Distributed Event-Triggered Bandit Convex Optimization With Time-Varying Constraints

Kunpeng Zhang, Xinlei Yi, Guanghui Wen, *Senior Member, IEEE*, Ming Cao, *Fellow, IEEE*, Karl H. Johansson, *Fellow, IEEE*, Tianyou Chai, *Life Fellow, IEEE*, and Tao Yang, *Senior Member, IEEE*

***Abstract*—This article considers the distributed bandit convex optimization problem with time-varying inequality constraints over a network of agents, where the goal is to minimize network regret and cumulative constraint violation. Existing distributed online algorithms solving this problem require that each agent broadcasts its decision to its neighbors at each iteration. However, communication resources are often limited. To better utilize communication resources, we propose a distributed event-triggered online primal–dual algorithm with two-point bandit feedback. Under several classes of appropriately chosen decreasing parameter sequences and nonincreasing event-triggered threshold sequences, we establish dynamic network regret and network cumulative constraint violation bounds. These bounds are comparable to the results achieved by distributed event-triggered online algorithms with full-information feedback. Finally, a numerical example is provided to verify the theoretical results.**

***Index Terms*—Bandit convex optimization, cumulative constraint violation, distributed optimization, event-triggered algorithm, time-varying constraints.**

Received 24 December 2024; accepted 27 March 2025. Date of publication 8 April 2025; date of current version 19 September 2025. This work was supported in part by the National Natural Science Foundation of China under Grant 62133003, Grant 62325304, and Grant U22B2046, in part by the Knut and Alice Wallenberg Foundation, in part by the Swedish Foundation for Strategic Research, and in part by the Swedish Research Council. Recommended by Associate Editor D. Ghose. *(Corresponding author: Tao Yang.)*

Kunpeng Zhang, Tianyou Chai, and Tao Yang are with the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China (e-mail: 2110343@stu.neu.edu.cn; tychai@mail.neu.edu.cn; yangtao@mail.neu.edu.cn).

Xinlei Yi is with the Shanghai Institute of Intelligent Science and Technology, National Key Laboratory of Autonomous Intelligent Unmanned Systems, and Frontiers Science Center for Intelligent Autonomous Systems, Ministry of Education, Shanghai 201210, China (e-mail: xinleiyi@tongji.edu.cn).

Guanghui Wen is with the School of Automation, Southeast University, Nanjing 210096, China (e-mail: ghwen@seu.edu.cn).

Ming Cao is with the Engineering and Technology Institute Groningen, Faculty of Science and Engineering, University of Groningen, 9747 AG Groningen, The Netherlands (e-mail: m.cao@rug.nl).

Karl H. Johansson is with the Division of Decision and Control Systems, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, 10044 Stockholm, Sweden, also with the Digital Futures, KTH Royal Institute of Technology, 10044 Stockholm, Sweden (e-mail: kallej@kth.se).

Digital Object Identifier 10.1109/TCNS.2025.3558791

I. Introduction

BANDIT convex optimization has drawn a growing attention due to its broad applications, such as online routing in data networks and online advertisement placement in web search [1]. Different from online convex optimization [2], [3], [4], [5], [6], [7], [8], where the decision maker receives full-information feedback for the loss function at each iteration (i.e., the loss function is revealed to the decision maker at each iteration), in bandit convex optimization, the decision maker receives bandit feedback for the loss function at each iteration (i.e., only the values of the loss function at some points are revealed to the decision maker at each iteration). In general, regret is a common performance metric [9], which measures the difference of the cumulative loss between the decision sequence selected by the decision maker and a comparator sequence. When each element of the comparator sequence is the offline optimal static decision, this metric is called a static regret [10], [11]. When the comparator sequence is the offline optimal dynamic decision sequence, this metric is called a dynamic regret.

Bandit convex optimization with time-invariant constraints is well studied. For example, Flaxman et al. [12] proposed a projection-based online gradient descent algorithm with one-point bandit feedback and established an $\mathcal{O}(T^{3/4})$ static regret bound for convex loss functions, where T is the total number of iterations. Agarwal et al. [13] proposed a projection-based online gradient descent algorithm by introducing the notable two-point bandit feedback and established an $\mathcal{O}(\sqrt{T})$ static regret bound for convex loss functions. Mahdavi et al. [14] considered the scenarios where constraints are characterized by static inequalities and introduced the idea of long-term constraints (i.e., inequality constraints are permitted to be violated but are fulfilled in the long run) to avoid projection operations onto the inequality constrained set due to the high computational complexity. In contrast, bandit convex optimization with time-varying constraints is studied in [15] and [16], where constraints are characterized by time-varying inequalities. Different from the time-invariant constraint setting, where the decision maker knows the constrained set in advance when he/she makes a decision, in the time-varying constraint setting, the decision maker has no a priori knowledge of the current inequality constrained set, and the information is revealed along with the values of the loss function after he/she makes a decision.

The aforementioned studies concentrate on centralized online algorithms with bandit feedback, which suffer a plethora of limitations, e.g., single point of failure and heavy communication and computation overheads [17], [18]. To deal with the limitations, distributed online algorithms with bandit feedback are developed in [19], [20], [21], [22], [23], [24], and [25]. Along the line of the time-invariant constraint setting, the authors in [19] and [20] successively proposed the projection-based distributed online algorithms with two-point and one-point bandit feedback. In the presence of feedback delays, Cao and Basar [22] proposed a projection-based distributed online algorithm with two-point bandit feedback and analyzed the impact of delay size on the algorithm performance. In addition, the authors in [21], [23], and [25] considered static inequality constraints and used the idea of long-term constraints to reduce the computational burden of projection operations. For the time-varying constraint setting, Yi et al. [24] proposed a distributed online primal–dual algorithm with two-point bandit feedback by using two-point stochastic subgradient approximations for both loss and inequality constraint functions at each iteration.

Note that in the above studies on distributed online algorithms with bandit feedback, all the decision makers require to collaboratively make decisions through local information exchange with their neighbors at each iteration. However, communication resources are so limited that frequent communication among the decision makers may cause network congestion. To better utilize communication resources, Cao and Basar [26] proposed a distributed event-triggered online algorithm with two-point bandit feedback, where each agent broadcasts the current local decision to its neighbors only if the norm of the difference between the decision and its last broadcasted decision is not less than the current event-triggering threshold. Moreover, sublinear static regret is achieved when the event-triggering threshold sequence is nonincreasing and converges to zero. By using one-point and two-point stochastic subgradient estimators, two distributed event-triggered online algorithms with delayed bandit feedback are developed in [27], and static regret bounds are established for the two algorithms.

The existing studies on distributed event-triggered online algorithms with bandit feedback do not consider inequality constraints. In this context, this article studies the distributed bandit convex optimization problem with time-varying constraints, where the decision makers receive bandit feedback for both loss and inequality constraint functions at each iteration. The contributions are summarized as follows.

1) This article proposes a distributed event-triggered online primal–dual algorithm with two-point bandit feedback by integrating event-triggered communication with the distributed online algorithm in [24]. Note that the introduction of event-triggered communication causes nontrivial challenges for performance analysis, which will be explained in detail in Remark 4. The proposed algorithm can be viewed as a bandit version of the distributed event-triggered online algorithm with full-information feedback in [8]. Their proofs are significantly different, which will also be explained in detail in Remark 4. Note that the authors in [26] and [27] do not consider inequality constraints and analyze static regret, and we consider time-varying inequality constraints and analyze dynamic regret. Moreover, the proposed algorithm uses bandit feedback for inequality constraint functions.
2) When the updating step-size sequence of local primal variables is appropriately designed based on the event-triggering threshold sequence (see Theorem 1 and Corollaries 1 and 2), this article establishes dynamic network regret and network cumulative constraint violation bounds for the proposed algorithm under a nonincreasing event-triggered threshold sequence (see Theorem 1). The bounds would be sublinear if the path length of the comparator sequence (i.e., the accumulated dynamic variation of the comparator sequence) grows sublinearly and the event-triggering threshold sequence converges to zero. In addition, this article also establishes dynamic network regret and network cumulative constraint violation bounds under the event-triggering threshold sequences produced by $\tau_t = 1/t^{\theta}$ with $\theta > 0$ (see Corollary 1) and $\tau_t = 1/c^t$ with $c > 1$ (see Corollary 2), respectively. These bounds are the same as the results achieved by the distributed event-triggered online algorithm with full-information feedback in [8]. If event-triggered communication is not considered, these bounds recover the results achieved by the centralized online algorithm with two-point bandit feedback in [16]. If inequality constraints are not considered, these dynamic network regret bounds recover the results achieved by the distributed event-triggered online algorithms with two-point bandit feedback in [26] and [27]. If event-triggered communication and inequality constraints are not considered, these dynamic network regret bounds recover the results achieved by the centralized online algorithm with two-point bandit feedback in [13] and the distributed online algorithm with two-point bandit feedback in [19].
3) When the updating step-size sequence of local primal variables is independently designed (see Theorem 2), this article establishes dynamic network regret and network cumulative constraint violation bounds. These bounds are the same as the results achieved by the distributed event-triggered online algorithm with full-information feedback in [8]. If event-triggered communication is not considered, these bounds recover the results achieved by the distributed online algorithm with two-point bandit feedback in [24]. Note that this article is among the first to establish dynamic network regret bounds (and network cumulative constraint violation bounds) for distributed event-triggered bandit convex optimization (with time-varying constraints).

The detailed comparison of this article to related studies is summarized in Table I.

The rest of this article is organized as follows. Section II presents the problem formulation and motivation. Section III proposes the distributed event-triggered online primal–dual algorithm with two-point bandit feedback and analyzes its performance. Section IV demonstrates a numerical simulation to verify the theoretical results. Finally, Section V concludes

TABLE I
COMPARISON OF THIS ARTICLE TO RELATED WORKS ON BANDIT CONVEX OPTIMIZATION

Reference	Problem type	Constraint type	Information feedback	Event-triggering	Regret type
[13]	Centralized	$g_t(x) \equiv \mathbf{0}_m$	Two-point bandit feedback for f_t	No	Static regret
[15]	Centralized	$g_t(x) \leq \mathbf{0}_m$ and Slater's condition	Two-point bandit feedback for f_t, and ∇g_t	No	Dynamic regret
[16]	Centralized	$g_t(x) \leq \mathbf{0}_m$	Two-point bandit feedback for f_t and g_t	No	Dynamic regret
[19]	Distributed	$g_t(x) \equiv \mathbf{0}_m$	Two-point bandit feedback for f_t	No	Static regret
[24]	Distributed	$g_t(x) \leq \mathbf{0}_m$	Two-point bandit feedback for f_t and g_t	No	Dynamic regret
[26]	Distributed	$g_t(x) \equiv \mathbf{0}_m$	Two-point bandit feedback for f_t	Yes	Static regret
[27]	Distributed	$g_t(x) \equiv \mathbf{0}_m$	Two-point bandit feedback for f_t	Yes	Static regret
This paper	Distributed	$g_t(x) \leq \mathbf{0}_m$	Two-point bandit feedback for f_t and g_t	Yes	Dynamic regret

this article. In Appendix A, we provide some useful lemmas, and in Appendix B, the detailed proof of Theorem 1 is provided.

Notations: $\mathbb{N}_+$, $\mathbb{R}$, $\mathbb{R}^p$, and $\mathbb{R}^p_+$ denote the sets of all positive integers, real numbers, p-dimensional, and nonnegative vectors, respectively. Given $m \in \mathbb{N}_+$, $[m]$ denotes the set $\{1,\dots,m\}$. 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. $\mathrm{col}(q_1,\dots,q_n)$ denotes the concatenated column vector of $q_i \in \mathbb{R}^{m_i}$ for $i \in [n]$. $\mathbb{B}^p$ and $\mathbb{S}^p$ denote the unit ball and sphere centered around the origin in $\mathbb{R}^p$ under Euclidean norm, respectively. $\mathbf{E}$ denotes the expectation. For a set $\mathbb{K} \in \mathbb{R}^p$ and a vector $x \in \mathbb{R}^p$, $\mathcal{P}_{\mathbb{K}}(x)$ denotes the projection of the vector x onto the set $\mathbb{K}$, i.e., $\mathcal{P}_{\mathbb{K}}(x) = \arg\min_{y\in\mathbb{K}} \|x-y\|^2$, and $[x]_+$ denotes $\mathcal{P}_{\mathbb{R}^p_+}(x)$. For a scalar function $f:\mathbb{R}^p \to \mathbb{R}$, let $\partial f(x) \in \mathbb{R}^p$ denote the subgradient of f at x. For a vector-valued function $f = [f_1,\dots,f_d]^T : \mathbb{R}^p \to \mathbb{R}^d$, its subgradient at x is denoted by $\partial f(x) = [\partial f_1(x), \dots, \partial f_d(x)] \in \mathbb{R}^{p\times d}$.

II. PROBLEM FORMULATION AND MOTIVATION

Consider the distributed bandit convex optimization problem with time-varying constraints. At iteration t, a network of n agents is modeled by a time-varying directed graph $\mathcal{G}_t = (\mathcal{V}, \mathcal{E}_t)$ with the agent set $\mathcal{V} = [n]$ and the edge set $\mathcal{E}_t \subseteq \mathcal{V}\times\mathcal{V}$. $(j,i) \in \mathcal{E}_t$ indicates that agent i can receive information from agent j. The sets of in-neighbors and out-neighbors of agent i are $\mathcal{N}_i^{\mathrm{in}}(\mathcal{G}_t) = \{j \in [n] | (j,i) \in \mathcal{E}_t\}$ and $\mathcal{N}_i^{\mathrm{out}}(\mathcal{G}_t) = \{j \in [n] | (i,j) \in \mathcal{E}_t\}$, respectively. An adversary first erratically selects n local convex loss functions $\{f_{i,t} : \mathbb{X} \to \mathbb{R}\}$ and n local convex constraint functions $\{g_{i,t} : \mathbb{X} \to \mathbb{R}^{m_i}\}$ for $i \in [n]$, where $\mathbb{X} \subseteq \mathbb{R}^p$ is a known set, and both m_i and p are positive integers. Then, the agents collaborate to select their local decisions $\{x_{i,t} \in \mathbb{X}\}$ without prior access to $\{f_{i,t}\}$ and $\{g_{i,t}\}$. At the same time, the values of $f_{i,t}$ and $g_{i,t}$ at the point $x_{i,t}$ as well as at other potential points are privately revealed to agent i. The goal of the network is to choose the decision sequence $\{x_{i,t}\}$ for $i \in [n]$ and $t \in [T]$ such that both network regret

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

and network cumulative constraint violation

$$\text{Net-CCV}(\{x_{i,t}\}) := \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T} \|[g_t(x_{i,t})]_+\| \tag{2}$$

increase sublinearly, where $y_{[T]} = (y_1,\dots,y_T)$ is a comparator sequence, $f_t(x) = \frac{1}{n}\sum_{j=1}^{n} f_{j,t}(x)$ and $g_t(x) = \mathrm{col}(g_{1,t}(x), \dots, g_{n,t}(x)) \in \mathbb{R}^m$ are the global loss and constraint functions of the network at iteration t, respectively, and $m = \sum_{i=1}^{n} m_i$.

Note that the network regret (1) measures the difference of the network-wide cumulative loss between the decision sequence $\{x_{i,t}\}$ and the comparator sequence $y_{[T]}$, which is also used in [8] and [24]. This regret is different from the used regrets in [19], [26], and [27] that measure the difference of the cumulative loss between the decision sequence $\{x_{i,t}\}$ of a single agent and the comparator sequence $y_{[T]}$. Because the established bounds of the used regret metrics in [19], [26], and [27] are uniform bounds for all the agents, we can compare the established bounds of network regret in this article with those in [19], [26], and [27]. In addition, the network cumulative constraint violation (2) is stricter than the common constraint violation, as used in [15] and [16]. For details, see the explanation in [24].

In the literature, there are two commonly used comparator sequences. One is the offline optimal dynamic decision sequence $\check{x}^*_{[T]} = (\check{x}^*_1, \dots, \check{x}^*_T)$, where $\check{x}^*_t \in \mathbb{X}$ is the minimizer of $f_t(x)$ subject to $g_t(x) \leq \mathbf{0}_m$. To guarantee that the offline optimal dynamic decision sequence $\check{x}^*_{[T]}$ always exists, we assume that for any $T \in \mathbb{N}_+$, the set of all the feasible decision sequences

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

is nonempty. In this case, $\text{Net-Reg}(\{x_{i,t}\}, \check{x}^*_{[T]})$ is called the dynamic network regret. Another comparator sequence is the offline optimal static decision sequence $\hat{x}^*_{[T]} = (\hat{x}^*, \dots, \hat{x}^*)$, 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 all $t \in [T]$. To guarantee that the offline optimal static decision sequence always exists, we assume that for any $T \in N_+$, the set of all the feasible static decision sequences

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

is nonempty. In this case, $\text{Net-Reg}(\{x_{i,t}\}, \hat{x}^*_{[T]})$ is called the static network regret.

In this article, the following assumptions are made, which are commonly adopted in distributed online convex optimization, see [24], [26], [27], [28], [29], and recent survey paper [18] and the references therein.

Assumption 1:

1) The set $\mathbb{X}$ is convex and closed. Moreover, the convex set $\mathbb{X}$ contains the ball of radius $r(\mathbb{X})$ and is contained in the ball of radius $R(\mathbb{X})$, i.e.,
$$r(\mathbb{X})\mathbb{B}^p \subseteq \mathbb{X} \subseteq R(\mathbb{X})\mathbb{B}^p. \tag{5}$$
2) For all $i \in [n]$ and $t \in \mathbb{N}_+$, the local loss functions $f_{i,t}$ and the local constraint function $g_{i,t}$ are convex, and there exists a constant F_1 such that
$$|f_{i,t}(x) - f_{i,t}(y)| \leq F_1 \tag{6a}$$
$$\|g_{i,t}(x)\| \leq F_1, x, y \in \mathbb{X}. \tag{6b}$$
3) For all $i \in [n]$ and $t \in \mathbb{N}_+$, the subgradients $\partial f_{i,t}(x)$ and $\partial g_{i,t}(x)$ exist, and there exists a constant F_2 such that
$$\|\partial f_{i,t}(x)\| \leq F_2 \tag{7a}$$
$$\|\partial g_{i,t}(x)\| \leq F_2, x \in \mathbb{X}. \tag{7b}$$

Assumption 2: For $t \in \mathbb{N}_+$, the time-varying directed graph $\mathcal{G}_t$ satisfies that the following holds.

1) There exists a constant $w \in (0, 1)$ such that $[W_t]_{ij} \geq w$ if $(j, i) \in \mathcal{E}_t$ or $i = j$, and $[W_t]_{ij} = 0$ otherwise.
2) The mixing matrix W_t is doubly stochastic, i.e., $\sum_{i=1}^n [W_t]_{ij} = \sum_{j=1}^n [W_t]_{ij} = 1 \; \forall i, j \in [n]$.
3) 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.

Assumption 1 implies that the local loss functions $f_{i,t}$ and the local constraint function $g_{i,t}$ are Lipschitz continuous on $\mathbb{X}$. Assumption 2 implies that the time-varying directed graph $\mathcal{G}_t$ does not need to be connected at each iteration. In addition, if the graph $\mathcal{G}_t$ is regular at each iteration, the mixing matrix W_t is doubly stochastic. For details on constructing such a graph, see [30] and [31].

The considered problem is studied in the work of Yi et al. [24], where they propose a distributed online primal–dual algorithm with two-point bandit feedback. Note that the algorithm requires that each agent broadcasts the current decision to its neighbors through the communication network at each iteration. However, communication resources are so limited that frequent communication between the agents may cause network congestion in many practical applications, e.g., sensor networks comprised of cheap sensors with small battery capacity [26]. To better utilize communication resources, this article proposes a distributed event-triggered online primal–dual algorithm with two-point bandit feedback by integrating event-triggered communication with the algorithm in [24] and establishes network regret and cumulative constraint violation bounds for the proposed algorithm. Based on these bounds, this article also discusses the impact of event-triggered threshold on the algorithm performance.

III. Distributed Event-Triggered Online Primal–Dual Algorithm With Two-Point Bandit Feedback

This section proposes the distributed event-triggered online primal–dual algorithm with two-point bandit feedback. The proposed algorithm can be viewed as an event-triggered version of the distributed online algorithm with two-point bandit feedback in [24], or a bandit version of the distributed event-triggered online algorithm with full-information feedback in [8]. This section also establishes network regret and cumulative constraint violation bounds for the proposed algorithm.

A. Algorithm Description

This proposed algorithm is presented in pseudocode as Algorithm 1, and its architecture is shown in Fig. 1. For $t \in [T]$ with $t \geq 2$ and $i \in [n]$, same as the distributed online algorithm with two-point bandit feedback in [24], Algorithm 1 uses the distributed consensus protocol (8) to compute $z_{i,t} \in \mathbb{X}$ for agent i via the time-varying directed graph $\mathcal{G}_t$, which estimates the average value of the local decisions of all agents $\frac{1}{n}\sum_{i=1}^n x_{i,t}$. Then, Algorithm 1 uses the primal–dual protocol (9a)–(9c) to update the local primal variable $x_{i,t} \in \mathbb{X}$ and dual variable $q_{i,t} \in \mathbb{R}_+^{m_i}$, where $\hat{\omega}_{i,t}$ is the updating direction of the local primal variable $x_{i,t}$, α_t and β_t are the updating step-sizes of the local primal variable $x_{i,t}$ and the local dual variable $q_{i,t}$, respectively, and γ_t is the regularization parameter used to influence the structure of the local decisions. Different from the distributed online algorithm with two-point bandit feedback in [24], Algorithm 1 uses the event-triggering check such that agent i broadcasts its current local decision $x_{i,t}$ to its neighbors only if the norm of the difference between the decision and its last broadcasted decision $x_{i,t-1}$ is not less than the current event-triggering threshold τ_t.

Note that different from the distributed event-triggered online algorithm with full-information feedback in [8], Algorithm 1 uses the values of the local loss function $f_{i,t}$ at $x_{i,t}$ and $x_{i,t} + \delta_t u_{i,t}$ to estimate the subgradient $\partial f_{i,t}(x_{i,t})$, and uses the values of the local constraint function $g_{i,t}$ at $x_{i,t}$ and $x_{i,t} + \delta_t u_{i,t}$ to estimate the subgradient $\partial[g_{i,t}(x_{i,t})]_+$ as the subgradients are unavailable in the bandit setting. The subgradient approximations follow the two-point stochastic subgradient estimator proposed in [13], which are given by

$$\hat{\partial} f_{i,t}(x_{i,t}) = \frac{p}{\delta_t}\left(f_{i,t}(x_{i,t} + \delta_t u_{i,t}) - f_{i,t}(x_{i,t})\right) u_{i,t} \in \mathbb{R}^p$$

$$\hat{\partial}[g_{i,t}(x_{i,t})]_+ = \frac{p}{\delta_t}\left([g_{i,t}(x_{i,t} + \delta_t u_{i,t})]_+ - [g_{i,t}(x_{i,t})]_+\right)^T \otimes u_{i,t} \in \mathbb{R}^{p \times m_i}$$

where $\delta_t \in (0, r(\mathbb{X})\xi_t]$ is an exploration parameter, $r(\mathbb{X})$ is a constant given in Assumption 1, $\xi_t \in (0, 1)$ is a shrinkage coefficient, and $u_{i,t} \in \mathbb{S}^p$ is a uniformly distributed random vector.

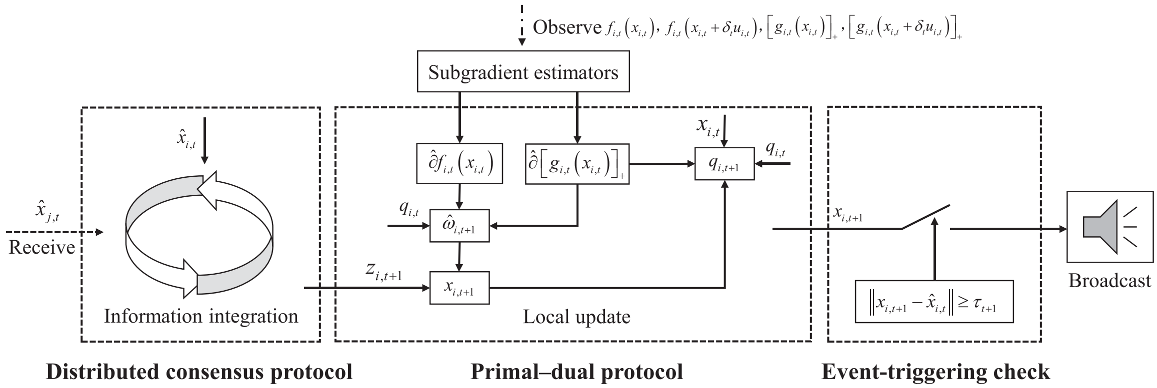


Fig. 1. Architecture of the distributed event-triggered online primal–dual algorithm with two-point bandit feedback.

Algorithm 1: Distributed Event-Triggered Online Primal–Dual Algorithm With Two-Point Bandit Feedback.

Input: constant $r(\mathbb{X})$, decreasing sequences $\{\alpha_t\} \subseteq (0,+\infty)$, $\{\beta_t\} \subseteq (0,+\infty)$, $\{\gamma_t\} \subseteq (0,+\infty)$, $\{\xi_t\} \subseteq (0,1)$, $\{\delta_t\} \subseteq (0, r(\mathbb{X})\xi_t]$, and a non-increasing sequence $\{\tau_t\} \subseteq (0,+\infty)$.

Initialize: $x_{i,1} \in (1-\xi_1)\mathbb{X}$, $\hat{x}_{i,1} = x_{i,1}$ and $q_{i,1} = \mathbf{0}_{m_i}$.
  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 $i \in [n]$.
  for $t = 1, \dots, T-1$
    for $i = 1, \dots, n$ in parallel
      Select vector $u_{i,t} \in \mathbb{S}^p$ independently and uniformly at random.
      Observe $f_{i,t}(x_{i,t})$, $f_{i,t}(x_{i,t}+\delta_t u_{i,t})$, $[g_{i,t}(x_{i,t})]_+$, and $[g_{i,t}(x_{i,t}+\delta_t u_{i,t})]_+$.
      Distributed consensus protocol:

$$z_{i,t+1} = \sum_{j=1}^{n} [W_t]_{ij} \hat{x}_{j,t}. \tag{8}$$

      Primal–dual protocol:

$$\hat{\omega}_{i,t+1} = \hat{\partial} f_{i,t}(x_{i,t}) + \hat{\partial}[g_{i,t}(x_{i,t})]_+ q_{i,t}, \tag{9a}$$

$$x_{i,t+1} = \mathcal{P}_{(1-\xi_{t+1})\mathbb{X}}(z_{i,t+1} - \alpha_{t+1}\hat{\omega}_{i,t+1}), \tag{9b}$$

$$q_{i,t+1} = \Big[(1-\beta_{t+1}\gamma_{t+1}) q_{i,t} + \gamma_{t+1}\Big([g_{i,t}(x_{i,t})]_+ + \Big(\hat{\partial}[g_{i,t}(x_{i,t})]_+\Big)^T (x_{i,t+1} - x_{i,t})\Big)\Big]_+. \tag{9c}$$

      Event-triggering check:
      if $\|x_{i,t+1} - \hat{x}_{i,t}\| \geq \tau_{t+1}$
        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: $\{x_{i,t}\}$.

B. Performance Analysis

We first appropriately design the parameter sequences for Algorithm 1 and establish dynamic network regret and network cumulative constraint violation bounds in the following theorem.

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

$$\alpha_t = \sqrt{\frac{\Psi_t}{t}}, \beta_t = \frac{1}{t^{\kappa}}, \gamma_t = \frac{1}{t^{1-\kappa}}$$

$$\xi_t = \frac{1}{t+1}, \delta_t = \frac{r(\mathbb{X})}{t+1} \quad \forall t \in \mathbb{N}_+ \tag{10}$$

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

$$\mathbf{E}[\text{Net-Reg}(\{x_{i,t}\}, y_{[T]})] = \mathcal{O}(T^{\kappa} + \sqrt{\Psi_T T} + \sqrt{\Psi_T{}^{-1} T P_T}) \tag{11}$$

$$\mathbf{E}[\text{Net-CCV}(\{x_{i,t}\})] = \mathcal{O}(T^{1-\kappa/2} + \Psi_T^{1/4} T^{3/4}) \tag{12}$$

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

The proof is given in Appendix B.

Remark 1: The updating step-size β_t of the dual variable and the regularization parameter γ_t are specifically designed to bound the term $\sum_{t=1}^{T} (\frac{1}{\gamma_t} - \frac{1}{\gamma_{t+1}} + \beta_{t+1})$ in (27) and (28) in Lemma 5, as shown in (32) in the proof of Theorem 1. In addition, unlike the distributed event-triggered online algorithm with full-information feedback in [8], Algorithm 1 utilizes two-point stochastic subgradient estimators to estimate the subgradients of local loss and constraint functions. Consequently, we must design the shrinkage coefficient ξ_t and the exploration parameter δ_t. Terms involving ξ_t and δ_t appear in (27) and (28), which are absent in [8]. These terms influence the dynamic network regret bound (11) and the network cumulative constraint violation bound (12), thereby affecting the performance of Algorithm 1.

Remark 2: The bounds (11) and (12) characterize the impact of event-triggered threshold τ_t on dynamic network regret and network cumulative constraint violation through Ψ_T. The larger the event-triggering threshold τ_t, the larger the static part of the bound (11) (i.e., $T^\kappa + \sqrt{\Psi_T T}$) and the bound (12), and the smaller the dynamic part (i.e., $\sqrt{\Psi_T{}^{-1} T P_T}$) of the bound (11).

Remark 3: Note that the dynamic network regret bound (11) is expressed in terms of the path-length P_T of the comparator sequence $y_{[T]}$. From the bound (11), we can infer that a restricted path-length P_T is required to achieve sublinear dynamic network regret bound. Specifically, the condition that the path-length P_T grows sublinearly over time is essential. In addition to the condition, sublinear dynamic network regret bound (11) and network cumulative constraint violation bound (12) would be established if the event-triggering threshold τ_t converges to zero, i.e., $\sum_{k=1}^{t} \tau_k$ grows sublinearly. The bounds (12) are the same as the results achieved by the distributed event-triggered online algorithm with full-information feedback in [8]. If event-triggered communication is not considered, i.e., $\tau_1 = 1$ and $\tau_t = 0$ for $t \in [2, T]$, these bounds recover the results achieved by the centralized online algorithm with two-point bandit feedback in [16]. If inequality constraints are not considered, i.e., given any $x \in \mathbb{X}$ and $g_{i,t}(x) \equiv \mathbf{0}_{m_i}$ for $i \in [n]$ and $t \in [T]$, the dynamic network regret bound (11) recovers the results achieved by the distributed event-triggered online algorithms with two-point bandit feedback in [26] and [27] if the path-length P_T of the comparator sequence is zero for any T, i.e., $P_T \equiv 0$ for any T. If event-triggered communication and inequality constraints are not considered, the dynamic network regret bound (11) recovers the results achieved by the centralized online algorithm with two-point bandit feedback in [13] and the distributed online algorithm with two-point bandit feedback in [19] if the path-length P_T of the comparator sequence is zero for any T. Note that we analyze dynamic regret, while the authors in [13], [19], [26], and [27] analyze static regret. However, if event-triggered communication is not considered, the network cumulative constraint violation bound (12) is larger than that achieved by the centralized online algorithm with two-point bandit feedback in [15]. This is reasonable since Chen and Giannakis [15] considered Slater's condition for constraint functions (i.e., there is a point that strictly satisfies inequality constraints), which is a sufficient condition for strong duality to hold [32]; moreover, the algorithm in [15] uses full-information feedback for constraint functions. In addition, the dynamic network regret bound (11) and the network cumulative constraint violation bound (12) do not recover the results achieved by the distributed online algorithm with two-point bandit feedback in [24] since the step-size α_t of the local primal variable is only a special case of that in [24]. If choosing $\alpha_t = 1/\sqrt{t}$ for the algorithm in [24], these bounds would recover the results achieved by the algorithm in [24].

Remark 4: The proof of Theorem 1 has substantial differences compared to that of [24, Thm. 3]. More specifically, different from the distributed bandit online algorithm without event-triggered communication in [24] where the agents broadcast the current local decisions at each iteration, in our Algorithm 1, the agents do not broadcast the current local decisions if the event-triggering condition is not satisfied. Thus, the local decision sequences produced by our Algorithm 1 are different from those produced by the algorithm in [24] although the updating rules are similar. This critical difference leads to challenges in theoretical proof because the result on the disagreement among agents cannot be directly used. To tackle these challenges, we analyze the difference between the produced local decision $x_{i,t}$ and the stored local decision $\hat{x}_{i,t}$ for running distributed consensus protocol at each iteration for agent i, $i \in [n]$, in the proof of Lemma 4. The analysis shows that the norm of the difference can be bounded by the current event-triggering threshold τ_t, i.e., $\|\hat{x}_{i,t} - x_{i,t}\| \le \tau_t$, regardless of whether the event-triggering condition is satisfied or not. The dynamic network regret bound (11) and the network cumulative constraint violation bound (12) are established in this way, and are thus subject to event-triggering threshold. In addition, the proof of Theorem 1 has significant differences compared to that of [8, Thm. 1]. Different from the distributed online algorithm in [8] where the accurate subgradients of local loss and constraint functions are directly used, in our Algorithm 1, two-point stochastic subgradient estimators are used to estimate the subgradients since they are unavailable in the bandit setting. However, there exist gaps between the estimators and the accurate subgradients although the estimators are unbiased subgradients of the uniformly smoothed versions of local loss and constraint functions. This causes that the property of subgradient, including convexity and Lipschitz continuity, cannot be directly used. To deal with this challenge, we use Lemma 1 where the relationship between the smooth functions and their original functions is established by utilizing the property of local loss and constraint functions, e.g., boundedness, convexity, and Lipschitz continuity, and reanalyze dynamic network regret and network cumulative constraint violation bounds.

Then, we consider two classes of explicit expressions for the event-triggering threshold τ_t. First, we select the event-triggering threshold sequence produced by $\tau_t = 1/t^\theta$, which is also adopted by the distributed online algorithms in [26], [27], [28], and [29]. We establish dynamic network regret and network cumulative constraint violation bounds in the following corollary.

Corollary 1: Under the same conditions as in Theorem 1 with $\tau_t = 1/t^\theta$ and $\theta > 0$, for any $T \in \mathbb{N}_+$ and any comparator sequence $y_{[T]} \in \mathcal{X}_T$, it holds that

$$\mathbf{E}[\text{Net-Reg}(\{x_{i,t}\}, y_{[T]})] = \begin{cases} \mathcal{O}(T^{\max\{\kappa, 1-\theta/2\}} + T^{\theta/2} P_T), & \text{if } 0 < \theta < 1 \\ \mathcal{O}\left(T^\kappa + \sqrt{T \log(T)} + \sqrt{\frac{T}{\log(T)}} P_T\right), & \text{if } \theta = 1 \\ \mathcal{O}(T^{\max\{\kappa, 1/2\}} + \sqrt{T} P_T), & \text{if } \theta > 1 \end{cases} \tag{13}$$

$$\mathbf{E}[\text{Net-CCV}(\{x_{i,t}\})] = \begin{cases} \mathcal{O}(T^{\max\{1-\kappa/2, 1-\theta/4\}}), & \text{if } 0 < \theta < 1 \\ \mathcal{O}(T^{1-\kappa/2} + T^{3/4} \log(T)^{1/4}), & \text{if } \theta = 1 \\ \mathcal{O}(T^{\max\{1-\kappa/2, 3/4\}}), & \text{if } \theta > 1. \end{cases} \tag{14}$$

Remark 5: The dynamic network regret bound (13) and the network cumulative constraint violation bound (14) would be sublinear if the path-length P_T grows sublinearly. Moreover, the larger the θ, the smaller the static part of the bounds (13) and (14), and the larger the dynamic part of the bound (13). If $\theta > 1$, these bounds recover the results achieved by the centralized online algorithm with two-point bandit feedback in [16]. In addition, if inequality constraints are not considered, the dynamic network regret bound (13) recovers the results achieved by the centralized online algorithm with two-point bandit feedback in [13], the distributed online algorithm with two-point bandit feedback in [19], and the distributed event-triggered online algorithm with two-point bandit feedback in [26] and [27] if the path-length P_T of the comparator sequence is zero for any T. In addition, we consider time-varying inequality constraints, while the authors in [13], [26] and [27] do not consider inequality constraints.

Second, we choose the event-triggering threshold sequence produced by $\tau_t = 1/c^t$, which is also adopted in distributed optimization with event-triggered communication; see, e.g., [33], [34], [35], [36], and [37]. We establish dynamic network regret and network cumulative constraint violation bounds in the following corollary.

Corollary 2: Under the same conditions as in Theorem 1 with $\tau_t = 1/c^t$ and $c > 1$, for any $T \in \mathbb{N}_+$ and any comparator sequence $y_{[T]} \in \mathcal{X}_T$, it holds that

$$\mathbf{E}[\text{Net-Reg}(\{x_{i,t}\}, y_{[T]})] = \mathcal{O}(T^{\max\{\kappa,1/2\}} + \sqrt{T}P_T) \quad (15)$$

$$\mathbf{E}[\text{Net-CCV}(\{x_{i,t}\})] = \mathcal{O}(T^{\max\{1-\kappa/2,3/4\}}). \quad (16)$$

Remark 6: The dynamic network regret bound (15) and the network cumulative constraint violation bound (16) recover the results in Corollary 1 with $\theta > 1$ and achieved by the centralized online algorithm with two-point bandit feedback in [16]. Moreover, if inequality constraints are not considered, the dynamic network regret bound (15) recovers the results achieved by the centralized online algorithm with two-point bandit feedback in [13], the distributed online algorithm with two-point bandit feedback in [19], and the distributed event-triggered online algorithm with two-point bandit feedback in [26] and [27] if the path-length P_T of the comparator sequence is zero for any T.

Note that the event-triggering threshold τ_t in (10) affects the updating step-size α_t of the local primal variable. To avoid that, we next show how to independently design α_t in the following theorem.

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

$$\alpha_t = \frac{\alpha_0}{t^{\theta_1}}, \beta_t = \frac{1}{t^{\theta_2}}, \gamma_t = \frac{1}{t^{1-\theta_2}}$$
$$\xi_t = \frac{1}{t+1}, \delta_t = \frac{r(\mathbb{X})}{t+1}, \tau_t = \frac{\tau_0}{t^{\theta_3}} \quad \forall t \in \mathbb{N}_+ \quad (17)$$

where $\alpha_0, \theta_1 \in (0,1), \theta_2 \in (0,1)$, and θ_3 are positive constants, and τ_0 is a nonnegative constant. Then, for any $T \in \mathbb{N}_+$ and any comparator sequence $y_{[T]} \in \mathcal{X}_T$

$$\mathbf{E}[\text{Net-Reg}(\{x_{i,t}\}, y_{[T]})]$$
$$= \begin{cases} \mathcal{O}\Big(\alpha_0 T^{1-\theta_1} + T^{\theta_2} + \frac{\tau_0}{\alpha_0}T^{1+\theta_1-\theta_3} \\ \quad + \frac{T^{\theta_1}(1+P_T)}{\alpha_0}\Big), & \text{if } \theta_1 < \theta_3 < 1+\theta_1 \\ \mathcal{O}\Big(\alpha_0 T^{1-\theta_1} + T^{\theta_2} + \frac{\tau_0}{\alpha_0}\log(T) \\ \quad + \frac{T^{\theta_1}(1+P_T)}{\alpha_0}\Big), & \text{if } \theta_3 = 1+\theta_1 \\ \mathcal{O}\Big(\alpha_0 T^{1-\theta_1} + T^{\theta_2} + \frac{\tau_0}{\alpha_0} \\ \quad + \frac{T^{\theta_1}(1+P_T)}{\alpha_0}\Big), & \text{if } \theta_3 > 1+\theta_1 \end{cases} \quad (18)$$

$$\mathbf{E}[\text{Net-CCV}(\{x_{i,t}\})]$$
$$= \begin{cases} \mathcal{O}(\sqrt{\alpha_0}T^{1-\theta_1/2} + T^{1-\theta_2/2} \\ \quad + \sqrt{\tau_0}T^{1-\theta_3/2}), & \text{if } \theta_1 < \theta_3 < 1 \\ \mathcal{O}\Big(\sqrt{\alpha_0}T^{1-\theta_1/2} + T^{1-\theta_2/2} \\ \quad + \sqrt{\tau_0 T\log(T)}\Big), & \text{if } \theta_3 = 1 \\ \mathcal{O}(\sqrt{\alpha_0}T^{1-\theta_1/2} + T^{1-\theta_2/2} \\ \quad + \sqrt{\tau_0 T}), & \text{if } \theta_3 > 1. \end{cases} \quad (19)$$

The proof is given in the online version [38] due to space limitations.

Remark 7: The bounds (18) and (19) indicate that the larger the initial value of the event-triggering threshold τ_0, the larger the dynamic network regret and network cumulative constraint violation bounds. Moreover, a larger τ_0 not only increases the event-triggering threshold τ_t, but also potentially reduces the total number of triggers.

Remark 8: The dynamic network regret bound (18) and the network cumulative constraint violation bound (19) are the same as the results achieved by the distributed event-triggered online algorithm with full-information feedback in [8], that is, in an average sense, our Algorithm 1 is as efficient as its full-information feedback version. If event-triggered communication is not considered, i.e., $\tau_0 = 0$, these bounds recover the results achieved by the distributed online algorithm with two-point bandit feedback in [24] when $\theta_1 = \theta_2$.

Remark 9: By replacing the comparator sequence $y_{[T]}$ with the offline optimal static decision sequence $\hat{x}^*_{[T]}$, we have $P_T \equiv 0$ for any T, and then the static network regret and cumulative constraint violation bounds for Algorithm 1 with corresponding parameter and event-triggered threshold sequences can be easily established based on the results in Theorems 1 and 2 and Corollaries 1 and 2, respectively. These bounds are the same as (11)–(16), (18), and (19) with $P_T = 0$, respectively. If inequality constraints are not considered, these static network regret bounds recover the results achieved by the distributed event-triggered online algorithms with two-point bandit feedback in [26] and [27]. If event-triggered communication and inequality constraints are not considered, these static network regret bounds recover the results achieved by the centralized online algorithm with two-point bandit feedback in [13] and the distributed online algorithm with two-point bandit feedback in [19].

IV. Numerical Example

To evaluate the performance of Algorithm 1, we consider a distributed online linear regression problem with time-varying

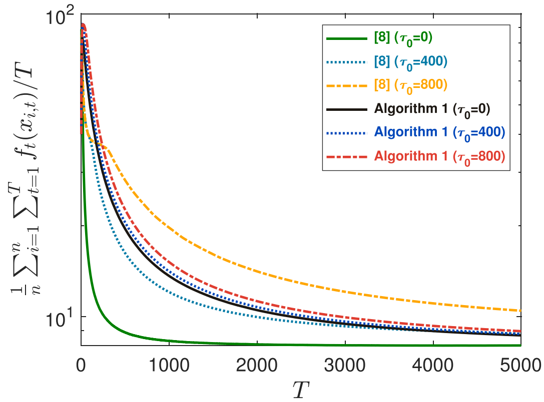


Fig. 2. Evolutions of $\frac{1}{n}\sum_{i=1}^n \sum_{t=1}^T f_t(x_{i,t})/T$.

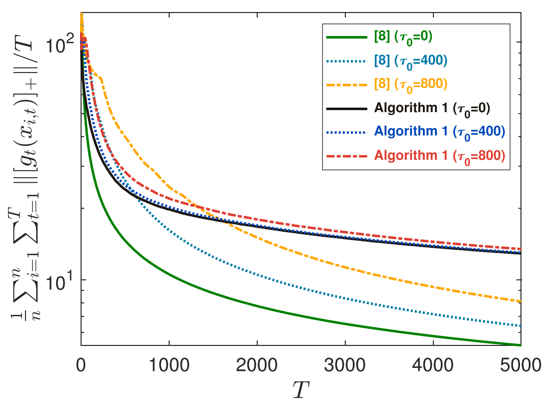


Fig. 3. Evolutions of $\frac{1}{n}\sum_{i=1}^n \sum_{t=1}^T \|[g_t(x_{i,t})]_+\|/T$.

linear inequality constraints over a network of n agents. At iteration t, the local loss and constraint functions are $f_{i,t}(x) = \frac{1}{2}(A_{i,t}x - \vartheta_{i,t})^2$ and $g_{i,t}(x) = B_{i,t}x - b_{i,t}$, respectively, where each component of $A_{i,t} \in \mathbb{R}^{q_i \times p}$ is randomly generated from the uniform distribution in the interval $[-1, 1]$, $\vartheta_{i,t} = A_{i,t}\mathbf{1}_p + \zeta_{i,t}$ with $\vartheta_{i,t} \in \mathbb{R}^{q_i}$ and $\zeta_{i,t}$ being a standard normal random vector, and each component 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 intervals $[0, 2]$ and $[0, 1]$, respectively. We set $n = 100$, $q_i = 4$, $p = 10$, $m_i = 2$, and $\mathbb{X} = [-5, 5]^p$. We use a time-varying undirected graph to model the communication topology. Specifically, at each iteration t, the graph is first randomly generated where the probability of any two agents being connected is 0.1. Then, to make sure that Assumption 2 is satisfied, we add edges $(i, i+1)$ for $i = 1, \dots, 24$ when $t \in \{4c+1\}$, edges $(i, i+1)$ for $i = 25, \dots, 49$ when $t \in \{4c+2\}$, edges $(i, i+1)$ for $i = 50, \dots, 74$ when $t \in \{4c+3\}$, and edges $(i, i+1)$ for $i = 75, \dots, 99$ when $t \in \{4c+4\}$, with c being a nonnegative integer. Moreover, let $[W_t]_{ij} = \frac{1}{n}$ if $(j, i) \in \mathcal{E}_t$ and $[W_t]_{ii} = 1 - \sum_{j=1}^n [W_t]_{ij}$. Note that there are no other distributed event-triggered online algorithms with bandit feedback to solve the considered problem due to the time-varying constraints. We compare our Algorithm 1 with the distributed event-triggered online algorithm with full-information feedback in [8].

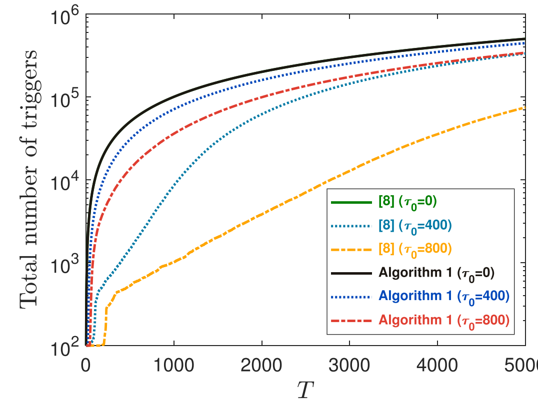


Fig. 4. Evolutions of total number of triggers.

Set $\alpha_t = 1/\sqrt{t}$, $\beta_t = 1/\sqrt{t}$, $\gamma_t = 1/\sqrt{t}$, and $\tau_t = \tau_0/t$ for our Algorithm 1 and the distributed event-triggered online algorithm in [8]. To explore the impact of different event-triggering threshold sequences on network regret and cumulative constraint violation, we select $\tau_0 = 0$, $\tau_0 = 400$, and $\tau_0 = 800$, respectively. With different values of τ_0, Figs. 2–4 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 demonstrate that as τ_0 increases, the average cumulative loss and the average cumulative constraint violation increase, while the total number of triggers decreases, which is consistent with Theorem 2. In addition, Figs. 2 and 3 also demonstrate that our Algorithm 1 has larger average cumulative loss and constraint violation than the algorithm in [8] under the same τ_0. That is reasonable, as the algorithm in [8] uses full-information feedback, while our Algorithm 1 uses two-point bandit feedback. However, the performance gap diminishes as τ_0 increases. Even when $\tau_0 = 800$, our Algorithm 1 outperforms the algorithm in [8] with a smaller average cumulative loss due to the significantly larger total number of triggers in our Algorithm 1 compared to that in the algorithm in [8], as demonstrated in Fig. 4.

V. Conclusion

This article considered the distributed bandit convex optimization problem with time-varying inequality constraints. To better utilize communication resources, we proposed the distributed event-triggered online primal–dual algorithm with two-point bandit feedback. We analyzed the network regret and cumulative constraint violation for the proposed algorithm under several classes of appropriately chosen decreasing parameter sequences and nonincreasing event-triggered threshold sequences. Our theoretical results were comparable to the results achieved by distributed event-triggered online algorithms with full-information feedback. In brief, this article broadened the applicability of distributed event-triggered online convex optimization to the regime with time-varying constraints. The future direction is to investigate compressed communication to reduce communication overhead at each iteration.

APPENDIX A
USEFUL LEMMAS

Some preliminary results are given in this section.

Lemma 1 (See [24, Lemma 8]): If Assumption 1 holds, then $\hat{f}_{i,t}(x)$ and $[\hat{g}_{i,t}(x)]_+$ are convex on $(1-\xi_t)\mathbb{X}$, and for any $i \in [n]$, $t \in \mathbb{N}_+$, $x \in (1-\xi_t)\mathbb{X}$, and $q \in \mathbb{R}^{m_i}_+$

$$\partial \hat{f}_{i,t}(x) = \mathbf{E}_{\mathfrak{U}_t}[\hat{\partial} f_{i,t}(x)] \tag{20a}$$

$$f_{i,t}(x) \le \hat{f}_{i,t}(x) \le f_{i,t}(x) + F_2\delta_t \tag{20b}$$

$$\|\hat{\partial} f_{i,t}(x)\| \le pF_2 \tag{20c}$$

$$\partial[\hat{g}_{i,t}(x)]_+ = \mathbf{E}_{\mathfrak{U}_t}\left[\hat{\partial}[g_{i,t}(x)]_+\right] \tag{20d}$$

$$\begin{aligned} q^T[g_{i,t}(x)]_+ &\le q^T[\hat{g}_{i,t}(x)]_+ \\ &\le q^T[g_{i,t}(x)]_+ + F_2\delta_t\|q\| \end{aligned} \tag{20e}$$

$$\|\hat{\partial}[g_{i,t}(x)]_+\| \le pF_2 \tag{20f}$$

$$\|[\hat{g}_{i,t}(x)]_+\| \le F_1 \tag{20g}$$

where $\hat{f}_{i,t}(x) = \mathbf{E}_{v\in\mathbb{B}^p}[f_{i,t}(x+\delta_t v)]$ and $[\hat{g}_{i,t}(x)]_+ = \mathbf{E}_{v\in\mathbb{B}^p}[[g_{i,t}(x+\delta_t v)]_+]$ with v being chosen uniformly at random, and $\mathfrak{U}_t$ is the σ-algebra induced by the independent and identically distributed variables $u_{1,t},\ldots,u_{n,t}$.

Lemma 2 (See [24, Lemma 4]): If Assumption 2 holds, then for all $i \in [n]$ and $t \in \mathbb{N}_+$, $\hat{x}_{i,t}$ generated by Algorithm 1 satisfy

$$\begin{aligned} \|\hat{x}_{i,t} - \bar{x}_t\| &\le \tau\lambda^{t-2}\sum_{j=1}^n \|\hat{x}_{j,1}\| + \frac{1}{n}\sum_{j=1}^n \|\hat{\varepsilon}^x_{j,t-1}\| + \|\hat{\varepsilon}^x_{i,t-1}\| \\ &\quad + \tau\sum_{s=1}^{t-2}\lambda^{t-s-2}\sum_{j=1}^n \|\hat{\varepsilon}^x_{j,s}\| \end{aligned} \tag{21}$$

where $\bar{x}_t = \frac{1}{n}\sum_{j=1}^n \hat{x}_{j,t}$ and $\hat{\varepsilon}^x_{i,t-1} = \hat{x}_{i,t} - z_{i,t}$.

Lemma 3 (See [24, Lemma 9]): Suppose Assumptions 1 and 2 hold, and $\gamma_t\beta_t \le 1$, $t \in \mathbb{N}_+$. For all $i \in [n]$ and $t \in \mathbb{N}_+$, the sequences $q_{i,t}$ generated by Algorithm 2 satisfy

$$\|\beta_t q_{i,t}\| \le \hat{\varpi}_1 \tag{22}$$

$$\begin{aligned} \Delta_{i,t}(\mu_i) &\le 2\hat{\varpi}_1^2\gamma_t + q^T_{i,t-1}\hat{b}_{i,t} - \mu_i^T[g_{i,t-1}(x_{i,t-1})]_+ \\ &\quad + \frac{1}{2}\beta_t\|\mu_i\|^2 + pF_2\|\mu_i\|\|x_{i,t} - x_{i,t-1}\| \end{aligned} \tag{23}$$

where $\hat{\varpi}_1 = F_1 + 2pF_2R(\mathbb{X})$, and $\hat{b}_{i,t} = [g_{i,t-1}(x_{i,t-1})]_+ + (\hat{\partial}[g_{i,t-1}(x_{i,t-1})]_+)^T(x_{i,t} - x_{i,t-1})$.

Next, we present the network regret bound at one iteration.

Note that since there exists the event-triggering check in our Algorithm 1, the local decisions and corresponding losses of the agents may be different with those of the distributed online algorithm with two-point bandit feedback in [24] although the updating rules are similar. We give a new bound for the average of network-wide loss at one iteration based on the behavior of Algorithm 1 under event-triggering check in the proof of the following lemma, which is critical to rederive the network regret bound at one iteration.

Lemma 4: Suppose Assumptions 1 and 2 hold. For all $i \in [n]$, let $\{x_{i,t}\}$ be the sequences generated by Algorithm 2 and $\{y_t\}$ be an arbitrary sequence in $\mathbb{X}$, then

$$\begin{aligned} &\frac{1}{n}\sum_{i=1}^n f_t(x_{i,t}) - f_t(y_t) \\ &\le \frac{1}{n}\sum_{i=1}^n q^T_{i,t}\left([g_{i,t}(y_t)]_+ - \mathbf{E}_{\mathfrak{U}_t}[\hat{b}_{i,t+1}]\right) - \frac{1}{n}\sum_{i=1}^n \frac{\mathbf{E}_{\mathfrak{U}_t}[\|\varepsilon^x_{i,t}\|^2]}{2\alpha_{t+1}} \\ &\quad + \frac{1}{n}\sum_{i=1}^n F_2\left(2\|\hat{x}_{i,t} - \bar{x}_t\| + p\mathbf{E}_{\mathfrak{U}_t}[\|x_{i,t} - x_{i,t+1}\|]\right) \\ &\quad + \frac{1}{2n\alpha_{t+1}}\sum_{i=1}^n \mathbf{E}_{\mathfrak{U}_t}\Big[\|\hat{y}_t - z_{i,t+1}\|^2 - \|\hat{y}_{t+1} - z_{i,t+2}\|^2 \\ &\quad + \|\hat{y}_{t+1} - \hat{x}_{i,t+1}\|^2 - \|\hat{y}_t - x_{i,t+1}\|^2\Big] \\ &\quad + \frac{1}{n}\sum_{i=1}^n F_2\left(R(\mathbb{X})\xi_t + \delta_t\right)\left(\|q_{i,t}\| + 1\right) + 2F_2\tau_t \end{aligned} \tag{24}$$

where $\mathfrak{U}_t$ is the σ-algebra induced by the independent and identically distributed variables $u_{1,t},\ldots,u_{n,t}$.

Proof: We first analyze the behavior of Algorithm 1 under event-triggering check.

In Algorithm 1, for any $t \in \mathbb{N}_+$, if $\|x_{i,t+1} - \hat{x}_{i,t}\| \ge \tau_{t+1}$, then $\|\hat{x}_{i,t+1} - x_{i,t+1}\| \le \tau_{t+1}$. If $\|x_{i,t+1} - \hat{x}_{i,t}\| < \tau_{t+1}$, then $\hat{x}_{i,t+1} = \hat{x}_{i,t}$ and we still have $\|\hat{x}_{i,t+1} - x_{i,t+1}\| \le \tau_{t+1}$. Therefore, we always have $\|\hat{x}_{i,t} - x_{i,t}\| \le \tau_t$ for any $t \ge 2$, $i \in [n]$. Recall that $\hat{x}_{i,1} = x_{i,1}$. Thus, $\|\hat{x}_{i,t} - x_{i,t}\| \le \tau_t$ for any $t \ge 1$, $i \in [n]$.

Next, we give the bound for the average of network-wide loss at one iteration.

From Assumption 1, for $i \in [n]$, $t \in \mathbb{N}_+$, and $x, y \in \mathbb{X}$, we have

$$|f_{i,t}(x) - f_{i,t}(y)| \le F_2\|x - y\| \tag{25a}$$

$$\|g_{i,t}(x) - g_{i,t}(y)\| \le F_2\|x - y\|. \tag{25b}$$

From (20b), (25a), and $\|\hat{x}_{i,t} - x_{i,t}\| \le \tau_t$, we have

$$\begin{aligned} &\frac{1}{n}\sum_{i=1}^n f_t(x_{i,t}) \\ &\le \frac{1}{n}\sum_{i=1}^n f_{i,t}(x_{i,t}) + \frac{F_2}{n^2}\sum_{i=1}^n\sum_{j=1}^n \|x_{i,t} - x_{j,t}\| \\ &\le \frac{1}{n}\sum_{i=1}^n \hat{f}_{i,t}(x_{i,t}) + \frac{2F_2}{n}\sum_{i=1}^n \|\hat{x}_{i,t} - \bar{x}_t\| + 2F_2\tau_t. \end{aligned} \tag{26}$$

It then follows from the proof of Lemma 10 of [24] that (24) holds.

Lemma 5: Suppose Assumptions 1 and 2 hold, and $\gamma_t\beta_t \le 1$, $t \in \mathbb{N}_+$. For all $i \in [n]$, let $\{x_{i,t}\}$ be the sequences generated by Algorithm 2. Then, for any comparator sequence $y_{[T]} \in \mathcal{X}_T$

$$\mathbf{E}[\text{Net-Reg}(\{x_{i,t}\}, y_{[T]})]$$

$$\leq 2(p+1)F_2\varpi_2 + 2\hat{\varpi}_1^2\sum_{t=1}^{T}\gamma_t + 10\hat{\varpi}_3\sum_{t=1}^{T}\alpha_t + F_2(\hat{\varpi}_2+2)\sum_{t=1}^{T}\tau_t + 2R(\mathbb{X})\sum_{t=1}^{T}\frac{\tau_{t+1}}{\alpha_{t+1}} + \frac{2R(\mathbb{X})^2}{\alpha_{T+1}} + \frac{2R(\mathbb{X})}{\alpha_T}P_T + \sum_{t=1}^{T}F_2(R(\mathbb{X})\xi_t+\delta_t)\left(\frac{\hat{\varpi}_1}{\beta_t}+1\right) + 2R(\mathbb{X})^2\sum_{t=1}^{T}\frac{\xi_t-\xi_{t+1}}{\alpha_{t+1}} - \frac{1}{2n}\sum_{t=1}^{T}\sum_{i=1}^{n}\left(\frac{1}{\gamma_t}-\frac{1}{\gamma_{t+1}}+\beta_{t+1}\right)\mathbf{E}[\|q_{i,t}\|^2] \quad (27)$$

$$\mathbf{E}\left[\frac{1}{n}\sum_{i=1}^{n}\left\|\sum_{t=1}^{T}[g_t(x_{i,t})]_+\right\|^2\right] \leq 4(p+1)nF_1F_2\varpi_2T + 2nF_1F_2(\hat{\varpi}_2+2)T\sum_{t=1}^{T}\tau_t + 2\left(\frac{1}{\gamma_1}+\sum_{t=1}^{T}(\beta_t+40\hat{\varpi}_3\alpha_t)\right)\left(nF_1T + 2(p+1)nF_2\varpi_2 + 2n\hat{\varpi}_1^2\sum_{t=1}^{T}\gamma_t + 20n\hat{\varpi}_3\sum_{t=1}^{T}\alpha_t + nF_2(\hat{\varpi}_2+2)\sum_{t=1}^{T}\tau_t + 2nR(\mathbb{X})\sum_{t=1}^{T}\frac{\tau_{t+1}}{\alpha_{t+1}} + \frac{2nR(\mathbb{X})^2}{\alpha_{T+1}} + 2nR(\mathbb{X})^2\sum_{t=1}^{T}\frac{\xi_t-\xi_{t+1}}{\alpha_{t+1}} + \sum_{t=1}^{T}nF_2(R(\mathbb{X})\xi_t+\delta_t)(\frac{\hat{\varpi}_1}{\beta_t}+1) - \frac{1}{2}\sum_{t=1}^{T}\sum_{i=1}^{n}(\frac{1}{\gamma_t}-\frac{1}{\gamma_{t+1}}+\beta_{t+1})\|q_{i,t}-\hat{\mu}_{ij}^0\|^2\right) \quad (28)$$

where $\hat{\varpi}_2 = 2\varpi_4 + 2p\varpi_4 + p$, $\hat{\varpi}_3 = 2F_2\varpi_3 + p^2\varpi_5$, and $\hat{\mu}_{ij}^0 = \frac{\sum_{t=1}^{T}[g_{i,t}(x_{j,t})]_+}{\frac{1}{\gamma_1}+\sum_{t=1}^{T}(\beta_t+40\hat{\varpi}_3\alpha_t)}$.

The proof is given in the online version [38] due to space limitations.

Appendix B
Proof of Theorem 1

Based on Lemma 5, we are now ready to prove Theorem 1.

1) For any constant $a \in [0,1)$ and $T \in \mathbb{N}_+$, it holds that

$$\sum_{t=1}^{T}\frac{1}{t^a} \leq 1 + \int_1^T \frac{1}{t^a}dt = \frac{T^{1-a}-a}{1-a} \leq \frac{T^{1-a}}{1-a}. \quad (29)$$

Form (29), we have

$$\sum_{t=1}^{T}\sqrt{\frac{\Psi_t}{t}} \leq \sqrt{\Psi_T}\sum_{t=1}^{T}\frac{1}{\sqrt{t}} \leq 2\sqrt{T\Psi_T}. \quad (30)$$

From Cauchy–Schwarz inequality, we have

$$\sum_{t=1}^{T}\frac{\tau_{t+1}}{\sqrt{\frac{\Psi_{t+1}}{t+1}}} \leq \sum_{t=1}^{T}\sqrt{\tau_{t+1}} \leq \sum_{t=1}^{T}\sqrt{\tau_t} \leq \sqrt{T\Psi_T}. \quad (31)$$

From (10), we have

$$\frac{t}{t^\kappa} - \frac{t+1}{(t+1)^\kappa} + \frac{1}{(t+1)^\kappa} = \frac{t}{t^\kappa} - \frac{t}{(t+1)^\kappa} > 0. \quad (32)$$

For any $T \in \mathbb{N}_+$, there exists a constant $H > 0$ such that

$$\sum_{t=1}^{T}\left(\frac{1}{t+1}-\frac{1}{t+2}\right)\sqrt{\frac{t+1}{\Psi_{t+1}}} \leq H\sqrt{\frac{1}{\Psi_2}}. \quad (33)$$

Combining (10), (27), and (29)–(33) yields

$$\mathbf{E}[\text{Net-Reg}(\{x_{i,t}\}, y_{[T]})] \leq 2(p+1)F_2\varpi_2 + \frac{2\hat{\varpi}_1^2}{\kappa}T^\kappa + 20\hat{\varpi}_3\sqrt{T\Psi_T} + F_2(\hat{\varpi}_2+2)\Psi_T + 2R(\mathbb{X})\sqrt{T\Psi_T} + 2\sqrt{2}R(\mathbb{X})^2\sqrt{\frac{T}{\Psi_T}} + F_2\left(R(\mathbb{X})+r(\mathbb{X})\right)\left(\frac{\hat{\varpi}_1}{\kappa}T^\kappa + \log(T)\right) + 2HR(\mathbb{X})^2\sqrt{\frac{1}{\Psi_2}} + 2R(\mathbb{X})\sqrt{\frac{T}{\Psi_T}}P_T \quad (34)$$

which gives (11).

2) Combining (10) and (28)–(33) yields

$$\mathbf{E}\left[\frac{1}{n}\sum_{i=1}^{n}\|\sum_{t=1}^{T}[g_t(x_{i,t})]_+\|^2\right] \leq 4(p+1)nF_1F_2\varpi_2T + 2nF_1F_2(\hat{\varpi}_2+2)T\Psi_T + 2\left(\frac{1}{\gamma_1}+\frac{T^{1-\kappa}}{1-\kappa}+80\hat{\varpi}_3\sqrt{T\Psi_T}\right)\left(nF_1T + 2(p+1)nF_2\varpi_2 + \frac{2n\hat{\varpi}_1^2}{\kappa}T^\kappa + 40n\hat{\varpi}_3\sqrt{T\Psi_T} + nF_2(\hat{\varpi}_2+2)\Psi_T + 2nR(\mathbb{X})\sqrt{T\Psi_T} + 2\sqrt{2}nR(\mathbb{X})^2\sqrt{\frac{T}{\Psi_T}} + 2nHR(\mathbb{X})^2\sqrt{\frac{1}{\Psi_2}} + nF_2\left(R(\mathbb{X})+r(\mathbb{X})\right)\left(\frac{\hat{\varpi}_1}{\kappa}T^\kappa+\log(T)\right)\right). \quad (35)$$

From Cauchy–Schwarz inequality, we have

$$\left(\frac{1}{n}\sum_{i=1}^{n}\|\sum_{t=1}^{T}[g_t(x_{i,t})]_+\|\right)^2 \leq \frac{1}{n}\sum_{i=1}^{n}\|\sum_{t=1}^{T}[g_t(x_{i,t})]_+\|^2. \quad (36)$$

We have

$$\sum_{t=1}^{T}\|[g_t(x_{i,t})]_+\| \leq \sqrt{m}\|\sum_{t=1}^{T}[g_t(x_{i,t})]_+\|. \quad (37)$$

Combining (35)–(37) yields (12).

References

[1] E. Hazan, "Introduction to online convex optimization," *Found. Trends Optim.*, vol. 2, no. 3–4, pp. 157–325, 2016.

[2] M. Zinkevich, "Online convex programming and generalized infinitesimal gradient ascent," in *Proc. Int. Conf. Mach. Learn.*, 2003, pp. 928–936.

[3] E. Hazan, A. Agarwal, and S. Kale, "Logarithmic regret algorithms for online convex optimization," *Mach. Learn.*, vol. 69, no. 2-3, pp. 169–192, 2007.

[4] A. Mokhtari, S. Shahrampour, A. Jadbabaie, and A. Ribeiro, "Online optimization in dynamic environments: Improved regret rates for strongly convex problems," in *Proc. IEEE Conf. Decis. Control*, 2016, pp. 7195–7201.

[5] J. Yuan and A. Lamperski, "Online convex optimization for cumulative constraints," in *Proc. Int. Conf. Adv. Neural Inf. Process. Syst.*, 2018, pp. 6140–6149.

[6] L. Zhang, S. Lu, and Z. Zhou, "Adaptive online learning in dynamic environments," in *Proc. Int. Conf. Adv. Neural Inf. Process. Syst.*, 2018, pp. 1323–1333.

[7] H. Guo, X. Liu, H. Wei, and L. Ying, "Online convex optimization with hard constraints: Towards the best of two worlds and beyond," in *Proc. Int. Conf. Adv. Neural Inf. Process. Syst.*, 2022, pp. 6426–6439.

[8] K. Zhang, X. Yi, Y. Li, M. Cao, T. Chai, and T. Yang, "Distributed online constrained convex optimization with event-triggered communication," *Eur. J. Control*, vol. 80, 2024, Art. no. 101042.

[9] S. Shalev-Shwartz, "Online learning and online convex optimization," *Found. Trends Mach. Learn.*, vol. 4, no. 2, pp. 107–194, 2012.

[10] T. Yang, L. Zhang, R. Jin, and J. Yi, "Tracking slowly moving clairvoyant: Optimal dynamic regret of online learning with true and noisy gradient," in *Proc. Int. Conf. Mach. Learn.*, 2016, pp. 449–457.

[11] 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 Trans. Autom. Control*, vol. 66, no. 10, pp. 4620–4635, Oct. 2021.

[12] A. D. Flaxman, A. T. Kalai, and H. B. McMahan, "Online convex optimization in the bandit setting: Gradient descent without a gradient," in *Proc. 16th Annu. ACM-SIAM Symp. Discrete Algorithms*, 2005, pp. 385–394.

[13] A. Agarwal, O. Dekel, and L. Xiao, "Optimal algorithms for online convex optimization with multi-point bandit feedback," in *Proc. Conf. Learn. Theory*, 2010, pp. 28–40.

[14] M. Mahdavi, R. Jin, and T. Yang, "Trading regret for efficiency: Online convex optimization with long term constraints," *J. Mach. Learn. Res.*, vol. 13, no. 1, pp. 2503–2528, 2012.

[15] T. Chen and G. B. Giannakis, "Bandit convex optimization for scalable and dynamic IoT management," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 1276–1286, Feb. 2019.

[16] X. Cao and K. R. Liu, "Online convex optimization with time-varying constraints and bandit feedback," *IEEE Trans. Autom. Control*, vol. 64, no. 7, pp. 2665–2680, Jul. 2019.

[17] T. Yang et al., "A survey of distributed optimization," *Annu. Rev. Control*, vol. 47, pp. 278–305, 2019.

[18] X. Li, L. Xie, and N. Li, "A survey on distributed online optimization and online games," *Annu. Rev. Control*, vol. 56, 2023, Art. no. 100904.

[19] D. Yuan, Y. Hong, D. W. C. Ho, and S. Xu, "Distributed mirror descent for online composite optimization," *IEEE Trans. Autom. Control*, vol. 66, no. 2, pp. 714–729, Feb. 2021.

[20] D. Yuan, B. Zhang, D. W. Ho, W. X. Zheng, and S. Xu, "Distributed online bandit optimization under random quantization," *Automatica*, vol. 146, 2022, Art. no. 110590.

[21] D. Yuan, A. Proutiere, and G. Shi, "Distributed online optimization with long-term constraints," *IEEE Trans. Autom. Control*, vol. 67, no. 3, pp. 1089–1104, Mar. 2022.

[22] X. Cao and T. Başar, "Decentralized online convex optimization with feedback delays," *IEEE Trans. Autom. Control*, vol. 67, no. 6, pp. 2889–2904, Jun. 2022.

[23] X. Cao and T. Başar, "Distributed constrained online convex optimization over multiple access fading channels," *IEEE Trans. Signal Process.*, vol. 70, pp. 3468–3483, 2022.

[24] X. Yi, X. Li, T. Yang, L. Xie, T. Chai, and K. H. Johansson, "Regret and cumulative constraint violation analysis for distributed online constrained convex optimization," *IEEE Trans. Autom. Control*, vol. 68, no. 5, pp. 2875–2890, May 2023.

[25] W. Yan and X. Cao, "Decentralized multi-task online convex optimization under random link failures," *IEEE Trans. Signal Process.*, vol. 72, pp. 622–635, 2024.

[26] X. Cao and T. Başar, "Decentralized online convex optimization with event-triggered communications," *IEEE Trans. Signal Process.*, vol. 69, pp. 284–299, 2021.

[27] M. Xiong, B. Zhang, D. Yuan, Y. Zhang, and J. Chen, "Event-triggered distributed online convex optimization with delayed bandit feedback," *Appl. Math. Comput.*, vol. 445, 2023, Art. no. 127865.

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

[29] K. Oakamoto, N. Hayashi, and S. Takai, "Distributed online adaptive gradient descent with event-triggered communication," *IEEE Trans. Control Netw. Syst.*, vol. 11, no. 2, pp. 610–622, Jun. 2024.

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

[31] A. Nedić and A. Olshevsky, "Distributed optimization over time-varying directed graphs," *IEEE Trans. Autom. Control*, vol. 60, no. 3, pp. 601–615, Mar. 2015.

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

[33] 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.

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

[35] L. Ding, Q. Han, X. Ge, and X. Zhang, "An overview of recent advances in event-triggered consensus of multiagent systems," *IEEE Trans. Cybern.*, vol. 48, no. 4, pp. 1110–1123, Apr. 2018.

[36] 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 Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 9, pp. 3112–3125, Sep. 2020.

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

[38] K. Zhang et al., "Distributed event-triggered bandit convex optimization with time-varying constraints," 2024, *arXiv:2406.14060*.

Kunpeng Zhang received the B.Eng. and M.Eng. degrees in automation and control engineering from Sichuan University, Chengdu, China, in 2018 and 2021, respectively. He is currently working toward the Ph.D. degree in control science and engineering with the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang, China.

His research interests include online optimization and distributed optimization.

Xinlei Yi received the B.S. degree in mathematics from the China University of Geoscience, Wuhan, China, in 2011, the M.S. degree in mathematics from Fudan University, Shanghai, China, in 2014, and the Ph.D. degree in electrical engineering from the KTH Royal Institute of Technology, Stockholm, Sweden, in 2020.

He was a Postdoc with the KTH Royal Institute of Technology from 2020 to 2022 and with the Lab for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA, USA, from 2022 to 2024. He is a tenure-track Professor with the Shanghai Institute of Intelligent Science and Technology, Tongji University, Shanghai. His current research interests include distributed online and optimization, online optimization, meta-learning, and graph neural networks.

Dr. Yi was selected as one of the four finalists for the 2021 European Systems and Control Ph.D. Thesis Award.

Guanghui Wen (Senior Member, IEEE) received the Ph.D. degree in mechanical systems and control from Peking University, Beijing, China, in 2012.

He is currently an Endowed Chair Professor and the Vice Dean with the School of Automation, Southeast University, Nanjing, China. His current research interests include autonomous intelligent systems, complex networked systems, distributed control and optimization, resilient control, and distributed reinforcement learning.

Dr. Wen was the recipient of the National Science Fund for Distinguished Young Scholars, the China Youth Science and Technology Award, the Australian Research Council Discovery Early Career Researcher Award, and the Asia Pacific Neural Network Society Young Researcher Award. He is currently a Technical Editor of IEEE/ASME Transactions on Mechatronics and an Associate Editor for IEEE Transactions on Control of Network Systems, IEEE Transactions on Industrial Informatics, IEEE Transactions on Neural Networks and Learning Systems, IEEE Transactions on Intelligent Vehicles, IEEE Transactions on Fuzzy Systems, IEEE Transactions on Systems, Man and Cybernetics: Systems, IEEE Journal of Emerging and Selected Topics in Industrial Electronics, IEEE Open Journal of the Industrial Electronics Society, and *Asian Journal of Control*. Prof. Wen has been named a Highly Cited Researcher by Clarivate Analytics since 2018. He is an IET Fellow.

Ming Cao (Fellow, IEEE) received the bachelor's and master degrees from Tsinghua University, Beijing, China, in 1999 and 2002, and the Ph.D. degree from Yale University, New Haven, CT, USA, in 2007.

From 2007 to 2008, he was a Research Associate with Princeton University, Princeton, NJ, USA. In 2006, he worked as a Research Intern with the IBM T.J. Watson Research Center, Yorktown Heights, NY, USA. Since 2016, he has been a Professor of networks and robotics with the Engineering and Technology Institute, University of Groningen, Groningen, The Netherlands, where he started as an Assistant Professor in 2008. His research interests include autonomous robots and multiagent systems, complex networks, and decision-making processes.

Dr. Cao was the 2017 and inaugural recipient of the Manfred Thoma medal from the International Federation of Automatic Control and the 2016 recipient of the European Control Award sponsored by the European Control Association. He is a Senior Editor for *Systems and Control Letters*, an Associate Editor for IEEE Transactions on Automatic Control, IEEE Transaction of Control of Network Systems, and *IEEE Robotics & Automation Magazine*.

Karl H. Johansson (Fellow, IEEE) received the M.Sc. degree in electrical engineering and the Ph.D. degree in automatic control from Lund University, Lund, Sweden, in 1992 and 1997, respectively.

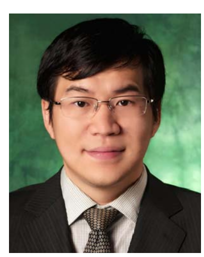
He is a Swedish Research Council Distinguished Professor in electrical engineering and computer science with the KTH Royal Institute of Technology, Stockholm, Sweden, and the Founding Director of Digital Futures. He has held visiting positions with UC Berkeley, Caltech, NTU, and other prestigious institutions. His research interests include networked control systems and cyber-physical systems with applications in transportation, energy, and automation networks.

Dr. Johansson was the recipient of numerous best paper awards and various other distinctions from IEEE, IFAC, and other organizations, for his scientific contributions. He was also the recipient of the triennial IFAC Young Author Prize, IEEE CSS Distinguished Lecturer, IFAC Outstanding Service Award, and IEEE CSS Hendrik W. Bode Lecture Prize. He has been awarded Distinguished Professor by the Swedish Research Council, Wallenberg Scholar by the Knut and Alice Wallenberg Foundation, and Future Research Leader by the Swedish Foundation for Strategic Research. His extensive service to the academic community includes being President of the European Control Association, IEEE CSS Vice President Diversity, Outreach and Development, and Member of IEEE CSS Board of Governors and IFAC Council. He has served on the editorial boards of *Automatica*, IEEE Transactions on Automatic Control, IEEE Transactions on Control of Network Systems, and many other journals. He has also been a Member of the Swedish Scientific Council for Natural Sciences and Engineering Sciences. He is Fellow of the Royal Swedish Academy of Engineering Sciences.

Tianyou Chai (Life Fellow, IEEE) received the Ph.D. degree in control theory and engineering from Northeastern University, Shenyang, China, in 1985.

In 1988, he became a Professor with Northeastern University. He is the Founder and Director with the Center of Automation, which became a National Engineering and Technology Research Center and a State Key Laboratory. From 2010 to 2018, he was the Director with the Department of Information Science, National Natural Science Foundation of China. He has authored or coauthored 356 peer reviewed international journal papers. His paper titled *Hybrid Intelligent Control for Optimal Operation of Shaft Furnace Roasting Process* was selected as one of three best papers for the Control Engineering Practice Paper Prize for 2011–2013. He has developed control technologies with applications to various industrial processes. His current research interests include modeling, control, optimization, and integrated automation and intelligence of complex industrial processes.

Dr. Chai was the recipient of five prestigious awards of National Natural Science, National Science and Technology Progress, and National Technological Innovation, the 2007 Industry Award for Excellence in Transitional Control Research from IEEE Multiple-Conference on Systems and Control, and the 2017 Wook Hyun Kwon Education Award from Asian Control Association, for his contributions. He is a Member of Chinese Academy of Engineering and IFAC Fellow.



Tao Yang (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from Washington State University, Pullman, WA, USA, in 2012.

From August 2012 to August 2014, he was an ACCESS Postdoctoral Researcher with the ACCESS Linnaeus Centre, Royal Institute of Technology, Stockholm, Sweden. He then joined the Pacific Northwest National Laboratory as a Postdoc, and was promoted to Scientist/Engineer II in 2015. From 2016 to 2019, he was an Assistant Professor with the Department of Electrical Engineering, University of North Texas, Denton, TX, USA. He is currently a Professor with the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang, China. His research interests include industrial artificial intelligence, integrated optimization and control, distributed control and optimization with applications to process industries, cyber physical systems, networked control systems, and multiagent systems.

Dr. Yang is an Associate Editor for IEEE Transactions on Control of Network Systems, IEEE Transactions on Control Systems Technology, and IEEE Transactions on Neural Networks and Learning System. He was the recipient of the Ralph E. Powe Junior Faculty Enhancement Award and Best Student Paper award (as an Advisor) of several international conferences.