SIAM J. CONTROL OPTIM.
Vol. 50, No. 4, pp. 2308–2334

MEAN FIELD GAMES FOR LARGE-POPULATION MULTIAGENT SYSTEMS WITH MARKOV JUMP PARAMETERS*

BING-CHANG WANG† AND JI-FENG ZHANG†

Abstract. In this paper, distributed games for large-population multiagent systems with random time-varying parameters are investigated, where the agents are coupled via their individual costs and the structure parameters are a family of independent Markov chains with identical generators. The cost function of each agent is a long-run average tracking-type functional with an unknown mean field coupling nonlinear term as "reference signal." To reduce the computational complexity, the mean field approach is applied to construct distributed strategies. The population statistics effect (PSE) is used to approximate the average effect of all the agents, and the distributed strategies are given through solving a Markov jump tracking problem. Here the PSE is a deterministic quantity and can be obtained by solving the Stackelberg equilibrium of an auxiliary two-player game. It is shown that the closed-loop system is uniformly stable, and the distributed strategies are asymptotically optimal in the sense of Nash equilibrium, as the number of agents grows to infinity. A numerical example is provided to demonstrate the procedure of designing the strategies as well as the influence of the heterogeneity intensity and the parameter jump rate of the agents on the closed-loop system.

Key words. mean field approach, multiagent system, distributed game, Markov jump parameter, optimal control

AMS subject classifications. 91A10, 91A23, 93E20, 68M14

DOI. 10.1137/100800324

1. Introduction.

In recent years, there has been a drastically increasing interest in control and optimization of multiagent systems (MASs) among the system and control community. The optimization problem of MASs in game-theoretic framework, though not investigated extensively, has attracted the attention of many researchers [3, 16, 18, 24, 25, 29, 32]. In this paper, we will study the problems of distributed games for large-population MASs, which are widely existent in economics, engineering, biology, and social sciences, for instance, production output planning in oligopoly markets [21], dynamic models of advertising competition [9, 10], wireless communication networks [6, 15], swarming and flocking [31, 13, 8], and voluntary vaccination games [2].

Mean field (MF) approaches play an important role in diverse fields of physics and chemistry (e.g., the derivation of Boltzman or Vlasov equations in the kinetic gas theory). Recently, the areas investigated by the MF approaches have been extended to game theory, economics, and finance. Several papers [14, 16, 18, 5] investigated the large-population stochastic differential games on wireless communications and gave ε-Nash equilibrium strategies by the Nash certainty equivalence methodology. Weintraub, Benkard, and Van Roy [36, 37] studied discrete-time large-scale stochastic dynamic games on microeconomics and introduced the notion of oblivious equilibrium via MF approximation. Lasry and Lions [22, 23] presented an MF model and a limit equation for economic and financial systems. The class of large-population games investigated via the MF approach has the following features: each agent is affected by the average interaction of all the other agents, while the individual influence of each

*Received by the editors June 28, 2010; accepted for publication (in revised form) May 10, 2012; published electronically August 23, 2012. This work was supported by the National Natural Science Foundation of China (under 60934006 and 61120106011).
http://www.siam.org/journals/sicon/50-4/80032.html

†Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China (wangbc@amss.ac.cn, jif@iss.ac.cn).

agent is negligible. The key idea of tackling this class of game problems is to transform the large population game into a single-agent optimal control problem by replacing the average effect of all the agents to a single agent with the macroscopic population information. This could make us transfer a high-dimensional problem into a low-dimensional problem and hence reduce the computational complexity dramatically.

There have been many works [16, 17, 18, 24, 25, 27, 19, 29, 38] investigating the large population MASs by using the MF approaches. The papers [16, 17, 18, 19] investigated stochastic differential games for large-population systems with agents coupled via their discounted costs. Li and Zhang [25, 24] considered decentralized games for large-population stochastic MASs with stochastic time-average coupled cost functions. Nourian et al. [29] considered leader-follower dynamic games, where trajectories of the leaders are unknown but in a given finite class of models. Zhang and Zhang [38] investigated the robust adaptive control of MASs with unknown parameters and unmodeled dynamics. However, to our knowledge, the systems with random time-varying parameters have not yet been investigated, which have a strong background in practical application and are definitely worth deeply studying.

For instance, consider an example of production processes in N companies. In a market, N companies supply the same products. The production line in each company consists of multiple machines. In the practical production, it is unavoidable that some machines fail. The production lines in all the companies are the same by and large, while the production lines in different companies operate independently. This kind of system can be characterized by an MAS with parameters modeled by a family of independent Markov chains with identical generators. The study of the similar fault-prone dynamics characterized by Markov jump systems may trace back to [33] and [4]. Through the development in several decades, many deep results have been obtained [34, 35, 20, 28, 12, 7] in this direction.

This paper is aimed at investigating distributed games for a class of large-population MASs, where the agents are coupled together via cost functions with nonlinear average terms and with time-varying parameters characterized by a sequence of independent Markov chains with identical generators. Although a range of mature techniques and results for Markov jump systems (particularly, for the linear case [28, 7]) have been developed, the results cannot be extended to MASs in parallel. The main reason is that in the MASs, not only agents have autonomy and self-governing capability, but also there may exist all kind of interactions (e.g., competition and cooperation) among the agents.

In this paper, we will use the MF approach to design the distributed strategies. To do so and analyze uniform stability of the closed-loop system and asymptotic optimality of the distributed strategies, we need to approximate the average effect of all the agents by the population statistics effect (PSE), which can be obtained by solving the equilibrium problem of an auxiliary two-player Stackelberg game. Compared with the time-invariant case, one key difficulty encountered here is to solve a fixed point problem of an operator. For the time-invariant case, the operator can be expressed explicitly and the existence of the fixed point problem is easy to solve. However, for the case of Markov jump parameters, the operator is very complicated and the fixed point problem is hard to analyze. To overcome this difficulty, we first construct a stochastic Lyapunov function and then use the differential inequality and its solution related to the Lyapunov function to study the contraction property of the operator and analyze the existence of the fixed point problem.

The remainder of this paper is organized as follows. In section 2, we describe the model and basic assumptions. In section 3, we first give the solution of Markov jump linear-quadratic-Gaussian (LQG) optimal tracking control problem and then obtain a fixed point equation for the PSE and the conditions under which the equation admits a unique solution. Based on these, we give a group of distributed control strategies. In section 4, we analyze the stability of closed-loop systems. In section 5, we show that distributed strategies are asymptotically optimal in the sense of the Nash equilibrium. In section 6, we give a detailed discussion for the case where the cost functions are with linear effect terms and simplify the condition given in section 3 to one easier to be verified. In section 7, we give a numerical example to demonstrate the procedure of the control design and the influence of the heterogeneity intensity and the parameter jump rate of the agents on the closed-loop system. In section 8, we provide a brief summary for the paper.

The following notation will be used in the paper. $\|\cdot\|$ denotes the Euclidean vector norm or matrix norm induced by Euclidean vector norm; I denotes an n-dimensional identity matrix; $\mathbf{1}$ denotes the m-dimensional column vector whose elements are 1; e_j denotes the column vector whose elements are 0 except that the jth element is 1. For a given matrix A, A^T denotes its transpose and $tr(A)$ denotes the trace of A. For any vector x with proper dimensions and symmetric matrix $Q \geq 0$, $\|x\|_Q = (x^T Q x)^{1/2}$. For an n-dimensional matrix P, $\lambda_1(P)$ denotes the minimum eigenvalue of P and $\lambda_n(P)$ denotes maximum eigenvalue of P. I_B denotes the indicator function of set B. $C([0,\infty),\mathbb{R}^n)$ denotes the class of n-dimensional continuous functions in $[0,\infty)$; $C_b([0,\infty),\mathbb{R}^n) = \{f \in C([0,\infty),\mathbb{R}^n) : \|f\|_\infty \triangleq \sup_{t\geq 0}\|f(t)\| < \infty\}$.

2. Problem formulation. In the paper, we consider the MAS described by the following dynamics:

$$dx_i(t) = A_{\theta_i(t)}x_i(t)dt + B_{\theta_i(t)}u_i(t)dt + h(t)dt + D_{\theta_i(t)}dW_i(t),\ 1 \leq i \leq N, \tag{2.1}$$

where $x_i \in \mathbb{R}^n$ and $u_i \in \mathbb{R}^r$ are the state and input of agent i, respectively, and $\{W_i(t), 1 \leq i \leq N\}$ is stochastic disturbance, a sequence of d dimensional standard Brownian motions. $h \in C_b([0,\infty),\mathbb{R}^n)$ is an external signal, reflecting the impact on the agent i by the environment. $\{\theta_i(t), 1 \leq i \leq N\}$ is a sequence of independent continuous-time Markov chains taking value in $S = \{1, 2, \ldots, m\}$ with the identical infinitesimal generator (transition rate matrix) $\Lambda = \{\lambda_{ij}, i, j = 1, \ldots, m\}$. $\{W_i(t)\}$ and $\{\theta_i(t)\}$ are independent of each other. The cost function of agent i is

$$\begin{aligned} &J_i(u_i, u_{-i}) \\ &= \limsup_{T\to\infty} \frac{1}{T} E \int_0^T \Big\{\big\|x_i(t) - \Phi[\Psi^{(N)}(x)]\big\|^2_{Q_{\theta_i(t)}} + \|u_i(t)\|^2_{R_{\theta_i(t)}}\Big\}dt, \end{aligned} \tag{2.2}$$

where $Q_j > 0,\ R_j > 0,\ j = 1, \ldots, m$,

$$u_{-i} = \{u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_N\},$$
$$\Psi^{(N)}(x) = \frac{1}{N}\sum_{i=1}^N \Psi(x_i(t)).$$

Here $\Phi[\Psi^{(N)}(x)]$ gives a measure of the average effect caused by all the agents and both Φ and Ψ are known functions from $\mathbb{R}^n$ to $\mathbb{R}^n$, called effect functions.

Remark 2.1. Model (2.1) with the cost function (2.2) has a wide economic background. For instance, let us continue to consider the example of the production processes in N companies mentioned in the introduction. In a market, an increasing number of companies distributed over diverse areas join together to serve an increasing number of customers. As it is known, supply-demand relationship affects market price. Hence, we can assume that the following relation holds:

$$p = \frac{1}{N}\sum_{i=1}^{N}\psi(x_i), \tag{2.3}$$

where p denotes the market price, x_i denotes the output level of company i, and $\psi(x) = \beta - \alpha x, \beta > 0, \alpha > 0$. Indeed, (2.3) is a variant form of the price model in [21] and [18].

Now, we assume each company raises its output level as the price goes up. Specifically, each company adjusts its output level to make it satisfy the following relationship:

$$x_i \approx \phi(p) = \phi\left(\frac{1}{N}\sum_{i=1}^{N}\psi(x_i)\right), \tag{2.4}$$

where ϕ is an increasing function. Thus, for company i, the following cost function needs to be investigated:

$$\limsup_{T\to\infty}\frac{1}{T}E\int_0^T\left\{\left[x_i - \phi\left(\frac{1}{N}\sum_{i=1}^{N}\psi(x_i)\right)\right]^2 + ru_i^2\right\}dt.$$

Here the penalty term $[x_i - \phi(\frac{1}{N}\sum_{i=1}^{N}\psi(x_i))]^2$ is introduced based on (2.4), and ru_i^2 $(r > 0)$ denotes the instantaneous cost of adjustment for agent i.

Remark 2.2. If for any $t \geq 0$, $\theta_i(t) \equiv \theta_i, 1 \leq i \leq N$, and Φ, Ψ are linear functions, then the system (2.1)–(2.2) degenerates to the model studied in [24]. If $\theta_i(t) \equiv \theta_i, 1 \leq i \leq N$ and the state variable x_i is 1-dimensional, then the system (2.1)–(2.2) degenerates to the model studied in [16]. Hence, this paper extends the models in the existing works [24, 16] from the time-invariant case to the case of random time-varying parameters.

The objective of this paper is to seek optimal distributed strategies for the system (2.1)–(2.2) within the setting of a noncooperative game. However, due to the particular structure of the system, constructing distributed Nash equilibrium strategies by using the usual method, if feasible, generally results in a quite high computational complexity. In this paper, we will use PSE to approximate the average effect of all the agents and then design a group of distributed strategies, which is proved to be asymptotically optimal in the sense of Nash equilibrium.

For convenience of reference, we list some assumptions which will be used in the paper as follows:

(A1) $\{x_i(0)\}$ are i.i.d. random variables (r.v.s) and independent of $\{W_i(t), 1 \leq i \leq N\}$ and $\{\theta_i(t), 1 \leq i \leq N\}$, $\max_i E\|x_i(0)\|^2 < \infty$. $\theta_i(0),\ 1 \leq i \leq N$ have identical distributions.

(A2) Given the positive matrices $N_1, \ldots, N_m$, the following coupled Riccati equations have a set of unique positive definite solutions $\{M_j, j = 1, \ldots, m\}$:

$$A_j^T M_j + M_j A_j - M_j B_j B_j^T M_j + \sum_{k=1}^{m}\lambda_{jk}M_k = -N_j, \quad j = 1, \ldots, m.$$

(A3) The functions Φ and Ψ are Lipschitz continuous with Lipschitz constants L_Φ and L_Ψ, i.e., for any $x, y \in \mathbb{R}^n$,

$$\|\Phi(x) - \Phi(y)\| \le L_\Phi \|x - y\|,$$
$$\|\Psi(x) - \Psi(y)\| \le L_\Psi \|x - y\|.$$

Remark 2.3. As a game problem, we need the assumption that each agent is rational. To be specific, we assume that (i) each agent optimizes its cost function and (ii) when making decisions, each agent assumes that the others are rational. This assumption ensures the rationality of the MF approximation.

Remark 2.4. Assumption (A2), which often appears in the study on stabilization of Markov jump systems, ensures the stability of the closed-loop system [20, 11, 28]. Indeed, assumption (A2) is equivalent to stochastic stabilizability [20] of the Markov jump system

$$\dot{y}(t) = A_{\theta_1(t)} y(t) + B_{\theta_1(t)} u(t),$$

i.e., for any $y(0) \in \mathbb{R}^n, \theta_1(0) \in S$, there exists a linear feedback law $u(t) = -L_{\theta_1(t)} y(t)$ and a positive definite matrix M such that

$$\lim_{T\to\infty} E\left[\int_0^T y^T(t) y(t) dt \big| y(0), \theta_1(0)\right] \le y^T(0) M y(0).$$

3. Construction of distributed strategies. In this section, we will construct a group of distributed strategies by using the MF approach.

Since each agent is weakly coupled with others and both $\{\theta_i(t)\}$ and $\{W_i(t)\}$ are independent of each other, from the MF theory and agents' rationalities we can take PSE as the average effect of all the agents to decouple the large-population system and then design distributed strategies. Generally speaking, PSE is deterministic but cannot be available beforehand. To overcome this difficulty, we first regard the PSE as a deterministic function, and all the agents use this function to construct their optimal control strategies. In this case, when the number of agents is sufficiently large, the average effect of all the agents in the closed-loop system should intuitively be close to the deterministic function.

Following this line, we shall obtain a fixed point equation, which under some conditions can be shown to have a solution. With this solution we can get the PSE of the closed-loop system. As a result, all the agents can design their own optimal strategy by using its own state and this macroscopic population information PSE. From the analysis below, one can see that to get the PSE, an auxiliary two-player game problem may be considered first, and then by solving the Stackelberg equilibrium of the game we can obtain the PSE.

Thus, we can construct distributed strategies in three steps: (I) solve the Markov jump optimal tracking problem with a deterministic reference signal; (II) give the macroscopic population information PSE approximating the average effect of all the agents; and (III) design distributed strategies.

3.1. Markov jump LQG tracking problem with a known reference signal. Consider the system

$$dx(t) = A_{\theta(t)} x(t) dt + B_{\theta(t)} u(t) dt + h(t) dt + D_{\theta(t)} dW(t), \tag{3.1}$$

where $x \in \mathbb{R}^n, \{W(t), t \ge 0\}$ is a d dimensional standard Brownian motion, amd $\{\theta(t), t \ge 0\}$ is a Markov chain with the infinitesimal generator Λ. The cost function is

$$J(u) = \limsup_{T\to\infty} \frac{1}{T} E \int_0^T \left\{ \|x(t) - f^b(t)\|^2_{Q_{\theta(t)}} + \|u(t)\|^2_{R_{\theta(t)}} \right\} dt, \tag{3.2}$$

where $f^b \in C_b([0,\infty),\mathbb{R}^n)$ is a known reference signal. The admissible control set is

$$\mathcal{U} = \Bigg\{ u \mid u(t) \in \sigma(x(s),\theta(s), s \le t), \quad E\|x(T)\| = o(\sqrt{T}), \\ E\int_0^T \|x(t)\|^2 dt = O(T), T \to \infty \Bigg\}.$$

THEOREM 3.1. *If assumption* (A1) *holds,* $R_j > 0$, $Q_j > 0$, $j = 1,\dots,m$, *then for the system* (3.1) *with the cost function* (3.2), *the following results hold:*

(i) *The coupled Riccati equations*

$$A_j^T K_j + K_j A_j + \sum_{k=1}^m \lambda_{jk} K_k - K_j B_j R_j^{-1} B_j^T K_j + Q_j = 0 \tag{3.3}$$

have a set of unique positive definite solutions K_j, $j = 1,\dots,m$.

(ii) $G^T + \Lambda \otimes I$ *is stable, i.e., all the eigenvalues of* $G^T + \Lambda\otimes I$ *have negative real parts, where* $G = diag\{G_1,\dots,G_m\}$, $G_j = A_j - B_j R_j^{-1} B_j^T K_j$.

(iii) *The differential equations*

$$\frac{dr_j(t)}{dt} + G_j^T r_j(t) + \sum_{k=1}^m \lambda_{jk} r_k(t) + K_j h(t) - Q_j f^b(t) = 0, j = 1,\dots,m, \tag{3.4}$$

admit a set of unique solutions in $C_b([0,\infty),\mathbb{R}^n)$

$$r_j(t) = e_j^T \otimes I \cdot \int_t^\infty \exp[-(G^T + \Lambda\otimes I)(t-s)]\big\{K[\mathbf{1}\otimes h(s)] - Q[\mathbf{1}\otimes f^b(s)]\big\} ds.$$

(iv) *The optimal control law* $u^* = arginf_{u\in\mathcal{U}} J(u)$ *is given by*

$$u^*(t) = -\sum_{j=1}^m R_j^{-1} B_j^T [I_{[\theta(t)=j]}(K_j x(t) + r_j(t))].$$

(v) *The optimal value of the cost function is*

$$J(u^*) = \limsup_{T\to\infty} \frac{1}{T} E \int_0^T \Big\{ {f^b}^T(t) Q_{\theta(t)} f^b(t) - 2r^T(\theta(t),t)h(t) \\ - r^T(\theta(t),t) B_{\theta(t)} R^{-1}_{\theta(t)} B^T_{\theta(t)} r(\theta(t),t) + tr(K_{\theta(t)} D_{\theta(t)} D^T_{\theta(t)}) \Big\} dt.$$

Moreover, if $\theta(t)$ *is ergodic and the stationary distribution is* $\{\pi_j\}$, *then*

$$J(u^*) = \sum_{j=1}^m \pi_j \limsup_{T\to\infty} \frac{1}{T} \int_0^T \Big\{ {f^b}^T(t) Q_j f^b(t) - 2r_j^T(t) h(t) \\ - r_j^T(t) B_j R_j^{-1} B_j^T r_j(t) + tr(K_j D_j D_j^T) \Big\} dt.$$

Proof. See Appendix A. ∎

3.2. Approximation of the PSE based on the MF approach. Theorem 3.1 solves a class of LQG optimal tracking problems of Markov jump system with a known reference signal. However, the reference signal in the cost function (2.2) is unknown and cannot be used directly to design control strategies. Thus, we first try to estimate the reference signal and then construct distributed strategies by using Theorem 3.1 and the MF approach.

To do so, we consider an auxiliary two-player Stackelberg hierarchical game. The state evolves according to

$$dx(t) = A_{\theta(t)}x(t)dt + B_{\theta(t)}u(t)dt + h(t)dt + D_{\theta(t)}dW(t), \tag{3.5}$$

where $W(t)$ and $\theta(t)$ are specified as section 3.1 and the initial state $x(0)$ is a random variable, which has the same distribution as $x_i(0)$. The cost functions of the two players are, respectively,

$$J_1(u,v) = \limsup_{T\to\infty} \frac{1}{T} E\int_0^T \left\{ \|x(t) - \Phi(v(t))\|^2_{Q_{\theta(t)}} + \|u(t)\|^2_{R_{\theta(t)}} \right\} dt, \tag{3.6}$$

$$J_2(u,v) = \limsup_{T\to\infty} \frac{1}{T} \int_0^T \|v(t) - E\Psi(x(t))\|^2 \, dt, \tag{3.7}$$

where $u \in \mathcal{U}$ is the strategy of player 1 and $v \in \mathcal{V} = C_b([0,\infty), \mathbb{R}^n)$ is the strategy of player 2. Our goal is to seek the Stackelberg equilibrium solution with player 2 acting as the leader.

Remark 3.2. The game problem (3.5)–(3.7) introduced is based on the MF approach. Player 1 is a representative agent in model (2.1). Player 2 is a virtual agent–"population." v is the population effect. Indeed, for the model of Remark 2.1, we may regard player 2 as "market" and v as price.

A usual way of seeking the Stackelberg equilibrium solution is the brute-force method [1]. Specifically, for every fixed $v \in \mathcal{V}$, we can obtain an optimal reaction strategy u^v and get a Stackelberg strategy for player 2 by minimizing $J_2(v, u^v)$ over $\mathcal{V}$.

Now we solve problem (3.5)–(3.7) by the brute-force method. First, for every fixed $v \in C_b([0,\infty), \mathbb{R}^n)$, by Theorem 3.1 we obtain the optimal tracking strategy for player 1

$$u^v(t) = -R^{-1}_{\theta(t)} B_{\theta(t)} [K_{\theta(t)} x(t) + r^v_{\theta(t)}(t)],$$

where

$$r^v_j(t) = e^T_j \otimes I \cdot \int_t^\infty \exp[-(G^T + \Lambda \otimes I)(t-s)] \big\{ K[\mathbf{1} \otimes h(s)] - Q[\mathbf{1} \otimes \Phi(v(s))] \big\} ds.$$

Substituting u^v_t into (3.5), we have the closed-loop equation

$$dx^v(t) = G_{\theta(t)} x^v(t)dt + h(t)dt - B_{\theta(t)} R^{-1}_{\theta(t)} B^T_{\theta(t)} r^v_{\theta(t)}(t)dt + D_{\theta(t)}dW(t). \tag{3.8}$$

For

$$J_2(u^v, v) = \limsup_{T\to\infty} \frac{1}{T} \int_0^T \|v(t) - E\Psi(x^v(t))\|^2 \, dt,$$

if the equation

$$E\Psi(x^v) = v \tag{3.9}$$

admits a solution v^*, then

$$J_2(u^{v^*}, v^*) = \inf_{v \in \mathcal{V}} J_2(u^v, v) = 0.$$

Thus, v^* is a Stackelberg equilibrium strategy for the leader (player 2). From the above analysis, it follows that $\Phi(v^*)$ is the closed-loop PSE that we are seeking.

However, an unsolved question is whether (3.9) admits a solution. For the convenience of statements, we first define an operator $\mathcal{T}: \; v \mapsto E\Psi(x^v)$, where $v \in C_b([0,\infty), \mathbb{R}^n)$. Since both Φ and Ψ are Lipschitz continuous, $\sup_{t \geq 0} \|(\mathcal{T}v)(t)\| < \infty$. Thus, $\mathcal{T}$ is an operator on $C_b([0,\infty), \mathbb{R}^n)$. From the definition of $\mathcal{T}$, (3.9) can be rewritten as

$$\mathcal{T}v = v. \tag{3.10}$$

Thus, the question becomes whether $\mathcal{T}$ has a fixed point. For the time-invariant case, the form of $\mathcal{T}$ can be given explicitly and the question can be answered comparatively more easily. However, for the case of Markov jump parameters, the question is quite complicated. Since the analytic solutions of linear differential equations with Markov jump parameters are hard to give, so is the explicit form of $\mathcal{T}$. To overcome this difficulty, here we first construct a stochastic Lyapunov function and with the help of a differential inequality and its solution related to this Lyapunov function show that the operator $\mathcal{T}$ is contractive and has a unique fixed point.

To obtain the existence for solutions of (3.9), we need the following assumption:

(A4) $\max_j \|B_j R_j^{-1} B_j^T\| \gamma_1 \gamma_2 < 1$, where

$$\begin{aligned}
\gamma_1 &= \sqrt{m} \max_j \lambda_n(Q_j) \int_0^\infty \|\exp[(G^T + \Lambda \otimes I)t]\| dt, \\
\gamma_2 &= \frac{2L_\Phi L_\Psi \max_j \sqrt{\lambda_n(K_j)}}{\min_j \sqrt{\lambda_1(K_j)} \min_j \{\lambda_1(N_j)/\lambda_n(K_j)\}}, \\
N_j &= Q_j + K_j B_j R_j^{-1} B_j^T K_j.
\end{aligned}$$

Remark 3.3. Generally speaking, assumption (A4) is difficult to verify directly. However, for a class of special jump system, 1-dimensional time-invariant system, by a straightforward calculation one can see that assumption (A4) holds if and only if

$$L_\Phi L_\Psi \leq 1 + \frac{A^2 R}{B^2 Q}.$$

Obviously, when $L_\Phi L_\Psi$ is small and R/Q is large, the above inequality is easy to hold. Besides, for general (nondegenerate) Markov jump cases, we shall provide a numerical example satisfying (A1)–(A4) in section 7.

THEOREM 3.4. *Under Assumptions* (A1)–(A4), (3.10) *admits a unique solution* v^* *in* $C_b([0,\infty), \mathbb{R}^n)$.

Proof. For any $f, g \in C_b([0,\infty), \mathbb{R}^n)$, let $r_j^f(t), x^f(t)$ $(r_j^g(t), x^g(t))$, $j = 1, \ldots, m$, denote the solution of (3.4) and (3.8) when replacing f^b and v by f (g). Let

$$\begin{aligned}
r^f(t) &= [r_1^f(t)^T, \ldots, r_m^f(t)^T]^T, \\
r^g(t) &= [r_1^g(t)^T, \ldots, r_m^g(t)^T]^T.
\end{aligned}$$

Then, from (3.4) it follows that

$$\frac{d(r^f - r^g)}{dt} + (G^T + \Lambda \otimes I)(r^f - r^g) - Q[\mathbf{1} \otimes \Phi(f - g)] = 0,$$

where $Q = diag\{Q_1, \dots, Q_m\}$. Using the same argument as in the proof of Theorem 3.1(iii), we can prove that the above differential equation has a unique solution in $C_b([0,\infty), \mathbb{R}^n)$, that is,

$$r^f(t) - r^g(t) = \int_t^\infty \exp[-(G^T + \Lambda \otimes I)(t-s)]Q[\mathbf{1} \otimes \Phi(f(s) - g(s))]ds.$$

Let $\Delta r(t) = r^f(t) - r^g(t), \Delta r_j(t) = r_j^f(t) - r_j^g(t), j = 1, \dots, m$. Then

$$\begin{aligned}\max_j \|\Delta r_j(t)\|_\infty \le \|\Delta r(t)\|_\infty &\le \sqrt{m} L_\Phi \max_j \|Q_j\| \|f-g\|_\infty \\ &\quad \times \int_t^\infty \|\exp[(G^T + \Lambda \otimes I)(s-t)]\| ds \\ &= L_\Phi \gamma_1 \|f-g\|_\infty.\end{aligned}$$

Now we calculate $\|\mathcal{T}f - \mathcal{T}g\|_\infty$. Let $\Delta x(t) = x^f(t) - x^g(t)$. Then, from (3.8) it follows that

$$d\Delta x(t) = G_{\theta(t)} \Delta x(t) dt - B_{\theta(t)} R_{\theta(t)}^{-1} B_{\theta(t)}^T \Delta r(\theta(t), t) dt. \tag{3.11}$$

By (3.3) we have

$$\begin{aligned}&K_j(A_j - B_j R_j^{-1} B_j^T K_j) + (A_j - B_j R_j^{-1} B_j^T K_j)^T K_j + \sum_{k=1}^m \lambda_{jk} K_k \\ &= -Q_j - K_j B_j R_j^{-1} B_j^T K_j, \quad j = 1, \dots, m.\end{aligned}$$

Let $N_j = Q_j + K_j B_j R_j^{-1} B_j^T K_j$. Obviously, $N_j > 0$ and

$$K_j G_j + G_j^T K_j + \sum_{k=1}^m \lambda_{jk} K_k = -N_j, j = 1, \dots, m,$$

where $G_j = A_j - B_j R_j^{-1} B_j K_j$. Let $V(j,x) = x^T K_j x$, $j = 1, \dots, m$ and $\mathcal{A}$ be the infinitesimal generator of $\{\theta(t), \Delta x(t)\}$ satisfying (3.11). Then

$$\begin{aligned}\mathcal{A}V(j,x) &= \lim_{s \to 0} \frac{1}{s} \big\{ E[V(\theta(t+s), \Delta x(t+s)) | \theta(t) = j, \Delta x(t) = x] - V(j,x) \big\} \\ &= \frac{\partial V}{\partial x}(j,x)(G_j x + B_j R_j^{-1} B_j^T \Delta r_j(t)) + \sum_k \lambda_{jk} V(k,x) \\ &= x^T \left(K_j G_j + G_j^T K_j + \sum_k \lambda_{jk} K_k \right) x + 2x^T K_j B_j R_j^{-1} B_j^T \Delta r_j,\end{aligned}$$

which leads to

$$\begin{aligned}&\frac{dEV(\theta(t), \Delta x(t))}{dt} \\ &= E\mathcal{A}V(\theta(t), \Delta x(t)) \\ &= -E[\Delta x^T(t) N_{\theta(t)} \Delta x(t)] - 2E[\Delta x^T(t) K_{\theta(t)} B_{\theta(t)} R_{\theta(t)}^{-1} B_{\theta(t)}^T \Delta r_{\theta(t)}(t)] \\ &\le -a_1 EV(\theta(t), \Delta x(t)) + 2E\left[\left\| \Delta x^T(t) \sqrt{K_{\theta(t)}} \right\| \left\| \sqrt{K_{\theta(t)}} B_{\theta(t)} R_{\theta(t)}^{-1} B_{\theta(t)}^T \Delta r_{\theta(t)}(t) \right\| \right] \\ &\le -a_1 EV(\theta(t), \Delta x(t)) + 2b_1 \|f-g\|_\infty \sqrt{EV(\theta(t), \Delta x(t))},\end{aligned}$$

where

$$a_1 = \min_j \frac{\lambda_1(N_j)}{\lambda_n(K_j)}, \quad b_1 = \max_j \sqrt{\lambda_n(K_j)}\|B_j R_j^{-1} B_j^T\| L_\Phi \gamma_1. \tag{3.12}$$

Noting $\Delta x(0) = 0$, from the above differential inequality it follows that

$$EV(\theta(t), \Delta x(t)) \leq \left\{ \int_0^t \exp[-\frac{a_1}{2}(t-s)] b_1 \|f-g\|_\infty ds \right\}^2 \leq \left(\frac{2b_1}{a_1}\|f-g\|_\infty \right)^2,$$

which implies

$$E\|\Delta x(t)\|^2 \leq \frac{EV(\theta(t), \Delta x(t))}{\min_j(\lambda_1(K_j))} \leq \frac{(\frac{2b_1}{a_1}\|f-g\|_\infty)^2}{\min_j(\lambda_1(K_j))}.$$

Thus, we have

$$\begin{aligned}\|\mathcal{T}f - \mathcal{T}g\|_\infty &= \sup_{t\geq 0} \|E[\Psi(x^f(t)) - \Psi(x^g(t))]\| \leq \sup_{t\geq 0} \left[E\|\Psi(x^f(t)) - \Psi(x^g(t))\|^2\right]^{\frac{1}{2}} \\ &\leq \sup_{t\geq 0} L_\Psi (E\|\Delta x(t)\|^2)^{\frac{1}{2}} \leq \frac{2L_\Psi b_1 \|f-g\|_\infty}{a_1 \min_j \sqrt{\lambda_1(K_j)}}.\end{aligned}$$

By assumption (A4) and (3.12), $\mathcal{T}$ is a contraction mapping on $C_b([0,\infty), \mathbb{R}^n)$, which yields the existence and uniqueness of the fixed point v^* such that $\mathcal{T}v^* = v^*$. ∎

3.3. Design of distributed strategies. From the analysis and conclusion of sections 3.1 and 3.2, we obtain a group of distributed strategies as follows:

$$u_i^*(t) = -\sum_{j=1}^m R_j^{-1} B_j^T I_{[\theta_i(t)=j]}[K_j x_i(t) + r_j^*(t)], \ 1 \leq i \leq N, \tag{3.13}$$

where K_j is the solution of (3.3),

$$r_j^*(t) = e_j^T \otimes I \cdot \int_t^\infty \exp[-(G^T + \Lambda \otimes I)(t-s)]\big\{K[\mathbf{1} \otimes h(s)] - Q[\mathbf{1} \otimes \Phi(v^*(s))]\big\} ds,$$

and v^* is the unique solution of (3.9).

By Theorem 3.1 and (3.9), $r_j^*(t)$ and K_j merely depend on $A_j, B_j, R_j, Q_j (j = 1, \ldots, m), \Phi, \Psi$, which are unrelated to the agents' states $x_i(t)$ in real time. Hence, (3.13) are distributed strategies. In addition, since $\mathcal{T}$ is a contraction, $v^* = \lim_{n\to\infty} \mathcal{T}^n v_0$ for any $v_0 \in C_b([0,\infty), \mathbb{R}^n)$. Thus, we can use an iterative algorithm to approximate the solution of (3.9).

4. Stability of closed-loop system. Substituting the control strategies (3.13) into (2.1), we have the closed-loop system equation

$$\begin{aligned} dx_i^*(t) &= G_{\theta_i(t)} x_i^*(t) dt - B_{\theta_i(t)} R_{\theta_i(t)}^{-1} B_{\theta_i(t)}^T r^*(\theta_i(t), t) \\ &\quad + h(t)dt + D_{\theta_i(t)} dW_i(t), \quad 1 \leq i \leq N. \end{aligned} \tag{4.1}$$

Now, we study the uniform stability of the closed-loop system in a time-average sense.

THEOREM 4.1. *If assumptions* (A1)–(A4) *hold, then there exists a constant* C_0 *independent of* i *and* N *such that the strategies* (3.13) *and the corresponding closed-loop system* (4.1) *satisfy*

$$\sup_{N\geq 1}\max_{1\leq i\leq N}\limsup_{T\to\infty}\frac{1}{T}E\int_0^T(\|x_i^*(t)\|^2+\|u_i^*(t)\|^2)dt\leq C_0. \tag{4.2}$$

Proof. By Theorem 3.1(iii) we have

$$\begin{aligned}\max_j\|r_j^*(t)\|_\infty &\leq \sqrt{m}(\max_j\lambda_n(K_j)+\max_j\lambda_n(Q_j))(\|h\|_\infty+\|\Phi(v^*)\|_\infty)\\ &\quad\times\int_0^\infty\|\exp[(G^T+\Lambda\otimes I)s]\|ds\triangleq M_r.\end{aligned} \tag{4.3}$$

Let $V(j,x)=x^TK_jx$, $j=1,\dots,m$, and $\mathcal{A}_i$ be the infinitesimal generator of $\{\theta_i(t), x_i^*(t), 1\leq i\leq N\}$ satisfying (4.1). Then

$$\begin{aligned}&\mathcal{A}_iV(j,x)\\ &\quad=\lim_{s\to 0}\frac{1}{s}\{E[V(\theta_i(t+s),x_i^*(t+s))|\theta_i(t)=j,x_i^*(t)=x]-V(j,x)\}\\ &\quad=\frac{\partial V}{\partial x}(j,x)[G_jx-B_jR_j^{-1}B_j^Tr_j^*(t)+h(t)]+\sum_{k=1}^m\lambda_{jk}V(k,x)+tr(K_jD_jD_j^T)\\ &\quad=-x^TN_jx-2x^TK_j[B_jR_j^{-1}B_jr_j^*-h]+tr(K_jD_jD_j^T).\end{aligned}$$

By Dynkin's formula [30] we have

$$\begin{aligned}\frac{dEV(\theta_i(t),x_i^*(t))}{dt}&=E\mathcal{A}_iV(\theta_i(t),x_i^*(t))\\ &=E\Big\{-x^T(t)N_{\theta_i(t)}x_i^*(t)+tr(K_{\theta_i(t)}D_{\theta_i(t)}D_{\theta_i(t)}^T)\\ &\qquad-2x^TK_{\theta_i(t)}[B_{\theta_i(t)}R_{\theta_i(t)}^{-1}B_{\theta_i(t)}^Tr^*(\theta_i(t),t)-h(t)]\Big\}\\ &\leq -\frac{1}{2}a_1EV(\theta_i(t),x_i^*(t))+b_2,\end{aligned}$$

where

$$\begin{aligned}a_1&=\min_j\frac{\lambda_1(N_j)}{\lambda_n(K_j)},\\ b_2&=2\max_j\lambda_n(K_j)\big(\|B_jR_j^{-1}B_j^T\|M_r+\|h\|_\infty\big)^2/a_1+\max_j tr(K_jD_jD_j^T).\end{aligned}$$

From this and the comparison principle it follows that

$$EV(\theta_i(t),x_i^*(t))\leq EV(\theta_i(0),x_i^*(0))e^{-\frac{1}{2}a_1t}+\frac{2b_2}{a_1}(1-e^{-\frac{1}{2}a_1t}),$$

which together with the definition of $EV(\theta_i(t),x_i^*(t))$ results in

$$\begin{aligned}E\|x_i^*(t)\|^2&\leq\frac{EV(\theta_i(t),x_i^*(t))}{\min_j(\lambda_1(K_j))}\\ &\leq\frac{\max_j\lambda_n(K_j)\max_iE\|x_i^*(0)\|^2+\frac{2b_2}{a_1}}{\min_j(\lambda_1(K_j))}\triangleq M_1.\end{aligned} \tag{4.4}$$

By (3.13) and (4.3) we have

$$E\|u_i^*(t)\|^2 \le \max_j \|R_j^{-1}\|^2\|B_j\|^2(2\max_j \|K_j\|^2 M_1 + 2M_r^2). \tag{4.5}$$

Thus, from (4.4) and (4.5) it follows that

$$\begin{aligned}&\limsup_{T\to\infty} \frac{1}{T}E\int_0^T (\|x_i^*(t)\|^2 + \|u_i^*(t)\|^2)dt \\ &\quad \le M_1 + \max_j \|R_j^{-1}\|^2\|B_j\|^2(2\max_j \|K_j\|^2 M_1 + 2M_r^2).\end{aligned} \tag{4.6}$$

Noting that M_r and M_1 are independent of i, N, take

$$C_0 = M_1 + \max_j \|R_j^{-1}\|^2\|B_j\|^2(2\max_j \|K_j\|^2 M_1 + 2M_r^2).$$

Then, by (4.6) one can get (4.2). ∎

5. Analysis of optimality.

We first analyze the property of the estimation for the average effect of population and then show the asymptotic optimality of distributed strategies (3.13).

LEMMA 5.1. *If assumptions* (A1)–(A4) *hold, then under the group of the strategies* (3.13)*, the closed-loop system* (4.1) *satisfies*

$$\lim_{N\to\infty}\limsup_{T\to\infty} \frac{1}{T}E\int_0^T \left\|\Phi(\Psi^{(N)}(x^*)) - \Phi(v^*)\right\|^2 dt = 0. \tag{5.1}$$

Proof. From Ψ's Lipschitz continuity and Theorem 4.1 it follows that for any $1 \le i \le N$,

$$\begin{aligned}E\|\Psi(x_i^*(t))\|^2 &= E\|\Psi(x_i^*(t)) - \Psi(0) + \Psi(0)\|^2 \\ &\le 2E\|\Psi(x_i^*(t)) - \Psi(0)\|^2 + 2\|\Psi(0)\|^2 \\ &\le 2L_\Psi^2 E\|x_i^*(t)\|^2 + 2\|\Psi(0)\|^2 \le 2L_\Psi^2 C_0 + 2\|\Psi(0)\|^2.\end{aligned}$$

Since $\{\theta_i(t), i = 1,\dots,N\}$ is a sequence of independent Markov chains, $\{W_i(t), i = 1,\dots,N\}$ is a sequence of independent Brownian motions, and $\{x_i(0), i = 1,\dots,N\}$ are independent of each other, by the closed-loop system equation (4.1) we get that $\{x_i^*(t), i = 1,\dots,N\}$ are independent of each other. Hence, noting $v^*(t) = E\Psi(x_i^*(t))$, we have

$$\begin{aligned}E\left\|\Phi(\Psi^{(N)}(x^*)) - \Phi(v^*)\right\|^2 &\le L_\Phi^2 E\left\|\Psi^{(N)}(x^*) - v^*]\right\|^2 \\ &\le \frac{\sum_{i=1}^N L_\Phi^2 E\|\Psi(x_i^*(t)) - v^*(t)]\|^2}{N^2} \\ &\le \frac{\sum_{i=1}^N L_\Phi^2 E\|\Psi(x_i^*(t))\|^2}{N^2} \\ &\le \frac{L_\Phi^2(2L_\Psi^2 C_0 + 2\|\Psi(0)\|^2)}{N} \to 0,\end{aligned} \tag{5.2}$$

which holds uniformly with repsect to t. Thus, (5.1) is true. ∎

Remark 5.2. Lemma 5.1 tells us that PSE is a good estimate of the average effect of all the agents in some sense and shows the rationality of the MF approach from a mathematical viewpoint.

Let

$$\mathcal{U}_{l,i} = \Big\{u_i \mid u_i(t) \in \sigma(x_i(s), \theta_i(s), s \le t), \quad E\|x_i(T)\| = o(\sqrt{T}),\\ E\int_0^T \|x_i(t)\|^2 dt = O(T),\ T \to \infty\Big\};$$

$$\mathcal{U}_{g,i} = \Big\{u_i \mid u_i(t) \in \sigma(x_i(s), \theta_i(s), 1 \le i \le N, s \le t), E\|x_i(T)\| = o(\sqrt{T}),\\ E\int_0^T \|x_i(t)\|^2 dt = O(T),\ T \to \infty\Big\}.$$

DEFINITION 5.3 (see [1]). *For a given $\varepsilon \ge 0$, a group of strategies $\{u_i \in \mathcal{U}_{l,i}, 1 \le i \le N\}$ is called an ε-Nash equilibrium with respect to the group of cost functions $\{J_i, 1 \le i \le N\}$ if for any $1 \le i \le N$,*

$$J_i(u_i, u_{-i}) \le \inf_{u_i' \in \mathcal{U}_{g,i}} J_i(u_i', u_{-i}) + \varepsilon.$$

Now we give the result of asymptotic optimality of the strategies (3.13).

THEOREM 5.4. *For the system* (2.1) *with the cost function* (2.2), *if assumptions* (A1)–(A4) *hold, then the group of strategies* (3.13) *is an ε-Nash equilibrium, where $\varepsilon = O(1/\sqrt{N})$.*

To prove Theorem 5.4, we need the following lemma, whose proof is in Appendix B.

LEMMA 5.5. *For the system* (2.1) *with the cost function* (2.2), *if assumptions* (A1)–(A4) *hold, then under the strategies* (3.13) *we have*

$$J_i(u_i^*, u_{-i}^*) \le J_i(u_i^*, v^*) + O(\varepsilon), \tag{5.3}$$

$$J_i(u_i^*, v^*) \le \inf_{u_i \in \mathcal{U}_{g,i}} J_i(u_i, u_{-i}^*) + O(\varepsilon) + O(N^{-1}), \tag{5.4}$$

where

$$\varepsilon = \left(\limsup_{T\to\infty} \frac{1}{T} E\int_0^T \left\|v^* - \Psi^{(N)}(x^*)\right\|^2 dt\right)^{\frac{1}{2}} = O\big(1/\sqrt{N}\big),$$

$$J_i(u_i^*, u_{-i}^*) = \limsup_{T\to\infty} \frac{1}{T} E\int_0^T \Big(\big\|x_i^* - \Phi\big[\Psi^{(N)}(x^*)\big]\big\|^2_{Q_{\theta_i(t)}} + \|u_i^*(t)\|^2_{R_{\theta_i(t)}}\Big)dt,$$

$$J_i(u_i^*, \Phi(v^*)) = \limsup_{T\to\infty} \frac{1}{T} E\int_0^T \Big(\|x_i^* - \Phi(v^*)\|^2_{Q_{\theta_i(t)}} + \|u_i^*(t)\|^2_{R_{\theta_i(t)}}\Big)dt,$$

$$J_i(u_i, u_{-i}^*) = \limsup_{T\to\infty} \frac{1}{T} E\int_0^T \left(\left\|x_i - \Phi\left[\frac{1}{N}\sum_{j\ne i} \Psi(x_j^*) + \Psi(x_i)\right]\right\|^2_{Q_{\theta_i(t)}}\right.\\ \left. + \|u_i^*(t)\|^2_{R_{\theta_i(t)}}\right)dt.$$

Here x_i is referred to the closed-loop solution of the system (2.1) *corresponding to some control strategy $u_i \in \mathcal{U}_{g,i}$.*

Proof of Theorem 5.4. From Lemma 5.5 it follows that

$$J_i(u_i^*, u_{-i}^*) \le \inf_{u_i \in \mathcal{U}_{g,i}} J_i(u_i, u_{-i}^*) + O(\varepsilon) + O(N^{-1}),$$

which together with Definition 5.3 leads to the conclusion. ∎

Remark 5.6. Theorem 5.4 implies that $\{u_i^*, i = 1, \ldots, N\}$ is an ε-Nash equilibrium. Noting that $\mathcal{U}_{g,i}$ is a centralized strategy set, by Definition 5.3 the group of distributed strategies $\{u_i^*, i = 1, \ldots, N\}$ is asymptotically optimal in Nash equilibrium sense with respect to the centralized strategy set.

6. Discussion on the case of costs with linear effect functions. In general, assumption (A4) of Theorem 3.4 is hard to verify directly. However, for some special case where the effect function Ψ is linear, we can get some conditions that are easier to verify.

In this section, for the system (2.1) with the cost function (2.2), we will consider the case where the effect function Ψ is linear. That is, the cost function for agent i is

$$\begin{aligned} &J_i(u_i, u_{-i}) \\ (6.1)\qquad &= \limsup_{T\to\infty} \frac{1}{T} E \int_0^T \Big\{ \big\| x_i(t) - \Phi(Hx^{(N)} + \beta) \big\|^2_{Q_{\theta_i(t)}} + \|u_i(t)\|^2_{R_{\theta_i(t)}} \Big\} dt, \end{aligned}$$

where $x^{(N)} = \frac{1}{N}\sum_{i=1}^N x_i(t), H \in \mathbb{R}^{n\times n}, \beta \in \mathbb{R}^n$.

To do so, we need the following assumptions:

(A1$'$) $\{x_i(0), 1 \le i \le N\}$ are independent r.v.s with the identical expectation and independent of $\{W_i(t)\}$ and $\{\theta_i(t)\}$. For any $1 \le i \le N$, $Ex_i(0) = x_0, E\|x_i(0)\|^2 < \infty$. $\{\theta_i(0)\}$ are independent r.v.s with the identical distribution $\{p_j^0, 1 \le j \le m\}$.

(A4$'$) $m\|H\|L_\Phi \max_j \|Q_j\| \max_j \|B_j R_j^{-1} B_j^T\| \big[\int_0^\infty \|\exp[(G^T + \Lambda \otimes I)t]\| dt\big]^2 < 1$.

Remark 6.1. Compared with (A1) and (A4), assumption (A1$'$) is more general and assumption (A4$'$) is easier to verify. Particularly, for a 1-dimensional time-invariant system, (A4$'$) holds if and only if

$$\|H\|L_\Phi \le 1 + \frac{A^2 R}{B^2 Q}.$$

Now we provide a group of distributed strategies for the system (2.1) with the cost function (6.1) by using the MF approach.

First, similar to section 3.2, we consider an auxiliary Stackelberg game with two players. The state also evolves according to (3.5). The cost functions of two players are, respectively,

$$J_1(u, v) = \limsup_{T\to\infty} \frac{1}{T} E \int_0^T \Big\{ \|x(t) - \Phi(v(t))\|^2_{Q_{\theta(t)}} + \|u(t)\|^2_{R_{\theta(t)}} \Big\} dt, \tag{6.2}$$

$$J_2(u, v) = \limsup_{T\to\infty} \frac{1}{T} \int_0^T \|v(t) - H \cdot Ex(t) - \beta\|^2 dt. \tag{6.3}$$

Now we solve problem (3.5), (6.2), and (6.3) by the brute-force method. First, for every fixed $v \in \mathcal{V} = C_b([0,\infty), \mathbb{R}^n)$, by Theorem 3.1 we get the optimal tracking strategy for player 1,

$$u^v(t) = -R^{-1}_{\theta(t)} B^T_{\theta(t)} (K_{\theta(t)} x(t) + r^v(\theta(t), t)), \tag{6.4}$$

where

$$\begin{aligned} r_j^v(t) = e_j^T \otimes I \cdot \int_t^\infty &\exp[-(G^T + \Lambda \otimes I)(t - s)] \\ &\times \big\{ K[\mathbf{1} \otimes h(s)] - Q[\mathbf{1} \otimes \Phi(v(s))] \big\} ds. \end{aligned} \tag{6.5}$$

Substituting (6.4) into (3.5), we get the closed-loop equation

$$dx(t) = G_{\theta(t)}x(t)dt - B_{\theta(t)}R^{-1}_{\theta(t)}B^T_{\theta(t)}r^v(\theta(t),t) + h(t)dt + D_{\theta(t)}dW(t).$$

If the equation

$$H \cdot Ex^v + \beta = v \tag{6.6}$$

admits a solution v^*, then

$$J_2(u^{v^*}, v^*) = \inf_{v\in\mathcal{V}} J_2(u^v, v) = 0.$$

Next we investigate the existence conditions for solutions of (6.6). Let $\varphi_j(t) = I_{[\theta(t)=j]}, j = 1,\dots,m$. Then, from [26] it follows that

$$\varphi_j(t) = \varphi_j(0) + \sum_{k=1}^m \int_0^t \lambda_{kj}\varphi_k(s)ds + M_j(t), j = 1,\dots,m,$$

where $M_j(t)$ is a square-integrable martingale. By Itô's formula we have

$$\begin{aligned} x(t)\varphi_j(t) = &\int_0^t \varphi_j(s)[G_{\theta(s)}x(s) + h(s) - B_{\theta(s)}R^{-1}_{\theta(s)}B^T_{\theta(s)}r^v(\theta(s),s)]ds \\ &+ \int_0^t x(s)\sum_{k=1}^m \lambda_{kj}\varphi_k(s)ds + \int_0^t \varphi_j(s)D_{\theta(s)}dW(s) \\ &+ \int_0^t x(s)dM_j(s) + \int_0^t D_{\theta(s)}d\langle W, M_j\rangle(s), \quad 1 \le j \le m, \end{aligned} \tag{6.7}$$

where $\langle W, M_j\rangle(t)$ is the cross-variation process between $W(t)$ and $M_j(t)$. Since $\{W(t)\}$ and $\{\theta(t)\}$ are independent of each other, we have $\langle W, M_j\rangle(t) = 0$. Letting $L_j(t) = E[x(t)\varphi_j(t)]$ and $p_j(t) = P(\theta(t) = j)$ and taking expectations on both sides of (6.7), we get

$$\begin{aligned} L_j(t) = &\int_0^t G_jL_j(s)ds + \int_0^t \sum_{k=1}^m \lambda_{kj}L_k(s)ds + \int_0^t p_j(s)h(s)ds \\ &- \int_0^t p_j(s)B_jR_j^{-1}B_j^Tr_j^v(s)ds, \end{aligned}$$

and hence,

$$\frac{dL_j(t)}{dt} = G_jL_j(t) + \sum_{k=1}^m \lambda_{kj}L_k(t) + p_j(t)h(t) - p_j(t)B_jR_j^{-1}B_j^Tr_j^v(t). \tag{6.8}$$

Let

$$\begin{aligned} L(t) &= [L_1^T(t),\dots,L_m^T(t)]^T, \\ p(t) &= [p_1(t),\dots,p_m(t)]^T, \\ BR^{-1}B^Tp(t) &= diag\{B_1R_1^{-1}B_1^Tp_1(t),\dots,B_mR_m^{-1}B_m^Tp_m(t)\}, \\ r^v(t) &= [r_1^v(t)^T,\dots,r_m^v(t)^T]^T, \\ K &= diag\{K_1,\dots,K_m\}. \end{aligned}$$

Then, by (6.8) we have

$$L(t) = L(0)\exp[(G + \Lambda^T \otimes I)t] + \int_0^t \exp[(G + \Lambda^T \otimes I)(t - s)][p(s) \otimes h(s) - BR^{-1}B^T p(s) r^v(s)]ds,$$

where $L(0) = [x_0^T p_1^0, \ldots, x_0^T p_m^0]^T$,

$$r^v(s) = \int_s^\infty \exp[(G^T + \Lambda \otimes I)(\tau - s)]\{K[\mathbf{1} \otimes h(\tau)] - Q[\mathbf{1} \otimes \Phi(v(\tau))]\}d\tau.$$

For any $v \in C_b([0,\infty), \mathbb{R}^n)$, define an operator $\mathcal{T}_1$,

$$(\mathcal{T}_1 v)(t) \triangleq \beta + H(\mathbf{1}^T \otimes I)\Bigg\{L(0)\exp[(G + \Lambda^T \otimes I)t] + \int_0^t \exp[(G + \Lambda^T \otimes I)(t - s)] \Bigg[p(s) \otimes h(s) + BR^{-1}B^T p(s) \times \int_s^\infty \exp[(G^T + \Lambda \otimes I)(\tau - s)]\big[Q(\mathbf{1} \otimes \Phi(v(\tau))) - K(\mathbf{1} \otimes h(\tau))\big]d\tau\Bigg]ds\Bigg\}. \tag{6.9}$$

Then, from Theorem 3.1(iii) it follows that $\sup_{t\geq 0} \|(\mathcal{T}_1 v)(t)\| < \infty$. Thus, $\mathcal{T}_1$ is an operator on $C_b([0,\infty), \mathbb{R}^n)$. By the definition of $L(t)$ and $\mathcal{T}_1$, (6.6) can be written as

$$\mathcal{T}_1 v = v. \tag{6.10}$$

Now we give the theorem of existence and uniqueness for v^*.

THEOREM 6.2. *Under assumptions* (A1′), (A2), (A3), *and* (A4′), (6.10) *admits a unique solution* v^* *in* $C_b([0,\infty), \mathbb{R}^n)$.

Proof. For any $y, z \in C_b([0,\infty), \mathbb{R}^n)$, from the definition of $\mathcal{T}_1$ it follows that

$$(\mathcal{T}_1 y)(t) - (\mathcal{T}_1 z)(t) = H(\mathbf{1}^T \otimes I)\int_0^t \exp[(G + \Lambda^T \otimes I)(t - s)]BR^{-1}B^T p(s) \times \int_s^\infty \exp[(G^T + \Lambda \otimes I)(\tau - s)]Q[\mathbf{1} \otimes (\Phi(y) - \Phi(z))]d\tau ds,$$

which renders

$$\begin{aligned}&\|(\mathcal{T}_1 y)(t) - (\mathcal{T}_1 z)(t)\| \\ &\quad \leq m\|H\|L_\Phi \max_j \|B_j R_j^{-1} B_j^T\| \max_j \|Q_j\| \|y - z\|_\infty \\ &\quad \times \int_0^t \|\exp[(G^T + \Lambda \otimes I)(t - s)]\| \int_s^\infty \|\exp[(G + \Lambda^T \otimes I)(\tau - s)]\| d\tau ds,\end{aligned}$$

and

$$\|(\mathcal{T}_1 y) - (\mathcal{T}_1 z)(t)\|_\infty \leq m\|H\|L_\Phi \max_j \|B_j R_j^{-1} B_j^T\| \max_j \|Q_j\| \times \left[\int_0^\infty \|\exp[(G^T + \Lambda \otimes I)t]\|dt\right]^2 \|y - z\|_\infty.$$

This together with assumption (A4′) implies that $\mathcal{T}_1$ is a contraction on $C_b([0,\infty), \mathbb{R}^n)$. Hence, (6.10) has a unique solution v^*. ∎

7. Numerical examples. First, we give a simulation for the practical example mentioned in the introduction and Remark 2.1.

For convenience of computation, we only consider the simplest case: the production line of each company merely has two states, "work" or "failure." The output level of the ith company evolves according to

$$dx_i(t) = a(\theta_i(t))x_i(t)dt + (u_i(t)+1)dt + dw_i(t), \tag{7.1}$$

where $x_i \in \mathbb{R}, 1 \le i \le N$, $\{\theta_i(t)\}$ is a sequence of independent Markov chains taking values in $\{1,2\}$ with the identical infinitesimal generator

$$\Lambda = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$

and $\{w_i(t)\}$ is a sequence of independent 1-dimensional standard Brownian motions. The cost function of the ith company is

$$\limsup_{T\to\infty} \frac{1}{T} E \int_0^T \Big\{[x_i - \phi(\beta - \alpha x^{(N)})]^2 + u_i^2\Big\} dt, \tag{7.2}$$

where ϕ is Lipschitz continuous with Lipschitz constant L_ϕ.

Take $a_1 = 0.1, a_2 = 0.7$. Then, the Riccati equation (3.3) has the unique positive solutions $k_1 = 0.8, k_2 = 0.6$. By direct computation, we have that if $|\alpha|L_\phi \in (0, 0.83]$, then both assumption (A4) and assumption (A4$'$) hold. Let $\phi(x) = 2x$, $\beta = 5, \alpha = 0.25$, and $\{x_i(0)\}$ be i.i.d. r.v.s with the normal distribution $N(2,2)$ and $\{\theta_i(0)\}$ be independent r.v.s satisfying $P(\theta_i(0) = 1) = 0.5$ and $P(\theta_i(0) = 2) = 0.5$. Then assumptions (A1)–(A4) hold.

For the above model and cost function, v^* can be solved in an explicit form. Now we give v^*.

From (6.5), (6.6), and (6.8) it follows that

$$\frac{dr}{dt} + Fr(t) - 2\begin{pmatrix} v(t) \\ v(t) \end{pmatrix} + \begin{pmatrix} 0.8 \\ 0.6 \end{pmatrix} = 0, \tag{7.3}$$

$$\frac{dL}{dt} = F^T L(t) - \frac{1}{2} r(t) + \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}, \tag{7.4}$$

$$v(t) = 5 - 0.25(1 \;\; 1)L(t), \tag{7.5}$$

where

$$F = G^T + \Lambda = \begin{pmatrix} -1.9 & 1 \\ 1 & -2.3 \end{pmatrix}.$$

Letting the derivative term be zero, we obtain the steady-state equation set

$$\begin{cases} Fr_\infty - 2\begin{pmatrix} v_\infty \\ v_\infty \end{pmatrix} = -\begin{pmatrix} 0.8 \\ 0.6 \end{pmatrix}, \\ FL_\infty - \frac{1}{2} r_\infty = -\begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}, \\ v_\infty = 5 - (0.25 \quad 0.25)L_\infty. \end{cases} \tag{7.6}$$

Solving (7.6) directly, we have $r_\infty = [-6.0393 \;\; -5.3679]^T, L_\infty = [3.3469 \;\; 2.8395]^T$. Let $\tilde{L}_t = L_t - L_\infty, \tilde{r}_t = r_t - r_\infty$. Then, by (6.5), (7.4), and (7.5) we have

$$\frac{d\tilde{L}}{dt} = \frac{dL}{dt} = F\tilde{L}(t) - \frac{1}{4}\int_t^\infty \exp[F(s-t)] \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \tilde{L}(s)ds. \tag{7.7}$$

Differentiating both sides of (7.7) leads to

$$\frac{d^2\tilde{L}}{dt^2} = F\frac{d\tilde{L}}{dt} + \frac{1}{4}F\int_t^\infty \exp[F(s-t)]\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\tilde{L}(s)ds + \frac{1}{4}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\tilde{L}(t),$$

which together with (7.7) gives

$$\frac{d^2\tilde{L}}{dt^2} = \left[F^2 + \frac{1}{4}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\right]\tilde{L}(t). \tag{7.8}$$

This has two solutions $e^{\sqrt{M}t}\tilde{L}_0$ and $e^{-\sqrt{M}t}\tilde{L}_0$ with

$$M = F^2 + \frac{1}{4}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 4.86 & -3.95 \\ -3.95 & 6.54 \end{pmatrix}.$$

By $v^* \in C_b([0,\infty),\mathbb{R})$ and (7.5), we should have

$$\tilde{L}_t = e^{-\sqrt{M}t}(L_0 - L_\infty).$$

Then

$$v_t^* = (1 \;\; 1)[L_\infty + e^{-\sqrt{M}t}(L_0 - L_\infty)].$$

In addition, from (7.3), (7.5), and (7.6) it follows that

$$\frac{d\tilde{r}}{dt} + F\tilde{r}(t) + \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\tilde{L}(t) = 0.$$

By Theorem 3.1(iii) we have

$$\tilde{r}(t) = \frac{1}{2}\int_t^\infty e^{F(s-t)}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}e^{-\sqrt{M}s}(L_0 - L_\infty)ds.$$

Thus, by (3.13) we get the distributed strategies

$$u_i^*(t) = -\sum_{j=1}^{2} I_{[\theta_i(t)=j]}(k_j x_i(t) + r_j^*(t)),\;\; 1 \le i \le N,$$

where $k_1 = 0.8, k_2 = 0.6$,

$$r_j^*(t) = e_j\left[\frac{1}{2}\int_t^\infty e^{F(s-t)}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}e^{-\sqrt{M}s}(L_0 - L_\infty)ds + r_\infty\right].$$

When the number of agents $N = 100$, the trajectories of closed-loop system are demonstrated in Figure 7.1. It can be seen that even though all the agents achieve an agreement, the trajectories of their states fluctuate within a certain range due to the influence of Markov jump parameters. The curves of v^* and $\beta - \alpha x^{(100)} = 5 - \sum_{i=1}^{100}\frac{x_i^*}{400}$ are shown in Figure 7.2. As the number of the agents grows from 1 to 500, the curve of ε is shown in Figure 7.3, where

$$\varepsilon(N) = \left(\limsup_{T\to\infty}\frac{1}{T}E\int_0^T \left\|v^* - (\beta - \alpha x^{(N)})\right\|^2 dt\right)^{\frac{1}{2}}.$$

From Figures 7.2 and 7.3, it can be seen that v^* is a good approximation of $\beta - \alpha x^{(N)}$.

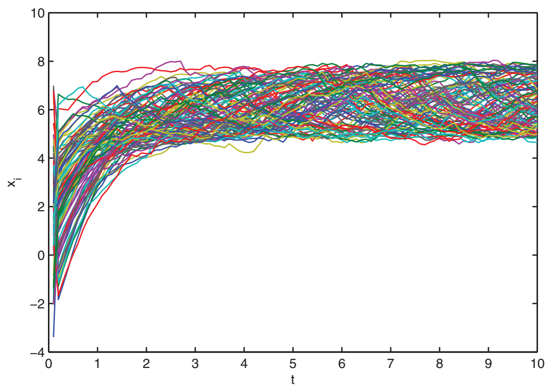

FIG. 7.1. *Trajectories of agents' states when* $N = 100$.

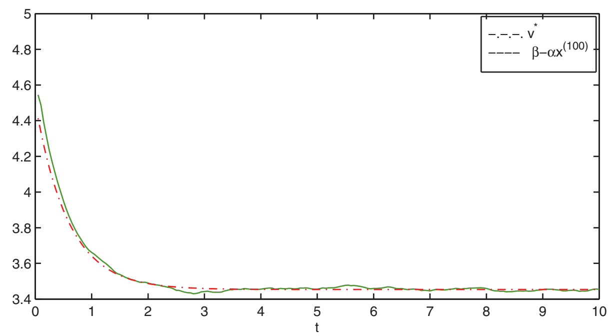

FIG. 7.2. *Curves of* v^* *and* $\beta - \alpha x^{(100)}$.

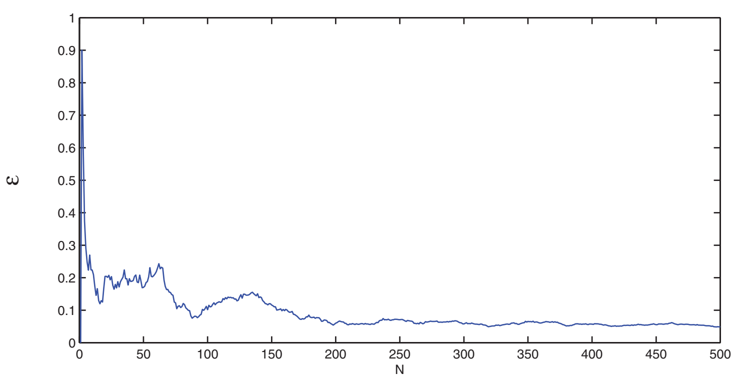

FIG. 7.3. *Curve of* ε *with respect to* N.

Next, taking the system (7.1)–(7.2) as an example, we analyze how the intensity of heterogeneity (i.e., the size of $|a_1 - a_2|$) and the jump rate of each agent's parameters affect the closed-loop system.

Case I. Taking the infinitesimal generator of $\{\theta_i(t), 1 \leq i \leq N\}$ as

$$\Lambda = \begin{pmatrix} -0.5 & 0.5 \\ 0.5 & -0.5 \end{pmatrix},$$

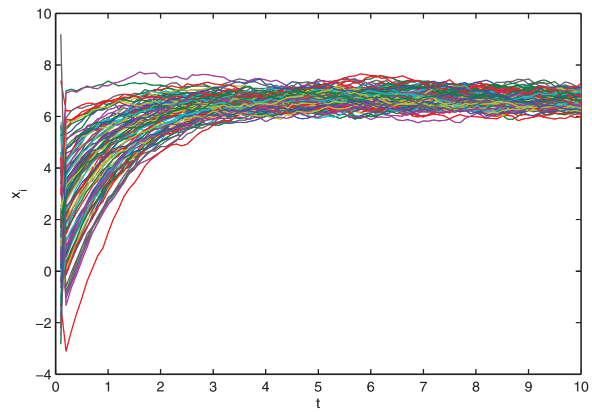

FIG. 7.4. *Trajectories of agents' states in Case I when* $N = 100$.

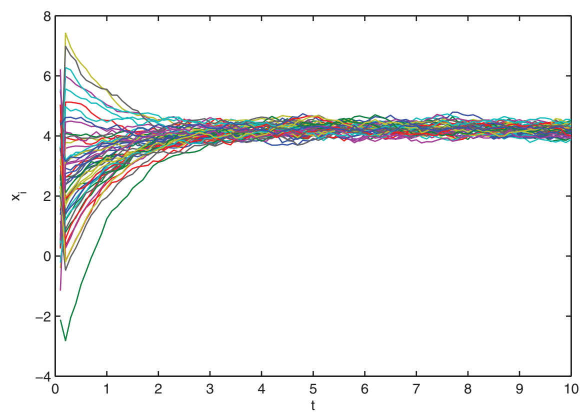

FIG. 7.5. *Trajectories of agents' states in Case II when* $N = 100$.

$a_1 = -0.078$ and $a_2 = -0.256$ when the number of agents $N = 100$; the trajectories of a closed-loop system are demonstrated in Figure 7.4.

Case II. Taking $\Lambda = \mathbf{0}$, $a_1 = a_2 = 0.4$ (time-invariant case), when the number of the agents $N = 100$, the trajectories of the closed-loop system are demonstrated in Figure 7.5.

From Figure 7.1, Figure 7.4, and Figure 7.5, we can see that, as the heterogeneity intensity is weak (i.e., $|a_1 - a_2|$ is small) and jump rate of the parameters is small, the consensus degree becomes high.

8. Concluding remarks. The paper studies the distributed games for large-population MASs with random time-varying parameters. We have provided a group of distributed strategies with respect to a set of tracking-type quadratic cost functions and proved the uniform stability of the closed-loop systems and asymptotic optimality of the distributed strategies in the sense of Nash equilibrium, which extends previous works [24, 16] from the time-invariant case to the case of Markov jump parameters. Through a numerical example, we show how to design the control strategies and how the heterogeneity intensity and the parameter transition rate of each agent affect the closed-loop system.

For large-population MASs, there are a lot of interesting problems worth investigating, for instance, how to design distributed strategies for the cases with unknown parameters or measurement errors and the case where measurement information is limited.

Appendix A. Proof of Theorem 3.1.

Proof of Theorem 3.1. The proof of (i) can be found in [20]. Thus, we only need to show (ii)–(iv).

(ii) Let $\varphi_j(t) = I_{[\theta(t)=j]}, j = 1, \ldots, m$. Then, by [26] we have

$$\varphi_j(t) = \varphi_j(0) + \sum_{k=1}^{m} \int_0^t \lambda_{kj}\varphi_k(s)ds + M_j(t),$$

where $M_j(t)$ is a square-integrable martingale. Let $y(t) \in \mathbb{R}^n$ satisfy the following equation

$$\frac{dy(t)}{dt} = G_{\theta(t)}y(t), \tag{A.1}$$

where $G_j = A_j - B_jR_j^{-1}B_j^TK_j, j = 1, \ldots, m$. Then, by Theorem 5 of [20], (A.1) is mean-square stable, i.e., for any initial state $y(0)$, $E\|y(t)\|^2 \to 0$, $t \to \infty$ holds. By Itô's formula we obtain

$$\begin{aligned} y(t)\varphi_j(t) = &\int_0^t G_{\theta(s)}\varphi_j(s)y(s)ds \\ &+ \int_0^t y(s)\sum_{k=1}^{m}\lambda_{kj}\varphi_k(s)ds + \int_0^t y(s)dM_j(s). \end{aligned} \tag{A.2}$$

Letting $S_j(t) = E(y(t)\varphi_j(t))$, $S(t) = (S_1^T(t), \ldots, S_m^T(t))^T$, and taking the expectations for both sides of (A.2), we have

$$\frac{dS(t)}{dt} = (G + \Lambda^T \otimes I)S(t).$$

Since (A.1) is mean-square stable, $S(t)$ is asymptotically stable, i.e., the above equation is asymptotically stable. Thus, the matrix $G + \Lambda^T \otimes I$ is stable.

(iii) Let

$$r(t) = [r_1(t)^T, \ldots, r_m(t)^T]^T,$$

$$K = diag(K_1, \ldots, K_m), \quad Q = diag(Q_1, \ldots, Q_m).$$

Then (3.4) can be rewritten as

$$\frac{dr(t)}{dt} + (G^T + \Lambda \otimes I)r(t) + K(\mathbf{1} \otimes h(t)) - Q(\mathbf{1} \otimes f^b(t)) = 0. \tag{A.3}$$

The general solution of (A.3) can be expressed as

$$\begin{aligned} r(t) = &\exp[-(G^T + \Lambda \otimes I)t]r(0) \\ &+ \int_0^t \exp[-(G^T + \Lambda \otimes I)(t-s)][Q(\mathbf{1} \otimes f^b(s)) - K(\mathbf{1} \otimes h(s))]ds. \end{aligned}$$

Since all the eigenvalues of $G^T + \Lambda \otimes I$ have negative real parts, there exist $\kappa > 0$ and $\rho > 0$ such that $\|\exp[(G^T + \Lambda \otimes I)t]\| \leq \kappa e^{-\rho t}, \forall t \geq 0$. Taking $r(0) = -\int_0^\infty \exp[(G^T +$

$\Lambda \otimes I)s][Q(\mathbf{1} \otimes f^b(s)) - K(\mathbf{1} \otimes h(s))]ds \triangleq r^*(0)$, we get

$$r(t) = -\int_t^\infty \exp[-(G^T + \Lambda \otimes I)(t-s)][Q(\mathbf{1} \otimes f^b(s)) - K(\mathbf{1} \otimes h(s))]ds.$$

This implies $r(t) \in C_b([0,\infty), \mathbb{R}^{nm})$. Since all the eigenvalues of $-(G^T + \Lambda \otimes I)$ have positive real parts, the solution of (A.3) is unique in $C_b([0,\infty), \mathbb{R}^{nm})$.

(iv)–(v) Let $\mathcal{A}_t$ be the infinitesimal generator of $(\theta(t), x(t))$ that satisfies (3.1). Then, by Dynkin's formula [30] we can get

$$\begin{aligned} &E\int_0^T x^T(t)Q_{\theta(t)}x(t)dt \\ &\quad = E\Bigg\{x^T(0)K_{\theta(0)}x(0) - x^T(T)K_{\theta(T)}x(T) \\ &\qquad + \int_0^T \big[x^T(t)K_{\theta(t)}B_{\theta(t)}R^{-1}_{\theta(t)}B^T_{\theta(t)}K_{\theta(t)}x(t) \\ &\qquad\qquad + 2x^T(t)K_{\theta(t)}(B_{\theta(t)}u(t) + h(t)) + tr(K_{\theta(t)}D_{\theta(t)}D^T_{\theta(t)})\big]dt\Bigg\} \end{aligned}$$

and

$$\begin{aligned} &E\int_0^T [K_{\theta(t)} - Q_{\theta(t)}f^b(t)h(t)]^T x(t)dt \\ &\quad = E\Bigg\{r^T(\theta(0),0)x(0) - r^T(\theta(T),T)x(T) \\ &\qquad + \int_0^T \big[(B_{\theta(t)}R^{-1}_{\theta(t)}B^T_{\theta(t)}K_{\theta(t)})^T x(t) + r^T(\theta(t),t)(B_{\theta(t)}u(t) + h(t))\big]dt\Bigg\}. \end{aligned}$$

This implies that

$$\begin{aligned} J(u) &= \limsup_{T\to\infty} \frac{1}{T} E\int_0^T \Big\{\|u(t) + R^{-1}_{\theta(t)}B^T_{\theta(t)}[K_{\theta(t)}x(t) + r(\theta(t),t)]\|^2_{R_{\theta(t)}} \\ &\qquad + \|f^b(t)\|^2_{Q_{\theta(t)}} - \|B^T_{\theta(t)}r(\theta(t),t)\|^2_{R^{-1}_{\theta(t)}} \\ &\qquad - 2r^T(\theta(t),t)h(t) + tr(K_{\theta(t)}D_{\theta(t)}D^T_{\theta(t)})\Big\}dt \\ &\geq \limsup_{T\to\infty} E\int_0^T [\|f^b(t)\|^2_{Q_{\theta(t)}} - \|B^T_{\theta(t)}r(\theta(t),t)\|^2_{R^{-1}_{\theta(t)}} - 2r^T(\theta(t),t)h(t) \\ &\qquad + tr(K_{\theta(t)}D_{\theta(t)}D^T_{\theta(t)})]dt \triangleq J^*. \end{aligned} \tag{A.4}$$

Taking $u^*(t) = -R^{-1}_{\theta(t)}B^T_{\theta(t)}[K_{\theta(t)}x(t) + r(\theta(t),t)]$, the closed-loop system becomes

$$\begin{aligned} dx(t) &= G_{\theta(t)}x(t)dt - B_{\theta(t)}R^{-1}_{\theta(t)}B^T_{\theta(t)}r(\theta(t),t) \\ &\quad + h(t)dt + D_{\theta(t)}dW(t). \end{aligned} \tag{A.5}$$

By [20], [28], and Theorem 3.1(iii), we have $E\|x(T)\| = o(\sqrt{T})$ and

$$\limsup_{T\to\infty} \frac{1}{T} E\int_0^T \|x(t)\|^2 dt < C,$$

which gives $u^* \in \mathcal{U}$. Thus, (iv) and the first part of (v) hold.

Now we prove the second part of (v). Let $p_j(t) = P(\theta(t) = j), j = 1, \dots, m$. Since π is the stationary distribution of $\{\theta(t)\}$, then $p_j(t) \to \pi_j (t \to \infty)$, i.e., for any $\varepsilon > 0$, there exist $t_0 > 0$ such that for any $t > t_0$, $|p_j(t) - \pi_j| \le \varepsilon$. For any $j = 1, \dots, m$, let

$$\begin{aligned} l(j,t) = {} & {f^b}^T(t) Q_j f^b(t) - 2r^T(j,t)h(t) \\ & - r^T(j,t) B_j R_j^{-1} B_j^T r(j,t) + tr(K_j D_j D_j^T). \end{aligned} \tag{A.6}$$

Then $J(u^*) = \limsup_{T\to\infty} \frac{1}{T} E \int_0^T l(\theta(t),t)dt$. Noting $El(\theta(t),t) = \sum_{j=1}^m l(j,t)p_j(t)$ and $L \triangleq \max_j \sup_{t>0} |l(j,t)| < \infty$, we can see that for all sufficiently large T,

$$\begin{aligned} & \left| \frac{1}{T} E \int_0^T l(\theta(t),t)dt - \frac{1}{T} \int_0^T \sum_{j=1}^m \pi_j l(j,t)dt \right| \\ & \le \frac{1}{T} \left(\int_0^{t_0} + \int_{t_0}^T \right) \sum_{j=1}^m |l(j,t)||p_j(t) - \pi_j|dt \\ & \le 2mL \frac{t_0}{T} + mL\varepsilon \frac{T - t_0}{T} \le (1 + mL)\varepsilon. \end{aligned}$$

Thus

$$J(u^*) = \sum_{j=1}^m \pi_j \limsup_{T\to\infty} \frac{1}{T} E \int_0^T l(j,t)dt.$$

By (A.6), the second part of (v) holds. □

Appendix B. Proof of Lemma 5.5.

To prove Lemma 5.5, we need the following lemma.

LEMMA B.1. *If assumptions* (A1)–(A4) *hold, then under the strategies* (3.13), *there exist constants* C_1 *and* C_2 *independent of* N *such that*

$$\max_{1\le i\le N} J_i(u_i^*, u_{-i}^*) \le C_1, \tag{B.1}$$

$$\max_{1\le i\le N} \limsup_{T\to\infty} \frac{1}{T} E \int_0^T \left\| \frac{1}{N} \sum_{j\ne i} \Psi(x_j^*(t)) \right\|^2_{Q_{\theta_i(t)}} dt \le C_2. \tag{B.2}$$

Proof. Since (B.1) can be derived easily from (B.2), we only need to prove (B.2). Noting

$$\begin{aligned} & \limsup_{T\to\infty} \frac{1}{T} E \int_0^T \left\| \frac{1}{N} \sum_{j\ne i} \Psi(x_j^*(t)) \right\|^2_{Q_{\theta_i(t)}} dt \\ & \le \limsup_{T\to\infty} \frac{1}{T} E \int_0^T \left(\frac{N-1}{N} \right)^2 \max_{l\le m} \|Q_l\| \frac{1}{N-1} \sum_{j\ne i} \left\| \Psi(x_j^*) \right\|^2 dt \\ & \le \max_{l\le m} \|Q_l\| \max_{1\le j\le N} \limsup_{T\to\infty} \frac{1}{T} E \int_0^T \left\| \Psi(x_j^*) \right\|^2 dt \\ & \le \max_{l\le m} \|Q_l\| (2L_\Psi^2 C_0 + 2\|\Psi(0)\|^2) \end{aligned}$$

and taking $C_2 = \max_{l\le m} \|Q_l\|(2L_\Psi^2 C_0 + 2\|\Psi(0)\|^2)$, we obtain (B.2).

Proof of Lemma 5.5. From

$$\begin{aligned} &J_i(u_i^*, u_{-i}^*) \\ &= \limsup_{T\to\infty} \frac{1}{T} E \int_0^T \Big(\big\| x_i^* - \Phi(v^*) + \Phi(v^*) - \Phi\big(\Psi^{(N)}(x^*)\big) \big\|^2_{Q_{\theta_i(t)}} + \|u_i^*(t)\|^2_{R_{\theta_i(t)}} \Big) dt \\ &\le J_i(u_i^*, \Phi(v^*)) + 2 \max_{l\le m} \|Q_l\| \varepsilon \left(\limsup_{T\to\infty} \frac{1}{T} E \int_0^T \|x_i^* - \Phi(v^*)\|^2 dt \right)^{\frac{1}{2}} \\ &\quad + \max_{l\le m} \|Q_l\| L_\Phi^2 (\varepsilon)^2 \\ &= J_i(u_i^*, \Phi(v^*)) + O(\varepsilon) \end{aligned}$$

it follows that (5.3) holds. Moreover, from the proof of Lemma 5.1, we have $\varepsilon = O(1/\sqrt{N})$.

Since $u_i^* \in \mathcal{U}_{l,i} \subseteq \mathcal{U}_{g,i}$, $\inf_{u_i\in\mathcal{U}_{g,i}} J_i(u_i, u_{-i}^*) \le J_i(u_i^*, u_{-i}^*)$. Thus, to prove (5.4), we only need to prove that for any $u_i \in \mathcal{U}'_{g,i} \triangleq \{u_i \in \mathcal{U}_{g,i} | J_i(u_i, u_{-i}^*) \le J_i(u_i^*, u_{-i}^*)\}$, the following inequality holds:

$$J_i(u_i^*, \Phi(v^*)) \le J_i(u_i, u_{-i}^*) + O(\varepsilon) + O(N^{-1}). \tag{B.3}$$

Let $u_i \in \mathcal{U}'_{g,i}$. Then, from

$$\limsup_{T\to\infty} \frac{1}{T} E \int_0^T \left\| x_i - \Phi\left[\frac{1}{N} \sum_{j\ne i} \Psi(x_j^*) + \frac{1}{N}\Psi(x_i) \right] \right\|^2_{Q_{\theta_i(t)}} dt \le J_i(u_i, u_{-i}^*) \le C_1$$

and Lemma B.1 it follows that

$$\limsup_{T\to\infty} \frac{1}{T} E \int_0^T \left\| x_i - \Phi\left[\frac{1}{N}\Psi(x_i) \right] \right\|^2_{Q_{\theta_i(t)}} dt \le 2C_1 + 2C_2 L_\Phi^2.$$

Note that $\lim_{N\to\infty} \frac{1}{N}\Psi(x_i) = 0$. Then, by the above inequality we get

$$\limsup_{T\to\infty} \frac{1}{T} E \int_0^T \|x_i - \Phi(0)\|^2_{Q_{\theta_i(t)}} dt \le 2C_1 + 2C_2 L_\Phi^2,$$

which implies that

$$\limsup_{T\to\infty} \frac{1}{T} E \int_0^T \|x_i\|^2_{Q_{\theta_i(t)}} dt \le 4C_1 + 4C_2 L_\Phi^2 + 2 \max_{l\le m} \|Q_l\| \Phi^2(0) \triangleq C_3. \tag{B.4}$$

By direct calculation, we have

$$\begin{aligned} J_i(u_i, u_{-i}^*) &= \limsup_{T\to\infty} \frac{1}{T} E \int_0^T \Bigg(\Bigg\| x_i - \Phi(v^*) + \Phi(v^*) \\ &\qquad - \Phi\left[\frac{1}{N} \sum_{j\ne i} \Psi(x_j^*) + \Psi(x_i) \right] \Bigg\|^2_{Q_{\theta_i(t)}} + \|u_i^*(t)\|^2_{R_{\theta_i(t)}} \Bigg) dt \\ &\ge J_i(u_i, \Phi(v^*)) + \limsup_{T\to\infty} \frac{2}{T} E \int_0^T (x_i - \Phi(v^*))^T Q_{\theta_i(t)} \\ &\quad \left[\Phi(v^*) - \Phi\big(\Psi^{(N)}(x^*)\big) + \Phi\big(\Psi^{(N)}(x^*)\big) - \Phi\left(\frac{1}{N} \sum_{j\ne i} \Psi(x_j^*) + \frac{1}{N}\Psi(x_i) \right) \right] dt \\ &\ge J_i(u_i^*, \Phi(v^*)) + I_1 + I_2, \end{aligned} \tag{B.5}$$

where

$$I_1 = \limsup_{T\longrightarrow\infty} \frac{2}{T} E \int_0^T (x_i - \Phi(v^*))^T Q_{\theta_i(t)} [\Phi(v^*) - \Phi\big(\Psi^{(N)}(x^*)\big)] dt,$$

$$I_2 = \limsup_{T\longrightarrow\infty} \frac{2}{T} E \int_0^T (x_i - \Phi(v^*))^T Q_{\theta_i(t)} \Bigg[\Phi\big(\Psi^{(N)}(x^*)\big) - \Phi\Bigg(\frac{1}{N}\sum_{j\neq i}\Psi(x_j^*) + \frac{1}{N}\Psi(x_i)\Bigg)\Bigg] dt.$$

By Schwarz inequality and (B.4), we have

$$\begin{aligned} \|I_1\| &\leq 2L_\Phi \max_{l\leq m}\sqrt{\|Q_l\|}\varepsilon\Bigg(\limsup_{T\to\infty}\frac{1}{T}E\int_0^T (2\|x_i\|^2_{Q_{\theta_i(t)}} + 2\|\Phi(v^*)\|^2_{Q_{\theta_i(t)}})dt\Bigg)^{\frac{1}{2}} \\ &\leq 2\sqrt{2}L_\Phi \max_{l\leq m}\sqrt{\|Q_l\|}\varepsilon(C_3 + \max_{l\leq m}\|Q_l\|\|\Phi(v^*)\|^2_\infty)^{1/2} \\ &= O(\varepsilon), \end{aligned} \tag{B.6}$$

$$\begin{aligned} \|I_2\| &\leq 2\Bigg(\limsup_{T\to\infty}\frac{1}{T}E\int_0^T (\|x_i - \Phi(v^*)\|^2_{Q_{\theta_i(t)}})dt\Bigg)^{1/2} \\ &\quad\times\Bigg(\limsup_{T\to\infty}\frac{1}{T}E\int_0^T (\|\Phi\big(\Psi^{(N)}(x^*)\big) \\ &\qquad - \Phi\Bigg(\frac{1}{N}\sum_{j\neq i}\Psi(x_j^*) + \frac{1}{N}\Psi(x_i)\Bigg)\|^2_{Q_{\theta_i(t)}}dt\Bigg)^{1/2} \\ &\leq \frac{4}{N}L_\Phi L_\Psi\big(C_3 + \max_{l\leq m}\|Q_l\|\|\Phi(v^*)\|^2_\infty\big)^{1/2}\big(C_3 + C_0\max_{l\leq m}\|Q_l\|\big)^{1/2} \\ &= O(1/N). \end{aligned} \tag{B.7}$$

Noting that (B.5) is equivalent to

$$J_i(u_i^*, \Phi(v^*)) \leq J_i(u_i, u_{-i}^*) - I_1 - I_2,$$

by (B.6) and (B.7) one can get (B.3). Thus, Lemma 5.5 is true. ☐

REFERENCES

[1] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, Academic Press, London, 1982.

[2] C. T. Bauch and D. J. D. Earn, *Vaccination and the theory of games*, Proc. Natl. Acad. Sci. USA, 101 (2004), pp. 13391–13394.

[3] D. Bauso, L. Giarré, and R. Pesenti, *Non-linear protocols for optimal distributed consensus in networks of dynamics agents*, Systems Control Lett., 55 (2006), pp. 918–928.

[4] J. D. Birdwell, *On reliable control systems design*, PhD dissertation, Electronic Systems Laboratory, Massachusetts Institute of Technology, report ESL-TH-821, 1978.

[5] P. E. Caines, *Mean field stochastic control*, in Proceedings of the 48th IEEE Conference on Decision and Control, Shanghai, China, 2009; also available online from http://www.cim.mcgill.ca/~arman/Shanghai2009/Shanghai2009.pdf.

[6] C. Y. Chong, S. P. Kumar, and B. A. Hamilton, *Sensor networks: Evolution, opportunities, and challenges*, Proc. IEEE, 91 (2003), pp. 1247–1256.

[7] O. L. V. Costa, M. D. Fragoso, and R. P. Marques, *Discrete-Time Markov Jump Linear Systems*, Springer-Verlag, London, 2005.

[8] F. Cucker and S. Smale, *Emergent behavior in flocks*, IEEE Trans. Automat. Control, 52 (2007), pp. 852–862.

[9] G. M. Erickson, *Differential game models of advertising competition*, European J. Oper. Res., 83 (1995), pp. 431–438.
[10] G. M. Erickson, *An oligopoly model of dynamic advertising competition*, European J. Oper. Res., 197 (2009), pp. 374–388.
[11] Y. Fang and K. A. Loparo, *Stabilization of continuous-time jump linear systems*, IEEE Trans. Automat. Control, 47 (2002), pp. 1590–1603.
[12] M. K. Ghosh, A. Arapostathis, and S. I. Marcus, *Ergodic control of switching diffusions*, SIAM J. Control Optim., 35 (1997), pp. 1952–1988.
[13] D. Gröbaum and A. Okubo, *Modeling social animal aggregations*, in Frontiers in Theoretical Biology, Lecture Notes in Biomathematics 100, Springer-Verlag, New York, 1994, pp. 296–325.
[14] M. Huang, P. E. Caines, and R. P. Malhamé, *Individual and mass behaviour in large population stochastic wireless power control problems: Centralized and Nash equilibrium solutions*, in Proceedings of the 42nd IEEE Conference on Decision and Control, Maui, HI, 2003, pp. 98–103.
[15] M. Huang, P. E. Caines, and R. P. Malhamé, *Uplink power adjustment in wireless communication systems: A stochastic control analysis*, IEEE Trans. Automat. Control, 49 (2004), pp. 1693–1708.
[16] M. Huang, R. P. Malhamé, and P. E. Caines, *Nash certainty equivalence in large population stochastic dynamic games: Connections with the physics of interacting particle systems*, in Proceedings of the 45th IEEE Conference Decision and Control, San Diego, CA, 2006, pp. 4921–4926.
[17] M. Huang, R. P. Malhamé, and P. E. Caines, *Large population stochastic dynamic games: Closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle*, Commun. Inf. Syst., 6 (2006), pp. 221–251.
[18] M. Huang, P. E. Caines, and R. P. Malhamé, *Large-population cost coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ε-Nash equilibria*, IEEE Trans. Automat. Control, 52 (2007), pp. 1560–1571.
[19] M. Huang, *Large-population LQG games involving a major player: The Nash certainty equivalence principle*, SIAM J. Control Optim., 48 (2010), pp. 3318–3353.
[20] Y. Ji and H. J. Chizeck, *Controllability, stabilizability, and continuous-time Markovian jump linear quadratic control*, IEEE Trans. Automat. Control, 35 (1990), pp. 777–788.
[21] V. E. Lambson, *Self-enforcing collusion in large dynamic markets*, J. Econom. Theory, 34 (1984), pp. 282–291.
[22] J.-M. Lasry and P.-L. Lions. *Jeux à champ moyen. I C Le cas stationnaire*, C. R. Math. Acad. Sci. Paris, 343 (2006), pp. 619–625.
[23] J.-M. Lasry and P.-L. Lions. *Mean field games*, Japan. J. Math., 2 (2007), pp. 229–260.
[24] T. Li and J. F. Zhang, *Asymptotically optimal decentralized control for large population stochastic multi-agent systems*, IEEE Trans. Automat. Control, 53 (2008), pp. 1643–1660.
[25] T. Li and J. F. Zhang, *Decentralized tracking-type games for multi-agent systems with coupled ARX models: Asymptotic Nash equilibria*, Automatica J. IFAC, 44 (2008), pp. 713–725.
[26] R. S. Lipster and A. N. Shiryaev, *Statistics of Random Processes*, vol. 3, Springer-Verlag, New York, 1977, pp. 329–340.
[27] C. Q. Ma, T. Li, and J. F. Zhang, *Linear quadratic decentralized dynamic games for large-population discrete-time stochastic multi-agent systems*, J. Systems Sci. Math. Sci., 27 (2007), pp. 464–480.
[28] M. Mariton, *Jump Linear Systems in Automatic Control*, Marcel Dekker, New York, 1990.
[29] M. Nourian, P. E. Caines, R. P. Malhame, and M. Huang, *Mean field LQG control in leader-follower stochastic multi-agent systems: Likelihood ratio based adaptation*, IEEE Trans. Automat. Control, 57 (2012), to appear.
[30] B. Øksendal, *Stochastic Differential Equations. An Introduction with Applications*, 6th ed., Springer-Verlag, Berlin, 2003.
[31] C. Reynolds, *Flocks, herds and schools: A distributed behavior model*, Computer Graphics, 21 (1987), pp. 25–34.
[32] E. Semsar-Kazerooni and K. Khorasani, *Multi-agent team cooperation: A game theory approach*, Automatica J. IFAC, 45 (2009), pp. 2205–2213.
[33] D. D. Siljak, *Dynamic reliability using multiple control systems*, in Proceedings of the 2nd Lawerence Symposium on System Decision Sciences, Berkeley, CA, 1978.
[34] D. D. Sworder, *Feedback control of a class of linear systems with jump parameters*, IEEE Trans. Automat. Control, 14 (1969), pp. 9–14.
[35] W. M. Wonham, *Random differential equations in control theory*, Probabilistic Methods in Applied Mathematics 2, A. T. Bharucha-Reid, ed., Academic Press, New York, 1971.

[36] G. Y. Weintraub, C. L. Benkard, and B. Van Roy, *Oblivious equilibrium: A mean field approximation for large-scale dynamic games*, in Advances in Neural Information Processing Systems, MIT Press, Cambridge, MA, 2005.
[37] G. Y. Weintraub, C. L. Benkard, and B. Van Roy, *Markov perfect industry dynamics with many firms*, Econometrica, 76 (2008), pp. 1375–1411.
[38] Q. Zhang and J. F. Zhang, *Adaptive tracking control for coupled stochastic linear multi-agent systems: Stability, optimality and robustness*, IEEE Trans. Automat. Control, submitted.