Mendel and Washburn results [5]. When $\phi_{22} \neq I$ and $w_2(j) \neq 0$ an algorithm similar to, but computationally simpler than Tanaka's results [17].

CONCLUSIONS

A new multistage approach to linear estimation has been developed. The optimal estimation problem is decomposed into two or more stages. The first stage is a Kalman filter with nominal values of process noise variance and initial state variance. The second stage Kalman filter, which contains any remaining process noise and initial state uncertainty, has a new system matrix and uses the innovations from the first stage Kalman filter as measurements. The two filters are combined to provide the overall optimal estimate.

REFERENCES

[1] D. G. Lainiotis, "Partitioned estimation algorithms, II: Linear estimation," *J. Inform. Sci.*, vol. 7, no. 3, pp. 317-340, 1974.

[2] ——, "A unifying framework for linear estimation: Generalized partitioned algorithms," *J. Inform. Sci.*, vol. 10, pp. 243-278, Apr. 1976.

[3] D. G. Lainiotis and K. S. Govindaraj, "Generalized partitioned algorithms," *J. Inform. Sci.*, vol. 15, pp. 169-185, 1978.

[4] B. Friedlander, T. Kailath, and L. Ljung, "Scattering theory and linear least squares estimation—II: Discrete-time problems," *J. Franklin Inst.*, vol. 301, no. 1 and 2, pp. 71-82, Jan.-Feb. 1976.

[5] J. M. Mendel and H. D. Washburn, "Multistage estimation of bias states in linear systems," *Int. J. Contr.*, vol. 28, no. 4, pp. 511-524, 1978.

[6] B. D. O. Anderson, "Stability properties of Kalman-Bucy filters," *J. Franklin Inst.*, vol. 291, pp. 137-144, Feb. 1971.

[7] D. Andrisani and C. F. Gau, "Recursive partitioning estimation," presented at the IEEE Mediterranean Electrotech. Conf., Athens, Greece, May 24-26, 1983.

[8] K. Watanabe, "Position and velocity estimation via pseudolinear partitioned tracking filter," personal communication, Apr. 6, 1983.

[9] J. M. Mendel, *Discrete Techniques of Parameter Estimation.* New York: Marcel Dekker, 1975.

[10] J. M. Mendel, "Multistage least squares parameter estimators," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 775-785, Dec. 1975.

[11] R. A. Singer and P. A. Frost, "On the relative performance of the Kalman and Wiener filters," *IEEE Trans. Automat. Contr.*, vol. AC-14, pp. 390-394, Aug. 1969.

[12] A. S. Willsky, M. G. Bello, D. A. Castanon, B. C. Levy, and G. C. Verghese, "Combining and updating of local estimates and regional maps along sets of one-dimensional tracks," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 799-831, Aug. 1982.

[13] M. E. Womble and J. E. Potter, "A prefiltering version of the Kalman filter with new numerical integration formulas for Riccati equations," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 378-381, June 1975.

[14] B. J. Eulrich, D. Andrisani, and D. G. Lainiotis, "Partitioning indentification algorithms," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 521-528, June 1980.

[15] K. Watanabe, "Optimal structurally partitioned filter for undisturbable stochastic systems, Part I: Basic theory," *Int. J. Syst. Sci.*, vol. 14, no. 10, pp. 1139-1188, 1983.

[16] A. W. Warren, "Partitioned state algorithms for recursive system indentification," in *Proc. 1983 Automat. Contr. Conf.*, San Francisco, CA, June 1983, pp. 742-747.

[17] A. Tanaka, "Parallel computation in linear discrete filtering," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 573-575, Aug. 1975.

Optimal Adaptive LQG Control for Systems with Finite State Process Parameters

P. E. CAINES AND H. F. CHEN

Abstract—**The situation where a totally observed process y_t is generated by a stochastic differential equation whose parameters evolve on a finite set $\{1, \cdots N\}$ according to a stochastic differential equation is considered. The optimal control law is sought with respect to quadratic loss functions on y_t and the control u_t. The auxiliary P.D.E. technique of Hijab [6] is used together with a nonlinear filter to obtain the solution whose existence depends upon that of a smooth solution to the auxiliary P.D.E. and strong solutions to the system S.D.E. under the given control inputs.**

Manuscript received September 27, 1983; revised April 11, 1984. Paper recommended by P. R. Kumar, Past Chairman of the Stochastic Control Committee. This work was supported by the Faculty of Engineering and the Faculty of Graduate Studies and Research, McGill University, the NSERC International Scientific Exchange Programme, and NSERC under Grant A1329, while the first author was visiting McGill University.

P. E. Caines is with the Canadian Institute for Advanced Research and the Department of Electrical Engineering, McGill University, Montreal, P.Q., Canada.

H. F. Chen is with the Institute of System Science, Academia Sinica, Beijing, China.

I. INTRODUCTION AND PROBLEM STATEMENT

Consider the situation where one wishes to design a regulator for a system but one has only inexact knowledge of the system parameters. This is commonly referred to as a parameter adaptive control problem. The solutions to date to this problem may be broadly classified as 1) stabilizing adaptive regulators for deterministic systems, 2) asymptotically stabilizing and optimizing adaptive regulators for stochastic systems, and 3) optimal adaptive regulators for stochastic systems.

As a result of the recent intense research activity, there is now a vast literature on these topics. (See, e.g., the proceedings of the IFAC workshop [7] for a representative set of current papers in this area.) Most of the work on adaptive control concerns systems with constant unknown parameters; this is with the exception of the results of Xie and Evans [5], Caines [1], Caines and Chen [2], and Chen and Caines [3], [4]. The first reference falls in category 1) and the other three in categories 2) and 3).

There is strong practical motivation to obtain adaptive control results for systems whose parameters vary in some deterministic or stochastic manner since one of the primary reasons for using adaptive controllers is the fact that control system parameters often drift from their initial values.

This paper is concerned with adaptive control problems in category 3) in the case where the system parameters evolve randomly. Chen and Caines [4] is the only other work we know of in this area.

For the case of constant unknown parameters, the optimal adaptive control problems of category 3) have often been called dual control problems.

A solution of the optimal adaptive control problem for linear systems with quadratic loss function was presented by Hijab [6]. He uses a dynamic programming formulation and assumes the existence of a smooth solution to an auxiliary partial differential equation. He also implicity requires that all of the candidate controls and the resulting optimal control generate a strong solution to the system equations.

Now the adaptive regulators in classes 1) and 2) typically generate parameter estimates via some recursive estimation scheme which can at best be asymptotically consistent (see, e.g., [3]). However, the optimal adaptive regulators of Hijab [6], Chen and Caines [4], and that presented in this paper require the exact solution of a set of filtering equations for the joint vector of the process y_t and the parameter vector θ_t. (We distinguish between the "state" process y_t and the parameter process θ_t by the property that the controls only influence the former process. Of course, a true state process for the system must be equivalent to the joint quantity (y_t, θ_t).)

The contribution of this paper is as follows: we consider the situation where a totally observed process y_t is generated by a stochastic differential equation into which the control u_t enters linearly and for which the parameters evolve on a finite set $\{1, \cdots, N\}$ according to a stochastic differential equation. We take quadratic loss functions of y_t and u_t and then we seek the optimal control law minimizing the expected loss over a time interval $[0, T]$. We use the technique of Hijab [6] which employs the solution to an auxiliary partial differential equation to complete the square in an expression for the total cost-to-go. In this paper this is used in combination with the appropriate nonlinear filter for θ_t (see [8] and [9]) in order to obtain the desired optimal adaptive LQG control. The main result of this paper may be viewed as a verification theorem, and an interesting problem is to find nontrivial examples for which a function S satisfying (1)-(3) in Section III exists.

In the formulation of an adaptive control law it is important to note what *a priori* information is required for its implementation. The optimal control law derived in this paper is a function of the parameters A_t, D_t, F_t and the functions $h(\cdot, \cdot, 1), \cdots, h(\cdot, \cdot, N)$, $B.(1), \cdots, B.(N)$ appearing in the stochastic differential equations (1), (3) below. If one wishes to relax this required *priori* information, one is forced to use some further level of adaptation, and so on.

An open problem is the asymptotic behavior of the optimal control law as the terminal time tends to infinity.

We now come to the precise formulation of the problem.

Let $(\Omega, \mathfrak{F}, P)$ be a probability space with an associated nondecreasing sequence of σ-algebras $\mathfrak{F}_t \subset \mathfrak{F}$. The stochastic system is observed as an m-dimensional vector y_t subject to a stochastic differential equation

$$dy_t = h(y_t, t, \theta_t)dt + B_t(\theta_t)u_t dt + F_t dw_t, \qquad E\|y_0\|^2 < \infty \tag{1}$$

where $(w_t, \mathfrak{F}_t)$ is an l-dimensional Wiener process, u_t is an r-dimensional control process, and the parameter θ_t is a random process taking values in the abstract set $\theta = \{1, \cdots, N\}$.

(Note that in (1) h and B are assumed to be known Borel measurable functions such that

$$h : \mathbb{R}^m \times \mathbb{R} \times \theta \to \mathbb{R}^m$$

and

$$B : \mathbb{R} \times \theta \to \mathbb{R}^{mr}.)$$

Set

$$\Phi(\theta_t) = [I_{[\theta_t = 1]}, \cdots, I_{[\theta_t = N]}]^T. \tag{2}$$

The process θ_t is assumed to evolve in such a way that $\Phi(\theta_t)$ satisfies the stochastic differential equation

$$\Phi(\theta_t) = \Phi(\theta_0) + \int_0^t A_s \Phi(\theta_s)ds + m_t \tag{3}$$

where $(m_t, \mathfrak{F}_t)$ is a square integrable martingale with quadratic variation matrix $\langle m, w \rangle_t$. It is always the case that $\langle m, w \rangle_t$ has a derivative process $D_t \triangleq d/dt \langle m, w \rangle_t$ (see [8, Theorem 5.3]) but we shall assume in addition that D_t is a deterministic process. Observe that if $D_t \neq 0$, then θ_t and m_t are, in general, dependent processes.

We remark that if θ_t is a finite state Markov process with stationary transition probabilities, i.e., $(dp_t/dt) = Ap_t$, where $p_t = E\Phi(\theta_t)$, then (3) is satisfied with $D_s \equiv 0$ and A_s a constant matrix.

Introducing the notation

$$H_1(y_t, t) = [h(y_t, t, 1) \cdots h(y_t, t, N)] \tag{4}$$

$$H_2(u_t, t) = [B_t(1)u_t \cdots B_t(N)u_t] \tag{5}$$

we can rewrite (1) as

$$dy_t = [H_1(y_t, t) + H_2(u_t, t)]\Phi(\theta_t)dt + F_t dw_t. \tag{6}$$

Concerning the matrix coefficients we assume that A_s, F_s are deterministic,

$$\|A_s\| \in L^1_{[0,T]}, \|F_s\|^2 \in L^2_{[0,T]}, F_s F_s^T \geqslant \delta I, \ \delta > 0$$

and that there are constants k_1, k_2, k_3 such that

$$\|D_t\| \leqslant k_1, \ \forall t \in [0, T]$$

$$\|H_1(y_t, t)\| \leqslant k_2(1 + \|y_t\|) \tag{7}$$

$$\|B_t(i)\| \leqslant k_3, \ \forall t \in [0, T], \qquad 1 \leqslant i \leqslant N. \tag{8}$$

The set $\mathcal{U}$ of admissible closed-loop controls consists of $\{u_t\}$ under which (6) has a strong solution and for which

1) u_t is $\mathfrak{F}_t^y$ $(\triangleq \sigma\{y_s, s \leqslant t\})$-measurable; and
2) $\sup_{0 \leqslant t \leqslant T} E\|y_t\|^2 < \infty$.

The problem is to find the control $\{u_t^*\} \in \mathcal{U}$ for which

$$EJ(u^*) = \inf_{\{u_t\} \in \mathcal{U}} EJ(u) \triangleq \inf_{\{u_t\} \in \mathcal{U}} \left\{ E(y_T - y_T^*)^T Q_0 (y_T - y_T^*) + E\int_0^T (y_t - y_t^*)^T Q_1 (y_t - y_t^*)dt + E\int_0^T u_t^T Q_2 u_t dt \right\} \tag{9}$$

whenever this exists, where y_t^* is a deterministic reference signal and

$$Q_0 \geqslant 0, \ Q_1 \geqslant 0, \ Q_2 > 0.$$

II. The Infinitesimal Generator Associated with the System

Set

$$E(\Phi(\theta_t)/\mathfrak{F}_t^y) \triangleq \hat{p}_t = \begin{bmatrix} \hat{p}_t^1 \\ \vdots \\ \hat{p}_t^N \end{bmatrix}, \ P_t = \begin{bmatrix} \hat{p}_t^1 & & 0 \\ & \ddots & \\ 0 & & \hat{p}_t^N \end{bmatrix}. \tag{10}$$

By using the nonlinear filtering equation for (3), (6) (see [8], [9]),

$$\begin{aligned} \hat{p}_t = \hat{p}_0 + \int_0^t A_s \hat{p}_s ds + \int_0^t \{D_s + E(\Phi(\theta_s)\Phi^T(\theta_s)/\mathfrak{F}_s^y)H_1^T + H_2^T) \\ - \hat{p}_s \hat{p}_s^T (H_1^T + H_2^T)\}(F_s F_s^T)^{-1/2} d\bar{w}_s \\ = \hat{p}_0 + \int_0^t A_s \hat{p}_s ds + \int_0^t [D_s + (P_s - \hat{p}_s \hat{p}_s^T)(H_1^T + H_2^T)\}](F_s F_s^T)^{-1/2} d\bar{w}_s \end{aligned} \tag{11}$$

where $(\bar{w}_t, \mathfrak{F}_t^y)$ is the Wiener process given by

$$d\bar{w}_t = (F_t F_t^T)^{-1/2}[dy_t - (H_1(y_t, t) + H_2(u_t, t))\hat{p}_t dt]. \tag{12}$$

The process y_t has the innovation representation

$$dy_t = [H_1(y_t, t) + H_2(u_t, t)]\hat{p}_t dt + (F_t F_t^T)^{-1/2} d\bar{w}_t \tag{13}$$

and (12), (13) are seen to form a system of (state plus observation) diffusion equations with completely observed state $\hat{p}_t$. The joint process $(\hat{p}_t, y_t)$ has the infinitesimal generator L defined by

$$\mathcal{L} = \hat{p}_t^T A_t^T \frac{\partial}{\partial \hat{p}_t} + \hat{p}_t^T [H_1^T(y_t, t) + H_2^T(u_t, t)] \frac{\partial}{\partial y_t} + \frac{1}{2} \operatorname{tr} G_t^T \begin{bmatrix} \dfrac{\partial^2}{\partial^2 \hat{p}_t} & \dfrac{\partial^2}{\partial \hat{p}_t \partial y_t} \\ \dfrac{\partial^2}{\partial \hat{p}_t \partial y_t} & \dfrac{\partial^2}{\partial^2 y_t} \end{bmatrix} G_t \tag{14}$$

where for $\mathcal{L}(f)$ the partial differential operations act only on f, and where $(\partial/\partial \hat{p}_t)$ and $(\partial/\partial y_t)$ denote the N and m component columns of partial differential operators with entries $(\partial/\partial \hat{p}_t^i)$ and $(\partial/\partial y_t^j)$, respectively, and $(\partial^2/\partial \hat{p}^2)$ and $(\partial^2/\partial \hat{p}_t \, \partial y_t)$ are the corresponding and matrices of second-order partial differential operators

$$G_t = G_t^1(\hat{p}_t) + G_t^2(\hat{p}_t, y_t),$$

$$G_t^1(\hat{p}_t, y_t) = \begin{bmatrix} G_t^{11} \\ (F_t F_t^T)^{1/2} \end{bmatrix}, \ G_t^2(\hat{p}_t, y_t) = \begin{bmatrix} (P_t - \hat{p}_t \hat{p}_t^T) H_2^T (F_t F_t^T)^{-1/2} \\ 0 \end{bmatrix},$$

$$G_t^{11} = [D_t - \hat{p}_t \hat{p}_t^T) H_1^T](F_t F_t^T)^{-1/2}.$$

Set

$$q_t^T = [\hat{p}_t^T y_t^T]$$

and denote by a_i, $1 \leqslant i \leqslant N$ and b_{ij}, $1 \leqslant i \leqslant N$, $1 \leqslant j \leqslant N$, respectively, the column vectors and the elements of the two matrices indicated below

$$[a_i \cdots a_n] \triangleq \left[(P_t - \hat{p}_t \hat{p}_t^T) \left(\frac{\partial^2}{\partial^2 \hat{p}_t} G_t^{11} (F_t F_t^T)^{-1/2} + \frac{\partial^2}{\partial \hat{p} \partial y_t} \right) \right]^T, \tag{15}$$

$$\begin{bmatrix} b_{11} & \cdots & b_{1N} \\ \vdots & \ddots & \vdots \\ b_{N1} & \cdots & b_{NN} \end{bmatrix} \triangleq (P_t - \hat{p}_t \hat{p}_t^T) \frac{\partial^2}{\partial^2 p_t} (P_t - \hat{p}_t \hat{p}_t^T). \tag{16}$$

We note that a_i and b_{ij} do not depend explicitly on the control u. Noting that

$$\hat{p}_t^T H_2^T(u_t, t) = \hat{p}_t^T \begin{bmatrix} u_t^T B_t^T(1) \\ \vdots \\ u_t^T B_t^T(N) \end{bmatrix} = u_t^T \sum_{i=1}^N B_t^T(i) \hat{p}_t^i$$

we then have

$$\begin{aligned}
&\operatorname{tr} G_t^\tau \frac{\partial^2}{\partial^2 q_t} G_t \\
&= \operatorname{tr} (G_t^1 + G_t^2)^\tau \frac{\partial^2}{\partial^2 q_t} (G_t^1 + G_t^2) \\
&= \operatorname{tr} G_t^{1\tau} \frac{\partial^2}{\partial^2 q_t} G_t^1 + 2 \operatorname{tr} (F_t F_t^\tau)^{-1/2} H_2 (P_t - \hat{p}_t \hat{p}_t^\tau) \\
&\quad \cdot \left(\frac{\partial^2}{\partial^2 \hat{p}_t} G_t^{11} + \frac{\partial^2}{\partial \hat{p}_t \partial y_t} (F_t F_t^\tau)^{1/2} \right) \\
&\quad + \operatorname{tr} (F_t F_t^\tau)^{-1/2} H_2 (P_t - \hat{p}_t \hat{p}_t^\tau) \frac{\partial^2}{\partial^2 \hat{p}_t} (P_t - \hat{p}_t \hat{p}_t^\tau) H_2^\tau (F_t F_t^\tau)^{-1/2} \\
&= \operatorname{tr} G_t^{1\tau} \frac{\partial^2}{\partial^2 q_t} G_t^1 + 2 \operatorname{tr} H_2^\tau \Bigg[(P_t - \hat{p}_t \hat{p}_t^\tau) \\
&\quad \cdot \left(\frac{\partial^2}{\partial^2 \hat{p}_t} G_t^{11} (F_t F_t^\tau)^{-1/2} + \frac{\partial^2}{\partial \hat{p}_t \partial y_t} \right) \Bigg]^\tau \\
&\quad + \operatorname{tr} [(F_t F_t^\tau)^{-1/2} B_t(1) u_t \cdots (F_t F_t^\tau)^{-1/2} \\
&\quad \cdot B_t(N) u_t] (P_t - \hat{p}_t \hat{p}_t^\tau) \frac{\partial^2}{\partial^2 \hat{p}_t} (P_t - \hat{p}_t \hat{p}_t^\tau) \\
&\quad \cdot [(F_t F_t^\tau)^{-1/2} B_t(1) u_t \cdots (F_t F_t^\tau)^{-1/2} B_t(N) u_t]^\tau \\
&= \operatorname{tr} G_t^{1\tau} \frac{\partial^2}{\partial^2 q_t} G_t^1 + 2 u_t^\tau \sum_{i=1}^N B_t^\tau(i) a_i \\
&\quad + u_t^\tau \sum_{i,j=1}^N b_{ij} B_t^\tau(i) (F_t F_t^\tau)^{-1} B_t(j) u_t.
\end{aligned}$$

Thus, we can write the infinitesimal generator $\mathfrak{L}$ as a quadratic form in u

$$\mathfrak{L} = L + u_t^\tau \left(\sum_{i=1}^N B_t^\tau(i) \hat{p}_t^i \frac{\partial}{\partial y_t} + \sum_{i=1}^N B_t^\tau(i) a_i \right) + \frac{1}{2} u_t^\tau \sum_{i,j=1}^N b_{ij} B_t^\tau(i) (F_t F_t^\tau)^{-1} B_t(j) u_t \tag{17}$$

where

$$L = \hat{p}_t^\tau A_t^\tau \frac{\partial}{\partial \hat{p}_t} + \hat{p}_t^\tau H_1^\tau(y_t, t) \frac{\partial}{\partial y_t} + \frac{1}{2} \operatorname{tr} G_t^{1\tau} \frac{\partial^2}{\partial^2 q_t} G_t^1.$$

III. Optimal Control

The optimal stochastic control problem is solved by the dynamic programming approach. The main assumption for this is the existence of a function $S(t, y, p)$ defined on $[0, T] \times \mathbb{R}^m \times R^N$ and taking values in $\mathbb{R}$ such that:

1) $S(t, y, p)$ is continuously differentiable in t and twice continuously differentiable in y and p,

2) there are constants k_4, k_5, and k_6 such that

$$|S(t, y, p)| < k_4(1 + \|y\|^2), \left\| \frac{\partial^2 S}{\partial^2 p} \right\| \leqslant k_5, \left\| \frac{\partial^2 S(t, y, p)}{\partial p \partial y} \right\| + \left\| \frac{\partial S}{\partial y} \right\| \leqslant k_6(1 + \|y\|)$$

with $\bar{Q}_2 \triangleq Q_2 - 2Nk_3^2 k_5 \delta^{-1} I > 0$,

3) $S(t, y, p)$ is the unique strong solution of the following differential equation

$$\begin{aligned}
&\frac{\partial S}{\partial t} + p^\tau A_t^\tau \frac{\partial S}{\partial p} + p^\tau H_1^\tau(y, t) \frac{\partial S}{\partial y} + \frac{1}{2} \operatorname{tr} G_t^{1\tau}(p, y) \frac{\partial^2 S}{\partial^2 q} G_t^1(p, y) \\
&= \frac{1}{4} \left[\sum_{i=1}^N B_t^\tau(i) \left(p_t^i \frac{\partial S}{\partial y} + a_i(S) \right) \right]^\tau \\
&\quad \cdot \left[Q_2 + \frac{1}{2} \sum_{i,j=1}^N b_{ij}(S) B_t^\tau(i) (F_t F_t^\tau)^{-1} B_t(j) \right]^{-1} \\
&\quad \cdot \left[\sum_{i=1}^N B_t^\tau(i) \left(p_t^i \frac{\partial S}{\partial y} + a_i(S) \right) \right] - (y - y_t^*)^\tau Q_1 (y - y_t^*)
\end{aligned} \tag{18}$$

with

$$S(t, y, p)|_{t=T} = (y - y_T^*)^\tau Q_0 (y - y_T^*), \tag{19}$$

where the indicated matrix inverse is assumed to exist on $[0, T] \times \mathbb{R}^m \times \mathbb{R}^p$ and where $G^1(p, y)$ is obtained from $G_t^1(\hat{p}_t, y_t)$ with $\hat{p}_t$, q_t, y_t replaced by p, q, y, respectively.

Using this auxiliary function S we are now in a position to solve the problem stated in Section I.

By Ito's formula we find that up to some stopping time s, $0 < s \leqslant T$

$$\begin{aligned}
&S(s, y_s, \hat{p}_s) \\
&= S(0, y_0, \hat{p}_0) + \int_0^s \Bigg[\frac{\partial S(t, y_t, \hat{p}_t)}{\partial t} + L(S(t, y_t, \hat{p}_t)) \\
&\quad + u_t^\tau \sum_{i=1}^N B_t^\tau(i) \left(\hat{p}_t^i \frac{\partial S(t, y_t, \hat{p}_t)}{\partial y_t} + a_i(S) \right) \\
&\quad + \frac{1}{2} u_t^\tau \sum_{i,j=1}^N b_{i,j}(S) B_t^\tau(i) (F_t F_t^\tau)^{-1} B_t(j) u_t \Bigg] dt + M(s)
\end{aligned} \tag{20}$$

where $M(s)$ is a zero-mean local martingale.

By (18), (20) it follows that

$$\begin{aligned}
&S(s, y_s, \hat{p}_s) \\
&= S(0, y_0, \hat{p}_0) + \int_0^s \Bigg\{ \frac{1}{4} \left[\sum_{i=1}^N B_t^\tau(i) \left(\hat{p}_t^i \frac{\partial S}{\partial y_t} + a_i(S) \right) \right]^\tau \Bigg[Q_2 \\
&\quad + \frac{1}{2} \sum_{i,j=1}^N b_{ij}(S) B_t^\tau(i) (F_t F_t^\tau)^{-1} B_t(j) \Bigg]^{-1} \left[\sum_{i=1}^N B_t^\tau(i) \left(\hat{p}_t^i \frac{\partial S}{\partial y_t} + a_i(S) \right) \right] \\
&\quad - (y_t - y_t^*)^\tau Q_1 (y_t - y_t^*) + u_t^\tau \sum_{i=1}^N B_t^\tau(i) \left(\hat{p}_t^i \frac{\partial S}{\partial y_t} + a_i(S) \right) \\
&\quad + \frac{1}{2} u_t^\tau \sum_{i,j=1}^N b_{ij}(S) B_t^\tau(i) (F_t F_t^\tau)^{-1} B_t(j) u_t \Bigg\} dt + M(s)
\end{aligned}$$

where $s = s_D \wedge T$, where s_D is the first exit time of y_s from some compact set $D \subset \mathbb{R}^m$. Since $M(s)$ has zero mean

$$\begin{aligned}
&EJ(u) \\
&= E(y_T - y_T^*)^\tau Q_0 (y_T - y_T^*) - ES(s, y_s, \hat{p}_s) + ES(0, y_0, \hat{p}_0) \\
&\quad + E \int_0^s \Bigg\{ u_t + \frac{1}{2} \left(Q_2 + \frac{1}{2} \sum_{i,j=1}^N b_{ij}(S) B_t^\tau(i) (F_t F_t^\tau)^{-1} B_t(j) \right)^{-1} \\
&\quad \cdot \left[\sum_{i=1}^N B_t^\tau(i) \left(\hat{p}_t^i \frac{\partial S}{\partial y_t} + a_i(S) \right) \right] \Bigg\}^\tau
\end{aligned}$$

$$\cdot\left[Q_2+\frac{1}{2}\sum_{i,j=1}^{N} b_{ij}(S)B_t^\tau(i)(F_tF_t^\tau)^{-1}B_t(j)\right]$$

$$\cdot\left\{u_t+\frac{1}{2}\left(Q_2+\frac{1}{2}\sum_{i,j=1}^{N} b_{ij}(S)B_t^\tau(i)(F_tF_t^\tau)^{-1}B_t(j)\right)^{-1}\right.$$

$$\left.\cdot\left[\sum_{i=1}^{N} B_t^\tau(i)\left(\hat{p}_t^i\frac{\partial S}{\partial y_t}+a_i(S)\right)\right]\right\}dt$$

$$+E\int_s^T \{(y_t-y_t^*)^\tau Q_1(y_t-y_t^*)+u_t^\tau Q_2u_t\}dt. \tag{21}$$

Since under any admissible control

$$\sup_{0\leqslant t\leqslant T} E\|y_t\|^2<\infty$$

it follows by condition 2) on S that we have

$$|S(t, y_t, \hat{p}_t)|\leqslant k_4(1+\|y_t\|^2).$$

Hence, by the dominated convergence theorem and (19), we obtain

$$ES(t, y_t, \hat{p}_t)\underset{t\to T}{\to} ES|_{t=T}=E(y_T-y_T^*)^\tau Q_0(y_T-y_T^*)$$

since

$$E\|y_t\|^2\underset{t\to T}{\to} E\|y_T\|^2$$

by the continuity of y_t.

Finally, letting D exhaust $\mathbb{R}^m$ we have $s\to T$ in (21) and we conclude that

$$EJ(u)$$

$$=ES(0, y_0,\hat{p}_0)+\int_0^T\left\{u_t+\frac{1}{2}\left[Q_2+\frac{1}{2}\sum_{i,j=1}^{N} b_{ij}(S)B_t^\tau(i)\right.\right.$$

$$\left.\cdot(F_tF_t^\tau)^{-1}B_t(j)\right]^{-1}$$

$$\left.\cdot\left[\sum_{i=1}^{N} B_t^\tau(i)\left(\hat{p}_t^i\frac{\partial S}{\partial y_t}+a_i(S)\right)\right]\right\}^\tau$$

$$\left[Q_2+\sum_{i,j=1}^{N} b_{ij}(S)B_t^\tau(i)(F_tF_t^\tau)^{-1}B_t(j)\right]$$

$$\cdot\left\{u_t+\frac{1}{2}\left[Q_2+\frac{1}{2}\sum_{i,j=1}^{N} b_{ij}(S)B_t^\tau(i)(F_tF_t^\tau)^{-1}B_t(j)\right]^{-1}\right.$$

$$\left.\cdot\left[\sum_{i=1}^{N} B_t^\tau(i)\left(\hat{p}_t^i\frac{\partial S}{\partial y_t}+a_i(S)\right)\right]\right\}dt. \tag{22}$$

Notice that $ES\,(0, y_0, \hat{p}_0)$ is independent of the control, and so the control

$$u_t=u_t^*\triangleq-\frac{1}{2}\left[Q_2+\frac{1}{2}\sum_{i,j=1}^{N} b_{ij}(S)B_t^\tau(i)(F_tF_t^\tau)^{-1}B_t(j)\right]^{-1}$$

$$\cdot\sum_{i=1}^{N} B_t^\tau(i)\left(\hat{p}_t^i\frac{\partial S(t, y_t, \hat{p}_t)}{\partial y_t}+a_i(S)\right) \tag{23}$$

makes EJ reach its minimum.

Theorem: If there exists a function $S(t, y, p)$ satisfying conditions 1)–3) given above and if the system (11), (13), with u_t defined by (23), has a strong solution, then u_t^* given by (23) is the optimal control and

$$EJ(u^*)=ES(0, y_0, \hat{p}_0).$$

Proof. The only part of the proof that remains to be given is to show that $\{u_t^*\}$ is admissible. It is clear that u_t^* is $\mathfrak{F}_t^y$-measurable, hence the only thing to verify is that under u_t^* we have

$$\sup_{0\leqslant t\leqslant T} E\|y_t\|^2<\infty.$$

Now by (23)

$$\|u_t^*\|\leqslant\frac{1}{2}\,k_3\|\bar{Q}_2^{-1}\|$$

$$\cdot\sum_{i=1}^{N}\left\{\left\|\frac{\partial S(t, y_t, \hat{p}_t)}{\partial y_t}\right\|+\frac{1}{\delta}\left\|\frac{\partial^2 S(t, y_t, \hat{p}_t)}{\partial^2\hat{p}_t}\right\|(\|D_t\|+\|H_1\|)\right.$$

$$\left.+\left\|\frac{\partial^2 S(t, y_t, \hat{p}_t)}{\partial\hat{p}_t\partial y_t}\right\|\right\}$$

$$\leqslant\frac{1}{2}\,k_3\|\bar{Q}_2^{-1}\|N\left[k_6(1+\|y_t\|)\right.$$

$$\left.+\frac{k_5}{\delta}\,(k_1+k_2(1+\|y_t\|))\right]\leqslant k_7(1+\|y_t\|) \tag{24}$$

where

$$k_7=\frac{1}{2}\,k_3N\|\bar{Q}_2^{-1}\|\left(k_6+\frac{k_2k_5}{\delta}+\frac{k_1k_5}{\delta}\right).$$

Further, from (6)–(8), $\|y_t\|^2$ may be estimated by

$$\|y_t\|^2\leqslant 4\left(\|y_0\|+\left\|\int_0^t F_sdw_s\right\|^2\right)$$

$$+8Tk_2^2\int_0^t(1+\|y_s\|^2)ds+4Tk_3^2\int_0^t\|u_s^*\|^2ds. \tag{25}$$

Combining (24), (25) we see that there are constants k_8, k_9 such that

$$\|y_t\|^2\leqslant 4\left(k_8+\|y_0\|^2+\left\|\int_0^t F_sdw_s\right\|^2\right)+k_9\int_0^t\|y_s\|^2ds,$$

and hence by the Bellman-Gromwall lemma

$$\|y_t\|^2\leqslant 4\left(k_8+\|y_0\|^2+\left\|\int_0^t F_sdw_s\right\|^2\right)$$

$$+k_9\int_0^t e^{k_9(t-s)}4\left(k_8+\|y_0\|^2+\left\|\int_0^s F_\lambda dw_\lambda\right\|^2\right)ds.$$

From this we have

$$\sup_{0\leqslant t\leqslant T} E\|y_t\|^2<\infty \text{ since } E\|y_0\|^2<\infty \text{ and since}$$

$$\cdot\,E\left\|\int_0^t F_sdw_s\right\|^2=\text{tr E}\int_0^t F_sdw_s\left(\int_0^t F_sdw_s\right)^\tau$$

$$=\text{tr}\int_0^t F_sF_s^\tau ds\leqslant\text{tr}\int_0^T F_sF_s^\tau ds<\infty.$$ □

REFERENCES

[1] P. E. Caines, "Stochastic adaptive control: Randomly varing parameters and continually distributed controls," IFAC Congress, Kyoto, Japan, Aug. 1981, in *Control Science and Technology for the Progress of Society*, H. Akashi, Ed. New York: Pergamon, 1981.

[2] P. E. Caines and H. F. Chen, "On the adaptive control of stochastic systems with random parameters: A counter-example," presented at the Workshop on Adaptive Control, Florence, Italy, Oct. 1982; also in *Récerche di Automatica*, to be published.
[3] H. F. Chen and P. E. Caines, "On the adaptive control of a class of systems with random parameters and disturbances," in *Adaptive Systems in Control and Signal Processing*, (Proc. IFAC Workshop, San Francisco, CA, June 1983), Y. D. Landau, M. Tomizuka, and D. M. Auslander, Eds. New York: Pergamon, 1984.
[4] ——, "Optimal stochastic control for discrete time systems with random parameters," Dep. Elec. Eng., McGill Univ., Montreal, P.Q., Canada, Res. Rep., 1983.
[5] R. Evans and X. Xianya, "Discrete time adaptive control for deterministic time-varying systems," Dep. Elec. Eng., Univ. Newcastle, New South Wales, Australia, Tech. Rep. EE8225, 1982.
[6] O. Hijab, "The adaptive LQG problem—Part 1," *IEEE Trans. Automat. Contr.*, vol. AC-28. no. 2, pp. 171–178, 1983.
[7] Y. D. Landau, M. Tomizuka, and D. M. Auslander, *Adaptive Systems in Control and Signal Processing* (Proc. IFAC Workshop, San Francisco, CA, June 1983). New York: Pergamon, 1984.
[8] R. S. Liptser and A. N. Shiryayev, *Statistics of Random Process I*. New York: Springer-Verlag, 1977.
[9] W. M. Wonham, "Some applications of stochastic differential equations to optimal non-linear filtering," *SIAM J. Contr.*, vol. 2, pp. 347–369, 1965.

The Strong Consistency of the Stochastic Gradient Algorithm of Adaptive Control

H. F. CHEN AND P. E. CAINES

Abstract—**By use of the technique of Chen [5] sufficient conditions are established for the strong consistency of the stochastic gradient (SG) algorithm for MIMO ARMAX stochastic systems without monitoring. This result is then used in conjunction with the method of disturbed adaptive controls introduced in [1], [2]. Hence it is shown that the SG algorithm generates strongly consistent parameter estimates while it is operating as a part of the SG algorithm of adaptive control of Goodwin, Ramadge, and Caines [6].**

I. Introduction

The stochastic gradient (SG) algorithm is probably the simplest method of parameter estimation for linear stochastic systems. It was used in [6] for the adaptive tracking control problem and later on in [2] for the adaptive tracking problem where disturbed controls were used for purposes of identification as explained below. In these papers the strong consistency of the SG algorithm was not established. In [4] the strong consistency of the estimates generated by the quasi-least-squares (QLS) method was proved for the case of a system subject to feedback control without monitoring.

In this note, by use of the technique given in [5], we first establish sufficient conditions for the strong consistency of the SG algorithm for MIMO stochastic systems without monitoring and then apply the method involving continually disturbed controls which was introduced in [1] and developed in [7] and [2]. Continually disturbing a system's control input provides a technique for ensuring that certain processes of regression vectors have the persistency of excitation property which is frequently required in recursive identification schemes (see, e.g., [9]).

The ordinary differential equation method and the associated hypotheses used in this paper should be compared to those (involving monitoring) used to obtain the related results of [8, Theorem 1]. Concerning proof techniques, we also remark that the stochastic Lyapunov—or super martingale—method used in the basic first lemma below is fundamental to the consistency proofs of [10] and [3], [4] and to all of the references of this note concerned with stochastic adaptive control.

It will be shown that the SG algorithm used in [7] and [2] does in fact generate strongly consistent estimates. Hence, it is not necessary to introduce a second algorithm in order to generate consistent estimates, as in [2], where the approximate maximum likelihood (AML) algorithm of [10] was used in addition to the SG algorithm.

We consider the MIMO system

$$y_n + A_1 y_{n-1} + \cdots + A_p y_{n-p} = B_1 u_{n-1} + \cdots + B_q u_{n-q} + w_n + C_1 w_{n-1} + \cdots + C_r w_{n-r} \tag{1}$$

where y_n, u_n, and w_n are m-, l-, and m-dimensional, respectively, and $y_i = 0$, $u_j = 0$, $w_k = 0$ for all $i < 0$, $j < 0$, $k < 0$.

Let $\mathfrak{F}_n$ be a family of nondecreasing σ-algebras; assume that w_n and u_n are $\mathfrak{F}_n$-measurable and that

$$E(w_n/\mathfrak{F}_{n-1}) = 0, \quad E(\|w_n\|^2/\mathfrak{F}_{n-1}) \leqslant k_o r_{n-1}^{\epsilon}, \qquad 0 \leqslant \epsilon < 1 \tag{2}$$

where k_o is a positive constant and r_n is defined below in (10).

A_i, B_j, C_k, $i = 1, \cdots, p$, $j = 1, \cdots, q$, $k = 1, \cdots, r$ are the unknown matrix coefficients to be estimated.

Let us write

$$A(z) = I + A_1 z + \cdots + A_p z^p \tag{3}$$

$$B(z) = B_1 + B_2 z + \cdots + B_q z^{q-1} \tag{4}$$

$$C(z) = I + C_1 z + \cdots + C_r z^r \tag{5}$$

where z denotes the unit backward shift operator.

We shall adopt the following notation:

$$\theta^\tau = [-A_1, \cdots, -A_p, B_1, \cdots, B_q, C_1, C_r] \; mx(mp + lq + mr) \tag{6}$$

$$\begin{aligned} \phi_n^\tau = [&y_n^\tau, y_{n-1}^\tau, \cdots, y_{n-p+1}^\tau, u_n^\tau, \cdots, u_{n-q+1}^\tau \\ &y_n^\tau - \phi_{n-1}^\tau \theta_{n-1}, \cdots, y_{n-r+1}^\tau - \phi_{n-r}^\tau \theta_{n-r}] \end{aligned} \tag{7}$$

$$\phi_n^{o\tau} = [y_n^\tau, y_{n-1}^\tau \cdots y_{n-p+1}^\tau, u_n^\tau \cdots u_{n-1+1}^\tau, w_n^\tau \cdots w_{n-r+1}^\tau] \tag{8}$$

where θ_n is the estimate for θ given by the SG algorithm

$$\theta_{n+1} = \theta_n + \frac{\phi_n}{r_n}(y_{n+1}^\tau - \phi_n^\tau \theta_n), \tag{9}$$

$$r_n = 1 + \sum_{i=1}^{n} \|\phi_i\|^2, \qquad r_o = 1 \tag{10}$$

with ϕ_{-1} and θ_o deterministic and arbitrarily chosen.

The difference between the SG algorithm and the QLS algorithm lies in the fact that the residual term $y_n^\tau - \phi_{n-1}^\tau \theta_{n-1}$ in the SG algorithm is replaced by the term $y_n^\tau - \phi_{n-1}^\tau \theta_n$ in the QLS algorithm; in other words, the *a priori* prediction error is replaced by the *a posteriori* prediction error.

Set

$$\tilde{\theta}_n = \theta - \theta_n, \tag{11}$$

and

$$\xi_n = y_n - w_n - \theta_{n-1}^\tau \phi_{n-1}. \tag{12}$$

Then we have

$$\begin{aligned} C(z)(y_n - w_n - \theta_{n-1}^\tau \phi_{n-1}) &= \{(y_n - C(z)w_n) \\ &\quad + (C(z) - I)(y_n - \theta_{n-1}^\tau \phi_{n-1})\} - \theta_{n-1}^\tau \phi_{n-1} \\ &= \theta^\tau \phi_{n-1} - \theta_{n-1}^\tau \phi_{n-1} = \tilde{\theta}_{n-1}^\tau \phi_{n-1}, \end{aligned}$$

hence

$$C(z)\xi_n = \tilde{\theta}_{n-1}^\tau \phi_{n-1} \tag{13}$$

Manuscript received March 29, 1984. This paper is based on a prior submission of August 15, 1983. This work was supported by the Faculty of Engineering and the Faculty of Graduate Studies and Research, McGill University, by the NSERC International Scientific Exchange Programme, and by NSERC under Grant A1329, while the first author was visiting McGill University for the academic year 1982–1983.

H. F. Chen is with the Institute of Systems Science, Academia Sinica, Beijing, China.
P. E. Caines is with the Department of Electrical Engineering, McGill University, Montreal, P.Q., Canada, and the Canadian Institute for Advanced Research.