Online Convex Optimization with Perturbed Constraints: Optimal Rates against Stronger Benchmarks

Víctor Valls
Yale University

George Iosifidis
Trinity College Dublin

Douglas J. Leith
Trinity College Dublin

Leandros Tassiulas
Yale University

Abstract

This paper studies Online Convex Optimization (OCO) problems where the constraints have additive perturbations that (i) vary over time and (ii) are not known at the time to make a decision. Perturbations may not be i.i.d. generated and can be used, for example, to model a time-varying budget or time-varying requests in resource allocation problems. Our goal is to design a policy that obtains sublinear regret and satisfies the constraints in the long-term. To this end, we present an online primal-dual proximal gradient algorithm that has $O(T^\epsilon \vee T^{1-\epsilon})$ regret and $O(T^\epsilon)$ constraint violation, where $\epsilon \in [0, 1)$ is a parameter in the learning rate. The proposed algorithm obtains optimal rates when $\epsilon = 1/2$, and can compare against a stronger comparator (the set of fixed decisions in hindsight) than previous work.

1 Introduction

The Online Convex Optimization (OCO) framework was introduced in (Zinkevich, 2003), and it is widely used to model applications such as spam filtering, portfolio selection, recommendation systems, among many others (Hazan, 2016). In short, OCO consists of a sequence of games where in each round $t \in \mathbb{N}$ an agent selects an action x_t from a convex set $X \subset \mathbb{R}^n$ and experiences a cost $f_t(x_t)$, where $f_t : \mathbb{R}^n \to \mathbb{R}$ is convex. Crucially, the cost function is not known at the time of making a decision, and it may even be selected by an adversary after the action has been played. The goal is to design a policy or algorithm that selects a sequence of actions $\{x_t\}$, $t = 1, \dots, T$ from X so that the regret

$$R_T := \sum_{t=1}^{T} f_t(x_t) - \min_{x \in X} \sum_{t=1}^{T} f_t(x) \qquad (1)$$

increases sublinearly i.e., $\limsup_{T\to\infty} R_T/T \le 0$. Hence, the incurred cost is asymptotically as good as the best fixed decision in hindsight.[1]

As it was shown in (Zinkevich, 2003), the online gradient descent (OGD) algorithm can obtain sublinear regret when the action set X is bounded and the convex cost functions f_t, $t \in \mathbb{N}$ have bounded subgradients. The algorithm consists of the update

$$x_{t+1} = \mathcal{P}_X(x_t - \alpha_t f_t'(x_t)), \qquad t = 1, 2, \dots \qquad (2)$$

where $\alpha_t = 1/\sqrt{t}$ is the learning rate (or, step size), $f_t'(x_t)$ a subgradient of the previous cost function at x_t, and $\mathcal{P}_X$ the Euclidean projection onto the convex set X. Note from Eq. (2) that x_{t+1} is selected using only information available at time t.

1.1 OCO with (online) long-term constraints

In the standard OCO framework, all the constraints must be satisfied in every iteration. However, in many important problems, it is only necessary to satisfy the constraints on average. For instance, in wireless communications a transceiver allocates power to maximize the probability that a message is received successfully (the power needed is not known a priori as it depends on the channel conditions, the behavior of other users, etc.) subject to average power transmission constraints. That is, the transceiver can occasionally exceed the average threshold, as long as its power budget is satisfied in the long run.

An extension of the above problem is the case of *online constraints*: constraints that vary over time and are not

[1] The regret captures the difference between the incurred cost and the cost obtained by an "offline" algorithm that has knowledge of all the cost functions from $t = 1, \dots, T$. The offline algorithm, however, can only choose one vector from X.

Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics (AISTATS) 2020, Palermo, Italy. PMLR: Volume 108.

known to the decision maker a priori. More specifically, in every time slot $t = 1, 2, \dots, T$ the decision maker (i) selects an action x_t; (ii) experiences a penalty $g_t(x_t)$; and (iii) learns the constraints $g_t : \mathbb{R}^m \to \mathbb{R}$. For these types of constraints, the goal is that the accumulated penalty (i.e., $\sum_{t=1}^T g_t(x_t)$) increases sublinearly in addition to having sublinear regret. Two motivating examples of OCO with online constraints are the following:

Online shortest path routing. Online shortest path routing with static constraints is a canonical example in OCO (Hazan, 2016, pp. 7). In brief, time is divided in slots $t \in \mathbb{N}$, and in each time slot an agent selects the amount of flow x_t to route over each link/path. The goal is to minimize the cost $f_t(x_t)$ and to satisfy the inequality constraint $Ax_t \preceq b$,[2] where $A \in \mathbb{R}^{m\times n}$ is a routing matrix and $b \in \mathbb{R}^m$ a request vector that indicates the supply/demand of flow (of material, traffic, information, etc.) at each of the nodes. The cost function $f_t(x_t)$ usually captures the latency from the source node to the sink. In the setting with online constraints, b is replaced by b_t, and we select x_t without prior knowledge on b_t. Hence, it is not possible to design an algorithm that guarantees that Ax_t is less than or equal to b_t in every iteration. This setting appears in practical queueing systems where resources (x_t) must be allocated before the real demand (b_t) can be observed (Georgiadis et al., 2006).[3]

Online advertising. Another example is online advertising with budget constraints (Liakopoulos et al., 2019). In this case, $x_t \in \mathbb{R}^n_+$ represents a vector of bids across n different websites, i.e., the price we are willing to pay per click. Each website decides where to place the ads of the different bidders, which affects the total number of clicks p_t. The goal is to maximize the total reward $\sum_{t=1}^T -f_t(x_t)$ and to keep the advertisement cost $\sum_{t=1}^T \langle p_t, x_t\rangle$ below the available budget c. The budget can be given, for example, in a daily, weekly, or monthly basis. Importantly, the problem is an online optimization because it is not possible to predict the number of clicks p_t, which affect the reward and constraint penalties. Furthermore, the number of clicks may be affected by an adversary (e.g., a bot) that aims to increasing our advertising cost.

In sum, in OCO with long-term constraints we have a sequence of convex penalty functions $\{g_t\}$, and the goal is that the average penalty $\frac{1}{T}\sum_{t=1}^T g_t(x_t)$ goes to zero as $T \to \infty$. Specific instances of this problem are when

[2] See Rockafellar (1984); Georgiadis et al. (2006) for an introduction to modeling different types of network flow problems. The constraint can also be written as an inequality, meaning that the supply is equal to the demand. The symbol $\preceq$ indicates entry-wise comparison of two vectors.

[3] See also the example in (Mannor et al., 2009, Sec. 8).

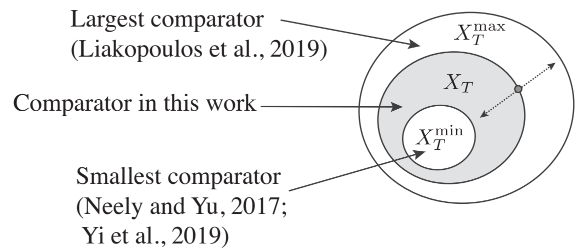

Figure 1: Illustrating the relationship between our comparator (X_T) and $X_T^{\max}$ and $X_T^{\min}$.

the constraints are static (i.e., $g_t = g_{t+1}$ for all $t \in \mathbb{N}$) as in Mahdavi et al. (2012a); Jenatton et al. (2016), or when the constraints are online (g_t is *not* known at the time to select an action x_t). In this work, we focus on the last type since it naturally encompasses a vast range of problems in communications, network control, and operations research.

1.2 Research question, previous work, and contributions

A fundamental question in OCO with online constraints is: *What is the comparator (i.e., the set) from which we select the best fixed decision in hindsight?* Or put differently: *Which is the largest comparator we can use to obtain sublinear regret and constraint violation?* The answer to these questions is non-trivial since, unlike standard OCO where the constraints are static, now the constraints vary over time and so the set of admissible fixed decisions in hindsight.

A natural choice of comparator would be

$$X_T^{\max} := \left\{ x \in C \mid \textstyle\sum_{t=1}^T g_t(x) \preceq 0 \right\},$$

where C is the set of actions that can be selected in every round. However, Mannor et al. (2009) showed that it is not possible, in general, to design an algorithm that obtains sublinear regret and constraint violation against a fixed decision in $X_T^{\max}$. Because of the latter, previous work (e.g., (Neely and Yu, 2017; Yi et al., 2019; Chen et al., 2017; Cao and Liu, 2019)) has proposed OCO algorithms with online constraints that compare their performance to the weaker comparator

$$X_T^{\min} := \{x \in C \mid g_t(x) \preceq 0 \text{ for all } t = 1, 2, \dots, T\}.$$

However, using $X_T^{\min}$ as a comparator is far from ideal. A fixed decision $x \in X_T^{\min}$ has to satisfy all the constraints, and an adversary can shrink $X_T^{\min}$ with a single attack. In the online advertisement problem, for example, a bot can affect the comparator by generating fictitious clicks in just one round. Hence, $X_T^{\min}$ is vulnerable in adversarial settings. Observe the latter is

Table 1: The comparator is the set from which the fixed decision in hindsight is selected. CV and LR stand for constraint violation and learning rate respectively. In (Jenatton et al., 2016), $\beta \in (0,1)$. In (Yu et al., 2017), the bounds and the comparator are defined in expectation. In Yi et al. (2019), $\kappa \in (0,1)$. The bounds in (Liakopoulos et al., 2019) correspond to selecting $K = T^{1-\frac{1}{T}}$ and $V = T^{1-\frac{1}{2T}}$ in (Liakopoulos et al., 2019, Theorem 1). In this work, $\epsilon \in [0,1)$. The works with (†) assume the Slater condition.

Paper	**Constraints**	**Comparator**	**Regret**	**CV**	**LR**
Mahdavi et al. (2012a)	static	$X^{\max}$	$O(T^{\{\frac{1}{2},\frac{2}{3}\}})$	$O(T^{\{\frac{3}{4},\frac{2}{3}\}})$	constant
Jenatton et al. (2016)	static	$X^{\max}$	$O(T^{\beta} \vee T^{1-\beta})$	$O(T^{1-\beta/2})$	adaptive
Yu et al. (2017)†	stochastic	$X^{\max}$	$O(\sqrt{T})$	$O(\sqrt{T})$	constant
Neely and Yu (2017)†	online	$X_T^{\min}$	$O(\sqrt{T})$	$O(\sqrt{T})$	constant
Yi et al. (2019)†	online	$X_T^{\min}$	$O(T^{\kappa} \vee T^{1-\kappa})$	$O(T^{\kappa} \vee T^{1-\kappa})$	adaptive
Liakopoulos et al. (2019)	online	$X_T^{\max}$	$O(T^{1-\frac{1}{2T}})$	$O(T^{1-\frac{1}{4T}})$	constant
This work†	online	$X_T^{\min} \subseteq X_T \subseteq X_T^{\max}$	$O(T^{\epsilon} \vee T^{1-\epsilon})$	$O(T^{\epsilon})$	adaptive

not the case for $X_T^{\max}$, where the size of the set depends on the whole sequence of constraints $\{g_t\}_{t=1}^T$ (i.e., a sequence of attacks).

Contributions: We propose an algorithm for solving OCO problems with online perturbed constraints that obtains (i) optimal rates and (ii) uses a comparator X_T that is *strictly* larger than the one used in the state of the art. In particular, we show that $X_T^{\min} \subseteq X_T \subseteq X_T^{\max}$ (see Fig. 1), and that $X_T^{\min} \subset X_T$ unless $X_T^{\min} = X_T^{\max}$ (which corresponds to the deterministic case where $g_t = g_{t+1}$ for all $t = 1, 2, \dots, T$). Our results improve upon previous work on OCO with online constraints that obtain optimal rates but use $X_T^{\min}$ as a comparator (Neely and Yu, 2017; Yi et al., 2019). The work by Liakopoulos et al. (2019) proposes an algorithm that can compare with $X_T^{\max}$, however, the performance/rate of the algorithm degrades as the horizon increases.[4] Our results also clarify how the comparator affects the regret bound and provide a unified expression for analyzing the regret in a variety of settings. Table 1.1 summarizes the differences between this paper and previous approaches. The works by Mahdavi et al. (2012a), Jenatton et al. (2016), Yu et al. (2017) consider long-term constraints that are fixed or stochastic (i.e., non-adversarial), and so it is possible to use $X^{\max}$ as a comparator.[5] We leave in Sec. 4 an extended technical discussion of previous work, the impossibility result by Mannor et al. (2009), and the Slater condition.

The key to our results relies on modeling online constraints using perturbations. Specifically, we let $g_t(x) := g(x) + b_t$, where $g : \mathbb{R}^n \to \mathbb{R}$ is a static convex constraint and $b_t \in \mathbb{R}^m$ an arbitrary perturbation (i.e., it does not need to be i.i.d. or have any other statistical property). For instance, in the online shortest path problem $g(x) = -Ax$, and b_t is the time varying supply/demand vector. In the online advertisement problem, $g(x)$ is an approximate model of the "expected" cost per click, and $b_t = g_t(x_t) - g(x)$ captures the difference between our model and the true penalty.[6]

The proposed algorithm (Algorithm 1) is based on an online primal-dual proximal gradient method and obtains $O(T^{\epsilon} \vee T^{1-\epsilon})$ regret and $O(T^{\epsilon})$ accumulated constraint violation, where $\epsilon \in [0,1)$ is a parameter in the learning rate. Our algorithm allows us to balance between regret and constraint violation by choosing ϵ accordingly. When $\epsilon = 1/2$, our bounds match the best well-known rates (Neely and Yu, 2017; Yi et al., 2019), but we can also obtain faster violation rate than $O(\sqrt{T})$. For example, with $\epsilon = 1/4$, we have $O(T^{3/4})$ regret and $O(T^{1/4})$ constraint violation. Another key characteristic of our algorithm is that the learning rate is adaptive, and so we do not need to fix in advance the time the algorithm will run (which is not known in many resource allocation problems). The latter is challenging technically since the set in which the dual variables exists is unbounded.

The rest of the paper is organized as follows. Sec. 2 presents the problem model and Sec. 3 the main technical results. In Sec. 4, we discuss related work, and in Sec. 5, we present a numerical example that shows the performance of the proposed algorithm against different comparators. The proof of our main result, Theorem 1, is in the supplementary material.

[4] In particular, select parameters $K = T^{1-\frac{1}{T}}$ and $V = T^{1-\frac{1}{2T}}$ as indicated in (Liakopoulos et al., 2019, Theorem 1). Then, $\limsup_{T\to\infty} R_T/T \le \limsup_{T\to\infty} O(T^{1-\frac{1}{2T}})/T = \lim_{T\to\infty} O(T^{-1/T}) = O(1)$, i.e., a constant.

[5] The comparator does not depends on T when the constraints are static. With stochastic constraints, the comparator is defined in expectation.

[6] For instance, an increasing function such that $g(0) = 0$. We can expect to have more clicks if we bid higher, and no clicks if we do not bid.

2 Problem Model

The standard OCO framework can be extended to encompass long-term constraints with additive perturbations as follows. Let C be a convex set that contains the admissible or implementable actions, and $g^{(j)}(x) : \mathbb{R}^n \to \mathbb{R}$, $j \in \{1, \ldots, m\}$ be a collection of convex constraints that need to be satisfied on average. Each constraint $g^{(j)}$ has an associated perturbation $b_t^{(j)}$ that varies with time. There is no need for the perturbations to be i.i.d., have zero mean, or any other statistical property. The only assumption we will make is that they satisfy a mild Slater condition, which is typical in convex optimization (see Sec. 3.2). In each round $t \in \mathbb{N}$, an agent must select an action $x_t \in C$ without knowledge of the cost function f_t or the perturbation b_t.

We proceed to define the feasible set, the regret, and the constraint violation measure. To keep notation short, we let $g := (g^{(1)}, \ldots, g^{(m)})$ and $b_t := (b_t^{(1)}, \ldots, b_t^{(m)})$. The time-varying feasible set or comparator (i.e., the set from which we select the best fixed decision in hindsight) is given by

$$X_T(\omega) := \{x \in C \mid g(x) + \omega \preceq 0\}, \tag{3}$$

where $b_T^{\text{avg}} \preceq \omega \preceq b_T^{\max}$ with $b_T^{\text{avg}} := \frac{1}{T}\sum_{t=1}^T b_t$, $b_T^{\max} := \max\{b_t, t = 1, \ldots, T\}$.[7] The use of a parameterized comparator is a key difference with previous work. Observe that if we select $\omega = b_T^{\max}$, then our comparator is equal to $X_T^{\min}$; as in (Neely and Yu, 2017). If $\omega = b_T^{\text{avg}}$, then our comparator is equal to $X_T(b_T^{\text{avg}})$, such as in (Liakopoulos et al., 2019). And if $b_T^{\text{avg}} = b_T^{\max}$, we have that $X_T^{\min} = X_T(\omega) = X_T^{\max}$, which corresponds to the case of static constraints (Mahdavi et al., 2012a). Note that by construction we always have that

$$X_T^{\min} \subseteq X_T(\omega) \subseteq X_T^{\max}.$$

Since we use a parameterized comparator, we also parameterize the regret

$$R_T(\omega) := \sum_{t=1}^T f_t(x_t) - \min_{x \in X_T(\omega)} \sum_{t=1}^T f_t(x). \tag{4}$$

We define the sum of the constraint violations as follows

$$V_T := \left\| \left[\sum_{t=1}^T g(x_t) + b_t \right]^+ \right\|, \tag{5}$$

where $[z]^+ := (\max\{0, z^{(1)}\}, \ldots, \max\{0, z^{(m)}\})$ is the projection of each of the components of vector $z \in \mathbb{R}^m$ onto the non-negative orthant. Note that V_T does not depend on ω. Recall we would like that V_T grows at most sublinearly with T so that $\lim_{T\to\infty} V_T/T = 0$. There is no requirement that $V_T = 0$ for any particular $T \in \mathbb{N}$. Finally, note that if the sum of the penalties inflicted by a constraint $j \in \{1, \ldots, m\}$ is non-positive (i.e., $\sum_{t=1}^T g^{(j)}(x_t) + b_t^{(j)} \le 0$), then that constraint does not contribute to the aggregate constraint violation V_T.

[7] The max is component-wise.

3 Main Results

3.1 Proposed algorithm and interpretation

The main technical contribution of the paper is Algorithm 1, with which we can obtain $O(T^\epsilon \vee T^{1-\epsilon})$ regret and $O(T^\epsilon)$ constraint violation. The comparator for which those rates hold will be discussed in detail after Theorem 1.

In short, to handle long-term constraints, we define a Lagrangian-type function

$$\mathcal{L}_t(x, y) = \langle f_t'(x_t), x\rangle + \langle y, g(x) + b_t\rangle, \tag{6}$$

where $f_t'(x_t)$ is a (sub)gradient of the cost function in the previous round, and $y \in \mathbb{R}_+^m$ a vector of dual variables. To streamline exposition, in the rest of the paper we will refer to $\mathcal{L}_t(x, y)$ simply as Lagrangian. Note that the Lagrangian is convex in x, concave in y, and that it depends on t as the objective function and constraints change in each round. The second term of the Lagrangian can be regarded as a penalty or as an adaptive regularizer that allows us to steer the decisions towards $X_T(\omega)$.

Algorithm 1 is based on a regularized primal-dual proximal gradient method, where we use the general Bregman divergence as the proximal term instead of the usual squared Euclidean distance; see, for example, (Beck and Teboulle, 2003). Recall the Bregman divergence associated with function ψ is defined as

$$B_\psi(a, b) = \psi(a) - \psi(b) - \langle \nabla\psi(b), a - b\rangle, \tag{7}$$

where ψ is usually assumed to be σ_ψ-strongly convex function and differentiable. The primal update ($\circ$) is equivalent to carrying out an (unconstrained) proximal gradient update with the regularization term $\langle y, g(x) + b_t\rangle$. The regularization or penalty term is updated via the dual update ($\bullet$), which can be regarded as applying a standard proximal gradient ascent since $(g(x_{t+1}) + b_{t+1}) \in \partial_y \mathcal{L}_{t+1}(x_{t+1}, y)$ for a fixed vector x_{t+1}. The following observation is crucial

$$\begin{aligned}\arg\min_{u\in C} \mathcal{L}_t(u, y_t) &= \arg\min_{u\in C} \langle f_t'(x_t), u\rangle + \langle y_t, g(u) + b_t\rangle \\ &= \arg\min_{u\in C} \langle f_t'(x_t), u\rangle + \langle y_t, g(u)\rangle,\end{aligned}$$

Algorithm 1

Input: Bregman functions ψ and φ; vector $f'_1 = 0$; set C.
Set: $x_1 \in C$; $y_1 = 0$; and $\epsilon \in [0, 1)$.
for $t = 1, 2, \dots$ **do**
  $\rho \leftarrow 1/t^\epsilon$
  ($\circ$) $x_{t+1} \leftarrow \arg\min_{u \in C}\{\mathcal{L}_t(u, y_t) + \frac{1}{\rho} B_\psi(u, x_t)\}$
  ($\bullet$) $y_{t+1} \leftarrow \arg\max_{v \in \mathbb{R}^m_+}\{\langle v, g(x_{t+1}) + b_{t+1}\rangle - \frac{1}{\rho} B_\varphi(v, y_t)\}$
  $f_{t+1} \leftarrow$ play action x_{t+1} and learn cost function
end for

That is, the primal update is oblivious to perturbation b_t. Hence, the perturbation is only relevant in our algorithm in the update of the dual variables.[8]

Finally, observe that we use step size ρ equal to $t^{-\epsilon}$ with $\epsilon \in [0, 1)$ for both updates, so there is no need to fix in advance the time horizon the algorithm will run. Note that when $\epsilon = 0$, then the algorithm corresponds to using constant step size $\rho = 1$. The algorithm's complexity depends on the structure of the constraints and the functions associated with the Bregman divergence terms in the primal and dual updates. When $g(x)$ is linear (e.g., $g(x) = Ax$) and ψ, φ equal to $\frac{1}{2}\|\cdot\|_2^2$, Algorithm 1 has the same complexity than previous work on OCO with long-term constraints (Mahdavi et al., 2012a; Jenatton et al., 2016; Yu et al., 2017). In particular, the primal and dual updates can be written as $x_{t+1} = \mathcal{P}_C(x_t - (\rho/2)\langle A^T, y_t\rangle)$ and $y_{t+1} = [y_t + (\rho/2)(g(x_{t+1}) + b_{t+1})]^+$.

3.2 Assumptions

Fix a norm $\|\cdot\|$. To establish the convergence of Algorithm 1, we need to make the following assumptions:

Bounded set. Set C is bounded. There exists a constant D such that $\|u - v\| \leq D$, $\forall u, v \in C$.

Bounded perturbation. $\|b_t\| \leq B$ for all $t \in \mathbb{N}$ for some constant $B \geq 0$.

Bounded subgradients. Let $\|\cdot\|_*$ denote the dual norm of $\|\cdot\|$. There exist constants F_*, G_*, G such that $\|f'_t(x)\|_* \leq F_*$, $\|g(x) + b_t\|_* \leq G_*$, $\|g(x) + b_t\| \leq G$ for all $x \in C$, $t \in \mathbb{N}$.

Slater condition. There exists a $\eta > 0$ such that $g(\hat{x}) + b_t + \eta \mathbb{1} \preceq 0$ for an $\hat{x} \in C$ and all $t \in \mathbb{N}$.

Bregman functions. ψ and φ are $\sigma_\psi, \sigma_\varphi$-strongly convex and L_ψ, L_φ-smooth. Also, φ is strictly increasing.

[8]This is due to the fact that y is the dual variable of the additive perturbation on the constraints. See Sec. D in the supplementary material for more details.

The first assumption is standard in OCO. The second and third assumptions are also standard in OCO and ensure that the subgradients used in the primal and dual updates are bounded. The Slater condition says that there is a set of actions that satisfy the constraints $g(x) + b_t$ strictly for all $t \in \{1, \dots, T\}$, and is key to ensure that the constraint violation V_T is sublinear. Importantly, the Slater condition assumption is mild in many problems. For example, when the perturbation b_t represents the budget available at time t, that budget has to be always positive—independently of whether we decide to spend more (i.e., violate the constraint). Finally, the assumption that function ψ and φ are strongly convex is also standard in the definition of the Bregman divergence. The additional assumption that ψ and φ are smooth (hence, ψ and φ are upper and lower bounded by a quadratic function) is to streamline exposition in the proofs.[9] The assumption that φ is strictly increasing is necessary to obtain the faster rates on the constraint violation when $\epsilon \in [0, \frac{1}{2})$. Note that all the conditions are satisfied, for example, by the squared Euclidean distance.

3.3 Bounds and discussion

The following theorem is the main result of the paper.

Theorem 1. *Consider Algorithm 1 and suppose the assumptions in Sec. 3.2 are satisfied. For any $\omega \in [b_T^{\text{avg}}, b_T^{\max}]$, the following bounds hold*

$$R_T(\omega) \leq \frac{\mathbb{S}}{\rho_T} + \mathbb{G}\sum_{t=1}^T \rho_t + \Delta(\omega) = O(T^\epsilon \vee T^{1-\epsilon}) + \Delta(\omega)$$

$$V_T \leq G + \frac{L_\varphi}{2\rho_T} E = O(T^\epsilon)$$

where $\mathbb{S} := \left(\frac{L_\psi}{2} D^2 + \frac{L_\varphi}{2} E^2\right)$, $\mathbb{G} := \left(\frac{2F_*^2}{\sigma_\psi} + \frac{2G_*^2}{\sigma_\varphi}\right)$,

$$\Delta(\omega) := \sum_{t=1}^T \langle y_{t+1}, b_{t+1} - \omega\rangle,$$

and E is a constant that does not depend on T. In particular, E captures the diameter of the set in which the dual variables are contained (i.e., $\|y_t\| \leq E$ for all $t = 1, \dots, T$) and is given by

$$E := \sqrt{\frac{L_\varphi}{\sigma_\varphi}\left(\frac{2\chi}{\eta}\right)^2 + \frac{2}{\sigma_\varphi}\chi}$$

where $\chi := \frac{6G_*^2}{\sigma_\varphi} + 3F_* D + \frac{L_\psi D^2}{2}$.

[9]Technically, all we need is that $B_\psi(u, v)$ is uniformly upper bounded for all $u, v \in C$. Such assumption is also made in previous work and elsewhere to streamline exposition; see, for example, (Mahdavi et al., 2012a, Lemma 10) or Duchi et al. (2011).

Interpretation of the regret bound: This consists of three terms. The first two are the typical terms in standard OCO[10], while the third term depends on the ω in the comparator; see Eq. (3). When ψ and φ are the squared Euclidean distance (i.e., $L_\psi, L_\varphi, \sigma_\psi, \sigma_\varphi = 2$), we have $R_T(\omega) \leq \frac{1}{\rho_T}\left(D^2+E^2\right)+(F_*^2+G_*^2)\sum_{t=1}^T \rho_t + \Delta(\omega)$. The first term is related to the size of the sets where the primal and dual variables are contained (i.e., D and E respectively) and is inversely proportional to the learning rate at time T (i.e., $\rho_T^{-1} = T^\epsilon$). The second term consists of the bounds on the subgradients of the cost functions (F_*) and the constraints (G_*) multiplied by $\sum_{t=1}^T \rho_t \leq 1+\int_1^T t^{-\epsilon}dt \leq 1+\int_0^T t^{-\epsilon}dt \leq 1+\frac{T^{1-\epsilon}}{1-\epsilon}$.[11] The bounds in Theorem 1 are of course useful if the constants are bounded, which is the case for D, F_* and G_* by standard OCO assumptions (see Sec. 3.2). However, for constant E we need more work. To show that this constant exists is one of the main technical challenge of the paper; we will discuss it in detail later in the section.

Term $\Delta(\omega)$: Our regret bound allows us to use different comparators, i.e., we can compare our algorithm to different hindsight policies. The impact of the comparator is captured in the term $\Delta(\omega)$. We proceed to show how $\Delta(\omega)$ is affected by the perturbations and ω. We study four cases:

Case (i) – Pessimistic: $\left(\omega = b_T^{\max}\right)$. In this case, we trivially have that $\Delta(\omega) \leq 0$ since $b_{t+1} - b_T^{\max} \preceq 0$ and $y_t \succeq 0$ for all $t = 1 \ldots, T$. Recall that by selecting $\omega = b_T^{\max}$, then $X_T(\omega) = X_T^{\min}$.

Case (ii) – Stochastic: $\big(b_t$ are i.i.d. random variables with $\mathbf{E}(b_t) = b$ for all $t = 1,\ldots,T\big)$. In this case, we have that

$$\mathbf{E}(\Delta(b)) = 0.$$

Let $b_t = b + \delta_t$, where δ_t is a vector with $\mathbf{E}(\delta_t) = 0$ (the all zeroes vector). By the linearity of the expectation and the fact that δ_t and y_t are independent, we have $\mathbf{E}(\Delta(b)) = \mathbf{E}(\sum_{t=1}^T \langle y_t, b_{t+1} - b\rangle) = \mathbf{E}(\sum_{t=1}^T \langle y_t, \delta_{t+1}\rangle) = \sum_{t=1}^T \langle y_t, \mathbf{E}(\delta_{t+1})\rangle = 0$.

This case can be seen as having static constraints (hence, $X^{\min} = X_T(b) = X^{\max}$ for all $T \in \mathbb{N}$) and an unbiased noise vector δ_t in the dual update (•). The bounds in Theorem 1 still hold in expectation, and the Slater condition must be satisfied also in expectation.

Case (iii) – Convex combination: $(\omega = (1-\rho_t)b_T^{\max} + \rho_t b_T^{\text{avg}})$. Note that ω is the convex combination of $b_T^{\max}$ and b_T^{avg} since $\rho_t \in (1/t, 1]$. In this case, the comparator $X_T(\omega)$ is in between $X_T^{\min}$ or $X_T^{\max}$. We can upper bound $\Delta(\omega)$ as follows

$$\begin{aligned}
\Delta(\omega) &= \textstyle\sum_{t=1}^T \langle y_{t+1}, b_{t+1} - \omega\rangle \\
&= \textstyle\sum_{t=1}^T \langle y_{t+1}, b_{t+1} - b_T^{\max} + \rho_t(b_T^{\max} - b_T^{\text{avg}})\rangle \\
&\leq \textstyle\sum_{t=1}^T \rho_t \langle y_{t+1}, b_T^{\max} - b_T^{\text{avg}}\rangle && (8)\\
&\leq \textstyle\sum_{t=1}^T \rho_t \|y_{t+1}\| \|b_T^{\max} - b_T^{\text{avg}}\| && (9)\\
&\leq 2EB \textstyle\sum_{t=1}^T \rho_t && (10)\\
&\leq 2EB\left(1 + \frac{T^{1-\epsilon}}{1-\epsilon}\right) && (11)
\end{aligned}$$

where Eq. (8) follows by dropping $\sum_{t=1}^T \langle y_{t+1}, b_{t+1} - b_T^{\max}\rangle$ (case (i)); Eq. (9) by Cauchy-Schwarz; Eq. (10) since $\|b_t\| \leq B$ and $\|y_t\| \leq E$; and Eq. (11) by the upper bound on $\sum_{t=1}^T \rho_t$ given in the previous section. We have arrived to the following corollary.

Corollary 1. *Consider the setup of Theorem 1 and select* $\omega_{\text{cvx}} = (1-\rho_t)b_T^{\max} + \rho_t b_T^{\text{avg}}$. *Then,* $R_T(\omega_{\text{cvx}}) \leq O(T^\epsilon \vee T^{1-\epsilon})$ *and* $V_T \leq O(T^\epsilon)$.

Finally, we note that since $\rho_t > 0$, then $X_T^{\min} \subset X_T$ unless $X_T^{\min} = X_T^{\max}$, i.e., our comparator is strictly larger than $X_T^{\min}$. The latter means that Algorithm 1 has optimal regret and constraint violation rates ($\epsilon = 0.5$) with respect to a stronger benchmark.

Case (iv) – General: $(\omega_\kappa \in \{s \in [b_T^{\text{avg}}, b_T^{\max}] \mid \Delta(s) \leq \kappa\}$ where $\kappa \geq 0)$. This case corresponds to enforcing $\Delta(\omega_\kappa)$ to be smaller than κ, which is always possible since by the first case $\Delta(b_T^{\max}) \leq 0 \leq \kappa$. Note that we could also allow κ to increase with T in order to obtain a larger comparator. For instance, we could set $\kappa = O(T^\epsilon \vee T^{1-\epsilon})$, and so $\Delta(\omega_\kappa)$ will grow at the same rate than the fastest of the two terms in the regret bound in Theorem 1. The choice of ω in case (iii) is an example that ensures that $\Delta(\omega_\kappa)$ grows at a rate of at most $O(T^{1-\epsilon})$.

Interpretation of the constraint violation bound: The bound on the accumulated constraint violation V_T consists of two terms. The first term is a constant related to the constraints, and the second term depends on constants E and L_φ, and is divided by the learning rate at time T. Hence, if $\epsilon = 0$ we have that $V_T \leq O(1)$; however, constant constraint violation comes at the price of the regret bound not being sublinear. Also, observe that for any ϵ in the range $[0, \frac{1}{2})$, the constraint violation has better rate than the regret. Finally, note again that the constraint violation rates do not depend on ω since G, L_φ and E are just constants.

Constant E: This constant is analogous to constant D, which measures the maximum distance between *any* two vectors in the *bounded* set C of primal variables; see Sec. 3.2. However, we cannot define E in the same

[10] See, for example, (Zinkevich, 2003, Theorem 1). More specifically, the regret bound in the second column on page 4, i.e., $R_T \leq \|F\|^2 \frac{1}{2\eta_T} + \frac{\|\nabla c\|^2}{2}\sum_{t=1}^T \eta_t$.

[11] We write $O(T^\epsilon \vee T^{1-\epsilon})$ instead of $O(T^\epsilon \vee \frac{T^{1-\epsilon}}{1-\epsilon})$ in Theorem 1 as the interesting range is when $\epsilon \in [0, \frac{1}{2}]$.

way as the dual variables exist in the nonnegative orthant (which is an *unbounded* set). Instead, we show that the difference between the vectors generated by the dual update in Algorithm 1 is bounded. Or equivalently, that the dual variables obtained with Algorithm 1 remains bounded for all $t \in \mathbb{N}$; see Lemma 7 in the supplementary material.

To ensure that $\|y_t\|$ is bounded for any $t \in \mathbb{N}$, we rely on the Slater condition. In brief, this condition requires that there exists an $x \in C$ such that $g(x) + b_t + \eta\mathbb{1} \preceq 0$ for some scalar $\eta > 0$, and ensures that the dual variables in Algorithm 1 are attracted to a bounded set within $\mathbb{R}^m_+$.[12] Constant η can be regarded, informally, as the minimum curvature constant of a strongly convex function. The technical challenge is to characterize the diameter of the set to which these dual variables are attracted since unlike standard optimization with a static objective function, in OCO the cost functions vary over time and, indirectly, the (bounded) sets to which the dual variables are attracted. See Proposition 2 and discussion in Section B.2 in the supplementary material.

Finally, we note that $E = O(D^2)$. For instance, when ψ and φ are the squared Euclidean distance, E becomes $\sqrt{2\chi^2\eta^{-2} + \chi}$ where $\chi := \frac{6G_*^2}{\sigma_\varphi} + 3F_*D + \frac{L_\psi D^2}{2}$. The latter implies that $R_T(\omega) \leq O(D^4(T^\epsilon \vee T^{1-\epsilon}))$ and $V_T \leq O(D^2 T^\epsilon)$.

Online convex optimization: Our analysis also allow us to recover the standard OCO bound when the constraints are always satisfied. We have the following corollary to Theorem 1.

Corollary 2. *Suppose that $X_T = C$ (i.e., $g(x)+b_t \preceq 0$ for all $x \in C$ and $t \in \mathbb{N}$). The regret bound becomes $R_T \leq \frac{1}{\rho_T}\left(\frac{L_\psi D^2}{2}\right) + \left(\frac{2F_*^2}{\sigma_\psi}\right)\sum_{t=1}^T \rho_t$.*

That is, when the constraints are always satisfied the dual variables will always be equal to zero and therefore $E = 0$ and $G_* = 0$.[13] Hence, by considering perturbed constraints in the learning problem we are adding $(2\rho_T)^{-1}L_\varphi E^2 + 2G_*^2\sigma_\varphi{}^{-1}\sum_{t=1}^T \rho_t$ to the bound of the standard regret in Corollary 2. Such symmetry is not available in previous works (Mahdavi et al., 2012a; Jenatton et al., 2016; Neely and Yu, 2017; Yu et al., 2017), and it appears in our work as Algorithm 1 can be regarded as applying OGD twice (see Lemma 3 in the supplementary material for the technical details). As a result, the constants in the usual OCO bound appear "duplicated".

Constrained convex optimization: Our results can also be applied to constrained optimization problems. The following corollary to Theorem 1 establishes the convergence of a constrained convex program with relaxed constraints and primal averaging.

Corollary 3. *Consider the setup of Theorem 1 where the objective function and constraints are constant (i.e., $f_t = f$ and $b_t = b$ for all $t \in \mathbb{N}$) and step size $\rho_t = \alpha t^{-\epsilon}$ with $\alpha > 0$. We have that*

$$\text{(i)} \quad f(\bar{x}_T) - f^\star \leq O\left(\frac{1}{\alpha T^{1-\epsilon}} + \frac{\alpha}{T^\epsilon}\right)$$

$$\text{(ii)} \quad \|[g(\bar{x}_T) + b]^+\| \leq O\left(\frac{1}{\alpha T^{1-\epsilon}}\right)$$

where $f^\star := \min_{x\in X} f(x)$ with $X = \{x \in C \mid g(x)+b \preceq 0\}$ and $\bar{x}_T := \frac{1}{T}\sum_{t=1}^T x_t$.

The result recovers the upper bound on the objective and constraint violation in Proposition 1 in (Nedić and Ozdaglar, 2009) when $\epsilon = 0$ (fixed step size), but also ensures that $f(\bar{x}_T) \to f^\star$ and $\bar{x}_T$ converges to a vector in X asymptotically as $T \to \infty$ for any $\epsilon \in (0, 1)$.

4 Related Work and Discussion

Long-term constraints: The first works of OCO with long-term constraints were motivated by the complexity of the projection step in OGD. In brief, when set X is composed of general convex constraints, the projection step involves solving a convex program that can be computationally burdensome. For example, projections onto the semidefinite cone. Expensive projections are dealt with in mainstream convex optimization by carrying them out only in the last iteration (Mahdavi et al., 2012b) or less often (Cotter et al., 2016), but that is not possible in OCO since every action incurs an instantaneous cost. The latter was noted by Mahdavi et al. (2012a), which formalized the OCO problem with static long-term constraints and proposed two algorithms that obtain $O(T^{\{\frac{1}{2},\frac{2}{3}\}})$ regret and $O(T^{\{\frac{3}{4},\frac{2}{3}\}})$ constraint violation respectively. Jenatton et al. (2016) extends the work in (Mahdavi et al., 2012a) by proposing an algorithm that can balance regret and constraint violation. In particular, the algorithm obtains $O(T^\beta \vee T^{1-\beta})$ regret and $O(T^{1-\beta/2})$ constraint violation where $\beta \in (0, 1)$ is a design parameter. Furthermore, the learning rate is adaptive, and so the algorithm can run for any time horizon. Finally, we note the recent work by Yuan and Lamperski (2018), which considers static long-term constraints where the penalties cannot cancel out.

Online constraints: This problem was addressed for the first time in Mannor et al. (2009) in an online

[12] This type of behavior is typical in *dual* subgradient methods. See, for example, Figure 8.2.6. in (Bertsekas et al. 2003). This is also discussed in detail in (Nedić and Ozdaglar, 2009); see Lemma 1.

[13] The fact that $G_* = 0$ follows by adding a slack variable s_t to change the inequality constraint to equality, i.e., $g(x) + b_t + s_t = 0$.

learning setting (not just OCO). The work showed that sublinear regret and constraint violation are not attainable in general when the comparator is equal to $X_T^{\max}$ (Mannor et al., 2009, Prop. 2.4).[14] In view of that negative result, the authors proposed to use a more restrictive regret benchmark, which is analogous to selecting a fixed decision from the smaller set $X_T^{\min}$ we consider in this paper. It is also shown that restricted comparator is tight for a special case with only one constraint; see Sec. 5 in (Mannor et al., 2009). However, that does not contradict our results since (Mannor et al., 2009) does not assume the Slater condition.

The impossibility result shown in (Mannor et al., 2009) motivated the work in (Neely and Yu, 2017) to consider OCO problems with online constraints,[15] and to propose the first algorithm to obtain $O(\sqrt{T})$ regret and constraint violation with respect to fixed decision in $X_T^{\min}$. The work in (Neely and Yu, 2017) also improves the bounds in previous works with long-term constraints, with the exception that it cannot handle linear equality constraints. Recently, Yi et al. (2019) have presented an algorithm for distributed OCO with online constraints and uses $X_T^{\min}$ as comparator. The bounds obtained are similar to ours in the sense that the step size is adaptive. However, we can further tradeoff regret for zero constraint violation and compare with a larger set (see Table 1.1).

Regarding online stochastic constraints, Yu et al. (2017) consider online constraints that are i.i.d. generated where the feasible set is defined in expectation and equal to $X^{\max}$. The proposed algorithm obtains $O(\sqrt{T})$ regret and constraint violation in expectation, and $O(\sqrt{T}\log(T)\log(\frac{1}{\delta}))$ regret and $O(\sqrt{T}\log(T)\log^{3/2}(\frac{1}{\delta}))$ constraint violation bounds that hold for every sample path with probability $1-\delta$, $\delta \in (0,1)$. The recent work by Wei et al. (2019) extends the previous work to handle stochastic linear equality constraints.

Slater condition: This condition is typical for non-linear convex constraints as otherwise the comparator would be a single point in space. The downside of requiring the Slater condition is that we cannot handle online linear equality constraints. Nonetheless, there is a broad class of problems where the constraints are inequalities and the Slater condition holds naturally. For instance, in the online advertising problem the Slater condition amounts to having the option of not bidding (i.e., spending zero) and so to remain under the budget. We are not aware of any algorithm that supports online linear inequalities constraints and obtains sublinear regret and constraint violation.

[14] In (Mannor et al., 2009, Sec. 2.4), the authors say *"... we do not see a way to reduce the problem of online learning with constraints to an online convex optimization problem, and given the results below, it is unlikely that such a reduction is possible."* We understand this refers to OCO without long-term constraints.

[15] See the discussion in Sec. I.A and last paragraph at the end on Sec. II.

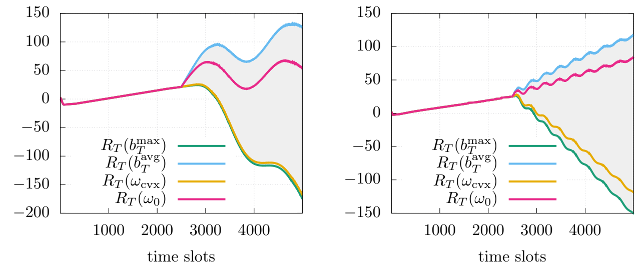

Figure 2: Illustrating the regret using different comparators for $\epsilon = 0.5$ (left) and $\epsilon = 0.25$ (right).

5 Numerical Example

We illustrate the results with a synthetic example that corresponds to the online advertising problem (see Sec. 1.1) with normalized variables and parameters. In each time slot $t = 1, 2, \ldots, T$ we select a vector $x_t \in C = [0,1]^n$ with $n = 10$. The cost and constraint functions are $\langle l_t, \cdot\rangle$ and $\langle a, \cdot\rangle + b_t - c$ respectively, where $l_t \in \{z \in [-1,0]^n \mid \mathbb{1}^T z = -1\}$, $a \in [0,1]^n$, and $b_t \in [0, c - 0.01]$ with $c = 1/n$. The cost function l_t is selected by an adversary to maximize the cost at time t, i.e., $l_t \in \arg\max_{l \in [-1,0]^n} \langle l, x_t\rangle$. The adversarial perturbation in the constraint is selected as follows: $b_t = c - 0.01$ for $t = 1, \ldots, T/2$, and $b_t = 0$ for $t = T/2+1, \ldots, T$. Note the Slater condition is satisfied, and that this choice of perturbations makes set $X_T^{\max}$ constant for the first $T/2$ slots, and then the set expands as fast as possible. That is, we are making the best fixed decision in hindsight to change quickly.

Fig. 2 shows the simulation results for $\omega \in \{b_T^{\max}, \omega_{\text{cvx}}, \omega_0, b_T^{\text{avg}}\}$ and $\epsilon = 0.5$ (left) and $\epsilon = 0.25$ (right).Both Bregman functions (ϕ, φ) are selected equal to the squared Euclidean distance. Observe from the figure that $R_T(b_T^{\max}) = R_T(\omega_{\text{cvx}}) = R_T(\omega_0) = R_T(b_T^{\text{avg}})$ for the first $T/2$ time slots since $X_T^{\max} = X_T^{\min}$ (the constraints are static), whereas for the remaining $T/2$ time slots $R_T(\omega_{\text{cvx}})$ and $R_T(\omega_0)$ are in between $R_T(b_T^{\max})$ and $R_T(b_T^{\text{avg}})$ as expected. However, note that $X(\omega_0)$ is a stronger comparator than $X(\omega_{\text{cvx}})$ since $R_T(\omega_0) \geq R_T(\omega_{\text{cvx}})$. Observe also that $R_T(\omega_{\text{cvx}})$ is negative for $t > T/2$, which means that the total cost is smaller than the cost of the best fixed decision in hindsight in $X_T(\omega_{\text{cvx}})$. This is not the behavior we would expect in an adversarial setting. With $X(\omega_0)$ as comparator, however, we obtain positive regret as we are indeed comparing to a stronger benchmark.

Acknowledgements

This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 795244.

This work was supported by Science Foundation Ireland under Grant No. 17/CDA/4760.

This work was supported by Science Foundation Ireland under Grant No. 16/IA/4610.

This research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence under Agreement Number W911NF-16-3-0001. This research was also supported by ARO (W911NF1810378) and ONR (N00014-19-1-2566).

References

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

Elad Hazan. *Introduction to online convex optimization*. Now Publishers, Inc., 2016.

R. Tyrrell Rockafellar. *Network flows and monotropic optimization*. Wiley, New York, NY, 1984.

Leonidas Georgiadis, Michael J. Neely, and Leandros Tassiulas. *Resource Allocation and Cross-Layer Control in Wireless Networks*. Foundatations and Trends in Optimization, 2006.

Shie Mannor, John N. Tsitsiklis, and Jia Yuan Yu. Online learning with sample path constraints. *Journal of Machine Learning Research*, pages 569–590, Mar 2009.

Nikolaos Liakopoulos, Apostolos Destounis, Georgios Paschos, Thrasyvoulos Spyropoulos, and Panayotis Mertikopoulos. Cautious regret minimization: Online optimization with long-term budget constraints. In *Proceedings of the 36th International Conference on Machine Learning*, 2019.

Mehrdad Mahdavi, Rong Jin, and Tianbao Yang. Trading regret for efficiency: Online convex optimization with long term constraints. *Journal of Machine Learning Research*, pages 2503–2528, Sep 2012a.

Rodolphe Jenatton, Jim C. Huang, and Cedric Archambeau. Adaptive algorithms for online convex optimization with long-term constraints. In *International Conference on Machine Learning*, 2016.

Hao Yu, Michael Neely, and Xiaohan Wei. Online convex optimization with stochastic constraints. In *Advances in Neural Information Processing Systems 30*, pages 1428–1438, 2017.

Xinlei Yi, Xiuxian Li, Lihua Xie, and Karl H. Johansson. Distributed online convex optimization with time-varying coupled inequality constraints, 2019.

Michael J Neely and Hao Yu. Online convex optimization with time-varying constraints. *arXiv preprint arXiv:1702.04783*, 2017.

T. Chen, Q. Ling, and G. B. Giannakis. An online convex optimization approach to proactive network resource allocation. *IEEE Transactions on Signal Processing*, 65(24):6350–6364, 2017.

X. Cao and K. J. R. Liu. Online convex optimization with time-varying constraints and bandit feedback. *IEEE Transactions on Automatic Control*, 64 (7):2665–2680, 2019.

Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 31(3):167 – 175, 2003.

John Duchi, Elad Hazan, and Yoram Singer. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12(Jul):2121–2159, 2011.

Angelia Nedić and Asuman Ozdaglar. Approximate primal solutions and rate analysis for dual subgradient methods. *SIAM Journal on Optimization*, 19(4): 1757–1780, 2009.

Mehrdad Mahdavi, Tianbao Yang, Rong Jin, Shenghuo Zhu, and Jinfeng Yi. Stochastic gradient descent with only one projection. In *Advances in Neural Information Processing Systems 25*, pages 494–502, 2012b.

Andrew Cotter, Maya Gupta, and Jan Pfeifer. A light touch for heavily constrained sgd. In *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 729–771, 2016.

Jianjun Yuan and Andrew Lamperski. Online convex optimization for cumulative constraints. In *Advances in Neural Information Processing Systems 31*, pages 6137–6146, 2018.

Xiaohan Wei, Hao Yu, and Michael J. Neely. Online primal-dual mirror descent under stochastic constraints, 2019.

Gong Chen and Marc Teboulle. Convergence analysis of a proximal-like minimization algorithm using bregman functions. *SIAM Journal on Optimization*, 3(3):538–543, 1993.

Lieven Vandenberghe. EE263C, UCLA. Lecture notes, 2016. URL `http://www.seas.ucla.edu/~vandenbe/ee236c.html`.

H. Uzawa. Iterative methods in concave programming. *Studies in Linear and Nonlinear Programming (K.*

Arrow, L. Hurwicz, and H. Uzawa, eds.), pages 154–165, 1958.

Dimitri P. Bertsekas, Angelia Nedić, and Asuman E. Ozdaglar. *Convex analysis and optimization*. Athena Scientific, 2003.

R. Tyrrell Rockafellar and Roger J-B Wets. *Variational Analysis*, volume 317. Springer: Grundlehren der Math. Wissenschaften., Berlin, 1998.

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

R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970.