Journal of Machine Learning Research 21 (2020) 1-24 Submitted 10/16; Revised 8/18; Published 2/20

A Low Complexity Algorithm with $O(\sqrt{T})$ Regret and $O(1)$ Constraint Violations for Online Convex Optimization with Long Term Constraints

Hao Yu EEYUHAO@GMAIL.COM
Department of Electrical Engineering
University of Southern California
Los Angeles, CA, 90089-2565, USA

Michael J. Neely MJNEELY@USC.EDU
Department of Electrical Engineering
University of Southern California
Los Angeles, CA, 90089-2565, USA

Editor: Csaba Szepesvari

Abstract

This paper considers online convex optimization over a complicated constraint set, which typically consists of multiple functional constraints and a set constraint. The conventional online projection algorithm (Zinkevich, 2003) can be difficult to implement due to the potentially high computation complexity of the projection operation. In this paper, we relax the functional constraints by allowing them to be violated at each round but still requiring them to be satisfied in the long term. This type of relaxed online convex optimization (with long term constraints) was first considered in Mahdavi et al. (2012). That prior work proposes an algorithm to achieve $O(\sqrt{T})$ regret and $O(T^{3/4})$ constraint violations for general problems and another algorithm to achieve an $O(T^{2/3})$ bound for both regret and constraint violations when the constraint set can be described by a finite number of linear constraints. A recent extension in Jenatton et al. (2016) can achieve $O(T^{\max\{\theta,1-\theta\}})$ regret and $O(T^{1-\theta/2})$ constraint violations where $\theta \in (0,1)$. The current paper proposes a new simple algorithm that yields improved performance in comparison to prior works. The new algorithm achieves an $O(\sqrt{T})$ regret bound with $O(1)$ constraint violations.

Keywords: online convex optimization, long term constraints, regret bounds, constraint violation bounds, low complexity

1. Introduction

Online optimization and learning is a multi-round process of making decisions in the presence of uncertainty, where a decision strategy should generally adapt decisions based on results of previous rounds (Cesa-Bianchi and Lugosi, 2006). Online convex optimization is an important subclass of these problems where the received loss function is convex with respect to the decision. At each round of online convex optimization, the decision maker is required to choose $\mathbf{x}(t)$ from a known convex set $\mathcal{X}$. After that, the convex loss function $f^t(\mathbf{x}(t))$ is disclosed to the decision maker. Note that the loss function can change arbitrarily every round t, with no probabilistic model imposed on the changes.

The goal of an online convex optimization algorithm is to select a good sequence $\mathbf{x}(t)$ such that the accumulated loss $\sum_{t=1}^{T} f^t(\mathbf{x}(t))$ is competitive with the loss of any fixed $\mathbf{x} \in \mathcal{X}$. To capture this, the T-round regret with respect to the best fixed decision is defined as follows:

$$\text{Regret}_T = \sum_{t=1}^{T} f^t(\mathbf{x}(t)) - \min_{\mathbf{x}\in\mathcal{X}} \sum_{t=1}^{T} f^t(\mathbf{x}). \tag{1}$$

The best fixed decision in hindsight $\mathbf{x}^* = argmin_{\mathbf{x}\in\mathcal{X}} \sum_{t=1}^{T} f^t(\mathbf{x})$ typically cannot be implemented. That is because it would need to be determined before the start of the first round, and this would require knowledge of the future $f^t(\cdot)$ functions for all $t \in \{1, 2, \ldots, T\}$. However, to avoid being embarrassed by the situation where our performance is significantly exceeded by a stubborn decision maker guessing $\mathbf{x}^*$ correctly by luck, a desired learning algorithm should have a small regret. Specifically, we desire a learning algorithm for which Regret_T grows sub-linearly with respect to T, i.e., the difference of average loss tends to zero as T goes to infinity when comparing the dynamic learning algorithm and a lucky stubborn decision maker.

For online convex optimization with loss functions that are convex and have bounded gradients[1], the best known regret is $O(\sqrt{T})$ and is attained by a simple online gradient descent algorithm (Zinkevich, 2003). At the end of each round t, Zinkevich's algorithm updates the decision for the next round $t+1$ by

$$\mathbf{x}(t+1) = \mathcal{P}_{\mathcal{X}}\left[\mathbf{x}(t) - \gamma \nabla f^t(\mathbf{x}(t))\right], \tag{2}$$

where $\mathcal{P}_{\mathcal{X}}[\cdot]$ represents the projection onto convex set $\mathcal{X}$ and γ is the step size.

Hazan et al. (2007) shows that better regret is possible under a more restrictive strong convexity assumption. However, Hazan et al. (2007) also shows that $\Omega(\sqrt{T})$ regret is unavoidable with no additional assumption.

In the case when $\mathcal{X}$ is a simple set, e.g., a box constraint, the projection $\mathcal{P}_{\mathcal{X}}[\cdot]$ is simple to compute and often has a closed form solution. However, if set $\mathcal{X}$ is complicated, e.g., set $\mathcal{X}$ is described via a number of functional constraints as $\mathcal{X} = \{\mathbf{x} \in \mathcal{X}_0 : g_k(\mathbf{x}) \leq 0, i \in \{1, 2, \ldots, m\}\}$, then equation (2) requires to solve the following convex program:

$$\begin{aligned} \text{minimize:} \quad & \|\mathbf{x} - [\mathbf{x}(t) - \gamma \nabla f^t(\mathbf{x}(t))]\|^2 && (3)\\ \text{such that:} \quad & g_k(\mathbf{x}) \leq 0, \forall k \in \{1, 2, \ldots, m\} && (4)\\ & \mathbf{x} \in \mathcal{X}_0 \in \mathbb{R}^n && (5) \end{aligned}$$

which can yield heavy computation and/or storage burden at each round. For instance, the interior point method (or other Newton-type methods) is an iterative algorithm and takes a number of iterations to approach the solution to the above convex program. The

1. In fact, Zinkevich's algorithm in (Zinkevich, 2003) can be extended to treat non-differentiable convex loss functions by replacing the gradient with the subgradient. The same $O(\sqrt{T})$ regret can be obtained as long as the convex loss functions have bounded subgradients. This paper also has the bounded gradient assumption in Assumption 1. This is solely for simplicity of the presentation. In fact, none of the results in this paper require the differentiability of loss functions. If any loss function is non-differentiable, we could replace the gradient with the subgradient and obtain the same regret and constraint violation bounds by replacing the bounded gradient assumption with the bounded subgradient assumption.

computation and memory space complexity at each iteration is between $O(n^2)$ and $O(n^3)$, where n is the dimension of $\mathbf{x}$.

As an attempt to reduce the projection complexity, Hazan and Kale (2012) use the Frank-Wolfe technique to replace the quadratic convex program (3)-(5) with simpler linear optimization with the same constraints (4)-(5). To completely circumvent the computational challenge due to constraint (4) in the projection operator, a variation of the standard online convex optimization, also known as online convex optimization with long term constraints, is first considered by Mahdavi et al. (2012). In this variation, complicated functional constraints $g_k(\mathbf{x}) \leq 0$ are relaxed to be soft long term constraints. That is, we do not require $\mathbf{x}(t) \in \mathcal{X}_0$ to satisfy $g_k(\mathbf{x}(t)) \leq 0$ at each round, but only require that $\sum_{t=1}^{T} g_k(\mathbf{x}(t))$, called constraint violations, grows sub-linearly. Mahdavi et al. (2012) proposes two algorithms such that one achieves $O(\sqrt{T})$ regret and $O(T^{3/4})$ constraint violations; and the other achieves $O(T^{2/3})$ for both regret and constraint violations when the set $\mathcal{X}$ can be represented by linear constraints. Further, Mahdavi et al. (2012) posed an open question of whether there exists a low complexity algorithm with an $O(\sqrt{T})$ bound on the regret and a better bound than $O(T^{3/4})$ on the constraint violations. Jenatton et al. (2016) recently extends the algorithm of Mahdavi et al. (2012) to achieve $O(T^{\max\{\theta,1-\theta\}})$ regret and $O(T^{1-\theta/2})$ constraint violations where $\theta \in (0,1)$ is a user-defined tradeoff parameter. By choosing $\theta = 1/2$ or $\theta = 2/3$, the $[O(\sqrt{T}), O(T^{3/4})]$ or $[O(T^{2/3}), O(T^{2/3})]$ regret and constraint violations of Mahdavi et al. (2012) are recovered. It is easy to observe that the best regret or constraint violations in Jenatton et al. (2016) are $O(\sqrt{T})$ under different β values. However, the algorithm of Jenatton et al. (2016) can not achieve $O(\sqrt{T})$ regret and $O(\sqrt{T})$ constraint violations simultaneously.

The current paper proposes a new algorithm that can achieve $O(\sqrt{T})$ regret and $O(1)$ constraint violations that do not grow with T; and hence yields improved performance in comparison to prior works (Mahdavi et al., 2012; Jenatton et al., 2016). The algorithm is the first to reduce the complexity associated with multiple constraints while maintaining $O(\sqrt{T})$ regret and achieving a constraint violation bound strictly better than $O(T^{3/4})$. Hence, we give a positive answer to the open question posed by Mahdavi et al. (2012). The new algorithm is related to a recent technique we developed for deterministic convex programming with a fixed objective function (Yu and Neely, 2017) and the drift-plus-penalty technique for stochastic optimization in dynamic queue networks (Neely, 2010). Our other paper (Yu et al., 2017) developed another algorithm with $O(\sqrt{T})$ regret and weaker $O(\sqrt{T})$ constraint violations for online convex optimization with stochastic constraints that include long term constraints as special cases.

Many engineering problems can be directly formulated as online convex optimization with long term constraints. For example, problems with energy or monetary constraints often define these in terms of long term time averages rather than instantaneous constraints. In general, we assume that instantaneous constraints are incorporated into the set $\mathcal{X}_0$; and long term constraints are represented via functional constraints $g_k(\mathbf{x})$. Two example problems are given as follows. More examples can be found in Mahdavi et al. (2012) and Jenatton et al. (2016).

- In the application of online display advertising (Goldfarb and Tucker, 2011; Ghosh et al., 2009), the publisher needs to iteratively allocate "impressions" to advertisers to optimize some online concave utilities for each advertiser. The utility is typically

unknown when the decision is made but can be inferred later by observing user click behaviors under the given allocations. Since each advertiser usually specifies a certain budget for a period, the "impressions" should be allocated to maximize advertisers' long term utilities subject to long term budget constraints.

- In the application of network routing in a neutral or adversarial environment, the decision maker needs to iteratively make routing decisions to maximize network utilities. Furthermore, link quality can vary after each routing decision is made. The routing decisions should satisfy the long term flow conservation constraint at each intermediate node so that queues do not overflow.

2. Online Convex Optimization with Long Term Constraints

This section introduces the problem of online convex optimization with long term constraints and presents our new algorithm.

2.1. Online Convex Optimization with Long Term Constraints

Let $\mathcal{X}_0$ be a closed convex set and $g_k(\mathbf{x}), k \in \{1, 2, \ldots, m\}$ be continuous convex functions. Denote the stacked vector of multiple functions $g_1(\mathbf{x}), \ldots, g_m(\mathbf{x})$ as $\mathbf{g}(\mathbf{x}) = [g_1(\mathbf{x}), \ldots, g_m(\mathbf{x})]^{\mathsf{T}}$. Define $\mathcal{X} = \{\mathbf{x} \in \mathcal{X}_0 : g_k(\mathbf{x}) \leq 0, k \in \{1, 2, \ldots, m\}\}$. Let $f^t(\mathbf{x})$ be a sequence of continuous convex loss functions which are determined by nature (or by an adversary) such that $f^t(\mathbf{x})$ is unknown to the decision maker until the end of round t. For any sequence $\mathbf{x}(t)$ yielded by an online algorithm, define $\sum_{t=1}^{T} f^t(\mathbf{x}(t)) - \min_{\mathbf{x}\in\mathcal{X}} \sum_{t=1}^{T} f^t(\mathbf{x})$ as the regret and $\sum_{t=1}^{T} g_k(\mathbf{x}(t)), k \in \{1, 2, \ldots, m\}$ as the constraint violations. The goal of online convex optimization with long term constraints is to choose $\mathbf{x}(t) \in \mathcal{X}_0$ for each round t such that both the regret and the constraint violations grow sub-linearly with respect to T. Throughout this paper, we use $\|\cdot\|$ to denote the Euclidean norm.

Assumption 1

- *The loss functions have bounded gradients on $\mathcal{X}_0$. That is, there exists $D > 0$ such that $\|\nabla f^t(\mathbf{x})\| \leq D$ for all $\mathbf{x} \in \mathcal{X}_0$ and all t.*

- *There exists a constant β such that $\|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{y})\| \leq \beta\|\mathbf{x} - \mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{X}_0$, i.e., $\mathbf{g}(\mathbf{x})$ is Lipschitz continuous with modulus β.*

Assumption 2 *There exists a constant G such that $\|\mathbf{g}(\mathbf{x})\| \leq G$ for all $\mathbf{x} \in \mathcal{X}_0$.*

Assumption 3 *There exists a constant R such that $\|\mathbf{x} - \mathbf{y}\| \leq R$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{X}_0$.*

Note that if $\mathcal{X}_0$ is bounded, then the existence of G follows directly from the compactness of set $\mathcal{X}_0$ and the continuity of $\mathbf{g}(\mathbf{x})$ and the existence of R follows directly from the boundedness of set $\mathcal{X}_0$.

Assumption 4 *There exists $\epsilon > 0$ and $\hat{\mathbf{x}} \in \mathcal{X}_0$ such that $g_k(\hat{\mathbf{x}}) \leq -\epsilon$ for all $k \in \{1, 2, \ldots, m\}$.*

Assumption 4, known as the Slater condition or the interior point condition, is a mild assumption in convex optimization (Boyd and Vandenberghe, 2004).

In Sections 2-3, we shall propose a new algorithm to achieve $O(\sqrt{T})$ regret and $O(1)$ constraint violations for online convex optimization with long term constraints under Assumptions 1-4 by assuming time horizon T is known in advance. In Section 4, further extensions on dealing with unknown time horizon T and relaxing Assumptions 3-4 will be discussed.

2.2. New Algorithm

Define $\tilde{\mathbf{g}}(\mathbf{x}) = \gamma \mathbf{g}(\mathbf{x})$ where $\gamma > 0$ is an algorithm parameter. Note that each $\tilde{g}_k(\mathbf{x})$ is still a convex function and $\tilde{\mathbf{g}}(\mathbf{x}) \leq \mathbf{0}$ if and only if $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$. The next lemma follows directly.

Lemma 1 *If online convex optimization with long term constraints satisfies Assumptions 1-4, then*

- $\|\tilde{\mathbf{g}}(\mathbf{x}) - \tilde{\mathbf{g}}(\mathbf{y})\| \leq \gamma\beta\|\mathbf{x} - \mathbf{y}\|$ *for all* $\mathbf{x}, \mathbf{y} \in \mathcal{X}_0$.
- $\|\tilde{\mathbf{g}}(\mathbf{x})\| \leq \gamma G$ *for all* $\mathbf{x} \in \mathcal{X}_0$.
- $\tilde{g}_k(\hat{\mathbf{x}}) \leq -\gamma\epsilon$ *for all* $k \in \{1, 2, \ldots, m\}$.

Now consider the following algorithm described in Algorithm 1. This algorithm chooses $\mathbf{x}(t+1)$ as the decision for round $t+1$ based on $f^t(\cdot)$ without knowing the cost function $f^{t+1}(\cdot)$. In this paper, we show that if the parameters γ and α are chosen to satisfy $\gamma = T^{1/4}$ and $\alpha = \frac{1}{2}[\beta^2 + 1]\sqrt{T}$, then Algorithm 1 achieves an $O(\sqrt{T})$ regret bound with $O(1)$ constraint violations.

Algorithm 1

Let $\gamma, \alpha > 0$ be constant parameters. Initialize $Q_k(0) = 0, \forall k \in \{1, 2, \ldots, m\}$. Choose arbitrary $\mathbf{x}(1) \in \mathcal{X}_0$. At the end of each round $t \in \{1, 2, 3, \ldots\}$, observe $f^t(\cdot)$ and do the following:

- Update virtual queue vector $\mathbf{Q}(t)$ via

$$Q_k(t) = \max\left\{-\tilde{g}_k(\mathbf{x}(t)), Q_k(t-1) + \tilde{g}_k(\mathbf{x}(t))\right\}, \forall k \in \{1, 2, \ldots, m\}.$$

- Choose $\mathbf{x}(t+1)$ that solves

$$\min_{\mathbf{x} \in \mathcal{X}_0} \left\{ [\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\mathbf{x} - \mathbf{x}(t)] + [\mathbf{Q}(t) + \tilde{\mathbf{g}}(\mathbf{x}(t))]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}) + \alpha\|\mathbf{x} - \mathbf{x}(t)\|^2 \right\}$$

  as the decision for the next round $t+1$, where $\nabla f^t(\mathbf{x}(t))$ is the gradient of $f^t(\mathbf{x})$ at point $\mathbf{x} = \mathbf{x}(t)$.

This algorithm introduces a virtual queue vector for constraint functions. The update equation of this virtual queue vector is similar to an algorithm recently developed by us for deterministic convex programs (with a fixed and known objective function) in Yu and

Neely (2017). However, the update for $\mathbf{x}(t+1)$ is different from Yu and Neely (2017). The role of $\mathbf{Q}(t)$ is similar to a Lagrange multiplier vector and its value is like a "queue backlog" of constraint violations. By introducing the virtual queue vector, we can characterize the regret and constraint violations for the new algorithm through the analysis of a drift-plus-penalty expression. The analysis of a drift-plus-penalty expression was originally considered in stochastic optimization for dynamic queueing systems where the decision at each round is made by observing the instantaneous cost function that changes in an i.i.d. manner (Neely, 2003, 2010). The algorithm developed in this paper is different from the conventional drift-plus-penalty algorithm in both the decision update and the virtual queue update. However, it turns out that the analysis of the drift-plus-penalty expression is also the key step to analyze the regret and the constraint violation bounds for online convex optimization with long term constraints.

Because of the $\|\mathbf{x}-\mathbf{x}(t)\|^2$ term, the choice of $\mathbf{x}(t+1)$ in Algorithm 1 involves minimization of a *strongly convex function* (strong convexity is formally defined in the next section). If the constraint functions $\mathbf{g}(\mathbf{x})$ are separable (or equivalently, $\tilde{\mathbf{g}}(\mathbf{x})$ are separable) with respect to components or blocks of $\mathbf{x}$, e.g., $\mathbf{g}(\mathbf{x})=\sum_{i=1}^{n}\mathbf{g}^{(i)}(x_i)$ or $\mathbf{g}(\mathbf{x})=\mathbf{A}\mathbf{x}-\mathbf{b}$, then the primal updates for $\mathbf{x}(t+1)$ can be decomposed into several smaller independent subproblems, each of which only involves a component or block of $\mathbf{x}(t+1)$. The next lemma further shows that the update of $\mathbf{x}(t+1)$ follows a simple gradient update in the case when $\mathbf{g}(\mathbf{x})$ is linear.

Lemma 2 *If $\mathbf{g}(\mathbf{x})$ is affine, i.e, $\mathbf{g}(\mathbf{x})=\mathbf{A}\mathbf{x}+\mathbf{b}$ for some matrix $\mathbf{A}$ and vector $\mathbf{b}$, then the update of $\mathbf{x}(t+1)$ at each round in Algorithm 1 is given by*

$$\mathbf{x}(t+1)=\mathcal{P}_{\mathcal{X}_0}\Big[\mathbf{x}(t)-\frac{1}{2\alpha}\mathbf{d}(t)\Big],$$

where $\mathbf{d}(t)=\nabla f^t(\mathbf{x}(t))+\sum_{k=1}^{m}[Q_k(t)+\tilde{g}_k(\mathbf{x}(t))]\nabla\tilde{g}_k(\mathbf{x}(t))$ and $\mathcal{P}_{\mathcal{X}_0}[\cdot]$ denotes projection onto set $\mathcal{X}_0$.

Proof Fix $t\geq\{0,1,\ldots\}$. Note that $\mathbf{d}(t)$ is a constant vector in the update of $\mathbf{x}(t+1)$. The projection operator can be interpreted as an optimization problem as follows:

$$\begin{aligned}
&\mathbf{x}(t+1)=\mathcal{P}_{\mathcal{X}_0}\Big[\mathbf{x}(t)-\frac{1}{2\alpha}\mathbf{d}(t)\Big]\\
\overset{(a)}{\Leftrightarrow}\ &\mathbf{x}(t+1)=argmin_{\mathbf{x}\in\mathcal{X}_0}\Big[\Big\|\mathbf{x}-[\mathbf{x}(t)-\frac{1}{2\alpha}\mathbf{d}(t)]\Big\|^2\Big]\\
\Leftrightarrow\ &\mathbf{x}(t+1)=argmin_{\mathbf{x}\in\mathcal{X}_0}\Big[\|\mathbf{x}-\mathbf{x}(t)\|^2+\frac{1}{\alpha}\mathbf{d}^{\mathsf{T}}(t)[\mathbf{x}-\mathbf{x}(t)]+\frac{1}{4\alpha^2}\|\mathbf{d}(t)\|^2\Big]\\
\overset{(b)}{\Leftrightarrow}\ &\mathbf{x}(t+1)=argmin_{\mathbf{x}\in\mathcal{X}_0}\Big[\sum_{k=1}^{m}[Q_k(t+1)+\tilde{g}_k(\mathbf{x}(t))]\tilde{g}_k(\mathbf{x}(t))+\mathbf{d}^{\mathsf{T}}(t)[\mathbf{x}-\mathbf{x}(t)]\\
&\qquad\qquad+\alpha\|\mathbf{x}-\mathbf{x}(t)\|^2\Big]
\end{aligned}$$

$$\overset{(c)}{\Leftrightarrow} \quad \mathbf{x}(t+1) = argmin_{\mathbf{x}\in\mathcal{X}_0}\Big[[\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\mathbf{x}-\mathbf{x}(t)] + \sum_{k=1}^{m}[Q_k(t)+\tilde{g}_k(\mathbf{x}(t))]\tilde{g}_k(\mathbf{x}(t)) + \sum_{k=1}^{m}[Q_k(t)+\tilde{g}_k(\mathbf{x}(t))][\nabla\tilde{g}_k(\mathbf{x}(t))]^{\mathsf{T}}[\mathbf{x}-\mathbf{x}(t)] + \alpha\|\mathbf{x}-\mathbf{x}(t)\|^2\Big]$$

$$\overset{(d)}{\Leftrightarrow} \quad \mathbf{x}(t+1) = argmin_{\mathbf{x}\in\mathcal{X}_0}\left[[\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\mathbf{x}-\mathbf{x}(t)] + [\mathbf{Q}(t)+\tilde{\mathbf{g}}(\mathbf{x}(t))]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}) + \alpha\|\mathbf{x}-\mathbf{x}(t)\|^2\right],$$

where (a) follows from the definition of the projection onto a convex set; (b) follows from the fact the minimizing solution does not change when we remove constant term $\frac{1}{4\alpha^2}\|\mathbf{d}(t)\|^2$, multiply positive constant α and add constant term $\sum_{k=1}^{m}[Q_k(t)+\tilde{g}_k(\mathbf{x}(t))]\tilde{g}_k(\mathbf{x}(t))$ in the objective function; (c) follows from the definition of $\mathbf{d}(t)$; and (d) follows from the identity $[\mathbf{Q}(t)+\tilde{\mathbf{g}}(\mathbf{x}(t))]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}) = \sum_{k=1}^{m}[Q_k(t)+\tilde{g}_k(\mathbf{x}(t))]\tilde{g}_k(\mathbf{x}(t)) + \sum_{k=1}^{m}[Q_k(t)+\tilde{g}_k(\mathbf{x}(t))][\nabla\tilde{g}_k(\mathbf{x}(t))]^{\mathsf{T}}[\mathbf{x}-\mathbf{x}(t)]$ for any $\mathbf{x}\in\mathbb{R}^n$, which further follows from the linearity of $\tilde{\mathbf{g}}(\mathbf{x})$. ■

2.3. Properties of Virtual Queues in Algorithm 1

In this subsection, we summarize important properties for virtual queues introduced in Algorithm 1. The virtual queue properties in this subsection are similar to those we developed in another context in Yu and Neely (2017). However, the virtual queue of the current paper is slightly different and we give proofs for completeness. These properties come from the algorithm design and rely on none of Assumptions 1-4.

Lemma 3 *In Algorithm 1, we have*

1. *At each round $t\in\{0,1,2,\ldots\}$, $Q_k(t)\geq 0$ for all $k\in\{1,2,\ldots,m\}$.*
2. *At each round $t\in\{1,2,\ldots\}$, $Q_k(t)+\tilde{g}_k(\mathbf{x}(t))\geq 0$ for all $k\in\{1,2\ldots,m\}$.*
3. *At round $t=1$, $\|\mathbf{Q}(1)\|^2 = \|\tilde{\mathbf{g}}(\mathbf{x}(1))\|^2$. At each round $t\in\{2,3,\ldots\}$, $\|\mathbf{Q}(t)\|^2 \geq \|\tilde{\mathbf{g}}(\mathbf{x}(t))\|^2$.*
4. *At each round $t\in\{1,2,\ldots\}$, $\|\mathbf{Q}(t)\| \leq \|\mathbf{Q}(t-1)\| + \|\tilde{\mathbf{g}}(\mathbf{x}(t))\|$.*

Proof

1. Fix $k\in\{1,2,\ldots,m\}$. The proof is by induction. Note that $Q_k(0)=0$ by initialization. Assume $Q_k(t)\geq 0$ for some $t\in\{0,1,2,\ldots\}$. We now prove $Q_k(t+1)\geq 0$. If $\tilde{g}_k(\mathbf{x}(t+1))\geq 0$, the virtual queue update equation of Algorithm 1 gives:

   $$Q_k(t+1) = \max\{-\tilde{g}_k(\mathbf{x}(t+1)), Q_k(t)+\tilde{g}_k(\mathbf{x}(t+1))\} \geq Q_k(t)+\tilde{g}_k(\mathbf{x}(t+1)) \geq 0.$$

   On the other hand, if $\tilde{g}_k(\mathbf{x}(t+1))<0$, then $Q_k(t+1) = \max\{-\tilde{g}_k(\mathbf{x}(t+1)), Q_k(t)+\tilde{g}_k(\mathbf{x}(t+1))\} \geq -\tilde{g}_k(\mathbf{x}(t+1)) > 0$. Thus, in both cases we have $Q_k(t+1)\geq 0$.

2. Fix $k\in\{1,2,\ldots,m\}$. Fix $t\in\{0,1,\ldots\}$. By the virtual queue update equation, we have $Q_k(t) = \max\{-\tilde{g}_k(\mathbf{x}(t)), Q_k(t-1)+\tilde{g}_k(\mathbf{x}(t))\} \geq -\tilde{g}_k(\mathbf{x}(t))$, which implies that $Q_k(t)+\tilde{g}_k(\mathbf{x}(t))\geq 0$.

3. Since we initialize $Q_k(0) = \mathbf{0}, \forall k$, we have $Q_k(1) = \max\{-\tilde{g}_k(\mathbf{x}(1)), \tilde{g}_k(\mathbf{x}(1))\} = |\tilde{g}_k(\mathbf{x}(1))|, \forall k$. Thus, $\|\mathbf{Q}(1)\|^2 = \|\tilde{\mathbf{g}}(\mathbf{x}(1))\|^2$. Fix $t \in \{2, 3, \ldots\}$ and $k \in \{1, 2, \ldots, m\}$. If $\tilde{g}_k(\mathbf{x}(t)) \geq 0$, then

$$\begin{aligned} Q_k(t) &= \max\{-\tilde{g}_k(\mathbf{x}(t)), Q_k(t-1) + \tilde{g}_k(\mathbf{x}(t))\} \\ &\geq Q_k(t-1) + \tilde{g}_k(\mathbf{x}(t)) \\ &\overset{(a)}{\geq} \tilde{g}_k(\mathbf{x}(t)) = |\tilde{g}_k(\mathbf{x}(t))|, \end{aligned}$$

where (a) follows from part 1. On the other hand, if $\tilde{g}_k(\mathbf{x}(t)) < 0$, then $Q_k(t) = \max\{-\tilde{g}_k(\mathbf{x}(t)), Q_k(t-1) + \tilde{g}_k(\mathbf{x}(t))\} \geq -\tilde{g}_k(\mathbf{x}(t)) = |\tilde{g}_k(\mathbf{x}(t))|$. Thus, in both cases, we have $Q_k(t) \geq |\tilde{g}_k(\mathbf{x}(t))|$. Squaring both sides and summing over $k \in \{1, 2, \ldots, m\}$ yields $\|\mathbf{Q}(t)\|^2 \geq \|\tilde{\mathbf{g}}(\mathbf{x}(t))\|^2$.

4. Fix $t \in \{0, 1, \ldots\}$. Define vector $\mathbf{h} = [h_1, \ldots, h_m]^{\mathsf{T}}$ by $h_k = |\tilde{g}_k(\mathbf{x}(t))|, \forall k \in \{1, 2, \ldots, m\}$. Note that $\|\mathbf{h}\| = \|\tilde{\mathbf{g}}(\mathbf{x}(t))\|$. For any $k \in \{1, 2, \ldots, m\}$, by the virtual update equation we have

$$\begin{aligned} Q_k(t) &= \max\{-\tilde{g}_k(\mathbf{x}(t)), Q_k(t-1) + \tilde{g}_k(\mathbf{x}(t))\} \\ &\leq |Q_k(t-1)| + |\tilde{g}_k(\mathbf{x}(t))| \\ &= Q_k(t-1) + h_k. \end{aligned}$$

Squaring both sides and summing over $k \in \{1, 2, \ldots, m\}$ yields $\|\mathbf{Q}(t)\|^2 \leq \|\mathbf{Q}(t-1) + \mathbf{h}\|^2$, which is equivalent to $\|\mathbf{Q}(t)\| \leq \|\mathbf{Q}(t-1) + \mathbf{h}\|$. Finally, by the triangle inequality $\|\mathbf{Q}(t-1) + \mathbf{h}\| \leq \|\mathbf{Q}(t-1)\| + \|\mathbf{h}\|$ and recalling that $\|\mathbf{h}\| = \|\tilde{\mathbf{g}}(\mathbf{x}(t))\|$, we have $\|\mathbf{Q}(t)\| \leq \|\mathbf{Q}(t-1)\| + \|\tilde{\mathbf{g}}(\mathbf{x}(t))\|$.

■

Lemma 4 *Let $\mathbf{Q}(t), t \in \{0, 1, \ldots\}$ be the sequence generated by Algorithm 1. For any $T \geq 1$, we have*

$$\sum_{t=1}^{T} g_k(\mathbf{x}(t)) \leq \frac{1}{\gamma} Q_k(T), \forall k \in \{1, 2, \ldots, m\}.$$

Proof Fix $k \in \{1, 2, \ldots, m\}$ and $T \geq 1$. For any $t \in \{1, 2, \ldots, T\}$ the update rule of Algorithm 1 gives:

$$Q_k(t) = \max\{-\tilde{g}_k(\mathbf{x}(t)), Q_k(t-1) + \tilde{g}_k(\mathbf{x}(t))\} \geq Q_k(t-1) + \tilde{g}_k(\mathbf{x}(t)).$$

Hence, $\tilde{g}_k(\mathbf{x}(t)) \leq Q_k(t) - Q_k(t-1)$. Summing over $t \in \{0, \ldots, T-1\}$ yields

$$\sum_{t=1}^{T} \tilde{g}_k(\mathbf{x}(t)) \leq Q_k(T) - Q_k(0) \overset{(a)}{\leq} Q_k(T).$$

where (a) follows from the initialization rule $Q_k(0) = 0$. This lemma follows by recalling that $\tilde{g}_k(\mathbf{x}) = \gamma g_k(\mathbf{x})$. ■

Let $\mathbf{Q}(t) = \big[Q_1(t), \ldots, Q_m(t)\big]^{\mathsf{T}}$ be the vector of virtual queue backlogs. Define $L(t) = \frac{1}{2}\|\mathbf{Q}(t)\|^2$. The function $L(t)$ shall be called a *Lyapunov function*. Define the Lyapunov drift as

$$\Delta(t) = L(t+1) - L(t) = \frac{1}{2}[\|\mathbf{Q}(t+1)\|^2 - \|\mathbf{Q}(t)\|^2]. \tag{6}$$

The Lyapunov drift in (6) was originally used in the drift-plus-penalty technique for stochastic optimization in dynamic queueing networks (Neely, 2010) and was recently used in Yu and Neely (2017) to develop fast $O(1/T)$ converging Lagrangian methods for deterministic optimization by controlling constraint violations through a virtul queue dynamic. While online convex optimization with long term constraints differs from both scenarios, it turns out the Lyapunov drift is still quite useful to jointly analyze regret and constraint violations.

Lemma 5 *At each round $t \in \{0, 1, 2, \ldots\}$ in Algorithm 1, an upper bound of the Lyapunov drift is given by*

$$\Delta(t) \leq [\mathbf{Q}(t)]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}(t+1)) + \|\tilde{\mathbf{g}}(\mathbf{x}(t+1))\|^2. \tag{7}$$

Proof Fix $t \in \{0, 1, \ldots, \}$. The virtual queue update equations $Q_k(t+1) = \max\{-\tilde{g}_k(\mathbf{x}(t+1)), Q_k(t) + \tilde{g}_k(\mathbf{x}(t+1))\}, \forall k \in \{1, 2, \ldots, m\}$ can be rewritten as

$$Q_k(t+1) = Q_k(t) + h_k(\mathbf{x}(t+1)), \forall k \in \{1, 2, \ldots, m\}, \tag{8}$$

where

$$h_k(\mathbf{x}(t)) = \begin{cases} \tilde{g}_k(\mathbf{x}(t+1)), & \text{if } Q_k(t) + \tilde{g}_k(\mathbf{x}(t+1)) \geq -\tilde{g}_k(\mathbf{x}(t+1)) \\ -Q_k(t) - \tilde{g}_k(\mathbf{x}(t+1)), & \text{else} \end{cases} \quad \forall k.$$

Fix $k \in \{1, 2, \ldots, m\}$. Squaring both sides of (8) and dividing by 2 yields:

$$\begin{aligned}
&\frac{1}{2}[Q_k(t+1)]^2 \\
=&\frac{1}{2}[Q_k(t)]^2 + \frac{1}{2}[h_k(\mathbf{x}(t+1))]^2 + Q_k(t)h_k(\mathbf{x}(t+1)) \\
=&\frac{1}{2}[Q_k(t)]^2 + \frac{1}{2}[h_k(\mathbf{x}(t+1))]^2 + Q_k(t)\tilde{g}_k(\mathbf{x}(t+1)) + Q_k(t)[h_k(\mathbf{x}(t+1)) - \tilde{g}_k(\mathbf{x}(t+1))] \\
\overset{(a)}{=}&\frac{1}{2}[Q_k(t)]^2 + \frac{1}{2}[h_k(\mathbf{x}(t+1))]^2 + Q_k(t)\tilde{g}_k(\mathbf{x}(t+1)) \\
&\quad - [h_k(\mathbf{x}(t+1)) + \tilde{g}_k(\mathbf{x}(t+1))][h_k(\mathbf{x}(t+1)) - \tilde{g}_k(\mathbf{x}(t+1))] \\
=&\frac{1}{2}[Q_k(t)]^2 - \frac{1}{2}[h_k(\mathbf{x}(t+1))]^2 + Q_k(t)\tilde{g}_k(\mathbf{x}(t+1)) + [\tilde{g}_k(\mathbf{x}(t+1))]^2 \\
\leq&\frac{1}{2}[Q_k(t)]^2 + Q_k(t)\tilde{g}_k(\mathbf{x}(t+1)) + [\tilde{g}_k(\mathbf{x}(t+1))]^2,
\end{aligned}$$

where (a) follows from the fact that $Q_k(t)[h_k(\mathbf{x}(t+1)) - \tilde{g}_k(\mathbf{x}(t+1))] = -[h_k(\mathbf{x}(t+1)) + \tilde{g}_k(\mathbf{x}(t+1))] \cdot [h_k(\mathbf{x}(t+1)) - \tilde{g}_k(\mathbf{x}(t+1))]$, which can be shown by considering $h_k(\mathbf{x}(t+1)) = \tilde{g}_k(\mathbf{x}(t+1))$ and $h_k(\mathbf{x}(t+1)) \neq \tilde{g}_k(\mathbf{x}(t+1))$. Summing over $k \in \{1, 2, \ldots, m\}$ yields

$$\frac{1}{2}\|\mathbf{Q}(t+1)\|^2 \leq \frac{1}{2}\|\mathbf{Q}(t)\|^2 + [\mathbf{Q}(t)]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}(t+1)) + \|\tilde{\mathbf{g}}(\mathbf{x}(t+1))\|^2.$$

Rearranging the terms yields the desired result. ■

3. Regret and Constraint Violation Analysis of Algorithm 1

This section analyzes the regret and constraint violations of Algorithm 1 for online convex optimization with long term constraints under Assumptions 1-4.

3.1. An Upper Bound of the Drift-Plus-Penalty Expression

Definition 1 (Strongly Convex Functions) *Let $\mathcal{X} \subseteq \mathbb{R}^n$ be a convex set. Function h is said to be strongly convex on $\mathcal{X}$ with modulus α if there exists a constant $\alpha > 0$ such that $h(\mathbf{x}) - \frac{1}{2}\alpha\|\mathbf{x}\|^2$ is convex on $\mathcal{X}$.*

By the definition of strongly convex functions, if $h(\mathbf{x})$ is convex and $\alpha > 0$, then $h(\mathbf{x}) + \alpha\|\mathbf{x} - \mathbf{x}_0\|^2$ is strongly convex with modulus 2α for any constant $\mathbf{x}_0$. The following lemma summarizes a useful property of the minimizer for a strongly convex function.

Lemma 6 (Corollary 1 in Yu and Neely (2017)) *Let $\mathcal{X} \subseteq \mathbb{R}^n$ be a convex set. Let function h be strongly convex on $\mathcal{X}$ with modulus α and $\mathbf{x}^{opt}$ be a global minimum of h on $\mathcal{X}$. Then, $h(\mathbf{x}^{opt}) \leq h(\mathbf{x}) - \frac{\alpha}{2}\|\mathbf{x}^{opt} - \mathbf{x}\|^2$ for all $\mathbf{x} \in \mathcal{X}$.*

Lemma 7 *Consider online convex optimization with long term constraints under Assumption 1. Let $\mathbf{x}^* \in \mathcal{X}_0$ be any fixed solution that satisfies $\mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}$, e.g., $\mathbf{x}^* = argmin_{\mathbf{x}\in\mathcal{X}} \sum_{t=1}^{T} f^t(\mathbf{x})$. Let $\gamma > 0$ and $\eta > 0$ be arbitrary. If $\alpha \geq \frac{1}{2}[\gamma^2\beta^2 + \eta]$ in Algorithm 1, then for all $t \geq 1$, we have*

$$\begin{aligned}&\Delta(t) + f^t(\mathbf{x}(t))\\ \leq& f^t(\mathbf{x}^*) + \alpha[\|\mathbf{x}^* - \mathbf{x}(t)\|^2 - \|\mathbf{x}^* - \mathbf{x}(t+1)\|^2] + \frac{1}{2}[\|\tilde{\mathbf{g}}(\mathbf{x}(t+1))\|^2 - \|\tilde{\mathbf{g}}(\mathbf{x}(t))\|^2] + \frac{1}{2\eta}D^2,\end{aligned}$$

where β and D are defined in Assumption 1.

Proof Fix $t \geq 1$. Note that part 2 of Lemma 3 implies that $\mathbf{Q}(t) + \tilde{\mathbf{g}}(\mathbf{x}(t))$ is componentwise nonnegative. Hence, $[\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\mathbf{x} - \mathbf{x}(t)] + [\mathbf{Q}(t) + \tilde{\mathbf{g}}(\mathbf{x}(t))]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x})$ is a convex function with respect to $\mathbf{x}$. Since $\alpha\|\mathbf{x} - \mathbf{x}(t)\|^2$ is strongly convex with respect to $\mathbf{x}$ with modulus 2α, it follows that

$$[\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\mathbf{x} - \mathbf{x}(t)] + [\mathbf{Q}(t) + \tilde{\mathbf{g}}(\mathbf{x}(t))]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}) + \alpha\|\mathbf{x} - \mathbf{x}(t)\|^2$$

is strongly convex with respect to $\mathbf{x}$ with modulus 2α.

Since $\mathbf{x}(t+1)$ is chosen to minimize the above strongly convex function, by Lemma 6, we have

$$\begin{aligned}&[\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\mathbf{x}(t+1) - \mathbf{x}(t)] + [\mathbf{Q}(t) + \tilde{\mathbf{g}}(\mathbf{x}(t))]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}(t+1)) + \alpha\|\mathbf{x}(t+1) - \mathbf{x}(t)\|^2\\ \leq& [\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\mathbf{x}^* - \mathbf{x}(t)] + [\mathbf{Q}(t) + \tilde{\mathbf{g}}(\mathbf{x}(t))]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}^*) + \alpha\|\mathbf{x}^* - \mathbf{x}(t)\|^2 - \alpha\|\mathbf{x}^* - \mathbf{x}(t+1)\|^2.\end{aligned}$$

Adding $f^t(\mathbf{x}(t))$ on both sides yields

$$\begin{aligned}
&f^t(\mathbf{x}(t)) + [\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\mathbf{x}(t+1) - \mathbf{x}(t)] + [\mathbf{Q}(t) + \tilde{\mathbf{g}}(\mathbf{x}(t))]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}(t+1)) \\
&+ \alpha\|\mathbf{x}(t+1) - \mathbf{x}(t)\|^2 \\
\leq& f^t(\mathbf{x}(t)) + [\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\mathbf{x}^* - \mathbf{x}(t)] + [\mathbf{Q}(t) + \tilde{\mathbf{g}}(\mathbf{x}(t))]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}^*) + \alpha\|\mathbf{x}^* - \mathbf{x}(t)\|^2 \\
&- \alpha\|\mathbf{x}^* - \mathbf{x}(t+1)\|^2 \\
\overset{(a)}{\leq}& f^t(\mathbf{x}^*) + \underbrace{[\mathbf{Q}(t) + \tilde{\mathbf{g}}(\mathbf{x}(t))]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}^*)}_{\leq 0} + \alpha[\|\mathbf{x}^* - \mathbf{x}(t)\|^2 - \|\mathbf{x}^* - \mathbf{x}(t+1)\|^2] \\
\overset{(b)}{\leq}& f^t(\mathbf{x}^*) + \alpha[\|\mathbf{x}^* - \mathbf{x}(t)\|^2 - \|\mathbf{x}^* - \mathbf{x}(t+1)\|^2],
\end{aligned}$$

where (a) follows from the convexity of function $f^t(\mathbf{x})$; and (b) follows by using the fact that $\tilde{g}_k(\mathbf{x}^*) \leq 0$ and $Q_k(t) + \tilde{g}_k(\mathbf{x}(t)) \geq 0$ (i.e., part 2 in Lemma 3) for all $k \in \{1, 2, \ldots, m\}$ to eliminate the term marked by an underbrace.

Rearranging terms yields

$$\begin{aligned}
&f^t(\mathbf{x}(t)) + [\mathbf{Q}(t)]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}(t+1)) \\
\leq& f^t(\mathbf{x}^*) + \alpha[\|\mathbf{x}^* - \mathbf{x}(t)\|^2 - \|\mathbf{x}^* - \mathbf{x}(t+1)\|^2] - \alpha\|\mathbf{x}(t+1) - \mathbf{x}(t)\|^2 \\
&- [\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\mathbf{x}(t+1) - \mathbf{x}(t)] - [\tilde{\mathbf{g}}(\mathbf{x}(t))]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}(t+1)).
\end{aligned} \tag{9}$$

For any $\eta > 0$, we have

$$\begin{aligned}
-[\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\mathbf{x}(t+1) - \mathbf{x}(t)] &\overset{(a)}{\leq} \|\nabla f^t(\mathbf{x}(t))\|\|\mathbf{x}(t+1) - \mathbf{x}(t)\| \\
&= [\frac{1}{\sqrt{\eta}}\|\nabla f^t(\mathbf{x}(t))\|][\sqrt{\eta}\|\mathbf{x}(t+1) - \mathbf{x}(t)\|] \\
&\overset{(b)}{\leq} \frac{1}{2\eta}\|\nabla f^t(\mathbf{x}(t))\|^2 + \frac{1}{2}\eta\|\mathbf{x}(t+1) - \mathbf{x}(t)\|^2 \\
&\overset{(c)}{\leq} \frac{1}{2\eta}D^2 + \frac{1}{2}\eta\|\mathbf{x}(t+1) - \mathbf{x}(t)\|^2,
\end{aligned} \tag{10}$$

where (a) follows from the Cauchy-Schwarz inequality; (b) follows from the basic inequality $ab \leq \frac{1}{2}(a^2 + b^2), \forall a, b \in \mathbb{R}$; and (c) follows from Assumption 1.

Note that $\mathbf{u}_1^{\mathsf{T}}\mathbf{u}_2 = \frac{1}{2}[\|\mathbf{u}_1\|^2 + \|\mathbf{u}_2\|^2 - \|\mathbf{u}_1 - \mathbf{u}_2\|^2]$ for any $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{R}^m$. Thus, we have

$$[\tilde{\mathbf{g}}(\mathbf{x}(t))]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}(t+1)) = \frac{1}{2}\big[\|\tilde{\mathbf{g}}(\mathbf{x}(t))\|^2 + \|\tilde{\mathbf{g}}(\mathbf{x}(t+1))\|^2 - \|\tilde{\mathbf{g}}(\mathbf{x}(t+1)) - \tilde{\mathbf{g}}(\mathbf{x}(t))\|^2\big]. \tag{11}$$

Substituting (10) and (11) into (9) yields

$$\begin{aligned}
&f^t(\mathbf{x}(t)) + [\mathbf{Q}(t)]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}(t+1))\\
\leq& f^t(\mathbf{x}^*) + \alpha[\|\mathbf{x}^* - \mathbf{x}(t)\|^2 - \|\mathbf{x}^* - \mathbf{x}(t+1)\|^2] + [\frac{1}{2}\eta - \alpha]\|\mathbf{x}(t+1) - \mathbf{x}(t)\|^2 + \frac{1}{2\eta}D^2\\
&+ \frac{1}{2}\|\tilde{\mathbf{g}}(\mathbf{x}(t+1)) - \tilde{\mathbf{g}}(\mathbf{x}(t))\|^2 - \frac{1}{2}\|\tilde{\mathbf{g}}(\mathbf{x}(t))\|^2 - \frac{1}{2}\|\tilde{\mathbf{g}}(\mathbf{x}(t+1))\|^2\\
\overset{(a)}{\leq}& f^t(\mathbf{x}^*) + \alpha[\|\mathbf{x}^* - \mathbf{x}(t)\|^2 - \|\mathbf{x}^* - \mathbf{x}(t+1)\|^2] + [\frac{1}{2}\gamma^2\beta^2 + \frac{1}{2}\eta - \alpha]\|\mathbf{x}(t+1) - \mathbf{x}(t)\|^2\\
&+ \frac{1}{2\eta}D^2 - \frac{1}{2}\|\tilde{\mathbf{g}}(\mathbf{x}(t))\|^2 - \frac{1}{2}\|\tilde{\mathbf{g}}(\mathbf{x}(t+1))\|^2\\
\overset{(b)}{\leq}& f^t(\mathbf{x}^*) + \alpha[\|\mathbf{x}^* - \mathbf{x}(t)\|^2 - \|\mathbf{x}^* - \mathbf{x}(t+1)\|^2] + \frac{1}{2\eta}D^2 - \frac{1}{2}\|\tilde{\mathbf{g}}(\mathbf{x}(t))\|^2 - \frac{1}{2}\|\tilde{\mathbf{g}}(\mathbf{x}(t+1))\|^2,
\end{aligned} \tag{12}$$

where (a) follows from the fact that $\|\tilde{\mathbf{g}}(\mathbf{x}(t+1)) - \tilde{\mathbf{g}}(\mathbf{x}(t))\| \leq \gamma\beta\|\mathbf{x}(t+1) - \mathbf{x}(t)\|$, which further follows from Lemma 1; and (b) follows from the fact that $\alpha \geq \frac{1}{2}[\gamma^2\beta^2 + \eta]$.

By Lemma 5, we have

$$\Delta(t) \leq [\mathbf{Q}(t)]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}(t+1)) + \|\tilde{\mathbf{g}}(\mathbf{x}(t+1))\|^2. \tag{13}$$

Summing (12) and (13) together yields

$$\begin{aligned}
&\Delta(t) + f^t(\mathbf{x}(t))\\
\leq& f^t(\mathbf{x}^*) + \alpha[\|\mathbf{x}^* - \mathbf{x}(t)\|^2 - \|\mathbf{x}^* - \mathbf{x}(t+1)\|^2] + \frac{1}{2}[\|\tilde{\mathbf{g}}(\mathbf{x}(t+1))\|^2 - \|\tilde{\mathbf{g}}(\mathbf{x}(t))\|^2] + \frac{1}{2\eta}D^2.
\end{aligned}$$

■

3.2. Regret Analysis

Theorem 1 *Consider online convex optimization with long term constraints under Assumption 1. Let $\mathbf{x}^* \in \mathcal{X}_0$ be any fixed solution that satisfies $\mathbf{g}(\mathbf{x}^*) \leq \mathbf{0}$, e.g., $\mathbf{x}^* = argmin_{\mathbf{x}\in\mathcal{X}} \sum_{t=1}^T f^t(\mathbf{x})$.*

1. *Let $\gamma > 0$ and $\eta > 0$ be arbitrary. If $\alpha \geq \frac{1}{2}[\gamma^2\beta^2 + \eta]$ in Algorithm 1, then for all $T \geq 1$, we have*

$$\sum_{t=1}^T f^t(\mathbf{x}(t)) \leq \sum_{t=1}^T f^t(\mathbf{x}^*) + \alpha\|\mathbf{x}^* - \mathbf{x}(1)\|^2 + \frac{1}{2\eta}D^2T.$$

2. *If $\gamma = T^{1/4}, \eta = \sqrt{T}$ and $\alpha = \frac{1}{2}[\beta^2 + 1]\sqrt{T}$ in Algorithm 1, then for all $T \geq 1$, we have*

$$\sum_{t=1}^T f^t(\mathbf{x}(t)) \leq \sum_{t=1}^T f^t(\mathbf{x}^*) + O(\sqrt{T}).$$

Proof Fix $T \geq 1$. Since $\alpha \geq \frac{1}{2}[\gamma^2\beta^2 + \eta]$, by Lemma 7, for all $t \in \{1, 2, \ldots, T\}$, we have

$$\begin{aligned}
&\Delta(t) + f^t(\mathbf{x}(t)) \\
\leq& f^t(\mathbf{x}^*) + \alpha[\|\mathbf{x}^* - \mathbf{x}(t)\|^2 - \|\mathbf{x}^* - \mathbf{x}(t+1)\|^2] + \frac{1}{2}[\|\tilde{\mathbf{g}}(\mathbf{x}(t+1))\|^2 - \|\tilde{\mathbf{g}}(\mathbf{x}(t))\|^2] + \frac{1}{2\eta}D^2.
\end{aligned}$$

Summing over $t \in \{1, 2, \ldots, T\}$ yields

$$\begin{aligned}
\sum_{t=1}^{T}\Delta(t) + \sum_{t=1}^{T} f^t(\mathbf{x}(t)) \leq& \sum_{t=1}^{T} f^t(\mathbf{x}^*) + \alpha\sum_{t=1}^{T}[\|\mathbf{x}^* - \mathbf{x}(t)\|^2 - \|\mathbf{x}^* - \mathbf{x}(t+1)\|^2] \\
&+ \frac{1}{2}\sum_{t=1}^{T}[\|\tilde{\mathbf{g}}(\mathbf{x}(t+1))\|^2 - \|\tilde{\mathbf{g}}(\mathbf{x}(t))\|^2] + \frac{1}{2\eta}D^2T.
\end{aligned}$$

Recalling that $\Delta(t) = L(t+1) - L(t)$ and simplifying summations yields

$$\begin{aligned}
&L(T+1) - L(1) + \sum_{t=1}^{T} f^t(\mathbf{x}(t)) \\
\leq& \sum_{t=1}^{T} f^t(\mathbf{x}^*) + \alpha\|\mathbf{x}^* - \mathbf{x}(1)\|^2 - \alpha\|\mathbf{x}^* - \mathbf{x}(T+1)\|^2 + \frac{1}{2}\|\tilde{\mathbf{g}}(\mathbf{x}(T+1))\|^2 - \frac{1}{2}\|\tilde{\mathbf{g}}(\mathbf{x}(1))\|^2 + \frac{1}{2\eta}D^2T \\
\leq& \sum_{t=1}^{T} f^t(\mathbf{x}^*) + \alpha\|\mathbf{x}^* - \mathbf{x}(1)\|^2 + \frac{1}{2}\|\tilde{\mathbf{g}}(\mathbf{x}(T+1))\|^2 - \frac{1}{2}\|\tilde{\mathbf{g}}(\mathbf{x}(1))\|^2 + \frac{1}{2\eta}D^2T.
\end{aligned}$$

Rearranging terms yields

$$\begin{aligned}
&\sum_{t=1}^{T} f^t(\mathbf{x}(t)) \\
\leq& \sum_{t=1}^{T} f^t(\mathbf{x}^*) + \alpha\|\mathbf{x}^* - \mathbf{x}(1)\|^2 + \frac{1}{2}\|\tilde{\mathbf{g}}(\mathbf{x}(T+1))\|^2 + L(1) - L(T+1) + \frac{1}{2\eta}D^2T \\
\overset{(a)}{=}& \sum_{t=1}^{T} f^t(\mathbf{x}^*) + \alpha\|\mathbf{x}^* - \mathbf{x}(1)\|^2 + \frac{1}{2}\|\tilde{\mathbf{g}}(\mathbf{x}(T+1))\|^2 - \frac{1}{2}\|\tilde{\mathbf{g}}(\mathbf{x}(1))\|^2 + \frac{1}{2}\|\mathbf{Q}(1)\|^2 \\
&- \frac{1}{2}\|\mathbf{Q}(T+1)\|^2 + \frac{1}{2\eta}D^2T \\
\overset{(b)}{\leq}& \sum_{t=1}^{T} f^t(\mathbf{x}^*) + \alpha\|\mathbf{x}^* - \mathbf{x}(1)\|^2 + \frac{1}{2\eta}D^2T
\end{aligned}$$

where (a) follows form the definition that $L(1) = \frac{1}{2}\|\mathbf{Q}(1)\|^2$ and $L(T+1) = \frac{1}{2}\|\mathbf{Q}(T+1)\|^2$; and (b) follows from the fact that $\|\mathbf{Q}(1)\|^2 = \|\tilde{\mathbf{g}}(\mathbf{x}(1))\|^2$ and $\|\mathbf{Q}(T+1)\|^2 \geq \|\tilde{\mathbf{g}}(\mathbf{x}(T+1))\|^2$, i.e., part 3 in Lemma 3.

Thus, the first part of this theorem follows. Note that if we let $\gamma = T^{1/4}$ and $\eta = \sqrt{T}$, then $\alpha = \frac{1}{2}[\beta^2+1]\sqrt{T} \geq \frac{1}{2}[\gamma^2\beta^2+\eta]$. The second part of this theorem follows by substituting

$\gamma = T^{1/4}, \eta = \sqrt{T}$ and $\alpha = \frac{1}{2}[\beta^2+1]\sqrt{T}$ into the first part of this theorem. Thus, we have

$$\begin{aligned}\sum_{t=1}^{T} f^t(\mathbf{x}(t)) \leq & \sum_{t=1}^{T} f^t(\mathbf{x}^*) + \frac{1}{2}[\beta^2+1]\|\mathbf{x}^* - \mathbf{x}(1)\|^2\sqrt{T} + \frac{1}{2}D^2\sqrt{T} \\ = & \sum_{t=1}^{T} f^t(\mathbf{x}^*) + O(\sqrt{T}).\end{aligned}$$

■

3.3. An Upper Bound of the Virtual Queue Vector

It remains to establish a bound on constraint violations. Lemma 4 shows this can be done by bounding $\|\mathbf{Q}(t)\|$.

Lemma 8 *Consider online convex optimization with long term constraints under Assumptions 1-4. At each round $t \in \{1,2,\ldots,\}$ in Algorithm 1, if $\|\mathbf{Q}(t)\| > \gamma G + \frac{\alpha R^2 + DR + 2\gamma^2 G^2}{\gamma\epsilon}$, where D, G and R are defined in Assumption 1 and ϵ is defined in Assumption 4, then $\|\mathbf{Q}(t+1)\| < \|\mathbf{Q}(t)\|$.*

Proof Let $\hat{\mathbf{x}} \in \mathcal{X}_0$ and $\epsilon > 0$ be defined in Assumption 4. Fix $t \geq 0$. Since $\mathbf{x}(t+1)$ is chosen to minimize $[\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\mathbf{x} - \mathbf{x}(t)] + [\mathbf{Q}(t) + \tilde{\mathbf{g}}(\mathbf{x}(t))]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}) + \alpha\|\mathbf{x} - \mathbf{x}(t)\|^2$, we have

$$\begin{aligned}&[\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\mathbf{x}(t+1) - \mathbf{x}(t)] + [\mathbf{Q}(t) + \tilde{\mathbf{g}}(\mathbf{x}(t))]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}(t+1)) + \alpha\|\mathbf{x}(t+1) - \mathbf{x}(t)\|^2 \\ \leq & [\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\hat{\mathbf{x}} - \mathbf{x}(t)] + [\mathbf{Q}(t) + \tilde{\mathbf{g}}(\mathbf{x}(t))]^{\mathsf{T}}\tilde{\mathbf{g}}(\hat{\mathbf{x}}) + \alpha\|\hat{\mathbf{x}} - \mathbf{x}(t)\|^2 \\ \overset{(a)}{\leq} & [\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\hat{\mathbf{x}} - \mathbf{x}(t)] - \gamma\epsilon\sum_{k=1}^{m}[Q_k(t) + \tilde{g}_k(\mathbf{x}(t))] + \alpha\|\hat{\mathbf{x}} - \mathbf{x}(t)\|^2 \\ \overset{(b)}{\leq} & [\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\hat{\mathbf{x}} - \mathbf{x}(t)] - \gamma\epsilon\|\mathbf{Q}(t) + \tilde{\mathbf{g}}(\mathbf{x}(t))\| + \alpha\|\hat{\mathbf{x}} - \mathbf{x}(t)\|^2 \\ \overset{(c)}{\leq} & [\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\hat{\mathbf{x}} - \mathbf{x}(t)] - \gamma\epsilon\big[\|\mathbf{Q}(t)\| - \|\tilde{\mathbf{g}}(\mathbf{x}(t))\|\big] + \alpha\|\hat{\mathbf{x}} - \mathbf{x}(t)\|^2,\end{aligned}$$

where (a) follows from the fact that $\tilde{g}_k(\hat{\mathbf{x}}) \leq -\gamma\epsilon, \forall k \in \{1,2,\ldots,m\}$, i.e., Lemma 1 and the fact that $Q_k(t) + \tilde{g}_k(\mathbf{x}(t)) \geq 0, \forall k \in \{1,2,\ldots,m\}$, i.e., part 2 in Lemma 3; (b) follows from the basic inequality $\sum_{i=1}^{m} a_i \geq \sqrt{\sum_{i=1}^{m} a_i^2}$ for any nonnegative vector $\mathbf{a} \geq \mathbf{0}$; and (c) follows from the triangle inequality $\|\mathbf{x} - \mathbf{y}\| \geq \|\mathbf{x}\| - \|\mathbf{y}\|, \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^m$.

Rearranging terms yields

$$\begin{aligned}
&[\mathbf{Q}(t)]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}(t+1)) \\
\leq& -\gamma\epsilon\|\mathbf{Q}(t)\| + \gamma\epsilon\|\tilde{\mathbf{g}}(\mathbf{x}(t))\| + \alpha\|\hat{\mathbf{x}} - \mathbf{x}(t)\|^2 - \alpha\|\mathbf{x}(t+1) - \mathbf{x}(t)\|^2 \\
&+ [\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\hat{\mathbf{x}} - \mathbf{x}(t)] - [\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\mathbf{x}(t+1) - \mathbf{x}(t)] - [\tilde{\mathbf{g}}(\mathbf{x}(t))]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}(t+1)) \\
\leq& -\gamma\epsilon\|\mathbf{Q}(t)\| + \gamma\epsilon\|\tilde{\mathbf{g}}(\mathbf{x}(t))\| + \alpha\|\hat{\mathbf{x}} - \mathbf{x}(t)\|^2 + [\nabla f^t(\mathbf{x}(t))]^{\mathsf{T}}[\hat{\mathbf{x}} - \mathbf{x}(t+1)] \\
&- [\tilde{\mathbf{g}}(\mathbf{x}(t))]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}(t+1)) \\
\overset{(a)}{\leq}& -\gamma\epsilon\|\mathbf{Q}(t)\| + \gamma\epsilon\|\tilde{\mathbf{g}}(\mathbf{x}(t))\| + \alpha\|\hat{\mathbf{x}} - \mathbf{x}(t)\|^2 + \|\nabla f^t(\mathbf{x}(t))\|\|\hat{\mathbf{x}} - \mathbf{x}(t+1)\| \\
&+ \|\tilde{\mathbf{g}}(\mathbf{x}(t))\|\|\tilde{\mathbf{g}}(\mathbf{x}(t+1))\| \\
\overset{(b)}{\leq}& -\gamma\epsilon\|\mathbf{Q}(t)\| + \gamma^2\epsilon G + \alpha R^2 + DR + \gamma^2 G^2, \qquad (14)
\end{aligned}$$

where (a) follows from Cauchy-Schwarz inequality and (b) follows from Assumption 1 and Lemma 1.

By Lemma 5, we have

$$\begin{aligned}
\Delta(t) \leq& [\mathbf{Q}(t)]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}(t+1)) + \|\tilde{\mathbf{g}}(\mathbf{x}(t+1))\|^2 \\
\overset{(a)}{\leq}& [\mathbf{Q}(t)]^{\mathsf{T}}\tilde{\mathbf{g}}(\mathbf{x}(t+1)) + \gamma^2 G^2 \\
\overset{(b)}{\leq}& -\gamma\epsilon\|\mathbf{Q}(t)\| + \gamma^2\epsilon G + \alpha R^2 + DR + 2\gamma^2 G^2,
\end{aligned}$$

where (a) follows from Lemma 1 and (b) follows from (14).

Thus, if $\|\mathbf{Q}(t)\| > \gamma G + \frac{\alpha R^2 + DR + 2\gamma^2 G^2}{\gamma\epsilon}$, then $\Delta(t) < 0$. That is, $\|\mathbf{Q}(t+1)\| < \|\mathbf{Q}(t)\|$. ■

Corollary 1 *Consider online convex optimization with long term constraints under Assumptions 1-4. At each round $t \in \{1, 2, \ldots, \}$ in Algorithm 1,*

$$\|\mathbf{Q}(t)\| \leq 2\gamma G + \frac{\alpha R^2 + DR + 2\gamma^2 G^2}{\gamma\epsilon},$$

where D, G and R are defined in Assumption 1 and $\epsilon > 0$ is defined in Assumption 4.

Proof

Note that $\|\mathbf{Q}(1)\| \overset{(a)}{\leq} \|\mathbf{Q}(0)\| + \|\tilde{\mathbf{g}}(\mathbf{x}(0))\| \overset{(b)}{\leq} \gamma G \leq 2\gamma G + \frac{\alpha R^2 + DR + 2\gamma^2 G^2}{\gamma\epsilon}$, where (a) follows from Lemma 3 and (b) follows from Lemma 1. We need to show $\|\mathbf{Q}(t)\| \leq 2\gamma G + \frac{\alpha R^2 + DR + 2\gamma^2 G^2}{\gamma\epsilon}$ for all rounds $t \geq 2$. This can be proven by contradiction as follows:

Assume that $\|\mathbf{Q}(t)\| > 2\gamma G + \frac{\alpha R^2 + DR + 2\gamma^2 G^2}{\gamma\epsilon}$ happens at some round $t \geq 2$. Let τ be the first (smallest) round index at which this happens, i.e., $\|\mathbf{Q}(\tau)\| > 2\gamma G + \frac{\alpha R^2 + DR + 2\gamma^2 G^2}{\gamma\epsilon}$. Note that $\tau \geq 2$ since we know $\|\mathbf{Q}(1)\| \leq 2\gamma G + \frac{\alpha R^2 + DR + 2\gamma^2 G^2}{\gamma\epsilon}$. The definition of τ implies that $\|\mathbf{Q}(\tau - 1)\| \leq 2\gamma G + \frac{\alpha R^2 + DR + 2\gamma^2 G^2}{\gamma\epsilon}$. Now consider the value of $\|\mathbf{Q}(\tau - 1)\|$ in two cases.

- If $\|\mathbf{Q}(\tau-1)\| > \gamma G + \frac{\alpha R^2+DR+2\gamma^2G^2}{\gamma\epsilon}$, then by Lemma 8, we must have $\|\mathbf{Q}(\tau)\| < \|\mathbf{Q}(\tau-1)\| \leq 2\gamma G + \frac{\alpha R^2+DR+2\gamma^2G^2}{\gamma\epsilon}$. This contradicts the definition of τ.

- If $\|\mathbf{Q}(\tau-1)\| \leq \gamma G + \frac{\alpha R^2+DR+2\gamma^2G^2}{\gamma\epsilon}$, then by part 4 in Lemma 3, we must have $\|\mathbf{Q}(\tau)\| \leq \|\mathbf{Q}(\tau-1)\|+\|\tilde{\mathbf{g}}(\mathbf{x}(\tau))\| \overset{(a)}{\leq} \gamma G+\frac{\alpha R^2+DR+2\gamma^2G^2}{\gamma\epsilon}+\gamma G = 2\gamma G+\frac{\alpha R^2+DR+2\gamma^2G^2}{\gamma\epsilon}$, where (a) follows from Lemma 1. This also contradicts the definition of τ.

In both cases, we have a contradiction. Thus, $\|\mathbf{Q}(t)\| \leq 2\gamma G + \frac{\alpha R^2+DR+2\gamma^2G^2}{\gamma\epsilon}$ for all round $t \geq 2$. ■

3.4. Constraint Violation Analysis

Theorem 2 *Consider online convex optimization with long term constraints under Assumptions 1-4. Let D, β, G, R and ϵ be defined in Assumptions 1-4.*

1. *For all $T \geq 1$, we have*

$$\sum_{t=1}^{T} g_k(\mathbf{x}(t)) \leq 2G + \frac{\alpha R^2 + DR}{\gamma^2\epsilon} + \frac{2G^2}{\epsilon}, \forall k \in \{1, 2, \ldots, m\}.$$

2. *If $\gamma = T^{1/4}$ and $\alpha = \frac{1}{2}[\beta^2+1]\sqrt{T}$ in Algorithm 1, then for all $T \geq 1$, we have*

$$\sum_{t=1}^{T} g_k(\mathbf{x}(t)) \leq 2G + \frac{\frac{1}{2}[\beta^2+1]R^2 + 2G^2}{\epsilon} + \frac{DR}{\epsilon\sqrt{T}}, \forall k \in \{1, 2, \ldots, m\}.$$

Proof Fix $T \geq 1$ and $k \in \{1, 2, \ldots, m\}$. By Lemma 4, we have

$$\sum_{t=1}^{T} g_k(\mathbf{x}(t)) \leq \frac{1}{\gamma}Q_k(T) \leq \frac{1}{\gamma}\|Q_k(T)\| \overset{(a)}{\leq} \frac{2\gamma}{\gamma}G + \frac{\alpha R^2 + DR + 2\gamma^2G^2}{\gamma^2\epsilon},$$

where (a) follows from Corollary 1. Thus, the first part of this theorem follows.

The second part of this theorem follows by substituting $\gamma = T^{1/4}$ and $\alpha = \frac{1}{2}[\beta^2+1]\sqrt{T}$ into the last inequality. ■

3.5. Performance Summary

Theorem 1 and Theorem 2 together imply that if we choose[2] $\gamma = T^{1/4}$ and $\alpha = \frac{1}{2}[\beta^2+1]\sqrt{T}$ in Algorithm 1, then we can achieve $O(\sqrt{T})$ regret and $O(1)$ constraint violations for online convex optimization with long term constraints under Assumptions 1-4.

2. More precisely, to achieve $O(\sqrt{T})$ regret and $O(1)$ constraint violations, it suffices to choose $\gamma = O(T^{1/4})$ and $\alpha = \frac{1}{2}\beta^2\gamma^2 + O(\sqrt{T})$.

Parts (1) of Theorems 1 and 2 suggest that our regret and constraint violations do not rely on the m, which is the number of constraints. However, the constraint violation bound does depend on G^2, where G is defined in Assumption 2. Note that G^2 in the worst case can grow linearly with respect to m.

Note that the $O(1)$ constraint violation bound proven in Theorem 2 is in terms of D, G, R and ϵ defined in Assumptions 1-4. However, the implementation of Algorithm 1 only requires the knowledge of β, which is known to us since the constraint function $\mathbf{g}(\mathbf{x})$ does not change. In contrast, the algorithms developed in Mahdavi et al. (2012) and Jenatton et al. (2016) have parameters that must be chosen based on the knowledge of D, which is usually unknown and can be difficult to estimate in an online optimization scenario.

4. Extensions

This section extends the analysis in the previous section by considering intermediate and unknown time horizon T and by relaxing Assumptions 2-4.

4.1. Intermediate Time Horizon T

Note that parts (1) of Theorems 1 and 2 hold for any T. For large T, choosing $\eta = T^{1/4}$ and $\alpha = \frac{1}{2}[\beta^2 + 1]\sqrt{T}$ yields the $O(\sqrt{T})$ regret bound and $O(1)$ constraint violations as proven in parts (2) of both theorems. For intermediate T, the constant factor hidden in the $O(\sqrt{T})$ bound can be important and the $O(1)$ constraint violation bound can be relatively large. If parameters in Assumptions 1-4 are known, we can obtain the best regret and constraint violation bounds by choosing γ and α as the solution to the following geometric program[3]:

$$\begin{aligned}
\min_{\eta,\gamma,\alpha,z} \quad & z \\
\text{s.t.} \quad & \alpha R^2 + 2\gamma^2 G^2 + \frac{1}{2\eta} D^2 T \leq z, \\
& 2G + \frac{\alpha R^2 + DR + 2\gamma^2 G^2}{\gamma^2 \epsilon} \leq z, \\
& \frac{1}{2}[\beta^2 \gamma^2 + \eta] \leq \alpha, \\
& \eta, \gamma, \alpha, z > 0.
\end{aligned}$$

3. By dividing the first two constraints by z and dividing the third constraint by α on both sides, this geometric program can be written into the standard from of geometric programs. Geometric programs can be reformulated into convex programs and can be efficiently solved. See Boyd et al. (2007) for more discussions on geometric programs.

In certain applications, we can choose γ and α to minimize the regret bound subject to the constraint violation guarantee by solving the following geometric program:

$$\begin{aligned} \min_{\eta,\gamma,\alpha} \quad & \alpha R^2 + 2\gamma^2 G^2 + \frac{1}{2\eta} D^2 T \\ \text{s.t.} \quad & 2G + \frac{\alpha R^2 + DR + 2\gamma^2 G^2}{\gamma^2 \epsilon} \leq z_0, \\ & \frac{1}{2}[\beta^2\gamma^2 + \eta] \leq \alpha, \\ & \eta, \gamma, \alpha > 0, \end{aligned}$$

where $z_0 > 0$ is a constant that specifies the maximum allowed constraint violation. Or alternatively, we can consider the problem of minimizing the constraint violation subject to the regret bound guarantee.

4.2. Unknown Time Horizon T

To achieve $O(\sqrt{T})$ regret and $O(1)$ constraint violations, the parameters γ and α in Algorithm 1 depend on the time horizon T. In the case when T is unknown, we can use the classical "doubling trick" to achieve $O(\sqrt{T})$ regret and $O(\log_2 T)$ constraint violations.

Suppose we have an online convex optimization algorithm $\mathcal{A}$ whose parameters depend on the time horizon. In the case when the time horizon T is unknown, the general doubling trick (Cesa-Bianchi and Lugosi, 2006; Shalev-Shwartz, 2011) is described in Algorithm 2. It is known that the doubling trick can preserve the order of algorithm $\mathcal{A}$'s regret bound in the case when the time horizon T is unknown. The next theorem summarizes that by using the "doubling trick" for Algorithm 1 with unknown time horizon T, we can achieve $O(\sqrt{T})$ regret and $O(\log_2 T)$ constraint violations.

Algorithm 2 The Doubling Trick (Cesa-Bianchi and Lugosi, 2006; Shalev-Shwartz, 2011)

- Let algorithm $\mathcal{A}$ be an algorithm whose parameters depend on the time horizon. Let $i = 1$.
- Repeat until we reach the end of the time horizon
  - Run algorithm $\mathcal{A}$ for 2^i rounds by using 2^i as the time horizon.
  - Let $i = i + 1$.

Theorem 3 *If the time horizon T is unknown, then applying Algorithm 1 with the "doubling trick" can yield $O(\sqrt{T})$ regret and $O(\log_2 T)$ constraint violations.*

Proof Let T be the unknown time horizon. Define each iteration in the doubling trick as a period. Since the i-th period consists of 2^i rounds, we have in total $\lceil \log_2 T \rceil$ periods, where $\lceil x \rceil$ denotes the smallest integer no less than x.

1. The proof of $O(\sqrt{T})$ regret is almost identical to the classical proof. By Theorem 1, there exists a constant C such that the regret in the i-th period is at most $C\sqrt{2^i}$. Thus, the total regret is at most

$$\begin{aligned}\sum_{i=1}^{\lceil \log_2 T\rceil} C\sqrt{2^i} =& C\frac{\sqrt{2}[1-\sqrt{2}^{\lceil \log_2 T\rceil}]}{1-\sqrt{2}}\\ =& \frac{\sqrt{2}C}{\sqrt{2}-1}[\sqrt{2}^{\lceil \log_2 T\rceil}-1]\\ \leq& \frac{\sqrt{2}C}{\sqrt{2}-1}\sqrt{2}^{1+\log_2 T}\\ \leq& \frac{2C}{\sqrt{2}-1}\sqrt{T}\end{aligned}$$

   Thus, the regret bound is $O(\sqrt{T})$ when using the "doubling trick".

2. The proof of $O(\log_2 T)$ constraint violations is simple. By Theorem 1, there exists a constant C such that the constraint violation in the i-th period is at most C. Since we have $\lceil \log_2 T\rceil$ periods, the total constraint violation is $C\lceil \log_2 T\rceil$.

■

4.3. Relaxing Assumptions 3-4

Note that previous works Mahdavi et al. (2012) and Jenatton et al. (2016) consider online convex optimization with long term constraints under Assumptions 1-3 and the additional assumption that all $f^t(\cdot)$ functions are bounded without imposing Assumption 4.

In this subsection, we show that if we are allowed to introduce the bounded $f^t(\cdot)$ assumption (formally defined in Assumption 5) used in Mahdavi et al. (2012) and Jenatton et al. (2016) , then Algorithm 1 can still achieve a superior performance without imposing Assumptions 3-4.

Assumption 5 *There exists a constant F such that $|f^t(\mathbf{x}) - f^t(\mathbf{y})| \leq F$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{X}_0$ and all $t \in \{1, 2, \ldots\}$.*

Recall that the $O(\sqrt{T})$ regret summarized in Theorem 1 holds regardless of Assumptions 2-4. Now, it remains to show the constraint violation of Algorithm 1 under Assumption 1 and Assumption 2, both of which are used in Mahdavi et al. (2012) and Jenatton et al. (2016).

Theorem 4 *Consider online convex optimization with long term constraints under Assumptions 1-2 and 5. If we choose $\gamma = T^{1/4}$ and $\alpha = \frac{1}{2}[\beta^2 + 1]\sqrt{T}$ in Algorithm 1, we have*

$$\sum_{t=1}^{T} g_k(\mathbf{x}(t)) \leq O(T^{1/4}), \forall k \in \{1, 2, \ldots, m\}.$$

Proof Fix $T \geq 1$. Let $\gamma > 0, \eta > 0$ be arbitrary. If $\alpha \geq \frac{1}{2}[\gamma^2\beta^2 + \eta]$, by Lemma 7, we have

$$\begin{aligned}&\Delta(t) + f^t(\mathbf{x}(t))\\ \leq& f^t(\mathbf{x}^*) + \alpha[\|\mathbf{x}^* - \mathbf{x}(t)\|^2 - \|\mathbf{x}^* - \mathbf{x}(t+1)\|^2] + \frac{1}{2}[\|\tilde{\mathbf{g}}(\mathbf{x}(t+1))\|^2 - \|\tilde{\mathbf{g}}(\mathbf{x}(t))\|^2] + \frac{1}{2\eta}D^2.\end{aligned}$$

Summing over $t \in \{1, 2, \ldots, T-1\}$ and rearranging terms yields

$$\begin{aligned}\sum_{t=1}^{T-1}\Delta(t) \leq& \sum_{t=1}^{T-1}\left[f^t(\mathbf{x}^*) - f^t(\mathbf{x}^t)\right] + \alpha\sum_{t=1}^{T-1}[\|\mathbf{x}^* - \mathbf{x}(t)\|^2 - \|\mathbf{x}^* - \mathbf{x}(t+1)\|^2]\\ &+ \sum_{t=1}^{T-1}\frac{1}{2}[\|\tilde{\mathbf{g}}(\mathbf{x}(t+1))\|^2 - \|\tilde{\mathbf{g}}(\mathbf{x}(t))\|^2] + \frac{1}{2\eta}[T-1]D^2\\ =& \sum_{t=1}^{T-1}\left[f^t(\mathbf{x}^*) - f^t(\mathbf{x}^t)\right] + \alpha\|\mathbf{x}^* - \mathbf{x}(1)\|^2 - \alpha\|\mathbf{x}^* - \mathbf{x}(T)\|^2\\ &+ \frac{1}{2}[\|\tilde{\mathbf{g}}(\mathbf{x}(T))\|^2 - \|\tilde{\mathbf{g}}(\mathbf{x}(1))\|^2] + \frac{1}{2\eta}[T-1]D^2\end{aligned}$$

Substituting $\Delta(t) = \frac{1}{2}\|\mathbf{Q}(t+1)\|^2 - \frac{1}{2}\|\mathbf{Q}(t)\|^2$ into it, simplifying the telescoping sum, and rearranging terms yields

$$\begin{aligned}\frac{1}{2}\|\mathbf{Q}(T)\|^2 \leq& \sum_{t=1}^{T-1}\left[f^t(\mathbf{x}^*) - f^t(\mathbf{x}^t)\right] + \alpha\|\mathbf{x}^* - \mathbf{x}(1)\|^2 - \alpha\|\mathbf{x}^* - \mathbf{x}(T)\|^2\\ &+ \frac{1}{2}\|\tilde{\mathbf{g}}(\mathbf{x}(T))\|^2 - \frac{1}{2}\|\tilde{\mathbf{g}}(\mathbf{x}(1))\|^2 + \frac{1}{2}\|\mathbf{Q}(1)\|^2 + \frac{1}{2\eta}[T-1]D^2\\ \overset{(a)}{=}& \sum_{t=1}^{T-1}\left[f^t(\mathbf{x}^*) - f^t(\mathbf{x}^t)\right] + \alpha\|\mathbf{x}^* - \mathbf{x}(1)\|^2 - \alpha\|\mathbf{x}^* - \mathbf{x}(T)\|^2\\ &+ \frac{1}{2}\gamma^2\|\mathbf{g}(\mathbf{x}(T))\|^2 + \frac{1}{2\eta}[T-1]D^2\\ \overset{(b)}{\leq}& [T-1]F + \alpha\|\mathbf{x}^* - \mathbf{x}(1)\|^2 + \frac{1}{2}\gamma^2G^2 + \frac{1}{2\eta}[T-1]D^2\end{aligned}$$

where (a) follows because $\|\tilde{\mathbf{g}}(\mathbf{x}(1))\|^2 = \|\mathbf{Q}(1)\|^2$ by Lemma 3; and (b) follows from Assumption 2 and Assumption 5.

Multiplying both sides by a factor of 2 and taking square roots on both sides yields

$$\|\mathbf{Q}(T)\| \leq \sqrt{2[T-1]F} + \sqrt{2\alpha}\|\mathbf{x}^* - \mathbf{x}(1)\| + \gamma G + D\sqrt{\frac{T-1}{\eta}} \tag{15}$$

Fix $k \in \{1, 2, \ldots, m\}$. By Lemma 4, we have

$$\begin{aligned}\sum_{t=1}^{T} g_k(\mathbf{x}(t)) \leq & \frac{1}{\gamma} Q_k(T) \\ \leq & \frac{1}{\gamma} \|Q_k(T)\| \\ \overset{(a)}{\leq} & \frac{\sqrt{2[T-1]F}}{\gamma} + \frac{\sqrt{2\alpha}}{\gamma} \|\mathbf{x}^* - \mathbf{x}(1)\| + G + \frac{D}{\gamma}\sqrt{\frac{T-1}{\eta}}\end{aligned}$$

where (a) follows from (15).

Note that if we let $\gamma = T^{1/4}$ and $\eta = \sqrt{T}$, then $\alpha = \frac{1}{2}[\beta^2 + 1]\sqrt{T} \geq \frac{1}{2}[\gamma^2\beta^2 + \eta]$. Substituting these values into the above equation yields

$$\begin{aligned}\sum_{t=1}^{T} g_k(\mathbf{x}(t)) \leq & \frac{\sqrt{2[T-1]F}}{T^{1/4}} + \frac{\sqrt{[\beta^2+1]\sqrt{T}}}{T^{1/4}} \|\mathbf{x}^* - \mathbf{x}(1)\| + G + \frac{D}{T^{1/4}}\sqrt{\frac{T-1}{\sqrt{T}}} \\ = & O(T^{1/4})\end{aligned}$$

■

Remark 1 *Theorem 1 and Theorem 4 together imply that if we choose $\gamma = T^{1/4}$ and $\alpha = \frac{1}{2}[\beta^2+1]\sqrt{T}$ in Algorithm 1, then we can achieve $O(\sqrt{T})$ regret and $O(T^{1/4})$ constraint violations for online convex optimization with long term constraints under Assumptions 1-2 and 5. This is still uniformly better than the best known $O(T^{\max\{\theta,1-\theta\}})$ regret and $O(T^{1-\theta/2})$ constraint violations for all $\theta \in (0,1)$ established in Jenatton et al. (2016) under Assumptions 1-3 and 5. We further note the parameters used to achieve $O(T^{1/4})$ constraint violations under Assumption 5 are identical to those used in Section 3 to achieve $O(1)$ constraint violations under Assumption 4. Thus, Algorithm 1 is very adaptive and its practical implementation can be blind to Assumption 4 or Assumption 5.*

5. Experiment

This section considers numerical experiments to verify the performance of our algorithm. Consider online convex optimization with loss functions $f^t(\mathbf{x}) = \mathbf{c}(t)^{\mathsf{T}}\mathbf{x}$, where $\mathbf{c}(t)$ is time-varying and unknown at round t; and constraint functions $\mathbf{A}\mathbf{x} \leq \mathbf{b}$. The constraint functions are only required to be satisfied in long term:

$$\limsup_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbf{A}\mathbf{x}(t) \leq \mathbf{b}.$$

The above problem formulation arises often in fields such as resource allocation, product planning, finance portfolio selection, network scheduling, load distribution, and so on (Ibaraki and Katoh, 1988). For example, consider a power grid network where the electricity

generation at each power plant is scheduled in real-time, e.g., hour-by-hour. In this problem, each component $\mathbf{x}_i$ corresponds to the amount of electricity produced by the i-th power plant. The time-varying loss function $f^t(\cdot)$, which represents the economic loss/reward depending on the real-time power demand, is in general unknown to the decision maker at the beginning of round t. Inequality constraint $\mathbf{Ax} \leq \mathbf{b}$ corresponds to constraints such as fuel consumption, man-power consumption, carbon emission and electricity scheduling.

In the numerical experiment, we assume $\mathbf{x} \in \mathbb{R}^2$, $\mathbf{A} \in \mathbb{R}^{3\times 2}$; each component of $\mathbf{x}$ satisfies the box constraint $\mathbf{x} \in \mathcal{X}_0$ where $\mathcal{X}_0 = \{\mathbf{x} : -1 \leq x_1 \leq 1, -1 \leq x_2 \leq 1\}$; and $T = 5000$. Each component of $\mathbf{A}$ is generated from the uniform distribution in the interval $[0, 1]$ and each component of $\mathbf{b}$ is generated from the uniform distribution in the interval $[0, 2]$. $\mathbf{A}$ and b are kept fixed for all rounds once they are generated. To simulate arbitrarily varying objective functions, at each round t, $\mathbf{c}(t) = \mathbf{c}^{(1)}(t) + \mathbf{c}^{(2)}(t) + \mathbf{c}^{(3)}(t)$ is randomly generated such that each component $\mathbf{c}^{(1)}(t)$ is from the uniform distribution in the interval $[-t^{1/10}, +t^{1/10}]$; each component of $\mathbf{c}^{(2)}(t)$ is from the uniform distribution in the interval $[-1, 0]$ when $t \in [1, 1500] \cup [2000, 3500] \cup [4000, 5000]$ and is from the uniform distribution in the interval $[0, 1]$ otherwise; and each element of $\mathbf{c}^{(3)}(t)$ is equal to $(-1)^{\mu(t)}$ where $\mu(t)$ is a random permutation of vector $[1 : 5000]$

We run our proposed Algorithm 1, our proposed Algorithm 1 with the doubling trick (without knowing $T = 5000$), Algorithm 1 in Mahdavi et al. (2012) and the Algorithm in Jenatton et al. (2016) with $\beta = 1/2$ and $\beta = 2/3$ over 1000 independent experiments generated from the above distribution setting. Figure 1 and Figure 2 plot the cumulative regret and the cumulative constraint violations (averaged over 1000 independent experiments), respectively. Figure 1 shows that the first 3 algorithms have similar regret since they are all proven to have $O(\sqrt{T})$ regret and the Algorithm in Jenatton et al. (2016) with $\beta = 2/3$ has the largest regret since it has $O(T^{2/3})$ regret. Figure 2 shows that our algorithm has the smallest constraint violation since the constraint violation of our algorithm is bounded by a constant and does not grow with T while the other algorithms have $O(T^{3/4})$ or $O(T^{2/3})$ constraint violations.

6. Conclusion

This paper considers online convex optimization with long term constraints, where functional constraints are only required to be satisfied in the long term. Prior algorithms in Mahdavi et al. (2012) can achieve $O(\sqrt{T})$ regret and $O(T^{3/4})$ constraint violations for general problems and achieve $O(T^{2/3})$ bounds for both regret and constraint violations when the constraint set can be described by a finite number of linear constraints. A recent extension in Jenatton et al. (2016) can achieve $O(T^{\max\{\theta,1-\theta\}})$ regret and $O(T^{1-\theta/2})$ constraint violations where $\theta \in (0, 1)$. This paper proposes a new algorithm that can achieve an $O(\sqrt{T})$ bound for regret and an $O(1)$ bound for constraint violations; and hence yields improved performance in comparison to the prior works (Mahdavi et al., 2012; Jenatton et al., 2016).

7. Acknowledgment

This work is supported in part by grant NSF CCF-1718477.

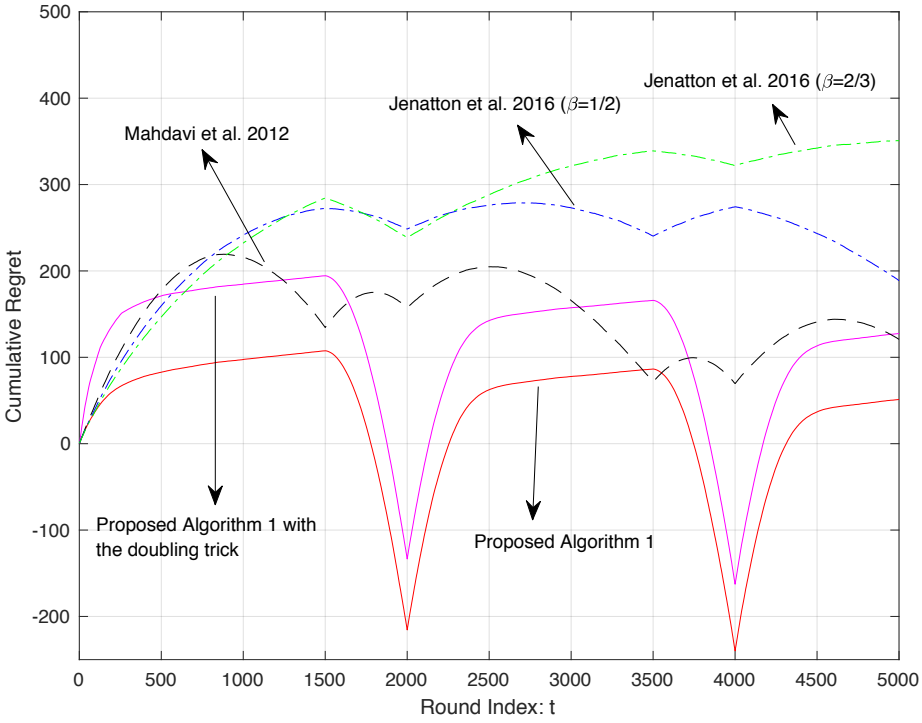

Figure 1: The cumulative regret.

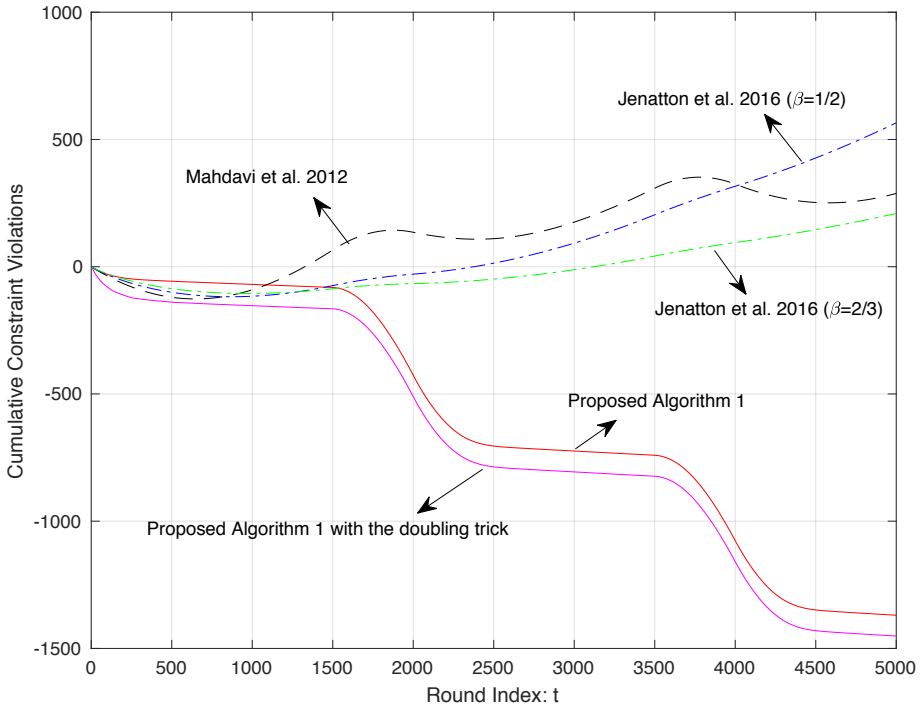

Figure 2: The cumulative constraint violations of the long term constraints.

References

Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

Stephen Boyd, Seung-Jean Kim, Lieven Vandenberghe, and Arash Hassibi. A tutorial on geometric programming. *Optimization and Engineering*, 8(1):67–127, 2007.

Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.

Arpita Ghosh, Preston McAfee, Kishore Papineni, and Sergei Vassilvitskii. Bidding for representative allocations for display advertising. In *Proceedings of International Workshop on Internet and Network Economics (WINE)*, 2009.

Avi Goldfarb and Catherine Tucker. Online display advertising: targeting and obtrusiveness. *Marketing Science*, 30(3):389–404, 2011.

Elad Hazan and Satyen Kale. Projection-free online learning. In *Proceedings of International Conference on Machine Learning (ICML)*, 2012.

Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69:169–192, 2007.

Toshihide Ibaraki and Naoki Katoh. *Resource Allocation Problems: Algorithmic Approaches*. MIT Press, 1988.

Rodolphe Jenatton, Jim Huang, and Cédric Archambeau. Adaptive algorithms for online convex optimization with long-term constraints. In *Proceedings of International Conference on Machine learning (ICML)*, 2016.

Mehrdad Mahdavi, Rong Jin, and Tianbao Yang. Trading regret for efficiency: online convex optimization with long term constraints. *Journal of Machine Learning Research*, 13(1):2503–2528, 2012.

Michael J. Neely. *Dynamic Power Allocation and Routing for Satellite and Wireless Networks with Time Varying Channels*. PhD thesis, Massachusetts Institute of Technology, 2003.

Michael J. Neely. *Stochastic network optimization with application to communication and queueing systems*. Morgan & Claypool Publishers, 2010.

Shai Shalev-Shwartz. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2):107–194, 2011.

Hao Yu and Michael J. Neely. A simple parallel algorithm with an $O(1/t)$ convergence rate for general convex programs. *SIAM Journal on Optimization*, 27(2):759–783, 2017.

Hao Yu, Michael Neely, and Xiaohan Wei. Online convex optimization with stochastic constraints. In *Advances in Neural Information Processing Systems (NIPS)*, 2017.

Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of International Conference on Machine Learning (ICML)*, 2003.