2016 IEEE 55th Conference on Decision and Control (CDC)
ARIA Resort & Casino
December 12-14, 2016, Las Vegas, USA

Online Optimization in Dynamic Environments: Improved Regret Rates for Strongly Convex Problems

Aryan Mokhtari, Shahin Shahrampour, Ali Jadbabaie, and Alejandro Ribeiro

Abstract— **In this paper, we address tracking of a time-varying parameter with unknown dynamics. We formalize the problem as an instance of online optimization in a dynamic setting. Using online gradient descent, we propose a method that sequentially predicts the value of the parameter and in turn suffers a loss. The objective is to minimize the accumulation of losses over the time horizon, a notion that is termed dynamic regret. While existing methods focus on convex loss functions, we consider strongly convex functions so as to provide better guarantees of performance. We derive a regret bound that captures the path-length of the time-varying parameter, defined in terms of the distance between its consecutive values. In other words, the bound represents the natural connection of tracking quality to the rate of change of the parameter. We provide numerical experiments to complement our theoretical findings.**

I. Introduction

Convex programming is a mature discipline that has been a subject of interest among scientists for several decades [1]–[3]. The central problem of convex programming involves minimization of a convex cost function over a convex feasible set. Traditional optimization has focused on the case that the cost function is time-invariant. However, in wide range of applications, the cost function *i*) varies over time, and *ii*) there is no prior information about the dynamics of the cost function. It is therefore important to develop *online* convex optimization techniques, which are adapted for non-stationary environments. The problem is ubiquitous in various domains such as machine learning, control theory, industrial engineering, and operations research.

Online optimization (learning) has been extensively studied in the literature of machine learning [4], [5], proving to be a powerful tool to model sequential decisions. The problem can be viewed as a game between a learner and an adversary. The learner (algorithm) sequentially selects actions, and the adversary reveals the corresponding convex losses to the learner. The term *online* captures the fact that the learner receives a streaming data sequence, and it processes that adaptively. The popular performance metric for online algorithms is called *regret*. Regret often measures the performance of algorithm versus a static benchmark [4]–[7]. For instance, the benchmark could be the optimal point of the temporal average of losses, had the learner known all the losses in advance. In a broad sense, when the benchmark is a fixed sequence, the regret is called *static*. Furthermore, improved regret bounds are derived for the case that the losses are *strongly* convex [8].

Recent works on online learning have investigated an important direction, which involves the notion of *dynamic* regret [6], [9]–[11]. The *dynamic* regret can be recognized in the form of the cumulative difference between the instantaneous loss and the minimum loss. Previous works on dynamic setting investigated convex loss functions. Motivated by the fact that in the static setting curvature gives advantage to the algorithm [8], we aim to demonstrate that *strong* convexity of losses yields an improved rate in dynamic setting.

Therefore, we consider the online optimization problem with strongly convex losses in dynamic setting, where the benchmark sequence varies without following any particular dynamics. We track the sequence using online gradient descent and prove that the dynamic regret can be bounded in terms of the *path-length* of the sequence. The path-length is defined in terms of the distance between consecutive values of the sequence. Interestingly, our result exhibits a smooth interpolation between the static and dynamic setting. In other words, the bound directly connects the tracking quality to the rate of change of the sequence. We further provide numerical experiments that verify our theoretical result.

Of particular relevance to our setup is Kalman filtering [12]. In the original Kalman filter, there are strong assumptions on the dynamical model, such as linear state-space and Gaussian noise model. However, we depart from the classical setting by assuming no particular dynamics for the states, and observing them only through the gradients of loss functions. Instead, we provide a worst-case guarantee which captures the trajectory of the states. We remark that the notion of dynamic regret is also related to adaptive, shifting, and tracking regret (see e.g. [13]–[18]) in the sense of including dynamics in the problem. However, each case represents a different notion of regret.

Proofs of results in this paper are available in [19].

II. Problem Formulation

We consider an online optimization problem where at each step t, a learner selects an action $\mathbf{x}_t \in \mathcal{X}$ and an adversary chooses a loss function $f_t : \mathcal{X} \to \mathbb{R}$. The loss associated with the action $\mathbf{x}_t$ and function f_t is given by $f_t(\mathbf{x}_t)$. Once the action is chosen, the algorithm (learner) receives the gradient $\nabla f_t(\mathbf{x}_t)$ of the loss at point $\mathbf{x}_t$.

In *static* online learning, we measure the performance of the algorithm with respect to a fixed reference $\mathbf{x} \in \mathcal{X}$ in

Work in this paper is supported by ARO W911NF-10-1-0388, NSF CAREER CCF-0952867, and ONR N00014-12-1-0997. Aryan Mokhtari, Alejandro Ribeiro, and Ali Jadbabaie are with the Department of Electrical and Systems Engineering at the University of Pennsylvania, Philadelphia, PA 19104 USA. (email: aryanm, aribeiro, jadbabai@seas.upenn.edu). Shahin Shahrampour is with the Department of Electrical Engineering at Harvard University, Cambridge, MA 02138 USA. (e-mail: shahin@seas.harvard.edu).

TABLE I: Summary of related works on dynamic online learning

Reference	Regret notion	Loss function	Regret rate
[6]	$\sum_{t=1}^{T} f_t(\mathbf{x}_t) - f_t(\mathbf{u}_t)$	Convex	$\mathcal{O}\left(\sqrt{T}(1 + C_T(\mathbf{u}_1, \ldots, \mathbf{u}_T))\right)$
[10]	$\sum_{t=1}^{T} f_t(\mathbf{x}_t) - f_t(\mathbf{u}_t)$	Convex	$\mathcal{O}\left(\sqrt{T}(1 + C'_T(\mathbf{u}_1, \ldots, \mathbf{u}_T))\right)$
[9]	$\sum_{t=1}^{T} \mathbb{E}\left[f_t(\mathbf{x}_t)\right] - f_t(\mathbf{x}_t^*)$	Convex	$\mathcal{O}\left(T^{2/3}(1 + V_T)^{1/3}\right)$
[9]	$\sum_{t=1}^{T} \mathbb{E}\left[f_t(\mathbf{x}_t)\right] - f_t(\mathbf{x}_t^*)$	Strongly convex	$\mathcal{O}\left(\sqrt{T(1 + V_T)}\right)$
[11]	$\sum_{t=1}^{T} f_t(\mathbf{x}_t) - f_t(\mathbf{x}_t^*)$	Convex	$\mathcal{O}\left(\sqrt{D_T + 1} + \min\left\{\sqrt{(D_T + 1)C_T}, [(D_T + 1)V_T T]^{1/3}\right\}\right)$
This work	$\sum_{t=1}^{T} f_t(\mathbf{x}_t) - f_t(\mathbf{x}_t^*)$	Strongly convex	$\mathcal{O}\left(1 + C_T\right)$

hindsight. In other words, we aim to minimize a *static regret* defined as

$$\mathbf{Reg}_T^s(\mathbf{x}) := \sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{x}), \tag{1}$$

and a particularly interesting value for $\mathbf{x}$ is $\mathbf{x}^* := \operatorname{argmin}_{\mathbf{x}' \in \mathcal{X}} \sum_{t=1}^{T} f_t(\mathbf{x}')$, i.e., the minimizer of the aggregate loss $\sum_{t=1}^{T} f_t$. A successful algorithm generates a set of actions $\{\mathbf{x}_t\}_{t=1}^{T}$ that yields to a sub-linear regret. Though appealing in various applications, static regret does not always serve as a comprehensive performance metric. For instance, static regret can be used in the context of static parameter estimation. However, when the parameter varies over time, we need to bring forward a new notion of regret.

In this work, we are interested to evaluate the algorithm with respect to a more stringent benchmark, which is the sequence of instantaneous minimizers. In particular, let $\mathbf{x}_t^* := \operatorname{argmin}_{\mathbf{x} \in \mathcal{X}} f_t(\mathbf{x})$ be the minimizer of the loss f_t associated with time t. Then, *dynamic* regret is defined as

$$\mathbf{Reg}_T^d(\mathbf{x}_1^*, \ldots, \mathbf{x}_T^*) := \sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{x}_t^*). \tag{2}$$

The dynamic regret in (2) captures how well the action $\mathbf{x}_t$ matches the optimal action $\mathbf{x}_t^*$ for each time t. It is well-known that in the worst-case, it is not possible to achieve a sub-linear dynamic regret, because drastic fluctuations in the minimum points can make the problem intractable. In this work, we would like to present a regret bound that maps the hardness of problem to variation intensity. We introduce a few complexity measures relevant to this context in the following section.

A. Measures of variation and bounds on dynamic regret

There are three common complexity measures to capture variations in the choices of the adversary. The first measure is the variation in losses V_T which is characterized by

$$V_T := \sum_{t=1}^{T} \sup_{\mathbf{x} \in \mathcal{X}} \left| f_t(\mathbf{x}) - f_{t-1}(\mathbf{x}) \right|. \tag{3}$$

The variation in losses V_T accumulates the maximum variation between the two consecutive functions f_t and f_{t-1} for any feasible point $\mathbf{x} \in \mathcal{X}$.

The second measure of interest in dynamic settings is the variation in gradients D_T which is measured by

$$D_T := \sum_{t=1}^{T} \left\| \nabla f_t(\mathbf{x}_t) - M_t \right\|^2, \tag{4}$$

where M_t is a causally predictable sequence available to the algorithm prior to time t [20], [21]. A simple choice is to select the previous gradient $M_t = \nabla f_{t-1}(\mathbf{x}_{t-1})$ [22], but M_t represents any predicted value for the next objective function gradient ∇f_t.

The third common measure to capture dynamics is the variation in the sequence of reference points $\mathbf{u}_1, \ldots, \mathbf{u}_T$. This variation is defined as the accumulation of the norm of the difference between subsequent reference points

$$C_T(\mathbf{u}_1, \ldots, \mathbf{u}_T) := \sum_{t=2}^{T} \|\mathbf{u}_t - \mathbf{u}_{t-1}\|, \tag{5}$$

The measure in (5) accumulates variations between two arbitrary consecutive reference points $\mathbf{u}_t$ and $\mathbf{u}_{t-1}$. Whenever the reference points are the optimal points in (2), i.e., whenever $\mathbf{u}_t = \mathbf{x}_t^*$, we drop the arguments in $C_T(\mathbf{x}_1^*, \ldots, \mathbf{x}_T^*)$ and simply use C_T to represent such variation. The variation C_T captures the difference between the optimal arguments of the two consecutive losses f_t and f_{t-1} over time. Another notion for the variation in reference points is defined as

$$C'_T(\mathbf{u}_1, \ldots, \mathbf{u}_T) := \sum_{t=2}^{T} \|\mathbf{u}_t - \Phi_t(\mathbf{u}_{t-1})\|. \tag{6}$$

where $\Phi_t(\mathbf{u}_{t-1})$ is the predicted reference point for step t evaluated at step $t-1$ by the learner [10]. If the prediction is $\Phi_t(\mathbf{u}_{t-1}) = \mathbf{u}_{t-1}$ we recover the measure in (5). In general, the variation $C'_T(\mathbf{u}_1, \ldots, \mathbf{u}_T)$ presents the variation of reference points with respects to a given dynamic $\Phi_t(\cdot)$.

The measures in (3)-(6) are different but largely compatible. They differ in that the comparisons are between functions in (3), gradients in (4) and a sequence of given – interesting is some sense, e.g., optimal – arguments in (5) and (6), but all of them yield qualitatively comparable verdicts. The comparisons in (4) and (6) further allow for the incorporation of a prediction if a model for the evolution of dynamics over time is available. The variation measures in (3)-(6) have been used to bound regret in different settings with results that we summarize in Table I. The work in

[6] uses online gradient descent (OGD) with a diminishing steprsize to establish a regret of order $\mathcal{O}(\sqrt{T}(1+C_T))$ when the losses are convex. In [10], the authors study the performance of mirror descent in the dynamic setting and establish a regret bound of order $\mathcal{O}(\sqrt{T}(1+C'_T))$ when the environment follows a adynamical model $\Phi_t(\cdot)$ and the loss functions are convex. The work in [9] evaluates the performance of OGD for the case when a noisy estimate of the gradient is available and an upper bound on V_T is assumed as prior knowledge. They establish regret bounds of order $\mathcal{O}(T^{2/3}(1+V_T)^{1/3})$ for convex loss functions and of order $\mathcal{O}(\sqrt{T(1+V_T)})$ for strongly convex loss functions. In [11], using optimistic mirror descent, the authors propose an adaptive algorithm which achieves a regret bound in terms of C_T, D_T, and V_T simultaneously, while they assume that the learner receives each variation measure online.

Motivated by the fact that in *static* regret problems *strong* convexity results in better regret bounds – order $\mathcal{O}(\log T)$ instead of order $\mathcal{O}(\sqrt{T})$, [8] – we study *dynamic* regret problems under strong convexity assumptions. Our contribution is to show that the regret associated with the OGD algorithm (defined in Section III, analyzed in Section IV) grows not faster than $1+C_T$,

$$\mathbf{Reg}_T^d(\mathbf{x}_1^*, \dots, \mathbf{x}_T^*) \leq \mathcal{O}\left(1+C_T\right). \quad (7)$$

This result improves the regret bound $\mathcal{O}(\sqrt{T}(1+C_T))$ for OGD when the functions f_t are convex but not necessarily strongly convex [6]. We remark that our algorithm assumes neither a prior knowledge nor an online feedback about C_T and the only available information for the learner is the loss function gradient $\nabla f_t(\mathbf{x}_t)$.

III. Online Gradient Descent

Consider the online learning problem for T iterations. At the beginning of each iteration t, the learner chooses the action $\mathbf{x}_t \in \mathcal{X}$ where $\mathcal{X}$ is a given convex set. Then, the adversary chooses a function f_t and evaluates the loss associated with the iterate $\mathbf{x}_t$ which is given by the difference $f_t(\mathbf{x}_t) - f_t(\mathbf{x}_t^*)$ where $\mathbf{x}_t^*$ is the minimizer of the function f_t over the set $\mathcal{X}$. The learner does not receive the loss $f_t(\mathbf{x}_t) - f_t(\mathbf{x}_t^*)$ associated with the action $\mathbf{x}_t$. Rather, she receives the gradient $\nabla f_t(\mathbf{x}_t)$ of the cost function f_t computed at $\mathbf{x}_t$. After receiving the gradient $\nabla f_t(\mathbf{x}_t)$, she uses this information to update the current iterate $\mathbf{x}_t$.

We consider a setting in which the learner uses the online gradient descent (OGD) method with a constant stepsize to update the iterate $\mathbf{x}_t$ using the released instantaneous gradient $\nabla f_t(\mathbf{x}_t)$. To be more precise, consider $\mathbf{x}_t$ as the sequence of actions that the learner chooses and define $\hat{\mathbf{x}}_t \in \mathcal{X}$ as a sequence of auxiliary iterates. At each iteration t, given the iterate $\mathbf{x}_t$ and the instantaneous gradient $\nabla f_t(\mathbf{x}_t)$, the learner computes the auxiliary variable $\hat{\mathbf{x}}_t$ as

$$\hat{\mathbf{x}}_t = \Pi_{\mathcal{X}}\left(\mathbf{x}_t - \frac{1}{\gamma}\nabla f_t(\mathbf{x}_t)\right), \quad (8)$$

where γ is a positive constant and $\Pi_{\mathcal{X}}$ denotes the projection onto the nearest point in the set $\mathcal{X}$, i.e, $\Pi_{\mathcal{X}}(\mathbf{y}) = \operatorname{argmin}_{\mathbf{x}\in\mathcal{X}} \|\mathbf{x}-\mathbf{y}\|$. Then, the action $\mathbf{x}_{t+1}$ is evaluated as

$$\mathbf{x}_{t+1} = \mathbf{x}_t + h(\hat{\mathbf{x}}_t - \mathbf{x}_t), \quad (9)$$

where h is chosen from the interval $(0,1]$. The online gradient descent method is summarized in Algorithm 1.

Algorithm 1 Online Gradient Descent

Require: Initial vector $\mathbf{x}_1 \in \mathcal{X}$, constants h and γ.
1: **for** $t = 1, 2, \dots, T$ **do**
2: Play $\mathbf{x}_t$
3: Observe the gradient of the current action $\nabla f_t(\mathbf{x}_t)$
4: Compute the auxiliary var.: $\hat{\mathbf{x}}_t = \Pi_{\mathcal{X}}(\mathbf{x}_t - \frac{1}{\gamma}\nabla f_t(\mathbf{x}_t))$
5: Compute the next action: $\mathbf{x}_{t+1} = \mathbf{x}_t + h(\hat{\mathbf{x}}_t - \mathbf{x}_t)$
6: **end for**

The updated action $\mathbf{x}_{t+1}$ in (9) can be written as $\mathbf{x}_{t+1} = (1-h)\mathbf{x}_t + h\hat{\mathbf{x}}_t$. Therefore, we can reinterpret the updated action $\mathbf{x}_{t+1}$ as a weighted average of the previous iterate $\mathbf{x}_t$ and the auxiliary variable $\hat{\mathbf{x}}_t$ which is evaluated using the gradient of the function f_t. It is worth mentioning that for a small choice of $h \approx 0$, $\mathbf{x}_{t+1}$ is close to the previous action $\mathbf{x}_t$, while the previous action $\mathbf{x}_t$ has less impact on $\mathbf{x}_{t+1}$ when h is close to 1.

The auxiliary variable $\hat{\mathbf{x}}_t$ is the result of applying projected gradient descent on the current iterate $\mathbf{x}_t$ as shown in (8). This update can be interpreted as minimizing a first-order approximation of the cost function f_t added to a proximal term $(\gamma/2)\|\mathbf{x}-\mathbf{x}_t\|^2$ as we show in the following proposition.

Proposition 1 *Consider the update in* (8)*. Given the iterate* $\mathbf{x}_t$*, the instantaneous gradient* $\nabla f_t(\mathbf{x}_t)$*, and the positive constant* γ*, the optimal argument of the optimization problem*

$$\tilde{\mathbf{x}}_t = \operatorname*{argmin}_{\mathbf{x}\in\mathcal{X}} \left\{\nabla f_t(\mathbf{x}_t)^T(\mathbf{x}-\mathbf{x}_t) + \frac{\gamma}{2}\|\mathbf{x}-\mathbf{x}_t\|^2\right\}, \quad (10)$$

is equal to the iterate $\hat{\mathbf{x}}_t$ *generated by* (8)*.*

The result in Proposition 1 shows that the updates in (8) and (10) are equivalent. In the implementation of OGD we use the update in (8), since the computational complexity of the update in (8) is lower than the complexity of the minimization in (10). On the other hand, the update in (10) is useful in the regret analysis of OGD that we undertake in the following section.

IV. Regret Analysis

We proceed to show that the dynamic regret $\mathbf{Reg}_T^d(\mathbf{x}_1^*, \dots, \mathbf{x}_T^*) = \mathbf{Reg}_T^d$ defined in (2) associated with the actions $\mathbf{x}_t$ generated by the online gradient descent algorithm in (9) has an upper bound on the order of the variation in the sequence of optimal arguments $C_T = \sum_{t=2}^T \|\mathbf{x}_t^* - \mathbf{x}_{t-1}^*\|$. In proving this result, we assume the following conditions are satisfied.

Assumption 1 *The functions* f_t *are strongly convex over the convex set* $\mathcal{X}$ *with constant* $\mu > 0$*, i.e.,*

$$f_t(\mathbf{x}) \geq f_t(\mathbf{y}) + \nabla f_t(\mathbf{y})^T(\mathbf{x}-\mathbf{y}) + \frac{\mu}{2}\|\mathbf{x}-\mathbf{y}\|^2, \quad (11)$$

for any $\mathbf{x}, \mathbf{y} \in \mathcal{X}$ *and* $1 \leq t \leq T$.

Assumption 2 *The gradients* ∇f_t *are Lipschitz continuous over the set* $\mathcal{X}$ *with constant* $L < \infty$*, i.e.,*

$$\|\nabla f_t(\mathbf{y}) - \nabla f_t(\mathbf{x})\| \leq L\|\mathbf{x} - \mathbf{y}\|, \tag{12}$$

for any $\mathbf{x}, \mathbf{y} \in \mathcal{X}$ *and* $1 \leq t \leq T$.

Assumption 3 *The gradient norm* $\|\nabla f_t\|$ *is bounded above by a positive constant* G *or equivalently*

$$\sup_{\mathbf{x} \in \mathcal{X}, 1 \leq t \leq T} \|\nabla f_t(\mathbf{x})\| \leq G. \tag{13}$$

According to Assumption 1, the instantaneous functions f_t are strongly convex over the convex set $\mathcal{X}$ which implies that there exists a unique minimizer $\mathbf{x}_t^*$ for the function f_t over the convex set $\mathcal{X}$. The Lipschitz continuity of the gradients ∇f_t in Assumption 2 is customary in the analysis of descent methods. Notice that we only assume for a fixed function f_t the gradients are Lipschitz continuous and we do not assume any conditions on the difference of two gradients associated with two different instantaneous functions. To be more precise, there is no condition on the norm $\|\nabla f_t(\mathbf{y}) - \nabla f_{t'}(\mathbf{x})\|$ where $t \neq t'$. The bound on the gradients norm in Assumption 3 is typical in the analysis of online algorithms for constrained optimization.

Our main result on the regret bound of OGD in dynamic settings is derived from the following proposition that bounds the difference $\|\mathbf{x}_{t+1} - \mathbf{x}_t^*\|$ in terms of the distance $\|\mathbf{x}_t - \mathbf{x}_t^*\|$.

Proposition 2 *Consider the online gradient descent method (OGD) defined by* (8) *and* (9) *or the equivalent* (10)*. Recall the definition of* $\mathbf{x}_t^*$ *as the unique minimizer of the function* f_t *over the convex set* $\mathcal{X}$*. If Assumptions 1 and 2 hold and the stepsize parameter* γ *in* (8) *is chosen such that* $\gamma \geq L$*, then the sequence of actions* $\mathbf{x}_t$ *generated by OGD satisfies*

$$\|\mathbf{x}_{t+1} - \mathbf{x}_t^*\| \leq \rho \, \|\mathbf{x}_t - \mathbf{x}_t^*\|, \tag{14}$$

where $0 \leq \rho := (1 - h\mu/\gamma)^{1/2} < 1$ *is a non-negative constant strictly smaller than* 1.

The result in Proposition 2 shows that the distance between the action $\mathbf{x}_{t+1}$ and the optimal argument $\mathbf{x}_t^*$ is strictly smaller than the difference between the previous action $\mathbf{x}_t$ and the optimal argument $\mathbf{x}_t^*$ at step t. The inequality in (14) implies that if the optimal arguments of the functions f_t and f_{t+1} which are $\mathbf{x}_t^*$ and $\mathbf{x}_{t+1}^*$, respectively, are not far away from each other the iterates $\mathbf{x}_t$ can track the optimal solution sequence $\mathbf{x}_t^*$. Notice that if in the left hand side of (14) instead of $\mathbf{x}_t^*$ we had the optimal argument $\mathbf{x}_{t+1}^*$ at step $t+1$, then we could show that the sequence of actions $\mathbf{x}_t$ generated by OGD asymptotically converges to the sequence of optimal arguments $\mathbf{x}_t^*$. Thus, the performance of OGD depends on the rate that the sequence of optimal arguments changes. This conclusion is formalized in the following Theorem.

Theorem 1 *Consider the online gradient descent method (OGD) defined by* (8) *and* (9) *or the equivalent* (10)*. Suppose that* h *is chosen from the interval* $(0, 1]$ *and the constant* γ *satisfies the condition* $\gamma \geq L$*, where* L *is the gradients Lipschitz continuity constant. If Assumptions 1 and 2 hold, then the sequence of actions* $\mathbf{x}_t$ *generated by OGD satisfies*

$$\sum_{t=1}^{T} \|\mathbf{x}_t - \mathbf{x}_t^*\| \leq K_1 \sum_{t=2}^{T} \|\mathbf{x}_t^* - \mathbf{x}_{t-1}^*\| + K_2, \tag{15}$$

where the constants K_1 *and* K_2 *are explicitly given by*

$$K_1 := \frac{\|\mathbf{x}_1 - \mathbf{x}_1^*\| - \rho\|\mathbf{x}_T - \mathbf{x}_T^*\|}{(1-\rho)}, \quad K_2 := \frac{1}{(1-\rho)}. \tag{16}$$

From Theorem 1, we obtain an upper bound for the aggregate variable error $\sum_{t=1}^{T} \|\mathbf{x}_t - \mathbf{x}_t^*\|$ in terms of the aggregate variation in the optimal arguments $C_T = \sum_{t=2}^{T} \|\mathbf{x}_t^* - \mathbf{x}_{t-1}^*\|$. This result matches the intuition that for the scenarios that the sequence of the optimal arguments $\{\mathbf{x}_t^*\}_{t=1}^{T}$ is not varying fast, the sequence of actions generated by the online gradient descent method can achieve a sublinear regret bound. In particular, when the optimal arguments are all equal to each other, i.e., when $\mathbf{x}_1^* = \cdots = \mathbf{x}_T^*$, the aggregate error $\sum_{t=1}^{T} \|\mathbf{x}_t - \mathbf{x}_t^*\|$ is bounded above by the constant K_2 which is independent of T. We use the bounded gradients assumption in (13) to translate the result in (15) into an upper bound for the dynamic regret $\mathbf{Reg}_T^d$ defined in (2).

Corollary 1 *Adopt the same definitions and hypothesis of Theorem 1 and furhter assume that the gradient norms* $\|\nabla f_t(\mathbf{x})\|$ *are upper bounded by the constant* G *for all* $\mathbf{x} \in \mathcal{X}$ *as in Assumption 3. Then, the dynamic regret* $\mathbf{Reg}_T^d$ *for the sequence of actions* $\mathbf{x}_t$ *generated by OGD is bounded above by*

$$\mathbf{Reg}_T^d \leq G K_1 \sum_{t=2}^{T} \|\mathbf{x}_t^* - \mathbf{x}_{t-1}^*\| + G K_2. \tag{17}$$

The result in Corollary 1 states that under the conditions that the functions f_t are strongly convex and their gradients are bounded and Lipschitz continuous, the dynamic regret $\mathbf{Reg}_T^d$ associated with the online gradient descent method satisfies the order bound that we previewed in (7). As already mentioned, this bound improves the OGD rate $\mathcal{O}(\sqrt{T}(1 + C_T))$ when the functions f_t are convex but not ncessarily strongly convex and the stepsize is diminishing [6].

Some interesting conclusions can be derived if we consider specific rates of variability:

Constant functions. If the functions are constant, i.e., if $f_t = f$ for all times, we have $C_T = 0$ and it follows that the regret grows at a rate $\mathcal{O}(1)$. This means that $\mathbf{x}_t$ converges to $\mathbf{x}^*$ and we recover a convergence proof for gradient descent.

Linearly decreasing variability. If the difference between consecutive arguments decreases as $1/t$, we have that $C_T = \mathcal{O}(\log T)$ and that the regret then grows at a logarithmic rate as well. Since this implies that the normalized regret grows not faster than $\mathbf{Reg}_T^d(\mathbf{x}_1^*, \ldots, \mathbf{x}_T^*)/T \leq \mathcal{O}(\log T/T)$ we must have that $\mathbf{x}_t$ converges to $\mathbf{x}_t^*$.

Decreasing variability. If the variability decreases as $1/t^{\alpha}$ with $\alpha \in (0,1)$, the regret is of order $\mathcal{O}(1+C_T) = \mathcal{O}(1+T^{1-\alpha})$. As before, this must imply that $\mathbf{x}_t$ converges to $\mathbf{x}_t^*$.

Constant variability. If the variability between functions stays constant, say $\|\mathbf{u}_t^* - \mathbf{u}_{t-1}^*\| \leq C$ for all T, the regret is of order $\mathcal{O}(1+C_T) = \mathcal{O}(1+CT)$. The normalized regret is then of order $\mathbf{Reg}_T^d(\mathbf{x}_1^*, \ldots, \mathbf{x}_T^*)/T \leq \mathcal{O}(C)$. This means that we have a steady state tracking error, where the tracking error depends on how different adjacent functions are.

V. Numerical Experiments

We numerically study the performance of OGD in solving a sequence of quadratic programming problems. Consider the decision variable $\mathbf{x} = [x_1; x_2] \in \mathbb{R}^2$ and the quadratic function f_t at time t which is defined as

$$f_t(\mathbf{x}) = f_t(x_1, x_2) = \rho\|x_1 - a_t\|^2 + \|x_2 - b_t\|^2 + c_t, \quad (18)$$

where a_t, b_t, and c_t are time-variant scalars and $\rho > 0$ is a positive constant. The coefficient ρ controls the condition number of the objective function f_t. In particular for $\rho > 1$, the problem condition number is equal to ρ. The convex set $\mathcal{X}$ is defined as $x_1^2 + x_2^2 = r^2$ which is the circle with center $[0; 0]$ and radius r. The radius r is chosen such that the optimal argument of the function f_t over $\mathbb{R}^2$, which is $[a_t, b_t]$, is not included in the set $\mathcal{X}$. This way we ensure that the constraint $\mathbf{x} \in \mathcal{X}$ is active at the optimal solution.

In our experiments we pick $\rho = 100$ to have a quadratic optimization problem with large condition number 100. Note that if we choose $\rho = 1$, the condition number of the function f_t is 1 and OGD can minimize the cost f_t in a couple of iterations. The constant γ in OGD is set as $\gamma = 2\rho$ in all experiments, since the Lipschitz continuity constant of gradients is $L = 2\rho$. Moreover, the OGD parameter h is set as $h = 1$ which implies $\mathbf{x}_{t+1} = \hat{\mathbf{x}}_t$.

To characterize the instantaneous performance of OGD we define $f_t(\mathbf{x}_t) - f_t(\mathbf{x}_t^*)$ as the instantaneous objective function error at time t. Further, we define $\mathbf{Reg}_t^d := \sum_{s=1}^t f_s(\mathbf{x}_s) - f_s(\mathbf{x}_s^*)$ as the dynamic regret up to step t. Likewise, we define $C_t := \sum_{s=2}^t \|\mathbf{x}_s^* - \mathbf{x}_{s-1}^*\|$ as the total optimal argument variation until step t.

We consider two different cases to study the regret bound of OGD in dynamic online settings. First, we consider a switching problem that the adversary switches between two quadratic functions after a specific number of iterations. Then, we study the case that the sequence of optimal arguments $\mathbf{x}_t^*$ changes at each iteration, while the difference $\|\mathbf{x}_t^* - \mathbf{x}_{t-1}^*\|$ diminishes as time progresses.

A. Switching problem

Consider the case that the adversary chooses between two functions where each of them is of the form of the quadratic function f_t in (18). In particular, consider the case that the adversary chooses the parameters a_t, b_t, and c_t from the two sets $\mathcal{S}^{(1)} = \{a, b, c\}$ and $\mathcal{S}^{(2)} = \{a', b', c'\}$. Therefore, at each iteration the adversary chooses either $f^{(1)} = \rho\|x_1 - a\|^2 + \|x_2 - b\|^2 + c$ or $f^{(2)} = \rho\|x_1 - a'\|^2 + \|x_2 - b'\|^2 + c'$. We run OGD for a fixed number of iterations $T = 100$ and assume that the adversary switches between the functions $f^{(1)}$ and $f^{(2)}$ every τ iterations.

In our experiments we set $a = -100$, $b = 0$, $c = 30$, $a' = 100$, $b' = 20$, and $c' = -50$. The convex set $\mathcal{X}$ is defined as $x_1^2 + x_2^2 = 50^2$ which is the circle with center $[0; 0]$ and radius 50. Note that this circle does not contain the points $[a; b] = [-100; 0]$ and $[a'; b'] = [100; 20]$. The optimal argument of $f^{(1)}$ and $f^{(2)}$ over the convex set $\mathcal{X}$ are $\mathbf{x}^{(1)^*} = [-50; 0]$ and $\mathbf{x}^{(2)^*} = [49.99; 0.19]$, respectively. We set the initial iterate $\mathbf{x}_0 = [0; 40]$ which is a feasible point for the set $\mathcal{X}$. We consider three different cases that $\tau = 4$, $\tau = 8$, and $\tau = 16$. The performance of OGD for the three different choices of τ are illustrated in Figure 1.

Figure 1a demonstrates the variable variation $C_t := \sum_{s=2}^t \|\mathbf{x}_s^* - \mathbf{x}_{s-1}^*\|$ over time t. For the case $\tau = 16$, the value of C_t increases every 16 iterations and the increment is equal to the norm of the difference between the optimal arguments of $f^{(1)}$ and $f^{(2)}$ which is $\|\mathbf{x}^{(1)^*} - \mathbf{x}^{(2)^*}\| = 100$. After $T = 100$ iterations, we observe 6 jumps which implies that the total variable variation is $C_T = 600$. Likewise, for the cases that $\tau = 8$ and $\tau = 4$, we observe 12 and 24 jumps in their corresponding plots, and the aggregate variable variations are $C_T = 1200$ and $C_T = 2400$, respectively.

Figure 1b showcases the instantaneous function error $f_t(\mathbf{x}_t) - f_t(\mathbf{x}_t^*)$ versus number of iterations t for $\tau = 16$, $\tau = 8$, and $\tau = 4$. In all of the cases, the sequence of errors $f_t(\mathbf{x}_t) - f_t(\mathbf{x}_t^*)$ converges linearly to 0 until the time the adversary switches the objective function f_t. By increasing the number of times that the adversary switches between the functions $f^{(1)}$ and $f^{(2)}$, the phase of linear convergence becomes shorter and the algorithm restarts more often.

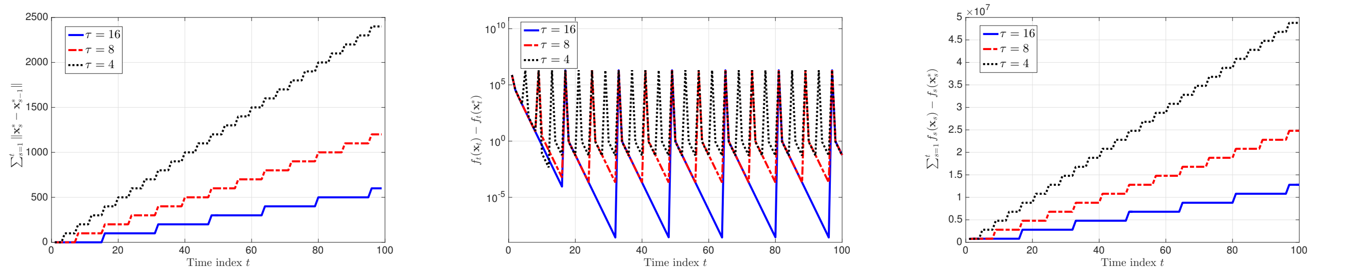
Thus, we expect to observe larger regret for the scenarios that τ is smaller. The dynamic regret $\mathbf{Reg}_t^d = \sum_{s=1}^t f_s(\mathbf{x}_s) - f_s(\mathbf{x}_s^*)$ versus the number of iterations t is shown in Figure 1c for $\tau = 16$, $\tau = 8$, and $\tau = 4$. As we expect, for the case that $\tau = 4$ the dynamic regret $\mathbf{Reg}_t^d$ grows faster. Note that the dynamic regret $\mathbf{Reg}_T^d$ of OGD for $\tau = 16$, $\tau = 8$, and $\tau = 4$ after $T = 100$ iterations are $\mathbf{Reg}_T^d = 1.28 \times 10^7$, $\mathbf{Reg}_T^d = 2.48 \times 10^7$, and $\mathbf{Reg}_T^d = 4.88 \times 10^7$, respectively.

Comparing the variation C_t and the dynamic regret $\mathbf{Reg}_T^d$ for the three cases $\tau = 16$, $\tau = 8$, and $\tau = 4$ shows that the growth patterns of C_t and $\mathbf{Reg}_T^d$ are similar. This observation is consistent with the theoretical result in Corollary 1 which indicates that the upper bound for the dynamic regret $\mathbf{Reg}_T^d$ is of order $\mathcal{O}(1 + C_T)$.

B. Diminishing variations

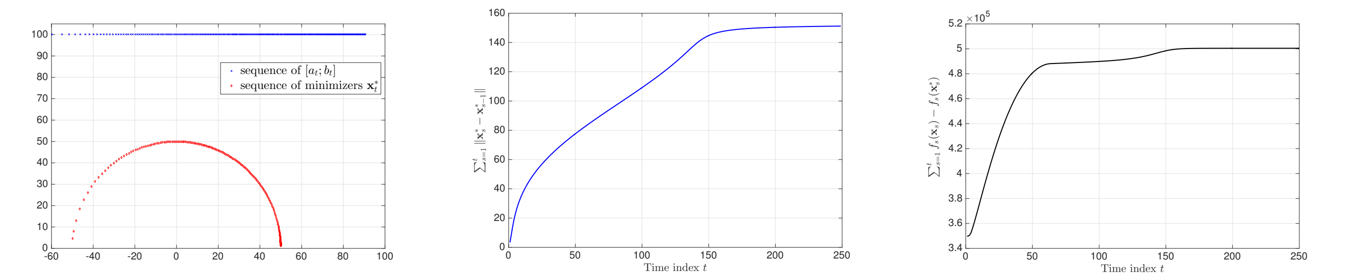
In this section we consider the case that the adversary picks a sequence of functions f_t as in (18) such that the sequence of optimizers $\mathbf{x}_t^*$ is convergent.

We use the parameters in Figure 1 except the total number of iterations which is set as $T = 250$. We set the initial values for $[a_1; b_1]$ as $[-60; 100]$. Further, we assume that b_t is time-invariant and for all steps t we have $b_t = b_1 = 100$. On the other hand, we assume that the parameter a_t



(a) (b) (c)

Fig. 1: Performance of OGD for the case that the adversary switches between two quadratic functions with different minimizers every τ iterations. The variation in the sequence of optimizers $C_t = \sum_{s=2}^{t} \|\mathbf{x}_s^* - \mathbf{x}_{s-1}^*\|$, instantaneous objective function error $f_t(\mathbf{x}_t) - f_t(\mathbf{x}_t^*)$, and dynamic regret $\mathbf{Reg}_t^d = \sum_{s=1}^{t} f_s(\mathbf{x}_s) - f_s(\mathbf{x}_s^*)$ are shown in Figures 1a, 1b, and 1c, respectively. In the case that τ is small and the adversary switches more often between the two quadratic functions, the variation C_t and the dynamic regret $\mathbf{Reg}_t^d$ grow faster. Moreover, the growth patterns of C_t and $\mathbf{Reg}_t^d$ are similar.



(a) (b) (c)

Fig. 2: Performance of OGD for the case that the adversary chooses a sequence of quadratic functions f_t where the sequence of optimal arguments $\mathbf{x}_t^*$ is convergent. Figure 2a illustrates the paths for the function parameters $[a_t; b_t]$ and the optimal arguments $\mathbf{x}_t^*$ over the set $\mathcal{X}$. The variation in the sequence of optimizers $C_t := \sum_{s=2}^{t} \|\mathbf{x}_s^* - \mathbf{x}_{s-1}^*\|$ and the dynamic regret $\mathbf{Reg}_t^d = \sum_{s=1}^{t} f_s(\mathbf{x}_s) - f_s(\mathbf{x}_s^*)$ are shown in Figures 2b and 2c, respectively. The sequence of optimal arguments is convergent in a way that the variation in the sequence of optimizers $C_t := \sum_{s=2}^{t} \|\mathbf{x}_s^* - \mathbf{x}_{s-1}^*\|$ is summable. Likewise, the dynamic regret $\mathbf{Reg}_T^d$ associated with the OGD method does not grow as the total number of iterations T increases.

changes as time passes and it satisfies the recursive formula $a_{t+1} = a_t + 5\sqrt{1/t}$. The sequence of parameters $[a_t, b_t]$, which are the optimal arguments of the function f_t over $\mathbb{R}^2$, is illustrated in Figure 2a. Moreover, the set of optimal arguments of f_t over the set $\mathcal{X}$, which are indicated by $\mathbf{x}_t^*$, are also demonstrated in Figure 2a. This plot shows that as time progresses and the difference between the functions f_t and f_{t-1} becomes less significant, the difference between the optimal arguments $\mathbf{x}_t^*$ and $\mathbf{x}_{t-1}^*$ diminishes. To formally study the variation in the sequence of optimal arguments $\mathbf{x}_t^*$, we demonstrate the variable variation $C_t := \sum_{s=2}^{t} \|\mathbf{x}_s^* - \mathbf{x}_{s-1}^*\|$ in terms of number of iterations t in Figure 2b. As we expect, the variation C_t converges as time progresses, since the difference $\|\mathbf{x}_t^* - \mathbf{x}_{t-1}^*\|$ is diminishing.

Figure 2c illustrates the dynamic regret $\mathbf{Reg}_t^d = \sum_{s=1}^{t} f_s(\mathbf{x}_s) - f_s(\mathbf{x}_s^*)$ of OGD in terms of number of iterations t. We observe that the dynamic regret of OGD follows the pattern of the variation C_t and converges to a constant value as time progresses. Thus, the dynamic regret $\mathbf{Reg}_T^d$ does not grow with the number of iterations T. This observation matches the theoretical result in Corollary 1 that the dynamic regret $\mathbf{Reg}_T^d$ converges to a constant value when the variation in optimal arguments $C_T := \sum_{s=2}^{T} \|\mathbf{x}_s^* - \mathbf{x}_{s-1}^*\|$ does not grow by the number of iterations T.

VI. Conclusions

This paper studies the performance of the online gradient descent (OGD) algorithm in online dynamic settings. We established an upper bound for the dynamic regret of OGD in terms of the variation in the sequence of optimal arguments $\mathbf{x}_t^*$ defined by $C_T = \sum_{t=2}^{T} \|\mathbf{x}_t^* - \mathbf{x}_{t-1}^*\|$. We showed that if the functions f_t chosen by the adversary are strongly convex, the online gradient descent method with a proper constant stepsize has a regret of order $\mathcal{O}(1 + C_T)$. This result indicates that the dynamic regret bound of OGD for strongly convex functions is significantly smaller than the regret bound of order $\mathcal{O}(\sqrt{T}(1 + C_T))$ for convex settings. Numerical experiments on a dynamic quadratic programming verified our theoretical result that the dynamic regret of OGD has an upper bound of order $\mathcal{O}(1 + C_T)$.

REFERENCES

[1] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[2] A. S. Nemirovski and D. B. Yudin, *Problem complexity and method efficiency in optimization*. Wiley (Chichester and New York), 1983.

[3] A. Beck and M. Teboulle, "Mirror descent and nonlinear projected subgradient methods for convex optimization," *Operations Research Letters*, vol. 31, no. 3, pp. 167–175, 2003.

[4] N. Cesa-Bianchi and G. Lugosi, *Prediction, learning, and games*. Cambridge University Press, 2006.

[5] S. Shalev-Shwartz, "Online learning and online convex optimization," *Foundations and Trends in Machine Learning*, vol. 4, no. 2, pp. 107–194, 2011.

[6] M. Zinkevich, "Online convex programming and generalized infinitesimal gradient ascent," in *International Conference on Machine Learning*, 2003.

[7] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of computer and system sciences*, vol. 55, no. 1, pp. 119–139, 1997.

[8] E. Hazan, A. Agarwal, and S. Kale, "Logarithmic regret algorithms for online convex optimization," *Machine Learning*, vol. 69, no. 2-3, pp. 169–192, 2007.

[9] O. Besbes, Y. Gur, and A. Zeevi, "Non-stationary stochastic optimization," *Operations Research*, vol. 63, no. 5, pp. 1227–1244, 2015.

[10] E. Hall and R. Willett, "Online convex optimization in dynamic environments," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 4, pp. 647–662, June 2015.

[11] A. Jadbabaie, A. Rakhlin, S. Shahrampour, and K. Sridharan, "Online optimization: Competing with dynamic comparators," in *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics*, 2015, pp. 398–406.

[12] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Fluids Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[13] M. Herbster and M. K. Warmuth, "Tracking the best expert," *Machine Learning*, vol. 32, no. 2, pp. 151–178, 1998.

[14] E. Hazan and C. Seshadhri, "Adaptive algorithms for online decision problems," in *Electronic Colloquium on Computational Complexity (ECCC)*, vol. 14, no. 088, 2007.

[15] N. Cesa-Bianchi, P. Gaillard, G. Lugosi, and G. Stoltz, "A new look at shifting regret," *CoRR abs/1202.3323*, 2012.

[16] A. Daniely, A. Gonen, and S. Shalev-shwartz, "Strongly adaptive online learning," in *Proceedings of the 32nd International Conference on Machine Learning (ICML-15)*, 2015, pp. 1405–1411.

[17] H. Luo and R. E. Schapire, "Achieving all with no parameters: Adanormalhedge," in *Proceedings of The 28th Conference on Learning Theory*, 2015, pp. 1286–1304.

[18] A. Koppel, A. Simonetto, A. Mokhtari, G. Leus, and A. Ribeiro, "Target tracking with dynamic convex optimization," in *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, 2015, pp. 1210–1214.

[19] A. Mokhtari, S. Shahrampour, A. Jadbabaie, and A. Ribeiro, "Online optimization in dynamic environments: Improved regret rates for strongly convex problems," *arXiv preprint arXiv:1603.04954*, 2016.

[20] A. Rakhlin and K. Sridharan, "Online learning with predictable sequences," in *Conference on Learning Theory*, 2013, pp. 993–1019.

[21] S. Rakhlin and K. Sridharan, "Optimization, learning, and games with predictable sequences," in *Advances in Neural Information Processing Systems*, 2013, pp. 3066–3074.

[22] C.-K. Chiang, T. Yang, C.-J. Lee, M. Mahdavi, C.-J. Lu, R. Jin, and S. Zhu, "Online optimization with gradual variations." in *Conference on Learning Theory*, 2012.