Keywords: Cost Cumulant Game Control, Nash Game, Neural Networks, Output-feedback, Statistical Control.

Abstract: This paper studies a finite horizon output-feedback game control problem where two players seek to optimize their system performance by shaping the distribution of their cost function through cost cumulants. We consider a two-player second cumulant nonzero-sum Nash game for a partially-observed linear system with quadratic cost function. We derive the near-optimal players strategy for the second cost cumulant function by solving the Hamilton-Jacobi-Bellman (HJB) equation. The results of the proposed approach are demonstrated by solving a numerical example.

1 INTRODUCTION

Game theory is the study of tactical interactions involving conflicts and cooperations among multiple decision makers called players with applications in diverse disciplines such as management, communication networks, electric power systems and control (Zhu et al., 2012), (Charilas and Panagopoulos, 2010), (Cruz et al., 2002). Stochastic differential game results from strategic interactions among players in a random dynamic system (Basar, 1999). In stochastic optimal control, there is a player and cost function to be optimized while in stochastic differential games, there are multiple players and separate cost function to be optimized by each player.

In most practical control engineering applications, not all the states are measurable. The system model may consists of unknown disturbances usually expressed as process noise while the inaccuracies in measurement are usually expressed as measurement noise. An approach to account for the unmeasurable states is to estimate those states using an estimator before utilizing the states in a controller in a feedback control system. This approach is part of a generalized method to analyzing linear stochastic systems by applying the concept of certainty equivalence principle (Van De Water and Willems, 1981) or related separation principle (Wonham, 1968). Bensoussan et al. (Bensoussan and Schuppen, 1985) investigated the stochastic optimal control problem for partially-observed system with exponential cost criterion and proved that separation theorem does not hold for such scenario. (Zheng, 1989) investigated both optimal and suboptimal approach to output feedback control for a linear system with quadratic cost function while the solvability of the necessary and sufficient conditions for the existence of a stabilizing output feedback solution for a continuous-time linear systems was studied in (Geromel et al., 1998). Aberkane et al. (Aberkane et al., 2008) investigated the output feedback solution for generalized stochastic hybrid linear systems and provided a dynamic system practical example. The infinite-horizon output feedback Nash game for a stochastic weakly-coupled system with state-dependent noise was studied in (Mukaidani et al., 2010). In addition, the necessary conditions for the existence of Nash equilibrium were given in (Mukaidani et al., 2010). Klompstra (Klompstra, 2000), extended risk-sensitive control to discrete time game theory and solved the Nash equilibrium for the partially observed state of a 2-player game.

In this paper, we are motivated to extend the above-referenced studies by considering higher-order statistics of cost function. In particular, we consider a second cumulant nonzero-sum Nash game for a partially-observed system of two players on a fixed time interval where the players shape the distribution of their cost cumulant function to improve system performance. This form of dynamic game finds application in satellite and mobile robot systems. The second cumulant of cost function is equivalent to the variance of the cost function. However, the optimization of cost function distribution through cost cumulant was initiated by Sain (Sain, 1966), (Sain and Liberty,
1971) while Won et al. (Won et al., 2010), extended the theory of cost cumulant to second, third and fourth cumulants for a nonlinear system with non-quadratic cost and derived the corresponding HJB equations.

The reminder of this paper is organized as follows. In Section 2, we state the mathematical preliminaries and formulate the second cumulant game problem. Section 3 states the necessary condition for the existence of Nash equilibrium solution while Section 4 derives the players strategy based on solving the coupled Hamilton-Jacobi-Bellman (HJB) equations which is the main result of this paper. Section 5 describes the numerical approximate method for solving the coupled HJB equations while a numerical example is solved in Section 6. Finally, the conclusions are given in Section 7.

2 PROBLEM FORMULATION

Consider a 2-player linear systems and measured output described by the linear Bôcher stochastic differential equation.

\[
\begin{align*}
\dot{x}(t) &= A(t)x(t)dt + \sum_{k=1}^{n} B_k(t)u_k(t)dt + G(t)dw_1(t), \\
\dot{y}(t) &= C(t)x(t)dt + D(t)dw_2(t),
\end{align*}
\]

where \( t \in [0,T] \), \( x(t) \in \mathbb{R}^n \) is the state, \( u_k(t) \in U_k \subset \mathbb{R}^m \) is the \( k \)-th player strategy, \( k = 1, 2 \) and \( w_1(t), w_2(t) \) are Gaussian random process defined on a probability space \((\Omega_0, F, \mathbb{P})\) where \( \Omega_0 \) is nonempty set, \( F \) is a \( \sigma \)-algebra of \( \Omega_0 \) and \( \mathbb{P} \) is a probability measure on \((\Omega_0, F)\). \( x(t_0) = x_0 \) is the initial state vector with covariance matrix \( P_0 \). The Gaussian random process \( w_1(t) \) has zero mean and covariance of \( E(\dot{w}_1(t)dw_1'(t)) = W_1(t)dt \) and additionally the Gaussian random process \( w_2(t) \) has zero mean and covariance of \( E(\dot{w}_2(t)dw_2'(t)) = W_2(t)dt \). The noise processes \( w_1(t) \) and \( w_2(t) \) are assumed independent with \( E(\dot{w}_1(t)dw_2'(t)) = E(\dot{w}_2(t)dw_1'(t)) = 0 \) assuming \( dw_1, dw_2 \) have same dimension. Let \( \Omega_0 = [0, T] \times \mathbb{R}^n \), \( \Omega_0 \) denote the closure of \( \Omega_0 \), i.e. \( \overline{\Omega_0} = T \times \mathbb{R}^n \). Assume there exist constants \( c_1, c_2 > 0 \in \mathbb{R} \) such that

\[
\|A(t)\| + \sum_{k=1}^{n} \|B_k\| \leq c_1, \quad \|G(t)\| \leq c_2,
\]

where \( A(\cdot), B_k(\cdot), C(\cdot), D(\cdot), G(\cdot) \) are elements of \( C^1([0,T]) \) with appropriate dimensions. Let a feedback strategy law be defined as \( u_k(t) = \mu_k(t, x(t)), t \in T \). Then, (1) can be written as

\[
\dot{x}(t) = f^{\mu_k}(x(t))dt + G(t)dw_1(t), x(t_0) = x_0,
\]

where \( f^{\mu_k}(x(t)) \) denotes \( A(t)x(t) + \sum_{k=1}^{n} B_k(t)u_k(t) \). There exist a bounded, borel measurable feedback strategy \( \mu_k(x) : \mathbb{R}^m \to U_k \) such that \( \mu_k(x) \) satisfies a global Lipschitz condition. i.e there exists a constant \( c_1 \) such that

\[
\|\mu_1(x_1) - \mu_2(x_2)\| \leq c_1 \|x_1 - x_2\|,
\]

\( \|\cdot\| \) is the Euclidean norm and \( x_1, x_2 \in \mathbb{R}^n \). Also, \( \mu_k(x) \) satisfies linear growth condition

\[
\|\mu_k(x)\| \leq c_2(1 + \|x\|).
\]

Then, if \( E\|x(t)\|^2 \) is finite, there is a unique solution to (1) which is a Markov diffusion process on \( \mathbb{R}^n \) (Fleming and Rishel, 1975). In order to assess performance of (1), consider the cost function \( J^k \) for the \( k \)-th player given as:

\[
J^k(t_0, x(t_0), \mu_1, \mu_2) = x'((t_0)Q_1x(t_F) + \int_{t_0}^{t_F} [x(s)Q(s)x(s) + \sum_{i=1}^{n} \mu_i(s)R_i(s)\mu_i(s)]ds,
\]

where \( k = 1, 2 \), \( Q(s), Q_i(s), R_i(...) \) are symmetric positive semi-definite and \( R_i(...) \) is symmetric positive definite, which can also be represented as

\[
J^k(t_0, x(t_0), \mu_1, \mu_2) = \int_{t_0}^{t_F} L^k(s,x,\mu_1, \mu_2)ds + \psi^k(x(t_F)),
\]

where \( k = 1, 2 \), \( L^k \) is the running cost, \( \psi^k \) is the terminal cost and \( L^k, \psi^k \) both satisfy polynomial growth condition. Let the state estimate be \( \hat{x}(t) \) and the state estimate error be \( \bar{x}(t) \) where \( x(t) \) is the state true value. Then, the state estimation error \( \bar{x}(t) \) is given as

\[
\bar{x}(t) = x(t) - \hat{x}(t).
\]

The filtered state estimate \( \hat{x}(t) \) is given as

\[
\dot{\hat{x}}(t) = A(t)\bar{x}(t)dt + \sum_{k=1}^{n} \left(B_k(t)u_k(t)\right)dt
\]

\[
+ K(t)\left(\bar{y}(t) - C(t)\hat{x}(t)dt\right).
\]

where \( K(t) \) is the Kalman Filter gain (Davis, 1977).

**Lemma 2.1.** The expected value of the cost function (6) conditioned on the \( \sigma \)-algebra generated by the measured output (1) can be rewritten as

\[
E\left\{J^k(t_0, x(t_0), \mu_1, \mu_2)\right\} = \int_{t_0}^{t_F} \left[ E\left(\bar{x}'(s)Q(s)\bar{x}(s)\right) + \text{tr}\left(Q(s)P(s)\right)\right]ds + \int_{t_0}^{t_F} \sum_{i=1}^{n} \mu_i(s)R_i(s)\mu_i(s)ds
\]

\[
+ E\left\{\bar{x}'(t_F)Q_J\bar{x}(t_F)\right\} + \text{tr}\left(Q_JP_J\right),
\]

where \( f^{\mu_k}(x(t)) \) denotes \( A(t)x(t) + \sum_{k=1}^{n} B_k(t)u_k(t) \).
where \( k = 1, 2 \). \( \dot{x}(t) = \xi_f, Q(.) , Q_f, \mathcal{P}(.) , \mathcal{P}_f \) are positive semi-definite, \( R_k(x) \) is positive definite for \( k = i \) and positive semi-definite for \( k \neq i \). \( \mathcal{P}(.) , \mathcal{P}_f \) are the state error estimate covariances.

**Proof.** See (Davis, 1977) for single player case, a two-player case follows similar derivation.

Furthermore, we utilize the backward evolution operator, \( \mathbb{O}^k(\mu_1, \mu_2) \), as defined in (Sain et al., 2000):

\[
\mathbb{O}^k(\mu_1, \mu_2) = \frac{\partial}{\partial t} + f'(t, x, \mu_1, \mu_2) \frac{\partial}{\partial x},
\]

\[
\mathbb{O}_2^k(\mu_1, \mu_2) = \frac{1}{2} \text{tr} \left( G(t) W_1(t) G(t) ' \frac{\partial^2}{\partial x^2} \right),
\]

with \( t = \text{tr} \) in (11). To study the cumulant game of cost function, the \( m \)-th moments of cost functions \( M_m^k \) of the \( k \)-th player is defined as:

\[
M_m^k(t, x, \mu_1, \mu_2) = E \left( (J^m)^{(t, x, \mu_1, \mu_2)} | x(t) = x \right),
\]

where \( m = 1, 2 \). The \( m \)-th cumulant cost function \( V^k_m(t, \hat{x}) \) of the \( k \)-th player is defined by (Smith, 1995):

\[
V^k_m(t, \hat{x}) = M_m^k - \sum_{i=0}^{m-2} \frac{(m-1)!}{i!(m-1-i)!} M_{m-i-1}^k - V^k_{i+1},
\]

where \( t \in T = [t_0, T], x_0 = x_0, \hat{x}(t) \in \mathbb{R}^n \). Next, we introduce some definitions.

**Definition 2.1.** A function \( M^k_1: \tilde{Q}_0 \rightarrow \mathbb{R}^+ \) is an admissible first moment cost function if there exists a strategy \( \mu_1 \) such that

\[
M^k_1(t, \hat{x}) = M^k_1(t, \hat{x}; \mu_1, \mu_2),
\]

for \( t \in T, \hat{x} \in \mathbb{R}^n, M^k_1 \in C^{1,2}(\tilde{Q}_0) \). Also, \( V^k_1 \) is the admissible first cumulant cost function for the \( k \)-th player related to the moment function through the moment-cumulant relation (13). In addition, \( \mu_0 \in U_{M^k_0}, V^k_1(t, \hat{x}) = V^k_1(t, \hat{x}; \mu_1, \mu_2) \).

**Definition 2.2.** A class of admissible strategy \( U_{M^k_0} \) is defined such that if \( \mu_0 \in U_{M^k_0} \subset \mathbb{R}^m \) then \( \mu_0 \) satisfies the equality of Definition 2.1 for \( M^k_0, M^k_1 \). It should be noted that first moment \( M^k_1 \) is the same as first cumulant \( V^k_1, M^k_0 = 1 \) and \( V^k_0 = 0 \).

**Definition 2.3.** Let \( V^k_1 \) be the \( k \)-th player admissible cumulant cost functions. The player strategy \( \mu^*_k \) is the \( k \)-th player equilibrium solution if it is such that

\[
V^k_1(t, \hat{x}) = V^k_1(t, \hat{x}; \mu^*_1, \mu^*_2) \leq V^k_1(t, \hat{x}; \mu_1, \mu_2),
\]

\[
V^k_2(t, \hat{x}) = V^k_2(t, \hat{x}; \mu^*_1, \mu^*_2) \leq V^k_2(t, \hat{x}; \mu_1, \mu_2),
\]

for all \( \mu_k \in U_{M^k} \), where the set \( \{ \mu^*_1, \mu^*_2 \} \) is a Nash equilibrium solution and the set \( \{ V^k_1, V^k_2 \} \) is the Nash equilibrium value set.

**Problem Definition.** Consider an open set \( Q \subset Q_0 \) and let the \( k \)-th player cost cumulant functions \( V^k_1(t, \hat{x}) \in C^{1,2}_p(Q) \cap \mathcal{C}(Q) \) be an admissible cumulant function where the set \( C^{1,2}_p(Q) \cap \mathcal{C}(Q) \) means that the function \( V^k_1 \) satisfy polynomial growth condition and is continuous in the first and second derivatives of \( Q \), and continuous on the closure of \( Q \). Assume the existence of a near-optimal strategy \( \mu^*_k \in U_{M^k} \) and near-optimal value function \( V^k_1(t, \hat{x}) \in C^{1,2}_p(Q) \cap \mathcal{C}(Q) \) for the \( k \)-th player. Thus, a 2-player second cumulant output feedback game problem is to find the Nash strategy \( \mu^*_k(t, \hat{x}) \) for the partially-observed linear state system with \( k = 1, 2 \) which results in the near-optimal 2\textsuperscript{nd} value function \( V^k_2(t, \hat{x}) \) given as

\[
V^k_2(t, \hat{x}) = \min_{\mu_k \in U_{M^k}} \left\{ V^k_1(t, \hat{x}; \mu^*_1, \mu^*_2) \right\},
\]

(16)

**Remarks.** To find the Nash equilibrium strategies \( \mu^*_1(t, \hat{x}), \mu^*_2(t, \hat{x}) \), we constrain the candidates of the near-optimal players strategy to \( U_{M^k}, U_{M^k} \), and the near-optimal value functions \( V^k_1(t, \hat{x}), V^k_2(t, \hat{x}) \) are found with the assumption that \( V^k_1(t, \hat{x}), V^k_2(t, \hat{x}) \), are admissible.

### 3 AD HOC OUTPUT FEEDBACK CUMULANT GAME

**Theorem 3:** From the full-state feedback statistical control in (Won et al., 2010), the minimal 2\textsuperscript{nd} value function \( V^k_2(t, x) \) for (1) with zero measurement noise satisfies the following HJB equation for the \( k \)-th player:

\[
0 = \min_{\mu_k \in U_{M^k}} \left\{ \mathbb{O}^k(\mu_1, \mu_2) \left[ V^k_2(t, x) \right] \right\} + \left( \frac{\partial V^k_2(t, x)}{\partial x} \right)' G(t) W_1(t) G(t)' \left( \frac{\partial V^k_2(t, x)}{\partial x} \right),
\]

(17)

with the terminal condition \( V^k_j(t_f, x_f) = 0, k = 1, 2, j = 1, 2 \). Assuming separation principle (Wonham, 1968), the minimal 2\textsuperscript{nd} value function \( V^k_2(t, \hat{x}) \) for (9) satisfies the following HJB equation for the \( k \)-th player:
\[ 0 = \min_{\mu_k \in \mathcal{U}_k} \left\{ O^k(\mu_k^1, \mu_k^2) \left[ V_{2,k}^{k}\left(t, \tilde{x}\right) \right] \right. \]
\[ + \left( \frac{\partial V_{1,k}^k(t, \tilde{x})}{\partial \tilde{x}} \right)' \left\{ K(t)W_2K(t)' \left( \frac{\partial V_{1,k}^k(t, \tilde{x})}{\partial \tilde{x}} \right) \right\}, \]
\]
with terminal condition \( V_{2,k}^k(t_F, \tilde{x}_F) = 0, k = 1, 2, j = 1, 2. K(t) \) is the Kalman filter gain. Since the first cost cumulant function \( V_{1,k}^k \) is admissible (def. 2.1), the following coupled equations are satisfied
\[ 0 = O^k(\mu_{-k}, \mu_k) \left[ V_{1,k}^k(t, \tilde{x}) \right] + M_0^k(t, \tilde{x}) L^k(t, \tilde{x}, \mu_{-k}, \mu_k), \]
\[ 0 = O^k(\mu_{-k}, \mu_k) \left[ V_{2,k}^k(t, \tilde{x}) \right] + \left( \frac{\partial V_{1,k}^k(t, \tilde{x})}{\partial \tilde{x}} \right)' \left\{ K(t)W_2K(t)' \left( \frac{\partial V_{1,k}^k(t, \tilde{x})}{\partial \tilde{x}} \right) \right\}, \]
where \( M_0^k = 1, O^k(.) \) is the backward operator and the first line of (22) follows from the classical HJB equation while the second line relates the second cumulant function with the first cumulant function in the HJB equation. Thus, converting (22) to unconstrained optimization problem gives
\[ 0 = \min_{\mu_{-k} \in \mathcal{M}_k} \left\{ O^k(\mu_{-k}, \mu_k) \left[ V_{1,k}^k(t, \tilde{x}) \right] \right. \]
\[ \left. + M_0^k(t, \tilde{x}) L^k(t, \tilde{x}, \mu_{-k}, \mu_k) + \gamma_{2,k}(t) O^k(\mu_{-k}, \mu_k) \left[ V_{1,k}^k(t, \tilde{x}) \right] \right\}, \]
(23)
where \( \gamma_{2,k} \) is the Lagrange multiplier. From backward operator (11) using (9), (10) and expanding (23) gives
\[ \min_{\mu_{-k} \in \mathcal{M}_k} \left\{ \left( \frac{\partial V_{1,k}^k(t, \tilde{x})}{\partial t} \right) + \tilde{x}(t)Q(t)\tilde{x}(t) \right. \]
\[ + \left. \frac{1}{2} \sum_{i=1}^n \mu_{i}(t)R_{ii}(t)\mu_{i}(t) \right\}, \]
\[ + \frac{1}{2} \gamma_{2,k}(t) \left( \frac{\partial V_{1,k}^k(t, \tilde{x})}{\partial x} \right)' \left( \frac{\partial V_{1,k}^k(t, \tilde{x})}{\partial x} \right) \]
\[ = \frac{1}{2} tr \left( K(t)W_2K(t)' \left( \frac{\partial V_{1,k}^k(t, \tilde{x})}{\partial x} \right) \right), \]
(24)
Minimizing (24) with respect to \( \mu_k(t, \tilde{x}) \) gives
\[ \mu_k^k(t, \tilde{x}) = -\frac{1}{2} R_{k\tilde{x}}^k B_{k}' \left( \frac{\partial V_{1,k}^k(t, \tilde{x})}{\partial \tilde{x}} \right) + \gamma_{2,k}(t) \left( \frac{\partial V_{1,k}^k(t, \tilde{x})}{\partial \tilde{x}} \right). \]
(25)
Remark. The strategy for the k-th player $\mu_k(t, \hat{x})$ derived from the coupled HJB equation (22) is suboptimal. In (22), the certainty equivalent principle has been extended to the second cumulant output feedback game where a Kalman filter is used for state estimation.

5 NUMERICAL APPROXIMATION METHOD

The analytical solutions of HJB equation (18) is difficult to find except for simple linear systems. Sandberg (Sandberg, 1998) showed that neural networks (NN) with time-varying weights can be utilized to approximate uniformly continuous time-varying functions. We are motivated by the work in (Chen et al., 2007) to extend NN approach to cost cumulant game. In this approach, NN is utilized to approximate the value function based on method of least squares on a pre-defined region. The value functions $V_{mk}^t(\hat{x})$ can be approximated as $V_{mk}^t(\hat{x}, x)$, where we introduce gaussian noise as $w_i(t)$ in state $x(t)$ on a compact set $\Omega \rightarrow \mathbb{R}^n$. Thus, we approximate the players value functions $V_{mk}^t$ as $V_{mk}^t(\hat{x}, x) = \sum_{i=1}^L w_i(t) \gamma_i(\hat{x})$ on $t$ on a compact set $\Omega \rightarrow \mathbb{R}^n$. We assume that $G$ and $D$ in (26) are $4 \times 1$ and $3 \times 1$ constant vectors given as $G = [1 \ 1 \ 1 \ 1]$, $D = [1 \ 1 \ 1]$ and $d w_1(t), d w_2(t)$ in (26) as a Gaussian process with mean $E\{d w_1(t)\} = E\{d w_2(t)\} = 0$ and covariance $E\{d w_1(t) d w_1(t)^t\} = 0.1$ and $E\{d w_2(t) d w_2(t)^t\} = 0.1$. In this example, we study a 2-player 2nd cumulant ad hoc output feedback Nash game. Here, we compute the suboptimal solution for the player strategy through solving the output feedback 2nd cumulant game problem constraint on the 1st cumulant cost function.

The first player cost function $J_1^t$ is

$$J_1^t(x_0, x(t_0), u_1(0), u_2(t_0)) = \int_{x_0}^{\psi_1(x(t_0), t_0)} \{x_1^2(t) + x_2^2(t) \} + x_1^2(t) + x_2^2(t) + u_1^2(t) dt + \psi_1(x(t), t)$$

(27)

where $\psi_1(x(t), t_0) = 0$ is the terminal cost and the second player cost function $J_2^t$ is

$$J_2^t(x_0, x(t_0), u_1(0), u_2(t_0)) = \int_{x_0}^{\psi_2(x(t_0), t_0)} \{x_1^2(t) + x_2^2(t) \} + x_1^2(t) + x_2^2(t) + u_2^2(t) dt + \psi_2(x(t), t)$$

(28)

where $\psi_2(x(t), t_0) = 0$ is the terminal cost. The activation function $A(t, \hat{x})$ for the value functions of the players are the same and based on (Chen and Jagannathan, 2008) which are formulated as

$$A(t, \hat{x}) = \sum_{i=0}^M \left( \sum_{j=1}^n x_j \right)^{2i}$$

(29)

where in (29), $M$ is an even-order of the approximation, $L$ is the number of hidden-layer neurons, $n$ is the system dimension.

The input function $A(t, \hat{x})$ (29) is

$$A(t, \hat{x}) = \left\{ x_1^2, x_2, x_1^2 x_2, x_1^2 x_3, x_1 x_2, x_1 x_3, x_2 x_3, x_2 x_4, x_3 x_4 \right\}^t.$$ 

(30)

We transform this problem as an innovative process (Kailath, 1968) in terms of state estimate using (8), (9), (10) and solve for the Kalman filter gain. For the NN series approximation, we choose a polynomial function (30) of up to second-order ($M = 2$) in state variable (i.e $x$ is 2nd order) with length $L = 10$. Higher order polynomial did not provide significant improvement in the approximation accuracy. In the simulation, the asymptotic stability region for states was arbitrarily chosen as $-5 \leq x_1 \leq 5, -5 \leq x_2 \leq 5, -5 \leq...$
$x_3 \leq 5$ and $-5 \leq x_4 \leq 5$. The final time $t_F$ was 20 s and $w_{11L}(t_F) - w_{21L}(t_F) = \{0\}$ and $w_{12L}(t_F) - w_{22L}(t_F) = \{0\}$. The initial condition was $x(t_0) = x_0 = [1 \ 1 \ 1 \ 1]^T$.

Figs. 1(a) to 1(c) show the first player neural network weights and value functions which are similar to the second player, hence only the first player plots are shown. Fig. 1(a), the neural network weights converge to constants. Plots 1(b) to 1(c) show the first and second value cumulant functions. From Fig. 1(b), it was observed that the value function $V_{11}^I$ increases with increase in $\gamma_{21}$ while from Fig. 1(c), it was observed that the value functions $V_{12}^I$ decreases as $\gamma_{21}$ increase. The Lagrange multipliers $\gamma_{2k}$ were selected as constants. The Nash suboptimal controls, $u_1$ and $u_2$ are shown in Fig. 2(a). It should be noted from Fig. 2(a), that the Nash suboptimal controls for the two players were solved for the 2nd cumulant game by selecting $\gamma_{21}$, $\gamma_{22}$ where the value functions are mini-
maximum which in our case were $\gamma_{21} = 0.001, \gamma_{22} = 0.001$. In addition, we have the design freedom in $\gamma_{2k}$ values selection to enhance system performance. From Figs. 2(b) to 2(c), the states converge to values close to zero.

7 CONCLUSIONS

In this paper, we analyzed an output feedback cumulant differential game control problem using cost cumulant optimization approach. We investigated a linear stochastic system with two players and derived a 2-player near-optimal strategies for the tractable auxiliary problem. The efficiency of our proposed method has been demonstrated using a numerical example where a neural network series method was applied to solve the HJB equations.

REFERENCES


