Prony Analysis Based Parameter Estimation of an NMR Signal of Blood Plasma for Cancer Detection*

P. K. Gale and J. W. Pierre

Department of Electrical Engineering
University of Wyoming
Laramie, WY 82071

ABSTRACT

Researchers have investigated the correlation between statistical differences in the aliphatic region of NMR processed blood plasma and malignant cancer. Previous efforts used nonlinear curve-fitting algorithms to indirectly estimate the nine components of interest in the methyl and methylene peaks of this region. A different method is described in this paper for directly estimating the nine components. A preprocessing algorithm is used to remove all components outside the aliphatic region. Prony's method and Kumaresan and Tufts' modification are then run, using an order reduction algorithm to identify the compounds of interest. These methods are applied to an NMR plasma sample and the results are compared to the nonlinear curve-fitting algorithm, showing a close relationship.

1. INTRODUCTION

Early detection of cancer is a beneficial factor influencing the prognosis of cancer patients. Medical researchers have recently used water-suppressed proton nuclear magnetic resonance spectroscopy (NMR) to study the aliphatic region of blood plasma. This region is dominated by the lipoprotein methyl and methylene resonances, which is composed of two high density lipids (HDL), two low density lipids (LDL), two very low density lipids (VLDL), two lactate peaks, and a protein resonance. In [1], NMR studies demonstrated differences in relaxation times of the aliphatic resonances if malignant tumors were present. Based on this observation, researchers have attempted to develop a reliable test for the presence of malignant tumors by analyzing the spectra in the aliphatic region.

In [2] and [3] the parameters of interest were indirectly solved for by first performing an FFT on the data obtained from the NMR, then a composite of Gaussian type or Lorentzian type line shapes that best fit the curve in the aliphatic region was obtained using nonlinear curve fitting. In [2], the locations of these resonances were assumed to be fixed, while in [3] they were allowed to vary in location and processed relative to the lactate peaks, but initial estimates were needed. Since the line shape of a damped complex sinusoid is Lorentzian in nature and the parameters were estimated based on a nonlinear curve fitting method, we investigated other methods of parameter estimation that would linearize the problem.

In [4], a method of directly estimating the parameters of exponentially damped sinusoids was described. Kumaresan and Tufts modified Prony's method by using a truncated SVD to increase the effective SNR in the data prior to finding the solution vector. In [5], this method was applied to NMR signals. Due to the fact that the NMR spectrum of blood plasma is composed of many components, the resulting Prony equations are unwieldy if the data is not preprocessed.

This paper focuses on a method of reducing the order of the original estimation while maintaining the spectral content of the aliphatic region. The resulting parameters, which are directly obtained from the NMR data, could then be used as a basis for an early test for cancer.

2. MODEL

The complex time domain output of the NMR spectrometer may be modeled as a sum of damped complex exponentials,

$$x(t) = \sum_{i=1}^{L} A_i \exp(j\omega_i t - t/T_{2i}^*) \qquad (1)$$

where L is the number of terms, A_i is the complex amplitude, ω_i is the frequency, and T_{2i}^* is the apparent spin-spin relaxation time for a particular

* The authors wish to thank the Whedon Foundation for Cancer Research for their generous support.

0-7803-2431-5/95 $4.00 © 1995 IEEE

component in the blood plasma. After sampling the data at a frequency of F_s = 8kHz,

$$x(n) = \sum_{i=1}^{L} A_i \exp[(j\omega_i - 1/T_{2i}^*)n/F_s]. \quad (2)$$

3. PREPROCESSING ALGORITHM

At this point, Prony's method and Kumaresan and Tufts' modification may be directly applied to the data. However, the signal contains many unwanted components and the least squares problem is too large to easily handle. The data was then preprocessed to remove some of the unwanted resonances. The signal was first modulated to shift the aliphatic region from its normal location from 1250 Hz to 1850 Hz down to -300 Hz to 300 Hz.

$$x(n)\exp(j\omega_o nT) = \sum_{i=1}^{L} A_i \exp[(j(\omega_i + \omega_o) - 1/T_{2i}^*)n/F_s] \quad (3)$$

A low pass 41st order Parks-McClellan FIR filter was applied to the modulated data. The filter was designed to have less than 0.4 dB of attenuation up to 250 Hz and 40 dB of attenuation at 500 Hz. It was assumed that the magnitude of the aliphatic parameters were unchanged. The resulting sum is reduced from L to K terms, where the first K terms are in the aliphatic region and (L-K) terms are assumed to be negligible after filtering, which simplifies the Least Squares solution. The signal was demodulated by the negative of the original frequency to produce

$$x(n) = \sum_{i=1}^{K} A_i \exp[(j\omega_i - 1/T_{2i}^*)n/F_s]. \quad (4)$$

Since the frequency spectrum now contains only signals from 1000 to 2000 Hz, the data could be decimated without having components or noise alias into the aliphatic region. The decimated data has the advantage of a larger time aperture for the same record length, resulting in an increase in resolution in the lower frequencies. The original data records are long (16k points), which is cumbersome for the original Prony or the Kumaresan and Tufts modification. The data was decimated by a factor of M=8, which corresponds to a sampling frequency of 1000 Hz, and the data was truncated to a shorter length.

$$x(Mn) = \sum_{i=1}^{K} A_i \exp[(j\omega_i - 1/T_{2i}^*)Mn/F_s] = \sum_{i=1}^{K} A_i \mu_i^n \quad (5)$$

where

$$\mu_i = \exp[(jw_i - 1/T_{2i}^*)M/F_s]. \quad (6)$$

The aliphatic region aliased to 250 Hz to 850 Hz as a result of the decimation. The points corresponding to the transient response of the digital filter were removed from the data set. The Prony method and the truncated SVD part of Kumaresan and Tufts modification of Prony's method was applied to this preprocessed data set to extract the desired parameters.

4. EXPERIMENTAL RESULTS

The following example of analysis of an actual plasma sample is given to see if Prony analysis can resolve the nine components of interest. A sample was treated with EDTA to prevent coagulation and was mixed with an equal quantity of deuterated water. This mixture was then processed using a JEOL 400 MHz spectrometer with water suppression. The sampling rate was 8 kHz and 16384 samples were taken. The entire spectrum of the data is shown in Figure 1.

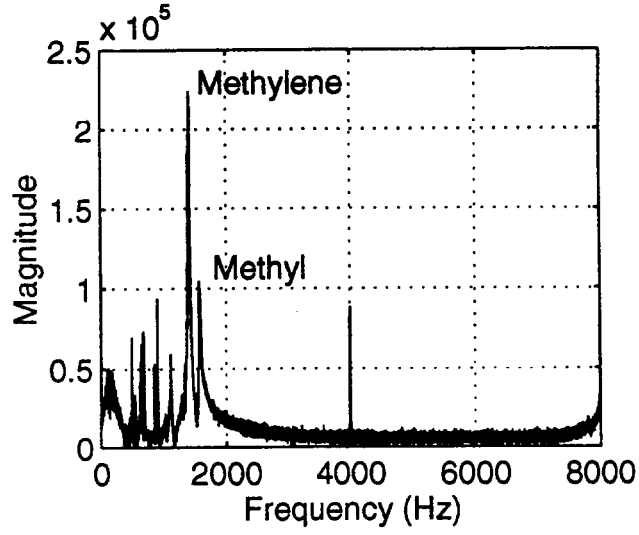

Figure 1 - Entire NMR Spectrum of blood plasma.

In addition to the methyl and methylene resonances many additional resonances are visible in the spectrum. The data was processed using three separate methods: direct Prony analysis, Prony analysis and Kumaresan and Tufts truncated SVD modification with preprocessing, and nonlinear curve-fitting from [3].

4.1 Direct Prony Analysis

The first 1200 data points were used, with an assumed order of 200 to overestimate the signal. Prony's method was directly applied to the data. An order reduction algorithm was applied to the data to identify the compounds in the aliphatic region. This approach failed to resolve seven of the nine components.

4.2 Data Preprocessing

The first 4096 data points in the set were used. Then as described above, the signal is modulated, low pass filtered, demodulated, and decimated by a factor of M=8. The spectrum after this preprocessing is shown in Figure 2.

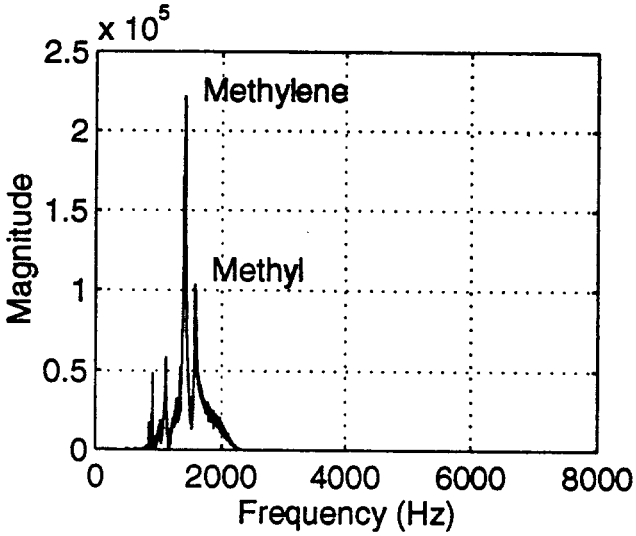

Figure 2 - Spectrum after data preprocessing.

An initial order of 100 was assumed for both Prony and modification.

4.2.1 Prony with Preprocessing

Prony's method was applied to the preprocessed data as described in 4.1, and an order reduction algorithm was applied to identify the nine components of interest.

4.2.2 KT Modification with Preprocessing

The parameters of the nine components were estimated using the Kumaresan and Tufts truncated SVD method. The largest 15 singular values were assumed to be significant based on observation. These largest singular values are due to the nine components of interest, components not completely filtered, and weak resonant low molecular weight components between the methyl and methylene resonances.

4.3 Nonlinear Curve-Fitting

An FFT was applied to the NMR data. The spectrum was converted to magnitude and frequency and the nonlinear curve fitting approach [3] was applied.

Figure 3 shows a polar plot of the estimated values of μ using the three methods. The estimates of f, T_2^* and A are shown in Table 1. The resonance of water occurs at 4.65 ppm, or 1860 Hz when a 400 Mhz NMR is used. All frequencies have been corrected for the water resonance, since the spectrum is shifted relative to this external reference.

5. CONCLUSIONS

As can be seen from the results, the estimates of the μ's are fairly consistent for the three approaches. Additionally, both Prony based algorithms had similar results for all frequencies and T_2^* values, while all three methods could estimate the lactate peaks consistently. The key result is that with the proposed preprocessing, the Prony method and the Kumaresan and Tufts SVD modification can resolve the nine components of interest without any initial parameter estimation or nonlinear curve-fitting.

The next steps will be to experiment with using different sets of decimated data from the same NMR plasma sample and tuning the preprocessing algorithm. Ultimately, these parameters will be used to try to predict the presence of colon cancer. This paper invites other researchers in the signal processing community to investigate the NMR parameter estimation problem for use in early cancer detection.

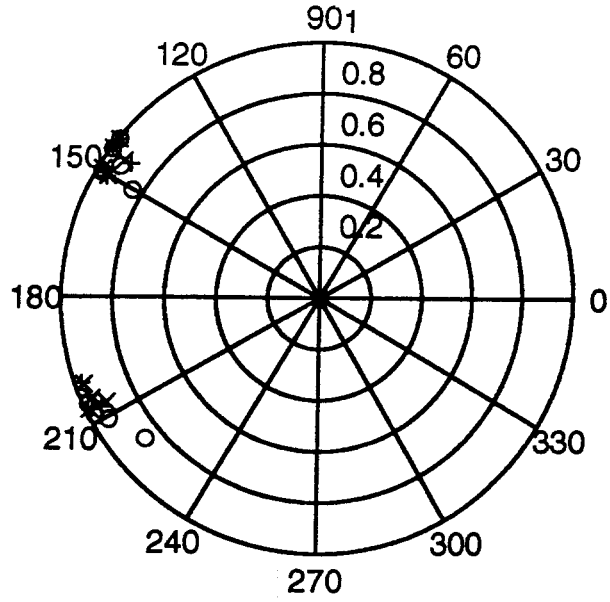

Figure 3 - Polar plot of estimated μ's: x - Prony's method, + - Kumaresan and Tufts method, and o - nonlinear curve fitting of the FFT method.

Table 1

	f (Hz)			T2* (sec)			A		
Methyl	Prony	KT	Curve-fit	Prony	KT	Curve-fit	Prony	KT	Curve-fit
HDL	285.92	284.94	249.55	0.011	0.016	0.007	0.945	1.000	0.393
LDL	286.19	285.84	275.00	0.059	0.040	0.015	0.033	0.136	0.191
VLDL	292.46	294.29	281.52	0.024	0.031	0.039	0.049	0.033	0.043
Protein	302.54	303.81	289.76	0.029	0.036	0.051	0.077	0.058	0.035
Methylene									
HDL	443.03	442.16	442.94	0.019	0.020	0.006	1.000	0.872	1.000
LDL	445.32	444.98	443.31	0.057	0.036	0.043	0.019	0.118	0.105
VLDL	457.23	457.88	452.66	0.013	0.010	0.014	0.449	0.947	0.423
Lactate 1	459.60	459.61	459.12	0.170	0.160	0.145	0.040	0.056	0.010
Lactate 2	466.21	466.24	466.45	0.213	0.214	0.190	0.038	0.049	0.007

References

[1] Fossel, E.T., Carr, J.M., and McDonagh, J., New England Journal of Medicine (1986): Vol. 315 (22), 1369-1376.

[2] Song, I., Kang, S., Kim, N., Im, J., and Min, B., Seoul Journal of Medicine (1989): 30(1), 31-36.

[3] Pappu, R., Johnson, T.W., Asplund, R.O., and Pierre, J.W., "Spectral estimation of Human Blood for Use in Cancer Detection," Proc. of the 36th Midwest Symposium on Circuits and Systems, August, 1993.

[4] Kumaresan, R. and Tufts, D.W., "Estimating the Parameters of Exponentially Damped Sinusoids and Pole Zero Modeling in Noise," IEEE Trans. on Acoustics, Speech, and Signal Processing, vol. 30, no. 6, pg. 833-840, December 1982.

[5] Barkhuijsen, H. De Beer, R., Bovee, W.M.M.J., and Van Ormondt, D., "Retrieval of Frequencies, Amplitudes, Damping Factors, and Phases from Time-Domain Signals Using a Linear Least-Squares Procedure," Journal of Magnetic Resonance, pp. 465-481, 1985.