Optimal Risk-Sensitive Control for Third Degree Polynomial Systems

Ma. Aracelia Alcorta G., Michael Basin, Sonia G. Anguiano R., Yosefat Nava A.

Abstract—The optimal exponential-quadratic control problem is considered for stochastic Gaussian systems with polynomial third degree drift terms and intensity parameters multiplying diffusion terms in the state equation. The closed-form optimal control algorithm is obtained using a quadratic value function as a solution to the corresponding Hamilton-Jacobi-Bellman equation. The performance of the obtained risk-sensitive regulator for stochastic third degree polynomial systems is verified in a numerical example, through comparing the exponential-quadratic criteria values for the optimal risk-sensitive control and third degree control algorithms. The simulation results reveal strong advantages in favor of the designed risk-sensitive algorithm in regard to the final criteria values for all values of the parameter $\varepsilon$.

I. INTRODUCTION

After the optimal linear stochastic control problem was solved (see [10], [7]), the optimal control theory for nonlinear stochastic systems is based on dynamic programming (Hamilton-Jacobi-Bellman) equation [7] and the maximum principle of Pontryagin [14]. A long tradition of the optimal control design was developed for nonlinear systems with respect to a quadratic Bolza-Meyer criterion (see, for example, [11]). The optimal control problems with respect to nontraditional criteria were also considered: the stochastic linear exponential-quadratic regulator (LEQR) problem was introduced in [9]. Further connection between the LEQR problem and $H_{\infty}$—control via a minimum entropy principle was given in [8]. Whittle ([18], [19]) considered problems on a finite-time horizon, using "small-noise" asymptotics. When the process being controlled is governed by stochastic differential equation, the Whittle’s formula for the optimal large-derivatives rate was obtained using partial differential equation viscosity solution method in [5]. Runolfsson [15], [16] used Donsker-Varadham-type large-derivations ideas to obtain a corresponding stochastic differential game for which the game payoff is an ergodic (expected average cost per unit time) criterion.

In [6], [13], [4], and [11] the risk-averse LEQR optimal control problem for a stochastic system with white Gaussian noises whose intensities depend on parameters was stated and solved using a value function, which is a viscosity solution to the dynamic programming equation (HJB). An advantage of risk-sensitive criteria is the robustness of the obtained solution with respect to noise level. Indeed, since the solution to the classical LQ problem is independent of noise level, it occurs to be too sensitive to parameter variations in noise intensity. On the other hand, the risk-sensitive problem assumes explicit presence of the small parameters in the criteria. This leads to a more robust solution, which correctly responds to parameter variations and results in close criterion values for both, large and small, parameter values.

This paper presents the explicit closed-form solutions to the optimal exponential-quadratic control problem for stochastic third degree polynomial systems including intensity parameters multiplying diffusion terms in the state equation. The optimal control algorithms are derived seeking quadratic value functions as solutions to the corresponding Hamilton-Jacobi-Bellman equations. Undefined parameters in the value functions are calculated through ordinary differential equations composed by collecting terms corresponding to each power of the state-dependent polynomial in the HJB equation. The closed-form risk-sensitive regulator equations are explicitly obtained in the control problem.

The performance of the obtained risk-sensitive regulator for stochastic third degree polynomial stochastic systems is verified in a numerical example against the conventional polynomial-quadratic regulator [2], through comparing the exponential-quadratic criteria values for both regulators. The simulation results reveal strong advantages in favor of the designed risk-sensitive algorithm in regard to the final criteria values uniformly for all considered values of the intensity parameters multiplying diffusion terms in the state equation. Tables of the criteria values and simulation graphs are included.

This paper is organized as follows. The optimal risk-sensitive control problem for stochastic systems of third degree with exponential-quadratic criterion is presented in Section 2 and Section 3 defines the optimal solution. A numerical application is presented in Section 4 and Section 5 presents the conclusions to this study.

II. OPTIMAL RISK-SENSITIVE STOCHASTIC CONTROL PROBLEM

Consider the following risk-sensitive control problem for a stochastic differential system:

$$dX = f(t,X(t),u(t))dt + \sqrt{\frac{\varepsilon}{2\gamma}}dW,$$

$$X(s) = x,$$

with respect to the exponential-quadratic cost criterion

$$I(s,X(t),u(t)) = \varepsilon logE_{s,x}exp\left\{\frac{1}{\varepsilon}\int_{s}^{T}L(t,X(t),u(t))dt + \psi(X_T)\right\},$$

(2)
where \( X(t) \) is the state at time \( t \), \( X(t) \in \mathbb{R}^n \), \( x \) is the initial state at time \( s \geq 0 \), \( f(t, X(t), u(t)) \) is a nonlinear function, which represents the nominal dynamics with control \( u(t) \) taking values in \( U \in \mathbb{R}^d \) and \( \{ W, F \} \) is an \( m \)-dimensional Brownian motion on the probability space \( (\Omega, F, P) \). The parameter \( \varepsilon \) is a measure of the risk-sensitivity and scales the diffusion term in (1), \( 0 \leq s \leq T < \infty \). \( T \) is a fixed terminal time, the parameter \( \gamma \) is a measure of the level of attenuation of the diffusion term in (1) (see [12] for more details), \( L(t, X(t), u(t)) \) is the quadratic running cost, and \( \psi(X_T) \) is the quadratic terminal cost. Define:

\[
A(s, X(t), u(t), \omega) = \int_s^T L(t, X(t), u(t)) dt + \psi(X_T),
\]

and

\[
J(s, X(t), u(t)) = E_{x,s} \exp\left[ -\frac{1}{\varepsilon} A(s, X(t), u(t), \omega) \right], \tag{3}
\]

so that

\[
I(s, X(t), u(t)) = \frac{\varepsilon \log I(s, X(t), u(t))}{\varepsilon} = E_{x,s} \exp\left[ -\frac{1}{\varepsilon} A(s, X(t), u(t), \omega) \right].
\]

Taking into account that the controller \( u(t) \) is minimizing, the following value function is considered:

\[
V(s, X(t)) = \inf_{u \in A_{x,s}} I(s, X(t), u(t)), \tag{4}
\]

where \( A_{x,s} \) is the set of progressively measurable controls with values in \( U \).

It is shown in [13] that under certain conditions, considering \( f(t, X(t), u(t)) \) a nonlinear function, \( V \) is a viscosity solution of the dynamical programming equation

\[
0 = V_x + \frac{\varepsilon}{2T^2} \sum_{x \in X} V_{xx} + \min_{u \in U} \{ f(x, X(t), u(t)) \nabla_x V + L(t, X(t), u(t)) + \frac{1}{2T^2} \nabla V + \}
\]

\[
V(X(t), T) = \psi(X_T).
\]

This paper shows that if \( f(t, X(t), u) = A(t) + A_1(t)X(t) + A_2(t)X(t)X(t)^T + A_3(t)X(t)X(t)^T X(t) + B(t)u(t), \) a viscosity solution \( V \) of the dynamical programming equation (5) can be explicitly found.

The optimal control problem is to find explicitly a viscosity solution \( V \) to the dynamic programming equation (5) when \( f(t, X(t), u) \) has the form third-degree polynomial, and to find the optimal control that minimizes the quadratic criterion \( J \) and the optimal trajectory \( x^* \), substituting \( u^* \) into the state equation. The conditions given in [13] for \( f, L, \psi, U \) are assumed throughout the paper. Those conditions remain true for a third degree polynomial function \( f(t, X(t), u) = A_1 + A_1(t)X(t) + A_2(t)X(t)X(t)^T + A_3(t)X(t)X(t)^T X(t) + B(t)u(t) \) in any interval \([0, T]\), where the corresponding solution of the equation (1) exists. Here, \( U \) is a compact subset of \( \mathbb{R}^d \), \( A(t) \in \mathbb{R}^d \), \( A_1(t) \in M_{n \times n} \), \( A_2(t) \) is a tensor of dimensions \( n \times n \times n \), \( A_3(t) \) is a tensor of dimensions \( n \times n \times n \times n \times n \).

As in [13], first consider the “cut off” problem, where the possibly unbounded functions \( f, L \) and \( \psi \) are replaced by bounded counterparts. We obtain analogous results for this “cut off” problem and then take a limit to obtain the desired result. The cut off functions, indexed by \( k \), are given by:

\[
f^k(t, x, u) = \begin{cases} f(t, x, u), & \text{if } |f(t, x, u)| \leq k \\ k, & \text{if } |f(t, x, u)| > k \end{cases}, \tag{5}
\]

In the same form for \( L(t, x, u, \psi), \)

\[
L^k(x) = \begin{cases} L(x), & \text{if } |L(x)| \leq k \\ k, & \text{if } |L(x)| > k \end{cases}, \tag{6}
\]

\[
\psi^k(x) = \begin{cases} \psi(x), & \text{if } |\psi(x)| \leq k \\ k, & \text{if } |\psi(x)| > k \end{cases}. \tag{7}
\]

The dynamic programming equation for \( \psi^k \) is given by:

\[
0 = V_x^k + \frac{\varepsilon}{2T^2} \sum_{x \in X} V_{xx} + \min_{u \in U} \{ f^k(t, X(t), u(t)) \nabla_x V + L^k(t, X(t), u(t)) + \frac{1}{2T^2} \nabla V + \}
\]

\[
V(X_T, T) = \psi^k(X_T). \tag{8}
\]

It is proved [13] that \( V^k \) is the unique, bounded, classical solution to (8), considering that \( f(t, X(t), u(t)) \) is nonlinear.

III. OPTIMAL RISK-SENSITIVE REGULATOR FOR THE THIRD DEGREE STATE

Taking into account that \( f(t, X(t), u) = A(t) + A_1(t)X(t) + A_2(t)X(t)X(t)^T + A_3(t)X(t)X(t)^T X(t) + B(t)u(t), \) the following state equation is obtained:

\[
\begin{align*}
\dot{X}(t) &= (A(t) + A_1(t)X(t) + A_2(t)X(t)X(t)^T + A_3(t)X(t)X(t)^T X(t) + B(t)u(t)) \n
V(X(t), T) &= \psi(X_T).
\end{align*}
\]

(9)

If \( L(t, X(t), u) = X(t)^T GX(t) + u(t)^T Ru(t), \) \( \psi(X(T)) = X_T^T \phi X, \) the exponential-quadratic cost criterion has the form:

\[
J(s, X(t), u) = \frac{\varepsilon}{2T^2} \int_s^T (X(t)^T GX(t) + 10) u(t)^T Ru(t) dt + |\psi|,
\]

where \( G, \phi \) are real symmetric positive semi-definite matrices, \( R \) is a real symmetric positive definite matrix.

**Theorem**

The solution to the stochastic control problem for the dynamical system (1) with the criterion (10) takes the form:

\[
\begin{align*}
\dot{P}(t) &= -2A_1(t)P(t) - 2A_2(t)P(t)X(t) + 2A_3(t)P(t)C(t) + \frac{1}{2}B(t)B(t)^T P(t) - \frac{1}{2}P(t)B(t)^T B(t) - 2I \\
C(t) &= -2P^{-1}(t)C(t)P(t) + A_0(t) - A_1(t)C(t) + \frac{1}{2}B(t)B(t)^T (P(t)C(t) - \frac{1}{2}P(t)C(t).
\end{align*}
\]

7047
where I is the identity matrix of dimensions \( n \times n \), \( L \) is as in (10). The optimal control law that minimizes the exponential-quadratic criterion (10) is given by:

\[
 u^*(t) = -\frac{1}{2}B^T(t)P(t)R^{-1}(X(t) - C(t)).
\]  

(12)

**Proof:** The value function is proposed:

\[
 V(t, X(t)) = \frac{1}{2}(X(t) - C(t))^T P(t)(X(t) - C(t)) + r(t) \quad (13)
\]

\( (C(t), P(t), r(t)) \) are functions of \( s \in [0, T], C(t) \in \mathbb{R}^n, P(t) \) is a matrix of dimension \( n \times n \) and \( r(t) \) is a scalar function) as a viscosity solution of the dynamic programming equation

\[
 0 = V_t + \frac{\epsilon}{2T^2} \sum V_{x_{ij}} + \min_{u \in \mathcal{U}} \{ (A(t) + A_1(t)X(t) + A_2(t)X(t)X(t)^T + A_3(t)X(t)X(t) + B(t)u(t))^T \nabla_V V + X(t)^T GX(t) + u^T(t)Ru(t) + \frac{\epsilon}{2T^2} \nabla V^T \nabla V \},
\]

\[
 V(X(t), T) = \psi(X(t)) = X^T(t)\phi(X(t)),
\]  

(14)

where \( V_t, V_{x_i} \) are the partial derivatives of \( V \) respect to \( t, x \), respectively, and \( \nabla V \) is the gradient of \( V \). Then, the partial derivatives of \( V \) are given by:

\[
 V_t = \frac{1}{2}(X(t) - C(t))^T \dot{P}(t)(X(t) - C(t)) + \dot{r}(t) - \frac{1}{2}C^T(t)P(t)(X(t) - C(t)) - \frac{1}{2}C(t)P(t)(X(t) - C(t)) - \frac{\epsilon}{2T^2} \sum P(t) + (A(t) + A_1(t)X(t) + A_2(t)X(t)X(t)^T + A_3(t)X(t)X(t) + B(t)(-\frac{1}{2}B(t)P(t)X(t) - C(t)))^T P(t)(X(t) - C(t)) + \frac{1}{2}(X(t) - C(t))^T P(t)(X(t) - C(t)) + \frac{\epsilon}{2T^2} (X(t) - C(t))^T P(t)(X(t) - C(t)).
\]  

(15)

Substituting (15) to the dynamic programming equation (14):

\[
 0 = \frac{1}{2}(X(t) - C(t))^T \dot{P}(t)(X(t) - C(t)) + \dot{r}(t) - \frac{1}{2}C^T(t)P(t)C(t) - \frac{1}{2}C(t)P(t)(X(t) - C(t)) + \frac{\epsilon}{2T^2} \sum P(t) + (A(t) + A_1(t)X(t) + A_2(t)X(t)X(t)^T + A_3(t)X(t)X(t) + B(t)(-\frac{1}{2}B(t)P(t)X(t) - C(t)))^T P(t)(X(t) - C(t)) + \frac{1}{2}(X(t) - C(t))^T P(t)(X(t) - C(t)) + \frac{\epsilon}{2T^2} (X(t) - C(t))^T P(t)(X(t) - C(t)).
\]  

(16)

Collecting the \( x^2, x^3 \) and \( x \) terms, the first equation in (11) is obtained. Similarly, collecting the \( x \) terms yields the second equation in (11). Note that the matrix \( P(t) \) is not symmetric, since the tensor \( A_3(t) \) is not symmetric with respect to its indices. Finally, collecting the terms independent of \( x \), the following equation is obtained

\[
 r(t) = \frac{3}{8} C(t)^T P(t)(B(t)B(t)^T)P(t)C(t) + \frac{1}{2} C(t)^T A(t) P(t) C(t) - C(t)^T A(t) P(t) C(t) - \frac{\epsilon}{2T^2} \sum P_{ij},
\]

where \( P_{ij} \) are the elements of the matrix \( P(t) \).

The optimal control law (12) that minimizes the quadratic criterion (10) is obtained from the condition:

\[
 \min_{u \in \mathcal{U}} \{ f^k(t, X(t), u(t))\nabla_V V + L^k(t, X(t), u(t)) + \frac{1}{2\sqrt{\gamma}} \nabla V^T \nabla V \}.
\]

IV. EXAMPLE: MONOAxaIAL SATELLITE

The risk-sensitive control equations for third degree polynomial systems will be applied to the problem of orientation of a monoaaxial satellite [17]. The description is as follows: a satellite rotates around a fixed axis without gravity. The rotation torques are produced by a system of mini-engines through a controlled explosion of gases in the opposite direction, as shown in Figure 3. The state equations for this model are given by:

\[
 X_1(t) = 0.5(1 + X_1^2(t))X_2(t) + \sqrt{\frac{\gamma}{2\sqrt{\gamma}}} dW_1(t),
\]  

(17)

\[
 X_2(t) = \frac{1}{\gamma} u(t) + \sqrt{\frac{\gamma}{2\sqrt{\gamma}}} dW_2(t),
\]

where \( X_1(t) \) represents the orientation angle of the satellite, measured with respect of a secondary axis which does not coincide with the principal one. \( X_2(t) \) is the angular velocity with respect to the principal axis. The control variable \( u \) represents the applied torque. \( J \) denotes the moment of inertia of the satellite disk, which is given by \( J = \frac{1}{2}mr^2 \), \( m \) is the mass of the satellite and \( r \) is the disk radius. In this case, \( m = 1, r = 1 \). The optimal control problem is to obtain the optimal control law \( u^*(t) \), that minimizes the energy produced by the torques, to reach a given orientation angle as defined by the criterion (10).

Applying the equations (11), (12) to the system (17), the following optimal risk-sensitive control equations are obtained:

\[
P_{11} = -P_{21} - \frac{1}{\gamma} (P_{11}^2 + P_{21}^2) - 2,
\]

(18)

\[
P_{12} = -P_{22} - \frac{1}{\gamma} (P_{11}P_{12} + P_{12}P_{21}) + 0.5P_{21}X_2X_1 + 0.5P_{21}C_2X_1,
\]

\[
P_{21} = \left( \frac{1}{2\sqrt{\gamma}} \right) (P_{21}P_{11} + P_{22}P_{21}) - \frac{1}{\gamma} (P_{21}P_{11} + P_{23}P_{21}),
\]

\[
P_{22} = \left( \frac{1}{2\sqrt{\gamma}} \right) (P_{21}P_{12} + P_{22}P_{22}) - \frac{1}{\gamma} (P_{12}^2 + P_{22}^2) - 2,
\]

\[
C_1 = \frac{-2}{P_{22}P_{11} - P_{22}P_{21}} \left[ P_{22}C_1 (P_{21} - \frac{1}{\gamma} (P_{11}^2 + P_{21}^2) - 2) + P_{22}C_2 (-P_{22} - \frac{1}{\gamma} (P_{11}^2 + P_{21}^2) - 2) + P_{21}X_2X_1 + 0.5P_{21}C_2X_1 - P_{21}C_1 \left( \frac{1}{2\sqrt{\gamma}} (P_{11}P_{12} + P_{12}P_{21}) \right) \right] - \frac{1}{\gamma} (P_{11}P_{12} + P_{21}P_{22}) - 0.5P_{21}C_2X_1 - P_{21}C_1 \left( \frac{1}{2\sqrt{\gamma}} (P_{11}P_{12} + P_{12}P_{21}) \right).
\]

(19)
The corresponding equations are given by:

\[ \dot{C}_2 = \frac{-2}{P_{12}P_{21}P_{11}} [-P_{21}C_1(-P_{21} - \frac{1}{\rho^2}(P_{11}^2 + P_{21}^2) - 2) - P_{22}C_2(-P_{22} - \frac{1}{\rho^2}(P_{11}P_{12} + P_{21}P_{22}) + 0.5P_{21}X_1X_1 + 0.5P_{22}C_2X_1) + P_{11}C_1(\frac{1}{\rho^2}(P_{11}P_{21} + P_{21}P_{22}))]
\]

where \( G, R \) are the identity matrices \( I_{2\times2} \), with the terminal conditions \( C_1(0.4) = 0, C_2(0.4) = 0, P_{11}(0.4) = 1, P_{22}(0.4) = P_{21}(0.4) = 0, P_{22}(0.4) = 1 \).

The optimal control law is given by:

\[
\begin{align*}
  u_2(t) & = \frac{-1}{2\rho^2}(P_{21}(X_1 - C_1) + P_{22}(X_2 - C_2)), \\
  u_1(t) & = 0.
\end{align*}
\]

The optimal values of the satellite orientation angle and angular velocity, respectively, \( X_1(t) \), \( X_2(t) \) are obtained by substituting the optimal control law (19) into (17):

\[
\begin{align*}
  X_1(t) & = 0.5(1 + X_1^2(t))X_2(t) + \sqrt{\frac{\epsilon}{2\rho^2}}dW_1(t), \\
  X_2(t) & = \frac{-1}{2\rho^2}(P_{21}(X_1(t) - C_1) + P_{22}(X_2(t) - C_2)) + \sqrt{\frac{\epsilon}{2\rho^2}}dW_2(t).
\end{align*}
\]

The initial conditions for \( X \) are: \( X_1(0) = 0.2, X_2(0) = 7 \).

The system formed by the equations (19), and (17) is simulated using Simulink in MatLab7 and applying Monte Carlo method. The performance of the designed risk-sensitive regulator is compared versus the third degree polynomial regulator [2], applied to the system (17), that is optimal with respect to the conventional quadratic criterion. The corresponding equations are given by:

\[
\begin{align*}
  \dot{Q}_{11} & = 1 - \frac{1}{\rho^2}Q_{12}Q_{21}, \\
  \dot{Q}_{12} & = -(0.5 + 0.5X_2X_1)Q_{11} - \frac{1}{\rho^2}Q_{12}Q_{22}, \\
  \dot{Q}_{21} & = -(0.5 + 1.5X_2X_1)Q_{11} - \frac{1}{\rho^2}Q_{22}Q_{21}, \\
  \dot{Q}_{22} & = 1 - (0.5 + 1.5X_2X_1)Q_{12} - (0.5 + 0.5X_2X_1)Q_{21} - \frac{1}{\rho^2}Q_{22}^2.
\end{align*}
\]

with terminal conditions \( Q_{11}(0.4) = -2, Q_{12}(0.4) = 0, Q_{21}(0.4) = 0, Q_{22}(0.4) = -2 \). The control law is given by:

\[
\begin{align*}
  u_2^*(t) & = 2(Q_{21}X_1(t) + q_{22}X_2(t)), \\
  u_1^*(t) & = 0,
\end{align*}
\]

where the corresponding satellite orientation angle \( X_1(t) \) and angular velocity \( X_2(t) \) satisfy the equations:

\[
\begin{align*}
  X_1(t) & = 0.5(1 + X_1^2(t))X_2(t) + \sqrt{\frac{\epsilon}{2\rho^2}}dW_1(t), \\
  X_2(t) & = \frac{1}{2}(Q_{21}X_1 + q_{22}X_2) + \sqrt{\frac{\epsilon}{2\rho^2}}dW_2(t).
\end{align*}
\]

Figures 1 and 2 show the graphs of the states \( X_1(t), X_2(t) \), the optimal control \( u^*(t) \), and the exponential-quadratic criterion \( J \) for both applied algorithms.

\section*{V. Conclusions}

This paper presents the optimal solution to the risk-sensitive control problem for third degree polynomial system in the presence of Gaussian white noise scaled by a parameter \( \epsilon \) multiplying the diffusion term. An exponential-quadratic cost function is used. The closed-form optimal control algorithm is obtained assigning a quadratic value function as a solution to the corresponding Hamilton-Jacobi-Bellman equation. Numerical simulations are conducted to make a comparison of the designed third degree risk-sensitive regulator versus the conventional quadratic criterion, considering the exponential-quadratic criterion values at the final time. The simulation results show definite advantage in favor of the risk-sensitive third degree optimal control algorithms for all parameter values \( \epsilon \) uniformly.

\section*{REFERENCES}

Fig. 1. Graphs of the state, control, and criterion corresponding to the risk-sensitive regulator for $\varepsilon = 1000$.

Fig. 2. Graphs of the state, control, and criterion corresponding to the third degree polynomial regulator for $\varepsilon = 1000$.


Fig. 3. Monoaxial satellite.