BAYESIAN SPECTRUM ANALYSIS ON QUADRATURE NMR DATA WITH NOISE CORRELATIONS

G. LARRY BRETTHORST
Department of Chemistry
Campus Box 1134
Washington University
1 Brookings Drive
St. Louis, MO 63130

Abstract. In NMR data analysis a great deal of prior information is available. We know, in general terms, what characteristic signal will be received, that for quadrature measurements it will be the same in both channels and that the noise is potentially correlated. We have shown in previous work [1], [2] that when prior information is incorporated into the analysis of data, the frequencies, decay rates, and amplitudes may be estimated much more precisely than by using the discrete Fourier transform directly. Here we extend the Bayesian analysis to include the quadrature nature of the data and noise correlations. We then show that in typical NMR data the frequencies and decay rates may be estimated with a precision several orders of magnitude better than directly from the discrete Fourier transform.

Introduction

In NMR, theory tells us that the free induction decay time series must be sinusoidal with exponential or Gaussian decay. When this information is incorporated into the spectral estimation problem, one may estimate the frequencies and decay rates much more accurately than directly from a discrete Fourier transform of the data [1], [2]. More importantly this information allows one to separate frequencies and decay rates that are too close for one to resolve using a discrete Fourier transform. The initial work [1] did not incorporate all of the information we possessed about NMR signals. We used the functional form of the signal, but we utilized the data as if two distinct measurements were available having the same frequencies and decay rates, but different amplitudes and phases. This gave $\sqrt{2}$ improvement in the parameter estimates. However, we have more information; in particular, we know that the signal in the second channel is 90° out of phase with that in the first channel. Also, we know that the noise is potentially correlated, and that the phases of all the sinusoids are typically the same. When more information is incorporated into a probability calculation, we expect that information to improve the estimates of the parameters. In this paper we specialize the Bayesian calculation to include quadrature, noise correlations, and phase coherences.

The General Model Equation

The basic model we are considering is: given a quadrature detected data set (*i.e.,* two data sets collected with a 90° phase difference), then the real data may be modeled as

$$d_R(t_i) = f_R(t_i) + n(\sigma, 0)$$

where $n(\sigma, 0)$ is a Gaussian noise component of mean zero and standard deviation σ, $f_R(t)$ is a model of the real signal, and the quadrature or imaginary data may be modeled as

$$d_I(t_i) = f_I(t_i) + n(\sigma, 0).$$

The basic problem we would like to solve is: "what are the best estimates of the parameters (frequencies and decay rates) hidden in f_R and f_I that one can make from the data and the prior information?" We will solve this problem using Bayesian probability theory and apply the calculation to several examples.

We write the model equations $f_R(t)$ and $f_I(t)$ as a sum over functions G_j and F_j such that

$$f_R(t) = \sum_{j=1}^{m} B_j G_j(\{\omega\}, t) \quad \text{and} \quad f_I(t) = \sum_{j=1}^{m} B_j F_j(\{\omega\}, t) \tag{1}$$

where m is the total number of model functions, B_j is the amplitude of the jth model function, and $G_j(\{\omega\}, t)$ and $F_j(\{\omega\}, t)$ are typically sinusoids with either exponential or Gaussian decay. The model functions F_j and G_j are functions of a continuous variable time t; however, the data have been sampled at discrete times $\{t_1, \cdots, t_N\}$. Additionally, the models are functions of other continuous parameters, which we collectively label $\{\omega\}$. These parameters are frequencies, decay rates or any other parameters which could be needed to model the data, for example the phase if it is the same on all of the sinusoids. Although the amplitudes $\{B\}$ are of substantial interest, for the purposes of analyzing the data, we wish to formulate the problem independently of these parameters to see what probability theory can tell us about the frequencies and decay rates. The quadrature information has been incorporated by assuming the amplitudes B_j are the same in both channel.

We would like to compute the posterior probability of the frequencies and decay rates, given the data D and the prior information I. This requires us to obtain two terms: the direct probability of the data and the prior probability of the parameters. We will compute the direct probability of the data first. Making the standard assumptions about the noise, the direct probability of the data is:

$$\begin{aligned} P(D|\{B\},\{\omega\},\sigma,\rho,I) &= (2\pi\sigma^2)^{-N}(1-\rho^2)^{-\frac{N}{2}} \\ &\times \exp\left\{-\sum_{i=1}^{N}\frac{[d_R(t_i)-f_R(t_i)]^2+[d_I(t-i)-f_I(t_i)]^2}{2\sigma^2(1-\rho^2)}\right\} \\ &\times \exp\left\{-2\rho\sum_{i=1}^{N}\frac{[d_R(t_i)-f_R(t_i)][d_I(t_i)-f_I(t_i)]}{2\sigma^2(1-\rho^2)}\right\} \end{aligned}$$

where ρ is the correlation coefficient – see Jeffreys [3] for a discussion of correlation, and [2], [4] for a discussion of when a Gaussian should be used to represent the noise. The symbol I in $P(D|\{B\},\{\omega\},\sigma,\rho,I)$ is there as a reminder that all probability distributions are computed based on our prior information I. Now substituting model Eq. (1) we have the direct probability of the data given the parameters:

$$P(D|\{\omega\},\{B\},\sigma,\rho,I) = (2\pi\sigma^2)^{-N}(1-\rho^2)^{-\frac{N}{2}}\exp\left\{-\frac{Q}{2\sigma^2(1-\rho^2)}\right\},$$

where

$$\begin{aligned} Q &\equiv d_R\cdot d_R - 2\rho d_R\cdot d_I + d_I\cdot d_I \\ &- 2\sum_{j=1}^{m}B_j[d_R\cdot G_j - \rho(d_R\cdot F_j + d_I\cdot G_j) + d_I\cdot F_j] \\ &+ \sum_{j=1}^{m}\sum_{k=1}^{m}B_jB_k[G_j\cdot G_k - \rho(G_j\cdot F_k + G_k\cdot F_j) + F_j\cdot F_k] \end{aligned}$$

and $(\cdot)$ means the sum over the discrete times: $d_I\cdot F_j \equiv \sum_{i=1}^{N} d_I(t_i)F_j(t_i)$.

Bayes' theorem tells us that the posterior probability of the nonlinear $\{\omega\}$ parameters, independently of the amplitudes, given the data and our prior information is

$$P(\{\omega\},\sigma,\rho|D,I) \propto \int d\{B\}P(\{B\},\{\omega\},\sigma,\rho|I)P(D|\{B\},\{\omega\},\sigma,\rho,I),$$

where $P(\{B\},\{\omega\},\sigma,\rho|D,I)$ is the posterior probability of the parameters, the direct probability of the data is $P(D|\{B\},\{\omega\},\sigma,\rho,I)$, and $P(\{B\},\{\omega\},\sigma,\rho|I)$ represents what was known about these parameters before we took the data and is called a prior probability. In this problem we assume that the data determine the parameters much more accurately than our prior information. Therefore, we assign a broad uninformative prior to the parameters: we use a uniform prior for the amplitudes and a Jeffreys prior for the variance.

Introducing the transformation

$$B_k = \sum_{j=1}^{m}\frac{A_je_{jk}}{\sqrt{\lambda_j}},\quad R_k = \sum_{j=1}^{m}\frac{G_je_{kj}}{\sqrt{\lambda_k}},\quad I_k = \sum_{j=1}^{m}\frac{F_je_{kj}}{\sqrt{\lambda_k}},$$

and

$$dB_1 \cdots dB_m = \lambda_1^{-\frac{1}{2}} \cdots \lambda_m^{-\frac{1}{2}} dA_1 \cdots dA_m$$

where e_{jk} is the kth component of the jth eigenvector of the interaction matrix

$$g_{jk} \equiv \sum_{i=1}^{N} G_j(t_i)G_k(t_i) - \rho(G_jF_k + G_kF_j) + F_j(t_i)F_k(t_i) \tag{2}$$

and λ_j is the jth eigenvalue, then the posterior probability of the parameters becomes

$$P(\{\omega\},\sigma,\rho|D,I) \propto \sigma^{-2N}(1-\rho^2)^{-\frac{N}{2}}\lambda_1^{-\frac{1}{2}} \cdots \lambda_m^{-\frac{1}{2}} \int_{-\infty}^{\infty} dA_1 \cdots dA_m \exp\left\{-\frac{Q'}{2\sigma^2(1-\rho^2)}\right\}$$

where

$$Q' = d_R \cdot d_R - 2\rho d_R \cdot d_I + d_I \cdot d_I - 2\sum_{j=1}^{m} A_j h_j + \sum_{j=1}^{m} A_j^2$$

and

$$h_j(\{\omega\},\rho) \equiv d_R \cdot R_j - \rho(d_R \cdot I_j + d_I \cdot R_j) + d_I \cdot I_j.$$

After completing the square in Q' and performing the m integrals, we have

$$\begin{aligned} P(\{\omega\},\sigma,\rho|D,I) &\propto \sigma^{m-2N}(1-\rho^2)^{-\frac{N-m}{2}}\lambda_1^{-\frac{1}{2}} \cdots \lambda_m^{-\frac{1}{2}} \\ &\times \exp\left\{-\frac{d_R \cdot d_R - 2\rho d_R \cdot d_I + d_I \cdot d_I - m\overline{h^2}}{2\sigma^2(1-\rho^2)}\right\} \end{aligned} \tag{3}$$

where

$$\overline{h^2} \equiv \frac{1}{m}\sum_{j=1}^{m} h_j^2.$$

If the variance of the noise σ^2 and the correlation coefficient ρ are known, then the problem is completed. The posterior probability of the frequencies and decay rates conditional on the data and our assumed knowledge of σ and ρ is

$$P(\{\omega\}|\sigma,\rho,D,I) \propto \lambda_1^{-\frac{1}{2}} \cdots \lambda_m^{-\frac{1}{2}} \exp\left\{\frac{m\overline{h^2}}{2\sigma^2(1-\rho^2)}\right\}. \tag{4}$$

But if σ is not known, then it too becomes a nuisance parameter to be removed by integration. Multiplying Eq. (3) by a Jeffreys prior and integrating with respect to σ, we obtain the posterior probability of the frequencies, decay rates, and the correlation coefficient ρ

$$P(\{\omega\},\rho|D,I) \propto \lambda_1^{-\frac{1}{2}} \cdots \lambda_m^{-\frac{1}{2}}(1-\rho^2)^{\frac{N-2m}{2}}\left[1 - \frac{2\rho d_R \cdot d_I + m\overline{h^2}}{d_R \cdot d_R + d_I \cdot d_I}\right]^{\frac{m-2N}{2}} \tag{5}$$

where $\overline{h^2}$ is a sufficient statistic for inferences about the $\{\omega\}$ parameters. Equation (5) is an exact result and does not depend on uniform sampling nor does it depend on the models being sinusoidal. Any quadrature data set that can be modeled by Eq. (1) can be used in these equations.

The Single Stationary Harmonic Frequency

What is to be gained from the use of Eq. (4) or (5) compared to a discrete Fourier transform of the data? The answer to this question is easily demonstrated by investigating one of the simplest quadrature spectral estimation problems: the single stationary harmonic frequency. Suppose we take

$$f_R(t) = B_1 \cos \omega t + B_2 \sin \omega t$$

as the model for the signal in the real channel and

$$f_I(t) = B_1 \sin \omega t - B_2 \cos \omega t$$

as the model for the signal in the imaginary channel. If the noise is uncorrelated, *i.e.*, $\rho = 0$, the interaction matrix, Eq. (2), becomes

$$g_{jk} = \begin{pmatrix} N & 0 \\ 0 & N \end{pmatrix}.$$

The R_j and I_j functions are given by

$$R_1 = \frac{\cos \omega t}{\sqrt{N}}, \quad R_2 = \frac{\sin \omega t}{\sqrt{N}}, \quad I_1 = \frac{\sin \omega t}{\sqrt{N}}, \quad I_2 = -\frac{\cos \omega t}{\sqrt{N}}.$$

The sufficient statistic $\overline{h^2}$ is given by

$$\begin{aligned} \overline{h^2} &= \frac{1}{2}\left[(R_1 \cdot d_R + I_1 \cdot d_I)^2 + (R_2 \cdot d_R + I_2 \cdot d_I)^2\right] \\ &= \frac{1}{2N}\left\{[C_R(\omega) + S_I(\omega)]^2 + [S_R(\omega) - C_I(\omega)]^2\right\} \end{aligned}$$

where

$$C_R(\omega) \equiv R_1 \cdot d_R = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} d_R(t_i) \cos \omega t_i$$

and

$$S_R(\omega) \equiv R_2 \cdot d_R = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} d_R(t_i) \sin \omega t_i$$

are the cosine and sine transforms of the real data, and $C_I(\omega)$ and $S_I(\omega)$ are the transforms for the imaginary data. The posterior probability of a stationary harmonic frequency ω, given the variance of the noise σ^2 and assuming the noise is uncorrelated, is

$$P(\omega|\sigma, D, I) \propto \exp\left\{\frac{[C_R(\omega) + S_I(\omega)]^2 + [S_R(\omega) - C_I(\omega)]^2}{2N\sigma^2}\right\}.$$

How does this compare to a discrete Fourier transform of the data? If we assume the data are the real and imaginary parts of a complex data set, then

$$d(t_i) = d_R(t_i) + i d_I(t_i).$$

Because

$$e^{-i\omega t} = \cos\omega t - i\sin\omega t,$$

the squared magnitude of the discrete Fourier transform may be written

$$\left| \sum_{k=1}^{N} \left[d_R(t_k) + i d_I(t_k) \right] e^{-i\omega t_k} \right|^2 = \left[C_R(\omega) + S_I(\omega) \right]^2 + \left[S_R(\omega) - C_I(\omega) \right]^2.$$

Up to the constant factor $1/2N$ the sufficient statistic $\overline{h^2}$ is the squared magnitude of a discrete Fourier transform of the complex data. Therefore, the discrete Fourier transform is essentially the natural logarithm of the posterior probability of a stationary harmonic frequency, given the variance of the noise σ^2, assuming the noise is uncorrelated, and assuming the channels are exactly 90° out of phase.

The implications of this are quite profound, because it means that only the highest peak in a discrete Fourier transform is of any importance for the estimation of a single stationary frequency, and then it is only the region around the maximum that is of importance. Moreover, the discrete Fourier transform will always interpret the data in terms of a single stationary harmonic frequency. If the data does not contain a single stationary harmonic frequency, or even if the data contain more than one stationary frequency, the discrete Fourier transform may give misleading or even incorrect results when compared to other more complex models. This is not because the discrete Fourier transform is wrong, but because it is answering what we should regard as the wrong question.

If we know that the signal consists of a single stationary harmonic frequency, how accurately can a frequency be estimated? We will assume that the data contain a single stationary sinusoid with no noise. Thus the accuracy estimates we derive will be optimistic in the sense that in real data, with a given noise variance σ^2, one would always make slightly worse frequency estimates than the ones we will derive. We take

$$d_R(t_i) = \hat{B}\cos\hat{\omega}t_i \quad \text{and} \quad d_I(t_i) = \hat{B}\sin\hat{\omega}t_i$$

as the signal in the real and imaginary channels, where $\hat{B}$ is the true amplitude of the sinusoid and $\hat{\omega}$ is the true frequency. We have set the phase of this sinusoid to zero. It will be obvious at the end of the calculation that the result for an arbitrarily phased sinusoid may be obtained by the replacement $\hat{B}^2 \to \hat{B}_1^2 + \hat{B}_2^2$. For uniformly sampled data we may take t_i to be integer or half integer, *i.e.*, $t_i = \{-T, -T+1, \cdots, T\}$ and

$2T+1=N$. The sufficient statistic $\overline{h^2}$ is

$$\begin{aligned}\overline{h^2} &= \frac{1}{2N}\left[\sum_{i=1}^{N}\hat{B}(\cos\hat{\omega}t_i\cos\omega t_i+\sin\hat{\omega}t_i\sin\omega t_i)\right]^2 \\ &\approx \frac{\hat{B}^2}{2N}\left[\frac{\sin\frac{N}{2}(\hat{\omega}-\omega)}{\sin\frac{1}{2}(\hat{\omega}-\omega)}\right]^2\end{aligned} \tag{6}$$

where we have explicitly performed the sum and have ignored terms of order one compared to N.

To estimate the accuracy of the frequency, we Taylor expand $\overline{h^2}$ in posterior probability

$$P(\omega|\sigma,D,I)\propto\exp\left\{\frac{\overline{h^2}}{\sigma^2}\right\}$$

around the maximum, and then make the (mean) $\pm$ (standard deviation) approximation. Around the maximum, the first derivative of $\overline{h^2}$ is zero, and the second is given by

$$\frac{\partial^2\overline{h^2}}{\partial\omega^2}\approx-\frac{\hat{B}^2N^3}{12}.$$

The Gaussian approximation to the posterior probability density is

$$P(\omega|\sigma,D,I)\approx\left(\frac{\hat{B}^2N^3}{24\pi\sigma^2}\right)^{\frac{1}{2}}\exp\left\{-\frac{\hat{B}^2N^3}{24\sigma^2}(\hat{\omega}-\omega)^2\right\},$$

from which we estimate the frequency to be

$$(\omega)_{\text{est}}=\hat{\omega}\pm\frac{\sigma}{|\hat{B}|}\sqrt{\frac{12}{N^3}},$$

or in Hertz

$$(f)_{\text{est}}=\hat{f}\pm\frac{\sigma}{2\pi|\hat{B}|T}\sqrt{\frac{12}{N}}\quad\text{Hz},$$

where T is now the total sampling time in seconds. The accuracy of the frequency estimate depends on the signal-to-noise ratio of the data, on the $\sqrt{N}$, and on the total sampling time T. The better the data, the better the estimate. If we double the number of data in the given sampling time we obtain the standard $\sqrt{2}$ improvement. However, if we sample two times longer, we pick up a factor of 2 from sampling longer and a factor of $\sqrt{2}$ from taking two times more data. Clearly for stationary frequencies taking data for a long time is the preferred way to sample the data.

In many NMR applications the discrete Fourier transform is taken directly as a frequency estimator. The accuracy is estimated from the full-width-at-half-maximum of the peak in the discrete Fourier transform. For the case just given, the squared magnitude of the discrete Fourier transform of the data (up to a constant) is given by Eq. (6). This has dropped to half its maximum value when the argument of the sine function has dropped to $\pi/4$:

$$\frac{N}{2}|\hat{\omega} - \omega| = \frac{\pi}{4}.$$

Thus for the discrete Fourier transform we find that the frequency estimate, in Hertz, is

$$(f)_{\text{est-dft}} = \hat{f} \pm \frac{1}{4T} \quad \text{Hz}$$

which neither depends on the magnitude of the signal nor the variance of noise σ^2.

Suppose we collect data for 1 second, with $\Delta T = 0.001$ seconds, collecting $N = 1000$ data values, and suppose we have RMS signal-to-noise ratio of $\hat{B}/\sqrt{2}\sigma = 1$. From the discrete Fourier transform we estimate the frequency to be

$$(f)_{\text{est-dft}} = \hat{f} \pm 0.25 \text{ Hertz},$$

and the Bayesian estimate is

$$(f)_{\text{est}} = \hat{f} \pm 0.012 \text{ Hertz}.$$

With signal-to-noise ratio of one, the Bayesian result is about 20 times better than the result from the discrete Fourier transform. If the signal-to-noise ratio were more typical of an NMR experiment, for example 100, then the Bayesian estimate would be more than three orders of magnitude better! Thus the probability analysis can estimate the frequency several orders of magnitude more precisely than a discrete Fourier transform directly. But this was in noiseless data. In practice, for frequency estimation, these procedures work at their theoretical best. However, the same cannot be said for other types of model functions. The reason frequency estimation is so accurate has to do with an interaction between the noise and the model functions. The oscillatory model functions and the noise tend to average to zero. When one computes the sufficient statistic, there is a sum of the model function times the data. Since the model and the noise are summing to zero separately, the sum of the product between the model and the noise tends to zero. This insures the projection of the model onto the noise is small compared to the projection of the model onto the signal, and the accuracy of the estimates are near the theoretical best. If the noise or the model did not average to zero, the accuracy estimates would be much worse.

The Single Frequency with Exponential Decay

In NMR the time series is typically the result of a complex chain of events: a sample is placed in a high magnetic field, and the nuclear spins are "excited" using a radio transmitter. These spins are then detected as they relax back to equilibrium. Using an RF antenna, the signal is amplified, split, mixed with a reference oscillator (a sine or cosine) oscillating with a frequency near the natural resonance of the sample, and low-pass filtered. The beats between the reference oscillator and the sample resonance are what is digitized and recorded. Because the signal in the two channels originated in the same physical event there is reason to expect the noise to be correlated. To give an understanding of what noise correlation can do for estimating the parameters we give a second example. We will use simulated data with noise correlations.

The data used in this example were generated from the following equations:

$$f_R(t_i) = 100\cos(0.3t_i + 1)\exp\{-0.01t_i\},$$

$$f_I(t_i) = 100\sin(0.3t_i + 1)\exp\{-0.01t_i\}.$$

To generate the data we first generated the signal from the above equations and then generated the noise. We generated the noise for the real channel from a Gaussian distributed random number generator with unit variance. To generate the noise for the imaginary channel we generated a second random number with unit variance and then added the noise from the real channel to this second random number. This was divided by $\sqrt{2}$ and then used as the noise in the imaginary channel. The noise in the two channels is, thus, slightly correlated.

The data and the discrete Fourier transform are displayed in Fig. 1. The data resemble an NMR signal which rapidly decays. There are $N = 512$ data values, and the signal-to-noise ratio is approximately 50. The discrete Fourier transform has a peak in the correct vicinity of the frequency. However, the width of the discrete Fourier transform is indicative of the decay rate, not the accuracy of the frequency estimate.

We now apply the results of this calculation to the data. The model we use is

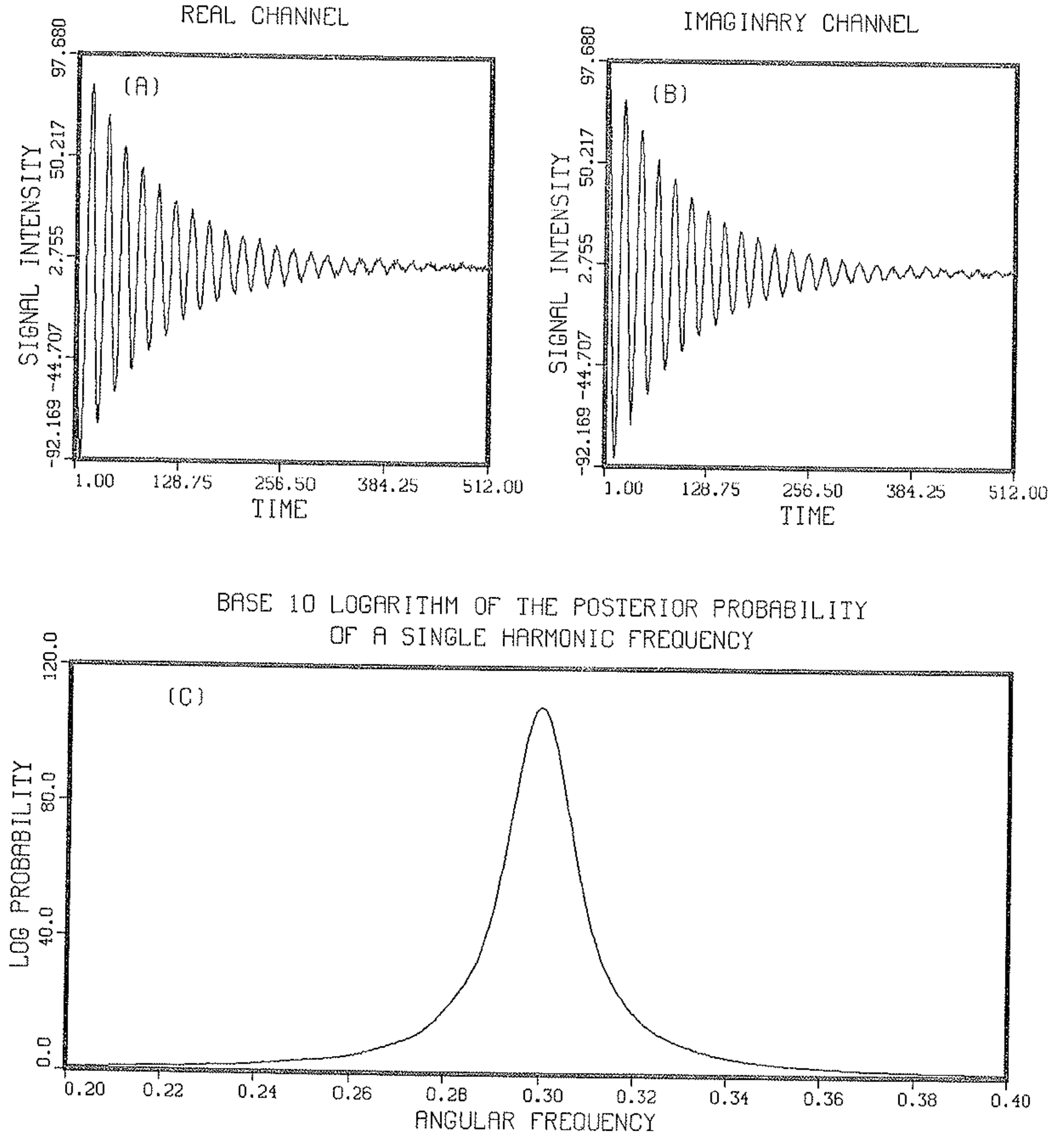
$$f_R(t) = B_1 \cos\omega t \exp\{-\alpha t\} + B_2 \sin\omega t \exp\{-\alpha t\}$$

for the real channel and

$$f_R(t) = B_1 \sin\omega t \exp\{-\alpha t\} - B_2 \cos\omega t \exp\{-\alpha t\}$$

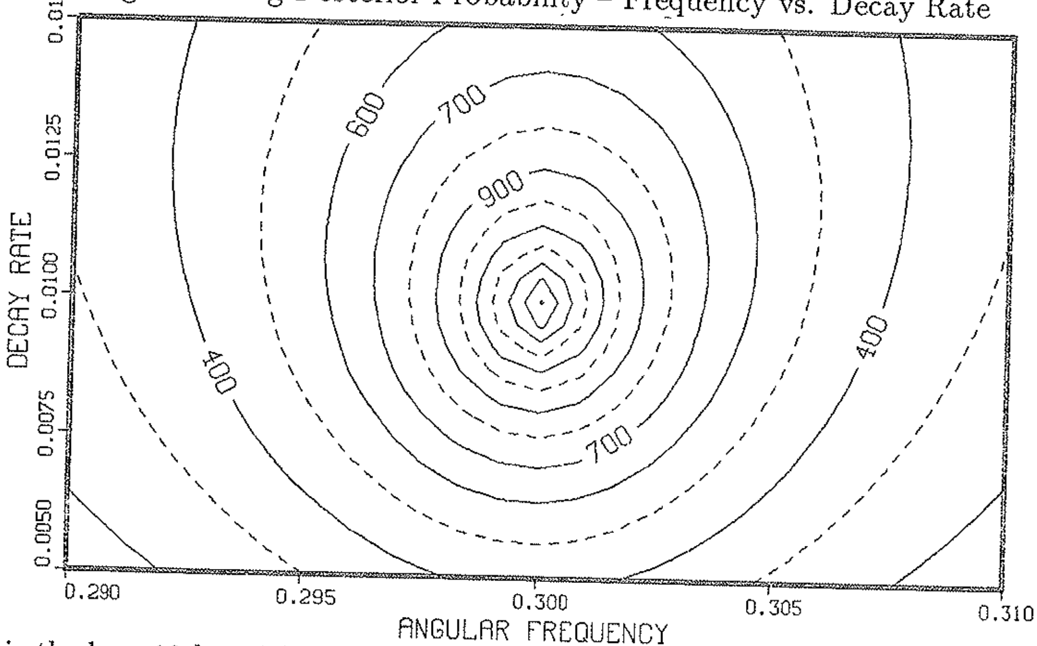
for the imaginary channel. After integrating out the amplitudes and variance of the noise, there are three remaining parameters to be estimated from the data: the frequency ω, the decay rate α, and the correlation coefficient ρ. We present the result of the calculation as three contour plots. First we plot the base 10 logarithm of the posterior probability of the frequency and decay rate while holding the correlation coefficient at its correct value. This is displayed in Fig. 2. We can see from this plot that there is a very sharp peak in the parameter space around the true value of the parameters. The normalization on this figure is irrelevant because of an interesting result, first noted by Jaynes [5]. If the contour lines are in increments of 1 (for example if the maximum posterior probability density were 100 and the contours be

Figure 1: The Computer Simulated Data and the Discrete Fourier Transform



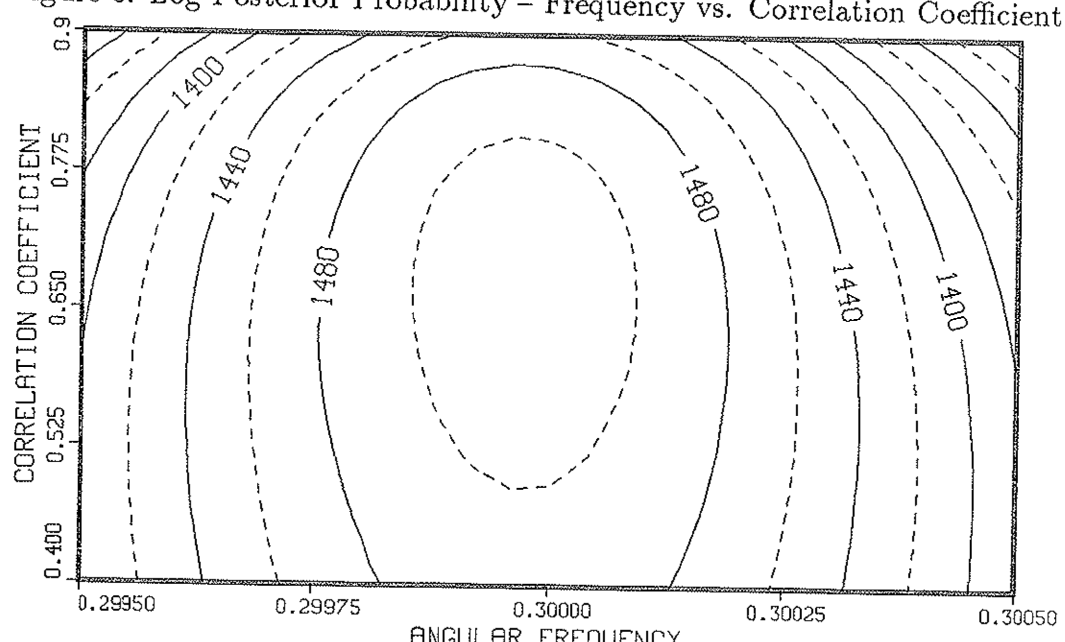
This computer simulated data (**A**) contain a single frequency which rapidly decays. The signal-to-noise ratio in these data is approximately 50. Now the discrete Fourier transform indicates the presence of a frequency in the right location. However, the width is indicative of the decay rate, not the accuracy of the estimate. Additionally, the discrete Fourier transform knows nothing of the noise correlations.

Figure 2: Log Posterior Probability – Frequency vs. Decay Rate



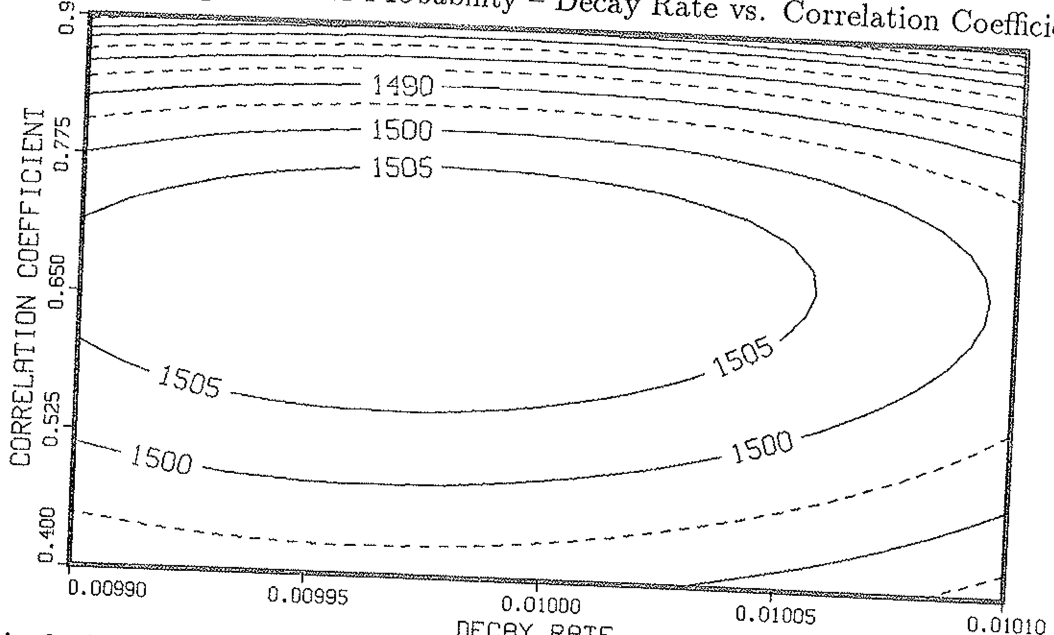
This is the base 10 logarithm of the posterior probability of the frequency and decay rate given the correlation coefficient. The total probability inside the highest contour is nearly 1. To write it out would require a decimal point followed by a string of approximately 200 nines.

Figure 3: Log Posterior Probability – Frequency vs. Correlation Coefficient



This is the base 10 logarithm of the posterior probability of the frequency and the correlation coefficient given the decay rate. The total probability inside the highest contour is nearly 1. Here it would require only a string of 20 nines to write it out.

Figure 4: Log Posterior Probability – Decay Rate vs. Correlation Coefficient



This is the base 10 logarithm of the posterior probability of the decay rate and the correlation coefficient given the frequency. The total probability inside the highest contour is approximately 0.99999.

labeled 99, 98, etc.), then for a 2D contour plot the first contour line contains 90% of the posterior probability, the second contour line contains 99% of the posterior probability, etc. Therefore, Fig. 2 represents an incredibly sharply peaked posterior probability density. The region inside the first contour line contains essentially all of the posterior probability. To write out the total probability enclosed by this contour would require a decimal point followed by a string of 200 nines. The second contour plot, Fig. 3, is the base 10 logarithm of the posterior probability of a frequency and the correlation coefficient given the true decay rate. Again there is a very sharp peak. The probability enclosed by the highest contour is approximately one; however, it would require only 20 nines to write it out. The third contour plot, Fig. 4, is of the base 10 logarithm of the posterior probability of the decay rate and the correlation coefficient given the true frequency. The probability enclosed by the highest contour here is only 0.99999.

Conclusions

In NMR a great deal of prior information is available about the time series. When this information is incorporated into the analysis of the data, the frequencies, decay rates, and amplitudes may be estimated several orders of magnitude better than by direct use of the discrete Fourier transform. Additionally, if the noise is correlated, substantial improvement in the estimation of the amplitudes, frequencies, and decay rates is possible.

Acknowledgments

This work supported by NIH grant GM-30331, J. J. H. Ackerman principal investigator. The encouragement of Dr. J. J. H. Ackerman and Professor E. T. Jaynes is greatly appreciated.

References

[1] Bretthorst G. L., (1987), Bayesian Spectrum Analysis and Parameter Estimation, Ph.D. thesis, Washington University, St. Louis, MO.; available from University Microfilms Inc., Ann Arbor, Mich.

[2] Bretthorst, G. L., (1988), Bayesian Spectrum Analysis and Parameter Estimation, in *Lecture Notes in Statistics,* Vol. 48, Springer-Verlag, New York, New York

[3] Jeffreys, H., (1939), Theory of Probability, Oxford University Press, London, (Later editions, 1948, 1961).

[4] Jaynes, E. T., (1987), "Bayesian Spectrum and Chirp Analysis," in Maximum Entropy and Bayesian Spectral Analysis and Estimation Problems, C. Ray Smith and G. J. Erickson, eds., D. Reidel, Dordrecht-Holland, pp. 1-37.

[5] Jaynes, E. T., (1988), private communications.