

Direct Updating of Intertemporal Criterion Functions for a Class of Adaptive Control Problems

LEIGH TESFATSION, MEMBER, IEEE

Abstract—Previous papers demonstrate the feasibility of direct criterion function updating for a class of adaptive control problems with sequential myopic expected cost objectives. In contrast, the present paper demonstrates the feasibility of direct criterion function updating for a class of adaptive control problems with intertemporal expected cost objectives. Specifically, the proposed criterion filter sequentially updates an initial estimate for expected cost plus cost-to-go in a dynamic programming framework. The control law generated by the filtered expected cost plus cost-to-go estimates is tested against the optimal control law for a linear-quadratic system with random state coefficients. Computer simulation results indicate that the total costs realized under the filter control law are approximately on par with the total costs realized under the optimal control law for the tested range of time horizons, cost function coefficients, and mean and standard deviation values for the random state coefficients.

I. INTRODUCTION

MANY CURRENTLY available state-space adaptive control techniques rely on the successive updating of state or state probability distribution estimates. Implementation of these techniques is often difficult, due to the extensive information and computation generally required for this updating (see [1], [2]).

For some systems, however, the estimation of states or state distributions has no intrinsic scientific value. The estimation is performed as an intermediate step towards optimal control selection. Ultimate interest focuses on system performance as measured by an expected cost criterion function. For such systems it may be conjectured that potentially significant reductions in computation and information requirements could be obtained if attention were to be shifted from the indirect updating of the criterion function, via state estimation or state distribution estimation, to the direct updating of the criterion function as the expectation of a random cost variable over an appropriate observation space.

The latter approach is investigated in [3]–[7] for a class of adaptive control problems with sequential myopic expected cost objectives. Instead of focusing on the state as a random control-dependent process whose characteristics are to be

estimated, the myopic cost function relevant for each period is viewed as a random function over an observation space, with control-conditioned expectation (the criterion function) to be directly estimated and updated in each period via a filtering operation on a vector of cost functions associated with past observations. Strong consistency is established for several specific linear criterion function filters, and small sample and asymptotic optimality properties are obtained for the generated control selections.

In contrast, the present paper will show that direct criterion filtering is also applicable for an interesting class of adaptive control problems with intertemporal expected cost objectives. The dynamic programming optimality (cost-to-go) equations can be directly updated in each period by means of a filtering operation on a vector of cost-to-go functions associated with past observations.

The organization of the paper is as follows. In Section II the intertemporal criterion-filtering approach is illustrated for a linear-quadratic control model with random state coefficients. In Section III a summary of results is included for a simulation study. The filter control law is shown to test well against the optimal control law for the model outlined in Section II, using a variety of different specifications for time-horizon, cost-function parameters, and state-coefficient mean and variance. An extension of criterion filtering methods to a more general class of adaptive control problems is outlined in Section IV, together with a variational principle for the adaptive updating of the criterion-filter weights. Concluding comments are given in Section V.

II. THE INTERTEMPORAL CRITERION-FILTERING APPROACH: AN ILLUSTRATIVE EXAMPLE

The intertemporal criterion-filtering approach is designed for the adaptive control of a certain class of nonlinear systems which have random state parameters exhibiting a possibly complicated dependence on time, control, and previous state values. For expositional clarity, the approach will be illustrated in the present section for a simple linear-quadratic control model with independent identically distributed state parameters. An extension of criterion filtering methods to the more general class of adaptive control problems will be discussed in Section V.

Consider the linear-quadratic control system described by equations

$$\begin{aligned} x_1 &= x_1^*, & \text{initial conditions;} \\ x_{n+1} &= \omega_n x_n + b_n \theta_n \equiv f_n(\omega_n, \theta_n, x_n), & 1 \leq n \leq N, \quad (1) \end{aligned}$$

Manuscript received June 19, 1978; revised October 30, 1978. This work was supported by the National Science Foundation under Grant ENG 77-28432. This paper is a revised version of "Direct Updating of Intertemporal Adaptive Control Criterion Functions," MRG Working Paper 7820, Department of Economics, University of Southern California.

The author is with the Department of Economics, University of Southern California, Los Angeles, CA 90007.

where $x_n \in R$ is the system state, $b_n \in R$ is a nonzero constant, $\theta_n \in R$ is the control, and the state coefficients $\omega_n \in R$ are independent and identically distributed drawings from a probability distribution $p: \mathcal{F} \rightarrow [0, 1]$ with finite first and second moments, defined over the σ -algebra \mathcal{F} consisting of all Borel subsets of R . The set \mathcal{L} of admissible feedback control laws for system (1) is assumed to be the set of all vectors $\theta = (\theta_1(\cdot), \dots, \theta_N(\cdot))$ of piecewise continuous functions $\theta_n: R \rightarrow R$. The control objective for system (1) is assumed to be the minimization of expected total cost:

$$E^N \left[\sum_{n=1}^N W_n(\omega_n, \theta_n(x_n), x_n) \right] \\ \equiv \int_{R^N} \left[\sum_{n=1}^N W_n(\omega_n, \theta_n(x_n), x_n) \right] p(d\omega_N) \cdots p(d\omega_1) \quad (2)$$

via selection of an admissible feedback control law $\theta \in \mathcal{L}$, where

$$W_n(\omega, \theta, x) \equiv c[\omega x + b_n \theta]^2 + q\theta^2 \quad (3)$$

for some constants $c > 0$ and $q \geq 0$. Intuitively, the controller's problem is to minimize the deviation of the state trajectory (x_1, \dots, x_N) from the null trajectory $(0, \dots, 0)$ by appropriate selection of possibly costly controls.

Letting $E[\cdot]$ denote expectation with respect to (R, \mathcal{F}, p) , a control law $\theta \in \mathcal{L}$ minimizes expected total cost (2) if and only if the following dynamic programming optimality equations are satisfied for almost every sequence $(\omega_1, \dots, \omega_N)$ [8; thm. 14.4, p. 101; thm. 15.2, p. 104; and thm. 17.6, p. 111]:

$$T_N^{\text{opt}}(x_N) \equiv \min_{\theta \in R} [E W_N(\omega_N, \theta, x_N)] = E[W_N(\omega_N, \theta_N(x_N), x_N)]; \\ T_n^{\text{opt}}(x_n) \equiv \min_{\theta \in R} [E W_n(\omega_n, \theta, x_n) + T_{n+1}^{\text{opt}} \circ f_n(\omega_n, \theta, x_n)] \\ = E[W_n(\omega_n, \theta_n(x_n), x_n) + T_{n+1}^{\text{opt}} \circ f_n(\omega_n, \theta_n(x_n), x_n)], \quad (4)$$

subject to

$$x_{n+1} = f_n(\omega_n, \theta_n(x_n), x_n), \quad 1 \leq n \leq N-1,$$

where $f_n(\cdot)$ is the state function defined by (1), and \circ denotes function composition. Assuming the distribution (R, \mathcal{F}, p) is known, (4) can in principle be solved to obtain the optimal control law.

Consider the more difficult problem in which the distribution (R, \mathcal{F}, p) is *a priori* unknown. The controller now faces an adaptive control problem, i.e., a problem of simultaneous learning and controlling. Many currently available state-space adaptive control techniques would focus on estimating the cost-to-go functions $T_n^{\text{opt}}(x)$, either by substituting estimates for the unknown state coefficients (certainty equivalence), or by substituting probability distribution estimates for the unknown underlying state coefficient probability distributions (Bayesian techniques, see [1] and [9]), or by specifying quadratic approximations for current plus future costs $W_n + T_{n+1}^{\text{opt}} \circ f_n$, with estimates then to be substituted in for the expectations and covariance matrices of the state coefficients (wide-sense adaptive dual control, see

[10]). As is well-documented, certainty equivalence techniques generally result in significant control performance loss, and Bayesian techniques are generally not computationally feasible. Although problems of modest size can be handled by wide-sense adaptive dual control techniques, the computational requirements are still quite extensive (see [10] and [2, ch. 9]).

In view of these difficulties, the exploration of alternative approaches to the adaptive control of system (1) seems worthwhile. Returning to the basic optimality equations (4), the crux of the problem in period n is to compute and recursively update the final cost-to-go function $T_N^{\text{opt}}(x)$ by means of the expectation operator $E[\cdot]$ in order to obtain the expected cost evaluations

$$C_n(\theta, x_n) \equiv E[W_n(\omega_n, \theta, x_n) + T_{n+1}^{\text{opt}} \circ f_n(\omega_n, \theta, x_n)] \quad (5)$$

relevant for period n . For adaptive control problems, the expectation operator $E[\cdot]$ is unknown and must be estimated. Intuitively, however, the estimation should not have to proceed by way of state coefficient probability distribution estimation, or even by the estimation of moments for this distribution. It is the expectation of cost upon which the controller's interest ultimately centers, not state coefficient probabilities or moments. Thus, significant reductions in computation and information requirements could possibly be realized if attention were to be shifted from the indirect recursive estimation of the expected cost evaluations (5), via state coefficient estimation or state coefficient distribution estimation, to the direct recursive estimation of (5) as the expectation of the random cost function

$$[W_n + T_{n+1}^{\text{opt}} \circ f_n](\cdot, \theta, x_n)$$

over the space R of possible state coefficient values ω_n .

To demonstrate the feasibility of this idea, consider the following "criterion-filtering" adaptive control algorithm which incorporates a procedure for directly generating estimates $C_n^{1*}(\theta, x_n)$ for the expected cost evaluations (5).

Period 1: Generate¹ a data set $\{\omega_{-M}, \dots, \omega_0\}$ from a prior probability distribution estimate for (R, \mathcal{F}, p) , and compute the prior optimality equation estimates

$$T_N^{1*}(x) \equiv \min_{\theta \in R} C_N^{1*}(\theta, x) \\ \equiv \min_{\theta \in R} \left(\frac{\sum_{i=0}^M W_N(\omega_{-i}, \theta, x)}{M+1} \right); \quad (6) \\ T_n^{1*}(x) \equiv \min_{\theta \in R} C_n^{1*}(\theta, x) \\ \equiv \min_{\theta \in R} \left(\frac{\sum_{i=0}^M [W_n(\omega_{-i}, \theta, x) + T_{n+1}^{1*} \circ f_n(\omega_{-i}, \theta, x)]}{M+1} \right), \quad (7)$$

¹ If actual observations are available for the parameters ω_n , these may be used in place of or in conjunction with simulated observations. More generally, the prior estimates $C_n^{1*}(\theta, x)$ can be specified directly, without reliance on actual or simulated data.

$x \in R, n \in \{1, \dots, N-1\}$. Select a control $\theta_1^* \equiv \theta_1^*(x_1^*) \in R$ for period 1 which satisfies $T_1^{1*}(x_1^*) = C_1^{1*}(\theta_1^*, x_1^*)$. Record the new initial state $x_2 = f_1(\omega_1, \theta_1^*, x_1^*)$ for period 2, and the state coefficient $\omega_1 = [x_2 - b_1\theta_1^*]/x_1^*$ for period 1.²

Period $n(2 \leq n \leq N)$: Compute the optimality equation estimates

$$T_N^{n*}(x) \equiv \min_{\theta \in R} C_N^{n*}(\theta, x) \\ \equiv \min_{\theta \in R} \left(\frac{S[C_N^{1*}(\theta, x)] + \sum_{j=1}^{n-1} W_N(\omega_j, \theta, x)}{S + n - 1} \right); \quad (8)$$

$$T_k^{n*}(x) \equiv \min_{\theta \in R} C_k^{n*}(\theta, x) \\ \equiv \min_{\theta \in R} \left(\frac{S[C_k^{1*}(\theta, x)] + \sum_{j=1}^{n-1} [W_k(\omega_j, \theta, x) + T_{k+1}^{n*} \circ f_k(\omega_j, \theta, x)]}{S + n - 1} \right); \quad (9)$$

$x \in R, k \in \{n, \dots, N-1\}$. Select a control $\theta_n^* \equiv \theta_n^*(x_n) \in R$ which satisfies

$$T_n^{n*}(x_n) = C_n^{n*}(\theta_n^*, x_n).$$

Record the new initial state $x_{n+1} = f_n(\omega_n, \theta_n^*, x_n)$ for period $n+1$, and the state coefficient $\omega_n = [x_{n+1} - b_n\theta_n^*]/x_n$ for period n .

The coefficient S entering into the optimality equation estimates (8) and (9) is a data compression constant reflecting the controller's confidence in the prior expected cost estimates $C_n^{1*}(\theta, x)$. Ignoring this coefficient for the moment by setting $S=1$, and letting $T_{N+1}^{n*}(x) \equiv T_{N+1}^{n*}(x) \equiv 0$, three steps are to be carried out in each period $n > 1$ in order to obtain updated estimates $C_n^{n*}(\theta, x_n)$ for the unknown expected cost evaluations $C_n(\theta, x_n)$ defined by (5). First, the expectation operation $E[\cdot]$ is approximated by an averaging operation over the past observations $\omega_1, \dots, \omega_{n-1}$; second, the cost-to-go term $T_{n+1}^{n*} \circ f_n$ is approximated by the recursively defined estimate $T_{n+1}^{n*} \circ f_n$; and finally, for each $\theta \in R$, the data-based estimate

$$\left(\frac{E_{j=1}^{n-1} [W_n(\omega_j, \theta, x_n) + T_{n+1}^{n*} \circ f_n(\omega_j, \theta, x_n)]}{n-1} \right) \quad (10)$$

for $C_n(\theta, x_n)$ is averaged with the prior estimate $C_n^{1*}(\theta, x_n)$ for $C_n^{n*}(\theta, x_n)$ to obtain $C_n^{n*}(\theta, x_n)$.

Stated somewhat differently, the estimate $C_n^{n*}(\theta, x_n)$ for the true but unknown expected cost $C_n(\theta, x_n)$ associated with the selection of control θ in period n is obtained by means of a linear filtering operation on a vector of n pseudo cost observations

$$(C_n^{1*}(\theta, x_n), [W_n + T_{n+1}^{n*} \circ f_n](\omega_1, \theta, x_n), \\ \dots, [W_n + T_{n+1}^{n*} \circ f_n](\omega_{n-1}, \theta, x_n)), \quad (11)$$

² If $x_n = 0$ for any $n \in \{1, \dots, N\}$, then the optimal control selection for period n and all subsequent periods is clearly 0. In the statement of the algorithm it is implicitly assumed that $x_n \neq 0$ for all $n \in \{1, \dots, N\}$.

each interpreted as a realization for the random cost variable

$$[W_n + T_{n+1}^{n*} \circ f_n](\cdot, \theta_n, x_n) \quad (12)$$

with expectation $C_n(\theta, x_n)$. Since by assumption the state coefficients $(\omega_1, \dots, \omega_N)$ are independent and identically distributed with finite expectation and variance, it follows by backward recursion that $C_n^{n*}(\theta, x_n)$ is a consistent estimate for $C_n(\theta, x_n)$.

Returning again to the data compression constant S which enters into the optimality equation estimates (8) and (9), notice as S varies from 0 to $M+1$ that the number of pseudo prior cost observations

$$W_n(\omega_{-i}, \theta, x) + T_{n+1}^{n*} \circ f_n(\omega_{-i}, \theta, x), \quad i \in \{0, \dots, M\}, \quad (13)$$

averaged into the expected cost estimate $C_n^{n*}(\theta, x)$ varies in effect from 0 to $M+1$. Specifically, when $S=0$, no pseudo prior cost observations (13) enter into $C_n^{n*}(\theta, x)$; when $S=1$, the pseudo prior cost observations (13) are compressed into a single piece of data $C_n^{1*}(\theta, x)$ before use; and, when $S=M+1$, each of the pseudo prior cost observations (13) enters into the estimate $C_n^{n*}(\theta, x)$ on an equal footing with all subsequent pseudo cost observations

$$W_n(\omega_j, \theta, x) + T_{n+1}^{n*} \circ f_n(\omega_j, \theta, x), \quad j \geq 1. \quad (14)$$

Higher values $S > M+1$ would indicate that the pseudo prior cost observations (13) are believed to be of better quality than the subsequently constructed pseudo cost observations (14).

Selection of the data compression constant S is one small part of the overall problem of optimal criterion-filter design. A variational principle for the selection of criterion-filter weight schemes will be discussed in Section IV.

One important question that must be asked concerns the computational requirements of the criterion-filtering algorithm. Criterion-filtering can be classified as an open-loop feedback adaptive control method, i.e., a feedback adaptive control method which directs the controller in each period n to ignore the fact that future observations will be made (see [9, pp. 241-245] and [2]). Open-loop feedback adaptive control methods have a great computational advantage over closed-loop adaptive control methods, such as Bayesian-dynamic programming, which attempt to take into account both past and potential future observations in each decision period. The principal computational advantage of the criterion-filtering method relative to many existing open-loop feedback adaptive control methods is the absence of required integrations. Although the computational burden is by no means eliminated, it is approximately reduced to the level of N deterministic dynamic programming problems.

Other important questions concern the optimality properties of the criterion-filtering approach to adaptive control.

- Does the filter-generated control law have the same general structure as the optimal control law?
- Do the filter-generated cost-to-go function estimates

have the same general structure as the true underlying cost-to-go functions?

- How do the state, control, single-period cost, and cost-to-go trajectories realized by use of the filter-generated control law compare to the state, control, single-period cost, and cost-to-go trajectories realized by use of the optimal control law?

In the following section a detailed analytical and computer simulation study of these questions will be carried out for the linear-quadratic control system (1), one of the few control problems for which the optimal feedback control law can be obtained in closed form.

III. SIMULATION RESULTS

The n th period optimal control $\theta_n^{\text{opt}}(x)$ and cost-to-go function $T_n^{\text{opt}}(x)$ for system (1) take the form

$$\theta_n^{\text{opt}}(x) = - \left(\frac{[E\omega]cb_n[1 + V(n+1)]}{cb_n^2[1 + V(n+1)] + q} \right) x, \quad (15)$$

$$T_n^{\text{opt}}(x) = cV(n)x^2, \quad (16)$$

where

$$V(N+1) = 0, \quad (17)$$

$$V(n) = [1 + V(n+1)]$$

$$\cdot \left[\text{var}(\omega) + (E\omega)^2 \left(\frac{q}{cb_n^2[1 + V(n+1)] + q} \right) \right], \quad (18)$$

$1 \leq n \leq N-1$ (see [11]). Here $E\omega$ and $\text{var}(\omega)$ denote the expectation and variance of the state coefficients ω_n with respect to (R, \mathcal{F}, p) .

Turning now to the adaptive control of system (1), suppose for simplicity that $M+1$ prior observations $\{\omega_{-M}, \dots, \omega_{-1}, \omega_0\}$ are available for the state coefficients ω_n , and the controller employs the criterion-filtering approach outlined in Section II with data compression constant S set equal to $M+1$. The controller's n th period optimality equation estimates then take the simple form

$$T_N^*(x) = \min_{\theta} \left(\frac{\sum_{j=-M}^{n-1} W_N(\omega_j, \theta, x)}{M+n} \right), \quad (19)$$

$$T_k^*(x) = \min_{\theta} \left(\frac{\sum_{j=-M}^{n-1} [W_k(\omega_j, \theta, x) + T_{k+1}^* \circ f_k(\omega_j, \theta, x)]}{M+n} \right), \quad n \leq k \leq N-1. \quad (20)$$

By straightforward arguments, entirely analogous to those used to derive (15) and (16), the n th period filter-generated control $\theta_n^*(x)$ and cost-to-go estimate $T_n^*(x)$ are explicitly given by

$$\theta_n^*(x) = - \left(\frac{[E_n\omega]cb_n[1 + K(n+1)]}{cb_n^2[1 + K(n+1)] + q} \right) x, \quad (21)$$

$$T_n^*(x) = cK(n)x^2, \quad (22)$$

where, for each $n \in \{1, \dots, N\}$,

$$K(N+1) = 0,$$

$$K(n) = [1 + K(n+1)]$$

$$\cdot \left[\text{var}_n(\omega) + [E_n\omega]^2 \left(\frac{q}{cb_n^2[1 + K(n+1)] + q} \right) \right],$$

$$E_n\omega = \left(\frac{\sum_{k=-M}^{n-1} \omega_k}{M+n} \right),$$

$$\text{var}_n(\omega) = \left(\frac{\sum_{j=-M}^{n-1} (\omega_j - E_n\omega)^2}{M+n} \right).$$

(23)

The first two questions posed in Section II concerning the optimality properties of the criterion-filtering approach can thus be answered in the affirmative for system (1); namely, the filter-generated control law $(\theta_1^*(x), \dots, \theta_N^*(x))$ and cost-to-go function estimates $(T_1^*(x), \dots, T_N^*(x))$ have the same general structure as the optimal control law $(\theta_1^{\text{opt}}(x), \dots, \theta_N^{\text{opt}}(x))$ and cost-to-go functions $(T_1^{\text{opt}}(x), \dots, T_N^{\text{opt}}(x))$.

The third and final optimality question posed in Section II is as follows: How do the control, state, single-period cost, and cost-to-go trajectories realized by use of the filter-generated control law compare to the corresponding trajectories realized by use of the optimal control law? To provide a partial answer to this question, comparative simulation studies were carried out on an IBM 370/158 for the optimal and criterion-filter control laws using the following range of parameter values:³

| | |
|--------------|---|
| x_1^* | initial state: 1 |
| b_n | control coefficient: 1 in each period n |
| c | cost function coefficient: 1 |
| q | cost function coefficient: 0 or 1 |
| N | time horizon: 10, 50, or 100 |
| M | number of prior observations: 10 or 100 |
| $E\omega$ | expected value of state coefficient: 0, 1, or 2 |
| $SD(\omega)$ | standard deviation of state coefficient: 0.5, 1.0, or 1.5 |

(24)

Complete trajectories were generated for states, controls, single-period costs, cost-to-go evaluations, mean estimates $E_n\omega$, and standard deviation estimates $SD_n(\omega) \equiv (\text{var}_n(\omega))^{1/2}$. Fifty trial runs were carried out for each parameter configuration $(N, M, q, E\omega, SD(\omega))$ using pseudo normally distributed random variables $N(E\omega, SD(\omega))$. Prior $N(E\omega, SD(\omega))$ variables $\omega_{-M}, \dots, \omega_0$ were generated separately for the filter case; otherwise, identical noise sequences were used in each trial run for both the filter and optimal cases. The averages of the resulting trajectories over the fifty trial runs were outputted for each parameter

³ Programs are available upon request.

TABLE I
OPTIMAL CONTROL: STATE, CONTROL, SINGLE-PERIOD COST, AND COST-TO-GO TRAJECTORIES*

Parameter Values: $q = 1, N = 10, E\omega = 2, SD(\omega) = 1.0$

| Period n | x_n | θ_n^{opt} | $W(\omega_n, \theta_n^{opt}, x_n)$ | $T_n^{opt}(x_n)$ |
|-----------------------------|-----------------|------------------|------------------------------------|-------------------|
| | σ | σ | σ | σ |
| 1 | 1.000 0.0 | -1.953 0.000 | 4.866 1.597 | 45.623 0.000 |
| 2 | -0.079 1.022 | 0.154 1.990 | 4.869 7.602 | 42.794 65.025 |
| 3 | -0.029 0.940 | 0.057 1.823 | 4.711 14.329 | 31.674 72.279 |
| 4 | -0.234 1.154 | 0.452 2.225 | 5.953 31.658 | 42.883 237.050 |
| 5 | 0.034 0.893 | -0.065 1.708 | 3.725 16.436 | 20.813 88.091 |
| 6 | -0.103 0.890 | 0.194 1.683 | 4.165 22.533 | 17.058 91.145 |
| 7 | -0.034 1.137 | 0.064 2.108 | 5.993 33.820 | 21.300 118.512 |
| 8 | -0.092 1.240 | 0.165 2.210 | 5.449 31.988 | 18.185 106.846 |
| 9 | 0.103 0.725 | -0.173 1.160 | 2.135 12.318 | 3.867 22.555 |
| 10 | 0.075 0.869 | -0.075 0.869 | 1.993 9.579 | 2.282 12.979 |
| 11 | -0.056 1.109 | | | |
| TOTAL REALIZED COST: 43.859 | | | | |

* All table entries represent averages over 50 trial runs.

configuration, together with their sample standard deviations.

Examples of two such simulations are presented in Tables I and II.

Comparing Tables I and II, the total costs realized by use of the filter and optimal control laws are very close. Since presumably it is total realized costs which ultimately interest the controller, Table III displays this information for simulations run with $N = 10, M = 10$, and various values for $q, E\omega$, and $SD(\omega)$.

As evidenced by Table III, the filter control law outperforms⁴ the optimal control law when the standard deviation $SD(\omega)$ is 1.5. For the more moderate case, $SD(\omega) = 1.0$, the optimal control law only slightly outperforms the filter control law by a factor of approximately 9/10. These results are particularly encouraging, for, as demonstrated in [11], system (1) is unstable whenever $SD(\omega) \geq 1.0$.

Intuitively, one would expect an increase in the number M of prior observations to improve the performance of the filter control law. As evidenced by Table IV, the effect on total costs of an increase from $M = 10$ to $M = 100$ is somewhat inconclusive for the short time horizon $N = 10$. For some parameter configurations, costs slightly increase, and for others costs slightly drop.

The beneficial effect of an increase in the number M of

⁴ The possibility of the filter control law outperforming the optimal control law in any given trial run arises because the optimal control law is only guaranteed to be best in a long-run average sense.

TABLE II
FILTER CONTROL: STATE, CONTROL, SINGLE-PERIOD COST, COST-TO-GO, MEAN ESTIMATE AND STANDARD DEVIATION ESTIMATE TRAJECTORIES*

Parameter Values: $q = 1, N = 10, E\omega = 2, SD(\omega) = 1.0$

| Period n | x_n | θ_n^* | $W(\omega_n, \theta_n^*, x_n)$ | $T_n^*(x_n)$ | $E_n\omega$ |
|-----------------------------|-----------------|-----------------|--------------------------------|---------------------|----------------|
| | σ | σ | σ | σ | $SD_n(\omega)$ |
| 1 | 1.000 0.0 | -1.937 0.308 | 4.867 1.742 | 135.744 274.961 | 2.024 0.972 |
| 2 | -0.063 1.007 | 0.064 1.888 | 4.610 6.752 | 184.719 495.580 | 2.010 0.971 |
| 3 | -0.134 1.011 | 0.257 1.902 | 6.504 22.220 | 185.017 636.539 | 2.018 0.970 |
| 4 | -0.248 1.661 | 0.400 2.868 | 9.382 47.592 | 368.697 2414.476 | 2.012 0.967 |
| 5 | 0.130 0.989 | -0.272 1.884 | 4.382 12.750 | 67.192 248.950 | 2.025 0.975 |
| 6 | -0.115 0.863 | 0.220 1.607 | 4.188 11.286 | 36.577 109.922 | 2.013 0.974 |
| 7 | 0.091 1.245 | -0.225 2.205 | 6.650 23.095 | 48.292 169.838 | 2.022 0.975 |
| 8 | -0.088 1.315 | 0.134 2.200 | 5.304 24.417 | 30.799 156.564 | 2.009 0.978 |
| 9 | 0.075 0.665 | -0.093 1.009 | 1.521 5.922 | 3.888 14.345 | 2.005 0.981 |
| 10 | 0.061 0.701 | -0.046 0.654 | 1.109 4.931 | 1.580 8.293 | 2.014 0.978 |
| 11 | 0.018 0.824 | | | | |
| TOTAL REALIZED COST: 48.519 | | | | | |

* All table entries represent averages over 50 trial runs.

TABLE III
TOTAL REALIZED COSTS: $N = 10$ AND $M = 10$ *

| Control | $SD(\omega)$ | $q = 0$ | | $q = 1$ | |
|---------|--------------|---------------------------------|---------------|---------------|---------------|
| | | $E\omega = 0, 1, \text{ or } 2$ | $E\omega = 0$ | $E\omega = 1$ | $E\omega = 2$ |
| Filter | 1 | 11.125 | 11.848 | 18.856 | 48.519 |
| Optimal | | 9.116 | 9.116 | 16.966 | 43.859 |
| Filter | 1.5 | 2,142.752 | 2,330.397 | 3,143.792 | 6,898.074 |
| Optimal | | 2,366.137 | 2,366.137 | 3,849.684 | 8,201.664 |

* All table entries represent averages over 50 trial runs.

TABLE IV
TOTAL REALIZED COSTS: $N = 10$ AND $M = 100$ *

| Control | $SD(\omega)$ | $q = 0$ | | $q = 1$ | |
|---------|--------------|---------------------------------|---------------|---------------|---------------|
| | | $E\omega = 0, 1, \text{ or } 2$ | $E\omega = 0$ | $E\omega = 1$ | $E\omega = 2$ |
| Filter | 1 | 9.396 | 9.476 | 19.229 | 49.259 |
| Optimal | | 9.116 | 9.116 | 16.966 | 43.859 |
| Filter | 1.5 | 2,219.750 | 2,248.261 | 3,945.809 | 8,706.934 |
| Optimal | | 2,366.137 | 2,366.1370 | 3,849.684 | 8,201.664 |

* All table entries represent averages over 50 trial runs.

TABLE V
TOTAL REALIZED COSTS: $N = 100$ AND $M = 10^*$

| | | q = 0 | | q = 1 | | |
|---------|----------------|----------------------------------|--|----------------|----------------|----------------|
| Control | SD(ω) | E $\omega = 0, 1, \text{ or } 2$ | | E $\omega = 0$ | E $\omega = 1$ | E $\omega = 2$ |
| Filter | 1 | 5.935 | | 6.121 | 13.313 | 34.366 |
| Optimal | | 2.996 | | 2.996 | 6.911 | 18.833 |
| Filter | 1.5 | 132,131.313 | | 142,809.500 | 265,980.750 | 653,426.125 |
| Optimal | | 83,622.125 | | 83,622.125 | 167,244.875 | 418,116.688 |

*All table entries represent averages over 50 trial runs.

TABLE VI
TOTAL REALIZED COSTS: $N = 10$ AND $M = 100^*$

| | | q = 0 | | q = 1 | | |
|---------|----------------|----------------------------------|--|----------------|----------------|----------------|
| Control | SD(ω) | E $\omega = 0, 1, \text{ or } 2$ | | E $\omega = 0$ | E $\omega = 1$ | E $\omega = 2$ |
| Filter | 1 | 2.811 | | 2.827 | 6.358 | 17.506 |
| Optimal | | 2.996 | | 2.996 | 6.911 | 18.833 |
| Filter | 1.5 | 86,923.000 | | 87,230.500 | 183,357.813 | 453,353.688 |
| Optimal | | 83,622.125 | | 83,622.125 | 167,244.875 | 418,116.688 |

*All table entries represent averages over 50 trial runs.

prior observations is much more in evidence for the longer time horizon $N = 100$. Comparing Table IV to Table V, filter cost performance deteriorates relative to optimal cost performance when the time horizon is increased from $N = 10$ to $N = 100$ and the number of prior observations is held constant at $M = 10$.

However, comparing Table V to Table VI, filter performance dramatically improves when the time horizon is $N = 100$ and the number of prior observations is increased from $M = 10$ to $M = 100$. Specifically, total costs realized under filter control are uniformly less than the total costs realized under optimal control when the standard deviation $SD(\omega)$ is 1.0, and filter and optimal total costs are approximately equal when $SD(\omega) = 1.5$.

Several uniformities are evident in Tables III–VI. The most obvious uniformity is that total realized costs for both filter and optimal control are independent of the state coefficient expectation $E\omega$ when $q = 0$. Examining (15) for the optimal control law, total realized costs should be approximately independent of $E\omega$ when $q = 0$ under optimal control, since then

$$W(\omega_n, \theta_n^{\text{opt}}(x_n), x_n) = [\omega_n - E\omega]^2 x_n^2, \quad (25)$$

$1 \leq n \leq N$, where $\omega_n - E\omega \sim N(0, \text{var}(\omega))$. Similarly, using the filter control law described by (21),

$$\begin{aligned} W_n(\omega_n, \theta_n^*(x_n), x_n) &= [\omega_n - E_n \omega]^2 x_n^2 \\ &\cong [\omega_n - E\omega]^2 x_n^2, \end{aligned} \quad (26)$$

$1 \leq n \leq N$. The central limit algorithm used to generate the pseudo normally distributed random variable ω_n yields exact independence, since all variables are generated from 0-expectation variables by simply adding the desired expectation.

A second uniformity in Tables III–VI is that total costs

TABLE VII
TOTAL REALIZED COSTS: $q = 1, E\omega = 2, M = 10^*$

| | | N | | |
|---------|----------------|-----------|-----------------|-------------|
| Control | SD(ω) | 10 | 50 | 100 |
| Filter | 0.5 | 4.887 | 4.559 | 5.045 |
| Optimal | | 4.832 | 4.233 | 4.807 |
| Filter | 1.0 | 48.519 | 111.684 | 34.366 |
| Optimal | | 43.859 | 34.429 | 18.833 |
| Filter | 1.5 | 6,898.074 | 161,301,280.000 | 693,426.125 |
| Optimal | | 8,201.664 | 11,745,118.000 | 418,116.688 |

*All table entries represent averages over 50 trial runs.

TABLE VIII
TOTAL REALIZED COSTS: $q = 1, E\omega = 2, M = 100^*$

| | | N | | |
|---------|----------------|-----------|----------------|-------------|
| Control | SD(ω) | 10 | 50 | 100 |
| Filter | 0.5 | 4.803 | 4.159 | 4.756 |
| Optimal | | 4.832 | 4.233 | 4.807 |
| Filter | 1.0 | 49.259 | 34.281 | 17.506 |
| Optimal | | 43.859 | 34.429 | 18.833 |
| Filter | 1.5 | 8,706.934 | 55,742,928.000 | 453,533.688 |
| Optimal | | 8,201.664 | 11,754,118.000 | 418,116.688 |

*All table entries represent averages over 50 trial runs.

exhibit a moderate increase along each row, as q and $E\omega$ increase, and total costs exhibit a dramatic increase along each column, as the standard deviation $SD(\omega)$ increases. The relative deviation between filter and optimal cost performance is fairly stable along rows, but not along columns. These observations suggest that a more detailed investigation should be made, focusing on the three key parameters N , M , and $SD(\omega)$. In Table VII, filter and optimal total realized costs are compared for the time horizon N varying from 10 to 100 and the state coefficient standard deviation varying from 0.5 to 1.5, with $(M, q, E\omega)$ held constant at $(10, 1, 2)$.

As evidenced by Table VII, filter cost performance is uniformly on par with optimal cost performance when the standard deviation $SD(\omega)$ is 0.5. In addition, filter cost performance is uniformly on par with or better than optimal cost performance when the time horizon N is 10. Relative filter cost performance deteriorates with increasing values for $SD(\omega)$ and N . However, as seen in Table VIII, an increase in the number M of prior observations significantly improves filter performance. Specifically, filter costs are now uniformly on par with optimal costs for $SD(\omega) = 0.5$ and for $SD(\omega) = 1.0$, and also for $SD(\omega) = 1.5$ when $N = 10$ or $N = 100$.

In summary, comparative simulation studies indicate that the total costs realized by use of the filter-generated control law are approximately the same as the total costs realized by use of the optimal control law when the state coefficient standard deviation $SD(\omega)$ is 0.5. A similar statement holds for the higher standard deviation levels 1.0 and 1.5 when $M = 100$. Increases in the cost function coefficient q and

state coefficient expectation $E\omega$ increase total realized costs for both the filter and the optimal control law, but do not appear to affect their relative cost performance. Overall, the filter control law performed very well in comparison to the optimal control law.

IV. EXTENSIONS

The criterion-filtering approach proposed in Section II for the adaptive control of system (1) can be modified to handle higher dimensional nonlinear systems with random parameters dependent on time, control, and previous state values (see [12]). In this section we wish to indicate the nature of the needed modifications for the time-dependent case and to suggest a general principle for the adaptive updating of the criterion-filter weights.

Consider a system described by equations of the form

$$\begin{aligned} x_1 &= x_1^*, & \text{initial conditions,} \\ x_{n+1} &= f_n(\omega_n, \theta_n, x_n), & 1 \leq n \leq N, \end{aligned} \quad (27)$$

where $x_n \in R^s$ is the system state, and the control θ_n is constrained to lie in an admissible control set $\Theta(n, x_n) \subseteq R^q$. Letting \mathcal{F} denote the σ -algebra consisting of all Borel subsets of R^m , it will be assumed that $\omega_n \in R^m$ is an observable random parameter governed by a nonstationary probability distribution (R^m, \mathcal{F}, p_n) . In addition, it will be assumed that the state function $f_n: R^m \times R^q \times R^s \rightarrow R^s$ is continuous and that the cost associated with any parameter, control, and state configuration $(\omega_n, \theta_n, x_n)$ for period n is measured by a continuous cost function $W_n: R^m \times R^q \times R^s \rightarrow R$.

The set \mathcal{L} of admissible feedback control laws for system (27) is assumed to be the set of all vectors $\theta = (\theta_1, \dots, \theta_N)$ of measurable functions $\theta_n: R^s \rightarrow R^q$ satisfying $\theta_n(x) \in \Theta(n, x)$ for each $x \in R^s$. The objective assumed for the controller of system (27) is to minimize the expected value of total cost

$$\sum_{n=1}^N W_n(\omega_n, \theta_n(x_n), x_n) \quad (28)$$

via selection of an admissible feedback control law $\theta \in \mathcal{L}$. Systems having form (27) with control objective (28) will, for simplicity, be said to have the *basic format*.

The basic format model characterizes a variety of socioeconomic control problems which have state parameters ω_n exhibiting a possibly complicated stochastic dependence on time, yet which have the compensating feature that both the parameters ω_n and the states x_n are observable. (An axiomatization for a version of the basic format model with state parameters dependent on control selections is provided in [13].)

Example—Classic Investment Problem [14], [15]: In each period $n \in \{1, \dots, N\}$ an investor must decide how to allocate his current wealth $x_n \in R_+$ between two investment opportunities A and B , the first which yields a positive or negative net return rate $\pm s_n$ ($0 < s_n \leq 1$) with probabilities p_n and $1 - p_n$, and the second which yields a net return rate r_n ($0 \leq r_n < s_n$) with probability 1. The objective of the investor is to maximize the expected utility of his total wealth x_{N+1} at

the end of stage N via feedback control. Initial wealth x_1^* in period 1 is assumed to be positive.

The investor's wealth x_{n+1} in period $n+1$ is a simple function of his wealth x_n in period n , the net return rate $\omega_n \in \{s_n, -s_n\}$ observed for investment opportunity A in period n , and the amount $\theta_n \in \Theta(n, x_n) \equiv [0, x_n]$ of his wealth he allocated to A in period n ; namely,

$$x_{n+1} = x_n + \omega_n \theta_n + r_n [x_n - \theta_n] \equiv f_n(\omega_n, \theta_n, x_n).$$

Letting $U: R_+ \rightarrow R_+$ denote the investor's utility of wealth function, and letting $W_n \equiv 0$, $1 \leq n \leq N-1$, and

$$\begin{aligned} W_N(\omega_N, \theta_N, x_N) \\ \equiv -U(x_N + \omega_N \theta_N + r_N [x_N - \theta_N]) = -U(x_{N+1}), \end{aligned}$$

this investment problem has the basic format.

Letting $E_n[\cdot]$ denote expectation with respect to (R^m, \mathcal{F}, p_n) , $1 \leq n \leq N$, an admissible feedback control law $\theta \in \mathcal{L}$ for the basic format model minimizes the expected value of total cost (28) if and only if the following dynamic programming optimality equations hold for almost every sequence $(\omega_1, \dots, \omega_N)$ [8; thm. 14.4, p. 101; thm. 15.2, p. 104; and thm. 17.6, p. 111]:

$$\begin{aligned} T_N^{\text{opt}}(x_N) &\equiv \min_{\theta \in \Theta(N, x_N)} E_N[W_N(\omega_N, \theta, x_N)] \\ &= E_N[W_N(\omega_N, \theta_N(x_N), x_N)], \\ T_n^{\text{opt}}(x_n) &\equiv \min_{\theta \in \Theta(n, x_n)} E_n[W_n(\omega_n, \theta, x_n) + T_{n+1}^{\text{opt}} \circ f_n(\omega_n, \theta, x_n)] \\ &= E_n[W_n(\omega_n, \theta_n(x_n), x_n) + T_{n+1}^{\text{opt}} \circ f_n(\omega_n, \theta_n(x_n), x_n)], \\ & \quad 1 \leq n \leq N-1. \end{aligned} \quad (29)$$

If the distributions (R^m, \mathcal{F}, p_n) are not *a priori* known, the controller is faced with an extremely difficult adaptive control problem. The following criterion-filtering algorithm suggests one possible means for obtaining estimates $T_n^{**}(x)$ for the optimal cost-to-go expressions $T_n^{\text{opt}}(x)$ in (29).

Period 1: For each $n \in \{1, \dots, N\}$, generate¹ a data set $\{\omega_{-M}(n), \dots, \omega_0(n)\}$ from a prior probability estimate for (R^m, \mathcal{F}, p_n) . Compute the prior optimality equation estimates

$$\begin{aligned} T_N^{1*}(x) &\equiv \min_{\theta \in \Theta(N, x)} C_N^{1*}(\theta, x) \\ &\equiv \min_{\theta \in \Theta(N, x)} \left(\frac{\sum_{i=0}^M W_N(\omega_{-i}(N), \theta, x)}{M+1} \right), \\ T_n^{1*}(x) &\equiv \min_{\theta \in \Theta(n, x)} C_n^{1*}(\theta, x) \\ &\equiv \min_{\theta \in \Theta(n, x)} \left(\frac{\sum_{i=0}^M [W_n(\omega_{-i}(n), \theta, x) + T_{n+1}^{1*} \circ f_n(\omega_{-i}(n), \theta, x)]}{M+1} \right), \end{aligned} \quad (30)$$

$x \in R^s$, $n \in \{1, \dots, N-1\}$. Select a control

$$\theta_1^* \equiv \theta_1^*(x_1^*) \in \Theta(1, x_1^*)$$

for period 1 which satisfies $T_1^{1*}(x_1^*) = C_1^{1*}(\theta_1^*, x_1^*)$. Record the observation ω_1 for period 1 and the new state $x_2 = f_1(\omega_1, \theta_1^*, x_1^*)$ for period 2.

Period n ($2 \leq n \leq N$): Compute the optimality equation estimates

$$\begin{aligned} T_N^{n*}(x) &\equiv \min_{\theta \in \Theta(N,x)} C_N^{n*}(\theta, x) \\ &\equiv \min_{\theta \in \Theta(N,x)} \left(H_{N,0} C_N^{1*}(\theta, x) \right. \\ &\quad \left. + \sum_{j=1}^{n-1} H_{Nj} W_N(\omega_j, \theta, x) \right) K_N^n, \end{aligned}$$

$$\begin{aligned} T_k^{n*}(x) &\equiv \min_{\theta \in \Theta(k,x)} C_k^{n*}(\theta, x) \\ &\equiv \min_{\theta \in \Theta(k,x)} \left(H_{k,0} C_k^{1*}(\theta, x) + \sum_{j=1}^{n-1} H_{kj} [W_k(\omega_j, \theta, x) \right. \\ &\quad \left. + T_{k+1}^{n*} \circ f_k(\omega_j, \theta, x)] \right) K_k^n, \quad (31) \end{aligned}$$

$x \in R^s, k \in \{n, \dots, N-1\}$. Select a control

$$\theta_n^* \equiv \theta_n^*(x_n) \in \Theta(n, x_n)$$

for period n which satisfies $T_n^{n*}(x_n) = C_n^{n*}(\theta_n^*, x_n)$. Record the observation ω_n for period n and the new initial state $x_{n+1} = f_n(\omega_n, \theta_n^*, x_n)$ for period $n+1$.

Comparing the cost-to-go estimate $T_k^{n*}(x)$ in (31) to the corresponding optimal cost-to-go function

$$\begin{aligned} T_k^{\text{opt}}(x) &\equiv \min_{\theta \in \Theta(k,x)} C_k(\theta, x) \\ &\equiv \min_{\theta \in \Theta(k,x)} E_k[W_k(\omega_k, \theta, x) + T_{k+1}^{\text{opt}} \circ f_k(\omega_k, \theta, x)], \end{aligned} \quad (32)$$

it is seen that $T_k^{n*}(x)$ is a minimized weighted average of the prior estimate $C_k^{1*}(\theta, x)$ for $C_k(\theta, x)$ and the n th period pseudo cost observations

$$W_k(\omega_j, \theta, x) + T_{k+1}^{n*} \circ f_k(\omega_j, \theta, x) \quad (33)$$

for $C_k(\theta, x)$. The term K_k^n in (31) is a normalization factor ensuring the filter weights sum to one. The weight H_{kj} in (31) measures the appropriateness of treating the pseudo cost observation (33) as if it were an actual observation on $C_k(\theta, x)$. Intuitively, the weight H_{kj} in (31) should be close to 1.0 if the distribution (R, \mathcal{F}, p_j) for the past observation ω_j is similar to the distribution (R, \mathcal{F}, p_k) for the currently relevant but unknown parameter ω_k , and close to 0.0 if the two distributions are very dissimilar.

Two key considerations in selecting a filter weight scheme H for (31) are consistency and rapid rate of convergence for the generated cost-to-go estimates $T_k^{n*}(x)$. The performance index used in the selection of the filter weight scheme should presumably reflect both objectives. In addition, since the controller has only partial prior structural information, the performance index should be adaptively updated as new observations become available. One filter weight scheme selection procedure which meets these specifications is as follows.

Suppose in period 1 the controller is able to span the possible model structures by means of a collection

$$\mathcal{H} = \{H^t \mid t \in T\} \quad (34)$$

of distinct filter weight schemes

$$H^t = \{H_{kj}^t \mid (k, j+1) \in Z^2\}, \quad Z \equiv \{1, 2, \dots\}. \quad (35)$$

For example, if the controller for system (27) believes the parameters (ω_n) are drawings from slowly time-evolving distributions (R^m, \mathcal{F}, p_n) , then \mathcal{H} might consist of a collection of fixed or fading-memory weight schemes; e.g., for $j \geq 1$,

$$H_{kj}^t = \frac{t}{t + |k-j|}, \quad (36)$$

and for $j = 0$,

$$H_{k0}^t = \begin{cases} \text{data compression coefficient corresponding to} \\ \text{the prior expected cost function } C_k^{1*}(\cdot, \cdot). \end{cases} \quad (37)$$

Alternatively, if the controller for system (27) believes the parameters (ω_n) are seasonally varying, with T denoting the set of possible season durations t , then in place of (36) he might specify

$$H_{kj}^t = \begin{cases} 1, & \text{if } k \equiv j \pmod{t}, \\ 0, & \text{otherwise.} \end{cases} \quad (38)$$

The weight scheme (38) is a simple 0-1 seasonal indicator. Less conservatively, the controller might specify an exponential weight scheme

$$H_{kj}^t = \exp(-d(k, j)), \quad (39)$$

where $d(\cdot, \cdot)$ is a suitable distance function on Z^2 satisfying $d(k, j) = 0$ if $k \equiv j \pmod{t}$. The scheme (39) is a smoothed version of (38) with broader support.

Intuitively, filter weight schemes (35) with relatively large support sets should result in rapid convergence of the criterion function estimates $C_k^{n*}(\theta, x)$, since numerous cost evaluations are averaged into each estimate. On the other hand, filter weight schemes (35) which assign positive weight in accordance with exact or near-exact exchangeability of time periods as conditioning events for parameter realizations, and hence which have relatively small support sets, should result in more consistent estimates. Thus, assuming t is the true season duration, one would expect the use of the large-support weight scheme (39) in place of the small-support weight scheme (38) to reduce the variability and increase the inconsistency of the resulting criterion function estimates. Support for this conjecture is provided by the well-known bias/variance trade-off associated with the use of spectral windows in the analogous problem of spectral density estimation.

Assuming a collection \mathcal{H} of filter weight schemes H^t has been specified in period 1, and a control-parameter sequence $(\theta_1, \omega_1, \dots, \theta_{n-1}, \omega_{n-1})$ has been recorded for some $n \geq 2$, it is proposed that the controller in period n select a weight

scheme H^i , $\hat{t} \in T$, which satisfies the following mean-squared error minimization problem:

$$\min_{t \in T} \sum_{k=1}^{n-1} \left[\left(A_k: \text{the actual cost} \right) \right. \\ \left. - \left(B_k^m: \text{the expected cost estimate} \right) \right]^2, \quad (40)$$

for period k generated by H^t

where

$$A_k \equiv A_k(\theta_k, x_k) \equiv W_k(\omega_k, \theta_k, x_k),$$

and

$$B_k^m \equiv B_k^m(\theta_k, x_k)$$

$$\equiv \left(H_{k0}^t \left[\frac{\sum_{i=0}^M W(\omega_{-i}(k), \theta_k, x_k)}{M+1} \right] \right. \\ \left. + \sum_{j=1}^{n-1} H_{kj}^t W_k(\omega_j, \theta_k, x_k) \right) K_k^m. \quad (41)$$

More generally, the controller in each period n could test the performance of the weight schemes H^t against all possible control-state pairs (θ, x) by considering the mean-squared error minimization problem

$$\min_{t \in T} \int_{R^s} \int_{R^q} \sum_{k=1}^{n-1} [A_k(\theta, x) - B_k^m(\theta, x)]^2 d\theta dx. \quad (42)$$

The filter optimality criterion (40) is not comprehensive, but it is computationally simple for finite index sets T . The filter optimality criterion (42) is more comprehensive, but it is also much more complex.

V. CONCLUSION

Adaptive control problems arise in all areas of the physical, social, and biological sciences. Nevertheless, the practical implementation of state-space adaptive control techniques has been hampered by computational and data storage difficulties arising from the apparent need to update and store estimates for states or state distributions.

In this paper an alternative approach to adaptive control is proposed for a class of adaptive control problems with intertemporal expected cost objectives. The intertemporal criterion function relevant for each period n is directly filtered without recourse to state estimation or state distribution estimation via Bayes' rule. The control law generated

by the filtered criterion function estimates is tested against the optimal control law for a linear-quadratic control model with random state coefficients. Computer simulation results indicate that the total costs realized under the filter control law are approximately on par with the total costs realized under the optimal control law for the tested range of time horizons, cost function coefficients, and mean and standard deviation values for the random state coefficients.

Further computer programs are currently being designed to test criterion filtering methods in the context of an adaptive multiteam decision problem in which the observations ω_n of one team are state and time-dependent controls implemented by a competing team acting in accordance with an unknown feedback control law (see [16]).

ACKNOWLEDGMENT

The author is grateful to R. E. Kalaba for helpful suggestions.

REFERENCES

- [1] M. Aoki, "Adaptive control theory: Survey and potential applications to decision processes," Invited Paper presented at the Stochastic Control Workshop, AIDS National Meeting, Chicago, IL, Oct. 1977.
- [2] G. N. Saridis, *Self-Organizing Control of Stochastic Systems*. New York: Marcel Dekker, 1977.
- [3] L. Tesfatsion, "Bayes' theorem for utility," discussion paper no. 76-65, Center for Economic Research, Univ. Minnesota, Minneapolis, Apr. 1976.
- [4] —, "A new approach to filtering and adaptive control," *J. Optimization Theory Appl.*, vol. 25, pp. 247-261, 1978.
- [5] —, "A new approach to filtering and adaptive control: Stability results," *Appl. Math. Comput.*, vol. 4, pp. 27-44, 1978.
- [6] —, "A new approach to filtering and adaptive control: Optimality results," *J. Cybernetics*, vol. 7, pp. 133-146, 1977.
- [7] R. E. Kalaba and L. Tesfatsion, "Two solution techniques for adaptive reinvestment: A small sample comparison," *J. Cybernetics*, vol. 8, pp. 101-111, 1978.
- [8] K. Hinderer, *Foundations of Non-Stationary Dynamic Programming with Discrete Time Parameter*. New York: Springer, 1970.
- [9] M. Aoki, *Optimization of Stochastic Systems*. New York: Academic, 1967.
- [10] Y. Bar-Shalom and E. Tse, "Caution, probing, and the value of information in the control of uncertain systems," *Annals of Economic and Social Measurements*, 5/3, pp. 323-336, 1976.
- [11] L. Tesfatsion, "Parameter sensitivity study for a linear quadratic control problem with random state coefficients," *Inf. Sciences*, vol. 15, pp. 101-125, 1978.
- [12] —, "Fuzzy filters for the direct estimation of adaptive control criterion functions," discussion paper, USC Dep. Economics, Aug. 1977.
- [13] —, "A conditional expected utility model for myopic decision makers," *Theory and Decision*, to be published.
- [14] K. J. Arrow, *Aspects of the Theory of Risk Bearing*. Yrjo Johnsson Foundation, Helsinki, 1965.
- [15] R. Bellman and R. E. Kalaba, "On the role of dynamic programming in statistical communication theory," *IRE Trans. Inform. Theory*, vol. IT-3, pp. 197-203, 1957.
- [16] L. Tesfatsion, "C³ modeling with symmetrical rationality," *Applied*