2024 European Control Conference (ECC)
June 25-28, 2024. Stockholm, Sweden

Distributed Online Constrained Convex Optimization with Event-Triggered Communication

Kunpeng Zhang[1], Xinlei Yi[2], Yuzhe Li[1], Ming Cao[3], Tianyou Chai[1] and Tao Yang[1]

Abstract— **This paper focuses on the distributed online convex optimization problem with time-varying inequality constraints over a network of agents, where the agents collaborate to minimize the cumulative network-wide loss over time through local information exchange. To reduce communication overhead between the agents, we propose a distributed event-triggered online primal–dual algorithm over a time-varying directed graph. With several classes of appropriately chosen decreasing parameter sequences and non-increasing event-triggered threshold sequences, we establish dynamic network regret and network cumulative constraint violation bounds. Finally, a numerical simulation example is provided to verify the theoretical results.**

I. INTRODUCTION

Distributed online convex optimization is a promising framework due to its powerful modeling capability for various problems in dynamic, uncertain and even adversarial environments. In distributed online convex optimization, the agents collaboratively make decisions without knowing their local loss functions at the current iteration, and then the time-varying local loss functions are privately revealed. The goal is to minimize the cumulative network-wide loss over time. In general, static regret is the standard performance metric to evaluate online algorithms, which measures the difference of the cumulative loss between the decision sequence and the optimal static decision in hindsight. Various distributed online algorithms with sublinear static regret have been developed, see, e.g., [1]–[10], recent survey paper [11] and references herein. For example, the authors of [1] develop a projection-based distributed online subgradient descent algorithm. By using Bregman divergence in lieu of Euclidean distance for projection, the authors of [2] propose a decentralized online mirror descent algorithm.

Note that the aforementioned distributed online algorithms require local information exchange between the agents via the underlying communication topology at each iteration, which may cause large amount of communication overhead. To overcome this limitation, by equipping event-triggered communication scheme to the algorithm proposed in [1], the authors of [12] propose two distributed event-triggered algorithms with full-information feedback and bandit feedback over a fixed undirected graph, respectively. Moreover, sublinear static regret is achieved for both algorithms when their event-triggering threshold sequences are non-increasing and converge to zero. By using one-point and two-point subgradient estimators respectively, two distributed event-triggered algorithms with bandit feedback are developed in [13] for the fix delayed bandit feedback case and sublinear static regret is established for the algorithms. The authors of [14] develop a distributed event-triggered algorithm based on the algorithm proposed [2] and sublinear static regret is achieved.

It is worth mentioning that the above algorithms which achieve performance close to the best static regret may perform poorly in terms of dynamic regret [15]. Dynamic regret is a more stringent performance metric, which measures the difference of the cumulative loss between the decision sequence and the best decision sequence selected by a clairvoyant that knows the sequence of loss functions in advance. By using the first and second moments of the gradient of the local loss functions, the authors of [16] develop the event-triggered algorithm with full-information feedback in [12] by imposing adaptive updating step-sizes, and analyze dynamic regret.

Most existing studies for online constrained convex optimization focus on the case where the feasible set is a simple closed convex set (a box or a ball). To cope with more complex scenarios, the authors of [17], [18] characterize the feasible set by inequality constraints and a simple closed convex set for centralized online constrained convex optimization. Note that projection operation is needed at each iteration in order to satisfy always inequality constraints, which results in heavy computation burden. Therefore, the authors treat these inequality constraints as long-term constraints, i.e., the inequality constraints are allowed to be violated but are satisfied in the long run. Inspired by [17], [18], a distributed online primal–dual algorithm with full-information feedback is proposed in [19], where agents need to share their local decisions with their neighboring agents via the underlying communication topology at each iteration. To reduce communication overhead, in this paper, we propose a distributed event-triggered online primal–dual algorithm over an uniformly jointly strongly connected time-varying directed graph by integrating event-triggered communication into the algorithm in [19], where each agent broadcasts its current local decision to its neighboring agents only if norm of the difference between the decision and the

*This work was supported by the National Key Research and Development Program of China under Grant 2022YFB3305904.

[1]K. Zhang, Y. Li, T. Chai and T. Yang are with the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China 2110343@stu.neu.edu.cn; {yuzheli; tychai; yangtao}@mail.neu.edu.cn

[2]X. Yi is with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, USA xinleiyi@mit.edu

[3]M. Cao is with the Engineering and Technology Institute Groningen, Faculty of Science and Engineering, University of Groningen, AG 9747 Groningen, The Netherlands m.cao@rug.nl

last broadcasted decision is not less than the current event-triggering threshold. In addition, base on several classes of appropriately chosen event-triggered threshold sequences, we analyzes the impact of event-triggering threshold on dynamic network regret and network cumulative constraint violation.

The contributions are as follows.

- To the best of our knowledge, this paper is among the first to consider time-varying inequality constraints for distributed online convex optimization with event-triggered communication. Compared to distributed event-triggered online algorithms [12]–[14], [16], which only consider a simple closed convex constrained set and an undirected fixed communication graph, we consider time-varying inequality constraints and a directed time-varying graph.
- In Theorem 1, we show that the proposed algorithm achieves sublinear dynamic network regret and network cumulative constraint violation if the path–length of the benchmark, the accumulated dynamic variation of the optimal decision sequence, grows sublinearly and the non-increasing event-triggering threshold sequence converges to zero. With two classes of natural decreasing event-triggering threshold sequences, in Corollaries 1 and 2, we respectively establish in sublinear dynamic network regret and network cumulative constraint violation bounds. These bounds recover the results achieved by the centralized online algorithm in [20], and the distributed event-triggered online algorithm in [16].
- In Theorem 2, by appropriately designing the parameter sequences, which avoid the impact of the event-triggering threshold on the updating step-sizes of the local primal variables, we establish sublinear dynamic network regret and network cumulative constraint violation bounds. These bounds recover the results achieved by the centralized online algorithms in [20], the distributed online algorithms in [19], and the distributed event-triggered online algorithm in [16].

The remainder of this paper is as follows. Section II presents the problem formulation and motivation. Section III proposes the distributed event-triggered online primal–dual algorithm and the performance metrics. Section IV analyzes the performance of the proposed algorithm. Section V demonstrates numerical simulations. Finally, Section VI concludes the paper. Due to the space limitations, we omit detailed proofs, which can be found in [21].

Notations: $\mathbb{N}_+$, $\mathbb{R}$, $\mathbb{R}^p$ and $\mathbb{R}^p_+$ denote the set of all positive integers, real numbers, p-dimensional vectors and nonnegative vectors, respectively. $[n]$ denotes the set $\{1,\cdot\cdot\cdot,n\}$ for any $n\in\mathbb{N}_+$. Given vectors x and y, x^T denotes the transpose of the vector x, and $\langle x,y\rangle$ and $x\otimes y$ denote the standard inner and Kronecker product of the vectors x and y, respectively. $\mathbf{0}_m$ denotes the m-dimensional column vector whose components are all 0. $\mathrm{col}(q_1,\cdot\cdot\cdot,q_n)$ denotes the concatenated column vector of $q_i\in\mathbb{R}^{m_i}$ for $i\in[n]$. For a set $\mathbb{K}\in\mathbb{R}^p$, $\mathcal{P}_{\mathbb{K}}(\cdot)$ denotes a projection operator, i.e., $\mathcal{P}_{\mathbb{K}}(x)=\arg\min_{y\in\mathbb{K}}\|x-y\|^2$, $\forall x\in\mathbb{R}^p$, and $[\cdot]_+$ denotes $\mathcal{P}_{\mathbb{R}^p_+}(\cdot)$. For a scalar function $f(x):\mathbb{R}^p\to\mathbb{R}$, $\partial f(x)$ denotes the subgradient of $f(x)$.

II. PROBLEM FORMULATION AND MOTIVATION

Consider a repeated game with T iterations over a network of n agents. At iteration t, the agents indexed by $i\in[n]$ exchange information with their neighboring agents via an underlying communication topology, and then select decisions $x_{i,t}\in\mathbb{X}$ without knowing the local loss functions $f_{i,t}:\mathbb{X}\to\mathbb{R}$ and constraint functions $g_{i,t}:\mathbb{X}\to\mathbb{R}^{m_i}$, where $\mathbb{X}\subseteq\mathbb{R}^p$ is known convex set to the agents. After that, the local loss functions $f_{i,t}$ and constraint functions $g_{i,t}$ are privately revealed. Accordingly, the agents suffer losses $f_{i,t}(x_{i,t})$. The goal of the network is to minimize the average of the network-wide loss accumulated over T iterations, i.e., $\frac{1}{n}\sum_{i=1}^n\sum_{t=1}^T f_t(x_{i,t})$. Note that at iteration t,

$$f_t(x)=\frac{1}{n}\sum_{j=1}^n f_{j,t}(x), \tag{1}$$

$$g_t(x)=\mathrm{col}\big(g_{1,t}(x),\cdot\cdot\cdot,g_{n,t}(x)\big), \tag{2}$$

are the global loss and constraint functions of the network, respectively.

Let the set of the feasible decision sequences

$$\mathcal{X}_T=\{(x_1,\cdot\cdot\cdot,x_T):g_t(x_t)\le\mathbf{0}_m,x_t\in\mathbb{X},\forall t\in[T]\}, \tag{3}$$

and the set of the feasible static decision sequences

$$\hat{\mathcal{X}}_T=\{(x,\cdot\cdot\cdot,x):g_t(x)\le\mathbf{0}_m,x\in\mathbb{X},\forall t\in[T]\}, \tag{4}$$

are non-empty, where $m=\sum_{j=1}^n m_j$.

The communication topology among agents is described by a time-varying directed graph $\mathcal{G}_t=(\mathcal{V},\mathcal{E}_t)$, where $\mathcal{V}=[n]$ is the set of agents and $\mathcal{E}_t\subseteq\mathcal{V}\times\mathcal{V}$ is the set of edges at iteration t. A directed edge $(j,i)\in\mathcal{E}_t$ implies that agent i can receive information from agent j at iteration t. The sets of in- and out-neighbors of agent i at iteration t are $\mathcal{N}_i^{\text{in}}(\mathcal{G}_t)=\{j\in[n]|(j,i)\in\mathcal{E}_t\}$ and $\mathcal{N}_i^{\text{out}}(\mathcal{G}_t)=\{j\in[n]|(i,j)\in\mathcal{E}_t\}$, respectively. The associated weight mixing matrix $W_t\in\mathbb{R}^{n\times n}$ satisfies that $[W_t]_{ij}>0$ if $(j,i)\in\mathcal{E}_t$ or $i=j$, and $[W_t]_{ij}=0$ otherwise.

In this paper, the following assumptions are made, which are commonly adopted in distributed online convex optimization, see, e.g., recent survey paper [11] and references herein.

Assumption 1. *(i) The set $\mathbb{X}$ is convex and closed. Moreover, it is bounded by a positive constant $R(\mathbb{X})$, i.e., for any $x\in\mathbb{X}$*

$$\|x\|\le R(\mathbb{X}). \tag{5}$$

(ii) For all $i\in[n]$, $t\in\mathbb{N}_+$, the local loss functions $f_{i,t}$ and constraint functions $g_{i,t}$ are convex, and there exists a positive constant F_1 such that

$$|f_{i,t}(x)-f_{i,t}(y)|\le F_1, \tag{6a}$$

$$\|g_{i,t}(x)\|\le F_1, x,y\in\mathbb{X}. \tag{6b}$$

(iii) For all $i\in[n]$, $t\in\mathbb{N}_+$, the subgradients $\partial f_{i,t}(x)$ and $\partial g_{i,t}(x)$ exist, and there exists a positive constant F_2 such

that

$$\|\partial f_{i,t}(x)\| \le F_2, \tag{7a}$$

$$\|\partial g_{i,t}(x)\| \le F_2, x \in \mathbb{X}. \tag{7b}$$

Assumption 2. *For $t \in \mathbb{N}_+$, the time-varying directed graph $\mathcal{G}_t$ satisfies that*
(i) There exists a constant $w \in (0,1)$, such that $[W_t]_{ij} \ge w$ if $[W_t]_{ij} > 0$.
(ii) The mixing matrix W_t is doubly stochastic, i.e., $\sum_{i=1}^{n} [W_t]_{ij} = \sum_{j=1}^{n} [W_t]_{ij} = 1$, $\forall i, j \in [n]$.
(iii) There exists an integer $B > 0$ such that the time-varying directed graph $(\mathcal{V}, \cup_{l=0}^{B-1} \mathcal{E}_{t+l})$ is strongly connected.

The authors of [19] consider the distributed online convex optimization problem with time-varying inequality constraints, and propose a distributed online primal–dual algorithm with full-information feedback, where the agents at each iteration need to share their decisions through a communication network. However, network resources are often limited. To reduce communication overhead, this paper integrates event-triggered communication into the algorithm.

III. DISTRIBUTED EVENT-TRIGGERED ONLINE PRIMAL–DUAL ALGORITHM

In this section, we propose a distributed event-triggered online primal–dual algorithm. Moreover, we present the performance metrics to evaluate the algorithm.

A. Algorithm Description

By integrating event-triggered communication into the algorithm with full-information feedback in [19], the distributed event-triggered online primal–dual algorithm is proposed from the perspective of each agent, which is presented in pseudo-code as Algorithm 1.

In Algorithm 1, for $t \in [T]$ with $t \ge 2$ and $i \in [n]$, by the distributed consensus protocol (8), agent i computes $z_{i,t} \in \mathbb{X}$ via the time-varying directed graph $\mathcal{G}_t$. In addition, by the primal–dual protocol (9)–(11), agent i updates its local primal variable $x_{i,t} \in \mathbb{X}$ and dual variable $q_{i,t} \in \mathbb{R}_+^{m_i}$, where $\omega_{i,t}$ is the updating direction of the local primal variable, α_t and β_t are the updating step-sizes of the local primal and dual variables, respectively, and γ_t is the regularization parameter. The current decision of agent i is broadcasted only if norm of the difference between the decision and the last broadcasted decision is not less than the current event-triggering threshold τ_t.

The following assumption is made for the event-triggering threshold, which is also adopted in [12]–[14], [16].

Assumption 3. *The event-triggering threshold sequence $\{\tau_t \ge 0\}$ satisfies $\tau_{t+1} \le \tau_t$ for all $t \ge 2$.*

B. Performance Metrics

Algorithm 1 Distributed Event-Triggered Online Primal–Dual Algorithm

```
Input: Decreasing and positive sequences $\{\alpha_t\}$, $\{\beta_t\}$, $\{\gamma_t\}$,
    and non-increasing and positive sequence $\{\tau_t\}$.
Initialize: For $i \in [n]$, initialize $x_{i,1} \in \mathbb{X}$, $\hat{x}_{i,1} = x_{i,1}$ and
    $q_{i,1} = \mathbf{0}_{m_i}$, and broadcast $\hat{x}_{i,1}$ to $\mathcal{N}_i^{\text{out}}(\mathcal{G}_1)$ and receive
    $\hat{x}_{j,1}$ from $j \in \mathcal{N}_i^{\text{in}}(\mathcal{G}_1)$.
  for $t = 1, \cdots, T-1$ do
    for $i = 1, \cdots, n$ in parallel do
      Observe $\partial f_{i,t}(x_{i,t})$, $\partial[g_{i,t}(x_{i,t})]_+$, and $[g_{i,t}(x_{i,t})]_+$;
      Distributed consensus protocol:
          $z_{i,t+1} = \sum_{j=1}^{n} [W_t]_{ij} \hat{x}_{j,t},$                     (8)
      Primal–dual protocol:
          $\omega_{i,t+1} = \partial f_{i,t}(x_{i,t}) + \partial[g_{i,t}(x_{i,t})]_+ q_{i,t},$        (9)
          $x_{i,t+1} = \mathcal{P}_{\mathbb{X}}(z_{i,t+1} - \alpha_{t+1}\omega_{i,t+1}),$               (10)
          $q_{i,t+1} = \Big[(1 - \beta_{t+1}\gamma_{t+1}) q_{i,t} + \gamma_{t+1}\Big([g_{i,t}(x_{i,t})]_+$
              $+ \big(\partial[g_{i,t}(x_{i,t})]_+\big)^T (x_{i,t+1} - x_{i,t})\Big)\Big]_+.$   (11)
      Event-triggering check:
      if $\|x_{i,t+1} - \hat{x}_{i,t}\| \ge \tau_{t+1}$ then
        Set $\hat{x}_{i,t+1} = x_{i,t+1}$, and broadcast $\hat{x}_{i,t+1}$ to
        $\mathcal{N}_i^{\text{out}}(\mathcal{G}_{t+1})$.
      else
        Set $\hat{x}_{i,t+1} = \hat{x}_{i,t}$, and do not broadcast.
      end if
    end for
  end for
Output: $\{x_{i,t}\}$.
```

We adopt network regret and cumulative constraint violation as performance metrics to evaluate Algorithm 1 as in [19], which are respectively defined as

$$\text{Net-Reg}(\{x_{i,t}\}, y_{[T]}) := \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T} f_t(x_{i,t}) - \sum_{t=1}^{T} f_t(y_t), \tag{12}$$

$$\text{Net-CCV}(\{x_{i,t}\}) := \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T} \|[g_t(x_{i,t})]_+\|, \tag{13}$$

where $y_{[T]} = (y_1, \cdots, y_T)$ is a benchmark.

Note that the network cumulative constraint violation (13) is appropriate to some applications where constraints have no cumulative nature, such as safety-critical applications. Moreover, because network cumulative constraint violation (13) avoids constraint violations at some iterations to be canceled out by strictly feasible decisions at other iterations, it is stricter than the network constraint violation adopted in [8], [9], [22] which takes the summation across rounds before the projection operation $[\cdot]_+$.

Moreover, we consider dynamic and static network regret, i.e., $\text{Net-Reg}(\{x_{i,t}\}, \check{x}_{[T]}^*)$ and $\text{Net-Reg}(\{x_{i,t}\}, \hat{x}_{[T]}^*)$. For dynamic network regret, the benchmark $\check{x}_{[T]}^* = (\check{x}_1^*, \cdots, \check{x}_T^*)$ is the optimal decision sequence, where $\check{x}_t^* \in \mathbb{X}$ is the minimizer of $f_t(x)$ subject to $c_t(x) \leq \mathbf{0}_m$. For static network regret, the benchmark $\hat{x}^*_{[T]} = (\hat{x}^*, \cdots, \hat{x}^*)$ is the optimal static decision sequence, where $\hat{x}^* \in \mathbb{X}$ is the minimizer of $\sum_{t=1}^{T} f_t(x)$ subject to $g_t(x) \leq \mathbf{0}_m$ for $t \in [T]$.

IV. Performance Analysis

In this section, we establish dynamic network regret and network cumulative constraint violation bounds for Algorithm 1.

Firstly, inspired by [12], [13], [19], we specially design the updating step-size sequences of the local primal and dual variables, and the regularization parameter sequence of Algorithm 1 in the following theorem.

Theorem 1. *Suppose Assumptions 1–3 hold. Let $\{x_{i,t}\}$ be the sequences generated by Algorithm 1 with*

$$\alpha_t = \sqrt{\frac{\Psi_t}{t}}, \beta_t = \frac{1}{t^\kappa}, \gamma_t = \frac{1}{t^{1-\kappa}}, \forall t \in \mathbb{N}_+, \tag{14}$$

where $\Psi_t = \sum_{s=1}^{t} \tau_s$, $\kappa \in (0,1)$ are constants. Then, for any $T \in \mathbb{N}_+$ and any comparator sequence $y_{[T]} \in \mathcal{X}_T$,

$$\text{Net-Reg}(\{x_{i,t}\}, y_{[T]}) = \mathcal{O}(T^\kappa + \sqrt{\Psi_T T} + \sqrt{\Psi_T{}^{-1} T P_T}), \tag{15}$$

$$\text{Net-CCV}(\{x_{i,t}\}) = \mathcal{O}(T^{1-\kappa/2} + \sqrt[4]{\Psi_T T^3}), \tag{16}$$

where $P_T = \sum_{t=1}^{T-1} \|y_{t+1} - y_t\|$ is the path–length of the benchmark $y_{[T]}$.

Remark 1. *Theorem 1 establishes dynamic network regret bound* (15) *and network cumulative constraint violation bound* (16) *for Algorithm 1. If the path–length of the benchmark grows sublinearly, and τ_t converges to zero, i.e., $\sum_{k=1}^{t} \tau_k$ grows sublinearly, then these bounds are sublinear. Moreover, note that $\sqrt{\Psi_T T}$ and $\sqrt{\Psi_T^{-1} T}$ are derived for α_t given by* (14)*. Due to the event-triggering threshold τ_t, the bound* (15) *is different from the well-known best regret bound $\mathcal{O}(\sqrt{T})$ achieved by the centralized online algorithm in [23].*

Remark 2. *When $\kappa = 1/2$, we have $\sqrt{\Psi_T T} > T^\kappa$ and $\sqrt[4]{\Psi_T T^3} > T^{1-\kappa/2}$. Therefore the bounds* (15) *and* (16) *become $\mathcal{O}(\sqrt{\Psi_T T} + \sqrt{\Psi_T^{-1} T P_T})$ and $\mathcal{O}(\sqrt[4]{\Psi_T T^3})$, respectively. Note that it follows from the bounds* (15) *and* (16) *that there exists a trade-off between dynamic network regret and network cumulative constraint violation, that is, as κ increases, dynamic network regret increases, while network cumulative constraint violation decreases.*

Remark 3. *The proof of Theorem 1 has substantial differences compared to the proof of Theorem 1 in [19]. More specifically, in our Algorithm 1, the agents broadcast the current local decisions only if the event-triggering condition is satisfied. Therefore, the resulting decision sequence is different with Algorithm 1 without event-triggered communication in [19] although the updating rules are same. This critical difference leads to challenges in theoretical proof because we need to reanalyse all results related to the local decisions, e.g., the disagreement among agents, the global loss and constraint, dynamic network regret and network cumulative constraint violation bounds. To tackle this challenge, we derive the upper bounds for the difference between the last broadcasted local decisions and the current local decisions using the current event-triggering threshold. Consequently, the established dynamic network regret bound* (15) *and network cumulative constraint violation bound* (16) *are subject to event-triggering threshold.*

Next, we select the event-triggering threshold sequence produced by $\tau_t = 1/t^\theta$ in the following corollary, which is also adopted by the distributed online algorithms in [12]–[14], [16] since $1/t^\theta$ naturally satisfies two conditions: 1) it is non-increasing; 2) it converges to zero.

Corollary 1. *Under the same conditions as in Theorem 1 with $\tau_t = 1/t^\theta$ and $\theta > 0$, for any $T \in \mathbb{N}_+$, it holds that*

$$\begin{aligned}&\text{Net-Reg}(\{x_{i,t}\}, y_{[T]})\\ &= \begin{cases} \mathcal{O}(T^{\max\{\kappa, 1-\theta/2\}} + T^{\theta/2} P_T), & \text{if } 0 < \theta < 1,\\ \mathcal{O}\big(T^\kappa + \sqrt{T\log(T)} + \sqrt{\frac{T}{\log(T)}} P_T\big), & \text{if } \theta = 1,\\ \mathcal{O}(T^{\max\{\kappa, 1/2\}} + T^{1/2} P_T), & \text{if } \theta > 1,\end{cases}\end{aligned} \tag{17}$$

$$\begin{aligned}&\text{Net-CCV}(\{x_{i,t}\})\\ &= \begin{cases} \mathcal{O}(T^{\max\{1-\kappa/2, 1-\theta/4\}}), & \text{if } 0 < \theta < 1,\\ \mathcal{O}(T^{1-\kappa/2} + \sqrt[4]{T^3 \log(T)}), & \text{if } \theta = 1,\\ \mathcal{O}(T^{\max\{1-\kappa/2, 3/4\}}), & \text{if } \theta > 1.\end{cases}\end{aligned} \tag{18}$$

Remark 4. *The bounds* (17) *and* (18) *are sublinear if the path–length P_T grows sublinearly. If $\theta > 1$, the bound* (17) *recover the results achieved by the centralized online algorithm in [20] and the distributed event-triggered online algorithm in [16]. In addition, our Algorithm 1 is able to handle (time-varying) inequality constraints, whereas the algorithms in [16], [20] are limited to a ball set and a box set, respectively.*

We then choose the event-triggering threshold sequence produced by $\tau_t = 1/c^t$ in the following corollary, which is also adopted in distributed optimization with event-triggered communication, see, e.g., [24]–[28].

Corollary 2. *Under the same conditions as in Theorem 1 with $\tau_t = 1/c^t$ and $c > 1$, for any $T \in \mathbb{N}_+$, it holds that*

$$\text{Net-Reg}(\{x_{i,t}\}, y_{[T]}) = \mathcal{O}(T^{\max\{\kappa, 1/2\}} + T^{1/2} P_T), \tag{19}$$

$$\text{Net-CCV}(\{x_{i,t}\}) = \mathcal{O}(T^{\max\{1-\kappa/2, 3/4\}}). \tag{20}$$

Remark 5. *The bounds* (19) *and* (20) *recover the results achieved in Corollary 1 with $\theta > 1$. Moreover, the bound* (19) *recover the results achieved by the centralized online algorithm in [20] and the distributed event-triggered online algorithm in [16].*

Note that in (14), the event-triggering threshold also affects the updating step-size α_t of the local primal variable in addition to communication between the agents. To avoid that, we appropriately design new parameter sequences for Algorithm 1 in the following theorem.

Theorem 2. *Suppose Assumptions 1–3 hold. Let $\{x_{i,t}\}$ be the sequences generated by Algorithm 1 with*

$$\alpha_t = \frac{\alpha_0}{t^{\theta_1}}, \beta_t = \frac{1}{t^{\theta_2}}, \gamma_t = \frac{1}{t^{1-\theta_2}}, \tau_t = \frac{\tau_0}{t^{\theta_3}}, \forall t \in \mathbb{N}_+, \quad (21)$$

where α_0, $\theta_1 \in (0,1)$, $\theta_2 \in (0,1)$, τ_0 and θ_3 are constants. Then, for any $T \in \mathbb{N}_+$ and any comparator sequence $y_{[T]} \in \mathcal{X}_T$,

$$\begin{aligned}&\text{Net-Reg}(\{x_{i,t}\}, y_{[T]})\\ &= \begin{cases} \mathcal{O}\big(\alpha_0 T^{1-\theta_1} + T^{\theta_2} + \frac{\tau_0}{\alpha_0} T^{1+\theta_1-\theta_3} \\ \quad + \frac{T^{\theta_1}(1+P_T)}{\alpha_0}\big), \quad \text{if } \theta_1 < \theta_3 < 1+\theta_1, \\ \mathcal{O}\big(\alpha_0 T^{1-\theta_1} + T^{\theta_2} + \frac{\tau_0}{\alpha_0}\log(T) \\ \quad + \frac{T^{\theta_1}(1+P_T)}{\alpha_0}\big), \quad \text{if } \theta_3 = 1+\theta_1, \\ \mathcal{O}\big(\alpha_0 T^{1-\theta_1} + T^{\theta_2} + \frac{\tau_0}{\alpha_0} \\ \quad + \frac{T^{\theta_1}(1+P_T)}{\alpha_0}\big), \quad \text{if } \theta_3 > 1+\theta_1, \end{cases}\end{aligned} \quad (22)$$

$$\begin{aligned}&\text{Net-CCV}(\{x_{i,t}\})\\ &= \begin{cases} \mathcal{O}(\sqrt{\alpha_0} T^{1-\theta_1/2} + T^{1-\theta_2/2} \\ \quad + \sqrt{\tau_0} T^{1-\theta_3/2}), \quad \text{if } \theta_1 < \theta_3 < 1, \\ \mathcal{O}\big(\sqrt{\alpha_0} T^{1-\theta_1/2} + T^{1-\theta_2/2} \\ \quad + \sqrt{\tau_0 T \log(T)}\big), \text{if } \theta_3 = 1, \\ \mathcal{O}(\sqrt{\alpha_0} T^{1-\theta_1/2} + T^{1-\theta_2/2} \\ \quad + \sqrt{\tau_0} T^{1/2}), \quad \text{if } \theta_3 > 1. \end{cases}\end{aligned} \quad (23)$$

Remark 6. *Theorem 2 establishes dynamic network regret bound* (22) *and network cumulative constraint violation bound* (23) *for Algorithm 1. If the path–length of the benchmark grows sublinearly, then these bounds are sublinear. Moreover, the bounds* (22) *and* (23) *show the impact of event-triggering threshold on dynamic network regret and network cumulative constraint violation, that is, the larger τ_0 is, the larger the bounds are. When $\tau_0 = 0$, i.e., without event-triggered communication, the bounds are the same as the results achieved by distributed online algorithms in [19] when choosing $\theta_1 = \theta_2$.*

Remark 7. *Note that to achieve dynamic network regret and network cumulative constraint violation, τ_t cannot be selected as a fixed positive constant in Theorems 1 and 2 due to the conditions that event-triggering threshold τ_t converges to zero in Theorem 1 and $\theta_3 > \theta_1$ in Theorem 2, respectively.*

Remark 8. *It should be pointed out that the smallest network cumulative constraint violation bounds in Corollaries 1 and 2 are $\mathcal{O}(T^{3/4})$, which are reduced to $\mathcal{O}(T^{1/2})$ in Theorem 2. Moreover, the smallest dynamic network regret bounds in Corollaries 1 and 2 are $\mathcal{O}(T^{1/2})$, which are recovered in Theorem 2. In addition, the bounds* (22) *and* (23) *recover the results achieved by the centralized online algorithms in [20], the distributed online algorithms in [19], and the distributed event-triggered online algorithm in [16].*

Remark 9. *By replacing the any benchmark $y_{[T]}$ with the optimal static decision sequence $\hat{x}^*_{[T]}$, we have $P_T \equiv 0$, and then the static network regret and cumulative constraint violation bounds for Algorithm 1 with corresponding parameter sequences can be easily established based on the results in Theorems 1 and 2, and Corollaries 1 and 2, respectively, which are the same as* (15)–(20), (22), *and* (23) *with $P_T \equiv 0$, respectively. The bounds recover the results achieved by the centralized online algorithm in [23], the distributed online algorithms in [7], [19], and the distributed event-triggered online algorithms in [12]–[14].*

V. NUMERICAL EXAMPLE

Consider a distributed online linear regression problem with time-varying linear inequality constraints over a network of n agents in [19]. At each iteration t, agent i for $i \in [n]$ accesses to the local loss and constraint functions, i.e., $f_{i,t}(x) = \frac{1}{2}\|A_{i,t}x - \vartheta_{i,t}\|^2$ and $g_{i,t}(x) = B_{i,t}x - b_{i,t}$, where each entry of $A_{i,t} \in \mathbb{R}^{q_i \times p}$ is randomly generated from the uniform distribution in the interval $[-1,1]$, $\vartheta_{i,t} = A_{i,t}\mathbf{1}_p + \zeta_{i,t}$, where $\zeta_{i,t}$ is a standard normal random vector, and each entry of $B_{i,t} \in \mathbb{R}^{m_i \times p}$ and $b_{i,t} \in \mathbb{R}^{m_i}$ is randomly generated from the uniform distribution in the interval $[0,2]$ and $[0,1]$, respectively. We set $n = 100$, $q_i = 4$, $p = 10$, $m_i = 2$, $\mathbb{X} = [-5,5]^p$. At each iteration t, we use an undirected random graph to model the underlying communication topology. Specifically, connections between the agents are random and the probability of two agents being connected is 0.1. To make sure that Assumption 2 is satisfied, we add edges $(i, i+1)$ for $i \in [n-1]$, and let $[W_t]_{ij} = \frac{1}{n}$ if $(j,i) \in \mathcal{E}_t$ and $[W_t]_{ii} = 1 - \sum_{j=1}^n [W_t]_{ij}$.

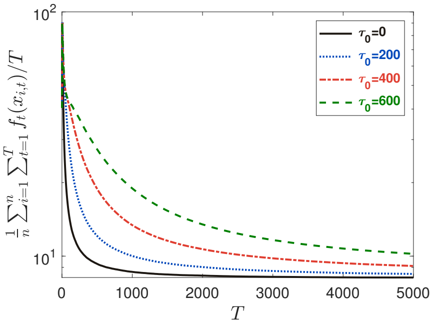

Fig. 1: Evolutions of $\frac{1}{n}\sum_{i=1}^n \sum_{t=1}^T f_t(x_{i,t})/T$.

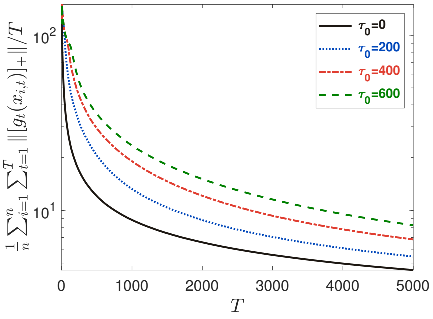

Fig. 2: Evolutions of $\frac{1}{n}\sum_{i=1}^n \sum_{t=1}^T \|[g_t(x_{i,t})]_+\|/T$.

Set $\alpha_t = 1/t^{1/2}$, $\beta_t = 1/t^{1/2}$, $\gamma_t = 1/t^{1/2}$ and $\tau_t = \tau_0/t$ for Algorithm 1. To explore the impact of different

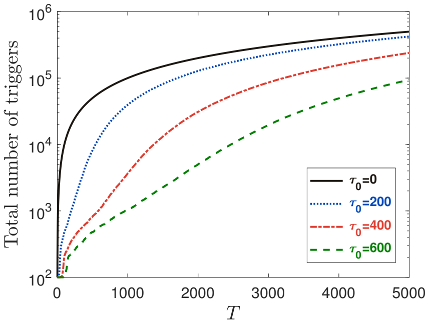

Fig. 3: Evolutions of total number of triggers.

event-triggering threshold sequences on network regret and cumulative constraint violation, we select $\tau_0 = 0$, $\tau_0 = 200$, $\tau_0 = 400$, and $\tau_0 = 600$, respectively. With different values of τ_0, Figs. 1–3 illustrate the evolutions of the average cumulative loss $\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T} f_t(x_{i,t})/T$, the average cumulative constraint violation $\frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T} \|[g_t(x_{i,t})]_+\|/T$, and total number of triggers, respectively. The results show that as τ_0 increases, the average cumulative loss and the average cumulative constraint violation increase, while the total number of triggers decreases, which are consistent with the theoretical results in Theorem 2.

VI. Conclusions

This paper considered the distributed online convex optimization problem with time-varying inequality constraints. We proposed the distributed event-triggered online primal–dual algorithm to reduce communication overhead for a time-varying directed graph. We analyzed the network regret and cumulative constraint violation for the proposed algorithm. Our theoretical results are comparable to the results achieved by the related centralized and distributed online algorithms and distributed event-triggered online algorithms when the event-triggering threshold is properly chosen. In the future, we will investigate the distributed event-triggered online primal–dual algorithm with bandit feedback as gradient information is unavailable in many real-world applications. Moreover, we will also consider to eliminate the need of doubly stochastic.

References

[1] F. Yan, S. Sundaram, S. 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, 2012.

[2] 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, 2017.

[3] S. Lee, A. Nedić, and M. Raginsky, "Stochastic dual averaging for decentralized online optimization on time-varying communication graphs," *IEEE Transactions on Automatic Control*, vol. 62, no. 12, pp. 6407–6414, 2017.

[4] W. Zhang, P. Zhao, W. Zhu, S. C. Hoi, and T. Zhang, "Projection-free distributed online learning in networks," in *International Conference on Machine Learning*, 2017, pp. 4054–4062.

[5] D. Yuan, D. W. Ho, and G. Jiang, "An adaptive primal-dual subgradient algorithm for online distributed constrained optimization," *IEEE Transactions on Cybernetics*, vol. 48, no. 11, pp. 3045–3055, 2017.

[6] D. Yuan, A. Proutiere, and G. Shi, "Distributed online linear regressions," *IEEE Transactions on Information Theory*, vol. 67, no. 1, pp. 616–639, 2021.

[7] ——, "Distributed online optimization with long-term constraints," *IEEE Transactions on Automatic Control*, vol. 67, no. 3, pp. 1089–1104, 2022.

[8] 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.

[9] 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, 2020.

[10] ——, "Distributed online convex optimization with an aggregative variable," *IEEE Transactions on Control of Network Systems*, vol. 9, no. 1, pp. 438–449, 2022.

[11] X. Li, L. Xie, and N. Li, "A survey on distributed online optimization and online games," *Annual Reviews in Control*, vol. 56, p. 100904, 2023.

[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] M. Xiong, B. Zhang, D. Yuan, Y. Zhang, and J. Chen, "Event-triggered distributed online convex optimization with delayed bandit feedback," *Applied Mathematics and Computation*, vol. 445, p. 127865, 2023.

[14] A. K. Paul, A. D. Mahindrakar, and R. K. Kalaimani, "Distributed online mirror descent algorithm with event triggered communication," *International Federation of Automatic Control*, vol. 55, no. 30, pp. 448–453, 2022.

[15] O. Besbes, Y. Gur, and A. Zeevi, "Non-stationary stochastic optimization," *Operations Research*, vol. 63, no. 5, pp. 1227–1244, 2015.

[16] K. Oakamoto, N. Hayashi, and S. Takai, "Distributed online adaptive gradient descent with event-triggered communication," *IEEE Transactions on Control of Network Systems*, 2023, DOI:10.1109/TCNS.2023.3294432.

[17] J. Yuan and A. Lamperski, "Online convex optimization for cumulative constraints," in *Advances in Neural Information Processing Systems*, 2018, pp. 6140–6149.

[18] X. Yi, X. Li, T. Yang, L. Xie, T. Chai, and K. H. Johansson, "Regret and cumulative constraint violation analysis for online convex optimization with long term constraints," in *International Conference on Machine Learning*, 2021, pp. 11 998–12 008.

[19] ——, "Regret and cumulative constraint violation analysis for distributed online constrained convex optimization," *IEEE Transactions on Automatic Control*, vol. 68, no. 5, pp. 2875–2890, 2023.

[20] E. C. Hall and R. M. Willett, "Online convex optimization in dynamic environments," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 4, pp. 647–662, 2015.

[21] K. Zhang, X. Yi, Y. Li, M. Cao, T. Chai, and T. Yang, "Dynamic regret and cumulative constraint violation analysis for distributed online constrained convex optimization with event-triggered communication," *arXiv:2311.01957*, 2023.

[22] 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.

[23] M. Zinkevich, "Online convex programming and generalized infinitesimal gradient ascent," in *International Conference on Machine Learning*, 2003, pp. 928–936.

[24] G. S. Seyboth, D. V. Dimarogonas, and K. H. Johansson, "Event-based broadcasting for multi-agent average consensus," *Automatica*, vol. 49, no. 1, pp. 245–252, 2013.

[25] D. Yang, W. Ren, X. Liu, and W. Chen, "Decentralized event-triggered consensus for linear multi-agent systems under general directed graphss," *Automatica*, vol. 69, pp. 242–249, 2016.

[26] L. Ding, Q. Han, X. Ge, and X. Zhang, "An overview of recent advances in event-triggered consensus of multiagent systems," *IEEE Transactions on Cybernetics*, vol. 48, no. 4, pp. 1110–1123, 2017.

[27] X. Ge, Q. Han, L. Ding, Y. Wang, and X. Zhang, "Dynamic event-triggered distributed coordination control and its applications: A survey of trends and techniques," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 9, pp. 3112–3125, 2020.

[28] R. Yang, L. Liu, and G. Feng, "Cooperative output tracking of unknown heterogeneous linear systems by distributed event-triggered adaptive control," *IEEE Transactions on Cybernetics*, vol. 52, no. 1, pp. 3–15, 2022.