Tina Memo No. 2003-008
Submitted to Journal of Computer Methods in Science and Engineering

Assessment of the Padé Approximant as a Method for Quantifying ^{1}H Magnetic Resonance Spectroscopic Data

D.C.Williamson and N.A.Thacker

Last updated
5 / 12 / 2003

Imaging Science and Biomedical Engineering Division,
Medical School, University of Manchester,
Stopford Building, Oxford Road,
Manchester, M13 9PT.

Assessment of the Padé Approximant as a Method for Quantifying ^{1}H Magnetic Resonance Spectroscopic Data

David C.Williamson and Neil A.Thacker.
Imaging Science and Biomedical Engineering, Stopford Building,
University of Manchester, Manchester, M13 9PT, U.K.
email: david.c.williamson@man.ac.uk

Submitted to Journal of Computer Methods in Science and Engineering

Abstract

The Padé approximant method is employed to analyse simulated and experimental MRS data. The results of the analysis are compared with results obtained using several existing time domain analysis algorithms; The black box HLSVD method, AMARES and a SIMPLEX data fitting using a Voigt lineshape model for the signal components. The results obtained using the Padé approximant method are at least as accurate as the other methods incorporating the Lorentzian lineshape approximation, however in the presence of Gaussian line broadening the Padé approximant also shares the same disadvantages. In the process of assessing the Padé approximant method, it became apparent that the Voigt lineshape model was generally superior to all the other methods considered. The analysis of experimental data obtained from a Perspex phantom demonstrated that all the methods considered were quite adequate, giving very similar results. This was perhaps a consequence of the care taken to acquire the data, which resulted in data that closely approximated the Lorentzian lineshape.

Introduction

The application of magnetic resonance spectroscopy (MRS) to clinical studies is somewhat hampered by the need for accurate quantitation of metabolite signals. In some case visual inspection of the MR spectra can provide insight into a particular condition, however in general more robust methods of peak quantification are required. Peak integration methods are commonly encountered, however they can be inaccurate, particularly in the presence of heavy noise or in cases where neighbouring peaks overlap. More sophisticated approaches include parameter estimation using the frequency domain [1] or the time domain [2] data.

Recently Belkic et al. [3] presented the Padé approximant (PA) as a method of generating magnetic resonance spectra with superior resolution to the standard discrete Fourier transform (DFT). The authors point out that the improved resolution is achieved because the PA is a parameter estimator, rather than an estimator of the spectral shape. In addition the PA was formulated within a signal decimation method used to reduce dimensionality of the data matrix. At the heart of this and many other parameter estimator methods encountered in the quantitation of MRS [4, 5, 6]. The signal decimation was employed to alleviate problems of ill conditioning encountered in the matrix mathematics.

In their paper, Belkic et al. [3] compare the Fourier and Padé spectra obtained from time signals that were generated both theoretically and experimentally. The experimental signals were obtained from NMR experiments and ion cyclotron resonance (ICR) experiments. In all the examples presented the spectra generated using the PA showed improved resolution in comparison to the analogous Fourier spectra. The authors present both noiseless and noisy signals and demonstrate the robust nature of the PA in this respect.

Though visual inspection of spectra can be qualitatively informative, the accurate estimation of peak amplitudes, which are directly proportional to the concentration of a particular metabolite, is required for quantitative assessment. Moreover the examples used by Belkic et al. [3] seem to conform to Lorentzian lineshapes, a situation which is rarely practically achievable in in vivo MRS. Exactly this point has been discussed by Marshall et al. [8] who compared the least squares fitting of Voigt lineshapes in the frequency domain to analogous fits of Gaussian and Lorentzian lineshapes. In addition the frequency domain fits were compared with the results obtained using the Hankel-Lanzcos Singular Value Decomposition (HLSVD) algorithm [5].

In this work we implement the PA in order to directly assess its ability to accurately quantify metabolite signals that are contaminated with white Gaussian noise and have been broadened due to imperfect shimming or susceptibility

variation. Two data sets are examined. The first is a simulated data set containing model signals for lactate, n-acetyl aspartate (NAA), creatine/phosphocreatine (Crn) and choline containing compounds (Cho). The second data set comprises seven spectra acquired from seven tubes in a phantom each of which contains a different concentration of acetate (Ace), creatine and choline. The results obtained using the PA are compared with results from HLSVD, AMARES [9] and a time domain fitting to Voigt lineshapes. Although a number of other black-box algorithms were available, e.g. Hankel Singular Value Decomposition (HSVD), Hankel Total Leaste Squares (HTLS), Hankel-Lanzcos total least squares (HLTLS), Linear Prediction Singular Value Decomposition (LPSVD), in the Java Magnetic Resonance User Interface (jMRUI) package, [10, 9], HSVD, HTLS and HLTLS are based on identical theoretical approached, the so called state space methods [4], [11] and differ only in their numerical implementation. HLSVD is chosen as the most efficient of these methods. No comparison is made to LPSVD since Belkic et al. [3] have demonstrated the mathematical relationship between the LP and PA methods. AMARES is currently the most robust and efficient non-linear least squares fit time domain quantitation algorithm. The method provides the possibility of using either Lorentzian or Gaussian lineshapes and the inclusion of extensive prior knowledge. At present it represents the state of the art in time domain analysis.

Methods and Materials

Theory

In this section we summaries the formulation of the PA with respect to free induction decays (FIDs) measured in MRS. The first part of the summary follows the general discussion about the PA given in Numerical Recipes [7] and the application of the PA to FIDs follows the Belkic et al. [3] discussion.

Consider a polynomial in the variable z such that

$$f(z) \;=\; \sum_{N=0}^{\infty} C_n z^n \tag{1}$$

and a rational function in the same variable.

$$R(z) \;=\; \frac{\sum_{k=0}^{M} a_k z^k}{1 \;+\; \sum_{k=0}^{N} b_k z^k} \tag{2}$$

$R(z)$ is said to be the Padé approximant of the series $f(z)$ if the following are true.

$$R(0) \;=\; f(0) \tag{3}$$

and

$$\frac{d^k}{dz^k} R(z)|_{z=0} \;=\; \frac{d^k}{dz^k} f(z)|_{z=0} \tag{4}$$

Equations (3) and (4) give rise to $M + N + 1$ equations for the unknown coefficients $a_0....a_M$ and $b_0....b_N$ The set of linear equations can be determined by equating equation (1) and equation (2) multiplying both sides by the denominator of equation (2) and equate all powers of z that have either a's or b's in their coefficients.

Now consider a a digitised time signal $\{c_n\}$, with $0 \leq n < N_D$ which comprises N_D data points, equidistantly sampled with a sampling rate of τ_D and having a total bandwidth equal to $2\pi/\tau_D$. The discretised Fourier integral $F(\omega)$ can be written as

$$F(\omega) \;=\; \sum_{n=0}^{N_D-1} c_n e^{(in\omega\tau_D)} \;=\; \sum_{n=0}^{N_D-1} c_n u^{-n} \tag{5}$$

Here, ω is the angular frequency and u is defined as $exp(-i\omega\tau_D)$. This expression reduces to the familiar discrete Fourier transform (DFT) when ω takes values at the Fourier grid, $\omega \;=\; 2\pi m/T$, with $0 \leq m < N_D$ and T is the total acquisition time. $F(\omega)$ is clearly a polynomial in u but it is well known to converge slowly with increasing length N_D.

To proceed it is assumed, for the time being, that the sum in equation (5) can be extended to infinity. The coefficients $\{c_n\}$ can be identified as coefficients of a Maclaurin series in the variable u^{-1} and applying the sum rule for geometric series it is possible to write

$$F_\infty(\omega) \;=\; \sum_{n=0}^{\infty} c_n u^{-n} \;=\; \sum_{k=1}^{K} d_k \sum_{n=0}^{\infty} (u_k/u)^n \;=\; \sum_{k=1}^{K} \frac{u d_k}{(u-u_k)} \;\equiv\; \frac{P_K(u)}{Q_K(u)} \tag{6}$$

Here the set of K peak parameters $\{d_k, \omega_k\}$ has been introduced, d_k is the amplitude of the kth peak, which has a frequency ω_k. $P_{K(u)}$ and $Q_{K(u)}$ are polynomials of order K

$$P_K(u) \;=\; \sum_{k=1}^{K} p_k u^k \qquad and \qquad Q_K(u) \;=\; \sum_{k=1}^{K} q_k u^k \tag{7}$$

and it is clear from equation (6) that the peak frequencies are the poles of $F_\infty(\omega)$, or more conveniently the zero's of the polynomial $Q_K(u)$.

The major draw back of equation (6) lies in the fact that the infinite sum on the right hand side of the equation (6) cannot be calculated. However, the PA is often used to estimate the sum of a power series when only a few coefficients of the series are known. In this case N_D coefficients are known. In addition the DFT is only accurate at the Fourier grid points, while in contrast the PA can be evaluated at arbitrary frequencies contained within the bandwidth of the data.

The determination of the coefficients of the polynomials, $P_K(u)$ and $Q_K(u)$ can be carried out as prescribed previously. Equation (5) is equated to the rational function in equation (6), both sides are multiplied by $Q_K(u)$ and coefficients of the like powers of u are compared. This leads to the following set of linear equations for the $\{q_k\}$ coefficients

$$c_n \;=\; \sum_{k=1}^{K} q_k c_{n+k} \quad , \quad (0 \le n < N_D - 1) \tag{8}$$

This set of equations is subsequently solved using SVD. The $\{p_k\}$ coefficients are determined explicitly from

$$p_k \;=\; \sum_{n=0}^{K-k} q_{k+n} c_n \quad , \quad (1 \le k < K) \tag{9}$$

At this point it is interesting to consider equation (8). When the elements are examined it is clear that the problem of determining the coefficients in equation (8) is identical to SVD of the Hankel matrix encountered in the state space methods, [2, 4, 5, 12, 13]. In these methods, optimal results are obtained when the submatrices of the Hankel matrix are square and all the data points are used at least once. Such considerations suggest that optimal results for the $\{q_k\}$ coefficients will be obtained when the order of $P(u)$ is equal to the order $Q(u)$, the so-called diagonal rational approximation and when the order $k \;=\; N_D/2$. Having determined the values of the $\{q_k\}$ coefficients it is clear the frequencies of the spectral peaks can be determined as the roots of the polynomial $Q_K(u)$ with the complex frequencies being given by

$$\omega_k \;=\; \frac{i}{\tau_D} ln(u_k) \tag{10}$$

The roots are determined in this work using Laguerre's method as prescribed in Numerical Recipes [7]. The corresponding amplitudes $\{d_k\}$ can be determined as the residues of the PA. For the case of distinct poles this gives the expression

$$d_k \;=\; \frac{P_K(u_k)}{u_k Q'_K(u_k)} \quad with \quad Q'_K(u) \;=\; dQ_k(u)/du \tag{11}$$

In the formulation of the PA and in the determination of the coefficients $\{p_k\}$ and $\{q_k\}$ no assumption has been made about the signal or the corresponding spectral lineshape. This is in contrast to the state space method where the signal is explicitly assumed to be a sum of damped exponentials. However, this advantage is not exploited and the determination of the peak parameters $\{d_k, \omega_k\}$ as distinct roots of the polynomial $Q_K(u)$ and the residue of the PA assumes a damped exponential form of the signal components. The generalisation of the PA is far from a trivial problem and remains a topic of current investigation. However, the fact that an explicit expression exists for the amplitudes $\{d_k\}$ presents a second advantage of the PA over other black box methods. All black-box methods, including the PA, will theoretically produces a number of solutions equal to the order of the data matrix in equation (8). A large number of these solutions will be due to noise, or non-physical in nature. In the state-space methods (and the linear prediction methods) the determination of the amplitudes happens through backsubstitution of the peak frequencies into a model for the signal and a least squares fitting to the data. The model requires the user to

determine the number of components in the model signal. The number of signal components that can be included is limited by the numerical practicalities of fitting. In the case of the PA the amplitude for **all** the solutions can be determined at and unwanted or spurious components can be eliminated based on the noise characteristics of the signal later.

Simulation

Synthetic FIDs were simulated comprising 1024 data points generated at 1 ms intervals. The signal was generated according to

$$S(t) \;=\; \sum_{j=1}^{n} A_j \; e^{-\alpha_j t - (\beta_j t)^2} e^{i(2\pi f_j t \; + \; \phi_j)} \tag{12}$$

where A_j is the amplitude of the j th metabolite, α_j is the Lorentzian damping coefficient and β_j the Gaussian damping coefficient, f_j and ϕ_j are the frequency and the phase of the j th metabolite.

Metabolites were simulated with chemical shifts of 1.26, 1.36, 2.02,3.01,3.22 and 3.91 ppm with the amplitudes 8,8,25,10,10,6.67 units. Representing, two lactate (Lac) peaks, N-acetylaspartate (NAA), two creatine (Crn) peaks and choline (Cho). Each of the metabolite was assigned a Lorentzian damping coefficient of 4 Hz corresponding to a T_2 value of 250 ms and zero phase. The Voigt model was chosen to simulate the effects of line broadening due to imperfect shimming and variations in susceptibility. A range of Gaussian damping coefficients were used ranging from 0 to 12 Hz, that corresponds to a range of frequency linewidths that might typically be encountered in *in vivo* and *in vitro* experiments [13]. In all 4 values of β were simulated, 0,4,8 and 12. In each case 200 FIDs were simulated in the presence of white noise, with the noise level being fixed a 10 % of the NAA amplitude.

Phantom details

The phantom used in these experiments is constructed from Perspex. Six tubes were arranged in a ring within the outer phantom tube, with a seventh at the centre of the phantom. The three principal peaks in human brain spectra, associated with N-acetyl- aspartate (NAA), creatine/phosphocreatine (Crn) and choline-containing compounds (Cho), were mimicked in the phantom using solutions of acetic acid, creatine monohydrate and choline chloride. Stock solutions of the three metabolites were made up in 0.09 M saline solutions containing 0.05 mM MnCl2 and 0.2 mM gadolinium DTPA, and the pH was adjusted to be 7.0. The combination of the relaxation agents is used to give relaxation times similar to those of water in brain tissues. The values obtained were: T_1 = 800 ms and T_2 = 260 ms. Different mixtures of metabolite solutions were made up in each of the six inner tubes; their exact concentrations are given in Table 1. The values in parentheses in the choline column are the relative proton concentrations, since choline contains three indistinguishable methyl groups per molecule. In order to estimate the error in quantitation, the signal noise is estimated from the last 10 % of data points in the FID. Random noise of the same level as the estimated value is added to the experimental FID, to generate 200 signals that differ only in the noise. The mean and standard deviation of these results are used to estimate the metabolite amplitude and quantitation error.

Acquisition Details

Single-voxel ^{1}H-spectroscopy was carried out on a 1.5 T Philips Gyroscan ACS NT scanner. A standard spectroscopy head coil was employed. The phantom was positioned in the isocentre of the B0 static field and a 2x2x2 cm3 VOI was localised at the centre of the phantom using the PRESS sequence with an echo time of 27 ms. After shimming of the VOI, 256 acquisition were collected each preceded by 3 CHESS water suppression pulses. Each acquisition comprised 512 data points collected with an acquisition time of 1 ms. Despite the CHESS water suppression the signal still includes a substantial residual water signal, this signal was removed using the SVD-filter method [14].

Quantitation

Both the simulated and phantom data were quantified using AMARES with Lorentzian and Gaussian lineshapes, HLSVD and the PA. The analysis with AMARES was repeated with the inclusion of the following prior knowledge. The signal frequencies were constrained to within 10 % of the expected value, the linewidths were constrained to

be less than 16 Hz and the relative phase of the individual signal components was fixed as zero. In addition the amplitude ratio of the two components of the lactate double was fixed at a value of 1. In addition, the data were quantified using a time domain least squares fitting to the model described in equation (12). The Fourier transform of this model describes a spectrum in which the peaks have the Voigt lineshape. The model was fitted to the data using the SIMPLEX algorithm [7] with the additional constraint that the Gaussian damping coefficients were the same for all metabolites. This amounts to the assumption that all metabolites in a given voxel will experience the same field inhomogeneities and/or susceptibility variations. The reasons for including the Voigt time-domain fitting were partly for completeness and partly to confirm the results of Marshall et al. [8] who observed improved spectral quantitation when fitting non-Lorentzian spectra with a linear approximation to the Voigt lineshape in the frequency domain.

Results and Discussion

The signal amplitudes obtained from quantifying the simulated spectra using all the algorithms discussed are collected in table 2. The standard deviations are calculated from quantifying 200 spectra for each value of the Gaussian broadening parameter, β. For $\beta = 0$, 4, 8, the results obtained using the PA, HLSVD and AMARES with a Lorentzian lineshape are identical within the estimated statistical error for all the signal components. Apparently at $\beta = 12$, HLSVD develops problems estimating the amplitudes of the two components associated with lactate. The remaining components appear to be identical with those obtained using AMARES and the PA. With the exception of the lactate components for $\beta = 12$, the only difference between these three methods appears to be standard deviation, which is slightly smaller for AMARES. The results presented in table 2 were obtained using AMARES without the inclusion of prior knowledge. The results obtained including prior knowledge constraints are not included since they were found to be very similar to the unconstrained results, with the exception that the standard deviations for the estimated amplitudes were considerably smaller. It should be noted that the constraints did prevent the introduction of an arbitrary frequency dependent phase modulation into the model solution. This is discussed in more detail later in this section.

For $\beta = 0$, the results of fitting a Voigt model are identical to those obtained using the PA, HLSVD and AMARES with Lorentzian lineshape. As the Gaussian broadening increases the results obtained using the Voigt fitting remain very close to the ideal values.

As observed by Marshall et al [8] in a previous study, the use of a Lorentzian lineshape to describe data that has a Voigt lineshape results in an over-estimation in the signal amplitudes. In fact taking the results from the PA as an example we observe an overestimation in the NAA amplitude that ranges from 0% for a pure Lorentzian line to over 16 % when the Gaussian broadening is $\beta = 12$. This presents an immediate problem in studies where the ultimate aim is absolute evaluation of metabolite concentrations. Fortunately, in most clinical investigations amplitude ratios are considered to be sufficiently sensitive measures of metabolite level changes. In Figure 1 the NAA amplitude obtained from quantitation, divided by the ideal noiseless value are plotted as a function of the Gaussian broadening. The similarity in the results from algorithms based on the Lorentzian lineshape assumption and the overestimation of the amplitude are clearly demonstrated. It is also clear that a Gaussian lineshape assumption results in an underestimation of the signal amplitude. It appears, from this plot that the bias introduced due to the incorrect lineshape varies in a linear fashion with respect to the line broadening. The solid line shows the best fit through all the results based on the Lorentzian approximation. This is an encouraging result, from the point of view of absolute quantitation, since systematic errors can be calibrated and corrected. Equally, the presence of only a systematic bias in signal amplitudes supports the idea that amplitude ratios may be used as reliable measures of relative metabolite levels. Figure 2 shows the reduced amplitudes obtained using the PA, for NAA, both Crn peaks and Cho. These components are chosen since the NAA peak and the second Crn peak are well separated from any other peaks in the simulated spectrum. On the other hand creatine and choline are sufficiently close so that there is significant overlap between the two peaks. The estimated amplitudes for the NAA and second Crn components show very similar variations with respect to the line broadening, further evidence that despite the bias, the amplitude ratios should be a stable measure. On the other hand the results for the first Crn signal and the Cho signal show very different behaviour. It is clear from this figure that the ratio of these peaks will not be stable with respect to Gaussian broadening. The implication of this result is that unless the signal frequencies are very well separated, it will be almost impossible to systematically correct the bias in the quantitation results. Nor can it be safely assumed that the amplitude ratios accurately reflect changes in metabolite levels. The loss of the systematic behaviour of the bias appears to be linked to the fitting process. Systematic errors are introduced when a single Lorentzian peak is fitted to a single Voigt peak, however when two Voigt peaks with the same relative phase are sufficiently close together, the best fit of a two Lorentzian model appears to occur when the Lorentzians are slightly out of phase. Such a result is physically unacceptable, since there is no ambiguity

in the physical interpretation of the signal phase. This raises a problem with both the PA and HLSVD where there is no mechanism for constraining the relative phase of two signal components in their current implementations. On the other hand, AMARES has been designed around the idea of constraining the fitting process using prior knowledge. Here, the relative phases of the different signal components can be constrained. Such a step is entirely valid, provided the signal components do not exhibit a genuine phase modulation, as is the case with glutamate, glutamine and the minor NAA peaks [15]. In such cases decomposing genuine phase modulation and the effects of fitting model lineshapes will be virtually impossible. Rather than trying to optimise a rather unsatisfactory model it is surely better to improve the model, which in this case involves fitting the data using Voigt lineshapes.

Marshall et al [13] came to the same conclusion with respect to the use of Lorentzian lineshapes. However, they also concluded that amplitude ratios obtained using the Gaussian lineshape assumption were not significantly different to the results obtained using an approximation to a Voigt lineshape. In contrast to this we find that the use of the Gaussian lineshape has similar problems to the use of a Lorentzian lineshape and can introduce additional problems with respect to signal component with similar chemical shifts. Figure 3 shows a plot of the reduced Crn amplitudes versus the reduced Cho amplitude for each of the 200 spectra quantified with $\beta = 12$. The results present are those obtained using AMARES with a Lorentzian lineshape (open squares), AMARES with a Gaussian lineshape (open triangles) and the Voigt lineshape (open circles). These methods were chosen since we wish to compare the lineshape assumption, rather than the quantitation method. The results for the PA and HLSVD lie very close to the AMARES results with Lorentzian lineshapes, although the spread of the data is slightly larger. The Lorentzian assumption shows bias for both the creatine and choline signals; however there does not appear to be a significant correlation between the estimated amplitudes. Using a linear regression, a correlation of, $r^2 = 0.18$ is obtained. Using the Voigt lineshape, the amplitude bias has been removed and the linear regression, shows no correlations, $r^2 = 0.06$. On the other hand, the results obtained using the Gaussian approximation, show little bias in estimated mean amplitude. However, the linear regression shows a negative correlation, slope = -0.93 and a correlation of, $r^2 = 0.85$. This negative correlation invalidates the use of either of these peaks as an internal reference, since an overestimation of one peak will be accompanied by and underestimation in the second peak. Taking the ratio of these two peaks will exaggerate the correlation, which in turn may mask actual changes in metabolite levels. Clearly this correlation could be minimised by the use of prior knowledge. This would be acceptable in the case of the lactate doublet where the amplitude ratios are known a priori. However, with regards to the creatine and choline signals, any constraint on the amplitude ratio would clearly bias the estimate obtained.

Taking the previous discussion into account, the amplitude of the acetate signal was chosen as the reference for the analysis of the phantom data. Figure 4 shows the ratio of the Crn and the acetate amplitudes plotted as a function of the ratio of the known Crn and acetate concentrations. The amplitudes are obtained using the PA (a), HLSVD (b), AMARES with a Lorentzian lineshape (c), AMARES with a Gaussian lineshape (d) and the Voigt fitting (e). A linear regression constrained to pass through zero is performed for each of the quantitation results. It is clear from Figure 4 that all the quantitation methods provide satisfactory results, with regression slopes close to 1 and correlations greater than 0.9. Though there is little to choose between the methods, it appears that the PA and HLSVD provide a slightly higher correlation than the other three methods. However, it is not be possible to definitively conclude that one method was measurably superior to any of the others. Results of similar accuracy were obtained for the choline signal in comparison with the acetate signal. Using the acetate signal in the tube 1 as a reference, we also see excellent agreement between the acetate signal and the acetate concentration

Conclusions

The PA recently described by Belkic et al. [3] has been compared to several time domain quantification algorithms commonly used to quantify ^{1}H MRS data. The results obtained from quantifying simulation and phantom data demonstrates that the PA is at least as accurate as the other algorithms. The PA falls into the category of black box methods and as such shares the advantages and disadvantages of such methods [2]. At the level of determining resonance frequencies the PA is mathematically identical to the HLSVD method. However, having obtained the frequencies, HLSVD requires a fitting step to determine the amplitudes, whereas the amplitudes are obtained from an analytical expression in the PA method. This permits the evaluation of all the amplitudes for all the signal components identified with no loss of computer time, rather than an arbitrarily chosen subset. Exploiting properties of the signal noise may then be used to eliminate non-physical components. This helps to minimise possible problems with failures of the fitting process. It may also be possible to generalise the PA to non-Lorentzian lineshapes, at which point the PA would gain an enormous advantage.

As a black box method the PA is not designed to exploit prior knowledge, which could be considered a disadvantage in comparison with AMARES. However, the data analysed in this work showed no difference in the accuracy of the amplitude estimates of the PA, HLSVD or AMARES, even when AMARES was implemented using prior

knowledge. What was quite apparent from the simulation data is that the assumption of a particular lineshape can make absolute quantification rather difficult, since the bias introduced by the lineshape approximation needs to be calibrated. The use of a Voigt lineshape removed the bias in the simulation data and would be expected to provide more accurate quantification in data where there is substantial line broadening due to field inhomogeneities or susceptibility differences. However, if data is collected carefully to minimise such effects the experimental lineshape may closely approximate a Lorentzian, in which case all the methods described in this work should perform well. This appears to be the case with the phantom data.

In agreement with Marshall and co-workers, [8, 13] and Slotboom et al [16] the use of the Voigt lineshape is recommended since it requires very little additional effort to include in the time domain model but allows for more accurate quantitation in situations where line broadening is unavoidable. In order to incorporate the advantage of the PA with the advantages of prior knowledge and a Voigt lineshape, a hybrid method is suggested in which the PA is used to identify frequencies and estimate the signal amplitudes. These results can be used to initialise a fitting routine, in which the frequencies are fixed, but linewidth, amplitude and phase may be fitted with or without prior knowledge constraints.

Acknowledgements

The authors would like to thank Prof. Dževad Belkić for helpful advice and discussion and also Marietta Scott for preparation of final results. One of us (DCW) would like to thank the Medical Research Council for the award of a special training fellowship in Bioinformatics.

Bibliography

1. Miersisov . And Ala-Korpela M. "MR spectroscopy quantitation: a review of frequency domain methods." NMR in Biomed, 14, 247-259, 2001..
2. Vanhamme L., Sudin T., Van Hecke P. and Van Huffel S., "MR spectroscopy quantitation: a review of time domain methods." NMR in Biomed, 14, 233- 246, 2001.
3. Belkic D. Dando P. A., Main J., and Taylor H. S. "Three novel high- resolution non-linear methods for fast signal processing." J. Chem. Phys. 113, 6542-6556, 2000. ; MAGMA, 15 (Suppl) 36-37, 2002
4. Barkhuysen H., de Beer R., van Ormondt D., Improved algorithm for non- iterative time-domain model fitting to exponentially damped magnetic resonance signals., J. Magn. Reson. 73, 553-557, 1987.
5. Pijnappel W. W. F., van den Boogaart A., de Beer R. and van Ormondt D.SVD-based quantification of magnetic resonance signals. J. Magn. Reson. 97, 122-134, 1992.
6. Koehl P., "Linear prediction spectral analysis of NMR data.", Prog. NMR Spectrosc., 34, 257-299, 1999.
7. Press W. H., Teukolsky S. A., Vetterling W. T., and Flannery B. P., "Numerical recipes in c" 2nd Ed. Cambridge University Press: Cambridge, 1992.
8. Marshall I., Higinbotham J., Bruce S. and Friese A. "Use of Voigt lineshape for quantification of in vivo ^{1}H spectra.", Magn. Reson. Med., 37, 651-657, 1997.
9. Mierisova S., van den Boogaart A., Tkac I., Van Hecke P., VanHamme L., and Liptaj T., New approach for quantitation of short echo time in vivo ^{1}H MR spectra of the brain using AMARES. NMR in Biomed, 11, 32-39, 1998.
10. Naressi A., Couturier C., Castang I.., de Beer R. and Graveron-Demilly D., Java-based graphical user interface for MRUI, a software package for quantitation of in vivo/medical magnetic resonance spectroscopy signals. Comp. Biol. Med.. 31, 269-286, 2001.
11. Naressi A., Couturier C., Devos J. M., Janssen M., Mangeat C., de Beer R. and Graveron-Demilly D., Java-based graphical user interface for the MRUI quantitation package.MAGMA. 12, 141-152, 2001.
12. Kung S. Y., Arun, K. S. and Bhaskar Rao D. V. State-space and singular value decomposition-based approximation methods for the harmonic retrieval problem., J. Opt. Soc Am. 73, 1799-1811, 1983
13. Marshall I., Wardlaw J., Cannon J., Slattery J., Stellar R. J., Reproducibility of metabolite peak areas in ^{1}H MRS of brain, Magn. Reson. Imaging, 14, 281- 291, 1996.
14. van den Boogaart A., van Ormondt D., Pijnappel W. W. F., de Beer R. and Ala-Korpela M., "Removal of the water resonance from ^{1}H magnetic resonance spectra." Mathematics in Signal Processing III, ed. J. G. McWhirter, Clarendon Press, Oxford, 175 195, 1994.
15. De Graaf R. A., "In vivo NMR spectroscopy; principles and techniques." Wiley: Chichester, 1999.
16. Slotboom, J., Boesch C., and Kries R., "Versatile frequency domain fitting using time domain models and prior knowledge.", Magn. Resonan. Med., 39, 899-911, 1998.

Tube	Acetate mM	Creatine mM	Choline mM
1	36	0	12(36)
2	36	36	0(0)
3	18	18	8(24)
4	30	18	6(18)
5	27	18	7(21)
6	27	21	6(18)
7	27	15	8(24)

Table 1: Metabolite Concentrations in each of the six phantom tubes. The values in parentheses in the choline column represent the proton concentration.

Metabolite	Padé	HLSVD	AMARES (Lorentzian)	AMARES (Gaussian)	Voigt
	$\alpha = 4$ Hz	$\beta = 0$ Hz	Noise = 10 %		
Lactate 1	8.06 ± 0.46	8.09 ± 0.49	8.05 ± 0.40	7.12 ± 0.29	8.02 ± 0.24
Lactate 2	8.00 ± 0.44	8.05 ± 0.45	8.01 ± 0.38	7.04 ± 0.26	8.01 ± 0.21
NAA	24.98 ± 0.48	24.99 ± 0.47	25.00 ± 0.42	21.46 ± 0.29	25.04 ±0.22
Creatine 1	10.07 ± 0.49	10.13 ± 0.50	10.09 ± 0.39	8.71 ± 0.32	10.02 ± 0.27
Choline	10.01 ± 0.47	10.03 ± 0.46	10.02 ± 0.45	8.71 ± 0.29	10.02 ± 0.27
Creatine 2	6.69 ± 0.39	6.76 ± 0.36	6.73 ± 0.34	5.83 ± 0.23	6.70 ± 0.22
	$\alpha = 4$ Hz	$\beta = 4$ Hz	Noise = 10 %		
Lactate 1	8.01 ± 0.46	8.24 ± 0.47	8.30 ± 0.41	7.47 ± 0.31	8.02 ± 0.26
Lactate 2	8.47 ± 0.60	8.55 ± 0.59	8.47 ± 0.48	7.38 ± 0.35	8.02 ± 0.20
NAA	26.42 ± 0.44	26.78 ± 0.44	26.96 ± 0.41	22.00 ± 0.32	25.04 ±0.31
Creatine 1	10.56 ± 0.55	10.75 ± 0.50	10.80 ± 0.39	9.02 ± 0.34	10.02 ± 0.31
Choline	10.29 ± 0.54	10.48 ± 0.52	10.52 ± 0.44	9.00 ± 0.31	10.02 ± 0.30
Creatine 2	6.93 ± 0.37	7.12 ± 0.37	7.12 ± 0.36	5.98 ± 0.27	6.70 ± 0.26
	$\alpha = 4$ Hz	$\beta = 8$ Hz	Noise = 10 %		
Lactate 1	7.35 ± 0.60	7.46 ± 0.58	8.36 ± 0.57	7.74 ± 0.46	7.98 ± 0.31
Lactate 2	8.93 ± 0.63	9.07 ± 0.64	8.85 ± 0.56	7.48 ± 0.40	7.96 ± 0.30
NAA	28.78 ± 0.59	29.03 ± 0.60	25.00 ± 0.55	23.01 ± 0.42	24.98 ± 0.31
Creatine 1	10.90 ± 0.66	11.17 ± 0.66	10.09 ± 0.60	9.32 ± 0.42	9.91 ± 0.37
Choline	10.60 ± 0.57	10.86 ± 0.58	10.02 ± 0.55	9.60 ± 0.41	10.01 ± 0.38
Creatine 2	7.09 ± 0.57	7.39 ± 0.62	6.73 ± 0.58	6.19 ± 0.46	6.59 ± 0.37
	$\alpha = 4$ Hz	$\beta = 12$ Hz	Noise = 10 %		
Lactate 1	8.93 ± 1.44	14.36 ± 4.48	8.24 ± 1.15	8.06 ± 1.54	8.04 ± 0.47
Lactate 2	9.51 ± 1.29	12.17 ± 3.36	9.35 ± 1.15	7.04 ± 0.26	8.06 ± 0.43
NAA	29.13 ± 2.61	31.43 ± 2.48	30.20 ± 0.70	23.70 ± 0.29	25.15 ±0.69
Creatine 1	11.25 ± 0.76	11.44 ± 0.84	12.11 ± 0.77	9.54 ± 0.32	10.02 ± 0.27
Choline	10.41 ± 0.74	10.99 ± 0.87	10.69 ± 0.68	9.83 ± 0.29	10.02 ± 0.27
Creatine 2	7.21 ± 0.39	7.88 ± 0.74	7.35 ± 0.67	7.35 ± 0.23	6.70 ± 0.22

Table 2: Metabolite amplitudes for Lactate, NAA, creatine and choline obtained using the PA, HLSVD, AMARES using both a Lorentzian and Gaussian lineshape model and a SIMPLEX fitting to a Voigt lineshape model. The Lorentzian linewidth and the noise were kept constant for the simulation data and the Gaussian line broadening was increased. For each set of parameters 200 spectra were generated and quantified. These results were used to calculate the means and standard deviations presented in this table.

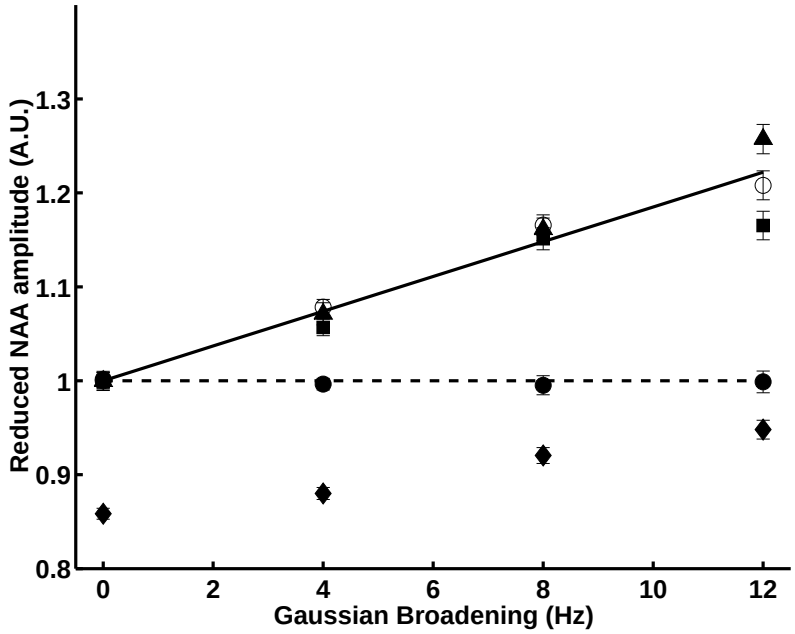


Figure 1: Figure 1. NAA amplitude reduced by the noiseless value plotted as a function of Gaussian line broadening β. The amplitudes were obtained using the PA (filled squares), AMARES using the Lorentzian lineshape approximation (filled triangles), AMARES using the Gaussian lineshape approximation (filled diamonds), HLSVD (open circles) and fitting a Voigt lineshape model (filled circles). The dashed line shows the ideal result of 1 for all values of the broadening parameter and the solid line represents the least squares fit through all the results obtained using the Lorentzian lineshape model, e.g., PA, HLSVD and AMARES. The error bars were obtained by quantifying 200 simulated spectra.

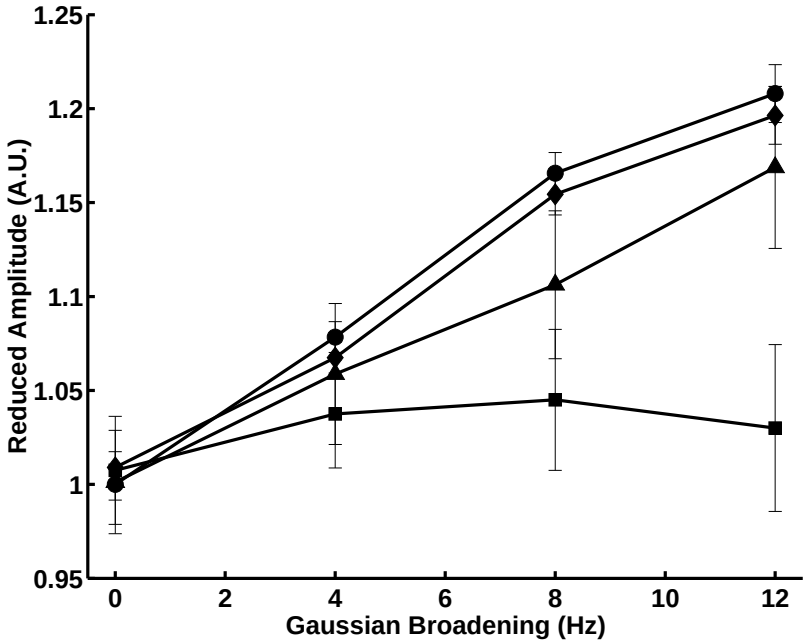


Figure 2: Figure 2. The reduced amplitudes of NAA (solid circle), Crn (solid square and solid diamond) and Cho (solid triangle), as a function of Gaussian line broadening β. The amplitudes were obtained using the PA and the error bars obtained by quantifying 200 simulated spectra. The lines have been drawn as a guide to the eye to highlight the trend in each of the data sets but do not interpolate the data.

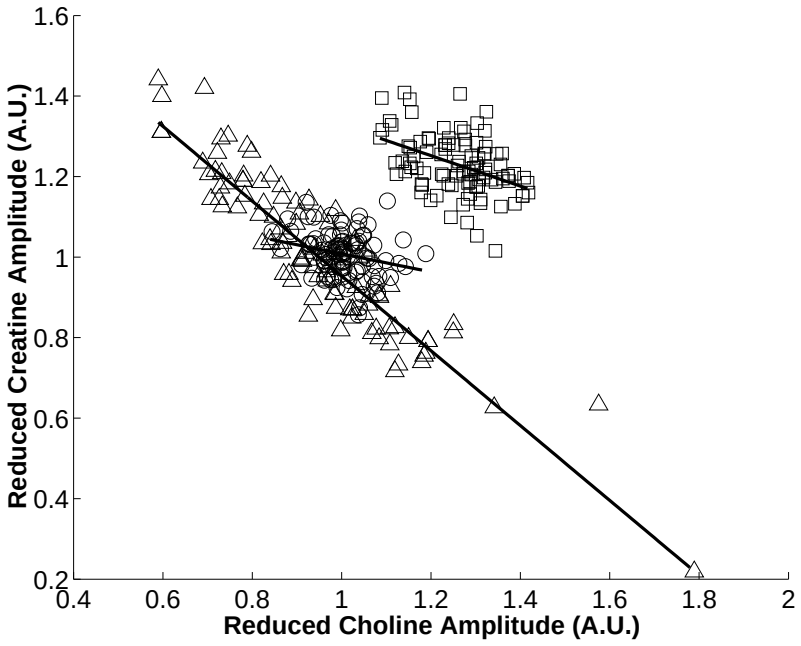
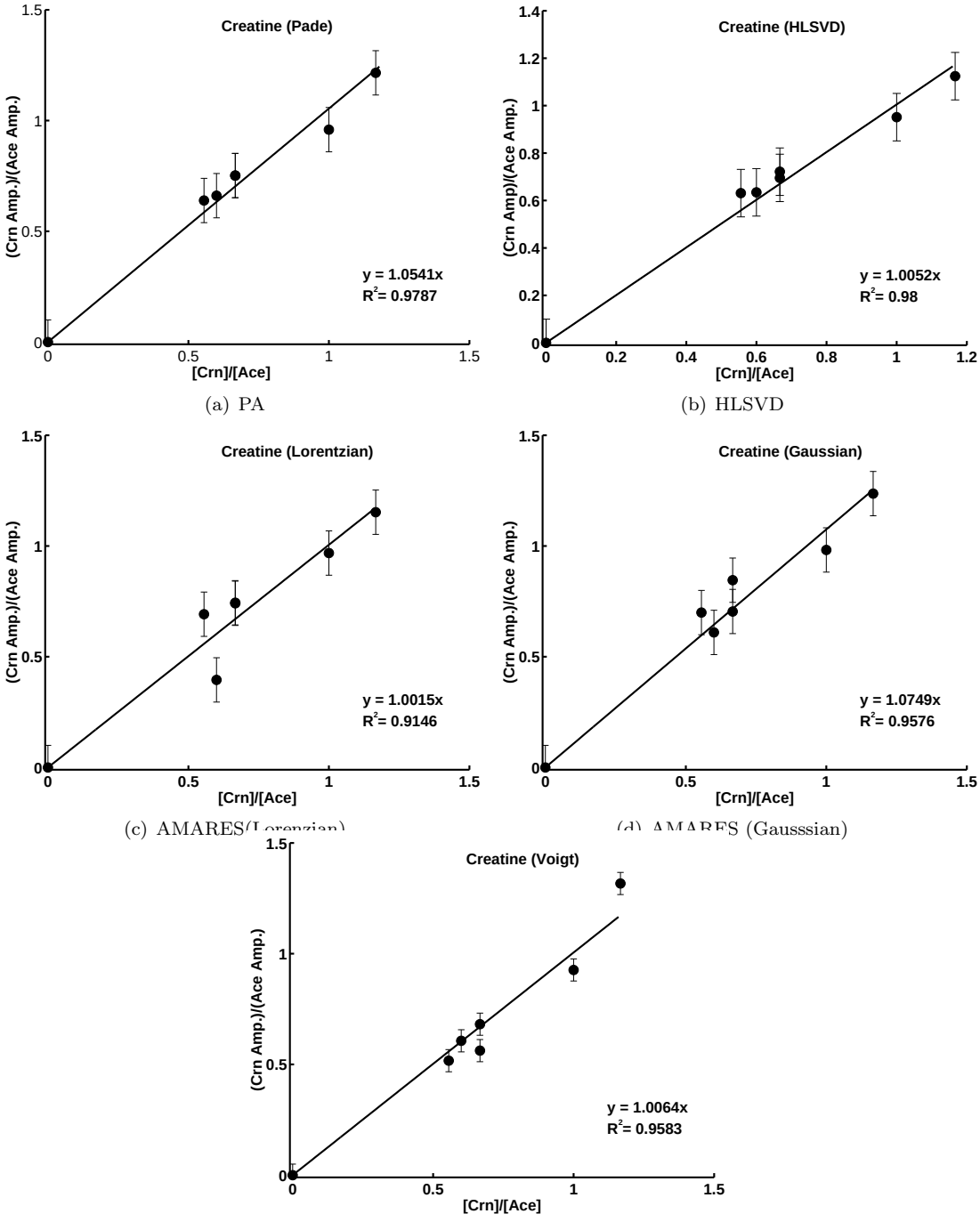


Figure 3: Figure 3. Reduced Crn amplitude plotted as a function of reduced Choline amplitude. Each point represents the results obtained from quantifying a single spectrum from the simulated data set. The black lines represent the results from linear regression analysis of each of the data sets. The amplitudes were obtained using AMARES with the Lorentzian lineshape model (open squares; $y = 1.71 - 0.37x$, $r^2 = 0.18$), AMARES with a Gaussian lineshape model (open triangles; $y = 1.88 - 0.93x$, $r^2 = 0.85$) and a SIMPLEX fitting of the data to a Voigt lineshape model (open circles; $y = 1.23 - 0.23x$, $r^2 = 0.06$).



(a) PA

(b) HLSVD

(c) AMARES(Lorenzian)

(d) AMARES (Gausssian)

(e) Simplex (Voigt)

Figure 4: Figure 4. Amplitude ratio of Crn and Acetate plotted as a function of the concentration ratio of the same metabolite pair. The plots include the results from each tube in a six tube phantom and the results of a linear regression analysis. The amplitudes were obtained using (a) the PA, (b) HLSVD, (c) AMARES with a Lorentzian lineshape model, (d) AMARES with a Gaussian lineshape model and (e) SIMPLEX fitting to a Voigt lineshape model.