J|A|C|S
ARTICLES
Published on Web 02/06/2002

Modern Spectrum Analysis in Multidimensional NMR Spectroscopy: Comparison of Linear-Prediction Extrapolation and Maximum-Entropy Reconstruction

Alan S. Stern, Kuo-Bin Li,† and Jeffrey C. Hoch*

Contribution from the Rowland Institute for Science, 100 Edwin H. Land Boulevard, Cambridge, Massachusetts 02142

Received July 9, 2001. Revised Manuscript Received October 22, 2001

Abstract: NMR spectroscopy is an inherently insensitive technique, and many challenging applications such as biomolecular studies operate at the very limits of sensitivity and resolution. Advances in superconducting magnet, cryogenic probe, and pulse sequence technologies have resulted in dramatic improvements in both sensitivity and resolution in the past decade. Conversely, the signal-processing method used most widely in NMR spectroscopy, extrapolation of the time domain signal by linear prediction (LP) followed by discrete Fourier transformation (DFT), was developed in the early 1980s and has not been subjected to detailed scrutiny for its impact on sensitivity and resolution. Here we report the first systematic investigation of the accuracy and precision of spectra obtained by LP extrapolation followed by DFT. We compare the results to spectra obtained by maximum-entropy (MaxEnt) reconstruction, which was developed contemporaneously to LP extrapolation but is not widely employed in NMR spectroscopy. Although it reduces truncation artifacts and increases the amplitudes of strong peaks, we find that LP extrapolation generates false-positive peaks and introduces frequency errors. These defects of LP extrapolation become less pronounced for longer data records and higher signal-to-noise ratio. MaxEnt generally yields more detectable peaks for a given number of data samples, more accurate peak frequencies, and fewer false-positive peaks than LP extrapolation. MaxEnt also permits the use of nonlinear sampling, which can give dramatic improvements in resolution. These results show that the use of MaxEnt together with nonlinear sampling, rather than LP extrapolation, could reduce the amount of instrument time required for adequate sensitivity and resolution by a factor of 2 or more.

Introduction

The discrete Fourier transform (DFT) played an enabling role in the development of modern nuclear magnetic resonance (NMR) spectroscopy.[1] Despite the importance of the DFT for spectrum analysis in NMR, its limitations are well known,[2] and a variety of modern methods of spectrum analysis have been developed (for reviews, see refs 3−7) that provide better results than the DFT. These methods are particularly attractive for multidimensional experiments, where the data acquisition time is directly proportional to the number of data samples in the indirect dimensions. Since they achieve higher resolution from shorter data records, they can also yield higher sensitivity by using the time savings to perform additional signal averaging. As the cost of state-of-the-art NMR spectrometers continues to soar, and the demand for higher sensitivity in biomolecular NMR experiments from structural genomics initiatives[8] and high-throughput screening of drug candidates[9] increases, the incentive to exploit these methods to reduce data acquisition times and improve spectral quality becomes more significant.

The most widely used method of modern spectrum analysis used for processing NMR data is to extrapolate the measured signal using linear prediction (LP) prior to discrete Fourier transformation. Considerable effort has been devoted to characterizing the precision and accuracy of spectrum analysis based on *parametric* LP spectrum analysis,[10−12] in which LP is used

* Address correspondence to this author. Fax: (617) 497-4627. E-mail: hoch@rowland.org.

† Current address: DigiGenomics Co. Ltd., 8F, No. 196, Zhou Z St., Neihu, Taipei, 114, Taiwan.

(1) Ernst, R. R. In *Computational Aspects of the Study of Biological Macromolecules by Nuclear Magnetic Resonance Spectroscopy*; Hoch, J. C., Poulsen, F. M., Redfield, C. R., Eds.; NATO ASI Series: Life Sciences 225; Plenum Press: New York, 1991; pp 1−25.
(2) Hoch, J. C.; Stern, A. S. *NMR Data Processing*; Wiley-Liss: New York, 1996.
(3) Stephenson, D. S. *Prog. NMR Spectrosc.* **1988**, *20*, 515−626.
(4) Hoch, J. C. *Methods Enzymol.* **1989**, *176*, 216−241.
(5) Hoffman, R. E.; Levy, G. C. *Prog. NMR Spectrosc.* **1991**, *23*, 211−258.
(6) Decanniere, C.; van Hecke, P.; Vanstapel, F.; Chen, H.; van Huffel, S.; van der Voort, C.; van Tongeren, B.; van Ormondt, D. *J. Magn. Reson.* **1994**, *B105*, 31−37.
(7) Koehl, P. *Prog. NMR Spectrosc.* **1999**, *34*, 257−299.
(8) Christendat, D.; Yee, A.; Dharamsi, A.; Kluger, Y.; Savchenko, A.; Cort, J. R.; Booth, V.; Mackereth, C. D.; Saridakis, V.; Ekiel, I.; Kozlov, G.; Maxwell, K. L.; Wu, N.; McIntosh, L. P.; Gehring, K.; Kennedy, M. A.; Davidson, A. R.; Pai, E. F.; Gerstein, M.; Edwards, A. M.; Arrowsmith, C. H. *Nat. Struct. Biol.* **2000**, *7*, 903−909.
(9) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Science* **1996**, *274*, 1531−1534.
(10) de Beer, R.; van Ormondt, D. In *In-Vivo Magnetic Resonance Spectroscopy I: Probeheads and Radiofrequency Pulses, Spectrum Analysis*; Rudin, M., Ed.; NMR Basic Principles and Progress 26; Springer-Verlag: Berlin, 1992; pp 201−248.
(11) Gesmar, H.; Led, J. J.; Abildgaard, F. *Prog. NMR Spectrosc.* **1990**, *22*, 255−288.
(12) Koehl, P.; Ling, C.; Lefèvre, J.-F. *J. Magn. Reson.* **1994**, *A109*, 32−40.

10.1021/ja011669o CCC: $22.00

to construct a model of the signal, rather than to extrapolate the data. To our knowledge, however, no comparable analysis has been conducted for spectra obtained by LP extrapolation. This curious state of affairs persists, despite the fact that LP extrapolation is used far more routinely in the analysis of multidimensional NMR data than is parametric LP.

In this paper we present the first quantitative assessment of the accuracy of NMR spectra computed using LP extrapolation. We compare the results with those obtained using maximum-entropy (MaxEnt) reconstruction, another modern, though less widely used method of spectrum analysis. The methods are compared by testing the accuracy and precision of spectra computed from synthetic data with known characteristics, using data records of varying size. Repeated realizations of pseudo-random noise were added to the synthetic data; in all, we analyzed 5940 spectra. The criteria used to measure spectral quality include signal-to-noise ratio (S/N), line width, the number of peaks detected above a noise-determined threshold, the number of false peaks, the precision of peak amplitudes, the accuracy of peak frequencies, and the ability to resolve closely spaced peaks. In almost every instance, MaxEnt reconstruction provides more accurate spectra, higher sensitivity and resolution, and fewer false-positive peaks than LP extrapolation. Our results show that the use of MaxEnt reconstruction instead of LP extrapolation can substantially increase the performance of multidimensional NMR experiments.

LP extrapolation is based on the principle that future values of certain types of signals can be expressed as linear combinations of past data values.[3,6,7] The LP coefficients are determined from the measured free induction decay (FID) and then used to extrapolate the FID beyond the measured time interval. The spectrum is computed by apodization and discrete Fourier transformation of the augmented data (consisting of measured and extrapolated values).[13−17] LP is equivalent to modeling the data as a sum of exponentially decaying sinusoids and is exact, provided that the signal conforms to this model and the number of LP coefficients is at least as large as the number of sinusoids.[2] Real data include noise, however, and in practice this can lead to difficulties. For LP to detect signal components reliably, even when noise is present, requires that the number of coefficients be rather larger than the number of sinusoids; this inevitably results in the propagation of nonsignal components in the data. Also, for the extrapolation to be numerically stable, there must not be signal components with negative decay rates, i.e., exponentially growing sinusoids. The presence of noise can result in LP coefficients corresponding to growing sinusoids, and using them to extrapolate the signal will result in numerical overflow, artifacts, or both. Growing components correspond to roots of the characteristic polynomial (a polynomial with coefficients given by the LP coefficients) lying outside the unit circle in the complex plane.[18] These roots are often displaced to the unit circle (corresponding to a nondecaying signal) or reflected inside the unit circle (corresponding to a decaying signal), and the LP coefficients are recomputed from the modified roots; this practice has no formal justification beyond ensuring stable, albeit incorrect, extrapolation.

Potential hazards of using LP extrapolation to augment experimental data have been noted previously. As pointed out by Tang and Norris,[17] noise present in the experimentally determined part of the FID is retained, and the last L points of the FID (where L is the number of LP coefficients), which have the lowest S/N, are used to start the LP extrapolation. They argued that the gains realized by LP extrapolation of the experimental FID amount to little more than reduction of truncation artifacts. A similar prediction was made by Stephenson,[3] who also emphasized the distinction between observation noise and prediction noise. He noted that at low S/N, the assumptions inherent in applying the LP model are no longer valid. These predictions were based mainly on consideration of the algorithm, and they have some validity. In fact, however, some gain in S/N does result from LP extrapolation of the FID, because LP preferentially propagates components at the frequencies corresponding to the zeros of the characteristic polynomial. This can be seen by computing LP coefficients for an FID and then using them to extrapolate a signal containing only noise. The result is that peaks appear in the DFT spectrum of LP-extrapolated noise at positions corresponding to the strong components of the original signal used to compute the LP coefficients. As we will show, gains in S/N for signal components using LP extrapolation are often offset by selective amplification of noise, resulting in false peaks.

In comparison to LP, MaxEnt reconstruction makes few assumptions about the characteristics of the signal.[19] Because the peaks are not assumed to have any particular characteristics, MaxEnt can be used for spectral analysis in solid-state NMR and other applications that have non-Lorentzian peak shapes. Instead of computing a spectrum directly from the data, MaxEnt constructs trial spectra, computes the hypothetical time domain signal that would give rise to that spectrum, and then compares this with the actual data for consistency. The hypothetical data are said to be consistent with the experimental data when the level of disagreement (computed using a χ^2 statistic, for example) is approximately equal to the noise level. From among all the consistent trial spectra, MaxEnt selects the one with the highest entropy. Advantages of this inverted approach to spectrum analysis are that the experimental data can be sampled at arbitrary times,[20,21] small amounts of missing or corrupted data can be tolerated,[2] and arbitrary functions can be stably deconvolved.[22,23] Disadvantages include the fact that nominally Lorentzian peak shapes can be distorted and peak intensities altered (although it is possible to compensate for this effect[24]).

The test data used in this study are synthetic two-dimensional signals with characteristics typical of the indirect dimensions of multinuclear experiments on proteins, i.e., corresponding to a two-dimensional cross section of a three- (or higher) dimensional experiment. We specifically examined the influence of

(13) Ni, F.; Scheraga, H. A. *J. Magn. Reson.* **1986**, *70*, 506−511.
(14) Tirendi, C. F.; Martin, J. F. *J. Magn. Reson.* **1989**, *81*, 577−585.
(15) Zeng, Y.; Tang, J.; Bush, C. A.; Norris, J. A. *J. Magn. Reson.* **1989**, *83*, 473−483.
(16) Olejniczak, E. T.; Eaton, H. L. *J. Magn. Reson.* **1990**, *87*, 628−632.
(17) Tang, J.; Norris, J. R. *J. Magn. Reson.* **1988**, *78*, 23−30.
(18) Kumaresan, R. *IEEE Trans. Acoust., Speech, Signal Process.* **1983**, *31*, 217−220.
(19) Wernecke, S. J.; D'Addario, L. R. *IEEE Trans. Comput.* **1977**, *C26*, 351−364.
(20) Barna, J. C. J.; Laue, E. D.; Mayger, M. R.; Skilling, J.; Worrall, S. J. P. *J. Magn. Reson.* **1987**, *73*, 69−77.
(21) Schmieder, P.; Stern, A. S.; Wagner, G.; Hoch, J. *J. Biomol. NMR* **1993**, *3*, 569−576.
(22) Hoch, J. C.; Stern, A. S. *Methods Enzymol.* **2001**, *338*, 159−178.
(23) Serber, Z.; Richter, C.; Moskau, D.; Böhlen, J.-M.; Gerfin, T.; Marek, D.; Häberli, M.; Baselgia, L.; Laukien, F.; Stern, A. S.; Hoch, J. C.; Dötsch, V. *J. Am. Chem. Soc.* **2000**, *122*, 3554−3555.
(24) Schmieder, P.; Stern, A. S.; Wagner, G.; Hoch, J. *J. Magn. Reson.* **1997**, *125*, 332−339.

the number and distribution of samples on spectral quality. Spectra computed using conventional discrete Fourier analysis or LP extrapolation are necessarily restricted to linearly spaced time samples, so that only the total number of samples in each dimension can be adjusted. MaxEnt reconstruction is free from such constraints, however, and for this method we examined the influence of nonlinear sampling schemes,[20,21] as well as the total number of samples, on spectral quality. Far from being consistently beneficial or even benign, our results indicate circumstances where LP extrapolation actually degrades the quality of spectra, compared to conventional discrete Fourier analysis. In contrast, MaxEnt reconstruction preserves the frequency accuracy and resistance to false-positive peaks of the DFT but provides much higher resolution, especially when used with nonlinear sampling. The results hold implications for the design of experimental sampling schemes, for the choice of signal-processing methods, and for automated analysis of spectra.

The use of LP extrapolation in NMR warrants a brief historical note. In a recent review by Koehl,[7] the first application of LP extrapolation in NMR was attributed to Tirendi and Martin.[14] An earlier example, however, was published by Ni and Scheraga,[13] and a nearly contemporaneous example was published by Zeng et al.[15] Closely related methods were described by Tang and Norris,[17] in which LP is used to extrapolate a model FID fit to the experimental data, rather than the experimental data itself. Proper attribution is made difficult by the fact that many early applications used nomenclature very different from that used today. Ni and Scheraga, for example, referred to their technique as "maximum-entropy extrapolation", because of its similarity to the Burg maximum-entropy method.[25] (The Burg method uses LP to extrapolate the autocorrelation function, which on Fourier transformation yields the power spectrum, rather than a complex spectrum. The differences between the Burg method and MaxEnt reconstruction are described in ref 2.)

Methods

Test Data Sets. A group of 10 master data sets was prepared, each containing synthetic sinusoids plus noise in a two-dimensional array mimicking the appearance of a t_1-t_2 cross section from a three-dimensional NMR experiment following Fourier transformation along the acquisition dimension. Each master data set contained 78 complex points in t_1 and 50 complex points in t_2, for a total of 3900 hypercomplex points. The simulated spectral widths were 7600 Hz in f_1 and 3198 Hz in f_2. Forty-nine sinusoids were included in each data set (all in-phase): 25 with amplitudes ranging from 60 to 84 (by 1) and 24 with amplitudes ranging from 105.5 to 600 (by 21.5). On occasion, we will refer to the 25 lower-amplitude signals as the "weak peaks". The decay rates in t_1 and t_2 corresponded to full widths at half-maximum (fwhm) of 30 and 50 Hz, respectively. Frequencies were assigned using a pseudo-random sequence with close overlaps filtered out and avoiding the edges of the spectral window so as to prevent wrap-around. Pseudo-random Gaussian noise having a root-mean-square (RMS) amplitude of 300 was added to each data set. Different random seeds were used for the frequencies and the noise in the various data sets. A contour plot of the zero-filled DFT spectrum for one of the data sets is shown in Figure 1.

A second group of 10 master data sets was also prepared, identical to the first (i.e., having the same frequencies and the same noise) except that the decay rate of the sinusoids in the t_1 dimension was set to zero, to simulate data from a constant-time experiment. Finally, a third group of 10 data sets was prepared, identical to the first except that each peak was replaced by a doublet split by 80 Hz in f_1, to assess the ability of the algorithms to resolve closely spaced peaks. Separate calculations were carried out using the three master data collections.

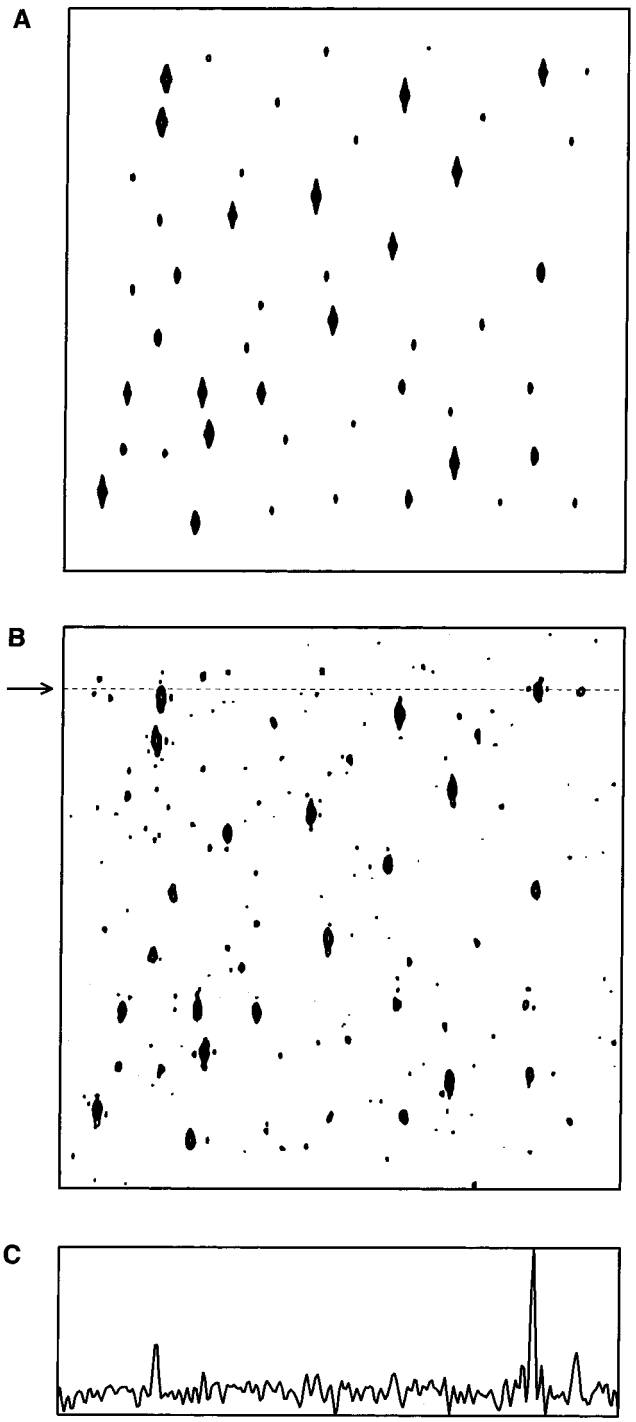

Figure 1. Contour plot for the DFT spectrum of a master data set without (A) and with (B) pseudo-random noise. The one-dimensional cross section in (C) is along the f_2 dimension at the position in f_1 indicated by the arrow and dashed line in (B).

Data Processing. Various subsets of each of the 30 master data sets were used as input for the spectrum reconstruction computations. The sizes of the final output spectra were all set to 256 × 256 points. Three types of spectra were computed: DFT alone (using zero-filling

(25) Burg, J. P. In *Modern Methods of Spectral Analysis*; Childers, D. G., Ed.; IEEE Press: New York, 1978; pp 34−48.

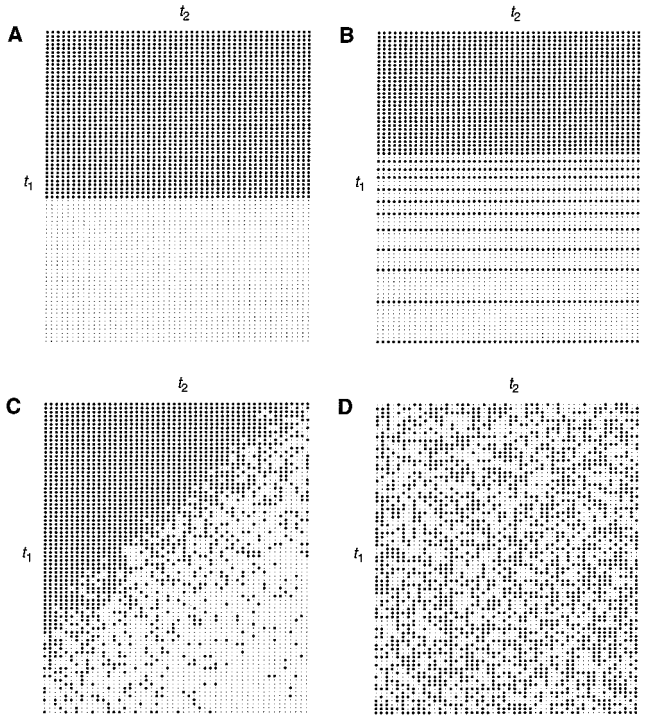

Figure 2. Typical sampling schedules for selecting subsets from a master data set, illustrated using large and small dots to represent samples used and not used, respectively. Linear sampling, used for all spectra employing the DFT or LP, and for linearly sampled MaxEnt reconstructions, is illustrated in (A). Here all points in t_2 from the master data set are used, and the number of points used in t_1 is varied. The nonlinear sampling schemes depicted employ exponential sampling in t_1 (B), exponential sampling in t_1 and t_2 (C), and random sampling in t_1 and t_2 (D).

and with or without apodization), extension in each of t_1 and t_2 by LP followed by DFT (with or without data mirroring and with or without apodization), and MaxEnt reconstruction in t_1 and t_2. Two types of apodization were used. For sensitivity enhancement, the window function consisted of an exponential decay corresponding to 30 Hz in t_1 and 15 Hz in t_2 for data sets with decaying signals; the exponential decay was set to 50 Hz in t_1 for data sets containing constant-time signals. For resolution enhancement, a 60°-shifted sinebell function was applied.

The test data subsets, or sampling schemes, are illustrated in Figure 2. For each scheme, reconstructions were performed using 500, 700, 900, ..., 3900 input points. Because LP and the DFT are not capable of handling nonlinearly sampled data, the linear sampling scheme used with these techniques consisted simply of all 50 points in t_2 together with various length initial segments of the full 78 points in t_1 (Figure 2A). Five sampling schemes were used with MaxEnt: linear sampling in t_1 (Figure 2A, the same as with LP and DFT), one-dimensional exponential sampling in t_1 (Figure 2B), one-dimensional random sampling from an exponential distribution in t_1 (not shown), random sampling from a two-dimensional exponential distribution in t_1 and t_2 (Figure 2C), and uniform random sampling in two dimensions (Figure 2D). Sample points for the 1D exponential scheme were selected as described previously.[21] The 1D random exponential sampling scheme resembles the 1D exponential scheme in that the samples in each column are distributed exponentially in t_1, but the random selection means that different columns have different points sampled. Sample density decay rates were set to 30 Hz in t_1 and 50 Hz in t_2.

Given input data d_0, ..., d_{M-1} (where M is the number of available points), the LP method involves finding coefficients a_1, ..., a_L (L is the prediction filter order) that best satisfy the prediction equations:

$$d_n = a_1 d_{n-1} + \cdots + a_L d_{N-L}, \quad n = L, ..., M-1 \tag{1}$$

Once determined, the prediction filter is used to extend the input data by iteratively applying eq 1 with n set equal to M, $M+1$, and so on. Our LP calculations used the singular value decomposition (SVD) method for determining the prediction coefficients.[2] The prediction filter order L was set to the maximum possible value in each dimension: one-half the number of available data points. Data mirroring (i.e., extension of the input data to negative time values) was performed according to Zhu and Bax.[26] Unstable growth of the output data was detected by checking whether the output points attained an absolute value more than twice as large as that of the largest input point; when this occurred, the prediction coefficients were altered by solving the characteristic polynomial

$$P(x) = x^L - a_1 x^{L-1} - \cdots - a_{L-1}x - a_L \tag{2}$$

and replacing roots larger than 1 in absolute value by their reflection about the unit circle in the complex plane. (In some of the spectra, for which only a small number of data points were used, this technique failed to detect some slowly growing sinusoids.)

Maximum-entropy reconstructions used a fully two-dimensional version[24] of the modified "Cambridge" algorithm.[27] The algorithm seeks to find the spectrum $f = f_0, ..., f_{N-1}$ (where N is the output size) that maximizes the "spin-1/2 entropy",[28,29]

$$S(\mathbf{f}) = \sum_{j=0}^{N-1} R(|f_j|) \tag{3}$$

$$R(x) = -\frac{x}{\mathrm{def}} \log\left(\frac{x/\mathrm{def} + \sqrt{4 + x^2/\mathrm{def}^2}}{2}\right) + \sqrt{4 + x^2/\mathrm{def}^2} \tag{4}$$

subject to the constraint that

$$C(\mathbf{f},\mathbf{d}) = \frac{1}{4M} \sum_{k=0}^{M-1} |\mathrm{IDFT}(\mathbf{f})_k - d_k|^2 \le \mathrm{aim}^2 \tag{5}$$

Here, IDFT is the inverse discrete Fourier transformation, def is a scale factor (set to 20), and aim is an estimate of the noise level in the data (set to 200). No decay kernel was deconvolved. Note that since the algorithm is two-dimensional, the j and k sums in eqs 3 and 5 range over both indices of the data sets (f_1, f_2 and t_1, t_2, respectively). The absolute values appearing in these equations are the square roots of the sum of the squares of all four hypercomplex data components, which accounts for the factor of 4 appearing in the denominator of eq 5. The entropy $S(\mathbf{f})$ of a spectrum is sometimes described as a measure of the amount of "missing information"; by this interpretation, the spectrum obtained by MaxEnt reconstruction is the one that has the least amount of information still consistent with the experimental data.[30] We take the somewhat more pragmatic view that the entropy is a useful regularizer that ensures smooth, stable reconstructions from noisy data.

Data Analysis. Peak maxima were identified by starting at the spectral position corresponding to the frequencies of a synthesized peak and conducting an uphill search, at each step comparing the current value with the eight nearest neighboring values. If the search failed to terminate after two steps, the peak was deemed undetectable. Peak amplitudes, line widths, and frequencies were determined by least-squares fitting a Lorentzian line to an f_1 cross section through the center of each peak. For peaks that were too small relative to the noise level, the fitting procedure did not converge; such peaks were not included in the statistical analyses. Reported signal heights are the average of the peak heights over all the fitted peaks in each spectrum. The noise

(26) Zhu, G.; Bax, A. *J. Magn. Reson.* **1990**, *90*, 405−410.
(27) Skilling, J.; Bryan, R. K. *Mon. Not. R. Astron. Soc.* **1984**, *211*, 111−124.
(28) Jones, J. A.; Hore, P. J. *J. Magn. Reson.* **1991**, *92*, 363−376.
(29) Hoch, J. C.; Stern, A. S.; Donoho, D. L.; Johnstone, I. M. *J. Magn. Reson.* **1990**, *86*, 236−246.
(30) Sibisi, S. *Nature* **1983**, *301*, 134−136.

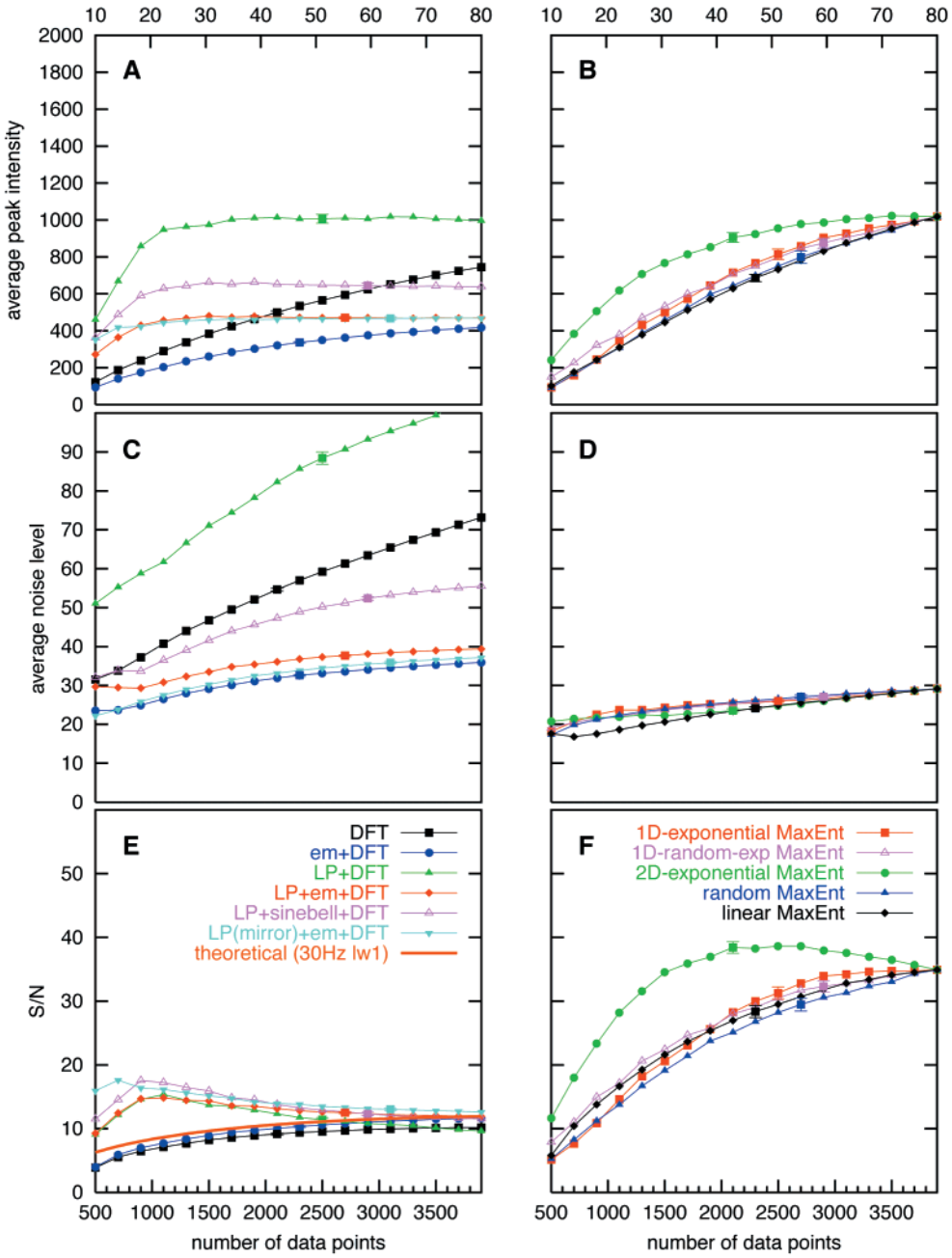

Figure 3. Signal, noise, and S/N amplitudes for spectra computed from the non-constant-time data sets. The lower horizontal axis label gives the total number of data points used in computing the spectrum; the upper label gives the corresponding number of points in t_1 for the linear sampling schemes (the same convention is used in Figures 4 and 6). Except when nonlinear sampling was employed (MaxEnt), all 50 points in t_2 from the master data sets were used. The thick curve in panel E shows the expected theoretical S/N (Appendix II) for a DFT spectrum.

level was computed as the RMS average value of the spectrum, excluding a square window 13 points across centered at each peak. False peaks were identified as all points not within any of the square peak windows, larger than a cutoff value, and larger than any other points within a 13 × 13 region; the cutoff value was set to 5 times the noise level. A "doublet splitting parameter" Δ, measuring the depth of the "valley" between components of a doublet, was calculated according to the formula

$$\Delta = 1 - \frac{C}{(A+B)/2} \qquad (6)$$

where A and B are the heights of the peaks comprising the doublet and C is the height of the intervening valley. All of the analysis was carried out separately for each spectrum, and the results were averaged over the 10 master data sets in each group to produce the reported values. Combined CPU timings for the spectral estimates and analyses are listed in Appendix I.

Results and Discussion

S/N and Sensitivity. S/N values for the zero-filled DFT, LP-extrapolated DFT, and MaxEnt spectra as functions of the number of data samples are shown in Figure 3 (panels E and F), along with separate plots for the RMS noise level and average signal amplitude (panels A−D). The theoretical result given by eq AII.3 (Appendix II) for the DFT spectrum of a decaying sinusoid plus random noise is also plotted in panel E of Figure 3. The S/N curves for the linear spectral estimates (DFT with and without apodization) closely resemble the

theoretical curve, monotonically increasing with increasing number of samples. As the sinusoids decay, eventually a point is reached beyond which additional data samples contribute more noise than signal, and the S/N begins to decrease. The predicted point of diminishing returns for our data is at 101 points in t_1, which is off scale in this plot. In contrast, S/N obtained using LP extrapolation actually decreased with increasing number of points. Panels A and C of Figure 3 show that this decrease resulted from an increase in the average noise level that was more rapid than the increase in the average signal level. The reason for this odd behavior is very simple. Once there are sufficiently many data points for LP to fit the signal peaks, adding more data will help make the fits more accurate but will not significantly affect the fitted intensities. The additional noise will increase the overall noise level, however, thereby decreasing the S/N value.

S/N in the MaxEnt spectra, on the other hand, exhibited a rather marked increase with increasing M (Figure 3F) and reached values that were substantially greater than those obtained using the DFT, with or without LP extrapolation. The average noise level in MaxEnt spectra was fairly insensitive to M (Figure 3D), and thus the increasing S/N resulted from increasing signal intensity (Figure 3B). It is interesting that with exponential sampling in both dimensions, the S/N was substantially higher and reached a maximum when M was around 2200 total points. This apparently results because more samples are concentrated at shorter aggregate times $t_1 + t_2$ for a given sample size than when linear sampling is used in t_2. Note that since the various nonlinear sampling schemes differ only in which sample points out of the master data sets they utilize, there is no difference among the MaxEnt reconstructions based on all 3900 input points. Consequently, in panels B, D, and F of Figure 3 (and also in later figures), the curves for the different sampling schemes converge at the right-hand edge.

S/N alone is not a reliable indicator of the sensitivity of nonlinear spectral estimates, such as LP extrapolation or MaxEnt reconstruction.[31] Sensitivity refers to the ability to distinguish between signal and noise; it implies not just detection of weak peaks, but also rejection of false or noise peaks. Panels A and B of Figure 4 display the number of false peaks using threshold peak detection for DFT-based spectra (with and without LP extrapolation) and for MaxEnt spectra, respectively. Exponential apodization and mirror-image extrapolation both resulted in dramatic reduction of the number of false peaks using LP, but for short samples LP extrapolation of any description resulted in more false peaks than using DFT alone. MaxEnt reconstruction, in contrast, was resistant to false peaks for all sample sizes when using linear, 2D exponential, or random sampling. For short samples, 1D exponential sampling yielded false peaks in numbers comparable to those obtained with LP extrapolation with exponential apodization. We believe that these false peaks to a great extent arise from the coherence (in f_2) of artifacts resulting from nonlinear sampling. When nonlinear sampling is applied in both dimensions, or when a degree of randomness is superimposed upon the nonlinear sampling in t_1 (the 1D random exponential sampling), this coherence disappears, and the number of false peaks diminishes.

(31) Donoho, D. L.; Johnstone, I. M.; Stern, A. S.; Hoch, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 5066−5068.

The plots in Figure 4C,D, showing the fraction of the weak peaks detected using the different spectral estimates, loosely resemble the curves for S/N, but there are differences. LP extrapolation resulted in a dramatic increase in the fraction of weak peaks detected, as well as a shift in the optimal sample size for detection of weak peaks to shorter samples. In contrast to their rejection of false peaks, apodization and mirror-image extrapolation had almost no impact on the detection of weak peaks. Detection of weak peaks using MaxEnt reconstruction increased nearly monotonically with increasing sample size, and 2D exponential sampling performed better than other sampling schemes. Overall, the highest scores for weak-peak detection using MaxEnt are comparable to those obtained using LP, although the best LP results occurred in a regime where there were significantly more false peaks than there were using MaxEnt.

Data from experiments that incorporate a constant-time evolution period do not decay in the corresponding time dimensions. In principle, the sensitivity for constant-time domains should not show a point of diminishing returns with additional samples. Using LP extrapolation, it should instead reach a plateau corresponding to convergence of the LP coefficients to the true values corresponding to the signal components. The plots for our reconstructions using the constant-time data sets do have a plateau in the average peak intensity starting around 45 sample points (not shown). The noise level does not plateau, but the slope diminishes near 50 samples. The net result is that S/N reaches a plateau fairly early, around 20 samples. The peak intensity and noise level for zero-filled DFT spectra do not plateau, since each additional sample contributes the same amount of signal and noise.

The number of false peaks detected when one of the time domains was constant time did not significantly differ from that obtained for non-constant-time data, whether using LP or MaxEnt (not shown). LP extrapolation invariably generated more false peaks than the DFT alone or MaxEnt reconstruction, except when nonlinear sampling was used only in one time domain. The trend observed for the fraction of weak peaks detected is qualitatively different, however. Using LP extrapolation, instead of reaching a point of diminishing returns, the number of peaks detected increased monotonically (Figure 4E). The qualitative trend for MaxEnt reconstruction is unchanged, but the fraction of weak peaks was larger for all sample sizes (Figure 4F). For constant-time data, the number of weak peaks detected was roughly comparable for LP extrapolation and MaxEnt reconstruction.

The distinctly different trends observed in threshold peak detection using LP extrapolation and MaxEnt reconstruction suggest that optimal experimental design for discriminating between signal and noise will depend on the technique used for spectrum analysis. With LP extrapolation, the optimal sample size for detecting weak peaks was around 30 samples in t_1, but LP extrapolation was also prone to generating false peaks for such short samples. The optimal sample size for discriminating signal peaks from noise is thus difficult to determine a priori and involves finding a tradeoff among peak detection, false peak rejection, and data acquisition time. Deciding the optimal sample size when using MaxEnt reconstruction is more straightforward. The optimal sensitivity to weak peaks occurred with large samples (near 70 samples in t_1), where MaxEnt was also highly

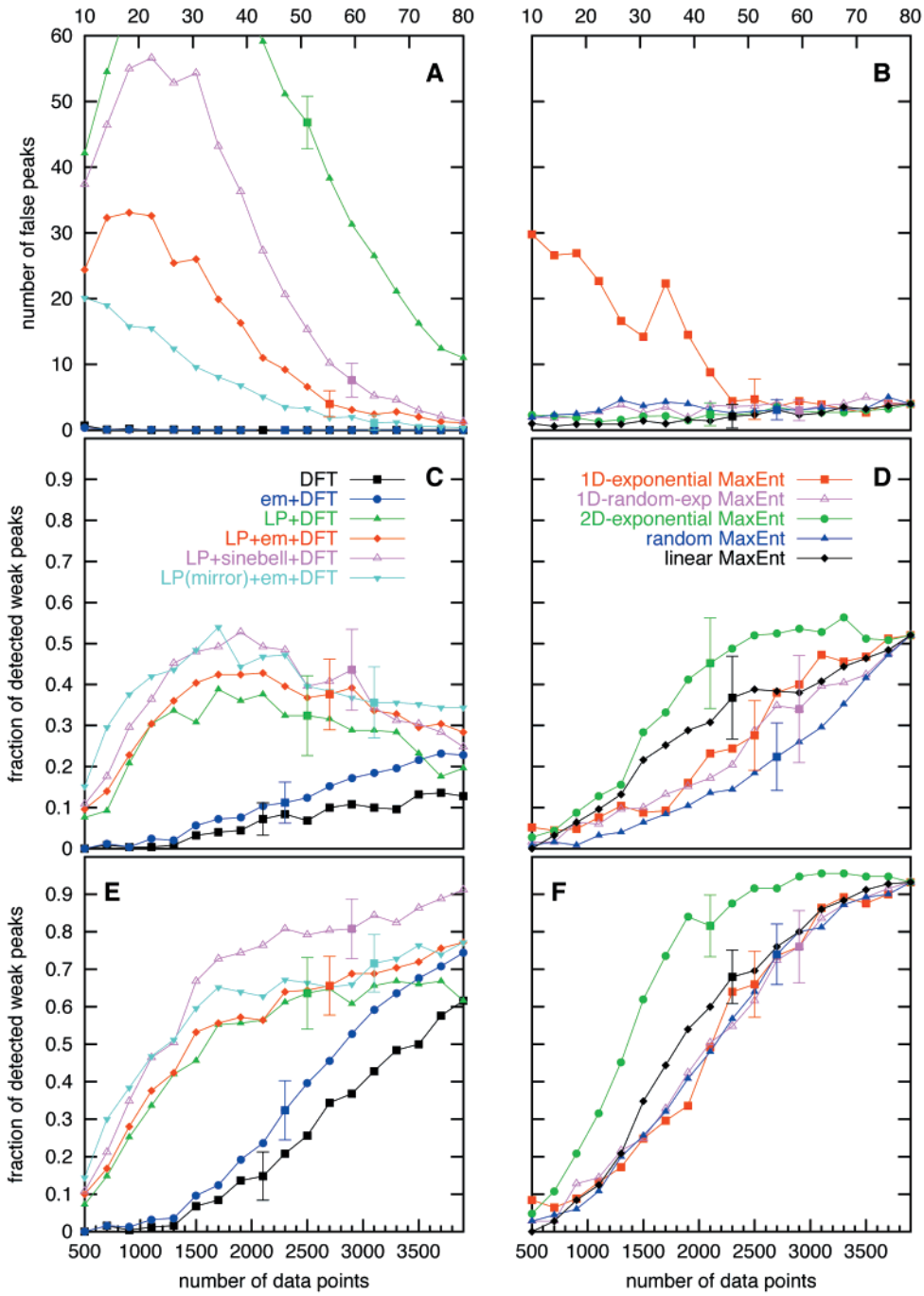

Figure 4. Results of threshold peak detection, indicating the number of false peaks detected (panels A and B) and the fraction of weak peaks detected for spectra from the non-constant-time data sets (panels C and D) and from the constant-time data sets (panels E and F).

resistant to false peaks. The tradeoff using MaxEnt is thus mainly between sensitivity and acquisition time.

The empirical sensitivity trends we observe are consistent with the algorithmic properties of the spectral estimates. LP extrapolation implicitly models the signal (containing both sinusoidal nuclear resonances and noise) as a sum of exponentially damped sinusoids. To help prevent the noise from biasing the spectral parameters (frequency, decay rate, and so on) of the sinusoids, the prediction order L is chosen to be larger than the expected number of resonances. Since L is equal to the number of sinusoids in the model, the extra sinusoids end up being fitted to the strongest noise components. The resulting LP coefficients propagate signal components with frequencies corresponding to the model, whether they are resonances or noise. Thus, the tendency of LP extrapolation to generate false peaks is attributable to the underlying model. Similarly, the resistance of MaxEnt reconstruction to false positives, even in regimes where it is maximally sensitive to signal components, can be understood from its mathematical foundation. In the limit where N (the number of frequencies in the spectral estimate) is equal to M (the number of time samples), it has been shown that MaxEnt reconstruction is numerically equivalent to a nonlinear transformation applied to the DFT spectrum.[31] This nonlinear transformation scales down the amplitude of the spectrum at all frequencies, with small-amplitude peaks scaled down more than large-amplitude peaks. Such a transformation

can only reduce the number of peaks above an RMS-noise-determined threshold. For the general case in which the number of points in the reconstructed spectrum is larger than the number of data samples, this relationship does not strictly hold, but the limiting behavior helps us to understand the resistance of MaxEnt reconstruction to false peaks.

The situation changes when MaxEnt is used to compute spectra for nonlinearly sampled data. The task here can approximately be described as deconvolving the spectrum of a sampling function (equal to 1 for every sampled time point and 0 for points not sampled) from the spectrum that would be obtained if the data were linearly sampled.[32] The presence of noise prevents perfect deconvolution, and the result is weak satellites around peaks at positions determined by the Fourier transform of the sampling function. When the dynamic range of the data is high, that is, when there are signal components spanning a wide range of amplitudes, these artifactual satellites can be confused with weak peaks.

When nonlinear sampling is used in only one dimension, and the same sampling schedule is used for all data columns, the sampling artifacts can align around peaks, creating extended artifacts. One way to minimize this effect is to use different sampling schedules for different columns of the data matrix, by introducing a degree of randomness that preserves the overall character of the schedule. This is one example of the more general phenomenon, that nonlinear sampling applied to one domain can have unexpected manifestations in other dimensions. Another example of this effect appears in Figure 3F (see above).

Amplitude Precision. False peaks and missing weak peaks result from the inability to distinguish signals from noise. For signal peaks that *are* detected, noise influences the accuracy of the amplitudes in the spectral estimate. In general, errors in the signal amplitudes can appear both as systematic bias and as random fluctuations that place a limit on the precision of any amplitude estimate. We previously showed that *in situ* calibration of MaxEnt spectra can remove the systematic bias, yielding intensities with accuracy comparable to that of DFT spectra in the context of relaxation and quantitative "J" experiments.[24] Here we examine the relative precision of amplitudes derived from different spectral estimates.

The relative precision (computed as the ratio of the RMS deviation of the amplitude divided by the average amplitude over the 10 master data sets) for spectra obtained using 54 sample points in t_1 is shown in Figure 5, plotted as a function of the input peak amplitude. For peaks detected in fewer than two spectra, the relative precision was assigned the value 0. For clarity, the weak peaks are represented by only two data points, computed as averages of the 13 smallest and 12 largest weak peaks, respectively.With the exception of unapodized LP extrapolation, the results using LP extrapolation (panel A of Figure 5) or MaxEnt reconstruction (panel B) were comparable. Both yielded more precise amplitudes for larger signal amplitudes, and both were less precise than Fourier transformation alone.

Line Widths, Resolution, and Frequency Accuracy. Although narrower line widths do not necessarily mean improved resolution, we will begin our characterization of LP's and MaxEnt's resolution by examining the line widths in the reconstructed spectra. Panels A and B in Figure 6 show the average measured widths, in the f_1 dimension, of all the peaks in the test spectra. (Although the digital resolution of the spectra is 29.7 Hz, the precision of the measured line widths is much better, because the analysis program used a nonlinear least-squares fitting procedure to estimate the peak parameters.) The decay rate of the signals in the test data corresponded to a line width of 30 Hz (fwhm); this value is accurately reflected by the reconstructions that used LP without any apodization, provided sufficient input points were used. As one would expect, the DFT-based reconstructions that did use exponential apodization yielded larger line widths.

The parametric model underlying LP shows up plainly in Figure 6A: once sufficient input data were present for the linear-prediction coefficients to match the actual signal components, the spectral estimates converged to fixed values for the line widths, and adding more data made essentially no difference. In contrast, the non-LP DFT spectral estimates resulted in peaks that became narrower as more input points were added. Reading down the right-hand side of the panel, one can distinguish two classes of spectral estimates: those that used exponential apodization and those that did not. Within each class, the LP method produced narrower lines than the non-LP method did.

With the exception of the linear schedule, the MaxEnt line widths shown in panel B are comparable to or better than the LP line widths in each case, and they are also rather insensitive to the number of input points. This is understandable, since the nonlinear sampling schemes all involve acquiring data for long t_1 delay times, even when the total number of points acquired is low. The information provided by these points constrained the reconstructions to match the actual decay rate much more closely than the reconstructions using linearly sampled data. In fact, the linear schedule produced results very similar to those of the plain DFT. (Note that the MaxEnt reconstruction did not deconvolve a decay from the signal.)

Just as important as the sharpness of the peaks, but often overlooked, is the accuracy of the peak frequencies. While frequency error is not usually considered an aspect of resolution, it is crucial for identifying connected peaks in multidimensional spectra. Frequency errors increase the ambiguity of correlations (peaks with a frequency in common) and thus increase the likelihood that correlations will be made in error or missed. The f_1 frequency errors of the reconstructed peaks are displayed in Figure 6C,D. Notice that all the LP reconstructions except for LP-mirror had larger errors than the DFT, although with more input data the errors became smaller. Interestingly, the unapodized and the sinebell-apodized LP reconstructions—which yielded the narrowest peaks—had the largest errors. Again with the exception of the linear sampling schedule, MaxEnt produced frequency errors that were considerably smaller than LP did for small data sets. For large data sets there was no significant difference in the extent of the errors. Just as in the line width analysis, MaxEnt with linear sampling gave results essentially the same as the DFT. The frequency error, shown in Figure 5C,D as a function of signal amplitude for spectra computed from 54 data samples in t_1, is smaller for more intense signals in each of the spectral estimates. Although the variation is large, the trend toward lower error is apparent and mirrors that observed for the amplitude precision (panels A and B of Figure 5). This trend is one tangible aspect of the connection between

(32) Schmieder, P.; Stern, A. S.; Wagner, G.; Hoch, J. *J. Biomol. NMR* **1994**, *4*, 483−490.

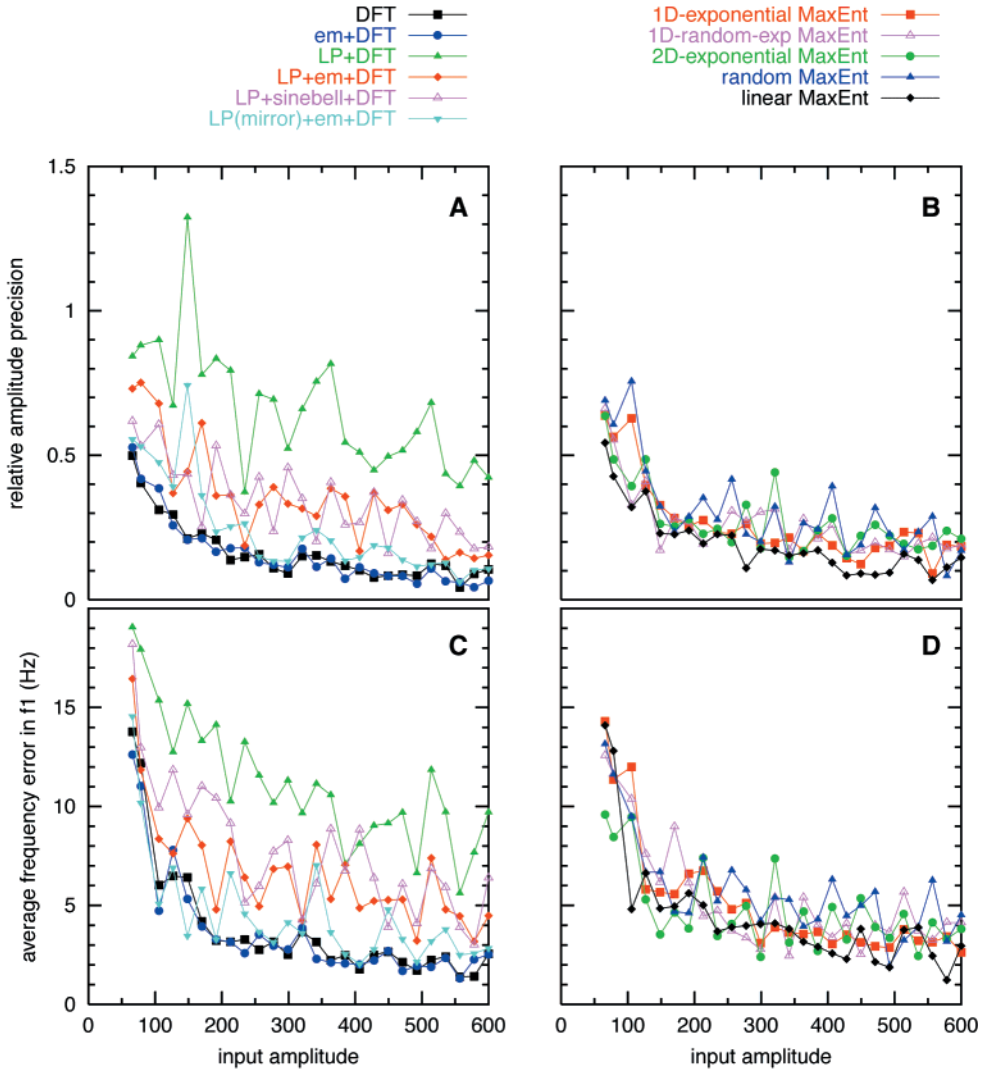

Figure 5. Relative precision (σ/average) of the amplitude estimates (panels A and B) and average frequency error in f_1 (panels C an D) as a function of the input signal amplitude, for spectral estimates computed using 54 data samples in t_1, or 2700 points total for the nonlinear sampling schemes.

sensitivity and resolution, and it shows that for high S/N, the deficiencies of LP extrapolation become less important.

A third way to characterize the resolution of a spectrum is to see how well it permits one to distinguish two closely spaced peaks. We defined a "splitting parameter" Δ (explained in the Methods section) in order to describe this property numerically. Plotted in Figure 6E,F are the average Δ values for the peaks in the spectra of our third test group, in which each peak was replaced with a doublet split in f_1. (The odd bump on the left side of the curve for unapodized LP is a result of the analysis program being confused by uncorrected exponentially growing sinusoids, as mentioned earlier.) The LP methods are clearly better at resolving doublets than are the DFT methods, but their resolving power drops quickly as the amount of input data is reduced; LP-mirror was the best behaved in this regard. The mirroring of the input data is undoubtedly the source of its advantage; it can make use of a data set that is effectively twice as long to improve the resolution. However, it lagged behind sinebell and unapodized LP when large amounts of data were available, because it employed exponential apodization. Had we used LP + mirror + sinebell, the results would probably have been close to the unapodized LP curve. The nonlinearly sampled MaxEnt reconstructions performed much better. In particular, 2D exponential sampling, 1D exponential sampling, and random sampling show only modest declines in their ability to separate closely spaced peaks even for the shortest data records. Linearly sampled MaxEnt gave results roughly comparable to those of DFT, performing slightly worse than LP for short data sets. Thus, the gains in resolution afforded by MaxEnt reconstruction arise mainly from the ability to use nonlinear sampling, and not MaxEnt in itself.

The influence of sample size on resolution for the constant-time data was virtually identical to that observed for the non-constant-time data (not shown).

Application to Experimental Data. Synthetic data sets provide an unambiguous standard for assessing the accuracy of spectral estimates, but they have limitations. Because synthetic data contain signal components that are rigorously sinusoidal and noise that is randomly distributed, the performance measured using synthetic data may overestimate the accuracy in the presence of nonrandom noise or signal components that are not perfect sinusoids. Since these assumptions are intrinsic to the method, it might be expected that spectra obtained using LP extrapolation, in particular, will not approach the accuracy reported for synthetic spectra. Although we cannot know a priori the exact characteristics of signal components in

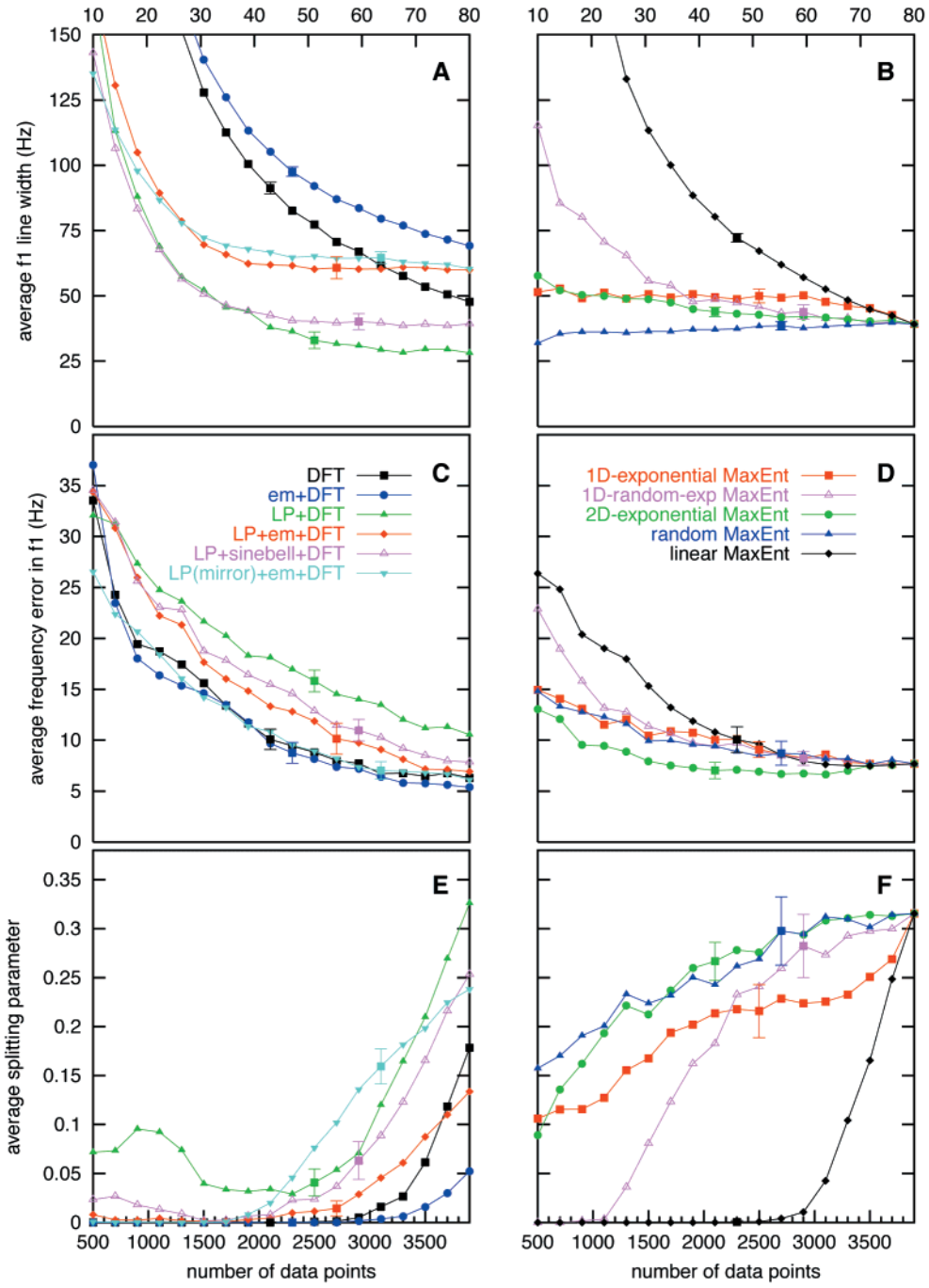

Figure 6. Measures of resolution as a function of sample size. Line widths in f_1 (panels A and B), average frequency error (panels C and D) for the non-constant-time data sets, and values of the "splitting parameter" Δ (as defined in the text, panels E and F) for the doublet non-constant-time data sets.

real data, we can use the fact that longer data records tend to provide more accurate frequencies to compare the relative performance of spectral estimates. Results for LP extrapolation and MaxEnt reconstruction of 40-, 64-, and 128-point data lengths in the ^{15}N (t_1) dimension of heteronuclear single quantum coherence (HSQC) data for the 23 kDa protein prolactin are compared with values derived from the DFT spectrum of 512-point (in t_1) data in Table 1 and Figure 7. The LP extrapolations were computed using the largest possible number of LP coefficients: 20, 32, and 64 for the 40-, 64-, and 128-point data sets, respectively. Using fewer LP coefficients tends to increase the frequency accuracy and decrease the number of the false-positive peaks, at the expense of detecting fewer real peaks (data not shown). The peaks listed in Table 1 lie in the spectral window corresponding to amide N−H resonances; a portion of this window is shown in the contour plots in Figure 7. Comparison of the contour plots of spectra computed using LP extrapolation (panel B) and MaxEnt reconstruction (panel C) dramatically reinforces the superiority of MaxEnt reconstruction that was apparent from the analysis of synthetic data. With the exception of the spectra computed using 128 points, for which LP extrapolation and MaxEnt reconstruction yielded comparable frequency errors in f_1, MaxEnt reconstruction performed far better than LP extrapolation in terms of total number of peaks detected above the noise threshold, the number of false peaks, and frequency error.

Table 1. Peak Counts and Frequency Fidelity for LP Extrapolation and MaxEnt Reconstruction of Truncated HSQC Data for Human Prolactin.

	Points						
	512	128		64		40	
	DFT	LP	MaxEnt	LP	MaxEnt	LP	MaxEnt
assigned peaks detected	182	158	174	111	170	80	169
peaks above a noise-determined threshold	182	158	174	108	170	59	167
false or unassigned peaks	16	485	23	172	18	101	17
avg frequency error (Hz)	0.00	1.35	1.40	2.30	1.49	2.38	1.65

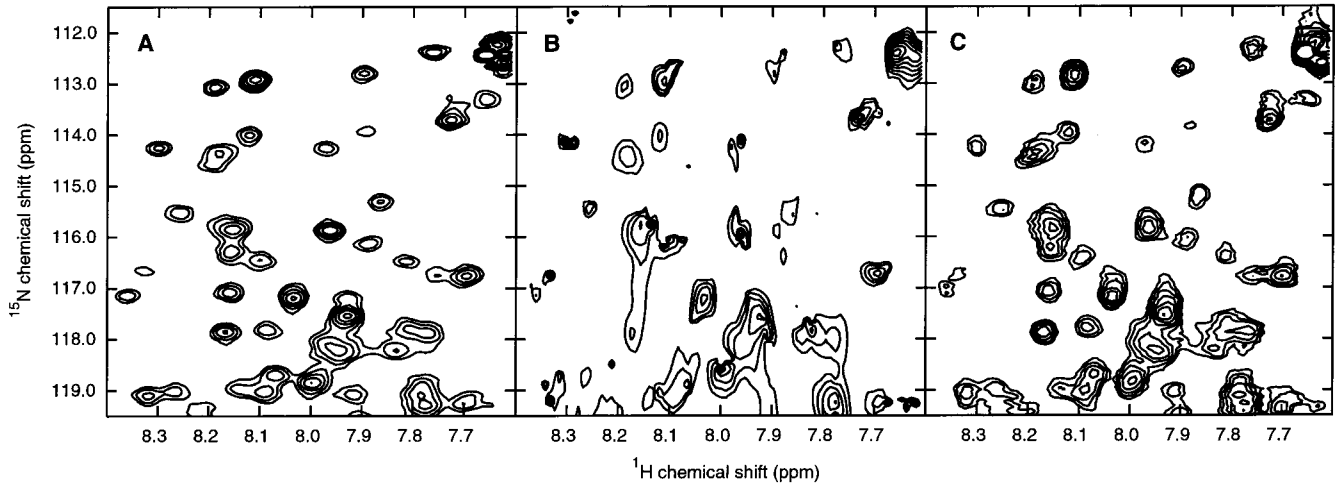

Figure 7. Contour plots for a part of the N−H region of the ^{15}N−^{1}H HSQC spectrum of human prolactin, utilizing (A) 512 data samples in t_1 and discrete Fourier transformation, (B) 64 points in t_1 with LP extrapolation to 512, 60°-shifted sinebell apodization, and discrete Fourier transformation, and (C) MaxEnt reconstruction using 64 samples in t_1, with exponential nonlinear sampling.

Concluding Remarks

The most important conclusions from this study are that LP extrapolation of FIDs can degrade the quality of the spectra and MaxEnt reconstruction performs better. The diminished accuracy of peak positions and false peaks that can result are particularly insidious defects, because they are not readily apparent on casual inspection, yet they can seriously compromise automated analysis. The dramatic reduction in truncation artifacts with LP extrapolation and the apparent increase in S/N (for intense peaks) that led to its widespread adoption in NMR do not correspond to improved accuracy of peak frequencies or improved sensitivity. In most respects, MaxEnt reconstruction is superior to LP extrapolation for improving the resolution of spectra from short data records. In particular, the ability of MaxEnt to accommodate nonlinearly sampled data can provide significant improvements in both sensitivity and resolution for short data sets, compared to linear sampling.

The reduction in the number of false peaks, and the improved amplitude precision and resolution that mirror-image extrapolation provides, indicate that many of the difficulties associated with LP extrapolation are due to numerical instability, rather than deficits in the LP model. Mirror-image extrapolation imposes a phase assumption, however, in addition to those inherent in the LP model. We have observed examples where the presence of even a single signal component that does not have the proper phase introduces significant frequency errors for other components of the spectrum, not just the component with the wrong phase. While these sorts of errors are beyond the scope of the present investigation, it is worth emphasizing that improvements of mirror-image over simple LP extrapolation for synthetic data do not necessarily imply robustness in the face of nonideal signals and noise that are characteristic of experimental data.

Our results also help to explain why the deficiencies of LP extrapolation are not reflected in the literature. LP extrapolation performs well when the S/N is high and data records are sufficiently long. Spectroscopists tend to collect data until the quality of the spectra are sufficient for their needs, regardless of the method used for spectrum analysis. Using LP extrapolation, however, requires that more time must be devoted to data collection in order to achieve sufficient spectral quality than would be necessary using MaxEnt reconstruction. For this reason we believe that the most significant benefit that will accrue from the use of MaxEnt reconstruction is reduction in the amount of time devoted to data acquisition, rather than an increase in resolution or sensitivity per se.

A lesson reinforced by this investigation is that the nonlinearity of modern methods of spectrum analysis demands special attention when addressing the question of spectral quality. Most importantly, the S/N ratio does not serve as a reliable indicator of sensitivity. While nonlinearity is the source of much of the power of modern spectral estimates, it places a burden on the spectroscopist to empirically evaluate the performance of a chosen method for the particular data being analyzed. *In situ* analysis, based on adding synthetic signals with known characteristics to the data prior to computing the spectrum, is a straightforward approach that is applicable to any method of spectrum estimation.[22,24]

The potential offered by MaxEnt reconstruction for higher resolution, better sensitivity, and more efficient use of valuable

Table 2. CPU Time (Minutes:Seconds) for Computing the Spectral Estimates and Peak Analyses

	500 (10)[a]	1500 (30)	2500 (50)	3500 (70)	3900 (78)
DFT	0:02	0:02	0:02	0:02	0:02
LP	0:24	0:31	0:56	2:32	4:45
LP mirror	0:33	0:44	1:19	3:21	5:49
MaxEnt/random	4:27	6:32	9:05	11:14	12:02
MaxEnt/2D exponential	5:34	9:52	11:29	11:57	12:02

[a] The first number indicates the total number of input points; the number in parentheses is the number of points in t_1 (for DFT and LP calculations only).

spectrometer time, not only for high-throughput applications of NMR such as structural proteomics[8] and drug candidate screening,[9] but for any endeavor employing multidimensional NMR, is substantial. When resolution in indirect dimensions is a determining factor, the use of MaxEnt reconstruction and nonlinear sampling can routinely reduce data collection time by a factor of 2, and often more. The failure to utilize state-of-the-art methods of spectrum analysis to process NMR data is not only an extravagant waste, but also a needless impediment to research.

Acknowledgment. This work was supported by the Rowland Institute for Science and by grants from the National Institutes of Health (GM 47467, G. Wagner, P.I.) and the National Science Foundation (MCB 9527181, G. Wagner, P.I.). We are grateful to Gerhard Wagner (Harvard Medical School) for his continuing interest and support. We thank Sandrina Kinet and Joseph Martial (Université de Liège) for the ^{15}N-labeled sample of prolactin.

Appendix I. CPU Times for the Spectral Estimates

Representative CPU times for computing a single two-dimensional spectrum and performing the peak analysis are given in Table 2. The times were measured using an SGI O2 computer, with a 174 MHz R10000 (IP32) central processing unit and 256 megabytes of random access memory. Note that the MaxEnt reconstructions exhibited a wide range of execution times (reflecting differing rates of convergence); in each class, the random sampling schedule used the least amount of CPU time and the 2D exponential used the most.

Appendix II. Analytical Derivation of S/N for a Decaying Sinusoid

For a signal containing a single exponentially decaying component with frequency equal to 0, the time domain signal measured at a time $n\Delta t$ can be represented as

$$d_n = A\, e^{-\pi L n \Delta t} \qquad \text{(AII.1)}$$

where A is the amplitude and L is the line width ($1/\pi L$ is the decay time for the signal envelope). The 0 Hz value (i.e., the peak value) of the Fourier transform of an M-point time domain signal zero-filled to N points is

$$\frac{1}{\sqrt{N}}\sum_{n=0}^{M-1} d_n = \frac{A}{\sqrt{N}}\sum_{n=0}^{M-1} z^n = \frac{A}{\sqrt{N}}\frac{1-z^M}{1-z} \qquad \text{(AII.2)}$$

where $z = e^{-\pi L \Delta t}$. If the RMS noise amplitude in the time domain data is σ, Parseval's theorem (2) implies that the noise level in the DFT spectrum is $\sigma(M/N)^{1/2}$. The S/N is then

$$\frac{A}{\sigma\sqrt{M}}\frac{1-z^M}{1-z} \qquad \text{(AII.3)}$$

demonstrating the dependence on the number of samples and the line width. The partial derivative with respect to M is

$$\frac{-A}{\sigma M^{3/2}(1-z)}((1-z^M)/2 + Mz^M \ln z) \qquad \text{(AII.4)}$$

The value of M for which this expression is 0, M_{max}, is both the number of points for which S/N is maximum and also the point of diminishing returns, because exceeding this value leads to degradation of S/N. For $L = 30$ Hz and $\Delta t = (1/7600)$ s, M_{max} is 101.

JA011669O