JOURNAL OF MAGNETIC RESONANCE **129,** 35–43 (1997)
ARTICLE NO. MN971244

Improved Method for Accurate and Efficient Quantification of MRS Data with Use of Prior Knowledge

Leentje Vanhamme,* Aad van den Boogaart,† and Sabine Van Huffel*

**Department of Electrical Engineering (ESAT), Katholieke Universiteit Leuven, Kard. Mercierlaan 94, 3001 Leuven, Belgium; and †Biomedical NMR Unit, Katholieke Universiteit Leuven, Gasthuisberg, 3000 Leuven, Belgium*

Received January 14, 1997; revised July 15, 1997

We introduce AMARES (*a*dvanced *m*ethod for *a*ccurate, *r*obust, and *e*fficient *s*pectral fitting), an improved method for accurately and efficiently estimating the parameters of noisy magnetic resonance spectroscopy (MRS) signals in the time domain. As a reference time domain method we take VARPRO. VARPRO uses a simple Levenberg–Marquardt algorithm to minimize the variable projection functional. This variable projection functional is derived from a general functional, which minimizes the sum of squared differences between the data and the model function. AMARES minimizes the general functional which improves the robustness of MRS data quantification. The newly developed method uses a version of NL2SOL, a sophisticated nonlinear least-squares algorithm, to minimize the general functional. In addition, AMARES uses a singlet approach for imposition of prior knowledge instead of the multiplet approach of VARPRO because this greatly extends the possibilities of the kind of prior knowledge that can be invoked. Other new features of AMARES are the possibility of fitting echo signals, choosing a Lorentzian as well as a Gaussian lineshape for each peak, and imposing lower and upper bounds on the parameters. Simulations, as well as *in vivo* experiments, confirm the better performance of AMARES compared to VARPRO in terms of accuracy, robustness, and flexibility.

INTRODUCTION

For medical diagnosis or biochemical analysis accurate and efficient quantification of magnetic resonance spectroscopy (MRS) signals is of utmost importance. MRS signals however are often characterized by a low signal-to-noise ratio and overlapping peaks. In these circumstances simple signal processing algorithms like numerical integration are not adequate. In recent years a number of more advanced time domain techniques based on a mathematical model function have been developed.

Noninteractive methods exist that are noniterative and computationally efficient and which can be fully automatic. A serious drawback however is the fact that only very limited prior knowledge can be incorporated in these algorithms and that the model function is restricted to a sum of complex exponentially damped sinusoids. Among this class of methods are the algorithms based on Kumaresan–Tufts' linear prediction (LP) method (*1*) combined with singular value decomposition (SVD) (*2*). Kung *et al.*'s state-space approach (*3*) combined with SVD (called HSVD (*4*)) is a more efficient and more accurate alternative to the LP methods as it circumvents the polynomial rooting and root selection. Rapid and more accurate variants of the state-space algorithms have been recently proposed (*5–8*), but the limitations to the imposition of prior knowledge about model function parameters are inherent to these types of methods. On the other hand, interactive methods exist that are iterative, with more user-involvement, in some cases computationally less efficient, but that do allow inclusion of prior knowledge. The algorithms fit the data to the nonlinear model function in a least-squares sense, leading to maximum likelihood parameter estimates in the case of white Gaussian noise. See (*9*) for an overview of time domain methods.

In this paper we introduce AMARES (*a*dvanced *m*ethod for *a*ccurate, *r*obust, and *e*fficient *s*pectral fitting), an improved interactive time domain method for accurate and efficient parameter estimation of MRS signals with use of prior knowledge. As a reference interactive method we take VARPRO (*10*), a method that has proven to be very reliable in recent years (*11, 12*). Our method improves VARPRO in several ways. First of all, we minimize a different functional than in VARPRO, leading to more robust quantification of MRS signals. In addition, we use a more sophisticated nonlinear least-squares (NLLS) algorithm which has a more robust behavior than the simple Levenberg–Marquardt

1090-7807/97 $25.00

(LM) algorithm used in VARPRO. The robustness is also increased by imposing lower and upper bounds on the parameters. This allows one to impose the physical requirement that all damping factors are positive. It is shown here that this strategy leads to greater robustness than the transformation of variables used in VARPRO. We also greatly extended the possibilities of imposing prior knowledge. This leads to more accurate parameter estimation since the precision of the quantitative analysis is increased by using all available information (*9*) and increases the flexibility for the user. In addition, using all available prior knowledge leads to an optimization problem in a smaller number of variables which can be solved more consistently in less time. As extra features of our new algorithm we made it possible to use a more general model function than in VARPRO and to fit entire spin echo signals instead of having to truncate the first part of the echo as in VARPRO. As a result our method allows for more robust, accurate, flexible, and efficient quantification of MRS signals.

In the first part of the paper we will explain the algorithmic differences between VARPRO and AMARES and we will point out the extended possibilities of imposing prior knowledge in AMARES. The differences in robustness and efficiency between VARPRO and AMARES are illustrated in the second part of the paper. In the third part we illustrate the extended possibilities of imposing prior knowledge using an *in vivo* signal.

METHODS

The most commonly used model function in MRS signal processing to fit N measured data points y_n is a sum of exponentially damped sinusoids (Lorentzian lines after FT),

$$y_n = \hat{y}_n + e_n = \sum_{k=1}^{K} a_k e^{j\phi_k} e^{(-d_k + j2\pi f_k)t_n} + e_n,$$
$$n = 0, 1, \ldots, N-1, \qquad [1]$$

where $j = \sqrt{-1}$, a_k is the amplitude, ϕ_k is the phase, d_k is the damping factor, and f_k is the frequency of the kth sinusoid ($k = 1, \ldots, K$); $t_n = n\Delta t + t_0$ with Δt the sampling interval (nonuniform sampling vectors are also valid), t_0 is the time between the effective time origin and the first data point to be included in the analysis, and e_n is complex white Gaussian noise. The caret on y indicates that this quantity represents the model function rather than the actual measurements. Other types of model line forms can be used; we will come back to this.

To obtain maximum likelihood estimates in the case of white Gaussian noise, one has the choice between minimizing two functionals. The first one, denoted by G throughout the paper, is straightforwardly derived using probability theory and is given by the formula

$$G(\mathbf{a}, \mathbf{d}, \mathbf{f}, \boldsymbol{\phi}, t_0) = \sum_{n=0}^{N-1} \left| y_n - \sum_{k=1}^{K} a_k e^{j\phi_k} e^{(-d_k + j2\pi f_k)t_n} \right|^2$$
$$= \|\mathbf{y} - \Psi\mathbf{l}\|^2, \qquad [2]$$

where $\mathbf{y} = [y_0, \ldots, y_{N-1}]^T$ is the signal vector, $\mathbf{l} = [a_1 e^{j\phi_1}, \ldots, a_K e^{j\phi_K}]^T$, and $\mathbf{a}$, $\mathbf{d}$, $\mathbf{f}$, and $\boldsymbol{\phi}$ are the vectors of amplitudes, dampings, frequencies, and phases, respectively. The superscript T denotes the transpose, $\|\cdot\|$ the Euclidean vector norm, and

$$\Psi = \begin{bmatrix} e^{(-d_1 + j2\pi f_1)t_0} & \cdots & e^{(-d_K + j2\pi f_K)t_0} \\ \vdots & \ddots & \vdots \\ e^{(-d_1 + j2\pi f_1)t_{N-1}} & \cdots & e^{(-d_K + j2\pi f_K)t_{N-1}} \end{bmatrix} \qquad [3]$$

is an $N \times K$ matrix of full rank.

The second functional is derived as follows. Suppose that the nonlinear parameters $\mathbf{f}$ and $\mathbf{d}$ are known; then the matrix Ψ can be computed and an estimate for the linear parameters $\mathbf{l}$ is obtained by solving a linear LS problem: $\hat{\mathbf{l}} = \Psi^{\dagger}\mathbf{y}$, with $\Psi^{\dagger} = (\Psi^T\Psi)^{-1}\Psi^T$, the pseudo-inverse of Ψ. Substituting this estimate into the original functional G of Eq. [2] results in

$$V(\mathbf{d}, \mathbf{f}, t_0) = \|\mathbf{y} - \Psi\Psi^{\dagger}\mathbf{y}\|^2, \qquad [4]$$

which is called the variable projection functional (*13*) and denoted by V from now on. In this way the amplitudes $\mathbf{a}$ and the phases $\boldsymbol{\phi}$ are eliminated. As a result we obtain a minimization problem where the number of variables is reduced but where the functional has become more complicated.

Both functionals consist of a sum of squared residuals and give rise to typical NLLS problems. The available biochemical prior knowledge can be expressed as a set of linear relations between parameters resulting in a minimization problem with linear equality constraints. These constraints are substituted in the original functional in order to obtain an unconstrained NLLS minimization problem. As a consequence, regardless of which of the two functionals in Eq. [2] or Eq. [4] is used, we always have to solve an unconstrained NLLS optimization problem.

The latter can be solved using local or global optimization theory. Global optimization has already been used in MRS (*14, 15*) but the main disadvantage of these methods is the poor computational efficiency. The methods are too time consuming to be useful in practical situations. On the other hand it is possible to obtain good starting values by peak picking for use in local optimization methods, resulting in an acceptable solution in a reasonable time in most circumstances. Therefore, in this paper we focus our attention on methods based on local optimization theory.

VARPRO: Algorithmic Aspects

VARPRO minimizes the variable projection functional V using a modified version of Osborne's Levenberg–Marquardt algorithm (*16*). The basic model function of Eq. [1] is extended so that one has the choice between a pure Gaussian ($g = 1$) or a pure Lorentzian ($g = 0$) lineshape for the *entire* signal:

$$y_n = \hat{y}_n + e_n = \sum_{k=1}^{K} a_k e^{j\phi_k} e^{(-b_k^2(1-g+gt_n)t_n)} e^{j2\pi f_k t_n} + e_n,$$
$$n = 0, 1, \ldots, N-1. \quad [5]$$

It is very important to note here that in VARPRO minimization is done w.r.t. the square roots of the damping factors b_k, i.e., $b_k^2 = d_k$ ($k = 1, \ldots, K$), to make sure that the damping is positive as required by physical considerations. This variable transformation works in most practical situations, but is not without danger as explained by Gill and coworkers (*17*) and as we will show under Simulations.

Different forms of multiplet prior knowledge can be imposed as relations between parameters of the same type; a detailed explanation is given under VARPRO: Multiplet Approach for Prior Knowledge.

With VARPRO it is possible to do frequency-selective quantification in the time domain (*18*). As a measure of precision we compute approximations to the Cramér–Rao lower bounds as explained in (*9*).

AMARES: Algorithmic Aspects

AMARES uses *dn2gb* (available from the Port library of netlib (*19*)), the most recent version of NL2SOL to minimize the general functional. NL2SOL is a secant method recommended in (*20*) and has the advantage of handling large-residual or very nonlinear problems better than a Levenberg–Marquardt algorithm (as used in VARPRO). An additional advantage of using *dn2gb* is the fact that the algorithm allows the user to specify upper and lower bounds on the variables. We can use this feature to impose the natural bounds on the variables. The physical requirement of a positive damping can thus easily be imposed. Likewise amplitudes (like the dampings) cannot become negative (negative amplitudes correspond to a 180° phase shift), whereas the upper and lower bounds on the frequencies are determined by the spectral width. Phases can be constrained to lie between −180° and +180°. In general it is recommended to add those extra bounds on the variables in order to ensure maximum accuracy and robustness (*17*). We show under Simulations that this new approach of ensuring the positiveness of the dampings leads to a more robust method than VARPRO.

AMARES fits the data to the following model function in a least-squares sense:

$$y_n = \hat{y}_n + e_n = \sum_{k=1}^{K} a_k e^{j\phi_k} e^{(-d_k(1-g_k+g_k t_n)t_n)} e^{j2\pi f_k t_n} + e_n,$$
$$n = 0, 1, \ldots, N-1. \quad [6]$$

The basic model function of Eq. [1] is extended so that one now has the choice between a pure Lorentzian ($g_k = 0$) or a pure Gaussian line form ($g_k = 1$) for every peak separately as opposed to VARPRO where the line form used must be the same for all peaks.

AMARES allows the imposition of all kinds of linear relations between individual parameters by using a singlet instead of a multiplet approach, thereby greatly increasing the prior knowledge that can be used in the data processing. This is explained in more detail under AMARES: Singlet Approach for Prior Knowledge and is illustrated under Quantification of an *in Vivo* Signal Using Prior Knowledge.

AMARES also offers the ability to fit echo signals, an echo being modeled as two FIDs back to back. The left and the right part of the echo are considered to have the same amplitudes, frequencies, and phases but different dampings. The dampings of the right and left part can however be linked to each other. In VARPRO the first part of the echo is truncated in order to be able to work with the model function of a FID signal. As a consequence, part of the signal and thus part of the information is destroyed.

Like in VARPRO the user has the possibility to perform frequency-selective quantification in the time domain (*18*). The approximations to the Cramér–Rao lower bounds are used as a measure of precision. Note that the Cramér–Rao bounds do not change by imposing simple bounds on the variables as proven in (*21*).

Computational Considerations

VARPRO and AMARES both make use of evaluations of the residuals and the Jacobian, the Jacobian being the matrix consisting of the first derivatives of the residuals with respect to the unknown parameters. In both methods the residuals and Jacobian are provided analytically. Evaluation of the residuals and the Jacobian of V (VARPRO) is far more complicated than in the case of G (AMARES). In VARPRO a simplification introduced by Kaufman (*22*) is used which results in a more efficient computation of the Jacobian of V.

To perform the actual minimization, the data and the model function are split in a real and imaginary part as

$$V = \|\mathbf{y}^V - \Psi^V \mathbf{l}^V\| = \|\mathbf{y}^V - \Psi^V \Psi^{V\dagger} \mathbf{y}\|, \quad [7]$$

where

$$\Psi^V = \begin{bmatrix} e^{-d_1 t_0}\cos(2\pi f_1 t_0) & -e^{-d_1 t_0}\sin(2\pi f_1 t_0) & \cdots & e^{-d_K t_0}\cos(2\pi f_K t_0) & -e^{-d_K t_0}\sin(2\pi f_K t_0) \\ e^{-d_1 t_0}\sin(2\pi f_1 t_0) & e^{-d_1 t_0}\cos(2\pi f_1 t_0) & \cdots & e^{-d_K t_0}\sin(2\pi f_K t_0) & e^{-d_K t_0}\cos(2\pi f_K t_0) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ e^{-d_1 t_{N-1}}\cos(2\pi f_1 t_{N-1}) & -e^{-d_1 t_{N-1}}\sin(2\pi f_1 t_{N-1}) & \cdots & e^{-d_K t_{N-1}}\cos(2\pi f_K t_{N-1}) & -e^{-d_K t_{N-1}}\sin(2\pi f_K t_{N-1}) \\ e^{-d_1 t_{N-1}}\sin(2\pi f_1 t_{N-1}) & e^{-d_1 t_{N-1}}\cos(2\pi f_1 t_{N-1}) & \cdots & e^{-d_K t_{N-1}}\sin(2\pi f_K t_{N-1}) & e^{-d_K t_{N-1}}\cos(2\pi f_K t_{N-1}) \end{bmatrix} \quad [8]$$

$$\mathbf{y}^V = \begin{bmatrix} \mathrm{Re}(y_0) \\ \mathrm{Im}(y_0) \\ \vdots \\ \mathrm{Re}(y_{N-1}) \\ \mathrm{Im}(y_{N-1}) \end{bmatrix}, \quad \mathbf{l}^V = \begin{bmatrix} a_1\cos(\phi_1) \\ a_1\sin(\phi_1) \\ \vdots \\ a_K\cos(\phi_K) \\ a_K\sin(\phi_K) \end{bmatrix} = \begin{bmatrix} \mathrm{Re}(c_1) \\ \mathrm{Im}(c_1) \\ \vdots \\ \mathrm{Re}(c_K) \\ \mathrm{Im}(c_K) \end{bmatrix} \quad [9]$$

with $c_k = a_k e^{j\phi_k}$ $(k = 1, \cdots K)$.

$$G = \|\mathbf{y}^G - \Psi^G \mathbf{l}^G\|, \quad [10]$$

where

$$\Psi^G = \begin{bmatrix} e^{-d_1 t_0}\cos(2\pi f_1 t_0 + \phi_1) & \cdots & e^{-d_K t_0}\cos(2\pi f_K t_0 + \phi_K) \\ e^{-d_1 t_0}\sin(2\pi f_1 t_0 + \phi_1) & \cdots & e^{-d_K t_0}\sin(2\pi f_K t_0 + \phi_K) \\ \vdots & \ddots & \vdots \\ e^{-d_1 t_{N-1}}\cos(2\pi f_1 t_{N-1} + \phi_1) & \cdots & e^{-d_K t_{N-1}}\cos(2\pi f_K t_{N-1} + \phi_K) \\ e^{-d_1 t_{N-1}}\sin(2\pi f_1 t_{N-1} + \phi_1) & \cdots & e^{-d_K t_{N-1}}\sin(2\pi f_K t_{N-1} + \phi_K) \end{bmatrix} \quad [11]$$

$$\mathbf{y}^G = \begin{bmatrix} \mathrm{Re}(y_0) \\ \mathrm{Im}(y_0) \\ \vdots \\ \mathrm{Re}(y_{N-1}) \\ \mathrm{Im}(y_{N-1}) \end{bmatrix}, \quad \mathbf{l}^G = \begin{bmatrix} a_1 \\ \vdots \\ a_K \end{bmatrix}. \quad [12]$$

$\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ denote the real and imaginary parts of a complex quantity.

In VARPRO and AMARES starting values for the frequencies and dampings need to be provided by the user. In AMARES, the starting values for the amplitudes and phases are computed by solving the LS problem $\hat{\mathbf{l}}^V = \Psi^{V\dagger}\mathbf{y}^V$, with the starting values for the frequencies and dampings inserted in Ψ^V.

VARPRO and AMARES are both written in Fortran77. They can be used as stand-alone programs, but they are also implemented within the MRUI software package (*23*), a graphical user interface to facilitate use of sophisticated spectral analysis routines in biomedical/biochemical laboratories and the clinic. AMARES is available from the authors upon request.

VARPRO: Multiplet Approach for Prior Knowledge

VARPRO puts peaks belonging to the same multiplet into one group. Amplitudes within one group can be left unconstrained or they can be linked to each other when the intensity ratios of the peaks are known. Dampings are treated in the same way, but with the extra possibility of fixing the damping to a value which must be the same for all peaks in that group. Frequencies within a group can be left unconstrained or they can be linked if the frequency splittings between neighboring peaks are known exactly. If the frequency splittings between neighboring peaks are unknown, but equal for all the peaks in the group, a new variable Δ is introduced. One then has f_1, $f_2 = f_1 + \Delta$, $f_3 = f_1 + 2\Delta$, $\ldots$, $f_k = f_1 + (k-1)\Delta$, for all k peaks belonging to the same group; f_k is the arbitrarily chosen reference peak, and neighboring peaks in the spectrum are denoted by consecutive numbers. In addition, the frequency of the first component (chosen arbitrarily) can be fixed together with the frequency splitting of the multiplet. In that case all the frequencies of a group are fixed. Note that it is impossible to fix the frequencies of the peaks of a certain group independently: the frequency differences between neighboring peaks are the same. The overall phase ϕ_k of a peak is considered to consist of an individual phase ϕ'_k and a so-called zero-order phase ϕ_0 which is the same for all peaks: $\phi_k = \phi'_k + \phi_0$. The individual phases can be left unconstrained, in which case the zero-order phase must be fixed. The phases of all peaks within a group can be fixed to a value (the same for all the peaks within the group) relative to the zero-order

phase, which can be estimated or kept fixed. Of all the groups with constraints on amplitudes one group can be taken as a reference and the other groups can be linked to the reference group. This means that the relative amplitude of the first component (chosen arbitrarily) of the reference group w.r.t. the first component of a linked group can be imposed. Groups with constrained dampings are treated in the same way. Groups with variable frequency splittings can also be linked in the same way, the only difference being that the frequency splitting ratio between the linked groups must be one.

We illustrate some of the above-mentioned possibilities using the ^{31}P simulation example. The peaks belonging to the β-ATP triplet are put into one group. The amplitude ratios are $a_1{:}a_2{:}a_3 = 1{:}2{:}1$. If we impose this prior knowledge in VARPRO and we leave the corresponding phases unconstrained, the program will automatically assume these phases to be equal (this is not the case when amplitudes are left unconstrained). This is equivalent to linking the complex amplitudes $a_k e^{j\phi_k}$ or in this example $\mathrm{Re}(c_1){:}\mathrm{Re}(c_2){:}\mathrm{Re}(c_3) = 1{:}2{:}1$ and $\mathrm{Im}(c_1){:}\mathrm{Im}(c_2){:}\mathrm{Im}(c_3) = 1{:}2{:}1$. Peaks 4 and 5 belong to the α-ATP doublet and are put in a second group. We know that $a_4{:}a_5 = 1{:}1$ and that $a_4{:}a_1 = 2{:}1$. VARPRO offers the possibility of imposing this prior knowledge between the two groups as long as the constraints on the phases are equal for both groups.

In the case of the β-ATP triplet we can impose a fixed frequency splitting of 16 Hz between the individual peaks within the triplet, leading to $f_2 = f_1 + 16$ and $f_3 = f_1 + 32$. Suppose however that we do not know the exact frequency splittings between the peaks but that we do know that these splittings are the same. We can express this by introducing a new variable Δ. We then get $f_2 = f_1 + \Delta$ and $f_3 = f_1 + 2\Delta$. We can also express that the frequency splitting in the α-ATP doublet is the same as that in the β-ATP triplet. Using VARPRO we express this as $f_5 = f_4 + \Delta$.

This implementation of prior knowledge in VARPRO already offers many possibilities of imposing various linear relations between parameters and has proven crucial in earlier studies (*12, 24, 25*). However, in recent years, more prior knowledge has become available and the present implementation of VARPRO can no longer satisfy all the needs (*11*). One example is formed by the six glycogen resonances in ^{13}C MRS, for which the relative frequency shifts are known but different. This can therefore not be implemented as a regular multiplet in the VARPRO algorithm. For the quantification of a ^{13}C MRS signal using AMARES, we refer to (*26*).

AMARES: Singlet Approach for Prior Knowledge

In AMARES we have chosen for a singlet approach, making an identical treatment of all parameters possible. This approach allows everything that was previously possible using the multiplet approach of VARPRO and much more.

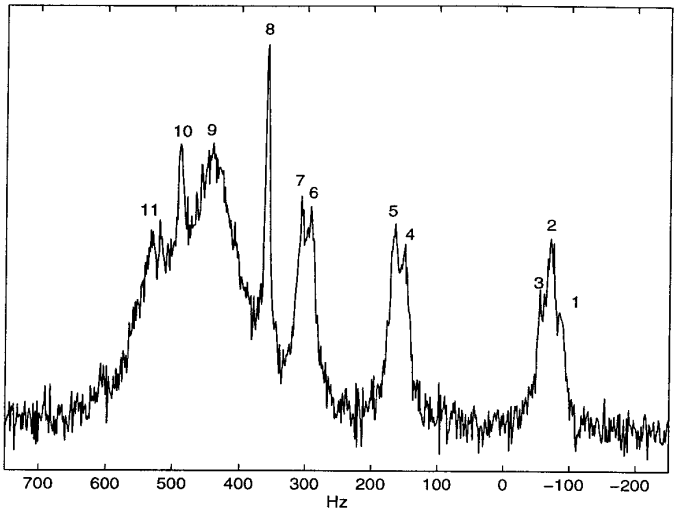

FIG. 1. Frequency domain representation of simulated ^{31}P MRS signal; the standard deviation of the noise is 15. See Table 1 for the parameter values of the signal.

Each of the parameters can be left unconstrained or kept fixed. One can impose a fixed shift or ratio w.r.t. any unconstrained parameter of the same type. Finally, it is possible to impose a *variable* shift or ratio w.r.t. any unconstrained or fixed parameter of the same type. These variable shifts or ratios can then be linked between different groups of peaks. Imposing a fixed shift or ratio and a variable shift or ratio deserves some more explanation; the other forms of prior knowledge are trivial to implement.

As an arbitrary example, take a signal consisting of three peaks. Imposing a fixed ratio is done in the following way, e.g., for the amplitudes $a_2 = a_1 x$, $a_3 = a_1 y$, $x, y \in \mathbb{R}$. x and y are values provided by the user. Imposing a variable ratio leads to the introduction of a new variable Δ, $a_2 = a_1(x\Delta)$, $a_3 = a_1(y\Delta)$, $x, y \in \mathbb{R}$. x and y are values provided by the user. Imposing a variable shift or ratio is useful only if the new variable is shared by at least two peaks, otherwise the total number of unknowns is not changed.

SIMULATIONS

In this section we demonstrate the differences between VARPRO and AMARES. To this end we perform a Monte Carlo study. The simulation signal we used was derived from an *in vivo* ^{31}P spectrum measured in the human brain and consisted of 256 complex data points and 11 exponentials (see Fig. 1 and Table 1). The ^{31}P peaks from brain tissue, from left to right phosphomonoesters, inorganic phosphate, phosphodiesters, phosphocreatine, γ-ATP, α-ATP, and β-ATP, can all be observed. From the noiseless signal 300 noisy realizations were generated with noise standard deviation σ_ν (on both the real and the imaginary parts). We used a low, intermediate, and high noise level ($\sigma_\nu = 5$, 15,

TABLE 1
Exact Parameter Values of the Simulated ^{31}P MRS Signal, Modeled by Eq. [1]

Peak k	f_k (Hz)	d_k (Hz)	a_k (a.u.)	θ_k (°)[a]
1	−86	50	75	135
2	−70	50	150	135
3	−54	50	75	135
4	152	50	150	135
5	168	50	150	135
6	292	50	150	135
7	308	50	150	135
8	360	25	150	135
9	440	285.7	1400	135
10	490	25	60	135
11	530	200	500	135

Note. $t_0 = 0$, $\Delta t = 0.333$. The values are based on the fit of an *in vivo* ^{31}P MRS signal.
[a] $\theta_k = \phi_k * 180/\pi$ expresses the phase in degrees.

TABLE 2
Sample Mean and Standard Deviation Values of Monte Carlo Simulation (300 Runs), No Prior Knowledge

σ_ν	Criterion	VARPRO	AMARES
5	% fail	0	0
	jev ± std	10.4 ± 1.0	11.5 ± 1.7
	fev ± std	13.5 ± 1.3	15.6 ± 2.4
	cpu ± std	2.6 ± 0.3	2.1 ± 0.3
15	% fail	12.3	0.33
	jev ± std	11.5 ± 3.8	14.7 ± 3.2
	fev ± std	15.8 ± 6.9	20.9 ± 5.8
	cpu ± std	2.9 ± 1.0	2.6 ± 0.6
25	% fail	13.0	7.0
	jev ± std	11.3 ± 4.2	15.0 ± 6.1
	fev ± std	14.1 ± 5.8	20.4 ± 9.7
	cpu ± std	2.8 ± 1.0	2.6 ± 1.1

Note. VARPRO and AMARES are compared. The computed quantities are % fail, the percentage of failures; jev, the average number of Jacobian evaluations; fev, the average number of functional evaluations; cpu, the average cpu time in seconds. std denotes the standard deviation.

25). A method is considered to fail if the minimization algorithm claims not to have found the solution (according to its corresponding stopping criterion) or if not all 11 peaks are resolved within specific intervals lying symmetrically around the exact frequencies. The half-widths of these intervals are 8.6, 7.3, 8.7, 3.2, 3.2, 3.5, 3.6, 0.7, 5.6, 2.4, and 7.8 Hz, respectively. These values are derived from the Cramér–Rao lower bounds on the frequencies at that noise standard deviation where the two intervals of two neighboring peaks in the β-ATP triplet touch.

For all our results we verified that these simple criteria guarantee that all methods find the same minimum if no failure occurs. As a result, root-mean-squared error, bias, and standard deviation of the parameter estimates computed by any method as a function of the noise level are the same and will not be shown.

Starting values for the dampings and the frequencies can be obtained by peak picking or by using a noninteractive method like HSVD. In a previous study (*27*) we found that for signals with a high signal-to-noise ratio HSVD is the best way of providing starting values. However, for low signal-to-noise ratio signals peak picking is preferred. Here only peak picking is considered. For every noise level used in the Monte Carlo simulation we randomly picked one signal out of 300 for which damping and frequency starting values were determined by peak picking. These starting values were then used to process all the signals affected by the same noise level.

We compared VARPRO and AMARES in terms of robustness and efficiency. The robustness could be assessed by the number of times a method fails. To compare the efficiency we looked at the average number of functional and Jacobian evaluations and the average overall CPU time. All experiments were performed on a SUN ULTRA2 (200 MHz).

In the first experiment no prior knowledge is imposed; hence four parameters for each of the 11 peaks are to be estimated. AMARES minimizes a problem in 44 variables, whereas VARPRO minimizes a problem in 22 variables. The results are displayed in Table 2. We see that VARPRO performs poorly in terms of failures compared to AMARES.

We did several tests to find the reason for this significantly poorer robustness of VARPRO. First, we changed the VARPRO code so that G was minimized instead of V. The results are displayed in Table 3. We see that the failure rate drops

TABLE 3
Sample Mean and Standard Deviation Values of Monte Carlo Simulation (300 Runs), No Prior Knowledge

σ_ν	Criterion	VARPRO-G	VARPRO-d	VARPRO-N
5	% fail	0	0	0
	jev ± std	9.9 ± 0.3	9.0 ± 0.3	8.5 ± 0.5
	fev ± std	10.9 ± 0.3	12.1 ± 0.5	11.5 ± 0.5
	cpu ± std	1.7 ± 0.1	2.2 ± 0.07	2.1 ± 0.1
15	% fail	0.33	0	0
	jev ± std	10.6 ± 1.3	10.2 ± 0.9	11.8 ± 1.5
	fev ± std	11.7 ± 1.7	14.4 ± 1.1	14.8 ± 2.0
	cpu ± std	1.8 ± 0.2	2.8 ± 0.2	2.7 ± 0.3
25	% fail	7.0	6.7	6.7
	jev ± std	12.7 ± 4.5	11.4 ± 3.5	12.3 ± 3.7
	fev ± std	14.5 ± 5.9	14.9 ± 3.6	15.1 ± 4.1
	cpu ± std	2.2 ± 0.8	2.9 ± 0.8	2.8 ± 0.8

Note. VARPRO-G is the original VARPRO code adapted to minimize the general functional G instead of the variable projection functional V. VARPRO-d is the original VARPRO code modified to minimize w.r.t. d_k, $k = 1 \cdots K$. VARPRO-N is the original VARPRO code where the LM algorithm is replaced by NL2SOL. See also the footnote to Table 2.

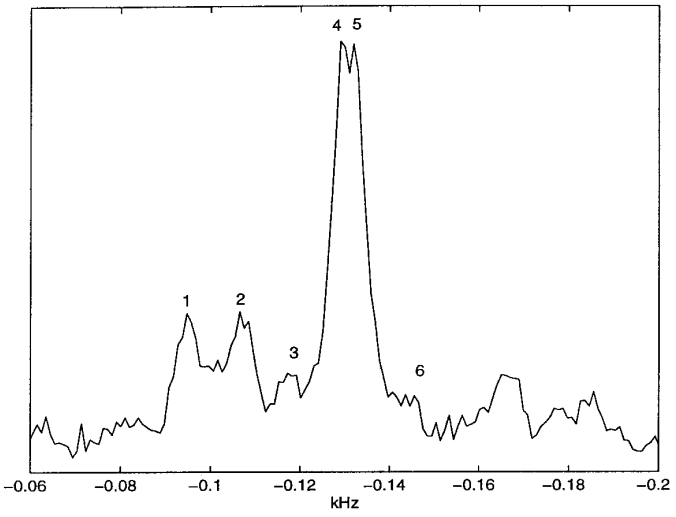

FIG. 2. *In vivo* spectrum of the prostate gland of a BPH patient. Peak 1 is choline, peak 2 is creatine, and peaks 3 to 6 constitute the citrate multiplet.

to a level comparable to that obtained with AMARES. This suggests that minimization of G is easier than that of V.

We see an equal improvement in failure rate when we take the VARPRO code and modify it so that minimization is done w.r.t. d_k, $k = 1, \ldots, K$, directly. The dampings are forced to be positive by imposing a penalty on the function if the algorithm computes damping estimates that are negative. Note however that this approach can lead to slow convergence in some cases as pointed out in (*28*).

To investigate the influence of the minimization algorithm used, we replaced the LM algorithm in VARPRO by the NL2SOL algorithm used in AMARES. The results are displayed in Table 3. Although the functional minimized is V and the minimization is done w.r.t. b_k, $k = 1, \ldots, K$, the failure rate drops to the same level as that obtained using AMARES. This clearly shows that AMARES is a more robust algorithm, able to deal with more difficult optimization problems.

When more prior knowledge is imposed or the starting values improve, leading to a reduction in complexity of the optimization problem, the differences in robustness between VARPRO and AMARES decrease. In these situations VARPRO can still be more efficient, due to the less complicated nature of the underlying optimization algorithm.

QUANTIFICATION OF AN *IN VIVO* SIGNAL USING PRIOR KNOWLEDGE

To illustrate the features of the new algorithm we tested it on a challenging data set provided by Dr. A. Heerschap of the University Hospital Nijmegen (see Acknowledgments). An *in vivo* spectrum was acquired from the prostatic gland of a benign prostatic hyperplasia patient, at 1.5 T using a PRESS localization sequence ($\tau_1 = 11$ ms, $\tau_2 = 67.5$ ms) and an endorectal coil for signal reception; 256 scans of 1024 complex FID data points were acquired from a volume of 8 cc, with a T_R of 1.6 s (see Fig. 2).

The goal of this measurement was the quantification of the citrate content, which is an important tool in the discrimination between prostate adenocarcinoma and benign prostatic hyperplasia (BPH). The citrate protons form a complicated AB-type multiplet, consisting of four Lorentzian lines with different phase. Relative intensities of the citrate lines and their phase deviations from the rest of the spectrum can be calculated (*29*) and were provided to us by Dr. M. van der Graaf (see Acknowledgments). It is also known that the linewidths of the main peaks should be equal as well as those of the side peaks.

To analyze the signal with VARPRO we put all the peaks in separate groups. We could have put the peaks of choline and creatine in one group but that would not have made any difference. The four peaks of citrate must be put in separate groups in order to impose the fixed (but different for each peak) phase shifts w.r.t. the zero-order phase. This implies that we cannot impose the known frequency shifts between the citrate peaks since these can only be imposed between peaks belonging to the same group. Amplitudes were linked

TABLE 4
Results of Analysis of the Citrate Signal with VARPRO

Peak k	$f_k \pm$ std (kHz)	$d_k \pm$ std (kHz)	$a_k \pm$ std (a.u.)	θ'_k (°)[a]
1	−0.0949 ± 0.0004	−0.0222 ± 0.0032	9.70 ± 1.30	0
2	−0.1069 ± 0.0003	−0.0251 ± 0.0040	10.42 ± 1.97	0
3	−0.1140 ± 0.0005	−0.0111 ± 0.0023	1.55 ± 0.05	−98.74
4	−0.1296 ± 0.0001	−0.0126 ± 0.0006	19.51 ± 0.64	35.60
5	−0.1317 ± 0.0001	−0.0129 ± 0.0006	19.51 ± 0.64	−35.60
6	−0.1472 ± 0.0005	−0.0111 ± 0.0023	1.55 ± 0.05	98.74

Note. Fitted zero-order phase (°) 225.6297 ± 25.8974. Fitted begin time (ms) 2.4339 ± 0.5528.
[a] $\theta'_k = \phi'_k * 180/\pi$ expresses the phase in degrees.

TABLE 5
Results of Analysis of the Citrate Signal with the New Algorithm

Peak k	$f_k \pm$ std (kHz)	$d_k \pm$ std (kHz)	$a_k \pm$ std (a.u.)	θ'_k (°)[a]
1	−0.0949 ± 0.0004	−0.0227 ± 0.0032	9.93 ± 1.27	0
2	−0.1068 ± 0.0003	−0.0240 ± 0.0037	9.95 ± 1.83	0
3	−0.1137 ± 0.0001	−0.0117 ± 0.0025	1.55 ± 0.05	−98.74
4	−0.1296 ± 0.0001	−0.0127 ± 0.0005	19.50 ± 0.63	35.60
5	−0.1317 ± 0.0001	−0.0127 ± 0.0005	19.50 ± 0.63	−35.60
6	−0.1477 ± 0.0001	−0.0117 ± 0.0025	1.55 ± 0.05	98.74

Note. Fitted zero-order phase (°) 221.1242 ± 25.2350. Fitted begin time (ms) 2.3313 ± 0.5380.

[a] $\theta'_k = \phi'_k * 180/\pi$ expresses the phase in degrees.

between the different groups and the individual phases of the peaks fixed relative to the zero-order phase. Since VARPRO allows only one overall linking for dampings between groups, we could link all four peaks to each other (theoretically incorrect), to the side peaks, or to the main peaks. In this example with the VARPRO method the linewidths of the small side peaks were linked, as these are more difficult to fit. Zero-order phase and begin time are estimated. The results are displayed in Table 4. Only with the new algorithm are we capable of imposing all the available prior knowledge. The results are displayed in Table 5.

The most important difference between the VARPRO results and those of the new method is that the latter can be obtained in a much more consistent fashion. This is entirely due to the increased use of important prior knowledge on the AB-type multiplet of citrate, in particular the chemical shift differences between its respective components. This is especially crucial at echo times where the side peaks reach their minimum intensities and sink into the noisy baseline (*30, 31*). With the combined prior knowledge of the chemical shift differences and amplitude (peak area) ratios with respect to the prominent main peaks of the quartet, the small side peaks can still be estimated reliably. Including the small side peaks in the fit is not only necessary for complete quantitation of the citrate multiplet, but also enhances the accuracy of the creatine fit, as their resonances overlap closely (see Fig. 3). Another advantage of the new method is that, in combination with all other prior knowledge, the linewidths of the side peaks can be linked to each other, and those of the main peaks can be linked to each other; see the results in Table 5. The main peaks were found to be a little broader than the side peaks, which corresponds to the expectations for the citrate multiplet (*32*). With the VARPRO method where we could only link the side peaks to each other, we obtained slightly different (although not significant, due to the high level of the noise) linewidths of the main peaks, which is undesired. Despite the slight differences in linewidths and chemical shifts between the two methods, the estimated amplitudes are identical for the citrate multiplet. This is mainly due to the available prior knowledge for the amplitude ratios with respect to the prominent main peaks, and to the precise phase constraints. As an aside, it can be noted that, perhaps because of the improved fitting of the left side peak of citrate, the fitting of creatine and choline is more according to our expectations for the new method. Their linewidths (damping factors) should not differ too much, due to the leveling effect on the effective linewidths of the magnetic field inhomogeneities and magnetic susceptibility effects. Creatine may indeed be a little broader than choline, as its T2 is shorter than that of choline (*33*).

CONCLUSIONS

We have presented AMARES, a new algorithm which outperforms the currently used reference time domain method VARPRO in two ways. First, we made an appropriate choice of the functional and the nonlinear least-squares algorithm. When estimating the parameters of a magnetic resonance spectroscopy signal in the time domain using optimization methods one has the choice between two functionals. On the one hand, we can minimize a general functional, consisting of the sum of squared differences between the data and the model function. On the other hand, a so-called variable projection functional can be derived from the general functional by eliminating all linear parameters and be optimized. We showed here that minimizing the general functional leads to a more robust determination of parameter estimates. In addition we showed that the functional should be minimized with respect to the damping with imposition of a positivity constraint instead of minimizing the functional

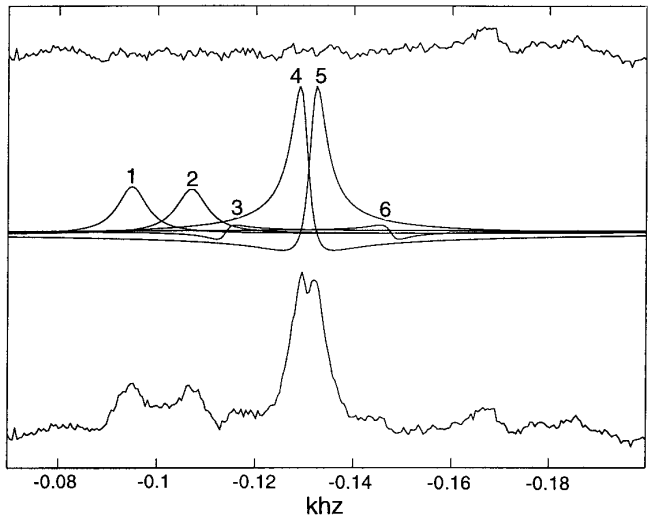

FIG. 3. Graphical result of the analysis of the citrate signal with the new algorithm. From bottom to top: the FT spectrum of the original signal, the individual Lorentzians (FT of fitted sinusoids), and the residual, which is the difference between the original signal and the reconstructed signal.

with respect to the square root of the damping as is done in VARPRO. In this way numerical problems are avoided.

Second, AMARES allows the inclusion of more prior knowledge about the signal parameters, the model function (Lorentz, Gauss), and the type of signal (FID or echo). This results in increased accuracy and user flexibility. In addition, imposing all prior knowledge leads to a problem in a smaller number of variables which can be solved in less time. Overall we obtain an algorithm which outperforms VARPRO in terms of accuracy, robustness, and flexibility.

ACKNOWLEDGMENTS

The authors thank Dr. Marinette van der Graaf and Dr. Arend Heerschap of the Department of Radiology, University Hospital Nijmegen St. Radboud, for supplying the example data set and the relevant molecular prior knowledge.

The first author is a Ph.D. student funded by the IWT (Flemish Institute for Support of Scientific–Technological Research in Industry). The third author is a research associate with the FWO (Fund for Scientific Research—Flanders). This work is supported by the Belgian Programme on Interuniversity Poles of Attraction (IUAP-4/2 & 24), initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture, by the EU Programme ''Human Capital and Mobility, Networks,'' Project CHRX-CT94-432, and by a Concerted Research Action (GOA) project of the Flemish Community, entitled ''Model-Based Information Processing Systems.''

REFERENCES

1. R. Kumaresan and D. W. Tufts, *IEEE Trans.* **ASSP 30,** 833 (1982).
2. H. Barkhuysen, R. de Beer, W. M. M. J. Bovée, and D. van Ormondt, *J. Magn. Reson.* **61,** 465 (1985).
3. S. Y. Kung, K. S. Arun, and D. V. Bhaskar Rao, *J. Opt. Soc. Am.* **73,** 1799 (1983).
4. H. Barkhuysen, R. de Beer, and D. van Ormondt, *J. Magn. Reson.* **73,** 553 (1987).
5. W. W. F. Pijnappel, A. van den Boogaart, R. de Beer, and D. van Ormondt, *J. Magn. Reson.* **97,** 122 (1992).
6. S. Van Huffel, H. Chen, C. Decanniere, and P. Van Hecke, *J. Magn. Reson. A* **110,** 228 (1994).
7. H. Chen, S. Van Huffel, C. Decanniere, and P. Van Hecke, *J. Magn. Reson. A* **109,** 46 (1994).
8. H. Chen, S. Van Huffel, D. van Ormondt, and R. de Beer, *J. Magn. Reson. A* **119,** 225 (1996).
9. R. de Beer and D. van Ormondt, *in* ''NMR Basic Principles and Progress'' (M. Rudin, Ed.), Vol. 26, p. 201, Springer-Verlag, Berlin (1992).
10. J. W. C. van der Veen, R. de Beer, P. R. Luyten, and D. van Ormondt, *Magn. Reson. Med.* **6,** 92 (1988).
11. C. Decanniere, P. Van Hecke, H. Chen, S. Van Huffel, C. van der Voort, B. van Tongeren, and D. van Ormondt, *J. Magn. Reson. B* **105,** 31 (1994).
12. A. van den Boogaart, F. A. Howe, L. M. Rodrigues, M. Stubbs, and J. R. Griffiths, *NMR Biomed.* **8,** 87 (1995).
13. G. H. Golub and V. Pereyra, *SIAM J. Numer. Anal.* **10,** 413 (1973).
14. F. S. DiGennaro and D. Cowburn, *J. Magn. Reson.* **96,** 582 (1992).
15. G. J. Metzger, M. Patel, and X. Hu, *J. Magn. Reson. B* **110,** 316 (1996).
16. M. R. Osborne, *in* ''Numerical Methods for Non-linear Optimization'' (Lootsma, Ed.), Academic Press, London (1972).
17. P. E. Gill, W. Murray, and M. H. Wright, ''Practical Optimization,'' Academic Press, San Diego (1988).
18. A. Knijn, R. de Beer, and D. van Ormondt, *J. Magn. Reson.* **97,** 444 (1992).
19. ftp://netlib.att.com.
20. J. E. Dennis and R. B. Schnabel, ''Numerical Methods for Unconstrained Optimisation and Nonlinear Equations,'' Prentice–Hall, Englewood Cliffs, NJ (1983).
21. J. D. Gorman and A. O. Hero, *IEEE Trans. Inform. Theory* **26,** 1285 (1990).
22. L. Kaufman, *BIT* **15,** 49 (1975).
23. A. van den Boogaart, S. Cavassila, L. Vanhamme, J. Totz, and P. Van Hecke, *in* ''Proceedings, 18th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.,'' Amsterdam, The Netherlands (1996). See also http://mrui-web.uab.es/mruiwww/mrui_hom.html.
24. A. van den Boogaart, M. Ala-Korpela, J. Jokisaari, and J. R. Griffiths, *Magn. Reson. Med.* **31,** 347 (1994).
25. M. Stubbs, A. van den Boogaart, C. L. Bashford, P. M. C. Miranda, L. M. Rodrigues, F. A. Howe, and J. R. Griffiths, *Biochim. Biophys. Acta* **1291,** 143 (1996).
26. L. Vanhamme, S. Van Huffel, and A. van den Boogaart, *in* ''Proceedings, ISMRM,'' 5th Scientific Meeting, pp. 1414, Vancouver, Canada (1997).
27. L. Vanhamme, A. van den Boogaart, and S. Van Huffel, *in* ''Proceedings, EUSIPCO-96,'' Trieste, Italy (1996).
28. J. J. Moré, B. S. Garbow, and K. E. Hillstrom, ''User Guide for Minpack-1,'' ANL-80-74, Argonne, IL (1980).
29. M. van der Graaf, A. van den Boogaart, W. M. M. J. Bovée, and A. Heerschap, *in* ''Proceedings, ISMRM,'' 5th Scientific Meeting, p. 1419, Vancouver, Canada (1997).
30. M. van der Graaf, G. J. Jager, and A. Heerschap, *MAGMA* **5,** 65 (1997).
31. M. van der Graaf, G. J. Jager, and A. Heerschap, *MAGMA* **4,** Suppl. 2, 320 (1996).
32. W. M. M. J. Bovée, *J. Magn. Reson.* **24,** 327 (1976).
33. A. van den Boogaart, ''The Use of Signal Processing Algorithms to Obtain Biochemically Relevant Parameters from Magnetic Resonance Data Sets,'' Ph.D. thesis, U. London (1995).