JOURNAL OF MAGNETIC RESONANCE. Series A 116, 1–9 (1995)

Comparison of Fourier and Bayesian Analysis of NMR Signals. II. Examination of Truncated Free Induction Decay NMR Data

JOHN J. KOTYK,*,† NORMAN G. HOFFMAN,*,‡ WILLIAM C. HUTTON,*
G. LARRY BRETTHORST,§ AND JOSEPH J. H. ACKERMAN§,†

*Monsanto Company, 800 North Lindbergh, St. Louis, Missouri 63167; and §Department of Chemistry, Campus Box 1134, Washington University, One Brookings Drive, St. Louis, Missouri 63130-4899

Received June 23, 1994; revised November 17, 1994

In a recent report [J. J. Kotyk *et al., J. Magn. Reson.* **98**, 483 (1992)], estimates of NMR frequency and amplitude parameters obtained from Bayesian probability theory were shown to be more precise and more accurate than those obtained from the discrete Fourier transform (DFT). This previous study examined the effects of varying signal-to-noise ratio and did not address the performance of either probability theory or the DFT as a function of acquisition time, i.e., truncation. Herein, a quantitative comparison between probability theory and the DFT is presented and discussed in terms of the accuracy and precision they provide in estimating the frequency and amplitude from truncated free induction decay data. For simplicity, data containing only a single frequency are examined. For frequency estimation, Bayesian probability theory gives either more precise and accurate estimates or exactly the same estimates as the DFT. This latter result only occurs when the theory indicates that the Bayesian procedure is functionally identical to the DFT. For amplitude estimation, the results presented herein are consistent with those cited in previous work. Even for the best DFT procedure, probability theory outperforms the DFT by a factor of two or more depending on the signal-to-noise ratio and the prior information supplied in the analysis. The amplitude estimates from probability theory are more precise and/or more accurate than the DFT results for all levels of truncation. Additionally, the only time the DFT amplitude estimates are more precise (less uncertain) are when they are inaccurate (give an incorrect answer). Thus, the use of probability theory offers significant improvement over the use of DFT in highly truncated NMR free induction decay data.

INTRODUCTION

In practical applications of Fourier-transform NMR spectroscopy, experimental time considerations often limit the acquisition time (AT) during which the free induction decay is discretely sampled. A short AT relative to the transverse-relaxation decay time constant (T_2^*) of the NMR signal or coherence (e.g., AT $< 5.0T_2^*$) leads to the collection of truncated time-domain data. In turn, a discrete Fourier transform (DFT) of this truncated FID data exhibits deleterious effects that can make frequency and amplitude estimation difficult. Such effects have been extensively documented (*1–3*). For example, frequency-domain NMR spectra of truncated FID data lose resolution, have distorted lineshapes, and exhibit sinc oscillations. The problems associated with truncation artifacts are more pronounced in multidimensional NMR experiments where the AT in higher dimensions is severely restricted. A typical AT for an indirect-detected dimension in a three-dimensional or four-dimensional NMR experiment is usually between 10 and 40 ms (*4, 5*), which can be extremely short compared to the life times of single-quantum and multiple-quantum coherence(s). In some cases, as few as 8 or 16 complex data points may be acquired. Such extreme truncation of the time-domain data results in poor estimates of the frequencies and amplitudes.

Several different techniques have been used to analyze truncated NMR FID data. The most common approach is to use a time-domain apodization filter (e.g., a decaying exponential, Gaussian, cosine, Hanning, or Hamming filter) in conjunction with a DFT. While some improvements in frequency-domain spectral presentation are obtained by apodization of FID data (e.g., reduction of sinc oscillations), apodization functions have the disadvantage of broadening resonances and distorting lineshapes. A second common approach is to use nonlinear-least-squares curve fitting in the frequency domain. Fitting routines seek to improve NMR frequency and amplitude estimates by modeling phase twists, baseline distortion, and other common artifacts (*6*). A third approach relies upon linear prediction to time-forward extend the original time-domain FID data. This extended FID is then analyzed using the DFT. Such procedures have been extensively studied by others (*7–12*). Like the DFT, linear prediction is usually not used to estimate parameters directly; rather, both procedures rely on frequency-domain peak picking and digital integration to obtain parameter estimates. A fourth, more recent approach is to analyze the time-domain FID data directly using probability theory. This ap-

† To whom correspondence should be addressed.

‡ Current address: Tapestry Computing, Inc., 514 Earth City Expressway, Suite 125, Earth City, Missouri 63045.

1064-1858/95 $12.00

proach has the advantage of not being directly affected by either truncation artifacts or baseline distortions. For the use of probability theory to be routinely accepted, its ability to obtain precise and accurate estimates of NMR frequencies and amplitudes must be demonstrated and directly compared to the DFT procedures.

Recently, Bayesian probability theory (BPT) has been applied to the problem of NMR parameter estimation (*13–23*). Bayesian analysis is a general inference procedure that readily incorporates prior information about the NMR experiment into the analysis and by some very simple and appealing criteria is guaranteed to provide the best estimates possible of the NMR signal parameters. A quantitative comparison between the BPT and the DFT was published in earlier work (*24*). That study examined the frequencies and amplitudes estimated from nontruncated NMR FID data containing a single frequency as a function of the signal-to-noise ratio (S/N) of the data. For both BPT and the DFT, the precision of the parameter estimates depended on whether the signal phase (θ) and T_2^* were incorporated into the analysis. As theory indicated, the two analysis procedures gave equivalent frequency estimates when both θ and T_2^* are known and used in the analysis. In these circumstances, the BPT frequency estimation procedure is equivalent to peak picking from the frequency-domain *absorption* spectrum generated using a matched exponential apodization filter. Similarly, when T_2^* is known and θ is not known, the BPT frequency-estimation procedure is equivalent to peak picking from the frequency-domain DFT *power* spectrum generated using a matched exponential apodization filter. Under conditions where T_2^* or θ are unknown, the DFT performs much worse or completely fails at S/N levels where BPT succeeds. Under these conditions, even when the DFT is successful, BPT provides a factor of 3 to 7 improvement in the precision of the frequency estimates. Unlike frequency estimation, no equivalency between BPT and the DFT procedures is observed for amplitude estimation. At time-domain S/N levels where both procedures work, BPT is substantially more precise (about threefold) than the DFT. Like the frequency-estimation procedures, when T_2^* is not known, DFT amplitude estimation fails at S/N levels where BPT is still successful. A more detailed discussion of these findings for nontruncated data is presented elsewhere (*24*).

As a continuation of earlier studies, we now report a quantitative comparison between the DFT and BPT under conditions of increasing FID data truncation. By extending the comparison to data where $AT \ll 5T_2^*$, the extreme effects of truncation can be examined. Although this work examines only the single-frequency case, these quantitative comparisons are also relevant for multiple well-separated frequencies. The single-frequency model was chosen in order to explore direct comparisons of the strengths and weaknesses of the two procedures and to avoid complicating factors that accompany the analysis of overlapping resonances. Under ideal conditions, the frequency-domain area under an NMR spectral absorption-mode peak is proportional to the initial time-domain amplitude of the NMR resonance. Hence, we will refer to both quantities as the NMR signal *amplitude.*

EXPERIMENTAL METHODS

Experimental Design

NMR frequency and amplitude estimates from simulated single-frequency FID data (frequency = 1000 Hz; amplitude = 100; noise standard deviation = 1; T_2^* = 0.2 s; linewidth = 1.5 Hz; sweep width = 65,536 Hz; 65,536 complex points) were obtained using the DFT and BPT procedures described below. The simulated data are equivalent to conditions where the $AT = 5T_2^*$. The data were analyzed without truncation (AT = 1.0 s) and at 50 increasing levels of truncation, where each increase in truncation resulted in a 15% loss of the number of complex data values. For example, the first truncated data set contains 55,492 complex points (i.e., 85% of the original 65K FID data; AT = 0.85 s), while the last truncated data set contains only 16 complex data points (AT = 0.244 ms). The advantage of using simulated data is that the true parameter values are known.

At each truncation level, 50 different individual sets of Gaussian white noise with fixed variance were generated (*25*) and individually added to the truncated FID data to generate a series of noise-containing FID data sets. Computer-generated noise used in this fashion gives statistically identical results to noise introduced by the spectrometer when the audio filter bandwidth is set to twice the spectral window (*24*). Each of the 2500 unique *combined data sets* (i.e., FID signal + Gaussian noise) contained a time-domain S/N of 100. This corresponds to a frequency-domain S/N of 1630 for the Fourier-transformed nontruncated data sets when a matched exponential filter was used. Relatively high S/N was chosen so as to explore the effects of increasing truncation, not the effects of decreasing S/N as was presented earlier (*24*). At each truncation factor, the standard deviation and variance of the NMR parameter estimates were computed from the 50 independent parameter estimates made (one each of frequency and amplitude). All calculations and data processing were performed using either automated C-shell scripts or FORTRAN programs developed in our laboratories.

Discrete Fourier-Transformation Procedures

Each combined data set (signal plus noise) at each truncation level was zero-padded to a Fourier number of 256K complex data values and then analyzed using a DFT to obtain both the frequency and amplitude estimates. Frequency estimates were obtained by peak picking the highest point in the frequency-domain NMR spectrum, while the integrated area or signal amplitudes were determined by summing the

spectrum over the region of the frequency-domain spectrum defined by the integration limits. As in our earlier work (*24*), ancillary measures were taken to help improve the robustness of the software-driven DFT procedures. This hidden or implicit use of prior information *is required* to prevent premature catastrophic failure of DFT procedures and is not inconsistent with the intuition guiding a trained spectroscopist.

Two DFT procedures using different sets of integration limits were used to obtain amplitude estimates. In the DFT_f procedure, the integration limits were fixed to include $\pm 12.67 \times$ (NMR signal linewidth), known for the nontruncated data in the presence of a matched exponential filter. This provides 95% of the integrated area in the absence of truncation. In the DFT_v procedure, the integration limits were varied as a function of the number of data points needed to give exactly 95% of the integrated area when a matched filter was applied. This required consideration of the actual (not idealized) frequency-domain lineshape. As discussed, use of variable integration limits based on the actual frequency-domain lineshape helps compensate for truncation effects. Setting the DFT integration limits (fixed or variable) in this fashion requires prior knowledge of T_2^* (or linewidth). Table 1 illustrates the changes in the DFT_v integration limits at selected truncation levels. Such an approach employing direct digital integration is obviously not generally feasible with actual multiresonance NMR FID data.

For the DFT, the following procedures were used: first, frequency estimates were obtained either by peak picking the absorption spectrum, when the phase, θ, was assumed known, or by peak picking a power spectrum when the phase, θ, was not given. All amplitude estimates were obtained by integration over the absorption spectrum. The phase of the absorption spectrum was set to either the "true" phase when θ is assumed known or the maximum-likelihood phase estimate obtained from the peak of the power spectrum. In addition to a T_2^* matched exponential filter, DFT estimates were obtained using one of the following common apodization filters: Gaussian, cosine, Hanning, Hamming, resolution enhancing, or an exponential (not T_2^* matched). In total, seven different apodization filters were employed. All apodization filters (except the T_2^* matched exponential filter) were set so that the amplitude of the last few data points in the truncated FID data were approximately 1/1000th the amplitude of the initial data point in the FID. The matched exponential filter was set based on the known T_2^* for the signal in the nontruncated data. Two frequency and four amplitude estimates (i.e., those with or without known phase for either fixed or variable integration limits) were obtained for each apodization filter for each of the 2500 free induction decays.

Bayesian Analysis Procedures

Each combined data set (signal plus noise) was analyzed using BPT to obtain the posterior probability distributions for the frequency and signal amplitude. A detailed description and theoretical basis for these procedures can be found elsewhere (*16–20*). Frequencies were estimated both in the absence and in the presence of θ and T_2^* prior information as indicated.

RESULTS AND DISCUSSION

The present work examines the ability of both BPT and the DFT to estimate the frequency and the amplitude of a single exponentially decaying sinusoid as a function of signal truncation. Comparisons are made between the BPT results and the DFT results obtained with various integration limits and apodization functions. The true parameter values were frequency = 1000 Hz and signal amplitude = 100. In the case of the DFT estimates, the integration limits were selected to give 95% of the known value. Any consistent deviation or scatter from these values reflects a lack of accuracy or precision in the analysis procedures caused by data truncation.

TABLE 1

Summary of NMR Parameters and DFT_v Integration Limits[a] at Various Truncation Levels

AT	Complex data points	Truncation level	No. of cycles digitized	DFT_v integration limits (pts)[b]	DFT_v integration limits (lw)[c]
1.0	65536	$5.0T_2^*$	1000	±152	± 12.67
0.20	13052	$1.0T_2^*$	200	±140	± 11.65
0.02	1314	$0.1T_2^*$	20	±129	± 10.65
0.002	130	$0.01T_2^*$	2	±1159	± 96.56
0.0002	16	$0.001T_2^*$	0.2	±9569	±797.42

[a] Variable integration limits were determined by applying a T_2^* matched exponential filter to the known exponentially decaying single-frequency signal, performing a DFT, and then finding the symmetric integration limits that enclosed 95% of the known and total integrated area.

[b] Integration limits of the fourfold zero-padded (65,536→262,144 complex points) frequency-domain spectrum. There are four points per hertz.

[c] Integration limits expressed in units of linewidth at half-height in the absence of truncation but in the presence of a T_2^* matched exponential filter; e.g., 1 lw = 3.0 Hz for the simulated signal used herein.

A nontruncated signal is defined in the present study as one that is digitized for at least $5T_2^*$. The degree of data truncation is discussed relative to the NMR signal in two ways. Most fundamentally, it can be presented in terms of the AT relative to the signal's T_2^* value, i.e., in units of T_2^*. Nontruncated data, for instance, is equivalent to collecting data with an AT = $5.0T_2^*$, while highly truncated data might be collected with an AT = $0.001T_2^*$. Additionally, the truncation level can be discussed in terms of the number of cycles in the simulated sinusoid. The nontruncated FID data set had 1000 cycles while the most highly truncated FID data set examined has only 0.2 cycles.

As described previously (*24*), inclusion of θ and T_2^* prior information markedly improves the parameter estimates obtained via DFT procedures but is less important in the BPT analysis. In fact, without T_2^* prior information, the DFT performance is so poor relative to BPT performance (see Ref. (*24*)) that it was decided to use T_2^* prior information for all DFT analyses in the present study. Such actions not only allow somewhat more practical comparisons to be made (the trained spectroscopist often knows the approximate value of T_2^*) but are also necessary for determining DFT integration limits.

Frequency Estimates

Figure 1 shows the BPT and the DFT frequency estimates plotted as a function of truncation. Both BPT and this DFT analysis give accurate frequency estimates with minimal scatter at truncation levels where AT > $0.03T_2^*$ (six cycles digitized). Both procedures are adversely affected by more severe truncation, i.e., the scatter increases markedly when AT < $0.03T_2^*$. At truncation levels beyond AT = $0.03T_2^*$, the precision of the BPT frequency estimates obtained without either θ or T_2^* prior information (Fig. 1A) is only slightly worse than that found for the DFT power spectrum given T_2^* (Fig. 1B). The BTP results given T_2^* (Fig. 1C) are identical to the DFT, as theory says they must be. Additional improvement in the DFT frequency estimates is observed when both T_2^* and θ prior information is used (Fig. 1D). Again theory indicates in this latter case that the BPT and the DFT results are identical. Use of apodization functions other than a matched exponential filter in the DFT analysis (data not shown) shows very little difference from the behavior displayed in Fig. 1B or Fig. 1D. Figure 1 demonstrates that both BPT and the DFT frequency estimates improve when more prior information is supplied and become identical when theory indicates they should. These findings are consistent with our previous study (*24*). Note that the small differences in the BPT and DFT estimates observed in the figures are due to the digital resolution implicit in the DFT. If BPT is restricted to that same digital resolution, the BPT results are identical to those obtained from the DFT.

The truncation of the FID data results in a progressive loss in the frequency-domain S/N level (excessive truncation

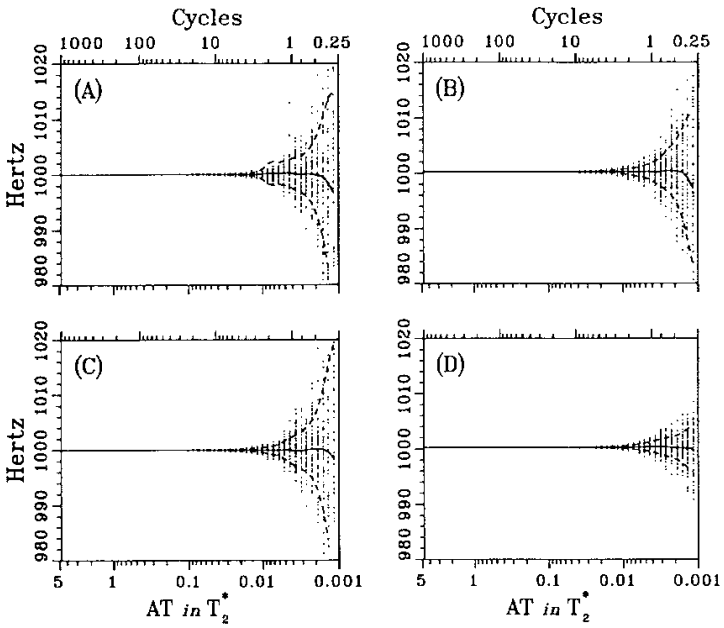

FIG. 1. Individual frequency estimates obtained using the BPT and the DFT analysis procedures plotted as a function of the FID data truncation level. Results are shown for (A) the BPT estimates obtained in the absence of T_2^* and θ prior information; (B) the DFT estimates obtained with T_2^*, but without θ prior information, i.e., from the frequency-domain power spectrum; (C) the BPT estimates obtained with T_2^*, but without θ prior information; and (D) the DFT frequency estimates obtained with T_2^* and θ prior information from the frequency-domain absorption spectrum. Fifty independent frequency estimates were made at each of 50 truncation levels for a total of 2500 individual estimates in each panel. The solid line shown in (A)–(D) represents the mean of the 50 individual frequency estimates obtained at each truncation level and is an indication of the accuracy of the estimation, while the upper and lower dashed lines represent one standard deviation from the mean, respectively, and reflect the precision of the estimation.

broadening decreases resonance lineshape height), but no loss in the time-domain S/N. Herein, even for the most truncated data, the S/N is sufficient to give frequency estimates that only deviate by 20 Hz from the true value of 1000 Hz. Thus, the precision of the frequency estimates is relatively good. The discrete nature of the fast Fourier transform degrades the performance of the DFT procedures. Effects due to coarse digitization even appear in the DFT frequency estimates (1000.25 Hz) obtained for the nontruncated data. Additionally, digitization also affects the amplitude estimates at all levels of truncation. These digitization effects appear because the lineshape changes as a function of the number of complex data values. The changing lineshape has the paradoxical affect of causing the average DFT area per frequency step either to decrease (so the integral is less than 95) or to increase (so the integral is larger than 95), depending on the degree of truncation. The size of the effect can be quite large. By comparison, neither the frequency-domain S/N level nor the discrete nature of the peak-picking routines are directly relevant to the BPT procedures. The mean BPT amplitude and frequency estimates do not display digitization artifacts like those observed in the DFT analysis.

One advantage of BPT is that, in addition to estimating the parameters, it also carries an estimate of the accuracy and precision of the estimated parameters. By performing a calculation similar to that done previously (*18, 19*), the accuracy and precision of the BPT frequency estimates at extreme levels of data truncation (linear approximation of exponential decay) can be expressed in terms of the experimental and signal parameters

$$(F)_{\text{est}} = \hat{F} \pm \frac{\sigma T_2^*}{8\hat{A}\sqrt{N}}, \qquad [1]$$

where $(F)_{\text{est}}$ is the estimated frequency, $\hat{F}$ is the true frequency, σ is the standard deviation for the noise, $\hat{A}$ is the true signal amplitude, and N is the number of discrete data values. Equation [1] shows that the precision of the BPT frequency estimates depends on, among other factors, the time-domain S/N level ($\hat{A}/\sigma$) and the square root of the number of data values gathered. Loss of accuracy in the BPT frequency-estimation results for constant time-domain S/N level is caused by a reduction in the number of data points. Also note that the BPT analysis is accurate in the sense that, as the noise level goes to zero, the BPT procedure always returns the true frequency. A more detailed discussion of the dependence of the BPT frequency estimates on the experimental parameters can be found in Refs. (*18, 19*).

Signal Amplitude Estimates

The BPT and DFT signal amplitude estimates obtained for the *nontruncated* FID data are summarized in Table 2. The BPT and the DFT (matched) procedures give the expected estimates of the signal amplitude, i.e., 100 and 95%, respectively. Not unexpectedly, application of apodization filters distorts the lineshape so that the fixed integration limits no longer define 95% of the integrated area of the NMR signal. Thus, differences in amplitude estimates for nontruncated data provide a relative measure of the incompatibility of the apodization filter and Lorentzian-based integration region. Such apodization-induced accuracy errors are also present, and in some cases greatly magnified, in truncated data analysis and have a significant effect upon the performance of the DFT amplitude estimation procedures (*vide infra*).

Figures 2 and 3 display the amplitude estimates for the BPT analysis and the DFT analysis using fixed and variable integration limits, respectively. In both figures, the abscissa scales are identical for all eight panels (A–H) and represent the AT in units of T_2^*. The ordinate scale in Fig. 2 and Fig. 3 varies depending on the analysis procedure. In most panels, the ordinate scale spans an amplitude range of 100 (Figs. 2B–2H and Figs. 3C–3H), while other panels display expanded ordinate scales: Fig. 2A spans a range of 8 and Figs. 3A and 3B each span a range of 4.

TABLE 2
Signal Amplitude Estimates (Mean) Obtained for Nontruncated Data Analysis

Method (apodization filter)[a]	Nontruncated amplitude[b]
BPT	99.995
DFT (matched)	95.002
DFT (unmatched)	98.369
DFT (Gaussian)	99.973
DFT (cosine)	97.488
DFT (Hanning)	97.488
DFT (Hamming)	97.487
DFT (enhancing)	100.804

[a] Functional forms for all DFT apodization functions were obtained from Ref. (*1*). All apodization filters other than the matched exponential filter were specifically set so that the amplitude of the last few data points in the truncated FID data were approximately 1/1000th the amplitude of the initial data point in the FID.

[b] Frequency-domain integration limits were set as is appropriate to recover 95% of the area for the nontruncated signal in the presence of a matched exponential filter.

A comparison of each procedure can be made at different truncation levels by examining both the mean and standard deviation of the parameter estimate. While the BPT analysis (Figs. 2A and 3A) gives reasonable amplitude estimates at all truncation levels, the performance of the DFT procedures depends strongly on the degree of data truncation, the integration limits, and the choice of apodization filter. In general, the DFT amplitude estimates at minor to moderate levels of data truncation yield reasonable results and follow the trend established for the nontruncated data in Table 2. At high truncation levels, all of the DFT procedures, with the possible exception of Fig. 3B, fail to yield accurate estimates of the signal amplitude.

As previously noted, seven different apodization functions were used in the DFT_f analysis. Of these seven, five were reasonably precise at moderate truncation levels (Figs. 2B, 2D, 2E, 2F, and 2G), but gave inaccurate or biased results at higher truncation levels. Two apodization functions (Figs. 2C and 2H) worked poorly at all levels of truncation. By contrast, the Bayesian estimate (Fig. 2A) worked well for all levels of truncation, exhibiting both accurate and precise results. The results from all the DFT procedures are inaccurate for high levels of truncation. The inaccuracies in the DFT_f analysis are caused by the fact that the integration limits are fixed, but the linewidth changes as a function of the truncation. This is most apparent when using a matched filter (Fig. 2B). Note how the amplitude estimate decreases and then increases and how finally most of the area moves completely out of the integration window. This behavior is caused both by the height of the peak and by the position of the sinc minima and maxima changing as a function of the number of complex data values.

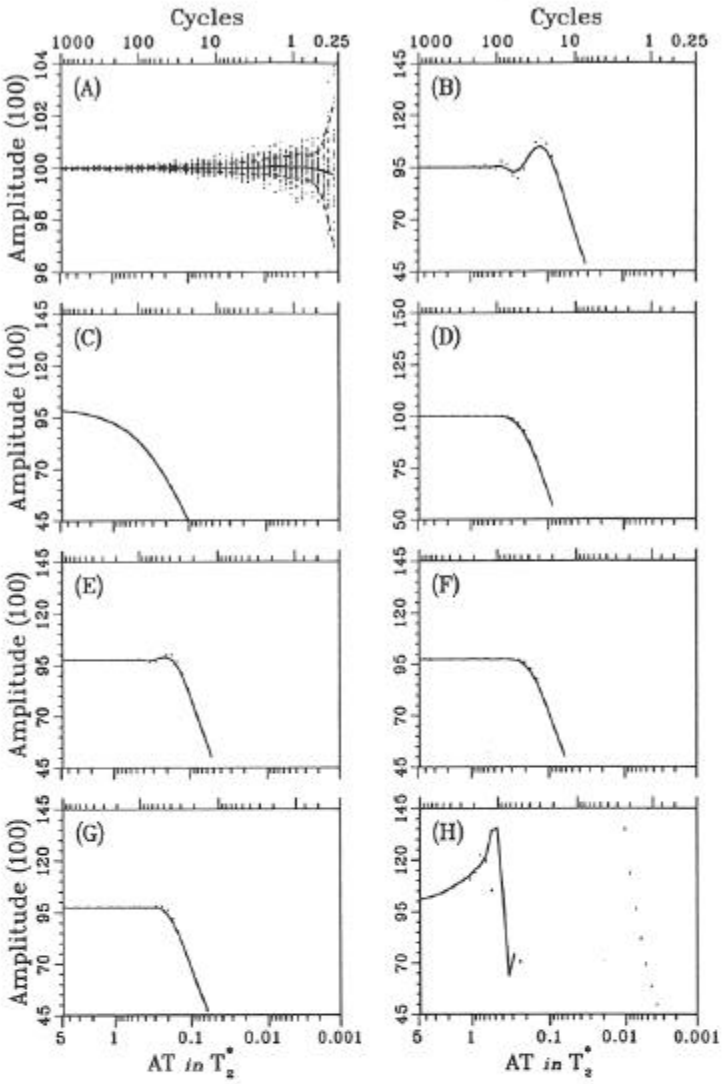

FIG. 2. Individual signal amplitude estimates obtained using the BPT and the DFT_f analysis procedures plotted as a function of the FID data truncation level. Results are shown for (A) the BPT procedure estimates obtained using T_2^* prior information (note expanded ordinate scale) and the DFT_f procedure estimates obtained using integration limits fixed as is appropriate to recover 95% of the area for the nontruncated signal in the presence of a matched exponential filter. These integration limits were then used with the following apodization filter functions: (B) T_2^* matched decaying exponential; (C) decaying exponential; (D) Gaussian; (E) cosine; (F) Hanning; (G) Hamming; and (H) resolution enhancing. Except for (B), all filters were applied such that the final point in the FID was reduced to approximately 1/1000th of the amplitude of the initial point in the FID. All DFT estimates were obtained with θ prior information, i.e., from the absorption spectrum. Fifty independent amplitude estimates were made at each of 50 truncation levels for a total of 2500 individual estimates in each panel. The solid line shown in (A)–(H) represents the mean of the 50 individual frequency estimates obtained at each truncation level and is an indication of the accuracy of the estimation, while the upper and lower dashed lines represent one standard deviation, respectively, and reflect the precision of the estimation.

To help compensate for this effect, amplitude estimates were obtained a second time using variable integration limits, i.e., DFT_v. These integration limits were set so that the matched filter integral enclosed at least 95% of the total area. To determine these limits, the simulated data were multiplied by a matched filter, transformed, and phased, and the smallest symmetric interval was found that enclosed at least 95% of the total amplitude. Because the DFT has finite resolution,

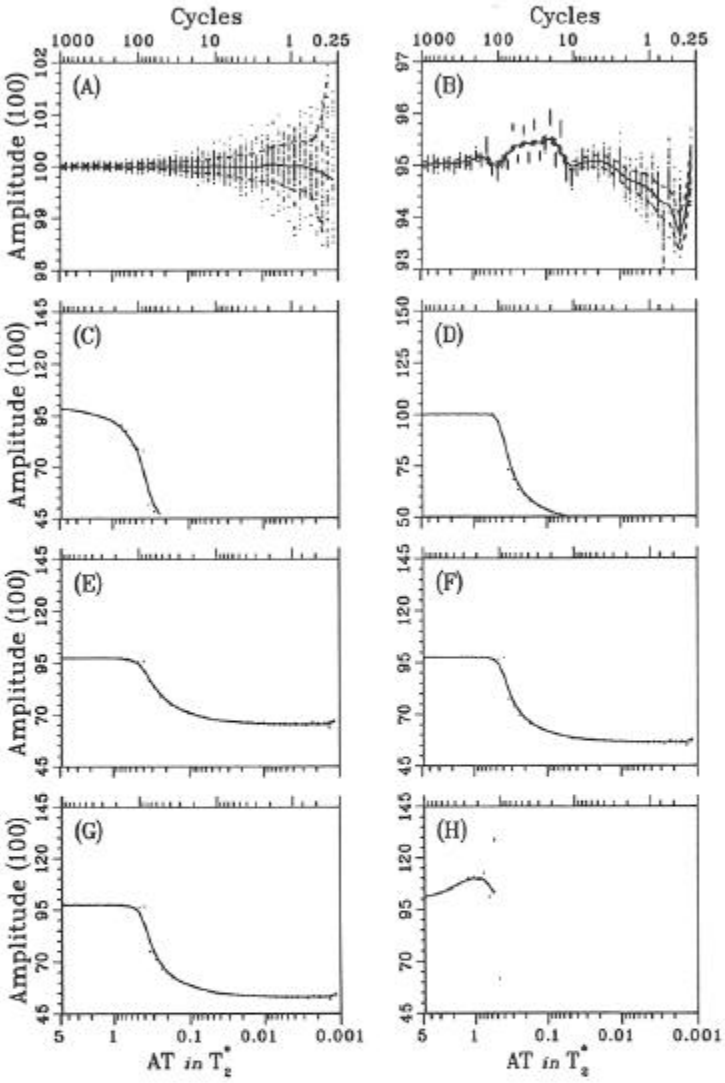

FIG. 3. Individual signal amplitude estimates obtained using the BPT and the DFT_v analysis procedures plotted as a function of the FID data truncation level. Results are shown for (A) the BPT procedure estimates obtained using T_2^* prior information (same as Fig. 2A but with further expansion of ordinate scale) and the DFT_v procedure estimates obtained using integration limits varied according to the true lineshape expected of a T_2^* matched exponentially apodized truncated signal. Panels represent the following apodization filter functions: (B) T_2^* matched decaying exponential (ordinate scale expanded equivalent to that of (A); (C) decaying exponential; (D) Gaussian; (E) cosine; (F) Hanning; (G) Hamming; and (H) resolution enhancing. Except for (B), all filters were applied such that the final point in the FID was reduced to approximately 1/1000th of the amplitude of the initial point in the FID. All DFT estimates were obtained with θ prior information, i.e., from the pure absorption spectrum. Fifty independent amplitude estimates were made at each of 50 truncation levels for a total of 2500 individual estimates in each panel. The solid line shown in (A)–(H) represents the mean of the 50 individual frequency estimates obtained at each truncation level and is an indication of the accuracy of the estimation, while the upper and lower dashed lines represent one standard deviation, respectively, and reflect the precision of the estimation.

the actual area enclosed using this approach is always a little more than 95%. The results of this analysis are shown in Figs. 3B–3H. Note the continued presence of digitization artifacts. These artifacts are unavoidable and are implicit in the discrete nature of the amplitude integral.

Table 3 summarizes the amplitude-estimation failure points for all the analytical procedures. For fixed integration limits and moderate levels of truncation, more accurate amplitude estimates are obtained using a Gaussian, cosine, Hanning, or Hamming filter than those obtained using a matched exponentially decaying filter. On the other hand, variable integration limits correct for sinc oscillation artifacts (Fig. 3B) and dramatically improve the accuracy of the matched exponential filter results, but show a decrease in the accuracy of the results for the other filters because an incorrect lineshape is assumed in setting the variable integration limits. In practice, meaningful DFT amplitude estimates can be obtained only at truncation levels less than those shown in Table 3.

A closer comparison of Figs. 2 and 3 and Table 3 show all DFT_v procedures, except DFT_v (matched), begin to fail at truncation levels lower than their corresponding DFT_f procedures. For instance, the DFT_v (Hanning) amplitude estimates (Fig. 3F) begin to fail when $AT \approx 0.9T_2^*$ (160 cycles digitized), whereas the DFT_f (Hanning) procedure (Fig. 2F) provides reasonable amplitude estimates until $AT \approx 0.3T_2^*$ (60 cycles digitized). This seemingly counterintuitive behavior for other apodization functions, i.e., poorer performance for DFT_v vs DFT_f, is observed because the variable integration limits used were specifically set for a matched exponential apodization function. Accurate amplitude determination using other apodization functions would require a different calculation for each apodization function to determine the appropriate integration limit. At high data truncation levels, where extremely large integration limits are used, some improvements in DFT_v amplitude estimates are seen for other apodization functions (Figs. 3E, 3F, and 3G). This is demonstrated by the estimates remaining on scale and leveling out at values closer to the known amplitude (compared to the DFT_f results shown in Fig. 2).

Lineshape distortion in the frequency domain is introduced by apodization and by truncation of the time-domain data. The net effect of lineshape distortion is to distribute the signal amplitude such that the integration limits set for a Lorentzian lineshape ($\pm 12.67 \times$ known NMR signal linewidth) are no longer appropriate. This leads to bias or systematic error in the accuracy of the DFT amplitude estimates which eventually results in the failure of the DFT procedures as shown in both Figs. 2 and 3. Unfortunately, at high levels of data truncation, impractically large variable integration limits are necessary to compensate for the changes in lineshape. For instance, to include 95% of the signal amplitude for a resonance with a true linewidth of 1.5 Hz, an integration region of 2392 Hz is required at $AT = 0.0002T_2^*$ (0.2 cycles digitized) when a matched filter is used, while only ±32 to 38 Hz is needed for $AT \geq 0.02T_2^*$ (20 cycles digitized). Although generally impractical in the presence of multiple resonances, variable integration limits account for the variability of the lineshape as a function of AT. However, even after accounting for the change in lineshape, variable integration is still effected by systematic digitization artifacts.

TABLE 3
Amplitude Estimation Failure Points for BPT and DFT Analysis Procedures Reported in Terms of T_2^*

Method[a] (apodization filter)	Using fixed integration limits[b] (cycles digitized)	Using variable integration limits[c] (cycles digitized)
BPT	$0.002T_2^*$ (0.4)	
DFT (matched)	$1.0T_2^*$ (200)	$0.002T_2^*$ (0.4)
DFT (unmatched)	—	—
DFT (Gaussian)	$0.4T_2^*$ (80)	$0.7T_2^*$ (140)
DFT (cosine)	$0.4T_2^*$ (80)	$0.9T_2^*$ (180)
DFT (Hanning)	$0.3T_2^*$ (60)	$0.9T_2^*$ (180)
DFT (Hamming)	$0.3T_2^*$ (60)	$0.8T_2^*$ (160)
DFT (enhancing)	—	—

[a] Functional forms for all DFT apodization functions were obtained from Ref. (*1*). All apodization filters other than the matched exponential filter were set so that the amplitude of the last few data points in the truncated FID data were approximately 1/1000th the amplitude of the initial data point in the FID.

[b] Frequency-domain integration limits were set as is appropriate to recover 95% of the area for the nontruncated signal in the presence of a matched exponential filter.

[c] Variable integration limits were chosen to bound 95% of the area for the actual DFT lineshape resulting from the matched exponentially filtered truncated data set.

In summary, accurate and precise amplitude estimates can only be obtained from the DFT for moderate levels of truncation. At high levels of truncation, all DFT procedures give amplitude estimates that are inaccurate. It is possible using variable integration limits specific for a given apodization function and truncation level to correct for these inaccuracies. However, digitization artifacts persist even here, which cause biased estimates to be obtained. By contrast, BPT amplitude estimation is free of artifacts and performs as expected.

In many applications where highly truncated data is collected, e.g., multidimensional experiments, this loss in amplitude accuracy may not be as disastrous as it first appears. Provided all resonances have roughly the same linewidth, strong apodization functions can be expected to dominate the linewidth and affect all resonances equally. In addition, such experiments often measure relative amplitudes which are compared to an internal control amplitude observed for a known resonance, e.g., as in the determination of distances derived from NOE cross-peak amplitudes/volumes. Thus,

to some extent, the loss in accuracy is compensated for by use of relative amplitude measurements.

Like the BPT frequency estimates, a straightforward calculation similar to that described earlier (*18, 19*) may be used to exhibit the dependence of the BPT amplitude estimates on the experimental parameters in highly truncated FID data (linear approximation of exponential decay). In this case one finds

$$(A)_{\text{est}} = \hat{A} + \frac{\sigma}{\sqrt{N}}, \qquad [2]$$

where $(A)_{\text{est}}$ is the estimated amplitude, $\hat{A}$ is the true amplitude, σ is the standard deviation of the noise, and N is the number of discrete data values. Thus, for constant time-domain S/N levels, the precision of the BPT amplitude estimate in Eq. [2] has the same dependence on N as BPT frequency estimates, Eq. [1]. While an increase in performance of the BPT amplitude (and BPT frequency) estimates might be expected at higher truncation levels by oversampling the data, in practice, these effects are tempered by the effects of audio filters. A more detailed discussion of the dependence of BPT amplitude estimates on experimental parameters can be found in Ref. (*20*).

Equations [1] and [2] demonstrate that the precision of the parameter estimate varies as one over the square root of the number of data values. Thus for a fixed value of the noise standard deviation, σ, one should expect less accurate results for both amplitude and frequency estimates as the number of data values decreases. For example, for a fixed noise standard deviation, the estimates obtained from a data set containing 128 data values are two times more precise than those obtained from a data set containing 32 data values, and four times more precise than those obtained from a data set containing only 8 data values. Of course, as the number of resonances increases, the number of parameters to be estimated increases. In the case of a very limited number of data values, these parameters will be almost unconstrained by the data, i.e., their values will be completely indeterminant unless one has strong prior information.

A closer comparison between BPT and DFT procedures can be made by examining Fig. 4 which plots the amplitude standard deviation vs the mean-amplitude estimate at selected truncation levels. Because variable integration limits can rarely be used in practice, the comparisons in Fig. 4 are made by using fixed integration limits. At all truncation levels shown, the BPT analysis provides more accurate and precise amplitude estimates than any of the DFT_f procedures. The mean BPT amplitude estimate falls exactly at the known amplitude of 100, while many of the DFT_f results give an estimate other than the expected value of 95. Consistent with earlier discussions (Table 3), the DFT_f(matched) procedure provides the most accurate DFT_f estimates at the first three truncation levels shown (Figs. 4A, 4B, and 4C). At higher truncation levels, the DFT_f procedures perform very poorly yielding amplitude estimates and standard deviations which are off the scale (Figs. 4D, 4E, and 4F). While a seemingly more precise amplitude estimate is observed for the DFT_f (matched) analysis at $\text{AT} = 0.1T_2^*$ (Fig. 4D), such precision is fortuitous in that the analysis has failed long before this level of truncation was reached (see Table 3 and Fig. 2B).

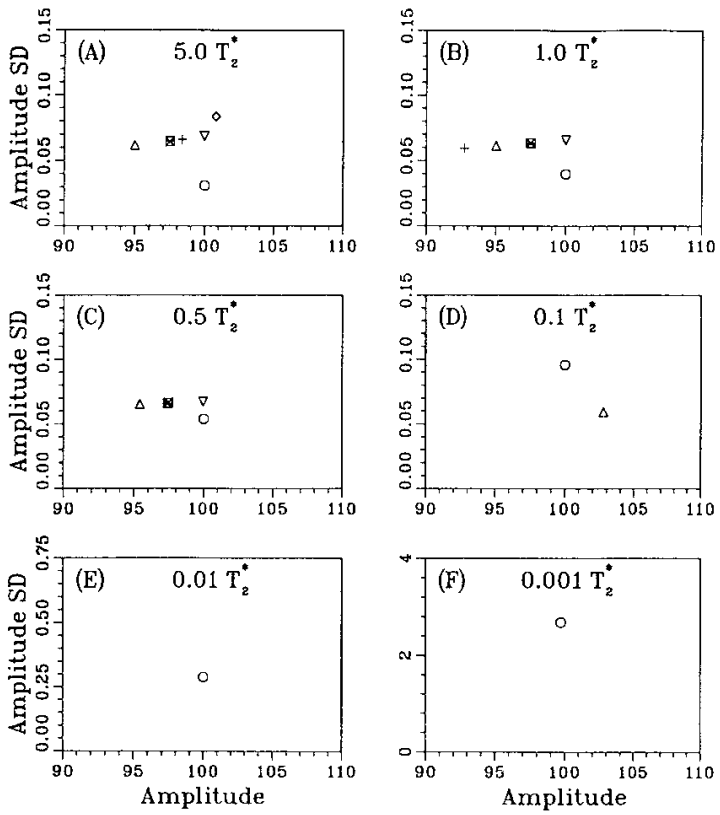

FIG. 4. Amplitude standard deviation vs mean amplitude estimate at selected FID data truncation levels for the BPT analysis (○) and the DFT_f procedures using a T_2^* matched exponential (△), a decaying exponential (+), a Gaussian (▽), a cosine (×), a Hanning (x), an Hamming (⊠), or an enhancing (◇) apodization filter. All DFT estimates were obtained with θ prior information, i.e., from the absorption spectrum. Fifty independent amplitude estimates were made at each of 50 truncation levels for a total of 2500 individual estimates in each panel. Accurate mean amplitude estimates for the BPT procedure should be 100, while accurate DFT procedures should yield a value of 95. While a seemingly more precise amplitude estimate is observed for the DFT_f(matched) analysis at $\text{AT} = 0.1T_2^*$ (D), such precision is fortuitous in that the analysis has failed long before this level of truncation was reached (see Table 3 and Fig. 2B).

SUMMARY AND CONCLUSIONS

Our results demonstrate that BPT offers distinct improvements over the DFT as a means to obtain precise and accurate estimates of the frequency and signal amplitude from single-frequency, truncated FID data. The source of improvement lies in fundamental differences between analysis using DFT frequency-domain and BPT time-domain procedures. The DFT relies on peak-picking and digital inte-

gration techniques and is greatly affected by distortions in lineshape and data truncation. BPT does not rely on frequency-domain peak-picking and lineshape integration limits. It is capable of incorporating prior information even in cases where it is impractical to do so using DFT procedures. The advantages observed using BPT exist for both nontruncated and truncated FID data, but are most pronounced for data with low frequency-domain S/N and/or high truncation levels.

When T_2^* prior information is supplied, BPT and the DFT procedures are equivalent for estimating frequency at all levels of data truncation. When θ is not known, both procedures reduce to peak picking a power spectrum, and when θ is known, both procedures reduce to peak picking an absorption spectrum. Such findings are consistent with earlier work (*24*) and are based both on theory and on empirically measured uncertainties determined from repetitive estimates of the NMR parameters. Regardless of the degree of data truncation, under practical conditions where neither θ nor T_2^* are known, BPT analysis can be expected to yield more accurate and precise estimates of the frequency than the DFT method as described in this and earlier work (*24*).

At increasing truncation levels, the BPT analysis gives more accurate and precise estimates of the signal amplitude. Only when θ and T_2^* prior information is used and integration limits are chosen based on the true lineshape does the accuracy and precision of the two procedures become comparable. Even then, the DFT results exhibit bias in the amplitude estimates. Obviously, this strategy would not be feasible in spectra where lineshapes overlap from multiple resonances. Inappropriately accounting for lineshape distortion in the frequency domain is the single major source of accuracy error in the DFT procedures. As truncation levels increase, all DFT procedures, except the DFT_v (matched) procedure, completely fail to provide accurate estimates of the signal amplitude. Use of apodization filters, other than a matched exponentially decaying filter, exacerbates the problems due to lineshape distortion and results in failure of the DFT procedure at moderate data truncation levels.

For practical applications, where T_2^* and θ are unknown, significantly more accurate and precise parameter estimates can be obtained using BPT. Furthermore, in contrast to direct DFT procedures, BPT analysis can be applied to signals containing closely spaced frequency components. In this regard, we anticipate that Bayesian analysis will become an important method for examining time-domain FID or interferogram NMR data.

ACKNOWLEDGMENTS

This work was supported by the Monsanto Company and, in part, by NIH Grant GM-30331.

REFERENCES

1. J. C. Lindon and A. G. Ferrige, *Prog. NMR Spectrosc.* **4,** 22 (1980).
2. R. R. Ernst, G. Bodenhausen, and A. Wokaun, "Principles of Nuclear Magnetic Resonance in One and Two Dimensions," Oxford Univ. Press, Oxford, 1987.
3. G. Otting, H. Wider, G. Wagner, and K. Wüthrich, *J. Magn. Reson.* **66,** 187 (1986).
4. M. Ikura, L. E. Kay, and A. Bax, *Biochemistry* **29,** 4659 (1990).
5. L. E. Kay, G. M. Clore, A. Bax, and A. M. Gronenborn, *Science* **249,** 411 (1990).
6. G. H. Weiss and J. A. Ferretti, *J. Magn. Reson.* **55,** 397 (1983).
7. D. S. Stephenson, *Prog. NMR Spectrosc.* **20,** 515 (1988).
8. S. Sibisi, *Nature (London)* **301,** 134 (1983).
9. S. Sibisi, J. Skilling, R. G. Brereton, E. D. Laue, and J. Staunton, *Nature (London)* **311,** 446 (1984).
10. E. D. Laue, J. Skilling, J. Staunton, S. Sibisi, and R. G. Brereton, *J. Magn. Reson.* **62,** 437 (1985).
11. J. F. Martin, *J. Magn. Reson.* **65,** 291 (1985).
12. F. Ni, G. C. Levy, and H. A. Scheraga, *J. Magn. Reson.* **66,** 385 (1986).
13. G. L. Bretthorst, *in* "Lecture Notes in Statistics" (J. Berger, S. Fienberg, J. Gani, K. Krickeberg, and B. Singer, Eds.), Vol. 48, Springer-Verlag, New York, 1988.
14. G. L. Bretthorst, C.-C. Hung, D. A. d'Avignon, and J. J. H. Ackerman, *J. Magn. Reson.* **79,** 369 (1988).
15. G. L. Bretthorst, J. J. Kotyk, and J. J. H. Ackerman, *Magn. Reson. Med.* **9,** 282 (1989).
16. G. L. Bretthorst, *J. Magn. Reson.* **88,** 533 (1990).
17. G. L. Bretthorst, *J. Magn. Reson.* **88,** 552 (1990).
18. G. L. Bretthorst, *J. Magn. Reson.* **88,** 571 (1990).
19. G. L. Bretthorst, *J. Magn. Reson.* **93,** 369 (1991).
20. G. L. Bretthorst, *J. Magn. Reson.* **98,** 501 (1992).
21. J. J. Neil and G. L. Bretthorst, *Magn. Reson. Med.* **29,** 642 (1993).
22. K. S. Vines, R. F. Evilia, and S. L. Whittenburg, *J. Magn. Reson.* **100,** 195 (1992).
23. K. S. Vines, R. F. Evilia, and S. L. Whittenburg, *J. Phys. Chem.* **97,** 4941 (1993).
24. J. J. Kotyk, N. G. Hoffman, W. C. Hutton, G. L. Bretthorst, and J. J. H. Ackerman, *J. Magn. Reson.* **98,** 483 (1992).
25. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, "Numerical Recipes: The Art of Scientific Computing," Cambridge Univ. Press, Cambridge, 1989.