APPENDIX CRAMER-RAO BOUNDS

The Cramer-Rao lower bounds for the parameters b_i and a_i are provided below. We assume that

$$p(z; a, b) = \frac{1}{(2\pi\sigma^2)^{n+1/2}} \cdot \exp \left\{ -\frac{1}{2\sigma^2} [z - p(b, a)]^T [z - p(b, a)] \right\}.$$

Then

$$I_{xy} = -E \left\{ \frac{\partial^2 \log p(z; a, b)}{\partial x \partial y} \right\}$$

where x and y are one of $\{a_1, \dots, a_p, b_0, \dots, b_q\}$. Then it may be shown that

$$I_{b_i b_m} = \frac{1}{\sigma^2} \sum_{i=0}^n \frac{4 \operatorname{Re} \{e^{j k \omega_i} B(e^{-j \omega_i})\} \operatorname{Re} \{e^{j m \omega_i} B(e^{-j \omega_i})\}}{|A(e^{j \omega_i})|^4}$$

$$I_{a_i b_m} = \frac{1}{\sigma^2} \sum_{i=0}^n \frac{4 |B(e^{j \omega_i})|^2 \operatorname{Re} \{e^{j k \omega_i} A(e^{-j \omega_i})\} \operatorname{Re} \{e^{j m \omega_i} B(e^{-j \omega_i})\}}{|A(e^{j \omega_i})|^6}$$

$$I_{a_i a_m} = \frac{1}{\sigma^2} \sum_{i=0}^n \frac{4 |B(e^{j \omega_i})|^4 \operatorname{Re} \{e^{j k \omega_i} A(e^{-j \omega_i})\} \operatorname{Re} \{e^{j m \omega_i} A(e^{-j \omega_i})\}}{|A(e^{j \omega_i})|^8}.$$

The matrix $[I_{xy}]$ is then inverted in order to find the covariance bound for the maximum likelihood estimates.

REFERENCES

- [1] J. Brittingham *et al.*, "Pole extraction from real-frequency information," *Proc. IEEE*, vol. 68, no. 2, Feb. 1980.
- [2] W. Cheney and D. Kincaid, *Numerical Mathematics and Computing*. Monterey, CA: Brooks/Cole, 1985, p. 470.
- [3] R. Kumaresan, "On a frequency domain analog of Prony's method," *IEEE Trans. Acoust., Speech, Signal Processing*, pp. 168-170, Jan. 1990.
- [4] R. Kumaresan, "Identification of rational transfer functions from frequency response samples," *IEEE Trans. Aerosp. Electron. Syst.*, pp. 925-934, Nov. 1990.
- [5] G. H. Golub and V. Pereyra, *SIAM J. Numer. Anal.*, pp. 413-432, Apr. 1973.
- [6] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, 1986.
- [7] L. R. Rajagopal and S. C. Dutta Roy, "A matrix approach for the coefficients of maximally flat FIR filter transfer functions," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 12, Dec. 1988.

A Newton-Based Ziskind-Wax Alternating Projection Algorithm Applied to Spectral Estimation

L. S. Biradar and V. U. Reddy

Abstract—In this correspondence, we present a Newton-based alternating projection (AP) algorithm for estimating the parameters of exponential signals in noise, and compare its performance with that of

modified forward-backward linear prediction/backward linear prediction (FBLP/BLP), total least squares (TLS), and iterative quadratic maximum likelihood (IQML) methods using computer simulations. In the case of undamped sinusoids, the AP algorithm yielded lower threshold SNR than the others, while in the case of damped sinusoids, its threshold SNR is same as that of TLS, but is 1 dB lower than that of MBLP and IQML.

I. INTRODUCTION

The problem of estimating the parameters of superimposed damped/undamped sinusoids in the presence of white noise arises in many practical applications. Further, in practice, we have to attempt this problem with a small number of data samples and possibly at low signal-to-noise ratio (SNR). Several methods based on linear and nonlinear least squares formulations have been recently proposed for the above problem. Notable among these are modified forward-backward/backward linear prediction (FBLP/MBLP) methods [1], [2], total least squares (TLS) method [3], iterative quadratic maximum likelihood (IQML) algorithm [4], and alternating projection (AP) algorithm [5]. Here, we present a Newton-based AP algorithm and compare its performance with that of [1]-[4].

The observed finite data, composed of uniformly spaced samples, satisfies a signal model

$$y(n) = \sum_{i=1}^M s_i z_i^n + v(n), \quad n = 0, 1, \dots, N-1 \quad (1)$$

with $z_i = e^{(\alpha_i + j\omega_i)}$, where M denotes the number of signals (damped if $\alpha_i < 0$ and undamped if $\alpha_i = 0$), s_i , α_i , and ω_i are the complex amplitude, damping factor, and frequency (normalized) of the i th signal, respectively. $\{v(n)\}$ is a zero mean, complex valued white noise process with independent real and imaginary parts, each having a variance $\sigma^2/2$.

The data given by (1) can be represented in a vector form as

$$y = A(z)s + v \quad (2)$$

where $y = [y(0), y(1), \dots, y(N-1)]^T$, $A(z) = [a(z_1), a(z_2), \dots, a(z_M)]$ with $a(z_i) = [1, z_i, z_i^2, \dots, z_i^{N-1}]^T$, $s = [s_1, s_2, \dots, s_M]^T$ and v is a vector of N noise samples. The nonlinear least squares estimation of s and z is formulated as

$$\min_{s, z} \|y - A(z)s\|^2 \quad (3)$$

with the constraint $|z_i| \leq 1$ for all i . Here, $\|\cdot\|$ denotes 2-norm. The AP algorithm of Ziskind and Wax [5] for solving (3) is as follows.

Denote

$$\hat{z}_i^{(0)} = [\hat{z}_1^{(0)} \hat{z}_2^{(0)} \dots \hat{z}_{i-1}^{(0)}]^T, \quad i = 2, 3, \dots, M.$$

Initialization:

$$\hat{z}_1^{(0)} = \arg \max_{z_1} \frac{a^+(z_1) \hat{R} a(z_1)}{a^+(z_1) a(z_1)}$$

for $i = 2$ to M

$$\hat{z}_i^{(0)} = \arg \max_{z_i} \frac{a^+(z_i) (I - P_{A(\hat{z}_i^{(0)})}) \hat{R} (I - P_{A(\hat{z}_i^{(0)})}) a(z_i)}{a^+(z_i) (I - P_{A(\hat{z}_i^{(0)})}) a(z_i)}$$

end.

Main Loop:

while

$$\|\hat{z}^{(k+1)} - \hat{z}^{(k)}\|_\infty > \epsilon (\|x\|_\infty = \max_i |x_i|)$$

Manuscript received June 9, 1991; revised May 31, 1992.
The authors are with the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore, 560 012, India.
IEEE Log Number 9206014.

for $i = 1$ to M

$$\hat{z}^{(k+1)} = \arg \max_{z_i} \frac{a^+(z_i) Q_1 a(z_i)}{a^+(z_i) Q_2 a(z_i)} \quad (4)$$

end.

Once we have the estimates $\{\hat{z}_i\}_{i=1}^M$, the damping factors, angular frequencies, and complex amplitudes can be obtained from

$$\alpha_i = \ln |z_i|, \quad \omega_i = \text{angle}(z_i), \quad \hat{s} = (A^+(z)A(z))^{-1} A^+(z)y.$$

II. NEWTON-RAPHSON SEARCH ALGORITHM

To speed up the convergence of the AP algorithm, we develop Newton-Raphson search technique for performing the maximization in (4). Consider the objective function

$$F(z_i) = \frac{a^+(z_i) Q_1 a(z_i)}{a^+(z_i) Q_2 a(z_i)} \quad (5)$$

where

$$Q_1 = (I - P_{A(z_i^{(k)})}) \hat{R} (I - P_{A(z_i^{(k)})}), \quad Q_2 = I - P_{A(z_i^{(k)})} \quad (6)$$

$$\hat{z}_i^{(k)} = [\hat{z}_1^{(k)}, \hat{z}_2^{(k)}, \dots, \hat{z}_{i-1}^{(k)}, \hat{z}_{i+1}^{(k)}, \dots, \hat{z}_M^{(k)}]^T \quad (7)$$

and $P_{A(z)}$ denotes the projection matrix which projects a vector onto the column space of $A(z)$.

A. Undamped Case

With $z_i^* = z_i^{-1}$, $F(z_i)$ can be expressed as

$$F(z_i) = \frac{\sum_{r=0}^{2N-2} b_r z_i^r}{\sum_{r=0}^{2N-2} c_r z_i^r} \quad (8)$$

where

$$b_r = \begin{cases} \sum_{i=1}^{r+1} Q_1(N-r-1+i, i), & 0 \leq r \leq N-1 \\ \sum_{i=1}^{2N-r-1} Q_1(i, r-3+i), & N \leq r \leq 2N-2 \end{cases} \quad (9)$$

and

$$c_r = \begin{cases} \sum_{i=1}^{r+1} Q_2(N-r-1+i, i), & 0 \leq r \leq N-1 \\ \sum_{i=1}^{2N-r-1} Q_2(i, r-3+i), & N \leq r \leq 2N-2 \end{cases} \quad (10)$$

with $Q_1(m, n)$ and $Q_2(m, n)$ denoting the (m, n) th elements of Q_1 and Q_2 , respectively. The Newton-Raphson iteration for seeking the maximum of (4) is then given by

$$\hat{z}_i^{(k+1)}(l+1) = \hat{z}_i^{(k+1)}(l) - \frac{\sum_{r=0}^{4N-6} d_r z_i^r \sum_{r=0}^{4N-4} e_r z_i^r}{\sum_{r=0}^{8N-12} f_r z_i^r} \bigg|_{z_i = \hat{z}_i^{(k+1)}(l)} \quad (11)$$

where

$$\begin{aligned} d_r &= \begin{cases} \sum_{i=0}^r (r+1-i) \{c_i b_{r+1-i} - b_i c_{r+1-i}\}, & 0 \leq r \leq 2N-3 \\ \sum_{i=0}^{2N-3} (i+1) \{c_{r-i} b_{i+1} - b_{r-i} c_{i+1}\}, & r = 2N-2 \end{cases} \\ &= \sum_{i=0}^{4N-5-r} \{r+1-(2N-2)+i\} [c_{(2N-2)-i} b_{r+1-(2N-2)+i} - b_{(2N-2)-i} c_{r+1-(2N-2)+i}], & 2N-1 \leq r \leq 4N-6 \end{aligned} \quad (12)$$

$$e_r = \begin{cases} \sum_{i=0}^r c_i c_{r-i}, & 0 \leq r \leq 2N-2 \\ \sum_{i=0}^{4N-4-r} c_{(2N-2)-i} c_{r-(2N-2)+i}, & 2N-1 \leq r \leq 4N-4 \end{cases} \quad (13)$$

and

$$f_r = \begin{cases} \sum_{i=0}^r (r+1-i) \{e_i d_{r+1-i} - d_i e_{r+1-i}\}, & 0 \leq r \leq 4N-7 \\ \sum_{i=0}^{4N-7} (i+1) e_{r-i} d_{i+1} - \sum_{i=0}^r (r+1-i) d_i e_{r+1-i}, & r = 4N-6 \\ \sum_{i=0}^{4N-7} (i+1) e_{r-i} d_{i+1} - \sum_{i=0}^{4N-6} (r+1-i) d_i e_{r+1-i}, & r = 4N-5 \\ \sum_{i=0}^{4N-7} (i+1) e_{r-i} d_{i+1} & \\ - \sum_{i=0}^{8N-11-r} \{r+1-(4N-6)+i\} d_{(4N-6)-i} e_{r+1-(4N-6)+i}, & r = 4N-4 \\ \sum_{i=0}^{8N-11-r} \{r+1-(4N-4)+i\} e_{(4N-4)-i} d_{r+1-(4N-4)+i} & \\ - \sum_{i=0}^{8N-11-r} \{r+1-(4N-6)+i\} d_{(4N-6)-i} e_{r+1-(4N-6)+i}, & 4N-3 \leq r \leq 8N-12. \end{cases} \quad (14)$$

In each iteration, $|\hat{z}_i^{(k+1)}(l)|$ is forced to unity keeping the angle same.

B. Damped Case

In this case we treat $F(\cdot)$ as a function of two variables, z_i and z_i^* , and express the Newton-Raphson iteration in two variables as

$$\begin{bmatrix} \hat{z}_i^{(k+1)}(l+1) \\ \hat{z}_i^{*(k+1)}(l+1) \end{bmatrix} = \begin{bmatrix} \hat{z}_i^{(k+1)}(l) \\ \hat{z}_i^{*(k+1)}(l) \end{bmatrix} - \mathbf{H}^{-1} \mathbf{h} \quad (15)$$

where $\mathbf{h}$ and $\mathbf{H}$ are the gradient vector and Hessian matrix of $F(z_i, z_i^*)$. The Newton-Raphson update for $\hat{z}_i^{(k+1)}(l+1)$ can then be obtained as

$$\hat{z}_i^{(k+1)}(l+1) = \hat{z}_i^{(k+1)}(l) - \frac{\beta_2[\{\beta_2(\beta_2\beta_5^* - \beta_1\beta_6^*) - 2\beta_4^*(\beta_2\beta_3^* - \beta_1\beta_4^*)\}(\beta_2\beta_3 - \beta_1\beta_4) - \{\beta_2(\beta_2\beta_7 - \beta_4^*\beta_3 - \beta_1\beta_8 - \beta_3^*\beta_4) + 2\beta_1\beta_4\beta_4^*\}(\beta_2\beta_3^* - \beta_1\beta_4^*)]}{\{[abs\{\beta_2(\beta_2\beta_5 - \beta_1\beta_6) - 2\beta_4(\beta_2\beta_3 - \beta_1\beta_4)\}]^2 - [\beta_2(\beta_2\beta_7 - \beta_4^*\beta_3 - \beta_1\beta_8 - \beta_3^*\beta_4) + 2\beta_1\beta_4\beta_4^*]^2\}} \quad (16)$$

where

$$\begin{aligned} \beta_1 &= \mathbf{a}^+(z_i) \mathbf{Q}_1 \mathbf{a}(z_i) & \beta_2 &= \mathbf{a}^+(z_i) \mathbf{Q}_2 \mathbf{a}(z_i) \\ \beta_3 &= \mathbf{a}^+(z_i) \mathbf{Q}_1 \frac{\partial \mathbf{a}(z_i)}{\partial z_i} & \beta_4 &= \mathbf{a}^+(z_i) \mathbf{Q}_2 \frac{\partial \mathbf{a}(z_i)}{\partial z_i} \\ \beta_5 &= \mathbf{a}^+(z_i) \mathbf{Q}_1 \frac{\partial^2 \mathbf{a}(z_i)}{\partial z_i^2} & \beta_6 &= \mathbf{a}^+(z_i) \mathbf{Q}_2 \frac{\partial^2 \mathbf{a}(z_i)}{\partial z_i^2} \\ \beta_7 &= \frac{\partial \mathbf{a}^+(z_i)}{\partial z_i^*} \mathbf{Q}_1 \frac{\partial \mathbf{a}(z_i)}{\partial z_i} & \beta_8 &= \frac{\partial \mathbf{a}^+(z_i)}{\partial z_i^*} \mathbf{Q}_2 \frac{\partial \mathbf{a}(z_i)}{\partial z_i} \end{aligned} \quad (17)$$

with

$$\begin{aligned} \frac{\partial \mathbf{a}(z_i)}{\partial z_i} &= [0, 1, 2z_i, 3z_i^2, \dots, (N-1)z_i^{(N-2)}]^T \\ \frac{\partial^2 \mathbf{a}(z_i)}{\partial z_i^2} &= [0, 0, 2 \cdot 1, 3 \cdot 2z_i, 4 \cdot 3z_i^2, \dots, (N-1)(N-2)z_i^{(N-3)}]^T \\ \frac{\partial \mathbf{a}^+(z_i)}{\partial z_i^*} &= [0, 1, 2z_i^*, 3z_i^{*2}, \dots, (N-1)z_i^{*(N-2)}] \\ &= \left(\frac{\partial \mathbf{a}(z_i)}{\partial z_i} \right)^+ \end{aligned} \quad (18)$$

If $|\hat{z}_i^{(k+1)}(l+1)| > 1$, then we replace it by $1/(\hat{z}_i^{*(k+1)}(l+1))$ and start the next iteration.

III. SIMULATION RESULTS

We used computer simulations to compare the performance of the AP method with that of earlier methods [1]–[4].

Example 1: Undamped case.

The data samples were generated from

$$y(n) = e^{j(2\pi f_1 n + \phi_1)} + e^{j(2\pi f_2 n + \phi_2)} + v(n),$$

$$n = 0, 1, \dots, 24$$

with $f_1 = 0.5$, $f_2 = 0.52$ (normalized frequencies), $\phi_1 = 0$ and $\phi_2 = \pi/4$. $\{v(n)\}$ was a sequence of independent and identically distributed zero mean complex Gaussian random variables with uncorrelated real and imaginary parts, each with a variance of $\sigma^2/2$. The variance of the noise, σ^2 , was chosen to give the desired signal-to-noise ratio (SNR), defined as $10 \log_{10} (1/\sigma^2)$. The AP method was initialized by searching over a linear grid with a res-

olution of 0.01 over the range of normalized frequency values in (0, 1). Subsequently, the AP iterations made use of the Newton-Raphson update given by (11). From the converged values of $\hat{z}_i$, we estimated the frequencies.

The mean square error (MSE) in the frequency estimates was evaluated from 50 Monte Carlo runs. This was repeated for different values of SNR and the results are plotted in Fig. 1. In the same figure, we also overlaid the MSE values obtained with the IQML method [4], and the MFBLP and TLS methods for a predictor order 18. We should point out here that we simulated all the algorithms using the same 50 data realizations. The MSE values corresponding to the Cramér-Rao (CR) bound were evaluated and overlaid in the same figure.

The plots in Fig. 1 show that the AP algorithm yields the lowest threshold SNR; 3 dB less than that of IQML and 6 dB less than that of MFBLP and TLS in the case of the estimate of $f_1(0.5)$, while the corresponding values are 1 and 7 dB in the case of the other frequency. Away from the threshold SNR, the performance of the AP and IQML methods is nearly same and is close to the CR bound, while that of MFBLP and TLS methods is poorer by 1 to 2 dB. Note that the TLS and MFBLP perform similarly, which is consistent with the results given in [3]. The MSE values lower than the CR bound, for the case of AP algorithm in the neighbourhood of the threshold SNR, may have resulted because of the small number of Monte Carlo experiments.

Example 2: Damped case.

The data samples were generated from

$$y(n) = e^{(\alpha_1 + j2\pi f_1)n} + e^{(\alpha_2 + j2\pi f_2)n} + v(n),$$

$$n = 0, 1, \dots, 24$$

with normalized frequencies $f_1 = 0.42$, $f_2 = 0.52$, and damping factors $\alpha_1 = -0.2$, $\alpha_2 = -0.1$. The model for $\{v(n)\}$ was the same as that in Example 1. The variance of the noise σ^2 was chosen to give the desired SNR, defined as $10 \log_{10} (1/\sigma^2)$. Note that the SNR here refers to that of zeroth sample and not to the average

Fig. 1. MSE in the frequency estimates for the undamped case. (a) Corresponds to $f_1 = 0.5$. (b) Corresponds to $f_2 = 0.52$.

Fig. 2. MSE in the frequency and damping factor estimates for the damped case. (a) Corresponds to $f_1 = 0.42$, $\alpha_1 = -0.2$. (b) Corresponds to $f_2 = 0.52$, $\alpha_2 = -0.1$.

over the entire data record. The AP algorithm was initialized by searching over a two-dimensional grid; with a resolution of 0.001 over the range of normalized frequency values in (0.38, 0.56), and 0.0005 over the range of $e^{-\alpha}$ in (0.75, 0.95). Note that the reason for searching over a truncated (but wide enough) two-dimensional grid was to reduce the computation time required to get the initial guess. Subsequently, the AP iterations made use of the Newton-Raphson update given by (16). From the converged values of $\hat{z}_i$, we estimated the frequencies and the damping factors.

The values of MSE in the frequency and damping factor estimates, evaluated from 50 Monte Carlo runs, are plotted in Fig. 2. In the same figure, we also plotted the respective MSE values obtained with IQML and MBLP and TLS methods for a predictor order 18. As in Example 1, we simulated all the algorithms using the same data realizations. The corresponding MSE values of CR bound are overlaid in the same figure. We may point out here that the TLS method was applied to the set of backward linear prediction equations.

We note from Fig. 2 that for both the frequency estimates, the threshold SNR of AP method is same as that of TLS but is 1 dB lower than that of IQML and MBLP. Away from the threshold SNR, the estimation accuracy of all the methods is nearly same and is close to CR bound. In the case of the estimate corresponding to lower damping factor, the performance of the AP and IQML methods is better than that of MBLP and TLS, while the TLS performed better than the others in the case of the higher damping factor.

IV. CONCLUSIONS

In this correspondence, we presented a Newton-based AP algorithm of Ziskind and Wax [5] and compared its performance with that of MFBP (or MBLP), TLS, and IQML methods using computer simulations.

ACKNOWLEDGMENT

The authors thank G. Sharma for providing the simulation results for the IQML case.

REFERENCES

- [1] D. W. Tufts and R. Kumaresan, "Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood," *Proc. IEEE*, vol. 70, pp. 975-989, Sept. 1982.
- [2] R. Kumaresan and D. W. Tufts, "Estimating the parameters of exponentially damped sinusoids and pole-zero modelling in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 30, pp. 833-840, Dec. 1982.
- [3] Md. A. Rahman and K. Yu, "Total least squares approach for frequency estimation using linear prediction," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 35, no. 10, pp. 1440-1454, Oct. 1987.
- [4] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, no. 5, pp. 1081-1089, Oct. 1986.
- [5] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 10, pp. 1553-1560, Oct. 1988.