



HAL
open science

Adaptive Algorithms for Online Convex Optimization with Long-term Constraints

Rodolphe Jenatton, Jim Huang, Cédric Archambeau

► **To cite this version:**

Rodolphe Jenatton, Jim Huang, Cédric Archambeau. Adaptive Algorithms for Online Convex Optimization with Long-term Constraints. 2015. hal-01247933

HAL Id: hal-01247933

<https://hal.science/hal-01247933>

Preprint submitted on 23 Dec 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Adaptive Algorithms for Online Convex Optimization with Long-term Constraints

Rodolphe Jenatton*
Amazon, Berlin*

Jim Huang†
Amazon, Seattle†

Cédric Archambeau*
{jenatton, huangjim, cedrica}@amazon.com

Abstract

We present an adaptive online gradient descent algorithm to solve online convex optimization problems with long-term constraints, which are constraints that need to be satisfied when accumulated over a finite number of rounds T , but can be violated in intermediate rounds. For some user-defined trade-off parameter $\beta \in (0, 1)$, the proposed algorithm achieves cumulative regret bounds of $\mathcal{O}(T^{\max\{\beta, 1-\beta\}})$ and $\mathcal{O}(T^{1-\beta/2})$ for the loss and the constraint violations respectively. Our results hold for convex losses and can handle arbitrary convex constraints without requiring knowledge of the number of rounds in advance. Our contributions improve over the best known cumulative regret bounds by Mahdavi, et al. (2012) that are respectively $\mathcal{O}(T^{1/2})$ and $\mathcal{O}(T^{3/4})$ for general convex domains, and respectively $\mathcal{O}(T^{2/3})$ and $\mathcal{O}(T^{2/3})$ when further restricting to polyhedral domains. We supplement the analysis with experiments validating the performance of our algorithm in practice.

1 Introduction

Online convex optimization (OCO) plays a key role in machine learning applications, such as adaptive routing in networks [2] and online display advertising [1]. In general, an OCO problem can be formulated as a sequential and repeated game between a learner and an adversary. In each round t , the learner first plays a vector $\mathbf{x}_t \in \mathcal{X} \subseteq \mathbb{R}^p$, where \mathcal{X} is a compact convex set corresponding to the set of possible solutions. The learner then incurs a loss $f_t(\mathbf{x}_t)$ for playing vector \mathbf{x}_t . The function $f_t(\mathbf{x}) : \mathcal{X} \rightarrow \mathbb{R}_+$ is defined by the adversary and can vary in each round, but it is assumed to

be convex. We say that a function $f_t : \mathbb{R}^d \mapsto \mathbb{R}_+$ is *strongly convex* with modulus $\sigma > 0$ if

$$f_t(\mathbf{x}) \leq f_t(\mathbf{y}) + \nabla f_t(\mathbf{x})^\top (\mathbf{x} - \mathbf{y}) - \frac{\sigma}{2} \|\mathbf{x} - \mathbf{y}\|^2.$$

for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$. We use the notation $\nabla f_t(\mathbf{x})$ to refer to any (sub-)gradient of f_t at \mathbf{x} . Furthermore, we say that f_t is *convex* if $\sigma = 0$.

The learner's objective is to generate a sequence of vectors $\mathbf{x}_t \in \mathcal{X}$ for $t = 1, 2, \dots, T$ that minimizes the cumulative regret over T rounds relative to the optimal vector \mathbf{x}^* :

$$\text{Regret}_T(\mathbf{x}^*) \triangleq \sum_{t=1}^T f_t(\mathbf{x}_t) - \sum_{t=1}^T f_t(\mathbf{x}^*). \quad (1)$$

The latter measures the difference between the cumulative loss of the learner's sequence of vectors $\{\mathbf{x}_t\}_{t=1}^T$ and the accumulated loss that would be incurred if the sequence of loss functions f_t would be known in advance and the learner could choose the best vector \mathbf{x}^* in hindsight.

Several algorithms have been developed over the past decade that achieve sub-linear cumulative regret in the OCO setting. The problem was formalized in the seminal work of [26], which presents an online algorithm based on projected gradient descent [3] that guarantees a cumulative regret of $\mathcal{O}(T^{1/2})$ when the set \mathcal{X} is convex and the loss functions are Lipschitz-continuous over \mathcal{X} . In [11] and [22], algorithms with logarithmic regret bounds were proposed for strongly convex loss functions. Notably, online gradient descent achieves an $\mathcal{O}(\log T)$ regret bound for strongly convex loss functions for appropriate choices of step size.

In the aforementioned work, the constraint on vector \mathbf{x}_t is assumed to hold in round t , such that a projection step is applied in every round to enforce the feasibility of each \mathbf{x}_t . For general convex sets \mathcal{X} , the projection step may require solving an auxiliary optimization problem, which can be computationally expensive (e.g., projections onto the semi-definite cone). More importantly, in practical applications, the learner may

in fact only be concerned with satisfying long-term constraints, that is, the *cumulative* constraint violations resulting from the sequence of vectors $\{\mathbf{x}_t\}_{t=1}^T$ should not exceed a certain amount by the final round T . An example of such an application is in online display advertising, where \mathbf{x}_t corresponds to a vector of ad budget allocations and the learner is primarily concerned in enforcing the long-term constraint that each ad consumes its budget in full over the lifetime of the ad. Another example is from wireless communications [18], where \mathbf{x}_t is a vector of power allocations across multiple devices, and the learner must satisfy average power consumption constraints per device.

In this work, we consider OCO problems where the learner is required to satisfy long-term constraints. For such problems, we will be concerned both with a) the learner’s cumulative regret as defined by (1) and b) the learner’s ability to satisfy long-term constraints – the notion of *long-term* shall be made more formal in Section 2.2. This class of problems was studied previously in [15, 16]. In particular, [16] considered online exponentially-weighted average in the case where the loss and constraints are linear, while [15] presented online algorithms based on projected subgradients and the mirror prox method [20]. The authors derived cumulative regret bounds for the cumulative loss and cumulative constraint violations respectively of $\mathcal{O}(T^{1/2})$ and $\mathcal{O}(T^{3/4})$ in the case of online projected subgradients, and respectively of $\mathcal{O}(T^{2/3})$ and $\mathcal{O}(T^{2/3})$ in the case of mirror prox. To our knowledge, these are the best-known regret bounds for OCO with long-term constraints. The analysis of [15] assumes the number of rounds is known ahead of time, which enables the authors to set the various constants in the algorithm that lead to the desired regret bounds. For the mirror prox method, it is additionally required that the constraint set \mathcal{X} is fully specified by a finite number of linear constraints.

The concept of long-term constraints also enables us to avoid the computation of potentially expensive projections onto the domain \mathcal{X} of \mathbf{x} in each round. This is closely related to recent work in stochastic optimization that aims to minimize the number of expensive projection steps [17]. The guarantees sought in our analysis have also similarities with the results obtained in the context of the online alternating direction method [23], where regret bounds are provided for the violation of equality constraints.

Contributions: Building on the work of [15], we propose an algorithm based on a saddle-point formulation of the OCO problem with long-term constraints, which is adaptive (i.e., the step sizes and the regularisation parameter depend on the round t). We show that the algorithm satisfies cumulative regret bounds

of $\mathcal{O}(T^{2/3})$ for both the cumulative loss and the cumulative constraint violations *without* requiring us to know the number of rounds T in advance to set the algorithmic hyperparameters. Unlike mirror prox, our method can deal with any convex constraint set \mathcal{X} , making it amenable to a wider range of OCO problems. Also, the algorithm we derive allows us to interpolate between the regret bounds of $\mathcal{O}(T^{1/2})$ and $\mathcal{O}(T^{3/4})$ from [15] and the above bound of $\mathcal{O}(T^{2/3})$, depending on how we wish to trade off between the cumulative regret associated to the loss and the long-term constraints. In addition to our analysis of regret bounds, we empirically validate our algorithm by comparing it to the methods of [15] on a) the online estimation of doubly stochastic matrices and b) the online learning of sparse logistic regression based on the elastic net penalty [27].

2 Online Convex Optimization with Long-term Constraints

2.1 Preliminaries

Consider m convex functions $g_j : \mathbb{R}^d \mapsto \mathbb{R}$ which induce a convex constraint set

$$\mathcal{X} \triangleq \left\{ \mathbf{x} \in \mathbb{R}^d : \max_{j \in \{1, \dots, m\}} g_j(\mathbf{x}) \leq 0 \right\}.$$

We assume that the set \mathcal{X} is bounded so that it is included in some Euclidean ball \mathcal{B} with radius $R > 0$ (to be further discussed in Section 5):

$$\mathcal{X} \subseteq \mathcal{B} \triangleq \left\{ \mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_2 \leq R \right\}.$$

Along with the functions g_j , we consider a sequence of convex functions $f_t : \mathbb{R}^d \mapsto \mathbb{R}_+$ such that

$$\max_{t \in \{1, \dots, T\}} \max_{\mathbf{x}, \mathbf{x}' \in \mathcal{B}} |f_t(\mathbf{x}) - f_t(\mathbf{x}')| \leq F \text{ for some } F > 0.$$

As is typically assumed in online learning [6], the functions g_j and f_t shall be taken to be Lipschitz continuous. We do not generally assume g_j and f_t to be differentiable. We further assume that for some finite $G > 0$ the (sub-)gradients of f and g_j are bounded

$$\begin{aligned} \max_{j \in \{1, \dots, m\}} \max_{\mathbf{x} \in \mathcal{B}} \|\nabla g_j(\mathbf{x})\|_2 &\leq G, \\ \max_{t \in \{1, \dots, T\}} \max_{\mathbf{x} \in \mathcal{B}} \|\nabla f_t(\mathbf{x})\|_2 &\leq G. \end{aligned}$$

We take the same constant G for, both, g_j and f_t for simplicity as we can always take the maximum between that of g_j and that of f_t . Finally, we assume that there exists a finite $D > 0$ such that the constraint functions are bounded over \mathcal{B} :

$$\max_{j \in \{1, \dots, m\}} \max_{\mathbf{x} \in \mathcal{B}} |g_j(\mathbf{x})| \leq D.$$

Finally, we note that the set of assumptions enumerated in this section are equivalent to the ones in [15].

2.2 Problem Statement

Let $\{\mathbf{x}_t\}_{t=1}^T$ be the sequence of vectors played by the learner and $\{f_t(\mathbf{x}_t)\}_{t=1}^T$ the corresponding sequence of incurred losses. We aim at efficiently solving the following optimization problem in online fashion:

$$\min_{\mathbf{x}_T \in \mathcal{B}} \left\{ f_T(\mathbf{x}_T) + \dots \right. \\ \left. + \min_{\mathbf{x}_2 \in \mathcal{B}} \left\{ f_2(\mathbf{x}_2) + \min_{\mathbf{x}_1 \in \mathcal{B}} f_1(\mathbf{x}_1) \right\} \right\} - \min_{\mathbf{x} \in \mathcal{X}} \sum_{t=1}^T f_t(\mathbf{x})$$

subject to the following *long-term constraint*:

$$\max_{j \in \{1, \dots, m\}} \sum_{t=1}^T g_j(\mathbf{x}_t) \leq 0.$$

3 Adaptive Online Algorithms based on a Saddle-point Formulation

Following [15], we consider a saddle-point formulation of the optimization problem. For any $\lambda \in \mathbb{R}^+$, $\mathbf{x} \in \mathcal{B}$ we define the following function:

$$\mathcal{L}_t(\mathbf{x}, \lambda) \triangleq f_t(\mathbf{x}) + \lambda g(\mathbf{x}) - \frac{\theta_t}{2} \lambda^2$$

where $g(\mathbf{x}) \triangleq \max_{j \in \{1, \dots, m\}} g_j(\mathbf{x})$ and $\{\theta_t\}_{t=1}^T$ is a sequence of positive numbers to be specified later. The role of g is to aggregate the m constraints into a single function. It otherwise preserves the same properties as those of individual g_j 's (sub-differentiability, bounded (sub-)gradients and bounded values; see Proposition 6 in [15] or Section 2.3 in [5] for a proof). In the saddle-point formulation (2), we will alternate between minimizing with respect to the primal variable \mathbf{x} and maximizing with respect to the dual parameter λ . A closer look at the function $\lambda \mapsto \mathcal{L}_t(\mathbf{x}, \lambda)$ indicates that we penalize the violation of the constraint $g(\mathbf{x}) \leq 0$ through, using $[u]_+ \triangleq \max\{0, u\}$,

$$\frac{1}{2\theta_t} [g(\mathbf{x})]_+^2 = \sup_{\lambda \in \mathbb{R}_+} \left[\lambda g(\mathbf{x}) - \frac{\theta_t}{2} \lambda^2 \right], \quad (2)$$

which is a penalty commonly used when going from constrained to unconstrained problems (e.g., Section 17.1 in [21]). Also, we can see from (2) that θ_t acts as a regularization parameter. We note at this point that in contrast to our method, [15] make use of a single θ that is constant in all rounds.¹

In the sequel, we study the following online algorithm where we alternate between primal-descent and dual-ascent steps:

- Initialize $\mathbf{x}_1 = \mathbf{0}$ and $\lambda_1 = 0$

¹More precisely, in [15], θ is equal to the product of a constant step size times a constant scaling factor.

- For $t \in \{1, \dots, T-1\}$:

$$\mathbf{x}_{t+1} = \Pi_{\mathcal{B}}(\mathbf{x}_t - \eta_t \nabla_{\mathbf{x}} \mathcal{L}_t(\mathbf{x}_t, \lambda_t)) \\ \lambda_{t+1} = \Pi_{\mathbb{R}^+}(\lambda_t + \mu_t \nabla_{\lambda} \mathcal{L}_t(\mathbf{x}_t, \lambda_t)),$$

where $\Pi_{\mathcal{C}}$ stands for the Euclidean projection onto the set \mathcal{C} , while $\{\eta_t\}_{t=1}^T$ and $\{\mu_t\}_{t=1}^T$ are sequences of non-negative step sizes that respectively drive the update of \mathbf{x} and λ . The algorithm is in the same vein as the ones proposed in [15, 13], but it is adaptive. The step sizes, which are different for the updates of \mathbf{x} and λ , are listed in Table 1. They result from the analysis we provide in the next section. We also derive sub-linear regret bounds associated to these instantiations of the sequences $\{\theta_t\}_{t=1}^T$, $\{\eta_t\}_{t=1}^T$ and $\{\mu_t\}_{t=1}^T$.

3.1 Main Results

We begin by listing three sufficient conditions for obtaining sub-linear regret bounds for the proposed algorithm:

(C1): For any $t \geq 2$, $\frac{1}{\mu_t} - \frac{1}{\mu_{t-1}} - \theta_t \leq 0$.

(C2): For any $t \geq 2$, $\eta_t G^2 + \mu_t \theta_t^2 - \frac{1}{2} \theta_t \leq 0$.

(C3): For some finite $U_{\eta} > 0$, $\sum_{t=2}^T \left[\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} - \sigma \right] \leq U_{\eta}$.

Conditions **C1** and **C3** impose constraints on the decreasing speed of the step sizes. We note that there is an asymmetry between μ_t and η_t : while we will always be able to control the norm of the variables \mathbf{x}_t (by design, they must lie in \mathcal{B}), the sequence $\{\lambda_t\}_{t=1}^T$ is not directly upper-bounded in the absence of further assumptions on the gradient of g , hence the most stringent condition **C1** is to avoid any dependencies on λ_t . Condition **C2** couples the behaviour of the three sequences to guarantee their validity. Finally, **C1**, **C2** and **C3** are expressed for $t \geq 2$ because of our choice for the initial conditions $\mathbf{x}_1 = \mathbf{0}$ and $\lambda_1 = 0$.

Our main result is described in the following theorem, whose proof is in Section 3.2:

Theorem 1. *Consider the convex case ($\sigma = 0$) and the corresponding instantiations of the sequences μ_t, η_t and θ_t for some $\beta \in (0, 1)$, as summarized in Table 1. It holds for any $T \geq 1$ that*

$$\sum_{t=1}^T \Delta f_t \leq \mathcal{R}_T^f \triangleq \left[RG + \frac{D^2}{6\beta RG} \right] T^\beta + \frac{2RG}{1-\beta} T^{1-\beta}$$

where $\Delta f_t \triangleq f_t(\mathbf{x}_t) - f_t(\mathbf{x}^*)$, and

$$\sum_{t=1}^T g(\mathbf{x}_t) \leq \sqrt{\frac{24RG}{1-\beta}} \left(\mathcal{R}_T^f + FT \right) T^{1-\beta}.$$

	Convex ($\sigma = 0$)	Strongly convex ($\sigma > 0$)
θ_t	$\frac{6RG}{t^\beta}$	$\frac{6G^2}{\sigma t^\beta}$
η_t	$\frac{R}{Gt^\beta}$	$\frac{1}{\sigma t}$
μ_t	$\frac{\theta_t(t+1)}{1-\beta}$	$\frac{\theta_t(t+1)}{\sigma}$
S_θ	$\frac{6RG}{1-\beta}T^{1-\beta}$	$\frac{6G^2}{\sigma(1-\beta)}T^{1-\beta}$
S_η	$\frac{R}{G(1-\beta)}T^{1-\beta}$	$\frac{1}{\sigma}(1 + \log(T))$
S_μ	$\frac{1}{6\beta RG}T^\beta$	$\frac{\sigma}{6\beta G^2}T^\beta$
U_η	$\frac{G}{R}T^\beta$	0
$\frac{1}{\eta_1} - \theta_1$	$6RG$	$\frac{6G^2}{\sigma}$
$\frac{\mu_1}{\eta_1} - \sigma$	$\frac{G}{R}$	0

Table 1: Parameter instantiations in different regimes (β denotes some number in $(0, 1)$).

For the strongly convex case ($\sigma > 0$), we also have valid instantiations of η_t, μ_t, θ_t (see Table 1), whose resulting cumulative regret bounds are tighter than those given in Theorem 1, but with the same leading terms. Theorem 1 can be stated in a simplified form, forgetting momentarily about the dependencies on $\{D, G, R, F\}$:

$$\sum_{t=1}^T \Delta f_t \leq \mathcal{O}(\max\{T^\beta, T^{1-\beta}\}),$$

$$\sum_{t=1}^T g(\mathbf{x}_t) \leq \mathcal{O}(T^{1-\beta/2}).$$

In particular, by setting $\beta = 2/3$, we obtain

$$\sum_{t=1}^T \Delta f_t \leq \mathcal{O}(T^{2/3}) \text{ and } \sum_{t=1}^T g(\mathbf{x}_t) \leq \mathcal{O}(T^{2/3}),$$

which matches the mirror prox guarantees of [15] while being valid for general convex constraint sets \mathcal{X} as opposed to just polyhedral constraint sets. Similarly, taking $\beta = 1/2$, we recover the regret bounds

$$\sum_{t=1}^T \Delta f_t \leq \mathcal{O}(T^{1/2}) \text{ and } \sum_{t=1}^T g(\mathbf{x}_t) \leq \mathcal{O}(T^{3/4}),$$

of Section 3.1 in [15]. We can, of course, also define novel trade-offs between loss and constraint violations, e.g., $\beta = 3/4$ with regret bounds of $\mathcal{O}(T^{3/4})$ and $\mathcal{O}(T^{5/8})$ respectively.

3.2 Analysis and Proofs

To analyze the above algorithm, we first introduce a series of lemmas. The analysis is analogous to that developed in [15], which we provide below for self-containedness. We begin by upper-bounding the variations of \mathcal{L}_t with respect to its two arguments. In particular, the following lemma takes advantage of the fact that the partial function $\lambda \mapsto \mathcal{L}_t(\mathbf{x}_t, \lambda)$ is not only concave as considered in [15], but *strongly* concave with parameter θ_t . This observation is at the origin of our improved regret bounds.

Lemma 1 (Upper bound of $\mathcal{L}_t(\mathbf{x}_t, \lambda) - \mathcal{L}_t(\mathbf{x}_t, \lambda_t)$). *For $\mathcal{L}_t(\mathbf{x}, \lambda)$ as defined above and for non-negative η_t, θ_t and μ_t , we have*

$$\mathcal{L}_t(\mathbf{x}_t, \lambda) - \mathcal{L}_t(\mathbf{x}_t, \lambda_t) \leq \frac{1}{2\mu_t} [b_t - b_{t+1}] - \frac{\theta_t}{2} b_t + \frac{\mu_t}{2} [\nabla_\lambda \mathcal{L}_t(\mathbf{x}_t, \lambda_t)]^2$$

where $b_t \triangleq (\lambda - \lambda_t)^2$.

Proof. Expanding $(\lambda - \lambda_{t+1})^2$ yields

$$\begin{aligned} (\lambda - \lambda_{t+1})^2 &= \left(\lambda - \Pi_{\mathbb{R}^+}(\lambda_t + \mu_t \nabla_\lambda \mathcal{L}_t(\mathbf{x}_t, \lambda_t)) \right)^2 \\ &\leq \left(\lambda - (\lambda_t + \mu_t \nabla_\lambda \mathcal{L}_t(\mathbf{x}_t, \lambda_t)) \right)^2 \\ &= (\lambda - \lambda_t)^2 - 2\mu_t(\lambda - \lambda_t) \nabla_\lambda \mathcal{L}_t(\mathbf{x}_t, \lambda_t) \\ &\quad + \mu_t^2 (\nabla_\lambda \mathcal{L}_t(\mathbf{x}_t, \lambda_t))^2 \end{aligned}$$

By strong concavity of $\mathcal{L}_t(\mathbf{x}_t, \lambda)$ with respect to λ ,

$$\mathcal{L}_t(\mathbf{x}_t, \lambda_t) - \mathcal{L}_t(\hat{\mathbf{x}}_t, \lambda) \leq (\lambda - \lambda_t) \nabla_\lambda \mathcal{L}_t(\mathbf{x}_t, \lambda_t) - \frac{\theta_t}{2} b_t.$$

Substituting the inequality for $\mu_t(\lambda - \lambda_t) \nabla_\lambda \mathcal{L}_t(\mathbf{x}_t, \lambda_t)$ completes the proof. \square

We omit the derivation for $\mathbf{x} \mapsto \mathcal{L}_t(\mathbf{x}, \lambda_t)$ that follows similar arguments. We now turn to a lower-bound of the variations of \mathcal{L}_t .

Lemma 2. *Let $\mathbf{x}^* = \arg \min_{\mathbf{x} \in \mathcal{X}} \sum_{t=1}^T f_t(\mathbf{x})$. Then*

$$\begin{aligned} \sum_{t=1}^T \mathcal{L}_t(\mathbf{x}_t, \lambda) - \mathcal{L}_t(\mathbf{x}^*, \lambda_t) &\geq \\ \sum_{t=1}^T \Delta f_t + \lambda \sum_{t=1}^T g(\mathbf{x}_t) - \frac{\lambda^2}{2} \sum_{t=1}^T \theta_t + \frac{1}{2} \sum_{t=1}^T \theta_t \lambda_t^2. \end{aligned}$$

Proof. We have $\mathcal{L}_t(\mathbf{x}_t, \lambda) - \mathcal{L}_t(\mathbf{x}^*, \lambda_t)$ equal to

$$f_t(\mathbf{x}_t) - f_t(\mathbf{x}^*) + \lambda g(\mathbf{x}_t) - \lambda_t g(\mathbf{x}^*) - \frac{\theta_t}{2} (\lambda^2 - \lambda_t^2).$$

We simply notice that $g(\mathbf{x}^*) \leq 0$ to obtain a lower bound $-g(\mathbf{x}^*) \sum_{t=1}^T \lambda_t$ to complete the proof after summing both sides over rounds $t = 1, \dots, T$. \square

Lemma 3. *Let $a_t \triangleq \|\mathbf{x}_t - \mathbf{x}\|^2$. For any $\sigma, \theta_t \geq 0$,*

$$\begin{aligned} \sum_{t=1}^T \frac{1}{2\eta_t} [a_t - a_{t+1}] - \frac{\sigma}{2} a_t &\leq \\ \frac{R^2}{2} \delta_\eta + \frac{1}{2} \sum_{t=2}^T a_t \left[\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} - \sigma \right], \\ \sum_{t=1}^T \frac{1}{2\mu_t} [b_t - b_{t+1}] - \frac{\theta_t}{2} b_t &\leq \\ \frac{\lambda^2}{2} \delta_\mu + \frac{1}{2} \sum_{t=2}^T b_t \left[\frac{1}{\mu_t} - \frac{1}{\mu_{t-1}} - \theta_t \right] \end{aligned}$$

where we have used $\delta_\eta \triangleq \frac{1}{\eta_1} - \sigma$ and $\delta_\mu \triangleq \frac{1}{\mu_1} - \theta_1$.

Proof. Shifting indices in the sums for terms that depend on $a_{t+1}/\eta_t, b_{t+1}/\mu_t$ and collecting terms that depend on a_t, b_t , we then use $a_1 = \|\mathbf{x}_1 - \mathbf{x}\|_2^2 = \|\mathbf{x}\|_2^2 \leq R^2$ and $b_1 = (\lambda - \lambda_1)^2 = \lambda^2$ to conclude. \square

We now present the key lemma of the analysis.

Lemma 4. [Cumulative regret bound] Let $\mathbf{x}^* = \arg \min_{\mathbf{x} \in \mathcal{X}} \sum_{t=1}^T f_t(\mathbf{x})$ and assume **C1**, **C2** and **C3** hold. Define $\mathcal{R}_T^f \triangleq \frac{R^2}{2} \delta_\eta + G^2 S_\eta + D^2 S_\mu + R^2 U_\eta$, where we have introduced $S_\eta \triangleq \sum_{t=1}^T \eta_t$, $S_\mu \triangleq \sum_{t=1}^T \mu_t$ and $S_\theta \triangleq \sum_{t=1}^T \theta_t$. Then, it holds that

$$\sum_{t=1}^T \Delta f_t \leq \mathcal{R}_T^f,$$

and

$$\sum_{t=1}^T g(\mathbf{x}_t) \leq \sqrt{2(S_\theta + \delta_\mu)(\mathcal{R}_T^f + FT)}.$$

Proof. By the triangle inequality, we have $\|\nabla_{\mathbf{x}} \mathcal{L}_t(\mathbf{x}, \lambda)\|_2^2 \leq 2G^2(1 + \lambda^2)$ and $(\nabla_{\lambda} \mathcal{L}_t(\mathbf{x}, \lambda))^2 \leq 2(D^2 + \theta_t^2 \lambda^2)$. We then combine Lemmas 1-3, starting from $\mathcal{L}_t(\mathbf{x}_t, \lambda) - \mathcal{L}_t(\mathbf{x}^*, \lambda_t) = \mathcal{L}_t(\mathbf{x}_t, \lambda) - \mathcal{L}_t(\mathbf{x}_t, \lambda_t) + \mathcal{L}_t(\mathbf{x}_t, \lambda_t) - \mathcal{L}_t(\mathbf{x}^*, \lambda_t)$, yielding

$$\sum_{t=1}^T \Delta f_t + \lambda \sum_{t=1}^T g(\mathbf{x}_t) - \frac{\lambda^2}{2} \left[S_\theta + \delta_\mu \right] \leq \mathcal{R}_T^f + \sum_{t=1}^T \lambda_t^2 \left[\eta_t G^2 + \mu_t \theta_t^2 - \frac{1}{2} \theta_t \right].$$

Maximizing the left-hand side with respect to $\lambda \in \mathbb{R}^+$, we obtain:

$$\sum_{t=1}^T \Delta f_t + \frac{[\sum_{t=1}^T g(\mathbf{x}_t)]_+^2}{2[S_\theta + \delta_\mu]} \leq \mathcal{R}_T^f + \sum_{t=1}^T \lambda_t^2 \left[\eta_t G^2 + \mu_t \theta_t^2 - \frac{1}{2} \theta_t \right].$$

The regret bound on the loss is obtained by using **C2** and $[\sum_{t=1}^T g(\mathbf{x}_t)]_+^2 / (2[S_\theta + \delta_\mu]) \geq 0$. The bound on constraint violations is obtained as above, but by substituting the lower bound $\sum_t \Delta f_t \geq -FT$. \square

In order to discuss the scaling of our regret bounds, we state the next simple lemma without proof

Lemma 5. Let $\beta \in (0, 1)$. Then $\sum_{t=1}^T \frac{1}{t^\beta} \leq \frac{T^{1-\beta}}{1-\beta}$.

With the above lemmas, we now prove Theorem 1:

Proof of Theorem 1. For the proposed choices of θ_t, μ_t and η_t , we can verify that **C1**, **C2** and **C3** hold. Here we focus on the convex case, the strongly convex one following along the same lines. First, we can easily see that **C1** is true as long as θ_t is non-increasing. Then, we can notice that, given the choice of μ_t , condition **C2**

is implied by the stronger condition $\eta_t \leq \frac{\theta_t}{6G^2}$ (satisfied by the choice of η_t and θ_t in Table 1). This results in

$$S_\mu = \sum_{t=1}^T \frac{1}{\theta_t(t+1)} \leq \sum_{t=1}^T \frac{t^\beta}{6RGt} \leq \frac{T^\beta}{6\beta RG}.$$

$$S_\eta = \sum_{t=1}^T \eta_t \leq \frac{R T^{1-\beta}}{G(1-\beta)}, \quad S_\theta = \sum_{t=1}^T \frac{6RG}{t^\beta} \leq \frac{6RG}{1-\beta} T^{1-\beta},$$

along with $1/\mu_1 - \theta_1 = 6RG$ and $1/\eta_1 - \sigma = G/R$. The term U_η can be obtained by summing the series $1/\eta_t - 1/\eta_{t-1} = (G/R)(t^\beta - (t-1)^\beta)$ over t , which directly simplifies by telescoping for the $\sigma = 0$ case, and is identically equal to zero for $\sigma > 0$. As a result, we obtain from Lemma 4

$$\sum_{t=1}^T \Delta f_t \leq \mathcal{R}_T^f \triangleq \left[RG + \frac{D^2}{6\beta RG} \right] T^\beta + \frac{RG}{1-\beta} T^{1-\beta} + \frac{RG}{2}$$

and for the constraint

$$\sum_{t=1}^T g(\mathbf{x}_t) \leq \sqrt{2(\mathcal{R}_T^f + FT) \left[\frac{6RG}{1-\beta} T^{1-\beta} + 6RG \right]}.$$

We obtain the desired conclusion by noticing that for any $T \geq 1$ and $\beta \in (0, 1)$, we have $\frac{T^{1-\beta}}{1-\beta} \geq 1$. \square

3.3 Towards No Violation of Constraints

We next show that our extension also applies to the more specific setting developed in Section 3.2 from [15], where additional assumptions on the gradient of g can translate into no constraint violations. For the sake of self-containedness, we briefly recall it. Assume that there exist $\gamma \geq 0$ and $r > 0$ such that the variations of g are lower bounded as

$$\min_{\mathbf{x} \in \mathbb{R}^d: g(\mathbf{x}) + \gamma = 0} \|\nabla g(\mathbf{x})\|_2 \geq r. \quad (3)$$

Denote $\mathcal{X}_\gamma \triangleq \{\mathbf{x} \in \mathbb{R}^d : g(\mathbf{x}) + \gamma \leq 0\} \subset \mathcal{X}$. It can then be shown that (see Theorem 7 in [15]):

$$\left| \sum_{t=1}^T f_t(\mathbf{x}^*) - \sum_{t=1}^T f_t(\mathbf{x}_\gamma) \right| \leq \frac{G}{r} T \gamma, \quad (4)$$

where \mathbf{x}^* and \mathbf{x}_γ are solutions of $\min_{\mathbf{x} \in \mathcal{X}} \sum_{t=1}^T f_t(\mathbf{x})$ and $\min_{\mathbf{x} \in \mathcal{X}_\gamma} \sum_{t=1}^T f_t(\mathbf{x})$ respectively. In words, the gap between the optimal value of the original optimization problem and that of the problem over \mathcal{X}_γ is well-controlled as a function of (γ, r) . Examples where (3) holds include the positive semi-definite cone, as described in Section 4 of [17]. For space limitation reasons, we state our result in a simplified form, only briefly sketching its proof that follows the same logic as that of the previous section.

Corollary 1. *Assume (3) holds. Consider the convex case ($\sigma = 0$) and some instantiations of the sequences μ_t, η_t and θ_t for some $\beta \in (0, 1)$, differing from Table 1 only up to numerical constants. There exist c_0 and c_1 depending on $\{D, G, R, F, r\}$ such that setting $\gamma \triangleq c_1 T^{-\beta/2}$, we have for any $T \geq c_0$*

$$\sum_{t=1}^T \Delta f_t \leq \mathcal{O}(\max\{T^\beta, T^{1-\beta}, T^{1-\beta/2}\}),$$

and no constraint violations $\sum_{t=1}^T g(\mathbf{x}_t) \leq 0$.

Sketch of proof. We can apply the same analysis as before to the function $g_\gamma(\mathbf{x}) \triangleq g(\mathbf{x}) + \gamma$, replacing D by $D + \gamma$ and adapting the constants in both **C2** (i.e., $\eta_t G^2 + \frac{3}{2}\mu_t \theta_t^2 - \frac{1}{2}\theta_t$) as well as for the instantiations of μ_t, η_t and θ_t . The regret bound on $\sum_{t=1}^T \Delta f_t$ is identical as earlier, with additional additive terms $3\gamma^2 S_\mu/2$ and $GT\gamma/r$ introduced as a result of (4). As for $\sum_{t=1}^T g(\mathbf{x}_t)$, the term $[\sum_{t=1}^T g(\mathbf{x}_t)]_+^2$ becomes here $[\sum_{t=1}^T g(\mathbf{x}_t) + \gamma T]_+^2$, which in turn leads to the same regret bound as previously stated, minus the contribution $-\gamma T$. We cancel out the constraint violations—scaling in $\mathcal{O}(T^{1-\beta/2})$ according to Theorem 1—by setting $\gamma = c_1 T^{-\beta/2}$. Note that c_0 is determined by examining when the extra term $3\gamma^2 S_\mu/2$ can be upper bounded by those in \mathcal{R}_T^f . \square

The regret bound presented in Corollary 1 is minimized for $\beta = 2/3$, leading to a regret of $\mathcal{O}(T^{2/3})$ with no constraint violations. This result extends Theorem 8 and Corollary 13 from [15] in that it holds for general convex domains \mathcal{X} (as opposed to only polyhedral ones).

4 Experiments

We ran two sets of experiments to validate the regret bounds obtained for our adaptive algorithms for OCO with long-term constraints and compare to the algorithms proposed in [15]. First, we examine the online estimation of doubly-stochastic matrices where the convex domain of interest \mathcal{X} is polyhedral but whose projection operator is difficult to compute [12, 10]. Second, we consider sparse online binary classification based on the elastic net penalty [27].

We shall refer to our adaptive online gradient descent (A-OGD) for convex f_t (i.e., $\sigma = 0$) as **Convex A-OGD** and for strongly convex f_t (i.e., $\sigma > 0$) as **Strongly convex A-OGD**, which enjoy the same regret guarantees of $\mathcal{O}(T^{2/3})$ for the loss and constraint. The method of [15] that handles general convex domains \mathcal{X} will be referred to as **Convex OGD**, while the mirror prox method analyzed in [15], which is only applicable to polyhedral domains, will be denoted by **Convex mirror prox**. The parameters of

Convex OGD and **Convex mirror prox** are instantiated according to [15].

Although we generate the sequence of losses $\{f_t\}_{t=1}^T$ stochastically in the experiments, we would like to emphasize that the regret bounds we obtain are also valid for adversarially-generated sequences, so that it is not required that $\{f_t\}_{t=1}^T$ are generated in i.i.d. fashion.

4.1 Doubly-Stochastic Matrices

Doubly-stochastic matrices appear in many machine learning and optimization problems, such as clustering applications [25] or learning permutations [12, 10]. Briefly, for a sequence of matrices $\{\mathbf{Y}_t\}_{t=1}^T$ in $\mathbb{R}^{p \times p}$, we solve the following optimization problem in an online fashion:

$$\min_{\mathbf{X} \in \mathbb{R}^{p \times p}} \sum_{t=1}^T \frac{1}{2} \|\mathbf{Y}_t - \mathbf{X}\|_F^2 \quad (5)$$

subject to the (linear) convex constraints

$$\mathbf{X} \geq \mathbf{0}, \quad \mathbf{X}\mathbf{1} = \mathbf{1} \text{ and } \mathbf{X}^\top \mathbf{1} = \mathbf{1}.$$

This problem can easily be mapped to OCO setting by assuming that the sequence $\{\mathbf{Y}_t\}_{t=1}^T$ is generated by random permutation matrices which are known to constitute the extreme points of the set of doubly-stochastic matrices [4]. We have $d = p^2$, $f_t(\mathbf{X}) = \frac{1}{2} \|\mathbf{Y}_t - \mathbf{X}\|_F^2$ and $m = p^2 + 4p$ to describe all the linear constraints, more specifically there are p^2 non-negativity constraints, along with $4p$ inequalities to model the $2p$ equality constraints. This leads to the following instantiations of the parameters controlling f_t and g : $R = \sqrt{p}$, $G = 2R$ and $D = R$. Note that we can apply a) **Strongly convex A-OGD** (since f_t is by construction strongly convex with parameter $\sigma = 1$), and b) **Convex mirror prox** since \mathcal{X} is polyhedral.

The cumulative regret for the loss and the long-term constraint are shown in Figures 1 and 2. They are computed over $T = 1000$ iterations with $d = 64$, and are averaged over 10 random sequences $\{\mathbf{Y}_t\}_{t=1}^T$ (the standard deviations are not shown since they are negligible). The offline solutions of (5) required for various $t \in \{1, \dots, T\}$ to compute the regret are obtained using CVXPY [8].

The results shown in Figure 1 and 2 indicate that although the cumulative regret bounds for **Strongly convex A-OGD** were not demonstrated to be tighter in our analysis than those for **Convex A-OGD**, they achieve a better cumulative regret for this problem, especially with respect to the long-term constraint. Also, while **Convex mirror prox** and **Convex A-OGD** should theoretically exhibit the same behavior, the results suggest that mirror prox is not able to decrease cumulative regret at the same

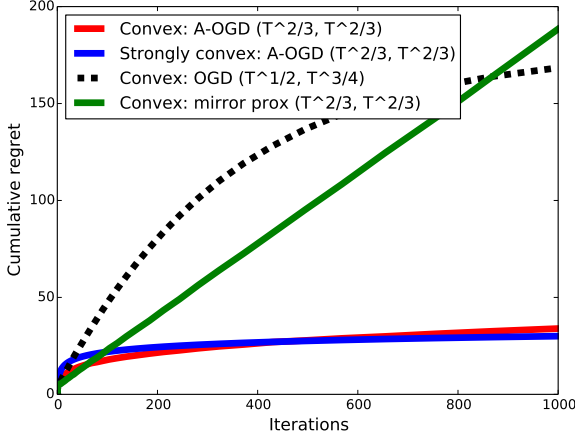


Figure 1: Cumulative regret of the loss function for the online estimation of doubly-stochastic matrices. We display the mean computed over 10 random sequences $\{\mathbf{Y}_t\}_{t=1}^T$. (Best viewed in colour.)

rate as our proposed method. We surmise that this may be due to the fact that the guarantees of **Convex mirror prox** proved in [15] only hold for very large T .²

4.2 Sparse Online Binary Classification

Next, we examine the application of sparse online binary classification. Our goal will be to minimize the log-loss subject to a constrained elastic-net penalty:³

$$\min_{\mathbf{x} \in \mathbb{R}^d: \|\mathbf{x}\|_1 + \frac{1}{2}\|\mathbf{x}\|_2^2 \leq \rho} \sum_{t=1}^T \log(1 + e^{-y_t \mathbf{x}^\top \mathbf{u}_t}), \quad (6)$$

where $\{y_t, \mathbf{u}_t\}_{t=1}^T$ denotes a sequence of label/feature-vector pairs and $\rho > 0$ is a parameter that measures the degree of the sparsity of the solutions of (6).

This problem is mapped to our formulation by setting $f_t(\mathbf{x}) = \log(1 + e^{-y_t \mathbf{x}^\top \mathbf{u}_t})$ with $m = 1$, $g(\mathbf{x}) = \|\mathbf{x}\|_1 + \frac{1}{2}\|\mathbf{x}\|_2^2 - \rho$, $R = \sqrt{1 + 2\rho} - 1$,⁴ $G = \max\{\sqrt{d} + R, \max_t \|\mathbf{u}_t\|_2\}$ and $D = \sqrt{d}R + R^2/2$. The sequences $\{y_t, \mathbf{u}_t\}_{t=1}^T$ are generated by drawing pairs at random with replacement.

We solve the above problem using the datasets `ijcnn1` and `covtype`, consisting respectively of 49,990 and 581,012 samples of dimension $d = 22$ and $d = 54$

²Theorem 12 from [15] requires $T \geq 164(m+1)^2$, which translates into $T > 10^7$ in our setting.

³Constraint formulations with sparsity-inducing terms are sometimes preferred over their penalized counterparts when they express some concrete physical budget, e.g., in the context of learning predictors with low-latency [24].

⁴The value for R is found by noticing that $\|\mathbf{x}\|_1 + \frac{1}{2}\|\mathbf{x}\|_2^2 \geq \|\mathbf{x}\|_2 + \frac{1}{2}\|\mathbf{x}\|_2^2$ and solving the resulting second-order polynomial inequality.

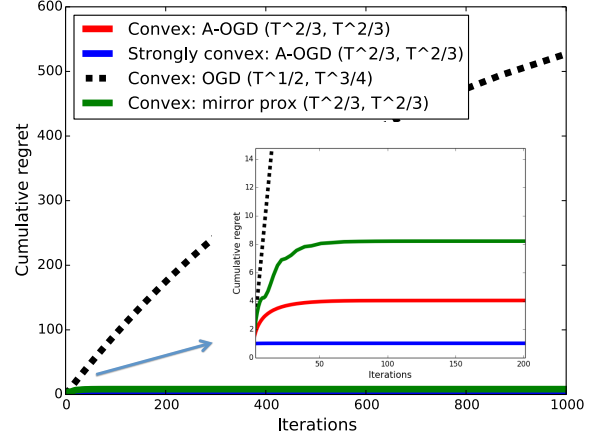


Figure 2: Cumulative regret of the long-term constraint for the estimation of doubly-stochastic matrices. We display the mean computed over 10 random sequences $\{\mathbf{Y}_t\}_{t=1}^T$. The embedded graph is a zoom of the area of interest in the original figure. (Best viewed in colour.)

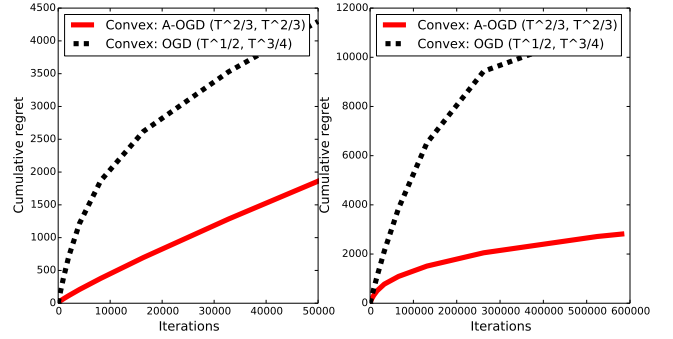


Figure 3: Cumulative regret of the loss function for the sparse online binary classification (left: `ijcnn1`, right: `covtype`). We display the mean computed over 10 random sequences $\{y_t, \mathbf{u}_t\}_{t=1}^T$.

each⁵. The parameter ρ is set to obtain approximately 30% of non-zero variables. Moreover, and in order to best display cumulative regret, we compute offline solutions of (6) for various $t \in \{1, \dots, T\}$ thanks to an implementation of [7].

The results are summarized in Figures 3 and 4 and represent an average over 10 random sequences $\{y_t, \mathbf{u}_t\}_{t=1}^T$ (the standard deviations are not shown since they are negligible). The number of iterations T is equal to the number of samples in each dataset. Interestingly, we observe that the constraint is not violated on average (i.e., via a negative cumulative regret) and the iterates \mathbf{x}_t remain feasible within the domain $\|\mathbf{x}\|_1 + \frac{1}{2}\|\mathbf{x}\|_2^2 \leq \rho$. This tendency is more pronounced for **Convex OGD**

⁵www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/binary.html

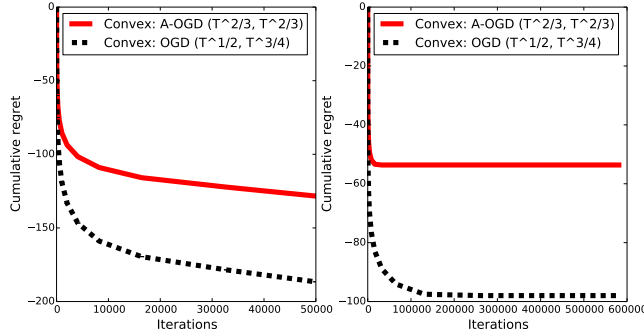


Figure 4: Cumulative regret of the constraint for the sparse online binary classification (left: `ijcnn1`, right: `covtype`). We display the mean computed over 10 random sequences $\{y_t, \mathbf{u}_t\}_{t=1}^T$.

since a closer inspection of the sequence $\{\eta_t\}_{t=1}^T$ shows numerical values smaller than those of our approach `Convex A-OGD` (by 2 to 3 orders of magnitude). As a result, starting from $\mathbf{x}_1 = \mathbf{0}$, we found that the iterates generated by `Convex OGD` do not approach the boundary of the domain, hence increasing regret on cumulative loss. We also note that the offline solutions of (6) always saturate the constraint. Although our analysis predicts that the cumulative regret of `Convex OGD` associated to the loss (i.e., $\mathcal{O}(T^{1/2})$) should be smaller than that associated to `Convex A-OGD` (i.e., $\mathcal{O}(T^{2/3})$), `Convex A-OGD` achieves here a lower cumulative regret. This observation may be explained by the same argument as that described previously, namely that the larger step sizes $\{\eta_t\}_{t=1}^T$ of `Convex A-OGD` enables us to make faster progress.

5 Discussion

In this section, we discuss several generalizations.

Broader families of step sizes. We have assumed that the updates of the primal variable \mathbf{x} are driven by a projected gradient step controlled through a single step size η_t . Following the ideas developed in [19, 9], we could analyze the regret guarantees of our algorithm when there is a vector of step sizes $\boldsymbol{\eta}_t$ that is given by a diagonal matrix $\text{Diag}(\boldsymbol{\eta}_t) \in \mathbb{R}^{d \times d}$, updating adaptively and separately each coordinate of \mathbf{x} .

Can we identify a better penalty? In the light of (2), it is tempting to ask whether we can find a penalty function that will lead to lower cumulative regret guarantees. To this end, we could for example introduce a smooth, 1-strongly-convex function ϕ with domain Ω . The saddle-point formulation of the new problem is then given by

$$\mathcal{L}_t(\mathbf{x}, \lambda) \triangleq f_t(\mathbf{x}) + \lambda g(\mathbf{x}) - \theta_t \phi(\lambda),$$

where $\{\theta_t\}_{t=1}^T$ is, as earlier, a sequence of non-negative numbers to be specified subsequently for any $\lambda \in \Omega, \mathbf{x} \in \mathcal{B}$. Interestingly, it can be shown that condition **C2** becomes a first-order nonlinear ordinary differential inequality in this setting, leading to

$$\eta_t G^2 \lambda^2 + \mu_t \theta_t^2 \left[\frac{d\phi}{d\lambda} \right]^2 - \theta_t \phi(\lambda) \leq 0, \text{ for all } \lambda \in \Omega.$$

Hence, the above differential inequality suggests a family of penalty functions which we could use. In particular, we see that ϕ must grow at least quadratically and stay greater than its squared first derivative, which rules out a softmax penalty like $\lambda \mapsto \log(1 + e^\lambda)$. Moreover, the maximization with respect to λ in the last step of Lemma 4 introduces the Moreau envelope [14] of the Fenchel conjugate of ϕ , namely

$$\phi_{S_\theta}^*(u) \triangleq \sup_{\lambda \in \Omega} \left[\lambda u - S_\theta \phi(\lambda) - \delta_\mu \frac{\lambda^2}{2} \right].$$

The goal is then to find a feasible penalty ϕ of which the inverse mapping $u \mapsto [\phi_{S_\theta}^*]^{-1}(u)$ would minimize the regret bound. For instance, the inverse mapping would scale as $u \mapsto \sqrt{u}$ when using the squared ℓ_2 norm over $\Omega = \mathbb{R}^+$. We defer to future work the study of this admissible family of penalties.

Which enclosing set for \mathcal{X} ? Our current analysis relies on the idea that instead of having to perform a projection on \mathcal{X} in each update (which could be computationally costly and perhaps intractable in some cases), we restrict the iterates \mathbf{x}_t to remain within a simpler convex set $\mathcal{B} \supseteq \mathcal{X}$. While we assumed de facto an Euclidean ball for \mathcal{B} , we could consider sets enclosing \mathcal{X} more tightly, while preserving the appealing computational properties. Having a principled methodology to choose \mathcal{B} and carefully assessing its impact on the regret bounds is an interesting avenue for future research.

References

- [1] Shipra Agrawal and Nikhil R. Devanur. Fast algorithms for online stochastic convex programming. In *SODA 2015 (ACM-SIAM Symposium on Discrete Algorithms)*. SIAM-Society for Industrial and Applied Mathematics, January 2015.
- [2] Baruch Awerbuch and Robert Kleinberg. Online linear optimization and adaptive routing. *Journal of Computer and System Sciences*, 74(1):97–114, 2008.
- [3] Dimitri P. Bertsekas and John N. Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods*. Prentice-Hall, Inc., 1989.

-
- [4] Garrett Birkhoff. Tres observaciones sobre el algebra lineal. *Univ. Nac. Tucumán Rev. Ser. A*, 5:147–151, 1946.
- [5] J. M. Borwein and A. S. Lewis. *Convex Analysis and Nonlinear Optimization: Theory and Examples*. Springer, 2006.
- [6] Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006.
- [7] Aaron Defazio, Francis Bach, and Simon Lacoste-Julien. Saga: A fast incremental gradient method with support for non-strongly convex composite objectives. Technical report, preprint arXiv:1407.0202, 2014.
- [8] Steven Diamond, Eric Chu, and Stephen Boyd. CVXPY: A Python-embedded modeling language for convex optimization, version 0.2. <http://cvxpy.org/>, May 2014.
- [9] J. Duchi, E. Hazan, and Y. Singer. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12:2121–2159, 2011.
- [10] Fajwel Fogel, Rodolphe Jenatton, Francis Bach, and Alexandre d’Aspremont. Convex relaxations for permutation problems. In *Advances in Neural Information Processing Systems*, pages 1016–1024, 2013.
- [11] Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192, 2007.
- [12] D. P. Helmbold and M. K. Warmuth. Learning permutations with exponential weights. *Journal of Machine Learning Research*, 10:1705–1736, 2009.
- [13] Alec Koppel, Felicia Y Jakubiec, and Alejandro Ribeiro. A saddle point algorithm for networked online convex optimization. In *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, pages 8292–8296. IEEE, 2014.
- [14] Claude Lemaréchal and Claudia Sagastizábal. Practical aspects of the moreau–yosida regularization: Theoretical preliminaries. *SIAM Journal on Optimization*, 7(2):367–385, 1997.
- [15] Mehrdad Mahdavi, Rong Jin, and Tianbao Yang. Trading regret for efficiency: online convex optimization with long term constraints. *Journal of Machine Learning Research*, 13(1):2503–2528, 2012.
- [16] Mehrdad Mahdavi, Tianbao Yang, and Rong Jin. Efficient constrained regret minimization. Technical report, preprint arXiv:1205.2265, 2012.
- [17] 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*, pages 494–502, 2012.
- [18] Shie Mannor and John N Tsitsiklis. Online learning with constraints. In *Learning Theory*, pages 529–543. Springer, 2006.
- [19] H Brendan McMahan and Matthew Streeter. Adaptive bound optimization for online convex optimization. In *Proceedings of the annual conference on Computational Learning Theory (COLT)*, page 244, 2010.
- [20] Arkadi Nemirovski. 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, 2004.
- [21] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, 2006.
- [22] Shai Shalev-Shwartz and Sham M Kakade. Mind the duality gap: Logarithmic regret algorithms for online optimization. In *Advances in Neural Information Processing Systems*, pages 1457–1464, 2009.
- [23] Huahua Wang and Arindam Banerjee. Online alternating direction method. Technical report, preprint arXiv:1306.3721, 2013.
- [24] Zhixiang Xu, Kilian Weinberger, and Olivier Chapelle. The greedy miser: Learning under test-time budgets. Technical report, preprint arXiv:1206.6451, 2012.
- [25] Ron Zass and Amnon Shashua. Doubly stochastic normalization for spectral clustering. In *Advances in Neural Information Processing Systems*, pages 1569–1576, 2006.
- [26] Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the International Conference on Machine Learning (ICML)*, 2003.
- [27] H. Zou and T. Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B*, 67(2):301–320, 2005.