JOURNAL OF MAGNETIC RESONANCE **100**, 202–207 (1992)

Improved Linear Prediction of Damped NMR Signals Using Modified "Forward–Backward" Linear Prediction

GUANG ZHU*,† AND AD BAX*

** Laboratory of Chemical Physics, National Institutes of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892; and †Chemical Physics Program, University of Maryland, College Park, Maryland 20742*

Received July 16, 1992

Linear prediction (LP) has become a standard tool for enhancing the appearance of multidimensional NMR spectra (*1–7*). In principle, the method can be used to calculate the frequencies, amplitudes, damping factors (linewidths), and phases of all components contained in the time-domain signal. In practice, however, the method is not very robust in the presence of noise. For this reason, a more conservative approach, where the time-domain signal is extended using (imperfect) linear prediction coefficients, has been much more popular. In the latter case, the appearance of the final spectrum frequently is dominated by the acquired time-domain data, but truncation at the end of the acquired time domain is minimized by elongating the acquired data by predicted data. The weight of these predicted data is decreased by appropriate digital filtering prior to Fourier transformation. Here we present a simple modification of this commonly used procedure, based on the "forward–backward" LP (FB-LP) method (*8, 9*), which improves significantly the quality of the prediction coefficients.

The principles of linear prediction have been described many times before (*1–10*), and only the points salient to the present discussion will be briefly reiterated. In the case of forward linear prediction, a data point x_n is expressed as a linear combination of its K preceding data points:

$$x_n = \sum_{k=1}^{K} c_k x_{n-k}. \qquad [1]$$

If N complex data points have been sampled, $N - K$ equations of type [1] can be generated, and the coefficients c_k can be determined by singular value decomposition (SVD), provided $K \leq N/2$. In the presence of noise, it is also important to ensure that $K > M$, where M is the number of frequency components in the time-domain signal. As M is generally unknown, K is typically set in the $N/4$–$N/3$ range. With the prediction coefficients (c_k) determined by SVD, the frequencies and damping factors of the signal components encoded in these coefficients can be calculated from the K roots ($z_1, \ldots, z_K$) of the polynomial

$$z^K - c_1 z^{K-1} - \cdots - c_K = 0. \qquad [2]$$

0022-2364/92 $5.00

The time-domain signal is then described by

$$x_n = \sum_{k=1}^{K} A_k z_k^{n-1}. \quad [3]$$

Each frequency component in the time series $x_1 \cdots x_N$ corresponds to one of the roots z_k; the amplitude and phase of each signal component is encoded in A_k, which can be determined by a second SVD calculation (*1*). Subsequently, a procedure known as "root reflection" is commonly used, which replaces z_k by $z_k/|z_k|^2$ if $|z_k| > 1$. This ensures that all signal components are decaying exponentials. After root reflection a new set of coefficients (c'_k) is obtained from the polynomial

$$\prod_{k=1}^{K} (z - z'_k), \quad [4]$$

where z'_k denotes the corrected (root-reflected) z_k value.

As was pointed out by Porat and Friedlander (*8*) and first used in NMR by Delsuc *et al.* (*9*), the LP can also be performed in a backward manner:

$$x_n = \sum_{k=1}^{K} d_k x_{n+k}. \quad [5]$$

In the same manner as was used for Eq. [1], the "backward" coefficients d_k can be calculated from [5] by SVD. Rooting of the new polynomial

$$v^K - d_1 v^{K-1} - \cdots - d_K = 0 \quad [6]$$

will then yield signal-related backward roots that after root reflection and complex conjugation[1] are identical (or nearly identical, in the presence of noise) to the signal-related roots of Eq. [2]. The noise-related roots generally do not show such a correlation (*8, 9*). It has been suggested that searching for the pairs of most similar roots, obtained from forward and root-reflected backward linear prediction, provides a robust means for distinguishing signal- and noise-related roots (*9*). The main problem with this procedure relates to the requirement for some prior knowledge about the system (the number of signal components) and to the fact that this search for closest pairs may result in missing roots for weak signal components.

Here we propose an alternative way to use the information contained in the backward linear prediction coefficients. As mentioned above, in the absence of noise, after root reflection and complex conjugation, the roots calculated from Eq. [6] are identical to those from Eq. [2], and the new coefficients recalculated (using Eq. [4]) from the complex conjugates of the root-reflected roots of Eq. [6] are therefore the same as those obtained from solving Eq. [2]. In practice, each set of prediction coefficients contains small random errors caused by noise, however. Simply averaging the two sets of coefficients results in a reduction of the random error present in these two sets of coefficients. The procedure is summarized in the flow diagram in Fig. 1.

[1] Complex conjugation is necessary because the reversal of the time axis in the backward linear prediction results in apparent frequencies that are opposite in sign to those in the forward direction.

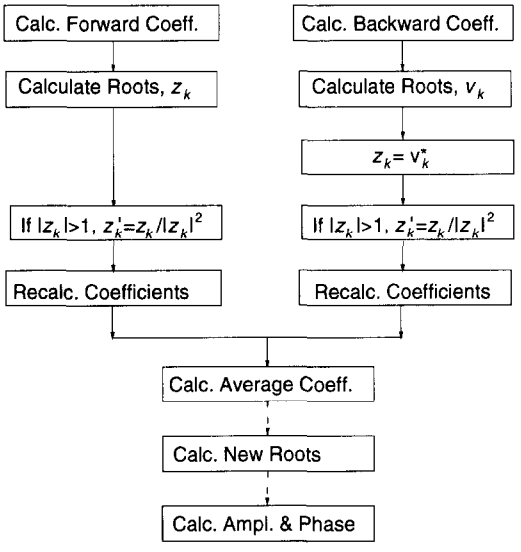

FIG. 1. Flow diagram of the forward–backward linear prediction procedure. The left part of the scheme is used in regular linear prediction; v_k^* is the complex conjugate of v_k. The last two steps, recalculation of the new roots from the averaged coefficients and determination of the amplitude and phase of the frequency components, are omitted when the method is used simply to extend the length of the FID.

Results of the new FB-LP method are illustrated both for simulated data and for a 2D cross section through a 3D data set. Simulated data consist of 16 complex data points, each separated by 1 ms dwell times (spectral width, 1000 Hz), containing three complex sinusoids with frequencies 160, 240, and 480 Hz and relative amplitudes $A_1:A_2:A_3$ = 1:1.5:3. The T_2 value of the simulated sinusoids was 50 ms. Different amounts of Gaussian noise with root-mean-square amplitude σ were added to the signal and the results of the LP calculations are reported for various signal-to-noise (S/N) ratios, defined as

$$S/N = 10 \log(A_1/\sigma). \qquad [7]$$

Table 1 shows the results of both the LP and the FB-LP methods for various values of S/N, using 1000 different sets of computer-generated noise for each value of σ. A score is counted only if all three frequency components have been correctly identified, i.e., if the frequencies derived from the LP for all three components fall within 5 Hz of their true values, the T_2 values are greater than 16 ms, the amplitudes fall within 30% of their true values, and the phase error is less than 30°. As expected (*1*), both the LP and the FB-LP methods perform optimally when the number of coefficients is 25–35% of the total number of time-domain data points. As also can be seen from Table 1, the FB-LP is approximately 1 dB more robust in identifying the correct signals than the regular LP method.

Figure 2 compares the results of the FB-LP and LP methods applied to a cross section through a 3D spectrum in which the frequency of $^{13}C_\delta H_3$ methyl protons of leucine residues in the protein Staphylococcal nuclease (S. nuclease) is correlated with

TABLE 1

Performance of the FB-LP and LP Methods for Correctly Identifying Three Signal Components from 1000 Data Sets Containing Additional Computer-Generated Gaussian Noise, as Described in the Text

		LP order				
S/N (dB)	Method	3	4	5	6	7
14	FB-LP	826	1000	1000	1000	997
14	LP	637	995	973	913	719
13	FB-LP	503	1000	1000	999	992
13	LP	281	993	979	911	744
12	FB-LP	208	997	1000	999	920
12	LP	69	959	975	910	718
11	FB-LP	34	992	994	998	984
11	LP	6	870	971	884	715
10	FB-LP	10	760	973	976	926
10	LP	1	622	930	863	661
9	FB-LP	0	501	910	930	854
9	LP	0	317	839	776	590
8	FB-LP	0	292	786	818	736
8	LP	0	113	670	651	481
7	FB-LP	0	103	626	669	545
7	LP	0	29	449	473	328

$C_\beta H_2$, $C_\gamma H$, and $C_\delta H_3$ proton signals via long-range ^1H–^{13}C J coupling. The 3D spectrum results from a 32 (t_1, ^{13}C) × 32 (t_2, ^1H) × 768 (t_3, ^1H) data matrix, with acquisition times of 26.5 (t_1), 11.9 (t_2), and 53.0 (t_3) ms. After 60°-shifted sine-bell apodization, zero filling, and Fourier transformation in the t_1 and t_3 dimensions, data in the t_2 dimension were extended from 32* to 256*, using 10 prediction coefficients, followed by apodization with a squared cosine bell (t_2) and t_2 Fourier transformation. Extending the data eightfold by linear prediction is more than what we typically use for most 3D data sets, even in cases such as the present one, which has very high S/N. For this high-S/N spectrum, both LP and FB-LP offer a comparable enhancement in resolution over the original spectrum (data not shown). However, as can be seen in Fig. 2A, numerous narrow artifacts appear in the slice processed with conventional LP, whereas in the FB-LP-processed slice such artifacts are much weaker. These artifacts correspond to z values that represent noise and that coincidentally fall very close to the unit circle, i.e., they result in nearly undamped sinusoidal extension of the time-domain data, giving rise to the narrow but sometimes quite intense spikes in the spectrum. Particularly when the LP is applied in two or more orthogonal frequency dimensions, such artifacts significantly decrease the apparent signal-to-noise ratio of the NMR spectrum.

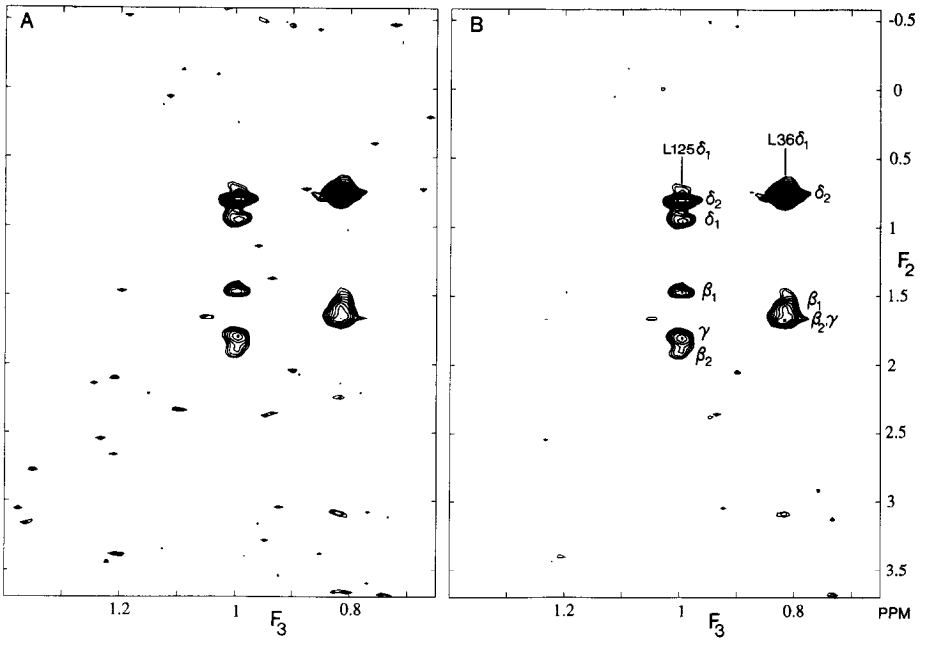

FIG. 2. Comparison of spectra obtained by extending the t_2 time domain eightfold using (A) regular LP and (B) FB-LP. In both cases, 10 coefficients were used and data were extended from 32 complex to 256 complex. Data shown correspond to an F_2/F_3 slice through a 3D spectrum of S. nuclease, enriched with ^{13}C in the C_δ position of Leu residues, showing correlations between methyl protons and other protons that have a significant ($\geq$2 Hz) J coupling to the methyl ^{13}C.

In contrast to the previously described mirror-image LP method (*10*), the FB-LP method does not require any prior knowledge about the NMR signal. It is applicable to both damped and undamped signals of unknown phase, but it cannot be used in addition to mirror-image LP, because backward and forward LP on the mirror-image data results in "backward" coefficients, d_k, that are the exact complex conjugates of the "forward" coefficients, c_k. Hence, the mirror-image LP method already makes implicit use of FB-LP. The "prior information" contained in the additional phase constraint ($\phi = 0$ at $t = 0$) in mirror-image LP makes this latter method superior to FB-LP for undamped signals of known phase. In addition, compared to LP or FB-LP, the mirror-image LP method makes it possible to determine double the number of frequency components for a signal consisting of a limited number of time-domain data points.

ACKNOWLEDGMENTS

We thank G. W. Vuister and S. Grzesiek for stimulating discussions and G. W. Vuister for making available the spectral data used for Fig. 2. G.Z. is supported by a fellowship from the Cooperative Graduate Program in Biophysics, sponsored by the Foundation for Advanced Education in the Sciences and the Graduate School of the University of Maryland at College Park. This work was supported by the Intramural AIDS-Targeted Anti-Viral Program of the Office of the Director of the National Institutes of Health.

REFERENCES

1. H. BARKHUIJSEN, R. DE BEER, W. M. M. J. BOVEE, AND D. VAN ORMONDT, *J. Magn. Reson.* **61,** 465 (1985).
2. H. GESMAR AND J. J. LED, *J. Magn. Reson.* **76,** 183 (1988).
3. J. HOCH, *in* "Methods in Enzymology" (N. Oppenheimer and T. L. James, Eds.), Vol. 176, p. 216, Academic Press, San Diego, 1989.
4. J. TANG AND J. R. NORRIS, *J. Magn. Reson.* **78,** 23 (1988).
5. Y. ZENG, J. TANG, C. A. BUSH, AND J. R. NORRIS, *J. Magn. Reson.* **83,** 473 (1989).
6. E. T. OLEJNICZAK AND H. L. EATON, *J. Magn. Reson.* **87,** 628 (1990).
7. J. J. LED AND H. GESMAR, *Chem. Rev.* **91,** 1413 (1991).
8. B. PORAT AND B. FRIEDLANDER, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-34,** 1336 (1986).
9. M. A. DELSUC, F. NI, AND G. C. LEVY, *J. Magn. Reson.* **73,** 548 (1987).
10. G. ZHU AND A. BAX, *J. Magn. Reson.* **90,** 405 (1990).