Scale-Free Algorithms for Online Linear Optimization

Francesco Orabona and Dávid Pál

Yahoo Labs, 11th Floor, 229 West 43rd Street, New York, NY 10036, USA
francesco@orabona.com and dpal@yahoo-inc.com

Abstract. We design algorithms for online linear optimization that have optimal regret and at the same time do not need to know any upper or lower bounds on the norm of the loss vectors. We achieve adaptiveness to norms of loss vectors by scale invariance, i.e., our algorithms make exactly the same decisions if the sequence of loss vectors is multiplied by any positive constant. Our algorithms work for any decision set, bounded or unbounded. For unbounded decisions sets, these are the first truly adaptive algorithms for online linear optimization.

1 Introduction

Online Linear Optimization (OLO) is a problem where an algorithm repeatedly chooses a point w_t from a convex decision set K, observes an arbitrary, or even adversarially chosen, loss vector ℓ_t and suffers loss $\langle \ell_t, w_t \rangle$. The goal of the algorithm is to have a small cumulative loss. Performance of an algorithm is evaluated by the so-called regret, which is the difference of cumulative losses of the algorithm and of the (hypothetical) strategy that would choose in every round the same best point in hindsight.

OLO is a fundamental problem in machine learning [3, 18]. Many learning problems can be directly phrased as OLO, e.g., learning with expert advice [9, 21, 2], online combinatorial optimization [7]. Other problems can be reduced to OLO, e.g. online convex optimization [18, Chapter 2], online classification and regression [3, Chapters 11 and 12], multi-armed problems [3, Chapter 6], and batch and stochastic optimization of convex functions [12]. Hence, a result in OLO immediately implies other results in all these domains.

The adversarial choice of the loss vectors received by the algorithm is what makes the OLO problem challenging. In particular, if an OLO algorithm commits to an upper bound on the norm of future loss vectors, its regret can be made arbitrarily large through an adversarial strategy that produces loss vectors with norms that exceed the upper bound.

For this reason, most of the existing OLO algorithms receive as an input—or explicitly assume—an upper bound B on the norm of the loss vectors. The input B is often disguised as the learning rate, the regularization parameter, or the parameter of strong convexity of the regularizer. Examples of such algorithms include the Hedge algorithm or online projected gradient descent with fixed learning rate. However, these algorithms have two obvious drawbacks.

Algorithm	**Decisions Set(s)**	**Regularizer(s)**	**Scale-Free**
Hedge [5]	Probability Simplex	Negative Entropy	No
GIGA [23]	Any Bounded	$\frac{1}{2}\|w\|_2^2$	No
RDA [22]	**Any**	**Any Strongly Convex**	No
FTRL-Proximal [11, 10]	Any Bounded	$\frac{1}{2}\|w\|_2^2+$ any convex func.	**Yes**
AdaGrad MD [4]	Any Bounded	$\frac{1}{2}\|w\|_2^2+$ any convex func.	**Yes**
AdaGrad FTRL [4]	**Any**	$\frac{1}{2}\|w\|_2^2+$ any convex func.	No
AdaHedge [15]	Probability Simplex	Negative Entropy	**Yes**
Optimistic MD [14]	$\sup_{u,v\in K} \mathcal{B}_f(u,v) < \infty$	**Any Strongly Convex**	**Yes**
NAG [16]	$\{u : \max_t \langle \ell_t, u\rangle \leq C\}$	$\frac{1}{2}\|w\|_2^2$	Partially[1]
Scale invariant algorithms [13]	**Any**	$\frac{1}{2}\|w\|_p^2+$ any convex func. $1 < p \leq 2$	Partially[1]
AdaFTRL **[this paper]**	Any Bounded	**Any Strongly Convex**	**Yes**
SOLO FTRL **[this paper]**	**Any**	**Any Strongly Convex**	**Yes**

Table 1. Selected results for OLO. Best results in each column are in bold.

First, they do not come with any regret guarantee for sequences of loss vectors with norms exceeding B. Second, on sequences where the norm of loss vectors is bounded by $b \ll B$, these algorithms fail to have an optimal regret guarantee that depends on b rather than on B.

There is a clear practical need to design algorithms that adapt automatically to norms of the loss vectors. A natural, yet overlooked, design method to achieve this type of adaptivity is by insisting to have a **scale-free** algorithm. That is, the sequence of decisions of the algorithm does not change if the sequence of loss vectors is multiplied by a positive constant.

A summary of algorithms for OLO is presented in Table 1. While the scale-free property has been looked at in the expert setting, in the general OLO setting this issue has been largely ignored. In particular, the AdaHedge [15] algorithm, for prediction with expert advice, is specifically designed to be scale-free. A notable exception in the OLO literature is the discussion of the "off-by-one" issue in [10], where it is explained that even the popular AdaGrad algorithm [4] is not completely adaptive; see also our discussion in Section 4. In particular, existing scale-free algorithms cover only some norms/regularizers and *only* bounded decision sets. The case of **unbounded decision sets**, practically the most interesting one for machine learning applications, remains completely unsolved.

Rather than trying to design strategies for a particular form of loss vectors and/or decision sets, in this paper we explicitly focus on the scale-free property. Regret of scale-free algorithms is proportional to the scale of the losses, ensuring optimal linear dependency on the maximum norm of the loss vectors.

The contribution of this paper is twofold. First, in Section 3 we show that the analysis and design of AdaHedge can be generalized to the OLO scenario

[1] These algorithms attempt to produce an invariant sequence of predictions $\langle w_t, \ell_t\rangle$, rather than a sequence of invariant w_t.

and to any strongly convex regularizer, in an algorithm we call AdaFTRL, providing a new and rather interesting way to adapt the learning rates to have scale-free algorithms. Second, in Section 4 we propose a new and simple algorithm, SOLO FTRL, that is scale-free and is the **first** scale-free online algorithm for unbounded sets with a non-vacuous regret bound. Both algorithms are instances of Follow The Regularized Leader (FTRL) with an adaptive learning rate. Moreover, our algorithms show that scale-free algorithms can be obtained in a "native" and simple way, i.e. without using "doubling tricks" that attempt to fix poorly designed algorithms rather than directly solving the problem.

For both algorithms, we prove that for bounded decision sets the regret after T rounds is at most $O(\sqrt{\sum_{t=1}^T \|\ell_t\|_*^2})$. We show that the $\sqrt{\sum_{t=1}^T \|\ell_t\|_*^2}$ term is necessary by proving a $\Omega(D\sqrt{\sum_{t=1}^T \|\ell_t\|_*^2})$ lower bound on the regret of any algorithm for OLO for any decision set with diameter D with respect to the primal norm $\|\cdot\|$. For the SOLO FTRL algorithm, we prove an $O(\max_{t=1,2,\ldots,T} \|\ell_t\|_* \sqrt{T})$ regret bound for any unbounded decision set.

Our algorithms are also **any-time**, i.e., do not need to know the number of rounds in advance and our regret bounds hold for all time steps simultaneously.

2 Notation and Preliminaries

Let V be a finite-dimensional real vector space equipped with a norm $\|\cdot\|$. We denote by V^* its dual vector space. The bi-linear map associated with (V^*, V) is denoted by $\langle \cdot, \cdot \rangle : V^* \times V \to \mathbb{R}$. The dual norm of $\|\cdot\|$ is $\|\cdot\|_*$.

In OLO, in each round $t = 1, 2, \ldots,$ the algorithm chooses a point w_t in the decision set $K \subseteq V$ and then the algorithm observes a loss vector $\ell_t \in V^*$. The instantaneous loss of the algorithm in round t is $\langle \ell_t, w_t \rangle$. The cumulative loss of the algorithm after T rounds is $\sum_{t=1}^T \langle \ell_t, w_t \rangle$. The regret of the algorithm with respect to a point $u \in K$ is

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

and the regret with respect to the best point is $\text{Regret}_T = \sup_{u \in K} \text{Regret}_T(u)$. We assume that K is a non-empty closed convex subset of V. Sometimes we will assume that K is also bounded. We denote by D its diameter with respect to $\|\cdot\|$, i.e. $D = \sup_{u,v \in K} \|u - v\|$. If K is unbounded, $D = +\infty$.

Convex Analysis. The *Bregman divergence* of a convex differentiable function f is defined as $\mathcal{B}_f(u, v) = f(u) - f(v) - \langle \nabla f(v), u - v \rangle$. Note that $\mathcal{B}_f(u, v) \geq 0$ for any u, v which follows directly from the definition of convexity of f.

The *Fenchel conjugate* of a function $f : K \to \mathbb{R}$ is the function $f^* : V^* \to \mathbb{R} \cup \{+\infty\}$ defined as $f^*(\ell) = \sup_{w \in K} (\langle \ell, w \rangle - f(w))$. The Fenchel conjugate of any function is convex (since it is a supremum of affine functions) and satisfies for all $w \in K$ and all $\ell \in V^*$ the *Fenchel-Young inequality* $f(w) + f^*(\ell) \geq \langle \ell, w \rangle$.

Algorithm 1 FTRL WITH VARYING REGULARIZER

Require: Sequence of regularizers $\{R_t\}_{t=1}^{\infty}$
1: Initialize $L_0 \leftarrow 0$
2: **for** $t = 1, 2, 3, \ldots$ **do**
3: $\quad w_t \leftarrow \operatorname{argmin}_{w \in K} \left(\langle L_{t-1}, w \rangle + R_t(w) \right)$
4: $\quad$ Predict w_t
5: $\quad$ Observe $\ell_t \in V^*$
6: $\quad L_t \leftarrow L_{t-1} + \ell_t$
7: **end for**

Monotonicity of Fenchel conjugates follows easily from the definition: If $f, g : K \to \mathbb{R}$ satisfy $f(w) \leq g(w)$ for all $w \in K$ then $f^*(\ell) \geq g^*(\ell)$ for every $\ell \in V^*$.

Given $\lambda > 0$, a function $f : K \to \mathbb{R}$ is called λ-*strongly convex* with respect to a norm $\| \cdot \|$ if and only if, for all $x, y \in K$,

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\lambda}{2} \|x - y\|^2 ,$$

where $\nabla f(x)$ is any subgradient of f at point x.

The following proposition relates the range of values of a strongly convex function to the diameter of its domain. The proof can be found in Appendix A.

Proposition 1 (Diameter vs. Range). *Let $K \subseteq V$ be a non-empty bounded closed convex subset. Let $D = \sup_{u,v \in K} \|u - v\|$ be its diameter with respect to $\| \cdot \|$. Let $f : K \to \mathbb{R}$ be a non-negative lower semi-continuous function that is 1-strongly convex with respect to $\| \cdot \|$. Then, $D \leq \sqrt{8 \sup_{v \in K} f(v)}$.*

Fenchel conjugates and strongly convex functions have certain nice properties, which we list in Proposition 2 below.

Proposition 2 (Fenchel Conjugates of Strongly Convex Functions). *Let $K \subseteq V$ be a non-empty closed convex set with diameter $D := \sup_{u,v \in K} \|u - v\|$. Let $\lambda > 0$, and let $f : K \to \mathbb{R}$ be a lower semi-continuous function that is λ-strongly convex with respect to $\| \cdot \|$. The Fenchel conjugate of f satisfies:*

1. *f^* is finite everywhere and differentiable.*
2. *$\nabla f^*(\ell) = \operatorname{argmin}_{w \in K} \left(f(w) - \langle \ell, w \rangle \right)$*
3. *For any $\ell \in V^*$, $f^*(\ell) + f(\nabla f^*(\ell)) = \langle \ell, \nabla f^*(\ell) \rangle$.*
4. *f^* is $\frac{1}{\lambda}$-strongly smooth i.e. for any $x, y \in V^*$, $\mathcal{B}_{f^*}(x, y) \leq \frac{1}{2\lambda} \|x - y\|_*^2$.*
5. *f^* has $\frac{1}{\lambda}$-Lipschitz continuous gradients i.e. $\|\nabla f^*(x) - \nabla f^*(y)\| \leq \frac{1}{\lambda} \|x - y\|_*$ for any $x, y \in V^*$.*
6. *$\mathcal{B}_{f^*}(x, y) \leq D \|x - y\|_*$ for any $x, y \in V^*$.*
7. *$\|\nabla f^*(x) - \nabla f^*(y)\| \leq D$ for any $x, y \in V^*$.*
8. *For any $c > 0$, $(c f(\cdot))^* = c f^*(\cdot / c)$.*

Except for properties 6 and 7, the proofs can be found in [17]. Property 6 is proven in Appendix A. Property 7 trivially follows from property 2.

Generic FTRL with Varying Regularizer. Our scale-free online learning algorithms are versions of the Follow The Regularized Leader (FTRL) algorithm with varying regularizers, presented as Algorithm 1. The following lemma bounds its regret.

Lemma 1 (Lemma 1 in [13]). *For any sequence $\{R_t\}_{t=1}^{\infty}$ of strongly convex lower semi-continuous regularizers, regret of Algorithm 1 is upper bounded as*

$$\text{Regret}_T(u) \leq R_{T+1}(u) + R_1^*(0) + \sum_{t=1}^{T} \mathcal{B}_{R_t^*}(-L_t, -L_{t-1}) - R_t^*(-L_t) + R_{t+1}^*(-L_t) \ .$$

The lemma allows data dependent regularizers. That is, R_t can depend on the past loss vectors $\ell_1, \ell_2, \ldots, \ell_{t-1}$.

3 AdaFTRL

In this section we generalize the AdaHedge algorithm [15] to the OLO setting, showing that it retains its scale-free property. The analysis is very general and based on general properties of strongly convex functions, rather than specific properties of the entropic regularizer like in AdaHedge.

Assume that K is bounded and that $R(w)$ is a strongly convex lower semi-continuous function bounded from above. We instantiate Algorithm 1 with the sequence of regularizers

$$R_t(w) = \Delta_{t-1} R(w) \quad \text{where} \quad \Delta_t = \sum_{i=1}^{t} \Delta_{i-1} \mathcal{B}_{R^*}\left(-\frac{L_i}{\Delta_{i-1}}, -\frac{L_{i-1}}{\Delta_{i-1}} \right) \ . \qquad (1)$$

The sequence $\{\Delta_t\}_{t=0}^{\infty}$ is non-negative and non-decreasing. Also, Δ_t as a function of $\{\ell_s\}_{s=1}^{t}$ is positive homogenous of degree one, making the algorithm scale-free.

If $\Delta_{i-1} = 0$, we define $\Delta_{i-1}\mathcal{B}_{R^*}(\frac{-L_i}{\Delta_{i-1}}, \frac{-L_{i-1}}{\Delta_{i-1}})$ as $\lim_{a \to 0^+} a\mathcal{B}_{R^*}(\frac{-L_i}{a}, \frac{-L_{i-1}}{a})$ which always exists and is finite; see Appendix B. Similarly, when $\Delta_{t-1} = 0$, we define $w_t = \operatorname{argmin}_{w \in K} \langle L_{t-1}, w \rangle$ where ties among minimizers are broken by taking the one with the smallest value of $R(w)$, which is unique due to strong convexity; this is the same as $w_t = \lim_{a \to 0^+} \operatorname{argmin}_{w \in K} (\langle L_{t-1}, w \rangle + aR(w))$.

Our main result is an $O(\sqrt{\sum_{t=1}^{T} \|\ell_t\|_*^2})$ upper bound on the regret of the algorithm after T rounds, without the need to know before hand an upper bound on $\|\ell_t\|_*$. We prove the theorem in Section 3.1.

Theorem 1 (Regret Bound). *Suppose $K \subseteq V$ is a non-empty bounded closed convex subset. Let $D = \sup_{x,y \in K} \|x - y\|$ be its diameter with respect to a norm $\|\cdot\|$. Suppose that the regularizer $R : K \to \mathbb{R}$ is a non-negative lower semi-continuous function that is λ-strongly convex with respect to $\|\cdot\|$ and is bounded from above. The regret of AdaFTRL satisfies*

$$\text{Regret}_T(u) \leq \sqrt{3} \max\left\{ D, \frac{1}{\sqrt{2\lambda}} \right\} \sqrt{\sum_{t=1}^{T} \|\ell_t\|_*^2 \,(1 + R(u))} \ .$$

The regret bound can be optimized by choosing the optimal multiple of the regularizer. Namely, we choose regularizer of the form $\lambda f(w)$ where $f(w)$ is 1-strongly convex and optimize over λ. The result of the optimization is the following corollary. Its proof can be found in Appendix C.

Corollary 1 (Regret Bound). *Suppose $K \subseteq V$ is a non-empty bounded closed convex subset. Suppose $f : K \to \mathbb{R}$ is a non-negative lower semi-continuous function that is 1-strongly convex with respect to $\|\cdot\|$ and is bounded from above. The regret of AdaFTRL with regularizer*

$$R(w) = \frac{f(w)}{16 \cdot \sup_{v \in K} f(v)} \qquad \textit{satisfies} \qquad \text{Regret}_T \le 5.3 \sqrt{\sup_{v \in K} f(v) \sum_{t=1}^{T} \|\ell_t\|_*^2} \, .$$

3.1 Proof of Regret Bound for AdaFTRL

Lemma 2 (Initial Regret Bound). *AdaFTRL, for any $u \in K$ and any $u \ge 0$, satisfies* $\text{Regret}_T(u) \le (1 + R(u))\, \Delta_T$.

Proof. Let $R_t(w) = \Delta_{t-1} R(w)$. Since R is non-negative, $\{R_t\}_{t=1}^{\infty}$ is non-decreasing. Hence, $R_t^*(\ell) \ge R_{t+1}^*(\ell)$ for every $\ell \in V^*$ and thus $R_t^*(-L_t) - R_{t+1}^*(-L_t) \ge 0$. So, by Lemma 1,

$$\text{Regret}_T(u) \le R_{T+1}(u) + R_1^*(0) + \sum_{t=1}^{T} \mathcal{B}_{R_t^*}(-L_t, -L_{t-1}) \, . \tag{2}$$

Since, $\mathcal{B}_{R_t^*}(u, v) = \Delta_{t-1} \mathcal{B}_{R^*}(\frac{u}{\Delta_{t-1}}, \frac{v}{\Delta_{t-1}})$ by definition of Bregman divergence and Part 8 of Proposition 2, we have $\sum_{t=1}^{T} \mathcal{B}_{R_t^*}(-L_t, -L_{t-1}) = \Delta_T$.

Lemma 3 (Recurrence). *Let $D = \sup_{u,v \in K} \|u - v\|$ be the diameter of K. The sequence $\{\Delta_t\}_{t=1}^{\infty}$ generated by AdaFTRL satisfies for any $t \ge 1$,*

$$\Delta_t \le \Delta_{t-1} + \min\left\{ D\|\ell_t\|_*, \ \frac{\|\ell_t\|_*^2}{2\lambda \Delta_{t-1}} \right\} \, .$$

Proof. The inequality results from strong convexity of $R_t(w)$ and Proposition 2.

Lemma 4 (Solution of the Recurrence). *Let D be the diameter of K. The sequence $\{\Delta_t\}_{t=0}^{\infty}$ generated by AdaFTRL satisfies for any $T \ge 0$,*

$$\Delta_T \le \sqrt{3} \max\left\{ D, \frac{1}{\sqrt{2\lambda}} \right\} \sqrt{\sum_{t=1}^{T} \|\ell_t\|_*^2} \, .$$

Proof of the Lemma 4 is deferred to Appendix C. Theorem 1 follows from Lemmas 2 and 4.

4 SOLO FTRL

The closest algorithm to a scale-free one in the OLO literature is the AdaGrad algorithm [4]. It uses a regularizer on each coordinate of the form

$$R_t(w) = R(w)\left(\delta + \sqrt{\sum_{s=1}^{t-1} \|\ell_s\|_*^2}\right).$$

This kind of regularizer would yield a scale-free algorithm *only* for $\delta = 0$. Unfortunately, the regret bound in [4] becomes vacuous for such setting in the unbounded case. In fact, it requires δ to be greater than $\|\ell_t\|_*$ for all time steps t, requiring knowledge of the future (see Theorem 5 in [4]). In other words, despite of its name, AdaGrad is not fully adaptive to the norm of the loss vectors. Identical considerations hold for the FTRL-Proximal in [11, 10]: the scale-free setting of the learning rate is valid only in the bounded case.

One simple approach would be to use a doubling trick on δ in order to estimate on the fly the maximum norm of the losses. Note that a naive strategy would still fail because the initial value of δ should be data-dependent in order to have a scale-free algorithm. Moreover, we would have to upper bound the regret in all the rounds where the norm of the current loss is bigger than the estimate. Finally, the algorithm would depend on an additional parameter, the "doubling" power. Hence, even guaranteeing a regret bound[2], such strategy would give the feeling that FTRL needs to be "fixed" in order to obtain a scale-free algorithm.

In the following, we propose a much simpler and better approach. We propose to use Algorithm 1 with the regularizer

$$R_t(w) = R(w)\sqrt{\sum_{s=1}^{t-1} \|\ell_s\|_*^2}\,,$$

where $R : K \to \mathbb{R}$ is any strongly convex function. Through a refined analysis, we show that the regularizer suffices to obtain an optimal regret bound for any decision set, bounded or unbounded. We call such variant Scale-free Online Linear Optimization FTRL algorithm (SOLO FTRL). Our main result is the following Theorem, which is proven in Section 4.1.

Theorem 2 (Regret of SOLO FTRL). *Suppose $K \subseteq V$ is a non-empty closed convex subset. Let $D = \sup_{u,v \in K} \|u - v\|$ be its diameter with respect to a norm $\|\cdot\|$. Suppose that the regularizer $R : K \to \mathbb{R}$ is a non-negative lower semi-continuous function that is λ-strongly convex with respect to $\|\cdot\|$. The regret of SOLO FTRL satisfies*

$$\operatorname{Regret}_T(u) \le \left(R(u) + \frac{2.75}{\lambda}\right)\sqrt{\sum_{t=1}^{T} \|\ell_t\|_*^2} + 3.5 \min\left\{\frac{\sqrt{T-1}}{\lambda}, D\right\} \max_{t \le T} \|\ell_t\|_*.$$

[2] For lack of space, we cannot include the regret bound for the doubling trick version. It would be exactly the same as in Theorem 2, following a similar analysis, but with the additional parameter of the doubling power.

When K is bounded, we can choose the optimal multiple of the regularizer. We choose $R(w) = \lambda f(w)$ where f is a 1-strongly convex function and optimize λ. The result of the optimization is Corollary 2; the proof is in Appendix D. It is similar to Corollary 1 for AdaFTRL. The scaling however is different in the two corollaries. In Corollary 1, $\lambda \sim 1/(\sup_{v \in K} f(v))$ while in Corollary 2 we have $\lambda \sim 1/\sqrt{\sup_{v \in K} f(v)}$.

Corollary 2 (Regret Bound for Bounded Decision Sets). *Suppose $K \subseteq V$ is a non-empty bounded closed convex subset. Suppose that $f : K \to \mathbb{R}$ is a non-negative lower semi-continuous function that is 1-strongly convex with respect to $\| \cdot \|$. SOLO FTRL with regularizer*

$$R(w) = \frac{f(w)\sqrt{2.75}}{\sqrt{\sup_{v \in K} f(v)}} \quad \textit{satisfies} \quad \text{Regret}_T \leq 13.3 \sqrt{\sup_{v \in K} f(v) \sum_{t=1}^{T} \|\ell_t\|_*^2} \,.$$

4.1 Proof of Regret Bound for SOLO FTRL

The proof of Theorem 2 relies on an inequality (Lemma 5). Related and weaker inequalities were proved by [1] and [6]. The main property of this inequality is that on the right-hand side C does *not* multiply the $\sqrt{\sum_{t=1}^{T} a_t^2}$ term. We will also use the well-known technical Lemma 6.

Lemma 5 (Useful Inequality). *Let $C, a_1, a_2, \ldots, a_T \geq 0$. Then,*

$$\sum_{t=1}^{T} \min \left\{ a_t^2 / \sqrt{\sum_{s=1}^{t-1} a_s^2}, \; C a_t \right\} \leq 3.5 C \max_{t=1,2,\ldots,T} a_t + 3.5 \sqrt{\sum_{t=1}^{T} a_t^2} \,.$$

Proof. Without loss of generality, we can assume that $a_t > 0$ for all t. Since otherwise we can remove all $a_t = 0$ without affecting either side of the inequality. Let $M_t = \max\{a_1, a_2, \ldots, a_t\}$ and $M_0 = 0$. We prove that for any $\alpha > 1$

$$\min \left\{ \frac{a_t^2}{\sqrt{\sum_{s=1}^{t-1} a_s^2}}, C a_t \right\} \leq 2\sqrt{1+\alpha^2} \left(\sqrt{\sum_{s=1}^{t} a_s^2} - \sqrt{\sum_{s=1}^{t-1} a_s^2} \right) + \frac{C\alpha(M_t - M_{t-1})}{\alpha - 1}$$

from which the inequality follows by summing over $t = 1, 2, \ldots, T$ and choosing $\alpha = \sqrt{2}$. The inequality follows by case analysis. If $a_t^2 \leq \alpha^2 \sum_{s=1}^{t-1} a_s^2$, we have

$$\min \left\{ \frac{a_t^2}{\sqrt{\sum_{s=1}^{t-1} a_s^2}}, C a_t \right\} \leq \frac{a_t^2}{\sqrt{\sum_{s=1}^{t-1} a_s^2}} = \frac{a_t^2}{\sqrt{\frac{1}{1+\alpha^2} \left(\alpha^2 \sum_{s=1}^{t-1} a_s^2 + \sum_{s=1}^{t-1} a_s^2 \right)}}$$

$$\leq \frac{a_t^2 \sqrt{1+\alpha^2}}{\sqrt{a_t^2 + \sum_{s=1}^{t-1} a_s^2}} = \frac{a_t^2 \sqrt{1+\alpha^2}}{\sqrt{\sum_{s=1}^{t} a_s^2}} \leq 2\sqrt{1+\alpha^2} \left(\sqrt{\sum_{s=1}^{t} a_s^2} - \sqrt{\sum_{s=1}^{t-1} a_s^2} \right)$$

where we have used $x^2/\sqrt{x^2+y^2} \le 2(\sqrt{x^2+y^2}-\sqrt{y^2})$ in the last step. On the other hand, if $a_t^2 > \alpha^2 \sum_{t=1}^{t-1} a_s^2$, we have

$$\begin{aligned}\min\left\{\frac{a_t^2}{\sqrt{\sum_{s=1}^{t-1} a_s^2}}, Ca_t\right\} &\le Ca_t = C\frac{\alpha a_t - a_t}{\alpha - 1} \le \frac{C}{\alpha-1}\left(\alpha a_t - \alpha\sqrt{\sum_{s=1}^{t-1} a_s^2}\right) \\ &= \frac{C\alpha}{\alpha-1}\left(a_t - \sqrt{\sum_{s=1}^{t-1} a_s^2}\right) \le \frac{C\alpha}{\alpha-1}\left(a_t - M_{t-1}\right) = \frac{C\alpha}{\alpha-1}\left(M_t - M_{t-1}\right)\end{aligned}$$

where we have used that $a_t = M_t$ and $\sqrt{\sum_{s=1}^{t-1} a_s^2} \ge M_{t-1}$.

Lemma 6 (Lemma 3.5 in [1]). *Let $a_1, a_2, \dots, a_T$ be non-negative real numbers. If $a_1 > 0$ then,*

$$\sum_{t=1}^{T} a_t/\sqrt{\sum_{s=1}^{t} a_s} \le 2\sqrt{\sum_{t=1}^{T} a_t} \ .$$

Proof (Proof of Theorem 2). Let $\eta_t = \frac{1}{\sqrt{\sum_{s=1}^{t-1} \|\ell_s\|_*^2}}$, hence $R_t(w) = \frac{1}{\eta_t} R(w)$. We assume without loss of generality that $\|\ell_t\|_* > 0$ for all t, since otherwise we can remove all rounds t where $\ell_t = 0$ without affecting regret and the predictions of the algorithm on the remaining rounds. By Lemma 1,

$$\text{Regret}_T(u) \le \frac{1}{\eta_{T+1}} R(u) + \sum_{t=1}^{T} \left(\mathcal{B}_{R_t^*}(-L_t, -L_{t-1}) - R_t^*(-L_t) + R_{t+1}^*(-L_t)\right) \ .$$

We upper bound the terms of the sum in two different ways. First, by Proposition 2, we have

$$\mathcal{B}_{R_t^*}(-L_t, -L_{t-1}) - R_t^*(-L_t) + R_{t+1}^*(-L_t) \le \mathcal{B}_{R_t^*}(-L_t, -L_{t-1}) \le \frac{\eta_t \|\ell_t\|_*^2}{2\lambda} \ .$$

Second, we have

$$\begin{aligned}&\mathcal{B}_{R_t^*}(-L_t, -L_{t-1}) - R_t^*(-L_t) + R_{t+1}^*(-L_t) \\ &= \mathcal{B}_{R_{t+1}^*}(-L_t, -L_{t-1}) + R_{t+1}^*(-L_{t-1}) - R_t^*(-L_{t-1}) \\ &\qquad + \langle \nabla R_t^*(-L_{t-1}) - \nabla R_{t+1}^*(-L_{t-1}), \ell_t \rangle \\ &\le \tfrac{1}{2\lambda}\eta_{t+1}\|\ell_t\|_*^2 + \|\nabla R_t^*(-L_{t-1}) - \nabla R_{t+1}^*(-L_{t-1})\| \cdot \|\ell_t\|_* \\ &= \tfrac{1}{2\lambda}\eta_{t+1}\|\ell_t\|_*^2 + \|\nabla R^*(-\eta_t L_{t-1}) - \nabla R^*(-\eta_{t+1} L_{t-1})\| \cdot \|\ell_t\|_* \\ &\le \frac{\eta_{t+1}\|\ell_t\|_*^2}{2\lambda} + \min\left\{\frac{1}{\lambda}\|L_{t-1}\|_* \left(\eta_t - \eta_{t+1}\right), D\right\} \|\ell_t\|_* \ ,\end{aligned}$$

where in the first inequality we have used the fact that $R_{t+1}^*(-L_{t-1}) \le R_t^*(-L_{t-1})$, Hölder's inequality, and Proposition 2. In the second inequality we have used properties 5 and 7 of Proposition 2. Using the definition of η_{t+1} we have

$$\frac{\|L_{t-1}\|_*(\eta_t - \eta_{t+1})}{\lambda} \leq \frac{\|L_{t-1}\|_*}{\lambda\sqrt{\sum_{i=1}^{t-1}\|\ell_i\|_*^2}} \leq \frac{\sum_{i=1}^{t-1}\|\ell_i\|_*}{\lambda\sqrt{\sum_{i=1}^{t-1}\|\ell_i\|_*^2}} \leq \frac{\sqrt{t-1}}{\lambda} \leq \frac{\sqrt{T-1}}{\lambda}.$$

Denoting by $H = \min\left\{\frac{\sqrt{T-1}}{\lambda}, D\right\}$ we have

$$\begin{aligned}
&\text{Regret}_T(u) \leq \frac{1}{\eta_{T+1}} R(u) + \sum_{t=1}^{T} \min\left\{\frac{\eta_t\|\ell_t\|_*^2}{2\lambda},\ H\|\ell_t\|_* + \frac{\eta_{t+1}\|\ell_t\|_*^2}{2\lambda}\right\} \\
&\leq \frac{1}{\eta_{T+1}} R(u) + \frac{1}{2\lambda}\sum_{t=1}^{T}\eta_{t+1}\|\ell_t\|_*^2 + \frac{1}{2\lambda}\sum_{t=1}^{T}\min\left\{\eta_t\|\ell_t\|_*^2,\ 2\lambda H\|\ell_t\|_*\right\} \\
&= \frac{1}{\eta_{T+1}} R(u) + \frac{1}{2\lambda}\sum_{t=1}^{T}\frac{\|\ell_t\|_*^2}{\sqrt{\sum_{s=1}^{t}\|\ell_t\|_*^2}} + \frac{1}{2\lambda}\sum_{t=1}^{T}\min\left\{\frac{\|\ell_t\|_*^2}{\sqrt{\sum_{s=1}^{t-1}\|\ell_s\|_*^2}},\ 2\lambda H\|\ell_t\|_*\right\}.
\end{aligned}$$

We bound each of the three terms separately. By definition of η_{T+1}, the first term is $\frac{1}{\eta_{T+1}} R(u) = R(u)\sqrt{\sum_{t=1}^{T}\|\ell_t\|_*^2}$. We upper bound the second term using Lemma 6 as

$$\frac{1}{2\lambda}\sum_{t=1}^{T}\frac{\|\ell_t\|_*^2}{\sqrt{\sum_{s=1}^{t}\|\ell_t\|_*^2}} \leq \frac{1}{\lambda}\sqrt{\sum_{t=1}^{T}\|\ell_t\|_*^2}\,.$$

Finally, by Lemma 5 we upper bound the third term as

$$\frac{1}{2\lambda}\sum_{t=1}^{T}\min\left\{\frac{\|\ell_t\|_*^2}{\sqrt{\sum_{s=1}^{t-1}\|\ell_s\|_*^2}},\ 2\lambda\|\ell_t\|_* H\right\} \leq 3.5H\max_{t\leq T}\|\ell_t\|_* + \frac{1.75}{\lambda}\sqrt{\sum_{t=1}^{T}\|\ell_t\|_*^2}\,.$$

Putting everything together gives the stated bound.

5 Lower Bound

We show a lower bound on the worst-case regret of any algorithm for OLO. The proof is a standard probabilistic argument, which we present in Appendix E.

Theorem 3 (Lower Bound). *Let $K \subseteq V$ be any non-empty bounded closed convex subset. Let $D = \sup_{u,v\in K}\|u-v\|$ be the diameter of K. Let A be any (possibly randomized) algorithm for OLO on K. Let T be any non-negative integer and let $a_1, a_2, \ldots, a_T$ be any non-negative real numbers. There exists a sequence of vectors $\ell_1, \ell_2, \ldots, \ell_T$ in the dual vector space V^* such that $\|\ell_1\|_* = a_1, \|\ell_2\|_* = a_2, \ldots, \|\ell_T\|_* = a_T$ and the regret of algorithm A satisfies*

$$\text{Regret}_T \geq \frac{D}{\sqrt{8}}\sqrt{\sum_{t=1}^{T}\|\ell_t\|_*^2}\,. \tag{3}$$

The upper bounds on the regret, which we have proved for our algorithms, have the same dependency on the norms of loss vectors. However, a gap remains between the lower bound and the upper bounds.

Our upper bounds are of the form $O(\sqrt{\sup_{v \in K} f(v) \sum_{t=1}^{T} \|\ell_t\|_*^2})$ where f is any 1-strongly convex function with respect to $\|\cdot\|$. The same upper bound is also achieved by FTRL with a constant learning rate when the number of rounds T and $\sum_{t=1}^{T} \|\ell_t\|_*^2$ is known upfront [18, Chapter 2]. The lower bound is $\Omega(D\sqrt{\sum_{t=1}^{T} \|\ell_t\|_*^2})$.

The gap between D and $\sqrt{\sup_{v \in K} f(v)}$ can be substantial. For example, if K is the probability simplex in $\mathbb{R}^d$ and $f(w) = \ln(d) + \sum_{i=1}^{d} w_i \ln w_i$ is the shifted negative entropy, the $\|\cdot\|_1$-diameter of K is 2, f is non-negative and 1-strongly convex w.r.t. $\|\cdot\|_1$, but $\sup_{v \in K} f(v) = \ln(d)$. On the other hand, if the norm $\|\cdot\|_2 = \sqrt{\langle \cdot, \cdot \rangle}$ arises from an inner product $\langle \cdot, \cdot \rangle$, the lower bound matches the upper bounds within a constant factor. The reason is that for any K with $\|\cdot\|_2$-diameter D, the function $f(w) = \frac{1}{2}\|w - w_0\|_2^2$, where w_0 is an arbitrary point in K, is 1-strongly convex w.r.t. $\|\cdot\|_2$ and satisfies that $\sqrt{\sup_{v \in K} f(v)} \leq D$. This leads to the following open problem (posed also in [8]):

> *Given a bounded convex set K and a norm $\|\cdot\|$, construct a non-negative function $f : K \to \mathbb{R}$ that is 1-strongly convex with respect to $\|\cdot\|$ and minimizes $\sup_{v \in K} f(v)$.*

As shown in [19], the existence of f with small $\sup_{v \in K} f(v)$ is equivalent to the existence of an algorithm for OLO with $\widetilde{O}(\sqrt{T \sup_{v \in K} f(v)})$ regret assuming $\|\ell_t\|_* \leq 1$. The $\widetilde{O}$ notation hides a polylogarithmic factor in T.

6 Per-Coordinate Learning

An interesting class of algorithms proposed in [11] and [4] are based on the so-called per-coordinate learning rates. As shown in [20], our algorithms, or in fact any algorithm for OLO, can be used with per-coordinate learning rates as well.

Abstractly, we assume that the decision set is a Cartesian product $K = K_1 \times K_2 \times \cdots \times K_d$ of a finite number of convex sets. On each factor K_i, $i = 1, 2, \ldots, d$, we can run any OLO algorithm separately and we denote by $\text{Regret}_T^{(i)}(u_i)$ its regret with respect to $u_i \in K_i$. The overall regret with respect to any $u = (u_1, u_2, \ldots, u_d) \in K$ can be written as

$$\text{Regret}_T(u) = \sum_{i=1}^{d} \text{Regret}_T^{(i)}(u_i) \ .$$

If the algorithm for each factor is scale-free, the overall algorithm is clearly scale-free as well. Using AdaFTRL or SOLO FTRL for each factor K_i, we generalize and improve existing regret bounds [11, 4] for algorithms with per-coordinate learning rates.

References

1. Auer, P., Cesa-Bianchi, N., Gentile, C.: Adaptive and self-confident on-line learning algorithms. Journal of Computer and System Sciences 64(1), 48–75 (2002)
2. Cesa-Bianchi, N., Freund, Y., Haussler, D., Helmbold, D.P., Schapire, R.E., Warmuth, M.K.: How to use expert advice. J. ACM 44(3), 427–485 (1997)
3. Cesa-Bianchi, N., Lugosi, G.: Prediction, Learning, and Games. Cambridge University Press Cambridge (2006)
4. Duchi, J., Hazan, E., Singer, Y.: Adaptive subgradient methods for online learning and stochastic optimization. J. Mach. Learn. Res. 12, 2121–2159 (2011)
5. Freund, Y., Schapire, R.E.: Large margin classification using the perceptron algorithm. Mach. Learn. 37(3), 277–296 (1999)
6. Jaksch, T., Ortner, R., Auer, P.: Near-optimal regret bounds for reinforcement learning. J. Mach. Learn. Res. 11, 1563–1600 (2010)
7. Koolen, W.M., Warmuth, M.K., Kivinen, J.: Hedging structured concepts. In: Proc. of COLT. pp. 93–105 (2010)
8. Kwon, J., Mertikopoulos, P.: A continuous-time approach to online optimization (February 2014), arXiv:1401.6956
9. Littlestone, N., Warmuth, M.K.: The weighted majority algorithm. Information and Computation 108(2), 212–261 (1994)
10. McMahan, H.B.: Analysis techniques for adaptive online learning (2014), arXiv:1403.3465
11. McMahan, H.B., Streeter, J.M.: Adaptive bound optimization for online convex optimization. In: Proc. of COLT. pp. 244–256 (2010)
12. Nemirovski, A., Yudin, D.B.: Problem complexity and method efficiency in optimization. Wiley (1983)
13. Orabona, F., Crammer, K., Cesa-Bianchi, N.: A generalized online mirror descent with applications to classification and regression. Mach. Learn. 99, 411–435 (2014)
14. Rakhlin, A., Sridharan, K.: Optimization, learning, and games with predictable sequences. In: Advances in Neural Information Processing Systems 26 (2013)
15. de Rooij, S., van Erven, T., Grünwald, P.D., Koolen, W.M.: Follow the leader if you can, hedge if you must. J. Mach. Learn. Res. 15, 1281–1316 (2014)
16. Ross, S., Mineiro, P., Langford, J.: Normalized online learning. In: Proc. of UAI (2013)
17. Shalev-Shwartz, S.: Online Learning: Theory, Algorithms, and Applications. Ph.D. thesis, Hebrew University, Jerusalem (2007)
18. Shalev-Shwartz, S.: Online learning and online convex optimization. Foundations and Trends in Machine Learning 4(2), 107–194 (2011)
19. Srebro, N., Sridharan, K., Tewari, A.: On the universality of online mirror descent. In: Advances in Neural Information Processing Systems (2011)
20. Streeter, M., McMahan, H.B.: Less regret via online conditioning (2010), arXiv:1002.4862
21. Vovk, V.: A game of prediction with expert advice. Journal of Computer and System Sciences 56, 153–173 (1998)
22. Xiao, L.: Dual averaging methods for regularized stochastic learning and online optimization. J. Mach. Learn. Res. 11, 2543–2596 (December 2010)
23. Zinkevich, M.: Online convex programming and generalized infinitesimal gradient ascent. In: Proc. of ICML. pp. 928–936 (2003)

A Proofs for Preliminaries

Proof (Proof of Proposition 1). Let $S = \sup_{u \in K} f(u)$ and $v^* = \operatorname{argmin}_{v \in K} f(v)$. The minimizer v^* is guaranteed to exist by lower semi-continuity of f and compactness of K. Optimality condition for v^* and 1-strong convexity of f imply that for any $u \in K$,

$$S \geq f(u) - f(v^*) \geq f(u) - f(v^*) - \langle \nabla f(v^*), u - v^* \rangle \geq \frac{1}{2}\|u - v^*\|^2 \ .$$

In other words, $\|u - v^*\| \leq \sqrt{2S}$. By triangle inequality,

$$D = \sup_{u,v \in K} \|u - v\| \leq \sup_{u,v \in K} (\|u - v^*\| + \|v^* - v\|) \leq 2\sqrt{2S} = \sqrt{8S} \ .$$

Proof (Proof of Property 6 of Proposition 2). To bound $\mathcal{B}_{f^*}(x, y)$ we add a non-negative divergence term $\mathcal{B}_{f^*}(y, x)$.

$$\begin{aligned} \mathcal{B}_{f^*}(x,y) &\leq \mathcal{B}_{f^*}(x,y) + \mathcal{B}_{f^*}(y,x) = \langle x - y, \nabla f^*(x) - \nabla f^*(y) \rangle \\ &\leq \|x - y\|_* \cdot \|\nabla f^*(x) - \nabla f^*(y)\| \leq D\|x - y\|_* \ , \end{aligned}$$

where we have used Hölder's inequality and Part 7 of the Proposition.

B Limits

Lemma 7. *Let K be a non-empty bounded closed convex subset of a finite dimensional normed real vector space $(V, \|\cdot\|)$. Let $R : K \to \mathbb{R}$ be a strongly convex lower semi-continuous function bounded from above. Then, for any $x, y \in V^*$,*

$$\lim_{a \to 0^+} a\mathcal{B}_{R^*}(x/a, y/a) = \langle x, u - v \rangle$$

where

$$u = \lim_{a \to 0^+} \operatorname*{argmin}_{w \in K} (aR(w) - \langle x, w \rangle) \quad \textit{and} \quad v = \lim_{a \to 0^+} \operatorname*{argmin}_{w \in K} (aR(w) - \langle y, w \rangle) \ .$$

Proof. Using Part 3 of Proposition 2 we can write the divergence

$$\begin{aligned} a\mathcal{B}_{R^*}(x/a, y/a) &= aR^*(x/a) - aR^*(y/a) - \langle x - y, \nabla R^*(y/a) \rangle \\ &= a\left[\langle x/a, \nabla R^*(x/a) \rangle - R(\nabla R^*(x/a))\right] \\ &\quad - a\left[\langle y/a, \nabla R^*(y/a) \rangle - R(\nabla R^*(y/a))\right] - \langle x - y, \nabla R^*(y/a) \rangle \\ &= \langle x, \nabla R^*(x/a) - \nabla R^*(y/a) \rangle - aR(\nabla R^*(x/a)) + aR(\nabla R^*(y/a)) \ . \end{aligned}$$

Part 2 of Proposition 2 implies that

$$\begin{aligned} u &= \lim_{a \to 0^+} \nabla R^*(x/a) = \lim_{a \to 0^+} \operatorname*{argmin}_{w \in K} (aR(w) - \langle x, w \rangle) \ , \\ v &= \lim_{a \to 0^+} \nabla R^*(y/a) = \lim_{a \to 0^+} \operatorname*{argmin}_{w \in K} (aR(w) - \langle y, w \rangle) \ . \end{aligned}$$

The limits on the right exist because of compactness of K. They are simply the minimizers $u = \operatorname{argmin}_{w \in K} -\langle x, w \rangle$ and $v = \operatorname{argmin}_{w \in K} -\langle y, w \rangle$ where ties in argmin are broken according to smaller value of $R(w)$.

By assumption $R(w)$ is upper bounded. It is also lower bounded, since it is defined on a compact set and it is lower semi-continuous. Thus,

$$
\begin{aligned}
&\lim_{a \to 0^+} a \mathcal{B}_{R^*}(x/a, y/a) \\
&= \lim_{a \to 0^+} \langle x, \nabla R^*(x/a) - \nabla R^*(y/a) \rangle - aR(\nabla R^*(x/a)) + aR(\nabla R^*(y/a)) \\
&= \lim_{a \to 0^+} \langle x, \nabla R^*(x/a) - \nabla R^*(y/a) \rangle = \langle x, u - v \rangle \ .
\end{aligned}
$$

C Proofs for AdaFTRL

Proof (Proof of Corollary 1). Let $S = \sup_{v \in K} f(v)$. Theorem 1 applied to the regularizer $R(w) = \frac{c}{S} f(w)$ and Proposition 1 gives

$$
\text{Regret}_T \leq \sqrt{3}(1+c) \max \left\{ \sqrt{8}, \frac{1}{\sqrt{2c}} \right\} \sqrt{S \sum_{t=1}^{T} \|\ell_t\|_*^2} \ .
$$

It remains to find the minimum of $g(c) = \sqrt{3}(1+c) \max\{\sqrt{8}, 1/\sqrt{2c}\}$. The function g is strictly convex on $(0, \infty)$ and has minimum at $c = 1/16$ and $g(\frac{1}{16}) = \sqrt{3}(1 + \frac{1}{16})\sqrt{8} \leq 5.3$.

Proof (Proof of Lemma 4). Let $a_t = \|\ell_t\|_* \max\{D, 1/\sqrt{2\lambda}\}$. The statement of the lemma is equivalent to $\Delta_T \leq \sqrt{3 \sum_{t=1}^T a_t^2}$ which we prove by induction on T. The base case $T = 0$ is trivial. For $T \geq 1$, we have

$$
\Delta_T \leq \Delta_{T-1} + \min \left\{ a_T, \ \frac{a_T^2}{\Delta_{T-1}} \right\} \leq \sqrt{3 \sum_{t=1}^{T-1} a_t^2} + \min \left\{ a_T, \frac{a_T^2}{\sqrt{3 \sum_{t=1}^{T-1} a_t^2}} \right\}
$$

where the first inequality follows from Lemma 3, and the second inequality from the induction hypothesis and the fact that $f(x) = x + \min\{a_T, a_T^2/x\}$ is an increasing function of x. It remains to prove that

$$
\sqrt{3 \sum_{t=1}^{T-1} a_t^2} + \min \left\{ a_T, \frac{a_T^2}{\sqrt{3 \sum_{t=1}^{T-1} a_t^2}} \right\} \leq \sqrt{3 \sum_{t=1}^{T} a_t^2} \ .
$$

Dividing through by a_T and making substitution $z = \frac{\sqrt{\sum_{t=1}^{T-1} a_t^2}}{a_T}$, leads to

$$
z\sqrt{3} + \min \left\{ 1, \frac{1}{z\sqrt{3}} \right\} \leq \sqrt{3 + 3z^2}
$$

which can be easily checked by considering separately the cases $z \in [0, \frac{1}{\sqrt{3}})$ and $z \in [\frac{1}{\sqrt{3}}, \infty)$.

D Proofs for SOLO FTRL

Proof (Proof of Corollary 2). Let $S = \sup_{v \in K} f(v)$. Theorem 2 applied to the regularizer $R(w) = \frac{c}{\sqrt{S}} f(w)$, together with Proposition 1 and a crude bound $\max_{t=1,2,\ldots,T} \|\ell_t\|_* \leq \sqrt{\sum_{t=1}^T \|\ell_t\|_*^2}$, give

$$\text{Regret}_T \leq \left(c + \frac{2.75}{c} + 3.5\sqrt{8} \right) \sqrt{S \sum_{t=1}^T \|\ell_t\|_*^2} \; .$$

We choose c by minimizing $g(c) = c + \frac{2.75}{c} + 3.5\sqrt{8}$. Clearly, $g(c)$ has minimum at $c = \sqrt{2.75}$ and has minimal value $g(\sqrt{2.75}) = 2\sqrt{2.75} + 3.5\sqrt{8} \leq 13.3$.

E Lower Bound Proof

Proof (Proof of Theorem 3). Pick $x, y \in K$ such that $\|x - y\| = D$. This is possible since K is compact. Since $\|x - y\| = \sup\{\langle \ell, x - y \rangle \ : \ \ell \in V^*, \|\ell\|_* = 1\}$ and the set $\{\ell \in V^* \ : \ \|\ell\|_* = 1\}$ is compact, there exists $\ell \in V^*$ such that

$$\|\ell\|_* = 1 \qquad \text{and} \qquad \langle \ell, x - y \rangle = \|x - y\| = D \; .$$

Let $Z_1, Z_2, \ldots, Z_T$ be i.i.d. Rademacher variables, that is, $\Pr[Z_t = +1] = \Pr[Z_t = -1] = 1/2$. Let $\ell_t = Z_t a_t \ell$. Clearly, $\|\ell_t\|_* = a_t$. The lemma will be proved if we show that (3) holds with positive probability. We show a stronger statement that the inequality holds in expectation, i.e. $\mathbf{E}[\text{Regret}_T] \geq \frac{D}{\sqrt{8}} \sqrt{\sum_{t=1}^T a_t^2}$. Indeed,

$$\begin{aligned}
\mathbf{E}\left[\text{Regret}_T\right] &\geq \mathbf{E}\left[\sum_{t=1}^T \langle \ell_t, w_t \rangle\right] - \mathbf{E}\left[\min_{u \in \{x,y\}} \sum_{t=1}^T \langle \ell_t, u \rangle\right] \\
&= \mathbf{E}\left[\sum_{t=1}^T Z_t a_t \langle \ell, w_t \rangle\right] + \mathbf{E}\left[\max_{u \in \{x,y\}} \sum_{t=1}^T -Z_t a_t \langle \ell, u \rangle\right] \\
&= \mathbf{E}\left[\max_{u \in \{x,y\}} \sum_{t=1}^T -Z_t a_t \langle \ell, u \rangle\right] = \mathbf{E}\left[\max_{u \in \{x,y\}} \sum_{t=1}^T Z_t a_t \langle \ell, u \rangle\right] \\
&= \frac{1}{2} \mathbf{E}\left[\sum_{t=1}^T Z_t a_t \langle \ell, x + y \rangle\right] + \frac{1}{2} \mathbf{E}\left[\left|\sum_{t=1}^T Z_t a_t \langle \ell, x - y \rangle\right|\right] \\
&= \frac{D}{2} \mathbf{E}\left[\left|\sum_{t=1}^T Z_t a_t\right|\right] \geq \frac{D}{\sqrt{8}} \sqrt{\sum_{t=1}^T a_t^2}
\end{aligned}$$

where we used that $\mathbf{E}[Z_t] = 0$, the fact that distributions of Z_t and $-Z_t$ are the same, the formula $\max\{a, b\} = (a + b)/2 + |a - b|/2$, and Khinchin's inequality in the last step (Lemma A.9 in [3]).