Single-Voxel Magnetic Resonance Spectroscopy: Assessment of the Detection Limits of Neurotransmitter Signals

Konstantin Edward Bosbach

University of Freiburg, Germany

visiting at the University of Oxford, UK

Supervisor: Prof Michael Bock
    assisted by Dr Johannes Fischer
(University of Freiburg, Germany)

In collaboration with:
    Prof Charlotte Stagg
    Dr William Clarke
(University of Oxford, UK)

With heartfelt thanks to the Lewintan family.

Abstract

Magnetic resonance spectroscopy (MRS) can quantify metabolite concentrations in the living human brain non-invasively. Unlike magnetic resonance imaging, which measures the abundant MR signal of ^{1}H atoms in water, the metabolites detected by MRS are much lower in concentration. As practical data acquisition times in humans are limited to less than an hour, MRS is a technique with limited signal-to-noise ratio (SNR). Still, in recent years MRS has been utilized to monitor small changes in neurotransmitters during brain activation. MRS spectra of different metabolites can have overlapping spectral resonances which makes it difficult to separate them by spectral fitting. As MRS data can be corrupted by noise and artefacts, it remains complex to quantify measurement uncertainties and to provide estimates for the required measurement setup to achieve a given measurement precision.

This thesis presents a Monte-Carlo simulation software to estimate the uncertainties of MRS-measured concentrations. Using in vivo MRS data acquired at 7 T, measurement settings are simulated to assess both signal quality and metabolite concentrations. By interpolating the simulated data, a wide range of measurement parameters can be quickly inspected. The software is used to perform both power calculations for functional MRS study designs, and to estimate the statistical power in potential measurement situations. It allows for planning of statistically powerful MRS studies, for clinical and neuroscientific purpose, to measure metabolite concentrations in the human brain.

For different study designs detecting Glutamate concentration changes, this method was used to compare tradeoffs in uncertainty and power calculation considerations. The results imply, that the study size can be reduced reliably, from 11 to 3 participants, by using event related paradigms instead of the established block studies.

Contents

List of Abbreviations

B_0 . . . . External magnetic field.

FID . . . . . . Free Induction Decay [signal].

fMRS . . . . . Functional Magnetic Resonance Spectroscopy.

GABA . . . . . γ-Aminobutyric acid.

γ . . . . Gyromagnetic Ratio.

¹H . . . . Hydrogen atom. Atomic number of 1.

I . . . . Spin quantum number.

μ . . . . Magnetic Moment.

M_0 . . . . Longitudinal equilibrium magnetisation.

MR . . . . . . Magnetic Resonance.

MRI . . . . . . Magnetic Resonance Imaging.

MRS . . . . . . Magnetic Resonance Spectroscopy.

NMR . . . . . Nuclear Magnetic Resonance.

noise-SD . . . Standard-Deviation of the signal noise.

RF . . . . Radio-Frequency [pulse].

SD . . . . Standard-Deviation of a normal distribution.

SNR . . . . . . Signal-to-Noise Ratio.

T . . . . Tesla, unit of magnetic flux density.

TE . . . . Echo Time.

TR . . . . Repetition Time.

t-test . . . . . Student's t-test.

1. Introduction

Although research has greatly furthered the understanding of brain processes in general, the application of this information in clinical care is lacking. By supplying easy to use tools, this thesis aims to close the gap between state-of-the-art-imaging techniques, and their practical application. [1]

Although the presented methods are applicable on a vast amount of metabolites, this thesis has a special interest in measuring γ-Aminobutyric acid (GABA). It is the main inhibitory neurotransmitter in the human central nervous system [2]; and has a wide functionality, even ranging to excitatory mechanisms in early embryological development [3].Alterations of GABA concentrations have been linked to many neurological deficits, such as autism [4] and dystonia [5]. Additionally, it has been used to treat epilepsy [6], and neuropathic pain [7].

This project provides code for evaluating the potential of magnetic resonance spectroscopy (MRS) applications, by estimating their sensitivity. The aim is to understand how sensitive MRS measurements are at determining brain metabolite concentrations non-invasively, in research and clinical work, allowing further development of measurements (and hopefully diagnosis) that are not deemed feasible today.

The phenomenon of Nuclear-Magnetic-Resonance (NMR, henceforth MR) is best known for its application in magnetic resonance imaging (MRI). Clinical MRI uses the MR signal of ^{1}H nuclei in the human body, to produce tissue images. However, the same MR technique can be used for in vivo spectral analysis, as well. This Magnetic-Resonance-Spectroscopy (MRS) provides specific information about the individual chemical environments of the inspected nuclei. Each molecule interacts with the signals of its measured nuclei, creating a distinctive signal spectrum. By

using those spectra, tissue signals can be disassembled into their contributing parts, returning an estimation of molecule concentrations.

This thesis uses Monte-Carlo Simulations, to evaluate the likeness of in vivo measurements of being successful. Monte-Carlo simulations use statistical simulation (i.e., repeatedly rolling dice), to get a profile of the underlying distribution (posterior distribution). More specifically, the `FSL-MRS` software toolbox is used to generate and analyse simulated patient brain spectra. In this way, it is predicted, how close a real measurement will come to correctly estimating the concentration. Using the characteristic parameters of a measurement, and by comparing the results of several runs, a better estimate of the measurement uncertainty can be given.

The function of GABA in physiological processes such as sleep [8], motor learning [9], and network structure functions [10] is of great interest, with GABA even being predicted to contribute to the understanding of self-consciousness [11].

Measuring GABA may help to understand and diagnose chronic diseases, and the recovery from severe illness. [12] For example, GABA MRS might make episodes of bipolar disorder predictable, thus improving preventive therapy options, benefitting patient quality of life, and saving on resource intense therapeutical intermissions of fully grown episode. [13]

In general, MRS measurements take longer and have lower resolution than MRI methods. Signal strengths in MRS are far weaker than in MRI, whilst noise is identical. This arises from substantially lower concentrations of metabolites, in the 1-20 millimolar range, compared with the approximately 20-55 molar water. As such, for a sufficient signal-to-noise ratio (SNR), MRS requires imaging of larger volumes or requires longer acquisition times to average signal. Thus, optimisation of MRS acquisition techniques is crucial of clinical implementation.

Especially the detection of GABA in clinical routine suffers from both its low in vivo concentration (ca. 1 millimolar), and its emitted spectrum: All 3 resonance spectra overlap with more intense resonances, e.g., creatine, glutamate, and macromolecules. All three resonances are coupled (j-coupling, see *Section 2.1.4*
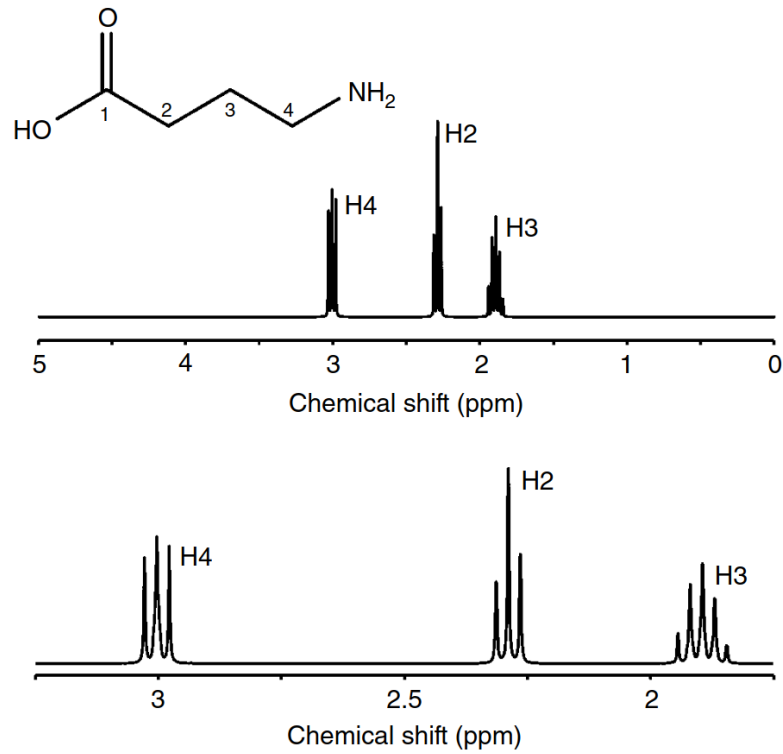
Figure 1.1: The GABA molecule contains three methylene groups, which contribute a total of six observable protons to the spectrum. Resonance spectra are emitted by the protons at the second carbon atom (C_2) of GABA (GABA-H2), as well as GABA-H3, and GABA-H4, appearing centred at 2.28 ppm, 1.89 ppm, and 3.00 ppm, respectively. [from 14, p.57]

Spectral Editing). Therefore, they produce complex and broad peaks (see Figure 1.1), which have a relatively low signal-to-noise ratio (SNR). [14, p.56 ff.]

With this thesis, I aim to create easy-to-use software, to enable medical personnel, as well as scientists, to estimate the sensitivity of MRS methods. By using in vivo data sets with the `FSL-MRS` library, in vivo MRS measurements are replicated, as they appear in clinical diagnostics and trials. By entering the characterised parameters of their measurement setup (such as noise and metabolite concentrations), the user can straightforwardly simulate the measurement conditions. As the software is able to return the unknown variables (e.g., the number of subjects needed to achieve a certain level of significance), users are able to better plan trials, diagnostics, and protocols.

2. Theory

This chapter introduces the basic principles of MRS and its application in GABA quantification, for which the software was developed. causes the effects we can witness with Magnetic Resonance Spectroscopy (MRS). Primary source for both is the textbook by de Graaf. [14, pp. 5-28] Further citation is added in place.

2.1 Nuclear Magnetic Resonance Spectroscopy

The in vivo concentrations of neurotransmitters in the human brain can be measured using MRS. Just as in other MR applications, like MRI, the MRS signal originates from the nuclear magnetic moment.

2.1.1 Nuclear Magnetisation

In 1922, the Stern-Gerlach experiment [15], was able to split a beam of silver atoms into two, by applying an external magnetic field. The results implied, that by applying the magnetic field on the beam, the intrinsic orientation of atoms is quantisised.[1] Isidor Rabi enhanced this technique to actively vary this atom orientation [16], developing the basics of today's MR.[2]

In vivo MR measurements are most commonly performed on hydrogen nuclei. ^{1}H is highly abundant in the human body, with water alone making up around 5-20% of fat, 50% of bone, and more than 70% of brain tissue.[17] Apart from water, ^{1}H exists in most molecules of the body, such as the brain metabolites tried to measure here.

---

[1]Otto *Stern* received the Nobel Price in 1943.
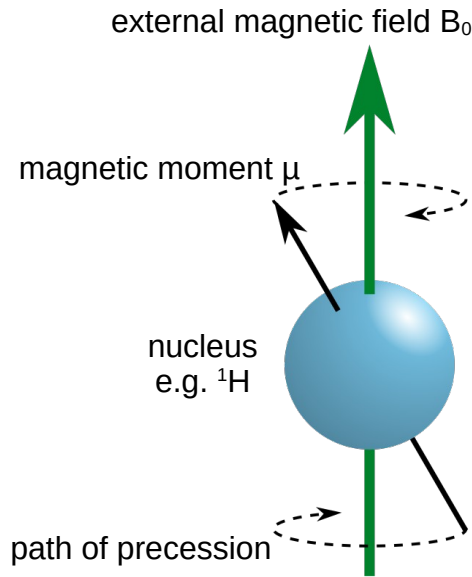[2]He received the Nobel Prize in 1944.

Figure 2.1: Precession of a particle with non-zero total spin, within an external field. [from 19]

According to their properties in relativistic quantum mechanics, hydrogen nuclei have a non-zero "spin" quantum number of $I = \frac{1}{2}$; it is associated with the magnetic moment μ. The intrinsic nuclear spin has no classical analogue.[3]

Figure 2.1 shows a nucleus (with $I \neq 0$) in an external magnetic field, B_0. The axis of magnetic moment μ precesses around the external field at a characteristic frequency, the Larmor frequency.

The Larmor frequency is given by

$$\nu_0 = \frac{\gamma}{2\pi} \cdot B_0 \,.\tag{2.1}$$

It is dependent on the nucleus-specific gyromagnetic ratio γ, and the strength of the external field B_0.

Samples contain many molecules, commonly quantified in moles (mol), each containing 6×10^{23} molecules. Usually, their orientation is arbitrary. However, if a B_0 field is applied, it interacts with the magnetic momenta of the nuclei, forcing them to align either parallel or anti-parallel along the magnetic field lines. It would be intuitive to assume the same amount of parallel and anti-parallel states. However, because the spins of different nuclei interact, there are more nuclei oriented in the

---

[3]"Two-valuedness not describable classically" it was first named by Wolfgang Pauli in 1924.[18]

same direction. This happens due to a statistical principle, and attributes for the MR signal origin. The Boltzmann distribution can be used to describe this longitudinal equilibrium magnetisation M_0, using a sample size of N particles, with I spin, as well as Planck constant h and Boltzmann constant k_B:

$$M_0 = \frac{N\gamma^2 h^2 I(I+1)}{16\pi^2 k_B T} \cdot B_0 \, . \tag{2.2}$$

To increase M_0, which is proportional to the signal strength, one can either decrease the temperature T, use nuclei with larger γ, or increase B_0 (the external magnetic field).

γ is fixed for a specific nucleus. The gyromagnetic ratio of hydrogen atoms equals $42.577\,\mathrm{MHz\,T^{-1}}$ [20]. Fortunately, hydrogen both possesses a greater γ compared to other nuclei with non-zero spin, and very high in vivo concentrations. Therefore, the majority of in vivo MRI and MRS experiments detect ${}^1\mathrm{H}$, though other spin-active nuclei e.g., carbon-13, oxygen-17 [21] and fluorine-19 are of interest, as well.

B_0 is limited by the hardware, with typical clinical routines operating at $B_0 = 1.5$ or $3\,\mathrm{T}$. Dedicated research sites may possess $7\,\mathrm{T}$ systems, with some pioneers operating at $10.5\,\mathrm{T}$. Apart from the primary increase of signal intensity, higher field strengths convey additional benefits for spectroscopy (see Section *2.1.3 Chemical Shift*).

2.1.2 Nuclear Induction

MR signals are measured via Faraday induction, as the precessing magnetic moment creates an alterating radio frequency (RF) field at the Larmor frequency. However, the signal induced by the equilibrium state is negligible, as its magnetic moment is parallel to the B_0 field. By using RF pulses at Larmor frequency, the orientation of the net precessing magnetic moment can be changed. Rotating the magnetic moment away from that parallel state is called excitation. Its amplitude is commonly referred to as flip angle, it is described by the following formula, both for degrees (°) and radian (rad):

$$\alpha = \gamma \int_0^T B_1(t')\mathrm{d}t' \, . \tag{2.3}$$

Where $B_1(t)$ is the (time-dependent) RF field, applied in the time interval [0,T]; γ the gyromagnetic ratio in $\mathrm{MHz\,T^{-1}}$. The maximum flip angle is at 90°, the orthogonal state.

As this field, emitted at Larmor Frequency, is detected, it causes the MR signal. The net magnetic moment will decay over time back to equilibrium state, governed by two (tissue dependent) relaxation times. It can be described by the Bloch Equations, formulated in this context, with $\vec{M}$ the vector of net magnetisation:

$$\frac{\mathrm{d}M_i(t)}{\mathrm{d}t} = \gamma \left(\vec{M} \times \vec{B} \right)_i ; \quad i \in \{x, y, z\} . \tag{2.4}$$

The respective solutions for the decay of a signal are [22]:

$$M_x(t) = M_0 e^{-t/T2} \sin(\omega t) \tag{2.5}$$

$$M_y(t) = M_0 e^{-t/T2} \cos(\omega t) \tag{2.6}$$

$$M_z(t) = M_0(1 - e^{-t/T1}) \tag{2.7}$$

with $\omega = \gamma B_0$. In z direction, the signal is dampened and returned to equilibrium state by T1, the *longitudinal relaxation time*. Its source is the interaction of the spin with its magnetic surroundings, transferring energy, giving it the name *spin-lattice relaxation*.

After initial excitation, which imparts a coherence to the individual spins in x and y direction, the magnetisation M_x and M_y will return to its initial zero-coherence state, governed by T2, the *transverse relaxation time*. T2 relies on various effects, one of it the spin-spin interaction, giving it the title *spin-spin relaxation*.

2.1.3 Chemical Shift

Different metabolites are involved in various neurochemistry processes of signal transmissions between neurons. Each neurotransmitter molecule possess a unique chemical structures. The different hydrogen nuclei in each metabolite experience different chemical and electronic bonding environments. The MRS signal arises

from the precession of a proton's magnetic moment, which occurs at a frequency depending on the external B_0 field. This magnetic field at the nucleus can be shielded by the surrounding electrons of the molecular orbitals. Those are assumed to form an electron cloud, consisting of classical particles moving around the nucleus. The circulation of these charged particles induces a field B_e, opposing the external magnetic field B_0. Elements with higher electronegativity pull the electrons from the hydrogen atom. With weaker shielding from electrons, B_e decreases, therefore increasing B_n, the total magnetic field acting on the nucleus. As described in Equation 2.1, this results in a higher Larmor frequency. For MRS, these small shifts in the signal frequency domain give rise to characteristic resonances in the spectrum. As those measured Larmor frequencies depend on the external B_0, a relative measure δ is used. Through this reference, chemical shifts are expressed in parts per million (ppm):

$$\delta = \frac{\eta - \eta_{\text{ref}}}{\eta_{\text{ref}}} \cdot 10^6 \tag{2.8}$$

where η_{ref} is the Larmor frequency of a defined reference molecule, such as tetramethylsilane (TMS).

This chemical shift has not only been used to distinguish molecules, but also to predict accurate properties, such as folding states of known proteins.[23]

Since the Larmor frequency increases linearly with B_0, higher field strengths increase the absolute shifts, that stem from the proton chemical environments. This, next to the signal increase described above, allows for better resolved spectroscopic measurements at higher B_0 field strengths.

2.1.4 Spectral Editing

Spectral editing is most commonly referring to techniques, that discriminate 'scalar-coupled' from 'uncoupled' spins. J-coupling (also 'scalar' or 'spin-spin' coupling) is caused by nuclei with magnetic moments, influencing each other through local electrons. Beneficially, the effect carries information about the molecule's bonds, with its effect size also being independent of the field strength. However, it splits

the resonance peaks, creating more complex spectra. [14, p.30 ff.] J-coupling editing can use prior knowledge about the individual spectra, to separate signals of low concentration metabolites, from overlapping stronger signals.[14, p.375 ff.]

However, this thesis inspects non-edited MRS, as does the in vivo data set it relies on (see Section 3.1.2), which is the most commonly applied type of MRS. The software developed in this thesis is applicable to edited MRS. Specifically MEGA-edited MRS edited with MEGA (*MEscher-GArwood* [24]) for understanding the uncertainty of fitting the GABA-plus-macromolecule peak.

2.1.5   Noise in MRS

Although physiological noise can cause errors, the main source for noise in MR is thermal, stemming from the probe (i.e., the patient) itself. The input noise, measured in standard deviations (noise-sd), like the signal-to-noise ratio (SNR), quantifies the measurement accuracy of the in vivo setup tried to simulate.

2.2   Signal Model

This Section describes the signal model used for fitting spectra. Fitting is used to extract meaningful concentration estimates from the MRS signal, and to distinguish neurotransmitter molecules. To visualise the problem, Figure 2.2 shows how the measured MRS spectrum is composed.

To analyse the complicated signal, forward fitting is used. It is one way to perform stepwise regression. To optimise their flow, forward fitting algorithms utilise the fact, that some input components only have minor influence on the fitting. While running, the current predictive potential of all components is estimated and compared. Therefore, the components can be prioritised to find the highest possible improvement on the current predictive potential of the fit in progress. This continues stepwise, until all components aspects are calculated to improve the fit less than a set threshold. [25, p.446 f.]
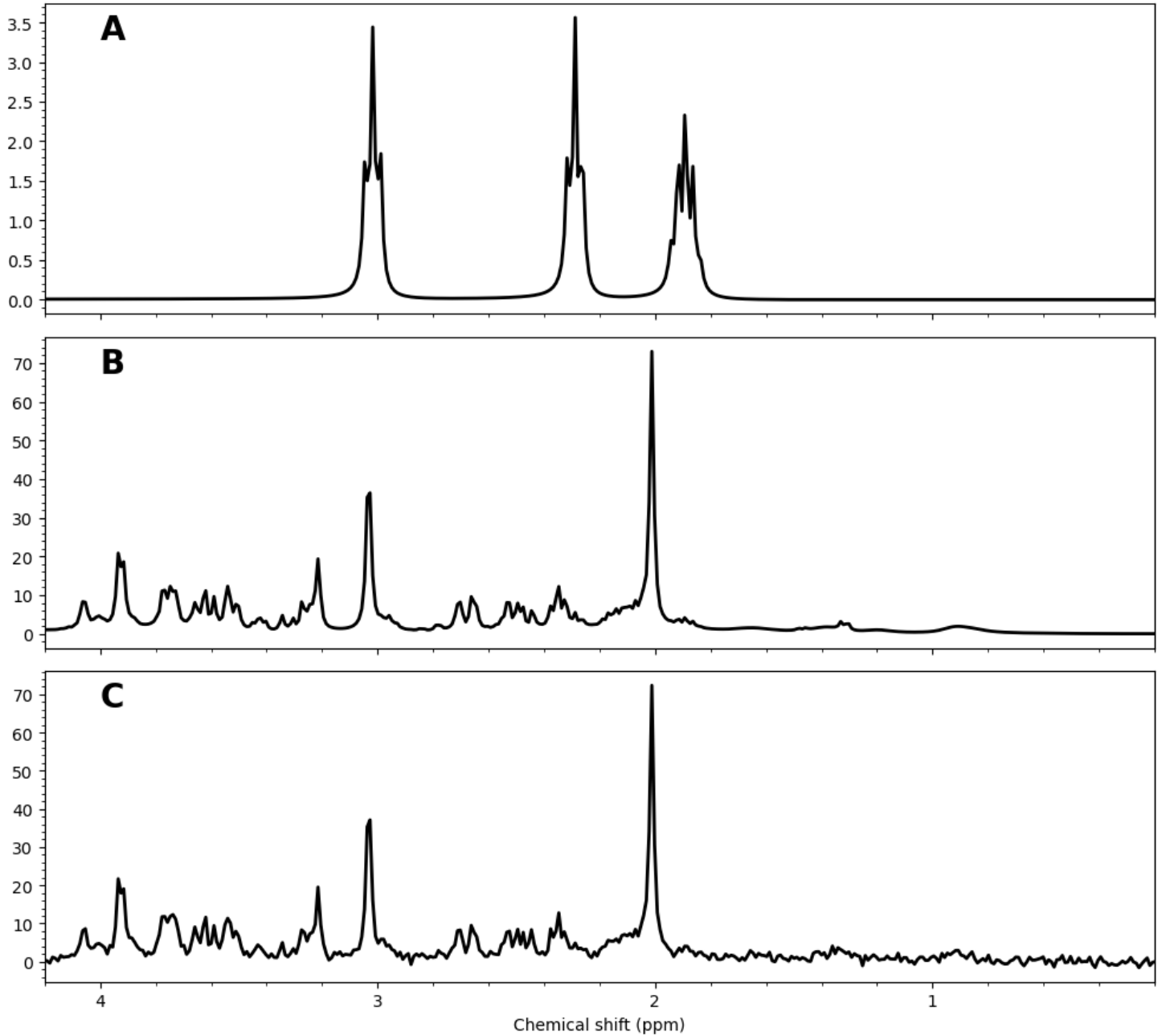
Figure 2.2: The stepwise composition of a MRS spectrum. (The reference data described in Section 3.1.2 is used.) (**A**) All metabolites send a unique signal, in this case GABA is depicted. (All other spectra are shown in Figure 3.2.) (**B**) The combination of all metabolite spectra, scaled by their respective concentration. As peaks of different spectra overlap, metabolites of low concentration like GABA become indistinguishable with the naked eye. (**C**) With a noise-SD of 1, this spectrum has considerable noise. Even though single peaks are recognisable, the fitting process for all 20 metabolites is complicated.

2.2.1   Basis

When fitting MRS signals, it is assumed, that each metabolite contributes a signal, that is scaled, broadened, and added to the others. Therefore, it is possible to subtract it lossless – the fitting process breaks down to finding the right proportions of containing spectra. This is the reason, why it is important to know the exact spectra forms of all contributing metabolites, called basis.

The basis is a set of free induction decay [signals] (FIDs), that each represent the ground truth of measuring a singulated metabolite – aiming to be free of noise and artefacts. They can be acquired measuring phantoms, but are usually simulated. [26] Section 3.1.2 describes the basis set used in this thesis. If peaks of metabolites overlap, their spectra and hence their fitting can be correlated. With selected metabolites, correlation is inspected in Subsection 3.2.1.

2.2.2   Fitting with FSL-MRS

This work uses the `FSL-MRS` library (version 1.1.0) [27], to both simulate and fit MRS signals. The library is a collection of python modules and wrapper scripts, for pre-processing and model fitting of MRS data. It is part of the FMRIB software library (`FSL`), released by the Wellcome Centre for Integrative Neuroimaging (Oxford, UK). `FSL-MRS` is part of the open-source FSL [28] software library, which in general provides a full range of analysis tools for MR, such as fMRI and diffusion brain imaging data.

Using FSL-MRS, this thesis performs forward fitting through linear combination of basis spectra, while optimizing with Bayesian statistics. This thesis uses the standard model of FSL-MRS [27, p.6 ff.]:

$$Y(\nu) = \mathrm{B}(\nu) + \underbrace{\exp\left[i\left(\phi_0 + \nu\phi_1\right)\right]}_{\alpha} \sum_{g=1}^{N_G} \sum_{l=1}^{N_g} C_{l,g} M_{l,g}\left(\nu; \gamma_g, \sigma_g, \epsilon_g\right) \qquad (2.9)$$

with

$$M_{l,g}\left(\nu; \gamma_g, \sigma_g, \epsilon_g\right) = \mathcal{F}\left\{m_{l,g}\left(t\right)\exp\left[-\left(\left(\gamma_g + \sigma_g^2 t\right) + i\epsilon_g\right)\right]\right\}. \qquad (2.10)$$

Here, ν is the frequency; B(ν) is an Nth-order complex polynomial to estimate the baseline. α is a term for global zeroth and first-order phase. The final sum describes the scaled, shifted, and broadened metabolite basis spectra $M_{l,g}\left(\nu; \gamma_g, \sigma_g, \epsilon_g\right)$. Thus $C_{l,g}$ are the concentration parameters, which are the target parameters of this study. To avoid overfitting, the metabolite line shapes can only be altered by shifting (ϵ) and broadening (γ, σ). Both can be flexibly applied on the metabolites, which are divided into the groups N_G, with the respective number N_g of metabolites in the group. $\mathcal{F}$ is the Fourier transform, and $m_{l,g}(t)$ is the inverse Fourier transform of $M_{l,g}\left(\nu; 0, 0, 0\right)$. The used polynomial baseline is second-order.

2.3   Monte Carlo Simulation

The Monte Carlo methods constitute a class of randomisation numerical simulations (which 'cast the dice'). This thesis will focus on its capability in generating pseudo-random data, which makes it invaluable for fields dealing with coupled degrees of freedom (e.g., physics of fluids or cellular solids).

This Section 2.3 is oriented at a textbook chapter *(15.6) Confidence Limits on Estimated Model Parameters.* [29, p.807 f.]

The theoretical problem is relying on a vector of underlying ground truth data $\mathbf{a}_{true}$, which is not known to the experimenter. Each measurement observes a new $\mathbf{a}_{(i)}$, deviated from $\mathbf{a}_{true}$ by noise. Below $\mathbf{a}_0$ is assumed to be an authentic measurement fit, whereas $\mathbf{a}_{(i)} = \{\mathbf{a}_{(1)}, \mathbf{a}_{(2)}, \dots\}$, are the results of repeated simulation.

Core concept of using Monte Carlo Simulations to fit $\mathbf{a}_{\text{true}}$, is to assume, that:

$$\mathbf{a}_{\text{true}} \approx \mathbf{a}_0 \qquad (2.11)$$

and the data points $\mathbf{a}_{(i)}$ are distributed in the same form around $\mathbf{a}_{true}$, as they are around $\mathbf{a}_0$. To meet this demand, foremost the distribution of values has to be assessed correctly, e.g., for many measurements as a standard normal deviation. The synthetic $\mathbf{a}_{(i)}$ will now amend formula (2.11), causing only a small offset to remain. Consequently, measuring $\mathbf{a}_{true}$, boils down to knowledge about $\mathbf{a}_{(i)}$ distribution, and its simulation.

As implementations may vary by demand, the core concept is drafted here:

1. Define scope of possible parameter variation in $\mathbf{a}$, based on $\mathbf{a_0}$.

2. Generate statistical data $\{\mathbf{a}_{(1)}, \mathbf{a}_{(2)}, \dots\}$, with desired probability density function.

3. Evaluate data points through deterministic criteria, retrieve result metric from the distribution of all $\mathbf{a}_{(i)} - \mathbf{a_0}$.

Its application is described in Section 3.1.

2.4 Power Calculations

Hypothesis testing is broadly used in research, to compare generated results with the investigated hypothesis (and therefore initial expectations). Whenever beneficial, known boundary conditions are used for more precise comparison. e.g., statistical distributions (such as normal distributions) can be inspected, when evaluating statistical models of observed data.

This section 2.4 cites the statistics textbook by Bortz. [25, p.122, p.133 ff.] In general, a quantity of experiment soundness is β: the probability, that the *null hypothesis H_0* (no assumption can be made) is right. In general research, it is the prime directive to try to disprove ones own theory.[4] However, the more practical approach is to describe the so-called *Power* $(1 - \beta)$. It is the probability of the observation confirming the initial hypothesis H_1 – i.e., the measurement results being close enough to the expectation to confirm them.

2.4.1 Student t-Test

The most popular *Student's t-Test*[5] compares two data samples for a power calculation, hence evaluating the validity of the initial hypothesis. The quality of

---

[4]The necessity of falsifiability in research due to epistemological demand is well established in present scientific routines. Many introductions can be found, e.g., with Philosopher Popper. [30]

[5]As William Gosset developed the t-test for the quality assurance of his employer Guinness' brewing processes, he had to use the pen name *Student*, to publish the now world known method.[31]

numerical implementations is examined well due to high demand, and its reliability is not commented here.

From two sample groups n_1 and n_2, the individual means μ_1 and μ_2 are taken. The null hypothesis H_0 is defined as:

$$H_0: \quad \mu_1 - \mu_2 = 0 \tag{2.12}$$

$$H_1: \quad \mu_1 - \mu_2 \neq 0 . \tag{2.13}$$

When repeatedly drawing two independent sets from two populations, a distribution of difference between those set-medians is generated. The null hypothesis (2.12) implies, that it should equal zero. More precisely, the mean values $\bar{x}_1$, $\bar{x}_2$ are realisations of the independent aleatory variables $\bar{X}_1$, $\bar{X}_2$. Here, $\bar{X}_1 - \bar{X}_2$ is the linear combination of the two separate numbers, their weighting being $+1/-1$.

The variance of this linear combination is known, resulting to:

$$\sigma^2_{(\bar{X}_1 - \bar{X}_2)} = (+1)^2 \, \sigma^2_{\bar{X}_1} + (-1)^2 \, \sigma^2_{\bar{X}_2} = \sigma^2_{\bar{X}_1} + \sigma^2_{\bar{X}_2} . \tag{2.14}$$

Additionally, the relation between population and sample variance is known:

$$\sigma^2_{\bar{X}_k} = \frac{\sigma^2_k}{n_k} . \tag{2.15}$$

Using this leads to the simplified:

$$\sigma_{(\bar{X}_1 - \bar{X}_2)} = \sqrt{\frac{\sigma^2_1}{n_1} + \frac{\sigma^2_2}{n_2}} . \tag{2.16}$$

Assuming both population variances are identical ($\sigma^2_1 = \sigma^2_2 = \sigma^2$) results in:

$$\sigma_{(\bar{X}_1 - \bar{X}_2)} = \sqrt{\sigma^2 \cdot \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} = \sigma \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} . \tag{2.17}$$

The shared population variance σ^2 can be estimated by using both samples:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n_1} (x_{i1} - \bar{x}_1)^2 + \sum_{i=1}^{n_1} (x_{i2} - \bar{x}_2)^2}{(n_1 - 1) + (n_2 - 1)} . \tag{2.18}$$

Resulting in a gauged standard error:

$$\hat{\sigma}_{(\bar{X}_1 - \bar{X}_2)} = \sqrt{\frac{\sum_{i=1}^{n_1} (x_{i1} - \bar{x}_1)^2 + \sum_{i=1}^{n_1} (x_{i2} - \bar{x}_2)^2}{(n_1 - 1) + (n_2 - 1)}} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} . \tag{2.19}$$

Or, if $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$ are pre-calculated:

$$\hat{\sigma}_{\left(\bar{X}_1-\bar{X}_2\right)} = \sqrt{\frac{(n_1-1)\cdot\hat{\sigma}_1^2 + (n_2-1)\cdot\hat{\sigma}_2^2}{(n_1-1)+(n_2-1)}}\cdot\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \; . \qquad (2.20)$$

Due to the *Central Limit Theorem*, $\bar{X}_1$ and $\bar{X}_2$ are normally distributed for $n_1+n_2 \geq 50$, causing $\bar{X}_1 - \bar{X}_2$ to be normally distributed, as well.[6]

To put $\bar{x}_1 - \bar{x}_2$ in context of $\mu_1 - \mu_2$, their relation to $\hat{\sigma}_{\left(\bar{X}_1-\bar{X}_2\right)}$ is being used:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\hat{\sigma}_{\left(\bar{X}_1-\bar{X}_2\right)}} \; . \qquad (2.21)$$

With the initial simplified H_0 assumption from formula 2.12:

$$\xrightarrow{\mu_1=\mu_2} t = \frac{\left(\bar{X}_1 - \bar{X}_2\right)}{\hat{\sigma}_{\left(\bar{X}_1-\bar{X}_2\right)}} \; . \qquad (2.22)$$

t now describes a normally distributed variable (for $n_1+n_2 \geq 50$). The t-distribution, however, helps to assess the data distribution for smaller sizes $n_1 + n_2 \lneq 50$.

The following limitations apply:

1. All subsets are assumed to be standard normal distributed.

2. The variance of both compared subsets have to be equal, although further mathematical research exists for compensation of alterations.

3. Both compared subsets have to be independent of each other. 'Dependence' or 'overlap' would originate, e.g., if the measurements for both sets are performed on the identical test subject (impartial if person or object).

For many violations of these assumptions, correction methods can be found. In literature discourse, t-test's behaviour is deemed resistant against model violations.[7] Still, by combining e.g., different sample sizes (e.g., $n_1 \ll n_2$), and diverging variances($\sigma_1 \neq \sigma_2$), results can be less reliable.

---

[6]For lower sizes, a t-curve distribution with impeded degrees of freedom $(n_1 + n_2 - 2)$ applies.
[7]Notably, this research on perturbation theory is influentially using the Monte Carlo Method, similar to this thesis application.

3. Materials and Methods

To estimate the sensitivity of MRS experiments, existing software for MRS spectra simulation was expanded to incorporate Monte-Carlo simulations, and power testing. The conceptual design of this software is presented first. It is extended by an explanation of the software flow for users. Finally, two types of experiments are described: One validating the generated data, the second answering the initial question of MRS sensitivity, for GABA.

3.1 Simulation Software

The new software estimates the sensitivity of in vivo MRS measurements. It uses the pre-existing MRS utility library `FSL-MRS`[1], to first simulate MRS spectra, and then analyse them. The difference between simulation input (in vivo reference) and analysed output (in silico data) is the measurement error. Using a Monte-Carlo method, a distribution of measurements is created, to estimate its mean and spread (i.e., uncertainty):

The MRS spectra of one MRS setup are based on a (unknown) ground truth $\mathbf{a}_{true}$. This set of characteristic parameters includes metabolite concentrations and signal nuisances (noise-SD, linewidth, shifting, phase, baseline). Three steps are performed, to retrieve the uncertainty of a setup:

1. A real measurement $\mathbf{a_0}$ (ignoring its noise) is defined as ground truth $\mathbf{a}_{true}$. The ideal spectrum therefore is $S(\mathbf{a_0})$ (without noise).

---

[1]See *Section 2.2.2 Fitting with FSL-MRS.*

2. n new spectra are simulated, by adding random noise N_i (its noise level being that of $\mathbf{a_0}$), to the ideal spectrum:

$$S(\mathbf{a}_{(i)}) = S(\mathbf{a_0}) + N_i ; \quad i \in \{1, 2, \ldots, n\} . \tag{3.1}$$

3. Analyse each simulated spectrum $S(\mathbf{a}_{(i)})$, to determine all n sets of parameters $\mathbf{a}_{(i)}$. Obtain the spread of all $\mathbf{a}_{(i)}$ around $\mathbf{a_0}$.

Expecting normal distribution of measurement result parameters, their mean and standard deviation can be determined. To reduce fit errors, a sufficient amount of data points (i.e., runs) is needed, as described in Section 3.1.1.

3.1.1 Workflow

The project code was written in python; its source code is available online, with exemplary scripts. [32] Configurable parameters are all metabolite concentrations and signal nuisances (noise-SD, line width, shifting, phase, baseline). To better inspect concentration changes, relative concentration units are available. Many default options are provided, e.g., basis set, and in vivo concentrations (as described below).

To simulate MRS, spectra are calculated[2] from the metabolite basis spectra, combined with the input default parameter set containing: metabolite concentrations, line-broadening, peak-shifts, zero and first order phase, and baseline parameters. Furthermore, complex white Gaussian noise was added[3] to the frequency domain spectrum.[4] The variance of the normally distributed noise can be altered to adjust the noise level in the generated data. Afterwards, the synthetic spectrum is analysed with `FSL-MRS`[5] (see *Section 2.2.2 Fitting with FSL-MRS*), returning the measured concentrations.

To enable larger sets of simulations, the software interface designed does not require grafic interaction, enabling the application in e.g., for computation clusters.

---

[2]Using the `FSL-MRS` function `synthetic_from_basis`.

[3]Additive white gaussion noise (AWGN)

[4]`synthetic_from_basis` utilizes the `numpy random.multivariate_normal` function, on both dimensions of the complex signal.

[5]Using the `FSL-MRS` function `MRS.fit`.

In the `FSL-MRS` library, CPU parallelisation is already implemented. However, all simulations for this thesis were performed on a laptop[6], with runtimes taking well below 20 seconds per spectrum simulation and analysis.

Convergence Criteria

In iterative calculations, iterations are stopped, once a convergence criterium is fulfilled.

As the Monte-Carlo simulation is used to determine the result distribution, convergence is reached, if the generated data allows for reliable fitting of its data. To determine, that parts of the data are coherent and confirm the remaining data, their individual distribution form is compared to the general result.

Therefore, the simulation is deemed convergent, if both mean and standard-deviation of their result normal distribution, are within a certain threshold to each other. Here, arbitrary values were used: The simulation is convergent, if the distribution means are within 1%, and the standard deviations are within 10% offset.

To apply this, the simulation is performed in subsets, with an increasing number of repetitions. Each subset is compared to the result of all previous subsets combined. Again arbitrarily, a factor of two was chosen, to increase the number of runs, between subsets.

In applying those thresholds and repetitions, a first experiment was performed in *Section 3.2.1 Data Generation and Validation* to determine the repetitions needed.

3.1.2 Dataset

To perform a sensitivity analysis on spectra which are representative for in vivo human MRS, all simulations in this work used a reference dataset provided by Barron and Koolschijn (WIN, University of Oxford). [33] The study consists of 22 healthy volunteers (mean age 22.8 ± 0.74 years, 4/18 males/females)[7], where

---

[6]CPU: intel i5-7200U, 4 threads at 2.6 GHz

[7]No significant effect of sex on either behavioural performance or MRS measures was found by the researchers.
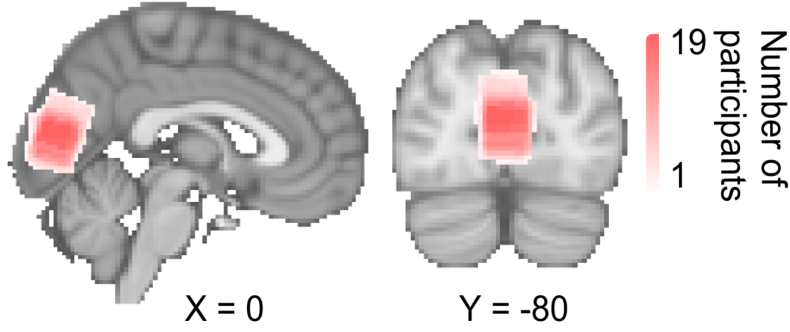
Figure 3.1: Position of the voxel ($2{\times}2{\times}2$ cm^3) for all patients of the presented data. [from 33]

19 produced valid data, as three had to be excluded[8].

The study was performed on a Siemens Magnetom 7T (Siemens Healthineers, Erlangen, Germany), using a 32-channel head coil (1Tx32Rx Head coil, Nova Medical Inc, WI, USA). The MRS data was acquired from one $2{\times}2{\times}2$ cm^3 voxel (see Figure 3.1), positioned in the primary visual cortex.

MRS data was acquired using a `semi-LASER` [34] sequence (TE = 36 ms, TR_{mrs} = 4 s) incorporating `VAPOR` water suppression and outer volume suppression. `semi-LASER` is a spin echo sequence, that uses adiabatic refocusing pulses to reduce chemical shift displacement errors, while being relatively insensitive to B1 inhomogeneity.

The `VAPOR` (*VAriable Pulse power and Optimized Relaxation delays*) method uses frequency-selective RF pulses, followed by 'crusher' (or 'spoiler') gradients. They selectively excite, and then dephase the water signal. One benefit of `VAPOR` is, that it can be used interleaved with another method used, the outer volume suppression. Here, spatially selective saturation pulses are placed tangentially around the skull prior to excitation. In doing so, the signals of those regions are suppressed, such as the signals of the subcutaneous fat. [35]

The data provided was pre-processed in `MRSpa` (CMRR, University of Minnesota). Specifically corrected for eddy currents, as well as frequency and phase alignment,

---

[8]One participant lacked combined fMRI-fMRS data due to time constraints during scanning. Two participants were excluded from the fMRI and fMRS analyses due to technical difficulties, resulting in the auditory cues not being fully audible during the inference test. [33]

alanine	(Ala)	glutamate	(Glu)
aspartate	(Asp)	glutamatetathione	(GSH)
ascorbate/vitamin C	(Asc)	myo-inositol	(myo-Ins)
γ-Aminobutyric acid	(GABA)	lactate	(Lac)
glycerophosphocholine	(GPC)	N-acetylaspartate	(NAA)
phosphocholine	(PCho)	N-acetylaspartylglutamate	(NAAG)
creatine	(Cr)	phosphoethanolamine	(PE)
phosphocreatine	(PCr)	scyllo-inositol	(scyllo-Ins)
glucose	(Glc)	taurine	(Tau)
glutamine	(Gln)		

Table 3.1: 19 metabolites represented in the basis set.

any residual water was removed using `HLSVD` (*Hankel Lanczos Singular Value Decomposition* [36]).

The simulation parameters of this thesis were then generated by averaging every parameter across the participants.

The set of basis spectra was provided by Clarke (WIN, University of Oxford). For its simulation, the `VESPA` (*Versatile Simulation, Pulses, and Analysis* [37]) package was applied, that uses the `pyGAMMA` library (Python wrapper of the `C++` *General Approach to Magnetic resonance Mathematical Analysis* [38]). For the contained stimulated spectra to meet the criteria of the in vivo data set, their number of points and bandwidth were downsampled to a value of 16384 points and 12000 Hz. Table 3.1 lists all metabolites, for which spectra were simulated. Additionally, macromolecules (Mac) were measured experimentally, by using an additional inversion pulse to measure metabolite nulled spectra in an equivalent subject cohort [33]. Both simulated and measured spectra forms are shown individually in Figure 3.2. By linear combination of the spectra, the MRS signal can be composed, as shown in Figure 2.2.
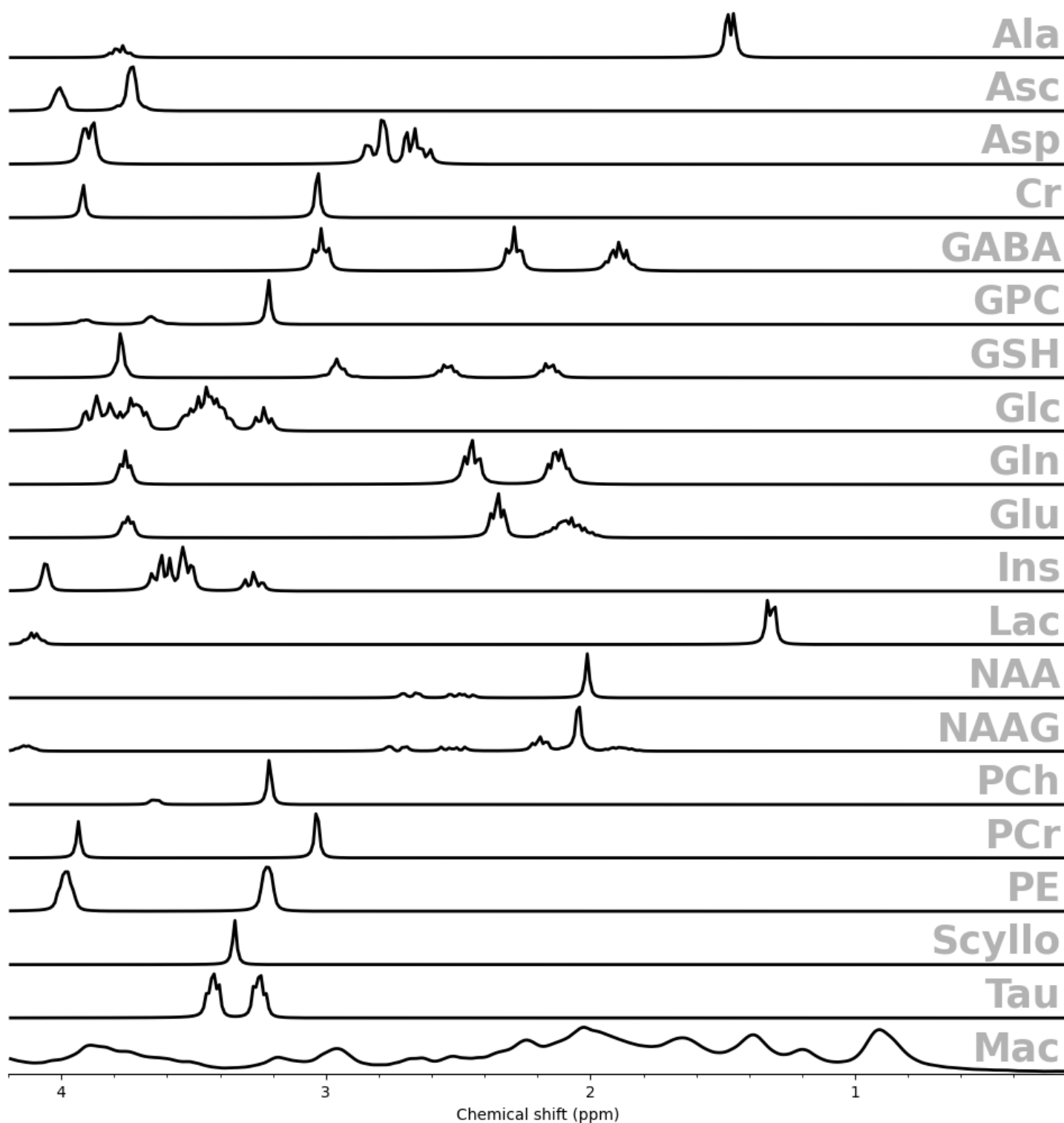
Figure 3.2: Spectra of the individual metabolites as represented within the basis set. To depict all peaks, the spectra do not share the same scale of magnitude.

3.2 Experiments

3.2.1 Data Generation and Validation

The Monte-Carlo simulations described above are the basis of all experiments. Based on the dataset described in Subsection 3.1.2, the simulated setup was varied. As practical approach, the simulations were not performed on demand, but as one set a priori, before inspecting it through the experiments below. Spectra were varied in quality, by changing the noise level, and the concentration of one metabolite of interest while all other parameters were set to in vivo values. More specifically, data was generated for 252 setups:

1. 4 metabolites of interest: NAA, GABA, Cr, glutamate.

2. 9 concentration steps relative to the in vivo concentration: 0.05, 0.2, 0.4, 0.6, 1, 1.5, 1.8, 2, 2.5.

3. 7 noise-SDs (σ of the noise baseline[9]): 0.01, 0.1, 0.5, 1, 2, 10, 25.

For this data, three validation tests were performed:

I. **The Convergence of Monte-Carlo Results**
   The convergence of the simulation (see Section 3.1.1 Convergence Criteria. Was inspected, to ensure adequate data quantities, across the different MRS configurations.

II. **Characteristics of Monte-Carlo Result Distributions**
   The characteristics of the result distributions are demonstrated. As described in Section 3.2.2, they are later used to interpolate data.

III. **Concentration Correlation of Metabolites with Similar Spectra**
   To assure accurate differentiation, metabolites with similar spectra are measured in comparison. Correlation of results would imply biased simulations.

---

[9]The sum of the unscaled basis sets were scaled to have a consistent norm.

3.2.2   Estimating Measurement Confidence

To determine the sensitivity of MRS measurements to concentration changes, the means and standard deviations were calculated. To choose multiple steps from a range, data interpolation is used, thereby circumventing additional Monte-Carlo simulations. Finally, t-tests determine the statistical power for detecting concentration changes. Using a statistical model, the required sample size (number of participants) to reach a certain power level can be calculated.

Data Interpolation

To calculate the concentration distribution of one specific measurement setup (target metabolite and concentration; and signal nuisances: noise-SD, linewidth, shifting, phase, baseline), the Monte-Carlo simulation can be used directly. However, using the full Monte-Carlo process for many points is highly resource intensive. Alternatively, interpolation can assess new setups – without simulating additional spectra.

To interpolate the distribution mean, the relative error had to be reviewed. As tested after the data generation, this coefficient depends linearly on the chosen noise. Hence, the correct coefficient is interpolated for the respective noise levels. Multiplying it with the simulation concentration input results in the expected result average.

As noise increases the measurement error, a direct connection between noise-SD and SD of the results can be made. For this simulation, a linear interpolation was performed, scaling both dimensions (noise-SD and result-SD) decade logarithmically.

With both mean and SD of the result distributions determined, t-tests are applied.

t-Test and Power Solving

t-Tests are performed to determine, whether two MRS setups are distinguishable (implying their difference is detectable). Representative sets of new data are created for both setups, by drawing random values[10] from the determined result normal

---

[10]Using the `numpy` function `random.normal`.

distributions. Different data sizes can be used, but both sets have to be equal in size, and generated independently. Comparing both sets, an independent t-test is performed.[11] For each performed pair, the statistical power (p-value) is calculated.

Conversely, single undefined parameters can be determined, using a statistical equation solver. It alters the unknown input parameter, with numerical optimisation, until the demand output is met.[12] In this thesis, it is used, to find the sample size needed, to reach a demand power level.

Glutamate Changes in fMRS

Both glutamate and GABA are key excitatory and inhibitory neurotransmitters. GABA is typically harder to measure and is as such exemplar for most of the work. However, glutamate is much better explored. A great interest lies in the development of functional-MRS (fMRS), which is an important growing area of MRS. While glutamate is easier to fit than GABA, fMRS is typically acquired faster, and has a lower SNR. Therefore, to compare study designs of fMRS, glutamate measurements are inspected.

---

[11]Assuming both sets have identical variances, it tests for the null hypothesis that both sets have identical average (expected) values. The function used is `scipy stats.ttest_ind`.

[12]The `statsmodels` function `TTestIndPower.solve_power` was used. It utilizes `scipy.optimize`, to find the value that satisfies the power equation.

4. Results

4.1 Simulation Software

The complete software for Monte-Carlo simulation and results generation can be obtained from the online repository. [32] The software's API is designed to be called within Python scripts or interactive Python environments. Preconfigured example setups are provided. Users are able to recreate a full range of MRS spectra. Measurement uncertainties of different neurotransmitters can be estimated, for single-subject diagnostics, and study cohorts.

4.2 Experiments

This section presents the simulation results. First, simulations are presented, validating the used simulation data. Afterwards, this information is used to estimate confidence levels for in vivo measurements.

4.2.1 Data Generation and Validation

As described in Section 3.2.1, synthetic data was generated for 252 different setups, varying concentrations and noise levels. The data integrity was validated with three tests. Additional graphs are in the appendix.

I. The Convergence of Monte-Carlo Results

Of all four metabolites (NAA, GABA, Cr, glutamate), Figure 4.1 shows the lowest signal metabolite, GABA. Monte-Carlo simulations for 3 different noise levels (noise-SD 0.08, 0.34 and 0.68) are inspected, where every simulation is depicted as series of subsets, that increase in their number of measurements. The amplitude is the mean

output deviation from the input concentration. The graph implies fast convergence, even below magnitudes of hundreds of measurements.

As the results of the three other metabolites were similar, the arbitrary choice was made, to simulate 200 measurements per measurement setup.
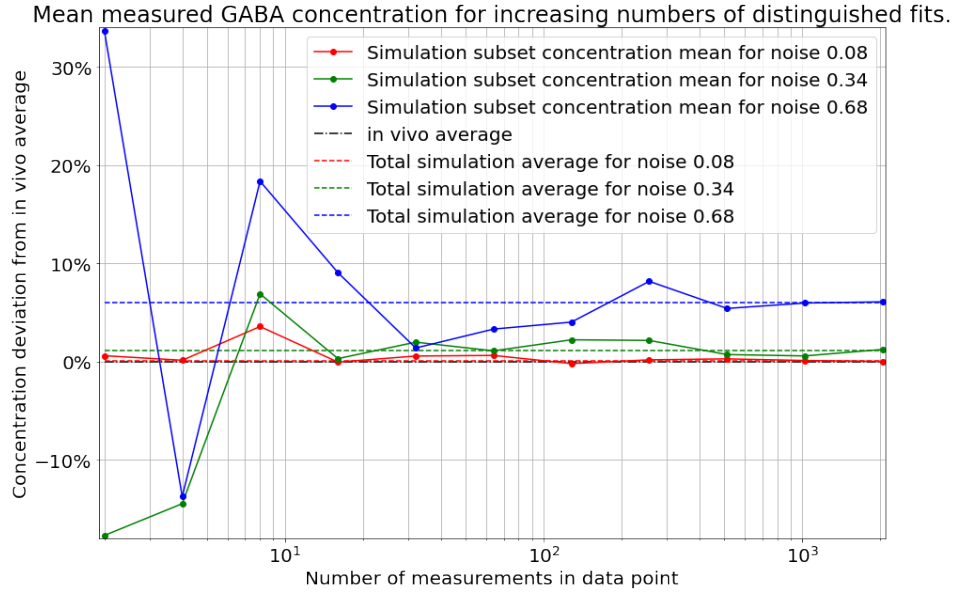
Figure 4.1: Rate of convergence at different noise levels. Doubling the number of measurements at every step, the mean concentration of each step is plotted (dots). Three noise-SD are tested, where 0.34 is the in vivo 7T data mean. For every simulation mean (dashed horizontal), the systemic difference is well below 7% concentration offset. (From [39])

II. Characteristics of Monte-Carlo Result Distributions

Exemplary correlations between fitted and input concentrations can be seen in Figure 4.2. On the left side, the correlation is linear, but deviates from the line of identity. As seen on the right, the slope deviation increases at higher noise levels. The graph confirms the expected linear correlation between fitted and input concentrations. Therefore, the first important result characteristic is the relative error, of the analysed concentrations. It is dependent on the noise level.

In Figure 4.3 the slope was subtracted from the data on the left. The spread of data around the linear fit is clearly visible, and it increases for higher noise levels.

Both the corrected mean (dependent on the relative error), and the spread, characterise the distribution of the Monte-Carlo results. This can be used to
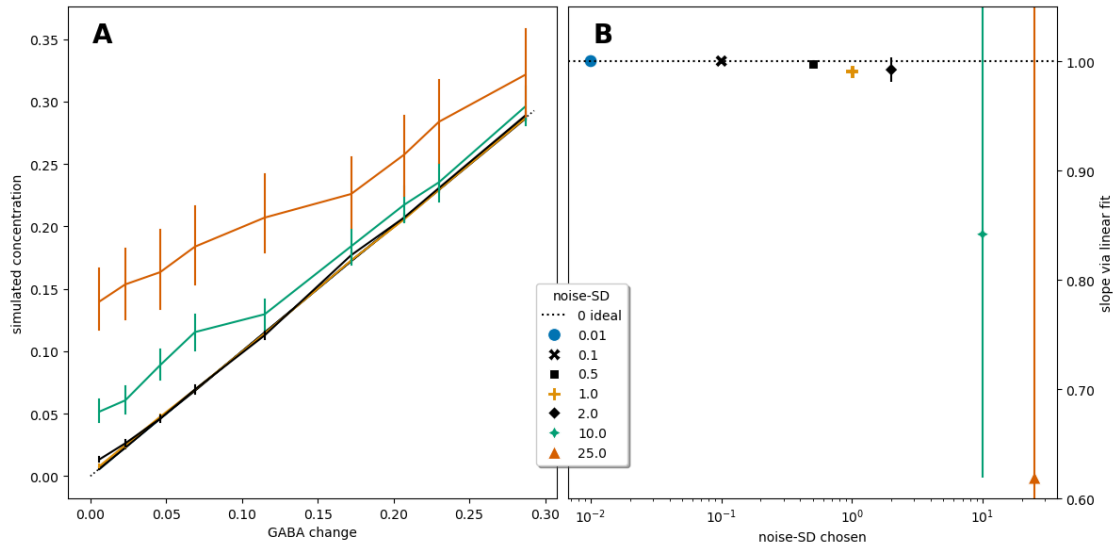
Figure 4.2: Simulation changes for different GABA concentrations. (**A**) Comparison of noises at different GABA input concentrations. (**B**) Slope of left side lines, assuming linear data.
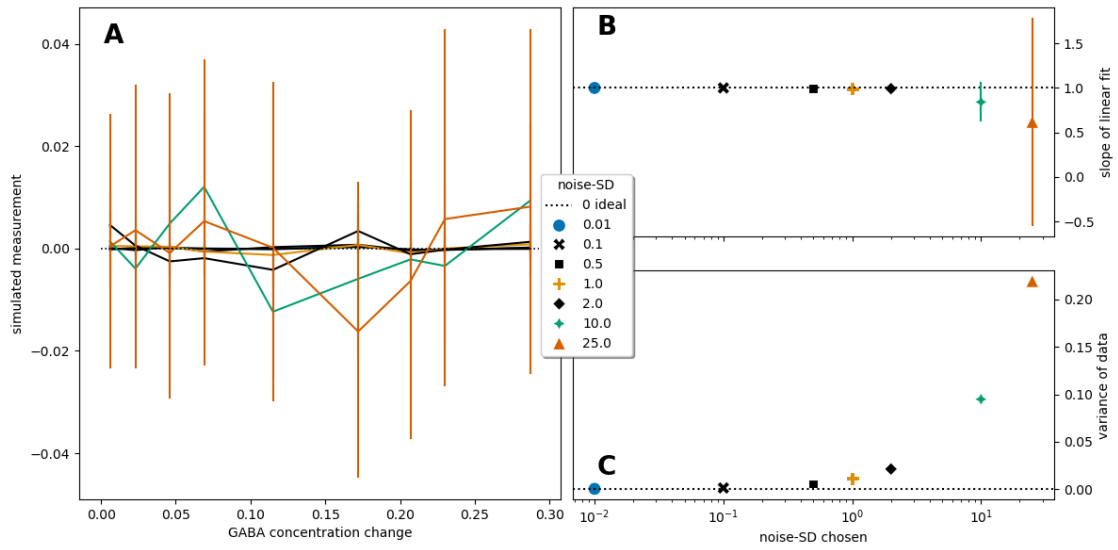
Figure 4.3: Simulation results for different GABA contexts. (**A**) Offset from linear fit. (**B**) Fit slope coefficient, compared to line of identity. (**C**) Variance of data around linear fit.

interpolate results, as mentioned in Section 3.2.2, and used for application below

in Section 4.2.2.

III. Concentration Correlation of Metabolites with Similar Spectra

To test, whether different metabolites can be differentiated, several metabolite pairs were simulated. Each pair contained one metabolite of interest, together with a second metabolite, that has an overlapping spectrum. In Figure 4.4, results are seen from the two strongly overlapping metabolites creatine (Cr) and phosphocreatine (PCr). To test separation, the fitting process of the metabolite of interest was inspected two times: First, as standalone analysis of its changing concentration. Then, the simulation is repeated, by measuring it, paired with its constant-concentration partner. For the combination of the metabolites (Cr+PCr) the graph is visibly shifted to higher values, as it is normalised to the ground truth creatine concentration (by the amount of added PCr). Still, the slope of both fitting processes are similar, so that the shift can be suppressed, if known. Notably, combining two metabolites (still as separate measurements) reduced the combined error, as noise is averaged with two independent values. Therefore, the fitted slope is closer to 1, when derived from two metabolite concentrations, instead of one. However, for the individual (but highly correlated) measurement, a clear bias is seen which underestimates the concentration. This bias is higher at low SNR where the peaks are less distinguishable due to noise contamination.

One other observation is, that as the Cr concentration approaches 0 (at 10% in vivo concentration), the simulation overestimates concentrations; even though for most configurations an underestimation is seen. The noise has a larger influence low values, shifting single measurements to negative values during simulation. As concentrations measure zero or more, the simulations are truncated (and set to) zero. Therefore, the result distribution is truncated on the near-zero side, resulting in an overestimation.

4.2.2 Estimating Measurement Confidence

Figure 4.5 depicts the confidence estimation for a range of setups. Changed parameters are the number of measurements, the concentration change, and the noise-SD. By averaging 1000 t-tests less fluctuant is seen.
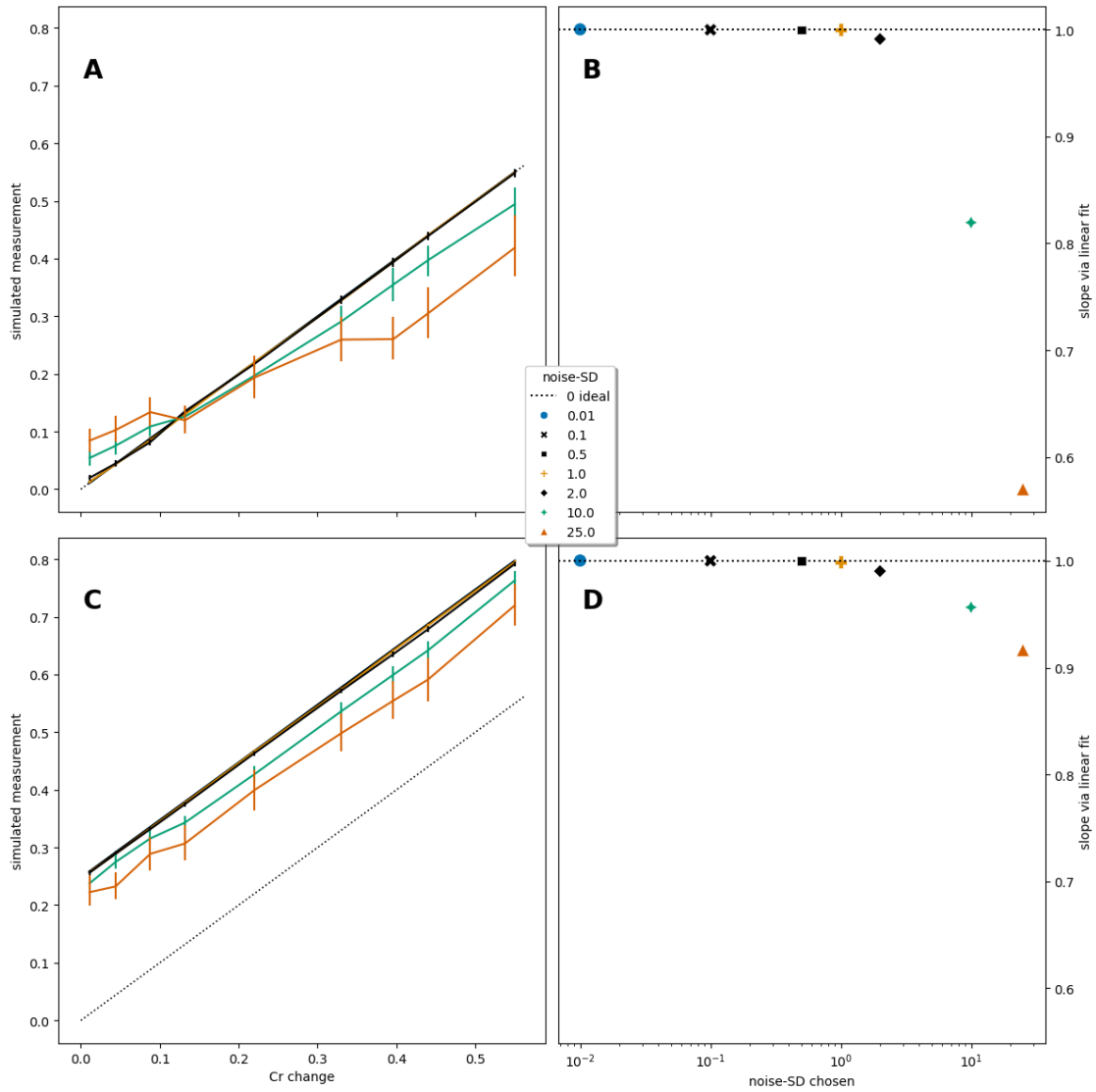
Figure 4.4: Changing Cr, comparing simulations of measuring single Cr, and Cr+PCr combined. (**A**) Fit of Cr concentration, with increasing Cr concentration input. (**B**) Slope of the simulation run, compared to the line of identity. (**C**),(**D**): Analogue for fitting Cr paired with PCr.

By using a statistical equation solver, the modelled data distribution was used, to retrieve the needed sample size, to detect a change. Figure 4.6 shows the same problem as 4.5, this time presenting the number of measurements needed to detect changes. Where the uncertainty was too high, the power calculation failed to converge, and an empty field is presented.
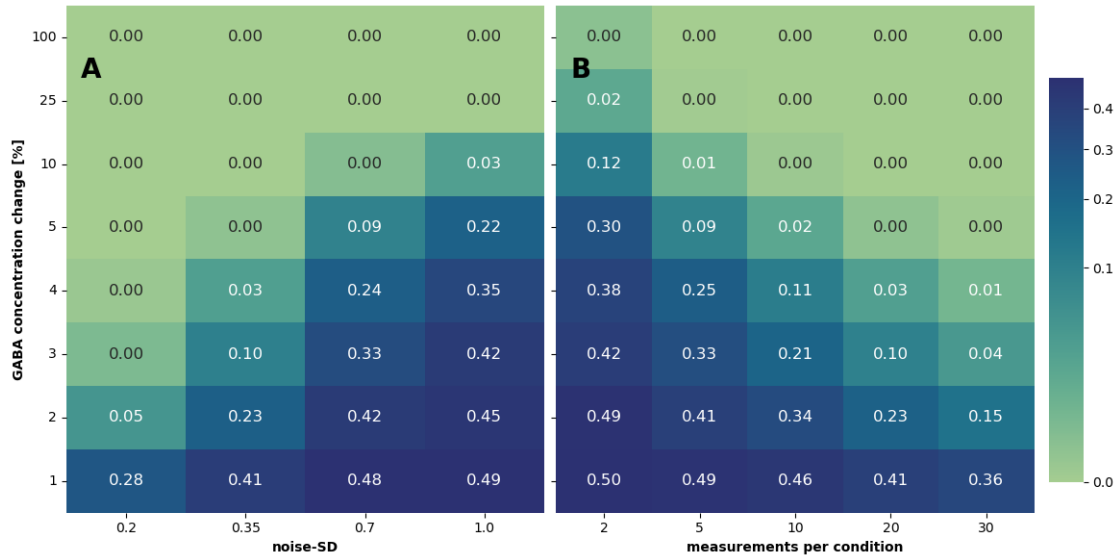
Figure 4.5: The probability of detecting GABA concentration changes is shown as p-value, for different setups. 1000 t-tests are averaged, to compare to a central input set (in vivo concentration, noise-SD 0.35). (**A**) Varying noise levels, constant number of measurements (20). (**B**) Varying numbers of measurement, constant noise-SD (0.35).

Glutamate Changes in fMRS

Finally, these techniques are applied, to design a study for detecting physiological changes in glutamate (Glu) concentrations. The in vivo glutamate concentration, as well as all other parameters, are assumed as given in the measurement setup. For comparison, the in vivo noise-SD is doubled and tripled. To determine the expected glutamate change, a meta review is used, comparing ^{1}H fMRS studies [40]. Different study designs are described, using different brain 'activation' paradigms, for each a different glutamate change is predicted. Table 4.1 shows the resulting study size requirement, as calculated by the thesis method.

To compare *event paradigm* to *block studies*, their size needed to detect the glutamate is taken. The upper bound of the confidence interval is used, optimising for a robust study result. The noise-SD chosen to be 0.7 as example (double the average of the simulation data set), as noise levels are higher for short acquisition times, as in fMRS. In this example, a *block study* needs 11 participants, where only 3 are needed for the *event paradigm* design.

In general, event paradigms might have higher noise levels. Still, in this
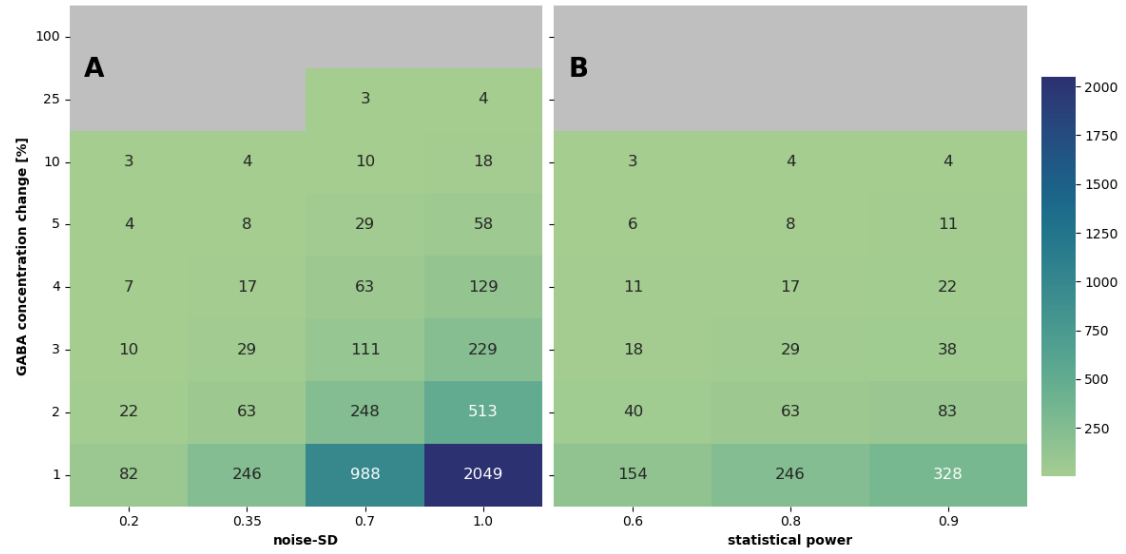
Figure 4.6: For different setups, the number of measurements needed to detect GABA concentration changes is shown. The change is compared to a reference setup (in vivo concentration, noise-SD 0.35), the significance level α is kept at constant 5%. To retrieve the minimum of needed measurements, the result is rounded to the next higher integer value. (**A**) Varying noise-SD, constant statistical power (0.8). (**B**) Varying statistical power, constant noise-SD (0.35).

simulation, if comparing the 3 needed subjects of the 0.7 noise-SD *event paradigm*, the *block study* still needs 4 subjects (one more), at the lowest 0.35 noise-SD.

Notably, *pain stimuli* have the highest predicted glutamate increase. However, the stimulation method can only be applied on some research.

Study Design	Glu increase predicted	noise-SD	measurements needed per condition	confidence interval $+\sigma$	$-\sigma$
neural activation	7.0% ($\pm$1.7%)	0.35	3	0	
		0.70	4	1	-1
		1	6	2	-2
block studies	4.7% ($\pm$1.5%)	0.35	3	1	0
		0.70	6	5	-2
		1	10	10	-3
event paradigms	13.4% ($\pm$3.6%)	0.35			
		0.70	3	0	
		1	3	1	0
visual stimuli	2.3% ($\pm$1.2%)	0.35	6	17	-2
		0.70	18	68	-8
		1	35	135	-18
pain stimuli	14.46% ($\pm$3.7%)	0.35			
		0.70			
		1	3	1	0

Table 4.1: Calculated cohort size requirements for a functional MRS study to detect predicted glutamate changes. Empty cells are settings with too high uncertainty, causing the power calculation convergence to fail.

5. Discussion and Conclusion

This thesis presents a software that, estimates uncertainties in MRS-measured metabolite concentrations. As example of application, the software was used to determine the required number of participants, for a study to detect (predicted physiological) glutamate changes at a certain significance and statistical power. Data from a 7 T MRI scanner and a `semi-LASER` sequence was used as a reference dataset to seed the simulations with realistic in vivo parameters. The results produced are credible, given they match established study design, though further discussion is useful.

As demonstrated in Section 4.2.2, with this tool, users can accurately predict the number of participants needed to measure a given variation in neurotransmitter concentration. They can explore multiple experimental eventualities, given the unknown effect size.

In case of fMRS study designs, the results heavily support the noticeable trend towards novelty event paradigm designs. The example emphasizes the benefits in comparing different tools. Still, before applying the method on fMRS, it has to be improved accordingly. Especially the signal model could be adapted (its spectra and time-dependent behaviour), to reassure the estimation reliability. However, those necessary changes are not expected to be tremendous, as the software is flexible in its setup simulation, as well as MRS and fMRS being similar in many aspects.

Simplifications were made for this thesis. The inspected parameters were noise-SD and concentration of the metabolite of interest. Meanwhile, additional signal nuisances, such as line width, shifts in peak frequencies, or baseline artefacts (caused by partially suppressed water) were ignored. As noise is the most relevant of those nuisance parameters, its presented variation should outweigh their influence;

verification would further the process. Additionally, the 7 T MRI measurement data was assumed to be artefact free, and homogenous in quality across participants. In a more realistic cohort, each subject's data has variable quality (noise and linewidth). This work doesn't simulate this effect, as the data parameter averages are used – in practice, this should be verified. In addition, it would help the future application precision, to consider physiological sources for noise and artefacts, despite the overall white noise of bigger magnitude.

Likewise, approximations were made for modelling the measurements. The relative simulation errors of MRS measurements were interpolated linearly (with logarithmic scaling) between the simulated data points. To increase accuracy, users could either perform full Monte-Carlo simulation for each data point of interest, or use a higher order interpolation.

Although convergence of the model was witnessed for every Monte-Carlo experiment, especially low SNR requires higher numbers of simulation iteration.

In future works, the software could be directed to novel, and more complex MRS types, namely diffusion MRS and fMRS applications. By extending the functionality for more complex spectra and combined spectra, not only additional techniques can be simulated, but their advantages compared, as well.

The `FSL-MRS` library is aimed to contain the new MRS functionalities of this thesis. As presented, those functions are easy to apply, yet provide valuable insight for research and diagnostics. Promising to solve many more adjacent problems, this software hopefully has wide applicability to users of MRS.

Appendices

Additional Result Graphs

II. Characteristics of Monte-Carlo Result Distributions
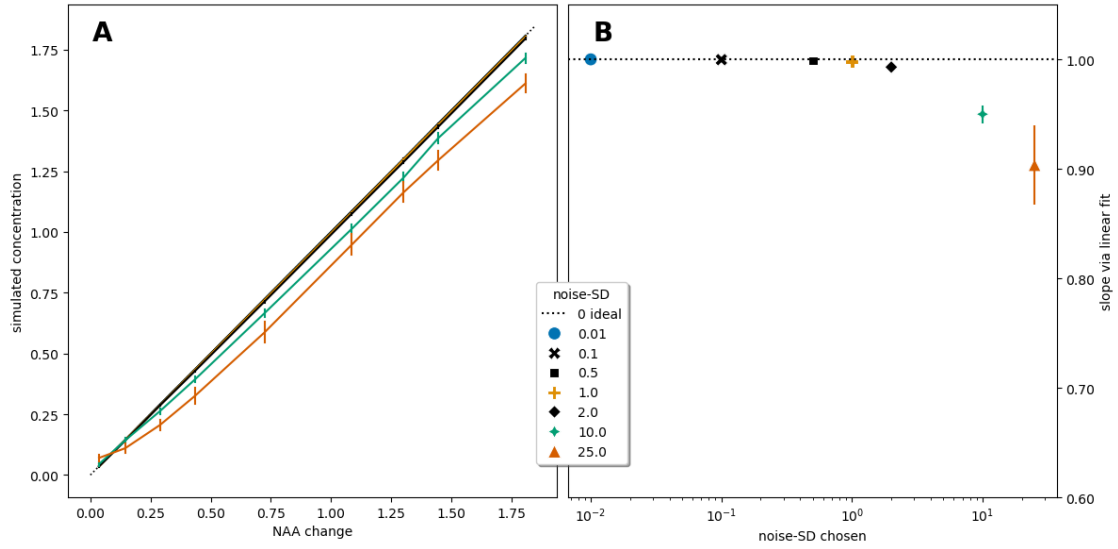
Figure 1: Simulation changes for different NAA concentrations. (**A**) Comparison of noises at different NAA input concentrations. (**B**) Slope of left side lines, assuming linear data.
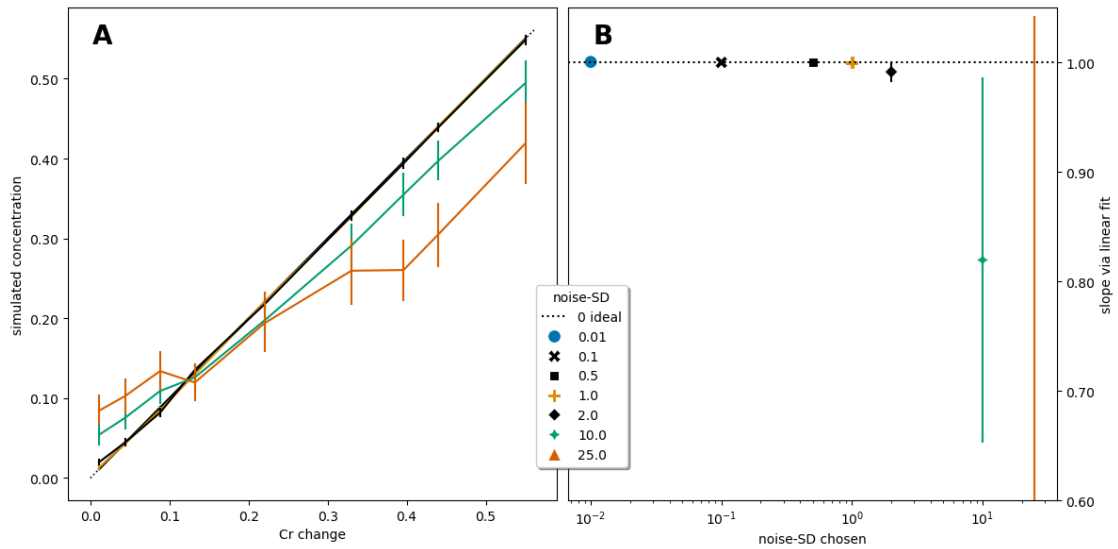
Figure 2: Simulation changes for different Cr concentrations. (**A**) Comparison of noises at different Cr input concentrations. (**B**) Slope of left side lines, assuming linear data.
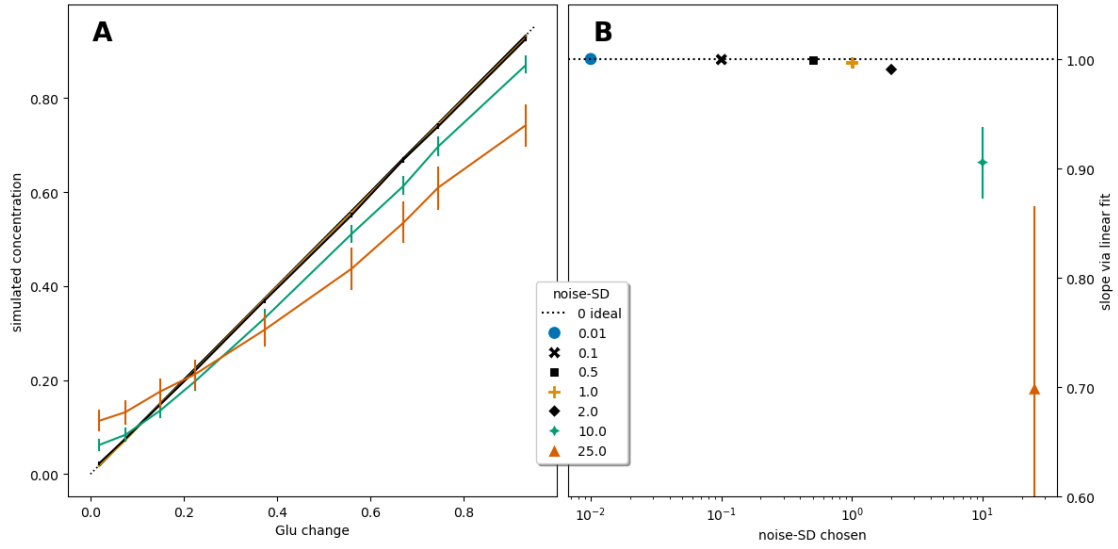
Figure 3: Simulation changes for different Glu concentrations. (**A**) Comparison of noises at different Glu input concentrations. (**B**) Slope of left side lines, assuming linear data.
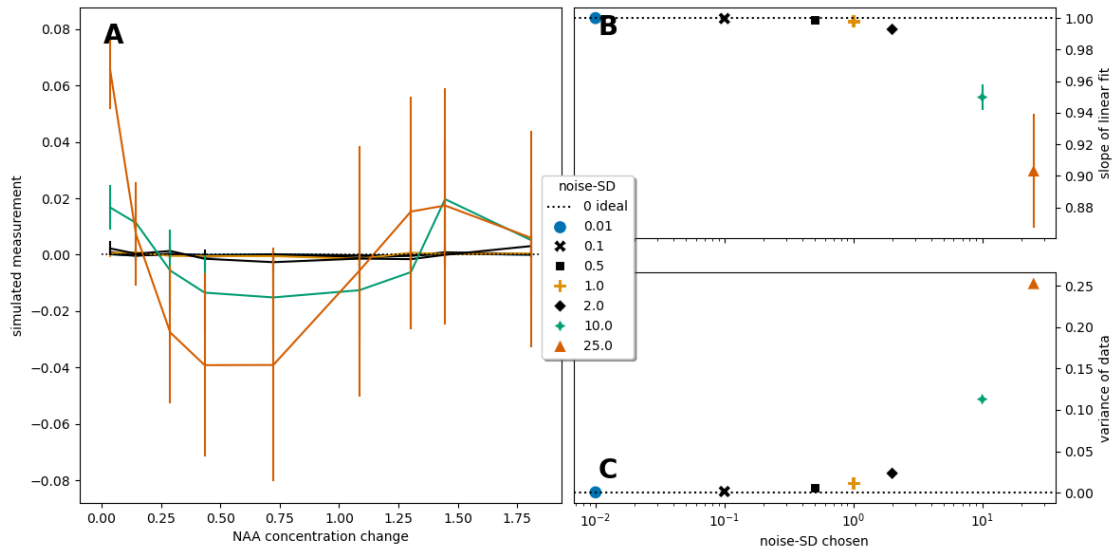
Figure 4: Simulation results for different NAA contexts. (**A**) Offset from linear fit. (**B**) Fit slope coefficient, compared to line of identity. (**C**) Variance of data around linear fit.
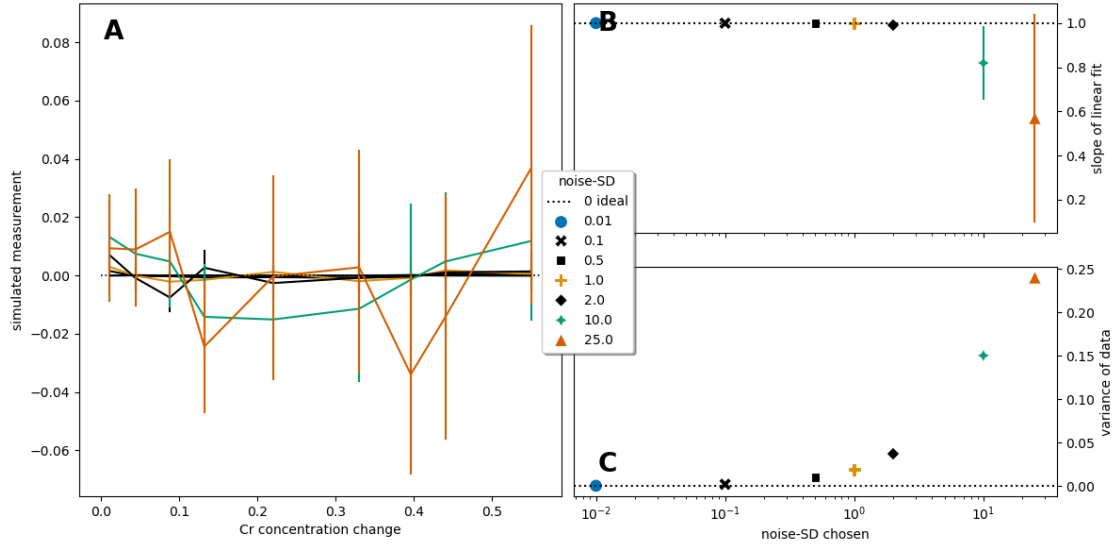
Figure 5: Simulation results for different Cr contexts. (**A**) Offset from linear fit. (**B**) Fit slope coefficient, compared to line of identity. (**C**) Variance of data around linear fit.
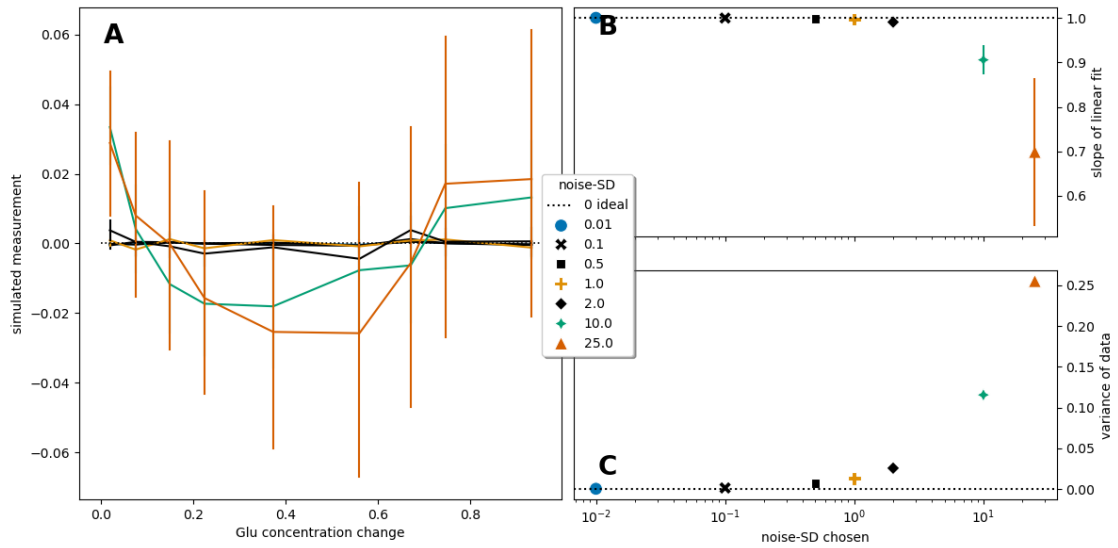
Figure 6: Simulation results for different Glu contexts. (**A**) Offset from linear fit. (**B**) Fit slope coefficient, compared to line of identity. (**C**) Variance of data around linear fit.

III. Concentration Correlation of Metabolites with Similar Spectra
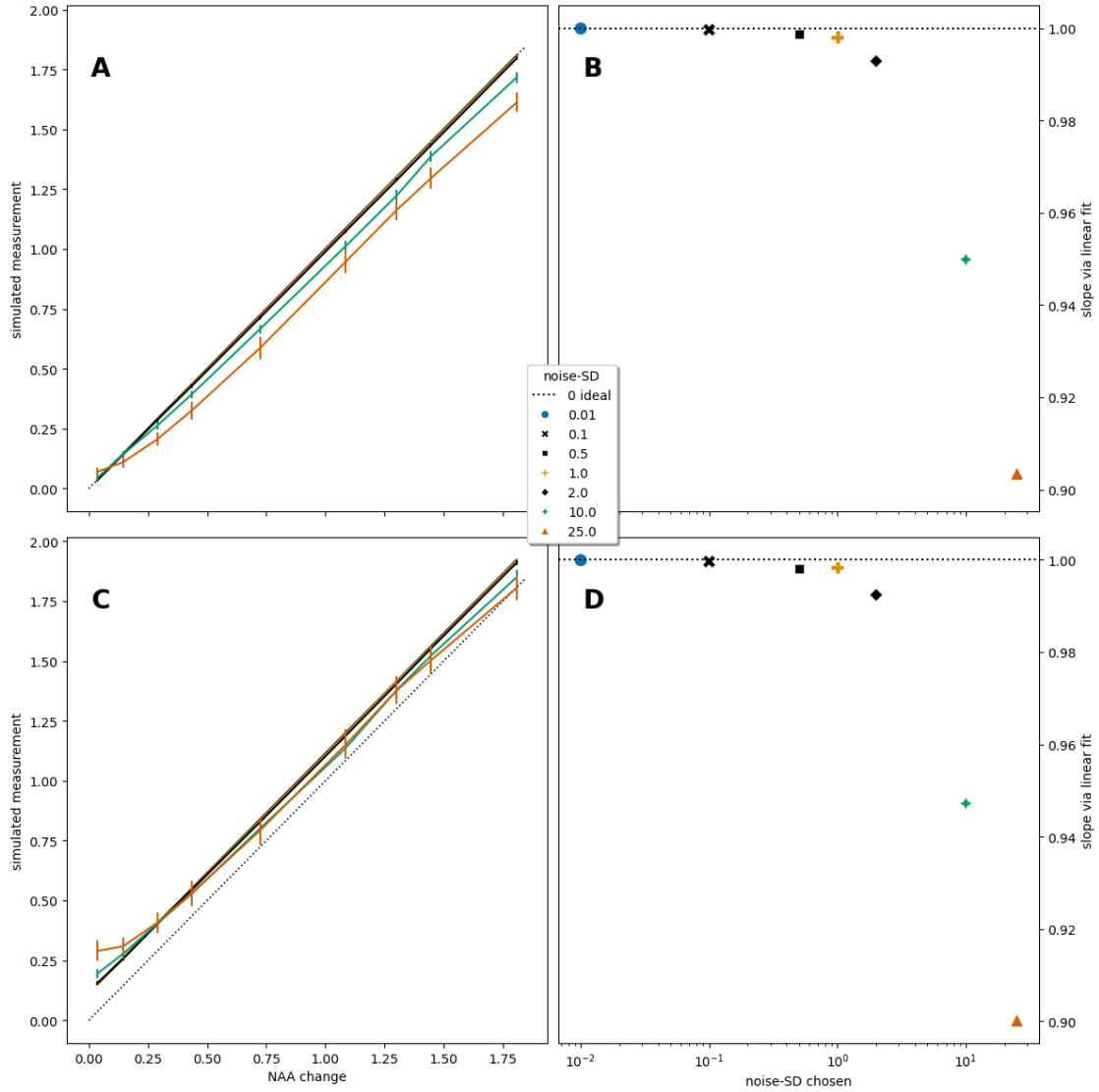
Figure 7: Changing NAA, comparing simulations of measuring single NAA, and NAA+GABA combined. (**A**) Fit of NAA concentration, with increasing NAA concentration input. (**B**) Slope of the simulation run, compared to the line of identity. (**C**),(**D**): Analogue for fitting NAA paired with GABA.
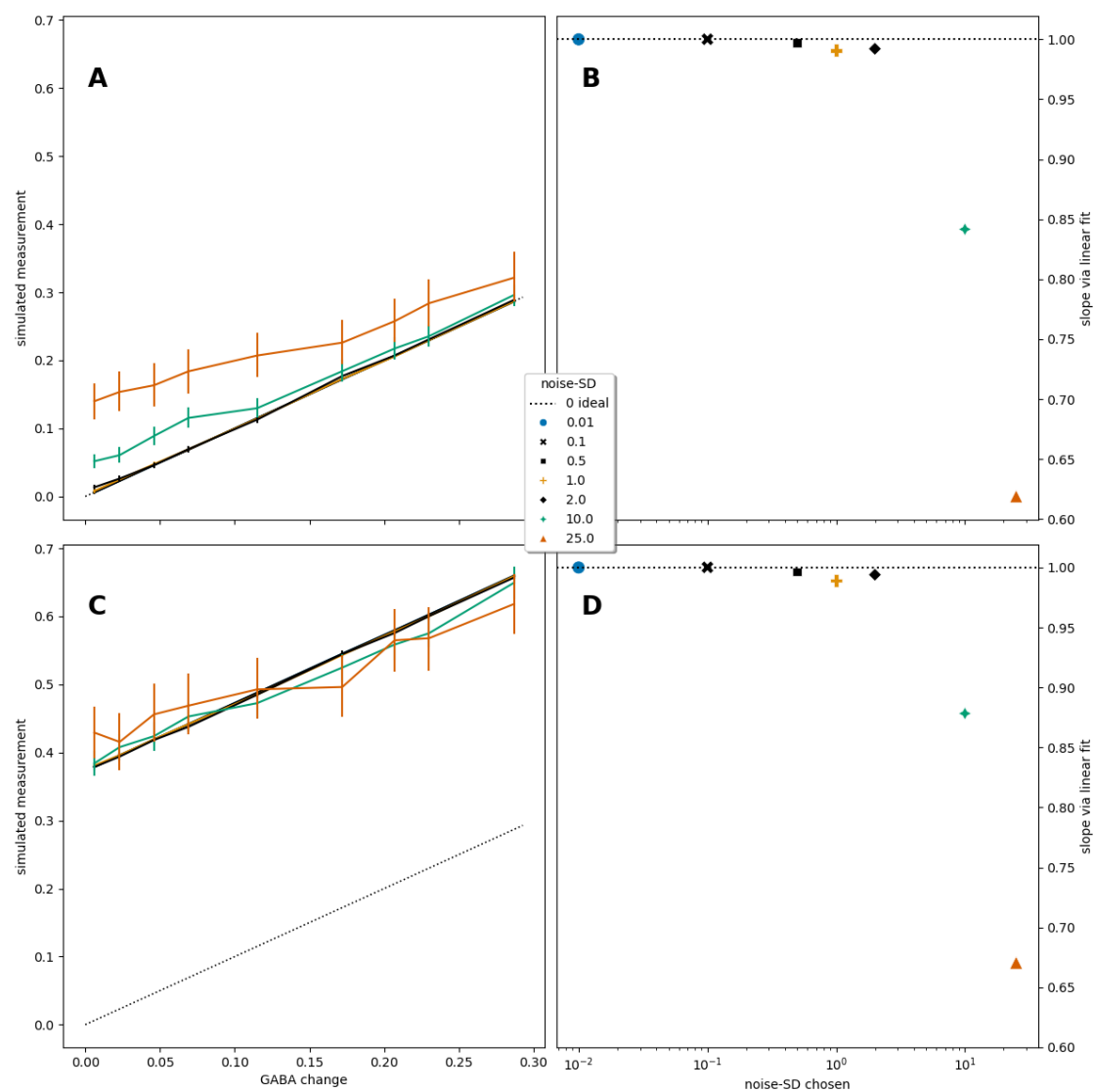
Figure 8: Changing GABA, comparing simulations of measuring single GABA, and GABA+Glu combined. (**A**) Fit of GABA concentration, with increasing GABA concentration input. (**B**) Slope of the simulation run, compared to the line of identity. (**C**),(**D**): Analogue for fitting GABA paired with Glu.

List of Figures

List of Tables

References

1. Wilson, M. *et al.* Methodological consensus on clinical proton MRS of the brain: Review and recommendations. *Magnetic Resonance in Medicine* **82,** 527–550. doi:`10.1002/mrm.27742` (Aug. 2019).

2. Fukuchi, M. *et al.* Excitatory GABA induces BDNF transcription via CRTC1 and phosphorylated CREB-related pathways in immature cortical cells. *Journal of Neurochemistry* **131,** 134–146. doi:`10.1111/jnc.12801` (Oct. 2014).

3. Ben-Ari, Y. Excitatory actions of gaba during development: the nature of the nurture. *Nature Reviews. Neuroscience* **3,** 728–739. doi:`10.1038/nrn920` (Sept. 2002).

4. Maier, S. *et al.* Increased prefrontal GABA concentrations in adults with autism spectrum disorders. *Autism Research: Official Journal of the International Society for Autism Research* **15,** 1222–1236. doi:`10.1002/aur.2740` (July 2022).

5. Balint, B. *et al.* Dystonia. *Nature Reviews. Disease Primers* **4,** 25. doi:`10.1038/s41572-018-0023-6` (Sept. 20, 2018).

6. Treiman, D. M. GABAergic mechanisms in epilepsy. *Epilepsia* **42 Suppl 3,** 8–12. doi:`10.1046/j.1528-1157.2001.042suppl.3008.x` (2001).

7. Colloca, L. *et al.* Neuropathic pain. *Nature reviews. Disease primers* **3,** 17002. doi:`10.1038/nrdp.2017.2` (Feb. 16, 2017).

8. Gottesmann, C. GABA mechanisms and sleep. *Neuroscience* **111,** 231–239. doi:`10.1016/s0306-4522(02)00034-9` (2002).

9. Kolasinski J *et al.* The dynamics of cortical GABA in human motor learning. *J Physiol* **597,** 271–282 (2019).

10. Stagg Cj *et al.* Local GABA concentration is related to network-level resting functional connectivity. *Elife* **3,** e01465 (Mar. 25, 2014).

11. Lou, H. C., Changeux, J. P. & Rosenstand, A. Towards a cognitive neuroscience of self-awareness. *Neuroscience and Biobehavioral Reviews* **83,** 765–773. doi:`10.1016/j.neubiorev.2016.04.004` (Dec. 2017).

12. Blicher, J. U. *et al.* GABA Levels Are Decreased After Stroke and GABA Changes During Rehabilitation Correlate With Motor Improvement. *Neurorehabilitation and Neural Repair* **29,** 278–286. doi:`10.1177/1545968314543652` (Mar. 2015).

13. Zhang, W. *et al.* Individual prediction of symptomatic converters in youth offspring of bipolar parents using proton magnetic resonance spectroscopy. *European Child & Adolescent Psychiatry* **30,** 55–64. doi:`10.1007/s00787-020-01483-x` (Jan. 1, 2021).

14. De Graaf, R. A. *In vivo NMR spectroscopy: principles and techniques* Third edition. 560 pp. (Wiley, Hoboken, NJ, 2019).

15. Gerlach, W. & Stern, O. Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld. *Zeitschrift für Physik* **9,** 349–352. doi:`10.1007/BF01326983` (Dec. 1922).

16. Rabi, I. I., Zacharias, J. R., Millman, S. & Kusch, P. A New Method of Measuring Nuclear Magnetic Moment. *Physical Review* **53,** 318–318. doi:`10.1103/PhysRev.53.318` (Feb. 15, 1938).

17. Pethig, R & Kell, D. B. The passive electrical properties of biological systems: their significance in physiology, biophysics and biotechnology. *Physics in Medicine and Biology* **32,** 933–970. doi:`10.1088/0031-9155/32/8/001` (Aug. 1, 1987).

18. Pauli, W. Über den Einfluß der Geschwindigkeitsabhängigkeit der Elektronenmasse auf den Zeemaneffekt. *Zeitschrift für Physik* **31,** 373–385. doi:`10.1007/BF02980592` (Feb. 1, 1925).

19. FbrG. *Larmor Precession* Jan. 7, 2021. doi:`https://commons.wikimedia.org/wiki/File:Precession_gamma.svg`.

20. Tiesinga, E., Mohr, P., Newell, D. B. & Taylor, B. 2018 CODATA RECOMMENDED VALUES OF THE FUNDAMENTAL CONSTANTS OF PHYSICS AND CHEMISTRY. *NIST* (June 3, 2019).

21. Borowiak, R., Groebner, J., Haas, M., Hennig, J. & Bock, M. Direct cerebral and cardiac 17O-MRI at 3 Tesla: initial results at natural abundance. *Magnetic Resonance Materials in Physics, Biology and Medicine* **27,** 95–99. doi:`10.1007/s10334-013-0409-0` (Feb. 2014).

22. Bloch, F. Nuclear Induction. *Physical Review* **70,** 460–474. doi:`10.1103/PhysRev.70.460` (Oct. 1, 1946).

23. Yao, J., Dyson, H. & Wright, P. E. Chemical shift dispersion and secondary structure prediction in unfolded and partly folded proteins. *FEBS Letters* **419,** 285–289. doi:`10.1016/S0014-5793(97)01474-9` (Dec. 15, 1997).

24. Mescher, M., Tannus, A., Johnson, M. & Garwood, M. Solvent Suppression Using Selective Echo Dephasing. *Journal of Magnetic Resonance, Series A* **123,** 226–229. doi:`10.1006/jmra.1996.0242` (Dec. 1996).

25. Bortz, J. *Statistik für Sozialwissenschaftler* 5., vollst. überarb. und aktualisierte Aufl. 836 pp. (Springer, Berlin Heidelberg, 1999).

26. Provencher, S. W. Automatic quantitation of localized in vivo 1H spectra with LCModel. *NMR in biomedicine* **14,** 260–264. doi:`10.1002/nbm.698` (June 2001).

27. Clarke, W. T., Stagg, C. J. & Jbabdi, S. FSL-MRS: An end-to-end spectroscopy analysis package. *Magnetic Resonance in Medicine* **85,** 2950–2964. doi:`10.1002/mrm.28630` (June 2021).

28. Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. FSL. *NeuroImage. 20 YEARS OF fMRI* **62,** 782–790. doi:`10.1016/j.neuroimage.2011.09.015` (Aug. 15, 2012).

29. *Numerical recipes: the art of scientific computing* (ed Press, W. H.) 3rd ed. OCLC: ocn123285342 (Cambridge University Press, Cambridge, UK ; New York, 2007). 1235 pp.

30. Popper, K. *The logic of scientific discovery* Special Indian Edition. 513 pp. (Routledge, London, 2010).

31. Pearson, E. S., Gosset, W. S., Plackett, R. L. & Barnard, G. A. *Student: a statistical biography of William Sealy Gosset* 142 pp. (Clarendon Press ; Oxford University Press, Oxford : New York, 1990).

32. Bosbach, K. E. *FSL-MRS Monte Carlo Experiment* `https://github.com/konbos/mrs_mce`. 2024.

33. Koolschijn, R. S. *et al.* Memory recall involves a transient break in excitatory-inhibitory balance. *eLife* **10** (eds Irish, M., Baker, C. I., Mullins, P. & Hutchinson, B.) e70071. doi:`10.7554/eLife.70071` (Oct. 8, 2021).

34. Scheenen, T. W. J., Klomp, D. W. J., Wijnen, J. P. & Heerschap, A. Short echo time ^{1}H-MRSI of the human brain at 3T with minimal chemical shift displacement errors using adiabatic refocusing pulses. *Magnetic Resonance in Medicine* **59,** 1–6. doi:`10.1002/mrm.21302` (Jan. 2008).

35. Tkáč, I. *et al.* Water and lipid suppression techniques for advanced ^{1}H MRS and MRSI of the human brain: Experts' consensus recommendations. *NMR in Biomedicine* **34,** e4459. doi:`10.1002/nbm.4459` (May 2021).

36. Laudadio, T., Mastronardi, N., Vanhamme, L., Van Hecke, P. & Van Huffel, S. Improved Lanczos Algorithms for Blackbox MRS Data Quantitation. *Journal of Magnetic Resonance* **157,** 292–297. doi:`10.1006/jmre.2002.2593` (Aug. 2002).

37. Soher, B. J. *et al.* Vespa: Integrated applications for RF pulse design, spectral simulation and MRS data analysis. *Magnetic Resonance in Medicine* **90,** 823–838. doi:`10.1002/mrm.29686` (Sept. 2023).

38. Smith, S. A., Levante, T. O., Meier, B. H. & Ernst, R. R. Computer Simulations in Magnetic Resonance. An Object-Oriented Programming Approach. *Journal of Magnetic Resonance, Series A* **106,** 75–105. doi:`10.1006/jmra.1994.1008` (Jan. 1, 1994).

39. Bosbach, W. A., Senge, J. F. & Dlotko, P. Mathematics in Medical Diagnostics - 2022 Proceedings of the 4th International Conference on Trauma Surgery Technology. In collab. with Senge *et al.* doi:`10.5281/ZENODO.7191419` (Aug. 20, 2022).

40. Mullins, P. G. Towards a theory of functional magnetic resonance spectroscopy fMRS : A meta-analysis and discussion of using MRS to measure changes in neurotransmitters in real time. *Scandinavian Journal of Psychology* **59,** 91–103. doi:`10.1111/sjop.12411` (Feb. 2018).

A. German Abstract

Mit Magnet-Resonanz-Spektroskopie (MRS) kann die Konzentration von Stoffwechselprodukten im Gehirn des lebenden Menschen non-invasiv bestimmt werden. Anders als die Magnet-Resonanz-Bildgebung (MRI), die das MR Signal von ^{1}H Atomen in Wasser misst, sind die Stoffwechselprodukte, die von MRS gemessen werden, in viel geringerer Konzentration. Da die Datenakquise im Menschen auf unter eine Stunde begrenzt ist, ist MRS eine Technik mit begrenztem Signal-zu-Rausch Verhältnis (SNR). Trotzdem konnte MRS in jüngerer Zeit benutzt werden, um kleine Veränderungen in Neurotransmitter-Konzentrationen während der Hirnaktivierung zu beobachten. MRS Spektren verschiedener Stoffwechselprodukte können überlappende Resonanzen haben, die das Separieren per Ausgleichsrechnung erschweren. Da MRS Daten von Rauschen und Artefakten verfälscht sein können, bleibt es kompliziert, die Messunsicherheit zu bestimmen, und die Anzahl benötigter Messungen für eine gegebene Messgenauigkeit zu bestimmen.

Diese Arbeit präsentiert eine Monte-Carlo Simulations-Software, um Konzentrations Messungenauigkeiten von MRS zu bestimmen. Durch die Verwendung von in vivo MRS Daten, die bei 7 T gewonnen wurden, werden Messbedingungen simuliert, um sowohl Signal Qualität, als auch Konzentrationen der Stoffwechselprodukte zu bewerten. Durch Interpolation der Simulationsdaten wird eine weite Auswahl an Messbedingungen schnell simuliert. Die Software erlaubt es dem Benutzer, sowohl Trennschärfen für funktionale MRS Studien-Designs zu berechnen, als auch die Trennschärfe von potentiellen Messungen zu bewerten. Sie erlaubt die Planung von statistisch ausdrucksvollen MRS Studien, für Anwendungen in Klinik und Neurowissenschaft, um im menschlichen Gehirn Konzentrationen von Stoffwechselprodukten zu messen.

Für verschiedene Studien-Designs, die Glutamat Konzentrationswechsel messen, wurde diese Methode verwendet, um Vorteile in Messungenauigkeit und Trennschärfe zu erwägen. Die Ergebnisse implizieren, dass die Studien-Größe zuverlässig reduziert werden kann, von 11 zu 3 Teilnehmern, durch das Verwenden von ereignisbasierten Designs, an Stelle der etablierten Blockstudien.

B. Declaration of Authorship

As required by the examination regulation:
Hereby I declare, that

(i) This thesis is my own work.

(ii) I have used no more sources and aids than those credited. All sources used have been quoted, according to the established rules of scientific conduct.

(iii) This thesis has not been submitted before, at any institution.

Moreover, no artificial intelligence has been used in the making of this thesis.

_____   _____
          Place & Date   Signature