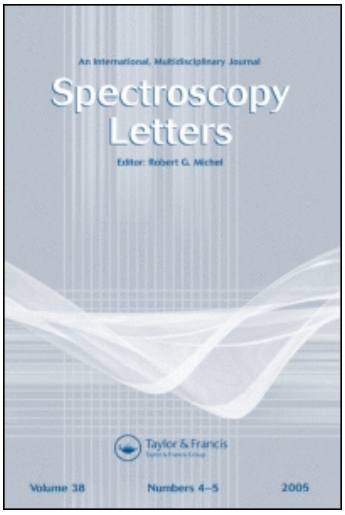

Spectroscopy Letters
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597299

Bayesian Deconvolution

Cynthia Rhode[a]; Scott Whittenburg[a]
[a] Department of Chemistry, University of New Orleans, New Orleans, Louisiana

To cite this Article Rhode, Cynthia and Whittenburg, Scott(1993) 'Bayesian Deconvolution', Spectroscopy Letters, 26: 6, 1085 — 1102
To link to this Article: DOI: 10.1080/00387019308011596
URL: http://dx.doi.org/10.1080/00387019308011596

SPECTROSCOPY LETTERS, 26(6), 1085-1102 (1993)

BAYESIAN DECONVOLUTION

Key Words: Bayesian probability, Deconvolution

Cynthia Rhode and Scott Whittenburg*
Department of Chemistry, University of New Orleans, New Orleans, Louisiana 70148

ABSTRACT

Bayesian probability has been applied to the problem of estimating the true spectral lineshape from an experimental peak with prior knowledge of the instrumental contribution. The results demonstrate that Bayesian deconvolution provides excellent estimates of the parameters defining the true spectral lineshape even from strongly overlapping or broadened peaks in the presence of significant noise levels.

INTRODUCTION

Signal recovery, namely deconvolution, spans several disciplines of experimental science. In all experiments where a signal is measured, the measuring instrument also adds something to the true signal causing a broadened signal. Elimination of that "something extra" is the art of deconvolution.

The desired signal is representable as a function $D(\omega)$ and the convolved observed measurement is

$$C(\omega) = \int T(\omega_0 - \omega) D(\omega) d\omega \qquad (1)$$

or

$$C = T \oplus D \tag{2}$$

where T is that "something extra" from the instrument. The recovery of D(ω) from C(ω) assuming $T(\omega_o-\omega)$ is known, is the process of deconvolution. In the time domain convolution is represented by

$$C(t) = T(t) * D(t) \tag{3}$$

Several deconvolution techniques have been developed, the most utilized of which are the Van Cittert[1] and modifications to the Van Cittert[2,3], and more recently Fourier self deconvolution[4]. Frieden[5] discusses many techniques in an extensive review of deconvolution. These techniques, while successful at deconvolution, do not yield directly the parameters of the model functions that describe the signal. Bayesian deconvolution is not only successful at deconvolving the signal, but also can directly yield values for the parameters that describe the deconvolved signal. The techniques also directly tells us about the accuracy of our estimated parameters. Bayesian deconvolution uses a mathematical probability theorem. The technique is based on recovering the most probable parameters of a model function that represents the data, given a set of data. Besides recovering the parameters, another advantage is that for a Lorentzian model function the amplitudes can be estimated as a fitting parameter, thereby decreasing valuable computational time.

THEORETICAL SECTION

Given a set of data

$$d_i = f(t_i) + e_i \tag{4}$$

where d_i is the data value at time t_i, and e_i is the noise at tha time, then we

can describe the time-domain signal by a function f(t) in the absence of noise,

$$f(t) = \Gamma * B\cos(\omega t)\exp[-\pi\Gamma t] + B\sin(\omega t)\exp[-\pi\Gamma t] \qquad (5)$$

which corresponds to two model functions. This model describes the general case of decaying sinusoid signals. In the frequency-domain, simple exponential decay corresponds to Lorentzian lineshapes. Gaussian lineshapes can be modeled with time-domain model functions with Gaussian decay. Data with a mixed Lorentzian/Gaussian frequency spectrum can be described by four model functions; two with simple exponential decay as in eq. 4 and two with Gaussian decay. But, our final model functions must reflect convolution, so we convolve the model functions with our chosen instrumental function. To describe the time-domain data we must compute the probabilities for 2*n+1 parameters (corresponding to A, ω, and Γ) in the data, where n is the number of frequencies in the frequency spectrum of the time-domain data. We proceed to evalute these parameters using Bayes' Theorem,

$$P(H,D|I) = P(D|I)P(H|D,I) = P(H|I)P(D|H,I) \qquad (6)$$

and

$$P(H|D,I) = \frac{P(H|I)P(D|H,I)}{P(D|I)} \qquad (7)$$

which is based on the fact that Aristotelian logic is commutative; so H and D are interchangeable. Our hypothesis H consists of the parameters of our model function, yielding

$$P(B,\omega,\Gamma|D,I) = \frac{P(B,\omega,\Gamma|I)P(D|B,\omega,\Gamma,I)}{P(D|I)} \qquad (8)$$

which says that the probability that the parameters are correct given the data, D and any prior information, I is equal to the product of the prior probability P(B, ω, Γ/I) and the direct probability of the data divided by the normalization probability of the data given only the prior information, P(D/I). Since we have no prior information about the values of the parameters, we can assume a uniform prior probability. Furthermore, we are not interested in a normalized probability so we will ignore the marginal probability. Finally, we will treat the amplitudes a nuisance parameters and integrate them out of the probability. The amplitudes will be back-calculated from the final results. Thus we wish to evaluate

$$P(\omega,\Gamma|D) = \int P(D|B,\omega,\Gamma)dB \qquad (9)$$

The probability that one of several mutually exclusive propositions is true, is the sum of their separate probabilities (sum rule of probability theory). We can solve the equation for the direct probability by first constructing the probability of the noise. The most conservative noise probability (least informative) is the result of maximum entropy,

$$P(e|\sigma) = (2\pi\sigma^2)^{-N/2}\exp[-\sum_{i=1}^{N} e(t_i)^2 / 2\sigma^2] \qquad (10)$$

for N points in the data and where σ is the noise variance. Solving for e_i in eq. 4 and substituting into eq. 10 yields the direct probability

$$P(\omega|D) = \int (2\pi\sigma^2)^{-N/2}\exp[-NQ / 2\sigma^2]dB \qquad (11)$$

where

$$Q = \overline{d^2}(\frac{2}{N})\sum_{i=1}^{N}\sum_{j=1}^{m} B_j d_i G_j(t_i) + (\frac{1}{N})\sum_{j=1}^{m}\sum_{k=1}^{m} g_{jk}B_jB_k \qquad (12)$$

where

$$\bar{d}^2 = (\frac{1}{N})\sum_{i=1}^{N} d^2(t_i) \tag{13}$$

and

$$g_{jk} = \sum_{i=1}^{N} G_j(t_i)G_k(t_i) \tag{14}$$

and where G_j and G_k are the model functions given in eq. 5. That is, $G_1(t) = \Gamma \cos(\omega t) \exp(-\Gamma\pi t)$ and $G_2(t)=\Gamma\sin(\omega t)\exp(-\Gamma\pi t)$. Eq. 14 can be solved by converting to orthonormal coordinates. We obtain the eigenvectors, e_{jk}, and the eigenvalues, λ_j, of g_{jk}. The orthonormal functions are then

$$H_j(t) = \lambda_j^{-1/2}\sum_{k=1}^{m} e_{jk}G_k(t) \tag{15}$$

In terms of the orthonormal functions, the model functions are

$$f(t) = \sum_{k=1}^{m} A_k H_k(t) \tag{16}$$

where the amplitudes are

$$A_k = \lambda_k^{-1/2}\sum_j B_j e_{kj} \tag{17}$$

and

$$B_k = \sum_j A_j e_{jk} / \lambda_j^{-1/2} \tag{18}$$

Integrating over the amplitudes and integrating over the noise variance yields

$$P(\omega,\Gamma|D) \propto (1-\frac{m\overline{h^2}}{N\overline{d^2}})^{(m-N)/2} \tag{19}$$

where

$$\overline{h^2} = m^{-1}\sum_j h_j^2 \tag{20}$$

and

$$h_j = \sum_{i=1}^{N} d_i H_j(t_i) \tag{21}$$

The probability that a resonance with frequency ω is in the data D given by eq. 19 is called the Student t-distribution. The maximun in the probability occurs at values of ω and Γ that maximize the sufficient statistic, h^2 The sufficient statistic is less sharply peaked than P(ω,Γ/D) and is, therefore, more amenable to digital optimization methods. The sufficient statistic is determined by projecting the model functions, $H_j(t_i)$, onto the data, d_j. In this work, a Lorentzian model function represents the data

$$L(\omega) = \sum_{i=1}^{N} \frac{A_i\Gamma_i^2}{\Gamma_i^2 + 4(\omega-\omega_{oi})^2} \tag{22}$$

where A is the amplitude, Γ is the full width at half height and ω_o is the frequency at maximum amplitude. The index i runs over the number of Lorentzians, N, in the data. This function can also be described in the time domain as

$$I(t) = \sum_{i=1}^{N} [\sin(2\pi\omega_0 t_i) + \cos(2\pi\omega_0 t_i)]\Gamma \exp(-\pi\Gamma t_i) \qquad (23)$$

The instrumental function can also be described by a Lorentzian lineshape, but centered at zero (ω_0 =0.0). The model function is convolved with an instrumental function in the time domain and Bayesian deconvolution is applied to recover the true signal from the data along with the parameters A, Γ, and ω.

Procedure

All of the computer programs required for conmputation are written in the C programming language on a Sun Sparcstation 1+. A pure Lorentzian lineshape is chosen to represent data in the frequency and time domains using eqs. 22 and 23, respectively. This generated data describes the sought-after deconvolved signal, D(ω) in eq. 1. Generation in both domains is done to compare the Bayesian output data to original frequency domain data. All of the actual Bayesian deconvolution computations are done on the time domain data. Many modern instruments are FT instruments, that is, the data is acquired in the time domain and the instrument transforms the data to the frequency domain for the user. Ultimately, one would perform the inverse fast Fourier transform on frequency domain data acquired from an instrument, then do Bayesian deconvolution. The generated data sets consisst of 1024 points. An instrumental function is also generated using a Lorentzian lineshape but with unit amplitude and with zero resonance frequency.

In order to represent convolved data from an instrument, each data set is convolved with the instrumental function. When convolving in the time domain, we simply multiply the Lorentzian function by the instrumental function point by point. Convolution in frequency or transform space is a little more tedious, requiring computing the sums of the products of the instrumental function with the Lorentzian as the instrumental function is swept over the Lorentzian function. This is represented mathematically by eq. 1. The effect of convolution of a peak with an instrumental contribution is a much broader and flatter spectrum.

In proceeding to apply Bayesian deconvolution, the convolved data set is read from a disk. Initial guesses for the parameters of frequency and linewidth are determined. We choose an initial linewidth of 2.0 and the frequency parameters are systematically chosen by examining the frequency domain data. A moving average of the intensities over 23 points is used to determine possible resonance frequencies. Model functions representing eq. 5 are set up using our guesses for the parameters. These model functions are then convolved with an instrumental function in order to represent G_j and G_k in eq. 14. We then proceed to maximize the sufficient statistic for our model functions by using conjugate gradient optimization of the parameters. Each time that the parameters are iterated, the model functions are again convolved with the instrumental function. Another peak is then added now giving us two more parameters to optimize. The new resonance frequency is chosen 5 frequency units away from the previous and the linewidth is chosen to be the same as the previous. Again the parameters for our new model functions are optimized using conjugate gradient optimization and a sufficient statistic is obtained. This sufficient statistic is compared to the previous one. The addition of another Lorentzian is repeated until the new sufficient statistic is no longer greater than the previous one. The parameters for the model functions that yield the maximum sufficient statistic represents the deconvolved signal.

RESULTS AND DISCUSSION

We have appplied Bayesian deconvolution initially to a simple data set where the peaks are well-separated. In order to provide a difficult test for the Bayesian deconvolution method, we have also chosen to study two close peaks with a sufficiently broad instrumental contribution so as to cause severe overlap. The parameters for the synthetic data sets are given in Table 1 and Table 2. The instrumental linewidth was 4.0. Various peak to peak noise levels were added to the data using a random number generator as indicated in Table 3.

Again, the true signal is convolved with the instrumental function to represent a convolved data set that would be obtained from an instrument. Bayes-

TABLE 1

Parameters Used to Generate Two Well-separated Lorentzians

Parameter	Peak1	Peak2
Amplitude	100.00	50.00
Linewidth	4.00	6.00
Frequency	45.00	55.00

TABLE 2

Parameters Used to Generate Two Overlapping Lorentzians

Parameter	Peak1	Peak2
Amplitude	100.00	50.00
Linewidth	4.00	6.00
Frequency	47.00	55.00

TABLE 3

Results of Parameters at Different Noise Levels, N, for Two Well-seperated Peaks. The Instrumental Linewidth is 4.0. The True Values are in Table 1.

Parameter	N=0.0	N=1.0	N=4.0	N=8.0
Amplitude				
Peak1	100.02	100.72	102.55	104.23
Error (%)	0.02	0.72	2.55	4.23
Peak2	50.03	49.09	46.4	43.35
Error (%)	0.06	1.82	7.2	13.3
Linewidth				
Peak1	4.00	4.02	4.08	4.15
Error (%)	0.00	0.50	2.00	3.75
Peak2	6.00	6.02	6.08	6.19
Error (%)	0.00	0.33	1.33	3.17
Frequency				
Peak1	45.00	45.00	44.99	44.97
Error (%)	0.00	0.00	0.02	0.07
Peak2	55.00	54.93	54.69	54.36
Error (%)	0.00	0.13	0.56	1.16

TABLE 4

Results of Parameters at Different Noise Levels, N, for Two Overlapping Lorentzians. The Instrumental Linewidth is 4.0. The True Values are in Table 2.

Parameter		N=0.0	N=1.0
Amplitude	Peak1	99.08	109.47
	Error (%)	0.92	9.47
	Peak2	50.08	52.64
		0.16	0.68
Linewidth	Peak1	3.99	4.14
	Error (%)	0.25	3.50
	Peak2	6.08	4.68
	Error (%)	1.33	17.00
Frequency	Peak1	46.99	47.06
	Error (%)	0.02	0.13
	Peak2	53.05	51.60
	Error (%)	0.09	2.64

ian deconvolution is applied and the resulting parameters describe the deconvolved lineshape.

Table 3 gives the results corresponding to the data set where the peaks are well-separated at four different noise levels of 0.0, 1.0, 4.0 and 8.0. Table 4 gives the results for two peaks that are close to one another corresponding to noise levels of 0.0 and 1.0. Sample spectra are shown in Figures 1 and 2.

Several limitations of the method are apparent in the results given in these Tables. When the signal-to-noise ratio of the data is very high, Bayesian deconvolution does a remarkable job of extracting the true signal from the convolved spectrum even when the peaks are unresolved. The source of the broadening can be due to either the overlap of the unconvolved peaks, the linewidth of the instrumental contribution, or a combination of these two effects. Because the linewidth information of the instrumental contribution is contained in the data file containing the experimentally measured or generated instrumental function the Bayesian algorithm can still regenerate a good

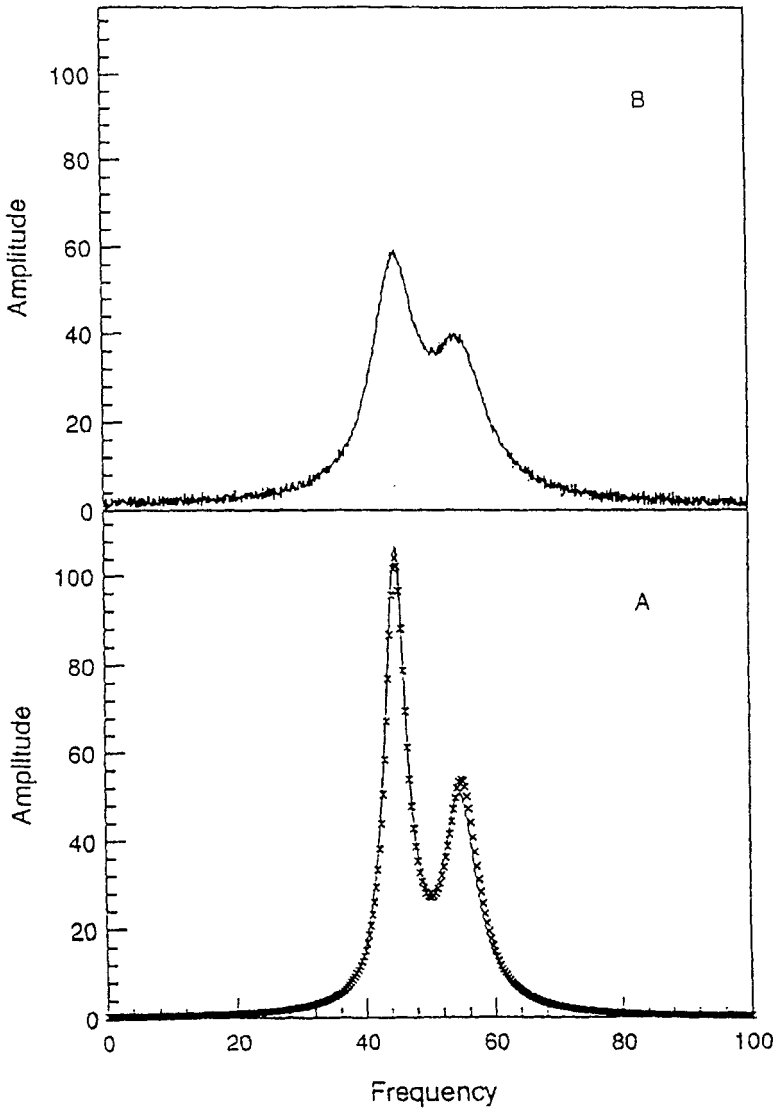

FIG. 1. Two well-resolved peaks with noise=4.0 and linewidth=4.0. (A) represents the Bayesian generated result (_) overlapping the true signal (x). (B) represents the convolved data.

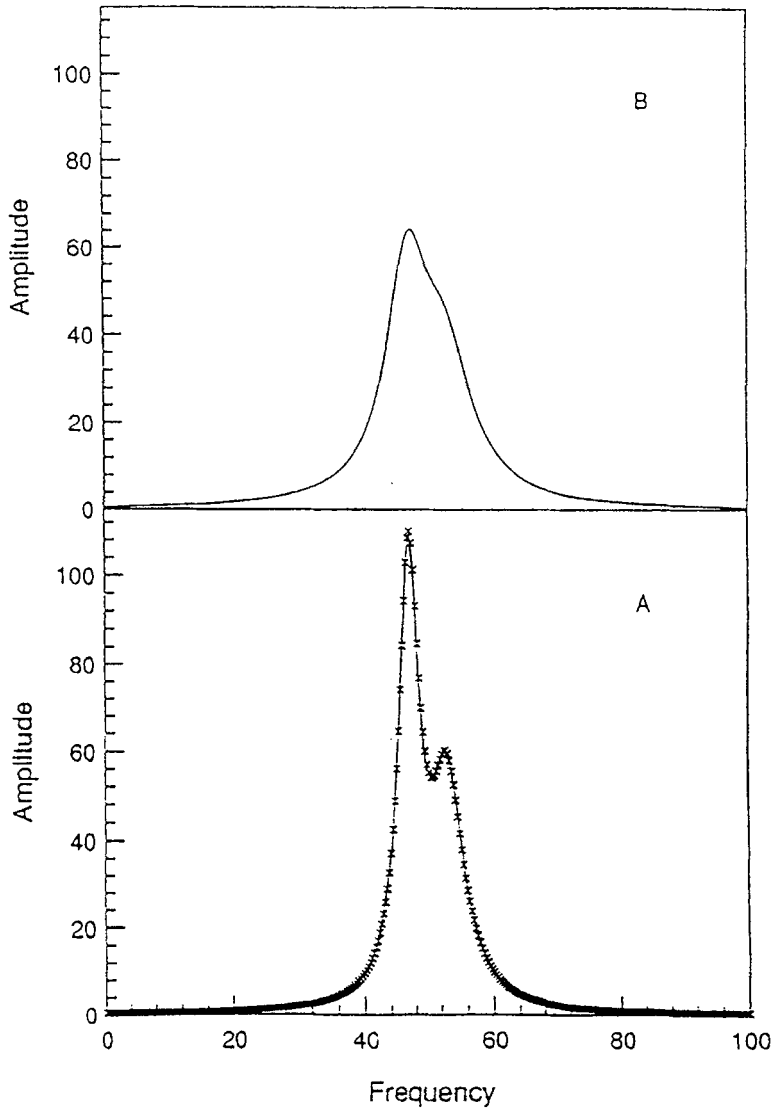

FIG. 2. Two overlapping peaks with noise=0.0 and linewidth=4.0. (A) represents the Bayesian generated result (_) overlapping the true signal (x). (B) represents the convolved data.

approximation to the unconvolved spectra as long as the convolved peaks are reasonably resolved. The Bayesian approach can deconvolve peaks in which one of the bands appears as a "shoulder", however, as the noise is increased, the uncertainty in the linewidth, and the resulting estimate of the amplitude, increases rapidly.

Brillouin Data

In Brillouin scattering a light wave has its frequency shifted by an amount equal to the frequency of a sound wave. Since energy must be conserved, the frequency of the incident photon is equal to the frequency of the scattered photon plus the frequency of the sound wave. The Brillouin spectrum consists of a central Rayleigh peak, usually centered at the incident laser frequency and a pair of shifted peaks, called the brillouin doublet. A Lorentzian lineshape best describes the peaks.

Bayesian deconvolution is applied to a Brillouin spectrum of formamide (300°C). Van Cittert deconvolution is also applied to the data and the results are compared to the Bayesian deconvolution.

The formamide data contains 512 data points. The signal is represented as a function of frequency; therefore, in order to apply Bayesian deconvolution, the data must be inverse Fourier transformed into the time-domain. The signal-to-noise ratio of the data set is approximately 19:1 for the two side bands and 40:1 for the central peak. The spectrum is shown in Figure 3.

An instrumental spectrum is also experimentally obtained by measuring the Brillouin spectrum of a fritted glass cell (so that the incident light is totally reflected). Instead of actually using the experimentally obtained instrumental spectrum, the linewidth (FWHM) of the spectrum is physically measured and used to generate an instrumental function. The instrumental linewidth is 8.0.

The parameters that represent the deconvolved lineshape obtained from Bayesian deconvolution are given in Table 5. The spectrum generated from these parameters is shown graphically in Figure 4. The model function (converted to the original coordinate system) that maximizes the sufficient statistic including convolution is shown in Figure 3.

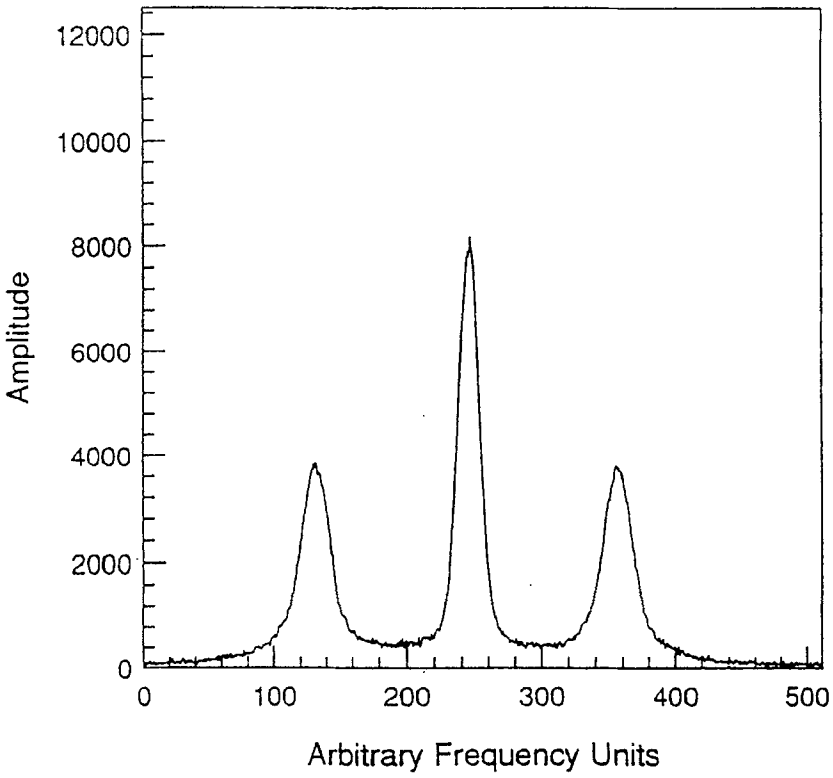

FIG. 3. Brillouin spectrum of formamide.

TABLE 5

Results of Bayesian Deconvolution of the Brillouin Spectrum of Formamide

Parameter	Peak1	Peak2	Peak3
Amplitude	5116.66	11962.03	4471.91
Linewidth	23.12	10.22	18.69
Frequency	131.29	242.55	353.6

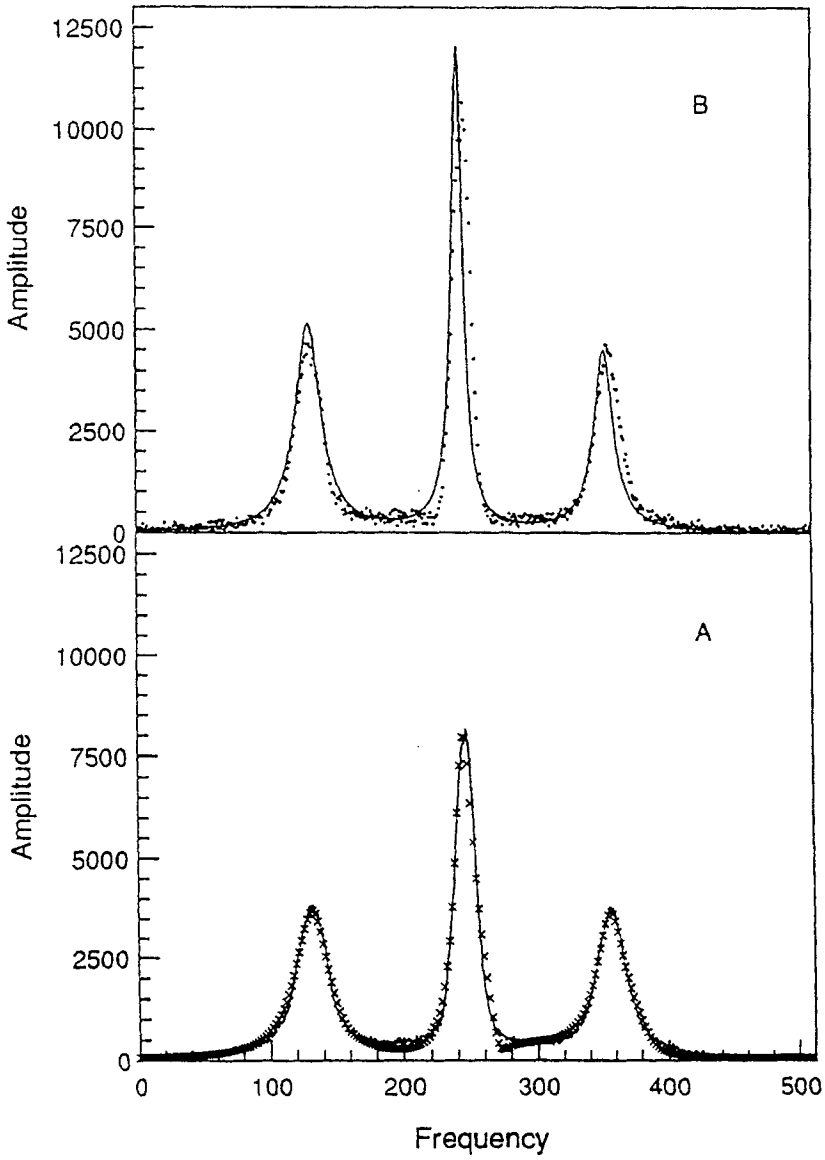

FIG. 4. Bayesian deconvolution of experimental Brillouin spectrum. (A) represents the Bayesian deconvolved spectrum. (B) represents the original data (_) and the Bayesian result convolved with the instrumental function (x).

The Van Cittert method of deconvolution can be represented by

$$D_n = D_{n-1} + (C - D_{n-1} \oplus T)$$

assuming D_1=C, where D_n is the deconvolved signal, D_{n-1} is the result of the previous iteration, C is the observed convolved data, and T is the instrumental function. First the observed signal is convolved with the instrumental function, subtracted from the observed signal and then used as the estimator for the next iteration. This process is continued until a convegence criterion is met. Convergence can be forced by multiplying by a mixing coefficient, α, yielding

$$D_n = D_{n-1} + \alpha(C - D_{n-1} \oplus T)$$

When applying the Van Cittert, three different values of the mixing coefficient, α, are chosen to examine how the choice effects the deconvolved result. These values of α are 0.05, 0.005, and 0.0005. The deconvolved spectrum demonstrating the best compromise between noise and linewidth is obtained at α=0.005. The resulting spectrum is also shown in Figure 4.

Since there is no way of knowing what our final result should be, we cannot calculate any errors in the parameters obtained. But since we do have confidence in the Bayesian deconvolution results of generated data with comparable signal-to-noise ratios, we can speculate that the Bayesian results are very good. Furthermore, after observing the spectrum of the model function that maximizes the sufficient statistic (Figure 3) and comparing it to the data, we can draw the conclusion that our final deconvolved spectrum is a good estimate.

Upon applying Van Cittert deconvolution to the Brillouin data, the algorithm does not converge, but is terminated after 300 iterations. Furthermore the results vary with different values of the initial mixing coefficient. These discrepancies can probably be attributed to the noise in the data, since we experience similar problems with the generated data sets that contain noise, but not with the noise-free data sets. Noise has been demonstrated by Blass and Halsey[6] to be one cause of Van Cittert not to converge. They determined

the "reasonable minimum requirement for deconvolvable data (by modified Van Cittert) to be a signal-to-noise ratio of 50:1 or greater".

CONCLUSION

Bayesian deconvolution provides a method for deconvolution of an instrumental function from a convolved spectrum while directly yielding deconvolved parameters such as peak height, linewidth, and area. The method is superior to the conventional Van Cittert approach, particularly at lower signal-to-noise ratios. The technique also handles peaks that are strongly-overlapping.

The instrumental signal can either be measured experimentally, or generated by estimating the linewidth. For spectroscopic methods for which the instrumental function can be experimentally measured, i.e. Brillouin or Raman Spectroscopy, the instrumental spectra can be stored into a data file and used directly in the convolution step of the algorithm. For spectroscopic methods in which this is not possible and approximate instrumental function may be generated and stored in a file. It is not necessary that the instrumental function be a simple Lorentzian or Gaussian lineshape.

Bayesian deconvolution is also superior to the Van Cittert in the case of deconvolving a real Brillouin data set. The modified Van Cittert leaves doubt as to the choice of the best value for the mixing coefficient. Furthermore, the best deconvolved spectrum has added moise from the deconvolution process, which will only increase the difficulty of obtaining the parameters from a fitting routine.

We have recently applied Bayesian analysis to multi-dimensional NMR. Several of the unique features of Bayesian analysis make it an ideal technique for multidimensional analysis. In future work we will apply Bayesian probability to multi-dimensional deconvolution. Bayesian deconvolution can also be applied to other lineshapes including Gaussian or mixed Gaussian/Lorentzian peaks. This approach will make Bayesian deconvolution especially valuable in other areas of spectroscopy such a FT-IR and Raman spectroscopy, and possibly, even in the field of chromatography.

REFERENCES

1. Van Cittert, P.H. Z. Phys. 1931; **69** : 298.

2. Jansson, P. A., Hunt, R. H., Plyer, E. K. J. Opt. Soc. Am. 1970; **60** : 184.

3. Jansson, P. A. *Deconvolution with Applications in Spectroscopy*, Orlando : Academic Press Inc., 1984.

4. Kauppien, J. K. Fourier Self-Deconvolution in Spectroscopy. In: Vanesse, G. ed. *Spectometric Techniques*. New York : Academic Press Inc. 1983 : 199-215.

5. Frieden, B. R. In: Huang, T.S. ed. *Picture Processing and Digital Filtering*. Berlin, New York: Springer-Verlag 1975 : 179-234.

6. Blass, W. E., Halsey, G. W. *Deconvolution of Absorption Spectra*. New York, London, Toronto, Sydney, San Franciso: Academic Press, 1981.

Date Received: December 23, 1992

Date Accepted: January 29, 1993