Adversarial Delays in Online Strongly-Convex Optimization

Daniel Khashabi  
Department of Computer Science  
Univ. of Illinois, Urbana-Champaign  
khashab2@illinois.edu

Kent Quanrud  
Department of Computer Science  
Univ. of Illinois, Urbana-Champaign  
quanrud2@illinois.edu

Amirhossein Taghvaei  
Coordinated Sciences Laboratory  
Univ. of Illinois, Urbana-Champaign  
taghvae2@illinois.edu

Abstract

We consider the problem of strongly-convex online optimization in presence of adversarial delays [1]; in a \( T \)-iteration online game, the feedback of the player’s query at time \( t \) is arbitrarily delayed by an adversary for \( d_t \) rounds and delivered before the game ends, at iteration \( t + d_t - 1 \). Specifically for online-gradient-descent algorithm we show it has a simple regret bound of \( O \left( \sum_{t=1}^{T} \log \left( 1 + \frac{d_t}{t} \right) \right) \). This gives a clear and simple bound without resorting any distributional and limiting assumptions on the delays. We further show how this result encompasses and generalizes several of the existing known results in the literature. Specifically it matches the celebrated logarithmic regret \( O \left( \log T \right) \) [2] when there are no delays (i.e. \( d_t = 1 \)) and the regret bound of \( O \left( \tau \log T \right) \) [3] for constant delays \( d_t = \tau \).

1 Introduction

Online learning [4–6] often is modelled as a \( T \)-round two-player game, a learner and an environment [\(^\ast\)] at iteration \( t \) the learner chooses an action \( x_t \in K \), where \( K \) is a compact set, e.g the unit \( \ell_1 \)-ball in the Euclidian space or a collection of \( n \) experts. Before the decision of the next round, the learner receives a feedback \( f_t \) from the environment and we incur the loss of \( \ell_t = f_t(x_t) \). While the total cost incurred in this online game is \( \sum_{t=1}^{T} \ell_t \), the goal is to minimize the difference between cumulative online cost, and the best single offline action: \( R(T) = \sum_{t=1}^{T} f_t(x_t) - \min_{x \in K} \sum_{t=1}^{T} f_t(x) \). The aforementioned quantity \( R(T) \) is usually referred to as the regret of the learner over this \( T \)-round game and it is a key definition in the study of online learning games.

There is large body of theoretical work for quantifying upper/lower bounds of this statistic. For example for a large class of settings this problem has shown to have the regret bound of \( R(T) = \Theta \left( \sqrt{T} \right) \) \( ^{\dagger} \) examples are hedge [7], follow-the-perturbed-leader [8]. See the existing comprehensive literature reviews for a complete list [4–6].

One of the setting which happens in some real scenarios and has enjoyed much attention is when the cost functions are strongly convex. One important example is the well-studied Support Vector Machine (SVM) model which has strongly convex (but non-smooth) objective function [9]. Although

\(^\ast\) Also referred to as “adversary”.

\(^\dagger\) Assuming some regularity conditions, for example boundedness of \( K \) and the \( f_t \).
the (offline) convex optimization literature has been around for several decades \cite{10,13} there is a relatively recent momentum on studying stronger results using more structural assumptions like strong-convexity (see for example \cite{14,18}). In the context of online convex optimization Hazan et al. \cite{2} presented first $O(\log T)$ regret bound for online strongly-convex optimization; later Hazan and Kale \cite{17} showed the tightness of this bound by providing a lower bound.

In the standard online learning model the player is assumed to receive feedback before making the decision for the next iteration. Unfortunately this assumption can be unrealistic and limiting in some applications. An example is online advertising: the goal of the learner is to maximize the number of clicks on the ads served to the users. In this model the feedback is whether the user has decided to click or ignore the ad. In reality there can be a significant delay between the time an ad is being displayed to the user and the feedback received, while this shouldn’t stop the advertising system from prediction.

Another important scenario involving delay is when the problem is distributed over a network. In this scenario the asynchronicity of the processors and other non-deterministic issues about the network connections naturally introduce delays in the process. This problem has long been studied in the (offline) optimization community (e.g. the seminal work of Bertsekas and Tsitsiklis \cite{12}). Recht et al. \cite{19} go past the issue of delays by studying a special case of this problem where the feedbacks are almost orthogonal in the strong-convexity setting. Similarly Langford et al. \cite{3}, Duchi et al. \cite{20} and McMahan and Streeter \cite{21} analyzed distributed variants of adaptive gradient descent \cite{22} under distributional assumptions on the delay (e.g. bounded expected delay); more recent developments can be found in \cite{23,28}.

There are several approaches for tackling delays in online learning. A simple approach adopted by several works with small variations (e.g. \cite{29,31}) is using undelayed algorithms for the delayed problem. The approach incrementally distributes the queries and their corresponding feedback into disjoint undelayed subsequences among multiple independent (undelayed) learning algorithms. Each round is delegated to a single instance of the learner from a pool of learners; the instance will wait for the feedback of the query before rejoining the pool. If there is no available learner, a new instance will be created. It is not hard to show that overall regret of this strategy equals the sum of regrets of individual learners. If $m$ is the expected maximum number of outstanding feedbacks, then Joulani et al. obtained a regret bound on the order of $m R(T/m)$ (in expectation), where $R(T)$ is the regret of an undelayed problem. Using the approach of Hazan and Kale one can easily get $O(m \log \frac{T}{m})$ for constant-delay online learning with strongly-convex loss functions.

Quanrud and Khashabi \cite{11} introduced the notion of adversarial delays for online convex optimization (i.e. no distributional assumptions for delays) and provided a combinatorial analysis to get an upper-bound of $O\left(\sqrt{\sum_{t=1}^{T} d_t}\right)$ on regret for online-gradient-descent \cite{32}, follow-the-perturbed-leader \cite{8} and online-mirror-descent \cite{11}. Around the same time Joulani et al. \cite{33} found the same regret upper-bound for follow-the-regularized-leader by extending ideas from McMahan and Streeter \cite{21}. There are also analysis of delays in online learning with partial feedback (multi-armed bandits) which are not directly relevant to this work; we refer the interested reader to \cite{31,34}.

We draw inspiration from the recent works \cite{1,33} and analyze the online-gradient-descent under extra assumption that the loss functions are strongly-convex. As a result of this extra assumption we get a strictly sharper upper-bound $O\left(\sum_{t=1}^{T} \log\left(1 + \frac{d_t}{T}\right)\right)$ for this problem compared to that of \cite{1,33}. Just like the setting in Joulani et al. \cite{33} and Quanrud and Khashabi \cite{11} feedbacks are not time-stamped, i.e. they do not contain the index of the round the loss function corresponds to. This rules out any issues related to synchronization which might come up in real applications. While time-stamping was required by the pooling strategies of Weinberger and Ordentlich \cite{29} and Joulani et al. \cite{33} in order to return the feedback to the appropriate learner. In this work we do not provide any general lower-bound for our result, however for some special cases we point out the known lower-bounds and show its tightness in these scenarios.

**Paper organization.** In the next section we provide the introductory definitions and background necessary for the main problem. Section 3 contains the main results, along with the proof of main
The following lemma will be useful. If \( f \) is differentiable at \( x \) then \( \partial f(x) \) has only one element denoted by \( \nabla f(x) \). Let \( d_t \in \mathbb{N} \) be the delay at round \( t \in [T] \). The feedback from round \( t \) is received at the end of round \( t + d_t - 1 \), and can be used in round \( t + d_t \). In the standard setting with no delays, \( d_t = 1 \) for all \( t \in [T] \). For each round \( t \), let \( F_t = \{ u \in [T] : u + d_u - 1 = t \} \) be the set of rounds whose feedback appears at the end of round \( t \). In the non-delayed settings we have \( F_t = \{ t \} \) for each \( t \).

**Convexity and strong convexity.** A function \( f : K \to \mathbb{R} \) is convex if
\[
f((1-\alpha)x + \alpha y) \leq (1-\alpha)f(x) + \alpha f(y) \quad \forall x, y \in K, \alpha \in [0, 1].
\]
If \( f \) is differentiable, then \( f \) is convex iff
\[
f(x) + \nabla f(x) \cdot (y - x) \leq f(y) \quad \forall x, y \in K. \tag{1}
\]
For convex function \( f \) but not necessarily differentiable, the subgradient of \( f \) at \( x \) is any vector that can replace \( \nabla f(x) \) in equation (1). The function \( f \) is strongly convex with modulus \( \sigma \) with respect to a norm \( \| \cdot \| \) if the function \( f(x) - \left(\sigma/2\right)\|x\|^2 \) is also convex. There are many equivalent properties for a function to be strongly convex; see, for example, [13, Section 2.1.3] or [35, Section 3.5]. In particular, for any two points \( x, y \in K \) and subgradient \( \nabla f|_x \in \partial f(x) \), we have
\[
f(y) \geq f(x) + \langle y - x, \nabla f|_x \rangle + \left(\sigma/2\right)\|y - x\|^2. \tag{2}
\]

**Projection to a convex set.** One important step in standard gradient-based algorithms is projection to the convex set \( K \). Define \( \pi_K \) to denote the map taking a point \( x \in \mathbb{R}^n \) to its closest point in the convex set \( K \) with respect to the Euclidian norm.

\[
\pi_K(x) \triangleq \arg \min_{y \in K} \|x - y\|_2
\]
The following lemma will be useful.

**Lemma 2.1.** Let \( K \) be a compact convex subset in \( \mathbb{R}^n \) and \( x \in \mathbb{R}^n \) a point, and let \( x' = \pi_K(x) \). Then
\[
\|x - y\|_2 \leq \|x' - y\|_2, \quad \forall y \in K.
\]

**Online gradient descent.** In online convex optimization, the input domain \( K \) is convex, and each cost function \( f_t \) is convex. For this setting Zinkevich’s algorithm online-gradient-descent [32] is as follows: The first point, \( x_1 \), is picked in \( K \) arbitrarily. After picking the \( t \)th point \( x_t \), online-gradient-descent computes the gradient \( \nabla f_t|_{x_t} \) of the loss function at \( x_t \), and chooses \( x_{t+1} = \pi_K(x_t - \eta \nabla f_t|_{x_t}) \) in the subsequent round, for some parameter \( \eta \in \mathbb{R}^+ \). Here, \( \pi_K \) is the projection that maps a point \( x' \) to its nearest point in \( K \). Zinkevich showed that, assuming the Euclidean diameter of \( K \) and the Euclidian norm of all the gradients \( \nabla f_t|_x \) are bounded, online-gradient-descent has an optimal regret bound of \( O\left(\sqrt{T}\right) \). Later Hazan and Kale [2] showed that by extra strong-convexity assumption on the loss function \( f_t \) one can get regret bound of \( O\left(\log T\right) \).

### 3 Main results

Following the setting of Hazan and Kale [2] we assume we are dealing with strongly convex loss functions. Also the feedbacks are delayed; hence at each round \( t \) we might receive more than one feedback. Following the natural common generalization [1][33] we apply the gradients the moment they are delivered. That is, we update \( x_{t+1} = x_t - \sum_{s \in F_t} \eta_s \nabla f_s|_{x_s} \) for some parameter \( \eta > 0 \), and then choose \((t+1)\)th point by projection \( x_{t+1} = \pi_K(x_{t+1}) \).
Theorem 3.1. Let $K$ be a compact convex set with diameter 1, let $f_1, \ldots, f_T$ be strongly-convex functions with modulus $\sigma$ over $K$ with $\|\nabla f_t(x)\|_2 \leq L$ for all $x \in K$ and $t \in [T]$. In the presence of adversarial delays $d_1, \ldots, d_T$ chosen arbitrarily in the span of $T$ rounds (i.e. $t + d_t \leq T$) online-gradient-descent selects points $x_1, \ldots, x_T \in K$ such that for all $y \in K$,

$$R(T) \leq \sum_{t=1}^{T} f_t(x_t) - \sum_{t=1}^{T} f_t(y) = O\left(\frac{L^2}{\sigma} \log T + \frac{L}{\sigma} \sum_{t=1}^{T} \log \left(1 + \frac{d_t}{t}\right)\right),$$

where $\eta_t$ is chosen as $\frac{2}{\sigma t|\mathcal{F}_t|}$.

There are a few interesting points in this regret bound: first, the delayed component is additive to base regret. This is compatible with Theorem 1 of Joulan et al. [33]. In addition, the effect of delays are normalized by round iteration: $\frac{d_t}{t}$; hence delays in the earlier rounds are more detrimental. Also note existence of $|\mathcal{F}_t|$ in $\eta_t$’s denominator is fine since whenever $|\mathcal{F}_t| = 0$, no update is done.

In Section 4 we show how this result encompasses and generalizes several of the existing known results in the literature, including the celebrated logarithmic regret of constant delays $d_t = \tau$ and regret bound of $O\left(\tau \log T\right)$ [3] for constant delays $d_t = \tau$.

Before proving the claim in Theorem 3.1 we provide the necessary lemmas. The first lemma models the delays (i.e. $d_t$) and regret bound of $O\left(\tau \log T\right)$ [3] when there are no delays (i.e. $d_t = 1$) and regret bound of $O\left(\tau \log T\right)$ [3] for constant delays $d_t = \tau$.

Before proving the claim in Theorem 3.1 we provide the necessary lemmas. The first lemma models the delays (i.e. $d_t$) and regret bound of $O\left(\tau \log T\right)$ [3] when there are no delays (i.e. $d_t = 1$) and regret bound of $O\left(\tau \log T\right)$ [3] for constant delays $d_t = \tau$.

Before proving the claim in Theorem 3.1 we provide the necessary lemmas. The first lemma models the delays (i.e. $d_t$) and regret bound of $O\left(\tau \log T\right)$ [3] when there are no delays (i.e. $d_t = 1$) and regret bound of $O\left(\tau \log T\right)$ [3] for constant delays $d_t = \tau$.

Lemma 3.2. Given the setting in Theorem 3.1, the regret function can be upper-bounded as following:

$$R(T) \leq \sum_{t=1}^{T} \left\|x_t - y\right\|^2 \left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} - \frac{\sigma^2}{2}\right) + L^2 \sum_{t=1}^{T} \frac{\eta_t}{2} \cdot |\mathcal{F}_t| + \sum_{t=1}^{T} \sum_{s \in \mathcal{F}_t} \nabla f_s(x_s) \cdot (x_s - x_{t,s}),$$

where $\frac{1}{\eta_0} \equiv 0$, $\frac{1}{\eta_T} \equiv \frac{2}{\sigma T}$ and $x_{t,s}$ is the partially updated version of $x_t$ using the instances before $s$ and delivered at time $t$ defined as $\mathcal{F}_{t,s} \equiv \{r \in \mathcal{F}_t : r < s\}$: $x_{t,s} \equiv x_t - \sum_{r \in \mathcal{F}_{t,s}} \eta_t \nabla f_r|x_r|$. 

Proof. Let $y = \arg \min_{x \in K}(f_1(x) + \cdots + f_T(x))$ be the best point in hindsight at the end of $T$ rounds. Fix $t \in [T]$ and consider the distance between $x_{t+1}$ and $y$. By Lemma 2.1 we know that $\|x_{t+1} - y\|_2 \leq \|x'_{t+1} - y\|_2$, where $x'_{t+1} = x_t - \sum_{s \in \mathcal{F}_t} \eta_t \nabla f_s|x_s|$. We split the sum of gradients applied in a single round and consider them one by one. Suppose $\mathcal{F}_t$ is nonempty, and fix $s \in \mathcal{F}_t$ to be the last index in $\mathcal{F}_t$. By Lemma 2.1 we have,

$$\|x_{t+1} - y\|_2^2 \leq \|x'_{t+1} - y\|_2^2 = \|x_{t,s'} - \eta_t \nabla f_{s'}|x_{s'}| - y\|_2^2 = \|x_{t,s'} - y\|_2^2 - 2\eta_t (\nabla f_{s'}|x_{s'}| \cdot (x_{t,s'} - y)) + \eta_t^2 \|\nabla f_{s'}|x_{s'}|\|_2^2.$$ 

Repeatedly unrolling the first term in this fashion gives

$$\|x_{t+1} - y\|_2^2 \leq \|x_t - y\|_2^2 - 2\eta_t \sum_{s \in \mathcal{F}_t} \nabla f_s|x_s| \cdot (x_{t,s} - y) + \sum_{s \in \mathcal{F}_t} \eta_t^2 \|\nabla f_s|x_s|\|_2^2.$$ 

For each $s \in \mathcal{F}_t$, by strong-convexity of $f$, we have,

$$-\nabla f_s|x_s| \cdot (x_{t,s} - y) = \nabla f_s|x_s| \cdot (x_t - x_{t,s}) = \nabla f_s|x_s| \cdot (y - x_t) + \nabla f_s|x_s| \cdot (x_s - x_{t,s}) \leq f_s(y) - f_s(x_t) - \sigma/2\|y - x_t\|^2 + \nabla f_s|x_s| \cdot (x_s - x_{t,s}).$$ 

By assumption, we also have $\|\nabla f_s|x_s|\|_2 \leq L$ for each $s \in \mathcal{F}_t$. With respect to the distance between $x_{t+1}$ and $y$, this gives,

$$\|x_{t+1} - y\|_2^2 \leq \|x_t - y\|_2^2 + 2\eta_t \sum_{s \in \mathcal{F}_t} \left(f_s(y) - f_s(x_t) - \sigma/2\|y - x_t\|^2 + \nabla f_s|x_s| \cdot (x_s - x_{t,s})\right) + L^2 \eta_t^2 |\mathcal{F}_t|.$$
Solving this inequality for the regret terms $\sum_{s \in F_t} f_s(x_s) - f_s(y)$,

$$\sum_{s \in F_t} (f_s(y) - f_s(x_s)) \leq \frac{1}{2\eta_t} \left[ \|x_t - y\|_2^2 - \|x_{t+1} - y\|_2^2 \right] - \frac{\sigma}{2} \sum_{s \in F_t} \|y - x_s\|^2 + \sum_{s \in F_t} \nabla f_s |_{x_s} \cdot (x_s - x_{t,s}) + L^2 \eta_t |F_t|, \forall t \in [T].$$

Now taking the sum of inequalities over all rounds $t \in [T]$, and re-ordering the elements we would get the desired result,

$$\sum_{t=1}^T (f_t(x_t) - f_t(y)) = \sum_{t=1}^T \sum_{s \in F_t} f_s(x_s) - f_s(y) \leq \sum_{t=1}^T \left[ \|x_t - y\|_2^2 \left( \frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} - \sigma/2 \right) \right] + \sum_{t=1}^T \sum_{s \in F_t} \nabla f_s |_{x_s} \cdot (x_s - x_{t,s}) + L^2 \sum_{t=1}^T \frac{\eta_t}{2} \cdot |F_t|. $$

\[ \blacksquare \]

The next lemma contains the upper-bound for the delay which is a key component of our result. Our proof technique is inspired by that of Quanrud and Khashabi \[\ref{lemma:delay-bound}\] Theorem 2.1, although the details of our combinatorial analysis are non-trivially different from theirs.

**Lemma 3.3.** Given the setting in Theorem 3.1 and for the choice of $\eta_t$, the delay expression in inequality \[\ref{delay-bound}\] $\sum_{t=1}^T \sum_{s \in F_t} \nabla f_s |_{x_s} \cdot (x_s - x_{t,s})$ is upper-bounded by $\frac{2L}{\eta} \sum_{t=1}^T \log (1 + \frac{d}{\eta t})$.

**Proof.** Using Cauchy-Schwarz inequality we get rid of the dot product,

$$\sum_{t=1}^T \sum_{s \in F_t} \nabla f_s |_{x_s} \cdot (x_s - x_{t,s}) \leq \sum_{t=1}^T \sum_{s \in F_t} \|\nabla f_s |_{x_s}\|_2 \|x_s - x_{t,s}\|_2 \leq L \sum_{t=1}^T \sum_{s \in F_t} \|x_s - x_{t,s}\|_2$$

It is easier to compare each $x_{t,s}$ with its older version $x_t$ rather that its later form $x_s$. Hence adding and subtracting $x_t$ inside $\|x_s - x_{t,s}\|_2$ and using the Lemma 2.1,

$$\|x_s - x_{t,s}\|_2 = \|x_s - x_t + x_t - x_{t,s}\|_2 \leq \|x_s - x_t\|_2 + \|x_t - x_{t,s}\|_2 \leq \|x_s - x'_t\|_2 + \|x_t - x_{t,s}\|_2$$

Also we simplify $\|x_s - x'_t\|_2$ based on the distance of the consecutive points,

$$\|x_s - x'_t\|_2 = \|x_s - x_{t-1} + x_{t-1} - x'_t\|_2 \leq \|x_s - x_{t-1}\|_2 + \|x_{t-1} - x'_t\|_2 \leq \|x_s - x'_{t-1}\|_2 + \|x_{t-1} - x'_t\|_2 \leq \sum_{r=s+1}^t \|x_{r-1} - x'_r\|_2$$

Putting these inequalities together we would get,

$$\|x_s - x_{t,s}\|_2 \leq \|x_t - x_{t,s}\|_2 + \sum_{r=s+1}^t \|x_{r-1} - x'_r\|_2$$

Now we sum this expression over time to get the regret term,

$$\sum_{t=1}^T \sum_{s \in F_t} \nabla f_s |_{x_s} \cdot (x_s - x_{t,s}) \leq \sum_{t=1}^T \sum_{s \in F_t} \|x_t - x_{t,s}\|_2 + \sum_{t=1}^T \sum_{s \in F_t} \sum_{r=s+1}^t \|x_{r-1} - x'_r\|_2 \leq \sum_{t=1}^T \sum_{s \in F_t} \sum_{p \in F_{t,s}} \eta_t \nabla f_p |_{x_p} + \sum_{t=1}^T \sum_{s \in F_t} \sum_{r=s}^{t-1} \sum_{q \in F_{t,r}} \eta_t \nabla f_q |_{x_q} \leq L \sum_{t=1}^T \sum_{s \in F_t} \eta_t |F_{t,s}| + L \sum_{t=1}^T \sum_{s \in F_t} \sum_{r=s}^{t-1} \eta_t |F_{t,r}|$$

(5)
The rest of the proof continues by making combinatorial arguments for counting and bounding each of the terms in the right-handside of the previous inequality.

We simplify the first term in equation (5). Consider the scenario in which $|F_t|$ feedbacks are received at time $t$, as shown in the following figure:

As we defined earlier $F_{t,s}$ is the subset of $F_t$ initiated before $s$. If $s$ is the first element of $F_t$, $|F_{t,s}| = 0$; similarly if $s$ is the final element of $F_t$, $|F_{t,s}| = |F_t| - 1$. Therefore the summation $\sum_{s\in F_t} |F_{t,s}|$ can simplified as $\sum_{i=0}^{T-1} i = T(T-1)/2$. Therefore, $\sum_{t=1}^{T} \sum_{s\in F_t} \eta_t |F_{t,s}| = \sum_{t=1}^{T} \eta_t |F_t|(|F_t| - 1)/2$. The rest of the first term can be simplified as follows.

$$\sum_{t=1}^{T} \sum_{s\in F_t} \eta_t |F_{t,s}| = \sum_{t=1}^{T} \eta_t |F_t| \frac{|F_t| - 1}{2} \leq \sum_{t=1}^{T} \eta_t |F_t| = \sum_{t=1}^{T} \eta_t |F_{t,s}|$$

where the last equality is true since $s + d_s = t$ holds for any $s \in F_t$ (hence dropping $|F_t|$ when writing in terms of $s$).

The second term is also simplified as follows,

$$\sum_{t=1}^{T} \sum_{s\in F_t} r \eta_t |F_{t,s}| = \sum_{t=1}^{T} \sum_{s=1}^{T} \eta_t |F_{t,s}|$$

Finally adding the two terms yields the following bound for the regret in (5),

$$\sum_{t=1}^{T} \sum_{s\in F_t} \eta_t |F_{t,s}| \cdot (x_t - x_{t,s}) \leq L \sum_{t=1}^{T} \sum_{s=1}^{T} \eta_t |F_{t,s}|$$

where $\sum_{s=1}^{T} \eta_t |F_{t,s}| \leq \frac{2L}{\sigma} \sum_{t=1}^{T} \sum_{s=1}^{T} \eta_t |F_{t,s}|$.

Here we give the proof of the main theorem. We put the result of Lemma 3.3 into Lemma 3.2 and follow the standard steps for bounding the undelayed problem [2].

**Proof of Theorem 3.1**: The delay term is bounded by $2L \sum_{t=1}^{T} \log (1 + \frac{d_t}{T})$ in Lemma 3.2. We bound each term of equation (3). We bound the first term:

$$\sum_{t=1}^{T} \left[ \frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} - \sigma/2 \right] = \frac{1}{\eta_T} - \frac{1}{\eta_0} - \sigma/2 T = \frac{\sigma}{2} T - 0 - \frac{\sigma}{2} T = 0.$$

Next we bound the summation in the second term of equation (4). Note that any for $t \in [T]$ that $|F_t| \neq 0$, $\frac{1}{T} \cdot |F_t| = \frac{1}{\sigma^2}$, given the choice of $\eta_t$ (Theorem 3.1); hence,

$$\sum_{t=1}^{T} \frac{\eta_t}{2} \cdot |F_t| \leq \frac{L^2}{\sigma} \left[ \sum_{t=1}^{T-1} \frac{1}{T} + \frac{|F_T|}{T} \right] \leq \frac{L^2}{\sigma} \left[ \log(T) + 1 \right].$$

Combining these with result of Lemma 3.2 would give the desired.

**4 Discussion**

We briefly study several of the interesting special cases of our result. As mentioned earlier we do not have a general lower-bound for our result; however for some of the special cases the lower-bounds are known. In what follows wherever possible we point out the lower-bounds as well. For simplicity we drop constants and the big-$O$ notation when using equation (3) for the rest of this section.
Undelayed setting. Hazan and Kale analyzed the undelayed setting \(^2\) Theorem 1) and obtain \(O(\log T)\) which matches the regret bound of \(^3\) Theorem 3.1) when \(d_t\) replaced with 1:

\[
R(T) \leq \log(T) + \sum_{t=1}^{T} \log(1 + 1/t) = \log(T) + \sum_{t=1}^{T} \log(t + 1) - \log t
\]

\[
= \log(T) + \log(T + 1) - \log 1 = O(\log T)
\]

This bound is proved to be tight by Hazan et al. \(^17\).

Constant delays. Suppose each delay \(d_t\) is bounded by some fixed value \(\tau\). The actual game is played for \(\tilde{T}\) rounds and \(\tau\) extra rounds until all feedbacks are received (hence in overall \(T = \tilde{T} + \tau\) rounds). The regret is upper-bounded by:

\[
R(T) \leq \log(\tilde{T} + \tau) + \sum_{t=1}^{\tilde{T}} \log(1 + \tau/t) = \log(\tilde{T} + \tau) + \log\left(\frac{\tilde{T} + \tau}{\tau}\right)
\]

\[
\leq \log(\tilde{T} + \tau) + \log\left(\frac{(\tilde{T} + \tau)\tau}{\tau!}\right)
\]

\[
\approx \log(\tilde{T} + \tau) + \tau \log(\tilde{T} + \tau) - \tau \log \tau = O\left(\tau \log \frac{T}{\tau}\right)
\]

This matches the \(O(\tau \log T/\tau)\) bound of Langford et al. \(^3\) Theorem 4). Also this is compatible with the black-box analysis of delayed learning \(^29\)\(^31\) which gives \(O(\tau \log T/\tau)\) regret bound.

Stochastic delays. Suppose the delays are sampled i.i.d. from some a distribution \(g(d)\) with mean \(\tau = \mathbb{E}_{d_t \sim g}[d_t]\) and with variance \(\lambda = \mathbb{V}_{d_t \sim g}[d_t]\). Then by applying Taylor expansion we can write:

\[
\mathbb{E}[R(T)] \leq \mathbb{E}\left[\sum_t \log(1 + d_t/t)\right] = \sum_t \mathbb{E}_{d_t \sim g}[\log(1 + d_t/t)]
\]

\[
\approx \sum_t \log \left(1 + \frac{\mathbb{E}_{d_t \sim g}[d_t]}{t}\right) - \frac{\mathbb{V}_{d_t \sim g}[d_t]}{2(t^2 + t\mathbb{E}_{d_t \sim g}[d_t])}
\]

\[
\approx \sum_t \log \left(1 + \frac{\tau}{t}\right) - \frac{\lambda}{2(1 + \tau)} \sum_t \frac{1}{t^2}
\]

which is again \(O(\tau \log(T/\tau))\) regret bound; hence only an assumption on the expected value of delays is enough to find closed-form regret bounds.

All feedbacks received in the final round. Since all the feedbacks are received in the very last round, in this scenario the online learner essentially learns nothing. Hence the regret is \(\Theta(T)\) (upper and lower bound). Plugging \(d_t = T - t\) into \(^3\) Theorem 3.1)

\[
R(T) \leq \log T + \sum_t \log(1 + (T - t)/t) = \log T + \sum_t \log(T/t) \approx \log T + \log \left(T/T^t/T!\right) = O(T)
\]

which is compatible with what we expected. In the last step we used Sterling’s approximation.

First \(k\) feedbacks received in the final round. This special case is interesting since in a sense, it interpolates between two special cases: undelayed setting \((k = 0\) and \(R(T) = \Theta(\log T)\)) and all-delayed setting \((k = T\) and \(R(T) = \Theta(T)\)). Rewriting the regret bound of \(^3\) Theorem 3.1)

\[
R(T) \leq \log T + \sum_{t \leq k} \log(1 + \frac{T - t}{t}) + \sum_{t > k} \log(1 + 1/t)
\]

\[
= \log T + \log\left(\frac{T^k}{k^k}\right) + \log \frac{T}{k} = O(\log T + k (\log T - \log k) + k).
\]

This suggests that only for \(k = o\left(\frac{T}{\log T}\right)\) the regret bound is sublinear (i.e. learning is possible), otherwise regret growth is linear with \(T\).
4.1 Toy Experiment

Doing a thorough experiment to verify our result is not trivial. On the other hand special cases like constants delays or stochastic delays with (fixed expected) are not very interesting as they are well-studied. In order to get empirical intuition about our theoretical results and partially verify it, we simulate the special “first $k$ feedbacks received in the final round”. Specifically we study the scenario in which first $k$ feedbacks are received in the final round. We choose $f_t$ as following:

$$f_t(x) = \begin{cases} 
\|x - 0.5\|^2 & t = \text{even} \\
\|x + 0.5\|^2 & t = \text{odd}.
\end{cases}$$

Figure 1 shows the regret growth as a function of number of rounds. Each graphs correspond a different value of $k$ with its corresponding $k$ shown on top of each graph. For $k$s close to $T$ the growth is almost linear with $T$, while for smaller $k$’s it is almost logarithmic. The red line corresponds to $k = \frac{T}{\log T} \approx 145$, almost where the transition between linear and logarithmic regret happens, as expected according to our theoretical analysis.

5 Conclusion

Along the line of recent efforts for finding general distribution free bounds for delayed online learning we provided a simple and intuitive regret bound for stochastic-gradient-descent. As a sanity check in special cases we showed that the bound matches the existing known results and it matches the known lower bounds.

Finding a general lower-bound remains an open problem for future work.

References

[1] K. Quanrud and D. Khashabi. Online learning with adversarial delays. In Adv. Neural Info. Proc. Sys. 29 (NIPS), pages 1270–1278, 2015.
[2] E. Hazan, A. Agarwal, and S. Kale. Logarithmic regret algorithms for online convex optimization. Mach. Learn. J., 69(2-3):169–192, 2007.
[3] J. Langford, A. J. Smola, and M. Zinkevich. Slow learners are fast. In Adv. Neural Info. Proc. Sys. 22 (NIPS), pages 2331–2339, 2009.
[4] A. Blum. On-line algorithms in machine learning. In A. Fiat and G. Woeginger, editors, Online algorithms, volume 1442 of LNCS, chapter 14, pages 306–325. Springer Berlin Heidelberg, 1998.
[5] S. Shalev-Shwartz. Online learning and online convex optimization. Found. Trends Mach. Learn., 4(2):107–194, 2011.
[6] E. Hazan. Introduction to online convex optimization. Internet draft available at http://ocobook.cs.princeton.edu 2015.
[7] Y. Freund and R. E. Schapire. A desicion-theoretic generalization of on-line learning and an application to boosting. In Comp. Learning Theory, pages 23–37. Springer, 1995.
[8] A. Kalai and S. Vempala. Efficient algorithms for online decision problems. J. Comput. Sys. Sci., 71:291–307, 2005. Extended abstract in Proc. 16th Conf. on Learning Theory (COLT), 2003.
[9] S. Shalev-Shwartz, O. Shamir, N. Srebro, and K. Sridharan. Stochastic convex optimization. In Proc. 22nd Conf. on Learning Theory (COLT), 2009.
[10] H. Robbins and S. Monro. A stochastic approximation method. The annals of mathematical statistics, pages 400–407, 1951.
[11] A.S. Nemirovski and D.B. Yudin. Problem complexity and method efficiency in optimization. Wiley-Interscience Series in Discrete Mathematics. John Wiley & Sons, 1983.
[12] D. P. Bertsekas and J. N. Tsitsiklis. Parallel and Distributed Computation: Numerical Methods. Prentice-Hall, 1989.
[13] Y. Nesterov. *Introductory lectures on convex optimization*, volume 87 of *Applied Optim*. Springer, 2004.

[14] A. Rakhlin, O. Shamir, and K. Sridharan. Making gradient descent optimal for strongly convex stochastic optimization. In *Proc. 29th Int. Conf. Mach. Learning* (ICML), 2012.

[15] A. Agarwal, P. L. Bartlett, P. Ravikumar, and M. J. Wainwright. Information-theoretic lower bounds on the oracle complexity of convex stochastic optimization. *IEEE Trans. on Info. Theory*, 58(5):3235–3249, 2012.

[16] O. Shamir and T. Zhang. Stochastic gradient descent for non-smooth optimization: Convergence results and optimal averaging schemes. In *Proc. 30th Int. Conf. Mach. Learning* (ICML), pages 71–79, 2013.

[17] E. Hazan and S. Kale. Beyond the regret minimization barrier: optimal algorithms for stochastic strongly-convex optimization. *J. Mach. Learn. Research*, 15(1):2489–2512, 2014.

[18] Y. Arjevani, S. Shalev-Shwartz, and O. Shamir. On lower and upper bounds for smooth and strongly convex optimization problems. *arXiv preprint arXiv:1503.06833*, 2015.

[19] B. Recht, C. Re, S. Wright, and F. Niu. Hogwild: a lock-free approach to parallelizing stochastic gradient descent. In *Adv. Neural Info. Proc. Sys. 24* (NIPS), pages 693–701, 2011.

[20] J. Duchi, M.I. Jordan, and B. McMahan. Estimation, optimization, and parallelism when data is sparse. In *Adv. Neural Info. Proc. Sys. 26* (NIPS), pages 2832–2840, 2013.

[21] H.B. McMahan and M. Streeter. Delay-tolerant algorithms for asynchronous distributed online learning. In *Adv. Neural Info. Proc. Sys. 27* (NIPS), pages 2915–2923, 2014.

[22] J. Duchi, E. Hazan, and Y. Singer. Adaptive subgradient methods for online learning and stochastic optimization. *J. Mach. Learn. Research*, 12:2121–2159, July 2011.

[23] J. Liu, S. J. Wright, C. Ré, V. Bittorf, and S. Sridhar. An asynchronous parallel stochastic coordinate descent algorithm. *J. Mach. Learn. Research*, 16:285–322, 2015.

[24] S. Chaturapruek, J. C. Duchi, and C. Ré. Asynchronous stochastic convex optimization: the noise is in the noise and SGD don’t care. In *Adv. Neural Info. Proc. Sys. 28* (NIPS), pages 1531–1539, 2015.

[25] S. J. Wright. Coordinate descent algorithms. *Math. Prog.*, 151(3–34), 2015.

[26] H. Mania, X. Pan, D. Papailiopoulos, B. Recht, K. Ramchandran, and M. I. Jordan. Perturbed iterate analysis for asynchronous stochastic optimization. *arXiv preprint arXiv:1507.06970*, 2015.

[27] S. Sra, A. W. Yu, M. Li, and A. J. Smola. Adadelay: Delay adaptive distributed stochastic convex optimization. *arXiv preprint arXiv:1508.05003*, 2015.

[28] B. Sirb and X. Ye. Decentralized consensus algorithm with delayed and stochastic gradients. *arXiv preprint arXiv:1604.05649*, 2016.

[29] M.J. Weinberger and E. Ordentlich. On delayed prediction of individual sequences. *IEEE Trans. Inf. Theory*, 48(7):1959–1976, 2002.

[30] C. Mesterharm. On-line learning with delayed label feedback. In *Alg. Learning Theory*, pages 399–413. Springer, 2005.

[31] P. Joulani, A. György, and C. Szepesvári. Online learning under delayed feedback. In *Proc. 30th Int. Conf. Mach. Learning* (ICML), volume 28, 2013.

[32] M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proc. 20th Int. Conf. Mach. Learning* (ICML), pages 928–936, 2003.

[33] P. Joulani, A. György, and C. Szepesvári. Delay-tolerant online convex optimization: Unified analysis and adaptive-gradient algorithms. In *Proc. 30th Conf. on Artificial Intelligence* (AAAI), pages 1744–1750, 2016.

[34] N. Cesa-Bianchi, C. Gentile, Y. Mansour, and A. Minora. Delay and cooperation in nonstochastic bandits. *arXiv preprint arXiv:1602.04741*, 2016.

[35] C. Zălinescu. *Convex analysis in general vector spaces*. World Scientific, 2002.