The Hidden Cost of Approximation in Online Mirror Descent

Ofir Schlisselberg OFIRS4@MAIL.TAU.AC.IL
Uri Sherman URISHERMAN@MAIL.TAU.AC.IL
Tel Aviv University
Tomer Koren TKOREN@TAUEX.TAU.AC.IL
Yishay Mansour MANSOUR.YISHAY@GMAIL.COM
Tel Aviv University and Google Research

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 [7, 23] 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; [15, 29]) is a generalization of the basic framework adapted to the more general online learning setup [37], where the goal of the learner is to minimize her *regret*, defined is 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 $t = 1, \ldots, T$ follow the update rule:

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

where ℓ_t is the loss at time t and D_R is the Bregman divergence associated with a regularizer $R \colon \mathcal{K} \to \mathbb{R}$ chosen by the learner. Notable instances of OMD include online gradient descent [37] and the well known multiplicative weights method [4, 12, 19], 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 [17, 26, 35]; generic online convex optimization algorithms that rely on OMD updates [1, 14, 16]; and algorithms defined over the simplex that use barrier regularization other than negative entropy, such as in adversarial bandits [2, 36] and portfolio selection [21]. 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 [10, 11, 27, 32]. (Due to space constraints, discussion of additional related work is deferred to Appendix A.)

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.

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 β, we establish a tight $\Theta(TD\sqrt{\beta\varepsilon}/\eta)$ bound on the excess regret due to ε-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 $\varepsilon = O(\eta^3/(T^2 d^2))$ 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 $\eta = \widetilde{O}(1/\sqrt{T})$, we show that a polynomially small error $\varepsilon = O(1/(dT^2))$ suffices for obtaining optimal regret with high probability, as 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 $\widetilde{\Omega}(T^{2/3})$ regret for any step size unless ε 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 algorithm [28], which re-optimizes against the cumulative loss at each round, and thus (at least in the linear case with an oblivious adversary), each optimization round is independent of the previous ones.

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. 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 5 and the discussion that follows).

The rest of this extended abstract focuses on our results for the general adversarial setting; the presentation of the results for the stochastic setting is deferred to the appendix due to space constraints. 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., [9]).

2. Preliminaries

We consider the standard online linear optimization setup, where at each round $t = 1, 2, \ldots, T$, 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:

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

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 $R : \mathcal{K} \to \mathbb{R}$ 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 *ε-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 ε additive error:

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

Regret bounds for OMD typically depend on the *diameter* of $\mathcal{K}$ with respect to the norm $\|\cdot\|$, given by $D = \max_{w,w'\in\mathcal{K}} \|w - w'\|$.

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

Definition 1 (coordinate separable barrier regularizers). We say $R\colon \mathcal{K} \to \mathbb{R}$ is a coordinate separable barrier[1] regularizer with parameter $\nu \ge 1$ (or simply a ν-barrier) if there exists a twice-differentiable function $r\colon [0,1] \to \mathbb{R}$ and $c_1, c_2 > 0$ such that:

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

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 ν will turn out to be directly associated with the robustness of OMD with ν-barrier regularization to approximation errors.

3. Overview of Results and Techniques

Smooth regularizers. We begin by establishing tight upper and lower bounds for approximate OMD with smooth regularizers,[2] 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.

Theorem 2. *Let $\mathcal{K} \subseteq \mathbb{R}^d$ be a convex set with diameter D, and let $R\colon \mathcal{K} \to \mathbb{R}$ be a β-smooth regularizer over $\mathcal{K}$. Then, for any loss sequence $\ell_1, \ldots, \ell_T \in [-1,1]^d$, the regret of any ε-approximate OMD trajectory compared to any $w \in \mathcal{K}$ is bounded as:*

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

1. Strictly speaking, these are barriers for the positive orthant in $\mathbb{R}^d$.
2. A function R is said to be β-smooth with respect to a norm $\|\cdot\|$ if its gradient is β-Lipschitz; $\|\nabla R(x) - \nabla R(y)\|_* \le \beta\|x - y\|$ for all $x, y \in \mathcal{K}$, where $\|\cdot\|_*$ is the norm dual to $\|\cdot\|$.

The proof follows the standard OMD analysis, replacing exact optimality with *approximate* optimality conditions. Indeed, for any β-smooth convex objective $f : \mathcal{K} \to \mathbb{R}$, if $f(\hat{w}) - \arg\min_{w\in\mathcal{K}} f(w) \le \varepsilon$, then one can show that (see Lemma 15):

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

Applying the above on ϕ_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 β compared to a similar result of [10]. 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 ε-approximate OMD over $\mathcal{K} = [0, D]$ with the β-smooth regularizer $R(\cdot) = \frac{\beta}{2}\|\cdot\|_2^2$. Then there exists a loss sequence, an ε-approximate OMD trajectory and $w \in \mathcal{K}$ such that:*

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

To see why this is true, consider the constant loss sequence $\ell_t = \min\left\{\sqrt{2\beta\varepsilon}/\eta, 1\right\}$ 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 ε-minimizer of ϕ_t; let w_{t+1}^* be the exact minimizer of ϕ_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 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.

Barrier regularizers. 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 ν-barrier (Definition 1) is unbounded since $r''(x) \to \infty$ as $x \to 0$. Interestingly, the robustness behavior of these barriers varies dramatically with ν: 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 $\mathcal{K} = \Delta_d$, $d = 2$, and R be the negative entropy over $\mathcal{K}$. Suppose that the approximation error satisfies $\varepsilon \ge 4\eta e^{-\eta T/3}$. Then there exists a sequence of losses $\ell_1, \ldots, \ell_T \in [0,1]^d$ for which there exists an ε-approximate OMD trajectory that suffers regret* $\mathrm{Regret}(w^*) = \Omega(T)$.

The exponential dependence of ε on the time horizon T is in fact tight: if ε is exponentially small in ηT, the standard regret guarantees are recovered (see Theorem 6 in Appendix B.1). 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.

We now turn our attention to ν-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 5. *Let $\mathcal{K} \subseteq \Delta_d$ be a polytope that contains the uniform distribution and the OMD is initialized there*[3]*, let $R\colon \mathcal{K} \to \mathbb{R}$ be a ν-barrier regularizer (cf. Definition 1) with $\nu > 1$ and $\eta \le \frac{1}{16c_1}$. If $\varepsilon \le \eta^4 \min\left\{\frac{1}{c_2}, c_2\right\}\left(\frac{16\eta T d + 2c_1(2d)^{\nu-1}}{c_1}\right)^{-\nu/(\nu-1)}$, then for any loss sequence $\ell_1, \ldots, \ell_T \in [-1,1]^d$, the regret of any ε-approximate OMD trajectory compared to any $w \in \mathcal{K}$ is bounded as:*

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

The principle underlying the analysis of Theorem 5 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:

$$r'(w_t) = r'(w_{t-1}) - \eta \ell_{t-1} = r'(w_1) - \eta \sum_{s=1}^{t-1} \ell_s,$$
$$\implies -\tfrac{1}{w_t^{\nu-1}} \ge -1 - \eta T \quad \implies \quad w_t^{\nu-1} \ge \tfrac{1}{\eta T}.$$

Thus, 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 $\log(w_t) \ge 0 - \eta T \implies w_t \ge e^{-\eta T}$. 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.

Improved robustness with stochastic losses. Theorem 5 presented above follows in principle by reducing to the smooth case; when the iterates remain bounded away from zero, the regularizer behaves as if it were smooth, and the same argument given for smooth regularizers applies. Theorem 4 builds on a similar idea but in the other direction; with negative entropy, the iterates may become exponentially small and thus follows an exponential lower bound. Interestingly, with stochastic losses we obtain a polynomial upper bound (Theorem 7) *despite the fact that the iterates may be exponentially close to zero*, where the effective smoothness is exponentially large. To achieve this, we show that the iterates may reach close to zero only when also approaching w^*, and that closeness to w^* counteracts the large effective smoothness in that region. For more details on the stochastic setting, see Appendix B.2.

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

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.

References

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] Amir Beck. *First-order methods in optimization*. SIAM, 2017.

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

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

[9] Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.

[10] 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.

[11] 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.

[12] 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.

[13] Vincent Guigues. Inexact stochastic mirror descent for two-stage nonlinear stochastic programs. *Mathematical Programming*, 187(1):533–577, 2021.

[14] 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.

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

[16] 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.

[17] 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.

[18] S Kabbadj. Inexact version of bregman proximal gradient algorithm. In *Abstract and Applied Analysis*, volume 2020, page 1963980. Wiley Online Library, 2020.

[19] Nick Littlestone and Manfred K Warmuth. The weighted majority algorithm. *Information and computation*, 108(2):212–261, 1994.

[20] 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.

[21] Haipeng Luo, Chen-Yu Wei, and Kai Zheng. Efficient online portfolio with logarithmic regret. *Advances in neural information processing systems*, 31, 2018.

[22] Arkadij Semenovič Nemirovskij and David Borisovich Yudin. Problem complexity and method efficiency in optimization, 1983.

[23] 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.

[24] 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.

[25] R Tyrrell Rockafellar. Monotone operators and the proximal point algorithm. *SIAM journal on control and optimization*, 14(5):877–898, 1976.

[26] 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.

[27] 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.

[28] Shai Shalev-Shwartz. 2012. doi: 10.1561/2200000018.

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

[30] 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.

[31] Marc Teboulle. A simplified view of first order methods for optimization. *Mathematical Programming*, 170(1):67–96, 2018.

[32] Silvia Villa, Saverio Salzo, Luca Baldassarre, and Alessandro Verri. Accelerated and inexact forward-backward algorithms. *SIAM Journal on Optimization*, 23(3):1607–1633, 2013.

[33] 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.

[34] 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.

[35] 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.

[36] 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.

[37] 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 A. Related Work

Mirror descent [7, 22] and the online convex optimization framework [37] 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 [6, 8, 15, 29]. 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; [5, 20], see also [31]). The BPGM is a generalization of the proximal gradient method [25] where a Bregman divergence replaces the norm proximity regularizer, and the objective is required to satisfy the weaker *relative smoothness* property [5]. 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., [3, 27, 32, 34]. Some works study the online case [11] with the euclidean regularizer, and some further generalize to the online BPGM but with smooth regularizers [10].

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 [18, 24, 30, 33]—which is to be contrasted with characterizing convergence in terms of the ad-hoc approximation errors. As one example, the work of Kabbadj [18] establishes that the inexact BPGM achieves the same rate of the exact version (aka NoLips; 5) as long as the approximation errors are smaller than the Bregman distance to the previous iterate. More recently, Yang and Toh [33] propose variants with several advantages at the expense of a somewhat more involved subproblem optimization procedure.

Finally, the work of Guigues [13] 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.

Appendix B. Additional main results

In this section, we present additional main results that were not included in the main text due to space constraints.

B.1. Negative entropy upper bound

Theorem 6. *Let $\mathcal{K} = \Delta_d$ and R be the negative entropy over $\mathcal{K}$. Assume $\eta \leq 1/16$ and $T \geq 3$, if $\varepsilon \leq \frac{1}{6d} e^{-\eta T/2} \min\left\{\eta^4, T^{-2}\right\}$, then for any loss sequence $\ell_1, \ldots, \ell_T \in [-1, 1]^d$, the regret of any*

ε-approximate OMD trajectory compared to any $w \in \mathcal{K}$ is bounded as:

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

The proof is deferred to Appendix G.

B.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 on 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 is uniform, $T \ge 256$, $\eta = \sqrt{\frac{\log(d)}{T}}$ and $\varepsilon \le \frac{\delta}{6d^2T^4}$. Then with probability $\ge 1 - \delta$ over the choice of an i.i.d. loss sequence $\ell_1, \ldots, \ell_T \in [-1, 1]^d$, the regret of any ε-approximate OMD trajectory compared to any $w \in \mathcal{K}$ is bounded as:*

$$\mathrm{Regret}(w) \le O\left(\sqrt{T \log(d)}\right)$$

However, this robustness does not extend to general domains. The geometry of the feasible set can reintroduce sensitivity: even with the same regularizer and similarly well-behaved stochastic losses, restricting the domain to a polytope 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 ε-approximate trajectory and $w \in \mathcal{K}$ such that:*

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

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

Detailed proofs of the theorems in this section are provided in Appendices G and I.

Appendix C. General Lemmas and definitions

Definition 9. We call $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$ an **exact OMD trajectory** if for every $t \in [T]$:

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

Definition 10. We call $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$ an ε-**approximate OMD trajectory** with some $\varepsilon > 0$ if for every $t \in [T]$ w_{t+1} is an ε-minimizer of $\eta\langle \ell_t,\, w\rangle + D_R(w_{t-1}, w)$.

Definition 11. Our assumptions about the regularizers are:

- There is a function $r : [0, 1] \to \mathbb{R}$ 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 12. We say that a function $F : \mathcal{W} \to \mathbb{R}$ ($\mathcal{W} \subseteq \mathbb{R}^d$) is coordinate-separated if there are functions $f_1, f_2, \ldots, f_d$ such that $F(w) = \sum_i f_i(w_i)$ for all $w \in \mathcal{W}$.

Definition 13. Let $F : \mathcal{W} \to \mathbb{R}$ 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 14. *Let $F : \mathcal{W} \to \mathbb{R}$ be a coordinate-separated function and Let $x_1, x_2 \in \mathcal{W}$. If β 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 15. *Let $\|\cdot\|$ be any norm, and let $f : \mathcal{W} \to \mathbb{R}$, and let $\hat{w}, w \in \mathcal{W}$ where $\hat{w}$ is an ε-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 β (with ℓ_1 or ℓ_2 norm, see Lemma 14) or any general β-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 16. *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 17. *Let $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$ be an ε-approximate trajectory above the simplex with $\eta \leq \frac{1}{4}$ and coordinate-separable regularizer. Let $h = \min\left\{r''(w_t^i), r''(w_{t+1}^i)\right\}$. Then for any $i \in [d]$:*

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

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

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

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

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

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 ith 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:

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

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 ≥ 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 $\|\tilde{w} - w_{t+1}\| = 2(w^i_{t+1} - w^i_t)$. Thus, from Holder:

$$\eta\langle \ell_t,\, w_{t+1} - \tilde{w}\rangle \geq -2\eta\left|w^i_{t+1} - w^i_t\right|$$

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

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

From Lemma 16 we get:

$$\left|w^i_t - w^i_{t+1}\right| < \frac{4\eta}{h} + \sqrt{\frac{\varepsilon}{h}}$$

■

Lemma 18. *Let $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$ be an ε-approximate trajectory above the simplex with $\eta \leq \frac{1}{16c_1}$ and ν-barrier regularizer. Let $t \in [T]$ and $i \in [d]$ be such that $\varepsilon \leq \frac{(w^i_t)^\nu}{16c_1}$, then:*

$$\begin{aligned}w^i_{t-1} &\geq \frac{1}{2}w^i_t \\ w^i_{t+1} &\geq \frac{1}{2}w^i_t\end{aligned}$$

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

We have:

$$\min\left\{r''(w^i_t), r''(w^i_{t-1})\right\} \geq \frac{c_1}{\max\left\{w^i_t, w^i_{t-1}\right\}^\nu} = \frac{c_1}{(w^i_t)^\nu} \geq \frac{c_1}{w^i_t}$$

From Lemma 17:

$$w^i_{t-1} \geq w^i_t - 4c_1\eta w^i_t - \sqrt{c_1\varepsilon w^i_t} \tag{6}$$

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

$$4c_1\eta w^i_t \leq \frac{w^i_t}{4} \tag{7}$$

From the assumption on ε and the fact that $r''(w^i_t) \geq w^i_t$:

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

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

Lemma 19 (Three-points identity). *For every differentiable function R:*

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

Proof.

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

■

Lemma 20 (OMD Helper).

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

Proof. From the strong convexity of R:

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

By Holder:

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

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

Appendix D. Smooth Regularizer

Theorem 21 (Restatement of Theorem 2). *Let $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$ be an ε-approximate trajectory above a convex set such that R is β-smooth, and let D be the diameter of the domain. Assume $\varepsilon \le D^2/2$, then for any $w \in \mathcal{K}$:*

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

Proof. From the strong convexity of R we have that $\beta \ge 1$, which means that from the assumptions $\varepsilon \le D^2\beta/2$. Then, from Lemma 15, for every t:

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

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 20:

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

■

Theorem 22 (Restatement of Theorem 3). *For every β, ε, there is an OMD ε-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 β-smooth and constant losses ($\ell_t = \ell$ for some ℓ 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 ε-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 ϕ_t) by differentiating and comparing to 0:

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

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 ℓ:

$$\ell \leq \frac{\sqrt{2\beta\varepsilon}}{\eta}$$
$$\iff \quad \frac{\eta^2\ell^2}{2\beta} \leq \varepsilon$$

Which means that w_{t-1} is an ε-minimizer. ■

Appendix E. Balance

All the lemmas in this section assumes ν-barrier regularizer.

E.1. General

Definition 23. Assume $\mathcal{K}$ is a polytope defined in standard form $\left\{w \in \mathbb{R}^d : Aw = b \wedge (w_i \geq 0, \forall i \in [d])\right\}$. 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 24.

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

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 25. *Assume $\mathcal{K}$ is a polytope. For some differentiable function $f : \mathcal{K} \to \mathbb{R}$, 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 α 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:

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

Which means that:

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

Which is our desired results. ∎

Lemma 26. *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 α such that both $w_{t'} + \alpha v$ and $w_{t'} - \alpha v$ is in the polytope . From Lemma 25 ($w^i_{t'} > 0$ 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. ∎

Lemma 27. *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 ε-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 $v^i \neq 0$, $\hat{w}^i_t \geq \psi$. We also assume that $\varepsilon \leq \psi/2$. we have:

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

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:

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

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 15, since $\hat{w}_t$ is an ε-minimizer of ϕ_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 $\psi > 0$:

$$\begin{aligned} -\sqrt{\frac{2c_2\varepsilon}{\psi^\nu}} &\leq \langle \eta \ell_t,\, v \rangle - B^v_{\hat{\gamma}}(t-1, t) \\ &= B^v_{\gamma}(t-1, t) - B^v_{\hat{\gamma}}(t-1, t) && \text{(Lemma 26)} \end{aligned}$$

Adding the induction assumption:

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

The last is from Lemma 24. ∎

E.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 28. *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 then $1/d$ 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 29. *Let $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$ be a k-balanced OMD trajectory with $w_1 = (1/d, 1/d, \ldots 1/d)$ 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 28, 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 26:

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

Denote:

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

From Lemma 28:

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

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:

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

Subtracting from both sides:

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

If $v^{\bar{i}} \leq 1/2d$ we have $w_t^{\bar{i}} \geq 1/2d$. 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 $v^{\bar{i}} \geq 1/2d$:

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

■

E.3. Simplex

Definition 30. 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 31. *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 32. *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 ε-approximate OMD trajectory.*

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

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

Proof. It is direct consequence of Lemma 27 for the case of $v_i = e_{i^*} - e_i$ (e_j is the jth element of the standard basis). ■

Appendix F. Lower bounds for negative entropy

Lemma 33. *Let $\gamma = (\{w_t\}_{t=1}^T, \{\ell_t\}_{t=1}^T, R, \eta)$ be an ε-approximate trajectory with $d = 2$ and ν-barrier regularizer. If for some coordinate i there is $\tau \in [T]$ such that $\frac{4\eta}{c_1}(w_\tau^i)^\nu \le \varepsilon$, then for any possible losses for $t \ge \tau$, 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 17 with $\varepsilon = 0$ we get:

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

Thus:

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

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

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

Which means that w_t is an ε-minimizer, as needed. ■

Lemma 34. *Assume for some $\alpha \ge T/2$, $\frac{1}{\eta}\log\left(\frac{4\eta}{\varepsilon}\right) \le \alpha$ with negative entropy regularizer, there is an instance above the simplex with α-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:

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

From Lemma 33, it is a valid error trajectory if for every $t \ge \tau$, $w_t^1 \le \frac{\varepsilon}{4\eta} \le \frac{1}{2}$. Thus, the regret for those steps is $\Omega(T - \tau)$. Adding the first τ steps we get a regret bound of $\Omega(T - 2\tau) \ge \Omega(T - 2\alpha)$. ■

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 ν-barrier regularizer. There is a set of constant losses for which there is an ε-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 ε-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 dth 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 ε-approximate minimizer for ϕ_t. Denote:

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

From Lemma 17 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 ε-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 33, it is a valid error trajectory if for every $3\tau \geq t \geq \tau$, $w_t = w_\tau$.

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

On the steps between 4τ and 5τ 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 33, 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 G. Proof of Theorems 6 and 7

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 ε-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 18:

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

From Lemma 17:

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

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'':

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

■

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 ε-approximate OMD trajectory such that there is $\xi > 0$ such that for every $t \in [T]$, $\hat{w}_t^{i^*} \geq \xi$. If $\varepsilon \leq \frac{\eta^4}{r''\left(\min\left\{\frac{\eta}{d}, \xi\right\}\right)}$, 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 $\xi' = \min\left\{\frac{\eta}{d}, \xi\right\}$, and let $S_t = \left\{i \neq i^* \;:\; \hat{w}_t^i \geq \xi'\right\}$ 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 then ξ', the effective smoothness is upper bounded by $r''(\xi')$ (since $r''(w) = 1/w^\nu$ for all $w \in (0,1]$). To use Lemma 15, 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 15 on ϕ_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:

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

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:

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

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 19:

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

Fix some coordinate $i\neq i^*$, and let $(s_1,t_1),(s_2,t_2),\ldots(s_n,t_n)$ be all enter and exit times for i to S_t. Namely, for every $j\in[n]$ and $s_j\le t\le t_j$, $i\in S_t$, 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 $\hat{w}_{t_j}^i\ge\xi'>\hat{w}_{s_j-1}^i$. Thus, by Lemma 37 we get that $D_r(0,\hat{w}_{s_j-1})\le D_r(0,\hat{w}_{t_j})$ which we apply on the RHS of the previous display to obtain:

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

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 17:

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

Thus:

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

Thus:

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

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

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

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

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

From Lemma 20:

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

∎

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 $T \ge 4$, γ is k-balanced and $\varepsilon \le \frac{1}{r''\left(\frac{1}{2de^{k+1}}\right)T^2}$.*

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

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 $\hat{w}_t^{i^*} \ge \hat{w}_s^{i^*}$ for some $s < t$ then the claim follows from the inductive assumption. If for every $i \ne i^*$, $\hat{w}_t^i \le \frac{1}{d}$ we have that $\hat{w}_t^{i^*} \ge \frac{1}{d}$ 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 $T \ge 3$ we have from the Lemma's assumptions and the induction assumptions that $\varepsilon \le \frac{1}{18r''(1/de^{k+1})} \le \frac{1}{16r''(\hat{w}_{t-1}^i)}$. From Lemma 18:

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

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

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 $\hat{w}_s^i \le \frac{1}{de^{k+1}}$. Again, since $T \ge 3$ we have $\varepsilon \le 1/16r''(\hat{w}_{s+1}^i)$, thus from Lemma 18:

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

Which means that for every $s' \in [s,t]$, $\hat{w}^i_{s'} \geq \frac{1}{2de^{k+1}}$. From Lemma 32 and our assumption on ε (see Definition 30 for the definition of B^i):

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

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

Recall that $\hat{w}^{i^*}_t < \hat{w}^{i^*}_s$, from Lemma 31:

$$\hat{w}^i_t \leq e^{B^i_{\hat{\gamma}}(s,t)}\hat{w}^i_s \leq e^{k+1}\hat{w}^i_s \leq e^{k+1}\frac{1}{de^{k+1}} = \frac{1}{d}$$

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

From Lemma 27:

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

Now, from Lemma 31 (recall that $\hat{w}^i_t > \frac{1}{d} = \hat{w}^i_1$):

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

which completes the inductive step and the proof. ∎

Lemma 41. *Let $\mathcal{K} = \Delta_d$ be the simplex, let $\{\ell_t\}_{t=1}^T$ be an α-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*

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

then the regret of any ε-approximate OMD trajectory is bounded as

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

Proof. From Lemma 26 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 $t \in [T]$, $\hat{w}^{i^*}_t \geq \frac{1}{de^{\alpha\eta+1}}$.

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. ∎

Implications for the main theorems.

Proof of Theorem 6: 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 41 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 $\eta\alpha \le \log\left(dT^2/\delta\right)$. Plugging this into Lemma 41 together with the fact that $D_R(w, w_1) \le \log(d)$ for all w yields the stochastic upper bound. ■

Appendix H. Proof of Theorem 5

Lemma 42. *For every ν-barrier regularizer rand $1 \ge w_2 \ge w_1 \ge 0$ we have:*

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

Proof.

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

■

Lemma 43. *Denote $\psi = \left(\frac{c_1}{8\eta T d + c_1 (2d)^{\nu-1}}\right)^{1/(\nu-1)}$. In the assumptions of Theorem 5, for every t, i:*

$$w_t^i \ge \psi$$

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

Notice that $\eta \le \frac{1}{16 c_1}$ and $\varepsilon \le \eta \psi^\nu \le \frac{(w_{t-1}^i)^\nu}{16 c_1}$, which means that from Lemma 18, we know for start that $w_t^i \ge \psi/2$. This means that for every $t' \in [1, t]$, $w_t^i \ge \psi/2$. One can see that $\varepsilon \le c_2 \psi/4$ which means that we can use Lemma 27 with $\psi/2$. Since the balance of an exact trajectory is always bounded by ηT, for every normalized $v \in \ker(A)$:

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

From Lemma 29 and the induction assumption, for every $i \in [d]$ and $t' \le t$:

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

If $w_t^i \leq 1/2d$ we can use Lemma 42:

$$\begin{aligned}\frac{c_1}{\left(w_i^{t'}\right)^{\nu-1}} &\leq 8\eta T d + c_1(2d)^{\nu-1}\\ \Rightarrow w_i^{t'} &\geq \left(\frac{c_1}{8\eta T d + c_1(2d)^{\nu-1}}\right)^{1/(\nu-1)} = \psi\end{aligned}$$

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

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

Which ends the induction step. ■

Proof of Theorem 5: Since the polytope is a subset of the simplex, the diameter is bounded by 2. From Lemma 43, the effective smoothness of the trajectory is bounded by $\beta \coloneqq c_2/\psi^\nu$. By the assumption about ε we have $\varepsilon \leq \eta^4/\beta$. From Lemma 15, for every t:

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

From here it is straightforward standard OMD arguments:

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

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

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

From Lemma 20:

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

■

Appendix I. Proof of Theorem 8

I.1. Polytope definition

The polytope is defined as $\left\{w \in \mathbb{R}^d : Aw = b \wedge (w_i \geq 0,\ \forall i \in [d])\right\}$ 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 $m \geq 128\log(2T)$ and $\varepsilon < 4\eta$ (obviously proof that works for small ε works for bigger).

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 $\left(\frac{1}{d}, \frac{1}{d}, \ldots, \frac{1}{d}\right)$ 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 44. $\{v_i\}_{i=1}^{m+1}$ *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 45. *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 44, 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. ■

I.2. General settings and hardness event

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

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 guassian with mean $\sqrt{\eta d} \le 1$ 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:

$$\sum_{i=4m+1}^{5m+1} \ell^i_{:\tau} \le 0$$
$$\sum_{i=4m+1}^{5m+1} \ell^i_t \le \frac{m}{16} \quad \forall t \in [T]$$

Lemma 46.

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

Proof. Denote $G = \sum_{i=4m+1}^{5m+1} \ell^i_{:\tau}$. Since G is a sum of gaussian random variables, it is also a gaussian random variable, denote its mean with μ and variance σ^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 $G \le 0$ is $\Theta(1)$.

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

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

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 44, it can be written as a linear combination of v. Denote the coefficients as α, namely:

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

I.3. Analysis

Lemma 47. *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 $w_t^{m+i} = w_t^{2m+i} = w_t^{m+i}$.

Notice that $\eta\langle \ell_{:\tau}, v_i \rangle = -3\eta\tau = -9$. From Lemma 26:

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

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 48. *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 47 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 26.

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 26:

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

From Lemma 26:

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

Since $w_\tau^i \leq \frac{1}{2d}$, we have $\frac{w_1^i}{w_\tau^i} \geq 2$. From Lemma 47, 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_τ^{5m+2} can only be altered with α_τ^{m+1}, this concludes the proof. ∎

Lemma 49. *For some $t \in [T]$, assume $\alpha_{t-1}^{m+1} \leq -\frac{1}{d} + e^{-m/8}$ and E. There is an ε-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 26 (as $\langle \ell_t, v_i \rangle$ has a constant value of -1).

From the second part of E and Lemma 26:

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

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 26:

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

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

By now we showed that if the tth 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 α_t^i for all $i \in [m]$ but change α_t^{m+1} to be equal to α_{t-1}^{m+1}, this will be an ε-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 50. *There is an ε-approximate trajectory that get a regret of:*

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

Proof. From Lemmas 48 and 49 we get that there is an ε-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. ■