

# Adaptive Continuous-Time Linear Quadratic Gaussian Control

T. E. Duncan, *Fellow, IEEE*, L. Guo, *Fellow, IEEE*, and B. Pasik-Duncan, *Senior Member, IEEE*

**Abstract**— The adaptive linear quadratic Gaussian control problem, where the linear transformation of the state  $A$  and the linear transformation of the control  $B$  are unknown, is solved assuming only that  $(A, B)$  is controllable and  $(A, Q_1^{1/2})$  is observable, where  $Q_1$  determines the quadratic form for the state in the integrand of the cost functional. A weighted least squares algorithm is modified by using a random regularization to ensure that the family of estimated models is uniformly controllable and observable. A diminishing excitation is used with the adaptive control to ensure that the family of estimates is strongly consistent. A lagged certainty equivalence control using this family of estimates is shown to be self-optimizing for an ergodic, quadratic cost functional.

**Index Terms**— Adaptive control, least-squares, linear-quadratic-Gaussian, linear stochastic systems, optimality.

## I. INTRODUCTION

THE LINEAR Gaussian control problem with an ergodic, quadratic cost functional is probably the most well-known ergodic control problem. Since, in the known parameter case, the optimal control can be easily computed and the existence of an invariant measure for the optimal system follows directly from the stability of the optimal system, it is a basic problem to solve for stochastic adaptive control where the system parameters are assumed to be unknown. For discrete-time linear systems it has been studied extensively, especially for autoregressive moving average with exogenous input (ARMAX) models (cf. [4] for many references) and for models with a finite parameter set (cf. [10]). While this adaptive control problem has been less studied for continuous-time linear systems [2], [3], [6], it is nonetheless an important problem as a model for physical systems that naturally evolve in continuous time and as an approximation for discrete-time sampled systems when the sampling rate is high.

A stochastic adaptive control problem can be posed as a stochastic optimal control problem with partial observations where the unknown parameters are considered as unobserved states of the stochastic system. However, with this approach no specific results have been obtained in general for the explicit controller design and the stability and the optimality of the closed-loop system because the methods to solve the optimal

control problem are difficult both analytically and computationally. These difficulties are usually avoided in adaptive control by using the certainty equivalence principle, which is a simple separation principle that determines the control by the following two steps: 1) construct the optimal control by assuming that the true values of the unknown parameters are available and 2) replace the unknown parameters in the control by their current (online) estimates. For the second step in this procedure, the standard least squares (LS) algorithm is usually used for the parameter estimation and it is natural to expect that the family of LS estimates is strongly consistent and that the control determined by this procedure is asymptotically optimal. However, since the system signals that are used in the LS algorithm are obtained from a complicated nonlinear stochastic equation, the well-known excitation conditions [11] that are needed to guarantee the consistency of the LS estimates are difficult to verify in general. For this reason the stability and the optimality of the adaptive linear-quadratic-Gaussian (LQG) control problem have not been resolved.

One approach to solving the adaptive LQG control problem is to use the weighted least squares (WLS) algorithms instead of the standard LS algorithms. By suitably choosing the weights in the WLS algorithm, it is possible to establish some convergence properties that are better than those of the standard LS algorithm, without requiring any stability and excitation properties of the closed-loop systems. Indeed, for discrete-time ARMAX models, a WLS scheme introduced by Bercu [1] for the identification of the unknown system has been shown to be convergent under no stability and excitation assumptions on the closed-loop systems [8], which makes it possible to solve an associated adaptive LQG control problem [8]. For continuous-time linear stochastic systems, a WLS scheme motivated by [1] has provided convergence and strong consistency under certain excitation conditions [7]. However, the associated adaptive LQG control problem is not solved in general [14].

In this paper, a complete solution to the continuous-time adaptive LQG control problem is given, using only the natural assumptions of controllability and observability. First, the WLS scheme is used to obtain a family of convergent estimates with no assumptions on the excitation and the stability of the closed-loop systems (the self-convergence property). Furthermore, the scheme is modified by a random regularization procedure introduced in [8] to obtain a uniformly controllable and observable family of estimated models. Finally, a diminishing excitation white noise is used in the controller design to obtain the strong consistency of the family of estimated

Manuscript received February 7, 1997; revised January 8, 1998 and October 13, 1998. Recommended by Associate Editor, G. G. Yin. This work was supported in part by the NSF under Grants DMS 9305936 and DMS 9623439 and the National Natural Science Foundation of China.

T. E. Duncan and B. Pasik-Duncan are with the Department of Mathematics, University of Kansas, Lawrence, KS 66045 USA.

L. Guo is with the Institute of Systems Science, Chinese Academy of Science, Beijing 100080, China.

Publisher Item Identifier S 0018-9286(99)07133-0.

models, and a lagged certainty equivalence control is used to obtain the optimality of the quadratic cost functional. The excitation is sufficient to include the identification of unknown deterministic linear systems as a special case.

The approach used here eliminates some other assumptions that have been previously used that are either not necessary for the control problem for a known system (e.g., the open-loop stability imposed in [4]) or difficult to verify (e.g., the uniform controllability of the estimated models assumed in [9]). Furthermore, this approach eliminates the need for random switchings or resettings which often occur in previous work such as [3] and [4].

In Section II the design procedure of the adaptive LQG control is presented and the main results on stability and optimality of the closed-loop systems are stated. In Section III the proof of stability is given, and in Section IV the proof for the strong consistency and the optimality are given. Finally, in Section V some concluding remarks are given and an Appendix is in Section VI.

## II. CONTROLLER DESIGN AND MAIN RESULTS

### A. The Optimal LQG Control

Let  $(X(t), t \geq 0)$  be the process that satisfies the stochastic differential equation

$$dX(t) = AX(t)dt + BU(t)dt + DdW(t) \quad (1)$$

where  $X(0) = X_0$ ,  $X(t) \in \mathbb{R}^n$ ,  $U(t) \in \mathbb{R}^m$ ,  $(W(t), \mathcal{F}_t; t \geq 0)$  is an  $\mathbb{R}^p$ -valued standard Wiener process, and  $(U(t), \mathcal{F}_t; t \geq 0)$  is a control from a family that is specified subsequently. The random variables are defined on a fixed complete probability space  $(\Omega, \mathcal{F}, P)$  and the filtration  $(\mathcal{F}_t, t \geq 0)$  is defined on this space and specified subsequently. It is assumed that the matrices  $A$  and  $B$  are unknown.

The objective is to design an admissible control process  $(U(t), \mathcal{F}_t; t \geq 0)$  so that the following ergodic cost functional for system (1) is minimized:

$$J(U) = \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T [X^T(t)Q_1X(t) + U^T(t)Q_2U(t)] dt \quad (2)$$

where  $Q_2^T = Q_2 > 0$  and  $Q_1^T = Q_1 \geq 0$ , and a control  $(U(t), t \geq 0)$  for system (1) is said to be *admissible* if it is adapted to  $(\mathcal{F}_t, t \geq 0)$  and

$$\limsup_{t \rightarrow \infty} \frac{\int_0^t |U(s)|^2 ds}{\int_0^t |X(s)|^2 ds} < \infty \quad \text{a.s.} \quad (3)$$

The following standard assumptions are made.

- A1)  $(A, B)$  is controllable.
- A2)  $(A, Q_1^{1/2})$  is observable.

It is well known that, under Assumptions A1) and A2), the optimal control for the known system in the family of admissible controls is a linear feedback expressed as

$$U^0(t) = -Q_2^{-1}B^TRX(t)$$

where  $R$  is the unique positive, symmetric solution of the following algebraic (control) Riccati equation

$$A^TR + RA - RBQ_2^{-1}B^TR + Q_1 = 0. \quad (4)$$

The corresponding minimal cost is

$$J(U^0) = \inf_U J(U) = \text{tr}(D^TRD) \quad \text{a.s.} \quad (5)$$

However, in the present case the optimal control law  $U^0(t)$  is not implementable because the system parameter pair  $(A, B)$  is not available. This is the motivation for constructing online estimates for  $(A, B)$  below.

### B. The WLS Estimation

To describe the estimation problem in a standard form let

$$\theta^T = [A \quad B] \quad (6)$$

and

$$\varphi(t) = \begin{bmatrix} X(t) \\ U(t) \end{bmatrix} \quad (7)$$

so that (1) can be rewritten as a linear regression

$$dX(t) = \theta^T \varphi(t)dt + DdW(t). \quad (8)$$

Now, the family of continuous-time WLS estimates,  $(\theta(t), t \geq 0)$ , is given by

$$d\theta(t) = a(t)P(t)\varphi(t)[dX^T(t) - \varphi^T(t)\theta(t)dt] \quad (9)$$

$$dP(t) = -a(t)P(t)\varphi(t)\varphi^T(t)P(t)dt \quad (10)$$

where  $\theta(0) = [A_0, B_0]^T$  and  $P(0) > 0$  are arbitrary deterministic values such that  $(A_0, B_0)$  is controllable and  $(A_0, Q_1^{1/2})$  is observable

$$a(t) = \frac{1}{f(r(t))} \quad (11)$$

$$r(t) = \|P^{-1}(0)\| + \int_0^t |\varphi(s)|^2 ds \quad (12)$$

and  $f \in \mathbb{F}$  with

$$\mathbb{F} = \left\{ f|f: \mathbb{R}_+ \rightarrow \mathbb{R}_+, f \text{ is slowly increasing} \right. \\ \left. \text{and } \int_c^\infty \frac{dx}{xf(x)} < \infty \text{ for some } c \geq 0 \right\} \quad (13)$$

where a function  $f$  is called slowly increasing if it is increasing and satisfies  $f \geq 1$  and  $f(x^2) = O(f(x))$  as  $x \rightarrow \infty$ .

*Remark 1:* A necessary condition for a function  $f \in \mathbb{F}$  is that  $f(x) = o(\log x)$  [8]. Some typical functions that are used in WLS algorithms are  $\log^{1+\delta} x$  and  $(\log x)(\log \log x)^{1+\delta}$ . In fact, the family of weights  $(a(t), t \geq 0)$  defined by (11) satisfies

$$a^{-1}(t) = f(r(t)) = O(\log^\ell r(t))$$

for some  $\ell > 0$  as  $t \rightarrow \infty$ .

It is shown in Lemma 1-1) in Section III that the convergence rate of the WLS algorithm can be characterized by  $P(t)$ , that is

$$\|\theta(t) - \theta\|^2 = O(\|P(t)\|).$$

By (10) the explicit solution of  $P(t)$  is

$$P(t) = \left[ P^{-1}(0) + \int_0^t a(s)\varphi(s)\varphi(s)^T ds \right]^{-1}$$

which clearly shows that  $P(t)$  is positive and nonincreasing so  $(P(t), t \geq 0)$  converges a.s. as  $t \rightarrow \infty$ .

It is also worth noting that the standard LS algorithm corresponds to the choice  $f(x) \equiv 1$ , which is excluded in (13). As explained in [8] (see also Lemma 1 in the next section), the main reason for using this class of weights is that the corresponding WLS algorithm can be guaranteed to have better asymptotic properties than the standard LS algorithm, which is used in adaptive control. The most notable advantage of the WLS algorithm over the standard LS algorithm is that the WLS algorithm has a self-convergence property, whereas the traditional LS algorithm does not. The terminology ‘‘self-convergence’’ is used here to describe the fact that the WLS algorithm converges almost surely to a finite random matrix (not necessarily the true system parameter), regardless of the excitation properties of the data used in the algorithm, which are usually needed for other algorithms.

Even though the WLS is self-convergent, there is still no guarantee that the family of estimated models provided by the WLS method is controllable. This is another difficulty in the adaptive LQG problem, which is overcome by using the method of random regularization.

### C. Random Regularization

In this section a random regularization method is used to modify the family of WLS estimates to ensure that this new family is uniformly controllable and observable. Initially the notions of uniform controllability and uniform observability are defined.

*Definition 1:* A family of linear system models  $(A(t), B(t); t \geq 0)$  is said to be uniformly controllable if there is a  $c > 0$  such that

$$\sum_{i=0}^{n-1} A^i(t)B(t)B^T(t)A^{iT}(t) \geq cI$$

for all  $t \in [0, \infty)$ .

A family of models  $(A(t), C(t); t \geq 0)$  is said to be uniformly observable if  $(A^T(t), C^T(t); t \geq 0)$  is uniformly controllable.

If  $(A(t), B(t); t \geq 0)$  is convergent, this definition coincides with the traditional definition of uniform controllability for time-varying linear systems in the literature.

To motivate the regularization of the family of WLS estimates  $(\theta(t), t \geq 0)$  given by (9) and (10) to obtain a uniformly controllable and observable family of estimates, it is useful to recall the certainty equivalence principle. Since this principle neglects the uncertainties of the parameter estimates in determining the control, it has the advantage of simplicity but the disadvantage of using elementary estimates. To refine the WLS estimates it is natural to adjust them by accounting for the level of uncertainty of the estimates. Since by Lemma 1 the estimation error  $(\theta(t) - \theta)$  is characterized asymptotically by  $P^{1/2}(t)$ , it is natural to regularize the WLS estimates in a direction determined by the information matrix  $P^{1/2}(t)$  [8], [13].

Let  $\bar{\theta}(t, x)$  be given by

$$\bar{\theta}(t, x) = \theta(t) - P^{1/2}(t)x \quad (14)$$

and denote

$$\bar{\theta}^T(t, x) = [A(t, x) \quad B(t, x)] \quad (15)$$

where  $x \in \mathcal{M}(n+m, n)$  and  $\mathcal{M}(n+m, n)$  denotes the family of  $(n+m) \times n$  real matrices.

Since for any fixed  $x \in \mathcal{M}(n+m, n)$ , the matrix sequence  $\{A(t, x), B(t, x), C, t \geq 0\}$  with  $C = Q_1^{1/2}$  is bounded, its uniform controllability and observability is equivalent to the uniform positivity of the families  $(F(t, x), t \geq 0)$  and  $(G(t, x), t \geq 0)$  where

$$F(t, x) = \det \left( \sum_{i=0}^{n-1} A^i(t, x)B(t, x)B^T(t, x)A^{iT}(t, x) \right). \quad (16)$$

and

$$G(t, x) = \det \left( \sum_{i=0}^{n-1} A^{iT}(t, x)C^T C A^i(t, x) \right). \quad (17)$$

Note that if  $x$  could be chosen as

$$x = P^{-1/2}(t)[\theta(t) - \theta]$$

then  $\bar{\theta}(t, x)$  would coincide with the true system parameter  $\theta$ , and by Assumptions A1) and A2) the family  $(A(t, x), B(t, x), C, t \geq 0)$  would be trivially uniformly controllable and observable. Unfortunately, such a choice of  $x$  is not feasible because  $\theta$  is unknown, so other methods must be used.

Using a method in [8], an  $x \in \mathcal{M}(n+m, n)$  is chosen that approximately maximizes  $f(t, x) = F(t, x)G(t, x)$ . This approximate maximization is achieved by a simple recursive random maximization procedure.

Let  $(\eta_k, k \in \mathbb{N})$  be a sequence of independent, identically distributed  $\mathcal{M}(n+m, n)$ -valued random variables that is independent of  $(W(t), t \geq 0)$  so that for each  $k \in \mathbb{N}$  the random variable  $\eta_k$  is uniformly distributed on the unit ball for a norm of the matrices. The maximization procedure is

recursively defined as follows:

$$\beta_0 = 0$$

$$\beta_k = \begin{cases} \eta_k, & \text{if } f(k, \eta_k) \geq (1 + \gamma)f(k, \beta_{k-1}) \\ \beta_{k-1}, & \text{otherwise} \end{cases} \quad (18)$$

where  $\gamma \in (0, \sqrt{2} - 1)$  is fixed and  $f(k, x) = F(k, x)G(k, x)$ .

Since at each iteration  $k$  only four determinants are required to be computed, the procedure is simpler than others [13]. The random sequence  $(\beta_k, k \in \mathbb{N})$  in the procedure provides an approximate maximization of  $f$  so that the sequence of estimates  $(\bar{\theta}_k, k \in \mathbb{N})$  given by

$$\bar{\theta}_k = \theta(k) - P^{1/2}(k)\beta_k \quad (19)$$

has the desired uniform controllability and observability properties where  $\theta(k)$  and  $P(k)$  are given by the WLS algorithm (9) and (10).

Finally, the family of continuous-time estimates  $(\hat{\theta}(t), t \geq 0)$  to be used for the adaptive control problem is simply a piecewise constant function induced by (19)

$$\hat{\theta}(t) = \bar{\theta}_k, \quad t \in (k, k+1] \quad (20)$$

where  $k \in \mathbb{N}$  and  $\bar{\theta}_k$  is given by (19).

#### D. The Main Results

For notational simplicity let the estimates given by (20) be expressed as

$$\hat{\theta}^T(t) = [A(t) \quad B(t)].$$

In the next section (see Lemma 2), it is verified that the family  $(A(t), B(t), Q_1^{1/2}; t \geq 0)$  is uniformly controllable and observable. Hence, the following stochastic algebraic (control) Riccati equation:

$$A^T(t)R(t) + R(t)A(t) - R(t)B(t)Q_2^{-1}B^T(t)R(t) + Q_1 = 0 \quad (21)$$

has a unique, adapted, symmetric, positive solution  $R(t)$  for each  $t \in [0, \infty)$  a.s.

Using  $R(t)$ , define a lagged certainty equivalence LQG control by

$$U(t) = -Q_2^{-1}B^T(t)R(t)X(t). \quad (22)$$

It is called "lagged" because a delay is present from (20), that is, by (20) and (21)

$$(B(t), R(t)) = (B(0), R(0))$$

for  $t \in [0, 1]$ , and

$$(B(t), R(t)) = (\bar{B}_k, \bar{R}_k)$$

for  $t \in (k, k+1]$  where  $\bar{\theta}_k = [\bar{A}_k, \bar{B}_k]^T$  is given by (19). This approach simplifies the task of computations of the controller (22) because the solution of the Riccati equation (21) is needed only at discrete-time instants  $t = 1, 2, 3, \dots$ .

The first theorem below states that using the above lagged certainty equivalence control (22), the solution of (1) is stable in the averaging sense.

*Theorem 1:* The process  $(X(t), t \geq 0)$  that is the solution of (1) with the adaptive control (22) is stable in the sense that

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T |X(s)|^2 ds < \infty \quad \text{a.s.} \quad (23)$$

The proof is given in Section III.

To obtain the optimality of the quadratic cost functional, it is necessary to obtain the strong consistency for the family of estimates  $(\hat{\theta}(t), t \geq 0)$  [10], [12]. For this, a diminishing excitation is added to the adaptive control (22), that is

$$U^*(t) = L_k X(t) + \gamma_k [V(t) - V(k)]$$

or

$$dU^*(t) = L_k dX(t) + \gamma_k dV(t) \quad (24)$$

for  $t \in (k, k+1]$  and  $k \in \mathbb{N}$  where  $U^*(0) \in \mathbb{R}^m$  is an arbitrary deterministic vector,

$$L_k = -Q_2^{-1}B^T(k)R(k) \quad (25)$$

and

$$\gamma_k^2 = \frac{\log k}{\sqrt{k}} \quad (26)$$

for  $k \geq 1$ . The process  $(V(t), t \geq 0)$  is an  $\mathbb{R}^m$ -valued standard Wiener process that is independent of  $(W(t), t \geq 0)$  and  $(\eta_k, k \in \mathbb{N})$ . Without loss of generality, the sub- $\sigma$ -algebra  $\mathcal{F}_t$  is defined as the  $P$ -completion of  $\sigma(X_0, W(s), \eta_j, V(s); s \leq t, j \leq t)$ .

Since the sequence of random processes  $(V(t) - V(k), k \in \mathbb{N}, t \in [k, k+1))$  are independent and have the same probability law and the sequence  $(\gamma_k, k \in \mathbb{N})$  converges to zero at a suitable rate, it is feasible that the diminishing excitation  $(\gamma_k[V(t) - V(k)], k \in \mathbb{N}, t \in [k, k+1))$  provides sufficient excitation for identification and is sufficiently small enough not to affect the optimality of the cost. The following theorem shows that the family of regularized WLS estimates is strongly consistent using the lagged certainty equivalence control with diminishing excitation. The proof is given in Section IV.

*Theorem 2:* Let  $(\hat{\theta}(t), t \geq 0)$  be the family of estimates given by (19) and (20) using the control (24) in (1). If A1) and A2) are satisfied, then

$$\lim_{t \rightarrow \infty} \hat{\theta}(t) = \theta \quad \text{a.s.} \quad (27)$$

where  $\theta$  is the true system parameter defined by (6).

*Remark 2:* It should be noted that the conditions of Theorem 2 are satisfied if  $D \equiv 0$  so that the identification of deterministic systems is included in this result.

Since the family of estimates  $(\hat{\theta}(t), t \geq 0)$  is strongly consistent, the self-optimality of the diminishingly excited lagged certainty equivalence control (24) can be verified as in [6], so the details are omitted.

*Theorem 3:* Let A1) and A2) be satisfied for the stochastic system (1) with the cost functional (2) where  $A$  and  $B$  are unknown. Then the adaptive control defined by (24)–(26) is admissible and optimal, that is

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T [X^T(t)Q_1 X(t) + U^{*\top}(t)Q_2 U^*(t)] dt = \text{tr}(D^T R D) \quad \text{a.s.}$$

where  $R$  is defined by (4).

### III. PROOF OF THEOREM 1

To prove Theorem 1, some basic properties of the WLS algorithm are used that are stated in Lemma 1 below.

*Lemma 1:* Let  $(\theta(t), P(t); t \geq 0)$  satisfy (9) and (10). Then the following properties are satisfied:

- 1)  $\sup_{t \geq 0} |P^{-1/2}(t)\hat{\theta}(t)|^2 < \infty$  a.s.;
- 2)  $\int_0^\infty a(t)|\hat{\theta}^T(t)\varphi(t)|^2 dt < \infty$  a.s.;
- 3)  $\lim_{t \rightarrow \infty} \theta(t) = \bar{\theta}$  a.s.;

where  $\hat{\theta}(t) = \theta(t) - \theta$  and  $\bar{\theta}$  is a matrix-valued random variable.

The proof of this lemma is given in the Appendix.

*Remark 3:* The standard LS algorithm does not have the three properties stated in Lemma 1 for the WLS algorithm, which is explained in [8] for the discrete-time case. These properties are the primary advantage to introducing the class of weights defined by (11)–(13) in constructing the WLS identification algorithm. If  $D = 0$ , then Lemma 1 holds surely, that is, the ‘‘a.s.’’ can be removed in the statement of Lemma 1.

*Remark 4:* For discrete-time ARMAX models the properties 1) and 2) in Lemma 1 are given by Bercu [1] and property 3) is given by Guo [8], and in continuous-time they are given by Gao [7]. The properties are verified here in the Appendix because the family of functions,  $\mathbb{F}$ , given by (13) [8], is somewhat different from the family in [1] that is used in [7], and the verification is somewhat different from [7].

Using Lemma 1, the following key result on the properties of the regularized estimates is verified.

*Lemma 2:* Let A1) and A2) be satisfied for system (1) with the quadratic cost (2). Then for any admissible control  $(U(t), t \geq 0)$ , the family of regularized WLS estimates  $(\hat{\theta}(t), t \geq 0)$  defined by (18)–(20) has the following properties.

- 1) Self-convergence, that is,  $\hat{\theta}(t)$  converges a.s. to a finite random matrix as  $t \rightarrow \infty$ .
- 2) The family  $(A(t), B(t), Q_1^{1/2}; t \geq 0)$  is uniformly controllable and observable where  $[A(t) \ B(t)] = \hat{\theta}^T(t)$ .
- 3) Semiconsistency, that is, as  $t \rightarrow \infty$

$$\int_0^t |(\hat{\theta}(s) - \theta)^T \varphi(s)|^2 ds = o(r(t)) + 0(1) \quad \text{a.s.}$$

where  $r(t)$  is defined by (12).

*Proof:* To verify 1) it is clear from Lemma 1-3) that only the convergence of  $(\beta_k, k \in \mathbb{N})$  is necessary to verify. Since  $f(k, x)$  is a polynomial in  $x$  it is continuous and therefore  $f(k, \cdot)$  is bounded on a compact set. This boundedness of  $f$  implies that the recursion in (18) terminates in a finite random number of iterations. In the proof of [8, Th. 2] an algorithm of the form (18) is used and the following are given: 1) the finite termination of  $(\beta_k, k \in \mathbb{N})$  and thereby the convergence almost surely of  $(\beta_k, k \in \mathbb{N})$ ; 2)  $\inf_{k \in \mathbb{N}} f(k, \beta_k) > 0$  a.s.; and 3)  $\lim_{k \rightarrow \infty} f(k, \beta_k) \geq (1/(1 + \gamma)) \max_{x \in K} [\lim_{k \rightarrow \infty} f(k, x)]$  a.s., where  $\gamma$  is given in (18) and  $K$  is the unit ball where  $\eta_k$  is uniformly

distributed. Property 2) follows from the fact that the elements  $x \in K$  where (16) and (17) are strictly positive are open and dense. Property 3) is the approximate maximization property of the algorithm (18).

To verify 2) it suffices from (20) to show that the sequence  $(\bar{A}_k, \bar{B}_k, Q; k \in \mathbb{N})$  is uniformly controllable and observable where  $(\bar{A}_k, \bar{B}_k) = \bar{\theta}_k^T$ . By the definition of  $f(k, x)$  it suffices to show that  $\inf_{k \in \mathbb{N}} f(k, \beta_k) > 0$  a.s., which by the algorithm (18) and the inequality for the approximate maximization follows by verifying

$$\max_{x \in K} \left[ \lim_{k \rightarrow \infty} f(k, x) \right] > 0 \quad \text{a.s.}$$

If  $g$  is given by

$$g(x) = \lim_{k \rightarrow \infty} f(k, x)$$

then  $g(x)$  is a polynomial in  $x$  so it suffices to show that  $g \not\equiv 0$ .

If  $x_k = P^{-1/2}(k)(\theta(k) - \theta)$  then by Lemma 1-1),  $(x_k, k \in \mathbb{N})$  is (almost surely) bounded and  $\theta = \theta(k) + P^{1/2}(k)x_k$ . By A1) and A2) and the definition of  $f(k, x)$ , it follows that  $f(k, x_k) = c > 0$  for all  $k \in \mathbb{N}$ . If  $(x_{k_l}, l \in \mathbb{N})$  is a subsequence of  $(x_k, k \in \mathbb{N})$  that converges to some  $x^*$  then  $g(x^*) = c > 0$  so  $g \not\equiv 0$ . This verifies 2).

To prove assertion 3), first note that if  $r(t)$  is bounded, then 3) clearly holds by the boundedness of  $(\hat{\theta}(t), t \geq 0)$  and the definition of  $r(t)$ . Hence, it is only necessary to consider the case where  $\lim_{t \rightarrow \infty} r(t) = \infty$ . In this case, by Lemma 1-2), the Kronecker lemma, and Remark 1, it follows that

$$\int_0^t |\hat{\theta}^T(s)\varphi(s)|^2 ds = o(f(r(t))) = o(r(t)) \quad \text{a.s.}$$

Consequently, by Lemma 1-3) and the definition of  $r(t)$  it follows that

$$\int_0^t |(\bar{\theta} - \theta)^T \varphi(s)|^2 ds = o(r(t)) \quad \text{a.s.} \quad (28)$$

Next, by (65) and (66) in the Appendix it follows that

$$\begin{aligned} & \int_0^t a(s)\varphi(s)^T P(s)\varphi(s) ds \\ &= \log |P^{-1}(t)| = O(\log r(t)). \end{aligned}$$

Hence by denoting  $P(\infty) = \lim_{t \rightarrow \infty} P(t)$ , it follows from (11) and Remark 1 that

$$\begin{aligned} & \int_0^t \varphi(s)^T P(\infty)\varphi(s) ds \\ & \leq \int_0^t \varphi(s)^T P(s)\varphi(s) ds \\ & \leq f(r(t)) \int_0^t a(s)\varphi(s)^T P(s)\varphi(s) ds \\ & = O(\log^{t+1} r(t)) = o(r(t)). \end{aligned} \quad (29)$$

Finally, denoting  $\hat{\theta}(\infty) = \lim_{t \rightarrow \infty} \hat{\theta}(t)$ , it follows by (19) Lemma 1-3) that

$$\hat{\theta}(\infty) = \bar{\theta} - P^{1/2}(\infty)\beta_\infty.$$

Hence the following inequalities hold:

$$\begin{aligned}
& \int_0^t |[\hat{\theta}(s) - \theta]^T \varphi(s)|^2 ds \\
&= \int_0^t |[\hat{\theta}(\infty) - \theta + \hat{\theta}(s) - \hat{\theta}(\infty)]^T \varphi(s)|^2 ds \\
&\leq 2 \int_0^t |(\hat{\theta}(\infty) - \theta)^T \varphi(s)|^2 ds \\
&\quad + 2 \int_0^t |(\hat{\theta}(s) - \hat{\theta}(\infty))^T \varphi(s)|^2 ds \\
&\leq 4 \int_0^t |(\bar{\theta} - \theta)^T \varphi(s)|^2 ds \\
&\quad + 4 \int_0^t |\beta_\infty^T P^{1/2}(\infty) \varphi(s)|^2 ds \\
&\quad + 2 \int_0^t |(\hat{\theta}(s) - \hat{\theta}(\infty))^T \varphi(s)|^2 ds \\
&= o(r(t))
\end{aligned}$$

where for the last step (12), (28), and (29) are used. Hence, the proof of Lemma 2 is completed.  $\square$

*Proof of Theorem 1:* By Lemma 2 there are random matrices  $A(\infty)$  and  $B(\infty)$  such that

$$\lim_{t \rightarrow \infty} A(t) = A(\infty) \quad \text{a.s.} \quad (30)$$

and

$$\lim_{t \rightarrow \infty} B(t) = B(\infty) \quad \text{a.s.} \quad (31)$$

Furthermore, the triple  $(A(\infty), B(\infty), Q_1^{1/2})$  is controllable and observable a.s.

Since the solution of the stochastic algebraic Riccati equation (21) is a smooth function of the parameters of the equation [5], there is a symmetric, positive  $\mathcal{L}(\mathbb{R}^n)$ -valued random variable  $R(\infty)$  such that

$$\lim_{t \rightarrow \infty} R(t) = R(\infty) \quad \text{a.s.}$$

If  $\Phi(t)$  is given by

$$\Phi(t) = A(t) - B(t)Q_2^{-1}B^T(t)R(t)$$

and

$$\Phi(\infty) = A(\infty) - B(\infty)Q_2^{-1}B^T(\infty)R(\infty)$$

then

$$\lim_{t \rightarrow \infty} \Phi(t) = \Phi(\infty)$$

and it is well known that both  $\Phi(t)$  and  $\Phi(\infty)$  are stable a.s.

By the stability of  $\Phi(t)$ , there is an  $\mathcal{L}(\mathbb{R}^n)$ -valued random matrix  $\bar{P}(t) > 0$  a.s. such that

$$\bar{P}(t)\Phi(t) + \Phi^T(t)\bar{P}(t) = -I \quad \text{a.s.} \quad (32)$$

Furthermore,  $\{\bar{P}(t), t \geq 0\}$  is bounded a.s. by the uniform stability of  $\{\Phi(t), 0 \leq t \leq \infty\}$ .

Note that

$$\begin{aligned}
dX(t) &= [A(t)X(t) + B(t)U(t) + \tilde{\delta}(t)]dt + DdW(t) \\
&= [\Phi(t)X(t) + \tilde{\delta}(t)]dt + DdW(t). \quad (33)
\end{aligned}$$

Then, applying Itô's formula to  $\langle \bar{P}(t)X(t), X(t) \rangle$ , and noting that  $P(t)$  is actually constant in any interval  $t \in (k, k+1]$ ,  $k \in \mathbb{N}$ , it follows that

$$\begin{aligned}
& d\langle \bar{P}(t)X(t), X(t) \rangle \\
&= 2\langle \bar{P}(t)X(t), \Phi(t)X(t) + \tilde{\delta}(t) \rangle dt \\
&\quad + \text{tr}[\bar{P}(t)DD^T]dt + 2\langle \bar{P}(t)X(t), DdW(t) \rangle \quad (34)
\end{aligned}$$

where  $\tilde{\delta}(t) = [\theta - \hat{\theta}(t)]^T \varphi(t)$ .

Now, by Theorem 2-3)

$$\begin{aligned}
\int_0^T |\tilde{\delta}(t)|^2 dt &= \int_0^T |[\theta - \hat{\theta}(t)]^T \varphi(t)|^2 dt \\
&= o(r(T)) + O(1) \quad \text{a.s.} \quad (35)
\end{aligned}$$

Furthermore, by the law of the iterated logarithm for Brownian motion and a time change of the stochastic integral [4, Lemma 12.3] it follows that

$$\begin{aligned}
& \left| \int_0^T \langle X(\sigma), P(\sigma)DdW(\sigma) \rangle \right| \\
&= O\left( \left[ \int_0^T |X(s)|^2 ds \right]^{(1/2)+\varepsilon} \right) \quad (36)
\end{aligned}$$

for each  $\varepsilon \in (0, 1/2)$ .

Integrating (34) and using (33), (35), and (36) it follows that

$$\begin{aligned}
& \langle \bar{P}(T)X(T), X(T) \rangle - \langle \bar{P}(0)X(0), X(0) \rangle + \int_0^T |X(s)|^2 ds \\
&\leq o(r(T)) + O(1) + O\left( \left[ \int_0^T |X(s)|^2 ds \right]^{(1/2)+\varepsilon} \right) \\
&\quad + \int_0^T \text{tr}[\bar{P}(t)DD^T]dt. \quad (37)
\end{aligned}$$

Since the control is given by (22), the desired inequality (23) follows from (37).  $\square$

#### IV. PROOF OF THEOREM 2

Augment the state  $X(t)$  with the control  $U(t)$  as follows:

$$\varphi(t) = \begin{bmatrix} X(t) \\ U(t) \end{bmatrix} \quad (38)$$

$$F_k = \begin{bmatrix} A & B \\ L_k A & L_k B \end{bmatrix} \quad (39)$$

$$G_k = \begin{bmatrix} D & 0 \\ L_k D & \gamma_k I \end{bmatrix} \quad (40)$$

$$\xi(t) = \begin{bmatrix} W(t) \\ V(t) \end{bmatrix}. \quad (41)$$

The stochastic differential equation for the augmented state process  $(\varphi(t), t \geq 0)$  is

$$d\varphi(t) = F_k \varphi(t) dt + G_k d\xi(t) \quad (42)$$

for  $t \in (k, k+1]$ .

The following result provides an inequality for a process that plays a central role in the verification of strong consistency.

*Lemma 3:* Let  $(H(t) - H(k), t \in (k, k + 1])$  be given by

$$H(t) - H(k) = \int_k^t \varphi(s)\varphi^T(s) ds. \quad (43)$$

Then

$$\begin{aligned} H(k+1) - H(k) &\geq \int_k^{k+1} e^{F_k(k+1-s)} \left[ G_k G_k^T \int_k^s d\tau + \int_k^s dg(\tau) \right] \\ &\cdot e^{F_k^T(k+1-s)} ds \quad \text{a.s.} \end{aligned} \quad (44)$$

where

$$dg(s) = \varphi(s) d\xi^T(s) G_k^T + G_k d\xi(s) \varphi^T(s) \quad (45)$$

for  $s \in (k, k + 1]$ .

*Proof:* Apply Itô's formula to  $(\varphi(t)\varphi^T(t), t \geq 0)$  and integrate this result to obtain for  $t \in [k, k + 1]$

$$\begin{aligned} \varphi(t)\varphi^T(t) &= \varphi(k)\varphi^T(k) + \int_k^t \varphi(s)\varphi^T(s) ds F_k^T \\ &+ F_k \int_k^t \varphi(s)\varphi^T(s) ds + G_k G_k^T (t - k) + g(t) - g(k). \end{aligned}$$

Thus

$$\begin{aligned} \frac{dH(t)}{dt} &= (H(t) - H(k))F_k^T + F_k(H(t) - H(k)) \\ &+ \varphi(k)\varphi^T(k) + G_k G_k^T (t - k) + g(t) - g(k). \end{aligned}$$

Integrating this differential equation and deleting the term that contains  $\varphi(k)\varphi^T(k)$  gives (44).  $\square$

The second integral on the right-hand side of (44) forms a martingale difference sequence.

*Lemma 4:* The process  $(M(k + 1) - M(k), k \in \mathbb{N})$  given by

$$\begin{aligned} M(k+1) - M(k) &= \int_k^{k+1} e^{F_k(k+1-s)} \int_k^s dg(\tau) e^{F_k^T(k+1-s)} ds \end{aligned} \quad (46)$$

is a martingale difference sequence and there is a random variable  $M$  such that

$$E[|M(k+1) - M(k)|^2 | \mathcal{F}_k] \leq M(|X(k)|^2 + 1) \quad \text{a.s.} \quad (47)$$

for all  $k \in \mathbb{N}$ .

*Proof:* By interchanging the order of integration in (46) it follows immediately that  $(M(k + 1) - M(k), k \in \mathbb{N})$  is a martingale difference sequence. Solving (42) for  $t \in [k, k + 1]$  it follows that

$$\varphi(t) = e^{F_k(t-k)} \varphi(k) + \int_k^t e^{F_k(t-s)} G_k d\xi(s).$$

Thus

$$\begin{aligned} E[|\varphi(t)|^2 | \mathcal{F}_k] &= |e^{F_k(t-k)} \varphi(k)|^2 + \int_k^t |e^{F_k(t-s)} G_k|^2 ds \\ &= O(|\varphi(k)|^2 + 1) \\ &= O(|X(k)|^2 + 1). \end{aligned}$$

For  $s \in [k, k + 1]$  it follows that

$$\begin{aligned} E \left[ \left| \int_k^s dg(\tau) \right|^2 | \mathcal{F}_k \right] &\leq 2|G_k|^2 E \left[ \left| \int_k^s \varphi(\tau) d\xi^T(\tau) \right|^2 | \mathcal{F}_k \right] \\ &\leq 2|G_k|^2 \int_k^s E[|\varphi(\tau)|^2 | \mathcal{F}_k] d\tau \\ &= O(|X(k)|^2 + 1). \end{aligned}$$

It is elementary from (46) that

$$|M(k+1) - M(k)|^2 \leq \int_k^{k+1} |e^{F_k(k+1-s)}|^4 \left| \int_k^s dg(\tau) \right|^2 ds.$$

The last two inequalities verify the inequality (47).  $\square$

*Lemma 5:* Let  $(X(t), t \geq 0)$  be the solution of (1) with the lagged certainty equivalence control (24). Then

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \int_0^t (|X(s)|^2 + |U^*(s)|^2) ds < \infty \quad \text{a.s.} \quad (48)$$

and

$$\limsup_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k |X(j)|^2 < \infty \quad \text{a.s.} \quad (49)$$

*Proof:* Since

$$\begin{aligned} E \sum_{k=2}^{\infty} \int_k^{k+1} \frac{\gamma_k^2 |V(t) - V(k)|^2}{\sqrt{k} \log^3 k} dt &\leq \sum_{k=2}^{\infty} \frac{1}{k \log^2 k} \int_k^{k+1} E|V(t) - V(k)|^2 dt \\ &= \sum_{k=2}^{\infty} \frac{1}{k \log^2 k} \int_k^{k+1} (t - k) dt < \infty \end{aligned}$$

it follows by the Kronecker lemma that

$$\limsup_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k \int_j^{j+1} \gamma_j^2 |V(t) - V(j)|^2 dt = 0 \quad \text{a.s.} \quad (50)$$

It is elementary to extend the stability result of Theorem 1 to the control given by (24). Using this extended Theorem 1, (48) is verified. To prove (49), first note that similar to (33) using the control (24)

$$dX(t) = [\Phi(t)X(t) + \tilde{\delta}(t)] dt + DdW_t + \gamma_k B(V(t) - V(k)) dt$$

for  $t \in (k, k + 1]$ .

Consequently

$$\begin{aligned} X(k+1) &= e^{\bar{\Phi}_k} X(k) + \int_k^{k+1} e^{(k+1-t)\bar{\Phi}_k} \\ &\cdot [\tilde{\delta}_t + \gamma_k B(V(t) - V(k))] dt + \int_k^{k+1} e^{(k+1-t)\bar{\Phi}_k} DdW_t \end{aligned}$$

where  $\bar{\Phi}_k$  denotes  $\Phi(t)$  for  $t \in (k, k + 1]$ .

Note that  $\bar{\Phi}_k$  converges a.s. and the family  $\{e^{\bar{\Phi}_k}, k \geq 0\}$  is a.s. uniformly stable. It is easy to conclude (49) by using (35), (48), (50), and the martingale difference property of the stochastic integral in the above equation.

Lemma 6: Let  $\overline{G}_k$  for  $k \in \mathbb{N}$  be given by

$$\overline{G}_k = \begin{bmatrix} D & 0 \\ L_k D & I \end{bmatrix} \quad (51)$$

and  $0 < \gamma_k < 1$  for  $k \in \mathbb{N}$ . Then

$$G_k G_k^T \geq \gamma_k^2 \overline{G}_k \overline{G}_k^T \quad (52)$$

for  $k \in \mathbb{N}$  where  $G_k$  is given by (40).

*Proof:* It follows directly from the definitions of  $G_k$  and  $\overline{G}_k$  that

$$G_k G_k^T = \gamma_k^2 \overline{G}_k \overline{G}_k^T + (1 - \gamma_k^2) \begin{bmatrix} D & 0 \\ L_k D & 0 \end{bmatrix} \begin{bmatrix} D & 0 \\ L_k D & 0 \end{bmatrix}^T.$$

Since the family of positive semidefinite matrices is a proper positive cone in the family of square matrices, it follows that

$$G_k G_k^T \geq \gamma_k^2 \overline{G}_k \overline{G}_k^T. \quad \square$$

The following result provides an asymptotic property of the other term in (44) that uses controllability.

Lemma 7: Let  $(N(k+1) - N(k), k \in \mathbb{N})$  be given by

$$\begin{aligned} & N(k+1) - N(k) \\ &= \int_k^{k+1} e^{F_k(k+1-s)} G_k G_k^T \int_k^s d\tau e^{F_k^T(k+1-s)} ds. \end{aligned} \quad (53)$$

Then

$$\liminf_{k \rightarrow \infty} \gamma_k^{-2} \lambda_{\min}(N(k+1) - N(k)) > c \quad \text{a.s.} \quad (54)$$

where  $\lambda_{\min}(M)$  is the smallest eigenvalue of  $M = M^T$

$$c = \lambda_{\min} \left( \int_0^1 e^{Ft} \overline{G} \overline{G}^T (1-t) e^{F^T t} dt \right) \quad (55)$$

and

$$\begin{aligned} F &= \begin{bmatrix} A & B \\ LA & LB \end{bmatrix} \\ \overline{G} &= \begin{bmatrix} D & 0 \\ LD & I \end{bmatrix} \end{aligned}$$

where  $L$  is a matrix-valued random variable such that  $\lim_{k \rightarrow \infty} L_k = L$  a.s.

*Proof:* By Lemma 6 it follows that

$$\begin{aligned} & N(k+1) - N(k) \\ & \geq \gamma_k^2 \int_k^{k+1} e^{F_k(k+1-s)} \overline{G}_k \overline{G}_k^T \int_k^s d\tau e^{F_k^T(k+1-s)} ds \\ & = \gamma_k^2 \int_0^1 e^{F_k t} \overline{G}_k \overline{G}_k^T (1-t) e^{F_k^T t} dt. \end{aligned} \quad (56)$$

Since  $\lim_{k \rightarrow \infty} L_k = L$  a.s. it follows that  $\lim_{k \rightarrow \infty} F_k = F$  a.s. and  $\lim_{k \rightarrow \infty} \overline{G}_k = \overline{G}$  a.s. By (56) it follows that

$$\begin{aligned} & \liminf_{k \rightarrow \infty} \gamma_k^{-2} \lambda_{\min}(N(k+1) - N(k)) \\ & \geq \lambda_{\min} \left( \int_0^1 e^{Ft} \overline{G} \overline{G}^T (1-t) e^{F^T t} dt \right). \end{aligned}$$

The right-hand side of this inequality is positive if and only if  $(F, \overline{G})$  is controllable. By elementary row operations,  $(F, \overline{G})$

is conjugate to  $(\begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} D & 0 \\ 0 & I \end{bmatrix})$ . In this latter form it is clear that  $(F, \overline{G})$  is controllable if  $(A, B)$  is controllable and  $D$  can be arbitrary.  $\square$

*Proof of Theorem 2:* By (10) it follows that

$$dP^{-1}(t) = a(t)\varphi(t)\varphi^T(t)dt.$$

Since the weight process  $(a(t), t \geq 0)$  is decreasing it follows that

$$\lambda_{\min}(P^{-1}(k)) \geq a(k) \lambda_{\min} \left( \int_0^k \varphi(s)\varphi^T(s) ds \right). \quad (57)$$

By Lemmas 3, 4, and 7 it easily follows that

$$H(k+1) - H(k) \geq N(k+1) - N(k) + M(k+1) - M(k)$$

so

$$\begin{aligned} & \lambda_{\min} \left( \int_0^k \varphi(s)\varphi^T(s) ds \right) \\ &= \lambda_{\min}(H(k)) \\ & \geq \lambda_{\min}(N(k) + M(k)) \\ & \geq \sum_{j=0}^{k-1} \lambda_{\min}\{N(j+1) - N(j)\} - |M(k)|. \end{aligned} \quad (58)$$

By Lemma 7 there are two positive random variables  $K_0$  and  $\Gamma$  such that

$$\gamma_k^{-2} \lambda_{\min}(N(k+1) - N(k)) \geq \Gamma$$

for all  $k \geq K_0$  so

$$\sum_{i=K_0}^k \lambda_{\min}(N(i+1) - N(i)) \geq \Gamma \sum_{i=K_0}^k \gamma_i^2 \geq \frac{\Gamma}{2} \sqrt{k} \log k. \quad (59)$$

By this inequality and

$$\limsup_{k \rightarrow \infty} \frac{1}{\sqrt{k}(\log k)^{(1/2)+\varepsilon}} M(k) = 0 \quad \text{a.s.} \quad (60)$$

for each  $\varepsilon \in (0, (1/2))$  that is obtained from the law of the iterated logarithm for Brownian motion it follows from (57) and (58) that

$$\lambda_{\min} \left( \int_{K_0}^k \varphi(s)\varphi^T(s) ds \right) \geq \frac{\Gamma}{2} \sqrt{k} \log k. \quad \text{a.s.} \quad (61)$$

By the stability of the closed-loop system given in Lemma 5 it follows that

$$r(k) = O(k) \quad \text{a.s.}$$

as  $k \rightarrow \infty$ . By Remark 1, (57), (58), and (61) it follows that if  $k \geq K_0$  then

$$\begin{aligned} \|P(k)\| &= \lambda_{\min}^{-1}(P^{-1}(k)) \\ & \leq \frac{a^{-1}(k)}{\lambda_{\min} \left( \int_0^k \varphi(s)\varphi^T(s) ds \right)} \\ & = O \left( \frac{\log^{\ell-1} k}{\sqrt{k}} \right) \quad \text{a.s.} \end{aligned} \quad (62)$$



as  $k \rightarrow \infty$ . By Lemma 1-1) it follows that

$$\begin{aligned} |\tilde{\theta}(k)|^2 &\leq \|P(k)\| \cdot \|P^{-1/2}(k)\tilde{\theta}(k)\|^2 = O(\|P(k)\|) \\ &= O\left(\frac{\log^{\ell-1} k}{\sqrt{k}}\right) \quad \text{a.s.} \end{aligned}$$

as  $k \rightarrow \infty$ . This inequality shows that

$$\lim_{t \rightarrow \infty} \theta(t) = \theta \quad \text{a.s.}$$

By (19) and (20) it is clear that

$$\lim_{k \rightarrow \infty} \hat{\theta}(k) = \theta \quad \text{a.s.}$$

because  $P(k) \rightarrow 0$  and  $(\beta(k), k \in \mathbb{N})$  is uniformly bounded.  $\square$

## V. CONCLUDING REMARKS

The continuous-time adaptive LQG control problem is a basic one in adaptive control. The major technical questions associated with this problem are: 1) finding a strongly consistent family of estimates of the unknown pair  $(A, B)$  so that this family is uniformly controllable and observable; 2) ensuring a strong solution of the closed-loop adaptive control system; and 3) determining a suitable continuous-time diminishing excitation for the system. In this paper, a complete solution is provided for this adaptive control problem. For question 1), a WLS method with random regularization is used; for question 2), a lagged certainty equivalence control is used; and for question 3), a diminishing family of increments of Brownian motion is used. No random switchings of the control are used, and since no condition is imposed on the linear transformation  $D$ , of the noise, the deterministic adaptive linear quadratic control problem is included as a special case.

A natural important generalization of this adaptive control problem is the problem where the state is only partially observed with noise. For discrete-time systems, some results for special partial observations are given in [8]. However, even in discrete time, the general partially observed adaptive LQG control problem has not been solved and further investigations are needed. Another important direction for generalization is the case where the unknown system parameter pair  $(A, B)$  is time-varying.

## APPENDIX

*Proof of Lemma 1:* Let  $Y(t) = \text{tr}[\tilde{\theta}^T(t)P^{-1}(t)\tilde{\theta}(t)]$  for  $t \geq 0$  and apply Itô's formula to  $(Y(t), t \geq 0)$  using (9) and (10)

$$\begin{aligned} dY(t) &= \text{tr}[d\tilde{\theta}^T(t)P^{-1}(t)\tilde{\theta}(t) + \tilde{\theta}^T(t)P^{-1}(t)d\tilde{\theta}(t)] \\ &\quad + \text{tr}[a^2(t)D\varphi^T(t)P(t)\varphi(t)D^T]dt \\ &\quad + \text{tr}[\tilde{\theta}^T(t)a(t)\varphi(t)\varphi^T(t)\tilde{\theta}(t)]dt \\ &= \text{tr}[DdW(t)a(t)\varphi^T(t)\tilde{\theta}(t) + \tilde{\theta}^T(t)a(t)\varphi(t)(DdW(t))^T] \\ &\quad - 2\text{tr}[\tilde{\theta}^T(t)a(t)\varphi(t)\varphi^T(t)\tilde{\theta}(t)]dt \\ &\quad + \text{tr}[a^2(t)D\varphi^T(t)P(t)\varphi(t)D^T]dt. \end{aligned} \quad (63)$$

By applying the law of the iterated logarithm for Brownian motion and a time change of the stochastic integrals [4, Lemma 12.3] and using the decreasing property of  $(a(t), t \geq 0)$ , the sum of the integrals of the first two terms is negative for  $S \geq T$  where  $T$  is a random time. Thus to prove the boundedness of  $Y(t)$  as  $t \rightarrow \infty$ , it suffices to prove the boundedness of the integral of the third term on the right-hand side of (63).

From (10) it follows that

$$P^{-1}(t + \Delta) = P^{-1}(t)(I + a(t)P(t)\varphi(t)\varphi^T(t)\Delta) + o(\Delta). \quad (64)$$

Let  $|A| = \det A$  for notational simplicity. From (64) it follows that

$$\begin{aligned} |P^{-1}(t + \Delta)| &= |P^{-1}(t)||I + a(t)P(t)\varphi(t)\varphi^T(t)\Delta| + o(\Delta) \\ &= |P^{-1}(t)|(1 + a(t)\varphi^T(t)P(t)\varphi(t)\Delta) + o(\Delta). \end{aligned}$$

Thus

$$\frac{d|P^{-1}(t)|}{|P^{-1}(t)|} = a(t)\varphi^T(t)P(t)\varphi(t)dt. \quad (65)$$

Since by Remark 1, the maximum eigenvalue of  $P^{-1}(t)$  is bounded by  $r(t)$  defined by (12), it is easy to verify the following inequality:

$$|P^{-1}(t)| \leq \lambda_{\max}^{n+m}(t) \leq [r(t)]^{n+m} \quad (66)$$

where  $\lambda_{\max}(t) = \lambda_{\max}(P^{-1}(t))$ . Now using (65) and the slowly increasing property of  $f(\cdot)$  given in (13) it follows that

$$\begin{aligned} &\int_c^t \text{tr}(a^2(s)D\varphi^T(s)P(s)\varphi(s)D^T)ds \\ &= \text{tr}(DD^T) \int_c^t a^2(s)\varphi^T(s)P(s)\varphi(s)ds \\ &\leq K \int_c^T \frac{d|P^{-1}(s)|}{|P^{-1}(s)|f(r(s))} \\ &\leq K_1 \int_c^\infty \frac{d|P^{-1}(s)|}{|P^{-1}(s)|f(|P^{-1}(s)|)} \\ &\leq K_1 \int_c^\infty \frac{dx}{xf(x)} < \infty. \end{aligned} \quad (67)$$

Thus it follows that for  $f \in \mathbb{F}$

$$E \int_c^\infty \text{tr}(a^2(s)D\varphi^T(s)P(s)\varphi(s)D^T)ds < \infty. \quad (68)$$

Thus by (63), assertions 1) and 2) are verified. To verify 3), recall the stochastic differential equation for  $(\tilde{\theta}(t), t \geq 0)$

$$\begin{aligned} d\tilde{\theta}(t) &= a(t)P(t)\varphi(t)[dX^T(t) - \varphi^T(t)\theta(t)dt] \\ &= a(t)P(t)\varphi(t)[- \varphi^T(t)\tilde{\theta}(t)dt + (DdW(t))^T] \end{aligned} \quad (69)$$

so that

$$\begin{aligned} \tilde{\theta}(t) &= \tilde{\theta}(0) - \int_0^T a(s)P(s)\varphi(s)\varphi^T(s)\tilde{\theta}(s)ds \\ &\quad + \int_0^T a(s)P(s)\varphi(s)(DdW(s))^T. \end{aligned} \quad (70)$$

Integrating the differential equation (10) and applying the trace, it follows that

$$\int_0^T a(s) |\varphi^T(s) P(s)|^2 ds \leq \text{tr}(P(0)) < \infty. \quad (71)$$

From this inequality and assertion 2) it easily follows that

$$\begin{aligned} & \left| \int_0^\infty a(s) P(s) \varphi(s) \varphi^T(s) \tilde{\theta}(s) ds \right| \\ & \leq \left[ \int_0^\infty a(s) |\tilde{\theta}^T(s) \varphi(s)|^2 ds \right]^{1/2} \\ & \cdot \left[ \int_0^\infty a(s) |P(s) \varphi(s)|^2 ds \right]^{1/2} < \infty. \quad (72) \end{aligned}$$

Furthermore, by (71) it follows that

$$E \int_0^\infty a(s) |\varphi^T(s) P(s)|^2 ds < \infty.$$

Thus by the martingale convergence theorem, the stochastic integral in (70) converges a.s. as  $t \rightarrow \infty$  to a real-valued random variable. Thus 3) is verified.  $\square$

#### REFERENCES

- [1] B. Bercu, "Weighted estimation and tracking for ARMAX models," *SIAM J. Contr. Optim.*, vol. 33, pp. 89–106, 1995.
- [2] P. E. Caines, "Continuous-time stochastic adaptive control: Non-explosions,  $\varepsilon$ -consistency and stability," *Syst. Contr. Lett.*, vol. 19, pp. 169–176, 1992.
- [3] H. F. Chen, T. E. Duncan, and B. Pasik-Duncan, "Stochastic adaptive control for continuous time linear systems with quadratic cost," *J. Appl. Math. Optim.*, vol. 34, pp. 113–138, 1996.
- [4] H. F. Chen and L. Guo, *Identification and Stochastic Adaptive Control*. Boston, MA: Birkhäuser, 1991.
- [5] D. Delchamps, "A note on the analyticity of the Riccati metric," *Algebraic and Geometric Methods in Linear System Theory, Lectures in Applied Mathematics*, vol. 18. Providence, RI: AMS, 1980.
- [6] T. E. Duncan and B. Pasik-Duncan, "Adaptive control of continuous time linear stochastic systems," *Math. Contr., Signals, Syst.*, vol. 3, pp. 45–60, 1990.
- [7] A. J. Gao, "Self-convergence of weighted least squares for continuous-time ARMAX model," *Ulam Quart.*, vol. III, 1996.
- [8] L. Guo, "Self-convergence of weighted least-squares with applications to stochastic adaptive control," *IEEE Trans. Automat. Contr.*, vol. 41, pp. 79–89, 1996.
- [9] P. A. Ioannou and J. Sun, *Robust Adaptive Control*. Upper Saddle River, NJ: Prentice-Hall, 1996.
- [10] P. R. Kumar, "Optimal adaptive control of linear-quadratic-Gaussian systems," *SIAM J. Contr. Optim.*, vol. 21, pp. 163–178, 1983.
- [11] T. L. Lai and C. Z. Wei, "Least-square estimation in stochastic regression models with applications to identification and control of dynamic systems," *Ann. Statist.*, vol. 10, pp. 154–166, 1982.
- [12] W. Lin, P. R. Kumar, and T. I. Seidman, "Will the self-tuning approach work for general cost criteria?," *Syst. Contr. Lett.*, vol. 6, pp. 77–85, 1985.
- [13] R. Lozano and X. H. Zhao, "Adaptive pole placement without excitation probing signals," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 47–58, 1994.
- [14] A. J. Gao and B. Pasik-Duncan, "Stochastic linear quadratic adaptive control for continuous-time first-order systems," *Syst. Contr. Lett.*, vol. 31, pp. 149–154, 1997.



**T. E. Duncan** (M'92–SM'96–F'99) received the B.E.E. degree from Rensselaer Polytechnic Institute, Troy, NY, in 1963 and the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, in 1964 and 1967, respectively.

He has held regular positions at the University of Michigan (1967–1971), the State University of New York, Stony Brook (1971–1974), and the University of Kansas, Lawrence (1974–present), where he is a Professor of mathematics. He has held visiting positions at the University of California, Berkeley (1969–1970), the University of Bonn, Germany (1978–1979), and Harvard University, Cambridge, MA (1979–1980) and shorter visiting positions at numerous other institutions.

Dr. Duncan is a Corresponding Editor of *SIAM Journal on Control and Optimization* and is a member of AMS, MAA, and SIAM.



**L. Guo** (M'88–F'99) was born in Shandong, China, in November 1961. He received the B.S. degree in mathematics from Shandong University in 1982 and the M.S. and Ph.D. degrees in control theory from the Chinese Academy of Sciences in 1984 and 1987, respectively.

He was a Postdoctoral Fellow at the Australian National University from 1987 to 1989. Since 1992 he has been a Professor with the Institute of Systems Science of the Chinese Academy of Sciences, where he is currently the Director of the Systems and Control Laboratory. He is the author of the book *Time-Varying Stochastic Systems: Stability, Estimation and Control* (in Chinese) and coauthor of the book *Identification and Stochastic Adaptive Control* (Boston, MA: Birkhäuser, 1991).

Dr. Guo is the recipient of the Young Scientist Award of China (1994), the Natural Science Prize of the Chinese Academy of Sciences (1994), the Young Author Prize of the International Federation of Automatic Control World Congress (1993), and (with his coauthor) the National Natural Science Prize of China (1987). He was an Associate Editor of the *SIAM Journal on Control and Optimization* from 1991 to 1993 and is currently a standing member of the Council of the Chinese Association of Automation and Vice Chairman of its Committee on Control Theory.



**B. Pasik-Duncan** (M'90–SM'96) received the M.S. degree from the Mathematics Department of Warsaw University in 1970 and the Ph.D. and Habilitation degrees from the Mathematics Department, Warsaw School of Economics in 1978 and 1986, respectively.

She was a staff member of the Mathematics Department, Warsaw School of Economics from 1970 to 1983. In 1984 she joined the University of Kansas, where she is a Professor of Mathematics. Her research interests include adaptive control and its applications, stochastic theory and its applications, mathematics of finance, numerical methods in control, manufacturing systems, distributed parameter systems, mathematical statistics, and mathematical education.

Dr. Pasik-Duncan served as Program Director of the SIAM activity Group on Control and Systems Theory and was a member of the SIAM Master Program Committee. She is an Associate Editor-at-Large of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL and is currently serving as Vice President for membership activities of the Control Systems Society. Her professional memberships include AMS, ASA, AWM, the Bernoulli Society, MAAA, the Polish Mathematical Society, SIAG/CST, and SIAM.