Journal of Magnetic Resonance **155,** 226–235 (2002)
doi:10.1006/jmre.2002.2521, available online at http://www.idealibrary.com on IDEAL®

A New Time-Domain Frequency-Selective Quantification Algorithm

Rocco Romano,*,[1] Andrea Motta,*,† Stefania Camassa,‡ Claudia Pagano,‡ Maria Teresa Santini§,* and Pietro Luigi Indovina*,‡

*Istituto Nazionale per la Fisica della Materia, Unità di Napoli, c/o Dipartimento di Scienze Fisiche, Università degli Studi di Napoli Federico II, Complesso Universitario di Monte S. Angelo, via Cinthia, I-80126 Naples, Italy; †Istituto di Chimica Biomolecolare del CNR, I-80078 Pozzuoli, Naples, Italy; ‡Dipartimento di Scienze Fisiche, Università degli Studi di Napoli Federico II, Complesso Universitario di Monte S. Angelo, via Cinthia, 80126 Naples, Italy; and §Laboratorio di Ultrastrutture, Istituto Superiore di Sanità, Viale Regina Elena 299, I-00161 Rome, Italy

Received July 30, 2001; revised December 26, 2001

In this paper a new time-domain frequency-selective quantification algorithm is presented. Frequency-selective quantification refers to a method that analyzes spectral components in a selected frequency region, ignoring all the other components outside. The algorithm, referred to as MeFreS (Metropolis Frequency-Selective), is based on rank minimization of an opportune Hankel matrix. The minimization procedure is satisfied by the down-hill simplex method, implemented with the simulated annealing method. MeFreS does not use any preprocessing step or filter to suppress nuisance peaks, but the signal model function is directly fitted. In this manner, neither inherent signal distortions nor estimation biases to be corrected occur. The algorithm was tested with Monte Carlo simulations. A comparison with VARPRO and AMARESw algorithms was carried out. Finally, two samples of known content from NMR data were quantified.

Key Words: **frequency selective algorithm; time domain algorithm; quantification in NMR; NMR spectroscopy.**

INTRODUCTION

Frequency-Selective (FS) quantification is a method that applies to spectral components of a selected frequency region, ignoring those outside the region of interest (*1*). Limiting the data analysis to a narrow region of an NMR spectrum is certainly a useful advantage. Frequency domain quantification of a Fast-Fourier Transform (FFT) NMR Free Induction Decay (FID), is frequency-selective by nature, as various spectral components are distributed along the frequency axis. However, the selective analysis in the frequency domain has some heavy limitations. Spectra require several manipulations (e.g., phase adjustment, baseline correction), and, for truncated time-domain data, the FFT needs to be deconvoluted before or during the quantitative data analysis (*2*). Both aspects are very limiting in 1*D* as well as in multidimensional NMR spectroscopy. Currently used FS methods can be divided in two principal categories: methods that can be used in combination with black-box time-domain procedures (*3, 4*) and methods based on model function fitting, which allow to impose prior knowledge (*1, 5–7*). Among model function fitting methods, some important examples can be considered: the ER-filter preprocessing method (*7*), and solvent suppression methods based on a number of preprocessing techniques to remove the influence of nuisance peaks (peaks that are in the same frequency region but are unwanted) prior to a time-domain model-fitting procedure (*8–11*). Furthermore, the use of a minimum-phase FIR filter to suppress nuisance peaks, followed by the fitting of an adapted time-domain model function, has been proposed (*5*).

In this paper, we present a time-domain frequency-selective quantification algorithm for quantitative analysis of NMR spectra. The MeFreS (Metropolis Frequency Selective) algorithm works in the time-domain with the advantage of being frequency-selective. Compared with the above methods, MeFreS does not use any preprocessing step or filter to suppress nuisance peaks, but the signal model function is directly fitted. In this manner, neither inherent signal distortions nor estimation biases to be corrected occur. Furthermore, when complex spectra such as those of *in vitro* cells are considered, the great number of components requires a frequency selective algorithm. In fact, the simultaneous analysis of too many parameters, for example with LP–SVD (Linear Prediction–Singular Value Decomposition), implies solving simultaneously a large set of linear equations, as well as rooting a high-order polynomial equation. This would require an unreasonable and often extremely excessive amount of computational time (*4*). Furthermore, fitting simultaneously a great number of components, and then of parameters, could likely give very large errors. FS and a priori knowledge render VARPRO (*1*) and AMARESw (*5*), two time-domain frequency selective iterative techniques, to be an efficient solution in the current state-of-the-art of *in vivo* NMR spectroscopy. However, AMARESw and VARPRO are expected to give good results for relatively well-separated peaks (*1, 5*). On the contrary, if nuisance peaks have a large amplitude or are close, in frequency, to the peaks of interest, the methods break down. In particular, this always happens when the assumption of "*large*" frequency separation no longer holds (*5*). Furthermore, as suggested in (*1*),

[1] To whom correspondence should be addressed. Fax: (+39) 081676346. E-mail: rocco.romano@na.infn.it.

1090-7807/02 $35.00

"should unwanted components visibly overlap with the wanted ones when viewed in the frequency domain, we recommend that such components be included in the fit." This means that often more than one component has to be simultaneously fitted, so reducing the advantage of frequency selectivity. Both VARPRO and AMARESw are implemented with a time-domain weighting function "consisting of a quarter wave sinusoid for the first (and last) twenty samples. . ." (*5*). This leads "to a loss of SNR (Signal to Noise Ratio) resulting in an increased variance of the parameter estimates" (*5*) and removes peaks with large linewidth.

The MeFreS algorithm was designed in order to overcome these limitations certainly relevant in high resolution NMR. For example, cell spectra contain a great number of overlapping signals well below the AMARESw and VARPRO breakdown frequency separation. Furthermore, a large linewidth is often encountered for many interesting signals that cannot be removed from the spectra. The MeFreS algorithm should give a solution to the above limitations. In fact, as no weighting function or filter is used, no signal is removed. Only one signal at a time can be fitted and consequently a maximum of four parameters at a time is fitted. Finally, it appears to be truly frequency selective as it allows to fit without bias a single signal which is close to large amplitude or to large linewidth spectral lines.

The MeFreS algorithm is based upon an opportune Hankel matrix rank minimization. The rank is determined by comparing singular values, obtained by Singular Value Decomposition (SVD) (*12–17*), of two opportune Hankel matrices (*18, 19*). The level of difference between the singular values is a function of K continuous parameters (K being the number of the unknown signals' parameters which must be fitted), while it assumes only a small, discrete number of values. Therefore, minimization of a discrete value function, defined in a continuous K-dimensional space (K-dimensional configuration space), is required. Being the function of a discrete-value function, it requires a multidimensional minimization procedure with only function evaluations (no derivatives), while the small number of assumed values needs a random choice of points to be tested. The first request can be satisfied by the *downhill simplex method* (*20, 21*), while the second can be satisfied by the *method of simulated annealing* (*22–24*). In particular, for the minimization of the function, we resorted to a *simulated annealing* for continuous K-dimensional configuration space (*25*), which uses a modification of the *downhill simplex method*. This procedure, due to Press *et al*. (*25*), has been modified to take into account the discrete value properties of the function. In order to test the MeFreS algorithm, we carried out Monte Carlo simulations, which demonstrate its ability to determine one signal in a selected range of frequencies, for different SNR. Furthermore, we compared MeFreS with VARPRO and AMARESw applying all algorithms to limit situations. Finally, we analyzed experimental NMR spectra of samples at known concentration to test the algorithm ability for quantitative evaluation.

RESULTS AND DISCUSSION

The Algorithm

A time-domain NMR FID can be modeled as a sum of complex exponentially decaying sinusoids,

$$x_n = \sum_{s=1}^{s=S} A_s \exp[i(\phi_0 + \phi_s)] \exp[(-\alpha_s + i2\pi\nu_s)t_n] + e_n \qquad [1]$$

where S is the number of sinusoids, and A_s, α_s, ν_s, and ϕ_s ($s = 1, 2, \ldots, S$) are the amplitude, decaying factor (in Hz), frequency (in Hz), and phase (in rad), respectively, of the sth sinusoid. The value of ϕ_0 is the zero-order phase, and e_n is complex white Gaussian noise. The number of complex data points is Np, and the discretely sampled time steps are $t_n = (n + \eta)\Delta t$, $n = 0, 1, \ldots, Np - 1$, with $t_0 = \eta\Delta t$ the begin time, or dead time of the spectrometer (*26, 27*). At this point, a FID $X(t)$ and a signal to be tested $\hat{X}(t)$ will be considered. They can be written, respectively, as

$$x_n = \exp(i\phi_0)\left[\sum_{s=1}^{s=P} P_s \exp(i\phi_s) \exp[(-\alpha_s + i2\pi\nu_s)t_n] + \sum_{s=P+1}^{s=S} A_s \exp(i\psi_s) \exp[(-\beta_s + i2\pi\upsilon_s)t_n]\right] + e_n \qquad [2]$$

and

$$\hat{x}_n = \sum_{s=1}^{s=P} \hat{P}_s \exp(i\hat{\phi}_s) \exp[(-\hat{\alpha}_s + i2\pi\hat{\nu}_s)t_n], \qquad [3]$$

where $n = 0, 1, \ldots, Np-1$, ϕ_0 is the zero-order phase, e_n is the complex Gaussian noise, and $t_n = (n + \eta)\Delta t$, with $t_0 = \eta\Delta t$, the begin time of the FID; S is the total number of sinusoids, $P \leq S$ is the number of signals to be tested (P can be even equal to 1), $P_s, \phi_s, \alpha_s, \nu_s$ ($s = 1, 2, \ldots, P$) are their amplitude, phase, damping factor, and frequency (in Hz), respectively, while A_s, ψ_s, β_s, and υ_s ($s = P + 1, P + 2, \ldots, S$) are the amplitude, phase, damping factor, and frequency (in Hz), respectively, of the remaining FID sinusoids; $\hat{P}_s$, $\hat{\phi}_s$, $\hat{\alpha}_s$, $\hat{\nu}_s$ ($s = 1, 2, \ldots, P$) are the amplitude, phase, damping factor, and frequency (in Hz), respectively, of the signal to be tested. We denote by

$$\delta\hat{x}_n = x_n - \hat{x}_n, \qquad [4]$$

where $n = 0, 1, \ldots, Np - 1$, and the Hankel matrix by

$$\Delta X(\hat{P}_s, \hat{\phi}_s, \hat{\alpha}_s, \hat{\nu}_s (s = 1, 2, \ldots, P)) = \begin{pmatrix} \delta\hat{x}_0 & \delta\hat{x}_1 & \cdots & \delta\hat{x}_{M-1} \\ \delta\hat{x}_1 & \delta\hat{x}_2 & \cdots & \delta\hat{x}_M \\ \vdots & \vdots & \vdots & \vdots \\ \delta\hat{x}_{L-1} & \delta\hat{x}_L & \cdots & \delta\hat{x}_{N-1} \end{pmatrix} \qquad [5]$$

with $N \leq Np$, L and M chosen greater than $S+P$ and subjected to the constraint $N = L + M - 1$. For $P_s \neq \hat{P}_s, \phi_s \neq \hat{\phi}_s, \alpha_s \neq \hat{\alpha}_s, \nu_s \neq \hat{\nu}_s\ (s = 1, 2, \ldots, P)$ these matrices will have a rank approximated at the most by $S + P$, because in ΔX $(\hat{P}_s, \hat{\phi}_s, \hat{\alpha}_s, \hat{\nu}_s\ (s = 1, 2, \ldots, P))$ there will be a number of independent complex decaying sinusoids equal at the most to the sum of the number of X and $\hat{X}$ complex decaying sinusoids. However, for $P_s = \hat{P}_s, \phi_s = \hat{\phi}_s, \alpha_s = \hat{\alpha}_s, \nu_s = \hat{\nu}_s\ (s = 1, 2, \ldots, P)$, the rank of $\Delta X(\hat{P}_s, \hat{\phi}_s, \hat{\alpha}_s, \hat{\nu}_s\ (s = 1, 2, \ldots, P))$ will be approximated at the most by $S - P$, because the P complex decaying sinusoids in X and $\hat{X}$ cancel each other out. In order to obtain the unknown parameters $P_s, \phi_s, \alpha_s, \nu_s\ (s = 1, 2, \ldots, P)$, the algorithm consists in calculating the minimum rank of the Hankel matrix in Eq. [5] when the parameters vary in a priori given intervals. The a priori given intervals can be rather easily obtained by roughly looking at the FFT of the data or by peak peaking. In fact, algorithm works well also with parameter intervals which are so large that a 100% parameter variation can occur. At this point two questions need to be considered. First, the rank of the Hankel matrix in Eq. [5], which is a discrete value function of the test parameters $\hat{P}_s, \hat{\phi}_s, \hat{\alpha}_s, \hat{\nu}_s\ (s = 1, 2, \ldots, P)$, must be calculated; second, the minimum of the rank discrete value function must be found.

Rank Determination

Given the time domain NMR FID, let us consider the $(N - M + 1) \times M$ Hankel matrix

$$X(N-M+1, M) = \begin{pmatrix} x_0 & x_1 & \cdots & x_{M-1} \\ x_1 & x_2 & \cdots & x_M \\ \vdots & \vdots & \vdots & \vdots \\ x_{N-M} & x_{N-M+1} & \cdots & x_{N-1} \end{pmatrix} \quad [6]$$

which consists of N uniformly sampled data points x_n, $n = 0, 1, \ldots, N-1$, with $(N-M+1)$ and M chosen greater than S. The SVD theorem (*14–17*) states that if X is an arbitrary $(N - M + 1) \times M$ complex valued matrix, then there exist unitary matrices $U((N - M + 1) \times (N - M + 1))$, $V(M \times M)$ and p ordered real numbers $(p = min((N - M + 1), M))$ $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p > 0$, such that

$$X = U \Sigma V^{\dagger}, \quad [7]$$

where $\Sigma((N-M+1)\times M)$ is such that $\Sigma = diag(\sigma_1, \sigma_2, \ldots, \sigma_p)$ and the † denotes Hermitian conjugation. The p numbers are the so-called singular values of the X matrix. If the X matrix has rank equal to S, only its first S singular values are greater than zero, that is, for a S rank X matrix one has

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_S > 0 \quad [8]$$

$$\sigma_{S+1} = \sigma_{S+2} = \cdots = \sigma_p = 0. \quad [9]$$

An X Hankel data matrix of a noiseless FID, comprising S complex decaying sinusoids, has rank equal to S (Kronecker's theorem (*12*)) because its elements are points of the FID and then they are obtained by a linear combination of S independent signal components. Each X Hankel matrix row (column) is a linear combination of the same S independent signal components, and then the X Hankel matrix can have only S independent rows (columns). The rank of the X Hankel matrix is S. As a consequence, the X Hankel matrix has only S singular values different from zero (*19*). If the FID is affected by noise, its X Hankel matrix becomes a full rank matrix because the noise destroys the linear dependence of the rows (columns): noise can be considered as the combination of infinite independent signals. However, if the SNR is not too low, that is if the signal amplitudes can be considered greater than noise amplitude, the signal related are very much greater than the noise related singular values and the rank of the X Hankel matrix still can be approximated by S. In particular, in the ordered singular values there will be a discontinuity between signal related and noise related singular values: a low limit for the singular values can be assumed and then an estimation for the rank of the X matrix can be obtained (*19, 29*). However, for low SNR, the discontinuity between signal related and noise related singular values is not so evident, so that it is not so easy to correctly make an assumption about the low limit and to obtain a correct rank estimation. In Fig. 1, the Hankel matrix singular values of a simulated FID ($SNR = 44.3$), containing 15 complex exponentially decaying sinusoids, are plotted. As it can be seen, the noise brings about an evident discontinuity in the range 10–11, while the rank of the Hankel matrix should be 15.

In order to calculate the rank matrix in Eq. [6], let us start to consider the singular value decomposition in Eq. [7]; denoted by $Tr(X)$ the X matrix trace, it is easy to show that

$$Tr(X(N - M + 1, M) * X^{\dagger}(N - M + 1, M)) = \sum_{i=1}^{p} \sigma_i^2, \quad [10]$$

where $N \leq Np$ is the number of sampled data points, M the matrix column number, and $N-M+1$ is the matrix row number. Furthermore, by simple calculations, it can be shown that fixed a given odd N, the $Tr(X(N - M + 1, M) * X^{\dagger}(N - M + 1, M))$ has the following properties:

1. $Tr(X(N - M + 1, M) * X^{\dagger}(N - M + 1, M))$ is a growing M function for $M \leq (N + 1)/2$;
2. $Tr(X(N-M+1, M)*X^{\dagger}(N-M+1, M))$ is a decreasing M function for $M \geq (N + 1)/2$;
3. $Tr(X(N-M+1, M)*X^{\dagger}(N-M+1, M))$ has a maximum for $M = (N + 1)/2$.

For a given odd N, let us consider two M values, M_1, M_2, with $S < M_2 < M_1 = (N + 1)/2$, being S the number of FID signals, and let us denote by $\sigma_{M_1}, \sigma_{M_2}$ the matrix singular values $X(N - M_1 + 1, M_1)$, $X(N - M_2 + 1, M_2)$, respectively. By Eq. [10] and

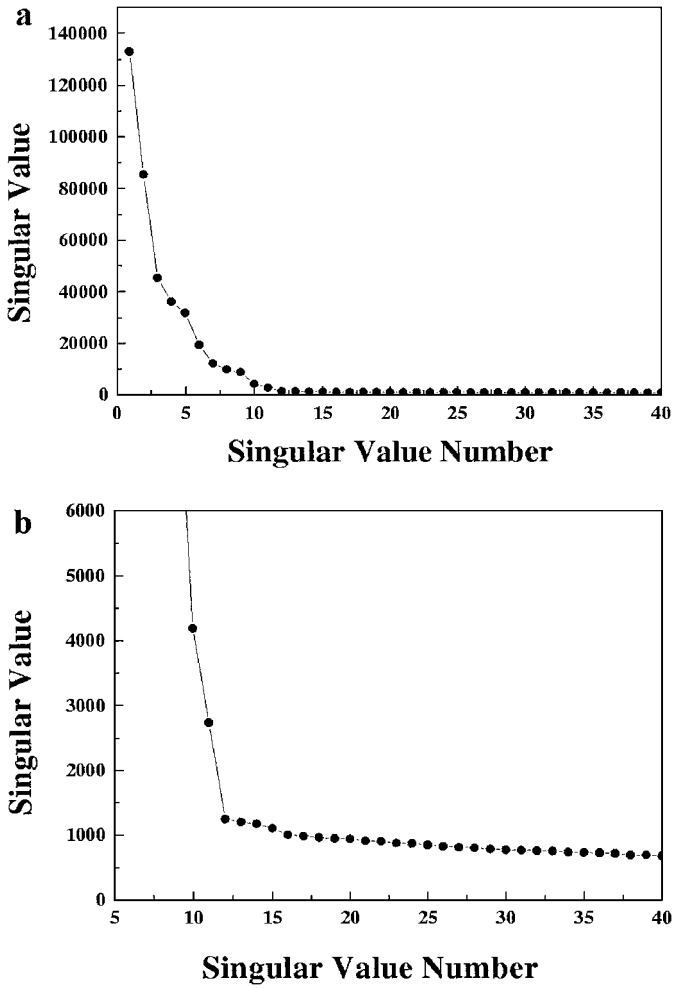

FIG. 1. Hankel matrix ordered singular values obtained by a simulated FID containing 15 complex exponentially decaying sinusoids ($SNR = 44.3$): (a) the first 40 ordered singular values; (b) zoom in of the last 35 ordered singular values in the range 5–40.

by Property 1, it follows that

$$\sum_{i=1}^{M_1} \sigma_{M_1 i}^2 > \sum_{i=1}^{M_2} \sigma_{M_2 i}^2. \qquad [11]$$

When a free noise FID is considered, both $X(N - M_1 + 1, M_1)$, $X(N - M_2 + 1, M_2)$ matrices will have S singular values, related to signals, different from zero, while the remaining are zero. So, when an FID affected by noise is considered, one expects that, due to Eq. [11], signal related singular values $\sigma_{M_1 i}$ are quite different from $\sigma_{M_2 i}$, $i = 1, 2, \ldots, S$, while noise related singular values are expected to be of the same order ($\sigma_{M_1 i} \simeq \sigma_{M_2 i}$, $i = S + 1, S + 2, \ldots, M_2$). If the absolute differences between $\sigma_{M_1 i}$ and $\sigma_{M_2 i}$, $i = 1, 2, \ldots, M_2$ are considered, there will be a discontinuity between absolute signal related singular value differences and noise related singular value differences. In order to obtain major differences between signal-related singular values, by Property 3, M_1 is chosen equal to $(N + 1)/2$, while, by Property 1, M_2 must be chosen such that $DNC \equiv M_1 - M_2 \gg 1$ with the condition $M_2 > S$. Furthermore, the first N data points are chosen because in the first points of a FID there is the major information about signals. In fact, all signals contribute to the first FID points, while in the last FID points only noise or not completely decayed components give contribute.

In Fig. 2a, the first 40 absolute singular values differences, relative to the above considered simulated FID, containing 15 complex exponentially decaying sinusoids ($SNR = 44.3$), are plotted. Singular values were obtained by SVD of two Hankel matrices having, respectively, $N = 199$, $M_1 = 100$, and three different DNC values ($DNC = 15, 17, 19$). As it can be seen, the last significant difference about singular values occurs at the 15th singular value, Fig. 2b. Then, the Hankel matrix rank in Eq. [6], obtained by the simulated FID, is 15. In order to obtain

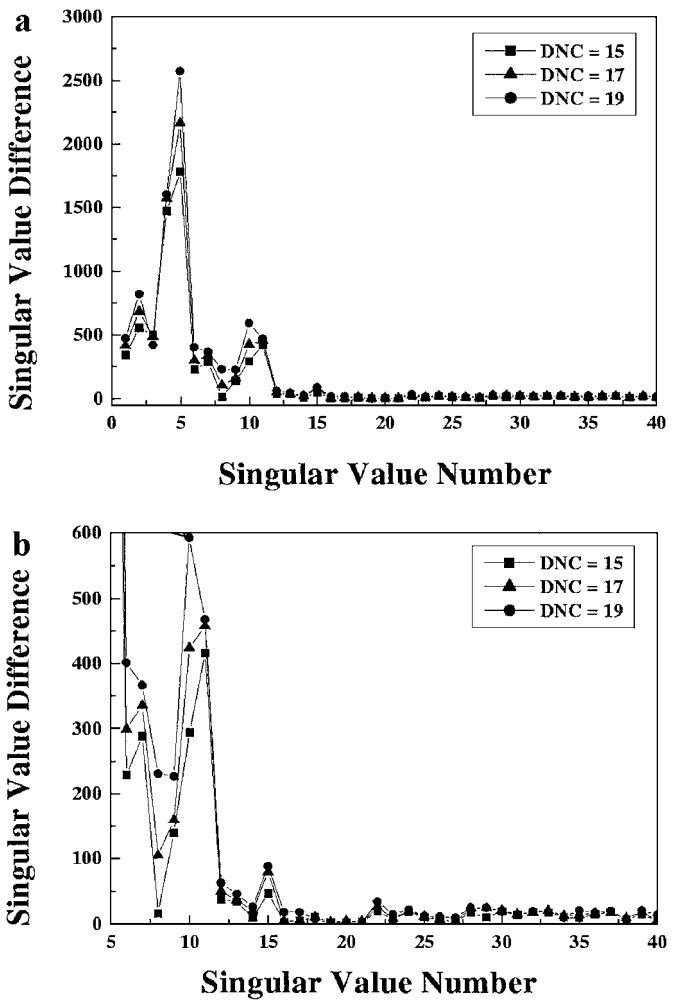

FIG. 2. Absolute difference of Hankel matrix ordered singular values obtained by a simulated FID containing 15 complex exponentially decaying sinusoids ($SNR = 44.3$): (a) the first 40 absolute singular value differences; (b) zoom in of the last 35 absolute singular value differences in the range 5–40.

the unknown parameters, the Hankel matrix rank in Eq. [6] must be minimized.

To obtain the index of the singular value difference, which separates signal related from noise related singular value differences, an automatic procedure was used. In particular, if $\delta_{p-\frac{p}{4}+1}$, $\delta_{p-\frac{p}{4}+2}$, $\delta_{p-\frac{p}{5}}$, ($\delta_j \equiv \sigma_{M_1 j} - \sigma_{M_2 j}$, $j = 1, 2, \ldots, M_2$, $p = M_2$) were difference values in the last quarter of singular value differences, certainly all noise related, denoted by

$$\tau = mean\left\{\delta_{p-\frac{p}{4}+1}, \delta_{p-\frac{p}{4}+2}, \ldots, \delta_{p-\frac{p}{5}}\right\} \qquad [12]$$

the index required was the last j, with $j = 1, 2, \ldots, p-p/5-1$, for which the condition $(\delta_j - \delta_{j+1}) \geq 2.5 * \tau$ was satisfied.

However, the correct rank determination is not a crucial point, because it is sufficient that the used rank determination is such that it is able to emphasize the passage from a major to a minor rank. So, the proposed rank determination is not an essential requirement and better rank determination procedures can only improve the algorithm.

The Minimization Procedure

The Hankel matrix rank in Eq. [6] is a K continuous parameter function (K being the number of the unknown signals' parameters to be fitted), defined in a continuous K-dimensional space (K-dimensional configuration space, which assumes only a little discrete number of values). The discrete-value nature of the function requires a multidimensional minimization procedure with only function evaluations (no derivatives), while a small number of assumed values needs a random choice of points to be tested. The second request can be satisfied by the *method of simulated annealing* (*22–24*), while the first request by the *downhill simplex method* (*20, 21*). The *method of simulated annealing* is a technique suitable for optimization problems, especially those with the desired global extremum hidden among many local extrema (*25*). This numerical method, due to Metropolis *et al*. (*24*), is based on an analogy with thermodynamics, specifically with the way that metals cool or anneal.

In the Metropolis algorithm, a simulated thermodynamic system is considered and it is assumed to change its configuration from energy E_1 to energy E_2 with probability $p = \exp[-(E_2 - E_1)/kT]$. In this manner, the system always accepts changes toward lower energy states, while the probability to accept changes toward greater energy states decreases with temperature lowering. If temperature is lowered sufficiently slowly, the system will be trapped in a global minimum energy state. In order to generate random state, the *downhill simplex method* (*20, 21*) was used. A simplex is a geometric figure that has one more vertex than the dimensions of the space in which it is defined. For example, a simplex on a plane is a triangle, the simplex used to fit one complex exponentially decaying sinusoid, which depends upon 4 parameters (a four-dimensional space), has 5 vertexes, and so on. Each vertex represents a state. To obtain the function minimum, the simplex was modified following the strategy described (*20, 21*). The implementation of the Metropolis procedure consists in adding a positive, logarithmically distributed random variable, proportional to the temperature T, to the stored function value associated with every vertex of the simplex, and in subtracting a similar random variable from the function value of every new point that is tried as a replacement point (*25*). This method always accepts a true downhill step, but sometimes accepts an uphill one, while in the limit $T \to 0$, reduces to the downhill simplex method. As a T decreasing schedule, the $T = T_0(1 - k/K)^\alpha$, where K is the total number of moves, k is the cumulative number of moves thus far, and α is a constant, say 1, 2, or 4 has been implied. The *amebsa* and *amotsa* routines proposed by Press *et al*. (*25*) were slightly modified to take into account the discrete value property of the rank function that had to be minimized. In particular, two vectors were introduced, in which all the best points found were memorized; and the going out condition was obtained controlling that all vector elements were equal. When the program terminated, the introduced vector contained all equal minimum found. The vector mean and standard deviation gave results with their respective errors.

Simulations Testing the Algorithm

In order to test the MeFreS algorithm, both simulated and experimental ^1H NMR FIDs of samples at known solute concentration were used. The simulated FIDs were generated by superposition of complex exponentially decaying sinusoids with additive Gaussian noise. The program was written in Matlab, and in order to generate random numbers, the *randn* Matlab function was used. Each noise realization FID consisted of the same 58 complex decaying sinusoids (Fig. 4a). The range of amplitudes, in arbitrary units, went from 50 to 13500, the decaying constant range was $(5 - 230)$ Hz, while the frequency range was from -1380 Hz to $+750$ Hz. In order to test the algorithm, seven chosen signals, selected from the 58 complex exponentially decaying sinusoids, were fitted for different SNR. The signals were chosen in such a manner that MeFreS ability in getting over the well known limits of other algorithms could be tested (Fig. 4b). In particular, we chose two very large signals (Freq. -762.1 Hz, Decay 210.5 Hz, Ampl. 1837 (black triangle right); Freq. -31.9 Hz, Decay 155.5 Hz, Ampl. 1245 (black triangle left)) one of which had also a component separated by less than 3 Hz (Freq. 28.8 Hz, Decay 67.7 Hz, Ampl. 962). Another component (Freq. 516.6 Hz, Decay 50.2 Hz, Ampl. 1330 (black square)) was chosen having a close component (at a frequency separation less than 8 Hz) with a greater amplitude (Freq. 507.5 Hz, Decay 65.4 Hz, Ampl. 1623). The fourth was chosen as a part of a doublet with a J separation of 10 Hz (Freq. 439.0 Hz, Decay 30.7 Hz, Ampl. 1633 (black triangle)). The fifth was chosen to be a very great amplitude signal (Freq. 235.8 Hz, Decay 26.2, Ampl. 6026 (black lozenge)) with very big near signals (amplitude greater then 3000 and at a frequency separation less then 5 Hz). The last two were always chosen in such a manner that they had near signals at a

frequency separation less then 7 Hz, but above all with a great number of near components (8 components in the range of 80 Hz) (Freq. 389.7 Hz, Decay 41.7, Ampl. 1716 (bullet); Freq. 306.7 Hz, Decay 17.8, Ampl. 884 (black triangle down)). These components were fitted, one by one and at different signal to noise ratio.

For each noise realization, the first 256 FID data points were used. Furthermore, for low SNR an exponential apodization function (broadening constant range (0.6–8.0) Hz) was applied (30). In Fig. 4a, together with the FFT of a noise realization simulated FID ($SNR = 23.3$), we report the FFT of the above cited 7 complex exponentially decaying sinusoids, fitted by the MeFreS algorithm one by one, and chosen among the 58 complex exponentially decaying sinusoids that formed the FID. The above 7 complex exponentially decaying sinusoids are compared, in Fig. 4b, with the FFT of the computed complex exponentially decaying sinusoids as fitted by MeFreS to the above total noisy FID ($SNR = 23.3$). As it can be seen (Fig. 4b), they are very similar, and the estimated frequencies, decays and amplitudes (filled symbols in Figs. 3a–3c) well correspond to computed data (empty symbols in Figs. 3a–3c) at different signal to noise ratio. The estimated frequencies, decays and amplitudes out of range in Figs. 3a–3c had a maximum bias less than 3% and were, in the error, consistent with the true parameter values. In order to compare MeFreS with both AMARESw and VARPRO, the same seven signals were fitted by AMARESw and VARPRO at different signal to noise ratio. In particular, the MRUI version 99.2 software package (http://www.mrui.uab.es/mrui/mrui_homePage.shtml) was used.

In Fig. 5, results about the fitted amplitudes of the above cited 7 signals at $SNR = 14.0 \pm 4.0$ are reported. We applied AMARESw, VARPRO, and MeFreS with the only a priori condition that phases were put equal to zero. To reduce the bias for AMARESw and VARPRO, the fit was extended to components in the 80 Hz range around the component of interest (generally the fit of 10, 12 components was necessary). Initial data for AMARESw and VARPRO were obtained by peak-peaking the region of interest of the frequency spectra. Simulations confirm the known AMARESw and VARPRO limitations, Fig. 5. As expected AMARESw and VARPRO amplitude estimation of the very large linewidth signals (Freq. −762.1 Hz, Decay 210.5 Hz, Ampl. 1837 (black triangle, right); Freq. −31.9 Hz, Decay 155.5 Hz, Ampl. 1245 (black triangle, left)) result very biased, while MeFreS estimation is unbiased. AMARESw and VARPRO give biased estimation also for the very large amplitude signal (Freq. 235.8 Hz, Decay 26.2, Ampl. 6026 (black lozenge)) very close to large amplitude signals (amplitudes greater than 3000 and at a frequency separation less than 5 Hz). As expected, also the signal (Freq. 389.7 Hz, Decay 41.7, Ampl. 1716 (bullet)), close to a great number of signals (8 components in the range of 80 Hz) and for which was necessary to fit simultaneously a great number of components, was estimated by AMARESw and VARPRO with a great

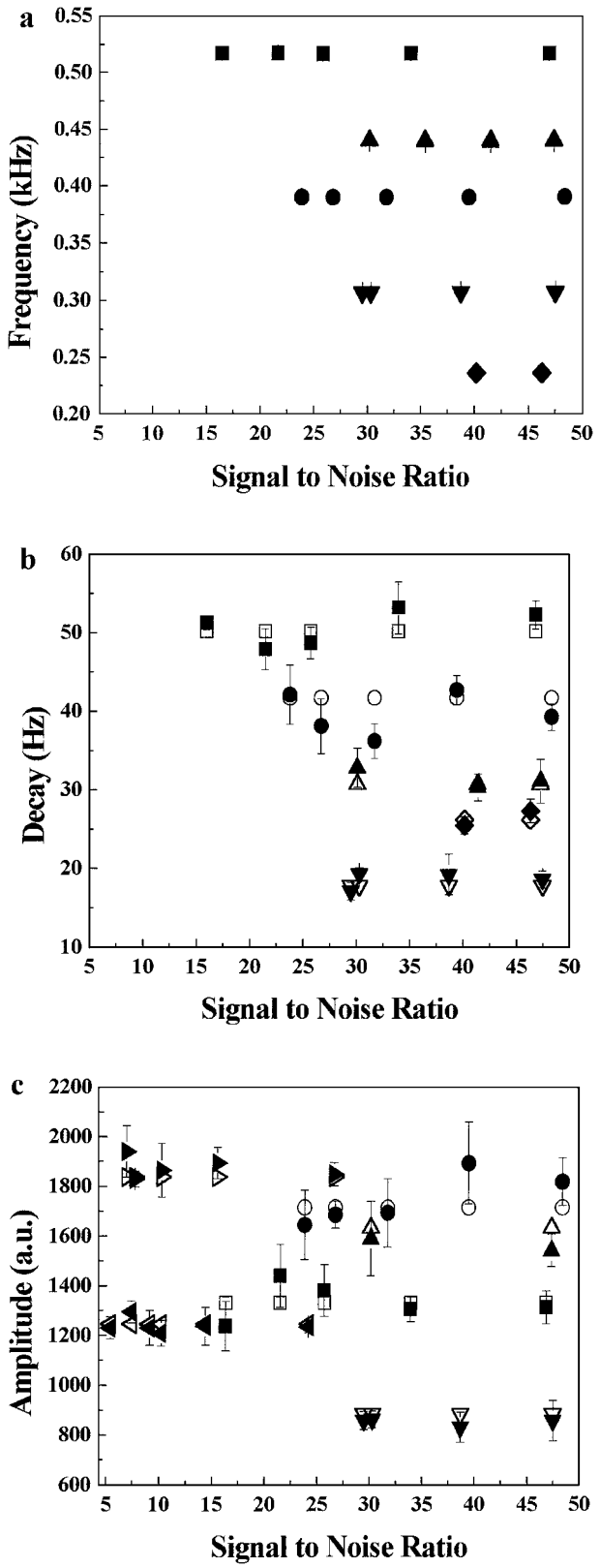

FIG. 3. Estimated (filled symbols) and true (empty symbols) spectral parameters of the seven chosen signals at different SNR. (a) Frequencies; (b) decays; (c) amplitudes. Each symbol identifies a signal as indicated in Fig. 4b.

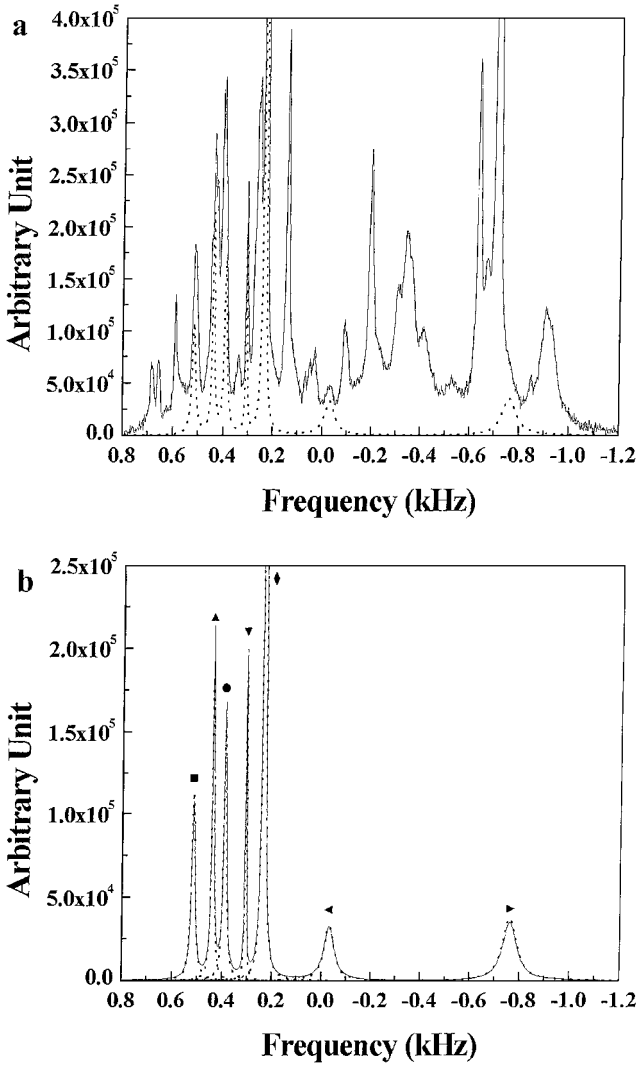

FIG. 4. (a) Simulated spectrum obtained from a simulated FID containing 58 complex exponentially decaying sinusoids ($SNR = 23.3$). Seven chosen signals fitted in the time domain by the MeFreS algorithm are also shown (dotted trace). (b) Simulated spectrum obtained from the seven chosen complex exponentially decaying sinusoids. Each line is labelled with a filled symbol. Also reported are fitted signals (dotted trace) in the time domain by the MeFreS algorithm.

bias. In both situations, MeFreS results unbiased. Finally, for signals which are expected to be well fitted by AMARESw and VARPRO (Freq. 516.6 Hz, Decay 50.2 Hz, Ampl. 1330 (black square); Freq. 439.0 Hz, Decay 30.7 Hz, Ampl. 1633 (black triangle); Freq. 306.7 Hz, Decay 17.8, Ampl. 884 (black triangle down)), all algorithms give comparable results. The same considerations are true for decays estimation (Fig. 6).

Spectra Quantification

In order to validate the MeFreS algorithm, two samples of known concentration were considered. In particular, the first sample (Sample A) contained (14 ± 1) μg of toluene

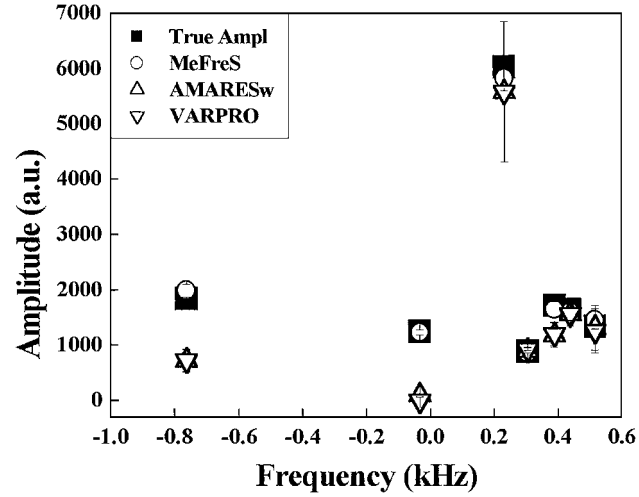

FIG. 5. True (filled symbols) and estimated (empty symbols) amplitudes of the seven chosen signals at SNR 14.0: (●) True amplitudes; (○) MeFreS estimated amplitudes; (△) AMARESw estimated amplitudes; (▽) VARPRO estimated amplitudes.

($MW = 92.14Da$) dissolved in (520 ± 2) μl of deuterated chloroform ($CDCl_3$, 99.96% deuterium content, $\rho = 1.48$ g/ml, $MW = 119.38Da$). The toluene concentration ($C_{toluene}$) was (0.3 ± 0.1) mmol/l, and $CDCl_3$ concentration (C_{CDCl_3}) was (12397.4 ± 182.2) mmol/l. Experimental spectra were obtained by using a Bruker DRX-500 spectrometer operating at 500 MHz. The FIDs were acquired with a 90 pulse and 64 transients of 2048 data points over ±3255.2 Hz. A reduced number of data points was used in order to avoid a complete decay of the FIDs to zero. Using the first 349 FID data points, the MeFreS algorithm was used to fit one by one the toluene (−1147.3 ± 0.1) Hz and $CDCl_3$ (1303.8 ± 0.2) Hz signals (Table 1).

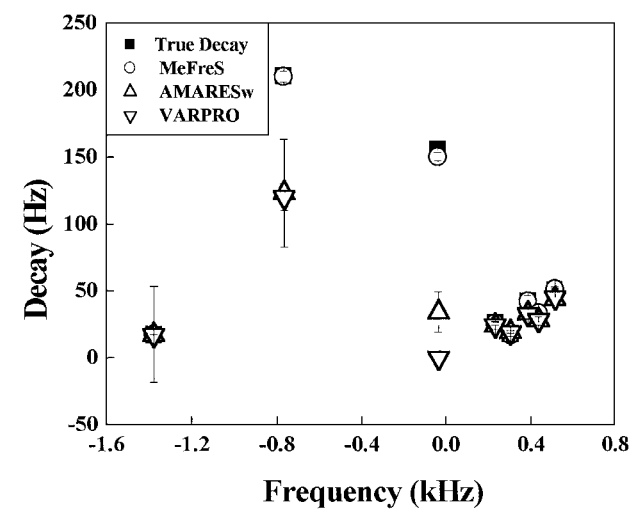

FIG. 6. True (filled symbols) and estimated (empty symbols) decays of the seven chosen signals at SNR 14.0. (●) True decays; (○) MeFreS estimated decays; (△) AMARESw estimated decays; (▽) VARPRO estimated decays.

TABLE 1
Sample A Fitted Signal Parameters

Frequency (Hz)	Amplitude (a.u.)	Decay (Hz)	Phase (rad)
-1147.3 ± 0.1	1246 ± 88	1.69 ± 0.06	-3.6 ± 0.1
1303.8 ± 0.2	6663 ± 431	1.24 ± 0.06	-1.03 ± 0.07

In Fig. 7, the FFT spectrum of Sample A is reported without any manipulation. Also reported is the FFT of the complex exponentially decaying sinusoids obtained with the MeFreS algorithm, one by one and without any manipulation (dashed and dotted traces). As it can be seen, fitted signals reproduce very well the experimental ones.

This is better appreciated in Figs. 8a and 8b that are expansions of the deuterated chloroform and toluene signals, respectively. Being the area under a Lorentzian line, obtained by FFT of a single complex exponentially decaying sinusoid, equal to one half of the amplitude, by using the results of Table 1 with a $CDCl_3$ reference concentration of (12397.4 ± 182.2) mmol/l, we obtained for toluene $C_{toluene} = (0.31 \pm 0.05)$ mmol/l, which well reproduces the experimental toluene concentration (0.3 ± 0.1) mmol/l. It is important to observe that such a result has been obtained by applying the MeFreS algorithm to a FID that has not fully decayed. In fact, in this case quantification in the frequency domain by using the usual integration routine is hampered by the presence of wiggles around each peak, which distort the baseline, thus causing a large error. In the Frequency Domain (FD), from peak integration we obtained a completely wrong concentration for toluene $FDC_{toluene} = (0.53 \pm 0.05)$ mmol/l. The second sample (Sample B) contained the tripeptide thyrotropin releasing hormone (TRH, Calbiochem, $MW = 363.4Da$), which

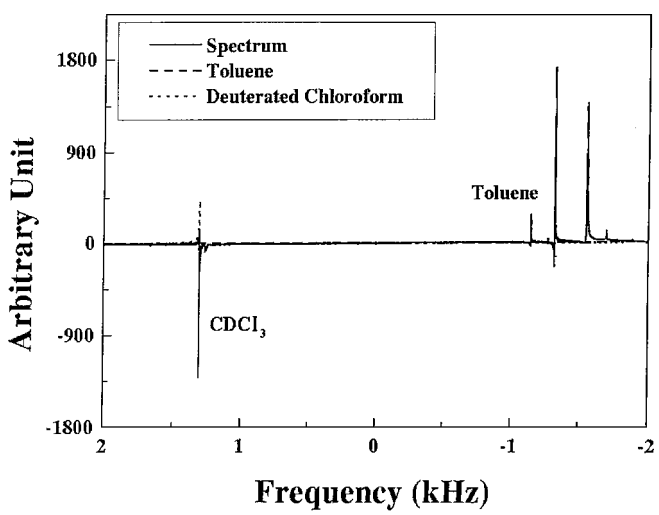

FIG. 7. Spectrum obtained by FFT of the Sample A. Neither phase adjustment nor baseline correction were applied ($SNR = 285$). Dashed and dotted traces refer to fitted toluene and deuterated chloroform respectively, as obtained with MeFreS algorithm.

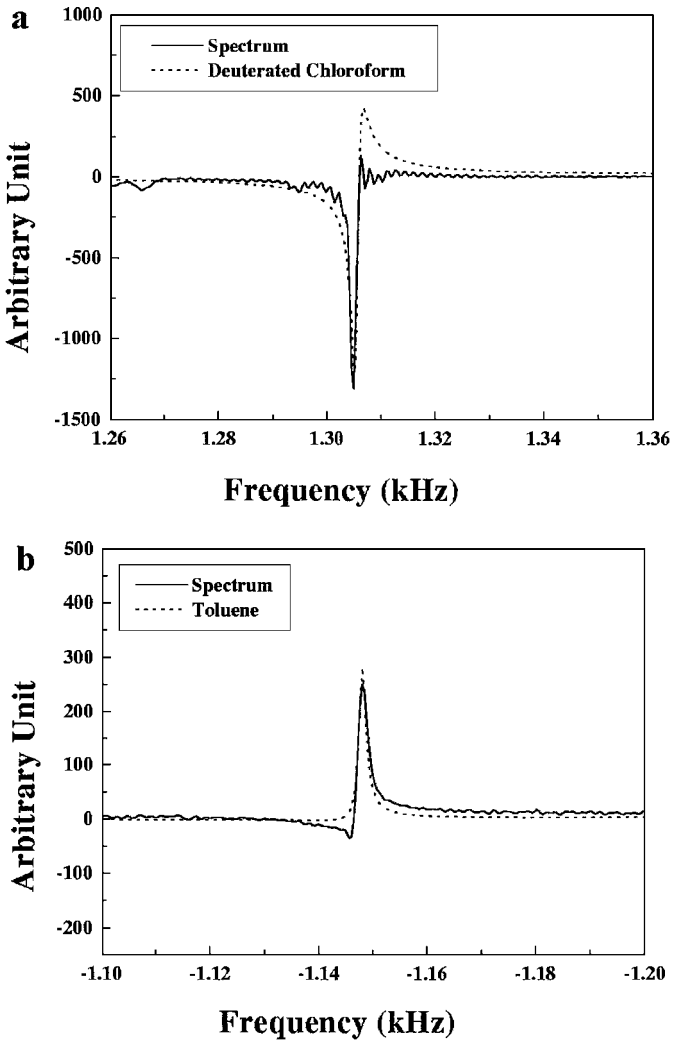

FIG. 8. (a) Zoom in around the deuterated chloroform of Fig. 7. In dot style, FFT of the fitted deuterated chloroform complex exponentially decaying sinusoid. (b) Zoom in around toluene of Fig. 7. In dot style, FFT of the fitted toluene complex exponentially decaying sinusoid.

controls the secretion of the thyroid stimulating hormone from the pituitary gland; 4.5 mg of TRH (pyrGlu-His-Pro) were dissolved in 400 μl of deuterated methanol (CD_3OD, 99.96%), and contained (10 ± 1) μl of a sodium trimethylsilyl[2,2,3,3-d4]propionate solution (TSP, 10 μmol/l), used as spectral reference. The proton spectrum of TRH was acquired on a Bruker DPX spectrometer operating at 300 MHz. The FIDs were

TABLE 2
Sample B Fitted Signal Parameters

Frequency (Hz)	Amplitude (a.u.)	Decay (Hz)	Phase (rad)
-1947.30 ± 0.04	1807 ± 80	0.93 ± 0.02	-1.2 ± 0.2
143.4 ± 0.2	19905 ± 974	3.2 ± 0.4	0.2 ± 0.1

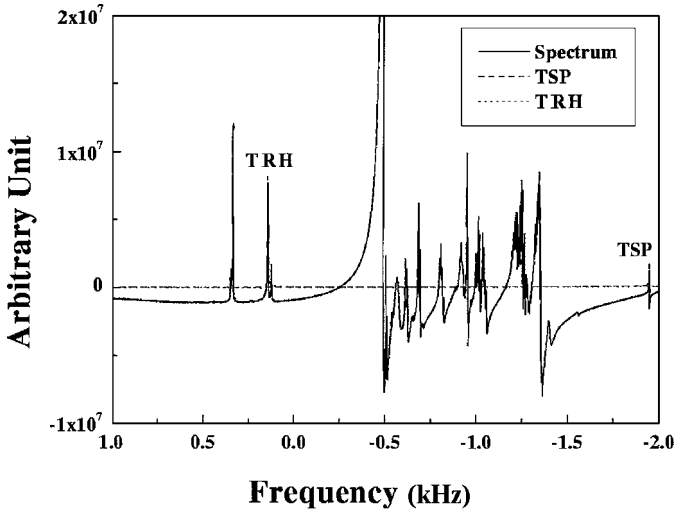

FIG. 9. Spectrum obtained by FFT of the FID of Sample B. Neither phase adjustment nor baseline correction was applied ($SNR = 336$).

accumulated with a 90 pulse and 64 scans of 8192 data points over a spectral window of ±2097.3 Hz. The frequency spectrum presents two isolated singlets in the aromatic region, corresponding to the His $C^{\omega 1}$ and the $C^{\delta 2}$ protons. Using the first 256 FID data points, the MeFreS algorithm was used to fit, one by one, TSP (-1947.30 ± 0.04) Hz and peptide His $C^{\delta 2}$ (143.4 ± 0.2) Hz signals (Table 2). Figure 9 reports the FFT spectrum of Sample B obtained without any manipulation, together with the FFT of the complex exponentially decaying sinusoids obtained by the MeFreS algorithm, one by one and without any manipulation (TRH, dotted line; TSP, dashed line). By using the results in Table 2, using a reference concentration for TSP of (0.24 ± 0.02) mmol/l, we obtained for TRH a concentration of (31.7 ± 5.6) mmol/l that well reproduces the experimental TRH concentration of (30.29 ± 0.02) mmol/l, while the concentration obtained from the integration in the frequency domain is (38.7 ± 3.6) mmol/l.

CONCLUSIONS

In this paper we have described MeFreS, a new algorithm for frequency-selective quantification of NMR parameters in the time domain. The results are very interesting as MeFreS avoids many of the drawbacks of selective analysis in the frequency domain. In particular, MeFreS works in the time domain, thus avoiding the usual manipulation of the FID (function multiplication, FFT, phase correction, baseline correction, etc.) for analysis in the frequency domain. The MeFreS algorithm does not use preprocessing steps and filter of nuisance peaks. It has the advantage of avoiding inherent distortions and estimation biases to be corrected. It correctly can fit a single component at a time, with a maximum of four parameters. It only uses rather weak prior knowledge assumptions obtainable at a glance from the FFT spectrum, or by a simple peak peaking. Furthermore, the prior knowledge of the chosen variation range to be assigned to each parameter can be as weak as 100%. Only a few minutes (from 3 to 7) for signal estimation were necessary on a 300 MHz PC. In order to verify its ability in avoiding the main drawbacks of other FS time-domain algorithms currently used, a comparison with AMARESw and VARPRO algorithms were made (MRUI, version 99.2 software package (http://www.mrui.uab.es/mrui/mrui_homePage.shtml)). Simulations confirm that MeFreS is able to correctly identify spectral parameters also in those cases when AMARESw and VARPRO are expected to fail. In particular, signals from high-resolution NMR simulated cell-type FID were successfully analyzed even in the presence of strong frequency-domain overlapping peaks, below the "allowed" frequency separation range. Large linewidth does not prevent MeFreS analysis, while AMARESw and VARPRO do not allow fitting of such lines, which are considered nuisance peaks (*8–11*). In order to fit a signal, MeFreS selects its single frequency and fits only that signal, while AMARESw and VARPRO need to fit all signals that fall in the range below the allowed frequency separation range. In addition, we have shown that MeFreS is able to determine the correct concentration of solutes in conditions very often encountered in NMR spectroscopy. It was able to find the correct toluene concentration (Sample A) in the presence of a truncated FID. It is well known that the FFT of a step function introduces wiggles in the frequency-domain spectrum, which strongly affect the signal intensity. This appears to be particularly important in fast acquisition (fast decomposing molecules or multidimensional NMR). Furthermore, the algorithm is able to correctly select the signal even in the presence of overlap, as in the case of His $C^{\delta 2}$ singlet (Sample B). Most importantly, The absence of filters and preprocessing steps makes the MeFreS algorithm naturally and directly extensible to multidimensional NMR experiments, on which we are currently working.

APPENDIX A

In this appendix the Hankel matrix trace properties (Properties 1, 2, and 3 from Results and Discussion) are demonstrated. By simple calculations, it can be shown that, for N odd, $M \leq (N+1)/2$,

$$\begin{aligned} &Tr(X(N-M+1, M) * X^{\dagger}(N-M+1, M)) \\ &\quad = \sum_{k=0}^{M-1} (k+1) * |x_k|^2 + \sum_{k=M}^{N-M} M * |x_k|^2 \\ &\quad + \sum_{k=N-M+1}^{N-1} (N-k) * |x_k|^2 \end{aligned} \qquad [A.1]$$

while, for $M \geq (N+1)/2$,

$$Tr(X(N-M+1,M)*X^{\dagger}(N-M+1,M)) = \sum_{k=0}^{N-M-1}(k+1)*|x_k|^2 + \sum_{k=N-M}^{M-1}(N-M+1) *|x_k|^2 + \sum_{k=M}^{N-1}(N-k)*|x_k|^2. \quad [A.2]$$

By Eq. [A.1], it follows that, for $M \le (N+1)/2$,

$$Tr(X(N-(M+1)+1,(M+1))*X^{\dagger}(N-(M+1)+1,(M+1))) - Tr(X(N-M+1,M)*X^{\dagger} \times (N-M+1,M)) = \sum_{k=M}^{N-M-1}|x_k|^2 \ge 0, \quad [A.3]$$

that is, for $M \le (N+1)/2$, $Tr(X(N-M+1,M)*X^{\dagger}(N-M+1,M))$ is an M increasing function (Property 1). In the same way, By Eq. [A.2], it follows that, for $M \ge (N+1)/2$,

$$Tr(X(N-M+1,M)*X^{\dagger}(N-M+1,M)) - Tr(X(N-(M+1)+1,(M+1)) *X^{\dagger}(N-(M+1)+1,(M+1))) = \sum_{k=N-M}^{M-1}|x_k|^2 \ge 0, \quad [A.4]$$

that is, for $M \ge (N+1)/2$, $Tr(X(N-M+1,M)*X^{\dagger}(N-M+1,M))$ is an M decreasing function (Property 2). Property 3 descends immediately, by maximum definition, from Properties 1 and 2.

REFERENCES

1. A. Knijn, R. De Beer, and D. van Ormondt, Frequency-selective quantification in the time domain, *J. Magn. Reson.* **97,** 444–450 (1992).
2. G. C. McKinnon, C. Burger, and P. Boesiger, Spectral baseline correction using CLEAN, *Magn. Reson. Med.* **13,** 145–149 (1990).
3. I. Dologlou, S. Van Huffel, and D. Van Ormondt, Metabolite-selective MRS data quantification with frequency prior knowledge, *J. Magn. Reson.* **130,** 238–243 (1998).
4. J. Tang and J. Norris, LP-ZOOM, a linear prediction method for local spectral analysis of NMR signals, *J. Magn. Reson.* **79,** 190–196 (1998).
5. L. Vanhamme, T. Sundin, P. Van Hecke, S. Van Huffel, and R. Pintelon, Frequency-selective quantification of biomedical magnetic resonance spectroscopy data, *J. Magn. Reson.* **143,** 1–16 (2000).
6. T. Sundin, L. Vanhamme, P. Van Hecke, I. Dologlou, and S. Van Huffel, Accurate quantification of ^1H spectra: From finite impulse response filter design for solvent suppression to parameter estimation, *J. Magn. Reson.* **139,** 189–204 (1999).
7. S. Cavassila, B. Fenet, A. van den Boogaart, C. Remy, A. Briguet, and D. Graveron-Demilly, ER-Filter: a preprocessing technique for frequency-selective time-domain analysis, *J. Magn. Reson. Anal.* **3,** 87–92 (1997).
8. A. van den Boogaart, D. van Ormondt, W. Pijnappel, R. de Beer, and M. Ala-Korpela, Removal of the water resonance from ^1H magnetic resonance spectra, *J. Magn. Reson.* **1,** 175–195 (1991).
9. K. Cross, Improved digital filtering technique for solvent suppression, *J. Magn. Reson.* **101,** 220–224 (1993).
10. G. Sobering, M. Kienlin, C. Moonen, P. van Zijl, and A. Bizzi, Post-acquisition reduction of water signals in proton spectroscopic imaging of the brain, *Proc. SMRM*, 171 (1991).
11. M. Deriche and X. Hu, Elimination of water signal by postprocessing, *J. Magn. Reson.* **101,** 229–232 (1992).
12. F. R. Gantmacher, "The Theory of Matrices," AMS Chelsea Publishing (1959).
13. G. H. Golub and C. F. Van Loan, "Matrix Computations," 3rd ed., John Hopkins University Press, Baltimore (1966).
14. C. L. Lawson and R. G. Hanson, "Solving Least Squares Problems," Prentice-Hall, Englewood Cliffs, NJ (1974).
15. G. E. Forsythe, M. A. Malcolm, and C. B. Moler, "Computer Methods for Mathematical Computations," Prentice-Hall, Englewood Cliffs, NJ (1977).
16. V. C. Klema and A. J. Laub, The singular value decomposition: Its computation and some applications, *IEEE Trans. Auto. Control.* **25,** 164–176 (1980).
17. J. Cadzow, Signal enhancement: A composite property mapping algorithm, *IEEE Trans. Acoustics Speech Signal Process.* **36,** 49–62 (1988).
18. M. K. Maple, "Digital Spectral Analysis with Applications," Prentice-Hall, Englewood Cliffs, NJ (1966).
19. R. Romano, M. T. Santini, and P. L. Indovina, A time-domain algorithm for NMR spectral normalization, *J. Magn. Reson.* **146,** 89–99 (2000).
20. J. A. Nelder and R. Mead, *Computer J.* **7,** 308–313 (1965).
21. M. S. Caceci and W. P. Cacheris, Fitting curves to data: The Simplex algorithm is the answer, *BYTE,* May, 340–362 (1984).
22. S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, *Science* **220,** 671–680 (1983).
23. S. Kirkpatrick, *J. Stat. Phys.* **34,** 975–986 (1984).
24. N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, *J. Chem. Phys.* **21,** 1087–1092 (1953).
25. W. H. Press, S. A. Teukolsky, W. T. Wetterling, and B. P. Flannery, "Numerical Recipes in C: the Art of Scientific Computing," 2nd ed., pp. 451–455, Cambridge Univ. Press, Cambridge, UK (1986).
26. A. van den Boogaart, M. Ala-Korpela, J. Jokisaari, and J. R. Griffiths, Time and frequency domain analysis of NMR data compared: An application to $1D^1H$ spectra of lipoproteins, *Magn. Reson. Med.* **31,** 347–358 (1994).
27. H. Chen, S. van Huffel, D. van Ormondt, and R. de Beer, Parameter estimation with prior knowledge of known signal poles for the quantification of NMR spectroscopy data in the time domain, *J. Magn. Reson. A* **119,** 225–234 (1996).
28. H. Chen, S. van Huffel, C. Decanniere, and P. van Hecke, A signal-enhancement algorithm for the quantification of NMR data in the time domain, *J. Magn. Reson. Series A* **109,** 46–55 (1994).
29. M. Lupu and D.Todor, Linear prediction and singular value decomposition in NMR signal analysis, *in* "Signal Treatment and Signal Analysis in NMR" (D. N. Rutledge Ed.), pp. 169–171, Elsevier, Amsterdam (1996).
30. J. C. Hoch and A. S. Stern, Using the DFT: Application to NMR, *in* "NMR Data Processing," pp. 40–50, Wiley–Liss, New York (1999).