Chapter 8

Maximum Entropy Methods in Multidimensional NMR

Jeffrey C. Hoch[1] and Mehdi Mobli[2]

[1]*University of Connecticut Health Center, Farmington, CT, USA*
[2]*University of Queensland, St. Lucia, Queensland, Australia*

8.1 INTRODUCTION

The discrete Fourier transform (DFT) played a seminal role in the development of modern nuclear magnetic resonance (NMR) spectroscopy.[1] Nevertheless, it has a number of well-known limitations. Chief among them is the difficulty of obtaining high resolution spectral estimates from short time records, because the ability to resolve signals with closely spaced frequencies is largely determined by the longest evolution time sampled. Padding the free induction decay (FID) with zeros to extend it in time increases the digital resolution of the DFT spectrum, at the expense of truncation artifacts ("sinc" wiggles) around components that have not decayed close to zero by the end of the sampling time. A veritable zoo of apodization functions[2] was developed to minimize these artifacts, but always at the expense of broader lines in the spectrum.

The ability to obtain accurate, high resolution spectral estimates from short data records is critical in many applications of NMR spectroscopy because the available sampling time is limited, for example due to sample instability or simply due to constraints on available instrument time. In practice, the latter is mainly encountered in multidimensional NMR experiments, where the data collection time is directly proportional to the number of data samples collected in the indirect time dimensions (indirect time dimensions correspond to time delays between RF pulses; real time is referred to as the *acquisition dimension*). Furthermore, at very high magnetic field, the competition between the goals of short data collection time and high resolution becomes more severe. The bandwidth spanned by the nuclear resonances increases linearly with field strength, necessitating a decrease in the time between samples in order to avoid aliasing.[2] This decrease necessitates an increase in the number of data samples just to maintain the same

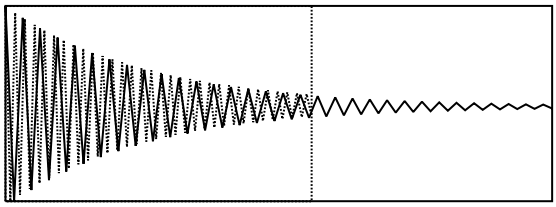

Figure 8.1. Sixty-four data points shown for a single sinusoid at 500 MHz (solid) and 900 MHz (dashed). The chemical shift for both signals would be identical, but the lower field allows for the collection of a longer (in time) data record thus minimizing truncation artifacts.

spectral resolution (Figure 8.1). Thus, an experiment conducted at 900 MHz must collect a factor of 9/5 more data in each dimension than the comparable experiment conducted at 500 MHz to maintain resolution. For a 3D experiment, this requires 3.3 times more data acquisition time; 5.8 times longer is required for a 4D experiment. In practice, these increases in data collection time are rarely realized.

Non-Fourier methods of spectrum analysis provide an avenue to high resolution spectral estimates from short data records. One of the first to enjoy widespread application in NMR was the use of linear prediction (LP) extrapolation to extend experimental FIDs beyond the sampled interval (see Chapter 10).[3] LP extrapolation implicitly assumes that the signal consists of exponentially decaying sinusoids. Nearly contemporaneous with the development of LP extrapolation, a nonparametric method of spectrum analysis called *maximum entropy* (*MaxEnt*) reconstruction was developed.[4] Unlike LP extrapolation, MaxEnt reconstruction makes no assumption regarding the nature of the signal, and thus it is applicable to signals with non-Lorentzian line shapes. MaxEnt reconstruction computes the spectrum $\boldsymbol{f}$ that maximizes a regularization functional (the entropy) subject to the constraint that the inverse DFT of the spectrum is consistent with the measured FID. Although solving this constrained optimization is computationally demanding, robust algorithms have been developed that are capable of computing large, multidimensional NMR spectra.[5,6] Over the past three decades, a host of additional non-Fourier methods of spectrum analysis have been developed, including maximum likelihood[7] and Bayesian methods,[8] the filter diagonalization method (FDM; see Chapter 9),[9] G-matrix Fourier transform (GFT),[10] back-projection reconstruction[11] (see Chapter 5), and multidimensional decomposition (MDD; see Chapter 6).[12] These methods span a continuum of assumptions about the nature of the signal, and restrictions (or lack thereof) on the characteristics of the data sampling. MaxEnt reconstruction lies at the extreme of making few assumptions about the signal, and furthermore can be applied to data collected in essentially arbitrary fashion. Thus, although it is not always the most powerful approach, it is the most general one. It also has been subject to more investigation, both theoretical and empirical, than the other methods, and our understanding of its behavior in various regimes of signal-to-noise (S/N) ratio, dynamic range, and signal decay (or lineshape) is better understood.[13–15] The focus of this article is the development, theory, and application of MaxEnt reconstruction in multidimensional NMR spectroscopy.

8.2 BACKGROUND

The concept of information entropy originated in the work of Claude Shannon on the information capacity of communication channels. The Shannon entropy,[16]

$$S = \sum_{i=1}^{N} p_i \log p_i \tag{8.1}$$

where p_i is the ith element of a probability distribution, describing the message or the image, is a measure of the amount of uncertainty or "missing" information at the receiving end of a communication channel. The maximum entropy principle applied to signal processing is that the method used to compute the spectrum from time series data should add the least amount of information. Alternatively, the information contained in a spectrum should be only that conveyed by the (available) data. An early practical application of the principle to spectrum analysis, and one of the first to be applied to NMR data,[17] is the method developed by John Burg.[18] His method approached the question of how to extrapolate a time series in a way that maximizes the Shannon entropy. The solution, which exploits the fact that the power spectrum is positive definite, involves fitting an LP model to the data and computing the power spectrum from the resulting coefficients. Although the Burg approach is not suitable for complex spectra or spectra containing positive and negative components, it

nevertheless demonstrated the power of the maximum entropy principle for signal processing.

A more powerful and general application of the maximum entropy principle was described in 1977 by Wernecke and D'Addario.[4] In their approach, trial spectra were generated and then tested for consistency with the time domain data following inverse Fourier transformation. Their approache amounts to solving a constrained optimization: maximizing the entropy (in the frequency domain) while enforcing consistency with the measured data in the time domain. This inverse approach enables essentially arbitrary sampling (unlike the Burg method that requires uniform sampling) and applications such as deconvolution. Unfortunately, the disparate natures of the time domain constraint and the frequency domain entropy objective function proved overly challenging for the conjugate-gradient approach employed by Wernecke and D'Addario, and practical application of their inverse approach required the development of more robust numerical optimization algorithms. The development of the "Cambridge" algorithm by Skilling and Bryan,[5] which is not merely robust but also highly efficient, effectively launched the modern application of the maximum entropy principle in NMR.

The earliest applications of MaxEnt reconstruction were to photometric images, in which the reconstructed image is everywhere positive or zero. With appropriate normalization, the Shannon entropy [equation (8.1)] applies. Modern NMR spectrometers employ (or emulate) phase-sensitive detection, yielding complex-valued spectra that furthermore can contain both positive and negative components. Clearly, the Shannon entropy cannot be applied to such spectra. The first attempt to apply the maximum entropy principle to reconstruct complex spectra adopted the approach of defining multiple channels (subspectra) that are constrained to be real and positive.[19] The Shannon entropy could be applied separately to each channel; the final spectrum was defined as the difference between the "positive" and the "negative" channels for the real and the imaginary parts of the spectrum:

$$\boldsymbol{f} = \boldsymbol{f}_+^r + \boldsymbol{f}_+^i - \boldsymbol{f}_-^r - \boldsymbol{f}_-^i \tag{8.2}$$

where each component of the spectrum is nonnegative. The net entropy is then simply the sum of the Shannon entropy for each component. Problems with this approach are that it imposes a phase bias and effectively doubles the number of variables to be optimized. Subsequently, Daniell and Hore[20] and Hoch *et al*.[21] separately (and by different approaches) arrived at an entropy functional suitable for phase-sensitive NMR data:

$$S(\boldsymbol{f}) = \sum_{n=0}^{N-1} \frac{|f_n|}{def} \log\left(\frac{|f_n|/def + \sqrt{4 + |f_n|^2/def^2}}{2}\right) - \sqrt{4 + |f_n|^2/def^2} \tag{8.3}$$

In the more elegant derivation by Hore *et al*., on the basis of the entropy of an ensemble of spin-1/2 particles, the parameter *def* is revealed to be related to the sensitivity of the NMR spectrometer.

A defining characteristic of MaxEnt reconstruction is that it is nonlinear. Although nonlinearity makes it a powerful tool for applications such as deconvolution, where linearity can result in instability, nonlinearity has important implications for quantitative applications such as nuclear Overhauser effect or relaxation measurements and for difference spectroscopy. Nonlinearity also means that standard metrics such as S/N ratio are no longer reliable indicators of sensitivity.[13] A consequence is that special precautions need to be taken when MaxEnt reconstruction is employed for quantitative applications. These are described later.

8.3 THEORY

Although the underpinnings of MaxEnt reconstruction can be traced to Shannon, the use of the maximum entropy principle to attack a broad range of statistical problems, including spectral reconstruction, can be largely attributed to Jaynes.[22] Formal derivations of the maximum entropy spectral reconstruction frequently treat the problem as one of finding the spectrum that maximizes the joint probability of the model (the spectrum) and the data, by invoking Bayes's rule with the entropy as prior and the χ^2 statistic as likelihood.[23] This approach has led some of its adherents to make statements such as "maximum entropy is the only logically consistent method of inference," and it is not surprising that to this day the topic of maximum entropy engenders noisy debates among statisticians.[24] We take the pragmatic view that entropy is a useful regularization functional that ensures smooth reconstructions, while the

constraint prevents overfitting. Here, we give only a summary of how the maximum entropy principle is applied to spectral reconstruction, without delving into its motivation. The MaxEnt reconstruction of the spectrum of a complex-valued time series $\boldsymbol{d}$ is the spectrum $\boldsymbol{f}$ that maximizes the entropy $S(\boldsymbol{f})$, subject to the constraint that the mock data $\boldsymbol{m}$, given by the inverse DFT of the spectrum, is consistent with the time series $\boldsymbol{d}$. Consistency is defined by the following condition:

$$C(\boldsymbol{f}, \boldsymbol{d}) \leqslant C_0 \tag{8.4}$$

where $C(\boldsymbol{f}, \boldsymbol{d})$ is the unweighted χ^2 statistic,

$$C(\boldsymbol{f}, \boldsymbol{d}) = \sum_{i=0}^{M-1} |m_i - d_i|^2 = \sum_{i=0}^{M-1} |\mathrm{i\ DFT}(\boldsymbol{f})_i - d_i|^2 \tag{8.5}$$

C_0 is an estimate of the noise level, and iDFT is the inverse DFT. The definition of the entropy $S(\boldsymbol{f})$ applicable to complex-valued spectra is given above [equation (8.3)]. In principle, the quantum-mechanical derivation prescribes the value of *def* (it depends on the sensitivity of the spectrometer and the number of spins in the sample), but it is more convenient to treat *def* as an adjustable parameter. Essentially, it determines the scale at which the nonlinearity of MaxEnt becomes pronounced. The MaxEnt solution is found by maximizing the objective function:

$$O = S(\boldsymbol{f}) - \lambda C(\boldsymbol{f}, \boldsymbol{d}) \tag{8.6}$$

where the value of the Lagrange multiplier λ is adjusted to obtain $C = C_0$. Standard optimizers such as the steepest descent or conjugate gradient are typically subject to premature convergence when applied to this problem. Efficient maximization of the objective function has been demonstrated using the "Cambridge" algorithm[5] and its extensions,[2] which employ concepts of variable metrics suitable for the entropy and the constraint applied to search step lengths in a subspace defined by first, second, and various mixed gradients of $S(\boldsymbol{f})$ and $C(\boldsymbol{f}, \boldsymbol{d})$. Details of the search algorithms are given elsewhere,[2] and codes that implement the algorithms are freely available.[25] The value of λ depends on the values of the parameters *def* and C_0 and on the data. An automated procedure for determining useful values for λ and *def* has been described.[26,27]

In general, MaxEnt reconstruction is nonlinear. The nonlinearity behaves much like an intensity-dependent scaling, in which weak spectral features are scaled down more than strong features.[13] If steps are taken to ensure that the nonlinearity is uniform across the entire spectrum, either by computing the MaxEnt reconstruction of the full nD spectrum, or piecewise reconstruction (e.g., row- or plane-wise) is performed using constant λ, then the nonlinearity can be calibrated and corrected.[28] Alternatively, the value of C_0 can be made small, tightly constraining the mock data to match the measured data. In this regime, MaxEnt reconstruction is linear, although the resulting spectra are overfit to the data.

8.4 NONUNIFORM SAMPLING OF NMR DATA

As discussed above, a principal advantage afforded by non-Fourier processing methods is the ability to obtain artifact-free spectral data from short data records (Figure 8.2). Although the ability of the maximum entropy method to suppress truncation artifacts is well established,[29] short data records inevitably lead to poor frequency resolution because two sinusoids oscillating at slightly different frequencies will not distinguish themselves significantly at short times.[30] Here, we must now distinguish a short data record, for which the maximum evolution time is short (see Chapters 9 and 10), from one that contains few samples but for which the maximum evolution time is sufficiently long to afford high resolution (see Chapters 5–7). The Nyquist sampling theorem, which dictates the maximum sampling interval sufficient to disambiguate signals spanning a given frequency bandwidth, and the requirement of the DFT for uniformly spaced samples, jointly conspire to ensure that short data records also have short maximum evolution times. Collecting samples at nonuniform sampling (NUS) intervals, however, one can sample long evolution times without having to collect all of the intervening data (as required by the Nyquist Theorem and the DFT). It is clear that this type of data is not well suited for processing using the DFT. In contrast, the inverse nature of MaxEnt reconstruction allows one to compare only those data points that have been experimentally sampled to those in the mock FID. This property of maximum entropy was indeed realized early in the application of maximum entropy to NMR, by Barna *et al*.,[31] who showed that by sampling nonuniformly from an exponential distribution

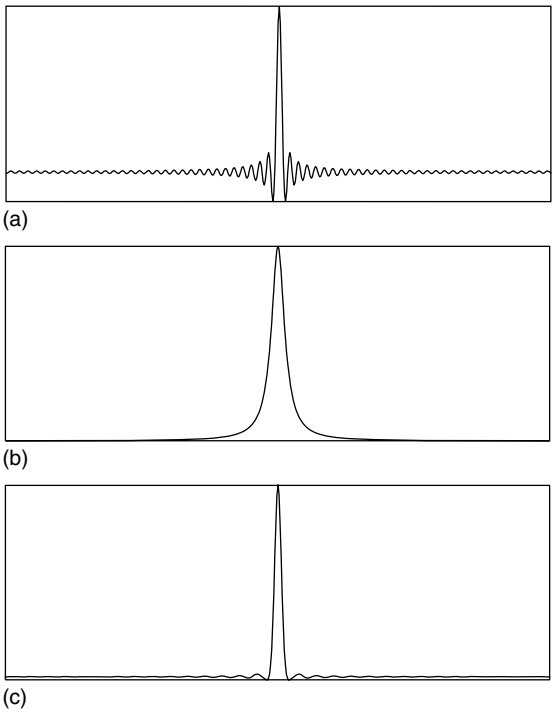

Figure 8.2. The effect of truncation in FT NMR. (a) The DFT of a truncated signal without apodization. (b) The same signal transformed after application of an exponential window function removing the truncation artifacts at the cost of signal broadening. (c) MaxEnt reconstruction of the same signal minimizing the truncation artifacts without signal broadening.

one could increase the spectral resolution in the same time it took to collect a conventional uniform DFT spectrum. The authors subsequently showed that by setting the omitted data values to "0", one could use the DFT to compute a spectrum. This nonuniform DFT (nuDFT) spectrum will, however, contain substantial sampling artifacts.[32] The authors suggested that these may be removed using the CLEAN algorithm, popularly applied in the field of astronomy. In a critical comparison of spectra obtained using nuDFT postprocessed by CLEAN with spectra obtained using MaxEnt reconstruction, they found that maximum entropy produced fewer artifacts. More recently, these algorithms have been applied to multidimensional data, and the connection between the many flavors of reduced dimensionality[33–37] and nuDFT[38] has been realized. It has been repeatedly shown that the artifacts caused by NUS are more satisfactorily removed using maximum entropy than by other postprocessing methods.[39] In the following section, we discuss the nature of the artifacts produced by nuDFT and approaches to suppress them.

8.5 NUS AND SAMPLING ARTIFACTS

From the earliest studies of processing of NUS data in NMR, it has been clear that although NUS is a powerful method for achieving high resolution in a short amount of time, robust methods are required to deal with the associated artifacts. These "sampling artifacts" or "sampling noise" may assume different shapes, patterns, and intensities depending on which evolution times are sampled. The distribution of the artifacts can, however, be predicted from knowledge of the sampled evolution times. If the sampled evolution times fall on a regular grid corresponding to some Nyquist interval, a "sampling spectrum" or "point spread function" (PSF) can be computed by applying the DFT to a sampling function that consists of the value "1" for each evolution time sampled from the grid and the value "0" for evolution times that are not sampled. The PSF will contain a signal at 0 Hz together with a set of artifacts due to incomplete sampling of the grid. The convolution theorem illuminates the connection between the PSF for a given set of sample times (called a *sampling schedule*) and the artifacts in the nuDFT spectrum. NUS data can be viewed as the element-by-element product (called the *array product*) of a uniformly sampled data set and the sampling schedule. According to the convolution theorem, the nuDFT spectrum is the convolution of the PSF with the uniform DFT spectrum (Figure 8.3). Thus in the nuDFT spectrum, a set of artifacts described by the PSF will be distributed about each peak in the experimental spectrum. The artifacts produced due to nonuniform sampling are deterministic and not random, suggesting it may be feasible to remove them through some deconvolution procedure. This is precisely what is done using the CLEAN algorithm. CLEAN (which is very similar to the idea of reference deconvolution[40]) works by finding the most prominent peaks in the spectrum and assuming that the sampling noise will be lower than this (as can be verified by the PSF) amplitude. A spectrum is now constructed (e.g., using a Lorentzian model) containing only the picked peak. Inverse DFT is then followed by multiplication of the resulting time domain signal with the sampling function. DFT of this signal

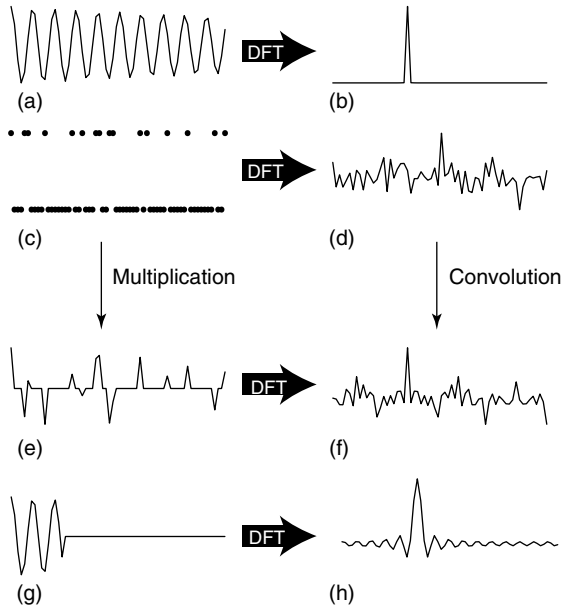

Figure 8.3. The DFT of a decaying sinusoid (a) and (b) and a sampling schedule (c) and (d) and their multiplication in the time domain (e), resulting in their convolution in the frequency domain (f). Uniform sampling of the same number of points (g) resulting in a truncated signal and associated artifacts (h).

will produce a spectrum with artifacts characteristic for the peak that was originally picked. This spectrum is now subtracted from the original spectrum. The procedure is repeated until some predetermined threshold is reached (generally the estimated noise level). The large artifacts associated with strong peaks are removed in this procedure, thus revealing weaker peaks. The procedure is slightly sensitive to the initial modeling of the peaks and works best in cases where S/N ratio is high and signal density is low.[32]

An alternative to deconvolution of sampling artifacts is to design the sampling schedule in such a way as to minimize the sampling artifacts in the region of interest.[38] Thus, this optimization would require some assumptions regarding the frequency distribution of the signals, and the sampling noise is moved rather than removed in this case.

Deconvolution of the PSF from a spectrum remains one of the principal tasks in dealing with sampling artifacts when using NUS. The other is the design of sampling schemes that minimize the artifacts in the first place.

8.6 NUS IN nD SPECTROSCOPY

Despite the development of magnets producing ultrahigh magnetic fields, the resolution afforded in one dimension is insufficient to resolve overlap of resonances for biological macromolecules. The ability to isotopically label biopolymers has enabled the development of a multitude of multidimensional experiments to overcome this problem. In contrast to the acquisition or real-time dimension, for which the time cost of collecting long data records is usually inconsequential because of the need to allow for relaxation, the time cost of collecting samples in the indirect dimensions is *directly proportional* to the number of samples. As discussed in the Introduction, the potential resolution afforded by ultrahigh field magnets is not realized in the indirect dimensions of multidimensional experiments when linear sampling is employed. Therefore, it is not surprising to see the explosion in methods applying various incarnations of NUS to 3D, 4D, and higher dimensional experiments.[41–45]

In the case of 2D data, a 1D sampling schedule is required (since the direct dimension is collected in real time), and the reconstruction can then be performed as a series of 1D reconstructions row-wise (perpendicular to the direct dimension). If the constraint statistic, C_0, is kept constant between rows, it can lead to variation in the weighting (λ) of this constraint and the entropy. This is due to natural variations in signal and noise distribution, which affect the entropy. This scenario will introduce small changes in the reconstruction between rows and may have a significant effect on peak shapes. By using a constant value for the weighting, λ, one can minimize the variation of the nonlinearity between rows.[28] A good estimate of λ can be made by finding representative rows where the constraint statistic $C(\boldsymbol{f}) = C_0$ is satisfied and using the value of λ found for these rows to perform the complete reconstruction.[45,46] The same basic strategy can be applied to higher dimensions, i.e., a 3D spectrum can be constructed as a series of 2D plane reconstructions. In a 3D experiment, it is generally desirable to distribute the sample points nonuniformly in both indirect dimensions, rather than one. When NUS is applied to only one dimension of a multidimensional spectrum, the sampling artifacts become "coherent" in the dimensions orthogonal to the NUS dimension, which can amplify their prominence.[15]

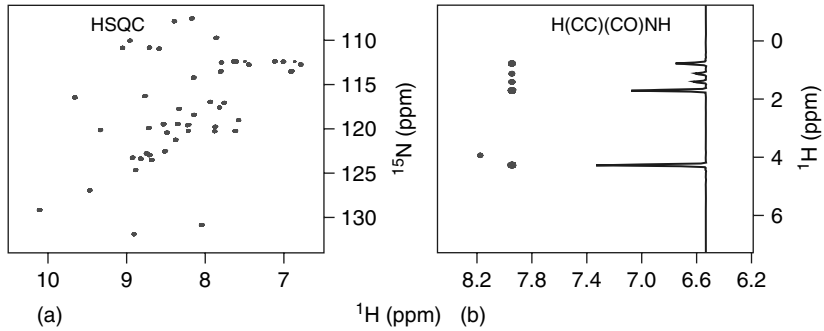

Figure 8.4. ^{15}N-^{1}H HSQC spectrum and indirect plane of an H(CC)(CO)NH experiment acquired on a 0.3 mM sample of 4.5 kDa toxin measured at 900 MHz equipped with a cryogenically cooled probed. Both spectra collected using NUS (~10% in each case) and reconstructed using the 1D (HSQC) and 2D (H(CC)(CO)NH) maximum entropy algorithm.

Figure 8.4 gives an example of the gains in resolution that can be obtained using NUS. These algorithms for reconstructing 2D and 3D data are currently being used routinely in several laboratories and shown to produce impressive results in a fraction of the time required by conventional sampling.

8.7 DESIGN OF SAMPLING SCHEDULES

In the pioneering work of Barna *et al*.,[47] samples were chosen from a uniform grid (with spacing corresponding to the Nyquist frequencies along each dimension, hence sometimes referred to as a *Nyquist grid*) according to an exponentially decaying sampling probability density. The resulting schedule samples the signal more frequently at short evolution times, when the signal is largest, and less frequently at long evolution times. This strategy invites comparison to the matched filter, which, as Ernst demonstrated,[67] is the apodization function that optimizes S/N of the DFT spectrum. The analogy to the matched filter can be extended to encompass experiments that elicit sine-modulated signals[48] or employ constant-time evolution periods.[49] This approach explicitly seeks to optimize sensitivity per unit time while maintaining resolution.

Methods employing coupled evolution times, in which two or more indirect evolution times are jointly incremented at fixed relative rates, are a subclass of NUS methods. The resulting sampling schemes correspond to collecting data along radial vectors emanating from zero time, and are employed by methods such as back projection reconstruction and GFT. Radial sampling schemes do not fall on a rectangular grid; however, it has been shown that on-grid approximations to radial sampling give essentially equivalent results.[50] Furthermore, it was demonstrated that back projection and MaxEnt reconstruction yield highly similar results, including the patterns and magnitudes of sampling artifacts, demonstrating that the sampling artifacts mainly result from the sampling scheme and not the method used to construct the spectrum. Recent approaches employing radial sampling focus on optimizing resolution, by selecting sets of radial vectors that will optimally resolve cross peaks based on prior knowledge of frequency distributions.[36,51]

In addition to designing sampling schedules to optimize sensitivity and resolution, another important criterion is minimizing sampling artifacts, which can be construed as false peaks. (Strictly speaking, this could also be characterized as optimizing sensitivity, because sensitivity is the ability to distinguish signal from noise.) Early studies that revealed the coherent nature of sampling artifacts when NUS is applied to only one dimension hinted at a general approach to suppress sampling artifacts: the introduction of randomness to cause sampling artifacts to "decohere".[15] Figure 8.5 vividly illustrates the ability of randomness to diminish sampling artifacts. The left-most panels show a MaxEnt reconstruction for radially sampled data (approximated on a Nyquist grid). The central and right-most panels depict the results of applying increasing amounts of random "dither" to the evolution times and the remarkable reduction of sampling artifacts. Randomness appears to be a useful adjunct to any NUS scheme.[40,52–55]

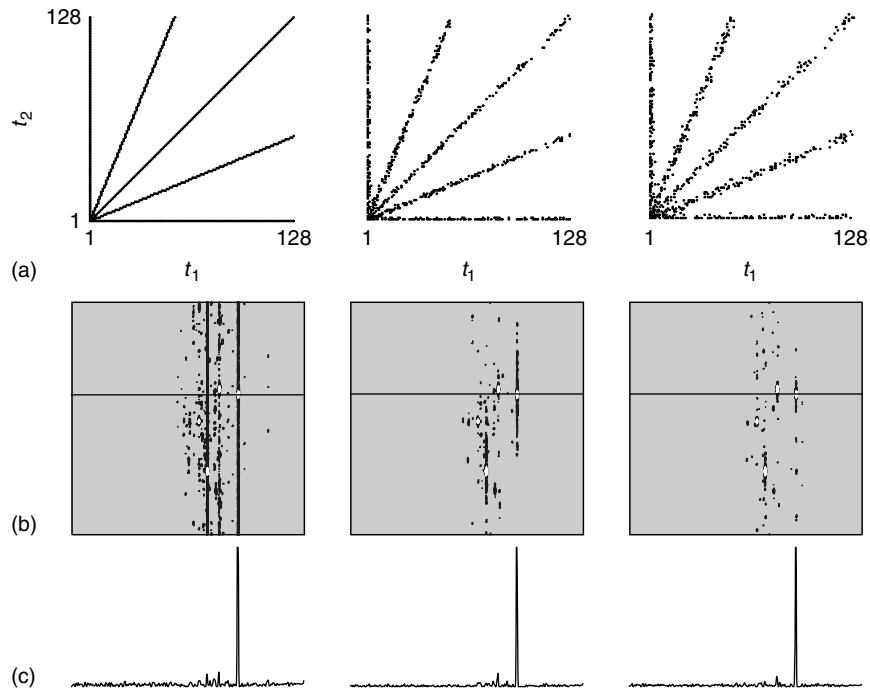

Figure 8.5. $^{1}H/^{13}C$ plane (^{15}N chemical shift 121.96 ppm) from the HNCO spectrum of ubiquitin, using data collected at 9.4 T (400 MHz for ^{1}H) on a Varian Inova instrument. Spectra were computed using MaxEnt reconstruction and radial sampling using five projections with different amounts of random "blurring" of the sampling schedule (RMS zero (none), 0.625 and 1.25, left to right). (a) Sampling schedule. (b) and (c) MaxEnt spectrum. White contour levels are plotted at multiples of 1.4 starting with 3% of the height of the highest peak. Black contours start with 0.5% of the height of the highest peak. 1D cross sections at the frequency indicated by the horizontal line on the contour plots are shown at the bottom.

When NUS is invoked, the Nyquist sampling theorem no longer applies; indeed, some sampling artifacts can be viewed as aliases. Bretthorst has pointed out that arbitrary sampling schemes, even those construed as "off-grid", can be viewed as subsets of regular grids with spacings determined by the precision (the number of significant digits) of the evolution times.[56] Using this ansatz elicits another strategy for minimizing sampling artifacts: selecting NUS evolution times from an oversampled grid. The effect is to shift NUS artifacts that correspond to aliases out of the spectral window containing the signals.[57]

Here, only a few general aspects of sampling schedule design have been discussed. This remains an area of active research and although a unified theory remains elusive, the general outlines of the ongoing efforts are clear from the above themes. It is also clear that the term *optimal* will have different implications and will involve trade-offs, depending on the goal (i.e., sensitivity vs resolution vs experiment time) and the context (including the number of dimensions and the S/N ratio).

8.8 MAXIMUM ENTROPY AND DECONVOLUTION

The convolution theorem holds that multiplication in the time domain is equivalent to deconvolution in the frequency domain. Linear deconvolution of a lineshape function, or convolution kernel, can be achieved by dividing the time domain data by the kernel. For kernels that approach or are equal to zero, this linear approach to deconvolution is unstable, leading (at minimum) to severe noise amplification.[58] In contrast, because MaxEnt reconstruction adopts an inverse approach, deconvolution is performed by

multiplication of the mock FID by the kernel, thus avoiding the instabilities associated with division by small numbers (or zero).[59–61] However, the form of the kernel being deconvolved must be known beforehand. The most common kernels deconvolved in NMR are line shapes (e.g., Lorentzian or Gaussian) and J-coupling (sine- or cosine-modulation). Line shape deconvolution for NMR via MaxEnt was proposed by Sibisi,[59] who applied MaxEnt to a 1D spectrum with varying linewidths. The reconstructions were performed using Lorentzian kernels with varying widths, and the results were used to create a 2D spectrum from a 1D dataset. Later work by Scheraga and colleagues proposed an algorithm for MaxEnt deconvolution of noisy data and suggested that a simultaneous gain in sensitivity and resolution was not possible otherwise.[58] Since these early studies, few examples of lineshape deconvolution using MaxEnt in NMR have been reported. In contrast, the use of MaxEnt to deconvolve J-couplings is more widely reported.[62–66] The method has found useful application for removing splittings that are constant, for example for removing a heteronuclear coupling, provided that the variation in couplings is comparable to the natural linewidth or smaller.[66] To remove splittings that vary substantially, it is necessary to construct a series of spectra using different values for the J-coupling.[62] The results can contain spurious peaks due to incomplete deconvolution, and more sophisticated iterative algorithms are generally warranted for extracting the values of coupling constants.[63–65]

8.9 CONCLUDING REMARKS

MaxEnt reconstruction is just one of many methods capable of processing NUS data. Some require very specific (though nonuniform) sampling patterns, whereas others can be used with essentially arbitrary sampling schemes. They span a continuum of assumptions about the signals contained in the data, ranging from no assumptions at all to the assumption that the signal consists of exponentially damped sinusoids plus random noise. Each has strengths and weaknesses, but direct critical comparison of different approaches has been elusive. Methods that make prior assumptions about the signal can outperform MaxEnt reconstruction, provided that the assumptions remain valid, but these methods can be less than robust when the assumptions are violated, even for a fraction of the signal(s). The robustness of MaxEnt reconstruction, together with its ability to accommodate essentially arbitrary sampling schemes and perform stable deconvolution, ensures that it will remain a valuable tool in multidimensional NMR spectroscopy, and its versatility will likely continue to find new applications.

RELATED ARTICLES IN THE ENCYCLOPEDIA OF MAGNETIC RESONANCE

Fourier Transform Spectroscopy

Maximum Entropy Reconstruction

Reference Deconvolution

REFERENCES

1. R. R. Ernst, Without computers—No modern NMR, in *'Computational Aspects of the Study of Biological Macromolecules by Nuclear Magnetic Resonance Spectroscopy'*, eds J. C. Hoch, F. M. Poulsen and C. Redfield, Plenum Press: New York, 1991, pp 1–25.
2. J. C. Hoch and A. S. Stern, *'NMR Data Processing'*, Wiley-Liss: New York, 1996.
3. F. Ni and H. A. Scheraga, *J. Magn. Reson.*, 1986, **70**, 506–511.
4. S. J. Wernecke and L. R. D'Addario, *IEEE Trans. Comput.*, 1977, **26**, 351–364.
5. J. Skilling and R. Bryan, *Mon. Not. R. Astron. Soc.*, 1984, **211**, 111–124.
6. J. C. Hoch and A. S. Stern, Maximum entropy reconstruction in NMR, in *'Encyclopedia of Magnetic Resonance'*, eds R. K. Harris and R. E. Wasylishen, John Wiley & Sons: Chichester. DOI:10.1002/9780470034590.emrstm0299.
7. R. A. Chylla and J. L. Markley, *J. Biomol. NMR*, 1995, **5**, 245–258.
8. G. L. Bretthorst, *J. Magn. Reson.*, 1990, **88**, 533–551.
9. V. A. Mandelshtam, H. S. Taylor, and A. J. Shaka, *J. Magn. Reson.*, 1998, **133**, 304–312.
10. S. Kim and T. Szyperski, *J. Am. Chem. Soc.*, 2003, **125**, 1385–1393.
11. Ē. Kupče and R. Freeman, *J. Am. Chem. Soc.*, 2003, **125**, 13958–13959.
12. V. Y. Orekhov, I. V. Ibraghimov, and M. Billeter, *J. Biomol. NMR*, 2001, **20**, 49–60.

13. D. L. Donoho, I. M. Johnstone, A. S. Stern, and J. C. Hoch, *Proc. Natl. Acad. Sci. U.S.A.*, 1990, **87**, 5066–5068.
14. J. A. Jones and P. J. Hore, *J. Magn. Reson.*, 1991, **92**, 363–376.
15. A. S. Stern, K.-B. Li, and J. C. Hoch, *J. Am. Chem. Soc.*, 2002, **124**, 1982–1993.
16. C. E. Shannon, *Bell Syst. Tech. J.*, 1948, **27**, 379–423.
17. J. F. Martin, *J. Magn. Reson.*, 1985, **65**, 291–297.
18. J. P. Burg, *'Maximum Entropy Spectrum Analysis'*, Stanford University, Stanford, CA, 1975.
19. J. C. J. Barna, E. D. Laue, M. R. Mayger, J. Skilling, and S. J. P. Worrall, *Biochem. Soc. Trans.*, 1986, **14**, 1262–1263.
20. G. J. Daniell and P. J. Hore, *J. Magn. Reson.*, 1989, **84**, 515–536.
21. J. C. Hoch, A. S. Stern, D. L. Donoho, and I. M. Johnstone, *J. Magn. Reson.*, 1990, **86**, 236–246.
22. E. T. Jaynes, Where do we stand on maximum entropy? in *'The Maximum Entropy Formalism'*, eds R. D. Levine and M. Tribus, MIT Press: Cambridge, MA, 1979, pp 15–118.
23. J. Skilling, Classic maximum entropy, in *'Maximum Entropy and Bayesian Methods'*, ed. J. Skilling, Kluwer: Norwell, MA, 1989, pp 45–52.
24. D. L. Donoho, I. M. Johnstone, A. S. Stern, and J. C. Hoch, *J. Roy. Stat. Soc. B*, 1992, **54**, 41–81.
25. A. S. Stern and J. C. Hoch, *The Rowland NMR Toolkit* http://rnmrtk.uchc.edu.
26. M. Mobli, M. W. Maciejewski, M. R. Gryk, and J. C. Hoch, *Nat. Methods*, 2007, **4**, 467–468.
27. M. Mobli, M. W. Maciejewski, M. R. Gryk, and J. C. Hoch, *J. Biomol. NMR*, 2007, **39**, 133–139.
28. P. Schmieder, A. S. Stern, G. Wagner, and J. C. Hoch, *J. Magn. Reson.*, 1997, **125**, 332–339.
29. S. Sibisi, J. Skilling, R. G. Brereton, E. D. Laue, and J. Staunton, *Nature*, 1984, **311**, 446–447.
30. S. J. Davies, C. Bauer, P. J. Hore, and R. Freeman, *J. Magn. Reson. (1969)*, 1988, **76**, 476–493.
31. J. C. J. Barna, E. D. Laue, M. R. Mayger, J. Skilling, and S. J. P. Worrall, *J. Magn. Reson. (1969)*, 1987, **73**, 69–77.
32. J. C. J. Barna, S. M. Tan, and E. D. Laue, *J. Magn. Reson.*, 1988, **78**, 327.
33. T. Szyperski, G. Wider, J. H. Bushweller, and K. Wüthrich, *J. Am. Chem. Soc.*, 1993, **115**, 9307–8308.
34. K. Ding and A. M. Gronenborn, *J. Magn. Reson.*, 2002, **156**, 262–268.
35. Ē. Kupče and R. Freeman, *Concepts Magn. Reson.*, 2004, **22A**, 4–11.
36. S. Hiller, F. Fiorito, and K. Wüthrich, *Proc. Natl. Acad. Sci. U.S.A.*, 2005, **102**, 10876–10888.
37. D. Malmodin and M. Billeter, *J. Am. Chem. Soc.*, 2005, **127**, 13486–13487.
38. K. Kazimierczuk, W. Kozminski, and I. Zhukov, *J. Magn. Reson.*, 2006, **179**, 323–328.
39. M. Mobli, A. S. Stern, and J. C. Hoch, *J. Magn. Reson.*, 2006, **182**, 96–105.
40. P. J. Bowyer, A. G. Swanson, and G. A. Morris, *J. Magn. Reson.*, 1999, **140**, 513–515.
41. D. A. Snyder, F. Zhang, and R. Brüschweiler, *J. Biomol. NMR*, 2007, **39**, 165–175.
42. D. Malmodin and M. Billeter, *J. Magn. Reson.*, 2005, **176**, 47–53.
43. I. Scholz, S. Jehle, P. Schmieder, M. Hiller, F. Eisenmenger, H. Oschkinat, and B. J. Van Rossum, *J. Am. Chem. Soc.*, 2007, **129**, 6682–6683.
44. K. Kazimierczuk, A. Zawadzka, and W. Kozminski, *J. Magn. Reson.*, 2008, **192**, 123–130.
45. M. Mobli, M. W. Maciejewski, M. R. Gryk, and J. C. Hoch, *Nat. Methods*, 2007, **4**, 467–468.
46. M. Mobli, M. W. Maciejewski, M. R. Gryk, and J. C. Hoch, *J. Biomol. NMR*, 2007, **39**, 133–139.
47. J. C. J. Barna, E. D. Laue, M. R. Mayger, J. Skilling, and S. J. P. Worrall, *J. Magn. Reson.*, 1987, **73**, 69.
48. P. Schmieder, A. S. Stern, G. Wagner, and J. C. Hoch, *J. Biomol. NMR*, 1993, **3**, 569–576.
49. P. Schmieder, A. S. Stern, G. Wagner, and J. C. Hoch, *J. Biomol. NMR*, 1994, **4**, 483–490.
50. M. Mobli, A. S. Stern, and J. C. Hoch, *J. Magn. Reson.*, 2006, **182**, 96–105.
51. H. R. Eghbalnia, A. Bahrami, M. Tonelli, K. Hallenga, and J. L. Markley, *J. Am. Chem. Soc.*, 2005, **127**, 12528–12536.
52. J. C. Hoch, M. W. Maciejewski, and B. Filipovic, *J. Magn. Reson.*, 2008, **193**, 317–320.
53. K. Kazimierczuk, A. Zawadzka, W. Kozminski, and I. Zhukov, *J. Biomol. NMR*, 2006, **36**, 157–168.
54. M. Misiak and W. Kozminski, *Magn. Reson. Chem.*, 2007, **45**, 171–174.
55. N. Pannetier, K. Houben, L. Blanchard, and D. Marion, *J. Magn. Reson.*, 2007, **186**, 142–149.

56. G. L. Bretthorst, *Concepts Magn. Reson.*, 2008, **32A**, 417–435.
57. M. W. Maciejewski, H. Z. Qui, I. Rujan, M. Mobli, and J. C. Hoch, *J. Magn. Reson.*, 2009, **199**(1), 88–93.
58. F. Ni, G. C. Levy, and H. A. Scheraga, *J. Magn. Reson. (1969)*, 1986, **66**, 385–390.
59. S. Sibisi, *Nature*, 1983, **301**, 134–136.
60. P. J. Hore, *J. Magn. Reson.*, 1985, **62**, 561–567.
61. E. D. Laue, J. Skilling, J. Staunton, S. Sibisi, and R. G. Brereton, *J. Magn. Reson.*, 1985, **62**, 437–452.
62. M. A. Delsuc and G. C. Levy, *J. Magn. Reson. (1969)*, 1988, **76**, 306–315.
63. J. A. Jones, D. S. Grainger, P. J. Hore, and G. J. Daniell, *J. Magn. Reson., Ser. A*, 1993, **101**, 162–169.
64. M. J. Seddon, A. G. Ferrige, P. N. Sanderson, and J. C. Lindon, *J. Magn. Reson., Ser. A*, 1996, **119**, 191–196.
65. V. Stoven, J. P. Annereau, M. A. Delsuc, and J. Y. Lallemand, *J. Chem. Inf. Comput. Sci.*, 1997, **37**, 265–272.
66. N. Shimba, A. S. Stern, C. S. Craik, J. C. Hoch, and V. Dötsch, *J. Am. Chem. Soc.*, 2003, **125**, 2382–2383.
67. R. R. Ernst, *Adv. Magn. Reson.*, 1996, **2**, 1–135.