515

Improved frequency resolution in multidimensional constant-time experiments by multidimensional Bayesian analysis

Roger A. Chylla and John L. Markley*

Department of Biochemistry, University of Wisconsin-Madison, 420 Henry Mall, Madison, WI 53706, U.S.A.

SUMMARY

The resolution of spectral frequencies in NMR data obtained from discrete Fourier transformation (DFT) along D constant-time dimensions can be improved significantly through extrapolation of the D-dimensional free induction decay (FID) by multidimensional Bayesian analysis. Starting from Bayesian probability theory for parameter estimation and model detection of one-dimensional time-domain data [Bretthorst, (1990) *J. Magn. Reson.*, **88**, 533–551; 552–570; 571–595], a theory for the D-dimensional case has been developed and implemented in an algorithm called BAMBAM (**BA**yesian **M**odel **B**uilding **A**lgorithm in **M**ultidimensions). BAMBAM finds the most probable sinusoidal model to account for the systematic portion of any D-dimensional stationary FID. According to the parameters estimated by the algorithm, the FID is extrapolated in D dimensions prior to apodization and Fourier transformation. Multidimensional Bayesian analysis allows for the detection of signals not resolved by the DFT alone or even by sequential one-dimensional extrapolation from mirror-image linear prediction prior to the DFT. The procedure has been tested with a theoretical two-dimensional dataset and with four-dimensional HN(CO)CAHA (Kay et al. (1992) *J. Magn. Reson.*, **98**, 443–450) data from a small protein (8 kDa) where BAMBAM was applied to the $^{13}C_\alpha$ and H_α constant-time dimensions.

INTRODUCTION

The estimation of spectral frequency, amplitude, phase and decay rate parameters from NMR time-domain data is accomplished conventionally by examination of the absorption spectrum of the discrete Fourier transform (DFT) of the data. The frequency resolution obtained from the DFT, however, is limited when acquisition times are short, as is the case for the indirectly detected dimensions of 3D and 4D NMR experiments. In many 4D experiments, for example, as few as eight complex points may be acquired along one or more of the indirectly detected dimensions.

---

*To whom correspondence should be addressed.

TABLE 1
RECEIVER TIME AND PHASE VALUES FOR VARIOUS METHODS OF ACQUISITION[a]

Method of acquisition	Position (x)	Time (t)	Phase (ϕ)
Redfield[b]	1	0	x (0)
	2	Δt	y ($\pi/2$)
	3	$2\Delta t$	x (0)
	4	$3\Delta t$	y ($\pi/2$)
TPPI[c]	1	0	x (0)
	2	Δt	y ($\pi/2$)
	3	$2\Delta t$	$-$x (π)
	4	$3\Delta t$	$-$y ($3\pi/2$)
States[d]	1	0	x (0)
	2	0	y ($\pi/2$)
	3	Δt	x (0)
	4	Δt	y ($\pi/2$)
States-TPPI[d]	1	0	x (0)
	2	0	y ($\pi/2$)
	3	Δt	$-$x (π)
	4	Δt	$-$y ($3\pi/2$)

[a] For each point y_i in the D-dimensional data matrix, there is a vector $X_i = [x_{i1}, x_{i2}, \ldots x_{iD}]$ which specifies the position of the point along dimension d = [1,2,...D] of the matrix. The vector X_i uniquely specifies the time and phase of y_i along each dimension according to the method of acquisition displayed in the table.
[b] Redfield and Kunz, 1975.
[c] Bodenhausen et al., 1980.
[d] States et al., 1982.

Extrapolation of the FID by linear prediction (Barkhuijsen et al., 1985) prior to the Fourier transform (Gesmar and Led, 1989; Zhu and Bax, 1990; Led and Gesmar, 1991; Kay et al., 1992) has been one approach to increasing the resolution of highly truncated NMR data (see Stephenson, 1988; Gesmar et al., 1990 for reviews). Although computationally stable and relatively fast, 1D linear prediction is sensitive to noise and is reliable only if the number of frequency components is less than one quarter of the number of available data points (Kumaresan and Tufts, 1982). Even with the use of constant-time experiments (Bax et al., 1979; Kessler et al., 1984) to increase the number of available data points through mirror-image time projection (Zhu and Bax, 1990), 1D linear prediction of highly truncated data can produce lineshape distortions and artifacts. A recently reported algorithm for simultaneous 2D linear prediction (Zhu and Bax, 1992) reduces these problems but only at a large computational expense.

Recently, Bayesian probability theory has been used to estimate relevant spectral information from 1D NMR data (Bretthorst, 1990a,b,c) with much greater resolution and accuracy than by

conventional Fourier transform processing (Kotyk et al., 1992). Through assigning probabilities to questions relating to parameter estimation and model selection, Bayesian analysis can construct the most probable mathematical model that is consistent with the NMR data and any known prior information. NMR data acquired with constant-time acquisition periods are particularly amenable to Bayesian analysis because the mathematical models describing such data need not include decay terms. We develop here an extension of 1D NMR parameter estimation by Bayesian analysis to the D-dimensional case. The theory is implemented in an algorithm for extrapolating D-dimensional stationary NMR data which, when used in conjunction with the Fourier transform, can significantly increase the frequency resolution obtainable from NMR experiments containing D constant-time dimensions. Preliminary results of this study have been presented elsewhere (Chylla and Markley, 1993).

THEORY

Bayesian probability theory as applied to the estimation of spectral parameters in 1D NMR data (Bretthorst, 1990a,b,c) has been extended here to D dimensions. Bretthorst's notation has been retained insofar as appropriate.

Consider a D-dimensional, rectangular matrix of time-domain NMR data in which the number of data points along dimensions $d \equiv [1, 2, ..., D]$ is defined by $n_d \equiv [n_1, n_2, ..., n_D]$ respectively. The total number of data points, N, is given by

$$N = \prod_{d=1}^{D} n_d \tag{1}$$

For convenience, the D-dimensional data matrix can be expressed as a 1D vector of N data points $Y \equiv [y_1, y_2, ..., y_N]$, where the position, time and phase in D-dimensional space of each data point y_i is defined by matrices X, T, and Φ respectively:

$$X \equiv \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{N1} \\ x_{12} & x_{22} & \cdots & x_{N2} \\ \cdots & \cdots & \cdots & \cdots \\ x_{1D} & x_{2D} & \cdots & x_{ND} \end{bmatrix} \quad T \equiv \begin{bmatrix} t_{11} & t_{21} & \cdots & t_{N1} \\ t_{12} & t_{22} & \cdots & t_{N2} \\ \cdots & \cdots & \cdots & \cdots \\ t_{1D} & t_{2D} & \cdots & t_{ND} \end{bmatrix} \quad \Phi \equiv \begin{bmatrix} \phi_{11} & \phi_{21} & \cdots & \phi_{N1} \\ \phi_{12} & \phi_{22} & \cdots & \phi_{N2} \\ \cdots & \cdots & \cdots & \cdots \\ \phi_{1D} & \phi_{2D} & \cdots & \phi_{ND} \end{bmatrix}$$

According to this notation, x_{id}, t_{id} and ϕ_{id} define, respectively, the position, time and phase of y_i along dimension d. These values are known parameters determined by the method of quadrature along d. Table 1 displays values for x, t and ϕ for several methods of quadrature assuming uniform data sampling.

A model describing the dataset may assume that the data can be separated into a systematic portion (signal) and a random portion (noise), both of which are functions of T and Φ:

$$y_i (T_i, \Phi_i) = f (T_i, \Phi_i) + e(T_i, \Phi_i), \ 1 \le i \le N \tag{2}$$

The systematic portion can be modeled as a linear combination of J signal functions:

$$f (T, \Phi) = \sum_{j=1}^{J} B_j U_j(T, \Phi, \Theta_j) \tag{3}$$

where $U_j(T,\Phi,\Theta_j)$ is the jth signal function, B_j is the amplitude of the jth signal function, and Θ_j represents the relevant nonlinear parameters (frequency, phase and decay rate) associated with U_j. The signal function U_j, for the case of D-dimensional time-domain NMR data, is a D-dimensional decaying sinusoid:

$$U_j(T_i,\Phi_i,\Theta_j) = \prod_{d=1}^{D} \cos(\omega_{jd}t_{id} + \phi_{id} + \theta_{jd})e^{-\alpha_{jd}t_{id}}, \; 1 \le i \le N; \; 1 \le j \le J \qquad (4)$$

where $U_j(T_i,\Phi_i,\Theta_j)$ is the value of the jth signal function at data point i, t_{id} and ϕ_{id} are the respective time and phase of the ith data point along dimension d, and ω_{jd}, θ_{jd}, and α_{jd} are the respective angular frequencies, phases and decay rates of the jth signal function along dimension d. For the special case of NMR data in which all D dimensions have been acquired by using a constant-time evolution period, the signal function U_j can be modeled by a stationary sinusoid without decay terms:

$$U_j(T_i,\Phi_i,\Theta_j) = \prod_{d=1}^{D} \cos(\omega_{jd}t_{id} + \phi_{id} + \theta_{jd}), \; 1 \le i \le N; \; 1 \le j \le J \qquad (5)$$

It is important to note the very different natures of ϕ_{id} and θ_{jd}. Although both represent phase terms in Eq. 5, ϕ_{id} is a known acquisition term determined by the method of quadrature; in contrast, θ_{jd} is an unknown parameter to be estimated from the data. To minimize the number of unknown, nonlinear parameters, θ can be expressed as an amplitude (linear) rather than as a phase (nonlinear) parameter. A sinusoid of amplitude B and phase ϕ is equivalent to the linear combination of two sinusoids differing in phase by $\frac{\pi}{2}$:

$$B \cos(\omega t + \phi) = A_1 \cos(\omega t) + A_2 \cos(\omega t + \frac{\pi}{2}) \qquad (6)$$

Application of this one-dimensional trigonometric identity to the D-dimensional Eq. 3 yields the following expression for $f(T,\Phi)$

$$f(T,\Phi) = \sum_{j=1}^{J} \sum_{k=1}^{K} A_{jk}V_{jk}(T,\Phi,W_j) \qquad (7)$$

where $K = 2^D$ is the number of amplitudes per sinusoid, A_{jk} is the kth amplitude of the jth sinusoid, $W_j \equiv [\omega_{j1}, \omega_{j2}, ..., \omega_{jD}]$, and $V_{jk}(T,\Phi,W_j)$, the kth signal function of the jth sinusoid, is given by

$$V_{jk}(T_i,\Phi_i,W_j) = \prod_{d=1}^{D} \cos(\omega_{jd}t_{id} + \phi_{id} + \psi_{kd}), \; 1 \le i \le N; \; 1 \le j \le J; \; 1 \le k \le K \qquad (8)$$

$$\Psi_{kd} = \left(\frac{\pi}{2}\right)[((k-1)/(2^{d-1})) \% 2], \; 1 \le d \le D; \; 1 \le k \le K \qquad (9)$$

The operators / and % represent the integer division and integer modulus operators respectively. Table 2 displays values of Ψ_{kd} for the cases of D = 1, 2, and 3. The parameters B (Eq. 3) and ϕ (Eq. 5) can be calculated from A (Eq. 7) according to

$$B_j = \sqrt{\sum_{k=1}^{K} A_{jk}^2}, \; 1 \le j \le J \qquad (10)$$

TABLE 2
KNOWN VALUES FOR Ψ_{kd} AS A FUNCTION OF k AND d

Dimensions	Ψ_{kd} [a]								
	d	k = 1				k = 2			
D = 1	1	0				$\pi/2$			
		k = 1		k = 2		k = 3		k = 4	
D = 2	1	0		$\pi/2$		0		$\pi/2$	
	2	0		0		$\pi/2$		$\pi/2$	
		k = 1	k = 2	k = 3	k = 4	k = 5	k = 6	k = 7	k = 8
D = 3	1	0	$\pi/2$	0	$\pi/2$	0	$\pi/2$	0	$\pi/2$
	2	0	0	$\pi/2$	$\pi/2$	0	0	$\pi/2$	$\pi/2$
	3	0	0	0	0	$\pi/2$	$\pi/2$	$\pi/2$	$\pi/2$

[a] There are $k = [1, 2, ...K = 2^D]$ phase terms along each dimension. Each Ψ_{kd} corresponds to the phase term appearing in dimension d of the kth basis function of Eq. 8.

$$\theta_{jd} = \arctan\left(\frac{A_{jz}}{A_{j1}}\right), \ z = (1 + 2^{d-1}); \ 1 \leq j \leq J; \ 1 \leq d \leq D \tag{11}$$

It is convenient to express the jk subscripts in terms of a single subscript, $m = 1 \ldots M$, where $M \equiv JK$ and m takes the M possible combinations of jk. With this change in notation, $A_m = A_{jk}$, V_m is given by

$$V_m(T,\Phi,W_j) = V_{jk}(T,\Phi,W_j), \ m = (k - 1)J + j \tag{12}$$

and $f(T,\Phi)$ is given by

$$f(T,\Phi) = \sum_{m=1}^{M} A_m V_m(T,\Phi,W_j), \ j = [(m - 1) / J] + 1 \tag{13}$$

Equation 13 is a model that is sufficient to describe the systematic portion of all D-dimensional NMR data that can be described by a sum of stationary sinusoids.

Given a specified set of (J + 1) possible models and $S \equiv \{f_0, f_1, f_2, ..., f_J\}$, Bayesian statistics can determine quantitatively which model f_j is the most probable given the data and any known prior information. This probability can be expressed as $P(D|f,\Theta,I)$, the global likelihood of the data D, given the form of the data f, the nonlinear parameters Θ associated with f, and any prior information I. Bretthorst (1990b) derived an expression for $P(D|f,\Theta,I)$ for any dataset that satisfies Eqs. 2 and 3 and whose noise carries a finite total power:

$$P(D \mid f,\Theta,I) \propto \Gamma\left(\frac{M}{2}\right)\Gamma\left(\frac{N-M}{2}\right)\left[\frac{M \ \overline{h^2}}{2}\right]^{-(M/2)}\left[\frac{N \ \overline{d^2} - M \ \overline{h^2}}{2}\right]^{(M-N)/2} \tag{14}$$

In the above equation, $\Gamma(x)$ is the gamma function, M is the number of signal functions, N is the total number of data points, $\overline{d^2}$ is the mean-square data value

$$\overline{d^2} = \frac{1}{N} \sum_{i=1}^{N} y_i^2 \tag{15}$$

$\overline{h^2}$ is the mean-square projection of the data unto a set of orthonormal basis functions $\mathbf{H}$,

$$\overline{h^2} = \frac{1}{M} \sum_{m=1}^{M} h_m^2 \tag{16}$$

$$h_m = \sum_{i=1}^{N} y_i H_m(T_i, \Phi_i), \ 1 \le m \le M \tag{17}$$

$$H_m(T_i, \Phi_i) = \frac{1}{\lambda_m} \sum_{l=1}^{M} e_{ml} V_l(T_i, \Phi_i), \ 1 \le m \le M \tag{18}$$

e_{ml} is the mth component of the lth eigenvector, and λ_m is the mth eigenvalue of the interaction matrix $\mathbf{G}_{lm}$ defined by

$$g_{lm} \equiv \sum_{i=1}^{N} V_l(T_i, \Phi_i) V_m(T_i, \Phi_i), \ g_{lm} \in \mathbf{G}_{lm}, 1 \le l \le M, \ 1 \le m \le M \tag{19}$$

The global likelihood of the data $\mathbf{P(D|f_j, \Theta_j, I)}$ depends upon the number of sinusoids J in the model ($M = 2^D J$) and the values of the nonlinear parameters Θ_j. For the special case of NMR data acquired with constant-time evolution periods, the nonlinear parameters are exclusively frequency parameters ($\Theta_j = \mathbf{W}_j$). $\mathbf{P(D|f, \Theta, I)}$ can be evaluated independent of the linear parameters A_m (Eq. 13). A_m are 'nuisance parameters' (Bretthorst, 1990a) which can be removed from the expression for $\mathbf{P(D|f, \Theta, I)}$ by integration over all possible values of A_m. The values of A_m, which maximize $\mathbf{P(D|f, \Theta, I)}$ for a given set of nonlinear parameters, are given by

$$A_m \equiv \sum_{l=1}^{M} \frac{h_l e_{lm}}{\sqrt{\lambda_l}}, \ 1 \le m \le M \tag{20}$$

Equation 13 can be used not only to estimate the nonlinear parameters that maximize $\mathbf{P(D|f_j, \Theta, I)}$ but also to estimate the number of signals present in the data. A dataset can be said to contain evidence for J signals if

$$\frac{\mathbf{P(D|f_{J+1}, \Theta, I)}}{\mathbf{P(D|f_J, \Theta, I)}} < 1 \quad \text{and} \quad \frac{\mathbf{P(D|f_J, \Theta, I)}}{\mathbf{P(D|f_{J-1}, \Theta, I)}} > 1,$$

where f_J refers to a model containing J signals. In view of the data and any prior information, application of Bayesian probability theory selects the simplest model that still accurately accounts for the systematic portion of the data.

Bayesian probability theory indicates the conditions under which the Fourier transform provides an accurate estimation of the frequencies present in NMR data. For a model with a fixed number of sinusoids, $\mathbf{P(D|f_J, \Theta, I)}$ increases as $\overline{h^2}$ increases. For $J = 1$ ($M = 2^D$) and $\Theta = \mathbf{W} = [\omega_1, \omega_2, \dots, \omega_D]$, it can be shown that

$$\overline{h^2} = \left(\frac{N}{2}\right)^D \sum_{m=1}^{M} T_m^2(\mathbf{W}), \ M = 2^D \tag{21}$$

$$T_m(\mathbf{W}) = \sum_{i=1}^{N} y_i \left[\prod_{d=1}^{D} \cos(\omega_d t_{id} + \phi_{id}) \right] \tag{22}$$

Equation 22 is simply the D-dimensional Fourier transform of the data and the quantity $\overline{h^2}$ is proportional to the power spectrum evaluated at $W = [\omega_1, \omega_2, \ldots, \omega_D]$. For a model consisting of a single stationary frequency, the position of $T_m(W)_{max}$ (maximum of the Fourier transform power spectrum) represents the values of W that are most 'probable' as expressed by the quantity $P(D|f_J, \Theta, I)$.

ALGORITHM

The preceding section has outlined a theory for evaluating the relative probability of different D-dimensional sinusoidal models of NMR data. This section applies that theory in an algorithm designed to construct the most probable sinusoidal model for a D-dimensional NMR dataset acquired under conditions where decay of the time-domain signals is negligible (e.g., constant-time evolution periods). We have dubbed the algorithm BAMBAM (BAyesian Model Building Algorithm in Multidimensions). BAMBAM will be explained through an example: Bayesian analysis of a synthetic 2D (8 × 8 complex) dataset with added random noise.

Consider a synthetic 2D matrix of NMR data Y, where eight complex points are 'acquired' along each dimension by using the procedure of States et al. (1982) for quadrature detection. This dataset can also be viewed as a linear array of ($N = (2 \times 8)^2 = 256$) time-domain data points, $Y = [y_1, y_2, y_3, \ldots, y_N]$. The time and phase along dimension d for point Y_i are given according to the States quadrature scheme outlined in Table 1. The time-domain dataset Y contains 15 sinusoids whose amplitudes and frequencies along each dimension are tabulated in Table 3. The phases and decay rates of the sinusoids are all zero. The frequencies in Table 3 are expressed in dimensionless units $-1 \le f \le 1$, which represent the angular frequency range $-\pi \le \omega \le \pi$. Y was constructed by calculating the time-domain data points according to the theoretical parameters in Table 3 and then by adding Gaussian noise with a mean of zero and a standard deviation of unity to each of the N data points.

BAMBAM applies a sequential strategy in order to find the most probable model for a given dataset. BAMBAM first calculates the quantity $P(D|f_J, \Theta, I)$ for a model containing no sinusoids ($J = 0$, $M = 1$), i.e. a model which is simply a constant. It then compares the value of $P(D|f_{J=0}, \Theta, I)$ to the posterior probability of a model containing one sinusoid ($J = 1$, $M = J \times 2^D = 1 \times 2^2 = 4$). If

$$\frac{P(D|f_{J=1}, \Theta, I)}{P(D|f_{J=0}, \Theta, I)} > 1,$$

then the most probable model contains at least one sinusoid; otherwise, BAMBAM 'concludes' that the most probable model is simply a constant and the algorithm terminates. In the former case, BAMBAM then calculates $P(D|f_J, \Theta, I)$ for a model containing two sinusoids ($J = 2$, $M = 2 \times 2^2 = 8$). BAMBAM repeats these steps until the termination criteria are met, namely that

$$\frac{P(D|f_{J+1}, \Theta, I)}{P(D|f_J, \Theta, I)} < 1 \quad \text{and} \quad \frac{P(D|f_J, \Theta, I)}{P(D|f_{J-1}, \Theta, I)} > 1.$$

With the synthetic dataset Y as an example, the following cases detail how $P(D|f_J, \Theta, I)$ is calculated for $J = 0, 1, 2, \ldots Z$.

Case (J = 0)

For a model containing no sinusoids, the systematic portion of the data is equivalent to a constant $f(\mathbf{T},\Phi) = A$ (Eq. 13), and the basis function $V_m(\mathbf{T}_i,\Phi_i,\Theta_j)$ (Eqs. 8 and 12) is equal to one. For this minimal model, the statistic $\overline{h^2}$ is equivalent to the square of the sum of the data

$$\overline{h^2} = \left[\sum_{i=1}^{N} y_i\right]^2$$ (Eqs. 16 and 17). The necessary quantities, $N = 256$, $M = 1$, $\overline{h^2}$, and $\overline{d^2}$ (calculated

according to Eq. 15), are known so that $P(\mathbf{D}|f_{J=0},\Theta,\mathbf{I})$ can be calculated according to Eq. 14. Because of the large absolute values of the exponents in Eq. 14, $P(\mathbf{D}|f_{J=0},\Theta,\mathbf{I})$ is more conveniently expressed in logarithmic form

$$\ln[P(\mathbf{D}|f,\Theta,\mathbf{I})] \propto \ln\left[\Gamma\left(\frac{M}{2}\right)\right] + \ln\left[\Gamma\left(\frac{N-M}{2}\right)\right] - \left(\frac{M}{2}\right)\ln\left[\frac{M\overline{h^2}}{2}\right] - \left[\frac{M-N}{2}\right]\ln\left[\frac{N\overline{d^2}-M\overline{h^2}}{2}\right]$$ (23)

Case (J = 1)

The next simplest model is one consisting of a single sinusoid ($J = 1$, $M = 4$). Equation 21 indicates that the most probable single frequency describing a dataset is given by the position of the maximum of the Fourier transform power spectrum ($T_m(W)_{max}$, Eq. 22). The $T_m(W)_{max}$ for

TABLE 3
COMPARISON BETWEEN THE THEORETICAL VALUES OF SIGNAL PARAMETERS AND THE VALUES DERIVED FROM BAYESIAN ESTIMATES OF A SYNTHETIC TWO-DIMENSIONAL DATASET WITH ADDED RANDOM NOISE[a]

Peak	Theoretical values			Bayesian estimates		
	A	ω_1	ω_2	A	ω_1	ω_2
a	3.00	0.4000	−0.7000	3.10	0.4053	−0.6875
b	8.00	−0.6000	−0.6000	7.70	−0.6067	−0.6007
c	5.00	−0.1000	−0.4000	4.95	−0.1321	−0.4026
d	4.50	−0.4000	−0.3000	4.72	−0.3818	−0.3229
e	5.00	0.2000	−0.3000	4.81	0.2035	−0.2822
f	6.00	0.0000	0.0000	6.00	−0.0089	−0.0051
g	2.00	−0.9500	0.0000	2.13	−0.9507	0.0042
h	4.00	0.3000	0.1000	4.70	0.3297	0.1158
i	4.00	0.4000	0.1500	4.65	0.3768	0.1441
j	8.00	−0.7000	0.3000	8.22	−0.7014	0.3074
k	4.00	−0.1000	0.4000	4.04	−0.0954	0.4011
l	6.00	−0.6000	0.4000	6.68	−0.6081	0.4218
m	6.00	−0.8000	0.5000	6.56	−0.7799	0.4883
n	5.00	−0.6000	0.6000	5.84	−0.6025	0.5826
o	3.50	0.4000	0.8000	3.51	0.4040	0.8135

[a] The theoretical time-domain dataset is an 8×8 complex array of data points ($N = (2 \times 8)^2 = 256$) consisting of 15 sinusoids (a–o) computed using the corresponding frequency and amplitude parameters listed in the table. The decay rate and phase along dimensions one and two of each sinusoid are zero. Random noise with a standard deviation of unity was added to each data point prior to Bayesian analysis. The frequencies ω_1 and ω_2 are displayed in the Nyquist interval $-1 \le f \le 1$ representing the angular frequency range $-\pi \le \omega \le \pi$.

dataset $\mathbf{Y}$ occurs at the frequency coordinate $[\omega_1 = -0.663\pi, \omega_2 = 0.313\pi]$. A model consisting of a single 2D sinusoid at this coordinate $[\omega_1, \omega_2]$ can be constructed by Eqs. 7–9, 12 and 13:

$$f_{J=1}(\mathbf{T}_i, \Phi_i) = A_1 \cos(\omega_1 t_{i1} + \phi_{i1})\cos(\omega_2 t_{i2} + \phi_{i2}) + A_2 \cos(\omega_1 t_{i1} + \phi_{i1} + \frac{\pi}{2})\cos(\omega_2 t_{i2} + \phi_{i2}) +$$
$$A_3 \cos(\omega_1 t_{i1} + \phi_{i1})\cos(\omega_2 t_{i2} + \phi_{i2} + \frac{\pi}{2}) + A_4 \cos(\omega_1 t_{i1} + \phi_{i1} + \frac{\pi}{2})\cos(\omega_2 t_{i2} + \phi_{i2} + \frac{\pi}{2}) \tag{24}$$

With the following shorthand:

$$\mathbf{C}_d \equiv \cos(\omega_d t_{id} + \phi_{id}) \tag{25}$$

$$\mathbf{S}_d \equiv \sin(\omega_d t_{id} + \phi_{id}) = \cos(\omega_d t_{id} + \phi_{id} + \frac{\pi}{2}) \tag{26}$$

$f_{J=1}$ can be rewritten as

$$f_{J=1}(\mathbf{T}_i, \Phi_i) = A_1\mathbf{C}_1\mathbf{C}_2 + A_2\mathbf{S}_1\mathbf{C}_2 + A_3\mathbf{C}_1\mathbf{S}_2 + A_4\mathbf{S}_1\mathbf{S}_2 \tag{27}$$

The interaction matrix for the single sinusoidal model (Eq. 19) is given by

$$\mathbf{G} = \begin{bmatrix} \sum_{i=1}^{N} C_1^2 C_2^2 & \sum_{i=1}^{N} C_1 S_1 C_2^2 & \sum_{i=1}^{N} C_1 S_1 C_2^2 & \sum_{i=1}^{N} C_1 S_1 C_2 S_2 \\ \sum_{i=1}^{N} C_1 S_1 C_2^2 & \sum_{i=1}^{N} S_1^2 C_2^2 & \sum_{i=1}^{N} C_1 S_1 C_2 S_2 & \sum_{i=1}^{N} S_1^2 C_2 S_2 \\ \sum_{i=1}^{N} C_1 S_1 C_2^2 & \sum_{i=1}^{N} C_1 S_1 C_2 S_2 & \sum_{i=1}^{N} C_1^2 S_2^2 & \sum_{i=1}^{N} C_1 S_1 S_2^2 \\ \sum_{i=1}^{N} C_1 S_1 C_2 S_2 & \sum_{i=1}^{N} S_1^2 C_2 S_2 & \sum_{i=1}^{N} C_1 S_1 S_2^2 & \sum_{i=1}^{N} S_1^2 S_2^2 \end{bmatrix} = \begin{bmatrix} \frac{N}{4} & 0 & 0 & 0 \\ 0 & \frac{N}{4} & 0 & 0 \\ 0 & 0 & \frac{N}{4} & 0 \\ 0 & 0 & 0 & \frac{N}{4} \end{bmatrix} \tag{28}$$

The eigenvectors and eigenvalues from the diagonal matrix $\mathbf{G}$ were used with the values $[\omega_1, \omega_2]$ to construct the orthogonal basis functions, $\mathbf{H}_{m=1-4}(\mathbf{T}_i, \Phi_i)$ (Eq. 18). The projections of the data onto the orthogonal basis functions $h_{m=1-4}$ and the mean-square value of this projection $\overline{h^2}$ were calculated according to Eqs. 17 and 16, respectively.

For a fixed value of M, $\overline{h^2}$ is a sufficient statistic to determine the values of the nonlinear parameters, ω_1 and ω_2, that represent the most probable frequencies contained in the data. The $\mathbf{T}_m(\mathbf{W})_{max}$ at $[\omega_1 = -0.663\pi, \omega_2 = 0.313\pi]$ was used as an estimate for the frequencies that maximize $\overline{h^2}$. The precision of this estimate was limited, however, by the size of the frequency grid over which the FT was performed. A more precise estimate was obtained from a two-parameter nonlinear maximization algorithm using $\overline{h^2}$ as the criterion for maximization. The nonlinear maximization yielded an $\overline{h^2}_{max}$ at the frequency coordinate $[\omega_1 = -0.673\pi, \omega_2 = 0.307\pi]$. The amplitude values $A_{m=1-4}$ consistent with this model were calculated from $h_{m=1-4}$ according to Eq. 20.

The values of $\overline{h^2}_{max}$, $\overline{d^2}$, M = 4 and N = 256 were used to calculate $\ln[P(\mathbf{D}|f_{J=1}, \Theta, \mathbf{I})]$, the logarithm of the posterior probability of the data, given a model containing a single sinusoid. A function $\Delta(J)$, defined as
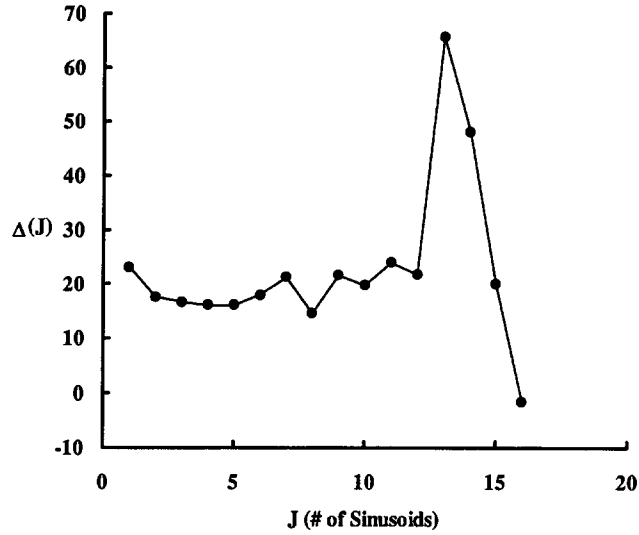
Fig. 1. Plot of the natural logarithm of the change in posterior probability of model J ($\Delta(J)$) vs. J. The expression for $\Delta(J)$ is given by Eq. 29 in the text. The sign of $\Delta(J)$ indicates the relative probability of model f_J compared to model f_{J-1}, positive if f_J is more probable than f_{J-1} and negative if f_J is less probable than f_{J-1}.

$$\Delta(J) = \ln[P(D|f_J,\Theta,I) - \ln[P(D|f_{J-1},\Theta,I) \tag{29}$$

indicates whether and to what extent a model containing J sinusoids is more or less probable than a model containing $(J-1)$ sinusoids. For the synthetic dataset **Y**, $\Delta(1) \approx 24$. This statistic indicates that the model containing a single sinusoid is e^{24} times more probable than the simple constant model in accounting for the systematic portion of **Y**. $\Delta(1)$ is thus strong evidence that the most probable model of **Y** contains at least one sinusoid.

Case (J = 2)

For a model containing J = 2 sinusoids, there are two sets of 2D frequency coordinates, $[\omega_{11},\omega_{12}]$, $[\omega_{21},\omega_{22}]$ and M = 8 amplitudes

$$f_{J=2}(T_i,\Phi_i) = A_1C_{11}C_{12} + A_2S_{11}C_{12} + A_3C_{11}S_{12} + A_4S_{11}S_{12} + \\ A_5C_{21}C_{22} + A_6S_{21}C_{22} + A_7C_{21}S_{22} + A_8S_{21}S_{22} \tag{30}$$

where

$$C_{jd} \equiv \cos(\omega_{jd}t_{id} + \phi_{id}) \tag{31}$$

$$S_{jd} \equiv \sin(\omega_{jd}t_{id} + \phi_{id}) = \cos(\omega_{jd}t_{id} + \phi_{id} + \frac{\pi}{2}) \tag{32}$$

The interaction matrix **G** for $f_{J=2}$ can be simplified to (terms below the diagonal are omitted because the matrix is symmetrical):

$$G = \begin{bmatrix}
\dfrac{N}{4} & 0 & 0 & 0 & \sum_{i=1}^{N} C_{11}C_{12}C_{21}C_{22} & \sum_{i=1}^{N} C_{11}C_{12}S_{21}C_{22} & \sum_{i=1}^{N} C_{11}C_{12}C_{21}S_{22} & \sum_{i=1}^{N} C_{11}C_{12}S_{21}S_{22} \\[2ex]
 & \dfrac{N}{4} & 0 & 0 & \sum_{i=1}^{N} S_{11}C_{12}C_{21}C_{22} & \sum_{i=1}^{N} S_{11}C_{12}S_{21}C_{22} & \sum_{i=1}^{N} S_{11}C_{12}C_{21}S_{22} & \sum_{i=1}^{N} S_{11}C_{12}S_{21}S_{22} \\[2ex]
 & & \dfrac{N}{4} & 0 & \sum_{i=1}^{N} C_{11}S_{12}C_{21}C_{22} & \sum_{i=1}^{N} C_{11}S_{12}S_{21}C_{22} & \sum_{i=1}^{N} C_{11}S_{12}C_{21}S_{22} & \sum_{i=1}^{N} C_{11}S_{12}S_{21}S_{22} \\[2ex]
 & & & \dfrac{N}{4} & \sum_{i=1}^{N} S_{11}S_{12}C_{21}C_{22} & \sum_{i=1}^{N} S_{11}S_{12}S_{21}C_{22} & \sum_{i=1}^{N} S_{11}S_{12}C_{21}S_{22} & \sum_{i=1}^{N} S_{11}S_{12}S_{21}S_{22} \\[2ex]
 & & & & \dfrac{N}{4} & 0 & 0 & 0 \\[2ex]
 & & & & & \dfrac{N}{4} & 0 & 0 \\[2ex]
 & & & & & & \dfrac{N}{4} & 0 \\[2ex]
 & & & & & & & \dfrac{N}{4}
\end{bmatrix} \qquad (33)$$

For a general model, the summations shown in Eq. 33 cannot be evaluated analytically; the summations are dependent upon the specific values of $[\omega_{11},\omega_{12}]$ and $[\omega_{21},\omega_{22}]$.

An estimate for $[\omega_{11},\omega_{12}]$ was obtained from the $f_{J=1}$ model, and an estimate for $[\omega_{21},\omega_{22}]$ was generated from the residual $R(2)$ between the original dataset and $f_{J=1}$. For a model consisting of J frequencies, the residual $R(J)$ is defined as

$$R(J) = Y - f_{J-1} \qquad (34)$$

The estimate $[\omega_{21} = -0.606\pi, \omega_{22} = -0.596\pi]$ was obtained from the $T_m(W)_{max}$ of the residual $R(2)$. A starting estimate for the nonlinear parameters comprising the most probable two-sinusoid model was thus given by the pair of coordinates $[\omega_{11} = -0.673\pi, \omega_{12} = -0.307\pi]$ and $[\omega_{21} = -0.606\pi, \omega_{22} = -0.596\pi]$. The eigenvectors and eigenvalues derived from G were used to calculate values for $H_{m=1-8}$, $h_{m=1-8}$ and $\overline{h^2}$ as done previously (Eqs. 16–18). The quantity $\overline{h^2}_{max}$ was found by a four-parameter nonlinear maximization of $[\omega_{11},\omega_{12}]$ and $[\omega_{21},\omega_{22}]$, which yielded $[\omega_{11} = -0.674\pi, \omega_{12} = -0.309\pi]$ and $[\omega_{21} = -0.598\pi, \omega_{22} = -0.587\pi]$. Values of $A_{m=1-8}$ consistent with this model were calculated according to Eq. 20. The value $\Delta(2)$ (Eq. 29) indicated that the model consisting of two sinusoids was e^{19} times more probable than the model consisting of a single sinusoid.

Case (J = Z)

BAMBAM determines the most probable values for the nonlinear parameters $([\omega_{11},\omega_{12}],[\omega_{21},\omega_{22}], ...[\omega_{Z1},\omega_{Z2}])$ for the case $J = Z$ in the same manner as the case $J = 2$. Each

model consists of Z 2D frequency coordinates which are used to define a basis set of $M = 4Z$ amplitudes and $4Z$ basis functions (Eqs. 7–9). The $T_m(W)_{max}$ value of the residual $R(Z)$ is used to obtain a starting frequency estimate for $[\omega_{Z1}, \omega_{Z2}]$, and the frequency estimates from $f_{J = 1-Z}$ yield starting values for $([\omega_{11}, \omega_{12}], [\omega_{21}, \omega_{22}], \ldots [\omega_{(Z-1)1}, \omega_{(Z-1)2}])$. From these nonlinear parameters, the interaction matrix G can be calculated numerically, and its eigenvectors and eigenvalues determine $H_{m = 1-M}$, $h_{m = 1-M}$, and $\overline{h^2}$ according to Eqs. 16–18. The search for $\overline{h^2}_{max}$ is accomplished through a $2Z$ nonlinear parameter maximization (see Appendix: Nonlinear Maximization of $\overline{h^2}$). The values for $A_{m = 1-M}$ consistent with the most probable nonlinear parameters (the frequencies $[\omega_{11}, \omega_{12}], [\omega_{21}, \omega_{22}], \ldots [\omega_{Z1}, \omega_{Z2}]$ which maximize $\overline{h^2}$) are calculated according to Eq. 20. The values obtained for $([\omega_{11}, \omega_{12}], [\omega_{21}, \omega_{22}], \ldots [\omega_{Z1}, \omega_{Z2}])$ and $A_{m = 1-M}$ define the new model $f_{J = z}$ according to Eqs. 7–9. The quantity $\Delta(Z)$ then determines the relative probabilities of models f_z and $f_{z + 1}$.

RESULTS

Synthetic dataset

The most probable model of dataset Y found by Bayesian analysis is presented in Table 3 next to the actual values that define the systematic portion of Y. The algorithm accurately predicted the number of sinusoids contained within the data. Figure 1 displays a plot of $\Delta(J)$ vs. J for $J = 1-16$. Bayesian statistics indicated that the 15-sinusoid model was e^{31} times more probable than the 14-sinusoid model and $e^{3.7}$ times more probable than the 16-sinusoid model. Although the 16-sinusoid model 'fits' the data better, the better fit was statistically unjustified by the greater number of basis functions required by the model. The inequality $\Delta(16) < 0$ (see Fig. 1) was the hallmark of the greater likelihood of the 15- vs. 16-sinusoid model.

In order to estimate the precision and the accuracy of each of the Bayesian-derived parameters in Table 3, identical analyses were performed upon a total of eight datasets. Each dataset was identical to dataset Y in its systematic component (signals a–o) but different in its random (noise) component. The accuracy and precision determined from the Bayesian estimates of the eight datasets are displayed in Table 4. The terms 'precision' and 'accuracy' as used here are defined as the root-mean-square (rms) deviation of each estimate from the mean value (of the eight datasets) and the 'true' value (left-hand portion of Table 3) respectively. The greater accuracy of the frequency vs. the amplitude estimates is not unexpected because the frequencies are much better determined by the data than are the amplitudes (Bretthorst, 1990c). The results of Tables 3 and 4 show clearly that, even though a zero-filled Fourier transform of dataset yields a frequency resolution of $(\dfrac{2\pi}{4 \times 8} = 0.0625\pi)$, Bayesian analysis was able to estimate the frequencies contained in dataset Y within an average accuracy of $0.013\pi \le \omega \le 0.018\pi$. Clearly, the actual frequency information contained in a highly-truncated dataset is much greater than is conventionally obtained from a zero-filled Fourier transform.

A graphical comparison of Bayesian analysis with the zero-filled Fourier transform (ZFT) and with the ZFT enhanced by 1D mirror-image linear prediction is shown in Fig. 2. Contour plots of the real frequency-domain spectrum of a noise-free (black contours, 64×64 complex) and a truncated (red contours, 8×8 complex) noise-added version of the same dataset are displayed in the figure. The noise-added, truncated dataset was processed by three different methods: (A) apodization and zero-filling, followed by Fourier transformation; (B) extrapolation of each dimension by separate 1D linear prediction prior to apodization, zero-filling and Fourier trans-

formation; and (C) extrapolation of both dimensions simultaneously by 2D Bayesian analysis prior to apodization, zero-filling and Fourier transformation (see legend to Fig. 2 for additional details). Processing of dataset Y with the simple zero-filled Fourier transform (Fig. 2A) resolved only 9 of the 15 actual frequencies contained within the data. Moreover, severe overlap distorted the positions of many of the frequencies and produced a large amplitude at ($\omega_1 \cong -\frac{\pi}{3}$, $\omega_2 \cong 0$) where no frequency actually exists. Extrapolation of the time-domain data by linear prediction (Fig. 2B) significantly improved the quality of the spectrum but still resolved only 12 of the 15 frequencies contained within the data. Extrapolation of the time-domain data by Bayesian analysis (Fig. 2C) prior to FT produced a real frequency-domain spectrum with all the frequencies of sinusoids a–o fully resolved within a 4% error. The amplitudes of signals h and i were only partially resolved by Bayesian analysis. They tended to be systematically overestimated in each of the eight datasets. This can be seen from the high accuracy-to-precision ratio for these two signals (Table 4).

Analysis of experimental data

To illustrate the applicability of Bayesian analysis to the processing of constant-time data in 3D and 4D experimental datasets, BAMBAM was used to extrapolate the 2D $H^\alpha C^\alpha$ planes from a 4D

TABLE 4
THE CALCULATED PRECISION AND ACCURACY DERIVED FROM BAYESIAN ESTIMATES OF A SERIES OF EIGHT SYNTHETIC TWO-DIMENSIONAL DATASETS WITH ADDED RANDOM NOISE[a]

Peak	Accuracy			Precision		
	A	ω_1	ω_2	A	ω_1	ω_2
a	0.13	0.0076	0.0133	0.13	0.0067	0.0086
b	0.15	0.0053	0.0056	0.11	0.0053	0.0054
c	0.39	0.0284	0.0059	0.31	0.0085	0.0059
d	0.30	0.0307	0.0147	0.24	0.0166	0.0066
e	0.17	0.0069	0.0169	0.12	0.0068	0.0038
f	0.16	0.0124	0.0077	0.14	0.0124	0.0040
g	0.14	0.0151	0.0088	0.12	0.0119	0.0035
h	1.87	0.0355	0.0192	1.24	0.0265	0.0187
i	1.60	0.0259	0.0118	0.80	0.0062	0.0054
j	0.45	0.0039	0.0105	0.28	0.0029	0.0057
k	0.16	0.0046	0.0042	0.16	0.0046	0.0041
l	0.89	0.0116	0.0127	0.54	0.0079	0.0080
m	0.49	0.0131	0.0093	0.29	0.0082	0.0074
n	0.71	0.0099	0.0211	0.40	0.0075	0.0139
o	0.13	0.0094	0.0170	0.08	0.0033	0.0114
ALL	0.74	0.0177	0.0129	0.45	0.0107	0.0085

[a] Each theoretical time-domain dataset is an 8×8 complex array of data points ($N = (2 \times 8)^2 = 256$) consisting of 15 sinusoids (a–o) computed using the corresponding frequency and amplitude parameters listed in Table 3. Random noise with a standard deviation of unity was added to each data point prior to Bayesian analysis. The eight datasets are thus identical with respect to their systematic components and differ only in their random (noise) components. The frequencies ω_1 and ω_2 are displayed in the Nyquist interval $-1 \le f \le 1$ representing the angular frequency range $-\pi \le \omega \le \pi$.
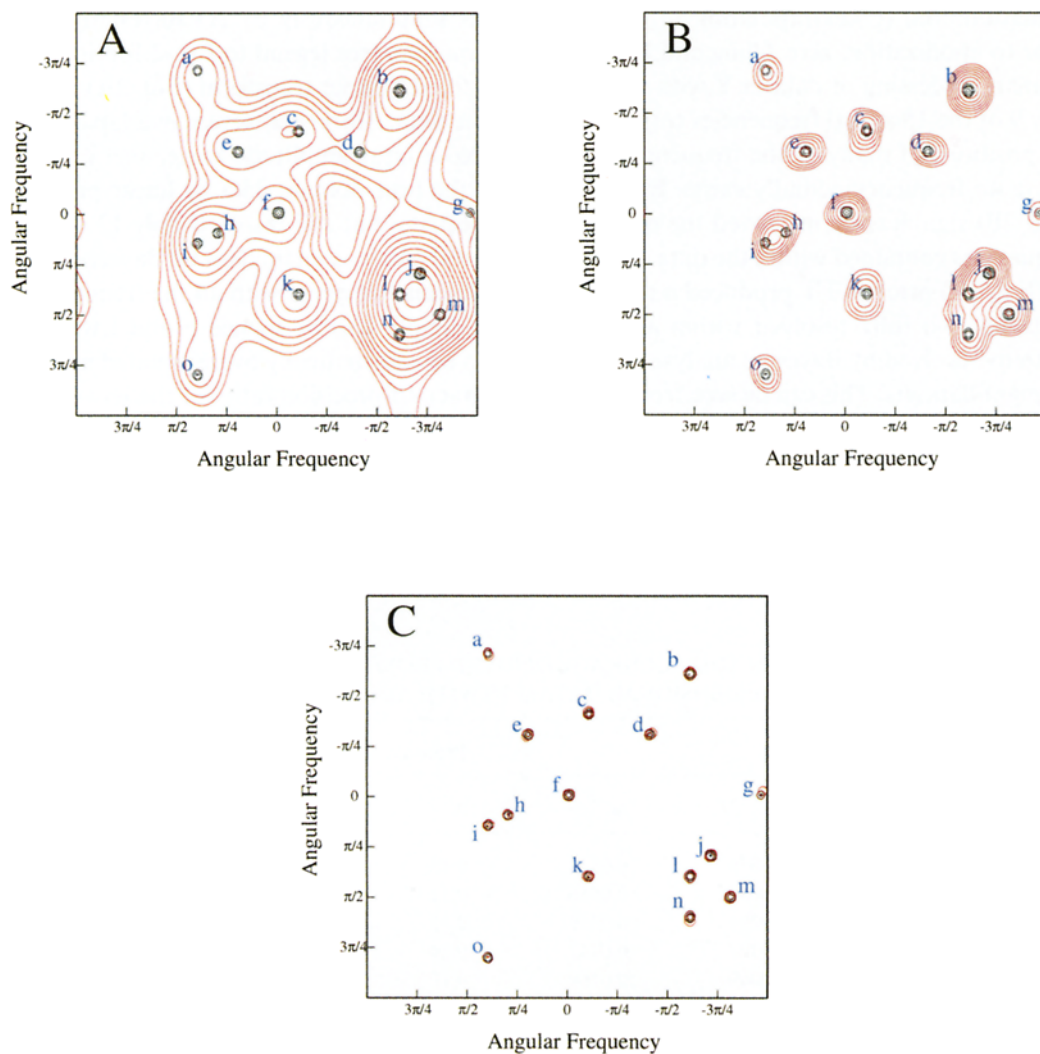
528

Fig. 2. Contour plots of the real frequency-domain spectrum of a noise-free (black contours) and a truncated, noise-added (red contours) version of the dataset defined by Table 3. The systematic portion of the dataset consisted of 15 sinusoids (a–o) with an amplitude range of 2–8. The noise-added dataset (8×8 complex points) was processed in three different ways prior to Fourier transformation: (A) Prior to Fourier transformation, each dimension was digitally filtered by a cosine-squared bell apodization function followed by zero-filling from 8 to 128 complex points. After Fourier transformation, the imaginary portion of the spectrum along each dimension was discarded; (B) Time domain t_2 was Fourier transformed. Mirror-image linear prediction was used to extend the length of the FID along t_1 from 8 to 15 complex points. Time domain t_1 was then digitally filtered with a cosine-squared bell apodization function and zero-filled from 15 to 128 complex points prior to Fourier transformation. After Fourier transformation, the imaginary portion of ω_1 was discarded. Frequency domain ω_2 was then inverse Fourier transformed and processed in the same manner as t_1; (C) The best estimates of the number, amplitude, frequency and phases of all detectable signals in the data were obtained by using the algorithm described in the text. The model derived from this Bayesian analysis was used to perform a 2D extrapolation of the (t_1,t_2) 8×8 complex dataset to a 64×64 complex array. Each dimension was then digitally filtered with a cosine-squared bell apodization function and zero-filled from 64 to 128 complex points prior to Fourier transformation. After Fourier transformation, the imaginary portion of the spectrum along each dimension was discarded.

HN(CO)CAHA experiment (Kay et al., 1992) performed upon uniformly $^{15}N-^{13}C$-labeled protein subunit c (DDT binding protein, 8 kDa) from E. coli. Although subunit c is a relatively small protein, it consists of predominantly α-helical regions and thus exhibits poor chemical-shift dispersion in both the backbone and side-chain NMR resonances. The dataset was collected on a Bruker 500 MHz spectrometer with a frequency resolution of ($256 \times 16 \times 8 \times 8$) complex points in the amide proton (acquisition, t_4), nitrogen (t_1), α-proton (t_2) and α-carbon (t_3) dimensions, respectively. The H^α and C^α (t_2,t_3) dimensions were acquired by using shared constant-time acquisition periods (Kay et al., 1992).

The acquisition dimension t_4 was processed using conventional apodization, zero-padding and Fourier transformation, followed by discarding of the imaginary portion of the frequency-domain data. The nitrogen t_1 dimension was then extrapolated from 16 to 24 complex points by linear prediction after Fourier transformation of the t_2 and t_3 dimensions. After apodization with a cosine-squared bell filter function, zero-filling to 64 complex points, Fourier transformation and discarding of the imaginary data points along the ω_1 dimension, the ω_2 and ω_3 dimensions were inverse Fourier transformed. The result of this series of operations was a 4D matrix consisting of 256 (ω_4) × 64 (ω_1) = 16 384 planes of data, each containing 8 × 8 complex points along dimensions t_2 and t_3 (H^α and C^α, respectively).

Each of the 256 × 64 (t_2,t_3) 2D FIDs was extrapolated from 8 × 8 to 32 × 32 complex points using multidimensional Bayesian analysis. The total processing time required to analyse the 16 384 planes was 23.1 h on a Silicon Graphics Indigo workstation. After extrapolation by Bayesian analysis, each of the (t_2,t_3) planes was apodized, zero-filled to 64 complex points and Fourier transformed. A single plane from the 4D dataset occurring at [ω_4 = 6.32 ppm, ω_1 = 120.6 ppm] was processed according to the three different methods (A, B and C) as described above, and the results are shown in Figs. 3A–C.

CONCLUSIONS

Bayesian probability theory applied to the analysis of stationary, multidimensional NMR data is a practical technique for extrapolating D-dimensional constant-time data prior to apodization, zero-filling and Fourier transformation. Bayesian analysis of constant-time data can resolve frequencies that conventional processing techniques and 1D mirror-image linear prediction cannot. Although a portion of the analysis requires a nonlinear optimization step, the optimized parameters interact only weakly with each other (partially orthogonal), and the initial values of the parameters are close to their optimal values, thus allowing stable and rapid convergence of the algorithm. Future work will be directed toward the addition of decay rates to the Bayesian models. This development would remove the constant-time constraint and allow simultaneous Bayesian analysis of all indirectly detected dimensions of any NMR experiment that can be modeled by the sum of simple D-dimensional decaying sinusoids. In the near future, however, computational limitations probably will prohibit simultaneous modeling of all dimensions in multidimensional NMR experiments of medium-sized and large proteins. The vast number of signals present in an entire dataset is too large to model according to the strategy outlined here. New approximations and a different approach will be required to reduce the size of the interaction matrices and the number of required Fourier transforms. If these barriers can be surmounted, complete estimation of frequencies, amplitudes, phases and decay rates of multidimensional
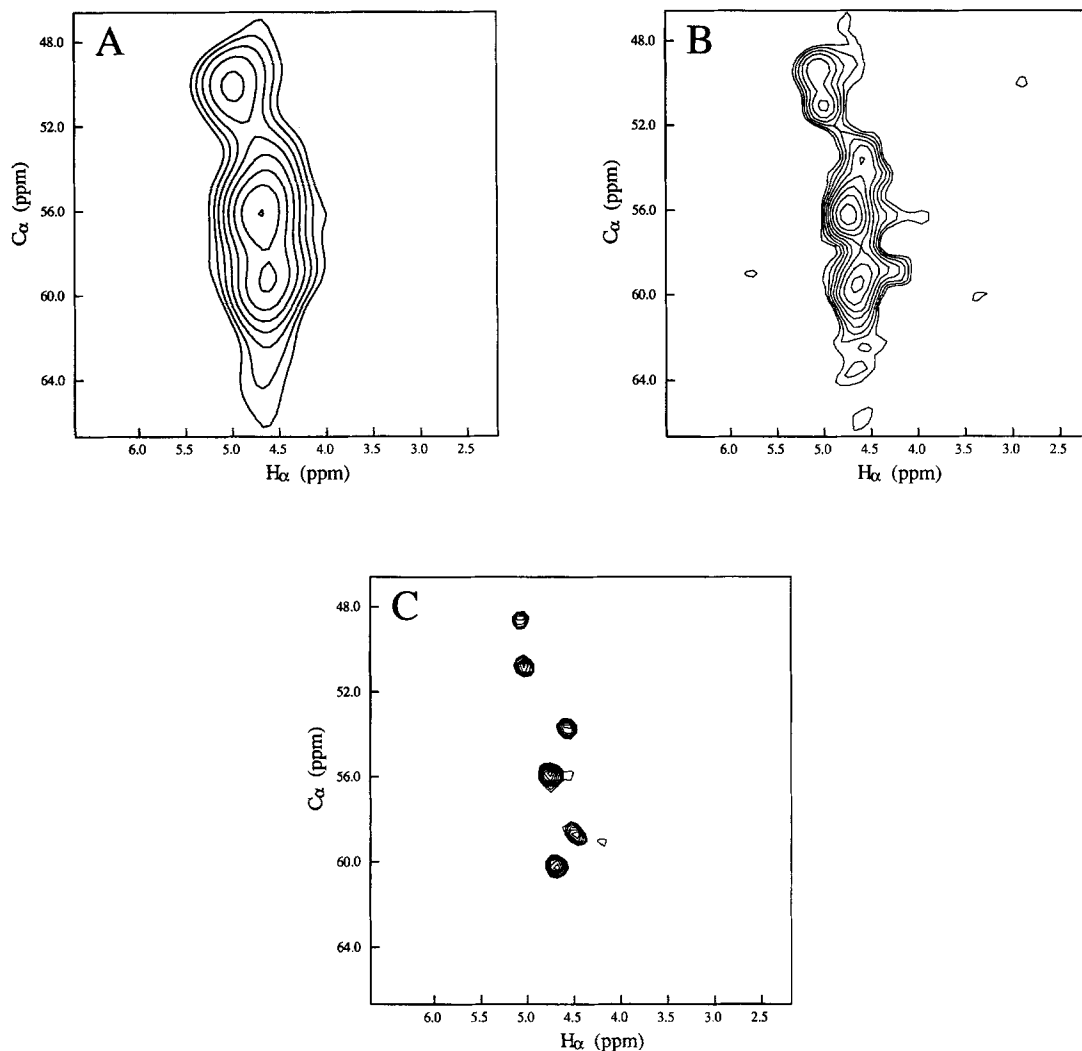
Fig. 3. Contour plots of the Fourier-transformed spectra (real portion) of the 2D $H^\alpha C^\alpha$ (ω_2, ω_3) plane at [ω_4 (H) = 6.32 ppm, ω_1 (N) = 120.59 ppm] from the 4D HN(CO)$C^\alpha H^\alpha$ experiment performed upon the ^{15}N-^{13}C uniformly labeled protein subunit c (DDT binding protein) from $E.$ $coli.$ The t_1 and t_4 (acquisition) dimensions were processed according to conventional methods described in the text. The 2D $H^\alpha C^\alpha$ (t_2, t_3) plane defined by [ω_4 (H) = 6.32 ppm, ω_1 (N) = 120.59 ppm] was then copied from the 4D matrix and processed according to three different methods: (A) Prior to Fourier transformation, each dimension was digitally filtered by a cosine-squared bell apodization function followed by zero-filling from 8 to 64 complex points. (B) Time dimension t_2 was Fourier transformed. Mirror-image linear prediction was used to extend the length of the FID along t_3 from 8 to 15 complex points. Time dimension t_3 was then digitally filtered with a cosine-squared bell apodization function and zero-filled from 15 to 64 complex points prior to Fourier transformation. Frequency dimension ω_2 was then inverse Fourier transformed and processed in the same manner as t_3. (C) The best estimates of the number, amplitude, frequency and phases of all detectable signals in the data were obtained by using the algorithm described in the text. The model derived from this Bayesian analysis was used to perform a 2D extrapolation of the (t_2, t_3) 8 × 8 complex dataset to a 32 × 32 complex array. Each dimension (t_2, t_3) was then digitally filtered with a cosine-squared bell apodization function and zero-filled from 32 to 64 complex points prior to Fourier transformation.

NMR data could be accomplished by Bayesian analysis with greater accuracy than by conventional processing techniques. In addition, the direct extraction of relevant NMR parameters from the data and the inherent ability of Bayesian analysis to perform signal recognition would facilitate automated approaches to the process of resonance assignment.

ACKNOWLEDGEMENTS

REFERENCES

Barkhuijsen, J., De Beer, W.M., Bovee, W.M.M.J. and Van Ormondt, D. (1985) *J. Magn. Reson.*, **61**, 465–481.
Bax, A., Mehlkopf, A.F. and Smidt, J. (1979) *J. Magn. Reson.*, **35**, 167–173.
Bodenhausen, G., Vold, R.L. and Vold, R.R. (1980) *J. Magn. Reson.*, **37**, 93–106.
Bretthorst, G.L. (1990a) *J. Magn. Reson.*, **88**, 533–551.
Bretthorst, G.L. (1990b) *J. Magn. Reson.*, **88**, 552–570.
Bretthorst, G.L. (1990c) *J. Magn. Reson.*, **88**, 571–595.
Chylla, R.A. and Markley, J.L. (1993) *J. Cell. Biochem.*, suppl. **17C**, 250 (Abs. L7-103).
Gesmar, H., Led, J.J. and Abildgaard, F. (1990) *Prog. NMR Spectrosc.*, **22**, 255–288.
Gesmar, H. and Led, J.J. (1989) *J. Magn. Reson.*, **83**, 53–64.
Kay, L.E., Wittekind, M., McCoy, M.A., Friedrichs, M.S. and Mueller, L. (1992) *J. Magn. Reson.*, **98**, 443–450.
Kessler, H., Griesinger, C., Zarbock, J. and Loosli, H.R. (1984) *J. Magn. Reson.*, **57**, 331–336.
Kotyk, J.J., Hoffman, N.G., Hutton, W.C., Bretthorst, G.L. and Ackerman, J.J.H. (1992) *J. Magn. Reson.*, **98**, 483–500.
Kumaresan, R. and Tufts, D.W. (1982) *Ieee Trans. Acoust. Speech Sign. Proces.*, **30**, 833–840.
Led, J.J. and Gesmar, H. (1991) *J. Biomol. NMR*, **1**, 237–246.
Redfield, A.G. and Kunz, S.D. (1975) *J. Magn. Reson.*, **19**, 250–254.
States, D.J., Haberkorn, R.A. and Ruben, D.J. (1982) *J. Magn. Reson.*, **48**, 286–292.
Stephenson, D.S. (1988) *Prog. NMR Spectrosc.*, **20**, 515–595.
Zhu, G. and Bax, A. (1992) *J. Magn. Reson.*, **98**, 192–199.
Zhu, G. and Bax, A. (1990) *J. Magn. Reson.*, **90**, 405–410.

APPENDIX

Nonlinear maximization of $\overline{h^2}$

For a D-dimensional dataset **Y** modeled by a fixed number of sinusoids (J), $\overline{h^2}$ is dependent upon the nonlinear parameters that comprise the basis functions. If the time-domain data are stationary, then the nonlinear parameters for a D-dimensional model containing J sinusoids are the set of frequencies ($[\omega_{11},\omega_{12},...,\omega_{1D}]$, $[\omega_{21},\omega_{22},...,\omega_{2D}]$, ..., $[\omega_{J1},\omega_{J2},...,\omega_{JD}]$). The quantity $\overline{h^2}_{max}$ is thus obtained from an optimization of the nonlinear parameters ($[\omega_{11},...\omega_{12},...,\omega_{1D}]$, $[\omega_{21},\omega_{22},...,\omega_{2D}]$, ..., $[\omega_{J1},\omega_{J2},...,\omega_{JD}]$).

The requirements for the nonlinear optimization are not rigorous. The parameters are nearly orthogonal, and the initial values of the parameters are close to their optimal values. The nonlinear optimization is performed as a series of iterations. For each iteration, an increment $\Delta\omega_{jd}$ is defined for each parameter ω_{jd}, the frequency of the jth sinusoid along dimension d. The value of $\Delta\omega_{jd}$ is set to $\frac{\pi}{n_d}$, where n_d is the number of data points sampled along dimension d. The value of $\overline{h}^2$ is calculated from the interaction matrix $\mathbf{G}$ (Eqs. 17–19) at three different points ($\omega_{jd} - \Delta\omega_{jd}$, ω_{jd}, and $\omega_{jd} + \Delta\omega_{jd}$) corresponding to ($\overline{h}^2_1$, $\overline{h}^2_2$, and $\overline{h}^2_3$), respectively. A new value for ω_{jd} is computed according to

$$
\omega_{jd} = \begin{cases} \omega_{jd} - \frac{3}{4}\Delta\omega_{jd}, & \overline{h}^2_1 > \overline{h}^2_2, \overline{h}^2_3 \\[2mm] \Delta\omega_{jd}\left[\dfrac{\sqrt{\overline{h}^2_1} - \sqrt{\overline{h}^2_3}}{2\sqrt{\overline{h}^2_1} + 4\sqrt{\overline{h}^2_2} + 2\sqrt{\overline{h}^2_3}}\right], & \overline{h}^2_2 > \overline{h}^2_1, \overline{h}^2_3 \\[2mm] \omega_{jd} + \frac{3}{4}\Delta\omega_{jd}, & \overline{h}^2_3 > \overline{h}^2_2, \overline{h}^2_1 \end{cases} \tag{A1}
$$

The middle term of Eq. A1 is a parabolic interpolation of ω_{jd} from the magnitude of $\Delta\omega_{jd}$ and the relative values $\overline{h}^2_1$, $\overline{h}^2_2$, and $\overline{h}^2_3$. If the middle condition $\overline{h}^2_2 > \overline{h}^2_1, \overline{h}^2_3$ is met, then the search for ω_{jd} is complete and ω_{jd} is replaced with its interpolated value. If either the first or third condition is true, ω_{jd} is set to ($\omega_{jd} - \frac{3}{4}\Delta\omega_{jd}$) or ($\omega_{jd} + \frac{3}{4}\Delta\omega_{jd}$), respectively, and the search for an optimum value of ω_{jd} will continue during the following iterations. The algorithm is complete when every nonlinear parameter ω_{jd} has met the second condition of Eq. A1.

Orthogonal approximation

For a D-dimensional model containing J sinusoids, there are JD frequency parameters and $M = JK$ ($K = 2^D$) basis functions. If all frequency parameters were optimized for a D = 2 model containing J = 15 sinusoids ($M = 15(2^2) = 60$), the size of the interaction matrix would be 60×60. For such large numbers of parameters and basis functions, the time required to calculate $\mathbf{G}$ and its eigenvectors and eigenvalues would be prohibitive. In practice, the orthogonality of the basis functions can be used to limit the number of frequencies and basis functions requiring optimization.

The orthogonality between two basis functions $\mathbf{V}_l$ and $\mathbf{V}_m$ can be quantified according to a parameter σ_{lm} given by

$$
\sigma_{lm} \equiv \frac{2g_{lm}}{(g_{ll} + g_{lm})} \tag{A2}
$$

where g_{lm} is defined according to Eq. 20. If $\sigma_{lm} = 0$, then $\mathbf{V}_l$ and $\mathbf{V}_m$ are completely orthogonal and the nonlinear parameters associated with each function can be optimized independently. In a model f_j containing $j = 1 \ldots J$ sinusoids, each sinusoid S_j is associated with K basis functions $[\mathbf{V}_{(j-1)k+1}, \ldots, \mathbf{V}_{jK}]$. A maximum interaction $\overline{\sigma}_{jk}$ can be defined between two sinusoids, S_j and S_k, which represents the maximum of the K^2 interactions between each of their associated basis functions.

$$
\overline{\sigma}_{jk} \equiv \left[\bigcup_{l=1^\circ}^{jK} \bigcup_{m=m^\circ}^{kK} \overline{\sigma}_{lm}\right]_{max} \quad 1^\circ \equiv (j-1)K + 1, \; m^\circ \equiv (k-1)K + 1 \tag{A3}
$$

In the BAMBAM algorithm, two sinusoids S_j and S_k are considered to be orthogonal to one another if $\bar{\sigma}_{jk} < 0.025$.

This orthogonality property allows BAMBAM to reduce the number of nonlinear parameters requiring optimization. For a given model containing J sinusoids $S_j = S_1, S_2, \ldots S_J$, all of the sinusoids that are orthogonal to a particular sinusoid, S_j, are eliminated from optimization of S_j. All of the nonlinear parameters associated with S_j are then simultaneously optimized on the nonorthogonal set of sinusoids. This is done for all J sinusoids until the middle condition of Eq. A1 is met on the first iteration for all JD frequency parameters.