

## NONLINEAR OPTIMAL STOCHASTIC CONTROL

### - SOME APPROXIMATIONS WHEN THE NOISE IS SMALL

by J.D. Perkins, Department of Chemical Engineering, University  
of Cambridge

and R.W.H. Sargent, Department of Chemical Engineering and  
Chemical Technology, Imperial College, London.

#### 1. Introduction

We consider the design of optimal controllers for the following nonlinear stochastic system:

##### System S1

The evolution of the "state" is described by

$$\dot{x}(t) = f(t, x(t), u(t), \xi(t)), \quad t \in [0, t_f], \quad (1)$$

$$\text{with measurements } y(t) = g(t, x(t), \eta(t)), \quad (2)$$

$$\text{controls } u(t) = h(t, I(t)) \in U, \quad (3)$$

$$\text{and performance index } J = E\{c^T x(t_f)\}. \quad (4)$$

Here  $x(t) \in E^n$  is an  $n$ -vector of "state" variables, and the initial "state"  $x(0) \in X(0)$  with probability one, where  $X(0) \subset E^n$  is a given convex, bounded set. The system noise,  $\xi(t) \in V$  w.p.l., and the measurement noise,  $\eta(t) \in W$  w.p.l., are both vector stochastic processes and  $V \subset E^r$ ,  $W \subset E^q$  are given convex, bounded sets of possible values.  $U \subset E^m$  is a given set of admissible values for the controls  $u(t)$ , which are vector-valued functions  $h(t, I(t))$  of the set of information  $I(t)$  available to the controller at time  $t$ ;  $I(t)$  is a specified subset of the a priori information  $I(0)$  and the measurements  $y(t)$ ,  $t \in [0, t_f]$ . In the expression for the performance index,  $c \in E^n$  is a given vector, and the expectation is taken over all possible initial conditions  $x(0)$  and all possible realizations of the noise processes  $\xi(t)$  and  $\eta(t)$ .

The control design problem is to find the function  $h(t, I(t))$ ,  $t \in [0, t_f]$  which minimizes the performance index  $J$  for specified information sets  $I(t)$ ,  $t \in [0, t_f]$ .

Following Joffe and Sargent<sup>(1)</sup>, we appeal to the theory of ordinary differential equations with Lebesgue-measurable input functions in order to obtain a set of assumptions which guarantee the existence and uniqueness of solutions to equation (1) for given initial condition  $x(0)$ , noise realization  $\xi(t)$  and admissible control  $u(t)$ ,  $t \in [0, t_f]$ :-

- (i) Each admissible control  $u(t)$  and each noise realization  $\xi(t)$  is a measurable function of  $t$  on the interval  $[0, t_f]$ .
- (ii) For each possible set of values  $x, u, \xi$ , the function  $f(t, x, u, \xi)$  is a measurable function of  $t$  on the interval  $[0, t_f]$ .
- (iii) For each  $t \in [0, t_f]$ ,  $f(t, x, u, \xi)$  is continuous in  $x, u$ , and  $\xi$  for all possible

values of these variables.

- (iv) There exists a function  $S(t)$ , summable on  $[0, t_f]$ , and a function  $\phi(z)$ , positive and continuous for  $z \geq 0$  but not summable on  $[0, \infty)$ , such that

$$||f(t, x, u, \xi)|| \leq S(t)/\phi(||x||)$$

for each possible  $t$ ,  $x$ ,  $u$ , and  $\xi$ .

- (v) There exists a function  $M(t)$ , summable on  $[0, t_f]$ , such that

$$|(x-x')^T(f(t, x, u, \xi) - f(t, x', u, \xi))| \leq M(t) \cdot ||x-x'||^2$$

for all possible  $t$ ,  $x$ ,  $x'$ ,  $u$  and  $\xi$ .

These conditions also ensure that the set  $X(t_f)$  of states reachable w.p.1. with some possible initial condition, control and noise realization, is a bounded set.

We note that by requiring  $\xi(t)$  to be measurable in  $t$  we exclude the possibility that  $\xi(t)$ , and hence also  $x(t)$ , are Markov processes. Thus  $x(t)$  does not represent a state in the stochastic sense, but for a given realization of the system noise it has all the properties required of a state vector for a deterministic system.

It is well known that systems with more general performance indices than (4) can often be put in the form of S1 by adjoining extra state variables with appropriate definitions. These systems can therefore be treated, provided that the adjoined variables and functions also satisfy the above conditions.

A variety of control problems can be posed for S1 by suitable choice of the information sets  $I(t)$ . For example this set may consist of just the current set of measurements  $y(t)$ , or perhaps past measurements at a sequence of times or over a specified time interval. Feed-forward control policies are obtained if  $I(t) = I(0)$ ,  $t \in [0, t_f]$ . Joffe and Sargent<sup>(1)</sup> considered the cases of feed-forward control, and feed-back control based on continuous exact measurement of the state  $x(t)$ . They also considered sub-optimal controls based on the functional form;

$$u(t) = h(t, x(t), \alpha(t)) \quad (5)$$

where  $h(t, x, \alpha)$  is a given heuristically derived function, and  $\alpha(t)$  is a vector of parameters which are chosen optimally.

Their approach was to assume that the noise  $\xi(t)$  and the initial state  $x(0)$  satisfy certain smallness conditions, and then to derive a deterministic model which predicts  $E\{x(t)\}$  for a specified control policy. In this paper we shall extend their treatment to deal with more general feed-back controls.

## 2. A Modelling Theorem

Suppose that we have a system as described in S1, and a mathematical model of this system which enables us to predict the performance index for any initial condition and any control law of the form of equation (3). More precisely, if  $J_s(h)$  is the value of the system performance index, using the control function  $h(t, I(t))$  with specified  $I(t)$ ,  $t \in [0, t_f]$  and initial condition  $x(0)$ , and  $J_m(h)$  is the value of the model performance index using the same control function, information sets and initial con-

dition, then there exists a  $\delta \geq 0$  such that

$$|J_S(h) - J_m(h)| \leq \delta \quad (6)$$

for all possible  $h(\cdot, \cdot)$ ,  $I(t)$  and  $x(0)$ .

Now suppose that we have solved the control design problem for the system for a given specification of the  $I(t)$ , obtaining the optimal system control function  $h_S(t, I(t))$  with corresponding performance index  $J_S(h_S)$ . Similarly, we solve the design problem for the model with the same specification of the  $I(t)$ , obtaining the optimal model control function  $h_m(t, I(t))$  with performance index  $J_m(h_m)$ . Then we have the following theorem:

Theorem 1

If the system and model satisfy condition (6), and if  $h_S(t, I(t))$  and  $h_m(t, I(t))$  are the optimal system and model control laws respectively for specified information sets  $I(t)$ ,  $t \in [0, t_f]$ , then

$$|J_S(h_m) - J_S(h_S)| \leq 2\delta, \quad |J_m(h_m) - J_S(h_S)| \leq 3\delta \quad (7)$$

Proof

From condition (6) we have

$$|J_S(h_m) - J_m(h_m)| \leq \delta, \quad |J_S(h_S) - J_m(h_S)| \leq \delta \quad (8)$$

It follows from the triangle inequality that

$$|J_S(h_m) - J_m(h_m) + J_m(h_S) - J_S(h_S)| \leq 2\delta$$

But from the optimality conditions  $J_S(h_S) \leq J_S(h_m)$  and  $J_m(h_m) \leq J_m(h_S)$ , and hence

$$|J_S(h_m) - J_S(h_S)| \leq 2\delta, \quad |J_m(h_S) - J_m(h_m)| \leq 2\delta$$

To obtain the second inequality in (7), we use the first part and condition (8):

$$\begin{aligned} |J_m(h_m) - J_S(h_S)| &= |J_m(h_m) - J_S(h_m) + J_S(h_m) - J_S(h_S)| \\ &\leq |J_m(h_m) - J_S(h_m)| + |J_S(h_m) - J_S(h_S)| \leq 3\delta \end{aligned}$$

Q.E.D.

This theorem tells us that the model optimal control law applied to the system will produce a performance index within  $O[\delta]$  of the true optimal result, and that the optimal performance index  $J_m(h_m)$  computed from the model is an  $O[\delta]$  approximation to the true optimal value.

It therefore remains to find good manageable models for nonlinear stochastic systems, and since  $J$  is an expected value it is worth looking for deterministic models for its prediction.

### 3. Deterministic Models for Small-Noise Stochastic Systems

If, in addition to the conditions listed in Section 1, the second partial derivatives of  $f(t, x, u, \xi)$  and  $g(t, x, \eta)$  with respect to  $x$ ,  $\xi$ ,  $\eta$  exist, and are Lipschitz continuous on  $X(t_f) \times V$  and  $X(t_f) \times W$  respectively, then it is possible to use Taylor expansions to obtain the following models:

Model M1

$$\dot{x}^*(t) = f(t, x^*(t), u(t), \xi^*(t)) \quad , \quad t \in [0, t_f] \quad (9)$$

$$y^*(t) = g(t, x^*(t), \eta^*(t)) \quad (10)$$

$$J^* = c^T x^*(t_f) \quad (11)$$

$$\text{where } \xi^*(t) = E\{\xi(t)\} \quad , \quad \eta^*(t) = E\{\eta(t)\} \quad (12)$$

Model M2

$$\dot{x}^*(t) = f(t, x^*(t), u(t), \xi^*(t)) + z(t) \quad , \quad t \in [0, t_f] \quad (13)$$

$$y^*(t) = g(t, x^*(t), \eta^*(t)) + \zeta(t) \quad (14)$$

$$J^* = c^T x^*(t_f) \quad (15)$$

$$\text{where } \dagger \quad z^k(t) = \frac{1}{2} \text{ trace } [f_{xx}^k P(t) + f_{\xi\xi}^k Q(t)] \quad (16)$$

$$\zeta^k(t) = \frac{1}{2} \text{ trace } [g_{xx}^k P(t) + g_{\eta\eta}^k R(t)] \quad (17)$$

$$\dot{P}(t) = f_{xx}^k P + P f_{xx}^T \quad (18)$$

$$\begin{aligned} Q(t) &= E\{\delta\xi(t) \cdot \delta\xi^T(t)\} \\ R(t) &= E\{\delta\eta(t) \cdot \delta\eta^T(t)\} \\ \xi^*(t) &= E\{\xi(t)\} \quad \delta\xi(t) = \xi(t) - \xi^*(t) \\ \eta^*(t) &= E\{\eta(t)\} \quad \delta\eta(t) = \eta(t) - \eta^*(t) \end{aligned} \quad (19)$$

We then have the following basic approximation theorem for Model M2:

Theorem 2

Suppose that, in addition to the assumptions stated previously for system S1, the following inequalities hold for some finite, positive constants A, B, C,  $\epsilon$ , and each  $t \in [0, t_f]$ :

$$\begin{aligned} \int_0^{t_f} E\{|\delta\xi(t)|^3\} dt &\leq t_f \epsilon^3 & E\{|\delta x(0)|^3\} &\leq A^3 \epsilon^3 \\ \int_0^{t_f} E\{|\delta\xi(t)|^2 |\delta\xi(t')|\} dt &\leq t_f \epsilon^3 & E\{|\delta x(0)|^2 |\delta\xi(t')|\} &\leq A^2 \epsilon^3 \\ \int_0^{t_f} E\{|\delta\xi(t)| \cdot |\delta\xi(t')|^2\} dt &\leq t_f \epsilon^3 & E\{|\delta x(0)| \cdot |\delta\xi(t')|^2\} &\leq A \epsilon^3 \\ \int_0^{t_f} |E\{\delta\xi(t) \cdot \delta\xi^T(t')\}| dt &\leq t_f \epsilon^3 & |E\{\delta x(0) \cdot \delta\xi^T(t')\}| &\leq A \epsilon^3 \\ E\{|\delta\eta(t)|^3\} &\leq B^3 \epsilon^3 & E\{|\delta x(t)| \cdot |\delta\eta(t)|^2\} &\leq AB^2 \epsilon^3 \\ |E\{\delta x(t) \cdot \delta\eta^T(t)\}| &\leq AB \epsilon^3 & E\{|\delta x(t)|^2 |\delta\eta(t)|\} &\leq A^2 B \epsilon^3 \\ |x^*(0) - E\{x(0)\}| &\leq C \epsilon^3 \quad , \quad |P(0) - E\{\delta x(0) \cdot \delta x^T(0)\}| &\leq C \epsilon^3 \quad (21) \\ \text{where } \delta x(t) &= x(t) - E\{x(t)\}. \end{aligned}$$

Then, for each admissible control  $u(t)$ ,  $t \in [0, t_f]$  and given initial conditions  $x^*(0)$ ,  $P(0)$ , there exists a unique solution to equations (13) - (19) of Model M2, and this

<sup>†</sup> Notation Subscripts denote differentiation with respect to the variables concerned, and superscripts denote elements of vectors. Arguments of functions, where omitted, are  $(t, x^*(t), u(t), \xi^*(t))$ .

solution satisfies the inequalities:

$$\begin{aligned} ||x^*(t) - E\{x(t)\}|| &\leq K_1 \varepsilon^3, \quad ||P(t) - E\{\delta x(t) \cdot \delta x^T(t)\}|| \leq K_2 \varepsilon^3 \\ ||y^*(t) - E\{y(t)\}|| &\leq K_3 \varepsilon^3, \quad |J^* - J| \leq K_1 \varepsilon^3 ||c|| \end{aligned} \quad (22)$$

where  $K_1$ ,  $K_2$  and  $K_3$  are finite positive numbers.

Joffe and Sargent<sup>(1)</sup> give the proof of this theorem without the results on  $y^*(t)$ , and on the assumption that  $C = 0$ . However the required extension of their proof is immediate and will not be given here. It is useful to note at this point that if  $C = 0$  then  $K_1$  and  $K_2$  are both directly proportional to  $t_f$ .

We also note that if  $X(0)$  consists of a single point there is no uncertainty in the initial state and its covariance matrix is zero. We may therefore set  $x^*(0) = x(0)$  and  $P(0) = 0$ , so that  $C$  can be taken as zero; it also follows from (18) that  $P(t) = 0$  for all  $t \in [0, t_f]$ , thereby removing a matrix differential equation from the model and simplifying  $z(t)$ .

The same proof shows that Model M1 provides  $O[\varepsilon^2]$  bounds for the quantities in (22), and again it is not necessary to generate  $P(t)$  in order to obtain  $x^*(t)$ .

#### 4. Optimal Controller Design

If the information sets  $I(t)$ ,  $t \in (0, t_f]$  are empty the control is simply a function of time, and Theorem 1 applies with optimal control policies  $h_s(t)$  and  $h_m(t)$  for the system and model respectively. Model M2 satisfies condition (6) with  $\delta = K_1 \varepsilon^3 ||c||$ , and it follows that the deterministic optimal control policy for Model M2 is within  $O[\varepsilon^3]$  of the true optimal policy for System S1. Similarly Model M1 approximates the true policy with an error of  $O[\varepsilon^2]$ .

Theorem 2 does not apply as it stands to general feed-back control laws as given by equation (3). However, if the information sets  $I(t)$  are written in terms of the measurements  $y(t)$ , which are in turn expressed in terms of  $x(t)$  through equation (2), then substitution of the resulting control law into equation (1) yields an equation of the same form as equation (1) with a given control policy  $u(t)$  substituted, and Theorem 1 then yields the general result that the optimal model control law is a suitable approximation to the true stochastic optimal law for any feed-back control law as defined in equation (3).

We note that the small-noise and smoothness assumptions of Theorem 2 must now apply to the substituted form of equation (1), and hence to the combined stochastic processes  $\xi(t)$  and  $\eta(t)$ . Because of the conditions already imposed on  $g(t, x, \eta)$ , the smoothness conditions will be satisfied if  $f(t, x, u, \xi)$  is twice differentiable in  $u$ , and  $h(t, I(t))$  is twice differentiable in its arguments, and all these second derivatives are Lipschitz continuous in the attainable domain of their arguments. Of course these strong smoothness assumptions may be unduly restrictive, and exclude control laws giving much more favourable performance. A further difficulty arises if the information set  $I(t)$  involves information at times other than the current time,

for the composite substituted equation then describes a generalized dynamical system falling outside the scope of the theory of differential equations used in Theorem 2. It turns out that these difficulties can be overcome, and a more general form of Theorem 2 obtained<sup>(2)</sup>, as will be shown in a separate paper. For the present discussion however, we shall restrict ourselves to the subclass of smooth control functions satisfying the above conditions, with the information sets  $I(t)$ ,  $t \in (0, t_f]$  either empty or consisting only of current measurements  $y(t)$ .

#### 4.1 Problems with Known Initial State

In an important class of problems the initial state  $x(0)$  is known to good accuracy - for example in problems of optimum start-up of a process, or change of regime from one steady state to another, or where the state variables can all be measured with negligible error.

In such cases the smallness conditions concerning the initial state are satisfied with  $A = 0$ , and we may set  $x^*(0) = x(0)$ ,  $P(0) = 0$  in Model M2, satisfying (21) with  $C = 0$ . The model then generates second-order estimates of expected quantities conditioned on the known initial state, and Theorems 1 and 2 show that the optimal model control law is a second-order approximation to the stochastic optimal control for the system for any feed-back control law.

However, we note that Model M2 is a deterministic system with given initial conditions, which generates for itself second-order approximations to the measurements  $y(t)$  for all  $t \in [0, t_f]$ . Thus the optimal feed-forward policy based on the given initial state will be within a margin of  $O[\varepsilon^3]$  of the optimal result for any feed-back control law, and it follows that little advantage is to be gained from the use of measurements in a feed-back controller in such a situation.

On the other hand, we noted in Section 3 that  $K_1$ , and hence the error, is directly proportional to  $t_f$ , so that the uncertainty grows with time. If the time interval of operation is large it may therefore be worth utilizing the extra information gained from measurements to compensate for this growth of uncertainty, and this brings us to the case where the initial state is no longer determined to high precision.

#### 4.2 Problems with Unknown Initial State

We now consider the situation where we know only the expected value of the initial state with its corresponding covariance matrix, conditioned on the initial information available  $I(0)$ .

If we set  $x^*(0) = E\{x(0) | I(0)\}$  and  $P(0) = E\{\delta x(0) \delta x^T(0) | I(0)\}$ , then Model M2 generates approximations to the corresponding conditional expectations, and we obtain the same result for the optimal model and system control laws for a given specification of the information sets  $I(t)$ . The design of the optimal feed-back control for the deterministic model is still much simpler than for the original stochastic system, but it is nevertheless a substantial problem, and it is worth seeking further insight.

If the information sets  $I(t)$ ,  $t \in (0, t_f]$  are empty, then the model control is simply a

feed-forward control policy, and the optimum policy for the given initial state is easily obtained by standard methods. With non-empty information sets it is clear that the optimal feed-back control must implicitly make use of the additional information to improve the estimate of the state, and it is of interest to examine the conjecture that an approximate form of Wonham's Separation Theorem<sup>(3)</sup> will apply to the nonlinear stochastic system. This would imply that the optimal controller can be split into an estimator, producing an estimate  $x^*(t)$  of the state from the measurements, and an optimal feed-back controller based on continuous observation of the state, which uses  $x^*(t)$  in place of the actual state.

In this connection we have the following theorem:

### Theorem 3

The performance of System S1 is compared under the two control laws

$$u(t) = h(t, x(t)) \quad (23a)$$

$$\text{and} \quad u(t) = h(t, x^*(t)) \quad (23b)$$

giving rise to states  $x_a(t)$  and  $x_b(t)$  respectively, where  $x^*(t) \in X(t_f)$  and  $h(t, x)$  are measurable functions of  $t$  on  $[0, t_f]$ , and  $h(t, x)$  is Lipschitz continuous in  $x$  on  $X(t_f)$ .

Then if the corresponding performance indices are  $J_a$  and  $J_b$  we have

$$\begin{aligned} E\{ \|x_a(t) - x_b(t)\| \} &\leq K \int_0^t E\{ \|x_b(\tau) - x^*(\tau)\| \} d\tau \\ |J_a - J_b| &\leq K \|c\| \int_0^{t_f} E\{ \|x_b(\tau) - x^*(\tau)\| \} d\tau \end{aligned} \quad (24)$$

for all  $t \in [0, t_f]$  and some  $K \in [0, \infty)$ .

### Proof

We compare performances with a given noise realization and initial state:

$$\begin{aligned} \dot{x}_a(t) &= f(t, x_a(t), h(t, x_a(t)), \xi(t)) \quad , \quad x_a(0) = x_0 \\ \dot{x}_b(t) &= f(t, x_b(t), h(t, x^*(t)), \xi(t)) \quad , \quad x_b(0) = x_0 \end{aligned} \quad (25)$$

From the properties of  $f(t, x, u, \xi)$  and  $h(t, x)$  we have w.p.1.

$$\begin{aligned} \|f(t, x_a, u, \xi) - f(t, x_b, u, \xi)\| &\leq L_x \|x_a - x_b\|, \\ \|f(t, x, u_a, \xi) - f(t, x, u_b, \xi)\| &\leq L_u \|u_a - u_b\|, \\ \|h(t, x_a) - h(t, x_b)\| &\leq L_h \|x_a - x_b\|, \end{aligned} \quad (26)$$

for some finite, positive constants  $L_x, L_u, L_h$ . Then from (25) and (26):

$$\begin{aligned} \frac{d}{dt} \|x_a(t) - x_b(t)\| &\leq \|f(t, x_a(t), h(t, x_a(t)), \xi(t)) - f(t, x_b(t), h(t, x_b(t)), \xi(t))\| \\ &\quad + \|f(t, x_b(t), h(t, x_b(t)), \xi(t)) - f(t, x_b(t), h(t, x^*(t)), \xi(t))\| \\ &\leq (L_x + L_u L_h) \|x_a(t) - x_b(t)\| + L_u L_h \|x_b(t) - x^*(t)\| \end{aligned} \quad (27)$$

Taking expectations of (27), noting that  $\|x_a(0) - x_b(0)\| = 0$ , and applying Halkin's "Generalized Gronwall Lemma"<sup>(1)</sup>, we obtain the first relation in (24) with

$$K = L_u L_h \exp(L_x + L_u L_h) t_f < \infty$$

and the second relation follows immediately from the definition of  $J$ .

Q.E.D.

Of course, Models M1 and M2 can themselves be used as estimators, and with additional smoothness conditions we can, for example, prove the following:

Theorem 4

The performance of System S1 is compared under the control laws given in (23a) and (23b), where  $h(t, x)$  is measurable in  $t$  and has continuous derivatives with respect to  $x$  on  $[0, t_f] \times X(t_f)$ , and  $x^*(t)$  is the solution for Model M1 with  $u(t)$  given by (23b) and initial condition  $x^*(0) = E\{x(0) | I(0)\}$ . Then with the small-noise conditions of (20):

$$||E\{x_a(t) - x_b(t)\}|| \leq K_1 \int_0^t ||E\{x_b(\tau) - x^*(\tau)\}|| d\tau + K_2 \varepsilon^2 \quad (28)$$

for some finite  $K_1$  and  $K_2$ , where all expectations are conditioned on the initial information  $I(0)$ .

Proof

$$\begin{aligned} \text{Write } \delta x_a(t) &= x_a(t) - x^*(t), \quad \delta x_b(t) = x_b(t) - x^*(t) \\ \delta \xi(t) &= \xi(t) - \xi^*(t), \quad \xi^*(t) = E\{\xi(t) | I(0)\} \end{aligned}$$

Using (26) with a similar Lipschitz condition for  $\xi$ :

$$||f(t, x, u, \xi) - f(t, x, u, \xi^*)|| \leq L_\xi ||\xi - \xi^*||, \quad (26')$$

we obtain from (9) and (25)

$$\begin{aligned} \frac{d}{dt} ||\delta x_a(t)|| &\leq (L_x + L_u L_h) ||\delta x_a(t)|| + L_\xi ||\delta \xi(t)||, \\ \frac{d}{dt} ||\delta x_b(t)|| &\leq L_x ||\delta x_b(t)|| + L_\xi ||\delta \xi(t)||, \end{aligned}$$

and using the Generalized Gronwall Lemma these yield

$$\begin{aligned} ||\delta x_a(t)|| &\leq \{ ||\delta x_a(0)|| + L_\xi \int_0^t ||\delta \xi(\tau)|| d\tau \} \cdot \exp(L_x + L_u L_h)t, \\ ||\delta x_b(t)|| &\leq \{ ||\delta x_b(0)|| + L_\xi \int_0^t ||\delta \xi(\tau)|| d\tau \} \cdot \exp(L_x t). \end{aligned} \quad (29)$$

Taking expectations of (29) and using (20) we obtain

$$\begin{aligned} E\{ ||\delta x_a(t)|| \} &\leq K_a \varepsilon, \quad K_a = (A + L_\xi t) \cdot \exp(L_x + L_u L_h)t, \\ E\{ ||\delta x_b(t)|| \} &\leq K_b \varepsilon, \quad K_b = (A + L_\xi t). \end{aligned} \quad (30)$$

Also from (20), (29) and (30), using the Hölder inequality for both expectations and integrals, we can obtain

$$\begin{aligned} E\{ ||\delta x_a(t)||^2 \} &\leq K_a^2 \varepsilon^3, \quad E\{ ||\delta x_b(t)||^2 \} \leq K_b^2 \varepsilon^2 \\ \int_0^t E\{ ||\delta x_a(t)|| \cdot ||\delta \xi(t)|| \} dt &\leq K_a \varepsilon^2, \quad \int_0^t E\{ ||\delta x_b(t)|| \cdot ||\delta \xi(t)|| \} dt \leq K_b \varepsilon^2 \end{aligned} \quad (31)$$

Now from (25), expanding about the values  $t$ ,  $x^*(t)$ ,  $\xi(t)$  we have

$$\begin{aligned} \dot{x}_a(t) &= f + (f_x + f_u h_x) \cdot \delta x_a(t) + v_a(t), \quad ||v_a|| \leq L_a ||\delta x_a(t)||^2 \\ \dot{x}_b(t) &= f + f_x \cdot \delta x_b(t) + v_b(t), \quad ||v_b|| \leq L_b ||\delta x_b(t)||^2 \end{aligned}$$

for some finite, positive  $L_a$ ,  $L_b$ . But we may now expand the derivatives  $f_x$ ,  $f_u$  about the value  $\xi^*(t)$ , and since the second derivatives  $f_{x\xi}$ ,  $f_{u\xi}$  are bounded there exists an  $L < \infty$  such that

$$\frac{d}{dt} \{x_a(t) - x_b(t)\} = (f_x^* + f_u^* h_x^*) \{x_a(t) - x_b(t)\} + f_u^* h_x^* \delta x_b(t) + w(t) \quad (32)$$

where  $f_x^*$ ,  $f_u^*$ ,  $h_x^*$  are evaluated for  $t$ ,  $x^*(t)$ ,  $\xi^*(t)$  - and hence are deterministic - and  $w(t)$  satisfies:



$$\|w(t)\| \leq L\{\|\delta x_a(t)\|^2 + \|\delta x_b(t)\|^2 + \|\delta x_a(t)\| \cdot \|\delta \xi(t)\| + \|\delta x_b(t)\| \cdot \|\delta \xi(t)\|\} \quad (33)$$

Taking expectations and then norms of (32) leads to

$$\frac{d}{dt} \|E\{x_a(t) - x_b(t)\}\| \leq (L_X + L_U L_H) \|E\{x_a(t) - x_b(t)\}\| + L_U L_H \|E\{\delta x_b(t)\}\| + E\{\|w(t)\|\} \quad (34)$$

and applying the Generalized Gronwall Lemma and using (33) and (31) finally yields (28) with  $K_1 = L_U L_H \exp(L_X + L_U L_H)t$  and  $K_2 = \{K_a + K_b + K_a^2 + K_b^2\} \cdot L \exp(L_X + L_U L_H)t_f$

Q.E.D.

From the analogue of Theorem 2 for Model M1 we note further that the integral on the right-hand side of (28) is  $O[\varepsilon^2]$ . A similar proof will show that Model M2 provides an  $O[\varepsilon^3]$  error for  $\|E\{x_a(t) - x_b(t)\}\|$  and hence for the difference in performance indices.

Theorem 3 is of interest since it does not depend on the form of the estimator and shows that the loss in performance of the "separated" controller depends only on the integral error of the estimate, allowing for example a poor estimate provided that the time over which it is used is short. Note however that the error also depends on the magnitude of the noise  $\xi(t)$  since  $x_b(\tau)$  is the actual state, and in fact, Theorem 3 places quite strong conditions on both noise and estimate to ensure a small loss in performance.

Condition (28) of Theorem 4 is weaker, but the small-noise conditions (20) in essence require that the uncertainty in the estimate of the initial state is small. This brings us back to the situation considered in Section 4.1, and indeed, since Models M1 and M2 as given in Section 3 simply predict states from the initial state estimate, without utilizing information from measurements, it is not surprising that we reach the same conclusion. However Theorem 4 is also of use if the state estimate is updated. Suppose, for example, that the state can be measured with negligible error at a sequence of times  $t_0 = 0, t_1, t_2, \dots$  and Model M1 is used for prediction of the state within these intervals, with  $x^*(t_k) = x_b(t_k)$  as initial condition. Then we may set  $A = 0$  and note from (30) that  $K_a$  and  $K_b$  are  $O[t]$ . If the integral conditions in (20) apply over each subinterval  $[t_k, t_{k+1}]$ , more careful integration of (34) then shows that  $\|E\{x_a(t_f) - x_b(t_f)\}\|$  is  $O[\varepsilon^2 t_f \Delta t]$ , where  $\Delta t$  is the duration of the largest interval between measurements. Hence, as would be expected, the loss in performance can be reduced by increasing the frequency of measurement.

If only noisy measurements are available, as in (2), then Model M1 or M2 can be used as a predictor in an estimator which makes use of these measurements. For example, if measurements  $y(t_k)$ ,  $k = 0, 1, 2, \dots$  are available, then an estimator based on Model M2 is as follows:

- (i) Given estimates  $x^*(t_{k-1}|I(t_{k-1}))$ ,  $P(t_{k-1}|I(t_{k-1}))$ , these are used as initial conditions with equations (13) - (19) to generate predictions  $x^*(t_k|I(t_{k-1}))$ ,  $P(t_k|I(t_{k-1}))$ .
- (ii) These predictions are updated using the measurements at  $t_k$  by the equations:
 
$$K(t_k) = P(t_k|I(t_{k-1})) \cdot (g_x^*)^T [g_x^* P(t_k|I(t_{k-1})) (g_x^*)^T + R(t_k)]^{-1}$$

$$P(t_k|I(t_k)) = P(t_k|I(t_{k-1})) - K(t_k)g_x^*P(t_k|I(t_{k-1})) \quad (35)$$

$$x^*(t_k|I(t_k)) = x^*(t_k|I(t_{k-1})) - K(t_k)(y(t_k) - y^*(t_k))$$

where  $g_x^*$  is the partial derivative with respect to  $x$  of the right-hand side of equation (14); in the iterated form of the estimator it is the arithmetic mean of the values at  $t_{k-1}$  and  $t_k$ .

The proof of Theorem 4 makes no use of the nature of the conditioning in the expectations, and hence the method of proof remains valid for the above estimator, the update of (35) serving to reduce the magnitude of the error due to the integral in (28). It is important to note that the update also reduces the variance of the state, and hence the variance of the performance index, so that feed-back increases the probability of the achieved performance being close to the expected optimal value.<sup>(2)</sup>

The analogy with the extended Kalman filter<sup>(4)</sup> is obvious, and a similar estimator and corresponding theorem can be developed for the case of continuous measurements. The small-noise assumptions concerning the initial state are still restrictive, and in essence are made necessary by the fact that comparison is made between use of the estimator and continuous exact measurement of the state. What is really required is a comparison with the true optimal controller based on the same information sets, but since all the past information is available to the controller this inevitably leads us to consideration of a generalized dynamical system, which is outside the scope of the present paper.

## 5. Conclusions

Stochastic control problems for a rather general class of nonlinear systems have been considered in this paper. Deterministic models for the prediction of expected performance have been presented, and it has been shown that optimal controllers designed using the models are good approximations to the true optimal controllers for the stochastic system. Some results have been presented on the use of an estimator coupled with an optimal controller based on continuous observation of the state.

The problems associated with the non-Markov nature of the system, and relaxation of the smoothness conditions imposed on the state and control laws will be the subject of a separate paper.

## Acknowledgments

One of the authors (J. D. Perkins) wishes to thank the Salters' Company for the award of a Scholarship.

## References

1. Joffe, B. L., and R. W. H. Sargent, "A Naïve Approach to the Optimal Control of Nonlinear Stochastic Systems", in "Recent Mathematical Developments in Control", D. J. Bell (Ed.), Academic Press (1973)
2. Perkins, J. D., "Optimal Stochastic Control of Continuous Processes - Some Approximations Applicable to Chemical Engineering Problems", Ph.D. Thesis, London (1976)
3. Wonham, W. M., "On the Separation Theorem of Stochastic Control", SIAM J. Control,

6 (2), 312 (1968)

4. Jazwinski, A. H., "Stochastic Processes and Filtering Theory", (Academic Press, New York, 1970)