

Gradient-Variation Bound for Online Convex Optimization with Constraints

Shuang Qiu¹, Xiaohan Wei², Mladen Kolar¹

¹ Booth School of Business, the University of Chicago

²Meta Platforms, Inc.

qius@uchicago.edu, ubimeteor@fb.com, mkolar@chicagobooth.edu

Abstract

We study online convex optimization with constraints consisting of multiple functional constraints and a relatively simple constraint set, such as a Euclidean ball. As enforcing the constraints at each time step through projections is computationally challenging in general, we allow decisions to violate the functional constraints but aim to achieve a low regret and cumulative violation of the constraints over a horizon of T time steps. First-order methods achieve an $\mathcal{O}(\sqrt{T})$ regret and an $\mathcal{O}(1)$ constraint violation, which is the best-known bound under the Slater’s condition, but do not take into account the structural information of the problem. Furthermore, the existing algorithms and analysis are limited to Euclidean space. In this paper, we provide an *instance-dependent* bound for online convex optimization with complex constraints obtained by a novel online primal-dual mirror-prox algorithm. Our instance-dependent regret is quantified by the total gradient variation $V_*(T)$ in the sequence of loss functions. The proposed algorithm works in *general* normed spaces and simultaneously achieves an $\mathcal{O}(\sqrt{V_*(T)})$ regret and an $\mathcal{O}(1)$ constraint violation, which is never worse than the best-known ($\mathcal{O}(\sqrt{T}), \mathcal{O}(1)$) result and improves over previous works that applied mirror-prox-type algorithms for this problem achieving $\mathcal{O}(T^{2/3})$ regret and constraint violation. Finally, our algorithm is computationally efficient, as it only performs mirror descent steps in each iteration instead of solving a general Lagrangian minimization problem.

Introduction

We study online convex optimization (OCO) with a sequence of loss functions f^1, f^2, \dots that arbitrarily vary over time. The decision maker chooses an action \mathbf{x}_t from a set \mathcal{X} and then observes the loss function f^t at each time step t . The goal is to minimize the regret over the T time steps, which is defined as

$$\text{Regret}(T) = \sum_{t=1}^T f^t(\mathbf{x}_t) - \sum_{t=1}^T f^t(\mathbf{x}^*), \quad (1)$$

where \mathbf{x}_t is the decision chosen in step t , and $\mathbf{x}^* \in \arg\min_{\mathbf{x} \in \mathcal{X}} \sum_{t=1}^T f^t(\mathbf{x})$ is the best decision in hindsight. The regret in (1) compares the sequence of decisions with

a best strategy \mathbf{x}^* in hindsight for all loss functions over T time steps to measure the performance of an online learning algorithm.

This problem has been extensively studied in existing work (Cesa-Bianchi, Long, and Warmuth 1996; Gordon 1999; Zinkevich 2003a; Hazan 2016a). Online mirror descent (OMD) is a commonly used first-order algorithm that subsumes online gradient descent (OGD) and achieves $\mathcal{O}(\sqrt{T})$ regret with a dependence on the dimension related to the chosen norm in the optimization space and logarithmic dependence on the probability simplex (Hazan 2016b). Existing work in OCO has also focused on characterizing an *instance-dependent* regret characterized by the notion of *gradient variation* (Chiang et al. 2012; Yang et al. 2014; Steinhardt and Liang 2014). Specifically, these works used first-order methods and characterized bounds on the regret in terms of the gradient-variation of the function gradient sequence. Compared to the \sqrt{T} -type bound, the gradient-variation bound explicitly takes into account the dynamics of the observed losses, which is the structural information of the problem. For example, Chiang et al. (2012) obtained a regret that scales as $(\sum_{t=1}^T \max_{\mathbf{x} \in \mathcal{X}} \|\nabla f^t(\mathbf{x}) - \nabla f^{t-1}(\mathbf{x})\|_2^2)^{1/2}$, which reduces to the $\mathcal{O}(\sqrt{T})$ regret bound only in the worst case and is better when the variation is small.

In this paper, we consider a more challenging OCO problem where the feasible set \mathcal{X} consists of not only a simple compact set \mathcal{X}_0 but also of K complex functional constraints,

$$\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^d : \mathbf{x} \in \mathcal{X}_0 \text{ and } g_k(\mathbf{x}) \leq 0, \forall k \in [K]\}, \quad (2)$$

where the k -th constraint function $g_k(\mathbf{x})$ is convex and differentiable. OMD with a simple projection does not work well in this problem, as projection onto the complex constraint set \mathcal{X} is usually computationally heavy. Rather than requiring each decision to be feasible, it is common to allow functional constraints to be slightly violated at each time step (Mahdavi, Jin, and Yang 2012; Jenatton, Huang, and Archambeau 2016; Yu, Neely, and Wei 2017; Chen, Ling, and Giannakis 2017; Liakopoulos et al. 2019) but require an algorithm to simultaneously maintain a sublinear regret and constraint violation. Specifically, in addition to (1), we also

Complex Constraint	Paper	Regret	Constraint Violation	Efficient	Space
✗	Standard OMD (Hazan 2016b)	\sqrt{T}	-	✓	General
	Chiang et al. (2012)	$\sqrt{V_2(T)} \vee L_f$			Euclidean Prob. Simp.
	Steinhardt and Liang (2014)	$\sqrt{B_{i^b}(T)} (\diamond)$			Euclidean
	Yang et al. (2014)	$\sqrt{V_*(T)} \vee L_f$			General
✓	Mahdavi, Jin, and Yang (2012)	$\frac{\sqrt{T}}{T^{2/3}}$	$\frac{T^{3/4}}{T^{2/3}}$	✓	Euclidean
	Jenatton, Huang, and Archambeau (2016)	$T^{\max\{\beta, 1-\beta\}}$	$T^{1-\beta/2}$		
	Yu, Neely, and Wei (2017)	\sqrt{T}	\sqrt{T}		
	Yuan and Lamperski (2018) (●)	$T^{\max\{\beta, 1-\beta\}}$	$T^{1-\beta/2}$		
	Yi et al. (2021) (●)	$T^{\max\{\beta, 1-\beta\}}$	$T^{(1-\beta)/2}$		
	Wei, Yu, and Neely (2020)	$\sqrt{T} (\star)$	$\sqrt{T} (\star)$	✓	General
	Yu and Neely (2020)	\sqrt{T}	<i>const.</i>	✗	Euclidean
	This work	$\sqrt{V_*(T)} \vee L_f (\star)$	<i>const.</i> (★)	✓	General

Table 1: Comparison with existing works. We use *const.* to denote a constant bound $\mathcal{O}(1)$. The parameter β satisfies $\beta \in (0, 1)$. “Complex Constraint” indicates whether a projection on the constraint set is computationally inefficient. “Efficient” indicates whether each round only involves gradient updates to compute the decision (see the discussion in Remark 1) such that the algorithm is computationally efficient. “Prob. Simp.” means that the bound is for the probability simplex case, which is one special case of the general space scenario. The quantity L_f is the Lipschitz constant for the gradient of the loss function, i.e., ∇f^t for any $t \geq 0$. We let (\diamond) indicate that the bound is only for the linear loss function, in which case we also have $L_f = 0$. We let (\star) indicate that a $\log T$ factor is imposed on the presented bound under the probability simplex setting. We let (\bullet) indicate another line of constrained OCO work, which is based on a different and stricter metric for constraint violation. We let $a \vee b$ denote $\max\{a, b\}$. In addition, $B_{i^b}(T)$ is another type of gradient-variation bound. See related work for a detailed description.

look at the following constraint violation

$$\text{Violation}(T, k) := \sum_{t=1}^T g_k(\mathbf{x}_t), \forall k \in \{1, \dots, K\}, \quad (3)$$

and aim to achieve a sublinear growth rate. In this setting, Yu and Neely (2020) achieved the best-known $\mathcal{O}(\sqrt{T})$ regret and $\mathcal{O}(1)$ constraint violation under the Slater’s condition, but was limited to the Euclidean space. Their algorithm solves a general Lagrangian minimization problem during each round, which requires costly inner loops to approximate the solutions. Yu, Neely, and Wei (2017) and Wei, Yu, and Neely (2020) both showed $\mathcal{O}(\sqrt{T})$ regret and constraint violation that applies to more general non-Euclidean spaces with stochastic constraints. However, the study of the instance-dependent bound for OCO with complex functional constraints remains unsatisfactory in the existing works. Our work aims to answer the following question:

Can we obtain a gradient-variation bound for OCO with complex constraints via efficient first-order methods?

We provide an affirmative answer to this question. We propose a novel online primal-dual mirror-prox method and prove a strong theoretical result that it can achieve a gradient-variation regret bound and simultaneously maintain the $\mathcal{O}(1)$ constraint violation in a general normed space. In the worst case, the upper bound matches the best-known ($\mathcal{O}(\sqrt{T}), \mathcal{O}(1)$) result under the Slater’s condition.

The mirror-prox algorithm (Nemirovski 2004) features two mirror-descent or gradient-descent steps bridged by an intermediate iterate, and it achieves $\mathcal{O}(1/T)$ convergence rate when minimizing deterministic smooth convex functions (Bubeck et al. 2015). Our bound also establishes $\mathcal{O}(1/T)$ convergence rate, measured by $\text{Regret}(T)/T$, when f^t is the same over $t \geq 1$, which is consistent with the result in Bubeck et al. (2015).

Obtaining our theoretical result is quite a challenge. First, although existing work using mirror-prox algorithms for the simple constraint setting can achieve the gradient-variation regret (Chiang et al. 2012), it is not obvious that this result holds in our setting due to the coupling of the primal and dual updates. According to Yu and Neely (2020), the regret bound depends on the drift of the dual iterates, which is only on the order of \sqrt{T} . Second, for the general non-Euclidean space setting, Wei, Yu, and Neely (2020) only achieved an $\mathcal{O}(\sqrt{T})$ constraint violation for stochastic constraints, and it is not obvious how to further improve this bound in a deterministic constraint setting. Moreover, Mahdavi, Jin, and Yang (2012) applied the mirror-prox algorithm to the constrained OCO and obtained an $\mathcal{O}(T^{2/3})$ regret and constrained violation, which is suboptimal. Our work provides a novel theoretical analysis for the drift-plus-penalty framework (Yu and Neely 2017) with gradient variation and a tight dual variable drift such that gradient-variation regret and constant constraint violation can be achieved.

Contributions. Our theoretical contributions are 3-fold:

- We propose a novel online primal-dual mirror-prox method that can simultaneously achieve an $\mathcal{O}(\max\{\sqrt{V_*(T)}, L_f\})$ regret and an $\mathcal{O}(1)$ constraint violation in a general normed space $(\mathcal{X}_0, \|\cdot\|)$ under Slater’s condition. Here L_f is the Lipschitz constant for ∇f^t and $V_*(T)$ is the gradient variation defined as

$$V_*(T) = \sum_{t=1}^T \max_{\mathbf{x} \in \mathcal{X}_0} \|\nabla f^t(\mathbf{x}) - \nabla f^{t-1}(\mathbf{x})\|_*^2, \quad (4)$$

where $\|\cdot\|_*$ is the dual norm w.r.t. $\|\cdot\|$, that is, $\|\mathbf{x}\|_* := \sup_{\|\mathbf{y}\| \leq 1} \langle \mathbf{x}, \mathbf{y} \rangle$. We can write $* = p$ with $p \geq 1$ for different dual norms. We further show that in the probability simplex case, only additional factors of $\log T$ occur.

- Even in the worst case, $V_*(T)$ is at the level of $\mathcal{O}(T)$ if $\|\nabla f^t(\mathbf{x})\|_*$ is bounded in \mathcal{X}_0 . Thus, our bound is never worse than the best-known $(\mathcal{O}(\sqrt{T}), \mathcal{O}(1))$ bound under the Slater’s condition and can be better when the variation is small. Our bound improves over the $(\mathcal{O}(T^{2/3}), \mathcal{O}(T^{2/3}))$ bound (Mahdavi, Jin, and Yang 2012) for OCO with complex constraints via the mirror-prox method. Our work also has a better constraint violation compared to $\mathcal{O}(\sqrt{T})$ in Wei, Yu, and Neely (2020) in the general normed space.
- Our method can be efficiently implemented in that our method only involves two mirror descent steps each round with the local linearization of the constraint functions. This is in stark contrast to Yu and Neely (2020) achieving the best-known rate, which requires solving a general Lagrangian minimization problem that involves entire constraint functions each round. See Table 1 for detailed comparisons.

Related Work

Online convex optimization (OCO) has been widely investigated. Various methods have been proposed to achieve an $\mathcal{O}(\sqrt{T})$ regret in different scenarios (Hazan 2016b). Beyond the \sqrt{T} -type regret, recent literature investigated instance-dependent bounds, where the upper bound on regret is expressed in terms of the variation defined by the sequence of the observed losses (Hazan and Kale 2010; Chiang et al. 2012; Yang et al. 2014; Steinhardt and Liang 2014). Hazan and Kale (2010) and Yang et al. (2014) studied regret characterized by variations of the loss function sequence $\{f^t\}_{t=1}^T$. Chiang et al. (2012) defined the gradient-variation as (4) and studied both linear and smooth convex losses in Euclidean space to get an $\mathcal{O}(\sqrt{V_2(T)})$ bound. Chiang et al. (2012) investigated the linear loss in the probability simplex setting, obtaining an $\mathcal{O}(\sqrt{V_\infty(T)})$ regret. Yang et al. (2014) analyzed the gradient-variation regret for the smooth convex losses and obtained an $\mathcal{O}(\sqrt{V_*(T)})$ bound in a general non-Euclidean space. Steinhardt and Liang (2014) achieved a different gradient-variation bound $\sqrt{B_{i^b}(T)}$ with $B_{i^b}(T) := \sum_{t=1}^T (\mathbf{z}_{t,i^b} - \mathbf{z}_{t-1,i^b})^2$ for the setting of linear losses $f^t(\mathbf{x}) = \langle \mathbf{x}, \mathbf{z}_t \rangle$, $t \in [T]$, where $i^b := \operatorname{argmin}_{i \in [d]} \sum_{t=1}^T \mathbf{z}_{t,i}$ with $\mathbf{z}_t \in \mathbb{R}^d$.

Our work is closely related to OCO with complex functional constraints, e.g., Mahdavi, Jin, and Yang (2012); Jenatton, Huang, and Archambeau (2016); Yu, Neely, and Wei (2017); Chen, Ling, and Giannakis (2017); Yuan and Lamperski (2018); Liakopoulos et al. (2019); Yi et al. (2021); Yu and Neely (2020). Mahdavi, Jin, and Yang (2012) proposed an primal-dual algorithm in Euclidean space that achieved an $\mathcal{O}(\sqrt{T})$ regret and an $\mathcal{O}(T^{3/4})$ constraint violation and further proposed a mirror-prox-type algorithm that obtained both $\mathcal{O}(T^{2/3})$ regret and constraint violation. Jenatton, Huang, and Archambeau (2016) obtained $\mathcal{O}(T^{\max\{\beta, 1-\beta\}})$ regret and $\mathcal{O}(T^{1-\beta/2})$ constraint violation through a primal-dual gradient-descent-type algorithm where $\beta \in (0, 1)$. Yu and Neely (2020) improved this bound and obtained $\mathcal{O}(\sqrt{T})$ regret and $\mathcal{O}(1)$ constraint violation. But it relied on solving a general Lagrangian minimization problem each round that is computationally inefficient. We remark that Yu and Neely (2020) gave the best-known bound under the Slater’s condition and our work uses the same setting. On the other hand, with a stricter constraint violation metric and without the Slater’s condition, Yuan and Lamperski (2018) obtained $\mathcal{O}(T^{\max\{\beta, 1-\beta\}})$ regret and $\mathcal{O}(T^{1-\beta/2})$ constraint violation and Yi et al. (2021) further improved the bound to $(\mathcal{O}(T^{\max\{\beta, 1-\beta\}}), \mathcal{O}(T^{(1-\beta)/2}))$ and $(\mathcal{O}(T^{\max\{\beta, 1-\beta\}}), \mathcal{O}(\sqrt{T}))$ where $\beta \in (0, 1)$. In addition, Yu, Neely, and Wei (2017) and Wei, Yu, and Neely (2020) showed both $\mathcal{O}(\sqrt{T})$ regret and constraint violation, but they studied the setting of stochastic constraints which subsumes the fixed constraint case as here. Note that Wei, Yu, and Neely (2020) studied constrained OCO in a general normed space other than Euclidean space as in other works.

Another line of work on OCO problems considered a different regret metric, namely the dynamic regret, which is distinct from the definition in (1) (Zinkevich 2003b; Hall and Willett 2013; Zhao et al. 2021; Yang et al. 2016; Zhao et al. 2020; Zhang, Lu, and Zhou 2018; Zhang et al. 2017, 2018; Baby and Wang 2022; Cheng et al. 2020; Chang and Shahrampour 2021; Hazan and Seshadhri 2007; Daniely, Gonen, and Shalev-Shwartz 2015; Besbes, Gur, and Zeevi 2015; Jadbabaie et al. 2015; Baby, Zhao, and Wang 2021; Zhao and Zhang 2021; Baby and Wang 2021; Goel and Wierman 2019; Mokhtari et al. 2016; Zhao, Wang, and Zhou 2022; Chen, Wang, and Wang 2019; Yi et al. 2021). The dynamic regret is defined as $\operatorname{Regret}(T) := \sum_{t=1}^T f^t(\mathbf{x}_t) - \sum_{t=1}^T f^t(\mathbf{x}_t^*)$, where the comparators are the minimizers for each individual loss function, i.e., $\mathbf{x}_t^* := \operatorname{argmin}_{\mathbf{x} \in \mathcal{X}} f^t(\mathbf{x})$, instead of the minimizer for the summation of all losses over T time slots as in our problem. Their regret bound is associated with the overall variation of the loss functions $\{f^t\}_{t=1}^T$ or the comparators $\{\mathbf{x}_t^*\}_{t=1}^T$ in T time steps. Thus, our setting and analysis are fundamentally different than the ones for the dynamic regret.

Problem Setup

Consider an OCO problem with long-term complex constraints. Suppose that the feasible set $\mathcal{X}_0 \subset \mathbb{R}^d$ is convex and compact, and that there are K long-term fixed constraints

$g_k(\mathbf{x}) \leq 0, k \in [K]$, which comprise the set \mathcal{X} defined as (2). At each round t , after generating a decision \mathbf{x}_t , the decision maker will observe a new loss function $f^t : \mathcal{X}_0 \mapsto \mathbb{R}$.¹ Our goal is to propose an efficient learning algorithm to generate a sequence of iterates $\{\mathbf{x}_t\}_{t \geq 0}$ within \mathcal{X}_0 , such that the regret $\text{Regret}(T)$ and constraint violation $\text{Violation}(T, k)$, $k \in [K]$, defined in (1) and (3) grow sublinearly w.r.t. the total number of rounds T .

We denote $\mathbf{g}(\mathbf{x}) = [g_1(\mathbf{x}), g_2(\mathbf{x}), \dots, g_K(\mathbf{x})]^\top$ as a vector constructed by stacking the values of the constraint functions at a point \mathbf{x} . We let $\|\cdot\|$ be a norm, with $\|\cdot\|_*$ denoting its dual norm. We let $[n]$ denote the set $\{1, 2, \dots, n\}$. We use $a \vee b$ to denote $\max\{a, b\}$. Then, for the constraints and loss functions, we make several common assumptions (Chiang et al. 2012; Hazan 2016b; Yu and Neely 2020).

Assumption 1. Assume that the set \mathcal{X}_0 , the functions f^t and g_k , $k \in [K]$ satisfy the following assumptions:

- The set \mathcal{X}_0 is convex and compact.
- The gradient of the loss function f^t is bounded: $\|\nabla f^t(\mathbf{x})\|_* \leq F, \mathbf{x} \in \mathcal{X}_0, t \geq 0$. Moreover, ∇f^t is L_f -Lipschitz continuous: $\|\nabla f^t(\mathbf{x}) - \nabla f^t(\mathbf{y})\|_* \leq L_f \|\mathbf{x} - \mathbf{y}\|, \mathbf{x}, \mathbf{y} \in \mathcal{X}_0, t \geq 0$.
- The constraint function g_k is bounded: $\sum_{k=1}^K |g_k(\mathbf{x})| \leq G, \mathbf{x} \in \mathcal{X}_0$. Moreover, g_k is H_k -Lipschitz continuous: $|g_k(\mathbf{x}) - g_k(\mathbf{y})| \leq H_k \|\mathbf{x} - \mathbf{y}\|, \mathbf{x}, \mathbf{y} \in \mathcal{X}_0$. We let $H := \sum_{k=1}^K H_k$. Furthermore, the gradient of the function g_k is L_g -Lipschitz continuous: $\|\nabla g_k(\mathbf{x}) - \nabla g_k(\mathbf{y})\|_* \leq L_g \|\mathbf{x} - \mathbf{y}\|, \mathbf{x}, \mathbf{y} \in \mathcal{X}_0$.

The boundedness of the gradient ∇f^t and the function g_k can be induced by the Lipschitz continuity properties of them plus the compactness of \mathcal{X}_0 when they are well-defined in \mathcal{X}_0 . We explicitly assume the boundedness of ∇f^t and g_k in b) and c) with extra constants for clarity.

Assumption 2 (Slater's Condition). There exists a constant $\varsigma > 0$ and a point $\tilde{\mathbf{x}} \in \mathcal{X}_0$ such that $g_k(\tilde{\mathbf{x}}) \leq -\varsigma$ for any $k \in [K]$.

This is a standard assumption used in constrained optimization (Boyd, Boyd, and Vandenberghe 2004) and is also commonly adopted for constrained OCO problems (Yu and Neely 2017; Yu, Neely, and Wei 2017; Yu and Neely 2020). It intuitively implies that the set formed by functional constraints $\mathbf{g}_k(\mathbf{x}) \leq 0, k \in [K]$, has a non-empty interior.

We further define the Bregman divergence $D(\mathbf{x}, \mathbf{y})$ for any $\mathbf{x}, \mathbf{y} \in \mathcal{X}_0$. Let $\omega : \mathcal{X}_0 \mapsto \mathbb{R}$ be a strictly convex function, which is continuously differentiable in the interior of \mathcal{X}_0 . The Bregman divergence is defined as

$$D(\mathbf{x}, \mathbf{y}) = \omega(\mathbf{x}) - \omega(\mathbf{y}) - \langle \nabla \omega(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle.$$

We call ω a distance generating function with modulus ρ w.r.t. the norm $\|\cdot\|$ if it is ρ -strongly convex, that is, $\langle \mathbf{x} - \mathbf{y}, \nabla \omega(\mathbf{x}) - \nabla \omega(\mathbf{y}) \rangle \geq \rho \|\mathbf{x} - \mathbf{y}\|^2$. Then the Bregman divergence defined by the distance generating function ω satisfies $D(\mathbf{x}, \mathbf{y}) \geq \frac{\rho}{2} \|\mathbf{x} - \mathbf{y}\|^2$. Some common examples for the Bregman divergence are listed below.

¹We use k to index the constraints and t to index the time step.

Algorithm 1: Online Primal-Dual Mirror-Prox Algorithm

- 1: **Initialize:** $\gamma > 0; \mathbf{x}_0, \mathbf{x}_1, \tilde{\mathbf{x}}_1 \in \mathcal{X}_0; Q_k(0) = 0, k \in [K]$
 - 2: **for** $t = 1, \dots, T$ **do**
 - 3: Update the dual iterate $Q_k(t)$ via (6).
 - 4: Update the primal iterate \mathbf{x}_t via (7).
 - 5: Play \mathbf{x}_t .
 - 6: Suffer loss $f^t(\mathbf{x}_t)$ and compute $\nabla f^t(\mathbf{x}_t)$.
 - 7: Update the intermediate iterate $\tilde{\mathbf{x}}_{t+1}$ for the next round via (8).
 - 8: **end for**
-

- If $\omega(\mathbf{x}) = \frac{1}{2} \|\mathbf{x}\|_2^2$ then the Bregman divergence is related to the squared Euclidean distance on \mathbb{R}^d , $D(\mathbf{x}, \mathbf{y}) = \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2$ and $\rho = 1$.
- If $\omega(\mathbf{x}) = -\sum_{i=1}^d \mathbf{x}_i \log \mathbf{x}_i$ for any $\mathbf{x} \in \mathcal{X}_0$ with \mathcal{X}_0 being a probability simplex

$$\Delta := \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_1 = 1 \text{ and } \mathbf{x}_i \geq 0, \forall i \in [d]\}, \quad (5)$$

then $D(\mathbf{x}, \mathbf{y}) = D_{\text{KL}}(\mathbf{x}, \mathbf{y}) := \sum_{i=1}^d \mathbf{x}_i \log(\mathbf{x}_i/\mathbf{y}_i)$ is the Kullback-Leibler (KL) divergence, where $\mathbf{x} \in \Delta$ and $\mathbf{y} \in \Delta^\circ := \Delta \cap \text{relint}(\Delta)$ with $\text{relint}(\Delta)$ denoting the relative interior of Δ . In this case, ω has a modulus $\rho = 1$ w.r.t. ℓ_1 norm $\|\cdot\|_1$ with the dual norm $\|\cdot\|_\infty$.

Algorithm

We introduce the proposed online primal-dual mirror-prox algorithm in Algorithm 1. At the t -th round, we let $Q_k(t)$ be the dual iterate updated based on the k -th constraint function $g_k(x), k \in [K]$. The dual iterate is updated as

$$Q_k(t) = \max\{-\gamma g_k(\mathbf{x}_{t-1}), Q_k(t-1) + \gamma g_k(\mathbf{x}_{t-1})\}. \quad (6)$$

We denote $\mathbf{Q}(t) = [Q_1(t), \dots, Q_K(t)]^\top \in \mathbb{R}^K$ as a dual variable vector, which can be viewed as a Lagrange multiplier vector whose entries are guaranteed to be non-negative by induction since $Q_k(0) = 0$ by initialization. From another perspective, $\mathbf{Q}(t)$ is a virtual queue for backlogging constraint violations. The primal iterate \mathbf{x}_t is the decision made at each round by the decision maker. To obtain this iterate, we use an online mirror-prox-type updating rule with the constraint functions:

- Incorporating the dual iterates $Q_k(t)$ and the gradient of constraints $\nabla g_k(\mathbf{x}_{t-1})$, Line 4 runs a mirror descent step based on the last decision \mathbf{x}_{t-1} and an intermediate iterate $\tilde{\mathbf{x}}_t$ generated in the last round, which is

$$\mathbf{x}_t = \underset{\mathbf{x} \in \mathcal{X}_0}{\text{argmin}} \langle \nabla f^{t-1}(\mathbf{x}_{t-1}), \mathbf{x} \rangle + \sum_{k=1}^K [Q_k(t) + \gamma g_k(\mathbf{x}_{t-1})] \langle \gamma \nabla g_k(\mathbf{x}_{t-1}), \mathbf{x} \rangle + \alpha_t D(\mathbf{x}, \tilde{\mathbf{x}}_t). \quad (7)$$

- After observing a new loss function f^t , Line 7 generates an intermediate iterate $\tilde{\mathbf{x}}_{t+1}$ for the next round, which is

$$\tilde{\mathbf{x}}_{t+1} = \underset{\mathbf{x} \in \mathcal{X}_0}{\text{argmin}} \langle \nabla f^t(\mathbf{x}_t), \mathbf{x} \rangle + \sum_{k=1}^K [Q_k(t) + \gamma g_k(\mathbf{x}_{t-1})] \langle \gamma \nabla g_k(\mathbf{x}_{t-1}), \mathbf{x} \rangle + \alpha_t D(\mathbf{x}, \tilde{\mathbf{x}}_t). \quad (8)$$

The above proposed updating steps can yield a gradient-variation regret without sacrificing the constraint violation bound, as shown in the following sections. Note that the primal iterates (7) and (8) are simple and can be cast as two standard mirror descent steps that incorporate the dual iterate. More specifically, letting $\mathbf{h}_{t-1} := \nabla f^{t-1}(\mathbf{x}_{t-1}) + \gamma \sum_{k=1}^K (Q_k(t) + \gamma g_k(\mathbf{x}_{t-1})) \nabla g_k(\mathbf{x}_{t-1})$, one can show that Line 4 can be rewritten as

$$\begin{aligned} \nabla \omega(\mathbf{y}_t) &= \nabla \omega(\tilde{\mathbf{x}}_t) - \alpha_t^{-1} \mathbf{h}_{t-1}, \\ \mathbf{x}_t &= \operatorname{argmin}_{\mathbf{x} \in \mathcal{X}_0} D(\mathbf{x}, \mathbf{y}_t). \end{aligned}$$

When choosing the Bregman divergence as $D(\mathbf{x}, \mathbf{y}) = \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2$, it further reduces to gradient descent based update with the Euclidean projection onto \mathcal{X}_0 , that is,

$$\mathbf{x}_t = \operatorname{Proj}_{\mathcal{X}_0} \{\tilde{\mathbf{x}}_t - \alpha_t^{-1} \mathbf{h}_{t-1}\}.$$

Remark 1. *The primal update of Yu and Neely (2020) is a minimization problem that involves the exact constraint function $g_k(\mathbf{x})$, which can be any complex form. Therefore, its primal update cannot reduce to the projected gradient descent step if $g_k(\mathbf{x})$ is not linear, leading to a high computational cost. As opposed to Yu and Neely (2020), our proposed updates are simple as discussed above and only rely on a local linearization of the constraint function $g_k(\mathbf{x})$ by its gradient at \mathbf{x}_{t-1} , which is $g_k(\mathbf{x}_{t-1}) + \langle \nabla g_k(\mathbf{x}_{t-1}), \mathbf{x} - \mathbf{x}_{t-1} \rangle$ (the constant term $g_k(\mathbf{x}_{t-1}) - \langle \nabla g_k(\mathbf{x}_{t-1}), \mathbf{x}_{t-1} \rangle$ conditioned on \mathbf{x}_{t-1} does not affect updates and is ignored in Algorithm 1). Thus, Algorithm 1 enjoys the advantage of lower computational complexity. Moreover, our algorithm is designed for a more general normed space.*

Main Results

We present the regret bound (Theorem 1) and the constraint violation bound (Theorem 2) for Algorithm 1. We first make the following standard assumption on the boundedness of the Bregman divergence for Algorithm 1.

Assumption 3. *There exists $R > 0$ such that $D(\mathbf{x}, \mathbf{y}) \leq R^2, \forall \mathbf{x}, \mathbf{y} \in \mathcal{X}_0$.*

Assumption 3 is sensible if the Bregman divergence is a well-defined distance on a compact set \mathcal{X}_0 , for example, $\frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2$. Then Assumption 3 implies that $\|\mathbf{x} - \mathbf{y}\|^2 \leq 2R^2/\rho, \forall \mathbf{x}, \mathbf{y} \in \mathcal{X}_0$, according to the relation that $D(\mathbf{x}, \mathbf{y}) \geq \frac{\rho}{2} \|\mathbf{x} - \mathbf{y}\|_2^2$.

Note that Assumption 3 does not hold when $D(\mathbf{x}, \mathbf{y})$ is the KL divergence with $\mathcal{X}_0 = \Delta$ being a probability simplex. If \mathbf{y} is close to the boundary of the probability simplex, $D(\mathbf{x}, \mathbf{y})$ can be arbitrarily large according to the definition of KL divergence so that $D(\mathbf{x}, \mathbf{y})$ is not bounded. We study this probability simplex setting to complement our theory by proposing Algorithm 2 and presenting the corresponding analysis.

Theorem 1 (Regret). *Under Assumptions 1, 2, and 3, setting $\eta = [V_*(T) + L_f^2 + 1]^{-1/2}$, $\gamma = [V_*(T) + L_f^2 + 1]^{1/4}$, and $\alpha_t = \max \{2\rho^{-1}(\gamma^2 L_g G + \eta L_f^2 + \eta^{-1} + \xi_t), \alpha_{t-1}\}$ with initializing $\alpha_0 = 0$ and defining*

$$\xi_t := \gamma L_g \|\mathbf{Q}(t)\|_1 + \gamma^2 (L_g G + H^2), \quad (9)$$

then Algorithm 1 ensures the following regret

$$\begin{aligned} \operatorname{Regret}(T) &\leq \mathcal{O}\left((1 + \varsigma^{-1}) \sqrt{V_*(T) + L_f^2 + 1} + \varsigma^{-1}\right) \\ &= \mathcal{O}\left(\sqrt{V_*(T)} \vee L_f\right), \end{aligned}$$

where \mathcal{O} hides constants $\operatorname{poly}(R, H, F, G, L_g, K, \rho)$.²

The setting of α_t guarantees that $\{\alpha_t\}_{t \geq 0}$ is a non-decreasing sequence such that $\alpha_{t+1} \geq \alpha_t$. Note that this setting is sensible since it in fact implies a non-increasing step size α_t^{-1} . For a clear understanding, consider a simple example: if $D(\mathbf{x}, \mathbf{y}) = \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2$ and all $g_k(\mathbf{x}) \equiv 0, \mathbf{x} \in \mathcal{X}_0$ such that we have an ordinary constrained online optimization problem, then (7) becomes $\mathbf{x}_t = \operatorname{Proj}_{\mathcal{X}_0} [\tilde{\mathbf{x}}_t - \alpha_t^{-1} \nabla f^{t-1}(\mathbf{x}_{t-1})]$ with α_t^{-1} being a non-increasing step size.

Theorem 2 (Constraint Violation). *Under Assumptions 1, 2, and 3, with the same settings of η, γ , and α_t as in Theorem 1, Algorithm 1 ensures the following constraint violation*

$$\operatorname{Violation}(T, k) \leq \mathcal{O}(1 + \varsigma^{-1}) = \mathcal{O}(1), \quad \forall k \in [K],$$

where \mathcal{O} hides the constant factor $\operatorname{poly}(R, H, F, G, L_f, L_g, K, \rho)$.

Theorem 1 and Theorem 2 can be interpreted as follows:

1. Regret is bounded by $\mathcal{O}(\sqrt{V_*(T)} \vee L_f)$, which can explicitly reveal the dependence of the regret on the gradient variation $V_*(T)$. Since $V_*(T) = \sum_{t=1}^T \max_{\mathbf{x} \in \mathcal{X}_0} \|\nabla f^t(\mathbf{x}) - \nabla f^{t-1}(\mathbf{x})\|_*^2 \leq 2FT$, then $\mathcal{O}(\sqrt{V_*(T)})$ reduces to $\mathcal{O}(\sqrt{T})$ in the worst case. This result indicates that we can achieve the gradient-variation bound for the constrained online convex optimization via a first-order method. Meanwhile, our constraint violation bound remains $\mathcal{O}(1)$ as in Yu and Neely (2020).
2. When $f^1 = f^2 = \dots = f^T$ such that $V_*(T) = 0$, we have $\mathcal{O}(1)$ regret and constraint violation, which is equivalent to the $\mathcal{O}(1/T)$ convergence rate (measured by $\operatorname{Regret}(T)/T$) of solving a smooth convex optimization, matching the result for mirror-prox algorithms (Bubeck et al. 2015). Our result also improves upon previous attempts using mirror-prox-type algorithms for OCO with complex constraints, which achieved a worse $\mathcal{O}(T^{2/3})$ regret and constraint violation (Mahdavi, Jin, and Yang 2012).
3. Moreover, our theorems hold in the general normed space, which covers Euclidean space as a special case. Compared to Wei, Yu, and Neely (2020) also for OCO in the general normed space but with stochastic constraints, our gradient-variation regret reduces to their $\mathcal{O}(\sqrt{T})$ regret in the worst case and our constraint violation bound $\mathcal{O}(1)$ improves over their $\mathcal{O}(\sqrt{T})$ result. The improvement in constraint violation results from exploiting the long-term fixed constraints in our setting rather than their stochastic constraints. Our work bridges this theoretical gap for the constrained OCO in the general space.

²Hereafter, we use $\operatorname{poly}(\dots)$ to denote a polynomial term composed of the variables inside the parentheses.

Remark 2. We set the hyperparameters by using the gradient-variation $V_*(T)$ following the existing work for the gradient-variation regret (Chiang et al. 2012; Yang et al. 2014). One potential research direction is to design an adaptive algorithm without using $V_*(T)$ for setting hyperparameters. We leave the problem of designing such adaptive algorithms for OCO as our future work.

Theoretical Analysis

Our analysis starts from the *drift-plus-penalty* expression with the drift-plus-penalty term defined as

$$\begin{aligned} \text{DPP}(t) := & \underbrace{\frac{1}{2} [\|\mathbf{Q}(t+1)\|_2^2 - \|\mathbf{Q}(t)\|_2^2]}_{\text{drift}} \\ & + \underbrace{\langle \nabla f^{t-1}(\mathbf{x}_{t-1}), \mathbf{x} \rangle + \alpha_t D(\mathbf{x}, \tilde{\mathbf{x}}_t)}_{\text{penalty}}. \end{aligned} \quad (10)$$

The drift term shows the one-step change of the vector $\mathbf{Q}(t)$ which is the backlog queue of the constraint functions. The penalty term is associated with a mirror descent step when observing the gradient of the loss function f^{t-1} . The drift-plus-penalty expression is investigated in recent papers on constrained online optimization problems (Yu and Neely 2017; Yu, Neely, and Wei 2017; Wei, Yu, and Neely 2020; Yu and Neely 2020). However, the techniques in our analysis for this expression are different. Our theoretical analysis makes a step toward understanding the drift-plus-penalty expression under mirror-prox-type algorithms. We develop a novel upper bound of the drift-plus-penalty term for the constraint online optimization problem at the point $\mathbf{x} = \mathbf{x}_t$ in the following lemma.

Lemma 1. *At the t -th round of Algorithm 1, for any $\gamma > 0$ and any $\mathbf{z} \in \mathcal{X}_0$, letting ξ_t as in (9), the drift-plus-penalty term admits the following bound*

$$\begin{aligned} \text{DPP}(t) \leq & \frac{\xi_t}{2} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|^2 + \frac{\gamma^2}{2} [\|\mathbf{g}(\mathbf{x}_t)\|_2^2 - \|\mathbf{g}(\mathbf{x}_{t-1})\|_2^2] \\ & + \alpha_t D(\mathbf{z}, \tilde{\mathbf{x}}_t) - \alpha_t D(\mathbf{z}, \tilde{\mathbf{x}}_{t+1}) - \alpha_t D(\tilde{\mathbf{x}}_{t+1}, \mathbf{x}_t) \\ & + \langle \nabla f^{t-1}(\mathbf{x}_{t-1}) - \nabla f^t(\mathbf{x}_t), \tilde{\mathbf{x}}_{t+1} \rangle + \langle \nabla f^t(\mathbf{x}_t), \mathbf{z} \rangle \\ & + \gamma \langle \mathbf{Q}(t) + \gamma \mathbf{g}(\mathbf{x}_{t-1}), \mathbf{g}(\mathbf{z}) \rangle. \end{aligned}$$

See Appendix A.2 for the proof. We use this lemma to obtain Lemma 2, Lemma 3, and Lemma 4.

Lemma 1 is proved by utilizing the dual update in (6) and the two mirror descent steps coupled with the dual iterates and the constraints in (7) and (8). As shown in Lemma 1, the upper bound of $\text{DPP}(t)$ contains the one-step gradient-variation term $\nabla f^{t-1}(\mathbf{x}_{t-1}) - \nabla f^t(\mathbf{x}_t)$, which will be further used to build an important connection with the gradient variation $V_*(T)$. It indicates that our proposed algorithm can reveal the gradient variation in the drift-plus-penalty term in contrast to the prior work where this term does not exist. Moreover, it also has the difference terms for $\|\mathbf{g}(\mathbf{x}_t)\|_2^2$ and $D(\mathbf{z}, \tilde{\mathbf{x}}_t)$, with which we will construct a telescoping summation later. The last term $\gamma \langle \mathbf{Q}(t) + \gamma \mathbf{g}(\mathbf{x}_{t-1}), \mathbf{g}(\mathbf{z}) \rangle$ together with Slater's condition and the dual updating rule is further utilized to bound the dual variable drift $\|\mathbf{Q}(T+1)\|_2$ and regret $\text{Regret}(T)$ when setting \mathbf{z} to be $\tilde{\mathbf{x}}$ or \mathbf{x}^* in Lemma 3 and Lemma 4.

Proof Sketches

We give proof sketches of the main theorems. Within this subsection, all the lemmas and the proofs are under Assumptions 1, 2, and 3.

Lemma 2. *At the t -th round of Algorithm 1, for any $\eta, \gamma > 0$ and any $\mathbf{z} \in \mathcal{X}_0$, setting α_t as in Theorem 1, the following inequality holds*

$$\begin{aligned} & \frac{1}{2} [\|\mathbf{Q}(t+1)\|_2^2 - \|\mathbf{Q}(t)\|_2^2] + \langle \nabla f^t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{z} \rangle \\ & \leq U_t - U_{t+1} + \frac{\eta}{2} \|\nabla f^{t-1}(\mathbf{x}_t) - \nabla f^t(\mathbf{x}_t)\|_*^2 \\ & \quad + (\alpha_{t+1} - \alpha_t) D(\mathbf{z}, \tilde{\mathbf{x}}_{t+1}) + \gamma \langle \mathbf{Q}(t) + \gamma \mathbf{g}(\mathbf{x}_{t-1}), \mathbf{g}(\mathbf{z}) \rangle, \end{aligned} \quad (11)$$

where we define the term $U_t := (\xi_t + \eta L_f^2) \|\mathbf{x}_{t-1} - \tilde{\mathbf{x}}_t\|^2 + \alpha_t D(\mathbf{z}, \tilde{\mathbf{x}}_t) - \gamma^2/2 \cdot \|\mathbf{g}(\mathbf{x}_{t-1})\|_2^2$.

Please see Appendix A.3 for the proof. Lemma 2 is obtained by rearranging the upper bound of $\text{DPP}(t)$ in Lemma 1 and properly setting the step size α_t such that the redundant terms, for example, $\|\mathbf{x}_t - \mathbf{x}_{t-1}\|^2$, are eliminated. Moreover, we now explicitly express the one-step gradient variation $\|\nabla f^{t-1}(\mathbf{x}_t) - \nabla f^t(\mathbf{x}_t)\|_*^2$ in the upper bound. The difference term for U_t can lead to a telescoping summation when taking summation over T slots on both sides of (11). We can also use Lemma 2 to derive the drift bound of $\|\mathbf{Q}(t)\|_2$ in Lemma 3 and the regret bound in Lemma 4 since (11) contains the difference term for $\|\mathbf{Q}(t)\|_2$ and also $\langle \nabla f^t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{z} \rangle$. Letting $\mathbf{z} = \mathbf{x}^*$, we have a relationship between this lemma and the regret as $\text{Regret}(T) = \sum_{t=1}^T f^t(\mathbf{x}_t) - \sum_{t=1}^T f^t(\mathbf{x}^*) \leq \sum_{t=1}^T \langle \nabla f^t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle$.

Based on Lemma 2, we obtain the following lemma.

Lemma 3. *Setting η, γ , and α_t as in Theorem 1, Algorithm 1 ensures*

$$\begin{aligned} \alpha_{T+1} & \leq \mathcal{O}((1 + \varsigma^{-1})[V_*(T) + L_f^2 + 1]^{1/2} + \varsigma^{-1}), \\ \|\mathbf{Q}(T+1)\|_2 & \leq \mathcal{O}((1 + \varsigma^{-1})[V_*(T) + L_f^2 + 1]^{1/4} + \varsigma^{-1}), \end{aligned}$$

where \mathcal{O} hides the constants $\text{poly}(R, H, F, G, L_g, K, \rho)$.

For this lemma, we develop a novel proof for the drift of the dual variable. Specifically, based on the relation in Lemma 2, the standard Slater's condition, and the dual updating rule, we derive an upper bound of $\|\mathbf{Q}(t)\|_2$ in terms of step size α_{T+1} by the proof of contradiction. Since the upper bound of α_{T+1} is also unknown, further with the step size setting α_t which is coupled with $\|\mathbf{Q}(t)\|_2$, we solve the upper bounds of α_{T+1} and $\|\mathbf{Q}(T+1)\|_2$ together. Our proof also depends on the elaborate setting of α_t and γ such that the upper bound has a favorable dependence on $V_*(T)$. Please see Appendix A.4 for a detailed proof of Lemma 3. This lemma shows that $\alpha_{T+1} \leq \mathcal{O}(\sqrt{V_*(T)})$ and $\|\mathbf{Q}(T+1)\|_2 \leq \mathcal{O}(V_*(T)^{1/4})$ after $T+1$ rounds of Algorithm 1, which is the key to obtaining the gradient variation regret bound and maintaining the $\mathcal{O}(1)$ constraint violation. Moreover, the bounds in Lemma 3 have a dependence on $1/\varsigma$, which reveals how Slater's condition will affect our regret and constraint violation bounds.

With Lemma 2, we obtain the following upper bound for the regret.

Lemma 4. For any $\eta, \gamma \geq 0$, setting α_t the same as in Theorem 1, Algorithm 1 ensures

$$\text{Regret}(T) \leq \mathcal{O}(\eta V_*(T) + L_f^2 \eta + \gamma^2 + \alpha_{T+1}),$$

where \mathcal{O} hides absolute constants $\text{poly}(R, H, G, L_g, \rho)$.

Letting $\mathbf{z} = \mathbf{x}^*$ and taking summation on both sides of (11), we can further prove Lemma 4 recalling that $\text{Regret}(T) = \sum_{t=1}^T f^t(\mathbf{x}_t) - \sum_{t=1}^T f^t(\mathbf{x}^*) \leq \sum_{t=1}^T \langle \nabla f^t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle$. Please see Appendix A.5 for the detailed proof. By the update rule of the dual variable, we have the following lemma for the constraint violation.

Lemma 5. For any $\eta > 0$, the updating rule of $Q_k(t)$ in Algorithm 1 ensures

$$\text{Violation}(T, k) \leq \gamma^{-1} \|\mathbf{Q}(T+1)\|_2.$$

Please see Appendix A.6 for a detailed proof. This lemma indicates that the upper bound of the constraint violation is associated with the dual variable drift.

Proof of Theorem 1 and Theorem 2 According to Lemma 4, by the settings of η and γ in Theorem 1, we have $\text{Regret}(T) \leq \mathcal{O}(\sqrt{V_*(T) + L_f^2 + 1} + \alpha_{T+1})$, which is due to $\eta V_*(T) + L_f^2 \eta \leq [V_*(T) + L_f^2] \eta \leq \sqrt{V_*(T) + L_f^2 + 1}$. Furthermore, combining the above inequality with the bound of α_{T+1} in Lemma 3 yields

$$\text{Regret}(T) \leq \mathcal{O}\left((1 + \varsigma^{-1}) \sqrt{V_*(T) + L_f^2 + 1} + \varsigma^{-1}\right).$$

According to Lemma 5 and the drift bound of $\mathbf{Q}(T+1)$ in Lemma 3, with the setting of γ , we have

$$\text{Violation}(T, k) \leq \gamma^{-1} \|\mathbf{Q}(T+1)\|_2 \leq \mathcal{O}(1 + \varsigma^{-1}),$$

where the second inequality follows from $1/(\varsigma\gamma) \leq 1/\varsigma$ since $1/\gamma \leq 1$. This completes the proof.

Extension to the Probability Simplex Case

In the probability simplex case, we have $\mathcal{X}_0 = \Delta$ where Δ denotes the probability simplex as in (5) and the Bregman divergence is the KL divergence, namely $D(\mathbf{x}, \mathbf{y}) = D_{\text{KL}}(\mathbf{x}, \mathbf{y})$. Thus, the norm $\|\cdot\|$ defined in this space is ℓ_1 norm $\|\cdot\|_1$ with the dual norm $\|\cdot\|_* = \|\cdot\|_\infty$ such that the gradient variation is measured by

$$V_\infty(T) = \sum_{t=1}^T \max_{\mathbf{x} \in \mathcal{X}_0} \|\nabla f^t(\mathbf{x}) - \nabla f^{t-1}(\mathbf{x})\|_\infty^2.$$

Then, the results in this section are expressed in terms of $V_\infty(T)$. Note that Assumption 3 is no longer valid, since $D_{\text{KL}}(\mathbf{x}, \mathbf{y})$ can tend to infinity by its definition if there is some entry $y_i \rightarrow 0$. Thus, we propose Algorithm 2 for the probability simplex case.

To tackle the challenge of unbounded KL divergence, we propose to mix iterate $\tilde{\mathbf{x}}_t$ with a vector $\mathbf{1}/d$ where $\mathbf{1} \in \mathbb{R}^d$ is an all-one vector, as shown in Line 3 of Algorithm 2. Intuitively, the mixing step is to push the iterates $\tilde{\mathbf{x}}_t$ slightly away from the boundary of Δ in a controllable way with a

Algorithm 2: Online Primal-Dual Mirror-Prox Algorithm on Probability Simplex

-
- 1: **Initialize:** $\gamma > 0; \nu \in (0, 1]; \mathbf{x}_0 = \mathbf{x}_1 = \tilde{\mathbf{x}}_1 = \mathbf{1}/d;$
 $Q_k(0) = 0, k \in [K]$
 - 2: **for** $t = 1, \dots, T$ **do**
 - 3: Mix the iterates: $\tilde{\mathbf{y}}_t = (1 - \nu)\tilde{\mathbf{x}}_t + \frac{\nu}{d}\mathbf{1}.$
 - 4: Update the dual iterate $Q_k(t)$ via (12).
 - 5: Update the primal iterate \mathbf{x}_t via (13).
 - 6: Play \mathbf{x}_t .
 - 7: Suffer loss $f^t(\mathbf{x}_t)$ and compute $\nabla f^t(\mathbf{x}_t)$.
 - 8: Update the intermediate iterate $\tilde{\mathbf{x}}_{t+1}$ for the next round via (14).
 - 9: **end for**
-

weight ν such that the KL divergence $D_{\text{KL}}(\mathbf{x}, \tilde{\mathbf{y}}_t)$ for any $\mathbf{x} \in \Delta$ will not be too large. Specifically, according to our theory, we set a suitable mixing weight as $\nu = 1/T$.

For the dual iterate $Q_k(t)$ in Algorithm 2, we have the same updating rule as in Algorithm 1, which is

$$Q_k(t) = \max\{-\gamma g_k(\mathbf{x}_{t-1}), Q_k(t-1) + \gamma g_k(\mathbf{x}_{t-1})\}. \quad (12)$$

For the primal iterate \mathbf{x}_t , the updating rule is now based on the new mixed iterate $\tilde{\mathbf{y}}_t$ and the probability simplex Δ , which is written as

$$\begin{aligned} \mathbf{x}_t = \underset{\mathbf{x} \in \Delta}{\text{argmin}} \quad & \langle \nabla f^{t-1}(\mathbf{x}_{t-1}), \mathbf{x} \rangle + \sum_{k=1}^K [Q_k(t) \\ & + \gamma g_k(\mathbf{x}_{t-1})] \langle \gamma \nabla g_k(\mathbf{x}_{t-1}), \mathbf{x} \rangle + \alpha_t D(\mathbf{x}, \tilde{\mathbf{y}}_t). \end{aligned} \quad (13)$$

The intermediate iterate is also updated with $\tilde{\mathbf{y}}_t$ as

$$\begin{aligned} \tilde{\mathbf{x}}_{t+1} = \underset{\mathbf{x} \in \Delta}{\text{argmin}} \quad & \langle \nabla f^t(\mathbf{x}_t), \mathbf{x} \rangle + \sum_{k=1}^K [Q_k(t) \\ & + \gamma g_k(\mathbf{x}_{t-1})] \langle \gamma \nabla g_k(\mathbf{x}_{t-1}), \mathbf{x} \rangle + \alpha_t D(\mathbf{x}, \tilde{\mathbf{y}}_t). \end{aligned} \quad (14)$$

Therefore, the updates of Algorithm 2 lead to the following regret upper bound and the constraint violation bound.

Theorem 3 (Regret). Under Assumptions 1 and 2, setting $\eta = [V_\infty(T) + L_f^2 + 1]^{-1/2}$, $\gamma = [V_\infty(T) + L_f^2 + 1]^{1/4}$, $\nu = 1/T$, and $\alpha_t = \max\{3(\eta L_f^2 + \gamma^2 L_g G) + 2/\eta + 3\xi_t, \alpha_{t-1}\}$ with $\alpha_0 = 0$, for $T > 2$ and $d \geq 1$, Algorithm 1 ensures the following regret

$$\text{Regret}(T) \leq \tilde{\mathcal{O}}(\sqrt{V_\infty(T)} \vee L_f),$$

where $\tilde{\mathcal{O}}$ hides constants $\text{poly}(H, F, G, L_g, K, 1/\varsigma)$ and the logarithmic factor $\log^2(Td)$.

Theorem 4 (Constraint Violation). Under Assumptions 1 and 2, with the same settings of η , γ , ν , and α_t as Theorem 1, Algorithm 1 ensures the following constraint violation for all $k \in [K]$

$$\text{Violation}(T, k) \leq \mathcal{O}(\log T) = \tilde{\mathcal{O}}(1),$$

where \mathcal{O} hides $\text{poly}(H, F, G, L_f, L_g, K, 1/\varsigma)$ and $\log d$, and $\tilde{\mathcal{O}}$ hides the logarithmic dependence on T .

The results of Theorems 3 and 4 show that Algorithm 2 only introduces an extra logarithmic factor $\log T$ in both regret and constraint violation bounds. The extra $\log T$ is incurred by the iterate mixing step in Line 3 of Algorithm 2, which guarantees that the iterates stay away from the boundary. We provide a novel drift-plus-penalty analysis for this algorithm which incorporates the $\log T$ factor with the gradient-variation bound. See Appendix B for a detailed proof. Compared to Wei, Yu, and Neely (2020) for the probability simplex setting, our gradient-variation regret reduces to their $\tilde{\mathcal{O}}(\sqrt{T})$ regret in the worst case, and our constraint violation bound $\tilde{\mathcal{O}}(1)$ improves over their $\tilde{\mathcal{O}}(\sqrt{T})$ result if applying their method to the setting of fixed constraints.

Conclusion

In this paper, we propose novel first-order methods for constrained OCO problems, which can achieve a gradient-variation bound $\mathcal{O}(\max\{\sqrt{V_*(T)}, L_f\})$ for the regret and $\mathcal{O}(1)$ bound for the constraint violation simultaneously in a general normed space $(\mathcal{X}_0, \|\cdot\|)$. In particular, our bound is never worse than the best-known $(\mathcal{O}(\sqrt{T}), \mathcal{O}(1))$ bound under the Slater's condition and can be much better when the variation $\sqrt{V_*(T)}$ is small.

References

- Baby, D.; and Wang, Y.-X. 2021. Optimal dynamic regret in exp-concave online learning. In *Conference on Learning Theory*, 359–409. PMLR.
- Baby, D.; and Wang, Y.-X. 2022. Optimal dynamic regret in proper online learning with strongly convex losses and beyond. In *International Conference on Artificial Intelligence and Statistics*, 1805–1845. PMLR.
- Baby, D.; Zhao, X.; and Wang, Y.-X. 2021. An optimal reduction of tv-denoising to adaptive online learning. In *International Conference on Artificial Intelligence and Statistics*, 2899–2907. PMLR.
- Besbes, O.; Gur, Y.; and Zeevi, A. 2015. Non-stationary stochastic optimization. *Operations research*, 63(5): 1227–1244.
- Boyd, S.; Boyd, S. P.; and Vandenberghe, L. 2004. *Convex optimization*. Cambridge university press.
- Bubeck, S.; et al. 2015. Convex optimization: Algorithms and complexity. *Foundations and Trends® in Machine Learning*, 8(3-4): 231–357.
- Cesa-Bianchi, N.; Long, P. M.; and Warmuth, M. K. 1996. Worst-case quadratic loss bounds for prediction using linear functions and gradient descent. *IEEE Transactions on Neural Networks*, 7(3): 604–619.
- Chang, T.-J.; and Shahrampour, S. 2021. On online optimization: Dynamic regret analysis of strongly convex and smooth problems. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, 6966–6973.
- Chen, T.; Ling, Q.; and Giannakis, G. B. 2017. An online convex optimization approach to proactive network resource allocation. *IEEE Transactions on Signal Processing*, 65(24): 6350–6364.
- Chen, X.; Wang, Y.; and Wang, Y.-X. 2019. Nonstationary stochastic optimization under L p, q-variation measures. *Operations Research*, 67(6): 1752–1765.
- Cheng, C.-A.; Lee, J.; Goldberg, K.; and Boots, B. 2020. Online learning with continuous variations: Dynamic regret and reductions. In *International Conference on Artificial Intelligence and Statistics*, 2218–2228. PMLR.
- Chiang, C.-K.; Yang, T.; Lee, C.-J.; Mahdavi, M.; Lu, C.-J.; Jin, R.; and Zhu, S. 2012. Online optimization with gradual variations. In *Conference on Learning Theory*, 6–1.
- Daniely, A.; Gonen, A.; and Shalev-Shwartz, S. 2015. Strongly adaptive online learning. In *International Conference on Machine Learning*, 1405–1411. PMLR.
- Goel, G.; and Wierman, A. 2019. An online algorithm for smoothed regression and lqr control. In *The 22nd International Conference on Artificial Intelligence and Statistics*, 2504–2513. PMLR.
- Gordon, G. J. 1999. Regret bounds for prediction problems. In *Proceeding of Conference on Learning Theory (COLT)*.
- Hall, E.; and Willett, R. 2013. Dynamical models and tracking regret in online convex programming. In *International Conference on Machine Learning*, 579–587. PMLR.
- Hazan, E. 2016a. Introduction to online convex optimization. *Foundations and Trends in Optimization*, 2(3–4): 157–325.
- Hazan, E. 2016b. Introduction to online convex optimization. *Foundations and Trends® in Optimization*, 2(3-4): 157–325.
- Hazan, E.; and Kale, S. 2010. Extracting certainty from uncertainty: Regret bounded by variation in costs. *Machine learning*, 80(2): 165–188.
- Hazan, E.; and Seshadhri, C. 2007. Adaptive algorithms for online decision problems. In *Electronic colloquium on computational complexity (ECCC)*, volume 14.
- Jadbabaie, A.; Rakhlin, A.; Shahrampour, S.; and Sridharan, K. 2015. Online optimization: Competing with dynamic comparators. In *Artificial Intelligence and Statistics*, 398–406. PMLR.
- Janatton, R.; Huang, J.; and Archambeau, C. 2016. Adaptive algorithms for online convex optimization with long-term constraints. In *International Conference on Machine Learning*, 402–411. PMLR.
- Liakopoulos, N.; Destounis, A.; Paschos, G.; Spyropoulos, T.; and Mertikopoulos, P. 2019. Cautious regret minimization: Online optimization with long-term budget constraints. In *International Conference on Machine Learning*, 3944–3952.
- Mahdavi, M.; Jin, R.; and Yang, T. 2012. Trading regret for efficiency: online convex optimization with long term constraints. *Journal of Machine Learning Research*, 13(Sep): 2503–2528.
- Mokhtari, A.; Shahrampour, S.; Jadbabaie, A.; and Ribeiro, A. 2016. Online optimization in dynamic environments: Improved regret rates for strongly convex problems. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, 7195–7201. IEEE.

- Nemirovski, A. 2004. Prox-method with rate of convergence $O(1/t)$ for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems. *SIAM Journal on Optimization*, 15(1): 229–251.
- Steinhardt, J.; and Liang, P. 2014. Adaptivity and optimism: An improved exponentiated gradient algorithm. In *International Conference on Machine Learning*, 1593–1601.
- Wei, X.; Yu, H.; and Neely, M. J. 2020. Online primal-dual mirror descent under stochastic constraints. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 4(2): 1–36.
- Yang, T.; Mahdavi, M.; Jin, R.; and Zhu, S. 2014. Regret bounded by gradual variation for online convex optimization. *Machine learning*, 95(2): 183–223.
- Yang, T.; Zhang, L.; Jin, R.; and Yi, J. 2016. Tracking slowly moving clairvoyant: Optimal dynamic regret of online learning with true and noisy gradient. In *International Conference on Machine Learning*, 449–457. PMLR.
- Yi, X.; Li, X.; Yang, T.; Xie, L.; Chai, T.; and Johansson, K. 2021. Regret and cumulative constraint violation analysis for online convex optimization with long term constraints. In *International Conference on Machine Learning*, 11998–12008. PMLR.
- Yu, H.; Neely, M.; and Wei, X. 2017. Online convex optimization with stochastic constraints. In *Advances in Neural Information Processing Systems*, 1428–1438.
- Yu, H.; and Neely, M. J. 2017. A Simple Parallel Algorithm with an $O(1/t)$ Convergence Rate for General Convex Programs. *SIAM Journal on Optimization*, 27(2): 759–783.
- Yu, H.; and Neely, M. J. 2020. A Low Complexity Algorithm with $O(\sqrt{T})$ Regret and $O(1)$ Constraint Violations for Online Convex Optimization with Long Term Constraints. *Journal of Machine Learning Research*, 21(1): 1–24.
- Yuan, J.; and Lamperski, A. 2018. Online convex optimization for cumulative constraints. In *Advances in Neural Information Processing Systems*, 6137–6146.
- Zhang, L.; Lu, S.; and Zhou, Z.-H. 2018. Adaptive online learning in dynamic environments. *Advances in neural information processing systems*, 31.
- Zhang, L.; Yang, T.; Yi, J.; Jin, R.; and Zhou, Z.-H. 2017. Improved dynamic regret for non-degenerate functions. *Advances in Neural Information Processing Systems*, 30.
- Zhang, L.; Yang, T.; Zhou, Z.-H.; et al. 2018. Dynamic regret of strongly adaptive methods. In *International conference on machine learning*, 5882–5891. PMLR.
- Zhao, P.; Wang, Y.-X.; and Zhou, Z.-H. 2022. Non-stationary online learning with memory and non-stochastic control. In *International Conference on Artificial Intelligence and Statistics*, 2101–2133. PMLR.
- Zhao, P.; and Zhang, L. 2021. Improved analysis for dynamic regret of strongly convex and smooth functions. In *Learning for Dynamics and Control*, 48–59. PMLR.
- Zhao, P.; Zhang, Y.-J.; Zhang, L.; and Zhou, Z.-H. 2020. Dynamic regret of convex and smooth functions. *Advances in Neural Information Processing Systems*, 33: 12510–12520.
- Zhao, Y.; Qiu, S.; Li, K.; Luo, L.; Yin, J.; and Liu, J. 2021. Proximal online gradient is optimum for dynamic regret: A general lower bound. *IEEE Transactions on Neural Networks and Learning Systems*.
- Zinkevich, M. 2003a. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of International Conference on Machine Learning (ICML)*.
- Zinkevich, M. 2003b. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th international conference on machine learning (icml-03)*, 928–936.