15th Triennial World Congress, Barcelona, Spain

INVERSE OPTIMAL CONTROLLER DESIGN FOR STRICT-FEEDBACK STOCHASTIC SYSTEMS WITH EXPONENTIAL-OF-INTEGRAL COST

Cheng Tang and Tamer Başar

Coordinated Science Laboratory, University of Illinois
1308 W. Main Street, Urbana, Illinois 61801-2307, USA
(cheng, tbasar)@decision.csl.uiuc.edu

Abstract: For a class of nonlinear stochastic systems in strict-feedback form, where the diffusion coefficients depend on the state, we obtain risk-sensitive state-feedback controllers which are both globally inverse optimal and locally sub-optimal. These controllers also lead to closed-loop system trajectories that are bounded in probability.

Keywords: Strict-feedback system; stochastic stability; risk-sensitive control; stochastic backstepping design.

1. INTRODUCTION

The last decade has witnessed considerable effort on stabilization and control of stochastic nonlinear systems. One way of approaching these problems, with an eye on robustness, is through a risk-sensitive formulation. This approach has received much attention, particularly in the light of the established relationship between risk-sensitive stochastic control (RSSC) problems and a particular class of stochastic zero-sum differential games (Whittle, 1990; Başar and Bernhard, 1995; Fleming and McEneaney, 1992; Runolfsson, 1994), both involving the solution of a particular Hamilton-Jacobi-Bellman (HJB) equation. Solving this HJB equation has presented a formidable task when the system dynamics are nonlinear, and this difficulty has driven the need to look into nonlinear systems exhibiting special structures, which might lead to constructive solutions of the HJB equation. One such structure is the *strict-feedback* form, which has been studied in (Pan and Başar, 1999). With some positive cost term $q(x)$ and no cost on control, a stochastic backstepping tool was developed, as a generalization of the backstepping methodology developed for deterministic systems (Krstić *et al.*, 1995), to obtain a controller that delivers any prespecified achievable long-term average cost, while leading to closed-loop system trajectories that are bounded in probability. In a related work (Krstić and Deng, 1998), the stochastic stabilization problem for strict-feedback systems has been considered, and an inverse optimal control law constructed using a quartic stochastic Lyapunov function, instead of the traditional quadratic one.

Noting that nonlinear systems in strict-feedback form admit controllable linearizations, and cost functions can admit quadratic approximations, a further development in this area has been to construct control laws that meet both a local design specification and a global one, with also positive cost on control. In the specific case when the diffusion coefficient does not depend on the state, a stochastic backstepping design can solve locally a related LEQG problem and globally an inverse RSSC problem, as shown in (Başar and Tang, 2000). This result can be viewed as a generalization of those developed for deterministic strict-feedback systems (Ezal *et al.*, 2000; Ezal *et al.*, 2001).

In this paper, we extend the previous results to encompass the larger class of strict-feedback stochastic systems where the diffusion coefficients are allowed to depend on the state. We construct a stabilizing state-feedback controller with appealing global and local optimality properties. The next section introduces the design problem, along with the notions of global and local optimality. In section 3, a state-feedback controller is constructed recursively using a stochastic backstepping tool. A numerical example illustrating the design concludes the paper.

2. PROBLEM FORMULATION

We consider the following stochastic nonlinear system in strict-feedback form:

$$
\begin{aligned}
dx_1 &= [x_2 + f_1(x_1)]dt + h_1'(x_1)dw_t \\
\vdots \quad & \quad \vdots \\
dx_{n-1} &= [x_n + f_{n-1}(x_{[n-1]})]dt + h_{n-1}'(x_{[n-1]})dw_t \\
dx_n &= [f_n(x_{[n]}) + b(x_{[n]})u]dt + h_n'(x_{[n]})dw_t
\end{aligned} \tag{1}
$$

where x is the n-dimensional state, u is the scalar control input, w is an r-dimensional standard vector Wiener process, and $x_{[k]} := (x_1, \dots, x_k)'$. The underlying probability space is the triple $(\Omega, \mathcal{F}, \mathbf{P})$. The functions $f_i : \mathbb{R}^i \to \mathbb{R}, h_i : \mathbb{R}^i \to \mathbb{R}^r, i = 1, \cdots, n$, and $b : \mathbb{R}^n \to \mathbb{R}$ are smooth, with $f_i(0) = 0, h_i(0) > 0$, and $b(x) > 0, \forall x$. $u(t) = \mu(x(t))$, where $\mu \in \mathcal{U}$, the set of all locally Lipschitz continuous state-feedback control laws. To facilitate the exposition, we rewrite (1) as

$$dx = [f(x) + G(x)u]dt + H(x)dw_t \tag{2}$$

with f, G, H appropriately defined. Note that here $H(x)$ depends on the state x, instead of being a constant as in (Başar and Tang, 2000).

Associated with system (1), we introduce an exponential-of-integral cost function

$$J = \limsup_{T\to\infty} \frac{2}{\theta T} \ln E[\exp(\frac{\theta}{2}\int_0^T q(x) + r(x)u^2 dt)] \tag{3}$$

where $\theta > 0$ is the risk-sensitive parameter, and $q(\cdot), r(\cdot)$ are nonnegative-definite (*n.n.d.*) continuous functions. One of our goals is to design $\mu^* \in \mathcal{U}$ so as to achieve global inverse optimality, i.e. to attain $J^* = \inf_{\mu\in\mathcal{U}} J(\mu)$.

Definition 1. A control law $\mu \in \mathcal{U}$ is *globally inverse optimal (g.i.o.)* for system (1) if it achieves the optimal value J^* of (3) for some $q(\cdot)$ and $r(\cdot)$, and some $\theta > 0$.

From RSSC theory, given that there exists a g.i.o. controller, then it can be expressed as

$$\mu^*(x) = -\frac{1}{2} r^{-1}(x) G'(x) V_x'(x) \tag{4}$$

where V is obtained as the solution of

$$
\begin{aligned}
J^* &= \min_u \Big\{ V_x(x)f(x) + V_x(x)G(x)u + q(x) + r(x)u^2 \\
&\quad + \frac{\theta}{4}|H'(x)V_x'(x)|^2 + \frac{1}{2}\mathrm{Tr}[V_{xx}(x)H(x)H'(x)] \Big\} \\
&= V_x(x)f(x) - \frac{1}{4}V_x(x)G(x)r^{-1}(x)G'(x)V_x'(x) \\
&\quad + \frac{\theta}{4}|H'(x)V_x'(x)|^2 + q(x) + \frac{1}{2}\mathrm{Tr}[V_{xx}(x)H(x)H'(x)]
\end{aligned} \tag{5}
$$

In addition to g.i.o. for the nonlinear system, we also wish to achieve local optimality or sub-optimality for a corresponding linearized system, with respect to some n.n.d. quadratic functions $x'Qx$ and Ru^2 in place of $q(\cdot)$ and $r(\cdot)u^2$ in (3). Toward this end, we rewrite (2) as:

$$
\begin{aligned}
dx = [&Ax + \tilde{f}(x) + Bu + \tilde{G}(x)u]dt \\
&+ (D + \tilde{H}(x))dw_t
\end{aligned} \tag{6}
$$

where $A = f_x(0)$, $B = G(0) =: (\,0\ \cdots\ 0\ b_0\,)'$ and $D = H(0) =: (d_1\ \cdots\ d_n)'$, with obvious definitions for the perturbation terms $\tilde{f}, \tilde{G}$ and $\tilde{H}$. Denote the unperturbed or linearized versions of x and u by x_ℓ and u_ℓ, respectively. Then, the linearized system is given by

$$dx_\ell = [Ax_\ell + Bu_\ell]dt + Ddw_t. \tag{7}$$

Note that (A, B) is a controllable pair by the structure of these matrices. Consider now the LEQG problem with dynamics (7) and cost function

$$J_\ell = \limsup_{T\to\infty} \frac{2}{\theta T} \ln E[\exp(\frac{\theta}{2}\int_0^T x'Qx + Ru^2 dt)] \tag{8}$$

where (A, Q) is observable. From LEQG theory, since (A, B) is controllable and (A, Q) is observable, there exists a finite number θ_ℓ^*, such that for each $\theta < \theta_\ell^*$, this risk-sensitive stochastic control problem admits the unique solution

$$\mu_\ell^*(x_\ell) = -R^{-1}B'Px_\ell, \tag{9}$$

where P is the minimal p.d. solution of

$$A'P + PA - P\left(BR^{-1}B' - \theta DD'\right)P + Q = 0. \tag{10}$$

Furthermore, the optimal cost is $J_\ell^* = \mathrm{Tr}(PDD')$ and $A - R^{-1}BB'P$ is Hurwitz.

Definition 2. Consider system (1) with its exponential-of-integral cost function (3), where the following relationship holds:

$$r(0) \le R, \quad q_{xx}(0) \le Q. \tag{11}$$

A g.i.o. control law $\mu \in \mathcal{U}$ is *locally sub-optimal (l.s.o.)* if for some R, Q satisfying (11), the linearized control law μ_ℓ is optimal for the corresponding LEQG problem (7)-(8). μ is *locally optimal* if μ_ℓ is optimal for the LEQG problem, with equality holding in (11) and J^* attained.

3. INVERSE OPTIMAL CONTROLLER DESIGN

A g.i.o. state feedback controller is constructed in two steps. First, the linearized system is considered, to which a backstepping-based design is applied. Then, this construction is extended to the original nonlinear system.

3.1 *Linear Optimal Design*

The linear optimal controller design is essentially the same as the one presented in (Başar and Tang, 2000; Ezal *et al.*, 2000). Consider the linear system (7). The solution to (5) is

$$V_\ell(x) = x'Px, \quad J_\ell^* = \mathrm{Tr}(DD'P). \tag{12}$$

Apply a coordinate transformation $z = Lx$ based on a Cholesky decomposition of P, $P = L'\Delta L$, where Δ is a diagonal matrix consisting of the positive constants

δ_i's, and L is a lower triangular matrix. This brings the GARE (10) into the form

$$\bar{A}'\Delta + \Delta\bar{A} - \Delta\left(BR^{-1}B' - \theta\bar{D}\bar{D}'\right)\Delta + \bar{Q} = 0 \qquad (13)$$

where $\bar{A} = LAL^{-1}$, $\bar{D} = LD$ and $\bar{Q} = (L')^{-1}QL^{-1}$. An important observation here is that $\bar{A}$ has the same structure as in the original coordinate system. This, together with the diagonal structure of Δ, leads to the important property that each principal minor of the GARE is itself a GARE for the corresponding minor of Δ. The transformed subsystem can be represented by

$$dz_{[k]} = \left[\bar{A}_{[k]}z_{[k]} + (0 \cdots z_{k+1})'\right]dt + \bar{D}_{[k]}dw_t$$

for $1 \le k < n$, where $z_{[k]} = L_{[k]}x_{[k]}$. This enables one to proceed with backstepping by recursively generating the value function

$$\bar{V}_\ell(z_{[n]}) = \sum_{i=1}^{n} \delta_i z_i^2 \equiv z'\Delta z, \qquad (14)$$

with $\bar{V}_{\ell i} = z'_{[i]}\Delta_{[i]}z_{[i]}$ being the value function for the i-th subsystem. In view of GARE

$$\bar{A}'_{[i]}\Delta_{[i]} + \Delta_{[i]}\bar{A}_{[i]} + \theta\Delta_{[i]}\bar{D}_{[i]}\bar{D}'_{[i]}\Delta_{[i]} + \bar{Q}_{[i]} = 0, \qquad (15)$$

one arrives at the following expression for $d\bar{V}_{\ell i}$:

$$d\bar{V}_{\ell i} = 2z'_{[i]}\Delta_{[i]}\bar{D}_{[i]}dw_t + \Big(\mathrm{Tr}[\Delta_{[i]}\bar{D}_{[i]}\bar{D}'_{[i]}] \\ -\theta z'_{[i]}\Delta_{[i]}\bar{D}_{[i]}\bar{D}'_{[i]}\Delta_{[i]}z_{[i]} - z'_{[i]}\bar{Q}_{[i]}z_{[i]} + 2z_i\delta_i z_{i+1}\Big)dt.$$

When $z_{i+1} = 0$, $\bar{V}_{\ell i}$ satisfies a HJB equation with cost $J_i = \mathrm{Tr}[\Delta_{[i]}\bar{D}_{[i]}\bar{D}'_{[i]}]$. Proceed with the steps from $i = 1$ to $n-1$, and at the last step let $z_n = x_n - \bar{\alpha}_{n-1}(z_{[n-1]})$. The actual control emerges as $x_{n+1} = b_0 u$, and the transformed system is now described by

$$dz(t) = (\bar{A}z(t) + Bu_l(t))dt + \bar{D}dw_t. \qquad (16)$$

This also leads to an expression for $d\bar{V}_\ell$ as in the solution of the LEQG problem:

$$d\bar{V}_\ell = 2z'\Delta\bar{D}dw_t + \left(\mathrm{Tr}[\Delta\bar{D}\bar{D}'] - \theta z'\Delta\bar{D}\bar{D}'\Delta z \\ -z'\bar{Q}z + 2z_n\delta_n b_0 u_\ell + R^{-1}b_0^2\delta_n^2 z_n^2\right)dt$$

with the optimal control being $\bar{\mu}_l(z) = -R^{-1}B'\Delta z$ and cost $J_l^* = \mathrm{Tr}[\Delta\bar{D}\bar{D}']$. When the controller $u_l = \bar{\mu}_l(z)$ is applied, the value function $\bar{V}_\ell$ satisfies a transformed HJB equation

$$\bar{V}_{\ell z}\bar{A}z - \frac{1}{4}\bar{V}_{\ell z}(BR^{-1}B' - \theta\bar{D}\bar{D}')\bar{V}_{\ell z} \\ +z'\bar{Q}z + \frac{1}{2}\mathrm{Tr}[\bar{V}_{\ell zz}\bar{D}\bar{D}'] = J_l^*. \qquad (17)$$

3.2 *Nonlinear Inverse Optimal Design*

Now, for the actual nonlinear system (2) in strict feedback form, we need to apply a different transformation from the linear design. The reason is that with the perturbation in (6) brought in, the second-order term appears in the corresponding HJB equation and cannot be eliminated by simply using the above backstepping method. The transformation we use now is a nonlinear mapping $z = \Phi(x)$, and this construction results in a lower-triangular diffeomorphism. Moreover, $\Phi(x) = Lx + \tilde{\Phi}(x)$, where Lx is linear part of this diffeomorphism, and $\tilde{\Phi}(x)$ contains only the higher-order terms. Combining with an appropriate choice of cost term $\bar{q}(z)$, we are able to obtain a globally optimal control law $u = \bar{\mu}(z)$ with respect to a RSSC problem, where the cost term is $\bar{q}(z) + \bar{r}(z)u^2$, similar to (3).

At the first step, define $z_1 = x_1$ and select $\bar{V}_1 = z'_{[1]}\Delta_{[1]}z_{[1]} = \delta_1 z_1^2$ as the value function. Then,

$$dz_1 = [\bar{a}_{[1]}z_1 + \hat{f}_1(z_1) + (x_2 - \bar{\alpha}_1(z_1))]dt \\ +\bar{h}'_1(z_1)dw_t \qquad (18)$$

where $\bar{a}_{[1]} = a_{[1]} + \bar{\alpha}_{[1]}$, $\hat{f}_1(z_1) = \tilde{f}_1(z_1) + \hat{\alpha}_1(z_1)$ and $\bar{h}_1(z_1) = h_1(z_1)$. The virtual control law is $\bar{\alpha}_1(z_1) = \bar{\alpha}_{[1]}z_1 + \hat{\alpha}_1(z_1)$, with $\bar{\alpha}_{[1]}$ given by the locally optimal backstepping design in previous subsection, and $\hat{\alpha}_1(\cdot)$ being an additional nonlinear term. Using GARE (15) with $i = 1$, we have

$$d\bar{V}_1 = 2z_1\delta_1\bar{h}'_1(z_1)dw_t + \big([\delta_1\bar{d}'_1\bar{d}_1] \\ -\theta z_1\delta_1\bar{h}'_1(z_1)\bar{h}_1(z_1)\delta_1 z_1 - \bar{q}_1(z_1) \\ +[\delta_1\bar{h}'_1(z_1)\bar{h}_1(z_1) - \delta_1\bar{d}'_1\bar{d}_1] - z_1\bar{Q}_{[1]}z_1 + \bar{q}_1(z_1) \\ +2z_1\delta_1\big(\hat{f}_1(z_1) + \frac{\theta}{2}\left(\bar{h}'_1(z_1)\bar{h}_1(z_1) - \bar{d}'_1\bar{d}_1\right)\delta_1 z_1 \\ +x_2 - \bar{\alpha}_1(z_1)\big)\big)dt.$$

Pick $\hat{\alpha}_1(z_1)$ to cancel out the nonlinear terms:

$$\hat{\alpha}_1(z_1) = -\tilde{f}_1(z_1) - \frac{\theta}{2}(\bar{h}'_1(z_1)\bar{h}_1(z_1) - \bar{d}'_1\bar{d}_1)\delta_1 z_1,$$

and let $c_1 = \hat{\alpha}_1(0)$ be the possible drift introduced by the nonlinear transformation. At this particular stage, $c_1 = 0$, which has been included for clarity and ease in the design at later stages. In addition, let $z_2 = x_2 - (\bar{\alpha}_1(z_1) - c_1)$ and

$$\bar{q}_1(z_1) = z_1\bar{Q}_{[1]}z_1 - \big(\delta_1\bar{h}'_1(z_1)\bar{h}_1(z_1) \\ -\delta_1\bar{d}'_1\bar{d}_1\big) + 2c_1\delta_1 z_1 + 2\bar{Q}_{[1]}^{-1}\delta_1^2 c_1^2. \qquad (19)$$

We have the following expression for $d\bar{V}_1$:

$$d\bar{V}_1 = 2z_1\delta_1\bar{h}'_1(z_1)dw_t + (J_1 - \bar{q}_1(z_1) \\ -\theta z_1\delta_1\bar{h}'_1(z_1)\bar{h}_1(z_1)\delta_1 z_1 + 2z_1\delta_1 z_2)\,dt, \qquad (20)$$

where $J_1 = \delta_1\bar{d}'_1\bar{d}_1 + 2\bar{Q}_{[1]}^{-1}\delta_1^2 c_1^2$. Then, $\bar{V}_1$ satisfies the HJB equation

$$\bar{V}_{1z}\left(\bar{A}_{[1]}z_1 + \hat{f}_1(z_1)\right) + \frac{\theta}{4}\bar{V}_{1z}\bar{h}'_1(z_1)\bar{h}_1(z_1)\bar{V}'_{1z} \\ +\bar{q}_1(z_1) + \frac{1}{2}\mathrm{Tr}[\bar{V}_{1zz}\bar{h}'_1(z_1)\bar{h}_1(z_1)] = J_1.$$

We repeat the preceding step from $k = 1$ to $i-1$, where at the i-th step, we define

$$z_i = \phi_i(x_{[i]}) = x_i - (\bar{\alpha}_{i-1}(z_{[i-1]}) - c_{i-1})$$

with $\bar{\alpha}_{i-1}(z_{[i-1]}) = \bar{\alpha}_{[i-1]}z_{[i-1]} + \hat{\alpha}_{i-1}(z_{[i-1]})$, $c_{i-1} = \hat{\alpha}_{i-1}(0)$ being the possible non-zero drift, and select the value function $\bar{V}_i$ to be

$$\bar{V}_i = \bar{V}_{i-1} + \delta_i z_i^2 = z'_{[i]}\Delta_{[i]}z_{[i]}.$$

Assume the following dynamics for $z_{[i-1]}$:

$$dz_{[i-1]} = \big[\bar{A}_{[i-1]}z_{[i-1]} + \hat{f}_{[i-1]}(z_{[i-1]}) + \big(0 \ \cdots \ x_i - \bar{\alpha}_{i-1}(z_{[i-1]})\big)'\big]\,dt + \bar{H}_{[i-1]}dw_t$$

and the following Itô differential for $\bar{V}_{i-1}$:

$$d\bar{V}_{i-1} = 2z'_{[i-1]}\Delta_{[i-1]}\bar{H}_{[i-1]}(z_{[i-1]})dw_t + \big(J_{i-1} - \bar{q}_{i-1}(z_{[i-1]}) + 2z_{i-1}\delta_{i-1}z_i - \theta z'_{[i-1]}\Delta_{[i-1]}\bar{H}_{[i-1]}\bar{H}'_{[i-1]}\Delta_{[i-1]}z_{[i-1]}\big)\,dt, \tag{21}$$

with the choice of J_{i-1} and $\bar{q}_{i-1}$ as

$$J_{i-1} = \mathrm{Tr}[\Delta_{[i-1]}\bar{D}_{[i-1]}\bar{D}'_{[i-1]}] + 2c'_{[i-1]}\Delta_{[i-1]}\bar{Q}^{-1}_{[i-1]}\Delta_{[i-1]}c_{[i-1]},$$
$$\bar{q}_{i-1} = z'_{[i-1]}\bar{Q}_{[i-1]}z_{[i-1]} - \Big(\mathrm{Tr}[\Delta_{[i-1]}\bar{H}'_{[i-1]}\bar{H}_{[i-1]}] - \mathrm{Tr}[\Delta_{[i-1]}\bar{D}'_{[i-1]}\bar{D}_{[i-1]}]\Big) + 2c'_{[i-1]}\Delta_{[i-1]}z_{[i-1]} + 2c'_{[i-1]}\Delta_{[i-1]}\bar{Q}^{-1}_{[i-1]}\Delta_{[i-1]}c_{[i-1]}$$

where $c_{[i-1]} = (\,c_1 \ \cdots \ c_{i-1}\,)'$ is the drift vector due to the nonlinear transformation. Then, letting

$$\bar{\alpha}_i(z_{[i]}) = \bar{\alpha}_{[i]}z_{[i]} + \hat{\alpha}_i(z_{[i]})$$

be the virtual control law for x_{i+1}, we get

$$dz_i = [\bar{a}_{[i]}z_{[i]} + \hat{f}_i(z_{[i]}) + (x_{i+1} - \bar{\alpha}_i(z_{[i]}))]dt + \bar{h}'_i(z_{[i]})dw_t \tag{22}$$

with $\bar{a}_{[i]}$ defined in the locally optimal design, and

$$\hat{f}_i(z_{[i]}) = \alpha_{[i]}\hat{\Psi}_{[i]}(z_{[i]}) + \tilde{f}_i(\Phi^{-1}_{[i]}(z_{[i]})) + \hat{\alpha}_i(z_{[i]}) - \bar{\alpha}_{i-1,i-1}c_{i-1} - \bar{\alpha}_{[i-1]}\hat{f}_{[i-1]} - \frac{\partial\hat{\alpha}_{i-1}}{\partial z_{[i-1]}}\hat{f}_{[i-1]} - \frac{\partial\hat{\alpha}_{i-1}}{\partial z_{[i-1]}}[\bar{A}_{[i-1]}z_{[i-1]} + (0 \ \cdots \ z_i - c_{i-1})'] - \frac{1}{2}\mathrm{Tr}\left[\frac{\partial^2\hat{\alpha}_{i-1}}{\partial z^2_{[i-1]}}\bar{H}_{[i-1]}(z_{[i-1]})\bar{H}'_{[i-1]}(z_{[i-1]})\right]$$
$$\bar{h}_i = h_i(\Phi^{-1}_{[i]}(z_{[i]})) - \frac{\partial\bar{\alpha}_{i-1}}{\partial z_{[i-1]}}\bar{H}_{[i-1]}(z_{[i-1]})$$

Here, the term $\hat{\Psi}_{[i]}(z_{[i]}) = \Phi^{-1}_{[i]}(z_{[i]}) - L^{-1}_{[i]}z_{[i]}$ contains the higher-order terms. After the above transformation, the $z_{[i]}$-subsystem becomes

$$dz_{[i]} = \big[\bar{A}_{[i]}z_{[i]} + \hat{f}_{[i]}(z_{[i]}) + \big(0 \ \cdots \ x_{i+1} - \bar{\alpha}_i(z_{[i]})\big)'\big]\,dt + \bar{H}_{[i]}dw_t. \tag{23}$$

Noting that $d\bar{V}_i = d\bar{V}_{i-1} + d(\delta_i z_i^2)$ and using (21), we can select $\hat{\alpha}_i(z_{[i]})$ to cancel out the nonlinear terms, i.e.,

$$\hat{\alpha}_i(z_{[i]}) = -\alpha_{[i]}\hat{\Psi}_{[i]}(z_{[i]}) - \tilde{f}_i(\Phi^{-1}_{[i]}(z_{[i]})) + \bar{\alpha}_{i-1,i-1}c_{i-1} + \bar{\alpha}_{[i-1]}\hat{f}_{[i-1]}(z_{[i-1]}) + \frac{\partial\hat{\alpha}_{i-1}}{\partial z_{[i-1]}}\hat{f}_{[i-1]}(z_{[i-1]}) + \frac{\partial\hat{\alpha}_{i-1}}{\partial z_{[i-1]}}[\bar{A}_{[i-1]}z_{[i-1]} + (0 \ \cdots \ z_i - c_{i-1})'] + \frac{1}{2}\mathrm{Tr}\left[\frac{\partial^2\hat{\alpha}_{i-1}}{\partial z^2_{[i-1]}}\bar{H}_{[i-1]}(z_{[i-1]})\bar{H}'_{[i-1]}(z_{[i-1]})\right] - \theta(\bar{h}'_i(z_{[i]})\bar{H}'_{[i-1]}(z_{[i-1]}) - \bar{d}'_i\bar{D}'_{[i-1]})\Delta_{[i-1]}z_{[i-1]} - \frac{\theta}{2}(\bar{h}'_i(z_{[i]})\bar{h}_i(z_{[i]}) - \bar{d}'_i\bar{d}_i)\delta_i z_i$$

In addition, pick $c_i = \hat{\alpha}_i(0)$ and

$$z_{i+1} = x_{i+1} - (\bar{\alpha}_i(z_{[i]}) - c_i)$$

as the nonlinear coordinate transformation for x_{i+1}. Select $\Delta\bar{q}_i(z_{[i]})$ such that

$$\bar{q}_i(z_{[i]}) = \bar{q}_{i-1}(z_{[i-1]}) + \Delta\bar{q}_i(z_{[i]}) = z'_{[i]}\bar{Q}_{[i]}z_{[i]} - \Big(\mathrm{Tr}[\Delta_{[i]}\bar{H}_{[i]}\bar{H}'_{[i]}] - \mathrm{Tr}[\Delta_{[i]}\bar{D}_{[i]}\bar{D}'_{[i]}]\Big) + 2c'_{[i]}\Delta_{[i]}z_{[i]} + 2c'_{[i]}\Delta_{[i]}\bar{Q}^{-1}_{[i]}\Delta_{[i]}c_{[i]}.$$

The form of $\bar{q}_i(z_{[i]})$ is now different from the one in (Başar and Tang, 2000) with extra terms brought in because of the state dependent diffusion coefficient $\bar{H}_{[i]}(z_{[i]})$ and the nonlinear transformation. Additionally, it results (by necessity) in a much more complex form of the nonlinear term $\hat{\alpha}_i(z_{[i]})$. These choices result in:

$$d\bar{V}_i = 2z'_{[i]}\Delta_{[i]}\bar{H}_{[i]}(z_{[i]})dw_t + \big(J_i - \bar{q}_i(z_{[i]}) - \theta z'_{[i]}\Delta_{[i]}\bar{H}_{[i]}(z_{[i]})\bar{H}'_{[i]}(z_{[i]})\Delta_{[i]}z_{[i]} + 2z_i\delta_i z_{i+1}\big)\,dt,$$

where

$$J_i = \mathrm{Tr}[\Delta_{[i]}\bar{D}_{[i]}\bar{D}'_{[i]}] + 2c'_{[i]}\Delta_{[i]}\bar{Q}^{-1}_{[i]}\Delta_{[i]}c_{[i]}.$$

Then, $\bar{V}_i$ satisfies a corresponding HJB equation

$$\bar{V}_{iz}\left(\bar{A}_{[i]}z_{[i]} + \hat{f}_{[i]}(z_{[i]})\right) + \frac{\theta}{4}\bar{V}_{iz}\bar{H}_{[i]}(z_{[i]})\bar{H}'_{[i]}(z_{[i]})\bar{V}'_{iz} + \bar{q}_i(z_{[i]}) + \frac{1}{2}\mathrm{Tr}[\bar{V}_{izz}\bar{H}_{[i]}(z_{[i]})\bar{H}'_{[i]}(z_{[i]})] = J_i.$$

At the final step of the backstepping procedure, we let

$$z_n = x_n - (\bar{\alpha}_{n-1}(z_{[n-1]}) - c_{n-1})$$

with $c_{n-1} = \hat{\alpha}_{n-1}(0)$, and choose

$$\bar{V}(z) = \bar{V}_{n-1}(z_{[n-1]}) + \delta_n z_n^2 = z'\Delta z \tag{24}$$

as the value function. The dynamics of z_n are given by (22) with $i = n$, $x_{n+1} = \bar{b}(z)u$ and $\bar{\alpha}_n = 0$. Therefore, we have obtained a coordinate transformation $z = \Phi(x)$ through a nonlinear backstepping procedure, and in this new coordinate system, the system is described by

$$dz = (\bar{A}z + \hat{f}(z) + \bar{G}(z)u)dt + \bar{H}(z)dw_t \tag{25}$$

and the Itô differential of value function $\bar{V}$ is

$$d\bar{V} = 2z'\Delta\bar{H}(z)dw_t + \Big\{\mathrm{Tr}[\Delta\bar{D}\bar{D}'] + 2c'_{[n-1]}\Delta_{[n-1]}\cdot \cdot\bar{Q}^{-1}_{[n-1]}\Delta_{[n-1]}c_{[n-1]} - \theta z'\Delta\bar{H}(z)\bar{H}'(z)\Delta z + \bar{r}^{-1}(z)\bar{b}(z)\delta_n^2 z_n^2 + \Delta\bar{q}_n(z) + \mathrm{Tr}[\delta_n\bar{h}'_n(z)\bar{h}_n(z)] - \bar{q}(z) - \mathrm{Tr}[\delta_n\bar{d}'_n\bar{d}_n] - z'Qz + z'_{[n-1]}Q_{[n-1]}z_{[n-1]} + 2z_n\delta_n\left(\theta\left[\bar{h}'_n(z)\bar{H}'_{[n-1]} - \bar{d}'_n\bar{D}'_{[n-1]}\right]\Delta_{[n-1]}z_{[n-1]} + \frac{\theta}{2}\left[\bar{h}'_n(z)\bar{h}_n(z) - \bar{d}'_n\bar{d}_n\right]\delta_n z_n + \hat{f}_n(z) + \bar{b}(z)u\right) + (R^{-1}b_0^2 - \bar{r}^{-1}(z)\bar{b}^2(z))\delta_n^2 z_n^2\Big\}\,dt \tag{26}$$

To achieve g.i.o., we need to find some n.n.d. functions $\bar{q}(z)$ and $\bar{r}(z)$ such that a form of HJB equation (5) holds, with the control law being

$$\bar{\mu}(z) = -(1/2)\bar{r}^{-1}(z)\bar{G}(z)\bar{V}_z = -\bar{r}^{-1}(z)\bar{b}(z)\delta_n z_n. \tag{27}$$

The desired Itô differential $d\bar{V}$ has the form

$$d\bar{V} = 2z'\Delta\bar{H}(z)dw_t + (J^* - \bar{q}(z) \\ -\theta z'\Delta\bar{H}(z)\bar{H}'(z)\Delta z - \bar{r}(z)^{-1}\bar{b}^2(z)\delta_n^2 z_n^2)\,dt \quad (28)$$

Comparing (26) with (28), this yields

$$\bar{q}(z) = \frac{1}{2}z'\bar{Q}z + \frac{1}{2}\left|z + 2\bar{Q}^{-1}\Delta c\right|^2_{\bar{Q}} - (\mathrm{Tr}[\Delta\bar{H}\bar{H}' \\ -\Delta\bar{D}\bar{D}']) + (\bar{r}^{-1}(z)\bar{b}^2 - R^{-1}b_0^2)\delta_n^2 z_n^2 - 2z_n\delta_n\bar{\eta}(z) \quad (29)$$

where $c_n = -\hat{f}_n(0)$, $c = (\,c_1 \;\cdots\; c_n\,)'$, and

$$\bar{\eta}(z) = \hat{f}_n(z) + c_n + \theta\left(\bar{h}_n'(z)\bar{H}_{[n-1]}' - \bar{d}_n'\bar{D}_{[n-1]}'\right)\cdot \\ \cdot\Delta_{[n-1]}z_{[n-1]} + \frac{\theta}{2}(\bar{h}_n'(z)\bar{h}_n(z) - \bar{d}_n'\bar{d}_n)\delta_n z_n,$$

and $\bar{r}(z) > 0$ is constructed such that $\bar{q}(z)$ is p.d. To achieve l.s.o., we want to ensure that condition (11) is satisfied in the z coordinate, which can be achieved by a judicious choice of $\bar{r}(z)$. Since the leading term in (29) is quadratic, we also need:

Assumption 1. For some $K > 0$, $\bar{H}(z)$ satisfies

$$\mathrm{Tr}[\Delta\bar{H}(z)\bar{H}'(z)] \le \mathrm{Tr}[\Delta\bar{D}\bar{D}'] + K|z|^2. \quad (30)$$

Further, $\mathrm{Tr}[\Delta\bar{H}(z)\bar{H}'(z)]$ is convex at the origin.

This technical assumption is necessary to bound the variance of the state, hence to make the cost term $\bar{q}(\cdot)$ be nonnegative, and thus to ensure the stabilizability of the system and existence of an optimal control law. The construction of $\bar{r}(\cdot)$ is of course not unique as in the deterministic case. In fact, one possible design here is the one given in (Ezal *et al.*, 2000). We do not give details here, but simply the main result.

Theorem 3. Consider the stochastic nonlinear system (1) with coordinate transformation $z = \Phi(x)$ through a nonlinear backstepping design. Let Assumption 1 hold, and $\bar{Q} > 2KI_n, R > 0$. Then, there exist a n.n.d. function $\bar{q}(z)$ and a strictly p.d. function $\bar{r}(z)$ satisfying (11) in the z coordinate, such that with the control law (27), the CL system is l.s.o. with respect to

$$\bar{J}_\ell = \limsup_{T\to\infty}\frac{2}{\theta T}\ln E[\exp(\frac{\theta}{2}\int_0^T(z'\bar{Q}z + Ru^2)dt)], \quad (31)$$

where $J_\ell^* = \mathrm{Tr}[\Delta\bar{D}\bar{D}']$, and g.i.o. with respect to

$$\bar{J} = \limsup_{T\to\infty}\frac{2}{\theta T}\ln E[\exp(\frac{\theta}{2}\int_0^T(\bar{q}(z) + \bar{r}(z)u^2)dt)], \quad (32)$$

where $J^* = J_\ell^* + 2c'\Delta\bar{Q}^{-1}\Delta c$. Furthermore, the CL system trajectory is bounded in probability.

Proof. The proof for the theorem except for the last statement on boundedness has already been outlined in the preceding derivation. We note that the condition $\bar{V}(z) \le c_1[\bar{q}(z) + \bar{r}(z)\bar{\mu}^2(z)] + c_2$ is satisfied for some positive constants c_1, c_2, where $\bar{V}(z)$ is the value function for the RSSC problem. It follows from (Pan and Başar, 1999) that the transformed system (25) is stable (bounded in probability).

Remark 4. By applying the stochastic backstepping method introduced in this section, we are able to obtain the state feedback control law (27) with desirable global and local properties, with the only constraint being condition (30) which restricts the growth of the diffusion coefficient. This condition is expressed in the transformed coordinate z, which is hard to check and verify. Further extensions of current work would be to relax this condition or convert it to the original coordinate system, and relate it to the matching conditions in nonlinear robust control problems.

Remark 5. The difference between the RSSC optimal cost J^* (with respect to system (1)) and the LEQG optimal cost J_ℓ^* (with respect to the linearized system (7)) is $2c'\Delta\bar{Q}^{-1}\Delta c$, which is dependent on the dynamics of the nonlinear system as well as the nonlinear coordinate transformation Φ. A challenging question is whether it is possible to design a control law such that this difference is made as small as possible.

Remark 6. When the diffusion coefficient does not depend on the state, and the additional nonlinear transformation terms $\hat{\alpha}_i$'s are homogeneous with their n-th order derivatives, one can verify that Assumption 1 is satisfied and $c = 0$. Thus in that case the constructed control law is both g.i.o. and locally optimal.

4. EXAMPLE

Consider the system:

$$dx_1 = (0.25x_1 - 0.00007x_1^3 + x_2)dt + 0.1x_1dw_1 + dw_2 \\ dx_2 = (0.25x_1 + x_1^2 + 0.25x_2 + x_2^2 + u)dt \\ +dw_1 - 0.1x_1dw_2$$

The linearized system is

$$dx_1 = (0.25x_1 + x_2)dt + dw_2 \\ dx_2 = (0.25x_1 + 0.25x_2 + u)dt + dw_1$$

with quadratic cost functional

$$J_1 = \limsup_{T\to\infty}\frac{2}{\theta T}\ln E[\exp(\frac{\theta}{2}\int_0^T\frac{3}{4}(x_1^2 + x_2^2) + 4u^2dt)]$$

In this case $\theta_\ell^* = 0.18$. Picking $\theta = 0.018$, the relevant solution of (10) is

$$P = \begin{bmatrix} 4.41258 & 4.81184 \\ 4.81184 & 7.77023 \end{bmatrix}.$$

The linear optimal controller is

$$u = -B'Px = -4.81184x_1 - 7.77023x_2. \quad (33)$$

If we apply this linear controller to the full-order nonlinear system, the CL system will not be globally stable. Now following the steps of the earlier derivation, we first rewrite P as $P = L'\Delta L$,

$$L = \begin{bmatrix} 1 & 0 \\ 0.619267 & 1 \end{bmatrix}, \quad \Delta = \begin{bmatrix} 1.43277 & 0 \\ 0 & 7.77023 \end{bmatrix},$$

and use the linear transform $z = Lx$, that is $z_1 = x_1$, $z_2 = x_2 + 0.619267x_1$. In the new coordinate system,

(33) is now expressed as $u = -B'\Delta z = -7.77023z_2$. Carrying out the nonlinear part of the design, we have $\hat{\alpha}_1 = 0$ which leads to the same coordinate transformation as

$$z = \Phi(x) = \begin{bmatrix} x_1 \\ 0.619267x_1 + x_2 \end{bmatrix}.$$

The nonlinear system is now described by

$$\begin{aligned} dz_1 &= [-0.3693z_1 - 0.0000716z_1^3 + z_2]dt \\ &\quad +0.1z_1dw_1 + dw_2 \\ dz_2 &= [(-0.13349 - 1.2385z_2)z_1 + 1.3835z_1^2 \\ &\quad +0.8693z_2 + z_2^2 - 0.00004436z_1^3 + u]dt \\ &\quad +(1 + 0.06193z_1)dw_1 + (0.6193 - 0.1z_1)dw_2. \end{aligned}$$

Note that $c_1 = c_2 = 0$, and Assumption 1 is satisfied, i.e. with $K = 0.25$, $\mathrm{Tr}[\Delta\bar{H}(z)\bar{H}(z)' - \Delta\bar{D}\bar{D}'] \leq K|z|^2$. Let

$$\bar{V}(z) = z'\Delta z = 1.43277z_1^2 + 7.77023z_2^2.$$

Then, using Itô's formula as well as GARE (13), we obtain a form of (28), i.e.

$$\begin{aligned} d\bar{V}(z) &= (0.28655z_1^2 + 15.5405z_2 + 0.96237z_1z_2)dw_1 \\ &\quad +(2.8655z_1 - 1.554z_1z_2 + 9.6237z_2)dw_2 \\ &\quad +[12.1828 - 0.00021z_1^4 - 0.0006894z_1^3z_2 \\ &\quad -(120.753\bar{r}^{-1}(z) - 13.5088)z_2^2 + 15.0519z_2^3]dt \\ &\quad +[15.5405z_2^3 + (-0.93632 + 21.5001z_2)z_1^2 \\ &\quad +(0.791z_2 - 19.2474z_2^2)z_1]dt \end{aligned}$$

The nonlinear feedback control law is given by

$$u_1 = \bar{\mu}(z) = -7.77023\bar{r}^{-1}(z)z_2, \qquad (34)$$

or, in the original coordinate system,

$$u_1 = -7.77023r^{-1}(x)(0.6193x_1 + x_2), \qquad (35)$$

where $r(x) = \bar{r}(\Phi(x_1))$. This controller is globally optimal for cost function (32), where

$$\begin{aligned} \bar{q}(z) &= (-14.3441 + 60.3764\bar{r}^{-1}(z))z_2^2 - 15.5405z_2^3 \\ &\quad -0.008353(-0.04259 + z_2)(2573.97 + z_2)z_1^2 \\ &\quad -0.0006894z_1^3z_2 - 0.9289z_1z_2 + 19.2474z_1z_2^2 \\ \bar{r}^{-1}(z) &= 0.25 + [-0.1088z_1 + 2.4303z_1^2 \\ &\quad +0.0001559z_1^3 + 0.2574z_2]_+ \end{aligned}$$

As shown in Fig. 1, $\bar{q}(z)$ is a p.d. function, and its second order term at the origin is indeed greater than $\underline{Q}$. Now, with $\underline{Q} > 2KI_2$ and from Theorem 1, the controller (35) achieves both l.s.o. and g.i.o., and the CL system is bounded in probability.

5. CONCLUSION

In this paper, we have studied the RSSC problem for strict-feedback stochastic nonlinear systems and constructed feedback control laws that are globally inverse optimal and locally suboptimal. The control laws further lead to system trajectories that are bounded in probability. These results hold under certain growth conditions on the system nonlinearities, which are expressed in a new coordinate system driven by the backstepping methodology. Further research is needed to relax these conditions and express them in the original coordinate system, as well as to obtain similar results for stochastic systems that are not in strict feedback form.

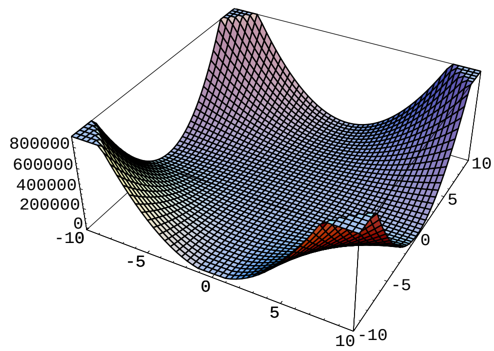

Fig. 1. $\bar{q}(z)$ function

6. REFERENCES

Başar, T. and C. Tang (2000). Locally optimal risk-sensitive controllers for strict-feedback nonlinear systems. *Journal of Optimization Theory and Applications* **105**, 521–541.

Başar, T. and P. Bernhard (1995). *H^∞-Optimal Control and Related Minimax Design*. 2nd ed.. Birkhauser. Boston, Massachusetts.

Ezal, K., P.V. Kokotović, A.R. Teel and T. Başar (2001). Disturbance attenuating output-feedback control of nonlinear systems with local optimality. *Automatica* **37**(6), 805–818.

Ezal, K., Z. Pan and P.V. Kokotović (2000). Locally optimal and robust backstepping design. *IEEE Trans. on Automat. Contr.* **45**(2), 260–271.

Fleming, W.H. and W.M. McEneaney (1992). Risk sensitive optimal control and differential games. In: *Stochastic Theory and Adaptive Control, Lecture Notes in Control and Information Sciences*. Vol. 184. pp. 185–197. Springer-Verlag. Berlin, Germany.

Krstić, M. and H. Deng (1998). *Stabilization of Nonlinear Uncertain Systems*. Springer.

Krstić, M., I. Kanellakopoulos and P.V. Kokotović (1995). *Nonlinear and Adaptive Control Design*. Wiley. New York, NY.

Pan, Z. and T. Başar (1999). Backstepping controller design for nonlinear stochastic systems under a risk-sensitive cost criterion. *SIAM Journal on Control and Optimization* **37**, 957–995.

Runolfsson, T. (1994). The equivalence between infinite-horizon optimal control of stochastic systems with exponential of integral performance index and stochastic differential games. *IEEE Transactions on Automatic Control* **39**, 1551–1563.

Whittle, P. (1990). *Risk-Sensitive Optimal Control*. Wiley. New York, NY.