

# The Hidden Cost of Approximation in Online Mirror Descent

**Ofir Schliesselberg**

*Tel Aviv University*

OFIRS4@MAIL.TAU.AC.IL

**Uri Sherman**

*Tel Aviv University*

URISHERMAN@MAIL.TAU.AC.IL

**Tomer Koren**

*Tel Aviv University and Google Research*

TKOREN@TAUEX.TAU.AC.IL

**Yishay Mansour**

*Tel Aviv University and Google Research*

MANSOUR.YISHAY@GMAIL.COM

**Editors:** Steve Hanneke and Tor Lattimore

## Abstract

Online mirror descent (OMD) is a fundamental algorithmic paradigm that underlies many algorithms in optimization, machine learning and sequential decision-making. The OMD iterates are defined as solutions to optimization subproblems which, oftentimes, can be solved only approximately, leading to an *inexact* version of the algorithm. Nonetheless, existing OMD analyses typically assume an idealized error free setting, thereby limiting our understanding of performance guarantees that should be expected in practice. In this work we initiate a systematic study into inexact OMD, and uncover an intricate relation between regularizer smoothness and robustness to approximation errors. When the regularizer is uniformly smooth, we establish a tight bound on the excess regret due to errors. Then, for barrier regularizers over the simplex and its subsets, we identify a sharp separation: negative entropy requires exponentially small errors to avoid linear regret, whereas log-barrier and Tsallis regularizers remain robust even when the errors are only polynomial. Finally, we show that when the losses are stochastic and the domain is the simplex, negative entropy regains robustness - but this property does not extend to all subsets, where exponentially small errors are again necessary to avoid suboptimal regret.

## 1. Introduction

Mirror Descent (Nemirovsky and Yudin, 1983; Beck and Teboulle, 2003) is a fundamental optimization paradigm that offers the flexibility to exploit the (typically non-Euclidean) intrinsic geometry of the optimization problem. The online variant (OMD; Shalev-Shwartz, 2012; Hazan, 2016) is a generalization of the basic framework adapted to the more general online learning setup Zinkevich (2003), where the goal of the learner is to minimize her *regret*, defined as the cumulative loss minus the loss of the best fixed decision in hindsight. Given a convex decision set  $\mathcal{K} \subset \mathbb{R}^d$ , an initialization  $w_1 \in \mathcal{K}$  and learning rate  $\eta > 0$ , the OMD steps  $t = 1, \dots, T$  follow the update rule:

$$w_{t+1} = \arg \min_{w \in \mathcal{K}} \{ \eta \langle \ell_t, w \rangle + D_R(w \| w_t) \}, \quad (1)$$

where  $\ell_t$  is the loss at time  $t$  and  $D_R$  is the Bregman divergence associated with a regularizer  $R: \mathcal{K} \rightarrow \mathbb{R}$  chosen by the learner. Notable instances of OMD include online gradient descent (Zinkevich, 2003) and the well known multiplicative weights method (Littlestone and Warmuth, 1994; Freund and Schapire, 1997; Arora et al., 2012), both of which are examples where the OMD update rule, namely the exact solution to the OMD subproblem Eq. (1), is given by a closed form expression (when operating over suitable decision sets).

However, in many cases of interest, the OMD update rule does not admit a closed form solution, and therefore demands employing an auxiliary iterative optimization procedure that only produces *approximate* minimizers of the respective OMD subproblems. Notable examples include reinforcement learning algorithms that optimize over occupancy measures, which form a polyhedral subset of the simplex (Zimin and Neu, 2013; Rosenberg and Mansour, 2019; Jin et al., 2020); generic online convex optimization algorithms that rely on OMD updates (Abernethy and Rakhlin, 2009; Hazan and Levy, 2014; Ito, 2020); and algorithms defined over the simplex that use barrier regularization other than negative entropy, such as in adversarial bandits (Abernethy et al., 2015; Zimmert and Seldin, 2021) and portfolio selection (Luo et al., 2018). Somewhat surprisingly, however, the existing literature lacks a systematic study of the effect these approximations have on the final regret guarantee, with prior art focusing on particular problem instances at best Schmidt et al. (2011); Villa et al. (2013); Dixit et al. (2019); Choi et al. (2023).

In this work, we initiate a systematic study into the robustness of OMD to approximations, aimed at understanding the interplay between regularization, quality of approximations, and regret. Our results uncover a direct link between robustness of inexact OMD and smoothness properties of the regularizer being used. For uniformly smooth regularizers, we establish that robustness to approximation errors is directly governed by the smoothness parameter. For the more prevalent non-smooth regularizer case, we demonstrate that OMD with negative entropy regularization is prone to incurring *linear* regret unless the approximation errors are made *exponentially small* in the number of steps; and in contrast, that for other barrier regularizers such as the log-barrier and Tsallis entropy, polynomially small errors suffice to obtain optimal regret. We then further investigate more carefully when non-robustness with the negative entropy arises. We show that when the losses are stochastic (i.i.d.), negative entropy over the simplex becomes robust and polynomially small errors are sufficient. On the other hand, we demonstrate this robustness may break even with i.i.d. losses when optimizing over a *subset* of the simplex, where again, exponentially small errors are necessary to avoid suboptimal regret.

### 1.1. Summary of contributions

In more detail, our contributions are summarized as follows.

- First, when the regularizer  $R$  is uniformly smooth over the domain  $\mathcal{K}$  with smoothness parameter  $\beta$ , we establish a tight  $\Theta(TD\sqrt{\beta\varepsilon}/\eta)$  bound on the excess regret due to  $\varepsilon$ -approximation errors, where  $D$  is the diameter of  $\mathcal{K}$  with respect to the relevant norm. E.g., for the typical setting  $\eta = \Theta(1/\sqrt{T})$  this implies that errors should be as small as  $\varepsilon = O(1/T^2)$  so as to recover optimal  $O(\sqrt{T})$  regret.
- We then move on to consider common non-smooth regularizers, such as the negative entropy, Tsallis entropies, and the log-barrier, focusing on the simplex and its subsets as decision sets. We observe a sharp dichotomy between the negative entropy and other regularizers in terms of robustness to approximations: on the one hand, for the negative entropy we show that an *exponentially small* error  $\varepsilon = \Omega(\eta e^{-\eta T})$  could already lead to *linear regret*, even when the domain is the simplex; and on the other hand, for Tsallis Entropies and the log-barrier over the simplex or a subset thereof, we prove that a *polynomially small* error, e.g  $\varepsilon = O(\eta^2/(T^2 d^2))$  for log-barrier, suffices for maintaining the same order of regret.
- Finally, we revisit the robustness to approximations with the negative entropy in the stochastic (i.i.d.) setting. Over the simplex and with  $\eta = \tilde{O}(1/\sqrt{T})$ , we show that a polynomially small error  $\varepsilon = O(1/(d^2 T^4))$  suffices for obtaining optimal regret with high probability, as

Regime	Decision set	Regularizer	Tolerated $\varepsilon$	Polynomial?
Adversarial	convex	$\beta$ -smooth	$\eta^4/\beta$	✓
Adversarial	simplex subset	$\nu$ -barrier ( $\nu > 1$ )	$\eta^4(\eta T d)^{-\nu/\nu-1}$	✓
Adversarial	simplex	negative entropy	$e^{-\eta T}$	✗
Stochastic	simplex	negative entropy	$d^{-2}T^{-4}$	✓
Stochastic	simplex subset	negative entropy	$e^{-1/\eta}$	✗

Table 1: Summary of contribution. The required  $\varepsilon$  column neglects low order terms.

opposed to the exponentially small error required in the non-stochastic case. However, this robustness does not extend more generally to proper subsets of the simplex: we construct a setting where OMD with negative entropy exhibits an excess term of  $\Omega(T\sqrt{\eta/\log(1/\varepsilon)})$  leading to  $\tilde{\Omega}(T^{2/3})$  regret for any step size unless  $\varepsilon$  is exponentially small in  $T$ .

At a conceptual level, our analysis reveals that compounding errors play a central role in OMD’s robustness to inexact updates. Since the per time step subproblem directly depends on the previous iterate, approximation errors propagate between rounds and lead to subtle optimization dynamics. This should be contrasted with the closely related Follow-The-Regularized-Leader (FTRL) algorithm (e.g., Shalev-Shwartz, 2012; Hazan, 2016), which re-optimizes against the cumulative loss at each round, and thus, each optimization round is independent of inexactness introduced in previous rounds. And indeed, for FTRL it is straightforward to prove that approximation errors have only a minor effect; for more details, see Appendix H<sup>1</sup>.

In addition, our results for the smooth case (Theorems 2 and 3) provide a tight characterization that is immediately applicable to a common technique where OMD is instantiated over a *shrunk* simplex (or subset thereof), where coordinates are bounded away from zero (as suggested by Dick et al. (2014)). In this case, a uniform bound for the smoothness parameter immediately follows as the regularizer domain becomes compact. Interestingly, our results for the non-smooth case reveal that while this technique may be necessary to cope with fragility of negative entropy (Theorem 4), it is not necessary for other barrier regularizers as they induce optimization dynamics where the iterates *naturally* stay bounded away from zero (see Theorem 6 and the discussion that follows).

Finally, we note that while our study focuses on the linear setup, all our results for the adversarial setting immediately carry to the general convex case via a standard reduction (e.g., Cesa-Bianchi and Lugosi 2006).

## 1.2. Related work

Mirror descent Nemirovskij and Yudin (1983); Beck and Teboulle (2003) and the online convex optimization framework (Zinkevich, 2003) have been central to the study of machine learning and optimization in the last decades. There exist many excellent books and surveys that provide thorough introductions to (online) mirror descent in its fundamental (i.e., exact, error free) form (Shalev-Shwartz, 2012; Bubeck, 2015; Hazan, 2016; Beck, 2017). Somewhat surprisingly, there hardly exist any works that study inexact mirror descent in the general stochastic or online setup.

In the classical (offline) optimization setup where the objective function is smooth, mirror descent coincides with a special case of the Bregman proximal gradient method (BPGM; Bauschke et al. (2017); Lu et al. (2018), see also Teboulle (2018)). The BPGM is a generalization of the

1. We note that even though the iterates of FTRL and OMD coincide in some cases in the exact case, it isn’t true for the approximate case.

proximal gradient method (Rockafellar, 1976) where a Bregman divergence replaces the norm proximity regularizer, and the objective is required to satisfy the weaker *relative smoothness* property (Bauschke et al., 2017). The BPGM and mirror descent coincide when the non-smooth part in the composite objective is the indicator function for the decision set. In contrast to online or stochastic mirror descent in the general case, inexact versions of the BPGM (and thus offline mirror descent in the smooth case) have been subject to several recent works. The majority of these study the Euclidean case (i.e., the proximal gradient method) with or without acceleration, e.g., Schmidt et al. (2011); Villa et al. (2013); Zhou and Pan (2022); Ahookhosh and Nesterov (2024). Some works study the online case Dixit et al. (2019) with the euclidean regularizer, and some further generalize to the online BPGM but with smooth regularizers (Choi et al., 2023).

There is also a recent line of works that study the (offline) BPGM in its general form (i.e., without making assumptions on the regularizer). These mostly focus on designing variants of the basic method that incorporate some mechanism to cope with the proximal subproblem approximation errors (Rebegoldi et al., 2018; Kabbadj, 2020; Stonyakin et al., 2021; Yang and Toh, 2025)—which is to be contrasted with characterizing convergence in terms of the ad-hoc approximation errors. As one example, the work of Kabbadj (2020) establishes that the inexact BPGM achieves the same rate of the exact version (aka NoLips; Bauschke et al., 2017) as long as the approximation errors are smaller than the Bregman distance to the previous iterate. More recently, Yang and Toh (2025) propose variants with several advantages at the expense of a somewhat more involved subproblem optimization procedure.

Finally, the work of Guigues (2021) is one of the only examples (to our best knowledge) of papers that study an inexact version of stochastic mirror descent, albeit one that relates to a particular (non-general) instantiation of the algorithm.

## 2. Preliminaries

We consider the standard online linear optimization setup, where at each round  $t = 1, 2, \dots, T$ , the learner selects a point  $w_t$  from a convex decision set  $\mathcal{K} \subset \mathbb{R}^d$ , and then observes a loss vector  $\ell_t \in [-1, 1]^d$ . The performance of the learner is measured in terms of her *regret* with respect to a fixed comparator point  $w \in \mathcal{K}$ , defined as follows:

$$\text{Regret}(w) = \sum_{t=1}^T \langle \ell_t, w_t \rangle - \sum_{t=1}^T \langle \ell_t, w \rangle.$$

We denote by  $w^* \in \arg \min_{w \in \mathcal{K}} \sum_{t=1}^T \langle \ell_t, w \rangle$  the best fixed decision in hindsight.

**Inexact Online Mirror Descent.** We let  $R : \mathcal{K} \rightarrow \mathbb{R}$  denote a differentiable regularizer which we assume to be 1-strongly convex w.r.t. a norm  $\|\cdot\|$ . The Bregman divergence associated with  $R$  is defined as:

$$D_R(w \| w') = R(w) - R(w') - \langle \nabla R(w'), w - w' \rangle.$$

We say that a sequence  $\{w_t\}_{t=1}^T$  is an  $\varepsilon$ -approximate OMD trajectory if, for every  $t$ ,  $w_{t+1}$  approximately minimizes the round  $t$  OMD objective (see Eq. 1)  $\phi_t(w) := \eta \langle \ell_t, w \rangle + D_R(w \| w_t)$ , up to  $\varepsilon$  additive error:

$$\phi_t(w_{t+1}) \leq \min_{w \in \mathcal{K}} \phi_t(w) + \varepsilon.$$

Regret bounds for OMD typically depend on the *diameter* of  $\mathcal{K}$  with respect to the norm  $\|\cdot\|$ , given by  $D = \max_{w, w' \in \mathcal{K}} \|w - w'\|$  and demand  $\|\ell_t\|_* \leq 1$  for all  $t \in [T]$ .

**Barrier Regularization.** A particular focus of this work is on prototypical barrier regularizers, used extensively in cases where  $\mathcal{K}$  is the probability simplex  $\Delta_d := \{p \in \mathbb{R}^d : p^i \geq 0, \sum_{i=1}^d p^i = 1\}$  (or a subset thereof).

**Definition 1 (coordinate separable barrier regularizers)** We say  $R: \mathcal{K} \rightarrow \mathbb{R}$  is a coordinate separable barrier<sup>2</sup> regularizer with parameter  $\nu \geq 1$  (or simply a  $\nu$ -barrier) if there exists a twice-differentiable function  $r: [0, 1] \rightarrow \mathbb{R}$  and  $c_1, c_2 > 0$  such that:

$$R(w) = \sum_{i=1}^d r(w^i), \quad \text{and} \quad \frac{c_1}{x^\nu} \leq r''(x) \leq \frac{c_2}{x^\nu} \quad \text{for all } x \in (0, 1].$$

These conditions ensure that the regularizer imposes a barrier-like growth as components of  $w$  approach zero, which plays a crucial role in controlling the optimization dynamics near the boundary of the positive orthant. This class captures several widely used regularizers, including:

- *Negative Entropy:*  $r(x) = x \log x$ , for which  $\nu = 1$ ;
- *Tsallis Entropy:*  $r(x) = \frac{x-x^q}{1-q}$  for  $q \in (0, 1)$ , where  $1 < \nu < 2$ ;
- *Log-Barrier:*  $r(x) = -\log x$ , which corresponds to  $\nu = 2$ .

The parameter  $\nu$  will turn out to be directly associated with the robustness of OMD with  $\nu$ -barrier regularization to approximation errors.

**Additional notation.** We denote by  $\ell_{t_1:t_2} = \sum_{t=t_1}^{t_2} \ell_t$  the cumulative loss vector over the interval  $[t_1, t_2]$ . For any vector  $v \in \mathbb{R}^d$ , we use  $v^i$  to denote its  $i$ -th coordinate. For example,  $\ell_t^i$  refers to the  $i$ -th component of the loss vector at time  $t$ , and  $w_t^i$  denotes the  $i$ -th component of the learner's decision at time  $t$ .

### 3. Smooth regularizers

We begin by establishing tight upper and lower bounds for approximate OMD with smooth regularizers,<sup>3</sup> over an arbitrary convex domain  $\mathcal{K} \subseteq \mathbb{R}^d$ . Our first theorem provides an upper bound that builds on the following key property of smooth functions: approximate minimization implies that first-order optimality conditions hold up to an error proportional to the square root of the sub-optimality times the smoothness parameter. The formal proofs for this section is given in Appendix B.

**Theorem 2** Let  $\mathcal{K} \subseteq \mathbb{R}^d$  be a convex set with diameter  $D$ , and let  $R: \mathcal{K} \rightarrow \mathbb{R}$  be a  $\beta$ -smooth regularizer over  $\mathcal{K}$ . Then, for any loss sequence  $\ell_1, \dots, \ell_T$  such that  $\|\ell_t\|_* \leq 1$  for all  $t \in [T]$ , the regret of any  $\varepsilon$ -approximate OMD trajectory with  $\varepsilon \leq D^2/2$  compared to any  $w \in \mathcal{K}$  is bounded as:

$$\text{Regret}(w) = O\left(\frac{1}{\eta} D_R(w, w_1) + T\eta + \frac{TD\sqrt{\beta\varepsilon}}{\eta}\right).$$

The proof follows the standard OMD analysis, replacing exact optimality with *approximate* optimality conditions. Indeed, for any  $\beta$ -smooth convex objective  $f: \mathcal{K} \rightarrow \mathbb{R}$ , if  $f(\hat{w}) - \arg \min_{w \in \mathcal{K}} f(w) \leq \varepsilon$ , then one can show that (see Lemma 22):

$$|\langle \nabla f(\hat{w}), w - \hat{w} \rangle| \leq D\sqrt{2\beta\varepsilon}. \tag{2}$$

2. Strictly speaking, these are barriers for the positive orthant in  $\mathbb{R}^d$ .

3. A function  $R$  is said to be  $\beta$ -smooth with respect to a norm  $\|\cdot\|$  if its gradient is  $\beta$ -Lipschitz;  $\|\nabla R(x) - \nabla R(y)\|_* \leq \beta\|x - y\|$  for all  $x, y \in \mathcal{K}$ , where  $\|\cdot\|_*$  is the norm dual to  $\|\cdot\|$ .

Applying the above on  $\phi_t$  for every  $t$ , and carrying the errors in the standard OMD analysis, gives the claimed result.

We note that Theorem 2 provides sharper dependence on  $\beta$  compared to a similar result of Choi et al. (2023). This bound is in fact tight, even in the simple case of OMD with Euclidean regularization and constant losses, as shown next.

**Theorem 3** *Let  $\beta, \varepsilon, D > 0$ , and consider  $\varepsilon$ -approximate OMD over  $\mathcal{K} = [0, D]$  with the  $\beta$ -smooth regularizer  $R(\cdot) = \frac{\beta}{2} \|\cdot\|_2^2$ . Then there exists a loss sequence, an  $\varepsilon$ -approximate OMD trajectory and  $w \in \mathcal{K}$  such that:*

$$\text{Regret}(w) = \Omega\left(\min\left\{\frac{TD\sqrt{\beta\varepsilon}}{\eta}, DT\right\}\right).$$

To see why this is true, consider the constant loss sequence  $\ell_t = \min\{\sqrt{2\beta\varepsilon}/\eta, 1\}$  for all  $t \in [T]$ , and initialize the trajectory at  $w_1 = D/2$ . Then for every  $t$ , the loss is small enough so that  $w_t$  itself is an  $\varepsilon$ -minimizer of  $\phi_t$ ; let  $w_{t+1}^*$  be the exact minimizer of  $\phi_t$ , then by direct computation:

$$\phi_t(w_{t+1}^*) = \eta \langle \ell_t, w_t \rangle - \varepsilon = \phi_t(w_t) - \varepsilon.$$

As a result, the approximation error might prevents any update from changing the iterate, so the trajectory remains fixed at  $w_t = w_1$  for all  $t$ . Consequently, the algorithm incurs the claimed regret. We note that the underlying reason the above argument works is that for the Euclidean regularizer, in the setting of Theorem 3, round  $t$  approximate optimality conditions (Eq. 2) are in fact tight.

## 4. Barrier regularizers

### 4.1. Adversarial losses

We next consider barrier regularizers the smoothness of which is not bounded uniformly over the domain  $\mathcal{K}$ . Indeed, the spectrum of the Hessian of any  $\nu$ -barrier (Definition 1) is unbounded since  $r''(x) \rightarrow \infty$  as  $x \rightarrow 0$ . Interestingly, the robustness behavior of these barriers varies dramatically with  $\nu$ : for negative entropy ( $\nu = 1$ ), exponentially small errors are required, whereas for log-barrier or Tsallis regularizers ( $\nu > 1$ ), polynomially small errors suffice. We begin with our lower bound for negative entropy given below.

**Theorem 4** *Let  $\mathcal{K} = \Delta_d$ ,  $d \geq 2$ , and  $R$  be the negative entropy over  $\mathcal{K}$ . Suppose that the approximation error satisfies  $\varepsilon \geq 4\eta e^{-\eta T/3}$ . Then there exists a sequence of losses  $\ell_1, \dots, \ell_T \in [0, 1]^d$  for which there exists an  $\varepsilon$ -approximate OMD trajectory that suffers regret  $\text{Regret}(w^*) = \Omega(T)$ .*

The key idea in the analysis of Theorem 4 is to exploit the fact that the *effective* smoothness of the regularizer—informally, the exact smoothness parameter on a given region—diverges at a rate inversely proportional to the iterate coordinates as they approach zero. Indeed, our construction is such that the coordinates of the iterate become as small as  $e^{-\eta T}$  (this follows from the closed form update equations), and thus reach the region of the domain where the effective smoothness is exponentially large. Then, when the errors are not exponentially small, the same mechanism as in the smooth-regularizer lower bound applies: the iterate becomes stuck even under constant losses, leading to linear regret. A similar argument also shows that  $\varepsilon$  must be polynomially small in  $d$ ; otherwise, the iterates can remain stuck at the initialization point (see Theorem 35). The exponential dependence of  $\varepsilon$  on the time horizon  $T$  is in fact tight: if  $\varepsilon$  is exponentially small in  $\eta T$ , the standard regret guarantees are recovered.

**Theorem 5** *Let  $\mathcal{K} = \Delta_d$  and  $R$  be the negative entropy over  $\mathcal{K}$ . Assume  $\eta \leq 1/16$  and  $T \geq 3$ , if  $\varepsilon \leq \frac{1}{6d}e^{-\eta T/2} \min\{\eta^4, T^{-2}\}$ , then for any loss sequence  $\ell_1, \dots, \ell_T \in [-1, 1]^d$ , the regret of any  $\varepsilon$ -approximate OMD trajectory compared to any  $w \in \mathcal{K}$  is bounded as:*

$$\text{Regret}(w) \leq \frac{1}{\eta} D_R(w, w_1) + O(\eta T),$$

where big-O hides only constant numerical factors.

The proof is deferred to Section 5.

We now turn our attention to  $\nu$ -barrier regularizers with  $\nu > 1$ . In this case, as it turns out, polynomially small errors suffice to *naturally* keep the iterates bounded away from zero (by a polynomial margin).

**Theorem 6** *Let  $\mathcal{K} \subseteq \Delta_d$  be a polytope that contains the uniform distribution and the OMD is initialized there,<sup>4</sup> let  $R: \mathcal{K} \rightarrow \mathbb{R}$  be a  $\nu$ -barrier regularizer (cf. Definition 1) with  $\nu > 1$  and  $\eta \leq 1/(16c_1)$ . If*

$$\varepsilon \leq \eta^4 \min\left\{\frac{1}{c_2}, c_2\right\} \left(\frac{16\eta T d}{c_1} + 2(2d)^{\nu-1}\right)^{-\frac{\nu}{\nu-1}},$$

then for any loss sequence  $\ell_1, \dots, \ell_T \in [-1, 1]^d$ , the regret of any  $\varepsilon$ -approximate OMD trajectory compared to any  $w \in \mathcal{K}$  is bounded as:

$$\text{Regret}(w) \leq \frac{1}{\eta} D_R(w, w_1) + O(\eta T).$$

The principle underlying the analysis of Theorem 6 is as follows. Consider for purposes of illustration the one-dimensional interval  $[0, 1]$  with  $w_1 = 1$ . In this setting the OMD updates require no projection and the iterate dynamics can be inspected more simply:

$$r'(w_t) = r'(w_{t-1}) - \eta \ell_{t-1} = r'(w_1) - \eta \sum_{s=1}^{t-1} \ell_s,$$

which implies

$$-\frac{1}{w_t^{\nu-1}} \geq -1 - \eta T \quad \implies \quad w_t \geq (\eta T)^{-\frac{1}{\nu-1}}.$$

Namely, the iterates can only shrink polynomially in  $T$ , and as a result the effective smoothness grows polynomially. This allows the use of approximate first-order optimality conditions in the standard OMD analysis, and the regret may be bounded using the standard OMD proof. Note that this comes in contrast to the negative entropy case ( $\nu = 1$ , Theorem 4) where a similar argument in this simplified setting gives  $\log(w_t) \geq 0 - \eta T \implies w_t \geq e^{-\eta T}$ . Finally, the simplified setting considered above we did not account for the possibility that the errors themselves can pull the iterates closer to the boundary. Evidently, the approximation errors may potentially drive the iterates toward zero even when the exact dynamics would not, which further complicates the analysis.

---

4. This assumption serves mainly to fix a natural starting point for OMD; a similar bound should hold for any reasonable initialization.

## 4.2. Improved robustness with stochastic losses

In the adversarial setting, we have seen that negative entropy requires exponentially small error to avoid linear regret, even over the simplex. Surprisingly, this fragility does not persist for stochastic losses over the full simplex. For i.i.d. stochastic losses, polynomially small approximation errors suffice to guarantee standard regret bounds with high probability.

**Theorem 7** *Let  $\mathcal{K} = \Delta_d$  and  $R$  be the negative entropy over  $\mathcal{K}$ . For any  $\delta > 0$ , suppose that  $w_1 = (1/d, 1/d, \dots, 1/d)$ ,  $T \geq 256$ ,  $\eta = \sqrt{\frac{\log(d)}{T}}$  and  $\varepsilon \leq \frac{\delta}{6d^2T^4}$ . Then with probability  $\geq 1 - \delta$  over the choice of an i.i.d. loss sequence  $\ell_1, \dots, \ell_T \in [-1, 1]^d$  the regret of any  $\varepsilon$ -approximate OMD trajectory compared to any  $w \in \mathcal{K}$  is  $O(\sqrt{T \log(d)})$ .*

However, this robustness does not extend to general domains. Even with the same regularizer and similarly stochastic losses, restricting the domain to a polyhedral subset of the simplex can cause suboptimal regret unless the approximation error is exponentially small.

**Theorem 8** *Consider approximate OMD with the negative entropy regularizer and stochastic losses. Then, there exists a polytope  $\mathcal{K} \subseteq \Delta_d$  and a distribution of losses such that for any  $\varepsilon > 0$ , there exists an  $\varepsilon$ -approximate trajectory and  $w \in \mathcal{K}$  such that:*

$$\mathbb{E}[\text{Regret}(w)] = \Omega\left(\frac{D_R(w, w_1)}{\eta} + T \sqrt{\frac{\eta}{\log(1/\varepsilon)}}\right).$$

One can see that any approximation error that is merely polynomial in  $T$  leads to a sub-optimal regret lower bound of  $\Omega(T^{2/3})$ , even under an optimally tuned learning rate.

## 5. Analysis overview

In this section we sketch the proofs of the results from Section 4. We begin in Section 5.1 by introducing the balance framework, which serves as a unifying tool throughout the analysis. The balance of an OMD trajectory quantifies how “well-behaved” it is, measuring how much noise or fluctuation it exhibits. Our framework provides analytic tools for computing or bounding the balance of a trajectory, relating it to a certain notion of balance of the loss sequence, and further translating balance to other properties of the trajectory.

Next, we apply the balance framework in two ways. In Section 5.2 we show that for barrier regularizers other than negative entropy ( $\nu > 1$ ), the iterates cannot approach zero too closely, which directly yields a polynomial-error upper bound (Theorem 6). Later, in Section 5.3, we analyze the negative entropy regularizer and establish a relationship between regret and balance of the loss sequence (Lemmas 14 and 15), leading to our main results for the adversarial (Theorems 4 and 5) and stochastic (Theorem 7) settings. Finally, the proof of Theorem 8 is given in Section 5.4.

### 5.1. Balance

To analyze the trajectory of OMD over general polytopes we introduce the notion of *Balance*. We assume the polytope is given in standard form:

$$\mathcal{K} = \{w \in \mathbb{R}^d : Aw = b, w_i \geq 0 \forall i \in [d]\},$$

where  $A \in \mathbb{R}^{m \times d}$  with  $m < d$  and  $b \in \mathbb{R}^m$  (see, e.g., Eq. 4.28 in [Boyd and Vandenberghe, 2004](#)).

**Definition 9** For every  $v \in \ker(A)$  and  $1 \leq t_1 < t_2 \leq T$  we define the balance of an OMD trajectory  $w_1, \dots, w_T \in \mathcal{K}$  with respect to  $t_1, t_2, v$  as follows:

$$B^v(t_1, t_2) = \langle \nabla R(w_{t_1}) - \nabla R(w_{t_2}), v \rangle.$$

If for every  $v \in \ker(A)$  and every  $t_1, t_2$  we have  $B^v(t_1, t_2) \leq k\|v\|$ , we say the trajectory is  $k$ -balanced w.r.t. the norm  $\|\cdot\|$ .

Our first lemma relates the variation of the loss sequence to the balance of the OMD iterates, which in turn will be used to establish properties of the trajectory leading to regret upper or lower bounds.

**Lemma 10** Assume the OMD trajectory is exact, then for every  $v \in \ker(A)$  and times  $t_1, t_2$ :

$$B^v(t_1, t_2) = \eta \langle \ell_{t_1:t_2}, v \rangle.$$

This motivates the definition that a sequence of losses  $\{\ell_t\}_{t=1}^T$  is  $\alpha$ -balanced w.r.t norm  $\|\cdot\|$  if for every  $v \in \ker(A)$  we have  $\langle \ell_{t_1:t_2}, v \rangle \leq \alpha\|v\|$ . It is immediate to verify that when this holds, the exact OMD trajectory is  $(\eta\alpha)$ -balanced. Notably, by working with the notion of loss balance, we obtain results that are later applicable to both the stochastic and adversarial settings.

The relation between loss balance and *approximate* trajectories ( $\varepsilon > 0$ ) is more nuanced, since the errors we want to control naturally scale with the smoothness parameter of the objective (which is unbounded in our case). To cope with this we introduce the notion of *effective smoothness*, which roughly corresponds to the smoothness parameter associated with the line segment between the exact and approximate OMD updates. Under the assumption that the iterates remain bounded away from zero, the effective smoothness remains finite and we may bound the difference between the balance of the exact and approximate OMD trajectories, as stated in our next lemma.

**Lemma 11** Let  $\{w_t\}_{t=1}^T, \{\hat{w}_t\}_{t=1}^T$  be an exact trajectory and an approximate trajectory with the same  $\nu$ -barrier regularizer, losses and  $\eta$ . Fix  $0 \leq t_1 \leq t_2 \leq T$  and  $v \in \ker(A)$  such that  $\|v\|_1 = 1$ . Let  $\psi > 0$  be such that for every  $t_1 \leq t \leq t_2$  and  $i \in [d]$  such that  $v^i \neq 0$ ,  $w_t^i \geq \psi$  and  $\varepsilon \leq c_2\psi/2$ . Then, we have:

$$\hat{B}^v(t_1, t_2) \leq B^v(t_1, t_2) + (t_2 - t_1) \sqrt{\frac{c_2\varepsilon}{\psi^\nu}},$$

where  $\hat{B}$  is the balance of the approximate trajectory.

Next, we introduce machinery that facilitates arguments going in the other direction; namely, that a balanced trajectory remains bounded away from zero. Our lemma below generalizes the argument given after Theorem 6 and bounds the first derivative of the regularizer in terms of the balance of the OMD trajectory. The bound on the first derivative may in turn be used to yield a bound on the actual iterate coordinates (such as in the special case discussed in the paragraph after Theorem 6).

**Lemma 12** Let  $\mathcal{K}$  be a polyhedral subset of the simplex. Assume the trajectory is  $k$ -balanced w.r.t to the  $L_1$ -norm and was initialized at the uniform distribution. Then, for every  $t \in [T], i \in [d]$ :

$$-r'(w_t^i) \leq 4kd - r'(1/2d).$$

When  $\mathcal{K}$  is the simplex, note that for every  $i, j \in [d]$ , the vector  $e_i - e_j$  belongs to  $\ker(A)$ , where  $e_i$  denotes the  $i$ th standard basis vector. Let  $i^*$  denote the optimal arm (the coordinate with the smallest cumulative loss). We write  $B^i$  to denote the balance with respect to the vector  $e_{i^*} - e_i$ . With this notation, we can state an additional lemma, relevant only when the decision space is the simplex, that bounds the iterate as a function of the balance.

**Lemma 13** *Let  $\mathcal{K} = \Delta_d$  and suppose that for some  $i \in [d]$  and  $t_1, t_2 \in [T]$  we have  $B^i(t_1, t_2) \leq k$ . Then*

$$w_{t_2}^i \geq w_{t_1}^i \Rightarrow w_{t_1}^{i*} \leq e^{k/c_1} w_{t_2}^{i*}, \quad w_{t_2}^{i*} \leq w_{t_1}^{i*} \Rightarrow w_{t_2}^i \leq e^{k/c_1} w_{t_1}^i.$$

These results allow us to translate control over balance into control over how far the coordinates of the trajectory can drift, which will be crucial in the later proofs. The full proofs for the lemmas in this section can be found in Appendix C.

### 5.2. Non-entropy barriers: Proof of Theorem 6

We first handle the case of non-entropy barrier regularizers ( $\nu > 1$ ), before turning in the next subsection to the negative entropy regularizer, which requires a separate treatment. The idea is to show that barrier regularizers with  $\nu > 1$  cannot drive the iterates exponentially close to zero, and as a result the relevant effective smoothness parameter grows only polynomially with  $T$ . Consequently, we can apply an argument similar to the regret bound in the smooth regularizer case from Theorem 2 to control the additional regret due to approximation errors. The full proof can be found in Appendix F.

**Proof of Theorem 6 (sketch)** Let

$$\psi := \left( \frac{c_1}{8\eta Td + c_1(2d)^{\nu-1}} \right)^{1/(\nu-1)}.$$

We prove by induction on  $t \in [T]$  that all coordinates remain bounded away from zero, namely  $w_t^i \geq \psi$  for all  $i \in [d]$ . Assume the claim holds up to step  $t - 1$ . Using Lemma 25, which bounds the step size of each iterate, we first show that  $w_t^i \geq \psi/2$ . Since the balance of an exact trajectory is always bounded by  $T\eta$ , Lemma 11 implies that for every normalized  $v \in \ker(A)$ ,

$$B^v(1, t) \leq T\eta + T\sqrt{c_2\varepsilon(2/\psi)^\nu} \leq 2T\eta.$$

Hence, the trajectory up to step  $t$  is  $2T\eta$ -balanced.

Applying Lemma 12, we have

$$-r'(w_t^i) \leq 8\eta Td - r'(1/2d) \quad \Longrightarrow \quad \frac{c_1}{(w_t^i)^{\nu-1}} \leq 8\eta Td + c_1(2d)^{\nu-1} \quad \Longrightarrow \quad w_t^i \geq \psi.$$

The first implication follows from Lemma 41, which provides a lower bound on the gradient difference for barrier regularizers. This completes the inductive step and establishes the result.  $\blacksquare$

### 5.3. Negative entropy over the simplex

With the negative entropy regularizer, robustness to approximation errors differs sharply between adversarial and stochastic settings. In this case, the iterates can approach zero exponentially fast, causing the effective smoothness to grow exponentially. However, under stochastic losses, the iterates do not become stuck despite approaching zero. The key lies in the balance of the loss sequence: while an arbitrary adversarial sequence can have balance as large as  $T$ , for i.i.d. stochastic losses the balance is bounded with high probability by roughly  $\sqrt{T}$ .

To capture this distinction, we establish two general lemmas that characterize the regret behavior as a function of the balance parameter. The first lemma shows that if the approximation error exceeds an exponential (in the balance) threshold, then linear regret can occur.

**Lemma 14** *Let  $d = 2$ ,  $\mathcal{K} = \Delta_d$  be the simplex, and let  $R(w) = \sum_{i=1}^d w_i \log w_i$  be the negative entropy regularizer. Then, for any  $\alpha \leq T/2$  there exists a sequence of  $\alpha$ -balanced losses  $\ell_1, \dots, \ell_T \in [0, 1]^d$  such that for any  $\varepsilon \geq 4\eta e^{-\eta\alpha}$ , there exists a trajectory that is  $\varepsilon$ -approximate w.r.t.  $(\mathcal{K}, R, \ell_{1..T})$  and has regret  $\Omega(T - 2\alpha)$ .*

The proof idea is that the approximation errors are large enough so that the iterate coordinates may reach the region close to zero where the iterate can become “stuck” due to additional subsequent errors (this idea was explained thoroughly after Theorem 4). This Lemma is the principle technical gradient in the proof of Theorem 4. Conversely, the second lemma shows that if the error is below this exponential threshold, standard regret bounds hold:

**Lemma 15** *Let  $\mathcal{K} = \Delta_d$  be the simplex, let  $\{\ell_t\}_{t=1}^T$  be  $\alpha$ -balanced loss sequence, and let  $R(w) = \sum_{i=1}^d w_i \log w_i$  be the negative entropy regularizer. Assume  $\eta \leq 1/16$  and  $T \geq 3$ , if the approximation error satisfies*

$$\varepsilon \leq \frac{1}{d \max\{6e^{\eta\alpha}, 1/\eta\}} \min\{\eta^4, 1/T^2\},$$

*then the regret of any  $\varepsilon$ -approximate OMD trajectory is bounded by  $\frac{1}{\eta} D_R(w, w_1) + O(T\eta)$ .*

**Proof (sketch)** Let  $i^*$  be the coordinate with the smallest cumulative loss. We prove the claim in two steps.

1. If the optimal arm coordinate  $i^*$  of all iterates is bounded away from zero, i.e.  $\forall t, w_t^{i^*} \geq \xi$ , and in addition  $\varepsilon \lesssim \eta^4 \xi$ , the regret bound follows (Lemma 39).
2. If the losses are  $\alpha$ -balanced and  $\varepsilon \leq \xi/(2T^2)$  for  $\xi = 1/(de^{\alpha\eta+1})$ , then the optimal arm coordinate is bounded away from zero throughout the trajectory, i.e.,  $\forall t, w_t^{i^*} \geq \xi$  (Lemma 40).

**Step 1.** In the classical OMD analysis, first-order optimality conditions are applied at every step. The gap in these conditions depends on the effective smoothness, which in turn reflects how close the coordinates are to zero. Since some coordinates may take very small values, we apply the optimality conditions only to those with  $w_t^i \geq \xi$ . Using a careful argument—based on the observation that coordinates close to zero contribute little to the overall regret—we extend the proof to all coordinates. An additional challenge arises because the set of small coordinates changes over time, which we handle using the monotonicity of the Bregman divergence (see Lemma 37)

**Step 2.** We now prove (2) by induction. Assume, for contradiction, that some step  $t$  is the first to have  $w_t^{i^*} < \xi$ . Then there must exist a coordinate  $i \neq i^*$  such that  $w_t^i \geq 1/d$ . We first show that for every  $s < t$ ,  $w_s^i \geq \xi$ . Suppose not, and let  $s$  be the last time for which  $w_s^i < \xi$ . We use Lemma 25, which bounds the step size of each iterate, to first establish that  $w_s^i \geq \xi/2$ . From Lemma 11,

$$B^i(s, t) \leq \alpha\eta + T\sqrt{r''(\xi/2)}\varepsilon = \alpha\eta + 1.$$

Applying the second part of Lemma 13 with  $(i, s, t)$  yields  $w_s^i \geq \xi$ , a contradiction. Hence,  $w_s^i \geq \xi$  for all  $s \leq t$ . We then use this to bound the balance from the beginning:

$$B^i(1, t) \leq \alpha\eta + T\sqrt{r''(\xi/2)}\varepsilon = \alpha\eta + 1.$$

Since  $w_t^i \geq w_1^i$ , the first part of Lemma 13 implies that  $w_t^{i^*} \geq \xi$ , completing the induction. ■

Together, these two lemmas provide a clean characterization: linear regret is unavoidable once the error exceeds an exponential threshold in  $\eta\alpha$ , while below this threshold optimal regret guarantees are preserved. Before applying the lemmas, let us note that when the polytope is the simplex itself, the vectors  $e_i - e_{i^*}$  for every  $i \in [d]$  form a basis of  $\ker(A)$ . Thus, if

$$\ell_{t_1:t_2}^i - \ell_{t_1:t_2}^{i^*} \leq \alpha \quad \forall i \in [d],$$

the loss sequence is  $\alpha$ -balanced. We now apply the lemmas to derive the main theorems.

**Proof of Theorem 4** Directly by applying Lemma 14 with  $\alpha = T/3$ . ■

**Proof of Theorem 5** Any adversarial sequence over the simplex is  $T/2$ -balanced: if one coordinate exceeds the best by more than  $T/2$ , it must actually be the best. Applying Lemma 15 with  $\alpha = T/2$  gives the desired upper bound. ■

**Proof of Theorem 7** For i.i.d. losses, Hoeffding’s inequality and union bounds implies that with probability at least  $1 - \delta$ , the balance is at most  $\alpha = O(\sqrt{T \log(dT^2/\delta)})$ , which means that  $\eta\alpha = O(\log(dT^2/\delta))$ . Plugging this into Lemma 15 together with the fact that  $D_R(w, w_1) \leq \log(d)$  for all  $w$  yields the stochastic upper bound. ■

#### 5.4. Negative entropy: lower bound with stochastic losses

Theorem 8 shows that when the decision set is a polytope subset of the simplex and the regularizer is negative entropy, approximation errors can lead to suboptimal regret unless  $\varepsilon$  is exponentially small even if the losses are stochastic. The key difference from the simplex is that we can design the constraints so that many coordinates lie in a single kernel direction, thus forcing them to move only together. Below we provide a high level sketch of the construction; see Appendix G for the detailed proof.

**Proof of Theorem 8 (sketch)** We begin by constructing a set  $S_0$  of  $\Theta(\log \frac{1}{\varepsilon})$  coordinates that have mean zero loss (with variance 1) and move in the same direction. We denote two additional coordinates by  $j$  and  $i$ ; coordinate  $j$  has *positive* mean loss of  $\Theta(\sqrt{\eta \log(1/\varepsilon)})$  and moves in the same direction as the aforementioned set, and coordinate  $i$  has zero loss but moves in the *opposite* direction of all others.

Since the set  $S_0 \cup \{j, i\}$  moves as a unit, only the sum of their losses influence the trajectory. An anti-concentration argument (see Lemma 45) shows that after  $\tau = \Theta(\log \frac{1}{\varepsilon})$  rounds, with constant probability this cumulative loss is non-positive. This, along with an additional component (Lemma 46) can be used to argue that  $S_0$  moves in the positive direction (i.e., away from zero; see Lemma 47), thus pushing  $i$  in the negative direction closer to zero. Importantly, the negative-entropy geometry forces the coordinate  $i$  to shrink exponentially fast, reaching  $w_t^i \approx \varepsilon$  after only  $t = \tau$  steps. Once this happens, approximation errors keep  $w_t^i \approx \varepsilon$  for all  $t > \tau$ , using the same stuck-iterate argument as in earlier lower bounds (see Lemma 48). As a result, coordinate  $j$  remains at value about  $1/\log(1/\varepsilon)$  for  $\Theta(T)$  rounds, leading to  $\Omega(T\sqrt{\eta/\log(1/\varepsilon)})$  regret. ■

## 6. Discussion

This work provides an analysis of how approximation errors affect Online Mirror Descent. We establish tight upper and lower regret bounds for smooth regularizers, showing that polynomially small errors suffice to maintain optimal regret. Moving beyond smoothness, we uncover a sharp

separation among barrier-type regularizers: with negative entropy, exponentially small errors are necessary to avoid linear regret, whereas log-barrier and Tsallis regularizers remain robust even with polynomially large errors. We further show that while negative entropy regains robustness under stochastic losses on the full simplex, this property fails on certain polyhedral subsets. Altogether, our results reveal a fundamental sensitivity of OMD to approximation accuracy, determined jointly by the geometry of the domain, the curvature of the regularizer, and the structure of the loss sequence. Furthermore, our work provides a detailed characterization of when precision is essential and when it is not.

A broader goal emerging from this work, left for future investigation, is to develop a comprehensive theory of inexact OMD for general regularizers and geometries. In particular, it would be valuable to characterize the robustness properties of self-concordant barrier regularizers over general convex domains.

### Acknowledgments

This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreements No. 101078075; 882396). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. This work received additional support from the Israel Science Foundation (ISF, grant numbers 3174/23; 1357/24), and a grant from the Tel Aviv University Center for AI and Data Science (TAD). This work was partially supported by the Deutsch Foundation. OS is also partially supported by the TAD Excellence Program for Doctoral Students in Artificial Intelligence and Data Science from the Tel Aviv University Center for AI and Data Science (TAD) and from the Israeli Council for Higher Education (CHE) Fellowship for Outstanding PhD Students in Data Science.

## References

- Jacob Abernethy and Alexander Rakhlin. Beating the adaptive bandit with high probability. In *2009 Information Theory and Applications Workshop*, pages 280–289, 2009. doi: 10.1109/ITA.2009.5044958.
- Jacob D Abernethy, Chansoo Lee, and Ambuj Tewari. Fighting bandits with a new kind of smoothness. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc., 2015.
- Masoud Ahookhosh and Yurii Nesterov. High-order methods beyond the classical complexity bounds: inexact high-order proximal-point methods. *Mathematical Programming*, 208(1):365–407, 2024.
- Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications. *Theory of computing*, 8(1):121–164, 2012.
- Heinz H Bauschke, Jérôme Bolte, and Marc Teboulle. A descent lemma beyond lipschitz gradient continuity: first-order methods revisited and applications. *Mathematics of Operations Research*, 42(2):330–348, 2017.
- Amir Beck. *First-order methods in optimization*. SIAM, 2017.
- Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 31(3):167–175, 2003.
- Stephen Boyd and Lieven Vandenberghhe. *Convex Optimization*. Cambridge University Press, 2004.
- Sébastien Bubeck. Convex optimization: Algorithms and complexity. *Foundations and Trends® in Machine Learning*, 8(3-4):231–357, 2015.
- Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.
- Woocheol Choi, Myeong-Su Lee, and Seok-Bae Yun. Inexact online proximal mirror descent for time-varying composite optimization. *arXiv preprint arXiv:2304.04710*, 2023.
- Travis Dick, András György, and Csaba Szepesvári. Online learning in markov decision processes with changing cost sequences. In *Proceedings of the 31st International Conference on International Conference on Machine Learning - Volume 32, ICML’14*, page I–512–I–520. JMLR.org, 2014.
- Rishabh Dixit, Amrit Singh Bedi, Ruchi Tripathi, and Ketan Rajawat. Online learning with inexact proximal online gradient descent algorithms. *IEEE Transactions on Signal Processing*, 67(5):1338–1352, 2019.
- Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1):119–139, 1997.
- Vincent Guigues. Inexact stochastic mirror descent for two-stage nonlinear stochastic programs. *Mathematical Programming*, 187(1):533–577, 2021.
- Elad Hazan. Introduction to online convex optimization. *Foundations and Trends® in Optimization*, 2(3-4):157–325, 2016.

- Elad Hazan and Kfir Y. Levy. Bandit convex optimization: towards tight bounds. In *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'14, page 784–792, Cambridge, MA, USA, 2014. MIT Press.
- Shinji Ito. An optimal algorithm for bandit convex optimization with strongly-convex and smooth loss. In Silvia Chiappa and Roberto Calandra, editors, *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pages 2229–2239. PMLR, 26–28 Aug 2020.
- Chi Jin, Tiancheng Jin, Haipeng Luo, Suvrit Sra, and Tiancheng Yu. Learning adversarial Markov decision processes with bandit feedback and unknown transition. In Hal Daumé III and Aarti Singh, editors, *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 4860–4869. PMLR, 13–18 Jul 2020.
- S Kabbadj. Inexact version of bregman proximal gradient algorithm. In *Abstract and Applied Analysis*, volume 2020, page 1963980. Wiley Online Library, 2020.
- Nick Littlestone and Manfred K Warmuth. The weighted majority algorithm. *Information and computation*, 108(2):212–261, 1994.
- Haihao Lu, Robert M Freund, and Yurii Nesterov. Relatively smooth convex optimization by first-order methods, and applications. *SIAM Journal on Optimization*, 28(1):333–354, 2018.
- Haipeng Luo, Chen-Yu Wei, and Kai Zheng. Efficient online portfolio with logarithmic regret. *Advances in neural information processing systems*, 31, 2018.
- Arkadij Semenovič Nemirovskij and David Borisovich Yudin. Problem complexity and method efficiency in optimization, 1983.
- A. S. Nemirovsky and D. B. Yudin. *Problem Complexity and Method Efficiency in Optimization*. Wiley-Interscience, New York, 1983. Translated and edited by E. R. Dawson.
- Simone Rebegoldi, Silvia Bonettini, and Marco Prato. A bregman inexact linesearch-based forward-backward algorithm for nonsmooth nonconvex optimization. In *Journal of Physics: Conference Series*, volume 1131, page 012013. IOP Publishing, 2018.
- R Tyrrell Rockafellar. Monotone operators and the proximal point algorithm. *SIAM journal on control and optimization*, 14(5):877–898, 1976.
- Aviv Rosenberg and Yishay Mansour. Online convex optimization in adversarial Markov decision processes. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 5478–5486. PMLR, 09–15 Jun 2019.
- Mark Schmidt, Nicolas Le Roux, and Francis Bach. Convergence rates of inexact proximal-gradient methods for convex optimization. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 24, pages 1458–1466, 2011.
- Shai Shalev-Shwartz. Online learning and online convex optimization. *Foundations and Trends® in Machine Learning*, 4(2):107–194, 2012.

- Fedor Stonyakin, Alexander Tyurin, Alexander Gasnikov, Pavel Dvurechensky, Artem Agafonov, Darina Dvinskikh, Mohammad Alkousa, Dmitry Pasechnyuk, Sergei Artamonov, and Victorya Piskunova. Inexact model: A framework for optimization and variational inequalities. *Optimization Methods and Software*, 36(6):1155–1201, 2021.
- Marc Teboulle. A simplified view of first order methods for optimization. *Mathematical Programming*, 170(1):67–96, 2018.
- Silvia Villa, Saverio Salzo, Luca Baldassarre, and Alessandro Verri. Accelerated and inexact forward-backward algorithms. *SIAM Journal on Optimization*, 23(3):1607–1633, 2013.
- Lei Yang and Kim-Chuan Toh. Inexact bregman proximal gradient method and its inertial variant with absolute and partial relative stopping criteria. *Mathematics of Operations Research*, 2025.
- Qiang Zhou and Sinno Jialin Pan. On acceleration for convex composite minimization with noise-corrupted gradients and approximate proximal mapping. *Journal of Machine Learning Research*, 23(223):1–59, 2022.
- Alexander Zimin and Gergely Neu. Online learning in episodic markovian decision processes by relative entropy policy search. In C.J. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 26. Curran Associates, Inc., 2013.
- Julian Zimmert and Yevgeny Seldin. Tsallis-inf: An optimal algorithm for stochastic and adversarial bandits. *Journal of Machine Learning Research*, 22(28):1–49, 2021.
- Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th international conference on machine learning (icml-03)*, pages 928–936, 2003.

## Appendix Structure

Appendix A provides general definitions and Lemmas used throughout the appendix. Appendix B contains the proofs for Section 3. Appendix C introduces the notion of *balance*, which is needed for the subsequent proofs. The remaining appendices establish the main technical arguments of the paper: Appendix D contains lower bounds for negative entropy (including the proof of Lemma 14 and additional results), Appendix E contains the proof of Lemma 15, Appendix F contains the proof of Theorem 6, and Appendix G contains the proof of Theorem 8.

### Appendix A. General Lemmas and definitions

**Definition 16** We call  $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$  an **exact OMD trajectory** if for every  $t \in [T]$ :

$$w_{t+1} = \operatorname{argmin}_{w \in \Delta_d} \eta \langle \ell_t, w \rangle + D_R(w_{t-1}, w)$$

**Definition 17** We call  $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$  an  $\varepsilon$ -**approximate OMD trajectory** with some  $\varepsilon > 0$  if for every  $t \in [T]$   $w_{t+1}$  is an  $\varepsilon$ -minimizer of  $\eta \langle \ell_t, w \rangle + D_R(w_{t-1}, w)$ .

**Definition 18** Our assumptions about the regularizers are:

- There is a function  $r : [0, 1] \rightarrow \mathbb{R}$  such that  $R$  is coordinate-separated with  $f_i = r$  for all  $i \in [d]$
- $r''$  is decreasing polynomially in  $[0, 1]$  and  $r''(w) \geq \frac{1}{w}$  for all  $w \in [0, 1]$ .

**Definition 19** We say that a function  $F : \mathcal{W} \rightarrow \mathbb{R}$  ( $\mathcal{W} \subseteq \mathbb{R}^d$ ) is coordinate-separated if there are functions  $f_1, f_2, \dots, f_d$  such that  $F(w) = \sum_i f_i(w_i)$  for all  $w \in \mathcal{W}$ .

**Definition 20** Let  $F : \mathcal{W} \rightarrow \mathbb{R}$  be a coordinate-separated function. Let  $w^1, w^2 \in \mathcal{W}$ , we say  $\beta \in \mathbb{R}$  is the **effective smoothness** of  $F$  w.r.t  $w_1, w_2$  if for every  $i \in [d]$  such that  $w_i^1 \neq w_i^2$  and  $\alpha \in [w_i^1, w_i^2]$ , we have  $f_i''(\alpha) \leq \beta$ .

**Lemma 21** Let  $F : \mathcal{W} \rightarrow \mathbb{R}$  be a coordinate-separated function and Let  $x_1, x_2 \in \mathcal{W}$ . If  $\beta$  is the effective smoothness of  $F$  w.r.t  $x_1, x_2$  we have for any  $w_1, w_2 \in [x_1, x_2]$ :

$$F(w_1) - F(w_2) - \langle \nabla F(w_2), w_1 - w_2 \rangle \leq \frac{\beta}{2} \|w_1 - w_2\|_2^2 \leq \frac{\beta}{2} \|w_1 - w_2\|_1^2$$

**Proof** The first inequality is directly from Taylor's theorem. The second is because generally  $\|\cdot\|_2 \leq \|\cdot\|_1$ . ■

**Lemma 22** Let  $\|\cdot\|$  be any norm, and let  $f : \mathcal{W} \rightarrow \mathbb{R}$ , and let  $\hat{w}, w \in \mathcal{W}$  where  $\hat{w}$  is an  $\varepsilon$ -minimizer of  $f$ . Assume that for all  $x, y \in [w, \hat{w}]$  it holds that:

$$f(y) - f(x) - \langle \nabla f(x), y - x \rangle \leq \frac{\beta}{2} \|y - x\|^2.$$

Then, we have:

$$\langle \nabla f(\hat{w}), w - \hat{w} \rangle \geq -\max \left\{ \|w - \hat{w}\| \sqrt{2\beta\varepsilon}, 2\varepsilon \right\}$$

Additionally, let  $D = \max_{w', w'' \in \mathcal{K}} \|w' - w''\|$  and assume  $\varepsilon \leq \frac{D^2\beta}{2}$ . We have:

$$\langle \nabla f(\hat{w}), w - \hat{w} \rangle \geq -D\sqrt{2\beta\varepsilon}$$

We note that this holds for coordinate-separated function with effective smoothness  $\beta$  (with  $\ell_1$  or  $\ell_2$  norm, see Lemma 21) or any general  $\beta$ -smooth function.

**Proof** From the assumptions of the Lemma, for any  $\gamma \in [0, 1]$ :

$$\begin{aligned} f(\hat{w} + \gamma(w - \hat{w})) &\leq f(\hat{w}) + \gamma \nabla f(\hat{w})(w - \hat{w}) + \gamma^2 \frac{\beta}{2} \|w - \hat{w}\|^2 \\ \nabla f(\hat{w})(w - \hat{w}) &\geq \frac{1}{\gamma} (f(\hat{w} + \gamma(w - \hat{w})) - f(\hat{w})) - \gamma \frac{\beta}{2} \|w - \hat{w}\|^2 \\ &\geq -\left( \frac{\varepsilon}{\gamma} + \gamma \frac{\beta}{2} \|w - \hat{w}\|^2 \right) \end{aligned}$$

Notice that if  $2\varepsilon \geq \|w - \hat{w}\|\sqrt{2\beta\varepsilon}$ , we have  $\varepsilon \geq \frac{\beta}{2}\|w - \hat{w}\|^2$  thus for  $\gamma = 1$ :

$$\nabla f(\hat{w})(w - \hat{w}) \geq -\left( \varepsilon + \frac{\beta}{2} \|w - \hat{w}\|^2 \right) \geq -2\varepsilon$$

Else, for  $\gamma = \frac{\sqrt{2\varepsilon}}{\sqrt{\beta}\|w - \hat{w}\|} \leq 1$ :

$$\nabla f(\hat{w})(w - \hat{w}) \geq \|w - \hat{w}\|\sqrt{2\beta\varepsilon}$$

If  $\varepsilon \leq \frac{D^2\beta}{2}$ , we have:

$$\begin{aligned} \sqrt{\varepsilon} &\leq \frac{D\sqrt{\beta}}{\sqrt{2}} \\ &= \frac{D\sqrt{2\beta}}{2} \\ \Leftrightarrow 2\varepsilon &\leq D\sqrt{2\beta\varepsilon} \\ \Rightarrow \langle \nabla f(\hat{w}), w - \hat{w} \rangle &\geq -D\sqrt{2\beta\varepsilon} \end{aligned}$$

■

**Lemma 23** If for some  $a, b, c > 0$  we have  $ax^2 - bx - c \leq 0$ , then  $x < \frac{b}{a} + \sqrt{\frac{c}{a}}$

**Proof** Assume  $x = \frac{b}{a} + \sqrt{\frac{c}{a}}$ , we have:

$$ax^2 - bx - c = \frac{b^2}{a} + 2b\sqrt{\frac{c}{a}} + c - \frac{b^2}{a} - b\sqrt{\frac{c}{a}} - c = b\sqrt{\frac{c}{a}} > 0$$

The minimum point of the parabola is at  $x = \frac{b}{2a}$ , so it only increases for  $x > \frac{b}{a} + \sqrt{\frac{c}{a}}$ .

■

**Lemma 24** Let  $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$  be an  $\varepsilon$ -approximate trajectory above the simplex with  $\eta \leq \frac{1}{4}$  and coordinate-separable regularizer. Let  $h = \min \{r''(w_t^i), r''(w_{t+1}^i)\}$ . Then for any  $i \in [d]$ :

$$|w_t^i - w_{t+1}^i| < \frac{4\eta}{h} + \sqrt{\frac{\varepsilon}{h}}$$

**Proof** Fix  $i \in [d]$ . We will prove for  $w_{t+1}^i \leq w_t^i$ . The proof for the other direction is identical.

Let  $i_1, \dots, i_m$  be an arbitrary set of coordinates that satisfies the following. For  $S := \{i_1, \dots, i_{m-1}\}$ ,  $i' := i_m$  it holds that:

$$\forall j \in (S \cup i') \quad w_{t+1}^j \geq w_t^j \quad (3)$$

$$\sum_{j \in S} w_{t+1}^j - w_t^j < w_t^i - w_{t+1}^i \quad (4)$$

$$\sum_{j \in (S \cup i')} w_{t+1}^j - w_t^j \geq w_t^i - w_{t+1}^i \quad (5)$$

Namely,  $S \cup i'$  is a set of coordinates that were increased in this step. The total increase of all the coordinates in  $S$  is less than the decreased in  $i$ , but with the increase of  $i'$  it is more than the decrease of  $i$ . Such coordinates exist since the difference that the  $i$ th coordinate was moved downward there must be a set of coordinates that upward to keep that sum of coordinate 1.

Denote  $\tilde{w}$  such that:

$$\begin{aligned} \forall j \in S \quad \tilde{w}^j &= w_t^j \\ \tilde{w}^i &= w_t^i \\ \tilde{w}^{i'} &= w_{t+1}^{i'} + \sum_{j \in (S \cup i)} w_{t+1}^j - w_t^j \\ \text{o.w. } \tilde{w}^j &= w_{t+1}^j \end{aligned}$$

From Equation (4) we have that  $\tilde{w}^{i'} < w_{t+1}^{i'}$ . From Equation (5) we have that  $\tilde{w}^{i'} \geq w_t^{i'}$ .

$\tilde{w}$  is a probability since all of its coordinates are  $\geq 0$  and:

$$\begin{aligned} \sum_{j \in [d]} \tilde{w}^j &= \sum_{j \in (S \cup i)} \tilde{w}^j + \sum_{j \notin (S \cup \{i, i'\})} \tilde{w}^j + \tilde{w}^{i'} \\ &= \sum_{j \in (S \cup i)} w_t^j + \sum_{j \notin (S \cup \{i, i'\})} w_{t+1}^j + \sum_{j \in (S \cup i)} w_{t+1}^j - w_t^j \\ &= \sum_{j \in [d]} w_{t+1}^j \\ &= 1 \end{aligned}$$

Since for all  $j \in S$  we have  $\tilde{w}^j = w_t^j$ , we have:

$$\sum_{j \in S} D_r(w_t^j, \tilde{w}^j) = 0 \leq \sum_{j \in S} D_r(w_t^j, w_{t+1}^j)$$

From Taylor inequality and the definition of  $h$ :

$$\begin{aligned} D_r(\tilde{w}^i, w_t^i) &= 0 \\ D_r(w_t^i, w_{t+1}^i) &\geq \frac{h}{2}(w_{t+1}^i - w_t^i)^2 \\ D_r(w_t^i, w_{t+1}^i) &\geq D_r(\tilde{w}^i, w_t^i) + \frac{h}{2}(w_{t+1}^i - w_t^i)^2 \end{aligned}$$

Since  $w_t^{i'} \leq \tilde{w}^{i'} < w_{t+1}^{i'}$  we have  $D_r(w_t^{i'}, \tilde{w}^{i'}) < D_r(w_t^{i'}, w_{t+1}^{i'})$ .

Since  $\tilde{w}^j = w_{t+1}^j$ , we have  $\sum_{j \notin (S \cup \{i, i'\})} D_r(w_t^j, \tilde{w}^j) = \sum_{j \notin (S \cup \{i, i'\})} D_r(w_t^j, w_{t+1}^j)$ .

Summing all we have:

$$D_R(w_t, w_{t+1}) - D_R(w_t, \tilde{w}) \geq \frac{h}{2}(w_{t+1}^i - w_t^i)^2$$

From the definition of  $\tilde{w}_i$  we have  $\|\tilde{w} - w_{t+1}\| = 2(w_{t+1}^i - w_t^i)$ . Thus, from Holder:

$$\eta \langle \ell_t, w_{t+1} - \tilde{w} \rangle \geq -2\eta |w_{t+1}^i - w_t^i|$$

Since  $w_{t+1}$  is an  $\varepsilon$ -minimizer of the OMD objective:

$$\begin{aligned} \varepsilon &\geq \eta \langle \ell_t, w_{t+1} - \tilde{w} \rangle + D_R(w_t, w_{t+1}) - D_R(w_t, \tilde{w}) \\ &\geq \frac{h}{2}(w_{t+1}^i - w_t^i)^2 - 2\eta |w_{t+1}^i - w_t^i| \end{aligned}$$

From Lemma 23 we get:

$$|w_t^i - w_{t+1}^i| < \frac{4\eta}{h} + \sqrt{\frac{\varepsilon}{h}}$$

■

**Lemma 25** *Let  $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$  be an  $\varepsilon$ -approximate trajectory above the simplex with  $\eta \leq \frac{1}{16c_1}$  and  $\nu$ -barrier regularizer. Let  $t \in [T]$  and  $i \in [d]$  be such that  $\varepsilon \leq \frac{(w_t^i)^\nu}{16c_1}$ , then:*

$$\begin{aligned} w_{t-1}^i &\geq \frac{1}{2}w_t^i \\ w_{t+1}^i &\geq \frac{1}{2}w_t^i \end{aligned}$$

**Proof** We will prove for  $w_{t-1}^i$  but the same proof goes for  $t+1$ . The interesting case is obviously  $w_{t-1}^i < w_t^i$ , so continuing assuming that.

We have:

$$\min \{r''(w_t^i), r''(w_{t-1}^i)\} \geq \frac{c_1}{\max \{w_t^i, w_{t-1}^i\}^\nu} = \frac{c_1}{(w_t^i)^\nu} \geq \frac{c_1}{w_t^i}$$

From Lemma 24:

$$w_{t-1}^i \geq w_t^i - 4c_1\eta w_t^i - \sqrt{c_1\varepsilon w_t^i} \tag{6}$$

Since  $\eta \leq 1/16c_1$ :

$$4c_1\eta w_t^i \leq \frac{w_t^i}{4} \tag{7}$$

From the assumption on  $\varepsilon$  and the fact that  $r''(w_t^i) \geq w_t^i$ :

$$\sqrt{c_1\varepsilon w_t^i} \leq \sqrt{\frac{(w_t^i)^2}{16}} = \frac{w_t^i}{4} \tag{8}$$

Placing Equations (7) and (8) in Equation (6) gives the desired results. ■

**Lemma 26 (Three-points identity)** For every differentiable function  $R$ :

$$\forall x, y, z : (\nabla R(z) - \nabla R(y)) \cdot (y - x) = D_R(x, z) - D_R(x, y) - D_R(y, z)$$

**Proof**

$$\begin{aligned} D_R(x, z) - D_R(x, y) - D_R(y, z) &= R(x) - R(z) - \nabla R(z) \cdot (x - z) \\ &\quad - R(x) + R(y) + \nabla R(y) \cdot (x - y) \\ &\quad - R(y) + R(z) + \nabla R(z) \cdot (y - z) \\ &= (\nabla R(z) - \nabla R(y)) \cdot (y - x) \end{aligned}$$

■

**Lemma 27 (OMD Helper)**

$$\ell_t \cdot (w_t - w_{t+1}) - \frac{1}{\eta} D_R(w_{t+1}, w_t) \leq \frac{\eta}{2} \|\ell_t\|_*^2$$

**Proof** From the strong convexity of  $R$ :

$$\frac{1}{\eta} D_R(w_{t+1}, w_t) \geq \frac{1}{2\eta} \|w_{t+1} - w_t\|^2$$

By Holder:

$$\begin{aligned} \ell_t \cdot (w_t - w_{t+1}) &\leq \|w_t - w_{t+1}\| \|\ell_t\|_* \\ &\leq \frac{1}{2\eta} \|w_t - w_{t+1}\|^2 + \frac{\eta}{2} \|\ell_t\|_*^2 \end{aligned}$$

We used the fact that  $ab \leq \frac{1}{2}a^2 + \frac{1}{2}b^2$  for every  $a, b \geq 0$ .

■

## Appendix B. Smooth Regularizer

**Theorem** [Restatement of Theorem 2] Let  $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$  be an  $\varepsilon$ -approximate trajectory above a convex set such that  $R$  is  $\beta$ -smooth, and let  $D$  be the diameter of the domain. Assume  $\varepsilon \leq D^2/2$ , then for any  $w \in \mathcal{K}$ :

$$\text{Regret}(w) \leq O\left(\frac{1}{\eta} D_R(w, w_1) + T\eta + \frac{TD\sqrt{\beta\varepsilon}}{\eta}\right)$$

**Proof** From the strong convexity of  $R$  we have that  $\beta \geq 1$ , which means that from the assumptions  $\varepsilon \leq D^2\beta/2$ . Then, from Lemma 22, for every  $t$ :

$$\langle \eta\ell_t + \nabla R(w_{t+1}) - \nabla R(w_t), w^* - w_{t+1} \rangle \geq -D\sqrt{2\beta\varepsilon}$$

From here it is straightforward standard OMD arguments:

$$\begin{aligned} \eta\ell_t \cdot (w_{t+1} - w^*) &\leq (\nabla R(w_{t+1}) - \nabla R(w_t)) \cdot (w^* - w_{t+1}) + D\sqrt{2\beta\varepsilon} \\ &= D_R(w^*, w_t) - D_R(w^*, w_{t+1}) - D_R(w_{t+1}, w_t) + D\sqrt{2\beta\varepsilon} \end{aligned}$$

Summing for all  $t \in [T]$ :

$$\sum_{t=1}^T \ell_t \cdot (w_{t+1} - w^*) \leq \frac{1}{\eta} D_R(w^*, w_1) - \frac{1}{\eta} \sum_{t=1}^T D_R(w_{t+1}, w_t) + \frac{TD\sqrt{2\beta\varepsilon}}{\eta}$$

From Lemma 27:

$$\text{Regret}(w^*) \leq O\left(\frac{1}{\eta} D_R(w^*, w_1) + T\eta + \frac{TD\sqrt{\beta\varepsilon}}{\eta}\right)$$

■

**Theorem** [Restatement of Theorem 3] For every  $\beta, \varepsilon$ , there is an OMD  $\varepsilon$ -approximate trajectory  $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$  above a convex set with diameter  $D$  with  $R$  being  $\beta$ -smooth and constant losses ( $\ell_t = \ell$  for some  $\ell$  for all  $t \in [T]$ ) that achieves a regret of

$$\Omega\left(\min\left(\frac{TD\sqrt{\beta\varepsilon}}{\eta}, DT\right)\right)$$

**Proof** Consider the domain  $[0, D]$ ,  $w_1 = \frac{D}{2}$ . The regularizer is  $R(w) = \frac{\beta}{2}w^2$ . The loss is  $\ell = \min\left\{\frac{\sqrt{2\beta\varepsilon}}{\eta}, 1\right\}$ .

We will now show by induction that  $w_t = w_1$  for all  $t$  is a valid  $\varepsilon$ -approximate trajectory. This trajectory suffers a loss of  $\Theta\left(\min\left(\frac{TD\sqrt{\beta\varepsilon}}{\eta}, DT\right)\right)$ , which means a same regret comparing to  $w^* = 0$ .

Assume true for  $t - 1$ , we will prove for  $t$ .

We start by finding the optimal  $w_t^*$  (the optimal solution for  $\phi_t$ ) by differentiating and comparing to 0:

$$\begin{aligned} \eta\ell + \beta(w_t^* - w_{t-1}) &= 0 \\ \iff w_t^* &= w_{t-1} - \frac{\eta}{\beta}\ell \end{aligned}$$

Placing it in the objective function:

$$\begin{aligned} \eta\ell w_t^* + \frac{\beta}{2}(w_t^* - w_{t-1})^2 &= \eta\ell\left(w_{t-1} - \frac{\eta}{\beta}\ell\right) + \frac{\beta}{2}\left(w_{t-1} - \left(w_{t-1} - \frac{\eta}{\beta}\ell\right)\right)^2 \\ &= \eta\ell w_{t-1} - \frac{\eta^2\ell^2}{\beta} + \frac{\eta^2\ell^2}{2\beta} \\ &= \eta\ell w_{t-1} - \frac{\eta^2\ell^2}{2\beta} \end{aligned}$$

Which means that the difference in the objective function between  $w_t^*$  and  $w_{t-1}$  is  $\frac{\eta^2\ell^2}{2\beta}$ .

From the definition of  $\ell$ :

$$\begin{aligned} \ell &\leq \frac{\sqrt{2\beta\varepsilon}}{\eta} \\ \iff \frac{\eta^2\ell^2}{2\beta} &\leq \varepsilon \end{aligned}$$

Which means that  $w_{t-1}$  is an  $\varepsilon$ -minimizer. ■

## Appendix C. Balance

All the lemmas in this section assumes  $\nu$ -barrier regularizer.

### C.1. General

**Definition 28** Assume  $\mathcal{K}$  is a polytope defined in standard form  $\{w \in \mathbb{R}^d : Aw = b \wedge (w_i \geq 0, \forall i \in [d])\}$ . For every  $v \in \ker(A)$ , denote the balance of an OMD trajectory w.r.t  $v$ :

$$B_\gamma^v(t_1, t_2) = \langle \nabla R(w_{t_1}) - \nabla R(w_{t_2}), v \rangle$$

Additionally, if for every  $v \in \ker(A)$  such that  $\|v\| \leq 1$  and  $t_1, t_2$  we have  $B_\gamma^v(t_1, t_2) \leq k$ , we say the trajectory is  $k$  balanced.

#### Lemma 29

$$B_\gamma^i(t_1, t_2) + B_\gamma^i(t_2, t_3) = B_\gamma^i(t_1, t_3)$$

#### Proof

$$\langle \nabla R(w_{t_1}) - \nabla R(w_{t_2}), v \rangle + \langle \nabla R(w_{t_2}) - \nabla R(w_{t_3}), v \rangle = \langle \nabla R(w_{t_1}) - \nabla R(w_{t_3}), v \rangle$$

■

**Lemma 30** Assume  $\mathcal{K}$  is a polytope. For some differentiable function  $f : \mathcal{K} \rightarrow \mathbb{R}$ , let  $w^*$  be the minimizer of  $f$  such that for all  $i \in [d]$ ,  $(w^*)^i > 0$ . For every  $w \in \mathcal{K}$  we have:

$$\langle \nabla f(w^*), w - w^* \rangle = 0$$

**Proof** Denote  $v = w - w^*$ . Since  $\min_i \hat{w}^i > 0$ , and  $v \in \ker(A)$  where  $A$  is the matrix of the polytope  $\mathcal{K}$ , there is an  $\alpha$  such that both  $w^* + \alpha v \in \mathcal{K}$  and  $w - \alpha v \in \Delta_d$ .

Since  $w^*$  is a minimizer, from first order optimality conditions:

$$\begin{aligned} \langle \nabla f(\hat{w}), w^* + \alpha v - w^* \rangle &\geq 0 \\ \langle \nabla f(\hat{w}), w^* - \alpha v - w^* \rangle &\geq 0 \end{aligned}$$

Which means that:

$$\begin{aligned} \langle \nabla f(\hat{w}), v \rangle &\geq 0 \\ \langle \nabla f(\hat{w}), -v \rangle &\geq 0 \end{aligned}$$

Which is our desired results. ■

**Lemma** [Restatement of Lemma 10] Let  $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$  be an exact OMD trajectory. For every  $v \in \ker(A)$  and times  $t_1, t_2$ :

$$B^v(t_1, t_2) = \eta \langle \ell_{t_1:t_2}, v \rangle$$

**Proof** Fix some  $t' \in [t_1, t_2]$ . There is some small  $\alpha$  such that both  $w_{t'} + \alpha v$  and  $w_{t'} - \alpha v$  is in the polytope. From Lemma 30 ( $w_{t'}^i > 0$  since the regularizer is undefined in 0):

$$\langle \ell_{t'-1} + \nabla R(w_{t'}) - \nabla R(w_{t'-1}), v \rangle = 0$$

Summing for all  $t' \in [t_1, t_2]$  gives the desired results.  $\blacksquare$

**Lemma** [Restatement of Lemma 11] Let  $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$  and  $\hat{\gamma} = (\{\hat{w}_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$  be an exact OMD trajectory and  $\varepsilon$ -approximate OMD trajectory.

Let  $0 \leq t_1 \leq t_2 \leq T$ ,  $v \in \ker(A)$  such that  $\|v\| = 1$  and let  $\psi > 0$  be such that for every  $t_1 \leq t \leq t_2$ , for all  $i \in [d]$  such that  $v^i \neq 0$ ,  $\hat{w}_t^i \geq \psi$ . We also assume that  $\varepsilon \leq c_2\psi/2$ . we have:

$$B_{\hat{\gamma}}^v(t_1, t_2) \leq B_{\gamma}^v(t_1, t_2) + (t_2 - t_1) \sqrt{\frac{c_2\varepsilon}{\psi^\nu}}$$

**Proof** We will prove it using induction for  $t \in [t_1, t_2]$ . The base  $t = t_1$  is trivial.

Assume true for  $t - 1$ , namely:

$$B_{\hat{\gamma}}^v(t_1, t - 1) \leq B_{\gamma}^v(t_1, t - 1) + (t - 1 - t_1) \sqrt{\frac{c_2\varepsilon}{\psi^\nu}}$$

From the assumptions of the lemma we have  $\hat{w}_t + \psi v \in \mathcal{K}$ . Additionally, the effective smoothness is  $\frac{c_2}{\psi^\nu}$ . From Lemma 22, since  $\hat{w}_t$  is an  $\varepsilon$ -minimizer of  $\phi_t$ :

$$\langle \eta \ell_t + \nabla R(\hat{w}_t) - \nabla R(\hat{w}_{t-1}), \psi v \rangle \geq -\max \left\{ \psi \sqrt{\frac{2c_2\varepsilon}{\psi^\nu}}, 2\varepsilon \right\}$$

Since  $\varepsilon \leq c_2\psi/2$  and from the definition of barrier regularizer:

$$\begin{aligned} \psi \sqrt{\frac{2c_2\varepsilon}{\psi^\nu}} &\geq \psi \sqrt{\frac{2c_2\varepsilon}{\psi}} \\ &= \sqrt{2c_2\varepsilon\psi} \\ &\geq 2\varepsilon \end{aligned}$$

Which means:

$$\langle \eta \ell_t + \nabla R(\hat{w}_t) - \nabla R(\hat{w}_{t-1}), \psi v \rangle \geq -\psi \sqrt{\frac{2c_2\varepsilon}{\psi^\nu}}$$

Dividing by  $\psi > 0$ :

$$\begin{aligned} -\sqrt{\frac{2c_2\varepsilon}{\psi^\nu}} &\leq \langle \eta \ell_t, v \rangle - B_{\hat{\gamma}}^v(t - 1, t) \\ &= B_{\gamma}^v(t - 1, t) - B_{\hat{\gamma}}^v(t - 1, t) \end{aligned} \quad (\text{Lemma 10})$$

Adding the induction assumption:

$$\begin{aligned} B_{\hat{\gamma}}^v(t_1, t - 1) + B_{\hat{\gamma}}^v(t - 1, t) &\leq B_{\gamma}^v(t_1, t - 1) + B_{\gamma}^v(t - 1, t) + \sqrt{\frac{2c_2\varepsilon}{\psi^\nu}} + (t - 1 - t_1) \sqrt{\frac{c_2\varepsilon}{\psi^\nu}} \\ \implies B_{\hat{\gamma}}^v(t_1, t) &\leq B_{\gamma}^v(t_1, t) + (t - t_1) \sqrt{\frac{c_2\varepsilon}{\psi^\nu}} \end{aligned}$$

The last is from Lemma 29.  $\blacksquare$

### C.2. Simplex subset

The lemmas in this section assumes that the polytope is a subset of the simplex. That is, for every  $w$  such that  $Aw = b$ ,  $\|w\|_1 = 1$ . Additionally, the primal norm is assumed to be  $L_1$  norm.

**Lemma 31** *Let  $v$  be a vector in the kernel of  $A$ . The sum of the elements of  $v$  is 0.*

**Proof** Denote  $w = w_1 + \frac{1}{d\|v\|_\infty}v$ . It is in the polytope - all the elements of  $\frac{1}{d\|v\|_\infty}v$  are smaller than  $1/d$  and thus the all the elements of  $w$  greater than 0, and since  $v$  is in the kernel of  $A$  we have:

$$Aw = Aw_1 + A\frac{1}{d\|v\|_\infty}v = Aw_1 = b$$

Thus, we have  $\|w\| = 1$ . Since also  $\|w_1\| = 1$ :

$$\frac{1}{d\|v\|_\infty} \sum_{i=1}^d v^i = \sum_{i=1}^d w^i - \sum_{i=1}^d w_1^i = 1 - 1 = 0$$

■

**Lemma** [Restatement of Lemma 12] *Let  $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$  be a  $k$ -balanced OMD trajectory with  $w_1 = (1/d, 1/d, \dots, 1/d)$  and coordinate-separated regularizer. For every  $t \in [T], i \in [d]$ :*

$$-r'(w_t^i) \leq \max \{4kd - r'(1/d), -r'(1/2d)\}$$

**Proof** Since  $w_1$  is uniform and Lemma 31, for every  $v \in \ker(A)$ :

$$\langle \nabla R(w_1), v \rangle = r'(1/d) \sum_{i=1}^d v^i = 0$$

Let  $v = w_1 - w_t \in \ker(A)$ . Notice that since  $\|w_t\|_1 = \|w_1\|_1 = 1$ , from triangle inequality  $\|v\| \leq 2$ . From Lemma 10:

$$\begin{aligned} 2k &\geq \langle -\nabla R(w_t), v \rangle \\ &= -\sum_{i=1}^d r'(w_t^i)v^i \\ &= -\sum_{i:v^i>0} r'(w_t^i)v^i - \sum_{i:v^i\leq 0} r'(w_t^i)v^i \end{aligned}$$

Denote:

$$\sum_{i:v^i>0} v^i = \alpha$$

From Lemma 31:

$$-\sum_{i:v^i\leq 0} v^i = \alpha$$

If  $v^i \leq 0$  it means that  $w_t^i \geq w_1^i = 1/d$ , thus:

$$- \sum_{i:v^i \leq 0}^d r'(w_t^i) v^i \geq -r'(1/d) \sum_{i:v^i \leq 0}^d v^i \geq r'(1/d) \alpha$$

We used the fact that from the convexity of  $r$ ,  $r'$  is monotonically increasing (as  $r'' \geq 0$ ).

Denote  $\bar{i} = \arg \min_{i \in [d]} w_t^i$ , we have:

$$\begin{aligned} 2k - \alpha r'(1/d) &\geq -r'(w_{\bar{i}}^{\bar{i}}) v^{\bar{i}} - \sum_{i:v^i > 0, i \neq \bar{i}}^d r'(w_t^i) v^i \\ &\geq -r'(w_{\bar{i}}^{\bar{i}}) v^{\bar{i}} - r'(1/d) \sum_{i:v^i > 0, i \neq \bar{i}}^d v^i \\ &= -r'(w_{\bar{i}}^{\bar{i}}) v^{\bar{i}} - r'(1/d) (\alpha - v^{\bar{i}}) \end{aligned}$$

Subtracting from both sides:

$$2k - r'(1/d) v^{\bar{i}} \geq -r'(w_{\bar{i}}^{\bar{i}}) v^{\bar{i}}$$

If  $v^{\bar{i}} \leq 1/2d$  we have  $w_{\bar{i}}^{\bar{i}} \geq 1/2d$ . Since  $r'$  is monotonically increasing, this means that for all  $i \in [d]$   $r'(w_t^i) \geq r'(1/2d)$  which concludes the proof. Else, dividing by  $v^{\bar{i}} \geq 1/2d$ :

$$-r'(w_t^{\bar{i}}) \leq 4kd - r'(1/d)$$

■

### C.3. Simplex

**Definition 32** We denote the balance of an OMD trajectory above the simplex  $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$  w.r.t to a coordinate  $i$  and  $0 \leq t_1 \leq t_2 \leq T$  to be:

$$B_\gamma^i(t_1, t_2) = r'(w_{t_1}^{i*}) - r'(w_{t_2}^{i*}) + r'(w_{t_2}^i) - r'(w_{t_1}^i)$$

We say that an OMD trajectory is  $k$ -balanced if, for every  $0 \leq t_1 \leq t_2 \leq T$  and coordinate  $i$ :

$$B_\gamma^i(t_1, t_2) \leq k$$

One can notice that it is a private case for the general polytope definition.

**Lemma [Restatement of Lemma 13]** Let  $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$  be an approximate OMD trajectory. Fix  $t_1, t_2 \in [T]$  and  $i \in [d]$  such that  $B_\gamma^i(t_1, t_2) \leq k$ . Then:

1. If  $w_{t_2}^i \geq w_{t_1}^i$  then  $e^{k/c_1} w_{t_2}^{i*} \geq w_{t_1}^{i*}$
2. If  $w_{t_2}^{i*} \leq w_{t_1}^{i*}$  then  $w_{t_2}^i \leq e^{k/c_1} w_{t_1}^i$

**Proof** We will prove the first statement and the second follows in just the same way.

Assume by contradiction that  $e^k w_{t_2}^{i*} < w_{t_1}^{i*}$ . Since  $w_{t_2}^i \geq w_{t_2}^i$  we have  $r'(w_{t_2}^i) \geq r'(w_{t_1}^i)$ , which means:

$$\begin{aligned}
 k &\geq B_\gamma^i(t_1, t_2) \\
 &= r'(w_{t_1}^{i*}) - r'(w_{t_2}^{i*}) + r'(w_{t_2}^i) - r'(w_{t_1}^i) \\
 &\geq r'(w_{t_1}^{i*}) - r'(w_{t_2}^{i*}) \\
 &= \int_{w_{t_2}^{i*}}^{w_{t_1}^{i*}} r''(w) dw \\
 &> \int_{w_{t_2}^{i*}}^{e^{k/c_1} w_{t_2}^{i*}} r''(w) dw && (r''(w) > 0) \\
 &\geq \int_{w_{t_2}^{i*}}^{e^{k/c_1} w_{t_2}^{i*}} \frac{c_1}{w} dw \\
 &= c_1 \left( \log \left( e^{k/c_1} w_{t_2}^{i*} \right) - \log \left( w_{t_2}^{i*} \right) \right) \\
 &= k
 \end{aligned}$$

Which is a contradiction  $k > k$ . ■

**Lemma 33** Let  $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$  and  $\hat{\gamma} = (\{\hat{w}_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$  be an optimal OMD trajectory and  $\varepsilon$ -approximate OMD trajectory.

Let  $0 \leq t_1 \leq t_2 \leq T$ ,  $i \in [d]$  and  $\psi > 0$  be such that for every  $t_1 \leq t \leq t_2$ ,  $\hat{w}_t^i \geq \psi$  and  $\hat{w}_t^{i*} \geq \psi$ . We also assume that  $\varepsilon \leq \psi/2$ . we have:

$$B_{\hat{\gamma}}^i(t_1, t_2) \leq B_\gamma^i(t_1, t_2) + (t_2 - t_1) \sqrt{r''(\psi)} \varepsilon$$

**Proof** It is direct consequence of Lemma 11 for the case of  $v_i = e_{i^*} - e_i$  ( $e_j$  is the  $j$ th element of the standard basis). ■

## Appendix D. Lower bounds for negative entropy

**Lemma 34** Let  $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$  be an  $\varepsilon$ -approximate trajectory with  $d = 2$  and  $\nu$ -barrier regularizer. If for some coordinate  $i$  there is  $\tau \in [T]$  such that  $\frac{4\eta}{c_1} (w_\tau^i)^\nu \leq \varepsilon$ , then for any possible losses for  $t \geq \tau$ , having  $w_t^i = w_\tau^i$  makes a valid error trajectory.

**Proof** We'll prove by induction. Assume true for  $w_t^i$ , we'll prove for  $w_{t+1}^i$ .

Denote  $\tilde{w}_{t+1}$  such that:

$$\tilde{w}_{t+1} = \arg \min_{w \in \Delta_2} \phi_t(w)$$

From Lemma 24 with  $\varepsilon = 0$  we get:

$$|w_t^i - \tilde{w}_{t+1}^i| \leq \frac{4\eta}{r''(w_t^i)} \leq \frac{4\eta}{c_1} (w_t^i)^\nu \leq \varepsilon$$

Thus:

$$\langle \ell_t, \tilde{w}_{t+1} - w_t \rangle \leq \varepsilon$$

Since by definition  $D_R(w_t, w_t) \leq D_R(\tilde{w}_{t+1}, w_t)$ , we get:

$$\phi(w_t) \leq \phi(\tilde{w}_{t+1}) + \varepsilon$$

Which means that  $w_t$  is an  $\varepsilon$ -minimizer, as needed. ■

**Lemma** [Restatement of Lemma 14] *Assume for some  $\alpha \geq T/2$ ,  $\frac{1}{\eta} \log\left(\frac{4\eta}{\varepsilon}\right) \leq \alpha$  with negative entropy regularizer, there is an instance above the simplex with  $\alpha$ -balanced losses that the regret achieved is  $\Omega(T - 2\alpha)$ .*

**Proof** We construct an instance with  $d = 2$  and  $(1, 0)$  losses for the first  $\tau = \frac{1}{\eta} \log\left(\frac{4\eta}{\varepsilon}\right)$  and then  $(0, 1)$ . Since  $\tau < T/2$  we have that the optimal coordinate is 1. We have:

$$w_\tau^1 \leq e^{-\eta\tau} = \frac{\varepsilon}{4\eta}$$

From Lemma 34, it is a valid error trajectory if for every  $t \geq \tau$ ,  $w_t^1 \leq \frac{\varepsilon}{4\eta} \leq \frac{1}{2}$ . Thus, the regret for those steps is  $\Omega(T - \tau)$ . Adding the first  $\tau$  steps we get a regret bound of  $\Omega(T - 2\tau) \geq \Omega(T - 2\alpha)$ . ■

We add another lower bound that shows an instance in which the optimal point in the optimal trajectory doesn't get close to 0 but still there is a linear regret.

**Theorem 35** *Assume  $\varepsilon \geq \frac{4\eta^2}{c_1 d^\nu}$  and  $\nu$ -barrier regularizer. There is a set of constant losses for which there is an  $\varepsilon$ -approximate OMD trajectory that achieves a regret of  $\Omega(T)$ .*

**Proof** The losses are  $\ell_t^d = 0$  and  $\ell_t^i = 1$  for  $i \in [d-1]$  for all  $t$ . We will show that having  $w_t = w_1$  for all  $t \in [T]$  is a valid  $\varepsilon$ -approximate OMD trajectory. Since  $w_1$  is the uniform distribution, the total loss is  $T - \frac{T}{d}$ . The optimal point is  $w^* = (0, \dots, 0, 1)$ , namely having 1 only in the  $d$ th coordinate, which gives a total loss of 0. Since  $T - \frac{T}{d} = \Omega(T)$  even for  $d = 2$ , this seals the proof.

We will now prove by induction that if  $w_t = w_1$ ,  $w_1$  is an  $\varepsilon$ -approximate minimizer for  $\phi_t$ . Denote:

$$\tilde{w}_{t+1} = \arg \min_{w \in \Delta_d} \phi_t(w)$$

From Lemma 24 with  $\varepsilon = 0$  we get:

$$\left| w_t^d - \tilde{w}_{t+1}^d \right| \leq \frac{4\eta}{r''(w_t^d)} \leq \frac{4\eta}{c_1 d^\nu} \leq \frac{\varepsilon}{\eta}$$

Since  $w_t^d = 1/d$ :

$$\tilde{w}_{t+1}^d \leq \frac{1}{d} + \frac{\varepsilon}{\eta}$$

Summing for all coordinates:

$$\frac{d-1}{d} - \frac{\varepsilon}{\eta} \leq \sum_{i=1}^{d-1} \tilde{w}_{t+1}^i = \langle \ell_t, \tilde{w}_{t+1} \rangle$$

Since  $\langle \ell_t, w_1 \rangle = \frac{d-1}{d}$  we have: We have:

$$\langle \eta \ell_t, \tilde{w}_{t+1} \rangle \geq \langle \eta \ell_t, w_1 \rangle - \varepsilon$$

Since by definition  $D_R(w_t, w_t) \leq D_R(\tilde{w}_{t+1}, w_t)$ , we get:

$$\phi_t(w_t) \leq \phi_t(\tilde{w}_{t+1}) + \varepsilon$$

which means that  $w_t = w_1$  is an  $\varepsilon$ -minimizer, as needed. ■

**Theorem 36** *Consider the following instance with negative entropy regularizer for some  $k \leq \frac{T\eta}{20}$ . For the first  $\frac{3k}{2\eta}$  steps, the loss is  $(0, 1)$ . Then, for the next  $\frac{k}{\eta}$  steps, the loss is  $(1, 0)$ . Then, for the rest ( $\geq \frac{3T}{4}$ ) of the steps, the loss is  $(0, 1)$ . There is an error OMD trajectory with  $\varepsilon = 4\eta e^{-k/2}$  that has a regret  $\Omega(T)$ .*

**Proof** After  $\tau = \frac{k}{2\eta}$  steps we have  $w_\tau^2 \leq \frac{\varepsilon}{4\eta}$ . From Lemma 34, it is a valid error trajectory if for every  $3\tau \geq t \geq \tau$ ,  $w_t = w_\tau$ .

On the steps between  $3\tau$  and  $4\tau$  we have a loss of  $(1, 0)$ . Since  $w_{3\tau} = w_\tau$ , we have that  $w_{4\tau} = (\frac{1}{2}, \frac{1}{2})$ . That is because this is what would have happen if those last  $\tau$  steps where after  $\tau$  (as the sum of losses for both coordinates is  $\tau$ ), and since we didn't move at all in  $\tau \leq t \leq 3\tau$  it is the same.

On the steps between  $4\tau$  and  $5\tau$  we assume no errors. Coordinate 1 does the same trajectory that coordinate 2 did in the beginning, so we have  $w_\tau^1 \leq \frac{\varepsilon}{4\eta}$ .

From Lemma 34, it is a valid error trajectory if for every  $T \geq t \geq 5\tau$ ,  $w_t = w_{5\tau} \leq \frac{\varepsilon}{4\eta}$ . Since this are  $3T/4$  steps, we have a regret of  $\Theta(T)$ .

For summary:

$$\begin{aligned} w_1 &= \left( \frac{1}{2}, \frac{1}{2} \right) \\ w_\tau &\approx \left( 1 - \frac{\varepsilon}{4\eta}, \frac{\varepsilon}{4\eta} \right) \\ w_{3\tau} &\approx \left( 1 - \frac{\varepsilon}{4\eta}, \frac{\varepsilon}{4\eta} \right) \\ w_{4\tau} &= \left( \frac{1}{2}, \frac{1}{2} \right) \\ w_{5\tau} &\approx \left( \frac{\varepsilon}{4\eta}, 1 - \frac{\varepsilon}{4\eta} \right) \\ w_T &\approx \left( \frac{\varepsilon}{4\eta}, 1 - \frac{\varepsilon}{4\eta} \right) \end{aligned}$$

■

## Appendix E. Proof of Lemma 15

**Lemma 37** *Let  $w_1, w_2 \in (0, 1]$  such that  $w_1 \leq w_2$ , then  $D_r(0, w_1) \leq D_r(0, w_2)$*

**Proof** Denote  $f(x) = D_r(0, x)$ . We have:

$$\begin{aligned} f(x) &= r(0) - r(x) + r'(x)x \\ f'(x) &= -r'(x) + r''(x)x + r'(x) = r''(x)x \geq 0 \end{aligned}$$

Which means that  $f$  is increasing in  $(0, 1]$ . ■

**Lemma 38** *Let  $\hat{\gamma} = (\{\hat{w}_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$  be an  $\varepsilon$ -approximate OMD trajectory with  $\eta \leq \frac{1}{4}$  and coordinate separable  $R$  with  $r''(w) = 1/w^\nu$ . For every  $i \in [d]$  and  $t \in [T]$  such that  $\varepsilon \leq \frac{\eta^2}{r''(\hat{w}_t)}$  we have:*

$$(\nabla \phi_t(\hat{w}_t^i)) \leq O(2^\nu \eta)$$

**Proof** Since  $\ell_t^i \leq 1$  we have  $\eta \ell_t^i \leq \eta$ , which means that we only need to prove:

$$r'(\hat{w}_t^i) - r'(\hat{w}_{t-1}^i) \leq O(2^\nu \eta)$$

Since  $r'$  is monotonically increasing it is trivial if  $\hat{w}_t^i \leq \hat{w}_{t-1}^i$ , continuing assuming  $\hat{w}_t^i > \hat{w}_{t-1}^i$ . We have  $\varepsilon \leq \eta^2 / r''(\hat{w}_t^i) \leq 1 / (16r''(\hat{w}_t^i))$ , so from Lemma 25:

$$\begin{aligned} \hat{w}_{t-1}^i &\geq \frac{1}{2} \hat{w}_t^i \\ \Leftrightarrow \frac{2^\nu}{(\hat{w}_t^i)^\nu} &\geq \frac{1}{(\hat{w}_{t-1}^i)^\nu} \\ \Leftrightarrow 2^\nu r''(\hat{w}_t^i) &\geq r''(\hat{w}_{t-1}^i) \end{aligned}$$

From Lemma 24:

$$\hat{w}_t^i - \hat{w}_{t-1}^i \leq \frac{4\eta}{r''(\hat{w}_t^i)} + \sqrt{\frac{\varepsilon}{r''(\hat{w}_t^i)}}$$

Which implies:

$$\begin{aligned} \varepsilon &\leq \frac{\eta^2}{\hat{w}_t^i} \leq \frac{\eta^2}{r''(\hat{w}_t^i)} \\ \Rightarrow \sqrt{\frac{\varepsilon}{r''(\hat{w}_t^i)}} &\leq \frac{\eta}{r''(\hat{w}_t^i)} \\ \Rightarrow \hat{w}_t^i - \hat{w}_{t-1}^i &\leq \frac{5\eta}{r''(\hat{w}_t^i)} \end{aligned}$$

From mean value theorem and monotonicity of  $r''$ :

$$\begin{aligned} r'(\hat{w}_t^i) - r'(\hat{w}_{t-1}^i) &\leq |\hat{w}_t^i - \hat{w}_{t-1}^i| \max_{w \in \{\hat{w}_t^i, \hat{w}_{t-1}^i\}} r''(w) \\ &\leq (\hat{w}_t^i - \hat{w}_{t-1}^i) r''(\hat{w}_{t-1}^i) \\ &\leq (\hat{w}_t^i - \hat{w}_{t-1}^i) 2^\nu r''(\hat{w}_t^i) \\ &\leq \frac{5\eta}{r''(\hat{w}_t^i)} 2^\nu r''(\hat{w}_t^i) \\ &\leq 5 \cdot 2^\nu \eta \\ &= O(2^\nu \eta) \end{aligned}$$

■

**Lemma 39** *Let  $\mathcal{K} = \Delta_d$  and  $\hat{\gamma} = (\{\hat{w}_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$  with  $\eta \leq \frac{1}{16}$ , coordinate separable  $R$  with  $r''(w) = 1/w^\nu$  and uniform initialization  $\hat{w}_1 = (1/d \dots 1/d)$  be an  $\varepsilon$ -approximate OMD trajectory such that there is  $\xi > 0$  such that for every  $t \in [T]$ ,  $\hat{w}_t^{i^*} \geq \xi$ . If  $\varepsilon \leq \frac{\eta^4}{r''(\min\{\frac{\eta}{d}, \xi\})}$ , its regret w.r.t any  $w \in \mathcal{K}$  is bounded by:*

$$\text{Regret}(w) \leq \frac{1}{\eta} D_R(w, \hat{w}_1) + O(2^\nu T \eta)$$

**Proof** Let  $\xi' = \min\{\frac{\eta}{d}, \xi\}$ , and let  $S_t = \{i \neq i^* : \hat{w}_t^i \geq \xi'\}$  for  $t \geq 2$ .

In every step  $t$  we set  $\tilde{w}_t$  to be:

$$\begin{aligned} \tilde{w}_t^i &= \hat{w}_t^i & i \notin S_t, i \neq i^* \\ \tilde{w}_t^i &= 0 & i \in S_t \\ \tilde{w}_t^{i^*} &= 1 - \sum_{i \notin S_t} \hat{w}_t^i \end{aligned}$$

Since the changes between  $\hat{w}_t$  and  $\tilde{w}_t$  are only in coordinates with value greater than  $\xi'$ , the effective smoothness is upper bounded by  $r''(\xi')$  (since  $r''(w) = 1/w^\nu$  for all  $w \in (0, 1]$ ). To use Lemma 22, we need to show that  $\varepsilon \leq D^2 r''(\xi')/2$  where  $D$  is the diameter w.r.t to  $L_1$  norm. Indeed, we have that  $r''(w^i) \geq 1$  for all  $w \in \Delta_d$  and  $i \in [d]$  and  $D = 2$ . By our assumptions it holds that  $\varepsilon \leq 1$ , hence  $\varepsilon \leq D^2 r''(\xi')/2$ . Thus, from Lemma 22 on  $\phi_t$ :

$$\langle \eta \ell_{t-1} + \nabla R(\hat{w}_t) - \nabla R(\hat{w}_{t-1}), \tilde{w}_t - \hat{w}_t \rangle \geq -2\sqrt{2r''(\xi')\varepsilon} \geq -2\eta^2.$$

Which means:

$$\begin{aligned} & \left( \eta \ell_{t-1}^{i^*} + \nabla R(\hat{w}_t)^{i^*} - \nabla R(\hat{w}_{t-1})^{i^*} \right) (\tilde{w}_t^{i^*} - \hat{w}_t^{i^*}) \\ & + \sum_{i \in S_t} \left( \eta \ell_{t-1}^i + \nabla R(\hat{w}_t)^i - \nabla R(\hat{w}_{t-1})^i \right) (0 - \hat{w}_t^i) \\ & \geq -2\eta^2. \end{aligned} \tag{9}$$

Notice that since  $\xi' \leq \frac{\eta}{d}$  we have that  $\sum_{i \notin S_t} \hat{w}_t^i \leq \eta$  which means  $1 - \tilde{w}_t^{i^*} \leq \eta$ . Additionally, from Lemma 38 we have that  $\nabla \phi(\hat{w}_t^{i^*}) \leq O(2^\nu \eta)$ . We have:

$$\left( \eta \ell_{t-1}^{i^*} + \nabla R(\hat{w}_t)^{i^*} - \nabla R(\hat{w}_{t-1})^{i^*} \right) (\tilde{w}_t^{i^*} - 1) = -O(2^\nu \eta^2)$$

Thus, Equation (9) can be written as:

$$\begin{aligned}
 & \left( \eta \ell_{t-1}^{i^*} + \nabla R(\hat{w}_t)^{i^*} - \nabla R(\hat{w}_{t-1})^{i^*} \right) (1 - \hat{w}_t^{i^*}) + \\
 & \quad \left( \eta \ell_{t-1}^{i^*} + \nabla R(\hat{w}_t)^{i^*} - \nabla R(\hat{w}_{t-1})^{i^*} \right) (\tilde{w}_t^{i^*} - 1) + \\
 & \quad \sum_{i \in S} \left( \eta \ell_{t-1}^i + \nabla R(\hat{w}_t)^i - \nabla R(\hat{w}_{t-1})^i \right) (0 - \hat{w}_t^i) \\
 & \geq -\eta^2 \\
 & \Rightarrow \\
 & \left( \eta \ell_{t-1}^{i^*} + \nabla R(\hat{w}_t)^{i^*} - \nabla R(\hat{w}_{t-1})^{i^*} \right) (1 - \hat{w}_t^{i^*}) + \sum_{i \in S} \left( \eta \ell_{t-1}^i + \nabla R(\hat{w}_t)^i - \nabla R(\hat{w}_{t-1})^i \right) (0 - \hat{w}_t^i) \geq -O(2^\nu \eta^2) \\
 & \eta \ell_{t-1}^{i^*} (\hat{w}_t^{i^*} - 1) + \eta \sum_{i \in S} \ell_{t-1}^i (\hat{w}_t^i - 0) \leq \\
 & \quad \left( \nabla R(\hat{w}_t)^{i^*} - \nabla R(\hat{w}_{t-1})^{i^*} \right) (1 - \hat{w}_t^{i^*}) + \sum_{i \in S} \left( \nabla R(\hat{w}_t)^i - \nabla R(\hat{w}_{t-1})^i \right) (0 - \hat{w}_t^i) + O(2^\nu \eta^2)
 \end{aligned}$$

From Lemma 26:

$$\begin{aligned}
 \eta \ell_{t-1}^{i^*} (\hat{w}_t^{i^*} - 1) + \eta \sum_{i \in S_t} \ell_{t-1}^i (\hat{w}_t^i - 0) & \leq D_r(1, \hat{w}_{t-1}^{i^*}) - D_r(1, \hat{w}_t^{i^*}) - D_r(\hat{w}_t^{i^*}, \hat{w}_{t-1}^{i^*}) \\
 & \quad + \sum_{i \in S_t} D_r(0, \hat{w}_{t-1}^i) - D_r(0, \hat{w}_t^i) - D_r(\hat{w}_t^i, \hat{w}_{t-1}^i) \\
 & \quad + O(2^\nu \eta^2)
 \end{aligned}$$

Fix some coordinate  $i \neq i^*$ , and let  $(s_1, t_1), (s_2, t_2), \dots, (s_n, t_n)$  be all enter and exit times for  $i$  to  $S_t$ . Namely, for every  $j \in [n]$  and  $s_j \leq t \leq t_j$ ,  $i \in S_t$ , and  $i \notin S_t$  otherwise. Hence,

$$\sum_{t: i \in S_t} D_r(0, \hat{w}_{t-1}^i) - D_r(0, \hat{w}_t^i) = \sum_{j=1}^n D_r(0, \hat{w}_{s_j-1}^i) - D_r(0, \hat{w}_{t_j}^i),$$

where the equality follows by telescoping the terms. Since  $s_j$  is enter time for coordinate  $i$ , we have that  $i \notin S_{s_j-1}$ , which means that  $\hat{w}_{s_j-1}^i < \xi'$ . On the other hand,  $i \in S_{t_j}$ , which means that  $\hat{w}_{t_j}^i \geq \xi' > \hat{w}_{s_j-1}^i$ . Thus, by Lemma 37 we get that  $D_r(0, \hat{w}_{s_j-1}^i) \leq D_r(0, \hat{w}_{t_j}^i)$  which we apply on the RHS of the previous display to obtain:

$$\sum_{t: i \in S_t} D_r(0, \hat{w}_{t-1}^i) - D_r(0, \hat{w}_t^i) \leq D_r(0, \hat{w}_{s_1-1}^i)$$

We now argue that for every  $i \in [d]$ ,  $i \in S_2$ , which means that  $s_1 = 2$ . Assume by contradiction that  $\hat{w}_2^i < \hat{w}_1^i$  (and thus  $r''(\hat{w}_2^i) > r''(\hat{w}_1^i)$ , from Lemma 24:

$$\begin{aligned}
 \hat{w}_2^i - \hat{w}_1^i & \leq \frac{\eta}{r''(1/d)} + \sqrt{\frac{\varepsilon}{r''(1/d)}} \\
 & \leq \frac{\eta}{d} + \sqrt{\frac{1}{16r''(\eta/d)r''(1/d)}} \\
 & \leq \frac{1}{4d} + \frac{1}{4d} \\
 & \Rightarrow \hat{w}_2^i \geq 1/2d
 \end{aligned}$$

Thus:

$$\sum_{t:i \in S_t} D_r(0, \hat{w}_{t-1}^i) - D_r(0, \hat{w}_t^i) \leq D_r(0, \hat{w}_1^i)$$

Thus:

$$\begin{aligned} & \sum_{t=2}^T \eta \ell_{t-1}^{i^*} (\hat{w}_t^{i^*} - 1) + \eta \sum_{i \in [d] \setminus i^*} \sum_{t:i \in S_t} \ell_{t-1}^i (\hat{w}_t^i - 0) \\ & \leq D_r(1, \hat{w}_1^{i^*}) - \sum_{t=2}^T D_r(\hat{w}_t^{i^*}, \hat{w}_{t-1}^{i^*}) + \sum_{i \in [d] \setminus i^*} D_r(0, \hat{w}_1^i) - \sum_{t=2}^T D_r(\hat{w}_t^i, \hat{w}_{t-1}^i) + O(2^\nu T \eta^2) \\ & = D_R(w^*, \hat{w}_1) - \sum_{t=2}^T D_R(\hat{w}_t, \hat{w}_{t-1}) + O(2^\nu T \eta^2) \end{aligned} \quad (10)$$

Additionally, since if  $i \notin S_t$  we have  $\hat{w}_t^i \leq \xi' \leq \frac{\eta}{d}$ , we can say:

$$\sum_{i \in [d] \setminus i^*} \sum_{t:i \notin S_t} \ell_{t-1}^i (\hat{w}_t^i - 0) \leq T \eta \quad (11)$$

Combining Equations (10) and (11) (recall that  $w^*$  has 1 in  $i^*$  and 0 in other coordinates):

$$\begin{aligned} & \sum_{t=2}^T \eta \ell_{t-1}^{i^*} (\hat{w}_t^{i^*} - 1) + \eta \sum_{i \in [d] \setminus i^*} \sum_{t=2}^T \ell_{t-1}^i (\hat{w}_t^i - 0) \leq D_R(w^*, \hat{w}_1) - \sum_{t=2}^T D_R(\hat{w}_t, \hat{w}_{t-1}) + O(2^\nu T \eta^2) \\ & \iff \sum_{t=2}^T \langle \eta \ell_{t-1}, \hat{w}_t - w^* \rangle \leq D_R(w^*, \hat{w}_1) - \sum_{t=2}^T D_R(\hat{w}_t, \hat{w}_{t-1}) + O(2^\nu T \eta^2) \\ & \iff \sum_{t=2}^T \langle \ell_{t-1}, \hat{w}_t - w^* \rangle \leq \frac{1}{\eta} D_R(w^*, \hat{w}_1) - \frac{1}{\eta} \sum_{t=2}^T D_R(\hat{w}_t, \hat{w}_{t-1}) + O(2^\nu T \eta) \end{aligned}$$

From Lemma 27:

$$\text{Regret}(w^*) \leq \frac{1}{\eta} D_R(w^*, \hat{w}_1) + O(2^\nu T \eta)$$

■

**Lemma 40** Let  $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$  and  $\hat{\gamma} = (\{\hat{w}_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$  be OMD trajectory and OMD error trajectory, and assume  $T \geq 4$ ,  $\gamma$  is  $k$ -balanced and  $\varepsilon \leq \frac{1}{r'' \left( \frac{1}{2de^{k+1}} \right)^{T^2}}$ .

Then, for every  $t \in [T]$ ,  $\hat{w}_t^{i^*} \geq \frac{1}{de^{k+1}}$ .

**Proof** We will prove by induction on  $t$ . Since  $\hat{w}_1^i = \frac{1}{d}$  the base case holds.

Assume the statement is true for  $t-1$  and we prove for  $t$ .

If  $\hat{w}_t^{i^*} \geq \hat{w}_s^{i^*}$  for some  $s < t$  then the claim follows from the inductive assumption. If for every  $i \neq i^*$ ,  $\hat{w}_t^i \leq \frac{1}{d}$  we have that  $\hat{w}_t^{i^*} \geq \frac{1}{d}$  and the claim follows trivially. Proceeding, we consider the case that for every  $s < t$ ,  $\hat{w}_t^{i^*} < \hat{w}_s^{i^*}$  and there is some  $i \in [d]$  such that  $\hat{w}_t^i > \frac{1}{d}$ .

Since  $T \geq 3$  we have from the Lemma's assumptions and the induction assumptions that  $\varepsilon \leq \frac{1}{18r''(1/de^{k+1})} \leq \frac{1}{16r''(\hat{w}_{t-1}^i)}$ . From Lemma 25:

$$\hat{w}_t^{i*} \geq \frac{\hat{w}_{t-1}^{i*}}{2}$$

From this and the inductive assumption we have that for all  $s \in [1, t]$ ,  $\hat{w}_s^{i*} \geq \frac{1}{2de^{k+1}}$ . (We now want to improve this statement to  $\hat{w}_t^{i*} \geq \frac{1}{de^{k+1}}$ .)

Fix  $i$  to be the coordinate for which  $\hat{w}_t^i > \frac{1}{d}$ . We'll show that for every  $s \in [1, t]$ ,  $\hat{w}_s^i > \frac{1}{de^{k+1}}$ . Assume by contradiction that  $s$  is the last time  $\hat{w}_s^i \leq \frac{1}{de^{k+1}}$ . Again, since  $T \geq 3$  we have  $\varepsilon \leq 1/16r''(\hat{w}_{s+1}^i)$ , thus from Lemma 25:

$$\hat{w}_s^i \geq \frac{1}{2}w_{s+1}^i > \frac{1}{2de^{k+1}}$$

Which means that for every  $s' \in [s, t]$ ,  $\hat{w}_{s'}^i \geq \frac{1}{2de^{k+1}}$ . From Lemma 33 and our assumption on  $\varepsilon$  (see Definition 32 for the definition of  $B^i$ ):

$$B_{\hat{\gamma}}^i(s, t) \leq B_{\gamma}^i(s, t) + T\sqrt{r''\left(\frac{1}{2de^{k+1}}\right)\varepsilon} \leq k + 1$$

(we used  $r''$  because in our case, that  $c_1 = c_2 = 1$ , it is the same).

Recall that  $\hat{w}_t^{i*} < \hat{w}_s^{i*}$ , from Appendix C.3:

$$\hat{w}_t^i \leq e^{B_{\hat{\gamma}}^i(s, t)}\hat{w}_s^i \leq e^{k+1}\hat{w}_s^i \leq e^{k+1}\frac{1}{de^{k+1}} = \frac{1}{d}$$

Which is a contradiction to  $\hat{w}_t^i > \frac{1}{d}$ . Now we can continue assuming that for all  $s \in [1, t]$ ,  $\hat{w}_s^i > \frac{1}{de^{k+1}}$ .

From Lemma 11:

$$B_{\hat{\gamma}}^i(1, t) \leq B_{\gamma}^i(1, t) + T\sqrt{r''\left(\frac{1}{2de^{k+1}}\right)\varepsilon} \leq k + 1$$

Now, from Appendix C.3 (recall that  $\hat{w}_t^i > \frac{1}{d} = \hat{w}_1^i$ ):

$$\hat{w}_t^{i*} \geq \frac{\hat{w}_1^{i*}}{e^{k+1}} = \frac{1}{de^{k+1}},$$

which completes the inductive step and the proof.  $\blacksquare$

**Lemma** [Restatement of Lemma 15] Let  $\mathcal{K} = \Delta_d$  be the simplex, let  $\{\ell_t\}_{t=1}^T$  be an  $\alpha$ -balanced loss sequence, and let  $R(w) = \sum_{i=1}^d w_i \log w_i$  be the negative entropy regularizer. Assume  $\eta \leq 1/16$  and  $T \geq 3$ , if the approximation error satisfies

$$\varepsilon \leq \frac{1}{d \max\{6e^{\eta\alpha}, 1/\eta\}} \min\{\eta^4, 1/T^2\},$$

then the regret of any  $\varepsilon$ -approximate OMD trajectory is bounded as

$$\text{Regret}(w) \leq \frac{1}{\eta} D_R(w, w_1) + O(T\eta).$$

**Proof** From Lemma 10 the optimal trajectory is  $\alpha\eta$ -balanced. Since  $\varepsilon \leq \frac{1}{16de^{\alpha\eta+1}T^2} = \frac{1}{r''(1/16de^{\alpha\eta+1})T^2}$ , from Lemma 40 for every  $t \in [T]$ ,  $\hat{w}_t^{i*} \geq \frac{1}{de^{\alpha\eta+1}}$ .

We also have that  $\varepsilon \leq \frac{\eta^4}{r''(\min\{\eta/d, 1/de^{\alpha\eta+1}\})}$ , hence from Lemma 39 with  $\xi = 1/de^{\alpha\eta+1}$  and  $\nu = 1$  we get the desired results.  $\blacksquare$

## Appendix F. Proof of Theorem 6

**Lemma 41** For every  $\nu$ -barrier regularizer  $r$  and  $1 \geq w_2 \geq w_1 \geq 0$  we have:

$$r'(w_2) - r'(w_1) = \frac{c_1}{w_1^{\nu-1}} - \frac{c_1}{w_2^{\nu-1}}$$

**Proof**

$$\begin{aligned} r'(w_2) - r'(w_1) &= \int_{w_1}^{w_2} r''(w) dw \\ &\geq \int_{w_1}^{w_2} \frac{c_1}{w^\nu} dw \\ &= \frac{c_1}{w_1^{\nu-1}} - \frac{c_1}{w_2^{\nu-1}} \end{aligned}$$

■

**Lemma 42** Denote  $\psi = \left( \frac{c_1}{8\eta T d + c_1 (2d)^{\nu-1}} \right)^{1/(\nu-1)}$ . In the assumptions of Theorem 6, for every  $t, i$ :

$$w_t^i \geq \psi$$

**Proof** One can see that the assumptions of the theorem are that  $\varepsilon \leq \eta^4 \min \left\{ c_2, \frac{1}{c_2} \right\} \left( \frac{\psi}{2} \right)^\nu$ . We will now prove by induction that for every  $t \in [T]$ ,  $w_t^i \geq \psi$ .

Notice that  $\eta \leq \frac{1}{16c_1}$  and  $\varepsilon \leq \eta \psi^\nu \leq \frac{(w_t^i - 1)^\nu}{16c_1}$ , which means that from Lemma 25, we know for start that  $w_t^i \geq \psi/2$ . This means that for every  $t' \in [1, t]$ ,  $w_t^i \geq \psi/2$ . One can see that  $\varepsilon \leq c_2 \psi/4$  which means that we can use Lemma 11 with  $\psi/2$ . Since the balance of an exact trajectory is always bounded by  $\eta T$ , for every normalized  $v \in \ker(A)$ :

$$\begin{aligned} B_\gamma^v(1, t) &\leq T\eta + T \sqrt{c_2 \varepsilon \frac{2^\nu}{\psi}} \\ &\leq 2\eta T \end{aligned}$$

From Lemma 12 and the induction assumption, for every  $i \in [d]$  and  $t' \leq t$ :

$$r'(1/2d) - r'(w_t^i) \leq 8\eta T d$$

If  $w_t^i \leq 1/2d$  we can use Lemma 41:

$$\begin{aligned} \frac{c_1}{(w_t^i)^{\nu-1}} &\leq 8\eta T d + c_1 (2d)^{\nu-1} \\ \Rightarrow w_t^i &\geq \left( \frac{c_1}{8\eta T d + c_1 (2d)^{\nu-1}} \right)^{1/(\nu-1)} = \psi \end{aligned}$$

Else, i.e if  $w_t^i \geq 1/2d$ , we have:

$$\begin{aligned}
 w_t^i &\geq 1/2d \\
 &= \left( \frac{1}{(2d)^{\nu-1}} \right)^{1/\nu-1} \\
 &= \left( \frac{c_1}{c_1(2d)^{\nu-1}} \right)^{1/\nu-1} \\
 &\geq \left( \frac{c_1}{8\eta Td + c_1(2d)^{\nu-1}} \right)^{1/\nu-1} \\
 &= \psi
 \end{aligned}$$

Which ends the induction step. ■

**Proof of Theorem 6:** Since the polytope is a subset of the simplex, the diameter is bounded by 2. From Lemma 42, the effective smoothness of the trajectory is bounded by  $\beta := c_2/\psi^\nu$ . By the assumption about  $\varepsilon$  we have  $\varepsilon \leq \eta^4/\beta$ . From Lemma 22, for every  $t$ :

$$\langle \eta \ell_t + \nabla R(w_{t+1}) - \nabla R(w_t), w^* - w_{t+1} \rangle \geq -2\sqrt{2\beta\varepsilon} = \Theta(\eta^2)$$

From here it is straightforward standard OMD arguments:

$$\begin{aligned}
 \eta \ell_t \cdot (w_{t+1} - w^*) &\leq (\nabla R(w_{t+1}) - \nabla R(w_t)) \cdot (w^* - w_{t+1}) + \Theta(\eta^2) \\
 &= D_R(w^*, w_t) - D_R(w^*, w_{t+1}) - D_R(w_{t+1}, w_t) + \Theta(\eta^2)
 \end{aligned}$$

Summing for all  $t \in [T]$ :

$$\begin{aligned}
 \eta \sum_{t=1}^T \ell_t \cdot (w_{t+1} - w^*) &\leq D_R(w^*, w_1) - \sum_{t=1}^T D_R(w_{t+1}, w_t) + \Theta(\eta^2 T) \\
 \implies \sum_{t=1}^T \ell_t \cdot (w_{t+1} - w^*) &\leq \frac{1}{\eta} D_R(w^*, w_1) - \frac{1}{\eta} \sum_{t=1}^T D_R(w_{t+1}, w_t) + \Theta(\eta T)
 \end{aligned}$$

From Lemma 27:

$$\text{Regret}(w^*) \leq O\left(\frac{1}{\eta} D_R(w^*, w_1) + \Theta(\eta T)\right)$$
■

## Appendix G. Proof of Theorem 8

### G.1. Polytope definition

The polytope is defined as  $\{w \in \mathbb{R}^d : Aw = b \wedge (w_i \geq 0, \forall i \in [d])\}$  for  $A, b$  defined below. Denote  $m = 16 \log\left(\frac{1}{\varepsilon}\right)$ , we have  $d = 5m + 2$ . Additionally, we for assume for convenience that  $\varepsilon \leq \min\{4\eta, 2/T^{\frac{8}{\varepsilon}}\}$  (obviously proof that works for small  $\varepsilon$  works for bigger). This means  $m \geq 128 \log(2T)$ .

The matrix  $A$  has  $4m + 1$  rows. The first  $4m$  rows are, for every  $i \in [m]$ :

$$\begin{aligned} A_i &= e_{m+i} + e_{2m+i} - 2e_{3m+i} \\ A_{m+i} &= e_{m+i} - e_{2m+i} \\ A_{2m+i} &= e_i + 3e_{m+i} + e_{4m+i} \\ A_{3m+i} &= e_{4m+i} - e_{5m+1} \end{aligned}$$

The last row is:

$$A_{4m+1} = e_{5m+1} + e_{5m+2}$$

And:

$$b = A \sum_{i=1}^d \frac{1}{d} e_i$$

Namely,  $b$  is defined such that the point  $(\frac{1}{d}, \frac{1}{d}, \dots, \frac{1}{d})$  is in the polytope. Denote this point as  $w_1$ , the OMD will always start from here.

Denote the following set of  $m + 1$  vectors,  $\{v_i\}_i^{m+1}$ :

$$\begin{aligned} v_i &= 3e_i - e_{m+i} - e_{2m+i} - e_{3m+i} \quad \forall i \in [m] \\ v_{m+1} &= \sum_{i=1}^m e_i + e_{5m+2} - \sum_{i=4m+1}^{5m+1} e_i \end{aligned}$$

**Lemma 43**  $\{v_i\}_{i=1}^{m+1}$  is a basis for  $\ker(A)$ .

**Proof** One can notice that  $A$  is already in echelon form, so it is full ranked, which means that  $\dim \ker(A) = m + 1$ . Additionally, Every vector of  $v$  has a non-zero coordinate that's zeroed in all other vectors of  $v$ , so  $v$  is linear independent, which means that we only need to show that each of the vectors indeed nulls  $A$ .

For every  $i \in [1, m]$ ,  $v_i$  has common non-zero coordinates only with  $A_i, A_{m+i}, A_{2m+i}, A_{3m+i}$ . One can easily see that it nulls them. As for  $v_{m+1}$ , it has common non-zero coordinates with  $A_i$  for every  $i \in [2m + 1, 4m + 1]$ , which again can be seen easily to nullify. ■

**Lemma 44** *The polytope is a subspace of the simplex*

**Proof** To be inside the simplex all points of the polytope should satisfy two conditions - all coordinates greater than 0 and the sum of coordinates should be 1. The first is by definition in this polytope.

Let  $w$  be some point in the polytope. Since  $Aw = b$  and  $Aw_1 = b$ ,  $w - w_1 \in \ker(A)$ . From Lemma 43, we can write:

$$w = w_1 + \sum_i \alpha_i v_i$$

For some  $\alpha_i \in \mathbb{R}$ .

All the vectors in  $v$  has the sum of their coordinates 0. Thus, the sum of coordinates of  $w$  is the same as  $w_1$ , concluding the proof. ■

## G.2. General settings and hardness event

Since we want to prove a lower bound of the form  $T \sqrt{\frac{\eta}{\log(\frac{1}{\varepsilon})}} = \Theta\left(T \sqrt{\frac{\eta}{d}}\right)$ , and there's a known lower bound for  $T\eta$ , we can assume  $d \leq \frac{1}{\eta}$ .

The losses for the first  $m$  coordinates is constant 0, for the  $[m+1, 4m]$  coordinates it's constant 1, for the  $[4m+1, 5m]$  coordinates it's gaussian with mean 0 and variance 1, for the  $5m+1$ th coordinate it is gaussian with mean  $\sqrt{\eta d} \leq 1$  and variance 1 and for the  $5m+2$ th coordinate it's constant 0.

Denote  $\tau = \frac{3}{\eta}$ . We define the hardness event  $E$  to be the following events:

$$\begin{aligned} \sum_{i=4m+1}^{5m+1} \ell_{i,\tau}^i &\leq 0 \\ \sum_{i=4m+1}^{5m+1} \ell_t^i &\leq \frac{m}{16} \quad \forall t \in [T] \end{aligned}$$

### Lemma 45

$$\Pr(E) = \Omega(1)$$

**Proof** Denote  $G = \sum_{i=4m+1}^{5m+1} \ell_{i,\tau}^i$ . Since  $G$  is a sum of gaussian random variables, it is also a gaussian random variable, denote its mean with  $\mu$  and variance  $\sigma^2$ . Simple calculation shows that  $\mu = \sqrt{d\eta}\tau = 2\sqrt{\frac{d}{\eta}}$  and  $\sigma^2 = \tau(m+1) = \frac{2(m+1)}{\eta}$ . Since  $m = \Theta(d)$ , we have that  $\mu = \Theta(\sigma)$ . It is a general attribute of a gaussian that in such case the probability of having  $G \leq 0$  is  $\Theta(1)$ .

Fix  $t \in [T]$ . Using Hoeffding inequality we have that w.p  $\frac{1}{2T}$ :

$$\sum_{i=4m+1}^{5m+1} \ell_t(i) \leq \sqrt{\frac{m+1}{2} \log(2T)}$$

Since  $\log(2T) \leq \frac{m}{128}$ :

$$\sum_{i=4m+1}^{5m+1} \ell_t(i) \leq \frac{m}{16}$$

Union bound on all  $t \in [T]$  concludes the proof. ■

For every  $t \in [T]$  we know that  $w_t - w_1 \in \ker(A)$ . From Lemma 43, it can be written as a linear combination of  $v$ . Denote the coefficients as  $\alpha$ , namely:

$$w_t = w_1 + \sum_{i=1}^{m+1} \alpha_t^i v_i$$

## G.3. Analysis

**Lemma 46** For every  $i \in [m]$ ,  $w_\tau^i \geq \frac{5}{2d}$

**Proof** Assume by contradiction that  $w_\tau^i \geq \frac{3}{d}$ . One can notice that the  $(m+i, 2m+i, 3m+i)$  are only in  $v_i$  with the same coefficient, which means that  $w_t^{m+i} = w_t^{2m+i} = w_t^{3m+i}$ .

Notice that  $\eta \langle \ell_{:\tau}, v_i \rangle = -3\eta\tau = -9$ . From Lemma 10:

$$\begin{aligned}
 -9 &= \langle \nabla R(w_1) - \nabla R(w_t), v_i \rangle \\
 &= 3 \log \left( \frac{w_1^i}{w_\tau^i} \right) + \log \left( \frac{w_\tau^{m+i}}{w_1^{m+i}} \right) + \log \left( \frac{w_\tau^{2m+i}}{w_1^{2m+i}} \right) + \log \left( \frac{w_\tau^{3m+i}}{w_1^{3m+i}} \right) \\
 &\geq 3 \log \left( \frac{1}{3} \right) + 3 \log \left( \frac{w_\tau^{m+i}}{w_1^{m+i}} \right) \\
 \iff -3 &\geq \log \left( \frac{1}{3} \right) + \log (dw_\tau^{m+i}) \\
 &= \log \left( \frac{dw_\tau^{m+i}}{3} \right) \\
 \iff w_\tau^{m+i} &\leq \frac{3}{e^3 d} \leq \frac{1}{6d}
 \end{aligned}$$

Which means that  $\alpha_i \geq \frac{5}{6d}$ . Additionally,  $\alpha_{m+1} \geq -\frac{1}{d}$ , since else it violates  $e_{5m+2} \geq 0$ . We have:

$$w_\tau^i = \frac{1}{d} + 3\alpha_i + \alpha_{m+i} \geq \frac{5}{2d}$$

■

**Lemma 47** Assume  $E$  and optimal trajectory, we have  $\alpha_\tau^{m+1} \leq -\frac{1}{d} + e^{m/8}$

**Proof** We'll first show that  $\alpha_\tau^{m+1} \leq 0$ . Assume the opposite by contradiction. This means that  $w_\tau^i \leq w_1^i$  for all  $i \in [4m+1, 5m+1]$  and  $w_\tau^{5m+2} \geq w_1^{5m+2}$ . Together with Lemma 46 this means that in the positive elements of  $v_{m+1}$  we have  $w_\tau^i > w_1^i$  and in the negative elements we have  $w_\tau^i < w_1^i$ . This means that  $\langle \nabla R(w_1) - \nabla R(w_t), v_{m+1} \rangle < 0$ . From  $E$  we have  $\langle \ell_{:\tau}, v_{m+1} \rangle \geq 0$ , which is a contradiction to Lemma 10.

We continue assuming  $\alpha_\tau^{m+1} \leq 0$ . Notice that  $\alpha_\tau^{m+1} \geq -\frac{1}{d}$  to satisfy the constraint  $w_\tau^{5m+2} \geq 0$ , so for every  $i \in [4m+1, 5m+1]$ , we have  $w_\tau^i \leq \frac{1}{2d}$ . From  $E$  and Lemma 10:

$$\begin{aligned}
 0 &\leq - \sum_{i=4m+1}^{5m+1} \ell_{:\tau}^i \\
 &= \sum_{i=1}^m \ell_{:\tau}^i + \ell_{:\tau}^{5m+2} - \sum_{i=4m+1}^{5m+1} \ell_{:\tau}^i \\
 &= \langle \ell_{:\tau}, v_{m+1} \rangle
 \end{aligned}$$

From Lemma 10:

$$\begin{aligned}
 0 &\leq \langle \nabla R(w_1) - \nabla R(w_t), v_{m+1} \rangle \\
 &= \sum_{i=1}^m \log \left( \frac{w_1^i}{w_\tau^i} \right) + \log \left( \frac{w_1^{5m+2}}{w_\tau^{5m+2}} \right) - \sum_{i=4m+1}^{5m+1} \log \left( \frac{w_1^i}{w_\tau^i} \right)
 \end{aligned}$$

Since  $w_\tau^i \leq \frac{1}{2d}$ , we have  $\frac{w_1^i}{w_\tau^i} \geq 2$ . From Lemma 46, we have  $\frac{w_1^i}{w_\tau^i} \leq 2.5$ . Thus:

$$m \log(2.5) - (m+1) \log(2) + \log\left(\frac{w_1^{5m+2}}{w_\tau^{5m+2}}\right) \geq 0$$

Since  $m \geq 8$ ,  $(m+1) \log(2) \leq m \log(2.2)$ . Thus:

$$\begin{aligned} m \log(2.5/2.2) + \log\left(\frac{w_1^{5m+2}}{w_\tau^{5m+2}}\right) &\geq 0 \\ \iff w_\tau^{5m+2} &\leq e^{m/8} \end{aligned}$$

Since  $w_\tau^{5m+2}$  can only be altered with  $\alpha_\tau^{m+1}$ , this concludes the proof.  $\blacksquare$

**Lemma 48** For some  $t \in [T]$ , assume  $\alpha_{t-1}^{m+1} \leq -\frac{1}{d} + e^{-m/8}$  and  $E$ . There is an  $\varepsilon$ -approximate step for which  $\alpha_t^{m+1} \leq -\frac{1}{d} + e^{-m/8}$ .

**Proof** First we show that for the optimal step,  $\alpha_t^{m+1} \leq -\frac{1}{d} + e^{-m/16}$

If  $\alpha_t^{m+1} \leq \alpha_{t-1}^{m+1}$  the proof concludes from the assumption. Continuing assuming the opposite. This means that for every  $i \in [4m+1, 5m+1]$ ,  $w_t^i \leq w_{t-1}^i$ .

Additionally, we'll show that for every  $i \in [1, m]$ ,  $w_t^i \geq w_{t-1}^i$ . Assume otherwise for some  $i$ , since  $\alpha_t^{m+1} \geq \alpha_{t-1}^{m+1}$  it means that  $\alpha_t^i \leq \alpha_{t-1}^i$ , which means that for every  $j \in \{m+i, 2m+i, 3m+i\}$  we have  $w_t^j \geq w_{t-1}^j$ . This means that  $\langle \nabla R(w_{t-1}) - \nabla R(w_t), v_i \rangle \geq 0$ , which contradicts Lemma 10 (as  $\langle \ell_t, v_i \rangle$  has a constant value of  $-1$ ).

From the second part of  $E$  and Lemma 10:

$$\begin{aligned} -\frac{m}{16} &\leq \eta \langle \ell_t, v_{m+1} \rangle \\ &= \langle \nabla R(w_{t-1}) - \nabla R(w_t), v_{m+1} \rangle \\ &= \sum_{i=1}^m \log\left(\frac{w_{t-1}^i}{w_t^i}\right) + \log\left(\frac{w_{t-1}^{5m+2}}{w_t^{5m+2}}\right) - \sum_{i=4m+1}^{5m+1} \log\left(\frac{w_{t-1}^i}{w_t^i}\right) \\ &\leq \log\left(\frac{w_{t-1}^{5m+2}}{w_t^{5m+2}}\right) \\ \iff w_t^{5m+2} &\leq w_{t-1}^{5m+2} e^{m/16} \\ &\leq e^{-m/16} \end{aligned}$$

Since the  $5m+2$ th coordinate is controlled only by  $v_{m+1}$ , this concludes the fact that  $\alpha_t^{m+1} \leq -\frac{1}{d} + e^{-m/16} = -\frac{1}{d} + \varepsilon$ , which means that  $\alpha_t^{m+1} \leq \alpha_{t-1}^{m+1} + \varepsilon$ .

Next we argue that in the optimal step, for every  $i \in [1, m]$ ,  $\alpha_t^i \geq \alpha_{t-1}^i$ . Assume otherwise for some  $i$ . This means that for every  $j \in \{m+i, 2m+i, 3m+i\}$  we have  $w_t^j \geq w_{t-1}^j$ . Additionally, it means that  $w_t^i \leq w_{t-1}^i + \varepsilon$ .

From Lemma 10:

$$\begin{aligned}
 -3\eta &= \eta \langle \ell_t, v_i \rangle \\
 &= \langle \nabla R(w_{t-1}) - \nabla R(w_t), v_i \rangle \\
 &= 3 \log \left( \frac{w_{t-1}^i}{w_t^i} \right) + \log \left( \frac{w_t^{m+i}}{w_{t-1}^{m+i}} \right) + \log \left( \frac{w_t^{2m+i}}{w_{t-1}^{2m+i}} \right) + \log \left( \frac{w_t^{3m+i}}{w_{t-1}^{3m+i}} \right) \\
 &\geq 3 \log \left( \frac{w_{t-1}^i}{w_t^i} \right) \\
 &\geq 3 \log \left( \frac{w_t^i - \varepsilon}{w_t^i} \right) \\
 \implies -\eta &\geq \log \left( 1 - \frac{\varepsilon}{4} \right) \geq -\frac{\varepsilon}{4} \\
 \varepsilon &\geq 4\eta
 \end{aligned}$$

Which contradicts our assumption that  $\varepsilon < 4\eta$ .

By now we showed that if the  $t$ th step is optimal, we have  $\alpha_{t-1}^{m+1} \leq \alpha_t^{m+1} \leq \alpha_{t-1}^{m+1} + \varepsilon$  and for all  $i \in [m]$ ,  $\alpha_t^i \geq \alpha_{t-1}^i$ .

We next argue that if we keep the same  $\alpha_t^i$  for all  $i \in [m]$  but change  $\alpha_t^{m+1}$  to be equal to  $\alpha_{t-1}^{m+1}$ , this will be an  $\varepsilon$ -approximate step.

First we notice that all  $w_t$  is now closer to  $w_{t-1}$ , which means that the bregman divergence only shrinks from that change. Indeed, coordinates  $[4m+1, 5m+2]$  are only getting closer from the change, coordinates  $[m+1, 4m]$  doesn't change (the change in  $v_{m+1}$  doesn't affect them). Finally, since for all  $i \in [m]$ ,  $\alpha_t^i \geq \alpha_{t-1}^i$ , we still have  $w_t^i \geq w_{t-1}^i$ , which means that those coordinates also got closer.

The proof concludes from the fact that from the second part of  $E$ , the first term in the objective can only be changed in  $\frac{m\eta\varepsilon}{16} < \varepsilon$ . ■

**Theorem 49** *There is an  $\varepsilon$ -approximate trajectory that get a regret of:*

$$\Omega \left( T \sqrt{\frac{\eta}{\log \left( \frac{1}{\varepsilon} \right)}} \right)$$

**Proof** From Lemmas 47 and 48 we get that there is an  $\varepsilon$ -approximate trajectory such that for every  $t \geq \tau$ ,  $\alpha_t^{m+1} \leq -\frac{1}{d} + e^{-m/8} \leq 0$ , which means that  $w_t^{5m+1} \geq \frac{1}{d}$ . The total expected loss of this coordinate is  $T\sqrt{d\eta}$ , which means that this trajectory suffers a loss of  $\Omega \left( T \sqrt{\frac{\eta}{\log \left( \frac{1}{\varepsilon} \right)}} \right)$ .

Now we only need to show that there is a point in the polytope that gets zero loss. Indeed, one can see that if  $\alpha^i = \frac{1}{d}$  for all  $i \in [1, m+1]$ , the point  $w = w_1 + \sum_i \alpha^i v_i$  has all coordinates with non-zero loss ( $[m+1, 4m], 5m+1$ ) to be zeroed. ■

## Appendix H. Approximate FTRL

In this section we analyze an approximate version of the Follow The Regularized Leader (FTRL) algorithm, analogous to the approximate OMD algorithm analyzed in the paper, shown in Algorithm 1. Our simple analysis shows that  $\varepsilon$ -approximate updates in FTRL gives rise to an additional

**Algorithm 1:** Approximate Follow The Regularized Leader

**Input:**  $\eta > 0$ , regularization function  $R$ , and a bounded, convex, and closed set  $\mathcal{K}$

$\tilde{w}_1 \leftarrow \arg \min_{w \in \mathcal{K}} \{R(w)\}$

**for**  $t = 1$  **to**  $T$  **do**

    Play  $\tilde{w}_t$  and observe loss  $\ell_t$

    Define

$$\phi_t(w) := \eta \sum_{s=1}^t \langle w, \ell_s \rangle + R(w)$$

    Compute  $\tilde{w}_{t+1}$  such that

$$\phi_t(\tilde{w}_{t+1}) \leq \min_{w \in \mathcal{K}} \phi_t(w) + \varepsilon$$

**end**

additive  $O(\sqrt{\varepsilon T})$  term in the regret, implying that polynomially small error suffices for optimal regret performance (i.e.,  $\varepsilon = O(1/T)$  for  $O(\sqrt{T})$  regret).

Our analysis closely follows the standard FTRL analysis, and specifically, arguments appearing in Hazan (2016). For convenience, we denote  $f_t(w) = \langle \ell_t, w \rangle$  for  $t \geq 1$  and  $f_0 = (1/\eta)R$ . Additionally, we denote  $w_{t+1}$  to be the exact minimizer of  $\phi_t$  over  $\mathcal{K}$ . We will need the following “be the leader” lemma, the proof of which can be found in Hazan (2016).

**Lemma 50 (Lemma 5.4 in Hazan, 2016)** *For every  $w^* \in \mathcal{K}$ :*

$$\sum_{t=0}^T f_t(w^*) \geq \sum_{t=0}^T f_t(w_{t+1}).$$

The key fact about the approximate minimizers  $\tilde{w}_t$  is the following.

**Lemma 51** *For every  $t \in [T]$ :*

$$\|\tilde{w}_t - w_{t+1}\| \leq 2\eta + \sqrt{2\varepsilon}.$$

**Proof** Since  $w_{t+1}$  is the minimizer of  $\phi_t$ , from first order optimality conditions:

$$\langle \nabla \phi_t(w_{t+1}), \tilde{w}_t - w_{t+1} \rangle \geq 0.$$

Thus:

$$\begin{aligned} D_{\phi_t}(\tilde{w}_t, w_{t+1}) &= \phi_t(\tilde{w}_t) - \phi_t(w_{t+1}) - \langle \nabla \phi_t(w_{t+1}), \tilde{w}_t - w_{t+1} \rangle \\ &\leq \phi_t(\tilde{w}_t) - \phi_t(w_{t+1}). \end{aligned}$$

Since a linear term doesn’t change the Bregman divergence, this also upper bounds  $D_R(\tilde{w}_t, w_{t+1})$ .

Thus:

$$\begin{aligned} D_R(\tilde{w}_t, w_{t+1}) &\leq \phi_t(\tilde{w}_t) - \phi_t(w_{t+1}) \\ &= \phi_{t-1}(\tilde{w}_t) - \phi_{t-1}(w_{t+1}) + \eta \langle \ell_t, \tilde{w}_t - w_{t+1} \rangle \\ &\leq \varepsilon + \eta \langle \ell_t, \tilde{w}_t - w_{t+1} \rangle \\ &\leq \varepsilon + \eta \|\ell_t\|_* \|\tilde{w}_t - w_{t+1}\| \\ &\leq \varepsilon + \eta \|\tilde{w}_t - w_{t+1}\|. \end{aligned}$$

From the 1-strong convexity of  $R$ :

$$\begin{aligned}
 D_R(\tilde{w}_t, w_{t+1}) &\geq \frac{1}{2} \|\tilde{w}_t - w_{t+1}\|^2 \\
 \implies 2\varepsilon + 2\eta \|\tilde{w}_t - w_{t+1}\| &\geq \|\tilde{w}_t - w_{t+1}\|^2 \\
 \implies \|\tilde{w}_t - w_{t+1}\| &\leq 2\eta + \sqrt{2\varepsilon}.
 \end{aligned}
 \tag{Lemma 23}$$

■

We can now prove the main result of this section:

**Theorem 52** *The regret of Algorithm 1 is bounded by:*

$$\text{Regret}(w^*) \leq \frac{R(w^*) - R(w_1)}{\eta} + (2\eta + \sqrt{2\varepsilon})T.$$

**Proof**

$$\begin{aligned}
 \text{Regret}(w^*) &= \sum_{t=1}^T \langle \ell_t, \tilde{w}_t - w^* \rangle \\
 &= \sum_{t=1}^T \langle \ell_t, \tilde{w}_t - w^* \rangle + \frac{R(w^*) - R(w^*)}{\eta} \\
 &\leq \sum_{t=1}^T \langle \ell_t, \tilde{w}_t - w_{t+1} \rangle + \frac{R(w^*) - R(w_1)}{\eta} && \text{(Lemma 50)} \\
 &\leq \sum_{t=1}^T \|\ell_t\|_* \|\tilde{w}_t - w_{t+1}\| + \frac{R(w^*) - R(w_1)}{\eta} \\
 &\leq (2\eta + \sqrt{2\varepsilon})T + \frac{R(w^*) - R(w_1)}{\eta}. && \text{(Lemma 51)}
 \end{aligned}$$

■