

# Maximum likelihood estimation in signal analysis of MR spectroscopy

Pia Löthgren  
*Faculty of Forest Sciences,  
Department of Forest Economics,  
Centre of Biostochastics,  
Umeå*

Licentiate Thesis  
Swedish University of Agricultural Sciences  
Umeå 2012

Acta Universitatis agriculturae Sueciae

ISSN 0348-2049

ISBN 978-91-576-9108-8

© 2012 Pia Löthgren, Umeå

Print: SLU Service/Repro, Umeå 2012

## Maximum likelihood estimation in signal analysis of MR spectroscopy

### Abstract

Proton magnetic resonance spectroscopy ( $^1\text{H}$  MRS) is used to determine the concentration of metabolites in organic tissues, or to study metabolic changes in a non-invasive way. The complex-valued magnetic resonance spectroscopy signals are assumed to be disturbed by additive white noise. The distributional properties of the stochastic noise are studied. A statistical model for the magnitude and phase of the Fourier transformed magnetic resonance spectroscopy signal is introduced. Maximum likelihood estimators of the distributional parameters of this model are derived and asymptotic properties such as consistency, asymptotic normality and efficiency of the estimators are verified. A simulation study is used to test the findings and the model is tested on magnetic resonance spectroscopy data from a spectroscopy phantom and human brain data.

*Keywords:* magnetic resonance spectroscopy, maximum likelihood estimation, Rice distribution, phase modeling

*Author's address:* Pia Löthgren, SLU, Centre of Biostochastics,  
901 83 Umeå, Sweden.

*E-mail:* Pia.Loethgren@slu.se



# Contents

<b>1</b>	<b>Introduction</b>	<b>7</b>
<b>2</b>	<b>Basic principals of magnetic resonance spectroscopy</b>	<b>9</b>
<b>3</b>	<b>Signal modeling</b>	<b>15</b>
3.1	MRS quantitation	19
3.2	Measured MRS signals	20
3.3	Difference signals	23
<b>4</b>	<b>The Rice distribution</b>	<b>25</b>
4.1	ML estimation of the Rice distribution	26
4.1.1	Consistency of the ML estimators for the Rice distribution	28
4.2	Modification to a three parameter Rice distribution	29
<b>5</b>	<b>A Phase-Magnitude model</b>	<b>32</b>
5.1	Phase modeling in MRI	34
5.2	ML estimation in the phase-magnitude model	34
5.3	Asymptotic properties	35
<b>6</b>	<b>Simulation study</b>	<b>38</b>
6.1	Methods	38
6.1.1	Properties of the estimates for different sample sizes and SNR	38
6.1.2	Asymptotic normality	38
6.1.3	Properties of the estimates for different true value of the phase $\psi$	39
6.1.4	Comparison with estimates for the Rice distribution	39
6.1.5	Comparison with $\psi$ estimates from the phase marginal distribution	40
6.2	Results	40
6.2.1	Properties for different sample sizes and SNR	40
6.2.2	Asymptotic normality	40
6.2.3	Properties for different true value of the phase $\psi$	43
6.2.4	Comparison with estimates for the Rice distribution	47
6.2.5	Comparison with $\psi$ estimates from the phase marginal distribution	47
6.3	Discussion	47
<b>7</b>	<b>Phantom data study</b>	<b>50</b>
7.1	Method	50
7.2	Results	52
7.2.1	Method I	52
7.2.2	Method II	53
7.3	Discussion	53

<b>8</b>	<b>Human brain data study</b>	<b>55</b>
8.1	Method	55
8.2	Results	56
8.3	Discussion	57
<b>9</b>	<b>Final words</b>	<b>58</b>

## 1 Introduction

Nuclear magnetic resonance (NMR) has become a routinely used method to determine the structure of organic compounds. While magnetic resonance imaging (MRI) uses the signal from hydrogen protons to form anatomic and spatio-temporal images in living organisms, proton magnetic resonance spectroscopy ( $^1\text{H}$  MRS) uses this information to determine the concentration of metabolites in the tissue examined, or to study metabolic changes in a non-invasive way. With MRS, we are able to

- study the metabolism and biochemistry of the brain in action,
- observe gene expressions or the activity of a particular membrane receptor, using molecular imaging with targeted MR contrast agents,
- diagnose certain metabolic disorders, especially those affecting the brain,
- provide information on tumor metabolism.

MRS signals measured in animals or human beings are called *in vivo* signals. They are characterized by a low signal-to-noise ratio (*SNR*), due to the low concentration of the metabolites and the limited measurement time, and by overlapping spectral components due to the low magnetic field strength used. Historically the challenge has been to quantify proton short echo-time spectra, which exhibit many metabolites and to estimate their concentrations. Analysis of MRS signals can be conducted either in time-domain [78] or in frequency-domain [45]. Recent research has shown that the direct processing of the signal in the native time-domain was as powerful and sometimes more simple than the traditional processing of the signal in the frequency-domain. An overview of time-domain and frequency-domain quantitation methods was given by Pouillet et al. [50], where advantages and drawbacks of these two families of quantitation methods were discussed.

Quantitation in time-domain offers useful features: Missing data points do not really hamper the quantitation [21] since MR measurements are made in time-domain, and it also enables one to automatically process water and background signals [8] with SVD-based methods. Therefore models developed for MRS are mostly defined in time-domain [14, 53, 78].

In clinical applications, however, the measured MRS signals are often visually interpreted in frequency-domain, where the data are transformed by the discrete Fourier transform (DFT). Frequency-domain methods can be divided into two classes. The *nonparametric* methods are based on integration of the peak area of the frequency-domain signal [44], and *parametric* methods rely on a model function for the metabolite peaks, which are

often based on a time-domain model [52, 66, 79]. The model functions in the frequency-domain are, in general, more complicated than in the time-domain and necessitate thereby more computation time. On the other hand, the frequency-domain allows a straightforward selection of a frequency interval. Only the points in the frequency region of interest are considered for quantitation, resulting in faster algorithms [51, 61, 68].

The main objective of this thesis is to study the distributional properties of the stochastic noise in the MRS signals. We consider the asymptotic properties of the maximum likelihood (ML) estimators of the parameters in the model function, given some prior knowledge. In Chapter 2 we offer a brief introduction to the basic principals of MRS. Chapter 3 is devoted to introduce some commonly used mathematical models with measurement error for MRS signals. Given the assumption that the complex-valued MRS signals are disturbed by additive complex-valued white noise we derive that the magnitude of the signal in frequency-domain follows the Rice [55] distribution. In Chapter 4 ML estimation of the parameters of this distribution is discussed, and consistency of the estimates is proved. However, in the analysis of the magnitude, we lose the original information given by the complex-valued signal. Thus, we introduce a phase-magnitude model in Chapter 5. Phase-magnitude modeling has previously been done for different types of MR imaging techniques [59, 60], but not yet for MRS. The distributional properties of the phase-magnitude model is derived and ML estimator of the distributional parameters are calculated. Asymptotic properties of the ML estimators are verified.

In Chapter 6 we conduct a thorough simulation study to verify our conclusion for the phase-magnitude model and its ML estimators. The model is then tested on real MRS scans of an MRS phantom, in Chapter 7, and on *in vivo* MRS brain scans from a healthy volunteer in Chapter 8.

This is a EU project financed by the EU Regional Development Fund, Center of Biostochastics (SLU), and the Center for Biomedical Engineering and Physics (UmU and LTU).

This research was conducted using the resources of High Performance Computing Center North (HPC2N).

## 2 Basic principals of magnetic resonance spectroscopy

The theory behind magnetic resonance techniques is complex. In this Chapter we will only offer a very brief, and simplified, overview of the basic principles. There are many textbooks in the subject. This overview is mostly taken from [13] and [10].

In any MR technique we are performing experiments on the nuclei of atoms, not on the electrons. In most MRI and MRS techniques we are interested in the  $^1\text{H}$  nuclei, which is a proton, possessing a spin of  $1/2$ . We can imagine the proton spinning around its axis, which generates a magnetic field. If protons are placed in an external magnetic field  $B_0$ , their spins will either align with  $B_0$ , with the lower energy state  $1/2$ , or opposite  $B_0$ , with the higher energy state  $-1/2$ , see Figure 2.1.

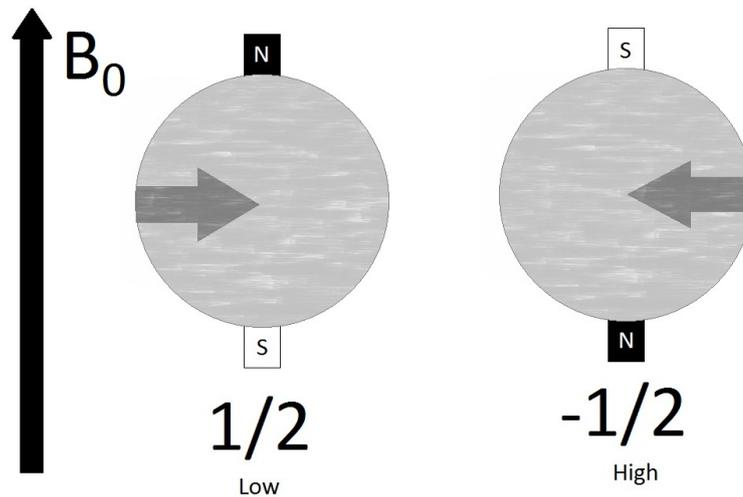


Figure 2.1: Protons in the external magnetic field  $B_0$  will either align their spin with, or opposite to,  $B_0$ .

We can imagine the protons as acting as tiny compass arrows. Let  $M$  denote the net magnetization [6]. In order to induce an MR signal, a second, rotation magnetic field,  $B_1$ , is applied perpendicular to  $B_0$ , which causes  $M$  to tip to alignment with  $B_1$ . This process is called an RF pulse and is done with an RF coil. When  $B_1$  is switched off,  $M$  spirals back to alignment with  $B_0$  and its original value of  $M_0$ , a process Bloch [6] denoted *relaxation*, see Figure 2.2. He introduced two time components,  $T_1$  and  $T_2$ , for the relaxation process.  $T_1$  describes the time for  $M_z$  to grow from 0 to  $1 - e^{-1}$ , about

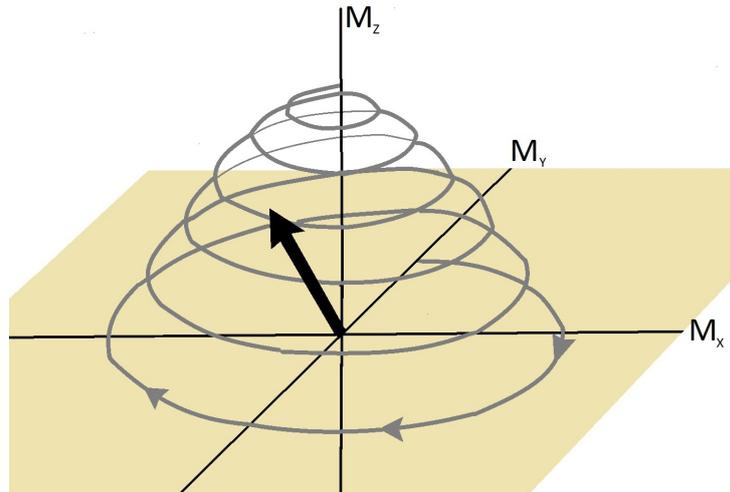


Figure 2.2: Relaxation after an RF pulse

63% of  $M_0$ .  $T_2$  represent the time for  $M_x$  or  $M_y$  to decay to  $e^{-1}$ , about 37% of  $M_0$ , see Figure 2.3. However, in practice we measure the decay rate  $T_2^*$ , which can be considered as an "observed"  $T_2$ . The  $T_2^*$  rate is faster than  $T_2$ , mostly due to inhomogeneity in the main magnetic field  $B_0$ . The  $T_2^*$  decay generates a so-called *free induction decay*, *FID*, signal. The signal is detected in a two channel quadrature coil, leading to two measured signals denoted *real* and *imaginary*.

The MRS signals are acquired in the complex time-domain. In practice a number of scans are done (usually 64 or 128) and then the means of the signals are computed. In order to visualize the signal the discrete Fourier transform (DFT) is taken and then the magnitude, or absolute values, of the signal. Figure 2.4 shows the real and imaginary part of a typical *in vivo* MRS signal and the real and imaginary part of the discrete Fourier transform of the same signal.

The magnitude spectrum consists of a collection of peaks, where each

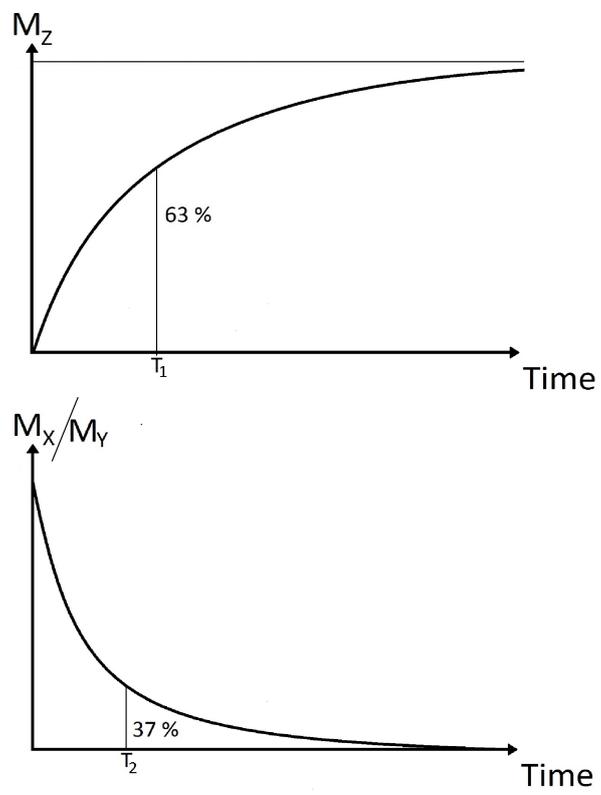


Figure 2.3: Upper:  $T_1$  is the time for  $M_z$  to grow from 0 to  $1 - e^{-1}$ . Lower:  $T_2$  is the time for  $M_x$  or  $M_y$  to decay to  $e^{-1}$

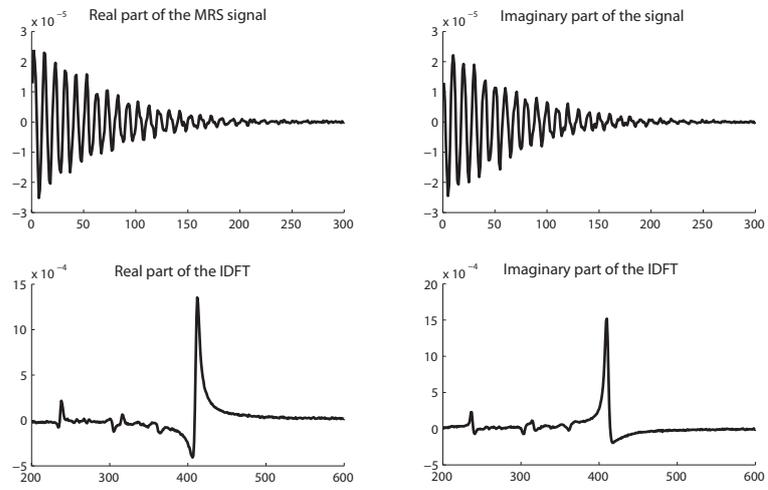


Figure 2.4: Upper left: Real part of MRS signal acquired in a healthy brain. Upper right: Imaginary part of the MRS signal. Lower left: Real part of the DFT of the MRS signal. Lower right: Imaginary part of the DFT of the MRS signal.

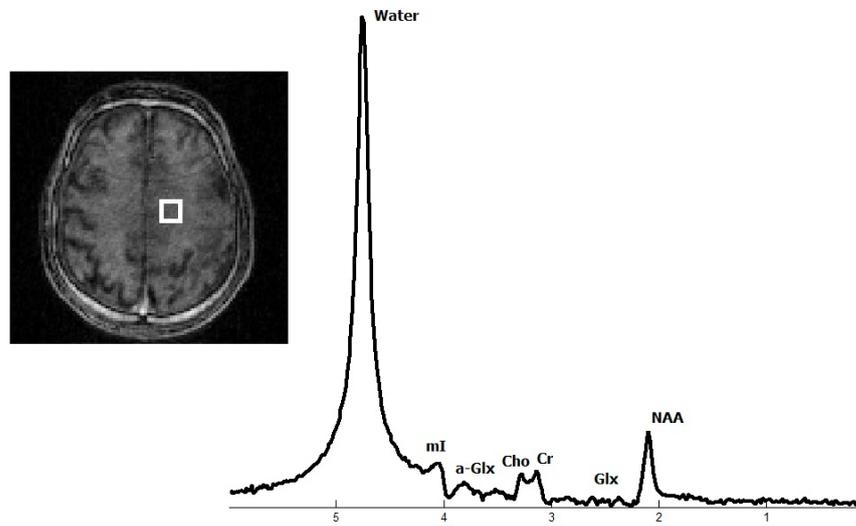


Figure 2.5: Typical MRS magnitude spectrum of a healthy brain.

peak may be described as the MR signal from magnetically equivalent protons [10, 5, 26, 34, 37, 46, 47, 65], for instance the protons in a specific metabolite. By costume the positions of the peaks in the frequency-domain are measured by *chemical shift*, which is the local magnetic field perceived by the specific proton group divided by the applied magnetic field  $B_0$ . Chemical shift is usually measured in *ppm* relative to a reference compound. For  $^1\text{H}$  MRS the reference is tetramethylsilane,  $\text{Si}(\text{CH}_3)_4$  [13].

Rubæk and Ross [10] gave a table, Table 2.1 for how different peaks positions may be used to identify the metabolites of the human brain. The water peak is set to have  $ppm = 4.76$  and hence the position of the other peaks follows.

Figure 2.5 shows the magnitude spectrum of at typical MRS signal from a healthy brain with some of the prominent peaks labeled. This Single Voxel Spectroscopy (SVS) is acquired in a  $2 \times 2 \times 2 \text{ cm}^3$  voxel.

The area under each peak represents the concentration of that specific metabolite. MRS *quantitation* aims to calculated these areas. We will return to this subject in Chapter 3.1. For now we will just note that in *in vivo* MRS the widths of the individual peaks are similar [10].

Since the actual scale on the *y*-axes in the MRS spectrum varies with different scan hardware peak hight is usually measured in ratio. Most commonly, ratios relative to  $Cr = 1$  are used.

The concentrations of the most common metabolites in the human brain are quite well known, as well as the changes in concentration due to many neurological diseases. Many of these values can be found in [10]. Figure 2.6 show the MRS spectrum from a brain tumor patient. The black line is the spectrum taken in the tumor. The grey line is the reference spectrum, taken in a healthy part of the patients brain. We can, for instance, see the decrease of the peak around  $ppm = 2$  for the tumor measurement, which might describe the expected decrease in N-acetylaspartate, NAA, in the tumor.

Resonance observed in normal MRS	Chemical shift (ppm)
N-acetylaspartate (first peak, $NAA_1$ )	2.02
$\beta, \gamma$ -Glutamine and glutamate ( $\beta, \gamma$ -Glx)	2.05 - 2.5
N-acetylaspartate (second peak, $NAA_2$ )	2.6
N-acetylaspartate (third peak, $NAA_3$ )	2.5
Total Creatine (Cr)	3.03
Total Choline (Cho)	3.22
<i>scyllo</i> -inositol (sI)	3.36
Glucose	3.43
<i>myo</i> -inositol (mI)	3.56
$\alpha$ -Glutamine and glutamate ( $\alpha$ -Glx)	3.65 - 3.8
Second peak of Glucose	3.8
Second peak of Gr	3.9
Second peak of mI	4.06

Table 2.1: Chemical shift of metabolites in MRS of the human brain.

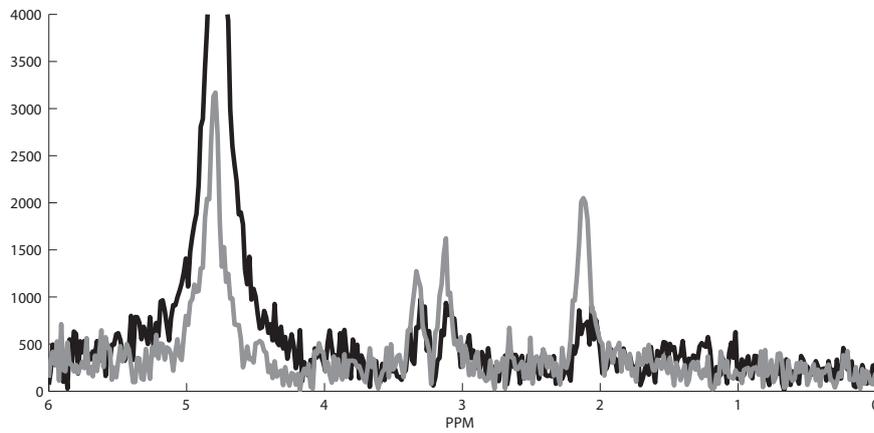


Figure 2.6: MRS measurement from patient with brain tumor. Black line: Tumor. Grey line: Reference measurement from healthy part of the patients brain.

### 3 Signal modeling

The model function most often used to represent the digitized MRS FID signal is the *Lorentzian* model, a sum of exponentially damped sinusoids in time-domain, contaminated with random noise:

$$y(t) = s(t) + \varepsilon(t) = \sum_{k=1}^K a_k e^{i\phi_k} e^{(-d_k + i\omega_k)t} + \varepsilon(t), \quad (3.1)$$

where  $K$  is the number of sinusoidal components (representing different resonances) in the signal,  $\omega_k \in (0, 2\pi]$  is the (angular) frequency of the  $k$ th sinusoid;  $a_k$  is its amplitude related to the concentration of the metabolite;  $d_k > 0$  is its damping (decaying) constant providing, among other things, information about its mobility and molecular environment;  $\phi_k$  is its phase; and  $i = \sqrt{-1}$ . The noise term  $\varepsilon(t)$  is usually assumed to be a random complex white noise with a circular Gaussian distribution. The term circular means that the real and imaginary parts of the noise are not correlated and have equal variance.  $\varepsilon(t)$  mainly consists of thermal noise, generated from the sample, coil and preamplifier and noise introduced by the remaining receiver and electronic components such as the analog-to-digital converter [24, 25]. The parameters in the model will in general change from voxel to voxel [23, 67]

The Lorentzian model (3.1) contains a number of parameters. The estimation of these is a well studied problem, for which a large number of solutions have appeared in both the MR literature and the signal processing literature (see e.g. [50] and many references therein). However, in MRS applications, prior knowledge concerning the spectral components is often present and needs to be incorporated into the model [69, 77]. As pointed out in [67], in virtually all spectroscopic imaging applications the compounds being imaged and their MR spectroscopic structure are known. More exactly, with reference to [67], the number of components (peaks)  $K$  is known. Although the frequencies and initial phases of the spectral lines present are not exactly known due to inhomogeneity of the main magnetic field, their differences are known. In other words we can write

$$\begin{aligned} \omega_k &= \omega_0 - \Delta\omega_k, \quad k = 1, \dots, K, \\ \phi_k &= \phi_0 - \Delta\phi_k, \quad k = 1, \dots, K, \end{aligned}$$

where  $\omega_0$  and  $\phi_0$  are unknown, but  $\{\Delta\omega_k, \Delta\phi_k, k = 1, \dots, K\}$  are known. Moreover, the damping constants  $d_k$  are also usually known with quite good accuracy. Based on this prior knowledge, model (3.1) can be simplified

to

$$y(t) = s(t) + \varepsilon(t) = \sum_{k=1}^K \Phi_k(t) a_k e^{i(\phi_0 + \omega_0 t)} + \varepsilon(t), \quad (3.2)$$

where

$$\Phi_k(t) = e^{-i\Delta\phi_k - (d_k + i\Delta\omega_k)t}$$

are given for  $k = 1, \dots, K$ .

Certainly, the above assumptions are valid under some conditions which must be evaluated for individual applications. As an example, Vanhamme et al. [78] illustrated the prior knowledge using adenosine triphosphate (ATP), and derived the following relations: a) the damping constants of all peaks are equal; b) the phases of all peaks are equal; c) the amplitudes relates with a fixed ratio; and d) the frequency differences between individual resonances are equal and known.

There are two alternative model functions that have received some attention. One is the *Gaussian* model:

$$y(t) = s(t) + \varepsilon(t) = \sum_{k=1}^K a_k e^{i\phi_k} e^{(-g_k t + i\omega_k)t} + \varepsilon(t),$$

and the another is the *Voigt* model [39, 64] which combines the Lorentzian and Gaussian model:

$$y(t) = s(t) + \varepsilon(t) = \sum_{k=1}^K a_k e^{i\phi_k} e^{(-d_k - g_k t + i\omega_k)t} + \varepsilon(t)$$

In the frequency-domain, the spectrum of the signal  $y(t)$ ,  $Y(\omega)$ , is the Fourier transform (FT) of model (3.1):

$$Y(\omega) = \int_0^{\infty} (s(t) + \varepsilon(t)) e^{-i\omega t} dt \triangleq S(\omega) + \Upsilon(\omega) \quad (3.3)$$

where the ideal MRS spectrum (i.e., without noise) is a sum of a series of

Lorentzian functions, one for each peak:

$$\begin{aligned}
S(\omega) &= \int_0^{\infty} s(t) e^{-i\omega t} dt \\
&= \int_0^{\infty} \sum_{k=1}^K a_k e^{i\phi_k} e^{(-d_k + i\omega_k)t} e^{-i\omega t} dt \\
&= \sum_{k=1}^K \frac{a_k}{d_k} e^{i\phi_k} \left( \frac{1}{1 + ((\omega - \omega_k)/d_k)^2} + i \frac{(\omega - \omega_k)/d_k}{1 + ((\omega - \omega_k)/d_k)^2} \right) \\
&\triangleq \sum_{k=1}^K e^{i\phi_k} (A_k(\omega) + iD_k(\omega))
\end{aligned}$$

Here  $A_k(\omega)$  and  $D_k(\omega)$  represent the absorption and the dispersion signal, respectively. This model describes a helix in the three-dimensional space comprised of the real-imaginary complex plane and the chemical shift/frequency axis [15]. Note that the bracketed terms are of form  $\frac{1}{1+x^2} + i \frac{x}{1+x^2}$ , representing a circle. The projection of the helix onto the complex plane forms a set of parametric equations that describes circles with centers at the real-imaginary coordinates  $(\frac{a_k}{2d_k}, 0)$  and radii  $\frac{a_k}{2d_k}$  for  $\phi_k = 0$ . The effect of  $\phi_k$  is simply to rotate the circle about the origin of the complex plane and move the circle's center. Thus the model spectrum is a set of circles, one for each peak, which form a series of nested curves in the complex plane resolved in the orthogonal dimension by the  $\omega$  parameter. Gabr et al. [15] summarized the relationships between the peak parameters in time-domain and frequency-domain as in Table 3.1. Because the trajectory of each peak in the complex plane is circular, one can model them with active circles that are contours that adaptively deform to best fit the model spectrum to the measured spectrum by minimizing a measure of the fit error while preserving their circular shape in the complex plane [15].

By assuming that phase correction has been applied properly, the dispersive part of the spectrum  $D_k(\omega)$  can be neglected, resulting in the *absorption mode* signal:

$$S(\omega) = \sum_{k=1}^K e^{i\phi_k} A_k(\omega) = \sum_{k=1}^K a_k e^{i\phi_k} \frac{d_k}{d_k^2 + (\omega - \omega_k)^2}, \quad (3.4)$$

where  $d_k$  stands for the half width as half height (HWHH). The height at the maximum of a Lorentzian function  $A_k$  is given as  $A_k(\omega_k) = a_k/d_k$ , and the area of  $A_k$  equals  $a_k \pi$  [31].

Parameter	Time	Frequency	Circle
$\omega_k$	Frequency	Peak position	Location of point density minimum of circle
$a_k$	Amplitude	Peak area	Radius times number of points in semicircle centered at $\omega_k$
$d_k$	Damping constant	Linewidth	Number of points in semicircle centered at $\omega_k$
$\phi_k$	Phase shift	Mode	Circle position/orientation
$a_k/d_k$	Amplitude/decay	Peak height	Circle diameter

Table 3.1: Relationship between the peak parameters in different domains

Even in the ideal case of model (3.3), the number of local maxima does not necessarily equal the number of single Lorentzian functions, due to effects of overlapping. Koh et al. [31] proposed an approach, called *Lorentzian Spectrum Reconstruction*, for peak identification and parameter approximation in order to automatically model an MRS spectrum as a superposition of single Lorentzian functions.

The amplitude spectral density, also called the *magnitude spectra*, of the signal  $y(t)$  is given by

$$|Y(\omega)| = \sqrt{\text{Re}(S(\omega) + \Upsilon(\omega))^2 + \text{Im}(S(\omega) + \Upsilon(\omega))^2}, \quad (3.5)$$

and the corresponding ideal magnitude spectra of Lorentzian types is

$$|S(\omega)| = \sqrt{\text{Re}(S(\omega))^2 + \text{Im}(S(\omega))^2}, \quad (3.6)$$

where

$$\begin{aligned} \text{Re}(S(\omega)) &= \sum_{k=1}^K (A_k(\omega) \cos \phi_k - D_k(\omega) \sin \phi_k), \\ \text{Im}(S(\omega)) &= \sum_{k=1}^K (A_k(\omega) \sin \phi_k - D_k(\omega) \cos \phi_k), \end{aligned}$$

In the special case of the absorption mode, the magnitude spectra is equivalent to the signal itself. Thus modeling the magnitude spectra directly will be an alternative way to estimate the parameters in the signal. This motivates the investigation of the distributional properties of  $|Y(\omega)|$ .

### 3.1 MRS quantitation

It is well known that characteristic resonance peaks at certain frequencies correspond to important brain metabolites [5, 10, 26, 34, 37, 46, 47, 65]. These peaks might be used as discriminatory features to distinguish tumor type, in particular for those regions of the spectrum which are clearly different between spectra of different tumor types. Thus, instead of using complete spectra, certain regions of the magnitude spectrum which are assumed to contain most of the information can be selected. For example, characteristic metabolites can be observed in the following regions of the magnitude MRS spectrum [12, 10, 36]: NAA (1.935-2.085 ppm), Cr (2.955-3.105 ppm), Cho (3.115-3.265 ppm), etc. The number of visible metabolites is larger in short echo time spectra than in long echo time spectra. Note that the selected regions are based on the metabolites that are assumed to be most characteristic according to prior knowledge available from field experts, and therefore the selection is subjective.

Because the amplitude of a resonance is proportional to the integral of the corresponding peak in the spectrum, the traditional approaches to estimate the metabolite concentration are based on peak integration [44]. The advantage of these methods is that no assumptions have been made concerning the lineshape of the signal. The major drawback is the low estimation accuracy. MR data from biological systems usually present severe difficulties for peak area integration, the most pronounced problems often being the extreme overlap of different resonances of interest, wavy baseline, low signal-to-noise ratio (*SNR*), and the discrete nature of the observed spectrum. In such cases, it is impossible to utilize straightforward integration techniques in a reliable manner and one has to make use of lineshape fitting methods, preferably those able to include prior knowledge [45].

Lineshape fitting methods are based on parametric frequency-domain models, such as the Lorentzian model. However, a simple exact analytical expression for the Gaussian or the Voigt model is not available, even if numerical approximations exist [19, 39]. For example, in [17, 39], approximated Voigt lineshapes have been proposed, and the spectra were fit with the Levenberg-Marquardt algorithm. In any case, the model functions in the frequency-domain are, in general, more complicated than in the time-domain, and necessitate thereby more computational time. Marshall et al. [38] show that the choice of the lineshape affects the metabolite peak areas and suggest the use of Gaussian lineshapes instead of Lorentzian lineshapes. The frequency-domain methods who only use the real part of the spectrum in their model, such as LCModel [52], require a very good phasing to get the spectrum in its absorption mode.

Many frequency-domain methods solve the non-linear least squares (NLLS) problem by local optimization techniques, in particular using the Levenberg-Marquardt algorithm (see, e.g. [52, 79]). Vanhamme [77] gave a formal proof of the statement that the solution of the NLLS problem in the frequency-domain is the same as that in the time-domain. Hence, time and frequency-domain fitting are equivalent from a theoretical point of view. But the data are always presented in the frequency-domain since this enables direct (visual) interpretation.

There have been other quantitation methods proposed for in vivo MR spectra, such as artificial neural network (ANN) [22, 30], principal component analysis (PCA) [70, 71, 72, 73], independent component analysis (ICA) [33], analysis of circles (CFIT) [15], time-domain frequency-domain fitting (TDFDF) [64], single value decomposition (SVD) [61], and the wavelet transform (WT) [17].

One of the main difficulties in quantitative analysis of MRS is that MRS signals appear to be contaminated with random noise. A number of work has been devoted to noise reduction, see e.g. [4, 11, 79]. In general, estimation of NMR spectral parameters, using e.g. maximum likelihood methods, is commonly based on the assumption of white complex Gaussian noise in the signal. Grae and Akke [18] tested the validity of this fundamental assumption. They showed that in general the noise in the sampled signal is not strictly white, even if the thermal noise in the receiver steps prior to digitization can be characterized as white Gaussian noise. They also showed that the noise correlation properties depend on the ratio between the sampling frequency and the filter cut-off frequency, as well as the filter characteristics, and identified the conditions that are expected to yield non-white noise in the sampled signal.

### 3.2 Measured MRS signals

Often the acquisition of the signal is only started after a time delay  $t_0$ , the receiver dead time. Therefore, the sampling time instances  $t_n$ ,  $n = 0, \dots, N - 1$ , can be written as  $t_n = n\Delta t + t_0$  with  $\Delta t$  as the sampling interval. The quantity  $\omega_k t_0$  is also called the first-order phase. Without loss of generality, we can assume that  $t_0 = 0$  and  $\Delta t = 1$ . Thus, model (3.1) and (3.2) with measurements becomes

$$y_n = s_n + \varepsilon_n = \sum_{k=1}^K a_k e^{i\phi_k} e^{(-d_k + i\omega_k)n} + \varepsilon_n, \quad (3.7)$$

and

$$y_n = s_n + \varepsilon_n = \sum_{k=1}^K \Phi_{k,n} a_k e^{i(\phi_0 + \omega_0 n)} + \varepsilon_n, \quad (3.8)$$

respectively, where

$$\Phi_{k,n} = e^{-i\Delta\phi_k - (d_k + i\Delta\omega_k)n}$$

are given for  $k = 1, \dots, K$  and  $n = 0, \dots, N-1$ . The parameters in the model can be estimated by minimizing

$$\sum_{n=0}^{N-1} |y_n - s_n|^2. \quad (3.9)$$

The solution of this NLLS problem provides the ML estimates of the parameters, in the case of circular and white Gaussian noise.

In frequency-domain, the measured data are transformed by the discrete Fourier transform (DFT),

$$\begin{aligned} Y_l &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} (s_n + \varepsilon_n) e^{-iv_l n} \\ &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \sum_{k=1}^K a_k e^{i\phi_k} e^{(-d_k + i\omega_k)n} e^{-iv_l n} + \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \varepsilon_n e^{-iv_l n} \\ &= \frac{1}{\sqrt{N}} \sum_{k=1}^K a_k e^{i\phi_k} \frac{1 - e^{(-d_k + i(\omega_k - v_l))N}}{1 - e^{-d_k + i(\omega_k - v_l)}} + \Upsilon_l \\ &\triangleq S_l + \Upsilon_l \end{aligned}$$

where  $v_l = 2\pi l/N$ ,  $l = 0, \dots, N-1$ . The magnitude spectra is given by

$$Z_l = |Y_l| = \sqrt{\text{Re}(S_l + \Upsilon_l)^2 + \text{Im}(S_l + \Upsilon_l)^2}. \quad (3.10)$$

Since the DFT is periodic, one can evaluate  $Y_l$  for other values of  $l$ . The convention in MRS is to evaluate  $Y_l$  for  $l = -\frac{N}{2}, \dots, \frac{N}{2}-1$  and to display the spectra from positive to negative frequency. After DFT each exponentially damped sinusoid gives rise to a (Lorentzian) peak in the obtained frequency-domain.

Consider the noise in frequency-domain, we have

$$\begin{aligned}
\Upsilon_l &= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \varepsilon_n e^{-i\nu_l n} \\
&= \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} (\operatorname{Re}(\varepsilon_n) \cos(\nu_l n) + \operatorname{Im}(\varepsilon_n) \sin(\nu_l n)) \\
&\quad + i \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} (\operatorname{Re}(\varepsilon_n) \sin(\nu_l n) + \operatorname{Im}(\varepsilon_n) \cos(\nu_l n)) \\
&\triangleq \operatorname{Re}(\Upsilon_l) + i \operatorname{Im}(\Upsilon_l)
\end{aligned}$$

From this it can be seen that the DFT of a circular complex white noise sequence with standard deviation  $\sigma$  for the real and imaginary parts is again circular complex white with the same standard deviation. The Gaussian distribution is retained because it is a finite linear combination of Gaussian variables. Therefore  $\Upsilon_l$  is still circular white Gaussian, with  $\operatorname{Re}(\Upsilon_l), \operatorname{Im}(\Upsilon_l) \sim N(0, \sigma^2)$ .

To obtain the parameters, one can minimize the difference between the DFT of the measured signal and the frequency-domain model function

$$\sum_{l=0}^{N-1} |Y_l - S_l|^2. \quad (3.11)$$

The solution of this NLLS problem is the same as the one obtained by minimizing (3.9) directly in the time-domain [77]. It is worth noting that in the early stage of model fitting in the frequency-domain, often a sampled version of the theoretical spectrum obtained by continuous Fourier transformation of a continuous time-domain signal was used to fit the measured data. Since there are several discrepancies between the DFT spectrum and the continuous FT spectrum, the results obtained as such are not optimal (see also [1, 16, 40] for further details). Examples of methods in which (3.11) is either solved directly or using some approximations can be found in [9, 22, 32, 48].

Since  $\Upsilon_l$  is circular white Gaussian, with  $\operatorname{Re}(\Upsilon_l), \operatorname{Im}(\Upsilon_l) \sim N(0, \sigma^2)$ , it follows that  $Y_l$  is complex-valued Gaussian distributed with

$$\operatorname{Re}(Y_l) \sim N(\operatorname{Re}(S_l), \sigma^2) \quad (3.12)$$

$$\operatorname{Im}(Y_l) \sim N(\operatorname{Im}(S_l), \sigma^2). \quad (3.13)$$

Hence,  $Z_l$ , in (3.10), is *Rican distributed* [55, 56] with non-centrality parameter

$$r_l = \sqrt{(\operatorname{Re}(S_l))^2 + (\operatorname{Im}(S_l))^2} \quad (3.14)$$

and scale parameter  $\sigma^2$ . This distribution is more generally known as the *non-central  $\chi$  distribution* [29] with two degrees of freedom.

### 3.3 Difference signals

In the clinical applications, patients are often measured by MRS at several occasions after the oncological treatments. One would like to analyze the effects of the treatments through possible changes in tumor metabolism and then make some diagnoses on the tumor development. For this purpose, we introduce a quantitation method based on the difference signals. Assume that two MRS signals,  $y^I(t)$  and  $y^{II}(t)$ , acquired from the same patient (and the same voxel) but at different time occasions, usually a couple of weeks between:

$$y^I(t) = s^I(t) + \varepsilon^I(t) \quad y^{II}(t) = s^{II}(t) + \varepsilon^{II}(t).$$

The difference signal:

$$y^D(t) = y^{II}(t) - y^I(t) = (s^{II}(t) - s^I(t)) + (\varepsilon^{II}(t) - \varepsilon^I(t)) = s^D(t) + \varepsilon^D(t)$$

will be used to analyze the changes in metabolites, such as amplitudes and phases.

According to model (3.8), if the Lorentzian model is used, the two acquired signals can be represented as follows:

$$\begin{aligned} y_n^I &= \sum_{k=1}^{K^I} \Phi_{k,n}^I a_k^I e^{i(\phi_0^I + \omega_0^I n)} + \varepsilon_n^I, \quad n = \dots, N^I - 1, \\ y_n^{II} &= \sum_{k=1}^{K^{II}} \Phi_{k,n}^{II} a_k^{II} e^{i(\phi_0^{II} + \omega_0^{II} n)} + \varepsilon_n^{II}, \quad n = \dots, N^{II} - 1, \end{aligned}$$

where  $\varepsilon_n^I$  and  $\varepsilon_n^{II}$  are assumed to be circular and white Gaussian noises with standard deviation  $\sigma^I$  and  $\sigma^{II}$  respectively.

As mentioned earlier in this Chapter, based on prior knowledge,  $\Phi_{k,n}^I$  and  $\Phi_{k,n}^{II}$  are given. We assume that the number of metabolites keeps unchanged ( $K^I = K^{II} = K$ ), the length of measured signals is the same ( $N^I = N^{II} = N$ ), and the standard deviations of the noises are constant ( $\sigma^I = \sigma^{II} = \sigma/2$ ). Suppose that our main interest is to detect the changes in the amplitudes (any significant changes in the metabolites will indicate the effect of oncological treatments). Then we can also assume that the frequency and the initial phase keep unchanged ( $\omega_0^I = \omega_0^{II} = \omega_0$ ,  $\phi_0^I = \phi_0^{II} = \phi_0$ ), though they are unknown and need to be estimated.

Thus, the measured difference signal can be formulated as follows:

$$y_n^D = s_n^D + \varepsilon_n^D, \quad n = 0, \dots, N-1$$

where

$$\begin{aligned} s_n^D &= \sum_{k=1}^K (\Phi_{k,n}^{II} a_k^{II} - \Phi_{k,n}^I a_k^I) e^{i(\phi_0 + \omega_0 n)} \\ \varepsilon_n^D &= \varepsilon_n^{II} - \varepsilon_n^I. \end{aligned}$$

By taking the discrete Fourier transform, the difference signal can be modeled in the frequency-domain

$$Y^D = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} (s_n^D + \varepsilon_n^D) e^{i\nu n} \triangleq S^D + \Upsilon^D$$

where  $\nu = 2\pi l/N$ ,  $l = 0, \dots, N-1$ . The magnitude spectra is given by

$$Z = |Y^D| = \sqrt{(\operatorname{Re}(S^D + \Upsilon^D))^2 + (\operatorname{Im}(S^D + \Upsilon^D))^2} \quad (3.15)$$

## 4 The Rice distribution

The next step is to investigate the distributional properties of the Rice distribution, and the focus will be on maximum likelihood estimation for the parameters involved in the noise distribution. We start by considering the non-centrality parameter  $r_l$  to be a single (unknown) constant, which means that we will not take the information in  $Re(S_l)$  and  $Im(S_l)$  into account at the first moment. We will however come back to this later. Since we are going to study the noise distribution of  $Z_l$  for each fixed frequency  $l$ , we skip the index  $l$  for simplicity.

The probability distribution function for a Rician distributed stochastic variable  $Z \sim Rice(r, \sigma^2)$  is

$$f_Z(z) = \frac{z}{\sigma^2} e^{-(z^2+r^2)/2\sigma^2} I_0\left(\frac{zr}{\sigma^2}\right), \quad z \geq 0 \quad (4.1)$$

where  $I_0$  is the modified Bessel function of the first kind [54, 49] of order 0. When  $r = 0$  (only noise is present) we obtain the Rayleigh distribution with mean  $\sigma\sqrt{\pi/2}$  and variance  $(2 - \pi/2)\sigma^2$ , which may be used to estimate the true noise power  $\sigma^2$ . When  $r$  is large the distribution can be approximated by  $N(\sqrt{r^2 + \sigma^2}, \sigma^2)$ .

It can be shown [63] that the  $p$ -th moment of  $Z$  is analytically expressed by

$$EZ^p = (2\sigma^2)^{p/2} \Gamma\left(1 + \frac{p}{2}\right) {}_1F_1\left(-\frac{p}{2}; 1; -\frac{r^2}{2\sigma^2}\right)$$

where  $\Gamma(\cdot)$  is the Gamma function and  ${}_1F_1(\cdot)$  is the confluent hypergeometric function [49]. The even moments are simple polynomials. For example,

$$\begin{aligned} EZ^2 &= r^2 + 2\sigma^2, \\ EZ^4 &= r^4 + 8r^2\sigma^2 + 8\sigma^4. \end{aligned}$$

However, the odd moments of  $Z$  are much more complex. For instance

$$EZ = \sigma \sqrt{\frac{\pi}{2}} e^{-r^2/4\sigma^2} \left( \left(1 + \frac{r^2}{2\sigma^2}\right) I_0\left(\frac{r^2}{4\sigma^2}\right) + \frac{r^2}{2\sigma^2} I_1\left(\frac{r^2}{4\sigma^2}\right) \right),$$

where  $I_1$  is the modified Bessel function of the first kind of order 1.

The Rice distribution can be well approximated by a Gaussian distribution at high  $SNR$ , defined by  $SNR = r/\sigma$ . When  $SNR \leq 1$ , the Rice distribution is far from being Gaussian. In Figure 4.1, the Rice distribution density function is plotted for different values of  $SNR$ .

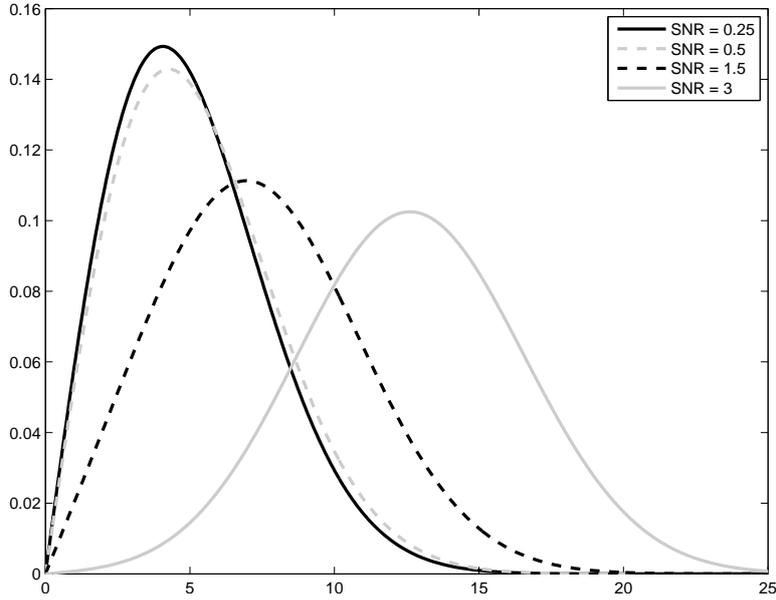


Figure 4.1: The Rice distribution density function for different values of  $SNR = r/\sigma$ .

#### 4.1 ML estimation of the Rice distribution

Given a sample of size  $N$  from the  $Rice(r, \sigma^2)$  distribution,  $Z_1, \dots, Z_N$ , the log-likelihood is given by

$$l(r, \sigma^2) = \sum_{i=1}^N \ln Z_i - N \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^N (Z_i^2 + r^2) + \sum_{i=1}^N \ln I_0 \left( \frac{Z_i r}{\sigma^2} \right)$$

Using the fact that  $I_0'(x) = I_1(x)$  [49], we can show that the maximum likelihood equations are

$$r - \frac{1}{N} \sum_{i=1}^N Z_i R(Z_i r / \sigma^2) = 0$$

where we let  $R(x) = I_1(x)/I_0(x)$  and

$$\sigma^2 - \frac{1}{2N} \sum_{i=1}^N (Z_i^2 - r^2) = 0.$$

Maximum likelihood estimators usually have good properties but unfortunately the ML equations above are not analytically solvable. They have to be solved numerically.

ML estimation of the parameters of the Rice distribution has been given some attention in literature. Meyer [43] looked at the non-central  $\chi$  distribution for  $\sigma^2 = 1$ . He actually looked at the non-central  $\chi^2$  distribution but this does not affect the ML estimates [29]. He showed that there exists a positive, unique solution to the likelihood equation for  $r$  if

$$\frac{1}{N} \sum_{i=0}^N Z_i^2 > 2, \quad (4.2)$$

otherwise  $r = 0$  is the only solution. He also showed that

$$\lim_{N \rightarrow \infty} P \left\{ \frac{1}{N} \sum_{i=0}^N Z_i^2 > 2 \right\} = 1$$

Anderson [2, 3] considered the non-central  $\chi$  distribution for a general  $\sigma^2$ . She provided a generalization of Meyer's condition (4.2)

$$\frac{1}{N} \sum_{i=1}^N Z_i^2 > 2\sigma^2 \quad (4.3)$$

and stated that there exists a unique solution to the ML equations if  $\sigma^2$  is greater than a small positive number. She provided expressions for the variance of, and the correlation between, the estimators of  $r$  and  $\sigma^2$  and the Cramér-Rao lower bounds (CRLB) for the estimates. Talukdar and Lawing [74] compared ML estimations and moment estimations of the  $SNR = r/\sigma$ , and found them equivalent. Sijbers et al. [63, 62] looked at MR Imaging data, and using Mean Square Error (MSE) and CRLB they compared ML estimators of the parameters of the Rice distribution on the absolute value of the data with estimators of the parameters of the Gaussian distribution of the complex valued data. They found that for low  $SNR < 3$ , and for images with non constant shift phase it is better to use the magnitude, Rician distributed, data. Gudbjartsson and Patz [20] showed that for  $SNR \geq 2$  the Rice distribution can be reasonably approximated by the Gaussian distribution, see Figure 4.1. They also argued that for lower  $SNR (< 2)$  it is better to work with the magnitude data than with the complex data.

#### 4.1.1 Consistency of the ML estimators for the Rice distribution

Let  $Z_1, \dots, Z_N$  be a sample of independent Rician distributed random variables with the true parameter pair  $\xi_0 = (r_0, \sigma_0^2)$ , and let  $\widehat{\xi}_N$  be a maximum likelihood estimator of  $\xi_0$ . The following theorem is a special case of a result given in Zhu et al. [80].

**Theorem 3.1:** *Assume that  $\xi_0$  is an interior point of  $\Xi = [a_1, a_2] \times [b_1, b_2]$ , where  $0 < a_1 < a_2 < \infty$  and  $0 < b_1 < b_2 < \infty$ . Then  $\widehat{\xi}_N$  converges in probability to  $\xi_0$  as  $n \rightarrow \infty$ .*

*Proof:* Let

$$M_N(\xi) = \frac{1}{N} \sum_{i=1}^N \log \frac{f_\xi(Z_i)}{f_{\xi_0}(Z_i)}, \quad \text{and} \quad M(\xi) = E \left( \log \frac{f_\xi(Z_i)}{f_{\xi_0}(Z_i)} \right)$$

By Theorem 5.7 in van der Vaart [75] it is sufficient to show for every  $\epsilon > 0$  that

$$\sup_{\xi: \|\xi - \xi_0\| \geq \epsilon} M(\xi) < M(\xi_0), \quad \text{and} \quad \sup_{\xi \in \Xi} |M_N(\xi) - M(\xi)| \xrightarrow{P} 0$$

as  $N \rightarrow \infty$ . The function  $\log f_\xi(z)$  is Lipschitz in  $\xi$ , i.e., there exists a measurable function  $m$  such that

$$\left| \log f_{\xi_1}(z) - \log f_{\xi_2}(z) \right| < m(z) \|\xi_1 - \xi_2\|, \quad \text{for every } \xi_1, \xi_2$$

and  $m$  is integrable. By Example 19.7 in van der Vaart [75] and Theorem 2.4.1 in van der Vaart and Wellner [76], this implies that  $\{\log f_\xi(z): \xi \in \Xi\}$  is Glivenko-Cantelli. Thus

$$\sup_{\xi \in \Xi} |M_N(\xi) - M(\xi)| \xrightarrow{P} 0, \quad \text{as } N \rightarrow \infty.$$

The function  $M(\xi)$  is continuous in  $\xi \in \Xi$ . Hence  $\sup_{\xi: \|\xi - \xi_0\| \geq \epsilon} M(\xi)$  is attained for some  $\xi \in \{\xi: \|\xi - \xi_0\| \geq \epsilon\}$ . Since  $M(\xi) = M(\xi_0)$  if and only if  $\xi = \xi_0$  (Lemma 5.35 in van der Vaart [75]), we have shown that

$$\sup_{\xi: \|\xi - \xi_0\| \geq \epsilon} M(\xi) < M(\xi_0).$$

This completes the proof.  $\square$

## 4.2 Modification to a three parameter Rice distribution

Recall that the non-centrality parameter (3.14) contains information both from  $Re(S)$  and  $Im(S)$ . The two-parameter Rice distribution that we considered above does not take this into account. We would like to modify the ordinary Rice distribution so that besides the variance or scale parameter  $\sigma^2$ , both real and imaginary parts of the original complex difference signal will be presented. We let  $\mu_1 = Re(S)$  and  $\mu_2 = Im(S)$  and suggest a modified Rice distribution with three parameters  $\mu_1$ ,  $\mu_2$  and  $\sigma^2$ . Its probability density function is given by

$$f(z|\mu_1, \mu_2, \sigma^2) = \frac{z}{\sigma^2} e^{-(z^2 + \mu_1^2 + \mu_2^2)/2\sigma^2} I_0 \left( \frac{z \sqrt{\mu_1^2 + \mu_2^2}}{\sigma^2} \right), \quad z \geq 0, \quad (4.4)$$

and given a sample of size  $N$  we obtain the log-likelihood as follows

$$\begin{aligned} l(\mu_1, \mu_2, \sigma^2 | Z_1, \dots, Z_N) &= \sum_{i=1}^N \ln Z_i - N \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^N Z_i^2 \\ &\quad - N \frac{\mu_1^2 + \mu_2^2}{2\sigma^2} + \sum_{i=1}^N \ln I_0 \left( \frac{Z_i \sqrt{\mu_1^2 + \mu_2^2}}{\sigma^2} \right). \end{aligned}$$

Direct optimization of this log-likelihood is not the easiest thing to do. Also, the three parameter Rice distribution is not identifiable since the parameters  $\mu_1$  and  $\mu_2$  are perfectly interchangeable in the term  $\sqrt{\mu_1^2 + \mu_2^2}$ . This means that we can not find consistent ML estimators of the parameters. However, given reasonable initial values for optimization we might get some reasonable results.

In [80], a *Rician regression* model and its related normal models was introduced to characterize noise distributions in various MRI modalities and to develop its associated estimation methods and diagnostic tools. The model is define by assuming that the MRI signal intensity follows a Rice distribution with a non-centrality parameter linked to the covariates of interest (such as the gradient directions and gradient strengths for acquiring diffusion tensor images (DTI)), that is,

$$Z_i \sim Rice(m_i(\beta), \sigma^2) \quad \text{and} \quad m_i = g(x_i, \beta) \quad (4.5)$$

where  $\{(S_i, x_i): i = 1, \dots, N\}$  denote the  $N$  measurements of the MRI intensity,  $S_i$ , and all the covariates,  $x_i$ , at a single voxel,  $\beta$  is a  $p \times 1$  parameter vector in  $\mathbb{R}^p$ , and  $g(\cdot, \cdot)$  is a known link function, which depends on the

particular MR imaging modalities (e.g., anatomical, functional, DTI, and so on). Because the density of the Rice distribution does not belong to the exponential family, the Rician regression model is not a special case of a generalized linear model [41].

Inspired by this model (4.5) for MR imaging data, we consider it for our MR spectroscopy data, by formally taking the link function in (4.5) as  $m_i = \sqrt{\mu_1^2 + \mu_2^2}$  for all  $i = 1, \dots, N$ . However, it should be noted that in model (4.5) the link function is designed for the magnitude and it does not distinguish the sources on which the magnitude is based.

Rician regression is done via an ECM algorithm, a version of the expectation maximization (EM) algorithm with a conditional maximization step, as derived by Meng and Rubin [42], and further developed by Zhu et al. [80]. The standard EM algorithm consists of two steps: the expectation (E) step and the maximization (M) step. The ECM algorithm differs from the EM algorithm in the M-step, where the latter has difficulty to find the solution in the case when the M-step does not have a closed form. The latent variable used in this particular ECM algorithm is a phase variable  $\Theta_i \in (-\pi, \pi)$ , introduced for each observed magnitude spectra  $Z_i$ , such that the joint density of  $(Z_i, \Theta_i)$  is given by

$$f(z_i, \theta_i | m_i(\beta), \sigma^2) = \frac{z_i}{2\pi\sigma^2} \cdot \exp\left(-\frac{m_i(\beta)^2 - 2z_i m_i(\beta) \cos \theta_i}{2\sigma^2}\right).$$

The ECM algorithm works quite well as long as the starting values are close to the true parameter values. However, if the starting values  $\mu_1^{(0)}$  and  $\mu_2^{(0)}$  are switched, the algorithm will give the wrong answer. This is of course due to the fact that the link function  $\sqrt{\mu_1^2 + \mu_2^2}$  is non-identifiable. Also, since the algorithm treats the phase  $\Theta$  as an unobserved variable, and due to the particular choice of link function, it can not leave the angle between the starting values, defined by

$$\psi^{(0)} = \arctan2\left(\frac{\mu_2^{(0)}}{\mu_1^{(0)}}\right),$$

where

$$\arctan2\left(\frac{y}{x}\right) = \begin{cases} \arctan(y/x), & \text{if } x > 0, \\ \arctan(y/x) + \pi, & \text{if } x < 0, y \geq 0, \\ \arctan(y/x) - \pi, & \text{if } x < 0, y < 0, \\ \pi/2, & \text{if } x = 0, y > 0, \\ -\pi/2, & \text{if } x = 0, y < 0, \\ \text{undefined}, & \text{if } x = 0, y = 0, \end{cases} \quad (4.6)$$

is the quadrant-specific inverse of the tangent, see for instance [27]. The ordinary arctan function is usually defined on  $(-\pi/2, \pi/2)$  but arctan2 is defined on the accurate interval  $(-\pi, \pi)$ .

The angle  $\psi^{(0)}$  might be close or far from the true phase between  $\mu_{1_0}$  and  $\mu_{2_0}$ ,  $\psi_0 = \arctan2(\mu_{2_0}/\mu_{1_0})$ , and the ECM can not correct for this error. It tries, but there will always be bias in the estimates. We need to incorporate the phase in our model in some way.

## 5 A Phase-Magnitude model

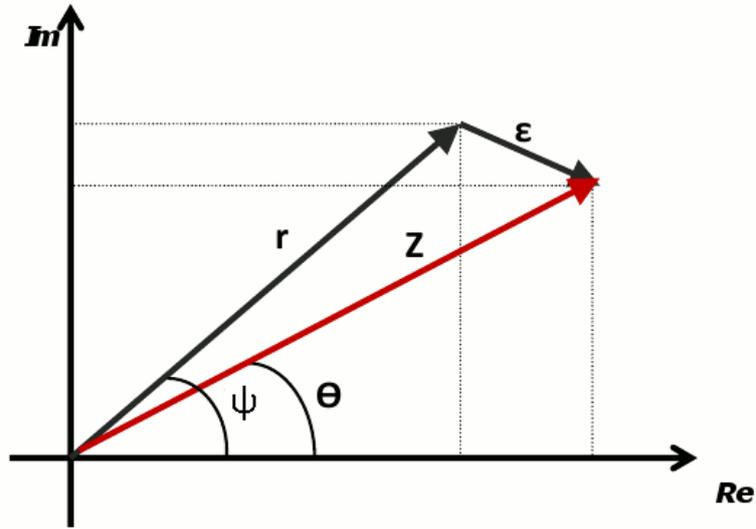


Figure 5.1: Complex valued signal disturbed by additive complex noise.

What happens when a complex-valued signal is disturbed by additive complex noise term? As seen in Figure 5.1, the observed signal magnitude  $Z$  and the observed signal phase  $\Theta$ , differs from the true, unobserved, signal magnitude  $r$  and phase  $\psi$ . That is, given the Gaussian assumptions on the stochastic noise in (3.12) and (3.13), we may suppose that the real and the imaginary noise corrupted channels are

$$\begin{aligned} U = \text{Re}(Y) &= r \cos \psi + \text{Re}(\Upsilon), \\ V = \text{Im}(Y) &= r \sin \psi + \text{Im}(\Upsilon), \end{aligned}$$

where  $\psi$  is the true, and unknown, phase between  $\text{Re}(S)$  and  $\text{Im}(S)$  such that  $\psi = \arctan2\left(\frac{\text{Im}(S)}{\text{Re}(S)}\right)$ . Hence

$$\begin{bmatrix} U \\ V \end{bmatrix} \sim N\left(\begin{bmatrix} r \cos \psi \\ r \sin \psi \end{bmatrix}, \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}\right),$$

and thus

$$f_{U,V}(u, v | r, \psi, \sigma^2) = \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{1}{2\sigma^2} \left((u - r \cos \psi)^2 + (v - r \sin \psi)^2\right)\right\}.$$

The polar transform gives the measured magnitude spectra  $Z = |Y|$ ,  $Z \in \mathbb{R}^+$ , and the measured phase  $\Theta$ ,  $\Theta \in (-\pi, \pi)$ , via

$$\begin{aligned} Z &= \sqrt{U^2 + V^2} \\ \Theta &= \arctan2\left(\frac{V}{U}\right) \end{aligned}$$

The joint density of  $(Z, \Theta)$  is thus

$$f_{Z,\Theta}(z, \theta | r, \psi, \sigma^2) = \frac{z}{2\pi\sigma^2} \cdot \exp\left\{-\frac{1}{2\sigma^2}(z^2 + r^2 - 2rz \cos(\theta - \psi))\right\}. \quad (5.1)$$

Since  $U$  and  $V$  are independent so are  $Z$  and  $\Theta$ .

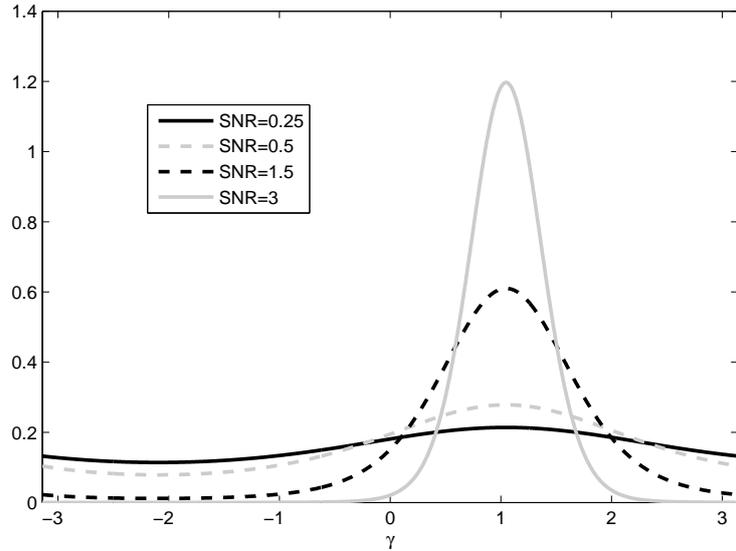


Figure 5.2: The marginal distribution density function of  $\Theta$  for different values of  $SNR = r/\sigma$ .

The marginal distribution of  $Z$  is the Rice distribution (4.1) and the marginal distribution of  $\Theta$  is

$$f_{\Theta}(\theta | r, \psi, \sigma^2) = \frac{1}{2\pi} e^{-r^2/2\sigma^2} \left(1 + \sqrt{\pi}\rho e^{\rho^2} (1 + \text{erf}(\rho))\right) \quad (5.2)$$

where  $\rho = \frac{r}{\sigma\sqrt{2}} \cos(\theta - \psi)$  and

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

is the error function [49]. Figure 5.2 shows examples of the marginal distribution  $f_{\Theta}(\theta|\psi, r, \sigma^2)$  for different  $SNR$ .

When  $r = 0$ , that is when only noise is present, we get that the phase follows a uniform distribution on  $(-\pi, \pi)$ ,  $f_{\Theta}(\theta) = \frac{1}{2\pi}$ .

As  $r$  gets really large the phase error  $\Theta - \psi$  becomes small and more concentrated around 0 and the marginal distribution of the phase may be approximated by  $N(0, \sigma/r)$ .

## 5.1 Phase modeling in MRI

In different types of MR phase imaging techniques the phases between the real and imaginary part of the signals are used to analyze flow in the tissue at hand. Modeling of both magnitude  $Z$  and phase  $\Theta$  in the analysis of MR imaging has been addressed by i.e. Bonny et al. [7], Rowe [57, 58] and Rowe and Logan [59, 60]. Bonny et al. [7] considered ML estimation of  $\sigma^2$  from magnitude measurements in background pixels in the image, where  $r = 0$  (Rayleigh distribution). They estimated  $r$  numerically from the Rice distribution (with  $\sigma^2$  known) and  $\psi$  numerically from the marginal distribution of  $\Theta$ . They found that the bias of their  $r$  and  $\psi$  estimates were decreasing by increasing  $N$  and  $SNR$ , and that the standard deviation of the estimates were decreasing by the factor  $N^{-1/2}$ .

Rowe and Logan [59, 60] and Rowe [57, 58] has done extensive work in phase and magnitude modeling for functional MRI (fMRI) and ML estimation of the parameters. They compared estimations from magnitude only model and from the phase-magnitude model and found that the models were comparable for larger values of the  $SNR$  but that the phase-magnitude model was superior for lower  $SNR$ .

## 5.2 ML estimation in the phase-magnitude model

Given a sample of size  $N$  from the joint distribution of  $(Z, \Theta)$

$$(Z_1, \Theta_1), \dots, (Z_N, \Theta_N) \sim f(z, \theta)$$

the log-likelihood is given by

$$l(r, \psi, \sigma^2) = \sum \ln Z_i - N \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum (Z_i^2 + r^2 - 2rZ_i \cos(\Theta_i - \psi)) \quad (5.3)$$

and the likelihood functions

$$\begin{aligned} \frac{\partial l}{\partial r} &= \frac{1}{\sigma^2} \sum (-r + Z_i \cos(\Theta_i - \psi)), \\ \frac{\partial l}{\partial \psi} &= \frac{1}{\sigma^2} \sum rZ_i \sin(\Theta_i - \psi), \\ \frac{\partial l}{\partial \sigma^2} &= -\frac{N}{\sigma^2} + \frac{1}{2\sigma^4} \sum (Z_i^2 + r^2 - 2rZ_i \cos(\Theta_i - \psi)). \end{aligned}$$

By setting the equations equal to zero we get the following set of solutions

$$\hat{\psi} = \arctan2\left(\frac{\sum Z_i \sin \Theta_i}{\sum Z_i \cos \Theta_i}\right), \quad (5.4)$$

$$\hat{r} = 0 \quad \text{or} \quad \hat{r} = \frac{1}{N} \sum Z_i \cos(\Theta_i - \hat{\psi}), \quad (5.5)$$

$$\hat{\sigma}^2 = \frac{1}{2N} \sum (Z_i^2 + \hat{r}^2 - 2\hat{r}Z_i \cos(\Theta_i - \hat{\psi})). \quad (5.6)$$

Hence, we have derived simple expressions for the estimates of the parameters of the distribution  $f(z, \theta | r, \psi, \sigma^2)$ . Next, we need to verify that these estimates have the desired asymptotic properties such as consistency, asymptotic normality and asymptotic efficiency.

### 5.3 Asymptotic properties

In order to verify the asymptotic properties for the estimates for  $\xi = (r, \psi, \sigma^2)$  in equations (5.4), (5.5) and (5.6), we choose to follow the regularity conditions stated by Lehman [35] for the multiparameter case:

- (A) *The distributions  $f(z, \theta | \xi)$  are distinct.*
- (B) *The distributions  $f(z, \theta | \xi)$  have common support.*
- (C) *The observations are independent and identically distributed (iid) with probability density  $f(z, \theta | \xi)$ .*

- (D) The parameter space  $\Xi$  contains an open subset  $\zeta$  of which the true parameter  $\xi_0$  is an interior point
- (E) There exists an open subset  $\zeta$  of  $\Xi$  containing the true parameter point  $\xi_0$  such that for almost all  $(z, \theta)$  the density  $f(z, \theta|\xi)$  admits all third derivatives  $(\partial^3 / \partial \xi_j \partial \xi_k \partial \xi_l) f(z, \theta|\xi)$  for all  $\xi \in \zeta$
- (F) The first and second logarithmic derivatives of  $f$  satisfy the equations

$$E_\xi \left( \frac{\partial}{\partial \xi_j} \ln f(z, \theta|\xi) \right) = 0, \quad \text{for } j = 1, 2, 3 \quad (5.7)$$

and

$$J_{jk}(\xi) = E_\xi \left( \frac{\partial}{\partial \xi_j} \ln f(z, \theta|\xi) \cdot \frac{\partial}{\partial \xi_k} \ln f(z, \theta|\xi) \right) \quad (5.8)$$

$$= E_\xi \left( -\frac{\partial^2}{\partial \xi_j \partial \xi_k} \ln f(z, \theta|\xi) \right) \quad (5.9)$$

- (G) The  $J_{jk}(\xi)$  are finite and the matrix  $J(\xi)$  is positive definite for all  $\xi \in \zeta$ , and hence the statistics

$$\frac{\partial}{\partial \xi_j} \ln f(z, \theta|\xi) \quad j = 1, 2, 3$$

are affinely independent with probability 1.

- (H) There exist functions  $M_{jkl}$  such that

$$\left| \frac{\partial^3}{\partial \xi_j \partial \xi_k \partial \xi_l} \ln f(z, \theta|\xi) \right| \leq M_{jkl}(z, \theta), \quad \text{for all } \xi \in \zeta \quad (5.10)$$

where

$$m_{jkl} = E_{\xi_0} (M_{jkl}(Z, \Theta)) < \infty, \quad \text{for all } j, k, l \quad (5.11)$$

Then this theorem follows [35]:

**Theorem 4.1:** Let  $(Z_1, \Theta_1), \dots, (Z_N, \Theta_N)$  be iid each with a density  $f(z, \theta|\xi)$  which satisfies assumptions (A)-(H) above. Then with probability 1 as  $N \rightarrow \infty$ , there exists solutions  $\hat{\xi}_N = \hat{\xi}_N(Z_1, \Theta_1, \dots, Z_N, \Theta_N)$  of the likelihood functions such that

- (i)  $\widehat{\xi}_{jN}$  is consistent for estimating  $\xi_j$ ,
- (ii)  $\sqrt{N}(\widehat{\xi}_N - \xi)$  is asymptotically normal with (vector) mean zero and covariance matrix  $J(\xi)^{-1}$ , and
- (iii)  $\widehat{\xi}_{jN}$  is asymptotically efficient in the sense that

$$\sqrt{N}(\widehat{\xi}_{jN} - \xi_j) \rightarrow N(0, J(\xi)_{jj}^{-1})$$

The distributions  $f(z, \theta | r, \psi, \sigma^2)$  are distinct, identifiable with respect to the parameters  $(r, \psi, \sigma^2)$  and have common support. We have assumed that the observations are iid from  $f(z, \theta | r, \psi, \sigma^2)$ , and that we have the parameter space  $\Xi = (0, \infty) \times (-\pi, \pi) \times (0, \infty)$  for  $(r, \psi, \sigma^2)$ .

Hence, assumptions (A)-(D) are satisfied.

The verification of assumption (E) is found in Appendix A. All third derivatives of  $f(z, \theta | r, \psi, \sigma^2)$  are quite easy to derive and exists for all parameters in the parameter space  $\Xi$ .

The verification of assumption (F) is found in Appendix B. The first part of the assumption is easily verified and the second part yield the following Fisher information matrix:

$$J = \begin{pmatrix} 1/\sigma^2 & 0 & 0 \\ 0 & r^2/\sigma^2 & 0 \\ 0 & 0 & 1/\sigma^4 \end{pmatrix} \quad (5.12)$$

Since  $r$  is finite and  $\sigma^2 > 0$ , all elements of  $J$  are finite. The eigenvalues of  $J$  are simply  $(1/\sigma^2, r^2/\sigma^2, 1/\sigma^4)$  and thus all eigenvalues are positive, and  $J$  is positive definite. Assumption (G) is satisfied.

The verification of assumption (H) is found in Appendix C. The existence of functions  $M_{jkl}$  such that  $m_{jkl} = E(M_{jkl}(Z, \Theta)) < \infty$ , for all  $j, k, l$ , is verified.

Consequently the assumptions of Theorem 4.1 are fulfilled and the solutions of the likelihood equations given in (5.4), (5.5) and (5.6) are consistent, asymptotically efficient and asymptotically normal with (vector) mean zero and covariance matrix

$$J^{-1} = \begin{pmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2/r^2 & 0 \\ 0 & 0 & \sigma^4 \end{pmatrix}.$$

## 6 Simulation study

In order to investigate the performance of the estimators  $(\widehat{\psi}, \widehat{r}, \widehat{\sigma}^2)$ , (5.4), (5.5) and (5.6), and their properties a thorough simulation study were performed.

The methods for the study are described in Chapter 6.1, the results are found in Chapter 6.2 and a discussion follows in 6.3.

### 6.1 Methods

#### 6.1.1 Properties of the estimates for different sample sizes and SNR

We let the true value of  $r_0$  mimic the values of the difference of the NAA/Cr ratio in normal tissues and pathologic tissue of in vitro MRS measurements [10], that is  $r_0 = 0.4$ . For a given  $SNR$  the true noise standard deviation is then calculated by  $\sigma_0 = r_0/SNR$ . The true phase  $\psi_0$  is given a randomly chosen value in  $(-\pi, \pi)$ .

For a given true parameter set  $(\psi_0, r_0, \sigma_0^2)$  we look at four different values of the  $SNR$ : 0.25, 0.5, 1.5 and 3. For each  $SNR$  we set the sample size  $N$  to  $2^1$ , then  $2^2$ , and so on, up to  $2^{15}$ .

The observations  $(Z_i, \Theta_i)$ ,  $i = 1, \dots, N$ , were simulated in two steps. First, two separate sequences  $\{Re_i: i = 1, \dots, N\}$  and  $\{Im_i: i = 1, \dots, N\}$ , corresponding to the real and imaginary part respectively, were simulated from the Gaussian distributions  $N(r_0 \cos \psi_0, \sigma_0^2)$  and  $N(r_0 \sin \psi_0, \sigma_0^2)$ , respectively. Then the square roots  $Z_i = \sqrt{Re_i^2 + Im_i^2}$  and the angles  $\Theta_i = \arctan2(Im_i/Re_i)$  were calculated. The estimators  $(\widehat{\psi}, \widehat{r}, \widehat{\sigma}^2)$  were calculated via (5.4), (5.5) and (5.6).

This procedure was repeated 1000 times for each true parameters  $(\psi_0, r_0, \sigma_0^2)$  set, and each value of the  $SNR$  and sample size  $N$ . Thus, we obtain 1000 estimates  $(\widehat{\psi}_k, \widehat{r}_k, \widehat{\sigma}_k^2)$ ,  $k = 1, \dots, 1000$ . The bias of the estimates were calculated by

$$\frac{1}{1000} \sum_{k=1}^{1000} \widehat{\psi}_k - \psi_0, \quad \frac{1}{1000} \sum_{k=1}^{1000} \widehat{r}_k - r_0 \quad \text{and} \quad \frac{1}{1000} \sum_{k=1}^{1000} \widehat{\sigma}_k^2 - \sigma_0^2. \quad (6.1)$$

Also, the variance and the mean square error of  $(\widehat{\psi}_k, \widehat{r}_k, \widehat{\sigma}_k^2)$  were calculated.

#### 6.1.2 Asymptotic normality

In (ii) of Theorem 4.1 we found that  $\sqrt{N}(\widehat{\xi}_N - \xi)$  is asymptotically normal with (vector) mean zero and covariance matrix  $J(\xi)^{-1}$  where the Fisher ma-

trix was given in (5.12). In order to verify this we investigate the asymptotic behaviors of

$$\begin{aligned} \sqrt{N}(\widehat{\psi}_N - \psi), \quad \text{where } \widehat{\psi}_N &= \widehat{\psi}((Z_1, \Theta_1), \dots, (Z_N, \Theta_N)), \\ \sqrt{N}(\widehat{r}_N - r), \quad \text{where } \widehat{r}_N &= \widehat{r}((Z_1, \Theta_1), \dots, (Z_N, \Theta_N)), \\ \sqrt{N}(\widehat{\sigma}_N^2 - \sigma^2), \quad \text{where } \widehat{\sigma}_N^2 &= \widehat{\sigma}^2((Z_1, \Theta_1), \dots, (Z_N, \Theta_N)), \end{aligned}$$

as  $N$  grows.

We have used the same simulated signals as in the previous section, and we will calculate the sample variances of the vectors  $(\widehat{\psi}_k - \psi_0)$ ,  $(\widehat{r}_k - r_0)$  and  $(\widehat{\sigma}_k^2 - \sigma_0^2)$ ,  $k = 1, \dots, 1000$ .

Additionally, for the largest value of the sample size,  $N = 2^{15}$ , the simulations are repeated 100 times, and for each of these times we have used the Jarque-Bera test [28] of the null hypothesis that the vectors  $\sqrt{N}(\widehat{\psi}_k - \psi_0)$ ,  $\sqrt{N}(\widehat{r}_k - r_0)$  and  $\sqrt{N}(\widehat{\sigma}_k^2 - \sigma_0^2)$ ,  $k = 1, \dots, 1000$ , are Gaussian distributed, against the alternative that they are not. The test returns the value 1 if the null hypothesis can be rejected at the 5% significance level and 0 if it can not.

### 6.1.3 Properties of the estimates for different true value of the phase $\psi$

The phase parameter  $\psi$  is defined on  $(-\pi, \pi)$ , that is almost the whole circle with radius  $r$ . But what happens when the parameter gets close to the boundaries of the parameter space? Due to how the estimator  $\widehat{\psi}$  is defined we would expect to get some problems with the estimates. In order to investigate what those problems are, we conduct simulations as follows:

We fix the true value of  $r$  to  $r_0 = 0.4$  and chose to work with one sample sizes, namely  $N = 2^5$ . We choose four different values of the  $SNR$ : 0.25, 0.5, 1.5 and 3, and for a given  $SNR$  the true noise standard deviation is calculated by  $\sigma_0 = r_0/SNR$ . The true phase  $\psi_0$  is let to vary in  $(-\pi, \pi)$ .

The observations  $(Z_i, \Theta_i)$ ,  $i = 1, \dots, N$ , were simulated in the same way as above. The estimators  $(\widehat{\psi}, \widehat{r}, \widehat{\sigma}^2)$  were calculated via (5.4), (5.5) and (5.6), and the procedure was then repeated until 1000 replicates of the estimates were obtained. The estimated bias, variance and MSE were calculated as described above.

### 6.1.4 Comparison with estimates for the Rice distribution

We would like to compare the ML estimators  $\widehat{r}$  and  $\widehat{\sigma}^2$  from the joint distribution of the phase-magnitude model  $f(z, \theta|r, \psi, \sigma^2)$  with the ML estimators for the marginal distribution of the magnitude, ie  $Rice(z|r, \sigma^2)$ .

We use the same simulation setup as above, with  $r_0 = 0.4$ ,  $\psi_0$  given a randomly chosen value in  $(-\pi, \pi)$ , and with  $\sigma_0 = r_0/SNR$  for a given  $SNR$ . The sample size is let to vary from  $2^1$  up to  $2^{12}$ . In order to keep the computational time down, we use a slightly lower highest value of  $N$  here. 1000 replicates of  $(Z_i, \Theta_i)$ ,  $i = 1, \dots, N$ , are simulated as before. The log likelihood of the Rice distribution for  $Z_i$  is optimized with the ordinary simplex algorithm. The estimators  $(\widehat{r}_k, \widehat{\sigma}_k^2)$ ,  $k = 1, \dots, 1000$ , are calculated. Bias and MSE of the estimates are calculated.

### 6.1.5 Comparison with $\psi$ estimates from the phase marginal distribution

We would like to compare the ML estimator  $\widehat{\psi}$  from the joint distribution  $f(z, \theta|r, \psi, \sigma^2)$  with the ML estimators for the marginal distribution of the phase, ie  $f(\theta|\psi, r, \sigma^2)$ .

We use the exact same setup as in the previous section. The log likelihood of the marginal distribution of the phase  $f(\theta)$  is optimized with the ordinary simplex algorithm.

## 6.2 Results

### 6.2.1 Properties for different sample sizes and SNR

Figure 6.1 shows a typical case of the behavior of the estimated biases (6.1) of  $(\widehat{\psi}_k, \widehat{r}_k, \widehat{\sigma}_k^2)$ ,  $k = 1, \dots, 1000$ . For the two lower value of the  $SNR$ , 0.25 and 0.5, there is considerable bias for sample sizes  $N \leq 5$ . However, the biases are decreasing as  $N$  is increasing, which supports our conclusion of asymptotic unbiased estimators from Chapter 5.3. For the higher values of the  $SNR$ , 1.5 and 3, the estimated biases are quite insignificant.

Figure 6.2 shows the estimated variance from the same simulation as in Figure 6.1. The estimated variances are high for  $SNR = 0.25$  and  $SNR = 0.5$  and the lower sample sizes, but they are decreasing as  $N$  increases. For  $SNR \geq 1.5$  the estimated variances are small.

The estimated mean square errors are dominated by the estimated variances, and hence follows the same pattern as they do. That is, high MSE for the lower values of the  $SNR$  and smaller  $N$ . As  $N$  increases, the MSE decreases. This supports our conclusions of asymptotic consistent estimators from Chapter 5.3.

### 6.2.2 Asymptotic normality

For  $SNR = 0.25$  Figure 6.4 shows the sample variance times  $N$  for the vectors  $(\widehat{\psi}_k - \psi_0)$ ,  $(\widehat{r}_k - r_0)$  and  $(\widehat{\sigma}_k^2 - \sigma_0^2)$ ,  $k = 1, \dots, 1000$ , and the values of the inverse of the Fisher matrix  $J^{-1}$  as given by our chosen true parameter

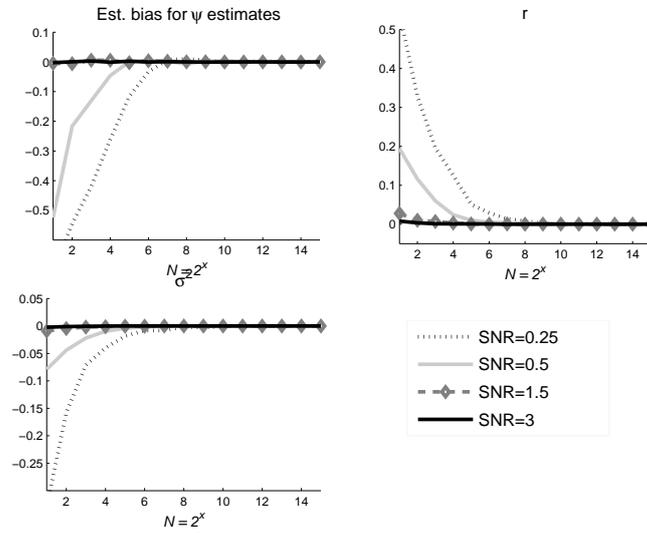


Figure 6.1: Upper left: Estimated bias for  $\hat{\psi}$ . Upper right: Estimated bias for  $\hat{r}$ . Lower left: Estimated bias for  $\hat{\sigma}^2$ .

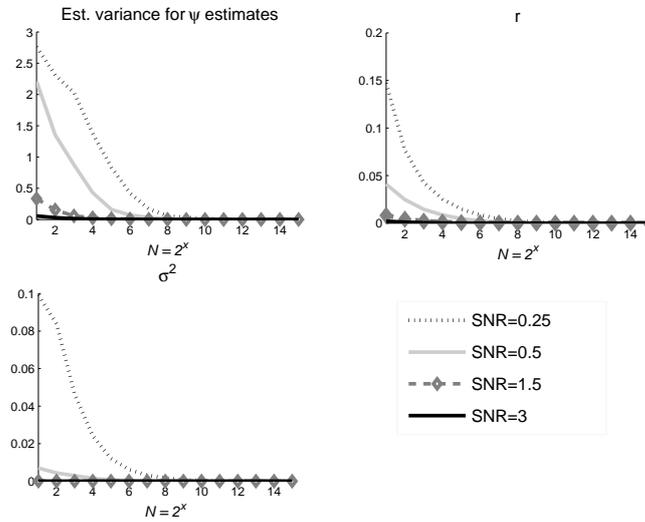


Figure 6.2: Upper left: Estimated variance for  $\hat{\psi}$ . Upper right: Estimated variance for  $\hat{r}$ . Lower left: Estimated variance for  $\hat{\sigma}^2$ .

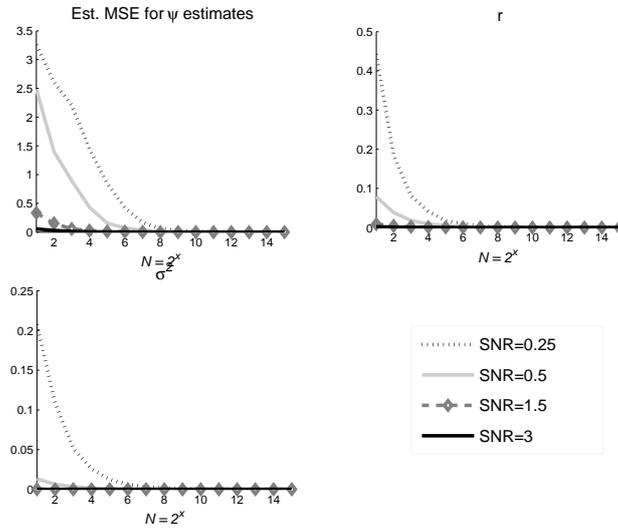


Figure 6.3: Upper left: Estimated MSE for  $\hat{\psi}$ . Upper right: Estimated MSE for  $\hat{r}$ . Lower left: Estimated MSE for  $\hat{\sigma}^2$ .

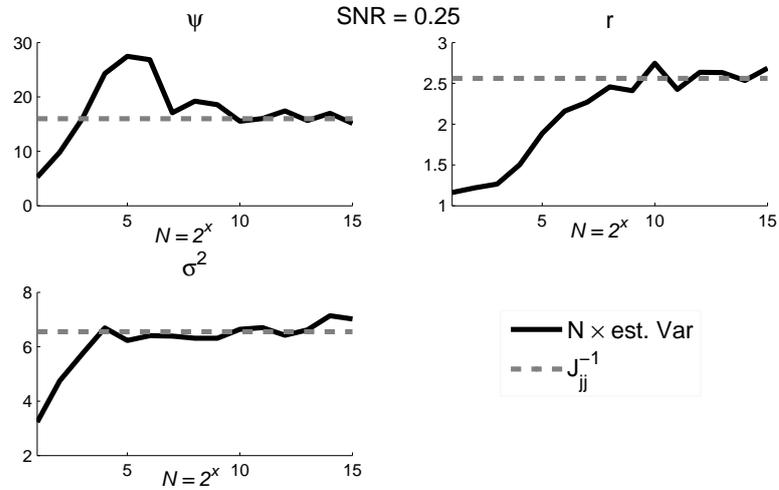


Figure 6.4:  $SNR = 0.25$ . Upper left:  $N$  times the sample variance of  $(\hat{\psi}_k - \psi_0)$ ,  $k = 1, \dots, 1000$  (solid line) and  $J_{11}^{-1}$  (dashed line). Upper right:  $N$  times the sample variance of  $(\hat{r}_k - r_0)$ ,  $k = 1, \dots, 1000$  (solid line) and  $J_{22}^{-1}$  (dashed line). Lower left:  $N$  times the sample variance of  $(\hat{\sigma}_k^2 - \sigma_0^2)$ ,  $k = 1, \dots, 1000$  (solid line) and  $J_{33}^{-1}$  (dashed line).

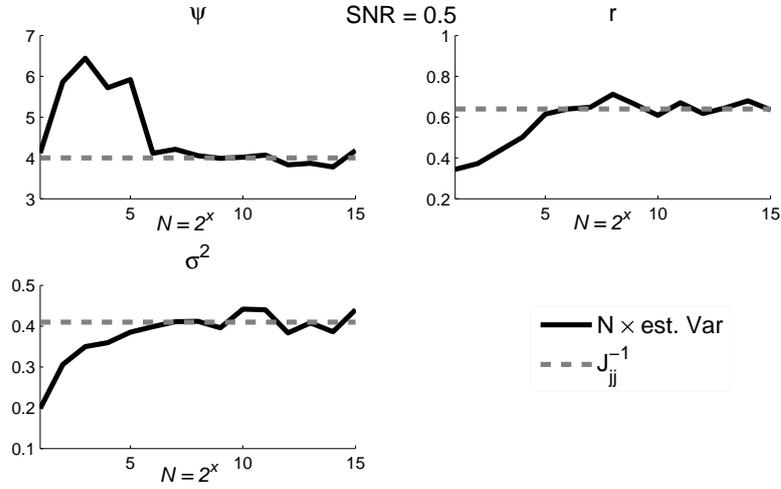


Figure 6.5:  $SNR = 0.5$ . Upper left:  $N$  times the sample variance of  $(\hat{\psi}_k - \psi_0)$ ,  $k = 1, \dots, 1000$  (solid line) and  $J_{11}^{-1}$  (dashed line). Upper right:  $N$  times the sample variance of  $(\hat{r}_k - r_0)$ ,  $k = 1, \dots, 1000$  (solid line) and  $J_{22}^{-1}$  (dashed line). Lower left:  $N$  times the sample variance of  $(\hat{\sigma}_k^2 - \sigma_0^2)$ ,  $k = 1, \dots, 1000$  (solid line) and  $J_{33}^{-1}$  (dashed line).

values. We can see that  $N$  times the sample variances does indeed approach the diagonal elements of  $J^{-1}$  as  $N$  grows. Figures 6.5 and 6.6 shows the same results for  $SNR = 0.5$  and  $SNR = 1.5$ , respectively. For  $SNR = 1.5$  we see that the convergence of  $N$  times the sample variances towards the diagonal elements of  $J^{-1}$  is quite faster than it was for the two lower values of the  $SNR$ .

Figure 6.7 shows the results from the 100 Jarque-Bera tests that was performed on vectors  $\sqrt{N}(\hat{\psi}_k - \psi_0)$ ,  $\sqrt{N}(\hat{r}_k - r_0)$  and  $\sqrt{N}(\hat{\sigma}_k^2 - \sigma_0^2)$ ,  $k = 1, \dots, 1000$ , for  $N = 2^{15}$  and  $SNR = 0.25$ . The test returns the value 1 if the null hypothesis can be rejected at the 5% significance level and 0 if it can not. We can note that only about 5% of the null hypotheses are reject. The assumption that the vectors are Gaussian does seem to be verified.

Thus, this verifies that the estimators  $(\hat{\psi}, \hat{r}, \hat{\sigma}^2)$  are asymptotically Gaussian distributed.

### 6.2.3 Properties for different true value of the phase $\psi$

In Figure 6.8 we see the resulting estimated biases from the simulations. Note that only the estimated bias of the estimator  $\hat{\psi}$  is affected when  $\psi_0$  gets

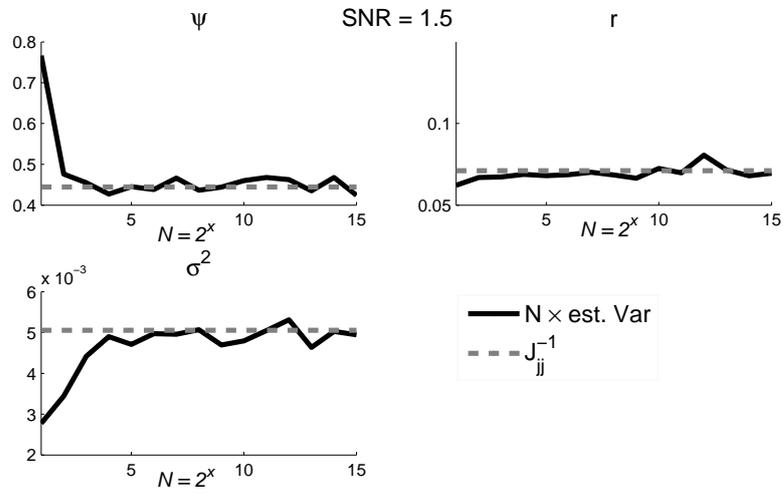


Figure 6.6:  $SNR = 1.5$ . Upper left:  $N$  times the sample variance of  $(\hat{\psi}_k - \psi_0)$ ,  $k = 1, \dots, 1000$  (solid line) and  $J_{11}^{-1}$  (dashed line). Upper right:  $N$  times the sample variance of  $(\hat{r}_k - r_0)$ ,  $k = 1, \dots, 1000$  (solid line) and  $J_{22}^{-1}$  (dashed line). Lower left:  $N$  times the sample variance of  $(\hat{\sigma}_k^2 - \sigma_0^2)$ ,  $k = 1, \dots, 1000$  (solid line) and  $J_{33}^{-1}$  (dashed line).

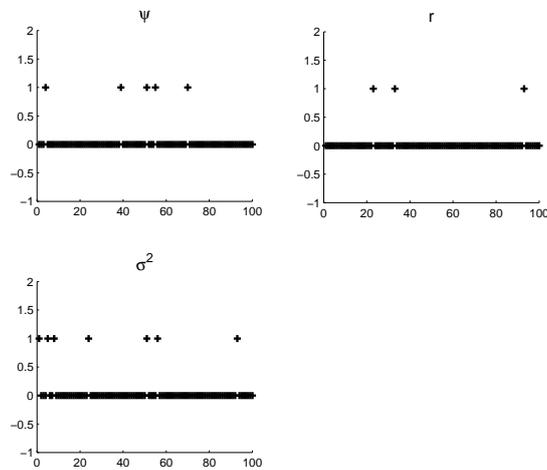


Figure 6.7: Results from 100 Jarque-Bera tests for  $N = 2^{15}$  and  $SNR = 0.25$ .

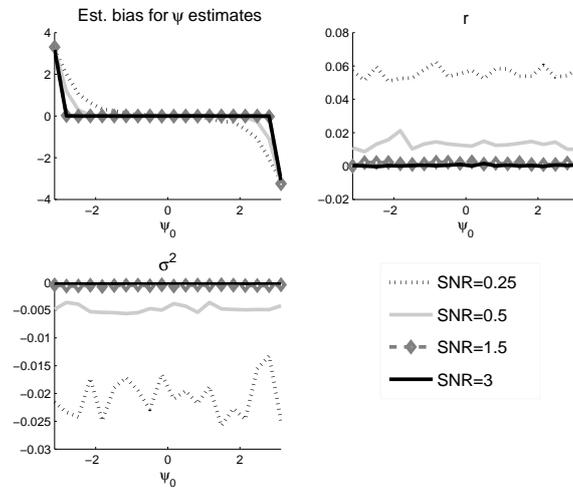


Figure 6.8: Sample size  $N = 2^5$ . Upper left: Bias of the mean of 1000  $\hat{\psi}$  estimates when the true  $\psi$  varies from  $-\pi$  to  $\pi$  and for different SNR. Upper right: Bias of the mean of 1000  $\hat{r}$  estimates. Lower left: Bias of the mean of 1000  $\hat{\sigma}^2$  estimates.

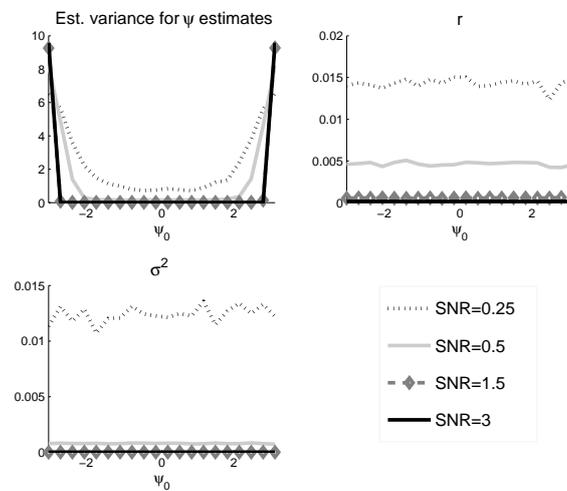


Figure 6.9: Sample size  $N = 2^5$ . Upper left: Variance of 1000  $\hat{\psi}$  estimates when the true  $\psi$  varies from  $-\pi$  to  $\pi$  and for different SNR. Upper right: Variance of 1000  $\hat{r}$  estimates. Lower left: Variance of 1000  $\hat{\sigma}^2$  estimates.

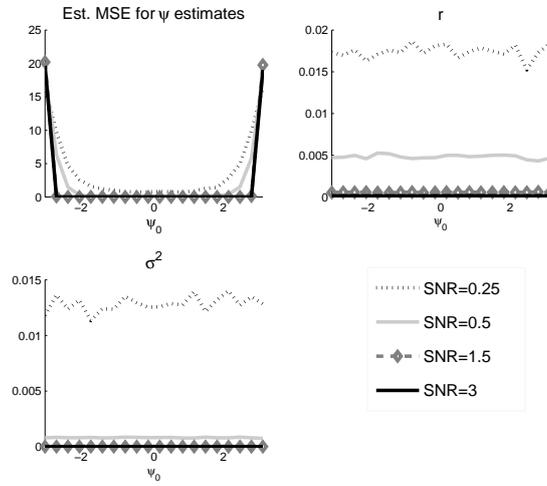


Figure 6.10: Sample size  $N = 2^5$ . Upper left: MSE of 1000  $\hat{\psi}$  estimates when the true  $\psi$  varies from  $-\pi$  to  $\pi$  and for different SNR. Upper right: MSE of 1000  $\hat{r}$  estimates. Lower left: MSE of 1000  $\hat{\sigma}^2$  estimates.

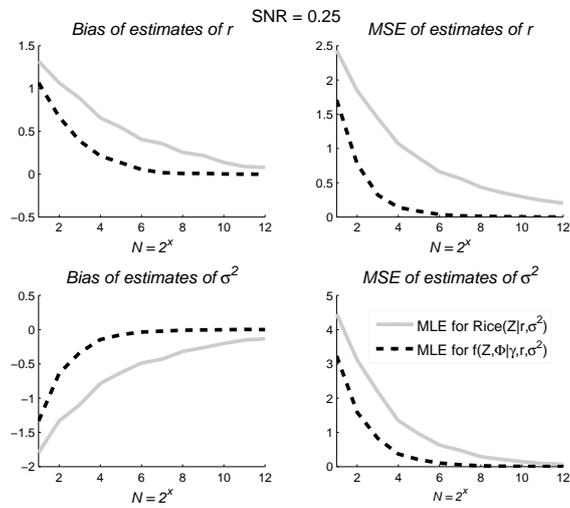


Figure 6.11:  $SNR = 0.25$ . Upper left: Estimated bias of ML estimates of  $r$  from  $Rice(Z)$  (dotted lined) and from  $f(Z, \Theta)$  (solid grey line). Upper right: MSE of the  $r$  estimates. Lower left: Estimated bias of ML estimates of  $\sigma^2$ . Lower right: MSE of the  $\sigma^2$  estimates.

close to  $-\pi$  and  $\pi$ . For  $SNR \geq 1.5$  only the actual boundary is problematic.

Figures 6.9 and 6.10 shows the estimated variance and MSE, respectively. Again, only the properties of the  $\hat{\psi}$  estimator when  $SNR = 0.25$  and  $SNR = 0.5$  are affected when  $\psi_0$  gets close to  $-\pi$  and  $\pi$

#### 6.2.4 Comparison with estimates for the Rice distribution

Figure 6.11 shows the resulting estimated bias and mean square errors when  $SNR = 0.25$ . We see that the estimates from  $f(z, \theta)$  have lower bias and lower MSE. For this value of the  $SNR$  the estimators  $\hat{r}$  and  $\hat{\sigma}^2$  are more efficient than those obtain from the Rice distribution. For  $SNR = 1.5$  the biases of the two estimation methods approach each other. The MSE for the estimates of the magnitude phase model is still slightly smaller for the smaller sample sizes.

#### 6.2.5 Comparison with $\psi$ estimates from the phase marginal distribution

Figure 6.13 shows the resulting estimated bias and MSE for  $SNR = 0.25$  and  $SNR = 3$ . We see that the estimates have similar bias. However the MSE is slightly lower for the lower  $N$  for the  $\hat{\psi}$  estimates than for the estimates of  $\psi$  from the marginal distribution of the phase.

### 6.3 Discussion

Our simulations have shown that the bias, variance and mean square error of the estimates  $(\hat{\psi}, \hat{r}, \hat{\sigma}^2)$  are decreasing as the sample size  $N$  or the  $SNR$  is increasing. This indicates that the estimators are asymptotically unbiased and consistent, which supports our conclusions from Chapter 5.3. We have seen that the estimators are asymptotically Gaussian distributed.

For  $SNR < 1$  and  $N \leq 2^5$  we have considerable bias for the estimates. In practice this might be a problem since clinical MRS data rarely are large.

If the true value of  $\psi$ ,  $\psi_0$  is close to  $-\pi$  or  $\pi$ , the  $\hat{\psi}$  estimator will have some issues with increasing bias, variance and MSE. This is probably due to the circular nature of the  $\psi$  parameter. We can note that, for a small value  $\epsilon$ ,  $-\pi + \epsilon$  and  $\pi - \epsilon$  are close to each other, on the circle. The estimators  $\hat{r}$  and  $\hat{\sigma}^2$  are not affected when  $\psi_0$  gets close to  $\pi$  or  $-\pi$ .

For low values of the  $SNR$  it is better and more efficient to estimate  $(r, \sigma^2)$  from the joint distribution of the magnitude and phase  $f(z, \theta)$  than to estimate them from the marginal distribution of the magnitude, ie the Rice distribution. The biases and mean square errors are smaller. Even for  $SNR = 3$  the mean square error is smaller for the estimates from the phase-magnitude distribution  $f(z, \theta)$ . Estimators of  $\psi$  from the marginal

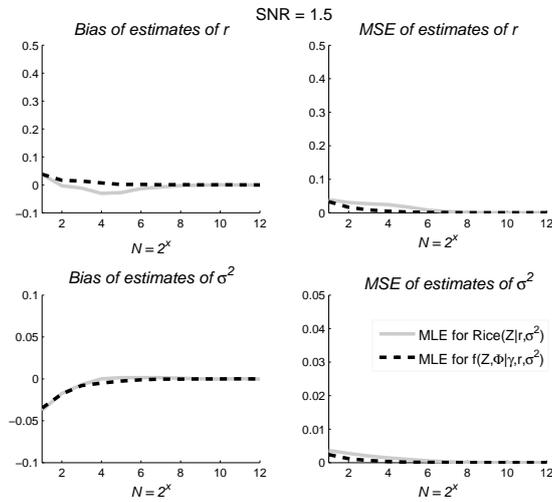


Figure 6.12:  $SNR = 1.5$ . Upper left: Estimated bias of ML estimates of  $r$  from  $Rice(Z)$  (dotted lined) and from  $f(Z, \Theta)$  (solid grey line). Upper right: MSE of the  $r$  estimates. Lower left: Estimated bias of ML estimates of  $\sigma^2$ . Lower right: MSE of the  $\sigma^2$  estimates.

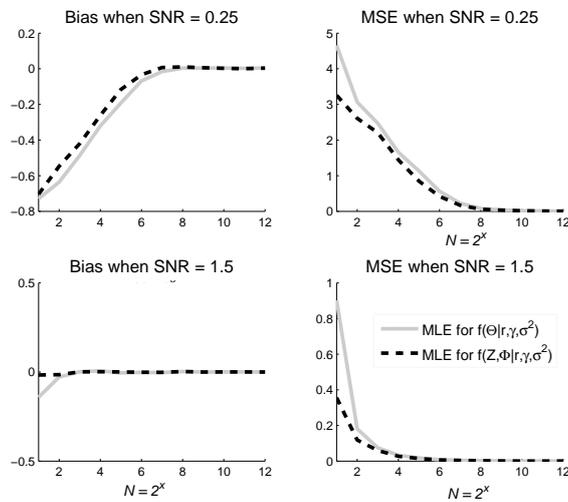


Figure 6.13: Upper left: Estimated bias of ML estimates of  $\psi$  from  $f(\Theta)$  (dotted lined) and from  $f(Z, \Theta)$  (solid grey line) when  $SNR = 0.25$ . Upper right: MSE when  $SNR = 0.25$ . Lower left: Estimated bias when  $SNR = 1.5$ . Lower right: MSE when  $SNR = 1.5$ .

distribution of the phase  $f(\theta)$  and from the phase-magnitude distribution  $f(z, \theta)$  have equivalent bias properties, but the estimators from the joint distribution  $f(z, \theta)$  have smaller mean square error.

Also, estimation from the phase-magnitude distribution  $f(z, \theta)$  does not involve any numerical optimization, unlike the estimation from the Rice distribution or from the marginal distribution of the phase. We have explicit expressions for the solutions of the likelihood equations and for the Fisher matrix. This does not exist for neither the Rice distribution, nor the marginal distribution of the phase.

## 7 Phantom data study

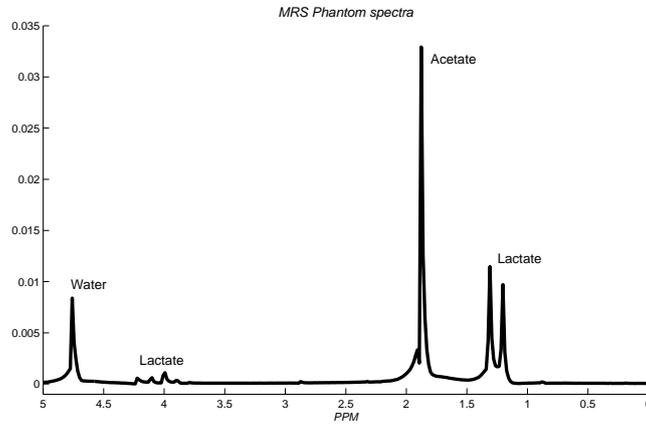


Figure 7.1: Spectra from MRS phantom.

We have 512 MRS scans from an spherical  $^1\text{H}$  MRS phantom, taken on a  $1.5\text{T}$  Siemens MR unit. The phantom is filled with an  $0.1\text{ M}$  solution of lithium lactate ( $\text{C}_3\text{H}_5\text{O}_3\text{Li}$ ) and sodium acetate ( $\text{C}_2\text{H}_3\text{O}_2\text{Na}$ ). The mean spectra of the scans can be seen in Figure 7.1. The double peaks around  $ppm = 1.2$  arises from the protons of the methyl group ( $\text{CH}_3$ ) of the lactate molecule. The quartet peaks around  $ppm = 4.1$  arise from the protons of the methylene group ( $\text{CH}$ ) of the lactate molecule. Ideally, there should be only one big peak at  $ppm = 1.9$ , from the methyl group of the acetate molecule. The smaller peak on the left side is an artefact, possibly due to inhomogeneity of the magnetic field.

The concentrations of the lactate and the acetate are the same and so the sums of the height of all the lactate peaks should be similar to the height of the acetate peak. Can our estimations confirm this statement?

### 7.1 Method

The first thing we would like to verify is whether or not the background noise in the signal is white. Unfortunately, as we can see from Figure 7.2, we do not have any region of the measurement without signal. This makes verification of the white noise assumption quite hard. We will however return to this matter in Chapter 8. For now, we will assume that the noise is white and carry on with estimation of the peak heights.

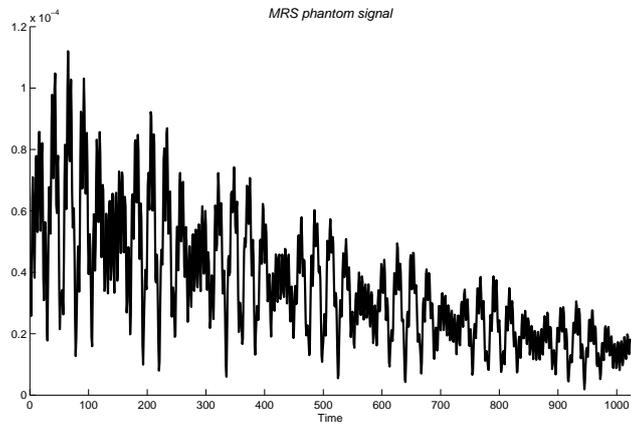


Figure 7.2: MRS signal from MRS phantom.

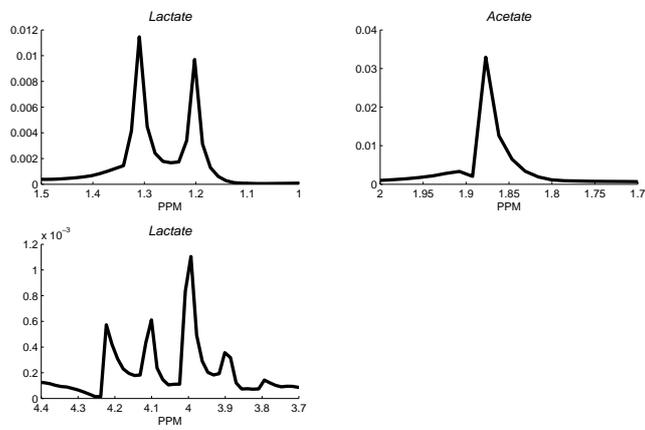


Figure 7.3: Upper left: The lactate doublet. Upper right: The acetate peak. Lower left: The lactate quartet.

There are in total seven peaks, for which we wish to estimate the heights, one peak corresponding to the acetate and seven peaks corresponding to the lactate, see Figure 7.3. We let  $Z_j$ ,  $j = 1, \dots, 7$  denote the observed peak height of peak number  $j$ , and let the numbering follow the chemical shift scale (ppm). That is, the lactate doublets at about  $ppm = 1.2$  and  $ppm = 1.3$ , will have the observed peak heights  $Z_1$  and  $Z_3$ , respectively, and so on for the next five peaks of interest in the spectrum.

Obtaining the peak heights measurements from the 512 different scans is not trivial. It can in fact be seen as quantitation of each separate scan, see Chapter 3.1. However, we know the approximate positions of the peaks and we will use this information to measure the peak heights. We will do this in two ways.

Method I: Identify the exact position of each of the seven peaks in the spectra of the mean of the scans, and then use the values on these positions in each of the separate 512 scans as measurements for the peak heights.

Method II: Use seven small intervals around the expected peak positions, and then use the maximum value in each interval as measurements.

The positions of the peaks in this phantom spectra is quite fixed within the separate scans, at least for the larger peaks. For the smaller peaks in the lactate quartet we have much lower  $SNR$  and the visual interpretation of the peak positions may shift between the different scans. The signals are digitalized and the peaks in the spectra are naturally not smooth, see Figure 7.3

For each of the seven peaks we obtain 512 complex, independent observations. We calculate the observed magnitudes and phases  $(Z_{j_i}, \theta_{j_i})$ ,  $i = 1, \dots, 512$ ,  $j = 1, \dots, 7$ , which are then used to calculate the estimates  $\widehat{\psi}_j$  and  $\widehat{r}_j$ .

In order to further investigate the behavior of our estimates, we use the ordinary Bootstrap method to resample 10000 replicates of the observations of  $(Z_{j_i}, \theta_{j_i})$ . The Bootstrap estimates and confidence intervals are calculated.

## 7.2 Results

### 7.2.1 Method I

The values of  $\widehat{\psi}_j$  and  $\widehat{r}_j$  for  $j = 1, \dots, 7$  are found in the third column in Table 7.1, the Bootstrap estimates in the fourth column and the Bootstrap confidence intervals in the fifth column. We see little difference in the values in the third and fourth column.

We now sum the peak heights estimates for the lactate peaks and get that  $\hat{r}_1 + \hat{r}_2 + \hat{r}_4 + \hat{r}_5 + \hat{r}_6 + \hat{r}_7 = 0.02381$ . For the acetate peak we have  $\hat{r}_3 = 0.03292$ . These two values are not quite equal, but not completely far off from each other either.

### 7.2.2 Method II

The values of  $\hat{\psi}_j$  and  $\hat{r}_j$  for  $j = 1, \dots, 7$  are found in the third column in table 7.2, the Bootstrap estimates in the fourth column and the Bootstrap confidence intervals in the fifth column.

For Method II we have that  $\hat{r}_1 + \hat{r}_2 + \hat{r}_4 + \hat{r}_5 + \hat{r}_6 + \hat{r}_7 = 9.8855 \cdot 10^{-3}$  for the lactate peaks and  $\hat{r}_3 = 3.6247 \cdot 10^{-5}$  for the acetate peak. The two values are not close to each other.

## 7.3 Discussion

The first method of obtaining the measurements, Method I, gives us quite reasonable results. The phantom spectra has little noise and sharp peaks that are easy to label and localize. The estimated peak heights of the lactate and acetate match quite well, specially if we take the variance estimations from the Bootstrap simulations into account.

Method II, on the other hand, gives us unreasonable results. This is clearly not the way to go to obtain the peak heights measurements.

Chemical shift	Estimator	Estimated value	Bootstrap estimate	Bootstrap CI
1.2	$\widehat{\psi}_1$	-0.9509	-0.9509	$[-0.9587, -0.9428]$
	$\widehat{r}_1$	0.0097	0.0097	$[9.6831, 9.7251] \cdot 10^{-3}$
1.3	$\widehat{\psi}_2$	2.7103	2.7103	$[2.7029, 2.7181]$
	$\widehat{r}_2$	$1.1467 \cdot 10^{-2}$	$1.1467 \cdot 10^{-2}$	$[1.1435, 1.1500] \cdot 10^{-2}$
1.85	$\widehat{\psi}_3$	-0.3126	-0.3126	$[-0.3214, -0.3036]$
	$\widehat{r}_3$	0.0329	0.0329	$[3.2887, 3.2962] \cdot 10^{-2}$
3.9	$\widehat{\psi}_4$	-2.4834	-2.4835	$[-2.4971, -2.4694]$
	$\widehat{r}_4$	$3.5639 \cdot 10^{-4}$	$3.5634 \cdot 10^{-4}$	$[3.5088, 3.6190] \cdot 10^{-4}$
4.0	$\widehat{\psi}_5$	-2.7780	-2.7780	$[-2.7849, -2.7705]$
	$\widehat{r}_5$	0.0011	0.0011	$[1.0984, 1.1104] \cdot 10^{-3}$
4.1	$\widehat{\psi}_6$	1.1890	1.1890	$[1.1768, 1.2017]$
	$\widehat{r}_6$	$6.1136 \cdot 10^{-4}$	$6.1135 \cdot 10^{-4}$	$[6.0604, 6.1656] \cdot 10^{-4}$
4.2	$\widehat{\psi}_7$	1.4558	1.4558	$[1.4418, 1.4697]$
	$\widehat{r}_7$	$5.7281 \cdot 10^{-4}$	$5.7284 \cdot 10^{-4}$	$[5.6743, 5.7816] \cdot 10^{-4}$

Table 7.1: Estimates of  $\psi$  and  $r$  for the seven peaks in the MRS phantom using Method I.

Chemical shift	Estimator	Estimated value	Bootstrap estimate	Bootstrap CI
1.2	$\widehat{\psi}_1$	-0.9509	-0.9508	$[-0.9589, -0.9430]$
	$\widehat{r}_1$	$9.7039 \cdot 10^{-3}$	$9.7040 \cdot 10^{-3}$	$[9.6831, 9.7255] \cdot 10^{-3}$
1.3	$\widehat{\psi}_2$	2.3006	2.2996	$[2.1640, 2.4414]$
	$\widehat{r}_2$	$3.6394 \cdot 10^{-5}$	$3.6511 \cdot 10^{-5}$	$[3.1322, 4.1052] \cdot 10^{-5}$
1.85	$\widehat{\psi}_3$	2.1390	2.1398	$[2.0027, 2.2750]$
	$\widehat{r}_3$	$3.6247 \cdot 10^{-5}$	$3.6370 \cdot 10^{-5}$	$[3.1540, 4.0836] \cdot 10^{-5}$
3.9	$\widehat{\psi}_4$	2.0436	2.0429	$[1.9017, 2.1854]$
	$\widehat{r}_4$	$3.6214 \cdot 10^{-5}$	$3.6314 \cdot 10^{-5}$	$[3.1372, 4.0667] \cdot 10^{-5}$
4.0	$\widehat{\psi}_5$	2.2434	2.2425	$[2.1105, 2.3824]$
	$\widehat{r}_5$	$3.7207 \cdot 10^{-5}$	$3.7311 \cdot 10^{-5}$	$[3.2276, 4.1774] \cdot 10^{-5}$
4.1	$\widehat{\psi}_6$	2.0121	2.0113	$[1.8822, 2.1442]$
	$\widehat{r}_6$	$3.6323 \cdot 10^{-5}$	$3.6405 \cdot 10^{-5}$	$[3.1507, 4.1054] \cdot 10^{-5}$
4.2	$\widehat{\psi}_7$	2.1117	2.1133	$[1.9760, 2.2436]$
	$\widehat{r}_7$	$3.5418 \cdot 10^{-5}$	$3.5489 \cdot 10^{-5}$	$[3.0864, 3.9909] \cdot 10^{-5}$

Table 7.2: Estimates of  $\psi$  and  $r$  for the seven peaks in the MRS phantom spectra using Method II.

## 8 Human brain data study

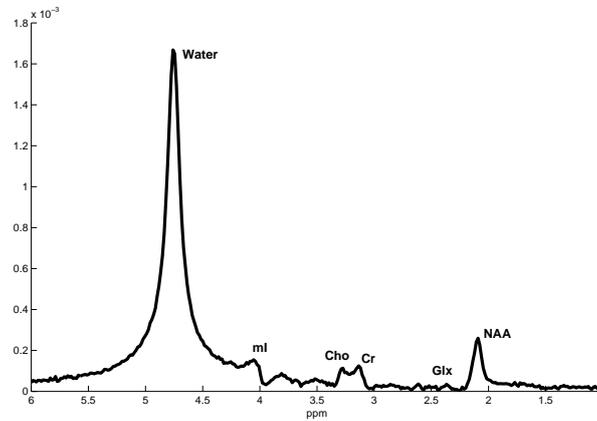


Figure 8.1: *in vivo* MRS spectra from healthy brain.

We have 128 MRS brain scans from a healthy volunteer, taken on a 1.5T Siemens MR unit, Figure 8.1. There are a great number of peaks present in the spectrum. We chose to focus our analysis of one of the metabolic ratios that are of interest for tumor diagnosis, namely the NAA/cr ratio. We would like to estimate this ratio in our spectra. By visual interpretation it appears to be around 2, which would concur with the value of this ratio in non-pathological grey tissue [10]. Can our estimations confirm this statement?

### 8.1 Method

First, we will verify the white noise assumption. As seen in Figure 8.2, we do have regions without signal in our scans, and hence we can investigate the distribution of the noise. We use the Jarque-Bera test [28] of the null hypothesis that the tails of the 128 scanned signals are Gaussian distributed with unknown mean and variance, against the alternative that they are not. The test returns the value 1 if the null hypothesis can be rejected at the 5% significance level and 0 if it can not.

Peak height measurements for the NAA peak,  $Z_1$ , and the Cr peak,  $Z_2$ , are obtained by Method I, as described in Chapter 7.1. We calculate the observed magnitudes and phases  $(Z_{j_i}, \theta_{j_i})$ ,  $i = 1, \dots, 128$ ,  $j = 1, 2$ , which are then used to calculate the estimates  $\widehat{\psi}_j$  and  $\widehat{r}_j$ . In order to further

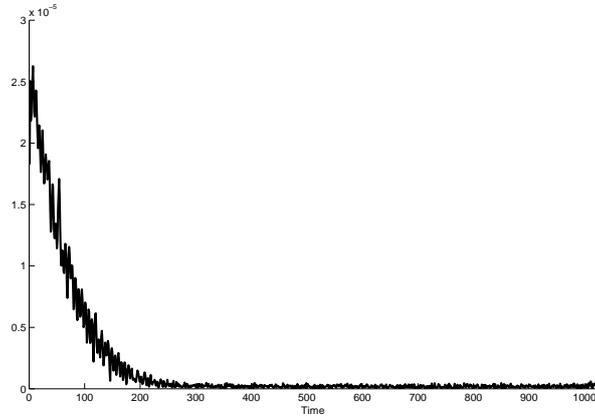


Figure 8.2: MRS signal from healthy brain.

investigate the behavior of our estimates, we use the ordinary Bootstrap method to resample 10000 replicates of the observations of  $(Z_{j_i}, \theta_{j_i})$ . The Bootstrap estimates and confidence intervals are calculated.

## 8.2 Results

The results from the Jarque-Bera test of tails of the 128 scanned signals are shown in Figure 8.3. Only about 5% of the null hypotheses are reject. The assumption that the background noise is white does seem to be verified.

	Estimator	Estimated value	Bootstrap estimate	Bootstrap CI
NAA	$\hat{\psi}$	1.0670	1.06748	[1.0224, 1.1117]
	$\hat{r}$	$2.5944 \cdot 10^{-4}$	$2.5952 \cdot 10^{-4}$	$[2.4889, 2.6984] \cdot 10^{-4}$
Cr	$\hat{\psi}$	-2.9945	-2.9932	[-3.0764, -2.9116]
	$\hat{r}$	$1.2470 \cdot 10^{-4}$	$1.2479 \cdot 10^{-4}$	$[1.1483, 13451] \cdot 10^{-4}$

Table 8.1: Estimates of  $\psi$  and  $r$  for the NAA peak and the Cr peak.

The values of  $\hat{\psi}_j$  and  $\hat{r}_j$  for  $j = 1, \dots, 2$  are found in the third column in table 8.1, the Bootstrap estimates in the fourth column and the Bootstrap confidence intervals in the fifth column. Again, there is little difference in the values in the third and fourth columns.

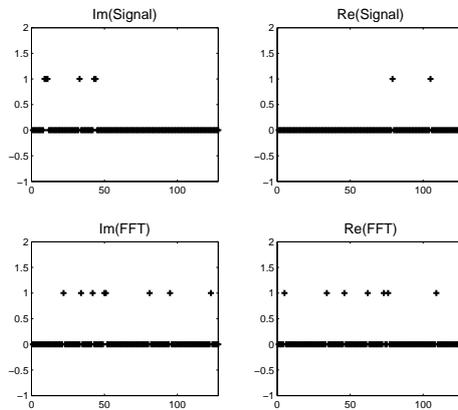


Figure 8.3: Upper left: Jarque-Bera test results from the real part of the tails of the 128 signals in time-domain. Upper right: Jarque-Bera test results from the imaginary part of the tails of the 128 signals in time-domain. Lower left: Jarque-Bera test results from the real part of the Fourier transform of the tails of the 128 signals. Lower right: Jarque-Bera test results from the imaginary part of the Fourier transform of the tails of the 128 signals.

We find that  $\widehat{r}_{NAA}/\widehat{r}_{Cr} = 2.0804$ , which is very near the expected value of about 2.

### 8.3 Discussion

Even thou human MRS spectra contains many overlapping peaks from many different metabolites, even thou we have low  $SNR$  and even thou we use a quite crude method to obtain our peak heights measurements, we do get reasonable results from our estimates. We have estimated the NAA/Cr ratio as around 2, which was what we expected.

## 9 Final words

In this thesis we have studied the distributional properties of the stochastic noise in MRS signals. We have introduced a phase-magnitude model for MRS signals and derived maximum likelihood estimators of the distributional parameters of this model. We have explicit expressions for the solutions of the likelihood equations and for the Fisher matrix. Consistency, asymptotic normality and efficiency of the estimators have been verified.

We have conducted simulation studies to test our findings. We have found that it is often better and more efficient to estimate  $(r, \sigma^2)$  from the joint distribution of the proposed phase-magnitude model, than the more conventional way to estimate them from the marginal distribution of the magnitude, ie the Rice distribution. The estimation from the phase-magnitude model does not require numerical optimization.

We have analyzed peak heights in MRS data from a MRS phantom and human brain data. We have found that our model works well for the analysis of these data sets.

Note that in the analysis of the phantom data and the *in vivo* data we had the signals from the original separate scans. Unfortunately, this is rarely the case with clinical patient data. The separate scans are not stored by the scanning software, only their mean values are. In order to thoroughly examine the difference signal model proposed in Chapter 3.3 and the ML estimation of the changes in metabolite concentration, we would need original separate clinical scans, taken at different times during the radiation therapy period. At this time, we do not have access to these types of data.

Further analysis of the phase-magnitude model for difference signals are needed when original separate scans are available.

The next step would then be to, not only estimate the peak heights of the magnitude spectra, but also take the peak widths into account, and also to incorporate the frequency dependence in our model.

## References

- [1] F. Abildgaard, H. Gesmar, and J. Led. Quantitative analysis of complicated nonideal fourier-transform nmr-spectra. *J. Magn. Reson.*, 79:78–89, 1988.
- [2] D. Andersson. The circular structural model. *J. R. Statist. Soc. B*, 43(2):131–141, 1981.
- [3] D. Andersson. Maximum likelihood estimation in the non-central chi distribution with unknown scale parameter. *Sankhya: The Indian J. Statist., Series B, Part 1*, 43:58–67, 1981.
- [4] J. Antoine, C. Chauvin, and A. Coron. Wavelets and related time-frequency techniques in magnetic resonance spectroscopy. *NMR Biomed.*, 14:265–270, 2001.
- [5] P. Barker and D. Lin. In vivo proton mr spectroscopy of the human brain. *Prog. NMR Spectrosc.*, 49:99–128, 2006.
- [6] F. Bloch. Nuclear induction. *Physical rev.*, 70(7-8), 1946.
- [7] J. Bonny, J. Renou, and M. Zanica. Optimal measurement of magnitude and phase from mr data. *J. Magn. Reson. Series B*, 113:136–144, 1996.
- [8] Y. Coenradie, R. de Beer, D. van Ormondt, H. Ratiney, S. Cavassila, and D. Graveron-Demilly. pages 248–254. RISC IEEE, 2002.
- [9] R. Corbett. How to perform automated curve fitting to in vivo 31p magnetic resonance spectroscopic data. *MAGMA*, 1:65–76, 1993.
- [10] E. R. Danielson and B. Ross. *Magnetic resonance spectroscopy diagnosis of neurological diseases*. Marcel Dekker, 1999.
- [11] H. C. de Greiff, R. Ramos-Garcia, and J. Lorenzo-Ginori. Signal denoising in magnetic resonance spectroscopy using wavelet transforms. *Concepts in Magnetic Resonance*, 14(6):388–401, 2002.
- [12] A. Devos, L. Lukas, J. Suykens, L. Vanhamme, A. Tate, F. Howe, C. Majós, A. Moreno-Torres, M. van der Graaf, C. Arús, and S. van Huffel. Classification of brain tumours using short echo time 1h mr spectra. *J. Magn. Reson.*, 170:164–175, 2004.
- [13] A. Elster and J. Burdette. *Questions and answers in magnetic resonance imaging*. Second edition.

- [14] C. Elster, R. Schubert, A. Link, M. Walzel, F. Seifert, and H. Rinneberg. Quantitative magnetic resonance spectroscopy: semi-parametric modeling and determination of uncertainties. *Magn. Reson. Med.*, 53, 2005.
- [15] R. Gabr, R. Ouwerkerk, and P. Bottomley. Quantifying in vivo mr spectra with circles. *J. Magn. Reson.*, 179:152–163, 2006.
- [16] H. Gesmar, J. Led, and F. Abildgaard. Improved methods for quantitative spectral analysis of nmr data. *Prog. NMR Spectrosc.*, 22:255–288, 1990.
- [17] P. Gillies, I. Marshall, M. Asplund, P. Winkler, and J. Higinbotham. Quantification of mrs data in the frequency domain using a wavelet filter - an approximated voigt lineshape model and prior knowledge. *NMR Biomed.*, 19(5):617–626, 2006.
- [18] H. Grage and M. Akke. A stastical analysis of nmr spectrometer noise. *J. Magn. Reson.*, 162:176–188, 2003.
- [19] J. Grivet. Accurate numerical approximation to the gauss-lorentz lineshape. *J. Magn. Reson.*, 125(1):102–106, 1997.
- [20] H. Gudbjartsson and S. Patz. The rician distribution of noisy mri data. *Magn Reson Med.*, 34(6):910–914, 1995.
- [21] J. Higinbotham and I. Marshall. Nmr lineshapes and lineshape fitting. *A. Rep. NMR Spectrosc.*, 43:59–120, 2001.
- [22] Y. Hiltunen, J. Kaartinen, J. Pulkkinen, A. Hakkinen, N. Lundbom, and R. Kauppinen. Quantification of human brain metabolites from in vivo 1h nmr magnitude spectra using automated artificial neural network analysis. *J. Magn Reson.*, 154(1):1–5, 2002.
- [23] R. Hoffman and G. Levy. Modern methods of nm data processing and data evolution. *Prog. Nucl. Magn. Reson. Spectrosc.*, 23(3):211–258, 1991.
- [24] D. Hoult. *Encyclopedia of Nuclear Magnetic Resonance*, pages 4256–4266. Wiley, 1996.
- [25] D. Hoult and R. Richards. The signal-to-noise ratio of the nuclear magnetic resonance experiment. *J. Magn Reson.*, 24:71–85, 1976.

- [26] F. Howe, S. Barton, S. Cudlip, M. Stubbs, D. Saunders, M. Murphy, P. Wilkins, K. Opstad, V. Doyle, M. McLean, B. Bell, and J. Griffiths. Metabolic profiles of human brain tumors using quantitative in vivo 1h magnetic resonance spectroscopy. *Magn. Reson. Med.*, 49:223–232, 2003.
- [27] S. Jammalamadaka and A. SenGupta. *Topics in circular statistics*. World Scientific, 2001.
- [28] C. Jarque and A. Bera. A test for normality of observations and regression residuals. *International Statistical Review.*, 55(2):163–172, 1987.
- [29] N. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions, Vol. 2*. Wiley, 1995.
- [30] J. Kaartinen, S. Mierisová, J. Oja, J. Usenius, R. Kauppinen, and Y. Hiltunen. Automated quantification of human brain metabolites by artificial neural network analysis from in vivo single-voxel 1h nmr spectra. *J. Magn Reson.*, 134(1):176–179, 1998.
- [31] H. Koh, S. Maddula, J. Lambert, R. Hergenröder, and L. Hildebrand. An approach to automated frequency-domain feature extraction in nuclear magnetic resonance spectroscopy. *J. Magn Reson.*, 201:146–156, 2009.
- [32] R. Kumerresan, C. Ramalingam, and D. van Ormondt. Estimating the parameters of nmr signals by transforming to the frequency domain. *J. Magn Reson.*, 89:562–567, 1990.
- [33] C. Ladroue, F. Howe, J. Griffiths, and A. Tate. Independent component analysis for automated decomposition of in vivo magnetic resonance spectra. *Magn. Reson. Med.*, 50:697–703, 2003.
- [34] X. Leclerc, T. Huisman, and A. Sorensen. The potential of proton magnetic resonance spectroscopy (1h) in the diagnosis and management of patients with brain tumors. *Curr. Opin. Oncol.*, 14:292–298, 2002.
- [35] E. Lehman. *Theory of Point Estimation*. Wiley, 1983.
- [36] L. Lukas, A. Devos, J. Suykens, L. Vanhamme, F. Howe, C. Majós, A. Moreno-Torres, M. van der Graaf, A. Tate, C. Arús, and S. van Huffel. Brain tumor classification based on long echo proton mrs signals. *Artificial Intelligence in Medicine*, 31:73–89, 2004.

- [37] C. Majós, J. Alonso, C. Aguilera, M. Serrallonga, J. Acebes, C. Arús, and J. Gili. Adult primitive neuroectodermal tumor: proton mr spectroscopic findings with possible applications for differential diagnosis. *Radiology*, 225:556–566, 2002.
- [38] I. Marshall, S. Bruce, J. Higinbotham, A. MacLulich, J. Wardlaw, K. Ferguson, and J. Seckl. Choice of spectroscopic lineshape model affects metabolite peak areas and area ratios. *Magn. Reson. Med.*, 44(4):646–649, 2000.
- [39] I. Marshall, J. Higinbotham, S. Bruce, and A. Freise. Use of voigt lineshape for quantification of in vivo 1h spectra. *Magn. Reson. Med.*, 37:651–657, 1997.
- [40] Y. Martin. A global approach to accurate and automatic quantitative analysis of nmr spectra by complex least-square curve fitting. *J. Magn Reson. Series A*, 111:1–10, 1994.
- [41] X. McCullagh and J. Nelder. *Generalized linear models*. Chapman and Hall, 1989.
- [42] X. Meng and D. Rubin. Maximum likelihood estimation via the ecm algorithm: A general framework. *Biometrika*, 80(2):267–278, 1993.
- [43] P. Meyer. The maximum likelihood estimate of the non-centrality parameter of a non-central  $\chi^2$  variate. *J. Amer. Statist. Assoc.*, 62:1258–1264, 1967.
- [44] R. Meyer, M. Fisher, S. Nelson, and T. Brown. Evaluation of manual methods for integration of in vivo phosphorus nmr spectra. *NMR Biomed.*, 1(3):131–135, 1988.
- [45] S. Mierisová and M. Ala-Korpela. Mr spectroscopy quantitation: a review of frequency domain methods. *NMR Biomed.*, 14:247–259, 2001.
- [46] S. Mukherji. *Clinical Applications of Magnetic Resonance Spectroscopy*. Wiley-Liss, 1998.
- [47] M. Murphy, A. Loosemore, A. Clifton, F. Howe, A. Tate, S. Cudlip, P. Wilkins, J. Griffiths, and B. Bell. The contribution of proton magnetic resonance spectroscopy (1h mrs) to clinical brain tumour diagnosis. *Br. J. Neurosurg.*, 16(4):329–334, 2002.
- [48] S. Nelson and T. Brown. The accuracy of quantification from 1d nmr spectra using the piqable algorithm. *J. Magn. Reson.*, 84:95–109, 1989.

- [49] F. Olver, D. Lozier, R. Boisvert, and C. Clark. *NIST Handbook of Mathematical Functions*. Cambridge University Press, 2010.
- [50] J. Pouillet, D. Sima, and S. van Huffel. Mrs signal quantitation: a review of time- and frequency-domain methods. *J. Magn. Reson.*, 195(2):134–144, 2008.
- [51] J. Pouillet, D. Sima, S. van Huffel, and P. van Hecke. Frequencyselective quantitation of short-echo time 1h magnetic resonance spectra. *J. Magn. Reson.*, 186(2):293–304, 2007.
- [52] S. Provencher. Estimation of metabolite concentrations from localized in vivo proton nmr spectra. *Magn. Reson. Med.*, 30(6):672–679, 1993.
- [53] 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.*, 18:1–13, 2005.
- [54] L. Råde and B. Westergren. *Mathematics handbook for science and engineering*. Studentlitteratur, 1995.
- [55] S. Rice. Mathematical analysis of random noise. *Bell System Technical Journal*, 24:46–156, 1945.
- [56] S. Rice. Statistical properties of a sine wave plus random noise. *Bell System Technical Journal*, 27:109–157, 1948.
- [57] D. Rowe. Modeling both the magnitude and phase of complex-valued fmri data. *NeuroImage*, 25:1310–1324, 2005.
- [58] D. Rowe. Parameter estimation in the magnitude-only and complex-valued fmri data models. *NeuroImage*, 25:1124–1132, 2005.
- [59] D. Rowe and B. Logan. A complex way to compute fmri activation. *NeuroImage*, 23:1078–1092, 2004.
- [60] D. Rowe and B. Logan. Complex fmri analysis with unrestricted phase is equivalent to a magnitude-only model. *NeuroImage*, 24:603–606, 2005.
- [61] N. Sandgren, Y. Selén, P. Stoica, and J. Li. Parametric methods for frequency-selective mr spectroscopy - a review. *J. Magn. Reson.*, 168(2):259–272, 2004.

- [62] J. Sijbers and A. den Dekker. Maximum likelihood estimation of signal amplitude and noise variance from mr data. *Magn. Reson. Med.*, 51:586–594, 2004.
- [63] J. Sijbers, A. den Dekker, P. Scheunders, and D. V. Dyck. Maximum-likelihood estimation of rician distribution parameters. *IEEE T. Med. Imaging*, 17(3):357–361, 1998.
- [64] J. Slotboom, C. Boesch, and R. Kreis. Versatile frequency domain fitting using time domain models and prior knowledge. *Magn. Reson. Med.*, 39:899–911, 1998.
- [65] I. Smith and L. Stewart. Magnetic resonance spectroscopy in medicine: clinical impact. *Prog. NMR Spectrosc.*, 40:1–34, 2002.
- [66] B. Soher, K. Young, V. Govindaraju, and A. Maudsley. Automated spectral analysis iii: application of in vivo proton mr spectroscopy and spectroscopic imaging. *Magn. Reson. Med.*, 40(6):822–831, 1998.
- [67] D. Spielman, P. Webb, and A. Macovski. A statistical framework for in vivo spectroscopic imaging. *J. Magn. Reson.*, 79:66–77, 1988.
- [68] P. Stoica, N. Sandgren, Y. Selén, L. Vanhamme, and S. van Huffel. Frequency-domain method based on the singular value decomposition for frequency-selective nmr spectroscopy. *J. Magn. Reson.*, 165(1):80–88, 2003.
- [69] P. Stoica and T. Sundin. Exact ml estimation of spectroscopic parameters. *J. Magn. Reson.*, 145:108–114, 2000.
- [70] R. Stoyanova and T. Brown. Nmr spectral quantitation by principal component analysis ii. determination of frequency and phase shifts. *J. Magn. Reson.*, 112(2):32–43, 1996.
- [71] R. Stoyanova and T. Brown. Nmr spectral quantitation by principal component analysis. *NMR Biomed.*, 14(4):271–277, 2001.
- [72] R. Stoyanova and T. Brown. Nmr spectral quantitation by principal component analysis iii. a generalized procedure for determination of lineshape variations. *J. Magn Reson.*, 154(2):163–175, 2002.
- [73] R. Stoyanova, A. Kuesel, and T. Brown. Application of principal component analysis for spectral quantitation. *J. Magn Reson.*, 115:265–269, 1995.

- [74] K. Talukdar and W. Lawing. Estimation of the parameters of the rice distribution. *J. Acoust. Soc. Am.*, 89(3):1193–1197, 1991.
- [75] A. van der Vaart. *Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, 1998.
- [76] A. van der Vaart and J. Wellner. *Weak Convergence and Empirical Processes with Applications to Statistics*. Springer, 1996.
- [77] L. Vanhamme. *Advanced Time-Domain Methods for Nuclear Magnetic Resonance Spectroscopy Data Analysis*. PhD thesis, Katholieke Universiteit, Leuven, Belgium, 1999.
- [78] L. Vanhamme, T. Sundin, P. van Hecke, and S. van Huffel. Mr spectroscopy quantitation: a review of time-domain methods. *NMR Biomed.*, 14:233–246, 2001.
- [79] K. Young, B. Soher, and A. Maudsley. Automated spectral analysis ii: application of wavelet shrinkage for characterization of non-parameterized signals. *Magn. Reson. Med.*, 40(6):816–821, 1998.
- [80] H. Zhu, Y. Li, J. Ibrahim, X. Shi, H. An, Y. Chen, W. Gao, D. Rowe, and B. Peterson. Regression models for identifying noise sources in magnetic resonance images. *J. Amer. Statist. Assoc.*, 104(486):623–637, 2009.

## Acknowledgements

Many thanks to my supervisor Jun Yu and my deputy supervisors Magnus Ekström and Nils Östlund for sharing all their knowledge, for their involvement, and for excellent supervising.

I wish to thank Jon Hauksson for reading and commenting on my work, for all the data, and for answering all my physics questions (even the really stupid ones).

Many thanks to Anders Garpebring for all the help with obtaining data.

I would also like to thank colleagues at the Centre of Biostochastics and the Department of Forest Economics.

Last, a thousand times thanks to my family and friends.

## Appendix A: Verification of Assumption (E)

We have

$$f_{Z,\Theta}(z, \theta | r, \psi, \sigma^2) = \frac{z}{2\pi\sigma^2} \cdot e^{-\frac{1}{2\sigma^2}(z^2+r^2-2rz\cos(\theta-\psi))}.$$

For simplicity we let  $x = -\frac{1}{2\sigma^2}(z^2+r^2-2rz\cos(\theta-\psi))$  and  $\tau = z\cos(\theta-\psi) - r$ . We get the following first derivatives with respect to  $r$ ,  $\psi$  and  $\sigma^2$

$$\begin{aligned}\frac{\partial f}{\partial r} &= \frac{z}{2\pi\sigma^4}\tau e^x, \\ \frac{\partial f}{\partial \psi} &= \frac{rz^2\sin(\theta-\psi)}{2\pi\sigma^4}e^x, \\ \frac{\partial f}{\partial \sigma^2} &= \frac{z}{2\pi\sigma^4}(-1-x)e^x.\end{aligned}$$

This gives the following second derivatives

$$\begin{aligned}\frac{\partial^2 f}{\partial r^2} &= \frac{z}{2\pi\sigma^6}(\tau^2 - \sigma^2)e^x, \\ \frac{\partial^2 f}{\partial r \partial \psi} &= \frac{z^2\sin(\theta-\psi)}{2\pi\sigma^6}(r\tau + \sigma^2)e^x, \\ \frac{\partial^2 f}{\partial r \partial \sigma^2} &= \frac{z}{2\pi\sigma^6}\tau(2\sigma^2 - x)e^x, \\ \frac{\partial^2 f}{\partial \psi^2} &= \frac{rz^2}{2\pi\sigma^6}(rz\sin^2(\theta-\psi) - \cos(\theta-\psi)\sigma^2)e^x, \\ \frac{\partial^2 f}{\partial \psi \partial \sigma^2} &= \frac{rz^2\sin(\theta-\psi)}{2\pi\sigma^6}(-\sigma^2 - x)e^x, \\ \frac{\partial^2 f}{\partial (\sigma^2)^2} &= \frac{z}{2\pi\sigma^6}(2 + 4x + x^2)e^x.\end{aligned}$$

Hence

$$\begin{aligned}
\frac{\partial^3 f}{\partial r^3} &= \frac{z}{2\pi\sigma^8} (\tau^3 - 3\sigma^2\tau) e^x, \\
\frac{\partial^3 f}{\partial r^2 \partial \psi} &= \frac{z^2 \sin(\theta - \psi)}{2\pi\sigma^8} (r\tau^2 + 2\sigma^2\tau - r\sigma^2) e^x, \\
\frac{\partial^3 f}{\partial r^2 \partial \sigma^2} &= \frac{z}{4\pi\sigma^8} (\sigma^2(4 + 2x) - 3\tau^2 - 2x\tau) e^x, \\
\frac{\partial^3 f}{\partial r \partial \psi^2} &= \frac{z^2 \cos(\theta - \psi)}{2\pi\sigma^6} (-\sigma^2 - r\tau) e^x + \frac{r z^3 \sin^2(\theta - \psi)}{2\pi\sigma^8} (2\sigma^2 + r\tau) e^x \\
\frac{\partial^3 f}{\partial r \partial \psi \partial \sigma^2} &= \frac{z^2 \sin(\theta - \psi)}{2\pi\sigma^8} (\sigma^2(-2 - x) + \tau(-3 - rx)) e^x \\
\frac{\partial^3 f}{\partial r \partial (\sigma^2)^2} &= \frac{z}{2\pi\sigma^8} \tau(6 + 6x + x^2) e^x, \\
\frac{\partial^3 f}{\partial \psi^3} &= \frac{r z^2 \sin(\theta - \psi)}{2\pi\sigma^8} (-\sigma^4 - 3r\sigma^2 \cos(\theta - \psi) + r^2 z^2 \sin^2(\theta - \psi)) e^x, \\
\frac{\partial^3 f}{\partial \psi^2 \partial \sigma^2} &= \frac{r z^2 \cos(\theta - \psi)}{2\pi\sigma^6} (2 + x) e^x + \frac{r^2 z^3 \sin^2(\theta - \psi)}{2\pi\sigma^8} (-3 - x) e^x, \\
\frac{\partial^3 f}{\partial \psi \partial (\sigma^2)^2} &= \frac{r 2z^2 \sin(\theta - \psi)}{2\pi\sigma^8} (6 + 6x + x^2) e^x, \\
\frac{\partial^3 f}{\partial (\sigma^2)^3} &= \frac{z}{2\pi\sigma^8} (-6 - 18x - 9x^2 - x^3) e^x.
\end{aligned}$$

Hence, assumption (E) is satisfied.

## Appendix B: Verification of assumption (F)

With

$$\ln f_{Z,\Theta}(z, \theta | r, \psi, \sigma^2) = \ln z - \ln 2\pi\sigma^2 - \frac{1}{2\sigma^2} (z^2 + r^2 - 2rz \cos(\theta - \psi))$$

and

$$\frac{\partial \ln f}{\partial r} = \frac{1}{\sigma^2} (z \cos(\theta - \psi) - r)$$

$$\frac{\partial \ln f}{\partial \psi} = \frac{1}{\sigma^2} r z \sin(\theta - \psi)$$

$$\frac{\partial \ln f}{\partial \sigma^2} = -\frac{1}{\sigma^2} + \frac{1}{2\sigma^4} (z^2 + r^2 - 2rz \cos(\theta - \psi))$$

we obtain

$$\begin{aligned} \mathbb{E}\left(\frac{\partial \ln f}{\partial r}\right) &= \frac{1}{\sigma^2} \mathbb{E}(Z \cos(\Theta - \psi)) - \frac{r}{\sigma^2} \\ &= \frac{1}{\sigma^2} \mathbb{E}(Z \cos \Theta \cos \psi + Z \sin \Theta \sin \psi) - \frac{r}{\sigma^2} \\ &= \frac{1}{\sigma^2} (\cos \psi \mathbb{E}(Z \cos \Theta) + \sin \psi \mathbb{E}(Z \sin \Theta)) - \frac{r}{\sigma^2} \\ &= \frac{1}{\sigma^2} (r \cos^2 \psi + r \sin^2 \psi) - \frac{r}{\sigma^2} = 0, \end{aligned}$$

where we have used the fact that  $\mathbb{E}(Z \cos \Theta) = r \cos \psi$  and  $\mathbb{E}(Z \sin \Theta) = r \sin \psi$ . In a similar way we can conclude that

$$\begin{aligned} \mathbb{E}\left(\frac{\partial \ln f}{\partial \psi}\right) &= \frac{r}{\sigma^2} \mathbb{E}(Z \sin(\Theta - \psi)) \\ &= \frac{r}{\sigma^2} (\cos \psi \mathbb{E}(Z \sin \Theta) - \sin \psi \mathbb{E}(Z \cos \Theta)) \\ &= \frac{r}{\sigma^2} (r \sin \psi \cos \psi - r \sin \psi \cos \psi) = 0, \end{aligned}$$

and, by observing that  $\mathbb{E}Z^2 = r^2 + 2\sigma^2$  and  $\mathbb{E}(Z \cos(\Theta - \psi)) = r$ , we obtain

$$\begin{aligned} \mathbb{E}\left(\frac{\partial \ln f}{\partial \sigma^2}\right) &= -\frac{1}{\sigma^2} + \frac{1}{2\sigma^4} \mathbb{E}(Z^2 + r^2 - 2rZ \cos(\Theta - \psi)) \\ &= -\frac{1}{\sigma^2} + \frac{1}{2\sigma^4} (r^2 + 2\sigma^2) + \frac{r^2}{2\sigma^4} - \frac{2r^2}{2\sigma^4} = 0 \end{aligned}$$

Hence the first part of assumption (F) is satisfied.

Now, let's direct our attention to the calculations of the Fisher information matrix  $J$ .

With

$$\frac{\partial \ln f}{\partial r} \cdot \frac{\partial \ln f}{\partial r} = \frac{r^2}{\sigma^4} - \frac{2r}{\sigma^4} Z \cos(\theta - \psi) + \frac{1}{\sigma^4} Z^2 \cos^2(\theta - \psi)$$

and using that  $(Z \sin \Theta)$  and  $(Z \cos \Theta)$  are independent, and that  $E(Z \sin \Theta)^2 = r^2 \sin^2 \psi + \sigma^2$  and  $E(Z \cos \Theta)^2 = r^2 \cos^2 \psi + \sigma^2$ , we get

$$\begin{aligned} E\left(\frac{\partial \ln f}{\partial r} \cdot \frac{\partial \ln f}{\partial r}\right) &= \frac{r^2}{\sigma^4} - \frac{2r}{\sigma^4} E(Z \cos(\Theta - \psi)) + \frac{1}{\sigma^4} E(Z^2 \cos^2(\Theta - \psi)) \\ &= \frac{r^2}{\sigma^4} - \frac{2r^2}{\sigma^4} + \frac{1}{\sigma^4} (\cos^2 \psi E(Z \cos \Theta)^2) + \frac{1}{\sigma^4} (\sin^2 \psi E(Z \sin \Theta)^2) \\ &\quad + \frac{2}{\sigma^4} (\sin \psi \cos \psi E(Z^2 \sin \Theta \cos \Theta)) \\ &= -\frac{r^2}{\sigma^4} + \frac{1}{\sigma^4} (\cos^2 \psi (r^2 \cos^2 \psi + \sigma^2)) + \frac{1}{\sigma^4} (\sin^2 \psi (r^2 \sin^2 \psi + \sigma^2)) \\ &\quad + \frac{2}{\sigma^4} (\sin \psi \cos \psi E(Z \sin \Theta) E(Z \cos \Theta)) \\ &= -\frac{r^2}{\sigma^4} + \frac{\sin^2 \psi + \cos^2 \psi}{\sigma^2} + \frac{r^2}{\sigma^4} (\sin^4 \psi + \cos^4 \psi + 2 \sin^2 \psi \cos^2 \psi) \\ &= -\frac{r^2}{\sigma^4} + \frac{1}{\sigma^2} + \frac{r^2}{\sigma^4} (\sin^2 \psi + \cos^2 \psi)^2 \\ &= \frac{1}{\sigma^2}. \end{aligned}$$

Since

$$\frac{\partial^2 \ln f}{\partial r^2} = -\frac{1}{\sigma^2},$$

we have that

$$E\left(-\frac{\partial^2 \ln f}{\partial r^2}\right) = \frac{1}{\sigma^2},$$

that is, that

$$J_{11} = E\left(\frac{\partial \ln f}{\partial r} \cdot \frac{\partial \ln f}{\partial r}\right) = E\left(-\frac{\partial^2 \ln f}{\partial r^2}\right) = \frac{1}{\sigma^2}.$$

We have

$$\frac{\partial \ln f}{\partial r} \cdot \frac{\partial \ln f}{\partial \psi} = \frac{r^2}{\sigma^4} z \sin(\theta - \psi) - \frac{r}{\sigma^4} z^2 \sin(\theta - \psi) \cos(\theta - \psi).$$

Note the result from above that  $E(Z \sin(\Theta - \psi)) = 0$ , and the fact that since  $(Z \sin \Theta)$  and  $(Z \cos \Theta)$  are independent and normally distributed, then  $(Z \sin(\Theta - \psi))$  and  $(Z \cos(\Theta - \psi))$  are also independent. Hence

$$\begin{aligned} E\left(\frac{\partial \ln f}{\partial r} \cdot \frac{\partial \ln f}{\partial \psi}\right) &= \frac{r^2}{\sigma^4} E(Z \sin(\Theta - \psi)) - \frac{r}{\sigma^4} E(Z^2 \sin(\Theta - \psi) \cos(\Theta - \psi)) \\ &= 0 - \frac{r}{\sigma^4} E(Z \sin(\Theta - \psi)) E(Z \cos(\Theta - \psi)) \\ &= 0. \end{aligned}$$

With

$$\frac{\partial^2 \ln f}{\partial r \partial \psi} = -\frac{1}{\sigma^2} z \sin(\theta - \psi),$$

we clearly have that

$$E\left(-\frac{\partial^2 \ln f}{\partial r \partial \psi}\right) = \frac{1}{\sigma^2} E(Z \sin(\Theta - \psi)) = 0,$$

that is, that

$$J_{12} = J_{21} = E\left(\frac{\partial \ln f}{\partial r} \cdot \frac{\partial \ln f}{\partial \psi}\right) = E\left(-\frac{\partial^2 \ln f}{\partial r \partial \psi}\right) = 0.$$

The next element in the Fisher matrix follows from

$$\frac{\partial \ln f}{\partial r} \cdot \frac{\partial \ln f}{\partial \sigma^2} = -\frac{z \cos(\theta - \psi) - r}{2\sigma^2} \left( \frac{z^2 + r^2 - 2rz \cos(\theta - \psi)}{2\sigma^4} - \frac{1}{\sigma^2} \right)$$

By noting above that  $E(Z \cos(\Theta - \psi) - r) = r - r = 0$  we obtain

$$\begin{aligned} E\left(\frac{\partial \ln f}{\partial r} \cdot \frac{\partial \ln f}{\partial \sigma^2}\right) &= -\frac{1}{2\sigma^2} E(Z \cos(\Theta - \psi) - r) \\ &\quad \cdot \left(\frac{EZ^2 + r^2 - 2rE(Z \cos(\Theta - \psi))}{2\sigma^4} - \frac{1}{\sigma^2}\right) \\ &= 0. \end{aligned}$$

With

$$\frac{\partial^2 \ln f}{\partial r \partial \sigma^2} = -\frac{1}{\sigma^2} (z \cos(\theta - \psi) - r),$$

we have

$$E\left(-\frac{\partial^2 \ln f}{\partial r \partial \sigma^2}\right) = \frac{1}{\sigma^2} E(z \cos(\theta - \psi) - r) = 0,$$

that is

$$J_{13} = J_{31} = E\left(\frac{\partial \ln f}{\partial r} \cdot \frac{\partial \ln f}{\partial \sigma^2}\right) = E\left(-\frac{\partial^2 \ln f}{\partial r \partial \sigma^2}\right) = 0.$$

From

$$\frac{\partial \ln f}{\partial \psi} \cdot \frac{\partial \ln f}{\partial \psi} = \frac{r^2 z^2 \sin^2(\theta - \psi)}{\sigma^4}$$

we have

$$\begin{aligned} E\left(\frac{\partial \ln f}{\partial \psi} \cdot \frac{\partial \ln f}{\partial \psi}\right) &= \frac{r^2}{\sigma^4} E(Z^2 \sin^2(\Theta - \psi)) \\ &= \frac{r^2}{\sigma^4} \sin^2 \psi E(Z \cos \Theta)^2 + \frac{r^2}{\sigma^4} \cos^2 \psi E(Z \sin \Theta)^2 \\ &\quad - \frac{2r^2}{\sigma^4} \sin \psi \cos \psi E(Z \sin \Theta)(Z \cos \Theta) \\ &= \frac{r^2}{\sigma^4} \sin^2 \psi (r^2 \cos^2 \psi - \sigma^2) + \frac{r^2}{\sigma^4} \cos^2 \psi (r^2 \sin^2 \psi - \sigma^2) \\ &\quad - \frac{2r^4}{\sigma^4} \sin^2 \psi \cos^2 \psi \\ &= \frac{r^2}{\sigma^2}, \end{aligned}$$

and from

$$\frac{\partial^2 \ln f}{\partial \psi^2} = -\frac{r}{\sigma^2} (z \cos(\theta - \psi)),$$

we have

$$\mathbb{E} \left( -\frac{\partial^2 \ln f}{\partial \psi^2} \right) = \frac{r}{\sigma^2} \mathbb{E}(Z \cos(\Theta - \psi)) = \frac{r^2}{\sigma^2},$$

that is

$$J_{22} = \mathbb{E} \left( \frac{\partial \ln f}{\partial \psi} \cdot \frac{\partial \ln f}{\partial \psi} \right) = \mathbb{E} \left( -\frac{\partial^2 \ln f}{\partial \psi^2} \right) = \frac{r^2}{\sigma^2}.$$

With

$$\frac{\partial \ln f}{\partial \psi} \cdot \frac{\partial \ln f}{\partial \sigma^2} = \frac{rz}{\sigma^2} \sin(\theta - \psi) \left( \frac{z^2 + r^2 - 2rz \cos(\theta - \psi)}{2\sigma^4} - \frac{1}{\sigma^2} \right)$$

we clearly have

$$\mathbb{E} \left( \frac{\partial \ln f}{\partial \psi} \cdot \frac{\partial \ln f}{\partial \sigma^2} \right) = 0,$$

and since

$$\frac{\partial^2 \ln f}{\partial \psi \partial \sigma^2} = -\frac{r}{\sigma^4} z \sin(\theta - \psi),$$

then

$$\mathbb{E} \left( -\frac{\partial^2 \ln f}{\partial \psi \partial \sigma^2} \right) = 0.$$

that is

$$J_{23} = J_{32} = \mathbb{E} \left( \frac{\partial \ln f}{\partial \psi} \cdot \frac{\partial \ln f}{\partial \sigma^2} \right) = \mathbb{E} \left( -\frac{\partial^2 \ln f}{\partial \psi \partial \sigma^2} \right) = 0.$$

Last, we have

$$\begin{aligned} \frac{\partial \ln f}{\partial \sigma^2} \cdot \frac{\partial \ln f}{\partial \sigma^2} &= \frac{1}{\sigma^4} - \frac{z^2 + r^2 - 2rz \cos(\theta - \psi)}{\sigma^6} \\ &\quad + \left( \frac{z^2 + r^2 - 2rz \cos(\theta - \psi)}{2\sigma^4} \right)^2 \end{aligned}$$

so, since  $EZ^2 = r^2 + 2\sigma^2$ ,  $EZ^4 = r^4 + 8r^2\sigma^2 + 8\sigma^4$  and  $E(Z \cos(\Theta - \psi))^2 = r^2 + \sigma^2$ , we obtain

$$\begin{aligned} E\left(\frac{\partial \ln f}{\partial \sigma^2} \cdot \frac{\partial \ln f}{\partial \sigma^2}\right) &= \frac{1}{\sigma^4} - \frac{1}{\sigma^6}EZ^2 - \frac{r^2}{\sigma^6} + \frac{2r}{\sigma^6}E(Z \cos(\Theta - \psi)) \\ &\quad + \frac{1}{4\sigma^8}EZ^4 + \frac{r^4}{4\sigma^8} + \frac{r^2}{2\sigma^8}EZ^2 \\ &\quad - \frac{r^3}{\sigma^8}E(Z \cos(\Theta - \psi)) - \frac{r}{\sigma^8}EZ^2E(Z \cos(\Theta - \psi)) \\ &\quad + \frac{r^2}{\sigma^8}E(Z \cos(\Theta - \psi))^2 = \frac{1}{\sigma^4} \end{aligned}$$

and with

$$\frac{\partial^2 \ln f}{\partial (\sigma^2)^2} = \frac{1}{\sigma^4} - \frac{z^2 + r^2 - 2rz \cos(\theta - \psi)}{\sigma^6},$$

we have

$$\begin{aligned} E\left(-\frac{\partial^2 \ln f}{\partial (\sigma^2)^2}\right) &= -\frac{1}{\sigma^4} + \frac{1}{\sigma^6}EZ^2 + \frac{r^2}{\sigma^6} - \frac{2r}{\sigma^6}E(Z \cos(\Theta - \psi)) \\ &= -\frac{1}{\sigma^4} + \frac{1}{\sigma^6}(r^2 + 2\sigma^2) + \frac{r^2}{\sigma^6} - \frac{2r^2}{\sigma^6} = \frac{1}{\sigma^4} \end{aligned}$$

Hence the Fisher matrix is

$$J = \begin{pmatrix} 1/\sigma^2 & 0 & 0 \\ 0 & r^2/\sigma^2 & 0 \\ 0 & 0 & 1/\sigma^4 \end{pmatrix}$$

and the second part of assumption (F) is satisfied.

## Appendix C: Verification of assumption (H)

We let

$$\zeta = \left(\frac{r_0}{2}, r_0 + 1\right) \times \left(\frac{-\pi - |\psi_0|}{2}, \frac{\pi + |\psi_0|}{2}\right) \times \left(\frac{\sigma_0^2}{2}, \sigma_0^2 + 1\right),$$

and note that  $(\psi_0, r_0, \sigma_0^2) \in \zeta$ .

With

$$\begin{aligned} \frac{\partial^3 \ln f}{\partial r^3} &= 0, \\ \frac{\partial^3 \ln f}{\partial r^2 \partial \psi} &= 0, \\ \frac{\partial^3 \ln f}{\partial r^2 \partial \sigma^2} &= \frac{1}{\sigma^4}, \\ \frac{\partial^3 \ln f}{\partial r \partial \psi^2} &= -\frac{1}{\sigma^2} z \cos(\theta - \psi), \\ \frac{\partial^3 \ln f}{\partial r \partial \psi \partial \sigma^2} &= -\frac{1}{\sigma^4} z \sin(\theta - \psi), \\ \frac{\partial^3 \ln f}{\partial r \partial (\sigma^2)^2} &= \frac{2}{\sigma^6} (z \cos(\theta - \psi) - r), \\ \frac{\partial^3 \ln f}{\partial \psi^3} &= -\frac{r}{\sigma^2} z \sin(\theta - \psi), \\ \frac{\partial^3 \ln f}{\partial \psi^2 \partial \sigma^2} &= \frac{r}{\sigma^4} z \cos(\theta - \psi), \\ \frac{\partial^3 \ln f}{\partial \psi \partial (\sigma^2)^2} &= \frac{2r}{\sigma^6} z \sin(\theta - \psi), \\ \frac{\partial^3 \ln f}{\partial (\sigma^2)^3} &= -\frac{2}{\sigma^6} + \frac{3(z^2 + r^2 - 2rz \cos(\theta - \psi))}{\sigma^8}. \end{aligned}$$

we may chose

$$\left| \frac{\partial^3 \ln f}{\partial r^3} \right| = \left| \frac{\partial^3 \ln f}{\partial r^2 \partial \psi} \right| = 0 = M_{111} = M_{112} = M_{121} = M_{211}$$

for which

$$m_{111} = m_{112} = m_{121} = m_{211} = E0 = 0 < \infty$$

and

$$\left| \frac{\partial^3 \ln f}{\partial r^2 \partial \sigma^2} \right| = \frac{1}{\sigma^4} < \frac{4}{\sigma_0^4} = M_{113} = M_{131} = M_{311},$$

where, since  $\sigma_0 > 0$

$$m_{113} = m_{131} = m_{311} = \frac{4}{\sigma_0^4} < \infty.$$

By noting that  $z \leq 1 + z^2$  we let

$$\left| \frac{\partial^3 \ln f}{\partial r \partial \psi^2} \right| = \left| -\frac{1}{\sigma^2} z \cos(\theta - \psi) \right| \leq \frac{2}{\sigma_0^2} z \leq \frac{2}{\sigma_0^2} (1 + z^2) = M_{122} = M_{212} = M_{221},$$

where

$$m_{122} = m_{212} = m_{221} = \frac{2}{\sigma_0^2} (1 + EZ^2) = \frac{2}{\sigma_0^2} (1 + r_0^2 + 2\sigma_0^2) < \infty,$$

and

$$\begin{aligned} \left| \frac{\partial^3 \ln f}{\partial r \partial \psi \partial \sigma^2} \right| &= \left| -\frac{1}{\sigma^4} z \sin(\theta - \psi) \right| \leq \frac{4}{\sigma_0^4} z \\ &\leq \frac{4}{\sigma_0^4} (1 + z^2) = M_{123} = M_{312} = M_{231} = M_{132} = M_{213} = M_{321}, \end{aligned}$$

where

$$m_{123} = m_{312} = m_{231} = m_{132} = m_{213} = m_{321} = \frac{4}{\sigma_0^4} (1 + r_0^2 + 2\sigma_0^2) < \infty,$$

and

$$\begin{aligned} \left| \frac{\partial^3 \ln f}{\partial r \partial (\sigma^2)^2} \right| &= \left| \frac{2}{\sigma^6} (z \cos(\theta - \psi) - r) \right| \leq \frac{16}{\sigma_0^6} (z + r_0 + 1) \\ &\leq \frac{16}{\sigma_0^6} (r_0 + 2 + z^2) = M_{133} = M_{313} = M_{331}, \end{aligned}$$

where

$$m_{133} = m_{313} = m_{331} = \frac{16}{\sigma_0^6} (r_0 + 2 + r_0^2 + 2\sigma_0^2) < \infty,$$

and

$$\left| \frac{\partial^3 \ln f}{\partial \psi^3} \right| = \left| -\frac{r}{\sigma^2} z \sin(\theta - \psi) \right| \leq \frac{2(r_0 + 1)}{\sigma_0^2} z \leq \frac{2(r_0 + 1)}{\sigma_0^2} (1 + z^2) = M_{222},$$

where

$$m_{222} = \frac{2(r_0 + 1)}{\sigma_0^2} (1 + r_0^2 + 2\sigma_0^2) < \infty,$$

and

$$\begin{aligned} \left| \frac{\partial^3 \ln f}{\partial \psi^2 \partial \sigma^2} \right| &= \left| \frac{r}{\sigma^4} z \cos(\theta - \psi) \right| \leq \frac{4(r_0 + 1)}{\sigma_0^4} z \\ &\leq \frac{4(r_0 + 1)}{\sigma_0^4} (1 + z^2) = M_{223} = M_{232} = M_{322}, \end{aligned}$$

where

$$m_{223} = m_{232} = m_{322} = \frac{4(r_0 + 1)}{\sigma_0^4} (1 + r_0^2 + 2\sigma_0^2) < \infty,$$

and

$$\begin{aligned} \left| \frac{\partial^3 \ln f}{\partial \psi \partial (\sigma^2)^2} \right| &= \left| \frac{2r}{\sigma^6} z \sin(\theta - \psi) \right| \leq \frac{16(r_0 + 1)}{\sigma_0^6} z \\ &\leq \frac{16(r_0 + 1)}{\sigma_0^6} (1 + z^2) = M_{233} = M_{323} = M_{332}, \end{aligned}$$

where

$$m_{233} = m_{323} = m_{332} = \frac{16(r_0 + 1)}{\sigma_0^6} (1 + r_0^2 + 2\sigma_0^2) < \infty.$$

Last, we let

$$\begin{aligned}
\left| \frac{\partial^3 \ln f}{\partial(\sigma^2)^3} \right| &= \left| -\frac{2}{\sigma^6} + \frac{3(z^2 + r^2 - 2rz \cos(\theta - \psi))}{\sigma^8} \right| \\
&\leq \frac{16}{\sigma_0^6} + \frac{48}{\sigma_0^8} (z^2 + (r_0 + 1)^2 + 2(r_0 + 1)z) \\
&= \frac{16}{\sigma_0^6} + \frac{48}{\sigma_0^8} ((r_0 + 1) + z)^2 \\
&\leq \frac{16}{\sigma_0^6} + \frac{48}{\sigma_0^8} ((r_0 + 2) + z)^2 \\
&= \frac{16}{\sigma_0^6} + \frac{48}{\sigma_0^8} ((r_0 + 2)^2 + 2(r_0 + 2)z^2 + z^4) = M_{333}
\end{aligned}$$

for which

$$m_{333} = \frac{16}{\sigma_0^6} + \frac{48}{\sigma_0^8} ((r_0 + 2)^2 + 2(r_0 + 2)(r_0^2 + 2\sigma_0^2) + (r_0^4 + 8r_0^2\sigma_0^2 + 8\sigma_0^4)) < \infty.$$

Hence, Assumption (H) is satisfied.

## Notations key

$B_0$  External magnetic field of a MR unit

$M$  Net magnetization

$B_1$  Second magnetic field, induced by the RF coil

$T_1$  Relaxation time for  $M_z$  to grow from 0 to  $1 - e^{-1}$

$T_2$  Relaxation time for  $M_x$  or  $M_y$  to decay to  $e^{-1}$

$T_2^*$  "Observed"  $T_2$

**FID** Free induction decay

*in vivo* Measurements in living tissue

**ppm** unit for chemical shift

**SVS** Single voxel spectroscopy

**NAA** N-acetylaspartate

**Cr** Creatine

$a_k$  Amplitude related to the concentration of the  $k$ th sinusoid

$\phi_k$  Phase related to the  $k$ th sinusoid

$\omega_k$  Frequency related to the  $k$ th sinusoid

$d_k$  Damping constant related to the  $k$ th sinusoid

$\varepsilon$  Circular white noise

$A_k(\omega)$  Absorption signal

$D_k(\omega)$  Dispersion signal

$\sigma^2$  Variance of the background white noise

**SNR** Signal-to-Noise ratio,  $SNR = r/\sigma$

$I_p(x)$  Modified Bessel function of the first kind of order  $p$

$R(x)$   $I_1(x)/I_0(x)$

$Z$  Observed signal magnitude

$\Theta$  Observed signal phase

$r$  Unobserved signal magnitude

$\psi$  Unobserved signal phase

$\xi$   $(\psi, r, \sigma^2)$

$\xi_0$   $(\psi_0, r_0, \sigma_0^2)$  true parameter value

$\hat{\xi}$   $(\hat{\psi}, \hat{r}, \hat{\sigma}^2)$  ML estimates of  $(\psi, r, \sigma^2)$  from the log likelihood of the magnitude phase distribution