ADCONIP 2014 Hiroshima

Nonlinear Stochastic Dynamic Games*

Muneomi Sagara[1], Hiroaki Mukaidani[2], Masaru Unno[3] and Hua Xu[4]

***Abstract*— In this paper, dynamic games for a class of infinite horizon for nonlinear stochastic system governed by Itô differential equation are investigated. Particularly, Pareto and Nash strategies are both discussed. After defining the equilibrium condition, the conditions for the existence of the strategy sets are given by means of solvability of cross-coupled Hamilton-Jacobi-Bellman equations (HJBEs). It is shown that these higher-order approximate strategy sets can be obtained by computing the recursive algorithm. A simple numerical example is given to show the reliability and usefulness of the considerable results.**

I. INTRODUCTION

Stochastic dynamic games for large-scale systems have become a popular research field in the last decade [1], [2], [3], [4], [5]. From a numerical point of view, similar problem for a stochastic system governed by Itô differential equations has been discussed [11], [12]. However, these existing results on dynamic game were investigated for the linear case. It is well known that nonlinear systems have important applications in engineering practice since they can be captured to represent various plants. Therefore, it can be observed that nonlinear stochastic dynamic games of stochastic Itô systems are still open and deserve further study.

Stochastic H_2/H_∞ control problems on the generalized Nash equilibrium strategy have become a popular research field [6], [7]. Although these results in [6], [7] are very efficient in theory, a practical issue that remains concerns the computation of the solutions of Hamilton-Jacobi-Bellman equations (HJBEs). In [8], an on-line adaptive control algorithm based on policy iteration reinforcement learning techniques to solve the multiplayer non-zero-sum game with infinite horizon for linear and nonlinear systems has been discussed. As a novel concept, the proposed adaptation algorithm is implemented via actor/critic approximated neural network (NN) for every player. Particularly, the required parameter can update on-line such that the solution of the coupled HJBEs converge. However, stochastic systems governed by Itô differential equation have not been considered. On the other hand, stochastic H_2/H_∞ control problem for nonlinear systems has been discussed [13]. Although the obtained result in this literature is very important theory and despite it being to compute a strategy set, the basic results for classical games are an open issue that remains to be considered.

In this paper, dynamic games for a class of infinite horizon for nonlinear stochastic system governed by Itô differential equation with multiple decision makers are investigated. Particularly, Pareto and Nash strategies as a main strategy are both treated. After establishing the existing conditions for the strategy set by means of solvability of cross-coupled Hamilton-Jacobi-Bellman equations (HJBEs), a novel recursive algorithm for solving HJBEs is established as compared with the existing result [13]. As a result, a higher-order approximate strategy set can be easily obtained iteratively. The main contributions of this paper are as follows. First, in order to establish the condition for the existence of the strategy set, several important preliminary results are introduced. After declaring the problem for the Pareto and Nash games, the strategy sets are derived via HJBE or HJBEs, respectively. It is newly shown that the Nash strategy set can be obtained by solving HJBEs iteratively. Compared with the existing stochastic optimal control problem [10], the multiple decision making problem is addressed. As a result, the consensus strategies for cooperative control can be solved. Finally, a numerical example is demonstrated to show the usefulness of the proposed strategy set.

Notation: The notations used in this paper are fairly standard. $\|v\|$ denotes the Euclidean norm of a real n-dimensional vector v. $\mathcal{L}^2_{\mathcal{F}}(\Re^+,\ \Re^q)$ denotes the space of non-anticipative stochastic processes $y(t) \in \Re^q$ with respect to an increasing σ-algebras F_t-measurable for every $t \geq 0$ satisfying $\mathbf{E}[\int_0^\infty \|y(t)\|^2 dt] < \infty$. $\mathbf{E}[\ \cdot\ ]$ stands for the conditional expectation operator. $\boldsymbol{C}^2(\Re^n)$ denotes the class of function twice continuously differential about $x \in \Re^n$. Finally, throughout this paper we have used the notation $\|x(t)\|^2_R$ instead of $x^T(t)Rx(t)$.

II. PRELIMINARY RESULTS

Consider the N-player nonlinear stochastic system governed by Itô differential equation defined by

$$dx(t) = \left[f(x(t)) + \sum_{j=1}^{N} g_j(x(t))u_j(t) \right] dt + r(x(t))dw(t), \quad (1)$$

*This work was supported in part by a Grant-in-Aid for Scientific Research (C)-23500017 from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

[1]Faculty of Administration, Prefectural University of Kumamoto, 3-1-100 Tsukide, Higashi-ku, Kumamoto, 862-8502 Japan. sagara@pu-kumamoto.ac.jp

[2]Institute of Engineering Hiroshima University, 1-4-1 Kagamiyama, Higashi-Hiroshima, 739-8527 Japan. mukaida@hiroshima-u.ac.jp.

[3]NTT FINANCE CORPORATION Seavans N(19th floor), 1-2-1 Shibaura Minato-ku, Tokyo, 105-6791, Japan. masaru.unno7@gmail.com

[4]Graduate School of Business Sciences, The University of Tsukuba, 3-29-1, Otsuka, Bunkyo-ku, Tokyo, 112-0012 Japan. xu@mbaib.gsbs.tsukuba.ac.jp

where $f(0) = 0$, $r(0) = 0$ and $x(0) = x^0$. $x(t) \in \Re^n$ represents the system state. $u_i(t) \in \mathcal{L}^2_{\mathcal{F}}(\Re^+, \ \Re^{m_i})$, $i = 1, \ \dots, N$ represent the i-th control inputs.

$w(t) \in \Re$ is a one-dimensional standard Wiener process defined in the filtered probability space $(\Omega, \ \mathcal{F}, \ \mathcal{P}, \ \mathcal{F}_t)$ with $\mathcal{F}_t = \sigma\{w(s) \ : \ 0 \le s \le t\}$ [6], [7]. f, g_i, h and k are assumed to be Borel measurable functions of suitable dimensions such that (1) has a unique strong solution on any finite interval $[0, \ T]$. Moreover, they are locally Lipschitz.

The cost functionals associated with each player are

$$\begin{aligned} & J_i(x^0, \ u_1, \ \dots, u_N) \\ &= \mathbf{E}\left[\int_0^\infty \left(\|x(t)\|^2_{Q_i} + \sum_{j=1}^N \|u_j(t)\|^2_{R_{ij}}\right) dt\right] \\ &= \mathbf{E}\left[\int_0^\infty \phi_i(x(t), \ u_1, \ \dots, u_N) dt\right], \end{aligned} \tag{2}$$

where $i = 1, \ \dots, N$ and $Q_i \ge 0$, $R_{ii} > 0$, $R_{ij} \ge 0$, $i \neq j$ are symmetric matrices.

It should be noted that since the function $r(x(t))$ depends only on the state $x(t)$, J_i can be defined without infinity expectation $\lim_{T\to\infty} 1/T\mathbf{E}[\int_0^T (\cdot) ds]$.

Some of definitions and useful lemmas with respect to stochastic Nash game are given.

Definition 1: [8] Feedback control strategies $u_i(t) = \mu_i(x)$, $i = 1, \ \dots, N$ are defined as admissible with respect to (2) on a set denoted by $\mu_i(x)$, if $\mu_i(x)$ is continuous, $\mu_i(x)$ stabilizes (1), and (2) is finite $\forall x^0$.

Definition 2: [9] The stochastically uncontrolled system governed by the Itô's equation $dx(t) = f(x(t))dt + r(x(t))dw$, $x(0) = x^0$, $f(0) = r(0) = 0$ is called exponentially mean square stable (EMSS), if there exist some positive scalars $\alpha > 0$ and $\beta > 0$ such that $\mathbf{E}[\|x(t)\|^2] \le \beta e^{-\alpha t}\|x(0)\|^2$ for all $t > 0$ and $x(0)$. Moreover, $x \equiv 0$ of the stochastically uncontrolled system is said to be asymptotically stable, if $\lim_{t\to\infty} \mathbf{E}[\|x(t)\|^2] = 0$.

The following lemma gives a solution of the infinite horizon stochastic nonlinear optimal regulator problem.

Lemma 1: [6], [7], [10] Consider the following regulator problem.

$$\min_u \ J(u) = \mathbf{E}\left[\int_0^\infty \left(\|x(t)\|^2_Q + \|u(t)\|^2_R\right) dt\right], \tag{3a}$$

$$\text{s.t. } dx(t) = [f(x(t)) + g(x(t))u(t)]dt + r(x(t))dw(t), \tag{3b}$$

where $Q \ge 0$ and $R > 0$ are symmetric matrices.

If there exists a nonnegative Lyapunov function $V_u \in \boldsymbol{C}^2(\Re^n)$ with

$$c_0\|x\|^2 \le V_u(x) \le c_1\|x\|^2, \quad c_0, \ c_1 > 0 \tag{4}$$

solving the following HJBE:

$$\frac{\partial V_u^T}{\partial x} f + \|x\|^2_Q - \frac{1}{4} \cdot \frac{\partial V_u^T}{\partial x} g R^{-1} g^T \frac{\partial V_u}{\partial x} + \frac{1}{2} r^T \frac{\partial^2 V_u}{\partial x^2} r = 0, \ V_u(0) = 0 \tag{5}$$

then we have

$$J \ge V_u(x^0) \tag{6}$$

with the optimal control

$$u(t) = u^*(t) = \mu^*(t) = -\frac{1}{2} R^{-1} g^T \frac{\partial V_u}{\partial x}. \tag{7}$$

Finally, the following result can be easily obtained by using the existing results [6], [7].

Corollary 1: Consider the following autonomous stochastic nonlinear system:

$$dx(t) = \tilde{f}(x(t))dt + \tilde{r}(x(t))dw(t). \tag{8}$$

If there exists a nonnegative Lyapunov function $\tilde{V} \in \boldsymbol{C}^2(\Re^n)$ with

$$\tilde{c}_0\|x\|^2 \le \tilde{V}(x) \le \tilde{c}_1\|x\|^2, \quad \tilde{c}_0, \ \tilde{c}_1 > 0 \tag{9}$$

solving the following HJBE:

$$\frac{\partial \tilde{V}^T}{\partial x} \tilde{f} + \|x\|^2 + \frac{1}{2} \tilde{r}^T \frac{\partial^2 \tilde{V}}{\partial x^2} \tilde{r} = 0, \ \tilde{V}(0) = 0 \tag{10}$$

then we have

$$\tilde{J} = \mathbf{E}\left[\int_0^\infty \|x(t)\|^2 dt\right] = \tilde{V}(x^0). \tag{11}$$

III. MAIN RESULTS

In this section, two strategies are discussed. The first one is Pareto strategy. The other one is Nash strategy.

A. Pareto Strategy

In this section, Pareto strategy set as one of the cooperative game theory is considered. It is assumed that each player wants to optimize its own cost described in (2). As the definition of Pareto efficient solution [16], let us combine the individual cost functions in (2) into a team cost function according to the following.

$$\begin{aligned} & J(x^0, \ u_1, \ \dots, u_N) := \sum_{j=1}^N \rho_j J_i(x^0, \ u_1, \ \dots, u_N) \\ &= \mathbf{E}\left[\sum_{j=1}^N \rho_j \int_0^\infty \left(\|x(t)\|^2_{Q_j} + \|u_j(t)\|^2_{R_{jj}}\right) dt\right], \\ & \sum_{j=1}^N \rho_j = 1, \ 0 < \rho_i < 1, \ i = 1, \ \dots, N, \end{aligned} \tag{12}$$

where $R_{ij} = 0$, $i \neq j$.

A Pareto solution is a set $(u_1, \ \dots, u_N)$, which minimizes $J(u_1, \ \dots, u_N)$. From the above problem, we obtain the following necessary optimality conditions.

Theorem 1: Suppose there exist a nonnegative-definite function $V \in \boldsymbol{C}^2(\Re^n)$ with the properties of

$$\bar{c}_0\|x\|^2 \le V(x) \le \bar{c}_1\|x\|^2, \quad \bar{c}_0, \ \bar{c}_1 > 0. \tag{13}$$

such that the following HJBE:

$$\begin{aligned} & \frac{\partial V^T}{\partial x} f(x(t)) + \bar{Q}_\rho - \frac{1}{4} \cdot \frac{\partial V^T}{\partial x} \bar{S}_\rho \frac{\partial V}{\partial x} \\ & + \frac{1}{2} r^T \frac{\partial^2 V}{\partial x^2} r = 0, \ V(0) = 0, \end{aligned} \tag{14}$$

where

$$\bar{S}_\rho := \sum_{j=1}^{N} g_j(\rho_j R_{jj})^{-1} g_j^T,\ \bar{Q}_\rho := \sum_{j=1}^{N} \rho_j \|x(t)\|_{Q_j}^2.$$

If HJBE (14) admits a strategy set $(K_1(x),\ \dots, K_N(x))$, then the infinite horizon dynamic game has a Pareto strategy set

$$u_i(t) = u_i^*(t) = K_i^*(x) = -\frac{1}{2}(\rho_i R_{ii})^{-1} g_i^T \frac{\partial V^*}{\partial x}. \quad (15)$$

Moreover, the closed-loop stochastic system is EMSS and $J^*(x^0,\ u_1^*,\ \dots, u_N^*) = V(x^0)$.

Proof: By using Itô's formula, (14) and completing the squares, the following equation holds.

$$\begin{aligned} & J_T(x^0,\ u_1,\ \dots, u_N) \\ &= \mathbf{E}\left[\sum_{j=1}^{N} \rho_j \int_0^T \left(\|x(t)\|_{Q_j}^2 + \|u_j(t)\|_{R_{jj}}^2\right) dt\right] \\ &= \mathbf{E}\left[\sum_{j=1}^{N} \rho_j \int_0^T \left(\|x(t)\|_{Q_j}^2 + \|u_j(t)\|_{R_{jj}}^2\right) dt + dV(x)\right] \\ &\quad + V(x^0) - V(x(T)) \\ &= \mathbf{E}\left[\int_0^T \sum_{j=1}^{N} [u_j - K_j(x)]^T (\rho_j R_{jj}) [u_j - K_j(x)] dt\right] \\ &\quad + V(x^0) - V(x(T)). \end{aligned} \quad (16)$$

Therefore, since $\lim_{T\to\infty} \mathbf{E}\|x(T)\| = 0$, if $u_i = K_i(x) = -1/2(\rho_i R_{ii})^{-1} g_i^T \partial V/\partial x$, it is easy to see that

$$\begin{aligned} & J^*(x^0,\ u_1^*,\ \dots, u_N^*) \\ &= \mathbf{E}\left[\sum_{j=1}^{N} \rho_j \int_0^\infty \left(\|x(t)\|_{Q_j}^2 + \|u_j(t)\|_{R_{jj}}^2\right) dt\right] \\ &= V(x^0). \end{aligned} \quad (17)$$

Hence, the proof is completed. ■

As an important application, a linear stochastic system and quadratic cost functions are considered. Assume that

$$f(x(t)) = Ax(t),\ g_i(x(t)) = B_i x(t),\ i = 1,\ \dots, N,$$
$$r(x(t)) = Cx(t).$$

Furthermore, let us define the following equation.

$$V = x^T(t) P x(t).$$

Therefore, Theorem 1 immediately yields the following Corollary as the linear time invariant case.

Corollary 2: Let us consider the following linear time-invariant stochastic systems.

$$dx(t) = \left[Ax(t) + \sum_{j=1}^{N} B_j u_j(t)\right] dt + Cx(t) dw(t), \quad (18)$$

with cost functionals (2).

Suppose that the following stochastic algebraic Riccati equation (SARE) have solutions $P \geq 0$.

$$PA + A^T P + C^T PC - PS_\rho P + Q_\rho = 0, \quad (19)$$

where

$$S_\rho = \sum_{j=1}^{N} B_j(\rho_j R_{jj})^{-1} B_j^T,\ Q_\rho = \sum_{j=1}^{N} \rho_j Q_j.$$

Then the infinite horizon Pareto strategy set is given below.

$$u_i^*(t) = -(\rho_i R_{ii})^{-1} B_i^T P x(t). \quad (20)$$

It should be noted that the obtained results are the same as the existing ones [15].

B. Nash Game

First, the definition of the stochastic Nash equilibrium is given by exploring the existing result.

Definition 3: Let us consider the nonlinear stochastic system expressed (1). Find an admissible state feedback strategies $u_i(t) = u_i^*(t) \in \mu_i(x)$, $u_i^*(0) = 0$ such that

$$\begin{aligned} & J_i(u_1^* \dots, u_N^*,\ x^0) \\ & \leq J_i(u_1^* \dots,\ u_{i-1}^*,\ u_i,\ u_{i+1}^*,\ \dots, u_N^*,\ x^0), \\ & \forall u_i(t) = \mu_i(x) \in \Re^{m_i}. \end{aligned} \quad (21)$$

The following result gives the nonlinear Nash strategy set.

Theorem 2: Suppose there exist a nonnegative-definite function $V_i \in \boldsymbol{C}^2(\Re^n)$ with the properties of

$$c_{i0}\|x\|^2 \leq V_i(x) \leq c_{i1}\|x\|^2,\quad c_{i0},\ c_{i1} > 0. \quad (22)$$

such that the following cross-coupled HJBEs:

$$\begin{aligned} & \frac{\partial V_i^T}{\partial x} \hat{f}_{-i}(x(t)) + \hat{Q}_{-i}(x) - \frac{1}{4} \cdot \frac{\partial V_i^T}{\partial x} g_i R_{ii}^{-1} g_i^T \frac{\partial V_i}{\partial x} \\ & + \frac{1}{2} r^T \frac{\partial^2 V_i}{\partial x^2} r = 0,\ V_i(0) = 0, \end{aligned} \quad (23)$$

where

$$\hat{f}_{-i}(x(t)) = f(x(t)) + \sum_{j=1,\ j\neq i}^{N} g_j(x(t)) \mu_j(t),$$

$$\hat{Q}_{-i}(x(t)) := \|x(t)\|_{Q_i}^2 + \sum_{j=1,\ j\neq i}^{N} \|\mu_j(x)\|_{R_{ij}}.$$

If the cross-coupled HJBEs (23) admits a strategy set $(V_1,\ \dots, V_N,\ \mu_1,\ \dots, \mu_N)$, then the infinite horizon Nash game has a strategy set

$$u_i(t) = \mu_i^*(t) = -\frac{1}{2} R_{ii}^{-1} g_i^T \frac{\partial V_i^*}{\partial x}. \quad (24)$$

Proof: In order to apply Lemma 1, the following regulator problem is considered.

$$\begin{aligned} & \text{minimize } J_i(u_i), \quad (25) \\ & \text{s.t. } dx(t) = [\hat{f}_{-i}(x(t)) + g_i(x(t)) u_i(t)] dt + r(x(t)) dw(t), \end{aligned}$$

where

$$\begin{aligned} & J_i(u_i) = J_i(\mu_1^*,\ \dots,\ \mu_{i-1}^*,\ \mu_i,\ \mu_{i+1}^*,\ \dots, \mu_N^*,\ x^0) \\ & = \mathbf{E}\left[\int_0^\infty [\hat{Q}_{-i}(x(t)) + \|\mu_i(t)\|_{R_{ii}}^2] dt\right]. \end{aligned}$$

If there exists a nonnegative Lyapunov function $V_i \in \boldsymbol{C}^2(\Re^n)$ with (22) solving the HJBE (23). Thus, we have

$$J_i(\mu_1^* \ldots , \ \mu_{i-1}^*, \ \mu_i, \ \mu_{i+1}^*, \ \ldots , \mu_N^*, \ x^0) \geq V_i(x^0) \quad (26)$$

with the optimal strategy (24) can be derived, respectively. This is the desired result. ■

Now we assume a linear system and a quadratic cost function. Assume that

$$f(x(t)) = Ax(t), \ g_i(x(t)) = B_i x(t), \ i = 1, \ \ldots , N,$$
$$r(x(t)) = Cx(t).$$

Thus, we have

$$V_i = x^T(t)P_i x(t), \ i = 1, \ \ldots , N.$$

Therefore, Theorem 2 immediately yields the following Corollary.

Corollary 3: Let us consider the following linear time-invariant stochastic systems.

$$dx(t) = \left[Ax(t) + \sum_{j=1}^{N} B_j u_j(t) \right] dt + Cx(t)dw(t), \quad (27)$$

with cost functionals (2).

Suppose that the following cross-coupled stochastic algebraic Riccati equations (CSAREs) have solutions $P_i \geq 0$, $i = 1, \ .. , N$.

$$P_i A_{-i} + A_{-i}^T P_i + C^T P_i C - P_i S_i P_i + Q_{-i} = 0, \quad (28)$$

where $i = 1, \ \ldots , N$,

$$A_{-i} = A + \sum_{j=1, \ j \neq i}^{N} S_j P_j, \ S_i = B_i R_{ii}^{-1} B_i^T,$$
$$Q_{-i} = Q_i + \sum_{j=1, \ j \neq i}^{N} P_j G_{ij} P_j.$$

Then the infinite horizon Nash strategy set is given below.

$$u_i(t) = \mu_i^* = -R_{ii}^{-1} B_i^T P_i x(t). \quad (29)$$

It should be noted that the obtained results are the same as the existing ones [14].

IV. SUCCESSIVE APPROXIMATION

In order to obtain the strategy sets, HJB or HJBEs should be solved. It is well known that it is hard and complicated to solve these equations. When the Pareto strategy set is computed, the existing result [10] seems to be reliable and useful. On the other hand, there is no algorithm to calculate the solution of HJBEs. Hence, the successive approximation algorithm for solving HJBEs (23) is newly given below.

Step 1. *Initialization: After linearizing the stochastic nonlinear differential equation* (1) *to* (27)*, choose*

$$u_i^{(0)}(t) = \mu_i^{(0)}(t) = -R_{ii}^{-1} B_i^T P_i x(t).$$

by solving CSAREs (28).

Step 2. *For* $k \geq 0$*, solve the following generalized Hamilton-Jacobi-Bellman Equations* (GHJBEs) *with respect to* $V_i^{(k+1)}$ $i = 1, \ \ldots , N$.

$$\begin{aligned}
&\frac{\partial V_i^{(k+1)T}}{\partial x} \hat{f}^{(k)}(x(t)) + \hat{Q}_{-i}^{(k)} \\
&- \frac{1}{4} \sum_{j=1, \ j \neq i}^{N} \frac{\partial V_j^{(k+1)T}}{\partial x} g_j R_{jj}^{-1} g_j^T \frac{\partial V_i^{(k)}}{\partial x} \\
&- \frac{1}{4} \sum_{j=1, \ j \neq i}^{N} \frac{\partial V_i^{(k)T}}{\partial x} g_j R_{jj}^{-1} g_j^T \frac{\partial V_j^{(k+1)}}{\partial x} \\
&- \frac{1}{2} \sum_{j=1, \ j \neq i}^{N} \frac{\partial V_j^{(k+1)T}}{\partial x} g_j R_{jj}^{-1} R_{ij} \mu_j^{(k)} \\
&- \frac{1}{2} \sum_{j=1, \ j \neq i}^{N} \mu_j^{(k)} R_{ij} R_{jj}^{-1} g_j^T \frac{\partial V_j^{(k+1)}}{\partial x} \\
&+ \frac{1}{2} r^T \frac{\partial^2 V_i^{(k+1)}}{\partial x^2} r = 0, \ V_i^{(k+1)}(0) = 0, \quad (30)
\end{aligned}$$

where

$$\hat{f}^{(k)}(x(t)) = f(x(t)) + \sum_{j=1}^{N} g_j(x(t)) \mu_j^{(k)}(t),$$

$$\begin{aligned}
\hat{Q}_{-i}^{(k)} &= \|x(t)\|_{Q_i}^2 - \frac{1}{2} \sum_{j=1, \ j \neq i}^{N} \mu_j^{(k)T}(t) g_j^T \frac{\partial V_i^{(k)}}{\partial x} \\
&- \frac{1}{2} \sum_{j=1, \ j \neq i}^{N} \frac{\partial V_i^{(k)T}}{\partial x} g_j \mu_j^{(k)}(t) \\
&+ \mu_i^{(k)T}(t) R_{ii} \mu_i^{(k)}(t) \\
&- \sum_{j=1, \ j \neq i}^{N} \mu_j^{(k)T}(t) R_{ij} \mu_j^{(k)}(t).
\end{aligned}$$

Step 3. *Compute* $u_i^{(k+1)}$.

$$u_i^{(k+1)} = \mu_i^{(k+1)*} = -\frac{1}{2} R_{ii}^{-1} g_i^T \frac{\partial V_i^{(k+1)}}{\partial x}. \quad (31)$$

Step 4. *Increment* $k \to k+1$ *and go to* **Step 2***, until the desired precision is attained simultaneously.*

It should be noted that it is still hard to solve the reduced-order HJBEs. Therefore, in order to overcome this drawback, the existing method based on the combination Galerkin spectral method and Chebyshev polynomials [13] would be helpful and informative.

If the linear case is considered, Newton's method is given below.

$$\begin{aligned}
&P_i^{(k+1)} A_i^{(k)} + A_i^{(k)T} P_i^{(k+1)} + Q_{-i}^{(k)} \\
&- \sum_{j=1, \ j \neq i}^{N} P_j^{(k+1)} S_j P_i^{(k)} - \sum_{j=1, \ j \neq i}^{N} P_i^{(k)} S_j P_j^{(k+1)} \\
&+ \sum_{j=1, \ j \neq i}^{N} [P_j^{(k)} G_{ij} P_j^{(k+1)} + P_j^{(k+1)} G_{ij} P_j^{(k)}] \\
&+ C^T P_i^{(k+1)} C = 0, \quad (32)
\end{aligned}$$

where $i = 1, \ldots, N$,

$$A_i^{(k)} = A - \sum_{j=1}^{N} S_j P_j^{(k)},$$

$$Q_{-i}^{(k)} = Q_i + \sum_{j=1,\ j\neq i}^{N} P_j^{(k)} S_j P_i^{(k)} + \sum_{j=1,\ j\neq i}^{N} P_i^{(k)} S_j P_j^{(k)}$$
$$+P_i^{(k)} S_i P_i^{(k)} - \sum_{j=1,\ j\neq i}^{N} P_j^{(k)} G_{ij} P_j^{(k)}.$$

It should be noted that the above mentioned algorithm is the same as the existing one [14] for the linear case.

V. NUMERICAL EXAMPLE

In order to demonstrate the efficiency of the proposed algorithm, a simple scalar example is given. The system and the cost functionals are given below.

$$dx(t) = \left[-2x(t) - \frac{1}{3}x^3 + u_1(t) + 2u_2(t)\right] dt$$
$$+0.1x(t)dw(t), \tag{33a}$$
$$J_1(x^0,\ u_1,\ u_2) = \mathbf{E}\left[\int_0^\infty (2\|x(t)\|^2 + \|u_1(t)\|^2)dt\right], \tag{33b}$$
$$J_2(x^0,\ u_1,\ u_2) = \mathbf{E}\left[\int_0^\infty (3\|x(t)\|^2 + \|u_2(t)\|^2)dt\right], \tag{33c}$$

where $R_{12} = R_{21} = 0$.

First, in order to obtain the initial strategies, the linear model of the stochastic system is established. The linearization model for the neighborhood $x(t) = 0$ as the equilibrium point is given below.

$$d\bar{x}(t) = [-2\bar{x}(t) + \bar{u}_1(t) + 2\bar{u}_2(t)]dt + 0.1\bar{x}(t)dw(t).$$

Therefore, the initial strategy set is given below by using of Newton's method (32).

$$u_1^{(0)}(t) = \mu_1^{(0)}(t) = -0.24982x(t), \tag{34a}$$
$$u_2^{(0)}(t) = \mu_2^{(0)}(t) = -0.94152x(t). \tag{34b}$$

On the other hand, using the proposed algorithm in section IV, the higher-order Nash strategy set is computed. Setting $k = 0$ at (30), the following equations can be obtained.

$$\frac{\partial V_1^{(1)}}{\partial x}\left(-2x - \frac{1}{3}x^3 + \frac{1}{2}\mu_1^{(0)} + 2\mu_2^{(0)}\right)$$
$$+2x^2 + 4\frac{\partial V_2^{(1)}}{\partial x}\mu_1^{(0)} + \frac{x^2}{200}\cdot\frac{\partial^2 V_1^{(1)}}{\partial x^2} + 4\mu_1^{(0)}\mu_2^{(0)} = 0,$$
$$V_1^{(1)} := \alpha_1 x^2,\ \alpha_1 > 0,$$
$$\frac{\partial V_2^{(1)}}{\partial x}\left(-2x - \frac{1}{3}x^3 + \mu_1^{(0)} + \mu_2^{(0)}\right)$$
$$+3x^2 + \frac{1}{2}\cdot\frac{\partial V_1^{(1)}}{\partial x}\mu_2^{(0)} + \frac{x^2}{200}\cdot\frac{\partial^2 V_2^{(1)}}{\partial x^2} + \mu_1^{(0)}\mu_2^{(0)} = 0,$$
$$V_2^{(1)} := \alpha_2 x^2,\ \alpha_2 > 0.$$

Finally, by continuing the same procedure, the following high-order approximate strategy set are calculated.

$$u_1^*(t) = -0.24968x(t) + O(x^2), \tag{35a}$$
$$u_2^*(t) = -0.94008x(t) + O(x^2). \tag{35b}$$

VI. CONCLUSIONS

Infinite horizon dynamic games for nonlinear stochastic system have been investigated. Particularly, Pareto and Nash strategy set were treated. It has been shown that the existing conditions were formulated by means of solvability of HJBE and HJBEs, respectively. Furthermore, these strategy sets can be obtained by solving HJBE and HJBEs. As compared with the existing result, a new recursive algorithm for solving HJBEs was given. As a result, the higher-order approximate strategy set can be obtained numerically. Finally, a numerical example was evaluated to validate and show the reliability of the proposed approach.

It is expected that the convergence proof of the proposed algorithm in section IV should be considered. Moreover, the avoidance of difficulty for carrying out of this algorithm should also be investigated. As other future investigations, the cost degradation is more interesting for using the approximate strategy set in the practical situation.

Although this paper investigates infinite horizon stochastic case, finite horizon case is more realistic. In this case, the condition that the closed-loop system is EMSS is not needed. Furthermore, in order to obtain the optimal control law of the finite horizon case, the algorithm that consists of the two essential steps in Four Step Scheme [17] for solving HJBE was developed. This literature would give another approximate strategy set. On the other hand, in order to adopt the difference between the approximate strategy and the optimal one, the hybrid controller via the neural networks [18] are suitable. The above mentioned these issues will be addressed in future investigations.

REFERENCES

[1] T. Basar, Nash equilibria of risk-sensitive nonlinear stochastic differential games, J. Optimization Theory and Applications, vol. 100, no. 3, 1999, pp. 479-498.

[2] M. Nourian, P. E. Caines, R. P. Malhamé and M. Huang Nash, social and centralized solutions to consensus problems via mean field control theory, IEEE Trans. Automatic Control, vol. 58, no. 3, 2013 (to appear).

[3] M. Nourian, P. E. Caines, R. P. Malhamé and M. Huang, Mean field LQG control in leader-follower stochastic multi-agent systems: Likelihood ratio based adaptation, IEEE Trans. Automatic Control, vol. 57, no. 11, pp. 2801-2816, 2012.

[4] M. Huang, P. E. Caines and R. P. Malhamé, Social optima in mean field LQG control: Centralized and decentralized strategies, IEEE Trans. Automatic Control, vol. 57, no. 7, pp. 1736-1751, 2012.

[5] M. Huang, P. E. Caines and R. P. Malhamé, Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behaviour and decentralized ε-Nash equilibria, IEEE Trans. Automatic Control, vol. 52, no. 9, pp. 1560-1571, 2007.

[6] W. Zhang, H. Zhang, and B. S. Chen, Stochastic H_2/H_∞ control with $(x,\ u,\ v)$-dependent noise: Finite horizon case, Automatica, vol. 42, no. 11, 2006, pp. 1891-1898.

[7] W. Zhang and G. Feng, Nonlinear stochastic H_2/H_∞ control with $(x,\ u,\ v)$-dependent noise: Infinite horizon case, IEEE Trans. Automatic Control, vol. 53, no. 5, 2008, pp. 1323-1328.

[8] K. G. Vamvoudakis and F. L. Lewis, Multi-player non-zero-sum games: Online adaptive learning solution of coupled Hamilton-Jacobi equations, Automatica, vol. 47, no. 8, 2011, pp. 1556-1569.
[9] R. Z. Has'minskii, *Stochastic Stability of Differential Equations*, Alphen, The Netherlands: Sijtjoff and Noordhoff, 1980.
[10] F. -Y. Wang and G. N. Saridis, On successive approximation of optimal control of stochastic dynamic systems, Int. Series in Operations Research & Management Science Volume 46, 2005, pp. 333-358.
[11] H. Mukaidani, Soft-constrained stochastic Nash games for weakly coupled large-scale systems, Automatica, vol. 45, no. 5, 2009, pp. 1272-1279.
[12] H. Mukaidani and H. Xu, Pareto optimal strategy for stochastic weakly coupled large scale systems with state dependent system noise, IEEE Trans. Automatic Control, vol. 54, no. 9, pp. 2244-2250, 2009.
[13] H. Mukaidani, Nonlinear stochastic H_2/H_∞ control with multiple decision makers, in Proc. 52nd Conf. Decision and Control, pp. 1186-1191, Florence, 2013.
[14] H. Mukaidani and V. Dragan, Stochastic Nash games for multimodeling systems, in Proc. American Control Conf., pp. 4404-4409, St. Louis, Missouri, 2009.
[15] H. Mukaidani and H. Xu, Pareto optimal strategy for stochastic weakly coupled large scale systems with state dependent system noise, IEEE Trans. Automatic Control, vol.54, no.9, pp. 2244-2450, 2009.
[16] J. C. Engwerda, *LQ Dynamic Optimization and Differential Games*, Chichester: John Wiley & Sons, 2005.
[17] J. Ma, P. Protter and J. Yong, Solving forward-backward stochastic differential equations explicitly -a four step scheme, Probability Theory and Related Fields, vol. 98, no. 3, pp. 339-359, 1994.
[18] Y. Iiguni, H. Sakai and H. Tokumaru, A nonlinear regulator design in the presence of system uncertainties using multi-layered neural networks, IEEE Trans. Neural Networks, vol. 2, no. 4, pp. 410-417, 1991.