

Distributed Online Optimization in Time-Varying Unbalanced Networks without Explicit Subgradients

Yongyang Xiong, Xiang Li, Keyou You, *Senior Member, IEEE*, and Ligang Wu, *Fellow, IEEE*

Abstract—This paper studies a distributed online constrained optimization problem over time-varying unbalanced digraphs without explicit subgradients. In sharp contrast to the existing algorithms, we design a novel consensus-based distributed online algorithm with a local randomized zeroth-order oracle and then rescale the oracle by constructing row-stochastic matrices, which aims to address the unbalancedness of time-varying digraphs. Under mild conditions, the average dynamic regret over a time horizon is shown to asymptotically converge at a sublinear rate provided that the accumulated variation grows sublinearly with a specific order. Moreover, the counterpart of the proposed algorithm when subgradients are available is also provided, along with its dynamic regret bound, which reflects that the convergence of our algorithm is essentially not affected by the zeroth-order oracle. Simulations on distributed targets tracking problem and dynamic sparse signal recovery problem in sensor networks are employed to demonstrate the effectiveness of the proposed algorithm.

Index Terms—Distributed algorithm, online constrained optimization, dynamic regret, time-varying networks.

I. INTRODUCTION

DISTRIBUTED optimization has been undoubtedly attracting a surge of attentions in recent years with the rapid development of large-scale networks. Many practical issues can be solved within the framework of distributed optimization problems, such as privacy preserving [1], [2], resource allocation [3]–[5], sensor fusion [6], just to mention a few. These applications promote the design of distributed algorithms such that a group of nodes cooperatively optimize the sum of their local cost functions via local communications. See [7]–[9] and the references therein.

Despite that distributed optimization algorithms has been studied extensively, the dynamic aspect of the problem has not been fully addressed where the cost functions may vary with time in an uncertain and even adversarial fashion. For example, time-varying cost functions frequently appear in machine learning, where data samples are observed in a sequential manner and the newly observed data samples result in new cost

functions. This inspires us to extend the distributed optimization algorithms to an online setting, in which the local cost functions vary with time and are only revealed to individuals after each node has made a decision. Amongst the existing distributed online optimization algorithms, the subgradient descent methods gain considerable attention [10]–[13]. More recently, the authors of [14]–[17] developed distributed online algorithms on the basis of primal-dual method, gradient tracking, mirror descent approach, and proximal gradient algorithm, respectively. The authors of [18] further introduced a primal-dual mirror descent algorithm to address distributed online problems with time-varying coupled inequality constraints over weight-balanced digraphs. Specifically, the algorithms in [10]–[15] focused on *static regret*, which can be used to measure the performance of estimating a static target in sensor networks. While the authors of [16]–[18] concentrated on *dynamic regret*, which is a more stringent metric than static regret since it allows the best decision in handsignht varies with time. Therefore, the dynamic regret can reflect the performance of tracking moving targets.

However, the aforementioned online algorithms highly rely on the assumption of doubly stochastic weight matrices, which is quite stringent and even deemed impracticable in applications, e.g., computer networks intrinsically operate with directionality and it is difficult to construct doubly stochastic weight matrices in a distributed manner. Extending distributed algorithms from weight balanced networks to general directed networks is non-trivial [19]–[26]. The authors of [27], [28] proposed distributed online optimization algorithms inspired by the push-sum based algorithm [19] and the surplus-based method [24], respectively. However, the former is incapable of tackling constrained optimization problems by combining with projected-based methods directly, and the latter involves a global parameter depending on weight matrices which should be known a priori.

In addition, the closed-form expression of cost function or gradient information may be not available in some scenarios, e.g., online source localization, online routing in data networks [29]. To relieve this bottleneck, zeroth-order algorithms have gained renewed interests in recent years [30]–[34], and been studied under distributed setting by combining the surplus-based method [35], the primal-dual method [36], the gradient tracking method [37], etc. The intuition of these distributed algorithms is constructing gradient estimators from function values and then substituting them for the true gradient under the assumption of doubly stochastic weight matrices or fixed networks. Though some efforts have been made to extend zeroth-order distributed algorithms to the online setting [11],

This work was supported in part by the National Natural Science Foundation of China (61525303, 41772377 and 61673130), Top-Notch Young Talents Program of China (Ligang Wu), and the Self-Planned Task of State Key Laboratory of Robotics and System (HIT) (SKLRS201806B). (*Corresponding author: Ligang Wu.*)

Y. Xiong, and K. You are with the Department of Automation, Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, P. R. China. This work was done when Y. Xiong was with the Department of Control Science and Engineering, Harbin Institute of Technology. E-mail: xiongyy@tsinghua.edu.cn; youky@tsinghua.edu.cn.

X. Li and L. Wu are with the Department of Control Science and Engineering, Harbin Institute of Technology, Harbin 150001, P. R. China. E-mail: lixiang_soa@hit.edu.cn; ligangwu@hit.edu.cn.

TABLE I
A COMPARISON OF OUR WORK WITH THE RELEVANT DISTRIBUTED ONLINE OPTIMIZATION ALGORITHMS

References	Time-Varying Networks	Unbalancedness	Constraint Set	Gradient-Free	Dynamic Regret
[10]	✓	✗	✗	✗	✗
[13], [14]	✓	✗	✓	✗	✗
[27]	✓	✓	✗	✗	✗
[11], [12]	✗	✗	✓	✓	✗
[28]	✗	✓	✓	✓	✗
[42]	✓	✓	✓	✗	✗
[44]	✗	✗	✗	✗	✓
[29]	✓	✗	✓	✓	✓
[40]	✓	✓	✓	✗	✓
[16]	✗	✗	✓	✗	✓
[15], [17], [18], [39]	✓	✗	✓	✗	✓
Our work	✓	✓	✓	✓	✓

[12], [28], they can only be applicable to fixed networks.

All the above motivate us to further explore an algorithm that has stronger adaptability. In this work, we study the distributed online optimization problem over time-varying unbalanced digraphs, under the settings where the decisions of nodes are constrained in a convex set and only the local cost function values are revealed to nodes. We compare this work with the state-of-the-art on distributed online optimization in Table I, and summarize our main contributions as follows:

- (1) We propose a novel distributed online constrained optimization algorithm over time-varying digraphs with a local randomized zeroth-order oracle. We rigorously analyze the dynamic regret of the algorithm. Our results show that the average dynamic regret over a time horizon converges to zero at a sublinear rate if the accumulated variation grows sublinearly with a specific order.
- (2) Inspired by the static push-sum protocols developed in [19], [20], [38], we present a novel methodology that provides a new perspective on the push-sum based distributed online algorithm [27]. Specifically, by dynamically constructing row-stochastic matrices and rescaling the zeroth-order oracle, our algorithm can be applicable to constrained distributed online optimization problems, while the push-sum based online algorithm in [27] is only feasible in the unconstrained case.
- (3) In comparison with the distributed online optimization algorithms [10], [16], [18], [39] relying highly on double stochastic weight matrices, our algorithm can be applicable to unbalanced networks with column-stochastic weight matrices, and allow nodes to utilize zeroth-order information in lieu of subgradients as well. In addition, compared with the fixed networks and static regret considered in [28], our algorithm can be applicable to time-varying networks with dynamic regret guarantees.

The rest of this paper is organized as follows. Section II describes the problem of interest. Section III presents the proposed algorithm. Section IV includes our main results on dynamic regret. Simulations are provided in Section V. We conclude this paper in Section VI.

Notations. The superscript ‘T’ denotes vector transposition; We use $[T] = \{1, \dots, T\}$ to denote a set of integers for $T \in \mathbb{N}_+$; The notation $[A]_{ij}$ or A_{ij} denotes the i, j -th element of matrix A ; $\mathbf{1}_n$ represents a column vector with its all

components equaling to one. We use $\mathcal{P}_\Omega(x)$ to denote the projection operation of a point x onto the set $\Omega \subset \mathbb{R}^m$, i.e., $\mathcal{P}_\Omega(x) = \arg \min_{w \in \Omega} \|w - x\|^2$. A subgradient of convex function f at x is represented by $g(x) \in \partial f(x)$, which satisfies $f(y) \geq f(x) + g^T(x)(y - x)$. For two functions f and h , the notation $f = \mathcal{O}(h)$ means that there exists a positive constant $\rho < \infty$ such that $f \leq \rho h$. $\mathbb{B}^m = \{x \in \mathbb{R}^m \mid \|x\| \leq 1\}$ and $\mathbb{S}^m = \{x \in \mathbb{R}^m \mid \|x\| = 1\}$ denote the unit ball and sphere, respectively. We denote by $\mathbb{E}(x)$ the expectation of x .

II. PRELIMINARIES AND PROBLEM FORMULATION

A. Graph Theory

We consider a network containing n nodes. A sequence of digraphs $\{\mathcal{G}_t(\mathcal{V}, \mathcal{E}_t)\}_{t \in [T]}$ is utilized to model the interactions among the nodes, where $\mathcal{V} = \{1, \dots, n\}$ represents the set of nodes and \mathcal{E}_t denotes the set of interaction links at t . $(i, j) \in \mathcal{E}_t$ implies that node j can receive information from node i at t . We denote $\mathcal{N}_{i,t}^+ = \{j : (j, i) \in \mathcal{E}_t\} \cup \{i\}$ and $\mathcal{N}_{i,t}^- = \{j : (i, j) \in \mathcal{E}_t\} \cup \{i\}$ as the in-neighbor and out-neighbor sets of node i , respectively. The digraph \mathcal{G}_t is strongly connected if there exists a path between any pair of distinct nodes. In addition, $\mathcal{A}_t = [a_{ij,t}] \in \mathbb{R}^{n \times n}$ is the weight matrix induced by \mathcal{G}_t , where $a_{ij,t} > 0$ if and only if $(j, i) \in \mathcal{E}_t$, and $a_{ij,t} = 0$ otherwise. Moreover, we say matrix \mathcal{A}_t is row-stochastic if $\mathcal{A}_t \mathbf{1}_n = \mathbf{1}_n$, column-stochastic if $\mathcal{A}_t^T \mathbf{1}_n = \mathbf{1}_n$, and doubly stochastic if $\mathcal{A}_t \mathbf{1}_n = \mathbf{1}_n$ and $\mathcal{A}_t^T \mathbf{1}_n = \mathbf{1}_n$ hold simultaneously. The following assumption is common in distributed optimization [19], [27], [40].

Assumption 1: The digraph \mathcal{G}_t , for $t \in [T]$, satisfies:

- (a) *Lower Bounds:* There exists a positive constant $\gamma \in (0, 1)$ that lower bounds all nonzero weights.
- (b) *Column-Stochasticity:* \mathcal{A}_t is column-stochastic.
- (c) *Uniformly Jointly Strongly Connected:* The digraphs $\{\mathcal{G}_t\}_{t=1}^\infty$ are uniformly jointly strongly connected, i.e., there exists a positive integer B such that the union digraph $\mathcal{G}(\mathcal{V}, \cup_{l=0, \dots, B-1} \mathcal{E}_{t+l})$ is strongly connected.

B. Problem Formulation

Consider a network of n nodes, which aim to collaboratively solve the following constrained optimization problem over a time horizon T :

$$\underset{x_t \in \Omega}{\text{minimize}} \sum_{t=0}^T F_t(x_t), \quad F_t(x) \triangleq \sum_{i=1}^n f_{i,t}(x) \quad (1)$$

where $f_{i,t}(x)$ is the local cost function of node $i \in \mathcal{V}$ at round $t \in [T]$. We consider the *online* and *bandit feedback* scenario wherein each node is only allowed to query the values of its local cost function after it has made its decision at each round. Since the global function F_t is not accessible to any node, each node needs to interact with its neighbors and make sequential decisions based on what it “thinks” the decisions that the whole network would make.

The *dynamic regret* is commonly adopted in the literature as a performance metric for the online algorithms (see e.g., [39]), which is defined as follows for node $j \in \mathcal{V}$

$$\mathcal{R}_j^d(T) = \sum_{t=0}^T F_t(x_{j,t}) - \sum_{t=0}^T F_t(x_t^*), \quad (2)$$

where $x_t^* \in \arg \min_{x \in \Omega} F_t(x)$. The dynamic regret (2) measures the difference between the cost incurred by node j 's decisions $\{x_{j,t}\}_{t \in [T]}$ against that of a time-varying clairvoyant. It is different from that of [29], [40], where the decisions of nodes are not required to reach consensus since the global cost function at round t is defined as $F_t(X) = \sum_{i=1}^n f_{i,t}(x_i)$ with $X \triangleq \text{col}(x_1, \dots, x_n)$ and $X_t^* \triangleq \text{col}(x_{1,t}^*, \dots, x_{n,t}^*)$. In contrast to the *static regret* [10] with fixed benchmark $x^* = \arg \min_{x \in \Omega} \sum_{t=0}^T F_t(x)$, the dynamic regret (2) allows the best decision varies with time, and thus can encompass the former as a special case. However, node $i \in \mathcal{V}$ can not access to $f_{i,t}$ when making decision $x_{i,t}$, which implies that the online algorithm is not able to track x_t^* well if x_t^* deviates from the past significantly. Therefore, a regularity measure named “path variation” [39] needs to be introduced to reflect the changes of successive minimizers

$$\mathcal{C}_T = \sum_{t=0}^T \|x_{t+1}^* - x_t^*\|. \quad (3)$$

The main objective of this work is to design a zeroth-order distributed online optimization algorithm over \mathcal{G}_t such that the dynamic regret (2) is upper bounded sublinearly with respect to T . We impose the following standard assumption in (distributed) bandit online optimization [13], [29], [32].

Assumption 2: For $i \in \mathcal{V}$ and $t \in [T]$, it satisfies

- (a) $\Omega \subseteq \mathbb{R}^m$ is a non-empty, convex and closed set. Moreover, there exist $r, R > 0$ such that

$$r\mathbb{B}^m \subseteq \Omega \subseteq R\mathbb{B}^m, \quad (4)$$

and r is known a priori.

- (b) The subgradient of $f_{i,t}$ is bounded over Ω , i.e., there exists a constant $G > 0$ such that $\|g_{i,t}(x)\| \leq G$ holds for $x \in \Omega$, where $g_{i,t}(x) \in \partial f_{i,t}(x)$.

III. ALGORITHM DEVELOPMENT

To develop the zeroth-order algorithm, we first define a smoothed version of $f_{i,t}(x)$, $i \in \mathcal{V}$, as follows

$$\hat{f}_{i,t}(x) \triangleq \mathbb{E}_{\zeta \in \mathbb{B}^m} [f_{i,t}(x + \mu\zeta_{i,t})], \quad x \in (1 - \xi)\Omega, \quad (5)$$

where $\zeta_{i,t}$ is a vector selected uniformly at random from \mathbb{B}^m , $\xi \in (0, 1)$ is a shrinkage parameter, $\mu \in (0, r\xi]$, and $(1 - \xi)\Omega$ is shorthand for $\{(1 - \xi)x : x \in \Omega\}$. Under these settings, it can be seen that $x + \mu\zeta_{i,t} \in \Omega$ for any $x \in (1 - \xi)\Omega$ and $\zeta_{i,t} \in \mathbb{S}^m$. In this paper, we adopt the following local zeroth-order oracle to avoid explicit subgradient calculations

$$\hat{g}_{i,t}(x) = \frac{m}{\mu} (f_{i,t}(x + \mu\zeta_{i,t}) - f_{i,t}(x)) \zeta_{i,t}, \quad (6)$$

where $x \in (1 - \xi)\Omega$, $\zeta_{i,t} \in \mathbb{S}^m$. In fact, $\hat{g}_{i,t}(x)$ is an unbiased gradient estimator of $\hat{f}_{i,t}(x)$. Notably, Gaussian random variables are used to construct zeroth-order oracles in [28], [35], [41], which cannot be applied in our setting. The reason is that Gaussian random variables do not have finite support such that the perturbation $x + \mu\zeta_{i,t}$ may lie outside of Ω .

In our algorithm, each node $i \in \mathcal{V}$ maintains a vector $x_{i,t} \in \mathbb{R}^m$ as well as a scalar $\phi_{i,t} \in \mathbb{R}$ at per round $t \in \mathbb{N}$, which are initialized with $x_{i,0} \in (1 - \xi)\Omega$ and $\phi_{i,0} = 1$, respectively. At per round t , each node i sends its decision $x_{i,t}$ and the scalar $\phi_{i,t}$ to its out-neighbors on the basis of \mathcal{G}_t , and then performs the following updates:

$$\phi_{i,t+1} = \sum_{j=1}^n a_{ij,t} \phi_{j,t}, \quad (7)$$

$$x_{i,t+1} = \mathcal{P}_{(1-\xi)\Omega} \left(\sum_{j=1}^n b_{ij,t} x_{j,t} - \frac{1}{\phi_{i,t+1}} \alpha_t \hat{g}_{i,t}(x_{i,t}) \right) \quad (8)$$

where α_t is the step size that will be specified later, $b_{ij,t} \triangleq \frac{1}{\phi_{i,t+1}} a_{ij,t} \phi_{j,t}$ is the entry at i -th row and j -th column of matrix \mathcal{B}_t . Under Assumption 1(b), it is straightforward to verify that $\mathcal{B}_t, \forall t \geq 0$, now becomes a row stochastic matrix by recalling (7) and the definition of $b_{ij,t}$. By this means, the term $\sum_{j=1}^n b_{ij,t} x_{j,t}$ in (8) is designed for consensus since our setting requires that each node can make sequential decisions to minimize the dynamic regret. The intuition of $\phi_{i,t+1}$ in (7) is that it is designed to rescale the zeroth-order oracle $\hat{g}_{i,t}(x_{i,t})$ in (8), which aims to address the unbalancedness issue of time-varying digraphs. In addition, the projection operation $\mathcal{P}_{(1-\xi)\Omega}(\cdot)$ is used to guarantee the feasibility of the quarry points, i.e., $x_{i,t} + \mu\zeta_{i,t} \in \Omega$ always holds for all $i \in \mathcal{V}$ and $t \in [T]$. Our algorithm is fully distributed that each node only requires the knowledge of itself as well as information from its immediate neighbors to carry out the above updates. The proposed algorithm is summarized in Algorithm 1.

Several features distinguish our algorithm from the relevant distributed (online) optimization methods over unbalanced digraphs. We summarize their connections as follows:

- (a) The \mathcal{AB} /push-pull algorithms [25], [26] and our algorithm all utilize row stochastic and column stochastic weight matrices simultaneously. Different from [25], [26] that the weight matrices are constructed independently, our algorithm only invokes a sequence of column

Algorithm 1 Zeroth-order Distributed Online OptimizationAlgorithm —from the view of node i

-
- 1: **Initialize:** Set $x_{i,0} \in (1 - \xi)\Omega$ and $\phi_{i,0} = 1$, $i \in \mathcal{V}$.
 - 2: **for** $t = 0, 1, 2, \dots, T$ **do**
 - 3: Broadcast $x_{i,t}$ and $\phi_{i,t}$ to nodes $j \in \mathcal{N}_{i,t}^-$.
 - 4: Calculate $\hat{g}_{i,t}(x_{i,t})$ via (6).
 - 5: Update auxiliary variable $\phi_{i,t}$ via (7).
 - 6: Update node's decision $x_{i,t}$ via (8).
 - 7: **end for**
-

stochastic matrices \mathcal{A}_t , and we obtain \mathcal{B}_t by performing the designated transformation on \mathcal{A}_t .

- (b) The push-sum based distributed online optimization algorithm in [27] will be invalid when it comes to problems with constraint set. The reason is that the balance properties guaranteed by taking specific ratios will be violated if we incorporate the projection operation into push-sum algorithms [19], [20], [38] directly. In contrast, our algorithm can deal with set constraint and be applicable to scenarios where the explicit subgradients of the cost functions are not available.
- (c) The static distributed algorithm in [22] also adjusts the gradient by a designed vector, which is essentially the estimate of left eigenvector of the weight matrix. However, it is not clear how to extend the algorithm to time-varying digraphs, since it is unrealistic to estimate the time-varying Perron vectors in a distributed manner. Our algorithm can be applicable to time-varying digraphs. Furthermore, each node i transmits the scalar $\phi_{i,t}$ at t instead of a vector, which considerably lightens the communication burden when compared with [22].

IV. CONVERGENCE ANALYSIS

In this section, we provide the theoretical analysis for the proposed algorithm. The main challenge in the convergence analysis lies in the combined effects of the projection operators, the time-varying unbalanced networks, and the embedded randomized zeroth-order oracles. In addition, the considered dynamic regret makes our analysis more challenging than that of the static regret scenario, since the cost functions and the minimizers are allowed to drift over time simultaneously in our setting. We begin with a few preliminary results, followed by the detailed analysis on dynamic regret of the proposed algorithm.

A. Preliminary results

Let \mathcal{F}_t denote the σ -field generated by the entire history of the random variables up to round t , i.e., $\mathcal{F}_t \triangleq \sigma(\zeta_{1,0}, \dots, \zeta_{n,0}, \dots, \zeta_{1,t}, \dots, \zeta_{n,t})$. The following Lemma collected from [29] outlines some important properties of $\hat{f}_{i,t}(x)$ and $\hat{g}_{i,t}(x)$.

Lemma 1: [29] Suppose Assumption 2 holds. For $i \in \mathcal{V}$, and $t \geq 0$, the following properties hold:

- (a) $\hat{f}_{i,t}(x)$ is convex and G -Lipschitz continuous on $(1 - \xi)\Omega$. Moreover,

$$f_{i,t}(x) \leq \hat{f}_{i,t}(x) \leq f_{i,t}(x) + \mu_i G. \quad (9)$$

for any $x \in (1 - \xi)\Omega$.

- (b) $\hat{f}_{i,t}(x)$ is differentiable on $(1 - \xi)\Omega$ even if $f_{i,t}(x)$ is not, and it satisfies

$$\nabla \hat{f}_{i,t}(x) = \mathbb{E}_{\zeta_{i,t} \in \mathbb{S}^m} [\hat{g}_{i,t}(x)], \quad x \in (1 - \xi)\Omega. \quad (10)$$

Moreover, $\nabla \hat{f}_{i,t}(x)$ is Lipschitz continuous on $(1 - \xi)\Omega$ with $L_i \triangleq mG/\mu_i$.

- (c) $\|\hat{g}_{i,t}(x)\| \leq mG$ holds for any $x \in (1 - \xi)\Omega$.

Recalling that \mathcal{B}_t is a row stochastic matrix for any $t \geq 0$, and it thus enjoys the following properties.

Lemma 2: [23] Suppose Assumptions 1 holds. For $s \geq t$, define $\mathcal{B}(s : t) := \mathcal{B}_{s-1} \cdots \mathcal{B}_t$ with the convention $\mathcal{B}(t : t) = I$ and $b_{ij}(s : t)$ being the entries of $\mathcal{B}(s : t)$. Then, there is a sequence of normalized vectors $\{\pi_t\}_{t \geq 0}$ with $\pi_t = [\pi_{1,t}, \dots, \pi_{n,t}]^T$ and $\mathbf{1}_n^T \pi_t = 1$, such that

- (a) There exists constants $C > 0$ and $\lambda \in (0, 1)$, such that $|b_{ij}(s : t) - \pi_{j,t}| \leq C\lambda^{s-t}$ holds for $i, j \in \mathcal{V}$;
- (b) There exists a constant $\beta \geq b_{\min}^{(n-1)B}$ such that $\pi_{i,t} \geq \beta$ for $i \in \mathcal{V}$ and $t \geq 0$, where $b_{\min} = \min_{i,j \in \mathcal{V}, t \geq 0} \{b_{ij,t} | j \in \mathcal{N}_{i,t}^+ \cup \{i\}\}$;
- (c) $\pi_t^T = \pi_{t+1}^T \mathcal{B}_t$.

The following lemma is collected from [20], which reveals that the scalar $\phi_{i,t}$ is bounded for $i \in \mathcal{V}$, $t \geq 0$.

Lemma 3: [20] Suppose Assumption 1 holds. Let $\{\phi_{i,t}\}_{t \geq 0}$, $\forall i \in \mathcal{V}$, be the sequence generated by (7). Then, there exists constants $\theta > 0$ and $\varpi > 0$ such that

$$\theta^{-1} \leq \phi_{i,t} \leq \varpi. \quad (11)$$

Remark 1: In fact, θ and ϖ in Lemma 3 can be precisely represented by $\gamma^{-2(n-1)B}$ and $n - \gamma^{2(n-1)B}$, respectively, where γ and B are constants defined in Assumption 1.

Furthermore, the following lemma establishes a relationship between $\phi_{i,t+1}$ and $\pi_{i,t+1}$, which plays a critical role in the dynamic regret analysis in the sequel.

Lemma 4: Let $\{\phi_{i,t+1}\}_{t \geq 0}$, $\forall i \in \mathcal{V}$, be the sequence generated by (7), and $\pi_{i,t+1}$ be the i -th component of π_{t+1} clarified in Lemma 2. Then, for a given time horizon T , the following inequality holds

$$\left| \frac{1}{n} \phi_{i,t+1} - \pi_{i,t+1} \right| \leq C\lambda^{T-t-1}. \quad (12)$$

Proof: See Appendix A. ■

B. Dynamic Regret Analysis

As detailed in the previous sections, the combination of projection operation and time-varying unbalanced networks brings new challenges in the dynamic regret analysis compared to the existing works. Particularly, we cannot concentrate on the evolution of the average process $\frac{1}{n} \sum_{i=1}^n x_{i,t}$ as many distributed algorithms [19], [27], [39] did. To cope with this challenge, we construct time-varying row-stochastic matrices \mathcal{B}_t , which enables us to turn attention from the average process to an auxiliary vector defined as

$$\bar{x}_t \triangleq \sum_{i=1}^n \pi_{i,t} x_{i,t} \quad (13)$$

with $\pi_t = [\pi_{1,t}, \dots, \pi_{n,t}]$ defined in Lemma 2. The following lemma establishes an upper bound on the expected disagreement between \bar{x}_t and $x_{i,t}$ for $i \in \mathcal{V}$.

Lemma 5: Suppose Assumptions 1-2 hold. Let $\{x_{i,t}\}_{t \in [T]}$ be the sequence generated by Algorithm 1. Then, for $i \in \mathcal{V}$, $t \in [T]$,

$$\begin{aligned} & \mathbb{E} [\|x_{i,t+1} - \bar{x}_{t+1}\|] \\ & \leq nm\theta CG \sum_{l=1}^{t+1} \lambda^{t+1-l} \alpha_{l-1} + C\lambda^{t+1} \sum_{j=1}^n \|x_{j,0}\|. \end{aligned} \quad (14)$$

Proof: See Appendix B. ■

Remark 2: Note that the first term on the right side of (14) can be calculated as follows

$$\begin{aligned} & \sum_{l=1}^{t+1} \lambda^{t+1-l} \alpha_{l-1} = \sum_{l=1}^{\lfloor t/2 \rfloor} \lambda^{t-l+1} \alpha_{l-1} + \sum_{l=\lfloor t/2 \rfloor + 1}^{t+1} \lambda^{t-l+1} \alpha_{l-1} \\ & = \lambda^{\lfloor t/2 \rfloor} \sum_{l=1}^{\lfloor t/2 \rfloor} \lambda^{\lfloor t/2 \rfloor - l + 1} \alpha_{l-1} + \sum_{l=\lfloor t/2 \rfloor + 1}^{t+1} \lambda^{t-l+1} \alpha_{l-1} \\ & \leq \xi^{\lfloor t/2 \rfloor} \frac{\lambda}{1-\lambda} \sup_{t \geq 0} \alpha_t + \frac{\lambda}{1-\lambda} \sup_{l \geq \lfloor t/2 \rfloor} \alpha_l, \end{aligned}$$

where $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ denote the floor function and ceiling function, respectively. If the step size α_t tends to zero, e.g., $\alpha_t = \mathcal{O}(\frac{1}{\sqrt{t+1}})$, then the two terms in the above inequality both tend to zero since $\xi^{\lfloor t/2 \rfloor} \rightarrow 0$ and $\sup_{l \geq \lfloor t/2 \rfloor} \alpha_l \rightarrow 0$ when $t \rightarrow \infty$. Thus we can obtain $\lim_{t \rightarrow \infty} \mathbb{E} [\|x_{i,t+1} - x_{j,t+1}\|] = 0$, which implies that all nodes can achieve consensus in expectation when $t \rightarrow \infty$.

To obtain the desired regret bound, it suffices to derive an upper bound on the expected optimality gap. However, x_t^* may lie outside of $(1-\xi)\Omega$, which bring new challenge to our analysis. To resolve this issue, we first establish a relation on $\mathbb{E}[f_{i,t}(\bar{x}_t) - f_{i,t}((1-\xi)x_t^*)]$, and then the desired result can be obtained by bounding the difference between $f_{i,t}((1-\xi)x_t^*) - f_{i,t}(x_t^*)$. The following result provides a bound on the weighted sum of the terms $\mathbb{E}[f_{i,t}(\bar{x}_t) - f_{i,t}(x_t^*)]$, which is pivotal to our dynamic regret analysis.

Theorem 1: Suppose Assumptions 1-2 hold. Let $\{x_{i,t}\}_{t \in [T]}$ be the sequence generated by Algorithm 1 with positive and non-increasing step size α_t . Then,

$$\begin{aligned} & \sum_{t=0}^T \sum_{i=1}^n \frac{\pi_{i,t+1}}{\phi_{i,t+1}} \mathbb{E} [f_{i,t}(\bar{x}_t) - f_{i,t}(x_t^*)] \\ & \leq 2nm\theta^2 RLCG \sum_{t=0}^T \sum_{l=1}^t \lambda^{t-l} \alpha_{l-1} + \frac{m^2\theta^2 G^2}{2} \sum_{t=0}^T \alpha_t \\ & \quad + 2nm^2\theta^2 CG^2 \sum_{t=0}^T \sum_{l=1}^{t+1} \lambda^{t+1-l} \alpha_{l-1} + \frac{2R^2}{\alpha_T} \\ & \quad + 2(RL + mG\lambda)\theta C \sum_{j=1}^n \|x_{j,0}\| \sum_{t=0}^T \lambda^t + \frac{2R}{\alpha_T} \mathcal{C}_T \\ & \quad + \mu\theta G(T+1) + \theta\xi GR(T+1). \end{aligned} \quad (15)$$

Proof: See Appendix C. ■

Albeit Theorem 1 has established a relation on the term $\mathbb{E}[f_{i,t}(\bar{x}_t) - f_{i,t}(x_t^*)]$, it should be noted here that this term is scaled by $\frac{\pi_{i,t+1}}{\phi_{i,t+1}}$ and we cannot determine the sign of the difference (i.e., $f_{i,t}(\bar{x}_t) - f_{i,t}(x_t^*)$) for particular $i \in \mathcal{V}$ since x_t^* is the minimizer of the global cost function rather than each local cost function, which bring new challenge in deriving the dynamic regret. Fortunately, we will soon find that the scalar $\frac{1}{n}\phi_{i,t+1}$ plays an important role in counteracting the impact of unbalancedness, and hence we can establish an upper bound of the dynamic regret (2) in the following theorem.

Theorem 2 (Dynamic Regret Bound): Suppose Assumptions 1-2 hold. For a given time horizon T , set $\mu = \frac{r}{\sqrt{T+1}}$ and $\xi = \frac{1}{\sqrt{T+1}}$. Let $\{x_{i,t}\}_{t \in [T]}$ be the sequence generated by Algorithm 1 with $\alpha_t = \frac{1}{m\sqrt{t+1}}$. Then,

$$\mathcal{R}_j^d(T) \leq \mathcal{K}_1 + \mathcal{K}_2 \sqrt{T+1} \quad (16)$$

where

$$\begin{aligned} \mathcal{K}_1 & = \frac{nC(2RL\theta + 2m\theta\lambda G + G)}{1-\lambda} \sum_{i=1}^n \|x_{i,0}\| + \frac{2n^2\theta CGR}{\lambda(1-\lambda)}, \\ \mathcal{K}_2 & = \frac{2n^2\theta CG}{1-\lambda} (2\theta RL + 2m\theta G + G) + nm\theta^2 G^2 \\ & \quad + 2nmR^2 + nr\theta G + n\theta GR + 2nmRC_T. \end{aligned} \quad (17)$$

Proof: See Appendix D. ■

Theorem 2 shows that if \mathcal{C}_T is bounded, then the average dynamic regret over T asymptotically converges zero at the rate of $\mathcal{O}(T^{-\frac{1}{2}})$, which matches the best static regret that can be achieved in literature for convex cost functions [10]–[13], [27], [42]. In fact, \mathcal{C}_T is not necessary to be bounded. If \mathcal{C}_T increases sublinearly with the rate ranged from zero to $\mathcal{O}(T^{\frac{1}{2}})$, then we have $\lim_{T \rightarrow \infty} \frac{\mathcal{C}_T \sqrt{T+1}}{T} = 0$, which implies that the average dynamic regret over T asymptotically converges to zero at a sublinear rate.

Remark 3: Zeroth order optimization algorithms usually suffer from deterioration in convergence as the problem dimension increases, which is a typical limitation of these algorithms in high dimension problems. From Lemma 1(c), it is readily seen that the upper bound of $\|\hat{g}_{i,t}(x)\|$ depends on the dimension m , which is essentially the penalty incurred by the use of zeroth-order oracle instead of the real subgradient. The dimension dependency of our algorithm is $\mathcal{O}(m)$, which is identical to that of [43], and better than $\mathcal{O}(m^2)$ in [11], [12], [28], [31], [32], [41]. The optimal dimension dependency $\mathcal{O}(\sqrt{m})$ is obtained in [33]. However, the algorithm in [33] is centralized, and the author only considered static regret rather than dynamic regret.

C. Discussion

In this section, we discuss the counterpart of Algorithm 1 when subgradients of local cost functions are available. In this scenario, a new algorithm can be directly obtained by utilizing the subgradient in lieu of the local randomized zeroth-

Algorithm 2 Distributed Online Optimization Algorithm —
 from the view of node i

- 1: **Input:** time horizon T ; for all $i \in \mathcal{V}$, set $x_{i,0} \in \Omega$ and $\phi_{i,0} = 1$.
 - 2: **for** $t = 0, 1, 2, \dots, T$ **do**
 - 3: Broadcast $x_{i,t}$ and $\phi_{i,t}$ to nodes $j \in \mathcal{N}_{i,t}^-$.
 - 4: Update auxiliary variable $\phi_{i,t}$ via (18).
 - 5: Update node's decision $x_{i,t}$ via (19).
 - 6: **end for**
 - 7: **Output:** $\{x_{i,t}\}_{t=1}^n$.
-

order oracle $\hat{g}_{i,t}(x_{i,t})$ in (8). Then, each node $i \in \mathcal{V}$ performs the following updates¹:

$$\phi_{i,t+1} = \sum_{j=1}^n a_{ij,t} \phi_{j,t}, \quad (18)$$

$$x_{i,t+1} = \mathcal{P}_\Omega \left(\sum_{j=1}^n b_{ij,t} x_{j,t} - \frac{1}{\phi_{i,t+1}} \alpha_t g_{i,t}(x_{i,t}) \right), \quad (19)$$

where $g_{i,t}(x_{i,t})$ denotes the subgradient of $f_{i,t}(x)$ evaluated at $x_{i,t}$. This algorithm is summarized in Algorithm 2.

Accordingly, the determined form of the dynamic regret (2) can be represented by $\tilde{\mathcal{R}}_j^d(T)$ as follows

$$\tilde{\mathcal{R}}_j^d(T) = \sum_{t=0}^T f_t(x_{j,t}) - \sum_{t=0}^T f_t(x_t^*), \quad (20)$$

Now we provide an upper bound of the dynamic regret (20) under Algorithm 2. The following theorem reveals that the average dynamic regret over time horizon T asymptotically converges to zero at a sublinear rate for convex cost functions provided that the accumulated variation grows sublinearly with the rate ranged from zero to $\mathcal{O}(T^{\frac{1}{2}})$, which reflects that the convergence of Algorithm 1 is essentially not affected by the incorporated zeroth-order oracle.

Theorem 3: Suppose Assumptions 1-2 hold. Let $\{x_{i,t}\}_{t \geq 0}$ be the sequence generated by Algorithm 2 with $\alpha_t = \frac{1}{\sqrt{t+1}}$. Then, for a given time horizon T , the dynamic regret defined in (20) satisfies

$$\tilde{\mathcal{R}}_j^d(T) \leq \mathcal{K}_3 + \mathcal{K}_4 \sqrt{T}, \quad (21)$$

where

$$\mathcal{K}_3 = \frac{(2\theta\lambda + \theta + 2)nCG}{1 - \lambda} \sum_{i=1}^n \|x_{i,0}\| + \frac{2n^2\theta CGR}{\lambda(1 - \lambda)},$$

$$\mathcal{K}_4 = \frac{2(3\theta + 2)n^2\theta CG^2}{1 - \lambda} + n\theta^2 G^2 + 2nRC_T + 2nR^2.$$

Proof: The proof follows the similar line of Theorem 1-2, we thus omit it. ■

¹In this paper, the notation of $x_{i,t}$ and $\phi_{i,t}$, will be kept the same across different algorithms to avoid notational clutter, and it is clear from context which method is in question.

V. NUMERICAL EXAMPLES

In this section, we numerically demonstrate the performance of the proposed algorithms. First, motivated by [39], we validate our theoretical findings by a numerical example. After that, we apply the proposed algorithms to a distributed target tracking problem, which has been widely investigated in literature, e.g., [16], [44]. Finally, we investigate the dynamic sparse signal recovery problem [17] and compare our algorithm with the existing ones.

A. Example 1

Consider a group of six agents modeled by the time-varying digraphs \mathcal{G}_t depicted in Fig. 1. The dynamic local cost function at time t is given by $f_{i,t}(\mathbf{x}) = \frac{i}{84}x_1^4 + \frac{i-1}{15}(x_1^2 + x_2^2) + \frac{2i+1}{4}x_1 - \frac{2(i-3)}{3}h(t)x_2$, $i \in \mathcal{V}$ and $\mathbf{x} = (x_1, x_2)^T$ is constrained in a box set $\Omega = \{-3 \leq x_1 \leq 2, 0 \leq x_2 \leq 3\}$. Thus, the global cost function at time t can be described as $f_t(\mathbf{x}) = \frac{1}{4}x_1^4 + x_1^2 + x_2^2 + 12x_1 - 2h(t)x_2$ with $\mathbf{x} \in \Omega$. We let $h(t) = \arctan(t/10)$. It is easy to verify that $\mathbf{x}_t^* = (-2, \arctan(t/10))$, $t \geq 0$. In our simulations, we set $\mu = 10^{-3}$ and $\xi = 0.02$. The graph \mathcal{G}_t changes by the order $\mathcal{G}_1 \rightarrow \mathcal{G}_2 \rightarrow \mathcal{G}_3 \rightarrow \dots$ throughout the whole process. For the weight matrix \mathcal{A}_t , we set $a_{ij,t} = 1/|\mathcal{N}_{j,t}^-|$ if $a_{ij,t} > 0$, and $a_{ii,t} = 1 - \sum_{j=1}^n a_{ji,t}$, where $|\mathcal{N}_{j,t}^-|$ denotes the number of elements in $\mathcal{N}_{j,t}^-$. Under these settings, all assumptions in this paper evidently hold. We run Algorithm 1 with $\alpha_t = \frac{1}{2\sqrt{t+1}}$. The trajectories of the optimal solution \mathbf{x}_t^* and the decision variables $x_{i,t}$, $i \in \mathcal{V}$, are shown in Fig. 2 with $x_{ij,t}$ being the j -th entry of the decision variable made by agent i at round t . It can be seen that all agents's decision variables approach the optimal solution \mathbf{x}_t^* .

Moreover, we compare the convergence performance between Algorithm 1 and Algorithm 2 under the same settings, except that Algorithm 1 adopts the local zeroth-order oracle (6) with $\mu = 10^{-3}$ and Algorithm 2 uses the explicit gradient of local cost functions. The maximum and minimum average dynamic regrets over time horizon T , defined as i.e., $\max_{j \in \mathcal{V}} \mathcal{R}_j(T)/T$ and $\min_{j \in \mathcal{V}} \mathcal{R}_j(T)/T$, are depicted in Fig. 3. The result is consistent with our theoretical results established in Theorem 2 and Theorem 3, and further verifies the effectiveness of our zeroth-order algorithm since the Algorithm 1 can achieve comparable performance with Algorithm 2 where explicit gradient information is used.

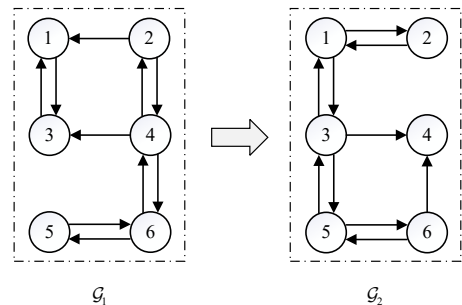


Fig. 1. The time-varying unbalanced digraphs.

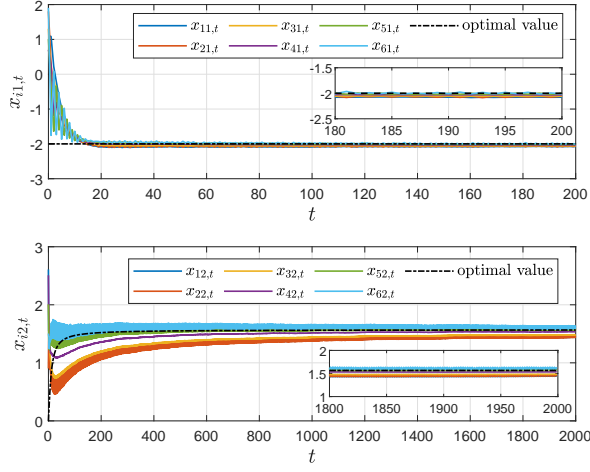


Fig. 2. Trajectories of the optimal solution x_t^* and the decisions $x_{i,t}$, $i \in \mathcal{V}$, generated by Algorithm 1.

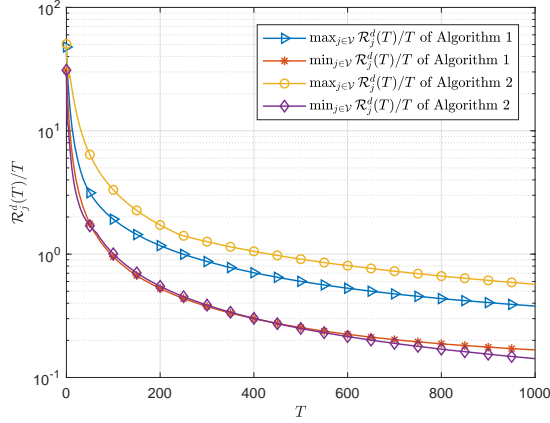


Fig. 3. Comparisons of the maximum and minimum dynamic regret over time horizon T between Algorithm 1 and Algorithm 2.

B. Example II: Targets Tracking

In this example, we consider a tracking problem where sensors collaborate to track moving targets. It is of practical significance to study distributed online algorithms over unbalanced networks as communications between sensors are usually directed. Moreover, dynamic regret is qualified to measure the tracking performance in contrast to static regret.

We consider a network consisting of $n = 6$ nodes, and the information sharing among the nodes is depicted in Fig. 1. All nodes aim to collaboratively track three time-varying signals via local communication. Specifically, each signal $x_{i,t}^* = [x_{i1,t}^*, x_{i2,t}^*]^T$, $i \in \{1, 2, 3\}$ is described as follows

$$x_{i,t}^* = \begin{bmatrix} x_{i1,t}^* \\ \dot{x}_{i1,t}^* \end{bmatrix} = \begin{bmatrix} \kappa_i \sin(\omega_i t + \nu_i) \\ \omega_i \kappa_i \cos(\omega_i t + \nu_i) \end{bmatrix} \quad (22)$$

where $x_{i1,t}^*$ is the position of target i , $\dot{x}_{i1,t}^*$ denotes the velocity of target i at time t , κ_i is the amplitude, ω_i represents the angular frequency, and ν_i is the phase of the target i . At time step t , each node $i \in \mathcal{V}$ observes $y_{i,t}$ via the measurement

model $y_{i,t} = C_i x_t^*$, where $x_t^* = [(x_{1,t}^*)^T, (x_{2,t}^*)^T, (x_{3,t}^*)^T]^T$ and $C_i \in \mathbb{R}^{1 \times 6}$ is the measurement matrix that is generated randomly. To track the moving targets, each node communicates with its neighbors aiming to minimize the global cost function $f_t(x) = \frac{1}{2} \sum_{i=1}^6 \|C_i x - y_{i,t}\|^2$.

In our simulations, the amplitude κ_i and the phase ν_i , for $i \in \mathcal{V}$, are random variables drawn from uniform distribution in $[0, 3]$ and $[0, \pi]$, respectively. We let the sampling frequency be 100Hz. Similar to the previous example, we verify the performance of Algorithm 1 with $\alpha_t = \frac{1}{2\sqrt{t+1}}$ and $\mu = 10^{-3}$. As shown in Fig. 4, the average dynamic regret over T calculated via any sequence of $\{x_{i,t}\}_{t=0}^T$, $i \in \mathcal{V}$, asymptotically converges at a sublinear rate, which is consistent with our results established in Theorem 2. Furthermore, we show the trajectories of the targets $x_{i1,t}^*$, $i \in \{1, 2, 3\}$, and the decisions $x_{3j,t}$ and $x_{4j,t}$, $j \in \{1, 3, 5\}$, in Fig. 5 with $x_{ij,t}$ being the j -th entry of the decision made by agent i at t . It can be observed that node 3 and node 4 both can track the three moving targets within the targets' small neighborhood. In fact, the tracking performance can be further improved by tuning the sampling frequency. However, utilizing the diminishing step-size prevents our algorithms from tracking quickly moving targets. It is thus of interest to develop distributed online optimization algorithms in time-varying unbalanced networks that admits more aggressive step sizes to obtain better tracking performance, which remains to be considered in our future work.

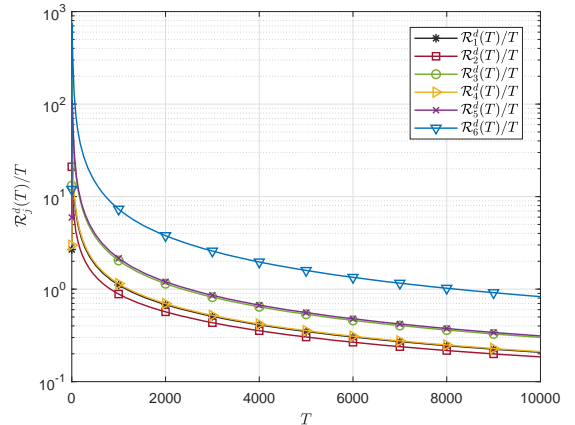


Fig. 4. Trajectories of the average dynamic regret over time horizon T of Algorithm 1.

C. Example III: The Dynamic Sparse Recovery Problem

In this example, we consider the dynamic sparse signal recovery problem with the goal of estimating a time-varying sparse parameter in a distributed manner. This problem has been widely investigated in signal processing literature. Specifically, we compare our algorithm with the ones in [16], [39], [44]. Note that all of the algorithms in these works can be only applicable to (fixed or time-varying) balanced networks. To investigate these algorithms within a unified framework, we consider a network with n sensors, which can

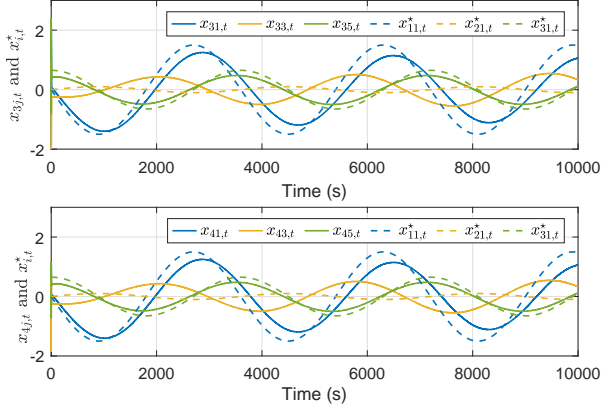


Fig. 5. Trajectories of the targets $x_{i1,t}^*$, $i \in \{1, 2, 3\}$, and the decision variables $x_{3j,t}$ and $x_{4j,t}$, $j \in \{1, 3, 5\}$, generated by Algorithm 1.

be represented by a complete graph. All sensors collectively solve the following optimization problem:

$$x_t^* = \arg \min_x f_t(x) \quad (23)$$

with

$$f_t(x) = \frac{1}{n} \sum_{i=1}^n \left(\|z_{i,t} - C_{i,t}x\|_2^2 + \gamma \|x\|_2^2 + \sigma \|x\|_1 \right), \quad (24)$$

where $x \in \mathbb{R}^m$, $C_{i,t} \in \mathbb{R}^{d \times m}$ is the observation matrix of sensor i at round t , γ and σ are constant regularization parameters used to prevent overfitting, and $z_{i,t} \in \mathbb{R}^d$ is the measurement given by

$$z_{i,t} = C_{i,t}\varpi_t + \vartheta_{i,t} \quad (25)$$

with $\varpi_t \in \mathbb{R}^m$ being the time-varying sparse signal of interest and $\vartheta_{i,t} \in \mathbb{R}^d$ being the noise. In this example, we set $n = 40$, $d = 3$, $m = 8$, $\gamma = \frac{1}{100d^2n}$ and $\sigma = \frac{1}{20d}$. The initial state of ϖ_t is chosen to be a sparse vector with 2 entries of value 1, and all other entries equal to zero. Let $\mathcal{S}_t = \{\kappa | [\varpi_t]_\kappa > 0\}$ be the support of ϖ_t , which is updated as follows:

$$\mathcal{S}_{t+1} = \begin{cases} \mathcal{S}_t, & \text{with probability } 1 - 1/t \\ \{\mathcal{S}_t \setminus \{\tau_t\}\} \cup \{\tau'_t\}, & \text{with probability } 1/t \end{cases} \quad (26)$$

where τ_t and τ'_t are randomly chosen from \mathcal{S}_t and the set $\{1, 2, \dots, m\} \setminus \mathcal{S}_t$, respectively. We add noise ς_t to ϖ_t and then normalize the obtained vector so that

$$\varpi_{t+1} = \frac{\varpi_t + \varsigma_t}{\|\varpi_t + \varsigma_t\|} \quad (27)$$

where $[\varsigma_t]_\kappa \sim \mathcal{N}(0, 1/t^2)$ for $\kappa \in \mathcal{S}_{t+1}$ and $[\varsigma_t]_\kappa = 0$ otherwise. By this means, the non-zero entries of ϖ_t are time-varying and their variations decay over time. The related parameters are properly selected such that the minimizer sequence $\{x_t^*\}_{t=1}^T$ varies slowly. We verify the performance of Algorithm 1 with $\alpha_t = \frac{1}{\sqrt{t+1}}$ and $\mu = 10^{-3}$. We run Algorithm 1 for 100 trials. As shown in Fig. 6, the average maximum and minimum dynamic regrets over T both converge sublinearly, which are consistent with our theoretical results.

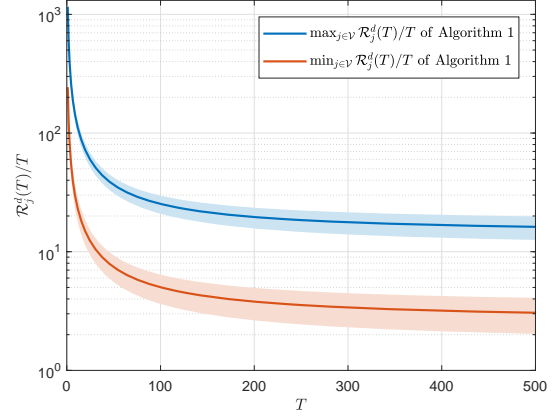


Fig. 6. Trajectories of the maximum and minimum dynamic regret over time horizon T of Algorithm 1.

We further compare our algorithm against the ones in [16], [39], [44]. For fair comparison, we adopt the average of $\frac{\mathcal{R}_j^d(T)}{T}$ over all nodes $j \in \mathcal{V}$ as the performance metric of the whole network. The results are depicted in Fig. 7, which reflect that the proposed Algorithm 1 can achieve comparable performance with the existing algorithms despite the presence of gradient estimate errors. Note that our algorithm can be applied to a broader family of networks since it does not require networks to be fixed or balanced in contrast to [16], [39], [44].

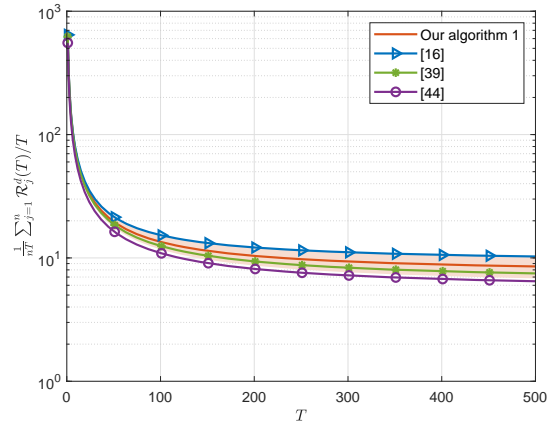


Fig. 7. Comparison of the time-average dynamic regret averaged over all nodes between our algorithms and the existing ones.

Finally, considering the fact that zeroth-order optimization algorithms usually suffer from deterioration in performance as the problem dimension increases, we investigate the influence of the problem dimension on the average dynamic regret of our algorithm. We run our algorithm for three different choices of the problem dimension m , i.e., $m = 6$, $m = 8$ and $m = 10$, respectively. The results are depicted in Fig. 8. It can be seen that the proposed zeroth-order optimization algorithm indeed achieves better performance with smaller problem dimension.

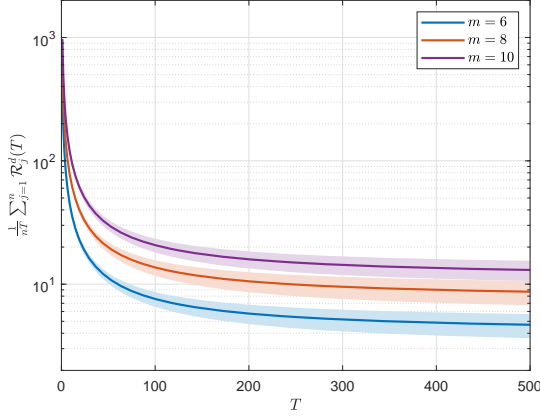


Fig. 8. The influence of the problem dimension m on the performance of Algorithm 1.

VI. CONCLUSIONS

In this paper, we have investigated the distributed online constrained optimization problems over time-varying unbalanced digraphs without explicit subgradients. To cope with the problem, a novel consensus-based distributed online optimization algorithm with a local randomized zeroth-order oracle has been proposed. The dynamic regret bound of the algorithm has been characterized under mild conditions, which showed that the algorithm can achieve comparable performance with its counterpart subgradient-based algorithm. Finally, simulations on distributed target tracking problem and dynamic sparse signal recovery problem in sensor networks have been conducted to verify the effectiveness of proposed algorithms. Future works can focus on deriving a tighter dynamic regret bound, and considering the scenario where inequalities constraints are involved in the problem. It is also of interest to develop dimension-insensitive zeroth-order distributed online optimization algorithms under mild assumptions.

APPENDIX A

Proof of Lemma 4. Recalling $\phi_{i,t+1} = \sum_{j=1}^n a_{ij,t} \phi_{j,t}$ and $b_{ij,t} = \frac{1}{\phi_{i,t+1}} a_{ij,t} \phi_{j,t}$, along with the fact that \mathcal{A}_t is a column matrix for $t \geq 0$, it is not hard to verify that $\phi_{t+1}^T \mathcal{B}_t = \phi_t^T$. Then, conducting the mathematical induction for ϕ_t yields that

$$\phi_{t+1}^T = \phi_t^T \mathcal{B}(T : t + 1) \quad (28)$$

On the other hand, by exploiting the column stochasticity of \mathcal{A}_t , we have

$$\mathbf{1}_n^T \phi_{t+1} = \mathbf{1}_n^T \mathcal{A}_t \phi_t = \dots = \sum_{i=1}^n \phi_{i,0} = n \quad (29)$$

Therefore, we can further obtain that

$$\begin{aligned} & \left| \frac{1}{n} \phi_{i,t+1} - \pi_{i,t+1} \right| \\ &= \left| \frac{1}{n} \sum_{l=1}^n \phi_{l,T} b_{li}(T : t + 1) - \frac{\pi_{i,t+1}}{n} \sum_{l=1}^n \phi_{l,T} \right| \end{aligned}$$

$$\begin{aligned} & \leq \frac{1}{n} \sum_{l=1}^n \phi_{l,T} |b_{li}(T : t + 1) - \pi_{i,t+1}| \\ & \leq C \lambda^{T-t-1} \end{aligned} \quad (30)$$

where Lemma 2(a) has been utilized to obtain the last inequality. The proof is thus completed. \blacksquare

APPENDIX B

Proof of Lemma 5. Define

$$v_{i,t} \triangleq \sum_{j=1}^n b_{ij,t} x_{j,t}$$

$$w_{i,t} \triangleq \mathcal{P}_{(1-\xi)\Omega} \left(v_{i,t} - \frac{1}{\phi_{i,t+1}} \alpha_t \hat{g}_{i,t}(x_{i,t}) \right) - v_{i,t}.$$

Then (8) can be rewritten as the following perturbed form:

$$x_{i,t+1} = \sum_{j=1}^n b_{ij,t} x_{j,t} + w_{i,t}. \quad (31)$$

Conducting the mathematical induction for (31) yields that

$$\begin{aligned} x_{i,t+1} &= \sum_{j=1}^n [\mathcal{B}(t+1 : 0)]_{ij} x_{j,0} \\ &+ \sum_{l=1}^{t+1} \sum_{j=1}^n [\mathcal{B}(t+1 : l)]_{ij} w_{j,l-1}. \end{aligned} \quad (32)$$

On the other hand, multiplying both sides of (31) by $\pi_{i,t+1}$ and then summing the obtained equality over $i \in \mathcal{V}$ yields that

$$\bar{x}_{t+1} = \bar{x}_t + \sum_{i=1}^n \pi_{i,t+1} w_{i,t}, \quad (33)$$

where Lemma 2(c) has been exploited to obtain this equality. Performing summations on (33) over $t \in \mathbb{N}$ leads to

$$\bar{x}_{t+1} = \bar{x}_0 + \sum_{l=1}^{t+1} \sum_{i=1}^n \pi_{i,l} w_{i,l-1}. \quad (34)$$

Then, combining (32) and (34) along with Lemma 2(a), it gives that

$$\begin{aligned} & \|x_{i,t+1} - \bar{x}_{t+1}\| \\ &= \sum_{j=1}^n \|[\mathcal{B}(t+1 : 0)]_{ij} - \pi_{j,0}\| \|x_{j,0}\| \\ &+ \sum_{l=1}^{t+1} \sum_{j=1}^n \|[\mathcal{B}(t+1 : l)]_{ij} - \pi_{j,l}\| \|w_{j,l-1}\| \\ &\leq C \lambda^{t+1} \sum_{j=1}^n \|x_{j,0}\| + C \sum_{l=1}^{t+1} \lambda^{t+1-l} \sum_{j=1}^n \|w_{j,l-1}\| \end{aligned} \quad (35)$$

Now, it remains to bound $\|w_{i,t}\|$. Note that $v_{i,t} \in \Omega$ is a convex combination of $x_{i,t} \in \Omega$, it thus follows that

$$\begin{aligned} \|w_{i,t}\| &\leq \left\| v_{i,t} - \frac{\alpha_t \hat{g}_{i,t}(x_{i,t})}{\phi_{i,t+1}} - v_{i,t} \right\| \\ &\leq \theta \alpha_t \|\hat{g}_{i,t}(x_{i,t})\| \\ &\leq m \theta G \alpha_t \end{aligned} \quad (36)$$

where the non-expansive projection property (see e.g., [45]) has been exploited to obtain the first inequality, and the last two inequalities follows from Lemma 3 and Lemma 1, respectively. It is straightforward to obtain (14) by combining (35) and (36). ■

APPENDIX C

Proof of Theorem 1. Define $\tilde{x}_t^* \triangleq (1 - \xi)x_t^*$. Obviously, $\tilde{x}_t^* \in (1 - \xi)\Omega$, and $\bar{x}_t \in (1 - \xi)\Omega$. We further define an auxiliary function D_t as

$$D_t \triangleq \frac{1}{2}(\bar{x}_t - \tilde{x}_t^*)^\top (\bar{x}_t - \tilde{x}_t^*) \quad (37)$$

Where \bar{x}_t has been defined in (13). Then, we consider the variation of D_t as follows

$$\begin{aligned} \Delta D_t &\triangleq D_{t+1} - D_t \\ &= -\frac{1}{2}\|\bar{x}_{t+1} - \bar{x}_t\|^2 + \frac{1}{2}\langle \tilde{x}_{t+1}^* + \tilde{x}_t^* - 2\bar{x}_{t+1}, \\ &\quad \tilde{x}_{t+1}^* - \tilde{x}_t^* \rangle + \langle \bar{x}_{t+1} - \bar{x}_t, \bar{x}_{t+1} - \tilde{x}_t^* \rangle \\ &\leq -\frac{1}{2}\|\bar{x}_{t+1} - \bar{x}_t\|^2 + 2R\|\tilde{x}_{t+1}^* - \tilde{x}_t^*\| \\ &\quad + \langle \bar{x}_{t+1} - \bar{x}_t, \bar{x}_{t+1} - \tilde{x}_t^* \rangle \end{aligned} \quad (38)$$

where the inequality follows from the boundedness of the constraint set. According to (33), the last term in (38) can be further expanded as follows

$$\begin{aligned} &\langle \bar{x}_{t+1} - \bar{x}_t, \bar{x}_{t+1} - \tilde{x}_t^* \rangle \\ &= \left\langle \sum_{i=1}^n \pi_{i,t+1} w_{i,t}, \bar{x}_{t+1} - \tilde{x}_t^* \right\rangle \\ &= \sum_{i=1}^n \pi_{i,t+1} \langle w_{i,t}, \bar{x}_{t+1} - x_{i,t+1} \rangle \\ &\quad + \sum_{i=1}^n \pi_{i,t+1} \langle w_{i,t}, x_{i,t+1} - \tilde{x}_t^* \rangle \end{aligned} \quad (39)$$

For the first term on the right side of (39), we have

$$\begin{aligned} &\sum_{i=1}^n \pi_{i,t+1} \langle w_{i,t}, \bar{x}_{t+1} - x_{i,t+1} \rangle \\ &\leq \sum_{i=1}^n \pi_{i,t+1} \|w_{i,t}\| \|\bar{x}_{t+1} - x_{i,t+1}\| \end{aligned} \quad (40)$$

Recalling (35)-(36), and taking expectation on both sides of (40), we can obtain

$$\begin{aligned} &\mathbb{E} \left[\sum_{i=1}^n \pi_{i,t+1} \langle w_{i,t}, \bar{x}_{t+1} - x_{i,t+1} \rangle \right] \\ &\leq m\theta CG\alpha_t \lambda^{t+1} \sum_{j=1}^n \|x_{j,0}\| \\ &\quad + nm^2\theta^2 CG^2 \alpha_t \sum_{l=1}^{t+1} \lambda^{t+1-l} \alpha_{l-1} \end{aligned} \quad (41)$$

where the fact that $\sum_{i=1}^n \pi_{i,t+1} = 1$ has also been exploited to obtain the above inequality.

For the last term on the right side of (39), we have

$$\begin{aligned} &\sum_{i=1}^n \pi_{i,t+1} \langle w_{i,t}, x_{i,t+1} - \tilde{x}_t^* \rangle \\ &= \sum_{i=1}^n \pi_{i,t+1} \left\langle w_{i,t} + \frac{\alpha_t \hat{g}_{i,t}(x_{i,t})}{\phi_{i,t+1}}, x_{i,t+1} - \tilde{x}_t^* \right\rangle \\ &\quad + \sum_{i=1}^n \pi_{i,t+1} \left\langle \frac{\alpha_t \hat{g}_{i,t}(x_{i,t})}{\phi_{i,t+1}}, \tilde{x}_t^* - x_{i,t+1} \right\rangle \end{aligned} \quad (42)$$

From Lemma 1 in [45], we know $(\mathcal{P}_{(1-\xi)\Omega}(x) - x)^\top (\mathcal{P}_{(1-\xi)\Omega}(x) - y) \leq 0$, for all $x \in \mathbb{R}^m$ and $y \in (1 - \xi)\Omega$. Hence, it can be observed that

$$\left\langle w_{i,t} + \frac{\alpha_t \hat{g}_{i,t}(x_{i,t})}{\phi_{i,t+1}}, x_{i,t+1} - \tilde{x}_t^* \right\rangle \leq 0 \quad (43)$$

That is to say, the first term on the right side of (42) is nonpositive. Now, we turn to the last term on the right side of (42), which can be calculated as follows

$$\begin{aligned} &\sum_{i=1}^n \pi_{i,t+1} \left\langle \frac{\alpha_t \hat{g}_{i,t}(x_{i,t})}{\phi_{i,t+1}}, \tilde{x}_t^* - x_{i,t+1} \right\rangle \\ &= \sum_{i=1}^n \pi_{i,t+1} \left\langle \frac{\alpha_t \hat{g}_{i,t}(x_{i,t})}{\phi_{i,t+1}}, \tilde{x}_t^* - \bar{x}_t \right\rangle \\ &\quad + \sum_{i=1}^n \pi_{i,t+1} \left\langle \frac{\alpha_t \hat{g}_{i,t}(x_{i,t})}{\phi_{i,t+1}}, \bar{x}_t - \bar{x}_{t+1} \right\rangle \\ &\quad + \sum_{i=1}^n \pi_{i,t+1} \left\langle \frac{\alpha_t \hat{g}_{i,t}(x_{i,t})}{\phi_{i,t+1}}, \bar{x}_{t+1} - x_{i,t+1} \right\rangle \end{aligned} \quad (44)$$

For the first term on the right side of (44), it can be calculated as follows

$$\begin{aligned} &\sum_{i=1}^n \pi_{i,t+1} \left\langle \frac{\alpha_t \hat{g}_{i,t}(x_{i,t})}{\phi_{i,t+1}}, \tilde{x}_t^* - \bar{x}_t \right\rangle \\ &= \sum_{i=1}^n \pi_{i,t+1} \alpha_t \left\langle \frac{\hat{g}_{i,t}(x_{i,t}) - \nabla \hat{f}_{i,t}(\bar{x}_t)}{\phi_{i,t+1}}, \tilde{x}_t^* - \bar{x}_t \right\rangle \\ &\quad + \sum_{i=1}^n \pi_{i,t+1} \alpha_t \left\langle \frac{\nabla \hat{f}_{i,t}(\bar{x}_t)}{\phi_{i,t+1}}, \tilde{x}_t^* - \bar{x}_t \right\rangle \end{aligned} \quad (45)$$

Taking expectation on the above equality with respect to \mathcal{F}_{t-1} , we can obtain

$$\begin{aligned} &\mathbb{E} \left[\sum_{i=1}^n \pi_{i,t+1} \left\langle \frac{\alpha_t \hat{g}_{i,t}(x_{i,t})}{\phi_{i,t+1}}, \tilde{x}_t^* - \bar{x}_t \right\rangle \middle| \mathcal{F}_{t-1} \right] \\ &= \sum_{i=1}^n \pi_{i,t+1} \alpha_t \left\langle \frac{\nabla \hat{f}_{i,t}(x_{i,t}) - \nabla \hat{f}_{i,t}(\bar{x}_t)}{\phi_{i,t+1}}, \tilde{x}_t^* - \bar{x}_t \right\rangle \\ &\quad + \sum_{i=1}^n \pi_{i,t+1} \alpha_t \left\langle \frac{\nabla \hat{f}_{i,t}(\bar{x}_t)}{\phi_{i,t+1}}, \tilde{x}_t^* - \bar{x}_t \right\rangle \\ &\leq \theta \alpha_t \sum_{i=1}^n \pi_{i,t+1} \left\| \nabla \hat{f}_{i,t}(x_{i,t}) - \nabla \hat{f}_{i,t}(\bar{x}_t) \right\| \|\tilde{x}_t^* - \bar{x}_t\| \\ &\quad + \sum_{i=1}^n \pi_{i,t+1} \alpha_t \left\langle \frac{\nabla \hat{f}_{i,t}(\bar{x}_t)}{\phi_{i,t+1}}, \tilde{x}_t^* - \bar{x}_t \right\rangle \end{aligned} \quad (46)$$

where Lemma 1(b) has been employed to obtain the equality, while the inequality follows from Lemma 3. Note that

$\|\tilde{x}_t^* - \bar{x}_t\| \leq 2R$. In addition, Lemma 1 implies that $\|\nabla \hat{f}_{i,t}(x_{i,t}) - \nabla \hat{f}_{i,t}(\bar{x}_t)\| \leq L\|x_{i,t} - \bar{x}_t\|$. Thus, (46) can be further manipulated as

$$\begin{aligned} & \mathbb{E} \left[\sum_{i=1}^n \pi_{i,t+1} \left\langle \frac{\alpha_t \hat{g}_{i,t}(x_{i,t})}{\phi_{i,t+1}}, \tilde{x}_t^* - \bar{x}_t \right\rangle \middle| \mathcal{F}_{t-1} \right] \\ & \leq 2\theta RL\alpha_t \sum_{i=1}^n \pi_{i,t+1} \|x_{i,t} - \bar{x}_t\| \\ & \quad + \alpha_t \sum_{i=1}^n \frac{\pi_{i,t+1}}{\phi_{i,t+1}} \left(\hat{f}_{i,t}(\tilde{x}_t^*) - f_{i,t}(\bar{x}_t) \right) \end{aligned} \quad (47)$$

where the convexity of $\hat{f}_{i,t}(x)$ and the fact that $f_{i,t}(x) \leq \hat{f}_{i,t}(x)$ for all $i \in \mathcal{V}$ and $t \geq 0$ in Lemma 1 (a) have been exploited to obtain the inequality. By recalling (33), the second term on the right side of (44) can be bounded as

$$\begin{aligned} & \sum_{i=1}^n \pi_{i,t+1} \left\langle \frac{\alpha_t \hat{g}_{i,t}(x_{i,t})}{\phi_{i,t+1}}, \bar{x}_t - \bar{x}_{t+1} \right\rangle \\ & \leq \frac{1}{2} \left\| \sum_{i=1}^n \pi_{i,t+1} \frac{\alpha_t \hat{g}_{i,t}(x_{i,t})}{\phi_{i,t+1}} \right\|^2 + \frac{1}{2} \|\bar{x}_t - \bar{x}_{t+1}\|^2 \\ & \leq \frac{1}{2} \theta^2 \alpha_t^2 \sum_{i=1}^n \pi_{i,t+1} \|\hat{g}_{i,t}(x_{i,t})\|^2 + \frac{1}{2} \|\bar{x}_t - \bar{x}_{t+1}\|^2 \end{aligned} \quad (48)$$

where the last inequality follows from Lemma 3, along with the fact that $\|\sum_{i=1}^n \pi_{i,t} x_i\|^2 \leq \sum_{i=1}^n \pi_{i,t} \|x_i\|^2$ for all $x_i \in \mathbb{R}^m$. Taking expectation on both sides of (48) yields that

$$\begin{aligned} & \mathbb{E} \left[\sum_{i=1}^n \pi_{i,t+1} \left\langle \frac{\alpha_t \hat{g}_{i,t}(x_{i,t})}{\phi_{i,t+1}}, \bar{x}_t - \bar{x}_{t+1} \right\rangle \right] \\ & \leq \frac{1}{2} m^2 \theta^2 G^2 \alpha_t^2 + \frac{1}{2} \mathbb{E} [\|\bar{x}_t - \bar{x}_{t+1}\|^2] \end{aligned} \quad (49)$$

The third term on the right side of (44) can be bounded as

$$\begin{aligned} & \sum_{i=1}^n \pi_{i,t+1} \left\langle \frac{\alpha_t \hat{g}_{i,t}(x_{i,t})}{\phi_{i,t+1}}, \bar{x}_{t+1} - x_{i,t+1} \right\rangle \\ & \leq \theta \alpha_t \sum_{i=1}^n \pi_{i,t+1} \|\hat{g}_{i,t}(x_{i,t})\| \|x_{i,t+1} - \bar{x}_{t+1}\| \end{aligned} \quad (50)$$

Recalling (35)-(36), and taking expectation on both sides of (50), we obtain

$$\begin{aligned} & \mathbb{E} \left[\sum_{i=1}^n \pi_{i,t+1} \left\langle \frac{\alpha_t \hat{g}_{i,t}(x_{i,t})}{\phi_{i,t+1}}, \bar{x}_{t+1} - x_{i,t+1} \right\rangle \right] \\ & \leq nm^2 \theta^2 CG^2 \alpha_t \sum_{l=1}^{t+1} \lambda^{t+1-l} \alpha_{l-1} \\ & \quad + m\theta CG \alpha_t \lambda^{t+1} \sum_{j=1}^n \|x_{j,0}\|. \end{aligned} \quad (51)$$

Now, taking expectation on both sides of (38) and applying the above relations, we can obtain

$$\begin{aligned} & \mathbb{E}[\Delta D_t] \\ & \leq \alpha_t \sum_{i=1}^n \frac{\pi_{i,t+1}}{\phi_{i,t+1}} \mathbb{E} \left[\hat{f}_{i,t}(\tilde{x}_t^*) - f_{i,t}(\bar{x}_t) \right] \end{aligned}$$

$$\begin{aligned} & + 2nm\theta^2 RLCG \alpha_t \sum_{l=1}^t \lambda^{t-l} \alpha_{l-1} + 2R \|x_{t+1}^* - x_t^*\| \\ & + 2nm^2 \theta^2 CG^2 \alpha_t \sum_{l=1}^{t+1} \lambda^{t+1-l} \alpha_{l-1} + \frac{1}{2} m^2 \theta^2 G^2 \alpha_t^2 \\ & + 2(RL + mG\lambda) \theta C \alpha_t \lambda^t \sum_{j=1}^n \|x_{j,0}\| \end{aligned} \quad (52)$$

Note that

$$\begin{aligned} - \sum_{t=0}^T \frac{\mathbb{E}[\Delta D_t]}{\alpha_t} & = \sum_{t=0}^T \frac{\mathbb{E}[D_t] - \mathbb{E}[D_{t+1}]}{\alpha_t} \\ & \leq \frac{\mathbb{E}[D_0]}{\alpha_0} + \sum_{t=1}^T \left(\frac{1}{\alpha_t} - \frac{1}{\alpha_{t-1}} \right) \mathbb{E}[D_t] \\ & \leq \frac{2}{\alpha_T} R^2 \end{aligned} \quad (53)$$

where the positivity of D_t , $t \geq 0$, has been utilized to obtain the first inequality, and the last inequality follows from Assumption 2. Then, dividing both sides of (52) by α_t and summing the new inequality over $t \in [T]$, we obtain

$$\begin{aligned} & \sum_{t=0}^T \sum_{i=1}^n \frac{\pi_{i,t+1}}{\phi_{i,t+1}} \mathbb{E} [f_{i,t}(\bar{x}_t) - f_{i,t}(\tilde{x}_t^*)] \\ & \leq 2nm\theta^2 RLCG \sum_{t=0}^T \sum_{l=1}^t \lambda^{t-l} \alpha_{l-1} + \frac{m^2 \theta^2 G^2}{2} \sum_{t=0}^T \alpha_t + \frac{2R^2}{\alpha_T} \\ & \quad + 2nm^2 \theta^2 CG^2 \sum_{t=0}^T \sum_{l=1}^{t+1} \lambda^{t+1-l} \alpha_{l-1} + \mu \theta G(1+T) \\ & \quad + 2(RL + mG\lambda) \theta C \sum_{j=1}^n \|x_{j,0}\| \sum_{t=0}^T \lambda^t + \frac{2R}{\alpha_T} \mathcal{C}_T \end{aligned} \quad (54)$$

Recalling the definition of \tilde{x}_t^* , we have

$$\begin{aligned} & f_{i,t}(\bar{x}_t) - f_{i,t}(x_t^*) \\ & = f_{i,t}(\bar{x}_t) - f_{i,t}(\tilde{x}_t^*) + f_{i,t}(\tilde{x}_t^*) - f_{i,t}(x_t^*) \\ & \leq f_{i,t}(\bar{x}_t) - f_{i,t}(\tilde{x}_t^*) + G \|\tilde{x}_t^* - x_t^*\| \\ & \leq f_{i,t}(\bar{x}_t) - f_{i,t}(\tilde{x}_t^*) + \xi GR \end{aligned} \quad (55)$$

Finally, (15) can be obtained by combining (54) and (55). \blacksquare

APPENDIX D

Proof of Theorem 2. Recalling (2), we have

$$\begin{aligned} \mathcal{R}_j^d(T) & = \sum_{t=0}^T \mathbb{E} [f_t(x_{j,t})] - \sum_{t=0}^T f_t(x_t^*) \\ & = \sum_{t=0}^T \sum_{i=1}^n \mathbb{E} [f_{i,t}(x_{j,t}) - f_{i,t}(\bar{x}_t)] \\ & \quad + \sum_{t=0}^T \sum_{i=1}^n \mathbb{E} [f_{i,t}(\bar{x}_t) - f_{i,t}(x_t^*)] \end{aligned} \quad (56)$$

For the first term on the right side of (56), we have

$$f_{i,t}(x_{j,t}) - f_{i,t}(\bar{x}_t) \leq g_{i,t}^T(x_{j,t})(x_{j,t} - \bar{x}_t)$$

$$\leq G\|x_{j,t} - \bar{x}_t\| \quad (57)$$

where Assumption 2(b) has been utilized to obtain the last inequality. Recalling (35)-(36) and taking expectation on both sides of (57), then the first term on the right side of (56) can be bounded as follows

$$\begin{aligned} & \sum_{t=0}^T \sum_{i=1}^n \mathbb{E}[f_{i,t}(x_{j,t}) - f_{i,t}(\bar{x}_t)] \\ & \leq n^2 m \theta C G^2 \sum_{t=0}^T \sum_{l=1}^t \lambda^{t-l} \alpha_{l-1} \\ & \quad + n C G \sum_{i=1}^n \|x_{i,0}\| \sum_{t=0}^T \lambda^t. \end{aligned} \quad (58)$$

To establish the desired dynamic regret bound, we now only need to bound the last term on the right side of (56). Note that the term $\mathbb{E}[f_{i,t}(\bar{x}_t) - f_{i,t}(x_t^*)]$ is scaled by $\frac{\pi_{i,t+1}}{\phi_{i,t+1}}$ in Theorem 1 and we cannot determine the sign of this term for particular $i \in \mathcal{V}$. Now, we calculate $-\frac{\pi_{i,t+1}}{\phi_{i,t+1}}(f_{i,t}(\bar{x}_t) - f_{i,t}(x_t^*))$ as follows

$$\begin{aligned} & -\frac{\pi_{i,t+1}}{\phi_{i,t+1}}(f_{i,t}(\bar{x}_t) - f_{i,t}(x_t^*)) \\ & = -\frac{\frac{1}{n}\phi_{i,t+1} + \pi_{i,t+1} - \frac{1}{n}\phi_{i,t+1}}{\phi_{i,t+1}}(f_{i,t}(\bar{x}_t) - f_{i,t}(x_t^*)) \\ & \leq \frac{|\frac{1}{n}\phi_{i,t+1} - \pi_{i,t+1}|}{\phi_{i,t+1}} |f_{i,t}(\bar{x}_t) - f_{i,t}(x_t^*)| \\ & \quad - \frac{1}{n}(f_{i,t}(\bar{x}_t) - f_{i,t}(x_t^*)) \\ & \leq \theta C \lambda^{T-t-1} G \|\bar{x}_t - x_t^*\| - \frac{1}{n}(f_{i,t}(\bar{x}_t) - f_{i,t}(x_t^*)) \end{aligned} \quad (59)$$

where the last inequality follows from Lemma 3-4 and Assumption 2(b). Thus, we can further obtain

$$\begin{aligned} & f_{i,t}(\bar{x}_t) - f_{i,t}(x_t^*) \\ & \leq n \frac{\pi_{i,t+1}}{\phi_{i,t+1}} (f_{i,t}(\bar{x}_t) - f_{i,t}(x_t^*)) + 2n\theta C G R \lambda^{T-t-1} \end{aligned} \quad (60)$$

Taking expectation on both sides of (60) and combining Theorem 1, we obtain

$$\begin{aligned} \mathcal{R}_j^d(T) & \leq n^2 m \theta C G (2\theta R L + G) \sum_{t=0}^T \sum_{l=1}^t \lambda^{t-l} \alpha_{l-1} \\ & \quad + 2n^2 m^2 \theta^2 C G^2 \sum_{t=0}^T \sum_{l=1}^{t+1} \lambda^{t+1-l} \alpha_{l-1} + \frac{2nR^2}{\alpha_T} \\ & \quad + (2RL\theta + 2m\theta\lambda G + G)nC \sum_{i=1}^n \|x_{i,0}\| \sum_{t=0}^T \lambda^t \\ & \quad + 2n^2 \theta C G R \sum_{t=0}^T \lambda^{T-t-1} + \frac{1}{2} n m^2 \theta^2 G^2 \sum_{t=0}^T \alpha_t \\ & \quad + \frac{2nR}{\alpha_T} \mathcal{C}_T + n\theta\mu G(T+1) + n\theta\xi G R(T+1) \end{aligned} \quad (61)$$

Now we turn to calculate the terms on the right side of (61), respectively. For the first term, we have

$$\sum_{t=0}^T \sum_{l=1}^t \lambda^{t-l} \alpha_{l-1} = \sum_{t=0}^{T-1} \lambda^t \sum_{l=0}^{T-t-1} \alpha_l \leq \frac{1}{1-\lambda} \sum_{l=0}^{T-1} \alpha_l \quad (62)$$

Analogously, we obtain $\sum_{t=0}^T \sum_{l=1}^{t+1} \lambda^{t+1-l} \alpha_{l-1} \leq \frac{1}{1-\lambda} \sum_{l=0}^T \alpha_l$. Note that $\sum_{t=0}^T \alpha_t \leq 1 + \int_0^{T+1} (t+1)^{-\frac{1}{2}} dt \leq 2\sqrt{T+1}$. Then, (16) can be obtained, which completes the proof. ■

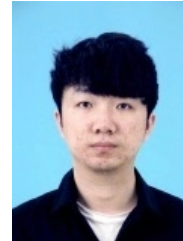
REFERENCES

- [1] Q. Li, R. Heusdens, and M. G. Christensen, "Privacy-preserving distributed optimization via subspace perturbation: A general framework," *IEEE Transactions on Signal Processing*, vol. 68, pp. 5983–5996, 2020.
- [2] Y. Xiong, J. Xu, K. You, J. Liu, and L. Wu, "Privacy preserving distributed online optimization over unbalanced digraphs via subgradient rescaling," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 3, pp. 1366–1378, 2020.
- [3] J. Zhang, K. You, and K. Cai, "Distributed dual gradient tracking for resource allocation in unbalanced networks," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2186–2198, 2020.
- [4] J. Xu, S. Zhu, Y. C. Soh, and L. Xie, "A dual splitting approach for distributed resource allocation with regularization," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 1, pp. 403–414, 2019.
- [5] Y. Xu, T. Han, K. Cai, Z. Lin, G. Yan, and M. Fu, "A distributed algorithm for resource allocation over dynamic digraphs," *IEEE Transactions on Signal Processing*, vol. 65, no. 10, pp. 2600–2612, 2017.
- [6] S. Zhu, C. Chen, J. Xu, X. Guan, L. Xie, and K. H. Johansson, "Mitigating quantization effects on distributed sensor fusion: A least squares approach," *IEEE Transactions on Signal Processing*, vol. 66, no. 13, pp. 3459–3474, 2018.
- [7] D. Jakovetić, J. M. F. Xavier, and J. M. F. Moura, "Convergence rates of distributed Nesterov-like gradient methods on random networks," *IEEE Transactions on Signal Processing*, vol. 62, no. 4, pp. 868–882, 2014.
- [8] K. You, R. Tempo, and P. Xie, "Distributed algorithms for robust convex optimization via the scenario approach," *IEEE Transactions on Automatic Control*, vol. 64, no. 3, pp. 880–895, 2019.
- [9] A. Reiszadeh, A. Mokhtari, H. Hassani, and R. Pedarsani, "An exact quantized decentralized gradient descent algorithm," *IEEE Transactions on Signal Processing*, vol. 67, no. 19, pp. 4934–4947, 2019.
- [10] D. Mateos-Núñez, and J. Cortés, "Distributed online convex optimization over jointly connected digraphs," *IEEE Transactions on Network Science and Engineering*, vol. 1, no. 1, pp. 23–37, 2014.
- [11] X. Cao, and T. Başar, "Decentralized online convex optimization based on signs of relative states," *Automatica*, vol. 129, doi: 10.1016/j.automatica.2021.109676, 2021.
- [12] X. Cao, and T. Başar, "Decentralized online convex optimization with event-triggered communications," *IEEE Transactions on Signal Processing*, vol. 69, pp. 284–299, 2021.
- [13] F. Yan, S. Sundaram, S. V. N. Vishwanathan, and Y. Qi, "Distributed autonomous online learning: Regrets and intrinsic privacy-preserving properties," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 11, pp. 2483–2493, 2013.
- [14] S. Lee, and M. M. Zavlanos, "Distributed primal-dual methods for online constrained optimization," in *Proceedings of American Control Conference*, Boston, USA, 2016, pp. 7171–7176.
- [15] X. Li, X. Yi and L. Xie, "Distributed online convex optimization with an aggregative variable," *IEEE Transactions on Control of Network Systems*, 2021.
- [16] S. Shahrampour, and A. Jadbabaie, "Distributed online optimization in dynamic environments using mirror descent," *IEEE Transactions on Automatic Control*, vol. 63, no. 3, pp. 714–725, 2018.
- [17] R. Dixit, A. S. Bedi, and K. Rajawat, "Online learning over dynamic graphs via distributed proximal gradient algorithm," *IEEE Transactions on Automatic Control*, 2020.
- [18] X. Yi, X. Li, L. Xie, and K. H. Johansson, "Distributed online convex optimization with time-varying coupled inequality constraints," *IEEE Transactions on Signal Processing*, vol. 68, pp. 731–746, 2020.
- [19] A. Nedić, and A. Olshevsky, "Distributed optimization over time-varying directed graphs," *IEEE Transactions on Automatic Control*, vol. 60, no. 3, pp. 601–615, 2015.
- [20] G. Scutari, and Y. Sun, "Distributed nonconvex constrained optimization over time-varying digraphs," *Mathematical Programming*, vol. 176, no. 1-2, pp. 497–544, 2019.
- [21] K. Cai, and H. Ishii, "Average consensus on arbitrary strongly connected digraphs with time-varying topologies," *IEEE Transactions on Automatic Control*, vol. 59, no. 4, pp. 1066–1071, 2014.
- [22] V. S. Mai, and E. H. Abed, "Distributed optimization over directed graphs with row stochasticity and constraint regularity," *Automatica*, vol. 102, pp. 94–104, 2019.

- [23] P. Xie, K. You, R. Tempo, S. Song, and C. Wu, "Distributed convex optimization with inequality constraints over time-varying unbalanced digraphs," *IEEE Transactions on Automatic Control*, vol. 63, no. 12, pp. 4331–4337, 2018.
- [24] C. Xi, and U. A. Khan, "Distributed subgradient projection algorithm over directed graphs," *IEEE Transactions on Automatic Control*, vol. 62, no. 8, pp. 3986–3992, 2017.
- [25] R. Xin, and U. A. Khan, "A linear algorithm for optimization over directed graphs with geometric convergence," *IEEE Control Systems Letters*, vol. 2, no. 3, pp. 315–320, 2018.
- [26] S. Pu, W. Shi, J. Xu, and A. Nedić, "Push-pull gradient methods for distributed optimization in networks," *IEEE Transactions on Automatic Control*, 2020.
- [27] M. Akbari, B. Ghahesifard, and T. Linder, "Distributed online convex optimization on time-varying directed graphs," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 3, pp. 417–428, 2017.
- [28] Y. Pang, and G. Hu, "Randomized gradient-free distributed online optimization with time-varying cost functions," in *2019 IEEE Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 4910–4915.
- [29] X. Yi, X. Li, T. Yang, L. Xie, T. Chai, and K. H. Johansson, "Distributed bandit online convex optimization with time-varying coupled inequality constraints," *IEEE Transactions on Automatic Control*, vol. 66, no. 10, pp. 4620–4635, 2021.
- [30] S. Liu, J. Chen, P. Y. Chen, and A. Hero, "Zeroth-order online alternating direction method of multipliers: Convergence analysis and applications," in *International Conference on Artificial Intelligence and Statistics (AISTATS)*. PMLR, 2018, pp. 288–297.
- [31] Y. Nesterov, and V. Spokoiny, "Random gradient-free minimization of convex functions," *Foundations of Computational Mathematics*, vol. 17, no. 2, pp. 527–566, 2017.
- [32] A. Agarwal, O. Dekel, and L. Xiao, "Optimal algorithms for online convex optimization with multi-point bandit feedback," in *Conference on Learning Theory (COLT)*. PMLR, 2010, pp. 28–40.
- [33] O. Shamir, "An optimal algorithm for bandit and zero-order convex optimization with two-point feedback," *Journal of Machine Learning Research*, vol. 18, no. 1, pp. 1703–1713, 2017.
- [34] J. C. Duchi, M. I. Jordan, M. J. Wainwright, and A. Wibisono, "Optimal rates for zero-order convex optimization: The power of two function evaluations," *IEEE Transactions on Information Theory*, vol. 61, no. 5, pp. 2788–2806, 2015.
- [35] Y. Pang, and G. Hu, "Randomized gradient-free distributed optimization methods for a multi-agent system with unknown cost function," *IEEE Transactions on Automatic Control*, vol. 65, no. 1, pp. 333–340, 2020.
- [36] D. Hajinezhad, M. Hong, and A. Garcia, "ZONE: Zeroth order non-convex multi-agent optimization over networks," *IEEE Transactions on Automatic Control*, 2019.
- [37] Y. Tang, J. Zhang, and N. Li, "Distributed zero-order algorithms for nonconvex multiagent optimization," *IEEE Transactions on Control of Network Systems*, vol. 8, no. 1, pp. 269–281, 2021.
- [38] A. Nedić, and A. Olshevsky, "Stochastic gradient-push for strongly convex functions on time-varying directed graphs," *IEEE Transactions on Automatic Control*, vol. 61, no. 12, pp. 3936–3947, 2016.
- [39] K. Lu, G. Jing, and L. Wang, "Online distributed optimization with strongly pseudoconvex-sum cost functions," *IEEE Transactions on Automatic Control*, vol. 65, no. 1, pp. 426–433, 2019.
- [40] X. Li, X. Yi, and L. Xie, "Distributed online optimization for multi-agent networks with coupled inequality constraints," *IEEE Transactions on Automatic Control*, vol. 66, no. 8, pp. 3575–3591, 2021.
- [41] D. Yuan, and D. W. C. Ho, "Randomized gradient-free method for multiagent optimization over time-varying networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 6, pp. 1342–1347, 2014.
- [42] S. Hosseini, A. Chapman, and M. Mesbahi, "Online distributed convex optimization on dynamic networks," *IEEE Transactions on Automatic Control*, vol. 61, no. 11, pp. 3545–3550, 2016.
- [43] Y. Tang, Z. Ren and N. Li, "Zeroth-order feedback optimization for cooperative multi-agent systems," in *2020 IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 3649–3656.
- [44] Y. Zhang, R. J. Ravier, M. M. Zavlanos, and V. Tarokh, "A distributed online convex optimization algorithm with improved dynamic regret," in *2019 IEEE Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 2449–2454.
- [45] A. Nedić, A. Ozdaglar, and P. A. Parrilo, "Constrained consensus and optimization in multi-agent networks," *IEEE Transactions on Automatic Control*, vol. 55, no. 4, pp. 922–938, 2010.



Yongyang Xiong received the B.S. degree in information and computational science, the M.E. and Ph.D. degrees in control science and engineering from Harbin Institute of Technology, Harbin, China, in 2012, 2014, and 2020, respectively. From 2017 to 2018, he was a Joint Ph.D. Student with the School of Electrical and Electronic Engineering, Nanyang Technological University (NTU), Singapore. He is currently a Postdoctoral Fellow with the department of automation, Tsinghua University, Beijing, China. His current research interests include networked control system, distributed optimization and learning, multi-agent reinforcement learning, and their applications.

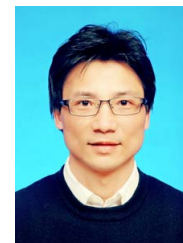


Xiang Li received B.E. degree in Automation and M.E. degree in Control Science and Engineering from Harbin Institute of Technology, Harbin, China, in 2015 and 2018, respectively. He is now pursuing his Ph.D. in the same major. His current research interests include image processing, pattern recognition, three-dimensional reconstruction and machine learning.



Keyou You received the B.S. degree in statistical science from Sun Yat-sen University, Guangzhou, China, in 2007 and the Ph.D. degree in electrical and electronic engineering from Nanyang Technological University (NTU), Singapore, in 2012. After briefly working as a Research Fellow at NTU, he joined Tsinghua University, Beijing, China in 2012 where he is currently an Associate Professor with the Department of Automation. He held visiting positions at Politecnico di Torino, Turin, Italy, the Hong Kong University of Science and Technology, Hong Kong, and the University of Melbourne, Parkville, VIC, Australia. His current research interests include networked control systems, distributed algorithms, and their applications.

Dr. You was a recipient of the Guan Zhaozhi Award at the 29th Chinese Control Conference in 2010, the CSC-IBM China Faculty Award in 2014, and the National Science Fund for Excellent Young Scholars in 2017. He was nominated for the National 1000-Youth Talent Program of China in 2014.



Ligang Wu (M'10-SM'12) received the B.S. degree in Automation from Harbin University of Science and Technology, China in 2001; the M.E. degree in Navigation Guidance and Control from Harbin Institute of Technology, China in 2003; the PhD degree in Control Theory and Control Engineering from Harbin Institute of Technology, China in 2006. From January 2006 to April 2007, he was a Research Associate in the Department of Mechanical Engineering, The University of Hong Kong, Hong Kong. From September 2007 to June 2008, he was a Senior Research Associate in the Department of Mathematics, City University of Hong Kong, Hong Kong. From December 2012 to December 2013, he was a Research Associate in the Department of Electrical and Electronic Engineering, Imperial College London, London, UK. In 2008, he joined the Harbin Institute of Technology, China, as an Associate Professor, and was then promoted to a Full Professor in 2012. Dr. Wu was the winner of the National Science Fund for Distinguished Young Scholars in 2015, and received China Young Five Four Medal in 2016. He was named as the Distinguished Professor of Yangtze River Scholar in 2017, and was named as the Highly Cited Researcher in 2015, 2016 and 2017.

Dr. Wu currently serves as an Associate Editor for a number of journals, including *IEEE TRANSACTIONS ON AUTOMATIC CONTROL*, *IEEE/ASME TRANSACTIONS ON MECHATRONICS*, *Information Sciences*, *Signal Processing*, and *IET Control Theory and Applications*. He is also an Associate Editor for the Conference Editorial Board, IEEE Control Systems Society. Dr. Wu has published 6 research monographs and more than 150 research papers in international referred journals. His current research interests include switched systems, stochastic systems, computational and intelligent systems, multidimensional systems, sliding mode control, and flight control.