JOURNAL OF MAGNETIC RESONANCE **128,** 30–41 (1997)
ARTICLE NO. MN971215

A Novel Detection–Estimation Scheme for Noisy NMR Signals: Applications to Delayed Acquisition Data

Yung-Ya Lin, Paul Hodgkinson,[1] Matthias Ernst,[2] and Alexander Pines[3]

Department of Chemistry, University of California, and Materials Sciences Division, Lawrence Berkeley Laboratory, Berkeley, California 94720

Received February 7, 1997; revised June 5, 1997

Many potentially interesting and useful classes of NMR experiments generate data for which conventional spectral estimation and quantification via the Fourier transform are unsatisfactory. In particular, recently introduced solid-state NMR experiments which involve long delays before data acquisition fall into this category, as the free induction decays are heavily ''truncated'' and have low signal-to-noise ratios. A novel detection–estimation scheme is introduced in order to analyze data from such experiments and others where the sensitivity is low and/or the data record is strongly damped or truncated. Based on the assumption of exponential data modeling, the number of signals present is first detected using criteria derived from information theory and the spectral parameters are then estimated using the matrix pencil method. Monte Carlo simulations and experimental applications are carried out to demonstrate its superior statistical and computational performances and its general applicability to delayed acquisition data. Over the range of noise levels investigated, it is found that this approach is essentially near-optimal in the sense that the estimated spectral parameters have biases almost equal to zero and variances very close to their theoretical Cramér–Rao lower bounds. Compared to the popular method of linear prediction with singular value decomposition, this method not only improves the estimation accuracy (by a factor of 2–4) with a lower ''breakdown'' signal-to-noise threshold (≈1.5 dB), but also reduces the computational cost by about an order of magnitude. It also holds great promise in effectively reducing truncation artifacts. It is concluded that this approach not only facilitates the analysis of delayed acquisition data, but can also become a valuable tool in the routine quantification of general NMR spectra. A listing of programs is also included in the Appendix.

INTRODUCTION

Conventional spectral estimation of NMR data is based on the Fourier transform (FT), which decomposes the time series into a sum of undamped sinusoidal oscillations. This can be done very efficiently using the fast Fourier transform. However, the Fourier transform is only strictly applicable to the limited subset of ''complete'' signals, i.e., $t = 0$ to ∞ (*1–8*). Fourier transformation of signals that are truncated, at either the start or the end of the decay, leads to familiar spectral distortions, baseline roll, and ''sinc-wiggles,'' respectively. Besides, there is no built-in mechanism for noise suppression. The linear nature of the Fourier transform implies that reducing these problems or improving apparent resolution can only be done at the expense of spectral resolution and/or sensitivity.

The goal of NMR spectral estimation is to obtain an estimate of the frequency response function of the underlying spin system from the measured free induction decay (FID). A particular FID can be characterized in terms of a model function with a set of free parameters. A crucial problem in NMR spectral estimation is, therefore, the detection of the signal model and the estimation of the spectral parameters (e.g., damping factor, frequency, amplitude, and phase). The difficulty of the detection–estimation problem is increased by the low sensitivity inherent in NMR spectroscopy. Because of the computational complexity and noise interference, the problem is usually solved in two steps. The model function is first chosen and verified on physical grounds or by statistical tests. After successful signal modeling, the free parameters of the signals are then estimated.

Detection theory refers to the selection of the physical or mathematical model that best describes the measured phenomena. The model function must be chosen with care; if the number of parameters is too large, many of them will be spurious, particularly if one must contend with noise, while too restrictive a model function leads to poor fitting of the data and systematic errors. Estimate statistics are usually better if the number of parameters is minimized. In NMR, the experimentally observed FID, $\boldsymbol{y} = [y_0, y_1, \ldots, y_{N-1}]^{\mathrm{T}}$, can be approximated by a sum of complex-valued noise-free signal $\boldsymbol{x} = [x_0, x_1, \ldots, x_{N-1}]^{\mathrm{T}}$ and additive noise perturbation $\boldsymbol{w} = [w_0, w_1, \ldots, w_{N-1}]^{\mathrm{T}}$, where ''T'' denotes matrix transpose and N is the number of complex data points. It is generally assumed that the elements of $\boldsymbol{w}$ are complex Gaussian random variables with zero mean, variance ρ, and uncorrelated real and imaginary parts. This *a priori* assump-

[1] Current address: Laboratoire de Stéréochemie et des Interactions Moléculaire, Ecole Normale Supérieure de Lyon, 69364 Lyon 07, France.

[2] Current address: Laboratory for Physical Chemistry, University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands.

[3] To whom correspondence should be addressed.

1090-7807/97 $25.00

tion of normality is not only mathematically convenient, but via the central limit theorem, it is often a good approximation of the real NMR circumstances.

The quality of spectral estimation can be improved by incorporating further information into the signal model. This is conventionally done by assuming that the signal can be decomposed into a set of exponentially damped oscillations,

$$\begin{aligned} y_n = x_n + w_n &= \sum_{i=1}^{M} |a_i| \exp(j\theta_i) \\ &\quad \times \exp[(-\alpha_i + j2\pi f_i)n] + w_n \\ &= \sum_{i=1}^{M} a_i z_i^n + w_n, \quad n = 0, 1, \ldots, N-1, \end{aligned} \tag{1}$$

where $|a_i|$, α_i, f_i, and θ_i represent the absolute amplitudes, damping factors (inverse time constants), frequencies, and phases of the M distinct exponentials, respectively; j is used to denote $\sqrt{-1}$. $z_i \equiv \exp(-\alpha_i + j2\pi f_i)$ is the ''signal pole'' and $a_i \equiv |a_i| \exp(j\theta_i)$ is the ''complex amplitude.'' This is in general a good assumption for liquid-state NMR, and for solid-state NMR with fast magic angle sample spinning (*9*). Based on this model function, the detection problem is then reduced to the determination of the number of signals M. Note that the assumption of exponential decay is not necessarily unduly restrictive. The application of such a model to nonexponentially decaying signals results in a mathematical, rather than a physical, analysis by expanding each spectral component into a sum of exponentials.

Within the past decade, much research activity has been focused on formulating and comparing alternative means of NMR spectral estimation (*2–8, 10*), driven by the promise of potentially superior spectral sensitivity and/or resolution in comparison to the conventional Fourier transform (albeit at the expense of greater computational complexity). In particular, parametric methods based on the exponential FID modeling, cf. Eq. [1], have attracted considerable interest. This incorporation of lineshape information should allow individual signals to be better resolved, both from each other and from the noise. In this paper, a combined detection–estimation scheme, ITMPM, based on information theory and the matrix pencil method, is introduced which improves the estimation performance and computational efficiency of the exponential FID modeling, relative to existing techniques based on linear prediction. Monte Carlo simulations are first carried out to verify the statistical superiority of ITMPM. Its applications to the experimental data from delayed acquisition measurements are then demonstrated. A fully automated program in MATLAB (*11*) is provided in the Appendix to minimize the user's implementation effort.

ESTIMATION BY THE MATRIX PENCIL METHOD

There is an important difference between the complex amplitude a_i and the signal pole z_i in the functional form of Eq. [1]; the signal model is linear in the first and nonlinear in the second. The complications of nonlinearity can be circumvented by invoking the linear prediction principle [e.g., linear prediction (*12, 13*), autoregressive modeling (*14, 15*)] or by employing matrix factorization techniques [e.g., state space formalism (*16, 17*), matrix pencil method (*18, 19, 21*)]. In particular, linear prediction with singular value decomposition (LPSVD) and related methods have been shown to be useful complements to the Fourier transform (*2–8, 12, 13, 22–25*). The principles of LPSVD have been extensively documented. In summary, one first solves the following linear prediction equations for the coefficients $\{c_i\}_{1 \leqslant i \leqslant L}$ of the prediction polynomial,

$$\begin{bmatrix} y_{L-1} & y_{L-2} & \cdots & y_0 \\ y_L & y_{L-1} & \cdots & y_1 \\ \vdots & \vdots & \ddots & \vdots \\ y_{N-2} & y_{N-3} & \cdots & y_{N-L-1} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_L \end{bmatrix} = \begin{bmatrix} y_L \\ y_{L+1} \\ \vdots \\ y_{N-1} \end{bmatrix}. \tag{2}$$

The roots of the prediction polynomial $P(z) = 1 - \Sigma_{i=1,L} c_i z^{-i}$ then produce the M signal poles $\{z_i\}_{1 \leqslant i \leqslant M}$. The success of LPSVD resides in the ability of linear prediction to extend truncated FIDs for resolution enhancement, or to estimate missing or corrupted initial data points for baseline and phase corrections (*26*). The corrupting effects of noise can be mitigated through a judicious combination of an overestimated prediction order ($L \gg M$, introducing extra $L - M$ noise-related poles) to account for the noise in the measurements, and SVD-based signal-subspace techniques to discriminate between signal and noise (*10*).

While the use of LPSVD for spectral quantification has many advantages, its major drawback is the considerably larger computational burden and higher algorithmic complexity (necessary to avoid numerical instability and overflow) mainly due to the high degree polynomial rooting. In addition, LPSVD exhibits a ''breakdown'' in estimation performance when the noise level increases beyond a certain threshold, e.g., Refs. (*2–8, 10, 12, 13, 27*). These factors make LPSVD less satisfactory, particularly when the signal-to-noise ratio (SNR) is low.

Unlike LPSVD, the recently proposed matrix pencil method, developed independently by Hua and Sarkar (*18, 19*) and by Kailath and co-workers (*21*), involves finding the signal poles, z_i, directly by solving a generalized eigenvalue problem. Its formulation is governed by the notion of pencil-of-functions and exploits the property of a matrix pencil constructed from the underlying FID. The mathematical entity ''matrix pencil'' refers to the linear combination of two matrices (say, $\boldsymbol{F}$ and $\boldsymbol{G}$) defined on a common domain, i.e., $\boldsymbol{F} + \lambda \boldsymbol{G}$. The eigenvalues of the matrix pencil are defined as the values of the scalar variable λ that decrease the rank of the matrix pencil.

Let $\boldsymbol{X}_0$ and $\boldsymbol{X}_1$ be two noise-free data matrices with dimension $(N - L) \times L$ defined by

$$\boldsymbol{X}_0 = \begin{bmatrix} x_{L-1} & x_{L-2} & \cdots & x_0 \\ x_L & x_{L-1} & \cdots & x_1 \\ \vdots & \vdots & \ddots & \vdots \\ x_{N-2} & x_{N-3} & \cdots & x_{N-L-1} \end{bmatrix},$$

$$\boldsymbol{X}_1 = \begin{bmatrix} x_L & x_{L-1} & \cdots & x_1 \\ x_{L+1} & x_L & \cdots & x_2 \\ \vdots & \vdots & \ddots & \vdots \\ x_{N-1} & x_{N-2} & \cdots & x_{N-L} \end{bmatrix}, \quad [3]$$

where L is called the pencil parameter. It follows from Eq. [1] that these matrices can be decomposed as

$$\boldsymbol{X}_0 = \boldsymbol{Z}_L \boldsymbol{B} \boldsymbol{Z}_R, \quad \boldsymbol{X}_1 = \boldsymbol{Z}_L \boldsymbol{B} \boldsymbol{Z} \boldsymbol{Z}_R, \quad [4]$$

where

$$\boldsymbol{Z}_L = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_M \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{N-L-1} & z_2^{N-L-1} & \cdots & z_M^{N-L-1} \end{bmatrix},$$

$$\boldsymbol{B} = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_M \end{bmatrix},$$

$$\boldsymbol{Z}_R = \begin{bmatrix} z_1^{L-1} & z_1^{L-2} & \cdots & 1 \\ z_2^{L-1} & z_2^{L-2} & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ z_M^{L-1} & z_M^{L-2} & \cdots & 1 \end{bmatrix},$$

$$\boldsymbol{Z} = \begin{bmatrix} z_1 & 0 & \cdots & 0 \\ 0 & z_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & z_M \end{bmatrix}, \quad [5]$$

$\boldsymbol{Z}_L$ and $\boldsymbol{Z}_R$ are Vandermonde matrices, and $\boldsymbol{B}$ and $\boldsymbol{Z}$ are diagonal matrices constructed from the complex amplitudes and signal poles, respectively.

Now consider the matrix pencil $\boldsymbol{X}_1 - \lambda \boldsymbol{X}_0$,

$$\begin{aligned} \boldsymbol{X}_1 - \lambda \boldsymbol{X}_0 &= \boldsymbol{Z}_L \boldsymbol{B} (\boldsymbol{Z} - \lambda \boldsymbol{I}_M) \boldsymbol{Z}_R \\ &= \boldsymbol{Z}_L \boldsymbol{B} \begin{bmatrix} z_1 - \lambda & 0 & \cdots & 0 \\ 0 & z_2 - \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & z_M - \lambda \end{bmatrix} \boldsymbol{Z}_R, \end{aligned} \quad [6]$$

where $\boldsymbol{I}_M$ is an $M \times M$ identity matrix. In general, the rank of the matrix pencil $\boldsymbol{X}_1 - \lambda \boldsymbol{X}_0$ is M. However, if $\lambda = z_i$ then the rank of $\boldsymbol{Z} - \lambda \boldsymbol{I}_M$ will be reduced to $M - 1$. In other words, each of the z_i will be a rank reducing number of the matrix pencil $\boldsymbol{X}_1 - \lambda \boldsymbol{X}_0$, and so, by definition, the set $\{z_i\}_{1 \leqslant i \leqslant M}$ can be identified with the M nonzero generalized eigenvalues of the matrix pair $(\boldsymbol{X}_1, \boldsymbol{X}_0)$,

$$\boldsymbol{X}_1 \boldsymbol{q}_i = z_i \boldsymbol{X}_0 \boldsymbol{q}_i, \quad [7]$$

where $\boldsymbol{q}_i$ is the eigenvector associated with the eigenvalue (and signal pole) z_i. The Moore–Penrose pseudo-inverse of a matrix is a generalization of the matrix inverse to the case where the matrix is not square and possibly of incomplete rank. Left multiplying Eq. [7] by the Moore–Penrose pseudo-inverse of $\boldsymbol{X}_0$, $\boldsymbol{X}_0^{\#}$, and using the property that $\boldsymbol{X}_0^{\#} \boldsymbol{X}_0 = \mathrm{I}$, it is clear that the generalized eigenvalues of $(\boldsymbol{X}_1, \boldsymbol{X}_0)$ can be solved by finding the M nonzero eigenvalues of the $L \times L$ matrix product $\boldsymbol{X}_0^{\#} \boldsymbol{X}_1$,

$$\boldsymbol{X}_0^{\#} \boldsymbol{X}_1 \boldsymbol{q}_i = z_i \boldsymbol{q}_i. \quad [8]$$

In order to apply such techniques to experimental data, it is important to account for the effects of noise corruption. In spectral estimation, the data matrix, covariance matrix, and autocorrelation matrix characterize the information contained in the observed signal. When constructed from noiseless signal, these matrices possess certain eigen-characteristics (e.g., rank, degeneracy, positive semidefinite) and/or matrix structures (e.g., Hermitian, Hankel, Toeplitz). Inevitable measurement noise, however, results in the loss of part (or all) of these theoretical matrix properties. Provided that the noise perturbation is not too large, a useful procedure for approximating the noiseless matrix is to find the matrix which possesses a set of chosen properties and lies closest to the empirical noisy matrix. The resulting matrix is a more accurate representation of the characteristics of the underlying signal than the original noisy matrix, where the corrupting effect of noise has been mitigated and a signal enhancement has been accomplished.

This can be successfully incorporated into the matrix pencil method. For the noisy data, we define $\boldsymbol{Y}_0$ and $\boldsymbol{Y}_1$ the same way as for $\boldsymbol{X}_0$ and $\boldsymbol{X}_1$ with x_i replaced by the noisy measurement y_i. It is clear that, while the noiseless data matrices $\boldsymbol{X}_0$ and $\boldsymbol{X}_1$ have a rank equal to the number of signal components (M), the noisy data matrices, $\boldsymbol{Y}_0$ and $\boldsymbol{Y}_1$, will in general be of full rank due to noise contamination. Singular value decomposition (SVD) provides a particularly useful tool for restoring the matrix-rank property. SVD, one of the most stable and computationally effective algorithms in the theory of matrix algebra, is a generalization of the eigenvalue decomposition for nonsquare matrices (*10, 14, 15, 28*). The SVD theorem states that for an arbitrary $(N - L) \times L$ matrix $\boldsymbol{Y}_0$ there exist positive real numbers $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_R > 0$ (the so-called singular values), an $(N - L) \times (N - L)$ unitary matrix $\boldsymbol{U} = [\boldsymbol{u}_1 \boldsymbol{u}_2 \cdots \boldsymbol{u}_{N-L}]$, and an $L \times L$ unitary matrix $\boldsymbol{V} = [\boldsymbol{v}_1 \boldsymbol{v}_2 \cdots \boldsymbol{v}_L]$ such that

$$\begin{aligned} \boldsymbol{Y}_0 &= \boldsymbol{U} \begin{bmatrix} \boldsymbol{\Sigma} \\ \boldsymbol{0} \end{bmatrix} \boldsymbol{V}^{\dagger} = \sum_{i=1}^{R} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{\dagger} \\ &= \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^{\dagger} + \sigma_2 \boldsymbol{u}_2 \boldsymbol{v}_2^{\dagger} + \cdots + \sigma_R \boldsymbol{u}_R \boldsymbol{v}_R^{\dagger}, \end{aligned} \quad [9]$$

where $\mathbf{\Sigma} = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_R)$ is an $R \times R$ diagonal matrix, $\mathbf{0}$ is a null matrix, ''†'' denotes conjugate transpose, and $R \leqslant \min(N - L, L)$ is defined as the rank of $\boldsymbol{Y}_0$. The matrix $\boldsymbol{Y}_0$ is thus constructed from R rank-one matrices weighted by their respective singular values. The matrix of rank M ($<R$) which lies closest, in a least-squares sense, to the original matrix is constructed using the first M principle singular values and the associated singular vectors [Eckart–Young theorem (*29*)], i.e.,

$$\hat{\boldsymbol{Y}}_0 = \boldsymbol{U}_M \mathbf{\Sigma}_M \boldsymbol{V}_M^{\dagger} = \sum_{i=1}^{M} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{\dagger}, \qquad [10]$$

where $\boldsymbol{U}_M = [\boldsymbol{u}_1 \boldsymbol{u}_2 \cdots \boldsymbol{u}_M]$, $\boldsymbol{V}_M = [\boldsymbol{v}_1 \boldsymbol{v}_2 \cdots \boldsymbol{u}_M]$, and $\mathbf{\Sigma}_M = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_M)$.

The Moore–Penrose pseudo-inverse can also be defined in terms of the SVD components of $\boldsymbol{Y}_0$,

$$\hat{\boldsymbol{Y}}_0^{\#} = \boldsymbol{V}_M \mathbf{\Sigma}_M^{-1} \boldsymbol{U}_M^{\dagger} = \sum_{i=1}^{M} \sigma_i^{-1} \boldsymbol{v}_i \boldsymbol{u}_i^{\dagger}. \qquad [11]$$

Hence, for noisy data, the M nonzero eigenvalues of the signal-enhanced $L \times L$ matrix product $\hat{\boldsymbol{Y}}_0^{\#} \boldsymbol{Y}_1$ give the estimates of the signal poles z_i, and hence the damping factors $\alpha_i = -\log|z_i|$, and frequencies $f_i = \arg(z_i)/2\pi$. It should be noted that further replacement of $\boldsymbol{Y}_1$ by its rank M counterpart $\hat{\boldsymbol{Y}}_1$ offers little advantage, due to the strong correlation of the noise between $\boldsymbol{Y}_0$ and $\boldsymbol{Y}_1$. Once $\{z_i\}_{1 \leqslant i \leqslant M}$ are known, the absolute amplitudes $|a_i|$ and phases $\theta_i = \arg(a_i)$ can be solved by a general linear least-squares analysis

$$\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{N-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_M \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{N-1} & z_2^{N-1} & \cdots & z_M^{N-1} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix} + \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{N-1} \end{bmatrix}. \qquad [12]$$

Note that the pencil parameter L (where $M < L < N$) will influence the quality of the results, in a similar way that LPSVD is affected by the choice of the polynomial degree (prediction order). A poor choice of prediction order or pencil parameter will limit the performance of either technique. It is empirically found that the optimal value for L ranges from $L = N/3$ for noisy signals to $L = N/2$ for signals with a higher SNR (*18, 19*).

The computational efficiency can be increased by noting that since $\hat{\boldsymbol{Y}}_0^{\#} \boldsymbol{Y}_1$ has rank $M < L$, $L - M$ of its eigenvalues are zero (*19*). The size of $\hat{\boldsymbol{Y}}_0^{\#} \boldsymbol{Y}_1$ can therefore be reduced before its eigenvalues are found. Substituting for $\hat{\boldsymbol{Y}}_0^{\#}$ from Eq. [11] into the eigenvalue equation, Eq. [8], for $\hat{\boldsymbol{Y}}_0^{\#} \boldsymbol{Y}_1$,

$$\boldsymbol{V}_M \mathbf{\Sigma}_M^{-1} \boldsymbol{U}_M^{\dagger} \boldsymbol{Y}_1 \boldsymbol{q}_i = z_i \boldsymbol{q}_i. \qquad [13]$$

Left multiplying by $\boldsymbol{V}_M^{\dagger}$ and using the unitary nature of $\boldsymbol{V}_M$, i.e., $\boldsymbol{V}_M \boldsymbol{V}_M^{\dagger} = \boldsymbol{V}_M^{\dagger} \boldsymbol{V}_M = \mathrm{I}$, we have

$$\mathbf{\Sigma}_M^{-1} \boldsymbol{U}_M^{\dagger} \boldsymbol{Y}_1 \boldsymbol{V}_M (\boldsymbol{V}_M^{\dagger} \boldsymbol{q}_i) = z_i (\boldsymbol{V}_M^{\dagger} \boldsymbol{q}_i). \qquad [14]$$

Now it can be seen that the estimates of $\{z_i\}_{1 \leqslant i \leqslant M}$ can be found by computing the eigenvalues of the much smaller $M \times M$ matrix $\mathbf{\Sigma}_M^{-1} \boldsymbol{U}_M^{\dagger} \boldsymbol{Y}_1 \boldsymbol{V}_M$.

DETECTION BASED ON INFORMATION THEORY

A potential difficulty encountered in practical applications of the matrix pencil and other SVD-based methods is the ambiguity in detecting the number of signal components, that is, in choosing the matrix rank M (*30*). The determination of M is crucial; too small a value of M results in information loss, while too large a value effectively incorporates more noise and generates spurious spectral features. For well-resolved spectra with reasonable SNR, the value of M can be determined from the sharp cutoff in magnitude of the singular values, or the number of resolved peaks in the FT spectrum above a predefined threshold. These criteria, however, become ill defined as the spacing between resonance frequencies, the number of data points, or the SNR decreases.

Various criteria have been proposed to address this problem. They may be classified into the following five categories: (i) subjective threshold settings of singular values (*28, 31, 32*), e.g., $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_M > \delta \geqslant \sigma_{M+1}$, where the threshold value δ is selected on an *ad hoc* basis; (ii) statistical threshold bounds on singular values (*30*), which is similar to (i) but the threshold is based on the theories of perturbations of singular values and statistical significance test; (iii) hypothesis test of likelihood ratios (*33, 34*), for each hypothesis the likelihood ratio statistic is compared to a subjective threshold level; (iv) matching of reconstruction residue with noise power (*35, 36*), consecutive reconstructions for various ranks are performed and the resulting error power is compared to the noise power; (v) information theory for model order selection (*37*). All these criteria appear intuitively reasonable and function effectively over various cases. Moreover, the information theoretic criteria, (v), have certain theoretical and computational advantages over the others. These criteria were originally introduced in the context of linear prediction by Akaike [AIC, Akaike information criterion (*38*)] and by Schwartz and Rissanen [MDL, minimum description length (*39, 40*)], and later adapted to exponential modeling by Wax and Kailath (*37*). Unlike the conventional approaches in categories (i) and (iii), AIC and MDL do not require any subjective threshold settings. Statistically determined criteria generally perform better than those empirically set by the user, as well as minimizing the necessity for user involvement. The principle and the derivation of these criteria can be found elsewhere (*14, 15*), but in summary, the optimal value of M is determined merely

by minimizing a discrete function of the singular values (which are already known from Eq. [9]), e.g.,

$$\mathrm{MDL}(k) = -\log\left[\frac{\left(\prod_{i=k+1}^{L}\sigma_i\right)^{N}}{\left(\frac{1}{L-k}\prod_{i=k+1}^{L}\sigma_i\right)^{(L-k)N}}\right] + \frac{1}{2}k(2L-k). \quad [15]$$

This is very different from the approaches in category (iv) where *a priori* knowledge of the noise power and extra computational cost for consecutive reconstructions are definitely required. In this work, MDL is used in preference to AIC as it has been shown to give consistent estimates of M, while AIC tends to overestimate the number of signals as the number of data points increases (*37*).

MONTE CARLO SIMULATIONS

The development of any signal processing protocol requires a rigorous statistical evaluation of its performance. Visual comparison of the results of a few realizations is insufficient to draw any general conclusions (*14, 15*). The bias, variance, and mean relative error (MRE) are statistical measures that are commonly used to quantify the performance,

$$\mathrm{bias}(\hat{\beta}) \overset{\text{def}}{=} \varepsilon(\hat{\beta}) - \beta \quad [16]$$

$$\mathrm{var}(\hat{\beta}) \overset{\text{def}}{=} \varepsilon\{[\hat{\beta} - \epsilon(\hat{\beta})]^2\} \quad [17]$$

$$\mathrm{MRE}(\hat{\beta}) \overset{\text{def}}{=} \varepsilon\left(\left|\frac{\hat{\beta}-\beta}{\beta}\right|\right) \times 100\%, \quad [18]$$

where ε is the expectation operator and $\hat{\beta}$ is the estimate of the spectral parameter whose true value is β. Ideally the bias should be negligibly small and the variance as close as possible to its theoretical lower limit, the Cramér–Rao lower bound (see below).

Analytical determination of these functions is generally intractable. Instead they can be estimated using Monte Carlo simulations in which the statistics are calculated using the same noise-free data set, but many different realizations of the noise. In this section, three simulation experiments are carried out to evaluate the performance of ITMPM, while LPSVD is used as a benchmark against which ITMPM can be compared. The MATLAB code for both ITMPM and the original Kumaresan–Tufts LPSVD algorithm with bias compensation (*12, 13*) is listed in the Appendix. The SNR is here defined as

$$\mathrm{SNR} \overset{\text{def}}{=} 10\log_{10}\frac{\sum_i |a_i|^2}{\rho}\ (\mathrm{dB}), \quad [19]$$

where ρ is the noise variance.

Example I: Single exponentially damped sinusoid. In this example, the synthesized FIDs are given by $y_n = |a|\exp(j\theta)\exp[(-\alpha + j2\pi f)n] + w_n$ with damping factor $\alpha = 0.1$, normalized frequency $f = 0.52$, absolute amplitude $|a| = 1.0$, phase $\theta = 0$, and $n = 0, \ldots, 24$. For nonparametric estimators (e.g., the Fourier transform), the noise can be observed directly in the spectrum, while for parametric estimators (e.g., LPSVD and ITMPM), the noise manifests itself as uncertainty in the estimated parameters, leading to a distribution of observed parameter values. Ideally the mean of this distribution corresponds to the true value of the parameter, i.e., the estimation method is *unbiased*, while the

TABLE 1
Bias ± Standard Deviation of the Estimated Spectral Parameters

SNR (dB)	Method	Damping factor α	Frequency f	Amplitude $\lvert a\rvert$	Phase θ
50	CRLB	0 ± 2.10 E-4	0 ± 3.35 E-5	0 ± 1.32 E-3	0 ± 1.32 E-3
	ITMPM	−7.34 E-7 ± 2.38 E-4	−1.30 E-6 ± 3.63 E-5	1.69 E-5 ± 1.41 E-3	2.74 E-5 ± 1.37 E-3
	LPSVD	5.83 E-4 ± 2.69 E-4	−1.21 E-6 ± 3.93 E-5	2.55 E-3 ± 1.53 E-3	2.50 E-5 ± 1.43 E-3
40	CRLB	0 ± 6.65 E-4	0 ± 1.06 E-4	0 ± 4.18 E-3	0 ± 4.18 E-3
	ITMPM	−9.99 E-6 ± 7.37 E-4	6.93 E-6 ± 1.13 E-4	−2.71 E-4 ± 4.58 E-3	−3.05 E-4 ± 4.16 E-3
	LPSVD	1.86 E-3 ± 8.35 E-4	5.56 E-6 ± 1.21 E-4	7.77 E-3 ± 4.84 E-3	−2.68 E-4 ± 4.38 E-3
30	CRLB	0 ± 2.10 E-3	0 ± 3.35 E-4	0 ± 1.32 E-2	0 ± 1.32 E-2
	ITMPM	1.57 E-4 ± 2.21 E-3	8.82 E-6 ± 3.50 E-4	9.26 E-4 ± 1.40 E-2	−5.81 E-4 ± 1.31 E-2
	LPSVD	5.98 E-3 ± 2.53 E-3	1.02 E-5 ± 3.85 E-4	2.55 E-2 ± 1.50 E-2	−6.18 E-4 ± 1.38 E-2
24	CRLB	0 ± 4.19 E-3	0 ± 6.68 E-4	0 ± 2.64 E-2	0 ± 2.64 E-2
	ITMPM	2.44 E-4 ± 4.40 E-3	6.46 E-5 ± 7.01 E-4	1.99 E-3 ± 2.71 E-2	−2.44 E-3 ± 2.71 E-2
	LPSVD	1.17 E-2 ± 5.42 E-3	5.70 E-5 ± 7.68 E-4	4.89 E-2 ± 2.92 E-2	−2.21 E-3 ± 2.85 E-2
18	CRLB	0 ± 8.37 E-3	0 ± 1.33 E-3	0 ± 5.26 E-2	0 ± 5.26 E-2
	ITMPM	2.69 E-3 ± 9.74 E-3	−2.54 E-5 ± 1.35 E-3	1.66 E-2 ± 5.52 E-2	4.88 E-4 ± 5.44 E-2
	LPSVD	2.52 E-2 ± 1.26 E-2	−3.22 E-5 ± 1.49 E-3	1.02 E-1 ± 5.92 E-2	5.83 E-4 ± 5.65 E-2

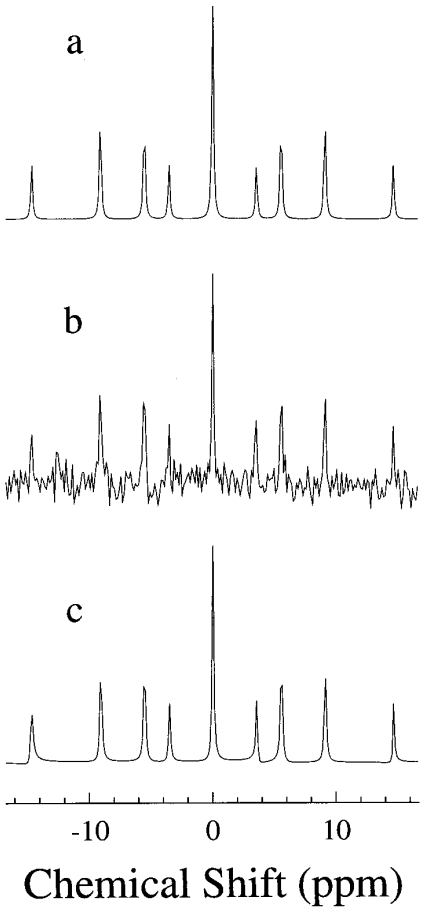

FIG. 1. The ^{19}F spectrum of *p*-fluorophenol calculated assuming weak coupling. The theoretical FIDs used in Example II and Example III are calculated by the NMR computer simulation package GAMMA (*60*). The equilibrium density matrix of the underlying spin system is first treated by a $(\pi/2)_y$ ideal pulse and then propagated in time governed by the isotropic chemical shift and scalar coupling Hamiltonian without relaxation effects. The FID, taken by a single-quantum quadrature detection operator sampling the *xy* magnetization, is apodized by an exponential decay function to simulate T_2 relaxation. (a) FT spectrum of the noiseless FID; (b) FT spectrum of a FID with SNR = 14.5 dB; (c) ITMPM spectrum of (b).

minimum value of its standard deviation is given by the Cramér–Rao lower bound (CRLB), also known as the minimum variance bound (*14, 15*). The calculation of these lower bounds is reasonably straightforward for signals distorted by uncorrelated Gaussian noise and parallels exactly the calculation of the covariance matrix for least-squares model fitting; indeed the CRLBs are identical to the standard deviations on parameter values returned by model fitting (in the absence of systematic error). The details of the calculation can be found elsewhere (*12, 19, 20*).

Table 1 summarizes the results of a Monte Carlo simulation using 500 noise realizations for each value of the SNR, using $M = 1$. It can be seen that the standard deviations (square root of the variance) for ITMPM are close to the theoretical lower bounds, and, unlike LPSVD, the bias on each ITMPM result is always much smaller than its standard deviation. For all SNRs tested here, ITMPM is more accurate and precise than LPSVD.

There are a couple of reasons why ITMPM can perform

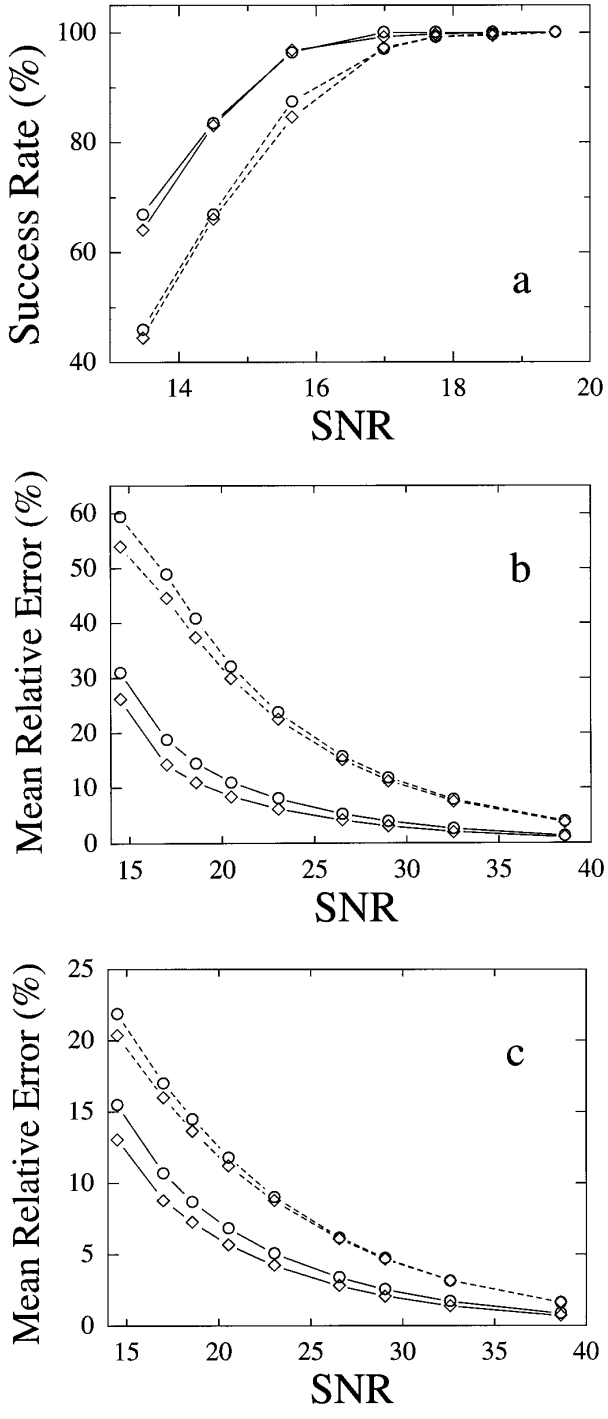

FIG. 2. Monte Carlo simulations (with 400 noise realizations) on the ^{19}F *p*-fluorophenol FIDs of Fig. 1 as a function of SNR. The curves show the results for ITMPM (solid line) and LPSVD (dashed line) using FID lengths of $N = 128$ (○) complex points (sampling from 0 to 2.56 T_2) and $N = 256$ (◇) points (sampling from 0 to 3.84 T_2); (a) success rate, (b) MRE of damping factors, and (c) MRE of amplitudes.

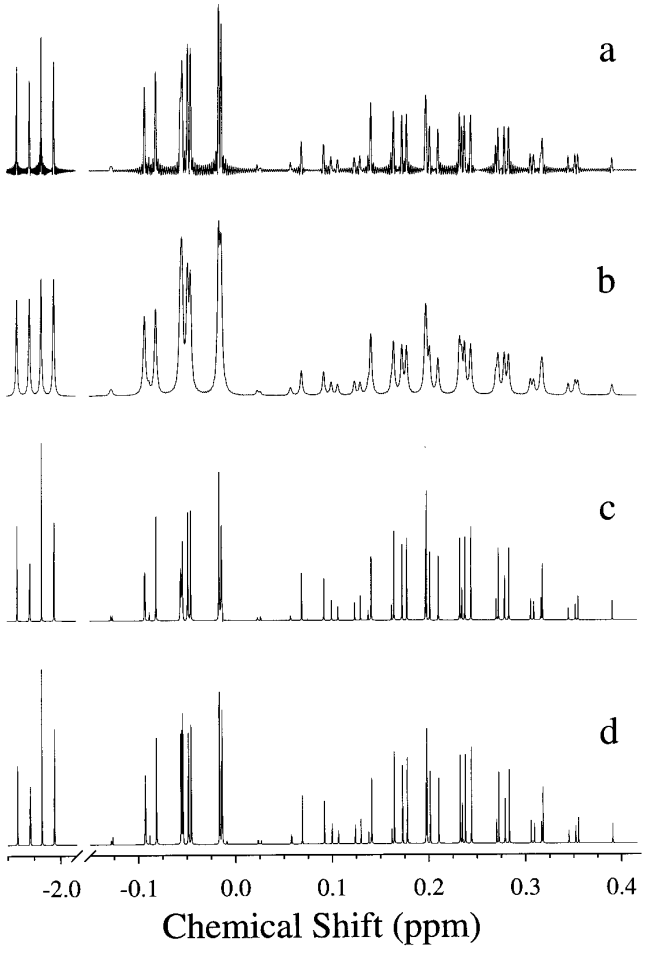

FIG. 3. Calculated ^{1}H spectra of glutamic acid in D_2O. (a) FT of the FID truncated to 384 data points and zero-filled to 4096 points (SNR = 54 dB); (b) FT spectrum of (a) apodized by multiplying the truncated FID by an exponential decay; (c) ITMPM spectrum of (a); (d) FT spectrum of the original FID of 4096 points and SNR = ∞.

better than LPSVD (*18, 19, 41, 42*). Despite their differences, both methods solve the following matrix prediction equation,

$$\hat{\boldsymbol{Y}}_0\boldsymbol{C} = \boldsymbol{Y}_1. \qquad [20]$$

The solution obtained by ITMPM is

$$\boldsymbol{C}_{\mathrm{ITMPM}} = \hat{\boldsymbol{Y}}_0^{\#}\boldsymbol{Y}_1 = [\hat{\boldsymbol{Y}}_0^{\#}\boldsymbol{y}_L, \hat{\boldsymbol{Y}}_0^{\#}\boldsymbol{y}_{L-1}, \ldots, \hat{\boldsymbol{Y}}_0^{\#}\boldsymbol{y}_1], \quad [21]$$

where $\boldsymbol{y}_i = [y_i, y_{i+1}, \ldots, y_{i+N-L-1}]^{\mathrm{T}}$ $(i = 1, 2, \ldots, L)$. Each column of $\boldsymbol{C}_{\mathrm{ITMPM}}$ (say, the jth column) is a solution to the equation

$$\hat{\boldsymbol{Y}}_0\boldsymbol{c} = \boldsymbol{y}_{L-j+1}. \qquad [22]$$

Note that, as $\hat{\boldsymbol{Y}}_0$ has a nontrivial null space, the least-squares solution to Eq. [22] is not unique. Of all the least-squares solution, the one with the minimum Euclidean norm is unique and is given by $\hat{\boldsymbol{Y}}_0^{\#}\boldsymbol{y}_{L-j+1}$. The minimum-norm choice has been shown to be an effective way to overcome the estimate sensitivity to noise perturbation (*43*). On the other hand, the solution provided by LPSVD is simply (cf. Eq. [2])

$$\boldsymbol{C}_{\mathrm{LPSVD}} = [\hat{\boldsymbol{Y}}_0^{\#}\boldsymbol{y}_L, \boldsymbol{J}_L, \boldsymbol{J}_{L-1}, \ldots, \boldsymbol{J}_2], \qquad [23]$$

where $\boldsymbol{J}_i$ is the $L \times 1$ vector with its ith element equal to 1 and all other elements zero. It is clear from Eq. [23] that only the first column of $\boldsymbol{C}_{\mathrm{LPSVD}}$ is the minimum-norm solution to Eq. [22] while the other columns are just trivial solutions. $\boldsymbol{C}_{\mathrm{ITMPM}}$, however, is the unique minimum-norm solution to the *full* matrix prediction equation of Eq. [20]; this results in better immunity to noise perturbation.

In addition, ITMPM distinguishes more reliably between signal and noise eigenvalues (poles). This is because $\boldsymbol{C}_{\mathrm{ITMPM}}$ has M eigenvalues at $\{z_i\}_{1\leq i\leq M}$ and $L - M$ extraneous zero eigenvalues, while $\boldsymbol{C}_{\mathrm{LPSVD}}$ has M eigenvalues at $\{z_i\}_{1\leq i\leq M}$ and $L - M$ extraneous eigenvalues that are nonzero and located inside the unit circle. (Recall that $\boldsymbol{C}_{\mathrm{LPSVD}}$ is in fact the companion matrix of the prediction polynomial in LPSVD, and solving the roots of a polynomial is equivalent to solving the eigenvalues of its companion matrix.) Consequently, LPSVD requires a tedious pruning step in discriminating the signal eigenvalues from the extraneous ones introduced by the noise. This discrimination becomes increasingly difficult as the noise level increases.

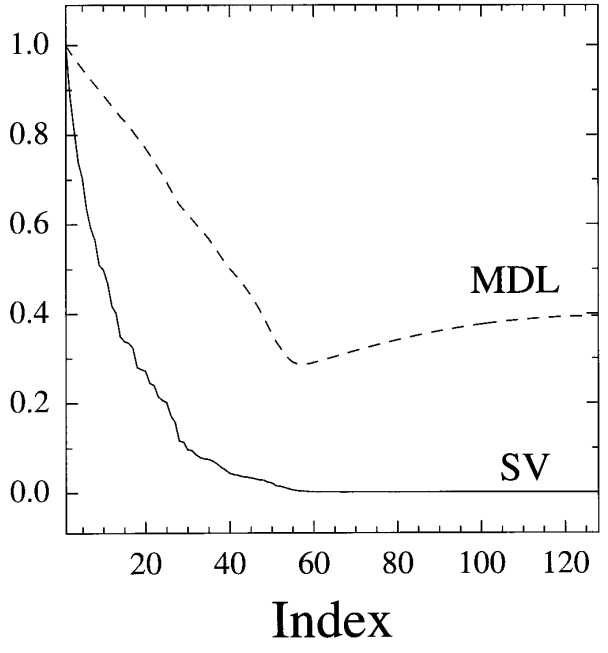

FIG. 4. The functional value of the MDL criterion (dashed line) and the magnitude of the singular values (solid line) (both are normalized) during the ITMPM analysis of Fig. 3c. The optimal matrix rank determined by MDL is $M = 57$.

Example II: ^{19}F spectrum of p-fluorophenol. Figure 1a shows the theoretical ^{19}F NMR spectrum of *p*-fluorophenol, a heteronuclear system containing five spin-$\frac{1}{2}$ species. Figure 1b shows the FT spectrum of a typical FID used in the Monte Carlo simulation (SNR = 14.5 dB) while (c) is the corresponding ITMPM spectrum. The improvement in sensitivity is substantial, although the effects of noise can still be seen as small errors in the phases of the lines as compared to the ideal spectrum, (a).

The results of Monte Carlo simulations using 400 noise realizations for various SNR with $M = 9$ are shown in Fig. 2. Figures 2b and 2c show the mean relative errors in the damping factors and amplitudes, respectively; both ITMPM and LPSVD give very reliable estimates of signal frequencies (*44*). Signals may not be correctly detected at low SNRs; for this example, detection is defined as successful if all the peaks are resolved within ± 0.6 ppm (about $\frac{1}{6}$ of the average peak spacing) of their correct chemical shifts. Figure 2a shows the failure rate as a function of SNR. It is apparent that ITMPM not only provides significantly more accurate estimates of the spectral parameters (by a factor of 2–4), but also has a lower SNR ''breakdown'' threshold (lower by $\approx$1.5 dB, as estimated from the success rate).

In practice, the usefulness of ''exotic'' spectral estimation methods, whatever their theoretical advantages, is limited if they involve significant computational overheads. ITMPM, however, is significantly faster than methods such as LPSVD, since no high degree polynomial rooting is needed. This is particularly true for general NMR FID lengths of about 1–2K points, where the number of floating-point operations (flops) is reduced by roughly an order of magnitude, as measured in their MATLAB implementations. ITMPM can still not compete, however, with the efficiency of the fast Fourier transform. The FFT requires computational time proportional to $O(N \log_2 N)$, whereas SVD is an $O[(N - L) \times L^2]$ process. Should computational efficiency be of great concern, rapid SVD of a Toeplitz or Hankel matrix can be accomplished by using the Lanczos algorithm (*45*) and/or explicitly exploiting the Toeplitz or Hankel structure of the data matrix (*46*). Moreover, the result of the ITMPM analysis is a table of spectral parameters which either can be used to construct a frequency/time domain spectrum or can be used directly without extra tedious manual operation for peak searching, curve fitting, and intensity integration. With the present processing protocol, results can be obtained in 4.5 min of CPU time for 1024 complex data points on a SGI Indigo workstation with a 100-MHz R4000 processor and a spectroscopist's intervention is needed only for data

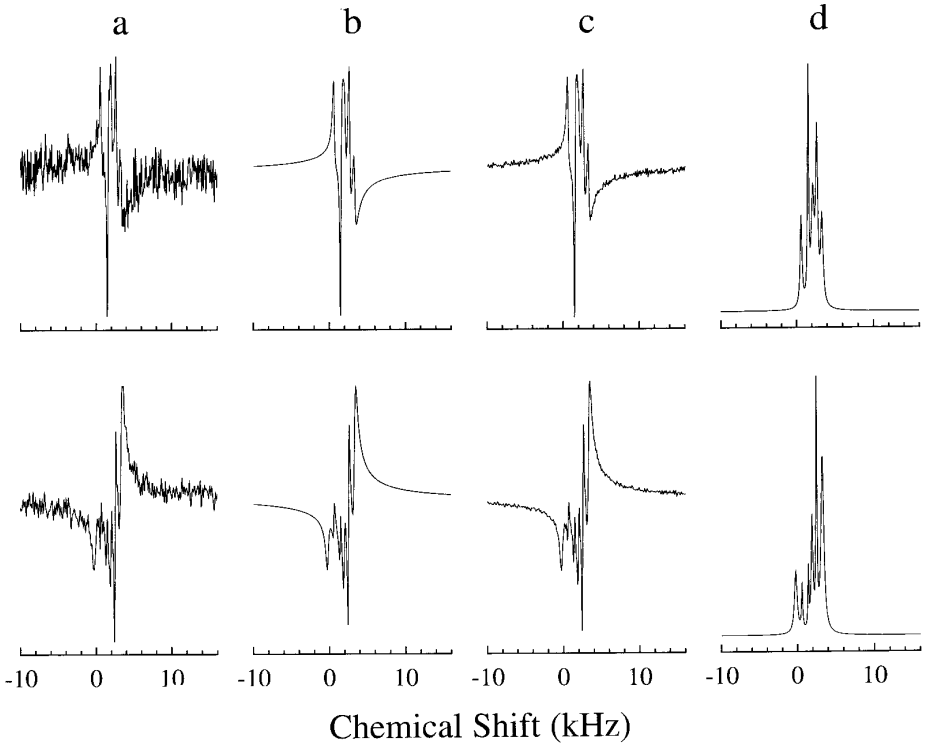

FIG. 5. Experimental spectra of static polycrystalline anthracene (upper row) and malonic acid (lower row) measured at room temperature on a Bruker AM-400 spectrometer with 231 points and initial delays of 600 μs (anthracene) and 856 μs (malonic acid). (a) FT spectra of the FIDs obtained by accumulating 10,240 (anthracene) and 8192 (malonic acid) scans; (b) ITMPM spectra of (a) with the optimal matrix rank, $M = 5$ (anthracene) and $M = 6$ (malonic acid), determined by the MDL criterion; (c) FT spectra of the FIDs obtained under the same experimental conditions as (a) but accumulating 198,865 (anthracene) and 117,112 (malonic acid) scans; (d) phase-corrected ITMPM spectra, obtained from (b) by setting all the phases to zero.

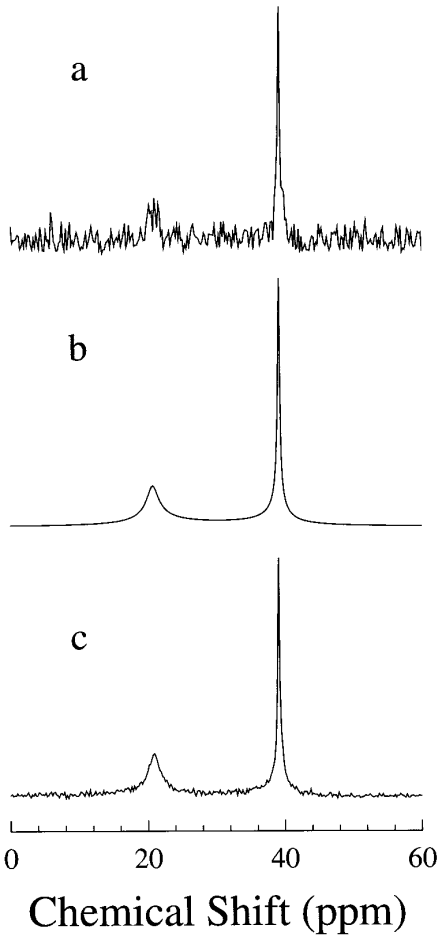

FIG. 6. Experimental spectra of an aqueous solution containing 0.20 *M* $CaCl_2$ and 0.10 *M* sodium EDTA obtained at 300 K on a Bruker AM-400 spectrometer. (a) FT spectra of the FID obtained by accumulating 50,000 scans; (b) ITMPM spectra of (a) with the optimal matrix rank, $M = 2$, determined by the MDL criterion; (c) FT spectra of the FIDs obtained under the same experimental conditions as (a) but accumulating 1,700,000 scans.

transfer (compared to 2 s and 40 min for FT and LPSVD, respectively).

Example III: ^1H spectrum of glutamic acid in D_2O. Classical spectral estimation using Fourier transformation of truncated data sets implicitly assumes that the unobserved data are zero. A blurred spectral estimate is a consequence of such an unrealistic assumption. It should be possible to obtain a better estimate by using *a priori* knowledge to construct a suitable model for the unobserved data. This would eliminate the need for window functions, along with the associated trade-offs of resolution and sensitivity. Moreover, by incorporating lineshape information, parametric methods, such as ITMPM, are able to resolve overlapped signals that cannot be separated in an FT spectrum.

This is illustrated in Fig. 3 using the calculated ^1H spectrum of glutamic acid in D_2O, whose ''true'' spectrum is shown in (d). The result of Fourier transforming the FID truncated to 384 complex data points (and zero-filled back to 4096) is shown in Fig. 3a. The details of the spectrum are obscured and distorted by sinc-wiggles from nearby stronger signals. Figure 3b shows the result of applying an exponential apodizing function to the FID prior to Fourier transformation. This effectively attenuates the sinc-wiggles, at the expense, however, of resolution. In contrast, applying ITMPM to the truncated FID gives a spectrum, Fig. 3c, in which the sinc-wiggles are suppressed without degrading resolution. This is particularly significant in multidimensional spectroscopy where the indirectly detected dimensions must often be severely truncated to keep the experiment time within reasonable bounds. Clearly ITMPM can effectively reduce the truncation artifacts.

Figure 4 plots the singular values obtained during the ITMPM analysis of Fig. 3c. It is clear from the gradual decrease of the singular values (solid line) that the separation of signal singular values from those associated with noise is not straightforward and subjective approaches are likely to be unsatisfactory. By contrast, the minimum of the MDL criterion (dashed line) is relatively well defined.

APPLICATIONS

This work has been prompted by our studies of the long-time behavior of free induction decays in solid-state NMR. Apparently high-resolution spectra (more characteristic of solution rather than solid-state NMR) have been obtained by using a sufficiently long delay between simple pulse excitation and data acquisition. This striking phenomenon has been observed in various experiments using static, magic angle spinning, and off-magic angle spinning samples, and has attracted a wide variety of explanations (*47–55*). Extensions of SPEDAS (single-pulse excitation delayed acquisition spectroscopy) to two-dimensional experiments involving ^1H COSY and multiquantum coherence using polycrystalline fumaric acid monoethyl ester (a ''rigid'' solid in which there is no motional averaging) have also been reported (*48, 50*).

Investigation of these experiments is hampered by the extremely large phase distortions and severe sensitivity losses (by factors of 10^{-2}–10^{-4}) that result from the long delay between excitation and acquisition. The use of signal averaging to improve the SNR is limited by the long T_1 spin–lattice relaxation times of solid-state samples. Hence, experimental restrictions on the total acquisition time result in low sensitivity, and potentially poor spectral resolution in multidimensional experiments (due to data truncation in the indirectly detected dimensions).

Clearly a more sophisticated spectral estimation method than the Fourier transform is required to analyze the results of these experiments. Although the nature of the signals is still currently under investigation, our experimental experience suggests that modeling as a sum of exponentials is a reasonable approximation (*56*). Figure 5a shows the FT spectra using 10,240 and 8192 scans for static polycrystalline anthracene (upper row)

and malonic acid (lower row), respectively. Despite the extensive signal averaging, the features of the spectrum are still difficult to distinguish from the high-level background noise. The corresponding ITMPM spectra are shown in Fig. 5b. Increasing the number of scans from 10,240 to 198,856 for anthracene and from 8192 to 117,112 for malonic acid gives the FT spectra shown in Fig. 5c. The difference between these ''true'' spectra and the ITMPM spectra derived from much noisier signals is very small. The ability of ITMPM to extract the signal information from relatively noisy FIDs greatly facilitates the study of the long-time behavior of solid-state NMR signals, particularly as the problems of phase correction are eliminated, cf. Fig. 5d.

ITMPM should also be important in other experiments where the FID is truncated or whenever the SNR is low, as demonstrated below. Figure 6a presents the FT spectrum of a noisy ^{43}Ca FID from an aqueous solution containing 0.20 *M* $CaCl_2$ (natural abundance) and 0.10 *M* sodium EDTA, and acquired with 50,000 scans. Its corresponding ITMPM spectrum is shown in Fig. 6b. From the chemical composition of the solution, it is known that the FID should consist of two exponential components corresponding to free Ca^{2+} and the Ca–EDTA complex. We have also performed another measurement under the same experimental conditions except for an increase of scans from 50,000 to 1,700,000. Its FT spectrum is depicted in Fig. 6c and will be used as the ''true'' spectrum for comparison. Again, the ITMPM spectrum serves as a very close approximation to the ''true'' spectrum.

CONCLUSIONS

An often recurring problem in NMR spectroscopy is how to improve spectral sensitivity and resolution. The Monte Carlo simulations described above indicate that this combined detection–estimation scheme, ITMPM, is able to simultaneously achieve these objectives. The deficiencies of the applications of the FT to imperfect NMR data can be diminished to a large extent at the expense of reasonable computational complexity. Compared to LPSVD, it has the advantages of greater computational efficiency, higher precision and accuracy of the estimated spectral parameters, and less tendency for spurious estimates at low signal-to-noise ratios. ITMPM is essentially near-optimal over the range of signal-to-noise ratios investigated; the parameter estimates have biases that are close to zero and standard deviations close to their Cramér–Rao lower bounds.

For experimental applications, ITMPM considerably facilitates our analysis of the delayed acquisition data, and increases the potential applicability of SPEDAS-like experiments to solid-state NMR, by providing phase-corrected spectra with much improved sensitivity and resolution. ITMPM should also be important in other experiments where the FID is truncated or whenever the SNR is low. This is often the case in one- and multidimensional NMR studies of low gamma nuclei, *in vivo* samples, and macromolecules of biological interest. Consequently, ITMPM could become a valuable quantification tool within NMR and other branches of Fourier spectroscopy.

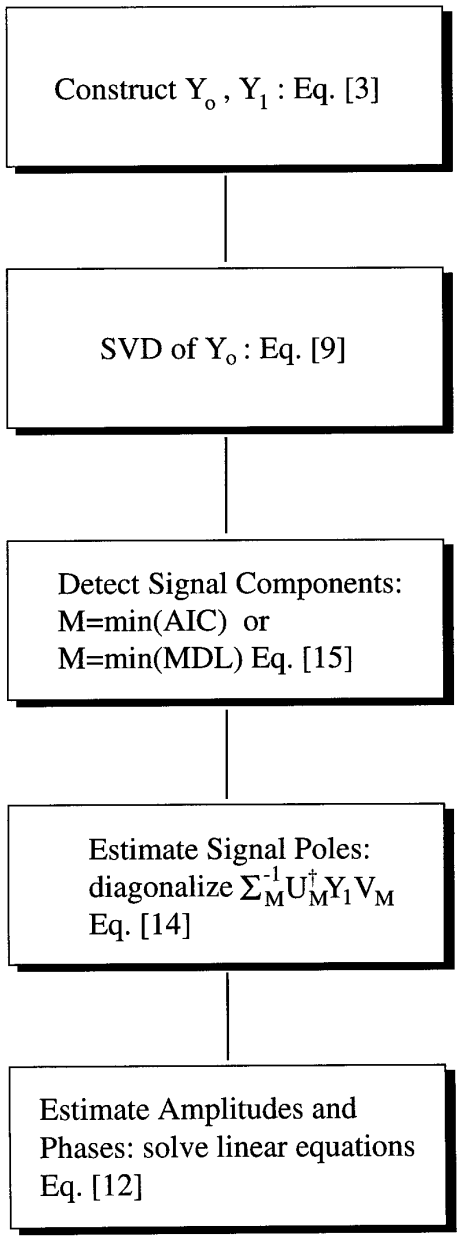

FIG. 7. The flowchart of the ITMPM algorithm.

APPENDIX

To assist in the implementation of ITMPM and related techniques, MATLAB programs for ITMPM and LPSVD are listed below; the flowchart for the ITMPM algorithm is depicted in Fig. 7. MATLAB has the advantage of providing easy

access to matrix software developed by the LINPACK (*57*) and EISPACK (*58*) projects, and allowing rapid code development and refinement (*59*).

```
function [para,M,itc]=itcmp(y,M)
% ITMPM information theoretic criteria and matrix pencil method
% function [para,M,itc]=itcmp(y,M)
% author:     Yung-Ya Lin, 4/15/96
% arguments:
%      y: complex vector, NMR FID time series
%      M: real scalar, number of signals or effective matrix rank
%         M=-1 using AIC; M=-2 using MDL; M >= 0 using the user's input value
%   para: real M*4 matrix, estimated damping factor, frequency, amplitude, phase
%    itc:  real vector, containing AIC or MDL function values
y=y(:); N=length(y);
L=floor(N/3);                          % pencil parameter
Y=toeplitz(y(L+1:N),y(L+1:-1:1));  % YO=Y(:,2:L+1), Y1=Y(:,1:L) Eq. [3]
[U,S,V]=svd(Y(:,2:L+1),0);             % singular value decomposition
S=diag(S);
itc=zeros(1,L);
if M==-1                               % determining M by AIC
   for k=0:L-1;
      itc(k+1)=-2*N*sum(log(S(k+1:L))) . . .
                  + 2*N*(L-k)*log((sum(S(k+1:L))/(L-k))) + 2*k*(2*L-k);
   end
   [tempY, tempI]=min(itc); M=tempI-1;
end
if M==-2                                        % determining M by MDL Eq. [16]
   for k=0:L-1;
      itc(k+1)=-N*sum(log(S(k+1:L))) . . .
                  + N*(L-k)*log((sum(S(k+1:L))/(L-k))) + k*(2*L-k)*log(N)/2;
   end
   [tempY, tempI]=min(itc); M=tempI-1;
end
s=log(eig(diag(1./S(1:M))*. . .              % signal pole z=exp(s)
   ((U(:,1:M)'*Y(:,1:L))*V(:,1:M))));
Z=zeros(N,M);
for k=1:M;     Z(:,k)=exp(s(k)).^[0:N-1].'; end;
a=Z\y;                                       % linear least squares analysis
para=[-real(s) imag(s)/2/pi abs(a) imag(log(a./abs(a)))];
return

function [para]=lpsvd(y,M)
% LPSVD linear prediction with singular value decomposition
% function [para]=lpsvd(y,M)
% author:    Yung-Ya Lin, 4/15/96
% reference: R. Kumaresan, D. W. Tufts IEEE Trans. Acoust. Speech Signal Processing
%            vol. ASSP-30, 837-840, 1982.
% arguments:
%      y: complex vector, NMR FID time series
%      M: real scalar, number of signals or effective matrix rank
%   para: real M*4 matrix, estimated damping factor, frequency, amplitude, phase
y=y(:);
N=length(y);                                 % # of complex data points in FID
L=floor(N*3/4);                              % linear prediction order L = 3/4*N
A=hankel(conj(y(2:N-L+1)),conj(y(N-L+1:N)));  % backward prediction data matrix
h=conj(y(1:N-L));                            % backward prediction data vector
[U,S,V]=svd(A,0);                            % singular value decomposition
clear A;
S=diag(S);
bias=mean(S(M+1:min([N-L,L])));              % bias compensation
 b=-V(:,1:M)*(diag(1./(S(1:M)-bias))*(U(:,1:M)'*h));
                                             % prediction polynomial coefficients
s=conj(log(roots([b(length(b):-1:1);1])));   % polynomial rooting
s=s(find(real(s)<0));                        % extract true signal poles
Z=zeros(N,M);
for k=1:M; Z(:,k)=exp(s(k)).^[0:N-1].'; end;
a=Z\y;                                       % linear least squares analysis
 para=[-real(s) imag(s)/2/pi abs(a) imag(log(a./abs(a)))];
 return
```

ACKNOWLEDGMENTS

This work was supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Materials Sciences Division of the U.S. Department of Energy under Contract DE-AC03-76SF00098.

REFERENCES

1. J. C. Lindon and A. G. Ferrige, *Prog. NMR Spectrosc.* **14,** 27 (1980).
2. R. de Beer and D. van Ormondt, *in* ''NMR Basic Principles and Progress'' (P. Diehl, E. Fluck, H. Gunther, R. Kosfeld, and J. Seelig, Eds.), Vol. 26, pp. 201, Springer-Verlag, Berlin/Heidelberg (1992).
3. R. E. Hoffman and G. C. Levy, *Prog. NMR Spectrosc.* **23,** 211 (1991).
4. J. J. Led and H. Gesmar, *Chem. Rev.* **91,** 1413 (1991).
5. R. Roy, B. G. Sumpter, G. A. Pfeffer, S. K. Gray, and D. W. Noid, *Phys. Rep.* **205,** 109 (1991).
6. H. Gesmar, J. J. Led, and F. Abildgaard, *Prog. NMR Spectrosc.* **22,** 255 (1990).
7. D. S. Stephenson, *Prog. NMR Spectrosc.* **20,** 515 (1988).
8. H. Barkhuijsen, R. de Beer, A. C. Drogendijk, D. van Ormondt, and J. W. C. van der Veen, *in* ''Proceedings, International School of Physics 'Enrico Fermi' on the 'Physics of NMR Spectroscopy in Biology and Medicine' '' (B. Maraviglia, Ed.), pp. 313, Italian Physical Society, 1988.
9. M. Mehring, ''Principles of High Resolution NMR in Solids,'' 2nd ed., Springer-Verlag, Berlin/New York, 1983.
10. A. van der Veen, E. F. Deprettere, and A. L. Swindlehurst, *Proc. IEEE* **81,** 1277 (1993).
11. The Mathworks, Inc., ''MATLAB Reference Guide,'' South Natick, MA (1992).
12. R. Kumaresan and D. W. Tufts, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-30,** 833 (1982).
13. H. Barkhuijsen, R. de Beer, W. M. M. J. Bovee, and D. van Ormondt, *J. Magn. Reson.* **61,** 465 (1985).
14. S. M. Kay, ''Modern Spectral Estimation,'' Prentice–Hall, Engle15. S. L. Marple, ''Digital Spectral Analysis with Applications,'' Prentice–Hall, Englewood Cliffs, NJ (1987).
16. S. Y. Kung, K. S. Arun, and D. V. Bhaskar Rao, *J. Opt. Soc. Am.* **73,** 1799 (1983).
17. H. Barkhuijsen, R. de Beer, and D. van Ormondt, *J. Magn. Reson.* **73,** 553 (1987).
18. Y. Hua and T. K. Sarkar, *IEEE Trans. Signal Process.* **39,** 892 (1991).
19. Y. Hua and T. K. Sarkar, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-38,** 814 (1990).
20. P. Hodgkinson and P. J. Hore, *Adv. Magn. Opt. Reson.* **20,** 187 (1997).
21. R. Roy, A. Paulraj, and T. Kailath, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-34,** 1340 (1986).
22. G. Zhu and A. Bax, *J. Magn. Reson.* **100,** 202 (1992).
23. J. Tang and J. R. Norris, *J. Chem. Phys.* **84,** 5210 (1986).
24. J. Tang and J. R. Norris, *Chem. Phys. Lett.* **131,** 252 (1986).
25. J. Tang, C. P. Lin, M. K. Bowman, and J. R. Norris, *J. Magn. Reson.* **62,** 167 (1985).
26. Y. K. Lee, R. L. Vold, G. L. Hoatson, Y.-Y. Lin, and A. Pines, *J. Magn. Reson. A* **112,** 112 (1995).
27. A. C. Kot, D. W. Tufts, and R. J. Vaccaro, *IEEE Trans. Signal Process.* **41,** 3174 (1993).
28. G. H. Golub and C. F. Van Loan, ''Matrix Computations,'' 2nd ed., Johns Hopkins Univ. Press, Baltimore, MD (1989).
29. C. Eckart and G. Young, *Psychometrika* **1,** 211 (1936).
30. K. Konstantinides and K. Yao, *IEEE Trans. Signal Process.* **36,** 757 (1988).
31. J. M. Chambers, ''Computational Methods for Data Analysis,'' Wiley, New York (1977).
32. J. A. Cadzow, B. Baseghi, and T. Hsu, *IEE Proc. Part F* **130,** 202 (1983).
33. M. S. Bartlett, *J. R. Stat. Soc. Ser. B* **16,** 296 (1954).
34. D. N. Lawley, *Biometrica* **43,** 128 (1956).
35. I. Dologlou and G. Carayannis, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-39,** 1681 (1991).
36. M. Shinnar and S. M. Eleff, *J. Magn. Reson.* **76,** 200 (1988).
37. M. Wax and T. Kailath, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-33,** 387 (1985).
38. H. Akaike, *IEEE Trans. Automat. Control* **AC-19,** 716 (1974).
39. G. Schwartz, *Ann. Stat.* **6,** 461 (1978).
40. J. Rissanen, *Automatica* **14,** 465 (1978).
41. Y.-M. Wang, H. Lee, and D. V. Apte, *Int. J. Imag. Syst. Tech.* **4,** 201 (1992).
42. H. Ouibrahim, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-37,** 133 (1989).
43. B. D. Rao, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-36,** 1026 (1988).
44. P. Koehl, C. Ling, and J. F. Lefèvre, *J. Magn. Reson. A* **109,** 32 (1994).
45. G. L. Millhauser, A. A. Carter, D. J. Schneider, J. H. Freed, and R. E. Oswald, *J. Magn. Reson.* **82,** 150 (1989).
46. W. W. F. Pijnappel, A. van den Boogaart, R. de Beer, and D. van Ormondt, *J. Magn. Reson.* **97,** 122 (1992).
47. S. Ding and C. A. McDowell, *J. Magn. Reson. A* **111,** 212 (1994).
48. S. Ding and C. A. McDowell, *J. Magn. Reson. A* **115,** 141 (1995).
49. S. Ding and C. A. McDowell, *J. Magn. Reson. A* **117,** 171 (1995).
50. S. Ding and C. A. McDowell, *J. Magn. Reson. A* **120,** 261 (1996).
51. S. Ding and C. A. McDowell, *Chem. Phys. Lett.* **255,** 151 (1996).
52. S. Ding and C. A. McDowell, *Chem. Phys. Lett.* **259,** 538 (1996).
53. B. C. Gerstein, J. Z. Hu, J. Zhou, C. Ye, M. Solum, R. Pugmire, and D. M. Grant, *Solid State NMR* **6,** 63 (1996).
54. J. Z. Hu, J. Zhou, F. Deng, H. Feng, N. Yang, L. Li, and C. Ye, *Solid State NMR* **6,** 85 (1996).
55. B. M. Fung, T. Dollase, M. L. Magnuson, and T.-H. Tong, *J. Magn. Reson. A* **123,** 56 (1996).
56. M. Engelsberg and I. J. Lowe, *Phys. Rev. B* **10,** 822 (1974).
57. J. J. Dongarra, C. B. Moler, J. R. Bunch, and G. W. Steward, ''LINPACK User's Guide,'' SIAM, Philadelphia (1979).
58. B. T. Smith, J. M. Boyle, J. J. Dongarra, B. S. Garbow, Y. Ikebe, V. C. Klema, and C. B. Moler, ''Matrix Eigensystem Routines—EISPACK Guide,'' Springer-Verlag, Berlin (1976).
59. B. P. O. van Tongeren, A. F. Mehlkopf, D. van Ormondt, M. Schramp, and R. de Beer, *J. Magn. Reson.* **100,** 619 (1992).
60. S. A. Smith, T. O. Levante, B. H. Meier, and R. R. Ernst, *J. Magn. Reson. A* **106,** 75 (1994).