Journal of Machine Learning Research 1 (2000) 1-48 Submitted 4/00; Published 10/00

Online Mirror Descent and Dual Averaging: Keeping Pace in the Dynamic Case

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
Department of Computer Science
University of British Columbia
Vancouver, BC V6T 1Z3, Canada

Editor:

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. Finally, we show how to effectively use dual-stabilization with composite cost functions with simple adaptations to both the algorithm and its analysis.

Keywords: online learning; mirror descent; dual averaging; stabilization; unknown time horizon.

1. Introduction

Online convex optimization (OCO) lies in 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 and each decision incurs a cost given by a convex function which is only revealed to the player after they makes 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 the best decision in hindsight. In this setting, algorithms for the player that attain sublinear regret (usually under mild conditions on the problem) with respect to the total number of rounds/decisions T are considered desirable.

Online mirror descent (OMD) and dual averaging (DA) are two important algorithm templates for OCO from which many classical online learning algorithms can be described as special cases (see Shalev-Shwartz, 2012 and McMahan, 2017 for some examples). When the total number T of decisions to be made is known in advance, OMD and DA achieve exactly the same regret bound when using the same constant learning rate (Hazan, 2016). 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 adaptive learning rate — while DA can guarantee sublinear regret bound $O(\sqrt{T})$ for any $T > 0$ (Nesterov, 2009), there are instances on which OMD suffers asymptotically linear $\Omega(T)$ regret (Orabona and Pál, 2018).

The aim of this paper is to introduce a *stabilization* technique that bridges the gap between OMD and DA with dynamic learning rates. We begin by giving almost identical abstract regret bounds (depending only on the Bregman divergence between iterates) for stabilized OMD and DA.

We provide simple and clean proofs showing that OMD with stabilization works similarly to DA in the following aspects:

- With the same adaptive learning rate, OMD with stabilization achieves exactly the same regret bound as DA. This mirrors the situation in which the time horizon is known in advance and both OMD and DA use a constant learning rate.

- For the problem of prediction with expert advice, OMD with stabilization matches the best known regret bound (even with the same constant and with a dynamic learning rate), which was originally achieved by DA (see Bubeck, 2011, Section 2.5 and Gerchinovitz, 2011, Proposition 2.1). In our technical report (Fang et al., 2021), we show that this regret bound is the best that can be achieved with learning rates of the form $c/\sqrt{t}$.

- For the problem of prediction with expert advice, we give a concise proof that OMD with stabilization can achieve a first-order regret bound. This means that the regret bound does not depend on T but rather on the cost of the best expert up to time T, which is no larger. Our analysis matches the best known analysis for DA (Bubeck, 2015; Gerchinovitz, 2011), though it is worse than the best known constant in the literature (Yaroshinsky et al., 2004).

- We formally compare the iterates generated by DA and OMD (with and without stabilization). This sheds light on the reasons why OMD behaves badly with dynamic learning rates. As a corollary of this comparison, we get simple sufficient conditions for the iterates of DA and (dual-)stabilized OMD to match. This mimics the behavior between DA and OMD when the learning rate is fixed.

- For composite functions, we show that a proximal variant of dual-stabilized OMD again achieves exactly the same regret bound as regularized DA (Xiao, 2010).

2. Related work

The concept of mirror descent originates from Nemirovski and Yudin (1983) and renewed interest in OMD began with a modern treatment given by Beck and Teboulle (2003). It then received a lot of attention from the optimization community due to the recent interest in first-order methods for large-scale problems. For example, see the works from Duchi et al. (2010); Allen-Zhu and Orecchia (2016) and for further details and references see the work of Beck (2017). DA is due to Nesterov (2009) and is motivated by the dissatisfying fact that the convergence of classical subgradient descent methods rely on decaying step sizes which ultimately give less weight to new information. This algorithm was later extended to regularized problems by Xiao (2010). DA is closely related to the *follow-the-regularized-leader* (FTRL) algorithm (Shalev-Shwartz, 2012). See also the works from Bubeck (2015), Hazan (2016), and McMahan (2017) for complete descriptions and further references. OMD and DA also grew a lot in popularity due to applications in online learning problems (Kakade et al., 2012; Audibert et al., 2014), and the fact that they generalize a wide range of online learning algorithms (Shalev-Shwartz, 2012; McMahan, 2017). Moreover, OMD and DA have been used to tackle problems in theoretical computer science such as graph sparsification (Allen-Zhu et al., 2015) and the k-server problem (Bubeck et al., 2018).

Unifying views of online learning algorithms have been shown to be useful for applications and have drawn recent attention. McMahan (2017) showed how to use adaptive regularization in the FTRL framework to derive many online learning algorithms, and compared the different ways OMD and FTRL deal with composite functions. Joulani et al. (2017) proposed a unified framework to analyze online learning algorithms under wildly different assumptions, extending even to the non-convex case. Juditsky et al. (2019) recently proposed a unified framework called unified mirror descent (UMD) that encompasses OMD and DA as special cases. In spite of these unifying frameworks, the differences between OMD and DA seemed to be overlooked and one might imagine that the algorithms had similar performance in all settings.

Only recently Orabona and Pál (2018) looked more closely at the difference between OMD and DA with time-varying learning rates. They presented counter examples to demonstrate that OMD with a dynamic learning rate could suffer from linear regret even under well-studied settings such as in the experts' problem, where the algorithm picks points in the simplex and the adversary picks linear functions whose gradients have ℓ_∞-norm at most 1. Although this may seem to contradict the well-known $O(\sqrt{T})$ regret bounds for OMD, it does not. These sub-linear bounds hold if the algorithm knows the time-horizon from the start or if the Bregman divergence (with respect to the mirror map) on the feasible set is bounded. However, the Kullback-Leibler divergence is *not* bounded on the simplex. In this paper we explain this phenomenon and show how the addition of stabilization to OMD fixes this problem.

For the problem of prediction with expert advice, by means of the *doubling trick*, Cesa-Bianchi et al. (1997) show an algorithm with a sublinear *anytime* regret bound, meaning a bound that holds at each round of the game. Improved anytime regret bounds were developed by Auer et al. (2002b), with a simplified description of the latter given by Cesa-Bianchi and Lugosi (2006, Section 2.3). Sublinear anytime regret bounds can also be derived from the original work on dual averaging in Nesterov (2009). Regret bounds that depend on the cost of the best expert (known as the first-order regret bound) can be traced back to the work by Cesa-Bianchi et al. (1997). Improved first-order regret bounds were given by Auer et al. (2002b) and the current best known first-order regret bound 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 $t \in \{1, 2, \ldots\}$ 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 which the algorithm has access to by a first order oracle, that is, for any $x \in \mathcal{X}$ the algorithm can compute $f_t(x)$ and a subgradient $\hat{g} \in \partial f_t(x) \coloneqq \{\hat{g} \in \mathbb{R}^n \mid f(z) \geq f(x) + \langle\, \hat{g},\, z - x \,\rangle \;\forall z \in \mathcal{X}\}$. We will assume[1] that all cost functions f_t in this text are subdifferentiable on $\mathcal{X}$, that is, meaning that the subdifferential $\partial f(x)$ is non-empty for all $x \in \mathcal{X}$.

The cost of the iteration at time t is defined as $f_t(x_t)$. In this setting the goal is to produce a sequence of proposals $\{x_t\}_{t \geq 1}$ that minimizes the *regret* against an unknown comparison point $z \in \mathcal{X}$ that has accrued up until time T:

$$\mathrm{Regret}(T, z) \;\coloneqq\; \sum_{t=1}^{T} f_t(x_t) - \sum_{t=1}^{T} f_t(z).$$

In this paper we are interested in the case where the algorithm does not know the time-horizon T in advance. This implies that our choices of parameters, including learning rates, cannot depend on T.

Often our results and proofs will make use of the *dual norm* of $\|\cdot\|$, defined by

$$\|z\|_* = \sup\left\{\langle z, x\rangle \mid x \in \mathbb{R}^n, \|x\| \leq 1\right\}.$$

Both dual averaging and online mirror descent are parameterized by a special convex function Φ, often referred as a regularizer or a mirror map (for $\mathcal{D}$ and $\mathcal{X}$), which among other properties needs[2] to be of Legendre type (Rockafellar, 1970, Chapter 26). Formally, throughout the paper we assume that the function $\Phi\colon \bar{\mathcal{D}} \to \mathbb{R}$ is a closed convex function such that $\operatorname{int} \bar{\mathcal{D}} \cap \operatorname{ri} \mathcal{X} \neq \varnothing$ (where $\operatorname{ri} \mathcal{X}$ denotes the relative interior of $\mathcal{X}$), and whose conjugate is differentiable on $\mathbb{R}^n$. Moreover, we also suppose that Φ is of Legendre type, which means that Φ is strictly convex on its domain[3] and essentially smooth, that is, for $\mathcal{D} \coloneqq \operatorname{int} \bar{\mathcal{D}}$ we have

- $\mathcal{D}$ is nonempty,
- Φ is differentiable on $\mathcal{D}$, and
- $\lim_{x \to \partial\mathcal{D}} \|\nabla \Phi(x)\| = +\infty$, where $\partial\mathcal{D}$ is the boundary of $\mathcal{D}$, i.e., $\partial\mathcal{D} \coloneqq \operatorname{cl} \mathcal{D} \setminus \mathcal{D}$.

The gradient of the mirror map $\nabla\Phi : \mathcal{D} \to \mathbb{R}^n$ and the gradient of its conjugate $\nabla\Phi^* : \mathbb{R}^n \to \mathcal{D}$ are mutually inverse bijections between the primal space $\mathcal{D}$ and the dual space $\mathbb{R}^n$. We will adopt the following notational convention. Any vector in the primal space

1. This holds, for example, if f is finite and convex on an open superset of $\mathcal{X}$ (Rockafellar, 1970, Theorem 23.4).
2. One may relax this condition in some cases. Bubeck (2011, § 5.2) discusses in depth the conditions needed on the mirror map.
3. In fact we only need Φ to be strictly convex on some convex subsets of the domain (Rockafellar, 1970, Chapter 26), but for the sake of simplicity we assume that Φ is strictly convex on its entire domain.

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:

$$\hat{x} \coloneqq \nabla\Phi(x) \qquad \text{and} \qquad x \coloneqq \nabla\Phi^*(\hat{x}) \qquad \text{for all letters } x. \tag{1}$$

Essential smoothness ensures not only that Φ is differentiable on the interior of its domain, but also that the slope of Φ increases arbitrarily fast near the boundary of its domain. The latter guarantees, at least intuitively, that the function is increasing near and in the direction of the boundary of its domain. This property is fundamental for mirror descent to be well-defined (although not essential for dual averaging) since it ensures that the Bregman projection onto $\mathcal{X}$ is attained by a point on $\mathcal{D}$ where Φ is differentiable (we give more details on this on Section 7), and uniqueness is a consequence of the strict convexity of Φ. Some mirror maps we shall look at are classical cases of the OCO literature such as the negative entropy $x \in \mathbb{R}^n_+ \mapsto \sum_{i=1}^n x_i \ln x_i$ and the squared 2-norm $\|\cdot\|_2^2$, and details on the reasons they are mirror maps can be found in the works of Shalev-Shwartz (2012); Bubeck (2011) and Bubeck (2015) (in particular, Bubeck, 2011, Section 5.2 discusses the properties of functions of Legendre type and why requiring the conjugate of the mirror map to be differentiable on the whole space is not necessary for mirror descent to be well-defined if one restricts the gradient steps in the dual space in some way).

Given a mirror map Φ, the Bregman divergence with respect to Φ is defined as

$$D_\Phi(x, y) \;\coloneqq\; \Phi(x) - \Phi(y) - \langle\, \nabla\Phi(y),\, x - y\,\rangle, \qquad \forall x \in \bar{\mathcal{D}}, \forall y \in \mathcal{D}. \tag{2}$$

Throughout this paper it will be convenient to use the notation

$$D_\Phi\big(\begin{smallmatrix} a \\ b \end{smallmatrix}; c\big) \;\coloneqq\; D_\Phi(a, c) - D_\Phi(b, c) \;=\; \Phi(a) - \Phi(b) - \langle\, \nabla\Phi(c),\, a - b\,\rangle. \tag{3}$$

In the important special case where $\Phi(x) = \frac{1}{2}\|x\|_2^2$, the Bregman divergence relates to the Euclidean distance, i.e., $D_\Phi(x, y) = \frac{1}{2}\|x - y\|_2^2$. When $\Phi(x) = \sum_{i=1}^n x_i \log x_i$, the Bregman divergence becomes the generalized Kullback-Leibler (KL) divergence. The projection operator induced by the Bregman divergence is $\Pi_{\mathcal{X}}^{\Phi}$ given by $\{\Pi_{\mathcal{X}}^{\Phi}\}(y) \coloneqq \arg\min\{D_\Phi(x, y) \mid x \in \mathcal{X}\}$ for any $y \in \mathcal{X} \cap \mathcal{D}$.

A general template for optimization in the mirror descent framework is shown in Algorithm 1. The two classical algorithms, online mirror descent and dual averaging, are incarnations of this, differing only in how the dual variable $\hat{y}_t$ is updated.

In this work, for a given initial point $x_1 \in \mathcal{X}$ of the player, we are interested in the case when $\sup_{z \in \mathcal{X}} D_\Phi(z, x_1)$ is bounded, which still allows $\sup_{z,x \in \mathcal{X}} D_\Phi(z, x)$ to be unbounded. In fact, in the Euclidean setting (i.e., $\Phi = \frac{1}{2}\|\cdot\|_2^2$), $\sup_{z \in \mathcal{X}} D_\Phi(z, x_1)$ is bounded if and only if the diameter of $\mathcal{X}$ given by $\sup_{x,y \in \mathcal{X}} \frac{1}{2}\|x - y\|_2^2$ is also bounded. However, for a general mirror map Φ, assuming $\sup_{z \in \mathcal{X}} D_\Phi(z, x_1)$ is bounded is strictly weaker than assuming $\sup_{x,y \in \mathcal{X}} D_\Phi(x, y)$ is bounded. This is the case for the well-known experts problem, where Φ is the negative entropy, D_Φ is the KL-divergence, $\mathcal{X}$ is the unit simplex, and $x_1 \coloneqq \frac{1}{n}\vec{1}$, where $\vec{1}$ is the vector in $\mathbb{R}^n$ with entries all set to 1. In this case, we have $\sup_{z \in \mathcal{X}} D_\Phi(z, x_1) \leq \ln n$ while $\sup_{z,x \in \mathcal{X}} D_\Phi(z, x) = +\infty$.

To conclude, throughout the paper we shall try to stick to the following naming convention: greek letters will denote scalars, lower-case roman letters will denote vectors, and capital caligraphic letters shall denote sets. Any deviations from this convention (or the hat notation as in (1)) are intended to follow other conventions in the literature.

Algorithm 1 Pseudocode for both online mirror descent and dual averaging with adaptive learning rate given by η_t on iteration t. These methods differ only in how the iterate $\hat{y}_{t+1}$ is updated.

Input: $x_1 \in \mathcal{X} \cap \mathcal{D}, \eta : \mathbb{N} \to \mathbb{R}_{>0}$.
for $t = 1, 2, \ldots$ **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**] $\hat{y}_{t+1} = \hat{x}_1 - \eta_t \sum_{i \le t} \hat{g}_i$
  $y_{t+1} = \nabla\Phi^*(\hat{y}_{t+1})$
  $x_{t+1} = \Pi_{\mathcal{X}}^{\Phi}(y_{t+1})$
end for

4. The relationship between OMD and DA

In this section, we present a detailed review of some known properties of OMD and DA. Our goal is to summarize known similarities and differences between the guarantees on the regret for these algorithms in the fixed-time and anytime settings.

4.1 OMD and DA with constant learning rate

When the time horizon T is known in advance, a constant learning rate that depends on T can be adopted in many algorithms for OCO to achieve sublinear regret. In particular, OMD and DA with the same fixed learning rate enjoy exactly the same regret bound.

Theorem 1 (Nesterov, 2009, Thm. 1, Hazan, 2016, Thm. 5.6) *Suppose that Φ is ρ-strongly convex with respect to a norm $\|\cdot\|$ and pick a constant learning rate $\eta_t \coloneqq \eta > 0$ for all $t \ge 1$. Let $\{x_t\}_{t\ge1}$ be the sequence of iterates generated by Algorithm 1. Then for any sequence of convex functions $\{f_t\}_{t\ge1}$ with $f_t \colon \mathcal{X} \to \mathbb{R}$ for each $t \ge 1$, the following bound holds for both OMD and DA updates,*

$$\operatorname{Regret}(T, z) \;\le\; \sum_{t=1}^{T} \frac{\eta \|\hat{g}_t\|_*^2}{2\rho} + \frac{D_\Phi(z, x_1)}{\eta}, \tag{4}$$

for any comparison point $z \in \mathcal{X}$.

Interestingly, though OMD and DA with constant learning rate have similar regret bounds, the proofs used to derive these bound tend to be quite different.

4.2 OMD and DA with dynamic learning rate

In the unknown time horizon scenario, a dynamic learning rate with $\eta_t \propto 1/\sqrt{t}$ is usually adopted in the literature of online learning (Beck and Teboulle, 2003; Zinkevich, 2003). Moreover, when the Bregman divergence (with respect to Φ) on the domain $\mathcal{X}$ is bounded, both OMD and DA with learning rate $\eta_t \propto 1/\sqrt{t}$ can achieve $O(\sqrt{T})$ regret bounds (with differing constants). However, when the Bregman divergence on $\mathcal{X}$ is unbounded, OMD is provably worse than DA as the next theorem shows.

Theorem 2 (Linear regret for OMD, Orabona and Pál, 2018, Thm. 3) *Set $\eta_t := 1/\sqrt{t}$ for each $t \geq 1$. Let $\{x_t\}_{t\geq 1}$ denote the sequence of iterates generated by Algorithm 1 with OMD update. There exists a sequence of convex 1-Lipschitz continuous functions $\{f_t\}_{t=1}^T$ and an initial point $x_1 \in \mathcal{X}$ such that*

$$D_\Phi(z, x_1) \ \ \textit{is bounded} \quad \textit{and} \quad \mathrm{Regret}(T, z) \ = \ \Omega(T).$$

In contrast, Algorithm 1 with the DA update can always guarantee sublinear regret bound $O(\sqrt{T})$ using a similar learning rates (which differ only by constants).

Moreover, folklore examples show that for offline 1-dimensional gradient descent (i.e., mirror descent with Euclidean regularization), a learning rate of either $o(1/\sqrt{t})$ or $\omega(1/\sqrt{t})$ cannot achieve regret $O(\sqrt{t})$ for all $t > 0$. Therefore OMD with learning rates of the form $t^{-\alpha}$ with $\alpha > 0$ may not have optimal regret guarantees when the Bregman divergence on $\mathcal{X}$ is unbounded. A natural question is if we can improve OMD to make it provably work with dynamic learning rates. In the next section we provide a fix for adaptive OMD through stabilization and later we show its connection to adaptive DA.

5. Stabilized OMD

As shown in Theorem 2, Orabona and Pál (2018) proved 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 chose functions so that picking z in 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 iterate computed by taking a gradient step). An illustration for dual stabilization is shown in Figure 1.

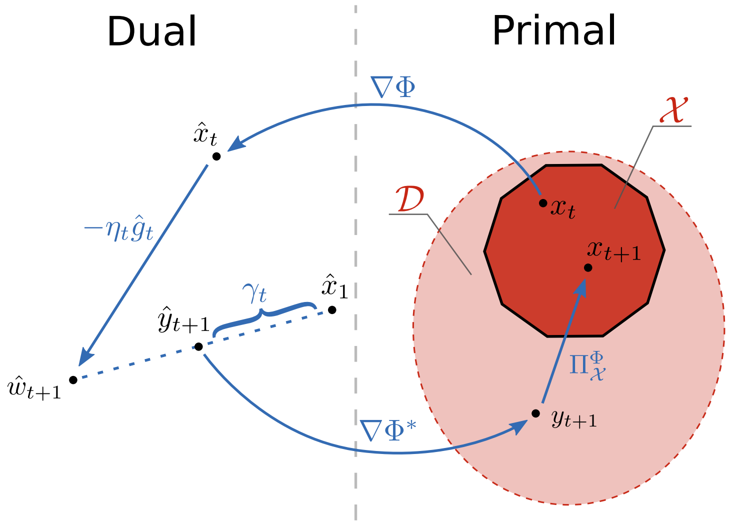

Figure 1: Illustration of the t-th iteration of DS-OMD.

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

Input: $x_1 \in \mathcal{X}, \eta : \mathbb{N} \to \mathbb{R}_+,\ \gamma : \mathbb{N} \to (0, 1]$

for $t = 1, 2, \ldots$ **do**

Incur cost $f_t(x_t)$ and receive $\hat{g}_t \in \partial f_t(x_t)$

$\hat{x}_t = \nabla\Phi(x_t)$ ▷ map primal iterate to dual space

$$\hat{w}_{t+1} = \hat{x}_t - \eta_t\hat{g}_t \qquad \triangleright \text{ gradient step in dual space} \tag{5}$$

$$\hat{y}_{t+1} = \gamma_t\hat{w}_{t+1} + (1 - \gamma_t)\hat{x}_1 \qquad \triangleright \text{ stabilization in dual space} \tag{6}$$

$y_{t+1} = \nabla\Phi^*(\hat{y}_{t+1})$ ▷ map dual iterate to primal space

$$x_{t+1} = \Pi^{\Phi}_{\mathcal{X}}(y_{t+1}) \qquad \triangleright \text{ project onto feasible region} \tag{7}$$

end for

After early versions of the paper were available, we were informed that ideas similar to primal stabilization had appeared in the Robust Optimistic Mirror Descent algorithm (Kangarshahi et al., 2018) and in the Twisted Mirror Descent(TMD) algorithm (György and Szepesvári, 2016). The setting of the former is different since they perform optimistic steps and their results are somewhat weaker in terms of constant factors and since they cannot handle Bregman projections. In the latter case, TMD is a meta-algorithm that has primal stabilization as a special case. But the authors only use primal stabilization for the experts' problem (György and Szepesvári, 2016, Example 6). In our work we extend the idea of primal stabilization for cases beyond the setting of prediction with expert advice.

5.1 Dual-stabilized OMD

Algorithm 2 gives pseudocode showing our modification of OMD to incorporate dual stabilization. Theorem 3 analyzes it without assuming strong convexity of Φ.

Theorem 3 (Regret bound for dual-stabilized OMD) *Assume we have $\eta_t \geq \eta_{t+1} > 0$ for each $t > 1$. 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 $f_t \colon \mathcal{X} \to \mathbb{R}$ 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 $z \in \mathcal{X}$,*

$$\operatorname{Regret}(T, z) \;\leq\; \sum_{t=1}^{T} \frac{D_\Phi\big(\genfrac{}{}{0pt}{}{x_t}{x_{t+1}}; w_{t+1}\big)}{\eta_t} + \frac{D_\Phi(z, x_1)}{\eta_{T+1}}. \tag{8}$$

Note that strong convexity of Φ is *not* assumed. As we will see in Section 6.1, the term $D_\Phi\big(\genfrac{}{}{0pt}{}{x_t}{x_{t+1}}; w_{t+1}\big)$ can be easily bounded by the dual norm of $\hat{g}_t$ 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 3]

Let $z \in \mathcal{X}$. The first step is the same as in the standard OMD proof.

$$\begin{aligned}
f_t(x_t) - f_t(z) &\leq \langle\, \hat{g}_t,\, x_t - z \,\rangle && \text{(subgradient ineq.)} \\
&= \frac{1}{\eta_t}\langle\, \hat{x}_t - \hat{w}_{t+1},\, x_t - z \,\rangle && \text{(by (5))} \\
&= \frac{1}{\eta_t}\big(D_\Phi(x_t, w_{t+1}) - D_\Phi(z, w_{t+1}) + D_\Phi(z, x_t)\big) && \text{(by Prop. 36).} \quad (9)
\end{aligned}$$

The next step exhibits the main point of stabilization. Without stabilization we would have $x_{t+1} = \Pi^\Phi_{\mathcal{X}}(w_{t+1})$ and $D_\Phi(z, w_{t+1}) \geq D_\Phi(z, x_{t+1}) + D_\Phi(x_{t+1}, w_{t+1})$ by Proposition 38, so (9) 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 need a claim that leads to telescoping terms by relating $D_\Phi(z, w_{t+1})$ to $D_\Phi(z, x_{t+1})$.

Claim 4 *Assume that $\gamma_t = \eta_{t+1}/\eta_t \in (0,1]$. Then*

$$(9) \;\leq\; \frac{D_\Phi\big(\genfrac{}{}{0pt}{}{x_t}{x_{t+1}}; w_{t+1}\big)}{\eta_t} + \underbrace{\Big(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\Big)}_{\textit{telescopes}} D_\Phi(z, x_1) + \underbrace{\frac{D_\Phi(z, x_t)}{\eta_t} - \frac{D_\Phi(z, x_{t+1})}{\eta_{t+1}}}_{\textit{telescopes}}.$$

Proof First we derive the inequality

$$\begin{aligned}
&\gamma_t\big(D_\Phi(z, w_{t+1}) - D_\Phi(x_{t+1}, w_{t+1})\big) \;+\; (1-\gamma_t) D_\Phi(z, x_1) \\
\geq\;& \gamma_t D_\Phi\big(\genfrac{}{}{0pt}{}{z}{x_{t+1}}; w_{t+1}\big) \;+\; (1-\gamma_t) D_\Phi\big(\genfrac{}{}{0pt}{}{z}{x_{t+1}}; x_1\big) && \text{(since } D_\Phi(x_{t+1}, x_1) \geq 0 \text{ and } \gamma_t \leq 1) \\
=\;& D_\Phi\big(\genfrac{}{}{0pt}{}{z}{x_{t+1}}; y_{t+1}\big) && \text{(by Proposition 37 and (6))} \\
\geq\;& D_\Phi(z, x_{t+1}) && \text{(by Proposition 38 and (7)).}
\end{aligned}$$

Rearranging and using $\gamma_t > 0$ yields

$$D_\Phi(z, w_{t+1}) \;\geq\; D_\Phi(x_{t+1}, w_{t+1}) - \Big(\frac{1}{\gamma_t} - 1\Big) D_\Phi(z, x_1) + \frac{1}{\gamma_t} D_\Phi(z, x_{t+1}). \tag{10}$$

Plugging this into (9) yields

$$\begin{aligned}
(9) \;&=\; \frac{1}{\eta_t}\Big(D_\Phi(x_t, w_{t+1}) - D_\Phi(z, w_{t+1}) + D_\Phi(z, x_t)\Big) \\
&\leq\; \frac{1}{\eta_t}\Bigg(D_\Phi(x_t, w_{t+1}) - D_\Phi(x_{t+1}, w_{t+1}) + \\
&\qquad \Big(\frac{1}{\gamma_t} - 1\Big)D_\Phi(z, x_1) - \frac{1}{\gamma_t}D_\Phi(z, x_{t+1}) + D_\Phi(z, x_t)\Bigg). \qquad \text{(by (10))}
\end{aligned}$$

The claim follows by the definition of γ_t. ■

The final step is very similar to the standard OMD proof. Summing (9) over t and using Claim 4 leads to the desired telescoping sum.

$$\begin{aligned}
&\sum_{t=1}^{T}\big(f_t(x_t) - f_t(z)\big) \\
\leq\;& \sum_{t=1}^{T}\left(\frac{D_\Phi(\substack{x_t \\ x_{t+1}}; w_{t+1})}{\eta_t} + \Big(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\Big)D_\Phi(z, x_1) + \frac{D_\Phi(z, x_t)}{\eta_t} - \frac{D_\Phi(z, x_{t+1})}{\eta_{t+1}}\right) \\
\leq\;& \sum_{t=1}^{T}\frac{D_\Phi(\substack{x_t \\ x_{t+1}}; w_{t+1})}{\eta_t} \;+\; \left(\frac{1}{\eta_1} + \sum_{t=1}^{T}\left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\right)\right) D_\Phi(z, x_1) \\
=\;& \sum_{t=1}^{T}\frac{D_\Phi(\substack{x_t \\ x_{t+1}}; w_{t+1})}{\eta_t} \;+\; \frac{D_\Phi(z, x_1)}{\eta_{T+1}}.
\end{aligned}$$

■

5.2 Primal-stabilized OMD

Algorithm 3 gives pseudocode showing our modification of OMD to incorporate primal stabilization. Interestingly, as we shall see in the next theorem, we need convexity of $D_\Phi(z, \cdot)$ for $z \in \mathcal{X}$ to get interesting bounds on the regret when using primal stabilization.

Algorithm 3 Online mirror descent with primal stabilization.

Input: $x_1 \in \mathbb{R}^n$, $\eta : \mathbb{N} \to \mathbb{R}$, $\gamma : \mathbb{N} \to \mathbb{R}$.
for $t = 1, 2, \ldots$ **do**
  Incur cost $f_t(x_t)$ and receive $\hat{g}_t \in \partial f_t(x_t)$
  $\hat{x}_t = \nabla\Phi(x_t)$ ▷ map primal iterate to dual space
  $\hat{w}_{t+1} = \hat{x}_t - \eta_t \hat{g}_t$ ▷ gradient step in dual space (11)
  $w_{t+1} = \nabla\Phi^*(\hat{w}_{t+1})$ ▷ map dual iterate to primal space (12)
  $y_{t+1} = \Pi^{\Phi}_{\mathcal{X}}(w_{t+1})$ ▷ project onto feasible region (13)
  $x_{t+1} = \gamma_t y_{t+1} + (1-\gamma_t)x_1$ ▷ stabilization in primal space (14)
end for

Theorem 5 (Regret bound for primal-stabilized OMD) *Assume $\eta_t \geq \eta_{t+1} > 0$ for each $t > 1$. Define $\gamma_t = \eta_{t+1}/\eta_t \in (0,1]$ for all $t \geq 1$. Let $\{x_t\}_{t\geq 1}$ be the sequence of iterates generated by Algorithm 3. Furthermore, assume that*

$$\textit{for all } z \in \mathcal{X}, \textit{ the map } x \mapsto D_\Phi(z,x) \textit{ is convex on } \mathcal{X}. \tag{15}$$

Then for any sequence of convex functions $\{f_t\}_{t\geq 1}$ with each $f_t : \mathcal{X} \to \mathbb{R}$,

$$\mathrm{Regret}(T,z) \;\leq\; \sum_{t=1}^{T} \frac{D_\Phi\big({x_t \atop y_{t+1}}; w_{t+1}\big)}{\eta_t} + \frac{D_\Phi(z,x_1)}{\eta_{T+1}} \quad \forall T > 0. \tag{16}$$

Proof [of Theorem 5]

Let $z \in \mathcal{X}$. The first step is identical to the proof of Theorem 3 since the update rule in (11) is exactly the same as (5). Therefore, we have that (9) holds, that is,

$$f_t(x_t) - f_t(z) \;\leq\; \frac{1}{\eta_t}\big(D_\Phi(x_t, w_{t+1}) - D_\Phi(z, w_{t+1}) + D_\Phi(z, x_t)\big).$$

Now instead of Claim 4 we use the next claim, which is similar to Claim 4 but replaces $D_\Phi\big({x_t \atop x_{t+1}}; w_{t+1}\big)$ with $D_\Phi\big({x_t \atop y_{t+1}}; w_{t+1}\big)$.

Claim 6 *Assume that $\gamma_t = \eta_{t+1}/\eta_t \in (0,1]$. Then*

$$(9) \;\leq\; \frac{D_\Phi\big({x_t \atop y_{t+1}}; w_{t+1}\big)}{\eta_t} + \underbrace{\Big(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\Big)}_{\textit{telescopes}} D_\Phi(z,x_1) + \underbrace{\frac{D_\Phi(z,x_t)}{\eta_t} - \frac{D_\Phi(z,x_{t+1})}{\eta_{t+1}}}_{\textit{telescopes}}.$$

Proof First, we derive the inequality

$$\begin{aligned}
&\gamma_t\big(D_\Phi(z,w_{t+1}) - D_\Phi(y_{t+1},w_{t+1})\big) \;+\; (1-\gamma_t)D_\Phi(z,x_1) \\
=\;& \gamma_t D_\Phi\big({z \atop y_{t+1}}; w_{t+1}\big) \;+\; (1-\gamma_t)D_\Phi(z,x_1) \\
\geq\;& \gamma_t D_\Phi(z,y_{t+1}) + (1-\gamma_t)D_\Phi(z,x_1) && \text{(by Prop. 38 and (13))} \\
\geq\;& D_\Phi(z,x_{t+1}) && \text{(by (14) and (15))}.
\end{aligned}$$

Rearranging and using $\gamma_t > 0$ yields

$$D_\Phi(z, w_{t+1}) \;\geq\; D_\Phi(y_{t+1}, w_{t+1}) - \Big(\frac{1}{\gamma_t} - 1\Big) D_\Phi(z, x_1) + \frac{1}{\gamma_t} D_\Phi(z, x_{t+1}). \tag{17}$$

Plugging this into (9) yields

$$\begin{aligned}
(9) \;&=\; \frac{1}{\eta_t}\Big(D_\Phi(x_t, w_{t+1}) - D_\Phi(z, w_{t+1}) + D_\Phi(z, x_t)\Big) \\
&\leq\; \frac{1}{\eta_t}\Bigg(D_\Phi(x_t, w_{t+1}) - D_\Phi(y_{t+1}, w_{t+1}) + \\
&\qquad \Big(\frac{1}{\gamma_t} - 1\Big) D_\Phi(z, x_1) - \frac{1}{\gamma_t} D_\Phi(z, x_{t+1}) + D_\Phi(z, x_t)\Bigg). \qquad \text{(by (17))}
\end{aligned}$$

The claim follows by the definition of γ_t. ■

The final step is very similar to the proof of Theorem 3. The only difference is that we are using Claim 6 instead of Claim 4 and we replace $D_\Phi(\substack{x_t \\ x_{t+1}}; w_{t+1})$ with $D_\Phi(\substack{x_t \\ y_{t+1}}; w_{t+1})$. Formally, summing (9) over t and using Claim 6 leads to the desired telescoping sum, that is,

$$\begin{aligned}
&\sum_{t=1}^{T} \big(f_t(x_t) - f_t(z)\big) \\
\leq\;& \sum_{t=1}^{T} \left(\frac{D_\Phi(\substack{x_t \\ y_{t+1}}; w_{t+1})}{\eta_t} + \Big(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\Big) D_\Phi(z, x_1) + \frac{D_\Phi(z, x_t)}{\eta_t} - \frac{D_\Phi(z, x_{t+1})}{\eta_{t+1}} \right) \\
\leq\;& \sum_{t=1}^{T} \frac{D_\Phi(\substack{x_t \\ y_{t+1}}; w_{t+1})}{\eta_t} \;+\; \left(\frac{1}{\eta_1} + \sum_{t=1}^{T} \left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t} \right) \right) D_\Phi(z, x_1) \\
=\;& \sum_{t=1}^{T} \frac{D_\Phi(\substack{x_t \\ y_{t+1}}; w_{t+1})}{\eta_t} \;+\; \frac{D_\Phi(z, x_1)}{\eta_{T+1}}.
\end{aligned}$$

■

5.3 Dual averaging

In this section, we show that Nesterov's dual averaging algorithm can be obtained from a small modification to dual-stabilized online mirror descent. Furthermore our proof of Theorem 3 can be adapted to analyze DA.

The main difference between DS-OMD and dual averaging is in the gradient step, as we now explain. In iteration t of DS-OMD, the gradient step in (5) is taken from $\hat{x}_t$, the dual counterpart of the iterate x_t:

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

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 $\hat{y}_1$ is defined to be $\hat{x}_1$.) The resulting gradient step is:

$$\text{Lazy gradient step:} \quad \hat{w}_{t+1} \;=\; \hat{y}_t - \eta_t \hat{g}_t. \tag{18}$$

As before, we set

$$\hat{y}_{t+1} \;=\; \gamma_t \hat{w}_{t+1} + (1-\gamma_t)\hat{x}_1 \tag{19}$$

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

Claim 7 *$\hat{w}_t = \hat{x}_1 - \eta_{t-1}\sum_{i<t}\hat{g}_i$ and $\hat{y}_t = \hat{x}_1 - \eta_t \sum_{i<t}\hat{g}_i$ for all $t > 1$.*

Thus, the algorithm with the lazy gradient step can be written as in Algorithm 4. This is equivalent to Algorithm 1 with the DA update, except that η_t in Algorithm 1 corresponds to η_{t+1} in Algorithm 4. In the next theorem we analyze DA with similar techniques to the ones used in Theorems 3 and 5.

Algorithm 4 Dual averaging with learning rate re-indexed as $\eta_2, \eta_3, \ldots$

```
Input: x_1 ∈ 𝒳, η : ℕ → ℝ_+, γ : ℕ → (0,1]
ŷ_1 = ∇Φ(x_1)
for t = 1, 2, ... do
    Incur cost f_t(x_t) and receive ĝ_t ∈ ∂f_t(x_t)
    ŷ_{t+1} = x̂_1 − η_{t+1} Σ_{i≤t} ĝ_i       ▷ dual averaging update
    y_{t+1} = ∇Φ*(ŷ_{t+1})                  ▷ map dual iterate to primal space
    x_{t+1} = Π^Φ_𝒳(y_{t+1})                 ▷ project onto feasible region
end for
```

Theorem 8 (Regret bound for dual averaging) *Assume we have $\eta_t \geq \eta_{t+1} > 0$ for each $t > 1$. Let $\{x_t\}_{t\geq 1}$ be the sequence of iterates generated by Algorithm 4. Then for any sequence of convex functions $\{f_t\}_{t\geq 1}$ with each $f_t : \mathcal{X} \to \mathbb{R}$,*

$$\text{Regret}(T, z) \;\leq\; \sum_{t=1}^{T} \frac{D_\Phi\big(\substack{x_t \\ x_{t+1}}; \nabla\Phi^*(\hat{x}_t - \eta_t \hat{g}_t)\big)}{\eta_t} + \frac{D_\Phi(z, x_1)}{\eta_{T+1}} \quad \forall T > 0. \tag{20}$$

Proof [of Theorem 8]

Let $z \in \mathcal{X}$. The first step is very similar to the proof of Theorem 3.

$$\begin{aligned}
f_t(x_t) - f_t(z) \;&\leq\; \langle\, \hat{g}_t,\, x_t - z \,\rangle && \text{(subgradient ineq.)} \\
&=\; \frac{1}{\eta_t} \langle\, \hat{y}_t - \hat{w}_{t+1},\, x_t - z \,\rangle && \text{(by (18))} \\
&=\; \frac{1}{\eta_t}\Big(D_\Phi(x_t, w_{t+1}) - D_\Phi(z, w_{t+1}) + D_\Phi\big(\substack{z \\ x_t}; y_t\big)\Big), && (21)
\end{aligned}$$

where in the last equation we have used Proposition 35 instead of Proposition 36.

As in the proof of Theorem 3, the next step is to relate $D_\Phi(z, w_{t+1})$ to $D_\Phi(z, y_{t+1})$ so that (21) can be bounded using a telescoping sum. The following claim is similar to Claim 4.

Claim 9 *Assume that* $\gamma_t = \eta_{t+1}/\eta_t \in (0, 1]$. *Then*

$$(21) \;\leq\; \frac{D_\Phi\big(\genfrac{}{}{0pt}{}{x_t}{x_{t+1}}; w_{t+1}\big)}{\eta_t} + \underbrace{\Big(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\Big)}_{telescopes} D_\Phi(z, x_1) + \underbrace{\frac{D_\Phi\big(\genfrac{}{}{0pt}{}{z}{x_t}; y_t\big)}{\eta_t} - \frac{D_\Phi\big(\genfrac{}{}{0pt}{}{z}{x_{t+1}}; y_{t+1}\big)}{\eta_{t+1}}}_{telescopes}.$$

Proof The first two steps are identical to the proof of Claim 4.

$$\begin{aligned}
&\gamma_t\big(D_\Phi(z, w_{t+1}) - D_\Phi(x_{t+1}, w_{t+1})\big) \;+\; (1-\gamma_t) D_\Phi(z, x_1) \\
\geq\;& \gamma_t D_\Phi\big(\genfrac{}{}{0pt}{}{z}{x_{t+1}}; w_{t+1}\big) \;+\; (1-\gamma_t) D_\Phi\big(\genfrac{}{}{0pt}{}{z}{x_{t+1}}; x_1\big) && \text{(since } D_\Phi(x_{t+1}, x_1) \geq 0 \text{ and } \gamma_t \leq 1) \\
=\;& D_\Phi\big(\genfrac{}{}{0pt}{}{z}{x_{t+1}}; y_{t+1}\big) && \text{(by Proposition 37 and (19)).}
\end{aligned}$$

Rearranging and using $\gamma_t > 0$ yields

$$D_\Phi(z, w_{t+1}) \;\geq\; D_\Phi(x_{t+1}, w_{t+1}) \;-\; \Big(\frac{1}{\gamma_t} - 1\Big) D_\Phi(z, x_1) \;+\; \frac{D_\Phi\big(\genfrac{}{}{0pt}{}{z}{x_{t+1}}; y_{t+1}\big)}{\gamma_t}. \tag{22}$$

Plugging this into (21) yields

$$\begin{aligned}
(21) \;=\;& \frac{1}{\eta_t}\Big(D_\Phi(x_t, w_{t+1}) - D_\Phi(z, w_{t+1}) + D_\Phi\big(\genfrac{}{}{0pt}{}{z}{x_t}; y_t\big)\Big) \\
\leq\;& \frac{1}{\eta_t}\Bigg(D_\Phi(x_t, w_{t+1}) - D_\Phi(x_{t+1}, w_{t+1}) + \\
& \Big(\frac{1}{\gamma_t} - 1\Big) D_\Phi(z, x_1) - \frac{D_\Phi\big(\genfrac{}{}{0pt}{}{z}{x_{t+1}}; y_{t+1}\big)}{\gamma_t} + D_\Phi\big(\genfrac{}{}{0pt}{}{z}{x_t}; y_t\big)\Bigg), && \text{by (22).}
\end{aligned}$$

The claim follows by the definition of γ_t. ■

Again the final step is very similar to the proof of Theorem 3. Summing (21) over t and using Claim 9 leads to the desired telescoping sum.

$$\begin{aligned}
& \sum_{t=1}^{T}\big(f_t(x_t) - f_t(z)\big) \\
\leq\;& \sum_{t=1}^{T}\Bigg(\frac{D_\Phi\big(\genfrac{}{}{0pt}{}{x_t}{x_{t+1}}; w_{t+1}\big)}{\eta_t} + \Big(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\Big) D_\Phi(z, x_1) + \frac{D_\Phi\big(\genfrac{}{}{0pt}{}{z}{x_t}; y_t\big)}{\eta_t} - \frac{D_\Phi\big(\genfrac{}{}{0pt}{}{z}{x_{t+1}}; y_{t+1}\big)}{\eta_{t+1}}\Bigg) \\
\leq\;& \sum_{t=1}^{T}\frac{D_\Phi\big(\genfrac{}{}{0pt}{}{x_t}{x_{t+1}}; w_{t+1}\big)}{\eta_t} \;+\; \left(\frac{1}{\eta_1} + \sum_{t=1}^{T}\Big(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\Big)\right) D_\Phi(z, x_1) \\
=\;& \sum_{t=1}^{T}\frac{D_\Phi\big(\genfrac{}{}{0pt}{}{x_t}{x_{t+1}}; w_{t+1}\big)}{\eta_t} \;+\; \frac{D_\Phi(z, x_1)}{\eta_{T+1}},
\end{aligned}$$

where for the second inequality we have also used that $D_\Phi(\substack{z\\x_1}; y_1) = D_\Phi(z, x_1)$ since $x_1 = y_1$. Thus, the above shows that

$$\mathrm{Regret}(T, z) \;\leq\; \sum_{t=1}^{T} \frac{D_\Phi(\substack{x_t\\x_{t+1}}; w_{t+1})}{\eta_t} + \frac{D_\Phi(z, x_1)}{\eta_{T+1}} \quad \forall T > 0. \tag{23}$$

Notice that (23) is syntactically identical to (8); the only difference is the definition of w_{t+1} in these two settings. However, in this section the definition of x_t has not yet been used! Doing so will provide an upper bound on (23), which is the conclusion of this theorem. To control $D_\Phi(\substack{x_t\\x_{t+1}}; w_{t+1})$, we will apply Proposition 39 as follows. Taking $p = y_t$, $\pi = x_t = \Pi_{\mathcal{X}}^{\Phi}(y_t)$, $v = x_{t+1}$ and $\hat{q} = \eta_t g_t$, we obtain

$$\begin{aligned}
D_\Phi(\substack{x_t\\x_{t+1}}; w_{t+1}) &= -D_\Phi(\substack{v\\\pi}; \nabla\Phi^*(\hat{p} - \hat{q})) && (\text{since } \hat{w}_{t+1} = \hat{y}_t - \eta_t \hat{g}_t = \hat{p} - \hat{q}) \\
&\leq -D_\Phi(\substack{v\\\pi}; \nabla\Phi^*(\hat{\pi} - \hat{q})) && (\text{by Proposition 39}) \\
&= D_\Phi(\substack{x_t\\x_{t+1}}; \nabla\Phi^*(\hat{x}_t - \eta_t \hat{g}_t)).
\end{aligned}$$

Plugging this into (23) completes the proof. ■

5.4 Remarks

Interestingly, the *doubling trick* (Shalev-Shwartz, 2012) on OMD can be viewed as an incarnation of stabilization. To see this, set $\eta_t := 1/\sqrt{2^{\lfloor \lg t \rfloor}}$ 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.

One should note that in Theorem 3 the stabilization parameter γ_t used in round $t \geq 1$ depends on the learning rates η_t and η_{t+1}, the latter of which is used in the *next* round. So the learning rate η_{t+1} must be known in iteration t in order to calculate γ_t appropriately. This subtlety will arise, for example, when we derive first-order regret bounds in Section 6.2.2 — here the learning rate is based on the subgradients of the past functions, not just on the iteration number. Reindexing the learning rates could fix the problem, but then the proof of Theorem 3 would look syntactically odd. Although this dependence on the future may seem unnatural, in Section 7 we shall see that under some mild conditions, stabilized OMD coincides exactly with DA with dynamic learning rates after reindexing. This matches the behavior observed between OMD and DA when the learning rates are fixed. In this sense, stabilization may seem as a natural way to fix OMD for dynamic learning rates.

6. 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.

6.1 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 3, 5, and 8.

Corollary 10 (Regret bound for stabilized OMD and DA) *Suppose that the mirror map 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 one of Algorithms 2, 3, or 4 (for Algorithm 3, the additional assumption* (15) *is required). Then, for all $T > 0$ and $z \in \mathcal{X}$,*

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

This is identical to the bound for dual averaging in Nesterov (2009, Eq. 2.15) (taking his $\lambda_i \coloneqq 1$ and his $\beta_i \coloneqq 1/\eta_i$). The proof is based on the following simple proposition, which bounds the Bregman divergence when Φ is strongly convex (Bubeck, 2015, pp. 300).

Proposition 11 *Suppose Φ is ρ-strongly convex on $\mathcal{X}$ with respect to $\|\cdot\|$. For any $x, x' \in \mathcal{X}$ and $\hat{q} \in \mathbb{R}^n$,*

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

Proof First we apply Proposition 36 with $a = x$, $b = x'$ and $d = \nabla\Phi^*(\hat{x} - \hat{q})$ to obtain

$$\begin{aligned}
D_\Phi\big(\tbinom{x}{x'}; d\big) &= \langle\, \hat{x} - \hat{d},\, x - x' \,\rangle - D_\Phi(x', x) \\
&= \langle\, \hat{q},\, x - x' \,\rangle - D_\Phi(x', x) && \text{(since } \hat{d} = \hat{x} - \hat{q}\text{)} \\
&\leq \|\hat{q}\|_* \|x - x'\| - \frac{\rho}{2}\|x - x'\|^2 && \text{(by the def. of dual norm and Prop. 33)} \\
&\leq \|\hat{q}\|_*^2 / 2\rho && \text{(by Fact 26).}
\end{aligned}$$

■

Now the proof of Corollary 10 is a simple application of the above proposition to the abstract regret bounds we have for each algorithm.

Proof [of Corollary 10] The regret bounds proven by Theorems 3, 5 and 8 all involve a summation with terms of a similar form.

$$\begin{aligned}
&\text{Theorem 3:} && D_\Phi\big(\tbinom{x_t}{x_{t+1}}; w_{t+1}\big) \\
&\text{Theorem 5:} && D_\Phi\big(\tbinom{x_t}{y_{t+1}}; w_{t+1}\big) \\
&\text{Theorem 8:} && D_\Phi\big(\tbinom{x_t}{x_{t+1}}; \nabla\Phi^*(\hat{x}_t - \eta_t \hat{g}_t)\big)
\end{aligned}$$

We may bound all three using Proposition 11. In all three cases we have $x_t, x_{t+1} \in \mathcal{X}$. In Theorem 5 we additionally have $y_{t+1} \in \mathcal{X}$. For Theorems 3 and 5 we have $w_{t+1} = \nabla\Phi^*(\hat{x}_t - \eta_t \hat{g}_t)$ by (5) and (11). Therefore all of these terms may be bounded using Proposition 11 with $x = x_t$ and $\hat{q} = \eta_t \hat{g}_t$, yielding the claimed bound. ■

6.2 Prediction with expert advice

Next consider the setting of prediction with expert advice. In this setting $\mathcal{X}$ is the simplex $\Delta_n \subset \mathbb{R}^n$, the mirror map is $\Phi(x) \coloneqq \sum_{i=1}^n x_i \log x_i$ for all $x \in \bar{\mathcal{D}} \coloneqq \mathbb{R}^n_{\geq 0}$ (taking $0 \ln 0 = 0$). Note that on $\mathcal{X}$ the mirror map Φ is the negative of the entropy function. The gradient of the mirror map and its conjugate are

$$\nabla\Phi(x)_i = \ln(x_i) + 1 \quad \text{and} \quad \nabla\Phi^*(\hat{x})_i = e^{\hat{x}_i - 1}, \qquad \forall x \in \mathcal{D}, \forall \hat{x} \in \mathbb{R}^n. \tag{24}$$

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

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

Note that the KL-divergence is convex in its second argument for any $b \in \mathcal{D} = \mathbb{R}^n_{>0}$ since the functions $-\ln(\cdot)$ and absolute value are both convex. This means that all the abstract regret bounds from Section 5 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:

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

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 3, 5, and 8.

Corollary 12 *Assume we have $\eta_t \geq \eta_{t+1} > 0$ for each $t > 1$. Define $\gamma_t \coloneqq \eta_{t+1}/\eta_t \in (0,1]$ for all $t \geq 1$. Let $x_1 \coloneqq \frac{1}{n}\vec{1}$ be the uniform distribution and $\{x_t\}_{t\geq 2}$ be the iterates produced by one of Algorithms 2, 3, or 4 in the setting of prediction with expert advice. Then, for all $T > 0$ and $z \in \mathcal{X}$,*

$$\mathrm{Regret}(T,z) \;\leq\; \sum_{t=1}^T \frac{\Lambda(x_t, \nabla\Phi^*(\hat{x}_t - \eta_t \hat{g}_t))}{\eta_t} + \frac{\ln n}{\eta_{T+1}}. \tag{25}$$

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

Proposition 13 $D_\Phi(\genfrac{}{}{0pt}{}{a}{b}; c) \leq \Lambda(a,c)$ *for* $a, b \in \mathcal{X}$, $c \in \mathcal{D}$.

Proof [of Corollary 12] Recall that D_{KL} is convex in its second argument, which allows us to use the bound (16) for primal-stabilized OMD. As in the proof of Corollary 10, we first observe that the regret bounds (8), (16) and (20) all have sums with terms of the form $D_\Phi(\genfrac{}{}{0pt}{}{x_t}{u_t}; \nabla\Phi^*(\hat{x}_t - \eta_t\hat{g}_t))$ for some irrelevant $u_t \in \mathcal{X}$, and hence may be bounded using Proposition 13. Finally, the standard inequality $\sup_{z\in\mathcal{X}} D_{\mathrm{KL}}(z, x_1) \leq \ln n$ completes the proof. ■

6.2.1 Anytime regret

From Corollary 12 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, Theorem 2.4) and Gerchinovitz (2011, Proposition 2.1). Moreover, in our technical report (Fang et al., 2021) we show that this is tight for DA in the case $n = 2$. By the equivalence of DA and DS-OMD in the experts' setting (Corollary 22 in Section 7), this regret bound is also tight for DS-OMD.

Corollary 14 *Define* $\eta_t = 2\sqrt{\ln(n)/t}$ *and* $\gamma_t = \eta_{t+1}/\eta_t \in (0, 1]$ *for each* $t \geq 1$. *Let* $\{f_t \coloneqq \langle \hat{g}_t, \cdot \rangle\}_{t \geq 1}$ *be such that* $\hat{g}_t \in [0, 1]^n$ *for all* $t \geq 1$. *Let* x_1 *be the uniform distribution* $\frac{1}{n}\vec{1}$ *and let* $\{x_t\}_{t \geq 2}$ *be as in one of Algorithms 2, 3, or 4 in the prediction with experts advice setting. Then,*

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

The proof follows from Corollary 12 and Hoeffding's Lemma, as shown below.

Lemma 15 (Hoeffding's Lemma, Cesa-Bianchi and 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 14] By (24) 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 15 for $s = -\eta_t$ yields

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

Plugging this and $\eta_t = 2\sqrt{\frac{\ln n}{t}}$ into (25), we obtain

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

by Fact 28 and concavity of square root. ∎

6.2.2 First-order regret bound

The regret bound described in Section 6.2.1 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 from 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 called this a "self-confident" learning rate.

Corollary 16 *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 $\frac{1}{n}\vec{1}$ and let $\{x_t\}_{t \geq 2}$ be as in one of Algorithms 2, 3, or 4 in the prediction with experts advice setting. 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 B.

Proposition 17 *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 16] Let $z \in \mathcal{X}$. From Corollary 12 and Proposition 17, we have

$$\sup_{z' \in \mathcal{X}} \sum_{t=1}^T \langle \hat{g}_t, x_t - z' \rangle \;\leq\; \sum_{t=1}^T \frac{\eta_t \langle \hat{g}_t, x_t \rangle}{2} + \frac{\ln n}{\eta_{T+1}}. \tag{26}$$

Denote the algorithm's total cost at time t by $A_t = \sum_{i \leq t} \langle \hat{g}_i, x_i \rangle$. Recall that 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 (26),

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

by Proposition 30 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 31 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. ∎

Comparing our bound with some existing results in the literature: our constant term of 2 obtained in Corollary 16 is better than the constant ($\sqrt{2}/(\sqrt{2} - 1)$) obtained by the doubling trick (Cesa-Bianchi and 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 due to Yaroshinsky et al. (2004). 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.

7. Comparing DS-OMD and DA

In this section we will write the iterates of dual-stabilized OMD in two equivalent forms. First we will write it in a proximal-like formulation similar to the mirror descent formulation by Beck and 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 by Nesterov (2009). This will allow us to intuitively understand why OMD does has poor results 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 and Teboulle (2003) showed 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 is a simple application of optimality conditions of (27). (For the sake of completeness we carefully state classical results on optimality conditions in Appendix C). Recall that the **normal cone** to a set $C \subseteq \mathbb{R}^n$ at $x \in \mathbb{R}^n$ is the set $N_C(x) \coloneqq \{ s \in \mathbb{R}^n \mid \langle s, y - x \rangle \leq 0 \ \forall y \in C \}$.

Proposition 18 *Let $\{f_t\}_{t \geq 1}$ be a sequence of convex functions with $f_t \colon \mathcal{X} \to \mathbb{R}$ for each $t \geq 1$. Assume we have $\eta_t \geq \eta_{t+1} > 0$ and $\gamma_t \in [0,1]$ for each $t \geq 1$. Let $\{x_t\}_{t \geq 1}$ and $\{\hat{g}_t\}_{t \geq 1}$ be as in Algorithm 2. Then, for any $t \geq 1$,*

$$\{x_{t+1}\} = \operatorname*{arg\,min}_{x \in \mathcal{X}} \Big(\gamma_t \big(\eta_t \langle \hat{g}_t, x \rangle + D_\Phi(x, x_t) \big) + (1 - \gamma_t) D_\Phi(x, x_1) \Big). \tag{27}$$

Proof Let $t \geq 1$ and let $F_t \colon \mathcal{D} \to \mathbb{R}$ be the function being minimized on the right-hand side of (27). By definition we have $x_{t+1} = \Pi^\Phi_{\mathcal{X}}(y_{t+1})$. By the optimality conditions for Bregman projections (see Lemma 42 in Appendix C),

$$x_{t+1} = \Pi^\Phi_{\mathcal{X}}(y_{t+1}) \iff \hat{y}_{t+1} - \hat{x}_{t+1} = -\nabla(D_\Phi(\cdot, y_{t+1}))(x_{t+1}) \in N_{\mathcal{X}}(x_{t+1}).$$

Referring to (2) we see that $\nabla(D_\Phi(\cdot, b))(a) = \hat{a} - \hat{b}$ for any $a, b \in \mathcal{D}$. Using this and the definitions from Algorithm 2 we get

$$\begin{aligned}
\hat{y}_{t+1} - \hat{x}_{t+1} &= \gamma_t(\hat{x}_t - \eta_t \hat{g}_t) + (1 - \gamma_t)\hat{x}_1 - \hat{x}_{t+1} \\
&= \gamma_t(\hat{x}_t - \hat{x}_{t+1} - \eta_t \hat{g}_t) + (1 - \gamma_t)(\hat{x}_1 - \hat{x}_{t+1}) \\
&= -\gamma_t \big(\nabla(D_\Phi(\cdot, x_t))(x_{t+1}) + \eta_t \hat{g}_t \big) - (1 - \gamma_t) \nabla(D_\Phi(\cdot, x_1))(x_{t+1}) \\
&= -\nabla F_t(x_{t+1}).
\end{aligned}$$

Thus, we have $-\nabla F_t(x_{t+1}) \in N_{\mathcal{X}}(x_{t+1})$. Again by classical optimality conditions for convex optimization we conclude that $x_{t+1} \in \operatorname{arg\,min}_{x \in \mathcal{X}} F_t(x)$, and strict convexity of Φ yields uniqueness of the minimizer, as desired. ■

As expected, when $\gamma_t = 1$ for each $t \geq 1$ the above theorem recovers exactly the OMD formulation from Beck and Teboulle (2003). Thus, the above result provides intuition to understand the effect of the stabilization step on the iterates of the algorithm: it biases the iterates toward points in $\mathcal{X}$ which are not too far (w.r.t. the Bregman Divergence) from x_1.

Despite their similar descriptions, Orabona and 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.

It is also 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 4, then for every $t \geq 0$, we have[4]

$$\{x_{t+1}\} = \arg\min_{x \in \mathcal{X}} \Big(\eta_{t+1} \sum_{i=1}^{t} \langle \hat{g}_i, x \rangle - \langle \hat{x}_1, x \rangle + \Phi(x) \Big). \tag{DA-Prox}$$

In the next theorem, proven in Appendix C, we write DS-OMD in a similar form, but with vectors from the normal cone of $\mathcal{X}$ creeping into the formula due to repeatedly mapping between the primal and dual spaces. The result by McMahan (2017, Theorem 11) is similar but slightly more intricate due to the use of time-varying mirror maps. Moreover, their result does not directly apply when we have stabilization.

Theorem 19 *Let $\{f_t\}_{t \geq 1}$ with $f_t : \mathcal{X} \to \mathbb{R}$ be a sequence of convex functions and let $\eta \colon \mathbb{N} \to \mathbb{R}_{>0}$ 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 $t \geq 0$*

$$\{x_{t+1}\} = \arg\min_{x \in \mathcal{X}} \Big(\sum_{i=1}^{t} \langle \eta_i \hat{g}_i + p_i, x \rangle - \langle \hat{x}_1, x \rangle + \Phi(x) \Big) \tag{OMD-Prox}$$

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

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

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

Theorem 19 is an easy consequence of the following proposition.

Proposition 20 *Let $\{f_t\}_{t \geq 1}$ with $f_t \colon \mathcal{X} \to \mathbb{R}$ be a sequence of convex functions and let $\eta \colon \mathbb{N} \to \mathbb{R}_{>0}$ be non-increasing. Let $\{x_t\}_{t \geq 1}$ and $\{\hat{g}_t\}_{t \geq 1}$ be as in Algorithm 2. Define $\gamma^{[i,t]} \coloneqq \prod_{j=i}^{t} \gamma_j$ for every $i, t \in \mathbb{N}$ with the convention $\prod_{j=i}^{t} \gamma_j = 1$ for $t < i$. Then, there are $\{p_t\}_{t \geq 1}$ with $p_t \in N_{\mathcal{X}}(x_t)$ for each $t \geq 1$ such that*

$$\{x_{t+1}\} = \arg\min_{x \in \mathcal{X}} \Big(\sum_{i=1}^{t} \gamma^{[i,t]} \langle \eta_i \hat{g}_i + p_i, x \rangle - \Big(\gamma^{[1,t]} + \sum_{i=1}^{t} \gamma^{[i+1,t]} (1 - \gamma_i) \Big) \langle \hat{x}_1, x \rangle + \Phi(x) \Big), \quad \forall t \geq 0. \tag{28}$$

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

Proof First of all, in order to prove (28) we claim it suffices to prove that there are $\{p_t\}_{t\geq 1}$ with $p_t \in N_{\mathcal{X}}(x_t)$ for each $t \geq 1$ such that

$$\hat{y}_{t+1} = -\sum_{i=1}^{t} \gamma^{[i,t]}(\eta_i \hat{g}_i + p_i) + \Big(\gamma^{[1,t]} + \sum_{i=1}^{t} \gamma^{[i+1,t]}(1-\gamma_i)\Big)\hat{x}_1, \qquad \forall t \geq 0. \tag{29}$$

To see the sufficiency of this claim, note that

$$\begin{aligned}
x_{t+1} = \Pi_{\mathcal{X}}^{\Phi}&(y_{t+1}) \\
\iff\ & \hat{y}_{t+1} - \hat{x}_{t+1} \in N_{\mathcal{X}}(x_{t+1}) && \text{(by Lemma 42)} \\
\iff\ & \hat{y}_{t+1} \in \partial(\Phi + \delta(\cdot \,|\, \mathcal{X}))(x_{t+1}) && (\partial(\delta(\cdot \,|\, \mathcal{X}))(x) = N_{\mathcal{X}}(x)) \\
\iff\ & x_{t+1} \in \operatorname*{arg\,max}_{x \in \mathbb{R}^n}\big(\langle \hat{y}_{t+1}, x\rangle - \Phi(x) - \delta(x \,|\, \mathcal{X})\big) && \text{(by Lemma 43)} \\
\iff\ & x_{t+1} \in \operatorname*{arg\,min}_{x \in \mathcal{X}}\big(-\langle \hat{y}_{t+1}, x\rangle + \Phi(x)\big).
\end{aligned}$$

The above together with (29) yields (28). Let us now prove (29) by induction on $t \geq 0$.

For $t = 0$, we have that (29) holds trivially. Let $t > 0$. By definition, we have $\hat{y}_{t+1} = (1-\gamma_t)(\hat{x}_t - \eta_t \hat{g}_t) + \gamma_t \hat{x}_1$. At this point, to use the induction hypothesis, we need to write $\hat{x}_t$ as a function of $\hat{y}_t$. From the definition of Algorithm 2, we have $x_t = \Pi_{\mathcal{X}}^{\Phi}(y_t)$. By Lemma 42, the latter holds if and only if $\hat{y}_t - \hat{x}_t \in N_{\mathcal{X}}(x_t)$. That is, there is $p_t \in N_{\mathcal{X}}(x_t)$ such that $\hat{x}_t = \hat{y}_t - p_t$. Plugging these facts together and using our induction hypothesis we have

$$\begin{aligned}
\hat{y}_{t+1} &= \gamma_t(\hat{x}_t - \eta_t \hat{g}_t) + (1-\gamma_t)\hat{x}_1 = \gamma_t(\hat{y}_t - \eta_t \hat{g}_t - p_t) + (1-\gamma_t)\hat{x}_1 \\
&\overset{\text{I.H.}}{=} \gamma_t\Bigg(-\sum_{i=1}^{t-1}\gamma^{[i,t-1]}(\eta_i \hat{g}_i + p_i) - \eta_t \hat{g}_t - p_t \\
&\qquad\qquad + \Big(\gamma^{[1,t-1]} + \sum_{i=1}^{t-1}\gamma^{[i+1,t-1]}(1-\gamma_i)\Big)\hat{x}_1\Bigg) + (1-\gamma_t)\hat{x}_1 \\
&= -\sum_{i=1}^{t}\gamma^{[i,t]}(\eta_i \hat{g}_i + p_i) + \Big(\gamma^{[1,t]} + \sum_{i=1}^{t}\gamma^{[i+1,t]}(1-\gamma_i)\Big)\hat{x}_1,
\end{aligned}$$

and this finishes the proof of (29). ∎

Proof [of Theorem 19] Define $\gamma^{[i,t]}$ for every $i, t \in \mathbb{N}$ as in Proposition 20. If $\gamma_t = 1$ for all $t \geq 1$, then $\gamma^{[i,t]} = 1$ for any $t, i \geq 1$. Moreover, if $\gamma_t = \frac{\eta_{t+1}}{\eta_t}$ for every $t \geq 1$, then for every $t, i \in \mathbb{N}$ with $t \geq i$ we have $\gamma^{[i,t]} = \frac{\eta_{t+1}}{\eta_i}$, which yields $\gamma^{[i,t]}(\eta_i \hat{g}_i + p_i) = \eta_{t+1}(\hat{g}_i + \frac{1}{\eta_i}p_i)$ and

$$\begin{aligned}
\gamma^{[1,t]} + \sum_{i=1}^{t}\gamma^{[i+1,t]}(1-\gamma_i) &= \frac{\eta_{t+1}}{\eta_1} + \sum_{i=1}^{t}\frac{\eta_{t+1}}{\eta_{i+1}}\Big(1 - \frac{\eta_{i+1}}{\eta_i}\Big) \\
&= \frac{\eta_{t+1}}{\eta_1} + \eta_{t+1}\sum_{i=1}^{t}\Big(\frac{1}{\eta_{i+1}} - \frac{1}{\eta_i}\Big) = 1.
\end{aligned}$$

■

With the above theorem, we may compare the iterates of DA, OMD, and DS-OMD by comparing the formulas (DA-Prox), (OMD-Prox), and (DSOMD-Prox). For the simple unconstrained case where $\mathcal{X} = \mathbb{R}^n$ we have $N_{\mathcal{X}}(x_t) = \{0\}$ for each $t \geq 1$, so both (DA-Prox) and (DSOMD-Prox) are identical. If the learning rate is constant, then all three formulas are equivalent. However, if the learning rate is not constant, OMD *is not* equivalent to the latter methods: changing the ordering of the subgradients $\hat{g}_1, \ldots, \hat{g}_t$ affects the choice of x_{t+1} in (OMD-Prox), while this does not happen in the other two formulas.

When $\mathcal{X}$ is an arbitrary convex set, DA and DS-OMD are not necessarily equivalent anymore due to the vectors p_i' in the normal cone of $\mathcal{X}$. However, if we know that the iterates live in the relative interior of $\mathcal{X}$, the next lemma shows that these vectors do not affect the set of minimizers in (DSOMD-Prox).

Lemma 21 *For any $\mathring{x} \in \operatorname{ri} \mathcal{X}$ we have $N_{\mathcal{X}}(\mathring{x}) = -N_{\mathcal{X}}(\mathring{x})$. In particular, for any $p \in N_{\mathcal{X}}(\mathring{x})$ we have $\langle p, x \rangle = \langle p, \mathring{x} \rangle$ for every $x \in \mathcal{X}$.*

Proof Let $\mathring{x} \in \operatorname{ri} \mathcal{X}$. For the sake of contradiction, suppose there is $p \in N_{\mathcal{X}}(\mathring{x})$ and $x \in \mathcal{X}$ such that $\langle p, x - \mathring{x} \rangle < 0$, that is, $-p \notin N_{\mathcal{X}}(\mathring{x})$. Since $\mathring{x}$ is in the relative interior of $\mathcal{X}$, there is $\mu > 1$ such that $x_\mu \coloneqq \mu\mathring{x} + (1-\mu)x \in \mathcal{X}$ (see Rockafellar, 1970, Thm. 6.4). Then,

$$\langle p, x_\mu - \mathring{x} \rangle = (1-\mu)\langle p, x - \mathring{x} \rangle > 0,$$

a contradiction since $p \in N_{\mathcal{X}}(\mathring{x})$. ■

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

Corollary 22 *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 $\mathcal{D} \cap \mathcal{X} \subseteq \operatorname{ri} \mathcal{X}$ and $x_1 = x_1'$, then $x_t = x_t'$ for each $t > 1$.*

Proof Let $t > 1$. Since $x_t = \Pi_{\mathcal{X}}^{\Phi}(y_t)$, where y_t is as in Algorithm 2, Lemma 42 implies $x_t \in \mathcal{D} \cap \mathcal{X} \subseteq \operatorname{ri} \mathcal{X}$. By Lemma 21 we have that the vectors on the normal cone in (DSOMD-Prox) do not affect the set of minimizers, which implies that (DA-Prox) and (DSOMD-Prox) are equivalent. ■

An important special case of the above corollary is the prediction with expert advice setting as in Section 6.2, where $\mathcal{D} = \mathbb{R}^n_{>0}$ and $\mathcal{X}$ is the simplex Δ_n. In this setting, $\mathcal{X} \cap \mathcal{D} = \{\, x \in (0,1)^d : \sum_{i=1}^n x_i = 1 \} = \operatorname{ri} \mathcal{X}$. 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 and Pál, 2018), whereas DA has sublinear regret.

8. Dual-Stabilized OMD for Composite Functions

In this section we extend the dual-stabilized OMD to the case where the functions revealed at each round are composite (Xiao, 2010; Duchi et al., 2010). More specifically, at each round $t \geq 1$ we see a function of the form $f_t + \Psi$, where f_t and Ψ are convex functions, but the latter is fixed and assumed to be "easy", i.e., we suppose we know how to efficiently compute points in $\arg\min_{x \in \mathcal{X}}(D_\Phi(x, \bar{x}) + \Psi(x))$. We could simply use the original dual-stabilized OMD in this setting, but this approach has some drawbacks. One issue is that subgradients of Ψ would end-up appearing in the regret bound from Theorem 3, which is not ideal: we want bounds that are unaffected by the "easy" function Ψ. Another drawback is that we would not be using the knowledge of the structure of the functions, which may result in sub-optimal performance. For example, one may take $\Psi = \|\cdot\|_1$ hoping for sparse iterates. Yet, blindly applying OMD (and, thus, linearizing Ψ) does not yield sparse iterates (McMahan, 2011). Finally, the analysis of dual-stabilized OMD adapted to the composite setting is an easy extension of the original analysis of Section 5. Usually, algorithms for the composite setting require a more intricate analysis, such as in the case of Regularized DA from Xiao (2010), or the use of powerful results, such as the duality between strong convexity and strong smoothness used by McMahan (2017). Duchi et al. (2010) give regret bound a whose analysis is somewhat simpler and perhaps resembles ours. Still, the techniques used in the latter do not directly apply when we use dual-stabilization.

In the composite setting we assume without loss of generality that $\mathcal{X} = \mathbb{R}^n$ since we may substitute Ψ by $\Psi + \delta(\cdot \,|\, \mathcal{X})$ where $\delta(x \,|\, \mathcal{X}) = 0$ if $x \in \mathcal{X}$ and is $+\infty$ anywhere else. The **(effective) domain** of Ψ is the set $\operatorname{dom}\Psi \subseteq \mathbb{R}^n$ of points where Ψ is finite. To avoid confusion and make the effect of Ψ explicit, we define the **Ψ-regret** of a sequence of functions $\{f_t\}_{t\geq 1}$ and iterates $\{x_t\}_{t \geq 1}$ (against a comparison point $z \in \operatorname{dom}\Psi$) by

$$\operatorname{Regret}^\Psi(T, z) \coloneqq \sum_{t=1}^{T} \big(f_t(x_t) + \Psi(x_t)\big) - \sum_{t=1}^{T} \big(f_t(z) + \Psi(z)\big), \quad \forall T \geq 0.$$

To adapt the dual stabilization method to this setting, we use the same idea as in Duchi et al. (2010). Namely, we modify the proximal-like formulation of dual stabilization from Proposition 18 so that we do not linearize (i.e., take the subgradient) of the function Ψ. This results in the following definition of dual-stabilized OMD for composite functions.

$$\{x_{t+1}\} \coloneqq \operatorname*{arg\,min}_{x \in \mathcal{X}} \Big(\gamma_t\big(\eta_t(\langle \hat{g}_t, x\rangle + \Psi(x)) + D_\Phi(x, x_t)\big) + (1 - \gamma_t) D_\Phi(x, x_1)\Big), \quad \forall t \geq 0. \tag{30}$$

This equation defines the algorithm in a proximal form. Due to the existence of Ψ, it is perhaps not obvious that it can also be written in pseudocode resembling Algorithm 2. Nevertheless, it can — Algorithm 5 presents pseudocode equivalent to (30). Interestingly, Ψ appears in the pseudocode only in the projection step. For the sake of conciseness, we defer the proof of equivalence between Algorithm 5 and (30) to our technical report, but it boils down to properly interpreting the optimality conditions of (30). In this new algorithm, we extend the definition of Bregman Projection and define the Ψ-Bregman projection by $\{\Pi_\Psi^\Phi(y)\} \coloneqq \arg\min_{x \in \mathbb{R}^n}(D_\Phi(x, y) + \Psi(y))$. The next lemma shows an analogue of the generalized Pythagorean theorem for the Ψ-Bregman Projection.

Algorithm 5 Dual-stabilized OMD with dynamic learning rate η_t and additional regularization function Ψ.

Input: $x_1 \in \arg\min_{x \in \mathbb{R}^n} \Psi(x),\ \eta : \mathbb{N} \to \mathbb{R}_+,\ \gamma : \mathbb{N} \to [0, 1]$
$\hat{y}_1 = \nabla\Phi(x_1)$
for $t = 1, 2, \ldots$ **do**
Incur cost $f_t(x_t)$ and receive $\hat{g}_t \in \partial f_t(x_t)$
$\hat{x}_t = \nabla\Phi(x_t)$ ▷ map primal iterate to dual space
$\hat{w}_{t+1} = \hat{x}_t - \eta_t \hat{g}_t$ ▷ gradient step in dual space (31)
$\hat{y}_{t+1} = \gamma_t \hat{w}_{t+1} + (1 - \gamma_t)\hat{x}_1$ ▷ stabilization in dual space (32)
$y_{t+1} = \nabla\Phi^*(\hat{y}_{t+1})$ ▷ map dual iterate to primal space
$\alpha_{t+1} = \eta_t \gamma_t$ ▷ compute scaling factor for Ψ
$x_{t+1} = \Pi^{\Phi}_{\alpha_{t+1}\Psi}(y_{t+1})$ ▷ project onto feasible region (33)
end for

Lemma 23 *Let $\alpha > 0$ and $\bar{y} \coloneqq \Pi^{\Phi}_{\alpha\Psi}(y)$. Then,*

$$D_\Phi(x, \bar{y}) + D_\Phi(\bar{y}, y) \leq D_\Phi(x, y) + \alpha(\Psi(x) - \Psi(\bar{y})), \qquad \forall x \in \mathbb{R}^n.$$

Proof By the optimality conditions for convex optimization, we have

$$\nabla\Phi(y) - \nabla\Phi(\bar{y}) \in \partial(\alpha\Psi)(\bar{y}).$$

Using the generalized triangle inequality for Bregman Divergences (Proposition 36) and the subgradient inequality, we get

$$D_\Phi(x, \bar{y}) + D_\Phi(\bar{y}, y) - D_\Phi(x, y) = \langle \nabla\Phi(y) - \nabla\Phi(\bar{y}), x - \bar{y} \rangle \overset{\text{(i)}}{\leq} \alpha(\Psi(x) - \Psi(\bar{y})).$$

where (i) follows from $\nabla\Phi(y) - \nabla\Phi(\bar{y}) \in \partial(\alpha\Psi)(\bar{y})$ and the convexity of $\alpha\Psi(\cdot)$. ■

Finally, a regret bound such as the one we have for dual-stabilized OMD also holds in this setting when using Algorithm 5.

Theorem 24 *Assume we have $\eta_t \geq \eta_{t+1} > 0$ for each $t > 1$. 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 $f_t \colon \mathbb{R}^n \to \mathbb{R}$ for each $t \geq 1$ and let $\Phi \colon \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be convex. Let $\{x_t\}_{t \geq 1}$ and $\{\hat{w}_t\}_{t \geq 2}$ be as in Algorithm 5. Then, for all $T > 0$ and $z \in \operatorname{dom}\Psi$,*

$$\operatorname{Regret}^{\Psi}(T, z) \ \leq \ \sum_{t=1}^{T} \frac{D_\Phi(\substack{x_t \\ x_{t+1}}; w_{t+1})}{\eta_t} + \frac{D_\Phi(z, x_1)}{\eta_{T+1}}, \quad \forall T > 0. \tag{34}$$

Proof (Sketch) The proof boils down to simple modifications to the original proof of Theorem 3. We mostly have to replace Claim 4 by the following claim, whose proof mimics the proof of the latter but replaces Proposition 38 with Lemma 23.

Claim 25 *Assume that* $\gamma_t = \eta_{t+1}/\eta_t \in (0,1]$. *Then*

$$
\begin{aligned}
&\frac{1}{\eta_t}\Big(D_\Phi(x_t, w_{t+1}) + D_\Phi(z, x_t) - D_\Phi\big(z, w_{t+1}\big)\Big) + \Psi(x_t) - \Psi(z) \\
&\leq \frac{D_\Phi\left({x_t \atop x_{t+1}}; w_{t+1}\right)}{\eta_t} \quad + \underbrace{\left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\right)}_{telescopes} D_\Phi(z, x_1) + \underbrace{\frac{D_\Phi(z, x_t)}{\eta_t} - \frac{D_\Phi(z, x_{t+1})}{\eta_{t+1}}}_{telescopes} + \underbrace{\Psi(x_t) - \Psi(x_{t+1})}_{telescopes}.
\end{aligned}
$$

Due to similarities with the proof of the previous regret bounds, we defer the complete proof to our technical report (Fang et al., 2021). ■

Acknowledgments

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.

Appendix A. Standard facts

A.1 Scalar inequalities

Fact 26 *For any* $a > 0$ *and* $b, x \in \mathbb{R}$*, we have* $-ax^2 + bx \leq b^2/4a$.

Fact 27 $e^{-x} \leq 1 - x + \frac{x^2}{2}$ *for* $x \geq 0$.

Fact 28 $\sum_{i=1}^{t} \frac{1}{\sqrt{i}} \leq 2\sqrt{t} - 1$ *for* $t \geq 1$.

Fact 29 $\log(x) \leq x - 1$ *for* $x \geq 0$.

The following proposition is a variant of an inequality that is frequently used in online learning; see, e.g., Auer et al. (2002b, Lemma 3.5), McMahan (2017, Lemma 4). Since well-known proofs can be easily adapted to prove the version used here, we defer a complete proof to our technical report (Fang et al., 2021).

Proposition 30 *Let* $u > 0$ *and* $a_1, a_2, \ldots, a_T \in [0, u]$*. Then*

$$\sum_{t=1}^{T} \frac{a_t}{\sqrt{u + \sum_{i<t} a_i}} \;\leq\; 2\sqrt{\sum_{t=1}^{T} a_t}.$$

We use following technical proposition in our proof of the first-order regret bounds from Corollary 16. Since the proof is not enlightening nor complicated, we also defer it to our technical report (Fang et al., 2021).

Proposition 31 *Let* $x, y, \alpha, \beta > 0$.

$$\text{If} \quad x - y \;\leq\; \alpha\sqrt{x} + \beta, \quad \text{then} \quad x - y \;\leq\; \alpha\sqrt{y} + \beta + \alpha\sqrt{\beta} + \alpha^2.$$

A.2 Bregman divergence properties

The following lemma collects basic facts regarding the a mirror map Φ (or simply a function of Legendre type) and the Bregman divergence it induces. See Cesa-Bianchi and Lugosi (2006, Lemma 11.5 and Proposition 11.1).

Lemma 32 *The mirror map* Φ *and the Bregman divergence it induces satisfy the following properties:*

- $D_\Phi(x, y)$ *is convex in* x.
- $\nabla\Phi(\nabla\Phi^*(z)) = z$ *and* $\nabla^*\Phi(\nabla\Phi(x)) = x$ *for all* x *and* z.
- $D_\Phi(x, y) = D_{\Phi^*}(\nabla\Phi(y), \nabla\Phi(x))$ *for all* x *and* y.

Proposition 33 *If* Φ *is* ρ*-strongly convex with respect to* $\|\cdot\|$ *then* $D_\Phi(x, y) \geq \frac{\rho}{2}\|x - y\|^2$.

A.2.1 Differences of Bregman divergences

Recall that in (3) we defined the notation

$$D_\Phi\left({a \atop b}; c\right) \;\coloneqq\; D_\Phi(a,c) - D_\Phi(b,c) \;=\; \Phi(a) - \Phi(b) - \langle\, \nabla\Phi(c),\, a-b \,\rangle.$$

This has several useful properties, which we now discuss.

Proposition 34 *$D_\Phi\left({a \atop b}; p\right)$ is linear in $\hat{p}$. In particular,*

$$D_\Phi\left({a \atop b}; \nabla\Phi^*(\hat{p}-\hat{q})\right) \;=\; D_\Phi\left({a \atop b}; p\right) + \langle\, \hat{q},\, a-b \,\rangle \qquad \forall \hat{q} \in \mathbb{R}^n.$$

Proof Immediate from the definition. ■

Proposition 35 *For all $a,b,c,d \in \mathcal{D}$,*

$$D_\Phi\left({a \atop b}; d\right) - D_\Phi\left({a \atop b}; c\right) \;=\; \langle\, \hat{c} - \hat{d},\, a-b \,\rangle \;=\; D_\Phi\left({a \atop b}; d\right) + D_\Phi\left({b \atop a}; c\right).$$

Proof The first equality holds from Proposition 34 with $\hat{p} = \hat{c}$ and $\hat{q} = \hat{c} - \hat{d}$. The second equality holds since $D_\Phi\left({b \atop a}; c\right) = -D_\Phi\left({a \atop b}; c\right)$. ■

An immediate consequence is the "generalized triangle inequality for Bregman divergence", such as in Bubeck (2015, Eq. (4.1)) or in Beck and Teboulle (2003, Lemma 4.1).

Proposition 36 *For all $a,b,d \in \mathcal{D}$,*

$$D_\Phi(a,d) - D_\Phi(b,d) + D_\Phi(b,a) \;=\; \langle\, \hat{a} - \hat{d},\, a-b \,\rangle$$

Proof Apply Proposition 35 with $c = a$ and use $D_\Phi(a,a) = 0$. ■

Proposition 37 *Let $a,b,c,u,v \in \mathbb{R}^n$ satisfy $\gamma\hat{a} + (1-\gamma)\hat{b} = \hat{c}$ for some $\gamma \in \mathbb{R}$. Then*

$$\gamma D_\Phi\left({u \atop v}; a\right) \;+\; (1-\gamma) D_\Phi\left({u \atop v}; b\right) \;=\; D_\Phi\left({u \atop v}; c\right).$$

Proof By definition of D_Φ, the claimed identity is equivalent to

$$\begin{aligned}\gamma\big(\Phi(u) - \Phi(v) - \langle\, \nabla\Phi(a),\, u-v \,\rangle\big) \;+\; (1-\gamma)\big(\Phi(u) - \Phi(v) - \langle\, \nabla\Phi(b),\, u-v \,\rangle\big) \\ =\; \big(\Phi(u) - \Phi(v) - \langle\, \nabla\Phi(c),\, u-v \,\rangle\big).\end{aligned}$$

This equality holds by canceling $\Phi(u) - \Phi(v)$ and by the assumption that $\nabla\Phi(c) = (1-\gamma)\nabla\Phi(a) + \gamma\nabla\Phi(b)$. ■

The following proposition is the "Pythagorean theorem for Bregman divergence". Recall that $\Pi^\Phi_\mathcal{X}(y) = \arg\min_{u \in \mathcal{X}} D_\Phi(u,y)$. A proof may be found in Bubeck (2015, Lemma 4.1).

Proposition 38 *Let $\mathcal{X} \subset \mathbb{R}^n$ be a convex set. Let $p \in \mathbb{R}^n$ and $\pi = \Pi^{\Phi}_{\mathcal{X}}(p)$. Then*

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

A generalization of the previous proposition can be obtained by using linearity.

Proposition 39 *Let $\mathcal{X} \subset \mathbb{R}^n$ be a convex set. Let $p \in \mathbb{R}^n$ and $\pi = \Pi^{\Phi}_{\mathcal{X}}(p)$. Then*

$$D_\Phi(\tfrac{v}{\pi}; \nabla\Phi^*(\hat{p} - \hat{q})) \;\geq\; D_\Phi(\tfrac{v}{\pi}; \nabla\Phi^*(\hat{\pi} - \hat{q})) \qquad \forall v \in \mathcal{X},\; \hat{q} \in \mathbb{R}^n.$$

Proof

$$\begin{aligned}
D_\Phi(\tfrac{v}{\pi}; \nabla\Phi^*(\hat{p} - \hat{q})) &= D_\Phi(\tfrac{v}{\pi}; p) + \langle\, \hat{q},\, v - \pi \,\rangle && \text{(by Proposition 34)} \\
&\geq D_\Phi(\tfrac{v}{\pi}; \pi) + \langle\, \hat{q},\, v - \pi \,\rangle && \text{(by Proposition 38)} \\
&= D_\Phi(\tfrac{v}{\pi}; \nabla\Phi^*(\hat{\pi} - \hat{q})) && \text{(by Proposition 34).}
\end{aligned}$$

■

Appendix B. Additional proofs for Section 6.2

An initial observation shows that Λ is non-negative in the experts' setting.

Proposition 40 $\Lambda(a, b) \geq 0$ *for all $a \in \mathcal{X}$, $b \in \mathcal{D}$.*

Proof Let us write $\Lambda(a, b) = -\sum_{i=1}^{n} a_i \ln \frac{b_i}{a_i} + \ln\left(\sum_{i=1}^{n} b_i\right)$. Since a is a probability distribution, we may apply Jensen's inequality to show that this expression is non-negative. ■

Proof [of Proposition 13] Since $a, b \in \mathcal{X}$ we have $\|a\|_1 = \|b\|_1 = 1$. Then

$$\begin{aligned}
D_\Phi(\tfrac{a}{b}; c) &= D_{\mathrm{KL}}(a, c) - D_{\mathrm{KL}}(b, c) \\
&= \big(D_{\mathrm{KL}}(a, c) + 1 - \|c\|_1 + \ln \|c\|_1\big) - \big(D_{\mathrm{KL}}(b, c) + 1 - \|c\|_1 + \ln \|c\|_1\big) \\
&= \Lambda(a, c) - \Lambda(b, c) && \text{(by definition of } \Lambda) \\
&\leq \Lambda(a, c) && \text{(by Proposition 40).}
\end{aligned}$$

■

Proof [of Proposition 17] Let $b = \nabla\Phi^*(\hat{a} - \eta\hat{q})$. By (24), $b_i = a_i \exp(-\eta\hat{q}_i)$. Then

$$\begin{aligned}
\Lambda(a, \nabla\Phi^*(\hat{a} - \eta\hat{q})) &= \sum_{i=1}^n a_i \ln(a_i/b_i) + \ln \|b\|_1 \\
&= \sum_{i=1}^n \eta a_i \hat{q}_i + \ln\Big(\sum_{i=1}^n a_i \exp(-\eta\hat{q}_i)\Big) \\
&\le \sum_{i=1}^n \eta a_i \hat{q}_i + \sum_{i=1}^n a_i \exp(-\eta\hat{q}_i) - 1 && \text{(by Fact 29)} \\
&\le \sum_{i=1}^n \eta a_i \hat{q}_i + \sum_{i=1}^n a_i \left(1 - \eta\hat{q}_i + \frac{\eta^2 \hat{q}_i^2}{2}\right) - 1 && \text{(by Fact 27)} \\
&\le \eta^2 \sum_{i=1}^n a_i \hat{q}_i/2,
\end{aligned}$$

using $\sum_{i=1}^n a_i = 1$ (since $a \in \mathcal{X}$) and $\hat{q}_i^2 \le \hat{q}_i$ (since $\hat{q} \in [0,1]^n$). ■

Appendix C. Additional proofs for Section 7

At many points throughout this section we will need to talk about optimality condition for problems where we minimize a convex function over a convex set. Such conditions depend on the *normal cone* of the set on which the optimization is taking place.

Lemma 41 (Rockafellar, 1970, Theorem 27.4) *Let $h\colon \mathcal{C} \to \mathbb{R}$ be a closed convex function such that $(\operatorname{ri}\mathcal{C}) \cap (\operatorname{ri}\mathcal{X}) \neq \varnothing$. Then, $x \in \arg\min_{z\in\mathcal{X}} h(z)$ if and only if there is $\hat{g} \in \partial h(x)$ such that $-\hat{g} \in N_{\mathcal{X}}(x)$.*

Using the above result allows us to derive a useful characterization of points that realize the Bregman projections. This result is similar to Bubeck (2015, Lemma 4.1) and we defer the complete proof to our technical report (Fang et al., 2021).

Lemma 42 *Let $y \in \mathcal{D}$ and $x \in \bar{\mathcal{D}}$. Then $x = \Pi_{\mathcal{X}}^{\Phi}(y)$ if and only if $x \in \mathcal{D} \cap \mathcal{X}$ and $\nabla\Phi(y) - \nabla\Phi(x) \in N_{\mathcal{X}}(x)$.*

Proof Suppose $x \in \mathcal{D} \cap \mathcal{X}$ and $\nabla\Phi(y) - \nabla\Phi(x) \in N_{\mathcal{X}}(x)$. Since $\nabla\Phi(y) - \nabla\Phi(x) = -\nabla(D_\Phi(\cdot, y))(x)$, by Lemma 41 we conclude that $x \in \arg\min_{z\in\mathcal{X}} D(z, y)$. Now suppose $x = \Pi_{\mathcal{X}}^{\Phi}(y)$. By Lemma 41 together with the definition of Bregman divergence, this is the case if and only if there is $-g \in \partial\Phi(x)$ such that $-(g - \nabla\Phi(y)) \in N_{\mathcal{X}}(x)$. Since Φ is of Legendre type we have $\partial\Phi(z) = \varnothing$ for any $z \notin \mathcal{D}$ (Rockafellar, 1970, Theorem 26.1). Thus, $x \in \mathcal{D}$ and $g = \nabla\Phi(x)$ since Φ is differentiable. Finally, $x \in \mathcal{X}$ by the definition of Bregman projection. ■

For some of the proofs from Section 7, we need to one last result about the relation of subgradients and conjugate functions which is worth stating in full.

Lemma 43 (Rockafellar, 1970, Theorem 23.5) *Let $f\colon \mathcal{X} \to \mathbb{R}$, let $x \in \mathcal{X}$ and let $\hat{y} \in \mathbb{R}^n$. Then $\hat{y} \in \partial f(x)$ if and only if x attains $\sup_{x\in\mathbb{R}^n}(\langle \hat{y}, x\rangle - f(x)) = f^*(\hat{y})$.*

References

Zeyuan Allen-Zhu and Lorenzo Orecchia. Linear coupling: An ultimate unification of gradient and mirror descent. November 2016. URL `https://arxiv.org/abs/1407.1537`.

Zeyuan Allen-Zhu, Zhenyu Liao, and Lorenzo Orecchia. Spectral sparsification and regret minimization beyond matrix multiplicative updates [extended abstract]. In *Proceedings of STOC'15*, pages 237–245. ACM, New York, 2015.

Jean-Yves Audibert, Sébastien Bubeck, and Gábor Lugosi. Regret in online combinatorial optimization. *Mathematics of Operations Research*, 39(1):31–45, 2014.

Peter Auer, Nicolò Cesa-Bianchi, Yoav Freund, and Robert E. Schapire. The nonstochastic multiarmed bandit problem. *SIAM Journal on Computing*, 32(1), 2002a.

Peter Auer, Nicolò Cesa-Bianchi, and Claudio Gentile. Adaptive and self-confident on-line learning algorithms. *Journal of Computer and System Sciences*, 64(1):48–75, 2002b.

Amir Beck. *First-order methods in optimization*, volume 25 of *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2017.

Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 31(3):167–175, 2003.

Sébastien Bubeck. Convex optimization: Algorithms and complexity. *Foundations and Trends in Machine Learning*, 8(3-4):231–357, 2015.

Sébastien Bubeck, Michael B. Cohen, Yin Tat Lee, James R. Lee, and Aleksander Madry. k-server via multiscale entropic regularization. In *Proceedings of STOC'18*, pages 3–16. ACM, New York, 2018.

Sbastien Bubeck. Introduction to online optimization, December 2011. unpublished.

Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006.

Nicolò Cesa-Bianchi, Yoav Freund, David Haussler, David P. Helmbold, Robert E. Schapire, and Manfred K. Warmuth. How to use expert advice. *Journal of the ACM*, 44(3), May 1997.

Nicolo Cesa-Bianchi, Yishay Mansour, and Gilles Stoltz. Improved second-order bounds for prediction with expert advice. *Machine Learning*, 66(2-3):321–352, 2007.

Steven de Rooij, Tim van Erven, Peter D. Grünwald, and Wouter M. Koolen. Follow the leader if you can, hedge if you must. *Journal of Machine Learning Research (JMLR)*, 15: 1281–1316, 2014.

John C. Duchi, Shai Shalev-shwartz, Yoram Singer, and Ambuj Tewari. Composite objective mirror descent. In *Proceedings of COLT*, pages 14–26, 2010.

Huang Fang, Nicholas J. A. Harvey, Victor S. Portella, and Michael P. Friedlander. Online mirror descent and dual averaging: keeping pace in the dynamic case, aug 2021. arXiv:2006.02585.

Sébastien Gerchinovitz. *Prediction of individual sequences and prediction in the statistical framework: some links around sparse regression and aggregation techniques*. PhD thesis, Université Paris-Sud, 2011.

András György and Csaba Szepesvári. Shifting regret, mirror descent, and matrices. In *Proceedings of ICML'16*, pages 2943–2951, 2016.

Elad Hazan. Introduction to online convex optimization. *Foundations and Trends in Optimization*, 2(3-4):157–325, 2016.

Pooria Joulani, András György, and Csaba Szepesvári. A modular analysis of adaptive (non-)convex optimization: Optimism, composite objectives, and variational bounds. In *Proceedings of ALT'17*, pages 681–720, 2017.

Anatoli B. Juditsky, Joon Kwon, and Eric Moulines. Unifying mirror descent and dual averaging. October 2019. URL `http://arxiv.org/abs/1910.13742`.

Sham M. Kakade, Shai Shalev-Shwartz, and Ambuj Tewari. Regularization techniques for learning with matrices. *Journal of Machine Learning Research (JMLR)*, 13:1865–1890, 2012.

Ehsan Asadi Kangarshahi, Ya-Ping Hsieh, Mehmet Fatih Sahin, and Volkan Cevher. Let's be honest: An optimal no-regret framework for zero-sum games. In *Proceedings of ICML'18*, pages 2493–2501, 2018.

H. Brendan McMahan. Follow-the-regularized-leader and mirror descent: Equivalence theorems and L1 regularization. In *Proceedings of AISTATS'11*, pages 525–533, 2011.

H. Brendan McMahan. A survey of algorithms and analysis for adaptive online learning. *Journal of Machine Learning Research*, 18:90:1–90:50, 2017.

Arkadi Nemirovski and David Yudin. *Problem Complexity and Method Efficiency in Optimization*. Wiley Interscience, 1983.

Yurii Nesterov. Primal-dual subgradient methods for convex problems. *Mathematical Programming*, 120(1):221–259, 2009.

Francesco Orabona and Dávid Pál. Scale-free online learning. *Theor. Comput. Sci.*, 716: 50–69, 2018.

R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, 1970.

Shai Shalev-Shwartz. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2):107–194, 2012.

Lin Xiao. Dual averaging methods for regularized stochastic learning and online optimization. *Journal of Machine Learning Research*, 11, 2010.

Rani Yaroshinsky, Ran El-Yaniv, and Steven S. Seiden. How to better use expert advice. *Machine Learning*, 55(3):271–309, 2004.

Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of ICML*, pages 928–936, 2003.