2019 IEEE 58th Conference on Decision and Control (CDC)
Palais des Congrès et des Expositions Nice Acropolis
Nice, France, December 11-13, 2019

Distributed Learning in Network Games: a Dual Averaging Approach

Shahriar Talebi†, Siavash Alemzadeh†, Lillian J. Ratliff‡, and Mehran Mesbahi†

Abstract— **In this paper, we propose a distributed no-regret learning algorithm for network games using a primal-dual method, i.e., dual averaging. With only locally available observations, we consider the scenario where each player optimizes a global objective, formed by local objective functions on the nodes of a given communication graph. Our learning algorithm for each player involves taking steps along their individual payoff gradients, dictated by local observations of the other player's actions. The output is then projected back–again locally–to the set of admissible actions for each player. We provide the regret analysis of this distributed learning algorithm for the case of a deterministic network that is subjected to two teams with distinct objectives, and obtain an $O(\sqrt{T}\log(T))$ regret bound. Our analysis indicates the key correlation between the rate of convergence and network connectivity that also appears in the distributed optimization setup via dual averaging. Furthermore, we show that the point of convergence of the proposed algorithm would be a Nash Equilibrium of the game. Finally, we showcase by an illustrative example the performance of our algorithm in relation to the size and connectivity of the network.**

I. INTRODUCTION

Networked systems analysis has been on the forefront of multi-disciplinary research over the past few years, with applications ranging from robotic swarms to biological networks. Common to such systems, a global objective is sought to be achieved based on local interactions–which in turn– require local decision-making that is inherently restricted by limited information exchange and prescribed set of admissible policies [1]. As such, many computationally efficient optimization algorithms [2]–[5] when implemented on a network have scaling issues as a function of the network size. Not surprisingly, there has been an extensive literature on exploiting the structure of an information-exchange graph in order to reduce the complexity of distributed methods built upon subgradient optimization [6], mirror descent and dual averaging [7], [8], and sequential decision-making [9]. In the meantime, while the aforementioned works mainly focus on a cooperative information sharing, many of real-world systems exhibit non-cooperative behaviors due to a plethora of reasons including intrusions/attacks, greedy agents, competition for constrained resources, misaligned incentives, or the adversarial nature of environments that necessitate a game-theoretical model.

† William E. Boeing Department of Aeronautics and Astronautics, University of Washington, Seattle WA, USA {shahriar, alems, mesbahi}@uw.edu.

‡ Department of Electrical Engineering, University of Washington, Seattle WA, USA {ratliffl}@ece.uw.edu.

The research of the authors has been supported by NSF grant SES-1541025 and Air Force Office of Scientific Research under grant number FA9550-16-1-0022.

Game theory has been successfully employed in non-cooperative decision-making [10]–[16]. As the key solution concept, Nash Equilibrium (NE) plays a central role in game theoretic analysis defined, in a nutshell, as a strategy from which no player has an incentive to deviate from unilaterally. Upon convexity and continuous differentiability assumptions, the NE can be characterized as the solution of a variational inequality (VI) problem [17], where its existence can be guaranteed under assumptions on action sets and the Jacobian of the game [17], [18]. Distributed Nash equilibrium seeking and multiagent network games refers to the class of algorithms that aim to learn a global NE via local information exchange [19]–[26], where each node of the network represents a player in the game.

In this work, we introduce a non-cooperative game between two teams, each consisting of players that interact over a network. While the goal of each team is to learn the global NE, players have no information from the other team, neither do they enjoy a global decision-making capability. Instead, only based on local observations, players choose a local strategy at each node and receive a local cost, resulting in learning the global NE in a distributed fashion. This setup resembles a scenario where each node in the network is subjected to actions that contribute to distinct objectives–in this sense, there is duality in the interactions between nodes in the network. For example, each node can represent a socio-economic entity, with objectives that are not completely aligned with each other (altruistic vs. profit-seeking). Of prime interest in this work is how algebraic and combinatorial properties of the network, such as its connectivity, contribute to the performance of distributed learning in the context of games. The choice of a primal/dual approach that is built around dual averaging [3], and its online implementation [4], is primarily motivated by this overarching objective; this choice is also consistent with how dual averaging has been used in the context of distributed optimization to underscore the "network-effect" on the convergence and optimality properties of distributed optimization, where there is no "duality" in the nodes' operation. However, in order to extend the work of [8] to the game setting, one has to pay a particular attention to the information structure, as for example, it would be unreasonable to assume information sharing with the opponents in the game setting.

The rest of the paper is organized as follows: In §II we provide a quick overview of mathematical tools that are used in the paper. In §III we introduce the problem setup and propose our method. In §IV we show that our algorithm enjoys an $O(\sqrt{T}\log(T))$ regret bound similar to the dual averaging methods which is well-known to be tight in black-box setting [2]. Furthermore, we analyze the point of convergence of the proposed algorithm under smoothness assumptions. Lastly, we provide an illustrative example in §V and also concluding remarks and future research directions in §VI.

II. Mathematical Preliminaries

We denote by $\mathbb{R}$ the set of real numbers. A column vector with n elements is designated as $v \in \mathbb{R}^n$, where v_i represents its ith element. The matrix $M \in \mathbb{R}^{p\times q}$ contains p rows and q columns with M_{ij} denoting the element in its ith row and jth column. The square matrix $N \in \mathbb{R}^{n\times n}$ is *symmetric* if $N^\top = N$, where $N^\top$ denotes the *transpose* of the matrix N. The $n \times 1$ vector of all ones is denoted by $\mathbb{1}$. A doubly stochastic matrix $P \in \mathbb{R}^{n\times n}$ is defined as a non-negative square matrix such that $\sum_j P_{ij} = \sum_j P_{jk} = 1$ for all i and k, and $\sigma_2(P)$ indicates the second largest singular value of P. We define $[n] = \{1, \dots, n\}$. The *Euclidean norm* of a vector $x \in \mathbb{R}^n$ is defined as $\|x\| = (x^\top x)^{1/2} = (\sum_{i=1}^n x_i^2)^{1/2}$ and the *dual norm* to $\|x\|$ is denoted by $\|x\|_* := \sup_{\|u\|=1}\langle x, u\rangle$ (with respect to the inner product $\langle ., .\rangle$); the *1-norm* is defined as $\|x\|_1 = \sum_{i=1}^n |x_i|$. A function f is convex if $f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta) f(y)$ for all $\theta \in (0, 1)$ and for all x, y in its convex domain, and g is a subgradient of f at point z if $f(y) \ge f(z) + \langle g, y - z\rangle$ for all y. The set of all subgradients of f at x is called its subdifferential and denoted by $\partial f(x)$.

III. Problem Setup

In this section we introduce the main framework of our analysis. Herein, we briefly mention relevant background material on dual averaging which is the workhorse of the proposed methodology for distributed learning on networks. We then proceed to propose the distributed setup followed by a two-player game-theoretic framework, which can naturally be generalized to a multi-player setting.

A. Dual Averaging

The dual averaging algorithm proposed by Nesterov [3] is a subgradient scheme for non-smooth convex problems. The primal-dual nature of this method generates a two sequence iterates $\{x(t), z(t)\}_{t=0}^\infty$ contained within $\mathcal{X}\times\mathbb{R}^d$ for a convex set $\mathcal{X}$, such that the update of $z(t)$ is responsible for averaging the support functions in the dual space, while the updates of $x(t)$ result from a scaled dynamic update between the primal and dual spaces. More precisely, after receiving the subgradient $g(t) \in \partial f(x(t))$ at iteration t, the algorithm is updated as,

$$\begin{aligned} z(t+1) &= z(t) + \gamma(t) g(t), \\ x(t+1) &= \Pi_{\mathcal{X}}^{\psi}(-z(t+1), \alpha(t)), \end{aligned} \tag{1}$$

where $\gamma(t) > 0$, $\{\alpha(t)\}_{t=0}^\infty$ is a positive non-increasing sequence, and

$$\Pi_{\mathcal{X}}^{\psi}(z, \alpha) := \text{argmin}_{x\in\mathcal{X}}\Big\{\langle -z, x\rangle + \frac{1}{\alpha}\psi(x)\Big\}, \tag{2}$$

is a generalized projection according to a strongly convex *prox-function* ψ.

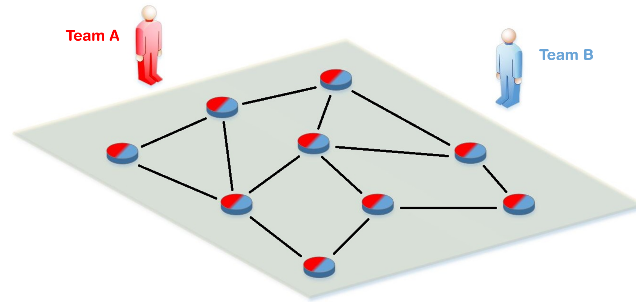


Fig. 1: Schematic of our game setup including two teams (red A and blue B), having a representative at every node.

B. Our Model

We consider the distributed learning problem for a game between two teams (players), both playing on a network consisting of n nodes interacting via a connected communication graph $\mathcal{G}$.[1] To this end, each team has a representative on each node, hence $2n$ members in total (Figure 1). The teams are grouped within the sets $\mathcal{I}_\ell = \{(\ell, 1), \dots, (\ell, n)\}$ for $\ell \in \{A, B\}$ and each team has a choice of action x_ℓ in a convex set $\mathcal{X}_\ell \subset \mathbb{R}^{d_\ell}$ that minimizes the following global cost,

$$f_\ell(x_A, x_B) = \tfrac{1}{n}\textstyle\sum_{i=1}^n f_{\ell,i}(x_A, x_B), \tag{3}$$

which is the average of its members' costs at each node i, denoted by $f_{\ell,i}$. As in the general game-theoretic setup, a solution concept of interest in our work is that of *Nash Equilibrium* (NE), defined as a strategy tuple (x_A^*, x_B^*) for which,

$$\begin{cases} f_A(x_A^*, x_B^*) \le f_A(x_A, x_B^*), & \forall x_A \in \mathcal{X}_A \\ f_B(x_A^*, x_B^*) \le f_B(x_A^*, x_B), & \forall x_B \in \mathcal{X}_B. \end{cases}$$

The goal of the distributed algorithm is to learn a global NE while players have no global decision-making capability.[2] Instead, the network $\mathcal{G}$ is assumed to be connected and players can communicate within their own respective teams according to the network structure.

To learn a global NE, each team updates the state of its nodes using a distributed dual averaging type method. The network-based information flow of our algorithm is related to [8], particularly as it pertains to the consensus error of the dual variable; however, in a non-cooperative game-theoretic setup, convergence of such iterative methods to NE is non-trivial due to the nature of equilibria and limited observation.

In the proposed algorithm, at each node i and iteration t, player $\ell \in \{A, B\}$ maintains an estimate of its team's action as $x_{\ell,i}(t)$. A communication protocol is designed for sharing dual variables among the nodes of each team, where node i updates its dual variable $z_{\ell,i}(t)$ using a convex combination of those of its neighboring teammates. Then the node maps $z_{\ell,i}(t)$ back to the set of admissible actions $\mathcal{X}_\ell$ followed by

[1] It is worth noting that different networks could be associated with each team with the communication graph representing their overlap.

[2] Note that in the team setting, a strategy tuple would be a NE if it is so according to all local cost pairs; however, the converse is not necessarily valid.

taking its local action $x_{\ell,i}(t)$. Subsequently, players observe the actions of the opponent at each node and locally obtain an estimate of the subgradient of their respective distributed cost. In this paper, under some regularity assumptions, we show that this process results in a sub-linear regret bound for each player.

C. Team-based Dual Averaging

We assume that the structure of $\mathcal{G}$ induces a doubly stochastic matrix P_ℓ available to each team, where $P_{\ell,ij} > 0$ if and only if nodes i and j are connected. Then at iteration t, each player $\ell \in \{A, B\}$ performs the following updates at each node $i \in V$,

$$\begin{aligned} z_{\ell,i}(t+1) &= \sum_{j \in \mathcal{N}_{\ell,i}} P_{\ell,ij} z_{\ell,j}(t) + \gamma_\ell(t) g_{\ell,i}(t), \\ x_{\ell,i}(t+1) &= \Pi^{\psi}_{\mathcal{X}_\ell}\big(- z_{\ell,i}(t+1), \alpha_\ell(t)\big), \end{aligned} \tag{4}$$

where $z_{\ell,i}$ and $x_{\ell,i}$ are the dual variable and the local action of player ℓ at node i respectively; $g_{\ell,i}$ on the other hand is a subgradient of the local cost $f_{\ell,i}$ at the local actions $x_{\ell,i}(t)$ as,

$$g_{\ell,i}(t) \in \partial_\ell f_{\ell,i}(x_{A,i}(t), x_{B,i}(t)), \tag{5}$$

where ∂_ℓ is the differential with respect to the action of player ℓ. Finally, $\alpha_\ell(t)$ and $\gamma_\ell(t)$ are sequences of positive stepsize with $\alpha_\ell(t)$ being non-increasing. Note that $x_{\ell,i}$ can be viewed as the local copy of x_ℓ at node i, and its updates require access to only the ith row of the matrix P_ℓ. We refer to the updates in (4) as *Team-based dual averaging (TDA)*. The proposed methodology is summarized in Algorithm 1. We define the *running local average* at node i, for player $\ell \in \{A, B\}$ as,

$$\hat{x}_{\ell,i}(t) \coloneqq \tfrac{1}{t} \textstyle\sum_{s=0}^{t} x_{\ell,i}(s). \tag{6}$$

In the sequel, we first show the sub-linear regret bound for TDA algorithm with appropriate choices of stepsize, and then analyze its point of convergence.

IV. Main Results

In this section, we present the main results of the paper and the corresponding analysis. We first make the following definition that is used in the subsequent analysis.

***Definition** 1:* A network game is called *Lipschitz convex* if the cost of player $\ell \in \{A, B\}$ at node $i \in [n]$ satisfies,

- $f_{\ell,i}(\,.\,, x_{-\ell}) : \mathcal{X}_\ell \mapsto \mathbb{R}$ is convex $\forall x_{-\ell} \in \mathcal{X}_{-\ell}$.
- $f_{\ell,i}(\,.\,, x_{-\ell}) \ : \ \mathcal{X}_\ell \mapsto \mathbb{R}$ is L_ℓ-Lipschitz continuous $\forall x_{-\ell} \in \mathcal{X}_{-\ell}$ such that $|f_{\ell,i}(x_\ell, x_{-\ell}) - f_{\ell,i}(x'_\ell, x_{-\ell})| \le L_\ell \|x_\ell - x'_\ell\|$, $\forall x_\ell, x'_\ell \in \mathcal{X}_\ell$.

Distinct from optimization problems where the notion of "optimality" plays a central role, in a game setup, the objective is seeking an equilibrium rather than an "optimal" solution. In general, in order to incorporate the interactions of players in convergence analysis of NE-learning algorithms, monotonicity has been known as a useful sufficient condition for problem "regularity", originally due to the seminal work of Rosen [18]. In contrast, with no further regularity assumptions, we will prove a sub-linear regret bound for

Algorithm 1 Team-based Dual Averaging

1: **Inputs**: For player ℓ
2: Local black-box oracle at node i to compute a subgradient of the local cost at any test point
3: Doubly stochastic matrix P_ℓ induced by the network structure
4: **Outputs**:
5: Estimates of NE $\hat{x}_{\ell,i}(t)$ at node i for player ℓ
6: **Initialize**:
7: $t = 0$
8: **For** Player $\ell \in \{A, B\}$ at node $i \in [n]$
9: $z_{\ell,i}(0) = 0$
10: Take random action $x_{\ell,i}(0)$
11: **while convergence**
12: **For** Player $\ell \in \{A, B\}$ at node $i \in [n]$
13: Observe the opponents local action and get
14: $g_{\ell,i}(t) \in \partial_\ell f_{\ell,i}(x_{A,i}(t), x_{B,i}(t))$
15: **For** Player $\ell \in \{A, B\}$ at node $i \in [n]$
16: Update the dual variable $z_{\ell,i}(t+1)$ via (4)
17: Compute and take action $x_{\ell,i}(t+1)$ via (4)
18: $t = t + 1$
19: **Return:** the *running average* $\hat{x}_{\ell,i}$ by (6)

TDA algorithm. Convergence of the action iterates to the unique NE can be established using similar monotonicity assumptions, and will be provided in our subsequent work.

A. Regret Analysis

First, we introduce two results that facilitate the basic convergence analysis. Herein, we borrow some tools from prior works, as such, we only mention the key steps that distinguish our contribution in the game setup. Furthermore, it is worth mentioning that even though we analyze TDA algorithm with $\gamma_\ell(t) = 1$, judicious choice of γ_ℓ can in fact improve the corresponding convergence rate.

Let us now define $\tilde{f}_A(\,.\,; t)$ and $\tilde{f}_B(\,.\,; t)$ as,

$$\begin{aligned} \tilde{f}_A(\,.\,; t) &\coloneqq \tfrac{1}{n} \textstyle\sum_{j=1}^{n} f_{A,j}(\,.\,, x_{B,j}(t)), \\ \tilde{f}_B(\,.\,; t) &\coloneqq \tfrac{1}{n} \textstyle\sum_{j=1}^{n} f_{B,j}(x_{A,j}(t), \,.\,), \end{aligned}$$

where $x_{A,i}(t)$ and $x_{B,j}(t)$ are the sequences generated by the TDA algorithm. Now for $\ell \in \{A, B\}$, we define $\mathcal{R}_{\ell,i}$ to be the regret of player ℓ for implementing the TDA algorithm at node i versus any fixed action x^*_ℓ, while the other player has committed to implement the TDA, i.e.,

$$\mathcal{R}_{\ell,i}(T) \coloneqq \textstyle\sum_{t=1}^{T} \tilde{f}_\ell(x_{\ell,i}(t); t) - \sum_{t=1}^{T} \tilde{f}_\ell(x^*_\ell; t). \tag{7}$$

***Lemma** 1:* Following Algorithm 1, suppose that player $\ell \in \{A, B\}$ has access to $g_{\ell,i}$ for $(\ell, i) \in \mathcal{I}_\ell$ at each node $i \in [n]$, and $\{\alpha_\ell(t)\}_{t=0}^{\infty}$ is a non-increasing sequence of positive stepsizes and $\gamma_\ell(t) = 1$. Then for any Lipschitz convex network game we have,

$$\begin{aligned} \mathcal{R}_{\ell,i}(T) \le{}& \tfrac{\psi(x^*_\ell)}{\alpha_\ell(T)} + \tfrac{L^2_\ell}{2} \textstyle\sum_{t=1}^{T} \alpha_\ell(t-1) \\ &+ L_\ell \textstyle\sum_{t=1}^{T} \alpha_\ell(t) \Big[\mathcal{Z}_{\ell,i}(t) + \tfrac{2}{n} \sum_{j=1}^{n} \mathcal{Z}_{\ell,i}(t) \Big], \end{aligned}$$

for any fixed $x^* = (x_A^*, x_B^*) \in \mathcal{X}_A \times \mathcal{X}_B$, where $\mathcal{Z}_{\ell,i}(t) = \|\bar{z}_\ell(t) - z_{\ell,i}(t)\|_*$ is the consensus error at each time with $\bar{z}_\ell(t) = \frac{1}{n}\sum_{i=1}^n z_{\ell,i}(t)$ as the averaging term in the dual space.

Proof: For each team define,

$$y_\ell(t) := \Pi_{\mathcal{X}_\ell}^{\psi}(-\bar{z}_\ell(t), \alpha(t)). \tag{8}$$

The doubly stochastic nature of P_ℓ implies an iterative form of $\bar{z}_\ell(t)$ as,

$$\bar{z}_\ell(t+1) = \bar{z}_\ell(t) + \tfrac{1}{n}\textstyle\sum_{j=1}^n g_{\ell,j}(t).$$

Hence, by the zero choice of the dual initialization we get,

$$y_\ell(t) = \text{argmin}_{x\in\mathcal{X}_\ell}\left\{\sum_{s=1}^{t-1}\langle\frac{1}{n}\sum_{i=1}^n g_{\ell,i}(s), x\rangle + \frac{1}{\alpha_\ell(t)}\psi(x)\right\}.$$

Now using the L_A-Lipschitz property of f_A we can show that,

$$\mathcal{R}_{A,i}(T) \le \textstyle\sum_{t=1}^T\left[\tilde{f}_A(y_A(t)) - \tilde{f}_A(x_A^*)\right] + \sum_{t=1}^T L_A\|y_A(t) - x_{A,i}(t)\|.$$

Also, by adding and subtracting $\sum_j f_{A,j}(x_{A,j}(t), x_{B,j}(t))$ to the first term and using convexity we have,

$$\mathcal{R}_{A,i}(T) \le \textstyle\sum_{t=1}^T \frac{1}{n}\sum_{j=1}^n\langle g_{A,j}(t), x_{A,j}(t) - x_A^*\rangle + \sum_{t=1}^T\frac{L_A}{n}\sum_{j=1}^n\left[\|y_A(t) - x_{A,j}(t)\| + \|y_A(t) - x_{A,i}(t)\|\right],$$

where $g_{A,j}(t) \in \partial_A f_{A,j}(x_{A,j}(t), x_{B,j}(t))$ and the L_A-Lipschitz condition is leveraged again. Then by adding and subtracting $y_A(t)$ in the inner-product we observe that,

$$\begin{aligned}\mathcal{R}_{A,i}(T) &\le \textstyle\sum_{t=1}^T \frac{1}{n}\sum_{j=1}^n\langle g_{A,j}(t), y_A(t) - x_A^*\rangle \\ &\quad+ \textstyle\sum_{t=1}^T\frac{L_A}{n}\sum_{j=1}^n\left[2\|y_A(t) - x_{A,j}(t)\| + \|y_A(t) - x_{A,i}(t)\|\right] \\ &\le \textstyle\sum_{t=1}^T\langle\frac{1}{n}\sum_{j=1}^n g_{A,j}(t), y_A(t) - x_A^*\rangle \\ &\quad+ \textstyle\frac{L_A}{n}\sum_{t=1}^T\alpha_A(t)\sum_{j=1}^n\left[2\mathcal{Z}_{A,j}(t) + \mathcal{Z}_{A,i}(t)\right],\end{aligned}$$

resulting from the L_A-Lipschitz property of f_A and α-Lipschitz continuity of the generalized projection $\Pi_{\mathcal{X}}^{\psi}(., \alpha)$ (which is a direct consequence of Lemma 1 in [3]) . Thereby, the lemma follows by applying Theorem 2 and Equation (3.3) in [3] (also restated as Lemma 3 in [8]) to the first term of the inequality above and using the fact that $\|g_{A,j}\|_* \le L_A$. Similar analysis results in the bound for $\mathcal{R}_{B,i}(T)$. ■

B. Choice of the Learning Rate $\alpha_\ell(t)$

In order to achieve convergence, an appropriate choice of stepsize (learning rate) $\alpha_\ell(t)$ is required. The next result shows how specific choices of $\alpha_\ell(t)$ result in practical bounds on $\mathcal{R}_{\ell,i}$ by essentially bounding the $\mathcal{Z}_{\ell,k}$ terms.

***Theorem** 1:* Under the notation adopted in Lemma 1, suppose that $\psi(x_\ell^*) \le R_\ell^2$. Then by choosing the stepsize,

$$\alpha_\ell(t) = \frac{R_\ell\sqrt{1-\sigma_2(P_\ell)}}{\sqrt{13}L_\ell\sqrt{t}},$$

for any Lipschitz convex network game we obtain,

$$\mathcal{R}_{\ell,i}(T) \le \sqrt{T}\log(T\sqrt{n})\frac{2\sqrt{13}R_\ell L_\ell}{\sqrt{1-\sigma_2(P_\ell)}}.$$

Proof: Stacking the updates of dual variables (4) into a matrix $Z_\ell = [z_{\ell,1} \ \dots \ z_{\ell,n}]$ and similarly $G_\ell = [g_{\ell,1} \ \dots \ g_{\ell,n}]$, for an undirected graph ($P_\ell^\top = P_\ell$) leads to,

$$Z_\ell(t+1) = Z_\ell(t)P_\ell + G_\ell(t).$$

Define $\Phi_\ell(t+1, s) = P_\ell^{t+1-s}$, where Φ is the transition matrix for a discrete-time linear system with state Z. Thus,

$$Z_\ell(t+1) = Z(s)\Phi_\ell(t+1, s) + \textstyle\sum_{r=s+1}^{t+1} G_\ell(r-1)\Phi_\ell(t+1, r),$$

for $0 \le s \le t$. Note that $\Phi_\ell(t+1, s)\mathbb{1} = \mathbb{1}$ and by definition $\bar{z}_\ell(t) = Z_\ell(t)\mathbb{1}/n$ and $z_{\ell,i}(t) = Z_\ell(t)e_i$, we obtain,

$$\bar{z}_\ell(t) - z_{\ell,i}(t) = \bar{z}_\ell(s) - Z(s)\Phi_\ell(t, s)e_i + \textstyle\sum_{r=s+1}^t G_\ell(r-1)[\mathbb{1}/n - \Phi_\ell(t, r)e_i].$$

For simplicity we assume that $z_{\ell,i}(0) = \bar{z}_\ell(0)$ (say by choosing $z_{\ell,i}(0) = 0$); then we have $\bar{z}_\ell(s) - Z(s)\Phi_\ell(t, s)e_i = 0$ at $s = 0$. This implies that,

$$\bar{z}_\ell(t) - z_{\ell,i}(t) = \textstyle\sum_{r=1}^t G_\ell(r-1)[\mathbb{1}/n - \Phi_\ell(t, r)e_i].$$

We can then proceed to bound this error as,

$$\|\bar{z}_\ell(t) - z_{\ell,i}(t)\|_* \le L_\ell\textstyle\sum_{r=1}^t\|\mathbb{1}/n - \Phi_\ell(t, r)e_i\|_1,$$

where we have used $\|g_{\ell,i}\|_* \le L_\ell$ and norm inequalities. Consider the following standard inequality ([27]),

$$\|\mathbb{1}/n - \Phi_\ell(t, r)e_i\|_1 \le \sqrt{n}\sigma_2(P_\ell)^{t+1-r}.$$

We say that r is "small" if $r \le t+1+\log(T\sqrt{n})/\log\sigma_2(P_\ell)$ (otherwise it is "large"). Then by splitting the sum, one would note that for the small r,

$$\|\mathbb{1}/n - \Phi_\ell(t, r)e_i\|_1 \le \frac{1}{T},$$

since $\sigma_2(P_\ell) < 1$. Otherwise, for the large r,

$$\|\mathbb{1}/n - \Phi_\ell(t, r)e_i\|_1 \le 2.$$

Then it can be shown that,

$$\mathcal{Z}_{\ell,i}(t) = \|\bar{z}_\ell(t) - z_{\ell,i}(t)\|_* \le 2L_\ell\frac{\log(T\sqrt{n})}{1-\sigma_2(P_\ell)}, \tag{9}$$

using $\log\sigma_2(P_\ell)^{-1} \ge 1 - \sigma_2(P_\ell)$. Hence from Lemma 1,

$$\mathcal{R}_{\ell,i}(T) \le \frac{\psi(x_\ell^*)}{\alpha_\ell(T)} + \frac{L_\ell^2}{2}\sum_{t=1}^T\alpha_\ell(t-1) + \frac{6L_\ell^2\log(T\sqrt{n})}{1-\sigma_2(P_\ell)}\sum_{t=1}^T\alpha_\ell(t).$$

Define the sequence $\{\alpha(t)\}_{t=0}^\infty$ as, $\alpha_\ell(t) = K_\ell/\sqrt{t}$, $\alpha_\ell(0) = 1$. Since $\psi(x_\ell^*) \le R_\ell^2$ and $\sum_{t=1}^T t^{-1/2} \le 2\sqrt{T} - 1$, choosing $K_\ell = R_\ell\sqrt{1-\sigma_2(P_\ell)}/\sqrt{13}L_\ell$ completes the proof. ■

C. Convergence of TDA Algorithm

A natural question on the performance of the TDA algorithm pertains to the convergence of the corresponding action iterates. However, it is known that the convergence of action iterates for this general class of algorithms cannot be guaranteed without further regularity assumptions. Nevertheless, with minimal continuity assumptions, our next result ensures that the point of convergence of TDA algorithm is in fact a NE. First a relevant definition.

***Definition** 2:* A network game is called *continuous* if the cost of player $\ell \in \{A, B\}$ at node $i \in [n]$ satisfies,

- $f_{\ell,i}(x_\ell, x_{-\ell})$ is continuously differentiable in x_ℓ,
- $f_{\ell,i}(x_\ell, x_{-\ell})$ and $\nabla f_{\ell,i}(x_\ell, x_{-\ell})$ are both continuous in the joint variable $(x_\ell, x_{-\ell})$.

***Theorem** 2:* Under the notation adopted in Theorem 1, for a continuous network game, if Algorithm 1 converges, i.e., $x_{\ell,i}(t) \to x^*_{\ell,i}$ as $t \to \infty$, then $x^*_{\ell,i} = x^*_\ell$ for all $i \in [n]$ and $(x^*_\ell, x^*_{-\ell})$ is a NE.

Proof: By continuous differentiability, we have that $g_{\ell,i} = \nabla f_{\ell,i}$ for all i and $g_{\ell,i} \to g^*_{\ell,i}$ due to the joint continuity. Now by α_ℓ-Lipschitz continuity of the generalized projection we have,

$$\|x_{\ell,i}(t) - x_{\ell,j}(t)\| \leq \alpha_\ell(t)\left[\mathcal{Z}_{\ell,i}(t) + \mathcal{Z}_{\ell,j}(t)\right].$$

From (9) and the choice of $\alpha_\ell(t) = K_\ell/\sqrt{t}$ we conclude that for all i, j, $\|x_{\ell,i}(t) - x_{\ell,j}(t)\| \to 0$ as $t \to \infty$, and thus $x^*_{\ell,i} = x^*_\ell$ for all i. Similarly, this implies $y_\ell(t) \to x^*_\ell$ since,

$$\|y_\ell(t) - x^*_\ell\| \leq \alpha_\ell(t)\mathcal{Z}_{\ell,i}(t) + \|x_{\ell,i}(t) - x^*_\ell\|.$$

The rest of the proof is by contradiction. Suppose $x^* = (x^*_\ell, x^*_{-\ell})$ is not a NE and define $G^*_\ell = G_\ell(x^*)$. Then by definition of NE in our setup and noting that $\nabla_\ell f_\ell(x^*) = G^*_\ell \mathbb{1}/n$, at least for one of the players (say player ℓ) we have the following (Proposition 1.4.2 in [17]),

$$\exists q_\ell \in \mathcal{X}_\ell \quad \text{s.t.} \quad \langle G^*_\ell \mathbb{1}/n, q_\ell - x^*_\ell \rangle < 0.$$

By continuity there exist a constant $c > 0$ and neighborhoods U, V of points x^*_ℓ, $G^*_\ell \mathbb{1}/n$, respectively, such that,

$$\langle G'_\ell \mathbb{1}/n, q_\ell - x'_\ell \rangle \leq -c, \quad \forall x'_\ell \in U, \quad \forall G'_\ell \text{ s.t. } G'_\ell \mathbb{1}/n \in V.$$

On the other hand, by definition of $\Pi^\psi_\mathcal{X}$ and y_ℓ as in (8), and strong convexity of ψ we can conclude that (Theorem 23.5 in [28]) $-\alpha_\ell(t)\bar{z}_\ell(t) \in \partial\psi(y_\ell(t))$, and therefore,

$$\psi(q_\ell) - \psi(y_\ell(t)) \geq -\alpha_\ell(t)\langle \bar{z}_\ell(t), q_\ell - y_\ell(t) \rangle.$$

Note that by convergence of $y_\ell(t)$, there exists N such that $y_\ell(t) \in U$ and $G_\ell(t)\mathbb{1}/n \in V$, $\forall t \geq N$. Furthermore, $\bar{z}_\ell(t) = \sum_{r=1}^{t-1} G_\ell(r)\mathbb{1}/n$ since $z_{\ell,i}(0) = \bar{z}_\ell(0)$; thus we can conclude that

$$\begin{aligned}&\psi(q_\ell) - \psi(y_\ell(t)) \\ &\geq \alpha_\ell(t)\textstyle\sum_{r=N}^{t-1} c - \alpha_\ell(t)\langle \sum_{r=1}^{N-1} G_\ell(r)\mathbb{1}/n, q_\ell - y_\ell(t) \rangle.\end{aligned}$$

Now as $t \to \infty$ the right hand side of the above inequality approaches positive infinity, which is a contradiction. ■

V. Example

In this section we illustrate the performance of our method for a two-team game on a network where each player has an objective at each node as follows,

$$f_{\ell,i}(x_\ell, x_{-\ell}) = \frac{1}{2}\|x_\ell - a_{\ell,i}\|^2 + \frac{1}{4}\langle x_\ell, x_{-\ell} - b_{\ell,i} \rangle,$$

where $a_{\ell,i}$ and $b_{\ell,i}$ are arbitrary prescribed parameters. Although the cost functions are only locally Lipschitz, they can be treated as a Lipschitz continuous function over any bounded (potentially large) domain. We have simulated the performance of TDA over complete, random 6-regular, and cycle graphs [29] each consisting of 50 nodes. Figure 3(b) shows the random 6-regular network used for our simulations. To illustrate the advantage of dual averaging in our algorithm, we compare the results with another distributed algorithm—referred to as *Team-based Mirror-Descent (TMD)*—which is an extension of *Distributed Mirror Descent* algorithm [7] adopted for our game model with $\psi(.) = \frac{1}{2}\|.\|^2$. Also, as in standard form, the stepsize of this algorithm $\beta_\ell(t) = K_\ell/t^{0.8}$ is chosen to be not summable but square summable where K_ℓ is as defined in Theorem 1. The TMD algorithm is detailed below with the same initialization as TDA. At iteration t, members of each team at each node observe the opponent's action $x_{-\ell,i}(t)$ locally and compute the local average estimate $v_{\ell,i}(t)$ as,

$$v_{\ell,i}(t) = \textstyle\sum_{j \in \mathcal{N}_{\ell,i}} P_{\ell,ij} x_{\ell,j}(t),$$

and get $g_{\ell,i}(t) \in \partial_\ell f_{\ell,i}(v_{\ell,i}(t), x_{-\ell,i}(t))$. Next, they project the average estimates back to the set of actions as,

$$x_{\ell,i}(t+1) = \text{Proj}_{\mathcal{X}_\ell}[v_{\ell,i}(t) - \beta_\ell(t) g_{\ell,i}(t)].$$

Next, we define the Normalized Average Error (NAE) as the normalized mean of the running average error from NE over all nodes of the network as follows,

$$\text{NAE}(t) = \frac{\sum_{i=1}^n \|\hat{x}_{A,i}(t) - x^*_A\| + \|\hat{x}_{B,i}(t) - x^*_B\|}{\sum_{i=1}^n \|\hat{x}_{A,i}(0) - x^*_A\| + \|\hat{x}_{B,i}(0) - x^*_B\|}. \quad (10)$$

Figure 2 shows the NAE at each iteration for both TDA and TMD algorithms. It can be noted that on networks with the same structure, the TDA method has a much better convergence rate as compared with TMD, primary due to the dual averaging nature of TDA. Clearly, as the connectivity of the network decreases from complete to cycle, convergence rate becomes slower. This is captured in Theorem 1 where lower connectivity in the network leads to an increase of $\sigma_2(P_\ell)$ closer to 1, resulting in a larger bound for each time horizon T.

Next, we examine the performance of TDA in terms of the required iteration for a given error (from the NE). To do so, we consider four complete graphs with 25, 50, 80, and 120 nodes and require the algorithm to achieve a NAE of less than 0.1. Figure 3 illustrates the number of iterations T needed for each network to achieve NAE$(T) < 0.1$. Since, the initialization of the algorithm is random, for each network, we have illustrated the mean and variance of the number of iterations required by 30 realizations. It is evident

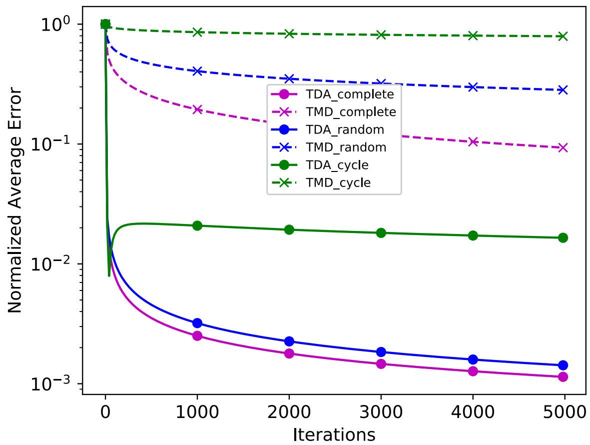
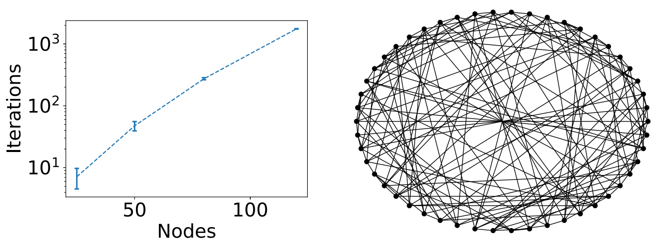


Fig. 2: NAE at each iteration for both TDA and TMD algorithms in complete, random 6-regular, and cycle graphs with 50 nodes.



(a)

(b)

Fig. 3: (a) Number of iterations needed for each network so that NAE is less than 0.1, (b) The random 6-regular network with 50 nodes simulated in the example.

that the number of required iterations to achieve the same error-bound, increases exponentially in this simulation. This behavior is partially due to the fact that TDA is a first-order method in a game setting and it is converging towards an equilibrium point, unlike distributed optimization scenario, where convergence is towards an attractive optimal point.

VI. Conclusion

In this paper, we have proposed a new model for network games where each player's cost is distributed over a network. The setup does not rely on global information sharing for the decision-making process. We have shown that dual averaging can be applied to the scenario where two teams, with distinct objectives, can coordinate with their respective team members, leading to a sub-linear regret bound.

Stochastic access to subgradient of cost functions as well as noisy observations of opponent's action is considered as the next immediate extensions of this work. Also with additional regularity conditions on the game structure, the convergence of action iterates to NE can be established and will be presented in our subsequent work. Finally, analysis of the stepsize for the dual variables and equipping the algorithm with feasible second-order information are yet other directions that can further improve the convergence behavior of this method close to the equilibrium.

References

[1] M. Mesbahi and M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*, vol. 33. Princeton University Press, 2010.

[2] A. Nemirovsky and D. Yudin, *Problem Complexity and Method Efficiency in Optimization.* Wiley, New York, 1983.

[3] Y. Nesterov, "Primal-dual Subgradient Methods for Convex Problems," *Mathematical Programming*, vol. 120, no. 1, pp. 221–259, 2009.

[4] L. Xiao, "Dual averaging methods for regularized stochastic learning and online optimization," *JMLR*, vol. 11, no. 1, pp. 2543–2596, 2010.

[5] M. Sedghi, G. Atia, and M. Georgiopoulos, "Robust manifold learning via conformity pursuit," *IEEE Signal Processing Letters*, vol. 26, pp. 425–429, March 2019.

[6] A. Nedić and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48–61, 2009.

[7] T. T. Doan, S. Bose, D. H. Nguyen, and C. L. Beck, "Convergence of the iterates in mirror descent methods," *IEEE Control Systems Letters*, vol. 3, no. 1, pp. 114–119, 2019.

[8] J. C. Duchi, A. Agarwal, and M. J. Wainwright, "Dual averaging for distributed optimization: Convergence analysis and network scaling," *IEEE Trans. on Automatic Control*, vol. 57, no. 3, pp. 592–606, 2012.

[9] S. Alemzadeh and M. Mesbahi, "Distributed q-learning for dynamically decoupled systems," in *2019 American Control Conference (ACC)*, pp. 772–777, July 2019.

[10] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. Academic Press, 1995.

[11] J. Engwerda, *LQ Dynamic Optimization & Differential Games*. John Wiley & Sons, 2005.

[12] L. Lambertini, *Differential Games in Industrial Economics*. Cambridge University Press, 2018.

[13] M. Simaan and J. B. Cruz, "On the Stackelberg strategy in nonzero-sum games," *Journal of Optimization Theory and Applications*, vol. 11, no. 5, pp. 533–555, 1973.

[14] S. Talebi, M. A. Simaan, and Z. Qu, "Cooperative, non-cooperative and greedy pursuers strategies in multi-player pursuit-evasion games," in *2017 IEEE Conference on Control Technology and Applications (CCTA)*, pp. 2049–2056, Aug 2017.

[15] P. Mertikopoulos and Z. Zhou, "Learning in games with continuous action sets and unknown payoff functions," *Mathematical Programming*, pp. 1–43, 2018.

[16] S. Talebi, M. A. Simaan, and Z. Qu, "Decision-making in complex dynamical systems of systems with one opposing subsystem," in *2019 18th European Control Conference (ECC)*, pp. 2789–2795, June 2019.

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

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

[19] N. Li and J. R. Marden, "Designing games for distributed optimization," *IEEE Journal on Selected Topics in Signal Processing*, vol. 7, no. 2, pp. 230–242, 2013.

[20] F. Salehisadaghiani and L. Pavel, "Distributed Nash equilibrium seeking in networked graphical games," *Automatica*, pp. 17–24, 1 2018.

[21] P. Frihauf, M. Krstic, and T. Basar, "Nash equilibrium seeking in noncooperative games," *IEEE Transactions on Automatic Control*, vol. 57, pp. 1192–1207, May 2012.

[22] T. Alpcan and T. Başar, *Distributed Algorithms for Nash Equilibria of Flow Control Games*, pp. 473–498. Boston, MA: Birkhäuser, 2005.

[23] T. Tatarenko, W. Shi, and A. Nedic, "Accelerated Gradient Play Algorithm for Distributed Nash Equilibrium Seeking," in *2018 IEEE Conference on Decision and Control*, pp. 3561–3566, IEEE, 12 2018.

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

[25] S. Grammatico, "Proximal dynamics in multiagent network games," *IEEE Transactions on Control of Network Systems*, vol. 5, pp. 1707–1716, Dec 2018.

[26] F. Parise and A. Ozdaglar, "A variational inequality framework for network games: Existence, uniqueness, convergence and sensitivity analysis," *Games and Economic Behavior*, no. 2010, pp. 1–43, 2019.

[27] P. Diaconis and D. Stroock, "Geometric bounds for eigenvalues of Markov chains," *Annals of Applied Prob.*, vol. 1, pp. 36–61, 2 1991.

[28] R. T. Rockafellar, *Convex Analysis*. Princeton University Press, 1970.

[29] J. Bondy and U. Murty, *Graph Theory*. Springer Publishing Company, Incorporated, 1st ed., 2008.