No-Regret Distributed Learning in Subnetwork Zero-Sum Games

Shijie Huang, Jinlong Lei, Member, IEEE, Yiguang Hong, Fellow, IEEE, Uday V. Shanbhag, and Jie Chen, Fellow, IEEE

Abstract—In this article, we consider a distributed learning problem in a subnetwork zero-sum game, where agents are competing in different subnetworks. These agents are connected through time-varying graphs where each agent has its own cost function and can receive information from its neighbors. We propose a distributed mirror descent algorithm for computing a Nash equilibrium and establish a sublinear regret bound on the sequence of iterates when the graphs are uniformly strongly connected and the cost functions are convex–concave. Moreover, we prove its convergence with suitably selected diminishing step sizes for a strictly convex–concave cost function. We also consider a constant step-size variant of the algorithm and establish an asymptotic error bound between the cost function values of running average actions and a Nash equilibrium. In addition, we apply the algorithm to compute a mixed-strategy Nash equilibrium in subnetwork zero-sum finite-strategy games, which have merely convex–concave (to be specific, multilinear) cost functions, and obtain a final-iteration convergence result and an ergodic convergence result, respectively, under different assumptions.

Index Terms—Distributed mirror descent, no-regret distributed learning, subnetwork zero-sum game.

I. INTRODUCTION

A NONCOOPERATIVE game is a framework for modeling conflict situations among agents with individual costs. Specifically, it studies the interaction of agents whose decisions are affected by the actions of others. Zero-sum games represent an important class of noncooperative games where the gain of one agent is equivalent to the loss of another. Some complex decision-making problems, such as power allocation in wireless communication [1], robust portfolio selection in finance [2], and robust matched filtering in signal processing [3], may be modeled by generalized zero-sum games, named as “subnetwork zero-sum” [4], where two subnetworks of agents are engaged in a zero-sum game. A core solution concept in noncooperative games is the Nash equilibrium (NE), and NE seeking problems have been extensively studied [5], [6], [7], [8].

Conventional NE seeking algorithms mainly address scenarios with full-decision information, where each agent may observe the actions of all its rivals. However, in practice, agents may have to make decisions based only on limited information, suggesting the development of distributed NE seeking algorithms. For example, in [9], distributed synchronous and asynchronous algorithms with diminishing step sizes have been provided, along with an error bound for a constant step size for strictly aggregative games. Furthermore, various fully distributed generalized NE seeking algorithms, based on operator splitting schemes, have been further designed for games with coupled constraints [10], [11], [12]. Specifically, Belgioioso et al. [13] have offered a comprehensive survey for operator-theoretic approaches. Moreover, in [14] and [15], a regularized distributed algorithm for merely monotone aggregative games and a gradient-free distributed algorithm for convex games with limited knowledge of cost function have been proposed, respectively. Distinct from these directions, a linearly convergent distributed gradient-response scheme for stochastic aggregative games has been introduced in [16]. In addition, continuous-time distributed algorithms via consensus-based approaches have also been analyzed in [17] and [18].

Much of the aforementioned work focused on the convergence of those proposed algorithms to an NE. Within game-theoretic learning algorithms, an important class of dynamics, known as no-regret dynamics [7], is dedicated to the online learning process. The no-regret framework for games has received a great amount of research attention due to its simplicity and celebrated connections with game-theoretic equilibrium concepts.
For finite games, no-regret learning algorithms are guaranteed to converge to a set of coarse correlated equilibria, whereas the connection between no-regret learning and NE remains limited. In fact, there exist no-regret algorithms that do not converge to an NE [19]. Fortunately, partial convergence results can be established for certain classes of games. For instance, in [20], Nesterov’s excessive gap technique was employed to propose a near-optimal no-regret algorithm, ensuring that the cost value converges to that of an NE for two-player zero-sum games with finite action sets. Subsequently, Rakhlin and Sridharan [21] enhanced the convergence rate by introducing a modified optimistic mirror descent algorithm. Similarly, Kangarshahi et al. [22] further proposed an optimal no-regret algorithm with a rate of $O\left(\frac{1}{T}\right)$. In addition to the convergence of cost value, sufficient conditions were provided in [23], under which the strategies generated by no-regret learning converges to an NE of variationally stable game, and the efforts in [24] led to an asynchronous no-regret learning algorithm for games with lossy feedback.

However, much of prior research has primarily concentrated on centralized no-regret learning algorithms. While numerous distributed NE seeking algorithms exist, it remains unclear whether these algorithms are no-regret schemes [24]. In fact, the only result on distributed online games that we are aware of was provided in [25]. Inspired by the distributed online optimization problem [26], [27], [28], [29], Lu et al. [25] introduced an online distributed algorithm to track the generalized NE in dynamic environments, establishing a sublinear regret bound. Unlike no-regret learning in games under consideration, Lu et al. [25] assumed a time-varying cost function, with the offline benchmark of the regret being the cost value of an NE.

The motivation of this article is to design effective no-regret distributed algorithms for subnetwork zero-sum games considered in [4] and [30], where agents in the same subnetwork collaborate for consensus, while playing antagonistic roles with the agents in the other subnetwork. The contributions of this article are summarized as follows.

1) We propose a distributed learning algorithm based on a mirror descent scheme and derive a regret bound for the algorithm, demonstrating its no-regret property under diminishing and constant step sizes. Moreover, we prove convergence to the unique NE in cases where the cost function is strictly convex–concave. To the best of our knowledge, there has been no theoretical result on no-regret distributed algorithms for computing an NE of subnetwork zero-sum games yet.

2) Although the proposed algorithm does not guarantee convergence to an NE under constant step sizes, we obtain an error bound akin to that in [31] showing that the running average actions generated by the algorithm provide approximate solutions to the NE seeking problem.

3) We further apply the algorithm to compute a mixed-strategy NE of subnetwork zero-sum finite-strategy games. We establish its final-iteration and ergodic convergence, different from [23], which provides an ergodic convergence result for finite two-person zero-sum games.

A preliminary version of our work appeared at the 2021 IEEE Conference on Decision and Control [32]. The current article makes the following extensions. First, we consider a broader class of subnetwork zero-sum games with continuous strategy sets and generalize the algorithm to a distributed mirror descent scheme. Second, we prove the convergence of the iterates to the unique NE for a strictly convex–concave cost function and establish an asymptotic error bound for a constant step-size variant of the generalized algorithm. Third, more simulations are carried out to discuss how algebraic connectivity of the communication network influences the regret bound. Finally, we provide all proofs omitted in the conference paper.

The rest of this article is organized as follows. In Section II, we formulate the no-regret distributed learning problem in subnetwork zero-sum games and propose a distributed mirror descent algorithm, while in Section III, we establish several useful lemmas and further give a regret bound analysis of the proposed algorithm. In Section IV, we prove that the algorithm converges to the NE under diminishing step sizes, and with constant step sizes, we establish an asymptotic error bound for the cost value of the averaged iterates. Then, in Section V, we apply the proposed algorithm to subnetwork zero-sum finite-strategy games. In Section VI, we provide simulations to verify our theoretical analysis. Finally, Section VII concludes this article.

Notations: Denote by $\mathbb{R}^n$ the $n$-dimensional real Euclidean space. For column vectors $x, y \in \mathbb{R}^n$, $\langle x, y \rangle$ denotes the inner product and $\|x\|_p$ denotes the $L_p$ norm. For $x \in \mathbb{R}^n$, $B(x, \epsilon)$ denotes a ball with $x$ the center and $\epsilon > 0$ the radius. For a norm $\|\cdot\|_1$ on $\mathbb{R}^n$, $\|y\|_1 : \sup \{\langle y, x \rangle : \|x\| \leq 1\}$ denotes the dual norm. $A_{ij}$ denotes the element in the $i$th row and $j$th column of matrix $A$. For a function $f(x_1, \ldots, x_N)$, denote $\nabla f$ as the gradient of $f$ and $\partial f$ as the subdifferential of $f$ with respect to $x_i$. A function $f(x_1, x_2)$ is said to be (strictly) convex–concave (concave–convex) if $f(x_1, x_2)$ is (strictly) convex (concave) in $x_1$ for any $x_2$ and (strictly) concave (convex) in $x_2$ for any $x_1$. A digraph (directed graph) is characterized by $G = (V, E)$, where $V = \{1, \ldots, n\}$ is the set of nodes and $E \subset V \times V$ is the set of edges, where $(j, i) \in E$ if agent $i$ can obtain information from agent $j$. Associated with graph $G$, there is a diagonal matrix $W = [w_{ij}] \in \mathbb{R}^{n \times n}$ with nonnegative elements, which satisfy that $w_{ij} > 0$ if and only if $(j, i) \in E$. A path from $i_1$ to $i_p$ is an alternating sequence $i_1 e_{i_1i_2} \cdots i_{p-1} e_{i_{p-1}i_p}$ such that $e_r = (i_r, i_{r+1}) \in E$ ($r = 1, \ldots, p - 1$). A digraph is strongly connected if there is a path between any pair of distinct nodes.

II. PROBLEM FORMULATION AND ALGORITHM

In this section, we formulate the no-regret distributed learning problem in a subnetwork zero-sum game and propose a distributed mirror descent algorithm.

A. Problem Formulation

Consider a zero-sum game between two subnetworks $\Sigma_1$ and $\Sigma_2$, composed of agents $\mathcal{V}_1 \triangleq \{v_1^1, \ldots, v_1^{n_1}\}$ and $\mathcal{V}_2 \triangleq \{v_2^1, \ldots, v_2^{n_2}\}$, respectively. Assume that the agents in the subnetwork $\Sigma_1$ and the agents in $\Sigma_2$ are, respectively, connected by time-varying directed graph sequences $G_1(k)$ and $G_2(k)$, where $G_1(k) = (\mathcal{V}_1, \mathcal{E}_1(k))$, $G_2(k) = (\mathcal{V}_2, \mathcal{E}_2(k))$, and $\mathcal{V}_1 \cap \mathcal{V}_2 = \emptyset$.\footnote{Unless otherwise specified, we use $\|\cdot\|$ to represent any possible $L_p$ $(p \geq 1)$ norm in this article.}
The interaction between $\Sigma_1$ and $\Sigma_2$ is modeled by a bipartite network $\Sigma_{12}$ (see Fig. 1). Here, $\Sigma_{12}$ is described by a time-varying graph sequence $\tilde{G}_{12}(t) = (V_1 \cup V_2, \tilde{E}_{12}(t))$, which means that $\tilde{E}_{12}(t) \subset \{(v_l^1, v_l^2) \mid v_l^1 \in V_1, v_l^2 \in V_{3-i}, t \in \{1, 2\}\}$. For $t = 1, 2$, and $v_l^1 \in V_1$, denote $N_l^t(v_l^1) = \{v_l^1 \mid (v_l^1, v_l^2) \in \tilde{E}_{12}(t)\}$ and $N_{12,l}^t(\cdot) = \{v_l^2 \mid (v_l^2, v_l^1) \in \tilde{E}_{12}(t)\}$ as its neighbors in $V_1$ and $V_{3-i}$ at time $t$, respectively.

For each $t \in \{1, 2\}$, the action set of $\xi_i$ is denoted by $X_i \subset \mathbb{R}^{m_i}$. Each subnetwork $\Sigma_i, i \in \{1, 2\}$, aims to choose an action $x_i \in X_i$ to minimize the following global cost:

$$f_l(x_1, x_2) = \frac{1}{n_l} \sum_{i=1}^{n_l} f_{l,i}(x_1, x_2), \quad l \in \{1, 2\}$$

which is the average of its agents’ costs at each node $v_l^i$, denoted by $f_{l,i}$. The subnetworks are engaged in a zero-sum game, namely, for any $x_i \in X_i, i \in \{1, 2\}$, we have

$$f_l(x_1, x_2) + f_{l,ij}(x_1, x_2) = 0.$$ 

The following concept is well known [30].

**Definition 1:** An action profile $x^* = (x_1^*, x_2^*)$ is an NE of the subnetwork zero-sum game if

$$x_1^* \in \operatorname{arg} \min_{x_1 \in \mathcal{X}_1} U(x_1, x_2^*)$$

and

$$x_2^* \in \operatorname{arg} \max_{x_2 \in \mathcal{X}_2} U(x_1^*, x_2)$$

where $U(x_1, x_2) \triangleq f_1(x_1, x_2)$.

We make the following assumptions on the action sets and the cost functions $f_{l,i}$.

**Assumption 1:** i) The sets $\mathcal{X}_1$ and $\mathcal{X}_2$ are compact and convex.

ii) For each agent $v_l^i \in V_1$, the cost $f_{l,ij}(\cdot, \cdot)$ is convex–concave over $\mathcal{X}_1 \times \mathcal{X}_2$. Similarly, for each agent $v_l^i \in V_2$, the cost $f_{l,ji}(\cdot, \cdot)$ is concave–convex over $\mathcal{X}_1 \times \mathcal{X}_2$.

iii) For each $v_l^i \in V_1, f_{l,ij}(x_1, x_2)$ in $L_{1,1}$-Lipschitz continuous in $x_1 \in X_1$ for any $x_2 \in X_2$ and $L_{1,2}$-Lipschitz continuous in $x_2 \in X_2$ for any $x_1 \in X_1$, i.e.,

$$|f_{l,ij}(x_1, x_2) - f_{l,ij}(x_1', x_2)| \leq L_{1,1} |x_1 - x_1'| \quad \forall x_2 \in X_2$$

and

$$|f_{l,ij}(x_1, x_2) - f_{l,ij}(x_1, x_2')| \leq L_{1,2} |x_2 - x_2'| \quad \forall x_1 \in X_1.$$ 

Similarly, for each $v_l^i \in V_2, f_{l,ji}(x_1, x_2)$ in $L_{2,1}$-Lipschitz continuous in $x_2 \in X_2$ for any $x_1 \in X_1$ and $L_{2,2}$-Lipschitz continuous in $x_1 \in X_1$ for any $x_2 \in X_2$.

Assumption 1 ensures the existence of an NE (see [30, Th. 2.5]). Moreover, if the cost function $U$ is further assumed to be strictly convex–concave, the NE is unique.

**Lemma 1:** Under Assumption 1, there exists an NE for the considered subnetwork zero-sum game.

In the two subnetworks, each agent only knows its own cost function. Within each subnetwork, agents can exchange information with their neighbors, i.e., $v_l^i$ can pass information to $v_l^j$ at time $t$ if $v_l^j \in N_{12,l}^t$. Moreover, each subnetwork can also obtain information about the other subnetwork via $\tilde{G}_{12}(t)$, i.e., $v_{l-1}^i$ can pass the information of $\Sigma_{3-i}$ to $v_l^i$ at time $t$ if $v_{l-1}^i \in N_{12,l}^t$. For brevity, we use $i \in \mathcal{V}_l$ to represent the agent $v_l^i$ later when there is no confusion. We make the following assumption on the communication networks.

**Assumption 2:** For $l = 1, 2$, the graph sequence $G_l(t)$ is uniformly jointly strongly connected (i.e., there exists a positive integer $B_l$ such that $\bigcup_{k \geq 0} [t + k + B_l - 1] G_l(t)$ is strongly connected for $k \geq 0$) and every agent in $\gamma_l$ has at least one neighbor in $\Sigma_{3-i}$ for all $t$. Moreover, the associated adjacency matrices $W_l(t) = [w_{l,ij}(t)]$ and $W_{12}(t) = [w_{12,ij}(t)]$ in $i)$ There exists a scalar $\eta \in (0, 1)$ such that $w_{l,ij}(t) \geq \eta$ when $j \in N_l^t(\cdot)$, and $w_{l,ij}(t) = 0$ otherwise; the same for $w_{12,ij}(t)$.  

ii) $\sum_{j=1}^{n_l} w_{l,ij}(t) = \sum_{j=1}^{n_l} w_{l,ij}(t) = 1$. 

iii) $\sum_{j=1}^{n_l} w_{l,ij}(t) = 1, i \in \mathcal{V}_l$.

The weight rules (i)-(iii) have been widely used in distributed optimization [33, 34] and distributed NE seeking [4, 9]. In addition, the assumption that every agent in $\Sigma_l$ has at least one neighbor in $\Sigma_{3-i}$ for all $t$ is required to ensure that every agent in $\Sigma_l$ can receive the information about $\Sigma_{3-i}$ at every time.

Suppose that at every time $t$, agent $i$ in $\mathcal{V}_l \subset \{1, 2\}$ maintains an estimate of its subnetwork’s action as $x_{l,i}(t)$ and receives the information from its adversarial subnetwork. In short, the information structure of each agent $i \in \mathcal{V}_l$ is

$$\{x_{l,i}(t), x_{l,j}(t), j \in N_l^t(i), x_{3-i,k}(t), k \in N_{12,l}(i)\} \leq t$$

where $x_{l,j}(t) \in N_l^t(i)$ is the information received from its own subnetwork $\Sigma_j$ and $x_{3-i,k}(t) \in N_{12,l}(i)$ is the information received from the adversary subnetwork $\Sigma_{3-i}$ at time $s$. For clarity, we mainly focus on subnetwork $\Sigma_1$ in the subsequent formulation. At time $t$, each agent $i$ in $\gamma_l$ receives the information about $\Sigma_2$ from the agents $j \in N_{12,l}^t(i)$ and forms an estimation for $\Sigma_2$’s action, which is denoted by

$$u_{l,i}(t) \triangleq \sum_{j \in N_{12,l}^t(i)} w_{l,ij}(t) x_{l,j}(t).$$

Then, agent $i$ obtains a cost $f_{l,1}(x_{l,1}(t), u_{l,2}(t))$. After $T$ time steps, the regret of $\Sigma_1$ associated with agent $i$ is defined as

$$R_{1,i}(T) = \sum_{t=1}^{T} f_l(1, u_{l,2}(t)) - \min_{x_2 \in \mathcal{X}_2} \sum_{t=1}^{T} f_l(x_1, u_{l,2}(t))$$

i.e., the maximum gain $\Sigma_1$ could be achieved by playing the single best fixed action in case the estimated sequence of $\Sigma_2$’s actions $\{u_{l,2}(t)\}_{t=1}^{T}$ and the cost functions were known in hindsight. For the case of infinite horizon ($T$ is unknown), an algorithm is no-regret for $\Sigma_1$ if for all $i$, $R_{1,i}(T) / T \to 0$ as $T \to \infty$. In the case of finite horizon ($T$ is known in advance), an algorithm is no-regret if we may choose a sufficiently large $T$ to make the average regret $R_{1,i}(T) / T$ small enough. It is desirable for each agent to adopt a no-regret learning algorithm since the cumulative cost produced by the action sequence employed approaches the minimum cumulative cost it could achieve when all observable information is known in advance. The goal of this article is to design a no-regret distributed learning algorithm that converges to an NE.

**Remark 1:** Different from the regret definition in the time-varying game model explored in [25], which selects NE as the comparator, our definition (3) relies solely on the actions.
taken by agents during the learning process. Moreover, while the definition in [25] measures convergence to the NE when the model degenerates to a static game, ours can assess the performance of a learning algorithm in low-information settings where agents improve their performance primarily by “learning through play” [35].

Remark 2: About the most relevant results, Srivastava et al. [33] proposed a distributed Bregman distance algorithm for saddle-point problems, while we consider a game between two subnetworks, and Talebi et al. [36] developed a team-based dual averaging algorithm for a game (not necessarily zero-sum) between two teams over a network and proved the convergence of cost values by introducing cross-monotonicity. In comparison, we provide the regret bound of the algorithm and prove the convergence of the action profiles to an NE.

B. Distributed Mirror Descent

For \( l = 1, 2 \), let \( \psi_l \) be a continuously differentiable \( \sigma_l - \)strongly convex function on \( X_l \), which means that
\[
\psi_l(y) \geq \psi_l(x) + \langle \nabla \psi_l(x), y - x \rangle + \frac{\sigma_l}{2} \| y - x \|^2 \quad \forall x, y \in X_l.
\]
Recall from [37] that the Bregman divergence associated with \( \psi_l \) is defined as
\[
D_{\psi_l}(x, y) \triangleq \psi_l(x) - \psi_l(y) - \langle \nabla \psi_l(y), x - y \rangle. \quad (4)
\]
Then, we design our algorithm based on the mirror descent algorithm [38], [39], which was further extended to solve the distributed optimization problem in [40] and [41]. At time \( t + 1 \), each agent \( i \in V_l \) \( (l \in \{1, 2\}) \) receives the estimates \( x_{1,j}(t) \) from \( j \in N^l_1(t) \) and the estimates \( x_{3-1,j}(t) \) from \( j \in N^l_{12}(t) \). Let \( \nu_{l,i}(t) \) and \( \nu_{3-1,i}(t) \) be the weighted average of the estimations from \( \Sigma_l \) and \( \Sigma_{3-1} \), respectively. Then, for \( l \in \{1, 2\} \), each agent \( i \in V_l \) receives the subgradients of the local costs \( f_{l,i}(x) \) at \( \nu_{l,i}(t) \) by
\[
g_{l,i}(t) \in \partial f_{l,i}(\nu_{l,i}(t), \nu_{2,2}(t)) \quad \text{and updates its estimate by the mirror descent scheme (8).}
\]
We summarize the procedures in Algorithm 1.

In order to describe the information flow in Algorithm 1, we take two agents (denoted by \( v_1^i \) and \( v_2^j \) in Fig. 1) as an example and give an illustration diagram in Fig. 2. When Algorithm 1 iterates from step \( t \) to step \( t + 1 \). The information flow on other agents is similar.

By Assumption 1(iii) and the definition of the dual norm, we have
\[
\| g_{1,i}(t) \|_* \leq L_{1,1}, \quad \| g_{2,i}(t) \|_* \leq L_{2,1} \quad (9)
\]
which was frequently used in the convergence analysis of distributed algorithms [4], [34]. For the subsequent analysis, we make the following assumption.

Assumption 3: For \( l = 1, 2 \), the Bregman divergence \( D_{\psi_l}(x, y) \) is convex in \( y \) and satisfies
\[
x_k \rightarrow x \Rightarrow D_{\psi_l}(x_k, x) \rightarrow 0. \quad (10)
\]
In fact, \( D_{\psi_l}(\cdot, y) \) is always strictly convex and the convexity assumption about \( y \) has been widely used in distributed optimization (see, e.g., [37] and [42]). Commonly used Bregman

Algorithm 1: Distributed Mirror Descent Algorithm.
\[
\text{Initialize: For } l \in \{1, 2\} \text{ let } \nu_{l,i}(0) \in X_l \text{ for each } i \in V_l.
\]
\[
\text{Iterate until } t \geq T:
\]

Communication and distributed averaging: For \( l \in \{1, 2\}, i \in V_l \),
\[
u_{l,i}(t) := \sum_{j \in N^l_1(t)} w_{l,i,j}(t) x_{l,j}(t) \quad (6)
\]
\[
u_{3-1,i}(t) := \sum_{j \in N^l_{12}(t)} w_{3-1,i,j}(t) x_{3-1,j}(t). \quad (7)
\]

Update of \( x_{l,i}(t) \): For \( l \in \{1, 2\} \),
\[
x_{l,i}(t + 1) = \arg \min_{x_i \in X_l} \left\{ \langle g_{l,i}(t), x_i - \nu_{l,i}(t) \rangle + \frac{1}{\alpha(t)} D_{\psi_l}(x_i, \nu_{l,i}(t)) \right\} \quad (8)
\]
where \( \{ \alpha(t) \}_{t=0}^{\infty} \) is a positive nonincreasing sequence.

Fig. 2. Diagram of information flow when Algorithm 1 iterates from \( t \) to \( t + 1 \).

III. REGRET ANALYSIS

We begin this section by establishing preliminary lemmas and then provide an upper bound on each agent’s regret.

A. Preliminary Analysis

First, we state a basic result about the Bregman divergence.

Lemma 2: Let \( \psi \) be a continuously differentiable \( \sigma \)-strongly convex function on \( X \). Then, the Bregman divergence defined
satisfies, for all \( t \geq 1 \)
\[
D_\psi(y, z) = (\nabla \psi(z) - \nabla \psi(x), y - z)
\]

(11)

\[
D_\psi(x, y) \geq \frac{\sigma}{2} \|x - y\|^2.
\]

(12)

Lemma 2 is widely used in the analysis of mirror descent algorithms [28], [37] and can be easily obtained from (4). We next state a result about distributed optimization [34].

Lemma 3: Let \( \Phi_l(t, s) = W_l(t) \left( W_l(t - 1) \cdots W_l(s) \right) (l = 1, 2) \) be the transition matrices for \( \Sigma_l \) (l = 1, 2). Suppose that Assumption 2 holds. Then, with \( B_l \) and \( \eta \) given by Assumption 1, for all \( t, s \) with \( t \geq s \geq 0 \), we have
\[
\left| \Phi_l(t, s)_{ij} - \frac{1}{n_l} \right| \leq \Gamma_l t^{1-s}, \quad l = 1, 2
\]

where \( \Phi_l(t, s)_{ij} \) denotes its element in the \( i \)-th row and \( j \)-th column, \( \Gamma_l = (1 - \eta/4n_l^2)^{-2} \), and \( \theta_t = (1 - \eta/4n_l^2)^{1/2} \).

Note that \( \Gamma_l \) and \( \theta_t \) increase as \( n_l \) grows, indicating that as the network size increases, the rate at which \( \Phi_l(t, s)_{ij} \) converges to \( 1/n_l \) slows down. Moreover, for smaller intercommunication interval bound \( B_l \), the convergence rate is faster.

Based on (8), we provide the following error bound.

Lemma 4: Let Assumption 1 hold. Suppose that \( v_{l,i}(t) \) and \( x_{l,i}(t + 1) \) are generated by Algorithm 1. Then, for each \( l \in \{1, 2\} \) and \( i \in \mathcal{V}_l \), we have
\[
\|v_{l,i}(t) - x_{l,i}(t + 1)\| \leq \frac{\alpha(t)}{\sigma_l} \|g_{l,i}(t)\|_\infty \leq L_{l,2} \frac{\alpha(t)}{\sigma_l}.
\]

(13)

Lemma 4 is a consequence of the optimality condition. See Appendix A for its proof. Furthermore, let \( \bar{x}_l(t) = \frac{1}{n_l} \sum_{i=1}^{n_l} x_{l,i}(t) \) be the average state of \( \Sigma_l \) at time \( t \), and then, the following error bounds are established.

Lemma 5: Let Assumptions 1 and 2 hold. Suppose that \( x_{1,i}(t), v_{1,i}(t), \) and \( u_{1,i}(t) \) are generated by Algorithm 1. Then, for each \( l \in \{1, 2\} \) and \( i \in \mathcal{V}_l \), for any \( t \),
\[
\|x_{l,i}(t) - \bar{x}_l(t)\| \leq H_l(t)
\]

(14)

\[
\|\bar{x}_l(t) - v_{l,i}(t)\| \leq H_l(t)
\]

(15)

\[
\|\bar{x}_l(t) - u_{l,i}(t)\| \leq H_l(t)
\]

(16)

with
\[
H_l(t) = n_l \Gamma_l t^{1-s} \Lambda_l + \sum_{i=1}^{L_l} \alpha(s - 1) + \frac{1}{\sigma_l} (n_l L_{l,1} \Gamma_l \sum_{i=1}^{L_l} \theta_t^{1-s} \Lambda_l)
\]

(17)

\[
\Lambda_l \triangleq \max_{i \in \mathcal{V}_l} \|x_{l,i}(0)\|.
\]

In Lemma 5, (14) is a fundamental result in the analysis of distributed mirror descent algorithms [37], [42], and (15) and (16) can be easily established based on (14). For completeness, we give its proof in Appendix A. Finally, we establish the following result similar to that of [28] and [42] for distributed optimization.

Lemma 6: Let Assumptions 1–3 hold. Suppose that \( x_{1,i}(t) \) and \( g_{l,i}(t) \) are generated by Algorithm 1. Then, for \( l \in \{1, 2\} \), \( i \in \mathcal{V}_l \), for any \( \bar{x}_l \in \mathcal{X}_l \), we have
\[
\frac{1}{n_l} \sum_{i=1}^{n_l} \left( g_{l,i}(t), x_{l,i}(t + 1) - \bar{x}_l \right) \leq \frac{Y_1}{\alpha(T)}
\]

(18)

where \( Y_1 \triangleq \max \left\{ D_\psi(\bar{x}_l, x_l) \mid \forall \bar{x}_l, x_l \in \mathcal{X}_l \right\} \).

By the optimality condition and (11), we obtain (17). The detailed proof is provided in Appendix A.

B. Regret Bound

Equipped with the above lemmas, we are ready to provide a regret bound of Algorithm 1.

Theorem 1: Under Assumptions 1–3, the regret defined by (3) can be bounded as
\[
\begin{align*}
R_1(t) &\leq \frac{4}{\sigma_1} \left( \left( \sum_{i=1}^{L_1} \alpha(t - 1) \Gamma_1 + \sum_{i=1}^{L_1} \alpha(t - 1) \right) \sum_{t=1}^{T} \alpha(t - 1) + \frac{\Theta_l^2}{\alpha(T)} \right) \\
&\quad + \frac{12}{\sigma_2} \left( \left( \sum_{i=1}^{L_1} \alpha(t - 1) \Gamma_2 + \sum_{i=1}^{L_1} \alpha(t - 1) \right) \sum_{t=1}^{T} \alpha(t - 1) \\
&\quad + 4 \sum_{t=1}^{T} \sum_{i=1}^{L_1} \alpha(t) \Gamma_1 \sum_{i=1}^{L_1} \alpha(t) \Gamma_2 \right) \frac{\alpha(t)}{\sigma_1} L_{1,1}^2.
\end{align*}
\]

Lemma 6: For the case of a general step-size \( \alpha(t) \), \( \sum_{t=0}^{\infty} \alpha(t) = \infty \), and \( \sum_{t=0}^{\infty} \alpha^2(t) < \infty \).

Proof: See Appendix A.

Theorem 1 provides an upper bound on the individual regret for each agent in subnetwork \( \Sigma_1 \) in the case of a general step-size sequence \( \{\alpha(t)\}_t \). Note that the impact of the communication network is incorporated in the constants \( \Gamma_l \) and \( \theta_t \). Moreover, a regret bound for each agent in \( \Sigma_2 \) can be similarly established. Next, we characterize the regret bound under two specific step-size sequences.

Corollary 1: Under the same conditions stated in Theorem 1, Algorithm 1 with a constant step size \( \alpha(t) = 1/\sqrt{T} \) and a diminishing step-size sequence \( \alpha(t) = t^{-\frac{1}{2} + \epsilon} \) (\( \epsilon \in (0, \frac{1}{2}) \)) yields \( R_1(t) \leq O(\sqrt{T}) \) and \( R_1(t) \leq O(T^{1-2\epsilon}) \), respectively.

This corollary follows by \( \sum_{t=0}^{\infty} \alpha(t) = 1 + \sum_{t=1}^{\infty} t^{-\frac{1}{2} + \epsilon} \leq T^{\frac{1}{2} - \epsilon} + 1/\alpha(T) = T^{\frac{1}{2} + \epsilon} \). Furthermore, Corollary 1 shows that both constant and diminishing step sizes can make Algorithm 1 no-regret, and the regret rate has the same order as that of distributed optimization [28].

Remark 3: Regret analysis in a static game model has already been considered in some early works [44], [45]. Compared to the distributed online optimization problem [28], the difficulty in our regret analysis lies in dealing with the interactions among agents from different subnetworks. Moreover, compared to the two-player game model [44], the difficulty is to resolve the absence of complete information since each agent can only exchange information with its neighbors in our network case.

IV. CONVERGENCE ANALYSIS

In this section, we study the convergence of Algorithm 1 under both diminishing and constant step sizes.

A. Diminishing Step Size

In this subsection, we adopt a diminishing step-size sequence in Algorithm 1 and make the following assumption.

Assumption 4: \( \sum_{t=0}^{\infty} \alpha(t) = \infty \) and \( \sum_{t=0}^{\infty} \alpha^2(t) < \infty \).

To facilitate the convergence analysis, we first state a well-known result about nonnegative sequences [46].

Lemma 7: Let \( \{a_1\}, \{b_1\}, \) and \( \{c_1\} \) be nonnegative sequences with \( \sum_{t=0}^{\infty} b_t < \infty \). If \( a_{t+1} \leq a_t + b_t - c_t \) holds for any \( t \), then \( a_t \) converges to a finite number and \( \sum_{t=0}^{\infty} c_t < \infty \).
Next, we provide a convergence result on the actual sequence of actions; namely, each paired sequence \( \{(x_{1,i}(t), x_{2,j}(t))\} \) converges to the NE.

**Theorem 2:** Suppose that Assumptions 1–4 hold and cost function \( U \) is strictly convex–concave. Then, Algorithm 1 generates a sequence that converges to the unique NE \( x^* = (x_1^*, x_2^*) \), i.e.,

\[
\lim_{t \to \infty} x_{1,i}(t) = x_1^*, \quad \lim_{t \to \infty} x_{2,j}(t) = x_2^* \quad \forall i \in \mathcal{V}_1, j \in \mathcal{V}_2.
\]

**(Proof):** See Appendix B. ■

Compared with [36], which established the convergence of the running average of action iterates under cross-monotonicity and strict monotonicity, Theorem 2 shows that the agents’ estimates of the subnetwork state can achieve consensus and converge to the NE. Moreover, since the diminishing step-size sequence in Corollary 1 satisfies Assumption 4, the following corollary can be directly obtained from Corollary 1 and Theorem 2.

**Corollary 2:** Under Assumptions 1–3, Algorithm 1 with a diminishing step-size sequence \( \alpha(t) = t^{-(\frac{1}{2}+\epsilon)} \) (\( \epsilon \in (0, \frac{1}{2}) \)) is a no-regret distributed updating policy, converging to the NE.

### B. Constant Step Size

Here, we consider a constant step-size version of Algorithm 1 (i.e., \( \alpha(t) \equiv \alpha \)). By Corollary 1, we obtain a tighter regret bound under the constant step size. However, since Assumption 4 does not hold in this situation, Theorem 2 is no longer applicable. Inspired by Nedić and Ozdaglar [31], we consider the following running average actions for each node \( i \) in \( \Sigma_i \):

\[
\hat{x}_{i}(t) = \frac{1}{t} \sum_{s=0}^{t-1} x_{i,s}(s), \quad \text{for } t \geq 1, l = 1, 2.
\]

Next, we prove that these averages provide an approximation of NE. Specifically, let \( (x_1^*, x_2^*) \) be an NE. Then, the following theorem establishes an asymptotic error bound between the cost of the averages and the cost of the NE.

**Theorem 3:** Suppose Assumptions 1–3 hold. Consider Algorithm 1 with \( \alpha(t) \equiv \alpha \). Then, for all \( t \geq 1 \), the averages \( \{\hat{x}_{1,i}(t), \hat{x}_{2,i}(t)\}, i \in \mathcal{V}_1, j \in \mathcal{V}_2 \), satisfy

\[
\begin{align*}
|U(\hat{x}_{1,i}(t), \hat{x}_{2,j}(t)) - U(x_1^*, x_2^*)| &\leq \sum_{l=1}^{2} \left( \frac{\gamma_l^2}{\sigma_l} + \frac{L^2}{\sigma_l} \alpha \right) + 4L(K_1 + K_2)\alpha \\
&\quad + 4L \sum_{s=0}^{t-1} \sum_{l=1}^{2} n_l \Gamma_l \theta_l^{t-1} A_l
\end{align*}
\]

where \( K_l \equiv \frac{1}{\alpha_l} \left( \frac{\gamma_l^2}{\sigma_l} + 2L \right) \) for \( l = 1, 2 \).

**(Proof):** See Appendix B. ■

Theorem 3 shows that the cost value of the averaged iterates \( U(\hat{x}_{1,i}(t), \hat{x}_{2,j}(t)) \) converges to \( U(x_1^*, x_2^*) \) within error level \( \left( \frac{2L^2}{\sigma_l^2} + 4L(K_1 + K_2)\alpha \right) \) with rate \( O(t^{-\frac{1}{2}+\epsilon}) \), which is comparable to the rate established in [31] for the centralized saddle point problems.

### V. SUBNETWORK ZERO-SUM FINITE-STRATEGY GAMES

In this section, we consider a subnetwork zero-sum finite-strategy game as a concrete application of Algorithm 1. We give a simplified algorithm in Section V-A with its convergence results in Section V-B.

#### A. Simplified Algorithm

For each subnetwork \( \Sigma_i \ (l \in \{1, 2\}) \), suppose that its (pure) action set is \( \mathcal{A}_l = \{a_1^{(l)}, \ldots, a_{M_l}^{(l)}\} \), with \( M_l \) being an integer. To obtain a continuous cost function and apply Algorithm 1, we consider its mixed strategy, which has also been studied in [23] and [44]. The mixed strategy, denoted by \( x_i \), belongs to the corresponding mixed-strategy set \( \mathcal{X}_l = \{x_i = (x_1^{(l)}, \ldots, x_{M_l}^{(l)}) \mid \sum_{p=1}^{M_l} x_i^{(p)} = 1, 0 \leq x_i^{(p)} \leq 1\} \).

Let \( f_{i,l}(a_1^{(p)}, a_2^{(q)}) \) be the cost value of agent \( i \) at the pure action profile \( (a_1^{(p)}, a_2^{(q)}) \), and then, in a slight abuse of notation, the expected cost value at the mixed-strategy profile \( (x_1, x_2) \) is a multilinear function defined as follows:

\[
U(x_1, x_2) = \frac{1}{n_1} \sum_{i=1}^{n_1} f_{i,l}(x_1, x_2) \triangleq \frac{1}{n_1} \sum_{i=1}^{n_1} \sum_{p=1}^{M_1} \sum_{q=1}^{M_2} x_1^{(p)} x_2^{(q)} f_{i,l}(a_1^{(p)}, a_2^{(q)}).
\]

Therefore, Assumption 1 holds, and the subgradients (5) are equal to the mixed cost vectors \( (f_{i,l}(a_1^{(p)}, a_2^{(q)}))_{a_1 \in A_1} \) and \( (f_{2,i}(a_1^{(p)}, x_2))_{a_2 \in A_2} \). Similarly, a strategy profile \( x^* = (x_1^*, x_2^*) \) is a mixed-strategy NE of a subnetwork zero-sum finite-strategy game if (2) holds.

Because \( \mathcal{X}_l \) is a simplex, we consider the negative entropy regularizer \( \psi_l(x_i) = \sum_{p=1}^{M_l} x_i^{(p)} \log \frac{x_i^{(p)}}{y_l^{(p)}} \). Then, \( \psi_l \) is 1-strongly convex with respect to \( l_1 \) norm (when there is no confusion, we use the norm \( || \cdot || \) in Sections II and III as \( l_1 \) norm), and the Bregman divergence is given by

\[
D_{\psi_l}(x_i, y_l) = \sum_{p=1}^{M_l} x_i^{(p)} \log \frac{x_i^{(p)}}{y_l^{(p)}}.
\]

Therefore, Assumption 3 holds. Moreover, we construct the following Lagrangian function:

\[
L_i(x_i, \lambda) = \sum_{p=1}^{M_l} \alpha(t) \eta_{i,l}^{(p)}(t) (x_i^{(p)} - \psi_{i,l}^{(p)}(t)) + \sum_{p=1}^{M_l} x_i^{(p)} (\log x_i^{(p)} - \log \psi_{i,l}^{(p)}(t)) + \lambda \left( \sum_{p=1}^{M_l} x_i^{(p)} - 1 \right)
\]

where \( \lambda \) is a multiplier. The partial derivatives of \( L_i \) are

\[
\begin{align*}
\frac{\partial L_i}{\partial x_i^{(p)}} &\equiv \alpha(t) \eta_{i,l}^{(p)}(t) + \log x_i^{(p)} + 1 - \log \psi_{i,l}^{(p)}(t) + \lambda \\
\frac{\partial L_i}{\partial \lambda} &\equiv \sum_{p=1}^{M_l} x_i^{(p)} - 1.
\end{align*}
\]
By setting the derivatives to 0, (8) is simplified as
\[
x \frac{(p)}{t+1} = \frac{x \frac{(p)}{t} \exp(-\alpha(t) g \frac{(p)}{t}(t))}{\sum_{p=1}^{M_N} x \frac{(p)}{t} \exp(-\alpha(t) g \frac{(p)}{t}(t))},
\]
\[p = 1, \ldots, M_N. \tag{22}\]

The complete learning algorithm is summarized in Algorithm 2, which can be viewed as a distributed version of the classic multiplicative-weight (MW) algorithm [44].

**Algorithm 2:** Distributed MW Algorithm.

*Initialize:* For \(l \in \{1, 2\}, \) let \(x \frac{(l)}{0}) = \frac{1}{M_l} (1, \ldots, 1) \in X_l.\)

*Iterate until* \(t \geq T:*

**Communication and distributed averaging:** For \(l \in \{1, 2\}, \) in \(V_l\)

compute the estimates \(x \frac{(l)}{t},(t)\) based on (6)

compute the estimates \(u \frac{(l)}{3-t,(l)}(t)\) based on (7)

**Update of** \(x \frac{(l)}{t},(t):\) For \(l \in \{1, 2\}, \) in \(V_l\)

compute the gradients \(y \frac{(l)}{1},(t) \in R^{M_l};\)

\(g \frac{(l)}{1},(t) = f \frac{(l)}{1},(t);(t), p = 1, \ldots, M_1 \tag{23}\)

\(g \frac{(l)}{2},(t) = f \frac{(l)}{2},(t);(t), a \frac{(l)}{2}, p = 1, \ldots, M_2 \tag{24}\)

update the estimates \(x \frac{(l)}{t+1},(t+1)\) by (22)

**Remark 4:** The update rule (22) displays the advantage of the distributed mirror descent algorithm compared with the distributed projected subgradient descent algorithm [4]. We may avoid calculating the projection onto a simplex by choosing a negative entropy regularizer in the mirror descent algorithm.

**B. Convergence Results**

Note that if \(U\) is not strictly convex–concave, then Theorem 2 cannot be directly applied. Bertsekas et al. [47] showed that even if the cost function \(U\) is bilinear and admits an interior NE, the iterates of the mirror descent are not convergent. Therefore, we need a stronger assumption to establish a final-convergence and distributed model. Through an analysis similar to that in [4], we obtain the following convergence result.

**Theorem 4:** Under Assumptions 2 and 4, if the set of NE contains an interior point, then Algorithm 2 generates a sequence that converges to a mixed-strategy NE for the subnetwork zero-sum finite-strategy game.

**Proof:** See Appendix C.

Moreover, motivated by Theorem 3, we can obtain the following result for the generalized running average actions \(\hat{x}_1, (t)\) defined as

\[\hat{x}_1, (t) = \frac{1}{\sum_{s=0}^{t-1} \alpha(s)} \sum_{s=0}^{t-1} \alpha(s) x_1, (s), \text{ for } t \geq 1, l = 1, 2.\]

**Theorem 5:** Under Assumptions 2 and 4, Algorithm 2 generates an ergodic average sequence \((\hat{x}_1, (t), \hat{x}_2, (t))\) that converges to the mixed-strategy NE set.

**Proof:** See Appendix C.

A similar ergodic convergence of dual averaging algorithm for finite two-person zero-sum games has been obtained in [23], and here, we extend the result to subnetwork zero-sum games by using a distributed mirror descent algorithm.

**VI. SIMULATIONS**

In this section, we provide two potential applications to illustrate the structure of subnetwork zero-sum games and verify no-regret property and convergence of the proposed algorithms.

**A. Network Interdiction**

Consider a network interdiction problem modeled by a two-player zero-sum game in [48]. We generalize it to a zero-sum game between two groups, called interdictors (\(\mathbf{I}\)) and evaders (\(\mathbf{E}\)) (corresponding to the subnetworks \(\Sigma_1\) and \(\Sigma_2\) in Fig. 1). Both groups consist of \(N = 10\) agents. Group \(\mathbf{E}\) attempts to traverse from node \(s\) to node \(t\) through a network \(G_0\), without being detected by group \(\mathbf{I}\). Group \(\mathbf{I}\) selects an arc in the network and sets up an inspection site there. Let \(\mathbf{I}_1\) and \(\mathbf{I}_2\) represent agent \(i\) in group \(\mathbf{I}\) and agent \(j\) in group \(\mathbf{E}\), respectively. For agent \(\mathbf{I}_1\), if \(\mathbf{E}\) passes arc \(k\), then \(\mathbf{I}_1\) detects \(\mathbf{E}\) with probability \(p_k\). Similarly, for agent \(\mathbf{E}_j\), if \(\mathbf{I}\) selects arc \(k\) to detect, \(\mathbf{E}_j\) is detected with probability \(q_k\). Let \(\mathbf{A}_1\) denote the set of all \(s - t\) paths (i.e., the pure strategy set of \(\mathbf{E}\)) and \(\mathbf{A}_2\) denote the arc set of network \(G_0\). Let \(x_p\) be the probability that \(\mathbf{E}\) selects path \(p\), and \(y_k\) be the probability that \(\mathbf{I}\) selects arc \(k\). The individual interdiction probabilities of groups \(\mathbf{I}\) and \(\mathbf{E}\) are defined as

\[f_1, (x, y) = \sum_{p \in P} \sum_{k \in A} x_p p_k d_p y_k = x^T A_1 y\]

\[f_2, (x, y) = \sum_{p \in P} \sum_{k \in A} x_p q_k d_p y_k = x^T B_1 y\]

respectively, where \(A_1 = [A_1, p_k] \triangleq [p_k]^T d_p\) is the payoff matrix of agent \(\mathbf{I}_1\), \(B_1 = [B_1, p_k] \triangleq [q_k]^T d_p\) is the cost matrix of agent \(\mathbf{E}_1\), and \(d_p = 1\) if path \(p\) includes arc \(k\) otherwise \(d_p = 0\). Assume that the average probability of group \(\mathbf{I}\) interdicting group \(\mathbf{E}\), \(\frac{1}{N} \sum_{i=1}^{N} f_1, (x, y)\), is equal to the average probability of \(\mathbf{E}\) being interdicted by \(\mathbf{I}\), \(\frac{1}{N} \sum_{i=1}^{N} f_2, (x, y)\). Let \(f(x, y)\) be the common average interdiction probability. The goal of \(\mathbf{I}\) is to maximize \(f(x, y)\), while the goal of \(\mathbf{E}\) is to minimize \(f(x, y)\). For example, in the network of Fig. 3, there are three \(s - t\) paths \(s \rightarrow i_1 \rightarrow t\), \(s \rightarrow i_2 \rightarrow t\), and \(s \rightarrow t\) and five arcs \((s, i_1), (i_1, t), (s, i_2), (i_2, t), \) and \((s, t)\). The mixed-strategy
sets of \( \mathbf{E} \) and \( \mathbf{I} \) are \( \mathcal{X} \triangleq \{ x \in \mathbb{R}^3 | \sum_{i=1}^{3} x_i = 1, x \geq 0 \} \) and \( \mathcal{Y} \triangleq \{ y \in \mathbb{R}^5 | \sum_{i=1}^{5} y_i = 1, y \geq 0 \} \).

Agents in group \( \mathbf{E} \) cooperate to choose a mixed strategy (i.e., a probability distribution) on \( \mathcal{A}_i \), while agents in group \( \mathbf{I} \) cooperate to choose a mixed strategy on \( \mathcal{A}_2 \). Given a pool of connected graphs, two graphs \( G_1(t) \) and \( G_2(t) \) are randomly selected from the pool at time \( t \). Suppose that each agent in groups \( \mathbf{E} \) and \( \mathbf{I} \) exchanges information with their neighbors through graphs \( G_1(t) \) and \( G_2(t) \), respectively. Furthermore, assume that the interaction network is one-to-one, i.e., agent \( \mathbf{E}_i \) only receives information of group \( \mathbf{I} \) from agent \( \mathbf{I}_i \). Set \( |\mathcal{A}_1| = 30 \) and \( |\mathcal{A}_2| = 60 \), and generate ten random matrices as the payoff matrices of agents in group \( \mathbf{I} \). Then, we check the regret bound and convergence of Algorithm 2 under different step sizes. In Fig. 4(a), we compare the average regret of agent \( \mathbf{E}_2 \) denoted by \( \mathcal{R}_{[2]}(T)/T \) with diminishing step sizes \( \alpha(t) = t^{-\frac{1}{2}}, t^{-\frac{2}{2}}, t^{-\frac{3}{2}} \). Algorithm 2 produces a smaller average regret when the decay rate of \( \alpha(t) \) is smaller, which verifies Corollary 1. Since the trajectories decrease very slowly in the later period, we choose a simulation horizon \( (T = 2000) \) sufficient to compare the downward trend of the average regret. Fig. 4(b) provides a plot of the average gap function \( \frac{1}{N} \sum_{i=1}^{N} (\hat{x}_i(t), \hat{y}_i(t)) \), defined as \( \frac{1}{N} \sum_{i=1}^{N} \max_{x,y} \sum_{j=1}^{N} (x_j(t))^T A_j y - \min_{x,y} \sum_{j=1}^{N} x^T A_j y \), which illustrates the ergodic convergence of Algorithm 2. Moreover, Fig. 4(b) also shows faster convergence under a slower step-size sequence.

**B. Power Allocation With Adversaries**

In this part, we use a strictly convex–concave game to verify the convergence to the unique NE. Consider a power allocation problem with adversaries over \( N = 6 \) Gaussian communication channels [30]. The communication rate of each channel depends on its signal power and Gaussian noise power. \( N \) agents connected by network \( \mathcal{S}_1 \) wish to suitably allocate the total signal power among the channels so as to maximize the total communication rate, while \( N \) adversaries connected by network \( \mathcal{S}_2 \) attempt to minimize the total communication rate by selecting noise powers. To be specific, denote by \( \{ch_1, ch_2, \ldots, ch_6\} \) the channels, \( \mathcal{S}_1 \) decides to allocate signal power \( x_1 \) to \( \{ch_1, ch_2\} \), signal power \( x_2 \) to \( \{ch_2, ch_3\} \), and signal power \( x_3 \) to \( \{ch_3, ch_6\} \). Similarly, \( \mathcal{S}_2 \) decides to allocate noise power \( y_1 \) to \( \{ch_1, ch_2\} \), noise power \( y_2 \) to \( \{ch_3, ch_4\} \), and noise power \( y_3 \) to \( \{ch_5, ch_6\} \). Both signal and noise powers satisfy a budget constraint, \( 2x_1 + 2x_2 + 2x_3 = 2 \) and \( 2y_1 + 2y_2 + 2y_3 = 2 \). Let \( x = (x_1, x_2, x_3) \) and \( y = (y_1, y_2, y_3) \), and take the objective function of agent \( i \) in \( \mathcal{S}_1 \) as the communication rate of channel \( ch_i \), which is defined as \( f_{1,i}(x, y) = \log \left( 1 + \frac{8x_i y_i}{\sigma(i) + y_i} \right) \), where \( \sigma = \{1, 2, 3, 1, 2, 3\} \) and \( b = \{1, 1, 2, 2, 3, 3\} \) are index vectors, and \( \sigma(i) = i \) is the receiver noise of channel \( ch_i \). For \( i \in \{1, 2, 3, 4, 5, 6\} \), the individual objective function of \( \mathcal{S}_2 \) is \( f_{2,i}(x, y) = -f_{1,i}(x, y) \). The goal of \( \mathcal{S}_1 \) is to select signal power \( x \) to maximize \( f(x, y) = \sum_{i=1}^{N} \log \left( 1 + \frac{8x_i y_i}{\sigma(i) + y_i} \right) \), which is a strictly concave–convex function. The goal of \( \mathcal{S}_2 \) is to select noise power \( y \) to minimize \( f(x, y) \).

For simplicity, let \( \mathcal{S}_1 \) and \( \mathcal{S}_2 \) be fixed networks with the same topology and be described by a randomly generated graph \( \mathcal{G}(t) \). Thus, the interaction network \( \mathcal{S}_{12} \) is one-to-one. Denote by \( x^t = (x_1^t, x_2^t, x_3^t) \) (\( y^t = (y_1^t, y_2^t, y_3^t) \)) the estimated signal (noise) power of channel \( ch_i \) in \( \mathcal{S}_1 \) (\( \mathcal{S}_2 \)). Take \( \alpha(t) = 1/\sqrt{T} \) and the entropy regularizer \( \psi_L(x) = \sum_{p=1}^{3} x_p \log x_p \) in Algorithm 1. The trajectories of the average action errors of \( \mathcal{S}_1 \) and \( \mathcal{S}_2 \) are plotted in Fig. 5. Here, we use a centralized mirror descent method to compute the NE \( (x^*, y^*) \). Fig. 5 shows that \( (x^t, y^t) \) converges to \( (x^*, y^*) \) and, thus, verifies Theorem 2.

Let us consider the effect of the algebraic connectivity of the communication graph \( \mathcal{G}(t) \) on the regret bound. Denote the algebraic connectivity of \( \mathcal{S}_1 \) as \( \lambda_2 \), which is defined as the second smallest eigenvalue of its Laplacian matrix. With \( \mathcal{S}_2 \) and \( \mathcal{S}_{12} \) unchanged, the trajectories of the average regret of channel 1 for \( \lambda_2 \in \{0.4, 0.8, 1.2\} \) are plotted in Fig. 6. Fig. 6 shows that \( \lambda_2 \) does not change the rate of average regret, while Algorithm 1 produces a smaller regret if \( \mathcal{S}_1 \) has a larger algebraic connectivity.
VII. CONCLUSION

In this article, we proposed a distributed mirror descent algorithm for NE seeking in subnetwork zero-sum games. First, we provided a regret analysis for the proposed algorithm under both diminishing and constant step sizes. We also proved its convergence to the NE under diminishing step sizes. Our analysis demonstrates that the proposed algorithm satisfies a no-regret property while converging to the NE. Moreover, we established an asymptotic error bound on the cost value of averaged iterates in the constant step-size case. Finally, we proved final-iteration convergence and ergodic convergence results, respectively, under different assumptions on the cost functions in subnetwork zero-sum finite-strategy games. In the future, we may leverage subnetwork zero-sum games as an avenue for modeling practical problems and extend the results to games with $K$ subnetworks or time-varying cost functions. It is also of interest to analyze the dynamic regret in subnetwork zero-sum games.

APPENDIX A

PROOFS OF SECTION III

Proof of Lemma 4: Applying the optimality condition for (8), we get that for each $l = 1, 2$, and any $x_l \in X_l$,

$$
\langle g_{l,i}(t) + \frac{1}{\alpha(t)}\nabla x_l D_{\psi_i}(x_l, v_{l,i}(t))|_{x_l=x_{1,i}(t+1)}, x_l
- x_{l,i}(t+1) \rangle \geq 0.
$$

Recalling the definition in (4), we derive

$$
(\nabla \psi_l(x_{1,i}(t+1)) - \nabla \psi_l(v_{1,i}(t))) + \alpha(t) g_{l,i}(t), x_l
- x_{l,i}(t+1) \geq 0.
$$

By setting $x_l = v_{l,i}(t)$ in (25), we obtain

$$
\langle \nabla \psi_l(x_{1,i}(t+1)) - \nabla \psi_l(v_{1,i}(t)) + \alpha(t) g_{l,i}(t)
- v_{l,i}(t+1) \rangle \leq 0.
$$

Therefore, from the strong convexity of $\psi_l$

$$
\alpha(t)||g_{l,i}(t)||_s ||v_{l,i}(t) - x_{l,i}(t+1)||
\geq \langle \alpha(t) g_{l,i}(t), v_{l,i}(t) - x_{l,i}(t+1) \rangle
\geq \langle \nabla \psi_l(x_{1,i}(t+1)) - \nabla \psi_l(v_{1,i}(t)), x_{l,i}(t+1) - v_{l,i}(t) \rangle
\geq \sigma_l ||v_{l,i}(t) - x_{l,i}(t+1)||^2
$$

which together with (9) yields (13).

Proof of Lemma 5: By (6),

$$
x_{l,i}(t) = v_{l,i}(t-1) - (v_{l,i}(t-1) - x_{l,i}(t))
= \sum_{j \in N_{l,i}(t-1)} w_{l,ij}(t-1)x_{l,j}(t-1) + p_{l,i}(t-1)
= \sum_{j=1}^{n_l} [\Phi_l(t-1, 0)]_{ij} x_{l,j}(0)
+ \sum_{s=1}^{t-1} \sum_{j=1}^{n_l} [\Phi_l(t-1, s)]_{ij} p_{li}(s-1) + p_{l,i}(t-1)
$$

with $p_{l,i}(t-1) \triangleq x_{l,i}(t) - v_{l,i}(t-1)$. As a result, the double stochasticity of $W(t)$ implies

$$
\bar{x}_l(t) = \frac{1}{n_l} \sum_{j=1}^{n_l} x_{l,j}(0) + \frac{1}{n_l} \sum_{s=1}^{t} \sum_{j=1}^{n_l} p_{l,j}(s-1).
$$

Therefore, Lemma 3 and (13) together yield

$$
\|x_{l,i}(t) - \bar{x}_l(t)\|
\leq \sum_{j=1}^{n_l} [\Phi_l(t-1, 0)]_{ij} - \frac{1}{n_l} \|x_{l,j}(0)\|
+ \sum_{s=1}^{t-1} \sum_{j=1}^{n_l} [\Phi_l(t-1, s)]_{ij} - \frac{1}{n_l} \|p_{l,j}(s-1)\|
+ \frac{1}{n_l} \sum_{j=1}^{n_l} p_{l,j}(t-1) - p_{l,i}(t-1)
\leq n_l \Gamma_l \theta_l^{t-1} \Lambda_l
+ \frac{1}{\sigma_l} \left( n_l \Gamma_l \sum_{s=1}^{t-1} \theta_l^{t-1-s} \alpha(s-1) + 2L_{l,1} \alpha(t-1) \right).
$$

Thus, (15) holds. Similarly, by $\sum_{j=1}^{n_l} w_{l,ij}(t) = 1$ and (14), we obtain (16).

Proof of Lemma 6: By setting $x_l = \bar{x}_l$ in (25), we have

$$
\langle \alpha(t) g_{l,i}(t), x_{l,i}(t+1) - \bar{x}_l \rangle
\leq \langle \nabla \psi_l(v_{l,i}(t)) - \nabla \psi_l(x_{l,i}(t+1), x_{l,i}(t+1) - \bar{x}_l \rangle
= D_{\psi_l}(\bar{x}_l, v_{l,i}(t)) - D_{\psi_l}(\bar{x}_l, x_{l,i}(t+1))
- D_{\psi_l}(x_{l,i}(t+1), v_{l,i}(t))
\leq D_{\psi_l}(\bar{x}_l, v_{l,i}(t)) - D_{\psi_l}(\bar{x}_l, x_{l,i}(t+1))
\leq \sum_{j=1}^{n_l} w_{l,ij}(t) D_{\psi_l}(\bar{x}_l, x_{l,j}(t)) - D_{\psi_l}(\bar{x}_l, x_{l,i}(t+1))
$$

where the equality follows from (11) with $x = v_{l,i}(t)$, $y = \bar{x}_l$, and $z = x_{l,i}(t+1)$, the second inequality holds since $D_{\psi_l}(x_{l,i}(t+1), v_{l,i}(t)) \geq 0$ by (12), and the last inequality follows from Assumption 3 and Jensen’s inequality since $v_{l,i}(t) = \sum_{j=1}^{n_l} w_{l,ij}(t)x_{l,j}(t)$ and $\sum_{j=1}^{n_l} w_{l,ij}(t) = 1$. Therefore, from
Assumption 1(i) and $\sum_{i=1}^{n_1} w_{i,i}(t) = 1$, we have

$$\frac{1}{n_1} \sum_{t=1}^{T} \sum_{i=1}^{n_1} (g_{i,i}(t), x_{i,i}(t+1) - \bar{x}_i)$$

$$\leq \frac{1}{n_1} \sum_{t=1}^{T} \frac{1}{\alpha(t)} \left[ \sum_{i=1}^{n_1} \left( \sum_{j=1}^{n_1} w_{i,j}(t) D_{\psi_i}(\bar{x}_i, x_{i,j}(t)) - D_{\psi_i}(\bar{x}_i, x_{i,i}(t+1)) \right) \right]$$

$$= \frac{1}{n_1} \sum_{t=1}^{T} \sum_{i=1}^{n_1} \frac{1}{\alpha(t)} [D_{\psi_i}(\bar{x}_i, x_{i,i}(t)) - D_{\psi_i}(\bar{x}_i, x_{i,i}(t+1))]$$

$$\leq \frac{1}{n_1} \sum_{t=1}^{T} \sum_{i=1}^{n_1} \left[ \frac{1}{\alpha(1)} D_{\psi_i}(\bar{x}_i, x_{i,i}(1)) + \sum_{t=2}^{T} D_{\psi_i}(\bar{x}_i, x_{i,i}(t)) \left( \frac{1}{\alpha(t)} - \frac{1}{\alpha(t-1)} \right) \right]$$

$$\leq \frac{1}{n_1} \sum_{t=1}^{T} \sum_{i=1}^{n_1} \left[ \frac{\Upsilon_i}{\alpha(1)} + \sum_{t=2}^{T} \left( \frac{\Upsilon_i}{\alpha(t)} - \frac{\Upsilon_i}{\alpha(t-1)} \right) \right] \leq \frac{\Upsilon_i}{\alpha(T)}.$$  \hspace{1cm} (27)

**Proof of Theorem 1:** Denote $\bar{x}_i(t) = \frac{1}{n_1} \sum_{s=1}^{n_1} x_{i,s}(t)$ and $x_i^o(t)$ as $\arg\min_{x_{i,s}\in X_i} \sum_{t=1}^{T} f_1(x_{1,s}, x_{2,s}(t))$. Recalling from Assumption 1(iii) that $f_1(x_1, x_2)$ is $L_{1,1}$-Lipschitz continuous in $x_1 \in X_1$ for any $x_2 \in X_2$, we have

$$R_1^{(i)}(T) = \frac{1}{n_1} \sum_{t=1}^{T} (f_1(x_{1,i}(t), u_{2,i}(t)) - f_1(x_{1,i}(t), u_{2,i}(t)))$$

$$= \sum_{t=1}^{T} \left( f_1(x_{1,i}(t), u_{2,i}(t)) - f_1(\bar{x}_i(t), u_{2,i}(t)) + f_1(\bar{x}_i(t), u_{2,i}(t)) - f_1(x_{1,i}(t), u_{2,i}(t)) \right)$$

$$\leq \sum_{t=1}^{T} \left( f_1(\bar{x}_i(t), u_{2,i}(t)) - f_1(x_{1,i}(t), u_{2,i}(t)) \right)$$

$$\leq L_{1,1} \sum_{i=1}^{T} \|x_{i,i}(t) - \bar{x}_i(t)\|.$$  \hspace{1cm} (28)

Similarly, by Assumption 1(iii),

$$A_{1,i}(t) = f_1(\bar{x}_i(t), u_{2,i}(t)) - f_1(x_{1,i}(t), \bar{x}_2(t))$$

$$+ f_1(x_{1,i}(t), \bar{x}_2(t)) - f_1(x_{1,i}(t), \bar{x}_2(t))$$

$$+ f_1(x_{1,i}(t), \bar{x}_2(t)) - f_1(x_{1,i}(t), u_{2,i}(t))$$

$$\leq 2L_{1,2}\|u_{2,i}(t) - \bar{x}_2(t)\|$$

$$+ f_1(\bar{x}_i(t), \bar{x}_2(t)) - f_1(x_{1,i}(t), \bar{x}_2(t)).$$  \hspace{1cm} (29)

As a result

$$R_1^{(i)}(T) \leq \sum_{t=1}^{T} (2L_{1,2}\|u_{2,i}(t) - \bar{x}_2(t)\| + L_{1,1}\|x_{i,i}(t)\| - \bar{x}_i(t)) + \sum_{t=1}^{T} B_i(t).$$  \hspace{1cm} (30)

Furthermore, Assumption 1(ii) and (iii) imply

$$B_i(t) \leq \frac{1}{n_1} \sum_{i=1}^{n_1} \left( f_1(x_{1,i}(t), \bar{x}_2(t)) - f_1(x_{1,i}(t), \bar{x}_2(t)) \right)$$

$$+ f_1(x_{1,i}(t), \bar{x}_2(t)) - f_1(v_{1,i}(t), \bar{x}_2(t))$$

$$+ f_1(v_{1,i}(t), \bar{x}_2(t)) - f_1(v_{1,i}(t), u_{2,i}(t))$$

$$+ f_1(v_{1,i}(t), u_{2,i}(t)) - f_1(x_{1,i}(t), \bar{x}_2(t))$$

$$+ f_1(v_{1,i}(t), u_{2,i}(t)) - f_1(x_{1,i}(t), u_{2,i}(t))$$

$$\leq \frac{1}{n_1} \sum_{i=1}^{n_1} \left[ L_{1,1}(\|x_{i,i}(t) - \bar{x}_i(t)\| + \|x_{i,i}(t) - v_{i,i}(t)\|)$$

$$+ 2L_{1,2}\|u_{2,i}(t) - \bar{x}_2(t)\| + \|g_{i,i}(t), v_{i,i}(t) - \bar{x}_2(t)\| \right].$$  \hspace{1cm} (31)

By the Cauchy–Schwarz inequality, we obtain

$$C_{1,i}(t) = \langle g_{i,i}(t), v_{i,i}(t) - x_{i,i}(t+1) \rangle$$

$$+ \langle g_{i,i}(t), x_{i,i}(t+1) - x_i^o \rangle$$

$$\leq \|g_{i,i}(t)\| \|v_{i,i}(t) - x_{i,i}(t+1)\|$$

$$+ \|g_{i,i}(t), x_{i,i}(t+1) - x_i^o \|$$

$$\leq L_{1,1}^2 \frac{\alpha(t)}{\sigma_1} + \langle g_{i,i}(t), x_{i,i}(t+1) - x_i^o \rangle.$$  \hspace{1cm} (32)

where the last inequality follows from Lemma 4 and (9). Consequently

$$R_1^{(i)}(T) \leq \sum_{t=1}^{T} (2L_{1,2}\|u_{2,i}(t) - \bar{x}_2(t)\| + L_{1,1}\|x_{i,i}(t) - \bar{x}_i(t)\|$$

$$+ \|x_{i,i}(t) - v_{i,i}(t)\| + 2L_{1,2}\|u_{2,i}(t) - \bar{x}_2(t)\|$$

$$+ \|g_{i,i}(t), x_{i,i}(t+1) - x_i^o \|$$

$$+ \sum_{t=1}^{T} \frac{\alpha(t)}{\sigma_1} L_{1,1}^2.$$  \hspace{1cm} (33)

Combining (14) and (15) yields

$$\|x_{i,i}(t) - v_{i,i}(t)\| \leq 2H_i(t).$$  \hspace{1cm} (34)

Substituting (14), (16), (17), and (34) into (33), we obtain

$$R_1^{(i)}(T) \leq \frac{4}{\sigma_1} \sum_{i=1}^{n_1} \left( n_1 L_{1,1}^2 \Gamma_1 \sum_{t=1}^{T} s^{\alpha(t)-1} \right) + \frac{8}{\sigma_2} \sum_{i=1}^{n_1} \left( 2L_{1,1}^2 \sum_{t=1}^{T} \alpha(t) - 1 \right)$$

$$+ \sum_{t=1}^{T} \frac{\alpha(t)}{\sigma_1} L_{1,1}^2 + \frac{8}{\sigma_2} L_{1,2} L_{2,1} \sum_{t=1}^{T} \alpha(t) - 1.$$
\[ + \frac{4}{\sigma^2} \sum_{i=1}^{T} \left( n_2 L_{1,2} \Gamma_2 \sum_{i=1}^{T-1} \theta_i^{t-1} \alpha(s-1) \right) \]
\[ + \frac{4}{\sigma^2} \sum_{i=1}^{2} n_i \Gamma_i \theta_i^{t-1} \Lambda_i + \frac{\gamma_i^2}{\alpha(T)} \] (35)

By exchanging the order of summation, for \( l \in \{1, 2\} \), we obtain
\[ \sum_{t=1}^{T} \sum_{s=1}^{T-1} \theta_i^{t-1} \alpha(s-1) \leq \frac{1}{1 - \theta_i} \sum_{t=1}^{T} \alpha(t-1). \] (36)

This combined with (35) produces (18).

**Appendix B**

**Proofs of Section IV**

**Proof of Theorem 2:** By the convexity of \( f_{1,i} \) with respect to \( x_1 \) and recalling that \( g_{1,i}(t) \in \partial f_{1,i}(v_{1,i}(t), u_{2,i}(t)) \), for all \( x_1 \in X_1 \), we have
\[ \langle g_{1,i}(t), x_1 - v_{1,i}(t) \rangle \leq f_{1,i}(x_1, u_{2,i}(t)) - f_{1,i}(v_{1,i}(t), u_{2,i}(t)) \]
\[ = f_{1,i}(x_1, u_{2,i}(t)) + f_{1,i}(v_{1,i}(t), u_{2,i}(t)) - f_{1,i}(x_1, \bar{x}_2(t)) \]
\[ + f_{1,i}(x_1, \bar{x}_2(t)) - f_{1,i}(x_1, \bar{x}_2(t)) \]
\[ + f_{1,i}(v_{1,i}(t), \bar{x}_2(t)) - f_{1,i}(v_{1,i}(t), u_2(t)) \]
\[ = L \| x_1(t) - \bar{x}_1(t) \| + 2 \| x_2(t) - \bar{x}_2(t) \| \]
\[ + f_{1,i}(v_{1,i}(t), \bar{x}_2(t)) - f_{1,i}(v_{1,i}(t), u_2(t)) \] (37)

where \( L = \max\{L_{1,1}, L_{1,2}, L_{2,1}, L_{2,2}\} \) and the last inequality follows from Assumption 1(iii). Moreover, by Young’s inequality \( (2a, b) \leq c\|a\|^2 + \frac{\sigma}{2} \|b\|^2 \) and \( \|g_{1,i}(t)\| \leq L \), we get
\[ \langle \alpha(t) g_{1,i}(t), v_{1,i}(t) - x_1(t) \rangle + 1 \]
\[ \leq \sigma^2 \| g_{1,i}(t) \|^2 + \frac{\sigma_i}{2} \| v_{1,i}(t) - x_1(t) \|^2 \] (38)

Then, with \( U(\cdot, \cdot) \equiv f_{1}(\cdot, \cdot) \), we obtain
\[ \sum_{i=1}^{n_1} \langle \alpha(t) g_{1,i}(t), x_1 - x_1(t) \rangle \]
\[ \leq \alpha(t) \sum_{i=1}^{n_1} (f_{1,i}(x_1, \bar{x}_2(t)) - f_{1,i}(\bar{x}_1(t), \bar{x}_2(t))) \]
\[ + \sigma_i^2 \sum_{i=1}^{n_1} \| v_{1,i}(t) - x_1(t) \|^2 \]
\[ + \alpha(t) L \sum_{i=1}^{n_1} e_{1,i}(t) \]
\[ = n_1 \alpha(t) U(x_1, \bar{x}_2(t)) - U(\bar{x}_1(t), \bar{x}_2(t)) \]
\[ + \sigma_i^2 \sum_{i=1}^{n_1} \| v_{1,i}(t) - x_1(t) \|^2 \] (39)

where \( e_{1,i}(t) = \| v_{1,i}(t) - \bar{x}_1(t) \| + 2 \| u_{2,i}(t) - \bar{x}_2(t) \|. \) On the other hand, from Lemma 2, we get
\[ (\nabla \psi_1(x_{1,1}(t+1)) - \nabla \psi_1(v_{1,1}(t)), x_1 - x_{1,1}(t+1)) \]
\[ = D_{\psi_1}(x_1, v_{1,1}(t)) - D_{\psi_1}(x_1, x_{1,1}(t+1)) \]
\[ - D_{\psi_1}(x_{1,1}(t+1), v_{1,1}(t)) \] (40)

and
\[ D_{\psi_1}(x_{1,1}(t+1), v_{1,1}(t)) \geq \sigma_i^2 \| x_{1,1}(t+1) - v_{1,1}(t) \|^2. \] (41)

Meanwhile, by Assumption 3 and Jensen’s inequality,
\[ D_{\psi_1}(x_1, v_{1,1}(t)) = D_{\psi_1}(x_1, \sum_{j=1}^{n_1} w_{1,ij} x_{1,j}(t)) \]
\[ \leq \sum_{j=1}^{n_1} w_{1,ij}(t) D_{\psi_1}(x_1, x_{1,j}(t)). \] (42)

Note by exchanging the summation order that
\[ \sum_{j=1}^{n_1} \sum_{i=1}^{n_1} w_{1,ij}(t) D_{\psi_1}(x_1, x_{1,j}(t)) \]
\[ = \sum_{j=1}^{n_1} \sum_{i=1}^{n_1} w_{1,ij}(t) D_{\psi_1}(x_1, x_{1,i}(t)). \]

Since \( \sum_{i=1}^{n_1} w_{1,ij}(t) = 1 \), we further get
\[ \sum_{j=1}^{n_1} \sum_{i=1}^{n_1} w_{1,ij}(t) D_{\psi_1}(x_1, x_{1,j}(t)) \]
\[ = \sum_{i=1}^{n_1} \sum_{j=1}^{n_1} w_{1,ij}(t) D_{\psi_1}(x_1, x_{1,i}(t)). \]

Thus, plugging (41) and (42) back in (40), we derive
\[ \sum_{i=1}^{n_1} (\nabla \psi_1(x_{1,1}(t+1)) - \nabla \psi_1(v_{1,1}(t)), x_1 - x_{1,1}(t+1)) \]
\[ \leq \sum_{i=1}^{n_1} (D_{\psi_1}(x_1, v_{1,1}(t)) - D_{\psi_1}(x_1, x_{1,1}(t+1))) \]
\[ - \sigma_i^2 \sum_{i=1}^{n_1} \| x_{1,1}(t) - v_{1,1}(t+1) \|^2. \] (43)

From (8) and the optimality condition, for all \( x_1 \in X_1 \), we have
\[ (\nabla \psi_1(x_{1,1}(t+1)) - \nabla \psi_1(v_{1,1}(t)) + \alpha(t) g_{1,i}(t), \]
\[ x_1 - x_{1,1}(t+1)) \geq 0. \] (44)

Summing up (39) and (43), with (44), we obtain
\[ 0 \leq \sum_{i=1}^{n_1} (D_{\psi_1}(x_1, x_{1,1}(t)) - D_{\psi_1}(x_1, v_{1,1}(t))) \]
\[ + n_1 \alpha(t) U(x_1, \bar{x}_2(t)) - U(\bar{x}_1(t), \bar{x}_2(t)) \]
\[ + \sigma_i^2 \sum_{i=1}^{n_1} \| v_{1,i}(t) - x_1(t) \|^2 \]
\[ + \alpha(t) L \sum_{i=1}^{n_1} e_{1,i}(t). \]

Rearranging the terms yields
\[
\begin{align*}
\sum_{i=1}^{n_1} D_{\psi_1}(x_1, x_{1,i}(t+1)) & \\
& \leq \sum_{i=1}^{n_1} D_{\psi_1}(x_1, x_{1,i}(t)) + \alpha(t)L \sum_{i=1}^{n_1} e_{1,i}(t) + \alpha^2(t) \frac{n_1 L^2}{2\sigma_1} \\
& \quad + n_1 \alpha(t)(U(x_1, \bar{x}_2(t)) - U(\bar{x}_1(t), \bar{x}_2(t))). \quad (45)
\end{align*}
\]

Similarly, for subnetwork \( \Sigma_2 \), we get
\[
\begin{align*}
\sum_{i=1}^{n_2} D_{\psi_2}(x_2, x_{2,i}(t+1)) & \\
& \leq \sum_{i=1}^{n_2} D_{\psi_2}(x_2, x_{2,i}(t)) + \alpha(t)L \sum_{i=1}^{n_2} e_{2,i}(t) + \alpha^2(t) \frac{n_2 L^2}{2\sigma_2} \\
& \quad + n_2 \alpha(t)(U(\bar{x}_1(t), \bar{x}_2(t)) - U(\bar{x}_1(t), \bar{x}_2(t))) \quad (46)
\end{align*}
\]

where \( e_{2,i}(t) = \|e_{2,i}(t) - \bar{x}_2(t)\| + 2\|u_{1,i}(t) - \bar{x}_1(t)\| \).

Let \( x^t = (x_1^t, x_2^t) \) be the NE. Consider the Lyapunov function
\[
V(t, x_1^t, x_2^t) \triangleq \frac{1}{n_1} \sum_{i=1}^{n_1} D_{\psi_1}(x_1^t, x_{1,i}(t)) \\
+ \frac{1}{n_2} \sum_{i=1}^{n_2} D_{\psi_2}(x_2^t, x_{2,i}(t)).
\]

Obviously
\[
\begin{align*}
V(t+1, x_1^t, x_2^t) & = V(t, x_1^t, x_2^t) - \alpha(t)(U(\bar{x}_1(t), x_2^t) - U(x_1^t, \bar{x}_2(t))) \\
& \quad + \alpha(t)L \sum_{i=1}^{n_1} \sum_{s=1}^{\sigma_1} e_{1,i}(t) + \alpha^2(t)L^2 \sum_{i=1}^{n_1} \frac{2}{2\sigma_1}. \quad (47)
\end{align*}
\]

Note by the definition of NE that
\[
U(\bar{x}_1(t), x_2^t) \geq U(x_1^t, x_2^t) \geq U(x_1^t, \bar{x}_2(t)) \quad (48)
\]

and \( \sum_{i=1}^{\infty} \alpha^2(t) < \infty \). Therefore, from Lemma 7, it remains to show
\[
\sum_{i=1}^{\infty} \alpha(t) \left( \sum_{s=1}^{\sigma_1} e_{1,i}(t) + \sum_{s=1}^{\sigma_2} e_{2,i}(t) \right) < \infty. \quad (49)
\]

Recalling the bounds in (14)–(16) gives
\[
e_{1,i}(t) \leq C_1 \sum_{s=1}^{t-1} \theta_1^{t-1-s} \alpha(s-1) + C_2 \sum_{s=1}^{t-1} \theta_2^{t-1-s} \alpha(s-1) \\
+ C_3 \alpha(t-1) + \sum_{l=1}^{2} 2n_1 \Gamma_l \theta_l^{t-1} \Lambda_l
\]

where \( C_1 = \frac{n_1 L^2}{\sigma_1} \), \( C_2 = \frac{2n_2 L^2}{\sigma_2} \), and \( C_3 = \frac{2L}{\sigma_1} + \frac{4L}{\sigma_2} \). Thus,
\[
\sum_{t=1}^{T} \alpha(t)e_{1,i}(t) \leq C_1 \sum_{t=1}^{T} \theta_1^{t-1-s} \alpha(s-1) + 2 \sum_{l=1}^{2} n_1 \Gamma_l \Lambda_l \sum_{t=1}^{T} \alpha(t)\theta_l^{t-1} \\
+ C_2 \sum_{t=1}^{T} \theta_2^{t-1-s} \alpha(s-1) + C_3 \sum_{t=1}^{T} \alpha^2(t-1).
\]

Similar to (36), we have
\[
\sum_{t=1}^{T} \sum_{s=1}^{\infty} \theta_l^{t-1-s} \alpha^2(s-1) \leq \frac{1}{1 - \theta_l} \sum_{t=1}^{T} \alpha^2(t-1).
\]

Also, \( \sum_{t=1}^{T} \alpha(t)\theta_l^t \leq \alpha(0) \sum_{t=1}^{T} \theta_l^{t-1} \leq \frac{\alpha(0)}{1 - \theta_l} \). Then
\[
\sum_{t=1}^{T} \alpha(t)e_{1,i}(t) \leq \left( \frac{C_1}{1 - \theta_1} + \frac{C_2}{1 - \theta_2} + C_3 \right) \sum_{t=1}^{T} \alpha^2(t-1) \\
+ 2 \sum_{l=1}^{2} n_1 \Gamma_l \alpha(0) \Lambda_l \frac{\alpha(0)}{1 - \theta_l} \quad (50)
\]

implying \( \sum_{t=1}^{\infty} \alpha(t)e_{1,i}(t) < \infty \). Similarly, \( \sum_{i=1}^{\infty} \alpha(t)e_{2,i}(t) < \infty \), i.e., (49) holds. By Lemma 7, \( V(t, x_1^t, x_2^t) \) converges to a finite number. Furthermore, by (47) and (48), we have
\[
0 \leq \sum_{t=0}^{\infty} \alpha(t)(U(x_1^t, x_2^t) - U(x_1^t, \bar{x}_2(t))) \quad (50)
\]

Therefore, \( \sum_{t=0}^{\infty} \alpha(t) = \infty \) yields \( \lim_{t \to \infty} U(x_1^t, x_2^t) - U(x_1^t, \bar{x}_2(t)) = 0 \). Hence, there exists a subsequence \( \{t_r\} \) such that
\[
\lim_{r \to \infty} U(x_1^t, \bar{x}_2(t_r)) = U(x_1^t, \bar{x}_2(t_r)) = \lim_{t \to \infty} U(x_1^t, \bar{x}_2(t_r)). \quad (51)
\]

Let \( (\bar{x}_1, \bar{x}_2) \) be a limit point of the bounded sequence \( \{x_1(t), x_2(t)\} \). Then, there exists a subsequence \( \{t_{r_p}\} \) such that \( \lim_{p \to \infty} x_1(t_{r_p}) = \bar{x}_1 \), and hence, by the continuity of \( U(\cdot, \cdot) \), \( U(x_1^t, \bar{x}_2(t_r)) = U(x_1^t, \bar{x}_2(t_{r_p})) \). From the strict convexity–concavity of \( U \), the NE is unique, i.e., \( (\bar{x}_1, \bar{x}_2) = (x_1^t, x_2^t) \). Using (14) and (4, Lemma 5.2), we obtain
\[
\lim_{p \to \infty} x_1(t_{r_p}) = \lim_{p \to \infty} \bar{x}_1(t_{r_p}) = x_1^t \quad \lim_{p \to \infty} x_2(t_{r_p}) = \lim_{p \to \infty} \bar{x}_2(t_{r_p}) = x_2^t.
\]

According to Assumption 3, \( \lim_{p \to \infty} V(t_{r_p}, x_1^t, x_2^t) = 0 \). Moreover, by the convergence of \( V(t, x_1^t, x_2^t) \), we have
\[
\lim_{t \to \infty} V(t, x_1^t, x_2^t) = \lim_{t \to \infty} V(t, x_1^t, x_2^t) = 0
\]

which is incorporated with (12) to obtain (19).

Proof of Theorem 3: By Assumptions 1(ii) and 1(iii), \( U \) is convex and Lipschitz continuous in \( x_1 \in A_1 \) for any \( x_2 \). With the Jensen’s inequality, we obtain
\[
U(\hat{x}_1(t), x_2) \leq \frac{1}{t} \sum_{s=0}^{t-1} U(x_1(s), x_2) \\
= \frac{1}{t} \sum_{s=0}^{t-1} U(\bar{x}_1(s), x_2) - U(\bar{x}_1(s), x_2) + U(x_1(s), x_2)) \\
\leq \frac{1}{t} \sum_{s=0}^{t} U(\bar{x}_1(s), x_2) + L \|\bar{x}_1(s) - x_1(s)\| \\
= \frac{1}{t} \sum_{s=0}^{t} \left[ \frac{1}{n_2} \sum_{i=1}^{n_2} (f_{2,i}(u_{1,i}(s), x_2) - f_{2,i}(\bar{x}_1(s), x_2) \\
- f_{2,i}(u_{1,i}(s), x_2)) + L \|\bar{x}_1(s) - x_1(s)\| \right] \\
\leq \frac{1}{t} \sum_{s=0}^{t} \left[ \frac{1}{n_2} \sum_{i=1}^{n_2} (f_{2,i}(u_{1,i}(s), x_2) + L\|u_{1,i}(s) - \bar{x}_1(s)\| \\
+ L\|\bar{x}_1(s) - x_1(s)\| \right].
\]
Moreover, from (5) and the convexity of $f_{2,i}$ with respect to $x_2$
\[-f_{2,i}(u_{1,i}(s), x_2)\]
\[= -f_{2,i}(u_{1,i}(s), x_2) + f_{2,i}(u_{1,i}(s), v_{2,i}(s))\]
\[= f_{2,i}(u_{1,i}(s), v_{2,i}(s))\]
\[\leq g_2(s), v_{2,i}(s) - x_2 - f_{2,i}(u_{1,i}(s), v_{2,i}(s)).\] (52)

Substituting (52) into (51) yields $U(x,t), x_2) \leq \frac{t}{t} \sum_{s=0}^{t} F_2(s)$ with $F_2(s)$ defined as
\[F_2(s) = \frac{1}{n_2} \sum_{i=1}^{n_2} \langle g_{2,i}(s), v_{2,i}(s) - x_2 \rangle\]
\[\quad - \frac{1}{n_2} \sum_{i=1}^{n_2} f_{2,i}(u_{1,i}(s), v_{2,i}(s))\]
\[\quad + \left[ \frac{1}{n_2} \sum_{i=1}^{n_2} (L||u_{1,i}(s) - \hat{x}_1(s)|| - ||L||\hat{x}_1(s) - x_{1,i}(s)||) + [L||\hat{x}_1(s) - x_{1,i}(s)||].\] (53)

Similarly, $-f_{1,i}(x_1, u_{2,i}(s)) \leq (g_{1,i}(s), v_{1,i}(s) - x_1) - f_{1,i}(v_{1,i}(s), u_{2,i}(s))$ and
\[U(x_1, \hat{x}_{2,j}(t)) \leq \frac{t}{t} \sum_{s=0}^{t-1} F_1(s) \text{with}\]
\[F_1(s) = \frac{1}{n_1} \sum_{i=1}^{n_1} \langle g_{1,i}(s), v_{1,i}(s) - x_1 \rangle\]
\[\quad - \frac{1}{n_1} \sum_{i=1}^{n_1} f_{1,i}(v_{1,i}(s), u_{2,i}(s))\]
\[\quad + \left[ \frac{1}{n_1} \sum_{i=1}^{n_1} (L||u_{2,i}(s) - \hat{x}_2(s)||) + L||\hat{x}_2(s) - x_{2,j}(s)||.\] (54)

Consequently, $U(x_1, \hat{x}_{2,j}(t)) \leq \frac{t}{t} \sum_{s=0}^{t-1} F_1(s)$ with
\[\leq \frac{1}{n_2} \sum_{i=1}^{n_2} \sum_{s=0}^{t} (L||\hat{x}_1(s) - u_{1,i}(s)|| + L||\hat{x}_2(s) - v_{2,i}(s)||\]
\[+ \frac{1}{n_1} \sum_{i=1}^{n_1} \sum_{s=0}^{t} (L||\hat{x}_1(s) - u_{1,i}(s)|| + L||\hat{x}_2(s) - v_{2,i}(s)||).\] (55)

According to Lemma 5 and $K_i = \frac{1}{\theta_1} (\frac{\gamma L_{\xi}}{\theta_1} + 2L)$, for $l = 1, 2$,
\[\|\hat{x}_l(s) - x_{l,i}(s)\| \leq K_l \alpha + n_1 \Gamma_{l-1} \alpha_1\] (56)
\[\|\hat{x}_l(s) - v_{l,i}(s)\| \leq K_l \alpha + n_1 \Gamma_{l-1} \alpha_1\] (57)
\[\|\hat{x}_l(s) - u_{l,i}(s)\| \leq K_l \alpha + n_1 \Gamma_{l-1} \alpha_1.\] (58)

Applying (57) and (58) in (55), we obtain
\[-\frac{1}{n_2} \sum_{i=1}^{n_2} f_{2,i}(u_{1,i}(s), v_{2,i}(s)) - \frac{1}{n_1} \sum_{i=1}^{n_1} f_{1,i}(v_{1,i}(s), u_{2,i}(s))\]
\[\leq 2L(K_1 + K_2) \alpha + 2L \sum_{s=0}^{t} n_1 \Gamma_{l-1} \alpha_1.\] (59)

Adding (53) and (54) and using (56), (58), and (59), we derive
\[U(x_1, \hat{x}_{2,j}(t)) - U(x_1, \hat{x}_{2,j}(t))\]
\[\leq \frac{1}{t} \sum_{s=0}^{t-1} \frac{1}{n_2} \sum_{i=1}^{n_2} (g_{2,i}(s), v_{2,i}(s) - x_2) + 4L(K_1 + K_2) \alpha\]
\[+ \frac{1}{t} \sum_{s=0}^{t-1} \frac{1}{n_1} \sum_{i=1}^{n_1} (g_{1,i}(s), v_{1,i}(s) - x_1)\]
\[+ 4L \sum_{s=0}^{t-1} n_1 \Gamma_{l-1} \alpha_1.\] (60)

By (17), for $l = 1, 2$, for all $x_l \in \mathcal{X}_l$, we have
\[\frac{1}{t} \sum_{s=0}^{t-1} \frac{1}{n_l} \sum_{i=1}^{n_l} (g_{l,i}(s), x_{l,i}(s+1) - x_l) \leq \frac{\gamma^2}{L_\alpha}.\]

This together with (13) yields
\[\frac{1}{t} \sum_{s=0}^{t-1} \frac{1}{n_l} \sum_{i=1}^{n_l} (g_{l,i}(s), v_{l,i}(s) - x_l)\]
\[= \frac{1}{t} \sum_{s=0}^{t-1} \frac{1}{n_l} \sum_{i=1}^{n_l} (g_{l,i}(s), v_{l,i}(s) - x_{l,i}(s+1))\]
\[+ \frac{1}{t} \sum_{s=0}^{t-1} \frac{1}{n_l} \sum_{i=1}^{n_l} (g_{l,i}(s), x_{l,i}(s+1) - x_l)\]
\[\leq \frac{\gamma^2}{L_\alpha} + \frac{L^2}{\delta_1} \alpha.\] (61)

Since $(x_1^*, x_2^*)$ is an NE,
\[\max_{x_2 \in \mathcal{X}_2} U(x_1, x_2) \geq U(x_1^*, x_2^*) \geq U(x_1^*, x_2^*)\]
\[\min_{x_1 \in \mathcal{X}_1} U(x_1, \hat{x}_{2,j}(t)) \leq U(x_1^*, x_{2,j}(t)) \leq U(x_1^*, x_{2,j}(t))\]

Substituting (61) into (60), we obtain
\[U(x_1, x_2^*) - U(x_1, x_2^*)\]
\[\leq \max_{x_1, x_2} |U(x_1, x_2) - U(x_1, \hat{x}_{2,j}(t))|\]
\[
\begin{align*}
&\leq \sum_{l=1}^{2} \left( \frac{\gamma^2}{\sigma \alpha} + \frac{L^2}{\sigma \lambda} \right) + 4L(K_1 + K_2)\alpha \\
&+ 4L \sum_{t=0}^{t-1} \sum_{l=1}^{2} n_l \Gamma_l \theta_l^{-1} \Lambda_l
\end{align*}
\]
which completes the proof.

**APPENDIX C**

**PROOFS OF SECTION V**

**Proof of Theorem 4:** Denote by \( \mathcal{X}_1^* \times \mathcal{X}_2^* \) the mixed-strategy NE set, and suppose that \((x_1^+, x_2^+)\) is an interior point of \( \mathcal{X}_1^* \times \mathcal{X}_2^* \). Then, there exists \( \epsilon > 0 \), such that
\[
\mathcal{B}(x_1^+, x_2^+, \epsilon) \subset \mathcal{X}_1^* \times \mathcal{X}_2^*.
\]

Let \((\hat{x}_1, \hat{x}_2)\) and \((\hat{x}_1, \hat{x}_2)\) be any two limit points of the sequence \(\{\tilde{x}_1(t), \tilde{x}_2(t)\}\). By (14), they are also the limit points of \(\{x_{1,i}(t), x_{2,i}(t)\}\), and we set the corresponding convergent subsequences as \(\{x_{1,i}(t), x_{2,i}(t)\}\) and \(\{x_{1,i}(t), x_{2,i}(t)\}\). Recall from the proof of Theorem 1 that the strict convexity–concavity of \(U\) is not used when proving that \(V(t, x_1^*, x_2^*)\) converges to a finite number. Thus, for all \((x_1^+, x_2^+)\) \(\in \mathcal{X}_1^* \times \mathcal{X}_2^*\), \((t, x_1^*, x_2^*)\) still converges to a finite number. Therefore, for all \((x_1, x_2)\) \(\in \mathcal{B}(x_1^+, x_2^+, \epsilon)\), by the continuity of the Bregman divergence, we have
\[
\lim_{r \to \infty} D_{\psi_1}(x_1, x_{1,i}(t_r)) = D_{\psi_2}(x_1, \hat{x}_1), \quad l = 1, 2.
\]
Furthermore, we get
\[
\lim_{r \to \infty} V(t_r, x_1, x_2) = \frac{1}{n_1} \sum_{i=1}^{n_1} \lim_{r \to \infty} D_{\psi_1}(x_1, x_{1,i}(t_r)) + \frac{1}{n_2} \sum_{i=1}^{n_2} \lim_{r \to \infty} D_{\psi_2}(x_2, x_{2,i}(t_r)) = D_{\psi_1}(x_1, \hat{x}_1) + D_{\psi_2}(x_2, \hat{x}_2).
\]
As a result
\[
D_{\psi_1}(x_1, \hat{x}_1) + D_{\psi_2}(x_2, \hat{x}_2) = \lim_{r \to \infty} V(t_r, x_1, x_2) = \lim_{s \to \infty} V(t_s, x_1, x_2) = D_{\psi_1}(x_1, \hat{x}_1) + D_{\psi_2}(x_2, \hat{x}_2).
\]
Set \(x_2 = x_2^+\) in (62), and then, for all \((x_1, \epsilon)\), we have
\[
D_{\psi_1}(x_1, \hat{x}_1) - D_{\psi_1}(x_1, \hat{x}_1) = D_{\psi_1}(x_1^+, \hat{x}_2) - D_{\psi_2}(x_2^+, \hat{x}_2).
\]
Taking the derivative with respect to \(x_1\) on both sides, we obtain
\[
\nabla \psi_1(x_1) - \nabla \psi_1(x_1) = \nabla \psi_1(x_1) - \nabla \psi_1(x_1).
\]
Therefore, by the strong convexity of \(\psi_1\), we have
\[
0 = (\nabla \psi_1(x_1) - \nabla \psi_1(x_1), \hat{x}_1 - \hat{x}_1) \geq \|\hat{x}_1 - \hat{x}_1\|^2.
\]

Moreover, since \(\sum_{t=0}^{\infty} \alpha(t) = \infty\),
\[
\lim_{t \to \infty} (U(\hat{x}_1(t), \hat{x}_2(t)) - U(x_1, \hat{x}_2(t))) \leq 0.
\]
The continuity of \(U\) yields \(U(\hat{x}_1, \hat{x}_2) \leq U(x_1, \hat{x}_2), \forall x_1 \in \mathcal{X}_1\). Similarly, by (46), \(U(\hat{x}_1, \hat{x}_2) \leq U(x_1, \hat{x}_2), \forall x_2 \in \mathcal{X}_2\). By the definition of NE, \((\hat{x}_1, \hat{x}_2) \in \mathcal{X}_1^* \times \mathcal{X}_2^*\).

**Proof of Theorem 5:** With analysis similar to that of Theorem 3 and recalling the definitions of \(F_2(s)\) and \(F_1(s)\) in (53) and (54), for any \((x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2\), we obtain
\[
-U(\hat{x}_1, \hat{x}_2) \leq \frac{1}{\sum_{s=0}^{t-1} \alpha(s)} \sum_{s=0}^{t-1} \alpha(s)F_2(s)
\]
and
\[
-U(\hat{x}_1, \hat{x}_2) \leq \frac{1}{\sum_{s=0}^{t-1} \alpha(s)} \sum_{s=0}^{t-1} \alpha(s)F_1(s)
\]
Adding (63) and (64) and using (55) yield
\[
U(\hat{x}_1, \hat{x}_2) - U(x_1, \hat{x}_2(s)) \leq \frac{1}{\sum_{s=0}^{t-1} \alpha(s)} \sum_{s=0}^{t-1} \alpha(s)g_2(s, v_2(s) - x_2)
\]

with \(e_1(s)\) defined in Theorem 2. From Lemmas 4 and 6, for any \((x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2\), we have
\[
\sum_{s=0}^{t-1} \alpha(s)g_1(s, v_1(s) - x_1)
\]

Furthermore, by (50), there exists a constant \(\hat{C} > 0\) such that
\[
\sum_{s=0}^{t-1} \alpha(s)e_1(s) \leq \hat{C} \sum_{s=0}^{t-1} \alpha^2(s).
\]

Similarly, \(\sum_{s=0}^{t-1} \alpha(s)\|x_1(s) - x_1(s)\| \leq \hat{C} \sum_{s=0}^{t-1} \alpha^2(s)\). Therefore, substituting (66) and (67) into (65), and using Assumption 4, we have
\[
\max_{x_2 \in \mathcal{X}_2} U(\hat{x}_1, x_2) - \min_{x_1 \in \mathcal{X}_1} U(x_1, \hat{x}_2) \to 0, \quad t \to \infty.
\]
Consider the following gap function:

$$
\epsilon(\hat{x}_1, \hat{x}_2) = U^* - \min_{x_1 \in K_1} U(x_1, \hat{x}_2) + \max_{x_2 \in K_2} U(\hat{x}_1, x_2) - U^*
$$

where $U^*$ is the cost value of NE points. Then, $\epsilon(\hat{x}_1, \hat{x}_2) \geq 0$, and the equality holds if and only if $(\hat{x}_1, \hat{x}_2)$ is an NE. By (68), $\epsilon(\hat{x}_1(t), \hat{x}_2(t)) > 0$, which implies that $(\hat{x}_1(t), \hat{x}_2(t))$ converges to the mixed-strategy NE set.

**References**

[1] S. Boyd, S. P. Boyd, and L. Vandenberghe. *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[2] D. Goldberg and G. Iyengar. "Robust portfolio selection problems," *Math. Oper. Res.*, vol. 28, no. 1, pp. 1–38, 2003.

[3] S. A. Kassam and H. V. Poor. "Robust techniques for signal processing: A survey," *Proc. IEEE*, vol. 73, no. 3, pp. 433–481, Mar. 1985.

[4] Y. Lou, Y. Hong, L. Xie, G. Shi, and K. H. Johansson. "Learning in games with lossy feedback," *in Proc. Int. Conf. Neural Inf. Process. Syst.*, 2018, pp. 1–11.

[5] K. Lu, G. Li, and L. Wang. "Online distributed algorithms for seeking generalized Nash equilibria in dynamic environments," *IEEE Trans. Autom. Control*, vol. 66, no. 5, pp. 2259–2266, May 2021.

[6] N. Eshraghi and B. Liang. "Distributed online optimization over a heterogeneous network with any-batch mirror descent," *in Proc. Int. Conf. Mach. Learn.*, 2020, pp. 2933–2942.

[7] A. Nedić, S. Lee, and M. Raginsky. "Decentralized online optimization with global objectives and local communication," *in Proc. Amer. Control Conf.*, 2015, pp. 4497–4503.

[8] S. Shahrampour and A. Jadbabaie. "Distributed online optimization in dynamic environments using mirror descent," *IEEE Trans. Autom. Control*, vol. 63, no. 3, pp. 714–725, Mar. 2018.

[9] D. Yuan, D. W. Ho, and G.-P. Jiang. "An adaptive primal-dual subgradient algorithm for online distributed constrained optimization," *IEEE Trans. Cybern.*, vol. 48, no. 11, pp. 3045–3055, Nov. 2018.

[10] B. Gharesifard and J. Cortes. "Distributed convergence to Nash equilibria in two-network zero-sum games," *Automatica*, vol. 49, no. 6, pp. 1683–1692, 2013.

[11] A. Nedić and A. Ozdaglar. "Subgradient methods for saddle-point problems," *J. Optim. Theory Appl.*, vol. 142, no. 1, pp. 205–228, 2009.

[12] S. Yuan, J. Lei, Y. Hong, and U. V. Shanbhag. "No-regret distributed learning in two-network zero-sum games," *in Proc. IEEE 56th Conf. Decis. Control*, 2021, pp. 924–929.

[13] K. Srivastava, A. Nedić, and D. Stipanović. "Distributed Bregman-distance algorithms for min-max optimization," *in Agent-Based Optimization*. Berlin, Germany: Springer, 2013, pp. 143–174.

[14] A. Nedić, A. Ozdaglar, and P. A. Parrilo. "Constrained consensus and optimization in multi-agent networks," *IEEE Trans. Autom. Control*, vol. 55, no. 4, pp. 922–938, Apr. 2010.

[15] Y.-G. Hsieh, K. Antonakopoulos, and P. Mertikopoulos. "Adaptive learning in continuous games: Optimal regret bounds and convergence to Nash equilibrium," *in Proc. Conf. Learn. Theory*, 2021, pp. 2388–2422.

[16] S. Talebi, S. Alemzadeh, L. J. Ratliff, and M. Mesbahi. "Distributed learning in network games: A dual averaging approach," *in Proc. IEEE 58th Conf. Decis. Control*, 2019, pp. 5544–5549.

[17] T. T. Doan, S. Bose, D. H. Nguyen, and C. L. Beck. "Convergence of the itertes in mirror descent methods," *IEEE Control Syst. Lett.*, vol. 3, no. 1, pp. 114–119, Jan. 2019.

[18] A. Nemirovski and D. B. Yudin. *Problem Complexity and Method Efficiency in Optimization*. Hoboken, NJ, USA: Wiley, 1983.

[19] A. Beck and M. Teboulle. "Mirror descent and nonlinear projected subgradient methods for convex optimization," *Oper. Res. Lett.*, vol. 31, no. 3, pp. 167–175, 2003.

[20] H. Wang, X. Liao, T. Huang, and C. Li. "Cooperative distributed optimization in multiagent networks with delays," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 45, no. 2, pp. 363–369, Feb. 2015.

[21] C. Xi, Q. Wu, and U. A. Khan. "Distributed mirror descent over directed graphs," arXiv:1412.5526, 2014.

[22] J. Li, G. Chen, Z. Dong, and Z. Wu. "Distributed mirror descent method for multi-agent optimization with delay," *Neurocomputing*, vol. 177, pp. 643–650, 2016.

[23] M. Bravo, D. Leslie, and P. Mertikopoulos. "Bandit learning in concave N-person games," *in Proc. Int. Conf. Neural Inf. Process. Syst.*, 2018, pp. 5661–5671.

[24] Y. Freund and R. E. Schapire. "Adaptive game playing using multiplicative weights," *Games Econ. Behav.*, vol. 29, no. 1–2, pp. 79–103, 1999.

[25] D. F. Porter and R. V. Vohra. "Calibrated learning and correlated equilibrium," *Games Econ. Behav.*, vol. 21, no. 1–2, 1997, Art. no. 40.

[26] B. T. Polyak. *Introduction to Optimization*. New York, NY, USA: Optimization Software, Inc., 1987.

[27] P. Mertikopoulos, B. Lecouturier, H. Zenati, C.-S. Foo, V. Chandrasekhar, and G. Piliouras. "Optimistic mirror descent in saddle-point problems: Going the extra (gradient) mile," *in Proc. Int. Conf. Learn. Represent.*, 2019, pp. 1–23.

[28] A. Washburn and K. Wood. "Two-person zero-sum games for network interdiction," *Oper. Res.*, vol. 43, no. 2, pp. 243–251, 1995.
Shijie Huang received the B.S. degree in mathematics from the Huazhong University of Science and Technology, Wuhan, China, in 2018 and the Ph.D. degree in systems theory from the Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China, in 2023. He is currently a Postdoctoral Researcher with the Delft Center of Systems and Control, TU Delft, The Netherlands. His research interests include distributed optimization, game theory and feedback optimization.

Jinlong Lei (Member, IEEE) received the B.E. degree in automation from the University of Science and Technology of China, Hefei, China, in 2011, and the Ph.D. degree in operations research and cybernetics from the Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China, in 2016. She is currently a Professor with the Department of Control Science and Engineering, Tongji University, Shanghai, China. She was a Postdoctoral Fellow with the Department of Industrial and Manufacturing Engineering, Pennsylvania State University, University Park, PA, USA, from 2016 to 2019. Her research interests include stochastic approximation, stochastic and distributed optimization, and network games under uncertainty.

Yiguang Hong (Fellow, IEEE) received the B.S. and M.S. degrees in mechanics from Peking University, Beijing, China, in 1987 and 1990, respectively, and the Ph.D. degree in control theory from the Chinese Academy of Sciences (CAS), Beijing, China, in 1993. He is currently a Professor and Deputy Director of the Shanghai Research Institute for Intelligent Autonomous Systems, Tongji University, Shanghai, China. He was a Professor with the Academy of Mathematics and Systems Science, CAS, and the Director of the Key Laboratory of Systems and Control, CAS. His current research interests include nonlinear control, multiautonomy, distributed optimization and game, machine learning, and social networks.

Dr. Hong is a recipient of the Guang Zhaozhi Award at Chinese Control Conference, the Young Author Prize of IFAC World Congress, the Young Scientist Award of CAS, the Youth Award for Science and Technology of China, and the National Natural Science Prize of China. He is the Editor-in-Chief for Control Theory and Technology. He serves or served as an Associate Editor for many journals, including IEEE TRANSACTIONS ON AUTOMATIC CONTROL, IEEE TRANSACTIONS ON CONTROL OF NETWORK SYSTEMS, and IEEE Control Systems Magazine. He was a Board of Governor of the IEEE Control Systems Society (CSS), the Chair of IEEE CSS Membership and Public Information Committee, and the Chair of the IEEE CSS Chapter Activities Committee. He is a Fellow of the Chinese Association for Artificial Intelligence and the Chinese Association of Automation.

Uday V. Shanbhag received the bachelor's degree in aerospace engineering from the Indian Institute of Technology Bombay, Mumbai, India, in 1993, the master's degree in operations research from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 1998, and the Ph.D. degree in operations research from the Department of Management Science and Engineering, Stanford University, Stanford, CA, USA, in 2006. He has held the Gary and Sheila Bello Chair in Industrial and Manufacturing Engineering with Pennsylvania State University, University Park, PA, USA, since 2016. His research interests include the development of analytical tools and scalable computational schemes for optimization and equilibrium problems, with a focus on addressing uncertainty, nonsmoothness, and nonconvexity.

Dr. Shanbhag was the recipient of the triennial A.W. Tucker Prize by the Mathematical Programming Society in 2006, the NCSA Faculty Fellowship in 2006, the Computational Optimization and Applications Best Paper Award (with advisor Walter Murray) in 2007, and the Best Theoretical Paper Award at Winter Simulation Conference in 2013 (with Angela Nedich and Farzad Yousfian). He was also the recipient of the National Science Foundation Faculty Early Career Development Award in 2012 from the Operations Research Program. He is on the Editorial Board of SIAM Journal of Optimization, Computational Optimization and Applications, and Optimization Letters. He was an Associate Editor for IEEE TRANSACTIONS ON AUTOMATIC CONTROL.