
Online mirror descent and dual averaging: keeping pace in the dynamic case

Huang Fang¹ Nicholas J. A. Harvey¹ Victor S. Portella¹ Michael P. Friedlander¹

Abstract

Online mirror descent (OMD) and dual averaging (DA)—two fundamental algorithms for online convex optimization—are known to have very similar (and sometimes identical) performance guarantees when used with a *fixed* learning rate. Under *dynamic* learning rates, however, OMD is provably inferior to DA and suffers a linear regret, even in common settings such as prediction with expert advice. We modify the OMD algorithm through a simple technique that we call *stabilization*. We give essentially the same abstract regret bound for OMD with stabilization and for DA by modifying the classical OMD convergence analysis in a careful and modular way that allows for straightforward and flexible proofs. Simple corollaries of these bounds show that OMD with stabilization and DA enjoy the same performance guarantees in many applications—even under dynamic learning rates. We also shed light on the similarities between OMD and DA and show simple conditions under which stabilized-OMD and DA generate the same iterates.

1. Introduction

Online convex optimization (OCO) lies at the intersection of machine learning, convex optimization, and game theory. In OCO, a player is required to make a sequence of online decisions over discrete time steps. Each decision incurs a cost given by a convex function that is only revealed to the player after they make that decision. The goal of the player is to minimize what is known as *regret*: the difference between the total cost and the cost of a competitor with

the benefit of hindsight. Letting T denote the number of decisions, the goal is for the player’s algorithm to ensure its regret is sublinear in T .

Online mirror descent (OMD) and dual averaging (DA) are two important algorithm templates for OCO from which many classical online learning algorithms can be derived as special cases; see Shalev-Shwartz (2012) and McMahan (2017) for examples. When the number T of decisions to be made is known in advance, the performance of OMD and DA (with properly chosen constant learning rates) can be shown to be very similar (Hazan, 2016). That is, they achieve essentially the same regret bound when using the same learning rate. However, when the number of decisions is *not* known a priori, there is a fundamental difference in the regret guarantees for OMD and DA with a similar dynamic (time-varying) learning rate: while DA can guarantee sublinear regret bound $O(\sqrt{T})$ for any $T > 0$ (Nesterov, 2009), there are instances for which OMD suffers asymptotically linear $\Omega(T)$ regret (Orabona & Pál, 2018). We summarize this discussion as follows.

Previously known fact. With a dynamic learning rate, OMD does *not* match the performance of DA.

The purpose of this paper is to introduce a *stabilization* technique that bridges the gap between OMD and DA with dynamic learning rates.

Main result (informal). With a dynamic learning rate, stabilized-OMD matches the performance of DA.

For a formal statement, see the abstract regret bounds in Theorems 4.1, 4.3, and 4.6. In Section 5 we give some applications: regret bounds with strongly convex mirror maps; for prediction with expert advice, anytime regret bounds with the best known constant, and a first-order regret bound. Also, in Section 6 we formally compare the iterates of DA and stabilized-OMD. This sheds light on the drawbacks of OMD with dynamic learning rate and why stabilization helps. To conclude, we derive simple conditions under which stabilized-OMD and DA generate exactly the same iterates. This is analogous to the relationship between OMD and DA with a fixed learning rate, and is evidence that stabilization may be a natural way to extend OMD to dynamic learning rates. Additionally, in Appendix I we

*Equal contribution ¹Department of Computer Science, University of British Columbia, Canada. Correspondence to: Huang Fang <hgfang@cs.ubc.ca>, Nicholas J. A. Harvey <nickhar@cs.ubc.ca>, Victor S. Portella <victorsp@cs.ubc.ca>, Michael P. Friedlander <mpf@cs.ubc.ca>.

adapt stabilized-OMD for the composite objective setting, generalizing a result of [Duchi et al. \(2010\)](#).

2. Related work

Mirror descent (MD) originated with [Nemirovski & Yudin \(1983\)](#). [Beck & Teboulle \(2003\)](#) give a modern treatment. Recent interest in first-order methods for large-scale problems have boosted the popularity of OMD. See, e.g., [Duchi et al. \(2010\)](#); [Allen-Zhu & Orecchia \(2016\)](#); [Beck \(2017\)](#). DA is due to [Nesterov \(2009\)](#) and was later extended to regularized problems by [Xiao \(2010\)](#). DA is closely related to the *follow-the-regularized-leader* (FTRL) algorithm. Standard references for these algorithms include [Shalev-Shwartz \(2012\)](#); [Bubeck \(2015\)](#); [Hazan \(2016\)](#); [McMahan \(2017\)](#). OMD and DA have seen an increase in popularity due to applications in online learning problems ([Kakade et al., 2012](#); [Audibert et al., 2014](#)) and since they generalize a wide range of online learning algorithms ([Shalev-Shwartz, 2012](#); [McMahan, 2017](#)).

Unifying views of online learning algorithms have been shown to be useful for applications and have drawn recent attention. [McMahan \(2017\)](#) uses FTRL with adaptive regularizers to derive many online learning algorithms. [Joulani et al. \(2017\)](#) propose a unified framework to analyze online learning algorithms, even for non-convex problems. Recently, [Juditsky et al. \(2019\)](#) proposed a unified framework called unified mirror descent (UMD) that encompasses OMD and DA as special cases.

Despite these unifying frameworks, the differences between OMD and DA seem to have been overlooked. Only recently, [Orabona & Pál \(2018\)](#), who looked more closely at the difference between OMD and DA, presented counter examples to demonstrate that OMD with dynamic learning rate could suffer from linear regret even under the well-studied settings as in the experts' problem.

For the problem of prediction with expert advice, [Cesa-Bianchi et al. \(1997\)](#) use the *doubling trick* to give an algorithm with a sublinear *anytime* regret bound, meaning a bound parameterized by T and that holds for all T . Improved anytime regret bounds were developed by [Auer et al. \(2002b\)](#); a simplified description of this result appears in by [Cesa-Bianchi & Lugosi \(2006, §2.3\)](#). Sublinear anytime regret bounds for DA follow directly from the analysis of [Nesterov \(2009\)](#). Other expositions include ([Bubeck, 2011, Theorem 2.4](#)) and ([Gerchinovitz, 2011, Proposition 2.1](#)).

First-order regret bounds are bounds that depend on the cost of the best expert instead of on the number of decisions T . Such bounds can be proven using the doubling trick, as shown by [Cesa-Bianchi et al. \(1997\)](#). First-order regret bounds without using the doubling trick were proven by [Auer et al. \(2002b\)](#). Improved constants are known; see,

e.g., [de Rooij et al. \(2014, Theorem 8\)](#). The best known first-order regret bound, in some settings, is from a sophisticated algorithm designed by [Yaroshinsky et al. \(2004\)](#).

3. Formal definitions

We consider the online convex optimization problem with unknown time horizon. For each time step $t \in \{1, 2, \dots\}$ the algorithm proposes a point x_t from a closed convex set $\mathcal{X} \subseteq \mathbb{R}^n$, and an adversary simultaneously picks a convex cost function f_t to which the algorithm has access via a first order oracle, that is, for any $x \in \mathcal{X}$ the algorithm can compute $f_t(x)$ and a subgradient $g \in \partial f_t(x) := \{g \in \mathbb{R}^n : f(z) \geq f(x) + \langle g, z - x \rangle \forall z \in \mathcal{X}\}$. This function penalizes the proposal x_t by the amount $f_t(x_t)$. The cost of the iteration at time t is defined as $f_t(x_t)$. The goal is to produce a sequence of proposals $\{x_t\}_{t \geq 1}$ that minimizes the regret against a unknown comparison point $z \in \mathcal{X}$ that has accrued up until time T :

$$\text{Regret}(T, z) := \sum_{t=1}^T f_t(x_t) - \sum_{t=1}^T f_t(z).$$

We consider the case where the algorithm does not know the time horizon T in advance. Hence any parameters of the algorithms, including learning rate, cannot depend on T .

We assume that each function in the sequence $\{f_t\}_{t \geq 1}$ is L -Lipschitz (continuous) over \mathcal{X} with respect to a norm $\|\cdot\|$, and we denote the dual norm of $\|\cdot\|$ by $\|\cdot\|_*$.

Both OMD and DA are parameterized by a mirror map (for \mathcal{X}), that is, a closed convex function of Legendre type ([Rockafellar, 1970, Chapter 26](#)) $\Phi: \bar{\mathcal{D}} \rightarrow \mathbb{R}$ whose conjugate is differentiable on \mathbb{R}^n and with $\bar{\mathcal{D}} \subseteq \mathbb{R}^n$ a convex set such that $\mathcal{D} \cap \text{ri } \mathcal{X} \neq \emptyset$, where $\mathcal{D} := \text{int } \bar{\mathcal{D}}$ and $\text{ri } \mathcal{X}$ is the relative interior of \mathcal{X} . The gradient of the mirror map $\nabla \Phi: \mathcal{D} \rightarrow \mathbb{R}^n$ and the gradient of its conjugate $\nabla \Phi^*: \mathbb{R}^n \rightarrow \mathcal{D}$ are mutually inverse bijections between the primal space \mathcal{D} and the dual space \mathbb{R}^n ([Rockafellar, 1970, Theorem 26.5](#)). We will adopt the following notational convention. Any vector in the primal space will be written without a hat, such as $x \in \mathcal{D}$. The same letter with a hat, namely \hat{x} , will denote the corresponding dual vector:

$$\hat{x} := \nabla \Phi(x) \quad \text{and} \quad x := \nabla \Phi^*(\hat{x}) \quad \text{for all letters } x.$$

Given a mirror map Φ , the Bregman divergence of $x \in \bar{\mathcal{D}}$ and $y \in \mathcal{D}$ w.r.t. Φ is defined by $D_\Phi(x, y) := \Phi(x) - \Phi(y) - \langle \nabla \Phi(y), x - y \rangle$. Throughout this paper it will be convenient to use the notation

$$D_\Phi\left(\begin{smallmatrix} a \\ b \end{smallmatrix}; c\right) := D_\Phi(a, c) - D_\Phi(b, c).$$

The projection operator induced by the Bregman divergence is $\Pi_{\mathcal{X}}^\Phi(y) := \arg \min\{D_\Phi(x, y) : x \in \mathcal{X}\}$.

A general template for optimization in the mirror descent

framework is shown in Algorithm 1. OMD and DA are incarnations of this framework, differing only in how the dual variable \hat{y}_t is updated.

Algorithm 1 Pseudocode for OMD and DA.

Input: $x_1 \in \mathcal{X} \cap \mathcal{D}, \eta : \mathbb{N} \rightarrow \mathbb{R}_{>0}$.
for $t = 1, 2, \dots$ **do**
 Incur cost $f_t(x_t)$ and receive $\hat{g}_t \in \partial f_t(x_t)$
 $\hat{x}_t = \nabla \Phi(x_t)$
[OMD update] $\hat{y}_{t+1} = \hat{x}_t - \eta_t \hat{g}_t$
[DA update] $\hat{y}_{t+1} = \hat{x}_1 - \eta_{t+1} \sum_{i \leq t} \hat{g}_i$
 $y_{t+1} = \nabla \Phi^*(\hat{y}_{t+1})$
 $x_{t+1} = \Pi_{\mathcal{X}}^{\Phi}(y_{t+1})$
end for

4. Stabilized-OMD

Orabona & Pál (2018) showed that OMD with the standard dynamic learning rate ($\eta_t \propto 1/\sqrt{t}$) can incur regret linear in T when the feasible set \mathcal{X} is unbounded Bregman divergence, that is, $\sup_{x, z \in \mathcal{X}} D_{\Phi}(z, x) = \infty$. We introduce a stabilization technique that resolves this problem, allowing OMD to support a dynamic learning rate and perform similarly to DA even when the Bregman divergence on \mathcal{X} is unbounded.

The intuition for the idea is as follows. Suppose $\mathcal{Z} \subseteq \mathcal{X}$ is a set of comparison points with respect to which we wish our algorithm to have low regret. Usually, we assume $\sup_{z \in \mathcal{Z}} D_{\Phi}(z, x_1)$ is bounded, that is, the initial point is not too far (with respect to the Bregman divergence) from any comparison point. Since $\sup_{z \in \mathcal{Z}} D_{\Phi}(z, x_1)$ is bounded (but not necessarily $\sup_{z \in \mathcal{Z}, x \in \mathcal{X}} D_{\Phi}(z, x)$), the point x_1 is the only point in \mathcal{X} that is known to be somewhat close (w.r.t. the Bregman divergence) to all the other points in \mathcal{X} . Thus, iterates computed by the algorithm should remain reasonably close to x_1 so that no other point $z \in \mathcal{Z}$ is too far from the iterates. If there were such a point z , an adversary could later choose functions so that picking z every round would incur low loss. At the same time, OMD would take many iterations to converge to z since consecutive OMD iterates tend to be close w.r.t. the Bregman divergence. That is, the algorithm would have high regret against z . To prevent this, the stabilization technique modifies each iterate x_t to mix in a small fraction of x_1 . This idea is not entirely new: it appears, for example, in the original Exp3 algorithm (Auer et al., 2002a), although for different reasons.

There are two ways to realize the stabilization idea.

Primal Stabilization. Replace x_t with a convex combination of x_t and x_1 .

Dual Stabilization. Replace \hat{y}_t with a convex combination of \hat{y}_t and \hat{x}_1 (Recall from Algorithm 1 that \hat{y}_t is the dual

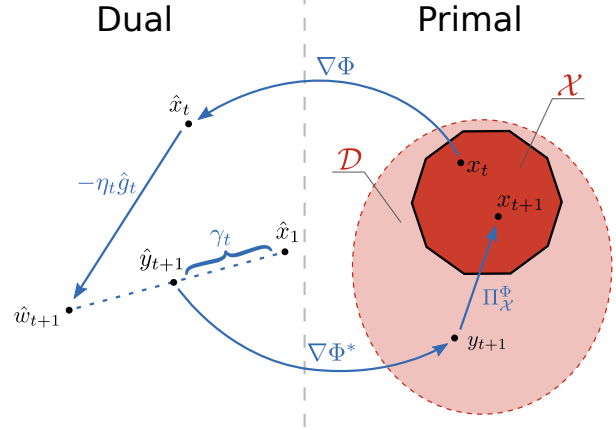


Figure 1: Illustration of the t -th iteration of DS-OMD.

iterate computed by taking a gradient step). An illustration for dual stabilization is shown in Figure 1.

After a draft of this paper was made publicly available, we were informed that an idea similar to primal stabilization had appeared in the Robust Optimistic Mirror Descent algorithm (Kangarshahi et al., 2018). Their setting is somewhat different since they perform optimistic steps. Furthermore, their results are somewhat weaker in terms of constant factors and since they cannot handle Bregman projections.

In this section we use many results regarding Bregman divergences (see Appendix A.2), and for ease of reference we will state the main results. Let $a, b, c \in \mathcal{D}$. A classic result is the three-point identity (Bubeck, 2015, §4)

$$D_{\Phi}(a, c) - D_{\Phi}(b, c) + D_{\Phi}(b, a) = \langle \hat{a} - \hat{c}, a - b \rangle. \quad (4.1)$$

If $\gamma \hat{a} + (1 - \gamma) \hat{b} = \hat{c}$ for some $\gamma \in \mathbb{R}$, then, for all $u, v \in \bar{\mathcal{D}}$,

$$\gamma D_{\Phi}(u; a) + (1 - \gamma) D_{\Phi}(u; b) = D_{\Phi}(u; c). \quad (4.2)$$

Finally, if $p \in \mathcal{D}$ and $\pi := \Pi_{\mathcal{X}}(p)$, then

$$D_{\Phi}(z; p) \geq D_{\Phi}(z; \pi) = D_{\Phi}(z, \pi) \quad \forall z \in \mathcal{X}. \quad (4.3)$$

4.1. Dual-stabilized OMD

Algorithm 2 gives pseudocode showing our modification of OMD to incorporate dual stabilization. Theorem 4.1 analyzes it without assuming strong convexity of Φ .

Theorem 4.1 (Regret bound for dual-stabilized OMD). Let $\eta : \mathbb{N} \rightarrow \mathbb{R}_{>0}$ be such that $\eta_t \geq \eta_{t+1}$ for all $t \geq 1$. Define $\gamma_t = \eta_{t+1}/\eta_t \in (0, 1]$ for all $t \geq 1$. Let $\{f_t\}_{t \geq 1}$ be a sequence of convex functions with $f_t : \mathcal{X} \rightarrow \mathbb{R}$ for each $t \geq 1$. Let $\{x_t\}_{t \geq 1}$ and $\{\hat{w}_t\}_{t \geq 2}$ be as in Algorithm 2. Then, for all $T > 0$ and $z \in \mathcal{X}$,

$$\text{Regret}(T, z) \leq \sum_{t=1}^T \frac{D_{\Phi}(x_t; w_{t+1})}{\eta_t} + \frac{D_{\Phi}(z, x_1)}{\eta_{T+1}}. \quad (4.7)$$

Algorithm 2 Dual-stabilized OMD (DS-OMD). The parameters γ_t control the amount of stabilization.

Input: $x_1 \in \mathcal{X}, \eta : \mathbb{N} \rightarrow \mathbb{R}_+, \gamma : \mathbb{N} \rightarrow (0, 1]$
for $t = 1, 2, \dots$ **do**
 Incur cost $f_t(x_t)$ and receive $\hat{g}_t \in \partial f_t(x_t)$
 $\hat{x}_t = \nabla \Phi(x_t)$
 $\hat{w}_{t+1} = \hat{x}_t - \eta_t \hat{g}_t$ (4.4)
 $\hat{y}_{t+1} = \gamma_t \hat{w}_{t+1} + (1 - \gamma_t) \hat{x}_1$ (4.5)
 $y_{t+1} = \nabla \Phi^*(\hat{y}_{t+1})$
 $x_{t+1} = \Pi_{\mathcal{X}}^{\Phi}(y_{t+1})$ (4.6)
end for

Note that strong convexity of Φ is *not* assumed. As we will see in Section 5.1, the term $D_{\Phi}(\hat{x}_t; w_{t+1})$ can be easily bounded when the mirror map is strongly convex. This yields sublinear regret for $\eta_t \propto 1/\sqrt{t}$, which is not the case for OMD when $\sup_{z \in \mathcal{Z}, x \in \mathcal{X}} D_{\Phi}(z, x) = +\infty$, where $\mathcal{Z} \subseteq \mathcal{X}$ is a fixed set of comparison points.

Proof (of Theorem 4.1).

The first step is the same as in the standard OMD proof. For all $z \in \mathcal{X}$, use the subgradient inequality to deduce

$$\begin{aligned} f_t(x_t) - f_t(z) &\leq \langle \hat{g}_t, x_t - z \rangle \\ &\stackrel{(i)}{=} \frac{1}{\eta_t} \langle \hat{x}_t - \hat{w}_{t+1}, x_t - z \rangle \\ &\stackrel{(ii)}{=} \frac{1}{\eta_t} (D_{\Phi}(\hat{x}_t; w_{t+1}) + D_{\Phi}(z, x_t)), \end{aligned} \quad (4.8)$$

where (i) follows from (4.4), and (ii) from (4.1).

The next step exhibits the main point of stabilization. Without stabilization we would have $x_{t+1} = \Pi_{\mathcal{X}}^{\Phi}(w_{t+1})$ and $D_{\Phi}(z, w_{t+1}) \geq D_{\Phi}(z, x_{t+1}) + D_{\Phi}(x_{t+1}, w_{t+1})$ by (4.3), so (4.8) would lead to a telescoping sum involving $D_{\Phi}(z, \cdot)$ if the learning rate were fixed. With a dynamic learning rate the analysis is trickier: we must obtain telescoping terms by relating $D_{\Phi}(z, w_{t+1})$ to $D_{\Phi}(z, x_{t+1})$. This is the purpose of the next claim.

Claim 4.2. Assume that $\gamma_t = \eta_{t+1}/\eta_t \in (0, 1]$. Then

$$\begin{aligned} (4.8) &\leq \frac{D_{\Phi}(\hat{x}_t; w_{t+1})}{\eta_t} + \underbrace{\left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t} \right)}_{\text{telescopes}} D_{\Phi}(z, x_1) \\ &\quad + \underbrace{\frac{D_{\Phi}(z, x_t)}{\eta_t} - \frac{D_{\Phi}(z, x_{t+1})}{\eta_{t+1}}}_{\text{telescopes}}. \end{aligned}$$

Proof. First we derive the inequality

$$\begin{aligned} &\gamma_t D_{\Phi}(\hat{x}_t; w_{t+1}) + (1 - \gamma_t) D_{\Phi}(z, x_1) \\ &\stackrel{(i)}{\geq} \gamma_t D_{\Phi}(\hat{x}_t; w_{t+1}) + (1 - \gamma_t) D_{\Phi}(\hat{x}_t; x_1) \end{aligned}$$

Algorithm 3 OMD with primal stabilization.

Input: $x_1 \in \mathbb{R}^n, \eta : \mathbb{N} \rightarrow \mathbb{R}, \gamma : \mathbb{N} \rightarrow \mathbb{R}$.
for $t = 1, 2, \dots$ **do**
 Incur cost $f_t(x_t)$ and receive $\hat{g}_t \in \partial f_t(x_t)$
 $\hat{x}_t = \nabla \Phi(x_t)$
 $\hat{w}_{t+1} = \hat{x}_t - \eta_t \hat{g}_t$
 $w_{t+1} = \nabla \Phi^*(\hat{w}_{t+1})$
 $y_{t+1} = \Pi_{\mathcal{X}}^{\Phi}(w_{t+1})$ (4.10)
 $x_{t+1} = \gamma_t y_{t+1} + (1 - \gamma_t) x_1$ (4.11)
end for

$$\begin{aligned} &= D_{\Phi}(\hat{x}_t; w_{t+1}) \quad (\text{by (4.2) and (4.5)}) \\ &\geq D_{\Phi}(z, w_{t+1}) \quad (\text{by (4.3) and (4.6)}), \end{aligned}$$

where (i) uses the fact that $D_{\Phi}(x_{t+1}, x_1) \geq 0$ and $\gamma_t \leq 1$. Rearranging and using $\gamma_t > 0$ yields

$$\begin{aligned} D_{\Phi}(z, w_{t+1}) &\geq D_{\Phi}(x_{t+1}, w_{t+1}) \\ &\quad - \frac{1 - \gamma_t}{\gamma_t} D_{\Phi}(z, x_1) + \frac{1}{\gamma_t} D_{\Phi}(z, x_{t+1}). \end{aligned} \quad (4.9)$$

Plugging this into (4.8) yields

$$\begin{aligned} (4.8) &= \frac{1}{\eta_t} [D_{\Phi}(x_t, w_{t+1}) - D_{\Phi}(z, w_{t+1}) + D_{\Phi}(z, x_t)] \\ &\stackrel{(4.9)}{\leq} \frac{1}{\eta_t} [D_{\Phi}(x_t, w_{t+1}) - D_{\Phi}(x_{t+1}, w_{t+1}) \\ &\quad + \left(\frac{1}{\gamma_t} - 1 \right) D_{\Phi}(z, x_1) - \frac{1}{\gamma_t} D_{\Phi}(z, x_{t+1}) + D_{\Phi}(z, x_t)]. \end{aligned}$$

The claim then follows by the definition of γ_t . \square

Summing the inequality from Claim 4.2 over $t \in [T]$ proves Theorem 4.1. For completeness we show these calculations in Appendix B. \square

4.2. Primal-stabilized OMD

Algorithm 3 gives pseudocode for the primal-stabilized OMD method, which has the following regret bound.

Theorem 4.3 (Regret bound for primal-stabilized OMD).

For all $t \geq 1$, let $\eta : \mathbb{N} \rightarrow \mathbb{R}_{>0}$ be such that $\eta_t \geq \eta_{t+1}$; define $\gamma_t = \eta_{t+1}/\eta_t \in (0, 1]$; and let $\{f_t\}_{t \geq 1}$ be a sequence of convex functions with $f_t : \mathcal{X} \rightarrow \mathbb{R}$. Let $\{x_t\}_{t \geq 1}, \{y_t\}_{t \geq 2}$ and $\{\hat{w}_t\}_{t \geq 2}$ be as in Algorithm 3. Furthermore, assume

$$\text{for all } z \in \mathcal{X}, \text{ the map } D_{\Phi}(z, \cdot) \text{ is convex on } \mathcal{X}. \quad (4.12)$$

Then, for all $T > 0$ and $z \in \mathcal{X}$,

$$\text{Regret}(T, z) \leq \sum_{t=1}^T \frac{D_{\Phi}(\hat{x}_t; w_{t+1})}{\eta_t} + \frac{D_{\Phi}(z, x_1)}{\eta_{T+1}}. \quad (4.13)$$

The proof is identical to the proof of Theorem 4.1, replacing $D_{\Phi}(\hat{x}_t; w_{t+1})$ with $D_{\Phi}(\hat{x}_t; y_{t+1})$ and replacing

Claim 4.2 with the following claim. (The complete proof of Theorem 4.3 can be found in Appendix C.)

Claim 4.4. Assume that $\gamma_t = \eta_{t+1}/\eta_t \in (0, 1]$. Then

$$(4.8) \leq \underbrace{\frac{D_{\Phi}(x_t; w_{t+1})}{\eta_t} + \left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\right) D_{\Phi}(z, x_1)}_{\text{telescopes}} + \underbrace{\frac{D_{\Phi}(z, x_t)}{\eta_t} - \frac{D_{\Phi}(z, x_{t+1})}{\eta_{t+1}}}_{\text{telescopes}}.$$

Proof. First, we derive the inequality

$$\begin{aligned} & \gamma_t D_{\Phi}(y_{t+1}; w_{t+1}) + (1 - \gamma_t) D_{\Phi}(z, x_1) \\ & \stackrel{(i)}{\geq} \gamma_t D_{\Phi}(z, y_{t+1}) + (1 - \gamma_t) D_{\Phi}(z, x_1) \\ & \stackrel{(ii)}{\geq} D_{\Phi}(z, x_{t+1}). \end{aligned}$$

where (i) follows from (4.3) and (4.10), and (ii) is by (4.11), (4.12) and $\gamma_t \in (0, 1]$. Rearranging and using $\gamma_t > 0$ yields

$$\begin{aligned} D_{\Phi}(z, w_{t+1}) & \geq D_{\Phi}(y_{t+1}, w_{t+1}) - \frac{1 - \gamma_t}{\gamma_t} D_{\Phi}(z, x_1) \\ & \quad + \frac{1}{\gamma_t} D_{\Phi}(z, x_{t+1}). \end{aligned} \quad (4.14)$$

Plugging this into (4.8) yields

$$\begin{aligned} (4.8) & = \frac{1}{\eta_t} \left(D_{\Phi}(x_t; w_{t+1}) + D_{\Phi}(z, x_t) \right) \\ & \stackrel{(4.14)}{\leq} \frac{1}{\eta_t} \left(D_{\Phi}(x_t, w_{t+1}) - D_{\Phi}(y_{t+1}, w_{t+1}) \right) \\ & \quad + \left(\frac{1}{\gamma_t} - 1 \right) D_{\Phi}(z, x_1) - \frac{1}{\gamma_t} D_{\Phi}(z, x_{t+1}) + D_{\Phi}(z, x_t). \end{aligned}$$

The claim follows by the definition of γ_t . \square

4.3. Dual averaging

In this section, we show that the DA algorithm can be obtained by a small modification of dual-stabilized online mirror descent. Furthermore our proof of Theorem 4.1 can be adapted to analyze this algorithm.

The main difference between DS-OMD and DA is in the gradient step. In iteration $t + 1$ of DS-OMD the gradient step is taken from \hat{x}_t (the dual counterpart of x_t):

$$\text{DS-OMD gradient step: } \hat{w}_{t+1} := \hat{x}_t - \eta_t \hat{g}_t.$$

Suppose that the algorithm is modified so that the gradient step is taken from \hat{y}_t , the dual point from iteration t before projection onto the feasible region (here define $\hat{y}_1 := \hat{x}_1$). The resulting gradient step is

$$\text{Lazy gradient step: } \hat{w}_{t+1} := \hat{y}_t - \eta_t \hat{g}_t. \quad (4.15)$$

As before, we set

$$\hat{y}_{t+1} := \gamma_t \hat{w}_{t+1} + (1 - \gamma_t) \hat{x}_1,$$

where $\gamma_t = \eta_{t+1}/\eta_t$. Then a simple inductive proof yields the following claim.

Claim 4.5. $\hat{w}_t = \hat{x}_1 - \eta_{t-1} \sum_{i < t} \hat{g}_i$ and $\hat{y}_t = \hat{x}_1 - \eta_t \sum_{i < t} \hat{g}_i$ for all $t > 1$.

Thus, DS-OMD with the lazy gradient step can be written as in Algorithm 1 with the DA update.

Theorem 4.6 (Regret bound for dual averaging). Let $\eta: \mathbb{N} \rightarrow \mathbb{R}_{>0}$ be such that $\eta_t \geq \eta_{t+1}$ for all $t \geq 1$. Define $\gamma_t = \eta_{t+1}/\eta_t \in (0, 1]$ for all $t \geq 1$. Let $\{f_t\}_{t \geq 1}$ be a sequence of convex functions with $f_t: \mathcal{X} \rightarrow \mathbb{R}$ for each $t \geq 1$. Let $\{x_t\}_{t \geq 1}$ and $\{\hat{g}_t\}_{t \geq 1}$ be as in as in Algorithm 1 with DA updates. Then, for all $T > 0$ and $z \in \mathcal{X}$,

$$\begin{aligned} & \text{Regret}(T, z) \\ & \leq \sum_{t=1}^T \frac{D_{\Phi}(x_t; \nabla \Phi^*(\hat{x}_t - \eta_t \hat{g}_t))}{\eta_t} + \frac{D_{\Phi}(z, x_1)}{\eta_{T+1}}. \end{aligned} \quad (4.16)$$

The proof parallels the proof of Theorem 4.1 and can be found in Appendix D.

4.4. Remarks

Interestingly, the *doubling trick* (Shalev-Shwartz, 2012) for OMD can be viewed as an incarnation of stabilization. To see this, set $\eta_t := 1/\sqrt{2^{\lceil \lg t \rceil}}$ and $\gamma_t := \mathbf{1}_{\{t \text{ is a power of } 2\}}$. Then, for each dyadic interval of length 2^ℓ , the first iterate is x_1 and a fixed learning rate $1/\sqrt{2^\ell}$ is used. Thus, with these parameters, Algorithm 2 reduces to the doubling trick.

Note that in Theorem 4.1 the stabilization parameter γ_t used in round $t \geq 1$ depends on the learning rates for rounds t and $t + 1$. Thus, to use stabilization as in Theorem 4.1 the learning rate for round t can depend on information available only up to round $t - 1$. This will be important when we derive first-order regret bounds in Section 5.2.2 where the learning rate depends on the past functions and iterates. Reindexing the learning rates could fix the problem, but then the proof of Theorem 4.1 would look syntactically odd. Although this ‘‘dependence on the future’’ may seem unnatural, in Section 6 we shall see that, under mild conditions, stabilized-OMD coincides with DA with dynamic learning rates. This extends the same behavior observed between OMD and DA when the learning rates are fixed. This may be seen as evidence that stabilization is a natural way to fix OMD for dynamic learning rates. Furthermore, McMahan (2017) shows this off-by-one difference among other algorithms for OCO and discusses the implications of this phenomenon.

5. Applications

In this section we show that stabilized-OMD and DA enjoy the same regret bounds in several applications that involve a dynamic learning rate.

5.1. Strongly-convex mirror maps

We now analyze the algorithms of the previous section in the scenario that the mirror maps are strongly convex. Let η_t, γ_t, f_t be as in the previous section. The next result is a corollary of Theorems 4.1, 4.3, and 4.6.

Corollary 5.1 (Regret bound for dual-stabilized OMD). Suppose that Φ is ρ -strongly convex on \mathcal{X} with respect to a norm $\|\cdot\|$. Let $\{x_t\}_{t \geq 1}$ be the iterates produced by Algorithms 1 with the DA update, or the update rules in Algorithms 2 or 3 (for Algorithm 3, the additional assumption (4.12) is required). Then, for all $T > 0$ and $z \in \mathcal{X}$,

$$\text{Regret}(T, z) \leq \sum_{t=1}^T \frac{\eta_t \|\hat{g}_t\|_*^2}{2\rho} + \frac{D_\Phi(z, x_1)}{\eta_{T+1}}.$$

This is identical to the bound for dual averaging in Nesterov (2009, Eq. 2.15) (taking his $\lambda_i := 1$ and his $\beta_i := 1/\eta_i$).

The proof is based on the following simple proposition, which bounds the Bregman divergence when Φ is strongly convex (Bubeck, 2015, pp. 300). The proof is given in Appendix E.

Proposition 5.2. Suppose Φ is ρ -strongly convex on \mathcal{X} with respect to $\|\cdot\|$. For any $x, x' \in \mathcal{X}$ and $\hat{q} \in \mathbb{R}^n$,

$$D_\Phi\left(\frac{x}{x'}; \nabla\Phi^*(\hat{x} - \hat{q})\right) \leq \|\hat{q}\|_*^2 / 2\rho.$$

Proof (of Corollary 5.1). The regret bounds proven by Theorems 4.1, 4.3 and 4.6 all involve a summation with terms of the form

$$\begin{aligned} 4.1 : D_\Phi\left(\frac{x_t}{x_{t+1}}; w_{t+1}\right) & \quad 4.3 : D_\Phi\left(\frac{x_t}{y_{t+1}}; w_{t+1}\right) \\ 4.6 : D_\Phi\left(\frac{x_t}{x_{t+1}}; \nabla\Phi^*(\hat{x}_t - \eta_t \hat{g}_t)\right). \end{aligned}$$

For Theorems 4.1 and 4.6, we have $x_{t+1} \in \mathcal{X}$, whereas for Theorem 4.3 we have $y_{t+1} \in \mathcal{X}$ by (4.10). For Theorems 4.1 and 4.3 we have $w_{t+1} = \nabla\Phi^*(\hat{x}_t - \eta_t \hat{g}_t)$ by (4.4) and (4.12). Therefore all of these terms may be bounded using Proposition 5.2 with $x = x_t$ and $\hat{q} = \eta_t \hat{g}_t$. This yields the claimed bound. \square

5.2. Prediction with expert advice

Next consider the setting of ‘‘prediction with expert advice’’. In this setting, $\bar{\mathcal{D}}$ is $\mathbb{R}_{>0}^n$, \mathcal{X} is the simplex $\Delta_n \subset \mathbb{R}^n$, and the mirror map is $\Phi(x) := \sum_{i=1}^n x_i \log x_i$. (On \mathcal{X} , Φ is the negative of the entropy function.) The gradient of the mirror map and its conjugate are

$$\nabla\Phi(x)_i = \ln(x_i) + 1 \quad \text{and} \quad \nabla\Phi^*(\hat{x})_i = e^{\hat{x}_i - 1}. \quad (5.1)$$

For any two points $a \in \bar{\mathcal{D}}$ and $b \in \mathcal{D}$, an easy calculation shows that $D_\Phi(a, b)$ is the generalized KL-divergence

$$D_{\text{KL}}(a, b) = \sum_{i=1}^n a_i \ln(a_i/b_i) - \|a\|_1 + \|b\|_1.$$

Note that the KL-divergence is convex in its second argument for any $b \in \mathcal{D} = \mathbb{R}_{>0}^n$ since the functions $-\ln(\cdot)$ and absolute value are both convex. This means that all the abstract regret bounds from Section 4 hold in this setting. Using them we will derive regret bounds for this setting with a little extra-work. As an intermediate step, we will derive bounds that use the following function:

$$\begin{aligned} \Lambda(a, b) & := D_{\text{KL}}(a, b) + \|a\|_1 - \|b\|_1 + \ln\|b\|_1 \\ & = \sum_{i=1}^n a_i \ln(a_i/b_i) + \ln\|b\|_1, \end{aligned}$$

which is a standard tool in the analysis of algorithms for the experts’ problem. For examples, see de Rooij et al. (2014, §2.1) and Cesa-Bianchi et al. (2007, Lemma 4).

The next result is a corollary of Theorems 4.1, 4.3, and 4.6.

Corollary 5.3. For all $t \geq 1$, let $\eta: \mathbb{N} \rightarrow \mathbb{R}_{>0}$ be such that $\eta_t \geq \eta_{t+1}$; define $\gamma_t = \eta_{t+1}/\eta_t \in (0, 1]$; and let $\{f_t\}_{t \geq 1}$ be a sequence of convex functions with $f_t: \mathcal{X} \rightarrow \mathbb{R}$. Let x_1 be the uniform distribution $\vec{1}/n$ and let $\{x_t\}_{t \geq 2}$ and $\{\hat{g}_t\}_{t \geq 1}$ be as in one of Algorithms 1 with the DA update, or the DA update in Algorithms 2 or 3. Then, for all $T > 0$ and $z \in \mathcal{X}$,

$$\text{Regret}(T, z) \leq \sum_{t=1}^T \frac{\Lambda(x_t, \nabla\Phi^*(\hat{x}_t - \eta_t \hat{g}_t))}{\eta_t} + \frac{\ln n}{\eta_{T+1}}. \quad (5.2)$$

The proof is a direct consequence of the following proposition, which is proven in Appendix F.

Proposition 5.4. $D_\Phi\left(\frac{a}{b}; c\right) \leq \Lambda(a, c)$ for $a, b \in \mathcal{X}$, $c \in \mathcal{D}$.

Proof (of Corollary 5.3). First, recall that D_{KL} is convex in its second argument, which allows us to use the bound from (4.13) for primal-stabilized OMD. As in the proof of Corollary 5.1, we first observe that the regret bounds (4.7), (4.13) and (4.16) all have sums with terms of the form $D_\Phi\left(\frac{x_t}{u_t}; \nabla\Phi^*(\hat{x}_t - \eta_t \hat{g}_t)\right)$ for some $u_t \in \mathcal{X}$ that may be bounded using Proposition 5.4. Finally, the standard inequality $\sup_{z \in \mathcal{X}} D_{\text{KL}}(z, x_1) \leq \ln n$ completes the proof. \square

5.2.1. Anytime regret

From Corollary 5.3 we now derive an anytime regret bound in the case of bounded costs. This matches the best known bound appearing in the literature; see Bubeck (2011, Theo-

rem 2.4) and Gerchinovitz (2011, Proposition 2.1). Moreover, in Appendix G we show that this is tight for DA.

Corollary 5.5. Define $\eta_t = 2\sqrt{\ln(n)/t}$ and $\gamma_t = \eta_{t+1}/\eta_t \in (0, 1]$ for all $t \geq 1$. Let $\{f_t := \langle \hat{g}_t, \cdot \rangle\}_{t \geq 1}$ be such that $\hat{g}_t \in [0, 1]^n$ for all $t \geq 1$. Let x_1 be the uniform distribution $\bar{1}/n$ and let $\{x_t\}_{t \geq 2}$ be as in one of Algorithms 1 with the DA update, or the DA update in Algorithms 2 or 3. Then,

$$\text{Regret}(T, z) \leq \sqrt{T \ln n}, \quad \forall T \geq 1, \forall z \in \mathcal{X}.$$

The proof follows from Corollary 5.3 and Hoeffding’s Lemma, as shown below.

Lemma 5.6 (Hoeffding’s Lemma (Cesa-Bianchi & Lugosi, 2006, Lemma 2.2)). Let X be a random variable with $a \leq X \leq b$. Then for any $s \in \mathbb{R}$,

$$\ln \mathbb{E}[e^{sX}] - s\mathbb{E}X \leq \frac{s^2(b-a)^2}{8}.$$

Proof (of Corollary 5.5). By (5.1) we have $\nabla\Phi^*(\hat{x}_t - \eta_t \hat{g}_t)_i = x_t(i) \exp(-\eta_t \hat{g}_t(i))$ for each $i \in [n]$. This together with Lemma 5.6 for $s = -\eta_t$ yields

$$\begin{aligned} & \Lambda(x_t, \nabla\Phi^*(\hat{x}_t - \eta_t \hat{g}_t)) \\ &= \eta_t \langle \hat{g}_t, x_t \rangle + \ln \left(\sum_{i=1}^n x_t(i) e^{-\eta_t \hat{g}_t(i)} \right) \leq \frac{\eta_t^2}{8}. \end{aligned}$$

Plugging this and $\eta_t = 2\sqrt{\frac{\ln n}{t}}$ into (5.2), we obtain

$$\begin{aligned} \text{Regret}(T) &\leq \sqrt{\ln n} \left(\frac{1}{4} \sum_{t=1}^T \frac{1}{\sqrt{t}} + \frac{\sqrt{T+1}}{2} \right) \\ &\leq \sqrt{\ln n} \left(\frac{2\sqrt{T}-1}{4} + \frac{\sqrt{T}+0.5}{2} \right) \\ &\leq \sqrt{T \ln n}, \end{aligned}$$

by Fact A.3 and sub-additivity of square root. \square

5.2.2. First-order regret bound

The regret bound described in Section 5.2.1 depends on \sqrt{T} ; this is known as a “zeroth-order” regret bound. In some scenarios the cost of the best expert up to time T can be far less than T . This makes the problem somewhat easier, and it is possible to improve the regret bound. Formally, let L_T^* denote the total cost of the best expert until time T . Then $L_T^* \leq T$ due to our assumption that all costs are at most 1. A “first-order” regret bound depends on $\sqrt{L_T^*}$ instead of \sqrt{T} .

The only modification to the algorithm is to change the learning rate. If the costs are “smaller than expected”, then intuitively time is progressing “slower than expected”. We will adopt an elegant idea of Auer et al. (2002b), which is to use the algorithm’s cost itself as a measure of the

progression of time, and to incorporate this into the learning rate. They call this a “self-confident” learning rate.

Corollary 5.7. Let $\{f_t := \langle \hat{g}_t, \cdot \rangle\}_{t \geq 1}$ be such that $\hat{g}_t \in [0, 1]^n$ for all $t \geq 1$. Define $\gamma_t = \eta_{t+1}/\eta_t \in (0, 1]$ and $\eta_t = \sqrt{\ln(n)/(1 + \sum_{i < t} \langle \hat{g}_i, x_i \rangle)}$ for all $t \geq 1$. Let x_1 be the uniform distribution $\bar{1}/n$ and let $\{x_t\}_{t \geq 2}$ be as in one of Algorithms 1 with the DA update, 2, or 3. Denote the minimum total cost of any expert up to time T as $L_T^* := \min_{j \in [n]} \sum_{t=1}^T \hat{g}_t(j)$. Then,

$$\text{Regret}(T, z) \leq 2\sqrt{\ln(n)L_T^*} + 8\ln n, \quad \forall T \geq 1, \forall z \in \mathcal{X}.$$

The main ingredient is the following alternative bound on Λ , which is proven in Appendix F.

Proposition 5.8. Let $a \in \mathcal{X}$, $\hat{q} \in [0, 1]^n$ and $\eta > 0$. Then $\Lambda(a, \nabla\Phi^*(\hat{a} - \eta\hat{q})) \leq \eta^2 \langle a, \hat{q} \rangle / 2$.

Proof (of Corollary 5.7). Let $z \in \mathcal{X}$. From Corollary 5.3 and Proposition 5.8, we have

$$\sup_{z' \in \mathcal{X}} \sum_{t=1}^T \langle \hat{g}_t, x_t - z' \rangle \leq \sum_{t=1}^T \frac{\eta_t \langle \hat{g}_t, x_t \rangle}{2} + \frac{\ln n}{\eta_{T+1}}. \quad (5.3)$$

Denote the algorithm’s total cost at time t by $A_t = \sum_{i \leq t} \langle \hat{g}_i, x_i \rangle$. Recall the total cost of the best expert at time T is $L_T^* = \min_{z' \in \Delta_n} \sum_{t=1}^T \langle \hat{g}_t, z' \rangle$ and the learning rate is $\eta_t = \sqrt{\ln(n)/(1 + A_{t-1})}$. Substituting into (5.3),

$$\begin{aligned} A_T - L_T^* &\leq \sqrt{\ln n} \left(\frac{1}{2} \sum_{t=1}^T \frac{\langle \hat{g}_t, x_t \rangle}{\sqrt{1 + A_{t-1}}} + \sqrt{1 + A_T} \right) \\ &\leq \sqrt{\ln n} \left(\sqrt{A_T} + \sqrt{A_T} + 1 \right) \end{aligned}$$

by Proposition A.5 with $a_i = \langle \hat{g}_i, x_i \rangle$ and $u = 1$. Rewriting the previous inequality, we have shown that

$$A_T - L_T^* \leq 2\sqrt{\ln(n)A_T} + \sqrt{\ln n}.$$

By Proposition A.7 we obtain

$$A_T - L_T^* \leq 2\sqrt{\ln(n)L_T^*} + \sqrt{\ln n} + 2(\ln n)^{3/4} + 4\ln n.$$

Since $A_T - L_T^* \geq \text{Regret}(T, z)$, the result follows. \square

Comparing our bound with some existing results in the literature: our constant term of 2 obtained in Corollary 5.7 is better than the constant $(\sqrt{2}/(\sqrt{2}-1))$ obtained by the doubling trick (Cesa-Bianchi & Lugosi, 2006, Exercise 2.8), and the constant $(2\sqrt{2})$ in Auer et al. (2002b) but worse than the constant $(\sqrt{2})$ of the best known first-order regret bound Yaroshinsky et al. (2004), which is obtained by a sophisticated algorithm. We also match the constant 2 of the Hedge algorithm from de Rooij et al. (2014, Theorem 8). Their result is actually more general; we could similarly

generalize our analysis, but that would deviate too far from the main purpose of this paper.

6. Comparing DS-OMD and DA

In this section we shall write the iterates of dual-stabilized OMD in two equivalent forms. First we shall write it in a proximal-like formulation similar to the mirror descent formulation in Beck & Teboulle (2003), shedding some light into the intuition behind dual-stabilization. We then write the iterates from DS-OMD in a form very similar to the original definition of DA in Nesterov (2009). The later will allow us to intuitively understand why OMD does not play well with dynamic step-size and to derive simple sufficient conditions under which DS-OMD and DA generate the same iterates, mimicking the relation between OMD and DA for a fixed learning rate.

Beck & Teboulle (2003) show that the iterate x_{t+1} for round $t + 1$ from OMD is the unique minimizer over \mathcal{X} of $\eta_t \langle \hat{g}_t, \cdot \rangle + D_\Phi(\cdot, x_t)$, where $\hat{g}_t \in \partial f_t(x_t)$. The next proposition extends this formulation to DS-OMD, recovering the result from Beck and Teboulle when $\gamma_t = 1$. The proof, which can be found in Appendix H, is a simple application of the optimality conditions of (6.1).

Proposition 6.1. Let $\{f_t\}_{t \geq 1}$ be a sequence of convex functions with $f_t: \mathcal{X} \rightarrow \mathbb{R}$ for each $t \geq 1$. Let $\eta: \mathbb{N} \rightarrow \mathbb{R}_{>0}$ and $\gamma: \mathbb{N} \rightarrow [0, 1]$. Let $\{x_t\}_{t \geq 1}$ and $\{\hat{g}_t\}_{t \geq 1}$ be as in Algorithm 2. Then, for any $t \geq 1$,

$$\{x_{t+1}\} = \arg \min_{x \in \mathcal{X}} \left(\gamma_t (\eta_t \langle \hat{g}_t, x \rangle + D_\Phi(x, x_t)) + (1 - \gamma_t) D_\Phi(x, x_1) \right). \quad (6.1)$$

In spite of their similar descriptions, Orabona & Pál (2018) showed that OMD and DA may behave in extremely different ways even on the well-studied experts' problem with similar choices of step-sizes. This extreme difference in behavior is not clear from the classical algorithmic description of these methods as in Algorithm 1. In the case of DA, it is well-known that DA can be seen as an instance of the FTRL algorithm; see Bubeck (2015, §4.4) or Hazan (2016, §5.3.1). More specifically, if $\{x_t\}_{t \geq 1}$ and $\{\hat{g}_t\}_{t \geq 1}$ are as in Algorithm 1 with the DA update, then for every $t \geq 0$,¹

$$\{x_{t+1}\} = \arg \min_{x \in \mathcal{X}} \left(\eta_{t+1} \sum_{i=1}^t \langle \hat{g}_i, x \rangle - \langle \hat{x}_1, x \rangle + \Phi(x) \right). \quad (6.2)$$

In the next theorem, proven in Appendix H, we write DS-OMD in a similar form, but with vectors from the normal cone of \mathcal{X} creeping into the formula due to the back and forth between the primal and dual spaces. Recall that the **normal cone** of \mathcal{X} at a point $x \in \mathcal{X}$ is the

set $N_{\mathcal{X}}(x) := \{p \in \mathbb{R}^n : \langle p, z - x \rangle \leq 0 \text{ for all } z \in \mathcal{X}\}$. The result in McMahan (2017, Theorem 11) is similar but slightly more intricate due to the use of time-varying mirror maps. Moreover, this result does not directly apply when we have stabilization.

Theorem 6.2. Let $\{f_t\}_{t \geq 1}$ with $f_t: \mathcal{X} \rightarrow \mathbb{R}$ be a sequence of convex functions and let $\eta: \mathbb{N} \rightarrow \mathbb{R}_{>0}$ be non-increasing. Let $\{x_t\}_{t \geq 1}$ and $\{\hat{g}_t\}_{t \geq 1}$ be as in Algorithm 2. Then, there are $\{p_t\}_{t \geq 1}$ with $p_t \in N_{\mathcal{X}}(x_t)$ for all $t \geq 1$ such that, if $\gamma_i = 1$ for all $i \geq 1$, then for all $t \geq 0$

$$\{x_{t+1}\} = \arg \min_{x \in \mathcal{X}} \left(\sum_{i=1}^t \langle \eta_i \hat{g}_i + p_i, x \rangle - \langle \hat{x}_1, x \rangle + \Phi(x) \right) \quad (6.3)$$

and if $\gamma_i = \frac{\eta_{i+1}}{\eta_i}$ for all $i \geq 1$, then for all $t \geq 0$

$$\{x_{t+1}\} = \arg \min_{x \in \mathcal{X}} \left(\eta_{t+1} \sum_{i=1}^t \langle \hat{g}_i + p'_i, x \rangle - \langle \hat{x}_1, x \rangle + \Phi(x) \right). \quad (6.4)$$

where $p'_t := \frac{1}{\eta_t} p_t \in N_{\mathcal{X}}(x_t)$ for every $t \geq 1$.

With the above theorem, we may compare the iterates of DA, OMD, and DS-OMD by comparing the formulas (6.2), (6.3), and (6.4). For the simple unconstrained case where $\mathcal{X} = \mathbb{R}^n$ we have $N_{\mathcal{X}}(x_t) = \{0\}$ for each $t \geq 1$ and DA and DS-OMD are identical. However, if the learning rate is not constant, OMD is *not* equivalent to the latter methods. In particular, if $\eta_t \propto \frac{1}{\sqrt{t}}$, (6.3) shows that the subgradients of the earlier-seen functions have a bigger weight on the iterates if compared to the subgradients of functions from later rounds. In other words, OMD may be sensitive to the ordering of the functions, and adversarial orderings may affect its performance.

When \mathcal{X} is an arbitrary convex set, DA and DS-OMD are not necessarily equivalent anymore due to the vectors from the normal cone of \mathcal{X} . If we know that the iterates live in the relative interior of \mathcal{X} , the next lemma (whose proof we give in Appendix H) shows that these vectors do not affect the set of minimizers from (6.4).

Lemma 6.3. For any $\hat{x} \in \text{ri } \mathcal{X}$ we have $N_{\mathcal{X}}(\hat{x}) = (-N_{\mathcal{X}}(\hat{x}) \cap N_{\mathcal{X}}(\hat{x}))$. In particular, for any $p \in N_{\mathcal{X}}(\hat{x})$ we have $\langle p, x \rangle = \langle p, \hat{x} \rangle$ for every $x \in \mathcal{X}$.

With this lemma, we can easily derive simple and intuitive conditions under which DS-OMD and DA are equivalent.

Corollary 6.4. Let $\mathcal{D} \subseteq \mathbb{R}^n$ be the interior of the domain of Φ , let $\{x_t\}_{t \geq 1}$ be the DS-OMD iterates as in Algorithm 2 and let $\{x'_t\}_{t \geq 1}$ be the DA iterates as in Algorithm 1 with DA updates. If $\mathcal{D} \cap \mathcal{X} \subseteq \text{ri } \mathcal{X}$, then $x_t = x'_t$ for each $t \geq 1$.

Proof. Let $t \geq 1$. Since $x_t = \Pi_{\mathcal{X}}^{\Phi}(y_t)$, where y_t is as in Algorithm 2, Lemma H.3 implies $x_t \in \mathcal{D} \cap \mathcal{X} \subseteq \text{ri } \mathcal{X}$. By

¹The $\langle \nabla \Phi(x_1), x \rangle$ term disappears if x_1 minimizes Φ on \mathcal{X} .

Lemma 6.3 we have that the vectors on the normal cone in (6.4) do not affect the set of minimizers, which implies that (6.2) and (6.4) are equivalent. \square

An important special case of the above corollary is the prediction with expert advice setting as in Section 5.2, where $\mathcal{D} = \mathbb{R}_{>0}^n$ and \mathcal{X} is the simplex Δ_n . In this setting, $\mathcal{X} \cap \mathcal{D} = \{x \in (0, 1)^d : \sum_{i=1}^n x_i = 1\} = \text{ri } \mathcal{X}$. By the previous corollary DS-OMD and DA produce the same iterates in this case even for dynamic learning rates. Classical OMD and DA were already known to be equivalent in the experts' setting for a *fixed* learning rate (Hazan, 2016, §5.4.2). In contrast, with a dynamic learning rate, the DA and OMD iterates are certainly different, since OMD with a dynamic learning rate may have linear regret (Orabona & Pál, 2018), whereas DA has sublinear regret.

7. Discussion

In this paper we modified OMD via *stabilization* in order to guarantee sublinear regret even when using the method with a dynamic learning-rate. We showed that (primal and dual) stabilized-OMD recover the regret bounds enjoyed by DA in the anytime setting, presented some applications of our results, and analyzed the similarities and differences between DS-OMD, OMD, and DA.

Our bounds for the problem of prediction with expert advice nearly match the current state-of-the-art. A distinctive feature of our proofs are their relative simplicity if compared to other results from the literature. It is our hope that the simplicity of our analysis framework allows it to be extended to other problems. Moreover, the modularity of our proofs allowed us to extend this analysis for DA, a fact interesting on its own since drastically different analysis techniques are usually used to analyze DA in the literature (such as the Follow the Leader-Be the Leader Lemma and optimality conditions of (6.2), see (Shalev-Shwartz, 2012, Section 2.3) for an example). This together with our analysis from Section 6 helps demystify the connections between DA and OMD, since in spite of having similar descriptions they had extremely different analyses and behaved wildly differently in some scenarios. We believe that a better understanding between the differences between DA and OMD will be helpful in future applications and in the design of new algorithms.

8. Acknowledgements

We thank Chris Liaw for pointing out a slight flaw in the proofs in an earlier draft of this paper. We also thank Francesco Orabona for suggesting the use of a slightly different definition of regret which allows for more nuanced statements of our results. We also express our gratitude for the detailed feedback given by the three anonymous

reviewers from ICML 2020.

References

- Allen-Zhu, Z. and Orecchia, L. Linear coupling: An ultimate unification of gradient and mirror descent. November 2016. URL <https://arxiv.org/abs/1407.1537>.
- Audibert, J.-Y., Bubeck, S., and Lugosi, G. Regret in online combinatorial optimization. *Mathematics of Operations Research*, 39(1):31–45, 2014.
- Auer, P., Cesa-Bianchi, N., Freund, Y., and Schapire, R. E. The nonstochastic multiarmed bandit problem. *SIAM Journal on Computing*, 32(1), 2002a.
- Auer, P., Cesa-Bianchi, N., and Gentile, C. Adaptive and self-confident on-line learning algorithms. *Journal of Computer and System Sciences*, 64(1):48–75, 2002b.
- Beck, A. *First-order methods in optimization*, volume 25 of *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2017.
- Beck, A. and Teboulle, M. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 31(3):167–175, 2003.
- Bubeck, S. Introduction to online optimization, December 2011. unpublished.
- Bubeck, S. Convex optimization: Algorithms and complexity. *Foundations and Trends in Machine Learning*, 8(3-4): 231–357, 2015.
- Cesa-Bianchi, N. and Lugosi, G. *Prediction, learning, and games*. Cambridge University Press, 2006.
- Cesa-Bianchi, N., Freund, Y., Haussler, D., Helmbold, D. P., Schapire, R. E., and Warmuth, M. K. How to use expert advice. *Journal of the ACM*, 44(3), May 1997.
- Cesa-Bianchi, N., Mansour, Y., and Stoltz, G. Improved second-order bounds for prediction with expert advice. *Machine Learning*, 66(2-3):321–352, 2007.
- de Rooij, S., van Erven, T., Grünwald, P. D., and Koolen, W. M. Follow the leader if you can, hedge if you must. *Journal of Machine Learning Research (JMLR)*, 15:1281–1316, 2014.
- Duchi, J. C., Shalev-shwartz, S., Singer, Y., and Tewari, A. Composite objective mirror descent. In *Proceedings of COLT*, pp. 14–26, 2010.

- Gerchinovitz, S. *Prediction of individual sequences and prediction in the statistical framework: some links around sparse regression and aggregation techniques*. PhD thesis, Université Paris-Sud, 2011.
- Hazan, E. Introduction to online convex optimization. *Foundations and Trends in Optimization*, 2(3-4):157–325, 2016.
- Joulani, P., György, A., and Szepesvári, C. A modular analysis of adaptive (non-)convex optimization: Optimism, composite objectives, and variational bounds. In *Proceedings of ALT'17*, pp. 681–720, 2017.
- Juditsky, A. B., Kwon, J., and Moulines, E. Unifying mirror descent and dual averaging. October 2019. URL <http://arxiv.org/abs/1910.13742>.
- Kakade, S. M., Shalev-Shwartz, S., and Tewari, A. Regularization techniques for learning with matrices. *Journal of Machine Learning Research (JMLR)*, 13:1865–1890, 2012.
- Kangarshahi, E. A., Hsieh, Y., Sahin, M. F., and Cevher, V. Let's be honest: An optimal no-regret framework for zero-sum games. In *Proceedings of ICML'18*, pp. 2493–2501, 2018.
- McMahan, H. B. A survey of algorithms and analysis for adaptive online learning. *Journal of Machine Learning Research*, 18:90:1–90:50, 2017.
- Nemirovski, A. and Yudin, D. *Problem Complexity and Method Efficiency in Optimization*. Wiley Interscience, 1983.
- Nesterov, Y. Primal-dual subgradient methods for convex problems. *Mathematical Programming*, 120(1):221–259, 2009.
- Orabona, F. and Pál, D. Scale-free online learning. *Theor. Comput. Sci.*, 716:50–69, 2018.
- Rockafellar, R. T. *Convex Analysis*. Princeton University Press, Princeton, 1970.
- Shalev-Shwartz, S. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2):107–194, 2012.
- Xiao, L. Dual averaging methods for regularized stochastic learning and online optimization. *Journal of Machine Learning Research*, 11, 2010.
- Yaroshinsky, R., El-Yaniv, R., and Seiden, S. S. How to better use expert advice. *Machine Learning*, 55(3):271–309, 2004.