CONTROL–THEORY AND ADVANCED TECHNOLOGY
Vol. 10, No. 4, Part 1, pp. 847–871, December, 1994

Contributed Paper

SUBOPTIMAL CONTROL OF NONLINEAR STOCHASTIC SYSTEMS*

G. N. SARIDIS[1] AND F.-Y. WANG[2]

Abstract. Theoretical procedures are developed for comparing the performance of arbitrarily selected admissible feedback controls among themselves with the optimal solution of a nonlinear optimal stochastic control problem. Iterative design schemes are proposed for successively improving the performance of a controller until a satisfactory design is achieved. Specifically, the exact design procedure is based on the generalized Hamilton-Jacobi-Bellman equation of the cost function of nonlinear stochastic systems, and the approximate design procedure for the infinite-time nonlinear stochastic regulator problem, is developed by using the upper and lower bounds of the cost functions. Stability of this problem is also considered. For a given controller, both the upper and lower bounds to its cost function can be obtained by solving a partial differential inequality. These bounds, constructed without actually knowing the optimal controller, are used as measure to evaluate the acceptability of suboptimal controllers. These results establish an approximation theory of optimal stochastic control and provide a practical procedure for selecting effective practical controls for nonlinear stochastic systems. An Entropy reformulation of the Generalized Hamilton-Jacobi-Bellman equation is also presented.

Key Words—Control systems synthesis, entropy, Hamilton-Jacobi-Bellman equation, infinite-time control, nonlinear stochastic systems, recursive design procedure, suboptimal stochastic control, stability, stochastic regulator problem.

1. Introduction

The problem of controlling a stochastic dynamic system, such that its behavior is optimal with respect to a performance cost, has received considerable attention over the past two decades. From a theoretical as well as practical point of view, it is desirable to obtain a feedback solution to the optimal control problem. In situations of linear stochastic systems with additive white Gaussian noise and quadratic performance indices (so-called LQG problems), the separation theorem is directly applicable, and the optimal control theory is well established (Aoki, 1967; Wonham, 1970; Kwakernaak and Sivan, 1972; Sage and White, 1977).

However, due to difficulties associated with the mathematics of stochastic processes, only fragmentary results are available for the design of optimal control of nonlinear stochastic systems. On the other hand, there is a need to design optimal and suboptimal controls for practical implementation in engineering applications (Panossian, 1988).

The objective of this paper is to develop an approximation theory that may be used to find feasible, practical solutions to the optimal control of nonlinear

* Received by the editors November 8, 1993 and in revised form April 22, 1994.
[1] Electrical, Computer, and Systems Engineering Department, School of Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180–3590, U.S.A.
[2] Department of Systems and Industrial Engineering, University of Arizona, Tucson, AZ 85721, U.S.A.

stochastic systems. To this end, the problem of stochastic control is addressed from an inverse point of view:

Given an arbitrary selected admissible feedback control, it is desirable to compare it to other feedback controls, with respect to a given performance cost, and to successively improve its design to converge to the optimal.

Various direct approximations of the optimal control have been widely studied for nonlinear deterministic systems (Rekasius, 1964; Leake and Liu, 1967; Saridis and Lee, 1979; Saridis and Balaram, 1986), and appeared to be more promising than the linearization type approximation methods that have met with limited success (Al'brekht, 1961; Lukes, 1969; Nishikawa et al., 1962). For stochastic systems, a method of successive approximation to solve the Hamilton-Jacobi-Bellman equation for a stochastic optimal control problem using quasilinearization, was proposed in Ohsumi (1984), but systematic procedures for the construction of suboptimal controllers were not established.

This paper presents a theoretical procedure to develop suboptimal feedback controllers for stochastic nonlinear systems (Wang and Saridis, 1992), as an extension of the Approximation Theory of Optimal Control developed by Saridis and Lee (1979) for deterministic nonlinear systems. The results are organized as follows. Section 2 gives the mathematical preliminaries of the stochastic optimal control problem. Section 3 describes major theorems that can be used for the construction of successively improved controllers. For the infinite-time stochastic regulator problem, a design theory using upper and lower bounds of the cost function is given in Sec. 4. Stability considerations of this problem are discussed in Sec. 5. Two proposed design procedures are outlined in Sec. 6, and illustrated with several examples. A reformulation of the problem using Entropy as a basic concept is given in Sec. 7. Conclusions summarize the paper in Sec. 8.

2. Problem Formulation

For the purpose of obtaining explicit expressions, and without loss of generality since the results are immediately generalizable, consider a nonlinear stochastic control system described by the following stochastic differential equation:

$$dx = f(t, x)dt + b(t, x)udt + g(t, x)dw, \quad t \in I \equiv [t_0, T], \tag{1}$$

where $x \in R^n$ is a vector of state of the stochastic system, $u \in \Omega_u \subset R^m$ is a control vector, Ω_u is a specified compact set of admissible controls, and $w \in R^k$ is a separable Wiener process. $f: I \times R^n \to R^n$, $b: I \times R^n \to R^{n \times m}$ and $g: I \times R^n \to R^{n \times k}$ are measurable system functions. Equation (1) was studied first by Itô (1951), and later, under less restrictive conditions, by Doob (1953), Dynkin (1953), Skorokhod (1965) and Kushner (1971). It is assumed that the feedback control law $u(t, x) \in \Omega_u$ satisfies the following conditions:

i) Linear Growth Condition,

$$\|f(t, x) + b(t, x)u(t, x)\| + \|g(t, x)\| \leq a(1 + \|x\|).$$

ii) Uniform Lipschitz Condition,

$$\|(f-bu)(t, x)-(f-bu)(t, y)\| + \|g(t, x)-g(t, y)\| \leq a\|x-y\|,$$

where (t, x), $(t, y) \in I \times R^n$, $\|\cdot\|$ is Euclidian norm operator, and a is some constant.

For a given *initial state* $x(t_0)=x_0$ (deterministic) and feedback control $u(t, x)$, the performance cost of the system (1) is defined as

$$J(u;\, t_0,\, x_0) = E\left\{ \int_{t_0}^{T} [L(t, x)+\|u\|^2]dt+\phi(T, x(T))/x(t_0)=x_0 \right\} \tag{2}$$

with nonnegative functions $L\colon I\times R^n \to R^1$ and $\phi\colon I\times R^n \to R^1$. J is also called the *Cost Function* of system (1).

The infinitesimal generator of the stochastic process specified by (1) is defined to be,

$$\mathscr{L}_u\Psi \equiv \frac{1}{2}\mathrm{tr}[g(t, x)g(t, x)^T\Psi_{xx}] + \Psi_x^T[f(t, x)+b(t, x)u], \tag{3}$$

where $\Psi\colon I\times R^n \to R^1$ has compact support and is continuous up to all its second order derivatives (Dynkin, 1953), and $(\cdot)^T$ and $\mathrm{tr}(\cdot)$ are transpose and trace oparators, respectively. The differential operators are defined as

$$(\cdot)_t = \frac{\partial(\cdot)}{\partial t}, \quad (\cdot)_x = \frac{\partial(\cdot)}{\partial x}, \quad (\cdot)_{xx} = \frac{\partial}{\partial x}\left(\frac{\partial(\cdot)}{\partial x}\right)^T.$$

A *pre-Hamiltonian* function of the system with respect to the given performance cost (2) and a control law $u(t, x)$ is defined as

$$H(x, \Psi_x, \Psi_{xx}, u, t) = L(t, x) + \|u\|^2 + \mathscr{L}_u\Psi. \tag{4}$$

The optimal control problem of stochastic systems can be stated now, as follows:

Optimal Stochastic Control Problem: For a given initial condition $(t_0, x_0)\in I\times R^n$, and the performance cost (2), find $u^*\in \Omega_u$, such that

$$V^*(t_0, x_0) \equiv J(u^*;\, t_0, x_0) = \inf_{u\in\Omega_u} J(u;\, t_0, x_0). \tag{5}$$

If it is assumed that the optimal control law, $u(x, t)^*$, exists and if the corresponding value function, $V^*(t, x)$, is sufficiently smooth, then u^* and V^* may be found by solving the well-known Hamilton-Jacobi-Bellman equation (Bellman, 1956),

$$\left.\begin{aligned} &V_t^* + \min\{\mathscr{L}_u V^* + L(t, x)+\|u\|^2\} = 0 \\ &V(T, x(T)) = \phi(T, x(T)) \end{aligned}\right\}. \tag{6}$$

Unfortunately, except in the case of linear quadratic Gaussian controls, where the problem has been solved (Wonham, 1970), a closed-loop form solution of the Hamilton-Jacobi-Bellman for solving the optimal stochastic control problem cannot be obtained in general when the system of Eq. (1) is nonlinear.

Instead, one may consider the optimal control problem relaxed to that of finding an admissible feedback control law, $u(x, t)$, that has an acceptable, but not necessarily optimal, cost function. This gives rise to a *stochastic suboptimal* *control solution* that could conceivably be solved with less difficulty than the original optimal stochastic control problem. The exact conditions for acceptability of a given cost function should be determined from practical considerations for the specific problem. The solution of the stochastic suboptimal control problem that converges successively to the optimal, is discussed in more detail in the next two sections.

3. An Approximation Theory of Stochastic Optimal Control

This section contains the main results of the appoximation theory for the solution of nonlinear stochastic control problems. Two theorems, one for the evaluation of performance of control laws and the other for the construction of lower and upper bounds of value functions, are established first. Then theoretical procedures which can lead to the iterative design of suboptimal controls are developed based on those two theorems.

Theorem 1. (Performance evaluation of a control law) Assume $V: I\times R^n \to R^1$ be an arbitrary function with continuous V, V_t, V_x and V_{xx} and satisfy the condition

$$\|V\| + \|V_t\| + \|x\|\|V_x\| + \|x\|^2\|V_{xx}\| < b(1+\|x\|^2), \tag{7}$$

where b is a suitable constant. Then the necessary and sufficient conditions for $V(t, x)$ to be the value function of an admissible fixed feedback control law $u(t, x)\in\Omega_u$, i.e.,

$$V(t, x) = E\left\{\int_t^T [L(t, x)+\|u\|^2]dt+\phi(T, x(T))/x(t)=x\right\}, \quad t\in I \tag{8}$$

are

$$V_t + \mathscr{L}_u V + L(t, x) + \|u\|^2 = 0, \tag{9}$$

$$V(T, x(T)) = \phi[T, x(T)]. \tag{10}$$

Proof. From (7), using *Itô's integration formula* (Itô, 1951), it follows that:

$$V(t, x) = E\left\{V(T, x(T)) - \int_t^T [\mathscr{L}_u V(\tau, x(\tau))+V_\tau(\tau, x(\tau))]d\tau/x(t)=x\right\}, \quad t\in I.$$

Therefore,

$$J(u; t, x) - V(t, x) = E\left\{\phi(T, x(T))-V(T, x(T))+\int_t^T[\mathscr{L}_u V(\tau, x(\tau)) + V_\tau(\tau, x(\tau))+L(\tau, x)+\|u(\tau, x)\|^2]d\tau/x(t)=x\right\}, \quad t\in I.$$

The sufficient condition results from the above equation. For the necessary condition, assume that $V(t, x)=J(u; t, x)$. Then, from the above equation, and for

$t = T$,

$$V(T, x(T)) = \phi(T, x(T)).$$

Therefore,

$$E\left\{\int_t^T [\mathscr{L}_u V(\tau, x(\tau)) + V_\tau(\tau, x(\tau)) + L(\tau, x) + \|u(\tau, x)\|^2] d\tau / x(t) = x\right\} = 0$$

has to be true for all $(t, x) \in I \times R^n$; hence,

$$V_t + \mathscr{L}_u V + L(t, x) + \|u\|^2 = 0,$$

which proves the necessary condition.

Remark 1: For a feedback controller that makes system (1) unstable, the associated $V(t, x)$ does not satisfy the assumptions of Theorem 1, particularly, condition Eq. (7).

Remark 2: The relation in Eqs. (9) and (10), called the *Generalized Hamilton-Jacobi-Bellman* equation for the stochastic control system (1), is reduced to the Hamilton-Jacobi-Bellman partial differential equation (6), when the control $u(x, t)$ is replaced by the optimal control $u^*(x, t)$.

Remark 3: The exact cost function for a given control law $u(x, t)$ is found by solving the *generalized Hamilton-Jacobi-Bellman* partial differential equation. This equation, though simpler and more general than the Hamilton-Jacobi-Bellman equation, is difficult to solve in a closed form.

Remark 4: For more general stochastic processes defined by the stochastic differential equation $dx = f(t, x, u)dt + g(t, x)dw$ and the general form of performance cost, one can show that Theorem 1 is still true.

Since it is generally difficult to find the exact cost functions satisfying Eqs. (9) and (10) of Theorem 1, the following theorem introduces a method of constructing the lower and upper bounds of the cost functions. This method can be used for the design of simpler suboptimal controllers based only on the upper bounds to value functions.

Theorem 2. (Lower and upper bounds of cost functions) For an admissible fixed feedback control law $u(t, x) \in \Omega_u$, and a continuous function $s(t, x)$, with $|s(t, x)| < \infty$ for all $(t, x) \in I \times R^n$. If the function $V(t, x)$ satisfies Eq. (7) with continuous V, V_t, V_x and V_{xx}, and

$$V_t + \mathscr{L}_u V + L(t, x) + \|u\|^2 \equiv \nabla V \leq s(t, x) \leq 0, \quad (\geq s(t, x) \geq 0), \tag{11}$$

$$V(T, x(T)) \geq \phi(T, x(T)), \quad (\leq \phi(T, x(T))), \tag{12}$$

then $V(t, x)$ is an upper (or a lower) bound of the cost function of system (1). That is,

$$V(t, x) \geq J(u; t, x) \; (\leq J(u; t, x)), \quad \forall (t, x) \in I \times R^n. \tag{13}$$

Proof. By a procedure similar to the proof of Theorem 1, it can be shown that

$$\begin{aligned}
J(u;t,x) &- V(t,x)\\
&= E\left\{\phi(T,x(T)) - V(T,x(T)) + \int_t^T [\mathscr{L}_u V(\tau, x(\tau))\right.\\
&\quad \left. + V_\tau(\tau, x(\tau)) + L(\tau, x) + \|u(\tau, x)\|^2]\, d\tau / x(t) = x\right\}\\
&= E\left\{\phi(T,x(T)) - V(T,x(T)) + \int_t^T \nabla V(\tau, x(\tau))\, d\tau / x(t) = x\right\}.
\end{aligned}$$

Therefore, from Eqs. (11), (12), it follows that,

$$\begin{aligned}
J(u;t,x) &- V(t,x)\\
&\le (\ge)\, E\left\{\int_t^T \nabla V(\tau, x(\tau))\, d\tau / x(t) = x\right\}\\
&\le (\ge)\, E\left\{\int_t^T \nabla s(\tau, x(\tau))\, d\tau / x(t) = x\right\}, \quad \forall (t,x) \in I \times R^n.
\end{aligned}$$

This completes the proof.

Remark 1: In general, the function $s(t, x)$ in this theorem does not need to be calculated. However, stating the inequality as in (11) gives an additional degree of flexibility that enables the determination of an upper (or a lower) bound to the cost function $J(u; t, x)$.

Remark 2: The function $V(t, x)$ in this theorem is the exact cost function for a system with a performance cost augmented by $-s(t, x) \ge 0$,

$$\begin{aligned}
I(u; t_0, x_0)&\\
= E&\left\{\int_{t_0}^T [L(t,x) + \|u\|^2 - s(t,x)]\,dt + \phi_s(T, x(T)) / x(t_0) = x_0\right\}, \qquad (14)
\end{aligned}$$

where $\phi_s(\cdot,\cdot)$ is a terminal cost function such that $\phi_s(T, x(T)) \ge (\le)$ $\phi(T, x(T))$.

Having established the two theorems for the evaluation of performance of a given feedback control law, it is necessary to develop algorithms to improve the control law. The followings Theorems 3–5 provide a theoretical procedure for designing the suboptimal feedback controllers based on the Theorem 1, while Theorem 6 presents a method for constructing upper and lower bounds to the optimal cost function, which can be used to evaluate the acceptability of suboptimal controllers.

Theorem 3. Given the admissible controls $u_1 \in \Omega_u$ and $u_2 \in \Omega_u$, with $V_1(t, x)$ and $V_2(t, x)$ be the corresponding cost functions satisfying Eqs. (7) and (8) for u_1 and u_2, respectively, define the Hamiltonian functions for $i = 1$ and 2,

$$H_{i\min} = H(x, V_{ix}, V_{ixx}, u_i^*, t) = L(t,x) + \|u\|^2 + \mathscr{L}_{u_i^*} V_i, \qquad (15)$$

where

$$u_i^*(t,x) = -\frac{1}{2} b(t,x)^T V_{ix}(t,x). \qquad (16)$$

It is shown that

$$V_1 \geq V_2, \tag{17}$$

when

$$V_{1t} + H_{1\min} \leq V_{2t} + H_{2\min}. \tag{18}$$

Proof. Let

$$\Delta V = V_2 - V_1, \quad \Delta V_t = V_{2t} - V_{1t},$$
$$\Delta V_x = V_{2x} - V_{1x}, \quad \Delta V_{xx} = V_{2xx} - V_{1xx},$$

then,

$$\begin{aligned} V_{2t} + H_{2\min} &= V_{1t} + \Delta V_t + L(t,x) + \frac{1}{2}\mathrm{tr}[g(t,x)g(t,x)^T V_{1xx}] \\ &\quad + \frac{1}{2}\mathrm{tr}[g(t,x)g(t,x)^T \Delta V_{xx}] + V_{1x}^T f(t,x) + \Delta V_x^T f(t,x) \\ &\quad - \frac{1}{4}\|b^T \Delta V_x\|^2 - \frac{1}{2} V_{1x}^T b b^T \Delta V_x \\ &= V_{1t} + H_{1\min} + \Delta V_t + \frac{1}{2}\mathrm{tr}[g(t,x)g(t,x)^T \Delta V_{xx}] \\ &\quad - \Delta V_x^T f(t,x) - \frac{1}{4}\|b^T \Delta V_x\|^2 - V_{1x}^T b\left(\frac{1}{2} b^T \Delta V_x\right) \\ &= V_{1t} + H_{1\min} - \frac{1}{4}\|b^T \Delta V_x\|^2 + \Delta V_t + \mathscr{L}_{u_1^*}\Delta V. \end{aligned}$$

Therefore,

$$\begin{aligned} &\Delta V_t + \mathscr{L}_{u_1^*}\Delta V \\ &= (V_{2t} + H_{2\min}) - (V_{1t} + H_{1\min}) + \frac{1}{4}\|b^T \Delta V_x\|^2, \end{aligned} \tag{19}$$

which, from assumption (18), implies that

$$\Delta V_t + \mathscr{L}_{u_1^*}\Delta V = \geq 0.$$

In addition, from Eq. (10) of Theorem 1,

$$\Delta V(T, x(T)) = V_2(T, x(T)) - V_1(T, x(T)) = 0.$$

However, applying Itô's integration formula to $\Delta V(t, x(T))$ along the trajectory generated by the control u_1^*, it follows that

$$\begin{aligned} &\Delta V(t,x) \\ &= -E\left\{ \int_t^T [\Delta V_\tau(\tau, x(\tau)) + \mathscr{L}_{u_1^*}\Delta V(\tau, x(\tau))]\, d\tau / x(t) = x \right\} \leq 0. \end{aligned}$$

Hence,

$$V_2(t, x) \leq V_1(t, x), \quad \forall (t, x) \in I \times R^n.$$

Remark: One should not try to find $\Delta V(t, x)$ by subtracting $V_1(t, x)$ from $V_2(t, x)$ directly based on their individual Itô's formulas evaluated along the trajectories generated by their corresponding controls. This is due to the fact that the two state trajectories, generated by $u_1(t, x)$ and $u_2(t, x)$ respectively, are different.

A combination of Theorems 1 and 3, where $V(t, x)$ represents the cost function of the system (1) when driven by control $u(t, x)$, yields an inequality that serves as a basis of suboptimal control algorithms, to iteratively reduce the cost of the performance of the system. This is outlined by the following theorem.

Theorem 4. Assume that there exist a control $u_1 \in \Omega_u$ and a corresponding function $V_1(t, x)$ satisfying Eqs. (7) and (8) of Theorem 1. If there exists a function $V_2(t, x)$ satisfying the same conditions of Theorem 1, its associated control $u_2 \in \Omega_u$, has been selected to satisfy

$$\left\| u_2 + \frac{1}{2} b^T \Delta V_{2x} \right\| \leq \left\| u_1 + \frac{1}{2} b^T \Delta V_{1x} \right\|, \tag{20}$$

then,

$$V_1 \geq V_2. \tag{21}$$

Proof. Since control u_1 and the corresponding value function V_1 must satisfy Eqs. (9) and (10), according to Theorem 1, it follows that for every $(t, x) \in I \times R^n$:

$$V_{1t} + L(t, x) + V_{1x}^T(f + bu_1) + \frac{1}{2}\operatorname{tr}[V_{1xx} gg^T] + \|u_1\|^2 = 0.$$

This can be rewritten as

$$V_{1t} + H_{1\min} + \left\| u_1 + \frac{1}{2} b^T V_{1x} \right\|^2 = 0,$$

that is,

$$V_{1t} + H_{1\min} = -\left\| u_1 + \frac{1}{2} b^T V_{1x} \right\|^2. \tag{22}$$

Similarly, one can find,

$$V_{2t} + H_{2\min} = -\left\| u_2 + \frac{1}{2} b^T V_{2x} \right\|^2. \tag{23}$$

Since,

$$\left\| u_2 + \frac{1}{2} b^T V_{2x} \right\|^2 \leq \left\| u_1 + \frac{1}{2} b^T V_{1x} \right\|^2.$$

it follows from Eqs. (22) and (23) that

$$V_{2t} + H_{2\min} \geq V_{1t} + H_{1\min}.$$

Hence, according to Theorem 2,

$$V_2(t, x) \leq V_1(t, x), \quad \forall (t, x) \in I \times R^n,$$

which proves the theorem.

Remark: Clearly, condition (20) in Theorem 4 is much easier to be tested than condition (18) in Theorem 3, since (20) does not involve the time derivative V_t and the infinitesimal generator $\mathscr{L}_u V$.

Based on Theorems 3 and 4, the following theorem establishes a sequence of feedback controls which successively improve the cost of performance of the system, and converge to the optimal feedback control.

Theorem 5. Let a sequence of pairs $\{u_i, V_i\}$ satisfy Eqs. (7), (8) of Theorem 1, and u_i be obtained by minimizing the pre-Hamiltonian function corresponding to the previous cost function V_{i-1}, that is,

$$u_i = -\frac{1}{2} b^T V_{i-1x}, \quad i = 1, 2, \cdots, \tag{24}$$

then the corresponding cost functions V_i, satisfies the inequality,

$$V_{i-1} \geq V_i, \quad i = 1, 2, \cdots. \tag{25}$$

Thus by selecting the pairs $\{u_i, V_i\}$ sequentially, in the above manner, the resulting sequence $\{V_i\}$ converges monotonically to the optimal cost function V^*, and the corresponding sequence $\{u_i\}$, converges to the optimal control u^*, associated with V^*.

Proof. Since the control u_i of (24) and the corresponding cost function V_i satisfy (9) and (10) of Theorem 1, it follows from (22) of Theorem 4 that

$$V_{it} + H_{i\min} = -\left\| u_i + \frac{1}{2} b^T V_{ix} \right\|^2 = -\left\| \frac{1}{2} b^T \Delta V_{ix} \right\|^2,$$

where $\Delta V_i = V_i - V_{i-1}$. Therefore, application of (19) of Theorem 2 yields

$$\begin{aligned} \Delta V_{it} + \mathscr{L}_{u_{i-1}^*} \Delta V_i &= (V_{it} + H_{i\min}) - (V_{i-1t} + H_{i-1\min}) + \frac{1}{4} \| b^T \Delta V_{ix} \|^2 \\ &= -\frac{1}{4} \| \Delta V_{ix}^T b \|^2 + \frac{1}{4} \| \Delta V_{i-1x}^T b \|^2 + \frac{1}{4} \| \Delta V_{ix}^T b \|^2 \\ &= \frac{1}{4} \| \Delta V_{i-1x}^T b \|^2 \geq 0. \end{aligned}$$

From (10),

$$\Delta V_i(T, x(T)) = V_i(T, x(T)) - V_{i-1}(T, x(T)) = 0,$$

hence, Itô's integration formula applied to ΔV_i, along the trajectory generated by u_{i-1}^* leads to the inequality,

$$\begin{aligned} &\Delta V_i(t, x) \\ &= -E\left\{ \int_t^T [\Delta V_{i\tau}(\tau, x(\tau)) + \mathscr{L}_{u_{i-1}^*} \Delta V_i(\tau, x(\tau))] \, d\tau / x(t) = x \right\} \leq 0, \end{aligned}$$

that is,

$$V_{i-1} \geq V_i,$$

which proves (25).

To show the convergence of the sequence, note that $\{V_i\}$ is a non-negative and monotonically decreasing sequence that satisfies (7). Therefore, the following limits exist:

$$\lim_{i \to \infty} V_i(t, x) = V^0(t, x) \tag{26}$$

and

$$\lim_{i \to \infty} V_{ix}(t, x) = V_x^0(t, x) \tag{27}$$

for all t and x, where V^0 is the limit of the cost functions. The corresponding limit of control sequence $\{u_i\}$ can be identified from (24) as,

$$\begin{aligned} u^0 = \lim_{i \to \infty} u_i(t, x) &= \lim_{i \to \infty} \left(-\frac{1}{2} b^T V_{i-1x}(t, x) \right) \\ &= -\left(\frac{1}{2} b^T V_x^0 \right)(t, x). \end{aligned} \tag{28}$$

Clearly, u^0 and V^0 thus obtained, still satisfy Eqs. (9) and (10) of Theorem 1. However, from the construction of control sequence $\{u_i\}$, u^0 minimizes the pre-Hamiltonian function associated with the value function V^0. In other words, u^0 and V^0 satisfy the Hamilton-Jacobi-Bellman equation for the optimal control of stochastic system (1)

$$V_t^* + \min_{u \in \Omega_u} \{ \mathscr{L}_u V + L(t, x) + \|u\|^2 \} = 0. \tag{29}$$

Hence,

$$u^0(t, x) = u^*(t, x) \quad \text{and} \quad V^0(t, x) = V^*(t, x), \quad \forall (t, x) \in I \times R^n \tag{30}$$

are the optimal control and the optimal value function of the stochastic control problem (5).

Remark 1: It follows from this theorem that the optimal feedback control u^* and the optimal cost function V^* are related by

$$u^*(t, x) = -\left(\frac{1}{2} b^T V_x^* \right)(t, x), \quad \forall (t, x) \in I \times R^n, \tag{31}$$

which is a relationship, that results from the minimization of the Hamiltonian function associated with the stochastic system (1).

Remark 2: As indicated by the conditions, in applying the theorem, assumptions must be made *a priori*, regarding the admissibility of the successively derived control laws and their corresponding value functions. However, for a nonlinear stochastic control system as in (1), the admissibility of the new control laws is not always easy to show.

Finally, the following theorem presents a method for the construction of an upper (or a lower) bound of the optimal cost function $V^*(t, x)$. Since the optimal cost function is extremely difficult to find, its upper (or lower) bounds can provide a practical measure to evaluate the effectiveness of the suboptimal controllers.

Theorem 6. Assume that there exists a function $V^s(t, x)$ satisfying condition (7) of Theorem 1, for which the associated control

$$u^s = -\frac{1}{2} b^T V_x^s(t, x) \tag{32}$$

is an admissible one. Then, $V^s(t, x)$ is an upper (or a lower) bound to the optimal cost function $V^*(t, x)$ of system (1), if it satisfies the following conditions:

$$V_t^s + \mathscr{L}_u V^s + L(t, x) + \|u\|^2 = s(t, x) \le 0 \quad (\ge 0), \tag{33}$$

$$V^s(T, x(T)) \ge \phi(T, x(T)) \quad (\le \phi(T, x(T))), \tag{34}$$

where $s(t, x)$ is continuous and $|s(t, x)| < \infty$, for all $(t, x) \in I \times R^n$.

Proof. From (31), and the Hamilton-Jacobi-Bellman equation, it is obvious that the optimal control and the optimal cost function, are related

$$V_t^* + H_{\min}^*(x, V_x^*, V_{xx}^*, u^*, t) = V_t^* + L(t, x) + \|u\|^2 + \mathscr{L}_{u^*} V^* = 0, \tag{35}$$

and similarly, for u^s and V^s,

$$V_t^s + H_{\min}^s(x, V_x^s, V_{xx}^s, u^s, t) = V_t^s + L(t, x) + \|u\|^2 + \mathscr{L}_{u_s} V^s = 0. \tag{36}$$

For $s(t, x) \le 0$, subtracting (35) from (36) yields

$$\Delta V_t + \mathscr{L}_{u_s} \Delta V = -s(t, x) + \frac{1}{4}\|b^T \Delta V_x\|^2 \ge 0,$$

where $\Delta V = V^*(t, x) - V^s(t, x)$. From assumption (34),

$$\begin{aligned} \Delta V(T, x(T)) &= V^*(T, x(T)) - V^s(T, x(T)) \\ &= \phi(T, x(T)) - V^s(T, x(T)) \le 0. \end{aligned}$$

Therefore, application of Itô's integration formula to $\Delta V(t, x)$ along the trajectory generated by control u^s yields,

$$\Delta V(t, x) \le -E\left\{ \int_t^T [\Delta V_\tau(\tau, x(\tau)) + \mathscr{L}_{u_s} \Delta V(\tau, x(\tau))] d\tau / x(t) = x \right\} \le 0.$$

So $V^s(t, x)$ is an upper bound to the optimal cost function $V^*(t, x)$. For $s(t, x) \ge 0$, subtracting (36) from (35) leads to,

$$\Delta V_t + \mathscr{L}_{u^*} \Delta V = s(t, x) + \frac{1}{4}\|b^T \Delta V_x\|^2 \ge 0,$$

where $\Delta V = V^s(t, x) - V^*(t, x)$. It can be shown, by using condition (34) and Itô's integration formula it follows that

$$\Delta V(t, x) \le -E\left\{\int_t^T [\Delta V_\tau(\tau, x(\tau)) + \mathscr{L}_{u^*}\Delta V(\tau, x(\tau))]\,d\tau / x(t) = x\right\} \le 0.$$

In this case, $V^s(t, x)$ is a lower bound to the optimal cost function $V^*(t, x)$.

Theorems, which lead to the design of simpler suboptimal controllers based on the upper and lower bounds of cost functions, may also be constructed. A more detailed discussion of such derivation for the infinite-time stochastic regulator problem, is given in the next section.

4. The Infinite-time Stochastic Regulator Problem

The infinite-time stochastic regulator problem is defined as a control problem for nonlinear stochastic system (1), with infinite duration $T \to \infty$. All state trajectories generated by admissible controls in Ω_u must be bounded uniformly in $I \times R^n$.

For the infinite-time stochastic regulator problem, and assuming that the system is stable, the Performance Cost exists and is defined as

$$J(u; t_0, x_0) = \lim_{T \to \infty} E\left\{\frac{1}{T}\int_{t_0}^T [L(t, x) + \|u\|^2]\,dt / x(t_0) = x_0\right\}. \tag{37}$$

A discussion of the stability of system (1) with Eq. (37) as $T \to \infty$, is given in the next section. Applying Itô's integration formula before the limit, the cost function becomes

$$V(t, x) = -E\left\{\int_t^T [\mathscr{L}_u V(\tau, x(\tau)) + V_\tau(\tau, x(\tau))]\,d\tau / x(t) = x\right\}, \quad t \in I, \tag{38}$$

where $V(t, x)$ satisfies $V(t, 0) = 0$ and (8) of Theorem 1 for all the possible state trajectories, which is true for all $u \in \Omega_u$.

All the theorems developed in the previous section are still valid for the infinite-time stochastic regulator problem, except that all the terminal conditions at $t = T$, in those theorems are no longer required. However, in this case, theorems can be constructed which can lead to the iterative design of simpler suboptimal controls based only on the upper and lower bounds of the cost functions. Since in general the upper and lower bounds can be obtained without solving the partial differential equation (9) of Theorem 1, those theorems have a great potential for application. Two of such theorems, corresponding to Theorems 3 and 4, are given in the sequel.

Theorem 7. Given admissible controls u_1 and $u_2 \in \Omega_u$, with $J_1(t, x)$ and $J_2(t, x)$ be their corresponding cost functions defined by (37), if there exist function pairs $V_1(t, x)$, $s_1(t, x) \ge 0$ and $V_2(t, x)$, $s_2(t, x) \le 0$ satisfying (11) of Theorem 2 for u_1 and u_2, respectively, then,

$$J_1 \ge J_2, \tag{39}$$

when

$$V_{1t} + H_{1\min} \leq V_{2t} + H_{2\min}. \tag{40}$$

Proof. Following the same procedure used in the proof for Theorem 3, one can show that

$$\Delta V_t + \mathscr{L}_{u_1^*}\Delta V = (V_{2t} + H_{2\min}) - (V_{1t} + H_{1\min}) + \frac{1}{4}\|b^T\Delta V_x\|^2 \geq 0,$$

where $\Delta V = V_2 - V_1$. Therefore, Itô's integration formula yields,

$$\Delta V(t, x) = -E\left\{\int_t^T [\Delta V_\tau(\tau, x(\tau)) + \mathscr{L}_{u_1^*}\Delta V(\tau, x(\tau))]\,d\tau / x(t) = x\right\} \leq 0,$$

hence,

$$V_2(t, x) \leq V_1(t, x), \quad \forall (t, x) \in I \times R^n,$$

which implies that

$$J_2(t, x) \leq \lim_{T\to\infty} \frac{1}{T} V_2(t, x) \leq \lim_{T\to\infty} \frac{1}{T} V_2(t, x) \leq J_1(t, x).$$

The next theorem is the counterpart of Theorem 4, and its proof can be carried out by the same procedure used in Theorem 4.

Theorem 8: Assume that there exist a control $u_1 \in \Omega_u$ and a function pair $\{V_1(t, x), s_1(t, x) \geq 0\}$ satisfying (11) of Theorem 2. If there exists a function pair $\{V_2(t, x), s_2(t, x) \leq 0\}$ satisfying the same condition of Theorem 2, of which the associated control $u_2 \in \Omega_u$, has been selected to satisfy

$$\left\| u_2 + \frac{1}{2} b^T V_{2x} \right\|^2 \leq \left\| u_1 + \frac{1}{2} b^T V_{1x} \right\|^2, \tag{41}$$

then,

$$J_1 \geq J_2, \tag{42}$$

where $J_1(t, x)$ and $J_2(t, x)$ are the cost functions of u_1 and u_2, respectively.

Note that neither Theorem 7 nor Theorem 8 is true for the stochastic system (1) with a cost function defined by (2).

5. Stability Considerations

Stability of the infinite-time approximate control problem, will be treated in a manner similar to the deterministic case:

It suffices to show that the Performance Index (37), *of the system* (1), *is bounded for all the controls generated by the Approximation Theory.*

Lemma 1. If Theorems 7 and/or 8 are satisfied, stability of the infinite horizon systems, driven by the subsequent controls generated by the approximation theory, is guarranteed if the first controller u_1 is selected to yield a bounded J_1.

The proof of the above Lemma is clearly established from the statements of Theorems 7 and/or 8.

In order to prove the stability of the system for the first controller, the following steps are appropriate for consideration:

1. System (1) must be Completely Controllable (CC), for all admissible controls u.
2. Since all the states are assumed available for measurement, system (1) is obviously Completely Observable (CO).
3. The Performance Cost (37) is bounded because, $J(u; t, x) = \lim_{T\to\infty} E\{1/T\int_t^T[L(\tau, x)+\|u\|^2]d\tau/x(t)=x\} = \lim_{T\to\infty} 1/T[V(t, x)]$.

From Itô's integration formula,

$$V(t, x) = -E\left\{\int_t^T [\mathscr{L}_u V(\tau, x(\tau))+V_\tau]d\tau/x(t)=x\right\},$$

where

$$\mathscr{L}_u V(t, x(t)) = \frac{1}{2}\operatorname{tr}[g(t, x)g(t, x)^T V_{xx}] + V_x^T[f(t, x)+b(t, x)u].$$

Select the first control u_1, to satisfy the previous theorems of the Approximation Theory, and the condition,

$$E[\mathscr{L}_u V(\tau, x(\tau))+V_\tau] \le M < \infty, \quad \forall\tau \in [t, T].$$

Then, using the Mean Value Theorem,

$$V(t, x) = E[\mathscr{L}_u V(\alpha, x(\alpha))+V_\alpha](T-t) \le M(T-t), \quad \alpha \in [t, T].$$

Then,

$$J(u; t, x) = \lim_{T\to\infty} E\left\{\frac{1}{T}\int_t^T[L(\tau, x)+\|u\|^2]\,d\tau/x(t)=x\right\} = \lim_{T\to\infty}\frac{1}{T}[V(t, x)]$$

$$= \lim_{T\to\infty}\frac{M(T-t)}{T} \le M < \infty. \tag{43}$$

The boundedness of the Performance Index $J(u; t, x)$ of the infinite-time problem, establishes the stability of the system for all the controllers derived from the Approximation Theory.

6. Design of Suboptimal Controllers

The optimal feedback control $u^*(t, x)$ and its associated $V^*(t, x)$ satisfying the Hamilton-Jacobi-Bellman equation, Eq. (6), obviously satisfy all the theorems developed in Sec. 3. However, in most of cases of nonlinear stochastic control systems, the optimal solution is very difficult, if not impossible, to implement either because the solution is unavailable or because some of the states are not available for measure. In both cases, the theory developed in Sec. 3 may serve to obtain controllers which can make the system stable, and then be successively modified to approximate the optimal solution. Upper and lower bounds of the value function of the nonlinear stochastic system may be used to evaluate the effectiveness of the approximation.

6.1 Exact design procedure This approach, based on the assumption that the cost function $V(t, x)$ for a control $u(t, x)$ can be found to satisfy Eqs. (9) and (10) of Theorem 1, may be implemented according to Theorems 3 and 4, by the following procedure:

1. Select a feedback control law $u_0(t, x)$ for system (1), set $i = 0$.
2. Find a $V_i(t, x)$ to satisfy Theorem 1 for $u_i(t, x)$.
3. Obtain a $u_{i+1}(t, x)$ and a $V_{i+1}(t, x)$ to satisfy Theorem 1, and Theorems 3 or 4 for u_i and V_i. u_{i+1} is an improved controller.
4. From Theorem 6, find a lower bound $V_L(t, x)$ to the optimal cost function $V^*(t, x)$, and then use $V_{i+1} - V_L$ as a measure to evaluate $u_{i+1}(t, x)$ as an approximation to the optimal control $u^*(t, x)$. If acceptable, stop.
5. If the approximation is not acceptable, repeat Step 2 by increasing index i by one and continue.

The improved controller u_{i+1} in Step 3 can also be constructed by using Theorem 5, if the corresponding cost function V_{i+1} can be obtained. When a lower bound to the optimal value function is difficult to find, Step 4 can be omitted and then the criterion for the acceptability of the approximation has to be determined based on other considerations.

Example 1. (Linear stochastic systems) In order to better comprehend the method, the design procedure of a suboptimal controller will be first applied to a linear stochastic system, the optimal solution of which is well known. The linear stochastic system is described by the following differential equation:

$$dx = A(t)x\,dt + B(t)u\,dt + G(t)\,dw.$$

The cost function of the system has the quadratic form

$$J(u; t_0, x_0) = E\left\{\int_{t_0}^{T} [x^T M(t)x + \|u\|^2]\,dt + x(T)^T \Phi x(T) / x(t_0) = x_0\right\}.$$

The infinitesimal generator of the linear stochastic process is

$$\mathscr{L}_u V(t, x) = \frac{1}{2}\operatorname{tr}[G(t)G(t)^T V_{xx}] + V_x^T[A(t)x + B(t)u].$$

Assume first a linear nonoptimal control,

$$u_1(t, x) = -K_1(t)x,$$

where $K_1(t)$ is a feedback matrix. The corresponding cost function is assumed to be

$$V_1(t, x) = s_1(t) + x^T S_1(t)x,$$

where s_1 and S_1 can be found by solving Eqs. (9) and (10) of Theorem 1, i.e.,

$$\frac{dS_1}{dt} + (A - BK_1)^T S_1 + S_1^T(A - BK_1) + M + K_1^T K_1 = 0, \quad S_1 = \Phi$$

and

$$s_1(t) = \int_t^T \mathrm{tr}[G(\tau)G(\tau)^T S_1(\tau)]\,d\tau.$$

The feedback law is improved by using Theorem 5. From Eq. (24),

$$\begin{aligned} u_i(t, x) &= -\frac{1}{2}B^T(t)V_{i-1x} \\ &= -B^T S_{i-1}(t)x = -K_i(t)x, \quad i \geq 2, \end{aligned}$$

and the corresponding cost function is assumed to be

$$V_i(t, x) = s_i(t) + x^T S_i(t)x, \qquad i \geq 2,$$

where s_i and S_i are determined by solving the equations,

$$\frac{dS_i}{dt} + (A - BK_i)^T S_i + S_i^T(A - BK_i) + M + K_i^T K_i = 0, \quad S_i(T) = \Phi,$$

$$s_i(t) = \int_t^T \mathrm{tr}[G(\tau)G(\tau)^T S_i(\tau)]\,d\tau.$$

As $i \to \infty$, $S_i(t)$ approaches S, the solution of the matrix Riccati equation, i.e.,

$$\frac{dS}{dt} + (A - BK)^T S + S^T(A - BK) + M + K^T K = 0,$$

$$S(T) = \Phi, \quad K = B^T S$$

and correspondingly, the control approaches to

$$u(t, x) = -B(t)^T S(t)x = -K(t)x,$$

which is the optimal control for the linear stochastic systems with quadratic performance criterion (Wonham, 1970).

This solution demonstrates the use of Theorem 5 to sequentially improve the control parameters towards the optimal values in a Linear Quadratic Gaussian system with well known solution.

Example 2. The second example illustrates the design method by the following nonlinear first-order stochastic system:

$$dx = -x\,dt + u\,dt + \frac{1}{2}x\,dw, \quad x(0) = x_0$$

with a cost function $J(u; t_0, x_0)$, selected to represent a minimun error, minimun input energy specification criterion, for a regulator control problem

$$J(u; t_0, x_0) = E\left\{\int_{t_0}^{\infty}[10x^2 + x^4 + u^2]\,dt / x(t_0) = x_0\right\}.$$

The infinitesimal generator of the stochastic process becomes

$$\mathscr{L}_u V(t, x) = \frac{1}{4}x^2 V_{xx} + (-x + u)V_x.$$

First assume a ***linear control law***,

$$u_1(x) = -ax, \quad a > 0.$$

The corresponding cost function is assumed to be

$$V_1(x) = s_1 x^2 + s_2 x^4.$$

Equation (9) of Theorem 1 yields

$$\left[10 + a^2 - s_1\left(2a + \frac{7}{4}\right)\right]x^2 + \left[1 - s_2\left(4a + \frac{5}{2}\right)\right]x^4 = 0,$$

which is true for

$$s_1 = \frac{40 + 4a^2}{8a + 7}, \quad s_2 = \frac{2}{8a + 5}, \quad a > 0.$$

Next, select a ***higher order control law***,

$$u_2(x) = -ax - bx^3, \quad b > 0.$$

Such a controller was selected to be of the same order as the partial derivative of the $V_1(x)$ cost function as per Theorem 5 suggests. The corresponding cost function is assumed to be,

$$V_2(x) = q_1 x^2 + q_2 x^4.$$

In this case, in order to satisfy Eq. (9), one must solve

$$\left[10 + a^2 - q_1\left(2a + \frac{7}{4}\right)\right]x^2 + \left[1 + 2ab - 2bq_1 - q_2\left(4a + \frac{5}{2}\right)\right]x^4 + (b^2 - 4bq_2)x^6 = 0,$$

which is true for

$$q_1 = \frac{40 + 4a^2}{8a + 7}, \quad q_2 = \frac{16a + 14}{675 - 16a}, \quad b = 4q_2,$$
$$0 < a < \frac{675}{16} = 42.1875.$$

To satisfy Theorem 4, controllers u_1 and u_2 must satisfy,

$$\| u^2 + q_1 x + 2q_2 x^3 \| \leq \| u_1 + s_1 x + 2 s_2 x^3 \|,$$

which yields

$$0 < a \leq \frac{\sqrt{689} - 7}{8} = 2.4061.$$

Their corresponding cost functions can be compared

$$V_1(x) = \frac{40 + 4a^2}{8a + 7}x^2 + \frac{2}{8a + 5}x^4,$$

$$V_2(x) = \frac{40+4a^2}{8a+7}x^2 + \frac{16a+14}{675-16a}x^4,$$

$$\Delta V(x) = V_2(x) - V_1(x) = \frac{32(4a^2+7a-40)}{(8a+5)(675-16a)}x^4 \geq 0.$$

Figure 1 illustrates V_1, V_2, and the corresponding performance improvement for $a = 0.2723$, and various initial states x.

If Theorem 5 is to be used for the above u_1 and V_1, u_2 must be selected according to

$$u_2(x) = -\frac{1}{2}V_{1x} = -s_1x - 2s_2x^3,$$

and a V_2 satisfying Eq. (9) of Theorem 1 exists, if

$$V_2(x) = q_1x^2 + q_2x^4,$$

$$q_1 = \frac{40+4s_1^2}{8s_1+7} = 2.7800,$$

$$q_2 = \frac{16s_1+14}{675-16s_1} = 0.1393,$$

$$a = 0.2723, \quad s_1 = 4.3904.$$

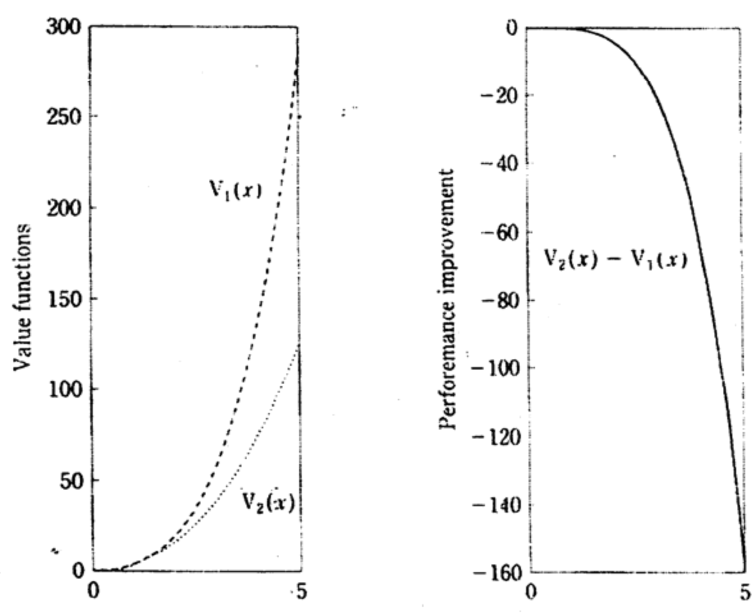

Fig. 1. Example 2 — Design with Theorem 4.

Comparing the cost functions, one finds that

$$V_1(x) = 4.3904x^2 + 0.2786x^4,$$

$$V_2(x) = 2.7800x^2 + 0.1393x^4,$$

$$\Delta V(x) = V_2(x) - V_1(x) = -1.6104x^2 - 0.1393x^4.$$

V_1, V_2 and the performance improvement are shown in Fig. 2, for various initial states x. Comparing Fig. 2 with Fig. 1, one can see that the performance improvement in this case is not as big as that made by the design using Theorem 4.

In both cases, since $\Delta V \leq 0$, the performance of the system has been improved by replacing a linear controller u_1 by a nonlinear controller u_2. All the above controllers make the origin an equilibrium point for the system.

6.2 Approximate design procedures for the regulator problem In many cases, the selection of a $V(t, x)$ to satisfy (9) and (10) of Theorem 1 is a very difficult task. In such a case approximate design procedures, which use the upper and lower bounds of the cost function obtained through Theorem 2 can be constructed. For the infinite time stochastic regulator problem, the following design procedure is proposed based on Theorems 7, 8 in Sec. 4:

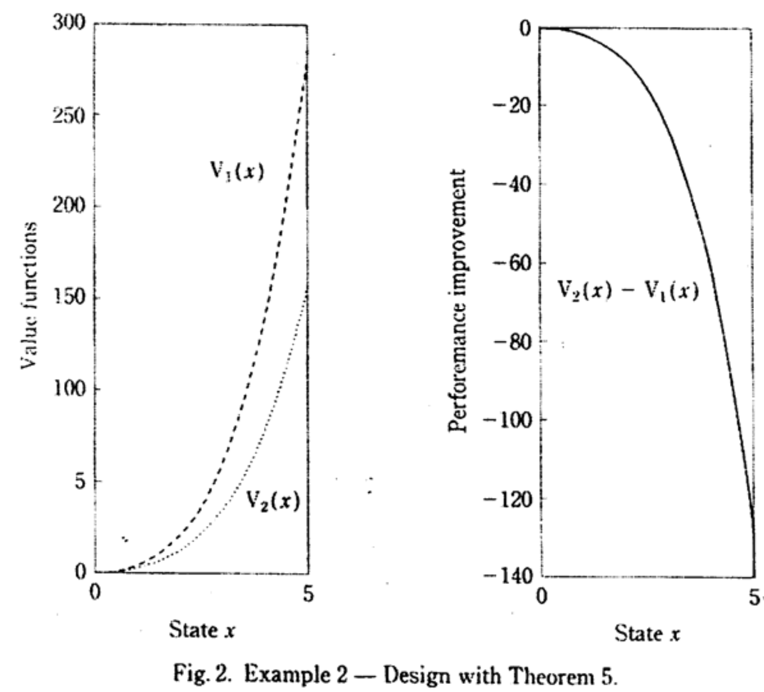

Fig. 2. Example 2 — Design with Theorem 5.

1. Select a feedback control law $u_0(t, x)$ for system (1), set $i = 0$.
2. For an $s_i(t, x) \geq 0$, find a $V_i(t, x)$ for u_i to satisfy Theorem 2 for a lower bound.
3. Obtain a $u_{i+1}(t, x)$, and for an $s_{i+1} \leq 0$, and find a $V_{i+1}(t, x)$ for u_{i+1} to satisfy Theorem 2 as an upper bound. $u_{i+1}(t, x)$, s_{i+1} and $V_{i+1}(t, x)$ found should also satisfy conditions (40) of Theorem 7 or (41) of Theorem 8 for the improvement of performance.
4. Using a lower bound to the optimal cost function, which is determined according to Theorem 6, the approximation of the optimal control can be measured. If acceptable, stop.
5. If the approximation is not acceptable, repeat Step 2 by increasing index i by one and continue.

Example 3. The design method is illustrated with the following nonlinear first-order stochastic regulator problem:

$$dx = x^3 dt + u dt + \frac{1}{2} x dw, \quad x(t_0) = x_0$$

with a cost function,

$$J(u; t_0, x_0) = \lim_{T \to \infty} E\left\{ \frac{1}{T} \int_{t_0}^{T} [10x^2 + x^4 + u^2] dt / x(t_0) = x_0 \right\}.$$

The infinitesimal generator of the stochastic process is

$$\mathscr{L}_u V(t, x) = \frac{1}{4} x^2 V_{xx} + (x^3 + u) V_x.$$

For a *linear control law,*

$$u_1(x) = -a_1 x, \quad a_1 > 0.$$

The lower bound of its cost function is assumed to be

$$V_1(x) = s_1 x^2 + s_2 x^4.$$

Application of Eq. (11) of Theorem 2 leads to

$$\left[10 + a_1^2 - s_1\left(2a_1 - \frac{1}{4}\right)\right] x^2 + \left[1 + 2s_1 - s_2\left(4a_1 - \frac{3}{2}\right)\right] x^4 + 4 s_2 x^6 \geq 0,$$

which is satisfied by

$$s^1 = \frac{40(1 - c_1) + 4a_1^2}{8a_1 - 1},$$

$$s_2 = \frac{2(1 - c_2) + 4s_1}{8a_1 - 3},$$

$$a_1 > \frac{3}{8} \quad \text{for any } 1 \geq c_1, \quad c_2 > 0.$$

For a *higher order control law,*

$$u_2(x) = -a_2 x - b_2 x^3, \quad a_2 > 0, \quad b_2 > 0,$$

the upper bound to its value function is assumed to be

$$V_2(x) = q_1 x^2 + q_2 x^4.$$

In this case, application of (11) yields,

$$\left[10 + a_2^2 - q_1\left(2a_2 - \frac{1}{4}\right)\right]x^2 + \left[1 + 2q_1 - 2(q_1 - a_2)b_2 - q_2\left(4a_2 - \frac{3}{2}\right)\right]x^4 + [b_2^2 - 4q_2(b_2 - 1)]x^6 \leq 0,$$

which is true for

$$q_1 = \frac{40(1+d_1) + 4a_2^2}{8a_2 - 1},$$

$$q_2 = \frac{(1+d_2)b_2^2}{4b_2 - 4},$$

$$1 + 2q_1 - 2(q_1 - a_2)b_2 - q_2\left(4a_2 - \frac{3}{2}\right) \leq 0, \quad a_2 > \frac{3}{8}, \quad b_2 > 1,$$

where $d_1, d_2 > 0$ are arbitrary.

Improvement of performance $\Delta V \leq 0$ occurs if Eq. (40) or Eq. (41) is satisfied, which leads to

$$s_1 = \frac{40(1-c_1) + 4a_1^2}{8a_1 - 1} \geq q_1 = \frac{40(1+d_1) + 4a_2^2}{8a_2 - 1},$$

$$s_2 = \frac{2(1-c_2) + 4s_1}{8a_1 - 3} \geq q_2 = \frac{(1+d_2) + b_2^2}{4b_2 - 4},$$

which, with the rest of the inequalities, produce acceptable values for a_1, a_2 and b_2. For example, one can show that

$$a_1 = 10, \quad a_2 = \frac{3}{2}, \quad b_2 = 2, \quad c_1 = c_2 = d_1 = d_2 = 0.1$$

is a set of the acceptable values. The lower and upper bounds of the value functions in this case are found as

$$V_1(x) = 5.5190x^2 + 0.3101x^4,$$

$$V_2(x) = 4.8182x^2 + 0.0688x^4,$$

$$\Delta V(x) = V_2(x) - V_1(x) = -0.7008x^2 - 0.2413x^4.$$

The results are illustrated in Fig. 3. Again, improvement of performance occurs by replacing a linear controller u_1 by a nonlinear controller u_2, of the same order as the partial derivative of $V_1(x)$. Note that in this case the actual cost function of control u_2 cannot be found by simply using the method applied in Example 2.

This approach has a great potential for application since one does not have to solve the partial differential equation of Theorem 1 every time.

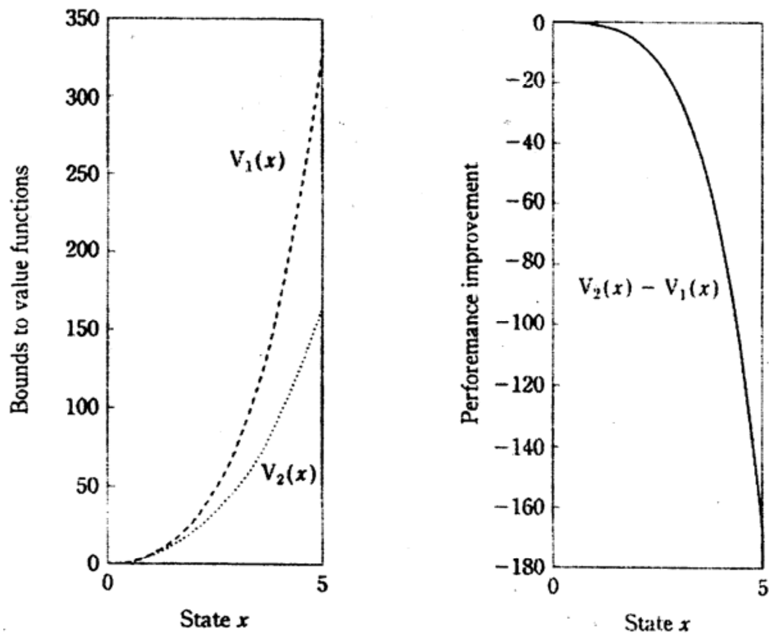

Fig. 3. Example 3 — Design with Theorem 8.

7. Entropy Reformulation of Suboptimal Control

A reformulation of the Approximation Theory using Entropy as the cost criterion, may provide an approach that integrates this method with the rest of the Generalized Control Theory created by Saridis (1988).

It was shown that the optimal control problem can be presented from an Entropy point of view, which provides a generalized formulation that answers many conceptual questions.

The Suboptimal Control problem, may also be reformulated using Jaynes's Principle of Maximun Entropy (Jaynes, 1957). A Generalized Hamilton-Jacobi-Bellman equation is derived for some admissible control $u(x)=u_i(x)$ using Jaynes's Principle of Maximun Entropy. Then the following theorem establishes the claim for system defined by Eq. (1).

Theorem 9. The performance cost $V(u_i(x); x, t)$, of system (1) for an admissible control $u_i(x)$, may be expressed by its associated Entropy $H(u_i(x, t))$, $i=1, \cdots$, derived by Jaynes's Principle of Maximun Entropy.

Proof. Derive $p(u(x, t))$ as Jaynes' Worst Entropy Probability Density Function:

$$p(u(x, t)) = \exp[-\lambda-\mu E\{V[u(x, t), x_0, t_0]\}]. \tag{44}$$

Then, the Associated Entropy is

$$H[u(x, t)] = \lambda + \mu E\{V[u(x), x_0, t_0]\}, \tag{45}$$

which may replace the Performance Criterion for Optimization. Jaynes' Principle of Maximun Entropy implies

$$\frac{dH}{dp} = 0.$$

The Incompressibility in time of Probability Density Condition implies

$$\frac{dp}{dt} = 0,$$

which for the selected $p(u(x, t))$ yields the ***Generalized Hamilton-Jacobi-Bellman*** equation for the Approximation Theory of stochastic systems

$$\left.\begin{aligned} &\frac{\partial V}{\partial t} + \frac{\partial V^T}{\partial x} f(x, u, t) + L(x, u, t) \\ &\quad + \frac{1}{2}\mathscr{L}\left[\operatorname{tr}\left(\frac{\partial}{\partial x}\left(\frac{\partial V}{\partial x}\right)^T gg^T\right)\right] = 0 \\ &V(T) = 0 \end{aligned}\right\}. \tag{46}$$

8. Conclusions

In this paper, an approximation theory of optimal control for nonlinear stochastic systems has been developed. The theory demonstrates the following points:

1. The iteration scheme and the theorem for the construction of upper and lower bounds of the cost functon, proposed by Saridis and Lee (1979) for deterministic systems, is generalized to the case of stochastic systems.
2. A successive design procedure using upper and lower bounds of the exact cost function has been developed for the infinite-time stochastic regulator problem. The determination of the upper and lower bounds requires the solution of a partial differential inequality instead of an equality. Therefore it provides a degree of flexibility in the design method over the exact design method.

Several examples are used to illustrate the application of the proposed approximation theory to stochastic control. It has been shown that in the case of linear quadratic Gaussian problems, the approximation theory leads to the exact solution of optimal control. Stability of the infinite-time suboptimal control problem was established under not very restrictive conditions, and stable sequences of controllers can be generated. Moreover, for the actual implementation of the design procedure, the computational problem has to be addressed. A reformulation of the suboptimal stochastic control problem, using Entropy as the basis for the derivation of the *Generalized Hamilton-Jacobi-Bellman* equation, is also given as an extension of the work by Saridis (1988).

9. References

Al'brekht, E.G. (1961). On the optimal stabilization of nonlinear systems. *J. Appl. Math. Mech. (ZPMM)*, **25**, 5, 1254–1266.

Aoki, M. (1967). *Optimization of Stochastic Systems*. Academic Press, N.Y.

Bellman, R. (1956). *Dynamic Programming*. Princeton University Press, Princeton, N.J.

Doob, J.L. (1953). *Stochastic Processes*. Wiley, N.Y.

Dynkin, E.B. (1953). *Markov Processes*. Academic Press, N.Y.

Itô, K. (1951). *On Stochastic Differential Equations*. Mem. Amer. Math. Soc., **4**.

Jaynes, E.T. (1957). Information theory and statistical mechanics. *Physical Review*, 4, 106.

Kushner, H. (1971). *Introduction to Stochastic Control*. Holt, Reinhardt and Winston, N.Y.

Kwakernaak, H. and R. Sivan (1972). *Linear Optimal Control Systems*. Wiley, N.Y.

Leake, R.J. and R.-W. Liu (1967). Construction of suboptimal control sequences. *SIAM J. Control*, **5**, 1, 54–63.

Lukes, D.L. (1969). Optimal regulation of nonlinear dynamical systems. *SIAM J. Control*, **7**, 1, 75–100.

Nishikawa, Y., N. Sannomiya and H. Itakura (1962). A method for suboptimal design of nonlinear feedback systems. *Automatica*, **7**, 6, 703–712.

Ohsumi, A. (1984). Stochastic control with searching a randomly moving target. *Proc. of American Control Conference*, San Diego, CA, 500–504.

Panossian, H.V. (1988). Algorithms and computational techniques in stochastic optimal control. C.T. Lenodes (ed.), *Control and Dynamic Systems*, **28**, 1, 1–55.

Rekasius, Z.V. (1964). Suboptimal design of intentionally nonlinear controllers. *IEEE Trans. Automatic Control*, **AC-9**, 4, 380–386.

Sage, A.P. and C.C. White (1977). *Optimum Systems Control*. Prentice-Hall, Englewood Cliffs, N.J.

Saridis, G.N. (1988). Entropy formulation for optimal and adaptive control. *IEEE Trans. Automatic Control*, **AC-33**, 8, 713–721.

Saridis, G.N. and J. Balaram (1986). Suboptimal control for nonlinear systems. *Control-Theory and Advanced Technology (C-TAT)*, **2**, 3, 547–562.

Saridis, G.N. and C.S.G. Lee (1979). An approximation theory of optimal control for trainable manipulators. *IEEE Trans. Systems, Man, and Cybernetics*, **SMC-9**, 3, 152–159.

Skorokhod, A.V. (1965). *Studies in the Theory of Random Processes*. Addison-Wesley, Reading, MA.

Wang, F. and G.N. Saridis (1992). Suboptimal control for nonlinear stochastic systems. *Proceedings of 31st Conference on Decision and Control*, Tucson, AZ, Dec.

Wonham, W.M. (1970). Ramdom differential equations in control theory. A.T. Bharucha-Reid(ed.), *Probabilistic Methods in Applied Mathematics*, Academic Press, N.Y.

George N. Saridis received the Diploma in Mechanical and Electrical Engineering from the National Technical University of Athens Greece in 1955, the MSEE in 1962 and Ph.D. degrees in 1965 from Purdue University, W. Lafayette, IN. In 1988, he was certified Manufacturing Engineer for Machine Vision by the Society of Manufacturing Engineers.

Since September, 1981, he has been Professor of the Electrical, Computer and Systems Engineering Department and Director of the Robotics and Automation Laboratories at Rensselaer Polytechnic Institute, in Troy, N.Y.

Dr. Saridis is a Fellow of IEEE and a member of the New York Academy of Science, ASME, SPIE, AAAS, ASEE and senior member of SME. In 1974 and 1981, he was appointed Vice-Chairman of the IFAC Committee on Education and in 1981–84 the Survey Paper Editor of *AUTOMATICA*.

In 1983–84, he was the *Founding President* of the IEEE Council of Robotics and Automation. He has written 6 books and over 350 book chapters, journal articles, conference papers and reports.

Dr. Saridis is the recipient of the IEEE Centennial Medal Award in 1984, and the IEEE Control Systems Society's Distinguished Member Award in 1989, and the certificate of the IEEE Society of Robotics and Automation as its Founding President.

Fei-Yue Wang received his B.E. degree in chemical engineering from Shandong Institute of Chemical Technology, Qingdao, China, M.S. degree in mechanics from Zhejiang University, Hangzhou, China and Ph.D. degree in computer and systems engineering from Rensselaer Polytechnic Institute, Troy, New York, in 1981, 1984 and 1990, respectively.

From 1984 to 1986, he was an instructor at the Department of Mechanics, Zhejiang University. In 1986, he was awarded Pao Yu-Kong and Pao Zao-Long Scholarship for Chinese Students for his academic achievements. From 1988 to 1990, he was with the NASA Center for Intelligent Robotic Systems for Space Exploration at RPI. He jointed the University of Arizona in 1990, where he is presently an Assistant Professor of the Systems and Industrial Engineering. In his previous work in mechanics and applied mathematics, he contributed significantly to the theoretical development of shells, plates, planes, three-dimensional elasticity and micropolar elasticity for both isotropic and anisotropic materials. He is the co-author of the book *Coordination Theory of Intelligent Machines: Applications in Intelligent Robotic Systems and CIM Systems*, and has published more than forty refereed journal articles and book chapters, and about fifty conference papers and technical reports in the areas of applied mathematics, mechanics, mechatronics, artificial intelligence, control theory, communication, robotics and automation, computer-integrated manufacturing, fuzzy logic, neural networks and intelligent machines. He is editor-in-chief of *Journal of Intelligent Control and Systems*, editor-in-chief of *World Scientific Book Series on Intelligent Control and Intelligent Automation*, associate editor of *Journal of Intelligent Automation & Soft Computing*, and Chair of the Working Group on Manufacturing System Design of IEEE Computer-Aided Control System Design Committee. His fields of interest include mechatronics, robotics and automation, computer vision, computer-integrated manufacturing, intelligent controls and intelligent systems.

Dr. Wang is a member of Sigma Xi, IEEE, Association for Computing Machinery (ACM) and National Council on Systems Engineering (NCoSE).