# Adaptive Tracking Games for Coupled Stochastic Linear Multi-Agent Systems: Stability, Optimality and Robustness

Qiang Zhang and Ji-Feng Zhang, *Senior Member, IEEE*

***Abstract*—Distributed adaptive tracking-type games are investigated for a class of coupled stochastic linear multi-agent systems with uncertainties of unknown structure parameters, external stochastic disturbances, unmodeled dynamics, and unknown agents' interactions. The control goal is to make the states of all the agents converge to a desired function of the population state average (PSA). Due to the fact that only local information is available for each agent, the control is distributed. For the time-invariant parameter case, the extended least-squares algorithm, Nash certainty equivalence (NCE) principle, and certainty equivalence (CE) principle are used to estimate the unknown parameters and the PSA term, and to design adaptive control, respectively. Under some mild conditions, it is shown that the closed-loop system is almost surely uniformly stable with respect to the population number $N$; the estimate for the PSA term is strongly consistent; the adaptive control is almost surely an asymptotic Nash equilibrium. When the dynamics of each agent contains time-varying parameters and unmodeled dynamics, the projected least mean square (LMS) algorithm, NCE principle, and CE principle are adopted to estimate the unknown time-varying parameters, and the unknown PSA term, and to design robust adaptive control, respectively. In addition to stability of the closed-loop system and consistency of the PSA estimate, the control law is shown to be robust Nash equilibrium with respect to the unmodeled dynamics, the variation of the unknown parameters, and the external disturbances. Two numerical examples are given to illustrate the methods and results of this paper.**

***Index Terms*—Adaptive control, distributed control, multi-agent systems, Nash equilibrium, robust adaptive control, stochastic dynamic game.**

## I. Introduction

### A. Motivation

DURING the last half century, adaptive control theory has attracted lots of attentions and successfully been used to many applications [3]–[6]. Adaptive schemes are of significance due to their effectiveness in dealing with various uncertainties, for instance, unknown structure parameters, external disturbances, unmodeled dynamics etc. When carrying out tasks such as cooperative control, distributed control, distributed estimation etc., a multi-agent system (MAS) is very susceptible to these uncertainties. Thus, it is also of great importance to study the adaptive control design for general MASs, and make them possess adaption capability to uncertainties.

Compared with the conventional case, the prominent characteristic of adaptive control of MASs is distributed rather than centralized. This is similar to the decentralized adaptive control of large-scale interconnected systems, on which many research results can be found, including adaptive stabilization, regulation and tracking control [7]–[14]. The key feature is that the communication information among the subsystems is not fully used to improve the control performance. In [7]–[11], stability conditions on system structure were investigated for various systems with different kind of disturbances and interconnections among subsystems, and decentralized adaptive stabilization controls were designed. In [12], [13], for the case where the global information, such as the target state/output of all the subsystems, is known, strictly decentralized asymptotic adaptive tracking controls were given for a class of large-scale interconnected systems. In [14], the issue on how to use the communication information among subsystems to enhance control performances was considered, and decentralized adaptive control was designed for a class of quite general large-scale systems with communication information among subsystems available.

Different from large-scale system, for MASs, on one hand, each agent has its own control goal, and certain capability of information collecting, processing, and communicating. When designing control, in addition to their own measurement information (control, state, output, etc.), their neighbors' information can be used. On the other hand, each agent has only limited capability on information collecting, processing, and communicating, and cannot get the global information of the whole system. Thus, their controllers have to be based on only the local information, and thus, are distributed. This brings us a basic problem, i.e., how to use only the local information to constructively design such a distributed controller that has the capability of both dealing with uncertainties (e.g., unknown parameters, unmodeled dynamics, external disturbances) and driving the whole MAS to a desired global optimal performance [15]–[27]. One research topic along this line is the consensus problem of MAS where connectivity properties are important for the information flow of the system [15]–[17]. For example, [28] uses a variable structure approach to solve both the distributed consensus tracking and swarm tracking problems under connectivity conditions. There are also other works to establish the

Manuscript received March 02, 2010; revised July 16, 2010, January 19, 2011, July 04, 2012, and January 07, 2013; accepted June 14, 2013. Date of publication June 21, 2013; date of current version October 21, 2013. This paper was presented in part at the 8th IEEE International Conference on Control and Automation and the 8th World Congress on Intelligent Control and Automation [1], [2]. This work was supported by the National Natural Science Foundation of China under Grants 61120106011 and 60934006. Recommended by Associate Editor M. Prandini.

Q. Zhang is with the International Cooperation Center for Economics and Technology, Ministry of Industry and Information Technology, Beijing 100846, China (e-mail: qzhang@amss.ac.cn).

J.-F. Zhang is with the Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China (e-mail: jif@iss.ac.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2013.2270869

global behavior without explicitly resorting to *a priori* connectivity conditions, by using other information or constraints [18], [19], [24]–[27]. Under the noncooperative individual-mass game framework, [24]–[27] considered the optimal LQG and tracking games for a class of large population stochastic multi-agent systems (LPSMASs), where the particular coupling structure of the individual dynamics and costs, the overall rationality hypotheses for the population, and the large population number are the key issues in the problem setup. In all these works, the system parameters are required to be known and deterministic, and the dynamic model of each agent does not contain any unmodeled dynamics.

Adaptive control of MASs, which can simultaneously cope with uncertainties such as unknown parameters, unmodeled dynamics, external disturbances, etc., and which make the whole system achieve a desired global control goal, has not been given sufficient attention. Only recently are some results on this topic gradually coming up [29]–[34]. Taking the LQG game of a simple first-order continuous-time LPSMAS as an example, [29] gave some intuitive suggestions on how to conduct parameter estimation and adaptive control design, and how to improve the transient estimation while avoid incurring over-energied control. This can be regarded as a preliminary work on adaptive control of LPSMASs, although lacking of concrete result and rigorous analysis. Reference [30] presented a comprehensive survey on the Nash certainty equivalence (NCE) based control of LPSMASs, including the adaptive LQG control of high-dimensional linear LPSMASs. This can actually be seen as a complement to the results in [29]. One basic characteristic of the above results on LPSMASs is that the population number is taken as a variable, and the asymptotic macro-properties of the whole system as the agent number increases to infinity are concerned. When the agent number is fixed, some nice results on the adaptive control of MASs can be found in [31]–[34]. Reference [31] considered the distributed adaptive tracking control for a class of coupled ARMAX MASs, and the signal to be tracked and some parameters (e.g., the high-frequency gain) of each agent are required to be known and deterministic. Reference [32] considered the robust adaptive consensus problem of MASs. They used the adaptive neural network scheme to approximate the uncertain dynamics, and a robustness signal to compensate for the approximation error and bounded external disturbances. Reference [33] considered the adaptive formation control problem by transforming it into nonlinear servomechanism problem, and designed an adaptive internal model based controller. Reference [34] considered the formation control and target tracking problems in a class of multi-agent systems with nonlinear and uncertain dynamics, and use a direct adaptive fuzzy control methodology to design control. However, both of the above three works did not provide rigorous theoretical analysis for the performance of the closed-loop system. In summary, when the dynamic model of each agent contains unknown structural parameters, unmodeled dynamics and external disturbances, and the performance index contains unknown coupled signal, the distributed adaptive control for general MASs is of great significance from the point view of both theory and practice, and worth thoroughly and deeply investigating.

### *B. Main Contributions*

Different from the conventional case [3]–[6], the adaptive control design for MASs needs to consider the interactions among agents, coupling terms in dynamics and performance indices, and the capability on information collection of each agent and so on. However, similar to the single-agent system case, the framework of the adaptive control of MASs should also include a parameter estimator to estimate the unknown parameters online and a distributed control law to achieve the given control tasks. Adaptive control can provide better capability in dealing with uncertainties, but may simultaneously introduce multiplicative nonlinearities into the closed-loop systems. This together with the interactions among agents makes the performance analysis of the closed-loop adaptive MASs much more difficult.

This paper is about to consider the adaptive optimal tracking task for a class of coupled stochastic linear MASs under the noncooperative individual-mass game framework [24]–[27], [29], [30]. For the case where the system uncertainties contain unknown constant parameters, stochastic disturbances and unknown coupled interactions, we will design a distributed adaptive policy to drive the agent system to a desired mass behavior via the individual-mass interactions, and prove the stability of the closed-loop system and the optimality of the designed control. If in addition, the uncertainties include unmodeled dynamics and slowly time-varying parameters, we design robust adaptive tracking control, and prove the stability of the closed-loop system and the robust optimality of the control. The interaction among agents lies in the following two aspects. One is the coupled term with its neighbors in the dynamics of each agent, including modeled part and unmodeled part; and the other is the unknown PSA term in the performance indices. The optimality of the performance indices is described by concepts of asymptotic Nash equilibrium [25], [27] and robust Nash equilibrium, respectively. Compared with the previous index-coupled individual-mass interplay system [24]–[27], [29], [30], in this paper we extend the results to weakly coupled systems with unknown parameters. Specifically, the main contributions of this paper can be summarized as follows.

1) The distributed adaptive tracking-type games and distributed robust adaptive tracking games are respectively obtained for stochastic linear MASs.
2) For the adaptive tracking games of stochastic linear time-invariant MASs, NCE principle and extended least-squares (ELS) algorithm are used to estimate the PSA term and the unknown constant parameters, respectively. A distributed adaptive tracking control is designed, under which the closed-loop system is almost surely uniformly stable with respect to (w.r.t.) the population number $N$; the estimate of PSA is strongly consistent; the adaptive control designed is almost surely asymptotic Nash equilibrium, and the performance indices converge to the optimal values with some convergence rate.
3) For the robust adaptive tracking games of stochastic linear MASs with unknown time-varying parameters, the NCE principle, projected least mean square (LMS) algorithm and CE principle are used to estimate the PSA term, the unknown time-varying parameters, and design the adaptive control. It is shown that the closed-loop system is stable, the estimate for the PSA term is consistent, and the adaptive control are robust Nash equilibrium w.r.t. the unmodeled dynamics, the parameter variation and the external disturbances.

### C. Organization of the Paper and Notations

The remainder of this paper is organized as follows. In Section II, we present the model and formulate the problem to be investigated. In Section III, we discuss the adaptive tracking games of MASs with unknown constant parameters. In Section IV, we consider a more general case with unmodeled dynamics and unknown time-varying parameters. In Section V, we use two numerical examples to illustrate the methods and results. In Section VI, we give some concluding remarks and further research topics.

The following notations will be used throughout this paper. $\mathbb{R}^m$ denotes the real $m$-dimensional space. For a given vector or matrix $X$, $X^T$ denotes its transpose, $tr(X)$ denotes its trace when $X$ is square, $\|X\|$ denotes the Euclid norm of vector $X$, or the $F$-norm of matrix $X$. For a given random variable $\xi$, $E\xi$ denotes the expectation of $\xi$, $E(\xi|\mathcal{F})$ denotes the conditional expectation of $\xi$ w.r.t. the $\sigma$-algebra $\mathcal{F}$. For a given sequence of nondecreasing $\sigma$-algebra $\{\mathcal{F}_t, t \geq 0\}$, and a sequence of random variables $\xi(t), t \geq 0$, we call $\{\xi(t), \mathcal{F}_t\}$ an adapted sequence, if for any $t \geq 0$, $\xi(t)$ is $\mathcal{F}_t$-measurable.

## II. Problem Statement

Within this and the next section we will consider the distributed optimal tracking-type games of $N$ agents, subjected to uncertainties of unknown constant parameters, stochastic disturbances and unknown coupled terms. The dynamics of agent $i$ ($i = 1, \ldots, N$) is described by the following coupled ARMAX model:

$$\begin{aligned} x_i(t+1) = & -\sum_{k=1}^{p_i} A_{ik} x_i(t-k+1) + \sum_{k=1}^{q_i} B_{ik} u_i(t-k+1) \\ & + \sum_{k=1}^{n_i} \sum_{j \in \mathcal{N}_i} G_{ijk} x_j(t-k+1) + w_i(t+1) \\ & + \sum_{k=1}^{r_i} C_{ik} w_i(t-k+1), \quad t \geq 0 \end{aligned} \tag{1}$$

where $x_i(t) \in \mathbb{R}^m$, $u_i(t) \in \mathbb{R}^m$ and $w_i(t) \in \mathbb{R}^m$ denote the state, control input and stochastic disturbance of agent $i$ at time $t$, $x_i(t) = 0$, $u_i(t) = 0$, $w_i(t) = 0$ for $t < 0$; $\{A_{ik}, 1 \leq k \leq p_i\}$, $\{B_{ik}, 1 \leq k \leq q_i\}$, $\{G_{ijk}, j \in \mathcal{N}_i, 1 \leq k \leq n_i\}$ are unknown parameter matrices with proper dimensions; $\mathcal{N}_i = \{n_{i1}, \ldots, n_{i,m_i}\}$ denotes neighbors of agent $i$; $\sum_{k=1}^{n_i} \sum_{j \in \mathcal{N}_i} G_{ijk} x_j(t-k+1)$ denotes the interaction term of agent $i$ with its neighbors, the coupling coefficients $G_{ijk}$ ($k = 1, \ldots, n_i$) describe the influence intensity of agent $j$ to agent $i$.

Here, we assume the network topology of the MAS is time invariant, i.e., $\mathcal{N}_1, \ldots, \mathcal{N}_N$ do not change with time. Moreover, we assume that the $i$th agent can measure the states of itself and its neighbors $j \in \mathcal{N}_i$ exactly.

For the above coupled stochastic MASs, we investigate the problem of distributed optimal tracking games via coupled performance indices or costs, which are described by

$$J_i^N(u_i, u_{-i}^N) = \limsup_{t \to \infty} \frac{1}{t} \sum_{k=0}^{t} \|x_i(k+1) - \Phi(\overline{x}_N(k))\|^2 \tag{2}$$

where $u_{-i}^N \triangleq (u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_N)$; $\Phi(\cdot) : \mathbb{R}^m \to \mathbb{R}^m$ is a Borel measurable function; $\overline{x}_N(k) \triangleq (1/N)\sum_{i=1}^N x_i(k)$ denotes the population state average (PSA) at time $k$.

System model (1) is the ARMAX model often used in the adaptive control field of single-agent system [3]–[6], the linear interaction model with its neighbors is common in the decentralized adaptive control of large-scale interconnected systems [12]–[14]. The model (1) with cost (2) has a wide background in biological, economic and engineering systems, for example, the particle swarm systems of [35], [36] and the resource allocation problem in wireless network [37]. In the particle swarm systems, the acceleration $v_i$ of agent $i$ is a linear dynamic equation of its own position $r_i$, velocity $v_i$, Gaussian white noise $w_i$, and the coupled interacting term $(1/N)\sum_{j=1}^N r_j$. In the resource allocation problem, each user, with the linear power adjustment dynamics, makes its own strategy $u_i$ ensure an adequate signal-to-interference ratio (SIR) $\Gamma_i = p_i/(\sum_{j=1}^N p_j + \nu)$, where $p_i$ is the power level of user $i$, $\nu$ is the constant system background noise intensity. Other motivating examples such as the production output adjustment problem in dynamical economic market can be found in [24] and the numerical example in Section V. In these examples, the mathematical expression of the PSA index coupling is determined by and varies with the specific group behavioral objectives of the whole agent system. For example, in the power allocation problem, the Borel measurable function $\Phi(x) = \gamma(x + \nu)$, where $\gamma$ is the target level of SIR. In the production output adjustment problem, $\Phi(\cdot)$ is approximately proportional to the current product price, which can be modeled as a nonlinear function of the overall production level $\sum_{j=1}^N x_j$ [38].

*Remark 1:* Different from the system model used in the decentralized adaptive control of large-scale interconnected systems, here we consider the individual-mass interacting stochastic performance indexes (2). Compared with the existing work [31], the problem discussed here has the following essential differences. First, $\overline{x}_N(k)$ in the performance index is unknown, and depends on the states of all the other agents. Second, the high-frequency gains $B_{i1}$ ($i = 1, \ldots, N$) are unknown, and need to be estimated.

For convenience of comparison, just as in [25], [27], we define the global-measurement-based admissible control set for agent $i$ by

$$\mathcal{U}_{g,i}^N = \left\{u | u(t) \in \sigma\left(\cup_{j=1}^N \sigma(x_j(s), 0 \leq s \leq t)\right)\right\}$$

and local-measurement-based admissible control set by

$$\mathcal{U}_{l,i}^N = \left\{u | u(t) \in \sigma\left(x_i(s), x_j(s), j \in \mathcal{N}_i, 0 \leq s \leq t\right)\right\}.$$

In order to quantitatively analyze the minimization of the stochastic performance indices $\{J_i^N(\cdot), 1 \leq i \leq N\}$, we introduce the definition of almost surely (a.s.) asymptotic Nash equilibrium.

*Definition 2.1 [27]:* For the system (1)–(2), a group of control sequences $\{U_N = \{u_i \in \mathcal{U}_i^N, 1 \leq i \leq N\}, N \geq 1\}$ is called a.s. asymptotic Nash equilibrium w.r.t. the corresponding index group $\{J_N = \{J_i^N, 1 \leq i \leq N\}, N \geq 1\}$, if there exists a sequence of non-negative random variables (r.v.s) $\{\epsilon_N, N = 1, 2, \ldots\}$ on a probability space $(\Omega, \mathcal{F}, P)$, such that $\epsilon_N \xrightarrow[N \to \infty]{} 0$ a.s., and $J_i^N(u_i, u_{-i}^N) \leq$

$\inf_{v_i \in \mathcal{U}_{g,i}^N} J_i^N(v_i, u_{-i}^N) + \epsilon_N$ a.s., $i = 1, 2, \ldots, N$, for sufficiently large $N$.

In the above definition, $\mathcal{U}_i^N$ denotes the admissible control set of agent $i$. Obviously, in distributed control situation, we have $\mathcal{U}_i^N = \mathcal{U}_{l,i}^N$, while in centralized control situation, $\mathcal{U}_i^N = \mathcal{U}_{g,i}^N$.

The purpose of this paper is to design a distributed adaptive tracking control, using states of itself and its neighbors for the system (1) with performance index (2), i.e., to find an adaptive control group in $\mathcal{U}_{l,i}^N$ to cope with the uncertainties, and analyze stability of the closed-loop system, and a.s. asymptotic Nash equilibrium of the designed control group w.r.t. the corresponding performance index group.

## III. Adaptive Tracking Control Design and Performance Analysis

In this section, we will design a distributed adaptive control for the MASs described by (1) and (2). To be specific, the parameter estimation algorithm is presented in Section III-A; the developed adaptive control is discussed in Section III-B; and properties of the closed-loop systems are analyzed in Section III-C.

### A. Estimation Algorithm

To proceed the design of adaptive tracking control and ensure good performance of the closed-loop system, we need the following assumptions on the system (1).

*A1):* $\{\{w_i(t), \mathcal{F}_t^i\}, i \geq 1\}$ is a group of independent martingale deference sequences on a probability space $(\Omega, \mathcal{F}, P)$, satisfying the following conditions:

$$\sup_{t\geq 0} E\left[\|w_i(t+1)\|^\beta \,|\mathcal{F}_t^i\right] < \infty \text{ a.s.,} \quad \text{for some } \beta > 2 \quad (3)$$

$$\lim_{t\to\infty} \frac{1}{t} \sum_{k=0}^{t} w_i(k) w_i^T(k) = R_\sigma \quad \text{a.s.} \quad (4)$$

where $\mathcal{F}_t^i \triangleq \sigma(w_i(s), 0 \leq s \leq t)$, $R_\sigma \in \mathbb{R}^{m\times m}$ is a positive definite matrix.

*A2):* $C_i^{-1}(e^{i\lambda}) + C_i^{-T}(e^{-i\lambda}) - I > 0, \forall \lambda \in [0, 2\pi]$, where $C_i(z) = I + C_{i1}z + C_{i2}z^2 + \cdots + C_{i,r_i}z^{r_i}$.

*A3):* $\det B_i(z) \neq 0$, $\forall z \in \mathcal{C} : |z| \leq 1$, where $B_i(z) = B_{i1} + B_{i2}z + B_{i3}z^2 + \cdots + B_{i,q_i}z^{q_i-1}$.

*A4):* $\{x_i(0), i \geq 1\}$ is an independent stochastic sequence with a common mathematical expectation $x_0 = Ex_i(0)$, and is independent of $\{\{w_i(t), t \geq 0\}, i \geq 1\}$.

*Remark 2:* Assumptions A1)–A3) are standard in adaptive control theory [4], [39]. Let the global $\sigma$-algebra at time $t$ ($t \geq 0$) be given by $\overline{\mathcal{F}}_t^N = \sigma(\mathcal{F}_0^N \cup (\cup_{i=1}^N \mathcal{F}_t^i))$, where $\mathcal{F}_0^N = \sigma(x_i(0), 1 \leq i \leq N)$. Then it is easy to verify that $\{\{w_i(t), \overline{\mathcal{F}}_t^N\}, 1 \leq i \leq N, N \geq 1\}$ satisfies A1), i.e., $\{w_i(t), \overline{\mathcal{F}}_t^N\}$ is a martingale difference sequence, satisfying (4) and $\sup_{t\geq 0} E\left[\|w_i(t+1)\|^\beta \,|\overline{\mathcal{F}}_t^N\right] < \infty$ a.s. Therefore, if $\{u_i(t), \overline{\mathcal{F}}_t^N\}$ is an adapted sequence, then the linear minimum variance estimation for $x_i(t+1)$ generated by (1) coincides with the minimum variance estimation $E[x_i(t+1)|\overline{\mathcal{F}}_t^N]$. A2) is the usual strictly positive condition. A3) is the minimum phase condition, which is necessary for adaptive tracking controls. A4) is a condition on initial states.

More generally, one can consider the case where the expectations of initial states $Ex_i(0)$ are different for each agent, but some statistical assumptions on $x_i(0)$ (for example, having a known distribution function with some properties) may be required as in [24]. The following lemma is key to the analysis of the ELS algorithm and the proofs of the theorems in this paper.

*Lemma 3.1 [4]:* Let $\{W_t, \mathcal{F}_t\}$ be a matrix martingale difference sequence, $\{M_t, \mathcal{F}_t\}$ an adapted random matrix sequence, satisfying $\|M_t\| < \infty$ a.s., $\forall t \geq 0$. If for some constant $\alpha \in (0, 2]$, $\sup_{t\geq 0} E[\|W_{t+1}\|^\alpha |\mathcal{F}_t] < \infty$ a.s., then $\sum_{k=0}^{t} M_k W_{k+1} = O\left(s_t(\alpha) \log^{1/\alpha+\eta}(s_t^\alpha(\alpha) + e)\right)$ a.s., $\forall \eta > 0$, where $s_t(\alpha) = \left(\sum_{k=0}^{t} \|M_k\|^\alpha\right)^{1/\alpha}$.

For convenience of expression, we rewrite (1) in the following form:

$$x_i(t+1) = \theta_i^T \varphi_i^0(t) + w_i(t+1), \quad t \geq 0$$

where

$$\begin{aligned} \theta_i = \big[&-A_{i1}, \ldots, -A_{i,p_i}, B_{i1}, \ldots, B_{i,q_i}, \\ &C_{i1}, \ldots, C_{i,r_i}, G_{i,n_{i1},1}, \ldots, G_{i,n_{i1},n_i}, \\ &\ldots, G_{i,n_{i,m_i},1}, \ldots, G_{i,n_{i,m_i},n_i}\big]^T \end{aligned} \quad (5)$$

$$\begin{aligned} \varphi_i^0(t) = \big[&x_i^T(t), \ldots, x_i^T(t-p_i+1), u_i^T(t), \\ &\ldots, u_i^T(t-q_i+1), w_i^T(t), \ldots, w_i^T(t-r_i+1), \\ &x_{n_{i1}}^T(t), \ldots, x_{n_{i1}}^T(t-n_i+1), \ldots, x_{n_{i,m_i}}^T(t), \\ &\ldots, x_{n_{i,m_i}}^T(t-n_i+1)\big]^T. \end{aligned} \quad (6)$$

We apply the ELS algorithm to estimate the unknown parameter matrix $\theta_i$:

$$\begin{aligned} \hat{\theta}_i(t+1) &= \hat{\theta}_i(t) + a_i(t)P_i(t)\varphi_i(t) \\ &\quad \cdot (x_i(t+1) - \hat{\theta}_i^T(t)\varphi_i(t))^T \quad (7) \\ P_i(t+1) &= P_i(t) - a_i(t)P_i(t)\varphi_i(t)\varphi_i^T(t)P_i(t) \quad (8) \\ a_i(t) &= \left[1 + \varphi_i^T(t)P_i(t)\varphi_i(t)\right]^{-1} \quad (9) \\ \varphi_i(t) &= \big[x_i^T(t), \ldots, x_i^T(t-p_i+1), u_i^T(t), \ldots, \\ &\quad u_i^T(t-q_i+1), \hat{w}_i^T(t), \ldots, \hat{w}_i^T(t-r_i+1), \\ &\quad x_{n_{i1}}^T(t), \ldots, x_{n_{i1}}^T(t-n_i+1), \ldots, x_{n_{i,m_i}}^T(t), \\ &\quad \ldots, x_{n_{i,m_i}}^T(t-n_i+1)\big]^T \quad (10) \\ \hat{w}_i(t) &= x_i(t) - \hat{\theta}_i^T(t)\varphi_i(t-1) \quad (11) \end{aligned}$$

where $\hat{\theta}_i(\cdot)$ is the estimate of $\theta_i$. Initial values $\hat{\theta}_i(0), \varphi_i(0) \neq 0$, $P_i(0) > 0$ can be arbitrarily taken.

*Remark 3:* In the design and analysis of decentralized adaptive control of large-scale systems, Lyapunov based method is commonly used, where the parameters are adjusted so that the tracking errors converge to zero. But here, for each agent we use an online (recursive) parameter estimator to estimate the unknown parameters, and design the adaptive control based on their fine estimations, i.e., the adaptive controller simultaneously performs the identification of the unknown plant.

*Remark 4:* As only local information is available for each agent, we cannot regard all the unknown parameter matrices of $N$ agents as a single parameter matrix to estimate, and use a centralized estimator as in [4] to design a controller and analyze the properties of the closed-loop system. Instead, an estimator has to be chosen by each agent $i$ ($i = 1, \ldots, N$), which may be different from each other due to their different initial values $\hat{\theta}_i(0)$, $P_i(0)$, and the regression vector $\varphi_i(t)$. From (7)–(11), the expression of each agent's estimator looks the same as in

[4], but from the discussion below, the analysis on the properties of the $N$ estimators and on the stability and optimality of the closed-loop systems are essentially different from [4]. The key difference lies in that the regressive matrix of each agent contains not only the state of itself but also the states of its neighbors.

Let

$$\alpha_i(t) \triangleq a_i(t)\|\tilde{\theta}_i^T(t)\varphi_i(t)\|^2,\ \delta_i(t) \triangleq\ tr(P_i(t) - P_i(t-1))$$
$$r_i(t) \triangleq 1 + \sum_{k=1}^{t} \varphi_i^T(k)\varphi_i(k),\ \tilde{\theta}_i(t) \triangleq\ \theta_i - \hat{\theta}_i(t). \quad (12)$$

Then by using the same technique as Lemma 1 in [39], we have the following results for algorithm (7)–(11).

*Lemma 3.2:* Under assumptions A1), A2) and A4), if $u_i(t)$ is $\overline{\mathcal{F}}_t^N$-measurable, then the ELS algorithm (7)–(11) has the following properties:

$$\|\tilde{\theta}_i(t+1)\|^2 = O\left(\frac{\log r_i(t)}{\lambda_i(t)}\right)$$
$$\sum_{k=1}^{t+1} \|\hat{w}_i(k) - w_i(k)\|^2 = O(\log r_i(t))$$
$$\sum_{k=1}^{t} \alpha_i(k) = O\left(\log r_i(t)\right) \text{ a.s.} \quad (13)$$

where $\overline{\mathcal{F}}_t^N$ is defined in Remark 2, $\lambda_i(t)$ is the minimum eigenvalue of $P_i^{-1}(t+1)$.

The conclusions of Lemma 3.2 are standard for the ELS algorithm, and can be proved in the same way as Lemma 1 in [39], only noticing that the $\sigma$-algebra under consideration at each time $t$ for each agent $i$, $i = 1, \ldots, N$, is $\overline{\mathcal{F}}_t^N$ rather than $\mathcal{F}_t^i$. A complete proof is very lengthy and almost the same as that of Lemma 1 in [39], thus, omitted here, due to the space limit.

### B. Adaptive Control Design

In this section, a distributed adaptive tracking control will be designed by using the CE principle, and the individual-mass interplay methodology, i.e., the mean field method [24]–[27], [29], [30] under the stochastic noncooperative game framework. In the setup of our game problem, the following rationality assumption is needed just as in [24]:

*Rationality Assumption:* 1) each agent optimizes its own cost function and 2) each agent assumes all other agents are being simultaneously rational when evaluating their competitive behavior.

We now give the design procedure of the control policy based on the individual-mass interplay methodology.

*Step 1: Estimate of the PSA Term $\overline{x}_N(k)$:* Denote the PSA estimate of agent $i$ by $f_i(k)$ ($i = 1, \ldots, N$, $k = 0, 1, \ldots, t$). Consider the distributed optimal tracking control problem of the systems (1) with cost function

$$\hat{J}_i^N(u_i, u_{-i}^N) = \limsup_{t\to\infty} \frac{1}{t} \sum_{k=0}^{t} \|x_i(k+1) - \Phi(f_i(k))\|^2.$$

In this case, from Remark 2 and the rationality assumption, the optimal control $u_i(t)$ of agent $i$ should satisfy:

$$\Phi(f_i(t)) = E\left[x_i(t+1)|\overline{\mathcal{F}}_t^N\right] = \theta_i^T\varphi_i^0(t), \quad t \geq 0 \quad (14)$$

where $\theta_i$ and $\varphi_i^0(t)$ are given by (5) and (6), respectively.

Then the closed-loop system of agent $i$ can be rewritten as

$$x_i(t+1) - \Phi(f_i(t)) = w_i(t+1), \quad i = 1, \ldots, N. \quad (15)$$

The PSA term at time $t+1$ is given by

$$\overline{x}_N(t+1) = \frac{1}{N}\sum_{i=1}^{N} \Phi(f_i(t)) + \frac{1}{N}\sum_{i=1}^{N} w_i(k+1)$$
$$f_i(0) = Ex_i(0).$$

Noticing that $\lim_{N\to\infty} (1/N)\sum_{i=1}^{N} w_i(k+1) = 0$, and assuming all agent can achieve the same estimate of the PSA term, i.e., $f_i(k) = f(k)$, $i = 1, \ldots, N$, then we can take $\Phi(f(t))$ as the estimate of $\overline{x}_N(t+1)$. In other words, when the population number $N$ is large enough, we can take the iterative solution $\{f(t), t \geq 0\}$ of

$$f(t+1) = \Phi(f(t)), \quad f(0) = x_0, \quad t \geq 1 \quad (16)$$

as the estimate of the PSA term.

Therefore, when the parameter matrix is known, by the NCE principle, (14) and (16) we can obtain the distributed optimal tracking control law for the systems (1)–(2). Denote it by $u_i^0(\cdot)$, which should satisfy:

$$\theta_i^T\varphi_i^0(t) = f(t+1), \quad t \geq 0. \quad (17)$$

In summary, the idea of the individual-mass methodology used in Step 1 is that if at time $t$, an estimate (i.e., $f(t)$) of the mass behavior (i.e., $\bar{x}_N(t)$) were given, the rationality would require each agent synthesize its individual cost based optimal control (i.e., $u_i(t)$) as a tracking action [see (14)]. Then, at time $t+1$, the resulted mass trajectory (i.e., $\bar{x}_N(t+1)$) would be used by the agents to iteratively update their estimates of the mass trajectory (i.e., $f(t+1)$), which can be proved to be very close to $\overline{x}_N(t+1)$ as the population number is sufficiently large [see (16) and the one above (16)].

*Step 2: Design of Adaptive Control:* When the parameter matrix is unknown, by the CE principle we should replace $\theta_i$ and $w_i(t)$ in (17) with their estimates $\hat{\theta}_i(t)$ and $\hat{w}_i(t)$ given by ELS algorithm (7)–(11), and then, obtain a distributed adaptive tracking control law for the systems (1)–(2). Denote it by $\hat{u}_i^*(\cdot)$, which should satisfy:

$$\hat{\theta}_i^T(t)\varphi_i(t) = f(t+1), \quad t \geq 0 \quad (18)$$

where $\hat{\theta}_i(t)$, $\varphi_i(t)$ and $f(t+1)$ are given by (7)–(11) and (16). From (18), when $\hat{B}_{i1}(t)$ is nonsingular, $\hat{u}_i^*(\cdot)$ can be equivalently written as

$$\hat{u}_i^*(t) = \hat{B}_{i1}^{-1}(t)\Bigg\{\Phi(f(t)) + \sum_{k=1}^{p_i} \hat{A}_{ik}(t)x_i(t-k+1)$$
$$- \sum_{k=2}^{q_i} \hat{B}_{ik}(t)\hat{u}_i^*(t-k+1) - \sum_{k=1}^{r_i} \hat{C}_{ik}(t)\hat{w}_i(t-k+1)$$
$$- \sum_{k=1}^{n_i} \sum_{j\in\mathcal{N}_i} \hat{G}_{ijk}(t)x_j(t-k+1)\Bigg\}.$$

To deal with the case where $\hat{B}_{i1}(t)$ is singular, we adopt the method of [39] to adjust $\hat{B}_{i1}(t)$:

$$\hat{B}^*_{i1}(t) = \begin{cases} \hat{B}_{i1}(t), & \text{if } \hat{B}^T_{i1}(t)\hat{B}_{i1}(t) \geq \frac{1}{\log r_i(t-1)} I, \\ \hat{B}_{i1}(t) + V_i(t)U_i^T(t)\frac{1}{\sqrt{\log r_i(t-1)}}, & \text{otherwise} \end{cases} \tag{19}$$

where $U_i(t)$, $V_i(t)$ are orthogonal matrices and obtained by the singular value decomposition [40] of $\hat{B}_{i1}(t)$, i.e., $\hat{B}_{i1}(t) = V_i(t)\begin{bmatrix}\Lambda_i(t) & 0 \\ 0 & 0\end{bmatrix}U_i^T(t)$, $\Lambda_i(t)$ is a positive definite diagonal matrix.

Thus, we obtain a distributed adaptive tracking control law for the case where the parameter matrix is unknown:

$$\begin{aligned} u_i^*(t) = \hat{B}^{*-1}_{i1}(t)\Bigg\{ & \Phi(f(t)) + \sum_{k=1}^{p_i}\hat{A}_{ik}(t)x_i(t-k+1) \\ & - \sum_{k=2}^{q_i}\hat{B}_{ik}(t)u_i^*(t-k+1) - \sum_{k=1}^{r_i}\hat{C}_{ik}(t)\hat{w}_i(t-k+1) \\ & - \sum_{k=1}^{n_i}\sum_{j\in\mathcal{N}_i}\hat{G}_{ijk}(t)x_j(t-k+1)\Bigg\} \end{aligned} \tag{20}$$

under which the dynamic equation of the closed-loop system can be written as

$$\begin{aligned} x_i(t+1) = {} & \tilde{\theta}_i^T(t)\varphi_i(t) + \theta_i^T\left(\varphi_i^0(t) - \varphi_i(t)\right) \\ & - \triangle\hat{B}_{i1}(t)u_i^*(t) + \Phi(f(t)) + w_i(t+1) \end{aligned} \tag{21}$$

where $\triangle\hat{B}_{i1}(t) \triangleq \hat{B}^*_{i1}(t) - \hat{B}_{i1}(t)$.

*Remark 5:* The mean field method adopted here does not rely on local interactions or connectivity properties for coordinating the multi-agent system. Each rational agent agrees to estimate the PSA term $\bar{x}_N(t)$ with the same value $f(t)$, which is derived from the same recursive (16), so that all agents then track the same reference signal $\Phi(f(t))$. The local information available to each agent on the connecting agents' state is not used for estimating the PSA but only to identify the equation governing its own dynamics. The possibility for the agents to locally infer the mean field effect is the key to the success of the NCE methodology. Here $\Phi(f(t))$ under the rationality assumption on all agents is important. The local forecast of the mean field evolution in various forms (for example, the equation system [24, (4.6)–(4.9)]). By this method, one can greatly reduce the complexity in both control computation and implementation since there is no necessity for each agent to collect the detailed state information from all other agents. Only macroscopic behavior matters.

### C. Closed-Loop System Analysis

In this subsection, we want to prove the closed-loop system (21) is uniformly stable w.r.t. $N$; the estimate of $\overline{x}_N(t)$ is strongly consistent; and the adaptive control group $\{U_N = \{u_i^* \in \mathcal{U}^N_{l,i}, 1 \leq i \leq N\}, N \geq 1\}$ is a.s. asymptotic equilibrium. To conduct the stability analysis, we need the following lemma.

*Lemma 3.3:* For the system (1), assume A1)–A4) hold, and the nonlinear iterative solution $\{f(t), t \geq 0\}$ of

$$f(t+1) = \Phi(f(t)) \quad \text{with } f(0) = x_0$$

is bounded. Then under the distributed adaptive tracking control (7)–(11), (16), and (20), we have

$$r(t) = O(t) \tag{22}$$

$$\begin{aligned} \sum_{k=0}^{t}\|\psi_i(k)\|^2 &= O(r^\delta(t)) + o(t) \\ &= o(t),\ \forall \delta \in \left(\frac{2}{\beta}, 1\right) \end{aligned} \tag{23}$$

where

$$\begin{aligned} r(t) &\triangleq \max\{r_1(t), \ldots, r_N(t)\} \\ \psi_i(k) &\triangleq \tilde{\theta}_i^T(k)\varphi_i(k) - (\triangle\hat{B}_{i1}(k))u_i^*(k) + \theta_i^T(\varphi_i^0(k) - \varphi_i(k)). \end{aligned} \tag{24}$$

*Proof:* The proof of this lemma is put into Appendix A. □

*Remark 6:* The proof idea of Lemma 3.3 is similar to Theorem 5.4 in [4], the key difficulty here stems from the state coupling of the $N$ agents, which requires us to analyze properties of the closed-loop systems of all the $N$ agents simultaneously. It is worth noticing that the $O(\cdot)$, $o(\cdot)$ in Lemma 3.3 depends on the population number $N$.

*Theorem 3.1:* Under the conditions of Lemma 3.3, the closed-loop system (21) is a.s. uniformly stable w.r.t. $N$. Specifically, $\sup_{N\geq 1}\max_{1\leq i\leq N}\limsup_{t\to\infty}(1/t)\sum_{k=1}^{t}\|x_i(k)\|^2 < \infty$ a.s.

*Proof:* For each agent $i$ ($i = 1, \ldots, N$), by (21) and (24) we have

$$\begin{aligned} &\frac{1}{t}\sum_{k=0}^{t}\|x_i(k+1) - \Phi(f(k))\|^2 \\ &= \frac{1}{t}\sum_{k=0}^{t}\|\psi_i(k) + w_i(k+1)\|^2 = \frac{1}{t}\sum_{k=0}^{t}\|\psi_i(k)\|^2 \\ &\quad + \frac{2}{t}\sum_{k=0}^{t}\psi_i^T(k)w_i(k+1) + \frac{1}{t}\sum_{k=0}^{t}\|w_i(k+1)\|^2. \end{aligned}$$

From Lemma 3.1, Lemma 3.3 and A1), for any $\epsilon > 0$,

$$\begin{aligned} \sum_{k=0}^{t}\psi_i^T(k)w_i(k+1) &= O\left(\left(\sum_{k=0}^{t}\|\psi_i(k)\|^2\right)^{1/2+\epsilon}\right) \\ &= o\left(t^{1/2+\epsilon}\right). \end{aligned} \tag{25}$$

Thus, by (23), (25), and A1) we have

$$\begin{aligned} &\lim_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}\|x_i(k+1) - \Phi(f(k))\|^2 \\ &= \lim_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}\|w_i(k+1)\|^2 = tr\left(R_\sigma\right), \quad \text{a.s.} \end{aligned}$$

This together with the boundness of $\{f(t), t \geq 0\}$ results in

$$\begin{aligned} &\limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}\|x_i(k+1)\|^2 \\ &\leq \lim_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}2\|x_i(k+1) - \Phi(f(k))\|^2 \\ &\quad + \limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}2\left\|\Phi(f(k))\right\|^2 \\ &= 2tr\left(R_\sigma\right) + 2\varpi < \infty, \quad \text{a.s.} \end{aligned}$$

where $\varpi = \limsup_{t\to\infty}(1/t)\sum_{k=0}^{t}\|\Phi(f(k))\|^2$. Noticing that $R_\sigma, \varpi$ are all independent of $N$, we obtain the desired result. □

*Remark 7:* Theorem 3.1 implies that, under the adaptive control (7)–(11), (16), and (20), the increase of the population does not affect the stability of the closed-loop system. This property is the prerequisite for discussing the a.s. asymptotic Nash equilibrium property of the adaptive control group.

Before analyzing and evaluating the performance of the distributed adaptive control (20), we first come to study the estimation error of $f(t)$ to the PSA term.

*Theorem 3.2:* For the system (1), assume A1)–A4) hold. Then under the adaptive control (7)–(11), (16), and (20), the closed-loop system has the following property:

$$\limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}\|\overline{x}_N(k)-f(k)\|^2 \le \frac{2tr(R_\sigma)}{N} \quad \text{a.s.} \tag{26}$$

*Proof:* By (21) we have

$$\overline{x}_N(t+1)-f(t+1)=\frac{1}{N}\sum_{i=1}^{N}\psi_i(t)+\frac{1}{N}\sum_{i=1}^{N}w_i(t+1). \tag{27}$$

By (23) we have

$$\frac{1}{t}\sum_{k=0}^{t}\left\|\frac{1}{N}\sum_{i=1}^{N}\psi_i(k)\right\|^2 \le \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{t}\sum_{k=0}^{t}\|\psi_i(k)\|^2\right) = o(1) \tag{28}$$

and by Theorem 4.1 of [27] we have

$$\limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}\left(\frac{1}{N}\sum_{i=1}^{N}w_i(k+1)\right)\times\left(\frac{1}{N}\sum_{i=1}^{N}w_i(k+1)\right)^T=\frac{1}{N}R_\sigma \quad \text{a.s.}$$

This together with (27) and (28) implies

$$\begin{aligned}\limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}&\|\overline{x}_N(k+1)-f(k+1)\|^2\\ &\le \limsup_{t\to\infty}\frac{2}{t}\sum_{k=0}^{t}\left\|\frac{1}{N}\sum_{i=1}^{N}\psi_i(k)\right\|^2\\ &\quad+\limsup_{t\to\infty}\frac{2}{t}\sum_{k=0}^{t}\left\|\frac{1}{N}\sum_{i=1}^{N}w_i(k+1)\right\|^2 = tr\left(\frac{2}{N}R_\sigma\right).\end{aligned}$$

From this, Theorem 3.2 can be easily obtained. □

*Remark 8:* From Theorem 3.2 we can show that the estimation of $f(t)$ to PSA term is consistent, as $N\to\infty$, i.e., $\lim_{N\to\infty}\limsup_{t\to\infty}(1/t)\sum_{k=0}^{t}\|\overline{x}_N(k)-f(k)\|^2=0$, a.s. This tells us that $f(t)$ can be regarded as the estimate of the PSA when the agent number is large, which is key to ensuring good properties of the adaptive control designed.

We now consider the a.s. asymptotic Nash equilibrium of the distributed adaptive tracking control group $\{U_N=\{u_i^*\in\mathcal{U}_{l,i}^N, 1\le i\le N\}, N\ge 1\}$ given by (7)–(11), (16), and (20) w.r.t. the corresponding performance index group $\{J_N=\{J_i^N, 1\le i\le N\}, N\ge 1\}$.

*Theorem 3.3:* For the system (1), $J_i^N$ $(i=1,2,\ldots,N)$ is defined by (2). Assume A1)–A4) hold, and $\Phi(\cdot)$ is $\mu$-Hölder continuous, i.e., $\|\Phi(x)-\Phi(y)\|\le\varrho\|x-y\|^\mu$, $\forall x,y\in\mathbb{R}$, where $\mu\in(0,1]$. Then under the distributed adaptive tracking controller (7)–(11), (16), and (20), the corresponding performance indices satisfy

$$J_i^N(u_i^*, u_{-i}^{*N})\le tr(R_\sigma)+2^{1+\mu}\varrho^2(tr(R_\sigma))^\mu N^{-\mu} \quad \text{a.s.}$$

*Proof:* By (2), (21), (24), and A1) we have

$$\begin{aligned}J_i^N&(u_i^*, u_{-i}^{*N})\\ &=\limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}\|\Phi(f(k))-\Phi(\overline{x}_N(k))+\psi_i(k)+w_i(k+1)\|^2\\ &\le tr(R_\sigma)+I_1^N+I_2^N\end{aligned} \tag{29}$$

where

$$I_1^N=\limsup_{t\to\infty}\frac{(1}{t)}\sum_{k=0}^{t}\|\Phi(f(k))-\Phi(\overline{x}_N(k))+\psi_i(k)\|^2$$

$$I_2^N=2\limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}(\Phi(f(k))-\Phi(\overline{x}_N(k))+\psi_i(k))^Tw_i(k+1).$$

From (23), (26), the $\mu$-Hölder continuity assumption of $\Phi(\cdot)$, and Jensen inequality [41], it follows that

$$\begin{aligned}I_1^N &\le 2\limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}\|\Phi(f(k))-\Phi(\overline{x}_N(k))\|^2\\ &\quad+2\limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}\|\psi_i(k)\|^2\\ &\le 2\varrho^2\left(\limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}\|f(k)-\overline{x}_N(k)\|^2\right)^\mu\\ &=2^{1+\mu}\varrho^2(tr(R_\sigma))^\mu N^{-\mu}.\end{aligned} \tag{30}$$

From this and Lemma 3.1, for any $\epsilon>0$ we have

$$\frac{1}{t}\sum_{k=0}^{t}(\Phi(f(k))-\Phi(\overline{x}_N(k))+\psi_i(k))^T\times w_i(k+1)=O\left(t^{-1/2+\epsilon}\right) \quad \text{a.s.}$$

Thus, $I_2^N=0$ a.s. This together with (29) and (30) leads to Theorem 3.3. □

*Corollary 3.1:* Under the condition of Theorem 3.3, the distributed adaptive control group $\{U_N=\{u_i^*\in\mathcal{U}_{l,i}^N, 1\le i\le N\}, N\ge 1\}$ given by (7)–(11), (16), and (20) is a.s. asymptotic Nash equilibrium w.r.t. the corresponding performance index group $\{J_N=\{J_i^N, 1\le i\le N\}, N\ge 1\}$. In addition, the convergence rate of each agent's performance index to the optimal value is $O(N^{-\mu})$.

*Proof:* Note that under the condition of Theorem 3.3, $\{\{w_i(t), \overline{\mathcal{F}}_t^N\}, 1\le i\le N, N\ge 1\}$ satisfies A1); the reference signal $\Phi(\overline{x}_N(t))$ and $x_j(t-k+1)$ $(j\in\mathcal{N}_i$, $k=1,\ldots,n_i)$ are all $\overline{\mathcal{F}}_t^N$-measurable; and $\Phi(\overline{x}_N(t))$ is uncorrelated with $w_i(t+1)$. Then, for any control group

$U_N = \{u_i \in \mathcal{U}_{g,i}^N, 1 \leq i \leq N\}$, by Theorem 3.6 in [4] we obtain

$$J_i^N(u_i, u_{-i}^N) \geq tr(R_\sigma) \quad \text{a.s.}, \quad i = 1, \ldots, N. \tag{31}$$

Particularly, we have that for any $u_i \in \mathcal{U}_{g,i}^N$, $J_i^N(u_i, u_{-i}^{*N}) \geq tr(R_\sigma)$ a.s., or equivalently, $\inf_{v_i \in \mathcal{U}_{g,i}^N} J_i^N(v_i, u_{-i}^{*N}) \geq tr(R_\sigma)$ a.s., $i = 1, \ldots, N$. This together with Theorem 3.3 leads to

$$\begin{aligned} tr(R_\sigma) &\leq \inf_{v_i \in \mathcal{U}_{g,i}^N} J_i^N(v_i, u_{-i}^{*N}) \leq J_i^N(u_i^*, u_{-i}^{*N}) \\ &\leq tr(R_\sigma) + 2^{1+\mu}\varrho^2(tr(R_\sigma))^\mu N^{-\mu} \quad \text{a.s.} \end{aligned}$$

Thus, by the Definition 2.1, the corollary is true. □

*Remark 9:* Similar to [27], the idea in this paper can be easily generalized to prove that the distributed adaptive control group $\{U_N = \{u_i^* \in \mathcal{U}_{l,i}^N, 1 \leq i \leq N\}, N \geq 1\}$ is asymptotic Nash equilibrium in probability w.r.t. the corresponding performance index group $\{J_N = \{J_i^N, 1 \leq i \leq N\}, N \geq 1\}$, under some mild conditions on the nonlinear function $\Phi(\cdot)$.

*Remark 10:* It is worth noticing that the information flow of each agent is embodied in the following two kinds of interactions. One is the individual-mass interaction, reflected by the index-coupled PSA term, and overcome by the individual-mass interplay methodology adequate for rational agents. The individual-mass interactions are comprehended from the game-theoretic viewpoint, and actually does not need communications with other agents [24]–[27], [29], [30]. The other information flow is the local communications with the neighbor agents, based on which parameter estimators are designed to cope with the weak couplings in each agent's dynamics. By using the estimates of the weak couplings and the certainty equivalence principle, the adaptive distributed controls are designed. We can see that it is the individual-mass interactions that play a key role in achieving the global behavior, i.e., the overall optimal tracking action, rather than the local communications with neighborhood.

## IV. Robust Adaptive Tracking Control and Performance Analysis

Due to the complexity of MASs, there always exists deviation between the mathematical model and the real dynamics of each agent, for instance, unmodeled dynamics, etc. At the same time, along with the evolution of itself and the environment, system parameters of each agent are always changing with time. Thus, it is more realistic to consider the following system model:

$$\begin{aligned} x_i(t+1) = &-\sum_{k=1}^{p_i} A_{ik}(t)x_i(t-k+1) + \sum_{k=1}^{q_i} B_{ik}(t)u_i(t-k+1) \\ &+ \sum_{k=1}^{n_i}\sum_{j\in\mathcal{N}_i} G_{ijk}(t)x_j(t-k+1) \\ &+ \omega_i(t+1) + \eta_i(t) \end{aligned} \tag{32}$$

where $x_i(t) \in \mathbb{R}^m$, $u_i(t) \in \mathbb{R}^m$, $\omega_i(t) \in \mathbb{R}^m$ denote the state, control input and external disturbances of agent $i$ at time $t$, $t \geq 0$, respectively, $x_i(t) = 0$, $u_i(t) = 0$, $w_i(t) = 0$ for $t < 0$; $\sum_{k=1}^{n_i}\sum_{j\in\mathcal{N}_i} G_{ijk}(t)x_j(t-k+1)$ denotes the interaction term of agent $i$ with its neighbors; $\eta_i(t) \in \mathbb{R}^m$ denotes the unmodeled dynamics; $\{A_{ik}(t), 1 \leq k \leq p_i\}$, $\{B_{ik}(t), 1 \leq k \leq q_i\}$, $\{G_{ijk}(t), j \in \mathcal{N}_i, 1 \leq k \leq n_i\}$ are unknown time-varying deterministic parameter matrices with proper dimensions.

Here, different from (1), the model (32) not only contains unmodeled dynamics, but also permits time-varying parameters. Obviously, when $G_{ijk}(t) = 0$, $k = 1, \ldots, n_i$, the interaction between $i$ and $j$ disappears, which leads the network topology of the MAS to be time-varying.

Based on the model (32), we investigate the robust tracking-type games via coupled performance indices, which are given by

$$\bar{J}_i^N(u_i, u_{-i}^N) = \limsup_{t\to\infty} \frac{1}{t}\sum_{k=0}^{t} E\|x_i(k+1) - \Phi(\overline{x}_N(k))\| \tag{33}$$

where definitions of $u_{-i}^N$, $\Phi(\cdot)$, $\overline{x}_N(k)$ are the same as (2).

*Remark 11:* Compared with (1), the model (32) can describe more general systems, which may include unknown time-varying parameters, unknown coupled interaction term, external disturbances and unmodeled dynamics, etc. Similar to $J_i^N(\cdot)$, the performance index $\bar{J}_i^N(\cdot)$ includes the unknown PSA term as well, but is deterministic. It is easier to conduct analysis than the stochastic case. From the discussions below, due to the difference of the parameter estimators, assumptions and the properties of the closed-loop systems, it is hard for us to treat Section III as a special case of this section.

In order to quantitatively analyze the minimization of the above performance indices $\{\bar{J}_i^N(\cdot), 1 \leq i \leq N\}$, we introduce the definition of $\epsilon$-Nash equilibrium.

*Definition 4.1 [24]:* For the system (32), (33) and control group $U_N = \{u_i \in \mathcal{U}_i^N, 1 \leq i \leq N\}$, if there exists $\epsilon \geq 0$, such that $\bar{J}_i^N(u_i, u_{-i}^N) \leq \inf_{v_i \in \mathcal{U}_{g,i}^N} \bar{J}_i^N(v_i, u_{-i}^N) + \epsilon$, $\forall i = 1, \ldots, N$, then we call the control group $U_N$ is $\epsilon$-Nash equilibrium w.r.t. the index group $\bar{J}_N = \{\bar{J}_i^N, 1 \leq i \leq N\}$. Particularly, if $\epsilon = 0$, then the control group $U_N$ is called Nash equilibrium w.r.t. the performance index group $\bar{J}_N$.

In this section, we will design a distributed optimal adaptive tracking policy for the systems (32) and (33), in order to handle the existing uncertainties, i.e., unmodeled dynamics, unknown time-varying parameters, unknown coupled interaction term, and external disturbances. Specifically, the adaptive estimation algorithm is given in Section IV-A, the adaptive control is constructively designed in Section IV-B, and the properties of the closed-loop system are analyzed in Section IV-C.

### A. Estimation Algorithm

To simplify the expression, we rewrite the system (32) in the following form:

$$x_i(t+1) = \theta_i^T(t)\phi_i(t) + \omega_i(t+1) + \eta_i(t), \quad t \geq 0 \tag{34}$$

where

$$\begin{aligned} \theta_i(t) = &[-A_{i1}(t), \ldots, -A_{i,p_i}(t), B_{i1}(t), \ldots, B_{i,q_i}(t), \\ &G_{i,n_{i1},1}(t), \ldots, G_{i,n_{i1},n_i}(t), \ldots, \\ &G_{i,n_{i,m_i},1}(t), \ldots, G_{i,n_{i,m_i},n_i}(t)]^T. \end{aligned} \tag{35}$$

$$\begin{aligned} \phi_i(t) = &[x_i^T(t), \ldots, x_i^T(t-p_i+1), u_i^T(t), \ldots, u_i^T(t-q_i+1), \\ &x_{n_{i1}}^T(t), \ldots, x_{n_{i1}}^T(t-n_i+1), \ldots, x_{n_{i,m_i}}^T(t), \\ &x_{n_{i1}}^T(t), \ldots, \ldots, x_{n_{i,m_i}}^T(t-n_i+1)]^T. \end{aligned} \tag{36}$$

To proceed the robust analysis of the systems (32)–(33), we need the the following assumptions.

*B1):* There exist $\rho_i \in (0,1)$, $C_{0i} > 0$ and $b_{i1} > 0$ such that $B_{i1}^T(t)B_{i1}(t) \geq b_{i1}^2 I$, $\forall t \geq 0$, and

$$\|B_i(t)\cdots B_i(k)\| \leq C_{0i}\rho_i^{t-k+1} \quad \forall t \geq k, \quad t,k = 0,1,\ldots$$

where

$$B_i(t) = \begin{bmatrix} -B_{i1}^{-1}(t)B_{i2}(t) & \cdots & -B_{i1}^{-1}(t)B_{iq_i}(t) & 0 \\ I & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & I & 0 \end{bmatrix}.$$

*B2):* There exist nonnegative constants $\delta, \delta_i$ $(i = 1,\ldots,N)$, such that

$$\limsup_{t-t_0\to\infty} \frac{1}{t-t_0} \sum_{k=t_0+1}^{t} \|\theta_i(k) - \theta_i(k-1)\|^2 \leq \delta_i^2 \leq \delta^2,\ \forall t_0 \geq 0.$$

*B3):* There is a known convex compact set $D_i$ such that $\theta_i(t) \in D_i$ for all $t \geq 0$, and $B_{i1}$ is nonsingular for all $\theta_i \in D_i$. Here $\theta_i$ and $B_{i1}$ are with the form of (35).

*B4):* $\{\{\omega_i(t), t \geq 0\}, i \geq 1\}$ is a group of deterministic or stochastic disturbance sequence on the probability space $(\Omega, \mathcal{F}, P)$, satisfying $E\|\omega_i(t)\|^2 \leq \sigma_\omega^2$, and there exists $\varepsilon_{i0} > 0$, such that for any $t \geq k \geq 0$, $E\exp(\sum_{j=k}^{t} \varepsilon_{i0}\|\omega_i(j)\|^2) \leq \exp(c_{i1}(t-k+1) + c_{i2})$, where $c_{i1}, c_{i2}, \sigma_\omega$ are positive constants.

*B5):* There exists a nonnegative sequence $\{\varepsilon_i(t)\}$ such that $\|\eta_i(t)\| \leq \varepsilon_i(t) f_i(t)$, $t \geq 0$, $i = 1,\ldots,N$, where $f_i(t)$ satisfies $f_i(t) = \gamma_i f_i(t-1) + \|\phi_i(t)\|$, $\gamma_i \in (0,1)$, $f_i(0) > 0$. The sequence $\{\varepsilon_i(t)\}$ is bounded and satisfies $\limsup_{t-t_0\to\infty}(1/(t-t_0))\sum_{k=t_0+1}^{t} \varepsilon_i^2(k) \leq \varepsilon_i \leq \varepsilon$, $\forall t_0 \geq 0$, where $\varepsilon, \varepsilon_i > 0$ are constants.

*B6):* $\{x_i(0), i \geq 1\}$ is an independent random variable sequence with a common mathematical expectation $x_0 = Ex_i(0)$.

*Remark 12:* The above assumptions are the same as single-agent case [42] except for B3) and B6). B1) is the generalization of the standard minimum phase condition in the constant parameter case. B2) describes the variation rate of the unknown parameter matrices $\{\theta_i(t), t \geq 0\}$ in the time average sense. In particular, when $\delta$ is small, it means that the system is slowly time-varying in the time average sense. The disturbance sequence $\{\{\omega_i(t), t \geq 0\}, i \geq 1\}$ in B4) contains all uniformly bounded noise sequence in sample path, and some classes of a.s. unbounded noise sequence [42]. B5) is a constraint on the divergence speed of the unmodeled dynamics, which includes both unmodeled part of each agent's dynamics and unmodeled part of interactions with its neighbors. B6) is a condition on initial states.

As it is well-known, the ELS algorithm is more adequate for the constant parameter case, but now we are facing the time-varying case. Thus, different from Section III, we will apply a projected LMS algorithm [42] to estimate the unknown time-varying parameter matrix $\theta_i(t)$ $(i = 1,\ldots,N)$:

$$\hat{\theta}_i(t+1) = \pi_{D_i}\left\{\hat{\theta}_i(t) + \frac{\phi_i(t)}{d_i + f_i^2(t)} \cdot \left(x_i(t+1) - \hat{\theta}_i^T(t)\phi_i(t)\right)^T\right\} \quad (37)$$

where $\hat{\theta}_i(t)$[1] is the estimate for $\theta_i(t)$; $\phi_i(t)$ and $f_i(t)$ are given by (36) and B5), respectively; $D_i$ is the convex compact set given in Assumption B3). $\pi_{D_i}\{x\}$ denotes the nearest point in $D_i$ to $x$, i.e., $\pi_{D_i}\{x\} = \operatorname{argmin}_{y\in D_i}\|x-y\|$; $\hat{\theta}_i(0) \in D_i$; $d_i > 0$ is a sufficiently large constant, specific constraint on which will be discussed below in Lemma 4.2.

*Remark 13:* Since $D_i$ is a convex compact set, there exists $R_i > 0$, such that $\|\theta_i\| \leq R_i$ for all $\theta_i \in D_i$. In addition, for any given $y \in \mathbb{R}^{m\times m(p_i+q_i+m_i n_i)}$, $\pi_{D_i}\{y\}$ is existent and unique; and for all $x \in D_i$, $\|x - \pi_{D_i}\{y\}\| \leq \|x-y\|$. This is key to analyzing the properties of the projected LMS algorithm (37).

For the projected LMS algorithm (37), we have the following estimates.

*Lemma 4.1:* Consider the system (34) and the estimation algorithm (37), if Assumptions B3) and B5) hold, then the estimation error has the following property:

$$\bar{\alpha}_i(t) \leq 2\left(\|\tilde{\theta}_i(t)\|^2 - \|\tilde{\theta}_i(t+1)\|^2\right)\frac{6\|\omega_i(t+1)\|^2}{d_i} + 6\varepsilon_i^2(t) + 2\delta_i(t+1)\left(\delta_i(t+1) + 4\sqrt{m}R_i\right)$$

where $R_i$ is the constant mentioned in Remark 13, and $\bar{\alpha}_i(t) = \|\tilde{\theta}_i^T(t)\phi_i(t)\|^2/(d_i + f_i^2(t))$, $\delta_i(t+1) = \|\theta_i(t+1) - \theta_i(t)\|$.

*Proof:* The key ideas and techniques for proving this lemma are the same as Lemma 5.1 in [42]. Here, due to high dimension, the proof is slightly different. See Lemma 6.3 of [43, p. 126]. □

### *B. Distributed Robust Adaptive Control Design*

In this subsection, we will constructively design a distributed robust adaptive tracking control by using the NCE principle [24], [27] and CE principle.

*Step 1: Estimate of the Unknown PSA Term $\overline{x}_N(k)$:* Assume the estimate of $\overline{x}_N(k)$ at time $k$ $(k = 0,1,\ldots,t)$ is $f(k)$. By the rationality assumption, the distributed robust tracking control $u_i(t)$ $(i = 1,\ldots,N)$ of the systems (32) and $\hat{\tilde{J}}_i^N(u_i, u_{-i}^N) = \limsup_{t\to\infty}(1/t)\sum_{k=0}^{t} E\|x_i(k+1) - \Phi(f(k))\|$ should satisfy

$$\Phi(f(t)) = \theta_i^T(t)\phi_i(t), \quad t \geq 0 \quad (38)$$

where $\theta_i(t)$ and $\phi_i(t)$ are defined by (35) and (36), respectively.

The closed-loop system of agent $i$ $(i = 1,\ldots,N)$ can be rewritten as $x_i(t+1) - \Phi(f(t)) = \omega_i(t+1) + \eta_i(t)$, $t \geq 0$. The PSA term at time $t+1$ is given by $\overline{x}_N(t+1) = \Phi(f(t)) + (1/N)\sum_{i=1}^{N}(\omega_i(t+1) + \eta_i(t))$, $f(0) = (1/N)\sum_{i=1}^{N} x_i(0)$.

From B6), when the population number $N$ is sufficiently large, the external disturbance and unmodeled dynamics are sufficiently small, this suggests us to take the iterative solution $\{f(t), t \geq 0\}$ of (16) as the estimate of the PSA term. Thus, when the parameter matrices $\{\theta_i(t),\ i = 1,\ldots,N\}$ are known, by NCE principle we can obtain a distributed tracking control for the system (32) and (33) based on (16) and (38). Denote it by $u_i^0(\cdot)$, which should satisfy

$$\theta_i^T(t)\phi_i(t) = f(t+1), \quad t \geq 0. \quad (39)$$

*Step 2: Design of Robust Adaptive Control:* When the time-varying parameter matrix $\theta_i(t)$ is unknown, by the CE principle

[1]To be consistent, symbols in this section will have the same meaning as those in Section III. For example, $\tilde{\theta}_i(t)$ denotes the parameter estimation error; $u_i^*(t)$ denotes the adaptive tracking control.

we replace the parameters in (39) with its estimate $\hat{\theta}_i(t)$ given by the projected LMS algorithm (37), and obtain a distributed robust tracking control. Denote it by $u_i^*(\cdot)$, which should satisfy $f(t+1) = \hat{\theta}_i^T(t)\phi_i(t)$, $\forall t \geq 0$, where $\phi_i(t)$ and $\{f(t)\}$ are given by (36) and (16), respectively. From the above equation and definition of $D_i$ $(i = 1,\ldots,N)$ in (37), the control $u_i^*(t)$ can be explicitly written as

$$u_i^*(t) = \hat{B}_{i1}^{-1}(t)\Bigg\{f(t+1) + \sum_{k=1}^{p_i}\hat{A}_{ik}(t)x_i(t-k+1) - \sum_{k=2}^{q_i}\hat{B}_{ik}(t)u_i^*(t-k+1) - \sum_{k=1}^{n_i}\sum_{j\in\mathcal{N}_i}\hat{G}_{ijk}(t)x_j(t-k+1)\Bigg\}. \quad (40)$$

The closed-loop system of agent $i$ $(i = 1,\ldots,N)$ under control (40) can be written as

$$x_i(t+1) = \tilde{\theta}_i^T(t)\phi_i(t) + f(t+1) + \omega_i(t+1) + \eta_i(t). \quad (41)$$

### C. Closed-Loop System Analysis

In this subsection, we will prove the closed-loop system (41) is stable and robust w.r.t. the system uncertainties; the consistency of the PSA estimate is robust w.r.t. the system uncertainties; and the control group $U_N = \{u_i^* \in \mathcal{U}_{l,i}^N,\ 1 \leq i \leq N\}$ is $\epsilon$-Nash equilibrium. To do so, we need the following results.

*Lemma 4.2:* For the system (32), (33), (37), (40), and (16), assume B1)–B5) hold, and the nonlinear iterative solution of $f(t+1) = \Phi(f(t))$ with $f(0) = x_0$ is bounded by a constant $M$ (i.e., $\|f(t)\| \leq M$). If $\varepsilon$ and $\delta$ are sufficiently small, and $\varepsilon_{i0}d_i$ is sufficiently large, such that $\bar{\lambda}(1-\gamma_i)\varepsilon_{i0}d_i \geq 48C_{2i}N$ and $\bar{\lambda}e^{c_1} \in (0,1)$, then we have

$$E\|\overline{L}_t\| = O(d), \quad \sup_t E\|\phi_i(t)\|^2 = O(d) \quad (42)$$

where $d = \max\{d_1,\ldots,d_N\}$, $\{\overline{L}_t\}$ is defined by

$$\overline{L}_t = \sum_{k=0}^{t}\bar{\lambda}^{t-k}\left(\|X(k)\|^2 + \sum_{i=1}^{N}\|\omega_i(k)\|^2 + \sum_{i=1}^{N}\|\eta_i(k-1)\|^2\right) \quad (43)$$

and satisfies the following recursive inequality:

$$\overline{L}_t \leq \left\{\bar{\lambda} + \sum_{i=1}^{N}\Delta_i(t-1)\right\}\overline{L}_{t-1} + \overline{\zeta}(t). \quad (44)$$

Here $X(k) = \left[x_1^T(k),\ldots,x_N^T(k)\right]^T$, $\bar{\lambda} \in (\max\{\gamma_i,\rho_i, i = 1,\ldots,N\},1)$

$$\Delta_i(j) = \frac{4C_{2i}}{1-\gamma_i}\bar{\alpha}_i(j-1) + \frac{14C_{2i}}{1-\gamma_i}\varepsilon_i^2(j-1)$$

$$c_1 = \frac{1}{\bar{\lambda}}\sum_{i=1}^{N}\left(\frac{38C_{2i}}{1-\gamma_i}\varepsilon_i + \frac{8C_{2i}}{1-\gamma_i}\delta_i^2 + \frac{32\sqrt{m}C_{2i}R_i}{1-\gamma_i}\delta_i + \frac{24c_{i1}C_{2i}}{(1-\gamma_i)d_i\varepsilon_{i0}}\right)$$

$$C_{1i} = \frac{2C_{0i}^2(4mb_{i1}^{-2}R_i^2+1)\|u_i(0)\|^2}{\rho_i^2(1-\rho_i)} + \frac{4mM^2}{b_{i1}^2(1-\bar{\lambda})}$$

$$C_{2i} = \max\Bigg\{(4mb_{i1}^{-2}R_i^2+1)p_i\bar{\lambda}^{-p_i+1} + \frac{10m\bar{\lambda}C_{0i}^2(4mb_{i1}^{-2}R_i^2+1)(p_iR_i^2\bar{\lambda}^{-p_i+1}+1)}{b_{i1}^2(1-\rho_i)(\bar{\lambda}-\rho_i)}, (4mn_ib_{i1}^{-2}R_i^2+1)n_i\bar{\lambda}^{-n_i+1} + \frac{10mn_i^2\bar{\lambda}C_{0i}^2R_i^2\rho_i^{-2n_i+2}(4mb_{i1}^{-2}R_i^2+1)}{b_{i1}^2(1-\rho_i)(\bar{\lambda}-\rho_i)}\Bigg\}$$

$$\overline{\zeta}(t) = \sum_{i=1}^{N}\Bigg\{2d_i\bar{\alpha}_i(t-1) + 6\|\omega_i(t)\|^2 + 6M^2 + 4\left(f_i^2(0)\gamma_i^{2t-2} + \frac{C_{1i}}{1-\gamma_i}\right)\bar{\alpha}_i(t-1) + 14\varepsilon_i^2(t-1)\left(f_i^2(0)\gamma_i^{2t-2} + \frac{C_{1i}}{1-\gamma_i}\right)\Bigg\}. \quad (45)$$

*Proof:* The proof of this lemma is put into Appendix B. □

*Remark 14:* Compared with the single-agent case, the key difficulty here stems from dealing with the state coupling of the $N$ agents, which requires us to analyze the closed-loop properties of all the $N$ agents simultaneously. It is worth noticing that the $O(\cdot)$ in Lemma 4.2 depends on the population number $N$; and the choice of $d_i$ needs the population number $N$ and $\bar{\lambda}$, which, in some sense, is a kind of global information. Of course, when $N$ is not exactly known, we can use an upper of $N$ to replace $N$ in the above equation.

*Remark 15:* From B2), the requirement of $\delta$ to be sufficiently small implies that the system (32) is slowly time-varying in the time average sense. From B5), the requirement of $\varepsilon$ to be sufficiently small implies that the unmodeled dynamics $\eta_i(t)$ is small w.r.t. $\|\phi_i(t)\|$.

*Theorem 4.1:* Under the conditions of Lemma 4.2, the closed-loop system (41) has the following property:

$$\limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}E\|x_i(k+1)\| < \infty. \quad (46)$$

*Proof:* For agent $i$ $(i = 1,\ldots,N)$, by (41) we have

$$\frac{1}{t}\sum_{k=0}^{t}E\|x_i(k+1) - \Phi(f(k))\| \leq \frac{1}{t}\sum_{k=0}^{t}E\|\tilde{\theta}_i^T(k)\phi_i(k)\| + \frac{1}{t}\sum_{k=0}^{t}E\|\omega_i(k+1)\| + \frac{1}{t}\sum_{k=0}^{t}E\|\eta_i(k)\|.$$

From B5) and Lemma 4.2 we have

$$Ef_i^2(t) = o(1) + O\left(\sum_{k=0}^{t}\gamma_i^{t-k}E\|\phi_i(k)\|^2\right) = O(1) + O(\sup_t E\|\phi_i(t)\|^2) = O(d). \quad (47)$$

This together with Lemma 4.1 and Jensen inequality [41] implies

$$\limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}E\|\tilde{\theta}_i^T(k)\phi_i(k)\| = \limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}E\left\{\bar{\alpha}_i(k)(d_i + f_i^2(k))\right\}^{1/2}$$

$$\le \left\{ \limsup_{t\to\infty} \frac{1}{t} \sum_{k=0}^{t} E\bar{\alpha}_i(k)(d_i + Ef_i^2(k)) \right\}^{1/2} = O(\sigma_\omega + \sqrt{(\varepsilon+\delta)d}). \tag{48}$$

From (47) and B5) we have

$$\frac{1}{t}\sum_{k=0}^{t} E\|\eta_i(k)\| = \frac{1}{t}\sum_{k=0}^{t} O(\varepsilon_i(k)Ef_i(k)) = O\left((\varepsilon d)^{1/2}\right). \tag{49}$$

From Lyapunov inequality [41] and B4) we have

$$\frac{1}{t}\sum_{k=0}^{t} E\|\omega_i(k+1)\| \le \frac{1}{t}\sum_{k=0}^{t}(E\|\omega_i(k+1)\|^2)^{1/2} \le \sigma_\omega \tag{50}$$

which together with (48), (49) and the boundness of sequence $\{f(t), t \ge 0\}$ gives

$$\begin{aligned}\limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t} E\|x_i(k+1))\| &\le \limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t} E\|x_i(k+1) - \Phi(f(k))\| \\ &\quad + \limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}\|\Phi(f(k))\| \\ &= O(\sigma_\omega + \sqrt{(\varepsilon+\delta)d}) + \varpi \end{aligned}\tag{51}$$

where $\varpi = \limsup_{t\to\infty}(1/t)\sum_{k=0}^{t}\|\Phi(f(k))\|$. This implies (46). □

*Remark 16:* From Theorem 4.1, it is hard to get the robust property of the uniform stability w.r.t. $N$ as in Theorem 3.1. The key difficulty lies in the close relation of the estimate in (51) with the population number $N$. But the asymptotic robust stability can be obtained by this theorem, which means no matter how large it is, the stability of the close-loop system (41) is robust w.r.t. external disturbances, slowly time-varying parameters and unmodeled dynamics.

Now we turn to study the consistency of the PSA estimate.

*Theorem 4.2:* For the system (32), assume B1)–B6) hold, and the nonlinear iterative solution of $f(t+1) = \Phi(f(t))$ with $f(0) = x_0$ is bounded. If $\varepsilon_{i0}d_i$ is sufficiently large, such that $\bar{\lambda}(1-\gamma_i)\varepsilon_{i0}d_i \ge 48C_{2i}N$ and $\bar{\lambda}e^{\bar{c}_1} \in (0,1)$, then under the distributed adaptive control (40), (37) and (16), the closed-loop system (41) has the following property:

$$\lim_{\substack{\sigma_\omega\to 0\\ \varepsilon\to 0\\ \delta\to 0}} \limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t} E\|\overline{x}_N(k) - f(k)\| = 0$$

where $\bar{c}_1 = (1/\bar{\lambda})\sum_{i=1}^{N} 24c_{i1}C_{2i}/(1-\gamma_i)d_i\varepsilon_{i0}$, and $C_{2i}$ is defined in (45).

*Proof:* By (41) we have

$$\overline{x}_N(t+1) - f(t+1) = \frac{1}{N}\sum_{i=1}^{N}\tilde{\theta}_i^T(t)\phi_i(t) + \frac{1}{N}\sum_{i=1}^{N}\eta_i(t) + \frac{1}{N}\sum_{i=1}^{N}\omega_i(t+1).$$

From (48) one can get

$$\begin{aligned}\limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t} E\|\frac{1}{N}\sum_{i=1}^{N}\tilde{\theta}_i^T(k)\phi_i(k)\| &\le \frac{1}{N}\sum_{i=1}^{N}\left(\limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}E\|\tilde{\theta}_i^T(k)\phi_i(k)\|\right) \\ &= O(\sigma_\omega + \sqrt{(\varepsilon+\delta)d}).\end{aligned}$$

This together with (49) and (50) leads to $\limsup_{t\to\infty}(1/t)\sum_{k=0}^{t}E\|\overline{x}_N(k) - f(k)\| = O(\sigma_\omega + \sqrt{(\varepsilon+\delta)d})$. Thus, the theorem holds for all given $N$. □

*Remark 17:* From the laws of large numbers, when the population number $N$ is sufficiently large, the estimation error of $\overline{x}_N(0)$ by $f(0)$ should be small. This may lead to a better estimate of $\overline{x}_N(t)$ by $f(t)$ when $N$ is large.

We now analyze the robust Nash equilibrium of the designed adaptive control.

*Definition 4.2:* For the system (32) and a control group $U_N = \{u_i \in \mathcal{U}_i^N, 1 \le i \le N\}$, we say the control group $U_N$ is robust Nash equilibrium w.r.t. the index group $\bar{J}_N = \{\bar{J}_i^N, 1 \le i \le N\}$, if $U_N$ is Nash equilibrium w.r.t. $\bar{J}_N$, as the uncertainties $\sigma_\omega \to 0$, $\varepsilon \to 0$, $\delta \to 0$.

*Theorem 4.3:* For the system (32), (33), assume B1)–B6) hold, and the nonlinear iterative solution of $f(t+1) = \Phi(f(t))$ with $f(0) = x_0$ is bounded. The nonlinear function $\Phi(\cdot)$ is $\mu$-Hölder continuous, i.e., $\|\Phi(x) - \Phi(y)\| \le \varrho\|x-y\|^\mu, \forall x, y \in \mathbb{R}$, where $\mu \in (0,1]$. If $\varepsilon$ and $\delta$ are sufficiently small, and $\varepsilon_{i0}d_i$ is sufficiently large, such that $\bar{\lambda}(1-\gamma_i)\varepsilon_{i0}d_i \ge 48C_{2i}N$ and $\bar{\lambda}e^{c_1} \in (0,1)$, then under the distributed adaptive control (40), (37), and (16), we have

$$\bar{J}_i^N(u_i^*, u_{-i}^{*N}) = O\Big((\sigma_\omega + \sqrt{(\varepsilon+\delta)d})^\mu \times \left[1 + (\sigma_\omega + \sqrt{(\varepsilon+\delta)d})^{1-\mu}\right]\Big) \tag{52}$$

where $\bar{\lambda}$, $c_1$, $C_{2i}$ $(i = 1, 2, \ldots, N)$ have the same meanings as in Lemma 4.2.

*Proof:* By (33) and (41) we have

$$\bar{J}_i^N(u_i^*, u_{-i}^{*N}) = \limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}E\|\Phi(f(k)) - \Phi(\overline{x}_N(k))\| + \limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}E\|\tilde{\theta}_i^T(k)\phi_i(k) + \eta_i(k) + \omega_i(k+1)\| \triangleq \bar{I}_1 + \bar{I}_2.$$

From the conditions of this theorem, Theorem 4.2 and Jensen inequality [41], we have

$$\bar{I}_1 \le \varrho\left\{\limsup_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t}E\|\overline{x}_N(k) - f(k)\|\right\}^\mu = O\left((\sigma_\omega + \sqrt{(\varepsilon+\delta)d})^\mu\right). \tag{53}$$

By (48), (49), and (50) we have $\bar{I}_2 = O(\sigma_\omega + \sqrt{(\varepsilon+\delta)d})$. This together with (53) implies (52). □

*Remark 18:* Under the global-measurement based control law $\hat{\theta}_i^T(t)\phi_i(t) = \Phi(\overline{x}_N(t))$ for agent $i$, and $u_{-i}^{*N}(t)$ for other agents, we can similarly prove that the global optimal performance index value is 0, when the uncertainties $\sigma_\omega$, $\varepsilon$ and $\delta$ tend to zero. From Theorem 4.3, we have $\bar{J}_i^N(u_i^*, u_{-i}^{*N}) \to 0$, as $\sigma_\omega \to 0$, $\varepsilon \to 0$, $\delta \to 0$. Thus, by Definition 4.2, the distributed control group $U_N$ is robust Nash equilibrium w.r.t. the corresponding performance index group $\bar{J}_N$. When $\varepsilon$, $\delta$, $\sigma_\omega$ are non-zero, by Definition 4.1, the designed adaptive control group $U_N$ is $\epsilon$-Nash equilibrium w.r.t. $\bar{J}_N$.

*Remark 19:* It is worth to considering the robustness issue when the initializations for the PSA estimate of all agents are different. The key difficulty to prove the PSA estimates of different agents to be consistent w.r.t. $N$ is that some conditions on $\Phi(\cdot)$ are needed to ensure the establishment of the estimates for the macroscopic dynamic behavior $\bar{L}_t$ in (44), and to make the initial PSA error for each agent not amplified in time average sense when it is iterated through the function $\Phi(\cdot)$.

## V. Numerical Examples

In this section, we will give two numerical examples to illustrate the results of Sections III and IV.

*Example 1:* Consider the production output adjustment for a dynamical market with $N$ firms supplying the same product [24], [38]. Different from [24], the production dynamics of each firm is modeled by a discrete time dynamic system, which is influenced by its neighbor firms. Here, neighbors of the firm $i$ can be interpreted as those joint firms distributed in different regions that have cooperation or coordination relationships with $i$. The production dynamics of the firm $i$ ($i = 1, \ldots, N$) is modeled by

$$x_i(t+1) = x_i(t) + bu_i(t) - \alpha\left(x_i(t) - \frac{1}{|\mathcal{N}_i|}\sum_{j\in\mathcal{N}_i} x_j(t)\right) + w_i(t+1)$$

where $x_i(t)$ and $u_i(t)$ denote the production output and adjustment control of the firm $i$, respectively; $\mathcal{N}_i = \{n_{i1}, \ldots, n_{i,m_i}\}$ denotes the joint firms of the firm $i$, $|\mathcal{N}_i|$ is the number of its joint firms; $\alpha(x_i(t) - (1/|\mathcal{N}_i|)\sum_{j\in\mathcal{N}_i} x_j(t))$ denotes the coordination relationship of the firm $i$ with its joint firms; $\{x_i(0),\ i = 1, \ldots, N\}$ are independent and identically distributed (i.i.d.) r.v.s with uniform distribution on $[x_0 - \bar{a}, x_0 + \bar{a}]$; $\{w_i(t),\ t \geq 0\}$ is a sequence of Gaussian white noise with distribution $\mathscr{N}(0, \sigma_w^2)$.

The following performance index [24] is expected to be optimized by the firm $i$ ($i = 1, \ldots, N$):

$$J_i^N(u_i, u_{-i}^N) = \limsup_{t\to\infty} \frac{1}{t}\sum_{k=0}^{t}\left(x_i(k+1) - \bar{\beta}p\right)^2$$

where $\bar{\beta} > 0$, $p = \bar{\eta} - \bar{\rho}\bar{x}_N(t)$ is the production price, $\bar{\eta}, \bar{\rho} > 0$.

The unknown parameter vector is $\theta_i = [1 - \alpha, b, \alpha/|\mathcal{N}_i|, \ldots, \alpha/|\mathcal{N}_i|]^T$. The coupling function in indexes (2) is $\Phi(x) = \bar{\beta}(\bar{\eta} - \bar{\rho}x)$. Obviously, $\Phi(\cdot)$ is $\mu$-Hölder continuous with $\mu = 1$, $\varrho = \bar{\rho}\bar{\beta}$.

Let $\alpha = 0.15$, $b = 2$, $\bar{\beta} = 0.95$, $\bar{\eta} = 400$, $\bar{\rho} = 0.98$, $x_0 = 200$, $\bar{a} = 15$, $\sigma_w^2 = 2$, $\mathcal{N}_i = \{i+1\}$, $i = 1, \ldots, N-1$, $\mathcal{N}_N = \{1\}$. Then, the above system satisfies the assumptions A1)–A4). Apply the parameter estimator (7)–(11) to estimate the unknown parameter vector. Initial values are taken as $\hat{\theta}_i(0) = 0$, $P_i(0) = I$, $i = 1, \ldots, N$. The estimate of the high-frequency gain $b$ is given by (19). The adaptive tracking control is designed as (20), where $f(t)$ is given by (16). When the number of firms $N = 100$, the trajectories of the closed-loop system are shown in Fig. 1.

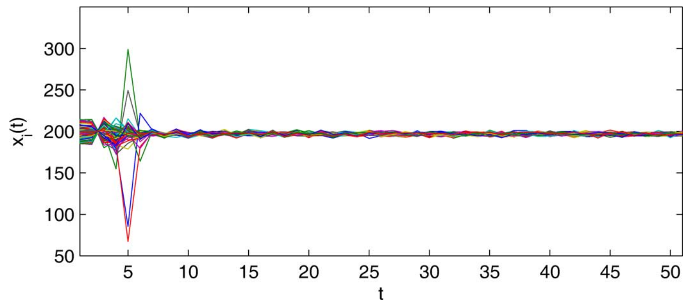


Fig. 1. Output trajectories of the firms, when $N = 100$.

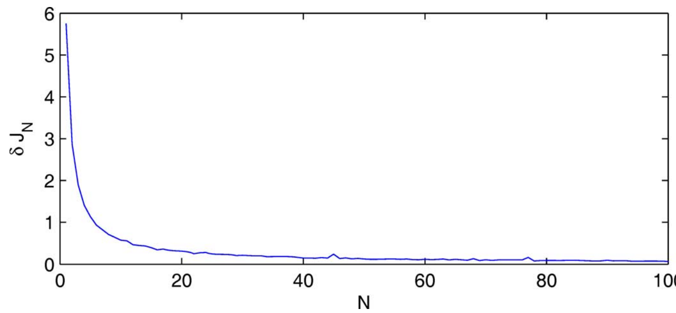


Fig. 2. Curve of $\delta J_N$ with respect to $N$.

Let the number of firms $N$ increase from 1 to 100, and denote the difference between the maximum of the performance indices and $\sigma_w^2$ by $\delta J_N$, the curve of $\delta J_N$ w.r.t. $N$ is shown in Fig. 2, from which one can find that, when $N$ increases, the performance index of each firm tends to $\sigma_w^2$. This is consistent with the conclusion of Theorem 3.3.

*Example 2:* Consider the following time-varying stochastic MAS described by

$$\begin{cases} x_i(t+1) = a_i(t)x_i(t) + b_i(t)u_i(t) + g_i(t)x_{i+1}(t) \\ \qquad\qquad +\omega_i(t+1) + \eta_i(t), \quad i = 1, \ldots, N-2, \\ x_{N-1}(t+1) = a_{N-1}(t)x_{N-1}(t) + b_{N-1}(t)u_{N-1}(t) \\ \qquad\qquad +g_{N-1}(t)x_1(t) + \omega_{N-1}(t+1) + \eta_{N-1}(t), \\ x_N(t+1) = a_N(t)x_N(t) + b_N(t)u_N(t) + \omega_N(t+1) \\ \qquad\qquad +\eta_N(t) \end{cases}$$

where initial states $x_i(0)$ are i.i.d. r.v.s with normal distribution $\mathscr{N}(4, 4)$. Take the time-varying parameters as follows: when $t > 0$ is even, $a_i(t) = 2 + -\delta/t$, $b_i(t) = 2 + -\delta/t$, $g_i(t) = 1 + -\delta/t$, $i = 1, \ldots, N$; when $t > 0$ is odd, $a_i(t) = 2 + ((t-1)\delta)/t$, $b_i(t) = 2 + ((t-1)\delta)/t$, $i = 1, \ldots, N$, $g_j(t) = 0$, $j = t \bmod (N-2) + 1$, $g_i(t) = 1 + ((t-1)\delta)/t$, otherwise. $\{\omega_i(t),\ t \geq 0\}$ is a sequence of Gaussian white noise, with distribution $\mathscr{N}(2, \sigma_\omega^2)$. $\eta_i(t)$ satisfies B5), $\varepsilon_i(t) \equiv \varepsilon$, $\gamma_i = 0.1$. Thus, the above parameters meet conditions B1)–B6). $R_i$ is taken as 6, $b_{i1}$ is taken as 1. The nonlinear coupling function in indexes (33) is $\Phi(x) = 5\sin(x) + 10$.

Apply estimation algorithm (37) to estimate the time-varying parameters, and the adaptive control is given by (40) and (16). Let $\sigma_\omega = 0.005$, $\varepsilon = 0.005$, $\delta = 0.005$, $N = 100$. Then, the trajectories of the closed-loop system are shown in Fig. 3; the estimate of $f(t)$ for the PSA term is shown in Fig. 4. We can see that, when $\sigma_\omega$, $\varepsilon$, $\delta$ are small, the designed adaptive

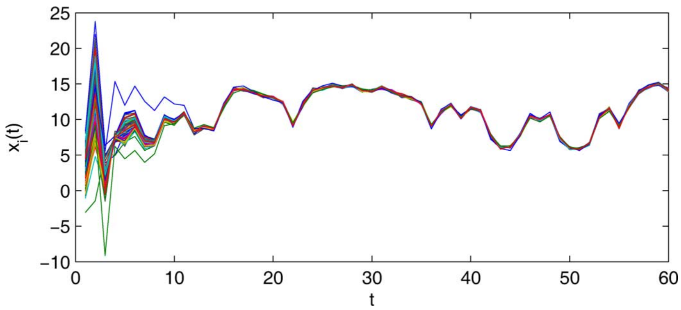


Fig. 3. State trajectories of the agents, when $N = 100$.

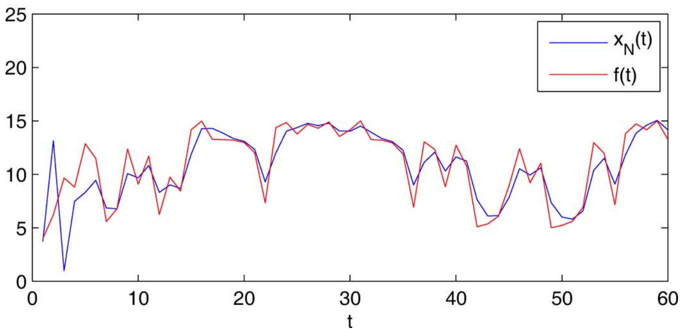


Fig. 4. Curves of $f(t)$ and $\overline{x}_N(t)$, when $N = 100$.

control works well in driving the agents to track the unknown coupling function $\Phi(\overline{x}_N(t))$. This is consistent with the result of Section IV.

It is obvious that agent $N$ is not connected to the other agents, thus, the connectivity of the network topology does not hold. However, from the simulation it can be seen that the global tracking behavior can still be established. This demonstrates the effectiveness of the individual-mass game methodology even under topologies of low connection characteristics.

## VI. Conclusion Remarks

Distributed adaptive tracking-type games for coupled stochastic linear MASs with various uncertainties is investigated. We first considered the time-invariant case subject to the uncertainties such as unknown structure parameters, external disturbances and unknown interactions among agents, and then, the case subject to unmodeled dynamics and unknown time-varying parameters, in addition to the external disturbances and unknown interactions among agents. For both of these two cases, we use ELS/LMS to estimate the unknown parameters, the NCE principle to estimate the unknown PSA term, and the CE principle to design adaptive control. Under some mild conditions, we proved that in the case without unmodeled dynamics, the closed-loop system is almost surely uniformly stable w.r.t. the population number $N$ and the performance indices converge to their optimal values; in the case with unmodeled dynamics, the closed-loop system is stable, and the adaptive control are robust Nash equilibrium w.r.t. the unmodeled dynamics, the parameter variation and the external disturbances.

For the adaptive control of MASs, many important issues are still open and worth investigating, such as the adaptive LQG games and adaptive controls in the case where communication channel is with noise, package loss, quantization error, etc.

## Appendix A
## Proof of Lemma 3.3

The proof idea of Lemma 3.3 comes from [4]. Before proceeding, we need the following two lemmas.

*Lemma A.1 [43]:* Under the assumption A1), for each agent $i$ $(i = 1, \ldots, N)$, we have $\sum_{k=1}^{t} \|x_i(k)\|^2 \to \infty$ a.s. and $\liminf_{t\to\infty} r_i(t)/t \geq tr(R_\sigma) > 0$ a.s., where $r_i(t)$ is defined by (12).

*Lemma A.2 [43]:* Under the notations and assumptions of Lemma 3.3, there is a positive stochastic sequence $\{L_t\}$ such that

$$\|x_i(t)\|^2 \leq L_t, \quad \forall t \geq 0, \quad i = 1, \ldots, N \tag{A1}$$
$$L_{t+1} \leq (\lambda + C\xi(t))\, L_t + C\zeta(t) \tag{A2}$$

where $\lambda \in (0, 1)$, $C > 0$, and

$$\xi(t) = \sum_{i=1}^{N} (\alpha_i^2(t)\delta_i^2(t) \log^2 r(t-1) + \alpha_i(t)\delta_i(t) + \|\triangle\hat{B}_{i1}(t)\|^2)$$

and

$$\zeta(t) = Nd(t)\log^4 r(t) + \log^5 r(t) d(t) = \max\{d_1(t), \ldots, d_N(t)\}.$$

*Proof of Lemma 3.3:* By Lemma A.2,we have

$$L_{t+1} \leq \lambda^{t+1}\left[\prod_{k=0}^{t}(1 + \lambda^{-1}C\xi(k))\right] L_0 + \sum_{k=0}^{t} \lambda^{t-k} C \left[\prod_{j=k+1}^{t} (1+\lambda^{-1}C\xi(j))\right] \zeta(k). \tag{A3}$$

From the third equation of (13), for any $\epsilon > 0$, there exists a constant $\delta > 0$ which is dependent on $N$, such that $\delta \sum_{k=0}^{t} \alpha_i(k) \leq (\epsilon/N) \log r_i(t)$ a.s. $\forall t \geq 0$. Since $\sum_{k=0}^{\infty} \delta_i(k) = \sum_{k=0}^{\infty}(trP_i(k) - trP_i(k+1)) \leq trP_i(0) < \infty$, we have $\delta_i(k) \xrightarrow[k\to\infty]{} 0$, and there exists sufficiently large $k_0 > 0$, which may depend on $N$, such that $(4/\delta)\,(C/\lambda)^{1/2} \sum_{j=k}^{\infty} \delta_i(j) \leq \epsilon/N, \quad \forall k \geq k_0$. This implies that, for any $t \geq k \geq k_0$

$$\begin{aligned}
&\prod_{j=k+1}^{t}\left(1 + \lambda^{-1}C\left(\sum_{i=1}^{N}\alpha_i^2(j)\delta_i^2(j)\right)\log^2 r(j-1)\right) \\
&\leq \prod_{j=k+1}^{t}\left[1+\left(\frac{\delta\alpha_1(j)}{2}\right)^2\right]\left[1+\lambda^{-1}C\left(\frac{2}{\delta}\delta_1(t)\log r(j-1)\right)^2\right] \\
&\cdots \prod_{j=k+1}^{t}\left[1+\left(\frac{\delta\alpha_N(j)}{2}\right)^2\right]\left[1+\lambda^{-1}C\left(\frac{2}{\delta}\delta_N(t)\log r(j-1)\right)^2\right] \\
&\leq r^{\epsilon}(t)\cdot\exp\{(\log r(t))\epsilon\} = r^{2\epsilon}(t), \quad \text{a.s.}
\end{aligned} \tag{A4}$$

In addition, by $1 + x \leq e^x$, $\forall x \geq 0$, we have

$$\prod_{j=k+1}^{t}\left(1 + \lambda^{-1}C\sum_{i=1}^{N}\alpha_i(j)\delta_i(j)\right) = O(r^{\epsilon}(t)) \text{ a.s. } \forall t \geq k \geq k_0. \tag{A5}$$

Since $\|\triangle\hat{B}_{i1}(t)\| \to 0$, there exists a constant $k_1 > 0$ such that $\sup_{j\geq k_1}\left\{1+(C/\lambda)\|\triangle\hat{B}_{i1}(j)\|^2\right\} < (2-\lambda)^{1/N}$

$$\prod_{j=k+1}^{t}\left(1+\lambda^{-1}C\sum_{i=1}^{N}\|\triangle\hat{B}_{i1}(j)\|^2\right) \leq (2-\lambda)^{t-k}, \quad \forall t\geq k\geq k_1. \quad \text{(A6)}$$

Thus, by (A4), (A5), (A6), and the definition of $\xi(t)$ we have

$$\prod_{j=k+1}^{t}(1+C\lambda^{-1}\xi(j)) = O(r^{3\epsilon}(t)(2-\lambda)^{t-k}) \quad \text{a.s.} \quad \forall t\geq k\geq \max\{k_0,k_1\}.$$

Substituting this into (A3) and from the arbitrariness of $\epsilon$, we obtain

$$L_{t+1} = O\left(r^{\epsilon}(t)d(t)\right), \quad \forall \epsilon > 0. \quad \text{(A7)}$$

By Lemma A.1, $t = O(r_i(t)) = O(r(t))$. Hence, $d(t) = O(t^{\delta}) = O(r^{\delta}(t))$, $\forall\delta\in(2/\beta,1)$. This together with (A7) and the arbitrariness of $\epsilon$ leads to $\|\varphi_i(t)\|^2 = O(r^{\delta}(t))$, $\forall\delta\in(2/\beta,1)$. Therefore, by the third equation of (13) we have

$$\sum_{k=0}^{t}\|\tilde{\theta}_i^T(k)\varphi_i(k)\|^2 = O\left(r^{\delta}(t)\right), \quad \forall\delta\in\left(\frac{2}{\beta},1\right) \quad \text{(A8)}$$

which together with A1), (13), and (21) implies

$$\begin{aligned}\sum_{k=0}^{t}\|x_i(k)\|^2 &= O\left(r^{\delta}(t)\right)+O(t)+o\left(\sum_{k=0}^{t}\|u_i^*(k)\|^2\right)\\ \sum_{k=0}^{t}\|\hat{w}_i\|^2 &= 2\sum_{k=0}^{t}\left[\|\hat{w}_i-w_i\|^2+\|w_i\|^2\right]\\ &= O(\log r(t))+O(t),\\ \sum_{k=0}^{t}\|u_i^*(k)\|^2 &= O\left(\sum_{k=0}^{t+1}\|x_i(k)\|^2\right)+O\left(\sum_{k=0}^{t+1}\|w_i(k)\|^2\right)\\ &\quad+O\left(\sum_{k=0}^{t}\|\overline{X}_i(k)\|^2\right)\end{aligned} \quad \text{(A9)}$$

and hence

$$\begin{aligned}&\sum_{i=1}^{N}\sum_{k=0}^{t}\|x_i(t)\|^2 = o(r(t))+O(t),\\ &\sum_{i=1}^{N}\sum_{k=0}^{t}\|u_i^*(k)\|^2 = o(r(t))+O(t), \ \forall\delta\in\left(\frac{2}{\beta},1\right).\end{aligned} \quad \text{(A10)}$$

Here the $O(\cdot)$ and $o(\cdot)$ are dependent on $N$. By (A9) and (A10) we have $r(t) = \max\{r_i(t), 1\leq i\leq N\} = o(r(t))+O(t)$. Thus, (22) is true. This together with (A8) and (A10) gives (23). □

## Appendix B
## Proof of Lemma 4.2

In order to prove Lemma 4.2, we need the following lemma.

*Lemma B.1:* Assume B1), B3), and B5) hold, the nonlinear iterative sequence $\{f(k)\}$ is bounded by $M$ (i.e., $\|f(k)\|\leq M$, $\forall k\geq 0$). If $\bar{\lambda}\in(\max\{\gamma_i,\rho_i,i=1,\ldots,N\},1)$, then under the adaptive control (40), (37), and (16), the sequence $\{\overline{L}_t\}$ defined in (43) satisfies (44), and we have the following estimation:

$$\sum_{k=0}^{t}\bar{\lambda}^{t-k}\|\phi_i(k)\|^2 \leq C_{1i}+C_{2i}\overline{L}_t, \quad \forall t\geq 0, \ i=1,\ldots,N. \quad \text{(B1)}$$

Here, $\rho_i$ and $\gamma_i$ are defined in B1) and B5), respectively; $C_{1i}$ and $C_{2i}$ are defined in (45).

*Proof:* The proof can be found in Lemma B.1 of [43]. □

*Proof of Lemma 4.2:* We first analyze the linear time-varying (44), and estimate $\overline{L}_t$.

From Lemma 4.1 we have

$$\begin{aligned}&E\exp\left\{\frac{2}{\bar{\lambda}}\sum_{i=1}^{N}\sum_{j=k+1}^{t}\Delta_i(j)\right\}\\ &\leq E\exp\left\{\sum_{i=1}^{N}\left[\frac{2}{\bar{\lambda}}\left(\frac{38C_{2i}}{1-\gamma_i}\varepsilon_i+\frac{8C_{2i}}{1-\gamma_i}\delta_i^2+\frac{32\sqrt{m}C_{2i}R_i}{1-\gamma_i}\delta_i\right)(t-k)\right.\right.\\ &\quad\left.\left.+\frac{48C_{2i}}{\bar{\lambda}(1-\gamma_i)d_i\varepsilon_{i0}}\sum_{j=k+1}^{t}\varepsilon_{i0}\|\omega_i(j)\|^2+C_{3i}\right]\right\},\end{aligned} \quad \text{(B2)}$$

where $C_{3i}$ is a constant dependent on $i$.

We now analyze the term

$$E\exp\left\{\sum_{i=1}^{N}\sum_{j=k+1}^{t}\frac{48C_{2i}}{\bar{\lambda}(1-\gamma_i)d_i\varepsilon_{i0}}\sum_{j=k+1}^{t}\varepsilon_{i0}\|\omega_i(j)\|^2\right\}.$$

By using the Hölder inequality $E\prod_{i=1}^{N}X_i \leq \prod_{i=1}^{N}\left(E|X_i|^N\right)^{1/N}$ we have

$$\begin{aligned}&E\exp\left\{\sum_{i=1}^{N}\sum_{j=k+1}^{t}\frac{48C_{2i}}{\bar{\lambda}(1-\gamma_i)d_i\varepsilon_{i0}}\sum_{j=k+1}^{t}\varepsilon_{i0}\|\omega_i(j)\|^2\right\}\\ &\leq\prod_{i=1}^{N}\left(E\exp\left\{\frac{48C_{2i}N}{\bar{\lambda}(1-\gamma_i)d_i\varepsilon_{i0}}\sum_{j=k+1}^{t}\varepsilon_{i0}\|\omega_i(j)\|^2\right\}\right)^{1/N}.\end{aligned} \quad \text{(B3)}$$

Notice that $\bar{\lambda}(1-\gamma_i)\varepsilon_{i0}d_i \geq 48C_{2i}N$ implies that $(48C_{2i}N/\bar{\lambda}(1-\gamma_i)d_i\varepsilon_{i0}) \in (0,1]$. Then, by Hölder inequality and the conditions $\bar{\lambda}(1-\gamma_i)\varepsilon_{i0}d_i\geq 48C_{2i}N$ and B4) we arrive at

$$\begin{aligned}&E\exp\left\{\frac{48C_{2i}N}{\bar{\lambda}(1-\gamma_i)d_i\varepsilon_{i0}}\sum_{j=k+1}^{t}\varepsilon_{i0}\|\omega_i(j)\|^2\right\}\\ &\leq\left(E\exp\left\{\sum_{j=k+1}^{t}\varepsilon_{i0}\|\omega_i(j)\|^2\right\}\right)^{48C_{2i}N/\bar{\lambda}(1-\gamma_i)d_i\varepsilon_{i0}}\\ &\leq\exp\left\{\frac{48C_{2i}N}{\bar{\lambda}(1-\gamma_i)d_i\varepsilon_{i0}}(c_{i1}(t-k)+c_{i2})\right\}.\end{aligned}$$

This together with (B3) gives

$$E\exp\left\{\sum_{i=1}^{N}\sum_{j=k+1}^{t}\frac{48C_{2i}}{\bar\lambda(1-\gamma_i)d_i\varepsilon_{i0}}\varepsilon_{i0}\|\omega_i(j)\|^2\right\}\le\exp\left\{\sum_{i=1}^{N}\frac{48c_{i1}C_{2i}}{\bar\lambda(1-\gamma_i)d_i\varepsilon_{i0}}(t-k)+\sum_{i=1}^{N}\frac{48c_{i2}C_{2i}}{\bar\lambda(1-\gamma_i)d_i\varepsilon_{i0}}\right\}. \quad (B4)$$

Therefore, by (B2) we have

$$E\exp\left\{\frac{2}{\bar\lambda}\sum_{i=1}^{N}\sum_{j=k+1}^{t}\Delta_i(j)\right\}\le\exp\left\{\sum_{i=1}^{N}\left[\frac{2}{\bar\lambda}\left(\frac{38C_{2i}}{1-\gamma_i}\varepsilon_i+\frac{8C_{2i}}{1-\gamma_i}\delta_i^2+\frac{32\sqrt{m}C_{2i}R_i}{1-\gamma_i}\delta_i\right)(t-k)+C_{3i}\right]+\sum_{i=1}^{N}\frac{48c_{i1}C_{2i}}{\bar\lambda(1-\gamma_i)d_i\varepsilon_{i0}}(t-k)+\sum_{i=1}^{N}\frac{48c_{i2}C_{2i}}{\bar\lambda(1-\gamma_i)d_i\varepsilon_{i0}}\right\}=\exp\{2c_1(t-k)+2c_2\} \quad (B5)$$

where

$$c_1=\frac{1}{\bar\lambda}\cdot\sum_{i=1}^{N}\left(\frac{38C_{2i}}{1-\gamma_i}\varepsilon_i+\frac{8C_{2i}}{1-\gamma_i}\delta_i^2+\frac{32\sqrt{m}C_{2i}R_i}{1-\gamma_i}\delta_i+\frac{24c_{i1}C_{2i}}{(1-\gamma_i)d_i\varepsilon_{i0}}\right)$$

$c_2=C_{3i}/2+\sum_{i=1}^{N}24c_{i2}C_{2i}/\bar\lambda(1-\gamma_i)d_i\varepsilon_{i0}$.

Noticing that $1+x\le e^x$, $\forall x>0$, we have

$$E\left(\prod_{j=k+1}^{t}\left[\bar\lambda+\sum_{i=1}^{N}\Delta_i(j)\right]\right)^2\le\bar\lambda^{2(t-k)}E\left(\prod_{j=k+1}^{t}\left[1+\frac{1}{\bar\lambda}\sum_{i=1}^{N}\Delta_i(j)\right]\right)^2\le\bar\lambda^{2(t-k)}E\exp\left\{\frac{2}{\bar\lambda}\sum_{j=k+1}^{t}\sum_{i=1}^{N}\Delta_i(j)\right\}$$

which together with (B5) gives

$$\left(E\left(\prod_{j=k+1}^{t}\left[\bar\lambda+\sum_{i=1}^{N}\Delta_i(j)\right]\right)^2\right)^{1/2}\le\bar\lambda^{t-k}\exp\left\{c_1(t-k)+c_2\right\}=e^{c_2}(\bar\lambda e^{c_1})^{t-k},\quad t\ge k,\ k\ge 0. \quad (B6)$$

By B4) and the elementary inequality $x\le\exp(ae^{-1}x^{1/a})$, $\forall a>0,\ x>0$, we have $E\Big((e/2)\varepsilon_{i0}\|\omega_i(t)\|^2\Big)^2\le E\exp(\varepsilon_{i0}\|\omega_i(t)\|^2)\le\exp(c_{i1}+c_{i2})$. This together with (B6) and Lemma B.1 renders

$$E\|\overline{L}_t\|=O\left(\sum_{k=0}^{t}\left(E\left(\prod_{j=k+1}^{t}[\bar\lambda+\sum_{i=1}^{N}\Delta_i(j)]\right)^2\right)^{1/2}\cdot\left(\sum_{i=1}^{N}(E\|\omega_i(k)\|^4)^{1/2}+1+d\right)\right)=O\left(\sum_{k=0}^{t}(\bar\lambda e^{c_1})^{t-k}\left[1+d+\sum_{i=1}^{N}\left(\frac{2}{e\varepsilon_{i0}}\exp(2^{-1}(c_{i1}+c_{i2}))\right)\right]\right)=O\left(\sum_{k=0}^{t}(\bar\lambda e^{c_1})^{t-k}(1+d)\right).$$

Here, the $O(\cdot)$ is related to $N$. From the above equation, if $\bar\lambda e^{c_1}\in(0,1)$, we have $E\|\overline{L}_t\|=O(d)$. Noticing (B1), we finally obtain (42). □

ACKNOWLEDGMENT

The authors would like to thank Prof. M. Huang, Dr. T. Li, and Prof. H.-B. Ma for their valuable discussions on this research topic, and the anonymous reviewers for their detailed and insightful comments.

REFERENCES

[1] Q. Zhang and J. Zhang, "Adaptive tracking-type games for coupled large population ARMAX systems," in *Proc. 8th IEEE Int. Conf. Control Autom.*, 2010, pp. 148–153.
[2] Q. Zhang and J. Zhang, "Robust adaptive control of coupled stochastic multi-agent systems with unmodeled dynamics," in *Proc. 8th World Congr. Intell. Control Autom.*, 2010, pp. 586–591.
[3] G. Goodwin and K. Sin, *Adaptive Filtering, Prediction and Control*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
[4] H. Chen and L. Guo, *Identification and Stochastic Adaptive Control*. Boston, MA: Birkhäuser, 1991.
[5] P. Caines, *Linear Stochastic Systems*. Hoboken, NJ: Wiley, 1987.
[6] P. Ioannou and B. Fidan, *Adaptive Control Tutorial*. Philadelphia, PA: SIAM, 2006.
[7] P. Ioannou, "Decentralized adaptive control of interconnected systems," *IEEE Trans. Autom. Control*, vol. 31, no. AC–4, pp. 291–298, Apr. 1986.
[8] D. Gavel and D. Siljak, "Decentralized adaptive control: Structural conditions for stability," *IEEE Trans. Autom. Control*, vol. 34, no. 4, pp. 413–426, Apr. 1989.
[9] C. Wen, "Decentralized adaptive regulation," *IEEE Trans. Autom. Control*, vol. 39, no. 10, pp. 2163–2166, Oct. 1994.
[10] J. T. Spooner and K. M. Passino, "Adaptive control of a class of decentralized nonlinear systems," *IEEE Trans. Autom. Control*, vol. 41, no. 1, pp. 280–284, Jan. 1996.
[11] S. Liu, J. Zhang, and Z. Jiang, "Decentralized adaptive output-feedback stabilization for large-scale stochastic nonlinear systems," *Autom.*, vol. 43, no. 2, pp. 238–251, 2007.
[12] K. Narendra and N. Oleng, "Decentralized adaptive control," in *Proc. Amer. Control Conf.*, 2002, vol. 5, pp. 3407–3412.
[13] K. Narendra and N. Oleng, "Exact output tracking in decentralized adaptive control systems," *IEEE Trans. Autom. Control*, vol. 47, no. 2, pp. 390–395, Feb. 2002.
[14] K. Narendra, N. Oleng, and S. Mukhopadhyay, "Decentralised adaptive control with partial communication," *IEE Proc.-Control Theory and Appl.*, vol. 153, no. 5, pp. 546–555, 2006.
[15] A. Jadbabaie, J. Lin, and A. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 988–1001, Jun. 2003.

[16] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, Sep. 2004.
[17] W. Ren and R. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 655–661, May 2005.
[18] F. Cucker and S. Smale, "Emergent behavior in flocks," *IEEE Trans. Autom. Control*, vol. 52, no. 5, pp. 852–862, May 2007.
[19] Z. Liu and L. Guo, "Synchronization of multi-agent systems without connectivity assumptions," *Autom.*, vol. 45, no. 12, pp. 2744–2753, 2009.
[20] V. Gazi, "Stability of a discrete-time asynchronous swarm with time-dependent communication links," *IEEE Trans. Syst., Man, Cybernet.: Part B*, vol. 38, no. 1, pp. 267–274, Feb. 2008.
[21] A. Samiloglu, V. Gazi, and A. Koku, S. Nolfi, Ed. *et al.*, "Asynchronous cyclic pursuit," in *Proc. SAB06*, 2006, vol. 4095, Lecture notes in Artificial Intelligence (LNAI), pp. 667–678.
[22] A. Samiloglu, V. Gazi, and A. Koku, "Comparison of three orientation agreement strategies in self-propelled particle systems with turn angle restrictions in synchronous and asynchronous settings," *Asian J. Control*, vol. 10, no. 2, pp. 212–232, 2008.
[23] E. Semsar-Kazerooni and K. Khorasani, "Optimal consensus algorithms for cooperative team of agents subject to partial information," *Autom.*, vol. 44, no. 11, pp. 2766–2777, 2008.
[24] M. Huang, P. Caines, and R. Malhamé, "Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized $\varepsilon$-Nash equilibria," *IEEE Trans. Autom. Control*, vol. 52, no. 9, pp. 1560–1571, Sep. 2007.
[25] T. Li and J. Zhang, "Asymptotically optimal decentralized control for large population stochastic multiagent systems," *IEEE Trans. Autom. Control*, vol. 53, no. 7, pp. 1643–1660, Jul. 2008.
[26] C. Ma, T. Li, and J. Zhang, "Linear quadratic decentralized dynamic games for large population discrete-time stochastic multi-agent systems," *J. Syst. Sci. Math. Scies*, vol. 27, no. 3, pp. 464–480, 2007.
[27] T. Li and J. Zhang, "Decentralized tracking-type games for multi-agent systems with coupled ARX models: Asymptotic Nash equilibria," *Autom.*, vol. 44, no. 3, pp. 713–725, 2008.
[28] Y. Cao and W. Ren, "Distributed coordinated tracking with reduced interaction via a variable structure approach," *IEEE Trans. Autom. Control*, vol. 57, no. 1, pp. 33–48, Jan. 2012.
[29] M. Huang, R. Malhamé, and P. Caines, "Nash strategies and adaptation for decentralized games involving weakly-coupled agents," in *Proc. 44th IEEE Conf. Decision and Control and 2005 Eur. Control Conf.*, 2005, pp. 1050–1055.
[30] P. E. Caines, "Mean field stochastic control," in *Bode Lecture of the 48th IEEE Conf. Decision and Control*, Shanghai, China, Dec. 18, 2009.
[31] H. Ma, K. Lum, and S. Ge, "DecentralizedÅström-Wittenmark self-tuning regulator of a multi-agent uncertain coupled ARMAX system," in *Proc. 2007 IEEE Multi-Conf. Syst. Control*, Singapore, 2007, pp. 363–368.
[32] Z. Hou, L. Cheng, and M. Tan, "Decentralized robust adaptive control for the multiagent system consensus problem using neural networks," *IEEE Trans. Syst., Man, Cybernet. B: Cybernet.*, vol. 39, no. 3, pp. 636–647, Jun. 2009.
[33] E. Gul and V. Gazi, "Adaptive internal model based formation control of a class of multi-agent systems," in *Proc. Amer. Control Conf.*, Baltimore, MD, USA, Jun.-Jul. 2010, pp. 4800–4805.
[34] S. Duran and V. Gazi, "Adaptive formation control and target tracking in a class of multi-agent systems," in *Proc. Amer. Control Conf.*, Baltimore, MD, USA, Jun.-Jul. 2010, pp. 75–80.
[35] A. Mikhailov and D. Zanette, "Noise-induced breakdown of coherent collective motion in swarms," *Phys. Rev. E*, vol. 60, no. 4, pp. 4571–4575, 1999.
[36] U. Erdmann, W. Ebeling, and A. Mikhailov, "Noise-induced transition from translational to rotational motion of swarms," *Phys. Rev. E*, vol. 71, no. 5, pp. 1–7, 2005.
[37] M. Huang, P. Caines, and R. Malhamé, "Uplink power adjustment in wireless communication systems: A stochastic control analysis," *IEEE Trans. Autom. Control*, vol. 49, no. 10, pp. 1693–1708, Oct. 2004.
[38] V. Lambson, "Self-enforcing collusion in large dynamic markets," *J. Econ. Theory*, vol. 34, no. 2, pp. 282–291, 1984.
[39] L. Guo and H. Chen, "The Åström-Wittenmark self-tuning regulator revistied and ELS-based adaptive trackers," *IEEE Trans. Autom. Control*, vol. 36, no. 7, pp. 802–812, Jul. 1991.
[40] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.
[41] Y. Chow and H. Teicher, *Probability Theory: Independence, Interchangeability, Martingales*. New York: Springer-Verlag, 1978.
[42] L. Guo, "On adaptive stabilization of time-varying stochastic systems," *SIAM J. Control Optim.*, vol. 28, pp. 1432–1451, 1990.
[43] Q. Zhang, "Distributed estimation and control of multi-agent systems in uncertain environment," Ph.D. dissertation, Chinese Acad. Sci., Beijing, China, 2012.

**Qiang Zhang** was born in Shandong, China, in 1985. He received the B.S. degree in mathematics from Shandong University, Shandong, China, in 2007, and the Ph.D. degree from the Institute of Systems Science (ISS), Academy of Mathematics and Systems Science (AMSS), Chinese Academy of Sciences (CAS), Beijing, China, in 2012.

He is now with the International Cooperation Center for Economics and Technology, the Ministry of Industry and Information Technology (MIIT), Beijing. His current research interests include complex system, economics, and management decision.

**Ji-Feng Zhang** (M'92–SM'97) received the B.S. degree in mathematics from Shandong University, Shandong, China, in 1985 and the Ph.D. degree from the Institute of Systems Science (ISS), Chinese Academy of Sciences (CAS), Beijing, China, in 1991.

Since 1985, he has been with the ISS, CAS, and now is the Director of the ISS. He is a Managing Editor of the *Journal of Systems Science and Complexity*, Deputy Editor-in-Chief of the *Journal of Systems Science and Mathematical Sciences*, the journal *Systems Engineering—Theory and Practice*, and the journal *Control Theory and Applications*. He is also an Associate Editor of several other journals, including the *SIAM Journal on Control and Optimization* and was a Deputy Editor-in-Chief of the *Journal of Automation*. His current research interests include system modeling and identification, adaptive control, stochastic systems, and multi-agent systems.

Dr. Zhang received the Second Prize of the National Natural Science Award of China in 2010, the Distinguished Young Scholar Fund from the National Natural Science Foundation of China in 1997, the First Prize of the Young Scientist Award of CAS in 1995, the Outstanding Advisor Award of CAS in 2007, 2008, and 2009, respectively. He was an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, the Program Chair of the 2012 IEEE Conference on Control Applications, the 9th World Congress on Intelligent Control and Automation, and the 30th Chinese Control Conference, and the General Co-Chair of the 32nd and 33rd Chinese Control Conference. He serves as a member of the Board of Governors of the IEEE Control Systems Society, a Vice President of the Systems Engineering Society of China, and the Director of the Technical Committee on Control Theory, Chinese Association of Automation.