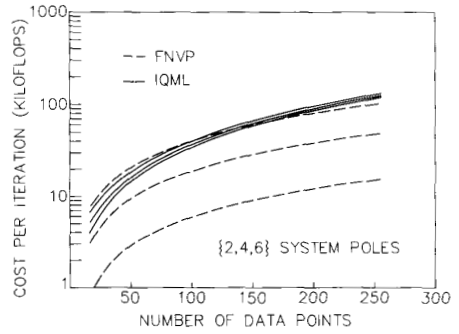

Fig. 1. Computational cost flops/iteration for stepped sine response modeling for transfer functions with 2, 4, and 6 system poles.

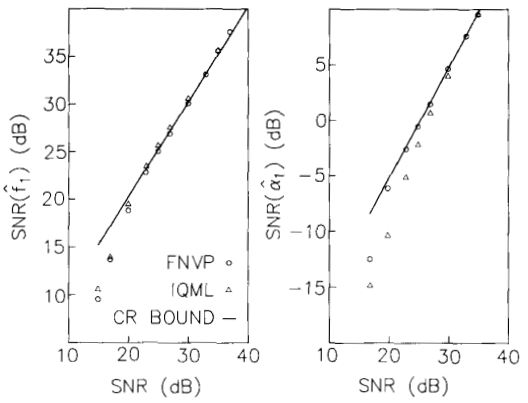

Fig. 2. Pole parameter estimate SNR $(\hat{\theta}) = 10 \log_{10}[\theta^2/\mathrm{MSE}(\hat{\theta})]$ versus the observed data SNR in decibels. Five hundred Monte Carlo trials. $N = 16$. Stepped sine response of a two-pole system with ideal parameters: $f_c = 1, f_c = 1, f_1 = 0.99804, \alpha_1 = -0.3927, Q = 8$, and $NTf_0 = 2$.

we introduce the relative mean squared error (MSE) of the estimates as an index of performance for the FNVP and IQML estimation strategies, and express it as a signal-to-error ratio or a signal-to-noise ratio in decibels:

$$\mathrm{SNR}(\hat{\theta}) = 10 \log_{10}[\theta^2/\mathrm{MSE}(\hat{\theta})]. \tag{5.2}$$

The pole parameter estimate signal-to-noise ratios, SNR $(\hat{f}_1)$ and SNR $(\hat{\alpha}_1)$, for the estimated resonant pole in our response signal, are compared to the CR bound in Fig. 2. The performance of FNVP and IQML relative to the CR bound are similar, but with FNVP continuing ML performance for damping estimates to lower SNR than IQML.

VI. Conclusions

We have extended the work of Golub and Pereyra [1] by deriving the Hessian matrix of the VPF, and implemented algorithm FNVP for estimating parameters in the semilinear model. For high SNR problems, the algorithm uses exact Newton steps only, while for data with lower SNR, the gradient and Hessian were integrated into the adaptive Newton algorithm HUMSL [13]. Simulation studies of estimating the parameters of a stepped sinusoid response of a resonant system were used to compare the estimation and cost performance of the FNVP and IQML algorithms. The computational cost of FNVP was shown to vary linearly with data length N, while the cost of IQML varied as $N \log_2 N$. In addition to the full Newton algorithm, we proposed and tested a deflation technique for introducing equality constraints on signal basis vectors. While the focus here has been on single snapshots of real data, the algorithms can be easily extended to multiple snapshots (e.g., see [4]) and complex data.

References

[1] G. H. Golub and V. Pereyra, "The differentiation of pseudoinverses and nonlinear least squares problems whose variables separate," *SIAM J. Numer. Anal.*, vol. 10, pp. 413–432, 1973.
[2] T. A. Parks, "Reducible nonlinear programming problems," Ph.D. dissertation, Rice University, Houston, TX, May 1985.
[3] J. D. George and P. L. Ainsleigh, "A constrained Newton algorithm for maximum likelihood modeling of sums of exponentials in noise," in *Proc. IEEE ICASSP 90*, Albuquerque, NM, vol. 5, pp. 2443–2446, Apr. 1990.
[4] Y. Bressler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, no. 5, pp. 637–640, Oct. 1986.
[5] R. Kumaresan, L. L. Scharf, and A. K. Shaw, "An algorithm for pole-zero modeling and spectral analysis," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, no. 3, pp. 637–640, June 1986.
[6] D. Starer and A. Nehorai, "Maximum likelihood estimation of exponential signals in noise using a Newton algorithm," in *Proc. 4th ASSP Workshop Spectrum Estimation Modeling*, Minneapolis, MN, Aug. 1988, pp. 240–245.
[7] P. L. Ainsleigh, "Maximum likelihood estimation for acoustic transducer calibration," M.S.E.E. thesis, Univ. South Florida, Tampa, FL, Dec. 1988.
[8] P. L. Ainsleigh and J. D. George, "Variable projection methods with application to exponential signals in white noise," NRL Memo. Rep. 6643, Naval Res. Lab., Orlando, FL, Dec. 1990.
[9] M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting," *IEEE Signal Processing*, vol. 39, no. 5, May 1991.
[10] L. Kaufman, "A variable projection method for solving separable nonlinear least squares problems," *BIT 15*, 1975.
[11] J. D. George, V. K. Jain, and P. L. Ainsleigh, "Estimating steady-state response of a resonant transducer in a reverberant environment," in *Proc. IEEE ICASSP 88*, New York, NY, vol. 5, Apr. 1988, pp. 2737–2740.
[12] J. D. George and D. M. Goodman, "Application of the output error system identification method to the calibration of underwater acoustic transducers," *Underwater Acoustic Processing*, Y. T. Chan, Ed. Kluwer, July 1988, pp. 39–45.
[13] D. M. Gay, "ALGORITHM 611—Subroutines for unconstrained minimization using a model/trust region approach," *ACM Trans. Math. Software*, vol. 9, no. 4, pp. 503–524, Dec. 1983.
[14] R. Kumar, "A fast algorithm for solving a Toeplitz system of equations," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, no. 1, pp. 254–267, Feb. 1985.

Multivariate ARMA Modeling by Scalar Algorithms

Mrityunjoy Chakraborty and Surendra Prasad

***Abstract*—In this correspondence, we have proposed an algorithm for multichannel autoregressive moving average (ARMA) modeling which uses scalar computations only and is well suited for parallel implemen-**

Manuscript received September 18, 1990; revised March 18, 1992.
The authors are with the Department of Electrical Engineering, Indian Institute of Technology, Delhi, New Delhi-110016, India.
IEEE Log Number 9206868.

tation. The given ARMA process is converted to an equivalent scalar, periodic ARMA process. The scalar autoregressive (AR) parameters are estimated by first deriving a set of modified Yule-Walker type equations and then solving them by a parallel, order recursive algorithm. The moving average (MA) parameters are estimated by a least squares method from the estimates of the input samples obtained via a high order, periodic AR approximation of the scalar process.

I. Introduction

In recent years, multichannel time series modeling has attracted a great deal of interest, especially in application areas like multisensor signal processing, identification of multivariate systems, parallel image processing, and multichannel power spectrum estimation. The direct extension of the single channel (i.e., scalar) algorithms to the multichannel case, however, involves extensive matrix operations and thus becomes computationally unattractive. As a result, various efforts ([1]–[5]) have been made over the past few years to derive computationally superior algorithms using scalar arithmetic only.

In this correspondence, we propose an algorithm to identify a multichannel autoregressive moving average (ARMA) model which employs scalar computations only and is well suited for implementation on a pipelined processor. Following a theorem by Pagano [6], the given d-channel ARMA process is converted to an equivalent periodically stationary, scalar ARMA process with the period equal to d.[1] The proposed algorithm estimates the ARMA parameters of the scalar process which are related to their multichannel counterparts by very simple relations. It is possible to establish the causality and causal invertibility of the scalar model in the light of innovation representation. The causality is required to derive the modified Yule-Walker (MYW) type equations for estimating the autoregressive (AR) parameters while the causal invertibility is used in estimating the input samples via a high order, periodic AR approximation of the scalar process. The MYW equations are solved by an order recursive algorithm which can be implemented on a set of d processors pipelined in a *circular manner*. The MA parameters are estimated from the estimates of the input samples by a least squares method.

II. Scalar Representation of Multivariate ARMA Processes

A. Relation Between a Multivariate and a Scalar, Periodic ARMA Process

Consider a d-variate, complex-valued, ARMA(p,q) process given by

$$\boldsymbol{x}(k) + \sum_{i=1}^{p} \boldsymbol{A}(i)\boldsymbol{x}(k-i) = \boldsymbol{w}(k) + \sum_{j=1}^{q} \boldsymbol{B}(j)\boldsymbol{w}(k-j), \quad (1)$$

where $\boldsymbol{A}(i)$'s and $\boldsymbol{B}(j)$'s are $d \times d$ matrices and $\boldsymbol{w}(k)$ is a zero-mean, d-variate input process with covariance matrix $\boldsymbol{Q}$, i.e., $E[\boldsymbol{w}(k+l)\boldsymbol{w}^{\dagger}(k)] = \boldsymbol{Q}\,\delta(l)$, where † denotes Hermitian transposition, E denotes expectation, and $\delta(l) = 1$ if $l = 0$ and zero otherwise. Without loss of generality, $\boldsymbol{Q}$ can be assumed to be positive definite. We then have the following Cholesky decomposition: $\boldsymbol{Q} = \boldsymbol{L}\boldsymbol{D}\boldsymbol{L}^{\dagger}$, where $\boldsymbol{L}$ is a unit lower triangular matrix and $\boldsymbol{D}$ is a diagonal matrix with real, positive diagonal entries. Introducing a new input $\boldsymbol{u}(k) = \boldsymbol{L}^{-1}\boldsymbol{w}(k)$ and substituting in (1), we obtain the following:

$$\boldsymbol{L}^{-1}\boldsymbol{x}(k) + \sum_{i=1}^{p} \boldsymbol{L}^{-1}\boldsymbol{A}(i)\boldsymbol{x}(k-i) = \boldsymbol{u}(k) + \sum_{j=1}^{q} \boldsymbol{L}^{-1}\boldsymbol{B}(j)\boldsymbol{L}\boldsymbol{u}(k-j). \quad (2)$$

Next introduce two scalar processes $y(n)$ and $v(n)$ related to $\boldsymbol{x}(k)$ and $\boldsymbol{u}(k)$, respectively, in the following manner: $y(j + d(k-1)) = x_j(k)$, $v(j + d(k-1)) = u_j(k)$, $j = 1, 2, \cdots, d$. Replacing the components of $\boldsymbol{x}(k)$ and $\boldsymbol{u}(k)$ in (2) by the corresponding elements of $y(n)$ and $v(n)$, respectively, and equating the rows of the LHS and the RHS, we observe that $y(n)$ is a periodically stationary ARMA process given by the following general form:

$$y(n) + \sum_{i=1}^{p_n} a_n(i)y(n-i) = v(n) + \sum_{j=1}^{q_n} b_n(j)v(n-j) \quad (3)$$

where

$$a_n(i) = a_{n+ld}(i), \qquad b_n(j) = b_{n+ld}(j),$$
$$p_n = p_{n+ld}, \qquad q_n = q_{n+ld}$$

for any integer l and $v(n)$ is periodically white, i.e., $E[v(n+m)v^*(n)] = \sigma_n^2\,\delta(m)$, where $\sigma_n^2\,\sigma_{n+ld}^2$ and * denotes complex conjugation. Equation (3) provides the scalar, periodic ARMA equivalent of the vector ARMA process. Using the fact that $\boldsymbol{x}(k)$ is a stationary process, it is easy to see that $y(n)$ has a periodic correlation function, i.e., $r(n, m) = r(n+d, m+d)$, where $r(n, m) = E[y(n)y^*(m)]$.

B. Innovation Model

Given a multivariate ARMA process $\boldsymbol{x}(k)$, there exists an infinite collection of models, each excited by a different white input producing the same output process $\boldsymbol{x}(k)$. There exists, however, a model which is a unique representation of the process $\boldsymbol{x}(k)$ and is known as the *innovation representation*, or, innovation model as the white input to the model, in this case, turns out to be the innovations of the process $\boldsymbol{x}(k)$. Let H be the Hilbert space of all zero-mean, finite-variance, complex-valued random variables with the following inner product: $\langle x, y \rangle = E(xy^*)$. Denote by $H(\boldsymbol{x}, k)$ the subspace of H spanned by the set

$$S(\boldsymbol{x}, k) = \{x_i(r) \mid i = 1, \cdots, d; \quad -\infty < r \le k\}$$

and let $x|_{\vartheta}$ denote the orthogonal projection of the random variable x on a subspace ϑ. Then the innovation corresponding to the ith channel at the time index k is defined as

$$z_i(k) = x_i(k) - x_i(k)|_{H(\boldsymbol{x}, k-1)}, \qquad i = 1, 2, \cdots, d. \quad (4)$$

We then have the following innovation representation of the given ARMA process:

$$\boldsymbol{x}(k) + \sum_{i=1}^{p} \boldsymbol{A}(i)\boldsymbol{x}(k-i) = z(k) + \sum_{j=1}^{q} \boldsymbol{B}'(j)z(k-j) \quad (5)$$

where $z(k) = [z_1(k), \cdots, z_d(k)]^t$ (superscript t denotes simple transposition). The innovation representation satisfies the following conditions: it is causal, stable and causally invertible and is uniquely identifiable under these conditions. It is then possible to establish the following.

Lemma 1: The Cholesky factorization of the innovation covariance matrix of a multivariate ARMA process generates the in-

[1]The periodic, scalar representation of a multichannel ARMA process was derived by the authors independently, who were unaware of the work of Cipra and Tlusty [7], who had earlier obtained a similar representation result.

novation sequence of the associated scalar, periodic ARMA process.

Proof: See the Appendix.

The above result implies that the scalar, periodic model derived from (5) represents the innovation model for $y(n)$. This provides the mathematical justification for the assumption made in the following section regarding the causality and the causal invertibility of the scalar model.

III. Parameter Estimation

A. AR Parameters

The causality of the model (3) implies that the current input at any index n is orthogonal to (i.e., uncorrelated with) all past outputs. Multiplying both sides of (3) by $y^*(n - q_n - j)$, $j = 1, 2, \cdots, p_n$, taking expectation and using the causality of $h(n, m)$, we get

$$\begin{aligned} &a_n(1)r(n-1, n-q_n-j) \\ &\quad + \cdots + a_n(p_n)r(n-p_n, n-q_n-j) \\ &\quad = -r(n, n-q_n-j) \qquad (j = 1, \cdots, p_n). \end{aligned} \tag{6}$$

Equation (6) provides the modified Yule-Walker (MYW) type equations corresponding to the index n. Altogether there will be d sets of the MYW-type equations, i.e., one each corresponding to the indices: $n, n-1, \cdots, n-d+1$. In [1], Sakai has proposed a parallel algorithm to solve the Yule-Walker type equations for a periodic AR process. The underlying matrices in his algorithm were Hermitian whereas the matrices associated with the MYW type equations are in general non-Hermitian. It is, however, possible to derive a similar algorithm for the non-Hermitian case also as is shown below.

In (6), we substitute q for q_n where $q = \max\{q_i \mid i = 1, \cdots, d\}$. The jth order solution for the AR parameters corresponding to the index n is obtained by solving

$$\boldsymbol{R}_n(j)\boldsymbol{a}_n(j) = [\alpha(n, j), 0, 0, \cdots, 0]^t \tag{7}$$

where $\boldsymbol{R}_n(j)$ is a $(j+1) \times (j+1)$ correlation matrix where

$$[\boldsymbol{R}_n(j)]_{l,m} = r(n-m+1, n-q-l+1), \qquad l, m = 1, 2, \cdots, (j+1)$$

and

$$\boldsymbol{a}_n(j) = [1, a_{n,j}(1), \cdots, a_{n,j}(j)]^t$$

and $\alpha(n, j)$ is a constant (unknown) given by

$$\begin{aligned} &r(n, n-q) + a_{n,j}(1)r(n-1, n-q) \\ &\quad + \cdots + a_{n,j}(j)r(n-j, n-q) = \alpha(n, j). \end{aligned} \tag{8}$$

Note that (7) is obtained by adding (8) to the set of equations given by (6). The periodicity of the scalar model implies that for any integer k, $\boldsymbol{R}_{n+kd}(j) = \boldsymbol{R}_n(j)$. Assume that the jth order solution for all the d indices is known. $\boldsymbol{R}_n(j+1)$ can then be segmented as follows:

$$\boldsymbol{R}_n(j+1) = \left[\begin{array}{c|c} \boldsymbol{R}_n(j) & \\ \hline & - \end{array}\right] = \left[\begin{array}{c|c} - & \\ \hline & \boldsymbol{R}_{n-1}(j) \end{array}\right]. \tag{9}$$

Introducing an auxiliary vector $\boldsymbol{c}_n(j) = [c_{n,j}(j), c_{n,j}(j-1), \cdots, c_{n,j}(1), 1]^t$, which satisfies

$$\boldsymbol{R}_n(j)\boldsymbol{c}_n(j) = [0, 0, \cdots, \beta(n, j)]^t \tag{10}$$

where

$$\begin{aligned} \beta(n, j) &= r(n, n-q-j)c_{n,j}(j) \\ &\quad + \cdots + r(n-j+1, n-q-j)c_{n,j}(1) \\ &\quad + r(n-j, n-q-j) \end{aligned} \tag{11}$$

we next seek the solution for $\boldsymbol{a}_n(j+1)$ and $\boldsymbol{c}_n(j+1)$ in the following forms:

$$\boldsymbol{a}_n(j+1) = [\boldsymbol{a}_n^t(j), 0]^t + K_n(j+1)[0, \boldsymbol{c}_{n-1}^t(j)]^t$$

and

$$\boldsymbol{c}_n(j+1) = [0, \boldsymbol{c}_{n-1}^t(j)]^t + K_n^t(j+1)[\boldsymbol{a}_n^t(j), 0]^t$$

where the scalars variables $K_n(j+1)$ and $K_n^t(j+1)$ have to be evaluated. Then, from (7), (9), and (10), we have

$$\begin{aligned} \boldsymbol{R}_n(j+1)\boldsymbol{a}_n(j+1) &= [\alpha(n, j), 0, 0, \cdots, 0, \delta(n, j)]^t \\ &\quad + K_n(j+1)[\tau(n-1, j), 0, 0, \\ &\quad \cdots, 0, \beta(n-1, j)]^t \end{aligned} \tag{12}$$

where

$$\begin{aligned} \delta(n, j) &= r(n, n-q-j-1) \\ &\quad + a_{n,j}(1)r(n-1, n-q-j-1) \\ &\quad + \cdots + a_{n,j}(j)r(n-j, n-q-j-1) \\ \tau(n-1, j) &= c_{n-1,j}(j)r(n-1, n-q) \\ &\quad + \cdots + c_{n-1,j}(1)r(n-j, n-q) \\ &\quad + r(n-j-1, n-q). \end{aligned} \tag{14}$$

Similarly,

$$\begin{aligned} &\boldsymbol{R}_n(j+1)\boldsymbol{c}_n(j+1) \\ &\quad = [\tau(n-1, j), 0, 0, \cdots, 0, \beta(n-1, j)]^t \\ &\quad + K_n^t(j+1)[\alpha(n, j), 0, 0, \cdots, 0, \delta(n, j)]^t. \end{aligned} \tag{15}$$

Comparing (12) and (15) with (7) and (10), respectively, the parameters $K_n(j+1)$ and $K_n^t(j+1)$ can then be obtained as

$$K_n(j+1) = -\delta(n, j)/\beta(n-1, j) \tag{16a}$$

$$K_n^t(j+1) = -\tau(n-1, j)/\alpha(n, j). \tag{16b}$$

Equations (8), (11), (13), (14) along with (16a) and (16b) result in a parallel algorithm which can be carried out by employing d processors, one each for the indices: $n, (n-1), \cdots, (n-d+1)$. The algorithm is listed in Table I. It is seen that the data flows between the processors in a circular manner. The algorithm requires $O(p^2d^2)$ operations per processor in contrast to the $O(p^2d^3)$ operations required by the usual multichannel Levinson-type algorithms.

B. MA Parameters

The MA parameters are estimated by estimating the input samples first. Using the causal invertibility of the periodic model and truncating the impulse response of the inverse filter up to a large but finite number of terms, say L, we obtain a high order periodic AR approximation of the scalar model. The high order AR parameters are estimated by the parallel algorithm of Sakai [1] and then used to filter the data to generate the input sample estimates. The

TABLE I
ALGORITHM FOR SOLVING THE MYW TYPE EQUATIONS. THE "DO-PAR-END-PAR" LOOP IS TO BE EXECUTED BY THE d PROCESSORS IN PARALLEL

Initialization ($j = 0$)

$\alpha(n, 0) = \beta(n, 0) = r(n, n - q), \quad n = 1, 2, \cdots, d.$

$a_{n,0}(0) = c_{n,0}(0) = 1, \quad n = 1, 2, \cdots, d.$

$p = \max\{p_n | n = 1, 2, \cdots, d\}.$

For $j = 0$ to $(p - 1)$ begin

do-par, $n = 1, 2, \cdots, d,$

$$\delta(n, j) = \sum_{l=0}^{j} a_{n,j}(l) r(n - l, n - q - j - 1);$$

$$\tau(n - 1, j) = \sum_{l=0}^{j} c_{n-1,j}(l) r(n - 1 - j + l, n - q);$$

$$K_n(j + 1) = -\delta(n, j)/\beta(n - 1, j);$$

$$K_n'(j + 1) = -\tau(n - 1, j)/\alpha(n, j);$$

For $i = 1$ to j begin

$$a_{n,j+1}(i) = a_{n,j}(i) + K_n(j + 1) c_{n-1,j}(j + 1 - i);$$

$$c_{n,j+1}(i) = c_{n-1,j}(i) + K_n'(j + 1) a_{n,j}(j + 1 - i);$$

end

$$a_{n,j+1}(0) = 1; \quad c_{n,j+1}(0) = 1;$$

$$a_{n,j+1}(j + 1) = K_n(j + 1); \quad c_{n,j+1}(j + 1) = K_n(j + 1);$$

$$\alpha(n, j + 1) = \alpha(n, j) + K_n(j + 1)\tau(n - 1, j);$$

$$\beta(n, j + 1) = \beta(n - 1, j) + K_n'(j + 1)\delta(n, j);$$

end-par

Transmit $\beta(n, j + 1)$ and $c_n(j + 1)$ from the nth processor to the $(n + 1)$th processor for $n = 1, 2, \cdots, (d - 1)$ and from the dth processor to the first processor for $n = d$.

end

MA parameters are then obtained by minimizing the following:

$$\epsilon^2(n) = \sum_{k=-L_1}^{L_2} \left| \left(y(n + kd) + \sum_{i=1}^{p_n} a_n(i) y(n + kd - i) - \sum_{j=1}^{q_n} b_n(j) e(n + kd - j) \right) \right|^2 \tag{17}$$

(in which L_1 and L_2 are to be chosen appropriately so as to cover the entire data record of length Nd to the maximum possible extent) and then solving the linear equations resulting from the minimization of $\epsilon^2(n)$ w.r.t. $b_n(j)$'s. The AR part in (17) is filtered out by replacing $a_n(i)$'s by their already estimated values.

IV. SIMULATION RESULTS

The algorithm was simulated on a real, two-variate, ARMA (1, 1) process $\boldsymbol{x}(k)$ with the following choice of the parameters $\boldsymbol{A}(1)$ and $\boldsymbol{B}(1)$ and the input covariance matrix $\boldsymbol{Q}$:

$$\boldsymbol{A}(1) = \begin{bmatrix} -0.85 & 0.75 \\ -0.65 & -0.55 \end{bmatrix}, \quad \boldsymbol{B}(1) = \begin{bmatrix} -1.9833 & 1.889 \\ -1.7 & 1.9833 \end{bmatrix},$$

$$\boldsymbol{Q} = \begin{bmatrix} 1.5625 & 1.5 \\ 1.5 & 1.5625 \end{bmatrix}.$$

A data record of 300 real, vector data points was taken and the model orders were assumed to be known *a priori*. The above choice of $\boldsymbol{A}(1)$ gives rise to poles at $0.9772 \exp(\pm j2\pi \cdot 0.123)$. The corresponding spectral plots are shown in Fig. 1 where $P_{11}(f)$ and $P_{22}(f)$ are the autospectra of $x_1(k)$ and $x_2(k)$, respectively; $P_{12}(f)$

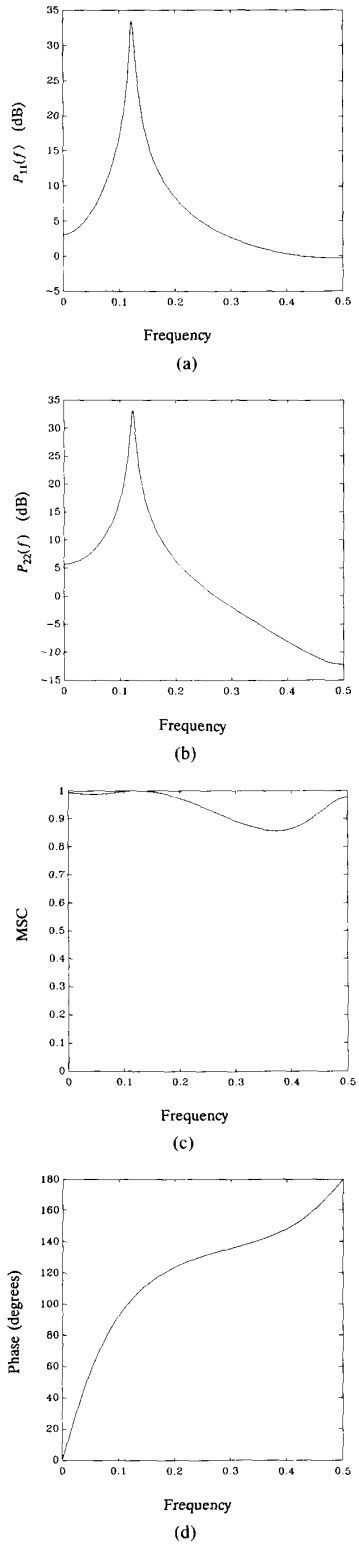

Fig. 1. (a) True autospectrum of the process $x_1(k)$ (decibels). (b) True autospectrum of the process $x_2(k)$ (decibels). (c) Original magnitude squared coherence (MSC). (d) Phase of the true cross power spectral density (degrees).

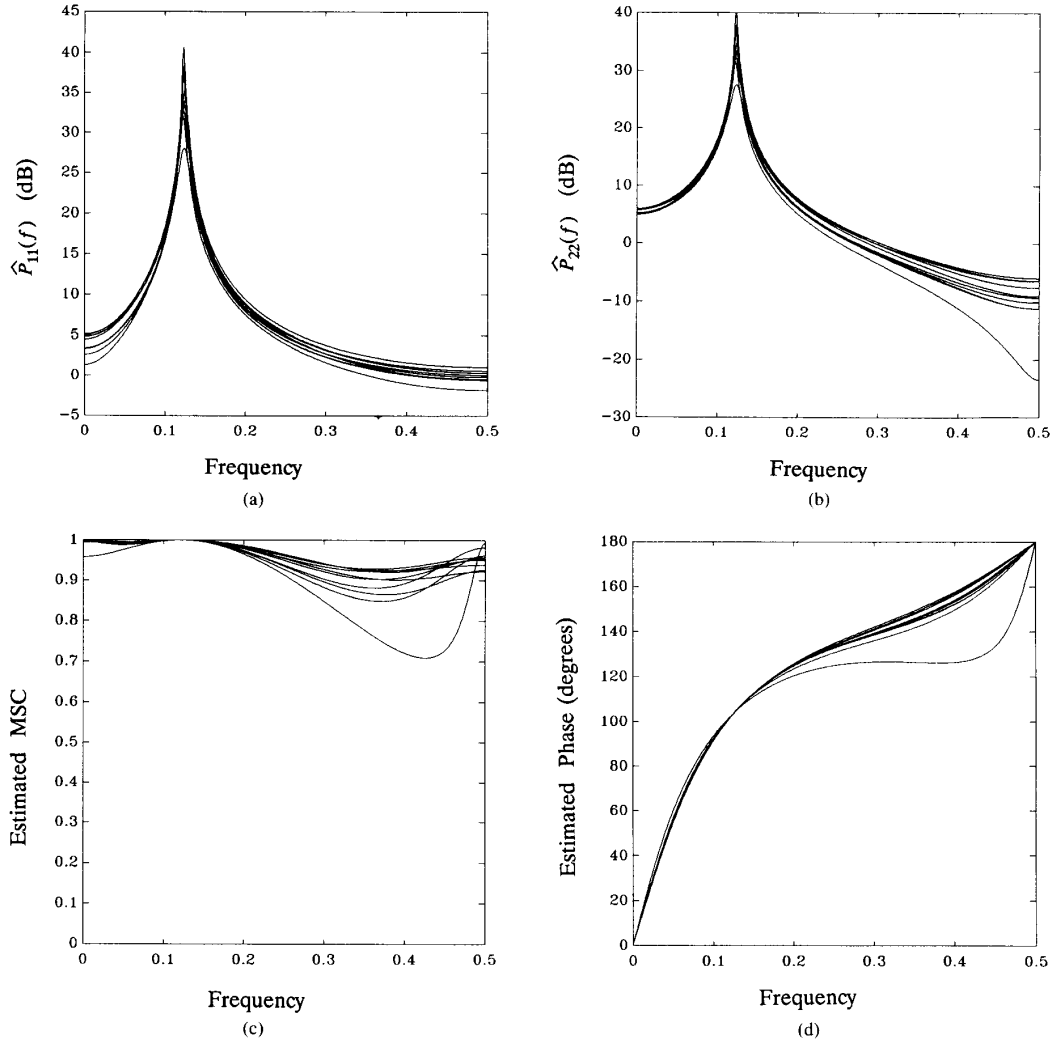

Fig. 2. Estimates of (a) the autospectrum of the process $x_1(k)$ (decibels), (b) the autospectrum of the process $x_2(k)$ (decibels), (c) the magnitude squared coherence (MSC), and (d) the phase of the cross power spectral density (degrees).

and $P_{21}(f)$ are the cross power-spectral densities between the two channels with $P_{12}(f) = P_{21}^*(f)$; and the magnitude-squared coherence (MSC) between the two processes $x_1(k)$ and $x_2(k)$ is defined as

$$|\gamma_{12}(f)|^2 = |P_{12}(f)|^2/[P_{11}(f)P_{22}(f)].$$

The MSC is bounded between 0 and 1. (The power-spectral-density matrix $\boldsymbol{P}_{xx}(f)$, where

$$[\boldsymbol{P}_{xx}(f)]_{i,j} = \boldsymbol{P}_{ij}(f), \qquad i, j = 1, 2$$

can be obtained as

$$\boldsymbol{P}_{xx}(f) = \boldsymbol{H}^*(f)\boldsymbol{Q}\boldsymbol{H}^t(f)$$

where

$$\boldsymbol{H}(z) = \left(\boldsymbol{I} + \sum_{i=1}^{p} \boldsymbol{A}(i)z^{-i}\right)^{-1}\left(\boldsymbol{I} + \sum_{j=1}^{q} \boldsymbol{B}(j)z^{-j}\right)$$

is the multichannel system function.) The scalar process $y(n)$ obtained from the two-variate ARMA process is periodically stationary with period equal to 2. The sample correlation matrix $\boldsymbol{R}(m)$ of the scalar process, where

$$[\boldsymbol{R}(m)]_{i,j} = r(m - j + 1, m - i + 1), \qquad i, j = 1, 2, \cdots, m$$

is calculated as $\boldsymbol{R}(m) = [\boldsymbol{Y}(m)\boldsymbol{Y}^t(m)]/M$, where $\boldsymbol{Y}(m)$ is the $m \times M$ data matrix with

$$[\boldsymbol{Y}(m)]_{i,j} = y(m - i + 2j - 1),$$
$$i = 1, 2, \cdots, m, \quad j = 1, 2, \cdots, M.$$

Here M is chosen to be sufficiently large so as to cover the recorded data to the maximum possible extent, i.e., $M = [(N - m)/2] + 1$, where N is the length of the data record and $[x]$ is an integer j such that $j \leq x < j + 1$. The model order chosen for the high order, periodic AR approximation of the scalar process was 20.

The simulation results are highlighted by displaying the estimated spectral plots. Fig. 2 shows 10 overlaid realizations of the

estimates of $P_{11}(f)$ and $P_{22}(f)$ (denoted respectively by $\hat{P}_{11}(f)$ and $\hat{P}_{22}(f)$, the MSC and the phase of the estimated cross power-spectral density. It is observed that the autospectra as well as the phase and coherence have been estimated reasonably well. The estimates follow the original ones closely near the pole frequency (i.e., 0.123) of the given model and some deviations from optimality can be noticed at frequencies far away from the pole frequency. This is, however, expected because the estimation of the AR parameters, which dictate the nature of the spectral plots near the pole frequency, involves less approximation in comparison to that of the MA parameters which have a dominant role in the nature of the plots at frequencies away from the pole frequency.

V. Conclusions

We have developed a method for identifying a multivariate ARMA process which employs scalar operations only and can be implemented on a pipelined processor.

Appendix

Proof of Lemma 1: The innovation corresponding to $y(n)$ at the index n is given by

$$e(n) = y(n) - y(n)|_{H(y,n-1)}. \tag{A1}$$

Also, it is easy to see that $H(\boldsymbol{x}, k-1) \equiv H(y, s-d)$, $s = kd$. Noting that

$$H(y, m) = \langle e(m) \rangle \oplus H(y, m-1)$$

(. $\oplus$. denotes the direct sum of the two associated vector spaces), we have the following direct sum decomposition of $H(y, s-d+r)$, $r = 1, 2, \cdots, d$:

$$H(y, s-d+r) = \langle e(s-d+r) \rangle \oplus \cdots \oplus \langle e(s-d+1) \rangle \oplus H(y, s-d) \tag{A2}$$

where the innovation $e(s)$ is given by (A1). It is easy to observe that $z_r(k)$, i.e., the rth component of the innovation vector $\boldsymbol{z}(k)$ of the multivariate process can be written as

$$z_r(k) = y(s-d+r) - y(s-d+r)|_{H(y,s-d)}, \quad s = kd. \tag{A3}$$

Now, from (A1), the innovation $e(s-d+r)$, $r = 1, \cdots, d$, is given by

$$e(s-d+r) = y(s-d+r) - y(s-d+r)|_{H(y,s-d+r-1)} \tag{A4}$$

where, from (A2),

$$y(s-d+r)|_{H(y,s-d+r-1)}$$

is obtained as

$$y(s-d+r)|_{H(y,s-d+r-1)} = y(s-d+r)|_{H(y,s-d)} + h_{r,1}e(s-d+1) + \cdots + h_{r,r-1}e(s-d+r-1). \tag{A5}$$

The $h_{r,i}$'s, $i = 1, \cdots, (r-1)$, in (A5) are given by

$$h_{r,i} = \langle y(s-d+r), e(s-d+i) \rangle / \langle e(s-d+i), e(s-d+i) \rangle. \tag{A6}$$

From (A4) and (A5) and making use of (A3), we get, for $r = 1, 2, \cdots, d$,

$$e(s-d+r) = z_r(k) - h_{r,1}e(s-d+1) - \cdots - h_{r,r-1}e(s-d+r-1), \quad s = kd. \tag{A7}$$

Now, by induction, we show that $e(s-d+r)$ is given by

$$e(s-d+r) = l_{r,1}z_1(k) + \cdots + l_{r,r-1}z_{r-1}(k) + z_r(k), \quad \text{where } l_{r,i} \in C, \ i = 1, \cdots, r-1, \quad \text{and } s = kd. \tag{A8}$$

Clearly, (A8) is satisfied for $r = 1$. Assume that (A8) is satisfied for r up to m, where $1 \le m \le d-1$. We will show that the result, i.e., (A8) is also valid for r up to $m+1$. Putting $r = (m+1)$ in (A7), substituting $e(s-d+i)$, $i = 1, \cdots, m$, by the corresponding expressions suggested by the RHS of (A8) and rearranging the terms, we obtain

$$e(s-d+m+1) = l_{m+1,1}z_1(k) + \cdots + l_{m+1,m}z_m(k) + z_{m+1}(k), \quad s = kd \tag{A9}$$

which conforms to the general form given by (A8) where the coefficients $\{l_{r,i}\}$ can be expressed in terms of $\{h_{j,i}\}$, $j = 1, \cdots, r$, $i = 1, \cdots, r-1$.

Defining the vector $\boldsymbol{e}(s) = [e(s-d+1), \cdots, e(s)]'$, (A8) can be written in the following form: $\boldsymbol{e}(s) = \boldsymbol{L}''\boldsymbol{z}(k)$, $s = kd$, where $\boldsymbol{L}''$ is a unit lower triangular matrix. It is easily seen that $\boldsymbol{e}(s)$ is a white process with diagonal covariance matrix and therefore the matrix $\boldsymbol{L}''$ performs the Cholesky factorization of the covariance matrix of $\boldsymbol{z}(k)$. This proves that the Cholesky factorization of the innovation covariance matrix of the multivariate process generates the innovation sequence of the associated scalar, periodic ARMA process.

References

[1] H. Sakai, "Circular lattice filtering using Pagano's method," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, no. 2, pp. 279-287, Apr. 1982.

[2] T. Kawase, H. Sakai, and H. Tokumaru, "Recursive least squares circular lattice and escalator estimation algorithms," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, no. 1, pp. 228-231, Feb. 1983.

[3] F. Ling and J. G. Proakis, "A generalized multichannel least squares lattice algorithm based on sequential processing stages," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, no. 32, pp. 381-390, Apr. 1984.

[4] H. Lev-Ari, "Modular architectures for adaptive multichannel lattice algorithms," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, no. 4, pp. 543-552, Apr. 1987.

[5] P. S. Lewis, "*QR*-based algorithms for multichannel adaptive least squares lattice filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 421-432, Mar. 1990.

[6] M. Pagano, "Periodic and multiple autoregressions," *Ann. Statist.*, vol. 6, pp. 1310-1317, Nov. 1978.

[7] T. Cipra and P. Tlusty, "Estimation in multiple autoregressive-moving average models using periodicity," *Time Series Anal.*, vol. 8, no. 3, pp. 293-300, 1987.