International Journal of Innovative Computing, Information and Control
Volume **13**, Number **1**, February **2017**
 ISSN 1349-4198
pp. **55–66**

DECENTRALIZED OPTIMAL CONTROL FOR THE MEAN FIELD LQG PROBLEM OF MULTI-AGENT SYSTEMS

Fangfang Zhang[1] and Wei Wang[2]

[1]School of Electrical Engineering
Zhengzhou University
No. 100, Kexue Road, Zhengzhou 450001, P. R. China
zhangfangfang@zzu.edu.cn

[2]School of Control Science and Engineering
Shandong University
No. 17923, Jingshi Road, Jinan 250061, P. R. China
w.wang@sdu.edu.cn

Received June 2016; revised October 2016

Abstract. *This paper investigates the decentralized optimal control of the linear quadratic Gaussian (LQG) problem in discrete-time stochastic multi-agent systems. The state equations of the subsystems are uncoupled and the individual cost function is coupled with the states of other agents. With the help of state aggregation technique and the mean field structure, we get the decentralized optimal controllers that each agent only uses its own state and an iterative function which may be computed off-line for the optimization of the individual cost function and the social cost function respectively. And then, we prove that as the number of subsystems increases to infinite, the losses of the decentralized controllers and the optimal cost function for the two optimal control problems will go to zero due to the approximation in the optimization. At last, an illustrative example is given.*
Keywords: Decentralized control, Mean field, Discrete-time system, Multi-agent system, Linear-quadratic Gaussian

1. **Introduction.** In recent years, the analysis and control design of multi-agent systems (MAS) have become very popular in the control communities due to the broad applications in many fields. As a special type of MAS, large numbers of stochastic multi-agent systems have many practical examples in the field of engineering, biological, social and economic system, such as wireless sensor network [1], large scale robotics [2], flocking and swarming in the biological system [3,4], and resource sharing and competing in the Internet [5].

Agents or subsystems in multi-agent systems usually are coupled in dynamic or cost function. Agents have the ability of self-governed, but communication ability is limited. Moreover, each agent cannot interact with all others and it is impossible that each agent obtains all the states of other agents. Sometimes, the topological structure of the system cannot be obtained, so the decentralized optimal control that each agent only uses its own state is necessary for many practical situations. For the multi-agent systems, in some social, economic, and engineering models, the individuals or agents involved have conflicting objectives and it is natural to consider optimization based upon individual cost function [8]; many authors [6-9] develop game theoretic approaches and get some useful solutions. Huang et al. [8] consider the continuous-time LQG problems in large population systems, where the agents evolve according to nonuniform dynamics and are coupled via their individual costs. The state aggregation technique is developed to obtain the decentralized controllers. This method establishes the stability property of the mass behavior and possesses an ϵ-Nash equilibrium property. And in other papers [10-13], they study the social optimal LQG control problems in mean field for continuous-time systems and get the centralized and decentralized optimal control and give the Nash certainty equivalence (NCE) principle and social certainty equivalence (SCE) principle. The mean field LQG control for leader-follower stochastic multi-agent systems is considered in [14] and the ϵ_N-Nash equilibrium property for both leaders and adaptive followers is proved.

In those papers, they all study the mean field LQG control for continuous-time systems and derive the centralized and decentralized control strategies. Many authors ([15-17] and therein) consider the discrete-time multi-agent systems. You and Xie [15] study the linear discrete-time multi-agent systems and get a necessary and sufficient condition for consensusability under a common control protocol. And Ma et al. [16] also consider the linear quadratic decentralized dynamic games for large population discrete-time stochastic multi-agent systems. The paper [17] studies the linear quadratic regulation problem for finite horizon linear discrete time-varying systems with delay in control input. In the actual applications, discrete system with finite time is more practical and easy to be realized by computer. Compared with [15-17], the decentralized optimal control of the linear quadratic Gaussian (LQG) problem in discrete-time stochastic multi-agent systems with finite horizon are considered, where reference [15] only considers consensus problem for multi-agent system, [16] studies the decentralized optimal dynamic games for large population discrete-time stochastic systems with infinite horizon and reference [17] considers the optimal control problem for linear discrete time-varying deterministic systems with input delay. The main contributions of this paper are that the explicit forms of the decentralized optimal controllers are given and the asymptotic equivalence of the optimal controllers for the individual cost function and the social cost function are proved.

The remainder of this paper is organized as follows. Section 2 presents the LQG control problems in discrete-time multi-agent systems with finite time. In Section 3, we give decentralized optimal controllers for the individual cost function and social cost function. The analysis of optimal control is given in Section 4. An illustrative example is given in Section 5. Section 6 gives the conclusions of the paper.

The following notations will be used in this paper: $\|\cdot\|$ denotes 2-norm of vector, $X'(x')$ denotes the transpose of matrix (vector) $X(x)$, $tr(P)$ denotes the trace of matrix P.

2. **The LQG Control Problem in Discrete-time System.** We consider a linear stochastic system with N agents in discrete finite time. The dynamics of agent i is described by $(k = 0, 1, \cdots, T-1)$

$$x_i(k+1) = Ax_i(k) + Bu_i(k) + w_i(k), \tag{1}$$

where $x_i(\cdot)$ is n_1 dimensional state vectors and $u_i(\cdot)$ is n_2 dimensional control vectors. T is the finite time step number and A, B are matrices with the compatible dimensions. $w_i(k)$ is independent white noise and $E\{w_i(k)\} = 0$, $E\{w_i(k)w_j(l)\} = \Lambda\delta_{ij}\delta_{kl}$. The initial states $\{x_i(0), i = 1, 2, \cdots, N\}$ are independent with each other and $\{w_i(k)\}$ and have the mean value $m_i(0)$ and covariance matrix Θ. All subsystems are independent.

The individual cost function of agent i is

$$\begin{aligned} J_i = E\Bigg\{ \sum_{k=0}^{T-1} \Bigg\{ \Bigg[x_i(k) - \frac{1}{N}\sum_{j\neq i}^{N} x_j(k)\Bigg]' Q \Bigg[x_i(k) - \frac{1}{N}\sum_{i\neq j}^{N} x_j(k)\Bigg] + u_i(k)'Ru_i(k) \Bigg\} \\ + \Bigg[x_i(T) - \frac{1}{N}\sum_{j\neq i}^{N} x_j(T)\Bigg]' F \Bigg[x_i(T) - \frac{1}{N}\sum_{j\neq i}^{N} x_j(T)\Bigg] \Bigg\}, \end{aligned} \tag{2}$$

where Q and F are symmetric, positive semi-definite matrices and R is symmetric, positive definite matrix. Let $x_{-i}^N(k) = \frac{1}{N}\sum_{j\neq i}^N x_j(k)$, $k = 0, 1, \cdots, T$. We call $x_{-i}^N(\cdot)$ the mean field term of agent i. The agents interact with each other through the coupling terms of the mean field. This type of model is a dynamically independent and cost-coupled system [8].

Sometimes, the subsystems or agents in the systems are cooperative and have the common objectives. And the social cost function of the system is defined [10,12] as

$$J_{soc}^{(N)} = \sum_{i=1}^{N} J_i, \tag{3}$$

where J_i is the individual cost function for agent i.

Remark 2.1. *In the individual cost function (2), these agents are individually incentive driven and noncooperative. In order to get the individual decentralized control solution, the Nash certainty equivalence (NCE) principle has been designated and many solutions have been got [8,9,13,16,17]. In the social cost function (3), these agents are cooperative and have the common objectives. Minimizing $J_{soc}^{(N)}$, from the point of view of an individual's control selection, it is necessary to maintain a delicate balance in reducing its own cost and also taking account of the social impact of such reductions. The social certainty equivalence (SCE) principle has also been got and Pareto optimality has been used in [10,12].*

We will study the following two problems that each agent optimizes its own cost J_i (P0) and the social cost $J_{soc}^{(N)}$ (P1) and looks for the decentralized optimal control strategies. Comparing with our solution, the centralized solution for the social cost optimization will also be discussed.

3. Decentralized Optimal Control.

To facilitate further analysis, we denote $\mathscr{F}_k^N = \sigma(x_i(0), w_i(s), i = 1, \cdots, N, s = 0, 1, \cdots, k)$ for $k = 0, 1, \cdots, T-1$, which are the σ-algebra generated by the initial condition $x_i(0)$ and the white noise up to time k. Denote the control set $\mathscr{U}_0 = \{(u_1, \cdots, u_N) | u_i(k,\omega) \text{ is adapted to } \mathscr{F}_k^N, i = 1, \cdots, N\}$.

Firstly, for agent i, the problem (P0) is to find the optimal control $u_i^*(k) \in \mathscr{U}_0$ that minimizes the individual performance index J_i. Using the principle of optimality and the Bellman's dynamic programming, we start with the finial state and work backwards. Let $x^N(k) = 1/N \sum\limits_{i=1}^{N} x_i(k)$. When the number of agents is large, it is rational that $x_{-i}^N(k) \approx x^N(k)$ is approximated by a function $\bar{x}(k)$ for $i = 1, \cdots, N$ and its optimality loss will be negligible in large population conditions. Using the state aggregation procedure, assume $\bar{x}(k)$ is known and it will be determined later, we will solve a linear quadratic tracking problem. Then individual cost for the agent i is approximated as follows:

$$\begin{aligned} \bar{J}_i = E\Bigg\{ & \sum_{k=0}^{T-1} \Big\{ [x_i(k) - \bar{x}(k)]' Q [x_i(k) - \bar{x}(k)] + u_i(k)' R u_i(k) \Big\} \\ & + [x_i(T) - \bar{x}(k)]' F [x_i(T) - \bar{x}(k)] \Bigg\}. \end{aligned} \tag{4}$$

For the agent i, using the solution of the linear tracking problem for the finite discrete-time stochastic system, we have the following theorem.

Theorem 3.1. *For $k = 0, 1, \cdots, T-1$, the decentralized optimal control of the state equation (1) with the approximate cost function (4) is*

$$u_i^*(k) = -\left(R + B' P_{k+1} B\right)^{-1} \left[B' P_{k+1} A x_i^*(k) - B' P_{k+1} s(k+1)\right], \tag{5}$$

where

$$P_k = Q - A' P_{k+1} B \left(R + B' P_{k+1} B\right)^{-1} B' P_{k+1} A + A' P_{k+1} A, \tag{6}$$

$$P_T = F, \tag{7}$$

if $I - B\left(R + B' P_{k+1} B\right)^{-1} B' P_{k+1}$ *is invertible*

$$s(k+1) = As(k), \tag{8}$$

$$s(0) = \frac{1}{N} \sum_{j=1}^{N} m_j(0). \tag{9}$$

Then the optimal cost function is

$$\bar{J}_i^* = [m_{i0} - s(0)]' P_0 [m_{i0} - s(0)] + tr\left\{P_0 \Theta + \sum_{k=0}^{T-1} P_{k+1} \Lambda\right\}. \tag{10}$$

Proof: We optimize from the finial state. For the stage $T-1$, the cost function of agent i is

$$\begin{aligned}
&\bar{J}_i(u(T-1)) \\
&= E\Big\{[x_i(T-1) - \bar{x}(T-1)]' Q [x_i(T-1) - \bar{x}(T-1)] + u_i(T-1)' R u_i(T-1) \\
&\quad + [x_i(T) - \bar{x}(T)]' F [x_i(T) - \bar{x}(T)]\Big\} \\
&= E\Big\{E\Big\{[x_i(T-1) - \bar{x}(T-1)]' Q [x_i(T-1) - \bar{x}(T-1)] + u_i(T-1)' R u_i(T-1) \\
&\quad + [x_i(T) - \bar{x}(T)]' F [x_i(T) - \bar{x}(T)]\Big\} \,|\mathscr{F}_{T-1}\Big\} \\
&= E\Big\{[x_i(T-1) - \bar{x}(T-1)]' Q [x_i(T-1) - \bar{x}(T-1)] + u_i(T-1)' R u_i(T-1) \\
&\quad + [Ax_i(T-1) + Bu_i(T-1) - \bar{x}(T)]' F [Ax_i(T-1) + Bu_i(T-1) - \bar{x}(T)]\Big\} + tr(F\Lambda).
\end{aligned}$$

Using the maximal principle, we can get the optimal control

$$u_i^*(T-1) = -\left(R + B' P_T B\right)^{-1} \left[B' P_T A x_i^*(T-1) - B' P_T s(T)\right], \tag{11}$$

where $P_T = F, s(T) = \bar{x}(T)$ and

$$P_{T-1} = Q - A' P_T B \left(R + B' P_T B\right)^{-1} B' P_T A + A' P_T A, \tag{12}$$

$$s(T-1) = \left(A' - A' P_T B \left(R + B' P_T B\right)^{-1} B'\right) s(T) + Qz^*(T-1), \tag{13}$$

and if $I - B\left(R + B' P_T B\right)^{-1} B' P_T$ is invertible (we will prove it later), we can get the following equation from Equations (11), (12), (13) and the definition of $\bar{x}(T)$

$$\begin{aligned}
s(T) &= \left[I - B\left(R + B' P_T B\right)^{-1} B' P_T\right]^{-1} \left[A - B\left(R + B' P_T B\right)^{-1} B' P_T A\right] s(T-1) \\
&= As(T-1).
\end{aligned} \tag{14}$$

Substituting the optimal control $u_i^*(T-1)$ into the cost function, by (12), we get

$$\bar{J}_i^*(T-1) = E\left\{[x_i(T-1) - \bar{x}(T-1)]' P_{T-1} [x_i(T-1) - \bar{x}(T-1)]\right\} + tr(P_T \Lambda). \tag{15}$$

The finial time $T-1$ is chosen arbitrarily, using mathematical induction, Equations (5)-(8) are right for any k, $k=0,1,\cdots,T-1$. For $k=0$, we have

$$u_i^*(0) = -\left(R + B'P_1B\right)^{-1}\left[B'P_1Ax_i(0) - B'P_Ts(1)\right], \tag{16}$$

$$P_0 = Q + A'P_1A - A'P_1B\left(R + B'P_1B\right)^{-1}B'P_1A, \tag{17}$$

$$s(1) = As(0), \tag{18}$$

$$s(0) = E\left\{\frac{1}{N}\sum_{j=1}^{N}x_j(0)\right\} = \frac{1}{N}\sum_{j=1}^{N}m_j(0). \tag{19}$$

And the finial optimal cost function is

$$\bar{J}_i^* = [m_{i0} - s(0)]'P_0[m_{i0} - s(0)] + tr\left\{P_0\Theta + \sum_{k=1}^{T}P_k\Lambda\right\}. \tag{20}$$

And then, we will prove that $I - B\left(R + B'P_{k+1}B\right)^{-1}B'P_{k+1}$, $k = 0,1,\cdots,T-1$ are invertible. It follows that R is invertible because R is positive definite matrix, let $W(k+1) = I - B\left(R + B'P_{k+1}B\right)^{-1}B'P_{k+1}$ and $V(k+1) = I + BR^{-1}B'P_{k+1}$, then we have

$$W(k+1)V(k+1) = V(k+1)W(k+1) = I,$$

therefore $W(k+1)$ are invertible. The proof of the theorem is completed. □

For the agent i, applying the optimal control law, the closed loop Equation (1) becomes

$$\begin{aligned} x_i^*(k+1) = &\left(A - B\left(R + B'P_{k+1}B\right)^{-1}B'P_{k+1}\right)x_i^*(k) \\ &+B\left(R + B'P_{k+1}B\right)^{-1}B'P_{k+1}s(k+1) + w_i(k),\ k = 0,1,\cdots,T-1. \end{aligned} \tag{21}$$

Denoting $\overline{x}_i(k) = Ex_i(k)$ and taking expectation on both sides of Equation (21) gives

$$\begin{aligned} \overline{x}_i^*(k+1) = &\left(A - B\left(R + B'P_{k+1}B\right)^{-1}B'P_{k+1}\right)\overline{x}_i^*(k) \\ &+B\left(R + B'P_{k+1}B\right)^{-1}B'P_{k+1}s(k+1), \end{aligned} \tag{22}$$

where $\overline{x}_i(0) = Ex_i(0) = m_i(0)$, and P_k are the solutions of the difference Riccati equation (6) and $s(k)$ are the solutions of the iterative equation (8).

Remark 3.1. *From (5), we know that the decentralized optimal controllers only depend on its own state and an iterative equation which can be solved off-line. The optimal state equations are given by (21).*

In the front of this section, the individuals or agents involved in the large systems have self-governed objectives and are individually incentive driven and noncooperative. Each agent optimizes its own cost function and does not consider the impact on other agents. Now we will study a different situation where the agents are cooperative and have a common objection. For minimizing the social cost $J_{soc}^{(N)}$, we also consider the decentralized strategies where each agent only uses local information. Let the set of the socially optimal controls be denoted by $(\hat{u}_1(\cdot),\cdots,\hat{u}_N(\cdot)) \in \mathscr{U}_0$. Suppose the state $\hat{x}_i(k)$ corresponds to $\hat{u}_i(k)$ and the admissible control set $\mathscr{U}_{0i}$ consists of all $u_i(k)$ adapted to $\mathscr{F}_k^N$. Then $\hat{u}_i(\cdot)$ is the unique optimal control for the following control problem (P2) [10]:

$$x_i(k+1) = Ax_i(k) + Bu_i(k) + w_i(k), \tag{23}$$

$$J_i^0(u_i(\cdot)) = J_{soc}^N(\hat{u}_1(\cdot), \cdots, \hat{u}_{i-1}(\cdot), u_i(\cdot), \hat{u}_{i+1}(\cdot), \cdots, \hat{u}_N(\cdot)), \tag{24}$$

for $k = 0, 1, \cdots, T-1$, where $J_i^0(u_i(\cdot))$ is to be minimized.

Due to the mean field coupling of the states in the J_{soc}^N, for any agent i, we give the reformulation of the cost function of the problem (P2). Since $(\hat{u}_1(\cdot), \cdots, \hat{u}_N(\cdot)) \in \mathscr{U}_0$. We know $(\hat{u}_1(\cdot), \cdots, \hat{u}_{i-1}(\cdot), \hat{u}_{i+1}(\cdot), \cdots, \hat{u}_N(\cdot))$ and $\hat{x}_j(\cdot), j \neq i$ have been specified in advance and will not change with $u_i(\cdot)$. After a simple calculation, we have

$$J_i^0(u_i) = E\left\{\sum_{k=0}^{T-1}\Big[L_1(x_i(k), \hat{x}_{-i}(k)) + u_i(k)' R u_i(k)\Big] + L_2\left(x_i(T), \hat{x}_{-i}(T)\right)\right\}, \tag{25}$$

where

$$\begin{aligned}
&L_1(x_i(k), x_{-i}(k)) \\
&= \frac{N^2+N-1}{N^2} x_i(k)' Q x_i(k) - \frac{N+2}{N} x_i(k)' Q \hat{x}_{-i}(k) \\
&\quad - \frac{N+2}{N}(\hat{x}_{-i}(k))' Q x_i(k) + M(\hat{x}_{-i}(k)),
\end{aligned}$$

$$\begin{aligned}
&L_2(x_i(T), x_{-i}(T)) \\
&= \frac{N^2+N-1}{N^2} x_i(T)' F x_i(T) - \frac{N+2}{N} x_i(T)' F \hat{x}_{-i}(T) \\
&\quad - \frac{N+2}{N}(\hat{x}_{-i}(T))' F x_i(T) + \bar{M}(\hat{x}_{-i}(T)).
\end{aligned}$$

Here, $M(\hat{x}_{-i}(k))$ and $\bar{M}(\hat{x}_{-i}(T))$ denote the terms not depending on $u_i(k)$ and $u_i(T)$. When N is sufficiently large, we also can approximate the term $\hat{x}_{-i}(k)$ by a deterministic function $\hat{\bar{x}}(k)$. Then the cost function (25) can be approximated by

$$\bar{J}_i^0(u_i) = E\left\{\sum_{k=0}^{T-1}\Big[\bar{L}_1\left(x_i(k), \hat{\bar{x}}(k)\right) + u_i(k)' R u_i(k)\Big] + \bar{L}_2\left(x_i(T), \hat{\bar{x}}(T)\right)\right\}, \tag{26}$$

where

$$\begin{aligned}
&\bar{L}_1\left(x_i(k), \hat{\bar{x}}(k)\right) \\
&= \left(1 + \frac{N-1}{N^2}\right) x_i(k)' Q x_i(k) - \frac{N+2}{N} x_i(k)' Q \hat{\bar{x}}(k) - \frac{N+2}{N}(\hat{\bar{x}}(k))' Q x_i(k),
\end{aligned}$$

$$\begin{aligned}
&\bar{L}_2\left(x_i(T), \hat{\bar{x}}(T)\right) \\
&= \left(1 + \frac{N-1}{N^2}\right) x_i(T)' F x_i(T) - \frac{N+2}{N} x_i(T)' F \hat{\bar{x}}(T) - \frac{N+2}{N}(\hat{\bar{x}}(T))' F x_i(T).
\end{aligned}$$

Minimizing the cost (26) with the sate equation (1), we have the following theorem.

Theorem 3.2. *For $k = 0, 1, \cdots, T-1$ and any agent i, the decentralized optimal controllers are*

$$\begin{aligned}
\hat{u}_i(k) = &-\left[\left(1 + \frac{N-1}{N^2}\right) B' \hat{P}_{k+1} B + R\right]^{-1} \left[\left(1 + \frac{N-1}{N^2}\right) B' \hat{P}_{k+1} A x_i(k) \right. \\
&\left. - \frac{N+2}{N} B' \hat{P}_{k+1} s(k+1)\right],
\end{aligned} \tag{27}$$

where

$$\hat{P}_k = Q + A' \hat{P}_{k+1} A$$

$$-\left(1+\frac{N-1}{N^2}\right)A'\hat{P}_{k+1}B\left(R+\left(1+\frac{N-1}{N^2}\right)B'\hat{P}_{k-1}B\right)^{-1}B'\hat{P}_{k+1}A, \quad (28)$$

$$\hat{P}_T = F, \quad (29)$$

and if $I-\left(1+\frac{N-1}{N^2}\right)B\left[R+\left(1+\frac{N-1}{N^2}\right)B'\hat{P}_{k+1}B\right]^{-1}B'\hat{P}_{k+1}$ *is invertible, then*

$$s(k+1) = As(k), \quad (30)$$

where

$$s(0) = \frac{1}{N}\sum_{j=1}^{N} m_j(0). \quad (31)$$

And the social optimal cost function for agent i *is*

$$\bar{J}_i^{0*}(\hat{u}_i) = [m_{i0}-s(0)]'\hat{P}_0[m_{i0}-s(0)] + tr\left\{\hat{P}_0\Theta + \sum_{k=0}^{T-1}\hat{P}_{k+1}\Lambda\right\}. \quad (32)$$

The proof of this theorem is similar with Theorem 3.1, and we omit it.

Remark 3.2. *We should point out that* $I-(1+\frac{N-1}{N^2})B[R+(1+\frac{N-1}{N^2})B'\hat{P}_{k+1}B]^{-1}B'\hat{P}_{k+1}$, $k=0,1,\cdots,T-1$ *are invertible. Let* $\hat{W}(k+1) = I-(1+\frac{N-1}{N^2})B[R+(1+\frac{N-1}{N^2})B'\hat{P}_{k+1}B]^{-1}B'\hat{P}_{k+1}$, *and* $\hat{V}(k+1) = I+(1+\frac{N-1}{N^2})BR^{-1}B'\hat{P}_{k+1}$, *and we can get*

$$\hat{W}(k+1)\hat{V}(k+1) = \hat{V}(k+1)\hat{W}(k+1) = I,$$

so $\hat{W}(k+1)$, $k=0,1,\cdots,T-1$ *are invertible.*

Remark 3.3. *It follows from (6) and (28) that the two Riccati equations are equivalent when* $N\to\infty$ *because* $\lim_{N\to\infty}(1+\frac{N-1}{N^2}) = 1$. *As* N *is large, the difference between the solutions of the two Riccati equations is also very small. From Theorems 3.1 and 3.2, by* $\lim_{N\to\infty}\frac{N-2}{N} = 1$, *we know that the decentralized optimal controllers and the optimal cost function of the individual optima and social optima are the same when* $N\to\infty$. *It can be seen from Theorem 3.2 that the decentralized optimal controller of the social optimal control for any agent* i *is related to the agent number of the systems.*

Remark 3.4. *The Riccati equations and the optimal controllers have the same forms with the deterministic optimal control, but the optimal cost function is different with the deterministic optimal control in Theorems 3.1 and 3.2 [18]. The optimal cost function has not only the quadratic forms of the initial values but also the stochastic quantity of the initial states and the noise of the system [19]. So the optimal cost function is bigger than the value of the deterministic optimal control.*

4. **Optimality Analysis.** In the above computation of the optimal control, $x_{-i}(k)$ is approximated by $\bar{x}(k)$, $\hat{x}_{-i}(k)$ is approximated by $\hat{\bar{x}}(k)$, for $i=1,2,\cdots,N$, $k=0,1,\cdots,T$, i.e.,

$$x_{-i}^N(k) = \sum_{j\neq i}^{N} x_j(k) \approx \sum_{j=1}^{N} x_j(k) = x^N(k), \quad (33)$$

$$\hat{x}_{-i}^N(k) = \sum_{j\neq i}^{N} \hat{x}_j(k) \approx \sum_{j=1}^{N} \hat{x}_j(k) = \hat{x}^N(k). \quad (34)$$

They will cause the error in the computation for the optimal computer of every step. Then we have the following theorem.

Theorem 4.1. *The error of the social cost function for agent i satisfies*

$$\lim_{N\to\infty} \delta J_{soc}(u_i(k)) = 0. \tag{35}$$

Proof: For any step k, $k = 0, 1, \cdots, T-1$, assume $\|x_i(k)\| < M$ for $i = 1, 2, \cdots, N$, M is a constant, and the error of individual cost function for agent i is

$$\begin{aligned}
\delta J_i(u_i(k)) &= \left|\bar{J}_i(u_i(k)) - J_i(u_i(k))\right| \\
&= \left|-\frac{1}{N}[x_i(k) - \bar{x}(k)]' Q x_i(k) - \frac{1}{N} x_i(k)' Q [x_i(k) - \bar{x}(k)] - \frac{1}{N^2} x_i(k)' Q x_i(k)\right| \\
&= \left|\frac{2}{N}(x_i(k) - \bar{x}(k))' Q x_i(k) + \frac{1}{N^2} x_i(k)' Q x_i(k)\right| \\
&\leq \frac{4}{N} M tr(Q) + \frac{1}{N^2} M^2 tr(Q).
\end{aligned}$$

Then we have

$$\lim_{N\to\infty} \delta J_i(k) = 0. \tag{36}$$

The error of the social cost function for agent i is

$$\begin{aligned}
\delta J_i^0(u_i(k)) &= \left|\bar{J}_i^0(u_i(k)) - J_i^0(u_i(k))\right| = \left|2\frac{N+2}{N^2} x_i(k)' Q x_i(k)\right| \\
&\leq 2\frac{N+2}{N^2} M^2 tr(Q).
\end{aligned}$$

When $N \to \infty$, we get

$$\lim_{N\to\infty} \delta J_{soc}(u_i(k)) = 0. \tag{37}$$

The proof of the theorem is completed. □

Remark 4.1. *It follows from Equations (36) and (37) that the losses of both individual cost and social cost for each agent caused by approximation of state aggregation are insignificant if the number of agent is large.*

Next, we consider the scalar model of the systems. Minimizing the social cost function $J_{soc}^{(N)}(u(\cdot))$ for the systems, the centralized optimal control can be obtained. Let $A = a$, $B = b$, $Q = q$, $R = r$ and $F = f$. Then the state equation for the system and the social cost function become

$$x(k+1) = \tilde{A}x(k) + \tilde{B}u(k) + w(k), \tag{38}$$

$$J_{soc}^{(N)}(u(\cdot)) = E\left\{\sum_{k=d}^{T-1}\left(x'(k)\tilde{Q}x(k) + u'(k)\tilde{R}u(k)\right) + x'(T)\tilde{F}x(T)\right\}, \tag{39}$$

where $x(k) = (x_1(k), x_2(k), \cdots, x_N(k))$, $u(k) = (u_1(k), u_2(k), \cdots, u_N(k))$, $w(k) = (w_1(k), w_2(k), \cdots, w_N(k))$, $\tilde{A} = aI_N$, $\tilde{B} = bI_N$, $\tilde{R} = rI_N$, $\tilde{Q} = (\tilde{q}_{ij})$, $\tilde{F} = (\tilde{f}_{ij})$, and

$$\tilde{q}_{ii} = \left(1 + \frac{N-1}{N^2}\right)q, \quad \tilde{q}_{ij} = -\frac{N+2}{N^2}q, \quad i \neq j,$$

$$\tilde{f}_{ii} = \left(1 + \frac{N-1}{N^2}\right)f, \quad \tilde{f}_{ij} = -\frac{N+2}{N^2}f, \quad i \neq j.$$

Minimizing the social cost function (39) under the state Equation (38), we can get the optimal controllers, for $k = 0, 1, \cdots, T-1$,

$$u_c^*(k) = -ab\left(rI_N + b^2\tilde{P}_{k+1}\right)^{-1}\tilde{P}_{k+1}x^*(k), \tag{40}$$

where

$$\tilde{P}_k = a^2\tilde{P}_{k+1} - a^2b^2\tilde{P}_{k+1}\left(rI_N + b^2\widetilde{P}_{k+1}\right)^{-1}\tilde{P}_{k+1} + \tilde{Q}, \tag{41}$$

$$\tilde{P}_T = \tilde{F}. \tag{42}$$

From Theorem 3.2, we can obtain the optimal controllers and Riccati equation of the scaler model for minimizing the social cost function, $k = 0, 1, \cdots, T-1$

$$\hat{u}_i(k) = \frac{1}{\left(1+\frac{N-1}{N^2}\right)b^2\tilde{p}_{k+1} + r}\left[\left(1+\frac{N-1}{N^2}\right)ab\tilde{p}_{k+1}x_i(k) - \frac{N+2}{N}b\tilde{p}_{k+1}s(k+1)\right], \tag{43}$$

where

$$\tilde{p}_k = q + a^2\tilde{p}_{k+1} - \frac{\left(1+\frac{N-1}{N^2}\right)a^2b^2\tilde{p}_{k+1}}{r + \left(1+\frac{N-1}{N^2}\right)b^2\tilde{p}_{k+1}}, \tag{44}$$

$$s(k+1) = as(k), \quad \tilde{p}_T = f. \tag{45}$$

Theorem 4.2. *It follows from Equations (40)-(45) that the decentralized optimal controllers for any agent i are the same with the centralized optimal controllers when $N \to \infty$.*

The proof of this theorem is simple and straight-forward, and we omit it. □

Then we analyze the difference of the solutions of the Riccati equations (6) and (28) for scalar model. For any k, suppose $m_1 < p_k < M_1, m_1 < \hat{p}_k < M_2$,

$$\begin{aligned} & p_k - \hat{p}_k \\ = & \frac{a^2rp_{k+1}}{r + b^2p_{k+1}} - \frac{a^2r\hat{p}_{k+1}}{r + \left(1+\frac{N-1}{N^2}\right)b^2\hat{p}_{k+1}} \\ = & \frac{a^2r^2\left(p_{k+1} - \hat{p}_{k+1}\right) + \frac{N-1}{N^2}a^2b^2rp_{k+1}\hat{p}_{k+1}}{\left(r + b^2p_{k+1}\right)\left(r + \left(1+\frac{N-1}{N^2}\right)b^2\hat{p}_{k+1}\right)}, \end{aligned}$$

it follows that

$$|p_k - \hat{p}_k| < \frac{a^2r^2}{(r+b^2m_1)(r+b^2m_2)}|p_{k+1} - \hat{p}_{k+1}| + \frac{\frac{N-1}{N^2}a^2b^2rM_1M_2}{(r+b^2m_1)(r+b^2m_2)}. \tag{46}$$

Because $p_{k+1} > 0$, $\tilde{p}_{k+1} > 0$, $r > 0$ and $p_T = \tilde{p}_T = f$, $\forall\varepsilon > 0$, if N is sufficient large and $|p_{k+1} - \tilde{p}_{k+1}| < \delta$, then $|p_k - \tilde{p}_k| < \varepsilon$. Therefore, by (46), we can get that $|p_k - \tilde{p}_k| = 0$ if $N \to \infty$ for $k = 0, 1, \cdots, T$. From (5) and (27), the optimal controllers $\hat{u}_i(k) = u_i^*(k)$ if $N \to \infty$. We also point out that the optimal individual cost function and social cost function for any agent i are the same when $N \to \infty$ because $p_k = \tilde{p}_k$ for $k = 0, 1, \cdots, T$.

Remark 4.2. *By the analysis of this section, if the number of agent is large, the differences of the optimal controllers and the optimal cost function of problem (P0) and (P1) are small. So in the computation of the decentralized optimal problems, we can only optimize the individual cost function and do not consider the impact on other agents if the number of agents is very large.*

5. **An Illustrative Example.**

We now provide one scaler model example to explain the solutions of Sections 3 and 4. Assume the system has 50 agents. For simplicity, let the system matrices be $A = 0.9$, $B = 1$ and the weighted matrices of the cost function be $Q = 1$, $R = 1.5$, $F = 2$ for agent i, $i = 1, \cdots, 50$. Suppose the initial states of all agents are given by $\mathcal{N}(0, 1)$, where $\mathcal{N}(0, 1)$ is the standard normal distribution. Without loss of generality, assume the covariance matrix of the white noise $\Lambda = 1$ and the time step number $T = 30$.

Based on the solutions of Theorems 3.1 and 3.2, for any agent i, we can get the decentralized optimal controller and optimal closed-loop state equation. The expectation of optimal states and the decentralized optimal controllers for any agent i are given in Figure 1, where $x1 = Ex_i^*(k)$, $x2 = E\hat{x}_i(k)$, $u1 = u_i^*(k)$, $u2 = \hat{u}_i$, $x3 = E\hat{x}_i(k) - Ex_i^*(k)$ and $u3 = \hat{u}_i - x_i^*(u)$. From this figure, we know that the difference of the expectation of optimal states and the difference of the optimal controllers gradually tend to 0 as the increase of k. By Figure 2, it is known that the relative error $dx(k)$ is only 2% for any agent i, where $dx(k) = [E\hat{x}_i(k) - Ex_i^*(k)]/Ex_i^*(k)$. Figure 3 shows different curves of the difference of the expectation of optimal states for different agent number N, where the number of the agents $N = 30, 50, 100$. It can be seen that the differences of the expectation of optimal states increase as the increase of k and decrease as the increase of N. By this example, we know that the states of the decentralized optimal control tend to be stable fast and the difference of the optimal states for individual optimal control and social optimal control is also small if the number of agents is large.

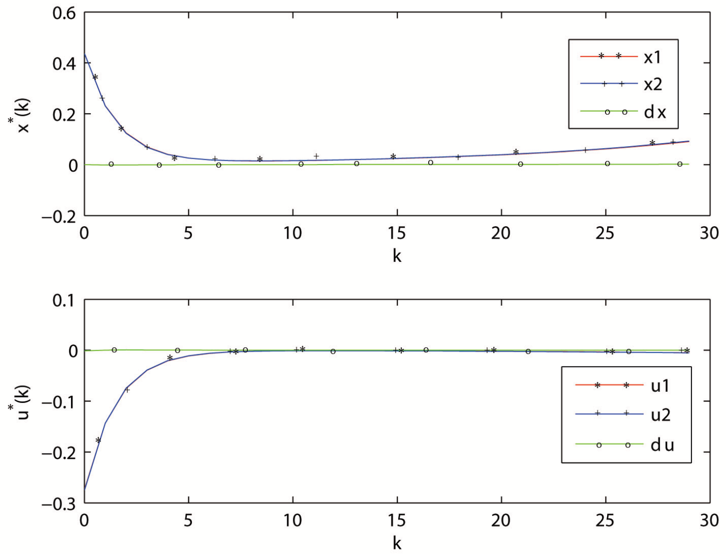

FIGURE 1. The curve of the optimal states and controllers

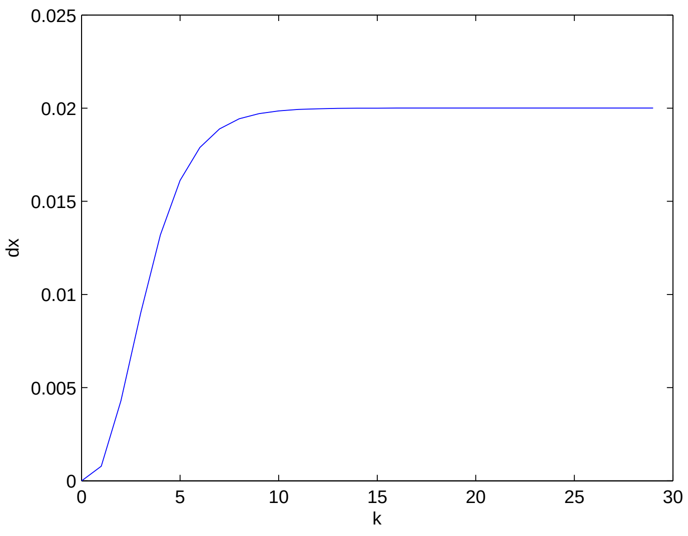

FIGURE 2. The curve of relative error of the optimal states

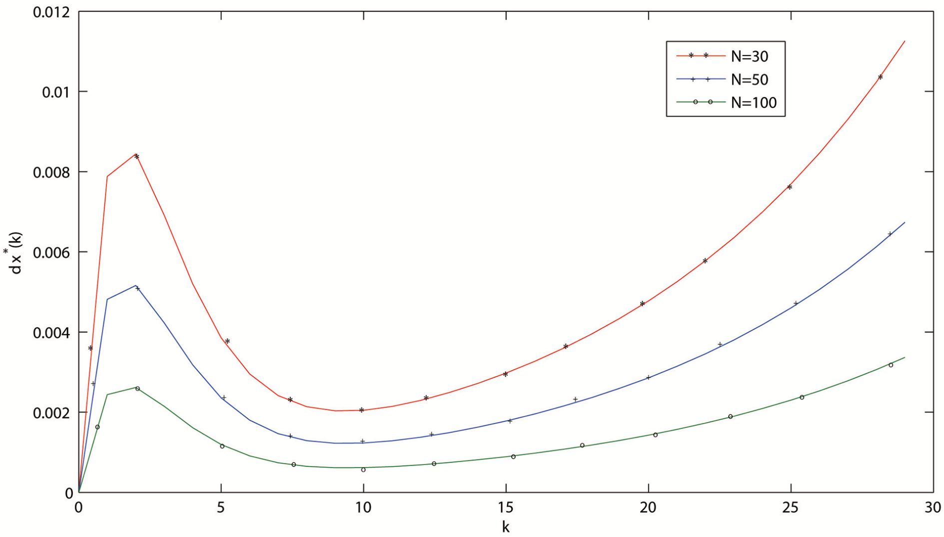

FIGURE 3. The curve of the optimal states for the different number N

6. Conclusions.

In this paper, the mean field LQG optimal control problem in discrete-time for large population stochastic system is considered. An approach of state aggregation is applied and the decentralized optimal controllers for individual optima and social optima are got. We get that the two solutions of the Riccati equations, the optimal controllers, and the cost function for each agent are the same as $N \to \infty$. We also prove that the losses of the individual cost function and the social cost function for each agent are equal to 0 when $N \to \infty$. We should point out that the approach may not work if the state equations or the cost function have different forms, and other approaches should be searched of the problem in the future.

Acknowledgment. This work is supported by the National Natural Science Foundation of China under Grants 61120106011, 61573220, and 61603345, the General Financial Grant from the China Postdoctoral Science Foundation (Grant No. 2012M582202), the Initial Scientific Research Fund of Young Teachers in Zhengzhou University (Grant No. F0000593), and the Key Scientific Research Projects of Henan Province (Grant No. 17A413003).

REFERENCES

[1] C. Y. Chong, S. P. Kumar and B. A. Hamilton, Sensor networks: Evolution, opportunities, and challenges, *Proc. of the IEEE*, vol.91, no.8, pp.1247-1256, 2003.
[2] J. H. Reif and H. Y. Wang, Social potential fields: A distributed behavioral control for autonomous robots, *Robotics and Autonomous Systems*, vol.27, pp.171-194, 1999.
[3] D. Grbaum and A. Okubo, Modeling social animal aggregations, *Frontiers in Theoretical Biology*, vol.100, pp.296-325, 1994.
[4] K. M. Passino, Biomimicry of bacterial foraging: For distributed optimization and control, *IEEE Control Systems Magazine*, vol.22, no.3, pp.52-67, 2002.
[5] E. Altman, T. Boulogne, R. El-Azouzi, T. Jimez and L. Wynter, A survey on networking games in telecommunications, *Computers & Operations Research*, vol.33, no.2, pp.286-311, 2006.
[6] Z. Dziong and L. G. Mason, Fair-efficient call admission control policies for broad-band networks – A game theoretic framework, *IEEE/ACM Transactions on Networking*, vol.4, no.1, pp.123-136, 1996.
[7] D. Fudenberg and D. K. Levine, *The Theory of Learning in Games*, MIT Press, Cambridge, MA, 1998.

[8] M. Y. Huang, P. E. Caines and R. P. Nash, Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ε Nash equilibria, *IEEE Transactions on Automatic Control*, vol.52, no.9, pp.1560-1571, 2007.
[9] M. Y. Huang, P. E. Caines and R. P. Nash, Certainty equivalence in large population stochastic dynamic games: Connections with the physics of interacting particle systems, *The 45th IEEE Conference on Decision and Control*, pp.4921-4926, 2006.
[10] M. Y. Huang, P. E. Caines and R. P. Nash, Social optima in mean field LQG control: Centralized and decentralized strategies, *The 47th Annual Allerton Conference on Communication, Control, and Computing*, Allerton, pp.322-329, 2009.
[11] M. Y. Huang, P. E. Caines and R. P. Malhamé, The NCE (mean field) principle with locality dependent cost interactions, *IEEE Transactions on Automatic Control*, vol.55, no.12, pp.2799-2805, 2010.
[12] M. Y. Huang, P. E. Caines and R. P. Malhamé, Social optima in mean field LQG control: Centralized and decentralized strategies, *IEEE Transactions on Automatic Control*, vol.57, no.7, pp.1736-1751, 2012.
[13] T. Li and J. F. Zhang, Asymptotically optimal decentralized control for large population stochastic multiagent systems, *IEEE Transactions on Automatic Control*, vol.53, no.7, pp.1643-1660, 2008.
[14] M. Nourian, P. E. Caines, R. P. Malhame et al., Mean field LQG control in leader-follower stochastic multi-agent systems: Likelihood ratio based adaptation, *IEEE Transactions on Automatic Control*, vol.57, no.11, pp.2801-2816, 2012.
[15] K. Y. You and L. H. Xie, Network topology and communication data rate for consensusability of discrete-time multi-agent systems, *IEEE Transactions on Automatic Control*, vol.56, no.10, pp.2262-2275, 2011.
[16] C. Q. Ma, T. Li and J. F. Zhang, Linear quadratic decentralized dynamic games for large population discrete-time stochastic multi-agent systems, *Journal of Systems Science and Mathematical Sciences*, vol.27, no.3, pp.464-480, 2007.
[17] J. Zhou and Q. Wang, Finite horizon linear quadratic regulation for linear discrete time-varying systems with single/multiple input delay(s), *Proc. of the 48th IEEE Conference on Decision and Control, 2009 Held Jointly with the 28th Chinese Control Conference*, Shanghai, China, pp.4050-4055, 2009.
[18] D. S. Naidu, *Optimal Control System*, CRC Press, Idaho State University, Pocatello, ID, USA, 2002.
[19] S. L. Guo, *Stochastic Control*, Tsinghua University Press, China, 1999.