This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2023.3237781

Online Distributed Learning for Aggregative Games with Feedback Delays

Pin Liu, Kaihong Lu, Feng Xiao, *Member, IEEE,* Bo Wei, and Yuanshi Zheng, *Senior Member, IEEE,*

***Abstract*— This paper proposes a new aggregative game (AG) model with feedback delays. The strategies of players are selected from given strategy sets and subject to global nonlinear inequality constraints. Both cost functions and constrained functions of players are time-varying, which reflects the changing nature of environments. At each time, each player only has access to its strategy set information, and the information of its current cost function and current constrained function is unknown. Due to feedback delays, the feedback information of corresponding cost functions and constrained functions is not revealed to players immediately after the selection of strategies. It would take a period of time for players to observe their feedback information. To address such an AG problem, a distributed learning algorithm is proposed with the local information from their neighbors and the delayed feedback information from environments, and it is applicable to time-varying weighted digraphs. We find that the two metrics of the algorithm grow sub-linearly with respect to the learning time. A simulation example is given to illustrate the performance of the proposed algorithm.**

***Index Terms*— Aggregative games, dynamic environments, feedback delays, generalized Nash equilibrium.**

I. INTRODUCTION

Game theory has been receiving increasing attention in the control community [1]–[4], mostly inspired by its various applications in multi-agent systems, such as smart grids and electric vehicles [5]–[8]. Generalized Nash equilibrium (GNE) seeking problems is one of the main problems in game theory [9]–[11]. With the development of the Cournot model, some literature focuses on AG problems where the cost functions of players depend on their strategies and the aggregation of the strategies of all players [12]–[17]. For example, to defend attacks, targets are surrounded by several robots whose cost functions depend on their own positions and the center of the positions of all robots. In games on networks, players only obtain the information from their neighbors [12], [13], [18], [19].

This work was supported by the National Natural Science Foundation of China (NSFC, Grant Nos. 62273145, 61903140, 62103169, 62250056).
Corresponding author: Feng Xiao.

Pin Liu and Feng Xiao are with the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources and the School of Control and Computer Engineering, Beijing 102206, China (e-mails: pin_liu@yeah.net; fengxiao@ncepu.edu.cn).

Kaihong Lu is with the College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao 266590, China (e-mail: khong_lu@163.com).

Bo Wei is with the School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China (e-mail: bowei@ncepu.edu.cn).

Yuanshi Zheng is with Shaanxi Key Laboratory of Space Solar Power Station System, School of Mechano-electronic Engineering, Xidian University, Xi'an 710071, China (e-mail:zhengyuanshi2005@163.com).

The main feature of the aforementioned work is that the cost functions of all players are time-invariant. However, environments are changing and uncertain. In such dynamic environments, cost functions are time-varying and are only available to players after the selection of strategies. The algorithms for seeking GNE of such games are called online or learning algorithms [20]. In [21], online algorithms were proposed to seek the GNE of noncooperative games in the presence of local strategy set constraints and time-invariant coupled constraints. Due to dynamic environments, coupled constraints are also time-varying and revealed to players after the selection of strategies. For example, in smart grids, the total energy supply is hard to be predicted due to the uncertainty of renewable energy such as solar and wind power [22]. In this work, we assume that players select their strategies from prescribed compact sets. The compactness of prescribed strategy sets is not a strong assumption for game problems [19], [21]. Lots of strategy constraints in practical applications, e.g., constrained rendezvous of unmanned aerial vehicles within a prescribed safety area and limited power generation of generators, can be formulated as compact strategy set constraints.

In the literature, see [21] and [23], online algorithms for noncooperative games in dynamic environments were proposed in the scenarios where the information of cost functions and constrained functions is revealed to players immediately after the selection of strategies. However, due to the computing overhead or inherent lag of observing the information, there exist feedback delays from the choice of a strategy to the observation of the corresponding feedback information from environments [24]. For example, in advertisement placement problems, there exist lags between the placement of an advertisement and its conversion. In game problems, feedback delays would affect the sub-linearity of metrics which are utilized to illustrate the performance of algorithms. The main difficulty in analyzing feedback delays is to quantify the impact of feedback delays.

In this paper, AG problems in dynamic environments are investigated. Different from [21] and [23] where noncooperative game problems in dynamic environments were addressed, we focus on AG problems. In addition, our studied model has some new features, such as feedback delays and time-varying digraphs. The main contributions are summarized as follows.

1) This work investigates AG problems in dynamic environments where both cost functions and constrained functions are time-varying. We propose a new learning algorithm in a distributed setting and apply it to a time-varying digraph.
2) Since, for each player, it takes a period of time from the selection of its strategies to the observation of the

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2023.3237781

corresponding feedback information from environments, feedback delays are investigated. The effects of feedback delays on the sub-linearity of metrics are quantified. In addition, the upper bounds of feedback delays are analyzed to guarantee the sub-linearity of metrics.

3) Time-varying learning rates effectively guarantee the proposed algorithms with non-zero learning rates. Existing works on feedback delay problems only focus on fixed learning rates with the learning time as the denominator [25]. When the learning time approaches to infinite, fixed learning rates would approach to zero.

The rest of this paper is organized as follows. In Section II, AG problems are formulated. In Section III, a distributed learning algorithm is proposed. Then, a simulation example is given in Section IV. Section V concludes this paper.

Notations. $\mathbb{R}(\mathbb{R}_+)$ denotes the set of (positive) real numbers. $\mathrm{col}(x_1, \cdots, x_N)$ denotes $[x_1^T, \cdots, x_N^T]^T$, where $x_1, \cdots, x_N$ are vectors. $\mathcal{O}(x)$ is a function that is linear with respect to x. $[z]_+$ denotes the projection of $z \in \mathbb{R}$ onto $\mathbb{R}_+$. $P_\Omega(x)$ denotes the projection of a vector $x \in \mathbb{R}^n$ onto set $\Omega \subset \mathbb{R}^n$.

II. Problem Formulation

We first introduce some basic concepts of graph theory [26]. Consider a time-varying digraph $\mathcal{G}_t = (\mathcal{V}, \mathcal{E}_t, \mathcal{A}_t)$ with N nodes. In consensus problems or distributed optimization problems, nodes represent agents [27]–[29]. In game problems, nodes represent players [12], [30]. $\mathcal{V} = \{1, ..., N\}$ denotes the node set. $\mathcal{E}_t$ denotes the edge set. If $(j, i) \in \mathcal{E}_t$, then node j is an in-neighbor of node i at t, which means that node i can receive the information from node j. a_{ij}^t denotes the weight that node i allocates to the edge $(j, i) \in \mathcal{E}_t$ at t. Matrix $\mathcal{A}_t = (a_{ij}^t)_{N\times N}$ denotes the adjacency matrix with $a_{ij}^t > 0$ if $(j, i) \in \mathcal{E}_t$, and $a_{ij}^t = 0$, otherwise. If there exists a directed path between any pair of nodes, then the digraph is strongly connected.

A. AG with and without feedback delays

Consider the AG problem consisting of N players in the time-varying digraph $\mathcal{G}_t$. Each node in digraph $\mathcal{G}_t$ represents a player. For $i \in \mathcal{V}$, denote the strategy of player i by $x_i \in \Omega_i$, where $\Omega_i \subset \mathbb{R}^{n_i}$ is the strategy set of player i, n_i denotes the strategy dimension of player i. $x_{-i} = \mathrm{col}(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_N)$. $x = (x_i, x_{-i}) \in \Omega \subset \mathbb{R}^n$, where $\Omega = \prod_{i\in\mathcal{V}} \Omega_i$ is the strategy set of all players and $n = \sum_{i=1}^N n_i$. In dynamic environments, at each time t, the private cost function $J_{i,t}(x_i, x_{-i}) : \mathbb{R}^{n_i} \times \mathbb{R}^{n-n_i} \to \mathbb{R}$ and the private constrained function $g_{i,t}(x_i) : \mathbb{R}^{n_i} \to \mathbb{R}$ of player i is unknown in advance and is only gradually revealed after that player i selects its strategy $x_{i,t}$ from Ω_i at t. Denote such an AG problem as $\mathcal{G}_N(\mathcal{G}_t, J_{i,t}, g_{i,t})$.

In $\mathcal{G}_N(\mathcal{G}_t, J_{i,t}, g_{i,t})$, the cost function of player i depends on its strategy x_i and the aggregation of the strategies of all players. For $i \in \mathcal{V}$, the specific cost function is given by $J_{i,t}(x_i, x_{-i}) = h_{i,t}(x_i, s(x))$, where the aggregation of the strategies of all players is given by $s(x) = \frac{1}{N}\sum_{j=1}^N \varphi_j(x_j)$ and $\varphi_i(x_i) : \mathbb{R}^{n_i} \to \mathbb{R}^m$ is the continuous nonlinear mapping of player i. The time-varying global inequality constraint is specified by $g_t(x) = \sum_{j=1}^N g_{j,t}(x_j) \le 0$. In game $\mathcal{G}_N(\mathcal{G}_t, J_{i,t}, g_{i,t})$, player i aims to minimize its own cost function at each time, i.e.,

$$\begin{aligned} &\min_{x_i} J_{i,t}(x_i, x_{-i}), \\ &s.t. \quad x_i \in \Omega_i, \quad \sum_{j=1}^N g_{j,t}(x_j) \le 0. \end{aligned} \tag{1}$$

Consider a practical example of the resource sharing in a network with N prosumers (players) [31]. Denote x_i as the volume of trade of prosumer i, where $i \in \{1, \cdots, N\}$. If $x_i > 0$, player i is a producer and injects resources into the network. If $x_i < 0$, player i is a consumer and buys resources from the network. Each player aims to solve the problem (1). The cost function of player i is given by $J_{i,t}(x_i, x_{-i}) = p_{i,t}(x_i) - q_t(\sum_{i=1}^N x_i)x_i$, where $p_{i,t}(\cdot)$ is the generation cost function of producer i or the disutility of consumer i. $q_t(\cdot)$ is the market price function. Due to dynamic environments, e.g., market policy adjustment, the price or the generation cost are time-varying, which implies that $J_{i,t}$ is time-varying. For player i, the volume of trade is chosen from a prescribed set $\Omega_i = [\underline{x}_i, \bar{x}_i]$. The injection into the network or the buy from the network results in the fluctuation of the network. To guarantee the stability of network, network constraint $\sum_{i=1}^N g_{i,t}(x_i) = Bx - b_t \le 0$ is needed, where $b_t \in \mathbb{R}$, $x = \mathrm{col}(x_1, \cdots, x_N)$ and B is a matrix. The controller of each prosumer adjusts it volume of trade to balance the network. As long as the injection or buy can be balanced in a long run, we say that the algorithm performs well.

Denote $x_{i,t}$ as the strategy of player i at time t. In $\mathcal{G}_N(\mathcal{G}_t, J_{i,t}, g_{i,t})$ without feedback delays, player i observes the feedback information of $J_{i,t}$ and $g_{i,t}$ immediately after the selection of $x_{i,t}$. Then, the observed information is utilized to obtain the next strategy $x_{i,t+1}$. However, in practical applications, the observation of the feedback information of $J_{i,t}$ and $g_{i,t}$ from environments may be delayed. In this paper, the AG problem with feedback delays is investigated. Consider the feedback delay τ satisfying $\tau < T$, where T is the learning time. At the time t, player i updates its strategy $x_{i,t}$. After player i selects its strategy $x_{i,t+\tau}$ at $t + \tau$, the information of $J_{i,t}$ and $g_{i,t}$ is first observed and utilized to update strategies after $t + \tau$.

B. Assumptions for the AG model

For digraph $\mathcal{G}_t$, the following assumption is given.

Assumption 1 ([32, Assumption 1]): The time-varying weighted digraph $\mathcal{G}_t = (\mathcal{V}, \mathcal{E}_t, \mathcal{A}_t)$ satisfies that

1) If $(j, i) \in \mathcal{E}_t$, then the weight $a_{ij}^t > a$, where $a \in (0, 1)$;
2) $\mathcal{A}_t$ is doubly stochastic, i.e., $\sum_{i=1}^N a_{ij}^t = 1$ for all $j \in \mathcal{V}$ and $\sum_{j=1}^N a_{ij}^t = 1$ for all $i \in \mathcal{V}$;
3) There exists a constant $Q > 0$ such that the digraph $(\mathcal{V}, \bigcup_{l=0,\cdots,Q-1} \mathcal{E}_{t+l})$ is strongly connected for all $t \ge 1$.

Denote $\chi = \Omega \cap \chi^g$, where $\chi^g = \{x \in \mathbb{R}^n | g_t(x) < 0\}$. For $i \in \mathcal{V}$, the mapping $H_{i,t}$ is defined by

$$H_{i,t}(x_i, z_i) = \left(\nabla_{x_i} h_{i,t}(\cdot, s) + \frac{1}{N}\nabla_s h_{i,t}(x_i, \cdot)\nabla\varphi_i\right)|_{s=z_i},$$

where $\nabla\varphi_i$ is the gradient of φ_i. Denote gradient vector $F_t(x)$ as $F_t(x) = \mathrm{col}(H_{1,t}(x_1, s(x)), \cdots, H_{N,t}(x_N, s(x)))$. For any $i \in \mathcal{V}$ and $t > 0$, some assumptions on strategy sets, cost functions and constrained functions are given as follows [12], [19], [21].

Assumption 2:

1) Ω_i is non-empty, convex and compact. χ is non-empty;
2) $J_{i,t}(x_i, x_{-i})$ is continuously differentiable with respect to x, Lipschtiz with respect to x_i and convex with respect to x_i for every x_{-i};
3) $g_{i,t}(x_i)$ is continuously differentiable and convex with respect to x_i;
4) $\varphi_i(x_i)$ is Lipschitz with respect to x_i.

Assumption 3:

1) $F_t(x)$ is strongly monotone, i.e., there exists $\mu^F > 0$ such that $(x-y)^T(F_t(x)-F_t(y)) \geq \mu^F\|x-y\|^2$ for all $x, y \in \Omega$;
2) $H_{i,t}(x_i, y)$ is Lipschitz with respect to y, i.e., there exists a constant $\mu^H > 0$ such that $\|H_{i,t}(x_i, y) - H_{i,t}(x_i, y')\| \leq \mu^H\|y - y'\|$, for all $x_i \in \Omega_i$, $y, y' \in \mathbb{R}^m$.

In Assumption 2, condition 1) guarantees that there exists $C_x > 0$ such that $\|x_i\| \leq C_x$ for all $x_i \in \Omega_i$ and $i \in \mathcal{V}$. Condition 2) in Assumption 2 guarantees that there exist $L_g > 0$ and $l_g > 0$ such that $|g_{i,t}(x_i)| \leq L_g$ and $\|\nabla_{x_i} g_{i,t}(x_i)\| \leq l_g$ for all $x_i \in \Omega_i$. Since $\varphi_i(x_i)$ is Lipschitz, there exists $\mu^\varphi > 0$ such that $\|\varphi_i(x_i) - \varphi_i(x_i')\| \leq \mu^\varphi\|x_i - x_i'\|$. Under Assumption 2, there exists $\mu^J > 0$ such that $\|J_i(x_i, x_{-i}) - J_i(x_i', x_{-i})\| \leq \mu^J\|x_i - x_i'\|$.

In this paper, we focus on the variational generalized Nash equilibrium (v-GNE) [33, Definition 3.10]. The v-GNE is a subclass of general Nash equilibrium and characterizes the same penalty to fulfill the coupling constraints for each player. By [34, Theorem 2.3.3], Assumptions 2 and 3 guarantee the existence and uniqueness of the v-GNE. $x_{i,t}^*$ is the v-GNE if and only if there exists a bounded Lagrange multiplier $y_t^* \leq C_y$, where $C_y > 0$, such that the following KKT conditions are satisfied [12], [21], [33].

$$
\begin{aligned}
x_{i,t}^* =& P_{\Omega_i}\Big(x_{i,t}^* - \alpha_t H_{i,t}\big(x_{i,t}^*, s(x_t^*)\big) - \alpha_t^2 \nabla g_{i,t}\big(x_{i,t}^*\big) y_t^*\Big),\\
y_t^* =& \Big[y_t^* + \sum_{i=1}^N g_{i,t}(x_{i,t}^*)\Big]_+, \quad i \in \mathcal{V}.
\end{aligned} \tag{2}
$$

C. Two regrets

In $\mathcal{G}_N(\mathcal{G}_t, J_{i,t}, g_{i,t})$, since $J_{i,t}$ and $g_{i,t}$ are revealed to player i only after the selection of $x_{i,t}$, it is impossible to obtain $x_{i,t}^*$ at each time t. Also, the inequality constraints cannot be satisfied at each time. Therefore, specific metrics, dynamic regrets and violations of constraints, are needed to define the performance of learning algorithms. Since player i aims to choose $x_{i,t}^*$ at each time t, the sequence $\{x_{i,t}^*\}_{i\in\mathcal{V}, t\geq 1}$ over the learning time T is utilized as the benchmark sequence. The first metric $\mathcal{R}_i^T$, say the dynamic regret [21], with respect to the benchmark is given by

$$
\mathcal{R}_i^T = \sum_{t=1}^T \big(J_{i,t}(x_{i,t}, x_{-i,t}^*) - J_{i,t}(x_{i,t}^*, x_{-i,t}^*)\big). \tag{3}
$$

Due to the unknown $J_{i,t}$ before choosing $x_{i,t}$, the difference between $J_{i,t}(x_{i,t}, x_{-i,t}^*)$ and $J_{i,t}(x_{i,t}^*, x_{-i,t}^*)$ exists, and thus $\mathcal{R}_i^T$ grows as t. In addition, another metric is needed to capture the violation of global inequality constraints. The second metric $\mathcal{R}_g^T$, say the violation of constraints, is given by

$$
\mathcal{R}_g^T = \Big[\sum_{t=1}^T \sum_{i=1}^N g_{i,t}(x_{i,t})\Big]_+. \tag{4}
$$

Time-varying global inequality constraints only need to be satisfied in a long run instead of at each time t. It is expected that the above-mentioned accumulations $\mathcal{R}_i^T$ and $\mathcal{R}_g^T$ are sub-linear with respect to the learning time T, i.e., $\lim_{T\to\infty} \mathcal{R}_i^T/T \to 0$ and $\lim_{T\to\infty} \mathcal{R}_g^T/T \to 0$.

III. Main Results

In this section, a distributed learning algorithm is proposed. We illustrate that metrics $\mathcal{R}_i^T$ and $\mathcal{R}_g^T$ are sub-linear with respect to the learning time T for every $i \in \mathcal{V}$.

Due to the feedback delay τ, player i cannot obtain the feedback information of $J_{i,t}$ and $g_{i,t}$ for $t \in [1, \cdots, \tau+1]$. Thus, for the first $\tau + 1$, strategies should be set in advance. After the selection of $x_{i,\tau+1}$ at the time $\tau + 1$, the feedback information begins to be observed. For $t \geq \tau + 1$, player i utilizes the information of $J_{i,t-\tau}$ and $g_{i,t-\tau}$ to update the strategies $x_{i,t+1}$ based on the distributed algorithm. Motivated by [25], for every $i \in \mathcal{V}$, we set $x_{i,1} = x_{i,2} = \cdots = x_{i,\tau+1} \in \Omega_i$ and $y_{i,1} = y_{i,2} = \cdots = y_{i,\tau+1} = 0$. In addition, we need to initialize that $\eta_{i,t} = \varphi_i(x_{i,t})$ for the first $\tau+1$. For $t \geq \tau+1$, the learning algorithm is given by

$$
\begin{aligned}
x_{i,t+1} =& (1 - \alpha_{t-\tau})x_{i,t} + \alpha_{t-\tau} P_{\Omega_i}\Big(x_{i,t}\\
& - \alpha_{t-\tau} H_{i,t-\tau}\big(x_{i,t-\tau}, \tilde{\eta}_{i,t-\tau}\big)\\
& - \alpha_{t-\tau}^2 \nabla g_{i,t-\tau}\big(x_{i,t-\tau}\big)\tilde{y}_{i,t}\Big),\\
\eta_{i,t+1} =& \tilde{\eta}_{i,t} + \varphi_i\big(x_{i,t+1}\big) - \varphi_i\big(x_{i,t}\big),\\
y_{i,t+1} =& \Big[\tilde{y}_{i,t} + \alpha_{t-\tau}\Big(g_{i,t-\tau}\big(x_{i,t-\tau}\big) - \alpha_{t-\tau}\tilde{y}_{i,t}\Big)\Big]_+,\\
\tilde{\eta}_{i,t} =& \sum_{j=1}^N a_{ij}^t \eta_{j,t}, \quad \tilde{y}_{i,t} = \sum_{j=1}^N a_{ij}^t y_{j,t},
\end{aligned} \tag{5}
$$

where $\alpha_{t-\tau} = \frac{1}{(t-\tau)^{l_\alpha}}$ is the learning rate. $\eta_{i,t}$ is the estimate of $s(x_t)$. $y_{i,t}$ is the estimate of Lagrange multiplier. $H_{i,t-\tau}$, $\nabla g_{i,t-\tau}$ and $g_{i,t-\tau}$ are delayed feedback information from environments. The updates of $x_{i,t}$ and $y_{i,t}$ are motivated by the primal-dual method. A penalty term $-\alpha_{t-\tau}^2\tilde{y}_{i,t}$ is employed so that the upper bound of $y_{i,t}$ can be obtained. The update of $\eta_{i,t}$ indicates that each player estimates the global information $s(x_t)$ only by using the information from itself and its neighbors. Note that the delayed information of cost functions and constrained functions from environments is employed to update $x_{i,t}$, $\eta_{i,t}$ and $y_{i,t}$.

Remark 1: The algorithm (5) is different from the one given in [21]. $\eta_{i,t}$ is employed to estimate the aggregation of strategies of all players in AGs. Constrained functions $g_{i,t}$

and $\nabla g_{i,t}$ are time-varying. Moreover, the delayed information of $J_{i,t}$ and $g_{i,t}$ is utilized in our algorithm.

Remark 2: Feedback delays were investigated in optimization problems [25]. Different from [25], we consider feedback delays in a distributed scenario. The information from environments is delayed for all players. In our algorithm (5), the learning rate α_t is time-varying. Even when $T \to \infty$, $\alpha_t \neq 0$ for all $t < T$.

It follows from Assumption 2 that there exists $k_1 > 0$ such that $k_1 = \sup_{x_t \in \Omega} \left\| H_{i,t}(x_{i,t}, \tilde{\eta}_{i,t}) \right\|$. In the following, the main result on the sub-linearity of $\mathcal{R}_i^T$ and $\mathcal{R}_g^T$ is given, which reflects the performance of learning algorithm (5).

Theorem 1: Suppose that Assumptions 1, 2 and 3 hold for the AG model $\mathcal{G}_N(\mathcal{G}_t, J_{i,t}, g_{i,t})$ described in (1). If the learning rate is given by $\alpha_t = \frac{1}{t^{l_\alpha}}$, where $0 < l_\alpha < \frac{1}{2}$, then, for every $i \in \mathcal{V}$ and $T \geq 1$, the upper bounds of the dynamic regrets in (3) and the violation of constraints in (4) are given by

$$\begin{aligned}\mathcal{R}_i^T \leq& \mathcal{O}(T^{\frac{1}{2}+l_\alpha}(\sqrt{\Theta^T+1}) + (\sqrt{\tau+1})T^{1-\frac{l_\alpha}{2}}) + \frac{\sqrt{k(\tau)T}}{\mu^J},\\ \mathcal{R}_g^T \leq& \mathcal{O}(T^{\frac{1}{2}+l_\alpha}(\sqrt{\Theta^T+1}) + (\sqrt{\tau+1})T^{1-\frac{l_\alpha}{2}})\\ &+ \sqrt{2N\mu^F k(\tau)T},\end{aligned}$$

where $\Theta^T = \sum_{i=1}^N \sum_{t=1}^T \|x_{i,t+1}^* - x_{i,t}^*\|$ denotes the variation of the Nash equilibrium. $k(\tau) = \big(2N\tau + 2N + \frac{N^2\gamma\theta}{1-\theta}\big)\frac{2C_x C_\eta \mu^H}{\mu^F} + \frac{2NC_x(2k_1+L_g l_g)\tau^2}{\mu^F}$.

Proof: The proof can be found in Appendix B.

Theorem 1 quantifies the effects of feedback delays and the variation of generalized Nash equilibria on the sub-linearity of $\mathcal{R}_i^T$ and $\mathcal{R}_g^T$. Large τ, such as $\tau = T-1$, cannot guarantee the sub-linearity of $\mathcal{R}_i^T$ and $\mathcal{R}_g^T$. If τ is sub-linear with respect to T^{l_α}, i.e., $\lim_{T\to\infty} \frac{\tau}{T^{l_\alpha}} = 0$, then $\lim_{T\to\infty} \frac{T^{1-\frac{l_\alpha}{2}}\sqrt{\tau+1}}{T} = 0$ and $\lim_{T\to\infty} \frac{T^{\frac{1}{2}}\sqrt{k(\tau)}}{T} = 0$. The upper bounds of $\mathcal{R}_i^T$ and $\mathcal{R}_g^T$ are also related to Θ_T, which implies that a drastic fluctuation of the v-GNE sequence $\{x_t^*\}_{t=0}^T$ may affect the sub-linearity of $\mathcal{R}_i^T$ and $\mathcal{R}_g^T$. If Θ^T is sub-linear with respect to T^{1-2l_α}, i.e., $\lim_{T\to\infty} \frac{\Theta^T}{T^{1-2l_\alpha}} = 0$, then $\lim_{T\to\infty} \frac{T^{\frac{1}{2}+l_\alpha}\sqrt{\Theta^T+1}}{T} = 0$. Since $l_\alpha < \frac{1}{2}$, the sub-linearity of τ and Θ^T with respect to T^{l_α} and T^{1-2l_α} guarantees that the proposed learning algorithm performs well. Note that the upper bounds of $\mathcal{R}_i^T$ and $\mathcal{R}_g^T$ with $\tau = 0$ coincide with the conclusion of Corollary 1 in [21]. In the case of static regrets, $\Theta^T = 0$ and the benchmark is given by $x_i^* = \arg\min_{x_i \in \Omega_i} \sum_{t=1}^T J_{i,t}(x_i, x_{-i}^*)$. Static regrets are another metrics reflecting the performance of learning algorithms.

Remark 3: The result of Theorem 1 is valid in the case where feedback delays are different for players, i.e., τ is replaced by τ_i, $i \in \mathcal{V}$. Then, the upper bounds of $\mathcal{R}_i^T$ and $\mathcal{R}_g^T$ are related to $\max_{i\in\mathcal{V}}\{\tau_i\}$. In this paper, we just consider a common feedback delay τ.

IV. A Simulation Example

In this section, we consider an example to illustrate the performance of the proposed distributed learning algorithm. $\mathcal{R}_i^t/t$ and $\mathcal{R}_g^t/t$ are plotted.

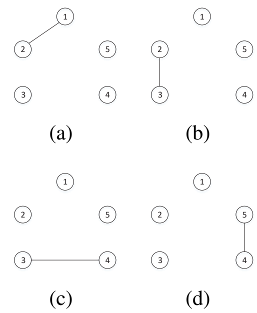

Fig. 1. 4 switching communication graphs.

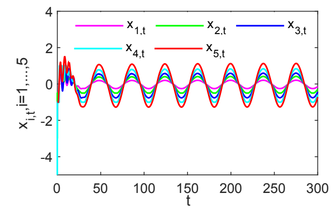

Fig. 2. Trajectories of $x_{i,t}$, $i \in \mathcal{V}$.

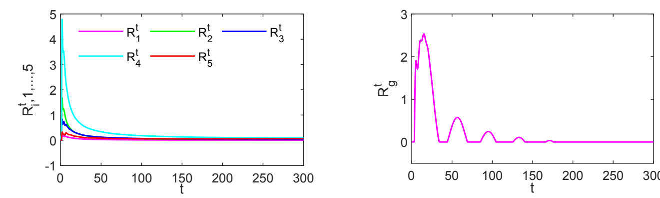

Fig. 3. Trajectories of $\mathcal{R}_i^t/t$ and $\mathcal{R}_g^t/t$, $i \in \mathcal{V}$.

Consider a time-varying network with five players labeled by index set $\mathcal{V} = \{1, 2, 3, 4, 5\}$. For $i \in \mathcal{V}$, the time-varying cost functions are given by

$$J_{i,t}(x_{i,t}, x_{-i,t}) = (x_{i,t})^2 - 0.5i\sin(\frac{t}{6})x_{i,t} + 0.5s(x_t)^2,$$

where $s(x_t) = \frac{1}{5}\sum_{j=1}^5 x_{j,t}$. The time-varying global constrained functions are given by $g_t(x_t) = \sum_{i=1}^5 x_{i,t} - \sum_{i=1}^5 l_{i,t} \leq 0$, where $l_{1,t} = l_{2,t} = l_{3,t} = l_{4,t} = 0.1\sin(\frac{t}{6})$ and $l_{5,t} = 0$. $\Omega_i = \{-5 \leq x_{i,t} \leq 5\}$, $i \in \mathcal{V}$. At GNE, $x_{i,t}^* = m_i \sin(\frac{t}{6})$, where $m_1 = 0.18$, $m_2 = 0.43$, $m_3 = 0.68$, $m_4 = 0.93$ and $m_5 = 1.18$.

The communication graph is considered with $Q = 4$, where the switching graphs are given in Fig. 1. $a_{ij}^t = 0.5$ if $(j,i) \in \mathcal{E}_t$; and $a_{ij}^t = 0$ otherwise. To satisfy Assumption 1, the diagonal entries are set as $a_{ii}^t = 1 - \sum_{j=1}^5 a_{ij}^t$, $i \in \mathcal{V}$.

Set $\tau = 1$ s, $T = 300$ s and $l_\alpha = \frac{1}{5}$. Initial values are given by $x_{1,t} = \eta_{1,t} = -1$, $x_{2,t} = \eta_{2,t} = -3$, $x_{3,t} = \eta_{3,t} = -2$, $x_{4,t} = \eta_{4,t} = -5$ and $x_{5,t} = \eta_{5,t} = -1$ for the first $\tau + 1$ s. Fig. 2 shows the trajectories of x_i. Fig. 3 shows the trajectories of time-average regrets $\mathcal{R}_i^t/t$ and the time-average violation $\mathcal{R}_g^t/t$, $i \in \mathcal{V}$. From Fig. 3, we observe that both $\mathcal{R}_i^t/t$ and $\mathcal{R}_g^t/t$ can converge to zero for $Q = 4$ and $\tau = 1$ s, which can corroborate the theoretical result of Theorem 1.

V. Conclusions

In this paper, the GNE seeking of AG problems was investigated in dynamic environments. A distributed learning algorithm was proposed with the local information and the delayed feedback information. The effects of both feedback delays and the variation of the Nash equilibrium on the sub-linearity of metrics have been quantified. An example has been utilized to illustrate the performance of the proposed learning algorithm.

VI. Appendix

A. Some prelimiary lemmas

To prove the main result of the sub-linearity of regrets, i.e., Theorem 1, we first list several lemmas related to the proposed learning algorithm (5).

Lemma 1 ([32, Lemma 1]): Suppose that Assumption 1 holds for the AG model $\mathcal{G}_N(\mathcal{G}_t, J_{i,t}, g_{i,t})$ described in (1). Consider sequences $\{\eta_{i,t}\}_{i\in\mathcal{V},t\geq 1}$ and $\{x_{i,t}\}_{i\in\mathcal{V},t\geq 1}$ generated by algorithm (5). Then, $s(x_t) = \bar{\eta}_t = \frac{1}{N}\sum_{i=1}^N \eta_{i,t} = \frac{1}{N}\sum_{i=1}^N \varphi_i(x_{i,t})$.

It follows from Assumption 2 that $\|\eta_{i,t}\|$, $\|\bar{\eta}_t\|$ and $\|\tilde{\eta}_{i,t}\|$ are bounded. It follows from Lemma 3 in [32] that there exists $C_\eta > 0$ such that $\|\eta_{i,t}\| \leq C_\eta$ for all $t > 0$ and $i \in \mathcal{V}$. Denote $\bar{\eta}_t = \frac{1}{N}\sum_{i=1}^N \eta_{i,t}$. Then we have $H_{i,t}(x_{i,t}, \bar{\eta}_{i,t}) = H_{i,t}(x_{i,t}, s(x_t))$. A result on the sequences updated by weighted averaging and a perturbed term is given as follows.

Lemma 2 ([35, Theorem 4.2]): Suppose that Assumption 1 holds for the AG model $\mathcal{G}_N(\mathcal{G}_t, J_{i,t}, g_{i,t})$ described in (1). Denote $\tilde{z}_{i,t} = \sum_{j=1}^N a_{ij}^t z_{j,t}$ and $\bar{z}_t = \frac{1}{N}\sum_{j=1}^N z_{j,t}$. The update of sequence $\{z_{i,t}\}_{i\in\mathcal{V},t\geq 1}$ is given by

$$z_{i,t+1} = \tilde{z}_{i,t} + \epsilon_{i,t}, \quad t \geq \tau + 1,$$

where $\epsilon_{i,t}$ is the perturbation at time t. Then,

$$\begin{aligned}&\|z_{i,t+1} - \bar{z}_{t+1}\| \leq N\gamma\theta_{t-\tau}\max_j \|z_{j,\tau+1}\| \\ &+ \gamma\sum_{l=\tau+2}^{T}\theta_{t+1-l}\sum_{j=1}^N\|\epsilon_{j,1}\| + \frac{1}{N}\sum_{j=1}^N\|\epsilon_{j,t+1}\| + \|\epsilon_{i,t+1}\|,\end{aligned}$$

where a and Q are given in Assumption 1, $\gamma = \left(1 - \frac{a}{2N^2}\right)^{-2}$ and $\theta = \left(1 - \frac{a}{2N^2}\right)^{\frac{1}{Q}}$.

From the learning algorithm (5) for every $i \in \mathcal{V}$, the upper bounds of $y_{i,t}$ and $\tilde{y}_{i,t}$ are obtained as follows.

Lemma 3: Suppose that Assumptions 1 and 2 hold for the AG model $\mathcal{G}_N(\mathcal{G}_t, J_{i,t}, g_{i,t})$ described in (1). Consider sequence $\{y_{i,t}\}_{i\in\mathcal{V},t\geq 1}$ generated by algorithm (5). Then, $y_{i,t} \leq \frac{L_g}{\alpha_{t-\tau}}$ and $\tilde{y}_{i,t} \leq \frac{L_g}{\alpha_{t-\tau}}$ for every $i \in \mathcal{V}$ and $t \geq \tau + 1$.

Proof: Since $y_{i,\tau+1} = 0$, $y_{i,\tau+1} \leq \frac{L_g}{\alpha_1}$ holds. Consider that $y_{i,t} \leq \frac{L_g}{\alpha_{t-\tau}}$ for $t \geq \tau + 1$. Then,

$$\begin{aligned}y_{i,t+1} =& \left[\tilde{y}_{i,t} + \alpha_{t-\tau}\left(g_{i,t-\tau}(x_{i,t-\tau}) - \alpha_{t-\tau}\tilde{y}_{i,t}\right)\right]_+ \\ \leq& \left[(1 - \alpha_{t-\tau}^2)\frac{L_g}{\alpha_{t-\tau}} + \alpha_{t-\tau}L_g\right]_+ \\ \leq& \frac{L_g}{\alpha_{t-\tau}} \leq \frac{L_g}{\alpha_{t+1-\tau}}.\end{aligned} \tag{6}$$

By induction, it follows from $y_{i,\tau+1} \leq \frac{L_g}{\alpha_1}$ and (6) that $y_{i,t} \leq \frac{L_g}{\alpha_{t-\tau}}$ for $t \geq \tau + 1$. We have that $\tilde{y}_{i,t+1} = \sum_{j=1}^N a_{ij}^t y_{j,t+1} \leq \frac{L_g}{\alpha_{t-\tau}}\sum_{j=1}^N a_{ij}^t \leq \frac{L_g}{\alpha_{t-\tau}}$, where the last inequality is based on Assumption 1. Then, we can conclude Lemma 3. ∎

In the following lemma, we obtain the upper bound related to the gap between the strategy $x_{i,t}$ and its delayed strategy $x_{i,t-\tau}$, $i \in \mathcal{V}$ and $t \geq \tau + 1$.

Lemma 4: Suppose that Assumption 2 holds for the AG model $\mathcal{G}_N(\mathcal{G}_t, J_{i,t}, g_{i,t})$ described in (1). Consider sequence $\{x_{i,t}\}_{i\in\mathcal{V},t\geq 1}$ generated by algorithm (5). Then, for every $i \in \mathcal{V}$ and $t \geq \tau + 1$,

$$\begin{aligned}&\textstyle\sum_{i=1}^N \|x_{i,t} - x_{i,t-\tau}\| \leq 2NC_x\tau, && \tau + 1 \leq t \leq 2\tau, \\ &\textstyle\sum_{i=1}^N \|x_{i,t} - x_{i,t-\tau}\| \leq 2NC_x\tau\alpha_{t-2\tau}, && 2\tau + 1 \leq t.\end{aligned}$$

Proof: From the distributed learning algorithm (5) and Assumption 2, for $t \geq \tau + 1$, we have that

$$\begin{aligned}&\|x_{i,t+1} - x_{i,t}\| \\ \leq& \Bigg\|\alpha_{t-\tau}P_{\Omega_i}\Bigg(x_{i,t} - \alpha_{t-\tau}H_{i,t-\tau}(x_{i,t-\tau}, \tilde{\eta}_{i,t-\tau}) \\ &- \alpha_{t-\tau}^2\nabla g_{i,t-\tau}(x_{i,t-\tau})\tilde{y}_{i,t}\Bigg) - \alpha_{t-\tau}x_{i,t}\Bigg\| \\ \leq& 2C_x\alpha_{t-\tau}.\end{aligned} \tag{7}$$

It follows from (7) and the non-expansiveness of projection operators that, for $t = 2\tau + 1, \cdots, \tau + T$,

$$\begin{aligned}&\sum_{i=1}^N \|x_{i,t} - x_{i,t-\tau}\| \\ \leq& \sum_{i=1}^N\sum_{j=0}^{\tau-1}\|x_{i,t-\tau+j+1} - x_{i,t-\tau+j}\| \\ \leq& \sum_{i=1}^N\sum_{j=t-2\tau}^{t-\tau-1} 2C_x\alpha_j \leq 2NC_x\tau\alpha_{t-2\tau}.\end{aligned}$$

For $t = \tau + 2, \cdots, 2\tau$, since $x_{i,1} = x_{i,2} = \cdots = x_{i,\tau+1}$, we have that $\sum_{i=1}^N \|x_{i,t} - x_{i,t-\tau}\| \leq 2NC_x\tau$. For $t = \tau + 1$, $\|x_{i,\tau+1} - x_{i,1}\| = 0$ implies that the above inequality holds. Therefore, we conclude Lemma 4. ∎

The upper bound related to the gap between the strategies and strategies at GNE is given as follows.

Lemma 5: Suppose that Assumptions 1, 2 and 3 hold for the AG model $\mathcal{G}_N(\mathcal{G}_t, J_{i,t}, g_{i,t})$ described in (1). Consider sequences $\{x_{i,t}\}_{i\in\mathcal{V},t\geq 1}$ and $\{y_{i,t}\}_{i\in\mathcal{V},t\geq 1}$ generated by algorithm (5). Then,

$$\begin{aligned}&\sum_{t=1}^T \|x_t - x_t^*\|^2 \\ &\leq \mathcal{O}(T^{2l_\alpha}\Theta^T + T^{2l_\alpha} + T^{1-l_\alpha} + \tau T^{1-l_\alpha}) + k(\tau) \\ &\quad - \sum_{t=1}^T\sum_{i=1}^N \frac{\alpha_t}{\mu^F}g_{i,t}(x_{i,t})\tilde{y}_{i,t+\tau}.\end{aligned} \tag{8}$$

Proof: The Jensen's inequality implies that $\varphi(\sum_i p_i z_i) \leq \sum_i p_i\varphi(z_i)$, where $p_i \geq 0$, $\sum_i p_i = 1$ and φ is convex. Then, for $t \geq \tau + 1$, it follows from the distributed learning algorithm

(2) and (5) that

$$\begin{aligned}
&\sum_{i=1}^{N}\|x_{i,t+1}-x_{i,t-\tau}^{*}\|^{2}\\
&=\sum_{i=1}^{N}\Big\|(1-\alpha_{t-\tau})(x_{i,t}-x_{i,t-\tau}^{*})+\alpha_{t-\tau}P_{\Omega_i}\Big(x_{i,t}\\
&\quad-\alpha_{t-\tau}H_{i,t-\tau}(x_{i,t-\tau},\tilde{\eta}_{i,t-\tau})\\
&\quad-\alpha_{t-\tau}^{2}\nabla g_{i,t-\tau}(x_{i,t-\tau})\tilde{y}_{i,t}\Big)-\alpha_{t-\tau}P_{\Omega_i}\Big(x_{i,t-\tau}^{*}\\
&\quad-\alpha_{t-\tau}H_{i,t-\tau}(x_{i,t-\tau}^{*},s(x_{t-\tau}^{*}))\\
&\quad-\alpha_{t-\tau}^{2}\nabla g_{i,t-\tau}(x_{i,t-\tau}^{*})y_{t-\tau}^{*}\Big)\Big\|^{2}\\
&=\|\Delta_1\|^{2}+\alpha_{t-\tau}^{3}\|\Delta_2\|+\alpha_{t-\tau}^{5}\|\Delta_3\|^{2}-2\alpha_{t-\tau}^{2}\Delta_1^{T}\Delta_2\\
&\quad-2\alpha_{t-\tau}^{3}\Delta_1^{T}\Delta_3+2\alpha_{t-\tau}^{4}\Delta_2^{T}\Delta_3 \qquad (9)
\end{aligned}$$

where $\Delta_1=x_{i,t}-x_{i,t-\tau}^{*}$, $\Delta_2=H_{i,t-\tau}(x_{i,t-\tau},\tilde{\eta}_{i,t-\tau})-H_{i,t-\tau}(x_{i,t-\tau}^{*},s(x_{t-\tau}^{*}))$ and $\Delta_3=\nabla g_{i,t-\tau}(x_{i,t-\tau})\tilde{y}_{i,t}-\nabla g_{i,t-\tau}(x_{i,t-\tau}^{*})y_{t-\tau}^{*}$.

It follows from Assumption 2 and Lemma 1 that $\|\eta_{i,t}\|$ is bounded and $H_{i,t}$ is continuous. Therefore, for $t\geq\tau+1$, $\sum_{i=1}^{N}\|\Delta_2\|^{2}\leq 4k_1^{2}N$. It follows from Assumption 2 and Lemma 3 that, for $t\geq 2\tau+1$,

$$\begin{aligned}
\alpha_{t-\tau}^{2}\sum_{i=1}^{N}\|\Delta_3\|^{2}&\leq\alpha_{t-\tau}^{2}\sum_{i=1}^{N}2l_g^{2}\left(\frac{L_g^{2}}{\alpha_{t-\tau}^{2}}+C_y^{2}\right)\\
&\leq 2Nl_g^{2}(L_g^{2}+C_y^{2}), \qquad (10)
\end{aligned}$$

where $\alpha_{t-\tau}<1$ is employed in the last inequality. Since $\tilde{y}_{i,t-\tau}=0$ for $t=\tau+1,\cdots,2\tau$, (10) holds. It follows from Assumption 2 and Lemma 3 that $\alpha_{t-\tau}\sum_{i=1}^{N}\Delta_2^{T}\Delta_3\leq 2Nk_1l_g(L_g+C_y)$. Under Assumptions 2, 3 and Lemmas 1 and 4, for $t\geq 2\tau+1$, we have that

$$\begin{aligned}
&\sum_{i=1}^{N}\Delta_1^{T}\Delta_2\\
&=\sum_{i=1}^{N}(x_{i,t-\tau}-x_{i,t-\tau}^{*})^{T}\Big(H_{i,t-\tau}(x_{i,t-\tau},\bar{\eta}_{t-\tau})\\
&\quad-H_{i,t-\tau}(x_{i,t-\tau}^{*},s(x_{t-\tau}^{*}))+H_{i,t-\tau}(x_{i,t-\tau},\tilde{\eta}_{i,t-\tau})\\
&\quad-H_{i,t-\tau}(x_{i,t-\tau},\bar{\eta}_{t-\tau}))\Big)+\sum_{i=1}^{N}(x_{i,t}-x_{i,t-\tau})^{T}\\
&\quad\times\Big(H_{i,t-\tau}(x_{i,t-\tau},\tilde{\eta}_{i,t-\tau})-H_{i,t-\tau}(x_{i,t-\tau}^{*},s(x_{t-\tau}^{*}))\Big)\\
&\geq\mu^{F}\|x_{t-\tau}-x_{t-\tau}^{*}\|^{2}-2C_x\mu^{H}\sum_{i=1}^{N}\|\tilde{\eta}_{i,t-\tau}-\bar{\eta}_{t-\tau}\|\\
&\quad-4NC_xk_1\tau\alpha_{t-2\tau}. \qquad (11)
\end{aligned}$$

Similarly, for $t=\tau+1,\cdots,2\tau$, we have that

$$\begin{aligned}
\sum_{i=1}^{N}\Delta_1^{T}\Delta_2\geq&\mu^{F}\|x_{t-\tau}-x_{t-\tau}^{*}\|^{2}-4NC_xk_1\tau\\
&-2C_x\mu^{H}\sum_{i=1}^{N}\|\tilde{\eta}_{i,t-\tau}-\bar{\eta}_{t-\tau}\|. \qquad (12)
\end{aligned}$$

From (5), $y_i\geq 0$ and $\tilde{y}_{i,t}\geq 0$ for every $i\in\mathcal{V}$ and $t\geq 1$. From (2), $x_{i,t-\tau}^{*}$ satisfies that $\sum_{i=1}^{N}g_{i,t-\tau}(x_{i,t-\tau}^{*})\leq 0$. Therefore, $\sum_{i=1}^{N}g_{i,t-\tau}(x_{i,t-\tau}^{*})\bar{y}_t\leq 0$. Moreover, for $t\geq 2\tau+1$, it follows from Assumption 2, Lemmas 3 and 4 that

$$\begin{aligned}
&\alpha_{t-\tau}\sum_{i=1}^{N}\Delta_1^{T}\Delta_3\\
&=\alpha_{t-\tau}\sum_{i=1}^{N}(x_{i,t-\tau}-x_{i,t-\tau}^{*})^{T}\nabla g_{i,t-\tau}(x_{i,t-\tau})\tilde{y}_{i,t}\\
&\quad+\alpha_{t-\tau}\sum_{i=1}^{N}(x_{i,t}-x_{i,t-\tau})^{T}\nabla g_{i,t-\tau}(x_{i,t-\tau})\tilde{y}_{i,t}\\
&\quad-\alpha_{t-\tau}\sum_{i=1}^{N}(x_{i,t}-x_{i,t-\tau}^{*})^{T}\nabla g_{i,t-\tau}(x_{i,t-\tau}^{*})y_{t-\tau}^{*}\\
&\geq\alpha_{t-\tau}\sum_{i=1}^{N}g_{i,t-\tau}(x_{i,t-\tau})\tilde{y}_{i,t}-L_g\alpha_{t-\tau}\sum_{i=1}^{N}\|\tilde{y}_{i,t}-\bar{y}_t\|\\
&\quad-2NC_xL_gl_g\tau\alpha_{t-2\tau}-2NC_xl_gC_y\alpha_{t-\tau}, \qquad (13)
\end{aligned}$$

where the last inequality follows from $(x_{i,t-\tau}-x_{i,t-\tau}^{*})^{T}\times\nabla g_{i,t-\tau}(x_{i,t-\tau})\geq g_{i,t-\tau}(x_{i,t-\tau})-g_{i,t-\tau}(x_{i,t-\tau}^{*})$. Similarly for $t=\tau+1,\cdots,2\tau$, we have that

$$\begin{aligned}
&\alpha_{t-\tau}\sum_{i=1}^{N}\Delta_1^{T}\Delta_3\\
&\geq\alpha_{t-\tau}\sum_{i=1}^{N}g_{i,t-\tau}(x_{i,t-\tau})\tilde{y}_{i,t}-L_g\alpha_{t-\tau}\sum_{i=1}^{N}\|\tilde{y}_{i,t}-\bar{y}_t\|\\
&\quad-2NC_xL_gl_g\tau-2NC_xl_gC_y\alpha_{t-\tau}. \qquad (14)
\end{aligned}$$

Based on inequalities (10), (11), (12), (13) and (14), summing (9) over $t=\tau+1,\cdots,\tau+T$, we have that

$$\begin{aligned}
&\sum_{t=\tau+1}^{\tau+T}\|x_{t-\tau}-x_{t-\tau}^{*}\|^{2}\\
&\leq\sum_{t=\tau+1}^{\tau+T}\sum_{i=1}^{N}\frac{1}{2\mu^{F}\alpha_{t-\tau}^{2}}(\|x_{i,t}-x_{i,t-\tau}^{*}\|^{2}-\|x_{i,t+1}-x_{i,t-\tau}^{*}\|^{2})\\
&\quad+\frac{2C_x\mu^{H}}{\mu^{F}}\sum_{t=\tau+1}^{\tau+T}\sum_{i=1}^{N}\|\tilde{\eta}_{i,t-\tau}-\bar{\eta}_{t-\tau}\|\\
&\quad+\sum_{t=\tau+1}^{\tau+T}\left(\frac{k_2}{2\mu^{F}}\alpha_{t-\tau}+\frac{\Psi\tau}{\mu^{F}}\right)+\frac{L_g}{\mu^{F}}\sum_{t=\tau+1}^{\tau+T}\sum_{i=1}^{N}\|\tilde{y}_{i,t}-\bar{y}_t\|\\
&\quad-\sum_{t=\tau+1}^{\tau+T}\sum_{i=1}^{N}\frac{\alpha_{t-\tau}}{\mu^{F}}g_{i,t-\tau}(x_{i,t-\tau})\tilde{y}_{i,t}, \qquad (15)
\end{aligned}$$

where $k_2=4k_1^{2}N+2Nl_g^{2}(L_g^{2}+C_y^{2})+4Nk_1l_g(L_g+C_y)+4NC_xl_gC_y$ and

$$\Psi=\begin{cases}2NC_x(2k_1+L_gl_g)\alpha_{t-2\tau}, & \tau+1\leq t\leq 2\tau,\\ 2NC_x(2k_1+L_gl_g), & 2\tau+1\leq t.\end{cases} \qquad (16)$$

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2023.3237781

The first term on the right side of (15) is bounded by

$$\begin{aligned}&\sum_{t=\tau+1}^{\tau+T}\sum_{i=1}^{N}\frac{1}{2\mu^F\alpha_{t-\tau}^2}\left(\|x_{i,t}-x_{i,t-\tau}^*\|^2-\|x_{i,t+1}-x_{i,t-\tau}^*\|^2\right)\\&\leq\frac{4C_x}{2\mu^F\alpha_{T-\tau}^2}\sum_{i=1}^{N}\sum_{t=\tau+1}^{\tau+T}\|x_{i,t+1-\tau}^*-x_{i,t-\tau}^*\|+\frac{2NC_x^2}{\mu^F\alpha_T^2}\\&=\mathcal{O}(T^{2l_\alpha}\Theta^T+T^{2l_\alpha}).\end{aligned}\tag{17}$$

Based on the definition of α_t, we have that

$$\sum_{t=1}^{T-\tau-1}\alpha_t\leq\frac{1}{1-l_\alpha}+\int_{t=1}^{T}t^{-l_\alpha}=\frac{T^{1-l_\alpha}}{1-l_\alpha}=\mathcal{O}(T^{1-l_\alpha}).\tag{18}$$

Then, $\sum_{t=\tau+1}^{\tau+T}\frac{k_2}{2\mu^F}\alpha_{t-\tau}=\mathcal{O}(T^{1-l_\alpha})$. To bound the second term on the right side of (15), Lemma 2 is utilized. We represent the iteration of $\eta_{i,t}$ in (5) as $\eta_{i,t+1}=\tilde{\eta}_{i,t}+\epsilon_{\eta_i,t}$, where $\epsilon_{\eta_i,t}=\varphi_i(x_{i,t+1})-\varphi_i(x_{i,t})$. It follows from (7) and Assumption 2, for $t\geq\tau+1$, we have that

$$\|\epsilon_{\eta_i,t}\|=\|\varphi_i(x_{i,t+1})-\varphi_i(x_{i,t})\|\leq 2\mu^\varphi C_x\alpha_{t-\tau}.\tag{19}$$

It follows from Lemma 2 that

$$\begin{aligned}&\sum_{t=\tau+1}^{T-1}\|\eta_{i,t+1}-\bar{\eta}_{t+1}\|\\&\leq\left(\frac{2\gamma\mu^\varphi C_xN\theta}{1-\theta}+4\mu^\varphi C_x\right)\sum_{t=1}^{T-\tau-1}\alpha_t\\&\quad+\frac{N\gamma\theta}{1-\theta}\max_j\|\eta_{j,\tau+1}\|.\end{aligned}\tag{20}$$

Hence, it follows from Assumption 1, (18), (19) and (20) that

$$\begin{aligned}&\frac{2C_x\mu^H}{\mu^F}\sum_{t=\tau+1}^{\tau+T}\sum_{i=1}^{N}\|\tilde{\eta}_{i,t-\tau}-\bar{\eta}_{t-\tau}\|\\&\leq\frac{2C_x\mu^H}{\mu^F}\Bigg(\sum_{t=\tau+1}^{2\tau}\sum_{i=1}^{N}\|\eta_{i,t-\tau}-\bar{\eta}_{t-\tau}\|\\&\quad+\sum_{i=1}^{N}\|\eta_{i,\tau+1}-\bar{\eta}_{\tau+1}\|+\sum_{t=\tau+2}^{T}\sum_{i=1}^{N}\|\eta_{i,t}-\bar{\eta}_t\|\Bigg)\\&\leq\left(2N\tau+2N+\frac{N^2\gamma\theta}{1-\theta}\right)\frac{2C_xC_\eta\mu^H}{\mu^F}+\mathcal{O}(T^{1-l_\alpha}).\end{aligned}\tag{21}$$

In addition, we present the iterate y_i in (5) as follows.

$$y_{i,t+1}=\tilde{y}_{i,t}+\epsilon_{y_i,t},$$

where $\epsilon_{y_i,t}=\left[\tilde{y}_{i,t}+\alpha_{t-\tau}\left(g_{i,t-\tau}(x_{i,t-\tau})-\alpha_{t-\tau}\tilde{y}_{i,t}\right)\right]_+-\tilde{y}_{i,t}$. It follows from Assumption 2 and Lemma 3 that $\|\epsilon_{y_i,t}\|\leq 2L_g\alpha_{t-\tau}$. Therefore, it follows from Lemma 2 that

$$\begin{aligned}&\frac{L_g}{\mu^F}\sum_{t=\tau+1}^{\tau+T}\sum_{i=1}^{N}\|\tilde{y}_{i,t}-\bar{y}_t\|\\&\leq\frac{L_g}{\mu^F}\left(\frac{2\gamma L_gN^2\theta}{1-\theta}+4L_gN\right)\sum_{t=1}^{T}\alpha_t=\mathcal{O}(T^{1-l_\alpha}).\end{aligned}\tag{22}$$

It follows from the definition of Ψ in (16) that

$$\begin{aligned}\frac{1}{\mu^F}\sum_{t=\tau+1}^{\tau+T}\Psi\tau=&\frac{1}{\mu^F}\Bigg(\sum_{t=\tau+1}^{2\tau}2NC_x(2k_1+L_gl_g)\tau\\&+\sum_{t=2\tau+1}^{\tau+T}2NC_x(2k_1+L_gl_g)\tau\alpha_{t-2\tau}\Bigg)\\=&\frac{2NC_x(2k_1+L_gl_g)\tau^2}{\mu^F}+\mathcal{O}(\tau T^{1-l_\alpha}).\end{aligned}\tag{23}$$

It follows from (15), (17), (21), (22) and (23) that

$$\begin{aligned}&\sum_{t=\tau+1}^{\tau+T}\|x_{t-\tau}-x_{t-\tau}^*\|^2\\&\leq\mathcal{O}(T^{2l_\alpha}\Theta^T+T^{2l_\alpha}+T^{1-l_\alpha}+\tau T^{1-l_\alpha})+k(\tau)\\&\quad-\sum_{t=\tau+1}^{\tau+T}\sum_{i=1}^{N}\frac{\alpha_{t-\tau}}{\mu^F}g_{i,t-\tau}(x_{i,t-\tau})\tilde{y}_{i,t}.\end{aligned}$$

Then, we obtain the result in Lemma 5. ■

To obtain the upper bounds of $\mathcal{R}_i^T$ and $\mathcal{R}_g^T$, the following relation is needed.

Lemma 6: Suppose that Assumption 2 holds. For $y>0$,

$$\begin{aligned}&-\alpha_t\sum_{t=1}^{T}\sum_{i=1}^{N}(\tilde{y}_{i,t+\tau}-y)^Tg_{i,t}(x_{i,t})\\&\leq\frac{N}{2}\left(1+\sum_{t=1}^{T}\alpha_t^2\right)y^2+\mathcal{O}(T^{1-2l_\alpha}).\end{aligned}\tag{24}$$

Proof: It follows from the learning algorithm (5) that

$$\begin{aligned}&\sum_{i=1}^{N}\|y_{i,t+1}-y\|^2\\&\leq\sum_{i=1}^{N}\|y_{i,t}-y\|^2+4NL_g^2\alpha_{t-\tau}^2+N\alpha_{t-\tau}^2\|y\|^2\\&\quad+2\alpha_{t-\tau}\sum_{i=1}^{N}(\tilde{y}_{i,t}-y)^Tg_{i,t-\tau}(x_{i,t-\tau}).\end{aligned}\tag{25}$$

Summing (25) over $t=\tau+1,\cdots,\tau+T$, we have that

$$\begin{aligned}&-\alpha_{t-\tau}\sum_{t=\tau+1}^{\tau+T}\sum_{i=1}^{N}(\tilde{y}_{i,t-\tau}-y)^Tg_{i,t-\tau}(x_{i,t-\tau})\\&\leq\frac{1}{2}\sum_{t=\tau+1}^{\tau+T}\sum_{i=1}^{N}\left(\|y_{i,t}-y\|^2-\|y_{i,t+1}-y\|^2\right)\\&\quad+\sum_{t=\tau+1}^{\tau+T}\left(2NL_g^2\alpha_{t-\tau}^2+\frac{N}{2}\alpha_{t-\tau}^2\|y\|^2\right).\end{aligned}\tag{26}$$

Since $y_{i,t}=0$ for all $t\leq\tau+1$, we have that

$$\begin{aligned}&\sum_{i=1}^{N}\sum_{t=\tau+1}^{\tau+T}\left(\|y_{i,t}-y\|^2-\|y_{i,t+1}-y\|^2\right)\\&\leq\sum_{i=1}^{N}\left(\|y_{i,\tau+1}-y\|^2-\|y_{i,\tau+T+1}-y\|^2\right)\leq Ny^2.\end{aligned}\tag{27}$$

It follows from the definition of α_t that $\sum_{t=\tau+1}^{\tau+T}\alpha_{t-\tau}^2\leq\frac{1}{1-2l_\alpha}+\int_{t=1}^{T}t^{-2l_\alpha}=\frac{T^{1-2l_\alpha}}{1-2l_\alpha}=\mathcal{O}(T^{1-2l_\alpha})$. Then, combing (26) and (27), we obtain the result in Lemma 6. ■

B. Proof of Theorem 1

Adding the inequalities (8) and (24), we have that

$$\begin{aligned}&\sum_{t=1}^{T}\|x_t-x_t^*\|^2+\frac{1}{\mu^F}\sum_{t=1}^{T}\alpha_t\sum_{i=1}^{N}g_{i,t}(x_{i,t})y\\&\le\mathcal{O}(T^{2l_\alpha}\Theta^T+T^{2l_\alpha}+T^{1-l_\alpha}+\tau T^{1-l_\alpha})+k(\tau)\\&\quad+\frac{N}{2\mu^F}\left(1+\sum_{t=1}^{T}\alpha_t^2\right)y^2.\end{aligned}\tag{28}$$

It follows from Assumption 2 that $\mathcal{R}_i^T\le\mu^J\sum_{t=1}^{T}\|x_{i,t}-x_{i,t}^*\|\le\mu^J\sqrt{T\sum_{t=1}^{T}\|x_t-x_t^*\|^2}$. Set $y=0$. Based on (28), we obtain the upper bound of $\mathcal{R}_i^T$.

Set $y=\frac{\left[\sum_{t=1}^{T}\alpha_t\sum_{i=1}^{N}g_{i,t}(x_{i,t})\right]_+}{N(1+\sum_{t=1}^{T}\alpha_t^2)}$. Then, we have that

$$\begin{aligned}&\frac{1}{\mu^F}\sum_{t=1}^{T}\alpha_t\sum_{i=1}^{N}g_{i,t}(x_{i,t})y-\frac{N(1+\sum_{t=1}^{T}\alpha_t^2)}{2\mu^F}\|y\|^2\\&=\frac{\left[\sum_{t=1}^{T}\alpha_t\sum_{i=1}^{N}g_{i,t}(x_{i,t})y\right]_+^2}{2N\mu^F(1+\sum_{t=1}^{T}\alpha_t^2)}\\&\ge\frac{\alpha_T^2(\mathcal{R}_g^T)^2}{2N\mu^F(1+\sum_{t=1}^{T}\alpha_t^2)}.\end{aligned}\tag{29}$$

Then, it follows from (28) and (29) that

$$\begin{aligned}(\mathcal{R}_g^T)^2\le&\frac{2N\mu^F(1+\sum_{t=1}^{T}\alpha_t^2)}{\alpha_T^2}\big(\mathcal{O}(T^{2l_\alpha}\Theta^T+T^{2l_\alpha}\\&+T^{1-l_\alpha}+\tau T^{1-l_\alpha})+k(\tau)\big).\end{aligned}$$

Then, we obtain the upper bound of $\mathcal{R}_g^T$. ∎

REFERENCES

[1] Y. Ho, A. Bryson, and S. Baron, "Differential games and optimal pursuit-evasion strategies," *IEEE Transactions on Automatic Control*, vol. 10, no. 4, pp. 385–389, 1965.

[2] A. W. Starr and Y.-C. Ho, "Nonzero-sum differential games," *Journal of optimization theory and applications*, vol. 3, no. 3, pp. 184–206, 1969.

[3] G. Owen, *Game Theory*. Emerald Group Publishing, 2013.

[4] J. R. Marden and J. S. Shamma, "Game theory and control," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, pp. 105–134, 2018.

[5] H. M. Soliman and A. Leon-Garcia, "Game-theoretic demand-side management with storage devices for the future smart grid," *IEEE Transactions on Smart Grid*, vol. 5, no. 3, pp. 1475–1485, 2014.

[6] A.-H. Mohsenian-Rad, V. W. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, "Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid," *IEEE Transactions on Smart Grid*, vol. 1, no. 3, pp. 320–331, 2010.

[7] A. Farraj, E. Hammad, A. A. Daoud, and D. Kundur, "A game-theoretic analysis of cyber switching attacks and mitigation in smart grid systems," *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 1846–1855, 2016.

[8] M. Pirani, E. Nekouei, H. Sandberg, and K. H. Johansson, "A game-theoretic framework for security-aware sensor placement problem in networked control systems," *IEEE Transactions on Automatic Control*, vol. 67, no. 7, pp. 3699–3706, 2021.

[9] G. Debreu, "A social equilibrium existence theorem," *Proceedings of the National Academy of Sciences*, vol. 38, no. 10, pp. 886–893, 1952.

[10] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave n-person games," *Econometrica: Journal of the Econometric Society*, pp. 520–534, 1965.

[11] P. T. Harker, "Generalized Nash games and quasi-variational inequalities," *European Journal of Operational Research*, vol. 54, no. 1, pp. 81–94, 1991.

[12] S. Liang, P. Yi, and Y. Hong, "Distributed Nash equilibrium seeking for aggregative games with coupled constraints," *Automatica*, vol. 85, pp. 179–185, 2017.

[13] M. Bianchi and S. Grammatico, "Continuous-time fully distributed generalized Nash equilibrium seeking for multi-integrator agents," *Automatica*, vol. 129, p. 109660, 2021.

[14] G. Belgioioso, A. Nedić, and S. Grammatico, "Distributed generalized Nash equilibrium seeking in aggregative games on time-varying networks," *IEEE Transactions on Automatic Control*, vol. 66, no. 5, pp. 2061–2075, 2021.

[15] D. Gadjov and L. Pavel, "Single-timescale distributed GNE seeking for aggregative games over networks via forward-backward operator splitting," *IEEE Transactions on Automatic Control*, vol. 66, no. 7, pp. 3259–3266, 2021.

[16] M. Ye, G. Hu, L. Xie, and S. Xu, "Differentially private distributed Nash equilibrium seeking for aggregative games," *IEEE Transactions on Automatic Control*, vol. 67, no. 5, pp. 2451–2458, 2022.

[17] X. Li, L. Xie, and Y. Hong, "Distributed aggregative optimization over multi-agent networks," *IEEE Transactions on Automatic Control*, vol. 67, no. 6, pp. 3165–3171, 2022.

[18] B. Franci and S. Grammatico, "Stochastic generalized Nash equilibrium-seeking in merely monotone games," *IEEE Transactions on Automatic Control*, vol. 67, no. 8, pp. 3905–3919, 2022.

[19] D. Gadjov and L. Pavel, "A passivity-based approach to Nash equilibrium seeking over networks," *IEEE Transactions on Automatic Control*, vol. 64, no. 3, pp. 1077–1092, 2019.

[20] M. Zinkevich, "Online convex programming and generalized infinitesimal gradient ascent," in *Proceedings of the 20th International Conference on Machine Learning (icml-03)*, 2003, pp. 928–936.

[21] K. Lu, G. Li, and L. Wang, "Online distributed algorithms for seeking generalized Nash equilibria in dynamic environments," *IEEE Transactions on Automatic Control*, vol. 66, no. 5, pp. 2289–2296, 2021.

[22] X. Cao and K. R. Liu, "Online convex optimization with time-varying constraints and bandit feedback," *IEEE Transactions on Automatic Control*, vol. 64, no. 7, pp. 2665–2680, 2018.

[23] M. Meng, X. Li, Y. Hong, J. Chen, and L. Wang, "Decentralized online learning for noncooperative games in dynamic environments," *arXiv preprint arXiv:2105.06200*, 2021.

[24] Y.-G. Hsieh, F. Iutzeler, J. Malick, and P. Mertikopoulos, "Multi-agent online optimization with delays: Asynchronicity, adaptivity, and optimism," *Journal of Machine Learning Research*, vol. 23, no. 78, pp. 1–49, 2021.

[25] X. Cao, J. Zhang, and H. V. Poor, "Constrained online convex optimization with feedback delays," *IEEE Transactions on Automatic Control*, vol. 66, no. 11, pp. 5049–5064, 2020.

[26] C. Godsil and G. F. Royle, *Algebraic Graph Theory*. Springer Science & Business Media, 2013.

[27] R. A. Freeman, P. Yang, and K. M. Lynch, "Stability and convergence properties of dynamic average consensus estimators," in *Proceedings of the 45th IEEE Conference on Decision and Control*. IEEE, 2006, pp. 338–343.

[28] T. Yang, X. Yi, J. Wu, Y. Yuan, D. Wu, Z. Meng, Y. Hong, H. Wang, Z. Lin, and K. H. Johansson, "A survey of distributed optimization," *Annual Reviews in Control*, vol. 47, pp. 278–305, 2019.

[29] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on automatic control*, vol. 49, no. 9, pp. 1520–1533, 2004.

[30] C. Sun and G. Hu, "Continuous-time penalty methods for Nash equilibrium seeking of a nonsmooth generalized noncooperative game," *IEEE Transactions on Automatic Control*, vol. 66, no. 10, pp. 4895–4902, 2021.

[31] Y. Su, F. Liu, Z. Wang, S. Mei, and Q. Lu, "Online distributed tracking of generalized Nash equilibrium on physical networks," *Autonomous Intelligent Systems*, vol. 1, no. 1, pp. 1–12, 2021.

[32] S. Lee and M. M. Zavlanos, "On the sublinear regret of distributed primal-dual algorithms for online constrained optimization," *arXiv preprint arXiv:1705.11128*, 2017.

[33] F. Facchinei and C. Kanzow, "Generalized Nash equilibrium problems," *Annals of Operations Research*, vol. 175, no. 1, pp. 177–211, 2010.

[34] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer, 2003.

[35] S. Sundhar Ram, A. Nedić, and V. V. Veeravalli, "A new class of distributed optimization algorithms: Application to regression of distributed data," *Optimization Methods and Software*, vol. 27, no. 1, pp. 71–88, 2012.