# Proximal Online Gradient Is Optimum for Dynamic Regret: A General Lower Bound

Yawei Zhao<sup>10</sup>, Shuang Qiu<sup>10</sup>, Kuan Li, Lailong Luo<sup>10</sup>, Jianping Yin<sup>10</sup>, and Ji Liu, Member, IEEE

Abstract—In online learning, the dynamic regret metric chooses the reference oracle that may change over time, while the typical (static) regret metric assumes the reference solution to be constant over the whole time horizon. The dynamic regret metric is particularly interesting for applications, such as online recommendation (since the customers' preference always evolves over time). While the online gradient (OG) method has been shown to be optimal for the static regret metric, the optimal algorithm for the dynamic regret remains unknown. In this article, we show that proximal OG (a general version of OG) is optimum to the dynamic regret by showing that the proved lower bound matches the upper bound. It is highlighted that we provide a new and general lower bound of dynamic regret. It provides new understanding about the difficulty to follow the dynamics in the online setting.

*Index Terms*—Dynamic regret, lower bound, online convex optimization, proximal online gradient (POG).

## I. INTRODUCTION

**O**NLINE learning [1]–[8] is a hot research topic for the last decade of years, due to its application in practices, such as online recommendation [9], online collaborative filtering [10], [11], moving object detection [12], and many others, as well as its close connection with other research areas, such as stochastic optimization [13], [14], image retrieval [15], multiple kernel learning [16], [17], and bandit problems [18]–[21], etc.

Manuscript received May 7, 2019; revised May 26, 2020, December 12, 2020, and May 5, 2021; accepted May 24, 2021. This work was supported by the National Key R&D Program of China under Grant No. 2018YFB1003203. (Yawei Zhao and Shuang Qiu contributed equally to this work.) (Corresponding author: Jianping Yin.)

Yawei Zhao and Jianping Yin are with the School of Cyberspace Security, Dongguan University of Technology, Dongguan 523808, China (e-mail: csyawei.zhao@gmail.com; jpyin@dgut.edu.cn).

Shuang Qiu is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: qiush@umich.edu).

Kuan Li is with the School of Cyberspace Security, Dongguan University of Technology, Dongguan 523808, China, and also with the Guangdong Key Laboratory of Intelligent Information Processing, Shenzhen 518052, China (e-mail: likuan@dgut.edu.cn).

Lailong Luo is with the Science and Technology on Information Systems Engineering Laboratory, National University of Defense Technology, Changsha 410073, China (e-mail: luolailong09@nudt.edu.cn).

Ji Liu is with the Department of Computer Science, University of Rochester, Rochester, NY 14620 USA (e-mail: ji.liu.uwisc@gmail.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TNNLS.2021.3087579.

Digital Object Identifier 10.1109/TNNLS.2021.3087579

The typical objective function in online learning is to minimize the (static) regret defined as follows:

$$\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \min_{\substack{\mathbf{x} \in \mathcal{X} \\ \text{the optimal reference}}} \sum_{t=1}^{T} f_t(\mathbf{x})$$
(1)

where  $\mathbf{x}_t$  is the decision made at step *t* after receiving the information before that (e.g.,  $\{\nabla f_s(\mathbf{x}_s), f_s(\mathbf{x}_s)\}_{s=1}^{t-1}$ ). The optimal reference is chosen at the point that minimizes the sum of all component functions up to time *T*. However, the way to decide the optimal reference may not fit some important applications in practice. For example, in the recommendation task,  $f_t(\mathbf{x})$  is the regret at time *t* decided by the *t*th coming customer and our recommendation strategy  $\mathbf{x}$ . Based on the definition of regret in (1), it implicitly assumes that the optimal recommendation strategy is constant over time, which is not necessarily true for the recommendation task (as well as many other applications) since the costumers' preference usually evolves over time.

Zinkevich [1] proposed to use the dynamic regret as the metric for online learning, which allows the optimal strategy changing over time. More specifically, it is defined by

$$\mathcal{R}_T^A := \sum_{t=1}^T f_t(\mathbf{x}_t) - \min_{\{\mathbf{y}_t\}_{t=1}^T \in \mathcal{L}_{D_0}^T} \sum_{t=1}^T f_t(\mathbf{y}_t)$$
(2)

where A denotes the algorithm that decides  $\mathbf{x}_t$  iteratively,  $\{\mathbf{y}_t\}_{t=1}^T$  is short for a sequence  $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T\}$ , and the dynamics upper bound  $\mathcal{L}_{D_0}^T$  is defined by

$$\mathcal{L}_{D_0}^T := \left\{ \{ \mathbf{y}_t \}_{t=1}^T : \sum_{t=1}^{T-1} \| \mathbf{y}_{t+1} - \mathbf{y}_t \| \le D_0 \right\}.$$
 (3)

It was shown that the dynamic regret of online gradient (OG) is bounded [1], [22], [23] by

$$\mathcal{R}_T^{\rm OG} \lesssim T\eta + \frac{1}{\eta} + \frac{D_0}{\eta} \tag{4}$$

where  $\eta$  represents the learning rate and  $\leq$  means "less than equal up to a constant factor." This reminds people to ask a few fundamental questions.

1) As we know, OG is optimum for the static regret (i.e., the case of  $D_0 = 0$ ), and the dependence on *T* is tight. However, is the dependence on the dynamics  $D_0$ 

2162-237X © 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

tight? In other words, is OG also optimal for dynamic regret?

- 2) Is this bound tight enough? If no, how to design a "smarter" algorithm to follow the dynamics?
- 3) How difficult is it to follow dynamics in online learning?

Although the dynamic regret receives more and more attention recently [23]–[28] and some successive studies claim to improve this result by considering specific functions types (e.g., strongly convex  $f_t$ ) or considering different definitions of dynamic regret, these fundamental questions still remain unsolved.

In this article, we consider a more general setup for the problem

$$f_t(\mathbf{x}) = F_t(\mathbf{x}) + H(\mathbf{x})$$
(5)

with  $F_t(\mathbf{x})$  and  $H(\mathbf{x})$  being only convex and closed and a more general definition for dynamic constraint in (6)

$$\mathcal{L}_{D_{\beta}}^{T} := \left\{ \{ \mathbf{y}_{t} \}_{t=1}^{T} : \sum_{t=1}^{T-1} t^{\beta} \cdot \| \mathbf{y}_{t+1} - \mathbf{y}_{t} \| \le D_{\beta} \right\}$$
(6)

where  $\beta$  and  $D_{\beta}$  are the predefined parameters to restrict the change of reference models over time. We show that the upper bound of the proximal online gradient (POG) algorithm can achieve

$$\mathcal{R}_T^{\text{POG}} \lesssim \sqrt{T} + \sqrt{T^{1-\beta} \cdot D_\beta} \tag{7}$$

where  $\lesssim$  means "less than equal up to a constant factor."

When  $\beta = 0$  and  $H(\mathbf{x}) \equiv 0$ , (7) recovers the dependence of T in (4). However, (7) still holds for proximal mapping when updating  $\mathbf{x}_t$ . When  $\beta > 0$ , since  $D_{\beta} < D_0 T^{\beta}$ , (7) is slightly better than the proved special case in (4).<sup>1</sup>

To understand the difficulty of following dynamics in online learning, we derive a lower bound (that measures the dynamic regret by the optimal algorithm) and show that the proved upper bound for POG matches the lower bound up to a constant factor, which indicates that POG is an optimal algorithm even for dynamic regret (not just for static regret).

# II. RELATED WORK

In this section, we outline and review the existing work about online learning problem with the regret in static and dynamic environments briefly.

# A. Static Regret

OG in the static environment has been extensively investigated for the last decade of years [2], [3], [29]. Specifically, when  $f_t$  is strongly convex, the regret of OG is  $\mathcal{O}(\log T)$ . When  $f_t(\cdot)$  is only convex, the regret of OG is  $\mathcal{O}((T)^{1/2})$ .

# B. Dynamic Regret

Zinkevich [1] obtained the regret in the order of  $\mathcal{O}(T\eta + (1/\eta) + (D_0/\eta))$  for the convex function  $f_t$ . Similarly, assume that the dynamic constraint is defined by  $\sum_{t=1}^{T-1} \|\mathbf{y}_{t+1} - \Phi(\mathbf{y}_t)\| \le D_0$ , where  $\Phi(\cdot)$  provides the prediction about the dynamic environment. When  $\Phi(\mathbf{y}_t)$  predicts the dynamic environment accurately, Hall and Willett [22], [23] obtained a better regret than [1].

In addition, assume that  $f_t$  is  $\alpha$  strongly convex and  $\beta$  smooth, and the dynamic constraint is defined by  $D^* := \sum_{t=1}^{T-1} \|\mathbf{y}_{t+1}^* - \mathbf{y}_t^*\|$ , where  $\mathbf{y}_t^* := \operatorname{argmin}_{\mathbf{y} \in \mathcal{X}} f_t(\mathbf{y})$ . Mokhtari *et al.* [24] obtained  $\mathcal{O}(D^*)$  regret. When querying noisy gradient, Bedi *et al.* [30] obtained  $\mathcal{O}(D^* + \mathcal{E})$  regret, where  $\mathcal{E}$  is the cumulative gradient error. Yang *et al.* [25] and Gao *et al.* [31] extended it for nonstrongly convex and nonconvex functions, respectively. Shahrampour and Jadbabaie [27] extended it to the decentrialized setting.<sup>2</sup> Furthermore, define  $S^* := \sum_{t=1}^{T-1} \|\mathbf{y}_{t+1}^* - \mathbf{y}_t^*\|^2$ , where  $\mathbf{y}_t^* := \operatorname{argmin}_{\mathbf{y} \in \mathcal{X}} f_t(\mathbf{y})$ . When querying  $\mathcal{O}(\kappa)$  with  $\kappa := (\beta/\alpha)$  gradients for every iteration, Zhang *et al.* [26] improved the dynamic regret to be  $\mathcal{O}(\min\{D^*, S^*\})$ . Comparing with the previous work, our analysis does not assume the differentiability and strong convexity of  $f_t$ .

Other regularities, including the functional variation [5], [32]–[34], the gradient variation [35], and the mixed regularity [28], [36], [37], have been investigated to bound the dynamic regret. Those different regularities cannot be compared directly because they measure different aspects of the variation in the dynamic environment. In this article, we use (6) to bound the regret, and it is the future work to extend our analysis to other regularities.

György and Szepesvári [38] studied a dynamic regret<sup>3</sup> in a slightly more general setting than (3) by relaxing the distance metric  $||\mathbf{y}_{t+1} - \mathbf{y}_t||$  to a general  $\ell_p$ -norm  $||\mathbf{y}_{t+1} - \mathbf{y}_t||_p$  with  $p \in (1, 2]$ . They obtain an upper bound  $\mathcal{O}((D_0/\eta) + (1/\eta) + T\eta)$  for an algorithm namely twisted mirror descent (TMD). When the dynamics  $D_0$  is known and can be used to set the learning rate  $\eta \propto ((D_0/T))^{1/2}$ , the upper bound becomes  $\mathcal{O}((D_0T + T)^{1/2})$  [38]. This result is essentially consistent with our upper bound, but we consider a different algorithm and a different generalization of the dynamic regret definition and provide a lower bound more importantly.

Recently, Zhang *et al.* [39] provided a lower bound of dynamic regret in the case of  $\beta = 0$ , and Abernethy *et al.* [40] and Orabona [41] presented lower bounds of static regret, that is,  $\beta = 0$  and  $D_0 = 0$ . Comparing with the known results, our lower bound holds for  $0 \le \beta < 1$  and  $D_{\beta} > 0$ , which, as far as we know, is the first lower bound for the dynamic regret. Besides, the previous results only hold for the differentiable function  $f_t$ , but our lower bound of dynamic regret still holds for nondifferentiable function  $f_t$ , e.g.,  $\ell_1$ -norm. In addition, Zhang *et al.* [39] led to much higher computational complexity than the work. The reason is that Zhang *et al.* [39] maintained  $\mathcal{O}(\log T)$  experts and thus led to  $\mathcal{O}(\log T)$  updates of the model at an iteration. Comparing with it, our method updates the model only once for every iteration.

<sup>&</sup>lt;sup>1</sup>This bound can be achieved by setting  $\eta \propto (1/(T)^{1/2}) + ((D_{\beta}/T))^{1/2}$ , which implies that  $D_{\beta}$  has to be known to tune the learning rate.

<sup>&</sup>lt;sup>2</sup>The definition of  $D^*$  is changed slightly in the decentrialized setting. <sup>3</sup>It is called shifting regret in [38]. To avoid the confusion with many papers that will be discussed in the following, the shifting regret in this article is defined in a different way from [38].

#### C. Shifting Regret (or Tracking Regret)

The *M*-shifting regret of an algorithm  $A \in \mathcal{A}$  is defined by [6], [38], [42]–[49]

$$\widetilde{\mathcal{R}}_{T}^{A} := \sum_{t=1}^{T} f_{t}(\mathbf{x}_{t}) - \min_{\{\mathbf{y}_{t}\}_{t=1}^{T} \in \mathcal{L}_{M}^{T}} \sum_{t=1}^{T} f_{t}(\mathbf{y}_{t})$$
(8)

where  $\mathcal{L}_{M}^{T} = \{\{\mathbf{y}_{t}\}_{t=1}^{T} : \sum_{t=1}^{T-1} \mathbb{1}\{\mathbf{y}_{t+1} \neq \mathbf{y}_{t}\} \leq M\}$ . Here, the dynamics is modeled by the number of changes of the reference sequence  $\{\mathbf{y}_{t}\}_{t=1}^{T}$ . The shifting regret is closely related to the dynamic regret and can be considered as a variation of dynamic regret and is usually studied in the setting of learning with expert advice. The result in [44] and [50] implies an upper bound  $\mathcal{O}((MT \log^{2} T)^{1/2})$  for the shifting regret. The results in [6] and [51] imply an improved upper bound to  $\mathcal{O}((MT \log T)^{1/2})$ . Note that those bounds are achieved under the condition that M is unknown, that is, M cannot be used to tune the learning rate  $\eta$ . In other words, those previous results about the shifting regret do not require knowledge of the dynamics.

#### **III. NOTATIONS AND ASSUMPTIONS**

In this section, we introduce notations and important assumptions for the online learning algorithm used throughout this article.

#### A. Notations

Throughout this article, we use the following notations.

- 1)  $\mathcal{A}$  represents the family of all possible online algorithms.
- 2)  $\mathcal{F}$  represents the family of loss functions available to the adversary, where for any loss function  $f_t \in \mathcal{F}$ :  $\mathcal{X} \subset \mathbb{R}^d \mapsto \mathbb{R}, f_t(\mathbf{x}) = F_t(\mathbf{x}) + H(\mathbf{x})$  satisfies Assumptions 1 and 2.  $\mathcal{F}^T$  denotes the function product space by  $\mathcal{F} \times \mathcal{F} \times \cdots \times \mathcal{F}$ .
- 3)  $\{\mathbf{u}_t\}_{t=1}^T$  represents a sequence of T vectors, namely,  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_T\}$ .  $\{f_t\}_{t=1}^T$  denotes a sequence of T functions, which is  $\{f_1, f_2, \dots, f_T\}$ .
- 4)  $\mathcal{R}_T^A$  is the regret for a loss function sequence  $\{f_t\}_{t=1}^T \in \mathcal{F}^T$  with a learning algorithm  $A \in \mathcal{A}$ , where A can be POG or OG.
- 5)  $\|\cdot\|_p$  denotes the  $\ell_p$ -norm.  $\|\cdot\|$  represents the  $\ell_2$ -norm by default.
- 6) ≤ means "less than equal up to a constant factor" and ≥ means "greater than equal up to a constant factor."
   ∂ represents the subgradient operator. E represents the mathematical expectation.

## **B.** Assumptions

We use the following assumptions to analyze the regret of the OG.

Assumption 1: Functions  $F_t : \mathcal{X} \subset \mathbb{R}^d \mapsto \mathbb{R}$  for all  $t \in [T]$  and  $H : \mathcal{X} \subset \mathbb{R}^d \mapsto \mathbb{R}$  are convex and closed but possibly nondifferential. In particular,  $f_t \in \mathcal{F}$  is defined as  $f_t(\mathbf{x}) = F_t(\mathbf{x}) + H(\mathbf{x})$ .

Assumption 2: The convex compact set  $\mathcal{X}$  is the domain for  $F_t$  and H, and  $\|\mathbf{x}-\mathbf{y}\|^2 \leq R$  for any  $\mathbf{x}, \mathbf{y} \in \mathcal{X}$ . Besides, for any  $\mathbf{x} \in \mathcal{X}$  and function  $F_t$ ,  $\|G_t(\mathbf{x})\|^2 \leq G$ , where  $G_t(\mathbf{x}) \in \partial F_t(\mathbf{x})$ .

Algorithm 1 POG			
D	<b>F1.</b> 1	•	

**Require:** The learning rate  $\eta_t$  with  $1 \le t \le T$ . 1: for t = 1, 2, ..., T do

2: Predict  $\mathbf{x}_t$ .

3: Observe the loss function  $f_t$  with  $F_t$  and H, and suffer loss  $f_t(\mathbf{x}_t) = F_t(\mathbf{x}_t) + H(\mathbf{x}_t)$ .

4: Query subgradient  $G_t(\mathbf{x}_t) \in \partial F_t(\mathbf{x}_t)$ .

5:  $\mathbf{x}_{t+1} = \mathbf{prox}_{H,\eta_t}(\mathbf{x}_t - \eta_t G_t(\mathbf{x}_t)).$ 

6: return 
$$\mathbf{x}_{T+1}$$

## IV. Algorithm

We use the POG for solving the online learning problem with  $f_t(\cdot)$  in the form of (5). The POG algorithm is a general version of OG for taking care of the regularizer component  $H(\cdot)$  in  $f_t(\cdot)$ . The complete POG algorithm is presented in Algorithm 1. Line 4 of Algorithm 1 is the proximal gradient descent step defined by

$$\mathbf{x}_{t+1} = \mathbf{prox}_{H,\eta_t}(\mathbf{x}_t - \eta_t G_t(\mathbf{x}_t))$$

where the proximal operator is defined as

$$\operatorname{prox}_{H,\eta_t}(\mathbf{x}') := \arg\min_{\mathbf{x}\in\mathcal{X}} \left\{ H(\mathbf{x}) + \frac{1}{2\eta_t} \|\mathbf{x} - \mathbf{x}'\|^2 \right\}$$

Therefore, the update of  $\mathbf{x}_{t+1}$  is also equivalent to

$$\mathbf{x}_{t+1} = \mathbf{prox}_{H,\eta_t}(\mathbf{x}_t - \eta_t G_t(\mathbf{x}_t))$$
  
=  $\arg\min_{\mathbf{x}\in\mathcal{X}} \langle G_t(\mathbf{x}_t), \mathbf{x} \rangle + \frac{1}{2\eta_t} \|\mathbf{x} - \mathbf{x}_t\|^2 + H(\mathbf{x})$ 

The POG algorithm reduces to the OG algorithm when  $H(\cdot)$  is a constant function.

#### V. THEORETICAL RESULTS

Recall that we now consider an online learning problem with a dynamic constraint

$$\mathcal{L}_{D_{\beta}}^{T} := \left\{ \left\{ \mathbf{y}_{t} \right\}_{t=1}^{T} : \sum_{t=1}^{T-1} t^{\beta} \cdot \|\mathbf{y}_{t+1} - \mathbf{y}_{t}\| \le D_{\beta} \right\}$$

which is more general comparing with the previous definition of the dynamic constraint  $\mathcal{L}_{D_0}^T$  defined in (3).

When  $\beta = 0$ ,  $\mathcal{L}_{D_{\beta}}^{T}$  reduces to the previous definition of the dynamic constraint. Comparing with the previous definition, when  $\beta \ge 0$ ,  $D_{\beta}$  allocates larger weights for the future parts of the dynamics than the previous parts.

*Remark 1:* It is worth noting that  $D_{\beta}$  is a predefined parameter to restrict the change of reference models.

In this section, we first present a lower bound, which was not well studied in previous literature to our best knowledge. Then, we prove an upper bound for the regret based on our general dynamic constraints via POG, which holds for a general dynamic regret, instead of  $\beta = 0$  shown in previous work. We will show that our proved upper bound matches the lower bound, implying the optimality of POG algorithm. A. General Lower Bound for Online Convex Optimization

Once we obtain the upper bound for dynamic regret via POG, namely  $\sup_{\{f_i\}_{i=1}^T \in \mathcal{F}^T} \mathcal{R}_T^{\text{POG}}$ , there still remains a question, whether our upper bound's dependence on  $D_\beta$  and T is tight enough or even optimal.

Unfortunately, to our best knowledge, this question has not been fully investigated in any existing literature, even for the case of the dynamic regret defined with  $D_0$ .

To answer this question, we attempt to explore the value of  $\sup_{\{f_i\}_{i=1}^T \in \mathcal{F}^T} \mathcal{R}_T^A$  for the optimal algorithm  $A \in \mathcal{A}$ , which is formally written as  $\inf_{A \in \mathcal{A}} \sup_{\{f_i\}_{i=1}^T \in \mathcal{F}^T} \mathcal{R}_T^A$ . If a lower bound for  $\inf_{A \in \mathcal{A}} \sup_{\{f_i\}_{i=1}^T \in \mathcal{F}^T} \mathcal{R}_T^A$  matches the upper bound in (10), then we can say that POG is optimum for dynamic regret in online learning.

Theorem 1: Assume that Assumptions 1 and 2 hold. For any  $0 \le \beta < 1$ , the lower bound for our problem with dynamic regret is

$$\inf_{A \in \mathcal{A}} \sup_{\{f_i\}_{i=1}^T \in \mathcal{F}^T} \mathcal{R}_T^A \gtrsim \sqrt{D_\beta \cdot T^{1-\beta}} + \sqrt{T}$$

where  $\mathcal{A}$  is the set of all possible learning algorithms.  $f_t(\mathbf{x}) = F_t(\mathbf{x}) + H(\mathbf{x}), \forall t \in [T]$ , with  $\{f_t\}_{t=1}^T \in \mathcal{F}^T$ .

The discussion for the lower bound is conducted in the following aspects.

- Insight: The lower bound in Theorem 1 can be interpreted by that for any algorithm, there always exists a problem (or a function sequence in F<sup>T</sup> such that the dynamic regret is not less than (D<sub>β</sub> · T<sup>1-β</sup>)<sup>1/2</sup> + (T)<sup>1/2</sup> up to a constant factor. It indicates that the lower bound matches with the upper bound shown in (10). This theoretical result implies that the POG is an optimal algorithm to find decisions in the dynamic environment defined by D<sub>β</sub> and our upper bound (shown in Section V-B) is also sufficiently tight. In addition, this lower bound also reveals the difficulty of following dynamics in online learning.
- Novelty: Zhang et al. [39] showed a lower bound for dynamic regret. Comparing with the known result, our lower bound has the following novelty.
  - 1) General Bound: Our lower bound holds for any  $0 \le \beta < 1$ , but the result in [39] only holds for the case of  $\beta = 0$ . When  $\beta > 0$ , it is the first work to show that the dynamic regret is  $\Omega((D_{\beta} \cdot T^{1-\beta})^{1/2} + (T)^{1/2}).$
  - 2) *Nondifferentiable*  $f_t$ : Our lower bound holds for the nondifferentiable function sequence  $\{f_t\}_{t=1}^T$ , but Zhang *et al.* [39] only held for the differentiable function sequence  $\{f_t\}_{t=1}^T$ .

### B. Upper Bound for a General Dynamic Regret ( $0 \le \beta < 1$ )

We provide the upper bound for the POG algorithm described in Algorithm 1 in the following. The complete proof is provided in the Appendix. It essentially follows the analysis framework for the OG algorithm. The main novelty lies that our analysis is more general than previous work. Our upper bound holds for a general dynamic regret, that is,  $0 \le \beta < 1$ , instead of  $\beta = 0$  in previous studies.

Theorem 2: Let  $0 \le \beta < 1$ . Choose the positive learning rate sequence  $\{\eta_t\}_{t=1}^T$  in Algorithm 1 to be nonincreasing. Under Assumptions 1 and 2, the following upper bound for the dynamic regret holds:

IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS

$$\sup_{\{f_t\}_{t=1}^T \in \mathcal{F}^T} \mathcal{R}_T^{\text{POG}} \leq \sqrt{R} \max_{\{\eta_t\}_{t=1}^T} \left\{ \frac{1}{\eta_t \cdot t^\beta} \right\} \cdot D_\beta + \frac{R}{2\eta_T} + \frac{G}{2} \sum_{t=1}^T \eta_t + H(\mathbf{x}_1) - H(\mathbf{x}_{T+1}).$$
(9)

To make the dynamic regret more clear, we choose the learning rate appropriately, which leads to the following result.

Corollary 1: For any  $0 \le \beta < 1$ , we choose an appropriate  $\gamma$  such that  $\gamma \ge \beta$  and  $0 \le \gamma < 1$ . Set the learning rate  $\eta_t$  by

$$\eta_t = t^{-\gamma} \cdot \sqrt{\frac{(1-\gamma)\left(2\sqrt{R}T^{2\gamma-\beta-1}D_\beta + RT^{2\gamma-1}\right)}{G}}$$

in Algorithm 1. Under Assumptions 1 and 2, we have

$$\sup_{\{f_t\}_{t=1}^T \in \mathcal{F}^T} \mathcal{R}_T^{\text{POG}} \lesssim \sqrt{D_\beta \cdot T^{1-\beta}} + \sqrt{T}.$$
 (10)

In order to compare the upper bound in (10) with existing results, we consider the special case that does not include the nonsmooth term  $H(\cdot)$  in the objective and a particular choice for  $\beta = 0$ . In such case, our upper bound is  $\mathcal{O}((TD_0)^{1/2} + (T)^{1/2})$ , which is consistent with the known regret [38], [39]. When  $\beta > 0$ , our upper bound is  $\mathcal{O}((T^{1-\beta}D_{\beta})^{1/2} + (T)^{1/2})$ , which extends the known result [38], [39]. In addition, the upper bound in [39] requires that the loss function  $f_t$  is differentiable. However, our upper bound still holds for nondifferentiable  $f_t$ .

1) Discussion About the Learning Rate: Corollary 1 holds under the condition that the dynamics  $D_{\beta}$  is known and can be used to tune the learning rate  $\eta_t$ . However, knowing the dynamics may be not realistic in the online setting, especially in the general dynamic environment. One of the promising extensions about the work is to investigate how to estimate the dynamics  $D_{\beta}$  based on the observed data for some specific online learning application scenarios.

2) Connections With M-Shifting Regret: Although the shifting regret defined in (8) is different from the dynamic regret considered in this article, it is worth noting that our result in (10) also implies an upper bound  $\mathcal{O}((MT)^{1/2} + (T)^{1/2})$ with respect to the shifting regret defined in (8).

Corollary 2: Set the learning rate  $\eta_t$  by

$$\eta_t = t^{-\gamma} \cdot \sqrt{\frac{(1-\gamma)\left(2RT^{2\gamma-1}M + RT^{2\gamma-1}\right)}{G}}$$

in Algorithm 1. Under Assumptions 1 and 2, we have

$$\sup_{f_t\}_{t=1}^T \in \mathcal{F}^T} \widetilde{\mathcal{R}}_T^{\text{POG}} \lesssim \sqrt{MT} + \sqrt{T}$$

where  $\widetilde{\mathcal{R}}_{T}^{\text{POG}}$  follows the definition in (8). The proof is provided in the Appendix, that is, when *M* in (8) is known and can be used to tune the learning rate  $\{\eta_t\}_{t=1}^T$ , the previous *M*-shifting regret has the similar bound with our results. However, when *M* is unknown, our result in (10) implies an



Fig. 1. Comparison of average loss between POG and Ader. (a) Synthetic data. (b) Usenet2. (c) Spam.

upper bound  $\mathcal{O}((T)^{1/2}M + (T)^{1/2})$ . Under this condition, the existing result for the shifting regret in [6] is  $(MT \log T)^{1/2}$ , which obtains better dependence on the unknown M. It is highlighted that the existing bound does not require the knowledge of the dynamics M.

### VI. EMPIRICAL STUDIES

In this section, we conduct experiments in the dynamic environment and show the effectiveness of our method to follow the dynamics.

#### A. Experimental Settings

In experiments, we conduct online logistic regression to test the performance of POG in the dynamic environment. The loss function at the *t*th round is

$$f_t(\mathbf{x}) = \log(1 + \exp(-\mathbf{y}_t \mathbf{A}_t^{\top} \mathbf{x})) + \frac{i}{2} \|\mathbf{x}\|^2$$

where  $\mathbf{A}_t$  and  $\mathbf{y}_t$  are the instance observed at the *t*th round and its label, respectively, and  $\iota = 10^{-3}$  is a given hyperparameter. In addition, we compare the performance of POG, that is, Algorithm 1, with the state-of-the-art method Ader [39] in the setting of  $\beta = 0$ . We evaluate the performance by measuring the average loss:  $(1/T) \sum_{t=1}^{T} f_t(\mathbf{x}_t)$ , instead of using the dynamic regret

$$\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \min_{\{\mathbf{y}_t\}_{t=1}^{T} \in \mathcal{L}_{D_0}^{T}} \sum_{t=1}^{T} f_t(\mathbf{y}_t)$$

directly. The reason is that the optimal reference points  $\{\mathbf{x}_t^*\}_{t=1}^T := \operatorname{argmin}_{\{\mathbf{y}_t\}_{t=1}^T \in \mathcal{L}_{D_0}^T} \sum_{t=1}^T f_t(\mathbf{y}_t)$  are the same for both POG and Ader.

We conduct experiments on a synthetic dataset and two real datasets, where the dynamic environment is due to that the data distribution of those datasets keeps changing over time. The synthetic dataset is generated as follows. A data matrix  $\mathbf{A} \in \mathbb{R}^{T \times 10}$  consists of *T* instances, where every instance is represented by a row of  $\mathbf{A}$ . Specifically, the *t*th row of  $\mathbf{A}$ , that is  $\mathbf{A}_t$ , represents the instance  $\mathbf{A}_t$  at the *t*th round. The elements of the instance  $\mathbf{A}_t$  are generated according to  $\mathbf{y}_t \in \{1, -1\}$ . When  $\mathbf{y}_t = 1$ ,  $\mathbf{A}_t$  is generated by sampling from a time-varying distribution  $N((1 + 0.5 \sin(t)) \cdot \mathbf{1}, \mathbf{I})$ . When  $\mathbf{y}_t = -1$ ,  $\mathbf{A}_t$  is generated by sampling from another time-varying distribution

 $N((-1+0.5 \sin(t)) \cdot \mathbf{1}, \mathbf{I})$ . In addition, the real public datasets include *usenet2*<sup>4</sup> (1500 samples) and *spam*<sup>5</sup> (9324 samples). Both *usenet2* and *spam* are "concept drift" datasets [52], for which the optimal model changes over time.

Finally, the dynamic budget  $D_0$  is fixed as  $D_0 = 10$ . The learning rate  $\eta_t$  is set to be  $\eta_t = (10^{-3}/(t)^{1/2})$ . Ader is an "expert" algorithm, where the number of experts is set optimally according to [39, Th. 3]. All step sizes used in Ader are set to be  $(10^{-3}/(T)^{1/2})$ .

## B. Numerical Results

As shown in Fig. 1, both POG and Ader can decrease the average loss effectively. Specifically, POG achieves significantly better performance than Ader for the synthetic data and the *usenet2* dataset, and both of them achieve a similar performance for the *spam* dataset. Since Ader has been proved to be an optimal online learning method to follow the dynamics in the setting of  $\beta = 0$  and  $D_0 > 0$ , POG can also be verified to be optimal in this setting.

## VII. CONCLUSION

The online learning problem with dynamic regret metric is particularly interesting for many real scenarios. Although the OG method has been shown to be optimal for the static regret metric, the optimal algorithm for the dynamic regret remains unknown. This article studies this problem from a theoretical perspective. We show that POG, a general version of OG, is optimum to the dynamic regret by showing that our proved lower bound matches the upper bound, which slightly improves the existing upper bound.

#### **APPENDIX: PROOFS**

In this section, we present the detailed proofs for the theorems in this article. In particular, some necessary lemmas used in proofs to theorems are placed in the Supplementary Materials.

In our proofs, we abuse the notations of  $\partial H(\mathbf{x})$  a little bit to represent any vector in the subgradient of  $H(\mathbf{x})$ .  $G_t(\mathbf{x})$  still represents any vector in  $\partial F_t(\mathbf{x})$ . We use

<sup>&</sup>lt;sup>4</sup>http://mlkd.csd.auth.gr/concept\_drift.html

<sup>&</sup>lt;sup>5</sup>http://mlkd.csd.auth.gr/concept\_drift.html

IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS

 $B_{\psi}(\mathbf{x}, \mathbf{y}) := \psi(\mathbf{x}) - \psi(\mathbf{y}) - \langle \psi(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle$  to denote the Bregman divergence with respect to the function  $\psi$ .

*Lemma 1:* Consider a sequence  $\{\mathbf{v}_t\}_{t=1}^T$ . For any  $t \in [T]$ , dimensions of  $\mathbf{v}_t \in \{\pm 1\}^d$  are independent identically distributed (i.i.d.) sampled from the Rademacher distribution. We have

$$\mathbb{E}_{\left\{\mathbf{v}_{t}\right\}_{t=1}^{T}}\left\|\sum_{t=1}^{T}\mathbf{v}_{t}\right\|_{1} \gtrsim d\sqrt{T}$$

Proof: We consider the left-hand side

$$\mathbb{E}_{\{\mathbf{v}_t\}_{t=1}^T} \left\| \sum_{t=1}^T \mathbf{v}_t \right\|_1 = \mathbb{E}_{\{\mathbf{v}_t\}_{t=1}^T} \sum_{i=1}^d \left| \sum_{t=1}^T \mathbf{v}_t(i) \right| \\
= d \cdot \mathbb{E}_{\{\mathbf{v}_t(1)\}_{t=1}^T} \left| \sum_{t=1}^T \mathbf{v}_t(1) \right| \tag{11}$$

where  $\mathbf{v}_t(i)$  denotes the *i*th dimension of  $\mathbf{v}_t$  and  $\{\mathbf{v}_t(1)\}_{t=1}^T := \{\mathbf{v}_1(1), \mathbf{v}_2(1), \dots, \mathbf{v}_T(1)\}$ . The second equality holds because every dimension of  $\mathbf{v}_t$  is independent of each other.

Consider the sequence  $\{\mathbf{v}_t\}_{t=1}^T$ . If the event, +1 is picked, happens *m* times with the probability  $P_m$ , then the event, -1 is picked, happens T - m times. Denote  $S_T := \sum_{t=1}^T \mathbf{v}_t(1)$ , and we have

$$S_T = m - (T - m) = 2m - T.$$

Denote  $S := \{-T, -T + 2, \dots, T - 2, T\}$  and  $S_T \in S$ . Thus, we have

$$P(S_T = 2m - T) = P_m = \frac{1}{2^T} \cdot \binom{T}{m}$$

and

$$\mathbb{E} |S_T| = \sum_{m=0}^{T} \frac{|2m - T|}{2^T} \cdot {\binom{T}{m}} \\ = \frac{1}{2^T} \cdot \sum_{m=0}^{T} \frac{|2m - T| \cdot T!}{m! \cdot (T - m)!}.$$

When T is even, denote T = 2J. Thus,

$$\begin{split} \mathbb{E} |S_{T}| \\ &= \frac{1}{2^{2J}} \cdot \sum_{m=0}^{T} \frac{|2m - 2J| \cdot (2J)!}{m! \cdot (2J - m)!} \\ &= \frac{(2J)!}{2^{2J}} \cdot \sum_{m=0}^{2J} \frac{|2m - 2J|}{m! \cdot (2J - m)!} \\ \stackrel{(1)}{=} \frac{(2J)!}{2^{2J-2}} \cdot \sum_{n=0}^{J} \frac{n}{(J + n)! \cdot (J - n)!} \\ &= \frac{1}{2^{2J-2}} \cdot \left( \sum_{n=0}^{J} (n + J) \binom{2J}{J + n} - \sum_{n=0}^{J} J \binom{2J}{J + n} \right) \right) \\ &= \frac{1}{2^{2J-2}} \cdot \left( \sum_{i=J}^{2J} i \binom{2J}{i} - \sum_{i=J}^{2J} J \binom{2J}{i} \right) \\ \stackrel{(2)}{=} \frac{1}{2^{2J-2}} \cdot \left( \sum_{i=J}^{2J} 2J \binom{2J-1}{i-1} - \sum_{i=J}^{2J} J \binom{2J}{i} \right) \end{split}$$

$$\begin{array}{l} \underbrace{\textcircled{3}}_{=} & \frac{2J}{2^{2J-2}} \cdot \left( \sum_{k=J-1}^{2J-1} \binom{2J-1}{k} - \frac{1}{4} \binom{2^{2J}}{2^{2J}} + \binom{2J}{J} \right) \right) \\ &= & \frac{2J}{2^{2J-2}} \cdot \left( \sum_{k=J}^{2J-1} \binom{2J-1}{k} + \binom{2J-1}{J-1} - \frac{1}{4} \binom{2^{2J}}{2^{2J}} + \binom{2J}{J} \right) \right) \\ \underbrace{\textcircled{4}}_{=} & \frac{2J}{2^{2J-2}} \cdot \left( 2^{2J-2} + \binom{2J-1}{J-1} - \frac{1}{4} \binom{2^{2J}}{2^{2J}} + \binom{2J}{J} \right) \\ &= & \frac{2J}{2^{2J-2}} \cdot \left( \binom{2J-1}{J-1} - \frac{1}{4} \binom{2J}{J} \right) \\ &= & \frac{2J}{2^{2J-2}} \cdot \left( \frac{1}{4} \binom{2J}{J} \right) \\ &= & \frac{2J}{4^{J}} \cdot \frac{(2J)!}{J! \cdot J!} \\ \underbrace{\textcircled{5}}_{\geq} & 2J \cdot \frac{1}{2\sqrt{J}} \\ &= & \sqrt{\frac{T}{2}}. \end{array}$$

Here, (1) holds due to

$$\frac{(2J)!}{2^{2J}} \cdot \sum_{m=0}^{2J} \frac{|2m-2J|}{m! \cdot (2J-m)!}$$

$$= \frac{(2J)!}{2^{2J}} \cdot \left(\sum_{m=0}^{J} \frac{2J-2m}{m! \cdot (2J-m)!} + \sum_{m=J+1}^{2J} \frac{2m-2J}{m! \cdot (2J-m)!}\right)$$

$$= \frac{(2J)!}{2^{2J}} \cdot \sum_{n_1=0}^{J} \frac{2n_1}{(J-n_1)! \cdot (J+n_1)!}$$

$$+ \frac{(2J)!}{2^{2J}} \cdot \sum_{n_2=1}^{J} \frac{2n_2}{(J+n_2)! \cdot (J-n_2)!}$$

$$= \frac{(2J)!}{2^{2J}} \cdot \sum_{n_1=0}^{J} \frac{2n_1}{(J-n_1)! \cdot (J+n_1)!}$$

$$+ \frac{(2J)!}{2^{2J-2}} \cdot \sum_{n_2=0}^{J} \frac{2n_2}{(J+n_2)! \cdot (J-n_2)!}$$

$$= \frac{(2J)!}{2^{2J-2}} \cdot \sum_{n=0}^{J} \frac{n}{(J-n)! \cdot (J+n)!}.$$
(2) holds because, for any  $0 \le k \le N$ 

$$k \binom{N}{k} = N \binom{N-1}{k-1}.$$

3 holds because

$$2^{2J} = \sum_{i=0}^{J} {\binom{2J}{i}} + \sum_{i=J}^{2J} {\binom{2J}{i}} - {\binom{2J}{J}}$$
$$= 2\sum_{i=J}^{2J} {\binom{2J}{i}} - {\binom{2J}{J}}.$$

(4) holds because

$$2^{2J-1} = \sum_{i=0}^{J-1} {\binom{2J-1}{i}} + \sum_{i=J}^{2J-1} {\binom{2J-1}{i}} = 2\sum_{i=J}^{2J-1} {\binom{2J-1}{i}}.$$

(5) holds because, for any n > 1

$$\frac{1}{4^n} \binom{2n}{n} \ge \frac{1}{2\sqrt{n}}.$$

When T is odd, we have

$$\mathbb{E} |S_T| = \mathbb{E} |S_{T-1} + \mathbf{v}_T(1)|$$
  

$$\geq \mathbb{E} |S_{T-1}| - \mathbb{E} |\mathbf{v}_T(1)|$$
  

$$= \mathbb{E} |S_{T-1}| - 1$$
  

$$= \sqrt{\frac{T}{2}} - 1.$$

Finally, we obtain

$$\mathbb{E}_{\mathbf{v}_t\}_{t=1}^T} \left\| \sum_{t=1}^T \mathbf{v}_t \right\|_1 \gtrsim d\sqrt{T}.$$

It completes the proof.

Proof of Theorem 1:

*Proof:* Let  $f_t(\mathbf{x}_t) = F_t(\mathbf{x}_t) + H(\mathbf{x}_t)$ , where  $F_t(\mathbf{x}_t) :=$  $\langle \mathbf{v}_t, \mathbf{x}_t \rangle$  and  $H(\mathbf{x}_t) = 0$  for all  $\mathbf{x}_t \in \mathcal{X}$ . Here,  $\mathbf{v}_t \in$  $\{+1, -1\}^d$  is a random vector with i.i.d. elements sampled from the Rademacher distribution.  $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_2 \leq 1\}$ and  $\mathcal{L}_{D_\beta}^T = \{\{\mathbf{y}_t\}_{t=1}^T : \sum_{t=1}^{T-1} t^\beta \cdot \|\mathbf{y}_{t+1} - \mathbf{y}_t\|_2 \leq D_\beta\}$ . Under this construction, for any given algorithm  $A \in \mathcal{A}$ , we have

$$\sup_{\{f_{t}\}_{t=1}^{T}} \mathcal{R}_{T}^{A} = \sup_{\{\mathbf{v}_{t}\}_{t=1}^{T}} \mathcal{R}_{T}^{A} \geq \mathbb{E}_{\{\mathbf{v}_{t}\}_{t=1}^{T}} \mathcal{R}_{T}^{A}$$

$$= \mathbb{E}_{\{\mathbf{v}_{t}\}_{t=1}^{T}} \sum_{t=1}^{T} f_{t}(\mathbf{x}_{t}) - \mathbb{E}_{\{\mathbf{v}_{t}\}_{t=1}^{T}} \left( \min_{\{\mathbf{y}_{t}\}_{t=1}^{T} \in \mathcal{L}_{D_{\beta}}^{T}} \sum_{t=1}^{T} f_{t}(\mathbf{y}_{t}) \right)$$

$$= \mathbb{E}_{\{\mathbf{v}_{t}\}_{t=1}^{T}} \sum_{t=1}^{T} \langle \mathbf{v}_{t}, \mathbf{x}_{t} \rangle - \mathbb{E}_{\{\mathbf{v}_{t}\}_{t=1}^{T}} \left( \min_{\{\mathbf{y}_{t}\}_{t=1}^{T} \in \mathcal{L}_{D_{\beta}}^{T}} \sum_{t=1}^{T} \langle \mathbf{v}_{t}, \mathbf{y}_{t} \rangle \right)$$

$$= 0 - \mathbb{E}_{\{\mathbf{v}_{t}\}_{t=1}^{T}} \left( \min_{\{\mathbf{y}_{t}\}_{t=1}^{T} \in \mathcal{L}_{D_{\beta}}^{T}} \sum_{t=1}^{T} \langle \mathbf{v}_{t}, \mathbf{y}_{t} \rangle \right)$$

$$= \mathbb{E}_{\{\mathbf{v}_{t}\}_{t=1}^{T}} \left( \max_{\{\mathbf{y}_{t}\}_{t=1}^{T} \in \mathcal{L}_{D_{\beta}}^{T}} \sum_{t=1}^{T} \langle \mathbf{v}_{t}, \mathbf{y}_{t} \rangle \right)$$

$$= \mathbb{E}_{\{\mathbf{v}_{t}\}_{t=1}^{T}} \left( \max_{\{\mathbf{y}_{t}\}_{t=1}^{T} \in \mathcal{L}_{D_{\beta}}^{T}} \sum_{t=1}^{T} \langle \mathbf{v}_{t}, \mathbf{y}_{t} \rangle \right)$$

$$(12)$$

(1) holds since the Rademacher distribution is a symmetric distribution.

Next, we try to estimate the lower bound  $\mathbb{E}_{\{\mathbf{v}_t\}_{t=1}^T} \max_{\{\mathbf{y}_t\}_{t=1}^T \in \mathcal{L}_{D_\beta}^T} \sum_{t=1}^T \langle \mathbf{v}_t, \mathbf{y}_t \rangle.$  One feasible solution of  $\mathbf{y}_t$  is constructed as follows. the lower bound of

- 1) Evenly split the sequence  $\{\mathbf{y}_t\}_{t=1}^T$  into two subsequences:  $\{\bar{\mathbf{y}}_t\}_{t=1}^{T_1} := \{\mathbf{y}_t\}_{t=1}^{(T/2)-1}$  and  $\{\hat{\mathbf{y}}_t\}_{t=1}^{T_2} := \{\mathbf{y}_t\}_{t=(T/2)}^T$ , where  $T_1 = T_2 = (T/2)$ .  $\{\mathbf{v}_t\}_{t=1}^T$  is also split into  $\{\bar{\mathbf{v}}_t\}_{t=1}^{T_1}$  and  $\{\hat{\mathbf{v}}_t\}_{t=1}^{T_2}$ .
- 2) Let all  $\|\mathbf{y}_t\|_2 \le (1/2)$ . 3) Evenly split  $\{\bar{\mathbf{y}}_t\}_{t=1}^{T_1}$  into  $N := [(D_\beta/T_1^\beta)]$ subsets  $\{\mathbf{y}_t\}_{t=1}^{(T_1/N)}, \{\mathbf{y}_t\}_{t=(T_1/N)+1}^{(2T_1/N)}, \{\mathbf{y}_t\}_{t=(2T_1/N)+1}^{(3T_1/N)}, \dots,$  $\{\mathbf{y}_t\}_{t=((N-1)T_1/N)+1}^{T_1}$ .
- 4) For the first subsequence  $\{\bar{\mathbf{y}}_t\}_{t=1}^{T_1}$ , within the *i*th subset, let the values in it be same, and denote it by  $\mathbf{u}_i$ . For the second subsequence  $\{\hat{\mathbf{y}}_t\}_{t=1}^{T_2}$ , let all values be  $\mathbf{u}_N$ .

5) Since elements in the second subsequence  $\{\hat{\mathbf{y}}_t\}_{t=1}^{I_2}$  have the same value **u**, the difference between two elements is 0. In addition, consider the first subsequence  $\{\bar{\mathbf{y}}_t\}_{t=1}^{I_1}$ . Elements in different subsets can be different such that  $\|\mathbf{u}_{i+1} - \mathbf{u}_i\| \le \|\mathbf{u}_{i+1}\| + \|\mathbf{u}_i\| \le 1$ . We have

$$\sum_{t=1}^{T-1} t^{\beta} \cdot \|\mathbf{y}_{t+1} - \mathbf{y}_t\| = \sum_{t=1}^{T_1-1} t^{\beta} \cdot \|\mathbf{y}_{t+1} - \mathbf{y}_t\| + 0$$
$$= \sum_{i=1}^{N-1} \|\mathbf{u}_{i+1} - \mathbf{u}_i\| \cdot \left(\frac{T_1}{N} \cdot i\right)^{\beta}$$
$$\leq T_1^{\beta} \sum_{i=1}^{N-1} \left(\frac{i}{N}\right)^{\beta}$$
$$\leq T_1^{\beta} (N-1)$$
$$\leq D_{\beta}.$$

It implies that  $\{\bar{\mathbf{y}}_t\}_{t=1}^{T_1}$  and  $\{\hat{\mathbf{y}}_t\}_{t=1}^{T_2}$  under our construction are feasible.

Based on the above steps, we have

$$\mathbb{E}_{\{\mathbf{v}_{t}\}_{t=1}^{T}} \left( \max_{\{\mathbf{y}_{t}\}_{t=1}^{T} \in \mathcal{L}_{D_{\beta}}^{T}} \sum_{t=1}^{T} \langle \mathbf{v}_{t}, \mathbf{y}_{t} \rangle \right)$$

$$= \mathbb{E}_{\{\bar{\mathbf{v}}_{t}\}_{t=1}^{T}} \max_{\{\mathbf{u}_{t}\}_{t=1}^{N}} \sum_{i=1}^{N} \left\langle \sum_{t=\frac{iT_{1}}{N}+1}^{U} \bar{\mathbf{v}}_{t}, \mathbf{u}_{i} \right\rangle$$

$$+ \mathbb{E}_{\{\bar{\mathbf{v}}_{t}\}_{t=1}^{T}} \max_{\mathbf{u}} \sum_{t=1}^{T_{2}} \langle \hat{\mathbf{v}}_{t}, \mathbf{u} \rangle$$

$$= \frac{1}{2} \mathbb{E}_{\{\bar{\mathbf{v}}_{t}\}_{t=1}^{T}} \max_{\mathbf{u}} \sum_{t=1}^{T_{2}} \langle \hat{\mathbf{v}}_{t}, \mathbf{u} \rangle$$

$$+ \frac{1}{2} \mathbb{E}_{\{\bar{\mathbf{v}}_{t}\}_{t=1}^{T}} \max_{\mathbf{z} \in \mathcal{X}} \sum_{t=1}^{T_{2}} \langle \hat{\mathbf{v}}_{t}, \mathbf{z} \rangle$$

$$\left( \bigcup_{t=\frac{1}{N}+1}^{U} \sum_{\{\bar{\mathbf{v}}_{t}\}_{t=1}^{T_{2}}} \sum_{\mathbf{z} \in \mathcal{X}} \sum_{t=1}^{T_{2}} \langle \hat{\mathbf{v}}_{t}, \mathbf{z} \rangle$$

$$\left( \bigcup_{t=\frac{1}{2}} \frac{1}{2} N \cdot \mathbb{E}_{\{\bar{\mathbf{v}}_{t}\}_{t=1}^{T_{1}}} \left\| \sum_{t=\frac{iT_{1}}{N}+1}^{U} \bar{\mathbf{v}}_{t} \right\| + \frac{1}{2} \mathbb{E}_{\{\bar{\mathbf{v}}_{t}\}_{t=1}^{T_{2}} \hat{\mathbf{v}}_{t} \right\|$$

$$\left( \bigcup_{t=\frac{1}{2}} \frac{1}{2} N \frac{1}{\sqrt{d}} \mathbb{E}_{\{\bar{\mathbf{v}}_{t}\}_{t=1}^{T_{1}}} \left\| \sum_{t=\frac{iT_{1}}{N}+1}^{U} \bar{\mathbf{v}}_{t} \right\| + \frac{1}{2\sqrt{d}} \mathbb{E}_{\{\bar{\mathbf{v}}_{t}\}_{t=1}^{T_{2}}} \hat{\mathbf{v}}_{t} \right\|_{1}$$

$$\left( \bigcup_{t=\frac{1}{2}} \frac{\sqrt{d}}{2} \cdot \sqrt{T_{1}N} + \frac{\sqrt{d}}{2} \cdot \sqrt{T_{2}}} \right)$$

$$(13)$$

Recall that  $\mathcal{X} = {\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_2 \le 1}$  in this example. ① holds due to  $\|\mathbf{x}\|_* = \|\mathbf{x}\|_2 = \max_{\|\mathbf{y}\| \le 1} \langle \mathbf{x}, \mathbf{y} \rangle$ . (2) holds due to  $\|\mathbf{x}\|_{1} \le (d)^{1/2} \|\mathbf{x}\|$ . (3) holds due to Lemma 1.

Since (14) holds for any algorithm  $A \in \mathcal{A}$ , we thus obtain

$$\inf_{A \in \mathcal{A}} \sup_{\{f_t\}_{t=1}^T \in \mathcal{F}^T} \mathcal{R}_T^A = \Omega\bigg(\sqrt{D_\beta \cdot T^{1-\beta}} + \sqrt{T}\bigg).$$

It completes the proof. Proof of Theorem 2:

IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS

*Proof:* For any sequence of T loss functions  $\{f_t\}_{t=1}^T \in \mathcal{F}^T$ , we have

$$\sum_{t=1}^{T} (F_t(\mathbf{x}_t) + H(\mathbf{x}_t) - F_t(\mathbf{y}_t) - H(\mathbf{y}_t))$$
  
=  $\sum_{t=1}^{T} (F_t(\mathbf{x}_t) + H(\mathbf{x}_{t+1}) - F_t(\mathbf{y}_t) - H(\mathbf{y}_t))$   
+  $H(\mathbf{x}_1) - H(\mathbf{x}_{t+1}).$ 

According to Lemma 3, we have

$$I_{0} = \sum_{t=1}^{T} (F_{t}(\mathbf{x}_{t}) + H(\mathbf{x}_{t+1}) - F_{t}(\mathbf{y}_{t}) - H(\mathbf{y}_{t}))$$

$$\leq \sum_{t=1}^{T} \frac{1}{2\eta_{t}} (\|\mathbf{y}_{t} - \mathbf{x}_{t}\|_{2}^{2} - \|\mathbf{y}_{t} - \mathbf{x}_{t+1}\|_{2}^{2})$$

$$+ \frac{1}{2} \sum_{t=1}^{T} \eta_{t} \|G_{t}(\mathbf{x}_{t})\|^{2}$$

$$\stackrel{(1)}{\leq} \sqrt{R} \sum_{t=1}^{T-1} \frac{1}{\eta_{t}} (\|\mathbf{y}_{t+1} - \mathbf{y}_{t}\|) + \frac{R}{2\eta_{T}} + \frac{G}{2} \sum_{t=1}^{T} \eta_{t}$$

$$\leq \sqrt{R} \max_{\{\eta_{t}\}_{t=1}^{T}} \left\{ \frac{1}{\eta_{t} \cdot t^{\beta}} \right\} \cdot D_{\beta} + \frac{R}{2\eta_{T}} + \frac{G}{2} \sum_{t=1}^{T} \eta_{t}.$$

① holds due to Lemma 4. Thus, we have

$$\sum_{t=1}^{T} (F_t(\mathbf{x}_t) + H(\mathbf{x}_t) - F_t(\mathbf{y}_t) - H(\mathbf{y}_t))$$

$$\leq \sqrt{R} \max_{\{\eta_t\}_{t=1}^{T}} \left\{ \frac{1}{\eta_t \cdot t^{\beta}} \right\} \cdot D_{\beta} + \frac{R}{2\eta_T} + \frac{G}{2} \sum_{t=1}^{T} \eta_t$$

$$+ H(\mathbf{x}_1) - H(\mathbf{x}_{T+1}). \tag{15}$$

Since (15) holds for any sequence of loss functions  $\{f_t\}_{t=1}^T \in \mathcal{F}^T$ , thus,

$$\sup_{\{f_t\}_{t=1}^T \in \mathcal{F}^T} \mathcal{R}_T^{\text{POG}} \leq \sqrt{R} \max_{\{\eta_t\}_{t=1}^T} \left\{ \frac{1}{\eta_t \cdot t^\beta} \right\} \cdot D_\beta + \frac{R}{2\eta_T} + \frac{G}{2} \sum_{t=1}^T \eta_t + H(\mathbf{x}_1) - H(\mathbf{x}_{T+1})$$

It completes the proof.

Proof of Corollary 1

*Proof:* Assume that  $\eta_t := t^{-\gamma} \cdot \sigma_1$ , where  $\sigma_1$  is a constant and does not depend on *t*. According to Theorem 2, when  $\gamma \ge \beta$ 

$$\max_{\{\eta_t\}_{t=1}^T} \left\{ \frac{1}{\eta_t \cdot t^{\beta}} \right\} = \frac{T^{\gamma - \beta}}{\sigma_1}$$

Substituting it into (9), we have

$$\mathcal{R}_{T}^{\text{POG}} \leq \frac{\sqrt{R}D_{\beta}}{\sigma_{1}}T^{\gamma-\beta} + \frac{R}{2\sigma_{1}}T^{\gamma} + \frac{G\sigma_{1}}{2}\sum_{t=1}^{T}t^{-\gamma} + H(\mathbf{x}_{1}) - H(\mathbf{x}_{T+1})$$

$$\stackrel{(1)}{=} \frac{\sqrt{RD_{\beta}}}{\sigma_{1}} T^{\gamma-\beta} + \frac{R}{2\sigma_{1}} T^{\gamma} + \frac{G\sigma_{1}}{2(1-\gamma)} T^{1-\gamma}$$
$$+ H(\mathbf{x}_{1}) - H(\mathbf{x}_{T+1}).$$

(1) holds due to  $0 \le \gamma < 1$  and Lemma 5. Choosing the optimal  $\sigma_1$  with

$$\sigma_1 = \sqrt{\frac{(1-\gamma)\left(2\sqrt{R}T^{2\gamma-\beta-1}D_{\beta} + RT^{2\gamma-1}\right)}{G}}$$

we have

$$\mathcal{R}_{T}^{\text{POG}} \leq \sqrt{\frac{2G\sqrt{R}D_{\beta}T^{1-\beta}}{1-\gamma}} + \sqrt{\frac{GRT}{4(1-\gamma)}} + H(\mathbf{x}_{1}) - H(\mathbf{x}_{T+1}) \\ \leq \sqrt{D_{\beta} \cdot T^{1-\beta}} + \sqrt{T}.$$

It completes the proof.

*Lemma 2:* The optimal reference points  $\{\mathbf{y}_t\}_{t=1}^T$  satisfying  $\sum_{t=1}^{T-1} \mathbb{1}\{\mathbf{y}_{t+1} \neq \mathbf{y}_t\} \leq M$  still satisfy  $\sum_{t=1}^{T-1} \|\mathbf{y}_{t+1} - \mathbf{y}_t\| \leq M(R)^{1/2}$ .

*Proof:* Denote  $a_t = \|\mathbf{y}_{t+1} - \mathbf{y}_t\|$  and  $\mathbf{a}_T = \{\mathbf{a}_t | t \in [T-1]\} \in \mathbb{R}^{T-1}$ . Note that  $\sum_{t=1}^{T-1} \mathbb{1}\{\mathbf{y}_{t+1} \neq \mathbf{y}_t\} = \|\mathbf{a}_T\|_0$ . Thus, for *M*-shifting regret,  $\sum_{t=1}^{T-1} \mathbb{1}\{\mathbf{y}_{t+1} \neq \mathbf{y}_t\} = \|\mathbf{a}_T\|_0 \leq M$ . When  $\beta = 0$ , we have T-1

$$\sum_{t=1} \|\mathbf{y}_{t+1} - \mathbf{y}_t\| = \|\mathbf{a}_T\|_1 \le \|\mathbf{a}_T\|_0 \sqrt{R} \le M\sqrt{R}.$$

The first inequality holds because, for any  $1 \le t \le T$ ,  $\|\mathbf{y}_{t+1} - \mathbf{y}_t\| \le (R)^{1/2}$ . It completes the proof.  $\Box$ *Proof of Corollary 2:* 

*Proof:* Replacing  $D_0$  by  $M(R)^{1/2}$  in Corollary 1, we have

$$\sup_{\{f_i\}_{i=1}^T \in \mathcal{F}^T} \mathcal{R}_T^{\text{POG}} \lesssim \sqrt{MT} + \sqrt{T}.$$

According to Lemma 2, we obtain

$$\sup_{\{f_i\}_{t=1}^T \in \mathcal{F}^T} \widetilde{\mathcal{R}}_T^{\text{POG}} \leq \sup_{\{f_i\}_{t=1}^T \in \mathcal{F}^T} \mathcal{R}_T^{\text{POG}} \lesssim \sqrt{MT} + \sqrt{T}.$$

It thus completes the proof.

*Lemma 3:* Given any sequence  $\{\mathbf{y}_t\}_{t=1}^T \in \mathcal{L}_{D_\beta}^T$  and setting any  $\eta_t > 0$  in Algorithm 1, we have

$$\sum_{t=1}^{T} (F_t(\mathbf{x}_t) + H(\mathbf{x}_{t+1}) - F_t(\mathbf{y}_t) - H(\mathbf{y}_t))$$
  
$$\leq \sum_{t=1}^{T} \frac{1}{2\eta_t} (\|\mathbf{y}_t - \mathbf{x}_t\|_2^2 - \|\mathbf{y}_t - \mathbf{x}_{t+1}\|_2^2) + \sum_{t=1}^{T} \frac{\eta_t}{2} \|G_t(\mathbf{x}_t)\|^2.$$

*Proof:* Define  $\psi(\mathbf{x}) := (1/2) \|\mathbf{x}\|^2$  and  $\mathbf{x}_{t+1} = \operatorname{argmin}_{\mathbf{x} \in \mathcal{X}} \langle G_t(\mathbf{x}_t), \mathbf{x} \rangle + (1/\eta_t) B_{\psi}(\mathbf{x}, \mathbf{x}_t) + H(\mathbf{x})$ , according to the optimal condition, for any  $\mathbf{x} \in \mathcal{X}$ , we have

$$0 \le \langle \mathbf{x} - \mathbf{x}_{t+1}, \eta_t G_t(\mathbf{x}_t) \rangle + \langle \mathbf{x} - \mathbf{x}_{t+1}, \nabla \psi(\mathbf{x}_{t+1}) - \nabla \psi(\mathbf{x}_t) + \eta_t \partial H(\mathbf{x}_{t+1}) \rangle.$$
(16)

Then, we have

$$\eta_t(F_t(\mathbf{x}_t) + H(\mathbf{x}_{t+1}) - F_t(\mathbf{y}_t) - H(\mathbf{y}_t)) \\ \leq \eta_t \langle \mathbf{x}_t - \mathbf{y}_t, G_t(\mathbf{x}_t) \rangle + \eta_t \langle \mathbf{x}_{t+1} - \mathbf{y}_t, \partial H(\mathbf{x}_{t+1}) \rangle$$

Authorized licensed use limited to: National Univ of Defense Tech. Downloaded on March 08,2022 at 13:20:24 UTC from IEEE Xplore. Restrictions apply.

ZHAO et al.: POG IS OPTIMUM FOR DYNAMIC REGRET

$$= \eta_t \langle \mathbf{x}_{t+1} - \mathbf{y}_t, G_t(\mathbf{x}_t) \rangle + \eta_t \langle \mathbf{x}_{t+1} - \mathbf{y}_t, \partial H(\mathbf{x}_{t+1}) \rangle + \eta_t \langle \mathbf{x}_t - \mathbf{x}_{t+1}, G_t(\mathbf{x}_t) \rangle \stackrel{(1)}{\leq} \langle \mathbf{y}_t - \mathbf{x}_{t+1}, \nabla \psi(\mathbf{x}_{t+1}) - \nabla \psi(\mathbf{x}_t) \rangle + \eta_t \langle \mathbf{x}_t - \mathbf{x}_{t+1}, G_t(\mathbf{x}_t) \rangle \stackrel{(2)}{=} B_{\psi}(\mathbf{y}_t, \mathbf{x}_t) - B_{\psi}(\mathbf{x}_{t+1}, \mathbf{x}_t) - B_{\psi}(\mathbf{y}_t, \mathbf{x}_{t+1}) + \eta_t \langle \mathbf{x}_t - \mathbf{x}_{t+1}, G_t(\mathbf{x}_t) \rangle \stackrel{(3)}{\leq} B_{\psi}(\mathbf{y}_t, \mathbf{x}_t) - B_{\psi}(\mathbf{y}_t, \mathbf{x}_{t+1}) + \frac{\eta_t^2}{2} \|G_t(\mathbf{x}_t)\|^2.$$

(1) holds due to (16). (2) holds due to three-point identity for Bregman divergence, which is, for any vectors  $\mathbf{x}$ ,  $\mathbf{y}$ , and  $\mathbf{z}$ 

$$B_{\psi}(\mathbf{x},\mathbf{y}) = B_{\psi}(\mathbf{x},\mathbf{z}) + B_{\psi}(\mathbf{z},\mathbf{y}) - \langle \mathbf{x} - \mathbf{z}, \nabla \psi(\mathbf{y}) - \nabla \psi(\mathbf{z}) \rangle.$$

(3) holds due to  $\psi(\mathbf{x}) = (1/2) \|\mathbf{x}\|_2^2$  so that  $B_{\psi}(\mathbf{x}_{t+1}, \mathbf{x}_t) = (1/2) \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2$ . Thus, we finally obtain

$$\sum_{t=1}^{T} (F_t(\mathbf{x}_t) + H(\mathbf{x}_{t+1}) - F_t(\mathbf{y}_t) - H(\mathbf{y}_t))$$

$$\leq \sum_{t=1}^{T} \frac{B_{\psi}(\mathbf{y}_t, \mathbf{x}_t) - B_{\psi}(\mathbf{y}_t, \mathbf{x}_{t+1})}{\eta_t} + \frac{1}{2} \sum_{t=1}^{T} \eta_t \|G_t(\mathbf{x}_t)\|^2$$

$$= \sum_{t=1}^{T} \frac{\|\mathbf{y}_t - \mathbf{x}_t\|_2^2 - \|\mathbf{y}_t - \mathbf{x}_{t+1}\|_2^2}{2\eta_t} + \sum_{t=1}^{T} \frac{\eta_t}{2} \|G_t(\mathbf{x}_t)\|^2.$$

It completes the proof.

*Lemma 4:* Given any sequence  $\{\mathbf{y}_t\}_{t=1}^T \in \mathcal{L}_{D_\beta}^T$  and setting a nonincreasing series  $0 < \eta_{t+1} \le \eta_t$  in Algorithm 1, we have

$$\sum_{t=1}^{T} \frac{1}{\eta_t} \left( -\|\mathbf{y}_t - \mathbf{x}_{t+1}\|^2 + \|\mathbf{y}_t - \mathbf{x}_t\|^2 \right)$$
$$\leq 2\sqrt{R} \sum_{t=1}^{T-1} \frac{1}{\eta_t} (\|\mathbf{y}_{t+1} - \mathbf{y}_t\|) + \frac{R}{\eta_T}.$$

*Proof:* According to the law of cosines, we have

$$-\|\mathbf{y}_{t} - \mathbf{x}_{t+1}\|^{2} + \|\mathbf{y}_{t+1} - \mathbf{x}_{t+1}\|^{2}$$

$$\leq 2\|\mathbf{y}_{t+1} - \mathbf{y}_{t}\|\|\mathbf{x}_{t+1} - \mathbf{y}_{t+1}\| - \|\mathbf{y}_{t+1} - \mathbf{y}_{t}\|^{2}$$

$$\leq 2\sqrt{R}\|\mathbf{y}_{t+1} - \mathbf{y}_{t}\| - \|\mathbf{y}_{t+1} - \mathbf{y}_{t}\|^{2}$$

$$\leq 2\sqrt{R}\|\mathbf{y}_{t+1} - \mathbf{y}_{t}\|.$$
(17)

Thus, we obtain

$$\sum_{t=1}^{T} \frac{1}{\eta_t} \left( -\|\mathbf{y}_t - \mathbf{x}_{t+1}\|^2 + \|\mathbf{y}_t - \mathbf{x}_t\|^2 \right)$$
  
=  $\sum_{t=1}^{T-1} \left( -\frac{1}{\eta_t} \|\mathbf{y}_t - \mathbf{x}_{t+1}\|^2 + \frac{1}{\eta_{t+1}} \|\mathbf{y}_{t+1} - \mathbf{x}_{t+1}\|^2 \right)$   
+  $\frac{1}{\eta_1} \|\mathbf{y}_1 - \mathbf{x}_1\|^2 - \frac{1}{\eta_T} \|\mathbf{y}_T - \mathbf{x}_{T+1}\|^2$   
 $\leq \sum_{t=1}^{T-1} \left( -\frac{1}{\eta_t} \|\mathbf{y}_t - \mathbf{x}_{t+1}\|^2 + \frac{1}{\eta_t} \|\mathbf{y}_{t+1} - \mathbf{x}_{t+1}\|^2 \right)$   
+  $\sum_{t=1}^{T-1} \left( \frac{1}{\eta_{t+1}} - \frac{1}{\eta_t} \right) \|\mathbf{y}_{t+1} - \mathbf{x}_{t+1}\|^2 + \frac{1}{\eta_1} \|\mathbf{y}_1 - \mathbf{x}_1\|^2$ 

$$\leq \sum_{t=1}^{T-1} \left( -\frac{1}{\eta_t} \| \mathbf{y}_t - \mathbf{x}_{t+1} \|^2 + \frac{1}{\eta_t} \| \mathbf{y}_{t+1} - \mathbf{x}_{t+1} \|^2 \right) \\ + R \sum_{t=1}^{T-1} \left( \frac{1}{\eta_{t+1}} - \frac{1}{\eta_t} \right) + \frac{R}{\eta_1} \\ \stackrel{\bigoplus}{\leq} 2\sqrt{R} \sum_{t=1}^{T-1} \frac{1}{\eta_t} (\| \mathbf{y}_{t+1} - \mathbf{y}_t \|) + \frac{R}{\eta_T}.$$

① holds due to (17). The proof is completed. Lemma 5: For any  $0 \le \gamma < 1$ , we have

$$\sum_{t=1}^{T} \frac{1}{t^{\gamma}} \le \frac{1}{1-\gamma} T^{1-\gamma}.$$

*Proof:* We will use a mathematical induction method to prove the result. Given  $0 \le \gamma < 1$ , it is trivial to verify that

$$\frac{1}{1^{\gamma}} = 1 \le \frac{1}{1 - \gamma}.$$

For an integer  $T_0$ , suppose that  $\sum_{t=1}^{T_0} (1/t^{\gamma}) \leq (1/1 - \gamma)T_0^{1-\gamma}$ . Then, we have

$$\begin{split} \sum_{t=1}^{T_0+1} \frac{1}{t^{\gamma}} &= \sum_{t=1}^{T_0} \frac{1}{t^{\gamma}} + \frac{1}{(T_0+1)^{\gamma}} \\ &\leq \frac{1}{1-\gamma} T_0^{1-\gamma} + \frac{1}{(T_0+1)^{\gamma}} \\ &= \frac{1}{1-\gamma} (T_0+1)^{1-\gamma} \left( \left( \frac{T_0}{T_0+1} \right)^{1-\gamma} + \frac{1-\gamma}{T_0+1} \right) \\ \stackrel{(\text{I})}{\leq} \frac{1}{1-\gamma} (T_0+1)^{1-\gamma} \left( 1 - \frac{1-\gamma}{T_0+1} - \frac{\gamma (1-\gamma)}{2(T_0+1)^2} + \frac{1-\gamma}{T_0+1} \right) \\ &\leq \frac{1}{1-\gamma} (T_0+1)^{1-\gamma} . \end{split}$$

① holds due to the Tylor expansion, that is,

$$\left(\frac{T_0}{T_0+1}\right)^{1-\gamma} = \left(1 - \frac{1}{T_0+1}\right)^{1-\gamma}$$
  
$$\leq 1 + (1-\gamma)\left(-\frac{1}{T_0+1}\right)$$
  
$$+ \frac{(1-\gamma)(-\gamma)}{2!} \frac{1}{(T_0+1)^2}.$$

It finally completes the proof.

## REFERENCES

- M. Zinkevich, "Online convex programming and generalized infinitesimal gradient ascent," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2003, pp. 928–935.
- [2] S. Shalev-Shwartz, "Online learning and online convex optimization," *Found. Trends Mach. Learn.*, vol. 4, no. 2, pp. 107–194, 2011.
- [3] E. Hazan, "Introduction to online convex optimization," Found. Trends Optim., vol. 2, nos. 3–4, pp. 157–325, 2016.
- [4] M. Mohri and S. Yang, "Accelerating online convex optimization via adaptive prediction," in *Proc. 19th Int. Conf. Artif. Intell. Statist.* (*AISTATS*), A. Gretton and C. C. Robert, Eds., vol. 51, May 2016, pp. 848–856.
- [5] L. Zhang, T. Yang, R. jin, and Z.-H. Zhou, "Dynamic regret of strongly adaptive methods," in *Proc. 35th Int. Conf. Mach. Learn. (ICML)*, Jul. 2018, pp. 5882–5891.

IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS

- [6] K.-S. Jun, F. Orabona, S. Wright, and R. Willett, "Improved strongly adaptive online learning using coin betting," in *Proc. 20th Int. Conf. Artif. Intell. Statist. (AISTATS)*, A. Singh and J. Zhu, Eds., vol. 54, Apr. 2017, pp. 943–951.
- [7] P. Jain, P. Kothari, and A. Thakurta, "Differentially private online learning," in *Proc. Int. Conf. Learn. Theory (COLT)*, vol. 23, 2012, pp. 1–34.
- [8] J.-Y. Wang, Y.-J. Zhou, D.-Q. Li, J.-G. Lv, and Q. Dong, "Fast projection-free algorithm for distributed online learning in networks," in *Proc. IEEE 18th Int. Conf. Commun. Technol. (ICCT)*, D. Precup and Y. W. Teh, Eds. Sydney, NSW, Australia: International Convention Centre, Oct. 2018, pp. 4054–4062.
- [9] S. Chaudhuri and A. T. Tewari, "Online learning to rank with feedback at the top," in *Proc. 19th Int. Conf. Artif. Intell. Statist.* (*AISTATS*), A. Gretton and C. C. Robert, Eds., vol. 51, May 2016, pp. 277–285.
- [10] C. Liu, T. Jin, S. C. H. Hoi, P. Zhao, and J. Sun, "Collaborative topic regression for online recommender systems: An online and Bayesian approach," *Mach. Learn.*, vol. 106, no. 5, pp. 651–670, May 2017.
- [11] B. Awerbuch and T. P. Hayes, "Online collaborative filtering with nearly optimal dynamic regret," in *Proc. 19th Annu. ACM Symp. Parallel Algorithms Architectures (SPAA)*, 2007, pp. 315–319.
- [12] V. Nair and J. J. Clark, "An unsupervised, online learning framework for moving object detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2004, pp. 317–325.
- [13] A. Rakhlin, K. Sridharan, and A. Tewari, "Online learning: Stochastic, constrained, and smoothed adversaries," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2011, pp. 1764–1772.
- [14] S. Liu, J. Chen, P.-Y. Chen, and A. Hero, "Zeroth-order online alternating direction method of multipliers: Convergence analysis and applications," in *Proc. 21st Int. Conf. Artif. Intell. Statist. (AISTATS)*, A. Storkey and F. Perez-Cruz, Eds., vol. 84, Apr. 2018, pp. 288–297.
- [15] X. Gao et al., "Sparse online learning of image similarity," ACM Trans. Intell. Syst. Technol., vol. 8, no. 5, pp. 64:1–64:22, Aug. 2017.
- [16] J. Lu, S. C. Hoi, J. Wang, P. Zhao, and Z.-Y. Liu, "Large scale online kernel learning," J. Mach. Learn. Res., vol. 17, no. 47, pp. 1–43, 2016.
- [17] Y. Shen, T. Chen, and G. Giannakis, "Online ensemble multi-kernel learning adaptive to non-stationary and adversarial environments," in *Proc. 21st Int. Conf. Artif. Intell. Statist. (AISTATS)*, A. Storkey and F. Perez-Cruz, Eds., vol. 84, Apr. 2018, pp. 2037–2046.
- [18] A. Flaxman, A. T. Kalai, and H. B. McMahan, "Online convex optimization in the bandit setting: Gradient descent without a gradient," in *Proc. ACM/SIAM Symp. Discrete Algorithms (SODA)*, 2005, pp. 1–12.
- [19] R. Arora, O. Dekel, and A. Tewari, "Online bandit learning against an adaptive adversary: From regret to policy regret," in *Proc. 29th Int. Conf. Mach. Learn. (ICML)*, 2012.
- [20] J. Kwon and V. Perchet, "Online learning and blackwell approachability with partial monitoring: Optimal convergence rates," in *Proc. 20th Int. Conf. Artif. Intell. Statist. (AISTATS)*, A. Singh and J. Zhu, Eds., vol. 54, Apr. 2017, pp. 604–613.
- [21] T. Kocák, G. Neu, and M. Valko, "Online learning with noisy side observations," in *Proc. 19th Int. Conf. Artif. Intell. Statist. (AISTATS)*, A. Gretton and C. C. Robert, Eds., vol. 51, May 2016, pp. 1186–1194.
- [22] E. C. Hall and R. Willett, "Dynamical Models and tracking regret in online convex programming," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2013, pp. 579–587.
- [23] E. C. Hall and R. M. Willett, "Online convex optimization in dynamic environments," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 4, pp. 647–662, Jun. 2015.
- [24] A. Mokhtari, S. Shahrampour, A. Jadbabaie, and A. Ribeiro, "Online optimization in dynamic environments: Improved regret rates for strongly convex problems," in *Proc. IEEE 55th Conf. Decis. Control* (CDC), Dec. 2016, pp. 7195–7201.
- [25] T. Yang, L. Zhang, R. Jin, and J. Yi, "Tracking slowly moving clairvoyant: Optimal dynamic regret of online learning with true and noisy gradient," in *Proc. 34th Int. Conf. Mach. Learn. (ICML)*, 2016, pp. 449–457.
- [26] L. Zhang, T. Yang, J. Yi, R. Jin, and Z.-H. Zhou, "Improved dynamic regret for non-degenerate functions," in *Proc. Neural Inf. Process. Syst.* (*NIPS*), 2017, pp. 1–25.

- [27] S. Shahrampour and A. Jadbabaie, "Distributed online optimization in dynamic environments using mirror descent," *IEEE Trans. Autom. Control*, vol. 63, no. 3, pp. 714–725, Mar. 2018.
- [28] A. Jadbabaie, A. Rakhlin, S. Shahrampour, and K. Sridharan, "Online optimization: Competing with dynamic comparators," in *Proc. Int. Conf. Artif. Intell. Statist. (AISTATS)*, 2015, pp. 398–406.
- [29] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," J. Mach. Learn. Res., vol. 12, pp. 2121–2159, Feb. 2011.
- [30] A. S. Bedi, P. Sarma, and K. Rajawat, "Tracking moving agents via inexact online gradient descent algorithm," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 202–217, Feb. 2018.
- [31] X. Gao, X. Li, and S. Zhang, "Online learning with non-convex losses and non-stationary regret," in *Proc. 21st Int. Conf. Artif. Intell. Statist.* (*AISTATS*), A. Storkey and F. Perez-Cruz, Eds., vol. 84, Apr. 2018, pp. 235–243.
- [32] R. Jenatton, J. Huang, and C. Archambeau, "Adaptive algorithms for online convex optimization with long-term constraints," in *Proc. The* 33rd Int. Conf. Mach. Learn. (ICML), vol. 48, Jun. 2016, pp. 402–411.
- [33] C. Zhu and H. Xu, "Online gradient descent in function space," 2015, arXiv:1512.02394. [Online]. Available: https://arxiv.org/abs/1512.02394
- [34] O. Besbes, Y. Gur, and A. J. Zeevi, "Non-stationary stochastic optimization," Oper. Res., vol. 63, no. 5, pp. 1227–1244, 2015.
- [35] C. K. Chiang *et al.*, "Online optimization with gradual variations," *J. Mach. Learn. Res.*, vol. 23, pp. 1–20, Jun. 2012.
- [36] T. Chen, Q. Ling, and G. B. Giannakis, "An online convex optimization approach to proactive network resource allocation," *IEEE Trans. Signal Process.*, vol. 65, no. 24, pp. 6350–6364, Dec. 2017.
- [37] R. Jenatton, J. Huang, D. Csiba, and C. Archambeau, "Online optimization and regret guarantees for non-additive long-term constraints," 2016, arXiv:1602.05394. [Online]. Available: https://arxiv.org/abs/1602.05394
- [38] A. György and C. Szepesvári, "Shifting regret, mirror descent, and matrices," in *Proc. 33rd Int. Conf. Mach. Learn. (ICML)*, 2016, pp. 2943–2951.
- [39] L. Zhang, S. Lu, and Z.-H. Zhou, "Adaptive online learning in dynamic environments," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2018, pp. 1323–1333.
- [40] J. Abernethy, P. L. Bartlett, A. Rakhlin, and A. Tewari, "Optimal strategies and minimax lower bounds for online convex games," in *Proc. Annu. Conf. Learn. Theory (COLT)*, 2008, pp. 1–19.
- [41] F. Orabona, "A modern introduction to online learning," 2019, arXiv:1912.13213. [Online]. Available: https://arxiv.org/abs/1912.13213
- [42] M. Herbster and M. K. Warmuth, "Tracking the best expert," *Mach. Learn.*, vol. 32, no. 2, pp. 151–178, 1998.
- [43] A. György, T. Linder, and G. Lugosi, "Tracking the best of many experts," in *Proc. Conf. Learn. Theory (COLT)*, 2005, pp. 204–216.
- [44] A. Gyorgy, T. Linder, and G. Lugosi, "Efficient tracking of large classes of experts," *IEEE Trans. Inf. Theory*, vol. 58, no. 11, pp. 6709–6725, Nov. 2012.
- [45] J. Mourtada and O.-A. Maillard, "Efficient tracking of a growing number of experts," 2017, arXiv:1708.09811. [Online]. Available: https://arxiv.org/abs/1708.09811
- [46] D. Adamskiy, W. M. Koolen, A. Chernov, and V. Vovk, "A closer look at adaptive regret," J. Mach. Learn. Res., vol. 17, no. 23, pp. 1–21, 2016.
- [47] C.-Y. Wei, Y.-T. Hong, and C.-J. Lu, "Tracking the best expert in nonstationary stochastic environments," in *Proc. Adv. Neural Inf. Process. Syst.*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, Eds., 2016, pp. 3972–3980.
- [48] N. Cesa-Bianchi, P. Gaillard, G. Lugosi, and G. Stoltz, "Mirror descent meets fixed share (and feels no regret)," in *Proc. NIPS*, 2012, p. 471.
- [49] M. Mohri and S. Yang, "Competing with automata-based expert sequences," in *Proc. 21st Int. Conf. Artif. Intell. Statist.*, A. Storkey and F. Perez-Cruz, Eds., vol. 84, Apr. 2018, pp. 1732–1740.
- [50] A. Daniely, A. Gonen, and S. Shalev-Shwartz, "Strongly adaptive online learning," in *Proc. 32nd Int. Conf. Mach. Learn. (ICML)*, 2015, pp. 1405–1411.
- [51] H. Luo and R. E. Schapire, "Achieving all with no parameters: Adanormalhedge," in *Proc. Conf. Learn. Theory (COLT)*, 2015, pp. 1286–1304.
- [52] I. Katakis, G. Tsoumakas, and I. Vlahavas, "Tracking recurring contexts using ensemble classifiers: An application to email filtering," *Knowl. Inf. Syst.*, vol. 22, no. 3, pp. 371–391, Mar. 2010.