Follow-the-Regularized-Leader with Adversarial Constraints

Ricardo N. Ferreira[1] and Cláudia Soares[1]

[1]NOVA School of Science and Technology

Abstract

Constrained Online Convex Optimization (COCO) can be seen as a generalization of the standard Online Convex Optimization (OCO) framework. At each round, a cost function and constraint function are revealed after a learner chooses an action. The goal is to minimize both the regret and cumulative constraint violation (CCV) against an adaptive adversary. We show for the first time that is possible to obtain the optimal $O(\sqrt{T})$ bound on both regret and CCV, improving the best known bounds of $O\left(\sqrt{T}\right)$ and $\tilde{O}\left(\sqrt{T}\right)$ for the regret and CCV, respectively.

1 Introduction

Consider a game where at each iteration $t \in \{1, \ldots, T\}$, an algorithm $\mathcal{A}$ has to make a decision and only after committing to that decision, a cost function f_t is revealed. The learner incurs the loss corresponding to his decision. This is the basic idea behind Online Learning. So, an algorithm $\mathcal{A}$ in Online Learning maps a game history to a decision in the decision set $\mathcal{K}$, i.e.,

Ricardo Ferreira, and Cláudia Soares are with NOVA School of Science and Technology (e-mail: rjn.ferreira@campus.fct.unl.pt, claudia.soares@fct.unl.pt).

This work is supported by NOVA LINCS (UIDB/04516/2020) with the financial support of FCT.IP. This research was carried out under Project "Artificial Intelligence Fights Space Debris" Nº C626449889-0046305 co-funded by Recovery and Resilience Plan and NextGeneration EU Funds, www.recuperarportugal.gov.pt.

$$x_t = \mathcal{A}\left(f_1, \ldots, f_{t-1}\right) \in \mathcal{K}.$$

So, we want our algorithm to minimize the cumulative cost of its decisions, especially in comparison with the best fixed decision in hindsight, $x^* = \arg\min_{x \in \mathcal{K}} \sum_{t=1}^{T} f_t(x)$. Thus, is used the metric called regret, which is defined as

$$Regret_T(\mathcal{A}) = \sum_{t=1}^{T} f_t(x_t) - \min_{x \in \mathcal{K}} \sum_{t=1}^{T} f_t(x). \tag{1}$$

Particularly, we want the regret to grow sublinearly. This means that, as the number of rounds increases, the difference between the average loss of the algorithm and the average loss of the best decision in hindsight tends to zero.

Online Convex Optimization is a special case of Online Learning, where we consider the cost functions f_t to be convex. The standard OCO framework has been extensively studied over the years [1, 2, 3, 4, 5, 6, 7, 8].

However, in real scenarios, most decisions depend on operational constraints that vary in time. Similarly, as it happens with the cost functions, we might need to commit to a decision before knowing what the constraints are. Thus, additionally to the static decision set $\mathcal{K}$, there can be some constraints in the form of $g_t(x) \leq 0$, such that g_t are convex functions $\forall t$. This framework is called Constrained Online Convex Optimization (or COCO, for short). In the literature, it is usually considered two scenarios: fixed constraints (i.e., $g_t(x) = g(x)$, $\forall t$) and adversarial constraints (where they can change at each round). In this paper, we focus on the latter. In particular, we focus on the problem with hard constraints, i.e., we do not assume that the decisions in some rounds can compensate for some constraint violations in other rounds [9, 10, 11, 12]. Contrarily, we resort to the metric of hard *cumulative constraint violation*:

$$CCV(T) \coloneqq \sum_{t=1}^{T} g_t^+(x_t), \tag{2}$$

such that $(\cdot)^+ = \max(\cdot, 0)$. In addition to bound the regret to grow sublinearly with the number of rounds, as in the standard OCO framework, we also want the cumulative constraint violation (or CCV, for short) to grow sublinearly with the number of rounds. Thus, the best decision in hindsight is the solution to the following optimization problem

$$
\begin{aligned}
&\underset{x \in \mathcal{K}}{\text{minimize}} \quad \sum_{t=1}^{T} f_t(x) \\
&\text{subject to} \quad g_t(x) \leq 0, \quad \text{for } t = 1, \ldots, T,
\end{aligned} \tag{3}
$$

where it is assumed that there is a fixed feasible decision that satisfies the constraints at every round. The regret is computed by comparing its cumulative cost against the cumulative cost of a fixed feasible action $x^* \in \mathcal{K}$, that satisfies all constraints g_t, for $t = 1, \ldots, T$. Thus, the regret can be rewritten as

$$
Regret_T(\mathcal{A}) = \sum_{t=1}^{T} f_t(x_t) - \sum_{t=1}^{T} f_t(x^*). \tag{4}
$$

such that $x^* \in \mathcal{K}^* := \{x \in \mathcal{K} \mid g_t(x) \leq 0, \forall t = 1, \ldots, T\}$.

1.1 Main Contributions

As stated before, in this work, we tackle the case of Constrained Online Convex Optimization (COCO), considering adversarial cost functions and adversarial constraint functions. We demonstrate that the Follow-the-Regularized-Leader algorithm applied to a surrogate function of the constrained problem attains optimal $O\left(\sqrt{T}\right)$ regret and CCV bounds, improving the best known bounds of $O\left(\sqrt{T}\right)$ and $\tilde{O}\left(\sqrt{T}\right)$ for the regret and CCV, respectively. As far as we know, this is the first work to attain optimal bounds for the COCO framework without additional assumptions, other than the convexity and Lipschitz continuity of the cost and constraint functions.

2 Related Work

Constrained Online Convex Optimization (COCO) can be seen as a generalization of the standard OCO framework, where we not only consider time-varying cost functions, as well as time-varying constraint functions, which are unknown to the decision-maker at the time of decision. The goal is to simultaneously bound the regret as well as the cumulative constraint violation. First works to tackle the problem of Constrained Online Convex Optimization have considered the case where the decisions in some rounds can compensate for some constraint violations in other rounds [9, 10, 11, 12], thus the goal is to bound the (soft) cumulative constraint violation defined as $\sum_{t=1}^{T} g_t(x_t)$.

However, for safety-critical applications, one may want to bound the "instantaneous" constraint violations, without assuming that these can be compensated by decisions at other times. Thus, in recent years, the scientific community has considered a stronger metric denominated hard *cumulative constraint violation*:

$$CCV(T) \coloneqq \sum_{t=1}^{T} g_t^+(x_t), \tag{5}$$

such that $(\cdot)^+ = \max(\cdot, 0)$. Additionally, in this work, we consider the most difficult setting of COCO, which considers time-varying constraints. This setting of adversarial time-varying constraints considering the hard cumulative constraint violation has been recently explored by the scientific community. A summary of the works developed in this area can be found in Table 1.

Table 1: Summary of the results on COCO for adversarial time-varying convex constraints and convex cost functions. "Conv-OPT" refers to solving a constrained convex optimization problem on each round. "Proj" refers to the Euclidean projection operation on the convex set $\mathcal{K}$.The term $\xi(T)$ was shown to be a worst-case complexity of $O(T)$, $c \in (0,1)$, and $\mathcal{V}$ denotes the distance between consecutively revealed constraint sets.

Method	Regret	CCV	Complexity
[13]	$O\left(T^{\max\{c,1-c\}}\right)$	$O\left(T^{1-c/2}\right)$	Proj
[14]	$O\left(\sqrt{T}\right)$	$O\left(T^{\frac{3}{4}}\right)$	Conv-OPT
[15]	$O\left(\sqrt{T}\right)$	$\tilde{O}\left(\sqrt{T}\right)$	Proj
[16]	$O\left(\xi(T)\right)$	$\tilde{O}\left(\xi(T)\right)$	Proj
[17]	$\tilde{O}\left(T^{\frac{3}{4}}\right)$	$O\left(T^{\frac{7}{8}}\right)$	Conv-OPT
[18]	$\tilde{O}\left(T^{\frac{3}{4}}\right)$	$\tilde{O}\left(T^{\frac{3}{4}}\right)$	Conv-OPT
[19]	$O\left(T^{\frac{3}{4}}\right)$	$\tilde{O}\left(T^{\frac{3}{4}}\right)$	Conv-OPT
[20]	$O\left(\sqrt{T}\right)$	$\tilde{O}\left(\sqrt{T}\right)$	Conv-OPT
[21]	$O\left(\sqrt{T}\right)$	$O\left(\min\{\mathcal{V}, \sqrt{T}\log T\}\right)$	Proj
Our Theorem 2	$O\left(\sqrt{T}\right)$	$O\left(\sqrt{T}\right)$	Conv-OPT

The basic assumptions in the COCO setting are the convexity and Lipschitz continuity of the cost and constraint functions. Following these assumptions, Yi et al. present a primal-dual algorithm, where, at each iteration, the authors consider a regularized version of the Lagrangian of an optimization problem considering the cost and constraint function revealed at that iteration. The authors present a regret and CCV bound of $O\left(T^{\max\{c,1-c\}}\right)$ and $O\left(T^{1-c/2}\right)$, respectively, which are dependent on a trade-off parameter $c \in (0,1)$ [13]. Guo et al. present the Rectified Online Optimization (RECOO) algorithm [14], which is based on a regularized approximation of Lagrangian. The authors resort to a first-order approximation of the revealed cost function and impose a minimum penalty price on the revealed constraint function. The authors demonstrate that the RECOO algorithm attains a regret and CCV bound of $O\left(\sqrt{T}\right)$ and $O\left(T^{\frac{3}{4}}\right)$, respectively.

More recently, considering the goal of minimizing both the regret and the CCV, Sinha and Vaze combine the two objectives [15], based on the drift-plus-penalty framework [22], and construct the surrogate function:

$$\hat{f}_t \coloneqq V\tilde{f}_t + \Phi'(Q(t))\tilde{g}_t, \tag{6}$$

where $V > 0$ and $\Phi : \mathbb{R}_+ \to \mathbb{R}_+$ is a non-decreasing convex potential (Lyapunov) function, such that $\Phi(0) = 0$. The functions $\tilde{f}_t$ and $\tilde{g}_t$ denote the original cost function f_t and constraint function g_t^+, respectively, scaled by a factor α. Moreover, the authors also introduce the Regret Decomposition Inequality, which, from known regret bounds of policies in the standard OCO framework, allows to easily obtain the regret and CCV bounds in the COCO framework for the original loss and constraint functions (in Section 3 we present a modified derivation of the Regret Decomposition Inequality). The authors show that the AdaGrad algorithm attains the optimal regret bound $O\left(\sqrt{T}\right)$ and near-optimal CCV bound $\tilde{O}\left(\sqrt{T}\right)$.

Based on the Regret Decomposition Inequality, different works have been exploring the applicability of known OCO policies for the COCO framework.

Lekeufack and Jordan present a meta-algorithm for Optimistic OCO (where the adversary is considered to be predictable) and achieve a regret and CCV bound of $O\left(\xi(T)\right)$ and $\tilde{O}\left(\xi(T)\right)$, respectively, where $\xi(T)$ was shown to be a worst-case complexity of $O(T)$ [16].

To avoid computationally intensive projections, different versions of the Adaptive Online Conditional Gradient algorithm have been explored [17, 18, 19], with the best regret and CCV bounds being $O\left(T^{\frac{3}{4}}\right)$ and $\tilde{O}\left(T^{\frac{3}{4}}\right)$, respectively. More recently, Lu et al. combine the Online Gradient Descent algorithm with adaptive step-sizes and infeasible projections via a Separation Oracle [20], and obtain the optimal regret bound $O\left(\sqrt{T}\right)$ and near-optimal CCV bound $\tilde{O}\left(\sqrt{T}\right)$.

Recently, Vaze and Sinha presents an algorithm that presents an instance dependent CCV bound, where $\mathcal{V}$ denotes the distance between consecutively revealed constraint sets, and shows that this bound is $\tilde{O}\left(\sqrt{T}\right)$ in the worst-case [21].

3 Preliminaries

Throughout the paper, we consider the decision set $\mathcal{K}$ to be a compact convex set, and denote by $\|\cdot\|_A$ the norm induced by the symmetric and positive-definite matrix A (i.e., $\|x\|_A = \sqrt{x^T A x}$). We denote the Euclidean norm by $\|\cdot\| = \|\cdot\|_I$, where I denotes the identity matrix. For simplicity, we use $\nabla f(x)$ to denote a subgradient of the function f at point x.

Our method is based on the Follow-the-Regularized-Leader (FTRL), therefore we consider a twice-differentiable, m-strongly convex and M-smooth regularization function $R : \mathbb{R}^n \to \mathbb{R}$, where $\nabla^2 R(x)$ denotes the second derivative of the regularization function R at point x. We follow the derivation used in [2], therefore we import some of the concepts:

Definition 1. *A function is M-smooth if $\forall x, y \in \mathcal{K}$,*

$$f(y) \leq f(x) + \nabla f(x)^T (y - x) + \frac{M}{2}\|y - x\|^2.$$

Definition 2. *A function is m-strongly convex if $\forall x, y \in \mathcal{K}$,*

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) + \frac{m}{2}\|y - x\|^2.$$

Additionally, we introduce a set of assumptions regarding the optimization problem:

Assumption 1. *The feasible set $\mathcal{K}$ is a compact convex set with diameter D such that $\|x - x'\| \leq D$, $\forall x, x' \in \mathcal{K}$.*

Assumption 2. *The loss functions $f_t : \mathcal{K} \to \mathbb{R}$ are convex and Lipschitz continuous with Lipschitz constant F such that $|f_t(x) - f_t(x')| \leq F\|x - x'\|$, $\forall x, x' \in \mathcal{K}$, $\forall t$.*

Assumption 3. *The constraint functions $g_t : \mathcal{K} \to \mathbb{R}$ are convex and Lipschitz continuous with Lipschitz constant G such that $|g_t(x) - g_t(x')| \leq G\|x - x'\|$, $\forall x, x' \in \mathcal{K}$, $\forall t$.*

Assumptions 1, 2 and 3 are the assumptions used in the literature, which express the convexity and Lipschitz continuity of the cost and constraint functions, as well as the convexity and compactness of the static decision set $\mathcal{K}$. From these assumptions, we can derive some results regarding the regularization function R and the local and dual local norms (defined in Definition 3).

Definition 3. *The norms $\|\cdot\|_t$ and $\|\cdot\|_t^*$ denote the local and dual local norms at iteration t, respectively. The local norm $\|\cdot\|_t$ is the norm induced by the second derivative of the regularization function at some point $z \in [x_t, x_{t+1}]$ between two consecutive decisions, i.e.,*

$$\|\cdot\|_t = \|\cdot\|_{\nabla^2 R(z)}.$$

Respectively, the dual local norm $\|\cdot\|_t^*$ is the norm induced by inverse of the second derivative of the regularization function at some point $z \in [x_t, x_{t+1}]$ between two consecutive decisions, i.e.,

$$\|\cdot\|_t^* = \|\cdot\|_{\nabla^{-2} R(z)}.$$

First, we start defining D_R, which denotes the diameter of the set $\mathcal{K}$ with respect to the function R, and prove that it is bounded.

Lemma 1. *Let D_R denote the diameter of the set $\mathcal{K}$ with respect to the function R, such that*

$$D_R = \sqrt{\max_{x,y \in \mathcal{K}} \{R(x) - R(y)\}}.$$

Then, D_R is bounded.

Proof. Since R is a real-valued convex function, then it is continuous. By Weierstrass theorem, as $\mathcal{K}$ is closed, $c = \min_{x \in \mathcal{K}} R(x)$ and $C = \max_{x \in \mathcal{K}} R(x)$ exist and are finite. Then $C - c$ is finite, and so is D_R. $\square$

Finally, we show that the dual local norms of the subgradients of the cost and constraint functions are bounded.

Lemma 2. *The dual local norm of the subgradient of the cost functions have a bound F_R, i.e.,* $\|\nabla f_t(x)\|_t^* \le F_R$, $\forall x \in \mathcal{K}$, $\forall t$.

Similarly, the dual local norm of the subgradient of the constraint functions has a bound G_R, i.e., $\|\nabla g_t(x)\|_t^* \le G_R$, $\forall x \in \mathcal{K}$, $\forall t$.

Proof. Since R is m-strongly convex, we have $\forall x \in \mathbb{R}^n$,

$$\nabla^{-2} R(x) \preceq \frac{1}{m} I.$$

Therefore, $\forall y \in \mathcal{K}$,

$$\|\nabla f_t(y)\|_{\nabla^{-2}R(z)} = \nabla f_t(y)^T \nabla^{-2} R(z) \nabla f_t(y) \le \frac{1}{m} \|\nabla f_t(y)\| \le \frac{F}{m} =: F_R,$$
$$\|\nabla g_t(y)\|_{\nabla^{-2}R(z)} = \nabla g_t(y)^T \nabla^{-2} R(z) \nabla g_t(y) \le \frac{1}{m} \|\nabla g_t(y)\| \le \frac{G}{m} =: G_R,$$

where the last inequalities derive from Assumptions 2 and 3. □

3.1 Regret Decomposition Inequality

In this section, we recapitulate the analysis of the regret and cumulative hard constraint violation, through the use of surrogate loss functions and the regret decomposition inequality. In this work, we based our derivation on the inequality presented by Wang et al. [19]. However, inspired by the work of Guo et al. [14], we impose a minimum penalty price on the constraint function g_t^+.

Let $Q(t)$ denote the CCV at iteration t, thus defined by the recursion rule $Q(t) = Q(t-1) + g_t^+(x_t)$, for all $t \ge 1$, with $Q(0) = 0$. Furthermore, define a new variable $P(t) = \max\{\rho, Q(t)\}$, such that $\rho > 0$. Now, consider a non-decreasing convex potential (Lyapunov) function $\Phi : \mathbb{R}_+ \to \mathbb{R}_+$, such that $\Phi(0) = 0$. By convexity,

$$\begin{aligned} \Phi(\beta P(t)) &\le \Phi(\beta P(t-1)) + \Phi'(\beta P(t))\beta \left[P(t) - P(t-1)\right] \\ &\le \Phi(\beta P(t-1)) + \Phi'(\beta P(t))\beta g_t^+(x_t). \end{aligned} \tag{7}$$

Thus,

$$\Phi(\beta P(t)) - \Phi(\beta P(t-1)) \le \Phi'(\beta P(t))\beta g_t^+(x_t). \tag{8}$$

From this result, inspired by the stochastic drift-plus-penalty framework [22], we construct the surrogate function:

$$\hat{f}_t \coloneqq V\beta f_t + \Phi'(\beta P(t))\beta g_t^+, \tag{9}$$

therefore, at every iteration, we impose a minimum penalty price on the constraint function, inducing a conservative decision. If we only considered $Q(t)$, as in previous works, then we could have $Q(t) = 0$, thus allowing overly optimistic decisions.

Let $x^* \in \mathcal{K}^*$ be any feasible decision of the problem in (3). Using the drift inequality and the surrogate loss functions defined in (8) and (9), respectively, and combining with the fact that $g_t(x^*) \leq 0, \forall t \geq 1$, then we have

$$\Phi(\beta P(t)) - \Phi(\beta P(t-1)) + V\beta\left(f_t(x_t) - f_t(x^*)\right) \leq \hat{f}_t(x_t) - \hat{f}_t(x^*), \qquad \forall t \geq 1.$$

By summing up from $t = 1$ to T, and remembering that $\Phi(0) = 0$, we arrive at the Regret Decomposition Inequality

$$\Phi(\beta P(T)) + V\beta Regret_T \leq Regret_T^*, \tag{10}$$

where $Regret_T \coloneqq \sum_{t=1}^T \left(f_t(x_t) - f_t(x^*)\right)$, as defined in (4), and $Regret_T^* \coloneqq \sum_{t=1}^T \left(\hat{f}_t(x_t) - \hat{f}_t(x^*)\right)$. Thus, the Regret Decomposition Inequality defined in (10) demonstrates that the Lyapunov function value of the CCV plus the regret obtained by the policy on the original function (in the left-hand side) is bounded by the regret obtained by the policy on the surrogate loss functions (in the right-hand side). Thus, by applying known policies on the surrogate loss functions, from their known regret bound in the standard OCO framework, we can easily obtain the regret and CCV bounds in the COCO framework for the original loss and constraint functions.

4 Follow-the-Regularized-Leader

In this section, we revisit the Follow-the-Regularized-Leader algorithm [2], which as been used to derive many different algorithms in the standard OCO framework. In this section, we analyze the regret of the Follow-the-Regularized-Leader (FTRL) algorithm in the COCO framework. In particular, we will analyze the regret bounds of the FTRL for the surrogate loss functions $\hat{f}_t$, and derive the regret and CCV bounds on the original functions by resorting to the Regret Decomposition Inequality in (10).

The following theorem presents a general regret bound of the RFTL algorithm for a set of convex functions $\hat{f}_t$ and a twice-differentiable, smooth, strongly-convex regularization function $R(x)$.

Algorithm 1 Follow-the-Regularized-Leader (FTRL)

Require: $T > 0$, $\rho, \eta, \beta, V > 0$, compact convex set $\mathcal{K}$, twice-differentiable, m-strongly convex and M-smooth regularization function $R(x)$.

$Q(0) = 0$

$x_1 = \arg\min_{x \in \mathcal{K}} \{R(x)\}$

for $t = 1, \dots, T$ **do**

Choose x_t and observe f_t, g_t.

$Q(t) = Q(t-1) + g_t^+(x_t)$

$P(t) = \max\{\rho, Q(t)\}$

$\hat{f}_t \coloneqq V\beta f_t + \Phi'(\beta P(t))\beta g_t^+$

Follow-the-Regularized-Leader update:

$$x_{t+1} = \arg\min_{x \in \mathcal{K}} \{\eta \sum_{s=1}^{t} \nabla \hat{f}_s(x_s)^T x + R(x)\}$$

end for

Theorem 1. *The FTRL algorithm attains for every comparator $u \in \mathcal{K}$ the following bound on the regret:*

$$Regret_T(FTRL) \leq 2\eta \sum_{t=1}^{T} \|\nabla \hat{f}_t(x_t)\|_t^{*2} + \frac{D_R^2}{\eta}. \tag{11}$$

A proof of this theorem can be found in [2]. So, resorting to the Regret Decomposition Inequality, we can derive the regret and CCV bounds on the original functions. First, we must bound the dual local norm of the subgradient of the surrogate functions $\hat{f}_t$.

Lemma 3. *The dual local norms of the subgradient of the surrogate functions $\hat{f}_t$ are bounded as follows*

$$\|\nabla \hat{f}_t(x_t)\|_t^* \leq V\beta F_R + \Phi'(\beta P(t))\beta G_R. \tag{12}$$

Proof. We can use the triangle inequality and homogeneity properties of the norm to obtain

$$\|\nabla \hat{f}_t(x_t)\|_t^* \leq V\beta \|\nabla f_t(x_t)\|_t^* + \Phi'(\beta P(t))\beta \|\nabla g_t^+(x_t)\|_t^*,$$

since $V, \beta > 0$ and Φ is a non-decreasing function, therefore $\Phi'(x) \geq 0, \forall x \geq 0$. By Lemma 2, we arrive at the desired result

$$\|\nabla \hat{f}_t(x_t)\|_t^* \leq V\beta F_R + \Phi'(\beta P(t))\beta G_R.$$

□

From this result, let us state two more lemmas that bound the regret and CCV on the original functions.

Lemma 4. *Let $\eta = \frac{1}{32\beta^2 G_R^2 T}$ and $\Phi(x) = x^2$. Then, the regret on the original functions can be bounded as*

$$Regret_T \leq \frac{F_R^2}{8G_R^2}\frac{V}{\beta} + \frac{32D_R^2 G_R^2 \beta T}{V}. \tag{13}$$

Proof. By combining Lemma 3 with Theorem 1, we can bound the regret on the surrogate loss functions as

$$\begin{aligned} Regret_T^* &= Regret_T(FTRL) \\ &\leq 2\eta \sum_{t=1}^{T} \big(V\beta F_R + \Phi'(\beta P(t))\beta G_R\big)^2 + \frac{D_R^2}{\eta} \\ &\leq 4\eta V^2 \beta^2 F_R^2 T + 4\eta\, \Phi'(\beta P(T))^2 \beta^2 G_R^2 T + \frac{D_R^2}{\eta}, \end{aligned} \tag{14}$$

where the last inequality results from the algebraic inequality $(a+b)^2 \leq 2(a^2+b^2)$, and the fact that $P(t)$ and the derivative of Φ are non-decreasing, therefore $\sum_{t=1}^T \Phi'(\beta P(t))\beta G_R \leq \sum_{t=1}^T \Phi'(\beta P(T))\beta G_R$. Applying the result in (14) in the Regret Decomposition Inequality in (10), we obtain

$$\Phi(\beta P(T)) + V\beta Regret_T \leq 4\eta V^2 \beta^2 F_R^2 T + 4\eta\, \Phi'(\beta P(T))^2 \beta^2 G_R^2 T + \frac{D_R^2}{\eta}. \tag{15}$$

Now, by considering $\Phi(x) = x^2$, and consequently $\Phi'(x) = 2x$, we obtain

$$\begin{aligned} &\beta^2 P(T)^2 + V\beta Regret_T \leq 4\eta V^2\beta^2 F_R^2 T + 16\eta\, \beta^2 P(T)^2 \beta^2 G_R^2 T + \frac{D_R^2}{\eta} \iff \\ &V\beta Regret_T \leq 4\eta V^2 \beta^2 F_R^2 T + \beta^2 P(T)^2 \left(16\eta\beta^2 G_R^2 T - 1\right) + \frac{D_R^2}{\eta}. \end{aligned} \tag{16}$$

By letting $\eta = \frac{1}{32\beta^2 G_R^2 T}$, we have that $16\eta\beta^2 G_R^2 T - 1 < 0$, thus we can further simplify and arrive at the desired result

$$Regret_T \leq \frac{F_R^2}{8G_R^2}\frac{V}{\beta} + \frac{32 D_R^2 G_R^2 \beta\, T}{V}. \tag{17}$$

$\square$

Lemma 5. *Let* $\eta = \frac{1}{32\beta^2 G_R^2 T}$, $\rho = 2\sqrt{FD\frac{V}{\beta}T}$ *and* $\Phi(x) = x^2$. *Then, the CCV on the original functions can be bounded as*

$$Q(T) \leq \frac{F_R}{\sqrt{2}G_R}\frac{V}{\beta} + \sqrt{128}\, D_R G_R \sqrt{T}. \tag{18}$$

Proof. Starting from the last inequality in (16), we have

$$Regret_T \leq 4\eta V \beta F_R^2 T + \frac{\beta}{V} P(T)^2 \left(16\eta\beta^2 G_R^2 T - 1\right) + \frac{D_R^2}{V\beta\,\eta}. \tag{19}$$

Trivially, we have $Regret_T \geq -FDT$. Thus, combining this result with the inequality in (19), we have

$$\frac{\beta}{V} P(T)^2 \left(1 - 16\eta\beta^2 G_R^2 T\right) \leq 4\eta V \beta F_R^2 T + FDT + \frac{D_R^2}{V\beta\,\eta}.$$

As in Lemma 4, let $\eta = \frac{1}{32\beta^2 G_R^2 T}$. Thus, we obtain

$$\begin{aligned} &\frac{\beta}{2V} P(T)^2 \leq \frac{F_R^2}{8G_R^2}\frac{V}{\beta} + FDT + \frac{32 D_R^2 G_R^2 \beta\, T}{V} \iff \\ &P(T)^2 \leq \frac{F_R^2}{4G_R^2}\frac{V^2}{\beta^2} + 2FD\frac{V}{\beta}T + 64 D_R^2 G_R^2 T. \end{aligned}$$

Remember that $P(T) = \max\{\rho, Q(T)\}$. By the inequality $\sqrt{a^2+b^2} \leq \sqrt{2}\max\{a,b\}$, we have $\sqrt{\frac{\rho^2 + Q(T)^2}{2}} \leq P(T)$, therefore

$$\rho^2 + Q(T)^2 \leq \frac{F_R^2}{2G_R^2}\frac{V^2}{\beta^2} + 4FD\frac{V}{\beta}T + 128 D_R^2 G_R^2 T.$$

Let $\rho = 2\sqrt{FD\frac{V}{\beta}T}$, and we have

$$Q(T)^2 \leq \frac{F_R^2}{2G_R^2}\frac{V^2}{\beta^2} + 128 D_R^2 G_R^2 T.$$

Since $Q(T) \geq 0$ and by the algebraic inequality $\sqrt{a+b} \leq \sqrt{a}+\sqrt{b}$, we obtain

$$Q(T) \leq \frac{F_R}{\sqrt{2}G_R}\frac{V}{\beta} + \sqrt{128}\, D_R G_R \sqrt{T}.$$

□

With these results, we are now ready to prove that the Follow-the-Regularized-Leader applied to the surrogate loss functions is able to obtain optimal regret and CCV bounds.

Theorem 2. *Let $V = T^{\frac{3}{4}}$, $\beta = T^{\frac{1}{4}}$, $\rho = 2\sqrt{FD\frac{V}{\beta}T}$ and $\eta = \frac{1}{32\beta^2 G_R^2 T}$. Consider the Lyapunov convex function $\Phi(x) = x^2$. Then, Algorithm 1 achieves the following regret and CCV bounds*

$$\begin{aligned} Regret_T &\leq \left(\frac{F_R^2}{8G_R^2} + 32 D_R^2 G_R^2\right) T^{\frac{1}{2}} \\ CCV &\leq \left(\frac{F_R}{\sqrt{2}G_R} + \sqrt{128}\, D_R G_R\right) T^{\frac{1}{2}}. \end{aligned}$$

Proof. From the results of Lemmas 4 and 5, by using $V = T^{\frac{3}{4}}$ and $\beta = T^{\frac{1}{4}}$, after some manipulation we arrive at the result. □

5 Conclusions

In this work, we achieve for the first time simultaneous optimal bounds on the regret and (hard) cumulative constraint violation, for the COCO framework without additional assumptions, other than the convexity of the functions and Lipschitz continuity of the cost and constraint functions. We demonstrate that the Follow-the-Regularized-Leader algorithm applied to a surrogate function that combines both the loss and constraint function at each iteration, is capable of obtaining these optimal bounds.

We recognize that the Follow-the-Regularized-Leader update can be complex and computationally expensive as it is necessary to solve an optimization problem at each round. However, we show that it is possible to obtain optimal bounds on the regret and CCV, thus opening new opportunities to explore more efficient algorithms capable of obtaining the optimal bounds on the COCO framework.

References

[1] Shai Shalev-Shwartz et al. Online learning and online convex optimization. *Foundations and Trends® in Machine Learning*, 4(2):107–194, 2012.

[2] Elad Hazan et al. Introduction to online convex optimization. *Foundations and Trends® in Optimization*, 2(3-4):157–325, 2016.

[3] Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th international conference on machine learning (icml-03)*, pages 928–936, 2003.

[4] John Duchi, Elad Hazan, and Yoram Singer. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12(null):2121–2159, July 2011. ISSN 1532-4435.

[5] Sébastien Bubeck, Nicolo Cesa-Bianchi, et al. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends® in Machine Learning*, 5(1):1–122, 2012.

[6] Elad Hazan and Satyen Kale. Projection-free online learning. In *Proceedings of the 29th International Coference on International Conference on Machine Learning*, ICML'12, page 1843–1850, Madison, WI, USA, 2012. Omnipress. ISBN 9781450312851.

[7] Elad Hazan and Edgar Minasyan. Faster projection-free online learning. In *Conference on Learning Theory*, pages 1877–1893. PMLR, 2020.

[8] Elad Hazan and Karan Singh. Boosting for online convex optimization. In *International Conference on Machine Learning*, pages 4140–4149. PMLR, 2021.

[9] 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, September 2012. ISSN 1532-4435.

[10] Rodolphe Jenatton, Jim Huang, and Cedric Archambeau. Adaptive algorithms for online convex optimization with long-term constraints. In Maria Florina Balcan and Kilian Q. Weinberger, editors, *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 402–411, New York, New York, USA, 20–22 Jun 2016. PMLR. URL `https://proceedings.mlr.press/v48/jenatton16.html`.

[11] Jianjun Yuan and Andrew Lamperski. Online convex optimization for cumulative constraints. *Advances in Neural Information Processing Systems*, 31, 2018.

[12] Hao Yu and Michael J. Neely. A low complexity algorithm with o(â^št) regret and o(1) constraint violations for online convex optimization with long term constraints. *Journal of Machine Learning Research*, 21(1):1–24, 2020. URL `http://jmlr.org/papers/v21/16-494.html`.

[13] Xinlei Yi, Xiuxian Li, Tao Yang, Lihua Xie, Tianyou Chai, and Karl Henrik Johansson. Regret and cumulative constraint violation analysis for distributed online constrained convex optimization. *IEEE Transactions on Automatic Control*, 68(5):2875–2890, 2023. doi: 10.1109/TAC.2022.3230766.

[14] Hengquan Guo, Xin Liu, Honghao Wei, and Lei Ying. Online convex optimization with hard constraints: Towards the best of two worlds and beyond. In *Advances in Neural Information Processing Systems*, 2022. URL `https://openreview.net/forum?id=rwdpFgfVpvN`.

[15] Abhishek Sinha and Rahul Vaze. Tight bounds for online convex optimization with adversarial constraints. In *ICML 2024 Workshop: Foundations of Reinforcement Learning and Control – Connections and Perspectives*, 2024. URL `https://openreview.net/forum?id=old9jd3TUu`.

[16] Jordan Lekeufack and Michael I Jordan. An optimistic algorithm for online convex optimization with adversarial constraints. *arXiv preprint arXiv:2412.08060*, 2024.

[17] Dan Garber and Ben Kretzu. Projection-free online convex optimization with time-varying constraints. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 14988–15005. PMLR, 21–27 Jul 2024. URL `https://proceedings.mlr.press/v235/garber24a.html`.

[18] Dhruv Sarkar, Aprameyo Chakrabartty, Subhamon Supantha, Palash Dey, and Abhishek Sinha. Projection-free algorithms for online convex optimization with adversarial constraints. *arXiv preprint arXiv:2501.16919*, 2025.

[19] Yibo Wang, Yuanyu Wan, and Lijun Zhang. Revisiting projection-free online learning with time-varying constraints. *arXiv preprint arXiv:2501.16046*, 2025.

[20] Yiyang Lu, Mohammad Pedramfar, and Vaneet Aggarwal. Order-optimal projection-free algorithm for adversarially constrained online convex optimization. *arXiv preprint arXiv:2502.16744*, 2025.

[21] Rahul Vaze and Abhishek Sinha. $O(\sqrt{T})$ static regret and instance dependent constraint violation for constrained online convex optimization. *arXiv preprint arXiv:2502.05019*, 2025.

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