An automated quantitation of short echo time MRS spectra in an open source software environment: AQSES

Jean-Baptiste Poullet Diana M. Sima Arjan W. Simonetti
Bart De Neuter Leentje Vanhamme Philippe Lemmerling
Sabine Van Huffel

ESAT-SISTA, K.U. Leuven

Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium

corresponding author:

Dr. Ir. Sabine Van Huffel

Tel. 016/32 17 03, Fax. 016/32 19 70

Email: sabine.vanhuffel@esat.kuleuven.be

September 14, 2006

Abstract

This paper describes a new quantitation method called AQSES for short echo time magnetic resonance spectra. This method is embedded in a software package available online from `www.esat.kuleuven.ac.be/sista/members/biomed/new/` with a graphical user interface, under an open source license, which means that the source code is freely available and easy to adapt to specific needs of the user. The quantitation problem is mathematically formulated as a separable nonlinear least squares fitting problem, which is numerically solved using a modified variable projection procedure. A macromolecular baseline is incorporated into the fit via nonparametric modeling, efficiently implemented using penalized splines. Unwanted components such as residual water are removed with a maximum-phase FIR filter. Constraints on the phases, dampings and frequencies of the metabolites can be imposed. AQSES has been tested on simulated MR spectra with several types of disturbances and on short echo time *in vivo* proton MR spectra. Results show that AQSES is robust, easy to use and very flexible.

Keywords: quantitation, MR spectroscopy, short echo time, metabolites

Abbreviations:

AQSES: Accurate Quantitation of Short Echo time domain Signals

FIR: Finite Impulse Response

HLSVD-PRO: Hankel-Lanczos Singular Value Decomposition with partial reorthogonalization

PM: Performance Measure

Lip1: Lipids at 1.3 ppm

Lip2: Lipids at 0.9 ppm

1 Introduction

Accurate quantitation of metabolites from short echo time *in vivo* magnetic resonance spectroscopy (MRS), such as proton spectra from the human brain, may be a very important aid in the correct noninvasive diagnosis of pathology. For example, magnetic resonance spectroscopic imaging can be of help in brain tumour diagnosis [1], but only if accurate quantitation of the metabolites of interest can be performed [2]. The development of easy-to-use quantitation software is a challenging task, and very important for the acceptance of spectroscopy in the clinic. This paper is devoted to a new quantitation method called AQSES (Automated Quantitation of Short Echo time MRS Spectra) and its implementation in the software package AQSES GUI (graphical user interface). The functionalities of AQSES GUI are described in the appendix.

During an *in vivo* NMR experiment, the measured time-domain signal consists of responses from all metabolites (including macromolecules), noise and partially suppressed water. Spectra of metabolites that are visible during *in vivo* spectroscopy can also be measured *in vitro* or they can be simulated using quantum mechanical knowledge, and these signals can be grouped in a database of metabolite profiles. An *in vivo* time-domain signal has, in theory, the shape of a sum of complex damped exponentials. Instead of modeling these individual lineshapes, the *in vivo* short echo time MRS signal can be modeled using (a selection of) the profiles in the database, such that the prior knowledge that relates individual peaks in the *in vivo* spectrum is implicitely imposed. The quantities of interest, the metabolite *concentrations*, can be estimated from the weighting coefficients (amplitudes) of the linearly combined *in vitro* profiles. The linear combination should allow small corrections in spectral parameters like frequency shifts, damping corrections and phase shifts, as well, since these parameters may vary from measurement to measurement [3].

Any quantitation method should also take into account the presence of a partially suppressed water resonance and the presence of the macromolecular baseline. The frequency region where the water is located is known, while the macromolecular baseline overlaps with the metabolites of interest. Typically, the quantitation methods assume that the water resonance is removed in a *preprocessing step*, using a frequency selective method.

For instance, the HSVD (Hankel singular value decomposition) method is based on the idea of first fitting the whole signal as a sum of complex damped exponentials and then subtracting all the components whose frequencies were estimated in the frequency region of the water resonance.

We propose the use of an alternative technique of dealing with the presence of water (and other resonances with relatively known frequency location). The method is based on finding a maximum-phase finite impulse response (FIR) filter optimized to suppress the resonances that are outside the frequency region of interest and to keep the region of interest (where the metabolites are visible) with minimal distortions. The quantitation algorithm aims then at fitting a filtered model to a filtered *in vivo* signal. This type of filter was originally proposed for use in the quantitation of long echo time MRS signals [4]. Our results show that it can be successfully applied to short echo time signals, as well.

Methods for taking into account the macromolecular baseline are very diverse: some baseline removal methods are applied as a preprocessing step [5, 6], others during quantitation [6, 7]; they can be applied to the time-domain signal [6] or to the Fourier-transformed frequency-domain signal [7]. So far, it is not clear which type of method is the best. We propose a new approach for estimating the baseline, whose main properties are: the baseline is estimated during the quantitation itself; the fitting takes place in the time domain, thus, the baseline is also a time series; the baseline is non parametrically reconstructed using *penalized splines*; the criterion that is monitored in order to prevent the situation that the baseline reconstructs part of the metabolites is the smoothness of its Fourier transform. Among the previous methods for baseline estimation, our method is most similar to the one in LCModel [7], but we point out the following differences:

- AQSES fits the whole *complex* signal in the *time-domain*, which is the data acquisition domain; LCModel fits the *real part of the frequency-domain* signal;
- AQSES uses a FIR filtering technique to remove undesired resonances from specific frequency ranges; the baseline is also passed through this filter;
- the baseline is nonparametrically modeled using splines: in AQSES, penalized splines are chosen, while in LCModel, smoothing splines are used. The advantage of penal-

ized splines [8] compared to smoothing splines is the fact that a smaller number of spline functions are used, simplifying the computation complexity.

Another recent contribution for MRS signal quantitation is the QUEST method [3, 6] provided within jMRUI [9]. It is worth mentioning the new features of AQSES compared to QUEST:

- AQSES uses the FIR filtering technique during quantitation, while QUEST performs water suppression as a preprocessing step;
- QUEST models the baseline using heuristic methods, where several steps are involved: truncation, partial fitting, subtraction, and final fitting. AQSES uses only one common optimization problem for the fitting of both the model and the baseline. As for the water removal, we conjecture that AQSES is thus less prone to accumulated errors, since its method is direct, and not a sequence of operations.
- In reference [6], an augmented Fisher information matrix (inspired by [10]) is used for the computation of confidence bounds for the variables of interest, but it is not clear how to choose the value for the number of effective parameters involved. In this respect, the semiparametric framework discussed in [11], based on asymptotic nonlinear regression theory, clarifies the way the confidence bounds can be automatically estimated for the procedure in AQSES.

The AQSES method comes with an advanced JAVA based – platform independent – user interface, which allows use in scientific as well as clinical environments. A major difference of AQSES compared to LCModel as well as QUEST is that the source code of AQSES GUI [12] is available under the Lesser Gnu Public License (GPL), an Open Source Initiative (OSI) approved open source license. This license forces that changes made to the framework of AQSES GUI by other parties are distributed under the same license. It also allows users to create their own plug-ins and distribute them under a different license, even as closed source.

The goal of this paper is to introduce the AQSES method and the AQSES GUI framework, and to discuss the performance of this quantitation method on simulations and real

data. In section 2.1, we describe the main features of AQSES. Its structure is meant to highlight the specificities of AQSES with respect to other existing quantitation methods. Section 2.2 aims at presenting the performances of AQSES through various experiments on simulated, in vitro and in vivo MR data. Next, the features of AQSES are discussed in section 4 and some conclusions are drawn. The AQSES GUI is described in detail in the appendix.

2 Theory and methods

2.1 Mathematical formulation

2.1.1 The model

We consider that we are given a "metabolite database" as a set $\{v_k, \text{ for } k = 1, \ldots, K\}$ of complex-valued time series of length m, representing *in vitro* measured NMR responses. An *in vivo* measured NMR signal y, which is another complex-valued time series of length m, satisfies the model

$$y(t) = \widehat{y}(t) + \varepsilon_t, \quad t = t_0, \ldots, t_{m-1}, \qquad \text{with} \tag{1}$$

$$\widehat{y}(t) = \sum_{k=1}^{K} \alpha_k \left(\zeta_k\right)^t v_k(t) + b(t) + w(t), \tag{2}$$

where $\alpha_k, \zeta_k \in \mathbb{C}$ are unknown parameters that account for amplitudes of the metabolites in the database and for the necessary corrections of the database signals. The complex amplitudes α_k and the complex values ζ_k can be written as (with $j = \sqrt{-1}$):

$$\alpha_k = a_k \exp(j\phi_k), \qquad \zeta_k = \exp(-d_k + jf_k), \tag{3}$$

where a_k are the real amplitudes, ϕ_k are the phase shifts, d_k are damping corrections, and f_k are frequency shifts. Our software implementation allows a more general model to

be used: Eddy current correction terms, as well as Gaussian or Voigt line shapes can be specified. For ease of exposition, we restrict the description in this paper to the classical Lorentzian line shapes, but we refer to our technical reports [13, 14] for details on the full formulation. In (1), the term ε_t denotes an unknown noise perturbation with zero mean, and in (2), $b(t)$ denotes the "baseline", and $w(t)$ denotes the water component.

Baseline modeling

The "baseline" $b(t)$ represents the chemical part corresponding to the unknown macromolecular components that are not included in the database. We characterize $b(t)$ by the assumption that its Fourier transformation is a smooth function. For the nonparametric modeling of the baseline, we construct a basis of splines [8, 15] and put the discretized splines as columns in a matrix A. An arbitrary nonlinear function can be approximated as a linear combination of spline functions. The coefficients in this linear combination are unknowns that must be identified. We denote these linear coefficients by $c_1, \ldots, c_n$ (or by $\mathbf{c} \in \mathbb{C}^n$, when stacked in a column vector). Thus, the discretization of a nonlinear function approximated with splines (in our case, the Fourier transform of the baseline b) can be written in matrix notation as the matrix-vector product $A \cdot \mathbf{c}$. Since the goal is to reconstruct a smooth baseline in the frequency-domain, while still fitting in the time domain, we transform the spline basis matrix A to the time domain, using the discrete inverse Fourier transform.

In order to fit the model and the smooth baseline at the same time, we consider the regularized nonlinear least squares criterion

$$\min \frac{1}{m} \sum_{t=t_0}^{t_{m-1}} |y(t) - \widehat{y}(t)|^2 + \lambda^2 \|D\mathbf{c}\|^2, \tag{4}$$

where $\widehat{y}$ is given in (2), but with the baseline b replaced by the formula $b = \mathcal{A} \cdot \mathbf{c}$, which is the inverse Fourier transformation of the frequency-domain baseline $A \cdot \mathbf{c}$. In (4), the whole term $\lambda^2 \|D\mathbf{c}\|^2$ is responsible for ensuring a certain degree of smoothness to the baseline b. The regularization matrix D is needed to measure the smoothness of the baseline in the

frequency domain. We can take D as a (combination of) discrete derivative operator(s) [8]. λ is a fixed *regularization (penalty) parameter*; the value that we give to λ controls the degree of smoothness; this value can be automatically selected using a generalized cross validation criterion (see [11] for more details).

Water removal

The term $w(t)$ refers to the residual water component (as well as other possible unwanted terms), whose frequency positions are relatively known. In AQSES, this term can be *filtered out* using the maximum-phase pass-band FIR filter from [4]. However, the region(s) that should be filtered out of the *in vivo* signal should be specified by the user. Ideally, this filter will suppress all the components in the specified frequency region(s) so that the spectrum becomes there smaller than an estimated noise level, while keeping the frequency region(s) of interest undistorted.

The FIR filter in [4] is automatically optimized in order to remove the water component from an *in vivo* NMR signal. The design of the filter is performed outside the actual fitting method of AQSES. Such a filter consists of a vector of coefficients; the length of the filter and the coefficients are optimized during the automatic filter design.

A FIR filter is a linear operator that commutes with the sum; however, it does not commute with the *modified* sum of metabolites, since the shifts and corrections on the spectral parameters involve nonlinear operations. When such an operation is applied to the measured signal, it must also be taken into account by the fitting model. In other words, a filtered measured signal will be fitted with a filtered model plus a filtered baseline. Thus, the filter is used explicitly during the iterative minimization in AQSES.

2.1.2 The quantitation method

Assuming that we neglect the filter for sake of simplicity, the nonlinear least squares problem (4) becomes

$$\min_{\alpha_k,\zeta_k,\mathbf{c}} \frac{1}{m} \sum_{i=0}^{m-1} \left| y(t_i) - \sum_{k=1}^{K} \alpha_k \left(\zeta_k\right)^{t_i} v_k(t_i) - \left(\mathcal{A}\mathbf{c}\right)_i \right|^2 + \lambda^2 \|D\mathbf{c}\|^2 =$$
$$\min_{\boldsymbol{\alpha},\boldsymbol{\zeta},\mathbf{c}} \frac{1}{m} \left\|\mathbf{y} - \Phi(\boldsymbol{\zeta})\boldsymbol{\alpha} - \mathcal{A}\mathbf{c}\right\|^2 + \lambda^2 \|D\mathbf{c}\|^2, \tag{5}$$

where $\mathbf{y}$ is a column vector containing $y(t_0), \ldots, y(t_{m-1})$, $\boldsymbol{\alpha}$ and $\boldsymbol{\zeta}$ are defined as K-dimensional column vectors from the respective variables α_k, ζ_k, and the $m \times K$ matrix $\Phi(\boldsymbol{\zeta})$ has elements of the form:

$$\Phi_{ik} = \left(\zeta_k\right)^{t_i} v_k(t_i) = \exp\left((-d_k + jf_k)t_i\right) \cdot v_k(t_i), \tag{6}$$

Problem (5) is a *separable* problem, where linear parameters $\boldsymbol{\alpha}$ and $\mathbf{c}$ can be projected out of the least squares problem, and only a smaller sized nonlinear least squares problem remains to be solved for the nonlinear variables $\boldsymbol{\zeta}$, via an iterative minimization algorithm of the Levenberg-Marquardt type [16]. This technique is called *variable projection* (VARPRO [17, 18]). The VARPRO method was already used in MRS problems [19, 20]; a historical note on the application of VARPRO to MRS data quantitation is given in Section 17 of the review paper [18]. The previous use of VARPRO in the above mentioned papers was restricted to long echo-time MRS signals and models of the type "sum of complex damped exponentials."

In the AQSES framework, the classical VARPRO method is modified [14] such that it is possible to impose *prior knowledge* in the form of upper and lower bounds on the nonlinear parameters, linear equalities between some variables of the same sort among d_k or f_k, or even some constraints on the linear parameters (non-negative amplitudes and equal phases). Normally, VARPRO allows the complex amplitudes α_k to take any complex value such that the residual fit is minimized. In terms of the real parameters, this means that the real amplitudes a_k could take any non-negative value, while the phase corrections ϕ_k would be unconstrained between $-\pi$ and π. AQSES allows an option of *equal phase corrections* for each metabolite, since it assumes that the database of metabolites is already constructed

to have metabolites in phase. In terms of the VARPRO implementation, the equal phase constraint is a difficult constraint, since it involves part of the linear variables. Sima and Van Huffel [14] deal with this problem by adding one common phase correction variable to the nonlinear set of variables, and effectively imposing a *non-negativity* condition on the real-valued amplitudes a_k. The solution is numerically computed thanks to a nonnegative least squares method instead of closed-form least squares.

The VARPRO method implemented in AQSES is much more efficient than optimizing the nonlinear least squares problem (4) directly, over all linear and nonlinear parameters without separation. This improvement in computation efficiency is more important than in the previous use of VARPRO for long echo time MRS [19], since there are many more linear parameters than nonlinear ones, as many as there are spline coefficients. The number of spline coefficients is typically one tenth of the number of data points. Another advantage of VARPRO is that it does not encounter numerical problems when some amplitudes a_k are nearly zero. Moreover, the Levenberg-Marquardt algorithm needs good initial values for its variables; we found that good initial values for all the nonlinear parameters – frequency and damping corrections – are zeros, which means that we start the optimization with no spectral corrections to the signals in the database. In the equal phase case, an initial guess for the common phase corrections is estimated from a preliminary optimization round of the free phase algorithm.

Cramer-Rao bounds, specially adapted for semiparametric nonlinear regression [11], can also be computed as a by-product of the quantitation procedure. They correspond to all the spectral parameters for the metabolites of interest (linear and nonlinear parameters, as well) and give an indication about the uncertainty of the final quantified parameters. If the given bounds are small enough relative to the corresponding parameter value, then it means that the computed value is reliable. If a large bound is found for a certain component, then the computed parameters might be unreliable. This is not due to a faulty minimization process; instead, it is due to poor signal-to-noise ratio, incomplete database of metabolites, or the inclusion of metabolites having similar spectra resulting in convergence problems. The risk of converging to local minima instead of global minima

is indeed higher with a poor signal-to-noise ratio. The Cramer-Rao bounds estimated in AQSES are proportional to the variance of the residue plotted in AQSES GUI. Therefore, it might be also helpful to inspect the residue to detect convergence problems.

2.2 Simulated examples and *in vivo* quantitation

The in vitro database used in this paper is described in the appendix.

Five experiments have been designed to test the robustness and the accuracy of AQSES. The **first** one emphasizes on the accuracy level that can be reached for the parameter estimation. A large amount of simulated signals were created for that purpose. Each simulated signal consisted of a linear combination of 8 metabolite profiles (6 metabolite profiles + 2 simulated lipid profiles) in the basis set, *i.e.*, myo-inositol (Myo), creatine (Cr), phosphocholine (PCh), glutamate (Glu), NAA, lactate (Lac), lipid at 1.3 ppm (Lip1), and lipid at 0.9 ppm (Lip2). These metabolites were chosen since they have been found to be important in the detection of many pathologies. Note that no disturbance components such as baseline, noise or water, were added to the simulated signals. The parameters for amplitude, damping, phase and frequency for each simulated signal were chosen in the following way: first, meaningful parameters were estimated from a set of 98 short echo time *in vivo* MRSI spectra acquired from normal brain tissue. Estimations were obtained using AQSES in a controlled way in which results were visually inspected and outliers removed. This was done to mimic a real world situation and to obtain sensible mean values and standard deviations (SD) for all parameters. Then, for the simulated signals, the amplitudes were restricted to the mean ±3*SD (only positive), the damping perturbation was restricted to the mean ±10 Hz. The phases were set variable between -45 and 45 degrees, but were the same for all profiles in one simulated signal. The frequency shifts were restricted to zero ±4.5 Hz. This resulted in a set of data (set 1) with 200 simulated spectra.

For each metabolite k in the simulated signal l, the true amplitudes ($a_{k,l}$) were compared to the estimated ones ($\tilde{a}_{k,l}$) obtained with AQSES by means of a performance measure (PM) defined as:

$$PM_k = 100\sqrt{\frac{\sum_{l=1}^{L}(a_{k,l} - \tilde{a}_{k,l})^2}{\sum_{l=1}^{L} a_{k,l}^2}}, \tag{7}$$

where L is the number of simulated signals in each set. A low PM reflects a high performance, and is a percentage measure of the difference between estimated and true amplitudes.

The **second** experiment extends the results of the first one for larger databases, *i.e.*, with more metabolite profiles. We took the same simulated spectra (set 1) as in the first experiment, but three more metabolite profiles were added to the basis set: taurine (Tau), alanine (Ala) and glucose (Glc). We chose these metabolites because they are known to be important metabolites that do not have strong correlation with the metabolites that were already inside the basis set.

The **third** experiment shows the influence of water, baseline and noise on the estimated amplitudes. We considered the same basis set as in the first experiment (using 8 metabolite profiles in AQSES). Five sets of simulated signals were constructed:

set 2 = set 1 with water resonance at 4.7 ppm,

set 3 = set 1 with low white noise (SNR = 25),

set 4 = set 1 with high noise (SNR = 7),

set 5 = set 1 with baseline distortion

set 6 = set 1 with water, baseline and high noise.

The baseline distortion was based on information from Table 1 in [21]; the baseline is the sum of gaussians referred to as lip3, lip4, lip5, mm2, mm3 and mm4 in that paper. The water profile has been extracted from an *in vivo* spectrum by means of HLSVD-PRO. The SNR is defined as the ratio of the reference peak height at 8.44 ppm and the standard deviation of the circular[1] white gaussian noise, both in the frequency domain.

[1] equal variances for real and imaginary part

To determine the influence of water, baseline and noise, the PM has been studied. As illustration, we plotted a simulated spectrum from set 4 and from set 6 in Figure 1.

In the **fourth** experiment, AQSES was validated using an *in vitro* sample. This test solution contained 9 metabolites of known concentrations: Cr, NAA, Glu, Gln, Myo, PCh, Glycerphosphocholine (GPCh), Tau and Lac. The basis set used in AQSES was: Cr, NAA, Glu, Myo, PCh, Tau and Lac. Glu was selected to fit the combination of Glu and Gln, since Glu and Gln have similar profiles and are highly correlated. Glu will be denoted by Glx in this experiment and in the last experiment to indicate that it fits the combination of Glu and Gln. The combination of PCh and GPCh was fitted similarly with PCh. The true and estimated proportions of metabolites have been compared, the proportion of metabolite k being the ratio of the concentration of metabolite k and the total concentration (all metabolites).

In the **last** experiment, *in vivo* NMR signals from a database containing MRSI spectra from normal tissue (122 spectra selected from data of four volunteers), gliomas of grade II (GII, 90 spectra selected from data of six patients) and Glioblastoma multiforma (GBM, 59 spectra selected from data of five patients) were processed. The goal of this part is to show that AQSES provides results in accordance to the literature. For the specific acquisition procedure and settings, see [22]. The basis set was identical to the one in the first experiment. We also display metabolic images obtained from a patient with a GBM (I-1285 from the INTERPRET database [23]) that underwent spectroscopic imaging. The metabolic images have been obtained by processing all spectra of the patient in batch mode by AQSES, using the same basis set as before.

The first three experiments are carried out assuming either no prior knowledge of the phases or that the metabolites are in phase. The last two experiments did not assume any prior knowledge of the metabolite phases.

3 Results

3.1 Robustness of AQSES

Figure 2 shows a boxplot of the distribution of the amplitudes found by AQSES for the *in vivo* spectra of healthy tissue that were used to estimate meaningful values for the simulated signals. The means are sensible values with respect to the literature (see, *e.g.*, [24]). The large variances may result from the fact that the spectra of healthy tissue originate from different brain locations, which have an influence on the metabolite concentration as shown in [24].

The **first** experiment shows that, in 95.5% of the cases, the parameter estimates are almost perfect (*i.e.*, the relative difference between the true and the simulated metabolite amplitudes is under 10^{-6}%). This result holds whether the metabolites are assumed to be in phase or not. Looking at each metabolite separately, the percentages of cases where the relative errors are higher than 1% are quite similar for both assumptions (rows 1 and 3 in Table 1). Table 1 indicates that Lac is misfit in most of the remaining cases (4% out of 4.5%). In the cases of imperfect parameter estimates, the relative error for NAA never exceeds 1% when we assume equal phases. In 4% out of 4.5% cases, the estimation is corrupted by the value of Lac which overlaps Lip1, resulting in larger amplitude estimation errors. Tables 2 and 3 report the PM corresponding to each set and to each metabolite, imposing respectively equal phases for all metabolites and letting the phases free. These values are obtained after discarding the signals that do not provide perfect parameter estimates (defined above), *i.e.*, 9 signals out of 200 (same signals in both tables). The first row of these tables shows that a higher degree of accuracy can be reached when equal phases are assumed.

The **second** simulated example shows the results when the number of metabolites in the basis set and in the simulated spectra are not the same. The percentage of almost perfect parameter estimates decreases to 92% when assuming the metabolites in phase, while it decreases to 92.5% when no assumptions about the phases are made. However, imposing equality of phases improves the accuracy of the fit (Tables 2 and 3, rows 2)

when considering the 191 signals that provide perfect parameter estimates in the first experiment. The goodness of the fit is now complicated by the overlap of metabolites such as Myo, Glc, PCh and Tau. Rows 2 and 4 in Table 1 illustrate the loss in accuracy compared to rows 1 and 3; when Glc, Ala and Tau are added to the basis set. The other metabolites, located at other frequency ranges, are less affected by the addition of Tau, Ala and Glc to the basis set. This shows that AQSES is not very sensitive to the choice of the basis set provided its components are not strongly correlated.

The **third** experiment investigates the robustness of AQSES against the addition of disturbance components such as noise, baseline and water resonances. The PM are reported in Tables 2 and Tables 3 for each simulation set (last four rows).

The extreme values are mainly present in the overlapping peaks, *i.e.*, Lac and Lip1. The maximum-phase filter removes satisfactorily the water component, if we inspect the PM in set 2 in Tables 2 and 3 (row 3). The PM for different noise levels confirms the stability of AQSES against noise. At low noise values, the PM does not increase dramatically for all metabolites, except for Lac, Lip1 and Lip2. At high noise, the PMs of Lip1 and Lip2 keep increasing strongly, while the PM of Lac is more stable. The fact that these metabolites are estimated with a lower accuracy may be explained by their low magnitude level. They are more embedded in noise than the other ones , or in other words, they have a smaller relative signal-to-noise ratio. The PMs of the 5 main metabolite profiles show stability against noise (columns 1 to 5 in Tables 2 and 3). The baseline affects each component, but mainly Myo and Glu. Myo and Glu are wider and therefore are more likely to be fitted by the baseline. Cr is less affected by the baseline. The PM of NAA, Cr and PCh remain under 10% in all sets with the assumption of equal phases (Table 2) and under 11% without this assumption (Table 3). The largest PM value is kept under 30% in Table 2 and under 33% in Table 3. In most of the cases, a slightly better accuracy is obtained by assuming that the metabolites are in phase.

Since these three experiments involve different models, we have studied the computation time required for each type of model. The results are reported in Table 4. For experiment 1, and 3 (sets 2, 3 and 4), we do not consider the baseline in the model and

the number of metabolite profiles in the basis set equals 8 (model 1 in Table 4). For experiment 2, the baseline is still not included in the model but the number of metabolite profiles in the basis set increases to 11 (model 2). For experiment 3 (sets 5 and 6), the baseline is incorporated in the model which takes into account the contribution of 8 metabolite profiles. The computation times of AQSES using 2 different assumptions have been compared: imposing equal phases for all metabolites or letting the phases free. The values reported in Table 4 are the average computation times for one spectrum. AQSES has been run on a Windows XP platform with a Pentium 4 (3 GHz CPU, 1 Gb RAM). The simplest case in terms of computation time is for model 1 when the phases are free with an average quantitation time per spectrum of 0.87 s (2'45" for 191 spectra). Assuming equal phases for all metabolites increases the computation time if the baseline is not considered in the model. This was expected since one variable (common phase) is added to the K real variables used for the amplitudes, while K complex variables are used for the complex amplitudes when the phases are free. Moreover, the amplitudes are restricted to positive numbers since negative amplitudes would mean a metabolite in opposite phase, contradicting the "equal phase" constraint (see [13] for more information). In the slowest case, *i.e.* the case when the database of metabolites is larger (11 metabolites), the computation time per spectrum is 4.03 s (12'49" for 191 spectra), which remains relatively small.

For the **fourth** experiment, the true and estimated proportions of the amplitudes of the metabolites in the test sample are reported in Figure 3. The relative errors between the true and the estimated proportions are relatively small (all $< 26\%$) and especially for the metabolites in higher concentration ($< 8\%$). Note that the estimation of Glx is quite good ($\approx 6\%$) while the individual estimates of Glu and Gln are not as good if we include their corresponding metabolite profile in the database ($\approx 25\%$, data not shown). This is due to the high degree of correlation between these two metabolite profiles. On the contrary, summing up the contributions of Glu and Gln for estimating Glx yields good results (also $\approx 6\%$, data not shown).

For the **last** experiment, we show in Figure 4 the averaged estimated amplitudes

obtained with AQSES, for 8 metabolites of the three classes normal, grade II and GBM. NAA and Cr decrease with the grade of the tumor, in agreement with [25]. A large difference occurs between the concentration of NAA in normal and glioma tissues. Myo, PCh and Glx exhibit much less variations. We note a small increase in Glx and PCh with the grade of glioma. When ordering comparable metabolites (*i.e.*, without Lip1 and Lip2, see Section 2.2) from the highest to the lowest in concentration, we have

for normal brain: NAA, Glx, Cr, Myo, Lac and PCh;

for GII: Glx, Lac, Cr, NAA, Myo, PCh; and

for GBM: Lac, Glx, PCh, Cr, NAA and Myo.

We note that Lac, Lip1 and Lip2 strongly increase with the grade of the tumor.

Metabolic images such as in Figure 5 can also be displayed. The computation time required by AQSES to provide the quantitation results is 2'50" (Windows XP, 1 Gb RAM, 3 GHz CPU). In this example, the affected region is clearly visible on the T2-weighted image (Figure 5a). The tumor region can also be identified from almost all metabolic images (see Figure 5 (d, f, g, h, and i)). The concentrations of NAA and Cr are lower in the tumor region, while the concentrations of Lac, Lip1 and Lip2 are larger in that region, which agrees with Figure 4 and with the literature (see, *e.g.*, [1] and [26]). Myo decreases in the tumor, but this contrast is less apparent than in the case of NAA or Cr. Also noticeable is the low Glx concentration in the region of the ventricle compared to the central part of the brain (ellipsoid in Figure 5e). An increase of PCh, present in the boundary between the malignant and healthy areas (ellipsoid in Figure 5e), could reflect an increase of membrane synthesis and accelerated cell proliferation (see, *e.g.*, [26]). The ellipsoids in Figure 5(g) and (i) correspond to a region that is slightly higher in intensity in the T2-weighted image. This region has also an increased concentration for Lip2 and Lac. The presence of Lip is thought to correspond to cellular and membrane breakdown corresponding to necrosis [1]. Lac is usually detected only under pathologic conditions, when energy metabolism is affected severely [27]. This shows that pathologic information, which is barely visible on MRI images, can be successfully extracted from metabolic maps.

4 Discussion

The VARPRO implementation of AQSES, using a modified Levenberg-Marquardt minimization algorithm, is important for fast convergence and reliable numerical computations and is less affected by local minima, since no starting values are needed for the linear parameters (including the parameters of the baseline). The nonlinear parameters (*e.g.*, frequency shifts or damping corrections) are initialized with zeros, which is the most reasonable choice in the context of short echo time quantitation. The classical VARPRO method leaves the linear part of the model completely free such that no constraints can be imposed on the amplitude and phase. But, in theory, the phases of the metabolites are equal. Constraining the phases to be equal implies some modifications of the classical VARPRO implementation as described in section 2.1. The results show that, in spite of large variations on all the variables to construct the simulations, a high accuracy level can be reached thanks to this algorithm (95.5% of the cases are almost perfect for noise-free signals).

Another important aspect is the choice of the basis set in AQSES. The basis set should contain the metabolites actually visible in the spectrum except for those that are too correlated to each other. Indeed, equivalent metabolite profiles can ruin their own individual parameter estimates as we have seen in the experiment 4 when Glu and Gln were both in the database. When the basis set contains more metabolites than actually contained in the spectrum, we showed that the amplitude estimates of the metabolites which were not actually in the signal (*i.e.*, Tau, Ala and Glc), were often close to zero (see experiment 2). These results indicate an added value of the VARPRO method implemented in AQSES over other methods. However, adding more metabolites in the basis set will increase the

complexity of the algorithm and consequently the computation time. The results show that the latter could be more than doubled by adding 3 metabolite profiles to the basis set of 8 profiles. Furthermore, the parameter variances will also increase, limiting the accuracy of the algorithm. In case of doubt about the presence of some metabolites, we recommend to add them to the basis set if they do not exhibit profiles too similar to the metabolite profiles already entered in the basis set. Large Cramer-Rao bounds displayed in AQSES GUI or the presence of obvious peaks in the plot of the residue may be an indicator of an incomplete basis set. In this respect, the basis set of 8 metabolite profiles used in this paper provides a good compromise although we do not claim that it is the best choice for any type of MRS spectra to be fitted. The protocol of the basis set should be similar to the one used to acquire the spectra to be fitted such that the parameter corrections (α_k, ζ_k) can be assumed as small, reducing the risk of convergence to a local minimum.

One of the underlying goals of a quantitation method is to separate the signal from the disturbance components. In this respect, the maximum-phase FIR filter provides satisfactory results. Its high efficiency allows its embedment in the iterative minimization in AQSES, resulting in more accurate parameter estimates. This filter, in addition to filtering out the unwanted components (including the noise) in the frequency region of no interest, removes partially the baseline. This reduces the importance of the regularization parameter, λ, that may influence substantially the results. The use of a semi-parametric model reduces the number of steps to obtain the parameter estimates, decreasing their variances and thus increasing the reliability of the results. The baseline remains a crucial issue during fitting due to its large correlation with the other metabolites. Our results show that (compared to the results of data set 1) the errors for all metabolites increase

substantially, with a smaller increase for Cr. This seems to be in agreement with the observation of Ratiney *et al.* [6], who noted that the background amplitude was correlated with all metabolite amplitudes except that of Cr. The algorithm used in this paper is robust against noise since the least squares problem that is solved is restricted to a linear combination of corrected metabolite profiles present in the signal; using metabolite profiles instead of individual peaks implies an increased robustness against noise, because correlations between spectral regions that are relatively far apart are taken into account.

5 Conclusion

In this paper we have described a new short echo time MR quantitation method AQSES embedded in a user-friendly software package. It provides more flexibility in exploiting prior knowledge than the classical VARPRO method without sacrifying its advantages in numerical accuracy and computational efficiency. The use of the maximum-phase FIR filter allows more robustness against disturbance components such as water, noise or baseline. Our method generates fast and accurate results on simulated data and *in vitro* samples. The results of batches of *in vivo* data that contain spectra from healthy, grade II gliomas and glioblastomas are in line with results described in literature. The results of a set of spectroscopic imaging data is in accordance with MRI images obtained from the same slice. Since the GUI is open source and written in a flexible language it has the potential to grow rapidly and be of high importance for the medical MR community.

6 Acknowledgements

The Biomedical Magnetic Resonance Research Group Radboud University Nijmegen Medical Center (`http://get.to/mrs`) and the EU funded projects BIOPATTERN (EU network of excellence; Contract No. FP6-2002-IST 508803), INTERPRET (EU shared-cost RTD project; Contract No. FP5-IST-1999-10310), eTUMOUR (FP6-2002-LIFESCIHEALTH; Contract No. 503094) and HEALTHagents (IST200427214) are gratefully acknowledged. Arend Heerschap and his group (Radboud University Nijmegen, Medical Center) are gratefully acknowledged for providing data. Lutgarde Buydens and her group are (Radboud University Nijmegen, Faculty of sciences) are gratefully acknowledged for having preprocessed the data.

Research supported by

Research Council KUL: GOA-AMBioRICS, CoE EF/05/006 Optimization in Engineering, several PhD/postdoc & fellow grants;

Flemish Government: FWO: PhD/postdoc grants, projects, FWO-G.0321.06 (Tensors/Spectral Analysis), G.0360.05 (EEG, Epileptic), G.0519.06 (Noninvasive brain oxygenation), research communities (ICCoS, ANMMM); **IWT:** PhD Grants;

Belgian Federal Science Policy Office: IUAP P5/22 ('Dynamical Systems and Control: Computation, Identification and Modelling');

A AQSES GUI

The AQSES framework

To increase the usability, AQSES has been embedded in an application framework which controls the optimization method [12]. In general, an application framework provides a

set of abstract classes and interfaces, called hot spots that can be combined to create an application. For AQSES, the framework provides hot spots for plug-ins to read or write, to preprocess, to quantitate, to export and to visualize MRS and MRSI data. The software is developed in Java and is platform independent (*e.g.*, Windows, Linux, OSX and Solaris). The graphical user interface (GUI) of the framework is written using the Swing library. AQSES and the HLSVD-PRO algorithms are implemented in FORTRAN 77. Both methods have been compiled for Windows and Linux and have been made accessible to the Java code, to make sure that the software can run on both platforms. The main optimization part of AQSES is carried out using an extension of the Levenberg-Marquardt algorithm, which accepts linear bounds constraints, in the DN2GB implementation written by David M. Gay [28]. FORTRAN is used since well-tested and optimized numerical libraries (*e.g.*, LAPACK, Blas) exist.

At the moment the AQSES GUI accepts MRS and MRSI data coming from

- Philips (SDAT/SPAR format),
- SIEMENS (RDA format),
- Matlab (The MathWorks, inc., Version 4) and
- text files.

It is easy to add plug-ins to read or write other formats such as DICOM by implementing some of the hot spots. After loading, the data can be visualized in the time or frequency domain and is presented to the user in a 3D graph. It is also possible to work with multiple data sets at the same time, which are stored in projects.

The AQSES GUI offers two types of projects, one called *AQSES database creator* to create personal databases, and another called *MRS* to process the signals. Any metabolite

profile can be loaded in the database. There is no restriction about the acquisition protocol or about the number of metabolite profiles in the database. All preprocessing tools are available in both projects, while the quantitation methods are only available in the *MRS* project. The preprocessing methods included at the moment in AQSES GUI are

- Eddy current correction (Klose [29]),
- HLSVD-PRO [30],
- zero filling,
- point truncation, and
- manual phase and frequency correction.

The quantitation methods available in the AQSES GUI at the moment are

- AQSES and
- HLSVD-PRO.

AQSES and HLSVD-PRO are the only methods in AQSES GUI whose codes are not open source. AQSES incorporates model functions with three possible lineshapes (Lorentzian, Gaussian and Voigt). Although this is implemented in the FORTRAN code, the user has presently only access via AQSES GUI to the Lorentzian lineshape, since the Gaussian and Voigt lineshapes have not been fully tested yet. The Eddy current correction term discussed in section 2.1 is also not available yet in AQSES GUI. Therefore, one should use some preprocessing methods (phase correction, Eddy current correction, etc) before applying AQSES. However, these preprocessing techniques might be not recommendable if the basis set of metabolite profiles and the MR spectra were acquired with the same

acquisition protocol such that it is reasonable to assume that they have undergone the same pattern of distortions.

The AQSES GUI is template driven, which means that scripts can be built that contain the settings for preprocessing and quantitation. Templates can be written from scratch or generated automatically in AQSES GUI when processing the signals. These templates can be saved and used to easily redo experiments or to start processing in batch mode. The GUI is intended to run in two different modes. The *clinical mode* hides most of the complexity by using templates, while the *research mode* provides the user with all features. In the *clinical mode*, the user is restricted to

- load, read, visualize and save data,
- process data **using existing templates**,
- visualize and save results.

The results of the quantitation methods are visualized in a separate window and can be exported to a Matlab file.

The database used in this paper

The *in vitro* metabolite profiles in this paper have been acquired on a 1.5 T Philips NT Gyroscan using a PRESS sequence with an echo time of 23 ms, and a PRESS box of $2 \times 2 \times 2\,\mathrm{cm}^3$. To each sample, two reference compounds were added, situated at 0.0 ppm (3-trimethylsilyl-1-propane-sulfonic acid or TSPS) and at 8.44 ppm (formate). One buffer solution of K2HPO4 and KH2PO4 was made including TSPS, formate, phosphate buffer compounds (except for the choline metabolite solutions) and NaOH for pH adjustment (brought to 7.2 pH). The solution was divided in small identical quantities, in which

the relevant metabolites were dissolved. All raw FIDs were Eddy current corrected using Klose's method in the "Database Creator" of the AQSES GUI, and the spectra were frequency shifted such that the second reference peak was at 8.44 ppm. All profiles were normalized or scaled with respect to the 8.44 ppm resonance in the sample of creatine, to ensure that absolute comparison between metabolites is possible. This was done as follows. All spectra were first filtered with HLSVD-PRO to keep only the reference peaks at 8.44 ppm. These were fitted using AQSES with the one of the creatine profile. Each metabolite profile was then normalized with respect to the amplitudes obtained in AQSES since the metabolite profiles should have the same reference peak at 8.44 ppm, its concentration being identical for all metabolite solutions. All spectra were finally normalized with respect to the concentration in the sample during the acquisition, the metabolite concentration being different from one metabolite solution to another. Metabolite profiles of lipids at 1.3 ppm and 0.9 ppm have been artificially created from the creatine resonance. Therefore, absolute values of these metabolites cannot be compared with the other metabolites. Their simulation was performed by removing the creatine signal from the spectrum with HLSVD-PRO, followed by addition of a single resonance at 1.3 or 0.9 ppm. Finally, all metabolite profiles have been processed by HLSVD-PRO (from 4.4 ppm to 7.0 ppm) to remove the residual water resonance and some artifacts within that region. As example, the N-acetylaspartate (NAA) profile is plotted in Figure 6.

At present, two databases are available online from

`www.esat.kuleuven.ac.be/sista/members/biomed/new/`:

Database 1: 16 metabolite profiles acquired on a 1.5 T Philips NT Gyroscan using a PRESS sequence with an echo time of 23 ms, and a PRESS box of $2 \times 2 \times 2\,\mathrm{cm}^3$ + simulated profiles for Lip1 and Lip2.

Database 2: 16 metabolite profiles acquired on a 1.5 T Siemens using a STEAM sequence with an echo time of 20 ms, and a STEAM box of $2 \times 2 \times 2\,\mathrm{cm}^3$ + simulated profiles for Lip1 and Lip2.

A document is also available, which describes the generation and the preprocessing of these databases. Note that AQSES can be combined with any other database, either generated via simulation or by in vitro acquisition. The databases given here are just representative examples.

References

[1] Nelson SJ. Magnetic resonance spectroscopic imaging. *IEEE Eng. Med. Biol.* 2004; **23**: 30–39.

[2] Howe FA, Barton SJ, Cudlip SA, Stubbs M, Saunders DE, Murphy M, Wilkins P, Opstad KS, Doyle VL, McLean MC, Bell BA, Griffiths JR. Metabolic profiles of human brain tumours using quantitative in vivo ^{1}H magnetic resonance spectroscopy. *Magn. Res. Med.* 2003; **49**: 223–232.

[3] Ratiney H, Sdika M, Coenradie Y, Cavassila S, van Ormondt D, Graveron-Demilly D. Time-domain semi-parametric estimation based on a metabolite basis set. *NMR Biomed.* 2004; **17**: 1–13.

[4] Sundin T, Vanhamme L, Van Hecke P, Dologlou I, Van Huffel S. Accurate quantification of ^{1}H spectra: from FIR filter design for solvent suppression to parameter estimation. *J. Magn. Reson.* 1999; **139**: 189–204.

[5] Golotvin S, Williams A. Improved baseline recognition and modeling of FT NMR spectra. *J. Magn. Reson.* 2000; **146**(1): 122–125.

[6] Ratiney H, Coenradie Y, Cavassila S, van Ormondt D, Graveron-Demilly D. Time-domain quantitation of ^{1}H short echo-time signals: Background accommodation. *MAGMA*. 2004; **16**: 284–296.

[7] Provencher SW. Estimation of metabolite concentrations from localized *in vivo* proton NMR spectra. *Magn. Res. Med.* 1993; **30**(6).

[8] Eilers PHC, Marx BD. Flexible smoothing with B-splines and penalties. *Stat. Sci.* 1997; **11**(2): 89–121.

[9] Magnetic Resonance User Interface in JAVA (jMRUI). http://www.mrui.uab.es/mrui/.

[10] Spall JC, Garner JP. Parameter identification of state-space models with nuisance parameters. *IEEE T. Aero. Elec. Sys.* 1990; **26**(6): 992–998.

[11] Sima DM, Van Huffel S. Regularized semiparametric model identification with application to NMR signal quantification with unknown macromolecular base-line. *J. Roy. Stat. Soc. B.* 2006; **68**(3), to appear.

[12] De Neuter B, Luts J, Vanhamme L, Lemmerling P, Van Huffel S. Java-based framework for processing and displaying short-echo-time magnetic resonance spectroscopy signals. *Comput Methods Programs Biomed*, to appear.

[13] Sima DM, Van Huffel S. AQSES$_{VP}$ - description of a variable projection implementation for nonlinear least squares with linear bounds constraints, applied to accurate quantification of short-echo time magnetic resonance spectroscopic signals. Technical Report 05-120, ESAT-SISTA report, ESAT Laboratory, K.U. Leuven, Belgium. 2005.

[14] Sima DM, Van Huffel S. Separable nonlinear least squares fitting with linear bound constraints and its application in magnetic resonance spectroscopy data quantification. *J. Comput. Appl. Math.* 2006, to appear.

[15] de Boor C. *A Practical Guide to Splines*. Springer, Berlin. 1978.

[16] Moré J. The Levenberg-Marquardt algorithm: Implementation and theory. In GA Watson (ed.) *Numerical Analysis: Proceedings of the Biennial (Lecture Notes in Mathematics)*, Springer Verlag, Dundee. 1978; 104–116.

[17] Golub GH, Pereyra V. The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate. *SIAM J. Numer. Anal.* 1973; **10**: 413–432.

[18] Golub GH, Pereyra V. Separable nonlinear least squares: the variable projection method and its applications. *Inverse Probl.* 2003; **19**(2): 1–26.

[19] van der Veen JWC, de Beer R, Luyten PR, van Ormondt D. Accurate quantification of *in vivo* PNMR signals using the variable projection method and prior knowledge. *Magn. Res. Med.* 1988; **6**: 92–98.

[20] Vanhamme L, Sundin T, Van Hecke P, Van Huffel S. MR spectroscopic quantitation: a review of time domain methods. *NMR Biomed.* 2001; **14**: 233–246.

[21] Seeger U, Klose U. Parameterized evaluation of macromolecules and lipids in proton MR spectroscopy of brain diseases. *Magn. Res. Med.* 2003; **49**: 19–28.

[22] Simonetti AW, Melsen W, van der Graaf M, Postma G, Heerschap A, Buydens LMC. A chemometric approach for brain tumor classsification using magnetic resonance imaging and spectroscopy. *Anal. Chem.* 2003; **75(20)**: 29–36.

[23] International network for pattern of tumours using magnetic resonance. http://carbon.uab.es/INTERPRET/interpret_home.shtml.

[24] Tong Z, Toshiaki Y, Yamaki T, Harada K, Houkin K. In vivo quantification of metabolites in normal brain and brain tumors by proton MR spectroscopy using water as an internal standard. *Magn. Reson. Imaging.* 2004; **22**: 1017–1024.

[25] Fountas KN, Kapsalaki EZ, Gotsis SD, Kapsalakis JZ. In vivo proton magnetic resonance spectroscopy of brain tumors. *Stereot. Funct. Neuros.* 2000; **74**: 83–94.

[26] Fan G, Sun B, Wu Z, Guo Q, Guo Y. In vivo single-voxel proton MR spectroscopy in the differentiation of high-grade gliomas and solitary metastases. *Clin. Radiol.* 2004; **59**: 77–85.

[27] Furuya S, Naruse S, Ide M, Morishita H, Kizu O, Ueda S, Maeda T. Evaluation of metabolic heterogeneity in brain tumors using ^{1}H-chemical shift imaging method. *Neurosurgery.* 1997; **10**: 25–30.

[28] Gay DM. A trust-region approach to linearly constrained optimization. In DF Griffiths (ed.) *Numerical Analysis Proceedings*, Springer-Verlag, Dundee. 1983; 72–105.

[29] Klose U. In vivo proton spectroscopy in presence of Eddy currents. *Magn. Res. Med.* 1990; **14**: 26–30.

[30] Laudadio T, Mastronardi N, Vanhamme L, Van Hecke P, Van Huffel S. Improved Lanczos algorithms for blackbox MRS data quantitation. *J. Magn. Reson.* 2002; **157**: 292–297.

Tables

Table 1: Percentages of cases (for each metabolite) for which the relative differences between the true and the estimated metabolite amplitudes is smaller than 1% (based on 200 signals for each set). The top (resp. bottom) two rows correspond to the equal phase constraint (resp. no constraints on the phases). The notations are explained in Section 2.2.

	Myo	PCh	Cr	Glu	NAA	Lac	Lip1	Lip2
exp. 1, Set 1	98.5	99	98.5	98.5	100	96	96.5	97
exp. 2, Set 1	92.5	95	95.5	97	98.5	93.5	94.5	96
exp. 1, Set 1	98.5	97.5	98	98	99.5	96	96	97
exp. 2, Set 1	94	94.5	97.5	95	98.5	93.5	94	95

Table 2: Performance measure (PM) for each metabolite of each simulation set (in percentage, see Equation 7 with $L = 191$). Equality of all metabolite phases has been imposed.

	Myo	PCh	Cr	Glu	NAA	Lac	Lip1	Lip2
exp. 1, Set 1	4.1E-12	1.4E-12	1.9E-12	1.35E-11	1.5E-12	1.81E-11	1.6E-12	1.08E-10
exp. 2, Set 1	1.72	1.21	0.07	0.09	0.01	0.65	0.77	0.19
exp. 3, Set 2	1.24	0.93	0.39	0.67	0.20	6.59	7.91	0.81
exp. 3, Set 3	5.76	4.04	2.87	4.33	2.48	15.18	19.85	6.89
exp. 3, Set 4	8.66	7.07	4.95	6.65	5.67	20.10	29.25	14.74
exp. 3, Set 5	9.51	2.67	1.76	24.39	6.25	10.79	14.01	6.26
exp. 3, Set 6	14.73	8.36	6.44	25.04	9.81	23.53	29.14	17.34

Table 3: Performance measure (PM) for each metabolite of each simulation set (in percentage, see Equation 7 with $L = 191$). No prior knowledge has been imposed.

	Myo	PCh	Cr	Glu	NAA	Lac	Lip1	Lip2
exp. 1, Set 1	6.95E-07	1.92E-07	1.01E-07	6E-07	2.07E-07	8.12E-07	8.51E-07	6.31E-07
exp. 2, Set 1	2.67	3.53	0.44	0.14	0.07	1.34	1.24	0.30
exp. 3, Set 2	1.20	1.40	0.35	0.67	0.46	10.47	6.48	1.26
exp. 3, Set 3	7.37	5.88	4.25	4.27	3.23	19.17	20.01	7.71
exp. 3, Set 4	10.81	9.77	6.22	7.74	5.97	23.00	34.44	15.32
exp. 3, Set 5	10.06	7.68	2.90	29.41	7.97	19.15	14.07	9.76
exp. 3, Set 6	13.45	11.00	7.41	29.71	10.44	24.65	32.44	18.28

Table 4: Average computation times (in seconds) per spectrum for each model, imposing equal phases for all metabolites or letting them free.

Models \ Prior knowledge	Equal phases	Free phases
Model 1: exp. 1, exp. 3 (Sets 2, 3, 4)	1.39	0.87
Model 2: exp. 2	4.03	2.74
Model 3: exp. 3 (Sets 5, 6)	3.29	3.32

Figure Captions

Figure 1. Simulated spectra from set 4 and set 6. The amplitudes are in arbitrary units.

Figure 2. Boxplot showing the quantitated amplitudes of *in vivo* spectra from healthy volunteers.

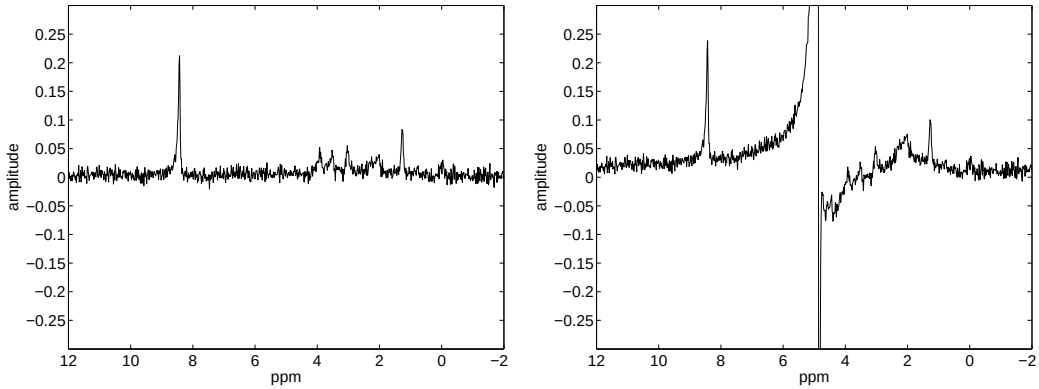
Figure 3. *In vitro* test sample results.

Figure 4. Averaged estimated amplitudes for normal tissues, GII and GBM, in arbitrary units.

Figure 5. Metabolic maps of a patient with a glioblastoma. The ellipsoids in (c), (g) and (i) indicate regions where we note a substantial increase of the corresponding metabolites compared to a region with normal tissue (high concentrations in red and low concentrations in blue with a linear scale). This information may particularly be useful since it barely appears on the T2-weighted image. The concentration of Glu is lower in the ventricles (one of the ventricle regions is indicated by the left ellipsoid in (e)). The ellipsoid on the right hand side in (e) encircles the region that lies between the ventricles.

Figure 6. NAA profile used in the basis set. The amplitudes are in arbitrary units. At 0.0 ppm and 8.44 ppm, the two reference compounds are visible. They are effectively removed by the FIR filter during quantitation.

Figures



(a) Set 4 (b) Set 6

Figure 1: Simulated spectra from set 4 and set 6. The amplitudes are in arbitrary units.

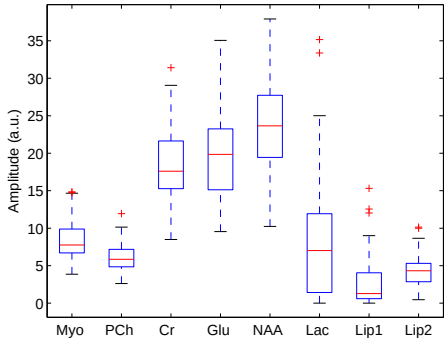


Figure 2: Boxplot showing the quantitated amplitudes of *in vivo* spectra from healthy volunteers.

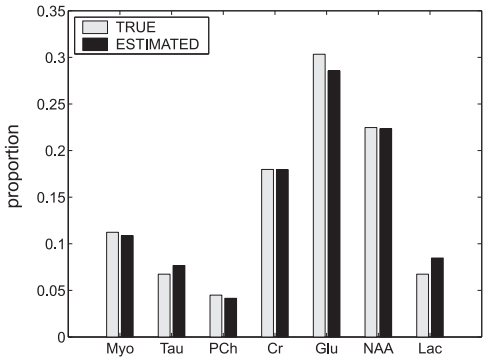


Figure 3: *In vitro* test sample results.

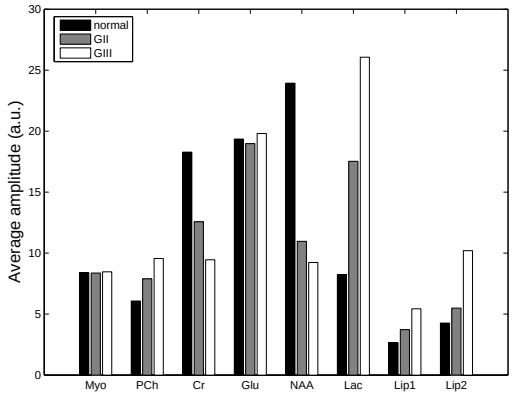
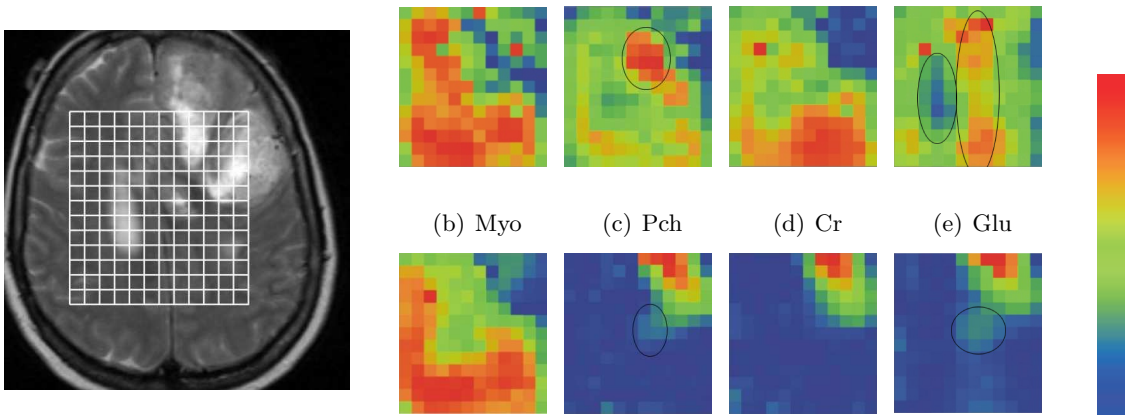


Figure 4: Averaged estimated amplitudes for normal tissues, GII and GBM, in arbitrary units.



(a) T_2 image

(b) Myo (c) Pch (d) Cr (e) Glu

(f) NAA (g) Lac (h) Lip1 (i) Lip2

Figure 5: Metabolic maps of a patient with a glioblastoma. The ellipsoids in (c), (g) and (i) indicate regions where we note a substantial increase of the corresponding metabolites compared to a region with normal tissue (high concentrations in red and low concentrations in blue with a linear scale). This information may particularly be useful since it barely appears on the T2-weighted image. The concentration of Glu is lower in the ventricles (one of the ventricle regions is indicated by the left ellipsoid in (e)). The ellipsoid on the right hand side in (e) encircles the region that lies between the ventricles.

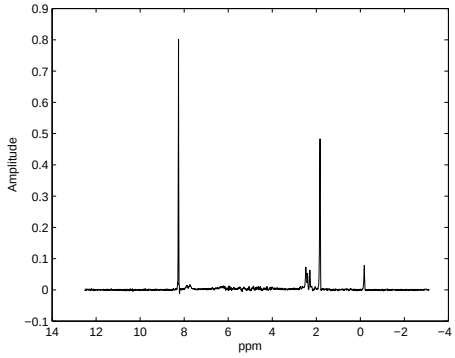


Figure 6: NAA profile used in the basis set. The amplitudes are in arbitrary units. At 0.0 ppm and 8.44 ppm, the two reference compounds are visible. They are effectively removed by the FIR filter during quantitation.