Nonlinear Second Cost Cumulant Control using Hamilton-Jacobi-Bellman Equation and Neural Network Approximation

Bei Kang, Student Member, IEEE, and Chang-Hee Won, Member, IEEE

Abstract—Cost cumulant control is an optimal control method applied to the stochastic systems. Unlike the deterministic systems, the cost function in a stochastic system is a random variable. We use statistical control to optimize the distribution of the cost function via cost cumulants. We analyze the first and second cost cumulant optimization problems for nonlinear stochastic systems. The Hamilton-Jacobi-Bellman (HJB) equations are derived with respect to each cumulant minimization case. Then a neural network approximation method is used to numerically solve the HJB equations, which in turn is used to find the optimal controller. Two examples including a nonlinear system and a linearized satellite attitude control system are presented. The first and second cumulant optimal controllers are found. Then the performance of the optimal controllers for the first and the second cumulants are compared through the examples.

I. INTRODUCTION

In a stochastic system, the system is coupled with noise or disturbances which are random variables. The optimal control of a stochastic system is to minimize a certain cost function which is associated with the system dynamic function. Because the system is stochastic, the cost function is viewed as a random variable. In statistical optimal control, we shape the distribution of the cost function via minimizing the cost cumulants.

Cost cumulant control was first studied by Sain [1]. Later Sain and Liberty investigated cost cumulant control for the LQG systems in [2]. The second cost cumulant control for a full-state-feedback system was given in [3]. Won extended the research to a general n-th cost cumulant control problem and derived the Hamilton-Jacobi-Bellman equations for each cumulant [4]. The significance of the cost cumulant control is that it flexibly shapes of the distribution for the given cost function by controlling different cumulants which carry different physical meanings. For example, the first cost cumulant represents the mean of the cost function. Stochastic control methods such as LQG control focus on the first cumulant control, i.e., to minimize the mean of the cost function. The second cumulant is the variance, which is the variation of the distribution of the cost function. The third cumulant, skewness, measures the departure from symmetry of the cost distribution. And the fourth cumulant, kurtosis, represents the flatness of the distribution. The importance of the cumulant decreases when the order increases. Therefore, in this paper, we focus on only the first two cumulants, i.e., we find the optimal controllers which minimize the mean and variance of a cost function.

Statistical optimal control problems can be formulated using Hamilton-Jacobi-Bellman (HJB) equations. It is well known that HJB equations are extremely difficult to solve, both numerically and analytically. Although there are few linear systems such as LQG that can be solved analytically, the general nonlinear systems and the higher order cumulant problems are extremely challenging. Even for a linear system, the higher order cumulant problem will yield a nonlinear controller [4]. For a full state feedback control affine system, which we will investigate in this paper, there is a method using series expansion to approximate the value function in HJB equation. Al’brekht in [5] first introduced the basic idea of this method where he used a power series expansion to approximate the value function in infinite time horizon, where the system function and the cost are deterministic functions. The approximated value function and the optimal controller are determined by finding the coefficients of the series expansion. That research was, however, limited to the infinite time horizon, where the coefficients are time invariant. The method of determining the coefficients of the series expansion is discussed in several papers. Y. Chen et al. in [6] proposed a method to convert HJB equation to a Riccati equation and a series of algebraic equations, which deals with the infinite time horizon deterministic systems. R. Beard in [7] proposed a Galerkin approximation method to solve the generalized Hamilton-Jacobi-Bellman (GHJB) equation for the deterministic system. T. Cheng et al. used the neural network method to solve the fixed-final time deterministic system, which extend the study to the finite time horizon [8]. P. V. Medagam et al. proposed a radial basis function neural network method for the output feedback system [9]. The advantage of the above mentioned method is that the system to be controlled can be nonlinear, with non-quadratic cost function. However, all of above references focus on the deterministic system. In this paper, we extend the neural network approximation method used in [8] to the nonlinear stochastic system to solve the HJB equations for the first and second minimal cost cumulants.

In Section II, we formulate the cost cumulant control problems and derive the Hamilton-Jacobi-Bellman equations with respect to the first and second cumulant. The series expansion approximation and neural network method for the HJB equation is discussed in Section III. In Section IV, two examples are analyzed, and the optimal controllers for the first and second cost cumulant minimization are calculated and compared. Conclusions are then given in the last section.
II. COST CUMULANT CONTROL PROBLEM FORMULATION AND HJB EQUATION DERIVATION

Consider a stochastic system equation with the following dynamics.
\[ dx(t) = \left[ g\left(x(t)\right) + B\left(x(t)\right)u(t) \right] dt + \sigma(t,x(t))dw(t), \]
where \( t \in [t_0, t_f] \), \( x(t) \in \mathbb{R}^n \), \( u \in U \subset \mathbb{R}^m \), \( dw(t) \) is a \( d \)-dimensional Gaussian random process with zero mean and covariance of \( W(t)dt \). The system control is given as \( u(t) = k(t, x(t)) \), \( t \in T \).

We assume system (1) is a control affine system, which means that the system is linear with respect to the control.

The system cost function is given in the following form,
\[ J(t, x(t); k) = \psi(x(t)) \]
\[ + \int_t^{t_f} \left[ l(s, x(s)) + k'(s, x(s))Rk(s, x(s)) \right] ds, \]
where \( \psi(x(t)) \) is the terminal cost, \( l(s, x(s)) \) is a positive definite function, \( R \) is the positive definite matrix. In order to derive the \( n \)-th cumulant equation, we make the following assumptions.

- \( L(t, x, k) = l(s, x(s)) + k'(s, x(s))Rk(s, x(s)) \) is continuously differentiable and satisfies the polynomial growth condition.
- \( g \) and \( \sigma \) satisfy the Lipschitz condition and the linear growth condition.

We look for an optimal controller \( u^*(t) = k^*(t, x) \), to minimize the first and second cost cumulant of the cost function (3) on \( t \in [t_0, t_f] \).

To study the cumulant, we first introduce the moments, which is defined as \( M_s(t, x, k) = E\left\{ J(t, x, k) \mid x(t) = x \right\} \). The \( n \)-th cumulant and the \( n \)-th moment are related by the following equation [10].
\[ \exp \left( \sum_{i=1}^{n} V_i \frac{t^i}{i!} \right) = \sum_{i=0}^{n} M_i \frac{t^i}{i!}. \]

The following definitions are necessary for the HJB equation derivations.

**Definition 2.1:** A function \( M_s: \mathbb{Q}_0 \rightarrow \mathbb{R}^+ \) is an admissible cost mean if there exists an admissible control law \( k \) such that \( M_s(t, x, k) = E_a\left\{ J(t, x, k, x(t, k)) \right\} \).

**Definition 2.2:** The admissible cost mean \( M_s \) defines a class of control laws \( K_{u_s} \) such that for each \( k \in K_{u_s}, k \) and \( M_s \) satisfy Definition 2.1.

Because we want to find the minimal cost cumulant, which are viewed as the value function of the statistical control problem, it is necessary to find the HJB equations associated with each cost cumulant, and then solve the HJB equation. Won in [4] derived the first and second cost cumulant HJB equations as given in the following lemma and theorems.

**Lemma 2.1:** The minimal first cost cumulant \( V_1(t, x) \) of the cost function (3) equals to the first moment \( M_1(t, x) \), which satisfies the following equation.
\[ \min_{k} O(k) [V_1(t, x)] + L(t, x, k) = 0, \]
where \( O(k) = \frac{\partial}{\partial t} + \left( g(t, x, k), \frac{\partial}{\partial x} \right) + \frac{1}{2} tr \left( \sigma W \sigma' \frac{\partial^2}{\partial x^2} \right) \) is the backward evolution operator.

**Theorem 2.1:** (Second cumulant HJB equation) Let \( M_s(t, x) \in C_{\bar{\rho}}^{1,2}(\mathbb{Q}_0) \) be an admissible mean cost function, and let \( M_s(t, x) \) induce a nonempty class \( K_{u_s} \) of admissible control laws. Assume the existence of an optimal control law \( k = k_s^* \) and an optimum value function \( V_s^*(t, x) \) is given as \( C_{\bar{\rho}}^{1,2}(\mathbb{Q}_0) \).

Then the minimal second cumulant (variance) function \( V_s^*(t, x) \) satisfies the following HJB equation.
\[ \min_{k} O(k) [V_s^*(t, x)] + \left\| \frac{\partial V_s^*(t, x)}{\partial x} \right\|_{\mathbb{Q}_0} = 0, \]
for \( (t, x) \in \mathbb{Q}_0 \), with the terminal condition \( V_s^*(t_f, x) = 0 \).

**Proof:** See [4].

**Remark:** It should be noted that the second cumulant HJB equation (6) has the first cumulant term \( V_1(t, x) \). When we look for the minimal variance \( V_s^*(t, x) \), the first cumulant \( V_1(t, x) \) is not necessarily the minimum, but is pre-specified in some certain domain. By substituting the optimal controller \( k_s^* \) back to equations (6) and (5), we will find and \( V_s^*(t, x) \) at the same time by solving equation (5) and (6) together.

We set our focus on the minimal second cumulant control problems, which requires us to solve HJB equation (5) together with equation (6). Because \( V_1^*_s(t, x) \) is pre-specified, which satisfies the HJB equations (5) intrinsically, we want to find \( V_s^*(t, x) \) to satisfy equation (6).

Using the method introduced in [3], the optimal controller \( k^*_s(t, x) \) for equation (6), is obtained as
\[ k^*_s(t, x) = - \frac{1}{2} R^{-1} B'(x) \left[ \frac{\partial V_s^*(t, x)}{\partial x} + \gamma_s \frac{\partial V_s^*(t, x)}{\partial x} \right], \]
where the parameter \( \gamma_s \) is defined as a time varying Lagrange multiplier. Now, substitute \( k^*_s(t, x) \) back to equations (5) and (6), we obtain the following equations.
\[ \frac{\partial V_s^*(t, x)}{\partial t} + g'(x) \left( \frac{\partial V_s^*(t, x)}{\partial x} \right) + l + \frac{1}{2} tr \left( \sigma W \sigma' \frac{\partial^2 V_s^*(t, x)}{\partial x^2} \right) \]
For a stochastic system, when \( \gamma_z(t) \) is defined, we can find \( V_1(t,x) \) and \( V_2(t,x) \) by solving equations (8) and (9) together, and then find the minimal second cumulant controller. The method we used to solve the HJB equations (8) and (9) will be discussed in next section.

### III. HJB EQUATION APPROXIMATION AND NEURAL NETWORK METHOD

HJB equations (8) and (9) are coupled with each other and are very difficult to solve directly. Sain et al. proposed an analytical method to solve the first two cumulants HJB equation for linear systems [3]. However, for nonlinear systems, there are no effective analytical methods. In this section, we use a polynomial series expansion to approximate the value functions in equations (8) and (9), and apply neural network method to find the solution of the HJB equations.

Neural network approximation method for HJB equation is discussed in [8], where Cheng et al. developed a method to find the neural network coefficients of the series expansion which approximates the value function, under the assumption that the value function is uniformly continuous on a compact set. He also proved the convergence of the neural network method for the deterministic systems.

Inspired by [8], we use the neural network idea to analyze the cost cumulant control problems. However, for the HJB equations given in the previous section, there are Gaussian noise term and the second order derivatives in the HJB equations, which make the system analysis different from the deterministic systems. Therefore, a modified neural network method is necessary to treat the problem.

The essence of the neural network method is shown in Fig. 1. A number of the neural network input functions are multiplied by the corresponding weights and then summed up to produce an output function, which will be the approximated value function and is the solution of the HJB equation. In this paper, we use a polynomial series expansion \( \overline{d}_x(x) = \{ \delta_1, \delta_2, \ldots, \delta_L \} \) as the neural network input functions. By determining the weights of the series expansion, we find the output functions, which are the approximations to the value functions \( V_i(t,x) \) and \( V_i^*(t,x) \).

![Fig. 1 Block diagram a neural network unit.](image-url)

The selection of the delta functions depends on the properties of the cost function and the approximating precision and is usually determined on an empirical basis. The weights will be found by solving the approximated HJB equations with the following procedures. First we use \( \delta_i'(x,t) = \overline{\delta}_i(x,t) \) to approximate the first cumulant value function \( V_i(t,x) \), and use \( \delta_i^*(x,t) = \overline{\delta}_i(x,t) \) to approximate \( V_i^*(t,x) \), where the vectors \( \overline{\delta}_i(x,t) \), \( \overline{\delta}_i(x,t) \) are the time varying coefficients of the polynomial series expansion \( \delta_i(x,t) \) and \( \delta_i^*(x,t) \) respectively. In this paper, we choose \( \delta_i(x,t) = \overline{\delta}_i(x,t) = \overline{\delta}_i(x) \). The subscript \( L \) represents the order of the polynomial series. The larger \( L \) leads to the closer approximation to the value functions. Then we substitute \( V_i'(t,x) = \overline{\delta}_i(x,t) \delta_i'(x,t) \) and \( V_i^*(t,x) = \overline{\delta}_i(x,t) \delta_i^*(x,t) \) back to (8) and (9), we obtain equations (10) and (11) as follows,

\[
\overline{\delta}_i(x,t) \delta_i'(x,t) + g'(x) \left( \nabla \delta_i'(x,t) \overline{\delta}_i(x,t) \right) + \frac{1}{2} \left( \nabla g'(x) \overline{\delta}_i(x,t) \delta_i'(x,t) \right) = 0, 
\]

\[
\overline{\delta}_i(x,t) \delta_i^*(x,t) + g'(x) \left( \nabla \delta_i^*(x,t) \overline{\delta}_i(x,t) \right) + \frac{1}{2} \left( \nabla g'(x) \overline{\delta}_i(x,t) \delta_i^*(x,t) \right) = 0, 
\]

where the vectors \( \overline{\delta}_i(x,t) \), \( \overline{\delta}_i(x,t) \) are the approximations to the first cumulant value function \( V_i(t,x) \) and \( V_i^*(t,x) \), respectively.

\( \overline{\delta}_i(x,t) \delta_i'(x,t) \) and \( \overline{\delta}_i(x,t) \delta_i^*(x,t) \) are the time varying coefficients of the polynomial series expansion \( \delta_i(x,t) \) and \( \delta_i^*(x,t) \) respectively. In this paper, we choose \( \delta_i(x,t) = \overline{\delta}_i(x,t) = \overline{\delta}_i(x) \). The subscript \( L \) represents the order of the polynomial series. The larger \( L \) leads to the closer approximation to the value functions. Then we substitute \( V_i'(t,x) = \overline{\delta}_i(x,t) \delta_i'(x,t) \) and \( V_i^*(t,x) = \overline{\delta}_i(x,t) \delta_i^*(x,t) \) back to (8) and (9), we obtain equations (10) and (11) as follows,

\[
\overline{\delta}_i'(x,t) + g'(x) \left( \nabla \delta_i'(x,t) \overline{\delta}_i(x,t) \right) + \frac{1}{2} \left( \nabla g'(x) \overline{\delta}_i(x,t) \delta_i'(x,t) \right) = 0, 
\]

where the vectors \( \overline{\delta}_i(x,t) \), \( \overline{\delta}_i(x,t) \) are the approximations to the first cumulant value function \( V_i(t,x) \) and \( V_i^*(t,x) \), respectively.

\( \overline{\delta}_i'(x,t) + g'(x) \left( \nabla \delta_i^*(x,t) \overline{\delta}_i(x,t) \right) + \frac{1}{2} \left( \nabla g'(x) \overline{\delta}_i(x,t) \delta_i^*(x,t) \right) = 0, 
\]
\[ + \frac{1}{2} \text{tr} \left( \sigma W \sigma' \frac{\partial \left( \nabla \delta_i(x)^T \nabla \bar{w}(t) \right)}{\partial x} \right) \]

\[ + \left( \nabla \delta_i(x)^T \nabla \bar{w}(t) \right)^T \sigma W \sigma' \left( \nabla \delta_i(x)^T \nabla \bar{w}(t) \right) = e_{z_i}(t, x). \] (11)

The values of \( e_{z_i}(t, x) \) and \( e_{z_i}(t, x) \) on the right hand side of the equations (10) to (11) are due to the approximation errors, which in other words, represents the difference between \( V_i(t, x) \) and \( V_a(t, x) \). We want to minimize this approximation error by assigning appropriate \( \bar{w}(t) \) and \( \bar{v}(t) \), such that \( V_{i\epsilon}(t, x) \) and \( V_{z\epsilon}(t, x) \) achieve the best approximation to \( V_i(t, x) \) and \( V_z(t, x) \).

The method of weighted residuals [11] is used in this paper to find the coefficients which minimize the approximation error. The method of weighted residuals is discussed in reference [7] and [8] in detail. We apply the method of weighted residuals to equations (10) and (11) and obtain the following ordinary differential equations.

\[ \ddot{\bar{w}}(t) = -\left( \delta_i(x) \delta_i(x)^T \right)^T \left( \nabla \delta_i(x)^T g(x) \nabla \delta_i(x) \right) \bar{w}(t) \]

\[ + \frac{1}{4} \left( \delta_i(x) \delta_i(x)^T \right)^T A \cdot \bar{w}(t) - \frac{\gamma^2}{4} \left( \delta_i(x) \delta_i(x)^T \right)^T B \cdot \bar{v}(t) \]

\[ - \left( \delta_i(x) \delta_i(x)^T \right)^T \delta_i(x) \delta_i(x)^T \bar{v}(t), \] (12)

and

\[ \ddot{\bar{v}}(t) = -\left( \delta_i(x) \delta_i(x)^T \right)^T \left( \nabla \delta_i(x)^T g(x) \nabla \delta_i(x) \right) \bar{v}(t) \]

\[ + \frac{1}{2} \left( \delta_i(x) \delta_i(x)^T \right)^T A \cdot \bar{v}(t) - \frac{\gamma^2}{2} \left( \delta_i(x) \delta_i(x)^T \right)^T B \cdot \bar{v}(t) \]

\[ - \frac{1}{2} \left( \delta_i(x) \delta_i(x)^T \right)^T D - \left( \delta_i(x) \delta_i(x)^T \right)^T E \cdot \bar{v}(t), \] (13)

where

\[ A = \sum_{i=1}^L w_i \left( \nabla \delta_i(x)^T B(t, x) R^2 B(t, x) \frac{\partial \delta_i(x)}{\partial x} \bar{w}(t) \right), \]

\[ B = \sum_{i=1}^L w_i \left( \nabla \delta_i(x)^T B(t, x) R^2 B(t, x) \frac{\partial \delta_i(x)}{\partial x} \bar{v}(t) \right), \]

\[ C = \text{tr} \left( \sigma W \sigma' \left( \nabla \delta_i(x)^T \bar{w}(t) \right) \right), \]

\[ D = \text{tr} \left( \sigma W \sigma' \left( \nabla \delta_i(x)^T \bar{v}(t) \right) \right), \]

\[ E = \sum_{i=1}^L w_i \left( \nabla \delta_i(x)^T \sigma W \sigma' \frac{\partial \delta_i(x)}{\partial x} \bar{w}(t) \right), \]

\( \langle \cdot, \cdot \rangle \) denotes the inner products, \( \Omega \) is the region where an admissible \( k'(t, x) \) exists.

The neural network method thus converts HJB equations (8) and (9) into the ordinary differential equations (12) and (13) with respect to the weights \( \bar{w}(t) \) and \( \bar{v}(t) \). Once we know the terminal condition, which is given as zero in Theorem 2.1, we are able to find \( \bar{w}(t) \) and \( \bar{v}(t) \) by integrating equations (12) and (13) backwards. Then we find the approximated value function \( V_i(t, x) \) and \( V_{z\epsilon}(t, x) \) and consequently, the first and second cumulant optimal controller.

In next section, we illustrate the procedure of solving the first and second cumulant minimization problem and neural network approximation method through some examples and compare the performance of the different cumulant controllers.

IV. EXAMPLES AND THE PERFORMANCE OF THE COST CUMULANT CONTROL

In this section, we will use the neural network method discussed in previous section to calculate the first and second cumulant optimal control for the two applications. The applications include one nonlinear single state system and one linearize satellite attitude control system, which demonstrate that neural network approximation is applicable to statistical control for both the nonlinear and linear systems. Example 1: Nonlinear system example.

Assuming that we have a nonlinear one dimensional system with the following dynamics,

\[ dx(t) = \sin(x)dt + u(t)dt + \sigma dw(t), \]

where \( \sigma W \sigma' = 0.09 \), and the cost function is given as

\[ J(t, x, u) = \int_0^\tau (x^2 + u^2) ds. \]

The terminal condition is assumed to be zero. Here \( L = x^2 + u^2 \) satisfies differentiable and polynomial growth condition, and \( \sin(x) \) and \( \sigma \) satisfy the Lipschitz condition and the linear growth condition. Thus we use the method described in Section II to analyze system (14) and derive the HJB equations for the first and second cumulant control using equation (8) and (9).

We then use the neural network method to approximate and solve the HJB equations. To choose the polynomial series expansion, because that the value functions must be positive definite, we use the following series expansion,

\[ \delta_i(x) = \{1, x^1, x^2, x^3, x^4, x^5, x^6, x^7\}. \]

Unlike \( \delta_i(x) \) chosen in deterministic systems, \( \delta_i(x) \) in stochastic system contains a constant term "1", which represents a state independent term in value functions.

The coefficients of the series expansions for the first and second cumulant HJB approximations are defined as \( \bar{w}(t) = \{w_1, w_2, w_3, w_4, w_5, w_6, w_7\} \) and \( \bar{v}(t) = \{v_1, v_2, v_3, v_4, v_5\} \) respectively.
When computing the minimal second cumulant control, we assume \( \gamma'_i(t) \) to be a constant such that \( \gamma'_i(t) = 1 \). The choice of \( \gamma'_i(t) \) manages the proportion of the contribution from the second cumulant to the total system cost, which is usually determined empirically depending on the systems. Applying the neural method described in the previous section, we then determine the value of \( v^L_1(t) \) and \( v^L_2(t) \). Due to the limited space, we will not give the detailed calculation procedures. After we find \( v^L_1(t, x) \), and \( v^L_2(t, x) \), we are able to calculate the optimal controller for each cost cumulant.

The state trajectories for the first and second cumulant control are plotted in Fig. 2.

Fig. 2. State trajectories of the first and second cumulant control.

It can be seen from Fig. 2, that the state \( x \) approaches the origin (equilibrium) faster under second cumulant control than first cumulant control.

**Example 2: Satellite attitude control.**

Consider a satellite attitude control system abstracted from the LEO satellite KOMPSAT [12, 13], where the satellite attitude is controlled by four reaction wheels and four thrusters. We assume that the gravity torque is small enough to be ignored. The solar radiation torque, magnetic torque, and aerodynamic torque are also ignored in this paper. We assume that the gravity torque is small enough to be ignored. The neural method described in the previous section, we then determine the value of \( v^L_1(t) \) and \( v^L_2(t) \). Due to the limited space, we will not give the detailed calculation procedures. After we find \( v^L_1(t, x) \), and \( v^L_2(t, x) \), we are able to calculate the optimal controller for each cost cumulant.

The state trajectories for the first and second cumulant control are plotted in Fig. 2.

\[
\dot{x} = A x + B u + E \xi,
\]

where matrices \( A \) and \( B \) are defined in [13], \( \xi \) is a Brownian motion with variance matrix \( W = 0.01 I_{5 \times 5} \). The terminal cost is assumed as zero. And the cost function is given in quadratic form

\[
J(t, x(t); k) = \int_t^k \left[ x(s) Q x(s) + k'(s, x(s)) R k(s, x(s)) \right] ds.
\]

For the first cumulant control, because the system is linearized, we use LQG method to find the optimal controller. For the second cumulant, we follow the same procedure as in Example 1. We derive the HJB equations for the first and second cost cumulant of the satellite attitude model, and then use the neural network method to find the corresponding minimal cumulant controllers.

First, we assume \( \gamma'_i(t) = 60 \), and solve for the value functions for the second cumulant optimal controller. The state trajectories of each cumulant for the satellite attitude control model are then calculated.

Selected angular velocities of the satellite, the reaction wheel angular velocities dynamics are plotted in Fig. 3 and Fig. 4 respectively.

\[
\dot{\theta} = \frac{1}{\cos \psi} \begin{bmatrix}
\cos \psi & -\cos \phi \sin \psi & \sin \phi \sin \psi \\
0 & \cos \phi & -\sin \phi \\
\sin \phi \cos \psi & \cos \phi \cos \psi & 0
\end{bmatrix} \theta + n \tag{18}
\]

The satellite attitude control system model is given by combining equations (15) to (18).

In this example, we will study how to move and keep the satellite at the preset orientation, which, represented by the Euler angle, is \( \{ \phi = 0^\circ; \theta = 0^\circ; \psi = 0^\circ \} \) with the presence of the Gaussian random noise.

Because we focus on the attitude control actions around the equilibrium point \( \{ \phi = 0^\circ; \theta = 0^\circ; \psi = 0^\circ \} \), we use a linearized model deriving from equations (15) to (18) to simplify the calculations.

Let \( \bar{x} = [\phi \; \theta \; \psi \; \omega_x \; \omega_y \; \omega_z \; \Omega_1 \; \Omega_2 \; \Omega_3 \; \Omega_4]^T \), where \( \phi, \theta \) and \( \psi \) are Euler angles of the satellite; \( \omega_x, \omega_y, \) and \( \omega_z \) are angular velocities of the satellite; \( \Omega_1, \Omega_2, \Omega_3, \) and \( \Omega_4 \) are the reaction wheel angular velocities. Then we have the following system dynamics

\[
\bar{x} = A \bar{x} + B u + E \xi,
\]

where matrices \( A \) and \( B \) are defined in [13], \( \xi \) is a Brownian motion with variance matrix \( W = 0.01 I_{15 \times 5} \). The terminal cost is assumed as zero. And the cost function is given in quadratic form

\[
J(t, x(t); k) = \int_t^k \left[ x(s) Q x(s) + k'(s, x(s)) R k(s, x(s)) \right] ds.
\]

For the first cumulant control, because the system is linearized, we use LQG method to find the optimal controller. For the second cumulant, we follow the same procedure as in Example 1. We derive the HJB equations for the first and second cost cumulant of the satellite attitude model, and then use the neural network method to find the corresponding minimal cumulant controllers.

First, we assume \( \gamma'_1(t) = 60 \), and solve for the value functions for the second cumulant optimal controller. The state trajectories of each cumulant for the satellite attitude control model are then calculated.

Selected angular velocities of the satellite, the reaction wheel angular velocities dynamics are plotted in Fig. 3 and Fig. 4 respectively.
cumulant controllers have the better performance compared to the first cumulant controllers and thus demonstrate the effectiveness of the second cumulant control for the stochastic systems. Furthermore, the results indicate the feasibility of using the proposed neural network method to solve the HJB equation for the nonlinear stochastic systems.

REFERENCES