



HAL
open science

A quantum mechanics-based approach for optimization of metabolite basis-sets : application to quantitation of HRMAS-NMR signals

Andrii Lazariev

► **To cite this version:**

Andrii Lazariev. A quantum mechanics-based approach for optimization of metabolite basis-sets : application to quantitation of HRMAS-NMR signals. Human health and pathology. Université Claude Bernard - Lyon I, 2011. English. NNT : 2011LYO10113 . tel-00843311

HAL Id: tel-00843311

<https://theses.hal.science/tel-00843311>

Submitted on 11 Jul 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE
présentée devant
l'UNIVERSITE CLAUDE BERNARD - LYON I
pour l'obtention
du DIPLOME DE DOCTORAT
(arrêté du 7 août 2006)

par

LAZARIEV Andrii

**A Quantum Mechanics-based approach for optimization of
metabolite basis-sets. Application to quantitation of
HRMAS-NMR signals**

Soutenue le 27 juin 2011

Spécialité: **Images et Systèmes**

Directeur de thèse: **Dr. Danielle GRAVERON-DEMILLY**

JURY: **Pr Abdul-Rahman ALLOUCHE**

Pr André BRIGUET

Pr Emmanuelle CANET-SOULAS

Dr Florence FAUVELLE

Dr Danielle GRAVERON-DEMILLY Directeur

Dr Dirk VAN ORMONDT Rapporteur

Dr Chantal REMY Rapporteur

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université

M. A. Bonmartin

Vice-président du Conseil d'Administration

M. le Professeur G. Annat

Vice-président du Conseil des Etudes et de la Vie Universitaire

M. le Professeur D. Simon

Vice-président du Conseil Scientifique

M. le Professeur J-F. Mornex

Secrétaire Général

M. G. Gay

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard

Directeur : M. le Professeur J. Etienne

Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux

Directeur : M. le Professeur F-N. Gilly

UFR d'Odontologie

Directeur : M. le Professeur D. Bourgeois

Institut des Sciences Pharmaceutiques et Biologiques

Directeur : M. le Professeur F. Locher

Institut des Sciences et Techniques de la Réadaptation

Directeur : M. le Professeur Y. Matillon

Département de formation et Centre de Recherche en Biologie Humaine

Directeur : M. le Professeur P. Farge

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies

Directeur : M. le Professeur F. Gieres

Département Biologie

Directeur : M. le Professeur F. Fleury

Département Chimie Biochimie

Directeur : M. le Professeur H. Parrot

Département GEP

Directeur : M. N. Siauve

Département Informatique

Directeur : M. le Professeur S. Akkouche

Département Mathématiques

Directeur : M. le Professeur A. Goldman

Département Mécanique

Directeur : M. le Professeur H. Ben Hadid

Département Physique

Directeur : Mme S. Fleck

Département Sciences de la Terre

Directeur : Mme le Professeur I. Daniel

UFR Sciences et Techniques des Activités Physiques et Sportives

Directeur : M. C. Collignon

Observatoire de Lyon

Directeur : M. B. Guiderdoni

Ecole Polytechnique Universitaire de Lyon 1

Directeur : M. P. Fournier

Ecole Supérieure de Chimie Physique Electronique

Directeur : M. G. Pignault

Institut Universitaire de Technologie de Lyon 1

Directeur : M. le Professeur C. Coulet

Institut de Science Financière et d'Assurances

Directeur : M. le Professeur J-C. Augros

Institut Universitaire de Formation des Maîtres

Directeur : M. R. Bernard

Remerciements

J'ai passé trois années agréables en effectuant les travaux de cette thèse. Chaque jour, j'étais entouré par l'ambiance amicale du laboratoire Creatis et aujourd'hui je voudrais remercier tous les membres.

Tout d'abord, j'imagine que ces lignes n'auraient jamais été écrites si je n'avais été accueilli par Madame Danielle Graveron-Demilly. Il est difficile de trouver les mots pour lui exprimer ma reconnaissance pour les heures, les jours et les mois de travail qui m'ont ouvert la porte de l'avenir. Je voudrais aussi souligner les efforts qu'elle a fournis pour me faire entrer dans le monde scientifique où j'ai fait connaissance avec de nombreux chercheurs du domaine.

Ainsi, je dois mentionner les discussions fructueuses avec Monsieur André Briguet concernant le texte de ma thèse. Je remercie Madame Chantal Rémy d'avoir été rapporteur et membre du jury, comme je remercie également Monsieur Dirk van Ormondt. Je lui serai toujours reconnaissant pour sa relecture du manuscrit de ma thèse ainsi que les divers échanges que nous avons pu avoir sur l'histoire et la culture européennes à travers nos conversations fréquentes sur ces sujets.

L'aide de Madame Florence Fauvelle ne sera jamais oubliée comme je n'oublierai pas celle de Monsieur Martial Piotto. Grâce à ces collaborations, j'ai obtenu les spectres expérimentaux indispensables à l'aboutissement de ma thèse. Je remercie Madame Emmanuelle Cannet-Soulas et Monsieur Abdul-Rahman Allouche d'avoir été membres de jury lors de la soutenance.

Je remercie tous mes collègues et amis qui m'ont aidé à améliorer la présentation de la thèse, tous ceux qui sont venus partager mes émotions pendant la soutenance et tous ceux qui m'entouraient et me soutenaient pendant ces années inoubliables.

Contents

| | |
|---|-----------|
| Summary | 1 |
| Notations | 3 |
| Introduction | 5 |
| Introduction | 5 |
| Préambule | 9 |
| 1 Nuclear Magnetic Resonance Spectroscopy | 13 |
| 1.1 Physics of NMR spectroscopy | 13 |
| 1.1.1 General overview | 13 |
| 1.1.2 Intramolecular nuclei characteristics | 19 |
| 1.1.3 Quantum mechanics modelling | 22 |
| 1.2 HRMAS Spectroscopy | 25 |
| 1.2.1 Chemical Shift Anisotropy | 25 |
| 1.2.2 Dipolar Coupling | 25 |
| 1.3 Main Metabolites | 27 |
| 1.3.1 Lactate | 27 |
| 1.3.2 Inositol | 28 |
| 1.3.3 Creatine | 28 |
| 1.3.4 Glycine | 28 |
| 1.3.5 Taurine | 28 |
| 1.3.6 Choline | 29 |
| 1.3.7 Glutamine | 29 |
| 1.3.8 Glutamate | 29 |
| 1.3.9 GABA | 29 |
| 1.3.10 N-acetyl-aspartate | 29 |
| 1.3.11 Ethanol | 30 |
| 1.3.12 Lipids | 30 |
| 1.4 Spectrum quantitation | 30 |
| 1.4.1 Methods of quantitation | 31 |
| 1.4.2 Quantitation with QUEST | 32 |
| 1.5 Influence of micro-environment in tissues on quantitation | 33 |

| | | |
|----------|---|------------|
| 2 | Speed-up of NMR QM simulations | 37 |
| 2.1 | Overview of the methods | 37 |
| 2.2 | Decomposition | 37 |
| 2.3 | Splitting of spin-systems | 39 |
| 2.3.1 | “Buffering” effect | 39 |
| 2.3.2 | Peak identification | 42 |
| 2.3.3 | Pseudo-intensities | 42 |
| 2.3.4 | Fragmentation procedure. The kernel | 44 |
| 2.4 | Results | 47 |
| 2.4.1 | Phantoms | 47 |
| 2.4.2 | Metabolites | 49 |
| 2.4.3 | Implementation details | 53 |
| 3 | Signal Processing basis-set optimization | 55 |
| 3.1 | Introduction | 55 |
| 3.2 | Stretching/Shrinking | 55 |
| 3.2.1 | Interpolation | 57 |
| 3.2.2 | Time- or frequency-domain | 59 |
| 3.2.3 | FID cumulated error estimation | 63 |
| 3.2.4 | Spectrum cumulated error estimation | 64 |
| 3.2.5 | Error comparison | 66 |
| 3.2.6 | Akima interpolation | 66 |
| 3.2.7 | Theoretical limits of the method | 68 |
| 3.3 | Splitting of spectra | 70 |
| 3.3.1 | Group selection | 70 |
| 3.3.2 | Theoretical limitations | 72 |
| 3.3.3 | Accuracy-time dilemma | 73 |
| 3.3.4 | Integer-valued shift | 75 |
| 3.3.5 | Strongly-coupled spins | 76 |
| 4 | Optimization based on Quantum Mechanics | 79 |
| 4.1 | Introduction | 79 |
| 4.2 | Method | 79 |
| 4.3 | Computing time costs | 81 |
| 4.4 | Algorithm implementation | 86 |
| 4.4.1 | Gradient descent method | 86 |
| 4.4.2 | Cost function | 86 |
| 4.4.3 | Descent coefficient | 87 |
| 4.4.4 | Optimization loop | 88 |
| 4.5 | Results | 88 |
| 4.5.1 | Simulated signals | 89 |
| 4.5.2 | Real spectra. Oligodendroglioma | 92 |
| 4.5.3 | Quantitation of cell spectra | 96 |
| 4.6 | Conclusion | 99 |
| 5 | General conclusions | 101 |

| | |
|---|------------|
| Bibliography | 104 |
| Publication list | 113 |
| Appendix | 115 |
| Interpolation errors | 115 |
| Integration to obtain Eq. 3.2.28 | 116 |
| Spectrum errors | 116 |
| Fragmentation implementation in jMRUI | 117 |

Summary

From day to day, the role of HRMAS (High-Resolution Magic Angle Spinning) Nuclear Magnetic Resonance Spectroscopy (NMRS) in medical diagnosis is increasing. This technique enables setting up metabolite profiles of *ex vivo* pathological and healthy tissue. Automatic spectrum quantitation enables monitoring of diseases. However for several metabolites, the values of chemical shifts of proton groups may slightly differ according to the micro-environment in the tissue or cells, in particular to its pH. This hampers accurate estimation of the metabolite concentrations mainly when using quantitation algorithms based on a metabolite basis-set.

The present work is devoted to the optimization of NMR metabolite basis set signals, particularly to the algorithms of chemical shift mismatch correction. Two signal processing (“warping”) methods were developed for simple and fast spectrum optimization: signal stretching/shrinking (resampling) and spectrum splitting. Then, another optimization method, QM-QUEST, coupling Quantum Mechanical simulation and quantitation algorithms was implemented. The latter provides more robust fitting while limiting user involvement and respects the correct fingerprints of metabolites. Its efficiency is demonstrated by accurately quantitating signals from tissue samples of human brains with oligodendroglioma, obtained at 11.7 Tesla and spectra of cells acquired at 9.4T by HRMAS-NMR.

As the necessity of fast NMR signal simulation based on Quantum Mechanics is raised in the thesis, a part of the work is dedicated to an approximate method speeding-up the calculations. The algorithm based on spin-system fragmentation could become an important part of the QM-QUEST optimization method and will be implemented as an option of simulation in NMR-SCOPE, module of the jMRUI software package.

Keywords: Nuclear Magnetic Resonance Spectroscopy, HRMAS-NMR, Quantum Mechanics simulation, optimization, quantitation.

Résumé

La Spectroscopie de Résonance Magnétique Nucléaire (RMN) Haute Résolution à l'angle magique (HRMAS) joue un rôle de plus en plus prépondérant pour le diagnostic médical. Cette technique permet d'établir les empreintes *ex vivo* des métabolites de tissus sains et pathologiques. Cependant, pour certains métabolites, les valeurs des déplacements chimiques des groupes de protons peuvent légèrement varier en fonction de l'environnement des tissus ou cellules, particulièrement de son acidité. Cet effet gêne l'estimation correcte des concentrations des métabolites lorsqu'on utilise des algorithmes fondés sur des bases de métabolites.

Ce travail est dévolu aux méthodes d'optimisation des bases de métabolites, notamment aux algorithmes de correction des changements des déplacements chimiques. Deux méthodes de traitement du signal ont été développées pour l'optimisation simple et rapide des signaux/spectres : contraction/expansion du signal moyennant ré-échantillonnage et fractionnement du spectre. Une autre méthode, QM-QUEST, conjuguant la simulation par Mécanique Quantique et la quantification, a été mise en œuvre. Cette dernière permet l'ajustement plus robuste des spectres en limitant l'implication de l'utilisateur et préserve les empreintes correctes des métabolites. Son efficacité est démontrée pour la quantification de spectres RMN de biopsies cérébrales humaines d'oligodendroglioma, obtenues à 11.7 Tesla et de spectres de cellules acquis à 9.4T par la technique RMN-HRMAS.

Etant donné la nécessité de simulation rapide des signaux RMN basée sur la Mécanique Quantique, une partie du travail est vouée à une méthode approchée accélérant la simulation. L'algorithme fondé sur la fragmentation du système de spins pourrait devenir une partie importante de la méthode d'optimisation QM-QUEST et sera mis en œuvre en tant qu'option de simulation de la méthode NMR-SCOPE, module du logiciel jMRUI.

Mots-clés : Spectroscopie de Résonance Magnétique Nucléaire, RMN-HRMAS, simulation par Mécanique Quantique, optimisation, quantification.

Notations

Greek letters

| | |
|----------|----------------------------|
| α | Damping factor |
| γ | Gyro-magnetic ratio |
| δ | Chemical shift |
| ν | Frequency |
| σ | Nuclear shielding constant |
| ϕ | Phase |
| ω | Angular frequency |

Latin letters

| | |
|-------------------|--|
| B_0 | Static magnetic field |
| B_1 | RF field |
| c_m | m th metabolite concentration |
| f | Frequency |
| \mathcal{H} | Hamiltonian |
| t_s | Sampling time |
| T_1, T_2, T_2^* | Relaxation times |

Abbreviations

| | |
|--------|---|
| AMARES | Advanced Method for Accurate, Robust and Efficient Spectral fitting |
| FFT | Fast Fourier Transform |
| FD | Frequency domain |
| HSVD | Hankel Singular Value Decomposition |
| HR | High-Resolution |
| LS | Least Squares |
| MAS | Magic Angle Spinning |
| MRS | Magnetic Resonance Spectroscopy |
| NMR | Nuclear Magnetic Resonance |

| | |
|-------|--|
| QUEST | Quantitation based on quantum estimation |
| SNR | Signal-to-Noise Ratio |
| TD | Time Domain |

Metabolite abbreviations

| | |
|------|----------------------------|
| Cho | Choline |
| Cr | Creatine |
| Eth | Ethanol |
| Glc | Glucose |
| Gln | Glutamine |
| Glu | Glutamate |
| Gly | Glycine |
| GPC | Glycerophosphoryl Choline |
| Lac | Lactate |
| Lip | Lipids |
| MM | Macromolecules |
| NAA | N-acetylaspartate |
| NAAG | N-Acetylaspartateglutamate |
| PCr | Phosphocreatine |
| tCr | Total Creatine (Cr+PCr) |

Other symbols

| | |
|-----------|--|
| $\ z\ $ | Module of complex number z |
| i | Square complex root of -1 ($i^2 = -1$) |
| \hat{p} | Estimated parameters |

Introduction

During the ages the sciences needed methods for analysis of chemical compounds. Evolving from one method to another, in the second half of XXth century mankind obtained a very powerful instrument based on Nuclear Magnetic Resonance (NMR) phenomena. NMR was first described and measured in molecular beams by Isidor Rabi in 1938 [1], and in 1944, Rabi was awarded the Nobel Prize in physics for this work [2]. In 1946, Felix Bloch and Edward Mills Purcell expanded the technique for use on liquids and solids, for which they shared the Nobel Prize in physics in 1952 [3]. They noticed that magnetic nuclei, like ^1H and ^{31}P , could absorb radio frequency energy when placed in a magnetic field of a strength specific to the identity of the nuclei. When this absorption occurs, the nucleus is described as being *in resonance*. Different atomic nuclei within a molecule resonate at different (radio) frequencies for the same magnetic field strength. The observation of such magnetic resonance frequencies of the nuclei present in a molecule enables any trained user to discover essential, chemical and structural information about the molecule.

Nowadays NMR is widely used in chemistry, medicine and other branches that need non-destructive and non-invasive analyses. The application of NMR best known to the general public is Magnetic Resonance Imaging (MRI) for medical diagnosis and MR Microscopy in research settings; however, it is also widely used in chemical studies, notably in NMR spectroscopy such as proton ^1H NMR, carbon ^{13}C NMR and phosphorus ^{31}P NMR. Biochemical information can also be obtained from living tissue (e.g. human brain tumours) with the technique known as *in vivo* Magnetic Resonance Spectroscopy (MRS). As one of the two major spectroscopic techniques used in metabolomics, NMR is used to generate metabolic fingerprints from biological fluids to obtain information about disease states or toxic insults.

With the development of computer sciences some tasks in the NMR domain were automated. One of the major of them is spectrum quantitation – computing of the relative concentration of the chemical compounds in the sample, which a spectrum has been taken from. Rather often, these analyses comprise the spectrum misplacement correction (which can appear due to an inexact reference peak position). The correction of such type is rather easy and accurate in low resolution experiments and analyses. However, for high resolution spectra (such as NMR spectroscopy of liquids or HRMAS spectroscopy) the small mismatches of chemical shifts become significant with respect to reduced line width. This work is devoted to some methods of correction of chemical shift mismatches prior to quantitation.

Goals and outline of the thesis

The thesis attempts to review some problems linked to NMR-peak shifts due to environment influence. As it will be shown, small changes in the environment acidity or in its temperature can cause variations of chemical shifts of the metabolites subject to NMR spectroscopy. Specific attention will be paid to spectrum quantitation based on simulated metabolite basis sets. The higher the resolution of the acquired spectrum, the finer must be the precision of the chemical shift data used for the spectrum simulation. Even small (of the order of some Hz) changes in chemical shifts can significantly hamper the quantitation. Different parts of this work describe various methods of chemical shift correction either based on signal processing or Quantum Mechanics.

Chapter 1

The first chapter introduces the main principles of Nuclear Magnetic Resonance and, particularly, NMR spectroscopy. Physical mechanisms are described in classical and Quantum Mechanical formalism, principal processes are explained. The main human brain metabolites are presented with their principal biochemical functions in the human organism. Some experiments proving the chemical shift dependence on the solution pH are provided.

Chapter 2

Improvement of spectrum simulation is proposed in this chapter to enable spectrum simulation of large spin systems. After having analyzed the main computing time costs for the simulation, a method is presented to decrease the spin system (thus – Hamiltonian) size. An approximate algorithm based on the coupling chain fragmentation is developed. Its application to the simulation of large spin systems will be discussed with analyses of the respective errors.

Chapter 3

The following chapter is devoted to two simple methods of basis-set spectrum correction. They are based on signal processing and are parts of so-called “*warping methods*” (see e.g. [4]). Thus correct metabolite fingerprints are not respected after the algorithm application. The margins of applicability of these methods are estimated, respective results are presented.

Chapter 4

The final chapter presents an accurate algorithm for basis-set spectrum optimization based on automatic chemical shift mismatch correction. The algorithm is based on the spectrum resimulation with the conventional approach of Quantum Mechanics [5]. The optimization is done as an iterative approach (gradient descent) to maximize the cross-correlation between the investigated and optimized

NMR basis-set signals. The algorithm is tested on HRMAS spectra of biopsies of human brains, with oligodendroglioma as well as on spectra of cellular HRMAS spectroscopy.

Préambule

Depuis les temps anciens, les sciences nécessitaient des méthodes d'analyse de composés chimiques. Évoluant d'une méthode à l'autre, dans la deuxième moitié du XXème siècle, l'homme obtint un outil très puissant fondé sur le phénomène de Résonance Magnétique Nucléaire (RMN). Pour la première fois, la RMN a été décrite par Isidor Rabi en 1938 [1] et, en 1944, l'auteur a reçu le Prix Nobel de physique pour ce travail [2]. En 1946, Felix Bloch et Edward Mills Purcell ont étendu la technique aux liquides et solides, en ayant, par la suite, partagé le Prix Nobel de l'année 1952 [3]. Ils ont remarqué que les noyaux magnétiques, comme le proton ^1H ou le phosphore ^{31}P , pouvaient absorber l'énergie de radiofréquence en présence d'un champ magnétique d'une intensité spécifique au type de noyau. Pendant cette absorption, les noyaux sont considérés comme étant "*en résonance*". Les différents noyaux atomiques dans une molécule résonnent à des fréquences différentes pour la même intensité du champ magnétique. L'observation de ces fréquences des noyaux présents dans une molécule permet à un utilisateur qualifié de découvrir l'information essentielle et structurelle de la molécule.

De nos jours, la RMN est très largement appliquée en chimie, médecine et dans d'autres domaines nécessitant des analyses non destructives et non invasives. L'application de la RMN la plus connue au monde, est l'Imagerie de Résonance Magnétique (IRM) pour la diagnostic médical et la Microscopie en RM. En même temps, la RMN s'utilise largement dans les études chimiques, notamment en spectroscopies RMN telles que la RMN des protons (^1H), RMN du carbone (^{13}C) et RMN du phosphore (^{31}P). L'information biomédicale peut aussi être obtenue à partir de tissus vivants (e.g. tumeurs cérébrales humaines) par la technique connue comme Spectroscopie de Résonance Magnétique (MRS) *in vivo*. Comme elle est une des deux techniques spectroscopiques principales pour la métabolomique, la RMN est utilisée pour obtenir les empreintes de métabolites des tissus et liquides biologiques. Cela permet d'obtenir l'information sur l'état de la maladie ou des influences de produits toxiques ou de médicaments.

Avec le développement des sciences informatiques, plusieurs tâches dans le domaine de la RMN ont été automatisées. L'une des principales est la quantification de spectres – le calcul des concentrations relatives des composés chimiques dans l'échantillon à partir du spectre. Très souvent, ces analyses comprennent la recalage du spectre (nécessaire parfois comme la conséquence de la position inexacte du pic de référence). Cette correction est assez simple et précise pour les expériences à basse résolution. Cependant, pour les spectres haute résolution (tels que ceux obtenus en spectroscopie des liquides ou en spectroscopie HRMAS) les petits décalages

des déplacements chimiques deviennent significatifs comparés à la largeur de raie. Ce travail est dévolu aux méthodes de correction des décalages de déplacements chimiques avant la quantification.

Buts et contenu de la thèse

La thèse vise à examiner certains problèmes liés aux changements de déplacements chimiques dans les spectres RMN en raison de l'influence de l'environnement. Il a été démontré que les petits changements de l'acidité ou de la température de l'environnement peuvent provoquer des variations des déplacements chimiques des métabolites étudiés par spectroscopie RMN. Une attention particulière sera portée à la quantification de spectres fondée sur les bases de métabolites simulées. Meilleure est la résolution du spectre acquis, plus la précision des déplacements chimiques doit être fine pour la simulation des spectres de base des métabolites. Même les petits changements (de l'ordre de quelques Hertz) peuvent gêner significativement la quantification. Les chapitres de cette thèse décrivent différentes méthodes de correction des déplacements chimiques fondées soit sur le traitement du signal soit sur la Mécanique Quantique.

Chapitre 1

Le premier chapitre introduit les principes généraux de la Résonance Magnétique Nucléaire et, particulièrement, la spectroscopie RMN. Les lois physiques fondamentales du phénomène sont présentées à l'aide du formalisme classique et de la mécanique quantique, les processus principaux sont expliqués. Les métabolites plus importants du cerveau humain sont présentés avec leurs fonctions biochimiques principales dans l'organisme humain. Certaines expériences prouvant la dépendance des déplacements chimiques en fonction du pH de la solution sont fournies.

Chapitre 2

Une méthode approchée de simulation de spectres est présentée dans ce chapitre, elle permet la simulation de grands systèmes de spins. Après avoir analysé les coûts de temps de calcul de la simulation, nous présentons une méthode permettant de réduire la taille du système de spins (ainsi que du Hamiltonien). Un algorithme approché fondé sur la fragmentation de la chaîne de spins couplés est développé. Son application à la simulation des grands systèmes des spins sera effectuée ainsi que l'analyse des erreurs respectives.

Chapitre 3

Le chapitre suivant est voué à deux méthodes de correction des bases de métabolites. Elles sont fondées sur le traitement du signal et font partie des méthodes dites

“*warping*” (à voir, par exemple [4]). Cependant, les empreintes correctes des métabolites ne sont plus respectées après application de l’algorithme. Les marges d’applicabilité de ces méthodes sont estimées, les résultats respectifs sont présentés.

Chapitre 4

Le chapitre final présente un algorithme précis pour l’optimisation des spectres de bases des métabolites fondé sur la correction automatique des décalages des déplacements chimiques. L’algorithme est fondé sur la resimulation du spectre en approche conventionnelle par Mécanique Quantique [5]. L’optimisation est faite lors d’une approche itérative (algorithme du gradient) maximisant la corrélation croisée entre les signaux RMN à quantifier et les signaux de la base à optimiser. L’algorithme a été testé avant la quantification de signaux HRMAS de biopsies cérébrales humaines d’oligodendrogliomas et sur des spectres HRMAS de cellules.

Chapter 1

Nuclear Magnetic Resonance Spectroscopy

1.1 Physics of NMR spectroscopy

1.1.1 General overview

Nuclear magnetic resonance spectroscopy [6, 7], most commonly known as NMR spectroscopy, is the name given to the technique which exploits the magnetic properties of certain nuclei. The most important applications for the organic chemistry are proton NMR and carbon-13 NMR spectroscopy.

Nuclear magnetic resonance phenomenon

Each nucleus has its property called spin quantum number I (or simply, spin) which can take only discrete values, namely: to be integer, half-integer or zero. The nuclei, which have a non-zero spin, present a magnetic moment μ that can be expressed as:

$$\mu = \gamma \left(I \frac{h}{2\pi} \right) \quad (1.1.1)$$

where γ is the gyromagnetic ratio: $\gamma = \frac{e}{2m}$, h is the Planck's constant, e is the proton charge.

The protons and neutrons have a spin quantum number $I = 1/2$. Having knowledge about an element's atomic number one can judge about its nuclear spin. If the atomic number and mass number are even, then the magnetic moments of protons (as the ones of neutrons) are compensated. In this case $I = 0$. This, for example, is the case for ${}^{12}_6\text{C}$ or ${}^{16}_8\text{O}$. If the mass number is odd and the atomic number is even (this means that the nucleus has an even number of protons and the same number of neutrons), its nucleus has an integer spin (${}^{14}_7\text{N}$, ${}^2_1\text{D}$). Finally, if the mass number is odd, the nucleus has a half-integer spin. These nuclei are particularly suitable for NMR spectroscopy: ${}^1_1\text{H}$, ${}^{19}_9\text{F}$, ${}^{31}_{15}\text{P}$, ${}^{13}_6\text{C}$, ${}^{15}_7\text{N}$. They all have $I = 1/2$ so no electric quadrupole.

Being placed into a magnetic field B_0 , each atom, which has a non-zero nuclear magnetic moment, can occupy $2I + 1$ different energy levels. For the simplest case of $I = 1/2$ they are $+\mu B_0$ and $-\mu B_0$, which correspond to equal and opposite orientations of the magnetic moment in the external magnetic field. Usually, these spin states are noted as α and β (see Fig. 1.1).

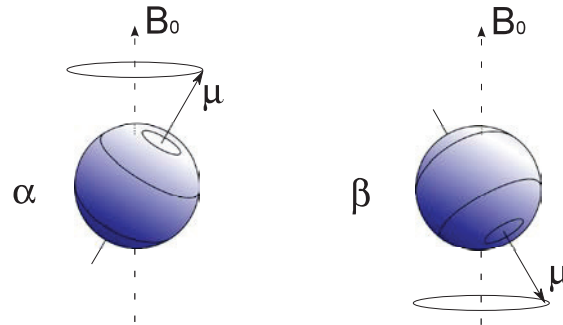


Figure 1.1: α - and β -states of a nucleus spin placed into an external magnetic field B_0 .

The energy difference of $2\mu B_0$ between possible states leads to the nuclear magnetic resonance condition. The nucleus will be subject to the NMR effect if, being placed into a static magnetic field B_0 , it is irradiated by an electromagnetic field with a frequency ν_0 which is

$$\nu_0 = \frac{2\mu B_0}{h} = \frac{\gamma B_0}{2\pi} \quad (1.1.2)$$

Because of Larmor precession, all magnetic moments will precess around the direction \vec{B}_0 with the frequency

$$\omega_0 = 2\pi\nu_0 = \gamma B_0 \quad (1.1.3)$$

This is expressed in the laboratory system of coordinates xyz . In the system $x'y'z'$, rotating with the frequency ω around the axe z (which corresponds to the direction of \vec{B}_0), we will be interested in the difference of the Larmor frequency and the frequency of the system: $\omega_0 - \omega$.

Magnetizations and relaxations

Let us now introduce the macroscopic magnetization:

$$\vec{M} = \sum \vec{\mu} \quad (1.1.4)$$

In the free state, the global magnetization of a sample is equal to zero: the thermal motion of the molecules leads to a chaotic orientation of the magnetic vectors. But being put into a static magnetic field, the sample obtains a magnetization according to Boltzmann equilibrium because of reorientation of magnetic moments along two possible orientations:

$$\frac{N^-}{N^+} = \exp\left(\frac{-E}{kT}\right) \quad (1.1.5)$$

where N^+ and N^- are respectively the numbers of spins parallel and antiparallel with the magnetic field, E is the potential energy of the spin in a field ($E = -\vec{\mu}\vec{B}$), k is the Boltzmann constant, and T is the absolute temperature.

The bigger the magnetic field strength, the bigger is the energy difference between the levels and the difference between the level populations. Schematically, this is illustrated in Fig. 1.2

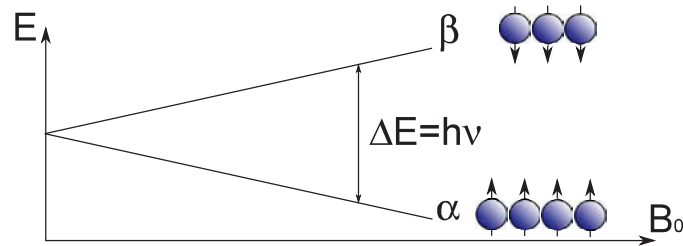


Figure 1.2: The nuclear spin energy diagram for $I = 1/2$. β -state is the high-energy state because of the antiparallel alignment between B_0 and μ_z . α -state is the low-energy state, thus its population is higher than the population of β -level.

While an RF-field B_1 , orthogonal to B_0 , with a resonance frequency $\omega = \omega_0$, is applied to a set of spins placed in B_0 , the vector of magnetization turns to the transverse plane xy . In other words: the vector of magnetization starts to precess around the direction of \vec{B}_1 perpendicular to \vec{B}_0 creating a transverse magnetization M_{xy} , leaving its z -component (longitudinal magnetization) to decrease (Fig. 1.3 a).

After switching the RF-field B_1 off, in the assumption of absence of any spin-spin interaction and of the perfect homogeneity of B_0 , the vector of magnetization would return to the direction of z -axis following the exponential law:

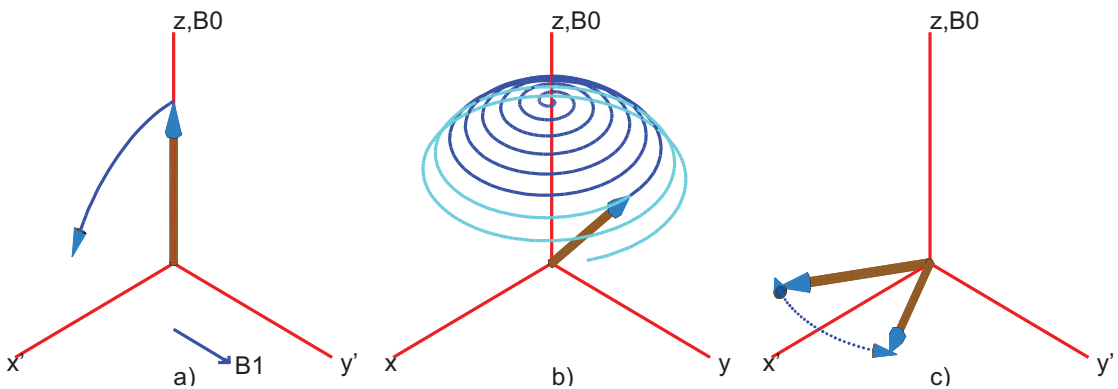


Figure 1.3: Spin behaviour (a) rotation, b) spin-lattice relaxation, c) spin-spin relaxation) in a magnetic field.

$$M_z(t) = M_0 \left[1 - \exp\left(-\frac{t}{T_1}\right) \right] \quad (1.1.6)$$

where M_0 is the equilibrium longitudinal magnetization and T_1 is a characteristic time of the process. The cause of this effect is the restoration of Boltzmann equilibrium (see Eq. 1.1.5). The time T_1 , present in Eq. 1.1.6, is called the characteristic time of spin-lattice relaxation.

This time is comparatively large relatively to the characteristic times of the other relaxation process present in the described system. T_2 describes the so-called spin-spin relaxation or transverse relaxation. This is an entropy-process which corresponds to the exchange of the energy between the spins:

$$M_{y'}(t) = M_{y'}(0) \exp\left(\frac{-t}{T_2}\right) \quad (1.1.7)$$

These two characteristic relaxation times – T_1 and T_2 – were introduced in the assumption of the totally homogeneous field B_0 . In the real experiment this condition can never be achieved. So, another relaxation time T_2^* is considered which takes into account two processes: spin-spin relaxation with the characteristic time T_2 and the dephasing of the spins due to the effects of inhomogeneity of B_0 (see Fig. 1.3, c):

$$\frac{1}{T_2^*} = \frac{1}{T_2} + \frac{1}{T_{2\text{inh}}} \quad (1.1.8)$$

The latter introduced relaxation time T_2^* is the smallest of all the mentioned ones. But the effects of dephasing due to $T_{2\text{inh}}$ are not irreversible. For example, the spin-echo experiment [8] allows to rephase the spins, dephased due to T_2^* , to obtain echo signals whose amplitudes decay with T_2 .

Schematically, the two decay times are presented in the spin-echo pulse sequence in Fig. 1.4

NMR signal

While the transverse magnetization in the plane xy exists and is not constant, its changes create the changes of magnetic flux through a chosen surface.

$$\Phi = \int_{\text{area}} \vec{B} d\vec{S} \quad (1.1.9)$$

If a coil corresponds to the introduced *area*, then, due to the changes of the flux, an electromotive force is induced into it:

$$\text{emf} = -B \frac{d}{dt} \Phi(t) \quad (1.1.10)$$

So, it is possible to read an electro-magnetic signal induced by a rotating magnetization vector in the transverse plane. If we read the signal induced in two orthogonal directions x and y , then, conventionally, the x component is the

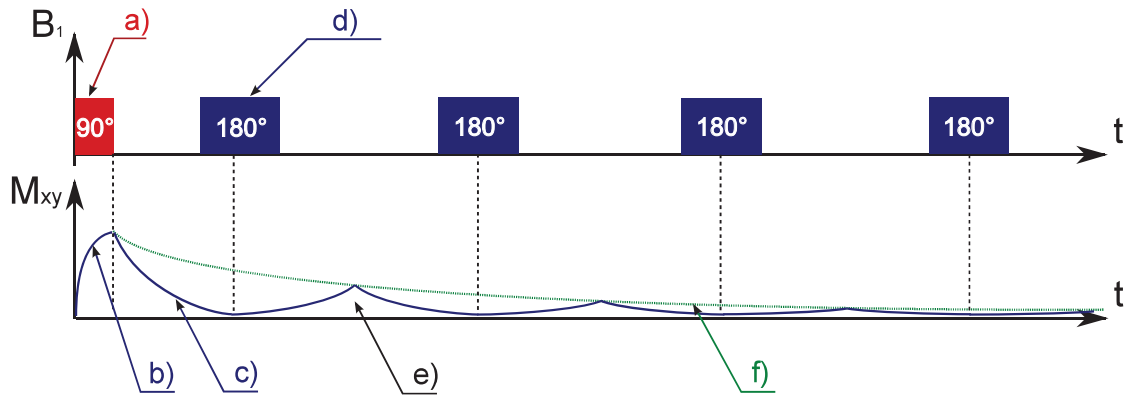


Figure 1.4: Spin echo sequence and relaxation times: a) excitation pulse; b) rotation of M_z , appearance of the transverse magnetization; c) dephasing due to inhomogeneities of B_0 , characteristic time T_2^* ; d) rephasing pulse; e) rephasing of the spins in the transverse plane, echo; f) spin-spin relaxation, characteristic time T_2 .

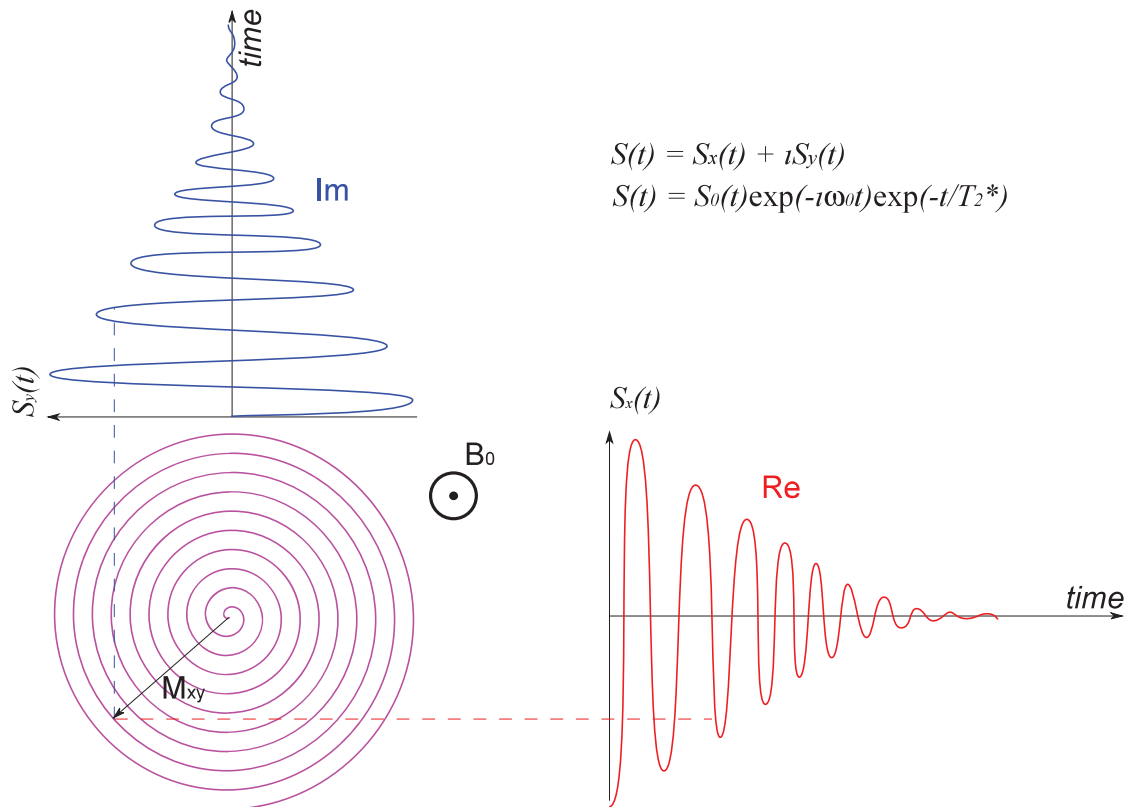


Figure 1.5: Complex representation of an NMR signal. T_2^* represents the characteristic time of the decay.

real part of a complex-represented signal and the y component is the imaginary part (see Fig. 1.5).

The signal is called free-induction decay, or, simply, FID.

Demodulation

The conservation of data of spin rotation in laboratory frame is redundant. Only the difference $\omega_0 - \omega$ between the Larmor frequency and the precession frequency of the particular vectors of magnetization is of interest. In fact, the “raw” FID-signal is a modulated signal with the modulation frequency ω_0 . The process of subtraction is called demodulation.

This is illustrated in Fig. 1.6. If one looks at the signal from the laboratory frame, the oscillations at a high frequency ω are observed. Let us firstly imagine that the magnetic field in the experiment is totally homogenous, thus all the vectors of magnetizations precess with the Larmor frequency $\omega = \omega_0$ (Fig. 1.6, top). After the demodulation of such signal, one obtains the free induction decay envelope: $\exp -t/T_2^*$ (Fig. 1.6, middle). If some frequencies in the sample are different from the Larmor one because of some inhomogenities of the external magnetic field, then the demodulated signal will show the oscillations at the frequencies $\omega - \omega_0$, which are much smaller than the Larmor frequency (see Fig. 1.6, bottom).

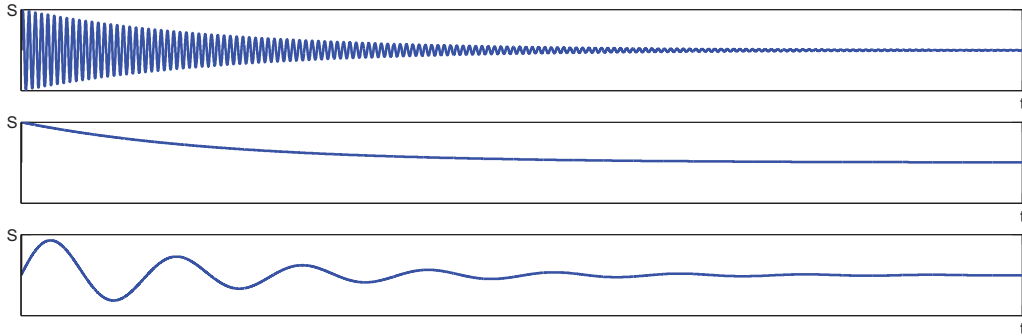


Figure 1.6: Top: an NMR signal seen from the laboratory frame; middle: demodulated signal from a magnetization vector which rotates at Larmor frequency; bottom: demodulated signal from a magnetization vector which rotates at a frequency, slightly different from the Larmor one.

The sum of these different demodulated oscillations will produce the total NMR signal:

$$s(t) = \sum_m s_m(t) = \sum_m c_m \exp\left(\frac{-t}{T_{2m}^*}\right) \exp[i(\omega t + \phi_m)] \quad (1.1.11)$$

where c_m is proportional to the m -th resultant vector of magnetization which (after being demodulated) oscillates with the frequency ω_m , having an initial phase ϕ_m and decaying with the characteristic time $1/T_{2m}^*$.

These signals are often represented as spectra, after being subject to a fast Fourier transform:

$$S(\omega_n) = \sum_{t_i=0}^{N-1} s(t_i) \exp\left(-\frac{2\pi}{N} i \omega_n t_i\right) \quad \omega = 0, \dots, \omega_N \quad (1.1.12)$$

Here we talk directly about the discrete but not the continuous Fourier transform, because the acquired signals are digitized.

A simulated signal, consisting of four spectral components, and its fast Fourier transform are presented in Fig. 1.7

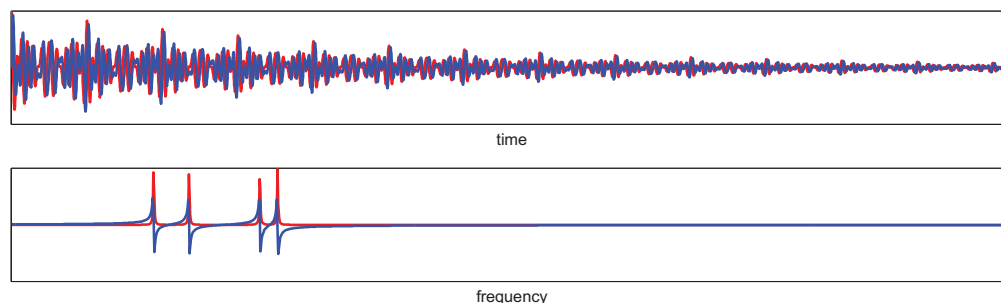


Figure 1.7: Complex signal representation: red – real part, blue – imaginary part. Top: A signal as a sum of four spectral components; bottom: its fast Fourier transform.

1.1.2 Intramolecular nuclei characteristics

Chemical shift

Depending on the molecule's structure each nucleus has its own property called chemical shift. The total magnetic field experienced by a nucleus includes local magnetic fields induced by currents of electrons in the molecular orbitals (note that electrons have a magnetic moment themselves).

Let us consider an atom with a symmetric electron shell. In an external magnetic field atom electrons precess around the axis which passes through the centre of the atom and is colinear with the field direction. This precession creates a secondary field \vec{B}' which, close to the nucleus, is opposite to the main field \vec{B}_0 . In turn, the field strength B' is proportional to the angular velocity of electrons' precession, so, to the strength of the external field:

$$B' = -\sigma B_0 \quad (1.1.13)$$

where σ is the shielding constant. This means that the nuclei are situated in an efficient field B_{eff} , caused by “screening” of the external field by the electrons:

$$B_{\text{eff}} = B_0 - \sigma B_0 = B_0(1 - \sigma) \quad (1.1.14)$$

Equation 1.1.14 can be rewritten to express σ :

$$\sigma = \frac{B_0 - B_{\text{eff}}}{B_0} = \frac{\Delta B}{B_0} \quad (1.1.15)$$

The electron distribution of the same type of nucleus (e.g. ^1H , ^{13}C) usually varies according to the local geometry (binding partners, bond lengths, angles

between bonds, ...), and it influences the local magnetic field at each nucleus. This is reflected in the spin energy levels (and resonance frequencies).

Using Eq. 1.1.14, the resonance condition of Eq. 1.1.2 can be modified to:

$$\nu = \frac{\gamma B_0}{2\pi} (1 - \sigma) \quad (1.1.16)$$

The chemical shift, labeled later as δ , shows the ratio between the difference of the resonance frequency and that of a reference peak to the resonance frequency. Since the numerator usually has the order of tens or hundreds hertz, and the order of denominator is of hundreds megahertz, delta is generally expressed in *ppm* which means “part per million”:

$$\delta = \frac{\nu - \nu_{\text{ref}}}{\nu_{\text{ref}}} \times 10^6 \quad (1.1.17)$$

The value of the chemical shift is given with respect to a reference frequency or to a reference sample – usually a molecule with a barely distorted electron distribution.

The detected frequencies for ^1H , ^{13}C , and ^{29}Si nuclei are usually referenced against TMS (tetramethylsilane, see Fig. 1.8). Because all twelve hydrogen atoms in a tetramethylsilane molecule are equivalent, its ^1H NMR spectrum consists of a singlet, which is assigned the chemical shift of zero. The majority of compounds studied by ^1H NMR spectroscopy absorb downfield of the TMS signal, thus there is usually no interference between the standard and the sample. For example, the water hydrogens have the chemical shift around 4.7ppm relatively to tetramethylsilane singlet.

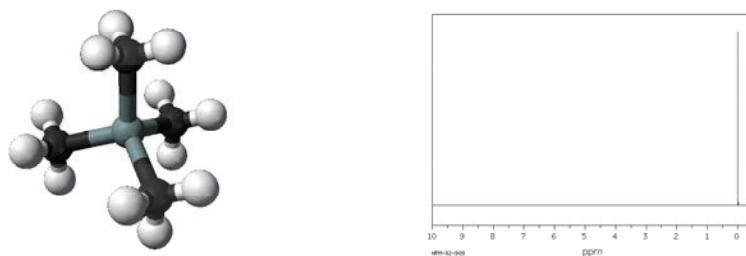


Figure 1.8: Tetramethylsilane spatial structure and its ^1H NMR spectrum according to the Spectral Database for Organic Compounds, SDBS (National Institute of Advanced Industrial Science and Technology, Japan)[9].

Proton-proton coupling

Some of the most useful information for structure determination in a one-dimensional NMR spectrum comes from J-coupling or scalar coupling (a special case of spin-spin coupling) between NMR active nuclei. This coupling arises from the interaction of different spin states through the chemical bonds of a molecule and results in the splitting of NMR signals. These splitting patterns can be complex

or simple and, likewise, can be straightforwardly interpretable or deceptive. This coupling provides detailed insight into the connectivity of atoms in a molecule.

Coupling to n equivalent (spin $1/2$) nuclei (the nuclei in a molecule which are transformed into each other by rotations, reflections, or combinations of these operations, leaving the molecule invariant) splits the signal into an $n + 1$ multiplet with intensity ratios following Pascal's triangle as described in Tab. 1.1. Coupling to additional spins will lead to further splitting of each component of the multiplet e.g. coupling to two different spin $1/2$ nuclei with significantly different coupling constants will lead to a doublet of doublets (abbreviation: *dd*). Note that coupling between nuclei that are chemically equivalent (that is, have the same chemical shift) has no effect on the NMR spectra and couplings between nuclei that are distant (usually, more than 3 bonds apart for protons in flexible molecules) are usually too small to cause observable splitting. Long-range couplings over more than three bonds can often be observed in cyclic and aromatic compounds, leading to more complex splitting patterns. In Tab. 1.1 some examples of splitting are presented.

For example, in the proton spectrum of ethanol $\text{CH}_3\text{-CH}_2\text{-OH}$, the CH_3 group is split into a triplet with an intensity ratio of 1:2:1 by the two neighboring CH_2 protons. Similarly, the CH_2 is split into a quartet with an intensity ratio of 1:3:3:1 by the three neighboring CH_3 protons. In principle, the two CH_2 protons would also be split again into a doublet to form a doublet of quartets by the hydroxyl proton, but intermolecular exchange of the acidic hydroxyl proton often results in a loss of coupling information.

| Multiplicity | Intensity Ratio |
|--------------|------------------|
| Singlet (s) | 1 |
| Doublet (d) | 1:1 |
| Triplet (t) | 1:2:1 |
| Quartet (q) | 1:3:3:1 |
| Quintet | 1:4:6:4:1 |
| Sextet | 1:5:10:10:5:1 |
| Septet | 1:6:15:20:15:6:1 |

Table 1.1: Relative intensities of the multiplet's peaks

The sequence of "Pascal triangle" for the intensity distribution within a multiplet is only valid for weakly coupled spins. The effects of strong coupling can lead to additional redistribution of the energy between the peaks. The Quantum Mechanical explanation of this effect will be provided in the following section.

Coupling combined with the chemical shift tell us not only about the chemical environment of the nuclei, but also the number of neighboring NMR active nuclei within the molecule. In more complex spectra with multiple peaks at similar chemical shifts or in spectra of nuclei other than hydrogen, coupling is often the only way to distinguish different nuclei.

1.1.3 Quantum mechanics modelling

The energy of an atom, or radical, containing unpaired electrons and nuclei with nonzero spins, can be expressed in terms of the Hamiltonian (energy) operator. To simplify the explanation, we will continue with the assumption of two-spin case. All the other cases can be derived similarly as, for example, it is presented in [5]. The Hamiltonian for two-spin system with isotropic coupling is

$$\mathcal{H} = \delta_1 I_{1Z} + \delta_2 I_{2Z} + J \vec{I}_1 \cdot \vec{I}_2, \quad (1.1.18)$$

considering $\hbar = 1$.

The spin Hamiltonian has a dimensionality $(2I_1 + 1)(2I_2 + 1) = 4$ for the case $I_1 = I_2 = 1/2$, and there are a corresponding number of basis function or ket vectors, $|I_1 I_2 m_1 m_2\rangle$, which will be written in the shortened notation $|m_1 m_2\rangle$, where m_1 and m_2 are the magnetic quantum numbers of the respective atoms ($m_1 = \pm\frac{1}{2}$, $m_2 = \pm\frac{1}{2}$):

$$\begin{aligned} \left| \frac{1}{2} \right\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \Rightarrow \alpha\text{-state}, \\ \left| -\frac{1}{2} \right\rangle &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Rightarrow \beta\text{-state}. \end{aligned} \quad (1.1.19)$$

Consequently, the ket vectors $|m_1 m_2\rangle$ are expressed by direct product expansion of these kets:

$$\begin{aligned} \left| \frac{1}{2} \frac{1}{2} \right\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ \left| \frac{1}{2} -\frac{1}{2} \right\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \\ \left| -\frac{1}{2} \frac{1}{2} \right\rangle &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \\ \left| -\frac{1}{2} -\frac{1}{2} \right\rangle &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \end{aligned} \quad (1.1.20)$$

One must calculate the matrix elements $\langle \phi_1 | \mathcal{H} | \phi_2 \rangle$, where ϕ_1 and ϕ_2 are the basis functions. The eigenfunctions ψ_j of the Hamiltonian will be linear combinations of these basis functions.

Each operator in the spin Hamiltonian is a 2×2 matrix. As a result Eq.

(1.1.18) becomes:

$$\mathcal{H} = \frac{1}{2} \left[\delta_1 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \delta_2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right] + \frac{J}{4} \left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right] \quad (1.1.21)$$

After all the calculations done one gets a 4×4 Hamiltonian matrix:

$$\mathcal{H} = \frac{1}{2} \begin{pmatrix} \delta_1 + \delta_2 + \frac{J}{2} & 0 & 0 & 0 \\ 0 & \delta_1 - \delta_2 + \frac{J}{2} & J & 0 \\ 0 & J & \delta_2 - \delta_1 - \frac{J}{2} & 0 \\ 0 & 0 & 0 & -\delta_2 - \delta_1 + \frac{J}{2} \end{pmatrix} \quad (1.1.22)$$

Each basis function is a product of the wave functions $|I, m\rangle$ of each nucleus defined as:

$$\begin{aligned} \mathbf{I}_z |I, m\rangle &= m |I, m\rangle \\ \mathbf{I}^2 |I, m\rangle &= I(I+1) |I, m\rangle \\ \mathbf{I}^\pm |I, m\rangle &= [(I \mp m)(I \pm m + 1)]^{1/2} |I, m \pm 1\rangle \end{aligned} \quad (1.1.23)$$

Here \mathbf{I}^\pm is the transition operator. So the basis functions will be:

$$|I_1, m_1 \pm 1\rangle \otimes |I_2, m_2 \pm 1\rangle \otimes \cdots \otimes |I_n, m_n \pm 1\rangle \quad (1.1.24)$$

where each ket corresponds to its particle. In the Tab. 1.2 the simplified notation is used:

$$\begin{aligned} |+\rangle &\text{ for } |1/2, 1/2\rangle, \quad \alpha\text{-state} \\ |-\rangle &\text{ for } |1/2, -1/2\rangle, \quad \beta\text{-state} \end{aligned} \quad (1.1.25)$$

After being diagonalized this matrix provides the eigenvalues, whose differences correspond to the frequencies of the spin system, and the set of eigenvectors, which will allow to find the amplitude for each transition.

| \mathcal{H} | $ ++\rangle$ | $ +-\rangle$ | $ -+\rangle$ | $ --\rangle$ |
|---------------|-------------------------------------|-------------------------------------|-------------------------------------|--------------------------------------|
| $ ++\rangle$ | $\delta_1 + \delta_2 + \frac{J}{2}$ | 0 | 0 | 0 |
| $ +-\rangle$ | 0 | $\delta_1 - \delta_2 + \frac{J}{2}$ | J | 0 |
| $ -+\rangle$ | 0 | J | $\delta_2 - \delta_1 - \frac{J}{2}$ | 0 |
| $ --\rangle$ | 0 | 0 | 0 | $-\delta_2 - \delta_1 + \frac{J}{2}$ |

Table 1.2: General form of the Hamiltonian for a two-spin system

Each row and column of this Hamiltonian matrix corresponds to a basis function (see Tab. 1.2). Let us have a look at the structure of the Hamiltonian matrix. It is partially diagonal, since it is in the form of Table 1.3.

| | | ++⟩ | +-⟩ | -+⟩ | --⟩ | |
|-----|-------|-------|-----|-----|-------|--|
| ++⟩ | 1 × 1 | 0 | 0 | 0 | 0 | |
| +-⟩ | 0 | 2 × 2 | | 0 | 0 | |
| -+⟩ | 0 | 0 | 0 | 0 | 0 | |
| --⟩ | 0 | 0 | 0 | 0 | 1 × 1 | |

Table 1.3: Block representation of the Hamiltonian.

The transition operator, described in (1.1.23) can be also rewritten in matrix form:

$$\begin{aligned} I^+ &= I_x + iI_y, \\ I^- &= I_x - iI_y. \end{aligned} \quad (1.1.26)$$

where

$$\begin{aligned} I_x &= \sum_{i,j=1}^{N_{\text{spins}}} \begin{cases} \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \text{for } j = i - 1, \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{otherwise;} \end{cases} \\ I_y &= \sum_{i,j=1}^{N_{\text{spins}}} \begin{cases} \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} & \text{for } j = i - 1, \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{otherwise.} \end{cases} \end{aligned} \quad (1.1.27)$$

For example, for two-spin cases this gives:

$$I^+ = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad I^- = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}. \quad (1.1.28)$$

The expressions (1.1.28) show a correspondence of the possible transitions between the states to the respective cell values. The zeros of the transition matrices show the impossibility of the transitions between the states for which

$$\sum \delta_{ik} \neq \pm 1 \quad (1.1.29)$$

where δ_{ik} represents +1 or -1 for the wave functions $|+\rangle$ or $|-\rangle$ respectively.

So, having now the I^+ and I^- operators and the eigenvectors ψ of the diagonalized Hamiltonian, we are able to calculate the probabilities (thus – intensities) of the respective transitions:

$$P_{i \leftrightarrow j} = (\langle \psi_i | (I^+ + I^-) | \psi_j \rangle)^2 \quad (1.1.30)$$

1.2 HRMAS Spectroscopy

Nowadays, medical diagnoses often exploit results obtained from High-Resolution Magic Angle Spinning Nuclear Magnetic Resonance (HRMAS-NMR) ^1H spectroscopy. This technique enables setting up metabolite profiles of *ex vivo* pathological and healthy tissue, i.e. biopsies [10–22]. There are two directionally dependent interactions commonly found in solid-state NMR [23]: the chemical shift anisotropy (CSA) and the internuclear dipolar coupling.

1.2.1 Chemical Shift Anisotropy

The chemical shift tensor can be diagonalized:

$$\delta = \begin{pmatrix} \delta_{11} & \delta_{12} & \delta_{13} \\ \delta_{21} & \delta_{22} & \delta_{23} \\ \delta_{31} & \delta_{32} & \delta_{33} \end{pmatrix} \rightarrow \begin{pmatrix} \delta'_{11} & 0 & 0 \\ 0 & \delta'_{22} & 0 \\ 0 & 0 & \delta'_{33} \end{pmatrix} \quad (1.2.1)$$

$$\begin{aligned} \nu_{CS} &= \frac{\gamma B_0}{2\pi} [1 - (\sigma_{11} \sin^2 \theta \cos^2 \phi + \sigma_{22} \sin^2 \theta \sin^2 \phi + \sigma_{33} \cos^2 \theta)] \\ &= \frac{\gamma B_0}{2\pi} (\delta_{11} \sin^2 \theta \cos^2 \phi + \delta_{22} \sin^2 \theta \sin^2 \phi + \delta_{33} \cos^2 \theta) \end{aligned} \quad (1.2.2)$$

The chemical shift tensor can be placed into its principal axis system (PAS) – the reference frame in which it is diagonal and all off-diagonal components are zero – by diagonalization. The eigenvalues that result – δ'_{11} , δ'_{22} and δ'_{33} – are called the principal components, and the eigenvectors are the direction cosines that relate the PAS to the frame in which the tensor was originally expressed. This is usually the lab frame (defined by the field direction) or the molecular frame (defined by the local molecular symmetry). The average value of the tensor elements is one third of the trace of the tensor, where the trace is the sum of the diagonal components:

$$\sigma_{\text{iso}} = \frac{1}{3}(\sigma_{11} + \sigma_{22} + \sigma_{33}) \quad \delta_{\text{iso}} = \frac{1}{3}(\delta_{11} + \delta_{22} + \delta_{33}) \quad (1.2.3)$$

The resulting chemical shift Hamiltonian includes terms which depend on the orientation of the tensor with respect to the magnetic field:

$$\begin{aligned} \mathcal{H}_{CS} &= \gamma B_0 [1 - (\sigma_{11} \sin^2 \theta \cos^2 \phi + \sigma_{22} \sin^2 \theta \sin^2 \phi + \sigma_{33} \cos^2 \theta)] \hat{I}_Z \\ &= \gamma B_0 [\delta_{11} \sin^2 \theta \cos^2 \phi + \delta_{22} \sin^2 \theta \sin^2 \phi + \delta_{33} \cos^2 \theta] \hat{I}_Z \end{aligned} \quad (1.2.4)$$

1.2.2 Dipolar Coupling

Magnetic dipole-dipole interaction, also called dipolar coupling, refers to the direct interaction between two magnetic dipoles [24]. The Hamiltonian of the interaction is as follows:

$$\mathcal{H} = -\frac{\mu_0}{4\pi r_{jk}^3} (3(\mathbf{m}_j \cdot \mathbf{e}_{jk})(\mathbf{m}_k \cdot \mathbf{e}_{jk}) - \mathbf{m}_j \cdot \mathbf{m}_k) \quad (1.2.5)$$

where e_{jk} is a unit vector parallel to the line joining the centers of the two dipoles, r_{jk} is the distance between two dipoles, m_k and m_j .

For two interacting nuclear spins:

$$\mathcal{H} = -\frac{\mu_0}{4\pi} \frac{\gamma_j \gamma_k}{r_{jk}^3} (3(\mathbf{I}_j \cdot \mathbf{e}_{jk})(\mathbf{I}_k \cdot \mathbf{e}_{jk}) - \mathbf{I}_j \cdot \mathbf{I}_k) \quad (1.2.6)$$

γ_j , γ_k and r_{jk} are gyromagnetic ratios of two spins and spin-spin distance respectively.

These interactions strongly depend on the molecular orientation and make it impossible to obtain high-resolved spectra. For example, the line width of the ^1H NMR resonance of water is of the order of 0.1Hz, while the line from a static sample of ice is approximately 100kHz wide, i.e. a million times broader[24].

If one considers the case of protons as dipoles, from Eq. 1.2.6, one obtains:

$$\mathcal{H} = -\frac{1}{2}d[I_1 I_2 - 3I_{1z} I_{2z}](1 - 3\cos^2 \theta) \quad (1.2.7)$$

where d is the dipolar coupling constant and θ is the angle between the direction of the magnetic field B_0 and the vector connecting two protons.

The isotropic reorientation of molecules in liquids reduces the spectral effects of the dipolar interaction because the integral

$$\int (1 - 3\cos^2 \theta) d\Omega \quad (1.2.8)$$

over the sphere vanishes.

The same effect may be obtained by choosing the angle

$$\theta = \arccos\left(\frac{1}{\sqrt{3}}\right). \quad (1.2.9)$$

called “magic angle”.

To see the application of this, we will assume a sample rotating in the magnetic field B_0 inclined with an angle α . One proton is placed at the origin. Denoting the angle between the directions to the second proton and the axis of rotation as β , one can deduce (as it shown in [24]):

$$3\cos^2 \theta - 1 = \frac{3}{2} \sin 2\alpha \sin 2\beta \cos \phi \quad (1.2.10)$$

where ϕ is the rotation angle: $\frac{d\phi}{dt} = \Omega$. As the angles α and β are constant and the average value of the cosines $\cos 2\phi$ and $\cos \phi$ are zero, we obtain:

$$\overline{3\cos^2 \theta - 1} = \frac{1}{2} (3\cos^2 \alpha - 1) (3\cos^2 \beta - 1) \quad (1.2.11)$$

Thus, if the value of α corresponds to the “magic” angle, then $\overline{3\cos^2 \theta - 1} = 0$, independently of the value of β which may differ for different pairs of protons. This means that all dipolar interactions between protons within the sample are then reduced to zero.

1.3 Main Metabolites

Our work being devoted to the optimization of the metabolite basis-sets, it is important to say some words about the most current metabolites. Let us take as an example a spectrum, presented by Cheng et al. in 1997 [25]. It is a Proton MR spectrum (400 MHz) of human brain tissue collected from superior temporal gyrus histologically determined to be mildly affected with Pick disease. Despite the fact that the tissue was not healthy (high Lactate concentration), the spectrum from Fig. 1.9 finely represents major metabolites, observed in HRMAS spectroscopy of the brain tissues.

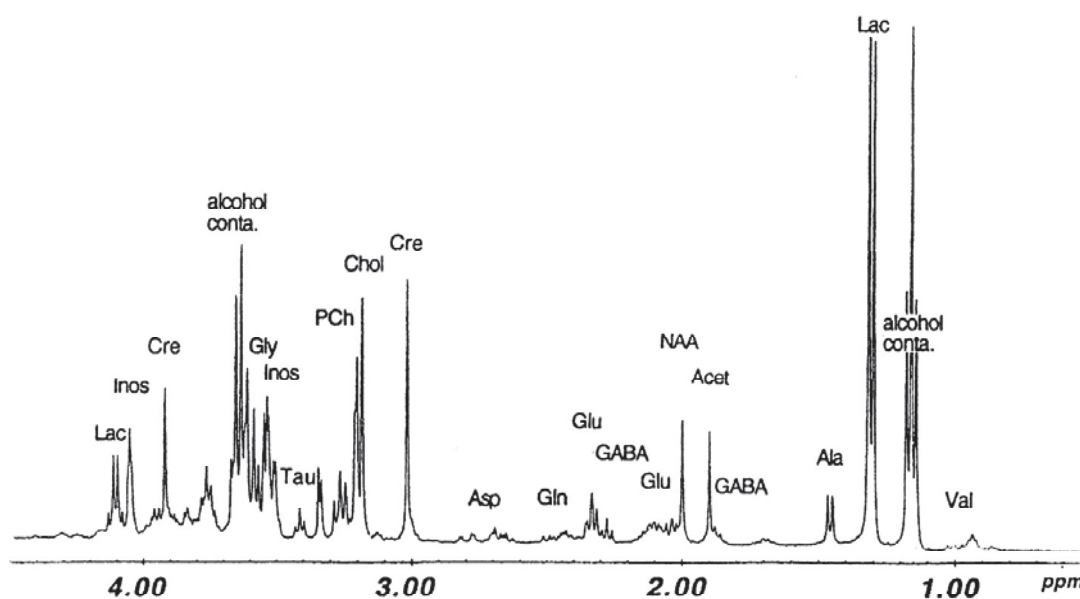


Figure 1.9: Proton MR spectra (400 MHz) of human brain tissue collected from superior temporal gyrus histologically determined to be mildly affected with Pick disease [25].

In this section, we will describe most of the metabolites, following the order in which they are presented in Fig. 1.9.

1.3.1 Lactate

Lactic acid is mostly present in the tissues as the ions $\text{CH}_3\text{CH}(\text{OH})\text{COO}^-$ of its salts – Lactates. Two well resolved peak groups are represented as a quadruplet at 4.1 ppm [26], corresponding to the signal of the CH_2 protons, and a doublet at 1.3 ppm [26], produced by the $^1\text{CH}_3$ proton group. In healthy organism lactic acid is mostly observed in muscular tissue, where it is produced in the hypoxia conditions, later being evacuated either by oxidation to pyruvate by well-oxygenated muscle cells, which is then directly used to fuel the Krebs cycle, either by Cori cycle, being converted to glucose via gluconeogenesis [26]. For brain tissues, markedly elevated lactate levels are common after severe head injury [27]. Acidoses of lactic acid as the consequences of fluid-percussion trauma were observed by Inao et al. in

[28]. Some studies represent Lactate as an important energy source for the human brain [29]

1.3.2 Inositol

In the particular case shown in Fig. 1.9 the ^2CH group of Myo-inositol represents the majority of the Inositol in the tissue. Inositol or cyclohexane-1,2,3,4,5,6-hexol is a cyclic compound $(-\text{CHOH})_6$, having nine possible stereoisomers. The most observed of them in brain tissues are myo- and scyllo-inositols. Commonly classified as a member of the vitamin B complex [30], it is discussed to be related to them because it is produced by the human body from glucose. In ^1H NMR spectroscopy it appears as the singlet of scyllo-inositol at 3.34 ppm and multiplets at 3.5 and 3.6 ppm with a recognized triplet at 4.05 ppm of myo-inositol [26].

1.3.3 Creatine

Creatine is one of the best-resolved in HR ^1H -NMR spectroscopy metabolites. Usually represented as tCr – total Creatine – which means the sum of the Creatine and Phosphocreatine. It has two singlets at 3.03 ppm and 3.92 ppm respectively. Concentrations in the human brain are approximately 4.0–5.5 mmol/kg of wet weight for Phosphocreatine, and 4.8–5.6 mmol/kg of wet weight for Creatine [31]. Because of its small variation in the case of brain pathologies, as well as due to the high SNR observed for their singlets, Cr and PCr are generally used as internal references for the estimation of the other metabolite concentrations.

1.3.4 Glycine

Glycine has the smallest molecular size of the twenty amino acids commonly found in proteins. In human brain it is normally biosynthesized through Serine from 3-Phosphoglyceric acid resulting in a concentration of approximately 1 mmol/kg of wet weight in humans [31]. It is reported to increase its concentration in tumor tissues and to have it significantly high in Glioblastomas, Ependynomas and Medulloblastomas [32].

1.3.5 Taurine

Taurine is seen as two neighbour triplets at 3.4 ppm and 3.25 ppm distorted due to the strong-coupling effects. It is one of the few known naturally occurring sulfonic acids. Taurine is remarkable, both for its high concentrations in animal tissues and for the variety of functions that have been ascribed to it. Its normal brain concentrations, observed in rat brain tissues are up to 5 mmol/kg [33]. High concentrations (up to 30 mM) in medulloblastomas have been reported [34].

1.3.6 Choline

Choline and its derivatives – Phosphoryl choline and Glycerophosphocholine – are observed as a large peak on low resolved spectra and can be recognized as independent singlets in the HRMAS spectra at 3.18 ppm, 3.20 ppm and 3.21 ppm respectively. Nowadays researches bind the Choline concentration incrementation with progressive cerebral glioma [35], medulloblastoma and adenoma [32] tumours. Normal concentrations are of the order 1-2 mmol/kg of wet weight and known to be nonuniformly distributed [36].

1.3.7 Glutamine

Glutamine (Gln) is the most abundant free amino acid in human blood [37]. Glutamine is synthesized by the enzyme glutamine synthetase from glutamate and ammonia. The most relevant glutamine-producing tissue is the muscle mass. Glutamine is also released, in small amounts, by the lung and the brain [38].

1.3.8 Glutamate

Glutamate (Glu) is often taken into account together with Glutamine, abbreviated as Glx in this case. As for the latter, Glutamate also has a role in protein synthesis, while the former is the most abundant excitatory neurotransmitter in the nervous system. Typical values for Gln and Glu are 12 mmol/kg and, respectively, 24 mmol/kg. Changes in the Gln/Glu ratio have been used as markers for cerebral ischemia, hepatic encephalopathy and Rett's syndrome [26].

1.3.9 GABA

γ -Aminobutyric acid (GABA) is a primary inhibitory neurotransmitter that is present in the brain at a concentration of approximately 1 mmol/kg. It is synthesized from glutamate using the enzyme L-glutamic acid decarboxylase and pyridoxal phosphate (which is the active form of vitamin B6) as a cofactor via a metabolic pathway called the GABA shunt. This process converts glutamate, the principal excitatory neurotransmitter, into the principal inhibitory neurotransmitter GABA. GABA synthesis is unique among neurotransmitters, having two separate isoforms of the rate-controlling enzyme, glutamic acid decarboxylase. The need for two separate genes on two chromosomes to control GABA synthesis is unexplained. Two metabolites of GABA are present in uniquely high concentrations in the human brain. Homocarnosine and pyrrolidinone have a major impact on GABA metabolism in the human brain. Both of these GABA metabolites have anticonvulsant properties and can have a major impact on cortical excitability [39].

1.3.10 N-acetyl-aspartate

N-acetyl-aspartate (NAA) is highly recognizable by the 2.01 ppm singlet, with other multiplets at 2.48 ppm, 2.67 ppm, 4.38 ppm and 7.82 ppm. Reported NAA

concentrations in the human brain are between 7-16 mmol/kg. NAAG is primarily detected via the 2.04 ppm resonance peak therefore appearing as a lineshape deformation of the 2.01 ppm NAA peak [40]. NAA is a concentrated, neuron-specific brain metabolite routinely used as a magnetic resonance spectroscopy marker for brain injury and disease [41]. In [42] NAA was presented as an acetyl source in the nervous system.

1.3.11 Ethanol

The spectrum presented in Fig. 1.9 has considerable peaks produced by ethanol molecules. Despite the fact that ethyl alcohol is not a metabolite of healthy brain in the normal state, it may appear rather often in spectra, obtained *ex vivo*, as a trace of biopsy procedure as it is in the present case [25].

1.3.12 Lipids

Lipids are present in the brain in their free form mainly due to cell and membrane degeneration and have been proposed as a marker for necrosis.

Because lipids have T_2 relaxation times much smaller than the main metabolites they produce very large spectral peaks, with the main contributions at 1.3 ppm and 0.9 ppm. Lipid concentrations are typically very small, but have been reported to drastically increase in the case of tumours. Another cause of high lipid peaks in MRS is subcutaneous fat tissue present in the skull, either because the tissue is part of the acquired voxel or because of signal contamination due to chemical shift artefacts [40].

1.4 Spectrum quantitation

Each chemical compound has its own “fingerprint” in the mixture of substances (see Fig. 1.10). Having knowledge about chemical shifts and spin-coupling data of a certain compound one can judge about its presence in the investigated material and, moreover, find out its relative concentration. This process is called quantitation of a spectrum.

There are different ways to segregate the quantitation methods. For example, methods are usually classified according to what input is used for the quantitation algorithm: time-domain (TD) (which is also the measurement domain) methods use the raw signal, frequency domain (FD) methods use a spectral estimate of the raw signal (mostly by DFT, while TD-FD approaches combine the two). Another classification, very common in MRS quantitation literature, concerns the presence of an underlying physical model in the quantitation algorithm [40]. Non-parametric methods, also called black-box methods, are based on a mathematical description of the signal, such as a decomposition by HSVD or FT. Due to the general possibility of complete or partial signal reconstruction from the mathematical decomposition, black-box methods have been mainly used in the recent years as preprocessing methods. Parametric methods, on the other hand, are based on

modelling of the acquisition process, and tend to give values directly related to physical measures.

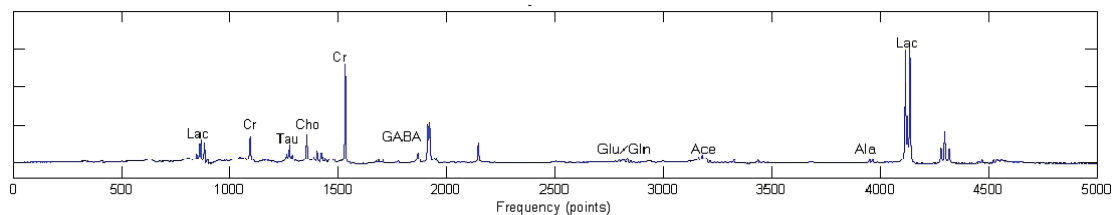


Figure 1.10: Region of interest of an HRMAS spectrum from a tissue sample of a human brain with an oligodendroglioma tumour, acquired at 11.7T.

1.4.1 Methods of quantitation

The need to monitor diseases and pharmaceutical follow-up requires automatic and reliable quantitation of HRMAS ^1H signals [17, 43–45]. Surely, nowadays the process of spectrum quantitation is well automatized [46–48]. But whichever algorithm is chosen it will need certain prior knowledge to be successful. For example, peak-finding algorithms [49, 50] need the ratios between the peaks on the respective positions. Others (e.g. [48, 51]), widely talked about in this work, use a basis-set of signals which needs to be simulated or acquired before the quantitation (e.g. Fig. 1.11)

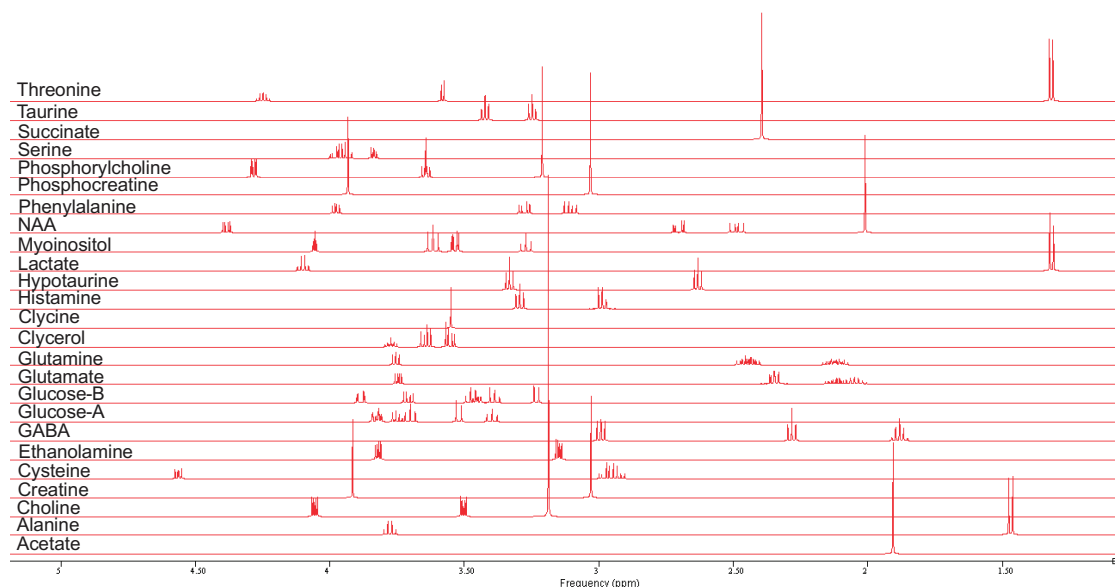


Figure 1.11: Typical basis-set of metabolite signals used for quantitation of HRMAS cerebral biopsy spectra (11.7T).

1.4.2 Quantitation with QUEST

Signal quantitation method QUEST was proposed by Ratiney et al. in 2005 [48]. In this method, a complex-valued time-domain model signal is written as a linear combination of the M weighted metabolite model \hat{x}_m of the basis-set:

$$\hat{x} = \exp(i\phi_0) \sum_{m=1}^M (c_m \hat{x}_m \exp((i\Delta\omega_m + \Delta\alpha_m)t) \exp(i\Delta\phi_m)) \quad (1.4.1)$$

c_m is proportional to the concentration of the metabolite m , $\Delta\alpha_m$, $\Delta\omega_m$, $\Delta\phi_m$ are small extra damping factors, angular frequency shifts and phase shifts enabling to automatically compensate for distortions due to the magnetic field heterogeneities with respect to the ideal signals of the metabolite basis-set. Passing these parameters through a non-linear least squares optimisation method the algorithm returns all found optimal values and the residue as a difference between the investigated signal and the sum of basis-set signals with respective found parameters.

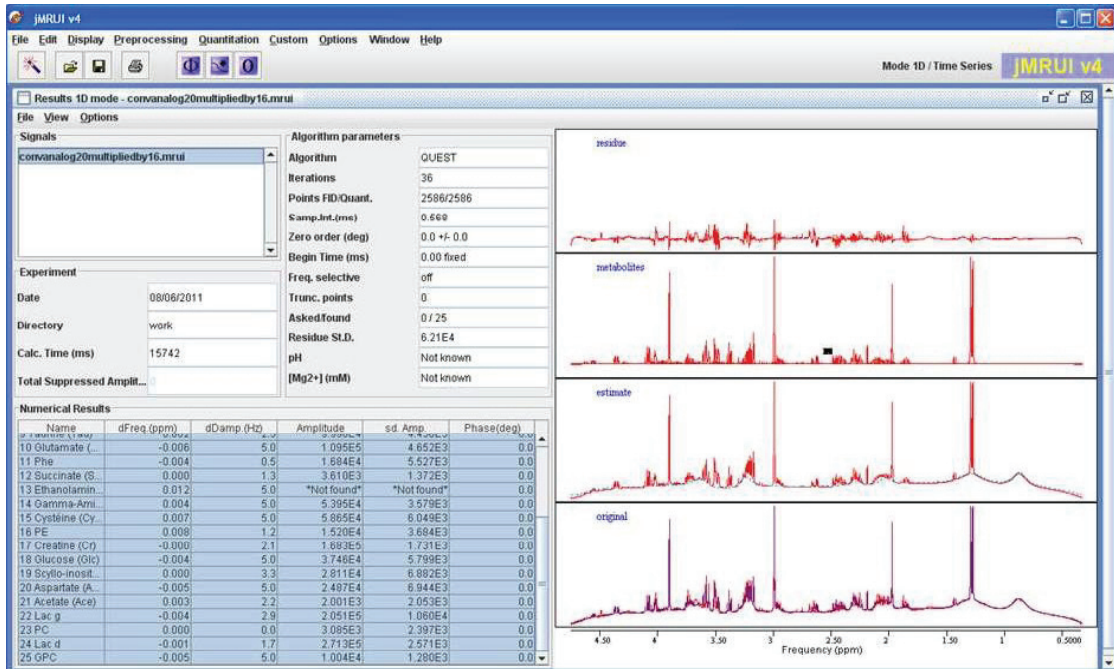


Figure 1.12: Example of a screenshot with results of quantitation obtained with the algorithm QUEST of jMRUI4 software package.

Nowadays [52] this method is a part of the NMR signal processing software package called jMRUI [53, 54].

It should be mentioned that in this case the extra signal damping is considered to have the exponential-decay shape, which is rather natural [55]:

$$f(x; x_0, \alpha) = \frac{1}{\pi} \left[\frac{\alpha}{(x - x_0)^2 + \alpha^2} \right], \quad (1.4.2)$$

where x_0 is the peak position (corresponds to the main frequency of the harmonic), and α is a scale parameter (corresponds to the linewidth of the mentioned harmonic).

1.5 Influence of micro-environment in tissues on quantitation

It is widely known that the exact values of the chemical shifts are not only the subject of the molecule configuration but also depend on the molecule's environment [56]. Its changes can significantly influence the electronic clouds' distribution and, due to this, affect the chemical shifts [57–59]. Thus, for example, the temperature dependence of certain peak positions enables temperature mapping on the images [52, 58, 60], the pH-dependence [56, 61, 62] gives a possibility to analyse acidity maps [63]. This dependence, which is a subject of interest of this work, can be simply demonstrated by providing several MR scans of the solution of the same substance at different pH values in the sample (similar works [61, 64]). A 0.1M solution of citric acid in D₂O was taken at five different fixed pH values. The pH of the sample was modified by adding a sodium hydroxyde D₂O solution. The acidity was measured by a voltmetric pH-meter HI9025 (pH-diapason 1.00..14.00, err.±0.01pH) calibrated by 1 point calibration on 7.01 pH buffer solution (HI 70007). The solution temperature was 24.9±0.1°C. The spectra were obtained on a Bruker-DRX300 spectrometer (300MHz). The dependence of the absolute peak positions is easily seen in Fig. 1.13. Also, the changes of the distances Δ between the peaks as a function of the pH-factor are evident. Moreover, the spectra show the redistribution of the peak intensities as a consequence of strongly-coupled protons. These data are presented in Tab. 1.4.

| | | | | | |
|-----------------|----------|----------|----------|----------|----------|
| pH | 9.47 | 6.21 | 5.00 | 3.40 | 2.26 |
| Amplitudes a.u. | 2:44 : 1 | 2.40 : 1 | 2.22 : 1 | 1.82 : 1 | 1.67 : 1 |
| Ratio | 2.44 | 2.4 | 2.21 | 1.82 | 1.66 |

Table 1.4: Relative values of peak's amplitudes in doublets of citric acid ¹H NMR spectrum respectively to environment acidity.

Environment influences the HRMAS spectra significantly and hampers the accurate quantitation. Basis sets simulated at certain conditions can be inappropriate for the quantitation of spectra obtained at other conditions [51, 65]. The spectra of the same tissue obtained at different conditions are presented in Fig. 1.14.

Whereas the temperature changes are usually easy controllable (e.g. the “temperature-at-acquisition” value is usually saved in the signal data file [66]), the pH value is rarely known. The acidity of the sample can differ accordingly to tissue of the sample, state of the tissue, temperature itself and others. This dependence can not be described in general case; however it can either be known from the prior knowledge (titration curves) or be predicted by quantum chemistry calculations.

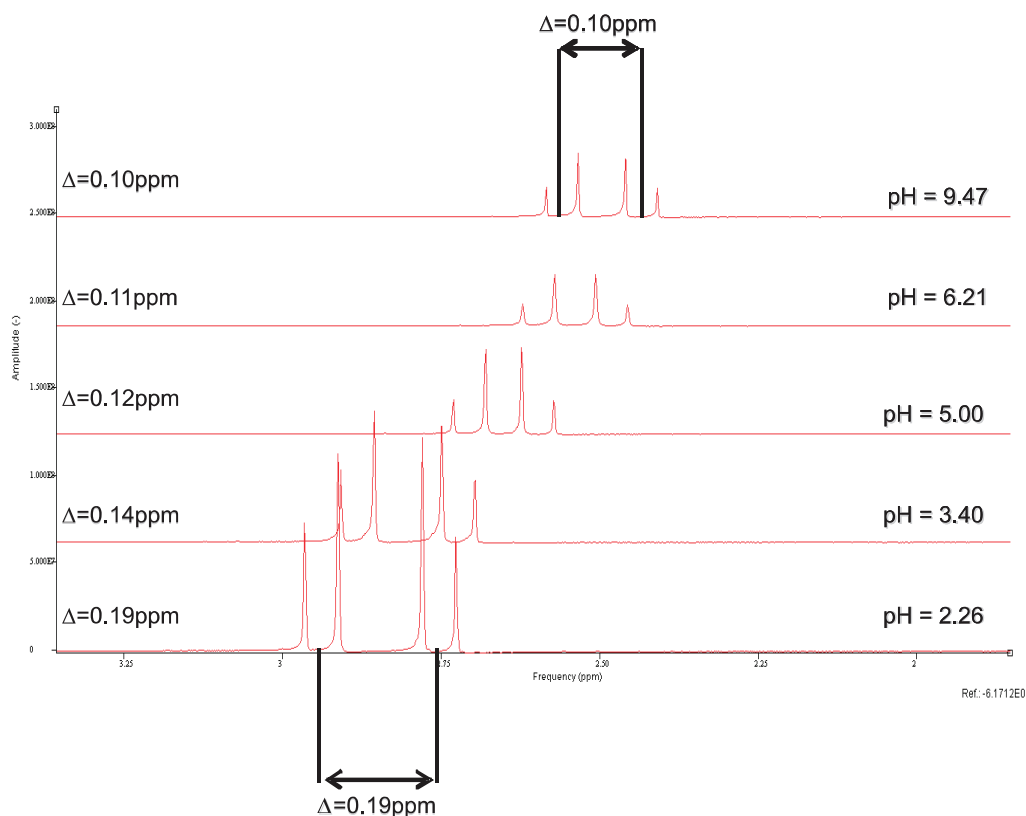


Figure 1.13: ^1H NMR spectrum of citric acid solution at different pH-factors.

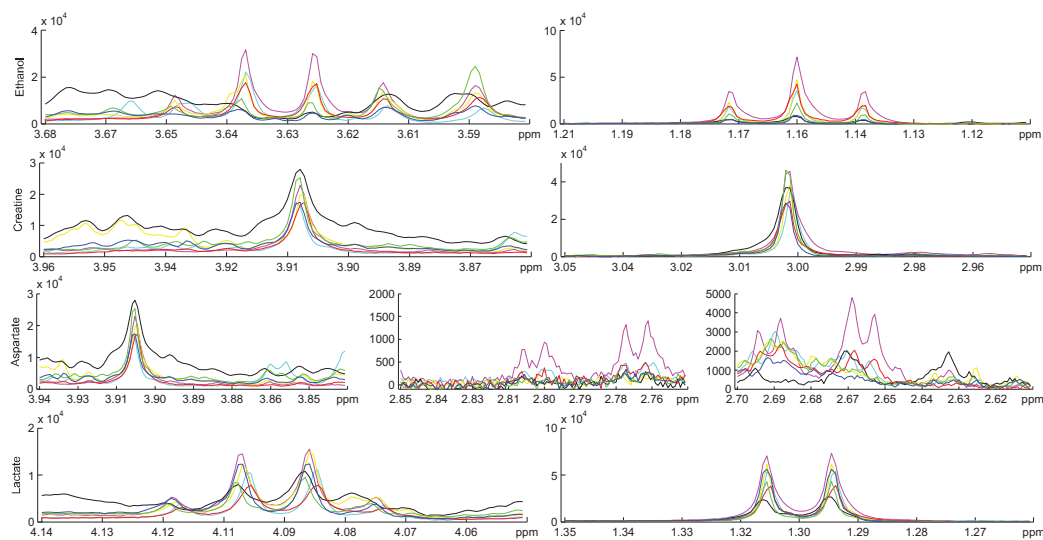


Figure 1.14: Zoom in the regions of interest of seven HRMAS NMR spectra of the brain tissue affected to oligodendroglioma obtained at same temperature conditions.

For example, the software package ACD Labs [67] enables prediction of NMR spec-

tra and even certain influences of the environment acidity can be provided such as absolutely negatively ionized, neutral or protonized molecule of the substance can be investigated (an example of HNMR-spectra for lactic acid estimation is presented in Fig. 1.15).

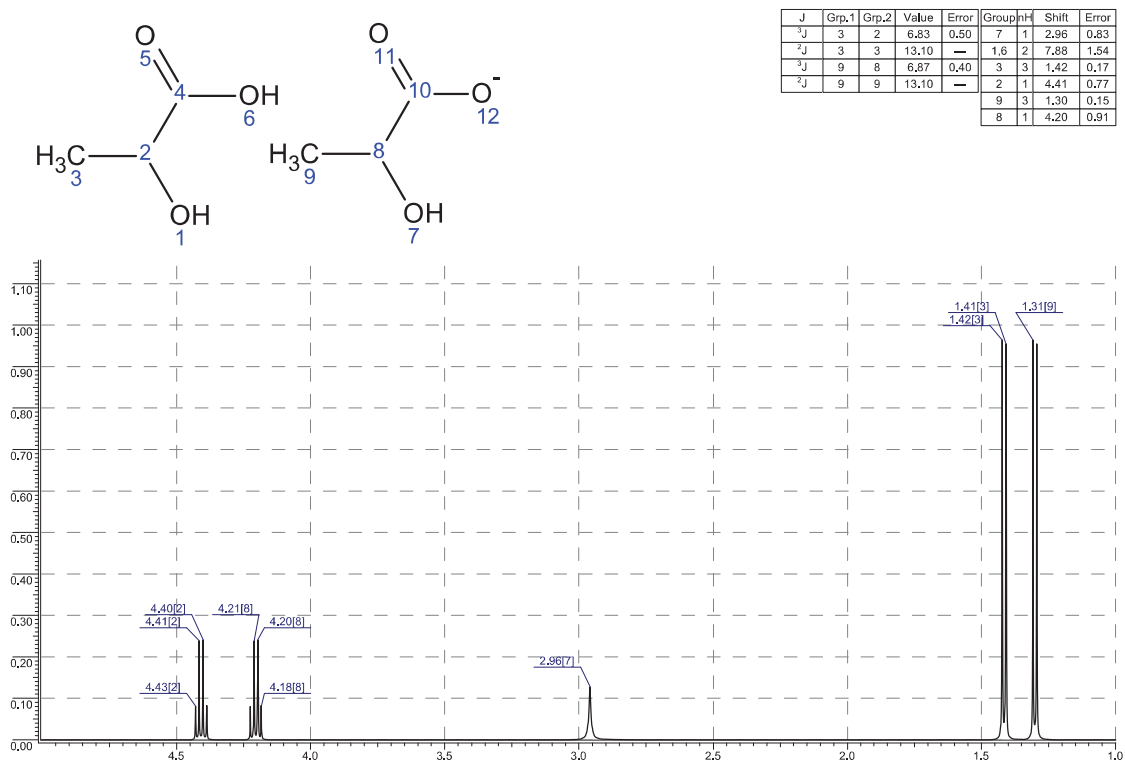


Figure 1.15: An example of ^1H NMR-spectrum estimation by ACD Labs HNMR-Pro [67] software as an overlap of the signals of a neutral and a deprotonized molecule of lactic acid at 500MHz.

Unfortunately, these investigations are not enough to predict peak positions with enough accuracy because the dependences are not linear [56] and their analytical functions are not known.

The common values for the peak position variations were obtained during the investigation of the series of spectra of human brain biopsies of oligodendroglioma tumors. Specific attention was paid to the Creatine (see Tab. 1.5) and Lactate peaks.

| # | Peak position | | Distance, ppm |
|---|---------------|-----------|---------------|
| | Right, ppm | Left, ppm | |
| 1 | 3.873 | 2.97 | 0.90304 |
| 2 | 3.9145 | 3.0088 | 0.90565 |
| 3 | 3.9116 | 3.0062 | 0.90539 |
| 4 | 3.9139 | 3.0081 | 0.90573 |
| 5 | 3.9196 | 3.0145 | 0.90513 |
| 6 | 3.922 | 3.016 | 0.90599 |
| 7 | 3.9218 | 3.0163 | 0.90547 |

| | | | |
|----|--------|--------|---------|
| 8 | 3.9233 | 3.0175 | 0.90573 |
| 9 | 3.924 | 3.0181 | 0.90591 |
| 10 | 3.9226 | 3.017 | 0.90556 |
| 11 | 3.9072 | 3.0021 | 0.90513 |
| 12 | 3.9167 | 3.0094 | 0.90721 |
| 13 | 3.9112 | 3.0059 | 0.9053 |
| 14 | 3.9169 | 3.0107 | 0.90626 |
| 15 | 3.9175 | 3.0117 | 0.90582 |
| 16 | 3.9114 | 3.0058 | 0.90565 |
| 17 | 3.922 | 3.0169 | 0.90513 |
| 18 | 3.9151 | 3.0097 | 0.90539 |
| 19 | 3.9245 | 3.019 | 0.90547 |
| 20 | 3.9239 | 3.0186 | 0.9053 |
| 21 | 3.9218 | 3.016 | 0.90573 |
| 22 | 3.9223 | 3.0167 | 0.90556 |
| 23 | 3.9207 | 3.0146 | 0.90617 |
| 24 | 3.9216 | 3.0164 | 0.90521 |
| 25 | 3.9236 | 3.018 | 0.90556 |
| 26 | 3.9169 | 3.0113 | 0.90565 |
| 27 | 3.9248 | 3.0191 | 0.90573 |
| 28 | 3.919 | 3.0141 | 0.90487 |
| 29 | 3.928 | 3.0225 | 0.90556 |
| 30 | 3.9168 | 3.0109 | 0.90591 |
| 31 | 3.9238 | 3.018 | 0.90582 |
| 32 | 3.9185 | 3.013 | 0.90547 |
| 33 | 3.919 | 3.0136 | 0.90539 |

Table 1.5: Creatine peak positions and distances between the peaks for a series of spectra from human brain biopsies with oligodendroglioma.

Chapter 2

Speed-up of NMR Quantum Mechanics simulations

2.1 Overview of the methods

The simulation procedure has significant computing times for the systems having a large number of spins. What is important, is that the computing time increases not linearly but exponentially as a function of the spin number N_{spins} :

$$t \propto 2^{2N_{\text{spins}}} \quad (2.1.1)$$

Certainly, the fact that independent simulations of n - and m -spin systems require less time than the simulation of $(n + m)$ -spin system raises the idea to find ways to split a spin system before the simulation of its NMR signal. Evidently, not all the molecules can be subject to such action. The majority of common brain metabolites has “dense” spectra (in the peak position sense) and must be simulated as entire spin systems. However some others present independent spin sub-structures (Creatine, Choline), either without ^1H -coupling between their elements (GPC), or with groups for which the weak-coupling approximation is applicable (Glucoses, Threonine). For such systems, there are some ways to optimize the signal simulation. This will be discussed in the following sections, starting from the simple decomposition and advancing to the uncoupling possibilities with a discussion about the gains of computing time, obtained by implementing these methods.

2.2 Decomposition

Analyzing different spin-systems, one can notice that in some molecular structures there are independent proton groups or even independent sub-systems. In this case, there is no necessity of putting all the spin-system in one Hamiltonian. One may just independently proceed the respective parts of the system into the simulation algorithm and simply add the output signals. A simple recursive algorithm has been developed for this procedure:

- setting initial *component-ID* for each group to 0,
- starting from the first spin, the component identifier binds all spins which have non-zero J-coupling value to the processed spin; then the procedure is recursively applied for all of mentioned spins (end condition for recursion: current spin is only coupled to the spins already marked as non-zero),
- repeating the previous step in the loop of counter *component-ID*,
- rearranging the processed system by constructing new spin systems, each one with the same *component-ID* value.

For example, for the phantom spin system, presented in Tab. 2.1, the algorithm separates it as it is presented in Tab. 2.2 and returns Tab. 2.3.

| σ , ppm | Multiplicity | J-coupling, Hz | | | | | |
|----------------|--------------|----------------|------|------|------|-----|------|
| 1.00 | 1 | 0.0 | 12.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 2.00 | 2 | 12.0 | 0.0 | 0.0 | 24.0 | 0.0 | 0.0 |
| 3.00 | 1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 36.0 |
| 4.00 | 2 | 0.0 | 24.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 5.00 | 1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 6.00 | 2 | 0.0 | 0.0 | 36.0 | 0.0 | 0.0 | 0.0 |

Table 2.1: A phantom spin structure with particularly coupled spins.

| σ , ppm | Multiplicity | J-coupling, Hz | | | | | |
|----------------|--------------|----------------|------|------|-----|------|------|
| 1.00 | 1 | 0.0 | 12.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 2.00 | 2 | 12.0 | 0.0 | 0.0 | 0.0 | 24.0 | 0.0 |
| 3.00 | 1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 36.0 |
| 4.00 | 2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 5.00 | 1 | 0.0 | 24.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 6.00 | 2 | 0.0 | 0.0 | 36.0 | 0.0 | 0.0 | 0.0 |

Table 2.2: The recursive assignment of the *component-ID* property to each spin. Here, the *component-ID* is presented as grey color intensity.

| σ ,ppm | M | J, Hz | | |
|---------------|---|-------|------|------|
| 1.00 | 1 | 0 | 12.0 | 0 |
| 2.00 | 2 | 12.0 | 0.0 | 24.0 |
| 5.00 | 1 | 0 | 24.0 | 0 |

| σ ,ppm | M | J, Hz | |
|---------------|---|-------|------|
| 3.00 | 2 | 0.0 | 36.0 |
| 6.00 | 1 | 36.0 | 0.0 |

| σ ,ppm | M | J, Hz |
|---------------|---|-------|
| 4.00 | 1 | 0.0 |

Table 2.3: Results of the decomposition procedure: three independent spin-systems.

This simple procedure is expected to have negligibly small time costs in execution compared to the simulation algorithm. So the expected computing time gains will be due to the decrease of the Hamiltonian size (in general, by a factor of two

for every spin detachment). One of the important features of this algorithm is its user independency. That means, firstly, that the small time cost enables to run it for all the spin systems to be passed to the simulation, secondly, the absence of any prior knowledge (initial values) needed for the algorithm to be executed. In the nearest future it could be implemented in a release of the jMRUI software package.

Unfortunately, it should be noticed, that in its first form, the algorithm of decomposition is applicable only to a few number of the common brain metabolites. Creatine, Choline, Phosphoryl Choline, Glycerophosphoryl Choline [26] are the more common ones (as brain metabolites) for which the algorithm provides computing time gains. It is especially true due to the fact that Choline moieties contain a nine-proton singlet which can not be separated in some implementations of Quantum Mechanics simulations. If we consider this possibility of splitting, we can obtain the time gains presented in Tab. 2.4. Despite the narrow field of

| Metabolite | Usual simulation (CA) | Decomposition + Simulation | Ratio |
|-----------------------|-----------------------|----------------------------|-------|
| Creatine | 0.219 s | 0.172 s | 0.785 |
| Choline | 0.250 s | 0.219 s | 0.876 |
| Phosphocholine | 0.266 s | 0.234 s | 0.951 |
| Glycerophosphocholine | 2765.3 s | 304.8 s | 0.110 |

Table 2.4: Computing time costs for simulations with and without using the proposed decomposition algorithm (CA: conventional approach).

its direct application, the method becomes important and absolutely necessary, when being applied to previously fragmented systems, discussed in the following section.

2.3 Splitting of spin-systems

2.3.1 “Buffering” effect

This section is devoted to the spin systems which have a specific property: a spin (or the spins, or the spin groups) in the “middle of the chain of coupled spins” can be considered as weakly coupled to each other. This is, for example, the case for Glycerophosphorylcholine, α - and β -Glucoses, Valine etc. [26] at high fields.

Let us step by step explain this specificity. First, we will consider the phantom spin structure presented in Tab. 2.5.

| | σ , ppm | M | J, Hz | | |
|---|----------------|---|-------|------|------|
| A | 3.00 | 1 | 0 | 23.0 | 0 |
| M | 2.00 | 3 | 23.0 | 0 | 12.0 |
| X | 1.00 | 1 | 0 | 12.0 | 0 |

Table 2.5: An AM_3X phantom parameters.

Important properties of the phantom are:

- to be an AMX,
- to have AM and MX spins coupled,
- to have A and X spins uncoupled.

All the parameters in the example given in Tab. 2.5 are subject to these conditions and are more or less real for a metabolite, investigated by ^1H NMR. Values of multiplicity are not principally important. In our investigation, we will keep the groups of 2ppm and 3ppm at fixed positions and the position of the remaining group (σ_X) will be variable (it is schematically marked by red color in Fig. 2.1). In the following explanation, we will give the names A, M and X for the groups from left to right (A – 3ppm, M – 2ppm, X – variable).

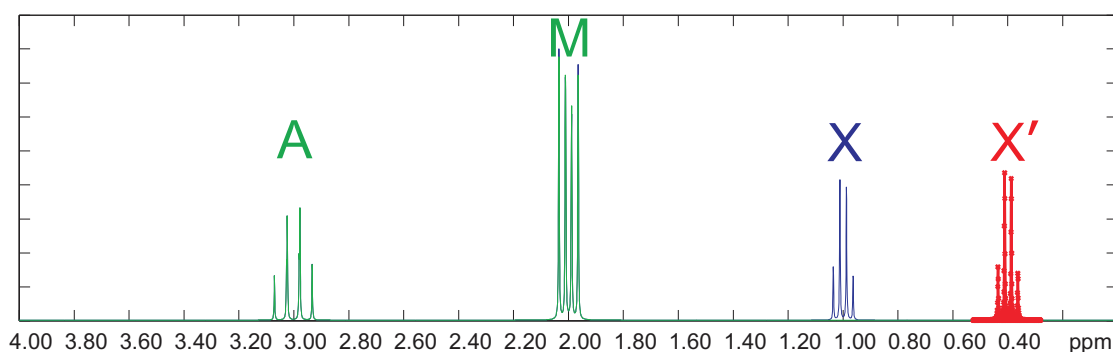


Figure 2.1: Spectra of the phantom corresponding to the spin parameters given in Tab. 2.5: blue – initial spectrum, red – group with variable position, green – spectrum of the system after displacement of the variable group.

Let us investigate the changes of intensity distribution in the multiplets of the fixed groups A and M (their “shapes”, in simple words) as a function of the position of the variable group X. The simulation was set to move the X multiplet along the frequency axis ($\sigma_X \in [0.4\text{ppm}..1.6\text{ppm}] \cap [2.4\text{ppm}..3.6\text{ppm}]$). The 0.4ppm area around the M-group is not taken under investigation to avoid AB structures. The spectral region of the M-multiplet was assumed to be $\pm 0.15\text{ppm}$ and the sum of absolute differences between the original and distorted spectra relatively to the integral area under the spectrum in the window has been taken as a measure of distortion.

The behaviour of the dependence is more or less as expected (Fig. 2.2). When the difference between the chemical shifts of M and X is small, the system has the behaviour of an AB-system, which significantly increases the relative error. When moving X away from the investigated group, the system becomes less and less coupled, decreasing the error. The zero-error point, observed on the red graph and absent on the blue one is absolutely normal, because it corresponds to the 1ppm position of the X-group which transforms the system to exactly match the initial one (Tab. 2.5).

There are no more important points in this investigation except the numerical values of the error. We emphasize that they need to be remembered to be compared with the analogous ones from the following simulation.

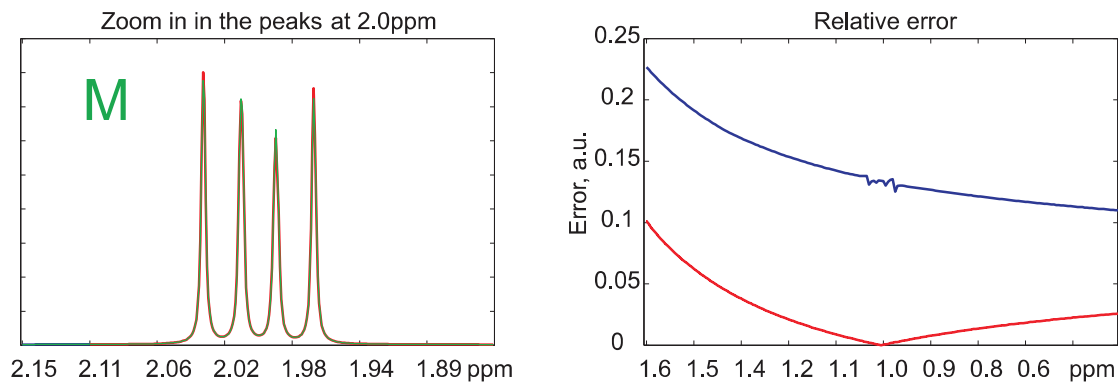


Figure 2.2: Left: original (green) and distorted (red) spectra in the surroundings of the 2ppm spectral group. Right: relative error between the spectra on left for the increasing (blue) and decreasing (red) displacement of the X-group position.

Now we move to the same investigation for the A-group. Remember that the A-spin is coupled to the second one (M) but has no proton-proton couplings with X. The shape of graph (Fig. 2.3) is rather familiar, with one exception: it is impossible to analyze the errors between the spectra when an additional group arrives in the window of interest (factually, the tentative to do that will lead to relative error of order $n/(n+m)$ where n and m are the multiplicities of the respective groups, which is much more than the investigated errors). So for the first analysis, the region around the +1ppm position has been cut off. The picture in other regions is already known. But the most important is the fact that, numerically, these errors are more than one order less than for the “buffering” group.

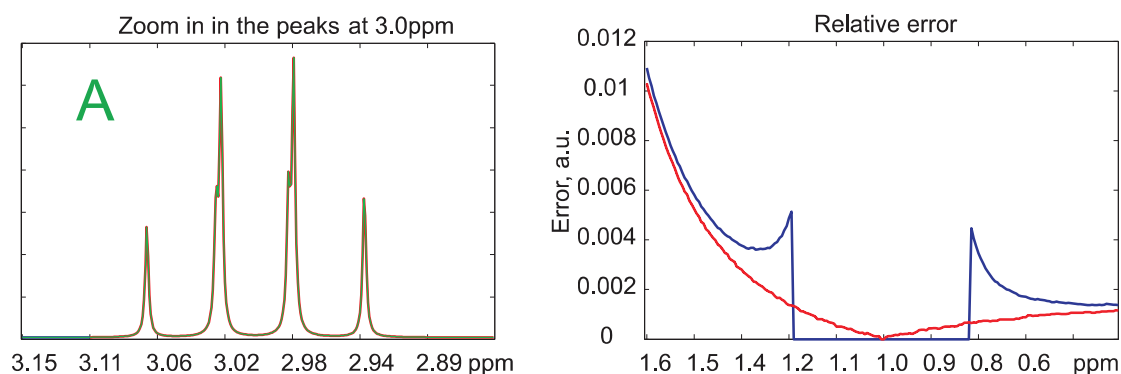


Figure 2.3: Left: original (green) and distorted (red) spectra in the surroundings of the 3ppm spectral group. Right: relative error between the spectra on left for the increasing (blue) and decreasing (red) displacement of the X-group position.

The intermediate conclusion is that *in strongly coupled spin-systems with consecutive coupling, the middle groups have a “buffering” effect not allowing the indirectly coupled spins to influence each other.*

2.3.2 Peak identification

It is not the same task that will be necessary for splitting algorithms (ch. “Optimization of the metabolites basis-sets”, sec. “Splitting of spectra”). For the following work, we need the exact information of the belonging of each frequency, returned by the simulation method, to a certain spin. This will be achieved in the following way. As it is known, the Hamiltonian is a function of two sets of data: chemical shifts and coupling constants. Furthermore, if the diagonal elements of the matrix are dependent on the both of them, the non-diagonal elements are either zeros, or linear combinations of the coupling terms only:

$$\mathcal{H} = \begin{pmatrix} f(\sigma, J) & 0 & 0 & \dots & 0 \\ 0 & f(\sigma, J) & f(J) & \dots & 0 \\ 0 & f(J) & f(\sigma, J) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & f(\sigma, J) \end{pmatrix} \quad (2.3.1)$$

For assumed AMX systems the values of non-diagonal elements are small relatively to the diagonal ones. This leads to the fact that if one considers the Hamiltonian for the same system but without couplings (Eq. 2.3.2) and runs a diagonalization procedure for both of them, then the *order of the eigenvalues corresponding to the respective diagonal elements will be the same in both data outputs.*

$$\mathcal{H} = \begin{pmatrix} f(\sigma, J) & 0 & 0 & \dots & 0 \\ 0 & f(\sigma, J) & 0 & \dots & 0 \\ 0 & 0 & f(\sigma, J) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & f(\sigma, J) \end{pmatrix} \quad (2.3.2)$$

Remembering, that the order of Hamiltonian construction is not random, but depends on the order of reading input data, one can infer the correspondence between eigenvalues and chemical shifts. That means, that each frequency, obtained as a difference between the respective eigenvalues, can be interpreted as

$$\nu_i = f(\sigma_i) + o(J_1..J_N), i = 1..N_{\text{freq}} \quad (2.3.3)$$

This enables to make a direct link between the frequency set and the input chemical shift data set. For example: input data are presented in Tab. 2.6, the algorithm output – in Tab. 2.7.

2.3.3 Pseudo-intensities

The pseudo-intensities are the additional data set to be introduced in the simulation procedure. By default, three data sets are needed to describe the behaviour of a spin system during a one pulse NMR experiment: chemical shifts, group multiplicities and coupling constants. Now, knowing at which group a peak belongs, it is helpful to introduce a parameter for the virtual spectral intensity of each group

| Resonance frequency = 1GHz | | | | |
|----------------------------|---|-------|------|------|
| σ ,ppm | M | J, Hz | | |
| 1.00 | 1 | 0 | 12.0 | 13.0 |
| 2.00 | 1 | 12.0 | 0 | 23.0 |
| 3.00 | 1 | 13.0 | 23.0 | 0 |

Table 2.6: An example spin system to be simulated by Quantum Mechanics with the conventional approach, at 1000MHz.

| Frequency | Amplitude | Group |
|-----------|-----------|-------|
| -3.0182 | 0.0002 | 3 |
| -2.0174 | 0.0003 | 2 |
| -1.0124 | 0.0003 | 1 |
| -1.9944 | 0.0002 | 2 |
| -0.9994 | 0.0003 | 1 |
| 0.0003 | 0.0000 | 0 |
| -2.9952 | 0.0003 | 3 |
| -2.0002 | 0.0000 | 2 |
| -1.0004 | 0.0002 | 1 |
| -4.0001 | 0.0000 | 0 |
| -3.0052 | 0.0002 | 3 |
| -2.0054 | 0.0003 | 2 |
| -0.9874 | 0.0002 | 1 |
| -1.9824 | 0.0002 | 2 |
| -2.9822 | 0.0003 | 3 |

Table 2.7: Output of the simulation procedure. The last column shows the spin group which produces the peak of the respective frequency.

called in the following “pseudo-intensity”. One should note, that for uncoupled spin systems this factor is identical to the multiplicity parameter (see the example given in Tab. 2.8). However, for the systems with coupled spins, multiplicity and pseudo-intensity have different physical meanings: multiplicity is an integer (strictly) and it changes the intensity distribution within the multiplets, while pseudo-intensity can be fractional and it corresponds to the total intensity of the spin group signal (see Tab. 2.9). In other words: *multiplicity of a group influences its signal but also splits the signals of neighbour groups; pseudo-intensity does only change the intensity of the signal of a certain group but it **does not split** the signals of neighbour groups into multiplets.* This is illustrated in Tab. 2.8 and Tab. 2.9.

In Tab. 2.8 an uncoupled spin system is presented. The vector of pseudo-intensities (denoted as “I”) is added into the simulation. One sees that changes of the multiplicities have the same effect as the changes of the pseudo-intensities – both spectra are the same. In Tab. 2.9, another spin system is presented. The chemical shifts are the same as in the system described in Tab. 2.8 but the spin groups are coupled ($J_{12} = 10\text{Hz}$). Changes of the multiplicities give an effect of

splitting (in the figure on the left, the multiplets in the spectrum are split in accurate correspondance to the A_3X_2 system) but the changes of pseudo-intensities influences the peak amplitude only (in the figure on the right the spectrum corresponds to the AX system).

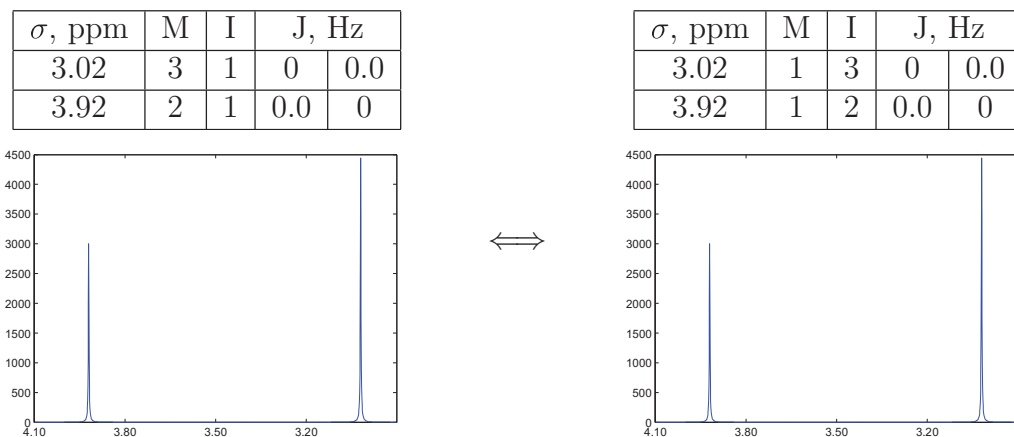


Table 2.8: Identity of pseudo-intensity and multiplicity for uncoupled spin systems.

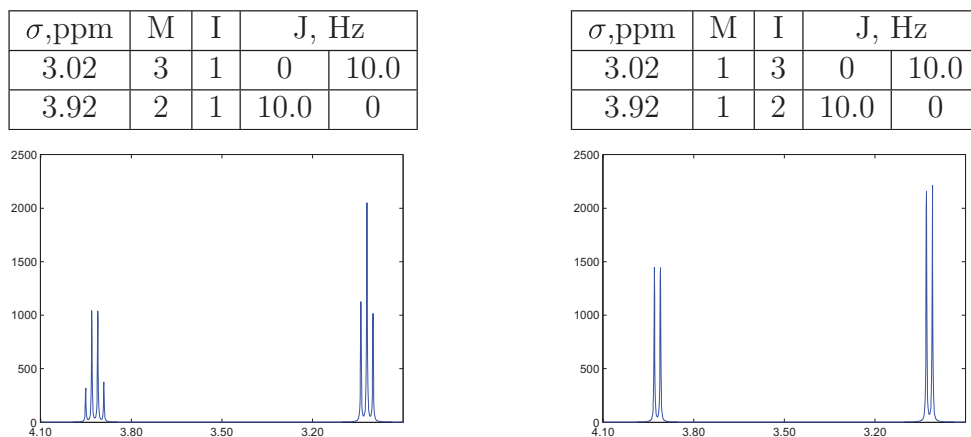


Table 2.9: Different physical meanings of the pseudo-intensity and multiplicity for coupled spin systems.

Both of these facts will be used in following algorithm implementation.

2.3.4 Fragmentation procedure. The kernel

Let us consider an hypothetical molecule with a CH-R chain. The chemical structure is totally unimportant for the following explanation. The important fact is the consecutive couplings between the set of nuclei in the molecule (Fig. 2.4, top):

$$J_{ij} = 0 \text{ for } |i - j| \neq 1 \quad (2.3.4)$$

Assuming the general case ($\sigma_1 \neq \sigma_2 \neq \sigma_3 \neq \sigma_4 \neq \sigma_5$, $J_{12} \neq J_{23} \neq J_{34} \neq J_{45}$) one can guess the spectrum of the molecule (Fig. 2.4, bottom).

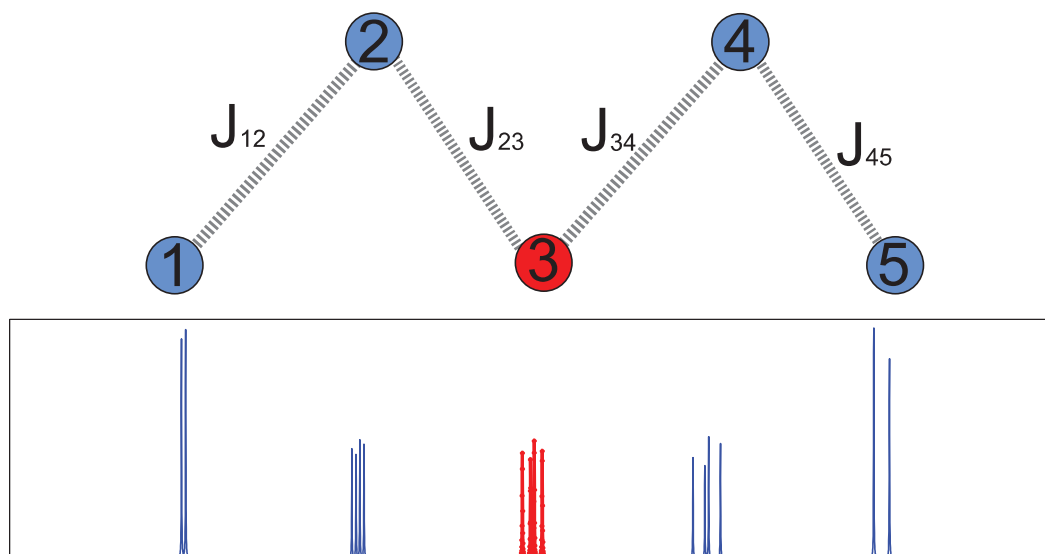


Figure 2.4: A coupled spin structure (chain) and its corresponding spectrum.

Specific attention will be paid to the third group marked in red. This is the chosen key point of the fragmentation procedure. Remember that (see subsection “Buffering effect”) due to its position in the structure it “buffers” the coupling effects, thus separating the molecule in two almost independent structures. All the peaks (and sets of frequencies and amplitudes) belonging to this group can be inferred.

Let us present the same data in the usual way (Tab. 2.10):

| Chemical Shift | Multiplicity | J-couplings | | | | |
|----------------|--------------|-------------|----------|----------|----------|----------|
| σ_1 | 1 | 0 | J_{12} | 0 | 0 | 0 |
| σ_2 | 1 | J_{12} | 0 | J_{23} | 0 | 0 |
| σ_3 | 1 | 0 | J_{23} | 0 | J_{34} | 0 |
| σ_4 | 1 | 0 | 0 | J_{34} | 0 | J_{45} |
| σ_5 | 1 | 0 | 0 | 0 | J_{45} | 0 |

Table 2.10: The coupled spin system shown in Fig. 2.4. σ -J-representation.

The kernel of the fragmentation procedure (which is always run relatively to a precise spin, so-called “key spin”) is to split the structure in three substructures:

- the “left” part – from the very beginning of the chain to the *key spin*,
- the middle part, consisting of the *key spin* and all the spins coupled with it,
- and the “right” part – from the *key spin* to the very end of the chain.

Schematically it is represented in Fig. 2.5. Firstly, the pseudo-intensity property, introduced in the previous subsection, will be used. Certainly, one must not have more than one contribution to the signal from the “copied” proton groups.

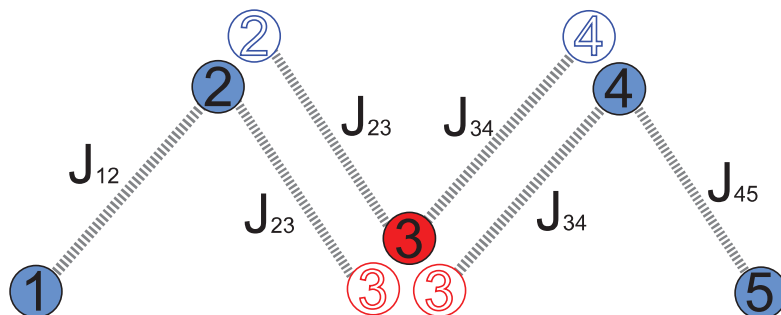


Figure 2.5: Representation of the coupled spins in the initial chain and its subchains.

In Fig. 2.5, these extra groups are repeated and are marked with white. Their pseudo-intensities were set to zero. Those groups, which are left with intensity of 1, conserve their colors.

Let us make some conclusions at this step:

- the total number of spins in the system is conserved,
- no any additional links were created,
- the couplings are not changed for all the spins that have non-zero intensity.

The table representation can be seen in Tab. 2.11.

| Chemical Shift | Multiplicity | Intensity | J-couplings | | | | | | | | | |
|----------------|--------------|-----------|-------------|----------|----------|----------|----------|----------|----------|----------|----------|---|
| σ_1 | 1 | 1 | 0 | J_{12} | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| σ_2 | 1 | 1 | J_{12} | 0 | 0 | 0 | 0 | J_{23} | 0 | 0 | 0 | 0 |
| σ_3 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | J_{23} | J_{34} | 0 |
| σ_4 | 1 | 1 | 0 | 0 | 0 | 0 | J_{45} | 0 | J_{34} | 0 | 0 | 0 |
| σ_5 | 1 | 1 | 0 | 0 | 0 | J_{45} | 0 | 0 | 0 | 0 | 0 | 0 |
| σ_3 | 1 | 0 | 0 | J_{23} | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| σ_3 | 1 | 0 | 0 | 0 | 0 | J_{34} | 0 | 0 | 0 | 0 | 0 | 0 |
| σ_2 | 1 | 0 | 0 | 0 | J_{23} | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| σ_4 | 1 | 0 | 0 | 0 | J_{34} | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Table 2.11: Fragmented spin system, ready to be decomposed.

The under-diagonal part of the coupling matrix is never considered due to the symmetry. Here it is presented filled for reader's convenience. The row colors are introduced to show the independent groups for future decomposition. Notice that the last two rows in the list above are the most important. As it was shown in Tab. 2.9, changing the pseudo-intensity of the group does not change the intensity distribution within the multiplet. That means, for this example, that the spins 2, 3 and 4 still behave as being coupled to their respective neighbours, despite the zero-intensity peaks, produced by some of them.

| CS | M | I | J | | |
|------------|---|---|----------|----------|----------|
| σ_1 | 1 | 1 | 0 | J_{12} | 0 |
| σ_2 | 1 | 1 | J_{12} | 0 | J_{23} |
| σ_3 | 1 | 0 | 0 | J_{23} | 0 |

| CS | M | I | J | | |
|------------|---|---|----------|----------|----------|
| σ_3 | 1 | 1 | 0 | J_{32} | J_{34} |
| σ_2 | 1 | 0 | J_{12} | 0 | 0 |
| σ_4 | 1 | 0 | J_{34} | 0 | 0 |

| CS | M | I | J | | |
|------------|---|---|----------|----------|----------|
| σ_4 | 1 | 1 | 0 | J_{45} | J_{34} |
| σ_5 | 1 | 1 | J_{45} | 0 | 0 |
| σ_3 | 1 | 0 | J_{34} | 0 | 0 |

Table 2.12: Fragmented and decomposed spin system to be passed to a Quantum Mechanical simulation procedure (Here, “CS” means “chemical shifts”).

The last step before the simulation is to recognize the system as three independent spin systems which is done by the decomposition algorithm. The result is presented in Tab. 2.12.

The computing time gains, even for this simple example, are evident: the Hamiltonian size for the coupled structure is 2^5 , so the matrix has $32 \times 32 = 1\text{K}$ elements and the simulation algorithm provides a signal comprising 210 frequencies, while after the fragmentation procedure three 8×8 matrices (192 elements each) return 55 frequencies (165 in total), many of them have zero intensity.

2.4 Results

This method provides advantages for more than four-spin systems. Simply, because the minimum possible output of the procedure is three systems of 2-3-2 spins, which, in time-costs resources, are equivalent to $2^4 + 2^6 + 2^4 = 96$ Hamiltonian elements plus $4 + 14 + 4 = 22$ frequencies. The four-spin general Hamiltonian has $16 \times 16 = 256$ Hamiltonian elements and produces 55 frequencies. The comparative computing time costs are presented on p. 82.

For the five-spin systems computing time gains are evident, so let us consider the accuracy of the simulation. In all the following simulations *all* the frequency-amplitude pairs returned by the simulation method are considered for the signal construction.

2.4.1 Phantoms

First tests are set on phantoms. The general conditions were to mimic an HRMAS experiment. A spin system is presented in Tab. 2.13. The transmitter frequency was chosen to 500Mhz. The apodization factor was set to 4Hz.

| | Chemical Shift, ppm | Multiplicity | J-couplings, Hz | | | | |
|-----|---------------------|--------------|-----------------|---|---|---|---|
| | 1 | 1 | 0 | 2 | 0 | 0 | 0 |
| A | 2 | 1 | 2 | 0 | 4 | 0 | 0 |
| B,M | 3 | 1 | 0 | 4 | 0 | 6 | 0 |
| X | 4 | 1 | 0 | 0 | 6 | 0 | 8 |
| | 5 | 1 | 0 | 0 | 0 | 8 | 0 |

Table 2.13: A coupled spin system. σ -J-representation.

AMX spin system

The third spin has a frequency variable from 2.2 ppm to 3.8 ppm (marked with red in Fig. 2.6). The goal of the simulation was to compute the relative errors and correlation changes between the exact simulated signal and the signal, simulated after fragmentation of the spin system at $KeySpin = 3$ with respect to the relative group positions. Results can be seen in Fig. 2.6.

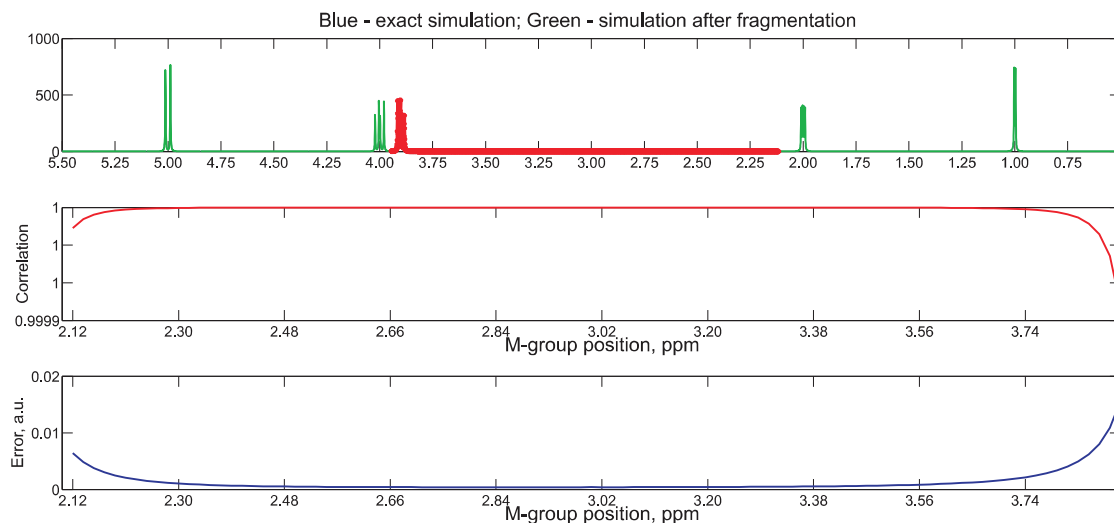


Figure 2.6: Error dependence for the AMX spectrum as a function of the M-group position.

The graphs are not symmetrical with respect to the point at 3 ppm, as expected. Despite the equal differences between the chemical shifts of the groups, their coupling constants are not equal, which makes the 4 ppm group showing an AB-behaviour when the moving group approaches 4 ppm. Important are the numerical values, which show negligibly small differences between the signals simulated with and without fragmentation. The relative error does not exceed 0.05% in a 0.8ppm interval. Zoom in in the [2.5ppm..3.3ppm] interval is presented in Fig. 2.7.

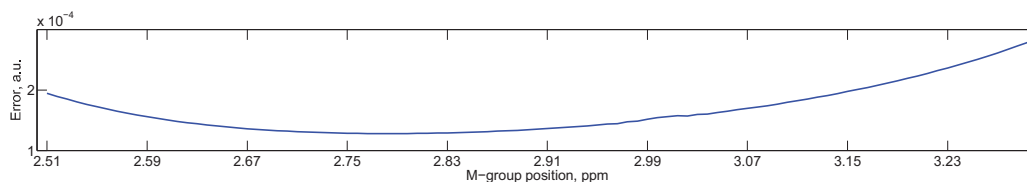


Figure 2.7: Zoom in in the [2.5ppm..3.3ppm] interval of the Fig. 2.6.

ABX spin system

The next investigation is the influence of the fragmentation procedure on the quality of simulation of an ABX spin system, fragmented at A point. For this, the

spin system from Tab. 2.13 was simulated at different (from 3.8ppm to 4.2ppm) positions of the third group with a fragmentation at the second one. Results are presented in Fig. 2.8.

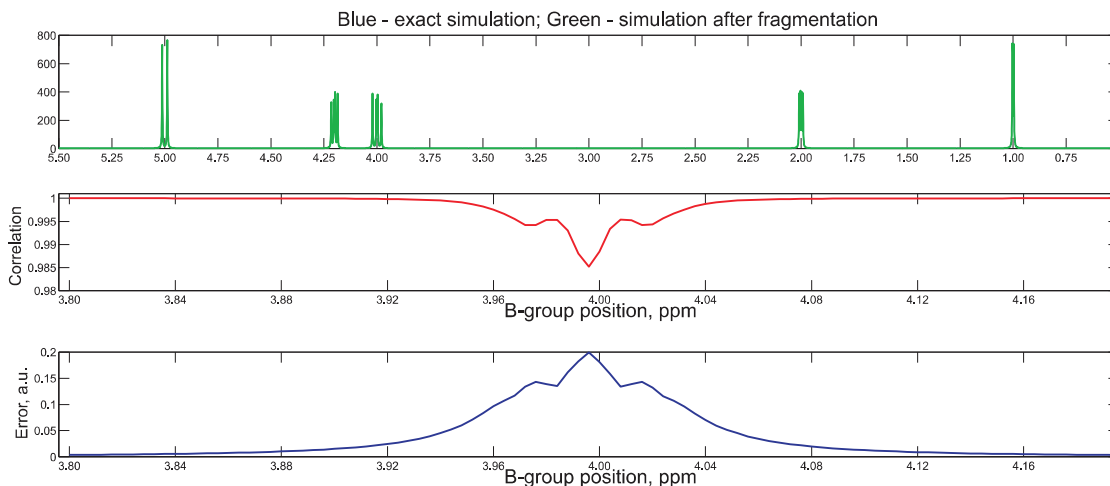


Figure 2.8: Simulation error dependence for the ABX spectrum as a function of the B group position. Top: blue – spectrum simulated with QM conventional approach, green – spectrum simulated after the spin system fragmentation; middle: correlation between the two simulated signals as a function of the B group position; bottom – error obtained as the sum of squares of differences between the two simulated signals as a function of the B group position.

The relative error is below one percent for the distances between the groups larger than 0.1ppm (see Fig. 2.9).

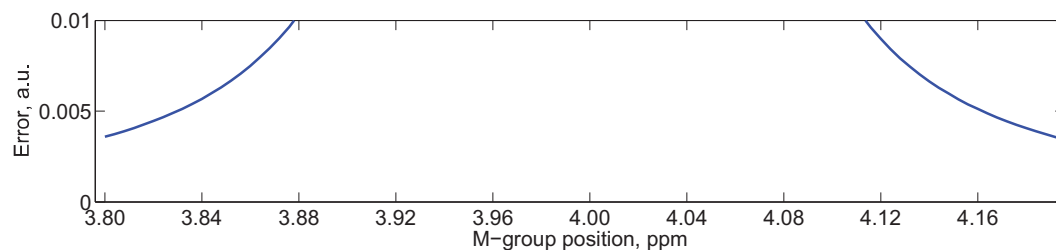


Figure 2.9: Zoom in in the [0..1%] interval of the relative simulation error given in the Fig. 2.8.

2.4.2 Metabolites

Our fragmentation method is then applied to spectrum simulation of metabolite spin-systems that are the subject of this work. Spectra were simulated twice: with and without fragmentation using the same Quantum Mechanics method. Initially all the simulations were done using the chemical shifts and J-coupling data from [26] at a magnetic field strength of 11.7T. For each simulation the computing time was recorded. Results are presented in Tab. 2.14, where the column “1-Corr”

represents the difference between 1 and the cross correlation of the two signals which, in its turn, is:

$$\text{Corr}(S_1, S_2) = \sum_{i=1}^N S_1^*(i) S_2(i). \quad (2.4.1)$$

| Metabolite | $\frac{\sum S_1(i) - S_2(i) }{\sum S_1(i) }$ | $1 - \text{Corr}$ | $t_1, \text{ s}$ | $t_2, \text{ s}$ | Ratio |
|--------------|--|-------------------|------------------|------------------|-------|
| Cho | 4.723e-013 | 2.220e-016 | 0.453 | 0.172 | 2.632 |
| PhC | 3.391e-013 | 0.000e+000 | 0.359 | 0.157 | 2.288 |
| Gly | 5.898e-002 | 5.299e-004 | 0.421 | 0.797 | 0.528 |
| Myo | 5.359e-002 | 4.718e-004 | 1.610 | 0.390 | 4.132 |
| Val | 5.881e-004 | 8.077e-008 | 39.78 | 44.95 | 0.885 |
| α -Gl | 6.172e-002 | 8.316e-004 | 6.703 | 0.359 | 18.52 |
| β -Gl | 5.626e-002 | 5.696e-004 | 6.594 | 0.328 | 20.10 |
| GPC | 1.477e-001 | 2.366e-003 | 2879 | 11.51 | 250.1 |

Table 2.14: Comparison between accuracies and computing time costs for the simulation of main brain metabolites. t_1 corresponds to the exact simulation with the conventional Quantum Mechanics approach; t_2 – to the simulation with the fragmentation procedure.

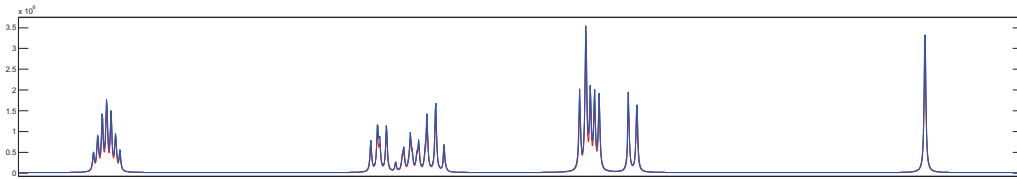


Figure 2.10: Glycero Phosphoryl Choline spectra: red – simulated with the conventional approach (2800 s), blue – simulated with the conventional approach after fragmentation (12 s).

The region of interest of GPC-spectrum (chosen as the worst in terms of accuracy) is presented in Fig. 2.10. Analyzing the data in the Tab. 2.14 one sees some points of attention. In the first two lines the errors between the simulations seem to be quite small. It is due to a specific case of the procedure implementation. Choline and phosphocholine have no spins convenient for the fragmentation. That means, that the fragmentation procedure executes only its last step – decomposition. As we already know, these metabolites are decomposable, so the results, obtained here, are more or less the same that the results seen in Tab. 2.4. Numerically they correspond to two different algorithm implementations used in the first and the present cases.

Other interesting results are those, where despite of using fragmentation methods, we observe an increase of computing time costs. The reason of this is masked in the scheme of coupling bounds in the structure. For example, Glycerol decomposition is schematically presented in Fig. 2.11.

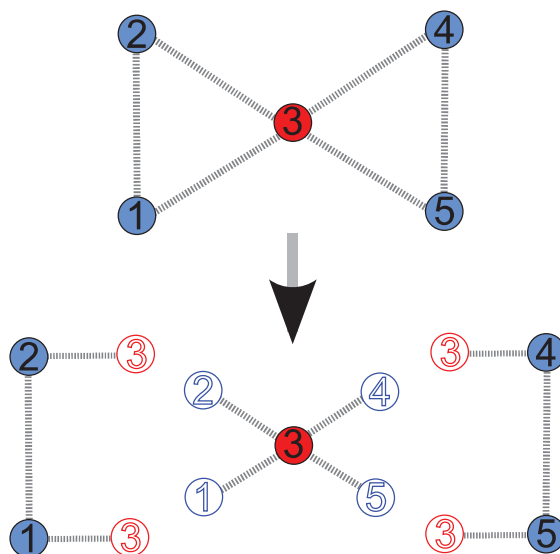


Figure 2.11: Glycerol fragmentation.

Execution of the fragmentation procedure for the Glycerol spin system did not decrease the number of spins in the largest structure. In other words, we added two new three-spin structures without gaining anything. Another time consuming aspect of the same origin is observed with Valine. So, the implementation of the fragmentation on these systems was totally senseless.

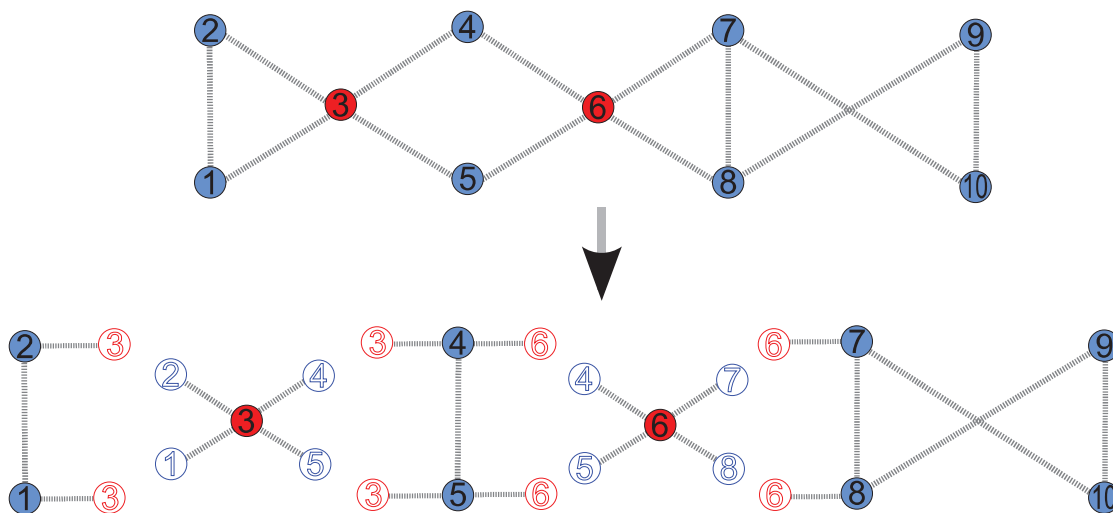


Figure 2.12: GlyceroPhosphoCholine spin system fragmentation.

Despite this Glycerol example, the fragmentation of GlyceroPhosphoCholine is still profitable. This can be understood after taking a look at Fig. 2.12. The necessity of fragmentation at point “6” is clear: the 10-spin structure (≈ 1 million Hamiltonian elements) is transformed into three structures with a maximum of 7 spins. But the first of them should be fragmented too (at point “3”) to decrease the maximum spin number to 6 (4K elements). This last step is not absolutely necessary, because it is an issue of quality/speed ratio. Values of errors and com-

puting time gain for the last fragmentation are presented in Tab. 2.14. In case of the fragmentation only at point “6” one obtains 14.9s as the simulation time, $1.369 \cdot 10^{-1}$ as the integral relative error, and $(1 - 5.45 \cdot 10^{-5})$ as the correlation value.

Choosing from the list in Tab. 2.14 the metabolites, for which the method is appropriate (Myoinositol, α -Glucose, β -Glucose) we can estimate the error dependence on the B_0 field. Respective graphs are represented in Fig. 2.13.

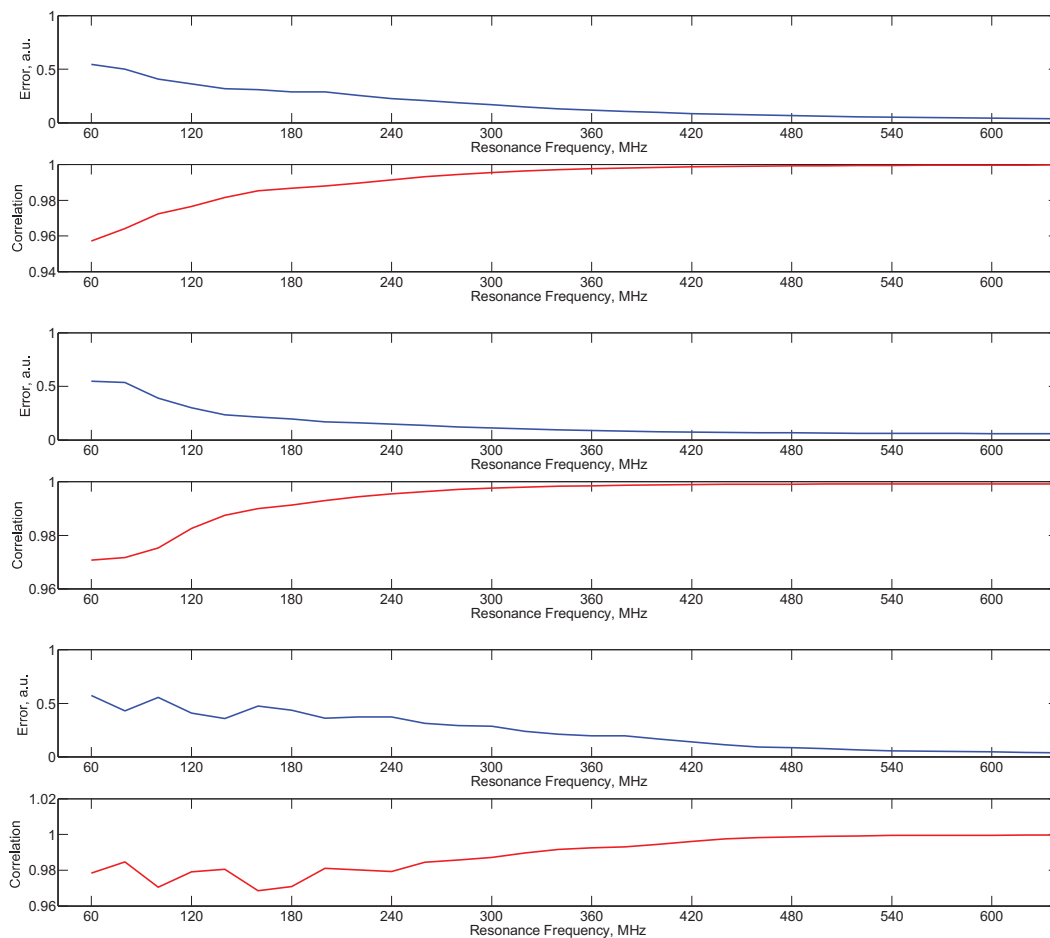


Figure 2.13: Errors and correlations, induced by the fragmentation method, as a function of the transmitter frequency.

Myoinositol is a cyclic compound (as in chemical sense as in coupling one). This means that fragmentation must be done in more than one point (otherwise one will only cut the ring but increase the total number of coupled spins). It is rather logic to choose the *KeySpins* at opposite sides of the ring, and for the provided simulations the most distanced (in spectral sense) groups at 4.04ppm and 3.26ppm were chosen.

For both Glucoses, the choice of the fragmentation point on the triplet at 3.38ppm (4th group selon Govindaraju et al.[26] interpretation) is natural.

2.4.3 Implementation details

The method was firstly implemented as MATLAB code, including all the steps and basing the simulation on a matrix multiplication procedure (creation of the Hamiltonian from the Pauli matrices, calculation of the transition operators). This implementation enables to work out the pseudo-intensity parameters of the spins, but is slow in comparison to the compiled programs.

The conventional approach simulation, done by A.-R.Allouche as a part of his Gabedit software [68], implemented in C, was used initially to do the fragmentation. But this implementation does not return the data of peak belongings to the groups. For the following analysis, use of MATLAB matrix realization is suggested.

It must be noticed that both described simulations are dedicated to get the signals as a response of the one pulse NMR sequence. But the simulations of NMR signals for more complicated sequences are also possible. One of the softwares, enabling to proceed different pulse sequences, has been mentioned already and is part of jMRUI package [52].

It is interesting that the current implementation of simulation by the Baker-Campbell-Hausdoff (BCH) formalism [69], implemented in jMRUI, has a vector of multiplicity values for a simulated spin system. In fact, the definition used in NMR-SCOPE corresponds to the pseudo-intensity previously described.

Chapter 3

Optimization of metabolite basis-sets based on Signal Processing

3.1 Introduction

The problem caused by chemical shift dependence on the environment being obvious, we will present in details the different methods of corrections of mismatches between the investigated spectrum and the spectra from the basis sets used for quantitation. Two simple methods – stretching/shrinking and splitting – will be first discussed. Their advantages and limitations will be established and discussed. It will be shown that using any of these methods leads to some spectrum shape distortions and corrupts the correct metabolite fingerprints. In chapter 4, a new method, based on the Quantum Mechanics conventional approach will be introduced. The latter method, despite the important computation time costs, is the only method which respects the correct metabolite fingerprints.

3.2 Stretching/Shrinking

If one needs to scale the time representation of a signal by a factor a , one needs to scale its spectral representation with the reciprocal factor $1/a$:

$$f(at) \quad \Rightarrow \quad \frac{1}{|a|} F(\omega/a) \quad (3.2.1)$$

where f and F signifies the function in the time domain and its Fourier transform respectively.

This time scaling property of the Fourier transform leads us to the idea to correct the spectrum mismatches by resimulation of the simulated basis-set spectra, used for quantitation, with another sampling interval. However, if the metabolite basis-set spectra, used for quantitation, are not simulated but acquired, this action is not possible anymore. To solve the problem of basis set optimization in this case, the stretching/shrinking method, based on the artificial signal resampling,

was introduced. Stretching/shrinking was proposed as a simply way to correct for the mismatches between the peak positions of the signal to be quantified and the metabolite basis-set signals. Starting from the method QUEST presented in (1.4.1), we propose to modify the metabolite basis-set signals \hat{x}^m , sensitive to pH changes, by just slightly stretching/shrinking (expansion/contraction) the corresponding spectrum. This is justified by the fact that to first order approximation a change of pH can produce a linear change of the width of the spectrum of a metabolite of an AB and AX system. Stretching/shrinking can easily be done by introducing a frequency scaling parameter κ_m and by directly using the inverse scale change of the \hat{x}_k^m time-domain signals (frequency scale expansion results in contraction of the time scale and vice versa). This avoids splitting of the metabolite basis-set signals of given metabolites into basis sub-components according to chemical groups and adding appropriate constraints (prior knowledge) to the parameters of the groups. Either κ_m can be estimated prior to quantitation by maximizing the correlation between the *ex vivo* signal and \hat{x}^m or by introducing κ_m in the quantitation procedure. In the Fig. 3.1 a schematic image of stretching for lactate spectrum is presented. In the example the extremely (pH=1) distorted (respectively to the normal one, pH \approx 7) Lactate spectrum was matched to the positions of the respective peaks of a human oligodendrogloma biopsy spectrum.

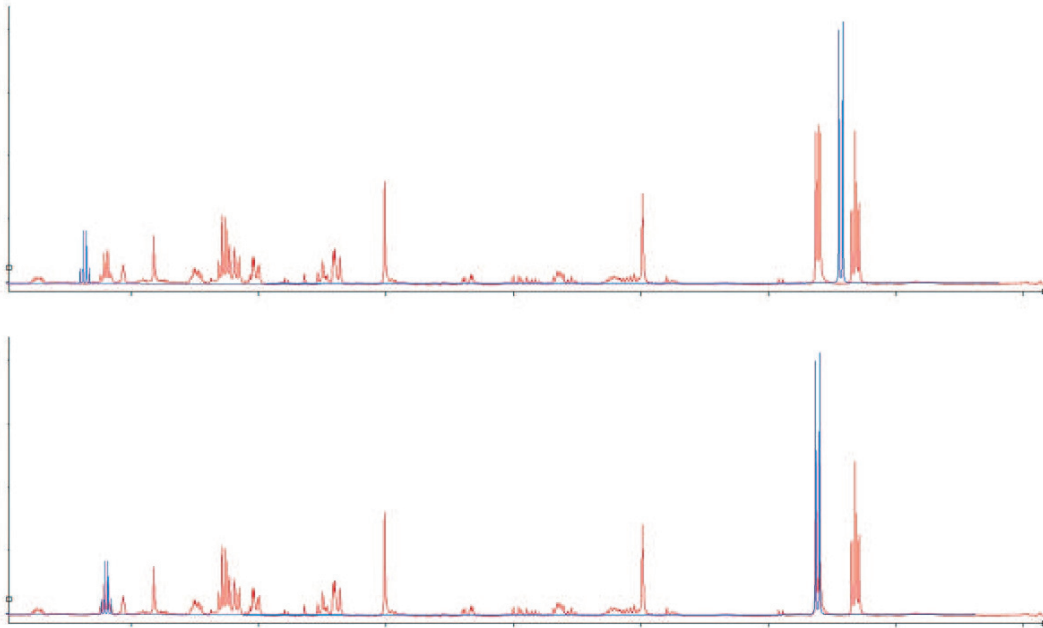


Figure 3.1: Matching of a simulated spectrum of lactic acid (blue) to the respective peaks in an acquired spectrum of a human oligodendrogloma biopsy spectrum (red).

Stretching/shrinking is well suited to improve quantitation of AX spin systems like lactate, Creatine, Ethanolamine, *etc.*. But, despite the simplicity of the method, there are some points of attention in its implementation which will be

discussed below.

3.2.1 Interpolation

To implement the stretching method, numerous out-of-sample values of the signal are necessary. The necessity of the approximation forces us to use a method of interpolation. Three different algorithms of interpolation were investigated and tested to choose the optimal one for the implementation in the stretching algorithm:

- Linear interpolation
- Cubic-spline interpolation
- FFT-interpolation

Linear interpolation

Linear interpolation [70] is the simplest (after the nearest-neighbour interpolation) method of interpolation. Using

$$f(x) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0) \quad (3.2.2)$$

the error is bounded by

$$|E_T| \leq \frac{(x_1 - x_0)^2}{8} \max_{x_0 \leq x \leq x_1} |f''(x)| \quad (3.2.3)$$

which shows the major dependence on the expected second derivative of the interpolated function which, in turn, makes clear the fact that the more “curved” the function is, the worse are the approximations made with simple linear interpolation.

Cubic-spline interpolation

Interpolation using natural cubic spline [71, 72] can be defined as

$$S_i(x) = \frac{z_{i+1}(x - x_i)^3 + z_i(x_{i+1} - x)^3}{6h_i} + \left(\frac{y_{i+1}}{h_i} - \frac{h_i}{6}z_{i+1}\right)(x - x_i) + \left(\frac{y_i}{h_i} - \frac{h_i}{6}z_i\right)(x_{i+1} - x) \quad (3.2.4)$$

where

$$h_i = x_{i+1} - x_i \quad (3.2.5)$$

The coefficients z_i can be found by solving this system of equations:

$$\begin{cases} z_0 = 0 \\ h_{i-1}z_{i-1} + 2(h_{i-1} + h_i)z_i + h_i z_{i+1} = 6 \left(\frac{y_{i+1} - y_i}{h_i} - \frac{y_i - y_{i-1}}{h_{i-1}} \right), i = 1, \dots, n-1 \\ z_n = 0 \end{cases} \quad (3.2.6)$$

The error for this interpolation depends on the fourth derivative of the interpolated function and could be bounded as

$$|E(t)| \leq \frac{5}{384} \|f^{(IV)}\| h^4 \quad (3.2.7)$$

for the natural spline interpolation and has the same order but five times less

$$|E(t)| \leq \frac{1}{384} \|f^{(IV)}\| h^4 \quad (3.2.8)$$

for cubic Hermite splines [73], where $\|f\|$ is the infinity norm of a function $f : \mathbb{R} \rightarrow \mathbb{R}$ over an interval $[a, b]$ is defined as

$$\|f\|_{\infty, [a, b]} = \max_{i \in [a, b]} |f(i)| \quad (3.2.9)$$

But, as it was presented by Ron [74], even with the assumption that evaluating the derivative is at least as costly as evaluating the function, we obtain that finding cubic Hermite splines over N subintervals is twice more expensive than finding cubic splines. That means that looking at time costs of the algorithm, we may choose at least twice more point cubic splines as we use for cubic Hermite splines for the same time cost. So, the normalized accuracy will be:

$$|E(t)| = \frac{5}{384} \|f^{(IV)}\| \left(\frac{h}{2}\right)^4 = \frac{5}{16 \cdot 384} \|f^{(IV)}\| h^4 \quad (3.2.10)$$

Thus, the fact that $5/16 < 1$ enables us not to take cubic Hermitic splines for our calculations.

Trigonometric interpolation

Taking into consideration that NMR signals are more or less periodic it seems rather logical to look for specific interpolation. That is why the fast Fourier transform interpolation [75] was studied. A trigonometric polynomial of degree n has the form

$$p(x) = a_0 + \sum_{m=1}^n a_m \cos(mx) + \sum_{m=1}^n b_m \sin(mx). \quad (3.2.11)$$

This expression contains $2n + 1$ coefficients, $a_0, a_1, \dots, a_n, b_1, \dots, b_n$ and we wish to compute those coefficients so that the function passes through N points:

$$p(x_k) = y_k, \quad k = 1, \dots, N. \quad (3.2.12)$$

Since the trigonometric polynomial is periodic with period 2π , it makes sense to assume that

$$0 \leq x_1 < x_2 < \dots < x_N < 2\pi. \quad (3.2.13)$$

The interpolation problem is now to find coefficients such that the trigonometric polynomial p satisfies the interpolation conditions.

Under the above conditions, there exists a solution to the problem for any given set of data points $\{x_k, p(x_k)\}$ as long as N , the number of data points, is not larger than the number of coefficients in the polynomial, i.e., $N \leq 2n + 1$ (a solution may or may not exist if $N > 2n + 1$ depending on the particular set of data points). Moreover, the interpolating polynomial is unique if and only if the number of adjustable coefficients is equal to the number of data points, i.e., $N = 2n + 1$.

The solution can be written in a form similar to the Lagrange formula for polynomial interpolation:

$$p(x) = \sum_{k=1}^{2n+1} y_k \prod_{m=1, m \neq k}^{2n+1} \frac{\sin \frac{1}{2}(x - x_m)}{\sin \frac{1}{2}(x_k - x_m)}. \quad (3.2.14)$$

This can be shown to be a trigonometric polynomial by employing the multiple-angle formula and other identities for $\sin \frac{1}{2}(x - x_m)$.

Comparison

These three methods of interpolation were studied respectively in terms of accuracy and time of execution.

The steps of the study were:

- to simulate a very accurate signal,
- to simulate a low-resolved signal,
- to interpolate the second signal to the number of points in the first one,
- to look at the difference between the original and interpolated signals and consequently at the interpolation quality criterion.

A simulated signal was set up by the quantum mechanics conventional approach [5] of Hamiltonian with the Zeeman and spin-spin terms. The simulated signal was the sum of signals with equal molar concentrations of the common cerebral metabolites (Cr, Lac, Ala, Myo, Asp) and the ethanol (Eth) often found as a trace of the biopsy procedure. The simulation parameters were chosen to fit usual HRMAS signals obtained at resonance frequency of 500MHz (magnetic field 11.9T), sampling interval 160msec, damping factor $4s^{-1}$. The number of points of the signal was chosen equal to 256k to consider this signal as an ideal signal. For the interpolated signal much more lower number of points was chosen starting from 8k (to subject the signal to the Nyquist criterion [76] for the highest peak frequencies). Typical results of these interpolations using built-in MATLAB functions are presented in Fig. 3.2.

3.2.2 Time- or frequency-domain

The subsequent idea was to investigate which domain would be better to process a signal. There are two points of attention here to be taken into account. Firstly,

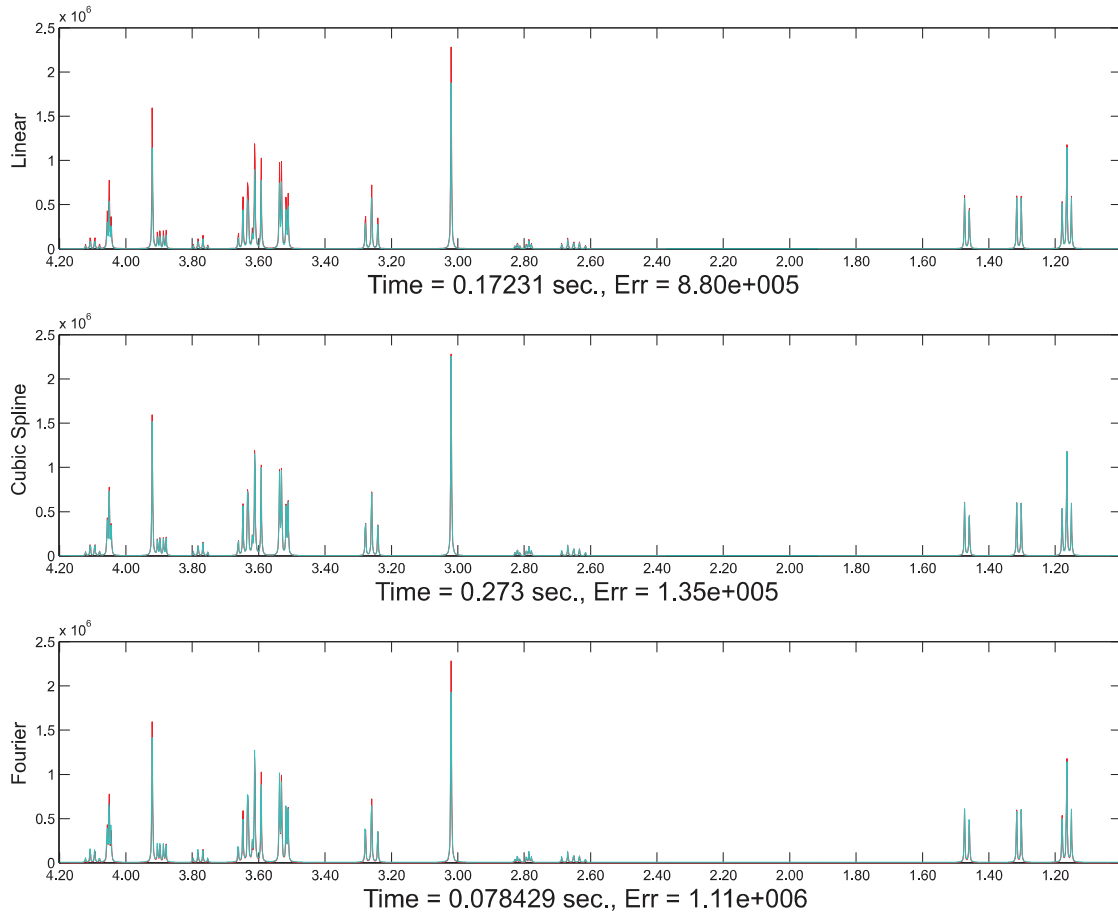


Figure 3.2: Spectra, times of execution and relative errors for different types of interpolation applied to a simulated HRMAS-spectrum (500MHz).

it is more or less logical to test the interpolation methods applied to the FFT of the signals, because the spectrum of a simulated signal approaches zero for the majority of the bandwidth:

$$\frac{d^n f}{d\nu^n} \approx 0, n = 0, 1, 2, \dots \quad \text{for} \quad \begin{cases} \omega \gg \omega_0, \\ \omega \ll \omega_0 \end{cases} \quad (3.2.15)$$

which undergoes the condition 3.2.3 of the error minimization (here, ω_0 signifies a peak position).

Secondly, it is easily seen that if for the expansion (in time-domain) procedure we have some redundant points which can be simply cut off, the contraction action will lead to an undetermined region of the signal. This fact is clearly represented in the Fig. 3.3.

A signal of about 14000 points was simulated for a Lactate sample with usual HRMAS parameters (500MHz, $t_s=0.16\text{ms}$, $\alpha=-4\text{Hz}$). Subsequently, a contraction procedure was applied with a stretching ratio of 0.999. The cubic-spline algorithm was used for the interpolation. Fig. 3.3 clearly shows the ambiguity of the signal extrapolation for the last points in this case as a consequence of the cubic interpolation (the last “*sine oscillation*” is prolonged by a cubic polynomial).

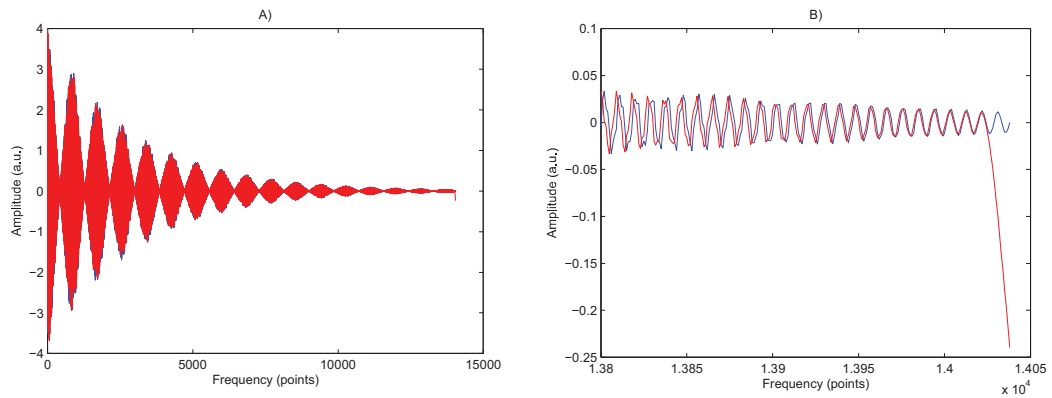


Figure 3.3: Ambiguity of the final pixels after application of the contraction procedure. Blue – original phantom signal, red – a signal contracted with a ratio 0.999. Last points' behaviour is the consequence of a cubic-spline interpolation.

Certainly, this artefact will have influence on the quantitation results. Simply, it can be represented by the distortion of the spectrum (see Fig. 3.4).

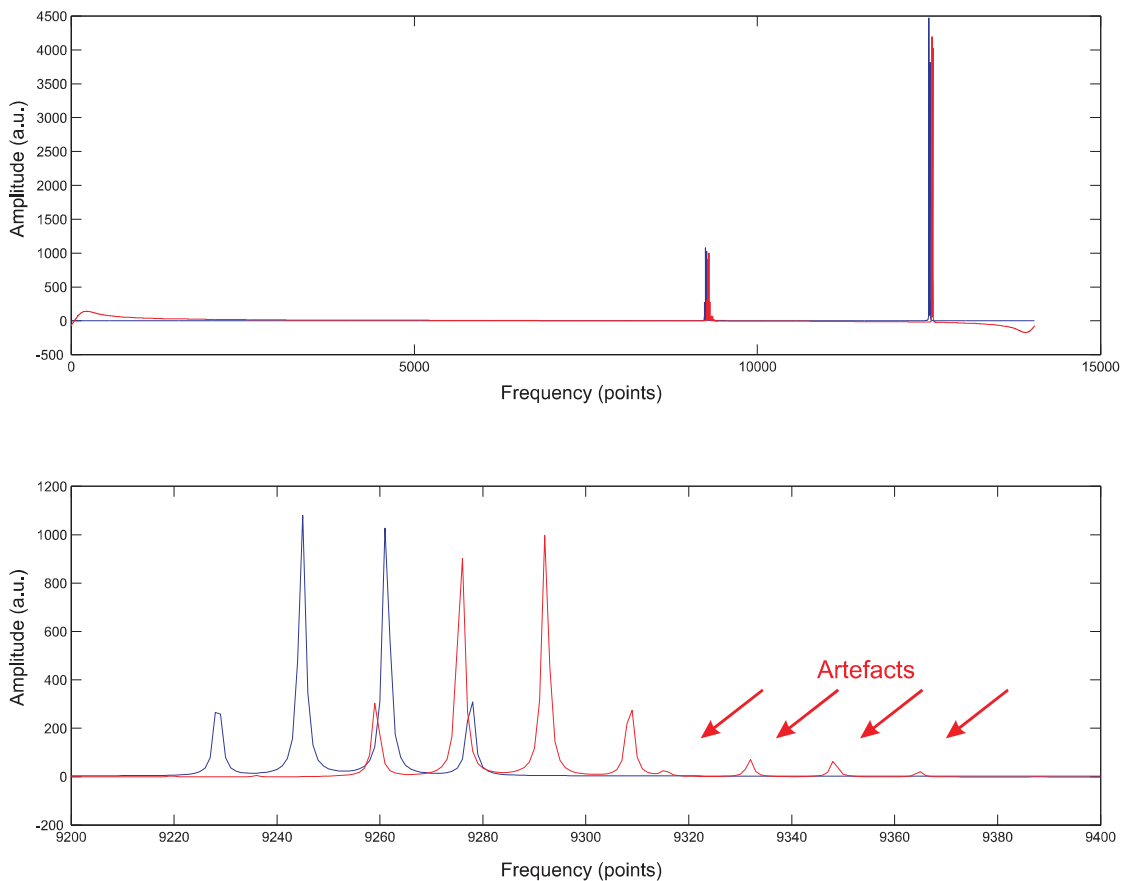


Figure 3.4: Distortion of the spectrum after applying the stretching procedure (with cubic spline interpolation) with a ratio of 0.995. Note the appeared artefacts easily seen in the bottom zoom in at right.

This distortion can be decreased by using other processing methods applied to the end of the FID signal. Firstly let us take into attention the fastest way which is the zero filling for the non-determined part of a signal. The time- and frequency-domain results of this procedure are presented in Fig. 3.5.

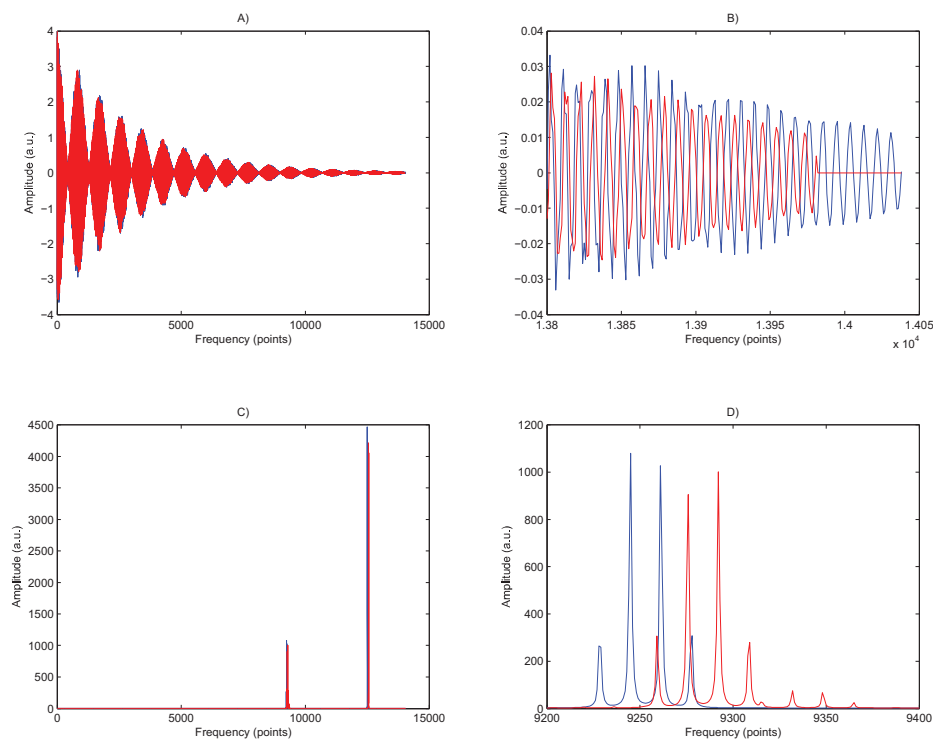


Figure 3.5: Zero-filling method for decreasing the contraction effects: A) Lactic acid phantom FID signal, B) Zoom in in the region of interest, C) Spectrum, D) Zoom in in the region of interest of the spectrum. The zero-filling leads to the spectrum distortion.

Also, rather satisfactory results can be obtained with leaving the undefined part of a signal unchanged. This means that the last points of the shrunk FID signal will be copied from the unstretched signal. The visual representation of this procedure is shown in Fig. 3.6.

Now the same procedures will be applied in the frequency domain. As it was noticed above, there are large “*low-information*” regions in the spectral domain. Having executed the procedures one can simply notice the advantage of this implementation. Computing times are more or less the same (that was expected due to the same number of points in FID and its FFT). The execution of FFT-interpolation is senseless in that case because, firstly, the spectrum is not periodic and, secondly, the Fourier transformation has been already done so now the implementation of linear or cubic interpolation is already the fact of interpolating the FID with the help of FFT.

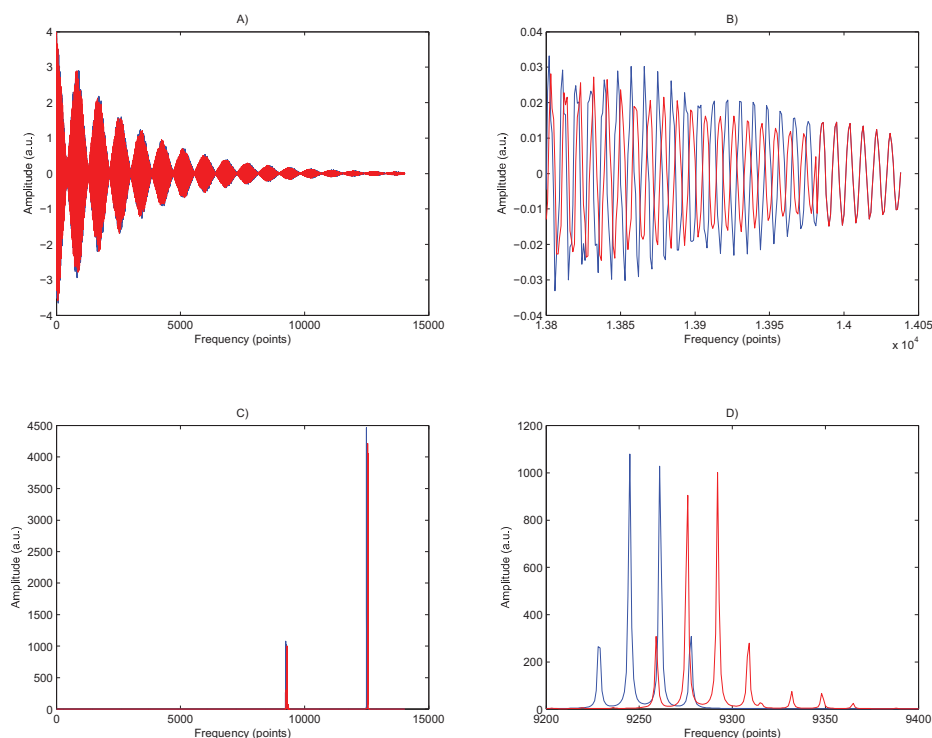


Figure 3.6: Conservation of the last points for the contraction procedure: A) Simulated lactic acid signal, B) Zoom in in the region of interest, C) Spectrum, D) Zoom in in the region of interest in the spectrum. Notice: the blue part at the end of the signal is not “absent”, it is just exactly overlapped by the red one.

The only analysis which should be done is the error analysis for the interpolation routine. As it was shown in 3.2.3 and in 3.2.7 the top limit of the error depends on the “curvature” of the interpolated function for the linear interpolation and on the fourth derivative for the cubic spline interpolation. Thus, now our goal will be to compare the derivative values for the normal signal and its FFT.

3.2.3 FID cumulated error estimation

For the general representation of a free induction decay with “zero” initial conditions ($t_0 = 0$, $\phi_0 = 0$)

$$f(t) = A \exp((i\omega - \alpha)t) \quad (3.2.16)$$

it is simple to calculate the derivatives with respect to t :

$$f''(t) = A(i\omega - \alpha)^2 \exp((i\omega - \alpha)t) \quad (3.2.17)$$

$$f^{(4)}(t) = A(i\omega - \alpha)^4 \exp((i\omega - \alpha)t) \quad (3.2.18)$$

For the FID error calculations, we will assume that $\omega \gg \alpha$ which is totally natural for HR signals. In this case for the total error estimation, we can start from one-period error calculation neglecting the decay function:

$$\begin{aligned}
E_{\text{lin}} &\leq \frac{\|f''\|_T}{8} h^2 T = 1/2 A \pi^2 \nu^2 t_s^2, \\
E_{\text{cub}} &\leq \frac{\|5f^{(4)}\|_T}{384} h^4 T = 120 A \pi^4 \nu^4 t_s^4
\end{aligned} \tag{3.2.19}$$

where t_s is the sampling interval, ν is the frequency, T - the period of signal.

The accurate derivations of formulae mentioned above and below are presented in the Appendix.

Taking into consideration that the total acquisition time is $N_p \cdot t_s$ (where N_p is the number of points in the processed signal) one can derive the number of periods during the acquisition

$$N_{\text{periods}} = N_p \cdot t_s \cdot \nu \tag{3.2.20}$$

which allows us to interpret the total error as

$$E_{\text{lin}} \leq \sum_{n=0}^{N_{\text{periods}}} 1/2 A \pi^2 \nu^2 t_s^2 e^{-\frac{n\alpha}{\nu}} \tag{3.2.21}$$

for linear interpolation and

$$E_{\text{cub}} \leq \sum_{n=0}^{N_{\text{periods}}} 120 A \pi^4 \nu^4 t_s^4 e^{-\frac{n\alpha}{\nu}} \tag{3.2.22}$$

for the cubic spline which can be respectively estimated assuming that $n \rightarrow \infty$ relatively to the scales of calculation:

$$\begin{aligned}
E_{\text{lin}} &\leq \frac{A \pi^2 \nu^3 t_s^2}{2\alpha}, \\
E_{\text{cub}} &\leq \frac{20 A \pi^4 \nu^5 t_s^3}{\alpha}
\end{aligned} \tag{3.2.23}$$

3.2.4 Spectrum cumulated error estimation

Let us consider the FFT interpolation error calculation. However, here we do not have a natural time interval to be considered as small one relatively to the other parameters. So it will be chosen as $d\omega$ – the element of spectral resolution – and later used for integration.

As it was noticed in Eq. 1.4.2 the shape of the peak which results of FID Fourier transform can be described as

$$f(\omega) = \frac{A}{\pi} \left(\frac{\alpha}{\omega^2 + \alpha^2} \right). \tag{3.2.24}$$

which will be used to compute the needed derivatives:

$$\frac{d^2 f}{d\omega^2} = \frac{2A\alpha}{\pi (\omega^2 + \alpha^2)^2} \left(\frac{4\omega^2}{\omega^2 + \alpha^2} - 1 \right) \tag{3.2.25}$$

$$\frac{d^4 f}{d\omega^4} = \frac{24A\alpha}{\pi(\omega^2 + \alpha^2)^3} \left(\frac{16\omega^4}{(\omega^2 + \alpha^2)^2} - \frac{12\omega^2}{\omega^2 + \alpha^2} + 1 \right) \quad (3.2.26)$$

Unfortunately, the obtained functions are too complicated to be integrated analytically. But it can be shown that it is limited by the easily integrated function $f(\omega)$ with respective coefficients and powers:

$$\begin{aligned} \frac{d^2 f}{d\omega^2} &< \frac{6f}{\omega^2 + \alpha^2} & \omega \in (-\infty.. \infty), \\ \frac{d^4 f}{d\omega^4} &< \frac{120f}{(\omega^2 + \alpha^2)^2} & \omega \in (-\infty.. \infty). \end{aligned} \quad (3.2.27)$$

Graphically it is presented in Fig. 3.7. These functions are easily integrable (the

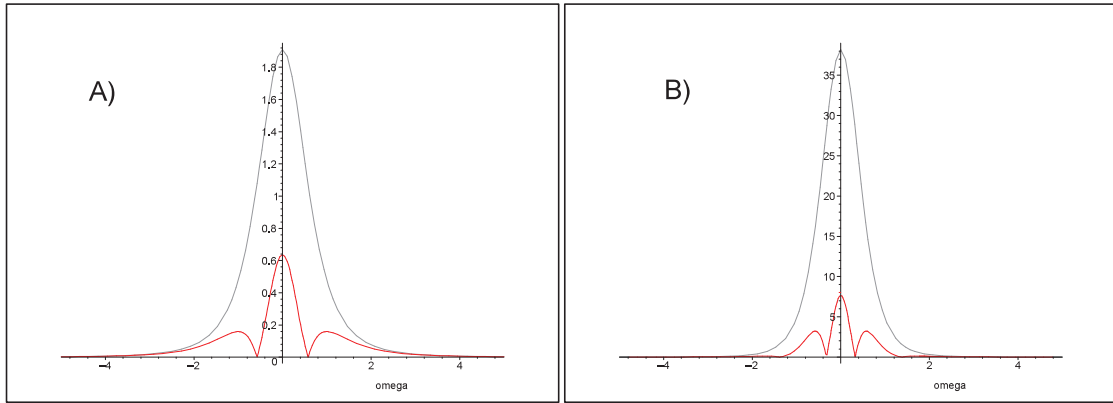


Figure 3.7: Visual representation of the bounding of the derivatives (grey) of the Lorentzian function superposed on the function itself (red). A) second derivative, B) fourth derivative.

integrals are given in the Appendix):

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{6f}{\omega^2 + \alpha^2} d\omega &= \frac{3A}{\alpha^2}, \\ \int_{-\infty}^{\infty} \frac{120f}{(\omega^2 + \alpha^2)^2} d\omega &= \frac{45A}{\alpha^4} \end{aligned} \quad (3.2.28)$$

So, after multiplication by the respective coefficient for error calculations, we can state that the integrated errors on the total bandwidth can be expressed as:

$$\begin{aligned} E_{\text{lin}} &\leq \frac{3A\pi^2}{2\alpha^2 N_p^2 t_s^2}, \\ E_{\text{cub}} &\leq \frac{75A\pi^4}{8\alpha^4 N_p^4 t_s^4} \end{aligned} \quad (3.2.29)$$

These calculations are presented in details in the Appendix.

3.2.5 Error comparison

Actually, the last multiplication by the axis scaling parameter is a little redundant because we are interested in the “*whole-signal cumulated error*” normalized to the bandwidth. That means that the error results obtained for time-domain will be normalized with respect to the decay time and the frequency-domain ones – to the bandwidth respectively. To make it clear, these comparisons are represented in Tab. 3.1.

| | Time domain | Frequency domain |
|-----------------------------|---|--|
| Linear interpolation | $\frac{A\pi^2\nu^2t_s}{2\alpha N_p}$ | $\frac{3A\pi^2}{2\alpha^2 N_p^2 t_s^2}$ |
| Spline interpolation | $\frac{20A\pi^4\nu^4t_s^3}{\alpha N_p}$ | $\frac{75A\pi^4}{8\alpha^4 N_p^4 t_s^4}$ |

Table 3.1: Comparison of the cumulated errors when using the linear and spline interpolations procedures on a signal represented in time and frequency domains.

For linear interpolation (approximately) the ratio between the respective errors

$$\frac{E_{\text{TD}}}{E_{\text{FD}}} \sim \nu^2 t_s^3 \alpha N_p > 1 \quad (3.2.30)$$

is the criterion of preference of frequency domain for the algorithm implementation. For common HR NMR signal parameters ($\nu \sim 10^3\text{Hz}$, $t_s \sim 10^{-4}\text{s}$, $\alpha \sim 1\text{Hz}$, $N_p \sim 10^4$) it leads to $10 > 1$ which is true, so the selection of frequency domain is useful. And for the natural cubic spline interpolation,

$$\frac{E_{\text{TD}}}{E_{\text{FD}}} \sim \nu^4 t_s^7 \alpha^3 N_p^3 < 1 \quad (3.2.31)$$

gives an advantage for time domain processing. But, looking at the orders of the compared values one can notice that this decision is not absolute and sometimes can vary accordingly to the task.

3.2.6 Akima interpolation

One specific case for which this subsection is dedicated is the case of low-resolved signal. Despite the fact that the “HR” in “HRMAS” has the meaning “high resolution”, the low-resolution signals remain of interest. The point is that such signals are highly resolved in physical sense but have lost their resolution after certain signal processing operations (as, for example, cutting a certain number of points). In this case it is extremely necessary to do a good interpolation to restore the maximum information. And also, taking into consideration that the frequency domain is chosen for signal processing, it will be very important to find the interpolation algorithm avoiding the cubic-spline artefacts. The most significant of them is the “oscillation” behaviour near the point for which the value is rather different from the average value of the region (see Fig. 3.8, top).

That is why the Akima interpolation [77] must have our attention. In this method the series of five consecutive points ([1..5]) are investigated. Firstly,

the slope in the middle point is calculated from the four known linear slopes ($\overline{12}, \overline{23}, \overline{34}, \overline{45}$):

$$t = \frac{|m_4 - m_3|m_2 + |m_2 - m_1|m_3}{|m_4 - m_3| + |m_2 - m_1|} \quad (3.2.32)$$

where $m_1, m_2, m_3,$ and m_4 are the slopes of the line segments $\overline{12}, \overline{23}, \overline{34}$ and $\overline{45}$, respectively.

Then, for each pair of points, we express the curve between them using the values of two ordinates and two slopes at the ends as a polynomial of, at most, degree three. The polynomial, though uniquely determined, can be written in several ways. As an example we shall give the following form:

$$y = p_0 + p_1(x - x_1) + p_2(x - x_1)^2 + p_3(x - x_1)^3 \quad (3.2.33)$$

where

$$\begin{aligned} p_0 &= y_1, \\ p_1 &= t_1, \\ p_2 &= \frac{3\frac{y_2 - y_1}{x_2 - x_1} - 2t_1 - t_2}{x_2 - x_1}, \\ p_3 &= \frac{t_1 + t_2 - 2\frac{y_2 - y_1}{x_2 - x_1} - 2t_1 - t_2}{(x_2 - x_1)^2}. \end{aligned} \quad (3.2.34)$$

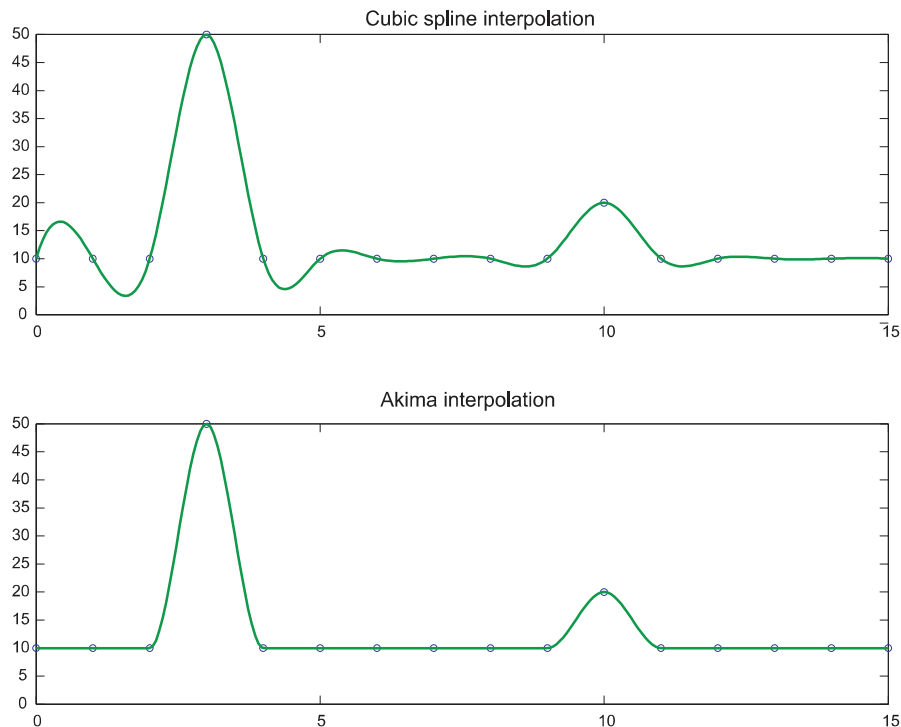


Figure 3.8: The cubic spline interpolation (top) in comparison to the Akima one (bottom)(MATLAB realisation by Shamsundar [78]).

And, finally, at each end of the curve, two more points have to be estimated from the given points. We assume for this purpose that the end point (x_3, y_3)

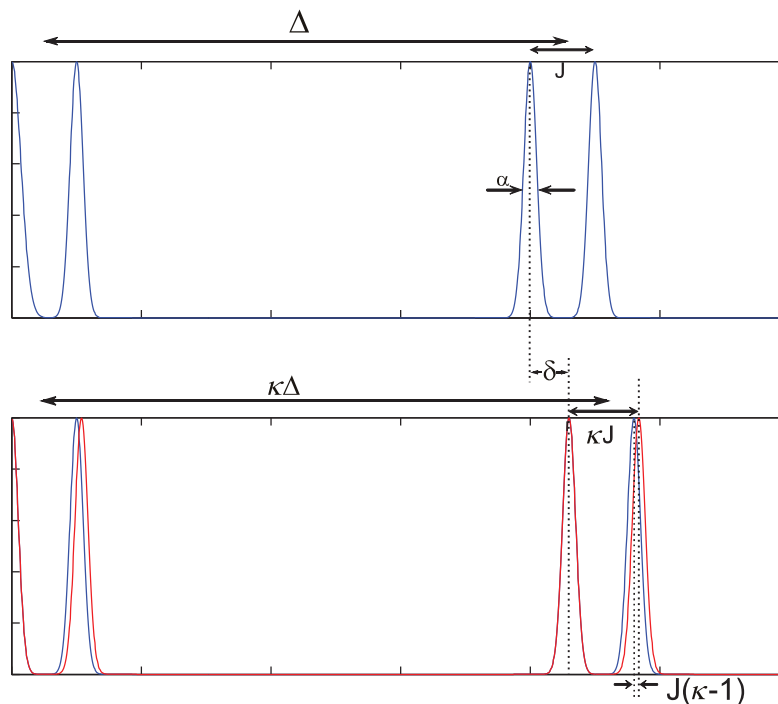


Figure 3.9: Visual representation of the stretching consequences: top – original signal, bottom – signal to be fit(blue) and stretched signal(red)

and two adjacent given points (x_2, y_2) and (x_1, y_1) , together with two more points (x_4, y_4) and (x_5, y_5) to be estimated, lie on a curve expressed by

$$y = g_0 + g_1(x - x_3) + g_2(x - x_3)^2, \quad (3.2.35)$$

where the g 's are constants. Assuming that

$$x_5 - x_3 = x_4 - x_2 = x_5 - x_1, \quad (3.2.36)$$

we can determine the ordinates y_4 and y_5 , corresponding to x_4 and x_5 , respectively, from Eq. 3.2.35. The results are

$$\frac{(y_5 - y_4)}{x_5 - x_4} - \frac{(y_4 - y_3)}{x_4 - x_3} = \frac{(y_4 - y_3)}{x_4 - x_3} - \frac{(y_3 - y_2)}{x_3 - x_2} = \frac{(y_3 - y_2)}{x_3 - x_2} - \frac{(y_2 - y_1)}{x_2 - x_1}. \quad (3.2.37)$$

The positive consequences of replacing the cubic spline interpolation by the Akima one are easily seen on the examples with the outliers. An example of the demonstrating interpolations is shown in Fig. 3.8, bottom.

3.2.7 Theoretical limits of the method

But, certainly, the method has some limitations in its implementation. They appear due to fact that small changes of chemical shifts do not induce changes of J-coupling values. But the stretching/shrinking method changes the distances between the peaks in each multiplet. This makes certain limitations to the stretching

values respectively to the J-coupling values and distances between the peaks. The stretching coefficient can be defined as $\kappa = (\Delta + \delta)/\delta$, where Δ is the chemical shift difference between two groups of spins, δ is the mismatch of the peaks to be corrected. Mathematically the limitation can be expressed as

$$J(\kappa - 1) < \alpha/2 \quad (3.2.38)$$

or

$$J \left(\frac{\Delta + \delta}{\Delta} - 1 \right) < \frac{\alpha}{2} \quad (3.2.39)$$

for the general case. This expression can easily be simplified to

$$\delta < \frac{\alpha\Delta}{2J} \quad (3.2.40)$$

Visually this effect is illustrated in the Fig. 3.9.

An analysis of matching this criterion for interesting metabolites has been provided in Tab. 3.2. All the calculations were done for a resonance frequency

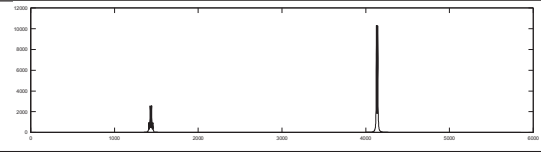
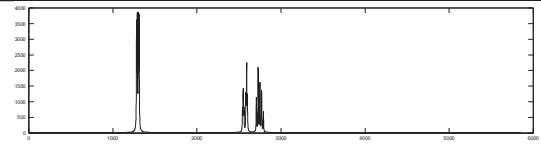
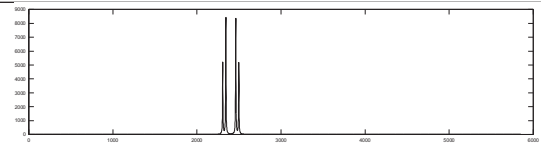
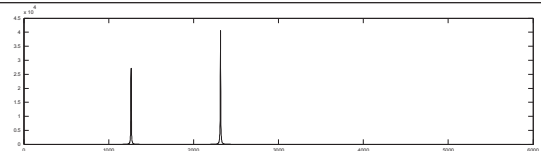
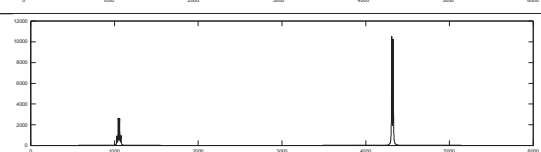
| Metabolite | Spectrum | AX-distance, ppm | δ_{\max} , Hz | δ_{\max} , ppm |
|------------|--|------------------|----------------------|-----------------------|
| Alanine |  | 2.31 | 160 | 0.32 |
| Aspartate |  | 1.09 | 31.3 | 0.063 |
| Citrate |  | 0.13 | 4.1 | 0.008 |
| Creatine |  | 0.89 | — | — |
| Lactate |  | 2.78 | 167 | 0.334 |

Table 3.2: Metabolites, well processed with the method. The limit of correction was calculated for a resonance frequency of 500MHz and a damping factor of 2Hz.

of 500MHz and a damping factor of 2Hz. A remarkable point in this table is the absence of any correction limit for the Creatine spectrum. The reason is very

simple: both groups of protons of Creatine produce singlets but not multiplets. So there is no distortion of the spectrum because of stretching. To be strict, we must notice that there is still a limitation of stretching for such types of spectra, related to possible peak shape distortion. Being stretched too much the signal will feel artificial damping factor changes. Fortunately, the stretching parameter, needed to make this effect considerable, is too large and thus is not a subject of investigation.

3.3 Splitting of spectra

Splitting of spectra, as method of correction, is evident. It is based on splitting of the spectrum in independent parts that can be moved independently along the frequency axis. This method has also numerous advantages for the correction of spectra consisting of more than two multiplets. It has not the disadvantages of multiplet shape distortion due to stretching/shrinking discussed previously and presented in table 3.2.

There are different methods for signal decomposition. The main task now will be to split automatically the spectrum in different groups and to determine the belonging of the peaks to a certain group.

3.3.1 Group selection

Because of the discrete nature of spectra and assuming Lorentzian lineshapes, we had to introduce the parameter d (see Fig. 3.10) greater than the half-height linewidth of peaks in order to not lose peak area when delimiting a peak. The normalized error between the total area and the integral on the interval $[-d; d]$ of a peak area as a function of d is given in Fig. 3.11.

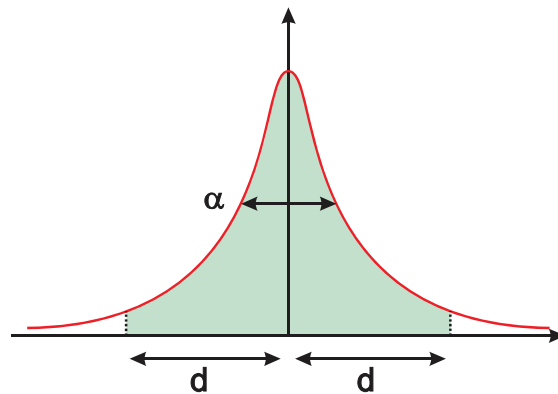


Figure 3.10: Lorentzian function with its linewidth and limits of integration presented.

For the following investigation this function will sometimes be read as the inverse: we will know the necessary peak size d as a function of the desired accuracy of the peak selection ζ :

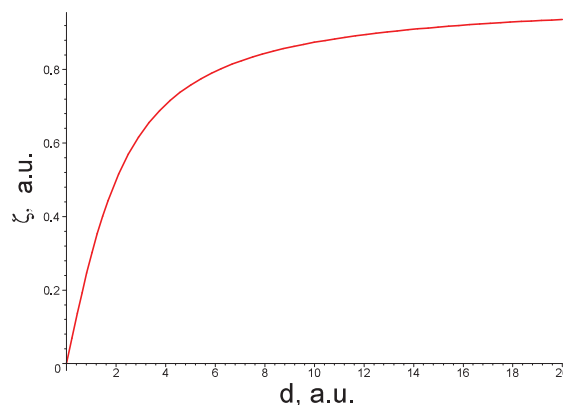


Figure 3.11: Plot of the selection accuracy ζ as a function of the chosen peak size d .

$$d = \alpha \tan\left(\frac{\zeta\pi}{2}\right) \quad (3.3.1)$$

where α signifies the half-height line width.

As the area of one peak is dependent on the accuracy of processing, the area of a group of peaks will also be dependent on it. If all the peak positions are known, we will start the procedure of group selection. In all cases, it starts from the selection of the most distanced peaks (Fig. 3.12). Recall that we work with the simulated spectrum and there is no noise distortion, so the peaks will really belong to the specified groups.



Figure 3.12: The most distanced peaks selection.

The second step is to define the middle point between the selected peaks. Working with the assumption of an AX-case, we assume that there will not be any peak *in the interval* $[+d; -d]$ *from the middle point*. The width parameter d is defined in Fig. 3.10 and computed in Eq. 3.3.1. So, after selecting the middle point, one can trust that both multiplets are on the different sides from it (Fig. 3.13-2).

The next peaks to be found are the two peaks closest to the middle point and belonging to each part of the spectrum. Fig. 3.13-3 shows how they are determined. This allows us to be sure that everything between the selected bounding peaks belongs to the respective groups. And, finally, we need to surround each group by the interval of precision d defined in Eq. 3.3.1. This is illustrated in Fig. 3.13-4.

An additional algorithm of selection must be introduced for three-group spectra where the fact of choosing middle point is senseless.

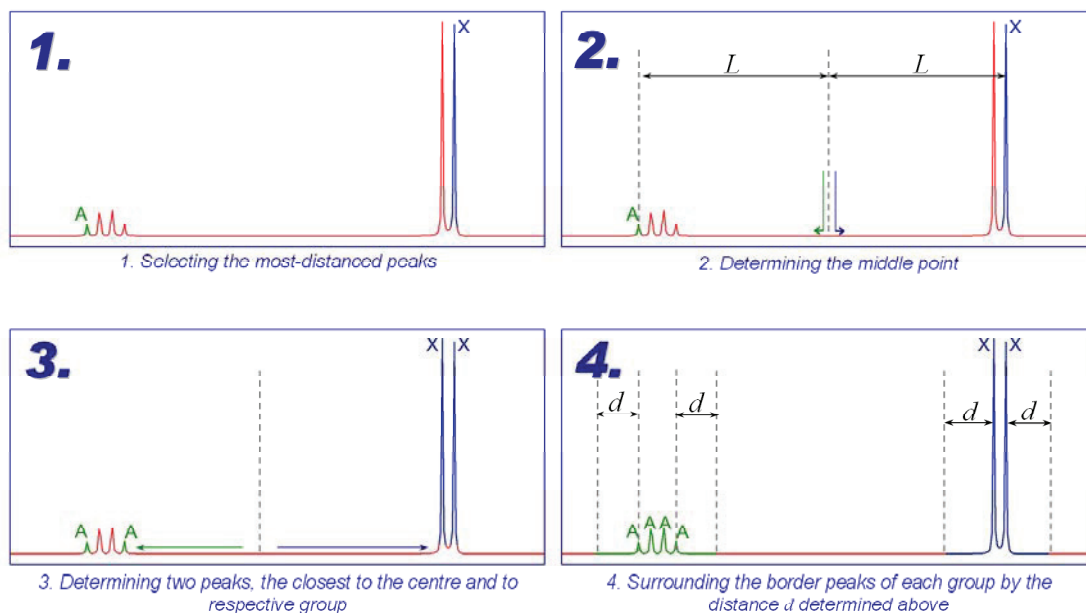


Figure 3.13: Group selection scheme. AX-case.

The initial idea is the same – to select two most distanced peaks in the spectrum (Fig. 3.12). We will consider all consecutive frequency intervals between the peaks choosing the *two largest* frequency intervals. This step is shown in Fig. 3.14-A. So now the border peaks of each group are known and, as in the previous case, we surround each group by an interval d to achieve the necessary precision of selection.

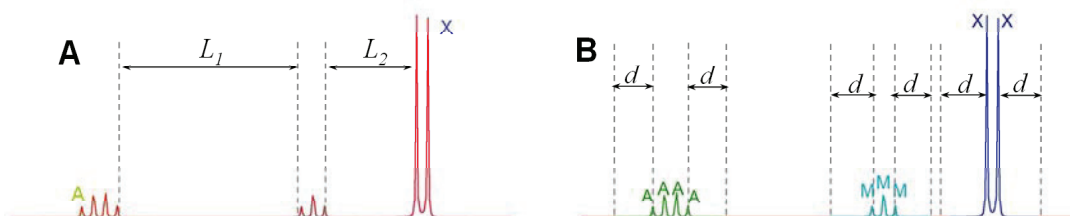


Figure 3.14: Group selection scheme. AMX-case.

3.3.2 Theoretical limitations

Certainly, this procedure has some limitations in its implementation. Sometimes, certain groups are too close to be separated with the necessary accuracy. Obviously, to make this selection, the distance between two border peaks of neighbour groups must be larger than $2d$. Mathematically this can be expressed as

$$\Delta - \frac{nJ_1 + mJ_2}{2} > d \quad (3.3.2)$$

where Δ is the distance between the centres of the multiplets, n and m are the multiplicities of the groups and J_1 and J_2 are the respective coupling constants.

We need to notice that in the general case J_1 and J_2 are different because the neighbour multiplets have different couplings.

If one expresses d , we obtain:

$$\Delta - \frac{nJ_1 + mJ_2}{2} > \alpha \tan\left(\frac{\zeta\pi}{2}\right) \quad (3.3.3)$$

And, neglecting the differences of coupling constants and taking into consideration the quadruplets as a maximum level of coupling (for some common cases as Lac, Cit, Ala etc.) one can estimate the limits of accuracy:

$$\zeta < \frac{2}{\pi} \arctan\left(\frac{\Delta - 3J}{\alpha}\right) \quad (3.3.4)$$

This limitation is visually represented in Fig. 3.15. The following plots (Fig. 3.16)

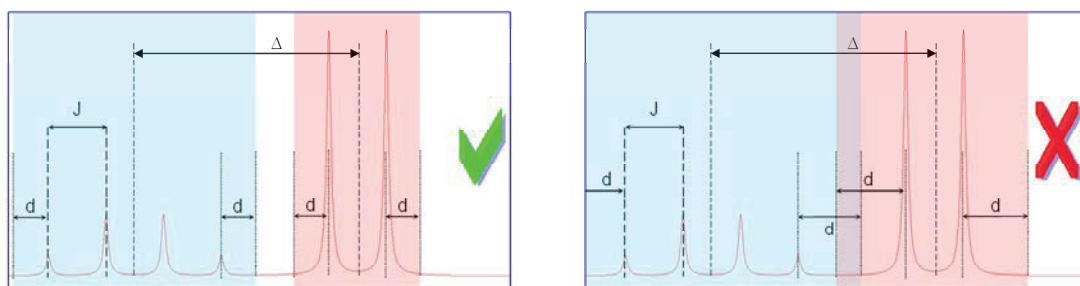


Figure 3.15: Visual representation of the splitting implementation limits.

illustrate dependencies of possible accuracy with respect to the parameters:

- Plot A - $\zeta(\alpha, \Delta)$, J is set to 10Hz ,
- Plot B - $\zeta(\Delta/\alpha, J/\alpha)$ – in the usual ranges for NMR spectra,
- Plot C - $\zeta(\Delta/\alpha)$ for $J/a = 2$.

3.3.3 Accuracy-time dilemma

As is evident, the accuracy parameter, has an influence on the width of the moved region of the spectrum. This could raise a question about its influence on the computing time of the algorithm. The empirical way to see this dependence was chosen.

Starting from the simulated spectrum (Lactic acid NMR spectrum at 500Hz artificially apodized by a Lorentzian with damping factor of 4Hz) another spectrum was simulated with the position of the 4.09ppm quadruplet artificially shifted to 4.11ppm . The second spectrum was subject to the splitting procedure implemented in the MATLAB code. The goal was to select the groups and to shift the respective group to its normal position (a known fixed value of -0.02ppm). Shifting of the group was implemented in the frequency domain. The obtained signal was

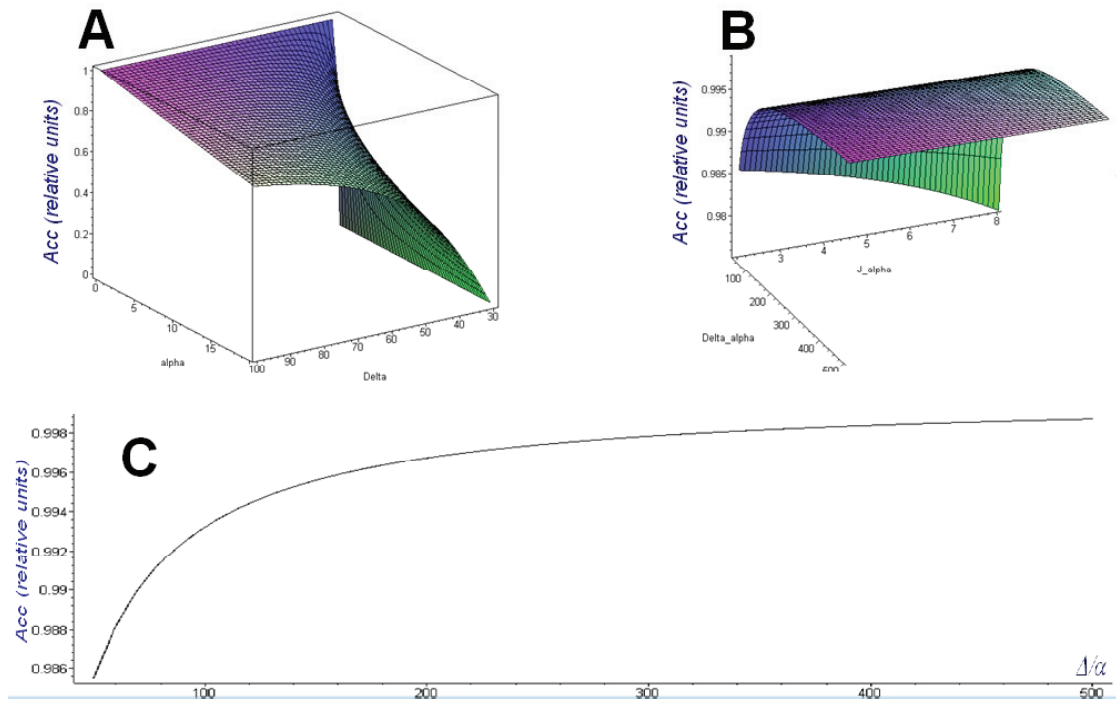


Figure 3.16: Some numerical estimation of the limits of the splitting implementation.

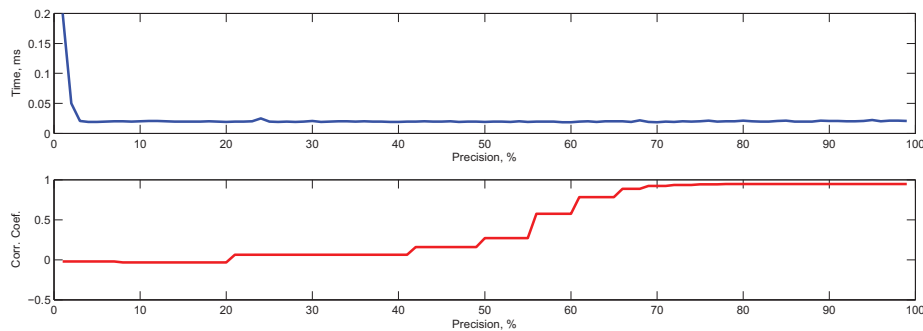


Figure 3.17: Computing time and method accuracy as a function of the group selection precision. Whole region.

compared to the initial one by maximizing their mutual correlation in the time domain. The results of this virtual experiment are presented in Fig. 3.17 and 3.18.

The behaviour of the precision graph is as predicted. It should be only noticed that the experimental function is not monotone which is quite normal, especially in the low-accuracy region. The reason is the discrete dependence of the number of selected points within $\pm d$ on the precision argument. This raises the fact that for different (but close) values of precision the same regions of spectrum might be selected. This can easily be seen by having a look at the values from the graph in Fig. 3.16-C. So these regions will produce the same errors in analysis of correlation.

But surprising was the dependence of the execution time against the size of se-

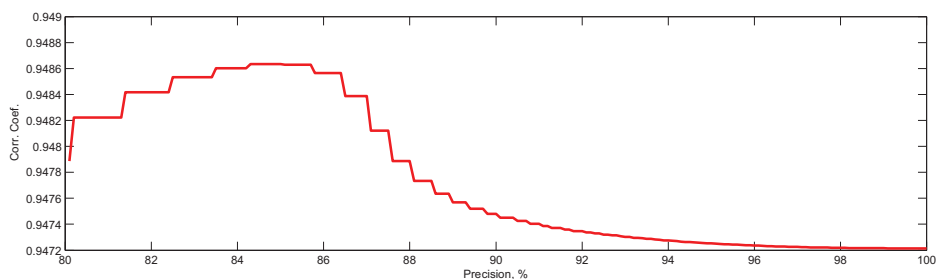


Figure 3.18: Method accuracy as a function of the group selection precision. Zoom in in the high-precision region.

lected region (precision of selection). The time of algorithm execution is incredibly weakly dependent on the accuracy. The only high-value point presented on the graph is an outlier and could not be avoided because of MATLAB-implementation properties (this additional time is spent for the first call of the procedure). This interesting fact will give us more freedom in future, granting a possibility of free choice of necessary accuracy.

A part of specific attention should be paid to the graph in Fig. 3.18 with zoom in in the last points. It appears that after some value of precision the accuracy of calculation starts to decrease. But this fact has a logical explanation. After achieving a certain value of the precision the width of the selected zone is so large that it “takes” points from other multiplets. The criterion Eq. 3.3.4 of the theoretical limits of the algorithm is not valid anymore.

3.3.4 Integer-valued shift

In the stretching/shrinking section, we had to resort to methods of interpolation. For the splitting algorithm we have the possibilities *to use or not to use interpolation*. We may shift a group with an integer number of points, so as to save time but to lose accuracy, or to find the exact value of the necessary shift of a group position but perform interpolation and have computing time which could be important. We will try to analyze it empirically, simulating the same phantom for which the computing time dependencies will occur.

This virtual experiment consists of simulating two lactate signals with a displacement of position of one group of 0.02ppm. Series of shifts of the group positions will be executed and computing time and accuracy values recorded. For the representation of Fig. 3.19 one thousand of possible shifts were done within the interval ± 0.01 ppm around the initial value (with the distance of $2 \cdot 10^{-5}$ ppm respectively). Accuracy of group selection was chosen to .998.

The values were chosen in the way to make each 20th displacement value integer (in points). In this case the algorithm skipped the interpolation procedure, directly replacing points positions. As it is seen from Fig. 3.19, the integer shifts are considerably (up to two times) faster than the interpolation ones. Also this graph allows us to estimate the minimum error to be allowed to consider shift values as integer-point and to round the shifting value.

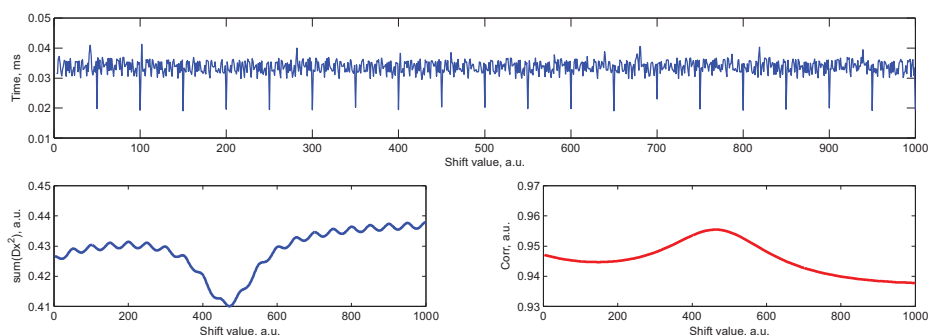


Figure 3.19: Time and accuracy dependence on the displacement. Integer-points shifts are up to two times faster.

3.3.5 Strongly-coupled spins

Due to the properties of the Splitting method, strongly-coupled protons need to be investigated in more details. The reason of this is the dependence of the distribution of the energy within the multiplet (in simple words: the *shape* of spectral group) on the distance between the groups.

Normally, if two strongly-coupled spin systems have the same values of coupling constants but different AB-distances in NMR spectra, they will produce multiplets which are similar but never totally “overlappable”. This can be understood if we appeal to the derivation in page 23. Because of the diagonalization procedure non-diagonal elements of the Hamiltonian will have an influence on the eigenvectors which will later influence the transition probabilities, and consequently the amplitudes of the peaks. The fact that the peak amplitudes change when varying δ/J value has already been mentioned in this work and this influence was presented in Tab. 1.4. So, the following subsection is to study this influence more in details.

To do that, two phantom spectra of type AB were simulated. The first one had a fixed AB distance ($\Delta = \delta_A - \delta_B$), the second one had a displacement of $\delta\Delta$ of one of the peaks. The multiplet signals in the second spectrum were taken independently (the method was explained more in detail in Ch. 2). After this, the spectral group of the mismatched group was shifted over δ and its position was optimized by minimization of the root mean square of the difference between the signals with standard MATLAB *lsqnonlin*. The residue after this minimization has been taken as a criterion:

$$\min \sum_{i=1}^{N_{\text{points}}} (S_1(i) - S_2(i))^2 \quad (3.3.5)$$

The variables were the mismatch δ and AB-distance Δ . So, the series of 3D-plots enable us to see the expected dependence (see Fig. 3.20).

Plots were built for different combinations of the multiplets, produced by the common NMR spin systems, such as: AB, AB₂, AB₃, A₂B₂, A₂B₃ and A₃B₃ (where the latter might be a little redundant because chemically it corresponds to a carbon-proton structure CH₃–CH₃ which is ethane by itself but not a part of

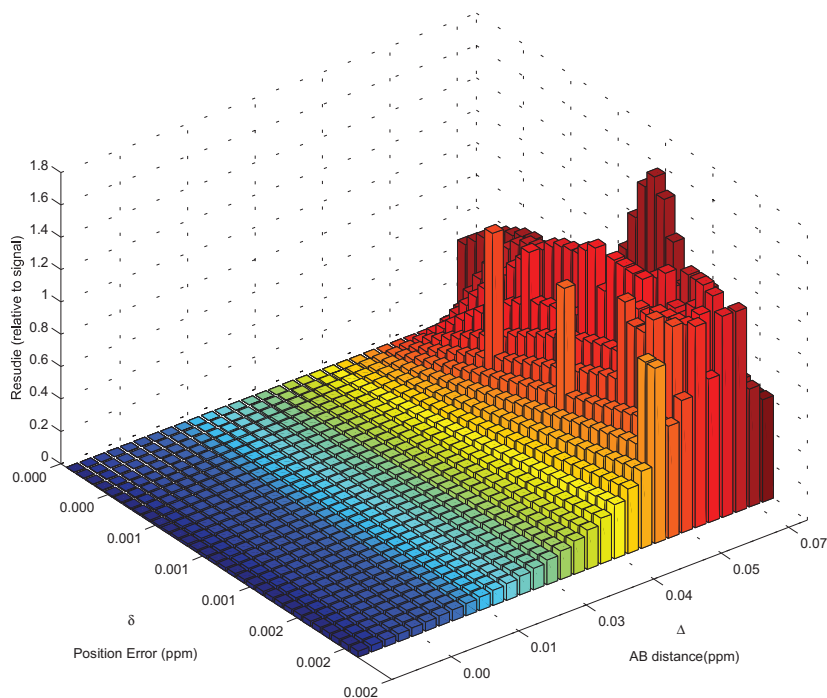


Figure 3.20: Relative mismatch of a strongly-coupled A_2B_2 signal and the shifted A_2B_2 signal after shifting correction of a group position as a function of the AB distance (in ppm) and the signal shift of A group position.

any common myo- or cerebral metabolite). These 3D-plots are more representative being plotted as maps, as presented in Fig. 3.21.

The necessity of using an optimization procedure (impossibility to do manual correction due to the great number of experiments) has some consequences. Firstly, some outliers (as, for example in Fig. 3.21- A_1B_1) are found. Secondly, maps of AB_m show non-monotone behavior in certain non expected regions.

But despite the presence of such artifacts, this simulation showed important results. By studying these maps, one can judge the possibility of using the discussed way of correction. It is easily seen that beyond a certain AB-distance, the error becomes negligibly small. For instance, for the A_2B_2 -case, presented in the Fig. 3.20, the relative error is less than 10% for all mismatches to be corrected if the distance between the groups is larger than 0.25ppm. Moreover, the data maps were acquired using the common HRMAS parameters, so one can estimate the expected errors for implementation of this method of correction.

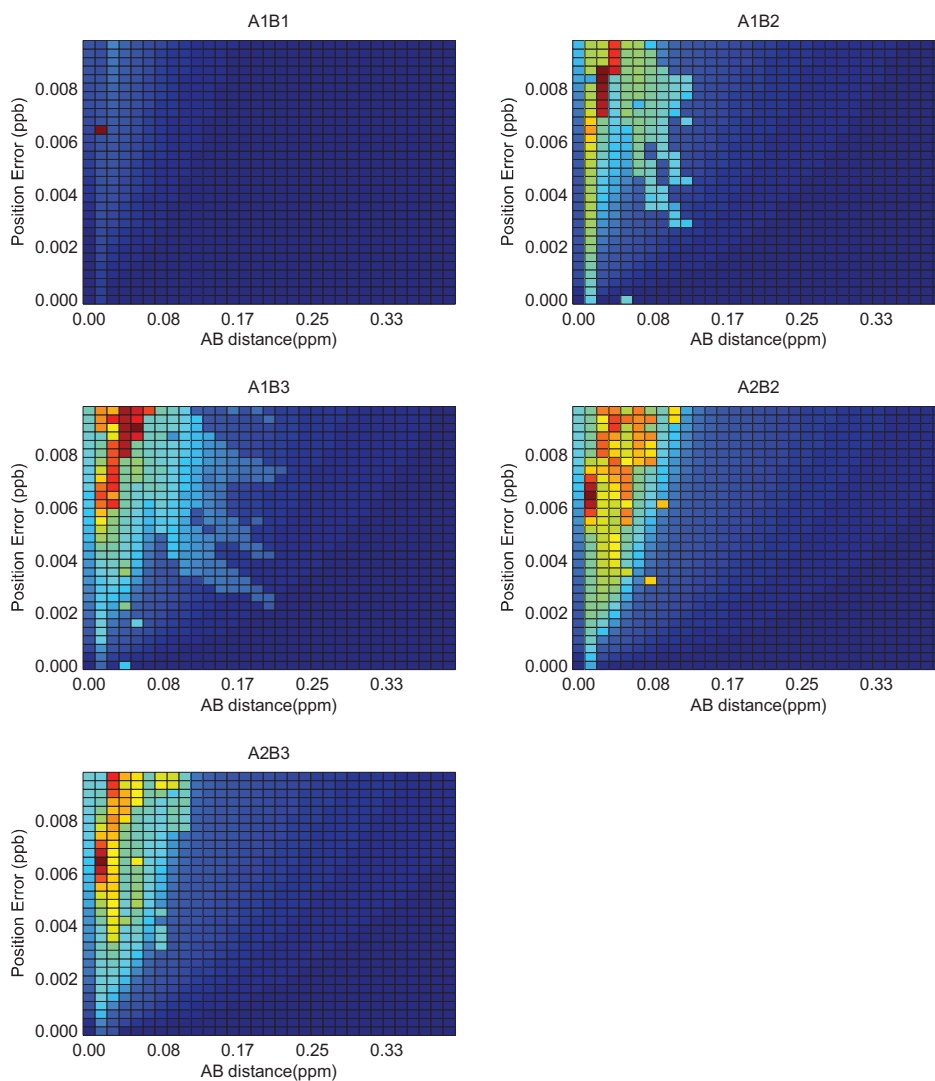


Figure 3.21: Strong coupling effect influence on the different structures. Maps of the residue related to the integral spectrum area. X-axis - AB distance in ppm, Y-axis - error corrected by the splitting algorithm.

Simple Correction of Chemical Shift Changes in Magnetic Resonance Spectroscopy Quantitation

A. Lazariiev*, M. Piotto^{†‡}, K. Elbayed[‡], I.J. Namer[‡], D. van Ormondt[§] and D. Graveron-Demilly*

*Laboratoire CREATIS-LRMN; CNRS UMR 5220; Inserm U630; Université Claude Bernard Lyon 1
Villeurbanne, France

Email: danielle.graveron@univ-lyon1.fr, Telephone: +33-472431049, Fax: +33-472448199

[†]Bruker BioSpin, Wissembourg, France

Email: martial.piotto@bruker.fr

[‡]Department of Biophysics and Nuclear Medicine, University Hospitals of Strasbourg, Strasbourg, France

Emails: elbayed@chimie.u-strasbg.fr, Izzie.Jacques.Namer@chru-strasbourg.fr

[§]Applied Physics; Delft University of Technology, Delft, Netherlands

Email: D.vanOrmondt@tudelft.nl

Abstract—High-resolution magic angle spinning (HRMAS) ¹H spectroscopy is playing an increasingly important role for diagnosis. This technique enables setting up metabolite profiles of *ex vivo* pathological and healthy tissue. Automatic quantitation of HRMAS signals provides reliable reference profiles to monitor diseases and pharmaceutical follow-up. Nevertheless, for several metabolites chemical shifts may slightly differ according to the micro-environment in the tissue or cells, in particular its pH. This hampers accurate estimation of the metabolite concentrations mainly when using quantitation algorithms based on a metabolite basis-set. In this work, we propose a user-friendly way to circumvent this problem based on stretching of the metabolite basis-set signals and maximization of the correlation between the HRMAS and basis-set spectra prior to quantitation.

Index Terms—Magnetic Resonance Spectroscopy, HRMAS, Quantitation, Stretching, jMRUI Software Package.

I. INTRODUCTION

High Resolution Magic Angle Spinning (HRMAS) Nuclear Magnetic Resonance (NMR) is playing an increasingly important role for diagnosis. This technique enables setting up metabolite profiles of *ex vivo* pathological and healthy tissue, *i.e.* biopsies [1–8]. Automatic quantitation of HRMAS signals [9–12] will provide reliable reference profiles to monitor diseases and pharmaceutical follow-up.

¹H HRMAS signals of *ex vivo* tissues contain several overlapping spectral components from many metabolites, see Fig 1. For several metabolites chemical shifts may slightly differ according to the micro-environment in the tissue or cells, in particular its pH, see *e.g.* [13]. This hampers accurate estimation of the metabolite concentrations mainly when using quantitation algorithms based on a metabolite basis-set [14–16]. The values of mismatch (around $2 \cdot 10^{-2}$ ppm) between the *ex vivo* signal and those of the metabolite basis-set are not significant for *in vivo* (low-resolution) spectroscopy. But for HRMAS spectroscopy, this increases the error of quantitation to tens percents for certain metabolites (*e.g.* Creatine, Lactate). In this work, we propose a user-friendly way to circumvent this problem based on stretching of some metabolite basis-set signals prior to quantitation with the method QUEST [15].

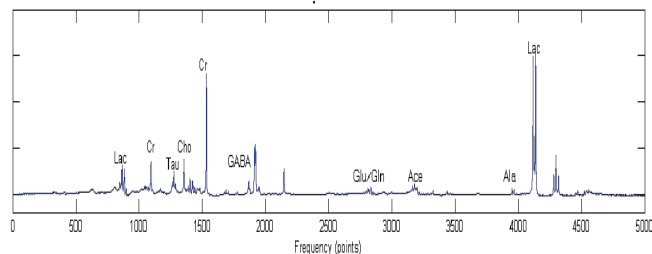


Fig. 1. Region of interest of an HRMAS spectrum from a tissue sample of a human brain with an oligodendroglioma, acquired at 11.7T.

II. METHOD

A. Quantitation with QUEST

In Magnetic Resonance Spectroscopy, quantitation based on a metabolite basis-set has become very popular in the last ten years. In this work, we used the method QUEST. The complex-valued time-domain model signal is written as a linear combination of the M weighted metabolite model \hat{x}^m – either quantum-mechanically simulated or *in vitro* measured – of the basis-set, see Fig. 2. The model samples, $\hat{x}_n, n = 1, 2, \dots, N$ where N is the number of data-points, can be written as

$$\hat{x}_n = \exp(i\phi_0) \times \sum_{m=1}^M a_m \hat{x}_n^m \exp[(\Delta\alpha_m + i\Delta\omega_m)t_n + i\Delta\phi_m], \quad (1)$$

where

- \hat{x}^m , m being a superscript, represents the metabolite basis-set signals.
- a_m are M amplitudes to be estimated. Note that these amplitudes represent the relative proportions of the M metabolites signals \hat{x}^m in the signal x rather than the amplitudes of individual spectral components.
- $\Delta\alpha_m, \Delta\omega_m, \Delta\phi_m$ represent small changes of the damping factors, angular frequencies, and phase shifts respectively. These changes – relative to the initial values in

the metabolite basis set – are included in the estimation procedure to automatically compensate for the effect of magnetic field inhomogeneities. Most often $\Delta\phi_m = 0$. Soft constraints on $\Delta\alpha_m$ and $\Delta\omega_m$ have been used in the minimization procedure,

- $t_n = nt_s + t_0, n = 1, 2, \dots, N$, are the sampling times, in which t_0 is the dead-time of the receiver – included in the estimation – and t_s the sampling interval,
- ϕ_0 is an overall phase, included in the estimation,
- $i^2 = -1$.

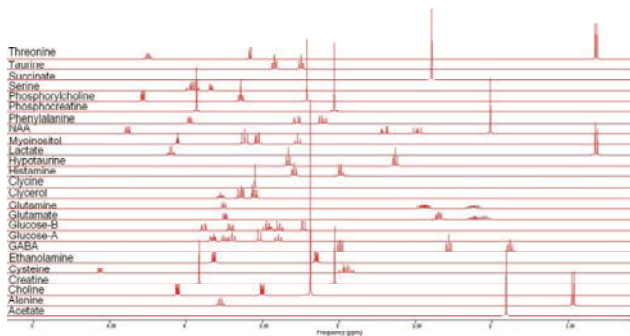


Fig. 2. Fourier transform of a metabolite basis-set at 11.7 Tesla, simulated by Quantum Mechanics with NMR-SCOPE for a one pulse sequence. This basis set was used in QUEST for quantitation of human brain HRMAS signals. Lorentzian lineshapes were used.

The basis-set signals were simulated directly in the time-domain with NMR-SCOPE which is based on the density matrix and product operator formalism [17]. Spin parameters were taken from [18]. We refined them. Twenty-five metabolites – Acetate (Ace), Alanine (Ala), Aspartate (Asp), Creatine (Cr), Choline (Cho), Cysteine (Cys), Ethanolamine (Eth), γ -amino-butyric acid (GABA), Glucose (Glc), Glutamate (Glu), Glutamine (Gln), Glycerol, Glycine (Gly), Histamine, Hypotaurine, Lactate (Lac), Myo-Inositol (mI), N-acetylaspartate (NAA), Phosphoryl-choline (PC), Phosphocreatine (PCr), Phenylalanine (Phe), Serine (Ser), Succinate (Suc), Taurine (Tau) – were included in the basis-set. Signals modelling the lipids (Lip) at 0.9 and 1.3 ppm were not included in the basis set, considering that their model function is insufficiently known. They are estimated with the background signal. The metabolite basis-set computed with NMR-SCOPE and used in QUEST for quantitation of human brain HRMAS signals is shown in Fig. 2. As mentioned above, for several metabolites chemical shifts can slightly differ according to the micro-environment in the tissue or cells, in particular with its pH. That means that prior knowledge based on the metabolite basis-set signals is not entirely correct anymore: the fingerprints of metabolites slightly differ.

B. Stretching of the Basis-set Signals

We propose to modify the metabolite basis-set signals \hat{x}^m , sensitive to pH changes, by just slightly stretching (expansion/contraction) the corresponding spectrum. This is illustrated in Fig. 3 where the lactate spectrum of the basis-set dif-

fers from the *ex vivo* spectrum. Stretching can easily be done by introducing frequency scaling parameters κ_m for given metabolites and by directly using the inverse scale change of the \hat{x}^m time-domain signals (frequency scale expansion results in contraction of the time scale and vice versa) [19]. This avoids splitting of the metabolite basis-set signals of given metabolites into basis sub-components according to chemical groups and adding appropriate constraints (prior knowledge) to the parameters of the groups. Either κ_m can be estimated prior to quantitation by maximizing the correlation between the *ex vivo* signal and \hat{x}^m or by introducing κ_m in the quantitation procedure.

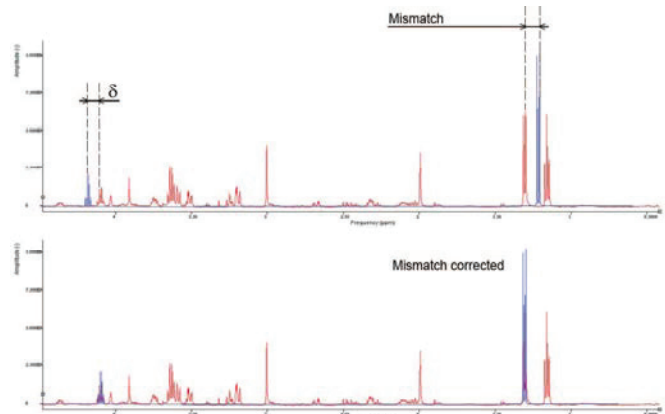


Fig. 3. Needs for stretching of a basis-set spectrum/signal. Red: oligodendrogloma spectrum, blue: Lactate basis-set spectrum presenting a mismatch with respect to the corresponding *ex vivo* Lactate spectrum.

C. Theoretical Limitations

Small changes of chemical shifts do not induce changes of J-coupling values, but the proposed method changes the distances between the peaks in each multiplet. This leads to some limitations of the stretching values with respect to the J-coupling values and distances between the peaks. A stretching coefficient can be defined as $\kappa = (\Delta + \delta)/\Delta$ where Δ corresponds to the distance between two multiplets (or to the difference between two chemical shifts) and δ corresponds to the frequency shift resulting from the stretching of the spectrum. Mathematically, the limitation can be expressed as $J(\kappa - 1) < \alpha/2$ where α represents the linewidth of a peak *i.e.* $J((\Delta + \delta)/\Delta - 1) < \alpha/2$ for the general case. This expression can easily be simplified to $\delta < \alpha\Delta/2J$. A schematic illustration of the stretching consequences on J coupling values is given in Fig. 4. It is then possible to establish the maximum limit of corrections possible, for instance, for Alanine, Aspartate, Citrate, Creatine and Lactate, see table I. These values were calculated for a resonance frequency of 500 MHz and a damping factor of 2 Hz corresponding to HRMAS spectra. As an example, for Alanine δ values less than or equal to 160 Hz can be corrected with the proposed method. For AB spectra, the distortions related to stretching of the signal became more important and lead to more significant

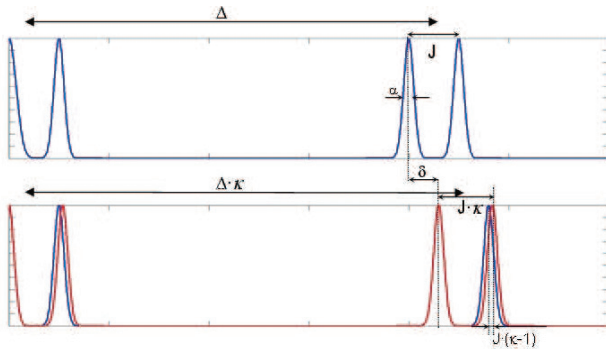


Fig. 4. Schematic illustration of the stretching consequences: Blue: spectrum with changeable peak positions, red: stretched spectrum.

limitations of the method's applications, *e.g.* the limit of correction of peak positions in the Citrate spectrum is about 4 Hz. Certain metabolites (*e.g.* Creatine) have no significant limitations as their proton groups are singlets. In this case, the artificial changes of the spectral linewidth is the only consequence of the method and it is insignificant.

TABLE I

METABOLITES WELL-PROCESSED WITH THE METHOD. THE LIMIT OF CORRECTION WAS CALCULATED FOR A RESONANCE FREQUENCY OF 500MHZ AND A DAMPING FACTOR OF 2HZ; - UNCOUPLED PROTONS.

| Metabolite | AX-distance, ppm | δ_{Max} in Hz/ppm |
|------------|------------------|--------------------------|
| Alanine | 2.31 | 160 / 0.32 |
| Aspartate | 1.09 | 31.3 / 0.063 |
| Citrate | 0.13 | 4.1 / 0.008 |
| Creatine | 0.89 | - |
| Lactate | 2.78 | 167 / 0.334 |

III. RESULTS

Thirty spectra from tissue samples of human brains with oligodendroglioma acquired at 11.7T, were quantitated. Mismatches between the basis-set signals and some *ex vivo* signals were observed. For the latter, adaptation of the basis-set signals was necessary. Results are illustrated on such a spectrum, see Fig. 1 and Fig. 5. Lactate, Creatine, Alanine and others signals of the basis set were automatically stretched prior to the quantitation procedure as mentioned above by maximizing the correlation between the *ex vivo* and basis-set signals. It can easily be seen that, thanks to stretching, one can independently "shift" the different multiplets of the spectrum. This enabled us to adapt artificially the chemical shifts variations due to pH. Stretching coefficients automatically found for Alanine, Creatine and Lactate for the considered oligodendroglioma and basis-set spectra, are given in Table II. Quantitation was significantly improved.

IV. CONCLUSION

We proposed a very simple method to account for chemical shift changes related to pH in quantitation methods based on a metabolite basis-set and used for quantitation of HR-MAS or high resolution signals. The method avoids splitting of the

TABLE II

STRETCHING COEFFICIENTS FOUND FOR ALANINE, CREATINE AND LACTATE FOR THE CONSIDERED OLIGODENDROGLIOMA AND BASIS-SET SPECTRA.

| Metabolite | Ala | Cr | Lac |
|------------------------|-------|-------|-------|
| Stretching coefficient | 1.004 | 0.994 | 1.009 |
| Shift, Hz | 5.9 | 0.55 | 0.32 |

metabolite basis-set signals of given metabolites into basis sub-components thus decreasing user involvement. It is well suited to improve quantitation of AB and AX spin systems like Citrate, Lactate, Creatine, Ethanolamine, etc.. For the more complicated spin systems like Glutamate and Glutamine, the method is less suited but still improves the quantitation results. A more accurate method based on Quantum Mechanics is under development and will be implemented in the version 4.x of the jMRUI software package [20].

ACKNOWLEDGMENT

The oligodendrogliomas spectra were acquired in the context of the CARMEN project. This work is supported by the EU Marie Curie Research Network, FAST, MRTNCT-2006-035801 and Philips Medical Systems, Best, NL.

REFERENCES

- [1] M. Swanson, A. Zektzer, Z. Tabatabai, J. Simko, S. Jarso, K. Keshari, L. Schmitt, P. Carroll, K. Shinohara, D. Vigneron, and J. Kurhanewicz, "Quantitative Analysis of Prostate Metabolites Using ^1H HR-MAS Spectroscopy," *Journal of Magnetic Resonance*, vol. 55, pp. 1257–1264, 2006.
- [2] B. Martinez-Granados, D. Monleón, M. C. Martinez-Bisbal, J. M. Rodrigo, J. del Olmo, P. Lluch, A. Ferrández, L. Martí-Bonmatí, and B. Celda, "Metabolite identification in human liver needle biopsies by high-resolution magic angle spinning ^1H NMR spectroscopy," *NMR in Biomedicine*, vol. 19, pp. 90–100, 2006.
- [3] T. Tsang, J. Griffin, J. Haselden, C. Fish, and E. Holmes, "Metabolic Characterization of Distinct Neuroanatomical Regions in Rats by Magic Angle Spinning ^1H Nuclear Magnetic Resonance Spectroscopy," *Magnetic Resonance in Medicine*, vol. 53, pp. 1018–1024, 2005.
- [4] M. A. Burns, J. L. Taylor, C. Wu, A. G. Zepeda, A. Bielecki, D. Cory, and L. L. Cheng, "Reduction of Spinning Sidebands in Proton NMR of Human Prostate Tissue With Slow High-Resolution Magic Angle Spinning," *Magnetic Resonance in Medicine*, vol. 54, pp. 34–42, 2005.
- [5] B. Sitter, S. Lundgren, T. F. Bathen, J. Halgunset, H. E. Fjosne, and I. S. Gribbestad, "Comparison of HR MAS MR spectroscopic profiles of breast cancer tissue with clinical parameters," *NMR in Biomedicine*, vol. 19, pp. 30–40, 2005.
- [6] M. C. Martínez-Bisbal, L. Martí-Bonmatí, J. Piquer, A. Revert, P. Ferrer, J. Llácer, M. Piotto, O. Assemet, and B. Celda, " ^1H and ^{13}C HR-MAS spectroscopy of intact biopsy samples ex-vivo and in-vivo ^1H MRS study of human high grade gliomas," *NMR in Biomedicine*, vol. 17, pp. 191–205, 2004.
- [7] F. Fauvelle, G. Testylier, H. Rabeson, D. Graveron-Demilly, A. Foquin, P. Carpentier, and F. Dorandeu, "Evolution of Mouse Brain Metabolism after a Convulsive Dose of Soman," in *Proc. Int. Soc. Magnetic Resonance in Medicine and European Society for Magnetic Resonance in Medicine and Biology, Joint Annual Meeting ISMRM-ESMRMB*, Berlin, Germany, 2007, p. 2416.
- [8] J. Lindon, O. Beckonert, E. Holmes, and J. Nicholson, "High-resolution magic angle spinning NMR spectroscopy: Application to biomedical studies," *Prog. Nucl. Magn. Reson. Spectrosc.*, vol. 55, pp. 79–100, 2009.
- [9] H. Rabeson, F. Fauvelle, H. Ratiney, C. Cudalbu, D. van Ormondt, and D. Graveron-Demilly, "Quantitation with QUEST of ER-Filtered HRMAS-NMR Signals," in *ProRISC, IEEE Benelux*, Veldhoven, The Netherlands, November 2006, pp. 171–175.

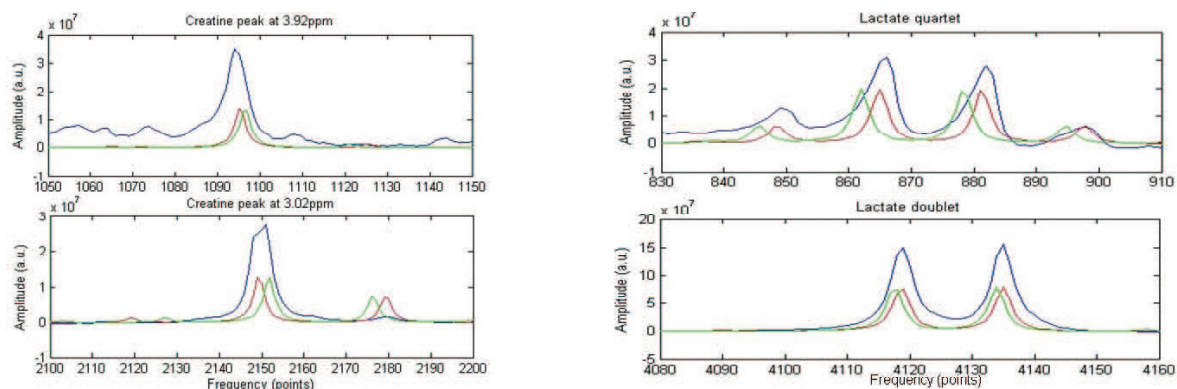


Fig. 5. Zoom in on the multiplets of Creatine and Lactate spectra. Blue: oligodendrogloma spectrum; green: initial basis-set spectrum and red: corrected (stretched) basis-set spectrum of Cr and Lac respectively.

- [10] H. Rabeson, H. Ratiney, F. Fauvelle, C. Cudalbu, S. Cavassila, A. Bernard, P. Giraudon, D. van Ormondt, and D. Graveron-Demilly, "Quantitation for *in vivo* and *ex vivo* NMR Spectroscopy," *Journal of Optoelectronics and Advanced Materials*, vol. 9, pp. 505–511, 2007.
- [11] G. Reynolds, M. Wilson, A. Peet, and T. N. Arvanitis, "An Algorithm for the Automated Quantitation of Metabolites in *in Vitro* NMR Signals," *Magn. Reson. Med.*, vol. 56, pp. 1211–1219, 2006.
- [12] F. DiCamillo, M. Swanson, J. Kurhanewicz, and S. Nelson, "Assessment of HR-MAS NMR Prostate Biopsy Tissue Spectra with Principal Component Analysis and Metabolite Quantification," in *Proc. Int. Soc. Magnetic Resonance in Medicine, 14th Scientific Meeting and Exhibition*, Seattle, USA, 2006, p. 2534.
- [13] M. van der Graaf and A. Heerschap, "Effect of cation binding on the proton chemical shifts and the spin-spin coupling constant of citrate," *Journal of Magnetic Resonance Ser. B*, vol. 112, pp. 58–62, 1996.
- [14] S. Provencher, "Estimation of Metabolite Concentrations from Localized *In Vivo* Proton NMR Spectra," *Magnetic Resonance in Medicine*, vol. 30, pp. 672–679, 1993.
- [15] H. Ratiney, M. Sdika, Y. Coenradie, S. Cavassila, D. van Ormondt, and D. Graveron-Demilly, "Time-Domain semi-parametric estimation based on a metabolite basis set," *NMR in Biomedicine*, vol. 18, pp. 1–13, 2005.
- [16] H. Ratiney, Y. Coenradie, S. Cavassila, D. van Ormondt, and D. Graveron-Demilly, "Time-Domain Quantitation of ^1H Short Echo-Time Signals: Background Accommodation," *Magn Reson Mater Phy*, vol. 16, pp. 284–296, 2004.
- [17] D. Graveron-Demilly, A. Diop, A. Briguet, and B. Fenet, "Product-operator algebra for strongly coupled spin systems," *Journal of Magnetic Resonance*, vol. A101, pp. 233–239, 1993.
- [18] V. Govindaraju, K. Young, and A. Maudsley, "Proton NMR chemical shifts and coupling constants for brain metabolites," *NMR in Biomedicine*, vol. 13, pp. 129–153, 2000.
- [19] E. O. Brigham, *The Fast Fourier Transform*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1974.
- [20] D. Stefan, F. DiCesare, A. Andrasescu, E. Popa, A. Lazariev, E. Vescovo, O. Strbak, S. Williams, Z. Starcuk, M. Cabanas, D. van Ormondt, and D. Graveron-Demilly, "Quantitation of Magnetic Resonance Spectroscopy signals: The jMRUI software package," *Measurement Science and Technology*, vol. 20, 2009.

Chapter 4

Optimization of metabolite basis-set signals based on a Quantum Mechanical approach

4.1 Introduction

In parallel to our work presented in Ch. 3, Giskeødegård et al. have recently studied (*Analytica Chimica Acta*, September 2010) the problem of the difference of spectra of the same chemical compounds acquired at different conditions, trying to correct mismatches between the spectra [4]. Developed methods were tested on tissues of cervical, breast and colon cancers. All five presented methods are called *warping* methods, meaning that the signal needs to be distorted in a certain way to get the necessary correction. This distortion implies a one-directional irreversible mapping of the frequencies on the new ones so as to improve the optimization criterion (correlation, in majority). Some of the methods use the stretching optimization that we previously reported in [79], others are new, showing promising possibilities. In *in vivo* MR Spectroscopy, chemical shift recalibration based on linear interpolation was recently proposed [80]. In this chapter, we will present a new method enabling to correct such mismatches without distorting the signal and keeping all the metabolite fingerprints correct.

4.2 Method

As we already pointed out above, both presented methods – stretching and splitting – still lead to some signal distortions and, sometimes, important limitations of implementation. Knowing the origin of signal distortion, one can arrive at the conclusion that the only way to get a non-distorted signal is to resimulate it using Quantum Mechanics with new parameters. This is the optimization algorithm, presented below.

It should be remarked that signal simulation based on Quantum Mechanics is a rather resource-intensive procedure. Consequently, not all the methods of signal simulation are applicable. In the optimization procedure, the conventional Quan-

tum Mechanical approach [5] is used to speed up the calculations. The method is based on the time-independent Schrödinger equation

$$\mathcal{H}\Psi = E\Psi \quad (4.2.1)$$

where Ψ are the wave functions and E the energy levels of \mathcal{H} . The frequencies and corresponding transition probabilities (thus: amplitudes) of the spectral components are obtained after diagonalization of \mathcal{H} which is taken equal to the strong coupling spin Hamiltonian (see Eq. 1.1.22).

Taking into account that the differences of the eigenvalues (energy) provide the frequencies, the final output of the method will be a $2 \times n_f$ -matrix of frequencies and their relative amplitudes (n_f represents the number of frequencies). A simple example of results for an AB_3 spectrum is presented in Tab. 4.1.

| i_f | Frequency, ppm | Amplitude, a.u. |
|-------|----------------|-----------------|
| 1 | 1.459399 | 2.981021 |
| 2 | 3.752972 | 1.018979 |
| 3 | 3.767391 | 0.807593 |
| 4 | 3.767391 | 0.198674 |
| 5 | 1.459445 | 0.063201 |
| 6 | 1.459445 | 0.930532 |
| 7 | 3.767391 | 0.198674 |
| 8 | 3.767391 | 0.807593 |
| 9 | 1.459445 | 0.930532 |
| 10 | 1.459445 | 0.063201 |
| 11 | 3.767527 | 1.006088 |
| 12 | 1.459446 | 3.974933 |
| 13 | 1.473954 | 3.018979 |
| 14 | 1.473909 | 0.963001 |
| 15 | 1.473909 | 0.043266 |
| 16 | 1.473909 | 0.043266 |
| 17 | 1.473909 | 0.963001 |
| 18 | 3.781855 | 0.993722 |
| 19 | 3.781855 | 0.993722 |
| 20 | 1.473909 | 4.025067 |
| 21 | 3.781990 | 0.993558 |
| 22 | 1.459491 | 2.981375 |
| 23 | 1.473865 | 3.018625 |
| 24 | 3.796364 | 0.981375 |

Table 4.1: An example of the output data from the quantum mechanics conventional approach simulation. $n_f = 24$.

The value n_f is the total number of frequencies provided by the algorithm. For the most general case the number of frequencies is strictly defined by the number of spins. But for real implementation, some frequencies have the same value (for

instance, the values ##14-17,20 in the example), so their intensities should be added, also some of them have negligible intensities (peaks 5,10,15-16 in Tab. 4.1) and they may be ignored.

The next step is to create the theoretical signal. Using the general form which corresponds to Eq. 1.4.1 one obtains:

$$S = \sum_{i_f=1}^{n_f} S_{i_f} = \sum_{i_f=1}^{n_f} A_{i_f} \exp((2\pi i f_i - \alpha)t) \quad (4.2.2)$$

where A_{i_f} and f_i are the respective amplitude and frequency outputs of the Quantum Mechanics algorithm and α is an artificial decaying factor. The signal in this form is considered to have a Lorentzian decay but any decay can be imposed as well. As the simulation corresponds to a 90°-pulse excitation, the signal is assumed to have the zero order phase equal to zero for all spectral components.

4.3 Computing time costs

The importance of computing time and/or quality of the algorithm becomes evident after the investigation of the computing time costs for the different steps of the algorithm. Three main steps should be reviewed separately:

- Hamiltonian matrix representation and its diagonalization,
- Matrix-product calculation to compute the transition probabilities (amplitudes),
- Signal calculations (Eq. 4.2.2).

For the first investigation, the matrix implementation of the algorithm in MATLAB was launched for different spin systems to be simulated. Three “step times” were recorded independently. Simulation parameters were chosen as usual mimicking HRMAS signals, the most important of them for this experiment – number of points in the FID – was chosen to 16k points. These results are presented in Fig. 4.1.

Some conclusions can be set after analyzing these data. Firstly, it is easily seen that the time, spent for Hamiltonian diagonalization is negligible with respect to other times. Secondly, the time costs for the computation of transitions between the energy levels and signal construction have different relative values for different spin systems.

Now, considering the behaviour the time function as a function of N_{spins} , we will ignore the values of:

- t_m – the time of double matrix multiplication for the transition,
- and t_{FID} – the time of N_{points} signal construction according to Eq. 1.1.11.

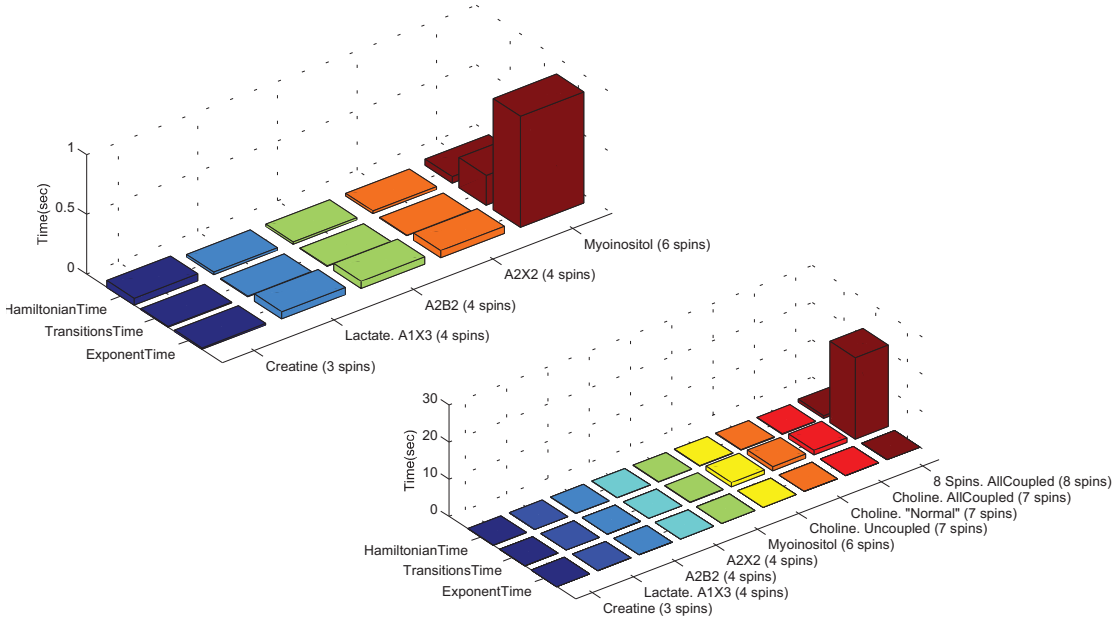


Figure 4.1: Computing times of simulation for the various steps of the Quantum Mechanics algorithm. Top-left – small spin systems, max.time ≈ 1 s; bottom-right – up to eight spin systems, max.time $\approx 10^2$ s.

The number of transition to be calculated increases with the Hamiltonian matrix size. An N_{spins} -spin system has a Hamiltonian matrix of $2^{N_{\text{spins}}} \times 2^{N_{\text{spins}}}$ size. Thus, this value will be a determinant parameter for the time needed for the transition calculation. Certainly, it will not be equal to the exact multiplication

$$t_{\text{tr}} = 2^{2N_{\text{spins}}} t_m \quad (4.3.1)$$

as it would be in the simplest algorithm implementation. There are several ways to decrease the number of calculated transitions. First, the Hamiltonian matrix and the complete transition matrix are symmetric so there is no need to calculate transitions from both sides of the diagonal. This decreases the total number of operations by a factor of two. Secondly, some changes in the order of Hamiltonian construction will make possible the block-by-block calculation of transitions. To do this, the Hamiltonian matrix must be arranged so that to increase the z component of the basis vectors as was done in [81].

This will be explained in more details for a three-spin system. Using the general form (Eq. 1.1.21) to construct a Hamiltonian one obtains a matrix of a form given in Eq. 4.3.2.

$$\mathcal{H} = \frac{1}{2} \begin{pmatrix}
 (\delta_1 + \delta_2 + \delta_3) + \frac{1}{2}(J_{12} + J_{23}) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & (\delta_1 + \delta_2 - \delta_3) + \frac{1}{2}(J_{12} - J_{23}) & J_{23} & J_{13} & 0 & 0 & 0 & 0 & 0 \\
 0 & J_{23} & (\delta_1 - \delta_2 + \delta_3) + \frac{1}{2}(-J_{12} + J_{13} - J_{23}) & J_{12} & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & (\delta_1 - \delta_2 - \delta_3) + \frac{1}{2}(-J_{12} - J_{13} + J_{23}) & 0 & J_{12} & J_{13} & 0 & 0 \\
 0 & 0 & 0 & 0 & (-\delta_1 + \delta_2 - \delta_3) + \frac{1}{2}(-J_{12} - J_{13} + J_{23}) & 0 & 0 & 0 & 0 \\
 0 & J_{13} & J_{12} & 0 & 0 & (-\delta_1 + \delta_2 - \delta_3) + \frac{1}{2}(-J_{12} + J_{13} - J_{23}) & 0 & 0 & 0 \\
 0 & 0 & 0 & J_{12} & 0 & 0 & (-\delta_1 + \delta_2 - \delta_3) + \frac{1}{2}(-J_{12} + J_{13} - J_{23}) & J_{23} & 0 \\
 0 & 0 & 0 & 0 & J_{12} & 0 & 0 & (-\delta_1 - \delta_2 + \delta_3) + \frac{1}{2}(J_{12} - J_{13} - J_{23}) & 0 \\
 0 & 0 & 0 & 0 & 0 & J_{13} & 0 & 0 & (-\delta_1 - \delta_2 - \delta_3) + \frac{1}{2}(J_{12} + J_{23})
 \end{pmatrix}$$

(4.3.2)

This form has no convenient block structure. And its basis vector order is as follows:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad (4.3.3)$$

The basis vectors are obtained by direct product expansion of the $|+\rangle$ and $|-\rangle$ kets introduced in Eq. 1.1.25 in the convention that

$$\begin{aligned} |+\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ |-\rangle &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned} \quad (4.3.4)$$

To overview the general structure of the Hamiltonian, we will introduce the following notation in the Hamiltonian matrix: \mathbf{D} – diagonal element, \mathbf{N} – non diagonal element. So the form is as presented in Tab. 4.2.

| | $ +++ \rangle$ | $ ++- \rangle$ | $ +-+ \rangle$ | $ +-- \rangle$ | $ -++ \rangle$ | $ -+- \rangle$ | $ - - + \rangle$ | $ - - - \rangle$ |
|-----------------|----------------|----------------|----------------|----------------|----------------|----------------|------------------|------------------|
| $\langle +++ $ | \mathbf{D} | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $\langle ++- $ | 0 | \mathbf{D} | \mathbf{N} | 0 | \mathbf{N} | 0 | 0 | 0 |
| $\langle +-+ $ | 0 | \mathbf{N} | \mathbf{D} | 0 | \mathbf{N} | 0 | 0 | 0 |
| $\langle +-- $ | 0 | 0 | 0 | \mathbf{D} | 0 | \mathbf{N} | \mathbf{N} | 0 |
| $\langle -++ $ | 0 | \mathbf{N} | \mathbf{N} | 0 | \mathbf{D} | 0 | 0 | 0 |
| $\langle -+- $ | 0 | 0 | 0 | \mathbf{N} | 0 | \mathbf{D} | \mathbf{N} | 0 |
| $\langle --+ $ | 0 | 0 | 0 | \mathbf{N} | 0 | \mathbf{N} | \mathbf{D} | 0 |
| $\langle --- $ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | \mathbf{D} |

Table 4.2: Schematic view of a non-arranged three-spin system Hamiltonian.

Arranging the basis function order according to the sum of states, one obtains the block-by-block structure of the Hamiltonian presented in Tab. 4.3 [81].

After this transformation, it could easily be seen (either from the matrix representation of the transition operator, or from the rules of the allowed transitions (Eq. 1.1.29)) that the transitions with a non-zero probability could appear only between two consecutive blocks. This important fact allows to decrease the number of matrix multiplication operations to be done during the transition probability calculation.

The computing time gains can be estimated taking into consideration, that the number of elements in the Hamiltonian increases as $2^{N_{\text{spins}}}$ (accordingly to Eq. 4.3.1), and the block diagonal in the arranged Hamiltonian corresponds to an

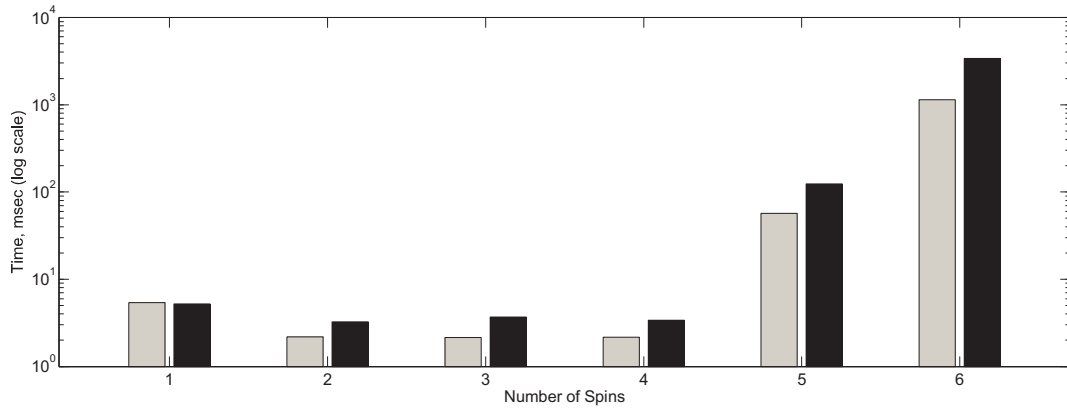


Figure 4.2: Comparison of the simulation time for non-arranged (black) and block-by-block (gray) arranged Hamiltonians.

4.4 Algorithm implementation

4.4.1 Gradient descent method

A gradient descent method [83] for a function F depending on an argument vector x with a descent coefficient λ

$$\vec{x}^{[j+1]} = \vec{x}^{[j]} - \lambda^{[j]} \nabla F(\vec{x}^{[j]}) \quad (4.4.1)$$

was chosen to optimize the peak positions of the metabolite basis-set. The steepest descent condition

$$\lambda^{[j]} = \operatorname{argmin}_{\lambda} F(\vec{x}^{[j]} - \lambda^{[j]} \nabla F(\vec{x}^{[j]})) \quad (4.4.2)$$

has not been chosen due to the very significant related computing time costs. So the fractional step was taken. This means that λ was decreased by certain factor k at each the step:

$$\begin{aligned} \lambda^{[j+1]} &= \frac{\lambda^{[j]}}{k}, \\ \lambda^{[j]} &= k^{-j} \lambda^{[0]}. \end{aligned} \quad (4.4.3)$$

4.4.2 Cost function

There were two possibilities to choose the cost function:

- sum of squares of point-by-point signal differences between the signal to be analyzed and the estimated signal;
- signal correlation function of the signal to be analyzed and the estimated signal

The least squares method often converged to local minima. Unfortunately, the correlation method has this risk at considerable distances from the cost function

minimum too. But the sum of squares criterion function showed this behaviour even at the very epsilon of the sought-for minimum point due to the beating effect at certain ratio of the frequencies of periodical signals. This effect, simulated for the lactic acid HR-NMR spectra phantom, is presented in Fig. 4.3.

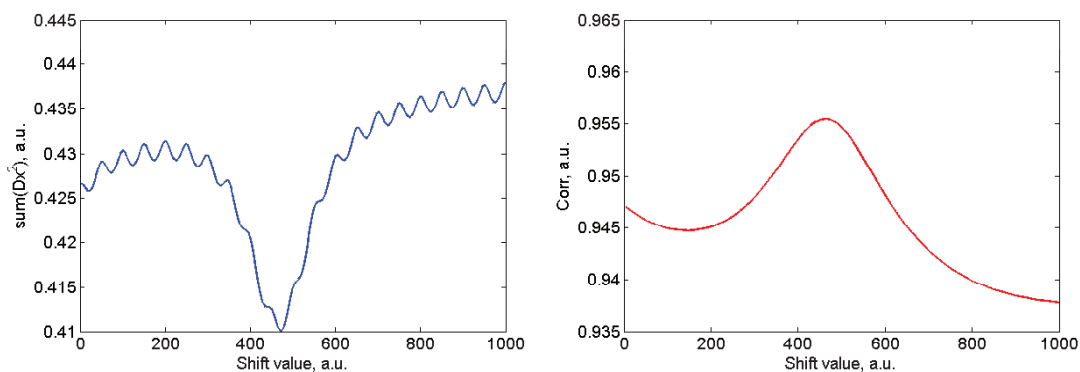


Figure 4.3: Oscillating behaviour of the least squares difference function due to the beating effect in comparison to the monotone character of the correlation function as a function of the shift value of one simulated signal respectively to the other one.

So the signal correlation function was chosen as the cost function.

4.4.3 Descent coefficient

Working in the mode of decreasing the descent coefficient λ (from Eq. 4.4.1) it is always difficult to choose the necessary initial value and final condition. If for the exit condition everything is determined by the desired precision of optimization

$$|\vec{x}^{[j+1]} - \vec{x}^{[j]}| < \varepsilon \quad (4.4.4)$$

then for the starting value of λ it is always good to find the optimum because for higher values one risks to miss the sought-for minimum and for the lower ones the number of iterations before reaching the exit condition increases.

Actually, the optimization algorithm was tested as pre-processing method. Theoretically, nothing hampers its implementation as a part of a quantitation method. But the number of degree of freedoms will increase significantly leading to time costs incompatible with normal times expected for a quantitation procedure.

Ex altera parte, the algorithm being described used as preprocessing *puts all the spectral groups at their correct positions corresponding to those observed in the quantitated spectrum*. In turn, this means that the correction value $\Delta\omega_m$ in Eq. 1.4.1 is no more necessary, so it will decrease by one the number of degrees of freedom for each signal during the QUEST-procedure. Of course, this is not an advantage of quality in the most general case. The consequences will be discussed.

4.4.4 Optimization loop

The optimization loop is schematically presented in Fig. 4.4. The main steps are:

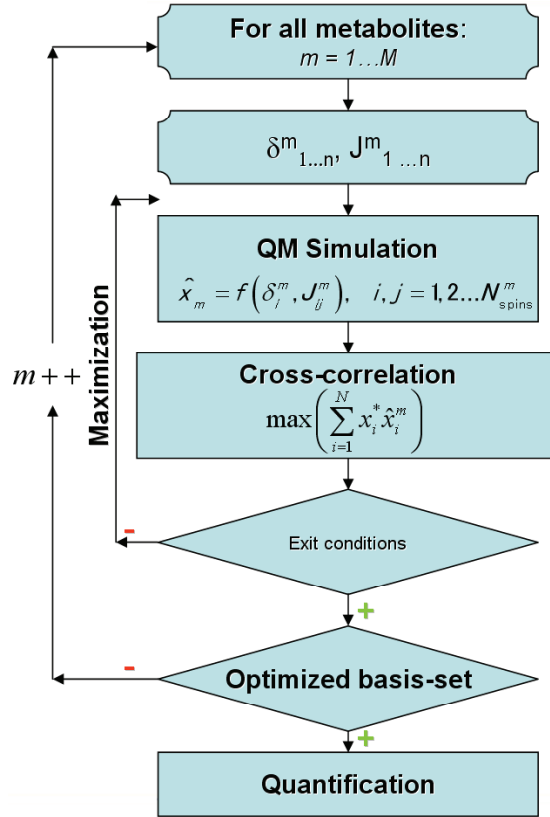


Figure 4.4: The optimization algorithm based on Quantum Mechanics resimulation.

- Simulation with NMR-SCOPE [52, 84] of the basis-set $\{\hat{x}^m, m = 1, 2, \dots, M\}$ using the initial spin parameters $\{\delta_i^m$ and $J_{ij}^m, i, j = 1, 2, \dots, N_{\text{spins}}^m\}$.
- Selection of metabolites that are subject to pH changes leading to M_{pH} metabolite signals to be optimized.
- For each of the M_{pH} metabolites:
Quantum Mechanical calculations and maximization of $\text{corr}(x, \hat{x}^m) \implies \{\delta_{i,\text{opt}}^m, i = 1, 2, \dots, N_{\text{spins}}^m\}$ and \hat{x}_{opt}^m
- Simulation with NMR-SCOPE of the optimized basis-set $\{\hat{x}_{\text{opt}}^m, m = 1, 2, \dots, M_{\text{pH}}\}$ using the spin parameters $\{\delta_{i,\text{opt}}^m$ and $J_{ij}^m\}, i, j = 1, 2, \dots, N_{\text{spins}}^m\}$.
- QUEST quantitation using the optimized basis-set $\{\hat{x}_{\text{opt}}^m, m = 1, 2, \dots, M\}$.

4.5 Results

The proposed method was first tested on simulated signals mimicking HRMAS-NMR signals and then on the real HRMAS-NMR signals.

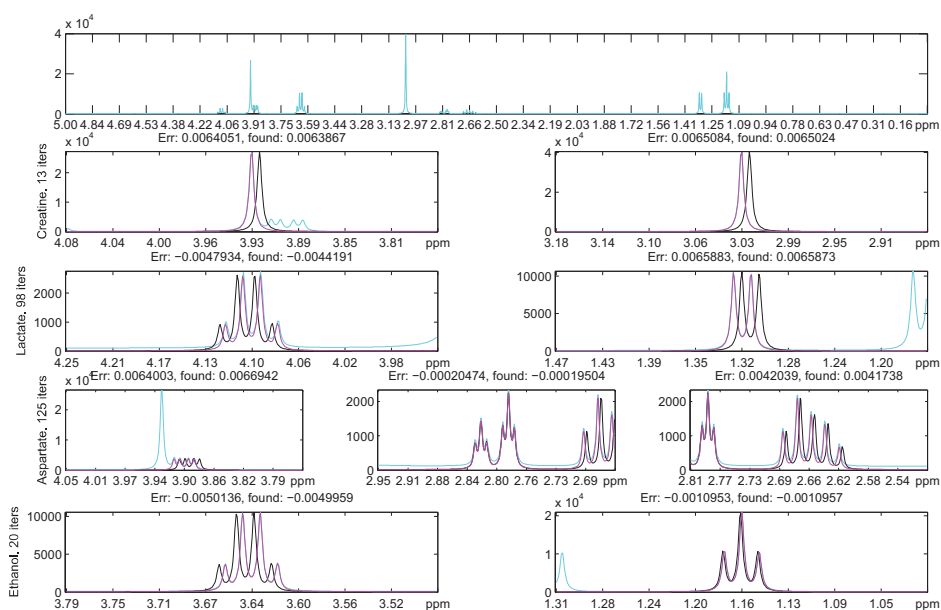


Figure 4.5: Optimization of the group positions for an HRMAS spectrum phantom. Legend: cyan – phantom signal, black – initial basis-set metabolite positions, magenta – optimized spectrum positions.

4.5.1 Simulated signals

Thirty three spectra from tissue samples of human brains with oligodendroglioma acquired at 11.7T (sampling interval: 0.143ms, $N = 16384$) were quantitated. The relevant part of the spectra was first extracted with ER-Filter [85] as mentioned in [17]. Then, each chemical shift was shifted with a random value limited to 0.01ppm (taken after overview of the series of HRMAS NMR signals).

Results are illustrated for an oligodendroglioma simulated spectrum of the series in Fig. 4.5. Lac, Cr, Asp, Eth and others signals of the basis set were automatically optimized prior to the quantitation procedure as mentioned above by maximizing the correlation between the *ex vivo* and basis-set signals. An Ethanol spectrum (A_2X_3) is also included despite its weak sensitivity to the pH of the environment.

The QM-method enables independent movements of the different multiplets of the spectrum, keeping all strong-coupling effects, contrary to methods which decompose the basis-set signals into sub-components. This enabled us to adapt the chemical shift variations due to simulated pH.

Series of simulations showed that with usual shift values (0.01ppm) the simulated AX, AMX spectra matched well the phantom ones. But even in these cases there were some conditions where a secondary minimum¹ was chosen by the optimization algorithm.

¹When we say “minimum” it has its general meaning in the optimization sense. As was shown before, here a “minimum” is, in fact, the maximum of correlation between the simulated and HR MAS signals.

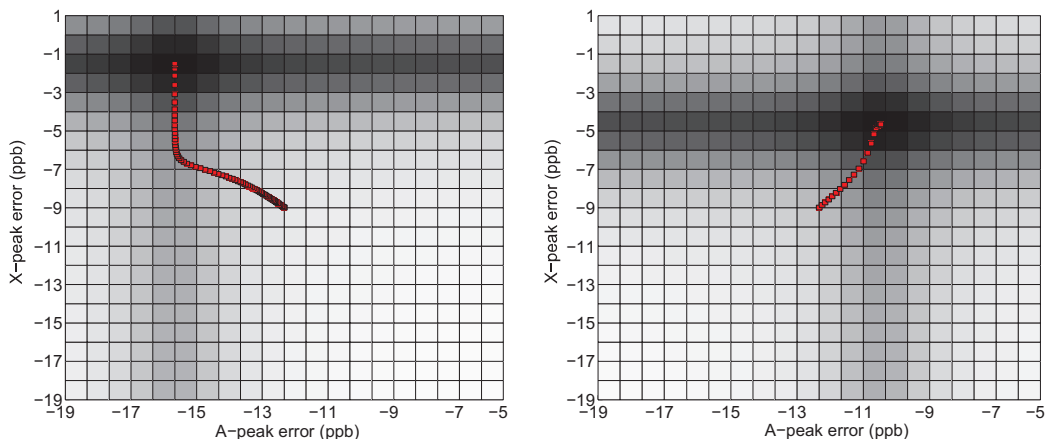


Figure 4.6: Gradient descent path for two simulated creatine signals. Colour intensity corresponds to the cost function value. Red – iterative values of the correction vector $\Delta\sigma_A, \Delta\sigma_X$

To interpret the results, a series of more detailed simulations was set. For the 2D-mode (the optimization problem with two degrees of freedom for AX-spectra) map plots were built for some simulations. Firstly a simulated AX-signal with an artificial displacement of both group positions was simulated. After this, the maps of the cost function were built in the ranges

$$\delta_{A,X} = \pm\Delta\delta_{\max} \quad (4.5.1)$$

where $\Delta\delta_{\max}$ is the maximum displacement expected – parameter limits for the optimization problem. Having launched the optimization algorithm with printing the parameter values at each iteration, one can follow the iterative trace of the run (see Fig. 4.6).

Optimization of a Lactate (AX_3) spectrum shows more interesting behaviour and needs additional analysis. The cost function map was built for ± 0.02 ppm shift ranges (for both groups) for a 500MHz resonance frequency. A gradient descent with a constant gradient feedback parameter was launched twice ($\lambda_1/\lambda_2 = 1/5$). For the low value of λ the algorithm missed the optimum point and for the high one it has diverged. The traces are presented in Fig. 4.7. For these simulations the errors were set manually (non-randomly) between -0.01 ppm and 0.015 ppm.

The reason of the appeared problem can be explained with a simple 1D example. Let us imagine that we need to find the minimum of the potential plotted in Fig. 4.8. This is not just an “any” smooth function with a minimum. Very important is the behaviour of its first derivative. At the initial point the derivative is low, it is high on the way to the optimal point and it is (as is normal) zero at the minimum. Choosing a gradient feedback for descent one will face the problem:

- low and constant values of the feedback coefficient hamper arrival at the minimum point within an adequate amount of time;
- high and constant values enable to start with an appreciable “speed” of approaching but lead to a risk of overshoot the optimal positions;

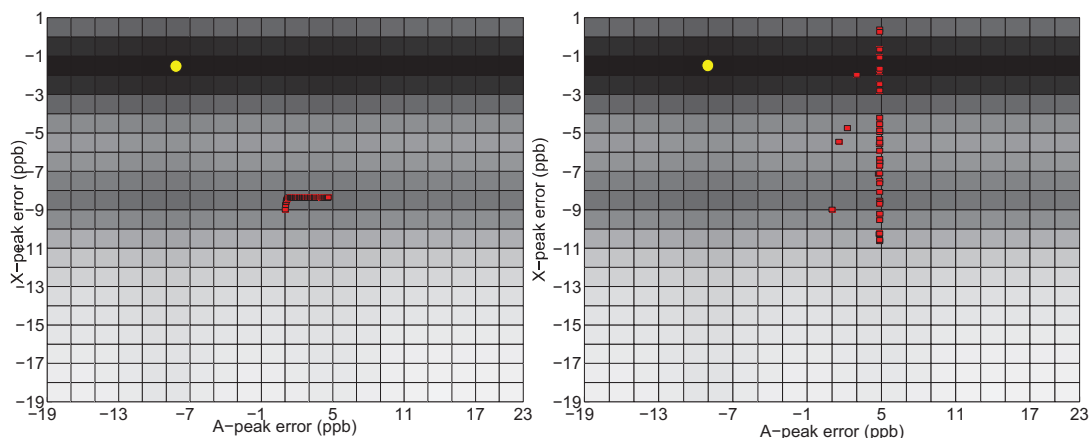


Figure 4.7: Gradient descent run for a Lactic acid spectrum simulation. Left: $\lambda = 30$ a.u., right: $\lambda = 150$ a.u.. The yellow marks signify the absolute optimum positions.

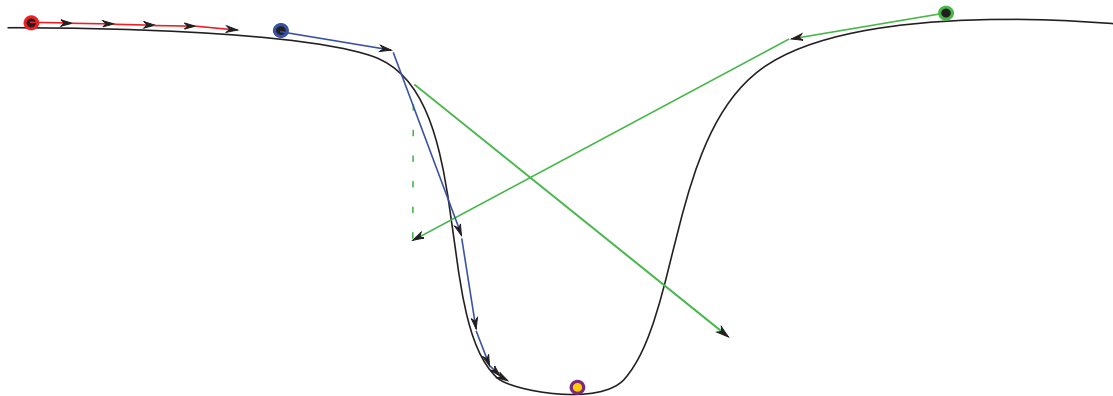


Figure 4.8: Different behaviours of the minimization algorithms for different λ . Red – low value, green – high value, blue – decreasing value, yellow/magenta – sought-for point.

- finally, the decreasing feedback coefficient must decrease enough to avoid the previous disadvantage but in this case it will not work properly in the region of the cost function around the optimum, where the gradient has a small absolute value.

The fact, that Lorentzian functions, very common for lineshapes in HR NMR, are subject to the problems described above, hampers successful optimization. Even for a case of steepest descent, the section of 2D map in the gradient direction rather often will resemble the schematic example in Fig. 4.8.

Some promising results (especially in time costs) were obtained with a decrement coefficient equal to 2 ($k = 2$ in Eq. 4.4.3). Results can be seen in Fig. 4.9.

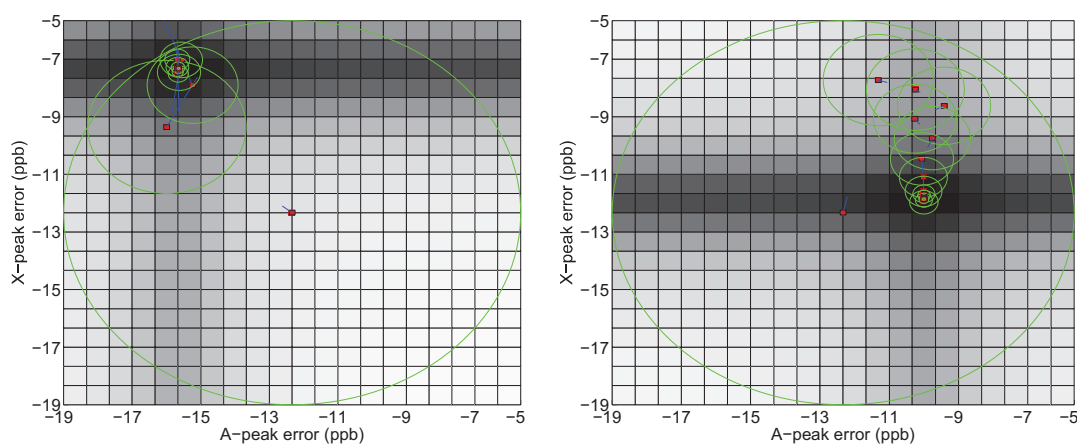


Figure 4.9: Creatine peak position optimization by the steepest descent. Decreasing step is used with the ratio coefficient of 2 (see Eq. 4.4.3). The green circles show the λ decrementation at each iteration.

4.5.2 Real spectra. Oligodendrogloma

Thirty three spectra of the brain biopsies were used to test the method on real HRMAS NMR signals. Keeping the parameters of the simulation the same as in the test on the phantom spectra, optimization was launched for twenty five metabolites: Acetate (Ace), Alanine (Ala), Aspartate (Asp), Creatine (Cr), Choline (Cho), Cysteine (Cys), Ethanolamine (Eth), γ -amino-butyric acid (GABA), Glucose (Glc), Glutamate (Glu), Glutamine (Gln), Glycerol, Glycine (Gly), Histamine (Hist), Hypotaurine (HTau), Lactate (Lac), Myo-Inositol (mI), N-acetyl-aspartate (NAA), Phosphoryl-choline (PC), Phosphocreatine (PCr), Phenylalanine (Phe), Serine (Ser), Succinate (Suc), Taurine (Tau) and Ethanol (Eth) as a trace of the biopsy procedure. Spin parameters were initially taken from [26] and we refined them to fit rodent brain HRMAS spectra [17].

Several series of optimizations showed that for some of the metabolites listed above it was difficult to optimize their chemical shifts. These are the metabolites of weak concentrations, having their fingerprint peaks in the region of 3.2-3.9 ppm, tightly covered with different spectral compounds. For instance, these weakly recognized groups are: Choline multiplet at 3.5ppm, Ethanolamine multiplet at 3.8ppm, Glucose and Glycerol peaks. On the other hand, for example, myo-Inositol spectrum, despite being into the mentioned region, is often recognizable due to its high relative concentration in the tissues. Some important problems also appeared due to the impossibility to distinguish Choline, Phosphocholine and Glycerophosphocholine peaks at 3.2ppm, the same applies to Creatine and Phosphocreatine peaks at 3.0ppm and 3.9ppm.

So now let us consider well optimized metabolites. It can easily be seen in Fig. 4.10 that the mismatches between the basis-set and HRMAS spectra have been reduced for Cr, Lac, Eth. Asp multiplets at 2.80 and 2.65 ppm have amplitudes of the same order as the noise level which makes the recognition of the peak positions more difficult.

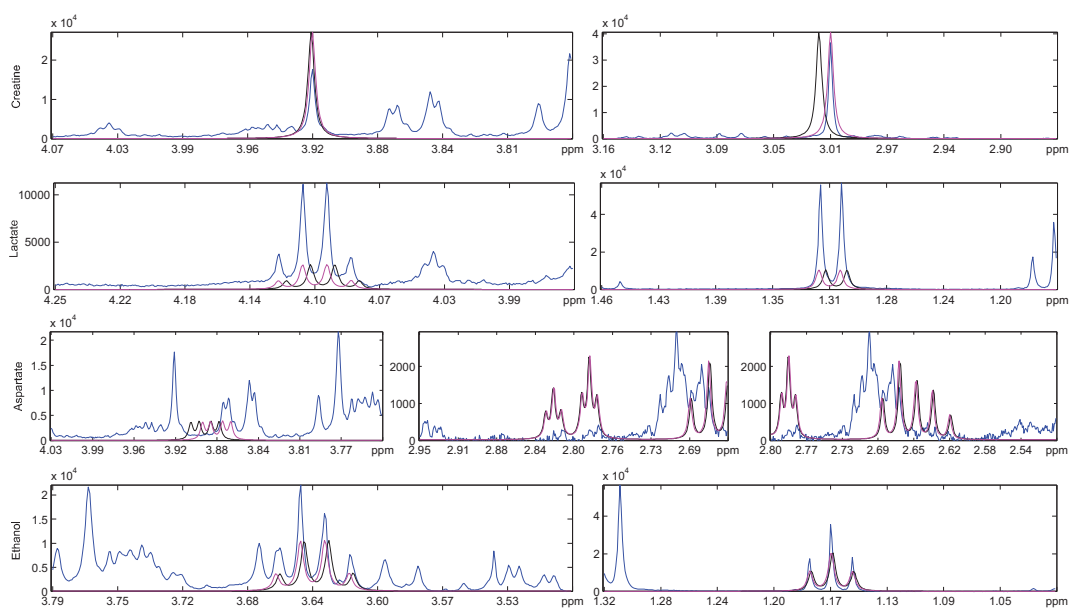


Figure 4.10: Zoom in on the Cr, Lac, Asp and Eth multiplet regions in an HRMAS spectrum from a tissue sample of a human brain with an oligodendroglioma, acquired at 11.7T and the basis-set ones (blue: raw spectrum; black: original basis-set spectra; magenta: optimized basis-set spectra).

The optimized chemical shifts for Lac and Cr for the thirty three spectra are shown in Fig. 4.11. The left plots represent the chemical shifts of two spin groups (or of two chosen groups for the Asp in Fig. 4.12). The red dots correspond to the starting values provided to the optimization procedure. In absence of pH-dependence, a straight line (corresponding to a global spectrum shift) would be expected. The “cloud points” represent independent changes of the chemical shifts of the different multiplets due to the environment changes. On the right, the series of optimized chemical shift differences are plotted as a function of the experiment number. Again, it can be easily seen that independent chemical shift corrections are needed to adapt the distance between the metabolite multiplets. Comparing the results obtained for Creatine and Lactate to those obtained for Aspartate, one can see the poor optimization quality for the latter, see Fig. 4.12. The Aspartate spectra after optimization are not (or slightly) changed. This is due to the small amplitude of Aspartate signal in the investigated spectra. For more complicated spin systems like glutamate and glutamine, the residue of the fit decreases but the commonly low concentrations of these metabolites do not allow us to conclude about the efficiency of the proposed method.

Results show an improvement of quantitation quality when using the optimized basis-set. As quality criterion of the method, we chose the estimated standard deviations on the amplitudes provided by QUEST for metabolites subject to pH changes. Comparing these values obtained using the initial and optimized basis-

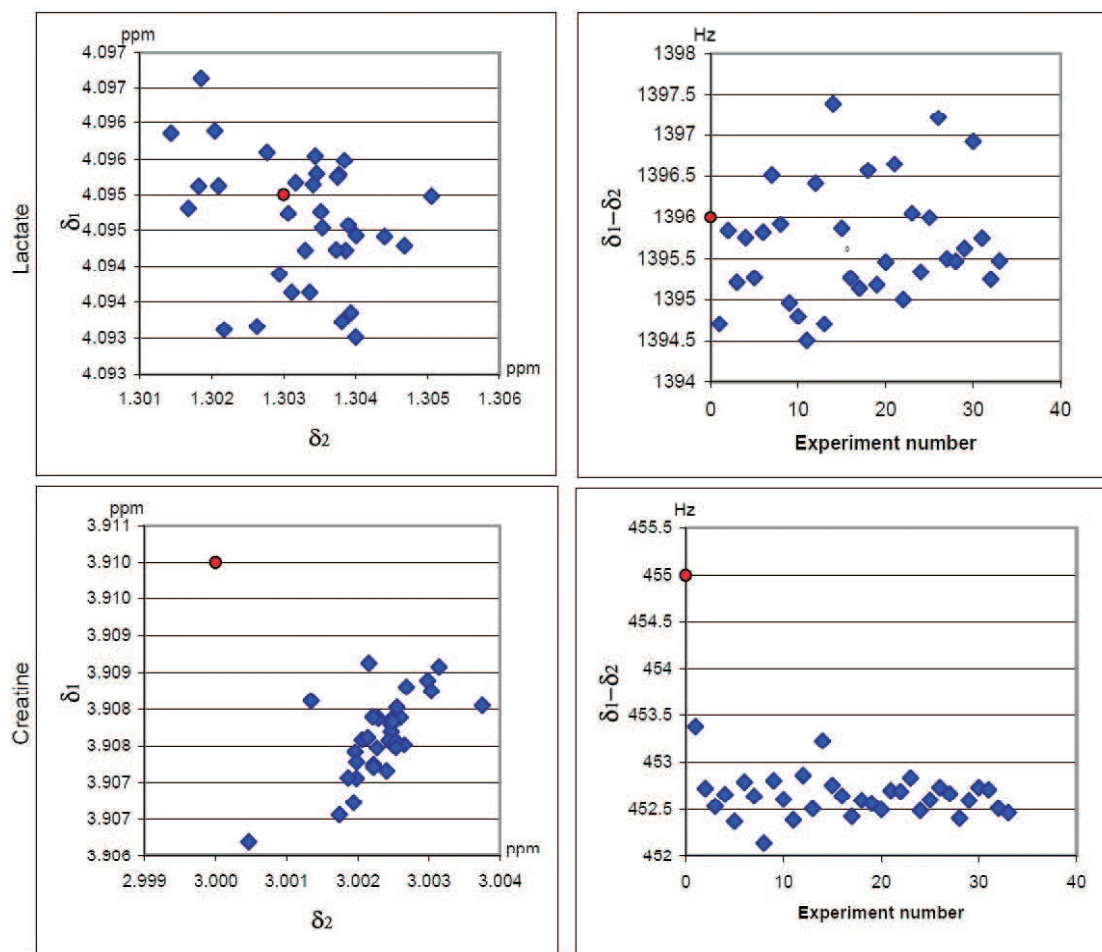


Figure 4.11: Optimized chemical shift values of Lac and Cr for thirty three spectra from tissue samples of human brains with oligodendroglioma acquired at 11.7T. Left: chemical shifts δ_1 and δ_2 of the two proton groups; Right: differences $\delta_1 - \delta_2$. The red dots correspond to the starting values provided to the optimization procedure.

sets respectively, one finds decreases of as much as 30-40%. Moreover, parameters $\Delta\omega_m$ of the model function (Eq. 1.4.1) become redundant after optimization of the chemical shifts, thus reducing the number of free parameters in the quantitation procedure.

To verify the nature of the corrected mismatches, an additional study was carried out. As the changes in chemical shifts are supposed to appear in part due to the pH changes of the environment, we used the Ethanol spectra as a reference because its CH_3 and CH_2 chemical shift values are weakly dependent on the environment acidity. A series of 15mmol/l aqueous Ethanol solutions were investigated. The pH factor of the solution was changed by adding citric acid and was measured with a HANNA Instrument pH-meter HI9025 (one point calibration). Obtained samples were analysed on a Bruker DRX400 NMR spectrometer. The D_2O peak was taken as the spectrum shift reference. Water was not suppressed (chosen Ethanol concentration is enough to clearly recognize all the CH_3 and CH_2

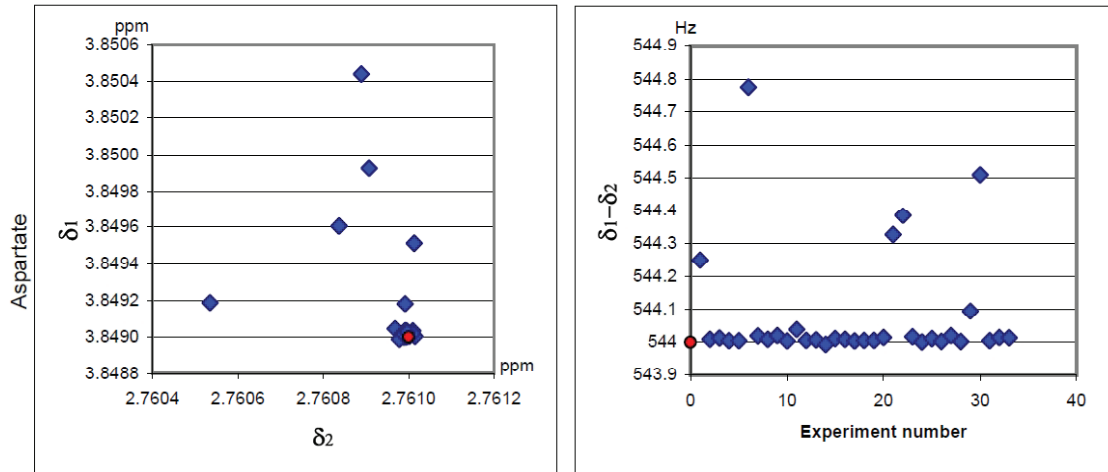


Figure 4.12: Optimized chemical shift values of Asp for thirty three spectra from tissue samples of human brains with oligodendroglioma acquired at 11.7T. Left: chemical shifts δ_1 and δ_2 of the two proton groups at about 2.8 and 3.9 ppm; Right: differences $\delta_1 - \delta_2$. The red dots correspond to the starting values provided to the optimization procedure.

peaks).

Obtained spectra were analyzed in order to find the differences between the chemical shifts of the two proton groups δ_1 and δ_2 . Spectra were artificially dephased to make one of the quadruplet strongest peaks slightly higher than the others. This had no significant influence on the triplet shape, so the middle triplet peak position was taken into account as the chemical shift of CH_3 group. Automatically the distances $\delta_1 - \delta_2$ were found. They are presented in Tab. 4.4. Having

| pH | $\delta_1 - \delta_2$, ppm |
|-------|-----------------------------|
| 3.15 | 2.4629 |
| 3.78 | 2.4629 |
| 4.6 | 2.4626 |
| 5.19 | 2.4634 |
| 5.5 | 2.4630 |
| 5.62 | 2.4624 |
| 5.75 | 2.4631 |
| 6.34 | 2.4638 |
| 10.9 | 2.4632 |
| 11.32 | 2.4625 |

Table 4.4: Differences $\delta_1 - \delta_2$ between the chemical shifts of the two proton groups of the Ethanol spectrum as a function of the solution pH.

found that the standard deviation on dependence of $\Delta\delta$ with pH is equal $4 \cdot 10^{-4}$ ppm, we may conclude that the Ethanol peak positions are very weakly dependent on the environment acidity. Moreover, this can be confirmed by the plot $\delta_1 - \delta_2$ as

a function of the pH of the solution (see Fig. 4.13). The absence of any monotony is another reason to consider the deviations as statistical.

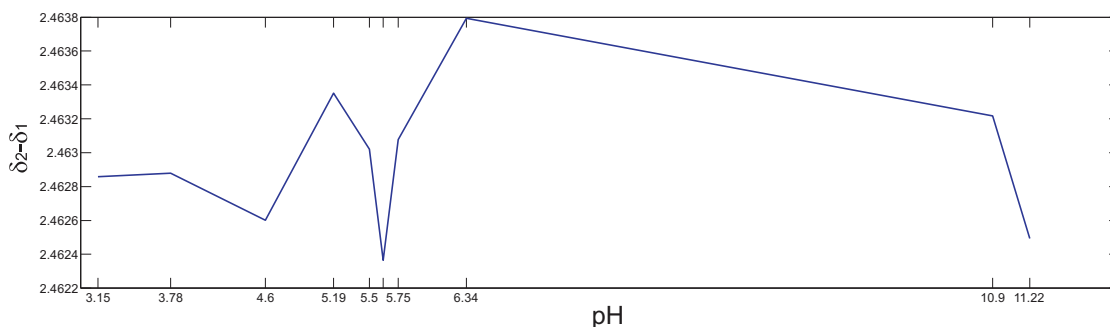


Figure 4.13: Differences between the chemical shifts of the two proton groups in the Ethanol spectrum as a function of pH.

Based on this knowledge, we analyzed the reasons of peak mismatches in the spectra of oligodendroglioma biopsies. To establish possible correlation between the obtained variations in chemical shifts across different metabolites among the thirty-three investigated biopsies, the differences in chemical shifts $\delta_1 - \delta_2$ of the two proton groups for Lac were sorted in ascending order, see Fig. 4.14, left. On the right are shown the differences in chemical shifts $\delta_1 - \delta_2$ of the two proton groups for Cr for the corresponding sorted experiments. A certain dependence exists. Pearson's r between the two series is 0.29 corresponding to a small correlation. But note that the trends of pH curves for different chemical shifts and various metabolites are different [56, 86]. The r between Eth and Lac, Eth and Cr series were also computed and found equal to 0.52 and 0.47 respectively showing a rather large correlation. That means that differences in chemical shift changes is not only due to pH. This can also be confirmed by the analysis of the chemical shift correction data distribution ($\delta_1 - \delta_2$). We used the chemical shifts corrections and looked at the distances between the peak. The regions of interest (1394.5, 1397.5Hz for Lactate, 452.2, 453.4Hz for Creatine) were divided into sub-intervals and numbers of occurrences of $\delta_1 - \delta_2$ in each sub-interval were counted. Plot of the histograms (Fig. 4.15) of the optimized chemical shift differences $\delta_1 - \delta_2$ set suggests the statistical nature of chemical shift mismatches. Distributions are rather close to Gaussian distributions.

4.5.3 Quantitation of cell spectra

The optimization of a metabolite basis set for quantitation of signals of cells appeared much more difficult than that of the tumour biopsies. The “difficulty” is due to the sensitivity of prior knowledge: firstly, to the global spectrum displacement (the position of the water peak, which the spectrometers are often tuned to, is pH dependent, and the acidity variations in cellular samples are more important than in the brain biopsy samples); secondly, to the starting values of chemical shifts (due to the same reason). Moreover, the background due to macromolecular

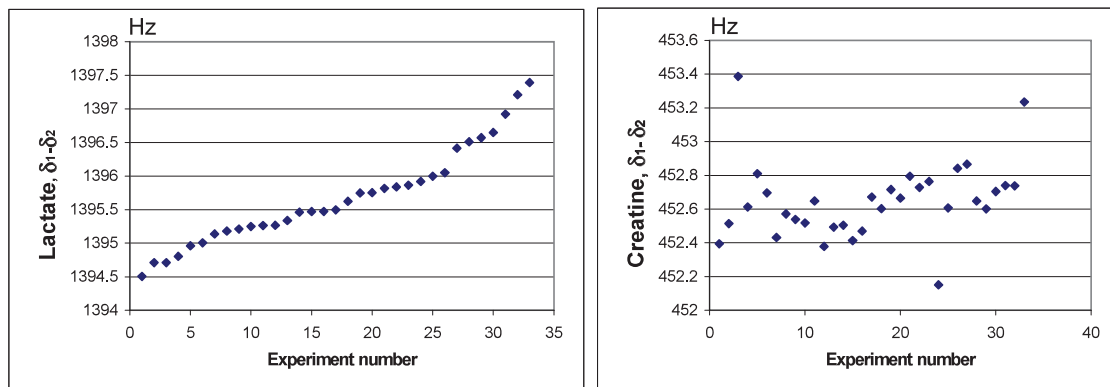


Figure 4.14: Left: optimized chemical shift differences $\delta_1 - \delta_2$ of the two proton groups at about 1.3 and 4.1 ppm of Lac for thirty three spectra from tissue samples of human brains with oligodendroglioma, acquired at 11.7T and sorted in ascending order. Right: differences $\delta_1 - \delta_2$ of the two proton groups at about 3.01 and 3.91 ppm for Cr corresponding to sorted experiments.

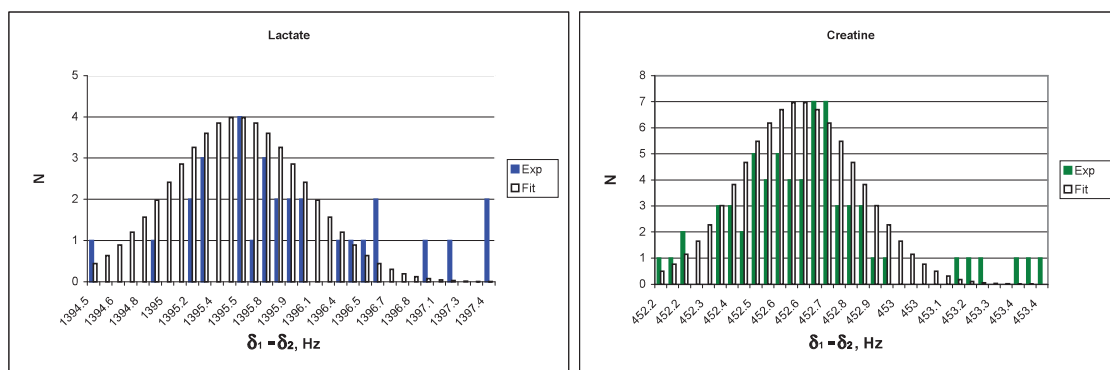


Figure 4.15: Data distributions of the Lactate and Creatine peak position corrections (coloured) and the corresponding Gaussian fits (white).

signals is much more significant, sometimes larger than the important reference peaks as it can be seen in Fig. 4.16.

Therefore, the majority of the tests were done in order to optimize the Creatine and Lactate basis-set signals, as these metabolites are the most recognizable and most dependent (due to presence of the acetic groups) on the acidity of the environment. The spectral regions of interest of these metabolites do not overlap the spectral components of other metabolites, so the initial conditions for the optimization procedure have relatively big margins. Obtained corrections are presented in Tab. 4.5. The maximal number of the loops in the optimization procedure was set to 200, thus the value of 200 in the “Iterations” column means the imposed exit from the loop. Results of optimization of Creatine and Lactate chemical shifts for a spectrum of cells are presented in Fig. 4.17.

As a conclusion to these optimizations, we must note their two main differences from the basis-set optimization for the brain tissue samples. First of them is the fact that the smaller T_2 times lead to larger peak linewidths. This means that

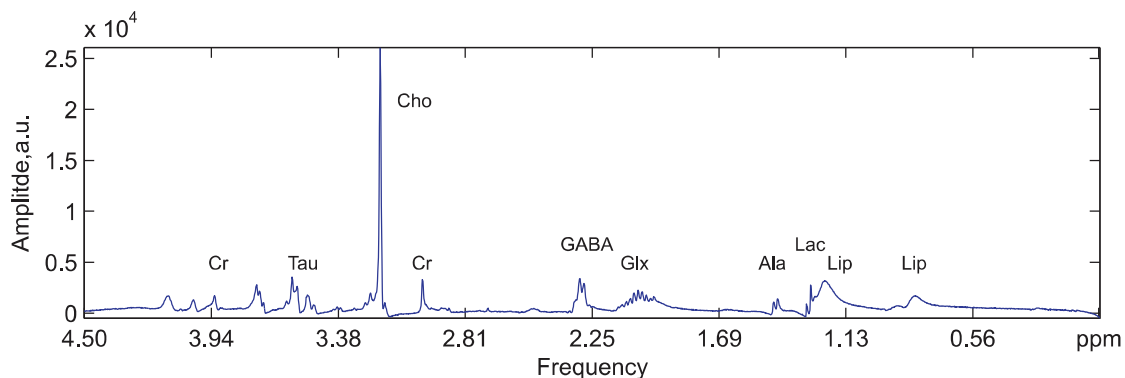


Figure 4.16: An NMR spectrum of a cellular sample, acquired at 400MHz. The only high-resolved peak is the Choline's singlet. Creatine and Alanine groups are also recognizable. The Lactate doublet is almost masked by the lipid peak at 1.3 ppm.

| Experiment | Metabolite | δ_1 | δ_2 | Iterations |
|------------|------------|------------|------------|------------|
| 1 | Creatine | 3.920859 | 2.986938 | 200 |
| 1 | Lactate | 4.093038 | 1.302997 | 2 |
| 2 | Creatine | 3.927632 | 3.004597 | 200 |
| 2 | Lactate | 4.093547 | 1.290304 | 163 |
| 3 | Creatine | 3.931533 | 2.998504 | 200 |
| 3 | Lactate | 4.093049 | 1.287755 | 67 |
| 4 | Creatine | 3.929975 | 3.019971 | 2 |
| 4 | Lactate | 4.093254 | 1.288410 | 163 |
| 5 | Creatine | 3.932643 | 2.999818 | 101 |
| 5 | Lactate | 4.093045 | 1.293825 | 22 |
| 6 | Creatine | 3.932555 | 3.009200 | 143 |
| 6 | Lactate | 4.093036 | 1.304134 | 23 |

Table 4.5: Chemical shift corrections for the series of six spectra from cellular samples.

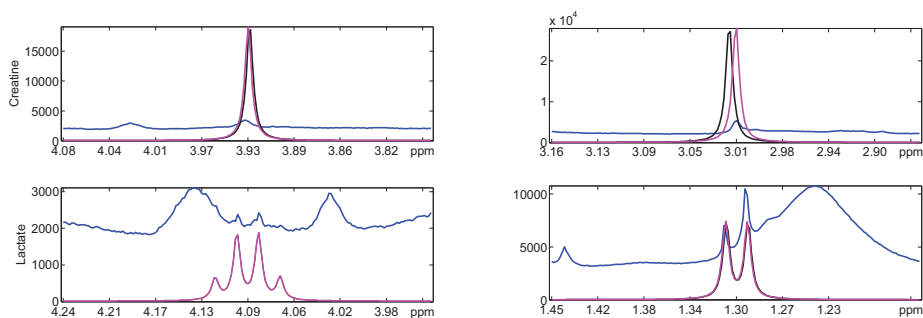


Figure 4.17: Zoom in on the Cr and Lac multiplet regions in an HR MAS spectrum from a cellular sample acquired at 400MHz and the basis-set spectra (blue: raw spectrum; black: original basis-set spectra; magenta: optimized basis-set spectra).

small displacements of the peaks due to pH variations might not be as significant (relative to the linewidth) as they were in brain biopsy spectrum quantitation. Secondly, the variations of the acidity are bigger than previously, so the effect is more significant. Whether to apply or not the optimization method – must be the user decision before the quantitation of each series.

4.6 Conclusion

We proposed a new Quantum Mechanics-based method accounting for pH/temperature-related chemical shift changes prior to quantitation. Despite the fact that the method is time consuming for large spin systems, it offers significant advantages: 1) it is the only method which respects the correct fingerprints of metabolites; 2) parameters $\Delta\omega_m$ of the model (Eq. 1.4.1) become redundant after optimization of the chemical shifts, thus reducing the number of free parameters in the quantitation procedure. The proposed method is well suited to improve quantitation of metabolites with well resolved spectra (lactate, creatine, aspartate, inositol, ethanol as a trace of the biopsy procedure, etc.). For more complicated spin systems like glutamate and glutamine, the method works well if these metabolites have sufficient concentrations. The proposed method will be implemented in the version 4 of the jMRUI software package.

Optimization of metabolite basis sets prior to quantitation in magnetic resonance spectroscopy: an approach based on quantum mechanics

A Lazariev¹, A-R Allouche², M Aubert-Frécon², F Fauvelle³, M Piotto^{4,5}, K Elbayed⁵, I-J Namer⁶, D van Ormondt⁷ and D Graveron-Demilly¹

¹ Laboratoire CREATIS-LRMN, CNRS UMR 5220, Inserm U1044, Université Claude Bernard Lyon 1, Villeurbanne, France

² LASIM, Université Claude Bernard Lyon 1, Villeurbanne, France

³ CRSSA/BCM, Grenoble, France

⁴ Bruker BioSpin, Wissembourg, France

⁵ Institut de Chimie, Strasbourg, France

⁶ Department of Biophysics and Nuclear Medicine, University Hospitals of Strasbourg, Strasbourg, France

⁷ Applied Physics, Delft University of Technology, Delft, The Netherlands

E-mail: danielle.graveron@univ-lyon1.fr

Received 16 December 2010, in final form 13 April 2011

Published DD MMM 2011

Online at stacks.iop.org/MST/22/000000

Abstract

High-resolution magic angle spinning (HRMAS) nuclear magnetic resonance (NMR) is playing an increasingly important role for diagnosis. This technique enables setting up metabolite profiles of *ex vivo* pathological and healthy tissue. The need to monitor diseases and pharmaceutical follow-up requires an automatic quantitation of HRMAS ¹H signals. However, for several metabolites, the values of chemical shifts of proton groups may slightly differ according to the micro-environment in the tissue or cells, in particular to its pH. This hampers the accurate estimation of the metabolite concentrations mainly when using quantitation algorithms based on a metabolite basis set: the metabolite fingerprints are not correct anymore. In this work, we propose an accurate method coupling quantum mechanical simulations and quantitation algorithms to handle basis-set changes. The proposed algorithm automatically corrects mismatches between the signals of the simulated basis set and the signal under analysis by maximizing the normalized cross-correlation between the mentioned signals. Optimized chemical shift values of the metabolites are obtained. This method, QM-QUEST, provides more robust fitting while limiting user involvement and respects the correct fingerprints of metabolites. Its efficiency is demonstrated by accurately quantitating 33 signals from tissue samples of human brains with oligodendroglioma, obtained at 11.7 tesla. The corresponding chemical shift changes of several metabolites within the series are also analyzed.

Keywords: magnetic resonance spectroscopy, HRMAS-NMR, quantitation, quantum mechanics, jMRUI software package

1. Introduction

High-resolution magic angle spinning (HRMAS) nuclear magnetic resonance (NMR) is playing an increasingly

important role for diagnosis. This technique enables setting up metabolite profiles of *ex vivo* pathological and healthy tissue, i.e. biopsies [1–10]. The need to monitor diseases and

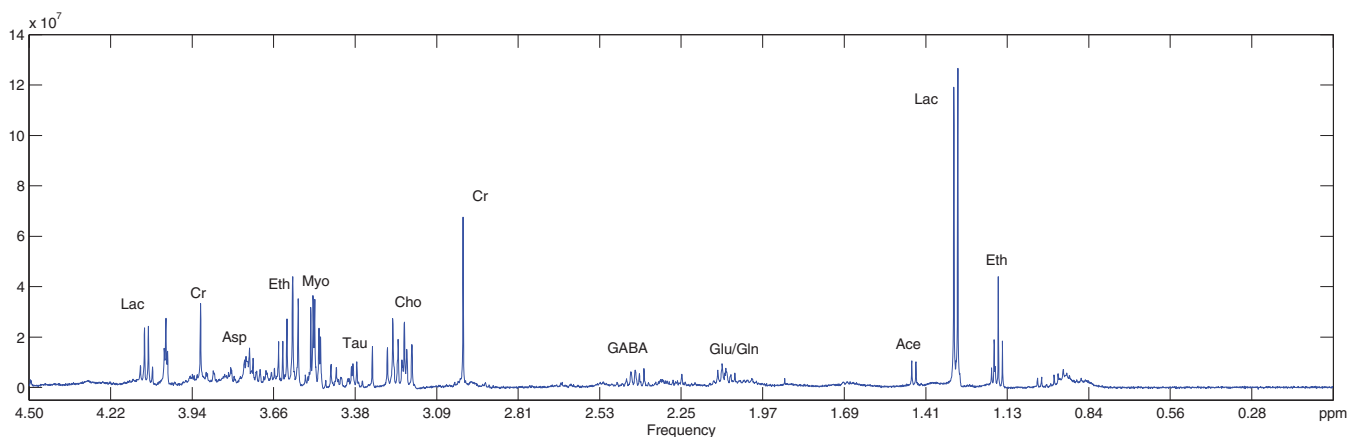


Figure 1. Region of interest of an HRMAS spectrum from a tissue sample of a human brain with an oligodendroglioma, acquired at 11.7 tesla.

pharmaceutical follow-up requires an automatic and reliable quantitation of HRMAS ^1H signals [11–14].

Proton HRMAS signals of *ex vivo* tissues contain several overlapping spectral components (many hundreds) from many metabolites, see figure 1. For many metabolites, the values of chemical shifts of proton groups may slightly differ according to the micro-environment in the tissue or cells, see e.g. [15–17]. In fact, it is widely known that the exact values of chemical shifts are not only subject to the molecular geometry but also depend on the molecular environment. Thus, for example, the temperature dependence of certain peak positions enables temperature mapping [18–20], the pH dependence gives the possibility of analyzing acidity maps [15]. The influence of temperature on chemical shifts can also be predicted by quantum chemistry calculations [21]. These changes of chemical shift values hamper the accurate estimation of the metabolite concentrations mainly when using quantitation algorithms based on a metabolite basis set [22–24]: the metabolite fingerprints are not correct anymore.

The values of mismatches (around some Hz) between the signal to be fitted and those of the metabolite basis set are often not significant for *in vivo* (low-resolution) spectroscopy. But for HRMAS spectroscopy, this increases the error of quantitation to tens of percents for certain metabolites (e.g. Creatine, Lactate). Several methods were proposed to circumvent this problem, see for instance [25] and references herein. Subdividing the basis-set signals of given metabolites into basis sub-components according to chemical groups and adding appropriate constraints (prior knowledge) to the parameters of the groups were proposed in [13]. To limit user involvement, a simple method for chemical shift correction based on signal processing and stretching/shrinking of the metabolite basis-set signals was previously reported in [26, 27]. In *in vivo* MR spectroscopy, chemical shift recalibration based on linear interpolation was recently proposed [28]. In this work, we propose a more accurate method QM-QUEST, based on coupling quantum mechanical simulations and quantitation algorithms to handle basis-set changes, thus respecting the correct fingerprints of metabolites. The method is applied to the quantitation

of 33 signals from tissue samples of human brains with oligodendroglioma, obtained at 11.7 tesla.

This paper is set up as follows. In section 2, we first describe the experiments and the quantitation method QUEST and then detail the quantum mechanics methods used to optimize the chemical shifts of metabolites. In section 3, our method is applied to automatically quantitate series of ^1H HRMAS signals from biopsies of human brain with oligodendroglioma. The obtained chemical shift changes of several metabolites within the series are analyzed. In section 4, results and the method are discussed.

2. Method

2.1. Experiments

1D ^1H HRMAS NMR spectra were obtained on a Bruker Avance 500 spectrometer operating at a proton frequency of 500.13 MHz using a 4 mm double resonance (^1H , ^{13}C) gradient HRMAS probe using a CPMG sequence as described previously [29]. The amount of tumor tissue used ranged from 16 to 20 mg. The whole HRMAS study was performed at 4 °C and was started immediately after the temperature inside the probe reached the equilibrium condition (5 min). Thirty-three spectra from tissue samples of human brains with oligodendroglioma were acquired (sampling interval: 0.143 ms, number of data points $N = 16\,384$).

2.2. Quantitation with QUEST

In magnetic resonance spectroscopy, quantitation based on a metabolite basis set has become very popular in the last 10 years. In this work, we used the method QUEST [23]. The complex-valued time-domain model signal is written as a linear combination of the M weighted metabolite models \hat{x}^m —either quantum mechanically simulated or *in vitro* measured—of the basis set, see figure 2. The model samples, \hat{x}_n , $n = 1, 2, \dots, N$, where N is the number of data points, can be written as

$$\hat{x}_n = \exp(i\phi_0) \sum_{m=1}^M a_m \hat{x}_n^m \exp[(\Delta\alpha_m + i\Delta\omega_m)t_n + i\Delta\phi_m], \quad (1)$$

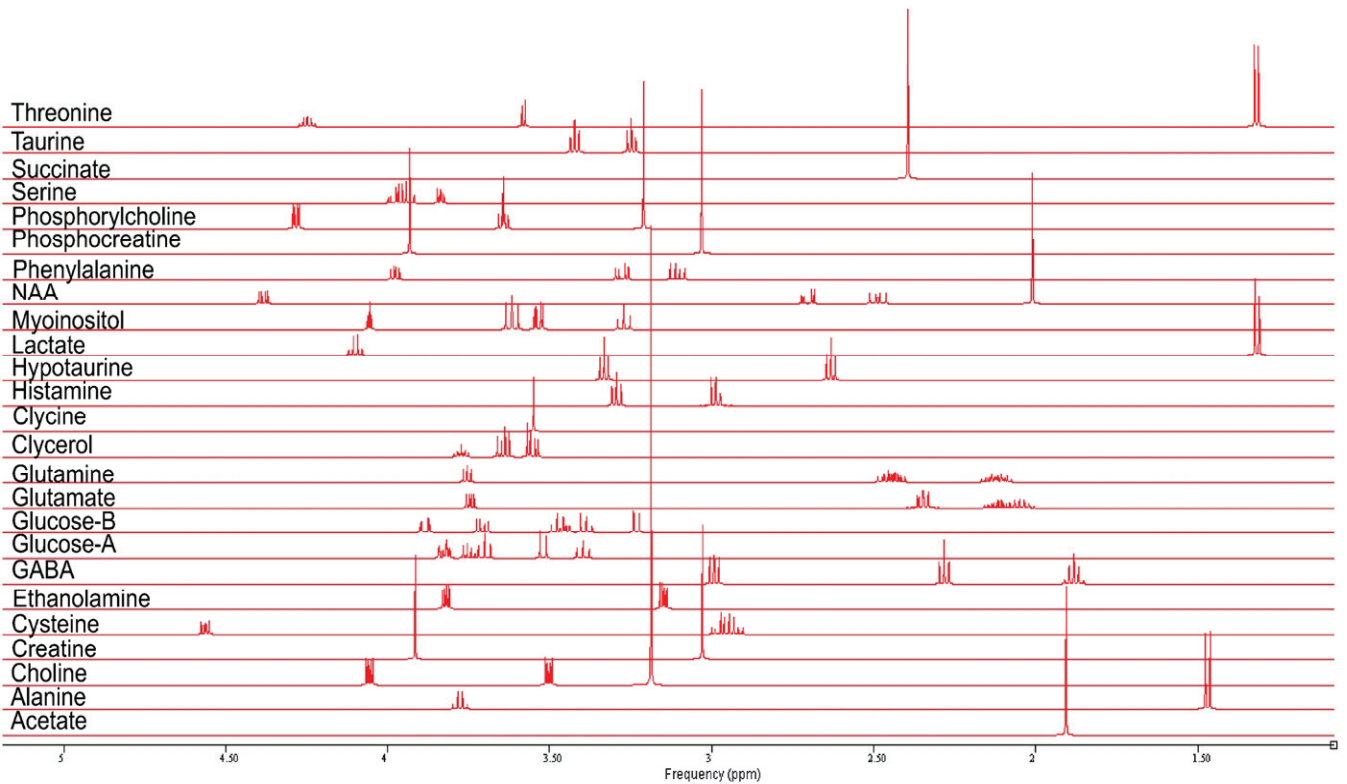


Figure 2. Fourier transform of a metabolite basis set at 11.7 tesla, simulated by quantum mechanics with NMR-SCOPE for a one-pulse sequence. This basis set was used in QUEST for the quantitation of human brain HRMAS signals. Lorentzian lineshapes were used.

where

- \hat{x}^m , m being a superscript, represents the metabolite basis-set signals.
- a_m are M amplitudes to be estimated. Note that these amplitudes represent the relative proportions of the M metabolite signals \hat{x}^m in the signal x rather than the amplitudes of individual spectral components.
- $\Delta\alpha_m$, $\Delta\omega_m$ and $\Delta\phi_m$ represent small changes of the damping factors, angular frequencies, and phase shifts, respectively. These changes—relative to the initial values in the metabolite basis set—are included in the estimation procedure to automatically compensate for the effect of magnetic field inhomogeneities. Most often $\Delta\phi_m = 0$. Soft constraints on $\Delta\alpha_m$ and $\Delta\omega_m$ have been used in the minimization procedure.
- $t_n = nt_s + t_0$, $n = 1, 2, \dots, N$, are the sampling times, in which t_0 is the dead-time of the receiver—fixed to 0 in this study—and t_s is the sampling interval.
- ϕ_0 is an overall phase, fixed to 0 in this study.
- $i^2 = -1$.

2.3. Optimization of the basis-set signals

Two quantum-mechanical simulators were used to simulate the basis sets. The first one is the NMR-SCOPE algorithm based on density matrix formalism [30], the second one, used in the optimization procedure, is based on the conventional approach [31].

2.3.1. Density matrix approach: NMR-SCOPE. NMR-SCOPE is based on the density matrix and the product operator (super operators) formalism, and commutator algebra. It handles a product-operator description of strongly coupled spin $\frac{1}{2}$ systems and enables to simulate signals of the metabolites in response to magnetic resonance pulse-sequences. It can handle various pulse sequences such as STEAM and PRESS and provides directly the time-domain signals. It is applicable to ^1H , ^{13}C , ^{19}F , ^{31}P , \dots , nuclei, and arbitrary field strength. The equation of motion of the density matrix operator $\rho(t)$ under the influence of a time-independent Hamiltonian \mathcal{H} is given by

$$\rho(t) = \exp(-i\mathcal{H}t)\rho(0)\exp(i\mathcal{H}t) \quad (2)$$

where $\rho(0)$ is the density operator at the beginning of the experiment. Then, the evolution of the density matrix product operators was computed according to equation (2) for each pulse-sequence event by adapting \mathcal{H} and t . During free-precession time intervals, \mathcal{H} was taken equal to the strong coupling spin Hamiltonian

$$\mathcal{H} = \sum_i^{N_{\text{spins}}} \delta_i I_{iz} + \sum_{i<j}^{N_{\text{spins}}} \sum_j^{N_{\text{spins}}} 2\pi J_{ij} \mathbf{I}_i \cdot \mathbf{I}_j, \quad (3)$$

where N_{spins} is the number of spins of the molecule/metabolite under consideration, I_{ix} , I_{iy} , I_{iz} represent the components of the spin \mathbf{I}_i , δ_i represent the chemical shifts and J_{ij} represent the coupling constants of the metabolite spins.

For radio-frequency (RF) pulses, \mathcal{H} is equal to

$$\mathcal{H}_{\text{RF}} = \theta \sum_i^{N_{\text{spins}}} (I_{ix} \cos \phi + I_{iy} \sin \phi), \quad (4)$$

where θ is the magnetization flip angle and ϕ is the phase of the RF pulses. The Hamiltonian for selective RF pulses also includes the free-precession Hamiltonian.

In the detection period, the sampled signal is directly simulated in the time-domain leading to \hat{x}^m .

2.3.2. Conventional approach. In the optimization procedure, the conventional quantum mechanical approach [31] is used to speed up the calculations as at this step, one needs signals with accurate frequencies. The method is based on the time-independent Schrödinger equation

$$\mathcal{H}\Psi = E\Psi \quad (5)$$

where Ψ are the wavefunctions and E are the energy levels of \mathcal{H} . The matrix representation of \mathcal{H} and eigenvectors in the spin system were obtained by working in the appropriate direct product space. Thus, for two spins $\frac{1}{2}$, $|\alpha\alpha\rangle$, $|\alpha\beta\rangle$, $|\beta\alpha\rangle$, $|\beta\beta\rangle$ were used as basis set, where α and β represent the $\frac{1}{2}$ and $-\frac{1}{2}$ states, respectively. The frequencies and corresponding amplitudes of the spectral components are obtained after diagonalization of the matrix representation of \mathcal{H} corresponding to equation (3) on the mentioned basis set thus enabling to compute \hat{x}^m .

2.3.3. Metabolite basis sets. The basis-set signals were simulated with NMR-SCOPE. Spin parameters were initially taken from [32] and we refined them to fit rodent brain HRMAS spectra [33]. Twenty-five metabolites—acetate (Ace), alanine (Ala), aspartate (Asp), creatine (Cr), choline (Cho), cysteine (Cys), ethanolamine (Eth), γ -amino-butyric acid (GABA), glucose (Glc), glutamate (Glu), glutamine (Gln), glycerol, glycine (Gly), histamine (Hist), hypotaurine (HTau), lactate (Lac), myo-inositol (mI), N-acetylaspartate (NAA), phosphoryl-choline (PC), phosphocreatine (PCr), phenylalanine (Phe), serine (Ser), succinate (Suc), taurine (Tau) and ethanol (Eth) as a trace of biopsy procedure—were included in the basis set. Signals modelling the lipids (Lip) at 0.9 and 1.3 ppm were not included in the basis set, considering that their model function is insufficiently known. They are estimated with the background signal.

As mentioned above, for several metabolites, chemical shifts can slightly differ according to the micro-environment in the tissue or cells, in particular with its pH. That means that prior knowledge based on the metabolite basis-set signals is not entirely correct anymore: the fingerprints of metabolites slightly differ.

2.3.4. Optimization of metabolite chemical shifts. The goal was to correct the possible mismatches between the signals of the simulated basis set and the experimental signal and to optimize the chemical shift values of the metabolites.

As mentioned above, the basis-set signals \hat{x}^m depend directly on the spin parameters (chemical shifts δ_i^m and coupling constants J_{ij}^m) of the metabolite m :

$$\hat{x}^m = f(\delta_i^m, J_{ij}^m), \quad i, j = 1, 2, \dots, N_{\text{spins}}^m. \quad (6)$$

We propose to modify the metabolite basis-set signals \hat{x}^m , sensitive to pH and/or temperature changes, before quantitation by maximizing the normalized cross-correlation between each of these signals and the investigated HRMAS signal x as a function of chemical shift values:

$$\text{corr}(x, \hat{x}^m) = \sum_{i=1}^N x_i^* \hat{x}_i^m \quad (7)$$

where $*$ indicates complex conjugation. This means that the chemical shifts δ_i^m , initially provided to the quantum mechanical simulation procedure for metabolite subject to pH and/or temperature changes, will be optimized and new $\delta_{i,\text{opt}}^m$ estimated. The normalized cross-correlation, chosen as the cost function, avoids signal normalization.

To the best of our knowledge, such work has never been reported. If fitting of spin parameters is a well-known problem in high-resolution NMR, the problem tackled here is much more complicated as the signals under analysis show rich biomedical profile from a mixture of metabolites with unknown concentrations and overlapping spectral components. Minimization of the residue between the fitted and theoretical signals of a metabolite, usually used in high-resolution NMR, could not be used anymore.

2.3.5. Algorithm scheme. The optimization scheme of MQ-QUEST can be summarized as follows:

- (1) simulation with NMR-SCOPE of the basis set $\{\hat{x}^m, m = 1, 2, \dots, M\}$ using the initial spin parameters $\{\delta_i^m$ and $J_{ij}^m, i, j = 1, 2, \dots, N_{\text{spins}}^m\}$;
- (2) selection of metabolites that are subject to pH changes leading to M_{pH} metabolite signals to be optimized;
- (3) for each of the M_{pH} metabolites: maximization of $\text{corr}(x, \hat{x}^m) \implies \{\delta_{i,\text{opt}}^m, i = 1, 2, \dots, N_{\text{spins}}^m\}$ and \hat{x}_{opt}^m ;
- (4) simulation with NMR-SCOPE of the optimized basis set $\{\hat{x}_{\text{opt}}^m, m = 1, 2, \dots, M_{\text{pH}}\}$ using the spin parameters $\{\delta_{i,\text{opt}}^m$ and $J_{ij}^m, i, j = 1, 2, \dots, N_{\text{spins}}^m\}$;
- (5) QUEST quantitation using the optimized basis set $\{\hat{x}_{\text{opt}}^m, m = 1, 2, \dots, M\}$.

Note that the conventional quantum mechanical simulation algorithm is inside the maximization procedure (based on the steepest descent algorithm) of the cost function. The optimization was performed in Matlab (©MathWorks) using the conventional quantum mechanical algorithm in C [34] and quantitation was performed with the jMRUI software package [20].

As a quality criterion of the method, we chose the estimated standard deviations based on the estimated Cramér-Rao lower bounds on the amplitudes provided by QUEST [23] for metabolites subject to pH and/or temperature changes.

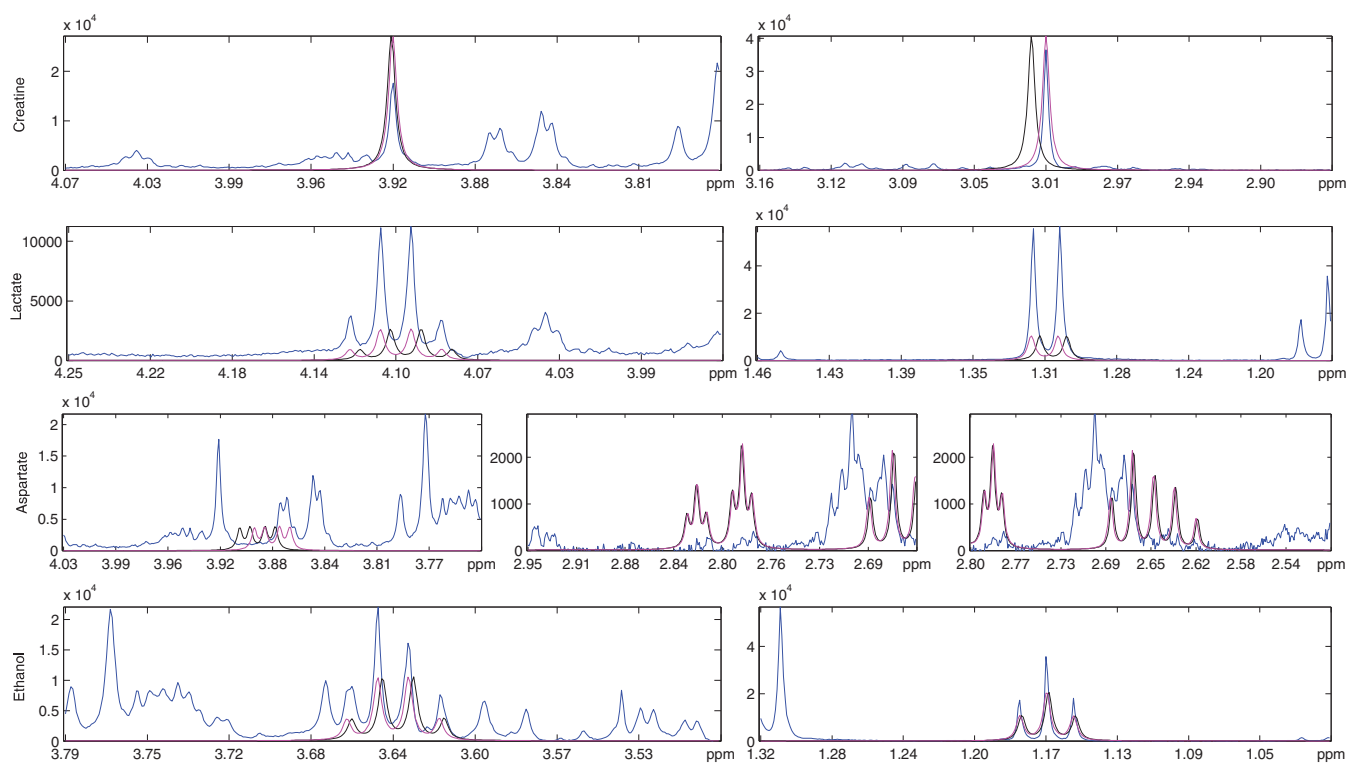


Figure 3. Zoom in on the Cr, Lac, Asp and Eth multiplet regions in an HRMAS spectrum from a tissue sample of a human brain with an oligodendroglioma, acquired at 11.7 tesla and the basis-set ones (blue: raw spectrum, black: original basis-set spectra, magenta: optimized basis-set spectra).

3. Results

The proposed method was first tested on simulated signals mimicking HRMAS-NMR ones (not reported in this paper) and then on HRMAS-NMR signals. Thirty-three spectra from tissue samples of human brains with oligodendroglioma, acquired at 11.7 tesla were quantitated. The region of interest of the spectra was first extracted with ER-Filter [35] as mentioned in [33]. Mismatches between the initial basis-set signals and some *ex vivo* signals were observed. For the latter, adaptation of the basis-set signals was necessary. Results are illustrated for an oligodendroglioma spectrum of the series in figure 3. Lac, Cr, Asp, Eth, the latter as a trace of the biopsy procedure, and other signals of the basis set were automatically optimized prior to the quantitation procedure as mentioned above by maximizing the correlation between the *ex vivo* and basis-set signals. The method enables independent movements of the different multiplets of the spectrum, keeping all strong-coupling effects, contrary to methods which subdivide the basis-set signals of given metabolites into sub-components. This enabled us to adapt the chemical shift variations. It can easily be seen in figure 3 that the mismatches between the basis-set and HRMAS spectra have been reduced for Cr, Lac, Eth. Asp multiplets at 2.80 and 2.65 ppm have amplitudes of the same order as the noise level which makes the recognition of peak positions more difficult.

The optimized chemical shifts for Lac and Cr for the 33 spectra are shown in figure 4. The left plots represent the chemical shifts of two spin groups (or of two chosen

groups for Asp in figure 5). The red dots correspond to the starting values provided to the optimization procedure. In the absence of pH-dependence, a straight line (corresponding to a global spectrum shift) would be expected. The ‘cloud points’ represent independent changes of the chemical shifts of the different multiplets due to the environment changes. Note that the Cr chemical shift at about 3 ppm is nearly the same for all biopsies, in agreement with [36]. On the right, the series of optimized chemical shift differences are plotted as a function of the experiment number. Again, it can be easily seen that independent chemical shift corrections are needed to adapt the distance between the metabolite multiplets. From the estimated chemical shifts of Cr in the series and the titration curves given in [36], one could cautiously estimate that the pH in these biopsies was about 7–8. Comparing the results obtained for Cr and Lac to those obtained for Asp, one can see the poor optimization quality for the latter, see figure 5. The Asp spectra after optimization are not (or slightly) changed. This is due to the small amplitude of the Asp signal in the investigated spectra. For more complicated spin systems such as Glu and Gln, the residue of the fit decreases but the commonly low concentrations of these metabolites do not allow us to conclude about the efficiency of the proposed method.

To establish possible correlation between the obtained variations in chemical shifts across different metabolites among the 33 investigated biopsies, the differences in chemical shifts $\delta_1 - \delta_2$ of the two proton groups for Lac were sorted in ascending order, see figure 6, left. On the right the differences in chemical shifts $\delta_1 - \delta_2$ of the two proton groups for Cr for

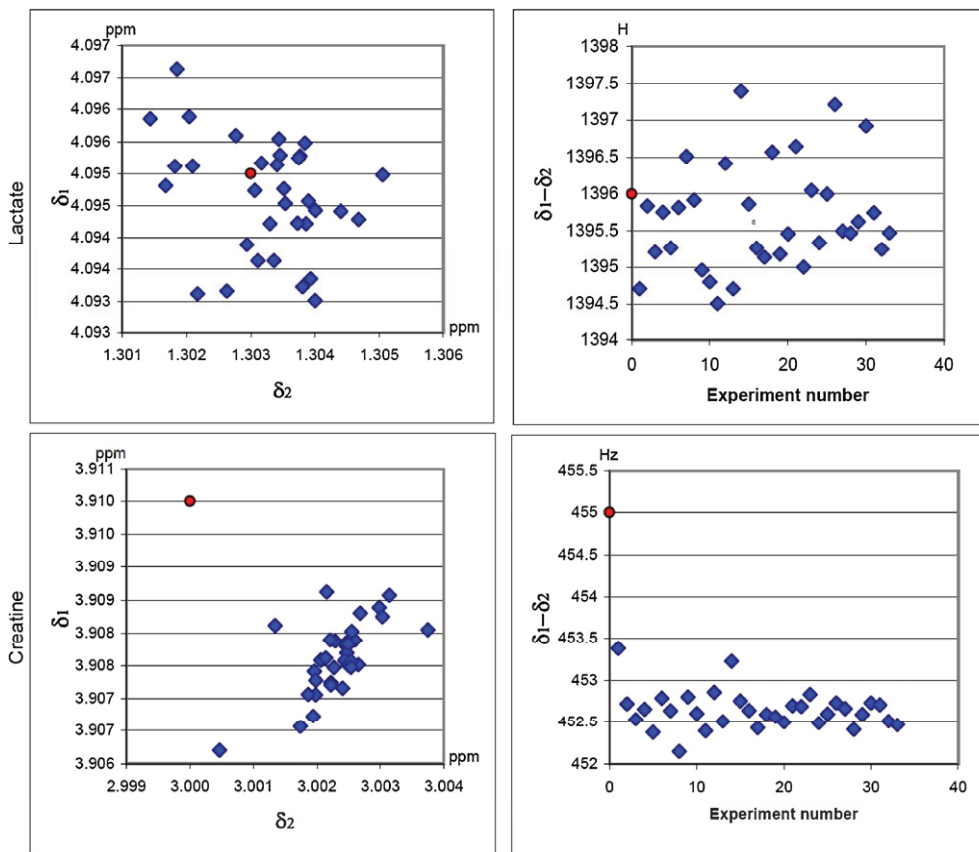


Figure 4. Optimized chemical shift values of Lac and Cr for 33 spectra from tissue samples of human brains with oligodendroglioma, acquired at 11.7 tesla. Left: chemical shifts δ_1 and δ_2 of the two proton groups; right: differences $\delta_1 - \delta_2$. The red dots correspond to the starting values provided to the optimization procedure.

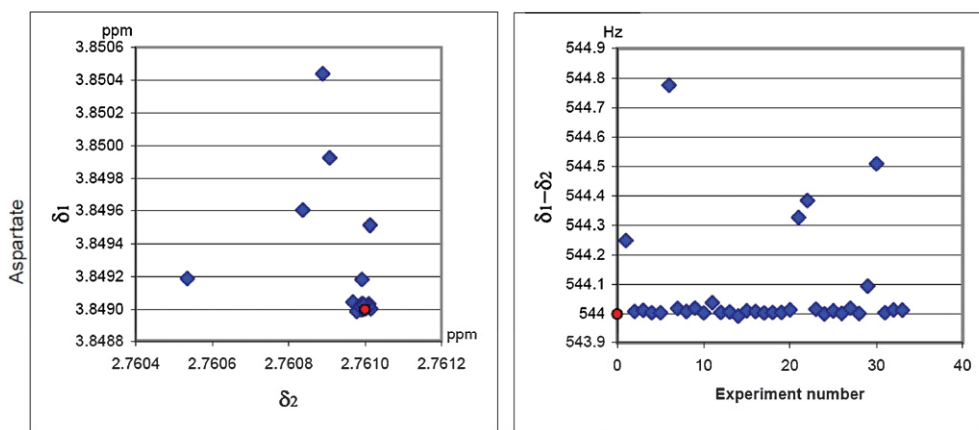


Figure 5. Optimized chemical shift values of Asp for 33 spectra from tissue samples of human brains with oligodendroglioma, acquired at 11.7 tesla. Left: chemical shifts δ_1 and δ_2 of the two proton groups at about 2.8 and 3.9 ppm; right: differences $\delta_1 - \delta_2$. The red dots correspond to the starting values provided to the optimization procedure.

the corresponding sorted experiments are shown. A certain dependence exists. The Pearson product-moment correlation r between the two series is 0.29 corresponding to a small correlation. But note that the trends of pH curves for different chemical shifts and various metabolites are different [16, 36]. Pearson's r between Eth and Lac, and Eth and Cr series were also computed and found equal to 0.52 and 0.47, respectively, showing a rather large correlation. That means that

differences in chemical shift changes are in part due to pH as expected.

Results show an improvement of quantitation quality when using the optimized basis set. As a quality criterion of the method, we chose the estimated standard deviations on the amplitudes provided by QUEST for metabolites subject to pH and/or temperature changes. Comparing these values obtained using the initial and optimized basis sets, one finds

Q1

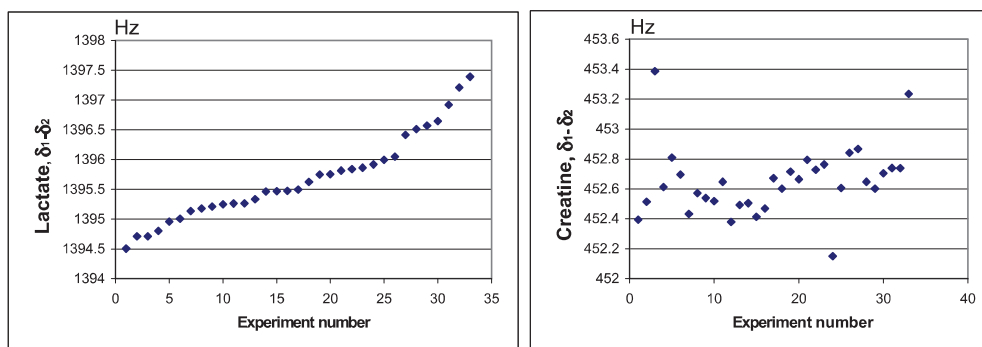


Figure 6. Left: optimized chemical shift differences $\delta_1 - \delta_2$ of the two proton groups at about 1.3 and 4.1 ppm of Lac for 33 spectra from tissue samples of human brains with oligodendroglioma, acquired at 11.7 tesla and sorted in ascending order. Right: differences $\delta_1 - \delta_2$ of the two proton groups at about 3 and 3.91 ppm for Cr corresponding to sorted experiments.

decreases of as much as 30–40%. Moreover, the parameters $\Delta\omega_m$ of the model function (equation (1)) become redundant after the optimization of the chemical shifts, thus reducing the number of free parameters in the quantitation procedure.

The computing time for the optimization procedure depends mainly on the spin number of the considered metabolite as the quantum mechanical simulation is iteratively computed in the cost function. It was about 8 s for Cr, 63 s for Lac, 44 s for Asp and 67 s for Eth, on a laptop PC with an Intel Core Duo with a 2.4 GHz processor and 2 GHz of RAM, running Windows XP.

4. Discussion

A novel method based on quantum mechanical simulations was developed that provides an *automatic* approach for the accurate quantitation of the metabolite concentrations in high-resolution signals from biopsies. As previously demonstrated, the method when combined with a quantitation algorithm (for instance, QUEST, leading to the method QM-QUEST) provides more robust fitting. In particular, it is well suited to improve the quantitation of metabolites with well-resolved spectra (Lac, Cr, Asp, mI, Eth as a trace of the biopsy procedure, etc) in HRMAS spectra. Moreover, it limits user involvement as only metabolites for which chemical shifts have to be optimized must be selected, whereafter optimized chemical shift values for these metabolites are provided automatically. Compared to methods based on subdividing the basis-set signals of given metabolites into basis sub-components according to chemical groups and adding appropriate constraints (prior knowledge) to the parameters of the groups [13], the proposed method limits user interaction and the number of parameters used in the quantitation routine. The simple method for chemical shift correction based on signal processing and stretching/shrinking of the metabolite basis-set signals, we previously reported in [26, 27], does not provide optimized chemical shifts.

The reliability achieved in the final fits was assessed using the estimated Cramér–Rao Bounds and shows decreases of as much as 30–40% for metabolites for which the basis-set signals have been optimized. For a given metabolite, results depend on the signal-to-noise ratio (SNR) related to the spin

numbers and concentration and on the overlap between the spectral components of other metabolites. For example, the scyllo-Inositol singlet appears near other metabolite multiplets around 3.3 ppm, which makes its chemical shift hardly optimizable. For several metabolites, the estimated changes in the chemical shifts were unique; this result differed from one biopsy to the other. Only the Cr chemical shift at about 3 ppm was nearly the same for all biopsies. No strong correlation has been observed between changes in the $\delta_1 - \delta_2$ of pairs of metabolites, indicating that these changes are due not only to pH.

Besides important improvements, the method has some limitations. For metabolites with very similar spectra such as Cr and PCr, it was not possible to adjust their chemical shifts separately and only total Cr was considered. Note that in most of the spectra of the series, PCr was not identifiable. For more complicated and overlapping spin systems such as Glu and Gln, α and β Glc, the method works well if these metabolites have sufficient concentrations. When the SNR is low, the quantum mechanical optimization procedure should be included into the quantitation algorithm. This would enable to quantitate signals with an optimized basis set but the computing time would substantially increase. As mismatches of peak positions between the basis-set signals and the signal under analysis are of the order of some Hz, the optimization algorithm is also sensitive to the presence of local minima of the correlation function.

The method would also be well suited for the accurate quantitation of prostate tissues, mainly because of citrate chemical shift changes according to pH [16].

5. Conclusion

We proposed a new quantum mechanics-based method accounting for pH- and/or temperature-related chemical shift changes prior to quantitation. Its advantages are as follows: (1) it is the only method which respects the correct fingerprints of metabolites; (2) it limits user involvement, (3) the parameters $\Delta\omega_m$ of the model (equation (1)) in QUEST become redundant after the optimization of the chemical shifts, thus reducing the number of free parameters in the quantitation procedure. The proposed method QM-QUEST is well suited

to improve the quantitation of metabolites with well-resolved spectra (lactate, creatine, aspartate, myo-inositol, ethanol as a trace of the biopsy procedure, etc). For more complicated spin systems such as glutamate and glutamine, the method works well if these metabolites have sufficient concentrations. The proposed method QM-QUEST will be implemented in the coming version 4.2 of the jMRUI software package.

Acknowledgments

The oligodendrogliomas spectra were acquired in the context of the CARMEN project. This work was supported by the EU Marie Curie Research Network, FAST, MRTNCT-2006-035801 and Philips Medical Systems, Best, NL.

Q2 References

- [1] Swanson M G *et al* 2006 Quantitative analysis of prostate metabolites using ^1H HR-mAS spectroscopy *J. Magn. Reson.* **55** 1257–64
- [2] Martínez-Granados B, Monleón D, Martínez-Bisbal M C, Rodrigo J M, Olmo J del, Lluch P, Ferrández A, Martí-Bonmatí L and Celda B 2006 Metabolite identification in human liver needle biopsies by high-resolution magic angle spinning ^1H NMR spectroscopy *NMR Biomed.* **19** 90–100
- [3] Tsang T M, Griffin J L, Haselden J, Fish C and Holmes E 2005 Metabolic characterization of distinct neuroanatomical regions in rats by magic angle spinning ^1H nuclear magnetic resonance spectroscopy *Magn. Reson. Med.* **53** 1018–24
- [4] Burns M A, Taylor J L, Wu C L, Zepeda A G, Bielecki A, Cory D and Cheng L L 2005 Reduction of spinning sidebands in proton NMR of human prostate tissue with slow high-resolution magic angle spinning *Magn. Reson. Med.* **54** 34–42
- [5] Sitter B, Lundgren S, Bathen T F, Halgunset J, Fjosne H E and Gribbestad I S 2005 Comparison of HR MAS MR spectroscopic profiles of breast cancer tissue with clinical parameters *NMR Biomed.* **19** 30–40
- [6] Martínez-Bisbal M C, Martí-Bonmatí L, Piquer J, Revert A, Ferrer P, Llácer J L, Piotta M, Assemat O and Celda B 2004 ^1H and ^{13}C HR-mAS spectroscopy of intact biopsy samples *ex vivo* and *in vivo* ^1H MRS study of human high grade gliomas *NMR Biomed.* **17** 191–205
- [7] Lindon J C, Beckonert O P, Holmes E and Nicholson J K 2009 High-resolution magic angle spinning NMR spectroscopy: application to biomedical studies *Prog. Nucl. Magn. Reson. Spectrosc.* **55** 79–100
- [8] Fauvelle F, Dorandeu F, Carpentier P, Foquin A, Rabeson H, Graveron-Demilly D, Arvers P and Testylier G 2010 Changes in mouse brain metabolism following a convulsive dose of soman: a proton HRMAS NMR study *Toxicology* **267** 99–111
- [9] Wright A J, Fellows G A, Griffiths J R, Wilson M, Bell B Anthony and Howe F A 2010 *Ex vivo* HRMAS of adult brain tumours: metabolite quantification and assignment of tumour biomarkers *Mol. Cancer* **9** 66
- [10] Sava A Croitor, Martínez-Bisbal M C, Van Huffel S, Cerda J M, Sima D M and Celda B 2011 *Ex vivo* high resolution magic angle spinning metabolic profiles describe intratumoral histopathological tissue properties in adult human gliomas *Magn. Reson. Med.* **65** 320–8
- [11] Rabeson H, Fauvelle F, Ratiney H, Cudalbu C, van Ormondt D and Graveron-Demilly D 2006 Quantitation with QUEST of ER-filtered HRMAS-NMR signals *ProRISC, IEEE Benelux* (Veldhoven, The Netherlands, November) pp 171–5
- [12] Rabeson H, Ratiney H, Fauvelle F, Cudalbu C, Cavassila S, Bernard A, Giraudon P, van Ormondt D and Graveron-Demilly D 2007 Quantitation for *in vivo* and *ex vivo* NMR spectroscopy *J. Optoelectron. Adv. Mater.* **9** 505–11
- [13] Reynolds G, Wilson M, Peet A and Arvanitis T N 2006 An algorithm for the automated quantitation of metabolites in *in vitro* NMR signals *Magn. Reson. Med.* **56** 1211–9
- [14] DiCamillo F P, Swanson M G, Kurhanewicz J and Nelson S J 2006 Assessment of HR-MAS NMR prostate biopsy tissue spectra with principal component analysis and metabolite quantification *Proc. Int. Soc. Magnetic Resonance in Medicine, 14th Scientific Meeting and Exhibition (Seattle, USA)* pp 2534
- [15] Hall L D and Talagala S L 1985 Mapping of pH and temperature distribution using chemical-shift-resolved tomography *J. Magn. Reson.* **65** 501–5
- [16] van der Graaf M and Heerschap A 1996 Effect of cation binding on the proton chemical shifts and the spin–spin coupling constant of citrate *J. Magn. Reson. Ser. B* **112** 58–62
- [17] Robert O, Sabatier J, Desoubzdanne D, Lalande J, Balaýssac S, Gilard V, Martino R and Malet-Martino M 2011 PH optimization for a reliable quantification of brain tumor cell and tissue extracts with ^1H NMR: focus on choline-containing compounds and taurine *Anal. Bioanal. Chem.* **399** 987–999
- [18] Ishihara Y, Calderon A, Watanabe H, Okamoto K, Suzuki Y, Kuroda K and Suzuki Y 1995 A precise and fast temperature mapping using water proton chemical shift *Magn. Reson. Med.* **34** 814–23
- [19] Baxter N J and Williamson M P 1997 Temperature dependence of ^1H chemical shifts in proteins *J. Biomol. NMR* **9** 359–69
- [20] Stefan D *et al* 2009 Quantitation of magnetic resonance spectroscopy signals: the jMRUI software package *Meas. Sci. Technol.* **20** 9 pp
- [21] Atieh Z, Allouche A R, Lazariev A, van Ormondt D, Graveron-Demilly D and Aubert-Frécon M 2010 DFT calculations of ^1H NMR chemical shifts, simulated and experimental spectra of sarcosine *Chem. Phys. Lett.* **492** 297–301
- [22] Provencher S W 1993 Estimation of metabolite concentrations from localized *in vivo* proton NMR spectra *Magn. Reson. Med.* **30** 672–9
- [23] Ratiney H, Sdika M, Coenradie Y, Cavassila S, van Ormondt D and Graveron-Demilly D 2005 Time-domain semi-parametric estimation based on a metabolite basis set *NMR Biomed.* **18** 1–13
- [24] Ratiney H, Coenradie Y, Cavassila S, van Ormondt D and Graveron-Demilly D 2004 Time-domain quantitation of ^1H short echo-time signals: background accommodation *Magn. Reson. Mater. Phys.* **16** 284–96
- [25] Giskeødegård G F, Bloemberg T G, Postma G, Sitter B, Tessem M-B, Gribbestad I S, Bathen T F and Buydens L M C 2010 Alignment of high resolution magic angle spinning magnetic resonance spectra using warping methods *Anal. Chim. Acta* **683** 1611
- [26] Lazariev A, Piotta M, Elbayed K, Namer I J, van Ormondt D and Graveron-Demilly D 2010 Simple correction of chemical shift changes in magnetic resonance spectroscopy quantitation *Proc. IEEE IST2010 (Thessaloniki, Greece, 1–2 July)* pp 365–8
- [27] Lazariev A, Fauvelle F, Piotta M, Elbayed K, Namer I J, van Ormondt D and Graveron-Demilly D 2010 Simple correction of chemical shift changes in magnetic resonance spectroscopy quantitation *ISMRM-ESMRMB 2010 (Stockholm, SE, 1–7 May)* pp 3318
- [28] Alusta P, Im I, Pearce B A, Beger R D, Kretzer R M, Buzatu D A and Wilkes J G 2010 Improving proton MR

- spectroscopy of brain tissue for noninvasive diagnostics
J. Magn. Reson. Imag. **32** 818–29
- [29] Piotta M, Moussallieh F M, Dillmann B, Imperiale A, Neuville A, Brigand C, Bellocq J P, Elbayed K and Namer I J 2009 Metabolic characterization of primary human colorectal cancers using high resolution magic angle spinning ^1H magnetic resonance spectroscopy *Metabolomics* **5** 292–301
- [30] Graveron-Demilly D, Diop A, Briguet A and Fenet B 1993 Product-operator algebra for strongly coupled spin systems *J. Magn. Reson. A* **101** 233–9
- [31] Poole C P and Farach H A 1972 *Theory of Magnetic Resonance* (New York: Wiley)
- [32] Govindaraju V, Young K and Maudsley A A 2000 Proton NMR chemical shifts and coupling constants for brain metabolites *NMR Biomed.* **13** 129–53
- [33] Rabeson H, Fauvelle F, Testylier G, Foquin A, Carpentier P, Dorandeu F, van Ormondt D and Graveron-Demilly D 2008 Quantitation with QUEST of brain HRMAS-nMR signals: application to metabolic disorders in experimental epileptic seizures *Magn. Reson. Med.* **59** 1266–73
- [34] Allouche A R 2011 Gabedit—a graphical user interface for computational chemistry softwares *J. Comput. Chem.* **32** 174–82
- [35] Cavassila S, Fenet B, van den Boogaart A, Rémy C, Briguet A and Graveron-Demilly D 1997 ER-Filter: a preprocessing technique for frequency-selective time-domain analysis *Magn. Reson. Anal.* **3** 87–92
- [36] Bales J R, Higham D P, Howe I, Nicholson J K and Sadler P J 1984 Use of high resolution proton nuclear magnetic resonance spectroscopy for rapid multi-component analysis of urine *Clin. Chem.* **30** 426

Chapter 5

General conclusions

The primary objectives of this work were to analyze the changes of NMR spectra due to environment variations, estimate the margins of this effect for HRMAS spectra of human brain biopsy samples and to develop methods to minimize undesirable concomitant spectrum distortions.

This work led to several results applicable to various NMR spectroscopy domains.

The first results belong to the spectrum quantitation subdomain and are dedicated to improve the quality of spectrum quantitation. As it was shown at the end of the third chapter, the developed methods for chemical shift optimization permit to decrease significantly the residue in the method QUEST thereby improving the quantitation quality. The methods of stretching/shrinking and splitting have shown good efficiency combined with an easy implementation and modest calculation resources requirements. If one needs to correct rapidly the peak positions in spectra such as Lactate or Creatine, these algorithms will be indispensable.

The method of optimization, based on Quantum Mechanics resimulation, enabled automatic actions, previously done only manually, adjusting step-by-step the parameters of the spin systems and resimulating the theoretical spectra again and again. The chemical shifts of numerous A_nX_m , AMX, and ABX spectra of basis-set metabolites can now be optimized before quantitation of HRMAS spectra. Currently implemented in MATLAB, the algorithm can be easily transcribed to Java, enabling its implementation in the jMRUI software. The program is implemented in modular approach. The spectrum simulation modules were successfully implemented and tested in MATLAB (which makes the program source-modifiable, cross-platform, but slower) and written in C and compiled into Win32-executable application (which makes the entire program platform-dependent but brings gains in computing time).

The described methods of optimization can increase the accuracy of further quantitation up to 30-40 percents (for high-resolved metabolite spectra) with computing time costs of minutes in the case of resimulation in Quantum Mechanics and of seconds in optimization by spectrum stretching/shrinking or splitting. These important improvements are not possible for all brain metabolites. Comparative analysis of Creatine, Lactate and Aspartate chemical shifts optimization results showed possible difficulties related to the optimization. A peak, being weakly re-

solved with respect to the noise level or to the neighbour proton peaks, can not always be found by the minimization algorithm.

An additional important result of the Quantum Mechanical optimization method (QM-QUEST) is that it enables to obtain optimized values of chemical shifts. Methods of stretching/shrinking and splitting enable only to adjust the basis-set signals to the analyzed ones. This means, that once a spectrum, for which the acquisition conditions are known, is processed, the values of corrected chemical shifts can be used in further investigations without losses of time due to repetition of the optimization procedure. Having a sufficient number of acquired spectra, one can create a database of the metabolite spectra. This database can be used directly in further quantitations or as a guide of prior knowledge in case of additional optimizations.

Another part of the work was dedicated to speed up Quantum Mechanical simulations. In Ch. 2, analyses of computation time costs for signal simulation were done. Different steps of the simulation process were investigated, such as Hamiltonian construction, Hamiltonian diagonalization and signal creation. The Hamiltonian matrix size increases exponentially as a function of the number of spins, thus preventing the simulation for large spin-system molecules. The method is based on virtual metabolite structure to minimize the number of J-coupling values. Its first part enables the decomposition of large spin structures into smaller parts and is an easily implemented method to save some simulation time. The simulation of Choline previously needed eight minutes on a 2.4 GHz processor computer. It takes now no more than some seconds. Being implemented in Java, the method can become a useful tool to accelerate the NMR-SCOPE simulations.

The next step, focused on losses of quality, is based on neglecting proton-proton J-coupling interactions of topologically distant spins. The losses of second order in accuracy led to computing time gains which increase exponentially with molecule size. The analyses of introduced inaccuracies were done and represented in maps, taking the chemical shifts and their relative differences as parameters. Being developed with the goal to extend the performance of the NMR-SCOPE module of the jMRUI software, the algorithm has been implemented in Java- and C-codes, enabling now to decrease the time of simulation of cyclic compounds (such as Inositols or Glucoses) to some seconds and to perform the simulations of spectra of metabolites such as Phosphocholine and Glycerophosphocholine, previously impossible to be simulated with the strong-coupling Hamiltonian. Generally, the precision losses of the method are not above 10% (calculated as integrated absolute difference between the ideal and fast-simulated spectra relative to the spectrum integral), and the correlation between the results of optimized simulation and the ideal one reaches 0.99 for the majority of common brain metabolites for magnetic field strengths higher than 7T. These fields (thus – frequencies above 300Mhz) are the subject of high-resolution magnetic resonance spectroscopy.

Parallel to the previously mentioned algorithm, numerous useful MATLAB functions were written to process the NMR spectra. Among them are the signal preprocessing procedures (time circular shift, phasing, ER-filter), all the opti-

mization methods (stretching, splitting etc.), simulation tools and output functions. The ensemble of proposed methods is well suited to improve quantitation of metabolites with well resolved spectra.

The proposed method QM-QUEST will be implemented in version 4 of the jMRUI software package.

Bibliography

- [1] I. I. Rabi, J. R. Zacharias, S. Millman, and P. Kusch. Milestones in Magnetic Resonance: ‘a new method of measuring nuclear magnetic moment’ . 1938. *J Magn Reson Imaging*, 2(2):131–133, 1992.
- [2] Biography of I. Rabi at Nobelprize.org, 1944. URL http://nobelprize.org/nobel_prizes/physics/laureates/1944/rabi-bio.html.
- [3] Aaron G. Filler. The History, Development and Impact of Computed Imaging in Neurological Diagnosis and Neurosurgery: CT, MRI, and DTI. *Nature Proceedings*, 2009.
- [4] G.F. Giskeødegård, T.G. Bloemberg, G. Postma, B. Sitter, M.B. Tessem, I.S. Gribbestad, T.F. Bathen, and L. Buydens. Alignment of high resolution magic angle spinning magnetic resonance spectra using warping methods. *Analytica Chimica Acta*, 2010. ISSN 0003-2670.
- [5] Charles P. Poole and Horacio A. Farach. *Theory of Magnetic Resonance*. John Wiley & Sons Inc, 1972. ISBN 0471693839.
- [6] Robin A. de Graaf. *In Vivo NMR Spectroscopy: Principles and Techniques*. Wiley, 1999. ISBN 0471983659.
- [7] E. Mark Haacke, Robert W. Brown, Michael R. Thompson, and Ramesh Venkatesan. *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. Wiley-Liss, 1999. ISBN 0471351288.
- [8] E. L. Hahn. Spin Echoes. *Physical Review*, 80(4):580–594, Nov 1950. doi: 10.1103/PhysRev.80.580.
- [9] National Institute of Advanced Industrial Science and Technology. Spectral Database for Organic Compounds, 2010. URL riodb01.ibase.aist.go.jp/sdbs/cgi-bin/cre_index.cgi.
- [10] M.G. Swanson, A.S. Zektzer, Z.L. Tabatabai, J. Simko, S. Jarso, K.R. Keshari, L. Schmitt, P.R. Carroll, K. Shinohara, D.B. Vigneron, and Kurhanewicz J. Quantitative analysis of prostate metabolites using ^1H HR-MAS spectroscopy. *Magnetic resonance in medicine*, 55(6):1257–1264, 2006. ISSN 1522-2594.

- [11] B. Martínez-Granados, D. Monleón, M.C. Martínez-Bisbal, J.M. Rodrigo, J. Olmo, P. Lluch, A. Ferrández, L. Martí-Bonmatí, and B. Celda. Metabolite identification in human liver needle biopsies by high-resolution magic angle spinning ^1H NMR spectroscopy. *NMR in Biomedicine*, 19(1):90–100, 2006. ISSN 1099-1492.
- [12] T.M. Tsang, J.L. Griffin, J. Haselden, C. Fish, and E. Holmes. Metabolic characterization of distinct neuroanatomical regions in rats by magic angle spinning ^1H nuclear magnetic resonance spectroscopy. *Magnetic resonance in medicine*, 53(5):1018–1024, 2005. ISSN 1522-2594.
- [13] M.A. Burns, J.L. Taylor, C.L. Wu, A.G. Zepeda, A. Bielecki, D. Cory, and L.L. Cheng. Reduction of spinning sidebands in proton NMR of human prostate tissue with slow high-resolution magic angle spinning. *Magnetic resonance in medicine*, 54(1):34–42, 2005. ISSN 1522-2594.
- [14] B. Sitter, S. Lundgren, T.F. Bathen, J. Halgunset, H.E. Fjosne, and I.S. Gribbestad. Comparison of HR MAS MR spectroscopic profiles of breast cancer tissue with clinical parameters. *NMR in Biomedicine*, 19(1):30–40, 2006. ISSN 1099-1492.
- [15] M.C. Martínez-Bisbal, L. Martí-Bonmatí, J. Piquer, A. Revert, P. Ferrer, J.L. Llacer, M. Piotto, O. Assemat, and B. Celda. ^1H and ^{13}C HR-MAS spectroscopy of intact biopsy samples ex vivo and in vivo ^1H MRS study of human high grade gliomas. *NMR in Biomedicine*, 17(4):191–205, 2004. ISSN 1099-1492.
- [16] A. Lazariev, A-R. Allouche, M. Aubert-Frécon, F. Fauvelle, M. Piotto, K. Elbayed, I.J. Namer, D. van Ormondt, and D. Graveron-Demilly. Optimization of metabolite basis-sets prior to quantitation in magnetic resonance spectroscopy: an approach based on quantum mechanics. *Measurement Science and Technology*, 2011. in press.
- [17] H. Rabeson, F. Fauvelle, G. Testylier, A. Foquin, P. Carpentier, F. Dorandeu, D. Van Ormondt, and D. Graveron-Demilly. Quantitation with QUEST of brain HRMAS-NMR signals: Application to metabolic disorders in experimental epileptic seizures. *Magnetic Resonance in Medicine*, 59(6):1266–1273, 2008. ISSN 1522-2594.
- [18] F. Fauvelle, G. Testylier, H. Rabeson, D. Graveron-Demilly, A. Foquin, P. Carpentier, and F. Dorandeu. Evolution of Mouse Brain Metabolism after a Convulsive Dose of Soman. In *Proc. Int. Soc. Magnetic Resonance in Medicine and European Society for Magnetic Resonance in Medicine and Biology, Joint Annual Meeting ISMRM-ESMRMB*, page 2416, 2007.
- [19] J.C. Lindon, O.P. Beckonert, E. Holmes, and J.K. Nicholson. High-resolution magic angle spinning NMR spectroscopy: Application to biomedical studies. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 55(2):79–100, 2009. ISSN 0079-6565.

- [20] F. Fauvelle, F. Dorandeu, P. Carpentier, A. Foquin, H. Rabeson, D. Graveron-Demilly, P. Arvers, and G. Testylier. Changes in mouse brain metabolism following a convulsive dose of soman: A proton HRMAS NMR study. *Toxicology*, 267(1-3):99–111, 2010. ISSN 0300-483X.
- [21] A. Croitor Sava, M.C. Martinez-Bisbal, S. Van Huffel, J.M. Cerda, D.M. Sima, and B. Celda. Ex vivo high resolution magic angle spinning metabolic profiles describe intratumoral histopathological tissue properties in adult human gliomas. *Magnetic Resonance in Medicine*, 2010. ISSN 1522-2594.
- [22] A.J. Wright, G.A. Fellows, J.R. Griffiths, M. Wilson, B.A. Bell, and F.A. Howe. Ex-vivo HRMAS of adult brain tumours: metabolite quantification and assignment of tumour biomarkers. *Molecular Cancer*, 9(1):66, 2010. ISSN 1476-4598.
- [23] Norman F. Ramsey. Magnetic shielding of nuclei in molecules. *Phys. Rev.*, 78(6):699–703, Jun 1950. doi: 10.1103/PhysRev.78.699.
- [24] Jacek W. Hennel and Jacek Klinowski. *New Techniques in Solid-State NMR (Topics in Current Chemistry)*. Springer, 2005. ISBN 3540221689.
- [25] L.L. Cheng, M.J. Ma, L. Becerra, T. Ptak, I. Tracey, A. Lackner, and R.G. Gonzalez. Quantitative neuropathology by high resolution magic angle spinning proton magnetic resonance spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 94(12):6408, 1997.
- [26] V. Govindaraju, K. Young, and A.A. Maudsley. Proton NMR chemical shifts and coupling constants for brain metabolites. *NMR in Biomedicine*, 13(3):129–153, 2000.
- [27] M. Menzel, E.M.R. Doppenberg, A. Zauner, J. Soukup, M.M. Reinert, and R. Bullock. Increased inspired oxygen concentration as a factor in improved brain tissue oxygenation and tissue lactate levels after severe human head injury. *Journal of Neurosurgery: Pediatrics*, 91(1), 1999.
- [28] S. Inao, A. Marmarou, G.D. Clarke, B.J. Andersen, P.P. Fatouros, and H.F. Young. Production and clearance of lactate from brain tissue, cerebrospinal fluid, and serum following experimental brain injury. *Journal of neurosurgery*, 69(5):736–744, 1988.
- [29] G. van Hall, M. Strømstad, P. Rasmussen, Ø. Jans, M. Zaar, C. Gam, B. Quistorff, N.H. Secher, and H.B. Nielsen. Blood lactate is an important energy source for the human brain. *Journal of Cerebral Blood Flow & Metabolism*, 29(6):1121–1129, 2009.
- [30] List of B vitamins. URL http://en.wikipedia.org/wiki/Vitamin_B#List_of_B_vitamins.

- [31] J.B. Pouillet. *Quantification and Classification of Magnetic Resonance Spectroscopic Data for Brain Tumor Diagnosis*. PhD thesis, Faculty of Engineering, K.U.Leuven, Nov 2008.
- [32] Y. Kinoshita and A. Yokota. Absolute concentrations of metabolites in human brain tumors using in vitro proton Magnetic Resonance Spectroscopy. *NMR in Biomedicine*, 10(1):2–12, 1997.
- [33] J.T. Brosnan and M.E. Brosnan. The sulfur-containing amino acids: an overview. *Journal of Nutrition*, 136(6):1636S, 2006.
- [34] Z. Tong, T. Yamaki, K. Harada, and K. Houkin. In vivo quantification of the metabolites in normal brain and brain tumors by proton MR spectroscopy using water as an internal standard. *Magnetic resonance imaging*, 22(5):735–742, 2004.
- [35] G. Tedeschi, N. Lundbom, R. Raman, S. Bonavita, J.H. Duyn, J.R. Alger, and G. Di Chiro. Increased choline signal coinciding with malignant degeneration of cerebral gliomas: a serial proton magnetic resonance spectroscopy imaging study. *Journal of Neurosurgery: Pediatrics*, 87(4), 1997.
- [36] Y. Wang and S.J. Li. Differentiation of metabolic concentrations between gray matter and white matter of human brain by in vivo ^1H Magnetic Resonance Spectroscopy. *Magnetic Resonance in Medicine*, 39(1):28–33, 1998.
- [37] J.T. Brosnan. Interorgan amino acid transport and its regulation. *The Journal of nutrition*, 133(6):2068S, 2003. ISSN 0022-3166.
- [38] P. Newsholme, MMR Lima, J. Procopio, TC Pithon-Curi, SQ Doi, RB Bazotte, and R. Curi. Glutamine and glutamate as vital metabolites. *Brazilian journal of medical and biological research*, 36(2):153–163, 2003. ISSN 0100-879X.
- [39] O.A.C. Petroff. Book Review: GABA and Glutamate in the Human Brain. *The Neuroscientist*, 8(6):562, 2002. ISSN 1073-8584.
- [40] E. Popa. *Algorithms for Handling Arbitrary Lineshape Distortions in Magnetic Resonance Spectroscopy and Spectroscopic Imaging*. PhD thesis, Université Claude Bernard Lyon 1, France, Jul 2010.
- [41] P.S. Ariyannur, J.R. Moffett, P. Manickam, N. Pattabiraman, P. Arun, A. Nitta, T. Nabeshima, C.N. Madhavarao, and A. Namboodiri. Methamphetamine-induced neuronal protein NAT8L is the NAA biosynthetic enzyme: Implications for specialized acetyl coenzyme A metabolism in the CNS. *Brain Research*, 1335:1–13, 2010.
- [42] V. Mehta and MAA Namboodiri. N-acetylaspartate as an acetyl source in the nervous system. *Molecular Brain Research*, 31(1-2):151–157, 1995.

- [43] H. Rabeson, F. Fauvelle, H. Ratiney, C. Cudalbu, S. Cavassila, D. van Ormondt, and D. Graveron-Demilly. Quantitation with QUEST of ER-filtered HRMAS NMR signals. *Proc. ProRISC, IEEE Benelux, Veldhoven, The Netherlands*, pages 171–175, 2006.
- [44] G. Reynolds, M. Wilson, A. Peet, and T.N. Arvanitis. An algorithm for the automated quantitation of metabolites in in vitro NMR signals. *Magnetic resonance in medicine*, 56(6):1211–1219, 2006. ISSN 1522-2594.
- [45] F.P. DiCamillo, M.G. Swanson, J. Kurhanewicz, and S.J. Nelson. Assessment of HR-MAS NMR Prostate Biopsy Tissue Spectra with Principal Component Analysis and Metabolite Quantification. In *Proc. Int. Soc. Magnetic Resonance in Medicine, 14th Scientific Meeting and Exhibition*, page 2534, 2006.
- [46] P. B. Barker, B. J. Soher, S. J. Blackband, J. C. Chatham, V. P. Mathews, and R. N. Bryan. Quantitation of proton NMR spectra of the human brain using tissue water as an internal concentration reference. *NMR Biomed*, 6(1): 89–94, 1993.
- [47] A.D. Bain, B.J. Fahie, T. Kozluk, and W.J. Leigh. Improvements in the quantitation of NMR spectra by the use of statistical methods. *Canadian Journal of Chemistry*, 69(8):1189–1192, 1991.
- [48] H. Ratiney, M. Sdika, Y. Coenradie, S. Cavassila, D. van Ormondt, and D. Graveron-Demilly. Time-domain semi-parametric estimation based on a metabolite basis set. *NMR Biomed*, 18(1):1–13, Feb 2005. doi: 10.1002/nbm.895. URL <http://dx.doi.org/10.1002/nbm.895>.
- [49] A. van den Boogaart. Quantitative Data Analysis of In Vivo MRS Data Sets. *Magnetic resonance in chemistry*, 35:S146–S152, 1997.
- [50] G. H. Golub and V. Pereyra. The differentiation of pseudo-inverses and non-linear least squares problems whose variables separate. *SIAM J. Numer. Anal.*, 10:413–432, 1973. URL <http://www.intermnet.ua.es/inteRMNet/c59/node32.html>.
- [51] S.W. Provencher. Estimation of metabolite concentrations from localized in vivo proton NMR spectra. *Magnetic Resonance in Medicine*, 30(6):672–679, 1993. ISSN 1522-2594.
- [52] D. Stefan, F. Di Cesare, A. Andrasescu, E. Popa, A. Lazariiev, E. Vescovo, O. Strbak, S. Williams, Z. Starcuk, M. Cabanas, D. van Ormondt, and D. Graveron-Demilly. Quantitation of magnetic resonance spectroscopy signals: the jMRUI software package. *Measurement Science and Technology*, 20: 104035, 2009. doi: 10.1088/0957-0233/20/10/104035.
- [53] A. Naressi, C. Couturier, JM Devos, M. Janssen, C. Mangeat, R. de Beer, and D. Graveron-Demilly. Java-based graphical user interface for the MRUI quantitation package. *Magnetic Resonance Materials in Biology, Physics, and Medicine*, 12(2-3):141–152, 2001.

- [54] A. Naressi, C. Couturier, I. Castang, R. de Beer, and D. Graveron-Demilly. 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, 2001.
- [55] Lorentzian function. URL http://en.wikipedia.org/wiki/Lorentzian_function.
- [56] M. van der Graaf and A. Heerschap. Effect of cation binding on the proton chemical shifts and the spin-spin coupling constant of citrate. *Journal of Magnetic Resonance-Series B*, 112(1):58–62, 1996. ISSN 1064-1866.
- [57] E.S. Stevens, N. Sugawara, G.M. Bonora, and C. Toniolo. Conformational analysis of linear peptides. 3. Temperature dependence of NH chemical shifts in chloroform. *Journal of the American Chemical Society*, 102(23):7048–7050, 1980.
- [58] Y. Ishihara, A. Calderon, H. Watanabe, K. Okamoto, Y. Suzuki, K. Kuroda, and Y. Suzuki. A precise and fast temperature mapping using water proton chemical shift. *Magnetic Resonance in Medicine*, 34(6):814–823, 1995.
- [59] N.J. Baxter and M.P. Williamson. Temperature dependence of ^1H chemical shifts in proteins. *Journal of biomolecular NMR*, 9(4):359–369, 1997.
- [60] Z Atieh, A-R. Allouche, A. Lazariiev, D. van Ormondt, D. Graveron-Demilly, and M. Aubert-Frécon. DFT calculation of ^1H chemical shifts and simulated and experimental NMR spectra for sarcosine. In *The 16th International Scientific Conference, Beirut Arab University, Beirut, Lebanon*, Nov 2009.
- [61] A. Lakatos, I. Bányai, P. Decock, and T. Kiss. Time-Dependent Solution Speciation of the Al^{III} -Citrate System: Potentiometric and NMR Studies. *European Journal of Inorganic Chemistry*, 2001(2):461–469, 2001.
- [62] O. Robert, J. Sabatier, D. Desoubzdanne, J. Lalande, S. Balayssac, V. Gilard, R. Martino, and M. Malet-Martino. pH optimization for a reliable quantification of brain tumor cell and tissue extracts with ^1H NMR: focus on choline-containing compounds and taurine. *Analytical and bioanalytical chemistry*, 399:987–999, 2011. ISSN 1618-2650.
- [63] L.D. Hall and S.L. Talagala. Mapping of pH and temperature distribution using chemical-shift-resolved tomography. *Journal of Magnetic Resonance*, 65(3):501–505, 1985.
- [64] RA Wevers, U. Engelke, U. Wendel, JG De Jong, FJ Gabreels, and A. Heerschap. Standardized method for high-resolution ^1H -NMR of cerebrospinal fluid. *Clinical chemistry*, 41:744, 1995.

- [65] H. Ratiney, Y. Coenradie, S. Cavassila, D. van Ormondt, and D. Graveron-Demilly. Time-domain quantitation of 1 H short echo-time signals: background accommodation. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 16(6):284–296, 2004. ISSN 0968-5243.
- [66] Bruker BioSpin GmbH, 2001. URL http://www.chem.wisc.edu/~cic/nmr/Guides/Bruker_AVANCE_manuals/Acquisition%20Reference%20XwinNMR%20%28April%202002%29%20-%20acqref.pdf.
- [67] ACD Labs, 2010. URL <http://www.acdlabs.com/>.
- [68] A.R. Allouche. Gabedit – A graphical user interface for computational chemistry softwares. *Journal of Computational Chemistry*, 32:174–182, 2011. ISSN 1096-987X.
- [69] N. Chandrakumar and S. Subramanian. *Modern techniques in high-resolution FT-NMR*. Springer-Verlag, New York, NY, 1987.
- [70] Erik Meijering. A Chronology of Interpolation: From Ancient Astronomy to Modern Signal and Image Processing. In *Proceedings of the IEEE*, volume 90, pages 319–342, Mar 2002. doi: 10.1109/5.993400.
- [71] Y. Nievergelt. *Splines in Single and Multivariable Calculus*, 1993.
- [72] S. McKinley and M. Levine. Cubic spline interpolation, 2006.
- [73] John H. Mathews and Kurtis D. Fink. *Numerical Methods Using MATLAB*. Simon & Schuster, 1998. ISBN 0132700425.
- [74] Amos Ron. Introduction to numerical methods. Oct 2006. URL <http://pages.cs.wisc.edu/~amos/412/>.
- [75] Trigonometric interpolation polynomial, 2010. URL http://en.wikipedia.org/wiki/Trigonometric_interpolation_polynomial.
- [76] H. Nyquist. Certain topics in telegraph transmission theory. *Proceedings of the IEEE*, 90(2):280–305, 2002.
- [77] Hiroshi Akima. A new method of interpolation and smooth curve fitting based on local procedures. *J. ACM*, 17(4):589–602, 1970. ISSN 0004-5411. doi: <http://doi.acm.org/10.1145/321607.321609>.
- [78] N. Shamsundar. Akima Interpolation. To interpolate smooth curve through given points on a plane., Jan 2006. URL <http://www.mathworks.com/matlabcentral/fileexchange/1814-akima-interpolation>.
- [79] A. Lazariiev, M. Piotto, K. Elbayed, I.J. Namer, D. van Ormondt, and D. Graveron-Demilly. Simple correction of chemical shift changes in Magnetic Resonance Spectroscopy quantitation. In *Imaging Systems and Techniques (IST), 2010 IEEE International Conference, Thessaloniki, Greece*, pages 365–368. IEEE, 2010.

- [80] P. Alusta, I. Im, B.A. Pearce, R.D. Beger, R.M. Kretzer, D.A. Buzatu, and J.G. Wilkes. Improving proton MR spectroscopy of brain tissue for noninvasive diagnostics. *Journal of Magnetic Resonance Imaging*, 32(4):818–829, 2010. ISSN 1522-2586.
- [81] Danielle Demilly. *Mise au point d'un programme d'analyse de spectres RMN. Application a des molécules hétérocycliques possédants un élément de symétrie*. PhD thesis, Faculté des sciences de l'Université de Lyon, France, Jul 1970.
- [82] Stirling's approximation. URL http://en.wikipedia.org/wiki/Stirling%27s_approximation.
- [83] Gradient descent. URL http://en.wikipedia.org/wiki/Gradient_descent.
- [84] D. Graveron-Demilly, A. Diop, A. Briguet, and B. Fenet. Product-operator algebra for strongly coupled spin systems. *Journal of Magnetic Resonance, Series A*, 101(3):233–239, 1993. ISSN 1064-1858.
- [85] S. Cavassila, B. Fenet, A. van den Boogaart, C. Rémy, A. Briguet, and D. Graveron-Demilly. ER-Filter: a preprocessing technique for frequency-selective time-domain analysis. *J. Magn. Reson. Anal*, 3:87–92, 1997.
- [86] J.R. Bales, D.P. Higham, I. Howe, J.K. Nicholson, and P.J. Sadler. Use of high-resolution proton nuclear magnetic resonance spectroscopy for rapid multi-component analysis of urine. *Clinical chemistry*, 30(3):426, 1984.

Publication list

Peer-reviewed papers in international journals

- A. Lazariiev, A-R. Allouche, M. Aubert-Frécon, F. Fauvelle, M. Piotta, K. Elbayed, I.J. Namer, D. van Ormondt, and D. Graveron-Demilly. Optimization of metabolite basis-sets prior to quantitation in magnetic resonance spectroscopy: an approach based on quantum mechanics. *Measurement Science and Technology*, 2011. in press.
- Z Atieh, A-R. Allouche, A. Lazariiev, D. van Ormondt, D. Graveron-Demilly, and M. Aubert-Frécon. DFT calculation of ^1H chemical shifts and simulated and experimental NMR spectra for sarcosine. *Chem. Phys. Letts*, 492:297–301, 2010.
- D. Stefan, F.D. Cesare, A. Andrasescu, E. Popa, A. Lazariiev, E. Vescovo, O. Strbak, S. Williams, Z. Starcuk, M. Cabanas, et al. Quantitation of magnetic resonance spectroscopy signals: the jMRUI software package. *Measurement Science and Technology*, 20:104035, 2009. doi:10.1088/0957-0233/20/10/104035.

Peer-reviewed IEEE Proceedings

- A. Lazariiev, M. Piotta, K. Elbayed, I.J. Namer, D. van Ormondt, and D. Graveron-Demilly. Simple correction of chemical shift changes in magnetic resonance spectroscopy quantitation. In *Imaging Systems and Techniques (IST), 2010 IEEE International Conference*, Thessaloniki, Greece, pages 365–368.

Peer-reviewed international conferences

Oral communications

- A. Lazariiev, A-R. Allouche, M. Aubert-Frécon, F. Fauvelle, M. Piotta, K. Elbayed, I.J. Namer, D. van Ormondt, and D. Graveron-Demilly. A Quantum Mechanics-based approach for optimization of metabolite basis-sets. Application to quantification of HRMAS-NMR signals. In *RITS GRAMM*, 4-7 April 2011, Rennes, France.

- A. Lazareiev, M. Piotto, K. Elbayed, I.J. Namer, D. van Ormondt, and D. Graveron-Demilly. Simple correction of chemical shift changes in magnetic resonance spectroscopy quantitation. In *Imaging Systems and Techniques (IST), 2010 IEEE International Conference*, 1-2 July 2010, Thessaloniki, Greece.

Posters

- A. Lazareiev, M. Piotto, K. Elbayed, I.J. Namer, D. van Ormondt, and D. Graveron-Demilly. Optimization of metabolite basis-sets prior to quantitation: a quantum mechanics approach. In *ISMRM 2011*, page 4771. Montréal, Canada. ISMRM, 7-13 May 2011. e-Poster.
- A. Lazareiev, M. Piotto, K. Elbayed, I.J. Namer, D. van Ormondt, and D. Graveron-Demilly. Simple correction of chemical shift changes in magnetic resonance spectroscopy quantitation. In *ISMRM-ESMRMB 2010*, Stockholm, Sweden, page 3318. May 2010. e-Poster.
- Z Atieh, A-R. Allouche, A. Lazareiev, D. van Ormondt, D. Graveron-Demilly, and M. Aubert-Frécon. DFT calculation of ^1H chemical shifts and simulated and experimental NMR spectra for sarcosine. In *The 16th International Scientific Conference, Beirut Arab University*, Beirut, Lebanon, Nov 2009.
- D. Stefan, A. Andrasescu, E. Popa, A. Lazareiev, O. Strbak, M. Cabañas, Z. Starcuk, D. van Ormondt, and D. Graveron-Demilly. jMRUI version 4 : A plug-in platform. In *Int. Soc. Magnetic Resonance in Medicine, ISMRM, 17th Scientific Meeting and Exhibition*, Honolulu, USA, page 4340. ISMRM, April 2009. e-Poster.

Others

- A. Lazareiev. Problems of pH-dependence of NMR spectra and possible solutions. In *Meeting of the European project FAST*, Brno, Czech Republic, June 2009. URL <http://www.fast-mariecurie-rtn-project.eu/>.

Appendix

Interpolation errors (Eq. 3.2.19)

Linear interpolation

The error of the linear interpolation on an interval $[a, b]$ is

$$E_{lin} \leq \frac{\|f''\|_{[a, a+h]} h^2}{8} \quad (5.0.1)$$

where

$$\|f\|_{[a, b]} = \max_{i \in [a, b]} |f(i)| \quad (5.0.2)$$

and h signifies the interpolation interval. For the equidistant points case (see Fig. 5.1):

$$h = \frac{b - a}{N} \quad (5.0.3)$$

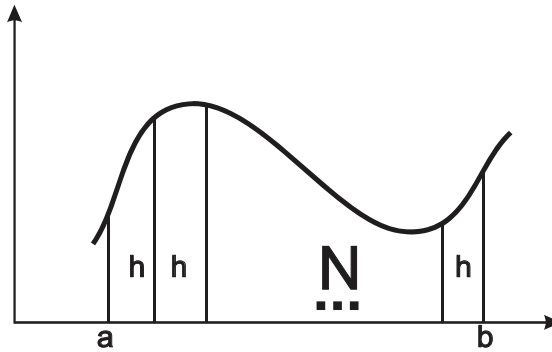


Figure 5.1: A function to be interpolated with N data points on the interval $[a, b]$

For the harmonically oscillating function with the angular frequency ω and the amplitude A , the maximum maximum estimation (for $b - a = T/2$, the limit of the Nyquist criterion) gives:

$$\|f''\|_{[a, b]} = A\omega^2 \quad (5.0.4)$$

and the error:

$$E_{lin} \leq \frac{A\omega^2 t_s^2}{8} \quad (5.0.5)$$

where t_s is the sampling time.

$$E_{lin} \leq \frac{A\pi^2\nu^2}{2}t_s^2 \quad (5.0.6)$$

Cubic spline interpolation

Similar calculations for the cubic spine interpolation lead to:

$$\begin{aligned} E_{cub} &\leq \frac{\|5f^{(4)}\|_T}{384}h^4 \\ E_{cub} &\leq \frac{5A\omega^4}{384}t_s^4 \end{aligned} \quad (5.0.7)$$

Integration to obtain Eq. 3.2.28

Eq. 3.2.28.1

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{6f}{\omega^2 + \alpha^2} d\omega &= \frac{6A\alpha}{\pi} \int_{-\infty}^{\infty} \frac{d\omega}{\omega^2 + \alpha^2} = \\ &\frac{6A\alpha}{\pi} \frac{1}{2\alpha^2} \left(\frac{\omega}{\omega^2 + \alpha^2} + \frac{1}{\alpha} \arctan \frac{\omega}{\alpha} \right) \Big|_{-\infty}^{\infty} = \\ &\frac{3A}{\alpha\pi} \left(0 + \frac{1}{\alpha} \left(\frac{\pi}{2} - \left(-\frac{\pi}{2} \right) \right) \right) = \frac{3A}{\alpha^2} \end{aligned} \quad (5.0.8)$$

Eq. 3.2.28.2

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{120f}{(\omega^2 + \alpha^2)^2} d\omega &= \frac{120A\alpha}{\pi} \int_{-\infty}^{\infty} \frac{d\omega}{(\omega^2 + \alpha^2)^3} = \\ &\frac{120A\alpha}{\pi} \frac{1}{2\alpha^2} \left(\frac{\omega}{(\omega^2 + \alpha^2)^2} + \frac{3\omega}{2\alpha^2(\omega^2 + \alpha^2)} + \frac{3}{2\alpha^3} \arctan \frac{\omega}{\alpha} \right) \Big|_{-\infty}^{\infty} = \\ &\frac{120A}{2\pi\alpha} \left(0 + 0 + \frac{3\pi}{2\alpha^3} \right) = \frac{45A}{\alpha^4} \end{aligned} \quad (5.0.9)$$

Spectrum errors (Eq. 3.2.29)

Linear interpolation

$$E_{lin} \leq \frac{1}{8} \cdot \frac{3A}{\alpha^2} \cdot h^2 \quad (5.0.10)$$

where

$$h = \frac{\omega_{\max}}{N_p} = \frac{2\pi\nu_{\max}}{N_p} = \frac{2\pi}{N_p t_s} \quad (5.0.11)$$

Thus,

$$E_{lin} \leq \frac{1}{8} \cdot \frac{3A}{\alpha^2} \cdot \frac{4\pi^2}{N_p^2 t_s^2} \quad (5.0.12)$$

$$E_{lin} \leq \frac{3A\pi^2}{2\alpha^2 N_p^2 t_s^2} \quad (5.0.13)$$

Spline interpolation

$$E_{cub} \leq \frac{5}{384} \cdot \frac{45A}{\alpha^4} \cdot h^4$$

$$E_{cub} \leq \frac{5}{384} \cdot \frac{45A}{\alpha^4} \cdot \frac{16\pi^4}{N_p^4 t_s^4} \quad (5.0.14)$$

$$E_{cub} \leq \frac{75A\pi^4}{8\alpha^4 N_p^4 t_s^4}$$

Fragmentation implementation in jMRUI

Current realization of simulation by BCH formalism, implemented in jMRUI, has a vector of multiplicity values for a simulated spin system. In fact, the definition used in NMR-SCOPE corresponds to the pseudo-intensity previously described in this work. This enables to use the NMR-SCOPE multiplicity values as the fragmentation procedure output.

Data, intercepted right before the start of simulation were transformed in a structure *Metabolite* with the fields:

- double[] Shifts
- int[] Multiplicity
- double[] Intensity
- double[][] Couplings
- Metabolite[] Components

So, the last field has the same representation and ready to be simulated. It is passed to the simulation procedure. Simulating one by one the “*Components*” of the spin-system from the respective array, one obtains the series of signals which will be added to get the final output.

The changed version of the sequence window interface of NMR-SCOPE is presented in Fig. 5.2.

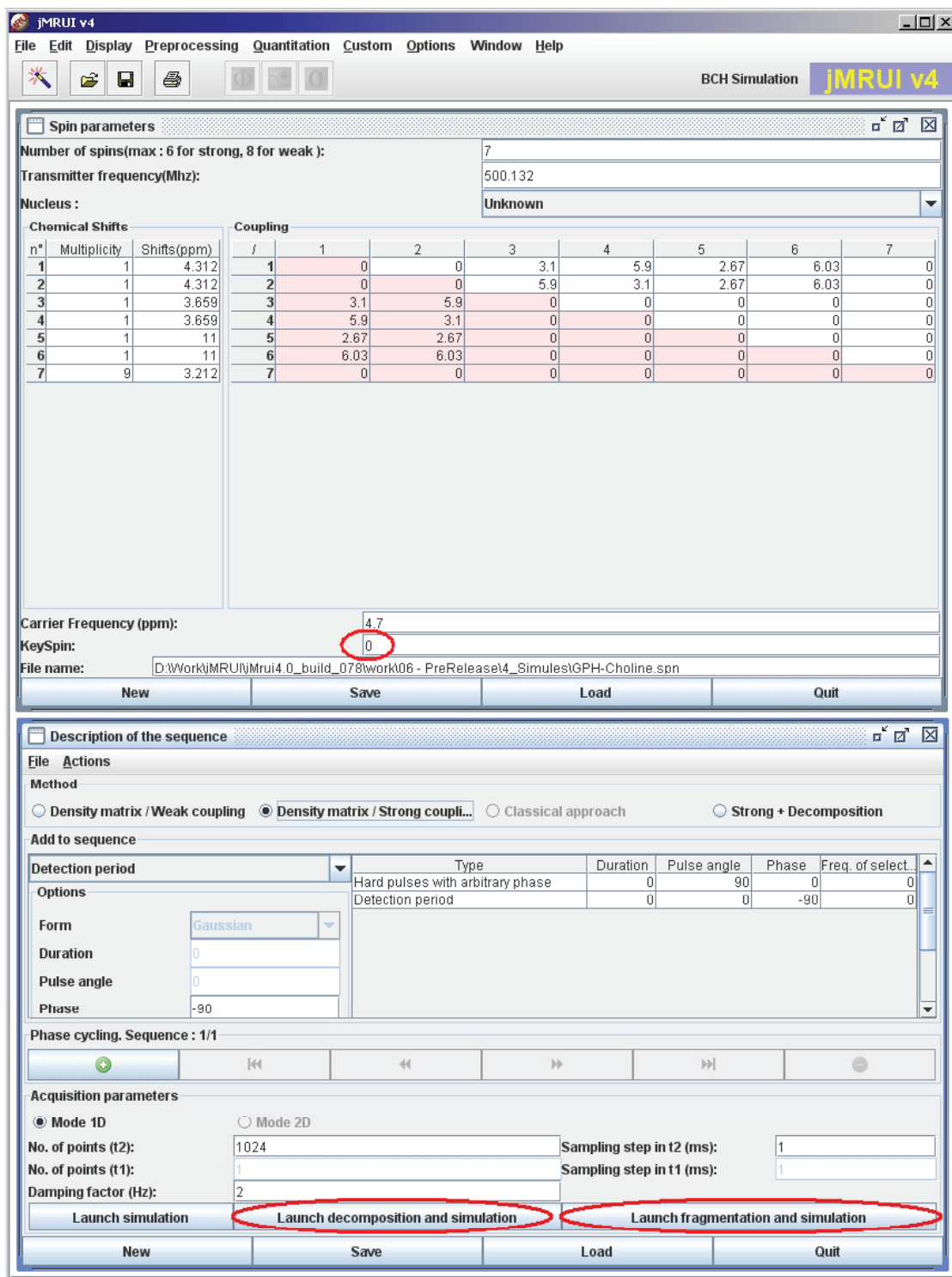


Figure 5.2: Screen shot of the NMR-SCOPE module of jMRUI. Please, note the added options of simulation: simple decomposition (suitable for Creatine, Choline, GPC) and fragmentation (for Glucose, myo-Inositol, PhC, GPC).