Network Aggregative Game in Unknown Dynamic Environment With Myopic Agents and Delay

Mohammad Shokri and Hamed Kebriaei, *Senior Member, IEEE*

***Abstract*—In this article, the framework of network aggregative game (NAG) is extended by considering the interaction of players in a dynamic environment whose model is not known to them. The cost function of each player depends on its own strategy, the aggregate strategy of its neighbor players in the network, and also the state of the environment. The state of the environment evolves with a dynamics that is affected by the strategies of players as an input to the model. The players of NAG are modeled as myopic selfish agents who observe the decision value of their neighbors and the state of the environment with some time delay and, then, adjust their decisions using a projected subgradient rule. The main contribution of this article is to provide conditions in which the strategies of myopic agents converge to the unique fixed-point Nash equilibrium point of the proposed dynamic NAG. Furthermore, as a case study, the framework is applied to a wireless network by modeling the service providers as the agents and evolution of service selection by users as the dynamics of the environment.**

***Index Terms*—Dynamic environment, myopic agents, network aggregative game (NAG), time delay.**

I. Introduction

Network games have been gaining importance among researchers in recent years [1]. In the network game framework, the cost function of each individual agent is influenced by other neighboring agents via a communication network. Each agent aims to find its optimal decision selfishly while communicating with its neighbors [2]. In particular, when the cost function of an agent is affected by an aggregation (e.g., weighted sum) of its neighbors' decisions, the problem is called network aggregative game (NAG). The NAG framework can be applied to a variety of applications, such as social networks [3], wireless cellular networks [4], public good provision [5], and microgrids [6].

In the literature, NAGs are investigated from different aspects. In [7] and [8], the best response dynamics is proposed for the decision making of the agents with convex cost function in a NAG framework. While each agent has access to decision information of its neighbors, the convergence of decisions to the Nash equilibrium point of NAG is proven. Moreover, some research works utilized an additive regularization term to the cost function of the agents in order to maintain the decision of each agent close to aggregate term of its neighbors while the best response of the agents is calculated using proximal methods [9], [10]. Besides, considering a limited computational ability of the agents, in [11], a subgradient method is utilized as a Nash seeking algorithm in a NAG with stochastic agents' communication and activeness.

In the aforementioned research works on NAG, the cost function of the agents is only affected by the neighbors' strategies. Nevertheless, in many situations, the agents interact in a dynamic environment whose state affects the cost function of the agents. Dynamic noncooperative game theory (DNG) is a well-known branch of game theory, which has been extensively studied in [12], and further developed in the seminal papers [13], [14]. In the setting of DNG, there exist a dynamic system/environment whose state affects the players' decisions and, in turn, the dynamics is also affected by decisions of the players [12]. In DNG, agents find their optimal trajectory over a time horizon as their strategy. To obtain the Nash equilibrium strategies of the players in standard DNGs, some coupled Hamilton–Jacobi–Bellman (HJB) or Bellman equations need to be solved in continuous-time or discrete-time systems, receptively. For DNGs with quadratic cost functions and linear-known dynamic, the closed-form solution of Nash strategies can be obtained as a linear state feedback thanks to solving a set of coupled Riccati equations [15].

However, HJB/Bellman equations cannot be solved analytically for non-LQ problems and also when the dynamics of environment/system in unknown [16]. Due to such difficulties, adaptive dynamic programming (ADP) is proposed to approximate both the value function and control inputs via actor–critic neural networks [17]. A few research works in this area have been dedicated to obtaining the Nash equilibrium point of discrete-time nonzero-sum DNGs with unknown dynamics [17], [18]. However, in this approach, the parameters of neural networks converge uniformly ultimate bounded around their optimal values and, therefore, ADP methods converge to a bound around the Nash equilibrium point. In addition, the ADP methods suffer from huge computational cost, which is not convenient with the bounded rationality and myopic behavior of the players in many decision-making problems [19], [20].

To the best of our knowledge, NAG has not been studied in dynamic environment up to now. The aim of our study is to set up a connection between DNG and NAG for the first time. In this article, we develop a dynamic NAG model in which the cost function of each agent is affected by the aggregated decision variables of its neighbors on a directed graph and also the state of the environment while the state evolves with a dynamic model. Neither the cost function of rival agents nor the dynamics of environment is known to an agent. Due to such lack of information, the agents are modeled as some myopic decision-makers who receive the latest decision value of their neighbors and also the last state value of the environment and, then, adjust their decision for one step ahead using the subgradient method. The motivation behind such naive expectation and myopic adjustment is taken from the fact that myopic fashion and bounded rationality can be fitted to human behavior in decision

Manuscript received October 12, 2020; revised February 12, 2021; accepted March 28, 2021. Date of publication April 5, 2021; date of current version March 29, 2022. This work was supported by the Institute for Research in Fundamental Sciences (IPM) under Grant CS 1400-4-451. Recommended by Associate Editor S. Grammatico. *(Corresponding author: Hamed Kebriaei.)*

Mohammad Shokri is with the School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran 1439957131, Iran (e-mail: mohammad.shokri.acct@gmail.com).

Hamed Kebriaei is with the School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran 1439957131, Iran and also with School of Computer Science, Institute for Research in Fundamental Sciences (IPM), Tehran 1953833511, Iran (e-mail: kebriaei@ut.ac.ir).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TAC.2021.3071017.

Digital Object Identifier 10.1109/TAC.2021.3071017

making without anticipation of the future [19], [20]. Furthermore, in the proposed model, the communication delays of data exchange among the agents and also from the environment to the agents are taken into account. We prove that under specific assumptions, which are mainly the input-to-state stability (ISS) of environment dynamics and strong monotonicity of augmented subgradient functions, the agents' strategies converge to the unique Nash equilibrium point of the NAG while the dynamics of environment reaches to its equilibrium point.

The contributions of this article can be summarized as follows.

1) We propose a new NAG framework in which the agents with private cost functions compete in an unknown dynamic environment.
2) We consider the time delay in data communication among the neighboring agents and also from the environment to the agents.
3) The conditions in which the proposed dynamic NAG admits a unique Nash equilibrium point are provided.
4) A Nash seeking algorithm based on the subgradient method is proposed, and its convergence to the unique Nash equilibrium point of the dynamic NAG is studied.

Notation: $\mathcal{X}=\prod_{n\in\mathcal{N}}\mathcal{X}_n$ shows Cartesian product of sets $\mathcal{X}_n$. $||x||$ and $x^\top$ are the norm-2 and the transpose of x, respectively. $\mathbf{col}(x_1,\ldots,x_N)=[x_1^\top,\ldots,x_N^\top]^\top$ denotes the column augmentation of vectors x_n. $x\odot y$ is elementwise product of x and y. $\vec{1}_N$ denotes vector of ones. $g(x_0)$ is the subgradient of $f(x):\mathcal{X}\to\mathbb{R}$ at x_0 if $\forall x\in\mathcal{X}: f(x_0)+(x-x_0)^\top g(x_0)\le f(x)$. $g(x):\mathbb{R}^n\to\mathbb{R}^n$ is strongly monotone with C if $\forall x_1,x_2:(g(x_2)-g(x_1))^\top(x_2-x_1)\ge C||x_2-x_1||^2$. $\Pi_{\mathcal{X}}(x)=\arg\min_{y\in\mathcal{X}}||y-x||^2$ shows the projection onto $\mathcal{X}$. $\int_{x_0}^{x}(g(v))^\top dv$ denotes the integral of $g(v):\mathbb{R}^n\to\mathbb{R}^n$.

II. Dynamic NAG

Consider a set $\mathcal{N}$ of N agents who interact with each other in a dynamic environment and aim to minimize their cost functions selfishly. The agents are connected to each other through a directed graph $G(\mathcal{N},\mathcal{E})$ where $\mathcal{E}$ is the set of edges. Defining $\mathcal{N}_n=\{m|(n,m)\in\mathcal{E}\}$ as the set of neighbors to agent $n\in\mathcal{N}$, the cost function of agent n is affected by an aggregate of its neighbors' strategies and the state of environment as follows:

$$J_n(x_n^t,\sigma_n(x_{-n}^t),y^t) \tag{1}$$

where $x_n^t\in\mathcal{X}_n\subset\mathbb{R}^M$ is the strategy of agent n at time t, and $x_{-n}^t=\mathbf{col}(x_1^t,\ldots,x_{n-1}^t,x_{n+1}^t,\ldots,x_N^t)\in\mathcal{X}_{-n}=\prod_{m\neq n}\mathcal{X}_m$. $\sigma_n(x_{-n}^t)$ is the aggregate strategy of the neighbors to agent n, which is defined as

$$\sigma_n(x_{-n}^t)=\sum_{n'\in\mathcal{N}\setminus\{n\}} w_{nn'}x_{n'}^t \tag{2}$$

where $w_{nn'}>0$ if $(n,n')\in\mathcal{E}$, otherwise, $w_{nn'}=0$, and $\sum_{n'\in\mathcal{N}\setminus\{n\}}w_{nn'}=1$ (agent n is considered nonneighbor with itself and $w_{nn}=0$). Also, $y^t\in\mathbb{R}^p$ stands for the environment's state at time t, which evolves with the following dynamics:

$$y^{t+1}=f(y^t,x^t) \tag{3}$$

where $x^t=\mathbf{col}(x_1^t,\ldots,x_N^t)\in\mathcal{X}=\prod_{n\in\mathcal{N}}\mathcal{X}_n$. $y^0\in\mathbb{R}^p$ is considered to be the initial value for the dynamics. The communication scheme among the agents and environment is shown in Fig. 1. Based on the interaction term $\sigma_n(x_{-n}^t)$, the problem can be formulated as the following game:

$$\mathcal{G}=\begin{cases}\text{Players}: \ \mathcal{N}\\ \text{Strategies}: n\in\mathcal{N}: x_n^t\in\mathcal{X}_n\\ \text{Cost Functions}: J_n(x_n^t,\sigma_n(x_{-n}^t),y^t)\\ \text{Environment Dynamics}: y^{t+1}=f(y^t,x^t).\end{cases} \tag{4}$$

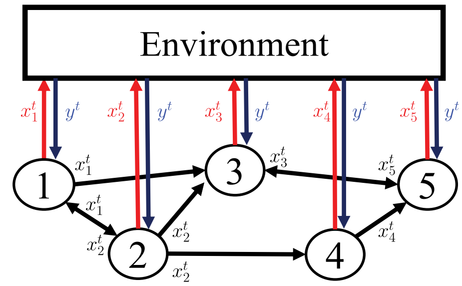

Fig. 1. Information scheme of game $\mathcal{G}$. The communication between the nodes can be unidirectional or bidirectional.

The equilibrium point of $\mathcal{G}$ is defined as follows.

Definition 1: (x^*,y^*) is a Fixed-point Nash Equilibrium (FNE) point of game $\mathcal{G}$ if $\forall n\in\mathcal{N}:\forall x_n\in\mathcal{X}_n$

$$J_n(x_n^*,\sigma_n(x_{-n}^*),y^*)\le J_n(x_n,\sigma_n(x_{-n}^*),y^*) \tag{5}$$

and also, $y^*=f(y^*,x^*)$, where $x^*=\mathbf{col}(x_1^*,\ldots,x_N^*)$ and $x_{-n}^*=\mathbf{col}(x_1^*,\ldots,x_{n-1}^*,x_{n+1}^*,\ldots,x_N^*)$. □

III. Information Structure and Decision Making

In the proposed framework, agents do not have information on the dynamic model of the environment and also the cost functions of other agents. We consider that agent n can observe the decision value of its neighbor agent $m\in\mathcal{N}_n$ with time delay τ_{nm}, denoted by $\tilde{x}_{n,m}^t=x_m^{t-\tau_{nm}}$, and also observes the state of the environment with delay τ_{n0}, denoted by $\tilde{y}_n^t=y^{t-\tau_{n0}}$. Let denote $\tilde{x}_{-n}^t=\mathbf{col}(\tilde{x}_{n,1}^t,\ldots,\tilde{x}_{n,n-1}^t,\tilde{x}_{n,n+1}^t,\ldots,\tilde{x}_{n,N}^t)$ and $z_n^t=\mathbf{col}(\tilde{x}_{n,1}^t,\ldots,\tilde{x}_{n,N}^t)$ where $\forall m\notin\mathcal{N}_n:\tilde{x}_{n,m}^t=x_m^t$. This means that we virtually let agent n receive the decision value of nonneighbor agents with zero delay. Nevertheless, this is made just for the simplicity of notation and according to (1), (2), and (6), such information is not used by agent n for decision making, since $\forall m\notin\mathcal{N}_n$ we have: $w_{nm}=0$.

The agents utilize the projected subgradient method to update their strategies as follows:

$$x_n^{t+1}=\Pi_{\mathcal{X}_n}(x_n^t-\alpha^t\tilde{d}_n^t) \tag{6}$$

where $\tilde{d}_n^t=h_n(x_n^t,\sigma_n(\tilde{x}_{-n}^t),\tilde{y}_n^t)$ is a sub-gradient of the cost function $J_n(x_n^t,\sigma_n(\tilde{x}_{-n}^t),\tilde{y}_n^t)$ with respect to x_n^t, and α^t is the step-size at time t. The procedure of agents' decision making is presented in Algorithm 1. At each time-step t, the agents observe decision value of their neighbors and also the state of environment with some time delay. Having this information, they calculate the subgradient of their cost function (i.e., $\tilde{d}_n^t$). Afterward, they update their decisions based on (6). Also, the state of environment evolves through (3), which is affected by the last decisions of the agents. Finally, the observation of agents from the updated state of environment is utilized for the next round of decision making.

IV. Convergence Analysis

In this section, the uniqueness of FNE of $\mathcal{G}$ and also, the convergence of Algorithm 1 to FNE of $\mathcal{G}$ are investigated under the following assumptions.

Assumption 1: For $n\in\mathcal{N}$, $\mathcal{X}_n$ is convex and compact. □

Assumption 2: α^t is nonincreasing, and it satisfies $\sum_{t=0}^{\infty}\alpha^t=\infty$ and $\sum_{t=0}^{\infty}(\alpha^t)^2<\infty$. □

Assumption 3: There exists finite $\bar{\tau}$ such that $\tau_{nm}<\bar{\tau}$ and $\tau_{n0}<\bar{\tau}$ $\forall n,m\in\mathcal{N}$. □

Algorithm 1: Distributed Sub-Gradient Algorithm for NAGs with Dynamic Environment and Delayed Information Exchange.

Initialize x_n^0, $\tilde{y}_n^0$, and $\tilde{x}_{n,m}^0$ for $n,m \in \mathcal{N}$, $t \leftarrow 0$
Given y^0
Iteration
Agents: $n \in \mathcal{N}$
$\tilde{x}_{n,m}^t \leftarrow x_m^{t-\tau_{nm}}$ for $m \in \mathcal{N}_n$, and $\tilde{y}_n^t \leftarrow y^{t-\tau_{n0}}$
$\tilde{\sigma}_n^t \leftarrow \sigma_n(\tilde{x}_{-n}^t)$
$\tilde{d}_n^t \leftarrow h_n(x_n^t, \tilde{\sigma}_n^t, \tilde{y}_n^t)$
$x_n^{t+1} \leftarrow \Pi_{\mathcal{X}_n}(x_n^t - \alpha^t \tilde{d}_n^t)$
Environment:
$y^{t+1} = f(y^t, x^t)$
$t \leftarrow t+1$

Assumption 4: $J_n(x_n, \sigma_n(x_{-n}), y)$ is convex w.r.t x_n. Further, there exists a bound γ_n such that $\forall x_n \in \mathcal{X}_n, \forall x_{-n} \in \mathcal{X}_{-n}, \forall y \in \mathbb{R}^p$: $||h_n(x_n, \sigma_n(x_{-n}), y)|| \leq \gamma_n$. □

Assumption 5: The augmented subgradient $g(x,y) = \mathbf{col}(h_1(x_1, \sigma_n(x_{-1}), y), \ldots, h_N(x_N, \sigma_N(x_{-N}), y))$ is strongly monotone w.r.t x, and Lipschitz w.r.t x and y, i.e.

$$(g(x,y) - g(x',y))^\top (x - x') \geq C||x - x'||^2 \tag{7a}$$

$$||g(x,y) - g(x',y')|| \leq L_x^a||x - x'|| + L_y^a||y - y'|| \tag{7b}$$

for every $x, x' \in \mathcal{X}$ and $y, y' \in \mathbb{R}^p$. □

Assumption 6: $f(y,x)$ is integrable over y, and contractive w.r.t. y and Lipschitz w.r.t. x, i.e., there exist constants $L_x^e > 0$ and $0 < L_y^e < 1$ such that

$$||f(y,x) - f(y',x')|| \leq L_x^e||x - x'|| + L_y^e||y - y'|| \tag{8}$$

for every $x \in \mathcal{X}$ and $y \in \mathbb{R}^p$. □

Remark 1: The ISS condition in Assumption 6 does not ensure the convergence of Algorithm 1 to FNE point of $\mathcal{G}$, since the fixed point of (3) is affected by the strategies of players (i.e., x) and x is also affected by state y. Hence, there exists a coupling between the Nash strategies of players and a fixed point of dynamics (3). Therefore, the Nash equilibrium point cannot be achieved at any arbitrary fixed point of (3). □

Lemma 1: Given any $y^0 \in \mathbb{R}^p$, there exists a $\mathcal{Y}(y^0)$ such that $\forall t \geq 0 : y^t \in \mathcal{Y}(y^0)$. □

Proof: Let x^p be an arbitrary point in $\mathcal{X}$. Based on the contractivity of the operator $f(y, x^p)$, there exists a unique fixed-point y^p such that $y^p = f(y^p, x^p)$. Also, based on boundedness of $\mathcal{X}$, there exists η such that $||x|| \leq \eta$. Using (8) for (x^t, y^t) and (x^p, y^p), we have $||y^{t+1} - y^p|| \leq L_x^e||x^t - x^p|| + L_y^e||y^t - y^p|| \leq 2L_x^e\eta + L_y^e||y^t - y^p||$. By applying this recursive inequality, $||y^t - y^p|| \leq 2L_x^e\eta\sum_{t'=0}^{t-1}(L_y^e)^{t'} + (L_y^e)^{t-1}r_0 \leq 2L_x^e\eta\sum_{t'=0}^{\infty}(L_y^e)^{t'} + r_0 \leq \frac{2L_x^e\eta}{1-L_y^e} + r_0$ where $r_0 = ||y^0 - y^p||$. Thus, there exists a convex and compact set $\mathcal{Y}(y^0) = \{||y^t - y^p|| \leq \frac{2L_x^e\eta}{1-L_y^e} + r_0\}$ such that $y^t \in \mathcal{Y}(y^0)$. □

Theorem 1: Under Assumptions 4–6, if

$$C(1 - L_y^e) > \left(\frac{L_y^a + L_x^e}{2}\right)^2 \tag{9}$$

$\mathcal{G}$ admits a unique FNE. □

Proof: Consider $J_0(y,x) = \frac{1}{2}y^\top y - F(y,x)$ where $F(y,x) = \int_{y_0}^{y}(f(\nu,x))^\top d\nu$. Hence, the subgradient of $J_0(y,x)$ with respect to y is obtained as $h_0(y,x) = y - f(y,x)$. From (8) and using Cauchy–Schwarz inequality, $(f(y,x) - f(y',x))^\top(y - y') \leq L_y^e||y - y'||^2$. Thus, we have

$$(h_0(y,x) - h_0(y',x))^\top(y - y') \geq (1 - L_y^e)||y - y'||^2. \tag{10}$$

Thus, $J_0(y,x)$ is strongly convex with respect to y. Clearly, the minimizer of $J_0(y,x)$ satisfies $h_0(y,x) = y - f(y,x) = 0$, and hence, it is also a fixed point of $f(\cdot, x)$. Now, consider $q(v) = \mathbf{col}(g(x,y), h_0(y,x))$ and $v = \mathbf{col}(x,y)$. Based on Assumptions 5 and 6, and also, (10), we have

$$\begin{aligned}
&(q(v) - q(v'))^\top(v - v') = (g(x,y) - g(x',y'))^\top(x - x') \\
&+ (h_0(y,x) - h_0(y',x'))^\top(y - y') \\
&= (g(x,y) - g(x',y) + g(x',y) - g(x',y'))^\top(x - x') \\
&+ (h_0(y,x) - h_0(y',x) + h_0(y',x) - h_0(y',x'))^\top(y - y') \\
&\geq C\delta_x^2 - L_y^a\delta_x\delta_y + (1 - L_y^e)\delta_y^2 - L_x^e\delta_x\delta_y \\
&= C\delta_x^2 + (1 - L_y^e)\delta_y^2 - (L_y^a + L_x^e)\delta_x\delta_y
\end{aligned}$$

where $\delta_x = ||x - x'||$, and $\delta_y = ||y - y'||$. Therefore, from the last equality, if $\Phi = (L_y^a + L_x^e)^2 - 4C(1 - L_y^e) < 0$, then $(q(v) - q(v'))^\top(v - v') \geq 0$, which proves the strong monotonicity of $h(v)$. Hence, uniqueness of FNE of $\mathcal{G}$ can be concluded from [19, Th. 2] with condition (9).

Lemma 2: Under Assumption 2 and 4, for (6), $||\nabla^\tau x^t|| \leq \gamma\tau\alpha^t$ where $\nabla^\tau x^t = x^{t+\tau} - x^t$, and $\gamma = \sum_{n\in\mathcal{N}}\gamma_n$. □

Proof: Clearly, $||\nabla^\tau x_n^t|| = ||\sum_{i=0}^{\tau-1} x_n^{t+i+1} - x_n^{t+i}|| \leq \sum_{i=0}^{\tau-1}||x_n^{t+i+1} - x_n^{t+i}||$. From Assumption 4, $||\tilde{d}_n^t|| \leq \gamma_n$. Using (6), Assumption 2, and nonexpansivity of projection operator, we have

$$\begin{aligned}
||\nabla^\tau x_n^t|| &\leq \sum_{i=0}^{\tau-1}||\Pi_{\mathcal{X}_n}(x_n^{t+i} - \alpha^{t+i}\tilde{d}_n^{t+i}) - x_n^{t+i}|| \\
&\leq \sum_{i=0}^{\tau-1}\alpha^{t+i}||\tilde{d}_n^{t+i}|| \leq \gamma_n\sum_{i=0}^{\tau-1}\alpha^{t+i} \leq \gamma_n\tau\alpha^t.
\end{aligned} \tag{11}$$

Thus, $||\nabla^\tau x^t|| \leq \sum_{n\in\mathcal{N}}||\nabla^\tau x_n^t|| \leq \sum_{n\in\mathcal{N}}\gamma_n\tau\alpha^t \leq \gamma\tau\alpha^t$.

Lemma 3: Under the conditions of Lemma 2, for (6), $||z_n^t - x^t|| \leq \gamma\bar{\tau}\alpha^{t-\bar{\tau}}$ for $t \geq \bar{\tau}$ is satisfied. □

Proof: For $t \geq \bar{\tau}$, based on (11) and Assumption 2, $||\tilde{x}_{n,m}^t - x_m^t|| = ||x_m^{t-\tau_{nm}} - x_m^t|| \leq \gamma_m\tau_{nm}\alpha^{t-\tau_{nm}} \leq \gamma_m\bar{\tau}\alpha^{t-\bar{\tau}}$. Thus, $||z_n^t - x^t|| \leq \sum_{m\in\mathcal{N}_n}||\tilde{x}_{n,m}^t - x_m^t|| \leq \bar{\tau}\alpha^{t-\bar{\tau}}\sum_{m\in\mathcal{N}_n}\gamma_m \leq \gamma\bar{\tau}\alpha^{t-\bar{\tau}}$.

Lemma 4: Under the conditions of Lemmas 1 and 2, for $t \geq \bar{\tau}, \forall n \in \mathcal{N}$, the following inequality

$$||\tilde{y}_n^t - y^t|| \leq 2\xi(L_y^e)^{t-\bar{\tau}} + L_x^e\gamma\bar{\tau}\sum_{t'=\bar{\tau}}^{t}(L_y^e)^{t-t'}\alpha^{t'-\bar{\tau}} \tag{12}$$

is satisfied for dynamics (3) where $\xi = \mathbf{max}_{y\in\mathcal{Y}(y^0)}(||y||)$. □

Proof: From Lemma 2, $||x^{t-\tau_{n0}-1} - x^{t-1}|| \leq \sum_{m\in\mathcal{N}}||x_m^{t-\tau_{n0}-1} - x_m^{t-1}|| \leq \tau_{n0}\alpha^{t-\tau_{n0}-1}\sum_{m\in\mathcal{N}}\gamma_m \leq \gamma\bar{\tau}\alpha^{t-\bar{\tau}}$. Based on the fact $\tilde{y}_n^t \in \mathcal{Y}(y^0)$, $||\tilde{y}_n^t - y^t|| \leq ||\tilde{y}_n^t|| + ||y^t|| \leq 2\xi$ for $t \geq 0$. In particular, for $t \geq \bar{\tau}$, by considering (8) with (y^{t-1}, x^{t-1}) and $(y^{t-\tau_{n0}-1}, x^{t-\tau_{n0}-1})$, we have

$$\begin{aligned}
&||\tilde{y}_n^t - y^t|| = ||y_n^{t-\tau_{n0}} - y^t|| \leq L_y^e||y_n^{t-\tau_{n0}-1} - y^{t-1}|| \\
&+ L_x^e||x^{t-\tau_{n0}-1} - x^{t-1}|| \leq L_y^e||y_n^{t-\tau_{n0}-1} - y^{t-1}|| + L_x^e\gamma\bar{\tau}\alpha^{t-\bar{\tau}}.
\end{aligned}$$

Therefore, (12) can be proven straightforwardly by expanding the aforementioned recursive inequality from t to $\bar{\tau}-1$ and using the fact that $||y_n^{t-\tau_{n0}} - y^t|| \leq 2\xi$ for $t = \bar{\tau}-1$.

Lemma 5: Under the conditions of Lemmas 3 and 4, there exists sequence S_n^t such that $\forall t \geq 0$ $\alpha^t ||(\tilde{d}_n^t - d_n^t)^\top \Delta x_n^t|| \leq S_n^t$, and $\sum_{t=0}^{\infty} S_n^t < \infty$ where $\Delta x_n^t = x_n^t - x_n^*$ and $d_n^t = h_n(x_n^t, \sigma_n(x_{-n}^t), y^t)$. □

Proof: Consider $\eta = \mathbf{max}_{x\in\mathcal{X}}(||x||)$. Clearly, $||\Delta x_n^t|| \leq ||x_n^t|| + ||x_n^*|| \leq 2\eta$. Based on definition of $g(.,.)$, $||h_n(x_n, \sigma_n(x_{-n}), y) - h_n(x'_n, \sigma_n(x'_{-n}), y')|| \leq ||g(x,y) - g(x',y')||$. Thus, from (7b) and by knowing $\tilde{x}_n^t = x_n^t$, we have $||\tilde{d}_n^t - d_n^t|| = ||h_n(x_n, \sigma_n(\tilde{x}_{-n}^t), \tilde{y}_n^t) - h_n(x_n^t, \sigma_n(x_{-n}^t), y_n^t)|| \leq ||g(z_n^t, \tilde{y}_n^t) - g(x^t, y^t)|| \leq L_x^a ||z_n^t - x_n^t|| + L_y^a ||\tilde{y}_n - y_n^t||$. Therefore, according to Lemmas 3 and 4, for $t \geq \bar{\tau}$, we have

$$\alpha^t || \left(\tilde{d}_n^t - d_n^t\right)^\top \Delta x_n^t || \leq 2\eta\alpha^t ||\tilde{d}_n^t - d_n^t|| \leq 2\eta L_x^a \alpha^t ||z_n^t - x_n^t|| + 2\eta L_y^a \alpha^t ||\tilde{y}_n - y_n^t|| \leq 2\eta L_x^a \gamma\bar{\tau}\alpha^t\alpha^{t-\bar{\tau}} + 4\eta\xi L_y^a (L_y^e)^{t-\bar{\tau}}\alpha^t + 2\eta L_y^a L_x^e \gamma\bar{\tau}\alpha^t \sum_{t'=\bar{\tau}}^{t} (L_y^e)^{t-t'}\alpha^{t'-\bar{\tau}} \triangleq S_n^t.$$

For $0 \leq t < \bar{\tau}$, from boundedness of $\tilde{d}_n^t$ and d_n^t, it can be considered that $\alpha^t ||(\tilde{d}_n^t - d_n^t)^\top \Delta x_n^t|| \leq 4\eta\gamma_n\alpha^t \triangleq S_n^t$. Hence, $\alpha^t ||(\tilde{d}_n^t - d_n^t)^\top \Delta x_n^t|| \leq S_n^t$ for each $t \geq 0$.

Now, by summation of S_n^t on $t \geq 0$, we have

$$\sum_{t=0}^{\infty} S_n^t = 4\eta\gamma_n \sum_{t=0}^{\bar{\tau}-1} \alpha^t + 2\eta L_x^a \gamma\bar{\tau} \sum_{t=\bar{\tau}}^{\infty} \alpha^t \alpha^{t-\bar{\tau}} + 4\eta\xi L_y^a \sum_{t=\bar{\tau}}^{\infty} (L_y^e)^{t-\bar{\tau}}\alpha^t + 2\eta L_y^a L_x^e \gamma\bar{\tau} \sum_{t=\bar{\tau}}^{\infty} \alpha^t \sum_{t'=\bar{\tau}}^{t} (L_y^e)^{t-t'} \alpha^{t'-\bar{\tau}}.$$

$\sum_{t=\bar{\tau}}^{\infty} \alpha^t \alpha^{t-\tau_{nm}} \leq \frac{1}{2}(\sum_{t=\bar{\tau}}^{\infty} (\alpha^t)^2 + (\alpha^{t-\tau_{nm}})^2) < \infty$. Similarly, $\sum_{t=\bar{\tau}}^{\infty} (L_y^e)^{t-\bar{\tau}}\alpha^t \leq \frac{1}{2}\sum_{t=\bar{\tau}}^{\infty}((L_y^e)^{2t-2\bar{\tau}} + (\alpha^t)^2) < \infty$. By exchanging the order of summation and knowing that α^t is nonincreasing, we have $\sum_{t=\bar{\tau}}^{\infty} \alpha^t \sum_{t'=\bar{\tau}}^{t} (L_y^e)^{t-t'} \alpha^{t'-\bar{\tau}} = \sum_{t=\bar{\tau}}^{\infty} (\sum_{t'=0}^{\infty} \alpha^{t+t'} (L_y^e)^{t'}) \alpha^{t-\bar{\tau}} \leq \sum_{t=\bar{\tau}}^{\infty} (\sum_{t'=0}^{\infty} (L_y^e)^{t'}) (\alpha^{t-\bar{\tau}})^2 = \sum_{t=\bar{\tau}}^{\infty} \frac{(\alpha^{t-\bar{\tau}})^2}{1-L_y^e} < \infty$. Therefore, $\sum_{t=0}^{\infty} S_n^t < \infty$. □

Theorem 2: Under Assumptions 2–6, and (9), Algorithm 1 converges to FNE of the game $\mathcal{G}$ from any initial condition $x_n^0 \in \mathcal{X}_n, \forall n \in \mathcal{N}$ and $y^0 \in \mathcal{Y}(y^0)$. □

Proof: According to [21, Proposition 1.5.8], $x_n^* = \Pi_{\mathcal{X}_n}(x_n^* - \alpha^t d_n^*)$ where $d_n^* = h_n(x_n^*, \sigma_n^*, y^*)$, and $\sigma_n^* = \sigma_n(x_{-n}^*)$. By using Lemma 5, and knowing that the projection operator is nonexpansive, we have

$$||\Delta x_n^{t+1}||^2 \leq ||\Delta x_n^t - \alpha^t \left(\tilde{d}_n^t - d_n^*\right)||^2 = ||\Delta x_n^t||^2 + (\alpha^t)^2 ||\tilde{d}_n^t - d_n^*||^2 - 2\alpha^t \left(\tilde{d}_n^t - d_n^*\right)^\top \Delta x_n^t \leq ||\Delta x_n^t||^2 + 4\gamma_n^2 (\alpha^t)^2 - 2\alpha^t \left(\tilde{d}_n^t - d_n^t\right)^\top \Delta x_n^t - 2\alpha^t \left(d_n^t - d_n^*\right)^\top \Delta x_n^t \leq ||\Delta x_n^t||^2 + 4\gamma_n^2(\alpha^t)^2 + 2S_n^t - 2\alpha^t \left(d_n^t - d_n^*\right)^\top \Delta x_n^t.$$

Therefore, by summation of over $n \in \mathcal{N}$, we get

$$||\Delta x^{t+1}||^2 \leq ||\Delta x^t||^2 + 4\sum_{n\in\mathcal{N}} \gamma_n^2 (\alpha^t)^2 + 2\sum_{n\in\mathcal{N}} S_n^t - 2\alpha^t \left(d^t - d^*\right)^\top \Delta x^t \tag{13}$$

where $d^t = \mathbf{col}(d_1^t, \ldots, d_N^t)$, $d^* = \mathbf{col}(d_1^*, \ldots, d_N^*)$, and $\Delta x^t = \mathbf{col}(\Delta x_1^t, \ldots, \Delta x_N^t)$. From Assumption 5, we have

$$-\left(d^t - d^*\right)^\top \Delta x^t = -\left(g(x^t, y^t) - g(x^*, y^*)\right)^\top \Delta x^t = -\left(g(x^t, y^t) - g(x^*, y^t)\right)^\top \Delta x^t - \left(g(x^*, y^t) - g(x^*, y^*)\right)^\top \Delta x^t \leq -C||\Delta x^t||^2 + L_y^a ||\Delta x^t|| ||\Delta y^t|| \tag{14}$$

where $\Delta y^t = y^t - y^*$. Let us consider the sequence $R^t = ||\Delta x^t|| ||\Delta y^t||$. Then, taking summation over telescopic inequality (13) from time 0 to ∞ and using (14) result

$$\lim_{t\to\infty} ||\Delta x^t||^2 \leq ||\Delta x^0||^2 + 4\sum_{n\in\mathcal{N}} \gamma_n^2 \sum_{t=0}^{\infty} (\alpha^t)^2 + 2\sum_{t=0}^{\infty}\sum_{n\in\mathcal{N}} S_n^t - \sum_{t=0}^{\infty} 2C\alpha^t ||\Delta x^t||^2 + 2L_y^a \sum_{t=0}^{\infty} \alpha^t R^t. \tag{15}$$

Clearly, $\Delta x^t = x_n^{t-1} - x_n^* + x_n^t - x_n^{t-1} = \Delta x^{t-1} + \nabla^1 x^{t-1}$, and $||\Delta y^t|| \leq 2\xi$. Thus, based on Lemma 2, we have

$$R^t = ||\Delta x^{t-1} + \nabla^1 x^{t-1}|| ||\Delta y^t|| \leq ||\Delta x^{t-1}|| ||\Delta y^t|| + ||\nabla^1 x^{t-1}|| ||\Delta y^t|| \leq ||\Delta x^{t-1}|| ||\Delta y^t|| + 2\xi\gamma\alpha^{t-1}. \tag{16}$$

Based on (8), by putting (y^{t-1}, x^{t-1}) and (y^*, x^*) into (y, x) and (y', x'), respectively, it can be deduced that

$$||\Delta y^t|| = ||f(y^{t-1}, x^{t-1}) - f(y^*, x^*)|| \leq L_x^e ||\Delta x^{t-1}|| + L_y^e ||\Delta y^{t-1}||. \tag{17}$$

Therefore

$$R^t \leq ||\Delta x^{t-1}|| \left(L_x^e ||\Delta x^{t-1}|| + L_y^e ||\Delta y^{t-1}||\right) + 2\gamma\xi\alpha^{t-1} = L_x^e ||\Delta x^{t-1}||^2 + L_y^e R^{t-1} + 2\gamma\xi\alpha^{t-1}. \tag{18}$$

By applying (18) recursively from $t' = 0$ to $t' = t$, we have

$$R^t \leq (L_y^e)^t R^0 + 2\gamma\xi \sum_{t'=0}^{t} (L_y^e)^{t-t'} \alpha^{t'} + L_x^e \sum_{t'=0}^{t} (L_y^e)^{t-t'} ||\Delta x^{t'}||^2$$

and hence

$$\sum_{t=0}^{\infty} \alpha^t R^t \leq \underbrace{R^0 \sum_{t=0}^{\infty} (L_y^e)^t \alpha^t + 2\gamma\xi \sum_{t=0}^{\infty} \alpha^t \sum_{t'=0}^{t} (L_y^e)^{t-t'} \alpha^{t'}}_{=T_1} + \underbrace{L_x^e \sum_{t=0}^{\infty}\sum_{t'=0}^{t} \alpha^t (L_y^e)^{t-t'} ||\Delta x^{t'}||^2}_{T_2}. \tag{19}$$

$\sum_{t=0}^{\infty} (L_y^e)^t \alpha^t \leq \frac{1}{2}\sum_{t=0}^{\infty}((L_y^e)^{2t} + (\alpha^t)^2) < \infty$. Furthermore, by replacing the order of summations, and knowing that α^t is decreasing, we have $\sum_{t=0}^{\infty} \alpha^t \sum_{t'=0}^{t} (L_y^e)^{t-t'} \alpha^{t'} = \sum_{t'=0}^{\infty} (\sum_{t=0}^{\infty} \alpha^{t'+t} (L_y^e)^t) \alpha^{t'} \leq \sum_{t'=0}^{\infty} (\sum_{t=0}^{\infty} (L_y^e)^t)(\alpha^{t'})^2 \leq \sum_{t'=0}^{\infty} \frac{(\alpha^{t'})^2}{1-L_y^e} < \infty$. Therefore, the term T_1 is bounded. In a same way, by replacing the order of summations of term T_2, and knowing that α^t is decreasing, it can be concluded that

$$\sum_{t=0}^{\infty}\sum_{t'=0}^{t} \alpha^t (L_y^e)^{t-t'} ||\Delta x^{t'}||^2 = \sum_{t=0}^{\infty} \left(\sum_{t'=0}^{\infty} \alpha^{t+t'} (L_y^e)^{t'}\right) ||\Delta x^t||^2 \leq \sum_{t=0}^{\infty} \left(\sum_{t'=0}^{\infty} (L_y^e)^{t'}\right) \alpha^t ||\Delta x^t||^2 = \sum_{t=0}^{\infty} \frac{\alpha^t ||\Delta x^t||^2}{1 - L_y^e}. \tag{20}$$

Thus, by putting (19) and (20) into (15), we have

$$\lim_{t\to\infty}||\Delta x^t||^2 \leq ||\Delta x^0||^2 + T_1 + 4\underbrace{\sum_{n\in\mathcal{N}}\gamma_n^2\sum_{t=0}^{\infty}(\alpha^t)^2}_{T_3}$$
$$+\,2\underbrace{\sum_{t=0}^{\infty}\sum_{n\in\mathcal{N}}S_n^t}_{T_4} - \underbrace{\sum_{t=0}^{\infty}2(C-\frac{L_y^aL_x^e}{1-L_y^e})\alpha^t||\Delta x^t||^2}_{T_5}. \tag{21}$$

Based on average inequality, $(\frac{L_y^a+L_x^e}{2}) \geq \sqrt{L_y^aL_x^e}$. Thus, from (9), $T_5 \geq 0$. Based on Lemma 5, T_4 is bounded. Because of the boundedness of T_1, T_3, and T_4, the term T_5 must be bounded. Hence, based on Assumption 2 and boundedness of T_5, $||\Delta x^t||^2$ must converge to 0 as $t \to \infty$. This results x^t converges to x^* as $t \to \infty$.

As for dynamic (3), x^t can be considered as an input that goes to x^* as $t \to \infty$. Hence, to prove the convergence of y^t to y^*, we consider the Lyapunov function $V(y^t) = ||\Delta y^t||$. Then, from (17), it can be concluded that $V(y^t) \leq L_y^e||\Delta y^{t-1}|| + L_x^e||\Delta x^{t-1}|| = V(y^{t-1}) - (1-L_y^e)||\Delta y^{t-1}|| + L_x^e||\Delta x^{t-1}||$. Thus, since $(1-L_y^e)||\Delta y^{t-1}||$ and $L_x^e||\Delta x^{t-1}||$ are $\mathcal{K}$-class functions, based on [22, Lemma 3.13], it can be deduced that $\lim_{t\to\infty}||\Delta y^t|| \leq \frac{L_x^e}{1-L_y^e}\lim_{t\to\infty}||\Delta x^t|| = 0$. As a result, y^t converges to y^* as $t \to \infty$. □

V. Simulation Results

As an application, the rate allocation and service selection in a wireless cellular network is studied [4]. Consider a network with N service providers (SPs) with set $\mathcal{N}$ who provide radio signal coverage M users, which subscribe each SP to connect with. In this network, the coverage region of SPs can overlap with each other, which causes the signal interference between SPs. Based on the Shannon formula, the rate of data transfer would be diminished due to the signaling of neighbors of each SP. The aggregative influence of neighboring SPs. The objective function of SP n can be modeled as follows [23]:

$$J_n(x_n^t, x_{-n}^t, y^t) = A_n Ln\left(1 + \frac{a_n x_n^t}{N_0 + \sum_{m\in\mathcal{N}_n} b_{nm}x_m^t}\right) - C_n(\eta M y_n^t - B_n\mu_n(1-x_n^t))^2 \tag{22}$$

where $y_n^t \in [0,1]$, and $x_n^t \in [0,1]$ are the subscribed user proportion and the bandwidth ratio of SP n at time t, respectively. $x_{-n}^t = \mathbf{col}(x_1^t,\ldots,x_{n-1}^t,x_{n+1}^t,\ldots,x_N^t)$, and $\mathcal{N}_n$ indicates the set of neighbors of SP n. Also, N_0, η, a_n, b_{nm}, A_n, and C_n are some positive constants. The first and second terms in (22) indicate the signal-to-interference-noise ratio, and performance discrepancy of SP n. The evolution of proportion of users who subscribe SP n is modeled at time t by the following replicator dynamic [24]:

$$y^{t+1} = y^t + \delta y^t \odot (\pi(y^t,x^t) - \bar{\pi}(y^t,x^t))$$
$$\pi_n(y_n^t, x_n^t) = \frac{B_n\mu_n(1-x_n^t)}{p_n M y_n^t} \tag{23}$$

where $\pi_n(y_n^t, x_n^t)$ is offered payoff function of SP n to users. Also, $x^t = \mathbf{col}(x_1^t,\ldots,x_N^t)$, $y^t = \mathbf{col}(y_1^t,\ldots,y_N^t)$, $\pi(y^t,x^t) = \mathbf{col}(\pi_1(y_1^t,x_1^t),\ldots,\pi_N(y_n^t,x_n^t)), \bar{\pi}(y^t,x^t) = (y^\top\pi(y^t,x^t))\vec{1}_N$, and δ is the time step. Based on SPs' offer, each user selects an SP that wants to communicate with. Furthermore, B_n, μ_n, and p_n are the maximum bandwidth, the achieved throughput, and service price of SP n, respectively. Considering SPs as the agents in our framework, they are involved in a NAG based on (22). Moreover, since the cost function of the SPs are influenced by the users' replicator dynamics (23), the whole problem can be modeled as a dynamic NAG. Consequently, the proposed game formulation (4) can be applied to the rate allocation problem for SPs.

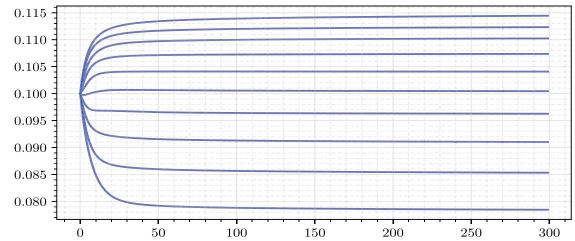

Fig. 2. Service selection of users.

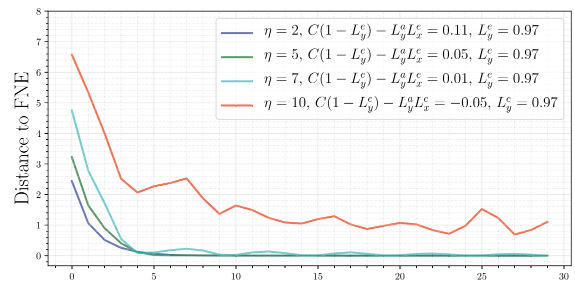

Fig. 3. Distance of SPs' strategies to FNE for different values of η.

For the numerical simulation, we consider a network of 10 SPs and 40 000 users. SPs are located randomly in a squared region with 8-km width and a pair of them is considered as neighbors, if the distance between them are lesser than 2˜km. Each SP has the following parameters: $\eta = 10$, $A_n = 80$, $C_n = 10^{-5}$, $p_n = 0.1$, and $\mu_n = 10^{-4}$. B_n of SPs are 6˜MHz, 12˜MHz,..., and 60˜MHz. Also, $N_0 = 4\times 10^{-21}$ and $\alpha^t = \frac{1}{t+10}$. w_{nm} is set reversely proportional to the distance between SP n and Sp m. All delay parameters τ_{nm} and τ_{n0} are considered to be selected from a Poisson random distribution with parameter 1. As shown in Fig. 2, at the first iteration, users are uniformly distributed to subscribe different SPs (10% for each SP), and finally, the proportion of subscribed users to different SPs converges to $y^* = [0.078, 0.085, 0.090, 0.096, 0.100, 0.104, 0.107, 0.110, 0.112, 0.114]$. To clarify Remark 1, the distance of SPs' strategies to the FNE point is depicted Fig. 3 by changing η. Based on (23), the constants in (9) are

$$L_y^e = 1 - \sum_{n\in\mathcal{N}}\frac{\delta B_n\mu_n}{p_nM}, L_x^e = \max\left(\frac{\delta B_n\mu_n}{p_nM}\right) + \sum_{n\in\mathcal{N}}\frac{\delta B_n\mu_n}{p_nM}$$
$$L_y^a = 2\eta M\max(C_n), C = \min\left(2C_nB_n^2\mu_n^2 - \frac{A_na_n^2}{(N_0+a_n)^2}\right).$$

As it is shown, by increasing the value of η, although (23) satisfies the ISS condition, however, since in the red plot, the condition in (9) is not satisfied, the strategies of SPs do not converge to the FNE point.

VI. Conclusion

In this article, NAGs were studied in a dynamic environment. The cost functions of the agents were influenced by not only their neighboring agents but also the state of an environment. Furthermore, the environment's state dynamics was affected by strategies of the agents. Letting the model of dynamics be unknown to the agents, a distributed algorithm was proposed that converges the unique FNE point in the presence of delayed information. To evaluate the proposed algorithm, the rate allocation and service selection in a wireless cellular network were investigated as a case study.

References

[1] Y. Bramoullé and R. Kranton, "Games played on networks," in *The Oxford Handbook of the Economics of Networks*, Y. Bramoullé, A. Galeotti, and B. Rogers, Eds. London, U.K.: Oxford Univ. Press, Apr. 2016, pp. 82–112.

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

[3] J. Ghaderi and R. Srikant, "Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate," *Automatica*, vol. 50, pp. 3209–3215, 2014.

[4] E. H. K. Zhu and D. Niyatoo, "Pricing, spectrum sharing, and service selection in two-tier small cell networks: A hierarchical dynamic game approach," *IEEE Trans. Mobile Comput.*, vol. 13, no. 8, pp. 857–869, Aug. 2014.

[5] Y. Bramoullé and R. Kranton, "Public goods in networks," *J. Econ. Theory*, vol. 135, no. 1, pp. 478–494, 2007.

[6] C. Zhao, U. Topcu, N. Li, and S. Low, "Design and stability of load-side primary frequency control in power systems," *IEEE Trans. Autom. Control*, vol. 59, no. 5, pp. 1177–1189, May 2014.

[7] F. Parise, B. Gentile, S. Grammatico, and J. Lygeros, "Network aggregative games: Distributed convergence to Nash equilibria," in *Proc. 54th IEEE Conf. Decis. Control*, Dec. 2015, pp. 2295–2300.

[8] F. Parise and A. Ozdaglar, "A variational inequality framework for network games: Existence, uniqueness, convergence and sensitivity analysis," *Games Econ. Behav.*, vol. 114, pp. 47–82, Mar. 2019.

[9] C. Cenedese, G. Belgioioso, S. Grammatico, and M. Cao, "Time-varying constrained proximal type dynamics in multi-agent network games," *Eur. Control Conf. (ECC)*, St. Petersburg, Russia, 2020, pp. 148–153,.

[10] S. Grammatico, "Proximal dynamics in multiagent network games," *IEEE Control Netw. Syst.*, vol. 5, no. 4, pp. 1707–1716, Dec. 2018.

[11] M. Shokri and H. Kebriaei, "Leader-follower network aggregative game with stochastic agents' communication and activeness," *IEEE Trans. Autom. Control*, vol. 65, no. 12, pp. 5496–5502, Dec. 2020.

[12] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, 2nd ed. Philadelphia, PA, USA: Soc. Ind. Appl. Math., 1998.

[13] J. Engwerda, *LQ Dynamic Optimization and Differential Games*. Hoboken, NJ, USA: Wiley, 2005.

[14] T. Başar and Q. Zhu, "Prices of anarchy, information, and cooperation in differential games," *Dyn. Games Appl.*, vol. 1, no. 1, pp. 50–73, Mar. 2011.

[15] H. Kebriaei and L. Iannelli, "Discrete-time robust hierarchical linear-quadratic dynamic games," *IEEE Trans. Autom. Control*, vol. 63, no. 3, pp. 902–909, Mar. 2018.

[16] Q. Wei, R. Song, and P. Yan, "Data-driven zero-sum neuro-optimal control for a class of continuous-time unknown nonlinear systems with disturbance using ADP," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 2, pp. 444–458, Feb. 2016.

[17] H. Zhang, H. Jiang, C. Luo, and G. Xiao, "Discrete-time nonzero-sum games for multiplayer using policy-iteration-based adaptive dynamic programming algorithms," *IEEE Trans. Cybern.*, vol. 47, no. 10, pp. 3331–3340, Oct. 2017.

[18] H. Jiang, H. Zhang, G. Xiao, and X. Cui, "Data-based approximate optimal control for nonzero-sum games of multi-player systems using adaptive dynamic programming," *Neurocomputing*, vol. 275, pp. 192–199, 2018.

[19] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave n-person games," *Econometrica*, vol. 33, no. 3, pp. 520–534, 1965.

[20] C. Camerer, "Bounded rationality in individual decision making," *Exp. Econ.*, vol. 1, no. 2, pp. 163–183, 1998.

[21] J.-S. P. F. Facchinei and J.-S. Pang, *Finite Dimensional Variational Inequalities and Complementarity Problems*, vol. 1, 1st ed. Berlin, Germany: Springer, 2003.

[22] Z.-P. Jiang and Y. Wang, "Input-to-state stability for discrete-time nonlinear systems," *Automatica*, vol. 37, no. 6, pp. 857–869, 2001.

[23] P. Semasinghe, E. Hossain, and S. Maghsudi, "Cheat-proof distributed power control in full-duplex small cell networks: A repeated game with imperfect public monitoring," *IEEE Trans. Commun.*, vol. 66, no. 4, pp. 1787–1802, Apr. 2018.

[24] X. Wang, B. Liu, X. Su, X. Xu, and L. Xiao, "Evolutionary game based heterogeneous wireless network selection with multiple traffics in 5G," in *Proc. 26th Int. Conf. Telecommun.*, Apr. 2019, pp. 80–84.