97*).

Identified Phantom Metabolites. The classifier identified all the phantom metabolites in almost every case. The exception is the result of Experiment c - refer to the result under **c** in *Table 5.7* - where a component that should have been classified as creatine was classified as phosphocreatine. This is a plausible misclassification because the metabolites are quite similar, as discussed in the previous experiment; the estimates are actually closer to phosphocreatine's QMP spectral frequencies.

Misclassifications. The misclassification of a creatine peak as phosphocreatine was mentioned above. The l-aspartic acid, NAAG, and phosphorylcholine misclassifications discussed in the previous experiment also occur here; the same analysis applies. The acetate classification of *Table 5.6*, is probably a noise peak classification considering its P_δ value; although acetate recurs a few times, the estimated spectral frequencies associated with its core QMP spectral frequency are quite dissimilar. Some of the spectral frequencies associated with the phantom metabolites have high λ values and low P_δ values, such classifications **might** be noise peak classifications.

Probability of Detection P_δ. The P_δ values of genuine classifications, especially those with low λ values (< 10), tend to be quite high.

Overlapping Instances. The overlapping instances, *Table 5.5*, are consistent and plausible, i.e. only metabolites that have overlapping structures (*Fig. F.5*, *F.6*, and *F.7*) and/or overlapping spectral frequencies (*Fig. F.1*, *F.2*, *F.3*, and *F.4*), overlap.

The Estimated Spectral Frequencies & Amplitudes. The graphs of *Fig. 5.6a* & *b* summarize how similar, or dissimilar, the estimated spectral frequencies and amplitudes, of all 4 experiments, are. The spectral frequency estimates, *Fig. 5.6a*, are quite similar. The amplitude estimates, *Fig. 5.6b*, vary more.

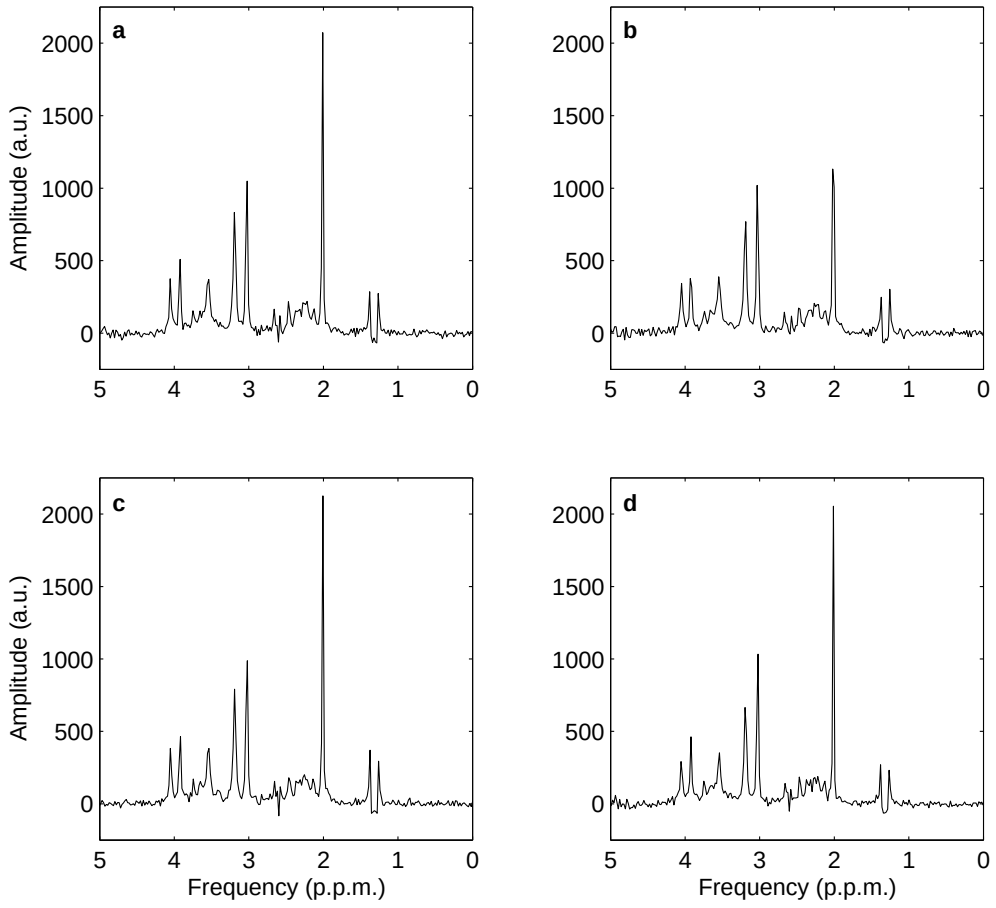
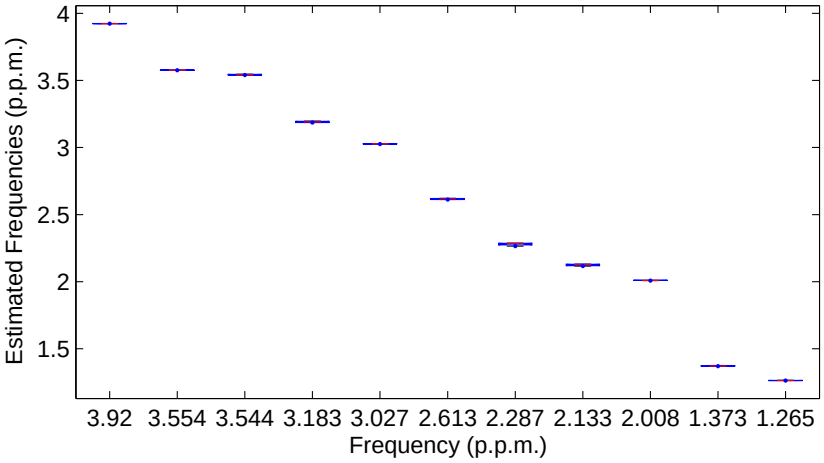
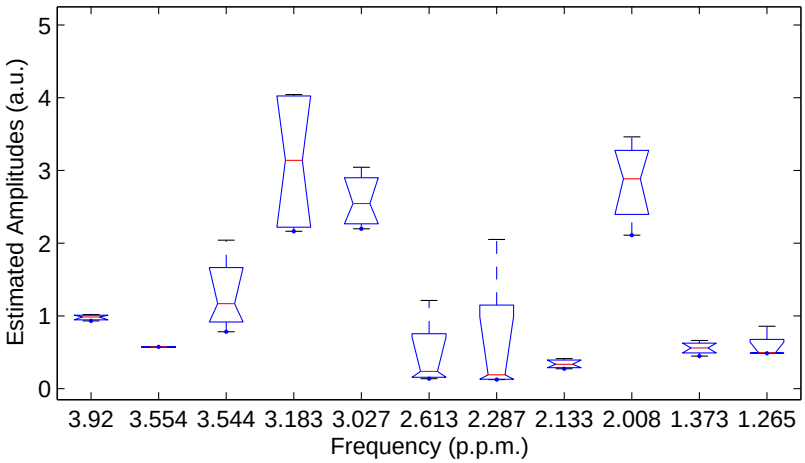


Figure 5.5: The Fourier spectra of the four, water filtered, NMRS signals acquired for the reproducibility assessment experiment; the scanning parameters per acquired signal are the same, as detailed in the text. The classifications per NMRS signal are detailed in the tables that follow (*Tables 5.6*, and *5.7*). In relation to (b), it is quite probable that the lower peak around the 2 p.p.m. mark, compared to the other graphs, is due to the relevant point falling between the two adjacent points.



(a) Comparing the FPT estimated frequencies with the original values.



(b) Comparing the FPT estimated amplitudes with the original values.

Figure 5.6: Reproducibility Assessment. The difference, using the standard deviation measure, between the estimated parameter values, of a classified estimated component, and the original values.

T_E = 35 ms, NSA = 128: **a**					T_E = 35 ms, NSA = 128: **b**				
C_k	f_n	$p(f_n \in C_k \mid \ldots)$	P_δ	λ	C_k	f_n	$p(f_n \in C_k \mid \ldots)$	P_δ	λ
Ace	1.908	1.000	0.340	1	Ch	3.186	0.308	0.990	1
Ch	3.196	0.361	0.997	1		3.518	0.107	0.758	3
	4.181	0.028	0.375	13		4.156	0.003	0.300	13
Cr	3.030	0.299	0.994	1	Cr	3.027	0.307	0.991	1
	3.924	0.032	0.968	2		3.923	0.004	0.965	2
LAsp	2.747	0.977	0.442	1	LGlc	2.288	0.640	0.614	1
	2.708	0.991	0.268	2		2.137	0.968	0.986	5
	2.837	0.124	0.631	3		2.481	0.463	0.840	10
LGlc	2.283	0.738	0.595	1	LLac	1.376	0.902	0.896	1
	2.116	0.038	0.888	2		1.263	0.901	0.907	2
LLac	1.372	0.916	0.941	1		4.474	0.026	0.239	3
	1.266	0.903	0.959	2	MI	3.546	0.003	0.951	1
	4.630	0.025	0.809	3		4.054	0.005	0.975	4
MI	3.539	0.481	0.989	1		3.821	0.144	0.110	16
	3.576	0.060	0.928	2	NAA	2.010	0.959	0.996	1
	4.052	0.124	0.931	4		2.613	0.151	0.621	2
	3.649	0.121	0.654	5		2.666	0.038	0.681	3
	3.298	0.029	0.919	20		2.582	0.231	0.743	4
	4.114	0.005	0.358	21	**NAAG**	2.057	0.881	0.354	1
	2.966	0.002	0.771	24		2.248	0.642	0.413	2
	4.007	0.116	0.882	25		2.194	0.231	0.745	3
NAA	2.008	0.975	0.990	1		1.979	0.176	0.972	5
	2.620	0.609	0.976	2		1.536	0.004	0.133	32
	2.652	0.651	0.645	3	**PCh**	3.203	0.968	0.820	1
	2.584	0.565	0.876	4		4.214	0.004	0.371	4

Table 5.6: The reproducibility assessment classification results: the first two NMRS signals. All the NMRS signals for this experiment were acquired using the same experiment parameter values. **Variables** $\longrightarrow$ C_k: Metabolite, f_n: estimated spectral frequency, $p(f_n \in C_k \mid \ldots)$: the conditional probability (*Eq. 4.1.8*), P_δ: probability of detection, λ: importance (*Definition 9*, *page 97*)

T_E = 35 ms, NSA = 128: **c**					T_E = 35 ms, NSA = 128: **d**				
C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ	C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ
Ch	3.193	0.755	0.991	1	Ch	3.186	0.308	0.998	1
LGlc	2.265	0.852	0.794	1		3.732	0.115	0.815	10
	2.131	0.992	0.856	2	Cr	3.025	0.390	0.996	1
	2.370	0.117	0.454	3		3.923	0.004	0.974	2
	2.224	0.250	0.551	4	**LAsp**	2.741	0.512	0.830	1
	1.974	0.237	0.924	18		2.704	0.775	0.718	2
	1.525	0.025	0.084	20		2.823	0.184	0.652	3
LLac	1.371	0.908	0.930	1		2.913	0.002	0.576	12
	1.263	0.901	0.911	2	LGlc	2.284	0.726	0.993	1
MI	3.543	0.309	0.978	1		2.127	0.646	0.858	2
	4.048	0.242	0.958	4		2.374	0.050	0.410	3
	3.661	0.287	0.729	5		2.231	0.596	0.735	4
	3.078	0.013	0.374	7		1.976	0.045	0.854	18
	3.796	0.001	0.194	14		1.688	0.022	0.098	20
NAA	2.009	0.953	0.995	1	LLac	1.372	0.916	0.940	1
	2.621	0.538	0.709	2		1.263	0.901	0.933	2
	2.662	0.083	0.824	3		4.743	0.114	0.518	3
	2.587	0.005	0.839	4	MI	3.541	0.075	0.978	1
	2.827	0.004	0.742	14		4.049	0.335	0.952	4
PCr	3.029	0.331	0.991	1		3.660	0.716	0.967	5
	3.927	0.001	0.951	2		3.376	0.070	0.088	12
					NAA	2.010	0.959	0.995	1
						2.612	0.261	0.872	2
						2.663	0.006	0.741	3
						2.579	0.015	0.731	4

Table 5.7: The reproducibility assessment classification results (continued): the other two signals. All the NMRS signals for this experiment were acquired using the same experiment parameter values. **Variables** $\longrightarrow$ C_k: Metabolite, f_n: estimated spectral frequency, $p(f_n \in C_k \vert \ldots)$: the conditional probability (*Eq. 4.1.8*), P_δ: probability of detection, λ: importance (*Definition 9*, *page 97*)

5.2.3 Experiment and Results: Stability Analysis

To assess the stability of the method, the GE Phantom Head was scanned 4 times using different echo times T_E: 35 ms, 60 ms, 95 ms, 145 ms. The T_E variable controls the time at which signal acquisition begins, and because a metabolite's signal decays with time, the acquirable metabolite signal intensity decreases as T_E increases. Therefore a decrease in the number of identifiable phantom metabolites is expected, but such a decrease should not lead to an increase in spurious classifications. If it does then the classification method is not stable. The other scanning parameters are voxel size 20 mm × 20 mm × 15 mm, and NSA = 128; refer to *Table D.1* for a detailed breakdown of the echo times.

Observations

Experiment	C_k	λ	p_{core}	C_k	λ	p_{core}
T_E = 35 ms	Cr	1,2	0.389			
	PCr	1	0.201			
T_E = 60 ms	Cr	1,2	0.358			
	PCr	1	0.403			
T_E = 95 ms	Cr	1,2	0.456	Ch	1	0.877
	PCr	1	0.373	GPC	1,6,9	0.003
				PCh	1,6	0.025
T_E = 145 ms	Cr	1,2	0.390			
	PCr		0.134			

Table 5.8: Overlapping Pairs: Stability analysis experiments. In each case the best of each pair is tabulated in *Tables 5.9 & 5.10*. **Variables** $\longrightarrow C_k$: Metabolite, λ: Importance (*Definition 9, page 97*). p_{core}: Core conditional probability (*Definition 8, page 97*).

Identified Phantom Metabolites. As the T_E increases, the number of identified phantom metabolites decreases. Myoinositol was not detected at T_E = 145 ms. Myoinositol might be detectable at 95 ms; some of the estimated spectral frequencies of *Spectrum 5.10a* are in the vicinity of the QMP spectral frequencies of MI, however this needs to be investigated further. In most cases MI is not detectable at long echo times [5, 43]. Most of the λ values per metabolite, and per experiment, are low − suggesting high plausibility. It is also in line with the fact that spectra acquired at T_E = 145 ms tend to be *cleaner* than those acquired at lower T_E (*Fig. 5.7*).

Misclassifications. The l-aspartic acid (LAsp), phosphorylcholine (PCh), and acetate (Ace) misclassifications discussed in the previous experiments also occur here; the same analysis applies. Most of the classifications associated with scylloinositol (SI) and taurine (Tau) are probably noise peak classifications, considering their low P_δ values. The one exception is the 3.152 p.p.m. taurine classification.

Probability of Detection P_δ. The P_δ values of genuine classifications, especially those with low (< 10) λ values, are initially (i.e. case $T_E = 35$ ms) high, but generally decrease with increasing T_E.

Overlapping Instances. The overlapping instances, *Table 5.8*, are consistent and plausible, i.e. only metabolites that have overlapping structures (*Fig. F.5*, *F.6*, and *F.7*) and/or overlapping spectral frequencies (*Fig. F.1*, *F.2*, *F.3*, and *F.4*), overlap.

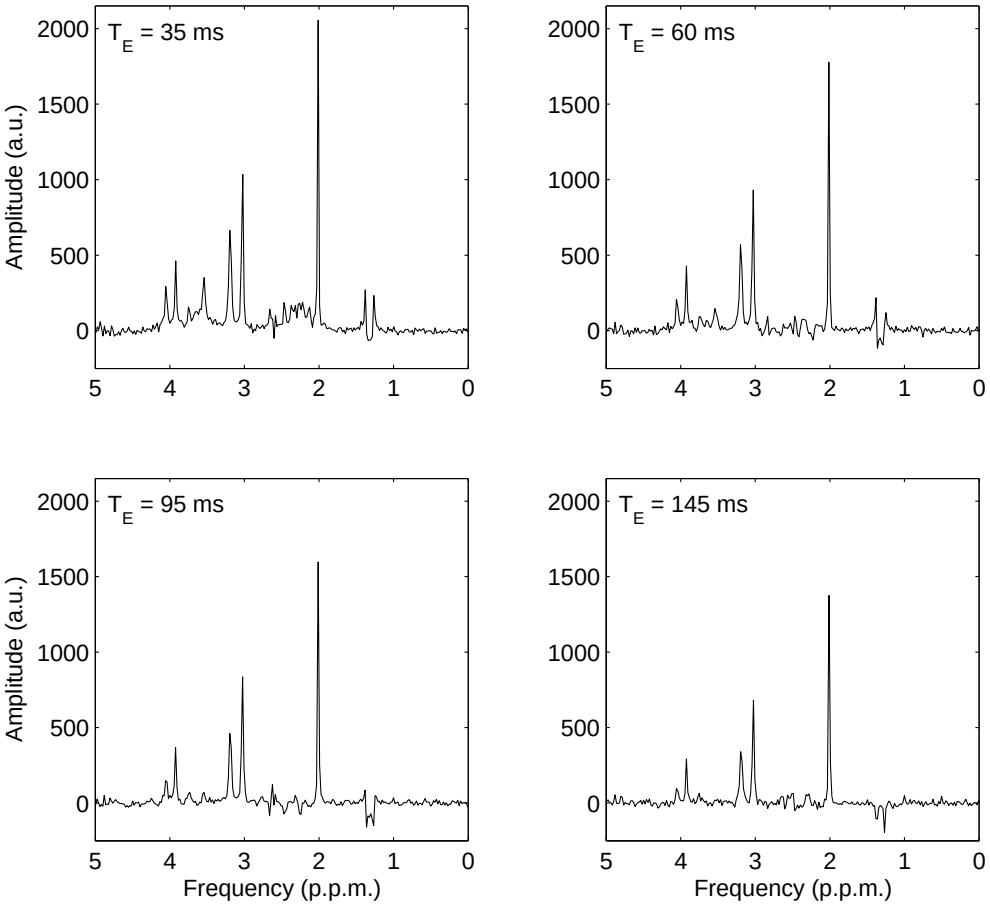


Figure 5.7: The Fourier spectra of the four, water filtered, NMRS signals acquired for the stability analysis experiment.

$T_E = 35$ ms					$T_E = 60$ ms				
C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ	C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ
Ch	3.186	0.308	0.998	1	Ch	3.185	0.972	0.983	1
	3.732	0.115	0.815	10		3.418	0.045	0.161	5
Cr	3.025	0.390	0.996	1	Cr	3.029	0.358	0.993	1
	3.923	0.004	0.974	2		3.923	0.000	0.978	2
LAsp	2.741	0.512	0.830	1	**LAsp**	2.708	0.983	0.625	1
	2.704	0.775	0.718	2		2.766	0.989	0.463	2
	2.823	0.184	0.652	3		2.799	0.032	0.949	4
	2.913	0.002	0.576	12		2.968	0.061	0.121	10
LGlc	2.284	0.726	0.993	1		2.521	0.003	0.596	11
	2.127	0.646	0.858	2		3.065	0.012	0.395	12
	2.374	0.050	0.410	3	LGlc	2.318	0.479	0.865	1
	2.231	0.596	0.735	4		2.366	0.002	0.637	6
	1.976	0.045	0.854	18		1.990	0.245	0.795	9
	1.688	0.022	0.098	20		1.924	0.006	0.570	19
LLac	1.372	0.916	0.940	1		2.433	0.008	0.966	21
	1.263	0.901	0.933	2	LLac	1.374	0.998	0.921	1
	4.743	0.114	0.518	3		1.264	0.999	0.919	2
MI	3.541	0.075	0.978	1		4.672	0.020	0.677	4
	4.049	0.335	0.952	4	MI	3.527	0.088	0.993	1
	3.660	0.716	0.967	5		3.551	0.020	0.141	2
	3.376	0.070	0.088	12		3.675	0.064	0.402	3
NAA	2.010	0.959	0.995	1		4.050	0.500	0.938	4
	2.612	0.261	0.872	2		3.601	0.802	0.944	6
	2.663	0.006	0.741	3		3.968	0.003	0.891	15
	2.579	0.015	0.731	4		4.517	0.083	0.111	16
					NAA	2.013	0.898	0.995	1
						2.475	0.328	0.922	2
						2.650	0.610	0.715	5
						2.231	0.624	0.752	6
					PCh	3.197	0.321	0.951	1
						4.170	0.039	0.228	4
					SI	3.333	0.986	0.445	1

Table 5.9: The stability analysis classification results: cases $T_E = 35$ ms, and 60 ms. **Variables** $\longrightarrow C_k$: Metabolite, f_n: estimated spectral frequency, $p(f_n \in C_k \vert \ldots)$: the conditional probability (*Eq. 4.1.8*), P_δ: probability of detection, λ: importance (*Definition 9*, *page 97*)

$T_E = 95$ ms					$T_E = 145$ ms				
C_k	f_n	$p(f_n \in C_k \mid \ldots)$	P_δ	λ	C_k	f_n	$p(f_n \in C_k \mid \ldots)$	P_δ	λ
Ch	3.190	0.877	0.993	1	**Ace**	1.853	0.179	0.352	1
Cr	3.027	0.456	0.992	1	Ch	3.189	0.959	0.980	1
	3.924	0.009	0.966	2		3.558	0.003	0.450	14
LGlc	2.346	0.003	0.949	1	Cr	3.023	0.390	0.966	1
	2.280	0.307	0.959	3		3.923	0.007	0.944	2
	2.118	0.426	0.568	4	**LAsp**	2.692	0.113	0.366	1
	2.018	0.796	0.985	6		2.837	0.177	0.498	2
	2.236	0.023	0.632	7		2.651	0.077	0.803	3
LLac	1.377	0.965	0.860	1	LGlc	2.271	0.125	0.825	1
	1.266	0.987	0.919	2		2.117	0.083	0.217	5
NAA	2.007	0.778	0.993	1		2.207	0.101	0.699	6
	2.656	0.536	0.920	2		2.324	0.007	0.812	8
	2.616	0.093	0.817	3	LLac	1.374	0.998	0.840	1
	2.481	0.380	0.213	4		1.266	1.000	0.810	2
	2.583	0.003	0.390	5	NAA	2.011	0.990	0.996	1
Tau	3.415	0.975	0.483	1		2.583	0.664	0.777	2
	3.152	0.038	0.919	3		2.617	0.312	0.245	4
	3.288	0.508	0.172	6		2.485	0.072	0.852	5

(a)

Table 5.10: The stability analysis classification results: cases $T_E = 95$ ms, and 145 ms. **Variables** $\longrightarrow$ C_k: Metabolite, f_n: estimated spectral frequency, $p(f_n \in C_k \mid \ldots)$: the conditional probability (*Eq. 4.1.8*), P_δ: probability of detection, λ: importance (*Definition 9*, *page 97*).

5.2.4 Experiment and Results: Reduced Voxel Size

In SVS, as discussed in *Section 2.7* (*page 29*), as the voxel size is decreased the magnetic homogeneity within the voxel improves. However, this is at the expense of Signal-to-Noise Ratio (SNR). To assess the classification method's response to the effects of smaller voxel sizes, the previous experiment was repeated using voxels of size 15 mm × 15 mm × 15 mm.

Observations

Experiment	C_k	λ	p_{core}	C_k	λ	p_{core}	C_k	λ	p_{core}
T_E = 35 ms	Cr	1,2	0.332						
	PCr	1	0.317						
T_E = 60 ms	Cr	1,2	0.479	PCh	1,4	0.058			
	PCr	1	0.206	Ch	1	0.023			
T_E = 95 ms	Cr	1,2	0.467	Ch	1,3	0.990	NAA	1,2,3,5	0.977
	GABA	1,2,4,14	0.194	PCh	1,2,6,9	0.003	NAAG	1,11	0.004
	PCr	1,2	0.171						
T_E = 145 ms	Cr	1,2	0.385	Ch	1,8	0.959			
	PCr	1	0.140	PCh	1	0.007			

Table 5.11: Overlapping Pairs: Reduced voxel size experiments. In each case the best of each pair is tabulated in *Tables 5.12 & 5.13*. **Variables** $\longrightarrow C_k$: Metabolite, λ: Importance (*Definition 9*, *page 97*). p_{core}: Core conditional probability (*Definition 8*, *page 97*).

Identified Phantom Metabolites. As a whole the results are poor compared to those of voxel size 20 mm × 20 mm × 15 mm. The smaller voxel size had quite an effect on the quality of the acquired signals, and hence on the quality of the mono-components passed on to the classifier. Consequently, misclassifications such as l-glutamine (LGle), which we have not come across before, occur a few times. The main surprise is the presence of a myoinositol (MI) classification per experiment. Altogether, the experiment shows that the classifier's performance is, unsurprisingly, dependent on the quality of the estimated decompositions it is analysing.
The l-serine (LSer) classification of the T_E = 35 ms results has been included to illustrate the degree of overlap amongst the metabolites. The spectral frequencies in the range [3.92 3.93] p.p.m. are key spectral frequencies of creatine (Cr), phosphocreatine (PCr) and l-serine (LSer). With regards to the current set-up, the results of creatine take precedence because l-serine is a *rarely detected* metabolite. However, if all three metabolites are present in a mixture, a signal decomposition should include three components within the region, hence separate classifications would not be a problem.

Misclassifications. A positive aspect of these results is that some of the incorrect classifications are plausible. For example, the components classified as glutamine (LGle) would mainly be the l-glutamic acid (LGlc) components if the decompositions passed on to the classifier were

more accurate. The components classified as phosphorylcholine (PCh) are either choline (Ch) components, considering their overlapping structures, or myoinositol (MI), based on the discussion of the first experiment (*Section 5.2.1, page 99*). Considering the generally low P_δ values of the taurine (Tau) and scylloinositol (SI) classifications, it is quite probable that they are noise peak classifications.

Probability of Detection P_δ. The P_δ values of genuine classifications, especially those with low λ values (< 10), are generally high.

Overlapping Instances. The overlapping instances, *Table 5.11*, are consistent and plausible, i.e. only metabolites that have overlapping structures (*Fig. F.5*, *F.6*, and *F.7*) and/or overlapping spectral frequencies (*Fig. F.1*, *F.2*, *F.3*, and *F.4*), overlap.

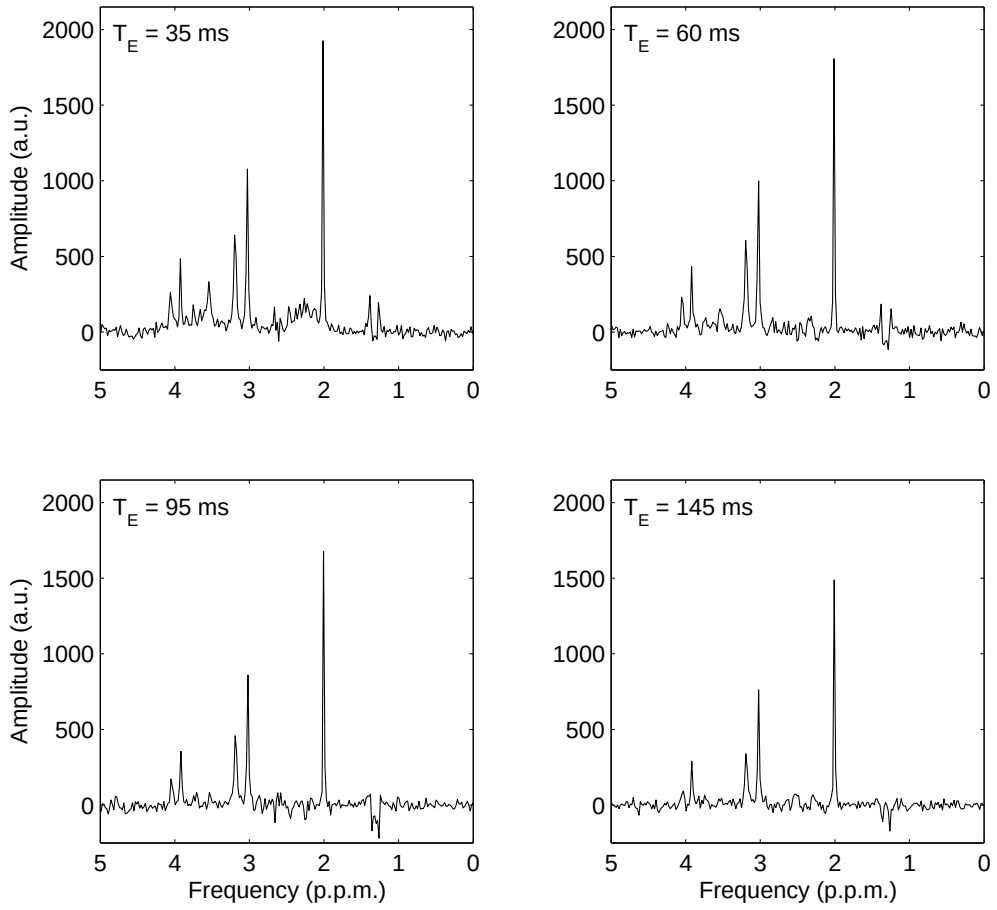


Figure 5.8: The Fourier spectra of the four, water filtered, NMRS signals acquired for the reduced voxel size experiment.

$T_E = 35$ ms					$T_E = 60$ ms				
C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ	C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ
Ch	3.191	0.894	0.981	1	Cr	3.024	0.479	0.982	1
	3.929	0.000	0.958	2		3.924	0.003	0.916	2
Cr	3.028	0.332	0.979	1	LGlc	2.317	0.455	0.755	1
LLac	1.373	0.920	0.837	1		2.518	0.257	0.713	17
	1.266	0.903	0.811	2	**LGle**	2.462	0.749	0.771	1
MI	3.532	0.186	0.835	1	LLac	1.373	0.998	0.879	1
	4.060	0.338	0.904	4		1.261	0.999	0.924	2
	3.644	0.018	0.588	5	MI	3.530	0.224	0.946	1
	2.911	0.002	0.388	24		4.051	0.493	0.951	4
NAA	2.013	0.830	0.991	1		3.112	0.107	0.101	5
	2.613	0.151	0.706	2	NAA	2.008	0.941	0.989	1
	2.659	0.006	0.433	3		2.574	0.183	0.166	3
	2.581	0.014	0.816	4		2.623	0.065	0.509	4
	2.374	0.003	0.378	16		2.670	0.060	0.298	5
	2.250	0.002	0.852	17		4.565	0.019	0.101	16
PCh	3.217	0.268	0.598	1	**PCh**	3.197	0.058	0.967	1
	3.684	0.017	0.481	2		4.205	0.003	0.143	4
	3.748	0.010	0.584	3	**Tau**	3.232	0.215	0.376	1
LSer†	3.929	0.992	0.958	1		3.443	0.015	0.230	2
	4.060	0.279	0.904	5		3.364	0.693	0.241	7
						3.053	0.039	0.514	10

Table 5.12: The voxel size 15 mm × 15 mm × 15 mm classification results: cases $T_E = 35$ ms, and 60 ms. †Rare. **Variables** $\longrightarrow$ C_k: Metabolite, f_n: estimated spectral frequency, $p(f_n \in C_k \vert \ldots)$: the conditional probability (*Eq. 4.1.8*), P_δ: probability of detection, λ: importance (*Definition 9, page 97*).

$T_E = 95$ ms					$T_E = 145$ ms				
C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ	C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ
Ch	3.187	0.990	0.972	1	Ch	3.189	0.959	0.965	1
	4.166	0.068	0.257	3		4.625	0.014	0.349	8
Cr	3.021	0.466	0.980	1	Cr	3.022	0.385	0.968	1
	3.916	0.038	0.923	2		3.919	0.048	0.897	2
LGlc	2.329	0.553	0.376	1	**LGle**	2.152	0.763	0.949	1
	2.424	0.813	0.241	2		2.430	0.245	0.147	2
	2.283	0.138	0.909	3		2.226	0.874	0.521	4
LGle	2.194	0.018	0.784	1		2.323	0.314	0.345	6
	2.225	0.005	0.832	3	LLac	1.358	1.000	0.508	1
LLac	1.372	0.967	0.883	1		1.264	1.000	0.394	2
	1.258	0.897	0.776	2	MI	4.037	0.047	0.614	1
	4.684	0.065	0.322	3	NAA	2.007	0.998	0.985	1
MI	3.528	0.664	0.784	1		2.583	0.664	0.639	2
	4.044	0.048	0.833	3		2.667	0.038	0.756	3
	3.788	0.013	0.083	14		2.475	0.497	0.808	5
NAA	2.005	0.977	0.985	1	**SI**	3.353	0.075	0.699	1
	2.654	0.646	0.561	2					
	2.610	0.071	0.582	3					
	2.578	0.041	0.617	5					

Table 5.13: The voxel size 15 mm × 15 mm × 15 mm classification results: cases $T_E = 95$ ms, and 145 ms. **Variables** $\longrightarrow$ C_k: Metabolite, f_n: estimated spectral frequency, $p(f_n \in C_k \vert \ldots)$: the conditional probability (*Eq. 4.1.8*), P_δ: probability of detection, λ: importance (*Definition 9*, *page 97*).

5.3 Analysis of Human Brain Data

Human brain single voxel spectroscopy scans, of two adult male volunteers, were conducted "at the SFC Brain Imaging Research Centre (www.sbirc.ed.ac.uk) using a protocol approved by the Local Research Ethics Committee (LREC/1998/3/1; 'Assessment of Magnetic Resonance Imaging with Healthy Volunteers')". To assess the stability of the method − as in *Section 5.2.3* (*page 113*) − w.r.t. human brain data, a region of each volunteer's brain was scanned 4 times using different echo times: 35 ms, 60 ms, 95 ms, 145 ms. As before, the T_E variable controls the time at which signal acquisition begins, and because a metabolite's signal decays with time, the acquirable metabolite signal intensity decreases as T_E increases. Therefore a decrease in the number of identifiable metabolites is expected. The other scanning parameters are voxel size 20 mm × 20 mm × 15 mm, and NSA = 128; refer to *Table D.1* for a detailed breakdown of the echo times.

5.3.1 Observations: Volunteer 1

The Fourier spectra of the water filtered NMRS signals acquired from Volunteer 1 are displayed in *Fig. 5.9*.

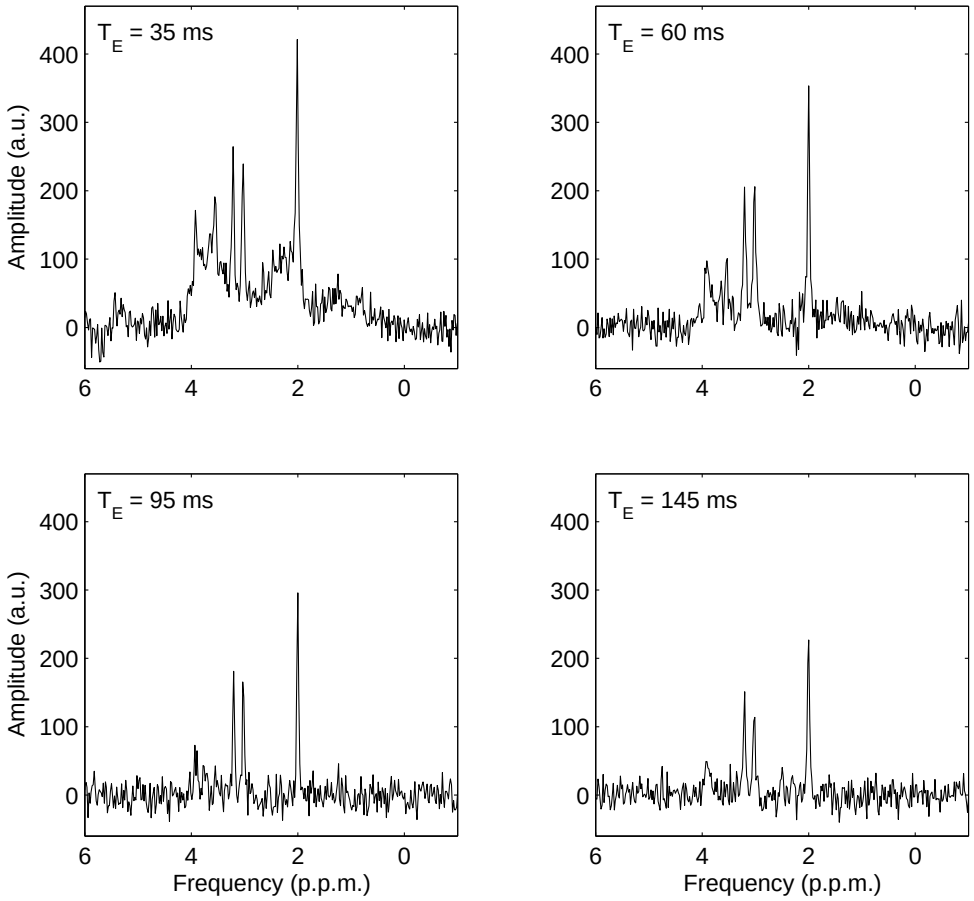


Figure 5.9: The Fourier spectra of the four, water filtered, NMRS signals acquired from the brain of Volunteer 1.

Experiment	C_k	λ	p_{core}	C_k	λ	p_{core}	C_k	λ	p_{core}
$T_E = 35$ ms	PCr	1	0.546	PCh	1,4	0.468			
	Cr	1	0.214	GPC	1	0.115			
				Ch	1,3,10,12,14	0.002			
$T_E = 60$ ms	PCr	1	0.607	PCh	1,5	0.649			
	Cr	1	0.203	Ch	1,13	0.140			
$T_E = 95$ ms	Cr	1,2	0.446	PCh	1,5	0.718	NAA	1,4,5	0.979
	PCr	1	0.388	Ch	1,3,11,13	0.154	NAAG	1,3,4	0.008
				GPC	1,18	0.053			
$T_E = 145$ ms	PCh	1,7	0.989						
	GPC	1	0.003						
	Ch	1,3	0.001						

Table 5.14: Overlapping Pairs: Volunteer 1. Only the best of each pair is tabulated in *Tables 5.15* & *5.16*. **Variables** $\longrightarrow C_k$: Metabolite, λ: Importance (*Definition 9*, *page 97*). p_{core}: Core conditional probability (*Definition 8*, *page 97*).

The Identified Metabolites. Presently, there is no way to confirm the identity of the metabolites at the source. Hence the main points that can be made, w.r.t. the results of *Tables 5.15* & *5.16*, is that a consistent set of metabolites were identified, and the observed metabolites are in line with those that are normally identified in the human brain [5, 43]. The following paragraph elaborates on these points.

Betaine (Bet), l lactate (LLac), NAA, and phosphorylcholine (PCh) classifications are associated with each acquired signal. This result is in line with [5], which notes that the four metabolites are detectable at long T_E. A creatine (Cr) or phosphocreatine (PCr) classification is associated with each acquired signal. Since creatine and phosphocreatine have similar spectral frequencies (as discussed in the *Analysis of Phantom Data* section, *page 99*), this behaviour was expected. Myoinositol (MI) classifications are associated with the $T_E = 35$ ms and $T_E = 145$ ms signals, but not the other two. Unfortunately, it seems the main myoinositol components might be amongst the genuine components that the decomposition method, the FPT, was unable to extract due to an insufficient number of well-sampled NMRS signal data points; refer to the *FPT* section (*page 42*). The inconsistent myoinositol classifications underscore an important point – *the classification method's performance depends on the quality of the data it receives*. Another odd (inconsistent) classification is the l glutamic acid (LGlc) w.r.t. the $T_E = 95$ ms only. L glutamic acid is one of the most abundant metabolites in the human brain [43]. Thus its detection is expected at $T_E = 35$ ms, at least. Hence it is quite possible that the main components of l glutamic are also amongst the genuine components that the FPT was unable to extract due to an insufficient number of well-sampled data points. A scylloinositol (SI) classification is associated with all the acquired signals except T_E =145 ms. This is a plausible trend, and hence a plausible result, because the acquirable signal of a metabolite decreases with increasing T_E. A similar analysis applies to NAAG. Some, or all, of the acetate (Ace) classifications are probably noise peak classifications. It should be noted that acetate's QMP

spectrum consists of a single peak at approximately 1.9 p.p.m., i.e. an area where metabolite spectral frequencies are rare (refer to *Fig. F.1*, *F.2*, *F.3*, and *F.4*). Hence it is quite easy for the classifier to mistake a noise component for a acetate peak in this area.

Probability of Detection P_δ. The P_δ values of most of the classifications, especially those with low λ values (< 10), are quite high. Hence the classifications are quite plausible.

Overlapping Instances. The overlapping instances, *Table 5.14*, are consistent and plausible, i.e. only metabolites that have overlapping structures (*Fig. F.5*, *F.6*, and *F.7*) and/or overlapping spectral frequencies (*Fig. F.1*, *F.2*, *F.3*, and *F.4*), overlap.

$T_E = 35$ ms					$T_E = 60$ ms				
C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ	C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ
Bet	3.265	0.805	0.973	1	Bet	3.271	0.113	0.144	1
	3.698	0.140	0.696	13	LLac	1.357	0.999	0.661	1
LLac	1.431	0.626	0.717	1		1.278	0.998	0.785	2
	1.254	0.852	0.937	2	NAA	2.008	0.941	0.993	1
NAA	2.008	0.975	0.999	1		2.494	0.002	0.867	2
	2.673	0.015	0.928	3		2.577	0.208	0.956	3
	2.584	0.565	0.975	4	NAAG	2.058	0.720	0.509	1
NAAG	2.043	0.804	0.951	1		1.939	0.207	0.984	2
	2.167	0.148	0.974	3		2.179	0.460	0.596	3
	2.525	0.022	0.944	12		1.877	0.051	0.979	4
	1.868	0.015	0.818	14		1.817	0.196	0.324	31
	2.106	0.052	0.645	19		2.109	0.015	0.969	32
PCr	3.033	0.546	0.998	1	PCr	3.034	0.607	0.974	1
PCh	3.215	0.468	0.996	1	PCh	3.198	0.649	0.950	1
	4.210	0.008	0.672	4		4.259	0.010	0.324	5
SI	3.341	0.994	0.696	1	SI	3.338	0.991	0.278	1
Tau	3.463	0.966	0.776	1	Tau	3.222	0.012	0.760	1
	3.305	0.999	0.918	3		3.409	0.007	0.601	2
MI	3.546	0.003	0.971	1		3.160	0.113	0.377	5
	4.059	0.500	0.552	4	Ace	1.860	0.941	0.302	1
	3.654	0.130	0.614	5					

Table 5.15: The classification results w.r.t. Volunteer 1: cases $T_E = 35$ ms, and 60 ms. **Variables** $\longrightarrow$ C_k: Metabolite, f_n: estimated spectral frequency, $p(f_n \in C_k \vert \ldots)$: the conditional probability (*Eq. 4.1.8*), P_δ: probability of detection, λ: importance (*Definition 9*, *page 97*).

$T_E = 95$ ms					$T_E = 145$ ms				
C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ	C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ
Ace	1.812	0.355	0.577	1	Ace	1.861	0.228	0.787	1
Bet	3.249	0.580	0.577	1	Bet	3.274	0.067	0.668	1
Cr	3.028	0.446	0.995	1	Cr	3.043	0.022	0.984	1
	3.915	0.013	0.683	2					
LLac	1.353	0.614	0.375	1	LLac	1.370	0.986	0.637	1
	1.227	0.710	0.775	2		1.255	1.000	0.756	2
NAA	2.008	0.979	0.997	1	NAA	2.008	0.998	0.997	1
	2.484	0.028	0.878	4		2.562	0.009	0.888	2
	2.555	0.011	0.962	5		2.619	0.171	0.677	4
PCh	3.198	0.718	0.985	1	PCh	3.207	0.989	0.991	1
	3.564	0.002	0.741	5		4.236	0.004	0.545	7
LGlc	2.332	0.451	0.802	1	MI	4.018	0.007	0.560	1
	1.957	0.013	0.968	18		4.052	0.527	0.899	2
	1.786	0.046	0.564	20		3.373	0.032	0.920	5
SI	3.346	0.946	0.736	1					

Table 5.16: The classification results w.r.t. Volunteer 1 (continued): cases $T_E = 95$ ms, and 145 ms. **Variables** $\longrightarrow$ C_k: Metabolite, f_n: estimated spectral frequency, $p(f_n \in C_k \vert \ldots)$: the conditional probability (*Eq. 4.1.8*), P_δ: probability of detection, λ: importance (*Definition 9*, *page 97*).

5.3.2 Observations: Volunteer 2

The Fourier spectra of the water filtered NMRS signals acquired from Volunteer 2 are displayed in *Fig. 5.10*.

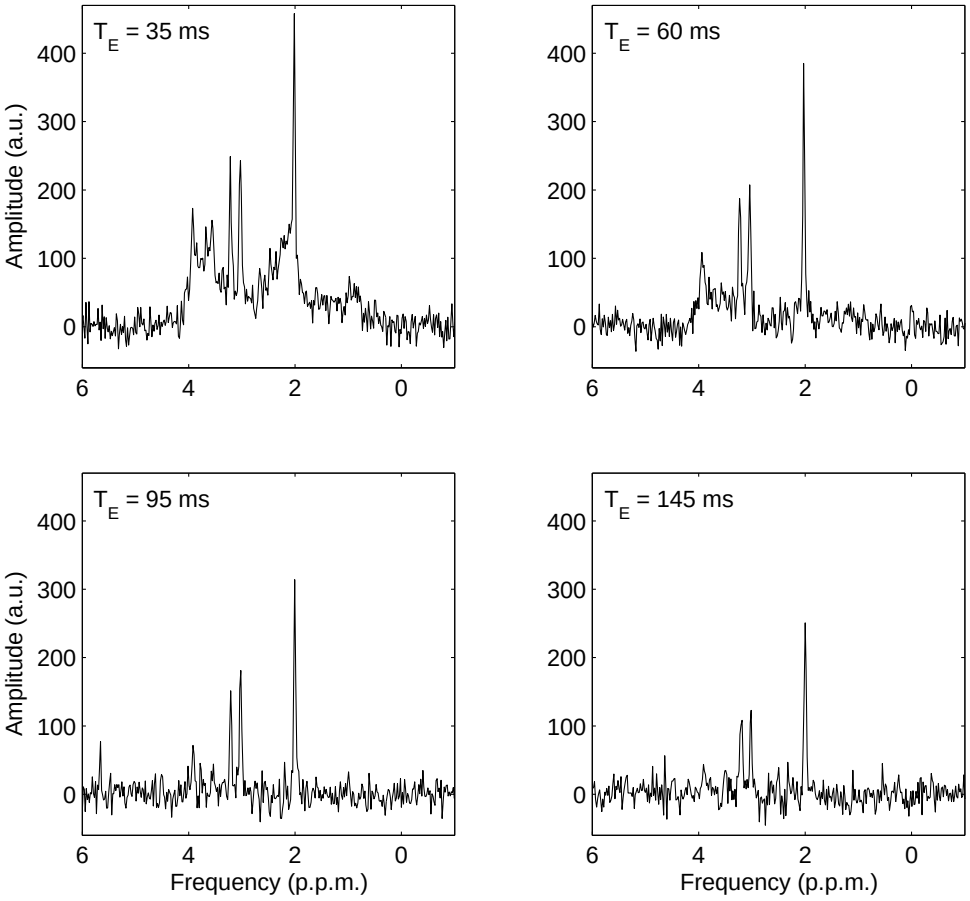


Figure 5.10: The Fourier spectra of the four, water filtered, NMRS signals acquired from Volunteer 2.

The Identified Metabolites. Foremost, the metabolites are in line with those that are normally observed in the human brain [5,43]. Secondly, a consistent and plausible set of metabolites were identified, as the following paragraph observes.

In most cases the same metabolite, or one of a similar structure, is either identified per signal or per $T_E \leq 95$ ms signal (*Tables 5.18 & 5.19*). The latter occurs, and is expected, because as T_E increases the acquirable signal intensity per metabolite decreases. Hence, and depending on a metabolite's decay rate and concentration, a metabolite might not be detectable at longer echo times [5,43]. For example, NAAG, l-glutamic acid (LGlc), and scylloinositol (SI) were identified in all the NMRS signals except the T_E =145 ms signal, a result that is in line with NMRS studies [5, 43]. In the case of NAAG and scylloinositol, both tend to occur in small amounts in the human brain [5,43], this leads to two possible situations w.r.t the current experiment, either

Experiment	C_k	λ	p_{core}	C_k	λ	p_{core}
$T_E = 35$ ms	LGlc	1	0.483	PCh	1,3	0.125
	GABA	1,8	0.173	GPC	1,17	0.025
$T_E = 60$ ms	Cr	1	0.311	PCh	1	0.412
	PCr	1	0.170	Ch	1	0.307
				GPC	1,2,4	0.010
$T_E = 95$ ms	Cr	1,2	0.246	PCh	1,8	0.966
	GABA	1,2,3,7, 14,15,17	0.175	GPC	1,3	0.027
$T_E = 145$ ms	PCh	1,2	0.329			
	GPC	1,10,13,15	0.113			

Table 5.17: Overlapping Pairs: Volunteer 2. Only the best of each pair is tabulated in *Tables 5.18* & *5.19*. **Variables** $\longrightarrow C_k$: Metabolite, λ: Importance (*Definition 9*, *page 97*). p_{core}: Core conditional probability (*Definition 8*, *page 97*).

- both have fully decayed before the T_E =145 ms signal is acquired, or
- the decomposition method was unable to detect their rather diminished components.

L glutamic acid, on the other hand, is present in the human brain in large amounts [43]. In fact l glutamic acid and NAA are two of the most abundant metabolites in the human brain, and their concentrations in the human brain are usually similar [43]. However, and unlike NAA, the spectrum of l glutamic acid consists of several low amplitude mono-components [43]. Hence, it is quite difficult to detect l glutamic acid at long echo times, e.g. T_E =145 ms. A creatine (Cr) or phosphocreatine (PCr) classification is associated with each acquired signal; as discussed within the *Analysis of Phantom Data* section (*Section 5.2*), this is an expected result because the two metabolites have similar spectral frequencies. Both phosphorylcholine (PCh) and choline (Ch) classifications are associated with the T_E =95 ms & T_E =145 ms signals, but only phosphorylcholine (PCh) was identified w.r.t. the other two signals. It is quite possible that the choline classifications are misclassifications because only very small, barely detectable, amounts of free choline exists in the human brain [43]. Most of the choline that is detectable in the human brain is due to the choline moieties of phosphorylcholine and GPC [43]. Hence, the phosphorylcholine classifications are plausible, whilst the choline classifications are questionable. However the phosphorycholine and myoinositol observations of the *Analysis of Phantom Data* section (*Section 5.2*) should be borne in mind; there it was noted that a myoinositol spectral frequency might exist around 3.20 p.p.m. There are Lactate (LLac), myoinositol (MI), and NAA classifications associated with each acquired signal. The presence of a myoinositol classification per signal, even though it is normally detected at short T_E, suggests that the volunteer probably has a high concentration of myoinositol. It also makes the scylloinositol classifications much more plausible because the concentration of scylloinositol (SI) seems to be "coupled" to that of myoinositol (MI) [43]; SI:MI $\approx$ 1:12. On the other hand, lactate and NAA are detectable at long T_E [5]. L glutamine (LGle) classifications are associated with the T_E = 60 ms, T_E =

95 ms, and $T_E = 145$ ms signals, but not the $T_E = 35$ ms signal. It is quite probable that the decomposition method, the FPT, was unable to extract all the genuine components from the T_E =35 ms signal; an expected FPT behaviour, especially in the case of noisy signals (*FPT, page 42*), which underscores the fact that the performance of the classification method depends on the quality of the data it receives. It is difficult to comment on the betaine (Bet), $T_E = 35$ ms, and taurine (Tau), $T_E = 60$ ms, classifications. Overall, a consistent and plausible set of metabolites were identified w.r.t. Volunteer 2, and a few more metabolites were identified in this case compared to Volunteer 1.

Probability of Detection P_δ. The P_δ values of most of the classifications, especially those with low λ values (< 10), are quite high. Hence the classifications are quite plausible.

Overlapping Instances The overlapping instances, *Table 5.17*, are consistent and plausible, i.e. only metabolites that have overlapping structures (*Fig. F.5*, *F.6*, and *F.7*) and/or overlapping spectral frequencies (*Fig. F.1*, *F.2*, *F.3*, and *F.4*), overlap.

$T_E = 35$ ms					$T_E = 60$ ms				
C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ	C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ
Ace	1.871	0.371	0.228	1	Cr	3.018	0.311	0.997	1
Bet	3.265	0.805	0.820	1	LGlc	2.346	0.001	0.776	1
	3.891	0.007	0.683	2		1.986	0.029	0.992	9
LAsp	2.746	0.823	0.634	1		1.711	0.930	0.338	22
	2.868	0.013	0.507	3	LGle	2.455	0.994	0.946	1
LGlc	2.280	0.483	0.665	1		2.410	0.033	0.905	4
LLac	1.360	0.882	0.785	1		2.120	0.002	0.558	6
	1.169	0.614	0.893	2		1.993	0.024	0.993	9
MI	3.533	0.544	0.999	1	LLac	1.357	0.999	0.494	1
	3.655	0.211	0.541	5		1.254	0.998	0.771	2
	4.019	0.004	0.722	6	MI	3.523	0.050	0.965	1
	4.089	0.249	0.775	10		3.674	0.080	0.809	3
	3.792	0.090	0.981	14		3.110	0.149	0.756	5
	3.854	0.003	0.961	16		4.038	0.353	0.338	8
NAA	2.002	0.382	0.998	1		3.815	0.024	0.838	13
	2.659	0.006	0.942	3	NAA	2.012	0.698	0.978	1
	2.562	0.009	0.810	4		2.574	0.183	0.528	3
NAAG	2.038	0.295	0.993	1		2.610	0.074	0.663	4
	2.233	0.300	0.984	2		2.267	0.012	0.918	6
PCr	3.042	0.193	0.997	1	NAAG	2.066	0.761	0.823	1
	3.924	0.014	0.997	2		2.187	0.016	0.882	3
PCh	3.220	0.125	0.880	1		2.308	0.001	0.146	12
	3.721	0.006	0.489	3	PCh	3.196	0.412	0.991	1
	4.207	0.002	0.541	4	SI	3.359	0.241	0.854	1
SI	3.349	0.532	0.967	1	Tau	3.237	0.533	0.954	1
						3.384	0.815	0.943	7

Table 5.18: The classification results w.r.t. Volunteer 2: cases $T_E = 35$ ms, and 60 ms. **Variables** $\longrightarrow C_k$: Metabolite, f_n: estimated spectral frequency, $p(f_n \in C_k \vert \ldots)$: the conditional probability (*Eq. 4.1.8*), P_δ: probability of detection, λ: importance (*Definition 9*, *page 97*).

$T_E = 95$ ms					$T_E = 145$ ms				
C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ	C_k	f_n	$p(f_n \in C_k \vert \ldots)$	P_δ	λ
Ace	1.866	0.930	0.688	1	Ace	1.901	1.000	0.513	1
Ch	3.168	0.250	0.774	1	Ch	3.192	0.633	0.987	1
Cr	3.020	0.246	0.996	1		3.510	0.095	0.513	2
	3.924	0.009	0.972	2		3.448	0.013	0.477	5
LAsp	2.739	0.011	0.704	1	Cr	3.020	0.256	0.988	1
	2.808	0.465	0.903	2		3.918	0.050	0.928	2
	2.662	0.012	0.905	4	GABA	1.979	0.661	0.980	1
LGlc	2.312	0.896	0.928	1	LAsp	2.693	0.911	0.827	1
	2.448	0.097	0.409	2		2.800	0.484	0.926	2
	2.129	0.974	0.912	4		2.669	0.006	0.613	3
LGle	2.174	0.283	0.924	1	LGle	2.152	0.763	0.391	1
LLac	1.354	0.623	0.234	1	LLac	1.413	0.755	0.777	1
	1.186	0.656	0.531	2		1.288	1.000	0.740	2
MI	3.528	0.664	0.891	1	MI	4.028	0.119	0.952	1
	4.057	0.062	0.704	3		3.371	0.038	0.613	5
	3.569	0.003	0.650	4		4.128	0.073	0.513	9
	3.434	0.002	0.758	9	NAA	2.008	0.998	0.993	1
NAA	2.008	0.979	0.998	1		2.530	0.004	0.795	2
	2.596	0.002	0.547	3		2.493	0.021	0.850	5
	2.504	0.003	0.986	4	PCh	3.212	0.329	0.964	1
NAAG	2.052	0.919	0.809	1		3.636	0.003	0.513	2
	2.200	0.013	0.495	4					
PCh	3.204	0.966	0.999	1					
	3.604	0.319	0.983	8					
SI	3.340	0.981	0.531	1					

Table 5.19: The classification results w.r.t. Volunteer 2 (continued): cases $T_E = 95$ ms, and 145 ms. **Variables** $\longrightarrow$ C_k: Metabolite, f_n: estimated spectral frequency, $p(f_n \in C_k \vert \ldots)$: the conditional probability (*Eq. 4.1.8*), P_δ: probability of detection, λ: importance (*Definition 9*, *page 97*).

5.4 Conclusions

The results of this chapter, especially those of the phantom, support the core conclusions of the previous chapter's test on synthetic data (*An Illustrative Example using Synthetic Data, page 86*), i.e.

- A classification method that exploits QMP data can be used to classify the estimated mono-components of an NMRS signal.
- By extension, QMP data is a good prior knowledge candidate.
- An estimated mono-component can be classified via its estimated spectral frequency only.

Additionally, the consistency of the classifications, w.r.t. both the phantom and the volunteers, illustrates that the classifications are not random, and that the classification method's potential w.r.t. human NMRS signals is promising. The absent metabolite classifications due to the absence of the relevant estimated mono-components − e.g. the absence of a Volunteer 2 l glutamine classification w.r.t. signal $T_E = 35$ ms − illustrates the dependence of the classification method's performance on the data it receives.

The P_δ and λ values are promising decision making tools. Most of the correct classifications, w.r.t. the phantom, tended to have high P_δ values and low λ values. The most plausible classifications w.r.t. the volunteers were also characterized by the same pattern. Hence P_δ and λ, together, might be helpful decision making tools.

With regard to user interaction, when a software package of the analysis approach is developed, a user only needs to provide a water filtered NMRS time series signal for analysis. The FPT's optimal setting depends on the length of the signal, which is automatically acquired. Classification is w.r.t. all the compounds present in the package, i.e. the user does not make any assumptions about the compounds present at a source.

In relation to the human brain data analysis, it is important to note that some of the rejected or unclassified mono-components might be mono-components of macromolecules or lipids. In fact, a few of the classified mono-components might be misclassified mono-components of macromolecules or lipids. Neither suggestions can be verified. Unfortunately, the classification method's QMP data does not include data from macromolecules or lipids because the actual identities of NMRS observable human brain macromolecules and lipids is unclear [91]. However the classification method's QMP data can be expanded as more knowledge about the metabolites, macromolecules, and lipids, present in the human is acquired.

The next chapter summarizes and concludes the thesis.

Chapter 6

Conclusions

At the beginning of this thesis three problems were identified w.r.t. NMRS signal analysis

1. The most widely used NMRS signal analysis methods are subjective and non-unique, i.e. they all depend on a user's choice of parameter values and/or prior knowledge data, and as the parameter values and/or prior knowledge data change, the results change.

2. Inefficient exploitation of metabolite prior knowledge. The interactive methods use prior knowledge via fitting methods, which is an ineffective approach because fitting methods are susceptible to under-fitting and over-fitting. This susceptibility is due to the fact that a fitting method finds a solution that fits the assumptions made about a problem, regardless of the physical implausibility of the solution.

3. An automatic classification method, for classifying the estimated components (decompositions) of decomposition based NMRS signal analysis methods, is not available.

Hence, the aim of the thesis was the development of an analysis approach that **(a)** relies on a unique, objective, and non-linear, signal analysis method that does not assume the nature of the signal being analysed, **(b)** exploits prior knowledge, but not via a fitting method, and **(c)** includes an automatic classification method if the appropriate signal analysis method is a decomposition based method.

In relation to **(a)**, the thesis reviewed and assessed the Fast Padé Transform [8]. The FPT was recently introduced to the field of NMRS signal analysis by Belkić [7, 18, 19]. The observed behaviour and performance of the FPT, and those of other decomposition based non-interactive analysis methods, is summarized later in this chapter. The FPT evaluates the mono-components of a multi-component signal, hence in relation to **(c)** a novel classification method was developed for classifying the estimated components. The classification method uses quantum mechanically predicted data to classify the estimated mono-components of a signal, i.e. the classification method also addresses objective **(b)**. The performance and characteristics of the classification method are summarized next.

6.1 Summary of Contributions

The contributions that directly, or indirectly, addressed the aforementioned problems are as follows

6.1.1 The Novel Classification Method.

The classification method introduced in *Chapter 4* (*A Novel Classification Method*) is the main original aspect of this thesis. It uses quantum mechanically predicted (QMP) metabolite data to classify the estimated mono-components of an NMRS signal, i.e. it exploits prior knowledge. Only the estimated spectral frequency of an estimated mono-component is required by the classifier in order to classify a mono-component. The extensive testing of the classification method led to the following conclusions

- The performance of the classifier is quite good, as the ROC graphs of the classifier illustrate (*An Illustrative Example using Synthetic Data*, *page 86*); and sometimes better if used in conjunction with an appropriate probability of detection threshold P_δ^τ.
- The good performance of the classification method suggests that QMP prior knowledge is a good prior knowledge candidate, and can be exploited in NMRS via non-fitting methods.
- The generalized likelihood ratio test (GLRT), which is used in detection theory for assessing whether a signal is genuine [59], can be used to reduce the number of noise components passed on to the classifier (*The Generalized Ratio Test*, *page 4.3*). When used appropriately it rarely rejects a genuine signal mono-component (e.g. *Table 4.3*).
- The classifier's classifications are consistent rather than random, i.e. the classifications w.r.t. NMRS signals from similar sources are similar, as the noise level analysis (*page 99*), the stability analysis (*page 113*), and the reproducibility assessment (*page 108*) phantom experiments illustrate. The results with respect to the human volunteers (*page 121*) are also consistent rather than random.
- The results summarizing overlapping metabolites are consistent and plausible, i.e. only metabolites with overlapping chemical structures (*Fig. F.5*, *F.6*, *F.7*) are amongst the phantom results (*Tables 5.1*, *5.5*, *5.8*), and the volunteers results (*Tables 5.14*, *5.17*).
- The classifier aids novelty detection. The classification method consistently classified specific estimated components of the signals acquired from a phantom (*Table C.1*) as metabolites that are not present in the phantom. A careful analysis of these results led to some interesting findings. For example, the structure of one of the non-phantom metabolites, l aspartic acid, strongly overlaps the structure of the phantom metabolite NAA (refer to *Fig. F.5* for the 2-dimensional structures), hence a slight overlap amongst their QMP spectral frequencies was expected. This is not the case (Fig, 5.1). Leading to the conclusion that the physico-chemical constants (*Tables F.2* & *F.4*, garnered from literature) used for the quantum mechanical predictions of NAA, are either inaccurate or incomplete.

- The performance of the classification method depends on the quality of the data that it receives.

The classification method was tested on FPT estimated mono-components. As already noted, the FPT has some important, and advantageous, properties: (1) it is not subjective, (2) it is a unique analysis method, (3) it is a non-linear analysis method, and (4) it does not assume the nature of the signal being analysed. The first two properties give it an advantage over the discussed interactive methods (*Interactive Methods*, *page 53*), and all four properties give it an advantage over the widely used non-interactive methods (*Non-Interactive Methods*, *page 38*). An additional property of the FPT is that it converges to its results.

The convergence property of the FPT is quite advantageous. It means that the limit of an estimated parameter value, if reached, will be equal to the original parameter value. However the limit of an estimated parameter value will be reached if and only if there are enough well-sampled data points, and if other properties, e.g. noise, do not slow down its convergence rate. This leads to an important question − can the FPT parameter estimates of a mono-component signal converge to a reasonable level of accuracy considering the **(a)** noise levels, and **(b)** lengths (usually 2048 data points) of well-sampled NMRS signals. The detailed mathematical, and comparative analysis, of *Chapter 3* addressed these questions to a certain level, and the conclusions are discussed next.

6.1.2 The Analysis of Non-interactive Methods.

It is rare for a review of methods to be conducted via a mathematical breakdown of the methods' algorithms. However the detailed mathematical review of the non-interactive methods (*Non-Interactive Methods*, *page 38*) highlights the advantage of this approach. Such an approach gives a clear view of the properties and/or algorithms that underly each method's strengths and weaknesses. This should make it easier for interested parties to develop any of the methods further. The observations made w.r.t. the non-interactive analysis methods are as follows

The Fast Padé Transform (FPT)

- The rate at which the estimated parameter values converge depends on the SNR, and noise complexity of a signal; convergence rate decreases with decreasing SNR and/or increasing noise complexity.
- The accuracy of the estimated parameters increases as the number of well-sampled signal data points increases because there are more points available for convergence.
- Once an estimated parameter value reaches its limit, i.e. the *correct* parameter value, all estimates thereafter are approximately equal, i.e. increasing the number of well-sampled signal data points beyond what is required to reach the limit of a parameter value does not change the parameter value.
- The FPT is good at detecting fast decaying signals. However the parameter values of fast decaying signals converge much more slowly than those of slow decaying signals. Hence much more well-sampled signal data points, than is currently the norm in NMRS, will

be required if the FPT is to be used for the detection of macromolecular & lipid signals. The number of well-sampled data points can be increased either by reducing the sampling time − thereby acquiring more data points within a time period, and hence retaining more continuous-signal information − or by increasing the length of the acquired signal without changing the sampling time. The latter will only be useful if a shorter signal means acquiring a signal that does not fully decay.

The SVD Based Methods

- The subjectivity and non-uniqueness of the SVD based methods are due to four parameters, termed structural parameters in this text, which affect the structure of the matrices and vectors which are directly or indirectly used to estimate the parameter values of an NMRS signal's components (*Singular value Decomposition Methods*, *page 38*). All four parameters depend on a user's input/assumptions.

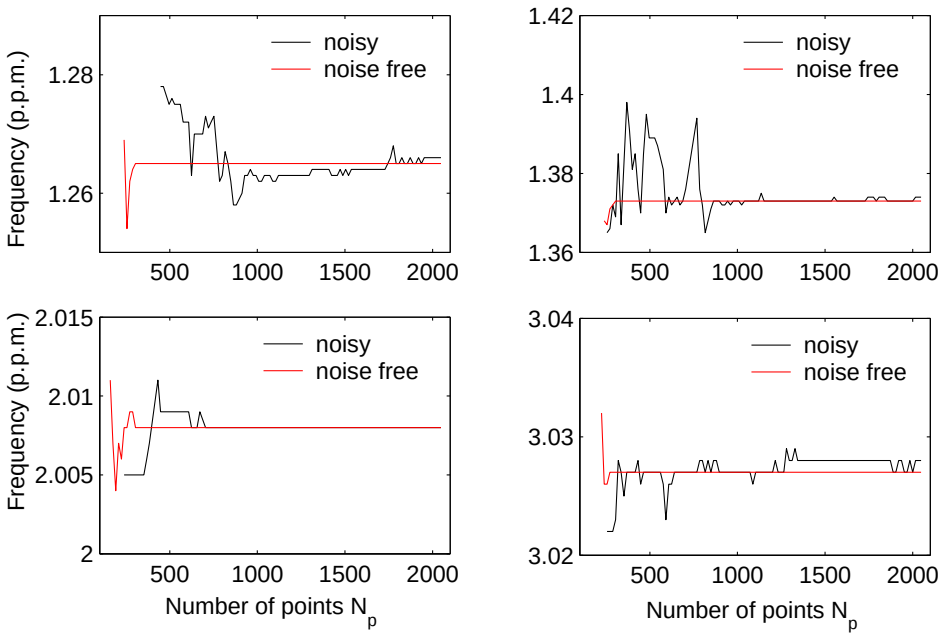
- The fitting that occurs per SVD based method, which are fitting methods, also depends on the four structural parameters, as the discussion of the SVD methods illustrates (*Section 3.1.1*). Implicitly, how well the noise and signal subspaces are separated is also structural parameter dependent.

- The values of these four parameters are constrained by the length of a well-sampled signal and/or how the methods are implemented in an NMRS signal analysis application. In the case of the jMRUI [74, 75] NMRS signal analysis application, the maximum number of assumed components per signal is 50.

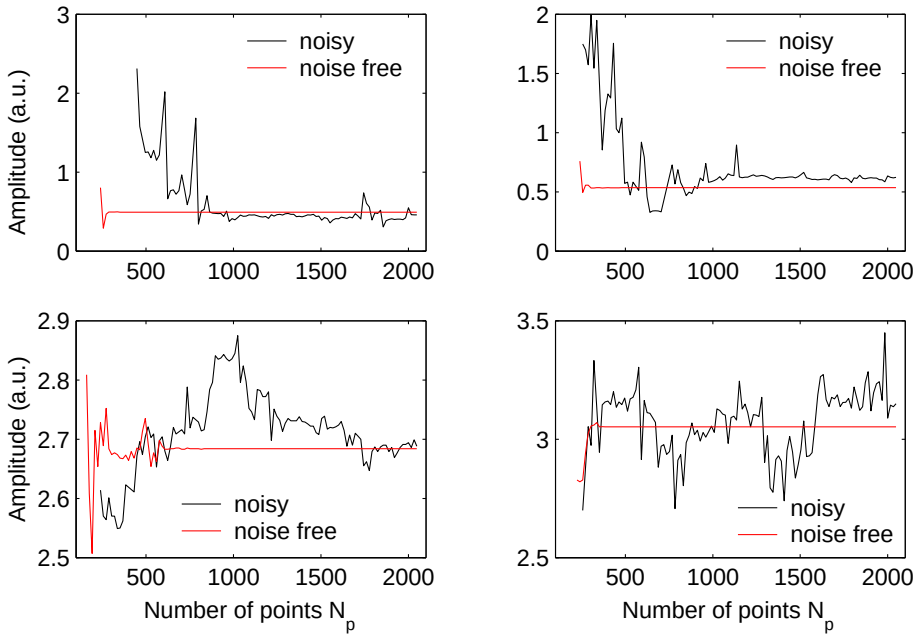
- Due to the above constraints, as the SNR of a signal decreases it becomes increasingly difficult for these methods to extract genuine, or physically meaningful, components from a signal. This is in line with the observations of [109].

- When the SNR of a signal is high, the level of accuracy of the estimates is quite good.

Due to the effect of noise on the convergence rate of the FPT's parameter values, e.g. *Fig. 6.1*, an FPT parameter value might not converge to its limit by the time a signal's data points have been exhausted. It is due to this noise and convergence rate dynamics of the FPT that the SVD based methods are sometimes more accurate than the FPT, especially when the SNR is high. However, as the noise level rises the number genuine components that the SVD based methods can deduce reduces rapidly (*The Effect of Noise on Decomposition Accuracy*, *page 66*), compared to the FPT. In fact the number of mono-components that the FPT can estimate, as the noise level rises, hardly changes (e.g. *Table 3.4*) − it is the accuracy of the estimates, due to slower convergence, that is affected. Hence, in order to make the best use of the FPT, longer well-sampled NMRS signals, than is currently the norm in NMRS studies (≈ 2048), might sometimes be required. Otherwise the FPT estimates, especially of fast decaying signals, might not be reasonably accurate.



(a)



(b)

Figure 6.1: The effect of noise on convergence rate. **(a)** −**Noise Free**: The noise free synthetic signal used in this text, *Table C.2*, was decomposed using the FPT, and the convergence behaviour of four of the estimated parameter values, every 16 points (i.e. step length 16), is recorded above. In each case the estimated parameters converged to the exact original value. −**Noisy**: The same noise free signal was corrupted with noise − the noisy signal's SNR is approximately 26.154dB − and the convergence behaviour of four parameters, similar to the previous four, is also recorded above. **(b)** The corresponding estimated amplitudes of the frequencies of **(a)**

6.2 Suggested Future Work

The following summarize future work proposals in relation to the observations and discussions of this thesis

Extensive Testing of the Classification Method The classification method needs to be tested much more extensively, and systematically, on human data, especially on data acquired from different NMR scanners. The latter will ensure that the classifier is not equipment specific. The use of the Generalized Likelihood Ratio Test (GLRT) also needs to be tested much more extensively. The development of a standard software package will definitely help.

A Comparison of Classification & Pattern Recognition Approaches Considering the fact that one of the proposed uses of NMRS in its early days was tissue characterization [30], it is not surprising that classification and pattern recognition has featured in NMRS studies in the past e.g. [39, 62, 73, 101]. However the aims and approaches of these studies are quite different from those of the classification method introduced in this thesis. The aims of the aforementioned studies, for example, were to

- differentiate between toxic/diseased spectra and non-toxic/healthy spectra [39, 62].
- classify tumour types via spectra patterns [101].
- assess the quality of acquired spectra by assessing their patterns within a spectral region [73].

i.e. it is a spectrum, rather than a mono-component, that is classified. And in contrast to the introduced classification method the classifier in each case was a trained classifier. The training data consisted of spectra that symptomized a state e.g. spectra representing tumorous and tumour free prostate tissue [62]. It would be worthwhile to compare the introduced classification method with that of a mono-component classifier whose classification features have been extracted from data sets acquired from different NMR scanners.

An Extensive Study of FPT Behaviour As previously noted, the FPT includes an internal validation via two complementary forms of the FPT: the FPT^{+}, and the FPT^{-}. The difference between these two forms is their convergence regions; FPT^{+} converges within the unit circle, FPT^{-} converges outwith the unit circle. Due to their different convergence regions each form converges to the *correct* parameter value from different ends, and during convergence the estimated values w.r.t. a specific parameter value are, in theory, the upper and lower boundaries of the *correct* parameter value [14]. Hence it might be possible to estimate much more accurate parameter values, especially w.r.t. noisy signals of limited length, if an appropriate interpolation method can be developed. This of course requires a much more in-depth understanding of convergence behaviour w.r.t. SNR, noise complexity, the length of well-sampled signals, and the characteristics of a signal's mono-component.

Appendix A

Algorithms

A.1 The Generalized Likelihood Ratio Test

Prior to classifying an estimated mono-component $\hat{\boldsymbol{x}}_n$ – the n^{th} $\mathcal{C}\times 1$ estimated mono-component w.r.t. the estimated parameter values $\hat{a}_n$, $\hat{\phi}_n$, $\hat{f}_n$ & $\hat{d}_n$ – the probability that $\hat{\boldsymbol{x}}_n$ is a signal is determined. This is achieved via the GLRT, a *composite hypothesis test statistic* (a hypothesis test statistic whose assumed distribution PDF, w.r.t. to an hypothesis, has at least one estimated parameter [59]) for discriminating between noise and genuine signals [59]. Consider the hypotheses

$$\begin{aligned}&\mathcal{H}_0 : \hat{\mathbf{x}}_n \text{ is a noise signal}\\ &\mathcal{H}_1 : \hat{\mathbf{x}}_n \text{ is a genuine mono-component time series signal of a metabolite}\end{aligned} \tag{A.1.1}$$

by the GLRT [59] $\hat{\boldsymbol{x}}_n$ is accepted under $\mathcal{H}_1$ if

$$\frac{p(\hat{\mathbf{x}}_n;\hat{a}_n,\hat{\phi}_n,\hat{f}_n,\hat{d}_n,\mathcal{H}_1)}{p(\hat{\mathbf{x}}_n;\mathcal{H}_0)} > \gamma, \quad \gamma \in \Re \tag{A.1.2}$$

whereby $p(\hat{\mathbf{x}}_n;\ldots,\mathcal{H}_i)$ is the probability that a signal $\hat{\boldsymbol{x}}_n$ is accepted under an hypothesis $\mathcal{H}_i$, and γ is a threshold value that depends on a pre-set probability of false alarm P_α. P_α (also known as Significance Level) is the probability of committing a Type I Error, i.e., and w.r.t. the defined hypotheses, the probability of accepting a signal under $\mathcal{H}_1$, albeit $\mathcal{H}_0$ is true. Hence a small P_α value should ensure that only a minimal number of noise signals are accepted as genuine signals. Before defining the relationship between P_α and γ, the probabilities of *Eq. A.1.2* are defined w.r.t. a distribution.

Consider a signal $x = 1$. Lets assume this signal is emitted a hundred times, and is noise-corrupted on detection. Lets also assume that the signals are independent, i.e. the emission or detection of one signal does not affect the emission or detection of another. If the noise is White Gaussian Noise (WGN), hence leading to a range of noisy signal values that are also normally distributed, e.g. *Fig. A.1*, then the natural logarithm of the left-hand-side quotient of *Eq. A.1.2* (in which case the signal is a vector rather than a scalar) is

$$\ln\left(\frac{p(\mathbf{x}_n;\hat{a}_n,\hat{\phi}_n,\hat{f}_n,\hat{d}_n,\mathcal{H}_1)}{p(\mathbf{x}_n;\mathcal{H}_0)}\right) = \frac{1}{\sigma^2} I(\hat{a}_n,\hat{\phi}_n,\hat{f}_n,\hat{d}_n) \tag{A.1.3}$$

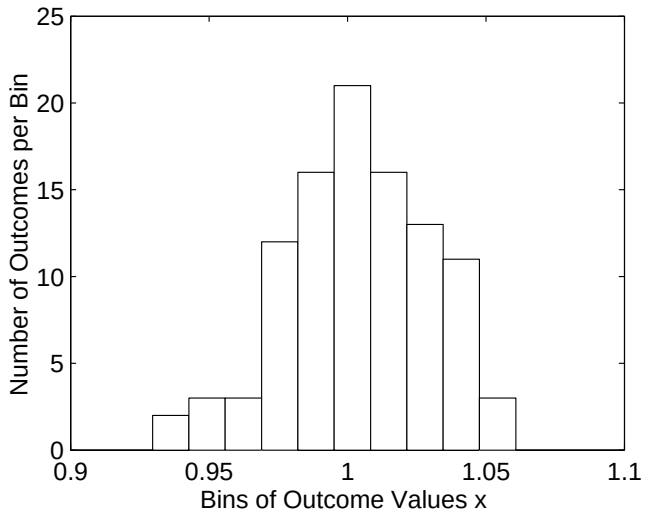


Figure A.1: The distribution of outcomes x, $\mathrm{x} \sim \mathcal{N}(1, 0.025^2)$

the variable σ is the standard deviation of the noise, and $I(\hat{a}_n, \hat{\phi}_n, \hat{f}_n, \hat{d}_n)$ or I_n is the periodogram of $\hat{\boldsymbol{x}}_n$; refer to Chapter 7 of [59] for a detailed proof of *Eq. A.1.3*. Hence $\hat{\boldsymbol{x}}_n$ is accepted under $\mathcal{H}_1$ if

$$max\, I_n > \sigma^2 \ln(\gamma) = \epsilon \tag{A.1.4}$$

The assumption of independence holds in NMRS. This is because the acquisition of an NMRS signal in one experiment is not affected by its acquisition in another, and by extension a metabolite's mono-components in one experiment do not depend on those of another; as long as the NMRS experiments are conducted according to standard practice, e.g. a reasonable time apart. Per experiment, the estimated mono-components are independent vectors since the spectral peaks of a metabolite depend on the chemical shift and coupling constant values **not** on each other. Hence, and in line with [59], a subset of

$$\begin{aligned} &I_1, \ldots, I_K \sim \chi^2_{\nu=2} \text{ under } \mathcal{H}_0 \\ &I_1, \ldots, I_K \sim \chi'^2_{\nu=2}(\lambda) \text{ under } \mathcal{H}_1 \end{aligned} \tag{A.1.5}$$

i.e. the corresponding $\mathbf{x}_n$ vectors in each case are either noise signals or genuine signals respectively. $\chi^2_{\nu=2}$ is the chi-squared PDF, and $\chi'^2_{\nu=2}(\lambda)$ is the non-central chi-squared PDF. λ is the non-centrality parameter, and in both cases ν is the degrees of freedom. *Eq. A.1.5* is based on the theory that if $\hat{\boldsymbol{x}}_1, \ldots, \hat{\boldsymbol{x}}_n$ are independent, and each $\hat{\boldsymbol{x}}_n \sim \mathcal{N}(\boldsymbol{\mu}_n, \sigma^2)$, then either $\sum \hat{\boldsymbol{x}}_n \sim \chi^2_\nu$ for cases $\boldsymbol{\mu}_n = \mathbf{0}$ or $\sum \hat{\boldsymbol{x}}_n \sim \chi'^2_\nu(\lambda)$ for cases $\boldsymbol{\mu}_n \neq \mathbf{0}$ [59,111]. Now the relationship between P_α and γ, via $\epsilon = \sigma^2 \ln(\gamma)$, can be defined [59]

$$P_\alpha = p\{I_n > \epsilon; \mathcal{H}_0\} \tag{A.1.6}$$

$$= Q_{\chi^2_2}\left(\frac{2\mathcal{C}\epsilon}{\sigma^2}\right) \tag{A.1.7}$$

$$= e^{(-\mathcal{C}\epsilon/\sigma^2)} \tag{A.1.8}$$

whereby $Q_{\chi^2_2}$ is the CDF of $\chi^2_{\nu=2}$. As previously noted, and for the defined hypotheses (*Eq. A.1.1*), P_α is the probability of committing a Type I Error, i.e. the probability of ac-

cepting a signal under $\mathcal{H}_1$, albeit $\mathcal{H}_0$ is true. However if a signal is accepted under $\mathcal{H}_0$ albeit $\mathcal{H}_1$ is true, $P(\mathcal{H}_0; \mathcal{H}_1)$, we have a Type II Error. The aim of any decision problem is to minimize $\mathrm{P}(\mathcal{H}_0; \mathcal{H}_1)$, or maximize $P_\delta = 1 - P(\mathcal{H}_0; \mathcal{H}_1)$, where P_δ is the probability of detection. P_δ is defined as [59]

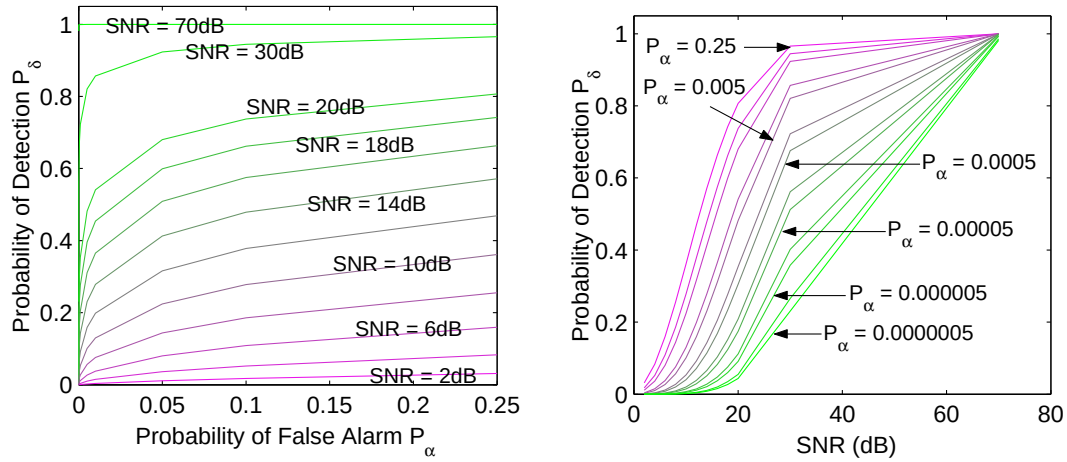
$$P_\delta = p\{I_n > \epsilon; \mathcal{H}_1\} \tag{A.1.9}$$

$$= Q_{\chi_2'^2(\lambda)}\left(2ln\frac{\mathcal{C}/2 - 1}{P_\alpha}\right) \tag{A.1.10}$$

whereby

$$\lambda = 10\log_{10}\frac{max(I_n)^2\mathcal{C}}{2\sigma^2} \tag{A.1.11}$$

The diagrams of *Fig. A.2* illustrate the behaviour of the GLRT.



(a) **(b)**

Figure A.2: The Generalized Likelihood Ratio Test. **(a)** The detection performance of the GLRT; these curves are also known as ROC curves. The SNR range, [0 70]dB, reflects the range observed in the experiments conducted for this project. The number of signal data points $\mathcal{C} = 2048$ for all curves. **(b)** The detection performance highlighting the range of P_δ per P_α over the experimentally observed range of signal-to-noise ratios using the same data values as **(a)**.

A.2 Minimization Algorithms

Let the objective function of a problem be

$$C = \|Y - F(\mathbf{x}, T)\|^2 \tag{A.2.1}$$

whereby Y and T are the observed data points, $\mathbf{x}$ is the set of parameters being estimated and F is the assumed functional form. The aim is to minimize C, which is the sum of the squared errors (SSE), w.r.t. the set of parameters $\mathbf{x}$. In general, the minimization of a function w.r.t. a set of parameters involves differentiating the function w.r.t. the parameters, and then evaluating the set of parameters that would lead to $\nabla C = \mathbf{0}$. The various minimization methods

or algorithms minimize different forms of the expression C. The methods discussed herein either minimize the function expressed above or its second order expansion about a point. The second order expansion of C, about a point $\mathbf{x}_0$, via Taylor's theorem is

$$\begin{aligned} C(\mathbf{x}) &= C(\mathbf{x}_0) + \frac{1}{1!}(\mathbf{x}-\mathbf{x}_0)^T \nabla C(\mathbf{x}_0) \\ &\quad + \frac{1}{2!}(\mathbf{x}-\mathbf{x}_0)^T \nabla^2 C(\mathbf{x}_0)(\mathbf{x}-\mathbf{x}_0) \end{aligned} \tag{A.2.2}$$

Let $R = Y - F(\mathbf{x}, T)$ i.e. $C(\mathbf{x}) = R(\mathbf{x})^T R(\mathbf{x})$ w.r.t. the parameters only, then

$$\frac{\partial C}{\partial \mathbf{x}} = \nabla C = \frac{\partial C}{\partial R}\frac{\partial R}{\partial \mathbf{x}} \tag{A.2.3}$$

whereby $\frac{\partial R}{\partial \mathbf{x}}$ is the Jacobian of R i.e. $J = \frac{\partial R}{\partial \mathbf{x}}$. Thus the expansion − *Eq. A.2.2* − can be expressed as

$$\begin{aligned} C(\mathbf{x}) &= R(\mathbf{x}_0)^T R(\mathbf{x}_0) + 2(\mathbf{x}-\mathbf{x}_0)^T J(\mathbf{x}_0)^T R(\mathbf{x}_0) \\ &\quad + (\mathbf{x}-\mathbf{x}_0)^T \left[J(\mathbf{x}_0)^T J(\mathbf{x}_0) + R(\mathbf{x}_0)^T \frac{\partial^2 R(\mathbf{x}_0)}{\partial \mathbf{x}^2} \right] (\mathbf{x}-\mathbf{x}_0) \end{aligned} \tag{A.2.4}$$

Differentiating with respect to the parameters leads to

$$\nabla C(\mathbf{x}) = 2J(\mathbf{x}_0)^T R(\mathbf{x}_0) + 2\left[J(\mathbf{x}_0)^T J(\mathbf{x}_0) + R(\mathbf{x}_0)^T \frac{\partial^2 R(\mathbf{x}_0)}{\partial \mathbf{x}^2} \right] (\mathbf{x}-\mathbf{x}_0) \tag{A.2.5}$$

Thus $\nabla C = \mathbf{0}$ implies

$$(\mathbf{x}-\mathbf{x}_0) = -\left[J(\mathbf{x}_0)^T J(\mathbf{x}_0) + R(\mathbf{x}_0)^T \frac{\partial^2 R(\mathbf{x}_0)}{\partial \mathbf{x}^2} \right]^{-1} J(\mathbf{x}_0)^T R(\mathbf{x}_0) \tag{A.2.6}$$

or

$$\mathbf{x}^{i+1} = \mathbf{x}^i - \left[J(\mathbf{x}^i)^T J(\mathbf{x}^i) + R(\mathbf{x}^i)^T \frac{\partial^2 R(\mathbf{x}^i)}{\partial \mathbf{x}^2} \right]^{-1} J(\mathbf{x}^i)^T R(\mathbf{x}^i) \tag{A.2.7}$$

Whereby i refers to the iteration number. Let

$$H(\mathbf{x}) = J(\mathbf{x})^T J(\mathbf{x}) \;\;\&\;\; \Upsilon(\mathbf{x}) = R(\mathbf{x})^T \frac{\partial^2 R(\mathbf{x})}{\partial \mathbf{x}^2} \tag{A.2.8}$$

then

$$\mathbf{x}^{i+1} = \mathbf{x}^i - \left[H(\mathbf{x}^i) + \Upsilon(\mathbf{x}^i) \right]^{-1} J(\mathbf{x}^i)^T R(\mathbf{x}^i) \tag{A.2.9}$$

The Gradient Descent (GD) Method

This method more or less minimizes the simplest, first order, form of C. The new and old set of parameters are related as follows

$$\mathbf{x}^{i+1} = \mathbf{x}^i - \alpha \nabla C \tag{A.2.10}$$

whereby

$$\nabla C = \frac{\partial C}{\partial R} J \tag{A.2.11}$$

and i refers to the iteration number. The iteration continues until the sum squared error (*Eq. A.2.1*), which is evaluated after each iteration, is zero or sufficiently small. This method is fast when far from a minimum point, but slow when close to a minimum. [40, 68]

The Newton Method

This method minimizes the second order form of C. In fact its algorithm is the same as *Eq. A.2.9* i.e.

$$\mathbf{x}^{i+1} = \mathbf{x}^i - \left[H(\mathbf{x}^i) + \Upsilon(\mathbf{x}^i)\right]^{-1} J(\mathbf{x}^i)^T R(\mathbf{x}^i)$$

The iteration continues until the sum squared error (*Eq. A.2.1*), which is evaluated after each iteration, is zero or sufficiently small. This method converges quickly near a minimum, but is unreliable far from it. [40, 68]

The Gauss-Newton (GN) Method

This method assumes that the values of the term

$$\Upsilon(\mathbf{x}) = R(\mathbf{x})^T \frac{\partial^2 R(\mathbf{x})}{\partial \mathbf{x}^2}$$

in Newton's algorithm are too small to have a significant effect on the progress of the algorithm. Hence the Gauss-Newton algorithm is

$$\mathbf{x}^{i+1} = \mathbf{x}^i - \left[H(\mathbf{x}^i)\right]^{-1} J(\mathbf{x}^i)^T R(\mathbf{x}^i) \tag{A.2.12}$$

Once again, the iteration continues until the sum squared error (*Eq. A.2.1*), which is evaluated after each iteration, is zero or sufficiently small.

The Levenberg-Marquardt (LM) Method

This method [72] is the minimization algorithm of VARPRO. This method is a hybrid of the Gradient Descent and Gauss-Newton methods. It takes advantage of the strength of each. For an initial set of parameter values far from their minimum point the Gradient Descent method minimizes the SSE at a much faster rate than the Gauss-Newton method. However, as the size of the SSE decreases, the rate at which GD algorithm minimizes the objective function decreases. On the other hand, the rate at which the GN method minimizes the SSE increases as the SSE decreases. The LM method switches between both methods depending on the size of the SSE after each iteration. Thus the LM algorithm is

$$\mathbf{x}^{i+1} = \mathbf{x}^i - \left[H(\mathbf{x}^i) + \lambda^2 diagonal(H)\right]^{-1} J(\mathbf{x}^i)^T R(\mathbf{x}^i) \tag{A.2.13}$$

The algorithm is an adaptation of Levenberg's algorithm in which the second term of the inverse expression is λ^2 rather than $\lambda^2 diagonal(H)$. A variety of methods have been used to evaluate per iteration. In the VARPRO metabolite quantification tool the method proposed by Osborne, 1972 [79] is used. In brief

If $\mathrm{SSE}^{i+1} < \mathrm{SSE}^i$

$\boldsymbol{x}^{i+1}$ is used for the next iteration.

λ is reduced by a pre-defined constant value < 1.

else

The $\boldsymbol{x}^{i+1}$, and other (i + 1) iteration values, are not used for the next iteration,

instead the $\boldsymbol{x}^{i}$ values are used to evaluate new $\boldsymbol{x}^{i+1}$ values, using a newλ value.

λ is increased by a pre-defined constant value > 1.

END

The LM method finds the local rather than global minimum [40].

The NL2SOL Method

This is the algorithm used by AMARES and AQSES. It is an odd mix of the Newton and GN methods. Basically, it does not assume that the $\Upsilon(\mathbf{x})$ term of *Eq. A.2.9* is insignificant. Instead $\Upsilon(\mathbf{x})$ is assigned an initial value e.g. $\Upsilon(\mathbf{x}) = \mathbf{0}$. Subsequent values are then evaluated via a number of proposed methods. For example, one of the method's algorithms is

$$\mathbf{x}^{i+1} = \mathbf{x}^{i} - \left[H(\mathbf{x}^{i}) + \lambda^{i} D^{2}(\mathbf{x}^{i})\right]^{-1} J(\mathbf{x}^{i})^{T} R(\mathbf{x}^{i}) \tag{A.2.14}$$

i.e. $\Upsilon(\mathbf{x}) = \lambda^{i} D^{2}(\mathbf{x}^{i})$. The term $D(\mathbf{x}^{i})$ is

$$D^{i} = \begin{pmatrix} d_1^i & 0 & \dots & 0 \\ 0 & d_2^i & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & d_p^i \end{pmatrix} \tag{A.2.15}$$

whereby $D^{i} \equiv D(\mathbf{x}_i)$ and

$$d_i^i = max\left\{\left[\left\|J_{.,l}(\mathbf{x}^{i})\right\|_2^2 + max\left\{0, \Upsilon_{ll}(\mathbf{x}^{i-1})\right\}\right]^{\frac{1}{2}}, 0.6 d_l^{i-1}\right\} \tag{A.2.16}$$

The subscript l refers to the l^{th} row, column or term, depending on the context e.g. $J_{.,l}$ implies the l^{th} column of the Jacobian. λ^{i} is evaluated via a method akin to that detailed in the Levenberg-Marquardt section above. An important property of this algorithm is the flexible $\Upsilon(\mathbf{x})$ update rule. Due to this property a parameter being estimated can be constrained. For example, if parameter x (of $\mathbf{x}$) is such that $x \in [a, b]$, a & $b \in \Re$, if after an iteration $x^{i+1} \notin [a, b]$ then x^{i+1} & $\Upsilon(x^{i})$ are rejected, and a new $\Upsilon(x^{i})$ is evaluated using a new λ^{i}.

Appendix B

Definitions

The Periodogram

The periodogram of a mono-component signal of length N, and sampling time τ, is defined as [83]

$$p(\omega_j) = \frac{1}{N}\left|\sum_{n=0}^{N-1} x_n e^{i\omega_j n\tau}\right|^2 \tag{B.0.1}$$

whereby

$$w_j = 2\pi j/N\tau, \qquad j = -N/2+1 : 1 : N/2-1$$

Normalization

The normalized form of a vector U is given by

$$\frac{2(U - min(U))}{max(U) - min(U)} - 1 \tag{B.0.2}$$

Mono-component

A mono-component, of a time series signal, is a signal formed by a single set of signal parameters, i.e. frequency, amplitude, decay, and phase parameters.

Multi-component Signal

In this text, a multi-component signal is a time series signal consisting of two or more mono-component time series signals.

Component

Sometimes, a decomposition method may not be able to resolve the overlapping mono-components, i.e. mono-components of the same frequency but different decay rates, of a multi-component signal. Hence some of the decomposition method's estimated mono-components may not [strictly] be mono-components. In order to highlight such possibilities, the term component rather than mono-component is used to refer to the estimated mono-components of relevant decomposition methods.

Appendix C

NMRS Data Sources

GE Head Phantom

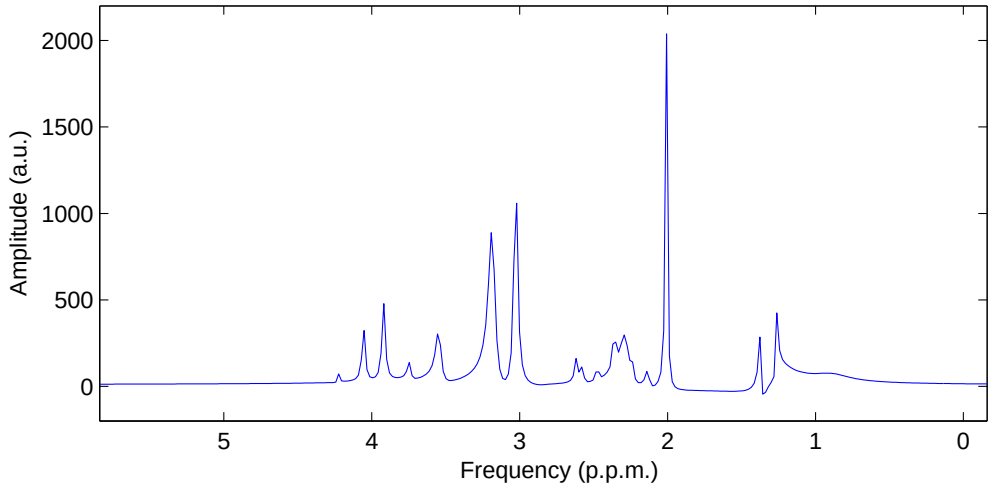
Chemical	Concentration (mM)
N-acetyl-l-aspartic Acid	12.5
Creatine [Hydrate]	10
Choline [Chloride]	3
Myo-inositol	7.5
L-Glutamate [Monosodium]	12.5
D L Lactate [Lithium Salt]	5

Table C.1: The chemical contents of the General Electric Phantom. Creatine Hydrate (Creatine + H_20), Choline Chloride (Choline + Chlorine + H_20), L-Glutamate Monosodium (Glutamate + Sodium + H_20), D L Lactate Lithium Salt (Lactate + Lithium). The phantom also includes the chemical compounds Monobasic Potassium Phosphate (50 mM), Sodium Hydroxide (56 mM), Sodium Azide (0.10%), and Magnavest (0.10%) for controlling or preserving a variety of properties, e.g. pH level.

Human Volunteers

All human brain studies were conducted "at the SFC Brain Imaging Research Centre (www.sbirc.ed.ac.uk) using a protocol approved by the Local Research Ethics Committee (LREC/1998/3/1; 'Assessment of Magnetic Resonance Imaging with Healthy Volunteers')"

Synthetic Data



(a)

	f (p.p.m.)	a	d (Hz)	ϕ (°)		f (p.p.m.)	a	d (Hz)	ϕ (°)
1	1.265	0.492	3.278	-33.000	**13**	2.287	2.051	16.096	21.000
2	1.373	0.534	3.677	37.000	**14**	2.133	0.274	6.696	24.565
3	2.008	2.681	2.855	5.948	**15**	2.369	0.061	1.246	0.000
4	2.613	0.299	4.821	10.582	**16**	2.238	0.157	4.340	9.000
5	2.475	0.144	1.954	2.616	**17**	3.744	0.219	5.473	22.762
6	2.573	0.155	2.273	1.751	**18**	2.348	0.335	5.950	34.000
7	4.228	0.053	1.688	0.000	**19**	0.890	1.000	62.518	-60.000
8	3.027	3.045	5.822	-2.608	**20**	1.300	2.340	38.956	-60.000
9	3.920	0.932	5.055	0.097	**21**	2.050	0.150	27.889	-60.000
10	3.183	4.043	10.919	19.000	**22**	2.240	0.160	35.127	-60.000
11	3.544	1.050	8.666	23.000	**23**	2.810	0.100	34.276	-60.000
12	4.055	0.620	4.837	3.500	**24**	0.890	1.000	62.518	60.000

(b)

Table C.2: (a) The Fourier transform of the artificial *noise free* time series signal X_{anf}, of length 2048, whose mono-component parameters are detailed in (b). The sampling time of X_{anf} is 0.4 ms. The first eighteen frequency f parameter values are the QMP spectral frequencies of Lactate (1,2), NAA ($3-7$), Creatine (8,9), Choline (10), Myoinositol (11,12), L Glutamate ($13-18$). The amplitude a, decay d, and phase ϕ parameter values are based on the analysis of signals acquired from the GE Phantom, which contains the same set of metabolites. The mono-components set 19 to 24 is a subset of the macromolecular and lipid mono-components defined by Seeger et al. [91]; the phase values of these mono-components are arbitrary because no phase values were included in [91]

Appendix D

Experiment and Simulation Parameters

Experiment Parameters

	Δ	ϑ	Δ	ϑ	Δ	ϑ	Δ	ϑ
Selective Pulses - s.c.	1.8	90	1.8	90	1.8	90	1.8	90
Free Precession	5.6	0	5.6	0	5.6	0	5.6	0
Selective Pulses - s.c.	5.2	180	5.2	180	5.2	180	5.2	180
Free Precession	12.3	0	24.8	0.0	42.3	0	67.3	0
Selective Pulses - s.c.	5.2	180	5.2	180	5.2	180	5.2	180
Free Precession	4.9	0	17.4	0	34.9	0	59.9	0.0
Detection Period	0	0	0	0	0	0	0	0
T_E (Sum)	35.0		60.0		95.0		145.0	

Table D.1: The NMRS scanning parameters per PRESS pulse sequence type. Per experiment the sampling time $\tau = 0.4$ ms, the length N of each acquired time series signal was 2048, the repetition time $T_R = 1.5s$, and various NSA values were used, hence they have been defined clearly in the text. Equipment: 1.5 tesla General Electric Signa Scanner (Western General Hospital, Edinburgh). Δ : Duration (ms). ϑ: Pulse Angle (°). s.c.: Sinus Cardinal.

Simulation Parameters

	Δ	ϑ	Δ	ϑ	Δ	ϑ	Δ	ϑ
Selective Pulses - s.c.	1.8	90	1.8	90	1.8	90	1.8	90
Free Precession (SC)	5.6	0	5.6	0	5.6	0	5.6	0
Selective Pulses - s.c.	5.2	180	5.2	180	5.2	180	5.2	180
Free Precession (SC)	12.3	0	24.8	0.0	42.3	0	67.3	0
Selective Pulses - s.c.	5.2	180	5.2	180	5.2	180	5.2	180
Free Precession (SC)	4.9	0	17.4	0	34.9	0	59.9	0.0
Detection Period	0	0	0	0	0	0	0	0
T_E (Sum)	35.0		60.0		95.0		145.0	

Table D.2: The NMR SCOPE [45, 74, 75, 105] simulation parameters, per pulse sequence type, used for simulating the time-dependent NMR behaviour of metabolites. The parameter values are similar to those of the real experiments − as tabulated above. 2048 complex data points − sampling step (time) 0.4 ms − were acquired from each simulated experiment. Δ : Duration (ms). ϑ: Pulse Angle (°). SC: Strong Coupling. s.c.: Sinus Cardinal. The phase and sinus cardinal (s.c.) shift value Nu_{rf} was always set to zero.

Appendix E

Additional Diagrams and Tables

	% Error (1. Noise Level 5%)					% Error (2. Noise Level 5%)				
	HSVD	HLSVD	HTLS	HLTLS	FPT	HSVD	HLSVD	HTLS	HLTLS	FPT
1.265	0.057	0.107	0.065	0.129	0.158	0.170	0.168	0.196	0.195	0.079
1.373	0.076	0.084	0.086	0.091	0.000	0.096	0.095	0.097	0.096	0.073
2.008	0.016	0.016	0.016	0.016	0.000	0.019	0.019	0.019	0.019	0.000
2.613	0.051	0.061	0.052	0.060	0.000	0.015	0.009	0.010	0.007	0.038
2.475	0.056	0.050	0.052	0.047	0.121	0.067	0.068	0.067	0.069	0.040
2.573	0.066	0.064	0.062	0.061	0.039	0.050	0.034	0.046	0.030	0.000
4.228	0.013	0.012	0.014	0.013	0.047					0.118
3.027	0.018	0.018	0.018	0.018	0.000	0.017	0.015	0.017	0.015	0.000
3.920	0.019	0.019	0.019	0.019	0.000	0.009	0.009	0.010	0.009	0.000
3.183	0.020	0.018	0.019	0.018	0.000	0.012	0.009	0.012	0.010	0.000
3.544	0.021	0.022	0.021	0.021	0.028	0.014	0.012	0.014	0.012	0.000
4.055	0.005	0.005	0.004	0.005	0.000	0.016	0.015	0.015	0.015	0.000
2.287	0.189	0.173	0.169	0.152	0.044	0.476	0.566	0.366	0.400	0.087
2.133	0.153	0.176	0.163	0.179	0.281	0.134	0.125	0.130	0.125	0.094
2.369									0.055	
2.238						0.133	0.091	0.116	0.071	0.045
3.744	0.025	0.025	0.026	0.025	0.053	0.067	0.074	0.065	0.071	0.160
2.348					0.000	1.063		0.884	0.564	1.022
0.890	4.104	2.406	5.679	3.132	5.955					3.820
1.300	0.668									
2.050					1.024					
2.240	0.118	0.123	0.134	0.133	0.045					
2.810					1.068	1.205		1.186		1.032
0.890										

Table E.1: Reproducibility assessment: *Section 3.3.3.* The percentage errors between the original frequencies and those estimated by the analysis methods.

	% **Error (3. Noise Level** 5%)					% **Error (4. Noise Level** 5%)				
	HSVD	HLSVD	HTLS	HLTLS	FPT	HSVD	HLSVD	HTLS	HLTLS	FPT
1.265	0.128	0.123	0.155	0.153	0.079	0.109	0.103	0.118	0.113	0.079
1.373	0.060	0.060	0.061	0.061	0.146	0.034	0.032	0.036	0.036	0.073
2.008	0.020	0.020	0.020	0.020	0.000	0.018	0.018	0.018	0.018	0.000
2.613	0.036	0.036	0.033	0.033	0.230	0.061	0.064	0.060	0.063	0.000
2.475	0.011	0.010	0.012	0.011	0.000	0.058	0.061	0.061	0.064	0.081
2.573	0.017	0.017	0.018	0.018	0.233	0.007	0.007	0.006	0.006	0.039
4.228	0.052	0.054	0.052	0.054		0.013	0.010	0.014	0.011	0.118
3.027	0.021	0.021	0.021	0.021	0.000	0.019	0.019	0.019	0.019	0.033
3.920	0.014	0.014	0.015	0.015	0.026	0.000	0.000	0.000	0.000	0.000
3.183	0.028	0.028	0.028	0.028	0.031	0.012	0.012	0.012	0.012	0.000
3.544	0.028	0.028	0.028	0.028	0.028	0.017	0.018	0.016	0.017	0.113
4.055	0.002	0.003	0.003	0.003	0.000	0.023	0.022	0.022	0.022	0.000
2.287	0.066	0.068	0.051	0.054		0.103	0.145	0.101	0.122	0.831
2.133	0.031	0.028	0.037	0.035	0.328	0.166	0.165	0.168	0.167	0.094
2.369										
2.238						0.199	0.225	0.186	0.213	
3.744	0.033	0.035	0.036	0.037	0.053	0.071	0.072	0.070	0.070	0.053
2.348						0.348	0.372	0.314	0.300	0.852
0.890	41.314	42.453	41.220	42.338	1.461	1.368	0.837	1.568	1.059	2.135
1.300					2.385		1.260			0.077
2.050										0.195
2.240	0.054	0.056	0.061	0.066	0.134					
2.810					0.819					1.281
0.890										

Table E.2: Reproducibility assessment: *Section 3.3.3.* The percentage errors between the original frequencies and those estimated by the analysis methods.

	% **Error (Noise Level** 5%)					% **Error (Noise Level** 10%)				
	HSVD	HLSVD	HTLS	HLTLS	FPT	HSVD	HLSVD	HTLS	HLTLS	FPT
1.265	0.029	0.026	0.035	0.035	0.079	0.002	0.001	0.096	0.096	0.079
1.373	0.050	0.051	0.045	0.045	0.073	0.231	0.238	0.219	0.229	0.146
2.008	0.017	0.017	0.018	0.018	0.000	0.020	0.019	0.019	0.019	0.000
2.613	0.040	0.039	0.039	0.039	0.000	0.117	0.108	0.110	0.108	0.000
2.475	0.026	0.026	0.022	0.022	0.081	0.084	0.087	0.081	0.081	0.000
2.573	0.010	0.010	0.012	0.012	0.039	0.066	0.052	0.036	0.029	0.194
4.228		0.010								0.237
3.027	0.016	0.016	0.016	0.016	0.000	0.023	0.024	0.024	0.025	0.033
3.920	0.030	0.030	0.030	0.030	0.026	0.010	0.010	0.007	0.008	0.000
3.183	0.023	0.023	0.022	0.022	0.031	0.044	0.046	0.046	0.047	0.063
3.544	0.008	0.007	0.007	0.007	0.056	0.017	0.014	0.014	0.013	0.056
4.055	0.009	0.009	0.009	0.009	0.000	0.020	0.020	0.023	0.022	0.025
2.287	0.060	0.061	0.060	0.061	0.000	0.024	0.282	0.250	0.346	
2.133	0.090	0.087	0.107	0.107	0.375	0.035	0.091	0.011	0.080	0.234
2.369					0.042					0.633
2.238									0.491	
3.744	0.002	0.003	0.002	0.002	0.080	0.138	0.149	0.149	0.158	0.000
2.348	0.399	0.401	0.380	0.380	0.085			0.464		0.341
0.890	0.758	0.939	0.153	0.260	2.584					2.472
1.300	0.189	0.271	0.229	0.221	1.231					2.154
2.050					0.488					0.537
2.240			0.029	0.030				0.259		
2.810					1.922					
0.890										

Table E.3: Estimation accuracy and increasing noise levels: *Section 3.3.4*. The percentage errors between the original frequencies and those estimated by the analysis methods.

	% **Error (Noise Level** 15%)					% **Error (Noise Level** 20%)				
	HSVD	HLSVD	HTLS	HLTLS	FPT	HSVD	HLSVD	HTLS	HLTLS	FPT
1.265			0.419	0.425	0.079					0.237
1.373	0.344	0.345	0.359	0.360	0.510	0.596	0.533	0.594	0.737	0.364
2.008	0.016	0.016	0.016	0.016	0.000	0.029	0.025	0.032	0.025	0.000
2.613	0.074	0.076	0.061	0.061	0.077					0.115
2.475			0.345	0.345	0.162					0.040
2.573					0.039	0.148		0.091	0.009	0.272
4.228										0.284
3.027	0.041	0.041	0.041	0.041	0.033	0.012	0.012	0.017	0.021	0.066
3.920	0.028	0.027	0.035	0.034	0.026	0.034	0.036	0.038	0.039	0.000
3.183	0.071	0.072	0.071	0.071	0.094	0.003	0.007	0.012	0.015	0.314
3.544	0.033	0.033	0.038	0.038	0.141	0.072	0.065	0.048	0.042	0.226
4.055	0.028	0.029	0.020	0.020	0.025		0.015	0.012	0.010	1.628
2.287	0.140	0.104	0.349	0.339	0.306				0.264	0.394
2.133	0.138	0.122	0.131	0.116	0.469					1.453
2.369										
2.238					0.492					1.296
3.744			0.400	0.524	0.855					1.469
2.348					0.681					
0.890	92.174	98.749	92.101	98.751	2.135	61.088	61.352	73.386	73.741	0.112
1.300										
2.050					1.317					
2.240										
2.810					0.320			2.492		0.071
0.890										

Table E.4: Estimation accuracy and increasing noise levels: *Section 3.3.4*. The percentage errors between the original frequencies and those estimated by the analysis methods.

Appendix F

Metabolite Constants

The developed classification method classifies estimated mono-component signals using the frequencies and probabilistic intensities predicted, per metabolite, by NMR quantum theory. These predictions depend on metabolite's set of chemical shifts δ and coupling constants J_{ab} [60]; in addition to the pulse sequence parameters of an experiment. The constants used for this project were obtained from a variety of sources, and the values and references are detailed in the tables below. *Table F.1* (structures) should be used in conjunction with *Table F.2*, and *Table F.3* (structures continued) should be used in conjunction with *Table F.4*.

	Metabolites	Chemical Formulae
1	Acetate	$\overset{4}{H_3}\mathrm{C}-\overset{2}{\mathrm{C}}(=\overset{3}{\mathrm{O}})-\overset{1}{O}H$
2	Betaine	$\overset{5}{H_3}\mathrm{C}-\overset{4}{\mathrm{N}}_+(\overset{7}{C}H_3,\overset{8}{C}H_3)-\overset{3}{C}H_2-\overset{2}{\mathrm{C}}(=\overset{1}{\mathrm{O}},\overset{6}{\mathrm{O}}_-)$
3	Choline	$\overset{5}{H_3}\mathrm{C}-\overset{4}{\mathrm{N}}_+(\overset{6}{C}H_3,\overset{7}{C}H_3)-\overset{3}{C}H_2-\overset{2}{C}H_2-\overset{1}{O}H$
4	Creatine	$\overset{1}{H}O-\overset{2}{\mathrm{C}}(=\overset{3}{\mathrm{O}})-\overset{4}{C}H_2-\overset{5}{\mathrm{N}}(\overset{6}{C}H_3)-\overset{7}{\mathrm{C}}(\overset{8}{\mathrm{N}}H_2)=\overset{9}{\mathrm{N}}H$
5	Ethanol	$\overset{3}{H_3}\mathrm{C}-\overset{2}{C}H_2-\overset{1}{O}H$
6	Ethanolamine	$\overset{1}{H}O-\overset{2}{C}H_2-\overset{3}{C}H_2-\overset{4}{\mathrm{N}}H_2$
7	GABA	$\overset{7}{H_2}\mathrm{N}-\overset{6}{C}H_2-\overset{5}{C}H_2-\overset{4}{C}H_2-\overset{2}{\mathrm{C}}(=\overset{3}{\mathrm{O}})-\overset{1}{O}H$
8	Glycerol	$\overset{6}{H}O-\overset{5}{C}H_2-\overset{2}{C}H(\overset{1}{O}H)-\overset{3}{C}H_2-\overset{4}{O}H$
9	Glycine	$\overset{5}{H_2}\mathrm{N}-\overset{4}{C}H_2-\overset{2}{\mathrm{C}}(=\overset{3}{\mathrm{O}})-\overset{1}{O}H$
10	Glutathione	$\overset{20}{\mathrm{O}}=\overset{17}{\mathrm{C}}(\overset{18}{\mathrm{O}}H)-\overset{11}{C}H(\overset{12}{\mathrm{N}}H_2)-\overset{10}{C}H_2-\overset{9}{C}H_2-\overset{5}{\mathrm{C}}(=\overset{6}{\mathrm{O}})-\overset{4}{\mathrm{N}}H-\overset{3}{C}H(\overset{2}{H_2}\mathrm{C}-\overset{1}{\mathrm{S}}H)-\overset{7}{\mathrm{C}}(=\overset{8}{\mathrm{O}})-\overset{13}{\mathrm{N}}H-\overset{14}{C}H_2-\overset{15}{\mathrm{C}}(\overset{16}{O}H)=\overset{19}{\mathrm{O}}$
11	GPC	$\overset{16}{H}O-\overset{15}{C}H_2-\overset{9}{C}H(\overset{10}{O}H)-\overset{8}{C}H_2-\overset{7}{O}-\overset{1}{\mathrm{P}}(=\overset{11}{\mathrm{O}},\overset{12}{\mathrm{O}}_-)-\overset{2}{O}-\overset{3}{C}H_2-\overset{4}{C}H_2-\overset{5}{\mathrm{N}}_+(\overset{13}{C}H_3,\overset{14}{C}H_3)-\overset{6}{C}H_3$
12	LAlanine	$\overset{6}{H_3}\mathrm{C}-\overset{4}{C}H(\overset{5}{\mathrm{N}}H_2)-\overset{2}{\mathrm{C}}(=\overset{3}{\mathrm{O}})-\overset{1}{O}H$
13	LAspartic	$\overset{9}{\mathrm{O}}=\overset{7}{\mathrm{C}}(\overset{8}{O}H)-\overset{6}{C}H_2-\overset{4}{C}H(\overset{5}{\mathrm{N}}H_2)-\overset{2}{\mathrm{C}}(=\overset{3}{\mathrm{O}})-\overset{1}{\mathrm{OH}}$
14	LGlutamic	$\overset{10}{\mathrm{O}}=\overset{8}{\mathrm{C}}(\overset{9}{O}H)-\overset{7}{C}H_2-\overset{6}{C}H_2-\overset{4}{C}H(\overset{5}{\mathrm{N}}H_2)-\overset{2}{\mathrm{C}}(=\overset{3}{\mathrm{O}})-\overset{1}{O}H$
15	LGlutamine	$\overset{10}{H_2}\mathrm{N}-\overset{8}{\mathrm{C}}(=\overset{9}{\mathrm{O}})-\overset{7}{C}H_2-\overset{6}{C}H_2-\overset{4}{C}H(\overset{5}{\mathrm{N}}H_2)-\overset{2}{\mathrm{C}}(=\overset{3}{\mathrm{O}})-\overset{1}{O}H$
16	LLactate	$\overset{1}{H}O-\overset{2}{C}H(\overset{3}{C}H_3)-\overset{4}{\mathrm{C}}(\overset{5}{O}H)=\overset{6}{\mathrm{O}}$

Table F.1: Annotated chemical structures. The blue numbers above the non-hydrogen atoms are used to identify the hydrogen atoms associated with the chemical shifts and coupling constants of *Table F.2*. For example, acetate's δ_4 (Table F.2) refers to the nuclei chemical shifts of the hydrogen atoms attached to [carbon] atom 4, and J_4 is the coupling constant between any pairing of these hydrogen atoms. ⟨ ⟩: Ring. =: Double Bond.

	Metabolite	Shifts (p.p.m.)	Coupling Constants (Hz)	References
1	Acetate	$\delta_4 = 1.9040$	$J_4 = 13$	[2, 43]
2	Betaine	$\delta_5 = \delta_7 = \delta_8 = 3.2540, \delta_3 = 3.8900$	$J_5 = J_7 = J_8 = 14.6, J_3 = 17$	[2, 104]
3	Choline	$\delta_5 = \delta_6 = \delta_7 = 3.1825, \delta_3 = 3.5010, \delta_2 = 4.0540$	$J_{3a,2a} = 3.14, J_{3a,2b} = 7.011, J_{3b,2a} = 6.979, J_{3b,2b} = 3.168$	[43], [1]
4	Creatine	$\delta_4 = 3.9200, \delta_6 = 3.027$	$J_4 = 18, J_6 = 13.71$	[2, 43]
5	Ethanol	$\delta_3 = 1.17, \delta_2 = 3.641$	$J_3 = 0.93, J_2 = -10.27, J_{3,2} = 7$	[1, 2, 104] [2]
6	Ethanolamine	$\delta_2 = 3.8184, \delta_3 = 3.1467$	$J_{2a,3a} = 3.897, J_{2a,3b} = 6.794, J_{2b,3a} = 6.694, J_{2b,3b} = 3.798$	[43]
7	GABA	$\delta_6 = 2.284, \delta_5 = 1.889, \delta_4 = 3.0128$	$J_{4a,5a} = 5.372, J_{4a,5b} = 7.127, J_{4b,5a} = 10.578, J_{4b,5b} = 6.982$	[43]
			$J_{5a,6b} = 7.432, J_{5b,6a} = 6.173, J_{5b,6b} = 7.933, J_{5a,6a} = 7.755$	
8	Glycerol	$\delta_{3a} = \delta_{5a} = 3.5522, \delta_{3b} = \delta_{5b} = 3.6402, \delta_2 = 3.7704$	$J_{3a,3b} = J_{5a,5b} = -11.715, J_{3a,2} = J_{2,5a} = 4.427, J_{3b,2} = J_{2,5b} = 6.485$	[43]
9	Glycine	$\delta_4 = 3.548$	$J_4 = -17.42$	[2, 43]
10	Glutathione	$\delta_{11} = 3.769, \delta_{10a} = 2.159, \delta_{10b} = 2.146, \delta_{9a} = 2.51, \delta_{9b} = 2.56$	$J_{11,10a} = 6.34, J_{11,10b} = 6.36, J_{10a,10b} = -15.48, J_{10a,9a} = J_{10b,9b} = 6.7$	[43]
		$\delta_4 = 8.1770, \delta_3 = 4.5608, \delta_{2a} = 2.9264, \delta_{2b} = 2.9747$	$J_{10a,9b} = J_{10b,9a} = 7.6, J_{3,2a} = 7.09, J_{3,2b} = 4.71, J_{2a,2b} = -14.06$	
		$\delta_{13} = 7.154, \delta_{14} = 3.769$	$J_{14} = 17.90$	
11	GPC	$\delta_{15a} = 3.605, \delta_{15b} = 3.672, \delta_9 = 3.903, \delta_{8a} = 3.871, \delta_{8b} = 3.946$	$J_{15a,9} = J_{9,8a} = 5.77, J_{15b,9} = J_{9,8b} = 4.53, J_{3a,4a} = J_{3b,4b} = 3.1$	[43]
		$\delta_3 = 4.312, \delta_4 = 3.659, \delta_{13} = \delta_4 = \delta_6 = 3.212$	$J_{3a,4b} = J_{3b,4a} = 5.9, J_{13} = J_{14} = J_6 = 14.6$	
12	LAlanine	$\delta_6 = 1.4667, \delta_4 = 3.7746$	$J_{4,6a} = 7.234, J_6 = -14.366$	[43]
13	LAspartic	$\delta_{6a} = 2.8011, \delta_{6b} = 2.6533, \delta_4 = 3.8914$	$J_6 = -17.426, J_{6a,4} = 3.647, J_{6b,4} = 9.107$	[43]
14	LGlutamic	$\delta_{7a} = 2.36, \delta_{7b} = 2.37, \delta_{6a} = 2.05, \delta_{6b} = 2.14, \delta_4 = 3.75$	$J_7 = -15.89, J_{7a,6a} = 6.43, J_{7a,6b} = 8.39, J_{7b,6a} = 8.47, J_{7b,6b} = 6.89$	[103]
			$J_{6a,6b} = -14.85, J_{6a,4} = 7.33, J_{6b,4} = 4.65$	
15	LGlutamine	$\delta_{7a} = 2.432, \delta_{7b} = 2.454, \delta_{6a} = 2.129, \delta_{6b} = 2.109, \delta_4 = 3.753$	$J_{7a,6a} = 9.165, J_{7a,6b} = 6.324, J_{7b,6a} = 6.347, J_{7b,6b} = 9.209, J_{6a,4} = 5.847$	
			$J_{6b,4} = 6.5, J_{7a,7b} = -15.371, J_{6a,6b} = -14.504$	[103]
16	LLactate	$\delta_2 = 4.1, \delta_3 = 1.32$	$J_2 = 13.1, J_{2,3} = 6.9333$	[2, 43, 116]

Table F.2: The chemical shifts, and coupling constants, per metabolite. The references [104], [116], and [115] each have an on-line database − Biological Magnetic Resonance Databank (www.bmrb.wisc.edu/), Human Metabolome Database (www.hmdb.ca/), and Institut für Organische Chemie (www-user.uni-bremen.de/ wie/) at Universität Bremen respectively. [2] Page 44 of [1]

	Metabolites	Chemical Formulae
17	LPhenylalanine	$\left\langle 1 \Leftarrow \overset{2}{C}H - \overset{3}{C}H = \overset{4}{C}H - \overset{5}{C}H = \overset{6}{C}H \rightarrow 1 \right\rangle \overset{1}{C} - \overset{7}{C}H_2 - \overset{8}{C}H(\overset{9}{N}H_2) - \overset{10}{C}(\overset{11}{O}H) = \overset{12}{O}$
18	LSerine	$\overset{1}{H}O - \overset{2}{C}H_2 - \overset{3}{C}H(\overset{4}{N}H_2) - \overset{5}{C}(\overset{6}{O}H) = \overset{7}{O}$
19	LThreonine	$\overset{1}{H}O - \overset{2}{C}H(\overset{5}{C}H_3) - \overset{3}{C}H(\overset{4}{N}H_2) - \overset{6}{C}(\overset{7}{O}H) = \overset{8}{O}$
20	LTryptophan	$\left\langle 1 \leftarrow \underset{\Downarrow 4}{\overset{5}{C}} - \overset{6}{C}H = \overset{7}{C}H - \overset{8}{C}H = \overset{9}{C}H - \underset{\Downarrow 5}{\overset{4}{C}} \rightarrow 3 \right\rangle \left\langle \rightarrow \overset{1}{N}H - \overset{2}{C}H = \overset{3}{C}\swarrow_{\rightarrow 10} \right\rangle \overset{10}{C}H_2 - \overset{11}{C}H(\overset{12}{N}H_2) - \overset{13}{C}(\overset{14}{O}H) = \overset{15}{O}$
21	LTyrosine	$\overset{7}{H}O - \overset{1}{C}\left\langle 1 \Leftarrow \overset{2}{C}H - \overset{3}{C}H = \underset{\downarrow 8}{\overset{4}{C}} - \overset{5}{C}H = \overset{6}{C}H \rightarrow 1 \right\rangle \overset{8}{C}H_2 - \overset{9}{C}H(\overset{2}{N}H_2) - \overset{11}{C}(\overset{12}{O}H) = \overset{13}{O}$
22	LValine	$\overset{8}{H_3}C - \overset{6}{C}H(\overset{7}{C}H_3) - \overset{4}{C}H(\overset{5}{N}H_2) - \overset{2}{C}(=\overset{3}{O}) - \overset{1}{O}H$
23	MSM	$\overset{5}{H_3}C - \overset{1}{S}(=\overset{2}{O},=\overset{3}{O}) - \overset{4}{C}H_3$
24	Myoinositol	$\left\langle 6 \leftarrow \overset{1}{C}H - \overset{2}{C}H - \overset{3}{C}H - \overset{4}{C}H - \overset{5}{C}H - \overset{6}{C}H \rightarrow 1 \right\rangle$ with $\overset{1}{\blacktriangle}OH$, $\overset{2}{\blacktriangle}OH$, $\overset{3}{\triangle}OH$, $\overset{4}{\blacktriangle}OH$, $\overset{5}{\triangle}OH$, $\overset{6}{\blacktriangle}OH$
25	NAA	$\overset{11}{H_3}C - \overset{6}{C}(=\overset{7}{O}) - \overset{5}{N}H \blacktriangleright \overset{2}{C}H(\overset{3}{C}(=\overset{4}{O}) - \overset{1}{O}H) - \overset{8}{C}H_2 - \overset{9}{C}(\overset{10}{O}H) = \overset{12}{O}$
26	NAAG	$\overset{20}{H_3}C - \overset{14}{C}(=\overset{15}{O}) - \overset{13}{N}H \blacktriangleright \overset{12}{C}H(\overset{17}{C}H_2 - \overset{18}{C}(\overset{19}{O}H) = \overset{21}{O}) - \overset{6}{C}(=\overset{7}{O}) - \overset{5}{N}H \triangleright \overset{4}{C}H(\overset{2}{C}(=\overset{3}{O}) - \overset{1}{O}) - \overset{8}{C}H_2 - \overset{9}{C}H_2 - \overset{10}{C}(\overset{11}{O}H) = \overset{16}{O}$
27	Phosphocreatine	$\overset{13}{O} = \overset{11}{C}(\overset{12}{O}H) - \overset{10}{C}H_2 - \overset{7}{N}(\overset{8}{C}H_3) - \overset{6}{C}(=\overset{9}{N}H) - \overset{5}{N}H - \overset{1}{P}(\overset{2}{O}H,\overset{3}{O}H) = \overset{4}{O}$
28	Phosphorylcholine	$\overset{10}{H_3}C - \overset{7}{N_+}(\overset{11}{C}H_3,\overset{8}{C}H_3) - \overset{6}{C}H_2 - \overset{5}{C}H_2 - \overset{4}{O} - \overset{1}{P}(=\overset{9}{O},\overset{2}{O}H) - \overset{3}{O}H$
29	Pyruvic	$\overset{6}{H_3}C - \overset{4}{C}(=\overset{5}{O}) - \overset{2}{C}(=\overset{3}{O}) - \overset{1}{O}H$
30	Scylloinositol	$\left\langle 6 \leftarrow \overset{1}{C}H - \overset{2}{C}H - \overset{3}{C}H - \overset{4}{C}H - \overset{5}{C}H - \overset{6}{C}H \rightarrow 1 \right\rangle$ with $\overset{1}{\triangle}OH$, $\overset{2}{\blacktriangle}OH$, $\overset{3}{\triangle}OH$, $\overset{4}{\blacktriangle}OH$, $\overset{5}{\triangle}OH$, $\overset{6}{\blacktriangle}OH$
31	Succinic	$\overset{1}{H}O - \overset{2}{C}(=\overset{3}{O}) - \overset{4}{C}H_2 - \overset{5}{C}H_2 - \overset{6}{C}(=\overset{8}{O}) - \overset{7}{O}H$
32	Taurine	$\overset{7}{H_2}N - \overset{6}{C}H_2 - \overset{5}{C}H_2 - \overset{1}{S}(=\overset{3}{O},=\overset{4}{O}) - \overset{2}{O}H$

Table F.3: Annotated chemical structures. The blue numbers above the non-hydrogen atoms are used to identify the hydrogen atoms associated with the chemical shifts and coupling coupling constants of *Table F.4*. For example, taurine's δ_6 (Table F.4) refers to the nuclei chemical shifts of the hydrogen atoms attached to [carbon] atom 6, and $J_{6a,5a}$ is the nuclei coupling constant between one of the hydrogen atoms attached to [carbon] atom 6 and one of the hydrogen atoms attached [carbon] atom 5. $\langle\,\rangle$: Ring. =: Double Bond. ▶: Up Stereo Bond. ▷: Down Stereo Bond

	Metabolite	Shifts (p.p.m.)	Coupling Constants (Hz)	References
17	LPhenylalanine	$\delta_{7a} = 3.2734, \delta_{7b} = 3.1049, \delta_8 = 3.9753$	$J_{7a,7b} = -14.573, J_{7a,8} = 5.209, J_{7b,8} = 8.013$	[43]
18	LSerine	$\delta_{2a} = 3.9379, \delta_{2b} = 3.9764, \delta_3 = 3.8347$	$J_{2a,2b} = -12.254, J_{2a,3} = 5.979, J_{2b,3} = 3.561$	[43]
19	LThreonine	$\delta_2 = 4.2464, \delta_5 = 1.3158, \delta_3 = 3.5785$	$J_{2,5} = 6.35, J_{2,3} = 4.917$	[43]
20	LTryptophan	$\delta_{10a} = 3.4739, \delta_{10b} = 3.2892, \delta_{11} = 4.0468$	$J_{10a,10b} = -15.368, J_{10a,11} = 4.851, J_{10b,11} = 8.145$	[43]
21	LTyrosine	$\delta_{8a} = 3.1908, \delta_{8b} = 3.0370, \delta_9 = 3.9281$	$J_{8a,8b} = -14.726, J_{8a,9} = 5.147, J_{8b,9} = 7.877$	[43]
22	LValine	$\delta_8 = 0.9764, \delta_7 = 1.0271, \delta_6 = 2.2577, \delta_4 = 3.5953$	$J_{6,4} = 4.405, J_{6,7} = 6.971, J_{6,7} = 7.071$	[43]
23	MSM	$\delta_5 = \delta_4 = 3.15$	$J_5 = J_4 = 12.70$	[2, 70]
24	Myoinositol	$\delta_1 = 4.0538, \delta_2 = 3.5217, \delta_3 = 3.6144, \delta_4 = 3.269$	$J_{1,2} = 3.006, J_{2,3} = 9.997, J_{3,4} = 9.485, J_{4,5} = 9.482, J_{5,6} = 9.998, J_{6,1} = 2.889$	[43]
		$\delta_5 = 3.6144, \delta_6 = 3.5217$		
25	NAA	$\delta_{11} = 2.008, \delta_5 = 7.8205, \delta_4 = 4.3817, \delta_{8a} = 2.6727$	$J_{4,5} = 6.4, J_{4,8a} = 3.861, J_{4,8b} = 9.821, J_{8a,8b} = -15.592$	[43]
		$\delta_{8b} = 2.4863$		
26	NAAG	$\delta_{20} = 2.06, \delta_{13} = 8.26, \delta_{12} = 4.63, \delta_{17a} = 2.54$	$J_{20} = -13.55, J_{13,12} = 7.32, J_{17} = -15.97, J_{17a,12} = 9.62, J_{17b,12} = 4.38$	[65]
		$\delta_{17b} = 2.74, \delta_5 = 7.95, \delta_4 = 4.14, \delta_{8a} = 2.06$	$J_{4,5} = 7.46, J_{4,8a} = 4.61, J_{4,8b} = 8.42, J_8 = -14.28, J_9 = -15.28$	
		$\delta_{8b} = 1.90, \delta_{9a} = 2.20, \delta_{9b} = 2.21$	$J_{8a,9a} = 10.56, J_{8a,9b} = 6.09, J_{8b,9a} = 4.9, J_{8b,9b} = 11.11$	
27	PPCr	$\delta_{10} = 3.93, \delta_8 = 3.029, \delta_9 = 6.581, \delta_5 = 7.296$	$J_{10} = 18.00, J_8 = 13.71$	[2, 43]
28	PPCh	$\delta_{10} = \delta_{11} = \delta_8 = 3.208, \delta_6 = 3.641, \delta_5 = 4.2805$	$J_{6a,5a} = 2.284, J_{6a,5b} = 7.326, J_{6b,5a} = 7.321, J_{6b,5b} = 2.235$	[43]
29	Pyruvic	$\delta_6 = 2.359$	$J_6 = 13$	[2, 43]
30	Scylloinositol	$\delta_1 = \delta_2 = \delta_3 = \delta_4 = \delta_5 = \delta_6 = 3.34$		[43]
31	Succinic	$\delta_4 = \delta_5 = 2.392$	$J_4 = J_5 = 12.00, J_{4,5} = 6.00$	[2, 43]
32	Taurine	$\delta_6 = 3.2459, \delta_5 = 3.4206$	$J_{6a,5a} = 6.742, J_{6a,5b} = 6.403, J_{6b,5a} = 6.464, J_{6b,5b} = 6.792$	[43]

Table F.4: The chemical shifts, and coupling constants, per metabolite. The references [104], [116], and [115] each have an on-line database – Biological Magnetic Resonance Databank (www.bmrb.wisc.edu/), Human Metabolome Database (www.hmdb.ca/), and Institut *für* Organische Chemie (www-user.uni-bremen.de/ wie/) at Universität Bremen respectively.

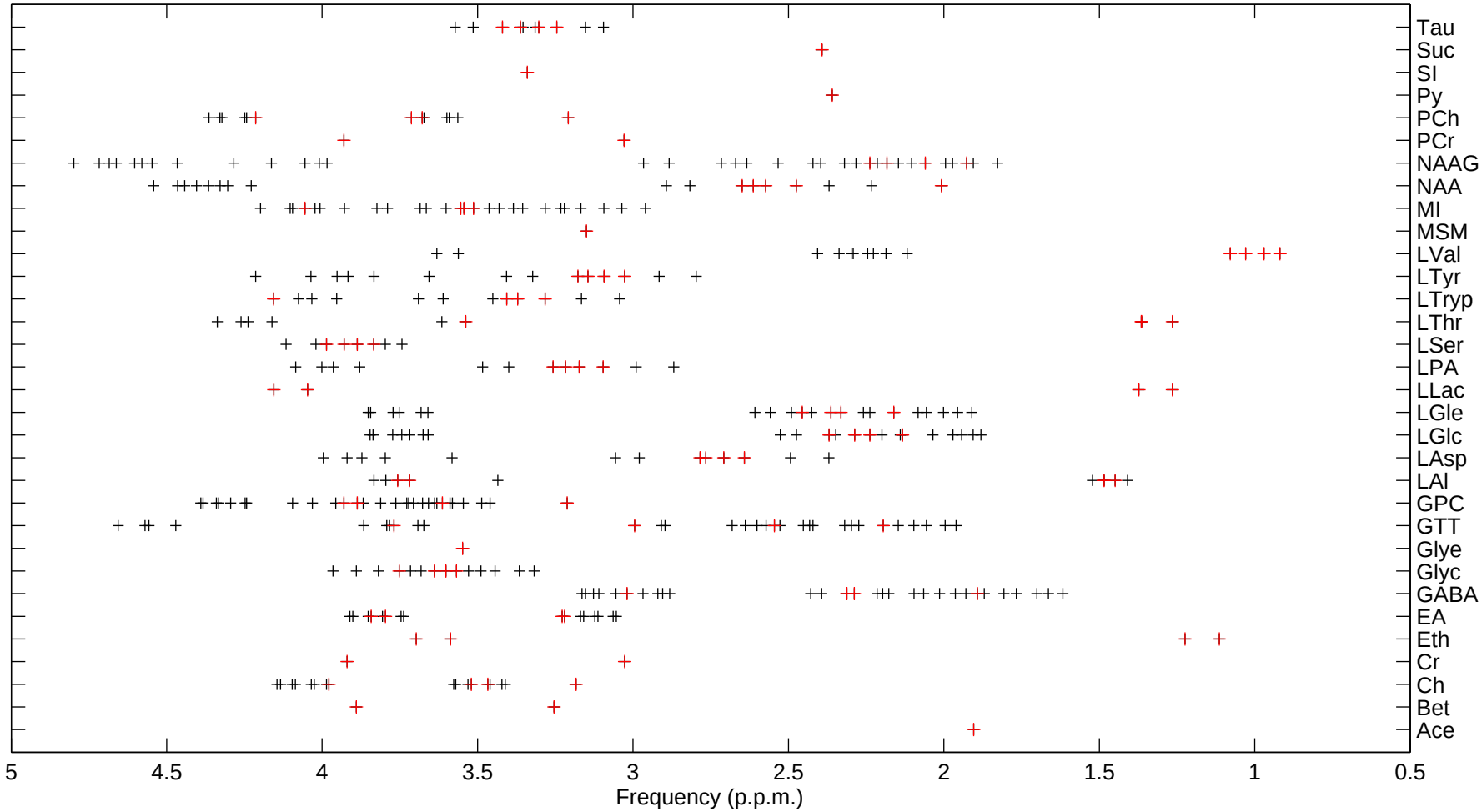


Figure F.1: Distribution of QMP spectral frequencies: case $T_E = 35$ ms. Refer to *Table D.2* for the simulation parameters for case $T_E = 35$ ms. Refer to the acronyms & abbreviations chapter for the full metabolite names. The four QMP spectral frequencies, or fewer if there are fewer QMP frequencies, with the highest QMP probabilistic intensities are indicated in red.

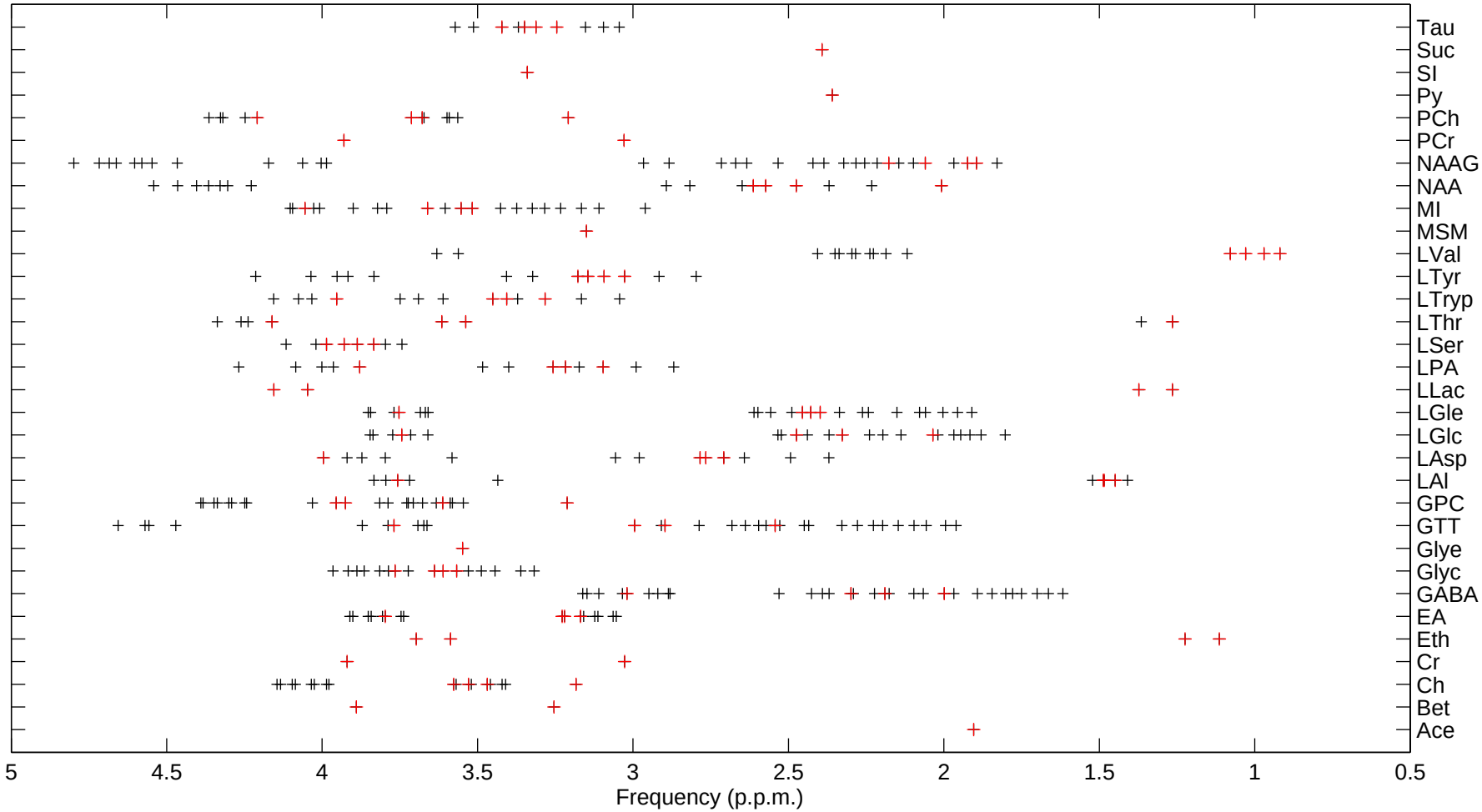


Figure F.2: Distribution of QMP spectral frequencies: case $T_E = 60$ ms. Refer to *Table D.2* for the simulation parameters for case $T_E = 60$ ms. Refer to the acronyms & abbreviations chapter for the full metabolite names. The four QMP spectral frequencies, or fewer if there are fewer QMP frequencies, with the highest QMP probabilistic intensities are indicated in red per metabolite.

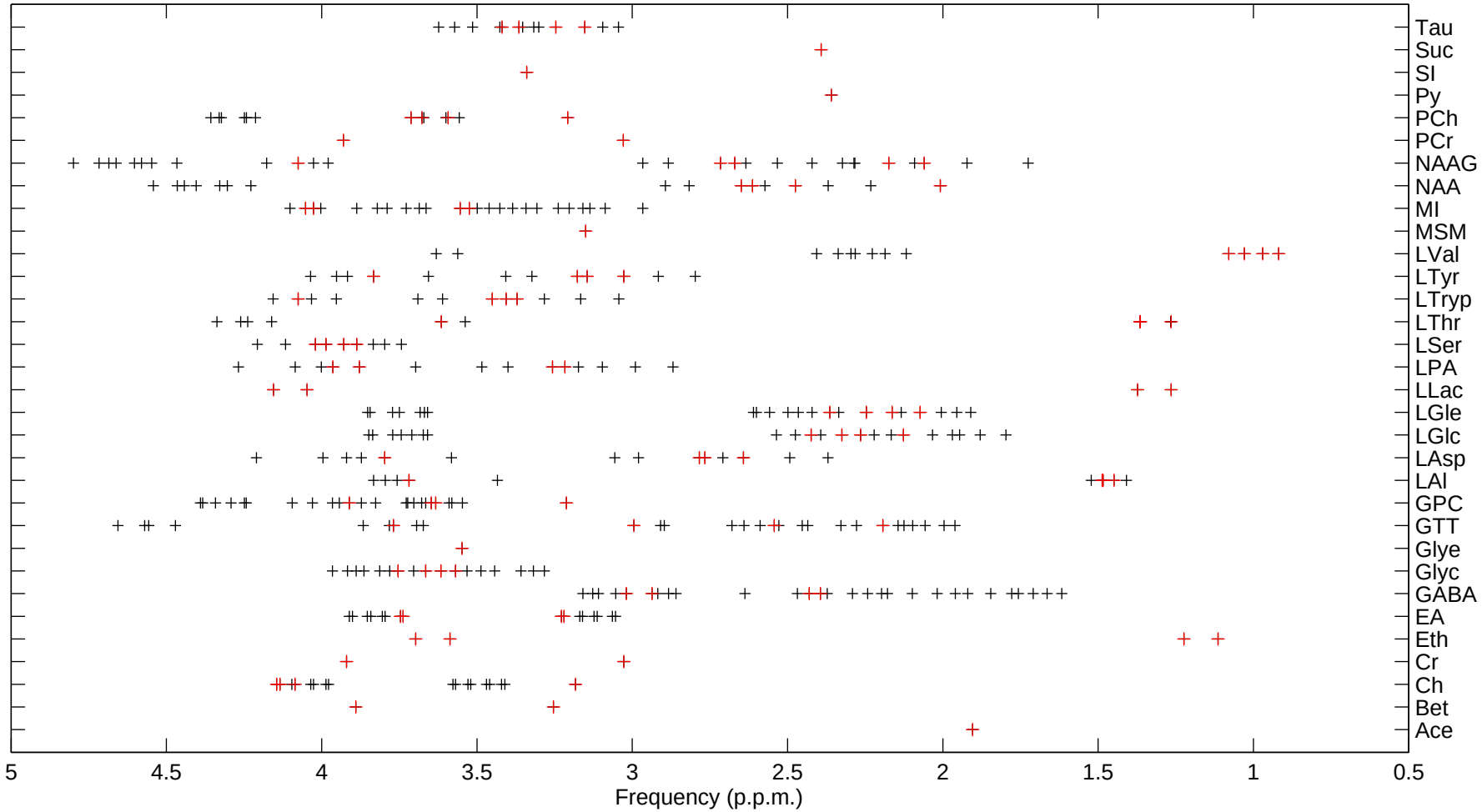


Figure F.3: Distribution of QMP spectral frequencies: case $T_E = 95$ ms. Refer to *Table D.2* for the simulation parameters for case $T_E = 95$ ms. Refer to the acronyms & abbreviations chapter for the full metabolite names. The four QMP spectral frequencies, or fewer if there are fewer QMP frequencies, with the highest QMP probabilistic intensities are indicated in red.

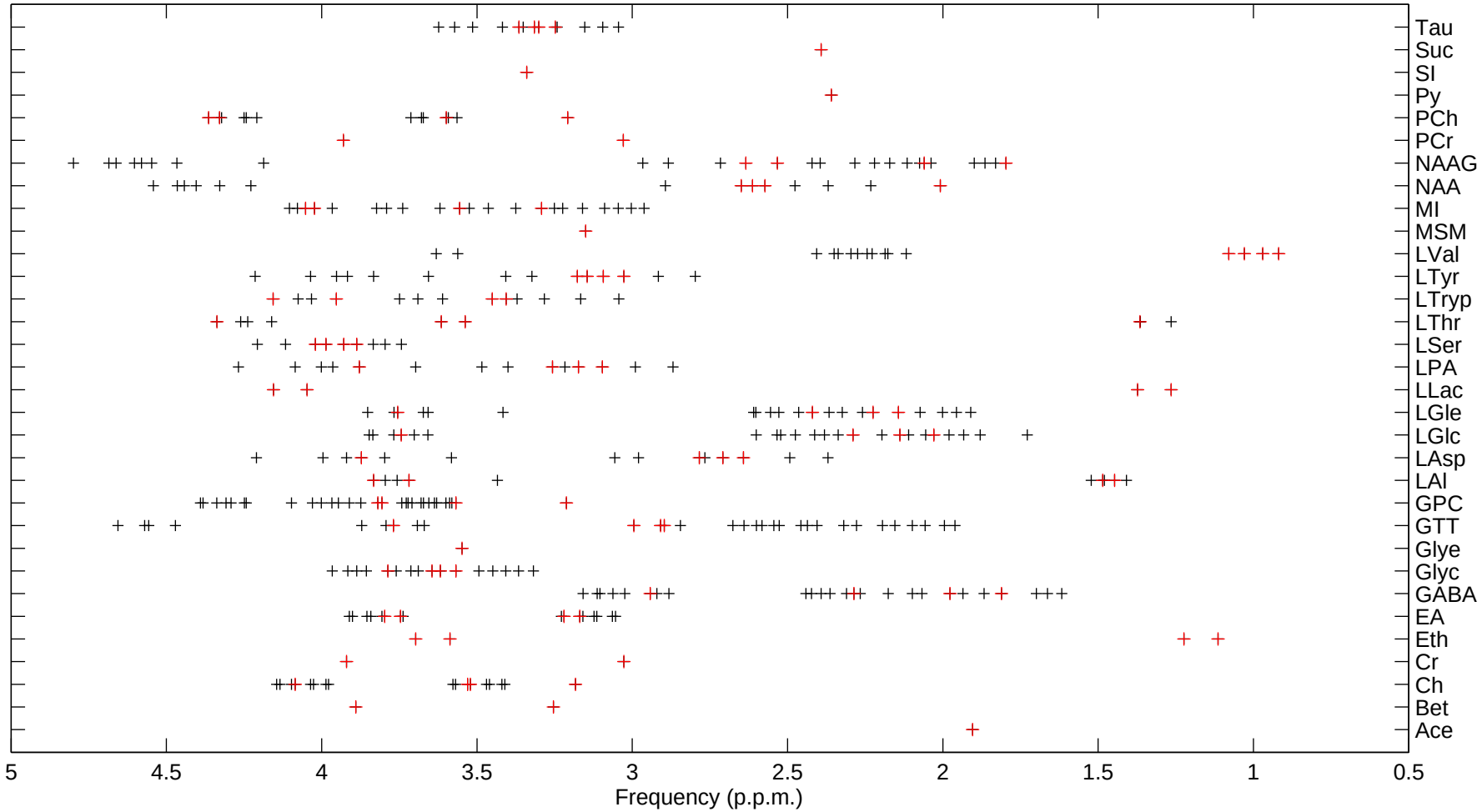


Figure F.4: Distribution of QMP spectral frequencies: case $T_E = 145$ ms. Refer to *Table D.2* for the simulation parameters for case $T_E = 145$ ms. Refer to the acronyms & abbreviations chapter for the full metabolite names. The four QMP spectral frequencies, or fewer if there are fewer QMP frequencies, with the highest QMP probabilistic intensities are indicated in red.

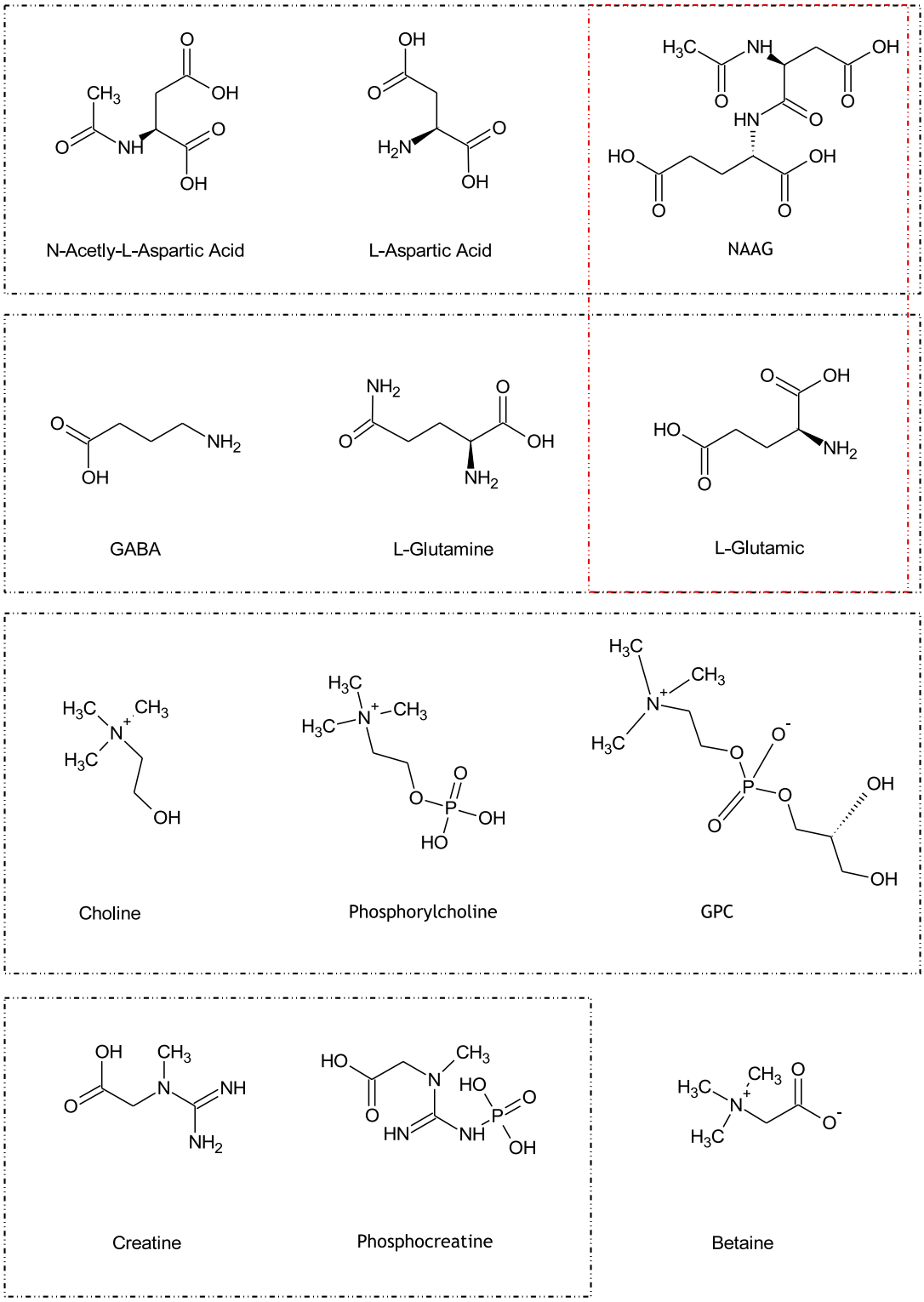


Figure F.5: These chemical structures, and those of *Fig. F.6* & *F.7*, are the 2-dimensional chemical structures of the metabolites, and other compounds, whose QMP data are used by the classification method. The metabolites whose structure and/or spectral frequencies overlap are grouped together. Some structures look similar but their spectra do not overlap at all, e.g. myoinositol and scylloinositol, because the orientations of their bonds in three-dimensional space are actually quite different. In such cases, the interaction between an external magnetic field and the physico-chemical constants of such metabolites can be quite different, and hence very different spectra. The spectral frequencies of **GABA and creatine** might sometimes overlap due to the partial structure overlap.

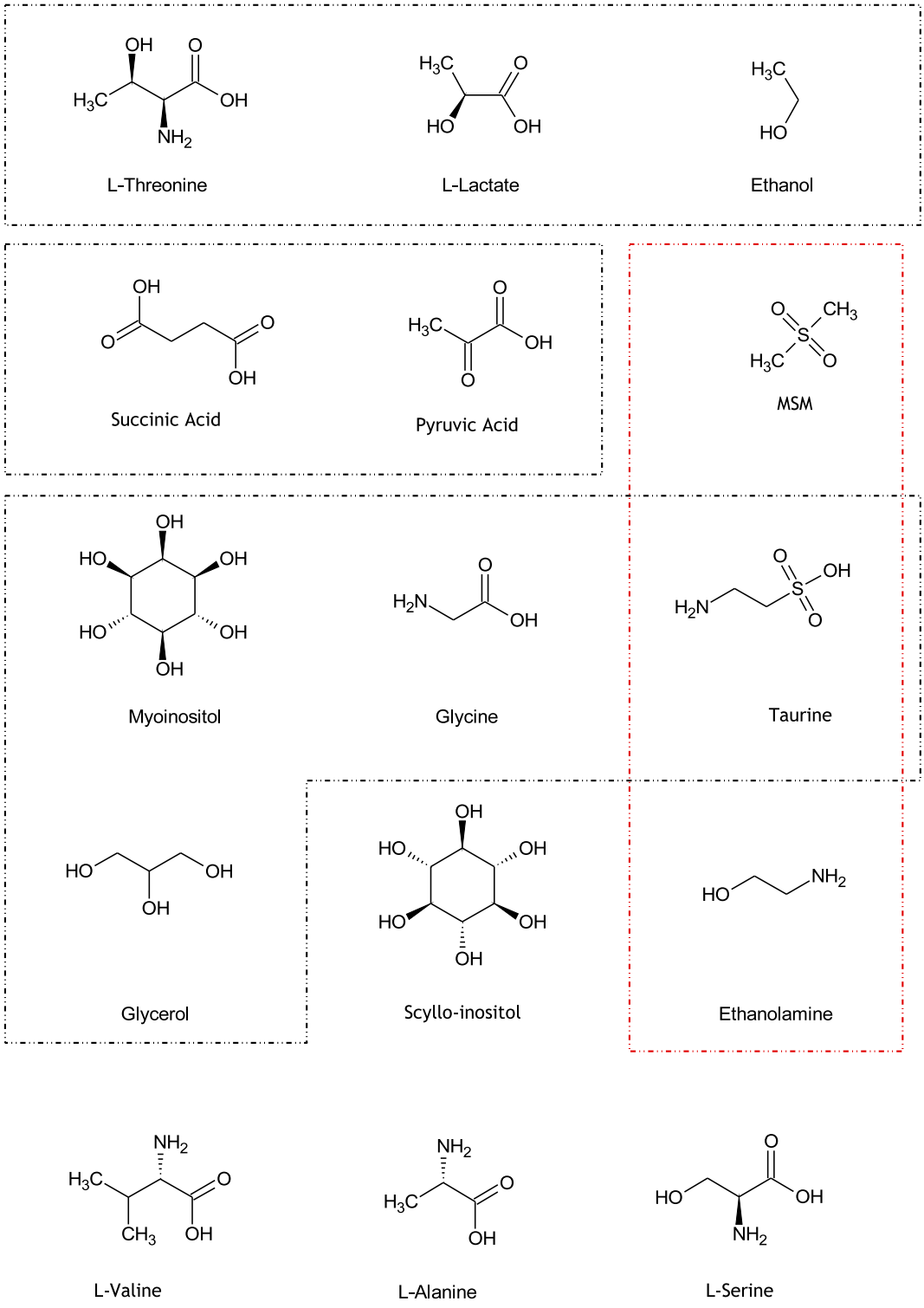


Figure F.6: Continued from *Fig. F.5*; the 2-dimensional chemical structures of the metabolites, and other compounds, whose QMP data are used by the classification method. The metabolites whose structure and/or spectral frequencies overlap are grouped together. Some structures look similar but their spectra do not overlap at all, e.g. myoinositol and scylloinositol, because the orientations of their bonds in three-dimensional space are actually quite different. In such cases, the interaction between an external magnetic field and the physico-chemical constants of such metabolites can be quite different, and hence very different spectra.

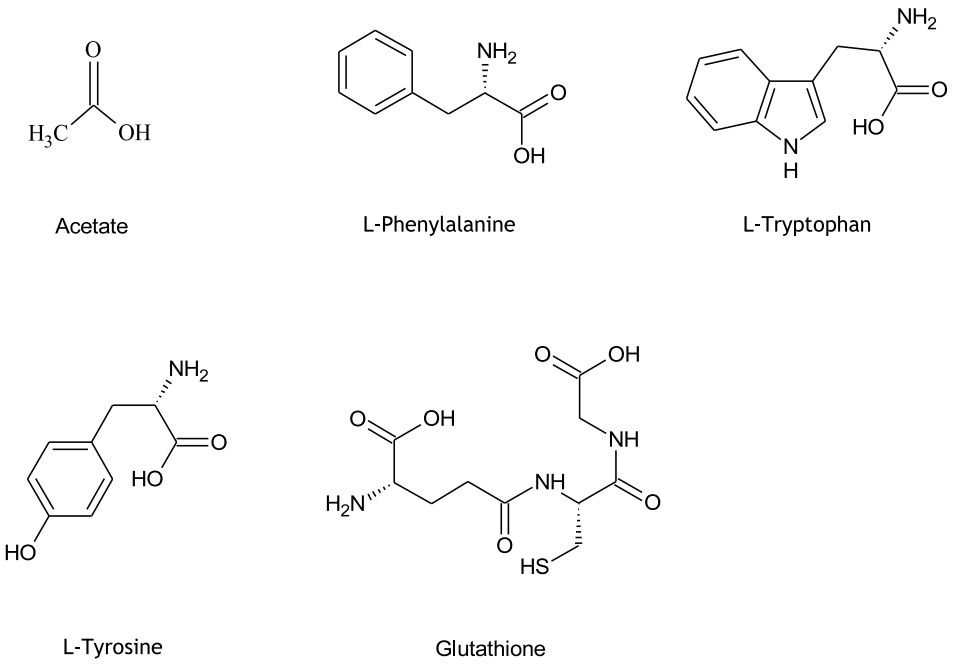


Figure F.7: Continued from *Fig. F.6*; the 2-dimensional chemical structures of the metabolites, and other compounds, whose QMP data are used by the classification method. The metabolites whose structure and/or spectral frequencies overlap are grouped together. Some structures look similar but their spectra do not overlap at all, e.g. myoinositol and scylloinositol, because the orientations of their bonds in three-dimensional space are actually quite different. In such cases, the interaction between an external magnetic field and the physico-chemical constants of such metabolites can be quite different, and hence very different spectra.

Bibliography

[1] Abraham, R. J., J. Fisher, and P. Loftus: 1988, *Introduction to NMR Spectroscopy.* John Wiley & Sons Ltd. Updated version of Proton and Carbon-13 NMR Spectroscopy by Abraham and Loftus.

[2] Advanced Chemistry Development, I.: 2007, 'ACD/Labs Release: 11.00 Product Version: 11.01 (Build 22585, 30 Oct 2007)'. Software Package.

[3] Allen, P. S. and R. B. Thompson: 1999, 'On the Localized Quantification of Metabolites with Coupled Spins'. *MAGMA* **9**(3), 159–163.

[4] Allouche, A. R., M. Aubert-Fré, and D. Graveron-Demilly: 2007, 'Quantum Chemistry-Based Spin Hamiltonian Parameters of GABA for Quantitation in Magnetic Resonance Spectroscopy'. *Physical Chemistry Chemical Physics* **9**, 3098–3103.

[5] Barker, P. and D. D. M. Lin: 2006, 'In Vivo Proton MR Spectroscopy of the Human Brain'. *Progress in Nuclear Magnetic Resonance Spectroscopy* **49**(2), 99–128.

[6] Belkić, D.: 2004a, 'Error analysis through residual frequency spectra in the fast Pad transform (FPT)'. *Nuclear Instruments & Methods In Physics Research* **525**(1–2), 379–386.

[7] Belkić, D.: 2004b, 'Strikingly Stable Convergence of the Fast Padé Transform (FPT) for High-Resolution Parametric and Non-Parametric Signal Processing of Lorentzian and Non-Lorentzian Spectra'. *Nuclear Instruments & Methods In Physics Research* **525**(1–2), 366–371.

[8] Belkić, D.: 2005, *Quantum-Mechanical Signal Processing and Spectral Analysis.* Institute of Physics.

[9] Belkić, D.: 2006a, 'Exact Quantification of Time Signals in Padé Based Magnetic Resonance Spectroscopy'. *Physics in Medicine and Biology* **51**, 2633–2670.

[10] Belkić, D.: 2006b, 'Exponential Convergence Rate (the spectral convergence) of the Fast Padé Transform for Exact Quantification in Magnetic Resonance Spectroscopy'. *Physics in Medicine and Biology* **51**, 6483–6512.

[11] Belkić, D.: 2006c, 'Mathematical Optimization of In Vivo NMR Chemistry Through the fast Padé Transform: Potential Relevance for Early Breast Cancer Detection by Magnetic Resonance Spectroscopy'. *Journal of Mathematical Chemistry* **40**(1), 85–103.

[12] Belkić, D. and K. Belkić: 2005a, 'Fast Padé Transform for Optimal Quantification of Time Signals from Magnetic Resonance Spectroscopy'. *International Journal of Quantum Chemistry* **105**(5), 493–510.

[13] Belkić, D. and K. Belkić: 2005b, 'The Fast Padé Transform in Magnetic Resonance Spectroscopy for Potential Improvements in Early Cancer Diagnostics'. *Physics in Medicine and Biology* **50**, 4385–4408.

[14] Belkić, D. and K. Belkić: 2006, 'In Vivo Magnetic Resonance Spectroscopy by the Fast Padé Transform'. *Physics in Medicine and Biology* **51**, 1049–1075.

[15] Belkić, D. and K. Belkić: 2008a, 'Mathematical Modeling of an NMR Chemistry Problem in Ovarian Cancer Diagnostics'. *Journal of Mathematical Chemistry* **43**, 395–425.

[16] Belkić, D. and K. Belkić: 2008b, 'Unequivocal Disentangling Genuine from Spurious Information in Time Signals: Clinical Relevance in Cancer Diagnostics through Magnetic Resonance Spectroscopy'. *Journal of Mathematical Chemistry* **44**, 884–912.

[17] Belkić, D. and K. Belkić: 2009, 'The General Concept of SignalNoise Separation (SNS): Mathematical Aspects and Implementation in Magnetic Resonance Spectroscopy'. *Journal of Mathematical Chemistry* **45**(3), 563–597.

[18] Belkić, D. and P. A. Dando: 2000, 'Three Novel High-Resolution Nonlinear Methods for Fast Signal Processing'. *The Journal of Chemical Physics* **113**(16), 6542–6556.

[19] Belkić, D., P. A. Dando, J. Main, H. S. Taylor, and S. K. Shin: 2000, 'Decimated Signal Diagonolization for Fourier Transform Spectroscopy'. *Journal of Physical Chemistry A* **104**(50), 11677–11684.

[20] Bishop, C. M.: 2006, *Pattern Recognition And Machine Learning*. Springer.

[21] Bottomley, P. A.: 1984, 'Selective Volume Method for Performing Localized NMR Spectroscopy'. *US Patent* **4480228**.

[22] Bottomley, P. A.: 1987, 'Spatial Localization in NMR Spectroscopy in Vivo'. *Annals of The New York Academy of Sciences* **508**, 333–348.

[23] Brannen, C.: 2007, *Operator Guide to the Standard Model*. Online, Liquafaction Corporation, Redmond, Washington, USA. 8 January 2007 version.

[24] Cavanagh, J., W. J. Fairbrother, A. G. Palmer III, and N. J. Skelton: 1996, *Protein NMR Spectroscopy, Principles and Practice*. Academic Press, Inc.

[25] Cerdan, S., C. A. Hansen, R. Johanson, T. Inubushi, and J. R. Williamson: 1986, 'Nuclear Magnetic Resonance Spectroscopic Analysis of Myo-Inositol Phosphates Including Inositol 1,3,4,5-Tetrakisphosphate'. *Journal of Biological Chemistry* **261**(31), 14676–14680.

[26] Claridge, T. D. W.: 1999, *High-Resolution NMR Techniques in Organic Chemistry*, Vol. 19 of *Tetrahedron Organic Chemistry Series*. Pergamon, Elsevier Science Ltd., 1st edition.

[27] Cobas, J. C. and F. J. Sardina: 2003, 'Nuclear Magnetic Resonance Data Processing. MestRe C: A Software Package for Desktop Computers'. *Concepts in Magnetic Resonance Part A* **19A**(2), 80–96. Includes NMRS background discussion.

[28] Corbett, R. J. T., A. R. Laptook, G. Tollefsbol, and B. Kim: 1995, 'Validation of a Noninvasive Method to Measure Brain Temperature In Vivo Using ^{1}H NMR Spectroscopy'. *Journal of Neurochemistry* **64**(3), 1224–1230.

[29] Cudalbu, C., S. Cavassila, H. Rabeson, D. van Ormondt, and D. Graveron-Demilly: 2008, 'Influence of Measured and Simulated Basis Sets on Metabolite Concentration Estimates'. *NMR in Biomedicine* **21**, 627–636.

[30] de Certaines, J. D., W. M. M. J. Bove, and F. Podo (eds.): 1992, *Magnetic Resonance Spectroscopy in Biology and Medicine: Functional and Pathological Tissue Characterization*. Pergamon Press.

[31] de Graaf, A. A., J. E. van Dijk, and B. W. M.: 1990, 'QUALITY: Quantification Improvement by Converting Lineshapes to the Lorentzian Type'. *Magnetic Resonance in Medicine* **13**(3), 343–357.

[32] de Graaf, R. A. and D. L. Rothman: 2001, 'In Vivo Detection and Quantification of Scalar Coupled ^{1}H NMR Resonances'. *Concepts in Magnetic Resonance* **13**(1), 32–76.

[33] Dennis, J. J. E., D. M. Gay, and R. E. Welsch: 1978, 'An Adaptive Nonlinear Least Squares Algorithm'. Technical Report 142, MIT.

[34] Dennis, J. J. E., D. M. Gay, and R. E. Welsch: 1981a, 'An Adaptive Nonlinear Least Squares Algorithm'. *ACM Transactions on Mathematical Software* **7**(3), 348–368.

[35] Dennis, J. J. E., D. M. Gay, and R. E. Welsch: 1981b, 'Algorithm 573 NL2SOL - An Adaptive Nonlinear Least-Squares Algorithm [E4]'. *ACM Transactions on Mathematical Software* **7**(3), 369–383.

[36] Dowsett, D. J., P. A. Kenny, and R. E. Johnston: 1998, *The Physics of Diagnostic Imaging*. Great Britain: Chapman & Hall Medical, 1st edition.

[37] Farr, R. F. and P. J. Allisy-Roberts: 1997, *Physics for Medical Imaging*. Great Britain: W. B. Saunders Company.

[38] Frahm, J., K. D. Merboldt, and W. Hänicke: 1987, 'Localized Proton Spectroscopy using Stimulated Echoes'. *Journal of Magnetic Resonance* **72**, 502–508. cf. Granot 1986.

[39] Gartland, K. P. R., C. R. Beddell, J. C. Lindon, and J. K. Nicholson: 1991, 'Application of Pattern Recognition Methods to the Analysis and Classification of Toxicological Data Derived from Proton Nuclear Magnetic Resonance Spectroscopy'. *Molecular Pharmacology* **39**, 629–642.

[40] Gershenfeld: 1999, *The Nature of Mathematical Modelling*. Cambridge University Press.

[41] Golub, G. H. and V. Pereyra: 1973, 'The Differentiation of Pseudo-Inverses and Nonlinear Least Squares Problems Whose Variables Separate'. *SIAM Journal on Numerical Analysis* **10**(2), 413–432.

[42] Golub, G. H. and C. Reinsch: 1970, 'Singular Value Decompostion and Least Squares Solutions'. *Numerische Mathematik* **14**(5), 403–420.

[43] Govindaraju, V., K. Young, and A. A. Maudsley: 2000, 'Proton NMR Chemical Shifts and Coupling Constants for Brain Metabolites'. *NMR in Biomedicine* **13**(3), 129–153.

[44] Granot, J.: 1986, 'Selected Volume Excitation using Stimulated Echoes (VEST): Applications to Spatially Localized Spectroscopy Imaging'. *Journal of Magnetic Resonance* **70**, 488–492. cf. Frahm 1987.

[45] Graveron-Demilly, D., A. Diop, A. Briguet, and B. Fenet: 1993, 'Product-Operator Algebra for Strongly Coupled Spin Systems'. *Journal of Magnetic Resonance, Series A* **101**(3), 233–239.

[46] Haacke, E. M., Z. P. Liang, and S. H. Izen: 1989, 'Constrained Reconstruction: a Superresolution, Optimal Signal-to-Noise Alternative to the Fourier Transform in Magnetic Resonance Imaging'. *Medical Physics* **16**(3), 388–397.

[47] Hahn, E. L.: 1950, 'Spin Echoes'. *Physical Review* **80**(4), 580–600.

[48] Higinbotham, J. and I. Marshall: 2001, *NMR Lineshapes and Lineshape Fitting Procedures*, Vol. 43 of *Annual Reports on NMR Spectroscopy*, Chapt. 2, pp. 60–116. Academic Press. Editor: Webb, G. A.

[49] Hore, P. J., J. A. Jones, and S. Wimperis: 2000, *NMR: The Toolkit*, No. 92 in Oxford Chemistry Primers. Oxford University Press.

[50] Hornak, J. P.: 2002, *The Basics of NMR*. Online, Magnetic Resonance Laboratory, Center for Imaging Science, Rochester Institute of Technology, 54 Lomb Memorial Drive, Rochester, NY 14623-5604.

[51] Hu, X., D. N. Levin, P. C. Lauterbur, and T. Spraggins: 1988, 'SLIM: Spectral Localization by Imaging'. *Magnetic Resonance in Medicine* **8**(3), 314–322.

[52] Huang, N. E., Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. Yen, C. C. Tung, and H. H. Liu: 1998, 'The Empirical Mode Decomposition and The Hilbert Spectrum for Nonlinear and Non-Stationary Time Series Analysis'. *Proceedings of the Royal Society of London. Series A* **454**(1971), 903–995.

[53] Jansen, J. F. A., W. H. Backes, K. Nicolay, and K. M. E.: 2006, '^1H MR Spectroscopy of the Brain: Absolute Quantification of Metabolites'. *Radiology* **240**(2), 2.

[54] Karplus, M.: 1959a, 'Contact Electron-Spin Coupling of Nuclear Magnetic Moments'. *The Journal of Chemical Physics* **30**(1), 11–15.

[55] Karplus, M.: 1959b, 'Interpretation of the Electron-Spin Resonance Spectrum of the Methyl Radical'. *The Journal of Chemical Physics* **30**(1), 15–18.

[56] Karplus, M.: 1963, 'Vicinal Proton Coupling in Nuclear Magnetic Resonance'. *Journal of the American Chemical Society* **85**(18), 2870–2871.

[57] Karplus, M. and D. H. Anderson: 1959, 'Valence-Bond Interpretation of Electron-Coupled Nuclear Spin Interactions; Application to Methane'. *The Journal of Chemical Physics* **30**(1), 6–10.

[58] Kaufman, L.: 1975, 'A Variable Projection Method for Solving Separable Nonlinear Least Squares Problems'. *BIT* **15**(1), 49–57.

[59] Kay, S. M.: 2003, *Fundamentals of Statistical Signal Processing Volum II: Detection Theory*, Vol. II of *Signal Processing*. Prentice Hall, korea edition.

[60] Keeler, J.: 2005, *Understanding NMR spectroscopy*. John Wiley & Sons Ltd.

[61] Keevil, S. F.: 2006, 'Spatial Localization in Nuclear Magnetic Resonance Spectroscopy'. *Physics in Medicine and Biology* **51**(16), R579–R636.

[62] Kelm, B. M., B. H. Menze, M. Zechmann, K. T. Baudendistel, and F. A. Hamprecht: 2007, 'Automated Estimation of Tumor Probability in Prostate Magnetic Resonance Spectroscopic Imaging: Pattern Recognition vs Quantification'. *Magnetic Resonance in Medicine* **57**, 150–159.

[63] Khalidov, I., D. van de Ville, M. Jacob, L. Lazeyras, and M. Unser: 2006, 'Improved MRSI With Field Inhomogeneity Compensation'. In: *Proceedings of SPIE*, Vol. 6144. Article Number 614467.

[64] Kranendonk, L. A., A. W. Caswell, and S. T. Sanders: 2007, 'Robust Method for Calculating Temperature, Pressure, and Absorber Mole Fraction from Broadband Spectra'. *Applied Optics* **46**(19), 4117–4124.

[65] Krawczyk, H. and W. Gradowska: 2003, 'Characterisation of the ^{1}H and ^{13}C NMR Spectra of N-acetylaspartylglutamate and its Detection in Urine from Patients with Canavan Disease'. *Journal of Pharmaceutical and Biomedical Analysis* **31**, 455–463.

[66] Kreis, R.: 2004, 'Issues of Spectral Quality in Clinical ^{1}H-Magnetic Resonance Spectroscopy and a Gallery of Artifacts'. *NMR in Biomedicine* **17**(6), 361–381.

[67] Kreis, R., J. Slotboom, L. Hofmann, and C. Boesch: 2005, 'Integrated Data Acquisition and Processing to Determine Metabolite Contents, Relaxation Times, and Macromolecule Baseline in Single Examinations of Individual Subjects'. *Magn Reson Med* **54**(4), 761–768.

[68] Kreyszig, E.: 1999, *Advanced Engineering Mathematics*. John Wiley & Sons, Inc., 8^{th} edition.

[69] Laudadio, T., N. Mastronardi, L. Vanhamme, P. van Hecke, and van Huffel S.: 2002, 'Improved Lanczos Algorithms for Blackbox MRS Data Quantitation'. *Journal of Magnetic Resonance* **157**(2), 292–297.

[70] Lin, A., C. H. Nguy, F. Shic, and B. D. Ross: 2001, 'Accumulation of Methylsulfonylmethane in the Human Brain: Identification by Multinuclear Magnetic Resonance spectroscopy'. *Toxicology Letters* **123**, 169–177.

[71] Lin, A., B. D. Ross, K. Harris, and W. Wong: 2005, 'Efficacy of Proton Magnetic Resonance Spectroscopy in Neurological Diagnosis and Neurotherapeutic Decision Making'. *The Journal of the American Society for Experimental NeuroTherapeutics* **2**(2), 197–214.

[72] Marquardt, D. W.: 1963, 'An Algorithm for Least-Squares Estimation of Nonlinear Parameters'. *Journal of the Society for Industrial and Applied Mathematics* **11**(2), 431–441.

[73] Menze, B. H., M. Kelm, M. Weber, P. Bachert, and F. A. Hamprecht: 2008, 'Mimicking the Human Expert: Pattern Recognition for an Automated Assessment of Data Quality in MR Spectroscopic Images'. *Magnetic Resonance in Medicine* **59**, 1457–1466.

[74] Naressi, A., C. Couturier, I. Castang, R. de Beer, and D. Graveron-Demilly: 2001a, 'Java-Based Graphical User Interface for MRUI, a Software Package for Quantitation of In Vivo/Medical Magnetic Resonance Spectroscopy Signals.'. *Computers in Biology and Medicine* **31**(4), 269–286.

[75] Naressi, A., C. Couturier, J. M. Devos, M. Janssen, C. Mangeat, R. de Beer, and D. Graveron-Demilly: 2001b, 'Java-Based Graphical User Interface for the MRUI Quantitation Package.'. *MAGMA* **12**(2), 141–152.

[76] Nazareth, L.: 1980, 'Some Recent Approaches to Solving Large Residual Nonlinear Least Squares Problems'. *SIAM Review* **22**(1), 1–11.

[77] Olver, P. J. and C. Shakiban: 2006, *Applied Linear Algebra*. Pearson Prentice Hall.

[78] Ordidge, R. J., A. Connelly, and J. A. B. Lohman: 1986, 'Image-Selected *In Vivo* Spectroscopy: A New Technique for Spatially Selective NMR Spectroscopy'. *Journal of Magnetic Resonance* **66**, 283–294.

[79] Osborne, M. R.: 1972, *Numerical Methods for Nonlinear Optimisation*, Chapt. Some Aspects of Nonlinear Least Squares Calculations, pp. 171–189. Academic Press.

[80] Pijnappel, W. W. F., A. van den Boogaart, R. de Beer, and D. van Ormondt: 1992, 'SVD-Based Quantification of Magnetic Resonance Signals'. *Journal of Magnetic Resonance* **97**(1), 122–134.

[81] Poullet, J., D. M. Sima, A. J. Simonetti, B. de Neuter, L. Vanhamme, P. Lemmerling, and S. van Huffel: 2006, 'An Automated Quantitation of Short Echo Time MRS Spectra in an Open Source Software Environment: AQSES'. *NMR in Biomedicine* **20**(5), 493–504. Online, December 2006.

[82] Poullet, J., D. M. Sima, and S. van Huffel: 2008, 'MRS Signal Quantitation: A Review of Time- and Frequency-Domain Methods'. *Journal of Magnetic Resonance* **195**, 134–144.

[83] Proakis, J. G. and D. G. Manolakis: 1988, *Introduction to Digital Signal Processing*. Maxwell Macmillan.

[84] Provencher, S.: 1993, 'Estimation of Metabolite Concentrations from Localized in Vivo Proton NMR Spectra'. *Magnetic Resonance in Medicine* **30**(6), 672–679.

[85] Provencher, S. W.: 1982a, 'A Constrained Regularization Method for Inverting Data Represented by Linear Algebraic or Intergral Equations'. *Computer Physics Communications* **27**, 213–227.

[86] Provencher, S. W.: 1982b, 'CONTIN: A General Purpose Constrained Regularization Program for Inverting Noisy Linear Algebraic and Intergral Equations'. *Computer Physics Communications* **27**, 229–242.

[87] Provencher, S. W.: 2001, 'Automatic Quantitation of Localized *in vivo* ^{1}H Spectra with LCModel'. *NMR in Biomedicine* **14**, 260–264.

[88] Ratiney, H., Y. Coenradie, S. Cavassila, D. van Ormondt, and D. Graveron-Demilly: 2004, 'Time-Domain Quantitation of ^{1}H Short Echo-Time Signals: Background Accommodation'. *MAGMA* **16**(6), 284–296.

[89] Ratiney, H., M. Sdika, Y. Coenradie, D. van Ormondt, and D. Graveron-Demilly: 2005, 'Time Domain Semi-Parametric Estimation Based On a Metabolite Basis Set'. *NMR in Biomedicine* **18**(1), 1–13.

[90] Schölkopf, B. and A. J. Smola: 2002, *Learning With Kernels*. The MIT Press.

[91] Seeger, U., U. Klose, I. Mader, W. Grodd, and T. Nägele: 2003, 'Parameterized Evaluation of Macromolecules and Lipids in Proton MR Spectroscopy of Brain Diseases'. *Magnetic Resonance in Medicine* **49**(1), 19–28.

[92] Seeger, U., I. Mader, T. Nägele, W. Grodd, O. Lutz, and U. Klose: 2001, 'Reliable Detection of Macromolecules in Single-Volume ^{1}H NMR Spectra of the Human Brain'. *Magnetic Resonance in Medicine* **45**(6), 948–954.

[93] Serrai, H., L. Nadal-Desbarats, H. Poptani, J. D. Glickson, and L. Senhadji: 2000, 'Lactate Editing and Lipid Suppression by Continuous Wavelet Transform Analysis: Application to Simulated and ^{1}H MRS Brain Tumor Time-Domain Data'. *Magnetic Resonance in Medicine* **43**(5), 649–656.

[94] Serrai, H., L. Senhadji, G. Wang, S. Akoka, and P. Stroman: 2003, 'Lactate Doublet Quantification and Lipid Signal Suppression Using a New Biexponential Decay Filter: Application to Simulated and ^{1}H MRS Brain Tumor Time-Domain Data'. *Magnetic Resonance in Medicine* **50**(3), 623–626.

[95] Sima, D. M. and S. van Huffel: 2006, 'Regularized Semiparametric Model Identification with Application to Nuclear Magnetic Resonance Signal Quantification with Unknown Macromolecular Base-line'. *Journal of the Royal Statistical Society, B* **68**(3), 383–409.

[96] Sima, D. M. and S. van Huffel: 2007, 'Separable Nonlinear Least Squares Fitting with Linear Bound Constraints and its Application in Magnetic Resonance Spectroscopy Data Quantification'. *Journal of Computational and Applied Mathematics* **203**(1), 264–278.

[97] Smith, S. A., T. O. Levante, B. H. Meier, and R. R. Ernst: 1994, 'Computer Simulations in Magnetic Resonance. An Object-Oriented Programming Approach'. *Journal of Magnetic Resonance, Series A* **106**(1), 75–105.

[98] Soher, B. J., K. Young, A. Bernstein, Z. Aygula, and A. A. Maudsley: 2007, 'GAVA; Spectral Simulation for in vivo MRS Applications'. *Journal of Magnetic Resonance* **185**, 291–299.

[99] Sørensen, O. W., G. W. Eich, M. H. Levitt, G. Bodenhausen, and R. R. Ernst: 1983, 'Product Operator Formulaism for the Description of NMR Pulse Experiments'. *Progress in Nuclear Magnetic Resonance Spectroscopy* **16**, 163–192.

[100] Sundin, T., L. Vanhamme, P. van Hecke, I. Dologlou, and S. van Huffel: 1999, 'Accurate quantification of ^{1}H Spectra: From Finite Impulse Response Filter Design for Solvent Suppression to Parameter Estimation'. *Journal of Magnetic Resonance* **139**(2), 189–204.

[101] Tate, A. R., C. Majós, A. Moreno, F. A. Howe, J. R. Griffiths, and C. Arús: 2003, 'Automated Classification of Short Echo Time in In Vivo ^{1}H Brain Tumor Spectra: A Multicenter Study'. *Magnetic Resonance in Medicine* **49**, 29–36.

[102] Thompson, R. B. and P. S. Allen: 1999, 'Sources of Variability in the Response of Coupled Spins to the PRESS Sequence and their Potential Impact on Metabolite Quantification'. *Magnetic Resonance in Medicine* **41**(6), 1162–1169.

[103] Thompson, R. B. and P. S. Allen: 2001, 'Response of Metabolites with Coupled Spins to the STEAM Sequence'. *Magnetic Resonance in Medicine* **45**(6), 955–965.

[104] Ulrich, E. L., H. Akutsu, J. F. Doreleijers, Y. Harano, Y. E. Ioannidis, J. Lin, M. Livny, S. Mading, D. Maziuk, Z. Miller, E. Nakatani, C. F. Schulte, D. E. Tolmie, R. K. Wenger, H. Yao, and J. L. Markley: 2008, 'BioMagResBank'. *Nucleic Acids Research* **36**, D402–D408.

[105] van den Boogaart, A., P. van Hecke, S. van Huffel, D. Graveron-Demilly, D. van Ormondt, and R. de Beer: 1996, 'MRUI: A Graphical User Interface for Accurate Routine MRS Data Analysis'. In: *Proceedings of the European Society of Magnetic Resonance in Medicine and Biology*. p. 318.

[106] van der Veen, J. W. C., R. de Beer, P. R. Luyten, and D. van Ormondt: 1988, 'Accurate Quantification of in Vivo ^{31}P NMR Signals Using the Variable Projection Method and Prior Knowledge'. *Magnetic Resonance in Medicine* **6**(1), 92–98.

[107] van Huffel, S., H. Chen, C. Decanniere, and P. van Hecke: 1994, 'Algorithm for Time-Domain NMR Data Fitting Based on Total Least Squares'. *Journal of Magnetic Resonance, Series A* **110**(2), 228–237.

[108] van Huffel, S. and J. Vandewalle: 1991, *The Total Least Squares Problem : Computational Aspects and Analysis*. Society for Industrial and Applied Mathematics, Philadelphia.

[109] Vanhamme, L., T. Sundin, P. V. Hecke, and S. V. Huffel: 2001, 'MR Spectroscopy Quantitation: A Review of Time-Domain Methods'. *NMR in Biomedicine* **14**(4), 233–246.

[110] Vanhamme, L., A. van den Boogart, and S. van Huffel: 1997, 'Improved Method for Accurate and Efficient Quantification of MRS Data with Use of Prior Knowledge'. *Journal of Magnetic Resonance* **129**(1), 35–43.

[111] Wasserman, L.: 2004, *All of Statistics: A Concise Course in Statistical Inference*, Springer Texts in Statistics. Springer, 1st edition. Corrected 2nd printing.

[112] Weishaupt, D., V. D. Köchli, and B. Marincek: 2006, *How Does MRI Work? An Introduction to the Physics and Function of Magnetic Resonance Imaging*. Springer-Verlag Berlin Heidelberg, 2nd edition.

[113] Williamson, D. C., H. Hawesa, N. A. Thacker, and S. R. Williams: 2006, 'Robust Quantification of Short Echo Time ^{1}H Magnetic Resonance Spectra using the Padé Approximant'. *Magnetic Resonance in Medicine* **55**, 762–771.

[114] Williamson, D. C. and N. A. Thacker: 2003, 'Assessment of the Padé Approximant as a Method for Quantifying ^1H Magnetic Resonance Spectroscopic Data'. *Journal of Computer Methods in Science and Engineering* **3**.

[115] Willker, W., J. Engelmann, A. Brand, and D. Leibfritz: 1996, 'Metabolite Identification in cell Extracts and Culture Media by Proton-Detected 2D-H, C-NMR Spectroscopy'. *Journal of Magnetic Resonance Analysis* **2**, 21–32.

[116] Wishart, D. S., D. Tzur, C. Knox, R. Eisner, A. C. Guo, N. Young, D. Cheng, K. Jewell, D. Arndt, S. Sawhney, C. Fung, L. Nikolai, M. Lewis, M. Coutouly, I. Forsythe, P. Tang, S. Shrivastava, K. Jeroncic, P. Stothard, G. Amegbey, D. Block, D. D. Hau, J. Wagner, J. Miniaci, M. Clements, M. Gebremedhin, N. Guo, Y. Zhang, G. E. Duggan, G. D. MacInnis, A. M. Weljie, R. Dowlatabadi, F. Bamforth, D. Clive, R. Greiner, L. Li, T. Marrie, B. D. Sykes, H. J. Vogel, and L. Querengesser: 2007, 'HMDB: the Human Metabolome Database'. *Nucleic Acids Research* **35**, D521–D526.

[117] Witteveen, J.: 1997, 'T_2 Relaxation due to Two-Level Field Fluctuations'. *Physical Review B* **55**(13), 8083–8085.