STOCHASTIC OPTIMAL CONTROL OF PARTIALLY OBSERVABLE NONLINEAR SYSTEMS

Tayfun Çimen[a] & Stephen P. Banks[b]

[a]*Turkish Naval Forces, Turkish Naval Research Center Command (TNRCC), Ankara Cad. No: 252, 34890 Pendik-Istanbul, Turkey*
[b]*Department of Automatic Control and Systems Engineering, The University of Sheffield, Mappin Street, Sheffield S1 3JD, UK*

Abstract: This paper presents a new theory for solving the continuous-time stochastic optimal control problem for a very general class of nonlinear (nonautonomous and nonaffine controlled) systems with partial state information. The proposed theory transforms the nonlinear problem into a sequence of linear-quadratic Gaussian (LQG) and time-varying problems, which converge (uniformly in time) under very mild conditions of *local Lipschitz continuity*. These results have been previously presented for deterministic nonlinear systems under perfect state measurements for finite horizons, but the present study shows how an additional class of nonlinear problems, involving partially observable stochastic systems, can be handled with the same theory. The method introduces an "*approximating sequence of Riccati equations*" (*ASRE*) to explicitly find the error covariance matrix and nonlinear time-varying optimal feedback controllers for such nonlinear systems, which is achieved using the framework of Kalman-Bucy filtering, separation principle and LQR theory. The paper shows a practical way of designing optimal feedback control systems for complex nonlinear stochastic problems using a combination of modern LQG estimation and LQ control-design methodologies.

Keywords: Nonlinear non-affine systems, Stochastic systems, Optimal control, Optimal filtering, Kalman filter, LQG control, Separation principle, Approximation theory, Iterative methods, Continuous-time systems, Time-varying systems.

1. INTRODUCTION

Nonlinear optimal control theory has advanced apace in recent years. Research effort has been directed towards theoretical challenges in deterministic optimal control of general nonlinear systems, which have resulted in the development of innovative and practically useful designs that significantly extend existing theory. Recently, a novel algorithm has been proposed in Çimen and Banks (2004a, b) for solving finite-time nonlinear deterministic optimal regulator and tracking control problems. The nonlinear optimal control problem is transformed into a sequence of linear-quadratic (LQ) and time-varying optimal control problems, which can be solved using well-known results from existing theory. Under very mild conditions of *local Lipschitz continuity*, the limit of the approximating sequences has been shown to converge *globally* in time. The proposed method does *not* involve any partial derivatives, and only requires solving an "***A**pproximating **S**equence of **R**iccati **E**quations*" (***ASRE***), which can be achieved by classical methods. These ASRE solutions converge to nonlinear time-varying *feedback* controllers for such nonlinear systems. The ASRE theory has been illustrated in designing autopilots for complex nonlinear models of practical real-world applications, including super-tankers and fighter aircraft systems. It has been shown that automatic finite-time nonlinear optimal ASRE feedback control systems provide very effective control, which is computationally simple to apply by using classical numerical techniques.

This paper presents an attack to the problem of dealing with partially observable systems subject to stochastic disturbance variables at the input and stochastic measurement errors at the output, where nonlinearities in the state and control input are explicitly taken into consideration. This is achieved through the theoretical development of an ASRE framework using classical linear stochastic optimal estimation and control theory. Thus this paper is a generalization of the deterministic ASRE framework in order to achieve similar results for a stochastic framework. For convenience, continuous-time models are treated to provide the basis of the appropriate optimal filter and optimal control design algorithms for general nonlinear dynamical systems, with a particular focus on nonlinear time-varying feedback design for optimal regulation in the presence of noise.

The paper is organized as follows. First, the notation adhered in the paper is presented in Section 2. A new technique is then introduced in Section 3 for designing an optimal filter for general nonlinear systems. The proposed method is the dual of ASRE control, and provides an explicit way of solving the nonlinear optimal filtering problem by using the well established theory of linear *Kalman-Bucy filtering*. This is achieved by transforming the nonlinear problem into a sequence of LTV approximations, which uniformly converge under very general and weak conditions of local Lipschitz continuity of the system dynamics. Optimal synthesis of general nonlinear (nonautonomous) stochastically disturbed nonaffine control systems is studied in Section 4. The stochastic *separation principle* of linear control theory is of central importance in formulating this problem. The solution is achieved by separating the stochastic nonlinear control problem into a problem of deterministic nonlinear control and a problem of stochastic nonlinear filtering, for which separation remains mathematically valid. The conclusions of this theoretical study are summarized in Section 5.

2. NOTATION

Let $\{\mathbf{x}(\omega,t); t\in[t_0,t_f]\}$ denote a stochastic (random) process whose state-space is an n-dimensional Euclidean space $\mathbb{R}^n$ (for $n\geq 1$) and whose index set t, referred to as "time", is an interval $[t_0,t_f]\in[0,\infty)$ of the real axis $\mathbb{R}^1$. Throughout the discussion to follow, the initial time t_0 and final time t_f will be kept *fixed*. As a rule, the argument ω of random vectors will be omitted in the paper. For brevity, the notation $\mathbf{x}_t$ will be written to denote the random state vector $\mathbf{x}(\omega,t)$ at any particular time $t\in[t_0,t_f]$, as is customary in probability theory. The conditional mean (that is, the unique unbiased estimate) of the probability distribution of the inaccessible n-dimensional state vector $\mathbf{x}_t$ is represented by $\hat{\mathbf{x}}_t = E\{\mathbf{x}_t\}$, which is a linear function of the measurement $\mathbf{y}_t\in\mathbb{R}^l$, $t_0\leq t\leq t_f$. The sequence using the proposed iterative LTV process for synthesizing optimal controls will be denoted by a superscript $[i]$ above the variable being iterated, where $i=0,1,2,\ldots$ so that the sequence is started with $i=0$.

3. NONLINEAR STOCHASTIC OPTIMAL ESTIMATION

The *Kalman filter* is an optimal estimator for the discrete-time LQG estimation problem, which is simple in form and powerful in effect. Analogous to the discrete-time case, *Kalman-Bucy filter* is the continuous-time equivalent of the Kalman filter (for details, see Anderson and Moore, 1979; Brammer and Siffling, 1989; Grewal and Andrews, 2001).

For the rest of the discussion in this section on nonlinear optimal filtering, suppose $\mathbf{u}_t=\mathbf{0}$ and thus consider a nonlinear stochastic dynamical system whose state $\mathbf{x}_t$ evolves in time according to

$$d\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t,t)dt + \mathbf{G}(\mathbf{x}_t,t)d\boldsymbol{\omega}_t,\quad E\{\mathbf{x}_{t_0}\} = \mathbf{x}_0 \quad (1)$$

for $t\geq t_0$. Suppose that Eq. (1) describes a physical system, the state $\mathbf{x}_t$ of which is not observable directly, but only through the nonlinear stochastic observation

$$d\mathbf{y}_t = \mathbf{g}(\mathbf{x}_t,t)dt + d\boldsymbol{\upsilon}_t,\quad \mathbf{y}_{t_0} = \mathbf{0}. \quad (2)$$

Here $\boldsymbol{\omega}_t$ and $\boldsymbol{\upsilon}_t$ are standard *Brownian motions* (*Wiener processes*) of appropriate dimensions on the interval $[t_0,t_f]$ whose respective formal time "derivatives" are $\mathbf{w}_t$ and $\mathbf{v}_t$. The plant and measurement noise models $\mathbf{w}_t$ and $\mathbf{v}_t$ are assumed zero-mean mutually correlated *Gaussian white noise* random processes (since this accords well with what occurs in practice) with known symmetric positive-semidefinite and positive-definite covariance matrices (strengths) $\tilde{\mathbf{Q}}(t)$ and $\tilde{\mathbf{R}}(t)$, respectively, and a known positive-semidefinite cross-covariance matrix $\tilde{\mathbf{S}}(t)$. The initial estimated value $\mathbf{x}_{t_0}$ is a biased Gaussian variate, with known mean $\mathbf{x}_0\in\mathbb{R}^n$ and known covariance matrix $\tilde{\mathbf{P}}(t_0)$, which is mutually uncorrelated with $\mathbf{w}_t$ and $\mathbf{v}_t$. In mathematical terms, these conditions become:

$$\left.\begin{aligned}
&E\{\mathbf{x}_{t_0}\} = \mathbf{x}_0\\
&E\left\{[\mathbf{x}_{t_0} - E\{\mathbf{x}_{t_0}\}][\mathbf{x}_{t_0} - E\{\mathbf{x}_{t_0}\}]^T\right\} = \tilde{\mathbf{P}}(t_0) = \tilde{\mathbf{P}}_0\\
&E\{\mathbf{w}_t\} = E\{\mathbf{v}_t\} = \mathbf{0}\\
&E\{\mathbf{w}_t\mathbf{w}_s{}^T\} = \tilde{\mathbf{Q}}(t)\delta(t-s)\\
&E\{\mathbf{v}_t\mathbf{v}_s{}^T\} = \tilde{\mathbf{R}}(t)\delta(t-s)\\
&E\{\mathbf{w}_t\mathbf{v}_s{}^T\} = \tilde{\mathbf{S}}(t)\delta(t-s)\\
&E\{\mathbf{x}_{t_0}\mathbf{w}_t{}^T\} = E\{\mathbf{x}_{t_0}\mathbf{v}_t{}^T\} = \mathbf{0}.
\end{aligned}\right\} \quad (3)$$

where $\delta(t)$ is the Dirac delta function. Note that the two random processes $\mathbf{w}_t$ and $\mathbf{v}_t$ are mutually *uncorrelated* if $\tilde{\mathbf{S}}(t)$ is zero.

The stochastic process and measurement models (1) and (2) represent a *linear* function of the additive disturbance $\mathbf{w}_t$ and measurement noise $\mathbf{v}_t$, respectively. On the other hand, $\mathbf{f}$, $\mathbf{g}$ and $\mathbf{G}$ are, in general, nonlinear functions of the state $\mathbf{x}_t$ (and/or control input $\mathbf{u}_t$) of the system. Applications involving nonlinear systems in this form generally require nonlinear optimal filters. The solution to the nonlinear filter problem was proposed by Kushner (1964) and is given by the *Kushner-Stratonovitch* stochastic PDE. This is a very awkward equation and its application to practical systems presents formidable numerical difficulties, which makes implementation very hard, if not impossible.

Since the general Kushner-Stratonovitch equation is so difficult to solve, approximation techniques have often been applied to nonlinear estimation problems to derive clearly suboptimal filters. More formal derivations of these nonlinear filters are given by Anderson and Moore (1979). However, common practice has been extensions of linear estimator methods for nonlinear problems, often using partial derivatives as linear approximations of nonlinear relations. The *linearized* and *extended Kalman filter* techniques are well-known and often used.

Now assuming that the origin $\mathbf{x}_t = \mathbf{0}$ is an isolated equilibrium point, that is $\mathbf{f}(\mathbf{0}, t) = \mathbf{0}$ and $\mathbf{g}(\mathbf{0}, t) = \mathbf{0}$, (1) and (2) can be represented in a (*nonunique*) factored state-space form

$$d\mathbf{x}_t = \mathbf{A}(\mathbf{x}_t, t)\mathbf{x}_t dt + \mathbf{G}(\mathbf{x}_t, t)d\boldsymbol{\omega}_t, \quad E\{\mathbf{x}_{t_0}\} = \mathbf{x}_0 \quad (4)$$

$$d\mathbf{y}_t = \mathbf{C}(\mathbf{x}_t, t)\mathbf{x}_t dt + d\boldsymbol{\upsilon}_t, \quad \mathbf{y}_{t_0} = \mathbf{0}. \quad (5)$$

where $\mathbf{A}$, $\mathbf{C}$ and $\mathbf{G}$ are continuous matrix-valued nonlinear functions. A sequence of LTV approximations can then be introduced to (4) and (5), which have the form (Banks and McCaffrey, 1998)

$$d\mathbf{x}_t^{[i]} = \mathbf{A}\left(\mathbf{x}_t^{[i-1]}, t\right)\mathbf{x}_t^{[i]}dt + \mathbf{G}\left(\mathbf{x}_t^{[i-1]}, t\right)d\boldsymbol{\omega}_t, E\{\mathbf{x}_{t_0}^{[i]}\} = \mathbf{x}_0 \quad (6)$$

$$d\mathbf{y}_t^{[i]} = \mathbf{C}\left(\mathbf{x}_t^{[i-1]}, t\right)\mathbf{x}_t^{[i]}dt + d\boldsymbol{\upsilon}_t, \quad \mathbf{y}_{t_0}^{[i]} = \mathbf{0} \quad (7)$$

for $i \geq 0$, where the iteration is initiated by assuming $\mathbf{x}_t^{[i-1]} = \mathbf{x}_0$ when $i = 0$. To prove convergence, an estimate for the bound on $\mathbf{\Phi}^{[i-1]}(t, t_0) - \mathbf{\Phi}^{[i-2]}(t, t_0)$ is required, where $\mathbf{\Phi}^{[i-1]}(t, t_0)$ and $\mathbf{\Phi}^{[i-2]}(t, t_0)$ are transition matrices generated by $\mathbf{A}(\mathbf{x}^{[i-1]}(t))$ and $\mathbf{A}(\mathbf{x}^{[i-2]}(t))$, respectively. This has been presented in Çimen and Banks (2004a), and is restated here as Lemma 2. First bear in mind the next Lemma.

Lemma 1 (Brauer, 1966, 1967)**.** The fundamental matrix $\mathbf{\Phi}^{[i-1]}(t, t_0)$ of the linear system

$$\dot{\mathbf{x}}^{[i]}(t) = \mathbf{A}\left(\mathbf{x}^{[i-1]}(t)\right)\mathbf{x}^{[i-1]}(t)$$

satisfies

$$\left\|\mathbf{\Phi}^{[i-1]}(t, t_0)\right\| \leq \exp\left[\int_{t_0}^{t} \mu\left(\mathbf{A}\left(\mathbf{x}^{[i-1]}(\tau)\right)\right) d\tau\right], \quad t \geq t_0$$

where the measure of the matrix $\mathbf{A}$, denoted by $\mu(\mathbf{A})$, is the logarithmic norm of $\mathbf{A}$ defined by

$$\mu(\mathbf{A}) \triangleq \lim_{h \to 0+} \left(\| \mathbf{I} + h\mathbf{A} \| - 1\right)/h.$$

Lemma 2 (Çimen and Banks, 2004a)**.** Suppose the following conditions are satisfied for finite $t \in [t_0, t_f]$:

(A1) $\mu(\mathbf{A}(\mathbf{x}, t)) \leq \mu_0$ for some finite constant μ_0 for all $\mathbf{x}$, and

(A2) Lipschitz continuity:

$$E\{\| \mathbf{A}(\mathbf{x}_1, t) - \mathbf{A}(\mathbf{x}_2, t) \|\} \leq \alpha E\{\| \mathbf{x}_1 - \mathbf{x}_2 \|\}, \forall \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$$

for some finite constant $\alpha > 0$.

Then

$$E\left\{\left\|\mathbf{\Phi}^{[i-1]}(t, t_0) - \mathbf{\Phi}^{[i-2]}(t, t_0)\right\|\right\} \leq L_1 \sup_{s \in [t_0, t]} E\left\{\left\|\mathbf{x}_s^{[i-1]} - \mathbf{x}_s^{[i-2]}\right\|\right\}$$

where $L_1 \triangleq \alpha(t - t_0)\exp[\mu_0(t - t_0)]$.

Now, using the variation of constants formula, on integrating (6) over $[t_0, t_f]$ the solution becomes

$$\mathbf{x}_t^{[i]} = \mathbf{\Phi}^{[i-1]}(t, t_0)\mathbf{x}_0 + \int_{t_0}^{t} \mathbf{\Phi}^{[i-1]}(t, s)\mathbf{G}\left(\mathbf{x}_s^{[i-1]}, s\right) d\boldsymbol{\omega}_s \quad (8)$$

and so $\mathbf{x}_t^{[i]} - \mathbf{x}_t^{[i-1]}$ can be written as

$$\begin{aligned} \mathbf{x}_t^{[i]} - \mathbf{x}_t^{[i-1]} &= \left[\mathbf{\Phi}^{[i-1]}(t, t_0) - \mathbf{\Phi}^{[i-2]}(t, t_0)\right]\mathbf{x}_0 \\ &+ \int_{t_0}^{t} \mathbf{\Phi}^{[i-1]}(t, s)\left[\mathbf{G}\left(\mathbf{x}_s^{[i-1]}, s\right) - \mathbf{G}\left(\mathbf{x}_s^{[i-2]}, s\right)\right] d\boldsymbol{\omega}_s \\ &+ \int_{t_0}^{t} \left[\mathbf{\Phi}^{[i-1]}(t, s) - \mathbf{\Phi}^{[i-2]}(t, s)\right]\mathbf{G}\left(\mathbf{x}_s^{[i-2]}, s\right) d\boldsymbol{\omega}_s. \end{aligned}$$

Under conditions (A1), (A2), and

(A3) $E\{\| \mathbf{x}_0 \|\} = c$,

(A4) $E\{\| \mathbf{G}(\mathbf{x}, t) \|\} \leq g_1$, $\forall \mathbf{x} \in \mathbb{R}^n$,

(A5) $E\{\| \mathbf{G}(\mathbf{x}_1, t) - \mathbf{G}(\mathbf{x}_2, t) \|\} \leq g_2 E\{\| \mathbf{x}_1 - \mathbf{x}_2 \|\}$, $\forall \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$,

for finite numbers $c \geq 0$ and $g_1, g_2 > 0$ with $t \in [t_0, t_f]$, using Lemma 1

$$\begin{aligned} E\left\{\left\|\mathbf{x}_t^{[i]} - \mathbf{x}_t^{[i-1]}\right\|\right\} &\leq E\left\{\left\|\mathbf{\Phi}^{[i-1]}(t, t_0) - \mathbf{\Phi}^{[i-2]}(t, t_0)\right\|\left\|\mathbf{x}_0\right\|\right. \\ &+ \int_{t_0}^{t} \left\|\mathbf{\Phi}^{[i-1]}(t, s)\right\|\left\|\mathbf{G}\left(\mathbf{x}_s^{[i-1]}, s\right) - \mathbf{G}\left(\mathbf{x}_s^{[i-2]}, s\right)\right\| d\boldsymbol{\omega}_s \\ &\left.+ \int_{t_0}^{t} \left\|\mathbf{\Phi}^{[i-1]}(t, s) - \mathbf{\Phi}^{[i-2]}(t, s)\right\|\left\|\mathbf{G}\left(\mathbf{x}_s^{[i-2]}, s\right)\right\| d\boldsymbol{\omega}_s\right\} \\ &\leq L_1 \sup_{s \in [t_0, t]} E\left\{\left\|\mathbf{x}_s^{[i-1]} - \mathbf{x}_s^{[i-2]}\right\|\right\} c \\ &+ \int_{t_0}^{t} \exp\left[\mu_0(t - t_0)\right] g_2 \left\|\mathbf{x}_s^{[i-1]} - \mathbf{x}_s^{[i-2]}\right\| d\boldsymbol{\omega}_s \\ &+ \int_{t_0}^{t} L_1 \sup_{s \in [t_0, t]} E\left\{\left\|\mathbf{x}_s^{[i-1]} - \mathbf{x}_s^{[i-2]}\right\|\right\} g_1 \, d\boldsymbol{\omega}_s. \end{aligned}$$

Suppose that $\xi_t^{[i]} \triangleq \sup_{s \in [t_0, t]} E\{\| \mathbf{x}_s^{[i]} - \mathbf{x}_s^{[i-1]} \|\}$. Then

$$\xi_t^{[i]} \leq L_2 \xi_t^{[i-1]} \quad (9)$$

for $t \in [t_0, t_f]$ where

$$\begin{aligned} L_2 &\triangleq L_1 c + \left\{\exp\left[\mu_0(t - t_0)\right] g_2 + L_1 g_1\right\} \int_{t_0}^{t} d\boldsymbol{\omega}_s \\ &= L_1 c + \left\{\exp\left[\mu_0(t - t_0)\right] g_2 + L_1 g_1\right\} (\boldsymbol{\omega}_t - \boldsymbol{\omega}_{t_0}) \end{aligned}$$

since every approximation of the integral $\int_{t_0}^{t} d\boldsymbol{\omega}_s$ by means of Riemann-Stieltjes sums of the form $S_n = \sum_{j=1}^{n} (\boldsymbol{\omega}_{t_j} - \boldsymbol{\omega}_{t_{j-1}})$, $t_0 \leq t_1 \leq \ldots \leq t_n = t$, leads to this value (Arnold, 1974). From (9), by induction, $\xi_t^{[i]}$ satisfies

$$\xi_t^{[i]} \leq L_2^{\,i-1} \xi_t^{[1]}. \quad (10)$$

Theorem 1. Under conditions (A1)-(A5) and provided that $| L_2 | < 1$, (4) has a unique solution on $[t_0, t_f]$, given by the limit of the solutions of the approximating equations (6) on $C([t_0, t_f]; \mathbb{R}^n)$.

Proof. The proof follows directly from (10) since this implies that $\mathbf{x}_t^{[i]}$ is a Cauchy sequence in the Banach space $C([t_0, t_f]; \mathbb{R}^n)$. The value of $\xi_t^{[i]}$ is calculated from (8). Using Lemma 1 and assumptions (A1)-(A5),

$$E\left\{\left\|\mathbf{x}_t^{[i]}\right\|\right\} \le e^{[\mu_0(t-t_0)]c} + E\left\{\int_{t_0}^{t} e^{[\mu_0(t-t_0)]} g_1 \, d\boldsymbol{\omega}_s\right\} = L_3$$

where $L_3 \triangleq \exp[\mu_0(t-t_0)]\{c + g_1(\boldsymbol{\omega}_t - \boldsymbol{\omega}_{t_0})\}$. Hence

$$\xi_t^{[1]} = \sup_{s\in[t_0,\,t]} E\left\{\left\|\mathbf{x}_s^{[1]} - \mathbf{x}_s^{[0]}\right\|\right\} \le E\left\{\left\|\mathbf{x}_s^{[1]}\right\|\right\} + E\left\{\left\|\mathbf{x}_s^{[0]}\right\|\right\} \le 2L_3 .$$

and thus, from (10), $\xi_t^{[i]} \le 2L_2^{\,i-1}L_3$. Therefore, if $|L_2|<1$ for $t\in[t_0,t_f]$, it follows that

$$\lim_{i\to\infty} E\left\{\left\|\mathbf{x}_t^{[i]} - \mathbf{x}_t^{[i-1]}\right\|\right\} = 0$$

and almost certainly $\lim_{i\to\infty} \mathbf{x}_t^{[i]} = \mathbf{x}_t$. □

Using similar ideas as above, it can be shown that $\lim_{i\to\infty} E\{\| \mathbf{y}_t^{[i]} - \mathbf{y}_t^{[i-1]} \|\} = 0$, that is, the unique solution of (5) is given by the limit of the approximating sequence (7) on $C([t_0,t_f];\mathbb{R}^l)$. Thus, using results from Çimen and Banks (2004a), the sequence of approximations (6) and (7) globally converge (uniformly in time) as $i\to\infty$. The general (nonautonomous) nonlinear system (4) with nonlinear stochastic observations (5) can therefore be represented by the sequence of approximations (6) and (7), respectively. Since each approximating problem in (6) and (7) is now linear, time-varying (with the exception of the first sequence), quadratic and Gaussian, the Kalman-Bucy filter algorithm for the LQG estimation problem can be applied to find an estimate $\hat{\mathbf{x}}_t$ of $\mathbf{x}_t$. The optimal ASRE filter for the nonlinear system is hence given by

Theorem 2 (Nonlinear optimal ASRE filter)**.** Suppose the functions $\mathbf{f}$, $\mathbf{g}$ and $\mathbf{G}$ are locally Lipschitz with respect to their arguments on $t\in[t_0,t_f]$. Provided that the origin is an equilibrium point, (1) and (2) can be written in factored form (4) and (5), respectively. The ASRE minimum variance estimate on the interval $[t_0,t_f]$ for the state $\mathbf{x}_t$ of the continuous-time stochastically disturbed nonlinear system (1), (2) with conditions (3) is then given by the stochastic differential equation

$$d\hat{\mathbf{x}}_t^{[i]} = \mathbf{A}\left(\hat{\mathbf{x}}_t^{[i-1]},t\right)\hat{\mathbf{x}}_t^{[i]}dt + \left[\tilde{\mathbf{P}}^{[i]}(t)\mathbf{C}^T\left(\hat{\mathbf{x}}_t^{[i-1]},t\right) + \tilde{\mathbf{S}}(t)\right] \times\tilde{\mathbf{R}}^{-1}(t)\left[d\mathbf{y}_t - \mathbf{C}\left(\hat{\mathbf{x}}_t^{[i-1]},t\right)\hat{\mathbf{x}}_t^{[i]}dt\right], \tag{11}$$

with $\hat{\mathbf{x}}_{t_0}^{[i]} = \mathbf{x}_0$. The covariance matrix $\tilde{\mathbf{P}}^{[i]}(t) = E\{\tilde{\mathbf{x}}_t^{[i]}\tilde{\mathbf{x}}_t^{[i]T}\}$ of the estimation error $\tilde{\mathbf{x}}_t^{[i]} = \mathbf{x}_t^{[i]} - \hat{\mathbf{x}}_t^{[i]}$ satisfies the ASRE

$$\begin{aligned}\dot{\tilde{\mathbf{P}}}^{[i]}(t) &= \mathbf{G}\left(\hat{\mathbf{x}}_t^{[i-1]},t\right)\tilde{\mathbf{Q}}(t)\mathbf{G}^T\left(\hat{\mathbf{x}}_t^{[i-1]},t\right) \\ &\quad + \mathbf{A}\left(\hat{\mathbf{x}}_t^{[i-1]},t\right)\tilde{\mathbf{P}}^{[i]}(t) + \tilde{\mathbf{P}}^{[i]}(t)\mathbf{A}^T\left(\hat{\mathbf{x}}_t^{[i-1]},t\right) \\ &\quad - \tilde{\mathbf{P}}^{[i]}(t)\mathbf{C}^T\left(\hat{\mathbf{x}}_t^{[i-1]},t\right)\tilde{\mathbf{R}}^{-1}(t)\mathbf{C}\left(\hat{\mathbf{x}}_t^{[i-1]},t\right)\tilde{\mathbf{P}}^{[i]}(t)\end{aligned} \tag{12}$$

for $i\ge 0$, where $\tilde{\mathbf{P}}^{[i]}(t_0) = \tilde{\mathbf{P}}_0$ (that is, the covariance matrix of $\mathbf{x}_{t_0}$) is given. For $i=0$, $\hat{\mathbf{x}}_t^{[i-1]}$ in (11) and (12) is set to $\mathbf{x}_0$.

Proof. By using methods similar to those presented in Çimen and Banks (2004a), the ASRE (12) can be shown to converge uniformly in time to solutions $\tilde{\mathbf{P}}(t)$ for $t\in[t_0,t_f]$. From Theorem 1, therefore, it follows that the approximating sequence (11) for $\hat{\mathbf{x}}_t^{[i]}$ will also converge globally in time to an estimate $\hat{\mathbf{x}}_t$ of $\mathbf{x}_t$ on $C([t_0,t_f];\mathbb{R}^n)$. □

Remark 1. ASRE theory in a deterministic setting (Çimen and Banks, 2004a, b) utilizes the actual state $\mathbf{x}_t$ in the sequence of approximations, such as $\mathbf{A}(\mathbf{x}^{[i-1]}(t),t)$. The same notion is used to transform nonlinear equations (4) and (5) into LTV ones (6) and (7). However, the reader should be aware of the use of the estimated state $\hat{\mathbf{x}}_t$ in the approximating sequences of Theorem 2. Therefore, $\mathbf{A}(\hat{\mathbf{x}}^{[i-1]}(t),t)$ must now be adopted for the ASRE filter problem.

Remark 2. The propagation over time of the probability distribution of the state of a nonlinear dynamic system is described by the Fokker-Planck nonlinear PDE. Certainly, the filter has to be infinite dimensional as a solution of the general Fokker-Planck equation. Theorem 2 provides a finite-dimensional approximation to this solution. Recall from conditions (3) that the input to the actual system is assumed Gaussian. Therefore, the output distribution will, in effect, be nonGaussian as a result of nonlinear dynamics. However, the ASRE filter, which admits Gaussian noise inputs, will clearly converge to solutions with Gaussian output distributions, due to the linear (time-varying) nature of the sequence of subproblems given by Theorem 2. This implies that the sequence of LTV systems of the ASRE filter will not, therefore, converge to the nonGaussian probability distribution corresponding to the nonlinear filtering density.

4. NONLINEAR STOCHASTIC OPTIMAL CONTROL

For linear systems, the solution to the stochastic optimal control problem can be achieved by separating the control system, the quadratic performance index, linear control law, Gaussian disturbance and measurement noise into a problem of *deterministic linear optimal control* and a problem of *stochastic linear optimal filtering* for which both topics have a fully developed theory. A controller and a filter can hence be designed completely separately from each other. The resulting optimal system is given by the closed-loop system consisting of the process, filter and controller, in which the optimal estimate $\hat{\mathbf{x}}_t$ of the state vector $\mathbf{x}_t$ is fed back to the controller. The separation theorem thus provides a method of designing a Kalman-Bucy filter while ignoring the control problem, and designing an LQG control strategy as though the system states were available, but in fact using the estimated states in the knowledge that the strategy is optimal in the same sense. Therefore, if the process noise corruption of a linear stochastic system is Gaussian, classical optimal regulator and tracking design strategies apply almost directly to the LQG control strategy, with appropriate changes in the cost function definitions. However, the separation principle does not hold in the strict sense for nonlinear systems or for nonGaussian noise. Nevertheless, in the synthesis of nonlinear stochastic control systems, there has been seldom any other choice available than to proceed as if separation were still valid. Even though this has at least been feasible for some nonlinear stochastic control problems, it is not a completely satisfactory method mathematically, with no guarantee for absolute optimization. A new method is now presented, which is mathematically valid for optimal control of partially observable nonlinear stochastically disturbed systems.

In Çimen and Banks (2004a), the ASRE theory has been developed to derive explicit solutions to the deterministic nonlinear optimal tracking problem for input-affine nonlinear systems. These solutions can be generalized to include nonlinearities in the control input $\mathbf{u}_t$, as in Çimen and Banks (2004b). In order to simplify mathematics, the formulation for nonlinear stochastic optimal control in the following discussion will be presented for the state-regulation problem, which can easily be extended to the more general tracking framework as in Çimen and Banks (2004a). Thus, consider a nonlinear stochastically disturbed nonaffine control system governed by the equation

$$d\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t, t)dt + \mathbf{G}(\mathbf{x}_t, \mathbf{u}_t, t)d\boldsymbol{\omega}_t,\ E\{\mathbf{x}_{t_0}\} = \mathbf{x}_0 \quad (13)$$

together with nonlinear partial-state information

$$d\mathbf{y}_t = \mathbf{g}(\mathbf{x}_t, \mathbf{u}_t, t)dt + d\boldsymbol{\upsilon}_t, \quad \mathbf{y}_{t_0} = \mathbf{0} \quad (14)$$

where the control input $\mathbf{u} \in \mathbb{R}^m$ is unconstrained. The minimum nonlinear control problem on a fixed finite-time interval $t_0 \le t \le t_f$ is given by a feedback control $\mathbf{u}_t$ which minimizes a Bolza-type cost criterion, represented here by a finite-time nonlinear (nonquadratic) and time-varying cost functional

$$J(\mathbf{u}) = \frac{1}{2}E\Big\{\mathbf{x}_{t_f}{}^T\mathbf{F}(\mathbf{x}_{t_f})\mathbf{x}_{t_f} + \int_{t_0}^{t_f}\left[\mathbf{x}_t{}^T\mathbf{Q}(\mathbf{x}_t, t)\mathbf{x}_t + \mathbf{u}_t{}^T\mathbf{R}(\mathbf{x}_t, t)\mathbf{u}_t\right]dt\Big\}. \quad (15)$$

Analogous to the LQR problem, $\mathbf{F}$ and $\mathbf{Q}$ are positive-semidefinite, $\mathbf{R}$ is positive-definite, and there are no control constraints associated with the nonlinear stochastic optimal regulator control problem. Thus, provided that $\mathbf{f}(\mathbf{0}, \mathbf{0}, t) = \mathbf{0}$ and $\mathbf{g}(\mathbf{0}, \mathbf{0}, t) = \mathbf{0}$, (13) and (14) can be represented in factored form

$$\begin{aligned} d\mathbf{x}_t &= \mathbf{A}(\mathbf{x}_t, \mathbf{u}_t, t)\mathbf{x}_t dt + \mathbf{B}(\mathbf{x}_t, \mathbf{u}_t, t)\mathbf{u}_t dt \\ &\quad + \mathbf{G}(\mathbf{x}_t, \mathbf{u}_t, t)d\boldsymbol{\omega}_t, \qquad E\{\mathbf{x}_{t_0}\} = \mathbf{x}_0 \end{aligned} \quad (16)$$

and

$$d\mathbf{y}_t = \mathbf{C}(\mathbf{x}_t, \mathbf{u}_t, t)\mathbf{x}_t dt + d\boldsymbol{\upsilon}_t, \quad \mathbf{y}_{t_0} = \mathbf{0}, \quad (17)$$

respectively, where the matrices $\mathbf{A}(\mathbf{x}, \mathbf{u}, t)$, $\mathbf{B}(\mathbf{x}, \mathbf{u}, t)$ $\mathbf{C}(\mathbf{x}, \mathbf{u}, t)$ and $\mathbf{G}(\mathbf{x}, \mathbf{u}, t)$ are jointly continuous in their arguments, as well as in the time variable t. Then, using methods similar to those presented in Section 3 and Çimen and Banks (2004a, b), Eqs. (15)-(17) can be replaced by the sequence of approximations

$$\begin{aligned} J^{[i]}(\mathbf{u}) = \frac{1}{2}E\Big\{&\mathbf{x}_{t_f}^{[i]T}\mathbf{F}\left(\mathbf{x}_{t_f}^{[i-1]}\right)\mathbf{x}_{t_f}^{[i]} + \int_{t_0}^{t_f}\Big[\mathbf{x}_t^{[i]T}\mathbf{Q}\left(\mathbf{x}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)\mathbf{x}_t^{[i]} \\ &+ \mathbf{u}_t^{[i]T}\mathbf{R}\left(\mathbf{x}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)\mathbf{u}_t^{[i]}\Big]dt\Big\}, \end{aligned} \quad (18)$$

$$\begin{aligned} d\mathbf{x}_t^{[i]} &= \mathbf{A}\left(\mathbf{x}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)\mathbf{x}_t^{[i]}dt + \mathbf{B}\left(\mathbf{x}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)\mathbf{u}_t^{[i]}dt \\ &\quad + \mathbf{G}\left(\mathbf{x}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)d\boldsymbol{\omega}_t, \\ E\left\{\mathbf{x}_{t_0}^{[i]}\right\} &= \mathbf{x}_0, \end{aligned} \quad (19)$$

$$d\mathbf{y}_t^{[i]} = \mathbf{C}\left(\mathbf{x}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)\mathbf{x}_t^{[i]}dt + d\boldsymbol{\upsilon}_t,\ \mathbf{y}_{t_0}^{[i]} = \mathbf{0} \quad (20)$$

for $i \ge 0$. The iteration is initiated at $i = 0$ with $\mathbf{x}_t^{[i-1]} = \mathbf{x}_0$ and $\mathbf{u}_t^{[i-1]} = \mathbf{0}$. The sequence of approximations (18)-(20) then converge globally in time as $i \to \infty$ (Çimen and Banks, 2004a, b).

Since each approximating problem (18)-(20) is LTV, quadratic and Gaussian, the stochastic separation principle is valid for each sequence of problems here. Hence the well-known results of LQG theory for stochastic optimal control of partially observable linear systems can now be applied to (18)-(20), thus extending the solution to the general case of continuous, nonlinear nonautonomous stochastic optimal regulation problem (13) and (14) with (15). The problem with partial-state information of the nonlinear system will be solved using the results of Section 3 for ASRE filtering. Therefore, the solution to the stochastic optimal control problem for partially observable nonlinear systems may be given by

Theorem 3 (ASRE separation theorem)**.** Consider the general (nonautonomous) nonlinear stochastic process (13) and measurement (14) together with conditions (3) and the nonlinear (nonquadratic and time-dependent) performance index (15) where $\mathbf{x}_t$, $\mathbf{u}_t$, $\mathbf{y}_t$, $\boldsymbol{\omega}_t$ and $\boldsymbol{\upsilon}_t$ assume values in Euclidean spaces of arbitrary dimensions, $\mathbf{Q}(\mathbf{x}, t)$, $\mathbf{F}(\mathbf{x})$ are symmetric positive-semidefinite, $\mathbf{R}(\mathbf{x}, t)$ is symmetric positive-definite, $\mathbf{x}_{t_0}$ is a normally distributed random vector with expectation $\mathbf{x}_0$, and $\mathbf{w}_t = \dot{\boldsymbol{\omega}}_t$ and $\mathbf{v}_t = \dot{\boldsymbol{\upsilon}}_t$ are vector-valued Gaussian white random processes. Suppose the origin is an equilibrium point so that $\mathbf{f}(\mathbf{0}, \mathbf{0}, t) = \mathbf{0}$ and $\mathbf{g}(\mathbf{0}, \mathbf{0}, t) = \mathbf{0}$. Then (13) and (14) can be written in a factored form (16) and (17), which, together with (15), can be replaced by the LTV systems (18)-(20). Hence, control and filtering can be separated from each other by analogy with the separation principle of linear stochastic control theory. The ASRE solution of the continuous-time nonlinear stochastic control problem is thus achieved by

(i) The deterministic optimal control law on the interval $[t_0, t_f]$, given by the limit of the approximating controls

$$\mathbf{u}_t^{[i]} = -\mathbf{R}^{-1}\left(\hat{\mathbf{x}}_t^{[i-1]}, t\right)\mathbf{B}^T\left(\hat{\mathbf{x}}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)\mathbf{P}^{[i]}(t)\hat{\mathbf{x}}_t^{[i]}$$

where $\mathbf{P}^{[i]}(t)$ is the solution of the backward-integrated ASRE

$$\dot{\mathbf{P}}^{[i]}(t) = -\mathbf{Q}\left(\hat{\mathbf{x}}_t^{[i-1]}, t\right) - \mathbf{P}^{[i]}(t)\mathbf{A}\left(\hat{\mathbf{x}}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)$$
$$-\mathbf{A}^T\left(\hat{\mathbf{x}}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)\mathbf{P}^{[i]}(t) + \mathbf{P}^{[i]}(t)\mathbf{S}\left(\hat{\mathbf{x}}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)\mathbf{P}^{[i]}(t),$$
$$\mathbf{P}^{[i]}(t_f) = \mathbf{F}\left(\hat{\mathbf{x}}_{t_f}^{[i-1]}\right).$$

with

$$\mathbf{S} \triangleq \mathbf{B}\left(\hat{\mathbf{x}}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)\mathbf{R}^{-1}\left(\hat{\mathbf{x}}_t^{[i-1]}, t\right)\mathbf{B}^T\left(\hat{\mathbf{x}}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right).$$

(ii) The stochastic optimal filter law

$$d\hat{\mathbf{x}}_t^{[i]} = \mathbf{A}\left(\hat{\mathbf{x}}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)\hat{\mathbf{x}}_t^{[i]}dt + \mathbf{B}\left(\hat{\mathbf{x}}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)\mathbf{u}_t^{[i]}dt$$
$$+\left[\tilde{\mathbf{P}}^{[i]}(t)\mathbf{C}^T\left(\hat{\mathbf{x}}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right) + \tilde{\mathbf{S}}(t)\right]\tilde{\mathbf{R}}^{-1}(t)$$
$$\times\left[d\mathbf{y}_t - \mathbf{C}\left(\hat{\mathbf{x}}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)\hat{\mathbf{x}}_t^{[i]}dt\right], \qquad \hat{\mathbf{x}}_{t_0}^{[i]} = \mathbf{x}_0$$

on the interval $[t_0, t_f]$, where the covariance matrix $\tilde{\mathbf{P}}^{[i]}(t)$ of the estimation error obeys the forward integrated ASRE

$$\dot{\tilde{\mathbf{P}}}^{[i]}(t) = \mathbf{G}\left(\hat{\mathbf{x}}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)\tilde{\mathbf{Q}}(t)\mathbf{G}^T\left(\hat{\mathbf{x}}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)$$
$$+\mathbf{A}\left(\hat{\mathbf{x}}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)\tilde{\mathbf{P}}^{[i]}(t) + \tilde{\mathbf{P}}^{[i]}(t)\mathbf{A}^T\left(\hat{\mathbf{x}}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)$$
$$-\tilde{\mathbf{P}}^{[i]}(t)\mathbf{C}^T\left(\hat{\mathbf{x}}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)\tilde{\mathbf{R}}^{-1}(t)\mathbf{C}\left(\hat{\mathbf{x}}_t^{[i-1]}, \mathbf{u}_t^{[i-1]}, t\right)\tilde{\mathbf{P}}^{[i]}(t),$$
$$\tilde{\mathbf{P}}^{[i]}(t_0) = \tilde{\mathbf{P}}_0.$$

Proof. The proof follows directly from Section 3 and the deterministic theory presented in Çimen and Banks (2004a, b). □

Remark 3. Similar to the sequence of approximations (11) and (12), the initial function $\hat{\mathbf{x}}_t^{[i-1]}$ (and $\hat{\mathbf{x}}_{t_f}^{[i-1]}$) when $i = 0$ is fixed at $\mathbf{x}_0$ for $t \in [t_0, t_f]$ in the ASRE control and filter equations of Theorem 3. In addition, the approximations for the nonlinear stochastic optimal control problem require an initial guess for $\mathbf{u}_t^{[i-1]}$ when $i = 0$, which is assumed $\mathbf{0}$.

Remark 4. For each LTV subproblem, the operations (i) and (ii) in Theorem 2 are independent in the sense that the ASRE filter does not depend in any way on the matrices $\mathbf{F}$, $\mathbf{Q}$, $\mathbf{R}$ defining the control problem, whereas the control function $\mathbf{u}_t$ does not depend on the "noise parameters" $\mathbf{G}$, $\tilde{\mathbf{P}}$, $\tilde{\mathbf{Q}}$, $\tilde{\mathbf{R}}$, $\tilde{\mathbf{S}}$. Thus, "separation" holds in each sequence of LTV approximations. This does not, of course, yield a global separation theorem for the nonlinear system. Theorem 3 only gives separation on each individual solution of the nonlinear system, and thus provides a method of approximating the separation principle to each solution trajectory.

5. CONCLUSIONS

A new theory has been proposed for finite-time optimal regulation of general nonautonomous nonlinear stochastic systems with nonquadratic performance criteria, where the origin of state-space is an equilibrium point of the nonlinear system. The nonlinear stochastic control problem has been solved using modern linear control design methodologies by replacing the general nonlinear equations representing the plant with a convergent sequence of LTV equations. The separation principle is valid here since the nonlinear problem is replaced by a sequence of LQG and time-varying approximations that converge under very mild conditions of local Lipschitz continuity. A new theory of separation principle has therefore been derived, which is valid for the optimal control of nonlinear stochastic systems. While this does not provide a global separation theorem for the nonlinear problem, the method is straightforward and practically realizable, and is achieved *without* having to use the mathematical tools of Itô stochastic integrals for continuous-time models. The simplicity of the ASRE approach is appealing compared with other methods that solve complex PDEs.

REFERENCES

Anderson, B.D.O. and J.B. Moore (1979). *Optimal Filtering*. Prentice-Hall, New Jersey.

Arnold, L. (1974). *Stochastic Differential Equations: Theory and Applications*. Wiley-Interscience, New York.

Banks, S.P. and D. McCaffrey (1998). Lie algebras, structure of nonlinear systems and chaotic motion, *International Journal of Bifurcation and Chaos*, **8** (7), 1437-1462.

Brammer, K. and G. Siffling, (1989). *Kalman-Bucy Filters*. Artech House, Norwood (MA).

Brauer, F. (1966). Perturbations of nonlinear systems of differential equations. *Journal of Mathematical Analysis and Applications*, **14**, 198-206.

Brauer, F. (1967). Perturbations of nonlinear systems of differential equations, II. *Journal of Mathematical Analysis and Applications*, **17**, 418-434.

Çimen, T. and S.P. Banks (2004a). Nonlinear optimal tracking control with application to super-tankers for autopilot design, *Automatica*, **40** (11), 1845-1863.

Çimen, T. and S.P. Banks (2004b). Global optimal feedback control for general nonlinear systems with nonquadratic performance criteria, *Systems and Control Letters*, **53** (5), 327-346.

Grewal, M.S. and A.P. Andrews (2001). *Kalman Filtering: Theory and Practice Using MATLAB*, Second Edition. Wiley-Interscience, New York.

Kushner, H.J. (1964). On the differential equations satisfied by conditional probability densities of Markov processes, with applications. *SIAM Journal of Control*, Ser. A, **2,** 106-119.