ON-LINE STOCHASTIC APPROXIMATION ALGORITHMS FOR CA^K-IDENTIFIABLE STRUCTURES

C. S. Kubrusly and M. D. Fragoso

Department of Electrical Engineering, Pontifícia Universidade Católica, R. Marques de S. Vicente 209, ZC-20, Rio de Janeiro, RJ 22453, Brazil

Abstract. The on-line applicability of stochastic approximation algorithms for identifying linear multivariable discrete-time dynamical systems operating in a stochastic environment is investigated. Several recursive algorithms, which are grouped into five disjoint classes, are applied as straightforward on-line identification procedures for CA^K-identifiable structures by using a general unified version of the explicit parameter lemma. This approach avoids any matrix inversion during the iterative identification process, does not require any specific type of probability distribution, and assumes acessibility only over the noisy observation sequence.

Keywords. Difference Equations. Discrete time systems. Identification. Iterative methods. On-line operation. Multivariable systems. Parameter Estimation. Stochastic approximation. Stochastic systems.

1. INTRODUCTION

CA^K-identifiable structures were recently introduced by Kubrusly (1978-b) in order to meet the following question: what kind of linear structures for multivariable dynamical systems operating in a stochastic environment (i.e, driven by random disturbances and observed through noisy measurements) are amenable to the use of the explicit parameter lemma? The explicit parameter lemma acts as an intermediate stage for the system identification problem, where the final parameter determination can be recursively achieved by stochastic approximation algorithms of the type discussed in Kubrusly (1978-a). Such a recursive identification procedure naturally requires the inversion of estimates for the state correlation matrix at each iteration, thus compromizing its on-line applicability.

In order to provide on-line recursive algorithms for CA^K-identifiable structures, we propose several new algorithms which can be thought of as different on-line versions of the above-mentioned identification scheme. These on-line algorithms are grouped into five disjoint classes according to their structural characteristcs.

The plan and content of the paper are as follows. The multivariable model under consideration is described in section 2.1 as well as the assumptions concerning the stochastic environment. Some basic results regarding identifiable structures and the explicit parameter lemma (according to Kubrusly(1978-b)) are briefly summarized in sections 2.2 and 2.3 since such concepts represent the key idea behind the approach used. In section 3 the recursive on-line parameter identification schemes (obtained by using stochastic approximation algorithms) are developed as approximate versions of the original "off-line" scheme which converges with probability one. Illustrative examples including a comparison with a previous single-variable approach proposed by Saridis and Stein (1968) are presented in section 4. The performance of the algorithms is analysed in section 5.

2. MODEL DESCRIPTION AND AUXILIARY RESULTS

2.1. Model description: Let $\mathcal{M}(\mathbb{R}^m,\mathbb{R}^n)$ denote the linear space of all matrices n by m, and consider a discrete-time dynamical system whose evolution is governed by the following time-invariant linear difference equation.

$$x(i+1) = A\,x(i) + B\,w(i);\ x(0) = x_o, \qquad (1)$$

where $\{x(i);\ i=0,1,2,\ldots\}$ is an $\mathbb{R}^n$-valued state sequence such that x_o is a second-order random vector, $\{w(i);i=0,1,2,\ldots\}$ is taken to be an $\mathbb{R}^p$-valued second-order random sequence with $p\leq n$, and $A \in \mathcal{M}(\mathbb{R}^n,\mathbb{R}^n)$, $B \in \mathcal{M}(\mathbb{R}^p,\mathbb{R}^n)$ are the system and input matrices, respectively. Let $\{z(i);\ i=0,1,2,\ldots\}$ be an $\mathbb{R}^m$-valued observation sequence, with $m\leq n$, described by the following measurement equation

$$z(i) = C\,x(i) + v(i) \qquad (2)$$

where the noise observation $\{v(i);i=0,1,2\ldots,\}$ is taken to be an $\mathbb{R}^m$-valued second-order random sequence, and $C \in \mathcal{M}(\mathbb{R}^n,\mathbb{R}^m)$ is the observation matrix. Furthermore, let $\{w(i)\}$ and $\{v(i)\}$ be wide-sense stationary random sequences satisfying the following assumptions

(A-1): $E\{w(i)\ v^*(j)\}=0 \quad \forall\ i,j = 0,1,2,\ldots$

(A-2): $E\{w(i)\ w^*(j)\}=R_w\delta(i-j)$

(A-3): $E\{v(i)\ v^*(j)\}=R_v(|i-j|)$

where E and the superscript * denote expectation and transpose as usual, $\delta(i-j)$ is the "Kronecker delta", $R_w \in \mathcal{M}(\mathbb{R}^p,\mathbb{R}^p)$ is symmetric positive-definite, and $R_v(k)$ is a symmetric positive-semidefinite matrix in $\mathcal{M}(\mathbb{R}^m,\mathbb{R}^m)$ which is supposed to be known for each $k=0,1,2,\ldots$

2.2. A definition of identifiable structures: Consider the standard notation (A,B,C) for the above described state-space representation and define

$$\Theta_k = CA^k \in \mathcal{M}(\mathbb{R}^n,\mathbb{R}^m)\ ;\ k=0,1,2,\ldots \tag{3}$$

$$\Omega_q = \left[\Theta_0^*\ \Theta_1^*\ldots\Theta_{q-1}^*\right]^* \in \mathcal{M}(\mathbb{R}^n,\mathbb{R}^{qm});\ q=1,2,\ldots \tag{4}$$

$$Y_k = \Theta_{k-1}\ B \in \mathcal{M}(\mathbb{R}^p,\mathbb{R}^m)\ ;\ k=1,2,\ldots \tag{5}$$

where Ω_q is the observability matrix when q=n, and $\{Y_k{}^q;k=1,2,\ldots\}$ are Markov parameters, regarding an arbitrary representation (A,B,C). Now assume that:

(i) the pair (A,B) is controllable,
(ii) the pair (A,C) is observable,
(iii) the system matrix A is asymptotically stable, and
(iv) there exists an integer $q \geq n/m$ such that the matrix Ω_q has at least one completely determined left inverse P_q.

Let q_o be the least integer satisfying (iv) and $K \geq q_o-1$ a positive integer defined for $q_o>1$. According to Kubrusly(1978-b) we have:

(a) (A,B,C) has a CA^K-identifiable structure if (v-a) $q_o>1$ and the Markov parameters Y_k are known for each $k=1,2,\ldots,K$.

(b) (A,B,C) has a CA^∞-identifiable structure if (v-b) either $q_o>1$ and $Y_k=0$ for each $k=1,2,\ldots,q_o-1$, or $q_o=1$.

(c) (A,B,C) has an iterative CA^k-identifiable structure if (v-c) $q_o>1$ and both B and Θ_{k-1}, $k=1,2,\ldots,q_o-1$, are known matrices.

2.3. The explicit parameter lemma: Regarding the system described in (1), (2), let $q \geq n/m$ be any integer such that Ω_q in (4) is left inversible[1]. Now consider the following random vectors in $\mathbb{R}^{qm}$ for each $i=0,1,2,\ldots$

$$w'(i) = (w_0(i)\ ,\ w_1(i),\ldots,w_{q-1}(i))$$

$$v'(i) = (v(i),\ v(i+1),\ldots,v(i+q-1))$$

$$z'(i) = (z(i),\ z(i+1),\ldots,z(i+q-1))$$

where $v(i)$, $z(i)$ (as in (2)), and

$$w_k(i) = \begin{cases} 0 & ;\ \text{if}\quad k=0 \\ \sum_{j=1}^{k} Y_j\ w(i+k-j) & ;\ \text{if}\ ,\ k=1,2,\ldots \end{cases}$$

are random vectors in $\mathbb{R}^m$, with Y_j as in (5)
Associated with such an integer q, let $P_q \in \mathcal{M}(\mathbb{R}^{qm},\mathbb{R}^n)$ be a left inverse of Ω_q and define for each $i,k=0,1,2,\ldots$

$$y(i)=P_q z'(i), \qquad \text{in } \mathbb{R}^n, \tag{6}$$

$$N_k=E\{z(i+k)y^*(i)\}\quad , \tag{7}$$

$$T_k=(E\{w_k(i)w'^*(i)\}+E\{v(i+k)v'^*(i)\})P_q^*\ , \tag{8}$$

$$M=E\{y(i)y^*(i)\}=\left[N_0^*\ N_1^*\ldots N_{q-1}^*\right]P_q^*\quad , \tag{9}$$

$$R=P_q(E\{w'(i)w'^*(i)\}+E\{v'(i)v'^*(i)\})P_q^*= \\ =\left[T_0^*\ T_1^*\ldots T_{q-1}^*\right]P_q^*\quad , \tag{10}$$

where $N_k,T_k \in \mathcal{M}(\mathbb{R}^n,\mathbb{R}^m)$ and $M,R \in \mathcal{M}(\mathbb{R}^n,\mathbb{R}^n)$. The following two results establish procedures for expressing the unknown parameter $\Theta_k=CA^k$ in terms of quantities which are completely known up to second-order moments of the observation sequence $\{z(i)\}$. For proof the reader is referred to Kubrusly (1978-b).

Lemma (L-1): Consider the system modelled in (1), (2) and suppose it operates in a wide-sense stationary stochastic environment satisfying assumptions (A-1)-(A-3).

(a) if (A,B,C) describes an observable controllable and asymptotically stable system, which is assumed to have reached the steady state, then Θ_k in (3) can be written as follows:

$$\Theta_k = S_k\ Q^{-1}\quad ;\quad k = 0,1,2,\ldots \tag{11}$$

where

$$S_k = N_k - T_k \in \mathcal{M}(\mathbb{R}^n,\mathbb{R}^m) \tag{12}$$

$$Q = M - R \in \mathcal{M}(\mathbb{R}^n,\mathbb{R}^n) \tag{13}$$

with N_k, T_k, M, and R as defined above for any integer $q \geq n/m$ such that Ω_q is left inversible.

(b) Moreover, let $q=q_o$ and $K \geq q_o-1$ be integers as defined in section 2.2. If either (A,B,C) has a CA^K-identifiable structure and the correlation matrix R_w is supposed to be known, or (A,B,C) has a CA^∞-identifiable structure; then T_k in (8) is completely

[1] Observability is a sufficient condition for the existence of a such q (cf. Kubrusly (1978-b))

determined (i.e, it does not depend on any unknown parameter) for each $k=0,1,\ldots,K$, or for all k, respectively; what implies that R in (10) is also a known quantity.

Lemma (L.2): Consider the model described in section 2.1 with known disturbance correlation matrix R_w. Suppose (A,B,C) has an iterative CA^k-identifiable structure, and let the quantities in (6)-(10) be associated with the integer q_o defined in section 2.2.

(a) The parameter Θ_k in (11) is fully determined up to M for each $k=0,1,\ldots,q_o-1$.

(b) Let K_o be any integer such that $K_o \geq q_o$. If M and $\{N_k \,;\, k = q_o, q_o+1,\ldots,K_o\}$ are available, then the sequence $\{\Theta_k; k=q_o,q_o+1,\ldots,K_o\}$ can be recursively identified according to the following algorithm.

$$Y_k = \Theta_{k-1} B$$

$$\psi_k = \left[R_v(k) \;\; Y_k R_w Y_1^* \;\cdots\; Y_k R_w Y_{q_o-1}^* \right]$$

$$\Sigma_k = \Sigma_{k-1}\, \Phi + \psi_k$$

$$T_k = \Sigma_k \;\; P_{q_o}^*$$

$$S_k = N_k - T_k$$

$$\Theta_k = S_k\, Q^{-1}$$

with known initial conditions given by

$$\Theta_{q_o-1} = S_{q_o-1}\, Q^{-1} = S_{q_o-1}\, [M-R]^{-1} \;,$$

$$\Sigma_{q_o-1} = E\{w_{q_o-1}(i) w'^*(i)\} + E\{v(i+q_o-1) v'^*(i)\}.$$

Here ψ_k and Σ_k are in $\mathcal{M}(\mathbb{R}^{q_o m},\mathbb{R}^m)$, and

$$\Phi = \left[\begin{array}{c|c} 0 & I_{m(q_o-1)} \\ \hline & 0 \end{array} \right] \;\varepsilon\; \mathcal{M}(\mathbb{R}^{q_o m},\mathbb{R}^{q_o m})$$

(where I_ℓ stands for the identity matrix in $\mathcal{M}(\mathbb{R}^\ell,\mathbb{R}^\ell)$).

3. IDENTIFICATION ALGORITHMS

Our aim is to develop recursive identification algorithms for multivariable linear systems, as modelled in section 2.1, with identifiable structures as defined in section 2.2. It can be shown (e.g. see Mayne (1972)) that for some canonical forms (A,B,C), with known structure but unknown coefficients, the matrix A can be completely specified in terms of the product CA^k. In other words, the identification of the unknown system matrix A can be directly obtained once $\Theta_k = CA^k$ in (11) is identified for some finite set of integers k. Under assumptions of lemmas (L-1) or (L-2) this is finally achieved when Q=M-R and S_k in (12), (13) are determined. Such unknown quantities can be recursively estimated through the following stochastic approximation algorithms

$$Q(i+1) = (1-\mu(i))Q(i) + \mu(i)\left[y(i)y^*(i) - R\right] \tag{14}$$

$$M(i+1) = (1-\mu(i))M(i) + \mu(i)y(i)y^*(i) \tag{15}$$

$$S_k(i+1) = (1-\lambda(i))S_k(i) + \lambda(i)\left[z(i+k)y^*(i) - T_k\right] \tag{16}$$

which converge in quadratic mean and with probability one to Q_k, M, and S_k respectively, provided that the standard stochastic approximation conditions (including those concerning the pre-selected real sequences $\{\mu(i)\,\varepsilon\,(0,1); i=0,1,2,\ldots\}$ and $\{\lambda(i)\,\varepsilon\,(0,1) \,;\, i=0,1,2,\ldots\}$) are fulfilled (e.g see Kubrusly (1978-a)).

Remark (R-1): It is important to notice that the above recursive algorithms can be carried out only if the matrices R and T_k are "a priori" known. According to lemmas (L-1) and (L-2), R and T_k are fully determined when the structure (A,B,C) under consideration is either CA^K, CA^∞, or iterative CA^k identifiable. ■

Remark (R-2): It can be shown (cf. Kubrusly (1978-b)) that $Q=E\{x(i)\;x^*(i)\}$ in (13), which is positive-definite since (A,B) is a controllable pair and R_w (the input disturbance correlation matrix) is positive-definite. Hence M=Q+R in (13) is positive-definite as well, since R in (10) is positive-semidefinite. It is also a simple matter to show that M(i) in (15) is symmetric positive-definite for each $i=0,1,\ldots$, once the initial condition M(0) is symmetric positive-definite. On the other hand, as far as the algorithms (14), (15) are concerned, it is also advisable that Q(i) and $[M(i)-R]$ become symmetric positive-definite at an early stage i_o. One can have $i_o=0$ by choosing appropriate symmetric positive-definite initial conditions Q(0) and M(0) in accordance with the known matrix R. ■

The following proposition provides two algorithms for identifying $\Theta_k = CA^k$, and they can be easily proved from (14)-(16) by using the same idea presented in Kubrusly and Curtain (1977). These algorithms will be referred to as "off-line" since they require a matrix inversion at each iteration.

Proposition (P-1): Define

$$\Theta_k(i) = S_k(i)\;\left[M(i) - R\right]^{-1} \tag{17}$$

$$\Theta_k'(i) = S_k(i)\;\; Q^{-1}(i) \tag{18}$$

where $Q_k(i)$, M(i), and $S_k(i)$ are as in (14)-(16), and the initial conditions $Q_k(0)$ and M(0) are properly choosen in order to ensure the existence of $Q^{-1}(i)$ and $[M(i)-R]^{-1}$ for each $i \geq i_o$, according to the above remark. Then both $\Theta_k(i)$ and $\Theta'_k(i)$ converge with probability one (w.p.1) to Θ_k as $i\to\infty$.

Proposition (P-2): The sequence $\{\Theta_k(i); i=0,1,2,\ldots\}$ in (17), which converge w.p.1 to Θ_k, can be written in the following form:

$$\Theta_k(i+1) = \frac{1-\lambda(i)}{1-\mu(i)}\Theta_k(i)\left[I-\mu(i)(I-RM^{-1}(i))y(i)y^*(i)M^{-1}(i+1)\right] + \lambda(i)\left[z(i+k)y^*(i)-T_k\right]M^{-1}(i+1)+\varepsilon(i,i+1), \quad (19)$$

where

$$M^{-1}(i+1) = \frac{1}{1-\mu(i)}\left[M^{-1}(i) - \frac{\mu(i)M^{-1}(i)y(i)y^*(i)M^{-1}(i)}{1-\mu(i)+\mu(i)y^*(i)M^{-1}(i)y(i)}\right], \quad (20)$$

$\{\mu(i) ; i=0,1,2,\ldots\}$ and $\{\lambda(i); i=0,1,2,\ldots\}$ are real sequences associated with $\{M(i)\}$ and $\{S_k(i)\}$ in (15), (16) respectively, and $\varepsilon(i,i+1)$ is given by

$$\varepsilon(i,i+1)=\Theta_k(i+1)RM^{-1}(i+1)- \frac{1-\lambda(i)}{1-\mu(i)}\Theta_k(i)RM^{-1}(i)$$
$$= \varepsilon_1(i,i+1) + \varepsilon_2(i,i+1), \quad (21)$$

with

$$\varepsilon_1(i,i+1)=\left[\Theta_k(i+1)- \frac{1-\lambda(i)}{1-\mu(i)}\Theta_k(i)\right]R\,M^{-1}(i+1),$$

$$\varepsilon_2(i,i+1)= \frac{1-\lambda(i)}{1-\mu(i)}\Theta_k(i)R\left[M^{-1}(i+1)-M^{-1}(i)\right].$$

Proof: By the matrix inversion lemma (or method of modification, cf. Householder (1964)) we have (20) from (15), which converge w.p.1 to M^{-1} as $i\to\infty$. Thus since

$$y^*(i)M^{-1}(i+1) = \frac{y^*(i)M^{-1}(i)}{1-\mu(i)+\mu(i)y^*(i)M^{-1}(i)y(i)},$$

we get by (16) and (20) that

$$S_k(i+1)M^{-1}(i+1) = \frac{1-\lambda(i)}{1-\mu(i)}S_k(i)M^{-1}(i)\left[I-\mu(i)y(i)y^*(i)M^{-1}(i+1)\right] + \lambda(i)\left[z(i+k)y^*(i)-T_k\right]M^{-1}(i+1).$$

But according to (17) we have

$S_k(i)M^{-1}(i) = \Theta_k(i)\left[I - R\,M^{-1}(i)\right]$. Then

$$\Theta_k(i+1)\left[I-R\,M^{-1}(i+1)\right] = \frac{1-\lambda(i)}{1-\mu(i)}\Theta_k(i)\left[I-RM^{-1}(i)\right]\left[I-\mu(i)y(i)y^*(i)M^{-1}(i+1)\right] + \lambda(i)\left[z(i+k)y^*(i)-T_k\right]M^{-1}(i+1),$$

and the desired result follows by a simple algebraic manipulation. □

The algorithm for Θ_k as written in (19) avoids the matrix inversion at each iteration required in (17). On the other hand the expression in (19) is not a recursive one, since $\varepsilon(i,i+1)$ depends on $\Theta_k(i+1)$. However, motivated by the particular form in which $\varepsilon(i,i+1) = \varepsilon_1(i,i+1) + \varepsilon_2(i,i+1)$ in (21) depends on the convergent (w.p.1) random sequences $\{\Theta_k(i)\}$ and $\{M^{-1}(i)\}$, we propose several modified versions for the algorithm in (19). These new recursive identification algorithms, which we refer to as on-line (as opposite to those in (17), (18)), are obtained from (19) by using the following approximations for $\varepsilon(i,i+1)$.

Algorithm (AL-1): $\varepsilon(i,i+1) \simeq 0$

Algorithm (AL-2): $\varepsilon(i,i+1) \simeq \varepsilon(i,i)$,

Algorithm (AL-3): $\varepsilon_1(i,i+1) \simeq 0$.

Algorithm (AL-4): $\varepsilon_1(i,i+1) \simeq \varepsilon_1(i,i)$.

Algorithm (AL-5): $\varepsilon_1(i,i+1) \simeq \varepsilon_1(i,i+1)\Big|_{\Theta_k(i+1)=\Theta_k(i)}$

Algorithm (AL-6): $\varepsilon(i,i+1) \simeq \varepsilon(i-1,i)$.

Algorithm (AL-7): $\varepsilon(i,i+1) \simeq \varepsilon(i,i+1)\Big|_{\substack{\Theta_k(i+1)=\Theta_k(i)\\ \Theta_k(i)=\Theta_k(i-1)}}$

Algorithm (AL-8): $\varepsilon_1(i,i+1) \simeq \varepsilon_1(i-1,i)$.

Algorithm (AL-9): $\varepsilon_1(i,i+1) \simeq \varepsilon_1(i,i+1)\Big|_{\substack{\Theta_k(i+1)=\Theta_k(i)\\ \Theta_k(i)=\Theta_k(i-1)}}$

Our goal here has been to provide suitable approximations for $\varepsilon(i,i+1)$ that do not depend on $\Theta_k(i+1)$. The above on-line recursive algorithms can be grouped into five classes according to the following rule: "if $\lambda(i)=\mu(i)$ for all i, then each class supplies an unique approximate version of (19)". So we have a classification for the on-line identification algorithms as follows:

Class	Algorithms	Remraks
I	(AL-1), (AL-2)	Approximations of the type $\varepsilon(i,i+1)\simeq 0$.
II	(AL-3), (AL-4), (AL-5)	Approximatons of the type $\varepsilon_1(i,i+1)\simeq 0$
III-a	(AL-6)	Approximations of the type $\varepsilon(i,i+1) \simeq \varepsilon(i-1,i)$ introducing a new delay characterized by the term $\Theta_k(i-1)$.
III-b	(AL-7), (AL-8)	
III-c	(AL-9)	

Remark (R-3): By using the matrix inversion lemma we got a recursive algorithm for M^{-1} in (20) from the algorithm for M in (15). The same technique can not be directly applied in order to provide a recursive algorithm for Q^{-1} (or $[M-R]^{-1}$) from (14) (or (15)) that does not require any matrix inversion; and this is due to the presence of the matrix R in (14) (an exception is made for the rather trivial case of $R=rr^*$, $r\in\mathbb{R}^n$). In this sense, the algorithm in (18) (or (17)) can not be converted into "on-line" without using approximations as above. ■

Remark (R-4): By wide-sense stationarity and time-invariability assumptions, the same quanties in (6)-(10) could have been defined by using any fixed integer time-shift ℓ. Hence $y(i)$ and $z(i+k)$ appearing in the algorithms (14)-(16), (19), (20) can be replaced by $y(i+\ell)$ and $z(i+k+\ell)$, respectively. ■

4. ILLUSTRATATIVE EXAMPLES

Consider the linear model described in section 2.1. Our first example investigates a particular single-variable version for the algorithms proposed in the preceding section. This is done in order to provide a comparison of effectiveness of different approaches, since the literature in system identification by stochastic approximation (e.g. see Saridis (1974)) is mainly concerned with single-variable (i.e, m=p=1) system. Next we consider two exemples for identifying multivariable systems with m=p=2. In both cases we have used well-known (asymptotically stable) canonical forms of the type discussed in Mayne (1972); and the following general assumptions for system simulation and identification were made: (1) the disturbances $\{w(i)\}$ and $\{v(i)\}$ were taken to be independent and normally distributed with zero mean such that $R_v(|i-j|) = \sigma_v^2 I_m \delta(i-j)$ and $R_w = \sigma_w^2 I_p$ in (A-2), (A-3) with $\sigma_w^2 = 1$ and $\sigma_v^2 = 0.25$; (2) $\lambda(i)=(i+0.5)^{-1}$, $\mu(i)=(i+1)^{-1}$; and (3) the square matrix-valued initial conditions were taken to be identity matrices, and null initial conditions were used otherwise.

Example (E-1) - Single-variable system: Let

$$A = \left[\begin{array}{c|c} 0 & I_{n-1} \\ \hline \multicolumn{2}{c}{a^*} \end{array}\right] \in \mathcal{M}(\mathbb{R}^n,\mathbb{R}^n), \quad a \in \mathbb{R}^n;$$

$$C^*=c=(1,0,\dots,0) \in \mathbb{R}^n; \quad B=b=(0,\dots,0,b_n)\neq 0 \in \mathbb{R}^n.$$

With $q_o=n$ and $P_n=I_n$ it is immediate to verify that (A,B,C) has a CA^∞-identifiable structure (if b_n is "a priori" known, then (A,B,C) has also a CA^n and an iterative CA^k-identifiable structure). Note that, in such a case and under the above assumptions, we have $T_n=0$ and $R=\sigma_v^2 I_n$ in (8), (10) with $q=q_o=n$. The identification of the matrix A (i.e, the identification of $\Theta_n=c^* A^n=a^*$) is then achieved by one of the algorithms in (17), (18), (AL-1)-(AL-9) with k=n. We shall also carry out, for comparison purposes, a previous on-line algorithm proposed by Saridi and Stein (1968) for identifying $\Theta_n=a^*$. Under the preceding assumptions and recalling remar (R-4) it can be written in following form:

$$\hat{\Theta}(i+n+1)=\hat{\Theta}(i)\Big[I-(y(i)y^*(i)-R)M'(i+n+1)\Big] + z(i+n)y^*(i)M'(i+n+1) \qquad (22)$$

where

$$M'(i+n+1) = M'(i) - \frac{M'(i)y(i)y^*(i)M'(i)}{1 + y^*(i)M'(i)y(i)} .$$

It is worth noting that the above algorithm behaves like those in class II. Actually, under the preceding assumptions, the algorithm (AL-3) can be written, after some algebric manipulation, as

$$\Theta_n(i+1) = \frac{1-\lambda(i)}{1-\mu(i)}\,\Theta_n(i)\Big[I-(y(i)y^*(i)-R)\mu(i)M^{-1}(i+1)\Big] + z(i+n)y^*(i)\lambda(i)M^{-1}(i+1). \qquad (23$$

With $\lambda(i) = \mu(i)$, for all i, the algorithms ir (22), (23) have essentially the same structure. Indeed the simulated results of (22) and those algorithms in class II are basical the same. Figures 1 to 3 show the evolutior of the square error $\tilde{a}(i)=||a(i)-a||^2$ where $a(i)=\Theta_n'(i)$, $a(i)=\Theta_n(i)$, and $a(i)=\hat{\Theta}(i)$ accor ing to the algorithms "off-line" in (18), on-line in (AL-1)-(AL-9), and on-line in (22), respectively. The following numerical values were used for system simulation: n=4 $a=(-0.656, 0.784, -0.18, 1)$, $b_4=1$.

Example (E-2) - Multivariable CA^∞-identifiable structure:

$$A=\left[\begin{array}{cc|cc} 0 & 1 & \multicolumn{2}{c}{0} \\ a_1 & a_2 & & \\ \hline 0 & 0 & 0 & 1 \\ a_3 & a_4 & a_5 & a_6 \end{array}\right], B=\begin{bmatrix} 0 & 0 \\ b_1 & b_2 \\ 0 & 0 \\ b_3 & b_4 \end{bmatrix}, C=\begin{bmatrix} 1&0&0&0 \\ 0&0&1&0 \end{bmatrix} .$$

Under controllability assumption it is a simple matter to show that (A,B,C) has a CA^∞-identifiable structure with $q_o = 2$ and $P_2 = \Omega_2$(if $\{b_i;\ i=1,2,3,4\}$ are "a priori" known, then (A,B,C) has also a CA^2 and an iterative CA^k-identifiable structure). Sir ce A is completely determined by

$$\Theta_2 = CA^2 = \begin{bmatrix} a_1 & a_2 & 0 & 0 \\ a_3 & a_4 & a_5 & a_6 \end{bmatrix}$$

its "off-line" and on-line identification can be carried out through algorithms $\Theta_2'(i)$ in (17), (18) and $\Theta_2(i)$ in (AL-1)-(AL-9), respectively, where $T_2 = 0$ and $R = \sigma_v^2 I_4$ in (8),(10) with $q=q_o=2$. Figure 4 shows the evolution of the square error $\tilde{a}(i) = ||a(i) - a||^2$ where $a(i)$ in $\mathbb{R}^6$ is obtained from $\Theta_2'(i)$ in (18) and $\Theta_2(i)$ in (AL-8) as $a=(a_1,\ldots,a_6) \in \mathbb{R}^6$ is naturally obtained from Θ_2. The following numerical values were used: $a=(-0.1, 0.8, 0.3, 0.4, -0.2, 0.5)$, $b_1=b_4=1$, and $b_2=b_3=0$.

Example (E-3) - Multivariable iterative CA^k-identifiable structure:

$$A = \left[\begin{array}{cc|ccc} 0 & 1 & & 0 & \\ a_1 & a_2 & & & \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ a_3 & a_4 & a_5 & a_6 & a_7 \end{array}\right], \quad B = \begin{bmatrix} 0 & 0 \\ b_1 & b_2 \\ 0 & 0 \\ 0 & 0 \\ b_3 & b_4 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

where $\{b_i;\ i=1,2,3,4\}$ are known coefficients satisfying the controllability assumption. It is easily verified that, with $q_o=3$ and $P_3 = \Omega_3^*|_{a_1=a_2=0}$, (A,B,C) has a CA^2 and an iterative CA^k-identifiable structure, but not a CA^3 or CA^∞-identifiable structure. In other words, it can be shown that $Y_1=0$ and Y_2 is entirely known in terms of B, what implies that T_2 in (8) and R in (10) are also known quantities. Actually, in this particular example, $R=\sigma_v^2 I_5$ and T_2 is a sparse matrix whose only non-zero entry is σ_v^2 at the last position. On the other hand Y_3 depends on a_1,a_2, and so does T_3 in (8). Since

$$\Theta_2 = CA^2 = \begin{bmatrix} a_1 & a_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\Theta_3 = CA^3 = \begin{bmatrix} a_1a_2 & a_1+a_2^2 & 0 & 0 & 0 \\ a_3 & a_4 & a_5 & a_6 & a_7 \end{bmatrix},$$

the complete identification of the system matrix A can be achieved by the following procedure:

1st STEP: Identify Θ_2 by using one of the algorithms $\Theta_2(i)$ in (17),(18), (AL-1) - (AL-9), since T_2 and R are known quantities.

2nd STEP: Determine $Y_3=\Theta_2 B$ and so T_3, according to the recursive scheme of lemma (L-2) or by definition (8), and then identity Θ_3 through one of the algorithms $\Theta_3(i)$ in (17),(18),(AL-1)-(AL-9).

Using the following numerical values for system simulation, $a=(a_1,\ldots a_7) = (0.3, 0.1, 0.7, 0.6, 0.2, -0.4, 0.5) \in \mathbb{R}^7$, $b_1 = b_4 = 1$, $b_2 = b_3 = 0$, the evolution of the square error $\tilde{a}(i)=||a(i)-a||^2$ is shown in figure 4. Here $a(i)=(\hat{a}_1,\hat{a}_2,a_3(i),\ldots,a_7(i))$ in $\mathbb{R}^7$ was obtained as follows: $(\hat{a}_1, \hat{a}_2) = (0{,}296, 0.096) \in \mathbb{R}^2$ is the result of 1st step through $\Theta_2(i)$ in (AL-8) up to the 10 000th iteration where $||(\hat{a}_1,\hat{a}_2) - (a_1, a_2)||^2 = 4.1\text{x}10^{-5}$, and $\{a_j(i)\ ;\ j=3,\ldots,7\}$ is the second row of $\Theta_3(i)$ in (AL-8) at the 2nd step.

5. CONCLUDING REMARKS

Nine recursive on-line identification schemes were proposed for CA^k-identifiable structures operating in a stochastic environment. The problem was approached by using stochastic approximation algorithms, which yielded straightforward on-line identification procedures applicable to multivariable systems.

The identifiable structures defined in section 2.2 enabled us to write down the unknown system parameters in an explicit form, $\Theta_k=Q^{-1} S_k$, as in lemma (L-1) and (L-2). In this way the system identification problem could be reduced to one of determining second-order moments of the observation process only (i.e, the matrices N_k and M in (7), (9)). By using some parametric estimation technique to finally solve the identification problem, estimates of the state correlation matrix Q(i) or M(i)-R in (14), (15) must be inverted at each iteration as in proposition (P-1). In order to avoid such a repetitive matrix inversion (and so to derive on-line identification schemes) the algorithms (AL-1) to (AL-9) were obtained by approximating in a recursive way the non-recursive version presented in proposition (P-2).

The major advantage of such algorithms over that proposed by Saridis and Stein (1968) is the applicability for multivariable systems. Another advantege of the algorithms proposed here concerns convergence speed: Although they have presented the same convergence rate and structural simplicity of the algorithm in (22) (cf. example (E-1)), they are much more amenable to accelerated versions by choosing a pair of optimal sequences $\{\lambda(i)\}$ and $\{\mu(i)\}$, which can be reflected by less time-consuming computer programmes.

Other features that have been presently accomplished are: (1) The identification method requires measurements only on the noisy observation sequence $\{z(i)\}$. This in fact may be very important in real cases where the input process is inaccessible. (2)Both B and R_w (the input disturbance parameters) do not need to known in case of CA^∞-identifiable structures. Examples (E-1) and (E-2) considered just particular cases of such structures, which were also shown to be the natural representation arising from standard discretization procedures for linear systems governed by partial differential equations (cf. Kubrusly (1978-c)).(3) The input disturbance and the observation noise sequences, $\{w(i)\}$ and $\{v(i)\}$, were not required to be independent; and no specific type of probability distribution was imposed.

Finally it is worth remarking on three points concerning the illustrative examples: (1) It was shown experimentally in (E-1) and (E-2) that the on-line algorithms (AL-1)-(AL-9) presented the same convergence rate of the "off-line" one in (18),whose convergence is proved w.p.1.(2) Example (E-3) differed considerably from the previous ones since CA^∞ (or even CA^3) identifiability was not accomplished there,thus becoming necessary to use the iterative CA^k-identifiability property through lemma (L-2). (3) In example (E-3) the identification procedure may not be appropriate for on-line applications, even using one of the algorithms (AL-1)-(AL-9), since it comprises two consecutive steps of recursive schemes. Investigations regarding the joint evolution of these two steps, where estimates $T_3(i)$ are directly obtained from each $\Theta_2(i)$, is a topic for further research.This can lead to on-line procedures where recursive algorithms for $\Theta_2(i)$ and $\Theta_3(i)$ are carried out simultaneously.

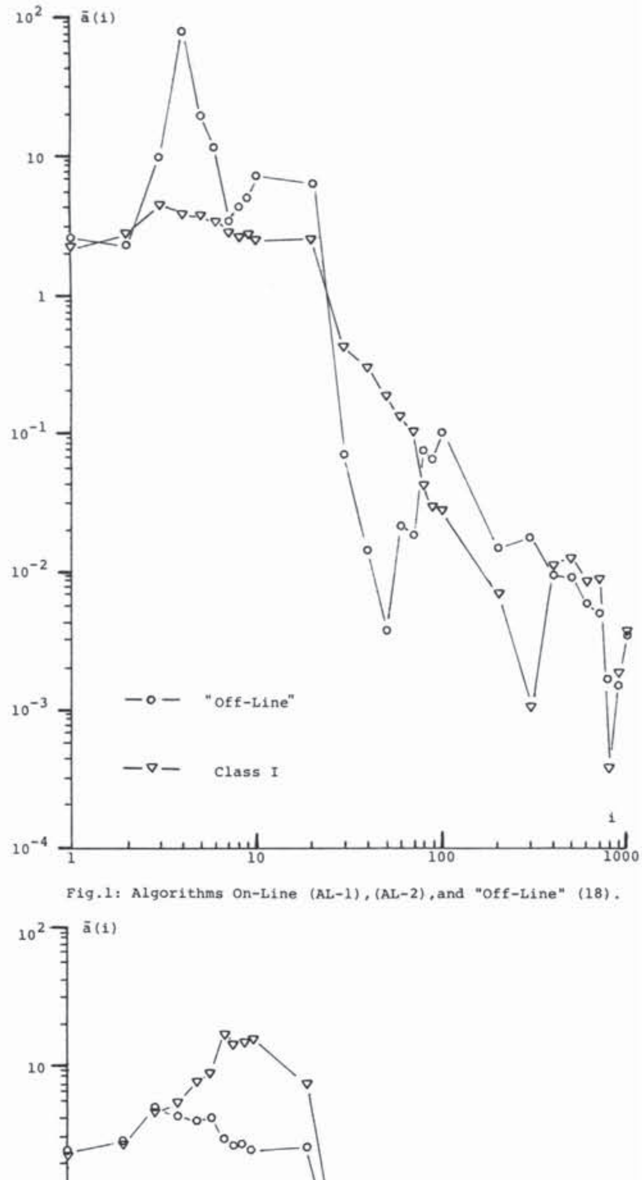

Fig.1: Algorithms On-Line (AL-1),(AL-2),and "Off-Line" (18).

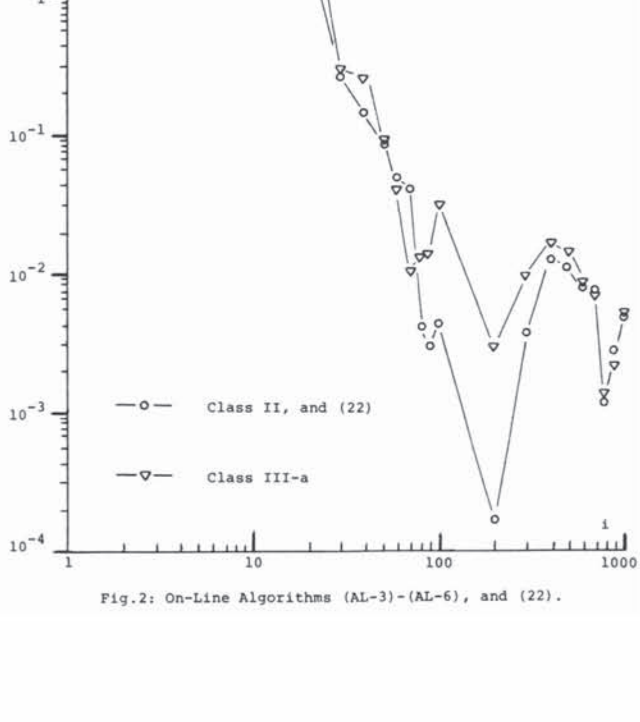

Fig.2: On-Line Algorithms (AL-3)-(AL-6), and (22).

ACKNOWLEDGMENT

The authors wish to thank Dr. A. C. Gadelha Vieira for his helpful comments.

REFERENCES

Householder, A.S. (1964):-"The Theory of Matrices in Numerical Analysis" , (Blaisdell).

Kubrusly, C. S. (1978-a):- "Applied Stochastic Approximation Algorithms in Hilbert Spaces", Int. J. Control,Vol. 28, Nº 1, pp. 23-31.

Kubrusly, C. S. (1978-b):-"On Identifiable

Structures in a Stochastic Environment", Com. GSC-09/78, Elect. Eng. Dep., Catholic Univ., Rio de Janeiro.

Kubrusly, C.S.(1978-c):-"Linear DPS Identification in a Stochastic Environment", Com.GSC 10/78, Elect. Eng. Dep.,Catholic Univ., Rio de Janeiro.

Kubrusly, C. S. and Curtain, R.F. (1977): "Identification of Noisy Distributed Parameter Systems Using Stochastic Approximation", Int. J. Control, Vol. 25, Nº 3, pp. 441-455.

Mayne, D. Q. (1972):-"Canonical Model for Identification of Linear Multivariable Systems", IEEE Trans. Autom. Control, Vol. AC-17, Nº 5, pp. 728-729.

Saridis, G. N. (1974):-"Stochastic Approximation Methods for Identification and Control - A Survey", IEEE Trans. Autom. Control, Vol. AC-19, Nº 6, pp. 798-809.

Saridis, G. N. Stein G. (1968)-"Stochastic Approximation Algorithms for Linear Discrete-Time System Identification", IEEE Trans. Autom. Control, Vol. AC-13, Nº 5, pp. 515-523.

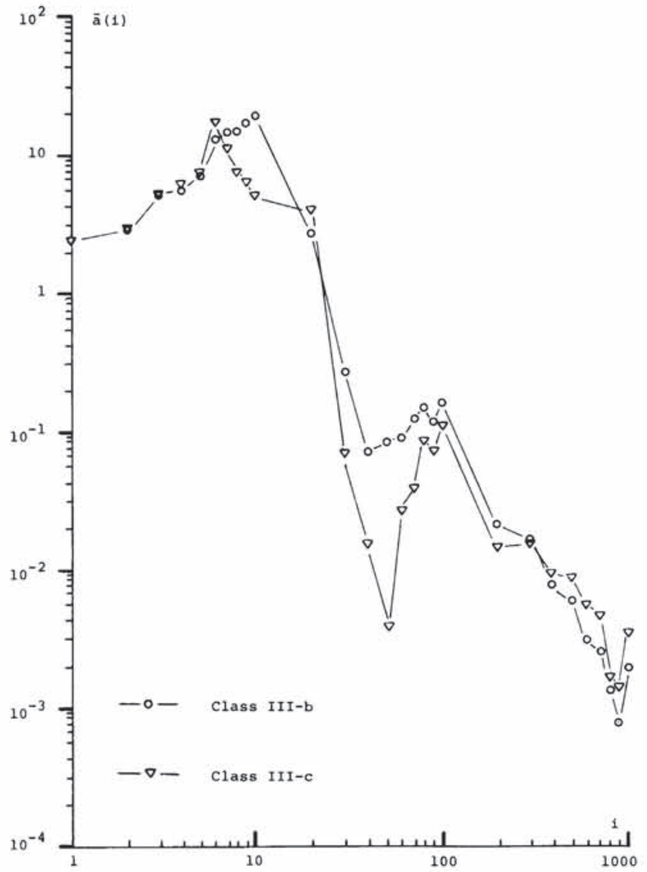

Fig.3: On-Line Algorithms (AL-7)-(AL-9).

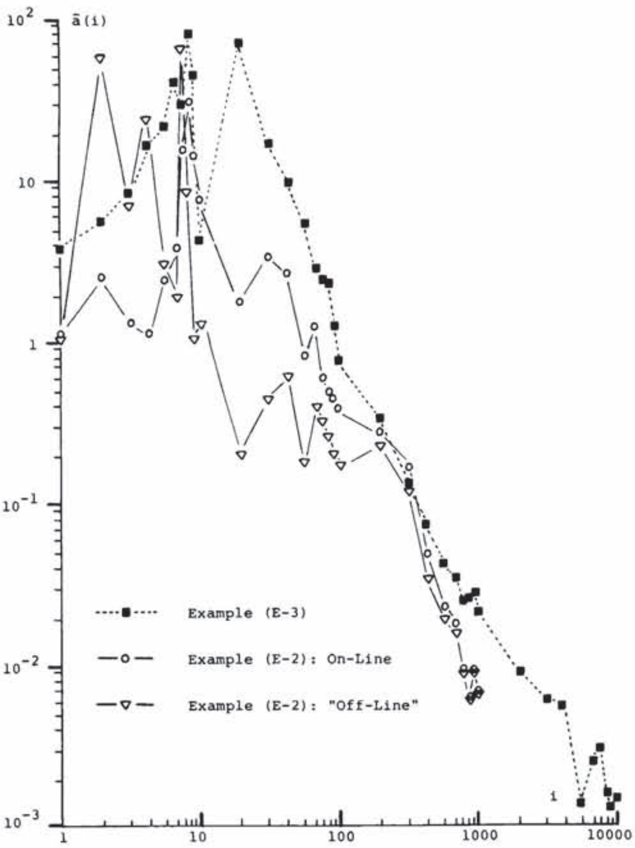

Fig.4: Multivariable Systems - Algorithms (AL-8) and (18).