From (4.7)–(4.10) the following closed-form nonlinear system of state equations is derived:

$$\dot{\zeta}_{i0} = C_0^i \zeta_{i0} + C_1^i \zeta_{i1} + C_2^i \zeta_{i2} \tag{4.13}$$

$$\zeta_{i0}(0) = 1; \ i = 1, \cdots, N$$

$$\dot{\zeta}_{i1} = C_1^i \zeta_{i2} + C_0^i \zeta_{i1} + C_2^i \left(3 \frac{\zeta_{i1}(t)\zeta_{i2}(t)}{\zeta_{i0}(t)} - 2 \frac{\zeta_{i1}^3(t)}{\zeta_{i0}^2(t)} \right) \tag{4.14}$$

$$\zeta_{i1}(0) = E\{x_0\}; \ i = 1, \cdots, N$$

$$\zeta_{i2} = C_0^i \zeta_{i2} + C_1^i \left(3 \frac{\zeta_{i1}(t)\zeta_{i2}(t)}{\zeta_{i0}(t)} - 2 \frac{\zeta_{i1}^3(t)}{\zeta_{i0}^2(t)} \right) + C_2^i \left(3 \frac{\zeta_{i2}^2(t)}{\zeta_{i0}(t)} - 2 \frac{\zeta_{i1}^4(t)}{\zeta_{i0}^3(t)} \right) \tag{4.15}$$

$$\zeta_{i2}(0) = E\{x_0^2\}.$$

The $3N$ state variables ζ_{iv} are the sufficient statistics for the first- and second-order filter approximations.

The three approximations to $q_j(x, t)$; $j = 2, \cdots, N$ are as follows:

$$q_j(x, t) \cong (p_j(0) + p_1(0)\lambda t)\tilde{q}_1(x, t) \tag{4.16a}$$

or

$$q_j(x, t) = p_j(0) q_j(x, t) + p_1(0)\lambda t e^{-\lambda t} \tilde{q}_1(x, t) \tag{4.16b}$$

or

$$q_j(x, t) \cong \left(p_j(0) + \frac{\lambda t}{2} \right) \tilde{q}_j(x, t) + p_1(0)(a_{j1}) \frac{\lambda t}{2} e^{-\lambda t} \tilde{q}_1(x, t). \tag{4.16c}$$

The output map is shown below for (4.16a) [a similar derivation can be used for (4.16b) and (4.16c)].

The normalizing factor approximation is

$$\hat{\Lambda}_t \cong \sum_{j=2}^{N} [p_j(0) + a_{j1} p_1(0)\lambda t]\zeta_{j0} + p_1(0) e^{-\lambda t} \zeta_{10}. \tag{4.17}$$

Then

$$p_1(t) \cong p_1(0) e^{-\lambda t} \zeta_{10} (\hat{\Lambda}_t)^{-1} \tag{4.18}$$

$$p_j(t) = [p_j(0)\zeta_{j0} + a_{j1} p_1(0) \lambda t e^{\lambda t} \zeta_{10}](\hat{\Lambda}_t)^{-1}; \qquad j = 2, \cdots, N. \tag{4.19}$$

V. Summary and Conclusions

This note analyzed the nonlinear filtering problem for a linear system with random structure governed by a FSCT Markov process. A characterization of the optimal mean-square filter is derived based on an explicit solution of the DMZ equation for the unnormalized conditional density. Various suboptimal filters can be derived by approximations of the integrals involved in the solution of the DMZ equation. One important conclusion that can be drawn is that the conditional density is given by a convergent power series expansion in terms of the FSCT Markov process transition rate parameters. Asymptotic expressions as suggested in the work of Blankenship *et al.* [5] can be analyzed rigorously within the framework developed. From an applications perspective, the work contained in the current note is important because it identifies the solution of the proposed nonlinear filtering problem in the context of solutions of known finite-dimensional filtering problems.

References

[1] O. Hijab, "Estimation of Markov processes," *Syst. Contr. Lett.*, vol. 1, no. 1, pp. 42–48, July 1981.

[2] V. Benes, "Exact finite dimensional filters for certain diffusions with nonlinear drift," *Stochastics*.

[3] O. Hijab, "The adaptive LQG problem—Part I," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 171–178, Feb. 1983.

[4] A. Makowski, "Results on the filtering problem for linear systems with non-gaussian initial conditions," in *Proc. 21st IEEE Conf. Decision Contr.*, Dec. 1982.

[5] G. Blankenship, C. Liu, and S. Marcus, "Asymptotic expansions and lie algebras for some nonlinear filtering problems," *IEEE Trans. Automat. Contr.*, vol. AC-28, July 1983.

[6] E. Pardoux, "Stochastic partial differential equations and filtering of diffusion processes," *Stochastics*, vol. 3, pp. 127–167, 1979.

[7] "Stochastic systems: The mathematics of filtering and identification and applications," (NATO Advanced Study Institute Series), M. Hazewinkel and J. Williams, Ed. Amsterdam, The Netherlands: Reidel, 1980.

[8] K. Loparo, S. Eckert, and Z. Roth, "An expansion of the conditional density for systems with random structure: The multivariable case," in preparation.

[9] M. Davis, "Pathwise solutions and multiplicative functionals in nonlinear filtering," in *Proc. 18th IEEE Conf. Decision Contr.*, Ft. Lauderdale, FL, 1979.

[10] M. Davis, "Pathwise nonlinear filtering," (NATO Advanced Study Institute Series), M. Hazewinkel and J. Williams, Ed. Amsterdam, The Netherlands: Reidel, 1980.

[11] S. Eckert, K. Loparo, and Z. Roth, "An application of nonlinear filtering to instrument failure detection in a pressurized water reactor," *Nuc. Technol.*, 1986, to be published.

An Asymptotically Efficient ARMA Estimator Based on Sample Covariances

PETRE STOICA and ARYE NEHORAI

Abstract—**An asymptotically efficient autoregressive moving-average (ARMA) spectral estimator is presented, based on the sample covariances of observed time series. The estimate of the autoregressive (AR) part is shown to be identical to the optimal instrumental variable (IV) estimator in [7] although derived here using a different approach. The moving-average (MA) spectral parameter estimate is new.**

I. Introduction

The subject of autoregressive moving-average (ARMA) parameter estimation has attracted considerable interest in various fields, such as engineering, statistics, econometrics, biometrics, and others. Since the maximum likelihood (ML) solution to this problem is known to be fairly complicated, various alternative solutions have been proposed. These solutions aim at preserving the asymptotic properties of the ML estimator, namely consistency and efficiency (see, e.g., [1], [4], [6], [7], [10], [11]).

In this note we propose a novel asymptotically efficient ARMA spectral estimator. The proposed estimator is based on a new method of approximating a statistically efficient ARMA estimator recently proposed by Porat and Friedlander [1], [2]. While the estimator of [1] is mainly of

Manuscript received May 8, 1986. The work of A. Nehorai was supported in part by the National Science Foundation under Grant DCI-8604351.

P. Stoica is with the Facultatea de Automatica, Institutul Politehnic Bucuresti, Bucharest, Romania.

A. Nehorai is with the Department of Electrical Engineering, Yale University, New Haven, CT 06520.

IEEE Log Number 8610719.

theoretical interest due to its difficult implementation, our proposed method leads to a computationally feasible scheme that preserves the asymptotic distribution of the former.

II. The Model and the Cost Function

Consider an ARMA process satisfying

$$A(q^{-1})y(t)=C(q^{-1})e(t) \tag{1a}$$

where $e(t)$ is white noise of zero mean and variance λ^2, $A(q^{-1})$ and $C(q^{-1})$ are polynomials of degree n in the unit delay operator q^{-1}

$$A(q^{-1})=1+a_1q^{-1}+\cdots+a_nq^{-n} \tag{1b}$$

$$C(q^{-1})=1+c_1q^{-1}+\cdots+c_nq^{-n}. \tag{1c}$$

The equal degree condition is imposed for convenience of notation.

Let

$$r_k=Ey(t)y(t+k) \qquad k=0,\ 1,\ 2,\ \cdots \tag{2a}$$

and define the unbiased sample covariances

$$\tilde{r}_k=\frac{1}{N-k}\sum_{t=1}^{N-k} y(t)y(t+k), \qquad \tilde{r}_{-k}=\tilde{r}_k \tag{2b}$$

where N denotes the number of data points. Introduce the parameter vector

$$\theta=[r_0 \cdots r_n a_1 \cdots a_n]^T. \tag{3}$$

It is easy to see that θ determines the ARMA spectrum completely (see, e.g., [4], [6]). In this note we shall derive an asymptotically efficient estimator for the ARMA spectral parameters (3).

Let

$$\rho=[\tilde{r}_0-r_0 \cdots \tilde{r}_m-r_m]^T \tag{4}$$

and denote the asymptotic (for $N \to \infty$) covariance of ρ by

$$\Sigma=\text{as cov }(\rho). \tag{5a}$$

An expression for Σ is presented, e.g., in [3], [6]. The ij entry of Σ is given by

$$\Sigma_{ij}=E[v(t-i)+v(t+i)]v(t-j) \qquad i,\ j=0,\ 1,\ \cdots,\ m \tag{5b}$$

$$v(t)=\lambda\frac{C^2(q^{-1})}{A^2(q^{-1})}e(t). \tag{5c}$$

From [1], [2] it follows that

$$\hat{\theta}=\arg\min_{\theta} V(\theta); \qquad V(\theta)=\rho^T\Sigma^{-1}\rho \tag{6}$$

is an asymptotically (for $N \to \infty$ and $m \to \infty$) efficient estimate of θ. Furthermore, it is shown in [1], [2] that $\hat{\theta}$ is the minimum variance estimate of θ, in the fairly large class of estimators which are based on the first m sample covariances. Thus, the estimator (6) has neat properties. However, it appears too difficult to use: $V(\theta)$ is a strongly nonlinear function of θ; moreover, the matrix Σ of large dimension $[(m+1)\times(m+1)]$ needs to be evaluated and inverted at each computation of the cost function in (6) (observe that Σ depends on θ). In the following we use (6) to derive a comparatively simple (linear) estimator, which preserves the asymptotic distribution of (6).

III. The Proposed Method

To derive our estimator we first rewrite the cost function $V(\theta)$ in an alternative form. Let

$$A=\begin{bmatrix} 1 & & & & & & 0 \\ & \ddots & & & & & \\ 0 & & 1 & & & & \\ 0 \;\; a_n & \cdots & a_1 & 1 & & & \\ & \ddots & & \ddots & \ddots & & \\ 0 & & & & & \ddots & \\ & & 0 & a_n & \cdots & a_1 & 1 \end{bmatrix} \quad (m+1)\times(m+1). \tag{7}$$

We have

$$V(\theta)=(A\rho)^T(A\Sigma A^T)^{-1}(A\rho). \tag{8}$$

Let

$$\tilde{R}=\begin{bmatrix} \tilde{r}_n & \cdots & \tilde{r}_1 \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \tilde{r}_{m-1} & \cdots & \tilde{r}_{m-n} \end{bmatrix} \qquad \tilde{r}=\begin{bmatrix} \tilde{r}_{n+1} \\ \cdot \\ \cdot \\ \cdot \\ \tilde{r}_m \end{bmatrix}. \tag{9}$$

Using the Yule–Walker equations $\Sigma_{i=0}^n a_i r_{n+k-i}=0$ for $k \geq 1$, we get

$$A\rho=\begin{bmatrix} \tilde{\beta}-\beta \\ \tilde{r}+\tilde{R}a \end{bmatrix} \tag{10a}$$

where

$$\beta=[r_0 \cdots r_n]^T \tag{10b}$$

$$a=[a_1 \cdots a_n]^T \tag{10c}$$

and where $\tilde{\beta}$ denotes the vector β made of $\{\tilde{r}_k\}$.

Next we turn to the calculation of

$$W \triangleq A\Sigma A^T. \tag{11}$$

Note first that for two stationary processes $x(t)$ and $w(t)$, we have

$$EA(q)x(t)\cdot A(q)w(t)=EA(q^{-1})x(t)\cdot A(q^{-1})w(t) \tag{12a}$$

$$Ex(t)\cdot A(q)w(t)=EA(q^{-1})x(t)\cdot w(t). \tag{12b}$$

Now

$$W=E\begin{bmatrix} v(t)+v(t) \\ \cdot \\ \cdot \\ \cdot \\ v(t-n)+v(t+n) \\ \hline A(q)v(t-n-1)+A(q^{-1})v(t+n+1) \\ \cdot \\ \cdot \\ \cdot \\ A(q)v(t-m)+A(q^{-1})v(t+m) \end{bmatrix} \cdot [v(t) \cdots v(t-n) : A(q)v(t-n-1) \cdots A(q)v(t-m)]. \tag{13}$$

From (12b), we get

$$\begin{aligned} &EA(q^{-1})v(t+n+i)\cdot A(q)v(t-n-j)\\ &=EA^2(q^{-1})v(t+n+i)\cdot v(t-n-j)\\ &=\lambda EC^2(q^{-1})e(t+n+i)\cdot v(t-n-j)=0 \qquad i,\ j\geq 1 \end{aligned} \tag{14}$$

and

$$\begin{aligned} &E[v(t-i)+v(t+i)]\cdot A(q)v(t-n-j)\\ &=Ev(t-i)\cdot[A(q)v(t-n-j)+A(q^{-1})v(t+n+j)]\\ &=EA(q^{-1})v(t-i)\cdot v(t-n-j)+Ev(t-i)\cdot A(q^{-1})v(t+n+j). \end{aligned} \tag{15}$$

Thus, W can be partitioned as

$$W=\begin{bmatrix} W_{11} & W_{12}\\ W_{12}^T & W_{22}\end{bmatrix} \tag{16a}$$

where

$$[W_{11}]_{ij}=\Sigma_{ij}=E[v(t-i)+v(t+i)]v(t-j) \qquad i,\ j=0,\ \cdots,\ n \tag{16b}$$

$$[W_{12}]_{ij}=EA(q^{-1})v(t-i)\cdot v(t-n-j)+Ev(t-i)\cdot A(q^{-1})v(t+n+j) \tag{16c}$$

$$i=0,\ \cdots,\ n \qquad j=1,\ \cdots,\ m-n$$

$$[W_{22}]_{ij}=EA(q^{-1})v(t-i)\cdot A(q^{-1})v(t-j) \qquad i,\ j=1,\ \cdots,\ m-n. \tag{16d}$$

In writing (16d) we made use of (12a).

We are now in a position to present the modified cost function $V(\theta)$ and the corresponding simple large-sample approximation of $\hat{\theta}$ in (6). From the above calculations it follows that $\hat{\theta}=[\hat{\beta}^T\hat{a}^T]^T$ is the minimizer of

$$V(\theta)=[(\tilde{\beta}-\beta)^T\ \ (\tilde{r}+\tilde{R}a)^T]W^{-1}\begin{bmatrix}\tilde{\beta}-\beta\\ \tilde{r}+\tilde{R}a\end{bmatrix}. \tag{17}$$

Differentiation of (17) gives

$$\frac{\partial V(\theta)}{\partial\beta}=-2[I\ \ 0]W^{-1}\begin{bmatrix}\tilde{\beta}-\beta\\ \tilde{r}+\tilde{R}a\end{bmatrix}+\begin{bmatrix}\mu^T(\theta)\dfrac{\partial W^{-1}}{\partial r_0}\mu(\theta)\\ \vdots\\ \mu^T(\theta)\dfrac{\partial W^{-1}}{\partial r_n}\mu(\theta)\end{bmatrix} \tag{18a}$$

$$\frac{\partial V(\theta)}{\partial a}=2[0\ \ \tilde{R}^T]W^{-1}\begin{bmatrix}\tilde{\beta}-\beta\\ \tilde{r}+\tilde{R}a\end{bmatrix}+\begin{bmatrix}\mu^T(\theta)\dfrac{\partial W^{-1}}{\partial a_1}\mu(\theta)\\ \vdots\\ \mu^T(\theta)\dfrac{\partial W^{-1}}{\partial a_n}\mu(\theta)\end{bmatrix} \tag{18b}$$

where

$$\mu(\theta)=[(\tilde{\beta}-\beta)^T\ \ (\tilde{r}+\tilde{R}a)^T]^T. \tag{18c}$$

As shown in [1], [2], $\hat{\theta}$ is a "root N consistent" estimate of θ, i.e., $|\hat{\theta}-\theta|=0(1/\sqrt{N})$ for large N. This readily implies that $\mu(\hat{\theta})=0(1/\sqrt{N})$. Thus, the second term in both (18a) and (18b) when evaluated at $\hat{\theta}$ is $0(1/N)$. With this observation, and recalling the following identity (see, e.g., [8]).

$$W^{-1}=\begin{bmatrix}0\\ I\end{bmatrix}W_{22}^{-1}[0\ \ I]+\begin{bmatrix}-I\\ W_{22}^{-1}W_{12}^T\end{bmatrix}(W_{11}-W_{12}W_{22}^{-1}W_{12}^T)^{-1}[-I\ \ W_{12}W_{22}^{-1}] \tag{19}$$

it follows from (18a) and (18b) that

$$\begin{cases} \tilde{\beta}-\hat{\beta}-W_{12}W_{22}^{-1}(\tilde{r}+\tilde{R}\hat{a})+0(1/N)=0 & (20a)\\ \tilde{R}^TW_{22}^{-1}(\tilde{r}+\tilde{R}\hat{a})+\tilde{R}^TW_{22}^{-1}W_{12}^T(W_{11}-W_{12}W_{22}^{-1}W_{12}^T)^{-1} & \\ \quad\cdot[W_{12}W_{22}^{-1}(\tilde{r}+\tilde{R}\hat{a})-\tilde{\beta}+\hat{\beta}]+0(1/N)=0 & (20b)\end{cases}$$

where W_{ij} are evaluated at $\hat{\theta}$. The equations above are still fairly complicated functions of $\hat{\theta}$. However, observe that due to (20a) the second term in (20b) is of order $1/N$. Further, let $\tilde{W}_{ij}$ be a root N consistent estimate of W_{ij}. Then, since $\tilde{r}+\tilde{R}\hat{a}=0(1/\sqrt{N})$ we get from (20a) and (20b)

$$\begin{cases} \hat{\beta}=\tilde{\beta}-\tilde{W}_{12}\tilde{W}_{22}^{-1}(\tilde{r}+\tilde{R}\hat{a})+0(1/N) & (21a)\\ (\tilde{R}^T\tilde{W}_{22}^{-1}\tilde{R})\hat{a}=-(\tilde{R}^T\tilde{W}_{22}^{-1}\tilde{r})+0(1/N). & (21b)\end{cases}$$

From the last two equations it follows immediately that a simple (linear) large-sample approximation of $\hat{\theta}$ can be obtained as

$$\begin{cases} \hat{\hat{a}}=-(\tilde{R}^T\tilde{W}_{22}^{-1}\tilde{R})^{-1}(\tilde{R}^T\tilde{W}_{22}^{-1}\tilde{r}) & (22a)\\ \hat{\hat{\beta}}=\tilde{\beta}-\tilde{W}_{12}\tilde{W}_{22}^{-1}(\tilde{r}+\tilde{R}\hat{\hat{a}}). & (22b)\end{cases}$$

It remains to discuss how to obtain the root N consistent estimates $\tilde{W}_{12}$, $\tilde{W}_{22}$ of W_{12}, W_{22}. These estimates have to be constructed from initial estimates $\tilde{\beta}$ and $\tilde{a}$ of β and a; with $\tilde{\beta}$ = sample covariances and $\tilde{a}$ obtained, for example, using the (overdetermined) Yule-Walker (YW) method [4]. It is easy to see from (16a)-(16d) that this can be done if we have estimates of $\{a_i\}$ and of $\gamma_k=Ev(t)v(t-k)$.

A possible simple procedure for estimating γ_k is as follows. Let

$$G(q^{-1})=C^2(q^{-1}),\ \ D(q^{-1})=A^2(a^{-1}),$$

$$x(t)=\frac{1}{\lambda D(q^{-1})}e(t)\ \text{ and }\ \delta_k=Ex(t)x(t-k).$$

Given $D(q^{-1})$, $\{\delta_k\}$ can be computed using standard techniques for computing the covariances of an AR process, for instance the "inverse" Levinson algorithm, see, e.g., [5], [8], [9] (note that λ is not needed, as $E[e(t)/\lambda]^2=1$). Next, observe that $v(t)=\lambda^2G(q^{-1})x(t)$ and thus

$$\begin{aligned}\gamma_k&=\lambda^4E\sum_{i,j=0}^{2n}g_ig_jx(t-i)x(t-j-k)\\ &=\lambda^4\sum_{i,j=0}^{2n}g_ig_j\delta_{j+k-i}=\lambda^4\sum_{p=-2n}^{2n}\left(\sum_{i=0}^{2n}g_ig_{i-p}\right)\delta_{k-p}\end{aligned} \tag{23}$$

where $g_0=1$ and $g_k=0$ for $k>2n$ or $k<0$. The sum $\lambda^4\sum_{i=0}^{2n}g_ig_{i+p}$ equals the covariance at lag p of $\lambda C^2(q^{-1})e(t)$ or the coefficient of z^p in $[\lambda^2C(z)C(z^{-1})]^2$. Estimation of $\lambda^2C(z)C(z^{-1})$ from $\{a_i\}_1^n$ and $\{r_i\}_0^n$ can be done as follows. Let

$$\lambda^2 C(z)C(z^{-1}) = \sum_{k=-n}^{n} b_k z^{-k}. \tag{24}$$

Then

$$\begin{aligned} b_k &= E[C(q^{-1})e(t) \cdot C(q^{-1})e(t-k)] \\ &= E[A(q^{-1})y(t) \cdot A(q^{-1})y(t-k)] \\ &= \sum_{i,j=0}^{n} a_i a_j r_{k+j-i} \\ &= \sum_{i=0}^{n} \sum_{j=0}^{n-k} a_i a_j r_{k+j-i}; \quad b_{-k} = b_k \quad k=0, \cdots, n \end{aligned} \tag{25}$$

where in the last equality we utilized the YW equations. Note that (25) is a function of $\{a_i\}_{i=1}^n$ and $\{r_k\}_{k=0}^n$ only.

In summary, the implementation of the ARMA spectral parameter estimator (22) introduced in this note, can be done in the following steps.

Step 1: Use (2b) and (22a) with $\tilde{W}_{22} \equiv I$, to obtain the initial (and in general, inefficient) estimates $\tilde{a}$ and $\tilde{\beta}$ of a and β.

Step 2: Insert $\tilde{a}$ and $\tilde{\beta}$ into (25) to obtain estimates $\{\tilde{b}_k\}$ of $\{b_k\}$; also use $\tilde{a}$ to estimate $\{\delta_k\}$. From $\{\tilde{b}_k\}$ and $\{\tilde{\delta}_k\}$ estimate $\{\gamma_k\}$ using (23). Next, use $\tilde{a}$ and $\{\tilde{\gamma}_k\}$ in (16c) and (16d) to determine estimates $\tilde{W}_{12}$ and $\tilde{W}_{22}$ of W_{12} and W_{22}.

Step 3: Compute asymptotically efficient estimates of a and β with (22), where $\tilde{W}_{12}$ and $\tilde{W}_{22}$ are provided from the previous step.

Note that m should be set to some large value to achieve maximum possible accuracy. Also, note that Steps 2 and 3 of the procedure above may be iterated, if desired, using the refined estimates $\hat{\hat{a}}$ and $\hat{\hat{\beta}}$ in Step 2 to improve the estimates of W_{12} and W_{22}.

IV. Concluding Remarks

It is of interest to compare the estimates (22a), (22b) with previous algorithms. The proposed estimate of the AR subvector of θ is identical to the optimal IV estimate of the AR parameters of an ARMA process introduced in [7]–[10] (cf. the OIV-1 method). The estimates of the β parameters have not appeared before in literature. The novel estimator of this note provides an asymptotically efficient estimate of the ARMA spectrum. This is in contrast to the IV formalism of [7], [10] which provides only an efficient estimate of the AR part in the spectrum.

Another ARMA spectral parameter estimator that can also be obtained as a large-sample approximation of (6) was recently proposed in [6]. The estimator of [6] seems to be the closest competitor to the estimator proposed in this note. Numerical evaluation of the performance for the latter estimator in the finite data case as well as comparison to the former estimator, are subjects left for future research.

References

[1] B. Porat and B. Friedlander, "Asymptotic accuracy of ARMA parameter estimation techniques based on sample covariances," in *Proc. 7th IFAC/IFORS Symp. Identification and Syst. Parameter Estimation*, York, U.K., 1985, pp. 963–968.

[2] B. Porat and B. Friedlander, "Bounds on the accuracy of Gaussian ARMA parameter estimation methods based on sample covariances," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 579–582, June 1986.

[3] B. Porat, "Some asymptotic properties of the sample covariances of Gaussian autorepressive moving average processes," *J. Times Series Anal.*, submitted for publication.

[4] J. A. Cadzow, "Spectral estimation: An overdetermined rational model equation approach," *Proc. IEEE*, vol. 70, pp. 907–938, 1982.

[5] B. Porat, "ARMA spectral estimation based on partial autocorrelations," *Circuits Syst. Signal Processing*, vol. 2, pp. 341–360, 1983.

[6] P. Stoica, B. Friedlander, and T. Söderström, "An approximate maximum likelihood approach to ARMA spectral estimation," Syst. Control Technol., Palo Alto, CA, Tech. Rep. 5498-03, 1985; also in *Proc. 24th Conf. on Decision Contr.*, Fort Lauderdale, FL, Dec. 1985, pp. 1276–1281.

[7] P. Stoica, T. Söderström, and B. Friedlander, "Optimal instrumental variable estimates of the AR parameters of an ARMA process," *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 1066–1074, Nov. 1985.

[8] T. Söderström and P. Stoica, *Instrumental Variable Methods for System Identification*. New York: Springer-Verlag, 1983.

[9] P. Stoica and T. Söderström, "Optimization with respect to covariance sequence parameters," *Automatica*, vol. 21, pp. 671–675, no. 6, Nov. 1985.

[10] P. Stoica, B. Friedlander, and T. Söderström, "Optimal instrumental variable multistep algorithms for estimation of the AR parameters of an ARMA process," Syst. Control Technol., Palo Alto, CA, Tech. Rep. 5498-04, 1984; also in *Proc. 24th IEEE Conf. Decision Contr.*, Fort Lauderdale, FL, Dec. 1985, pp. 1087–1092.

[11] *Proc. IEEE*, Special issue on Spectral Estimation, vol. 70, no. 9, Sept. 1982.

An Adaptation Algorithm for Parallel Model Reference Adaptive Bilinear Systems

Y. K. Yeo and D. C. Williams

Abstract—**An adaptation algorithm is presented for parallel model reference adaptive bilinear systems. The output error converges asymptotically to zero and the parameter estimates are bounded for stable reference models. The convergence criterion depends only upon the input sequence and *a priori* estimates of the maximum parameter values. A passivity condition, which is generally difficult to verify, is not required.**

I. Introduction

Convergence criteria which do not require an SPR condition have been derived previously for parallel model reference adaptive linear systems [1], [2]. Although the criteria are presented in different forms, they are shown to be equivalent in [2]. Here we present an adaptation algorithm and derive a convergence criterion for parallel model reference adaptive bilinear systems. The criterion does not require a passivity condition. We follow the simple algebraic approach of Altay [2].

II. Adaptation Algorithm and Convergence Criterion

Identification Problem

The plant to be identified is described by a single-input single-output bilinear equation:

$$y(k) = \sum_{i=1}^{n} [a_i y(k-i) + b_i y(k-i)u(k-i) + c_i u(k-i)] = p^T x(k-1) \tag{1}$$

where y is the output and u is the input. The model to be adjusted is

$$\begin{aligned} \hat{y}(k) &= \sum_{i=1}^{n} [\hat{a}_i(k)\hat{y}(k-i) + \hat{b}_i(k)\hat{y}(k-i)u(k-i) + \hat{c}_i(k)u(k-i)] \\ &= \hat{p}^T(k) w(k-1) \end{aligned} \tag{2}$$

Manuscript received October 22, 1985; revised March 31, 1986 and June 10, 1986. This work was supported by the Auburn University Engineering Experiment Station and the Auburn University Advanced Manufacturing Technology Center.

The authors are with the Department of Chemical Engineering, Auburn University, Auburn, AL 36847.

IEEE Log Number 8610442.