Available online at www.sciencedirect.com
ScienceDirect
Journal of Magnetic Resonance 183 (2006) 123–133
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Bayesian and information theory analysis of MAS sideband patterns in spin 1/2 systems

Joseph R. Sachleben

Department of Chemistry, Otterbein College, Westerville, OH 43081, USA

Received 11 April 2006; revised 3 July 2006
Available online 8 September 2006

Abstract

Bayesian statistics and information theory are used to analyze the reliability of extracting chemical shift parameters from spinning sideband patterns of spin 1/2 systems. Efficient code has been written to calculate the two-dimensional posterior probability as a function of the chemical shift anisotropy, δ, and the asymmetry parameter, η, given the sideband intensities and the signal-to-noise ratio. This method has the advantage of assuming only that the noise in the sideband intensities is distributed as a Gaussian. It assumes nothing about the distribution of the values of parameters δ and η, which are shown in some cases to be highly non-Gaussian. The utility of Bayesian analysis is demonstrated on 1D slow-spinning MAS spectra and on sideband patterns extracted from 2D PASS spectra. Previous investigations have shown that there is an optimal range of spinning frequencies for determining δ. In this study, information theory is used to determine the signal-to-noise ratio dependence of the entropy in δ, η, and total entropy in spinning sideband spectra. The entropy is a measure of the information content of a probability distribution. When the entropy is zero, there is perfect information on a system, while if it is infinite, there is no information on the system. It is found that for all values of η and for signal-to-noise ratios in the range 50–1000, an entropy minimum in ν_δ/ν_r occurs for values $1.5 \leqslant \nu_\delta/\nu_r \leqslant 3$. In the same range of signal-to-noise ratios, the entropy in η is a monotonically decreasing function of ν_δ/ν_r. The global information content of a spinning sideband pattern (i.e., the total entropy) is dependent on the signal-to-noise ratio and has an optimal value at $\nu_\delta/\nu_r \approx 2$ at a signal-to-noise ratio of 50 and increasing to $\approx$2.5 for signal-to-noise ratios of 1000. Finally, the increase of information in a sideband pattern as a function of the number of sidebands used in the analysis is examined. Most of the information about δ and η is contained in the five central sidebands; i.e., sidebands -2 to 2.

Keywords: Magic angle spinning; Errors; Bayesian analysis; Chemical shift tensor

1. Introduction

The goal of measuring NMR spectra is to determine the basic parameters that give rise to the spectrum and to extract physical or chemical information about the system from these parameters. For instance, by determining the chemical shift anisotropy (CSA) of a spin 1/2 nucleus in a solid, one can determine the symmetry of the local nuclear environment and thus the bond hybridization and coordination number of the site and the dynamics of that region of the solid. Ascertaining the certainty with which the parameters are determined is a critical step in determining the reliability of results and extracting chemical information.

In solid-state NMR, we are often interested in determining three parameters for each chemical environment: the isotropic chemical shift (ν_{iso}), the chemical shift anisotropy (δ), and the asymmetry parameter (η) [1,2]. These parameters can be extracted from both the static powder patterns and the intensities of magic angle spinning (MAS) sideband patterns. The occurrence of spinning sideband patterns was originally described by Mariq and Waugh [3] and by Herzfeld and Berger [4]. These papers described the source of the spinning sidebands in the MAS spectrum and demonstrated that the CSA parameters could be determined from the intensities of the sidebands. Herzfeld and Berger [4] supplied a series of graphs of relative sideband intensity

E-mail address: jsachleben@otterbein.edu

1090-7807/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2006.07.014

as a function of CSA parameters for different sideband orders that allows the graphical extraction of the CSA parameters. The difficulties in establishing precise tensor values when η is close to zero [5] and a method of estimating the standard deviation in the tensor parameters [6] were discussed in the literature in the late 1980s and early 1990s. Olivieri analyzed the errors in δ and η by inverting the Hessian matrix of the sidebands with respect to the parameters [7]. Hodgkinson and Emsley [8] provided a rigorous analysis of the uncertainty of the CSA parameters in the limit that the errors are distributed as a Gaussian and demonstrated that it is quicker to determine these parameters with a given uncertainty from MAS sideband patterns than from a static powder pattern. They showed at one signal-to-noise ratio that there is an optimal value of the ratio of the chemical shift anisotropy to the rotation frequency, $\upsilon_\delta/\upsilon_r$, where $\upsilon_\delta = \delta\upsilon_0$ and υ_0 is the Larmor frequency of the nucleus, which minimizes the error in δ. In their simulations, they find that the optimal value is $\upsilon_\delta/\upsilon_r = 2.56$.

In the following, the aim is to remove some of the limitations of previous work and extend the analysis to the amount of information present in a spectrum of given signal-to-noise ratio. Unfortunately, until the last 10 years, performing Bayesian statistical analysis and rigorously extracting confidence intervals of CSA parameters has been difficult due to the non-linearity of the problem and limitation in computation power. In this time-period, Bayesian analysis has been used to address many problems in NMR spectroscopy. Starting with the pioneering work of Bretthorst, who used Bayesian analysis to analyze liquid-state NMR spectra, [9–12] it has been applied to many diverse NMR problems including automated signal recognition in 2D NMR, [13] estimation of dynamic parameters from relaxation data, [14] analysis of protein sidechain rotamer preferences, [15] identification of helix content in solid peptides, [16] and to extract CSA parameters from spinning sideband patterns [17]. In this paper, an efficient method of calculating the errors in CSA parameters from MAS sideband patterns is demonstrated. Previous analyses of errors either explicitly or implicitly assumed that the errors in the parameters were distributed as a Gaussian. In the following, a Bayesian analysis approach has been used to calculate the posterior probability that the data are described by a set of values of the parameters. This approach allows a map of probability versus CSA parameters and peaks in this plot are "good fits" to the measured sideband intensities. For this analysis, it is assumed that the noise in the spectrum is white, i.e., not dependent on frequency, and distributed in Gaussian fashion about the true value, but it is not assumed that the errors in the parameters are distributed as a Gaussian. It will be shown that in some cases the probability distribution for the parameter is decidedly non-Gaussian. In addition, if the frequency dependence of the noise is known from the action of a filter, for instance, it is easily included in this formalism. Efficient C code has been developed to implement the Bayesian analysis of sideband patterns and examples with 1D ^{13}C MAS and 2D ^{13}C PASS spectra will be shown. This paper also examines the information content of MAS spectra as a function of the signal-to-noise ratio and the number of sidebands analyzed.

2. Theoretical and computational

Bayesian statistics gives a fundamentally different interpretation of probability than does traditional statistics. In traditional statistics, probability is the long-term relative frequency with which an event occurs. Since it takes an infinite number of measurements to determine this long-term relative frequency, traditional statistics concentrates on determining this population distribution from a small sample taken from it. Bayesian statistics does not assume the existence of an unmeasured and immeasurable population distribution. Instead, it interprets the probability as a "degree of belief" that something is true based on the evidence acquired [18]. In the case of spinning sideband patterns in magic angle spinning NMR spectroscopy, the probability of the parameters δ and η given the set of sideband intensities, $\tilde{\mathbf{I}}$, the signal-to-noise ratio, SN, and the background information, BI is of interest. This posterior probability is represented by $P(\delta,\eta|\tilde{\mathbf{I}},\mathrm{SN},\mathrm{BI})$. Bayes' Theorem relates this probability to other more easily determined probabilities by

$$P(\delta,\eta|\tilde{\mathbf{I}},\mathrm{SN},\mathrm{BI}) = \frac{P(\tilde{\mathbf{I}}|\delta,\eta,\mathrm{SN},\mathrm{BI})P(\delta,\eta|\mathrm{BI})}{P(\tilde{\mathbf{I}}|\mathrm{SN},\mathrm{BI})}. \tag{1}$$

$P(\tilde{\mathbf{I}}|\mathrm{SN},\mathrm{BI})$ is the probability of the sideband intensities given the background information and the signal-to-noise ratio. This term, called the evidence, is only a part of a normalization factor in the current application. $P(\delta,\eta|\mathrm{BI})$ is the prior probability; that is, the probability of δ and η given the background information [18]. By choosing the experimental conditions correctly, one already has information about δ and η. For instance, if the sweepwidth is chosen properly, then it is known that—$\mathrm{sw}/2 \leqslant \delta \leqslant \mathrm{sw}/2$ and it is always known that $0 \leqslant \eta \leqslant 1$. In this report, a uniform prior probability is assumed, which says that the probability that the parameter is in some range is just the inverse of the range, or

$$\begin{aligned} P(\delta,\eta|\mathrm{BI}) &= \frac{1}{(\delta_{\mathrm{upper}}-\delta_{\mathrm{lower}})(\eta_{\mathrm{upper}}-\eta_{\mathrm{lower}})} \\ &= \frac{1}{(\delta_{\mathrm{upper}}-\delta_{\mathrm{lower}})} = \frac{1}{\mathrm{sw}}, \end{aligned} \tag{2}$$

where the range in η is assumed to be from 0 to 1 and δ is within the sweepwidth of the experiment. For the analysis of any given set of data, this is also just part of the normalization constant. This allows us to rewrite Eq. (1) as

$$P(\delta,\eta|\tilde{\mathbf{I}},\mathrm{SN},\mathrm{BI}) = NP(\tilde{\mathbf{I}}|\delta,\eta,\mathrm{SN},\mathrm{BI}), \tag{3}$$

where N is the normalization constant.

Eq. (3) shows that the posterior probability is proportional to the likelihood function, $P(\tilde{\mathbf{I}}|\delta,\eta,\mathrm{SN},\mathrm{BI})$; the

probability of the sideband intensities given δ and η, the signal-to-noise ratio and the background information [18]. Assuming that the noise in each sideband is distributed as a Gaussian, the probability that a measured sideband intensity, I_i, is described by some value of δ and η given the signal to noise ratio and the background information is

$$P(I_i|\delta,\eta,\mathrm{SN},\mathrm{BI}) = \frac{1}{\sigma_i\sqrt{2\pi}}\exp\left(-\frac{(I_i - I_i^{\mathrm{Sim}})^2}{2\sigma_i^2}\right). \quad (4)$$

As long as the sidebands of different orders are well separated, they can be measured independently, which makes the likelihood function the product of the probabilities for each sideband,

$$P(\tilde{\mathbf{I}}|\delta,\eta,\mathrm{SN},\mathrm{BI}) = \prod_{i=N_{\mathrm{SB}}^{\mathrm{Min}}}^{N_{\mathrm{SB}}^{\mathrm{Max}}} \frac{1}{\sigma_i\sqrt{2\pi}}\exp\left(-\frac{(I_i - I_i^{\mathrm{Sim}})^2}{2\sigma_i^2}\right). \quad (5)$$

This allows us to rewrite the posterior probability as

$$P(\delta,\eta|\tilde{\mathbf{I}},\mathrm{SN},\mathrm{BI}) = N\exp\left(-\frac{1}{2}\chi^2\right), \quad (6)$$

where

$$\chi^2 = \sum_{i=N_{\mathrm{SB}}^{\mathrm{Min}}}^{N_{\mathrm{SB}}^{\mathrm{Max}}} \frac{(I_i - I_i^{\mathrm{Sim}})^2}{\sigma_i^2} \quad (7)$$

and $N_{\mathrm{SB}}^{\mathrm{Min}}$, $N_{\mathrm{SB}}^{\mathrm{Max}}$ are the minimum and maximum values of the sideband order analyzed. Analyzing MAS spectra using these equations is equivalent to examining the two-dimensional χ^2 for the fit of δ and η; however, the analysis of posterior probability has several advantages. The first of these is the intuitiveness of examining a probability distribution especially within the Bayesian interpretation as a "degree of belief" in the determined parameters. Second is the ease of converting the posterior probability into a confidence interval. Finally, the probability distribution can be mapped into the entropy, which allows the comparison of the information content of spinning sideband patterns acquired under different conditions.

A one-dimensional $N\%_x$ confidence interval for a parameter x is the minimum interval of the parameter containing $N\%_x$ of the area under the posterior probability. In a multidimensional case, the posterior probability must be projected onto the axis of interest, and then the confidence interval is found on the resulting one-dimensional posterior probability. In the analysis of MAS sideband patterns, this process can be represented as

$$N\%_\delta = 100\int_{\delta_{\min}}^{\delta_{\max}}\int_0^1 P(\delta,\eta|\tilde{\mathbf{I}},\mathrm{SN},\mathrm{BI})\,\mathrm{d}\eta\,\mathrm{d}\delta \quad (8)$$

for δ, and as

$$N\%_\eta = 100\int_{\eta_{\min}}^{\eta_{\max}}\int_{-\infty}^{\infty} P(\delta,\eta|\tilde{\mathbf{I}},\mathrm{SN},\mathrm{BI})\,\mathrm{d}\delta\,\mathrm{d}\eta \quad (9)$$

for η, where the posterior probability is normalized to 1 and the intervals $\Delta\delta = \delta_{\max} - \delta_{\min}$ and $\Delta\eta = \eta_{\max} - \eta_{\min}$ are of minimum length. Eq. (8) finds the limits of the confidence interval in δ by projecting over all possible values of η and then integrating from $\delta_{\min}$ over δ until the area between $\delta_{\min}$ and $\delta_{\max}$ equals $N\%_\delta/100$. This defines $\delta_{\max}$ for a given value of $\delta_{\min}$. $\delta_{\min}$ is then varied to minimize the distance between $\delta_{\min}$ and $\delta_{\max}$, which defines the confidence interval. The confidence interval for η defined in Eq. (9) is found similarly. Such confidence intervals can easily be found numerically.

Eqs. (6) and (7) assume that the noise in the sideband intensity is distributed as a Gaussian, but assume nothing about the distribution of the values of δ and η. For instance, in cases where the problem is underdetermined, i.e., an insufficient number of sideband intensities have been measured, multimodal probability distributions result. The distribution of η is commonly non-Gaussian due to the finite range of the possible values of η. In all the data analyzed in this paper, the noise is assumed to be constant as a function of frequency. With modern NMR equipment utilizing good digital filtering this is normally an appropriate assumption; however, if there is significant frequency dependence of the noise this can be included into Eq. (7) by using different values of σ_i for each sideband intensity. In this case, the simulated sideband intensities would also need to be corrected by the measured filter function.

In order to efficiently simulate the posterior probability distribution, an efficient method of comparing the simulated spinning sideband intensities with the measured is needed. Every sideband pattern is uniquely determined by the parameters $\upsilon_\delta/\upsilon_r$ and η, allowing the efficient simulation of a lookup table of sideband patterns. Using Eqs. (1), (6), (7), and the lookup table of sideband intensities, the two-dimensional posterior probability function can be quickly generated from the measured sideband intensities. This function gives the probability that pairs of values of δ and η describe the measured sideband intensities. Peaks in this map indicate regions of values of δ and η that have a high probability of describing the data. Source code for these programs is available upon request from the author. Fig. 1a and b shows that this method correctly extracts values of δ and η. A simulated MAS sideband pattern ($\delta = 80.0$ ppm, $\eta = 0.400$, $\nu_{\mathrm{rot}} = 1500$ Hz, spectrometer frequency = 100 MHz, and SN = 240) was analyzed with the Bayesian program. The resulting posterior probability distribution, Fig. 1b, was calculated in less than 1 min and shows a single strong probability peak near the appropriate values of δ and η. The values determined at the 95% confidence level are $\delta = 80.2 \pm 1.0$ ppm and $\eta = 0.394 \pm 0.025$, indicating that the true values are properly bracketed and that this procedure is both accurate and computationally efficient.

The optimal conditions to acquire spinning sideband patterns in order to produce the most information in the least time were analyzed. For this comparison, the total signal-to-noise, SN, was defined as the ratio between the noise at the end of the FID to the signal in its first point, or, equivalently, as the ratio of the noise between the side-

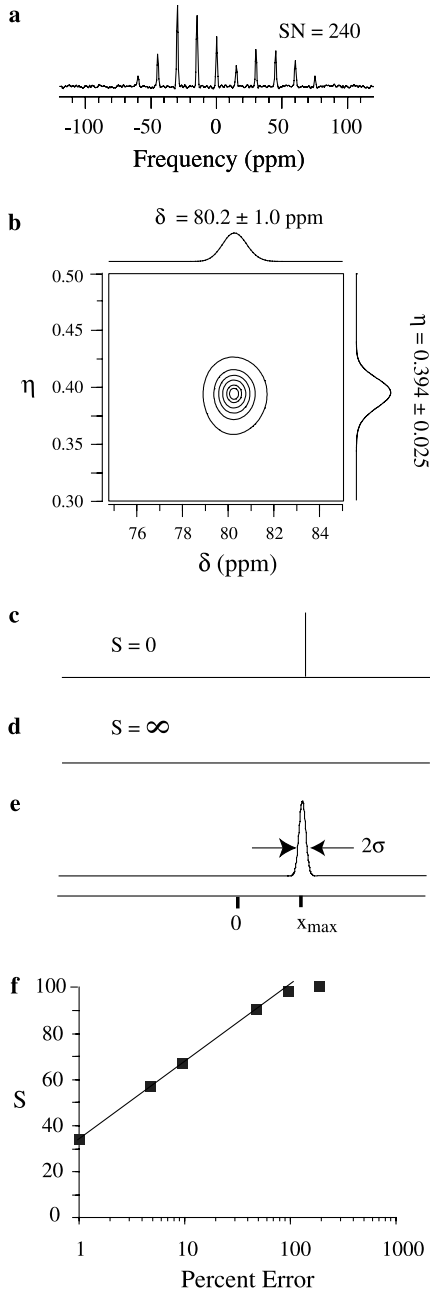

Fig. 1. (a) Simulated spinning sideband pattern with the following parameters $\nu_{iso} = 0$ ppm, $\delta = 80.0$ ppm, $\eta = 0.400$, $\nu_{rot} = 1500$ Hz, Larmor frequency for nucleus = 100 MHz, and the signal-to-noise ratio = 240. The signal-to-noise ratio is the total sideband intensity to the noise. (b) Two-dimensional posterior probability distribution for the simulated sideband pattern in (a). The 95% confidence intervals for δ and η show that the Bayesian analysis properly brackets the original parameters used for the simulation. (c) δ-Function probability distribution with a value of x_{max}. Such a distribution has an entropy of zero. (d) An infinitely diffuse distribution function, i.e., a distribution that is perfectly flat but whose area is 1. Such a distribution has infinite entropy since there is no information on the value of the x. (e) A Gaussian distribution of the parameter x with a width of 2σ and a peak at x_{max}. (f) Simulation of the entropy as a function of the logarithm of the percent error defined as Percent error $= 100\frac{\text{std}(x)}{x_{max}}$. Since the simulation was performed over a finite range in x, the maximum entropy was defined as 100. Notice that the entropy scales linearly with the logarithm of the percent error until the peak is broader than the range, which results in a scaled entropy of approximately 100.

bands to the total intensity in all the sidebands. In order to conveniently compare the error in different experimental situations, i.e., for various values of $\upsilon_\delta/\upsilon_r$, η, and SN, posterior probability distributions must be reduced to a single number that reflects the information content of the distribution. The entropy, S, of a one-dimensional function is defined as

$$S = -\int p\ln(p)\mathrm{d}x = -\sum_{i=1}^{N} p_i \ln(p_i)\Delta x_i, \tag{10}$$

where the sum is over all elements of the probability distribution. For a multidimensional probability distribution, the entropy is the multidimensional integral of $-p\ln p$. The entropy maps the probability distribution into a single number and allows the comparison of the information content in different probability distributions. The smaller the entropy, the more certain one is of the result. For instance, if the result is perfectly certain, then the probability distribution is a delta function at the value of the parameter, Fig. 1c. In this case, the entropy has a value of 0. In the case where the probability distribution is flat over all space but the area under it integrates to one, Fig. 1d, the entropy is infinite, i.e., there is no information about the values of the parameters. We use the entropy to examine the information content of spinning sideband patterns under different experimental circumstances.

As an empirical measure of the entropy, the step size in δ and η is ignored, and entropy is calculated by

$$S = -\sum_{i=1}^{N} p_i \ln(p_i) \tag{11}$$

where p_i is normalized posterior probabilities, defined by $\sum_{i=1}^{N} p_i = 1$. This definition of the entropy will depend on the number of steps in δ and η. To account for this, an entropy scale for the investigation of experimental parameters on information content has been defined. In this entropy scale, the maximum entropy is 100, which corresponds to no information on the value of the parameter(s), while the minimum entropy is 0 which corresponds to perfect information, which corresponds to only one value of δ and η describing the data. Since the probability distributions are found over a finite range of parameters with a discrete step size, a flat distribution does not lead to infinite entropy. In the case where all values in the two-dimensional plane are equally probable, $P_i = 9.98 \times 10^{-7}$ under the conditions of our simulations, indicating that our data have equal probability of being described in the region of our simulation. In this case, $S = -NP_i\ln(P_i) = -\ln(9.98 \times 10^{-7}) = 13.818$, which is the maximum value of the entropy for the 2D simulations. For the one-dimensional case where all values are equally probable, the maximum entropy is given by $S = -\ln(9.99 \times 10^{-4}) = 6.9087$. These values have been normalized to 100 in the entropy simulations to ease the interpretation of the results. Fig. 1e and f shows the entropy of a Gaussian probability distribution as a function of the percent error, which is defined as

$$\text{Percent error} = 100\frac{\text{std}(x)}{x_{\max}}, \quad (12)$$

where std(x)and $x_{\max}$ are the standard deviation and average position of the Gaussian. Fig. 1f shows the entropy as a function of the log of the percent error. The entropy scales linearly with the log of the percent error, demonstrating that it acts as a measure of the error. The relationship between the entropy and the percent error is complicated in the general case; however, it is clear that the smaller the entropy the smaller the error on the value of the parameter(s) determined.

3. Methods

3.1. NMR

^{119}Sn MAS spectra of Sn(II) oxalate (Aldrich) were acquired at 67.1 MHz on a Chemagnetics CMX180 NMR spectrometer with a rotor synchronized Hahn echo experiment. Two hundred and fifty-six free induction decays with 512 complex points were acquired over a sweep width of 100 kHz and with a recycle delay of 10 s. MAS spectra with spinning frequencies of 5.5 and 6.2 kHz were acquired in order to assign an isotropic chemical shift of Sn(II) oxalate of −871 ppm. Spectra were externally referenced to SnO whose isotropic shift is −199 ppm with respect to $Sn(CH_3)_4$.

The ^{1}H decoupled ^{13}C 2D PASS spectrum [19] of the yttrium glutarate MIL-8 [20] was acquired with 2 ms cross polarization, CW decoupling (50 kHz decoupling field), and a spinning frequency of 750 Hz. Sixteen sideband phase altered spectra with 1024 complex points in each were acquired. These spectra were copied and appended to the file eight times, which after shearing, zero-filling in the MAS time domain, and Fourier transformation leads to a spectrum with the infinite spinning frequency MAS spectrum along one axis and the sidebands along the other. Appending the copies of the fid to the sideband dimension interpolates zeros between the sidebands. All spectra and posterior probability functions were processed and viewed with Philip Grandinetti's program RMN [21].

3.2. Computational

All simulations presented in this paper were run on a Silicon Graphics O2, a Pentium III PC running Linux, or a MacIntosh G3 iBook using code written in C by the author. A table of sideband intensities as a function of $\upsilon_\delta/\upsilon_{\text{rot}}$ and η were simulated in the region $0 \leqslant \upsilon_\delta/\upsilon_r \leqslant 7$ and $0 \leqslant \eta \leqslant 1$ with 1001 steps of size 0.007 and 0.001, respectively. MAS sideband patterns were simulated by dividing one rotor period into $2n$ steps, where n is the highest order sideband desired, and then propagating the initial density matrix over one rotor period. Powder averaging was performed by using a 3-angle set with 3722 members generated by the ZCW algorithm, which was kindly supplied by Malcolm Levitt [22]. Sideband intensities were produced by Fast Fourier Transforming the data. The resulting table contains 17 sidebands (orders −8 to 8) for each value of $\upsilon_\delta/\upsilon_r$ and η. Posterior probability functions were calculated by comparing sideband intensities of the experimental data with the 1,002,001 simulated sideband patterns using the maximum likelihood predictor and the input signal-to-noise ratio of the spectrum. In a typical run of the Bayesian program, a file of sideband intensities and the rms noise of the spectrum are input and a two-dimensional, δ versus η posterior probability function is calculated and visualized using RMN. RMN is then used to project the data onto the two parameter axes. These 1D distributions are then used as input files for a program that calculates the confidence interval. The posterior probability is normalized so that the sum over all the simulated points is 1. Projections onto the two axes give one-dimensional posterior probabilities from which confidence intervals in δ and η can be calculated. One-dimensional N% confidence intervals were found by minimizing the interval on the parameter axis under the constraint that N% of the area under the posterior probability was found in that interval.

4. Results and discussion

To test Bayesian analysis on real samples, the spinning sideband patterns of a 1D, ^{119}Sn MAS spectrum of Sn(II) oxalate and sideband patterns of the carbonyl carbons of the yttrium glutarate MIL-8 taken from a ^{13}C 2D PASS spectrum of the sample have been analyzed. Fig. 2a shows the ^{119}Sn MAS spectrum of Sn(II) oxalate. The signal-to-noise ratio (total intensity of all sidebands divided by the noise between the sidebands) of this spectrum is 226. The isotropic shift is −871 ppm. Fig. 2b shows the simulated sideband pattern using the most probable values of δ and η from the Bayesian analysis. Fig. 2c shows the 2D posterior probability distribution function and the 1D projections. From these it was found that $\delta = -624.8 \pm 7.6$ ppm and the most probable value of η was 0.098 but the confidence interval indicates that $\eta \leqslant 0.145$ with 95% confidence. Notice in particular that the probability distribution for η is not distributed as a Gaussian because the minimum value of η is 0. Fig. 3a shows the ^{13}C 2D PASS spectrum of the microporous yttrium glutarate MIL-8. MIL-8 has an array of yttrium(III) ions supported by the dicarboxylic glutaric acid. The glutaric acid molecules arrange themselves between the lattice of metal ions in such a way as to form one-dimensional channels from which water can be exchanged [20,23]. Fig. 3b–d show the extracted spinning sideband patterns of the three resolved carbonyl peaks of the sample as well as their posterior probability distributions and the simulated sideband patterns for the most probable value of the parameters. While the isotropic shifts of these three sites are very similar, the resolution is sufficient for the three carbonyl sites to be distinguished and their CSA parameters extracted. Fig. 3b–d shows that the CSA parameters clearly

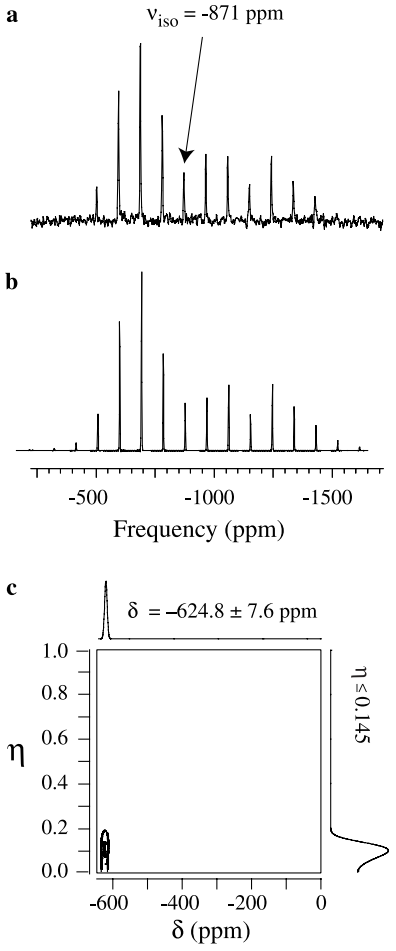

Fig. 2. Bayesian analysis of the ^{119}Sn spinning sideband pattern of Sn(II) oxalate acquired on a CMX180 NMR spectrometer at a Larmor frequency of 67.1 MHz acquired with a rotor synchronized Hahn echo sequence. (a) The ^{119}Sn MAS spectrum acquired at a spinning frequency of 6.2 kHz. The signal-to-noise ratio (intensity of all sidebands/noise) of this spectrum is 218. The isotropic peak was identified by running a second experiment at a spinning frequency of 5.5 kHz. (b) The simulation of the most probable sideband pattern determined from the posterior probability function for δ and η for the spectrum in (a). The most probable value of δ and η were determined to be -624.8 ppm and 0.098, respectively. (c) Posterior probability distribution for δ and η. The distribution shows a single peak that gives $\delta = -624.8 \pm 7.6$ ppm and $\eta \leqslant 0.145$ at 95% confidence.

distinguish the three nuclear environments of the three carbonyl sites in the sample and that there is excellent agreement between the most probable sideband patterns and the data. In cases where there is correlation in the error between δ and η (data not shown), the correlation is indicated by an oval shape of the peak in the probability distribution that is not aligned with either axis. The tilt of the major axis of the oval is related to the correlation coefficient of δ and η. These examples demonstrate that Bayesian analysis is a very powerful technique for the analysis of solid-state NMR data. The errors determined by this method are much lower than initially expected. Relative errors in δ of about 1% have been determined in these examples, while for η the error was less than 5%, in the case where it could

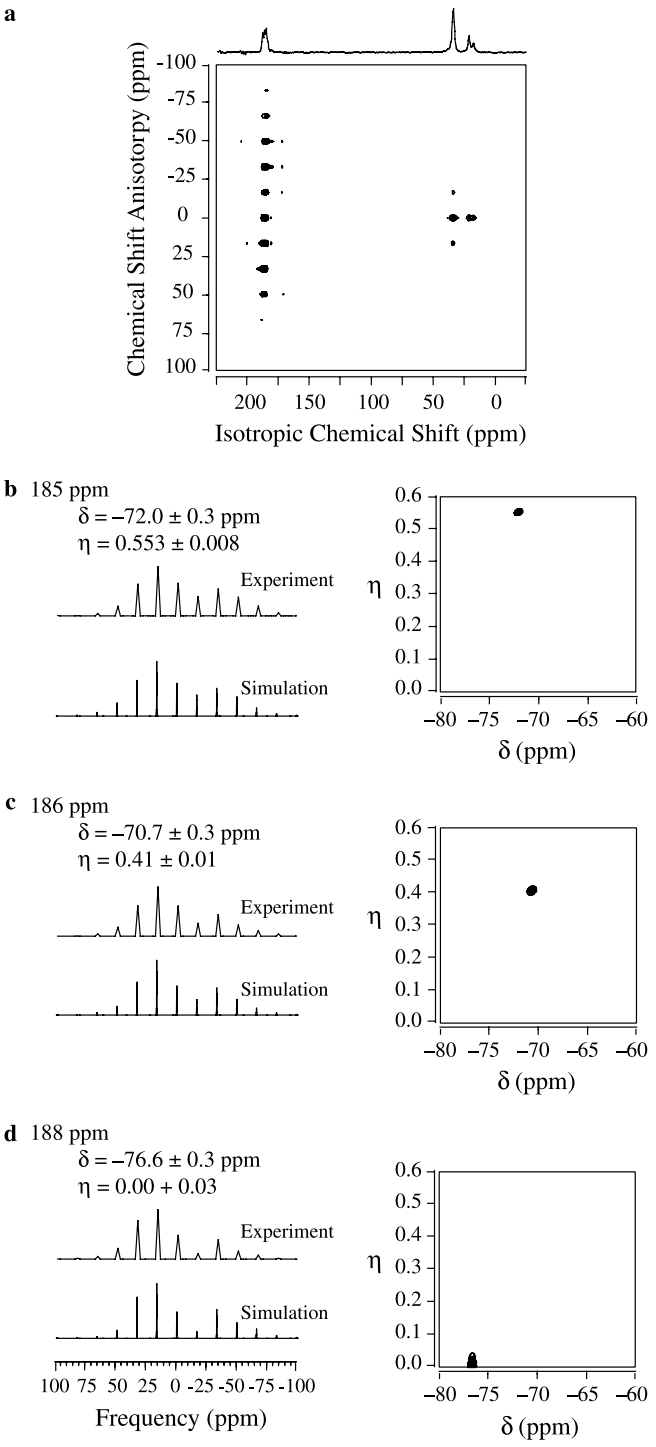

Fig. 3. Bayesian analysis of the ^{13}C spinning sidebands patterns of the carbonyl carbons of the yttrium glutarate MIL-8 extracted from the 2D PASS experiment. (a) The 2D PASS spectrum of yttrium MIL-8. The horizontal axis corresponds to the infinite spinning speed MAS spectrum, while the vertical axis contains only the sideband patterns. (b–c) Bayesian analysis of the three carbonyl peaks at the isotropic shifts 185, 186, and 188 ppm. It is clear from the posterior probability distributions and the resulting confidence intervals that the chemical shift anisotropies of these three sites are significantly different.

be determined. (In cases where $\eta = 0$, the relative error is meaningless.) These values are consistent with those seen in the simulations shown in Fig. 1.

In the rest of this paper, the effects of different experimental conditions on the information content of an MAS spinning sideband pattern are examined. In Fig. 4, the dependence of the entropy of the δ, η projections of the posterior probability—these are called the entropy in δ and η—as well as the total entropy in the full 2D posterior probability as a function of the values of $\upsilon_\delta/\upsilon_r$, η, and the signal-to-noise ratio are examined. The entropy maps of δ, left column in Fig. 4, show a minimum at values of $\upsilon_\delta/\upsilon_r$ between about 1.5 and 3.0 for all values of η and the signal-to-noise ratio. This indicates that the error in δ can be minimized if $\upsilon_\delta/\upsilon_r$ is in this range. This is consistent with the result of Hodgkinson and Emsley which indicates that the optimal value of $\upsilon_\delta/\upsilon_r$ was 2.56 [8]. The entropy maps of η, middle column in Fig. 4, show a monotonically decreasing function of $\upsilon_\delta/\upsilon_r$ for all values of η and signal-to-noise ratios examined. This indicates that the accuracy of η increases as more sidebands are present in the spectrum independent of the signal-to-noise ratio and the value of η. Thus, to determine η accurately, it is desirable to have as many sidebands as possible in the spectrum, but this limits the accuracy of δ. By examining the total entropy of the 2D probability distribution, the optimal relative spinning frequency, $\upsilon_\delta/\upsilon_r$, that gives the most information possible in the spectrum can be determined; this is the relative spinning frequency that simultaneously allows the determina-

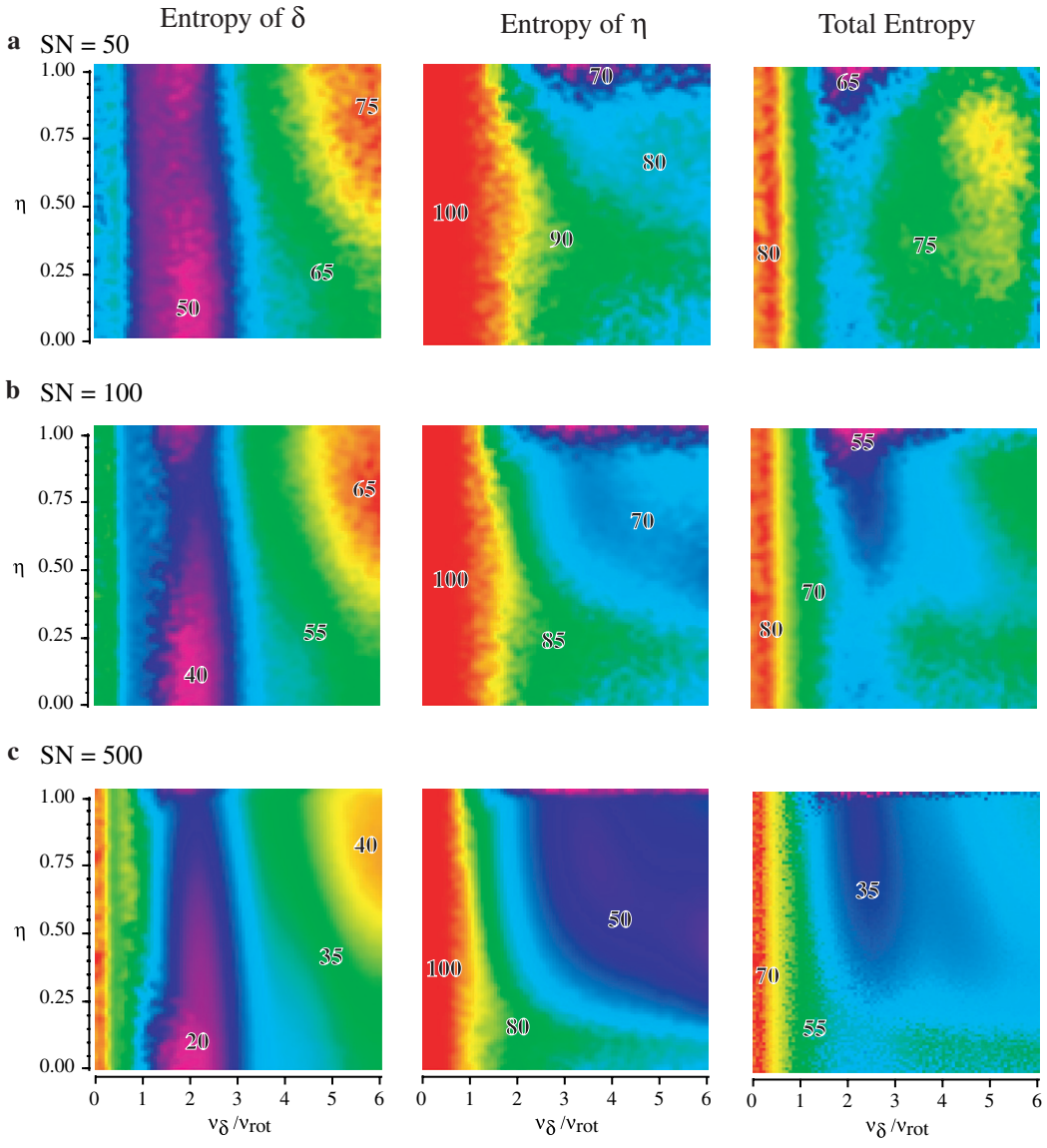

Fig. 4. The entropy of δ, η, and total entropy as a function of $\upsilon_\delta/\upsilon_r$, η, and the signal-to-noise ratio (SN). Approximate values of the entropy are placed on the graphs to allow comparisons. For all simulated signal-to-noise ratios, there is a minimum in the entropy of δ for values of $\upsilon_\delta/\upsilon_r$ between 1 and 3, while the entropy in η is a monotonically decreasing function of $\upsilon_\delta/\upsilon_r$ for fixed value of η. The total entropy shows a minimum like the entropy in δ, but it is shifted due to the effects of η. The minimum now occurs for values of $\upsilon_\delta/\upsilon_r$ between 2 and 3, with a slight signal-to-noise dependence. At lower signal-to-noise ratios, lower values of $\upsilon_\delta/\upsilon_r$ are indicated (δ/ν_{rot} closer to 2), while at higher signal-to-noise, $\upsilon_\delta/\upsilon_r$ closer to 3 are acceptable. The numbers in these plots indicate the value of the entropy where 0 means perfect information and 100 means no information about the CSA parameters.

tion of the most accurate values of δ and η in a single spectrum at a given signal-to-noise ratio. The right column in Fig. 4 shows that there is a minimum in the total entropy; however, it is much more shallow than that seen in the δ entropy. There is also a slight signal-to-noise ratio dependence of the optimal relative spinning frequency. At SN $\approx$ 50, the entropy minimum occurs at values of $\upsilon_\delta/\upsilon_r \approx 2$. At higher signal-to-noise ratios, SN $\approx$ 500, the entropy minimum occurs at $\upsilon_\delta/\upsilon_r \approx 2.5$. For noiseless signals, SN $\approx \infty$, any finite spinning frequency will lead to perfect information about the values of δ and η. In practical situations where the signal-to-noise ratio is between 50 and 500, the spinning frequency should be adjusted so that $\upsilon_\delta/\upsilon_r$ is between 2 and 2.5 to produce the most accurate values of δ and η in the least amount of time. This corresponds to having significant sideband intensity in between 5 and 7

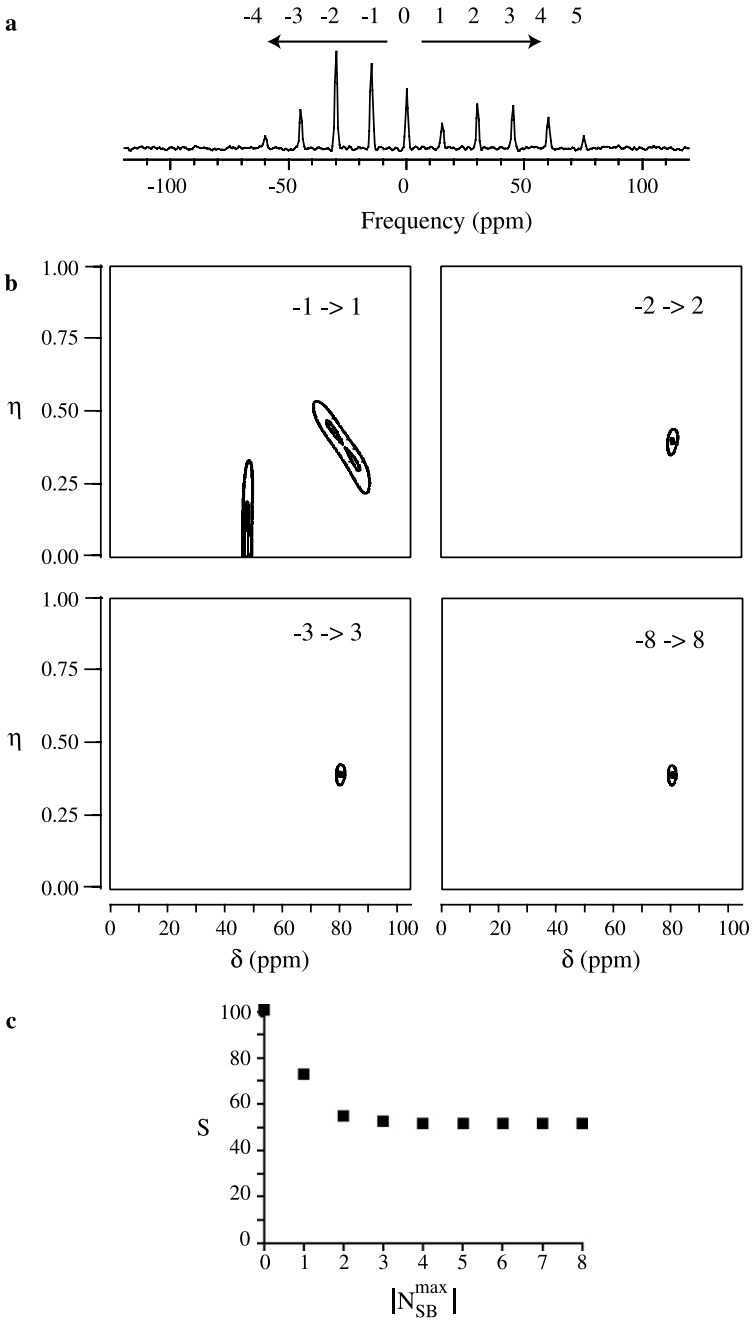

Fig. 5. Information content as a function of the number of inner sidebands analyzed. (a) The same simulated spectrum as in Fig. 1 with the sideband order (n) labeled. ($\upsilon_{\text{sideband}} = \upsilon_{\text{iso}} + n\upsilon_{\text{rot}}$). (b) Posterior probability distributions for different numbers of sidebands analyzed. When only the sidebands -1, 0, and 1 are analyzed, a trimodal distribution results. Once the sidebands -2, -1, 0, 1, and 2 are analyzed, a monomodal distribution is observed that brackets the values of δ and η used to simulate the spectrum in (a). More sidebands result in narrowing the peak in the posterior probability distribution. (c) Plot of the entropy as a function of maximum sideband order analyzed, $|N_{\text{SB}}^{\max}|$. Most of the information in the simulated spectrum is contained in the central 5 sidebands with orders from -2 to 2.

spinning sidebands (for 5 sidebands, sideband orders −2 to 2; for 7 sidebands, sideband order −3 to 3). The variation in optimal relative spinning frequency with signal-to-noise ratio suggests that at lower SN it is better to spin so that only 5 sidebands are observed whereas in more easily observed samples, it is better to spin at a slightly lower frequency so that 7 sidebands are visible. The choice of the optimal relative spinning frequency will be determined by the ease with which the spectrum can be acquired. For samples with long relaxation times or very low gyromagnetic ratios, it is better to choose slightly faster spinning frequencies, $\upsilon_\delta/\upsilon_r \approx 2$, because of the difficulty in acquiring high signal-to-noise spectra. For samples where the signal-to-noise ratio is not a problem, slightly lower spinning frequencies, $\upsilon_\delta/\upsilon_r \approx 2.5$, will allow the quickest acquisition of accurate values of δ and η.

In some samples, it is not possible to measure all potentially observable sidebands due to spectral overlap or other sample related problems. For this reason, the information content as a function of the number of sidebands has been analyzed. In Fig. 5, the 2D posterior probability distribution and the information content of the spinning sideband pattern changes as more inner sidebands are analyzed. Fig. 5a shows the same simulated sideband pattern as analyzed in Fig. 1 with the sideband orders labeled. Fig. 5b shows the progression of the 2D posterior probability function as more sidebands are analyzed. When sidebands −1 to 1 are analyzed, the probability distribution is trimodal with peaks near $\delta = 48$ and $\eta = 0$, $\delta = 77$ and $\eta = 0.45$, and $\delta = 85$ and $\eta = 0.3$. When sidebands −2 to 2 are analyzed the probability distribution becomes monomodal with $\delta = 79$ and $\eta = 0.4$, close to the values used to simulate the spectrum. As more sidebands are analyzed the distribution becomes narrower but it is clear that most of the information is present in the −2 to 2 sidebands. This is confirmed in Fig. 5c, where the entropy of these distributions is plotted as a function of maximum sideband order, $|N_{SB}^{max}|$, analyzed. When only the centerband is analyzed, there is no information on the CSA parameters so the entropy has the maximum value of 100. As more sidebands are added to the analysis, the entropy quickly drops to the limiting value of 51. When the −2 to 2 sidebands are analyzed, the entropy has already reached a value of about 55 indicating that most of the information present in the sideband pattern has been extracted.

Fig. 6 shows the dependence of the total entropy on the number of inner sidebands, $\upsilon_\delta/\upsilon_r$, and the signal-to-noise ratio. At a signal-to-noise ratio of 50, for $\upsilon_\delta/\upsilon_r$ between 1 and 3, the entropy approaches a constant value for 3 or more central sidebands (sideband orders −1 to 1) indicating that most of the information is in the central 3 sidebands at these relative spinning frequencies. For $\upsilon_\delta/\upsilon_r$ between 4 and 5.5, the number of sidebands necessary to get most of the information present in the spectrum increases to 5 (sideband orders −2 to 2), but the entropy and thus the uncertainty in the CSA parameters has slightly increased in this region. For $\upsilon_\delta/\upsilon_r$ from 5.5 to 6 (the limit of my simulation), another entropy minimum is observed but now at least 7 (sideband orders −3 to 3) inner sidebands need to be analyzed. As the signal-to-noise ratio increases to 250 and then 500, three significant changes occur in these entropy plots. First, the minimum value of $\upsilon_\delta/\upsilon_r$ to minimize the entropy increases to above 1.5. Second, the secondary minimum for $\upsilon_\delta/\upsilon_r$ greater than 5.5 does not appear. Third, a new entropy minimum occurs at $\upsilon_\delta/\upsilon_r$ between 2.5 and 3. In this region, most of the CSA information is contained in the central 5 sidebands. This con-

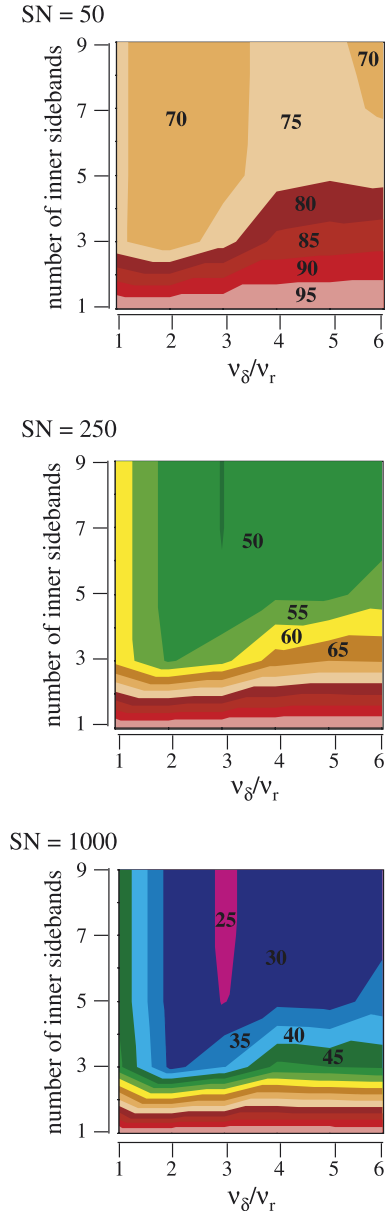

Fig. 6. The entropy as a function of $\upsilon_\delta/\upsilon_{rot}$, the number of sidebands analyzed, and the signal-to-noise ratio. Sidebands were always analyzed symmetrically about the isotropic peak so that the number of inner sidebands $= 2|N_{SB}^{max}| + 1$. The information content depends on the spinning frequency, the signal-to-noise ratio and the number of sidebands. The numbers in these plots indicate the value of the entropy where 0 means perfect information and 100 means no information about the CSA parameters.

firms the previous argument that at low signal-to-noise ratios increasing the spinning frequency so that $\upsilon_\delta/\upsilon_r$ is about 2 increases the reliability of the CSA parameters; however, at higher signal to noise ratios, lower frequency sample reorientation can be used.

5. Conclusions

This paper demonstrates that Bayesian analysis is a powerful and intuitive method of interpreting spinning sideband patterns in solid-state NMR. It produces a map of probability as a function of δ and η where a peak in the probability indicates likely values of the parameters. One-dimensional confidence intervals are easily determined from these posterior probability maps. Such an analysis requires no assumption about the distribution of δ and η, only that the noise in the sideband intensity is distributed as a Gaussian. Thus, non-Gaussian distributions were observed for the value of η in some samples because of its limited range. In addition, multimodal distributions were seen when the number of sidebands was limited. Such a situation could occur in samples where overlapping sidebands and fast relaxation times conspire to prevent the measurement of all sidebands and the use of 2D-PASS. Such multimodal distributions should not be interpreted as "false minima" in the χ^2 distribution, but rather as the current information is not sufficient to uniquely define that values of δ and η. Bayesian analysis in such a situation can allow the extraction of some information that could then further be limited by chemical information about the sample. Ideally, this method could be used to extract the information from an entire 2D-PASS in the time domain. Such a processing procedure would avoid the shearing and Fourier transformation procedure used currently to produce a spectrum, but would instead skip straight from the time-domain data to a map of the probability distribution of the chemically relevant chemical shift parameters. This is currently not practical due to computer limitations, which will disappear as computers continue to improve.

Analysis of the information content of spinning sideband patterns consistently shows that the spinning frequency should be adjusted such that $\upsilon_\delta/\upsilon_r$ is between 2 and 3 for all signal-to-noise ratios investigated. At lower signal-to-noise ratios, slightly faster spinning frequencies are indicated, $\upsilon_\delta/\upsilon_r$ closer to 2, while at higher signal-to-noise ratios, lower spinning frequencies will result in the most accurate determinations of δ and η per unit experiment time. Simulations demonstrate that most of the information about the values of δ and η is in the central five sidebands and that this information is dependent on the signal-to-noise ratio and the value of $\upsilon_\delta/\upsilon_r$.

The Bayesian model fitting presented in this paper is equivalent to a full analysis of the χ^2 surface of a fit to the CSA parameters. Both analyses provide unbiased estimators of the parameter values, and contours on both surfaces can be defined that bound parameter values with a given confidence. Many times, in traditional least squares fitting, i.e., when a full χ^2 surface is not analyzed, the χ^2 surface is assumed to be parabolic near its minimum. This is equivalent to assuming that the parameters are distributed as a Gaussian, which this paper has shown is not always true. The Cramer-Rao inequality [8] can be used to define a lower bound on the uncertainty without any assumptions about the shape of the χ^2 surface, but the true uncertainty can be larger than this value. While traditional least squares and Bayesian analysis provide the same information, Bayesian analysis provides an intuitively interpretable probability distribution which can easily be mapped into the entropy so that the effects of experimental parameters can be investigated. In addition, the Bayesian approach can be extended so that situations that are less well defined can be approached; i.e., when hypothesis testing is necessary or when a distribution of parameters occurs.

Acknowledgments

I thank the Materials Research Laboratory at the University of California, Santa Barbara, the Campus Chemical Instrument Center at the Ohio State University, and the Department of Chemistry and Biochemistry at Otterbein College, for their support during the course of this research. Anthony Cheetham and Gérard Férey supplied the MIL-8 sample. I would like to thank Ted Clark, Janet Gaba, Philip Grandinetti, Claus Jeppesen, and Theodore Stanford for useful comments and discussions.

References

[1] U. Haeberlen, High Resolution NMR in Solids: Selective Averaging, Academic Press, New York, 1976.
[2] M. Mehring, High Resolution NMR Spectroscopy in Solids, Springer, Berlin, 1983.
[3] M.M. Maricq, J.S. Waugh, NMR in rotating solids, J. Chem. Phys. 70 (1979) 3300–3316.
[4] J. Herzfeld, A.E. Berger, Sideband intensities in NMR spectra of samples spinning at the magic angle, J. Chem. Phys. 73 (1980) 6021–6030.
[5] N.J. Clayden, C.M. Dobson, L.-Y. Lian, D.J. Smith, Chemical-shift tensor analyses and simulations of slow-spinning MAS NMR spectra, J. Magn. Reson. 69 (1986) 476–487.
[6] D. Fenzke, B. Maess, H. Pfeifer, A novel method to determine the principal values of a chemical-shift tensor from MAS NMR powder spectra, J. Magn. Reson. 88 (1990) 172–176.
[7] A.C. Olivieri, Rigorous statistical analysis of errors in chemical-shift-tensor components obtained from spinning sidebands in solid-state NMR, J. Magn. Reson. A 123 (1996) 207–210.
[8] P. Hodgkinson, L. Emsley, The reliability of the determination of tensor parameters by solid-state nuclear magnetic resonance, J. Chem. Phys. 107 (1997) 4808–4816.
[9] G.L. Bretthorst, J.J. Kotyk, J.J.H. Ackerman, P-31 NMR Bayesian spectral-analysis of rat-brain invivo, Magn. Reson. Med. 9 (1989) 282–287.
[10] G.L. Bretthorst, Bayesian-analysis. 1. Parameter-estimation using quadrature NMR models, J. Magn. Reson. 88 (1990) 533–551.
[11] G.L. Bretthorst, Bayesian-analysis. 3. Applications to NMR signal-detection, model selection, and parameter-estimation, J. Magn. Reson. 88 (1990) 571–595.
[12] G.L. Bretthorst, Bayesian-analysis. 5. Amplitude estimation for multiple well-separated sinusoids, J. Magn. Reson. 98 (1992) 501–523.

[13] C. Antz, K.P. Neidig, H.R. Kalbitzer, A general bayesian method for an automated signal class recognition in 2D NMR-spectra combined with a multivariate discriminant-analysis, J. Biomol. NMR 5 (1995) 287–296.
[14] M. Andrec, G.T. Montelione, R.M. Levy, Estimation of dynamic parameters from NMR relaxation data using the Lipari-Szabo model-free approach and bayesian statistical methods, J. Magn. Reson. 139 (1999) 408–421.
[15] R.L. Dunbrack, F.E. Cohen, Bayesian statistical analysis of protein side-chain rotamer preferences, Protein Sci. 6 (1997) 1661–1681.
[16] H.W. Long, R. Tycko, Biopolymer conformational distributions from solid-state NMR: α-helix and 3_{10}-helix contents of a helical peptide, J. Am. Chem. Soc. 120 (1998) 7039–7048.
[17] R. Havlin, M. McMahon, R. Srinivasan, H. Le, E. Oldfield, Solid-state NMR and density functional investigation of carbon-13 shielding tensors in metal–olefin complexes, J. Phys. Chem. A 101 (1997) 8908–8913.
[18] D.S. Sivia, Data Analysis: A Bayesian Tutorial, Clarendon Press, Oxford, 1996.
[19] O.N. Antzutkin, S.C. Shekar, M.H. Levitt, 2-Dimensional side-band separation in magic-angle-spinning NMR, J. Magn. Reson. A 115 (1995) 7–19.
[20] F. Serpaggi, G. Férey, Hybrid open frameworks (MIL-n). Part 4: Synthesis and crystal structure of MIL-8, a series of lanthanide glutarates with an open framework, $[Ln(H_20)]_2[O_2C(CH_2)_3CO_2]_34\text{-}H_2O$, J. Mater. Chem. 8 (1998) 2737–2741.
[21] Courtesy of Professor P. Grandinetti, Ohio State University.
[22] M. Edén, M.H. Levitt, Computation of orientational averages in solid state NMR by Gaussian spherical quadrature, J. Magn. Reson. 132 (1998) 220–239.
[23] F. Serpaggi, T. Luxbacher, A.K. Cheetham, G. Férey, Dehydration and rehydration processes in microporous rare-earth dicarboxylates: a study by thermogravimetry, thermodiffractometry and optical spectroscopy, J. Solid State Chem. 145 (1999) 580–586.