View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 72, 122–149 (1979)

Algorithms for Adaptive Stochastic Control for a Class of Linear Systems*

M. TODA

International Institute for Advanced Study of Social Information Science, Fujitsu Ltd., Tokyo, Japan

AND

R. V. PATEL

Department of Electrical Engineering, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

Submitted by M. Aoki

This paper is concerned with the control of linear, discrete-time, stochastic systems with unknown control gain parameters. Two suboptimal adaptive control schemes are derived: One is based on underestimating future control and the other is based on overestimating future control. Both schemes require little on-line computation and incorporate in their control laws some information on estimation errors. The performance of these laws is studied by Monte Carlo simulations on a computer. Two single-input, third-order systems are considered, one stable and the other unstable, and the performance of the two adaptive control schemes is compared with that of the scheme based on enforced certainty equivalence and the scheme where the control gain parameters are known.

1. INTRODUCTION

Problems of controlling systems under uncertainty have long attracted the attention of many control theorists and engineers because of their importance in practical control systems. Since the work of Bellman [1], the stochastic adaptive control approach has been useful in treating such problems ([2]; also see [3] for a survey). For state space models, the optimization approach for stochastic adaptive control has been studied extensively. However, explicit solutions have been obtained for only a limited class of problems, for example, the well known certainty equivalence solution of the standard linear quadratic Gaussian

* This work was carried out at NASA, Ames Research Center, under the NRC Post-doctoral Research Associateship program.

0022-247X/79/110122-28$02.00/0
Copyright © 1979 by Academic Press, Inc.
All rights of reproduction in any form reserved.

problem. Although more general problems have been conceptually solved (i.e., requiring formal solutions of functional equations), explicit forms of the optimal control laws (if they exist) have yet to be obtained. In order to overcome the difficulties in solving the functional equations, many suboptimal schemes have been proposed [3]. Most of them incorporate approximations for some features of adaptive control. However, except for the ad hoc scheme where the certainty equivalence principle is enforced (this scheme will be called the CE law), they usually require a considerable amount of on-line computation, which can often be prohibitive. For example, the control law based on the dual control approach in [4], which exhibits an active learning property, requires extensive on-line compuation to evaluate future observation programs. The open loop optimal feedback control law (OLOF) ignores future measurements but incorporates some information concerning the uncertainty (covariances of estimation errors) in its control algorithm [5–7]. In this sense, this scheme was called "cautious" in [3]. The OLOF law still requires numerical optimization techniques on-line.

The purpose of this study is to investigate two suboptimal schemes which require little on-line computation but incorporate the effects of estimation errors in their control laws, and to study the performance of these laws by Monte Carlo simulations on a computer. We consider discrete-time linear stochastic systems with unknown control gain parameters (essentially the same class of problems as that treated in [6]). Admittedly, this class of systems is small in practice. However, we believe that because of their conceptual simplicity and computational efficiency, the two laws derived in this paper may provide a suitable framework for treating the more general problem, i.e., when the system state and control gain matrices are both unknown.

One of the control laws is based on *u*nder*e*stimating *f*uture *c*ontrol, hence called the UEFC law, and the other is based on *o*ver*e*stimating *f*uture *c*ontrol, the OEFC law. Two single-input, third-order systems (one stable and the other unstable) are simulated, and the performance of the UEFC and OEFC laws is compared with that of the CE law and the law where the control gain parameters are known. The sensitivity of the performance of the four laws is studied for various levels of initial uncertainties in the states and the control gain parameters.

This paper is organized as follows: Section 2 defines the notation. A precise definition of the problem is given in Section 3. Section 4 presents the results of the application of Kalman filter theory for the optimal estimation problem. We derive the UEFC and OEFC laws in Section 5, and Section 6 shows the results of the Monte Carlo simulations. Section 7 concludes with remarks on this study.

2. Notation

The transpose of a matrix X (vector x) is denoted by $X^{\mathrm{T}}(x^{\mathrm{T}})$. The trace of a square matrix X is denoted by $\mathrm{tr}(X)$. The matrices I_n and $0_{m,n}$ denote the

n-dimensional identity matrix and the $m \times n$ null matrix, respectively; the subscripts will be dropped when there is no ambiguity. The notation $X > 0$ ($X \geqslant 0$) denotes a positive definite (semidefinite) matrix X, and $X > Y$ ($X \geqslant Y$) implies $X - Y > 0$ ($X - Y \geqslant 0$). The Kronecker product of matrices X and Y is denoted by $X \otimes Y$. The mn-dimensional row and column string vectors of an $m \times n$ matrix X are denoted by $\mathrm{rs}(X)$ and $\mathrm{cs}(X)$; i.e.,

$$[\mathrm{rs}(X)]^{\mathrm{T}} \equiv (x_{R1}^{\mathrm{T}} x_{R2}^{\mathrm{T}} \cdots x_{Rm}^{\mathrm{T}}),$$
$$[\mathrm{cs}(X)]^{\mathrm{T}} \equiv (x_{C1}^{\mathrm{T}} x_{C2}^{\mathrm{T}} \cdots x_{Cn}^{\mathrm{T}}),$$

where $x_{Ri}^{\mathrm{T}}(x_{Ci})$ is the ith row (column) vector of X.

The (conditional) expectation of a random vector x (given Y) is denoted by $E[x]$ ($E[x \mid Y]$). The notation $x \sim N(\bar{x}, X)$ means that a random vector x has Gaussian distribution with mean $\bar{x}$ and covariance X. Statements with "a.s." imply that they hold with probability 1.

Symbols with subscript or superscript "U" ("O") pertain to algorithms for UEFC (OEFC).

3. Problem Statement

We consider a standard finite-stage discrete-time linear stochastic control problem with a quadratic performance index. The system dynamics and measurement relations are described by

$$x(k+1) = Ax(k) + Bu(k) + D\xi(k), \tag{1}$$

$$y(k+1) = Cx(k+1) + \eta(k+1), \qquad k = 0, 1, \ldots, N-1, \tag{2}$$

where the state $x(k)$, the control $u(k)$, the measurement $y(k)$, and the plant noise $\xi(k)$ are vectors of dimensions n, m, l, and r, respectively. The matrices A, C, and D are of appropriate dimensions and are assumed to be known. The $n \times m$ control matrix B is a random matrix[1] with

$$b \sim N(\bar{b}, P_b), \qquad b \equiv \mathrm{rs}(B).$$

The other primary random variables are

$$x(0) \sim N(\bar{x}_0, P_0),$$
$$\xi(k) \sim N(0, Q(k)),$$
$$\eta(k) \sim N(0, R(k)), \qquad R(k) > 0;$$

[1] For simplicity of derivation, we assume that B is a constant matrix. The extension of our results to the case with linearly varying B as in [6] is straightforward.

$\xi(k)$ and $\eta(k)$ are mutually independent white noise sequences, and both are independent of b and $x(0)$; b and $x(0)$ are also mutually independent.

The performance measure we wish to minimize is given by

$$J = E\left[\sum_{k=0}^{N-1} J(k)\right] = E\left\{\sum_{k=0}^{N-1}\left[x(k+1)^{\mathrm{T}} S(k+1)\, x(k+1) + u(k)^{\mathrm{T}} \Lambda(k)\, u(k)\right]\right\}, \tag{3}$$

where $S(k+1) \geqslant 0$ and $\Lambda(k) > 0$. Admissible control laws are causal; i.e.,

$$u(k) = u(k, Y(k), U(k-1)),$$

where $Y(k) \equiv [y(1),..., y(k)]$ and $U(k-1) \equiv [u(0),..., u(k-1)]$; $u(0)$ must be a function of prior information on the system.

4. Estimation

Since the system equations (1) and (2) are linear in the random vector $x(k)$ and random matrix B, Kalman filter theory can be applied to modified system equations to obtain the optimal minimum variance estimates.

Applying Lemma A1 in Appendix I, we get

$$Bu(k) = I_n Bu(k) = [I_n \otimes u(k)^{\mathrm{T}}]\, b. \tag{4}$$

We can write the following system equations for the augmented state vector $z(k)^{\mathrm{T}} \equiv [x(k)^{\mathrm{T}}\, b^{\mathrm{T}}]$:

$$z(k+1) = F(k)\, z(k) + G\xi(k), \tag{5}$$

$$y(k+1) = Hz(k+1) + \eta(k+1), \tag{6}$$

where[2]

$$F(k) \equiv \begin{bmatrix} A & I_n \otimes u(k)^{\mathrm{T}} \\ 0_{nm,n} & I_{nm} \end{bmatrix}, \qquad G \equiv \begin{bmatrix} D \\ 0_{nm,r} \end{bmatrix}, \tag{7}$$

$$H \equiv [C \quad 0_{l,nm}]. \tag{8}$$

[2] If we arrange the vectors of B columnwise we obtain augmented system equations of the same form as (5)–(8), except that $F(k)$ is given by

$$F(k) \equiv \begin{bmatrix} A & u(k)^{\mathrm{T}} \otimes I_n \\ 0_{nm,n} & I_{nm} \end{bmatrix}.$$

The augmented state vector for this case is $z(k)^{\mathrm{T}} \equiv [x(k)^{\mathrm{T}} b_c^{\mathrm{T}}]$, where $b_c \equiv \mathrm{cs}(B)$. The row string arrangement in (5)–(8) is preferred in order to facilitate backward optimization (see Section 5).

Application of Kalman filter theory to the linear equations (5) and (6) yields the following optimal minimum variance estimate:

$$\hat{z}(k+1 \mid k+1) = F(k)\,\hat{z}(k \mid k) + K(k+1)\,[y(k+1) - HF(k)\,\hat{z}(k \mid k)], \quad (9)$$

$$K(k+1) = P(k+1 \mid k)\,H^{\mathrm{T}}[HP(k+1 \mid k)\,H^{\mathrm{T}} + R(k)]^{-1}, \quad (10)$$

$$P(k+1 \mid k) = F(k)\,P(k \mid k)\,F^{\mathrm{T}}(k) + GQ(k)\,G^{\mathrm{T}}, \quad (11)$$

$$P(k+1 \mid k+1) = [I_n - K(k+1)\,H]\,P(k+1 \mid k), \quad (12)$$

$$\hat{z}(0 \mid 0) = \begin{bmatrix} \bar{x}_0 \\ \bar{b} \end{bmatrix}, \qquad P(0 \mid 0) = \begin{bmatrix} P_0 & 0_{n,nm} \\ 0_{nm,n} & P_b \end{bmatrix},$$

where $\hat{z}(k \mid k) \equiv E[z(k) \mid Y(k)]$, $\hat{z}(k+1 \mid k) \equiv E[z(k+1) \mid Y(k)] = F(k)\,\hat{z}(k \mid k)$, and

$$P(k \mid k) = E[\{z(k) - \hat{z}(k \mid k)\}\{z(k) - \hat{z}(k \mid k)\}^{\mathrm{T}} \mid Y(k)], \quad (13)$$

$$P(k+1 \mid k) = E[\{z(k+1) - \hat{z}(k+1 \mid k)\}\{z(k+1) - \hat{z}(k+1 \mid k)\}^{\mathrm{T}} \mid Y(k)]. \quad (14)$$

We partition $\hat{z}(i \mid k)$ and $P(i \mid k)$ as

$$\hat{z}(i \mid k) \equiv \begin{bmatrix} \hat{x}(i \mid k) \\ \hat{b}(i \mid k) \end{bmatrix}, \qquad P(i \mid k) \equiv \begin{bmatrix} \pi_1(i \mid k) & \pi_3(i \mid k)^{\mathrm{T}} \\ \pi_3(i \mid k) & \pi_2(i \mid k) \end{bmatrix}, \quad (15)$$

where $\hat{x}(i \mid k)$ is an n-dimensional vector, and $\pi_1(i \mid k)$ and $\pi_2(i \mid k)$ are $n \times n$ and $nm \times nm$ matrices, respectively.

5. Feedback Control Laws

It is well known that the control laws which solve the optimization problem are the formal solutions of the functional equation [2].

$$J_k^* = \min_{u(k)} J_k, \qquad k = N-1,\ldots,0, \quad (16)$$

where

$$J_k \equiv E[J(k) + J_{k+1}^* \mid Y(k)], \qquad J_N^* \equiv 0.$$

However, closed form solutions of the backward optimization are not available, and various suboptimal schemes have been proposed (see, for example, [3] for a survey of such schemes). Some of the schemes [4, 6] require a considerable amount of on-line computation at each stage k. We derive here two feedback laws which do not require lengthy on-line computations. The two laws are obtained by carrying out the backward optimization (16) approximately. In the following derivations of the control laws, the time indices will be dropped for brevity when there is no ambiguity in notation.

5.1. *Control Law Based on Underestimating Future Control Efforts* (UEFC)

This control law is derived by underestimating the effects of future control. The backward "suboptimization" proceeds as follows:

Last stage: $k = N - 1$. Since $J_N \equiv 0$, it is easy to obtain the quadratic cost-to-go functional,

$$\begin{aligned} J_{N-1} &= E[x(N)^{\mathrm{T}} S(N)\, x(N) + u(N-1)^{\mathrm{T}} \Lambda(N-1)\, u(N-1) \mid Y(N-1)] \\ &= u(N-1)^{\mathrm{T}} \Lambda(N-1)\, u(N-1) \\ &\quad + \operatorname{tr}\{S(N)\, E[Bu(N-1)\, u(N-1)^{\mathrm{T}} B^{\mathrm{T}} \mid Y(N-1)]\} \\ &\quad + 2\operatorname{tr}\{A^{\mathrm{T}} S(N)\, E[Bu(N-1)\, x(N-1)^{\mathrm{T}} \mid Y(N-1)]\} \\ &\quad + \alpha(N-1) + \beta(N-1), \end{aligned} \tag{17}$$

where

$$\alpha(N-1) \equiv \operatorname{tr}\{A^{\mathrm{T}} S(N)\, A E[x(N-1)\, x(N-1)^{\mathrm{T}} \mid Y(N-1)]\}, \tag{18}$$

$$\beta(N-1) \equiv \operatorname{tr}[D^{\mathrm{T}} S(N)\, DQ(N-1)], \tag{19}$$

are independent of $u(N-1)$.

Recalling (4), we can rewrite the second and third terms as

$$\begin{aligned} &\operatorname{tr}\{SE[Buu^{\mathrm{T}}B^{\mathrm{T}} \mid Y]\} \\ &\quad = \operatorname{tr}\{S(I_n \otimes u^{\mathrm{T}})\, E[bb^{\mathrm{T}} \mid Y]\, (I_n \otimes u)\} \\ &\quad = \operatorname{tr}\{S(N)\, [I_n \otimes u(N-1)^{\mathrm{T}}]\, M_2(N-1 \mid N-1)\, [I_n \otimes u(N-1)]\}, \end{aligned} \tag{20}$$

$$\begin{aligned} &\operatorname{tr}\{A^{\mathrm{T}} SE[Bux^{\mathrm{T}} \mid Y]\} \\ &\quad = \operatorname{tr}\{A^{\mathrm{T}} S[I_n \otimes u^{\mathrm{T}}]\, E[bx^{\mathrm{T}} \mid Y]\} \\ &\quad = \operatorname{tr}\{A^{\mathrm{T}} S(N)\, [I_n \otimes u(N-1)^{\mathrm{T}}]\, M_3(N-1 \mid N-1)\}, \end{aligned} \tag{21}$$

where the M_j's are defined by

$$\begin{aligned} M(i \mid k) &\equiv \begin{bmatrix} M_1(i \mid k) & M_3(i \mid k)^{\mathrm{T}} \\ M_3(i \mid k) & M_2(i \mid k) \end{bmatrix} \equiv E\left[\begin{matrix} x(i)\, x(i)^{\mathrm{T}} & x(i)\, b^{\mathrm{T}} \\ bx(i)^{\mathrm{T}} & bb^{\mathrm{T}} \end{matrix} \,\middle|\, Y(k)\right] \\ &= \begin{bmatrix} \pi_1(i \mid k) + \hat{x}(i \mid k)\, \hat{x}(i \mid k)^{\mathrm{T}} & \pi_3(i \mid k)^{\mathrm{T}} + \hat{x}(i \mid k)\, \hat{b}(i \mid k)^{\mathrm{T}} \\ \pi_3(i \mid k) + \hat{b}(i \mid k)\, \hat{x}(i \mid k)^{\mathrm{T}} & \pi_2(i \mid k) + \hat{b}(i \mid k)\, \hat{b}(i \mid k)^{\mathrm{T}} \end{bmatrix}. \end{aligned} \tag{22}$$

Applying Lemma A2 to (20) and (21), we have

$$\begin{aligned}\operatorname{tr}\{S(I_n \otimes u^{\mathrm{T}})\, M_2(I_n \otimes u)\} \\ &= [\operatorname{cs}(I_n \otimes u)]^{\mathrm{T}}\, (S \otimes M_2)\, \operatorname{cs}(I_n \otimes u) \\ &= u(N-1)^{\mathrm{T}}\, [\Gamma^{\mathrm{T}}(S(N) \otimes M_2(N-1 \mid N-1))\, \Gamma]\, u(N-1), \end{aligned} \tag{23}$$

$$\begin{aligned}\operatorname{tr}\{A^{\mathrm{T}}S(I_n \otimes u^{\mathrm{T}})\, M_3\} &= \operatorname{tr}\{(I_n \otimes u^{\mathrm{T}}) M_3 A^{\mathrm{T}} S\} \\ &= [\operatorname{cs}(I_n \otimes u)]^{\mathrm{T}}\, \operatorname{cs}(M_3 A^{\mathrm{T}} S) \\ &= \{\Gamma^{\mathrm{T}}\, \operatorname{cs}[M_3(N-1 \mid N-1)\, A^{\mathrm{T}} S(N)]\}^{\mathrm{T}}\, u(N-1), \end{aligned} \tag{24}$$

where the following identity was used to obtain the final expressions:

$$\begin{aligned} \operatorname{cs}(I_n \otimes u) &= \Gamma u, \\ \Gamma^{\mathrm{T}} &\equiv [I_m 0_{m,nm} I_m 0_{m,nm} I_m ,\ldots, 0_{m,nm} I_m]. \end{aligned}$$

Note that Γ is an $n^2 m \times m$ matrix.

Thus, (17), (23), and (24) yield

$$\begin{aligned} J_{N-1} &= u(N-1)^{\mathrm{T}}\, [\Lambda(N-1) + \theta(N-1)]\, u(N-1) \\ &\quad + 2w(N-1)^{\mathrm{T}}\, u(N-1) + \alpha(N-1) + \beta(N-1), \end{aligned} \tag{26}$$

where

$$\theta(N-1) \equiv \Gamma^{\mathrm{T}}[S(N) \otimes M_2(N-1 \mid N-1)]\, \Gamma, \tag{27}$$

$$w(N-1) \equiv \Gamma^{\mathrm{T}} \operatorname{cs}[M_3(N-1 \mid N-1)\, A^{\mathrm{T}} S(N)]. \tag{28}$$

Therefore, the optimal control law $u^*(N-1)$ and the associated cost-to-go are given by

$$u^*(N-1) = -[\Lambda(N-1) + \theta(N-1)]^{-1}\, w(N-1), \tag{29}$$

$$\begin{aligned} J^*_{N-1} &= -w(N-1)^{\mathrm{T}}\, [\Lambda(N-1) + \theta(N-1)]^{-1}\, w(N-1) + \alpha(N-1) \\ &\quad + \beta(N-1). \end{aligned} \tag{30}$$

Note that $\theta(N-1) \geqslant 0$ a.s., since $S(N) \geqslant 0$ and $M_2(N-1 \mid N-1) > 0$ a.s. (see Lemma A3 in Appendix I). Hence $\Lambda(N-1) + \theta(N-1) > 0$ and invertible a.s., since $\Lambda(N-1) > 0$.

Stage $k = N-2$. The functional relation (16) yields

$$\begin{aligned} J_{N-2} &= E[J(N-2) + J^*_{N-1} \mid Y(N-2)] \\ &= E[-w(N-1)^{\mathrm{T}}\, \{\Lambda(N-1) + \theta(N-1)\}^{-1}\, w(N-1) \mid Y(N-2)] \\ &\quad + E[J(N-2) + \alpha(N-1) \mid Y(N-2)] + \beta(N-1). \end{aligned} \tag{31}$$

Since $Y(N-1) = \{Y(N-2), y(N-1)\}$, from (18),

$$\begin{aligned} &E[\alpha(N-1) \mid Y(N-2)] \\ &\quad = E[E\{x(N-1)^{\mathrm{T}} A^{\mathrm{T}} S(N) A x(N-1) \mid Y(N-1)\} \mid Y(N-2)] \\ &\quad = E[x(N-1)^{\mathrm{T}} A^{\mathrm{T}} S(N) A x(N-1) \mid Y(N-2)]. \end{aligned}$$

Therefore, it is straightforward to obtain

$$\begin{aligned} J_{N-2}^{\mathrm{U}} &\equiv E[J(N-2) + \alpha(N-1) \mid Y(N-2)] + \beta(N-1) \\ &= u(N-2)^{\mathrm{T}} [\Lambda(N-2) + \theta_{\mathrm{U}}(N-2)]\, u(N-2) \\ &\quad + 2w_{\mathrm{U}}(N-2)^{\mathrm{T}} u(N-2) + \alpha_{\mathrm{U}}(N-2) + \beta_{\mathrm{U}}(N-2), \end{aligned} \tag{32}$$

where

$$\theta_{\mathrm{U}}(N-2) \equiv \Gamma^{\mathrm{T}}[V_{\mathrm{U}}(N-2) \otimes M_2(N-2 \mid N-2)]\, \Gamma, \tag{33}$$

$$w_{\mathrm{U}}(N-2) \equiv \Gamma^{\mathrm{T}} \operatorname{cs}[M_3(N-2 \mid N-2)\, A^{\mathrm{T}} V_{\mathrm{U}}(N-2)], \tag{34}$$

$$\alpha_{\mathrm{U}}(N-2) \equiv \operatorname{tr}\{A^{\mathrm{T}} V_{\mathrm{U}}(N-2)\, A E[x(N-2)\, x(N-2)^{\mathrm{T}} \mid Y(N-2)]\}, \tag{35}$$

$$\beta_{\mathrm{U}}(N-2) \equiv \beta(N-1) + \operatorname{tr}[D^{\mathrm{T}} S(N-1)\, D Q(N-2)], \tag{36}$$

$$V_{\mathrm{U}}(N-2) \equiv S(N-1) + A^{\mathrm{T}} S(N)\, A. \tag{37}$$

The difficulty in optimization lies in evaluating the first term in (31), since $\theta(N-1)$ and $w(N-1)$ are the complicated random matrix and vector, respectively, depending on $u(N-2)$. In this control law the term is neglected in order to simplify the backward optimization. Note that the term is nonpositive a.s., since $\Lambda(N-1) + \theta(N-1) > 0$ a.s. This term originates from the first two terms in (26) (with the optimal law $u^*(N-1)$ in (29)), and accounts for the amount of reduced cost due to the control at stage $N-1$. Hence the omission of this term means that the control law at $N-2$ is designed by neglecting the control effect at $N-1$ (the term $E[\alpha(N-1) \mid Y(N-2)]$ which is not neglected accounts for the cost due to the free motion from $N-1$ to N). Although this approximation may seem somewhat ad hoc, the resulting control law requires little on-line computation and shows good performance in the simulated examples, as will be observed in Section 6.

With the above simplification, we have control law $u_{\mathrm{U}}(N-2)$ which minimizes (32) and the associated cost-to-go functional J_{N-2}^{U}:

$$u_{\mathrm{U}}(N-2) = -[\Lambda(N-2) + \theta_{\mathrm{U}}(N-2)]^{-1}\, w_{\mathrm{U}}(N-2), \tag{38}$$

$$\begin{aligned} I_{N-1}^{*} \leqslant J_{N-2}^{\mathrm{U}} &= -w_{\mathrm{U}}(N-2)^{\mathrm{T}} [\Lambda(N-2) + \theta_{\mathrm{U}}(N-2)]^{-1}\, w_{\mathrm{U}}(N-2) \\ &\quad + \alpha_{\mathrm{U}}(N-2) + \beta_{\mathrm{U}}(N-2). \end{aligned} \tag{39}$$

Algorithm for UEFC

By proceeding with the simplification described for stage $N-2$, we obtain the control law for a general stage k:

$$u_U(k) = -[\Lambda(k) + \theta_U(k)]^{-1} w_U(k), \tag{40}$$

where

$$\theta_U(k) \equiv \Gamma^T[V_U(k) \otimes M_2(k \mid k)]\, \Gamma \geqslant 0, \qquad \text{a.s.}, \tag{41}$$

$$w_U(k) \equiv \Gamma^T \operatorname{cs}[M_3(k \mid k)\, A^T V_U(k)], \tag{42}$$

$$V_U(k) = S(k+1) + A^T V_U(k+1)\, A, \qquad k = N-1,\dots, 0, \tag{43}$$
$$V_U(N) \equiv 0,$$

and Γ and $M_i(k \mid k)$ are defined by (25) and (22), respectively.

Remarks. 1. Since $V_U(k)$ can be computed off-line by (43), this control law requires no on-line recursive computation, but computation of only $\theta_U(k)$ and $w_U(k)$ to obtain $u_U(k)$.

2. Note that $\theta_U(k)$ and $w_U(k)$ are functions of $\pi_2(k \mid k)$ and $\pi_3(k \mid k)$, measures of estimation error, as well as $\hat{x}(k \mid k)$ and $\hat{b}(k \mid k)$ (see Eq. (22)). In this sense UEFC is cautious like OLOF [3].

3. As mentioned above for stage $N-2$, $\Lambda(k) + \theta_U(k) > 0$ and invertible a.s., hence (40) provides a well-defined control law a.s.

5.2. *Control Law Based on Overestimating Future Control Efforts* (OEFC)

Stage $k = N-2$. The UEFC law was obtained by neglecting the term due to the control efforts at stage $N-1$ because of the difficulty in approximating the term in a simple manner. Here we bound the term, the first (negative) term in (31), from below, thereby obtaining a control law (OEFC) by overestimating the control efforts at stage $N-1$.

LEMMA. *The first term in* (31) *can be bounded as*

$$\begin{aligned} &-w(N-1)^T\,[\Lambda(N-1) + \theta(N-1)]^{-1}\, w(N-1) \\ &\quad \geqslant -\operatorname{tr}\{[\Lambda(N-1) + \theta(N-1)]^{-1}\, \theta(N-1)\}\, \alpha(N-1) \qquad \text{a.s.} \end{aligned} \tag{44}$$

Proof. Using $S(N) \geqslant 0$ and $M(N-1 \mid N-1) > 0$ in Lemma A4 in Appendix I, we have

$$\begin{bmatrix} S(N) \otimes M_1(N-1 \mid N-1) & S(N) \otimes M_3(N-1 \mid N-1)^T \\ S(N) \otimes M_3(N-1 \mid N-1) & S(N) \otimes M_2(N-1 \mid N-1) \end{bmatrix} \geqslant 0. \tag{45}$$

We define

$$\Psi \equiv \begin{bmatrix} \operatorname{tr}\{A^{\mathrm{T}}S(N)\,AM_1(N-1 \mid N-1)\} & [\operatorname{cs}\{M_3(N-1 \mid N-1)\,A^{\mathrm{T}}S(N)\}]^{\mathrm{T}} \\ \operatorname{cs}\{M_3(N-1 \mid N-1)\,A^{\mathrm{T}}S(N)\} & S(N) \otimes M_2(N-1 \mid N-1) \end{bmatrix}, \tag{46}$$

then

$$\begin{aligned} \alpha(N-1) &= \operatorname{tr}[A^{\mathrm{T}}S(N)\,AM_1(N-1 \mid N-1)] = \operatorname{tr}(M_1A^{\mathrm{T}}SA) \\ &= \operatorname{tr}(SAM_1A^{\mathrm{T}}) = [\operatorname{cs}(A^{\mathrm{T}})]^{\mathrm{T}}\,(S \otimes M_1)\operatorname{cs}(A^{\mathrm{T}}), \end{aligned}$$

where Lemma A2 was used to obtain the last equality. Also from Lemma A1

$$\operatorname{cs}[M_3(N-1 \mid N-1)\,A^{\mathrm{T}}S(N)].$$

Therefore,

$$\begin{aligned} \Psi &= \begin{bmatrix} \{\operatorname{cs}(A^{\mathrm{T}})\}^{\mathrm{T}}\,(S \otimes M_1)\operatorname{cs}(A^{\mathrm{T}}) & \{(S \otimes M_3)\operatorname{cs}(A^{\mathrm{T}})\}^{\mathrm{T}} \\ (S \otimes M_3)\operatorname{cs}(A^{\mathrm{T}}) & S \otimes M_2 \end{bmatrix} \\ &= \begin{bmatrix} \{\operatorname{cs}(A^{\mathrm{T}})\}^{\mathrm{T}} & 0_{1,n^2m} \\ 0_{n^2m,n^2} & I_{n^2m} \end{bmatrix} \begin{bmatrix} S \otimes M_1 & S \otimes M_3{}^{\mathrm{T}} \\ S \otimes M_3 & S \otimes M_2 \end{bmatrix} \begin{bmatrix} \operatorname{cs}(A^{\mathrm{T}}) & 0_{n^2,n^2m} \\ 0_{n^2m,1} & I_{n^2m} \end{bmatrix}. \end{aligned}$$

Hence, on noting (45), we have

$$\Psi \geqslant 0$$

and an application of Lemma A4 to (46) yields

$$\begin{aligned} &\operatorname{cs}[M_3(N-1 \mid N-1)\,A^{\mathrm{T}}S(N)]\,\{\operatorname{cs}[M_3(N-1 \mid N-1)\,A^{\mathrm{T}}S(N)]\}^{\mathrm{T}} \\ &\quad \leqslant \alpha(N-1)\,[S(N) \otimes M_2(N-1 \mid N-1)]. \end{aligned} \tag{47}$$

Thus, from (28)

$$\begin{aligned} &w(N-1)^{\mathrm{T}}\,[\Lambda(N-1) + \theta(N-1)]^{-1}\,w(N-1) \\ &\quad = \operatorname{tr}\{(\Lambda + \theta)^{-1}\,\Gamma^{\mathrm{T}}\operatorname{cs}(M_3A^{\mathrm{T}}S)\,[\operatorname{cs}(M_3A^{\mathrm{T}}S)]^{\mathrm{T}}\,\Gamma\} \\ &\quad \leqslant \operatorname{tr}\{(\Lambda + \theta)^{-1}\,\Gamma^{\mathrm{T}}\alpha(S \otimes M_2)\,\Gamma\} \\ &\quad = \operatorname{tr}\{[\Lambda(N-1) + \theta(N-1)]^{-1}\,\theta(N-1)\}\,\alpha(N-1), \end{aligned}$$

where (47), $\Lambda + \theta > 0$ a.s., and Lemma A5 were used to obtain the inequality. This completes the proof.

Using the above lemma and (31), we have a lower bound for J_{N-2}:

$$\begin{aligned} &J_{N-2} \\ &\quad \geqslant J^{\mathrm{U}}_{N-2} - E[\operatorname{tr}\{[\Lambda(N-1) + \theta(N-1)]^{-1}\,\theta(N-1)\}\,\alpha(N-1) \mid Y(N-2)], \end{aligned} \tag{48}$$

where $\theta(N-1)$ and $\alpha(N-1)$ are the random matrix and variable, respectively, given $Y(N-2)$, and no simple expression is available for the second term. As can be observed in (27), $\theta(N-1)$ is a function of $M_2(N-1 \mid N-1)$, the estimate of bb^T (a constant random matrix) at $N-1$. In order to proceed with the analysis in a simple manner, $\theta(N-1)$ is replaced by its estimate,

$$\begin{aligned}\hat{\theta}(N-1 \mid N-2)& \\ &= E[\theta(N-1) \mid Y(N-2)] = \Gamma^T[S(N) \otimes M_2(N-2 \mid N-2)]\,\Gamma,\end{aligned} \tag{49}$$

which is a function of $Y(N-2)$. Therefore, (48) is approximated by

$$\begin{aligned}J^O_{N-2} \equiv J^U_{N-2} - \operatorname{tr}\{&[\Lambda(N-1) + \hat{\theta}(N-1 \mid N-2)]^{-1} \\ &\times \hat{\theta}(N-1 \mid N-2)\, E[\alpha(N-1) \mid Y(N-2)]\}.\end{aligned} \tag{50}$$

For (32)–(37) and (50), we have the following cost-to-go expression for OEFC:

$$\begin{aligned}J^O_{N-2} = u(N-2)^T &[\Lambda(N-2) + \theta_O(N-2)]\, u(N-2) \\ &+ 2w_O(N-2)^T u(N-2) + \alpha_O(N-2) + \beta_O(N-2),\end{aligned} \tag{51}$$

where

$$\theta_O(N-2) \equiv \Gamma^T[V_0(N-2) \otimes M_2(N-2 \mid N-2)]\,\Gamma, \tag{52}$$

$$w_O(N-2) \equiv \Gamma^T \operatorname{cs}[M_3(N-2 \mid N-2)\, A^T V_O(N-2)], \tag{53}$$

$$\alpha_O(N-2) \equiv \operatorname{tr}[A^T V_O(N-2)\, A M_1(N-2 \mid N-2)], \tag{54}$$

$$\begin{aligned}\beta_O(N-2) &\equiv \beta_U(N-2), \\ V_O(N-2) &\equiv S(N-1) + \epsilon(N-1 \mid N-2)\, A^T S(N)\, A,\end{aligned} \tag{55}$$

$$\begin{aligned}\epsilon(N-1 \mid N-2)& \\ &\equiv 1 - \operatorname{tr}\{[\Lambda(N-1) + \hat{\theta}(N-1 \mid N-2)]^{-1}\, \hat{\theta}(N-1 \mid N-2)\}.\end{aligned} \tag{56}$$

Therefore, the control law OEFC which minimizes J^O_{N-2} is given by

$$u_O(N-2) = -[\Lambda(N-2) + \theta_O(N-2)]^{-1}\, w_O(N-2), \tag{57}$$

$$\begin{aligned}J^O_{N-2} = -w_O(N-2)^T &[\Lambda(N-2) + \theta_O(N-2)]^{-1}\, w_O(N-2) \\ &+ \alpha_O(N-2) + \beta_O(N-2).\end{aligned} \tag{58}$$

Algorithm for OEFC

Since expression (58) for J^O_{N-2} has the same quadratic form as (39) for J^U_{N-2}, it is easy to obtain the OEFC control law for a general stage k:

$$u_0(k) = -[\Lambda(k) + \theta_0(k)]^{-1}\, w_0(k), \tag{59}$$

where

$$\theta_0(k) \equiv \Gamma^{\mathrm{T}}[V_0(k) \otimes M_2(k \mid k)]\, \Gamma, \tag{60}$$

$$w_0(k) \equiv \Gamma^{\mathrm{T}}\, \mathrm{cs}[M_3(k \mid k)\, A^{\mathrm{T}} V_0(k)]. \tag{61}$$

The matrix $V_0(k)$ is computed by the following (backward) recursive formula:

$$V(i \mid k) = S(i+1) + \epsilon(i+1 \mid k)\, A^{\mathrm{T}} V(i+1 \mid k)\, A, \qquad i = N-1, N-2, \ldots, k \tag{62}$$

$$V_0(k) \equiv V(k \mid k), \qquad V(N \mid k) \equiv 0_{n,n}, \tag{63}$$

$$\epsilon(i+1 \mid k) \equiv 1 - \mathrm{tr}\{[\Lambda(i+1) + \hat{\theta}(i+1 \mid k)]^{-1}\, \hat{\theta}(i+1 \mid k)\}, \tag{64}$$

$$\hat{\theta}(i+1 \mid k) \equiv \Gamma^{\mathrm{T}}[V(i+1 \mid k) \otimes M_2(k \mid k)]\, \Gamma. \tag{65}$$

Remarks. 1. The OEFC algorithm has the same structure as the UEFC law given by (40)–(43), where $\epsilon(i+1 \mid k) \equiv 1$ (compare (43) with (62)).

2. The OEFC law requires more on-line computation than the UEFC law, since $\hat{\theta}(i+1 \mid k)$ depends on $M_2(k \mid k) = E[bb^{\mathrm{T}} \mid Y(k)]$ and (62) must be recursively computed for each stage k.

6. Examples

A computer simulation study was performed to evaluate the performance of the UEFC and OEFC control laws. The two systems selected are single-input third-order systems, and are essentially the same as those in [6]; one is a stable system and the other is an unstable system. The performance of the laws for Monte Carlo runs is statistically compared with the certainty equivalence law (CE) and the optimal control law when B is known (called the LQG algorithm—the solution of the standard LQG problem). The sensitivity of performance of the four algorithms is studied for various levels of initial uncertainties (P_b and P_0).

The system matrices common to the two systems are

$$C = [1 \quad 0 \quad 0], \qquad D^{\mathrm{T}} = [0.2 \quad 0.4 \quad 0.6],$$
$$\bar{x}_0^{\mathrm{T}} = [1 \quad 1 \quad 1], \qquad Q(k) = 0.01, \qquad R(k) = 0.09,$$
$$S(k+1) = I_3, \qquad \Lambda(k) = 1.$$

Unless specified otherwise, the processes simulated have 20 stages ($N = 19$) and the sample mean M_J of the performance measure $\sum_{k=0}^{N-1} J(k)$ has been computed for 20 Monte Carlo runs.

In this section we summarize the simulation results and discuss the properties of the UEFC and OEFC Laws. The reader is referred to [8] for more data and details.

6.1. *Stable System*

The system matrices are given by

$$A = \begin{bmatrix} 1 & 0.2 & 0.0 \\ 0 & 1.0 & 0.2 \\ -1 & -1.4 & 0.4 \end{bmatrix}, \qquad \bar{B} = \bar{b} = \begin{bmatrix} 0.0 \\ 0.0 \\ -0.4 \end{bmatrix},$$

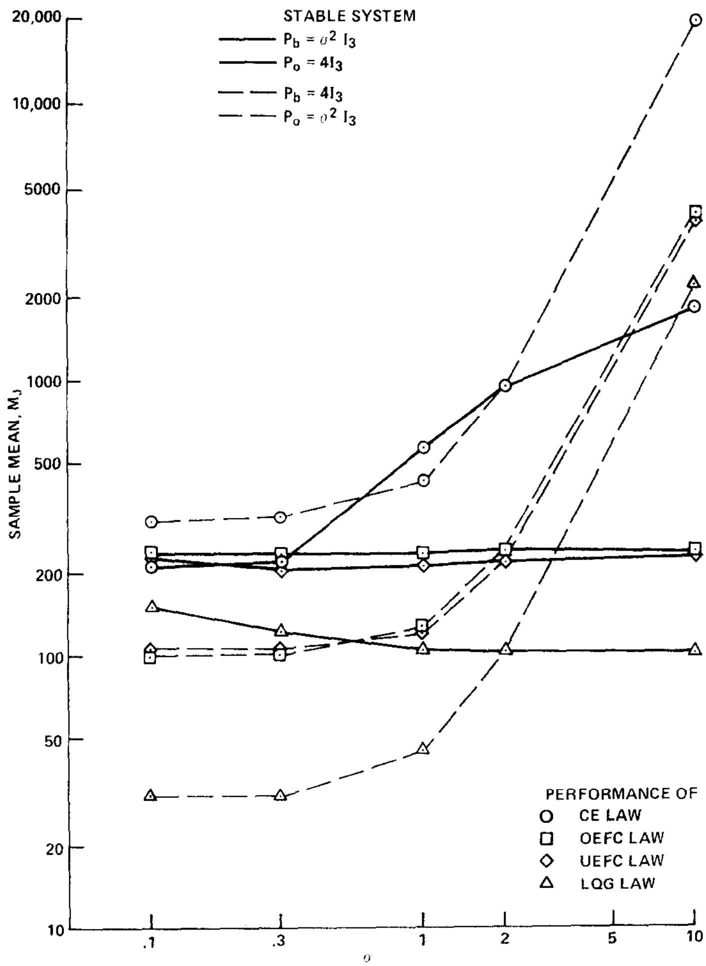

FIG. 1. Dependence of Monte Carlo performance on P_b and P_0 (sample mean M_J).

where A has eigenvalues 0.8 and $0.8 \pm 0.4j$. The sample mean M_J of the four algorithms (UEFC, OEFC, CE, and LQG) for the Monte Carlo runs is plotted in Fig. 1. The heavy lines in Fig. 1 correspond to M_J for $P_b \equiv \sigma^2 I_3$ with $P_0 = 4I_3$, and the dashed lines for $P_0 \equiv \sigma^2 I_3$ with $P_b = 4I_3$. The abscissa σ indicates the level of prior uncertainty for each case. For each of the 20 runs, $B = b$ and $x(0)$ are randomly generated by the distributions $b \sim N(\bar{b}, P_b)$ and $x(0) \sim N(\bar{x}_0, P_0)$. In order to see the normalized performance of the suboptimal laws, the ratio

$$r_J \equiv \frac{M_J \text{ for a suboptimal law}}{M_J \text{ for the LQG law}}$$

is plotted in Fig. 2 for various σ's.

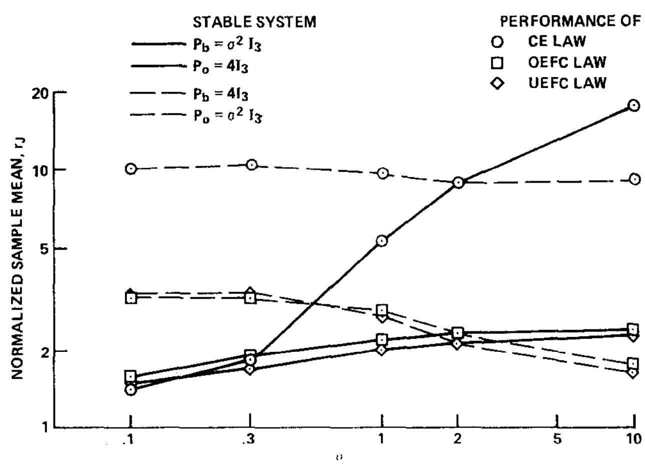

FIG. 2. Dependence of normalized sample mean r_J on P_b and P_0.

Observations. 1. In Fig. 1 the performance of UEFC and OEFC remains almost the same as σ increases for $P_b = \sigma^2 I_3$ with $P_0 = 4I_3$, whereas the CE performance becomes considerably worse (the heavy lines). This is to be expected, since both UEFC and OEFC take the error of estimates into consideration and are cautious in implementing control, while CE does not consider such uncertainty (see Remark 2 following Eq. (43)).

2. The normalized performance of the three suboptimal laws is rather insensitive to variations in P_0; however, r_J decreases slightly as P_0 increases (the dashed lines in Fig. 2). This is because the uncertainty in $\bar{x}_0$ (P_0), which is common to the four laws (including the LQG law), becomes comparatively more dominant than the uncertainty in $\bar{b}(P_b = 4I_3)$ as σ increases, and as a result the performance degradation due to unknown B tends to decrease.

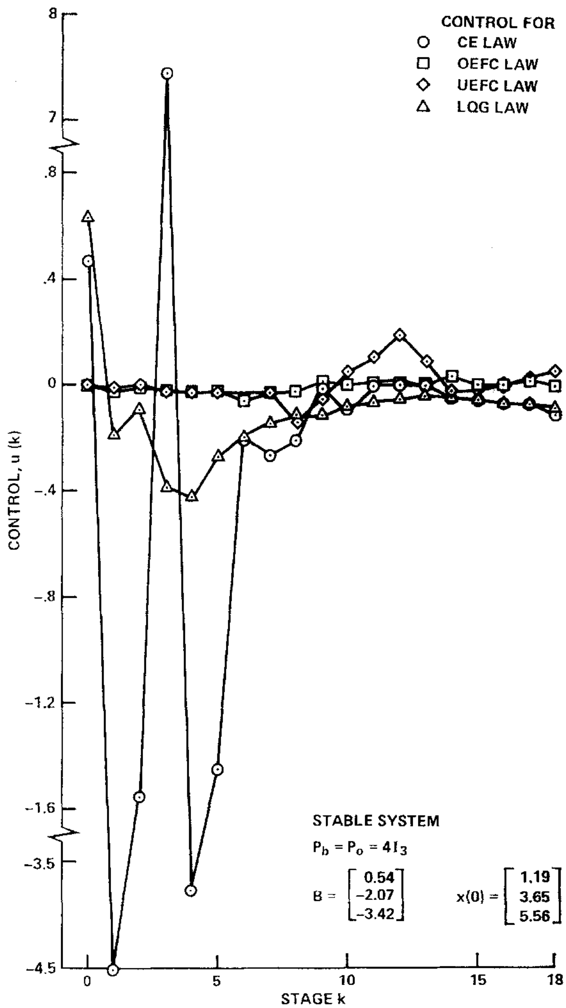

FIG. 3. Time history of control $u(k)$ for the four control laws.

3. Considering that the performance of the LQG law is impossible to attain and that the *optimal* law with unknown B is worse than the LQG law (the optimal law with known B), the performance of the UEFC and OEFC laws ($r_J \simeq 1.5$–3, Fig. 2) is good, especially since little on-line computation is required.

In order to study further the characteristics of the UEFC and OEFC laws, the time histories of the four laws for a representative run are plotted in

Figure 3: Control $u(k)$
Figure 4: Estimate $\hat{b}(k \mid k) \equiv [\hat{b}_1 \quad \hat{b}_2 \quad \hat{b}_3]^{\mathrm{T}}$

Figures 5–7: Estimate $\hat{x}(k \mid k) \equiv [\hat{x}_1 \quad \hat{x}_2 \quad \hat{x}_3]^T$

Figure 8: Instantaneous cost $J(k)$.

For this run $P_b = P_0 = 4I_3$, the true values of B and $x(0)$ are

$$B^T = [0.54 \quad -2.07 \quad -3.42], \qquad x(0)^T = [1.19 \quad 3.65 \quad 5.56],$$

and the performance measure $\sum_{k=0}^{N-1} J(k)$ is 404, 787, 880, and 4301 for the LQG, UEFC, OEFC, and CE laws, respectively.

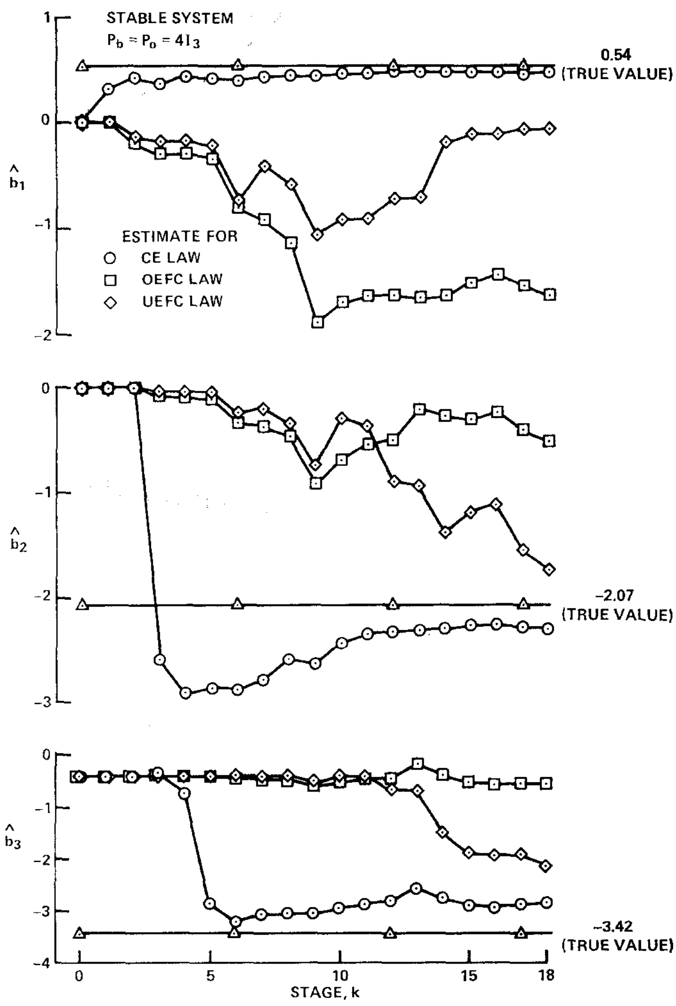

FIG. 4. Time history of estimate $\hat{b}(k \mid k)$ for the four control laws.

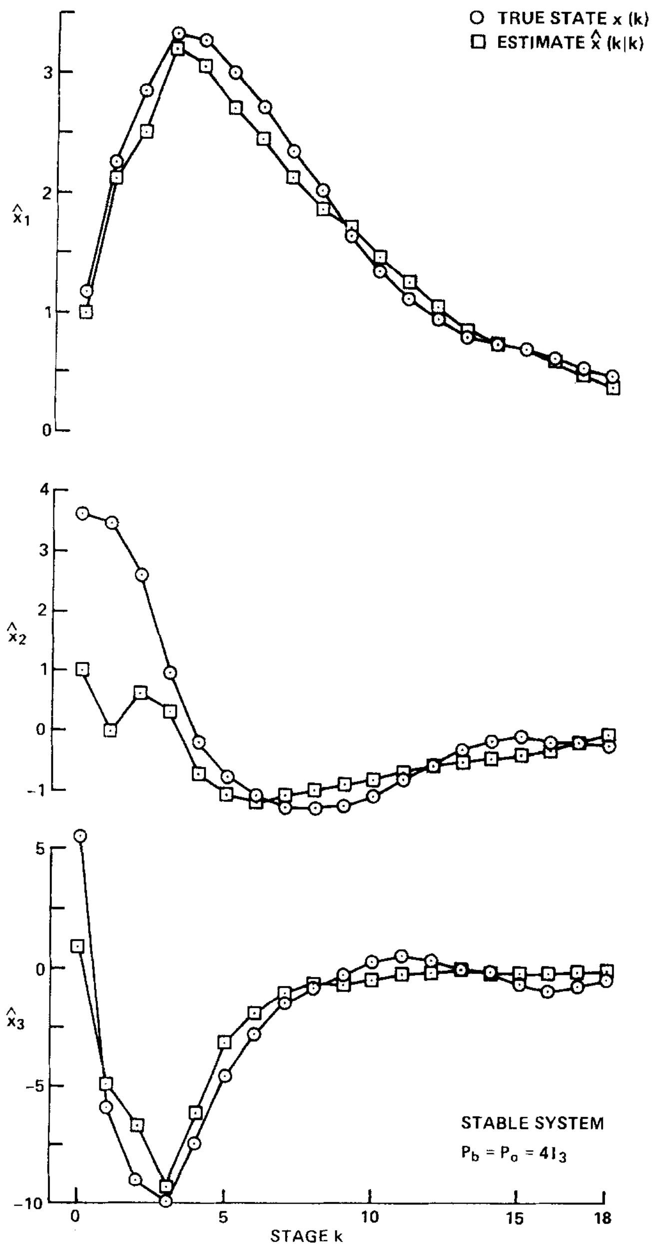

FIG. 5. Time history of true state $x(k)$ and estimate $\hat{x}(k \mid k)$ for LQG law.

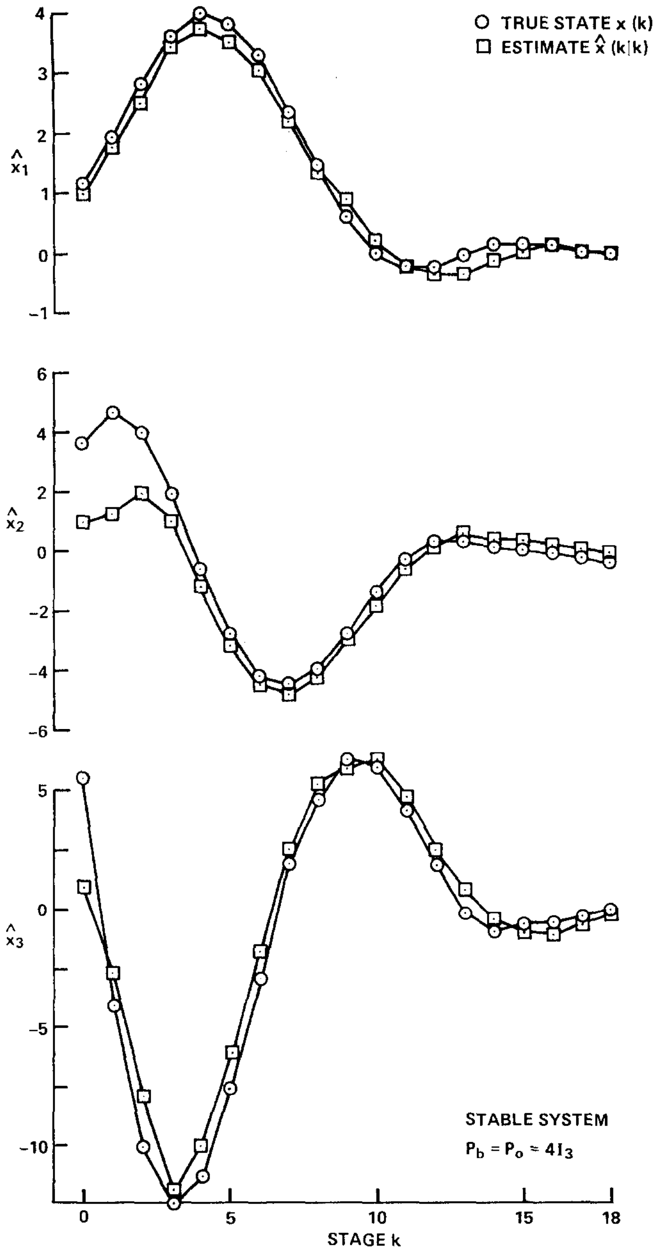

FIG. 6. Time history of true state $x(k)$ and estimate $\hat{x}(k \mid k)$ for UEFC law.

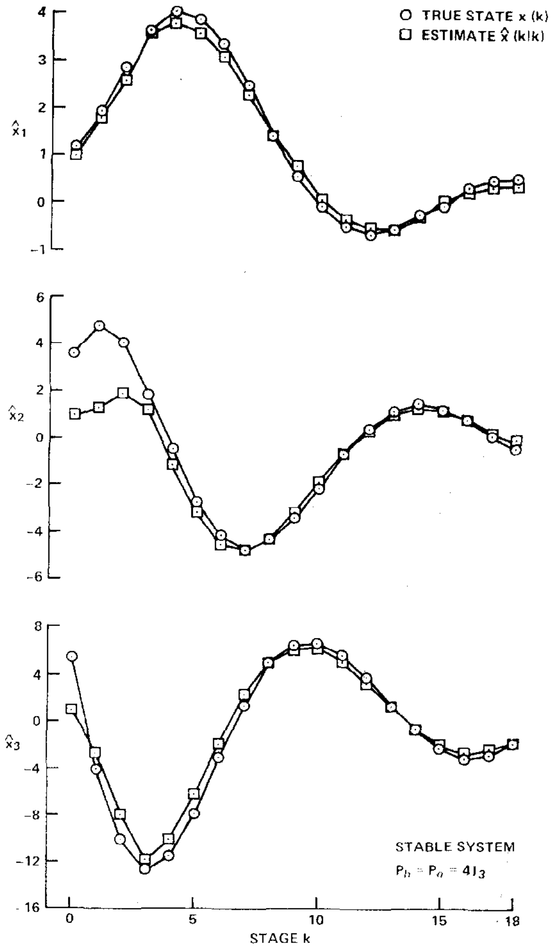

FIG. 7. Time history of true state $x(k)$ and estimate $\hat{x}(k \mid k)$ for OEFC law.

Observations. The characteristics of the three suboptimal laws are clearly shown in these figures. The CE law erroneously exerts large control in the beginning ($k = 0 - 5$ in Fig. 3), thereby incurring large costs (Fig. 8). The large control accidentally results in fast learning of B (Fig. 4), and less cost $J(k)$ than the UEFC and OEFC laws at later stages ($k \geqslant 7$). Both UEFC and OEFC are cautious and very little control energy is implemented in the beginning ($k \leqslant 7$ in Fig. 3), when larger estimation errors are expected (see Remark 2 following Eq. (43)). Since UEFC underestimates future control efforts, it is less cautious than OEFC and exerts more control at $k = 8 - 14$ than OEFC, thereby attaining better cost (Fig. 8) and better estimate $\hat{b}(k \mid k)$ (Fig. 4). Note that the estimation of $x(k)$ for UEFC and OEFC is very good (compare Figs. 6 and 7 with Fig. 5), although the estimate $\hat{b}(k \mid k)$ is not as good as CE.

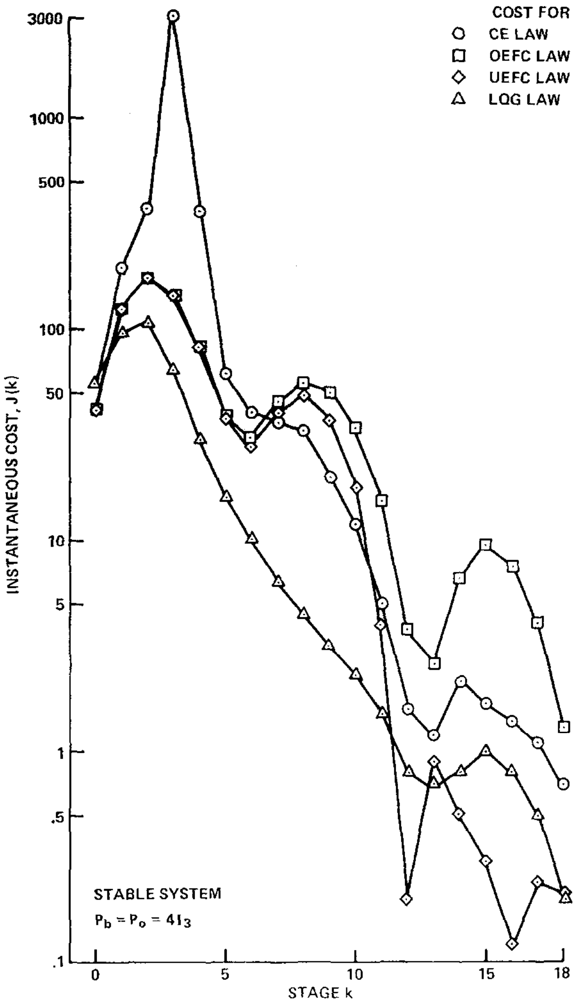

FIG. 8. Time history of instantaneous cost $J(k)$ for the four control laws.

Since the fast learning property of the CE law was observed for unknown B, additional simulations were performed with various numbers of stages (N varied) to compare the performance of the four laws. The normalized performance r_j of the simulations is plotted in Fig. 9 for various N's.

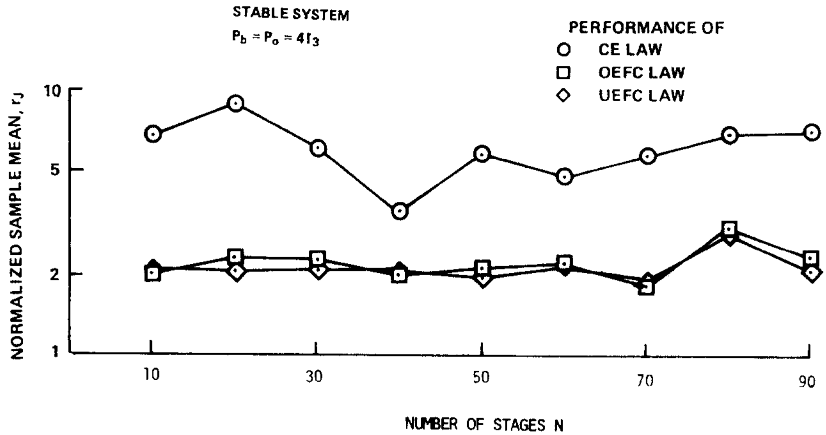

FIG. 9. Dependence of normalized sample mean r_J on the number of stages N.

Observation. The comparative performance of the three suboptimal laws in Fig. 9 does not vary significantly for different numbers of stages, and it is evident that even for large N's the CE law does not perform as well as the UEFC and OEFC laws. This observation suggests that the fast learning of B of the CE law in Fig. 4 is accidental and does not pay off even for large N's.

6.2. *Unstable System*

The system matrices are given by

$$A = \begin{bmatrix} 1 & 0.2 & 0.0 \\ 0 & 1.0 & 0.2 \\ 1 & -0.6 & 0.8 \end{bmatrix}, \qquad \bar{B} = \bar{b} = \begin{bmatrix} 0.0 \\ 0.0 \\ -0.2 \end{bmatrix},$$

where A has eigenvalues 1.2 and $0.8 \pm 0.4j$. As for the stable system, the performance of the four algorithms for 20 Monte Carlo runs is plotted in Figs. 10 and 11: The sample mean M_J in Fig. 10 and the normalized sample mean r_J in Fig. 11 for various P_b's and P_0's.

Observations. 1. The characteristics of the three suboptimal laws are very similar to those observed for the stable system.

2. The performance of the OEFC law is somewhat worse than that in the stable case, whereas the UEFC law performs consistently well. The CE law performs better than the cautious OEFC and UEFC laws for small P_b ($\sigma = 0.1$ and 0.3; i.e., when there is little uncertainty in $\bar{b}$).

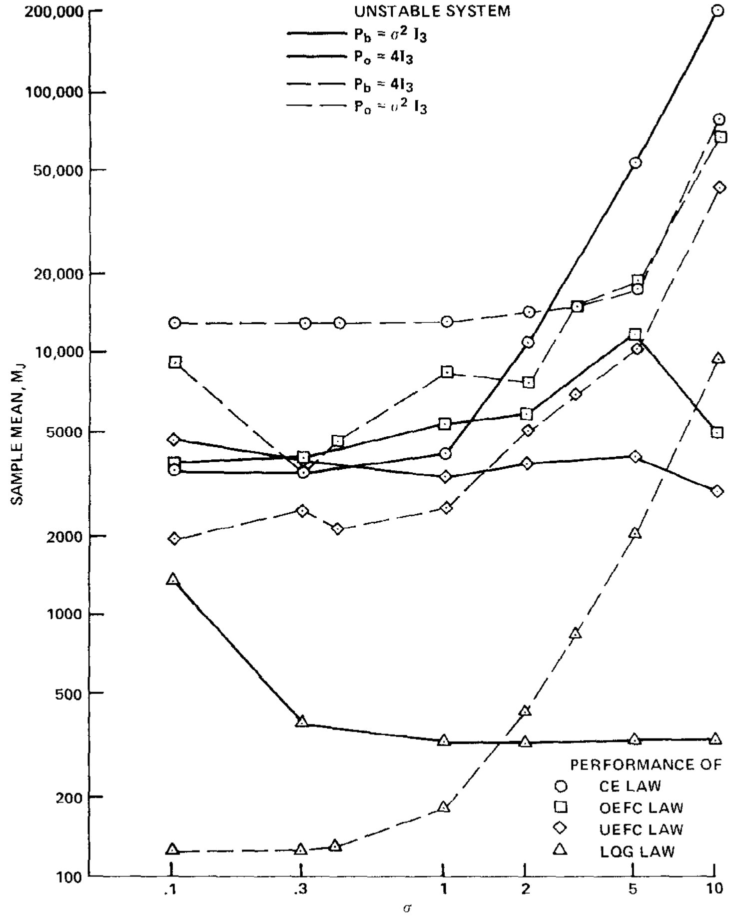

FIG. 10. Dependence of Monte Carlo performance on P_b and P_0 (sample mean M_J).

7. Conclusions

We have considered a discrete-time linear stochastic adaptive control system with unknown control gain matrix (B). Two suboptimal control laws have been derived: the UEFC law based on the underestimation of future control and the OEFC law based on the overestimation on future control. These laws require little on-line computation and at the same time incorporate some information on the estimation errors, hence they are in the category of "cautious" controls as classified by Wittenmark [3]. Two single-input third-order systems have been simulated to compare the Monte Carlo performance of the laws with

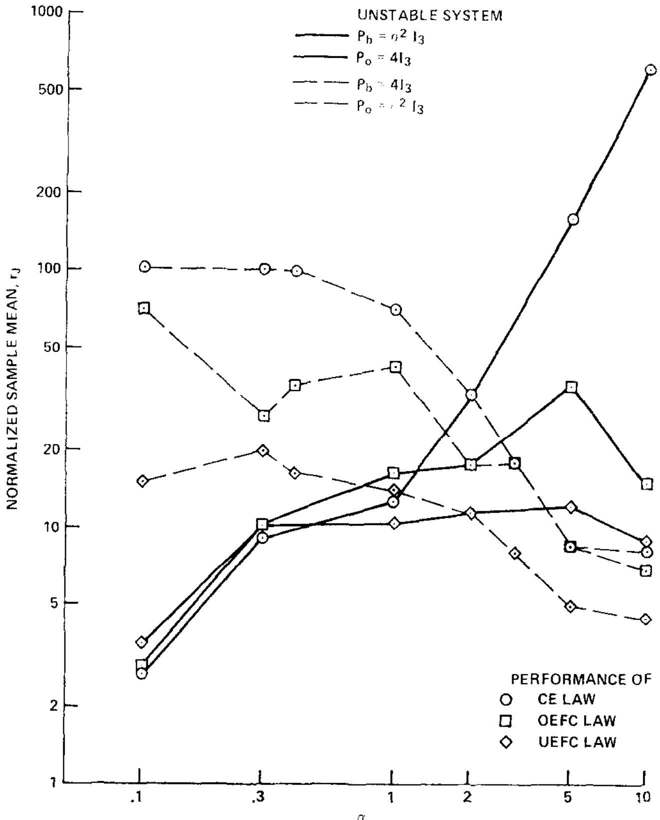

FIG. 11. Dependence of normalized sample mean r_j on P_b and P_0.

that of the CE and LQG laws. The dependence of the performance of the four laws on P_b and P_0 (the initial uncertainties on the state x and the control gain B) has been studied. The results indicate that the UEFC and OEFC laws perform much better than the CE law with only a little extra computation being required.

Admittedly, the class of systems considered in this study is small. However, the UEFC and OEFC laws derived for this class are conceptually simple and computationally efficient, and may provide a suitable framework for treating the more general class, where the system matrix (A) as well as the control gain matrix (B) are unknown. Further research is envisaged in this direction.

APPENDIX I

The identities and inequalities used to derive the estimation and control laws in the preceding sections are collected and proved where necessary. The matrices involved in the following lemmas are assumed to be conformable.

LEMMA A1.

$$\mathrm{cs}(ABC) = (C^{\mathrm{T}} \otimes A)\,\mathrm{cs}(B), \tag{A1}$$

$$\mathrm{rs}(ABC) = (A \otimes C^{\mathrm{T}})\,\mathrm{rs}(B). \tag{A2}$$

LEMMA A2.

$$\mathrm{tr}(AB) = [\mathrm{cs}(A^{\mathrm{T}})]^{\mathrm{T}}\,\mathrm{cs}(B), \tag{A3}$$

$$\mathrm{tr}(AC^{\mathrm{T}}BC) = [\mathrm{cs}(C)]^{\mathrm{T}}\,(A \otimes B^{\mathrm{T}})\,\mathrm{cs}(C). \tag{A4}$$

For the proofs of (A1), (A3), and (A4), the reader is referred to [9]. The identity (A1) is due to [10]. The proof of (A2) is straightforward and is omitted

LEMMA A3. *If*

$$A \geqslant 0 \quad \textit{and} \quad B \geqslant 0, \qquad \textit{then} \qquad A \otimes B \geqslant 0. \tag{A5}$$

If

$$A > 0 \quad \textit{and} \quad B > 0, \qquad \textit{then} \qquad A \otimes B > 0. \tag{A6}$$

Proof. Since A and B are symmetric, $A \otimes B$ is symmetric. The eigenvalues of $A \otimes B$ are $\lambda_i \mu_j$, where λ_i and μ_j are the eigenvalues of A and B, respectively [11, p. 235]. Since $A \geqslant 0$ and $B \geqslant 0$, $\lambda_i \geqslant 0$ and $\mu_j \geqslant 0$, hence

$$\lambda_i \mu_j \geqslant 0, \qquad \forall i, j.$$

This implies that $A \otimes B \geqslant 0$. The proof of (A6) is similar.

LEMMA A4. *If $A \geqslant 0$ and $B \equiv \left[\begin{smallmatrix} B_1 & B_3^{\mathrm{T}} \\ B_3 & B_2 \end{smallmatrix}\right] > 0$, where B_1 and B_2 are square matrices of dimensions m and l, respectively, then*

$$B_2 - B_3 B_1^{-1} B_3^{\mathrm{T}} > 0 \tag{A7}$$

and

$$C \equiv \begin{bmatrix} A \otimes B_1 & A \otimes B_3^{\mathrm{T}} \\ A \otimes B_3 & A \otimes B_2 \end{bmatrix} \geqslant 0. \tag{A8}$$

If $B \geqslant 0$ and B_1 is a scalar, then

$$B_1 B_2 \geqslant B_3 B_3^{\mathrm{T}}. \tag{A9}$$

Proof. Since $B > 0$, $B_1 > 0$ and invertible,

$$B = \begin{bmatrix} I_m & 0_{m,l} \\ B_3 B_1^{-1} & I_l \end{bmatrix} \begin{bmatrix} B_1 & 0_{m,l} \\ 0_{l,m} & B_2 - B_3 B_1^{-1} B_3^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} I_m & B_1^{-1} B_3^{\mathrm{T}} \\ 0_{l,m} & I_l \end{bmatrix} > 0,$$

which implies that $B_2 - B_3 B_1^{-1} B_3^{\mathrm{T}} > 0$.

For the case $B \geqslant 0$, (A9) clearly holds if $B_1 = 0$. If $B_1 > 0$, we obtain $B_2 - B_3 B_1^{-1} B_3^{\mathrm{T}} \geqslant 0$, which implies (A9).

To prove (A8) we assume that A is an n-dimensional matrix and let $\hat{A} = A + \epsilon I_n$; then from (A6) $\hat{A} \otimes B_1 > 0$ and is invertible, since $\hat{A} > 0$ and $B_1 > 0$. Therefore,

$$\begin{aligned} \hat{C} &\equiv \begin{bmatrix} \hat{A} \otimes B_1 & \hat{A} \otimes B_3^{\mathrm{T}} \\ \hat{A} \otimes B_3 & \hat{A} \otimes B_2 \end{bmatrix} \\ &= D \begin{bmatrix} \hat{A} \otimes B_1 & 0_{nm,nl} \\ 0_{nl,nm} & \hat{A} \otimes B_2 - (\hat{A} \otimes B_3)(\hat{A} \otimes B_1)^{-1} (\hat{A} \otimes B_3^{\mathrm{T}}) \end{bmatrix} D^{\mathrm{T}}, \end{aligned} \tag{A10}$$

where

$$D \equiv \begin{bmatrix} I_{nm} & 0_{nm,nl} \\ (\hat{A} \otimes B_3)(\hat{A} \otimes B_1)^{-1} & I_{nl} \end{bmatrix}.$$

Using identities for inverses and products of Kronecker products [9], we can easily write

$$(\hat{A} \otimes B_3)(\hat{A} \otimes B_1)^{-1} = I_n \otimes B_3 B_1^{-1},$$

$$\hat{A} \otimes B_2 - (\hat{A} \otimes B_3)(\hat{A} \otimes B_1)^{-1}(\hat{A} \otimes B_3^{\mathrm{T}}) = \hat{A} \otimes (B_2 - B_3 B_1^{-1} B_3^{\mathrm{T}}).$$

Therefore, from (A10)

$$\begin{aligned} C = \lim_{\epsilon \to 0} \hat{C} &= \begin{bmatrix} I_{nm} & 0_{nm,nl} \\ I_n \otimes B_3 B_1^{-1} & I_{nl} \end{bmatrix} \begin{bmatrix} A \otimes B_1 & 0_{nm,nl} \\ 0_{nl,nm} & A \otimes (B_2 - B_3 B_1^{-1} B_3^{\mathrm{T}}) \end{bmatrix} \\ &\times \begin{bmatrix} I_{nm} & I_n \otimes B_1^{-1} B_3^{\mathrm{T}} \\ 0_{nl,nm} & I_{nl} \end{bmatrix}. \end{aligned} \tag{A11}$$

From (A5) and (A7), $A \otimes B_1 \geqslant 0$ and $A \otimes (B_2 - B_3 B_1^{-1} B_3^{\mathrm{T}}) \geqslant 0$, hence (A11) implies $C \geqslant 0$.

LEMMA A5. *If $A \geqslant 0$, $B \geqslant C$, then*

$$\mathrm{tr}[AB] \geqslant \mathrm{tr}[AC]. \tag{A12}$$

Proof. $\mathrm{tr}(AB) - \mathrm{tr}(AC) = \mathrm{tr}[A(B - C)]$. Since $B \geqslant C$, $(B - C)^{1/2}$ exists and

$$\mathrm{tr}[A(B - C)] = \mathrm{tr}[(B - C)^{1/2} A (B - C)^{1/2}].$$

Since $A \geqslant 0$, $(B - C)^{1/2} A (B - C)^{1/2} \geqslant 0$. Consequently, $\mathrm{tr}[A(B - C)] \geqslant 0$.

APPENDIX II: List of Symbols

General

X^{T}	Transpose of a matrix X
$\mathrm{tr}(X)$	Trace of a square matrix X
I_n	n-Dimensional identity matrix
$0_{m,n}$	$m \times n$ Null matrix
$X > 0$	Matrix X is positive definite
$X \geqslant 0$	Matrix X is positive semidefinite
$X \otimes Y$	Kronecker product of X and Y
$\mathrm{rs}(X)$	Row string vector of X
$\mathrm{cs}(X)$	Column string vector of X
$E[x]$	Expectation of a random vector x
$E[x \mid Y]$	Conditional expectation of a random vector x given Y
$x \sim N(\bar{x}, X)$	x Has Gaussian distribution with mean $\bar{x}$ and covariance matrix X
a.s.	With probability 1 (almost surely)
CE	Control law with enforced certainty equivalence
OLOF	Open loop optimal feedback control law
LQG	Optimal control law when system parameters are known
UEFC	Control law underestimating future control
OEFC	Control law overestimating future control
Subscript or superscript "U" ("O")	Variables pertain to algorithms for UEFC (OEFC).

Symbols for Derivation

k	Time index
N	Number of stages
$x(k)$	n-Dimensional state vector at k
$u(k)$	m-Dimensional control vector at k
$y(k)$	l-Dimensional measurement vector at k
$\xi(k)$	r-Dimensional plant noise vector at k
$\eta(k)$	l-Dimensional measurement noise vector at k
A	$n \times n$ System matrix
B	$n \times m$ Control gain matrix
C	$l \times n$ Measurement matrix
D	$n \times r$ Plant noise gain matrix

b	Row string vector of B	
$\bar{b}$	A priori mean of b	
P_b	A priori covariance matrix of b	
$\bar{x}_0$	Mean of $x(0)$	
P_0	Covariance matrix of $x(0)$	
$Q(k)$	Covariance matrix of $\xi(k)$	
$R(k)$	Covariance matrix of $\eta(k)$	
J	Performance measure	
$J(k)$	Instantaneous cost at k	
$S(k+1)$	Positive semidefinite weighting matrix for $x(k+1)$	
$\Lambda(k)$	Positive definite weighting matrix for $u(k)$	
$Y(k)$	Measurement data up to time k	
$U(k-1)$	Past controls up to time $k-1$	
$z(k)$	$(n+nm)$-Dimensional augmented state vector	
$F(k)$	$(n+nm) \times (n+nm)$-Augmented state matrix	
G	$(n+nm) \times r$-Augmented plant noise gain matrix	
H	$l \times (n+nm)$-Augmented measurement matrix	
$\hat{z}(i \mid k)$	Conditional mean of $z(i)$ given $Y(k)$	
$P(i \mid k)$	Covariance of estimation error $\hat{z}(i \mid k) - z(i)$	
$\pi_j(i \mid k)$	Submatrix of $P(i \mid k)$ $(j = 1, 2, 3)$	
J_k	Cost-to-go at k	
J_k^*	Optimal cost-to-go	
$M(i \mid k)$	Conditional mean of $z(i)\,z(i)^{\mathrm{T}}$ given $Y(k)$	
$M_j(i \mid k)$	Submatrix of $M(i \mid k)$ $(j = 1, 2, 3)$	
Γ	$n^2m \times m$ Matrix composed of I_m and $0_{m,nm}$	
$\theta_{\mathrm{U}}(k)$	$m \times m$ Matrix	Variables for UEFC algorithm
$w_{\mathrm{U}}(k)$	m-Dimensional vector	Variables for UEFC algorithm
$V_{\mathrm{U}}(k)$	$n \times n$ Matrix	Variables for UEFC algorithm
$\alpha_{\mathrm{U}}(k)$	Scalar	Variables for UEFC algorithm
$\beta_{\mathrm{U}}(k)$	Scalar	Variables for UEFC algorithm
$\theta_0(k)$	$m \times m$ Matrix	Variables for UEFC algorithm
$w_0(k)$	m-Dimensional vector	Variables for OEFC algorithm
$V_0(k)$	$n \times n$ Matrix	Variables for OEFC algorithm
$\alpha_0(k)$	Scalar	Variables for OEFC algorithm
$\beta_0(k)$	Scalar	Variables for OEFC algorithm
$\epsilon(i+1 \mid k)$	Scalar	Variables for OEFC algorithm
$\hat{\theta}(i+1 \mid k)$	$m \times m$ Matrix	Variables for OEFC algorithm

Additional Symbols for Monte Carlo Simulations

M_J	Sample mean of performance measure
r_J	Normalized sample mean of suboptimal laws
σ	A priori standard deviation for $x(0)$ or b
$\hat{b}_j$	jth element of estimate $\hat{b}(k \mid k)$ $(j = 1, 2, 3)$
$\hat{x}_j$	jth element of estimate $\hat{x}(k \mid k)$ $(j = 1, 2, 3)$.

References

1. R. Bellman, "Adaptive Processes—A Guided Tour," Princeton Univ. Press, Princeton, N. J., 1961.
2. M. Aoki, "Optimization of Stochastic Systems," Academic Press, New York, 1967.
3. B. Wittenmark, Stochastic adaptive control methods: a survey, *Internat. J. Control* **21** (1975), 705–730.
4. E. Tse and Y. Bar-Shalom, An actively adaptive control for linear systems with random parameters via the dual control approach, *IEEE Trans. Automatic Control* **AC-18** (1973), 109–117.
5. S. Dreyfus, Some types of optimal control of stochastic systems, *SIAM J. Control* **2** (1964), 120–134.
6. E. Tse and M. Athans, Adaptive stochastic control for a class of linear systems, *IEEE Trans. Automatic Control* **AC-17** (1972), 38–51.
7. R. Ku and M. Athans, On the adaptive control of linear systems using the Open-Loop-Feedback-Optimal approach, *IEEE Trans. Automatic Control* **AC-18** (1973), 489–493.
8. M. Toda and R. V. Patel, "Algorithms for Adaptive Stochastic Control for a Class of Linear Systems," NASA TM-73,240, 1977.
9. H. Neudecker, Some theorems on matrix differentiation with special reference to Kronecker matrix products, *J. Amer. Statist. Assoc.* **64** (1969), 953–963.
10. D. H. Nissen, A note on the variance of a matrix, *Econometrica* **36** (1968), 603–604.
11. R. Bellman, "Introduction to Matrix Analysis," 2nd ed., McGraw-Hill, New York, 1970.