Dual Effect, Certainty Equivalence, and Separation in Stochastic Control

YAAKOV BAR-SHALOM, MEMBER, IEEE, AND EDISON TSE, MEMBER, IEEE

Abstract—In this paper the various policies in fixed end-time stochastic control are discussed first. The emphasis is on the differences between the feedback and closed-loop policies. It is shown how the closed-loop policy has the important property that it can be actively adaptive, while the feedback policy can only be passively adaptive. The feature of being actively adaptive is possible when the control has a dual effect, i.e., in addition to its effect on the state it affects the state uncertainty. The intimate connection between the neutrality (lack of dual effect) and certainty equivalence properties for a class of problems is proved. This new result is then used to widen the class of problems for which it was previously known that the certainty equivalence property holds.

I. INTRODUCTION

In the first part of this paper a discussion of the possible policies in the fixed end-time control of stochastic systems is presented. The various classes of policies differ in their information patterns—the availability of past observations and the possible usage of information about the future observations. Since the controls are non-anticipative, the only information about the future observations a controller can use is the probability distribution of the corresponding observation random variables. Loosely speaking, this is called "the future observation program and the associated statistics." It is shown that the incorporation of the future observation program and the associated statistics is what makes the difference between a feedback control policy and a closed-loop control policy. It is only the latter one that takes into account the


[24] Michael Athans (S'58-M'61-SM'69-F'73) for a photograph and biography see page 30 of the February issue of this TRANSACTIONS.
possible benefit to be derived from the future observations—it "knows" that the loop will be closed after each future measurement. The distinction between the feedback and closed-loop policies has not been sufficiently stressed in the literature and is believed to be especially important when one is looking for suboptimal algorithms. In such a case it is most desirable to obtain, if possible, a control scheme from the closed-loop class rather than the feedback class. These control policies, along with others are defined in Section II. It is pointed out that a consequence of the above mentioned difference between a feedback policy and a closed-loop policy is the following: the former is passively adaptive while the latter can be actively adaptive. The active learning feature of a closed-loop control policy can be utilized when the control has the so-called dual effect (Feldbaum [12]).

The control is said to have a dual effect when, in addition to its effect on the state of the system, it affects the uncertainty of the system's state. If the control cannot affect this uncertainty, the system is called neutral. When such a dual effect is present, the control can be used to improve the estimation, which ultimately helps to achieve the control objective. In this way a closed-loop stochastic control regulates its learning as required by the control objective. Another widely encountered property of a stochastic control, the certainty equivalence, and a weaker version of it, the separation property are also discussed in Section II.

The connection between the neutrality and the certainty equivalence properties is the topic of Section III. It is shown that for a class of problems the optimal (closed-loop) stochastic control has the certainty equivalence property if and only if the system is neutral. This new result is then used in Section III to widen the class of problems for which it was previously known that the certainty equivalence property holds.

II. Classes of Stochastic Control Policies, the Dual Effect and Certainty Equivalence

The control problem is defined as follows. The state of the system at time \( k \), \( x_k \), evolves according to the equation

\[
x_{k+1} = f(x_k, u_k, \omega_k), \quad k = 0, 1, \ldots, N - 1
\]

(2.1)

where \( u_k \) is the control applied at time \( k \) and \( \omega_k \) the process noise, and \( x_k \) are random variables. The number of stages, \( N \), is assumed to be given.1 At time \( k \), prior to applying the control, an observation might be obtained. The corresponding measurement is given by the equation

\[
y_k = h(x_k, \omega_k), \quad k = 0, 1, \ldots, N
\]

(2.2)

where \( \omega_k \), the measurement noise, is a random variable. All the above variables are, in general, vector-valued with appropriate dimensions. Denoting the underlying prob-

\footnote{The reason for limiting the present discussion to fixed end-time problems is the fact that in a free end-time stochastic control problem the definition of the end time, in a probabilistic sense (stopping time), can be done in a number of ways (see, e.g., Kushner [15]) and this is beyond the scope of the present work.}

\[ J_\delta = E[\mathcal{L}^n[X_{N+1}, U_{N+1}^{-1}]] \]  

(2.3)

where \( \mathcal{L}^n \) is a real-valued function,

\[ U_{N+1}^{-1} \triangleq \{ u_i \}_{i=0}^{N-1} \]  

(2.4)

and \( X_{N+1} \) is defined similarly.

The arguments of the minimization of the cost as given above are the controls \( U_{N+1}^{-1} \) to be applied during the \( N \)-stage control process.

As will be seen in the sequel, the various control policies to be discussed differ in the availability of past measurements and the future observation program with its associated statistics, to be defined next.

The set of observations from time \( i \) to time \( j \), when the input sequence \( U_{N-1}^{-1} \) has been applied to the system, is denoted by

\[ Y_{i:j} = Y_i \{ u_i U_{i-1}^{-1} \} \triangleq \{ y_i \}_{i=0}^{j}, \quad 1 \leq i \leq j \leq N. \]  

(2.5)

Since the probability distribution of \( x_0 \) is available, one can assume that there is no measurement at \( k = 0 \) and therefore \( Y_i \) is defined for \( i \geq 1 \). Note also that the last measurement, \( y_N \), is irrelevant, since no more controls will be applied.

The knowledge about the dynamics is denoted

\[ \mathbb{D} \triangleq \{ f(\cdot, \cdot, \cdot, \cdot) \}_{\omega_0}^{N-1}. \]  

(2.6)

The knowledge about the measurement system between times \( i \) and \( j \), called the measurement program, is

\[ \mathbb{M}_{i:j} \triangleq \{ h(\cdot, \cdot, \cdot, \cdot) \}_{\omega_0}^{j}, \quad 1 \leq i \leq j \leq N - 1. \]  

(2.7)

If the lower subscript for \( X \) or \( U \) is zero and for \( \mathbb{M} \) or \( Y \) is one, it will be dropped for simplicity in the sequel.

The joint probability measure induced by the random variables \( x_0, u_0, \ldots, u_{N-1}, y_1, \ldots, y_N \) will be represented by the symbol

\[ \mathcal{S} \triangleq dP(\omega_0, \ldots, \omega_{N-1}, y_1, \ldots, y_N) \]  

(2.8a)

while

\[ \mathcal{S} \triangleq dP(\omega_0, \ldots, \omega_{N-1}). \]  

(2.8b)

Even though this is an abuse of language, one shall call \( \mathcal{S} \) the "statistics" of the corresponding random variables.

The controller is assumed to be causal, i.e., \( u_k \) cannot be a function of \( Y_{k+1}^{-1} \) or any of its subsets. It will be also assumed that, when computing the present control \( u_k \), perfect knowledge of all past controls is available. Therefore \( u_k \) can be, at most, \( (\mathcal{P}, \mathcal{UN}^{-1}) \)-measurable; this is the "real-time" information. The "off-line" kinds of information are given by (2.6)–(2.8).

\footnote{If there is no confusion, the arguments of \( Y_{i:j} \) will be dropped for simplicity.}
Witsenhausen [30] discussed the various information patterns in stochastic control related to the sharing of the available data among the various "control stations." The aspect to be discussed in the following relates to the information about the future observations. Obviously, a causal control is constrained to be a function of only the observations that have already been obtained; however, as will be seen later, knowledge of the statistics of the future observations plays a key role in stochastic control.

With the above notations one can define the following classes of policies in fixed end-time stochastic control:

1) The open-loop policy. In this case the control has the following form:

\[ u^{o_2} = u^{o_2}[d, s^1, \ldots, s^{k}, N - 1] \]

i.e., no measurement knowledge is available for the controller.

2) The feedback policy. In this case at every time \( k \), \( Y^k \) is available for the computation of the control but no knowledge about the future measurements is available, i.e.,

\[ u^{s_k} = u^{s_k}[Y^k, U^{k-1}, d, s^1, \ldots, s^{k-1}, N - 1] \]

be the 4th central moment of \( x_{i,t} \) conditioned upon the \( \sigma \)-algebra generated by \( Y^k \) and \( U^{k-1} \). Then, the control is said to have no dual effect of order \( r (r \geq 2) \).

\[ E[M_{x,t}^{4r-1}(u_{x}, U^{r-1}), U^{k-1}] = E[M_{x,t}^{4r-1}(u_{x}^{o}), P(u_{o})] \] 

where

\[ P(u_{o}) \Delta P(u_{o}, U^{r-1} = 0) = P(u_{o}, u_{o} = \ldots = u_{k-1} = 0) \]

are the measurements of the corresponding autonomous system. In other words, the expected future uncertainty is \( \bar{F} \)-measurable, i.e., it is not affected by the control with probability one.

Conversely, if (2.14) does not hold for some \( r \geq 2 \), i.e., if the control can affect, with nonzero probability, one such central moment, then the control has a dual effect. The lack of dual effect has been called neutrality by Feldbaum [12].

There are two aspects in which the closed-loop policy differs from the feedback policy

1) Caution. In a stochastic control problem, due to the inherent uncertainties, the controller has to be "cautious" not to increase the effect of the existing uncertainties on the cost. However, the closed-loop controller, since it "knows" that future observations will be available and corrective actions based upon them will be taken, will exercise less "caution." It was this aspect of a stochastic controller that Dreyfus [10] used when he showed via an example a numerical improvement in the performance of the CL versus OLOF policy.

2) Probing or Active Learning. When the dual effect is present, the control can "help" in learning (estimation) by decreasing the uncertainty about the state. Therefore, the closed-loop control, which takes into account the future observation program and statistics, as pointed out in (2.12), has the capability of active learning [26], [27] when the dual effect exists. A feedback controller, even though

\[ M_{x,t}^{4r-1}(Y^k(u_{x}, U^{r-1}), U^{k-1}) - E(M_{x,t}^{4r-1}(Y^k(u_{x}, U^{r-1}), U^{k-1}), Y^k(u_{x}, U^{r-1})) \]

is implicitly assumed that if a control has a dual effect this helps decrease some uncertainty. This is the case when there are unknown parameters in the system and a closed-loop type control will then "probe" to "learn" them. Otherwise, added "caution" is required.

The m-measurement feedback does this partially.

Joint moments should also be included but due to the notational complexity they are omitted.

8 The m-measurement feedback does this partially.
it "learns" by using the measurements, it does not actively "help" the learning. This learning can be called, therefore, passive, or accidental, and the corresponding control policy passively adaptive, as opposed to the closed-loop control which is actively adaptive [27].

A sequential allocation procedure for a class of resource allocation problems where the decision maker can reduce the uncertainty that affects the results of the allocation has been presented in [7]. This procedure is of the one-measurement feedback class—the present decision depends upon the "value" of the information to be obtained from the next observation.

The closed-loop optimal control sequence is obtained by applying the principle of optimality (Bellman [5]) as follows:

\[
J_{c}^{\text{clo}} = \min_{u} \frac{1}{\beta} \min_{x_{n}} E\left[ \sum_{n=0}^{N-1} \mathcal{L}_{n}[x_{n},u_{n}] \right] \quad \text{where} \quad \beta = Y^{N-1},U^{N-1}] \quad \text{and} \quad \text{the, replacing } x_{n} \text{ by its estimate} \quad \hat{x}_{n1a} = E[x_{n} | Y^{n-1},U^{n-1}] \quad \text{(2.21)}
\]

Note that this control is of the feedback type rather than the closed-loop type—it makes use of the available observations but does not account for the future observations.

In a control problem it is said that the certainty equivalence (CE) property holds if the closed-loop optimal control has the same form as the deterministic optimal control with \( \hat{x}_{n1a} \) replaced by \( \hat{x}_{n1a} \), i.e.,

\[
u_{c}^{\text{clo}} = \phi_{n}(\hat{x}_{n1a}). \quad \text{(2.23)}
\]

In general, (2.22) is only an ad hoc control procedure (see, e.g., [1], [23], [30]).

The separation property (see, e.g., [1], [2], [30]) is a weaker one than the certainty equivalence. The closed-loop optimal control has the separation property if it depends on the data only via \( \hat{x}_{n1a} \)

\[
u_{c}^{\text{clo}} = \psi_{n}(\hat{x}_{n1a}) \quad \text{(2.24)}
\]

where the function \( \psi_{n} \) can be different from \( \phi_{n} \) obtained in the deterministic case (2.20). It can be easily seen that certainty equivalence implies separation but not the other way around. A problem in which the optimal control has the separation property but not the certainty equivalence was studied in [22].

The main result of this paper, to be presented in the next section, is that, for a class of problems, the certainty equivalence property holds if and only if the control has no dual effect (i.e., the system is neutral). The importance of this result lies in the fact that it gives the conditions under which the certainty equivalence control (2.22) coincides with the closed-loop optimal control. The sufficiency of neutrality for certainty equivalence has been suggested by Patchell and Jacobs [18].

III. THE CONNECTION BETWEEN THE DUAL EFFECT OF THE CONTROL AND THE CERTAINTY EQUIVALENCE PROPERTY

Consider the multidimensional system with linear dynamics and additive white, but not necessarily Gaussian noise

\[ z_{n+1} = F_{x}z_{n} + G_{x}u_{n} + v_{n} \quad \text{(3.1)} \]

where \( F_{x} \) and \( G_{x} \) are known matrices of appropriate dimensions.

\[
E v_{n} = 0 \quad \text{(3.24)}
\]

\[
E v_{n}^2 = V_{n} \delta_{n} \quad \text{(3.26)}
\]
and the general measurement model
\[ y_k = h_k(x_k, u_k) \]  
(3.3)

where \( u_k \) is the measurement noise with known but arbitrary statistics. The only restrictive assumption on the measurement noise sequence is that it is independent of the process noise. The cost to be minimized is assumed to be quadratic
\[ J_0 = E\left\{ x_0'Q_0x_0 + \sum_{i=0}^{N-1} x_i'Q_i x_i + u_i'R_i u_i \right\} \]  
(3.4)

where \( Q_i \geq 0 \) and \( R_i > 0 \). This guarantees the existence of the solution. For the control problem, defined by (3.1)–(3.4), the following result can be stated.

**Theorem:** The optimal stochastic control (i.e., closed-loop) for the system with linear dynamics (3.1) with white process noise (3.2), measurement equation (3.3), and cost (3.4) has the certainty equivalence property for all \( Q_i \geq 0, R_i > 0 \) if and only if, the control has no dual effect of second order, i.e., the updated covariance \( \Sigma_{1k} \) is not a function of the past control sequence \( U^{k-1} \), for all \( k \).

Similarly to (2.14) this requirement can be written as follows:
\[ E\left[ \Sigma_{1k} | Y'(\omega, U^{k-1}), U^{k-1} \right] = E\left[ \Sigma_{1k} | P(\omega) \right], \]
\[ \text{a.s.,} (\omega), \forall U^{-1}, \forall j \leq k \forall k. \]  
(3.5)

**Proof:** Sufficiency: It will be proved by induction that the optimal cost-to-go will be of the form
\[ J_{k+1}^* = E\left[ x_{k+1}'P_{k+1}x_{k+1} + \alpha_{k+1} \right] \]  
(3.6)

where \( P_{k+1} \) is a constant nonnegative definite matrix independent of \( U^{-1} \) and \( \alpha_{k+1} \) is independent of \( U^k \) in the sense of (2.14), i.e.,
\[ E\left[ \alpha_{k+1} | Y'(\omega, U^{-1}), U^k \right] = E\left[ \alpha_{k+1} | Y'(\omega) \right], \]
\[ \text{a.s.,} (\omega), \forall U^{-1}, \forall j \leq k. \]  
(3.7)

For \( k + 1 = N \), it is easily seen that (3.6) is true with
\[ P_N = Q_N; \alpha_N = 0. \]  
(3.8)

Assuming that (3.6) and (3.7) are true and using the stochastic dynamic programming equation (2.18), one has
\[ J_k^* = \min J_k^* = \min \left\{ E\left[ x_k'Q_k x_k + u_k'R_k u_k + J_{k+1}^* | Y_k, U^{k-1} \right] \right\} \]
\[ = \min \left\{ E\left[ x_k'Q_k x_k + u_k'R_k u_k + J_{k+1}^* | Y_k, U^{k-1} \right] \right\} \]
\[ + E\left[ x_{k+1}'(u_k)P_{k+1}x_{k+1}(u_k) | Y_k, U^{k-1} \right] \]
\[ + E\left[ \alpha_{k+1} | Y_k, U^{k-1} \right] \]
\[ = \min \left\{ E\left[ x_k'Q_k x_k + F_kP_{k+1}x_k + u_k'G_kP_{k+1}G_k u_k \right. \right. \]
\[ + u_k'R_k G_kP_{k+1}G_k u_k \]
\[ + u_k'R_k G_kP_{k+1}G_k u_k \]
\[ + \left. \left. E\left[ \alpha_{k+1} | Y_k, U^{k-1} \right] \right. \right\} \]
\[ + \left. \left. \text{tr} \left( V_kP_{k+1} \right) + E\left[ \alpha_{k+1} | Y_k, U^{k-1} \right] \right. \right\} \]  
(3.9)

where (3.2), the whiteness of \( u_k \), and its independence of the measurement noises have been utilized. In view of this
\[ E\left[ \alpha_{k+1} | Y_k, U^{k-1} \right] = 0 \]  
(3.9a)
\[ E\left[ u_k | Y_k, U^{k-1} \right] = 0. \]  
(3.9b)

The optimal \( u_k^* \) which minimizes the right-hand side of (3.9) is
\[ u_k^* = -\left( R_k + G_kP_{k+1}G_k \right)^{-1} G_kP_{k+1}E\left[ x_k | Y_k, U^{k-1} \right] \]
\[ = L_kx_k. \]  
(3.10)

If \( u_k \) depends on, for instance, \( u_k \), then (3.9a) would not hold, in general, and \( u_k^* \) would not be given by (3.10). Therefore, this independence is needed to have, in general, the CE property hold. Substituting (3.10) into (3.9) and going through some algebraic manipulations, one has \( J_k^* \) of the form (3.6) with \( P_k \) and \( \alpha_k \), satisfying the following recursion relationships:
\[ P_k = F_k'P_{k+1} - \alpha_kG_kR_kG_kP_{k+1} + \alpha_kG_kR_kG_kP_{k+1} - \alpha_kG_kR_kG_kP_{k+1} \]
\[ + Q_k \]  
(3.11)
\[ \alpha_k = \text{tr} \left( F_k'P_{k+1}G_kR_kG_kP_{k+1} \right) \]
\[ + V_kP_{k+1} + E\left[ \alpha_{k+1} | P_k \right] \Delta g_k(R_k, \Sigma_{1k}) \]
\[ + E\left[ \alpha_{k+1} | P_k \right]. \]  
(3.12)

From (3.8), (3.10), and (3.11), it is readily seen that \( L_k \) is the optimal feedback gain as in the deterministic version of the problem (see, e.g., [3]). Subsequently, (3.5) and (3.12) imply that \( \alpha_k \) indeed has property (3.7), i.e., it is independent of \( U^{k-1} \), and thus, the induction proof of sufficiency is completed. Therefore, the certainty equivalence property holds when the system is "neutral" in the sense of (3.5), i.e., there is no dual effect of second order.

**Necessity:** To prove necessity, one has to show that if the optimal control is given by (3.10) for \( k = N - 1, \cdots, 0 \), this implies that \( \Sigma_{1k} \) is independent of \( u_k \) for all \( j < k \) for all \( k \). Assume that \( u_k^* \) is given by (3.10) for \( k = N - 1, \cdots, j + 1 \). Then
\[ J_{j+1}^* = E\left[ x_{j+1}'P_{j+1}x_{j+1} | Y_{j+1}, U^j \right] + \beta_{j+1} \]  
(3.13)

where
\[ \beta_{j+1} = \sum_{k=j+1}^{N-1} E\left[ g_k(R_k, \Sigma_{1k}) | Y_{j+1}, U^j \right] \]  
(3.14)

and \( g_k \) is as defined in (3.12). We know that \( u_k^* \) is obtained by minimizing
\[ J_k = E\left[ x_k'Q_k x_k + u_k'R_k u_k + J_{k+1}^* | Y_k, U^{k-1} \right] \]  
(3.15)

and it is easy to see that its expression will be given by (3.10) with \( k = j \) only if
1) \( \beta_{j+1} \) is independent of \( u_j \).
2) \( \beta_{j+1} \) is a function of \( u_j \) but assumes its minimum at \( u_j^* \) as given by (3.10) with \( k = j \).

However, notice that \( u_j^* \) is a function of \( R_j \), which does not enter into \( \beta_{j+1} \) according to (3.14); therefore, since we assumed that certainty equivalence holds for all \( R_j > 0 \),
2) cannot happen. Furthermore, the only way \( u_j \) might enter into \( \beta_{j+1} \) is via \( \Sigma_{1k}, k \geq j + 1 \) and, hence, \( \Sigma_{1k} \) is
independent of \( u_j, j > k \). Since this is true for \( j = 0, \ldots, N - 1 \), it follows that \( \Sigma \) is independent of \( U^{k-1} \), i.e., it is necessary that the control have no dual effect of second order.

This completes the proof.

**IV. Discussion and Examples**

Consider the following properties that characterize various versions of the stochastic control problem considered in Section III.

1. Linear dynamics with zero-mean white additive process noise independent of the measurement noise.
2. Fixed end-time, quadratic cost.
3. Gaussian process noise.
4. Linear measurement with additive noise.
5. White measurement noise sequence.
6. Gaussian measurement noise.

The Linear-Quadratic-Gaussian (LQG) problem [4], defined by properties 1–6 above, is well-known to have the certainty equivalence property (Joseph and Tou [14], Gunekel and Franklin [13]). Root [19] has proven that one can relax the Gaussian requirements on both the process and measurement noises, i.e., for the problem defined by 1), 2), 4), 5) the certainty equivalence property holds, too. It was also shown [28] that, in most cases, the process noise is still Gaussian, but the measurement noise is not white, i.e., 1)-4), 6) the optimal control is certainty equivalent. It is interesting to point out that in continuous time certainty equivalence was proven (see [8]) under the usual assumptions weakened by allowing correlation between the process and measurement noises.

A first new class of problems for which the theorem of Section III can be used to show that the CE property holds is the one defined by 1), 2), 4). Consider the system (3.1) with the measurements

\[
y_k = H_k x_k + w_k
\]

(4.1)

with no restrictions whatsoever on the measurement noise \( w_k \). Similarly to Wonham [31], let

\[
x_k = \tilde{x}_k + \bar{x}_k
\]

(4.2)

where \( \tilde{x}_k \) is the state of the autonomous part of system (3.1), i.e.,

\[
\dot{x}_k = F_k \tilde{x}_k + v_k
\]

(4.3)

with the corresponding part of the observation

\[
\bar{y}_k = H_k \tilde{x}_k + w_k
\]

(4.4)

and \( \bar{x}_k \) is the state of the forced, noiseless, part of the system (3.1)

\[
\bar{x}_k = F_k \bar{x}_k + G_k u_k
\]

(4.5)

with

\[
\bar{y}_k = H_k \bar{x}_k
\]

(4.6)

The initial state of the forced part of the system is taken as \( \bar{x}_0 = 0 \).

It can be easily shown by superposition, that, since everything is linear, \( \bar{x}_k \) obeys (3.1) and

\[
y_k = \bar{y}_k + \hat{y}_k
\]

(4.8)

Notice that \( \bar{x}_k \) and \( \hat{y}_k \) are exactly known at every time (because they are \( U^{k-1} \)-measurable).

Therefore,

\[
E \{ x_k | Y^k, U^{k-1} \} = E \{ \bar{x}_k | Y^k, U^{k-1} \} + \hat{\bar{x}}_k
\]

\[
= E \{ \bar{x}_k | P^k \} + \hat{\bar{x}}_k
\]

and

\[
\text{cov} \{ x_k | Y^k, U^{k-1} \} = \text{cov} \{ \bar{x}_k | P^k \}
\]

(4.10)

i.e., the control is neutral and the CE property holds.

A second example will illustrate how in a problem defined by 1), 2), i.e., with partially nonlinear measurement system, one has the certainty equivalence property. Let the system be

\[
x_{k+1} = \begin{bmatrix} x_{k+1}^1 \\ x_{k+1}^2 \\ x_{k+1}^3 \\ x_{k+1}^4 \\ x_{k+1}^5 \end{bmatrix} = \begin{bmatrix} F_{11}^k & F_{12}^k \\ 0 & F_{22}^k \end{bmatrix} \begin{bmatrix} x_k^1 \\ x_k^2 \end{bmatrix} + \begin{bmatrix} G_k^1 \\ 0 \end{bmatrix} u_k + \begin{bmatrix} w_k^1 \\ w_k^2 \end{bmatrix}
\]

(4.11)

and the measurement

\[
y_k = \begin{bmatrix} y_k^1 \\ y_k^2 \end{bmatrix} = \begin{bmatrix} H_k x_k^1 + w_k^1 \\ h_k (x_k^3, x_k^4, x_k^5) \end{bmatrix}
\]

(4.12)

i.e., the part of the state denoted by \( x_k^4 \) is measured via a nonlinear device. However, since \( x_k^4 \) is not a function of the past controls—this follows from the particular form of \( F_k \) in this problem—it is readily apparent that the control is neutral. In general, for a system which is stabilizable [32] but not completely controllable, if there is a nonlinear observation on only the uncontrollable subspace, then CE holds. Therefore, even in certain systems with nonlinear observations, the certainty equivalence property holds.

**V. Conclusion**

There is a distinction between feedback and closed-loop policies in stochastic control and it is only the closed-loop policy that takes into account the possible estimation benefit to be derived from future observations. This leads to the active learning feature of the closed-loop control policy that can be utilized when the control has a dual effect. There is an intimate connection between the dual effect of the control and the certainty equivalence property. The class of problems for which it was previously known that the certainty equivalence property (and, hence, separation) holds has been expanded.

**Acknowledgment**

The authors would like to thank Dr. A. Segall for valuable comments and criticism and Prof. R. Sivan for stimulating discussions.
REFERENCES


Yakov Bar-Shalom (S’63–M’66) for a photograph and biography see page 7 of the February issue of this TRANSACTIONS.