

# Regret bounded by gradual variation for online convex optimization

Tianbao Yang · Mehrdad Mahdavi ·  
Rong Jin · Shenghuo Zhu

Received: date / Accepted: date

**Abstract** Recently, it has been shown that the regret of the Follow the Regularized Leader (FTRL) algorithm for online linear optimization can be bounded by the total variation of the cost vectors rather than the number of rounds. In this paper, we extend this result to general online convex optimization. In particular, this resolves an open problem that has been posed in a number of recent papers. We first analyze the limitations of the FTRL algorithm in [16] when applied to online convex optimization, and extend the definition of variation to a gradual variation which is shown to be a lower bound of the total variation. We then present two novel algorithms that bound the regret by the gradual variation of cost functions. Unlike previous approaches that maintain a single sequence of solutions, the proposed algorithms maintain two sequences of solutions that make it possible to achieve a variation-based regret bound for online convex optimization.

To establish the main results, we discuss a lower bound for FTRL that maintains only one sequence of solutions, and a necessary condition on smoothness of the cost functions for obtaining a gradual variation bound. We extend the main results three-fold: (i) we present a general method to obtain a gradual variation bound measured by general norm; (ii) we extend algorithms to a class of online non-smooth optimization with gradual variation bound; and

---

Tianbao Yang  
NEC Laboratories America  
E-mail: tyang@nec-labs.com  
(The most part of the work was done when the author was at Michigan State University)

Mehrdad Mahdavi and Rong Jin  
Department of Computer Science and Engineering  
Michigan State University  
E-mail: mahdavim@cse.msu.edu, rongjin@cse.msu.edu

Shenghuo Zhu  
NEC Laboratories America  
E-mail: zsh@nec-labs.com

(iii) we develop a deterministic algorithm for online bandit optimization in multipoint bandit setting.

**Keywords** online convex optimization · regret bound · gradual variation · bandit

## 1 Introduction

We consider the general online convex optimization problem [26] which proceeds in trials. At each trial, the learner is asked to predict the decision vector  $\mathbf{x}_t$  that belongs to a bounded closed convex set  $\mathcal{P} \subseteq \mathbb{R}^d$ ; it then receives a cost function  $c_t(\cdot) : \mathcal{P} \rightarrow \mathbb{R}$  and incurs a cost of  $c_t(\mathbf{x}_t)$  for the submitted solution. The goal of online convex optimization is to come up with a sequence of solutions  $\mathbf{x}_1, \dots, \mathbf{x}_T$  that minimizes the regret, which is defined as the difference in the cost of the sequence of decisions accumulated up to the trial  $T$  made by the learner and the cost of the best fixed decision in hindsight, i.e.

$$\text{Regret}_T = \sum_{t=1}^T c_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T c_t(\mathbf{x}).$$

In the special case of linear cost functions, i.e.  $c_t(\mathbf{x}) = \mathbf{f}_t^\top \mathbf{x}$ , the problem becomes online linear optimization. The goal of online convex optimization is to design algorithms that predict, with a small regret, the solution  $\mathbf{x}_t$  at the  $t$ th trial given the (partial) knowledge about the past cost functions  $c_\tau(\cdot), \tau = 1, \dots, t-1$ .

Over the past decade, many algorithms have been proposed for online convex optimization, especially for online linear optimization. As the first seminal paper in online convex optimization, Zinkevich [26] proposed a gradient descent algorithm with a regret bound of  $O(\sqrt{T})$ . When cost functions are strongly convex, the regret bound of the online gradient descent algorithm is reduced to  $O(\log T)$  with appropriately chosen step size [13]. Another common methodology for online convex optimization, especially for online linear optimization, is based on the framework of Follow the Leader (FTL) [17]. FTL chooses  $\mathbf{x}_t$  by minimizing the cost incurred by  $\mathbf{x}_t$  in all previous trials. Since the naive FTL algorithm fails to achieve sublinear regret in the worst case, many variants have been developed to fix the problem, including Follow the Perturbed Leader (FTPL) [17], Follow the Regularized Leader (FTRL) [1], and Follow the Approximate Leader (FTAL) [13]. Other methodologies for online convex optimization introduce a potential function (or link function) to map solutions between the space of primal variables and the space of dual variables, and carry out primal-dual update based on the potential function. The well-known Exponentiated Gradient (EG) algorithm [20] or multiplicative weights algorithm [12] belong to this category. We note that these different algorithms are closely related. For example, in online linear optimization, the potential-based primal-dual algorithm is equivalent to FTRL algorithm [16].

Generally, most previous studies of online convex optimization bound the regret in terms of the number of trials  $T$ . However, it is expected that the regret should be low in an unchanging environment or when the cost functions are somehow correlated. Specifically, the tightest rate for the regret should depend on the variance of the sequence of cost functions rather than the number of rounds  $T$ . An algorithm with regret in terms of variation of the cost functions is expected to perform better when the cost sequence has low variation. Therefore, it is of great interest to derive a variation-based regret bound for online convex optimization in an adversarial setting.

As it has been established as a fact that the regret of a natural algorithm in a stochastic setting, i.e. the cost functions are generated by a stationary stochastic process, can be bounded by the total variation in the cost functions [16], devising algorithm in fully adversarial setting, i.e. the cost functions are chosen completely adversarially, is posed as an open problem in [5]. The concern is whether or not it is possible to derive a regret bound for an online optimization algorithm by the variation of the observed cost functions. Recently [16] made a substantial progress in this route and proved a variation-based regret bound for online *linear* optimization by tight analysis of FTRL algorithm with an appropriately chosen step size. A similar regret bound is shown in the same paper for prediction from expert advice by slightly modifying the multiplicative weights algorithm. In this work, we aim to take one step further and contribute to this research direction by developing algorithms for general framework of online *convex* optimization with variation-based regret bounds.

A preliminary version of this paper appeared at the Conference on Learning Theory (COLT) [8] as the result of a merge between two papers as discussed in a commentary written by Satyen Kale [18]. We highlight the differences between this paper and [8]. We first motivate the definition of a gradual variation by showing that the total variation bound in [16] is not necessary small when the cost functions change slowly, and the gradual variation is smaller than the total variation. We then extend FTRL to achieve a regret bounded by the gradual variation in Subsection 2.1. The second algorithm in Subsection 2.2 is similar to the one that appeared in [8]. Sections 3 and 4 contain generalizations of the second algorithm to non-smooth functions and a bandit setting, respectively, which appear in this paper for the first time.

In the remaining of this section, we first present the results from [16] for online linear optimization and discuss their potential limitations when applied to online convex optimization which motivates our work.

## 1.1 Online Linear Optimization

Many decision problems can be cast into online linear optimization problems, such as prediction from expert advice [6], online shortest path problem [25]. [16] proved the first variation-based regret bound for online linear optimization problems in an adversarial setting. Hazan and Kale's algorithm for online

**Algorithm 1** Follow The Regularized Leader for Online Linear Optimization

- 
- ```

1: Input:  $\eta > 0$ 
2: for  $t = 1, \dots, T$  do
3:   If  $t = 1$ , predict  $\mathbf{x}_t = 0$ 
4:   If  $t > 1$ , predict  $\mathbf{x}_t$  by

```

$$\mathbf{x}_t = \arg \min_{\mathbf{x} \in \mathcal{P}} \sum_{\tau=1}^{t-1} \mathbf{f}_\tau^\top \mathbf{x} + \frac{1}{2\eta} \|\mathbf{x}\|_2^2$$

- ```

5:   Receive cost vector  $\mathbf{f}_t$  and incur loss  $\mathbf{f}_t^\top \mathbf{x}_t$ 
6: end for

```
- 

linear optimization is based on the framework of FTRL. For completeness, the algorithm is shown in Algorithm 1. At each trial, the decision vector  $\mathbf{x}_t$  is given by solving the following optimization problem:

$$\mathbf{x}_t = \arg \min_{\mathbf{x} \in \mathcal{P}} \sum_{\tau=1}^{t-1} \mathbf{f}_\tau^\top \mathbf{x} + \frac{1}{2\eta} \|\mathbf{x}\|_2^2,$$

where  $\mathbf{f}_t$  is the cost vector received at trial  $t$  after predicting the decision  $\mathbf{x}_t$ , and  $\eta$  is a step size. They bound the regret by the variation of cost vectors defined as

$$\text{VAR}_T = \sum_{t=1}^T \|\mathbf{f}_t - \mu\|_2^2, \quad (1)$$

where  $\mu = 1/T \sum_{t=1}^T \mathbf{f}_t$ . By assuming  $\|\mathbf{f}_t\|_2 \leq 1, \forall t$  and setting the value of  $\eta$  to  $\eta = \min(2/\sqrt{\text{VAR}_T}, 1/6)$ , they showed that the regret of Algorithm 1 can be bounded by

$$\sum_{t=1}^T \mathbf{f}_t^\top \mathbf{x}_t - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T \mathbf{f}_t^\top \mathbf{x} \leq \begin{cases} 15\sqrt{\text{VAR}_T} & \text{if } \sqrt{\text{VAR}_T} \geq 12 \\ 150 & \text{if } \sqrt{\text{VAR}_T} \leq 12 \end{cases}. \quad (2)$$

From (2), we can see that when the variation of the cost vectors is small (less than 12), the regret is a constant, otherwise it is bounded by the variation  $O(\sqrt{\text{VAR}_T})$ . This result indicates that online linear optimization in the adversarial setting is as efficient as in the stationary stochastic setting.

## 1.2 Online Convex Optimization

Online convex optimization generalizes online linear optimization by replacing linear cost functions with non-linear convex cost functions. It has found applications in several domains, including portfolio management [3] and online classification [19]. For example, in online portfolio management problem, an investor wants to distribute his wealth over a set of stocks without knowing the market output in advance. If we let  $\mathbf{x}_t$  denote the distribution on the

---

**Algorithm 2** Follow The Regularized Leader (FTRL) for Online Convex Optimization
 

---

- 1: **Input:**  $\eta > 0$
- 2: **for**  $t = 1, \dots, T$  **do**
- 3:   If  $t = 1$ , Predict  $\mathbf{x}_t = 0$
- 4:   If  $t > 1$ , Predict  $\mathbf{x}_t$  by

$$\mathbf{x}_t = \arg \min_{\mathbf{x} \in \mathcal{P}} \sum_{\tau=1}^{t-1} \mathbf{f}_\tau^\top \mathbf{x} + \frac{1}{2\eta} \|\mathbf{x}\|_2^2$$

- 5:   Receive cost function  $c_t(\cdot)$  and incur loss  $c_t(\mathbf{x}_t)$
  - 6:   Compute  $\mathbf{f}_t = \nabla c_t(\mathbf{x}_t)$
  - 7: **end for**
- 

stocks and  $\mathbf{r}_t$  denote the price relative vector, i.e.  $r_t[i]$  denote the ratio of the closing price of stock  $i$  on day  $t$  to the closing price on day  $t-1$ , then an interesting function is the logarithmic growth ratio, i.e.  $\sum_{t=1}^T \log(\mathbf{x}_t^\top \mathbf{r}_t)$ , which is a concave function to be maximized. Similar to [16], we aim to develop algorithms for online convex optimization with regret bounded by the variation in the cost functions.

Before presenting our algorithms, we first show that directly applying the FTRL algorithm to general online convex optimization may not be able to achieve the desired result. To extend FTRL for online convex optimization, a straightforward approach is to use the first order approximation for convex cost function, i.e.,  $c_t(\mathbf{x}) \simeq c_t(\mathbf{x}_t) + \nabla c_t(\mathbf{x}_t)^\top (\mathbf{x} - \mathbf{x}_t)$ , and replace the cost vector  $\mathbf{f}_t$  in Algorithm 1 with the gradient of the cost function  $c_t(\cdot)$  at  $\mathbf{x}_t$ , i.e.,  $\mathbf{f}_t = \nabla c_t(\mathbf{x}_t)$ . The resulting algorithm is shown in Algorithm 2. Using the convexity of  $c_t(\cdot)$ , we have

$$\sum_{t=1}^T c_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T c_t(\mathbf{x}) \leq \sum_{t=1}^T \mathbf{f}_t^\top \mathbf{x}_t - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T \mathbf{f}_t^\top \mathbf{x}. \quad (3)$$

If we assume  $\|\nabla c_t(\mathbf{x})\|_2 \leq 1, \forall t, \forall \mathbf{x} \in \mathcal{P}$ , we can apply Hazan and Kale's variation-based bound in (2) to bound the regret in (3) by the variation

$$\text{VAR}_T = \sum_{t=1}^T \|\mathbf{f}_t - \mu\|_2^2 = \sum_{t=1}^T \left\| \nabla c_t(\mathbf{x}_t) - \frac{1}{T} \sum_{\tau=1}^T \nabla c_\tau(\mathbf{x}_\tau) \right\|_2^2. \quad (4)$$

To better understand  $\text{VAR}_T$  in (4), we rewrite  $\text{VAR}_T$  as

$$\begin{aligned} \text{VAR}_T &= \sum_{t=1}^T \left\| \nabla c_t(\mathbf{x}_t) - \frac{1}{T} \sum_{\tau=1}^T \nabla c_\tau(\mathbf{x}_\tau) \right\|_2^2 = \frac{1}{2T} \sum_{t,\tau=1}^T \|\nabla c_t(\mathbf{x}_t) - \nabla c_\tau(\mathbf{x}_\tau)\|_2^2 \\ &\leq \frac{1}{T} \sum_{t=1}^T \sum_{\tau=1}^T \|\nabla c_t(\mathbf{x}_t) - \nabla c_t(\mathbf{x}_\tau)\|_2^2 + \frac{1}{T} \sum_{t=1}^T \sum_{\tau=1}^T \|\nabla c_t(\mathbf{x}_\tau) - \nabla c_\tau(\mathbf{x}_\tau)\|_2^2 \\ &= \text{VAR}_T^1 + \text{VAR}_T^2. \end{aligned} \quad (5)$$

We see that the variation  $\text{VAR}_T$  is bounded by two parts:  $\text{VAR}_T^1$  essentially measures the smoothness of individual cost functions, while  $\text{VAR}_T^2$  measures the variation in the gradients of cost functions. Let us consider an easy setting when all cost functions are identical. In this case,  $\text{VAR}_T^2$  vanishes, and  $\text{VAR}_T$  is equal to  $\text{VAR}_T^1/2$ , i.e.,

$$\begin{aligned}\text{VAR}_T &= \frac{1}{2T} \sum_{t,\tau=1}^T \|\nabla c_t(\mathbf{x}_t) - \nabla c_\tau(\mathbf{x}_\tau)\|^2 = \frac{1}{2T} \sum_{t,\tau=1}^T \|\nabla c_t(\mathbf{x}_t) - \nabla c_t(\mathbf{x}_\tau)\|^2 \\ &= \frac{\text{VAR}_T^1}{2}.\end{aligned}$$

As a result, the regret of the FTRL algorithm for online convex optimization may still be bounded by  $O(\sqrt{T})$  regardless of the smoothness of the cost function.

To address this challenge, we develop two novel algorithms for online convex optimization that bound the regret by the variation of cost functions. In particular, we would like to bound the regret of online convex optimization by the variation of cost functions defined as follows

$$\text{GV}_T = \sum_{t=1}^{T-1} \max_{\mathbf{x} \in \mathcal{P}} \|\nabla c_{t+1}(\mathbf{x}) - \nabla c_t(\mathbf{x})\|_2^2. \quad (6)$$

Note that the variation in (6) is defined in terms of gradual difference between individual cost function to its previous one, while the variation in (1) [16] is defined in terms of total difference between individual cost vectors to their mean. Therefore we refer to the variation defined in (6) as *gradual variation*<sup>1</sup>, and to the variation defined in (1) as *total variation*. It is straightforward to show that when  $c_t(\mathbf{x}) = \mathbf{f}_t^\top \mathbf{x}$ , the gradual variation  $\text{GV}_T$  defined in (6) is upper bounded by the total variation  $\text{VAR}_T$  defined in (1) with a constant factor:

$$\sum_{t=1}^{T-1} \|\mathbf{f}_{t+1} - \mathbf{f}_t\|_2^2 \leq \sum_{t=1}^{T-1} 2\|\mathbf{f}_{t+1} - \mu\|_2^2 + 2\|\mathbf{f}_t - \mu\|_2^2 \leq 4 \sum_{t=1}^T \|\mathbf{f}_t - \mu\|_2^2.$$

On the other hand, we can not bound the total variation by the gradual variation up to a constant. This is verified by the following example:  $\mathbf{f}_1 = \dots = \mathbf{f}_{T/2} = \mathbf{f}$  and  $\mathbf{f}_{T/2+1} = \dots = \mathbf{f}_T = \mathbf{g} \neq \mathbf{f}$ . The total variation in (1) is given by

$$\text{VAR}_T = \sum_{t=1}^T \|\mathbf{f}_t - \mu\|_2^2 = \frac{T}{2} \left\| \mathbf{f} - \frac{\mathbf{f} + \mathbf{g}}{2} \right\|_2^2 + \frac{T}{2} \left\| \mathbf{g} - \frac{\mathbf{f} + \mathbf{g}}{2} \right\|_2^2 = \Omega(T),$$

while the gradual variation defined in (6) is a constant given by

$$\text{GV}_T = \sum_{t=1}^{T-1} \|\mathbf{f}_{t+1} - \mathbf{f}_t\|_2^2 = \|\mathbf{f} - \mathbf{g}\|_2^2 = O(1).$$

<sup>1</sup> This is also termed as *deviation* in [8].

Based on the above analysis, we claim that the regret bound by gradual variation is usually tighter than total variation. Before closing this section, we present the following lower bound for FTRL whose proof can be found in [8].

**Theorem 1** *The regret of FTRL is at least  $\Omega(\min(GV_T, \sqrt{T}))$ .*

The theorem motivates us to develop new algorithms for online convex optimization to achieve a gradual variation bound of  $O(\sqrt{GV_T})$ .

The remainder of the paper is organized as follows. We present in Section 2 the proposed algorithms and the main results. Section 3 generalizes the results to the special class of non-smooth functions that is composed of a smooth and a non-smooth component. Section 4 is devoted to extending the proposed algorithms to the setting where only partial feedback about the cost functions are available, i.e., online bandit convex optimization with a variation-based regret bound. Finally, in Section 5, we conclude this work and discuss a few open problems.

## 2 Algorithms and Main Results

Without loss of generality, we assume that the decision set  $\mathcal{P}$  is contained in a unit ball  $\mathcal{B}$ , i.e.,  $\mathcal{P} \subseteq \mathcal{B}$ , and  $0 \in \mathcal{P}$  [16]. We propose two algorithms for online convex optimization. The first algorithm is an improved FTRL and the second one is based on the mirror prox method [21]. One common feature shared by the two algorithms is that both of them maintain two sequences of solutions: decision vectors  $\mathbf{x}_{1:T} = (\mathbf{x}_1, \dots, \mathbf{x}_T)$  and searching vectors  $\mathbf{z}_{1:T} = (\mathbf{z}_1, \dots, \mathbf{z}_T)$  that facilitate the updates of decision vectors. Both algorithms share almost the same regret bound except for a constant factor. To facilitate the discussion, besides the variation of cost functions defined in (6), we define another variation, named *extended gradual variation*, as follows

$$\text{EGV}_{T,2}(\mathbf{y}_{1:T}) = \sum_{t=0}^{T-1} \|\nabla c_{t+1}(\mathbf{y}_t) - \nabla c_t(\mathbf{y}_t)\|_2^2 \leq \|\nabla c_1(\mathbf{y}_0)\|_2^2 + \text{GV}_T, \quad (7)$$

where  $c_0(\mathbf{x}) = 0$ , the sequence  $(\mathbf{y}_0, \dots, \mathbf{y}_T)$  is either  $(\mathbf{z}_0, \dots, \mathbf{z}_T)$  (as in the improved FTRL) or  $(\mathbf{x}_0, \dots, \mathbf{x}_T)$  (as in the prox method) and the subscript 2 means the variation is defined with respect to  $\ell_2$  norm. When all cost functions are identical,  $\text{GV}_T$  becomes zero and the extended variation  $\text{EGV}_{T,2}(\mathbf{y}_{1:T})$  is reduced to  $\|\nabla c_1(\mathbf{y}_0)\|_2^2$ , a constant independent from the number of trials. In the sequel, we use the notation  $\text{EGV}_{T,2}$  for simplicity. In this study, we assume smooth cost functions with Lipschitz continuous gradients, i.e., there exists a constant  $L > 0$  such that

$$\|\nabla c_t(\mathbf{x}) - \nabla c_t(\mathbf{z})\|_2 \leq L\|\mathbf{x} - \mathbf{z}\|_2, \forall \mathbf{x}, \mathbf{z} \in \mathcal{P}, \forall t. \quad (8)$$

**Algorithm 3** Improved FTRL for Online Convex Optimization

- 
- 1: **Input:**  $\eta \in (0, 1]$
  - 2: **Initialization:**  $\mathbf{z}_0 = \mathbf{0}$  and  $c_0(\mathbf{x}) = 0$
  - 3: **for**  $t = 1, \dots, T$  **do**
  - 4:   Predict  $\mathbf{x}_t$  by

$$\mathbf{x}_t = \arg \min_{\mathbf{x} \in \mathcal{P}} \left\{ \mathbf{x}^\top \nabla c_{t-1}(\mathbf{z}_{t-1}) + \frac{L}{2\eta} \|\mathbf{x} - \mathbf{z}_{t-1}\|_2^2 \right\}$$

- 5:   Receive cost function  $c_t(\cdot)$  and incur loss  $c_t(\mathbf{x}_t)$
- 6:   Update  $\mathbf{z}_t$  by

$$\mathbf{z}_t = \arg \min_{\mathbf{x} \in \mathcal{P}} \left\{ \mathbf{x}^\top \sum_{\tau=1}^t \nabla c_\tau(\mathbf{z}_{\tau-1}) + \frac{L}{2\eta} \|\mathbf{x}\|_2^2 \right\}$$

7: **end for**

---

Our results show that for online convex optimization with  $L$ -smooth cost functions, the regrets of the proposed algorithms can be bounded as follows

$$\sum_{t=1}^T c_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T c_t(\mathbf{x}) \leq O\left(\sqrt{\text{EGV}_{T,2}}\right) + \text{constant}. \quad (9)$$

*Remark 1* We would like to emphasize that our assumption about the smoothness of cost functions is necessary<sup>2</sup> to achieve the variation-based bound stated in (9). To see this, consider the special case of  $c_1(\mathbf{x}) = \dots = c_T(\mathbf{x}) = c(\mathbf{x})$ . If the bound in (9) holds for any sequence of convex functions, then for the special case where all the cost functions are identical, we have

$$\sum_{t=1}^T c(\mathbf{x}_t) \leq \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T c(\mathbf{x}) + O(1),$$

implying that  $\widehat{\mathbf{x}}_T = (1/T) \sum_{t=1}^T \mathbf{x}_t$  approaches the optimal solution at the rate of  $O(1/T)$ . This contradicts the lower complexity bound (i.e.  $O(1/\sqrt{T})$ ) for any optimization method which only uses first order information about the cost functions [22, Theorem 3.2.1]. This analysis indicates that the smoothness assumption is necessary to attain variation based regret bound for general online convex optimization problem. We would like to emphasize the fact that this contradiction holds when only the gradient information about the cost functions is provided to the learner and the learner *may* be able to achieve a variation-based bound using second order information about the cost functions, which is not the focus of the present work.

## 2.1 An Improved FTRL Algorithm for Online Convex Optimization

---

<sup>2</sup> Gradual variation bounds for a special class of non-smooth cost functions that are composed of a smooth component and a non-smooth component are discussed in section 3.



The improved FTRL algorithm for online convex optimization is presented in Algorithm 3. Note that in step 6, the searching vectors  $\mathbf{z}_t$  are updated according to the FTRL algorithm after receiving the cost function  $c_t(\cdot)$ . To understand the updating procedure for the decision vector  $\mathbf{x}_t$  specified in step 4, we rewrite it as

$$\mathbf{x}_t = \arg \min_{\mathbf{x} \in \mathcal{P}} \left\{ c_{t-1}(\mathbf{z}_{t-1}) + (\mathbf{x} - \mathbf{z}_{t-1})^\top \nabla c_{t-1}(\mathbf{z}_{t-1}) + \frac{L}{2\eta} \|\mathbf{x} - \mathbf{z}_{t-1}\|_2^2 \right\}. \quad (10)$$

Notice that

$$\begin{aligned} c_t(\mathbf{x}) &\leq c_t(\mathbf{z}_{t-1}) + (\mathbf{x} - \mathbf{z}_{t-1})^\top \nabla c_t(\mathbf{z}_{t-1}) + \frac{L}{2} \|\mathbf{x} - \mathbf{z}_{t-1}\|_2^2 \\ &\leq c_t(\mathbf{z}_{t-1}) + (\mathbf{x} - \mathbf{z}_{t-1})^\top \nabla c_t(\mathbf{z}_{t-1}) + \frac{L}{2\eta} \|\mathbf{x} - \mathbf{z}_{t-1}\|_2^2, \end{aligned} \quad (11)$$

where the first inequality follows the smoothness condition in (8) and the second inequality follows from the fact  $\eta \leq 1$ . The inequality (11) provides an upper bound for  $c_t(\mathbf{x})$  and therefore can be used as an approximation of  $c_t(\mathbf{x})$  for predicting  $\mathbf{x}_t$ . However, since  $\nabla c_t(\mathbf{z}_{t-1})$  is unknown before the prediction, we use  $\nabla c_{t-1}(\mathbf{z}_{t-1})$  as a surrogate for  $\nabla c_t(\mathbf{z}_{t-1})$ , leading to the updating rule in (10). It is this approximation that leads to the variation bound. The following theorem states the regret bound of Algorithm 3.

**Theorem 2** *Let  $c_t(\cdot), t = 1, \dots, T$  be a sequence of convex functions with  $L$ -Lipschitz continuous gradients. By setting  $\eta = \min \{1, L/\sqrt{\text{EGV}_{T,2}}\}$ , we have the following regret bound for Algorithm 3*

$$\sum_{t=1}^T c_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T c_t(\mathbf{x}) \leq \max \left( L, \sqrt{\text{EGV}_{T,2}} \right).$$

*Remark 2* Comparing with the variation bound in (5) for the FTRL algorithm, the smoothness parameter  $L$  plays the same role as  $\text{VAR}_T^1$  that accounts for the smoothness of cost functions, and term  $\text{EGV}_{T,2}$  plays the same role as  $\text{VAR}_T^2$  that accounts for the variation in the cost functions. Compared to the FTRL algorithm, the key advantage of the improved FTRL algorithm is that the regret bound is reduced to a constant when the cost functions change only by a constant number of times along the horizon. Of course, the extended variation  $\text{EGV}_{T,2}$  may not be known apriori for setting the optimal  $\eta$ , we can apply the standard doubling trick [6] to obtain a bound that holds uniformly over time and is a factor at most 8 from the bound obtained with the optimal choice of  $\eta$ . The details are provided in Appendix A.

To prove Theorem 2, we first present the following lemma.

**Lemma 1** Let  $c_t(\cdot), t = 1, \dots, T$  be a sequence of convex functions with  $L$ -Lipschitz continuous gradients. By running Algorithm 3 over  $T$  trials, we have

$$\begin{aligned} \sum_{t=1}^T c_t(\mathbf{x}_t) &\leq \min_{\mathbf{x} \in \mathcal{P}} \left[ \frac{L}{2\eta} \|\mathbf{x}\|_2^2 + \sum_{t=1}^T c_t(\mathbf{z}_{t-1}) + (\mathbf{x} - \mathbf{z}_{t-1})^\top \nabla c_t(\mathbf{z}_{t-1}) \right] \\ &\quad + \frac{\eta}{2L} \sum_{t=0}^{T-1} \|\nabla c_{t+1}(\mathbf{z}_t) - \nabla c_t(\mathbf{z}_t)\|_2^2. \end{aligned}$$

With this lemma, we can easily prove Theorem 2 by exploring the convexity of  $c_t(\mathbf{x})$ .

*Proof (of Theorem 2)* By using  $\|\mathbf{x}\|_2 \leq 1, \forall \mathbf{x} \in \mathcal{P} \subseteq \mathcal{B}$ , and the convexity of  $c_t(\mathbf{x})$ , we have

$$\min_{\mathbf{x} \in \mathcal{P}} \left\{ \frac{L}{2\eta} \|\mathbf{x}\|_2^2 + \sum_{t=1}^T c_t(\mathbf{z}_{t-1}) + (\mathbf{x} - \mathbf{z}_{t-1})^\top \nabla c_t(\mathbf{z}_{t-1}) \right\} \leq \frac{L}{2\eta} + \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T c_t(\mathbf{x}).$$

Combining the above result with Lemma 1, we have

$$\sum_{t=1}^T c_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T c_t(\mathbf{x}) \leq \frac{L}{2\eta} + \frac{\eta}{2L} \text{EGV}_{T,2}.$$

By choosing  $\eta = \min(1, L/\sqrt{\text{EGV}_{T,2}})$ , we have the regret bound claimed in Theorem 2.

The Lemma 1 is proved by induction. The key to the proof is that  $\mathbf{z}_t$  is the optimal solution to the strongly convex minimization problem in Lemma 1, i.e.,

$$\mathbf{z}_t = \arg \min_{\mathbf{x} \in \mathcal{P}} \left[ \frac{L}{2\eta} \|\mathbf{x}\|_2^2 + \sum_{\tau=1}^t c_\tau(\mathbf{z}_{\tau-1}) + (\mathbf{x} - \mathbf{z}_{\tau-1})^\top \nabla c_\tau(\mathbf{z}_{\tau-1}) \right]$$

*Proof (of Lemma 1)* We prove the inequality by induction. When  $T = 1$ , we have  $\mathbf{x}_1 = \mathbf{z}_0 = 0$  and

$$\begin{aligned} &\min_{\mathbf{x} \in \mathcal{P}} \left[ \frac{L}{2\eta} \|\mathbf{x}\|_2^2 + c_1(\mathbf{z}_0) + (\mathbf{x} - \mathbf{z}_0)^\top \nabla c_1(\mathbf{z}_0) \right] + \frac{\eta}{2L} \|\nabla c_1(\mathbf{z}_0)\|_2^2 \\ &\geq c_1(\mathbf{z}_0) + \frac{\eta}{2L} \|\nabla c_1(\mathbf{z}_0)\|_2^2 + \min_{\mathbf{x}} \left\{ \frac{L}{2\eta} \|\mathbf{x}\|_2^2 + (\mathbf{x} - \mathbf{z}_0)^\top \nabla c_1(\mathbf{z}_0) \right\} = c_1(\mathbf{z}_0) = c_1(\mathbf{x}_1). \end{aligned}$$

where the inequality follows that by relaxing the minimization domain  $\mathbf{x} \in \mathcal{P}$  to the whole space. We assume the inequality holds for  $t$  and aim to prove it for  $t + 1$ . To this end, we define

$$\begin{aligned} \psi_t(\mathbf{x}) &= \left[ \frac{L}{2\eta} \|\mathbf{x}\|_2^2 + \sum_{\tau=1}^t c_\tau(\mathbf{z}_{\tau-1}) + (\mathbf{x} - \mathbf{z}_{\tau-1})^\top \nabla c_\tau(\mathbf{z}_{\tau-1}) \right] \\ &\quad + \frac{\eta}{2L} \sum_{\tau=0}^{t-1} \|\nabla c_{\tau+1}(\mathbf{z}_\tau) - \nabla c_\tau(\mathbf{z}_\tau)\|_2^2. \end{aligned}$$

According to the updating procedure for  $\mathbf{z}_t$  in step 6, we have  $\mathbf{z}_t = \arg \min_{\mathbf{x} \in \mathcal{P}} \psi_t(\mathbf{x})$ . Define  $\phi_t = \psi_t(\mathbf{z}_t) = \min_{\mathbf{x} \in \mathcal{P}} \psi_t(\mathbf{x})$ . Since  $\psi_t(\mathbf{x})$  is a  $(L/\eta)$ -strongly convex function, we have

$$\begin{aligned} \psi_{t+1}(\mathbf{x}) - \psi_{t+1}(\mathbf{z}_t) &\geq \frac{L}{2\eta} \|\mathbf{x} - \mathbf{z}_t\|_2^2 + (\mathbf{x} - \mathbf{z}_t)^\top \nabla \psi_{t+1}(\mathbf{z}_t) \\ &= \frac{L}{2\eta} \|\mathbf{x} - \mathbf{z}_t\|_2^2 + (\mathbf{x} - \mathbf{z}_t)^\top (\nabla \psi_t(\mathbf{z}_t) + \nabla c_{t+1}(\mathbf{z}_t)). \end{aligned}$$

Setting  $\mathbf{x} = \mathbf{z}_{t+1} = \arg \min_{\mathbf{x} \in \mathcal{P}} \psi_{t+1}(\mathbf{x})$  in the above inequality results in

$$\begin{aligned} \psi_{t+1}(\mathbf{z}_{t+1}) - \psi_{t+1}(\mathbf{z}_t) &= \phi_{t+1} - (\phi_t + c_{t+1}(\mathbf{z}_t) + \frac{\eta}{2L} \|\nabla c_{t+1}(\mathbf{z}_t) - \nabla c_t(\mathbf{z}_t)\|_2^2) \\ &\geq \frac{L}{2\eta} \|\mathbf{z}_{t+1} - \mathbf{z}_t\|_2^2 + (\mathbf{z}_{t+1} - \mathbf{z}_t)^\top (\nabla \psi_t(\mathbf{z}_t) + \nabla c_{t+1}(\mathbf{z}_t)) \\ &\geq \frac{L}{2\eta} \|\mathbf{z}_{t+1} - \mathbf{z}_t\|_2^2 + (\mathbf{z}_{t+1} - \mathbf{z}_t)^\top \nabla c_{t+1}(\mathbf{z}_t), \end{aligned}$$

where the second inequality follows from the fact  $\mathbf{z}_t = \arg \min_{\mathbf{x} \in \mathcal{P}} \psi_t(\mathbf{x})$ , and therefore  $(\mathbf{x} - \mathbf{z}_t)^\top \nabla \psi_t(\mathbf{z}_t) \geq 0, \forall \mathbf{x} \in \mathcal{P}$ . Moving  $c_{t+1}(\mathbf{z}_t)$  in the above inequality to the right hand side, we have

$$\begin{aligned} &\phi_{t+1} - \phi_t - \frac{\eta}{2L} \|\nabla c_{t+1}(\mathbf{z}_t) - \nabla c_t(\mathbf{z}_t)\|_2^2 \tag{12} \\ &\geq \frac{L}{2\eta} \|\mathbf{z}_{t+1} - \mathbf{z}_t\|_2^2 + (\mathbf{z}_{t+1} - \mathbf{z}_t)^\top \nabla c_{t+1}(\mathbf{z}_t) + c_{t+1}(\mathbf{z}_t) \\ &\geq \min_{\mathbf{x} \in \mathcal{P}} \left\{ \frac{L}{2\eta} \|\mathbf{x} - \mathbf{z}_t\|_2^2 + (\mathbf{x} - \mathbf{z}_t)^\top \nabla c_{t+1}(\mathbf{z}_t) + c_{t+1}(\mathbf{z}_t) \right\} \\ &= \min_{\mathbf{x} \in \mathcal{P}} \left\{ \underbrace{\frac{L}{2\eta} \|\mathbf{x} - \mathbf{z}_t\|_2^2 + (\mathbf{x} - \mathbf{z}_t)^\top \nabla c_t(\mathbf{z}_t) + c_{t+1}(\mathbf{z}_t)}_{\rho(\mathbf{x})} + \underbrace{(\mathbf{x} - \mathbf{z}_t)^\top (\nabla c_{t+1}(\mathbf{z}_t) - \nabla c_t(\mathbf{z}_t))}_{r(\mathbf{x})} \right\}. \end{aligned}$$

To bound the right hand side, we note that  $\mathbf{x}_{t+1}$  is the minimizer of  $\rho(\mathbf{x})$  by step 4 in Algorithm 3, and  $\rho(\mathbf{x})$  is a  $L/\eta$ -strongly convex function, so we have

$$\rho(\mathbf{x}) \geq \rho(\mathbf{x}_{t+1}) + \underbrace{(\mathbf{x} - \mathbf{x}_{t+1})^\top \nabla \rho(\mathbf{x}_{t+1})}_{\geq 0} + \frac{L}{2\eta} \|\mathbf{x} - \mathbf{x}_{t+1}\|_2^2 \geq \rho(\mathbf{x}_{t+1}) + \frac{L}{2\eta} \|\mathbf{x} - \mathbf{x}_{t+1}\|_2^2.$$

Then we have

$$\begin{aligned} \rho(\mathbf{x}) + c_{t+1}(\mathbf{z}_t) + r(\mathbf{x}) &\geq \rho(\mathbf{x}_{t+1}) + c_{t+1}(\mathbf{z}_t) + \frac{L}{2\eta} \|\mathbf{x} - \mathbf{x}_{t+1}\|_2^2 + r(\mathbf{x}) \\ &= \underbrace{\frac{L}{2\eta} \|\mathbf{x}_{t+1} - \mathbf{z}_t\|_2^2 + (\mathbf{x}_{t+1} - \mathbf{z}_t)^\top \nabla c_t(\mathbf{z}_t) + c_{t+1}(\mathbf{z}_t)}_{\rho(\mathbf{x}_{t+1})} + \frac{L}{2\eta} \|\mathbf{x} - \mathbf{x}_{t+1}\|_2^2 + r(\mathbf{x}) \end{aligned}$$

Plugging above inequality into the inequality in (12), we have

$$\begin{aligned} & \phi_{t+1} - \phi_t - \frac{\eta}{2L} \|\nabla c_{t+1}(\mathbf{z}_t) - \nabla c_t(\mathbf{z}_t)\|_2^2 \\ & \geq \frac{L}{2\eta} \|\mathbf{x}_{t+1} - \mathbf{z}_t\|_2^2 + (\mathbf{x}_{t+1} - \mathbf{z}_t)^\top \nabla c_t(\mathbf{z}_t) + c_{t+1}(\mathbf{z}_t) \\ & \quad + \min_{\mathbf{x} \in \mathcal{P}} \left\{ \frac{L}{2\eta} \|\mathbf{x} - \mathbf{x}_{t+1}\|_2^2 + (\mathbf{x} - \mathbf{z}_t)^\top (\nabla c_{t+1}(\mathbf{z}_t) - \nabla c_t(\mathbf{z}_t)) \right\} \end{aligned}$$

To continue the bounding, we proceed as follows

$$\begin{aligned} & \phi_{t+1} - \phi_t - \frac{\eta}{2L} \|\nabla c_{t+1}(\mathbf{z}_t) - \nabla c_t(\mathbf{z}_t)\|_2^2 \\ & \geq \frac{L}{2\eta} \|\mathbf{x}_{t+1} - \mathbf{z}_t\|_2^2 + (\mathbf{x}_{t+1} - \mathbf{z}_t)^\top \nabla c_t(\mathbf{z}_t) + c_{t+1}(\mathbf{z}_t) \\ & \quad + \min_{\mathbf{x} \in \mathcal{P}} \left\{ \frac{L}{2\eta} \|\mathbf{x} - \mathbf{x}_{t+1}\|_2^2 + (\mathbf{x} - \mathbf{z}_t)^\top (\nabla c_{t+1}(\mathbf{z}_t) - \nabla c_t(\mathbf{z}_t)) \right\} \\ & = \frac{L}{2\eta} \|\mathbf{x}_{t+1} - \mathbf{z}_t\|_2^2 + (\mathbf{x}_{t+1} - \mathbf{z}_t)^\top \nabla c_{t+1}(\mathbf{z}_t) + c_{t+1}(\mathbf{z}_t) \\ & \quad + \min_{\mathbf{x} \in \mathcal{P}} \left\{ \frac{L}{2\eta} \|\mathbf{x} - \mathbf{x}_{t+1}\|_2^2 + (\mathbf{x} - \mathbf{x}_{t+1})^\top (\nabla c_{t+1}(\mathbf{z}_t) - \nabla c_t(\mathbf{z}_t)) \right\} \\ & \geq \frac{L}{2\eta} \|\mathbf{x}_{t+1} - \mathbf{z}_t\|_2^2 + (\mathbf{x}_{t+1} - \mathbf{z}_t)^\top \nabla c_{t+1}(\mathbf{z}_t) + c_{t+1}(\mathbf{z}_t) \\ & \quad + \min_{\mathbf{x}} \left\{ \frac{L}{2\eta} \|\mathbf{x} - \mathbf{x}_{t+1}\|_2^2 + (\mathbf{x} - \mathbf{x}_{t+1})^\top (\nabla c_{t+1}(\mathbf{z}_t) - \nabla c_t(\mathbf{z}_t)) \right\} \\ & = \frac{L}{2\eta} \|\mathbf{x}_{t+1} - \mathbf{z}_t\|_2^2 + (\mathbf{x}_{t+1} - \mathbf{z}_t)^\top \nabla c_{t+1}(\mathbf{z}_t) + c_{t+1}(\mathbf{z}_t) - \frac{\eta}{2L} \|\nabla c_{t+1}(\mathbf{z}_t) - \nabla c_t(\mathbf{z}_t)\|_2^2 \\ & \geq c_{t+1}(\mathbf{x}_{t+1}) - \frac{\eta}{2L} \|\nabla c_{t+1}(\mathbf{z}_t) - \nabla c_t(\mathbf{z}_t)\|_2^2, \end{aligned}$$

where the first equality follows by writing  $(\mathbf{x}_{t+1} - \mathbf{z}_t)^\top \nabla c_t(\mathbf{z}_t) = (\mathbf{x}_{t+1} - \mathbf{z}_t)^\top \nabla c_{t+1}(\mathbf{z}_t) - (\mathbf{x}_{t+1} - \mathbf{z}_t)^\top (\nabla c_{t+1}(\mathbf{z}_t) - \nabla c_t(\mathbf{z}_t))$  and combining with  $(\mathbf{x} - \mathbf{z}_t)^\top (\nabla c_{t+1}(\mathbf{z}_t) - \nabla c_t(\mathbf{z}_t))$ , and the last inequality follows from the smoothness condition of  $c_{t+1}(\mathbf{x})$ . Since by induction  $\phi_t \geq \sum_{\tau=1}^t c_\tau(\mathbf{x}_\tau)$ , we have  $\phi_{t+1} \geq \sum_{\tau=1}^{t+1} c_\tau(\mathbf{x}_\tau)$ .

## 2.2 A Prox Method for Online Convex Optimization

In this subsection, we present a prox method for online convex optimization that shares the same order of regret bound as the improved FTRL algorithm. It is closely related to the prox method in [21] by maintaining two sets of vectors  $\mathbf{x}_{1:T}$  and  $\mathbf{z}_{1:T}$ , where  $\mathbf{x}_t$  and  $\mathbf{z}_t$  are computed by *gradient mappings*

**Algorithm 4** A Prox Method for Online Convex Optimization

- 
- 1: **Input:**  $\eta > 0$
  - 2: **Initialization:**  $\mathbf{z}_0 = \mathbf{x}_0 = \mathbf{0}$  and  $c_0(\mathbf{x}) = 0$
  - 3: **for**  $t = 1, \dots, T$  **do**
  - 4:   Predict  $\mathbf{x}_t$  by

$$\mathbf{x}_t = \arg \min_{\mathbf{x} \in \mathcal{P}} \left\{ \mathbf{x}^\top \nabla c_{t-1}(\mathbf{x}_{t-1}) + \frac{L}{2\eta} \|\mathbf{x} - \mathbf{z}_{t-1}\|_2^2 \right\}$$

- 5:   Receive cost function  $c_t(\cdot)$  and incur loss  $c_t(\mathbf{x}_t)$
- 6:   Update  $\mathbf{z}_t$  by

$$\mathbf{z}_t = \arg \min_{\mathbf{x} \in \mathcal{P}} \left\{ \mathbf{x}^\top \nabla c_t(\mathbf{x}_t) + \frac{L}{2\eta} \|\mathbf{x} - \mathbf{z}_{t-1}\|_2^2 \right\}$$

- 7: **end for**
- 

using  $\nabla c_{t-1}(\mathbf{x}_{t-1})$ , and  $\nabla c_t(\mathbf{x}_t)$ , respectively, as

$$\begin{aligned} \mathbf{x}_t &= \arg \min_{\mathbf{x} \in \mathcal{P}} \frac{1}{2} \left\| \mathbf{x} - \left( \mathbf{z}_{t-1} - \frac{\eta}{L} \nabla c_{t-1}(\mathbf{x}_{t-1}) \right) \right\|_2^2 \\ \mathbf{z}_t &= \arg \min_{\mathbf{x} \in \mathcal{P}} \frac{1}{2} \left\| \mathbf{x} - \left( \mathbf{z}_{t-1} - \frac{\eta}{L} \nabla c_t(\mathbf{x}_t) \right) \right\|_2^2 \end{aligned}$$

The detailed steps are shown in Algorithm 4, where we use an equivalent form of updates for  $\mathbf{x}_t$  and  $\mathbf{z}_t$  in order to compare to Algorithm 3. Algorithm 4 differs from Algorithm 3: (i) in updating the searching points  $\mathbf{z}_t$ , Algorithm 3 updates  $\mathbf{z}_t$  by the FTRL scheme using *all* the gradients of the cost functions at  $\{\mathbf{z}_\tau\}_{\tau=1}^{t-1}$ , while Algorithm 4 updates  $\mathbf{z}_t$  by a prox method using a *single* gradient  $\nabla c_t(\mathbf{x}_t)$ , and (ii) in updating the decision vector  $\mathbf{x}_t$ , Algorithm 4 uses the gradient  $\nabla c_{t-1}(\mathbf{x}_{t-1})$  instead of  $\nabla c_{t-1}(\mathbf{z}_{t-1})$ . The advantage of Algorithm 4 compared to Algorithm 3 is that it only requires to compute one gradient  $\nabla c_t(\mathbf{x}_t)$  for each loss function; in contrast, the improved FTRL algorithm in Algorithm 3 needs to compute the gradients of  $c_t(\mathbf{x})$  at two searching points  $\mathbf{z}_t$  and  $\mathbf{z}_{t-1}$ . It is these differences mentioned above that make it easier to extend the prox method to a bandit setting, which will be discussed in section 5.

The following theorem states the regret bound of the prox method for online convex optimization.

**Theorem 3** *Let  $c_t(\cdot), t = 1, \dots, T$  be a sequence of convex functions with  $L$ -Lipschitz continuous gradients. By setting  $\eta = (1/2) \min \{1/\sqrt{2}, L/\sqrt{\text{EGV}_{T,2}}\}$ , we have the following regret bound for Algorithm 4*

$$\sum_{t=1}^T c_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T c_t(\mathbf{x}) \leq 2 \max \left( \sqrt{2}L, \sqrt{\text{EGV}_{T,2}} \right).$$

We note that compared to Theorem 2, the regret bound in Theorem 3 is slightly worse by a factor of 2.

To prove Theorem 3, we need the following lemma, which is the Lemma 3.1 in [21] stated in our notations.

**Lemma 2 (Lemma 3.1 [21])** Let  $\omega(\mathbf{z})$  be a  $\alpha$ -strongly convex function with respect to the norm  $\|\cdot\|$ , whose dual norm is denoted by  $\|\cdot\|_*$ , and  $D(\mathbf{x}, \mathbf{z}) = \omega(\mathbf{x}) - (\omega(\mathbf{z}) + (\mathbf{x} - \mathbf{z})^\top \omega'(\mathbf{z}))$  be the Bregman distance induced by function  $\omega(\mathbf{x})$ . Let  $Z$  be a convex compact set, and  $U \subseteq Z$  be convex and closed. Let  $\mathbf{z} \in Z$ ,  $\gamma > 0$ , Consider the points,

$$\mathbf{x} = \arg \min_{\mathbf{u} \in U} \gamma \mathbf{u}^\top \xi + D(\mathbf{u}, \mathbf{z}), \quad (13)$$

$$\mathbf{z}_+ = \arg \min_{\mathbf{u} \in U} \gamma \mathbf{u}^\top \zeta + D(\mathbf{u}, \mathbf{z}), \quad (14)$$

then for any  $\mathbf{u} \in U$ , we have

$$\gamma \zeta^\top (\mathbf{x} - \mathbf{u}) \leq D(\mathbf{u}, \mathbf{z}) - D(\mathbf{u}, \mathbf{z}_+) + \frac{\gamma^2}{\alpha} \|\xi - \zeta\|_*^2 - \frac{\alpha}{2} [\|\mathbf{x} - \mathbf{z}\|^2 + \|\mathbf{x} - \mathbf{z}_+\|^2]. \quad (15)$$

In order not to have readers struggle with complex notations in [21] for the proof of Lemma 2, we present a detailed proof in Appendix A which is an adaption of the original proof to our notations.

Theorem 3 can be proved by using the above Lemma, because the updates of  $\mathbf{x}_t, \mathbf{z}_t$  can be written equivalently as (13) and (14), respectively. The proof below starts from (15) and bounds the summation of each term over  $t = 1, \dots, T$ , respectively.

*Proof (of Theorem 3)* First, we note that the two updates in step 4 and step 6 of Algorithm 4 fit in the Lemma 2 if we let  $U = Z = \mathcal{P}$ ,  $\mathbf{z} = \mathbf{z}_{t-1}$ ,  $\mathbf{x} = \mathbf{x}_t$ ,  $\mathbf{z}_+ = \mathbf{z}_t$ , and  $\omega(\mathbf{x}) = \frac{1}{2} \|\mathbf{x}\|_2^2$ , which is 1-strongly convex function with respect to  $\|\cdot\|_2$ . Then  $D(\mathbf{u}, \mathbf{z}) = \frac{1}{2} \|\mathbf{u} - \mathbf{z}\|_2^2$ . As a result, the two updates for  $\mathbf{x}_t, \mathbf{z}_t$  in Algorithm 4 are exactly the updates in (13) and (14) with  $\mathbf{z} = \mathbf{z}_{t-1}, \gamma = \eta/L$ ,  $\xi = \nabla c_{t-1}(\mathbf{z}_{t-1})$ , and  $\zeta = \nabla c_t(\mathbf{x}_t)$ . Replacing these into (15), we have the following inequality for any  $\mathbf{u} \in \mathcal{P}$ ,

$$\begin{aligned} \frac{\eta}{L} (\mathbf{x}_t - \mathbf{u})^\top \nabla c_t(\mathbf{x}_t) &\leq \frac{1}{2} (\|\mathbf{u} - \mathbf{z}_{t-1}\|_2^2 - \|\mathbf{u} - \mathbf{z}_t\|_2^2) \\ &\quad + \frac{\eta^2}{L^2} \|\nabla c_t(\mathbf{x}_t) - \nabla c_{t-1}(\mathbf{x}_{t-1})\|_2^2 - \frac{1}{2} (\|\mathbf{x}_t - \mathbf{z}_{t-1}\|_2^2 + \|\mathbf{x}_t - \mathbf{z}_t\|_2^2) \end{aligned}$$

Then we have

$$\begin{aligned} \frac{\eta}{L} (c_t(\mathbf{x}_t) - c_t(\mathbf{u})) &\leq \frac{\eta}{L} (\mathbf{x}_t - \mathbf{u})^\top \nabla c_t(\mathbf{x}_t) \leq \frac{1}{2} (\|\mathbf{u} - \mathbf{z}_{t-1}\|_2^2 - \|\mathbf{u} - \mathbf{z}_t\|_2^2) \\ &\quad + \frac{2\eta^2}{L^2} \|\nabla c_t(\mathbf{x}_{t-1}) - \nabla c_{t-1}(\mathbf{x}_{t-1})\|_2^2 + \frac{2\eta^2}{L^2} \|\nabla c_t(\mathbf{x}_t) - \nabla c_t(\mathbf{x}_{t-1})\|_2^2 \\ &\quad - \frac{1}{2} (\|\mathbf{x}_t - \mathbf{z}_{t-1}\|_2^2 + \|\mathbf{x}_t - \mathbf{z}_t\|_2^2) \\ &\leq \frac{1}{2} (\|\mathbf{u} - \mathbf{z}_{t-1}\|_2^2 - \|\mathbf{u} - \mathbf{z}_t\|_2^2) + \frac{2\eta^2}{L^2} \|\nabla c_t(\mathbf{x}_{t-1}) - \nabla c_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\ &\quad + 2\eta^2 \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 - \frac{1}{2} (\|\mathbf{x}_t - \mathbf{z}_{t-1}\|_2^2 + \|\mathbf{x}_t - \mathbf{z}_t\|_2^2), \end{aligned}$$

where the first inequality follows the convexity of  $c_t(\mathbf{x})$ , and the third inequality follows the smoothness of  $c_t(\mathbf{x})$ . By taking the summation over  $t = 1, \dots, T$  with  $\mathbf{z}^* = \arg \min_{\mathbf{u} \in \mathcal{P}} \sum_{t=1}^T c_t(\mathbf{u})$ , and dividing both sides by  $\eta/L$ , we have

$$\begin{aligned} \sum_{t=1}^T c_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T c_t(\mathbf{x}) &\leq \frac{L}{2\eta} + \frac{2\eta}{L} \sum_{t=0}^{T-1} \|\nabla c_{t+1}(\mathbf{x}_t) - \nabla c_t(\mathbf{x}_t)\|_2^2 \\ &+ \underbrace{\sum_{t=1}^T 2\eta^2 \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 - \sum_{t=1}^T \frac{1}{2} (\|\mathbf{x}_t - \mathbf{z}_{t-1}\|_2^2 + \|\mathbf{x}_t - \mathbf{z}_t\|_2^2)}_{B_T} \end{aligned} \quad (16)$$

We can bound  $B_T$  as follows:

$$\begin{aligned} B_T &= \frac{1}{2} \sum_{t=1}^T \|\mathbf{x}_t - \mathbf{z}_{t-1}\|_2^2 + \frac{1}{2} \sum_{t=2}^{T+1} \|\mathbf{x}_{t-1} - \mathbf{z}_{t-1}\|_2^2 \\ &\geq \frac{1}{2} \sum_{t=2}^T (\|\mathbf{x}_t - \mathbf{z}_{t-1}\|_2^2 + \|\mathbf{x}_{t-1} - \mathbf{z}_{t-1}\|_2^2) \\ &\geq \frac{1}{4} \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 = \frac{1}{4} \sum_{t=1}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \end{aligned} \quad (17)$$

where the last equality follows that  $\mathbf{x}_1 = \mathbf{x}_0$ . Plugging the above bound into (18), we have

$$\begin{aligned} \sum_{t=1}^T c_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T c_t(\mathbf{x}) &\leq \frac{L}{2\eta} + \frac{2\eta}{L} \sum_{t=0}^{T-1} \|\nabla c_{t+1}(\mathbf{x}_t) - \nabla c_t(\mathbf{x}_t)\|_2^2 \\ &+ \sum_{t=1}^T \left(2\eta^2 - \frac{1}{4}\right) \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \end{aligned} \quad (18)$$

We complete the proof by plugging the value of  $\eta$ .

### 2.3 A General Prox Method and Some Special Cases

In this subsection, we first present a general prox method to obtain a variation bound defined in a general norm. Then we discuss four special cases: online linear optimization, prediction with expert advice, and online strictly convex optimization. The omitted proofs in this subsection can be easily duplicated by mimicking the proof of Theorem 3, if necessary with the help of previous analysis as mentioned in the appropriate text.

---

**Algorithm 5** A General Prox Method for Online Convex Optimization
 

---

- 1: **Input:**  $\eta > 0, \omega(\mathbf{z})$   
 2: **Initialization:**  $\mathbf{z}_0 = \mathbf{x}_0 = \min_{\mathbf{z} \in \mathcal{P}} \omega(\mathbf{z})$  and  $c_0(\mathbf{x}) = 0$   
 3: **for**  $t = 1, \dots, T$  **do**  
 4:   Predict  $\mathbf{x}_t$  by

$$\mathbf{x}_t = \arg \min_{\mathbf{x} \in \mathcal{P}} \left\{ \mathbf{x}^\top \nabla c_{t-1}(\mathbf{x}_{t-1}) + \frac{L}{\eta} D(\mathbf{x}, \mathbf{z}_{t-1}) \right\}$$

- 5:   Receive cost function  $c_t(\cdot)$  and incur loss  $c_t(\mathbf{x}_t)$   
 6:   Update  $\mathbf{z}_t$  by

$$\mathbf{z}_t = \arg \min_{\mathbf{x} \in \mathcal{P}} \left\{ \mathbf{x}^\top \nabla c_t(\mathbf{x}_t) + \frac{L}{\eta} D(\mathbf{x}, \mathbf{z}_{t-1}) \right\}$$

7: **end for**

---

### 2.3.1 A General Prox Method

The prox method, together with Lemma 2 provides an easy way to generalize the framework based on Euclidean norm to a general norm. To be precise, let  $\|\cdot\|$  denote a general norm,  $\|\cdot\|_*$  denote its dual norm,  $\omega(\mathbf{z})$  be a  $\alpha$ -strongly convex function with respect to the norm  $\|\cdot\|$ , and  $D(\mathbf{x}, \mathbf{z}) = \omega(\mathbf{x}) - (\omega(\mathbf{z}) + (\mathbf{x} - \mathbf{z})^\top \omega'(\mathbf{z}))$  be the Bregman distance induced by function  $\omega(\mathbf{x})$ . Let  $c_t(\cdot), t = 1, \dots, T$  be  $L$ -smooth functions with respect to norm  $\|\cdot\|$ , i.e.,

$$\|\nabla c_t(\mathbf{x}) - \nabla c_t(\mathbf{z})\|_* \leq L \|\mathbf{x} - \mathbf{z}\|.$$

Correspondingly, we define the extended gradual variation based on the general norm as follows:

$$\text{EGV}_T = \sum_{t=0}^{T-1} \|\nabla c_{t+1}(\mathbf{x}_t) - \nabla c_t(\mathbf{x}_t)\|_*^2. \quad (19)$$

Algorithm 5 gives the detailed steps for the general framework. We note that the key differences from Algorithm 4 are:  $\mathbf{z}_0$  is set to  $\min_{\mathbf{z} \in \mathcal{P}} \omega(\mathbf{z})$ , and the Euclidean distances in steps 4 and 6 are replaced by Bregman distances, i.e.,

$$\begin{aligned} \mathbf{x}_t &= \arg \min_{\mathbf{x} \in \mathcal{P}} \left\{ \mathbf{x}^\top \nabla c_{t-1}(\mathbf{x}_{t-1}) + \frac{L}{\eta} D(\mathbf{x}, \mathbf{z}_{t-1}) \right\}, \\ \mathbf{z}_t &= \arg \min_{\mathbf{x} \in \mathcal{P}} \left\{ \mathbf{x}^\top \nabla c_t(\mathbf{x}_t) + \frac{L}{\eta} D(\mathbf{x}, \mathbf{z}_{t-1}) \right\}. \end{aligned}$$

The following theorem states the variation-based regret bound for the general norm framework, where  $R$  measure the size of  $\mathcal{P}$  defined as

$$R = \sqrt{2(\max_{\mathbf{x} \in \mathcal{P}} \omega(\mathbf{x}) - \min_{\mathbf{x} \in \mathcal{P}} \omega(\mathbf{x}))}.$$



**Theorem 4** Let  $c_t(\cdot), t = 1, \dots, T$  be a sequence of convex functions whose gradients are  $L$ -smooth continuous,  $\omega(\mathbf{z})$  be a  $\alpha$ -strongly convex function, both with respect to norm  $\|\cdot\|$ , and  $\text{EGV}_T$  be defined in (19). By setting  $\eta = (1/2) \min \{\sqrt{\alpha}/\sqrt{2}, LR/\sqrt{\text{EGV}_T}\}$ , we have the following regret bound

$$\sum_{t=1}^T c_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T c_t(\mathbf{x}) \leq 2R \max \left( \sqrt{2}LR/\sqrt{\alpha}, \sqrt{\text{EGV}_T} \right).$$

### 2.3.2 Online Linear Optimization

Here we consider online linear optimization and present the algorithm and the gradual variation bound for this setting as a special case of proposed algorithm. In particular, we are interested in bounding the regret by the gradual variation

$$\text{EGV}_{T,2}^f = \sum_{t=0}^{T-1} \|\mathbf{f}_{t+1} - \mathbf{f}_t\|_2^2,$$

where  $\mathbf{f}_t, t = 1, \dots, T$  are the linear cost vectors and  $\mathbf{f}_0 = 0$ . Since linear functions are smooth functions that satisfy the inequality in (8) for any positive  $L > 0$ , therefore we can apply Algorithm 4 to online linear optimization with any positive value for  $L$ <sup>3</sup>. The regret bound of Algorithm 4 for online linear optimization is presented in the following corollary.

**Corollary 1** Let  $c_t(\mathbf{x}) = \mathbf{f}_t^\top \mathbf{x}, t = 1, \dots, T$  be a sequence of linear functions. By setting  $\eta = \sqrt{1/(2\text{EGV}_{T,2}^f)}$  and  $L = 1$  in Algorithm 4, then we have

$$\sum_{t=1}^T \mathbf{f}_t^\top \mathbf{x}_t - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T \mathbf{f}_t^\top \mathbf{x} \leq \sqrt{2\text{EGV}_{T,2}^f}.$$

*Remark 3* Note that the regret bound in Corollary 1 is stronger than the regret bound obtained in [16] for online linear optimization due to the fact that the gradual variation is smaller than the total variation.

### 2.3.3 Prediction with Expert Advice

In the problem of prediction with expert advice, the decision vector  $\mathbf{x}$  is a distribution over  $m$  experts, i.e.,  $\mathbf{x} \in \mathcal{P} = \{\mathbf{x} \in \mathbb{R}_+^m : \sum_{i=1}^m x_i = 1\}$ . Let  $\mathbf{f}_t \in \mathbb{R}^m$  denote the costs for  $m$  experts in trial  $t$ . Similar to [16], we would like to bound the regret of prediction from expert advice by the gradual variation defined in infinite norm, i.e.,

$$\text{EGV}_{T,\infty}^f = \sum_{t=0}^{T-1} \|\mathbf{f}_{t+1} - \mathbf{f}_t\|_\infty^2.$$

<sup>3</sup> We simply set  $L = 1$  for online linear optimization and prediction with expert advice.

Since it is a special online linear optimization problem, we can apply Algorithm 4 to obtain a regret bound as in Corollary 1, i.e.,

$$\sum_{t=1}^T \mathbf{f}_t^\top \mathbf{x}_t - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T \mathbf{f}_t^\top \mathbf{x} \leq \sqrt{2\text{EGV}_{T,2}^f} \leq \sqrt{2m\text{EGV}_{T,\infty}^f}.$$

However, the above regret bound scales badly with the number of experts. We can obtain a better regret bound in  $O(\sqrt{\text{EGV}_{T,\infty}^f \ln m})$  by applying the general prox method in Algorithm 5 with  $\omega(\mathbf{x}) = \sum_{i=1}^m x_i \ln x_i$  and  $D(\mathbf{x}, \mathbf{z}) = \sum_{i=1}^m x_i \ln(z_i/x_i)$ . The two updates in Algorithm 5 become

$$\begin{aligned} x_t^i &= \frac{z_{t-1}^i \exp([\eta/L]f_{t-1}^i)}{\sum_{j=1}^m z_{t-1}^j \exp([\eta/L]f_{t-1}^j)}, i = 1, \dots, m \\ z_t^i &= \frac{z_{t-1}^i \exp([\eta/L]f_t^i)}{\sum_{j=1}^m z_{t-1}^j \exp([\eta/L]f_t^j)}, i = 1, \dots, m. \end{aligned}$$

The resulting regret bound is formally stated in the following Corollary.

**Corollary 2** *Let  $c_t(\mathbf{x}) = \mathbf{f}_t^\top \mathbf{x}$ ,  $t = 1, \dots, T$  be a sequence of linear functions in prediction with expert advice. By setting  $\eta = \sqrt{(\ln m)/\text{EGV}_{T,\infty}^f}$ ,  $L = 1$ ,  $\omega(\mathbf{x}) = \sum_{i=1}^m x_i \ln x_i$  and  $D(\mathbf{x}, \mathbf{z}) = \sum_{i=1}^m x_i \ln(x_i/z_i)$  in Algorithm 5, we have*

$$\sum_{t=1}^T \mathbf{f}_t^\top \mathbf{x}_t - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T \mathbf{f}_t^\top \mathbf{x} \leq \sqrt{2\text{EGV}_{T,\infty}^f \ln m}.$$

*Remark 4* By noting the definition of  $\text{EGV}_{T,\infty}^f$ , the regret bound in Corollary 2 is  $O\left(\sqrt{\sum_{t=0}^{T-1} \max_i |f_{t+1}^i - f_t^i| \ln m}\right)$ , which is similar to the regret bound obtained in [16] for prediction with expert advice. However, the definitions of the variation are not exactly the same. In [16], the authors bound the regret of prediction with expert advice by  $O\left(\sqrt{\ln m \max_i \sum_{t=1}^T |f_t^i - \mu_t^i|^2 + \ln m}\right)$ , where the variation is the maximum total variation over all experts. To compare the two regret bounds, we first consider two extreme cases. When the costs of all experts are the same, then the variation in Corollary 2 is a standard gradual variation, while the variation in [16] is a standard total variation. According to the previous analysis, a gradual variation is smaller than a total variation, therefore the regret bound in Corollary 2 is better than that in [16]. In another extreme case when the costs at all iterations of each expert are the same, both regret bounds are constants. More generally, if we assume the maximum total variation is small (say a constant), then  $\sum_{t=0}^{T-1} |f_{t+1}^i - f_t^i|$  is also a constant for any  $i \in [m]$ . By a trivial analysis  $\sum_{t=0}^{T-1} \max_i |f_{t+1}^i - f_t^i| \leq m \max_i \sum_{t=0}^{T-1} |f_{t+1}^i - f_t^i|$ , the regret bound in Corollary 2 might be worse up to a factor  $\sqrt{m}$  than that in [16].

*Remark 5* It was shown in [8], both the regret bounds in Corollary 1 and Corollary 2 are optimal because they match the lower bounds for a special sequence of loss functions. In particular, for online linear optimization if all loss functions but the first  $T_k = \sqrt{\text{EGV}_{T,2}^f}$  are all-0 functions, then the known lower bound  $\Omega(\sqrt{T_k})$  [22] matches the upper bound in Corollary 1. Similarly, for prediction from expert advice if all loss functions but the first  $T'_k = \sqrt{\text{EGV}_{T,\infty}^f}$  are all-0 functions, then the known lower bound  $\Omega(\sqrt{T'_k \ln m})$  [6] matches the upper bound in Corollary 2.

### 2.3.4 Online Strictly Convex Optimization

In this subsection, we present an algorithm to achieve a logarithmic variation bound for online strictly convex optimization [14]. In particular, we assume the cost functions  $c_t(\mathbf{x})$  are not only smooth but also strictly convex defined formally in the following.

**Definition 1** For  $\beta > 0$ , a function  $c(\mathbf{x}) : \mathcal{P} \rightarrow \mathbb{R}$  is  $\beta$ -strictly convex if for any  $\mathbf{x}, \mathbf{z} \in \mathcal{P}$

$$c(\mathbf{x}) \geq c(\mathbf{z}) + \nabla c(\mathbf{z})^\top (\mathbf{x} - \mathbf{z}) + \beta (\mathbf{x} - \mathbf{z})^\top \nabla c(\mathbf{z}) \nabla c(\mathbf{z})^\top (\mathbf{x} - \mathbf{z}) \quad (20)$$

It is known that such a defined strictly convex function include strongly convex function and exponential concave function as special cases as long as the gradient of the function is bounded. To see this, if  $c(\mathbf{x})$  is a  $\beta'$ -strongly convex function with a bounded gradient  $\|\nabla c(\mathbf{x})\|_2 \leq G$ , then

$$\begin{aligned} c(\mathbf{x}) &\geq c(\mathbf{z}) + \nabla c(\mathbf{z})^\top (\mathbf{x} - \mathbf{z}) + \beta' (\mathbf{x} - \mathbf{z})^\top (\mathbf{x} - \mathbf{z}) \\ &\geq c(\mathbf{z}) + \nabla c(\mathbf{z})^\top (\mathbf{x} - \mathbf{z}) + \frac{\beta'}{G^2} (\mathbf{x} - \mathbf{z})^\top \nabla c(\mathbf{z}) \nabla c(\mathbf{z})^\top (\mathbf{x} - \mathbf{z}), \end{aligned}$$

thus  $c(\mathbf{x})$  is a  $(\beta'/G^2)$  strictly convex. Similarly if  $c(\mathbf{x})$  is exp-concave, i.e., there exists  $\alpha > 0$  such that  $h(\mathbf{x}) = \exp(-\alpha c(\mathbf{x}))$  is concave, then  $c(\mathbf{x})$  is a  $\beta = 1/2 \min(1/(4GD), \alpha)$  strictly convex (c.f. Lemma 2 in [14]), where  $D$  is defined as the diameter of the domain.

Therefore, in addition to smoothness and strict convexity we also assume all the cost functions have bounded gradients, i.e.,  $\|\nabla c_t(\mathbf{x})\|_2 \leq G$ . In [8], a logarithmic gradual variation regret bound has been proved. For completeness, we present the algorithm in the framework of general prox method and summarize the results and analysis.

To derive a logarithmic gradual variation bound for online strictly convex optimization, we need to change the Euclidean distance function in Algorithm 4 to a generalized Euclidean distance function. Specifically, at trial  $t$ , we let  $H_t = I + \beta G^2 I + \beta \sum_{\tau=0}^{t-1} \nabla c_\tau(\mathbf{x}_\tau) \nabla c_\tau(\mathbf{x}_\tau)^\top$  and use the generalized Euclidean distance  $D_t(\mathbf{x}, \mathbf{z}) = \frac{1}{2} \|\mathbf{x} - \mathbf{z}\|_{H_t}^2 = \frac{1}{2} (\mathbf{x} - \mathbf{z})^\top H_t (\mathbf{x} - \mathbf{z})$  in updating

$\mathbf{x}_t$  and  $\mathbf{z}_t$ , i.e.,

$$\begin{aligned}\mathbf{x}_t &= \arg \min_{\mathbf{x} \in \mathcal{P}} \left\{ \mathbf{x}^\top \nabla c_{t-1}(\mathbf{x}_{t-1}) + \frac{1}{2} \|\mathbf{x} - \mathbf{z}_{t-1}\|_{H_t}^2 \right\} \\ \mathbf{z}_t &= \arg \min_{\mathbf{x} \in \mathcal{P}} \left\{ \mathbf{x}^\top \nabla c_t(\mathbf{x}_t) + \frac{1}{2} \|\mathbf{x} - \mathbf{z}_{t-1}\|_{H_t}^2 \right\},\end{aligned}\quad (21)$$

To prove the regret bound, we can prove a similar inequality as in Lemma 2 by applying  $\omega(\mathbf{x}) = 1/2\|\mathbf{x}\|_{H_t}^2$ , which is stated as follows

$$\begin{aligned}\nabla c_t(\mathbf{x}_t)^\top (\mathbf{x}_t - \mathbf{z}) &\leq D_t(\mathbf{z}, \mathbf{z}_{t-1}) - D_t(\mathbf{z}, \mathbf{z}_t) \\ &+ \|\nabla c_t(\mathbf{x}_t) - \nabla c_{t-1}(\mathbf{x}_{t-1})\|_{H_t^{-1}}^2 - \frac{1}{2} [\|\mathbf{x}_t - \mathbf{z}_{t-1}\|_{H_t}^2 + \|\mathbf{x}_t - \mathbf{z}_t\|_{H_t}^2].\end{aligned}$$

Then by applying inequality in (20) for strictly convex function, we obtain the following

$$\begin{aligned}c_t(\mathbf{x}_t) - c_t(\mathbf{z}) &\leq D_t(\mathbf{x}, \mathbf{z}_{t-1}) - D_t(\mathbf{x}, \mathbf{z}_t) - \beta \|\mathbf{x}_t - \mathbf{z}\|_{h_t}^2 \\ &+ \|\nabla c_t(\mathbf{x}_t) - \nabla c_{t-1}(\mathbf{x}_{t-1})\|_{H_t^{-1}}^2 - \frac{1}{2} [\|\mathbf{x}_t - \mathbf{z}_{t-1}\|_{H_t}^2 + \|\mathbf{x}_t - \mathbf{z}_t\|_{H_t}^2],\end{aligned}\quad (22)$$

where  $h_t = \nabla c_t(\mathbf{x}_t)^\top \nabla c_t(\mathbf{x}_t)$ . It remains to apply the analysis in [8] to obtain a logarithmic gradual variation bound, which is stated in the following corollary and its proof is deferred to Appendix C.

**Corollary 3** *Let  $c_t(\mathbf{x}), t = 1, \dots, T$  be a sequence of  $\beta$ -strictly convex and  $L$ -smooth functions with gradients bounded by  $G$ . We assume  $8dL^2 \geq 1$ , otherwise we can set  $L = \sqrt{1/(8d)}$ . An algorithm that adopts the updates in (21) has a regret bound by*

$$\sum_{t=1}^T c_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T c_t(\mathbf{x}) \leq \frac{1 + \beta G^2}{2} + \frac{8d}{\beta} \ln \max(16dL^2, \beta \text{EGV}_{T,2}),$$

where  $\text{EGV}_{T,2} = \sum_{t=0}^{T-1} \|\nabla c_{t+1}(\mathbf{x}_t) - \nabla c_t(\mathbf{x}_t)\|_2^2$  and  $d$  is the dimension of  $\mathbf{x} \in \mathcal{P}$ .

### 3 Online Non-Smooth Optimization with Gradual Variation Bound

All the results we have obtained so far rely on the assumption that the cost functions are smooth. Additionally, at the beginning of Section 2, we showed that for general non-smooth functions when the only information presented to the learner is the first order information about the cost functions, it is impossible to obtain a regret bound by gradual variation. However, in this section, we show that a gradual variation bound is achievable for a special class of non-smooth functions that is composed of a smooth component and a non-smooth component.

We consider two categories for the non-smooth component. In the first category, we assume that the non-smooth component is a fixed function and is relatively easy such that the composite gradient mapping can be solved without too much computational overhead compared to gradient mapping. A common example that falls into this category is to consider a non-smooth regularizer. For example, in addition to the basic domain  $\mathcal{P}$ , one would enforce the sparsity constraint on the decision vector  $\mathbf{x}$ , i.e.,  $\|\mathbf{x}\|_0 \leq k < d$ , which is important in feature selection. However, the sparsity constraint  $\|\mathbf{x}\|_0 \leq k$  is a non-convex function, and is usually implemented by adding a  $\ell_1$  regularizer  $\lambda\|\mathbf{x}\|_1$  to the objective function, where  $\lambda > 0$  is a regularization parameter. Therefore, at each iteration the cost function is given by  $c_t(\mathbf{x}) + \lambda\|\mathbf{x}\|_1$ . To prove a regret bound by gradual variation for this type of non-smooth optimization, we first present a simplified version of the general prox method and show that it has the exactly same regret bound as stated in Subsection 2.3.1, and then extend the algorithm to the non-smooth optimization with a fixed non-smooth component.

In the second category, we assume that the non-smooth component can be written as an explicit maximization structure. In general, we consider a time-varying non-smooth component, present a primal-dual prox method, and prove a min-max regret bound by gradual variation. When the non-smooth components are equal across all trials, the usual regret is bounded by the min-max bound plus a variation in the non-smooth component. To see an application of min-max regret bound, we consider the problem of online classification with hinge loss and show that the number of mistakes can be bounded by a variation in sequential examples.

Before moving to the detailed analysis, it is worth mentioning that several pieces of works have proposed algorithms for optimizing the two types of non-smooth functions as described above to obtain an optimal convergence rate of  $O(1/T)$  [24, 23]. Therefore, the existence of a regret bound by gradual variation for these two types of non-smooth optimization does not violate the contradictory argument in Section 2.

### 3.1 Online Non-Smooth Optimization with a Fixed Non-smooth Component

#### 3.1.1 A Simplified Version of Algorithm 5

In this subsection, we present a simplified version of Algorithm 5, which is the foundation for us to develop the algorithm for non-smooth optimization.

The key trick is to replace the domain constraint  $\mathbf{x} \in \mathcal{P}$  with a non-smooth function in the objective. Let  $\delta_{\mathcal{P}}(\mathbf{x})$  denote the indicator function of the domain  $\mathcal{P}$ , i.e.,

$$\delta_{\mathcal{P}}(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in \mathcal{P} \\ \infty, & \text{otherwise} \end{cases}$$

Then the proximal gradient mapping for updating  $\mathbf{x}$  (step 4) in Algorithm 5 is equivalent to

$$\mathbf{x}_t = \arg \min_{\mathbf{x}} \mathbf{x}^\top \nabla c_{t-1}(\mathbf{x}_{t-1}) + \frac{L}{\eta} D(\mathbf{x}, \mathbf{z}_{t-1}) + \delta_{\mathcal{P}}(\mathbf{x}).$$

By the first order optimality condition, there exists a sub-gradient  $\mathbf{v}_t \in \partial \delta_{\mathcal{P}}(\mathbf{x}_t)$  such that

$$\nabla c_{t-1}(\mathbf{x}_{t-1}) + \frac{L}{\eta} (\nabla \omega(\mathbf{x}_t) - \nabla \omega(\mathbf{z}_{t-1})) + \mathbf{v}_t = 0. \quad (23)$$

Thus,  $\mathbf{x}_t$  is equal to

$$\mathbf{x}_t = \arg \min_{\mathbf{x}} \mathbf{x}^\top (\nabla c_{t-1}(\mathbf{x}_{t-1}) + \mathbf{v}_t) + \frac{L}{\eta} D(\mathbf{x}, \mathbf{z}_{t-1}). \quad (24)$$

Then we can change the update for  $\mathbf{z}_t$  to

$$\mathbf{z}_t = \arg \min_{\mathbf{x}} \mathbf{x}^\top (\nabla c_t(\mathbf{x}_t) + \mathbf{v}_t) + \frac{L}{\eta} D(\mathbf{x}, \mathbf{z}_{t-1}). \quad (25)$$

The key ingredient of above update compared to step 6 in Algorithm 5 is that we explicitly use the sub-gradient  $\mathbf{v}_t$  that satisfies the optimality condition for  $\mathbf{x}_t$  instead of solving a domain constrained optimization problem. The advantage of updating  $\mathbf{z}_t$  by (25) is that we can easily compute  $\mathbf{z}_t$  by the first order optimality condition, i.e.,

$$\nabla c_t(\mathbf{x}_t) + \mathbf{v}_t + \frac{L}{\eta} (\nabla \omega(\mathbf{z}_t) - \nabla \omega(\mathbf{z}_{t-1})) = 0. \quad (26)$$

Note that equation (23) indicates  $\mathbf{v}_t = -\nabla c_{t-1}(\mathbf{x}_{t-1}) - \nabla \omega(\mathbf{x}_t) + \nabla \omega(\mathbf{z}_{t-1})$ . By plugging this into (26), we reach to the following simplified update for  $\mathbf{z}_t$ ,

$$\nabla \omega(\mathbf{z}_t) = \nabla \omega(\mathbf{x}_t) + \frac{\eta}{L} (\nabla c_{t-1}(\mathbf{x}_{t-1}) - \nabla c_t(\mathbf{x}_t)).$$

The simplified version of Algorithm 5 is presented in Algorithm 6.

*Remark 6* We make three remarks for Algorithm 6. First, the searching point  $\mathbf{z}_t$  does not necessarily belong to the domain  $\mathcal{P}$ , which is usually not a problem given that the decision point  $\mathbf{x}_t$  is always in  $\mathcal{P}$ . Nevertheless, the update can be followed by a projection step  $\mathbf{z}_t = \min_{\mathbf{x} \in \mathcal{P}} D(\mathbf{x}, \mathbf{z}'_t)$  to ensure the searching point also stay in the domain  $\mathcal{P}$ , where we slightly abuse the notation  $\mathbf{z}'_t$  in  $\nabla \omega(\mathbf{z}'_t) = \nabla \omega(\mathbf{x}_t) + \frac{\eta}{L} (\nabla c_{t-1}(\mathbf{x}_{t-1}) - \nabla c_t(\mathbf{x}_t))$ .

Second, the update in step 6 can be implemented by [6, Chapter 11]:

$$\mathbf{z}_t = \nabla \omega^* (\nabla \omega(\mathbf{x}_t) + \frac{\eta}{L} (\nabla c_{t-1}(\mathbf{x}_{t-1}) - \nabla c_t(\mathbf{x}_t))),$$

---

**Algorithm 6** A Simplified General Prox Method for Online Convex Optimization
 

---

- 1: **Input:**  $\eta > 0, \omega(\mathbf{z})$   
 2: **Initialization:**  $\mathbf{z}_0 = \mathbf{x}_0 = \min_{\mathbf{z} \in \mathcal{P}} \omega(\mathbf{z})$  and  $c_0(\mathbf{x}) = 0$   
 3: **for**  $t = 1, \dots, T$  **do**  
 4:   Predict  $\mathbf{x}_t$  by

$$\mathbf{x}_t = \arg \min_{\mathbf{x} \in \mathcal{P}} \left\{ \mathbf{x}^\top \nabla c_{t-1}(\mathbf{x}_{t-1}) + \frac{L}{\eta} D(\mathbf{x}, \mathbf{z}_{t-1}) \right\}$$

- 5:   Receive cost function  $c_t(\cdot)$  and incur loss  $c_t(\mathbf{x}_t)$   
 6:   Update  $\mathbf{z}_t$  by solving

$$\nabla \omega(\mathbf{z}_t) = \nabla \omega(\mathbf{x}_t) + \frac{\eta}{L} (\nabla c_{t-1}(\mathbf{x}_{t-1}) - \nabla c_t(\mathbf{x}_t)).$$

7: **end for**

---

where  $\omega^*(\cdot)$  is the Legendre-Fenchel conjugate of  $\omega(\cdot)$ . For example, when  $\omega(\mathbf{x}) = 1/2 \|\mathbf{x}\|_2^2$ ,  $\omega^*(\mathbf{x}) = 1/2 \|\mathbf{x}\|_2^2$  and the update for the searching point is given by

$$\mathbf{z}_t = \mathbf{x}_t + (\eta/L)(\nabla c_{t-1}(\mathbf{x}_{t-1}) - \nabla c_t(\mathbf{x}_t));$$

when  $\omega(\mathbf{x}) = \sum_i x_i \ln x_i$ ,  $\omega^*(\mathbf{x}) = \log [\sum_i \exp(x_i)]$  and the update for the searching point can be computed by

$$[\mathbf{z}_t]_i \propto \mathbf{x}_i \exp(\eta/L[\nabla c_{t-1}(\mathbf{x}_{t-1}) - \nabla c_t(\mathbf{x}_t)]), \quad s.t. \quad \sum_i [\mathbf{z}_t]_i = 1.$$

Third, the key inequality in (15) for proving the regret bound still hold for  $\zeta = \nabla c_t(\mathbf{x}_t) + \mathbf{v}_t$ ,  $\xi = \nabla c_{t-1}(\mathbf{x}_{t-1}) + \mathbf{v}_t$  by noting the equivalence between the pairs (24, 25) and (13, 14), which is given below:

$$\begin{aligned} \frac{\eta}{L} (\nabla c_t(\mathbf{x}_t) + \mathbf{v}_t)^\top (\mathbf{x}_t - \mathbf{x}) &\leq D(\mathbf{x}, \mathbf{z}_{t-1}) - D(\mathbf{x}, \mathbf{z}_t) \\ &+ \frac{\gamma^2}{\alpha} \|\nabla c_t(\mathbf{x}_t) - \nabla c_{t-1}(\mathbf{x}_{t-1})\|_*^2 - \frac{\alpha}{2} [\|\mathbf{x}_t - \mathbf{z}_{t-1}\|^2 + \|\mathbf{x} - \mathbf{z}_t\|^2], \forall \mathbf{x} \in \mathcal{P}. \end{aligned}$$

where  $\mathbf{v}_t \in \partial \delta_{\mathcal{P}}(\mathbf{x}_t)$ . As a result, we can apply the same analysis as in the proof of Theorem 3 to obtain the same regret bound in Theorem 4 for Algorithm 6. Note that the above inequality remains valid even if we take a projection step after the update for  $\mathbf{z}'_t$  due to the generalized pythagorean inequality  $D(\mathbf{x}, \mathbf{z}_t) \leq D(\mathbf{x}, \mathbf{z}'_t), \forall \mathbf{x} \in \mathcal{P}$  [6].

### 3.1.2 A Gradual Variation Bound for Online Non-Smooth Optimization

In spirit of Algorithm 6, we present an algorithm for online non-smooth optimization of functions  $c_t(\mathbf{x}) = \widehat{c}_t(\mathbf{x}) + g(\mathbf{x})$  with a regret bound by gradual

---

**Algorithm 7** A General Prox Method for Online Non-Smooth Optimization with a Fixed Non-Smooth Component

---

- 1: **Input:**  $\eta > 0, \omega(\mathbf{z})$   
 2: **Initialization:**  $\mathbf{z}_0 = \mathbf{x}_0 = \min_{\mathbf{z} \in \mathcal{P}} \omega(\mathbf{z})$  and  $c_0(\mathbf{x}) = 0$   
 3: **for**  $t = 1, \dots, T$  **do**  
 4:   Predict  $\mathbf{x}_t$  by

$$\mathbf{x}_t = \arg \min_{\mathbf{x} \in \mathcal{P}} \left\{ \mathbf{x}^\top \nabla c_{t-1}(\mathbf{x}_{t-1}) + \frac{L}{\eta} D(\mathbf{x}, \mathbf{z}_{t-1}) + g(\mathbf{x}) \right\}$$

- 5:   Receive cost function  $c_t(\cdot)$  and incur loss  $c_t(\mathbf{x}_t)$   
 6:   Update  $\mathbf{z}_t$  by

$$\mathbf{z}'_t = \nabla \omega^* \left( \nabla \omega(\mathbf{x}_t) + \frac{\eta}{L} (\nabla c_{t-1}(\mathbf{x}_{t-1}) - \nabla c_t(\mathbf{x}_t)) \right)$$

and  $\mathbf{z}_t = \min_{\mathbf{x} \in \mathcal{P}} D(\mathbf{x}, \mathbf{z}'_t)$

- 7: **end for**
- 

variation  $\text{EGV}_T = \sum_{t=0}^{T-1} \|\nabla c_{t+1}(\mathbf{x}_t) - \nabla c_t(\mathbf{x}_t)\|_*^2$ . The trick is to solve the composite gradient mapping:

$$\mathbf{x}_t = \arg \min_{\mathbf{x} \in \mathcal{P}} \mathbf{x}^\top \nabla c_{t-1}(\mathbf{x}_{t-1}) + \frac{L}{\eta} D(\mathbf{x}, \mathbf{z}_{t-1}) + g(\mathbf{x})$$

and update  $\mathbf{z}_t$  by

$$\nabla \omega(\mathbf{z}_t) = \nabla \omega(\mathbf{x}_t) + \frac{\eta}{L} (\nabla c_{t-1}(\mathbf{x}_{t-1}) - \nabla c_t(\mathbf{x}_t)).$$

Algorithm 7 shows the detailed steps and Corollary 4 states the regret bound, which can be proved similarly.

**Corollary 4** Let  $c_t(\cdot) = \widehat{c}_t(\cdot) + g(\cdot), t = 1, \dots, T$  be a sequence of convex functions where  $\widehat{c}_t(\cdot)$  are  $L$ -smooth continuous w.r.t  $\|\cdot\|$  and  $g(\cdot)$  is a non-smooth function,  $\omega(\mathbf{z})$  be a  $\alpha$ -strongly convex function w.r.t  $\|\cdot\|$ , and  $\text{EGV}_T$  be defined in (19). By setting  $\eta = (1/2) \min \{ \sqrt{\alpha}/\sqrt{2}, LR/\sqrt{\text{EGV}_T} \}$ , we have the following regret bound

$$\sum_{t=1}^T c_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T c_t(\mathbf{x}) \leq 2R \max \left( \sqrt{2}LR/\sqrt{\alpha}, \sqrt{\text{EGV}_T} \right).$$

### 3.2 Online Non-Smooth Optimization with an Explicit Max Structure

In previous subsection, we assume the composite gradient mapping with the non-smooth component can be efficiently solved. Here, we replace this assumption with an explicit max structure of the non-smooth component.

In what follows, we present a primal-dual prox method for such non-smooth cost functions and prove its regret bound. We consider a general setting, where



the non-smooth functions  $c_t(\mathbf{x})$  has the following structure:

$$c_t(\mathbf{x}) = \widehat{c}_t(\mathbf{x}) + \max_{\mathbf{u} \in \mathcal{Q}} \langle A_t \mathbf{x}, \mathbf{u} \rangle - \widehat{\phi}_t(\mathbf{u}), \quad (27)$$

where  $\widehat{c}_t(\mathbf{x})$  and  $\widehat{\phi}_t(\mathbf{u})$  are  $L_1$ -smooth and  $L_2$ -smooth functions, respectively, and  $A_t \in \mathbb{R}^{m \times d}$  is a matrix used to characterize the non-smooth component of  $c_t(\mathbf{x})$  with  $-\widehat{\phi}_t(\mathbf{u})$  by maximization. Similarly, we define a dual cost function  $\phi_t(\mathbf{u})$  as

$$\phi_t(\mathbf{u}) = -\widehat{\phi}_t(\mathbf{u}) + \min_{\mathbf{x} \in \mathcal{P}} \langle A_t \mathbf{x}, \mathbf{u} \rangle + \widehat{c}_t(\mathbf{x}). \quad (28)$$

We refer to  $\mathbf{x}$  as the primal variable and to  $\mathbf{u}$  as the dual variable. To motivate the setup, let us consider online classification with hinge loss  $\ell_t(\mathbf{w}) = \max(0, 1 - y_t \mathbf{w}^\top \mathbf{x}_t)$ , where we slightly abuse the notation  $(\mathbf{x}_t, y_t)$  to denote the attribute and label pair received at trial  $t$ . It is straightforward to see that  $\ell_t(\mathbf{w})$  is a non-smooth function and can be cast into the form in (27) by

$$\ell_t(\mathbf{w}) = \max_{\alpha \in [0,1]} \alpha(1 - y_t \mathbf{w}^\top \mathbf{x}_t) = \max_{\alpha \in [0,1]} -\alpha y_t \mathbf{x}_t^\top \mathbf{w} + \alpha.$$

To present the algorithm and analyze its regret bound, we introduce some notations. Let  $F_t(\mathbf{x}, \mathbf{u}) = \widehat{c}_t(\mathbf{x}) + \langle A_t \mathbf{x}, \mathbf{u} \rangle - \widehat{\phi}_t(\mathbf{u})$ ,  $\omega_1(\mathbf{x})$  be a  $\alpha_1$ -strongly convex function defined on the primal variable  $\mathbf{x}$  w.r.t a norm  $\|\cdot\|_p$  and  $\omega_2(\mathbf{u})$  be a  $\alpha_2$ -strongly convex function defined on the dual variable  $\mathbf{u}$  w.r.t a norm  $\|\cdot\|_q$ . Correspondingly, let  $D_1(\mathbf{x}, \mathbf{z})$  and  $D_2(\mathbf{u}, \mathbf{v})$  denote the induced Bregman distance, respectively. We assume the domains  $\mathcal{P}$ ,  $\mathcal{Q}$  are bounded and matrices  $A_t$  have a bounded norm, i.e.,

$$\begin{aligned} \max_{\mathbf{x} \in \mathcal{P}} \|\mathbf{x}\|_p &\leq R_1, & \max_{\mathbf{u} \in \mathcal{Q}} \|\mathbf{u}\|_q &\leq R_2 \\ \max_{\mathbf{x} \in \mathcal{P}} \omega_1(\mathbf{x}) - \min_{\mathbf{x} \in \mathcal{P}} \omega_1(\mathbf{x}) &\leq M_1 \\ \max_{\mathbf{u} \in \mathcal{Q}} \omega_2(\mathbf{u}) - \min_{\mathbf{u} \in \mathcal{Q}} \omega_2(\mathbf{u}) &\leq M_2 \\ \|A_t\|_{p,q} = \max_{\|\mathbf{x}\|_p \leq 1, \|\mathbf{u}\|_q \leq 1} \mathbf{u}^\top A_t \mathbf{x} &\leq \sigma. \end{aligned} \quad (29)$$

Let  $\|\cdot\|_{p,*}$  and  $\|\cdot\|_{q,*}$  denote the dual norms to  $\|\cdot\|_p$  and  $\|\cdot\|_q$ , respectively. To prove a variational regret bound, we define a gradual variation as follows:

$$\begin{aligned} \text{EGV}_{T,p,q} &= \sum_{t=0}^{T-1} \|\nabla \widehat{c}_{t+1}(\mathbf{x}_t) - \nabla \widehat{c}_t(\mathbf{x}_t)\|_{p,*}^2 + (R_1^2 + R_2^2) \sum_{t=0}^{T-1} \|A_t - A_{t-1}\|_{p,q}^2 \\ &\quad + \sum_{t=0}^{T-1} \|\nabla \widehat{\phi}_{t+1}(\mathbf{u}_t) - \nabla \widehat{\phi}_t(\mathbf{u}_t)\|_{q,*}^2. \end{aligned} \quad (30)$$

Given above notations, Algorithm 8 shows the detailed steps and Theorem 5 states a min-max bound.

---

**Algorithm 8** General Prox Method for Online Non-smooth Optimization with an Explicit Max Structure

---

- 1: **Input:**  $\eta > 0, \omega_1(\mathbf{z}), \omega_2(\mathbf{v})$   
2: **Initialization:**  $\mathbf{z}_0 = \mathbf{x}_0 = \min_{\mathbf{z} \in \mathcal{P}} \omega_1(\mathbf{z}), \mathbf{v}_0 = \mathbf{u}_0 = \min_{\mathbf{v} \in \mathcal{Q}} \omega_2(\mathbf{v})$  and  $\widehat{c}_0(\mathbf{x}) = \widehat{\phi}_0(\mathbf{u}) = 0$   
3: **for**  $t = 1, \dots, T$  **do**  
4:   Update  $\mathbf{u}_t$  by

$$\mathbf{u}_t = \arg \max_{\mathbf{u} \in \mathcal{Q}} \left\{ \mathbf{u}^\top (A_{t-1} \mathbf{x}_{t-1} - \nabla \widehat{\phi}_{t-1}(\mathbf{u}_{t-1})) - \frac{L_2}{\eta} D_2(\mathbf{u}, \mathbf{v}_{t-1}) \right\}$$

- 5:   Predict  $\mathbf{x}_t$  by

$$\mathbf{x}_t = \arg \min_{\mathbf{x} \in \mathcal{P}} \left\{ \mathbf{x}^\top (\nabla \widehat{c}_{t-1}(\mathbf{x}_{t-1}) + A_{t-1}^\top \mathbf{u}_{t-1}) + \frac{L_1}{\eta} D_1(\mathbf{x}, \mathbf{z}_{t-1}) \right\}$$

- 6:   Receive cost function  $c_t(\cdot)$  and incur loss  $c_t(\mathbf{x}_t)$   
7:   Update  $\mathbf{v}_t$  by

$$\mathbf{v}_t = \arg \max_{\mathbf{u} \in \mathcal{Q}} \left\{ \mathbf{u}^\top (A_t \mathbf{x}_t - \nabla \widehat{\phi}_t(\mathbf{u}_t)) - \frac{L_2}{\eta} D_2(\mathbf{u}, \mathbf{v}_{t-1}) \right\}$$

- 8:   Update  $\mathbf{z}_t$  by

$$\mathbf{z}_t = \arg \min_{\mathbf{x} \in \mathcal{P}} \left\{ \mathbf{x}^\top (\nabla \widehat{c}_t(\mathbf{x}_t) + A_t^\top \mathbf{u}_t) + \frac{L_1}{\eta} D_1(\mathbf{x}, \mathbf{z}_{t-1}) \right\}$$

- 9: **end for**
- 

**Theorem 5** Let  $c_t(\mathbf{x}) = \widehat{c}_t(\mathbf{x}) + \max_{\mathbf{u} \in \mathcal{Q}} \langle A_t \mathbf{x}, \mathbf{u} \rangle - \widehat{\phi}_t(\mathbf{u}), t = 1, \dots, T$  be a sequence of non-smooth functions. Assume  $\widehat{c}_t(\mathbf{x}), \widehat{\phi}_t(\mathbf{u})$  are  $L = \max(L_1, L_2)$ -smooth functions and the domain  $\mathcal{P}, \mathcal{Q}$  and  $A_t$  satisfy the boundness condition as in (29). Let  $\omega(\mathbf{x})$  be a  $\alpha_1$ -strongly convex function w.r.t the norm  $\|\cdot\|_p$ ,  $\omega(\mathbf{u})$  be a  $\alpha_2$ -strongly convex function w.r.t. the norm  $\|\cdot\|_q$ , and  $\alpha = \min(\alpha_1, \alpha_2)$ . By setting  $\eta = \min(\sqrt{M_1 + M_2}/(2\sqrt{\text{EGV}_{T,p,q}}), \sqrt{\alpha}/(4\sqrt{\sigma^2 + L^2}))$  in Algorithm 8, then we have

$$\begin{aligned} & \max_{\mathbf{u} \in \mathcal{Q}} \sum_{t=1}^T F_t(\mathbf{x}_t, \mathbf{u}) - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T F_t(\mathbf{x}, \mathbf{u}_t) \\ & \leq 4\sqrt{M_1 + M_2} \max \left( 2\sqrt{\frac{(M_1 + M_2)(\sigma^2 + L^2)}{\alpha}}, \sqrt{\text{EGV}_{T,p,q}} \right). \end{aligned}$$

To facilitate understanding, we break the proof into several Lemmas. The following Lemma is by analogy with Lemma 2.

**Lemma 3** Let  $\theta = \begin{pmatrix} \mathbf{x} \\ \mathbf{u} \end{pmatrix}$  denote a single vector with a norm  $\|\theta\| = \sqrt{\|\mathbf{x}\|_p^2 + \|\mathbf{u}\|_q^2}$  and a dual norm  $\|\theta\|_* = \sqrt{\|\mathbf{x}\|_{p^*}^2 + \|\mathbf{u}\|_{q^*}^2}$ . Let  $\omega(\theta) = \omega_1(\mathbf{x}) + \omega_2(\mathbf{u})$ ,  $D(\theta, \zeta) =$

$D_1(\mathbf{x}, \mathbf{u}) + D_2(\mathbf{z}, \mathbf{v})$ . Then

$$\begin{aligned} & \eta \begin{pmatrix} \nabla_{\mathbf{x}} F_t(\mathbf{x}_t, \mathbf{u}_t) \\ -\nabla_{\mathbf{u}} F_t(\mathbf{x}_t, \mathbf{u}_t) \end{pmatrix}^\top \begin{pmatrix} \mathbf{x}_t - \mathbf{x} \\ \mathbf{u}_t - \mathbf{u} \end{pmatrix} \leq D(\theta, \zeta_{t-1}) - D(\theta, \zeta_t) \\ & + \eta^2 (\|\nabla_{\mathbf{x}} F_t(\mathbf{x}_t, \mathbf{u}_t) - \nabla_{\mathbf{x}} F_{t-1}(\mathbf{x}_{t-1}, \mathbf{u}_{t-1})\|_{p,*}^2) \\ & + \eta^2 (\|\nabla_{\mathbf{u}} F_t(\mathbf{x}_t, \mathbf{u}_t) - \nabla_{\mathbf{u}} F_{t-1}(\mathbf{x}_{t-1}, \mathbf{u}_{t-1})\|_{q,*}^2) \\ & - \frac{\alpha}{2} (\|\mathbf{x}_t - \mathbf{z}_t\|_p^2 + \|\mathbf{u}_t - \mathbf{v}_t\|_q^2 + \|\mathbf{x}_t - \mathbf{z}_{t-1}\|_p^2 + \|\mathbf{u}_t - \mathbf{v}_{t-1}\|_q^2). \end{aligned}$$

*Proof* The updates of  $(\mathbf{x}_t, \mathbf{u}_t)$  in Algorithm 8 can be seen as applying the updates in Lemma 2 with  $\theta_t = \begin{pmatrix} \mathbf{x}_t \\ \mathbf{u}_t \end{pmatrix}$  in place of  $\mathbf{x}$ ,  $\zeta_t = \begin{pmatrix} \mathbf{z}_t \\ \mathbf{v}_t \end{pmatrix}$  in place of  $\mathbf{z}_+$ ,  $\zeta_{t-1} = \begin{pmatrix} \mathbf{z}_{t-1} \\ \mathbf{v}_{t-1} \end{pmatrix}$  in place of  $\mathbf{z}$ . Note that  $\omega(\theta)$  is a  $\alpha = \min(\alpha_1, \alpha_2)$ -strongly convex function with respect to the norm  $\|\theta\|$ . Then applying the results in Lemma 2 we can complete the proof.

Applying the convexity of  $F_t(\mathbf{x}, \mathbf{u})$  in terms of  $\mathbf{x}$  and the concavity of  $F_t(\mathbf{x}, \mathbf{u})$  in terms of  $\mathbf{u}$  to the result in Lemma 3, we have

$$\begin{aligned} & \eta (F_t(\mathbf{x}_t, \mathbf{u}) - F_t(\mathbf{x}, \mathbf{u}_t)) \\ & \leq D_1(\mathbf{x}, \mathbf{z}_{t-1}) - D_1(\mathbf{x}, \mathbf{z}_t) + D_2(\mathbf{u}, \mathbf{v}_{t-1}) - D_2(\mathbf{u}, \mathbf{v}_t) \\ & \quad + \eta^2 \|\nabla \widehat{c}_t(\mathbf{x}_t) - \nabla \widehat{c}_{t-1}(\mathbf{x}_{t-1}) + A_t^\top \mathbf{u}_t - A_{t-1}^\top \mathbf{u}_{t-1}\|_{p,*}^2 \\ & \quad + \eta^2 \|\nabla \widehat{\phi}_t(\mathbf{u}_t) - \nabla \widehat{\phi}_{t-1}(\mathbf{u}_{t-1}) + A_t \mathbf{x}_t - A_{t-1} \mathbf{x}_{t-1}\|_{q,*}^2 \\ & \quad - \frac{\alpha}{2} (\|\mathbf{x}_t - \mathbf{z}_{t-1}\|_p^2 + \|\mathbf{x}_t - \mathbf{z}_t\|_p^2) - \frac{\alpha}{2} (\|\mathbf{u}_t - \mathbf{v}_{t-1}\|_q^2 + \|\mathbf{u}_t - \mathbf{v}_t\|_q^2) \quad (31) \\ & \leq D_1(\mathbf{x}, \mathbf{z}_{t-1}) - D_1(\mathbf{x}, \mathbf{z}_t) + D_2(\mathbf{u}, \mathbf{v}_{t-1}) - D_2(\mathbf{u}, \mathbf{v}_t) \\ & \quad + 2\eta^2 \|\nabla \widehat{c}_t(\mathbf{x}_t) - \nabla \widehat{c}_{t-1}(\mathbf{x}_{t-1})\|_{p,*}^2 + 2\eta^2 \|A_t^\top \mathbf{u}_t - A_{t-1}^\top \mathbf{u}_{t-1}\|_{p,*}^2 \\ & \quad + 2\eta^2 \|\nabla \widehat{\phi}_t(\mathbf{u}_t) - \nabla \widehat{\phi}_{t-1}(\mathbf{u}_{t-1})\|_{q,*}^2 + 2\eta^2 \|A_t \mathbf{x}_t - A_{t-1} \mathbf{x}_{t-1}\|_{q,*}^2 \\ & \quad - \frac{\alpha}{2} (\|\mathbf{x}_t - \mathbf{z}_{t-1}\|_p^2 + \|\mathbf{x}_t - \mathbf{z}_t\|_p^2) - \frac{\alpha}{2} (\|\mathbf{u}_t - \mathbf{v}_{t-1}\|_q^2 + \|\mathbf{u}_t - \mathbf{v}_t\|_q^2) \end{aligned}$$

The following lemma provides tools for proceeding the bound.

**Lemma 4**

$$\|\nabla \widehat{c}_t(\mathbf{x}_t) - \nabla \widehat{c}_{t-1}(\mathbf{x}_{t-1})\|_{p,*}^2 \leq 2\|\nabla \widehat{c}_t(\mathbf{x}_t) - \nabla \widehat{c}_{t-1}(\mathbf{x}_t)\|_{p,*}^2 + 2L^2\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_p^2$$

$$\|\nabla \widehat{\phi}_t(\mathbf{u}_t) - \nabla \widehat{\phi}_{t-1}(\mathbf{u}_{t-1})\|_{q,*}^2 \leq 2\|\nabla \widehat{\phi}_t(\mathbf{u}_t) - \nabla \widehat{\phi}_{t-1}(\mathbf{u}_t)\|_{q,*}^2 + 2L^2\|\mathbf{u}_t - \mathbf{u}_{t-1}\|_q^2$$

$$\|A_t \mathbf{x}_t - A_{t-1} \mathbf{x}_{t-1}\|_{q,*}^2 \leq 2R_1^2 \|A_t - A_{t-1}\|_{p,q}^2 + 2\sigma^2 \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_p^2$$

$$\|A_t^\top \mathbf{u}_t - A_{t-1}^\top \mathbf{u}_{t-1}\|_{p,*}^2 \leq 2R_2^2 \|A_t - A_{t-1}\|_{p,q}^2 + 2\sigma^2 \|\mathbf{u}_t - \mathbf{u}_{t-1}\|_q^2$$

*Proof* We prove the first and the third inequalities. Another two inequalities can be proved similarly.

$$\begin{aligned}
& \|\nabla\widehat{c}_t(\mathbf{x}_t) - \nabla\widehat{c}_{t-1}(\mathbf{x}_{t-1})\|_{p,*}^2 \\
& \leq 2\|\nabla\widehat{c}_t(\mathbf{x}_t) - \nabla\widehat{c}_{t-1}(\mathbf{x}_t)\|_{p,*}^2 + 2\|\nabla\widehat{c}_{t-1}(\mathbf{x}_t) - \nabla\widehat{c}_{t-1}(\mathbf{x}_{t-1})\|_{p,*}^2 \\
& \leq 2\|\nabla\widehat{c}_t(\mathbf{x}_t) - \nabla\widehat{c}_{t-1}(\mathbf{x}_t)\|_{p,*}^2 + 2L^2\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_p^2
\end{aligned}$$

where we use the smoothness of  $\widehat{c}_t(\mathbf{x})$ .

$$\begin{aligned}
\|A_t\mathbf{x}_t - A_{t-1}\mathbf{x}_{t-1}\|_{q,*}^2 & \leq 2\|(A_t - A_{t-1})\mathbf{x}_t\|_{q,*}^2 + 2\|A_{t-1}(\mathbf{x}_t - \mathbf{x}_{t-1})\|_p^2 \\
& \leq 2R_1^2\|A_t - A_{t-1}\|_{p,q}^2 + 2\sigma^2\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_p^2
\end{aligned}$$

**Lemma 5** For any  $\mathbf{x} \in \mathcal{P}$  and  $\mathbf{u} \in \mathcal{Q}$ , we have

$$\begin{aligned}
& \eta(F_t(\mathbf{x}_t, \mathbf{u}) - F_t(\mathbf{x}, \mathbf{u}_t)) \\
& \leq D_1(\mathbf{x}, \mathbf{z}_{t-1}) - D_1(\mathbf{x}, \mathbf{z}_t) + D_2(\mathbf{u}, \mathbf{v}_{t-1}) - D_2(\mathbf{u}, \mathbf{v}_t) \\
& \quad + 4\eta^2 \left( \|\nabla\widehat{c}_t(\mathbf{x}_t) - \nabla\widehat{c}_{t-1}(\mathbf{x}_t)\|_{p,*}^2 + \|\nabla\widehat{\phi}_t(\mathbf{u}_t) - \nabla\widehat{\phi}_{t-1}(\mathbf{u}_t)\|_{q,*}^2 \right. \\
& \quad \left. + (R_1^2 + R_2^2)\|A_t - A_{t-1}\|_{p,q}^2 \right) \\
& \quad + 4\eta^2\sigma^2\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_p^2 + 4\eta^2L^2\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_p^2 - \frac{\alpha}{2}(\|\mathbf{x}_t - \mathbf{z}_t\|_p^2 + \|\mathbf{x}_t - \mathbf{z}_{t-1}\|_p^2) \\
& \quad + 4\eta^2\sigma^2\|\mathbf{u}_t - \mathbf{u}_{t-1}\|_q^2 + 4\eta^2L^2\|\mathbf{u}_t - \mathbf{u}_{t-1}\|_q^2 - \frac{\alpha}{2}(\|\mathbf{u}_t - \mathbf{v}_t\|_q^2 + \|\mathbf{u}_t - \mathbf{v}_{t-1}\|_q^2).
\end{aligned}$$

*Proof* The lemma can be proved by combining the results in Lemma 4 and the inequality in (31).

$$\begin{aligned}
& \eta(F_t(\mathbf{x}_t, \mathbf{u}) - F_t(\mathbf{x}, \mathbf{u}_t)) \\
& \leq D_1(\mathbf{x}, \mathbf{z}_{t-1}) - D_1(\mathbf{x}, \mathbf{z}_t) + D_2(\mathbf{u}, \mathbf{v}_{t-1}) - D_2(\mathbf{u}, \mathbf{v}_t) \\
& \quad + 2\eta^2\|\nabla\widehat{c}_t(\mathbf{x}_t) - \nabla\widehat{c}_{t-1}(\mathbf{x}_{t-1})\|_{p,*}^2 + 2\eta^2\|A_t^\top\mathbf{u}_t - A_{t-1}^\top\mathbf{u}_{t-1}\|_{p,*}^2 \\
& \quad + 2\eta^2\|\nabla\widehat{\phi}_t(\mathbf{u}_t) - \nabla\widehat{\phi}_{t-1}(\mathbf{u}_{t-1})\|_{q,*}^2 + 2\eta^2\|A_t\mathbf{x}_t - A_{t-1}\mathbf{x}_{t-1}\|_{q,*}^2 \\
& \quad - \frac{\alpha}{2}(\|\mathbf{x}_t - \mathbf{z}_{t-1}\|_p^2 + \|\mathbf{x}_t - \mathbf{z}_t\|_p^2) - \frac{\alpha}{2}(\|\mathbf{u}_t - \mathbf{v}_{t-1}\|_q^2 + \|\mathbf{u}_t - \mathbf{v}_t\|_q^2) \\
& \leq D_1(\mathbf{x}, \mathbf{z}_{t-1}) - D_1(\mathbf{x}, \mathbf{z}_t) + D_2(\mathbf{u}, \mathbf{v}_{t-1}) - D_2(\mathbf{u}, \mathbf{v}_t) \\
& \quad + 4\eta^2\|\nabla\widehat{c}_t(\mathbf{x}_t) - \nabla\widehat{c}_{t-1}(\mathbf{x}_t)\|_{p,*}^2 + 4\eta^2L^2\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_p^2 + 4\eta^2\sigma^2\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_p^2 \\
& \quad + 4\eta^2R_1^2\|A_t - A_{t-1}\|_{p,q}^2 + 4\eta^2R_2^2\|A_t - A_{t-1}\|_{p,q}^2 \\
& \quad + 4\eta^2\|\nabla\widehat{\phi}_t(\mathbf{u}_t) - \nabla\widehat{\phi}_{t-1}(\mathbf{u}_t)\|_{q,*}^2 + 4\eta^2L^2\|\mathbf{u}_t - \mathbf{u}_{t-1}\|_q^2 + 4\eta^2\sigma^2\|\mathbf{u}_t - \mathbf{u}_{t-1}\|_q^2 \\
& \quad - \frac{\alpha}{2}(\|\mathbf{x}_t - \mathbf{z}_{t-1}\|_p^2 + \|\mathbf{x}_t - \mathbf{z}_t\|_p^2) - \frac{\alpha}{2}(\|\mathbf{u}_t - \mathbf{v}_{t-1}\|_q^2 + \|\mathbf{u}_t - \mathbf{v}_t\|_q^2)
\end{aligned}$$

$$\begin{aligned}
&\leq D_1(\mathbf{x}, \mathbf{z}_{t-1}) - D_1(\mathbf{x}, \mathbf{z}_t) + D_2(\mathbf{u}, \mathbf{v}_{t-1}) - D_2(\mathbf{u}, \mathbf{v}_t) \\
&\quad + 4\eta^2 \left( \|\nabla \widehat{c}_t(\mathbf{x}_t) - \nabla \widehat{c}_{t-1}(\mathbf{x}_t)\|_{p,*}^2 + \|\nabla \widehat{\phi}_t(\mathbf{u}_t) - \nabla \widehat{\phi}_{t-1}(\mathbf{u}_t)\|_{q,*}^2 \right. \\
&\quad \left. + (R_1^2 + R_2^2) \|A_t - A_{t-1}\|_{p,q}^2 \right) \\
&\quad + 4\eta^2 \sigma^2 \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_p^2 + 4\eta^2 L^2 \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_p^2 - \frac{\alpha}{2} (\|\mathbf{x}_t - \mathbf{z}_t\|_p^2 + \|\mathbf{x}_t - \mathbf{z}_{t-1}\|_p^2) \\
&\quad + 4\eta^2 \sigma^2 \|\mathbf{u}_t - \mathbf{u}_{t-1}\|_q^2 + 4\eta^2 L^2 \|\mathbf{u}_t - \mathbf{u}_{t-1}\|_q^2 - \frac{\alpha}{2} (\|\mathbf{u}_t - \mathbf{v}_t\|_q^2 + \|\mathbf{u}_t - \mathbf{v}_{t-1}\|_q^2).
\end{aligned}$$

*Proof (of Theorem 5)* Taking summation the inequalities in Lemma 5 over  $t = 1, \dots, T$ , applying the inequality in (17) twice and using  $\eta \leq \sqrt{\alpha}/(4\sqrt{\sigma^2 + L^2})$ , we have

$$\begin{aligned}
&\sum_{t=1}^T F_t(\mathbf{x}_t, \mathbf{u}) - \sum_{t=1}^T F_t(\mathbf{x}, \mathbf{u}) \leq 4\eta \text{EGV}_{T,p,q} + \frac{M_1 + M_2}{\eta} \\
&= 4\sqrt{M_1 + M_2} \max \left( 2\sqrt{\frac{(M_1 + M_2)(\sigma^2 + L^2)}{\alpha}}, \sqrt{\text{EGV}_{T,p,q}} \right). \quad (32)
\end{aligned}$$

We complete the proof by using  $\mathbf{x}^* = \arg \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T F_t(\mathbf{x}, \mathbf{u})$  and  $\mathbf{u}^* = \arg \max_{\mathbf{u} \in \mathcal{Q}} \sum_{t=1}^T F_t(\mathbf{x}, \mathbf{u})$ .

As an immediate result of Theorem 5, the following Corollary states a regret bound for non-smooth optimization with a fixed non-smooth component that can be written as a max structure, i.e.,  $c_t(\mathbf{x}) = \widehat{c}_t(\mathbf{x}) + [g(\mathbf{x}) = \max_{\mathbf{u} \in \mathcal{Q}} \langle A\mathbf{x}, \mathbf{u} \rangle - \widehat{\phi}(\mathbf{u})]$ .

**Corollary 5** *Let  $c_t(\mathbf{x}) = \widehat{c}_t(\mathbf{x}) + g(\mathbf{x}), t = 1, \dots, T$  be a sequence of non-smooth functions, where  $g(\mathbf{x}) = \max_{\mathbf{u} \in \mathcal{Q}} \langle A\mathbf{x}, \mathbf{u} \rangle - \widehat{\phi}(\mathbf{u})$ , and the gradual variation  $\text{EGV}_T$  be defined in (19) w.r.t the dual norm  $\|\cdot\|_{p,*}$ . Assume  $\widehat{c}_t(\mathbf{x})$  are  $L$ -smooth functions w.r.t  $\|\cdot\|$ , the domain  $\mathcal{P}, \mathcal{Q}$  and  $A$  satisfy the boundedness condition as in (29). If we set  $\eta = \min(\sqrt{M_1 + M_2}/(2\sqrt{\text{EGV}_T}), \sqrt{\alpha}/(4\sqrt{\sigma^2 + L^2}))$  in Algorithm 8, then we have the following regret bound*

$$\begin{aligned}
&\sum_{t=1}^T c_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T c_t(\mathbf{x}) \\
&\leq 4\sqrt{M_1 + M_2} \max \left( 2\sqrt{\frac{(M_1 + M_2)(\sigma^2 + L^2)}{\alpha}}, \sqrt{\text{EGV}_T} \right) + V(g, \mathbf{x}_{1:T}),
\end{aligned}$$

where  $\widehat{\mathbf{x}}_T = \sum_{t=1}^T \mathbf{x}_t/T$  and  $V(g, \mathbf{x}_{1:T}) = \sum_{t=1}^T |g(\mathbf{x}_t) - g(\widehat{\mathbf{x}}_T)|$  measures the variation in the non-smooth component.

*Proof* In the case of fixed non-smooth component, the gradual variation defined in (30) reduces the one defined in (19) w.r.t the dual norm  $\|\cdot\|_{p,*}$ . By

using the bound in (32) and noting that  $c_t(\mathbf{x}) = \max_{\mathbf{u} \in \mathcal{Q}} F_t(\mathbf{x}, \mathbf{u}) \geq F_t(\mathbf{x}, \mathbf{u}_t)$ , we have

$$\sum_{t=1}^T \left( \widehat{c}_t(\mathbf{x}_t) + \langle A\mathbf{x}_t, \mathbf{u} \rangle - \widehat{\phi}(\mathbf{u}) \right) \leq \sum_{t=1}^T c_t(\mathbf{x}) + 4\sqrt{M_1 + M_2} \max \left( 2\sqrt{\frac{(M_1 + M_2)(\sigma^2 + L^2)}{\alpha}}, \sqrt{\text{EGV}_T} \right).$$

Therefore

$$\sum_{t=1}^T \left( \widehat{c}_t(\mathbf{x}_t) + g(\widehat{\mathbf{x}}_T) \right) \leq \sum_{t=1}^T c_t(\mathbf{x}) + 4\sqrt{M_1 + M_2} \max \left( 2\sqrt{\frac{(M_1 + M_2)(\sigma^2 + L^2)}{\alpha}}, \sqrt{\text{EGV}_T} \right).$$

We complete the proof by complementing  $\widehat{c}_t(\mathbf{x}_t)$  with  $g(\mathbf{x}_t)$  to obtain  $c_t(\mathbf{x}_t)$  and moving the additional term  $\sum_{t=1}^T (g(\widehat{\mathbf{x}}_T) - g(\mathbf{x}_t))$  to the right hand side.

*Remark 7* Note that the regret bound in Corollary 5 has an additional term  $V(g, \mathbf{x}_{1:T})$  compared to the regret bound in Corollary 4, which constitutes a tradeoff between the reduced computational cost in solving a composite gradient mapping.

To see an application of Theorem 5 to an online non-smooth optimization with time-varying non-smooth components, let us consider the example of online classification with hinge loss. At each trial, upon receiving an example  $\mathbf{x}_t$ , we need to make a prediction based on the current model  $\mathbf{w}_t$ , i.e.,  $\widehat{y}_t = \mathbf{w}_t^\top \mathbf{x}_t$ , then we receive the true label of  $\mathbf{x}_t$  denoted by  $y_t \in \{+1, -1\}$ . The goal is to minimize the total number of mistakes across the time line  $M_T = \sum_{t=1}^T I(\widehat{y}_t y_t \leq 0)$ . Here we are interested in a scenario that the data sequence  $(\mathbf{x}_t, y_t), t = 1, \dots, T$  has a small gradual variation in terms of  $y_t \mathbf{x}_t$ . To obtain such a gradual variation based mistake bound, we can apply Algorithm 8. For the purpose of deriving the mistake bound, we need to make a small change to Algorithm 8. At the beginning of each trial, we first make a prediction  $\widehat{y}_t = \mathbf{w}_t^\top \mathbf{x}_t$ , and if we make a mistake  $I(\widehat{y}_t y_t \leq 0)$  then we proceed to update the auxiliary primal-dual pair  $(\mathbf{w}'_t, \beta_t)$  similar to  $(\mathbf{z}_t, \mathbf{v}_t)$  in Algorithm 8 and the primal-dual pair  $(\mathbf{w}_{t+1}, \alpha_{t+1})$  similar to  $(\mathbf{x}_{t+1}, \mathbf{u}_{t+1})$  in Algorithm 8, which are given explicitly as follows:

$$\begin{aligned} \beta_t &= \prod_{[0,1]} (\beta_{t-1} + \eta(1 - \mathbf{w}_t y_t \mathbf{x}_t)), & \mathbf{w}'_t &= \prod_{\|\mathbf{w}\|_2 \leq R} (\mathbf{w}'_{t-1} + \eta \alpha_t y_t \mathbf{x}_t) \\ \alpha_{t+1} &= \prod_{[0,1]} (\beta_t + \eta(1 - \mathbf{w}_t y_t \mathbf{x}_t)), & \mathbf{w}_{t+1} &= \prod_{\|\mathbf{w}\|_2 \leq R} (\mathbf{w}'_t + \eta \alpha_t y_t \mathbf{x}_t). \end{aligned}$$

Without loss of generality, we let  $(\mathbf{x}_t, y_t), t = 1, \dots, M_T$  denote the examples that are predicted incorrectly. The  $F_t$  function is written as  $F_t(\mathbf{w}, \alpha) = \alpha(1 -$

$y_t \mathbf{w}_t^\top \mathbf{x}_t$ ). Then for a total sequence of  $T$  examples, we have the following bound by assuming  $\|\mathbf{x}_t\|_2 \leq 1$  and  $\eta \leq 1/2\sqrt{2}$

$$\sum_{t=1}^{M_T} F_t(\mathbf{w}_t, \alpha) \leq \sum_{t=1}^{M_T} \ell(y_t \mathbf{w}_t^\top \mathbf{x}_t) + \eta \sum_{t=0}^{M_T-1} (R^2 + 1) \|y_{t+1} \mathbf{x}_{t+1} - y_t \mathbf{x}_t\|_2^2 + \frac{R^2 + \alpha^2}{2\eta}.$$

Since  $y_t \mathbf{w}_t^\top \mathbf{x}_t$  is less than 0 for the incorrectly predicted examples, if we set  $\alpha = 1$  in the above inequality, we have

$$\begin{aligned} M_T &\leq \sum_{t=1}^{M_T} \ell(y_t \mathbf{w}_t^\top \mathbf{x}_t) + \eta \sum_{t=0}^{M_T-1} (R^2 + 1) \|y_{t+1} \mathbf{x}_{t+1} - y_t \mathbf{x}_t\|_2^2 + \frac{R^2 + 1}{2\eta} \\ &\leq \sum_{t=1}^{M_T} \ell(y_t \mathbf{w}_t^\top \mathbf{x}_t) + \sqrt{2}(R^2 + 1) \max(2, \sqrt{\text{EGV}_{T,2}}). \end{aligned}$$

which results in a gradual variational mistake bound, where  $\sqrt{\text{EGV}_{T,2}}$  measures the gradual variation in the incorrectly predicted examples. To end the discussion, we note that one may find applications of a small gradual variation of  $y_t \mathbf{x}_t$  in time series classification. For instance, if  $\mathbf{x}_t$  represent some medical measurements of a person and  $y_t$  indicates whether the person observes a disease, since the health conditions usually change slowly then it is expected that the gradual variation of  $y_t \mathbf{x}_t$  is small. Similarly, if  $\mathbf{x}_t$  are some sensor measurements of an equipment and  $y_t$  indicates whether the equipment fails or not, we would also observe a small gradual variation of the sequence  $y_t \mathbf{x}_t$  during a time period.

#### 4 Variation Bound for Online Bandit Convex Optimization

Online convex optimization becomes more challenging when the learner only receives partial feedback about the cost functions. One common scenario of partial feedback is that the learner only receives the cost  $c(\mathbf{x}_t)$  at the predicted point  $\mathbf{x}_t$  but without observing the entire cost function  $c_t(\mathbf{x})$ . This setup is usually referred to as bandit setting, and the related online learning problem is called online bandit convex optimization.

Before describing our result for bandit setting, we give a quick review of the literature on online bandit convex optimization. [11] presented a modified gradient descent approach for online bandit convex optimization that attains  $O(T^{3/4})$  regret bound. The key idea of their algorithm is to compute the stochastic approximation of the gradient of cost functions by single point evaluation of the cost functions. This regret bound is later improved to  $O(T^{2/3})$  [4, 9] for online bandit linear optimization. More recently, [10] proposed an inefficient algorithm for online bandit linear optimization with the optimal regret bound  $O(\text{poly}(d)\sqrt{T})$  based on multi-armed bandit algorithm. The key disadvantage of [10] is that it is not computationally efficient. Abernethy, Hazan, and Rakhlin [1] presented an efficient randomized algorithm with an optimal

regret bound  $O(\text{poly}(d)\sqrt{T})$  that exploits the properties of self-concordant barrier regularization. For online bandit convex optimization, [2] proposed optimal algorithms in a multi-point bandit setting, in which multiple points can be queried for the cost values. With multiple queries, they show that the modified online gradient descent algorithm [2] can give an  $O(\sqrt{T})$  expected regret bound.

Recently Hazan et al [15] extended the FTRL algorithm to online bandit *linear* optimization and obtained a variation-based regret bound in the form of  $O(\text{poly}(d)\sqrt{\text{VAR}_T \log(T)} + \text{poly}(d \log(T)))$ , where  $\text{VAR}_T$  is the total variation of the cost vectors. We continue this line of work by proposing algorithms for general online bandit convex optimization with a variation-based regret bound. We present a deterministic algorithm for online bandit convex optimization by extending Algorithm 4 to a multi-point bandit setting, and prove the variation-based regret bound, which is optimal when the variation is independent of the number of trials. In our bandit setting, we assume we are allowed to query  $d+1$  points around the decision point  $\mathbf{x}_t$ . We pose the problem of further reducing the number of point evaluations to a constant number that is independent of the dimension as an open problem.

#### 4.1 A Deterministic Algorithm for Online Bandit Convex Optimization

To develop a variation bound for online bandit convex optimization, we follow [2] by considering the multi-point bandit setting, where at each trial the player is allowed to query the cost functions at multiple points. We propose a deterministic algorithm to compete against the *completely adaptive* adversary that can choose the cost function  $c_t(\mathbf{x})$  with the knowledge of  $\mathbf{x}_1, \dots, \mathbf{x}_t$ . To approximate the gradient  $\nabla c_t(\mathbf{x}_t)$ , we query the cost function to obtain the cost values at  $c_t(\mathbf{x}_t)$ , and  $c_t(\mathbf{x}_t + \delta \mathbf{e}_i)$ ,  $i = 1, \dots, d$ , where  $\mathbf{e}_i$  is the  $i$ th standard base in  $\mathbb{R}^d$ . Then we compute the estimate of the gradient  $\nabla c_t(\mathbf{x}_t)$  by

$$g_t(\mathbf{x}_t) = \frac{1}{\delta} \sum_{i=1}^d (c_t(\mathbf{x}_t + \delta \mathbf{e}_i) - c_t(\mathbf{x}_t)) \mathbf{e}_i.$$

It can be shown that [2], under the smoothness assumption in (8),

$$\|g_t(\mathbf{x}_t) - \nabla c_t(\mathbf{x}_t)\|_2 \leq \frac{\sqrt{d}L\delta}{2}. \quad (33)$$

To prove the regret bound, besides the smoothness assumption of the cost functions, and the boundness assumption about the domain  $\mathcal{P} \subseteq \mathcal{B}$ , we further assume that (i) there exists  $r \leq 1$  such that  $r\mathcal{B} \subseteq \mathcal{P} \subseteq \mathcal{B}$ , and (ii) the cost function themselves are Lipschitz continuous, i.e., there exists a constant  $G$  such that

$$|c_t(\mathbf{x}) - c_t(\mathbf{z})| \leq G\|\mathbf{x} - \mathbf{z}\|_2, \forall \mathbf{x}, \mathbf{z} \in \mathcal{P}, \forall t.$$



**Algorithm 9** Deterministic Online Bandit Convex Optimization

- 
- 1: **Input:**  $\eta, \alpha, \delta > 0$
  - 2: **Initialization:**  $\mathbf{z}_0 = \mathbf{0}$  and  $c_0(\mathbf{x}) = 0$
  - 3: **for**  $t = 1, \dots, T$  **do**
  - 4:   Compute  $\mathbf{x}_t$  by

$$\mathbf{x}_t = \arg \min_{\mathbf{x} \in (1-\alpha)\mathcal{P}} \left\{ \mathbf{x}^\top g_{t-1}(\mathbf{x}_{t-1}) + \frac{G}{2\eta} \|\mathbf{x} - \mathbf{z}_{t-1}\|_2^2 \right\}$$

- 5:   Observe  $c_t(\mathbf{x}_t), c_t(\mathbf{x}_t + \delta \mathbf{e}_i), i = 1, \dots, d$
- 6:   Update  $\mathbf{z}_t$  by

$$\mathbf{z}_t = \arg \min_{\mathbf{x} \in (1-\alpha)\mathcal{P}} \left\{ \mathbf{x}^\top g_t(\mathbf{x}_t) + \frac{G}{2\eta} \|\mathbf{x} - \mathbf{z}_{t-1}\|_2^2 \right\}$$

- 7: **end for**

---

For our purpose, we define another gradual variation of cost functions by

$$\text{EGV}_T^c = \sum_{t=0}^{T-1} \max_{\mathbf{x} \in \mathcal{P}} |c_{t+1}(\mathbf{x}) - c_t(\mathbf{x})|. \quad (34)$$

Unlike the gradual variation defined in (7) that uses the gradient of the cost functions, the gradual variation in (34) is defined according to the values of cost functions. The reason why we bound the regret of Algorithm 8 by the gradual variation defined in (34) by the values of the cost functions rather than the one defined in (7) by the gradient of the cost functions is that in the bandit setting, we only have point evaluations of the cost functions. The following theorem states the regret bound for Algorithm 9.

**Theorem 6** *Let  $c_t(\cdot), t = 1, \dots, T$  be a sequence of  $G$ -Lipschitz continuous convex functions, and their gradients are  $L$ -Lipschitz continuous. By setting*

$$\delta = \sqrt{\frac{4d \max(\sqrt{2}G, \sqrt{\text{EGV}_T^c})}{(\sqrt{d}L + G(1 + 1/r))T}}, \quad \eta = \frac{\delta}{4d} \min \left\{ \frac{1}{\sqrt{2}}, \frac{G}{\sqrt{\text{EGV}_T^c}} \right\}, \quad \text{and } \alpha = \delta/r,$$

*we have the following regret bound for Algorithm 9*

$$\sum_{t=1}^T c_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T c_t(\mathbf{x}) \leq 4 \sqrt{\max(\sqrt{2}G, \sqrt{\text{EGV}_T^c}) d(dL + G/r)T}.$$

*Remark 8* Similar to the regret bound in [2](Theorem 9), Algorithm 9 also gives the optimal regret bound  $O(\sqrt{T})$  when the variation is independent of the number of trials. Our regret bound has a better dependence on  $d$  (i.e.,  $d$ ) compared with the regret bound in [2] (i.e.,  $d^2$ ).

*Proof* Let  $h_t(\mathbf{x}) = c_t(\mathbf{x}) + (g_t(\mathbf{x}_t) - \nabla c_t(\mathbf{x}_t))^\top \mathbf{x}$ . It is easy seen that  $\nabla h_t(\mathbf{x}_t) = g_t(\mathbf{x}_t)$ . Followed by Lemma 2, we have for any  $\mathbf{z} \in (1 - \alpha)\mathcal{P}$ ,

$$\begin{aligned} \frac{\eta}{G} \nabla h_t(\mathbf{x}_t)^\top (\mathbf{x}_t - \mathbf{z}) &\leq \frac{1}{2} (\|\mathbf{z} - \mathbf{z}_{t-1}\|_2^2 - \|\mathbf{z} - \mathbf{z}_t\|_2^2) + \frac{\eta^2}{G^2} \|g_t(\mathbf{x}_t) - g_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\ &- \frac{1}{2} (\|\mathbf{x}_t - \mathbf{z}_{t-1}\|_2^2 + \|\mathbf{x}_t - \mathbf{z}_t\|_2^2) \end{aligned}$$

Taking summation over  $t = 1, \dots, T$ , we have,

$$\begin{aligned}
& \sum_{t=1}^T \frac{\eta}{G} \nabla h_t(\mathbf{x}_t)^\top (\mathbf{x}_t - \mathbf{z}) \leq \frac{\|\mathbf{z} - \mathbf{z}_0\|_2^2}{2} + \sum_{t=1}^T \frac{\eta^2}{G^2} \|g_t(\mathbf{x}_t) - g_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\
& - \sum_{t=1}^T \frac{1}{2} (\|\mathbf{x}_t - \mathbf{z}_{t-1}\|_2^2 + \|\mathbf{x}_t - \mathbf{z}_t\|_2^2) \\
& \leq \frac{\|\mathbf{z} - \mathbf{z}_0\|_2^2}{2} + \sum_{t=1}^T \frac{\eta^2}{G^2} \|g_t(\mathbf{x}_t) - g_{t-1}(\mathbf{x}_{t-1})\|_2^2 - \sum_{t=1}^T \frac{1}{4} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \\
& \leq \frac{1}{2} + \sum_{t=1}^T \frac{\eta^2}{G^2} \|g_t(\mathbf{x}_t) - g_{t-1}(\mathbf{x}_{t-1})\|_2^2 - \sum_{t=1}^T \frac{1}{4} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \\
& \leq \frac{1}{2} + \sum_{t=1}^T \frac{2\eta^2}{G^2} \|g_t(\mathbf{x}_t) - g_t(\mathbf{x}_{t-1})\|_2^2 + \sum_{t=1}^T \frac{2\eta^2}{G^2} \|g_t(\mathbf{x}_{t-1}) - g_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\
& - \sum_{t=1}^T \frac{1}{4} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \\
& \leq \frac{1}{2} + \frac{2\eta^2}{\delta^2 G^2} \sum_{t=1}^T \left\| \sum_{i=1}^d (c_t(\mathbf{x}_t + \delta \mathbf{e}_i) - c_t(\mathbf{x}_{t-1} + \delta \mathbf{e}_i)) \mathbf{e}_i - (c_t(\mathbf{x}_t) - c_t(\mathbf{x}_{t-1})) \mathbf{e}_i \right\|_2^2 \\
& + \frac{2\eta^2}{\delta^2 G^2} \sum_{t=1}^T \left\| \sum_{i=1}^d (c_t(\mathbf{x}_{t-1} + \delta \mathbf{e}_i) - c_{t-1}(\mathbf{x}_{t-1} + \delta \mathbf{e}_i)) \mathbf{e}_i - (c_t(\mathbf{x}_{t-1}) - c_{t-1}(\mathbf{x}_{t-1})) \mathbf{e}_i \right\|_2^2 \\
& - \sum_{t=1}^T \frac{1}{4} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2.
\end{aligned}$$

where the second inequality follows (17). Next, we bound the middle two terms in R.H.S of the above inequality.

$$\begin{aligned}
& \sum_{t=1}^T \left\| \sum_{i=1}^d (c_t(\mathbf{x}_t + \delta \mathbf{e}_i) - c_t(\mathbf{x}_{t-1} + \delta \mathbf{e}_i)) \mathbf{e}_i - (c_t(\mathbf{x}_t) - c_t(\mathbf{x}_{t-1})) \mathbf{e}_i \right\|_2^2 \\
& \leq \sum_{t=1}^T 2d \sum_{i=1}^d (|c_t(\mathbf{x}_t + \delta \mathbf{e}_i) - c_t(\mathbf{x}_{t-1} + \delta \mathbf{e}_i)|^2 + |c_t(\mathbf{x}_t) - c_t(\mathbf{x}_{t-1})|^2) \\
& \leq \sum_{t=1}^T 4d^2 G^2 \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2,
\end{aligned}$$

and

$$\begin{aligned}
& \sum_{t=1}^T \left\| \sum_{i=1}^d (c_t(\mathbf{x}_{t-1} + \delta \mathbf{e}_i) - c_{t-1}(\mathbf{x}_{t-1} + \delta \mathbf{e}_i)) \mathbf{e}_i - (c_t(\mathbf{x}_{t-1}) - c_{t-1}(\mathbf{x}_{t-1})) \mathbf{e}_i \right\|_2^2 \\
& \leq \sum_{t=1}^T 2d \sum_{i=1}^d (|c_t(\mathbf{x}_{t-1} + \delta \mathbf{e}_i) - c_{t-1}(\mathbf{x}_{t-1} + \delta \mathbf{e}_i)|^2 + |c_t(\mathbf{x}_{t-1}) - c_{t-1}(\mathbf{x}_{t-1})|^2) \\
& \leq \sum_{t=1}^T 4d^2 \max_{\mathbf{x} \in \mathcal{P}} |c_t(\mathbf{x}) - c_{t-1}(\mathbf{x})|^2.
\end{aligned}$$

Then we have

$$\begin{aligned}
& \sum_{t=1}^T \frac{\eta}{G} \nabla h_t(\mathbf{x}_t)^\top (\mathbf{x}_t - \mathbf{z}) \leq \frac{1}{2} + \frac{8d^2 \eta^2}{\delta^2} \sum_{t=1}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + \frac{8d^2 \eta^2}{\delta^2 G^2} \sum_{t=1}^T \max_{\mathbf{x} \in \mathcal{P}} |c_t(\mathbf{x}) - c_{t-1}(\mathbf{x})|^2 \\
& - \sum_{t=1}^T \frac{1}{4} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \\
& \leq \frac{1}{2} + \frac{8d^2 \eta^2}{\delta^2 G^2} \sum_{t=1}^T \max_{\mathbf{x} \in \mathcal{P}} |c_t(\mathbf{x}) - c_{t-1}(\mathbf{x})|^2
\end{aligned}$$

where the last inequality follows that  $\eta \leq \delta/(4\sqrt{2}d)$ . Then by using the convexity of  $h_t(\mathbf{x})$  and dividing both sides by  $\eta/G$ , we have

$$\sum_{t=1}^T h_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{P}} h_t((1-\alpha)\mathbf{x}) \leq \frac{G}{2\eta} + \frac{8\eta d^2}{G\delta^2} \text{EVAR}_T^c \leq \frac{4d}{\delta} \max\left(\sqrt{2}G, \sqrt{\text{EVAR}_T^c}\right)$$

Following the the proof of Theorem 8 in [2], we have

$$\begin{aligned}
& \sum_{t=1}^T c_t(\mathbf{x}_t) - \sum_{t=1}^T c_t(\mathbf{x}) \leq \sum_{t=1}^T h_t(\mathbf{x}_t) - \sum_{t=1}^T h_t(\mathbf{x}) + \sum_{t=1}^T c_t(\mathbf{x}_t) - h_t(\mathbf{x}_t) - c_t(\mathbf{x}) + h_t(\mathbf{x}) \\
& \leq \sum_{t=1}^T h_t(\mathbf{x}_t) - \sum_{t=1}^T h_t(\mathbf{x}) + \sum_{t=1}^T (g_t(\mathbf{x}_t) - \nabla c_t(\mathbf{x}_t))^\top (\mathbf{x} - \mathbf{x}_t) \\
& \leq \sum_{t=1}^T h_t(\mathbf{x}_t) - \sum_{t=1}^T h_t(\mathbf{x}) + \sqrt{d}L\delta T
\end{aligned}$$

where the last inequality follows from the following facts:

$$\begin{aligned}
& \|g_t(\mathbf{x}_t) - \nabla c_t(\mathbf{x}_t)\|_2 \leq \frac{\sqrt{d}L\delta}{2} \\
& \|\mathbf{x} - \mathbf{x}_t\| \leq 2
\end{aligned}$$

Then we have

$$\sum_{t=1}^T c_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T c_t((1-\alpha)\mathbf{x}) \leq \frac{4d}{\delta} \max\left(\sqrt{2}G, \sqrt{\text{EVAR}_T^c}\right) + \sqrt{d}L\delta T$$

By the Lipschitz continuity of  $c_t(\mathbf{x})$ , we have

$$\sum_{t=1}^T c_t((1-\alpha)\mathbf{x}) \leq \sum_{t=1}^T c_t(\mathbf{x}) + G\alpha T$$

The we get

$$\sum_{t=1}^T c_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T c_t(\mathbf{x}) \leq \frac{4d}{\delta} \max\left(\sqrt{2}G, \sqrt{\text{EVAR}_T^c}\right) + \delta\sqrt{d}LT + \alpha GT$$

Plugging the stated values of  $\delta$  and  $\alpha$  completes the proof.

## 5 Conclusions and Discussions

In this paper, we proposed two novel algorithms for online convex optimization that bound the regret by the gradual variation of cost functions. The first algorithm is an improvement of the FTRL algorithm, and the second algorithm is based on the prox method. Both algorithms maintain two sequence of solution points, a sequence of decision points and a sequence of searching points, and share the same order of regret bound up to a constant. The prox method only requires to keep tracking of a single gradient of each cost function, while the improved FTRL algorithm needs to evaluate the gradient of each cost function at two points and maintain a sum of up-to-date gradients of the cost functions. We also extended the prox method to a general framework that yields a gradual variation bound with the variation defined by a general norm and also to a multi-point bandit setting. We discuss several special cases, including online linear optimization, predict with expert advice, online strictly convex optimization. We also developed a simplified prox method using a composite gradient mapping for non-smooth optimization with a fixed non-smooth component and a primal-dual prox method for non-smooth optimization with the non-smooth component written as a max structure.

Finally, it has been brought to our attention that as we prepare the final version of the present work, Chiang et al. [7] published a paper at the Conference on Learning Theory in 2013, which extends the prox method into a two-point bandit setting and achieves a similar regret bound in expectation as that in the full setting, i.e.,  $O(d^2\sqrt{\text{EGV}_{T,2}\ln T})$  for smooth convex cost functions and  $O(d^2\ln(\text{EGV}_{T,2} + \ln T))$  for smooth and strongly convex cost functions, where  $\text{EGV}_{T,2}$  is the gradual variation defined on the gradients of the cost functions. We would like to make a thought-provoking comparison between our regret bound and their regret bound for online bandit convex optimization with smooth cost functions. First, the gradual variation in our

bandit setting is defined on the values of the cost functions in contrast to that defined on the gradients of the cost functions. Second, we query the cost function  $d$  times in contrast to 2 times in their algorithms, and as a tradeoff our regret bound has a better dependence on the number of dimensions (i.e.,  $O(d)$ ) than that (i.e.,  $O(d^2)$ ) of their regret bound. Third, our regret bound has an annoying factor of  $\sqrt{T}$  in comparison with  $\sqrt{\ln T}$  in theirs. Therefore, some open problems are how to achieve a lower order of dependence on  $d$  than  $d^2$  in the two-point bandit setting, and how to remove the factor of  $\sqrt{T}$  while keeping a small order of dependence on  $d$  in our multi-point bandit setting; and studying the two different types of gradual variations for bandit settings is a future work as well.

**Acknowledgements** We thank the reviewers for their immensely helpful and thorough comments.

## References

1. Abernethy, J., Hazan, E., Rakhlin, A.: Competing in the dark: An efficient algorithm for bandit linear optimization. In: Proceedings of the 21st Annual Conference on Learning Theory, pp. 263–274 (2008)
2. Agarwal, A., Dekel, O., Xiao, L.: Optimal algorithms for online convex optimization with multi-point bandit feedback. In: Proceedings of the 23rd Annual Conference on Learning Theory, pp. 28–40 (2010)
3. Agarwal, A., Hazan, E., Kale, S., Schapire, R.E.: Algorithms for portfolio management based on the newton method. In: Proceedings of the 23rd international conference on Machine learning, pp. 9–16 (2006)
4. Awerbuch, B., Kleinberg, R.D.: Adaptive routing with end-to-end feedback: distributed learning and geometric approaches. In: Proceedings of the 36th ACM Symposium on Theory of Computing, pp. 45–53 (2004)
5. Bianchi, N.C., Mansour, Y., Stoltz, G.: Improved Second-Order bounds for prediction with expert advice. In: Proceedings of the 18th Annual Conference on Learning Theory, vol. 3559, pp. 217–232 (2005)
6. Cesa-Bianchi, N., Lugosi, G.: Prediction, Learning, and Games. Cambridge University Press, New York, NY, USA (2006)
7. Chiang, C.K., Lee, C.J., Lu, C.J.: Beating bandits in gradually evolving worlds. In: COLT, pp. 210–227 (2013)
8. Chiang, C.K., Yang, T., Lee, C.J., Mahdavi, M., Lu, C.J., Jin, R., Zhu, S.: Online optimization with gradual variations. In: COLT, pp. 6.1–6.20 (2012)
9. Dani, V., Hayes, T.P.: Robbing the bandit: less regret in online geometric optimization against an adaptive adversary. In: Proceedings of the 17th Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 937–943 (2006)
10. Dani, V., Hayes, T.P., Kakade, S.: The price of bandit information for online optimization. In: Proceedings of the Twenty-First Annual Conference on Neural Information Processing Systems (2007)
11. Flaxman, A.D., Kalai, A.T., McMahan, H.B.: Online convex optimization in the bandit setting: gradient descent without a gradient. In: Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms, pp. 385–394 (2005)
12. Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. In: Proceedings of the 2nd European Conference on Computational Learning Theory, pp. 23–37. Springer-Verlag, London, UK (1995)
13. Hazan, E., Agarwal, A., Kale, S.: Logarithmic regret algorithms for online convex optimization. Machine Learning **69**, 169–192 (2007)

14. Hazan, E., Agarwal, A., Kale, S.: Logarithmic regret algorithms for online convex optimization. *Mach. Learn.* **69**, 169–192 (2007)
15. Hazan, E., Kale, S.: Better algorithms for benign bandits. In: *Proceedings of the 20th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 38–47 (2009)
16. Hazan, E., Kale, S.: Extracting certainty from uncertainty: regret bounded by variation in costs. *Machine Learning* **80**(2-3), 165–188 (2010)
17. Kalai, A., Vempala, S.: Efficient algorithms for online decision problems. *Journal of Computer and System Sciences* **71**, 291–307 (2005)
18. Kale, S.: Commentary on "online optimization with gradual variations". *Journal of Machine Learning Research - Proceedings Track* **23**, 6.21–6.24 (2012)
19. Kivinen, J., Smola, A.J., Williamson, R.C.: *Online Learning with Kernels*. *IEEE Transactions on Signal Processing* **52**, 2165–2176 (2004)
20. Kivinen, J., Warmuth, M.K.: Additive versus exponentiated gradient updates for linear prediction. In: *Proceedings of the 27th annual ACM symposium on Theory of computing, Proceedings of the Twenty-Seventh Annual ACM Symposium on Theory of Computing*, pp. 209–218. ACM, New York, NY, USA (1995)
21. Nemirovski, A.: 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**, 229–251 (2005)
22. Nesterov, Y.: *Introductory Lectures on Convex Optimization: A Basic Course (Applied Optimization)*, 1 edn. Springer Netherlands (2004)
23. Nesterov, Y.: Excessive gap technique in nonsmooth convex minimization. *SIAM Journal on Optimization* **16**, 235–249 (2005)
24. Nesterov, Y.: Smooth minimization of non-smooth functions. *Math. Program.* **103**, 127–152 (2005)
25. Takimoto, E., Warmuth, M.K.: Path kernels and multiplicative updates. *Journal Machine Learning Research* **4**, 773–818 (2003)
26. Zinkevich, M.: Online convex programming and generalized infinitesimal gradient ascent. In: *Proceedings of the 20th International Conference on Machine Learning*, pp. 928–936 (2003)

## Appendix A. Regret Bound which holds uniformly over time

As mentioned in Remark 2, the algorithms presented rely on the previous knowledge of the gradual variation  $EGV_{T,2}$  to tune the learning rate  $\eta$  to obtain the optimal bound. In this appendix, we show that the Algorithm 3 can be used as a black-box to achieve the same regret bound but without any prior knowledge of the  $EGV_{T,2}$ . We note that the analysis here is not specific to Algorithm 3 and it is general enough to be adapted to other algorithms in the paper too.

The main idea is to run the algorithm in epochs with a fixed learning rate  $\eta_k = \eta_0/2^k$  for  $k$ th epoch where  $\eta_0$  is a fixed constant and will be decided by analysis. We denote the number of epochs by  $K$  and let  $b_k$  denote the start of  $k$ th epoch. We note that  $b_{K+1} = T + 1$ . Within  $k$ th epoch, the algorithm ensures that the inequality  $\eta_k \sum_{t=b_k}^{b_{k+1}-1} \|\nabla c_{t+1}(\mathbf{z}_t) - \nabla c_t(\mathbf{z}_t)\|_2^2 \leq L^2 \eta_k^{-1}$  holds. To this end, the algorithm computes and maintains the quantity  $\sum_{s=b_k}^t \|\nabla c_{s+1}(\mathbf{z}_s) - \nabla c_s(\mathbf{z}_s)\|_2^2$  and sets the beginning of new epoch to be  $b_{k+1} = \min_t \sum_{s=b_k}^t \|\nabla c_{s+1}(\mathbf{z}_s) - \nabla c_s(\mathbf{z}_s)\|_2^2 > L^2 \eta_k^{-1}$ , i.e., the first iteration for which the invariant is violated. We note that this decision can only be made after seeing the  $t$ th cost function. Therefore, we burn the first iteration

of each epoch which causes an extra regret of  $KL^2$  in the total regret. From the analysis we have:

$$\begin{aligned} \sum_{t=1}^T c_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T c_t(\mathbf{x}) &\leq \sum_{k=1}^K \left[ \sum_{t=b_k}^{b_{k+1}-1} c_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=b_k}^{b_{k+1}-1} c_t(\mathbf{x}) \right] \quad (35) \\ &\leq \sum_{k=1}^K \frac{L}{2\eta_k} + \frac{\eta_k}{2L} \text{EGV}_{b_k:b_{k+1}-1} + KL^2 \\ &\leq \sum_{k=1}^K \frac{L}{2\eta_k} + \frac{L}{2\eta_k} + K = 2L \sum_{k=1}^K \eta_k^{-1} + KL^2 \end{aligned}$$

where the first inequality follows the analysis of algorithm for each epoch, the last inequality follows the invariant maintained within each phase and the constant  $KL^2$  is due to burning the first iteration of each epoch. We now try to upper bound the last term. We first note that  $\sum_{k=1}^K \eta_k^{-1} = \sum_{k=1}^{K-1} \eta_0^{-1} 2^k + \eta_0^{-1} 2^K = \eta_0^{-1} (2^K - 1) + \eta_0^{-1} 2^K \leq \eta_0^{-1} 2^{K+1}$ . Furthermore, from  $b_K$ , we know that  $\eta_{K-1} \sum_{t=b_{K-1}}^{b_K} \|\nabla c_{t+1}(\mathbf{z}_t) - \nabla c_t(\mathbf{z}_t)\|_2^2 \geq L^2 \eta_{K-1}^{-1}$  since the  $b_K$  is the first iteration within epoch  $K-1$  which violates the invariant. Also, from the monotonicity of gradual variation one can obtain that  $\eta_{K-1} \text{EGV}_{T,2} \geq \eta_{K-1} \sum_{t=b_{K-1}}^{b_K} \|\nabla c_{t+1}(\mathbf{z}_t) - \nabla c_t(\mathbf{z}_t)\|_2^2 \geq L^2 \eta_{K-1}^{-1}$  which indicates  $\eta_{K-1}^{-1} \leq \sqrt{\text{EGV}_{T,2}/L}$ . Putting these together, from (35) we obtain:

$$\sum_{t=1}^T c_t(\mathbf{x}_t) - \min_{\mathbf{x} \in \mathcal{P}} \sum_{t=1}^T c_t(\mathbf{x}) \leq 2L \frac{2^{K+1}}{\eta_0} + KL^2 \leq 8\sqrt{\text{EGV}_{T,2}} + KL^2.$$

It remains to bound the number of epochs in terms of  $\text{EGV}_{T,2}$ . A simple idea would be to set  $K$  to be  $\lceil \log \text{EGV}_{T,2} \rceil + 1$ , since it is the maximum number of epochs that could exist. Alternatively, we can also bound  $K$  in terms of  $\sqrt{\text{EGV}_{T,2}}$  which worsen the constant factor in the bound but results in a bound similar to one obtained by setting optimal  $\eta$ .

## Appendix B. Proof of Lemma 2

By using the definition of Bregman distance  $D(\mathbf{u}, \mathbf{z})$ , we can write equations (13) and (14) as

$$\begin{aligned} \mathbf{x} &= \arg \min_{\mathbf{u} \in U} \mathbf{u}^\top (\gamma\xi - \omega'(\mathbf{z})) + \omega(\mathbf{u}), \\ \mathbf{z}_+ &= \arg \min_{\mathbf{u} \in U} \mathbf{u}^\top (\gamma\zeta - \omega'(\mathbf{z})) + \omega(\mathbf{u}), \end{aligned}$$

by the first order optimality condition, we have

$$(\mathbf{u} - \mathbf{x})^\top (\gamma\xi - \omega'(\mathbf{z}) + \omega'(\mathbf{x})) \geq 0, \forall \mathbf{u} \in U, \quad (36)$$

$$(\mathbf{u} - \mathbf{z}_+)^\top (\gamma\zeta - \omega'(\mathbf{z}) + \omega'(\mathbf{z}_+)) \geq 0, \forall \mathbf{u} \in U. \quad (37)$$

Applying (36) with  $\mathbf{u} = \mathbf{z}_+$  and (37) with  $\mathbf{u} = \mathbf{x}$ , we get

$$\begin{aligned}\gamma(\mathbf{x} - \mathbf{z}_+)^\top \xi &\leq (\omega'(\mathbf{z}) - \omega'(\mathbf{x}))^\top (\mathbf{x} - \mathbf{z}_+), \\ \gamma(\mathbf{z}_+ - \mathbf{x})^\top \zeta &\leq (\omega'(\mathbf{z}) - \omega'(\mathbf{z}_+))^\top (\mathbf{z}_+ - \mathbf{x}).\end{aligned}$$

Summing up the two inequalities, we have

$$\gamma(\mathbf{x} - \mathbf{z}_+)^\top (\xi - \zeta) \leq (\omega'(\mathbf{z}_+) - \omega'(\mathbf{x}))^\top (\mathbf{x} - \mathbf{z}_+).$$

Then

$$\begin{aligned}\gamma \|\xi - \zeta\|_* \|\mathbf{x} - \mathbf{z}_+\| &\geq -\gamma(\mathbf{x} - \mathbf{z}_+)^\top (\xi - \zeta) \geq (\omega'(\mathbf{z}_+) - \omega'(\mathbf{x}))^\top (\mathbf{z}_+ - \mathbf{x}) \\ &\geq \alpha \|\mathbf{z}_+ - \mathbf{x}\|^2.\end{aligned}\quad (38)$$

where in the last inequality, we use the strong convexity of  $\omega(\mathbf{x})$ .

$$\begin{aligned}D(\mathbf{u}, \mathbf{z}) - D(\mathbf{u}, \mathbf{z}_+) &= \omega(\mathbf{z}_+) - \omega(\mathbf{z}) + (\mathbf{u} - \mathbf{z}_+)^\top \omega'(\mathbf{z}_+) - (\mathbf{u} - \mathbf{z})^\top \omega'(\mathbf{z}) \\ &= \omega(\mathbf{z}_+) - \omega(\mathbf{z}) + (\mathbf{u} - \mathbf{z}_+)^\top \omega'(\mathbf{z}_+) - (\mathbf{u} - \mathbf{z}_+)^\top \omega'(\mathbf{z}) - (\mathbf{z}_+ - \mathbf{z})^\top \omega'(\mathbf{z}) \\ &= \omega(\mathbf{z}_+) - \omega(\mathbf{z}) - (\mathbf{z}_+ - \mathbf{z})^\top \omega'(\mathbf{z}) + (\mathbf{u} - \mathbf{z}_+)^\top (\omega'(\mathbf{z}_+) - \omega'(\mathbf{z})) \\ &= \omega(\mathbf{z}_+) - \omega(\mathbf{z}) - (\mathbf{z}_+ - \mathbf{z})^\top \omega'(\mathbf{z}) + (\mathbf{u} - \mathbf{z}_+)^\top (\gamma\zeta + \omega'(\mathbf{z}_+) - \omega'(\mathbf{z})) - (\mathbf{u} - \mathbf{z}_+)^\top \gamma\zeta \\ &\geq \omega(\mathbf{z}_+) - \omega(\mathbf{z}) - (\mathbf{z}_+ - \mathbf{z})^\top \omega'(\mathbf{z}) - (\mathbf{u} - \mathbf{z}_+)^\top \gamma\zeta \\ &= \underbrace{\omega(\mathbf{z}_+) - \omega(\mathbf{z}) - (\mathbf{z}_+ - \mathbf{z})^\top \omega'(\mathbf{z}) - (\mathbf{x} - \mathbf{z}_+)^\top \gamma\zeta}_{\epsilon} + (\mathbf{x} - \mathbf{u})^\top \gamma\zeta,\end{aligned}$$

where the inequality follows from (37). We proceed by bounding  $\epsilon$  as:

$$\begin{aligned}\epsilon &= \omega(\mathbf{z}_+) - \omega(\mathbf{z}) - (\mathbf{z}_+ - \mathbf{z})^\top \omega'(\mathbf{z}) - (\mathbf{x} - \mathbf{z}_+)^\top \gamma\zeta \\ &= \omega(\mathbf{z}_+) - \omega(\mathbf{z}) - (\mathbf{z}_+ - \mathbf{z})^\top \omega'(\mathbf{z}) - (\mathbf{x} - \mathbf{z}_+)^\top \gamma(\zeta - \xi) - (\mathbf{x} - \mathbf{z}_+)^\top \gamma\xi \\ &= \omega(\mathbf{z}_+) - \omega(\mathbf{z}) - (\mathbf{z}_+ - \mathbf{z})^\top \omega'(\mathbf{z}) - (\mathbf{x} - \mathbf{z}_+)^\top \gamma(\zeta - \xi) \\ &\quad + (\mathbf{z}_+ - \mathbf{x})^\top (\gamma\xi - \omega'(\mathbf{z}) + \omega'(\mathbf{x})) - (\mathbf{z}_+ - \mathbf{x})^\top (\omega'(\mathbf{x}) - \omega'(\mathbf{z})) \\ &\geq \omega(\mathbf{z}_+) - \omega(\mathbf{z}) - (\mathbf{z}_+ - \mathbf{z})^\top \omega'(\mathbf{z}) - (\mathbf{x} - \mathbf{z}_+)^\top \gamma(\zeta - \xi) - (\mathbf{z}_+ - \mathbf{x})^\top (\omega'(\mathbf{x}) - \omega'(\mathbf{z})) \\ &= \omega(\mathbf{z}_+) - \omega(\mathbf{z}) - (\mathbf{x} - \mathbf{z})^\top \omega'(\mathbf{z}) - (\mathbf{x} - \mathbf{z}_+)^\top \gamma(\zeta - \xi) - (\mathbf{z}_+ - \mathbf{x})^\top \omega'(\mathbf{x}) \\ &= [\omega(\mathbf{z}_+) - \omega(\mathbf{x}) - (\mathbf{z}_+ - \mathbf{x})^\top \omega'(\mathbf{x})] + [\omega(\mathbf{x}) - \omega(\mathbf{z}) - (\mathbf{x} - \mathbf{z})^\top \omega'(\mathbf{z})] - (\mathbf{x} - \mathbf{z}_+)^\top \gamma(\zeta - \xi) \\ &\geq \frac{\alpha}{2} \|\mathbf{x} - \mathbf{z}_+\|^2 + \frac{\alpha}{2} \|\mathbf{x} - \mathbf{z}\|^2 - \gamma \|\mathbf{x} - \mathbf{z}_+\| \|\zeta - \xi\|_* \\ &\geq \frac{\alpha}{2} \{\|\mathbf{x} - \mathbf{z}_+\|^2 + \|\mathbf{x} - \mathbf{z}\|^2\} - \frac{\gamma^2}{\alpha} \|\zeta - \xi\|_*^2,\end{aligned}$$

where the first inequality follows from (36), the second inequality follows from the strong convexity of  $\omega(\mathbf{x})$ , and the last inequality follows from (38). Combining the above results, we have

$$\gamma(\mathbf{x} - \mathbf{u})^\top \zeta \leq D(\mathbf{u}, \mathbf{z}) - D(\mathbf{u}, \mathbf{z}_+) + \frac{\gamma^2}{\alpha} \|\zeta - \xi\|_*^2 - \frac{\alpha}{2} \{\|\mathbf{x} - \mathbf{z}_+\|^2 + \|\mathbf{x} - \mathbf{z}\|^2\}.$$



### Appendix C. Proof of Corollary 3

We first have the key inequality in (22): for any  $\mathbf{z} \in \mathcal{P}$

$$\begin{aligned} c_t(\mathbf{x}_t) - c_t(\mathbf{z}) &\leq D_t(\mathbf{z}, \mathbf{z}_{t-1}) - D_t(\mathbf{z}, \mathbf{z}_t) - \beta \|\mathbf{x}_t - \mathbf{z}\|_{h_t}^2 \\ &\quad + \|\nabla c_t(\mathbf{x}_t) - \nabla c_{t-1}(\mathbf{x}_{t-1})\|_{H_t^{-1}}^2 - \frac{1}{2} [\|\mathbf{x}_t - \mathbf{z}_{t-1}\|_{H_t}^2 + \|\mathbf{x}_t - \mathbf{z}_t\|_{H_t}^2]. \end{aligned}$$

Taking summation over  $t = 1, \dots, T$ , we have

$$\begin{aligned} \sum_{t=1}^T c_t(\mathbf{x}_t) - \sum_{t=1}^T c_t(\mathbf{z}) &\leq \underbrace{\sum_{t=1}^T (D_t(\mathbf{z}, \mathbf{z}_{t-1}) - D_t(\mathbf{z}, \mathbf{z}_t))}_{A_t} - \underbrace{\sum_{t=1}^T \beta \|\mathbf{x}_t - \mathbf{z}\|_{h_t}^2}_{C_t} \\ &\quad + \underbrace{\sum_{t=1}^T \|\nabla c_t(\mathbf{x}_t) - \nabla c_{t-1}(\mathbf{x}_{t-1})\|_{H_t^{-1}}^2}_{S_t} - \underbrace{\sum_{t=1}^T \frac{1}{2} [\|\mathbf{x}_t - \mathbf{z}_{t-1}\|_{H_t}^2 + \|\mathbf{x}_t - \mathbf{z}_t\|_{H_t}^2]}_{B_t} \end{aligned}$$

Next we bound each term individually. First,

$$\sum_{t=1}^T A_t = D_1(\mathbf{z}, \mathbf{z}_0) - D_T(\mathbf{z}, \mathbf{z}_T) + \sum_{t=1}^{T-1} (D_{t+1}(\mathbf{z}, \mathbf{z}_t) - D_t(\mathbf{z}, \mathbf{z}_t))$$

Note that  $D_1(\mathbf{z}, \mathbf{z}_0) = \frac{1}{2}(1 + \beta G^2) \|\mathbf{z}\|_2^2 \leq \frac{1}{2}(1 + \beta G^2)$  for any  $\mathbf{z} \in \mathcal{P}$ , and  $D_{t+1}(\mathbf{z}, \mathbf{z}_t) - D_t(\mathbf{z}, \mathbf{z}_t) = \frac{\beta}{2} \|\mathbf{z} - \mathbf{z}_t\|_{h_t}^2$ , therefore

$$\sum_{t=1}^T A_t \leq \frac{1}{2}(1 + \beta G^2) + \sum_{t=1}^T \frac{\beta}{2} \|\mathbf{z} - \mathbf{z}_t\|_{h_t}^2$$

Then

$$\begin{aligned} \sum_{t=1}^T (A_t - C_t) &\leq \frac{1}{2}(1 + \beta G^2) + \sum_{t=1}^T \left[ \frac{\beta}{2} \|\mathbf{z} - \mathbf{z}_t\|_{h_t}^2 - \beta \|\mathbf{x}_t - \mathbf{z}\|_{h_t}^2 \right] \\ &\leq \frac{1}{2}(1 + \beta G^2) + \sum_{t=1}^T [\beta \|\mathbf{z} - \mathbf{x}_t\|_{h_t}^2 + \beta \|\mathbf{x}_t - \mathbf{z}_t\|_{h_t}^2 - \beta \|\mathbf{x}_t - \mathbf{z}\|_{h_t}^2] \\ &\leq \frac{1}{2}(1 + \beta G^2) + \sum_{t=1}^T \beta \|\mathbf{x}_t - \mathbf{z}_t\|_{h_t}^2 \leq \frac{1}{2}(1 + \beta G^2) + \sum_{t=1}^T \|\mathbf{x}_t - \mathbf{z}_t\|_{H_t}^2 \end{aligned}$$

Noting the updates in (21) and from inequality in (38), we can get

$$\|\mathbf{x}_t - \mathbf{z}_t\|_{H_t} \leq \|\nabla c_t(\mathbf{x}_t) - \nabla c_{t-1}(\mathbf{x}_{t-1})\|_{H_t^{-1}}$$

Next, we bound  $\sum_{t=1}^T B_t$ .

$$\begin{aligned}
\sum_{t=1}^T B_t &= \frac{1}{2} \sum_{t=0}^{T-1} \|\mathbf{x}_{t+1} - \mathbf{z}_t\|_{H_{t+1}}^2 + \frac{1}{2} \sum_{t=1}^T \|\mathbf{x}_t - \mathbf{z}_t\|_{H_t}^2 \\
&\geq \frac{1}{2} \sum_{t=1}^{T-1} \|\mathbf{x}_{t+1} - \mathbf{z}_t\|_{H_t}^2 + \frac{1}{2} \sum_{t=1}^{T-1} \|\mathbf{x}_t - \mathbf{z}_t\|_{H_t}^2 \\
&\geq \frac{1}{4} \sum_{t=1}^{T-1} \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_{H_t}^2 \geq \frac{1}{4} \sum_{t=0}^{T-1} \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2 - \frac{1}{4} \|\mathbf{x}_1 - \mathbf{x}_0\|_2^2 \\
&\geq \frac{1}{4} \sum_{t=0}^{T-1} \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2,
\end{aligned}$$

where the last inequality follows that  $\mathbf{x}_0 = \mathbf{x}_1 = 0$ . Therefore,

$$\begin{aligned}
\sum_{t=1}^T c_t(\mathbf{x}_t) - \sum_{t=1}^T c_t(\mathbf{z}) &\leq \frac{1}{2} (1 + \beta G^2) + 2 \sum_{t=1}^T \|\nabla c_t(\mathbf{x}_t) - \nabla c_{t-1}(\mathbf{x}_{t-1})\|_{H_t^{-1}} \\
&\quad - \frac{1}{4} \sum_{t=0}^{T-1} \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2
\end{aligned}$$

To proceed, we need the following lemma.

**Lemma 6 (Lemma 19 [8])**

$$\sum_{t=1}^T \|\nabla c_t(\mathbf{x}_t) - \nabla c_{t-1}(\mathbf{x}_{t-1})\|_{H_t^{-1}} \leq \frac{4d}{\beta} \ln \left( 1 + \frac{\beta}{4} \sum_{t=1}^T \|\nabla c_t(\mathbf{x}_t) - \nabla c_{t-1}(\mathbf{x}_{t-1})\|_2^2 \right)$$

Thus,

$$\begin{aligned}
\sum_{t=1}^T \|\nabla c_t(\mathbf{x}_t) - \nabla c_{t-1}(\mathbf{x}_{t-1})\|_{H_t^{-1}} &\leq \frac{4d}{\beta} \ln \left( 1 + \frac{\beta}{4} \sum_{t=1}^T \|\nabla c_t(\mathbf{x}_t) - \nabla c_{t-1}(\mathbf{x}_{t-1})\|_2^2 \right) \\
&\leq \frac{4d}{\beta} \ln \left( 1 + \frac{\beta}{4} \sum_{t=1}^T \|\nabla c_t(\mathbf{x}_t) - \nabla c_t(\mathbf{x}_{t-1}) + \nabla c_t(\mathbf{x}_{t-1}) - \nabla c_{t-1}(\mathbf{x}_{t-1})\|_2^2 \right) \\
&\leq \frac{4d}{\beta} \ln \left( 1 + \frac{\beta}{2} \sum_{t=1}^T L^2 \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + \frac{\beta}{2} \sum_{t=1}^T \|\nabla c_t(\mathbf{x}_{t-1}) - \nabla c_{t-1}(\mathbf{x}_{t-1})\|_2^2 \right) \\
&\leq \frac{4d}{\beta} \ln \left( 1 + \frac{\beta}{2} \sum_{t=0}^{T-1} L^2 \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2 + \frac{\beta}{2} \sum_{t=0}^{T-1} \|\nabla c_{t+1}(\mathbf{x}_t) - \nabla c_t(\mathbf{x}_t)\|_2^2 \right) \\
&\leq \frac{4d}{\beta} \ln \left( 1 + \frac{\beta}{2} \sum_{t=0}^{T-1} L^2 \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2 + \frac{\beta}{2} \text{EGV}_{T,2} \right)
\end{aligned}$$

Then,

$$\begin{aligned} \sum_{t=1}^T c_t(\mathbf{x}_t) - \sum_{t=1}^T c_t(\mathbf{z}) &\leq \frac{1}{2}(1 + \beta G^2) + \frac{8d}{\beta} \ln \left( 1 + \frac{\beta}{2} \sum_{t=0}^{T-1} L^2 \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2 + \frac{\beta}{2} \text{EGV}_{T,2} \right) \\ &\quad - \frac{1}{4} \sum_{t=0}^{T-1} \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2 \end{aligned}$$

Without loss of generality we assume  $8dL^2 \geq 1$ . Next, let us consider two cases. In the first case, we assume  $\beta \text{EGV}_{T,2} \leq 16dL^2$ . Then

$$\begin{aligned} &\frac{8d}{\beta} \ln \left( 1 + \frac{\beta}{2} \sum_{t=0}^{T-1} L^2 \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2 + \frac{\beta}{2} \text{EGV}_{T,2} \right) \\ &\leq \frac{8d}{\beta} \ln \left( 1 + \frac{\beta}{2} \sum_{t=0}^{T-1} L^2 \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2 + 8dL^2 \right) \\ &\leq \frac{8d}{\beta} \ln \left( \frac{\beta}{2} \sum_{t=0}^{T-1} L^2 \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2 + 16dL^2 \right) \\ &= \frac{8d}{\beta} \left[ \ln 16dL^2 + \ln \left( \frac{\frac{\beta}{2} \sum_{t=0}^{T-1} L^2 \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2}{16dL^2} + 1 \right) \right] \\ &\leq \frac{8d}{\beta} \ln 16dL^2 + \frac{8d \frac{\beta}{2} \sum_{t=0}^{T-1} L^2 \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2}{\beta 16dL^2} = \frac{8d}{\beta} \ln 16dL^2 + \frac{\sum_{t=0}^{T-1} \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2}{4} \end{aligned}$$

where the last inequality follows  $\ln(1+x) \leq x$  for  $x \geq 0$ . Then we get

$$\sum_{t=1}^T c_t(\mathbf{x}_t) - \sum_{t=1}^T c_t(\mathbf{z}) \leq \frac{1}{2}(1 + \beta G^2) + \frac{8d}{\beta} \ln 16dL^2$$

In the second case, we assume  $\beta \text{EGV}_{T,2} \geq 16dL^2 \geq 2$ , then we have

$$\begin{aligned} &\frac{8d}{\beta} \ln \left( 1 + \frac{\beta}{2} \sum_{t=0}^{T-1} L^2 \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2 + \frac{\beta}{2} \text{EGV}_{T,2} \right) \\ &\leq \frac{8d}{\beta} \ln \left( \frac{\beta}{2} \sum_{t=0}^{T-1} L^2 \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2 + \beta \text{EGV}_{T,2} \right) \\ &\leq \frac{8d}{\beta} \left[ \ln(\beta \text{EGV}_{T,2}) + \ln \left( \frac{\frac{\beta}{2} \sum_{t=0}^{T-1} L^2 \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2}{\beta \text{EGV}_{T,2}} + 1 \right) \right] \\ &= \frac{8d}{\beta} \ln(\beta \text{EGV}_{T,2}) + \frac{8d \frac{\beta}{2} \sum_{t=0}^{T-1} L^2 \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2}{\beta \text{EGV}_{T,2}} \\ &= \frac{8d}{\beta} \ln(\beta \text{EGV}_{T,2}) + \frac{4dL^2 \sum_{t=0}^{T-1} \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2}{\beta \text{EGV}_{T,2}} \\ &\leq \frac{8d}{\beta} \ln(\beta \text{EGV}_{T,2}) + \frac{\sum_{t=0}^{T-1} \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2}{4} \end{aligned}$$

where the last inequality follows  $\beta\text{EGV}_{T,2} \geq 16dL^2$ . Then we get

$$\sum_{t=1}^T c_t(\mathbf{x}_t) - \sum_{t=1}^T c_t(\mathbf{z}) \leq \frac{1}{2}(1 + \beta G^2) + \frac{8d}{\beta} \ln(\beta\text{EGV}_{T,2})$$

Thus, we complete the proof by combining the two cases.

Next, we prove Lemma 6. We need the following lemma, which can be proved by using Lemma 6 [14] and noting that  $|I + \sum_{\tau=1}^t \mathbf{u}_\tau \mathbf{u}_\tau^\top| \leq (1 + \sum_{t=1}^T \|\mathbf{u}_t\|_2^2)^d$ , where  $|\cdot|$  denotes the determinant of a matrix.

**Lemma 7** *Let  $\mathbf{u}_t \in \mathbb{R}^d, t = 1, \dots, T$  be a sequence of vectors. Let  $V_t = I + \sum_{\tau=1}^t \mathbf{u}_\tau \mathbf{u}_\tau^\top$ . Then,*

$$\sum_{t=1}^T \mathbf{u}_t^\top V_t^{-1} \mathbf{u}_t \leq d \ln \left( 1 + \sum_{t=1}^T \|\mathbf{u}_t\|_2^2 \right)$$

To prove Lemma 6, we let  $\mathbf{v}_t = \nabla c_t(\mathbf{x}_t), t = 1, \dots, T$  and  $\mathbf{v}_0 = 0$ . Then  $H_t = I + \beta G^2 I + \beta \sum_{\tau=0}^{t-1} \mathbf{v}_\tau \mathbf{v}_\tau^\top$ . Note that we assume  $\|\nabla c_t(\mathbf{x})\|_2 \leq G$ , therefore

$$\begin{aligned} H_t &\geq I + \beta \sum_{\tau=1}^t \mathbf{v}_\tau \mathbf{v}_\tau^\top \geq I + \frac{\beta}{2} \sum_{\tau=1}^t (\mathbf{v}_\tau \mathbf{v}_\tau^\top + \mathbf{v}_{\tau-1} \mathbf{v}_{\tau-1}^\top) \\ &\geq I + \frac{\beta}{4} \sum_{\tau=1}^t (\mathbf{v}_\tau - \mathbf{v}_{\tau-1})(\mathbf{v}_\tau - \mathbf{v}_{\tau-1})^\top = V_t \end{aligned}$$

Let  $\mathbf{u}_t = (\sqrt{\beta}/2)(\mathbf{v}_t - \mathbf{v}_{t-1})$ , then  $V_t = I + \sum_{\tau=1}^t \mathbf{u}_\tau \mathbf{u}_\tau^\top$ . By applying the above lemma, we have

$$\frac{\beta}{4} \sum_{t=1}^T (\mathbf{v}_t - \mathbf{v}_{t-1})^\top V_t^{-1} (\mathbf{v}_t - \mathbf{v}_{t-1}) \leq d \ln \left( 1 + \frac{\beta}{4} \sum_{t=1}^T \|\mathbf{v}_t - \mathbf{v}_{t-1}\|_2^2 \right)$$

Thus,

$$\begin{aligned} \sum_{t=1}^T (\mathbf{v}_t - \mathbf{v}_{t-1})^\top H_t^{-1} (\mathbf{v}_t - \mathbf{v}_{t-1}) &\leq \sum_{t=1}^T (\mathbf{v}_t - \mathbf{v}_{t-1})^\top V_t^{-1} (\mathbf{v}_t - \mathbf{v}_{t-1}) \\ &\leq \frac{4d}{\beta} \ln \left( 1 + \frac{\beta}{4} \sum_{t=1}^T \|\mathbf{v}_t - \mathbf{v}_{t-1}\|_2^2 \right) \end{aligned}$$