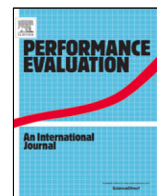




Contents lists available at ScienceDirect

## Performance Evaluation

journal homepage: [www.elsevier.com/locate/peva](http://www.elsevier.com/locate/peva)

# Simultaneously achieving sublinear regret and constraint violations for online convex optimization with time-varying constraints<sup>☆</sup>

Qingsong Liu<sup>a</sup>, Wenfei Wu<sup>b</sup>, Longbo Huang<sup>a</sup>, Zhixuan Fang<sup>a,c,\*</sup>

<sup>a</sup> Institute for Interdisciplinary Information Sciences, Tsinghua University, China

<sup>b</sup> School of Electronics Engineering and Computer Science, Peking University, China

<sup>c</sup> Shanghai Qj Zhi Institute, Shanghai, China

## ARTICLE INFO

### Article history:

Available online 6 October 2021

### Keywords:

Constrained optimization  
Online convex optimization  
Online network resource allocation  
Online job scheduling

## ABSTRACT

In this paper, we develop a novel virtual-queue-based online algorithm for online convex optimization (OCO) problems with long-term and time-varying constraints and conduct a performance analysis with respect to the dynamic regret and constraint violations. We design a new update rule of dual variables and a new way of incorporating time-varying constraint functions into the dual variables. To the best of our knowledge, our algorithm is the first parameter-free algorithm to simultaneously achieve sublinear dynamic regret and constraint violations. Our proposed algorithm also outperforms the state-of-the-art results in many aspects, e.g., our algorithm does not require the Slater condition. Meanwhile, for a group of practical and widely-studied constrained OCO problems in which the variation of consecutive constraints is smooth enough across time, our algorithm achieves  $O(1)$  constraint violations. Furthermore, we extend our algorithm and analysis to the case when the time horizon  $T$  is unknown. Finally, numerical experiments are conducted to validate the theoretical guarantees of our algorithm, and some applications of our proposed framework will be outlined.

© 2021 Elsevier B.V. All rights reserved.

## 1. Introduction

Online Convex Optimization (OCO) with long-term constraints has become one of the most popular online learning frameworks in recent years due to its powerful modeling capability for various problems such as network routing [1], online display advertising [2], and resources management [3]. In the formulation of OCO with long-term constraints, the agent wants to minimize the accumulated loss while satisfying the constraints as much as possible in the long-term. Most existing works consider the scenarios where the constraints are time-invariant [4,5]. However, *time-varying constraints* arise in many practical applications in which the underlying time-varying system is dynamic and uncertain, e.g., smart grid with uncertain renewable energy supply [6] and data centers with dynamic user demands [7]. Thus, this paper considers recently proposed *OCO framework with long-term and time-varying constraints* [8–10], which is more general and practical than the one with time-variant constraints setting.

<sup>☆</sup> The work of Longbo Huang is supported in part by the Technology and Innovation Major Project of the Ministry of Science and Technology of China under Grant 2020AAA0108400 and 2020AAA0108403.

\* Corresponding author at: Institute for Interdisciplinary Information Sciences, Tsinghua University, China.

E-mail addresses: [liu-qs19@mails.tsinghua.edu.cn](mailto:liu-qs19@mails.tsinghua.edu.cn) (Q. Liu), [wenfeiwu@pku.edu.cn](mailto:wenfeiwu@pku.edu.cn) (W. Wu), [longbohuang@tsinghua.edu.cn](mailto:longbohuang@tsinghua.edu.cn) (L. Huang), [zfang@mail.tsinghua.edu.cn](mailto:zfang@mail.tsinghua.edu.cn) (Z. Fang).

**Table 1**  
Comparison of performance bounds for OCO with long-term and time-varying constraints w.r.t. dynamic benchmark.

Reference	Regret (R)	Constraint violations (C)	Parameter-free	Slater condition-free <sup>a</sup>	Simultaneous sublinear R&C
[3]	$O(\max\{V_x T^a, V_g T^a, T^{1-a}\})$	$O(T^{1-a})$	✓	×	×
[11]	$O(T^{\frac{7}{8}} V_x)$	$O(\max\{T^{\frac{15}{16}}, T^{\frac{7}{8}} V_x\})$	✓	✓	×
[9]	$O(V_x T^{\frac{3}{4}})$	$O(T^{\frac{3}{4}})$	✓	×	×
[9]	$O(V_x T^{\frac{1}{2}})$	$O(T^{\frac{1}{2}})$	✓	×	×
[8]	$O(V_x^{\frac{1}{2}} T^{\frac{1}{2}})$	$O(V_x^{\frac{1}{4}} T^{\frac{3}{4}})$	×	×	✓
Thm.1	$O(\max\{\sqrt{TV_x}, V_g\})$	$O(\max\{\sqrt{T}, V_g\})$	✓	✓	✓
Thm.1	$O(\sqrt{TV_x})$	$O(\max\{T^{\frac{3}{4}}, V_g\})$	✓	✓	✓

<sup>a</sup>[8] assumes a slightly stronger Slater condition.

### 1.1. Prior work

OCO with long-term and time-invariant constraints has been extensively studied in the past few years. This branch of literature usually focuses on the minimization of the static regret. [12] first studied OCO with long-term and time-invariant constraints and developed an online algorithm with sublinear static regret and accumulated constraint violations. Later, [13,14] improved the performance bounds in [12]. These bounds are further improved in the recent work [4,5], where a state-of-the-art static regret and constraint violations upper bounds are shown under the assumption of the Slater condition. However, the setting of time-invariant constraints means the constraints will be learned by the agent easily, and hence does not capture the scenarios in which the underlying environment is dynamic and uncertain.

**The Time-varying constraints.** To overcome the limitations above, recent advances in OCO with long-term constraints considered the time-varying constraints and usually adopt a more practical but challenging metric, the dynamic regret. In this setting, a crucial challenge is to achieve sublinear dynamic regret and constraint violation simultaneously. [8] studied OCO with long-term and time-varying constraints both in full-information setting and bandit setting with two-point feedback. It is the first work to simultaneously achieve sublinear dynamic regret and constraint violations. But the performance bounds attained in [8] are only valid when the order of the accumulated variations of the environment is known to the agent in advance, i.e., parameter-dependent. For parameter-free results, [3] analyzed the performance of a modified online saddle-point (MOSP) method and showed that sublinear dynamic regret and constraint violation may be achieved if the accumulated variations of the environment are sublinear. Later [11] improves upon it in terms of fewer assumptions but incurs a degradation of the performance. [9] proposed a variant of MOSP method for bandit setting with two-point feedback and established the state-of-the-art performance upper bounds. However, all these parameter-free methods do not always guarantee the sublinear regret and constraint violations simultaneously, even given the accumulated variations of the environment is sublinear. Besides, most of them assume the Slater condition holds while it is not true in many scenarios. We list these works in Table 1.

Most related to our work is [4,5], which developed virtual-queue-based online algorithms and achieved the best performance bounds on static regret for the time-invariant constraints setting and the time-varying constraints setting, respectively. These results provide an inspiring insight for OCO with long-term constraints. However, a challenging question remains *if a virtual-queue-based algorithm can improve the state-of-the-art performance on OCO with long-term and time-varying constraints in terms of dynamic regret, and achieve sublinear regret and constraint violations simultaneously under only common assumptions*. The answer is yes and our main contributions are summarized in the following part.

### 1.2. Contributions

We summarize our main contributions as follows.

- We develop and analyze a novel *parameter-free* virtual-queue-based algorithm for OCO with long-term and time-varying constraints. Specifically, we prove that our algorithm achieves sublinear dynamic regret and constraint violations *simultaneously* without the Slater condition. The dynamic regret and constraint violations bounds of our developed algorithm outperform the state-of-the-art in many aspects. See also Table 1 for details.
- We show that when the variation of consecutive constraints is smooth enough across time, which holds in many practical applications [3], our algorithm can achieve  $O(1)$  constraint violations.
- To the best of our knowledge, we are the first to consider the unknown time horizon case for OCO with long-term and time-varying constraints. Furthermore, our algorithm with a doubling trick can still preserve the order of performance bounds when the time horizon is unknown.
- We outline some examples of applications, and fit them in the framework of OCO with long-term and time-varying constraints.

## 2. Problem setup

In this section, we first introduce the formulation of OCO problems with long-term and time-varying constraints. Then, we present the assumptions in our paper, which are widely-adopted.

### 2.1. Formulation

In each round  $t$ , the agent incurs a loss function  $f_t$  and a constraint requirement  $\mathbf{g}_t$ , i.e., the agent wants to make a decision  $x_t \in \chi$  to minimize the loss  $f_t(x_t)$  while satisfying  $\mathbf{g}_t(x_t) \leq \mathbf{0}$ , where  $\mathbf{g}_t(x)$  is defined as  $[g_{t,1}(x), g_{t,2}(x), \dots, g_{t,K}(x)]^T$ . In this paper, we assume that  $f_t(x)$  and  $g_{t,i}(x)$  are defined over a closed convex set  $\chi \subseteq \mathbb{R}^n$ . Denote  $\{f_t(x)\}_{t=1}^\infty$  and  $\{\mathbf{g}_t(x)\}_{t=1}^\infty$  as the sequence of the time-varying loss functions and constraint functions, respectively. Thus, the agent's goal is to compute the  $x_t^*$  defined as follows:

$$x_t^* = \arg \min_{x \in \chi} \{f_t(x) | \mathbf{g}_t(x) \leq \mathbf{0}\}.$$

However, solving this problem is challenging in the online setting since the information about the loss and constraint functions is unknown a priori to the agent. In particular, since  $\mathbf{g}_t$  is unknown a priori, the constraint  $\mathbf{g}_t(x_t) \leq \mathbf{0}$  is hard to be satisfied in every time slot  $t$ . Rather, previous work [3,8,9] allows instantaneous constraints to be violated at each round, but tries to satisfy the constraints in the long run. In other words, the agent wants to ensure the long-term constraint of  $\sum_{t=1}^T \mathbf{g}_t(x_t) \leq \mathbf{0}$  over some given period of length  $T$ . This type of long-term constraint is appropriate in many applications (e.g., smart grid with renewable energy supply [8]). Thus, we aim to solve the following online optimization problem:

$$\min_{\{x_t\}_{t=1}^T} \sum_{t=1}^T f_t(x_t), \quad s.t. \quad \sum_{t=1}^T \mathbf{g}_t(x_t) \leq \mathbf{0}. \quad (\text{P1})$$

Solving problem (P1) exactly is still impossible in the online setting, since the information about the  $f_t$  and  $\mathbf{g}_t$  is unknown before the action  $x_t$  is chosen. Instead, our goal is to make the total loss  $\sum_{t=1}^T f_t(x_t)$  as low as possible compared to the total loss incurred by the benchmark sequence  $\{x_t^*\}_{t=1}^T$  ( $x_t^*$  is commonly termed the per-slot minimizer since  $x_t^* = \arg \min_{x \in \chi, \mathbf{g}_t(x) \leq \mathbf{0}} f_t(x)$ ) and meanwhile, to ensure that  $\sum_{t=1}^T \mathbf{g}_t(x_t)$  is not too large, i.e., the long-term constraint is not violated too much. Therefore, for any sequence  $\{x_t\}_{t=1}^T$  yielded by online algorithms, we define the dynamic regret and the constraint violations, respectively as follow,

$$\begin{aligned} \text{Regret} &= \sum_{t=1}^T f_t(x_t) - \sum_{t=1}^T f_t(x_t^*), \\ \text{Vio}_k &= \sum_{t=1}^T g_{t,k}(x_t), \quad k \in \{1, 2, \dots, K\}. \end{aligned} \quad (1)$$

In this paper, we consider the dynamic regret and constraint violations as the performance metrics. We emphasize that the definition of dynamic regret and constraint violations in (1) are prevalent and widely adopted in the literature [3,8,9,11]. Our goal is to choose  $x_t$  in each round  $t$  such that both the dynamic regret and constraint violations grow sub-linearly with respect to the time horizon  $T$ . Note that the regret defined in (1) may be negative but this also makes sense. This is because we aim to minimize the total cost defined in (P1) as small as possible, while the comparator sequence can be arbitrarily given. The significance of the regret bound guarantee is to make sure the total cost incurred by the agent does not exceed that incurred by a comparator sequence too much, and we would like to see the appearance of the negative regret, i.e, the total cost incurred by the agent is smaller than that incurred by a comparator sequence. Indeed, negative regret is very common in the standard OCO in terms of universal dynamic regret in which the comparator sequence is arbitrary given [15–17].

Intuitively, the performance bounds of any online algorithm should depend on how drastically  $\{f_t\}$  and  $\{\mathbf{g}_t\}$  vary across time, that is, the temporal variations of  $\{f_t\}$  and  $\{\mathbf{g}_t\}$ . Thus we need to quantify the temporal variations of the dynamic environment. Specifically, we need to quantify the temporal variations of functions sequence. There are mainly two kinds of regularities used in the literature of constrained OCO [3,8–11,18].

- Path-length: the accumulated variation of per-slot minimizers  $\{x_t^*\}$

$$V_x = \sum_{t=2}^T \|x_t^* - x_{t-1}^*\|.$$

- Function variation: the accumulated variation of consecutive constraints

$$V_g = \sum_{t=2}^T \sup_{x \in \chi} \|\mathbf{g}_t(x) - \mathbf{g}_{t-1}(x)\|.$$

The reason we define the accumulative variation  $V_x$  with respect to  $x_t^*$  is that it can quantify the temporal variations of the entire dynamic environment including loss functions  $\{f_t\}$  and constraint functions  $\{g_t\}$  since  $x_t^* = \arg \min_{x \in \chi} \{f_t(x) | g_t(x) \leq \mathbf{0}\}$ . While other definitions of it like  $\sum_{t=2}^T \max_{x \in \chi} |f_t - f_{t-1}|$  can only quantify the temporal variations of the loss functions.

We let  $\|\cdot\|$  be the Euclidean norm throughout this paper. In general, it is challenging to achieve sublinear performance bounds for any online algorithm unless regularity measures are sublinear; that is, the optimization problem is feasible. For example, a non-oblivious adversary may choose a new objective function  $f_t$  and constraint function  $g_t$  such that the current per-slot minimizer  $x_t^*$  is at least  $O(1)$  distance away from the selected action  $x_t$  at each round  $t$  (i.e., the accumulative variations are the order  $T$ ). In such case, any online algorithm cannot track the per-slot minimizers sequence  $\{x_t^*\}$  well and guarantee the sublinear dynamic regret/constraint violations.

## 2.2. Assumptions

After specifying the problem, here we introduce some assumptions in this paper, which are also common in the literature of constraint OCO [4,8,9].

**Assumption 1.** We make following assumptions with respect to feasible set  $\chi$ , objective functions  $\{f_t(x)\}_{t=1}^T$  and constraint functions  $\{g_t(x)\}_{t=1}^T$ :

- The feasible set  $\chi$  is closed, convex, and compact with diameter  $R$ , i.e.,  $\forall x, y \in \chi$ , it holds that  $\|x - y\| \leq R$ .
- The loss functions and constraint functions are convex, and bounded on  $\chi$ , i.e., there exists a positive constant  $F$  such that  $\max\{|f_t(x)|, \|g_t(x)\|\} \leq F, \forall x \in \chi, t$ .
- The gradients of  $g_{k,t}$  and  $f_t$  are upper-bounded by  $G$  over  $\chi$ , i.e.,  $\max\{\|\nabla f_t(x)\|, \|\nabla g_{k,t}(x)\|\} \leq G, \forall x \in \chi, k, t$ . This is equivalent to  $g_t$  is Lipschitz continuous with parameter  $\beta$  ( $\beta = KG$ ), i.e.,  $\|g_t(x) - g_t(y)\| \leq \beta \|x - y\|, \forall x, y \in \chi, t$ .

Under [Assumption 1](#), we study problem (P1) in the full-information setting; that is, at round  $t$ , the agent can observe the complete loss and constraint functions after the decision  $x_t$  is submitted. In the following sections, we will propose a virtual-queue-based parameter-free algorithm and show that it simultaneously achieves sublinear regret and constraint violations without the Slater condition.

## 3. Algorithm

In this section, we propose a novel virtual-queue-based algorithm, VQB, which is illustrated in [Algorithm 1](#). It introduces a sequence of dual variables  $\{\lambda(t)\}$ , which is also called virtual queue. The purpose of introducing the virtual queues is that we can characterize the regret and constraint violations through the drift-plus-penalty expression and then analyze the regret and the constraint violations based on it. Similar ideas of updating dual variables based on the virtual queues are adopted in several very recent works (e.g., [4,5]) for OCO with long-term and time-invariant constraints.

---

### Algorithm 1 VQB

---

- 1: **Initialize:**  $\alpha_1, \gamma_0 > 0, g_0 = \lambda(0) = 0$ , and  $x_1 \in \chi$ .
  - 2: **for** round  $t = 1 \dots T - 1$  **do**
  - 3:   Update the dual iterate  $\lambda(t)$ :
  - 4:    $\lambda(t) = \max\{\lambda(t-1) + \gamma_{t-1} g_{t-1}(x_t), -\gamma_{t-1} g_{t-1}(x_t)\}$
  - 5:   Update the primal iterate that satisfies:
  - 6:    $x_{t+1} = \arg \min_{x \in \chi} \nabla f_t(x_t)^T (x - x_t) + [\lambda(t) + \gamma_{t-1} g_{t-1}(x_t)]^T (\gamma_t g_t(x)) + \alpha_t \|x - x_t\|^2$
  - 7:   Choose the action  $x_{t+1}$
  - 8: **end for**
- 

But there are some differences between our algorithm and theirs. First, in order to ensure both regret and constraint violations are simultaneously sublinear for the time-varying constraints setting, we design a new way of involving instantaneous per-slot constraint violation into the virtual queues and decision sequence update. Moreover, the learning rates of our algorithm, i.e.,  $\alpha_t$  and  $\gamma_t$  are time-varying, while the learning rates of algorithm in [1,4,5] are unchanged in the whole time horizon. Therefore, our algorithm needs a new regret and constraint violation analysis due to the new update rule of virtual queues and the time-varying parameters. We will show more details about it in the theoretical analysis part of [Section 4](#).

Here we elaborate on the novelty and intuition of the entire algorithmic approach of VQB. Note that if there are no constraints  $\{g_t\}$  (i.e.,  $g_t = \mathbf{0}$ ), then VQB has  $\lambda_t = \mathbf{0}, \forall t$  and becomes the OGD algorithm, which has been commonly used in classic OCO with learning rate  $\eta = \frac{1}{2\alpha_t}$  since

$$x_{t+1} = \arg \min_{x \in \chi} \underbrace{\nabla f_t(x_t)^T (x - x_t) + \alpha_t \|x - x_t\|^2}_{\text{penalty}} = \Pi_\chi(x_t - \frac{1}{2\alpha_t} \nabla f_t(x_t)). \quad (2)$$

We call the term marked by an underbrace in (2) the penalty. Hence, the OGD algorithm is to minimize the penalty term and is a special case of VQB. In our algorithm VQB, if we define  $\mathbf{Q}(t) = \lambda(t) + \gamma_{t-1} \mathbf{g}_{t-1}(x_t) = \max\{\lambda(t-1) + 2\gamma_{t-1} \mathbf{g}_{t-1}(x_t), 0\}$  to be the vector of virtual queue backlogs and define Lyapunov drift  $\Delta(t) = \frac{1}{2} \|\mathbf{Q}(t+1)\|^2 - \frac{1}{2} \|\mathbf{Q}(t)\|^2$ , then the intuition behind VQB is to choose  $x_{t+1}$  to minimize an upper bound of the following expression. (Since  $x_{t+1}$  has not been determined at round  $t$ , we replace  $\mathbf{g}_t(x_{t+1})$  with  $\mathbf{g}_t(x)$  in  $\Delta(t)$  and omit the constant term.)

$$\underbrace{\Delta(t)}_{\text{drift}} + \underbrace{\nabla f_t(x_t)^T (x - x_t) + \alpha_t \|x - x_t\|^2}_{\text{penalty}}.$$

Thus, the intention is to minimize penalty plus the Lyapunov drift, which is a natural method in stochastic network optimization incorporated with the stability condition (e.g., [19–21]). The drift term  $\Delta(t)$  could be used to evaluate the constraint violations and is closely related to the virtual queues. The penalty term includes the regularization term  $\|x_t - x_{t-1}\|^2$  which could smoothen the difference between the coherent actions and make the whole expression strongly-convex. The remaining term describes the optimization problem.

Our algorithm also has a close connection with the saddle point methods proposed in the literature of constrained OCO [3,12], which also incorporates dual variables to the decision-making process. For example, in Algorithm 1,  $\lambda(t) = \max\{\lambda(t-1) + \gamma_{t-1} \mathbf{g}_{t-1}(x_t), -\gamma_{t-1} \mathbf{g}_{t-1}(x_t)\}$  is a virtual queue vector for the constraint violations. The role of  $[\lambda(t) + \gamma_{t-1} \mathbf{g}_{t-1}(x_t)]^T$  is similar to a dual variable vector in saddle point-typed OCO algorithms. The main differences between our algorithm and theirs is the update of dual variables and the way of incorporating constraint functions into the dual variables (e.g., our algorithm uses a virtual queue to track the constraint violation, and the dual variables in our algorithm are adaptively adjusted by the per-time slot constraint violation). These differences render our algorithm some advantages over saddle point methods in terms of performance guarantees.

#### 4. Results

In this section, we first present the analysis and major theoretical results of our algorithm. Next, we extend our results to the case when the time horizon is unknown and the variation of consecutive constraints is smooth enough across time, which captures many practical scenarios and has been frequently considered in [3,22,23].

##### 4.1. Main results

Within this subsection, we present the upper bounds on the dynamic regret and constraint violations for VQB.

**Theorem 1.** Consider OCO problem (P1) under Assumption 1, let  $\{x_t^*\}_{t=1}^T$  be the per-slot minimizers sequence which satisfies  $x_t^* = \operatorname{argmin}_{x \in \mathcal{X}} \mathbf{g}_t(x) \leq 0$ .

- (Case 1) Setting  $\alpha_t = \sqrt{\frac{T}{R + \sum_{i \leq t} \|x_i^* - x_{i-1}^*\|}}$  and  $\gamma_t^2 = \frac{1}{2\beta^2} \frac{1}{\sqrt{2R}}$  in VQB, then we have the following performance upper bounds

$$\begin{aligned} \text{Regret} &\leq O(\max\{\sqrt{TV_x}, V_g\}), \\ \text{Vio}_k &\leq O(\max\{\sqrt{T}, V_g\}), \forall k = 1, 2, \dots, K. \end{aligned} \tag{3}$$

- (Case 2) Setting  $\alpha_t = \sqrt{\frac{T}{R + \sum_{i \leq t} \|x_i^* - x_{i-1}^*\|}}$  and  $\gamma_t^2 = \frac{1}{2\beta^2} \frac{1}{\sqrt{2R}} \frac{1}{\sqrt{t+1}}$  in VQB, then we have the following performance upper bounds

$$\begin{aligned} \text{Regret} &\leq O(\sqrt{TV_x}), \\ \text{Vio}_k &\leq O(\max\{T^{\frac{3}{4}}, V_g\}), \forall k = 1, 2, \dots, K. \end{aligned} \tag{4}$$

There are several advantages stated as following that makes our results outperform previous studies. First, Theorem 1 implies that VQB can guarantee sublinear regret and constraint violations simultaneously, as long as the accumulated variations of the environment are sublinear, i.e.,  $V_x = o(T)$  and  $V_g = o(T)$ . Previous studies listed in Table 1 do not always simultaneously guarantee the sublinear performance bounds since they introduce the  $o(T)V_x$  or  $o(T)V_g$  term in their performance bounds, which may be at least of the order  $T$  even the optimization problem is feasible, i.e.,  $\max\{O(V_x), O(V_g)\} = o(T)$ .

Second, the dynamic regret upper bound guaranteed by both two cases of Theorem 1 could match the state-of-the-art dynamic regret bound  $O(\sqrt{TV_x})$  in general OCO [15,17,24], when the path-length of the benchmark sequence is  $V_x$ .

Moreover, our algorithm is parameter-free, that is, the parameters in our algorithm do not require prior information of the regularities (e.g.,  $V_x$  or  $V_g$ ). Meanwhile, Theorem 1 holds whether the Slater condition holds or not. The theoretical results of most previous study are valid either under the Slater condition, or the order of the regularities are known prior to the learner. Only [11] is both parameter-free and independent of the assumption of the Slater condition, however, it introduced degraded performance bounds and cannot guarantee the sublinear regret and constraint violations simultaneously. Readers could see Table 1 for the detailed comparisons.

We compare the performance bounds of our algorithm with the previous studies listed in Table 1. When  $V_x$  is not too large (e.g.,  $V_x = o(\sqrt{T})$ ), the regret and constraint violations bounds presented in the first case of Theorem 1 are all no worse than the state-of-the-art results, i.e.,  $O[\sqrt{TV_x}, \sqrt{T}]$  and  $O[\sqrt{TV_x}, V_x^{1/4}T^{3/4}]$ , established in [8,9], respectively. Besides, the dynamic regret bound presented in the second case of Theorem 1 is superior to all existing works, and the corresponding constraint violations are also strictly sublinear when the optimization problem is feasible.

**Proof sketch of Theorem 1.** Within this subsection, we give a proof sketch of Theorem 1. All the proof details of listed lemmas could be found in the Appendix A. Since the drift-plus-penalty expression characterizes the dynamic regret expression, and we can translate the bounds of virtual queues  $\{\lambda(t)\}$  into bounds of constraint violations. Thus, our proof starts with the analysis of virtual queues properties and drift-plus-penalty expression, that is, the Lyapunov drift term  $\Delta(t) = \frac{1}{2}[\|\lambda(t+1)\|^2 - \|\lambda(t)\|^2]$  plus the penalty term  $f_t(x_t)$ , which is associated with the loss value after choosing an action. First, we present the main properties for virtual queues  $\{\lambda(t)\}_{t=1}^T$  introduced in Algorithm 1 and Lyapunov drift term.

**Lemma 1 (Properties of Virtual Queues).** In Algorithm 1, we have the following properties for virtual queues  $\lambda(t)$  and Lyapunov drift term  $\Delta(t)$ :

1.  $\lambda(t) \geq \mathbf{0}$
2.  $\lambda(t) + \gamma_{t-1}\mathbf{g}_{t-1}(x_t) \geq \mathbf{0}$
3.  $\|\lambda(t)\| \geq \gamma_{t-1}\|\mathbf{g}_{t-1}(x_t)\|$
4.  $\gamma_{t-1}\mathbf{g}_{t-1}(x_t) \leq \lambda(t) - \lambda(t-1)$ , furthermore,  $\|\lambda(t)\| - \|\lambda(t-1)\| \leq \gamma_{t-1}\|\mathbf{g}_{t-1}(x_t)\|$
5.  $\Delta(t) \leq \gamma_t[\lambda(t)]^T \mathbf{g}_t(x_{t+1}) + \gamma_t^2 \|\mathbf{g}_t(x_{t+1})\|^2$

The proof of this lemma is motivated by [4,5]. However, due to our new algorithm, different constraints setting and fewer assumptions, our proof techniques are slightly different from theirs. Then we present the upper bound of drift-plus-penalty expression in the following lemma.

**Lemma 2 (Upper Bound of the Drift-Plus-Penalty Expression).** Under Assumption 1, let  $\delta > 0$  and  $\{\alpha_t\}_{t=1}^T, \{\gamma_t\}_{t=1}^T$  be any positive non-increasing sequences, if  $2\gamma_t \leq \gamma_{t-1} + \gamma_{t+1}$  holds for all  $t$ , then VQB ensures that

$$f_t(x_t) + \Delta(t) \leq \alpha_t \|x_t^* - x_t\|^2 - \alpha_{t+1} \|x_{t+1} - x_{t+1}^*\|^2 + 4R\alpha_t \|x_{t+1}^* - x_t^*\| + (\beta^2 \gamma_{t-1}^2 + \frac{\delta}{2} - \alpha_t) \|x_{t+1} - x_t\|^2 + \frac{1}{2\delta} G^2 + \frac{1}{2} \gamma_t \gamma_{t+1} \|\mathbf{g}_t(x_{t+1})\|^2 - \frac{1}{2} \gamma_{t-1} \gamma_t \|\mathbf{g}_{t-1}(x_t)\|^2 + \gamma_{t-1} \gamma_t \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\|^2 + f_t(x_t^*). \tag{5}$$

This is the key lemma in our theoretical analysis, which is used to yield the eventual bounds of regret and virtual queues. Next, we bound the dynamic regret as follows based on Lemma 2.

**Lemma 3 (Regret Bound).** Under Assumption 1, for arbitrary  $\delta > 0$  which satisfies  $\alpha_t \geq \beta^2 \gamma_{t-1}^2 + \frac{\delta}{2}$ , if  $\gamma_t \leq \gamma_{t+1}, \alpha_t \leq \alpha_{t+1}$  and  $2\gamma_t \leq \gamma_{t-1} + \gamma_{t+1}$  hold for all  $t$ , then VQB ensures that

$$\sum_{t=1}^T f_t(x_t) \leq \sum_{t=1}^T f_t(x_t^*) + \alpha_1 R^2 + 4R \sum_{t=1}^T \alpha_t \|x_{t+1}^* - x_t^*\| + \frac{TG^2}{2\delta} + \frac{1}{2} \gamma_T \gamma_{T+1} \|\mathbf{g}_T(x_{T+1})\|^2 + \frac{1}{2} \|\lambda(1)\|^2 + 2F \sum_{t=1}^T \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\|. \tag{6}$$

Here we define  $x_{T+1}^* = x_T^*$ .

We further bound the eventual length of virtual queues in the following lemma based on Lemma 2.

**Lemma 4.** Under Assumption 1, setting  $\delta, \alpha_t$  and  $\gamma_t$  to be the same as Lemma 3, then VQB ensures that

$$\|\lambda(T)\| \leq 2\sqrt{F(T-1)} + \sqrt{2\alpha_1 R^2} + \sqrt{\frac{(T-1)G^2}{\delta}} + \gamma_{T-1} \|\mathbf{g}_{T-1}(x_T)\| + 2\sqrt{2R \sum_{t=1}^{T-1} \alpha_t \|x_{t+1}^* - x_t^*\|} + 2\sqrt{F \sum_{t=1}^{T-1} \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\|}. \tag{7}$$

This is another critical lemma in our theoretical analysis that could be used to yield the constraint violations' upper bounds. We derive the upper bounds of the constraint violations in the following two lemmas.

**Lemma 5.** For any non-increasing sequence  $\{\gamma_t\}$ , VQB ensures that

$$\sum_{t=1}^T \mathbf{g}_{k,t}(x_t) \leq \frac{\|\lambda(T)\|}{\gamma_T} + V_g, \quad \forall k = 1, 2, \dots, K. \tag{8}$$

Recall that [Lemma 4](#) bounds the virtual queue length. Thus combining this lemma with [Lemma 5](#), we can bound the constraint violations in the following lemma.

**Lemma 6** (Constraint Violations' Bounds). *Setting  $\delta$ ,  $\alpha_t$  and  $\gamma_t$  to be the same as [Lemma 3](#), then VQB ensures that*

$$\begin{aligned} \sum_{t=1}^T g_{k,t}(x_t) &\leq \frac{2}{\gamma_T} \sqrt{F(T-1)} + \frac{1}{\gamma_T} \sqrt{2\alpha_1 R^2} + V_g + \frac{\gamma_{T-1}}{\gamma_T} \|\mathbf{g}_{T-1}(x_T)\| \\ &+ \frac{2}{\gamma_T} \sqrt{2R \sum_{t=1}^{T-1} \alpha_t \|x_{t+1}^* - x_t^*\|} + \frac{2}{\gamma_T} \sqrt{F \sum_{t=1}^{T-1} \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\|} + \frac{G}{\gamma_T} \sqrt{\frac{T-1}{\delta}}. \end{aligned} \quad (9)$$

According to [Lemmas 3](#) and [6](#), with parameters stated in [Theorem 1](#), we could prove the theoretical results of [Theorem 1](#). First we consider the **Case 1** in [Theorem 1](#), by the setting of  $\alpha_t$  and according to [Lemma 13](#), we can obtain

$$\begin{aligned} \sum_{t=1}^T \alpha_t \|x_{t+1}^* - x_t^*\| &= \sum_{t=1}^T \sqrt{\frac{T}{R + \sum_{i \leq t} \|x_i - x_{i-1}\|}} \|x_{t+1}^* - x_t^*\| \\ &= \sqrt{T} \sum_{t=1}^T \frac{\|x_{t+1}^* - x_t^*\|}{\sqrt{R + \sum_{i \leq t} \|x_i - x_{i-1}\|}} \leq 2\sqrt{T} \sqrt{\sum_{t=0}^T \|x_{t+1}^* - x_t^*\|} = 2\sqrt{TV_x}. \end{aligned} \quad (10)$$

By the setting of  $\gamma_t$  and according to [Lemma 12](#), we also have

$$\begin{aligned} \sum_{t=1}^T \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\| &= \frac{1}{2\beta^2} \frac{1}{\sqrt{2R}} \sum_{t=1}^T \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\| \\ &\leq \frac{1}{2\beta^2} \frac{1}{\sqrt{2R}} \sum_{t=1}^T \max_{x \in \mathcal{X}} \|\mathbf{g}_{t-1}(x) - \mathbf{g}_t(x)\| = \frac{1}{2\beta^2} \frac{1}{\sqrt{2R}} V_g. \end{aligned} \quad (11)$$

Setting  $\delta = \frac{1}{2} \sqrt{\frac{T}{R+V_x}}$ , it is easy to verify that  $\alpha_t \geq \beta^2 \gamma_{t-1}^2 + \frac{\delta}{2}$ ,  $2\gamma_t \leq \gamma_{t-1} + \gamma_{t+1}$ , and both  $\{\alpha_t\}$  and  $\{\gamma_t\}$  are non-increasing sequences. Thus combining [Lemma 2](#) with [\(10\)](#), [\(11\)](#) and rearranging terms yields

$$\begin{aligned} &\sum_{t=1}^T f_t(x_t) - \sum_{t=1}^T f_t(x_t^*) \\ &\leq 4R \sum_{t=1}^T \alpha_t \|x_{t+1}^* - x_t^*\| + \alpha_1 R^2 + \frac{TG^2}{2\delta} + \frac{1}{2} \gamma_T \gamma_{T+1} \|\mathbf{g}_T(x_{T+1})\|^2 + 2F \sum_{t=1}^T \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\| + \frac{1}{2} \|\lambda(1)\|^2 \\ &\leq 8R\sqrt{TV_x} + \sqrt{\frac{T}{R}} R^2 + \frac{TG^2}{2\delta} + \frac{1}{4\beta^2} \frac{1}{\sqrt{2R}} F^2 + \frac{F}{\beta^2} \frac{V_g}{\sqrt{2R}} + \frac{1}{2} \|\lambda(1)\|^2 \\ &\stackrel{(a)}{\leq} 8R\sqrt{TV_x} + \sqrt{\frac{T}{R}} R^2 + TG^2 \sqrt{\frac{R+V_x}{T}} + \frac{1}{4\beta^2} \frac{1}{\sqrt{2R}} F^2 + \frac{F}{\beta^2} \frac{V_g}{\sqrt{2R}} + \frac{1}{2} \|\lambda(1)\|^2 \\ &= 8R\sqrt{TV_x} + R^{2/3} \sqrt{T} + G^2 \sqrt{T(R+V_x)} + \frac{1}{4\beta^2} \frac{1}{\sqrt{2R}} F^2 + \frac{F}{\beta^2} \frac{V_g}{\sqrt{2R}} + \frac{1}{2} \|\lambda(1)\|^2 \\ &= O(\max\{\sqrt{TV_x}, V_g\}), \end{aligned} \quad (12)$$

where (a) holds since we set  $\delta = \frac{1}{2} \sqrt{\frac{T}{R+V_x}}$ . According to [Lemma 5](#) and [Assumption 1](#), we have

$$\begin{aligned} \sum_{t=1}^T g_{t,k}(x_t) &\leq \frac{2}{\gamma_T} \sqrt{F(T-1)} + \frac{1}{\gamma_T} \sqrt{2\alpha_1 R^2} + \frac{\gamma_{T-1}}{\gamma_T} \|\mathbf{g}_{T-1}(x_T)\| + \frac{G}{\gamma_T} \sqrt{\frac{T-1}{\delta}} \\ &+ \frac{2}{\gamma_T} \sqrt{2R \sum_{t=1}^{T-1} \alpha_t \|x_{t+1}^* - x_t^*\|} + \frac{2}{\gamma_T} \sqrt{F \sum_{t=1}^{T-1} \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\|} + V_g \\ &\stackrel{(a)}{\leq} \sqrt{\frac{1}{2\sqrt{2R}\beta^2}} [2\sqrt{F(T-1)} + R\sqrt{2\sqrt{\frac{T}{R}}} + G\sqrt{2\frac{(T-1)\sqrt{R+V_x}}{\sqrt{T}}}] \end{aligned}$$



$$\begin{aligned}
 &+ F + \sqrt{\frac{2}{\sqrt{2R}\beta^2}} [\sqrt{4R\sqrt{TV_x}} + \sqrt{\frac{F}{2\beta^2} \frac{V_g}{\sqrt{2R}}}] + V_g \\
 &\stackrel{(b)}{\leq} \sqrt{\frac{1}{2\sqrt{2R}\beta^2}} [2\sqrt{F(T-1)} + R\sqrt{2\sqrt{\frac{T}{R}}} + G\sqrt{2R\sqrt{T(1+T)}}] \\
 &+ F + \sqrt{\frac{2}{\sqrt{2R}\beta^2}} [\sqrt{4R\sqrt{RT}} + \sqrt{\frac{F}{2\beta^2} \frac{V_g}{\sqrt{2R}}}] + V_g = O(\max\{\sqrt{T}, V_g\}),
 \end{aligned} \tag{13}$$

where (a) is due to (10) and (11); (b) is due to  $V_x \leq RT$ . For the **Case 2** in Theorem 1, since the settings of  $\alpha_t$  in both two cases are the same, we can also derive that

$$\sum_{t=1}^T \alpha_t \|x_{t+1}^* - x_t^*\| \leq 2\sqrt{TV_x}. \tag{14}$$

For term  $\sum_{t=1}^T \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\|$ , according to Lemma 12 and by the setting of  $\gamma_t$ , we have

$$\begin{aligned}
 \sum_{t=1}^T \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\| &= \frac{1}{2\beta^2} \frac{1}{\sqrt{2R}} \sum_{t=1}^T \frac{1}{\sqrt{t}} \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\| \\
 &\leq \frac{F}{\beta^2} \frac{1}{\sqrt{2R}} \sum_{t=1}^T \frac{1}{\sqrt{t}} \leq \frac{F}{\beta^2} \frac{\sqrt{2T}}{\sqrt{R}}.
 \end{aligned} \tag{15}$$

Setting  $\delta = \frac{1}{2}\sqrt{\frac{T}{R+V_x}}$ , it is easy to verify that  $\alpha_t \geq \beta^2 \gamma_{t-1}^2 + \frac{\delta}{2}$ ,  $2\gamma_t \leq \gamma_{t-1} + \gamma_{t+1}$ , and both  $\{\alpha_t\}$  and  $\{\gamma_t\}$  are non-increasing sequences. Based on Lemma 2, combining (14), (15) and Assumption 1 gives

$$\begin{aligned}
 &\sum_{t=1}^T f_t(x_t) - \sum_{t=1}^T f_t(x_t^*) \\
 &\leq 4R \sum_{t=1}^T \alpha_t \|x_{t+1}^* - x_t^*\| + \alpha_1 R^2 + \frac{TG^2}{2\delta} + \frac{1}{2} \gamma_T \gamma_{T+1} \|\mathbf{g}_T(x_{T+1})\|^2 + 2F \sum_{t=1}^T \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\| + \frac{1}{2} \|\lambda(1)\|^2 \\
 &\leq 8R\sqrt{TV_x} + R^{3/2}\sqrt{T} + TG^2\sqrt{\frac{R+V_x}{T}} + \frac{F^2}{4\beta^2} \frac{1}{\sqrt{2R}} \frac{1}{\sqrt{T}} + \frac{2F^2}{\beta^2} \frac{\sqrt{2T}}{\sqrt{R}} + \frac{1}{2} \|\lambda(1)\|^2 = O(\sqrt{TV_x}).
 \end{aligned} \tag{16}$$

Furthermore, based on Lemma 5, we obtain the bounds of constraint violations as follows

$$\begin{aligned}
 \sum_{t=1}^T g_{t,k}(x_t) &\leq \frac{2}{\gamma_T} \sqrt{F(T-1)} + \frac{1}{\gamma_T} \sqrt{2\alpha_1 R^2} + \frac{\gamma_{T-1}}{\gamma_T} \|\mathbf{g}_{T-1}(x_T)\| + \frac{G}{\gamma_T} \sqrt{\frac{T-1}{\delta}} \\
 &+ \frac{2}{\gamma_T} \sqrt{2R \sum_{t=1}^{T-1} \alpha_t \|x_{t+1}^* - x_t^*\|} + \frac{2}{\gamma_T} \sqrt{F \sum_{t=1}^{T-1} \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\|} + V_g \\
 &\stackrel{(a)}{\leq} T^{1/4} \sqrt{\frac{1}{2\sqrt{2R}\beta^2}} [2\sqrt{F(T-1)} + R\sqrt{2\sqrt{\frac{T}{R}}} + G\sqrt{2\frac{(T-1)\sqrt{R+V_x}}{\sqrt{T}}}] \\
 &+ (\frac{T+1}{T})^{1/4} F + T^{1/4} \sqrt{\frac{2}{\sqrt{2R}\beta^2}} [\sqrt{4R\sqrt{TV_x}} + \frac{F}{\beta} \sqrt{\frac{\sqrt{2T}}{\sqrt{R}}}] + V_g \\
 &\stackrel{(b)}{\leq} T^{1/4} \sqrt{\frac{1}{2\sqrt{2R}\beta^2}} [2\sqrt{F(T-1)} + R\sqrt{2\sqrt{\frac{T}{R}}} + G\sqrt{2R\sqrt{T(1+T)}}] \\
 &+ 2^{1/4} F + T^{1/4} \sqrt{\frac{2}{\sqrt{2R}\beta^2}} [\sqrt{4R\sqrt{TV_x}} + \frac{F}{\beta} \sqrt{\frac{\sqrt{2T}}{\sqrt{R}}}] + V_g = O(\max\{T^{3/4}, V_g\}),
 \end{aligned} \tag{17}$$

where (a) follows from (14) and (15); (b) holds by the fact that  $V_x \leq RT$  and  $\frac{T+1}{T} \leq 2$ . It completes the proof.



**Remark 1.** When feasible set  $\chi$  is time-varying, i.e.,  $\chi(t)$ , our algorithm VQB is valid and we can verify that its corresponding theoretical results also hold.

#### 4.2. Slater condition

In the previous section, we have shown that as long as the optimization problem is feasible, our algorithm could simultaneously achieve sublinear dynamic regret and constraint violation with only limited common assumptions in the literature of constrained OCO. Meanwhile, [3] pointed out that in many practical constrained OCO problems, the variation of consecutive constraints is smooth across time. Thus, we examine whether the smoothness of the dynamic environment's temporal variations can lead to better bounds of constraint violations for VQB. Within this subsection, we consider a slightly stronger Slater condition that has been considered in [3]. We will show that our variant of VQB, illustrated in Algorithm 2, could guarantee an  $O(1)$  constraint violations under this assumption. The difference between VQB and Algorithm 2 is the way of incorporating constraints into the virtual queues updates and decision iterations, i.e., Algorithm 2 uses  $\mathbf{g}_t(x_t)$  instead of  $\mathbf{g}_{t-1}(x_t)$  to update the dual iterate  $\lambda(t)$  and primal iterate  $x_t$  compared with VQB. Technically, the update step in Algorithm 2 can yield much lower constraint violations when the variation of consecutive constraints is smooth across time, as shown in Theorem 2. First, we give the definition of the Slater condition.

---

#### Algorithm 2

---

- 1: **Initialize:**  $\alpha, \gamma > 0, \mathbf{g}_0 = \lambda(0) = 0$ , and  $x_1 \in \chi$ .
  - 2: **for** round  $t = 1 \dots T - 1$  **do**
  - 3:   Update the dual iterate  $\lambda(t)$ :
  - 4:    $\lambda(t) = \max\{\lambda(t - 1) + \gamma \mathbf{g}_t(x_t), -\gamma \mathbf{g}_t(x_t)\}$
  - 5:   Update the primal iterate that satisfies:
  - 6:    $x_{t+1} = \arg \min_{x \in \chi} \nabla f_t(x_t)^T(x - x_t) + [\lambda(t) + \gamma \mathbf{g}_t(x_t)]^T(\gamma \mathbf{g}_t(x)) + \alpha \|x - x_t\|^2$
  - 7:   Choose the action  $x_{t+1}$
  - 8: **end for**
- 

**Assumption 2 (Slater Condition).** There exists  $\epsilon > 0$  and  $\hat{x} \in \chi$  such that  $\mathbf{g}_t(\hat{x}) \leq -\epsilon I, \forall t$ .

Assumption 2 is known as the interior point condition or Slater condition, which is also used widely in the literature of OCO with time-varying constraints [3,9,10,18]. Based on Assumption 2, we next introduce a slightly stronger Slater condition assumption, which is valid in many practical scenarios [3,22,23].

**Assumption 3.** The Slater constant  $\epsilon$  is larger than the maximum variation of consecutive constraints, i.e.,  $\epsilon > \bar{V}_g = \max_{t, x \in \chi} \|\mathbf{g}_{t+1}(x) - \mathbf{g}_t(x)\|$ .

Note that this assumption was adopted in [3], which is valid when the region defined by  $\{x|x \in \chi, \mathbf{g}_t(x) \leq \mathbf{0}\}$  is large enough, or the variation of consecutive constraints is smooth enough across time.

With a similar intuition of VQB, if we define  $\mathbf{Q}(t) = \lambda(t) + \gamma_t \mathbf{g}_t(x_t) = \max\{\lambda(t - 1) + 2\gamma_t \mathbf{g}_t(x_t), \mathbf{0}\}$  as the vector of virtual queue backlogs and let parameters  $\alpha_t, \gamma_t$  be time-invariant, Algorithm 2 also chooses  $x_{t+1}$  to minimize an upper bound of the following expression:

$$\underbrace{\Delta(t)}_{\text{drift}} + \underbrace{\nabla f_t(x_t)^T(x - x_t) + \alpha_t \|x - x_t\|^2}_{\text{penalty}}.$$

The reason we replace  $\mathbf{g}_{t-1}(x_t)$  with  $\mathbf{g}_t(x_t)$  in Algorithm 2 is motivated by the observation that  $\mathbf{g}_t(x_t)$  could be directly accumulated into queue  $\lambda(t)$  (recall that  $\lambda(t) = \max\{\lambda(t - 1) + \gamma \mathbf{g}_t(x_t), -\gamma \mathbf{g}_t(x_t)\}$ ) and we intend to have small queue backlogs when the variation of consecutive constraint functions is smooth across time. This is important for a much tighter analysis of the constraint violations under the strongly Slater condition. If we use  $\mathbf{g}_t(x_t)$  instead of  $\mathbf{g}_{t-1}(x_t)$  in VQB, we could get  $\gamma \mathbf{g}_{t,k}(x_t) \leq \lambda_k(t) - \lambda_k(t - 1)$ . Then, we can characterize the constraint violations only by the bounds of virtual queues  $\{\lambda(t)\}$  without the term  $V_g$  (comparing with Lemma 5), i.e.,  $\sum_{t=1}^T \mathbf{g}_{t,k}(x_t) \leq \frac{\lambda_k(T)}{\gamma} \leq \frac{\|\lambda(T)\|}{\gamma}, \forall k$ . In such case, the length of virtual queues  $\{\|\lambda(t)\|\}_{t=1}^T$  is upper bounded by a constant under the strongly Slater condition (Lemma 11). Then we could obtain an  $O(1)$  bound of constraint violations, shown in the following theorem.

**Theorem 2.** Under Assumptions 1-3, setting  $\alpha = T^a$  and  $\gamma^2 = \frac{1}{2\beta^2} T^a$  in Algorithm 2, the dynamic regret and constraint violations are upper bounded by

$$\begin{aligned} \text{Regret} &\leq O(\max\{T^a V_x, T^a V_g, T^{1-a}\}), \\ \text{Vio}_k &\leq O(1), \forall k = 1, 2, \dots, K. \end{aligned} \tag{18}$$

In particular, the performance upper bounds become  $O(\max\{\sqrt{T}V_x, \sqrt{T}V_g\})$  and  $O(1)$  if we set  $\alpha = 2\beta^2 \gamma^2 = \sqrt{T}$ .

Note that our performance bounds established by Algorithm 2 are strictly better than [3] under the same assumptions. Besides, the constraint violations for Algorithm 2 can decrease to  $O(1)$  when the variations of consecutive constraints are smooth enough across time.

**Proof sketch of Theorem 2.** Here we give a proof sketch of Theorem 2. All the proof details of listed lemmas could be found in Appendix B. Note that both  $\{\alpha_t\}$  and  $\{\gamma_t\}$  are constant sequences in Algorithm 2, thus here we omit the subscript  $t$ . Similar as Lemma 1, in Algorithm 2, we have the following lemma for the properties of virtual queues and Lyapunov drift term.

**Lemma 7.** In Algorithm 2, at each round  $t$ , we have

1.  $\lambda(t) \geq \mathbf{0}$
2.  $\lambda(t) + \gamma \mathbf{g}_t(x_t) \geq \mathbf{0}$
3.  $\|\lambda(t)\| \geq \gamma \|\mathbf{g}_{t-1}(x_t)\|$
4.  $\gamma \mathbf{g}_t(x_t) \leq \lambda(t) - \lambda(t-1)$ , furthermore,  $\|\lambda(t)\| - \|\lambda(t-1)\| \leq \gamma \|\mathbf{g}_t(x_t)\|$
5.  $\Delta(t) \leq \gamma [\lambda(t)]^T \mathbf{g}_{t+1}(x_{t+1}) + \gamma^2 \|\mathbf{g}_{t+1}(x_{t+1})\|_2^2$

The proof of this Lemma is similar as the proof of Lemma 1 and hence we omit the details.

**Lemma 8.** Under Assumptions 1–3, setting  $\alpha, \gamma, \delta$  such that  $\alpha \geq \frac{1}{2}(\beta^2\gamma^2 + \delta)$ , then Algorithm 2 ensures that

$$\begin{aligned} & f_t(x_t) + \Delta(t) \\ & \leq f_t(x_t^*) + \alpha \|x_t^* - x_t\|^2 - \alpha \|x_{t+1} - x_{t+1}^*\|^2 + 4R\alpha \|x_{t+1}^* - x_t^*\|^2 + \gamma^2 F \|\mathbf{g}_{t+1}(x_{t+1}) - \mathbf{g}_t(x_{t+1})\| \\ & \quad + \frac{1}{2\delta} G^2 + \frac{1}{2} \gamma^2 \|\mathbf{g}_{t+1}(x_{t+1})\|^2 - \frac{1}{2} \gamma^2 \|\mathbf{g}_t(x_t)\|^2 + \gamma \|\lambda(t)\| \|\mathbf{g}_{t+1}(x_{t+1}) - \mathbf{g}_t(x_{t+1})\|. \end{aligned} \quad (19)$$

Take a similar derivation process as the proof of Theorem 1, we also characterize the regret and constraint violations through the bound of drift-plus-penalty expression stated above. Therefore, we bound the dynamic regret and constraint violations in the following lemmas, respectively.

**Lemma 9.** Under Assumptions 1–3, setting  $\alpha, \gamma, \delta$  such that  $\alpha \geq \frac{1}{2}\beta^2\gamma^2 + \frac{1}{2}\delta$ , then Algorithm 2 ensures that

$$\sum_{t=1}^T f_t(x_t) \leq \sum_{t=1}^T f_t(x_t^*) + \alpha \|x_1^* - x_1\|^2 + 4R\alpha V_x + \frac{TG^2}{2\delta} + \gamma^2 F V_g + \frac{1}{2} \|\lambda(1)\|^2 + \gamma \max_t \|\lambda(t)\| V_g. \quad (20)$$

**Lemma 10.** In Algorithm 2, we have

$$\sum_{t=1}^T g_{t,k}(x_t) \leq \frac{\lambda_k(T)}{\gamma} \leq \frac{\|\lambda(T)\|}{\gamma}, \quad \forall k \in \{1, 2, \dots, K\}. \quad (21)$$

Note that the above two lemmas show that the final bounds of dynamic regret and constraint violations can be obtained by bounding the  $\lambda(t)$ . Since we are allowed to introduce the assumption of strongly Slater condition, we will show that the length of virtual queues is upper bounded by a constant in this case. Hence we adopt different techniques for the virtual queues analysis compared with the proof of Lemma 4, and bound them by accomplishing the following lemma.

**Lemma 11.** In Algorithm 2, we have

$$\|\lambda(t)\| \leq \gamma F + \frac{GR + \gamma^2 \epsilon F + 2\gamma^2 F^2 + \alpha R^2}{\gamma(\epsilon - \bar{V}_g)}, \quad \forall t. \quad (22)$$

Finally, based on the above lemmas, we prove Theorem 2 as follows. Setting  $\delta = \frac{1}{2}T^a$ , and it is easy to verify that  $\alpha \geq \frac{1}{2}\beta^2\gamma^2 + \frac{\delta}{2}$ . Combining Lemma 9 with Lemma 11, we have

$$\begin{aligned} & \sum_{t=1}^T f_t(x_t) - \sum_{t=1}^T f_t(x_t^*) \\ & \leq \alpha \|x_1^* - x_1\|^2 + 4R\alpha V_x + \frac{TG^2}{2\delta} + \gamma^2 F V_g + \frac{1}{2} \|\lambda(1)\|^2 + \gamma \max_t \|\lambda(t)\| V_g \\ & \leq \sum_{t=1}^T f_t(x_t^*) + \alpha R^2 + 4R\alpha V_x + \frac{TG^2}{2\delta} + \gamma^2 F V_g + \frac{1}{2} \|\lambda(1)\|^2 + (\gamma^2 F + \frac{GR + \gamma^2 \epsilon F + 2\gamma^2 F^2 + \alpha R^2}{(\epsilon - \bar{V}_g)}) V_g \\ & = O(\max\{T^a V_x, \gamma^2 V_g, T^{1-a}\}) = O(\max\{T^a V_x, V_g T^a, T^{1-a}\}). \end{aligned} \quad (23)$$

Furthermore, combining Lemma 10 with Lemma 11, we can obtain

$$\begin{aligned} \sum_{t=1}^T g_{t,k}(x_t) &\leq \frac{\|\lambda(T)\|}{\gamma} \leq F + \frac{GR + \gamma^2 \epsilon F + 2\gamma^2 F^2 + \alpha R^2}{\gamma^2(\epsilon - \bar{V}_g)} \\ &= O\left(\frac{\alpha}{\gamma^2}\right) = O(T^{1-a}), \quad \forall k = 1, 2, \dots, K. \end{aligned} \tag{24}$$

In particular, when setting  $\alpha = \frac{1}{2}\sqrt{T}$  and  $\gamma^2 = \frac{1}{2\beta^2}\sqrt{T}$ , the performance upper bounds become  $O(\max\{\sqrt{TV_x}, \sqrt{T}V_g\})$  and  $O(1)$ . This completes the proof.

### 4.3. Unknown time horizon T

In this subsection, we extend our algorithm and analyze the case when time horizon  $T$  is unknown. Recall that the parameters of VQB, i.e.,  $\gamma_t$  and  $\alpha_t$ , and previous methods for OCO with long-term and time-varying constraints all depend on the time horizon  $T$ , while the total rounds  $T$  are not known prior to the learner in many practical scenarios. In such cases, we use the doubling trick strategy to tune the parameters for our algorithm. To the best of our knowledge, we are the first to consider the unknown time horizon case for OCO with long-term and time-varying constraints. For any online Algorithm A whose parameters depend on the time horizon  $T$ , the doubling trick is described in Algorithm 3.

---

#### Algorithm 3 The Doubling Trick for Online Algorithm A

---

- 1: Let  $i = 1$ .
  - 2: **while** not reach the end of the time horizon **do**
  - 3:   Reset A with parameters chosen for  $T = 2^i$ .
  - 4:   Run A for  $2^i$  rounds.
  - 5:   Let  $i = i + 1$ .
  - 6: **end while**
- 

**Theorem 3.** Under Assumption 1, for any unknown time horizon  $T$ , run VQB until reaching the end of the time horizon. Let  $t_i = 2^i$  be the index of the first round of  $i$ -th epoch.

- (Case 1) Setting  $\alpha_t = \sqrt{\frac{2^i}{R + \sum_{t_i \leq i \leq t} \|x_t^* - x_{t-1}^*\|}}$  and  $\gamma_t^2 = \frac{1}{2\beta^2} \frac{1}{\sqrt{2R}}$  for  $t \in [t_i, t_{i+1} - 1]$ , then VQB with doubling trick ensures
 
$$\begin{aligned} \text{regret} &\leq O(\max\{\sqrt{TV_x}, V_g\}), \\ \text{Vio}_k &\leq O(\max\{\sqrt{T}, V_g\}), \quad \forall k = 1, 2, \dots, K. \end{aligned} \tag{25}$$

- (Case 2) Setting  $\alpha_t = \sqrt{\frac{2^i}{R + \sum_{t_i \leq i \leq t} \|x_t^* - x_{t-1}^*\|}}$  and  $\gamma_t^2 = \frac{1}{2\beta^2} \frac{1}{\sqrt{2R}} \frac{1}{\sqrt{t-t_i+2}}$  for  $t \in [t_i, t_{i+1} - 1]$ , then VQB with doubling trick ensures
 
$$\begin{aligned} \text{regret} &\leq O(\sqrt{TV_x}), \\ \text{Vio}_k &\leq O(\max\{T^{\frac{3}{4}}, V_g\}), \quad \forall k = 1, 2, \dots, K. \end{aligned} \tag{26}$$

Theorem 3 shows that our algorithm with the doubling trick can still preserve the order of dynamic regret and constraint violations bounds even though the time horizon  $T$  is unknown. Note that our algorithm adapts to the doubling trick because of the property of parameter-free, while parameter-dependent methods (e.g., [8]) cannot do this.

**Proof of Theorem 3.** Here we give a proof of Theorem 3. For the Case 1 in Theorem 3, since the  $i$ -th epoch consists of at most  $2^i$  rounds, the unknown time horizon  $T$  is divided into  $N = \lceil \log_2 T \rceil$  epochs. Let  $\Delta_x^i = \sum_{t=t_i}^{t-t_i-1} \|x_t^* - x_{t-1}^*\|$  and  $\Delta_g^i = \sum_{t=t_i}^{t-t_i-1} \sup_{x \in \mathcal{X}} \|g_t(x) - g_{t-1}(x)\|$ . By Theorem 1, in the  $i$ -th epoch there exists a constant  $C$  such that the dynamic regret and constraint violations are at most  $\max\{C\sqrt{T}\Delta_x^i, C\Delta_g^i\}$  and  $\max\{C\sqrt{2^i}, C\Delta_g^i\}$  respectively. The final bound could be obtained by summing the individual bounds over all the epochs. Therefore, we could upper bound the total dynamic regret and constraint violations as follows

$$\begin{aligned} \text{Regret} &\leq \max\left\{ \sum_{i=0}^N C\sqrt{2^i \Delta_x^i}, \sum_{i=0}^N C\Delta_g^i \right\} \stackrel{(a)}{\leq} \max\left\{ C\sqrt{\sum_{i=0}^N \Delta_x^i} \sqrt{\sum_{i=0}^N 2^i}, C\sum_{i=0}^N \Delta_g^i \right\} \\ &\leq \max\{C\sqrt{V_x} \sqrt{2^{N+1}} - 1, C\sum_{i=0}^N \Delta_g^i\} = \max\{C\sqrt{V_x} \sqrt{2^{\lceil \log_2 T \rceil + 1}} - 1, CV_g\} \\ &\leq \max\{C\sqrt{V_x} \sqrt{2^{\log_2 T + 2}}, CV_g\} = \max\{2C\sqrt{TV_x}, CV_g\}, \end{aligned} \tag{27}$$

**Table 2**  
Parameters of our algorithm and baselines.

Methods	Parameters
Baseline [8]	$\delta = 8nC^2 + 1, \eta = \frac{2}{\sqrt{T}}$
Baseline [3]	$\alpha = \mu = T^{\frac{1}{3}}$
Baseline [11]	$\delta = 1, \lambda_1 = 4\sqrt{2}T^{\frac{1}{8}}$
Baseline [9]	$\mu = T^{-\frac{1}{2}}, \alpha = 2T^{-\frac{1}{2}}$
VQB(Case 1)	$\alpha_t = \sqrt{\frac{T}{R + \sum_{i \leq t} \ x_i^* - x_{i-1}^*\ }}, \gamma_t^2 = \frac{1}{2\beta^2} \frac{1}{\sqrt{2R}}$
VQB(Case 2)	$\alpha_t = \sqrt{\frac{T}{R + \sum_{i \leq t} \ x_i^* - x_{i-1}^*\ }}, \gamma_t^2 = \frac{1}{2\beta^2} \frac{1}{\sqrt{2R}} \frac{1}{\sqrt{t+1}}$

where (a) is due to the Cauchy–schwarz inequality. And the total constraint violations are at most

$$\begin{aligned} \text{Vio}_k &\leq \max\left\{\sum_{i=0}^N C\sqrt{2^i}, \sum_{i=0}^N C\Delta_g^i\right\} = \max\left\{\frac{C}{\sqrt{2}-1}[\sqrt{2}^{N+1}-1], \text{CV}_g\right\} \\ &\leq \max\left\{\frac{2C}{\sqrt{2}-1}\sqrt{2}^{\log_2 T}, \text{CV}_g\right\} = \max\left\{\frac{2C}{\sqrt{2}-1}\sqrt{T}, \text{DV}_g\right\}, \forall k = 1, 2, \dots, K. \end{aligned} \quad (28)$$

For the **Case 2**, we conduct similar analysis as Case 1. By [Theorem 1](#), in the  $i$ -th epoch there exists a constant  $D$  such that the dynamic regret and constraint violations are at most  $D\sqrt{T\Delta_x^i}$  and  $\max\{D2^{\frac{3}{4}i}, D\Delta_g^i\}$  respectively. According to [\(27\)](#), the total regret is still at most the order of  $\sqrt{TV_x}$  without changing. For the total constraint violations, we also have

$$\begin{aligned} \text{Vio}_k &\leq \max\left\{D\sum_{i=0}^N 2^{\frac{3}{4}i}, \sum_{i=0}^N D\Delta_g^i\right\} \leq \max\left\{\frac{D}{2^{3/4}-1}(2^{\frac{3}{4}[\log_2 T]}-1), \text{DV}_g\right\} \\ &\leq \max\left\{\frac{2D}{2^{3/4}-1}2^{\frac{3}{4}\log_2 T}, \text{DV}_g\right\} = \max\left\{\frac{2D}{2^{3/4}-1}T^{3/4}, \text{DV}_g\right\}, \forall k = 1, 2, \dots, K. \end{aligned} \quad (29)$$

This completes the proof.

## 5. Numerical experiments

In this section, we conduct numerical experiments to validate the theoretical performance of our algorithm. Specifically, we consider the online ridge regression (ORR) problem [\[25\]](#) as the numerical example. We compare the time-averaged regrets and constraint violations of our algorithm with previous work in two different dynamic environments. The problem formulation of ORR at round  $t$  is as follows.

$$\begin{aligned} &\text{Minimize } \sum_{i=1}^n (x_i^T p_{i,t} + b - q_{i,t})^2 \\ &\text{s.t. } \|x_t\| \leq a_t \end{aligned} \quad (30)$$

Here  $\{p_{i,t}, q_{i,t}\}_{i=1}^n$  are the training data at round  $t$ , and  $a_t$  characterizes the  $t$ -th round constraint on the  $l_2$  norm of the decision variable, i.e., weight vector. We define  $\{x \mid \|x\|_\infty \leq C, x \in R^k\}$  as the feasible set. The above ORR formulation could be applied in accurate and reliable forecasting of traffic in intelligent transportation systems [\[26\]](#). The training data  $\{p_{i,t}, q_{i,t}\}_{i=1}^n$  and constraint  $a_t$  may not be known prior to the agent at round  $t$  due to the delayed arrival training data.

**Experimental setting.** At round  $t$ , we generate the parameters  $\{p_{i,t}, q_{i,t}, \forall i\}$ ,  $a_t$  and the per-slot minimizer  $x_t^*$  in the following way. Let  $x_t^* = \Pi_x(x_{t-1}^* + \tau_t)$ , where each entry of  $\tau_t$  is a uniform random variable, sampled from a time-varying set  $B_t$  (we will specialize it later). Then we generate  $a_t$  and  $\{p_{i,t}, q_{i,t}\}_{i=1}^n$  as follows. (i)  $p_{i,t} = p_{i,t-1} + u_{i,t}$ , where each entry of  $u_{i,t}$  is i.i.d, uniformly sampled from set  $B_t$ . (ii)  $q_{i,t} = p_{i,t}^T x_t^* + b$ . (iii)  $a_t = \|x_t^*\|$ .

Next, we introduce the baselines [\[3,8,9,11\]](#) for comparison. The algorithms in [\[3,11\]](#) are based on MOSP method. Although [\[9\]](#) only considered the bandit setting, their algorithm and theoretical guarantees are also valid in the full-information setting. Meanwhile, note that the theoretical guarantees in [\[8\]](#) are valid only when the agent has prior knowledge of  $V_x$  (or the order of it). For fair comparison, we set the learning rates in their algorithm to be parameter-free, and obtain the  $O(V_x T^{1/2})$  regret and  $O(\max\{V_x^{1/2} T^{1/2}, T^{3/4}\})$  constraint violations. Finally, we introduce our experimental details. In our experiment, we let  $n = k = 5$ ,  $C = 7$ . The parameters of our algorithm and other baselines are presented in [Table 2](#).

**Results and analysis.** We first consider the case when  $V_x = V_g = O(\log(T))$ . To do this, we set  $B_t$  to be  $[-\frac{1}{2t}, \frac{1}{2t}]$ . From [Fig. 1\(a\)](#) and (b), we can see that our algorithm VQB achieves lowest time-averaged regret  $\frac{\text{Regret}(t)}{t}$  and constraint violation  $\frac{\text{Vio}(t)}{t}$ , which validates our theoretical results. Moreover, we can also see that the regrets achieved by VQB under two parameter settings are very close, which is consistent with the theoretical results in [Theorem 1](#) that the regret upper bounds between them are identical by noting that  $V_g = O(\log(T))$ .

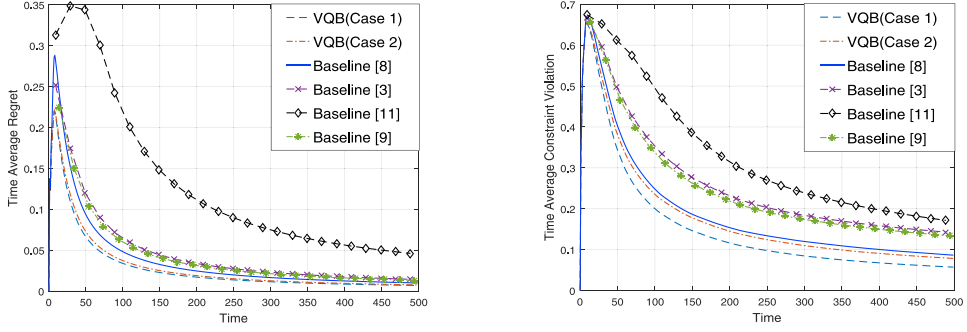


Fig. 1. (a) Evolutions of  $Regret(t)/t$ ; (b) Evolutions of  $Vio(t)/t$ .

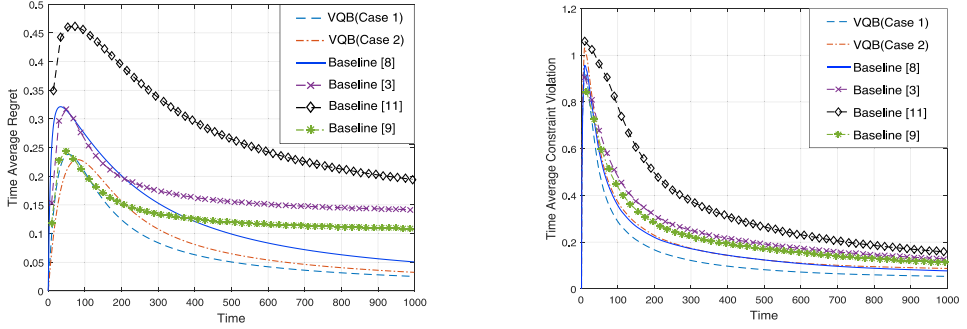


Fig. 2. (a) Evolutions of  $Regret(t)/t$ ; (b) Evolutions of  $Vio(t)/t$ .

We also consider the case when  $V_x = V_g = O(\sqrt{T})$ . To do this, we set  $B_t$  to be  $[-\frac{1}{2\sqrt{t}}, \frac{1}{2\sqrt{t}}]$ . In this case, the regret bounds of all baselines are at least the order of  $T$ . From Fig. 2, we notice that all methods can guarantee sublinear constraint violation in this case, which matches the theoretical results listed in Table 1. Fig. 2 also shows that VQB can achieve simultaneous sublinear regret and constraint violation, while other baselines [3,9,11] cannot, which matches their theoretical results. We observe that baseline [8] achieves a near sublinear regret in this setting, yet this may not always be the case due to its  $O(T)$  regret bound, or the performance bounds established by [8] may not be tight. Besides, the regrets of the VQB are better, which also coincides our theoretical bounds in this case.

## 6. Applications

In this section, we show several applications of our formulation to diverse problems across resource allocation and job scheduling. We emphasize that none of these applications would be possible without an algorithm that simultaneously achieves sublinear regret and constraint violations, which is not attainable with previous approaches.

### 6.1. Online network resource allocation

Within this subsection, we consider an online resource allocation problem over a cloud network [3,27]. The network consists of mapping nodes  $\mathcal{J} = \{1, \dots, J\}$  and data centers  $\mathcal{K} = \{1, \dots, K\}$ . We use a directed graph  $\mathcal{G} = (\mathcal{I}, \epsilon)$  to represent it, where  $\mathcal{I} = \mathcal{J} \cup \mathcal{K}$ ,  $|\mathcal{I}| = J + K$ , and  $|\epsilon| = E$ .  $\epsilon$  includes all the links which connect mapping nodes with data centers, and the “virtual” exogenous edges coming out of the data centers. At each time slot  $t$ , each mapping node  $j$  receives a data request  $b_t^j$  from exogenous user, and schedules  $x_t^{jk}$  workload to data center  $k$ . Each data center  $k$  serves workload  $y_t^k$  based on its source availability. We assume each node (including data center and mapping node) could buffer the unserved workloads into its local queue. Next we describe the workflow of the overall system, which is illustrated in Fig. 3. Specifically, at each time slot  $t$ , mapping node  $j$  has an exogenous workload  $b_t^j$  plus that stored in its local queue  $q_t^j$ , then it schedules workload  $x_t^{jk}$  to data center  $k$ . Data center  $k$  has a received workload of the amount of  $\sum_{j=1}^J x_t^{jk}$  plus that stored in its local queue  $q_t^{j+k}$ , and serves an amount of workload  $y_t^k$ .

We define a resource allocation vector  $\mathbf{x}_t = [x_t^{11}, \dots, x_t^{JK}, y_t^1, \dots, y_t^K]^T \in R_+^E$ , and load arrival vector  $\mathbf{b}_t = [b_t^1, \dots, b_t^J, 0, \dots, 0]^T \in R_+^I$  to represent the exogenous load arrival rates of all nodes at time slot  $t$ . We also define  $I \times E$  node-incidence matrix  $\mathbf{A}$ , where  $(i, j)$ -th entry  $\mathbf{A}_{(i,j)} = 1$  if link  $j$  enters node  $i$ , or  $\mathbf{A}_{(i,j)} = -1$  if link  $j$  leaves node  $i$ ,

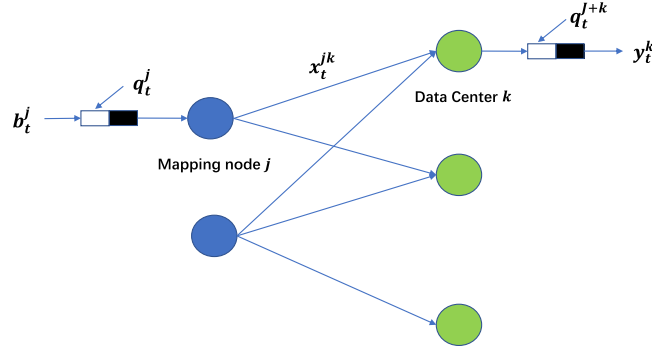


Fig. 3. The overview of the system.

otherwise  $\mathbf{A}_{(i,j)} = 0$ . Hence the vector  $\mathbf{A}\mathbf{x}_t + \mathbf{b}_t$  represents the aggregate workloads of all nodes. There is service residual at node  $i$  if  $(\mathbf{A}\mathbf{x}_t + \mathbf{b}_t)_i > 0$ , otherwise the current load of node  $i$  exceeds its service capacity. At each time slot  $t$ , the queue length vector of all nodes is given by  $\mathbf{q}_t = [q_t^1, \dots, q_t^{J+K}]^T$ , and its update rule of  $\mathbf{q}_t$  is  $\mathbf{q}_{t+1} = (\mathbf{q}_t + \mathbf{A}\mathbf{x}_t + \mathbf{b}_t)^+$ . We denote  $B_{jk}$  be the maximum bandwidth of link  $(j, k)$ , and  $C_k$  be the resource capacity of data center  $k$ . Thus the feasible set is  $\chi = \{x | 0 \leq x \leq c\}$ , where  $c = [B_{11}, \dots, B_{JK}, C_1, \dots, C_K]$ .

Here we formulate the accumulated cost of the overall system. We divide it into two parts, the one is power cost, the other is bandwidth cost. The power cost characterizes the energy price and renewable generation, and the bandwidth cost characterizes the transmission delay. The power cost of each data center  $k$  at time slot  $t$  is  $f_{t,k}(y_t^k)$ . The bandwidth cost of link  $(j, k)$  is  $f_{t,jk}(x_t^{jk})$ . Both of them are unknown before the resource allocation at time slot  $t$ . Hence at each time slot  $t$ , the instantaneous cost of the overall system is

$$f_t(\mathbf{x}_t) = \sum_{k \in \mathcal{K}} f_{t,k}(y_t^k) + \sum_{j \in \mathcal{J}} \sum_{k \in \mathcal{K}} f_{t,jk}(x_t^{jk}) \quad (31)$$

Our goal is to minimize the accumulated cost of the overall system while ensuring all workloads are served, shown in the following optimization problem  $\mathbf{P}_1$ :

$$\begin{aligned} \mathbf{P}_1 : \quad & \min_{\{\mathbf{x}_t \in \chi\}} \sum_{t=1}^T f_t(\mathbf{x}_t) \\ \text{s.t.} \quad & \mathbf{q}_{t+1} = [\mathbf{q}_t + \mathbf{A}\mathbf{x}_t + \mathbf{b}_t]^+, \forall t, \\ & \mathbf{q}_1 = \mathbf{q}_{T+1} = 0, \end{aligned} \quad (32)$$

where the initial queue length is given by  $\mathbf{q}_1$ , and  $\mathbf{q}_{T+1} = 0$  implies that all workloads should be served before the end of scheduling horizon  $T$ . However, solving  $\mathbf{P}_1$  is generally challenging using traditional methods since the future workload arrivals are not known a priori. Therefore, we could relax the first constraint in  $\mathbf{P}_1$  as follows:

$$\mathbf{q}_{T+1} \geq \mathbf{q}_T + \mathbf{A}\mathbf{x}_T + \mathbf{b}_T \geq \dots \geq \mathbf{q}_1 + \sum_{t=1}^T (\mathbf{A}\mathbf{x}_t + \mathbf{b}_t) \Rightarrow \sum_{t=1}^T (\mathbf{A}\mathbf{x}_t + \mathbf{b}_t) \leq \mathbf{q}_{T+1} - \mathbf{q}_1 = 0. \quad (33)$$

Then we transform  $\mathbf{P}_1$  into the following optimization problem  $\mathbf{P}_2$ :

$$\mathbf{P}_2 : \quad \min_{\{\mathbf{x}_t \in \chi\}} \sum_{t=1}^T f_t(\mathbf{x}_t) \quad \text{s.t.} \quad \sum_{t=1}^T (\mathbf{A}\mathbf{x}_t + \mathbf{b}_t) \leq 0, \quad (34)$$

which could be solved by our framework of OCO with long-term and time-varying constraints.

## 6.2. Online job scheduling

Within this subsection, we consider an online job scheduling problem [28–30], in which the computing cluster consists of multiple servers with heterogeneous computation resources. Specifically, consider a computing cluster, and it consists of  $M$  servers, which indexed from 1 to  $M$ . We assume server  $i$  has  $C_i$  CPU cores and can process multiple jobs simultaneously unless the total demand of its execution jobs exceeds  $C_i$ . Time is slotted and job  $j$  arrives the cluster at time slot  $a_j$ . The total number of jobs is  $N$ . Each incoming job joins a global queue managed by a scheduler, waiting to be assigned to an available server(s) for execution in the subsequent time slots. At the beginning of each time slot, the scheduler has to decide which job(s) to schedule and which sever(s) assigned to it(them). We assume job  $j$  requires  $d_j$  CPU cores to run

and  $p_j$  units of time to finish when its demand of  $d_j$  CPU cores are fully satisfied, where both  $d_j$  and  $p_j$  are integers and will be reported to the scheduler once job  $j$  arrives at the cluster. Thus the quantity  $v_j = p_j d_j$  is the volume of job  $j$ . We also assume that preemption and migration are allowed, i.e., a running job can be check-pointed, preempted, and then recovered on the same server or on a different server.

We denote  $u_j^i(t)$  be the number of CPU cores allocated to job  $j$  on server  $i$  at time-slot  $t$ . By time-division-multiplexing of CPU cores, any job could also be processed even if it is allocated fewer than  $d_j$  CPU cores, but it needs to take more than  $p_j$  time-slots to finish. We say job  $j$  finishes when its completion time  $c_j$  satisfies  $\sum_{t=a_j+1}^{c_j} \sum_{i=1}^M u_j^i(t) \geq d_j p_j$ , that is, its volume is fully served, and the flowtime of job  $j$  is  $c_j - a_j$ . We assume that any job  $j$  cannot benefit from the extra number of cores (i.e., it is allocated more than  $d_j$  cores). To avoid the waste of resources, we have  $\sum_{i=1}^M u_j^i(t) \leq d_j$ . The online scheduler strikes a balance between fairness and job latency. Hence, like [30], we adopt the  $l_k$  norm of job flowtime [30] to represent job's "cost". Then our goal is to minimize the sum of  $l_k$  norm of all jobs' flowtime while satisfying some constraints, shown in the following optimization problem  $\mathbf{P}_3$ :

$$\begin{aligned} \mathbf{P}_3 : \quad & \min_{\{u_j^i(t)\}} \sum_{j=1}^N (c_j - a_j)^k \\ \text{s.t.} \quad & \sum_{t=a_j+1}^{c_j} \sum_{i=1}^M u_j^i(t) \geq d_j p_j, \forall j, \\ & \sum_i u_j^i(t) \leq d_j, \forall j, t, \\ & \sum_{j:t>a_j} u_j^i(t) \leq C_i, \forall i, t, \\ & u_j^i(t) \in N, \end{aligned} \tag{35}$$

where the third constraint means the total number of allocated CPU cores on server  $i$  cannot exceed its capacity  $C_i$  at any time slot. However, it can be verified that  $\mathbf{P}_3$  is NP-hard since it is an integer programming problem. Hence, here we could adopt the approximation algorithm to solve  $\mathbf{P}_3$ . Specifically, we approximate  $l_k$  norm of flowtime  $(c_j - a_j)^k$  with its fractional job flowtime counterpart [31], that is,

$$(c_j - a_j)^k \approx \sum_{t=a_j+1}^M \sum_{i=1}^M ((t - a_j)^k / p_j + p_j^{k-1}) u_j^i(t) / d_j.$$

We define  $y_j(t) = \sum_i u_j^i(t) / d_j(t)$  be the total CPU cores rates allocated to job  $j$  at time slot  $t$ . We could transform  $\mathbf{P}_3$  into the following optimization problem  $\mathbf{P}_4$ :

$$\begin{aligned} \mathbf{P}_4 : \quad & \min_{\{u_j^i(t)\}} \sum_{t=1} \sum_{j:t \geq a_j+1} ((t - a_j)^k / p_j + p_j^{k-1}) y_j(t) \\ \text{s.t.} \quad & \sum_{t=a_j+1} y_j(t) \geq p_j, \forall j, \\ & 0 \leq y_j(t) \leq 1, \forall j, t, \\ & \sum_{j:t>a_j} d_j y_j(t) \leq \sum_{i=1}^M C_i, \forall t, \\ & d_j y_j(t) \in N, \forall j, t. \end{aligned} \tag{36}$$

We denote by  $OPT_{\mathbf{P}_3}^*$  and  $OPT_{\mathbf{P}_4}^*$  the optimal objective values of optimization problem  $\mathbf{P}_3$  and  $\mathbf{P}_4$  respectively, then by using the same argument as [31], we have  $OPT_{\mathbf{P}_4}^* \leq 2OPT_{\mathbf{P}_3}^*$ . Next we show that we could use the framework of OCO with long-term and time-varying constraints to solve problem  $\mathbf{P}_4$ .

**Solve  $\mathbf{P}_4$  using the framework of OCO with long-term and time-varying constraints.** We formulate the third constraint in  $\mathbf{P}_4$  as the short-term constraint which needs to be satisfied strictly at each time slot. We also notice that the first constraint could be formulated as the long-term constraint. Thus, we separate the first constraint in  $\mathbf{P}_4$  into each time slot constraint:

$$g_{t,j}(y_j(t)) = \frac{p_j}{T - a_j} - y_j(t) \leq 0, \tag{37}$$

where  $T$  is the predicted completion time for all jobs, which could be known or estimated ahead of time in many scenarios. The per time slot constraint (37) could be violated in some time slots but the accumulated constraint violations should



be controlled. We relax the integer constraint of  $y_j(t)$  and define the feasible set of it as:

$$\chi(t) = \{\mathbf{y} \mid \sum_{j:t>a_j} d_j y_j \leq \sum_{i=1}^M C_i, \mathbf{0} \leq y_j \leq 1.\}.$$

Indeed, we could still make the resultant online algorithm satisfies the integer constraint in the sequel. Therefore, the optimization problem  $\mathbf{P}_4$  could be transformed into the following optimization problem  $\mathbf{P}_5$ :

$$\begin{aligned} \mathbf{P}_5 : \quad & \min_{\mathbf{y}(t) \in \chi(t)} f_t(\mathbf{y}(t)) = \sum_{t=1}^T \sum_{j:t \geq a_j+1} ((t - a_j)^k / p_j + p_j^{k-1}) y_j(t) \\ & \text{s.t.} \quad \sum_{t=a_j+1}^T g_{t,j}(y_j(t)) \leq 0, \forall j, \end{aligned} \quad (38)$$

which could be solved by our framework of OCO with long-term and time-varying constraints. As stated before, although feasible set  $\chi$  is a time-varying set, our algorithm is valid and the corresponding theoretical results also hold.

## 7. Conclusion and future work

In this paper, we develop and analyze a novel algorithm for OCO with long term and time-varying constraints. To the best of our knowledge, our algorithm is the first parameter-free algorithm to simultaneously achieve sublinear dynamic regret and violation under common assumptions. We then extend our algorithm and analysis to some practical cases. For future work, It is a good direction to investigate sharper performance bounds for OCO with long-term and time-varying constraints. Moreover, whether incorporating other properties, like strong convexity and smoothness, can lead to better performance bounds is also an open question.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Proofs for Section 4.1

### A.1. Preliminary lemmas

**Lemma 12.** For any  $t \geq 1$ , we have

$$\sum_{i=1}^t \frac{1}{\sqrt{i}} \leq 2\sqrt{t} - 1. \quad (A.1)$$

**Lemma 13** (Proposition A.5 in [32]). Let  $R > 0$  and any real numbers  $x_1, x_2, \dots, x_T \in [0, R]$ , then we have

$$\sum_{t=1}^T \frac{x_t}{\sqrt{R + \sum_{i<t} x_i}} \leq 2 \sqrt{\sum_{t=1}^T x_t}. \quad (A.2)$$

### A.2. Proof of Lemma 1

**Proof.**

1. We prove the inequality (1) by induction. Assume  $\lambda(\tau) \geq \mathbf{0}$  holds for all  $\tau \in \{0, 1, \dots, t\}$ , then  $\forall k$  we consider two cases.

**Case 1:** If  $g_{k,t}(x_{t+1}) \geq 0$ , then we have

$$\lambda_k(t+1) = \max\{\lambda_k(t) + \gamma_t g_{k,t}(x_{t+1}), -\gamma_t g_{k,t}(x_{t+1})\} \geq \lambda_k(t) + \gamma_t g_{k,t}(x_{t+1}) \geq 0.$$

**Case 2:** If  $g_{k,t}(x_{t+1}) < 0$ , then we have

$$\lambda_k(t+1) = \max\{\lambda_k(t) + \gamma_t g_{k,t}(x_{t+1}), -\gamma_t g_{k,t}(x_{t+1})\} \geq -\gamma_t g_{k,t}(x_{t+1}) \geq 0.$$

Thus,  $\lambda(t) \geq \mathbf{0}$  holds for  $\forall t$ .

2. Since  $\lambda(t) = \max\{\lambda(t-1) + \gamma_{t-1}\mathbf{g}_{t-1}(x_t), -\gamma_{t-1}\mathbf{g}_{t-1}(x_t)\} \geq -\gamma_{t-1}\mathbf{g}_{t-1}(x_t)$ , then we can derive that  $\lambda(t) + \gamma_{t-1}\mathbf{g}_{t-1}(x_t) \geq 0, \forall t$ .
3. It is obvious that (3) holds if  $t = 1$ , then for  $t \geq 2$  and  $\forall k$  we consider two cases.

**Case 1:** If  $g_{k,t}(x_{t+1}) \geq 0$ , then we have

$$\begin{aligned}\lambda_k(t) &= \max\{\lambda_k(t-1) + \gamma_{t-1}g_{k,t-1}(x_t), -\gamma_{t-1}g_{k,t-1}(x_t)\} \\ &\geq \lambda_k(t-1) + \gamma_{t-1}g_{k,t-1}(x_t) \geq \gamma_{t-1}|g_{k,t-1}(x_t)|.\end{aligned}$$

**Case 2:** If  $g_{k,t}(x_{t+1}) < 0$ , then we have

$$\lambda_k(t) = \max\{\lambda_k(t-1) + \gamma_{t-1}g_{k,t-1}(x_t), -\gamma_{t-1}g_{k,t-1}(x_t)\} \geq -\gamma_{t-1}g_{k,t-1}(x_t) = \gamma_{t-1}|g_{k,t-1}(x_t)|.$$

Thus, we have  $\lambda_k(t) \geq \gamma_{t-1}|g_{k,t-1}(x_t)|, \forall t$ . Squaring both sides of it and summing over  $k$ , we obtain  $\|\lambda(t)\|^2 \geq \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t)\|^2$ , which is equivalent to the inequality (3).

4. Since  $\lambda(t) = \max\{\lambda(t-1) + \gamma_{t-1}\mathbf{g}_{t-1}(x_t), -\gamma_{t-1}\mathbf{g}_{t-1}(x_t)\} \geq \lambda(t-1) + \gamma_{t-1}\mathbf{g}_{t-1}(x_t)$ , then we have  $\gamma_{t-1}\mathbf{g}_{t-1}(x_t) \leq \lambda(t) - \lambda(t-1)$ . Furthermore,

$$\begin{aligned}\lambda_k(t) &= \max\{\lambda_k(t-1) + \gamma_{t-1}g_{k,t-1}(x_t), -\gamma_{t-1}g_{k,t-1}(x_t)\} \\ &\leq |\lambda_k(t-1)| + |\gamma_{t-1}g_{k,t-1}(x_t)| = \lambda_k(t-1) + \gamma_{t-1}|g_{k,t-1}(x_t)|.\end{aligned}$$

Squaring both sides of the above inequality and summing over  $k$ , we obtain

$$\|\lambda(t)\|^2 \leq \|\lambda(t-1) + \gamma_{t-1}\mathbf{g}_{t-1}(x_t)\|^2 \Leftrightarrow \|\lambda(t)\| \leq \|\lambda(t-1) + \gamma_{t-1}\mathbf{g}_{t-1}(x_t)\|.$$

By the triangle inequality we have  $\|\lambda(t)\| \leq \gamma_{t-1}\|\mathbf{g}_{t-1}(x_t)\| + \|\lambda(t-1)\|$ .

5. According to the above inequality  $\|\lambda(t)\| \leq \gamma_{t-1}\|\mathbf{g}_{t-1}(x_t)\| + \|\lambda(t-1)\|$ , we have

$$\begin{aligned}\|\lambda(t+1)\| &\leq \gamma_t\|\mathbf{g}_t(x_{t+1})\| + \|\lambda(t)\| \\ \Rightarrow \|\lambda(t+1)\|^2 &\leq \|\lambda(t)\|^2 + 2\gamma_t[\lambda(t)]^T \mathbf{g}_t(x_{t+1}) + \gamma_t^2 \|\mathbf{g}_t(x_{t+1})\|^2\end{aligned}$$

Rearranging terms yields the inequality (5).  $\square$

### A.3. Proof of Lemma 2

**Proof.** Since  $\nabla f_t(x_t)^T(x - x_t) + [\lambda(t) + \gamma_{t-1}\mathbf{g}_{t-1}(x_t)]^T(\gamma_t\mathbf{g}_t(x)) + \alpha_t\|x - x_t\|^2$  is a  $2\alpha_t$ -strong convex function with respect to  $x$  and  $x_{t+1}$  minimizes this expression over  $x$ , we have

$$\begin{aligned}\nabla f_t(x_t)^T(x_{t+1} - x_t) + [\lambda(t) + \gamma_{t-1}\mathbf{g}_{t-1}(x_t)]^T(\gamma_t\mathbf{g}_t(x_{t+1})) + \alpha_t\|x_{t+1} - x_t\|^2 \\ \leq \nabla f_t(x_t)^T(x_t^* - x_t) + [\lambda(t) + \gamma_{t-1}\mathbf{g}_{t-1}(x_t)]^T(\gamma_t\mathbf{g}_t(x_t^*)) + \alpha_t\|x_t^* - x_t\|^2 - \alpha_t\|x_{t+1} - x_t^*\|^2 \\ \stackrel{(a)}{\leq} \nabla f_t(x_t)^T(x_t^* - x_t) + \alpha_t\|x_t^* - x_t\|^2 - \alpha_t\|x_{t+1} - x_t^*\|^2,\end{aligned}\tag{A.3}$$

where (a) follows from the fact that  $\mathbf{g}_t(x_t^*) \leq 0$  and Lemma 13. Adding  $f_t(x_t)$  on both sides of (A.3) and using the convexity of  $f_t$ , we have

$$\begin{aligned}f_t(x_t) + \nabla f_t(x_t)^T(x_{t+1} - x_t) + [\lambda(t) + \gamma_{t-1}\mathbf{g}_{t-1}(x_t)]^T(\gamma_t\mathbf{g}_t(x_{t+1})) + \alpha_t\|x_{t+1} - x_t\|^2 \\ \leq f_t(x_t) + \nabla f_t(x_t)^T(x_t^* - x_t) + \alpha_t\|x_t^* - x_t\|^2 - \alpha_t\|x_{t+1} - x_t\|^2 \\ \leq f_t(x_t^*) + \alpha_t\|x_t^* - x_t\|^2 - \alpha_t\|x_{t+1} - x_t^*\|^2.\end{aligned}\tag{A.4}$$

Rearranging terms in (A.4), we have

$$\begin{aligned}f_t(x_t) + [\lambda(t)]^T(\gamma_t\mathbf{g}_t(x_{t+1})) - f_t(x_t^*) \\ \leq \alpha_t\|x_t^* - x_t\|^2 - \alpha_t\|x_{t+1} - x_t^*\|^2 - \alpha_t\|x_{t+1} - x_t\|^2 - \gamma_{t-1}\gamma_t[\mathbf{g}_{t-1}(x_t)]^T \mathbf{g}_t(x_{t+1}) - \nabla f_t(x_t)^T(x_{t+1} - x_t) \\ \stackrel{(a)}{\leq} \alpha_t\|x_t^* - x_t\|^2 - \alpha_t\|x_{t+1} - x_t^*\|^2 - \alpha_t\|x_{t+1} - x_t\|^2 - \gamma_{t-1}\gamma_t[\mathbf{g}_{t-1}(x_t)]^T \mathbf{g}_t(x_{t+1}) + \|\nabla f_t(x_t)\|\|x_{t+1} - x_t\| \\ \stackrel{(b)}{\leq} \alpha_t\|x_t^* - x_t\|^2 - \alpha_t\|x_{t+1} - x_t^*\|^2 - \alpha_t\|x_{t+1} - x_t\|^2 - \gamma_{t-1}\gamma_t[\mathbf{g}_{t-1}(x_t)]^T \mathbf{g}_t(x_{t+1}) + \frac{1}{2\delta}\|\nabla f_t(x_t)\|^2 + \frac{\delta}{2}\|x_{t+1} - x_t\|^2 \\ \stackrel{(c)}{\leq} \alpha_t\|x_t^* - x_t\|^2 - \alpha_t\|x_{t+1} - x_t^*\|^2 - \alpha_t\|x_{t+1} - x_t\|^2 - \gamma_{t-1}\gamma_t[\mathbf{g}_{t-1}(x_t)]^T \mathbf{g}_t(x_{t+1}) + \frac{1}{2\delta}G^2 + \frac{\delta}{2}\|x_{t+1} - x_t\|^2,\end{aligned}\tag{A.5}$$

where (a) holds by the Cauchy-Schwarz inequality; (b) comes from the AM-GM inequality; (c) holds due to Assumption 1. Based on Assumption 1, we note that

$$\begin{aligned}\|x_t^* - x_t\|^2 - \|x_{t+1} - x_t^*\|^2 &= \|x_t^* - x_t\|^2 - \|x_{t+1} - x_{t+1}^* + x_{t+1}^* - x_t^*\|^2 \\ &\leq \|x_t^* - x_t\|^2 - \|x_{t+1} - x_{t+1}^*\|^2 - \|x_{t+1}^* - x_t^*\|^2 - 2\|x_{t+1} - x_{t+1}^*\|\|x_{t+1}^* - x_t^*\| \\ &\leq \|x_t^* - x_t\|^2 - \|x_{t+1} - x_{t+1}^*\|^2 + 4R\|x_{t+1}^* - x_t^*\|.\end{aligned}\tag{A.6}$$

And

$$\begin{aligned}
 & -[\mathbf{g}_{t-1}(x_t)]^T \mathbf{g}_t(x_{t+1}) = -\frac{1}{2} \|\mathbf{g}_{t-1}(x_t)\|^2 - \frac{1}{2} \|\mathbf{g}_t(x_{t+1})\|^2 + \frac{1}{2} \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_{t+1})\|^2 \\
 & = -\frac{1}{2} \|\mathbf{g}_{t-1}(x_t)\|^2 - \frac{1}{2} \|\mathbf{g}_t(x_{t+1})\|^2 + \frac{1}{2} \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t) + \mathbf{g}_t(x_t) - \mathbf{g}_t(x_{t+1})\|^2 \\
 & \stackrel{(a)}{\leq} -\frac{1}{2} \|\mathbf{g}_{t-1}(x_t)\|^2 - \frac{1}{2} \|\mathbf{g}_t(x_{t+1})\|^2 + \frac{1}{2} [2\|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\|^2 + 2\|\mathbf{g}_t(x_t) - \mathbf{g}_t(x_{t+1})\|^2] \\
 & \stackrel{(b)}{\leq} -\frac{1}{2} \|\mathbf{g}_{t-1}(x_t)\|^2 - \frac{1}{2} \|\mathbf{g}_t(x_{t+1})\|^2 + \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\|^2 + \beta^2 \|x_{t+1} - x_t\|^2,
 \end{aligned} \tag{A.7}$$

where (a) follows from the AM-GM inequality; (b) holds by the Lipschitz continuity of  $\mathbf{g}_t$  (Assumption 1). Substituting (A.6) and (A.7) into (A.5) we obtain

$$\begin{aligned}
 & f_t(x_t) + [\lambda(t)]^T (\gamma \mathbf{g}_t(x_{t+1})) - f_t(x_t^*) \\
 & \leq \alpha_t \|x_t^* - x_t\|^2 - \|x_{t+1} - x_{t+1}^*\|^2 + 4R \|x_{t+1}^* - x_t^*\| + (\beta^2 \gamma_t \gamma_{t-1} + \frac{\delta}{2} - \alpha_t) \|x_{t+1} - x_t\|^2 \\
 & + \frac{1}{2\delta} G^2 - \frac{1}{2} \gamma_{t-1} \gamma_t \|\mathbf{g}_{t-1}(x_t)\|^2 - \frac{1}{2} \gamma_{t-1} \gamma_t \|\mathbf{g}_t(x_{t+1})\|^2 + \gamma_{t-1} \gamma_t \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\|^2 \\
 & \stackrel{(a)}{\leq} \alpha_t \|x_t^* - x_t\|^2 - \alpha_{t+1} \|x_{t+1} - x_{t+1}^*\|^2 + 4R \alpha_t \|x_{t+1}^* - x_t^*\| + (\beta^2 \gamma_{t-1} + \frac{\delta}{2} - \alpha_t) \|x_{t+1} - x_t\|^2 \\
 & + \frac{1}{2\delta} G^2 - \frac{1}{2} \gamma_{t-1} \gamma_t \|\mathbf{g}_{t-1}(x_t)\|^2 - \frac{1}{2} \gamma_{t-1} \gamma_t \|\mathbf{g}_t(x_{t+1})\|^2 + \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\|^2,
 \end{aligned} \tag{A.8}$$

where (a) comes from the fact that both  $\{\alpha_t\}$  and  $\{\gamma_t\}$  are non-increasing sequence. According to Lemma 1 and adding Lyapunov drift term on both sides of (A.8) yields

$$\begin{aligned}
 & f_t(x_t) + \Delta(t) - f_t(x_t^*) \\
 & \leq \alpha_t \|x_t^* - x_t\|^2 - \alpha_{t+1} \|x_{t+1} - x_{t+1}^*\|^2 + 4R \alpha_t \|x_{t+1}^* - x_t^*\| + (\beta^2 \gamma_{t-1} + \frac{\delta}{2} - \alpha_t) \|x_{t+1} - x_t\|^2 \\
 & + \frac{1}{2\delta} G^2 - \frac{1}{2} \gamma_{t-1} \gamma_t \|\mathbf{g}_{t-1}(x_t)\|^2 + (\gamma_t^2 - \frac{1}{2} \gamma_{t-1} \gamma_t) \|\mathbf{g}_t(x_{t+1})\|^2 + \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\|^2 \\
 & \stackrel{(a)}{\leq} \alpha_t \|x_t^* - x_t\|^2 - \alpha_{t+1} \|x_{t+1} - x_{t+1}^*\|^2 + 4R \alpha_t \|x_{t+1}^* - x_t^*\| + (\beta^2 \gamma_{t-1} + \frac{\delta}{2} - \alpha_t) \|x_{t+1} - x_t\|^2 \\
 & + \frac{1}{2\delta} G^2 - \frac{1}{2} \gamma_{t-1} \gamma_t \|\mathbf{g}_{t-1}(x_t)\|^2 + \frac{1}{2} \gamma_t \gamma_{t+1} \|\mathbf{g}_t(x_{t+1})\|^2 + \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\|^2,
 \end{aligned} \tag{A.9}$$

where (a) is due to the fact that  $2\gamma_t \leq \gamma_{t-1} + \gamma_{t+1}$ . This completes the proof.  $\square$

#### A.4. Proof of Lemma 3

**Proof.** According to Lemma 2, taking a telescoping sum over  $t = 1, \dots, T$ , we obtain

$$\begin{aligned}
 & \sum_{t=1}^T f_t(x_t) + \sum_{t=1}^T \Delta(t) \leq \sum_{t=1}^T f_t(x_t^*) + 4R \sum_{t=1}^T \alpha_t \|x_{t+1}^* - x_t^*\| + \alpha_1 \|x_1^* - x_1\|^2 + \frac{TG^2}{2\delta} \\
 & + \frac{1}{2} \gamma_T \gamma_{T+1} \|\mathbf{g}_T(x_{T+1})\|^2 + \sum_{t=1}^T \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\|^2 + \sum_{t=1}^T (\beta^2 \gamma_{t-1}^2 + \frac{\delta}{2} - \alpha_t) \|x_{t+1} - x_t\|^2 \\
 & \stackrel{(a)}{\leq} \sum_{t=1}^T f_t(x_t^*) + 4R \sum_{t=1}^T \alpha_t \|x_{t+1}^* - x_t^*\| + \alpha_1 R^2 + \frac{TG^2}{2\delta} \\
 & + \frac{1}{2} \gamma_T \gamma_{T+1} \|\mathbf{g}_T(x_{T+1})\|^2 + 2F \sum_{t=1}^T \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\|,
 \end{aligned} \tag{A.10}$$

where (a) holds by Assumption 1 and the fact that  $\alpha_t \geq \beta^2 \gamma^2 + \frac{\delta}{2}$ . Rearranging terms yields

$$\begin{aligned}
 & \sum_{t=1}^T f_t(x_t) \leq \sum_{t=1}^T f_t(x_t^*) + 4R \sum_{t=1}^T \alpha_t \|x_{t+1}^* - x_t^*\| + \alpha_1 R^2 + \frac{TG^2}{2\delta} \\
 & + \frac{1}{2} \gamma_T \gamma_{T+1} \|\mathbf{g}_T(x_{T+1})\|^2 + 2F \sum_{t=1}^T \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\| + \frac{1}{2} \|\lambda(1)\|^2.
 \end{aligned} \tag{A.11}$$

It completes the proof.  $\square$

## A.5. Proof of Lemma 4

**Proof.** According to Lemma 2, taking a telescoping sum over  $t = 1, \dots, T - 1$  and using the fact that  $\alpha_t \geq \beta^2 \gamma_{t-1}^2 + \frac{\delta}{2}$ , we obtain

$$\begin{aligned} \sum_{t=1}^{T-1} f_t(x_t) + \sum_{t=1}^{T-1} \Delta(t) &\leq \sum_{t=1}^{T-1} f_t(x_t^*) + 4R \sum_{t=1}^{T-1} \alpha_t \|x_{t+1}^* - x_t^*\| + \alpha_1 R^2 \\ &+ \frac{(T-1)G^2}{2\delta} + \frac{1}{2} \gamma_{T-1} \gamma_T \|\mathbf{g}_{T-1}(x_T)\|^2 + 2F \sum_{t=1}^{T-1} \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\|. \end{aligned} \quad (\text{A.12})$$

Rearranging terms and multiplying both sides by 2 yields

$$\begin{aligned} \|\lambda(t)\|^2 &\leq 2\left(\sum_{t=1}^{T-1} f_t(x_t^*) - \sum_{t=1}^{T-1} f_t(x_t)\right) + 8R \sum_{t=1}^{T-1} \alpha_t \|x_{t+1}^* - x_t^*\| + 2\alpha_1 R^2 \\ &+ \frac{(T-1)G^2}{\delta} + \gamma_{T-1} \gamma_T \|\mathbf{g}_{T-1}(x_T)\|^2 + 4F \sum_{t=1}^{T-1} \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\| \\ &\stackrel{(a)}{\leq} 2F(T-1) + 8R \sum_{t=1}^{T-1} \alpha_t \|x_{t+1}^* - x_t^*\| + 2\alpha_1 R^2 + \frac{(T-1)G^2}{\delta} \\ &+ \gamma_{T-1}^2 \|\mathbf{g}_{T-1}(x_T)\|^2 + 4F \sum_{t=1}^{T-1} \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\|, \end{aligned} \quad (\text{A.13})$$

where (a) holds by Assumption 1 and the fact that  $\gamma_T \leq \gamma_{T-1}$ . Taking the square root of both sides and using the fact that  $\sqrt{\sum_i a_i} \leq \sum_i \sqrt{a_i}$ ,  $\forall a_i \geq 0$ . Then we obtain

$$\begin{aligned} \|\lambda(t)\| &\leq \sqrt{2F(T-1)} + 2 \sqrt{2R \sum_{t=1}^{T-1} \alpha_t \|x_{t+1}^* - x_t^*\|} + \sqrt{2\alpha_1 R^2} \\ &+ \sqrt{\frac{(T-1)G^2}{\delta}} + \gamma_{T-1} \|\mathbf{g}_{T-1}(x_T)\| + 2 \sqrt{F \sum_{t=1}^{T-1} \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\|}. \end{aligned} \quad (\text{A.14})$$

It completes the proof.  $\square$

## A.6. Proof of Lemma 5

**Proof.** According to Lemma 1, we have  $\gamma_{t-1} \mathbf{g}_{t-1,k}(x_t) \leq \lambda_k(t) - \lambda_k(t-1) \Leftrightarrow \mathbf{g}_{t-1,k}(x_t) \leq \frac{\lambda_k(t)}{\gamma_{t-1}} - \frac{\lambda_k(t-1)}{\gamma_{t-1}}$ . Adding  $\mathbf{g}_{t,k}(x_t)$  on both sides of it and summing over  $t$  yields

$$\begin{aligned} \sum_{t=1}^T \mathbf{g}_{t-1,k}(x_t) + \mathbf{g}_{t,k}(x_t) &\leq \sum_{t=1}^T \mathbf{g}_{t,k}(x_t) + \sum_{t=1}^T \frac{\lambda_k(t)}{\gamma_{t-1}} - \frac{\lambda_k(t-1)}{\gamma_{t-1}} \\ &\stackrel{(a)}{\leq} \sum_{t=1}^T \mathbf{g}_{t,k}(x_t) + \sum_{t=1}^T \frac{\lambda_k(t)}{\gamma_t} - \frac{\lambda_k(t-1)}{\gamma_{t-1}} \leq \sum_{t=1}^T \mathbf{g}_{t,k}(x_t) + \frac{\lambda_k(T)}{\gamma_T} \\ &\Rightarrow \sum_{t=1}^T \mathbf{g}_{t,k}(x_t) \leq \sum_{t=1}^T \mathbf{g}_{t,k}(x_t) - \mathbf{g}_{t-1,k}(x_t) + \frac{\lambda_k(T)}{\gamma_T} \\ &\stackrel{(b)}{\leq} \sum_{t=1}^T \max\{|\mathbf{g}_{t,k}(x) - \mathbf{g}_{t-1,k}(x)|\} + \frac{\lambda_k(T)}{\gamma_T} \stackrel{(c)}{\leq} \frac{\|\lambda(T)\|}{\gamma_T} + V_{\mathbf{g}}, \quad \forall k = 1, 2, \dots, K, \end{aligned} \quad (\text{A.15})$$

where (a) comes from the fact that  $\gamma_t$  is non-increasing with respect to  $t$ ; (b) follows from the fact that  $|\mathbf{g}_{t,k}(x_t) - \mathbf{g}_{t-1,k}(x_t)| \leq \max\{|\mathbf{g}_{t,k}(x) - \mathbf{g}_{t-1,k}(x)|\}$ ; (c) is due to the fact that  $|\mathbf{g}_{t,k}(x) - \mathbf{g}_{t-1,k}(x)| \leq \|\mathbf{g}_t(x) - \mathbf{g}_{t-1}(x)\|_2$  and  $\lambda_k(T) \leq \|\lambda(T)\|$ . It completes the proof.  $\square$

## A.7. Proof of Lemma 6

**Proof.** Combining Lemma 4 with Lemma 5, we have

$$\begin{aligned} \sum_{t=1}^T g_{t,k}(x_t) &\leq \frac{\|\lambda(T)\|}{\gamma_T} + V_g \leq \frac{2}{\gamma_T} \sqrt{F(T-1)} + \frac{1}{\gamma_T} \sqrt{2\alpha_1 R^2} + \frac{\gamma_{T-1}}{\gamma_T} \|\mathbf{g}_{T-1}(x_T)\| + \frac{G}{\gamma_T} \sqrt{\frac{T-1}{\delta}} \\ &+ \frac{2}{\gamma_T} \sqrt{2R \sum_{t=1}^{T-1} \alpha_t \|x_{t+1}^* - x_t^*\|} + \frac{2}{\gamma_T} \sqrt{F \sum_{t=1}^{T-1} \gamma_{t-1}^2 \|\mathbf{g}_{t-1}(x_t) - \mathbf{g}_t(x_t)\|} + V_g. \end{aligned} \quad (\text{A.16})$$

It completes the proof.  $\square$

## Appendix B. Proofs for Section 4.2

Note that both  $\{\alpha_t\}$  and  $\{\gamma_t\}$  are constant sequences in Algorithm 2, thus here we omit the subscript  $t$ . In the following, we give the complete proof of all lemmas in Section 4.2.

## B.1. Proof of Lemma 8

**Proof.** We conduct a similar derivation process as the proof of Lemma 1, note that  $\nabla f_t(x_t)^T(x - x_t) + [\lambda(t) + \gamma \mathbf{g}_t(x_t)]^T(\gamma \mathbf{g}_t(x)) + \alpha \|x - x_t\|^2$  is a  $2\alpha$ -strong convex function with respect to  $x$  and  $x_{t+1}$  minimizes this expression over  $x$ , we have

$$\begin{aligned} &\nabla f_t(x_t)^T(x_{t+1} - x_t) + [\lambda(t) + \gamma \mathbf{g}_t(x_t)]^T(\gamma \mathbf{g}_t(x_{t+1})) + \alpha \|x_{t+1} - x_t\|^2 \\ &\leq \nabla f_t(x_t)^T(x_t^* - x_t) + [\lambda(t) + \gamma \mathbf{g}_t(x_t)]^T(\gamma \mathbf{g}_t(x_t^*)) + \alpha \|x_t^* - x_t\|^2 - \alpha \|x_{t+1} - x_t^*\|^2 \\ &\stackrel{(a)}{\leq} \nabla f_t(x_t)^T(x_t^* - x_t) + \alpha \|x_t^* - x_t\|^2 - \alpha \|x_{t+1} - x_t^*\|^2, \end{aligned} \quad (\text{B.1})$$

where (a) follows from the fact that  $\mathbf{g}_t(x_t^*) \leq \mathbf{0}$  and Lemma 8. Next we add  $f_t(x_t)$  on both sides of (B.1) and use the convexity of  $f_t$ , then we obtain

$$\begin{aligned} &f_t(x_t) + \nabla f_t(x_t)^T(x_{t+1} - x_t) + [\lambda(t) + \gamma \mathbf{g}_t(x_t)]^T(\gamma \mathbf{g}_t(x_{t+1})) + \alpha \|x_{t+1} - x_t\|^2 \\ &\leq f_t(x_t) + \nabla f_t(x_t)^T(x_t^* - x_t) + \alpha \|x_t^* - x_t\|^2 - \alpha \|x_{t+1} - x_t^*\|^2 \\ &\leq f_t(x_t^*) + \alpha \|x_t^* - x_t\|^2 - \alpha \|x_{t+1} - x_t^*\|^2. \end{aligned} \quad (\text{B.2})$$

Rearranging terms in (B.2), we have

$$\begin{aligned} &f_t(x_t) + [\lambda(t)]^T(\gamma \mathbf{g}_t(x_{t+1})) \\ &\leq f_t(x_t^*) + \alpha \|x_t^* - x_t\|^2 - \alpha \|x_{t+1} - x_t^*\|^2 - \alpha \|x_{t+1} - x_t\|^2 - \gamma^2 [\mathbf{g}_t(x_t)]^T \mathbf{g}_t(x_{t+1}) - \nabla f_t(x_t)^T(x_{t+1} - x_t) \\ &\stackrel{(a)}{\leq} f_t(x_t^*) + \alpha \|x_t^* - x_t\|^2 - \alpha \|x_{t+1} - x_t^*\|^2 - \alpha \|x_{t+1} - x_t\|^2 - \gamma^2 [\mathbf{g}_t(x_t)]^T \mathbf{g}_t(x_{t+1}) + \|\nabla f_t(x_t)\| \|x_{t+1} - x_t\| \\ &\stackrel{(b)}{\leq} f_t(x_t^*) + \alpha \|x_t^* - x_t\|^2 - \alpha \|x_{t+1} - x_t^*\|^2 - \alpha \|x_{t+1} - x_t\|^2 - \gamma^2 [\mathbf{g}_t(x_t)]^T \mathbf{g}_t(x_{t+1}) \\ &+ \frac{1}{2\delta} \|\nabla f_t(x_t)\|^2 + \frac{\delta}{2} \|x_{t+1} - x_t\|^2 \\ &\stackrel{(c)}{\leq} f_t(x_t^*) + \alpha \|x_t^* - x_t\|^2 - \alpha \|x_{t+1} - x_t^*\|^2 - \alpha \|x_{t+1} - x_t\|^2 - \gamma^2 [\mathbf{g}_t(x_t)]^T \mathbf{g}_t(x_{t+1}) \\ &+ \frac{1}{2\delta} G^2 + \frac{\delta}{2} \|x_{t+1} - x_t\|^2. \end{aligned} \quad (\text{B.3})$$

where (a) holds by the Cauchy–Schwarz inequality; (b) comes from the AM–GM inequality; (c) holds due to Assumption 1. Recall that we have the following inequality stated before

$$\|x_t^* - x_t\|^2 - \|x_{t+1} - x_t^*\|^2 \leq \|x_t^* - x_t\|^2 - \|x_{t+1} - x_{t+1}^*\|^2 + 4R \|x_{t+1}^* - x_t^*\|. \quad (\text{B.4})$$

Furthermore, we have

$$\begin{aligned} -[\mathbf{g}_t(x_t)]^T \mathbf{g}_t(x_{t+1}) &= -\frac{1}{2} \|\mathbf{g}_t(x_t)\|^2 - \frac{1}{2} \|\mathbf{g}_t(x_{t+1})\|^2 + \frac{1}{2} \|\mathbf{g}_t(x_t) - \mathbf{g}_t(x_{t+1})\|^2 \\ &\stackrel{(a)}{\leq} -\frac{1}{2} \|\mathbf{g}_t(x_t)\|^2 - \frac{1}{2} \|\mathbf{g}_t(x_{t+1})\|^2 + \frac{1}{2} \beta^2 \|x_{t+1} - x_t\|^2, \end{aligned} \quad (\text{B.5})$$

where (a) holds by the Lipschitz continuity of  $\mathbf{g}_t$  (Assumption 1). Substituting (B.4) and (B.5) into (B.3) we obtain

$$\begin{aligned}
& f_t(x_t) + [\boldsymbol{\lambda}(t)]^T (\gamma \mathbf{g}_t(x_{t+1})) \\
& \leq f_t(x_t^*) + \alpha_t [\|x_t^* - x_t\|^2 - \|x_{t+1} - x_{t+1}^*\|^2 + 4R\|x_{t+1}^* - x_t^*\|] + \left(\frac{1}{2}\beta^2\gamma^2 + \frac{\delta}{2} - \alpha\right) \|x_{t+1} - x_t\|^2 \\
& + \frac{1}{2\delta}G^2 - \frac{1}{2}\gamma^2\|\mathbf{g}_t(x_t)\|^2 - \frac{1}{2}\gamma^2\|\mathbf{g}_t(x_{t+1})\|^2 \\
& \stackrel{(a)}{\leq} f_t(x_t^*) + \alpha\|x_t^* - x_t\|^2 - \alpha\|x_{t+1} - x_{t+1}^*\|^2 + 4R\alpha\|x_{t+1}^* - x_t^*\| + \frac{1}{2\delta}G^2 - \frac{1}{2}\gamma^2\|\mathbf{g}_t(x_t)\|^2 - \frac{1}{2}\gamma^2\|\mathbf{g}_t(x_{t+1})\|^2,
\end{aligned} \tag{B.6}$$

where (a) holds since  $\alpha \geq \frac{1}{2}\beta^2\gamma^2 + \frac{1}{2}\delta$ . Based on Lemma 7, adding Lyapunov drift term on both sides of (B.6) and rearranging terms yields

$$\begin{aligned}
& f_t(x_t) + \Delta(t) \\
& \leq f_t(x_t^*) + \alpha\|x_t^* - x_t\|^2 - \alpha\|x_{t+1} - x_{t+1}^*\|^2 + 4R\alpha\|x_{t+1}^* - x_t^*\| + \frac{1}{2\delta}G^2 + \frac{1}{2}\gamma^2\|\mathbf{g}_{t+1}(x_{t+1})\|^2 \\
& - \frac{1}{2}\gamma^2\|\mathbf{g}_t(x_{t+1})\|^2 + \frac{1}{2}\gamma^2\|\mathbf{g}_{t+1}(x_{t+1})\|^2 - \frac{1}{2}\gamma^2\|\mathbf{g}_t(x_t)\|^2 + \gamma[\boldsymbol{\lambda}(t)]^T (\mathbf{g}_{t+1}(x_{t+1}) - \mathbf{g}_t(x_{t+1})) \\
& \stackrel{(a)}{\leq} f_t(x_t^*) + \alpha\|x_t^* - x_t\|^2 - \alpha\|x_{t+1} - x_{t+1}^*\|^2 + 4R\alpha\|x_{t+1}^* - x_t^*\| + \frac{1}{2\delta}G^2 \\
& + \gamma^2F\|\mathbf{g}_{t+1}(x_{t+1}) - \mathbf{g}_t(x_{t+1})\| + \frac{1}{2}\gamma^2\|\mathbf{g}_{t+1}(x_{t+1})\|^2 - \frac{1}{2}\gamma^2\|\mathbf{g}_t(x_t)\|^2 + \gamma[\boldsymbol{\lambda}(t)]^T (\mathbf{g}_{t+1}(x_{t+1}) - \mathbf{g}_t(x_{t+1})) \\
& \stackrel{(b)}{\leq} f_t(x_t^*) + \alpha\|x_t^* - x_t\|^2 - \alpha\|x_{t+1} - x_{t+1}^*\|^2 + 4R\alpha\|x_{t+1}^* - x_t^*\|^2 + \frac{1}{2\delta}G^2 \\
& + \gamma^2F\|\mathbf{g}_{t+1}(x_{t+1}) - \mathbf{g}_t(x_{t+1})\| + \frac{1}{2}\gamma^2\|\mathbf{g}_{t+1}(x_{t+1})\|^2 - \frac{1}{2}\gamma^2\|\mathbf{g}_t(x_t)\|^2 + \gamma\|\boldsymbol{\lambda}(t)\|\|\mathbf{g}_{t+1}(x_{t+1}) - \mathbf{g}_t(x_{t+1})\|,
\end{aligned} \tag{B.7}$$

where (a) is due to  $\|\mathbf{g}_{t+1}(x_{t+1}) - \mathbf{g}_t(x_{t+1})\| \leq 2F$ ; (b) holds by the Cauchy-Schwarz inequality. It completes the proof.  $\square$

### B.2. Proof of Lemma 9

**Proof.** According to Lemma 8, taking a telescoping sum over  $t = 1, \dots, T$ , we obtain

$$\begin{aligned}
& \sum_{t=1}^T f_t(x_t) + \sum_{t=1}^T \Delta(t) \leq \sum_{t=1}^T f_t(x_t^*) + \alpha\|x_1^* - x_1\|^2 + 4R \sum_{t=1}^T \alpha\|x_{t+1}^* - x_t^*\| \\
& + \frac{TG^2}{2\delta} + \gamma^2FV_{\mathbf{g}} + \frac{1}{2}\gamma^2\|\mathbf{g}_{T+1}(x_{T+1})\|^2 - \frac{1}{2}\gamma^2\|\mathbf{g}_1(x_1)\|^2 + \gamma \sum_{t=1}^T \|\boldsymbol{\lambda}(t)\|\|\mathbf{g}_{t+1}(x_{t+1}) - \mathbf{g}_t(x_{t+1})\| \\
& \leq \sum_{t=1}^T f_t(x_t^*) + \alpha\|x_1^* - x_1\|^2 + 4R\alpha V_x + \frac{TG^2}{2\delta} + \gamma^2FV_{\mathbf{g}} + \gamma \max_t \|\boldsymbol{\lambda}(t)\|V_{\mathbf{g}}.
\end{aligned} \tag{B.8}$$

Here we define  $\mathbf{g}_{T+1} = \mathbf{g}_T$ , rearranging terms of the above inequality yields

$$\begin{aligned}
& \sum_{t=1}^T f_t(x_t) \leq \sum_{t=1}^T f_t(x_t^*) + \alpha\|x_1^* - x_1\|^2 + 4R\alpha V_x + \frac{TG^2}{2\delta} + \gamma^2FV_{\mathbf{g}} + \frac{1}{2}\|\boldsymbol{\lambda}(1)\|^2 - \frac{1}{2}\|\boldsymbol{\lambda}(T+1)\|^2 + \gamma \max_t \|\boldsymbol{\lambda}(t)\|V_{\mathbf{g}} \\
& \leq \sum_{t=1}^T f_t(x_t^*) + \alpha\|x_1^* - x_1\|^2 + 4R\alpha V_x + \frac{TG^2}{2\delta} + \gamma^2FV_{\mathbf{g}} + \frac{1}{2}\|\boldsymbol{\lambda}(1)\|^2 + \gamma \max_t \|\boldsymbol{\lambda}(t)\|V_{\mathbf{g}}.
\end{aligned} \tag{B.9}$$

This completes the proof.  $\square$

### B.3. Proof of Lemma 10

**Proof.** Recall that Lemma 7 implies that  $\gamma \mathbf{g}_{t,k}(x_t) \leq \lambda_k(t) - \lambda_k(t-1)$ ,  $\forall k$ . Telescoping it over  $t$  yields

$$\begin{aligned}
& \sum_{t=1}^T \gamma \mathbf{g}_{t,k}(x_t) \leq \lambda_k(T) - \lambda_k(0), \quad \forall k \in \{1, 2, \dots, K\}. \\
& \Rightarrow \sum_{t=1}^T \mathbf{g}_{t,k}(x_t) \leq \frac{\lambda_k(T)}{\gamma} \leq \frac{\|\boldsymbol{\lambda}(T)\|}{\gamma}, \quad \forall k \in \{1, 2, \dots, K\}.
\end{aligned} \tag{B.10}$$

It completes the proof.  $\square$

## B.4. Proof of Lemma 11

**Proof.** According to the strong convexity of  $\nabla f_t(x_t)^T(x - x_t) + [\lambda(t) + \gamma \mathbf{g}_t(x_t)]^T(\gamma \mathbf{g}_t(x)) + \alpha \|x - x_t\|^2$  with respect to  $x$  and recalling that  $x_{t+1}$  minimizes this expression over  $\mathcal{X}$ , we have

$$\begin{aligned}
& \nabla f_t(x_t)^T(x_{t+1} - x_t) + [\lambda(t) + \gamma \mathbf{g}_t(x_t)]^T(\gamma \mathbf{g}_t(x_{t+1})) + \alpha \|x_{t+1} - x_t\|^2 \\
& \leq \nabla f_t(x_t)^T(\hat{x} - x_t) + [\lambda(t) + \gamma \mathbf{g}_t(x_t)]^T(\gamma \mathbf{g}_t(\hat{x})) + \alpha \|\hat{x} - x_t\|^2 - \alpha \|x_{t+1} - \hat{x}\|^2 \\
& \stackrel{(a)}{\leq} \nabla f_t(x_t)^T(\hat{x} - x_t) - \gamma \epsilon \sum_{k=1}^K [\lambda_k(t) + \gamma \mathbf{g}_{t,k}(x_t)] + \alpha \|\hat{x} - x_t\|^2 - \alpha \|x_{t+1} - \hat{x}\|^2 \\
& \stackrel{(b)}{=} \nabla f_t(x_t)^T(\hat{x} - x_t) - \gamma \epsilon \|\lambda(t) + \gamma \mathbf{g}_t(x_t)\|_1 + \alpha \|\hat{x} - x_t\|^2 - \alpha \|x_{t+1} - \hat{x}\|^2 \\
& \stackrel{(c)}{\leq} \nabla f_t(x_t)^T(\hat{x} - x_t) - \gamma \epsilon \|\lambda(t) + \gamma \mathbf{g}_t(x_t)\| + \alpha \|\hat{x} - x_t\|^2 - \alpha \|x_{t+1} - \hat{x}\|^2 \\
& \stackrel{(d)}{\leq} \nabla f_t(x_t)^T(\hat{x} - x_t) - \gamma \epsilon [\|\lambda(t)\| - \|\gamma \mathbf{g}_t(x_t)\|] + \alpha \|\hat{x} - x_t\|^2 - \alpha \|x_{t+1} - \hat{x}\|^2,
\end{aligned} \tag{B.11}$$

where (a) is due to the Slater condition (Assumption 2); (b) holds since  $\lambda_k(t) + \gamma g_{t,k}(x_t) \geq 0$ ,  $\forall k$ ; (c) holds due to the fact that  $\|x\|_1 \geq \|x\|$  for any vector  $x \in \mathcal{X}$ ; (d) holds by the triangle inequality  $\|u - v\| \geq \|u\| - \|v\|$ ,  $\forall u, v \in \mathcal{X}$ . Based on Lemma 7, we add Lyapunov drift term  $\Delta(t)$  on both sides and rearranging terms yields

$$\begin{aligned}
\Delta(t) & \leq \nabla f_t(x_t)^T(\hat{x} - x_t) - \nabla f_t(x_t)^T(x_{t+1} - x_t) - \gamma^2 \mathbf{g}_t(x_t)^T \mathbf{g}_t(x_{t+1}) - \gamma \epsilon [\|\lambda(t)\| - \|\gamma \mathbf{g}_t(x_t)\|] \\
& + \alpha \|\hat{x} - x_t\|^2 - \alpha \|x_{t+1} - \hat{x}\|^2 - \alpha \|x_{t+1} - x_t\|^2 + \gamma^2 \|\mathbf{g}_{t+1}(x_{t+1})\|^2 + \gamma [\lambda(t)]^T(\mathbf{g}_{t+1}(x_{t+1}) - \mathbf{g}_t(x_{t+1})) \\
& \leq \nabla f_t(x_t)^T(\hat{x} - x_{t+1}) - \gamma \epsilon \|\lambda(t)\| + \gamma^2 \epsilon \|\mathbf{g}_t(x_t)\| + \gamma^2 \|\mathbf{g}_{t+1}(x_{t+1})\|^2 - \gamma^2 \mathbf{g}_t(x_t)^T \mathbf{g}_t(x_{t+1}) \\
& + \alpha \|\hat{x} - x_t\|^2 - \alpha \|x_{t+1} - \hat{x}\|^2 - \alpha \|x_{t+1} - x_t\|^2 + \gamma [\lambda(t)]^T(\mathbf{g}_{t+1}(x_{t+1}) - \mathbf{g}_t(x_{t+1})) \\
& \leq \nabla f_t(x_t)^T(\hat{x} - x_{t+1}) - \gamma \epsilon \|\lambda(t)\| + \gamma^2 \epsilon \|\mathbf{g}_t(x_t)\| + \gamma^2 \|\mathbf{g}_{t+1}(x_{t+1})\|^2 - \gamma^2 \mathbf{g}_t(x_t)^T \mathbf{g}_t(x_{t+1}) \\
& + \alpha \|\hat{x} - x_t\|^2 + \gamma [\lambda(t)]^T(\mathbf{g}_{t+1}(x_{t+1}) - \mathbf{g}_t(x_{t+1})) \\
& \stackrel{(a)}{\leq} \nabla f_t(x_t)^T(\hat{x} - x_{t+1}) - \gamma \epsilon \|\lambda(t)\| + \gamma^2 \epsilon F + \gamma^2 F^2 - \gamma^2 \mathbf{g}_t(x_t)^T \mathbf{g}_t(x_{t+1}) \\
& + \alpha \|\hat{x} - x_t\|^2 + \gamma [\lambda(t)]^T(\mathbf{g}_{t+1}(x_{t+1}) - \mathbf{g}_t(x_{t+1})) \\
& \stackrel{(b)}{\leq} \|\nabla f_t(x_t)\| \|\hat{x} - x_{t+1}\| - \gamma \epsilon \|\lambda(t)\| + \gamma^2 \epsilon F + \gamma^2 F^2 + \gamma^2 \|\mathbf{g}_{t-1}(x_t)\| \|\mathbf{g}_t(x_{t+1})\| \\
& + \alpha \|\hat{x} - x_t\|^2 + \gamma \|\lambda(t)\| \|\mathbf{g}_{t+1}(x_{t+1}) - \mathbf{g}_t(x_{t+1})\| \\
& \stackrel{(c)}{\leq} -\gamma \epsilon \|\lambda(t)\| + GR + \gamma^2 \epsilon F + \gamma^2 F^2 + \gamma^2 F^2 + \alpha R^2 + \gamma \|\lambda(t)\| \|\mathbf{g}_{t+1}(x_{t+1}) - \mathbf{g}_t(x_{t+1})\| \\
& = -\gamma \epsilon (\epsilon - \|\mathbf{g}_{t+1}(x_{t+1}) - \mathbf{g}_t(x_{t+1})\|) \|\lambda(t)\| + GR + \gamma^2 \epsilon F + 2\gamma^2 F^2 + \alpha R^2 \\
& \leq -\gamma \epsilon (\epsilon - \max_t \max_{x \in \mathcal{X}} \{\|\mathbf{g}_{t+1}(x) - \mathbf{g}_t(x)\|\}) \|\lambda(t)\| + GR + \gamma^2 \epsilon F + 2\gamma^2 F^2 + \alpha R^2,
\end{aligned} \tag{B.12}$$

where (a) holds by Assumption 1; (b) is due to the Cauchy-Schwarz inequality; (c) follows from Assumption 1; the last inequality holds since  $\epsilon > \hat{V}_g = \max_t \max_{x \in \mathcal{X}} \{\|\mathbf{g}_{t+1}(x) - \mathbf{g}_t(x)\|\}$  (Assumption 3). Here we define  $\hat{\epsilon} = \epsilon - \max_t \max_{x \in \mathcal{X}} \{\|\mathbf{g}_{t+1}(x) - \mathbf{g}_t(x)\|\}$ . Next we perform a reduction to absurdity process to prove this Lemma. Recall that  $\lambda_k(1) = 0$  for all  $k \in \{1, 2, \dots, K\}$  implies that  $\|\lambda(1)\| = 0 \leq \gamma F + \frac{GR + 2\gamma^2 F^2 + \alpha R^2 + \gamma^2 \epsilon F}{\gamma \hat{\epsilon}}$ . Assume that there exists a  $\tau \geq 2$  such that  $\|\lambda(\tau)\| > \gamma F + \frac{GR + 2\gamma^2 F^2 + \alpha R^2 + \gamma^2 \epsilon F}{\gamma \hat{\epsilon}}$  and  $\|\lambda(t)\| \leq \gamma F + \frac{GR + 2\gamma^2 F^2 + \alpha R^2 + \gamma^2 \epsilon F}{\gamma \hat{\epsilon}}$  for all  $t < \tau$ , then we consider two cases about the value of  $\|\lambda(\tau - 1)\|$ .

**Case 1:** If  $\|\lambda(\tau - 1)\| > \frac{GR + 2\gamma^2 F^2 + \alpha R^2 + \gamma^2 \epsilon F}{\gamma \hat{\epsilon}}$ , then we can derive  $\Delta(\tau - 1) < 0$ . According to (B.12), we have

$$\begin{aligned}
\|\lambda(\tau)\| & < \|\lambda(\tau - 1)\| \leq \frac{GR + 2\gamma^2 F^2 + \alpha R^2 + \gamma^2 \epsilon F}{\gamma \hat{\epsilon}} \\
& < \gamma F + \frac{GR + 2\gamma^2 F^2 + \alpha R^2 + \gamma^2 \epsilon F}{\gamma \hat{\epsilon}},
\end{aligned} \tag{B.14}$$

which contradicts the definition of  $\tau$ .

**Case 2:** If  $\|\lambda(\tau - 1)\| \leq \frac{GR + 2\gamma^2 F^2 + \alpha R^2 + \gamma^2 \epsilon F}{\gamma \hat{\epsilon}}$ , then according to Lemma 7 we have

$$\|\lambda(\tau)\| \leq \|\lambda(\tau - 1)\| + \gamma \|\mathbf{g}_{t-1}(x_t)\| \leq \gamma F + \frac{GR + 2\gamma^2 F^2 + \alpha R^2 + \gamma^2 \epsilon F}{\gamma \hat{\epsilon}}, \tag{B.15}$$

which also contradicts the definition of  $\tau$ . Hence  $\|\lambda(t)\| \leq \gamma F + \frac{GR + 2\gamma^2 F^2 + \alpha R^2 + \gamma^2 \epsilon F}{\gamma \hat{\epsilon}}$  holds for all  $t > 1$ . It completes the Proof.  $\square$



## References

- [1] H. Yu, M. Neely, X. Wei, Online convex optimization with stochastic constraints, in: *Advances in Neural Information Processing Systems*, 2017, pp. 1428–1438.
- [2] E. Hazan, Introduction to online convex optimization, 2019, arXiv preprint arXiv:1909.05207.
- [3] T. Chen, Q. Ling, G.B. Giannakis, An online convex optimization approach to proactive network resource allocation, *IEEE Trans. Signal Process.* 65 (24) (2017) 6350–6364.
- [4] H. Yu, M.J. Neely, A low complexity algorithm with  $O(\sqrt{T})$  regret and  $O(1)$  constraint violations for online convex optimization with long term constraints, *J. Mach. Learn. Res.* 21 (1) (2020) 1–24.
- [5] S. Qiu, X. Wei, Beyond  $O(\sqrt{T})$  regret for constrained online optimization: Gradual variations and mirror prox, 2020, arXiv preprint arXiv:2006.12455.
- [6] Y. Zhang, M.H. Hajiesmaili, S. Cai, M. Chen, Q. Zhu, Peak-aware online economic dispatching for microgrids, *IEEE Trans. Smart Grid* 9 (1) (2016) 323–335.
- [7] Z. Liu, I. Liu, S. Low, A. Wierman, Pricing data center demand response, *ACM SIGMETRICS Perform. Eval. Rev.* 42 (1) (2014) 111–123.
- [8] X. Cao, K.R. Liu, Online convex optimization with time-varying constraints and bandit feedback, *IEEE Trans. Automat. Control* 64 (7) (2018) 2665–2680.
- [9] T. Chen, G.B. Giannakis, Bandit convex optimization for scalable and dynamic IoT management, *IEEE Internet of Things J.* 6 (1) (2019).
- [10] X. Yi, X. Li, T. Yang, L. Xie, T. Chai, K.H. Johansson, Distributed bandit online convex optimization with time-varying coupled inequality constraints, *IEEE Trans. Automat. Control* (2020).
- [11] T. Chen, Q. Ling, Y. Shen, G.B. Giannakis, Heterogeneous online learning for “Thing-Adaptive” fog computing in IoT, *IEEE Internet Things J.* 5 (6) (2018) 4328–4341.
- [12] M. Mahdavi, R. Jin, T. Yang, Trading regret for efficiency: online convex optimization with long term constraints, *J. Mach. Learn. Res.* 13 (1) (2012) 2503–2528.
- [13] R. Jenatton, J. Huang, C. Archambeau, Adaptive algorithms for online convex optimization with long-term constraints, in: *International Conference on Machine Learning*, 2016, pp. 402–411.
- [14] J. Yuan, A. Lamperski, Online convex optimization for cumulative constraints, in: *Advances in Neural Information Processing Systems*, 2018, pp. 6137–6146.
- [15] P. Zhao, Y.-J. Zhang, L. Zhang, Z.-H. Zhou, Dynamic regret of convex and smooth functions, *Adv. Neural Inf. Process. Syst.* 33 (2020).
- [16] L. Zhang, Online Learning in Changing Environments.
- [17] L. Zhang, S. Lu, Z.-H. Zhou, Adaptive online learning in dynamic environments, in: *Advances in Neural Information Processing Systems*, 2018, pp. 1323–1333.
- [18] P. Sharma, P. Khanduri, L. Shen, D.J. Bucci Jr., P.K. Varshney, On distributed online convex optimization with sublinear dynamic regret and fit, 2020, arXiv preprint arXiv:2001.03166.
- [19] L. Huang, M.J. Neely, Utility optimal scheduling in processing networks, *Perform. Eval.* 68 (11) (2011) 1002–1021.
- [20] L. Huang, S. Moeller, M.J. Neely, B. Krishnamachari, LIFO-backpressure achieves near-optimal utility-delay tradeoff, *IEEE/ACM Trans. Netw.* 21 (3) (2012) 831–844.
- [21] L. Huang, X. Liu, X. Hao, The power of online learning in stochastic network optimization, in: *The 2014 ACM International Conference on Measurement and Modeling of Computer Systems*, 2014, pp. 153–165.
- [22] H. Xu, Y. Liu, W.C. Lau, J. Guo, A. Liu, Efficient online resource allocation in heterogeneous clusters with machine variability, in: *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, IEEE, 2019, pp. 478–486.
- [23] M. Amiri, Towards Enhancing QoE for Software Defined Networks Based Cloud Gaming Services (Ph.D. thesis), Université d’Ottawa/University of Ottawa, 2019.
- [24] Y. Zhao, S. Qiu, J. Liu, Proximal online gradient is optimum for dynamic regret, 2018, arXiv preprint arXiv:1810.03594.
- [25] P. Arce, L. Salinas, Online ridge regression method using sliding windows, in: *2012 31st International Conference of the Chilean Computer Science Society*, IEEE, 2012, pp. 87–90.
- [26] J. Haworth, J. Shawe-Taylor, T. Cheng, J. Wang, Local online kernel ridge regression for forecasting of urban travel times, *Transp. Res. C* 46 (2014) 151–178.
- [27] T. Chen, A.G. Marques, G.B. Giannakis, DGLB: Distributed stochastic geographical load balancing over cloud networks, *IEEE Trans. Parallel Distrib. Syst.* 28 (7) (2016) 1866–1880.
- [28] S. Im, M. Naghshnejad, M. Singhal, Scheduling jobs with non-uniform demands on multiple servers without interruption, in: *IEEE INFOCOM 2016-the 35th Annual IEEE International Conference on Computer Communications*, IEEE, 2016, pp. 1–9.
- [29] S. Im, J. Kulkarni, B. Moseley, K. Munagala, A competitive flow time algorithm for heterogeneous clusters under polytope constraints, in: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2016)*, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2016.
- [30] Y. Liu, H. Xu, W.C. Lau, Online job scheduling with resource packing on a cluster of heterogeneous servers, in: *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, IEEE, 2019, pp. 1441–1449.
- [31] S. Im, J. Kulkarni, B. Moseley, Temporal fairness of round robin: Competitive analysis for  $l_k$ -norms of flow time, in: *Proceedings of the 27th ACM Symposium on Parallelism in Algorithms and Architectures*, 2015, pp. 155–160.
- [32] H. Fang, N.J. Harvey, V.S. Portella, M.P. Friedlander, Online mirror descent and dual averaging: keeping pace in the dynamic case, 2020, arXiv preprint arXiv:2006.02585.



**Qingsing Liu** received the B.Eng. degree in electronic engineering from Tsinghua University, China, in 2019. Now he is currently pursuing the Ph.D. degree with the Institute for Interdisciplinary Information Sciences (IIIS) of Tsinghua University. His research interests include online learning, and networked and computer systems modeling and optimization. He has published several papers in ICCCN, Infocom, Performance.



**Wenfei Wu** is an assistant professor at Peking University. Dr. Wu got his Ph.D. from the University of Wisconsin-Madison in 2015, and his research interest is in networked systems. Dr. Wu has published 36 papers, and some of them are in top conferences, including NSDI, INFOCOM, SIGKDD, etc. His Ph.D. work virtual network diagnosis was awarded the best student paper in SoCC'13. His work on 5G transport layer design was also awarded the best paper runner-up in IPCCC'19. And his work on programmable switch accelerated machine learning systems got the best paper award in NSDI'21.



**Longbo Huang** is an associate professor at the Institute for Interdisciplinary Information Sciences (IIIS) at Tsinghua University, Beijing, China. He has been a visiting scholar at the LIDS lab at MIT and at the EECS department at UC Berkeley, and a visiting professor at the Chinese University of Hong Kong, Bell-labs France, and Microsoft Research Asia (MSRA). He was a visiting scientist at the Simons Institute for the Theory of Computing at UC Berkeley in Fall 2016. He received the outstanding teaching award from Tsinghua university in 2014 and the ACM SIGMETRICS Rising Star Research Award in 2018. His research interests are in the areas of stochastic modeling and analysis, reinforcement learning and control, optimization and machine learning, and big data analytics.



**Zhixuan Fang** is a tenure-track assistant professor at the Institute for Interdisciplinary Information Sciences (IIIS) at Tsinghua University, Beijing, China. He mainly focuses on the design and analysis of multi-agent systems and networked systems. He received his Ph.D. degree in computer science from Tsinghua University, China, in 2018, and his B.S. degree in physics from Peking University, China, in 2013.