

# Adaptive Mean Field Games for Large Population Coupled ARX Systems with Unknown Coupling Strength

Tao Li · Ji-Feng Zhang

Published online: 29 May 2013  
© Springer Science+Business Media New York 2013

**Abstract** This paper is concerned with decentralized tracking-type games for large population multi-agent systems with mean-field coupling. The individual dynamics are described by stochastic discrete-time auto-regressive models with exogenous inputs (ARX models), and coupled by terms of the unknown population state average (PSA) with unknown coupling strength. A two-level decentralized adaptive control law is designed. On the high level, the PSA is estimated based on the Nash certainty equivalence (NCE) principle. On the low level, the coupling strength is identified based on decentralized least squares algorithms and the estimate of the PSA. The decentralized control law is constructed by combining the NCE principle and Certainty equivalence (CE) principle. By probability limit theory, under mild conditions, it is shown that: (a) the closed-loop system is stable almost surely; (b) as the number of agents increases to infinity, the estimates of both the PSA and the coupling strength are asymptotically strongly consistent and the decentralized control law is an almost sure asymptotic Nash-equilibrium.

**Keywords** Mean field game · Decentralized game · Adaptive control · Adaptive game · ARX model · Nash certainty equivalence principle

## 1 Introduction

The research on multi-agent dynamic games has a long history in the control community. A good survey of non-cooperative dynamic games can be found in Basar and Olsder [4]. In recent years, the dynamic game theory gets new inspiration and renews its vitality in network control and multi-agent systems. In the framework of dynamic games, a lot of researchers considered flow control, routing control, and multi-agent cooperation problems [1–3, 5,

---

T. Li (✉) · J.-F. Zhang

Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, No. 55, Zhongguancundonglu, Beijing, 100190, China  
e-mail: [litao@amss.ac.cn](mailto:litao@amss.ac.cn)

J.-F. Zhang  
e-mail: [jjf@iss.ac.cn](mailto:jjf@iss.ac.cn)

14, 15, 17]. For distributed multi-agent systems, generally speaking, there is no centralized control station and each agent has only limited sensing and communication ability, so control design is always required to be decentralized. In a decentralized control framework, the control input of each agent can only use the local state, or under certain circumstances, include those of others in its sensing/communication neighborhood.

Recently, Huang, Caines, and Malhamé did pioneering work on decentralized stochastic games for a kind of individual-population interacting multi-agent systems with mean-field coupling [16, 18–22], which have wide application background in biological, engineering, and economic systems [6, 14, 29, 34]. In these kinds of systems, the number of agents is quite large. Each agent is driven by stochastic noises and interacts with all other agents via the population state average (PSA). The interactions between individual states and the PSA exist in both the dynamic equation and the cost function of every agent. For a given agent, the impact of any other single agent is so small that can be neglected, however, that of the overall population is significant enough for its evolution. Though the agents are coupled with the PSA, the PSA cannot be used for the individual control design, since it is unknown for any given agent. This is an essential difficulty of the decentralized control design for decentralized mean field games. To overcome this difficulty, Huang, Caines, and Malhamé proposed the methodology called the Nash Certainty Equivalence (NCE) principle. In the NCE principle, the PSA is properly approximated by its mean field approximation, a deterministic signal, which is then used for the individual control design instead of the PSA. This principle is similar in spirit to the well-known Certainty Equivalence (CE) principle adopted in adaptive control, where the unknown parameters are estimated, and the estimates are used as the true parameters to construct the control laws. The LQG mean field games with scalar agent models and deterministic discounted cost functions are studied in Huang, Malhamé, and Caines [16, 18, 20, 22]. Li and Zhang [26, 27] introduced the concepts of asymptotic Nash equilibria in the probability sense and extended to the cases with state space or ARX dynamic models and stochastic ergodic cost functions. The mean field method is also developed independently in Lasry and Lions [25] and in Weintraub and Benkard [31–33] by using the concept of oblivious equilibrium. The mean field control for Markov decision problems is considered in Tembine et al. [30]. Now, decentralized mean field games have been extended to non-linear dynamic models and the case with heterogeneous agents [13, 21, 23].

For decentralized mean field games, most of the relevant literature assumes precise dynamic models of agents. However, in real systems, there may be parametric uncertainties or unmodelled dynamics in agents' models due to various kinds of unknown or uncertain factors in the environment. Generally speaking, the parametric uncertainties can be divided into two categories: unknown local parameters, which contain the information of local environment, and unknown global parameters, which are shared by all agents. In this paper, we assume the local dynamics of each agent is precisely known, but the common coupling strength between the individual state and the PSA, which is a global parameter, is unknown.

To eliminate the model uncertainties, each agent can exploit its learning ability to perfect its dynamic model by measured data step by step. By using the individual online learner or identifier, each agent uses its estimate for the coupling strength to construct its individual control law, which aims at optimizing its cost function. Therefore, the overall system emerges as a large population decentralized adaptive game. In these kinds of adaptive games, there are two estimation processes. One is the estimation for the PSA and the other is the identification for the unknown coupling strength. A key difficulty lies in that there is a product term of the unknown PSA and the unknown coupling strength in each agent's dynamic equation. So, if traditional identification algorithms were used, then the regression vector

would contain the PSA as a component in each agent's identification algorithm. However, we know that the PSA is unavailable for each individual. Intuitively, the estimation signal for the PSA can be used to construct the identification algorithms instead of the PSA. Unfortunately, this may result in the coupling between the estimation process for the PSA and that for the unknown coupling strength. Decentralized adaptive games for individual-population interacting systems are considered firstly in Huang, Malhamé, and Caines [19] and Kizilkale and Caines [24]. In Kizilkale and Caines [24], the dynamic equations of agents are uncoupled and the local dynamic parameters are unknown, while in Huang, Malhamé, and Caines [19], the dynamics of agents are coupled, but the precise value of the PSA is used in the identification algorithm. In brief, the coupling between the two estimation processes, which is a key difficulty in decentralized mean field adaptive games does not exist in Huang, Malhamé, and Caines [19] and Kizilkale and Caines [24]. To our best knowledge, up to now there is no relevant literature, which involves the case where both the PSA and the coupling strength are unknown.

For decentralized adaptive mean field games, there are some fundamental problems that have to be studied.

- (1) Whether the closed-loop system is stable, that is, the states of all agents are kept bounded as time goes on. And if the answer is affirmative, whether the stability can be retained when the number  $N$  of agents increases to infinity.
- (2) Whether the estimate of the PSA is strongly consistent or whether the estimation error for the PSA converges to zero with respect to some metric almost surely as  $N$  tends to infinity? If the answer is affirmative, what is the convergence rate?
- (3) Whether the identification algorithm for the coupling strength is strongly consistent or the estimation errors are bounded? If the estimation errors are bounded, can we ensure that the bound converges to zero as  $N$  tends to infinity and get the convergence rate?
- (4) Whether the decentralized control law designed is asymptotically optimal almost surely, or is an almost sure asymptotic Nash-equilibrium? If the answer is affirmative, what is the convergence rate of the sub-optimal cost function of each agent to the optimal value, as  $N$  tends to infinity?

From above, it can be seen that the large population decentralized adaptive mean field game is essentially different from traditional adaptive control for single agent systems [7, 9, 11], and the solutions to the convergence problems (1)–(4) cannot be found in the existing theoretical framework. It is worth pointing out that this kind of adaptive control is also essentially different from decentralized adaptive games for large-scale systems [28, 35, 36]. In decentralized adaptive games for large-scale systems, there is no need for the regression vector of each agent's identification algorithm to contain any unknown states, and thus, there is no coupling between the identification for unknown parameters and the estimation for unknown states. This large population decentralized adaptive mean field game is closely related to robust adaptive control. The identification algorithm for the coupling strength, which involves the estimate for unknown states, can be viewed as the identification algorithm for a dynamic system with unmodelled dynamics. However, the assumptions on the unmodelled dynamics in robust adaptive control [8, 10] become closed-loop assumptions and are hard to be verified in the decentralized control framework.

In this paper, we consider the decentralized adaptive mean field game for individual-population interacting stochastic multi-agent systems. The dynamic equation of each agent is described by a discrete-time ARX model, and coupled by terms of the PSA with unknown coupling strength. Each agent has a group tracking type cost function, also coupled by the PSA. Firstly, based on the NCE principle, the PSA is estimated by some deterministic signal.

Secondly, the estimation of the PSA is used to construct the decentralized Least square (LS) identification algorithm for the coupling strength. Finally, the estimates of the PSA and the coupling strength are both used to construct the decentralized control law based on the NCE and the CE principles. By the stochastic Lyapunov method, we analyze the decentralized LS algorithm and then by the probability limit theory, under mild conditions, we get the following convergence results of the closed-loop system:

- (i) The closed-loop system is stable almost surely and the states of agents will retain bounded as  $N$  tends to infinity.
- (ii) As  $N$  tends to infinity, the estimation error for the PSA converges to zero with the rate  $O(1/N)$  almost surely.
- (iii) As  $N$  tends to infinity, the identification error for the unknown coupling strength converges to zero with the rate  $O(1/\sqrt{N})$ .
- (iv) The decentralized control law designed is an almost sure asymptotic Nash equilibrium, and the cost function of each agent is almost surely asymptotically optimal with the convergence rate  $O(1/N)$  given that all other agents also employ the strategy specified by the asymptotic Nash equilibrium.

The remainder of this paper is organized as follows. In Sect. 2, the system model and the decentralized game problem are formulated. In Sect. 3, a detailed design procedure of the two-level decentralized adaptive control law is presented, which is based on the NCE and the CE principles. In Sect. 4, the main results of this paper, regarding the stability of the closed-loop system, the asymptotic optimality of the control law, the asymptotic consistency of the estimations for the PSA, and the coupling strength are presented. In Sect. 5, a numerical example is used to demonstrate our theoretic results. In Sect. 6, some concluding remarks are given.

The following notation will be used throughout this paper.  $\mathbb{R}$  denotes the set of all real numbers. For a given random variable (r.v.)  $\xi$  on a probability space  $(\Omega, \mathcal{F}, P)$ ,  $E(\xi)$  denotes the mathematical expectation of  $\xi$ . For a family  $\{\xi_\lambda, \lambda \in \Lambda\}$  of real-valued r.v.s,  $\sigma(\xi_\lambda, \lambda \in \Lambda)$  denotes the  $\sigma$ -algebra  $\sigma(\{\xi_\lambda \in B\}, B \in \mathcal{B}, \lambda \in \Lambda)$ , where  $\mathcal{B}$  denotes the one-dimensional Borel sets. For a sequence  $\{\mathcal{F}_t, t \geq 0\}$  of non-decreasing  $\sigma$ -algebras and a sequence  $\{\xi(t), t \geq 0\}$  of r.v.s, we say  $\xi(t)$  is adapted to  $\mathcal{F}_t$  or  $\{\xi(t), \mathcal{F}_t\}$  is an adapted sequence, if for any  $t \geq 0$ ,  $\xi(t)$  is  $\mathcal{F}_t$  measurable.

## 2 Problem Formulation

We consider a system of  $N$  agents denoted by  $\mathbf{S}^N$ . The dynamic equation of agent  $i$  is given by

$$x_i^N(t + 1) = g_i(x_i^N(t), t) + u_i^N(t) + \alpha \bar{x}_N(t) + \omega_i(t + 1), \quad t = 0, 1, \dots, 1 \leq i \leq N, \quad (1)$$

where  $x_i^N \in \mathbb{R}$ ,  $u_i^N \in \mathbb{R}$  are the state and control input, respectively;  $\bar{x}_N(t) \triangleq \frac{1}{N} \sum_{j=1}^N x_j^N(t)$  is the PSA;  $\omega_i(t) \in \mathbb{R}$  is the random noise;  $g_i(\cdot, \cdot): \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  is a known Borel measurable function;  $\alpha \in \mathbb{R}$  is the unknown coupling parameter satisfying  $|\alpha| < 1$ . Note that model (1) is just the scalar version of the dynamic model considered in [26], and what is different here the coupling strength  $\alpha$  is unknown.

For model (1), we have the following assumptions:

(A1)  $\{\{\omega_i(t), \mathcal{F}_t^i\}, 1 \leq i \leq N, N \geq 1\}$  is a family of independent martingale difference sequences defined on a probability space  $(\Omega, \mathcal{F}, P)$  with the following properties: there

exist constants  $\sigma > 0$  and  $\beta > 2$ , such that

$$\sup_{t \geq 0} E[|\omega_i(t)|^\beta | \mathcal{F}_{t-1}^i] < \infty \quad \text{a.s.}$$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T [\omega_i(t)]^2 = \sigma^2 \quad \text{a.s.}$$

where  $\mathcal{F}_t^i \triangleq \sigma(\omega_i(s), 0 \leq s \leq t)$ .

(A2)  $\{x_i^N(0), 1 \leq i \leq N, N \geq 1\}$  is independent of  $\{\{\omega_i(t), \mathcal{F}_t^i\}, i \geq 1\}$ , with a common mathematical expectation  $x_0 \triangleq E x_i^N(0) < \infty$ .

The cost function of agent  $i$  is given by

$$J_i^N(u_i^N, u_{-i}^N) = \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T [x_i^N(t+1) - \Phi(t, \bar{x}_N(t))]^2, \tag{2}$$

where  $u_{-i}^N = (u_1^N, \dots, u_{i-1}^N, u_{i+1}^N, \dots, u_N^N)$ ,  $\Phi(t, x) : [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$  is a Borel measurable function.

With regards to the cost function, the following assumptions will be involved in the closed-loop analysis:

(A3) The solution of the non-linear iteration  $x(t+1) = \Phi(t, x(t))$  with  $x(0) = x_0$  satisfies

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T x^2(t) < \infty.$$

(A4) The solution of the non-linear iteration  $x(t+1) = \Phi(t, x(t))$  with  $x(0) = x_0$  satisfies

$$\lim_{T \rightarrow \infty} \frac{x^2(T)}{1 + \sum_{t=0}^{T-1} x^2(t)} = 0.$$

It can be easily verified that if  $\Phi(t, x) = x, t \in [0, \infty), x \in \mathbb{R}$ , then both (A3) and (A4) hold.

For convenience of citation, for agent  $i$ , we denote the global-measurement-based admissible control set by

$$\mathcal{U}_{g,i}^N \triangleq \left\{ u_i^N \mid u_i^N(t) \text{ is adapted to } \sigma \left( \bigcup_{j=1}^N \sigma(x_j^N(s), 0 \leq s \leq t) \right) \right\},$$

the local-measurement-based admissible control set by

$$\mathcal{U}_{l,i}^N \triangleq \{ u_i^N \mid u_i^N(t) \text{ is adapted to } \sigma(x_i^N(s), 0 \leq s \leq t) \},$$

and admissible control set by  $\mathcal{U}_i^N$ . The so-called decentralized game means that agent  $i$  synthesizes  $u_i^N$  only based on the local measurement (i.e.  $\mathcal{U}_i^N = \mathcal{U}_{l,i}^N$ ) to minimize its cost function  $J_i^N(u_i^N, u_{-i}^N)$ . We denote a control group of the sequence  $\mathbf{S}^N$  of systems by  $\mathbf{U}^N = \{u_i^N, 1 \leq i \leq N\}$ , and its associated cost function group by  $\mathbf{J}^N = \{J_i^N(u_i^N, u_{-i}^N), 1 \leq i \leq N\}$ .

To characterize the asymptotic optimality of the decentralized control law with respect to the stochastic cost functions, we introduce the concept of almost sure asymptotic Nash-equilibrium given in Li and Zhang [26].

**Definition 1** For system (1), a sequence of control groups  $\{\mathbf{U}^N = \{u_i^N, 1 \leq i \leq N\}, N \geq 1\}$  is called an almost sure asymptotic Nash-equilibrium with respect to the associated sequence of cost function groups  $\{\mathbf{J}^N = \{J_i^N, 1 \leq i \leq N\}, N \geq 1\}$ , if there exists a sequence of non-negative r.v.s  $\{\epsilon_N(\omega), N \geq 1\}$  on the probability space  $(\Omega, \mathcal{F}, P)$ , such that  $\epsilon_N \rightarrow 0$  a.s., as  $N \rightarrow \infty$ , and for sufficiently large  $N$ ,

$$J_i^N(u_i^N, u_{-i}^N) \leq \inf_{v_i \in \mathcal{U}_{g,i}^N} J_i^N(v_i, u_{-i}^N) + \epsilon_N, \quad \text{a.s., } i = 1, 2, \dots, N. \tag{3}$$

By Theorem 2.1 of [26], we know that  $\inf_{v_i \in \mathcal{U}_{g,i}^N} J_i^N(v_i, u_{-i}^N) = \sigma^2$ . In the following, we will design decentralized control law  $\{\mathbf{U}^N, N \geq 1\}$ , such that the closed-loop system satisfies

$$J_i^N(u_i^N, u_{-i}^N) \leq \sigma^2 + o(1), \quad N \rightarrow \infty, \text{ a.s.,}$$

that is, the sequence of control groups  $\{\mathbf{U}^N, N \geq 1\}$  is an almost sure asymptotic Nash-equilibrium.

### 3 Control Design

Firstly, we make a review of the results with known coupling strength.

#### 3.1 The Case with Known Coupling Strength

For the centralized control law design, the control of agent  $i$  depends on the PSA  $\bar{x}_N$ , while for the design of the decentralized control law, the PSA is unknown. If the coupling strength  $\alpha$  is known, we may use the NCE principle to design the decentralized control law. Firstly, we construct an estimate  $f(t)$  of the PSA with the following property: if every agent takes  $f(t)$  as the estimate of the PSA, and according to  $f(t)$ , makes the optimal decision, then the expectation of the closed-loop PSA is just  $f(t)$  or convergent to it when  $N$  increases to infinity. Secondly, if the  $f(t)$  with the above property indeed exists, then we can construct the decentralized control law by using  $f(t)$  instead of  $\bar{x}_N(t)$ .

Based on the NCE principle, we now design the decentralized control law.

The auxiliary equation of agent  $i$  is given by

$$\hat{x}_i^N(t + 1) = g_i(\hat{x}_i^N(t), t) + \hat{u}_i^N(t) + \alpha f(t) + \omega_i(t + 1), \quad t \geq 0, i = 1, 2, \dots, N, \tag{4}$$

with a tracking-type cost function:

$$J_i^N(\hat{u}_i^N) = \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T [\hat{x}_i^N(t + 1) - \Phi(t, f(t))]^2.$$

In this case, the optimal control is obviously

$$\hat{u}_i^N(t) = \Phi(t, f(t)) - g_i(\hat{x}_i^N(t), t) - \alpha f(t). \tag{5}$$

Substituting control (5) into the model (4), we have

$$E\widehat{x}_i^N(t + 1) = \Phi(t, f(t)), \quad E\widehat{x}_i^N(0) = x_0. \tag{6}$$

As mentioned above, the mathematical expectation of the closed-loop PSA ought to be  $f(t)$ , that is,

$$\frac{1}{N} \sum_{j=1}^N E\widehat{x}_j^N(t) = f(t), \quad t \geq 0. \tag{7}$$

Therefore, the unique solution of the auxiliary system (6) and (7) can be used as the estimate of the PSA. We denote it by  $f^*(t)$ , which is iteratively given by

$$f^*(t + 1) = \Phi(t, f^*(t)), \quad t \geq 0, \quad f^*(0) = x_0. \tag{8}$$

By (5) and the NCE principle, the control law for agent  $i$  can be taken as

$$u_i^0(t) = \Phi(t, f^*(t)) - g_i(x_i^N(t), t) - \alpha f^*(t). \tag{9}$$

Here and hereafter, the superscript  $N$  of  $u_i^{0N}(t)$  is omitted for conciseness of expression. Comparing (9) with the centralized control law, it can be seen that  $\bar{x}_N$  in (9) is replaced by  $f^*$  for control design.

As shown in [26], we can prove the asymptotic consistency of the estimate  $f^*$  for the PSA, the stability and asymptotic optimality of the closed-loop system under the control law (9). We have the following theorems [26]).

**Proposition 1** *For system (1), if Assumptions (A1)–(A2) hold, then under the control law (9), the closed-loop system has the following properties:*

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T [\xi_N(t)]^2 = \frac{\sigma^2}{N(1 - \alpha^2)} \quad a.s., \tag{10}$$

where

$$\xi_N(t) = \bar{x}_N(t) - f^*(t) \tag{11}$$

is the estimation error for the PSA.

**Proposition 2** *For system (1), if Assumptions (A1)–(A3) hold, then under the control law (9), the closed-loop system satisfies*

$$\sup_{N \geq 1} \max_{1 \leq i \leq N} \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T [x_i^N(t)]^2 < \infty, \quad a.s.$$

**Proposition 3** *For system (1) with cost function (2), if Assumptions (A1)–(A2) hold and there exists a constant  $\gamma > 0$  such that  $|\Phi(t, x) - \Phi(t, y)| \leq \gamma|x - y|, \forall x, y \in \mathbb{R}, t \geq 0$ , then under the control law (9), the associated cost function group satisfies*

$$J_i^N(u_i^0, u_{-i}^0) \leq \sigma^2 + \epsilon_N \quad a.s., \quad i = 1, 2, \dots, N, \tag{12}$$

where

$$\epsilon_N = \frac{2\sigma^2(\alpha^2 + \gamma^2)}{N(1 - \alpha^2)}.$$

### 3.2 The Case with Unknown Coupling Strength

If the coupling strength  $\alpha$  is unknown, then the control law (9) is unavailable. Naturally, one might think that based on the model (1), agent  $i$  could use the recursive LS algorithm to estimate  $\alpha$ :

$$\bar{\alpha}_i(t + 1) = \bar{\alpha}_i(t) + (1 + P(t)\bar{x}_N^2(t))^{-1} \{P(t)\bar{x}_N(t)[x_i^N(t + 1) - u_i^N(t) - g_i(x_i^N(t), t) - \bar{\alpha}_i(t)\bar{x}_N(t)]\}, \tag{13}$$

$$P(t + 1) = P(t) - \frac{P^2(t)\bar{x}_N^2(t)}{1 + \bar{x}_N^2(t)P(t)}, \tag{14}$$

then by CE principle, instead of  $\alpha$ , the estimation  $\bar{\alpha}_i(t)$  could be used to construct the control law

$$u_i^N(t) = \Phi(t, f^*(t)) - g_i(x_i^N(t), t) - \bar{\alpha}_i(t)f^*(t), \quad i = 1, 2, \dots, N. \tag{15}$$

However, the control law (13)–(15) is not decentralized due to the use of the PSA in the identification algorithm. Since the PSA  $\bar{x}_N(t)$  is unknown for agent  $i$ , we use  $f^*(t)$ , which is the estimation of  $\bar{x}_N(t)$  based on the NCE principle, to construct the identification algorithm of agent  $i$ :

$$\alpha_i(t + 1) = \alpha_i(t) + (1 + P(t)f^{*2}(t))^{-1} \{P(t)f^*(t)[x_i^N(t + 1) - u_i^N(t) - g_i(x_i^N(t), t) - \alpha_i(t)f^*(t)]\}, \quad \alpha_i(0) = \alpha_0, \tag{16}$$

$$P(t + 1) = P(t) - \frac{P^2(t)f^{*2}(t)}{1 + f^{*2}(t)P(t)}, \quad P(0) = P_0, \tag{17}$$

where  $\alpha_0$  and  $P(0) = P_0$  are initial conditions to be designed, and  $f^*(t)$  is computed off-line by (8). The identification algorithm (16) and (17) is decentralized, since it only uses the local state and input of each agent. Then by CE principle, we use the estimate  $\alpha_i(t)$  to construct the control law:

$$u_i^N(t) = \Phi(t, f^*(t)) - g_i(x_i^N(t), t) - \alpha_i(t)f^*(t), \quad i = 1, 2, \dots, N. \tag{18}$$

It can be seen that the control law (16)–(18) is decentralized and it is designed based on both the NCE and the CE principles.

*Remark 1* Here, a decentralized two-level control scheme is used for adaptive mean field adaptive games. On the high level, the PSA is estimated based on the NCE principle. On the low level, the coupling strength is identified based on the decentralized LS algorithms and the estimate of the PSA. The decentralized control law is constructed by combining the NCE and the CE principles.



### 4 Closed-Loop Analysis

In this section, we analyze the identification algorithm, stability and optimality of the closed-loop system and the consistency of the estimates for the PSA and the coupling strength.

#### 4.1 Identification Algorithm Analysis

From the model (1) and (16), one gets

$$\tilde{\alpha}_i(t + 1) = \left(1 - \frac{P(t)f^{*2}(t)}{1 + P(t)f^{*2}(t)}\right)\tilde{\alpha}_i(t) - \frac{P(t)f^*(t)(\alpha\xi_N(t) + \omega_i(t + 1))}{1 + P(t)f^{*2}(t)}, \tag{19}$$

where  $\tilde{\alpha}_i(t) \triangleq \alpha - \alpha_i(t)$  is the estimation error for the coupling strength  $\alpha$ , and  $\xi_N(t)$  is the estimation error for the PSA given by (11).

Denote  $V_i(t + 1) = \tilde{\alpha}_i^2(t + 1)P^{-1}(t + 1)$ ,  $r(T) = e + \sum_{t=0}^T f^{*2}(t)$ . From (17), we know that

$$P^{-1}(t + 1) = P^{-1}(t) + f^{*2}(t). \tag{20}$$

Then summing the above equation from both sides, we have

$$P^{-1}(t + 1) = P_0^{-1} + \sum_{k=0}^t f^{*2}(k), \tag{21}$$

From (19) and (20), we have

$$\begin{aligned} V_i(t + 1) &= V_i(t) - \frac{\tilde{\alpha}_i^2(t)f^{*2}(t)}{1 + P(t)f^{*2}(t)} - 2\frac{\tilde{\alpha}_i(t)f^*(t)(\alpha\xi_N(t) + \omega_i(t + 1))}{1 + P(t)f^{*2}(t)} \\ &\quad + \frac{P(t)f^{*2}(t)(\alpha\xi_N(t) + \omega_i(t + 1))^2}{1 + P(t)f^{*2}(t)}. \end{aligned} \tag{22}$$

For the identification algorithm (16) and (17), we have the following results, which are important for the closed-loop analysis of the decentralized control law.

**Theorem 1** *If Assumption (A1) holds and*

$$\limsup_{T \rightarrow \infty} \frac{f^{*2}(T)}{P_0^{-1} + \sum_{t=0}^{T-1} f^{*2}(t)} < C, \quad \text{for some } C > 0, \tag{23}$$

*then the identification algorithm (16) and (17) have the following properties:*

(i)

$$\begin{aligned} V_i(T + 1) + \sum_{t=0}^T \frac{(\tilde{\alpha}_i(t)f^*(t) + \alpha\xi_N(t))^2}{1 + P(t)f^{*2}(t)} &\leq \alpha^2 \frac{2C + 1}{C + 1} \sum_{t=0}^T \xi_N^2(t) + o\left(\sum_{t=0}^T \xi_N^2(t)\right) \\ &\quad + o\left(\sum_{t=0}^T \frac{(\tilde{\alpha}_i(t)f^*(t) + \alpha\xi_N(t))^2}{1 + P(t)f^{*2}(t)}\right) \end{aligned}$$

$$+ O(\ln r(T)), \quad a.s. \tag{24}$$

(ii)

$$\begin{aligned} \tilde{\alpha}_i^2(n+1) &\leq \alpha^2 \frac{2C+1}{C+1} \frac{\sum_{t=0}^T \xi_N^2(t)}{r(T)} + o\left(\frac{\sum_{t=0}^T \xi_N^2(t)}{r(T)}\right) \\ &\quad + O\left(\frac{\ln r(T)}{r(T)}\right), \quad a.s., \quad i = 1, 2, \dots, N, \end{aligned} \tag{25}$$

(iii)

$$\begin{aligned} &\sum_{t=0}^T \frac{(\tilde{\alpha}_i(t)f^*(t) + \alpha\xi_N(t))^2}{1 + P(t)f^{*2}(t)} \\ &\leq \frac{\alpha^2(2C+1)}{(1-\delta)(C+1)} \sum_{t=0}^T \xi_N^2(t) \\ &\quad + o\left(\sum_{t=0}^T \xi_N^2(t)\right) + O(\ln r(T)), \quad a.s. \quad \forall \delta \in (0, 1), \quad i = 1, 2, \dots, N. \end{aligned} \tag{26}$$

*Proof* From (23), we have

$$\frac{P(t)f^{*2}(t)}{1 + P(t)f^{*2}(t)} \leq \frac{C}{1+C}, \quad \forall t \geq t_1, \text{ for some } t_1 > 0. \tag{27}$$

Then summing both sides of (22) from  $t = t_1$  to  $t = T$ , we get

$$\begin{aligned} V_i(T+1) &+ \sum_{t=t_1}^T \frac{\tilde{\alpha}_i^2(t)f^{*2}(t)}{1 + P(t)f^{*2}(t)} + 2 \sum_{t=t_1}^T \frac{\tilde{\alpha}_i(t)f^*(t)\alpha\xi_N(t)}{1 + P(t)f^{*2}(t)} + \sum_{t=t_1}^T \frac{\alpha^2\xi_N^2(t)}{1 + P(t)f^{*2}(t)} \\ &= V_i(t_1) + \sum_{t=t_1}^T \frac{\alpha^2\xi_N^2(t)}{1 + P(t)f^{*2}(t)} - 2 \sum_{t=t_1}^T \frac{(\tilde{\alpha}_i(t)f^*(t) + \alpha\xi_N(t))\omega_i(t+1)}{1 + P(t)f^{*2}(t)} \\ &\quad + 2 \sum_{t=t_1}^T \frac{\alpha\xi_N(t)\omega_i(t+1)}{1 + P(t)f^{*2}(t)} + 2 \sum_{t=t_1}^T \frac{P(t)f^{*2}(t)(\alpha\xi_N(t) + \omega_i(t+1))^2}{1 + P(t)f^{*2}(t)}. \end{aligned} \tag{28}$$

From the above equation, (23) and Lemma A.1, we have

$$\begin{aligned} V_i(T+1) &+ \sum_{t=t_1}^T \frac{(\tilde{\alpha}_i(t)f^*(t) + \alpha\xi_N(t))^2}{1 + P(t)f^{*2}(t)} \\ &\leq V_i(t_1) + \alpha^2 \frac{2C+1}{C+1} \sum_{t=0}^T \xi_N^2(t) + O\left(\left(\sum_{t=t_1}^T \frac{(\tilde{\alpha}_i(t)f^*(t) + \alpha\xi_N(t))^2}{1 + P(t)f^{*2}(t)}\right)^{1/2+\epsilon}\right) \\ &\quad + O\left(\left(\sum_{t=0}^T \xi_N^2(t)\right)^{1/2+\epsilon}\right) + 2 \sum_{t=0}^T \frac{P(t)f^{*2}(t)\omega_i^2(t+1)}{1 + P(t)f^{*2}(t)}. \end{aligned} \tag{29}$$

From (21), we have

$$\sum_{t=0}^T \frac{P(t)f^{*2}(t)}{1 + P(t)f^{*2}(t)} = \sum_{t=0}^T \frac{f^{*2}(t)}{P_0^{-1} + \sum_{k=0}^t f^{*2}(k)}, \tag{30}$$

and then

$$\begin{aligned} \sum_{t=0}^T \frac{P(t)f^{*2}(t)}{1 + P(t)f^{*2}(t)} &= \sum_{t=0}^T \frac{P^{-1}(t+1) - P^{-1}(t)}{P^{-1}(t+1)} \\ &\leq \sum_{t=0}^T \int_{P^{-1}(t)}^{P^{-1}(t+1)} \frac{dx}{x} \\ &= \ln P^{-1}(T+1) + \ln P_0 = O(\ln r(T)). \end{aligned} \tag{31}$$

Denote  $\mathcal{F}_t = \sigma(\cup_{j=1}^N \mathcal{F}_t^j)$ . For any given  $\nu \in (2, \min\{\beta, 4\}]$ , by Cr inequality, we have

$$\begin{aligned} \sup_{t \geq 0} E[|\omega_t^2(t+1) - E(\omega_t^2(t+1)|\mathcal{F}_t)|^{\nu/2}|\mathcal{F}_t] &\leq \sup_{t \geq 0} E[|\omega_t(t+1)|^\nu|\mathcal{F}_t] \\ &\quad + \sup_{t \geq 0} E[|E(\omega_t^2(t+1)|\mathcal{F}_t)|^{\nu/2}|\mathcal{F}_t], \end{aligned}$$

which together with Assumption (A1) and Lyapunov inequality leads to

$$\begin{aligned} \sup_{t \geq 0} E[|\omega_t^2(t+1) - E(\omega_t^2(t+1)|\mathcal{F}_t)|^{\nu/2}|\mathcal{F}_t] &\leq \sup_{t \geq 0} E[|\omega_t(t+1)|^\nu|\mathcal{F}_t] + \sup_{t \geq 0} |E(\omega_t^2(t+1)|\mathcal{F}_t)|^{\nu/2} \\ &\leq 2 \sup_{t \geq 0} E[|\omega_t(t+1)|^\nu|\mathcal{F}_t] \\ &\leq 2 \left( \sup_{t \geq 0} E[|\omega_t(t+1)|^\beta|\mathcal{F}_t^i] \right)^{\nu/\beta} < \infty, \quad \text{a.s.} \end{aligned} \tag{32}$$

Then by Lemma A.1, (30) and (31), for any given  $\epsilon > 0$ , noting that  $0 \leq \frac{f^{*2}(t)}{P_0^{-1} + \sum_{k=0}^t f^{*2}(k)} \leq 1$ , we have

$$\begin{aligned} \sum_{t=0}^T \frac{P(t)f^{*2}(t)\omega_t^2(t+1)}{1 + P(t)f^{*2}(t)} &= \sum_{t=0}^T \frac{P(t)f^{*2}(t)}{1 + P(t)f^{*2}(t)} (\omega_t^2(t+1) - E(\omega_t^2(t+1)|\mathcal{F}_t)) \\ &\quad + \sum_{t=0}^T \frac{P(t)f^{*2}(t)}{1 + P(t)f^{*2}(t)} E(\omega_t^2(t+1)|\mathcal{F}_t) \\ &\leq \sup_{t \geq 0} E(\omega_t^2(t+1)|\mathcal{F}_t^i) \sum_{t=0}^T \frac{P(t)f^{*2}(t)}{1 + P(t)f^{*2}(t)} \\ &\quad + O\left(\left(\sum_{t=0}^T \left(\frac{P(t)f^{*2}(t)}{1 + P(t)f^{*2}(t)}\right)^{\nu/2}\right)^{2/\nu+\epsilon}\right) \end{aligned}$$

$$= O\left(\sum_{t=0}^T \frac{f^{*2}(t)}{P_0^{-1} + \sum_{k=0}^t f^{*2}(k)}\right) + O(1) = O(\ln r(T)). \tag{33}$$

which together with (29) leads to (i) and (iii). Combining (i) and (iii), we get (ii). □

*Remark 2* By (11), the model (1) can be rewritten as

$$\widehat{x}_i^N(t + 1) = g_i(\widehat{x}_i^N(t), t) + \widehat{u}_i^N(t) + \alpha f^*(t) + \alpha \xi_N(t) + \omega_i(t + 1), \quad t \geq 0, i = 1, 2, \dots, N.$$

So, (16), (17), and (18) can be viewed as the identification algorithm and adaptive control law for model

$$\widehat{x}_i^N(t + 1) = g_i(\widehat{x}_i^N(t), t) + \widehat{u}_i^N(t) + \alpha f^*(t) + \omega_i(t + 1), \quad t \geq 0, i = 1, 2, \dots, N,$$

with  $\alpha \xi_N(t)$  as the unmodelled dynamics. It can be seen that  $\alpha \xi_N(t)$  contains the states of all other agents, due to decentralized information pattern, the conditions on unmodelled dynamics used in robust adaptive control [8, 10] cannot be used here.

### 4.2 Stability, Optimality, and Consistency

Substitute the control (18) to the model (1), we get the closed-loop equation of agent  $i$

$$x_i^N(t + 1) = \widetilde{\alpha}_i(t) f^*(t) + \alpha \xi_N(t) + \Phi(t, f^*(t)) + \omega_i(t + 1). \tag{34}$$

Summate the above equation for  $i = 1, 2, \dots, N$ , then by (8) we know that  $\xi_N(t)$  satisfies the following recursive equation:

$$\xi_N(t + 1) = \alpha \xi_N(t) + f^*(t) \frac{1}{N} \sum_{j=1}^N \widetilde{\alpha}_j(t) + \frac{1}{N} \sum_{j=1}^N \omega_j(t + 1). \tag{35}$$

From (19) and (35), it can be seen that the dynamic equation (35) of the estimation error and the dynamic equation (19) of the identification error are coupled together. Below is the main result of this paper.

**Theorem 2** *If Assumptions (A1)–(A4) hold, then for the system (1), under the control (8), (16), (17), and (18), we have*

(i) *The estimate for PSA is asymptotically consistent:*

$$\limsup_{T \rightarrow \infty} \|\xi_N\|_T^2 = O(1/N), \quad a.s., \tag{36}$$

where  $\|\xi_N\|_T = \sqrt{\frac{1}{T} \sum_{t=0}^T \xi_N^2(t)}$ .

(ii) *The closed-loop system is almost surely uniformly stable:*

$$\sup_{N \geq 1} \max_{1 \leq i \leq N} \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T [x_i^N(t)]^2 < \infty \quad a.s. \tag{37}$$

(iii) *Furthermore, if there exists  $\gamma > 0$ , such that for any  $x, y \in \mathbb{R}$  and  $t \geq 0$ , we have  $|\Phi(t, x) - \Phi(t, y)| \leq \gamma|x - y|$ , then  $\{U^N = \{u_i(t), 1 \leq i \leq N\}, N \geq 1\}$  is an almost sure*

asymptotic Nash equilibrium with respect to the associated sequence of cost function groups, and the cost function of each agent is almost surely asymptotically optimal with the convergence rate  $O(N^{-1})$  given that all other agents also employ the strategy specified by the asymptotic Nash equilibrium:

$$\max_{1 \leq i \leq N} J_i^N(u_i^N, u_{-i}^N) \leq \sigma^2 + \frac{2\sigma^2(\gamma^2 + \alpha^2)}{N(1 - \alpha^2)}, \quad \text{a.s.} \tag{38}$$

*Proof* Take positive real numbers  $\epsilon \in (0, \frac{1-\alpha^2}{2\alpha^2})$  and  $\delta \in (0, 1 - \alpha^2(2\epsilon + 1))$ . From Assumption (A4), we know that

$$\limsup_{T \rightarrow \infty} \frac{f^{*2}(T)}{P_0^{-1} + \sum_{t=0}^{T-1} f^{*2}(t)} < \epsilon, \tag{39}$$

and similar to (27), we have

$$\frac{P(t)f^{*2}(t)}{1 + P(t)f^{*2}(t)} \leq \frac{\epsilon}{1 + \epsilon}, \quad \forall t \geq t_\epsilon, t_\epsilon > 0. \tag{40}$$

From (8) and Assumption (A3), we know that  $r(T) = O(T)$ ,  $n \rightarrow \infty$ . Then by (39), (40), and (iii) of Theorem 1, we have

$$\begin{aligned} \sum_{t=0}^T (\tilde{\alpha}_i(t)f^*(t) + \alpha\xi_N(t))^2 &\leq \frac{\alpha^2(2\epsilon + 1)}{1 - \delta} \sum_{t=0}^T \xi_N^2(t) \\ &\quad + o\left(\sum_{t=0}^T \xi_N^2(t)\right) + o(T), \quad i = 1, 2, \dots, N, \end{aligned} \tag{41}$$

which together with (35) leads to

$$\begin{aligned} \sum_{t=0}^T \xi_N^2(t + 1) &\leq \frac{\alpha^2(2\epsilon + 1)}{1 - \delta} \sum_{t=0}^T \xi_N^2(t) + o\left(\sum_{t=0}^T \xi_N^2(t)\right) \\ &\quad + o(n) + \sum_{t=0}^T \left(\frac{1}{N} \sum_{j=1}^N \omega_j(t + 1)\right)^2. \end{aligned} \tag{42}$$

From the above and Assumption (A1), we get

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T \xi_N^2(t) \leq \frac{\sigma^2}{N(1 - \mu(\alpha, \epsilon, \delta))}, \quad \text{a.s.}, \tag{43}$$

where  $\mu(\alpha, \epsilon, \delta) \triangleq \frac{\alpha^2(2\epsilon+1)}{1-\delta}$ . This together with (41) leads to

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T (\tilde{\alpha}_i(t)f^*(t) + \alpha\xi_N(t))^2 \leq \frac{\sigma^2\mu(\alpha, \delta)}{N(1 - \mu(\alpha, \epsilon, \delta))}, \quad \text{a.s.} \tag{44}$$

Furthermore, by (34), Assumption (A3) and Lemma A.1, we have (ii).

From (34), (2), and Assumption (A1), it follows that

$$\begin{aligned}
 J_i^N(u_i^N, u_{-i}^N) &= \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T [\tilde{\alpha}_i(t) f^*(t) + \alpha \xi_N(t) + \Phi(t, f^*(t)) - \Phi(t, \bar{x}_N(t)) \\
 &\quad + \omega_i(t + 1)]^2 = I_1^N + I_2^N + \sigma^2,
 \end{aligned}
 \tag{45}$$

where

$$I_1^N = 2 \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T [\tilde{\alpha}_i(t) f^*(t) + \alpha \xi_N(t) + \Phi(t, f^*(t)) - \Phi(t, \bar{x}_N(t))] \omega_i(t + 1),$$

and

$$\begin{aligned}
 I_2^N &= \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T [\tilde{\alpha}_i(t) f^*(t) + \alpha \xi_N(t) + \Phi(t, f^*(t)) - \Phi(t, \bar{x}_N(t))]^2 \\
 &\leq 2 \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T [\tilde{\alpha}_i(t) f^*(t) + \alpha \xi_N(t)]^2 + 2\gamma^2 \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T \xi_N^2(t),
 \end{aligned}$$

which together with (45), (43), (44), and Lemma A.1 leads to

$$\max_{1 \leq i \leq N} J_i^N(u_i^N, u_{-i}^N) \leq \sigma^2 + \frac{2\sigma^2(\gamma^2 + \mu(\alpha, \epsilon, \delta))}{N(1 - \mu(\alpha, \epsilon, \delta))}, \quad \text{a.s.}
 \tag{46}$$

Letting  $\epsilon$  and  $\delta$  go to zero in (46) and (43), we get (iii), and

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T \xi_N^2(t) \leq \frac{\sigma^2}{N(1 - \alpha^2)}, \quad \text{a.s.},
 \tag{47}$$

which gives (i). □

*Remark 3* Compared (36) and (38) with (10) and (12), it is shown that for the case with unknown coupling strength, under the adaptive control law designed, the convergence rates of the estimation error for PSA and the cost function of each agent to the best response value are the same as those for the case with known coupling strength.

*Remark 4* From Theorem 2, we can see that to ensure the control law to be an asymptotic Nash equilibrium, the consistency of the identification for the coupling strength  $\alpha$  is not necessary. This is similar to the case of LS based adaptive tracker (Guo and Chen [12]).

In the following theorem, under certain excitation condition on the non-linear iteration, we get the asymptotic consistency of the identification algorithm, that is, the upper limit of the identification error vanishes as the number  $N$  of agents increases to infinity. We need the following assumption.

(A5) The solution of the non-linear iteration  $x(t + 1) = \Phi(t, x(t))$  with  $x(0) = x_0$  satisfies

$$\liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T x^2(t) > 0.$$

**Theorem 3** *If Assumptions (A1)–(A5) hold, then for the system (1), under the control (8), (16), (17), and (18), the closed-loop system satisfies*

$$\limsup_{t \rightarrow \infty} \tilde{\alpha}_i^2(t) \leq \frac{\alpha^2 \sigma^2}{N(1 - \alpha^2) \underline{f}}, \quad i = 1, 2, \dots, N, \text{ a.s.,}$$

where  $\underline{f} = \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T (f^*(t))^2 > 0$ .

*Proof* By (ii) of Theorem 1 and Assumption (A4), we have

$$\begin{aligned} \tilde{\alpha}_i^2(T + 1) &\leq \alpha^2 \frac{2\epsilon + 1}{\epsilon + 1} \frac{\sum_{t=0}^T \xi_N^2(t)}{r(T)} + o\left(\frac{\sum_{t=0}^T \xi_N^2(t)}{r(T)}\right) + O\left(\frac{\ln r(T)}{r(T)}\right), \quad \text{a.s.,} \\ i &= 1, 2, \dots, N, \quad \forall \epsilon > 0. \end{aligned} \tag{48}$$

From Assumption (A5), we know that there exists  $c_0 > 0$ , such that  $r(T) \geq c_0 T$  for sufficiently large  $T$ , which together with (48), (i) of Theorem 2, and Assumption (A5) leads to

$$\limsup_{t \rightarrow \infty} \tilde{\alpha}_i^2(t) \leq \frac{\alpha^2 \sigma^2}{N(1 - \alpha^2) \underline{f}} \frac{2\epsilon + 1}{\epsilon + 1}, \quad \text{a.s., } i = 1, 2, \dots, N.$$

Letting  $\epsilon$  go to zero, we get the conclusion of this theorem. □

### 5 Numerical Example

In this section, a numerical example is given to demonstrate stability of the closed-loop system and that the decentralized control law is designed to be an asymptotic Nash equilibrium.

The dynamic equation for the  $i$ th agent is given by

$$x_i^N(t + 1) = x_i^N(t) + u_i^N(t) + 0.5\bar{x}_N(t) + \omega_i(t + 1),$$

where the initial value  $x_i^N(0)$  has the normal distribution  $N(5, 1)$ ,  $\{\omega_i(t), t \geq 0\}$  is a sequence of Gaussian white noise with distribution  $N(0, 2)$ . So  $\alpha = 0.5$ ,  $\sigma = 2$ . The non-linear coupling function in the cost function is  $\Phi(t, x) = 5 \sin(x) + 6$ . The initial parameter estimate is taken as  $a_0 = 2$  and  $P_0 = 10$ .

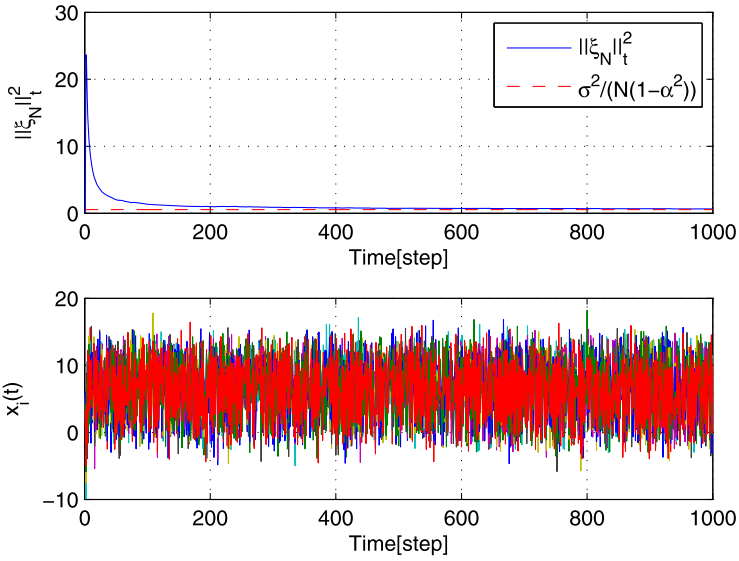
From (18), the decentralized control law is taken as

$$u_i^N(t) = 5 \sin(f^*(t)) + 6 - x_i^N(t) - 0.5 f^*(t), \quad i = 1, 2, \dots, N, \tag{49}$$

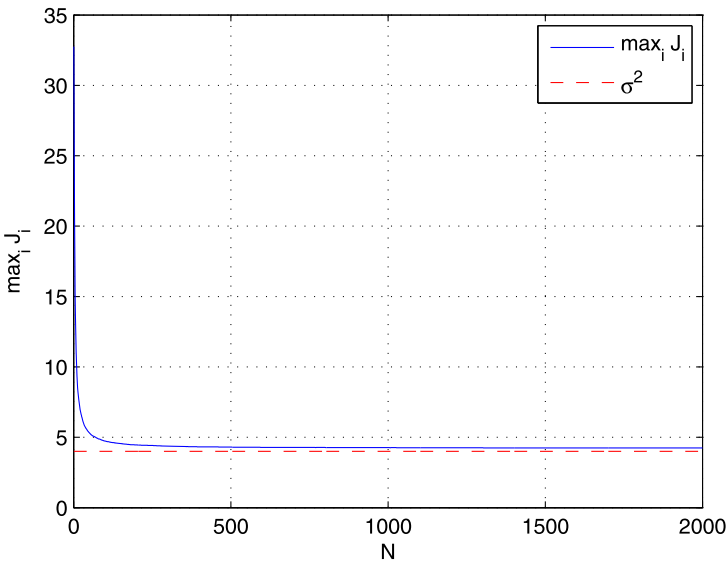
where  $f^*(t)$  is iteratively given by

$$f^*(t + 1) = 5 \sin(f^*(t)) + 6, \quad t \geq 0, \quad f^*(0) = 5,$$

with (16) and (17) as the decentralized LS identifiers. Let  $N = 10$ . The evolution of the estimation error of PSA, the states of agents are shown in Fig. 1. It can be seen that  $\|\xi_{10}\|_t^2$  converges to  $\frac{\sigma^2}{N(1 - \alpha^2)} = 8/15$ , as  $t \rightarrow \infty$  and the closed-loop system is stable. We let  $N$  vary from 1 to 2000. The evolution of the maximum values of the cost functions with respect to  $N$  is shown in Fig. 2. It can be seen that the maximum values of the cost functions converges to the optimal value  $\sigma^2 = 25$ , as  $N \rightarrow \infty$ .



**Fig. 1** Trajectories of  $\|\xi_{10}\|_t^2$  and  $x_i(t)$ ,  $i = 1, 2, \dots, 10$  with respect to  $t$



**Fig. 2** Trajectory of  $\max_{1 \leq i \leq N} J_i^N$  with respect to  $N$

### 6 Conclusions

A decentralized adaptive tracking-type game has been considered in this paper for individual-population interacting systems, in which each agent interacts with the overall population via the PSA in the individual dynamics and cost function. The coupling strength with the PSA is unknown. A two-level adaptive control law is designed based on the NCE



and the CE principles. Firstly, the PSA is estimated based on the NCE principle, then the estimation of the PSA is used to construct the decentralized LS identification algorithm for the coupling strength; finally, the estimates for the coupling strength and the PSA are used to construct the decentralized control law based on the NCE and the CE principles. It is shown that under mild conditions, in probability one, the closed-loop system is stable, the decentralized control law is an asymptotic Nash equilibrium, and the estimates are asymptotically consistent as the number of agents goes to infinity.

Here, as a preliminary research for this direction, we give a framework for this problem and consider the case with scalar dynamic models of agents. For the LS estimation, we know that the estimation error  $\|\tilde{\theta}(t)\|^2 = O\left(\frac{\ln(1+\sum_{i=0}^t \|\phi_i\|^2)}{\lambda_{\min}(\sum_{i=0}^{t+1} \phi_i \phi_i^T)}\right)$ . If the regression vector  $\phi_i$  is a scalar variable, then the LS algorithm is easily consistent, however, for the case with high dimensional regression vectors, the consistency requires additional excitation conditions [7]. So for adaptive control systems, there are essential differences between the scalar system models and high dimensional models. For future research, the extension to general linear models may be interesting. Another important issue is to consider the case with both unknown local and global parameters, which may be more widely applicable, but much more difficult.

**Acknowledgements** The research of Tao Li and Ji-Feng Zhang was supported by the National Natural Science Foundation of China under Grants 60934006 and 61004029.

### Appendix

**Lemma A.1** (Chen and Guo [7]) *Let  $\{X(t), \mathcal{F}_t\}$  be a matrix martingale difference sequence,  $\{M(t), \mathcal{F}_t\}$  an adapted sequence of random matrices,  $\|M(t)\| < \infty, \forall t \geq 0$ . If*

$$\sup_{t \geq 0} E[\|X(t)\|^\alpha | \mathcal{F}_{t-1}] < \infty \quad a.s.,$$

for some  $\alpha \in (0, 2]$ , then as  $T \rightarrow \infty$ ,

$$\sum_{t=0}^T M(t)X(t+1) = O(s_T(\alpha) \ln^{1/\alpha+\eta}(s_T^\alpha(\alpha) + e)) \quad a.s., \quad \forall \eta > 0,$$

where

$$s_T(\alpha) = \left( \sum_{t=0}^T \|M(t)\|^\alpha \right)^{1/\alpha}.$$

### References

1. Altman E, Basar T (1998) Multi-user rate-based flow control. *IEEE Trans Commun* 46(7):940–949
2. Altman E, Wynter L (2004) Equilibrium, games and pricing in transportation and telecommunication networks. *Networks and Spacial Issue on Crossovers Between Transportation and Telecommunication Modelling* 4(1):7–21
3. Altman E, Basar T, Srikant R (2002) Nash equilibria for combined flow control and routing in networks: asymptotic behavior for a large number of users. *IEEE Trans Autom Control* 47(6):917–930
4. Basar T, Olsder GJ (1982) *Dynamic noncooperative game theory*. Academic Press, London
5. Bauso D, Giarré L, Pesenti R (2006) Non-linear protocols for optimal distributed consensus in networks of dynamics agents. *Syst Control Lett* 55(11):918–928

6. Breban R, Vardavas R, Blower S (2007) Mean-field analysis of an inductive reasoning game: application to influenza vaccination. *Phys Rev E* 76:031127
7. Chen HF, Guo L (1991) Identification and stochastic adaptive control. Birkhäuser, Boston
8. Chen HF, Guo L (1991) A robust adaptive controller. *IEEE Trans Autom Control* 33(11):1035–1043
9. Duncan TE, Guo L, Pasik-Duncan B (1999) Adaptive continuous-time linear quadratic Gaussian control. *IEEE Trans Autom Control* 44(9):1653–1662
10. Guo L (1993) Time-varying stochastic systems. Ji Lin Science and Technology Press, Ji Lin
11. Guo L (1996) Self-convergence of weighted least-squares with applications to stochastic adaptive control. *IEEE Trans Autom Control* 41(1):79–89
12. Guo L, Chen HF (1991) The Åström-Wittenmark self-tuning regulator revised and ELS-based adaptive trackers. *IEEE Trans Autom Control* 36(7):802–812
13. Huang M (2010) Large-population LQG games involving a major player: the Nash certainty equivalence principle. *SIAM J Control Optim* 48(5):3318–3353
14. Huang M, Caines PE, Malhamé RP (2003) Individual and mass behaviour in large population stochastic wireless power control problem: centralized and Nash equilibrium solutions. In: Proceedings of the 42nd conference on decision and control, Maui, Hawaii, December 9–12, pp 98–103
15. Huang M, Caines PE, Malhamé RP (2004) Uplink power adjustment in wireless communication systems: a stochastic control analysis. *IEEE Trans Autom Control* 49(10):1693–1708
16. Huang M, Caines PE, Malhamé RP (2004) Large-population cost-coupled LQG problems: generalizations to non-uniform individuals. In: Proceedings of the 43rd conference on decision and control, Nassau, Bahamas, December 14–17, pp 3453–3458
17. Huang M, Malhamé RP, Caines PE (2004) On a class of large-scale cost coupled Markov games with applications to decentralized power control. In: Proceedings of the 43rd conference on decision and control, Nassau, Bahamas, December 14–17, pp 2830–2835
18. Huang M, Malhamé RP, Caines PE (2005) Nash equilibria for large-population linear stochastic systems of weakly coupled agents. In: Boukas EK, Malhamé RP (eds) Analysis, control and optimization of complex dynamic systems. Springer, New York, pp 215–252. Chap. 9
19. Huang M, Malhamé RP, Caines PE (2005) Nash strategies and adaptation for decentralized games involving weakly-coupled agents. In: Proceedings of the 44th IEEE conference on decision and control and the European control conference 2005, Seville, Spain, December 12–15, pp 1050–1055
20. Huang M, Malhamé RP, Caines PE (2006) Nash certainty equivalence in large population stochastic dynamic games: connections with the physics of interacting particle systems. In: Proceedings of the 45th IEEE conference on decision and control, San Diego, CA, USA, December 13–15, pp 4921–4926
21. Huang M, Malhamé RP, Caines PE (2006) Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun Inf Syst* 6(3):221–251
22. Huang M, Caines PE, Malhamé RP (2007) Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized  $\epsilon$ -Nash equilibria. *IEEE Trans Autom Control* 52(9):1560–1571
23. Huang M, Caines PE, Malhamé RP (2007) An invariance principle in large population stochastic dynamic games. *J Syst Sci Complex* 20(2):162–172
24. Kizilkale AC, Caines PE (2013) Mean field stochastic adaptive control. *IEEE Trans Autom Control* 58(4):905–920
25. Lasry JM, Lions PL (2007) Mean field games. *Jpn J Math* 2(1):229–260
26. Li T, Zhang JF (2008) Decentralized tracking-type games for multi-agent systems with coupled ARX models: asymptotic Nash equilibria. *Automatica* 44(3):713–725
27. Li T, Zhang JF (2008) Asymptotically optimal decentralized control for large population stochastic multi-agent systems. *IEEE Trans Autom Control* 53(7):1643–1660
28. Ma HB (2009) Decentralized adaptive synchronization of a stochastic discrete-time multiagent dynamic model. *SIAM J Control Optim* 48(2):859–880
29. McNamara JM, Houston AI, Collins EJ (2001) Optimality models in behavioral biology. *SIAM Rev* 43(3):413–466
30. Tembine H, Boudec JYL, El-Azouzi R, Altman E (2009) Mean field asymptotics of Markov decision evolutionary games and teams. In: Proceedings of international conference on game theory for networks (GameNets 2009), Istanbul, Turkey, May 13–15, pp 140–150
31. Weintraub GY, Benkard CL, Roy BV (2005) Oblivious equilibrium: a mean field approximation for large-scale dynamic games. *Advances in neural information processing systems*. MIT Press, Cambridge
32. Weintraub GY, Benkard CL, Roy BV (2007) Industry dynamics: from elemental to aggregate models. *Relation* 10(1.130):4948
33. Weintraub GY, Benkard CL, Roy BV (2008) Markov perfect industry dynamics with many firms. *Econometrica* 76(6):1375–1411

34. Yin H, Mehta PG, Meyn SP, Shanbhag UV (2010) Learning in mean-field oscillator games. In: Proceeding of the 49th IEEE conference on decision and control, Atlanta, GA, USA, December 15-17, pp 3125–3132
35. Zhang Q, Zhang JF (2010) Adaptive tracking-type games for coupled large population ARMAX systems. In: Proceedings of the 8th IEEE international conference on control and automation, Xiamen, China, June 9–11, pp 148–153
36. Zhang Q, Zhang JF (2010) Robust adaptive control of coupled stochastic multi-agent systems with unmodeled dynamics. In: Proceedings of the 8th world congress on intelligent control and automation, Jinan, China, July 6–9, pp 586–591