A GENERAL CLASS OF ADAPTIVE ALGORITHMS FOR APPROACHABILITY, REGRET AND OPTIMIZATION

JOON KWON

Abstract. Blackwell's approachability is a general online learning framework that contains regret minimization and many variants as special cases. The basic algorithm in this context is Blackwell's algorithm. Despite its simplicity, it demonstrates remarkable adaptivity to unknown problem characteristics in a parameter-free fashion, and excellent empirical performance. We develop a wide and unifying class of approachability algorithms that share a similar parameter-free and *Blackwell-like adaptive character*, offering a novel theoretical insight into the favorable characteristics of Blackwell's algorithm and its known variants. We then introduce several key innovations to apply Blackwell's algorithm to unconstrained regret minimization. The resulting regret minimizers compete with state-of-the-art bounds and also inherit the stronger *Blackwell-like adaptive character*. In turn, these algorithms are applied to convex optimization, achieving optimal dependence on the distance to the solution, on Lipschitz regularity and on Hölder-smoothness, without prior knowledge. Similarly, application to fixed point problems yields iterations that optimally adapt to the distance to the set of fixed points and to regularity properties analogous to Hölder-smoothness. Beyond those properties, the resulting algorithms bring to convex optimization and fixed point problems the *Blackwell-like adaptive character*, which we argue is stronger, illustrating the potential of using Blackwell's approachability for defining general iterative methods.

Contents

1. Introduction 1
2. Blackwell's approachability with convex cone targets 8
3. From regret to approachability 10
4. FTRL-type algorithms 12
5. OMD-type algorithms 16
6. Regret minimization on the simplex 20
7. From approachability to unconstrained regret minimization 22
8. Bounded payoff vectors 25
9. Unbounded payoff vectors 26
10. Convex optimization 29
11. Fixed point iterations 31
12. Conclusion and perspectives 33
References 34
Appendix 39

1. Introduction

1.1. **Blackwell's approachability.** Blackwell's approachability (Blackwell, 1956b) is an online learning framework that encompasses external and internal/swap regret minimization (Hart and Mas-Colell, 2000; Blum and Mansour, 2007), regret minimization with global costs (Even-Dar et al.,

Date: November 13, 2026.

2009; Kwon, 2021), with variable stage duration (Mannor and Shimkin, 2008), with fairness constraints (Chzhen et al., 2021), asymptotic calibration (Foster, 1999), online resource allocation, including load balancing and capacity pooling (Even-Dar et al., 2009; Zhong et al., 2018). Extensions of the framework also address adaptive regret and sleeping experts (Lee et al., 2022). See Perchet (2014) for a survey.

The framework is a repeated play between a Decision Maker and Nature. Let $d \geqslant 1$. At each stage $t \geqslant 1$, the Decision Maker (resp. Nature) chooses an action $a_t \in \mathcal{A}$ (resp. $b_t \in \mathcal{B}$), producing the *outcome vector* $r_t \coloneqq \rho(a_t, b_t) \in \mathbb{R}^d$, where $\rho\colon \mathcal{A} \times \mathcal{B} \to \mathbb{R}^d$ is the *outcome function*. Traditionally, the Decision Maker seeks to make the average outcome converge to a given set $\mathcal{C} \subset \mathbb{R}^d$ called the *target*. Whether this can be guaranteed by the Decision Maker depends on the outcome function and the target. For bounded outcome functions, Blackwell (1956b) gave a sufficient condition, now called *Blackwell's condition*, and a simple algorithm, now called *Blackwell's algorithm*, for guaranteeing such convergence. When the condition is satisfied, the convergence can be guaranteed at speed $1/\sqrt{T}$.

Blackwell's algorithm is strikingly simple, and its convergence guarantee holds with absolutely no parameter to be tuned, and in particular, no prior knowledge of the magnitude of the outcome vectors or of the time horizon. As will be discussed, this *parameter-free and adaptive character* is highly unusual and nontrivial to achieve in online learning and related problems such as first-order optimization.

These very favorable characteristics have also been observed in practice, in particular when Blackwell's algorithm is applied to the important problem of minimizing regret on the simplex through a reduction proposed by Hart and Mas-Colell (2000). In this context, Blackwell's algorithm is called Regret Matching (RM). Although standard analysis of RM has a suboptimal dependence on the dimension[1], RM and especially its variant RM$^+$ demonstrate excellent empirical performance[2]. This suggests that there may be more to uncover regarding the adaptive properties of Blackwell's algorithm. Through the counterfactual regret minimization approach, RM and RM$^+$ underlie the CFR and CFR$^+$ methods (Zinkevich et al., 2007; Tammelin et al., 2015), which played a central role in the development of superhuman poker AIs[3] (Moravčík et al., 2017; Brown and Sandholm, 2018), further demonstrating the excellent practical performance of approachability-based methods. Our initial goal is to deepen the theoretical understanding of the adaptive properties of Blackwell's algorithm.

Besides RM and RM$^+$, a few other approachability-based algorithms possess this parameter-free and adaptive character: the Polynomially Weighted Average Forecaster (Cesa-Bianchi and Lugosi, 2003), its RM$^+$-type variant (D'Orazio, 2020, Section 3.3), CBA and CBA$^+$ (Grand-Clément and Kroer, 2024). Note that RM and CBA are special cases of Blackwell's algorithm.

One of our contributions is a general scheme for constructing approachability algorithms for conic targets that enjoy similar parameter-free and adaptive guarantees.

1.2. **Online linear optimization.** An online learning framework that has fundamental ties with Blackwell's approachability is Online Linear Optimization (OLO). In the present work, both problems will feed one another: (i) OLO will be used to construct approachability algorithms and then (ii) approachability will be used in turn to define new parameter-free and adaptive algorithms for unconstrained OLO.

[1] $\sqrt{d}$ instead of the optimal $\sqrt{\log d}$ attained by the Exponential Weights Algorithm (Vovk, 1990; Kivinen and Warmuth, 1997; Cesa-Bianchi et al., 1997)

[2] In experiments on random matrix games for example, RM$^+$ outperforms AdaHedge (van Erven et al., 2011; De Rooij et al., 2014), an Exponential Weights variant with optimal $\sqrt{\log d}$ dependence, that is adaptive to the realized ℓ^∞-norm of the payoff vectors (Grand-Clément and Kroer, 2024).

[3] Such as DeepStack and Libratus that defeated professional players at heads-up no-limit Texas hold'em.

OLO is described as follows. Let $\mathcal{X}$,$\mathcal{U} \subset \mathbb{R}^d$ be nonempty, with $\mathcal{X}$ being convex. At each stage $t \geqslant 1$,

- the Decision Maker chooses[4] $x_t \in \mathcal{X}$,
- Nature chooses a *payoff vector* $u_t \in \mathcal{U}$ and reveals it to the Decision Maker,
- the Decision Maker receives payoff $\langle u_t, x_t \rangle$.

For a time-horizon $T \geqslant 1$ and a *comparator* $x \in \mathcal{X}$, the cumulative *regret* is

$$\sum_{t=1}^{T} \langle u_t, x - x_t \rangle .$$

This quantity is the difference between the payoff of the fixed action x and the Decision Maker's payoff against the same payoff sequence. Traditionally in the online learning literature, $\mathcal{X}$ is assumed compact, so the worst-case regret

$$\max_{x \in \mathcal{X}} \sum_{t=1}^{T} \langle u_t, x - x_t \rangle \tag{1}$$

is finite, and is the quantity of interest. If $\mathcal{U}$ is also bounded, there exist algorithms that guarantee an average regret bound of order $O(1/\sqrt{T})$ (Cesa-Bianchi and Lugosi, 2006; Shalev-Shwartz, 2012; Hazan, 2016; Orabona, 2026). Such *no-regret* guarantees are provided, for instance, by suitable Follow the Regularized Leader (FTRL) and Online Mirror Descent (OMD) algorithms—see also the unified UMD family (Juditsky et al., 2023). Standard fixed-parameter versions use the time horizon and a bound on the payoff magnitudes to tune their learning rates. On the unit simplex, an important special case of both families is the Exponential Weights Algorithm (Vovk, 1990; Kivinen and Warmuth, 1997; Cesa-Bianchi et al., 1997).

Unbounded domains are also relevant, particularly the *unconstrained case* $\mathcal{X} = \mathbb{R}^d$, because OLO algorithms can be applied to unconstrained optimization. In the unbounded setting, the supremum over comparators corresponding to (1) need not be finite. Nevertheless, FTRL and OMD algorithms give *comparator-dependent* bounds of the following form:

$$\sum_{t=1}^{T} \langle u_t, x - x_t \rangle \leqslant B_T(x), \quad x \in \mathcal{X}. \tag{2}$$

The simplest example is unconstrained Online Gradient Descent (OGD): given $x_1 \in \mathbb{R}^d$ and a parameter[5] $\gamma > 0$, its update at each stage $t \geqslant 1$ is

$$x_{t+1} = x_t + \gamma u_t.$$

For any comparator $x \in \mathbb{R}^d$, the identity

$$\|x_{t+1} - x\|_2^2 = \|x_t - x\|_2^2 + 2\gamma \langle u_t, x_t - x \rangle + \gamma^2 \|u_t\|_2^2$$

yields, by summation, the following regret bound:

$$\sum_{t=1}^{T} \langle u_t, x - x_t \rangle \leqslant \frac{\|x_1 - x\|_2^2}{2\gamma} + \frac{\gamma}{2} \sum_{t=1}^{T} \|u_t\|_2^2, \quad T \geqslant 1. \tag{3}$$

The above is an example of the *two typical regret bound terms* that need to be balanced by choosing the best possible parameter γ given the available information. When this regret bound is applied to an optimization problem with an unknown solution x_*, only the comparator $x = x_*$ is considered,

[4] Formally, an algorithm for the Decision Maker is a sequence of maps $\sigma = (\sigma_t)_{t \geqslant 1}$ where for each step $t \geqslant 1$, the map $\sigma_t : \mathcal{U}^{t-1} \to \mathcal{X}$ prescribes the action to choose as a function of previously observed payoff vectors, meaning $x_t = \sigma_t(u_1, \ldots, u_{t-1})$.

[5] also called *step-size*, or *learning rate*, depending on the context

and the dependence on the distance to the solution (here $\|x_1 - x_*\|_2^2$) is a key concern. Besides, our construction of parameter-free and adaptive approachability algorithms will rely on OLO on nontrivial convex cones, which are unbounded, strengthening our interest in unbounded domains.

Note that an algorithm may guarantee the regret bound (2) for every $x \in \mathcal{X}$ while its actions remain in a strict subset of $\mathcal{X}$. For example, with strictly positive initial weights, the Exponential Weights Algorithm always plays in the relative interior of the simplex, yet competes with every point of the simplex. This distinction will matter for our applications to unconstrained OLO and optimization. We call $\mathcal{X}$ the *comparator domain*, or simply the *domain*.

1.3. **Parameter-free regret minimization.** A central issue with (fixed-parameter) FTRL and OMD algorithms is that they are suboptimal both theoretically and empirically when used untuned. The same issue arises when they are applied to (stochastic) first-order optimization, and is well-known to e.g. deep learning practitioners.

The OGD bound (3) suggests the following classical thought experiment. Fix a comparator $x \neq x_1$ and suppose that $\sum_{t=1}^T \|u_t\|_2^2 > 0$. If T, the initial distance $\|x - x_1\|_2$ and the latter sum were *known in advance*, choosing

$$\gamma = \frac{\|x_1 - x\|_2}{\sqrt{\sum_{t=1}^T \|u_t\|_2^2}}$$

would guarantee

$$\sum_{t=1}^T \langle u_t, x - x_t \rangle \leqslant \|x_1 - x\|_2 \sqrt{\sum_{t=1}^T \|u_t\|_2^2}. \tag{4}$$

This *tuning in hindsight* improves the quadratic dependence on the initial distance to a linear one, and replaces the sum of squared payoff norms by its square root. Without this information, one may tune γ using some available estimates of these quantities (if any). A central goal of parameter-free regret minimization is to achieve regret bounds that are close to (4) simultaneously for all comparators $x \in \mathbb{R}^d$ (Streeter and McMahan, 2012; McMahan and Orabona, 2014; Orabona and Pal, 2016; Cutkosky and Orabona, 2018; Foster et al., 2018; Cutkosky, 2019b; Mhammedi and Koolen, 2020; Zhang and Cutkosky, 2022; Cutkosky and Mhammedi, 2024; Zhang et al., 2024). Note that the AdaGrad algorithms (McMahan and Streeter, 2010; Duchi et al., 2011) achieve optimal adaptivity to $\sqrt{\sum_{t=1}^T \|u_t\|_2^2}$ but not to the initial distance to the comparator.

In the present work, the algorithms first defined for approachability will be applied to unconstrained OLO and yield algorithms that inherit Blackwell-like adaptive and parameter-free guarantees that will turn out to be novel and stronger than those of existing algorithms.

1.4. **Approachability via OLO.** The first important relation between OLO and approachability is that OLO algorithms can be converted into approachability algorithms for conic targets. This was first proposed by Abernethy et al. (2011) and then extended by Kwon (2016, 2021) and Farina et al. (2021). See also Perchet (2015) for a conversion of the Exponential Weights algorithm to approachability.

The works of Abernethy et al. (2011) and Kwon (2016, 2021) first choose convex compact sets $\mathcal{X}$ that contain all directions of the polar cone of the target, define OLO algorithms (e.g. some FTRL or OMD algorithm) with comparator-domain $\mathcal{X}$, and take advantage of regret bounds of the form (1). The resulting approachability algorithms have convergence bounds that can better accommodate various assumptions on the outcome vectors, but require parameters tuned using a known upper bound on these vectors[6], as do FTRL and OMD algorithms in their native OLO

[6]except in special cases such as Kwon (2021, Theorem 26), Farina et al. (2021, Theorem 3) and Flaspohler et al. (2021, Lemma 7 and Corollary 9)

context. Importantly, Farina et al. (2021, Proposition 2) extends Abernethy et al. (2011) to allow the auxiliary OLO algorithm to be defined on the whole polar cone. The algorithms defined in the present work will rely on an extension of all three mentioned works, that allows the use of dense subcones of the polar cone.

1.5. **OLO via approachability.** Very successful applications of approachability are the already mentioned RM and RM$^+$ algorithms, which rely on the reduction of regret minimization on the simplex to an approachability problem with the negative orthant as the target set (Hart and Mas-Colell, 2000) (see Section 6 for details). We may call it the *orthant reduction* for short. Unfortunately, although very simple and efficient, this reduction does not extend to OLO on arbitrary domains except in very special cases—see e.g. Chakrabarti et al. (2024).

The orthant reduction also admits potential-based generalizations of Blackwell's algorithm (Hart and Mas-Colell, 2001; Cesa-Bianchi and Lugosi, 2003). These constructions include polynomial and exponential potentials and yield a broader family of regret minimizers. Lagrangian Hedging extends this potential-based approach to online convex programs (Gordon, 2007), and D'Orazio and Huang (2021) develop optimistic and adaptive variants.

However, an alternative reduction of OLO on the simplex to approachability dates back to the origins of Blackwell's approachability and uses a lifted outcome space obtained by adding a scalar coordinate. Section 7.2 below describes a modern variant of this reduction which differs from the original only in a minor way (Blackwell, 1956a). An important advantage of this approach, which we may call *Blackwell's reduction* for short, is that it easily extends to OLO on any convex compact domain as discussed in Perchet (2014), and this is how Grand-Clément and Kroer (2024) obtain the CBA and CBA$^+$ OLO algorithms.

Both the orthant and Blackwell's reductions allow us to define OLO algorithms that inherit the favorable parameter-free and adaptive character of Blackwell's algorithm, which the usual FTRL and OMD algorithms for OLO do not possess.

We now turn to unbounded domains (and in particular the unconstrained case $\mathcal{X} = \mathbb{R}^d$). A natural question is whether Blackwell's reduction can be extended to yield similar approachability-based OLO algorithms for unbounded domains. The main difficulty is that the corresponding target set no longer satisfies Blackwell's condition. We overcome this obstacle by introducing a *relaxed* Blackwell's condition and defining an auxiliary OLO algorithm on a dense subcone of the polar cone of the target. The resulting algorithms will constitute strong contributions to the topic of parameter-free regret minimization.

A peculiar aspect of reducing OLO to approachability is that, for defining algorithms for the auxiliary approachability problem, we may use, as explained above, OLO algorithms! Although this may sound like the snake is biting its own tail, going through the approachability formulation will prove beneficial, as it provides a general approach for defining parameter-free and adaptive algorithms.[7]

1.6. **First-order gradient methods.** The fundamental relation between OLO and first-order gradient methods has been known for a long time. The Mirror Descent algorithms from Nemirovski and Yudin (1983) and Beck and Teboulle (2003) were adapted to OLO by Warmuth and Jagota (1998) and Zinkevich (2003).

The Dual Averaging algorithms[8] (Nesterov, 2009) were originally introduced for convex optimization. They were then studied as OLO algorithms under the name Follow the Regularized Leader

[7]It is nevertheless possible to present such a construction without the approachability interpretation, as in the follow-up paper (Grand-Clément et al., 2026).

[8]Dual Averaging first appeared in Nesterov's CORE Discussion Paper 2005/67 and was later published in journal form as Nesterov (2009).

(FTRL) by Shalev-Shwartz (2007), Shalev-Shwartz and Singer (2007), Abernethy et al. (2008), and Hazan and Kale (2010).

Conversely, the classical conversion scheme from Cesa-Bianchi et al. (2004) turns any OLO regret bound into a convex optimization guarantee. Several variants and extensions were proposed (Levy, 2017; Cutkosky, 2019a; Defazio et al., 2024), including online-to-nonconvex conversion schemes (Cutkosky et al., 2023).

An important example of algorithms that were first introduced as regret minimizers and then applied to first-order optimization is the AdaGrad family (McMahan and Streeter, 2010; Duchi et al., 2011). Its strong theoretical and empirical properties encouraged the introduction of many variants such as RMSprop (Tieleman and Hinton, 2012), Adam (Kingma and Ba, 2015), AdamW (Loshchilov and Hutter, 2019) that are state-of-the-art deep learning optimizers.

Parameter-free regret minimization has also led to first-order methods that adapt to the distance to the solution and, depending on the method and the setting, adapt to gradient magnitudes and/or noise (Orabona and Pal, 2016; Orabona and Tommasi, 2017; Jun and Orabona, 2019; Chen et al., 2022). Online-to-batch conversions combined with optimism and acceleration can also adapt to (Hölder-)smoothness while retaining guarantees for nonsmooth objectives (Cutkosky, 2019a). Other approaches that do not rely on regret also give various adaptive first-order methods for convex optimization (Carmon and Hinder, 2022; Ivgi et al., 2023; Defazio and Mishchenko, 2023; Khaled et al., 2023; Mishchenko and Defazio, 2024; Rodomanov et al., 2024; Kreisler et al., 2024; Ren et al., 2025; Li and Lan, 2026; Moshtaghifar et al., 2026).

The parameter-free unconstrained OLO algorithms defined in the present work will also be converted into simple first-order optimization algorithms that inherit a Blackwell-like adaptive character, which we argue is stronger than the various types of adaptivity mentioned above.

1.7. **Fixed point iterations with nonexpansive maps.** Fixed points are a general formulation for constructing iterative methods for solving e.g. linear systems, ODEs and PDEs, and various optimization problems. Beyond the very favorable contractive case, an important class of problems is described with nonexpansive maps[9]. See Bauschke and Combettes (2017) and Ryu and Yin (2022) for monographs presenting the theory of monotone operators and splitting methods, which is built upon this simple setting.

It was recently noticed by Kwon (2026) that OLO algorithms can be used to solve fixed point problems defined by nonexpansive maps. In particular, converting the unconstrained Online Gradient Descent recovers the classical Krasnoselskii–Mann (KM) iteration (Mann, 1953; Krasnosel'skii, 1955).

A highlight from Kwon (2026) is the conversion of AdaGrad regret minimizers into fixed point iterations with the following adaptive character. Suppose that there exists $L > 0$ such that $F_L := I + \frac{1}{L}(F - I)$ is nonexpansive, where I denotes the identity map. This map has the same fixed points as F, which need not be nonexpansive. The AdaGrad-based iteration with F offers convergence guarantees similar to KM with F_L, without prior knowledge of L. This result suffers from one important limitation: the bounds do not offer the optimal dependence on the initial distance to the set of fixed points.

We will convert our approachability-based parameter-free unconstrained OLO algorithms to obtain fixed-point iterations that simultaneously adapt, without prior information, to (i) the initial distance to the set of fixed points and to (ii) regularity properties that are more general than nonexpansiveness of the above map F_L and are analogous to Hölder smoothness.

1.8. **Contributions.** We consider in Section 2 an approachability framework which allows for outcome functions with no regularity assumption. In this context, we introduce a relaxed variant

[9] meaning 1-Lipschitz continuous

of Blackwell's condition for convex conic targets. When the outcome function is bounded, we prove that it is a sufficient condition for the Decision Maker to possess an algorithm guaranteeing convergence to the target set.

The second contribution, in Section 3 is a conversion scheme from OLO algorithms to approachability algorithms for convex conic targets, that extends Abernethy et al. (2011), Kwon (2016),Kwon (2021), and Farina et al. (2021) to the relaxed variant of Blackwell's condition. It allows for the auxiliary OLO algorithm to be defined on whole subcones of the polar cone of the target, which will be key in deriving the parameter-free and adaptive properties that we seek.

The third contribution is the study in Sections 4 and 5 of general approachability algorithms obtained by applying our new conversion scheme with FTRL and OMD algorithms, respectively. An important highlight is the case where the auxiliary FTRL (resp. OMD) algorithm is defined on a whole cone with a (partly) 2-positively-homogeneous regularizer (resp. mirror map). A new analysis technique then demonstrates that the *two typical regret bound terms* (mentioned above) automatically balance, without the need to tune a parameter. This yields a wide class of approachability algorithms that share a *parameter-free and adaptive character* similar to Blackwell's original algorithm, while providing a novel theoretical insight. Notable special cases include (i) a general OGD counterpart to Blackwell's algorithm that contains RM^+ and CBA^+ and (ii) the Polynomially Weighted Average Forecaster (Cesa-Bianchi and Lugosi, 2003) and (iii) its RM^+-like counterpart (D'Orazio, 2020), demonstrating the unifying character of our construction.

Our fourth contribution applies the above algorithms to unconstrained OLO by extending Blackwell's reduction to this setting using the relaxed Blackwell's condition (Section 7). Introducing carefully constructed regularizers on a dense subcone of the polar cone of the target, this approach yields two new unconstrained OLO algorithms: (I) one in Section 8 for payoff vectors with a known bound on their norms, and (II) the other in Section 9 for arbitrary and possibly unbounded payoff vectors with no prior knowledge on their magnitude. Both algorithms compete with state-of-the-art guarantees in terms of regret bounds that explicitly depend on the norms of the comparator and payoff vectors. They are, however, much simpler in their definition, analysis, and bound expressions, and they are easily implementable. We also argue that the Blackwell-like adaptive character that these algorithms inherit provides stronger guarantees than existing algorithms.

Finally, the above algorithm (I) is applied to various optimization problems. In Section 10, it is converted into a first-order convex optimization algorithm using the normalized gradients approach (Levy, 2017; Orabona, 2023). The analysis offers nonaccelerated guarantees that adapt simultaneously to Lipschitz continuity, (Hölder) smoothness, and initial distance to the solution, up to logarithmic factors and additive terms in the distance dependence, with no prior knowledge of these quantities. These guarantees also retain the Blackwell-like adaptive character. Inspired by Kwon (2026), algorithm (I) is also converted in Section 11 into fixed point iterations for a given operator $F\colon \mathbb{R}^d \to \mathbb{R}^d$, with guarantees that are optimally adaptive, up to logarithmic factors and additive terms in the distance dependence, to (i) the distance to the set of fixed points, (ii) scaling factors $L > 0$ such that $I + \frac{1}{L}(F - I)$ is nonexpansive, and also to (iii) regularity properties that are analogous to Hölder smoothness. The combination of properties (i) and (ii) is novel, and adaptivity to (iii) is also novel.

Together, these results provide a new approach for designing practical iterative methods with strong adaptive guarantees, motivated by the empirical success of existing Blackwell-based regret minimizers.

1.9. **Related work.** The follow-up paper (Grand-Clément et al., 2026) further extends Blackwell's reduction of OLO to approachability to work with arbitrary closed convex domains, and chooses

regularizers that are different from ours to adapt to non-Euclidean norms, and to recover coin-betting algorithms (Orabona and Pal, 2016) such as Krichevsky–Trofimov, further demonstrating the unifying character of this approach.

1.10. **Definitions and notation.** Let $d \geqslant 1$ be an integer. We equip $\mathbb{R}^d$ with its usual inner product $\langle x, y\rangle = \sum_{i=1}^d x_i y_i$. For $1 \leqslant p < +\infty$, write $\|x\|_p = (\sum_{i=1}^d |x_i|^p)^{1/p}$ and $\|x\|_\infty = \max_{1\leqslant i\leqslant d} |x_i|$. For a norm $\|\cdot\|$, its dual norm is $\|y\|_* = \sup_{\|x\|\leqslant 1} \langle y, x\rangle$, and B denotes its closed unit ball. For a nonempty set $\mathcal{X} \subset \mathbb{R}^d$, set

$$\operatorname{dist}_{\|\cdot\|}(y, \mathcal{X}) = \inf_{x\in\mathcal{X}} \|y - x\| .$$

We write dist_p for the distance induced by $\|\cdot\|_p$. When $\mathcal{X}$ is closed and convex, $\Pi_{\mathcal{X}}(y)$ denotes the Euclidean projection of y onto $\mathcal{X}$.

We denote by $\operatorname{cl}\mathcal{X}$, $\operatorname{int}\mathcal{X}$ and $\operatorname{rel\,int}\mathcal{X}$ the closure, interior and relative interior of $\mathcal{X}$. For $y \in \mathbb{R}^d$, its positive part is $y_+ = (\max\{y_i, 0\})_{1\leqslant i\leqslant d}$. The same notation applies to scalars. Set $\mathbb{1} = (1, \dots, 1) \in \mathbb{R}^d$ and

$$\Delta_d = \left\{ a \in \mathbb{R}^d_+, \sum_{i=1}^d a_i = 1 \right\} .$$

For a function $f\colon \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$, its domain is $\operatorname{dom} f = \{x \in \mathbb{R}^d, f(x) < +\infty\}$. It is said to be proper if $\operatorname{dom} f \neq \varnothing$. For a proper convex function f and $x \in \operatorname{dom} f$, its subdifferential is

$$\partial f(x) = \left\{ y \in \mathbb{R}^d,\ \forall x' \in \mathbb{R}^d,\ f(x') \geqslant f(x) + \left\langle y, x' - x \right\rangle \right\} ,$$

and $\operatorname{dom}\partial f = \{x \in \operatorname{dom} f, \partial f(x) \neq \varnothing\}$. Its Legendre–Fenchel transform, or convex conjugate, is

$$f^*(y) = \sup_{x\in\mathbb{R}^d} \{\langle y, x\rangle - f(x)\} .$$

For a convex set $\mathcal{X}$, the convex indicator $I_{\mathcal{X}}$ is zero on $\mathcal{X}$ and $+\infty$ outside $\mathcal{X}$. If f is differentiable at $x \in \operatorname{int}\operatorname{dom} f$, its Bregman divergence is defined, for $x' \in \operatorname{dom} f$, by

$$D_f(x', x) = f(x') - f(x) - \left\langle \nabla f(x), x' - x \right\rangle .$$

A convex cone is a nonempty set closed under addition and multiplication by nonnegative scalars. In particular, it contains zero. The polar cone of a set $\mathcal{X} \subset \mathbb{R}^d$ is

$$\mathcal{X}^\circ = \left\{ y \in \mathbb{R}^d,\ \forall x \in \mathcal{X},\ \langle y, x\rangle \leqslant 0 \right\} .$$

A function f is positively homogeneous of degree 2 if $f(\lambda x) = \lambda^2 f(x)$ for all $x \in \mathbb{R}^d$ and $\lambda > 0$.

2. Blackwell's approachability with convex cone targets

We present an approachability framework for target sets that are closed convex cones[10]. Unlike what is often considered in the literature, we assume neither compactness of the Decision Maker's action set nor any regularity of the outcome function. We introduce a relaxed variant of Blackwell's condition and prove that, for bounded outcome functions, it is a sufficient condition for the existence of an algorithm for the Decision Maker that guarantees convergence of the average outcome to the target set. When the Decision Maker's action set is compact and the outcome function is continuous in that action, the relaxed condition is proved to be equivalent to the traditional one.

[10]This restriction is motivated as follows. First, all applications of approachability we are aware of naturally define targets as cones. Secondly, our analysis technique will strongly rely on the cone structure to achieve the kind of parameter-free and adaptive guarantees that we seek. Finally, for a given outcome function $\rho\colon \mathcal{A} \times \mathcal{B} \to \mathbb{R}^d$ and an arbitrary nonempty closed convex target $\mathcal{T} \subset \mathbb{R}^d$, one can define the following conic lift in $\mathbb{R}^d \times \mathbb{R}$: $\tilde{\rho}(a, b) = (\rho(a, b), 1)$ and $\mathcal{C} = \operatorname{cl}(\mathbb{R}_+(\mathcal{T} \times \{1\}))$.

Let $\mathcal{A}, \mathcal{B}$ be two nonempty sets with no particular structure, and $\rho\colon \mathcal{A}\times\mathcal{B} \to \mathbb{R}^d$. The elements of $\mathcal{A}$ (resp. $\mathcal{B}$) are called the *actions* of the Decision Maker (resp. of Nature). ρ is called the *outcome function*. At step $t \geqslant 1$,

- the Decision Maker chooses action[11] $a_t \in \mathcal{A}$,
- Nature chooses action $b_t \in \mathcal{B}$,
- outcome vector $r_t \coloneqq \rho(a_t, b_t) \in \mathbb{R}^d$ is revealed to the Decision Maker.

For all $T \geqslant 1$, denote $R_T = \sum_{t=1}^{T} r_t$. Traditionally, the goal is to define algorithms for the Decision Maker which guarantee some bound on the distance of the average (or cumulative) outcome vector to a given closed convex cone $\mathcal{C} \subset \mathbb{R}^d$ called the *target*.

Without further assumptions, convergence to the target may be impossible, since the outcome function can keep all outcomes away from it. A standard sufficient condition is the existence of an *oracle*. We also introduce a *relaxed* variant. Both involve the polar cone $\mathcal{C}^\circ$. See Appendix B for definition and properties.

Definition 2.1. (i) A function $\mathfrak{a}\colon \mathcal{C}^\circ \to \mathcal{A}$ is an *oracle* if

$$\forall x \in \mathcal{C}^\circ,\ \forall b \in \mathcal{B}, \quad \langle \rho(\mathfrak{a}(x), b), x\rangle \leqslant 0.$$

(ii) Let $\mathcal{K} \subset \mathbb{R}^d$ be a dense subcone of $\mathcal{C}^\circ$. A function $\mathfrak{a}\colon \mathcal{K} \to \mathcal{A}$ is a *relaxed oracle* if

$$\forall x \in \mathcal{K},\ \forall b \in \mathcal{B}, \quad \langle \rho(\mathfrak{a}(x), b), x\rangle \leqslant 0.$$

The existence of an *oracle* is traditionally called *Blackwell's condition* (Blackwell, 1956b). Our application to unconstrained OLO in Section 7, where no oracle exists, motivates relaxed oracles defined only on a dense subcone of $\mathcal{C}^\circ$. For bounded ρ, the following proposition shows that a relaxed oracle suffices for convergence of the average outcome to the target. If, in addition, $\mathcal{A}$ is compact and ρ is continuous in its first variable, both oracle conditions characterize this convergence.

Proposition 2.2. *Suppose that ρ is bounded and consider the following propositions.*

(i) An oracle exists.

(ii) A relaxed oracle exists.

(iii) There exists an algorithm for the Decision Maker that guarantees

$$\operatorname{dist}_2\left(\frac{1}{T}\sum_{t=1}^{T} r_t,\ \mathcal{C}\right) \leqslant \frac{\|\rho\|_2}{\sqrt{T}} + O\left(\frac{1}{T}\right),$$

as $T \to +\infty$, uniformly over all sequences of Nature's actions.

(iv) There exists an algorithm for the Decision Maker that guarantees

$$\operatorname{dist}_2\left(\frac{1}{T}\sum_{t=1}^{T} r_t,\ \mathcal{C}\right) \xrightarrow[T\to+\infty]{} 0,$$

against any sequence of Nature's actions.

Then, (i) implies (ii), which in turn implies (iii), which implies (iv). If moreover, $\mathcal{A}$ is a compact topological space and $a \mapsto \rho(a, b)$ is continuous for every $b \in \mathcal{B}$, the four propositions are equivalent.

Proof. (i) obviously always implies (ii), as an oracle is also a relaxed oracle. Assertion (iii) implies (iv) because ρ is bounded. Let us prove that (ii) implies (iii). Let $\mathcal{K}$ be a dense subcone of $\mathcal{C}^\circ$ and $\mathfrak{a}\colon \mathcal{K} \to \mathcal{A}$ be a relaxed oracle. Set $R_0 = 0$. At each step $t \geqslant 1$, since $\mathcal{K}$ is dense in $\mathcal{C}^\circ$, there exists $x_t \in \mathcal{K}$ such that

$$\|x_t - \Pi_{\mathcal{C}^\circ}(R_{t-1})\|_2 \leqslant 1/t^2, \tag{5}$$

[11] Formally, an algorithm of the Decision Maker is a sequence $(\sigma_t)_{t\geqslant 1}$ where for each step $t \geqslant 1$, the map $\sigma_t\colon (\mathbb{R}^d)^{t-1} \to \mathcal{A}$ prescribes the action to choose as a function of past outcome vectors, meaning $a_t = \sigma_t(r_1, \dots, r_{t-1})$.

where $\Pi_{\mathcal{C}^\circ}$ denotes the Euclidean projection onto $\mathcal{C}^\circ$ (which is well-defined, because $\mathcal{C}^\circ$ is a closed convex set according to Proposition B.4). An algorithm can then choose $a_t = \mathfrak{a}(x_t)$. Denote $\Pi_t = \Pi_{\mathcal{C}}(R_t)$. Then,

$$\begin{aligned}\min_{r\in\mathcal{C}} \|R_t - r\|_2^2 \leqslant \|R_t - \Pi_{t-1}\|_2^2 &= \|R_{t-1} + r_t - \Pi_{t-1}\|_2^2 \\ &= \|R_{t-1} - \Pi_{t-1}\|_2^2 + 2\langle R_{t-1} - \Pi_{t-1}, r_t\rangle + \|r_t\|_2^2 .\end{aligned}$$

The first term of the above last expression is simply $\min_{r\in\mathcal{C}} \|R_{t-1} - r\|_2^2$, and the last term is bounded by $\|\rho\|_2^2$, where

$$\|\rho\|_2 = \sup_{\substack{a\in\mathcal{A}\\ b\in\mathcal{B}}} \|\rho(a,b)\|_2 .$$

The dot product can be bounded from above using Moreau's decomposition theorem (from Theorem B.7, which gives $R_t = \Pi_t + \Pi_{\mathcal{C}^\circ}(R_t)$) and inequality (5) as

$$\begin{aligned}\langle R_{t-1} - \Pi_{t-1}, r_t\rangle = \langle \Pi_{\mathcal{C}^\circ}(R_{t-1}), r_t\rangle &= \langle \Pi_{\mathcal{C}^\circ}(R_{t-1}) - x_t, r_t\rangle + \langle x_t, r_t\rangle \\ &\leqslant \|\Pi_{\mathcal{C}^\circ}(R_{t-1}) - x_t\|_2 \|r_t\|_2 \leqslant \frac{\|\rho\|_2}{t^2},\end{aligned}$$

where $\langle x_t, r_t\rangle \leqslant 0$ because $a_t = \mathfrak{a}(x_t)$ by definition of the algorithm and $\mathfrak{a}$ is a relaxed oracle. Summing over $t = 1, \dots, T$ gives

$$\min_{r\in\mathcal{C}} \|R_T - r\|_2^2 \leqslant T\|\rho\|_2^2 + 2\|\rho\|_2 \sum_{t=1}^{T} \frac{1}{t^2}.$$

Taking the square root and dividing by T gives the convergence of R_T/T to $\mathcal{C}$.

Suppose now that $\mathcal{A}$ is compact and ρ is continuous in its first variable. If no oracle exists, there exists $x \in \mathcal{C}^\circ \setminus \{0\}$ such that $g(a) = \sup_{b\in\mathcal{B}} \langle \rho(a,b), x\rangle$ is positive for every $a \in \mathcal{A}$. Since g is lower semicontinuous on the compact set $\mathcal{A}$, $\delta = \min_{a\in\mathcal{A}} g(a) > 0$. After each action a_t, Nature can choose b_t such that $\langle \rho(a_t, b_t), x\rangle \geqslant \delta/2$. Consequently,

$$\operatorname{dist}_2(R_T/T, \mathcal{C}) \geqslant \frac{\langle R_T/T, x\rangle}{\|x\|_2} \geqslant \frac{\delta}{2\|x\|_2},$$

which contradicts (iv). Thus (iv) implies (i), and the four assertions are equivalent. □

In the remainder of the paper, $\mathcal{C}$ will be a given closed convex cone, $\mathcal{K}$ a dense subcone of $\mathcal{C}^\circ$ and, unless mentioned otherwise, $\mathfrak{a}\colon \mathcal{K} \to \mathcal{A}$ an associated relaxed oracle.

3. From regret to approachability

Inspired by Abernethy et al. (2011), Kwon (2016, 2021), and Farina et al. (2021), we present a more general scheme for converting an OLO algorithm into an approachability algorithm, which allows for the use of auxiliary OLO algorithms defined on dense subcones of the polar cone. This will be key in obtaining approachability algorithms possessing the *parameter-free and adaptive* character that we seek. First note that the distance (with respect to an arbitrary norm) to the target cone $\mathcal{C}$ can be written as follows.

Lemma 3.1 (Shimkin, 2016, Kwon, 2016, 2021)**.** *Let $\|\cdot\|$ be a norm in $\mathbb{R}^d$. Then,*

$$\operatorname{dist}_{\|\cdot\|_*}(y, \mathcal{C}) = \max_{\substack{x\in\mathcal{C}^\circ\\ \|x\|\leqslant 1}} \langle y, x\rangle .$$

Using the above, we can see that as soon as a nonnegative upper bound of the form

$$\left\langle \sum_{t=1}^{T} r_t, x \right\rangle \leqslant B_T(x), \tag{6}$$

is established for all $T \geqslant 1$ and all $x \in \mathcal{X}_0$, where $\mathcal{X}_0 \subset \mathcal{C}^\circ$ is a set containing all directions of a dense subset of $\mathcal{C}^\circ$, the distance of the average outcome vector to the target is controlled as follows

$$\operatorname{dist}_{\|\cdot\|_*}\left(\frac{1}{T}\sum_{t=1}^{T} r_t,\ \mathcal{C}\right) = \max\left\{0,\ \sup_{x\in\mathcal{X}_0\setminus\{0\}} \left\langle \frac{1}{T}\sum_{t=1}^{T} r_t, \frac{x}{\|x\|} \right\rangle\right\} \leqslant \max\left\{0,\ \sup_{x\in\mathcal{X}_0\setminus\{0\}} \frac{B_T(x)}{T\,\|x\|}\right\}.$$

If the above last upper bound vanishes as $T \to +\infty$, the average outcome vector converges to the target. We now turn to the lemma that makes the connection between $\left\langle \sum_{t=1}^{T} r_t, x \right\rangle$ and some OLO regret.

Lemma 3.2. *Let $\mathcal{K} \subset \mathbb{R}^d$ be a dense subcone of $\mathcal{C}^\circ$, $\mathfrak{a}\colon \mathcal{K} \to \mathcal{A}$ a relaxed oracle, $(x_t)_{t\geqslant 1}$ a sequence in $\mathcal{K}$, $(b_t)_{t\geqslant 1}$ a sequence in $\mathcal{B}$, and for all $t \geqslant 1$, $r_t = \rho(\mathfrak{a}(x_t), b_t)$. Let $T \geqslant 1$.*

(i) For all $x \in \mathbb{R}^d$,

$$\left\langle \sum_{t=1}^{T} r_t, x \right\rangle \leqslant \sum_{t=1}^{T} \langle r_t, x - x_t \rangle .$$

(ii) Consequently, if there exists $\mathcal{X} \subset \mathcal{C}^\circ$ and $B_T\colon \mathcal{X} \to \mathbb{R}_+$ such that

$$\sum_{t=1}^{T} \langle r_t, x - x_t \rangle \leqslant B_T(x), \quad x \in \mathcal{X},$$

then

$$\left\langle \sum_{t=1}^{T} r_t, x \right\rangle \leqslant \inf_{\substack{\lambda>0 \\ \lambda x\in\mathcal{X}}} \frac{B_T(\lambda x)}{\lambda}.$$

Proof. For all $t \geqslant 1$, the first inequality follows from the relaxed-oracle property of $\mathfrak{a}$:

$$\langle r_t, x_t \rangle = \langle \rho(\mathfrak{a}(x_t), b_t), x_t \rangle \leqslant 0.$$

Then, the second inequality follows by replacing x by λx (for all $\lambda > 0$ such that $\lambda x \in \mathcal{X}$), dividing by λ, and then taking the infimum over $\lambda > 0$. □

Lemma 3.2 converts auxiliary OLO regret bounds into bounds of the form (6). The inequality (ii) is especially useful when a meaningful regret bound on $\sum_{t=1}^{T} \langle r_t, \lambda x - x_t \rangle$ holds for every $\lambda > 0$, with every λx a valid comparator, as when the comparator domain is a convex cone. This is the key to obtaining parameter-free and adaptive guarantees for a large class of algorithms.

Our conversion scheme can be summarized as follows. For a given relaxed oracle $\mathfrak{a}\colon \mathcal{K} \to \mathcal{A}$,

- choose a convex comparator domain $\mathcal{X} \subset \mathcal{C}^\circ$,
- choose an auxiliary OLO algorithm on $\mathcal{X}$ that outputs actions in $\mathcal{K}$,
- use it with $(r_t)_{t\geqslant 1}$ as payoff vectors,
- get auxiliary actions $(x_t)_{t\geqslant 1}$,
- choose actions $a_t = \mathfrak{a}(x_t)$ $(t \geqslant 1)$ in the initial approachability problem.

Then, Lemma 3.2 applies and every regret guarantee for the auxiliary OLO algorithm yields an upper bound on $\left\langle \sum_{t=1}^{T} r_t, x \right\rangle$ for all $x \in \mathcal{X}$.

4. FTRL-type algorithms

We now apply the previous conversion scheme to FTRL algorithms (Nesterov, 2009; Shalev-Shwartz and Singer, 2007; Shalev-Shwartz, 2007) (see also Gordon, 1999b,a for an early version). Importantly, regularizers with positive homogeneity of degree 2 will yield a wide class of parameter-free and adaptive guarantees similar to Blackwell's original algorithm while providing a new theoretical insight. In particular, when such a regularizer also enjoys suitable strong convexity, the corresponding bound will adapt to the realized outcome norms without tuning a learning rate. Special cases will include Blackwell's algorithm (Blackwell, 1956b) (and thus the RM (Hart and Mas-Colell, 2000) and CBA (Grand-Clément and Kroer, 2024) algorithms) and the Polynomially Weighted Average Forecaster (Cesa-Bianchi and Lugosi, 2003), illustrating the unifying character of this construction. The algorithms we obtain below are also related to potential-based approachability algorithms proposed in Hart and Mas-Colell (2001), Cesa-Bianchi and Lugosi (2003), and Gordon (2007), as the Legendre–Fenchel transform of the regularizer plays the role of a (possibly time-dependent) potential.

Definition 4.1. A function $h\colon \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is a *pre-regularizer* if it is proper, strictly convex, and lower semicontinuous. Moreover, if $\operatorname{dom} h^* = \mathbb{R}^d$, then h is said to be a *regularizer*.

The above notation h^* denotes the Legendre–Fenchel transform:

$$h^*(y) = \sup_{x \in \mathbb{R}^d} \{\langle y, x\rangle - h(x)\}, \quad y \in \mathbb{R}^d,$$

whose properties are recalled in Appendix A for completeness. The following proposition gives several common sufficient conditions for a pre-regularizer to be a regularizer.

Proposition 4.2 (Juditsky et al., 2023, Proposition 2)**.** *Let h be a pre-regularizer.*

(i) If $\operatorname{dom} h$ is bounded, then h is a regularizer.
(ii) If h is differentiable on $\mathcal{D}_h \coloneqq \operatorname{int} \operatorname{dom} h$ and $\nabla h(\mathcal{D}_h) = \mathbb{R}^d$, then h is a regularizer.
(iii) If h is strongly convex, then h is a regularizer.

The definition of FTRL algorithms involves the gradient of h^*, whose existence is given by the following proposition.

Proposition 4.3 (Differentiability of h^*)**.** *Let h be a regularizer. Then, h^* is differentiable on $\mathbb{R}^d$, ∇h^* takes values in $\operatorname{dom} \partial h \subset \operatorname{dom} h$, and*

$$\nabla h^*(y) = \operatorname*{arg\,max}_{x \in \mathbb{R}^d} \{\langle y, x\rangle - h(x)\}.$$

We can now give the definition of a general FTRL-based approachability algorithm. Let $\mathcal{X}, \mathcal{K} \subset \mathcal{C}^\circ$ be nonempty convex sets with $\mathcal{K}$ being a cone, $\mathfrak{a}\colon \mathcal{K} \to \mathcal{A}$ a relaxed oracle, and $(h_t)_{t\geqslant 1}$ be a sequence of regularizers such that for all $t \geqslant 1$,

$$\operatorname{dom} h_t = \mathcal{X} \quad \text{and} \quad \operatorname{dom} \partial h_t \subset \mathcal{X} \cap \mathcal{K}.$$

Then, the associated algorithm is given by

$$\text{(FTRL)} \qquad a_t = \mathfrak{a}\left(\nabla h_t^*\left(\sum_{s=1}^{t-1} r_s\right)\right), \quad t \geqslant 1,$$

where for all $t \geqslant 1$, $r_t = \rho(a_t, b_t)$ with $(b_t)_{t\geqslant 1}$ being the actions in $\mathcal{B}$ chosen by Nature. The above is well-defined because $\mathfrak{a}$ is defined on $\mathcal{K}$, and thanks to Proposition 4.3, ∇h_t^* takes values in $\operatorname{dom} \partial h_t$ which is a subset of $\mathcal{K}$ by assumption. Such an algorithm will be said to be defined by a *constant regularizer* in cases where $(h_t)_{t\geqslant 1}$ is a *constant sequence*.

Example 4.4 (Blackwell's algorithm)**.** Let $\mathfrak{a}$ be an oracle. Blackwell's algorithm, in the context of a closed convex cone target, is defined as

$$a_t = \mathfrak{a}\left(\sum_{s=1}^{t-1} r_s - \Pi_{\mathcal{C}}\left(\sum_{s=1}^{t-1} r_s\right)\right), \quad t \geqslant 1, \tag{7}$$

which, using Moreau's decomposition (Theorem B.7), rewrites as

$$\begin{aligned} a_t &= \mathfrak{a}\left(\Pi_{\mathcal{C}^\circ}\left(\sum_{s=1}^{t-1} r_s\right)\right) = \mathfrak{a}\left(\operatorname*{arg\,min}_{x \in \mathcal{C}^\circ} \left\| x - \sum_{s=1}^{t-1} r_s \right\|_2^2\right) \\ &= \mathfrak{a}\left(\operatorname*{arg\,max}_{x \in \mathbb{R}^d} \left\{ \left\langle \sum_{s=1}^{t-1} r_s, x \right\rangle - \frac{1}{2}\|x\|_2^2 - I_{\mathcal{C}^\circ}(x) \right\}\right) \\ &= \mathfrak{a}\left(\nabla h^*\left(\sum_{s=1}^{t-1} r_s\right)\right), \end{aligned}$$

where the regularizer is $h := \frac{1}{2}\|\cdot\|_2^2 + I_{\mathcal{C}^\circ}$, function $I_{\mathcal{C}^\circ}$ being the convex indicator function of set $\mathcal{C}^\circ$. Therefore, Blackwell's algorithm is indeed a special case of (FTRL)[12].

Example 4.5 (ℓ^p regularization)**.** Let $p > 1$. An important special case is the algorithm corresponding to constant regularizer $h = \frac{1}{2}\|\cdot\|_p^2 + I_{\mathcal{C}^\circ}$. In the context of regret minimization on the simplex (see Section 6), this will recover the *Polynomially Weighted Average Forecaster* (Cesa-Bianchi and Lugosi, 2003).

Theorem 4.6. *Let $T \geqslant 1$ and $x \in \mathcal{X}$.*

(i) Then,

$$\left\langle \sum_{t=1}^{T} r_t, x \right\rangle \leqslant \inf_{\lambda > 0} \frac{h_{T+1}(\lambda x) + h_{T+1}^*\left(\sum_{t=1}^{T} r_t\right)}{\lambda}.$$

(ii) If h_{T+1} is positively homogeneous of degree 2,

$$\left\langle \sum_{t=1}^{T} r_t, x \right\rangle \leqslant 2\sqrt{h_{T+1}(x) \cdot h_{T+1}^*\left(\sum_{t=1}^{T} r_t\right)}.$$

Moreover, let $\|\cdot\|$ be a norm and B the associated closed unit ball. If $\mathcal{X} = \mathcal{C}^\circ$,

$$\operatorname{dist}_{\|\cdot\|_*}\left(\frac{1}{T}\sum_{t=1}^{T} r_t,\ \mathcal{C}\right) \leqslant \frac{2}{T}\sqrt{\left(\sup_{x \in \mathcal{C}^\circ \cap B} h_{T+1}(x)\right) \cdot h_{T+1}^*\left(\sum_{t=1}^{T} r_t\right)}.$$

In this distance bound, we use the convention $0 \cdot (+\infty) = 0$.

(iii) If for all $t \geqslant 1$, $h_{t+1} \geqslant h_t$ and h_t is K_t-strongly convex for a norm $\|\cdot\|_{(t)}$, then

$$h_{T+1}^*\left(\sum_{t=1}^{T} r_t\right) \leqslant h_1^*(0) + \sum_{t=1}^{T} \frac{\|r_t\|_{(t*)}^2}{2K_t},$$

where $\|\cdot\|_{(t*)}$ denotes the dual norm of $\|\cdot\|_{(t)}$.

[12] See Kwon (2016, 2021) for a slightly different connection between Blackwell's algorithm and FTRL. Independently, in the special case of regret minimization on the simplex, the connection between RM and FTRL was noted in an unpublished document by Niccolò Cesa-Bianchi dated October 14, 2015.

Proof. Our general approach combines classical FTRL regret bounds (see, e.g., Orabona, 2013, Lemma 1 and its proof, and Proposition C.2 for increasing regularizers) with Lemma 3.2 above. We here present an alternative proof. Using the Fenchel–Young inequality (Proposition A.1), for all $\lambda > 0$,

$$\left\langle \sum_{t=1}^{T} r_t, \lambda x \right\rangle \leqslant h_{T+1}(\lambda x) + h_{T+1}^* \left(\sum_{t=1}^{T} r_t \right),$$

where $h_{T+1}(\lambda x)$ may be infinite. Dividing by λ and taking the infimum gives (i).

We now turn to (ii). If h_{T+1} is positively homogeneous of degree 2, for all $x \in \mathcal{X}$, $h_{T+1}(\lambda x) = \lambda^2 h_{T+1}(x)$. Moreover, both h_{T+1} and h_{T+1}^* are nonnegative by Proposition A.7. Therefore, if either $h_{T+1}(x)$ or $h_{T+1}^* \left(\sum_{t=1}^{T} r_t \right)$ is zero,

$$\inf_{\lambda > 0} \left\{ \lambda \cdot h_{T+1}(x) + \frac{h_{T+1}^* \left(\sum_{t=1}^{T} r_t \right)}{\lambda} \right\} = 0,$$

and otherwise it is attained for $\lambda = \sqrt{h_{T+1}^* \left(\sum_{t=1}^{T} r_t \right) / h_{T+1}(x)}$ and equals

$$2 \sqrt{h_{T+1}(x) \cdot h_{T+1}^* \left(\sum_{t=1}^{T} r_t \right)}.$$

In the case where the common domain $\mathcal{X}$ of the regularizers is $\mathcal{C}^\circ$, the upper bound on the distance to the target is obtained by taking the maximum over $x \in \mathcal{C}^\circ \cap B$ and applying Lemma 3.1.

Regarding (iii), first note that assumption $h_{t+1} \geqslant h_t$ implies $h_{t+1}^* \leqslant h_t^*$. Besides, remember that the algorithm writes

$$R_t = \sum_{s=1}^{t} r_s, \quad a_{t+1} = \mathfrak{a}(\nabla h_{t+1}^*(R_t)), \quad t \geqslant 0.$$

By definition of relaxed oracle, $\langle r_t, \nabla h_t^*(R_{t-1}) \rangle \leqslant 0$ for all $t \geqslant 1$. Using also $h_{t+1}^* \leqslant h_t^*$ gives

$$\begin{aligned}
h_{T+1}^*(R_T) &= h_1^*(0) + \sum_{t=1}^{T} \left(h_{t+1}^*(R_t) - h_t^*(R_{t-1}) \right) \\
&\leqslant h_1^*(0) + \sum_{t=1}^{T} \left(h_t^*(R_t) - h_t^*(R_{t-1}) - \langle r_t, \nabla h_t^*(R_{t-1}) \rangle \right) \\
&= h_1^*(0) + \sum_{t=1}^{T} D_{h_t^*}(R_t, R_{t-1}) \\
&\leqslant h_1^*(0) + \sum_{t=1}^{T} \frac{\|r_t\|_{(t*)}^2}{2K_t},
\end{aligned}$$

where $D_{h_t^*}$ denotes the Bregman divergence associated with h_t^*, and where the last inequality simply applies the consequence of strong convexity from Proposition A.5. ☐

Combining assumptions (ii) and (iii) of the above theorem in the case of a constant 2-positively homogeneous regularizer h that is K-strongly convex with respect to a given norm $\|\cdot\|$, the guarantee writes

$$\left\langle \sum_{t=1}^{T} r_t, x \right\rangle \leqslant \sqrt{\frac{2h(x)}{K} \sum_{t=1}^{T} \|r_t\|_*^2}.$$

This guarantee requires no prior knowledge on the norm of the outcome vectors, nor on the time-horizon T, and adapts to the realized sum of the squared outcome vector norms. In the context of approachability, this solves the classical *two regret bound terms that need to be balanced through tuning* issue, and provides a generic construction of parameter-free and adaptive algorithms. Bounds that scale with the square root of the sum of squared norms of the realized outcome vectors have appeared in Cesa-Bianchi and Lugosi (2003, Corollary 1) for the Polynomially Weighted Average Forecaster in the context of regret minimization on the simplex and in Grand-Clément and Kroer (2024) and Kwon et al. (2024, Theorem 3.1) for Blackwell's algorithm with conic targets. The more general bound (ii) retains $h^*(R_T)$, which can be much smaller than its upper bound $\frac{1}{2K}\sum_{t=1}^{T}\|r_t\|_*^2$ and thus provides a somewhat *stronger adaptivity*. This aspect will be discussed in particular in Sections 8 and 9 in the context of unconstrained OLO where $h_{T+1}^*(R_T)$ will be, in some cases, smaller by a factor of nearly $1/T$ than its explicit upper bound.

This section emphasizes 2-positively homogeneous regularizers, which yield parameter-free and adaptive guarantees. In Sections 8 and 9, we will also consider regularizers in which only some terms are 2-positively homogeneous, while still achieving strong adaptivity for unconstrained OLO.

In the case of Blackwell's algorithm (presented in Example 4.4 above), the corresponding regularizer $\frac{1}{2}\|\cdot\|_2^2 + I_{\mathcal{C}^\circ}$ has $\mathcal{C}^\circ$ as domain, is indeed positively homogeneous of degree 2, and is 1-strongly convex for $\|\cdot\|_2$ (see Proposition A.6). The following corollary follows.

Corollary 4.7 (Blackwell's algorithm)**.** *Assume that* $\mathfrak{a}$ *is an oracle and let* $T \geqslant 1$.

(i) For all $x \in \mathcal{C}^\circ$*, Blackwell's algorithm* (7) *guarantees,*

$$\left\langle \frac{1}{T}\sum_{t=1}^{T} r_t, x \right\rangle \leqslant \frac{\|x\|_2}{T} \left\| \Pi_{\mathcal{C}^\circ}\left(\sum_{t=1}^{T} r_t\right)\right\|_2 \leqslant \frac{\|x\|_2}{T}\sqrt{\sum_{t=1}^{T}\|r_t\|_2^2}.$$

(ii) In particular,

$$\operatorname{dist}_2\left(\frac{1}{T}\sum_{t=1}^{T} r_t,\ \mathcal{C}\right) = \frac{1}{T}\left\| \Pi_{\mathcal{C}^\circ}\left(\sum_{t=1}^{T} r_t\right)\right\|_2 \leqslant \frac{1}{T}\sqrt{\sum_{t=1}^{T}\|r_t\|_2^2}.$$

(iii) If there exists $M > 0$ *such that* $\|r_t\|_2 \leqslant M$ *for all* $t \geqslant 1$*, then*

$$\operatorname{dist}_2\left(\frac{1}{T}\sum_{t=1}^{T} r_t,\ \mathcal{C}\right) \leqslant \frac{M}{\sqrt{T}}.$$

Proof. For $h = \frac{1}{2}\|\cdot\|_2^2 + I_{\mathcal{C}^\circ}$, Proposition C.1 and Moreau's decomposition give

$$h^*(y) = \frac{1}{2}\operatorname{dist}_2(y,\mathcal{C})^2 = \frac{1}{2}\left\|\Pi_{\mathcal{C}^\circ}(y)\right\|_2^2.$$

The first assertion follows from Theorem 4.6, since h is positively homogeneous of degree 2, is 1-strongly convex with respect to $\|\cdot\|_2$, and satisfies $h^*(0) = 0$. The equality in the second assertion also uses that $\mathcal{C}$ is a cone. The last assertion follows from $\sum_{t=1}^{T}\|r_t\|_2^2 \leqslant TM^2$. □

Corollary 4.8 (FTRL with ℓ^p regularization)**.** *Let $1 < p \leqslant 2$, $q = p/(p-1)$, and consider regularizer $h = \frac{1}{2}\|\cdot\|_p^2 + I_{\mathcal{C}^\circ}$. Write dist_q for distance in the norm $\|\cdot\|_q$. Assume that $\mathfrak{a}$ is an oracle and let $x \in \mathcal{C}^\circ$ and $T \geqslant 1$.*

(i) The (FTRL) algorithm with constant regularizer h guarantees

$$\left\langle \frac{1}{T}\sum_{t=1}^{T} r_t, x \right\rangle \leqslant \frac{\|x\|_p}{T} \operatorname{dist}_q\left(\sum_{t=1}^{T} r_t, \mathcal{C}\right) \leqslant \frac{\|x\|_p}{T}\sqrt{\frac{1}{p-1}\sum_{t=1}^{T}\|r_t\|_q^2}.$$

(ii) In particular, if $d \geqslant 2$, $q = \max\{2, 2\log d\}$ and $p = q/(q-1)$,

$$\left\langle \frac{1}{T}\sum_{t=1}^{T} r_t, x \right\rangle \leqslant \frac{\|x\|_p}{T}\sqrt{2e\log d\sum_{t=1}^{T}\|r_t\|_\infty^2}.$$

Moreover, if there exists $M > 0$ such that $\|r_t\|_\infty \leqslant M$ for all $t \geqslant 1$,

$$\left\langle \frac{1}{T}\sum_{t=1}^{T} r_t, x \right\rangle \leqslant M\|x\|_p\sqrt{\frac{2e\log d}{T}}.$$

Proof. By Proposition C.1, $h^*(y) = \frac{1}{2}\operatorname{dist}_q(y,\mathcal{C})^2$, $h^*(0) = 0$, and h is $(p-1)$-strongly convex with respect to $\|\cdot\|_p$. Combining these identities with Theorem 4.6 gives the first assertion. Besides, the norms are compared as follows. For $t \geqslant 1$,

$$\frac{\|r_t\|_q^2}{p-1} \leqslant (q-1)d^{2/q}\|r_t\|_\infty^2 = e(2\log d - 1)\|r_t\|_\infty^2 \leqslant 2e\log d\ \|r_t\|_\infty^2$$

in the case $d \geqslant 3$ and $\|r_t\|_2^2 \leqslant 2\|r_t\|_\infty^2$ when $d = 2$. The result follows. □

5. OMD-type algorithms

We now apply the conversion scheme from Section 3 to Online Mirror Descent (OMD). The most basic example of OMD is Online (Projected) Gradient Descent: we call the corresponding approachability algorithm the Greedy Blackwell algorithm, which in turn contains RM^+ and CBA^+ as special cases[13]. In addition to OMD being a central family of regret minimizers, one important motivation for converting them for approachability is the practical success of such algorithms in the context of solving games[14]. Similarly, Grand-Clément and Kroer (2024) demonstrate strong empirical performance for CBA^+ (over CBA) in their saddle-point settings (Grand-Clément and Kroer, 2024).

OMD algorithms are defined with the help of *mirror maps*. We here use the assumptions and properties from Juditsky et al. (2023). Analogously to FTRL-based algorithms, 2-positively homogeneous mirror maps will yield parameter-free and adaptive guarantees.

Throughout the section, $\mathcal{X} \subset \mathbb{R}^d$ will be a nonempty closed convex set.

Definition 5.1. Let $H\colon \mathbb{R}^d \to \mathbb{R}\cup\{+\infty\}$. Denote $\mathcal{D}_H \coloneqq \operatorname{int}\operatorname{dom} H$. H is a *mirror map* compatible with $\mathcal{X}$ if

(i) H is lower semicontinuous, strictly convex, and differentiable on $\mathcal{D}_H$,
(ii) the gradient of H takes all possible values, i.e. $\nabla H(\mathcal{D}_H) = \mathbb{R}^d$,
(iii) $\mathcal{X} \subset \operatorname{cl}\mathcal{D}_H$,
(iv) $\mathcal{X} \cap \mathcal{D}_H \neq \varnothing$.

[13] The link between RM^+ and OGD was established in Farina et al. (2021) and independently in Flaspohler et al. (2021). On the other hand, CBA^+ was explicitly defined with the help of OGD (Grand-Clément and Kroer, 2024).

[14] For extensive-form games in particular, RM^+ is a building block of CFR^+, which has shown substantial empirical improvements over vanilla CFR (which is based on RM) (Tammelin, 2014)

Denote by D_H its Bregman divergence:

$$D_H(x', x) = H(x') - H(x) - \left\langle \nabla H(x), x' - x \right\rangle, \quad x' \in \operatorname{dom} H,\ x \in \mathcal{D}_H.$$

Proposition 5.2 (Juditsky et al., 2023, Proposition 1)**.** *Let H be a mirror map compatible with $\mathcal{X}$, H^* its Legendre–Fenchel transform. Then,*

(i) $\operatorname{dom} H^* = \mathbb{R}^d$,
(ii) H^* *is differentiable on* $\mathbb{R}^d$,
(iii) $\nabla H^*(\mathbb{R}^d) = \mathcal{D}_H$,
(iv) For $x \in \mathcal{D}_H$ *and* $y \in \mathbb{R}^d$, $\nabla H^*(\nabla H(x)) = x$ *and* $\nabla H(\nabla H^*(y)) = y$.

Proposition 5.3 (OMD iteration is well-defined, Juditsky et al., 2023, Theorem 1)**.** *Let H be a mirror map compatible with $\mathcal{X}$, $x \in \mathcal{D}_H$, $u \in \mathbb{R}^d$ and consider*

$$x' = \operatorname*{arg\,max}_{x'' \in \mathcal{X}} \left\{ \left\langle u, x'' \right\rangle - D_H(x'', x) \right\}.$$

Then, $x' \in \mathcal{X} \cap \mathcal{D}_H$ and can also be written

$$\begin{aligned} x' &= \operatorname*{arg\,max}_{x'' \in \mathcal{X}} \left\{ \left\langle \nabla H(x) + u, x'' \right\rangle - H(x'') \right\} \\ &= \operatorname*{arg\,min}_{x'' \in \mathcal{X}} D_H(x'', \nabla H^*(\nabla H(x) + u)). \end{aligned}$$

Let $\mathcal{K}$ be dense convex subcone of $\mathcal{C}^\circ$ and $\mathfrak{a}\colon \mathcal{K} \to \mathcal{A}$ be a relaxed oracle. Let H be a mirror map compatible with $\mathcal{X}$ and such that $\mathcal{X} \cap \mathcal{D}_H \subset \mathcal{K}$. The associated approachability algorithm is defined as follows. Let $x_1 \in \mathcal{X} \cap \mathcal{D}_H$, $a_1 = \mathfrak{a}(x_1)$ and for all $t \geqslant 1$,

$$\text{(OMD)} \qquad x_{t+1} = \operatorname*{arg\,max}_{x \in \mathcal{X}} \left\{ \langle \nabla H(x_t) + r_t, x \rangle - H(x) \right\} \quad \text{and} \quad a_{t+1} = \mathfrak{a}(x_{t+1}).$$

In the case of the Online Gradient Descent (OGD) regret minimizer, we call the corresponding approachability algorithm the *Greedy Blackwell* algorithm. It is an extension of both RM^+ and CBA^+.

Example 5.4 (Greedy Blackwell algorithm)**.** Assume $\mathcal{X} = \mathcal{C}^\circ$ and that $\mathfrak{a}\colon \mathcal{C}^\circ \to \mathcal{A}$ is an oracle. Then the Greedy Blackwell algorithm writes $x_1 = 0$ and for $t \geqslant 1$,

$$x_{t+1} = \Pi_{\mathcal{C}^\circ}(x_t + r_t) \qquad \text{and} \qquad a_{t+1} = \mathfrak{a}(x_{t+1}), \tag{8}$$

which can be rewritten, using $H = \frac{1}{2} \left\| \cdot \right\|_2^2$ as

$$\begin{aligned} x_{t+1} &= \operatorname*{arg\,min}_{x \in \mathcal{C}^\circ} \frac{1}{2} \left\| x_t + r_t - x \right\|_2^2 = \operatorname*{arg\,min}_{x \in \mathcal{C}^\circ} \left\{ - \langle x_t + r_t, x \rangle + \frac{1}{2} \left\| x \right\|_2^2 \right\} \\ &= \operatorname*{arg\,max}_{x \in \mathcal{C}^\circ} \left\{ \langle \nabla H(x_t) + r_t, x \rangle - H(x) \right\}, \end{aligned}$$

which corresponds to (OMD).

Although it is possible to consider OMD with time-dependent mirror maps (analogously to FTRL with time-dependent regularizers), the corresponding regret bounds involve Bregman divergences between points x_t and comparators $x \in \mathcal{X} \cap \operatorname{dom} H$, for which estimates are rarely available when the comparator-domain $\mathcal{X} \cap \operatorname{dom} H$ is unbounded. In particular, we will use 2-positively homogeneous mirror maps on cone comparator domains, which need not be bounded. We thus restrict, for simplicity, to mirror maps that are constant over time.

Theorem 5.5. *Let $x \in \mathcal{C}^\circ$ and $T \geqslant 1$.*

(i) Then,

$$\left\langle \sum_{t=1}^{T} r_t, x \right\rangle \leqslant \inf_{\substack{\lambda>0 \\ \lambda x \in \mathcal{X} \cap \operatorname{dom} H}} \frac{D_H(\lambda x, x_1) + \sum_{t=1}^{T} D_{H^*}(\nabla H(x_t) + r_t, \nabla H(x_t))}{\lambda},$$

where D_H and D_{H^*} denote Bregman divergences, and $\inf \varnothing = +\infty$ by convention.

(ii) Suppose that $\mathcal{X} \cap \operatorname{dom} H$ is a convex cone, that H is positively homogeneous of degree 2, and that $x_1 = 0$. Then, for $x \in \mathcal{X} \cap \operatorname{dom} H$,

$$\left\langle \sum_{t=1}^{T} r_t, x \right\rangle \leqslant 2\sqrt{H(x)\left(\sum_{t=1}^{T} D_{H^*}(\nabla H(x_t) + r_t, \nabla H(x_t))\right)}.$$

Moreover, let $\|\cdot\|$ be a norm and B the associated closed unit ball. If $\mathcal{X} = \mathcal{C}^\circ$,

$$\operatorname{dist}_{\|\cdot\|_*}\left(\frac{1}{T}\sum_{t=1}^{T} r_t,\ \mathcal{C}\right) \leqslant \frac{2}{T}\sqrt{\left(\max_{x \in \mathcal{C}^\circ \cap B} H(x)\right)\sum_{t=1}^{T} D_{H^*}(\nabla H(x_t) + r_t, \nabla H(x_t))}.$$

(iii) Let $K > 0$. If H is K-strongly convex for a norm $\|\cdot\|$, then for all $t \geqslant 1$,

$$D_{H^*}(\nabla H(x_t) + r_t, \nabla H(x_t)) \leqslant \frac{1}{2K}\|r_t\|_*^2.$$

Proof. Let $x \in \mathcal{C}^\circ$ and $\lambda > 0$. Applying Lemma 3.2 together with the regret bound for OMD from Proposition D.2 and the properties of H gives the general guarantee (i). We now turn to (ii). Under these assumptions, according to Proposition A.7, $H(0) = 0$ and $\nabla H(0) = 0$, so for all $x \in \mathcal{X}$,

$$D_H(\lambda x, 0) = \lambda^2 H(x).$$

Taking the infimum over $\lambda > 0$ then gives the bound. The upper bound (iii) on the Bregman divergences follows from Proposition A.5. □

Corollary 5.6 (Greedy Blackwell algorithm)**.** *Assume $\mathcal{X} = \mathcal{C}^\circ$ and that $\mathfrak{a}\colon \mathcal{C}^\circ \to \mathcal{A}$ is an oracle. Let $H = \frac{1}{2}\|\cdot\|_2^2$ on $\mathbb{R}^d$ and $T \geqslant 1$. The Greedy Blackwell algorithm from Example 5.4 guarantees the following bounds.*

(i) For all $x \in \mathcal{C}^\circ$,

$$\left\langle \frac{1}{T}\sum_{t=1}^{T} r_t, x \right\rangle \leqslant \frac{\|x\|_2}{T}\sqrt{2\sum_{t=1}^{T} D_{H^*}(x_t + r_t, x_t)} = \frac{\|x\|_2}{T}\sqrt{\sum_{t=1}^{T}\|r_t\|_2^2}.$$

(ii) In particular,

$$\operatorname{dist}_2\left(\frac{1}{T}\sum_{t=1}^{T} r_t,\ \mathcal{C}\right) \leqslant \frac{1}{T}\sqrt{2\sum_{t=1}^{T} D_{H^*}(x_t + r_t, x_t)} = \frac{1}{T}\sqrt{\sum_{t=1}^{T}\|r_t\|_2^2}.$$

(iii) If there exists $M > 0$ such that $\|r_t\|_2 \leqslant M$ for all $t \geqslant 1$, then

$$\operatorname{dist}_2\left(\frac{1}{T}\sum_{t=1}^{T} r_t,\ \mathcal{C}\right) \leqslant \frac{M}{\sqrt{T}}.$$

Proof. Apply Theorem 5.5 with $\mathcal{X} = \mathcal{C}^\circ$, $x_1 = 0$ and the positively homogeneous mirror map $H = \frac{1}{2}\|\cdot\|_2^2$, which corresponds to the Greedy Blackwell algorithm as explained in Example 5.4. Since $H^* = H$ and $\nabla H(x_t) = x_t$,

$$D_{H^*}(x_t + r_t, x_t) = \frac{1}{2}\|r_t\|_2^2.$$

The distance bound follows by taking the supremum over $x \in \mathcal{C}^\circ$ with $\|x\|_2 \leqslant 1$ and applying Lemma 3.1. □

Corollary 5.7 (OMD with ℓ^p mirror map)**.** *Let $1 < p \leqslant 2$, $q = p/(p-1)$ and $H = \frac{1}{2}\|\cdot\|_p^2$ on $\mathbb{R}^d$. Assume $\mathcal{X} = \mathcal{C}^\circ$ and that $\mathfrak{a}\colon \mathcal{C}^\circ \to \mathcal{A}$ is an oracle. Let $T \geqslant 1$. The (OMD) algorithm with mirror map H and $x_1 = 0$ guarantees the following bounds.*

(i) For all $x \in \mathcal{C}^\circ$,

$$\left\langle \frac{1}{T}\sum_{t=1}^T r_t, x \right\rangle \leqslant \frac{\|x\|_p}{T}\sqrt{2\sum_{t=1}^T D_{H^*}\left(\nabla H(x_t) + r_t, \nabla H(x_t)\right)} \leqslant \frac{\|x\|_p}{T}\sqrt{\frac{1}{p-1}\sum_{t=1}^T \|r_t\|_q^2}.$$

(ii) In particular,

$$\operatorname{dist}_q\left(\frac{1}{T}\sum_{t=1}^T r_t,\ \mathcal{C}\right) \leqslant \frac{1}{T}\sqrt{2\sum_{t=1}^T D_{H^*}\left(\nabla H(x_t) + r_t, \nabla H(x_t)\right)} \leqslant \frac{1}{T}\sqrt{\frac{1}{p-1}\sum_{t=1}^T \|r_t\|_q^2}.$$

(iii) If $d \geqslant 2$, $q = \max\{2, 2\log d\}$ and $p = q/(q-1)$, then, for all $x \in \mathcal{C}^\circ$,

$$\left\langle \frac{1}{T}\sum_{t=1}^T r_t, x \right\rangle \leqslant \frac{\|x\|_p}{T}\sqrt{2e\log d\sum_{t=1}^T \|r_t\|_\infty^2},$$

and

$$\operatorname{dist}_q\left(\frac{1}{T}\sum_{t=1}^T r_t,\ \mathcal{C}\right) \leqslant \frac{1}{T}\sqrt{2e\log d\sum_{t=1}^T \|r_t\|_\infty^2}.$$

Moreover, if there exists $M > 0$ such that $\|r_t\|_\infty \leqslant M$ for all $t \geqslant 1$, then

$$\left\langle \frac{1}{T}\sum_{t=1}^T r_t, x \right\rangle \leqslant M\|x\|_p\sqrt{\frac{2e\log d}{T}}$$

for all $x \in \mathcal{C}^\circ$, and

$$\operatorname{dist}_q\left(\frac{1}{T}\sum_{t=1}^T r_t,\ \mathcal{C}\right) \leqslant M\sqrt{\frac{2e\log d}{T}}.$$

Proof. The mirror map H is positively homogeneous of degree 2 and $(p-1)$-strongly convex with respect to $\|\cdot\|_p$ by Proposition D.1, and $H^* = \frac{1}{2}\|\cdot\|_q^2$. Theorem 5.5 gives the first bound and the upper bound on its sum of Bregman divergences. Taking the supremum over $x \in \mathcal{C}^\circ$ with $\|x\|_p \leqslant 1$ gives the distance bound by Lemma 3.1. For the logarithmic bounds, the norm comparison in the proof of Corollary 4.8 gives

$$\frac{\|r_t\|_q^2}{p-1} \leqslant 2e\log d\ \|r_t\|_\infty^2.$$

□

6. Regret minimization on the simplex

In this section, we illustrate the unifying character of our approach in the context of regret minimization on the simplex. We first recall the reduction, due to Hart and Mas-Colell (2000), of OLO on the simplex to an approachability problem. The Blackwell and Greedy Blackwell algorithms then yield the celebrated Regret Matching (RM) (Hart and Mas-Colell, 2001) and Regret Matching$^+$ (RM$^+$) (Tammelin, 2014) algorithms respectively. The ℓ^p-FTRL and ℓ^p-OMD algorithms from Corollaries 4.8 and 5.7 yield the Polynomially Weighted Average Forecaster (Cesa-Bianchi and Lugosi, 2003) and Polynomial RM$^+$ (D'Orazio, 2020; Flaspohler et al., 2021). We call these algorithms ℓ^p-RM and ℓ^p-RM$^+$, respectively, to emphasize their relation to RM and RM$^+$. Parameter-free and adaptive regret guarantees are recovered as immediate corollaries. For $1 < p \leqslant 2$, we recover the regret bound of Cesa-Bianchi and Lugosi (2003, Corollary 1) in Proposition 6.2. For ℓ^p-RM$^+$, we sharpen in Proposition 6.3 the regret bounds that follow from D'Orazio and Huang (2021, Theorem 4 and equation (24)) and Flaspohler et al. (2021, Lemmas 7–8 and Theorem 6)[15].

Denote $\mathbb{1} = (1, \dots, 1) \in \mathbb{R}^d$. In this section, $(a_t)_{t\geqslant 1}$ and a denote actions and comparators in the OLO problem on the simplex. As always, the payoff vectors are denoted by $(u_t)_{t\geqslant 1}$.

Proposition 6.1 (Orthant reduction, Hart and Mas-Colell, 2000)**.** *Consider action sets $\mathcal{A} = \Delta_d$, $\mathcal{B} = \mathbb{R}^d$, and outcome function $\rho\colon (a, u) \mapsto u - \langle u, a\rangle\, \mathbb{1}$, and target set $\mathcal{C} = \mathbb{R}^d_-$. Let $a_1 \in \Delta_d$.*

(i) The following map is an oracle.

$$\mathfrak{a}(x) = \begin{cases} x/\|x\|_1 & \text{if } x \neq 0 \\ a_1 & \text{if } x = 0, \end{cases} \qquad x \in \mathbb{R}^d_+.$$

(ii) Let $(x_t)_{t\geqslant 1}$ and $(u_t)_{t\geqslant 1}$ sequences in $\mathbb{R}^d_+$ and $\mathbb{R}^d$ respectively, let $a_t = \mathfrak{a}(x_t)$ and $r_t = \rho(a_t, u_t)$ for all $t \geqslant 1$. Then, for all $T \geqslant 1$, and $a \in \Delta_d$

$$\sum_{t=1}^T \langle u_t, a - a_t\rangle = \left\langle \sum_{t=1}^T r_t, a \right\rangle.$$

The proof is in Appendix E.1.

For $y \in \mathbb{R}^d$, denote y_+ the corresponding vector obtained by taking the positive part in each component, in other words,

$$y_+ = \Big(\max(y_i, 0)\Big)_{1\leqslant i\leqslant d}.$$

Proposition 6.2 (ℓ^p-RM)**.** *Let $1 < p \leqslant 2$, $q = p/(p-1)$, $(u_t)_{t\geqslant 1}$ be a sequence in $\mathbb{R}^d$, and $a_1 \in \Delta_d$. Set $R_0 = 0$ and define recursively, for all $t \geqslant 1$,*

$$\begin{aligned} a_t &= \begin{cases} \dfrac{(R_{t-1})_+^{q-1}}{\left\|(R_{t-1})_+^{q-1}\right\|_1} & \text{if } (R_{t-1})_+ \neq 0, \\ a_1 & \text{otherwise,} \end{cases} \\ r_t &= u_t - \langle u_t, a_t\rangle\, \mathbb{1}, \\ R_t &= R_{t-1} + r_t, \end{aligned}$$

where powers of vectors are taken component-wise. Then for all $T \geqslant 1$,

$$\max_{a\in\Delta_d} \sum_{t=1}^T \langle u_t, a - a_t\rangle \leqslant \|(R_T)_+\|_q \leqslant \sqrt{(q-1)\sum_{t=1}^T \|u_t - \langle u_t, a_t\rangle\, \mathbb{1}\|_q^2}.$$

[15] without delay or optimistic predictions

Proof. Let $H = \frac{1}{2}\|\cdot\|_p^2$ and $h = H + I_{\mathbb{R}^d_+}$. Proposition C.1 shows that applying the oracle from Proposition 6.1 to $\nabla h^*(R_{t-1})$ gives the stated action, since the scalar factor cancels. The same proposition gives $h^*(R_T) = \frac{1}{2}\|(R_T)_+\|_q^2$ and $h^*(0) = 0$. By Proposition A.6, h is $(p-1)$-strongly convex for $1 < p \leqslant 2$. Theorem 4.6, inequality (iii), then yields

$$\frac{1}{2}\|(R_T)_+\|_q^2 \leqslant \frac{q-1}{2}\sum_{t=1}^T \|r_t\|_q^2 .$$

By Proposition 6.1,

$$\max_{a\in\Delta_d}\sum_{t=1}^T \langle u_t, a - a_t\rangle = \max_{1\leqslant i\leqslant d}(R_T)_i \leqslant \|(R_T)_+\|_q .$$

Combining the two inequalities proves the regret bound. □

Proposition 6.3 (ℓ^p-RM$^+$)**.** *Let $1 < p \leqslant 2$, $q = p/(p-1)$, $(u_t)_{t\geqslant 1}$ be a sequence in $\mathbb{R}^d$, and $a_1 \in \Delta_d$. Set $x_1 = 0$ and define recursively, for all $t \geqslant 1$,*

$$a_t = \begin{cases} \dfrac{x_t^{q-1}}{\left\|x_t^{q-1}\right\|_1} & \text{if } x_t \neq 0, \\ a_1 & \text{otherwise,} \end{cases}$$
$$r_t = u_t - \langle u_t, a_t\rangle \mathbb{1},$$
$$x_{t+1} = (x_t + r_t)_+.$$

Then,

$$\delta_t := \|x_t + r_t\|_q^2 - \|(x_t + r_t)_+\|_q^2 \geqslant 0,$$

and for all $T \geqslant 1$,

$$\max_{a\in\Delta_d}\sum_{t=1}^T \langle u_t, a - a_t\rangle \leqslant \|x_{T+1}\|_q \leqslant \sqrt{\sum_{t=1}^T \Big((q-1)\|u_t - \langle u_t, a_t\rangle \mathbb{1}\|_q^2 - \delta_t\Big)}.$$

Proof. We can prove by induction that $x_{T+1} \geqslant \sum_{t=1}^T r_t$ coordinate-wise. Proposition 6.1 therefore gives, for each $a \in \Delta_d$,

$$\sum_{t=1}^T \langle u_t, a - a_t\rangle = \left\langle \sum_{t=1}^T r_t, a \right\rangle \leqslant \langle x_{T+1}, a\rangle \leqslant \|x_{T+1}\|_q .$$

Let $H = \frac{1}{2}\|\cdot\|_p^2$ on $\mathbb{R}^d$ and $h = H + I_{\mathbb{R}^d_+}$. Introduce the auxiliary primal iterates $\xi_t = \nabla H^*(x_t)$. By Propositions C.1 and D.1, $\nabla h^*(x_t + r_t) = \nabla H^*((x_t + r_t)_+) = \xi_{t+1}$, and Proposition 5.2 gives $\nabla H(\xi_t) = x_t$. Applying the oracle from Proposition 6.1 to ξ_t gives a_t, including when $\xi_t = 0$, so ξ_t are the OMD iterates in (OMD). Also $\langle \nabla H^*(x_t), r_t\rangle = 0$, since ξ_t is a nonnegative multiple of a_t (or is zero). Proposition A.6 and Theorem 5.5, inequality (iii), give

$$D_{H^*}(x_t + r_t, x_t) \leqslant \frac{q-1}{2}\|r_t\|_q^2 .$$

By the definition of δ_t and the Bregman divergence,

$$\|x_{t+1}\|_q^2 - \|x_t\|_q^2 = 2D_{H^*}(x_t + r_t, x_t) - \delta_t.$$

Summing proves the state bound. □

Corollary 6.4 (Dimension-dependent bounds for ℓ^p-RM and ℓ^p-RM$^+$)**.** *Let $1 < p \leqslant 2$, $q = p/(p-1)$, and consider either algorithm. For all $T \geqslant 1$, the following bounds hold.*

(i) If $\|u_t\|_\infty \leqslant L$ *for all* $t \geqslant 1$*, then*

$$\max_{a \in \Delta_d} \sum_{t=1}^{T} \langle u_t, a - a_t \rangle \leqslant 2Ld^{1/q}\sqrt{(q-1)T}.$$

(ii) For $d \geqslant 2$*, choosing* $q = \max\{2, 2\log d\}$ *and* $p = q/(q-1)$ *gives*

$$\max_{a \in \Delta_d} \sum_{t=1}^{T} \langle u_t, a - a_t \rangle \leqslant \sqrt{2e \log d \sum_{t=1}^{T} \|u_t - \langle u_t, a_t \rangle \mathbb{1}\|_\infty^2}.$$

Under the assumption of assertion (i), this bound is at most $2L\sqrt{2eT\log d}$*.*

Proof. The first bound follows from $\|u_t - \langle u_t, a_t \rangle \mathbb{1}\|_q \leqslant d^{1/q} \|u_t - \langle u_t, a_t \rangle \mathbb{1}\|_\infty \leqslant 2Ld^{1/q}$, dropping the nonnegative deficit for RM^+. Then apply the norm comparison in Corollary 4.8 to the regret bounds above. The last bound is the specialization of assertion (i) to this choice of q. □

7. From approachability to unconstrained regret minimization

We now turn to unconstrained OLO. We first describe the regret bounds sought in this context and compare the algorithms constructed in Sections 8 and 9 with the state of the art. We then recall *Blackwell's reduction* of OLO on *convex compact domains* to approachability and extend it to *unconstrained* OLO using the relaxed oracles introduced in Section 2. The next two sections apply this reduction with carefully chosen regularizers.

Until Blackwell's reduction is introduced in Section 7.2, we use $(x_t)_{t\geqslant 1}$ and x for actions and comparators. From that point onward, as in the previous section, $(a_t)_{t\geqslant 1}$ and a denote actions and comparators in the OLO problem of interest, while $(x_t)_{t\geqslant 1}$ and x denote those in the auxiliary OLO problem used to define the approachability algorithm. As always, the payoff vectors are denoted by $(u_t)_{t\geqslant 1}$.

7.1. Unconstrained OLO.

As mentioned in Section 1, for a given time-horizon $T \geqslant 1$, and a given comparator $x \in \mathbb{R}^d$, unconstrained OGD with initial point $x_1 = 0$ and *tuned with hindsight* achieves the following regret bound

$$\sum_{t=1}^{T} \langle u_t, x - x_t \rangle \leqslant \|x\|_2 \sqrt{\sum_{t=1}^{T} \|u_t\|_2^2}.$$

The Holy Grail of parameter-free regret minimization is to achieve a similar regret bound for all time horizons and all comparators simultaneously, with no hindsight and no parameter to tune. This ideal goal is not achievable exactly: for comparator $x = 0$ for instance, the above exact guarantee can only hold for $x_t = 0$ (for all $t \geqslant 1$). But this easily implies that there exist unit vectors such that the regret grows linearly in T, whereas the regret bound is expected to grow as $\sqrt{T}$. This indicates that any meaningful algorithm must in particular have some tradeoff regarding the regret with respect to the origin. We restrict this discussion to regret bounds that depend only on T, $\|x\|_2$, $V_T \coloneqq \sum_{t=1}^{T} \|u_t\|_2^2$, $G_T \coloneqq \max_{1\leqslant t\leqslant T} \|u_t\|_2$, and possibly some user-chosen parameters.

An important line of work (Streeter and McMahan, 2012; Orabona, 2013, 2014; McMahan and Orabona, 2014; Orabona and Pal, 2016; Foster et al., 2017, 2018) focuses on the simpler (yet already very difficult) case where payoff vectors have Euclidean norm at most 1. Most bounds in this context depend on $\|x\|_2$, T, and possibly a user-chosen parameter $\varepsilon > 0$, which corresponds to the maximum regret that the algorithm allows at the origin. In the present work, we focus on

bounds that explicitly depend on $V_T \coloneqq \sum_{t=1}^{T} \|u_t\|_2^2$ instead of simply T. The state-of-the-art such guarantee is

$$O\left(\varepsilon + \|x\|_2 \left(\sqrt{(1+V_T)\log\left(1+\|x\|_2\sqrt{1+V_T}/\varepsilon\right)} + \log\left(1+\|x\|_2\sqrt{1+V_T}/\varepsilon\right)\right)\right),$$

where O ignores universal multiplicative constants and $\log\log$ factors, and is achieved by Jacobsen and Cutkosky (2022). Alternatively, Section 7 of the 2024 arXiv version of the same paper[16] yields a regret bound which, at the origin, may grow with V_T:

$$O\left(\|x\|_2 \log\left(1+\|x\|_2\right) + \sqrt{V_T} + \|x\|_2\sqrt{(1+V_T)\log\left(1+\|x\|_2\right)}\right).$$

Our approachability-based algorithm in Section 8 below provides the following slightly improved bound:

$$O\left(\|x\|_2 \log\left(1+\|x\|_2\right) + \sqrt{V_T} + \|x\|_2\sqrt{V_T\log\left(1+\|x\|_2\right)}\right).$$

The more challenging setting is where arbitrary payoff vectors are allowed, with no prior knowledge on their magnitude. Two types of guarantees have been proposed in the literature. The first (Cutkosky and Boahen, 2017; Mhammedi and Koolen, 2020) is of the form

$$\tilde{O}\left(G_T\left(\sqrt{T} + \|x\|_2^3\right) + \|x\|_2\sqrt{V_T}\right),$$

where notation $\tilde{O}$ considers the regime where T, $\|x\|_2$, V_T and G_T are large and ignores universal multiplicative constants and quantities that are logarithmic in the former quantities. In this work, we focus on the other type, that was proposed by Cutkosky and Mhammedi (2024) and that is of the form

$$\tilde{O}\left(G_T^2 + \|x\|_2^2 + \|x\|_2\sqrt{V_T}\right).$$

More precisely, with logarithms, Cutkosky and Mhammedi (2024) provide a bound that can be summarized as[17]

$$O\left(\|x\|_2^2\log(1+\|x\|_2) + G_T^2\log(1+G_T) + \sqrt{1+V_T} + \|x\|_2\sqrt{V_T\log(1+\|x\|_2)}\right).$$

Our approachability-based algorithm in Section 9 below does not uniformly improve the above guarantee but achieves the similar

$$O\left(\|x\|_2^2 + G_T^2(1+\log(1+V_T/G_T)) + \left(1+\|x\|_2\sqrt{\log(1+\|x\|_2)}\right)\sqrt{V_T}\right),$$

where $G_T^2\log(1+V_T/G_T)$ is defined to be zero when $G_T = 0$. One of the benefits over the very ingenious but highly complicated construction from Cutkosky and Mhammedi (2024) is the simplicity of our algorithm and its streamlined analysis.

Beyond bounds that explicitly depend on T, $\|x\|_2$, $\sum_{t=1}^{T}\|u_t\|_2^2$ and G_T, we will also argue that our algorithms enjoy a somewhat stronger guarantee that inherits the *Blackwell-like adaptive character*, which in some situations is tighter by almost a factor $\sqrt{T}$. We believe this is the core novelty and advantage of our approachability-based OLO algorithms.

[16] See the first display in the proof of Proposition 2 and the regularizer estimate in Appendix G.1 of arXiv:2203.00444v4, with $G = \varepsilon = \alpha_t = 1$.

[17] See Theorem 16 of that paper in the case $\psi(w) = w^2$, $\gamma = 1$, $p = 0$, $\varepsilon = 1$ and $h_1 = 1$.

7.2. **Blackwell's reduction for convex compact domains.** Long before Hart and Mas-Colell (2000), Blackwell (1956a) already proposed a reduction of regret minimization on the simplex to an approachability problem with outcome vectors in the lifted space $\mathbb{R}^d \times \mathbb{R}$. Consider the outcome function

$$\rho(a, u) = (u, -\langle u, a\rangle), \quad a \in \Delta_d,\ u \in \mathbb{R}^d,$$

and the target set

$$\mathcal{C} = \left\{(z, \zeta) \in \mathbb{R}^d \times \mathbb{R},\ \sup_{a \in \Delta_d} \langle z, a\rangle + \zeta \leqslant 0\right\}.$$

The latter is indeed a closed convex cone and its polar cone is $\mathcal{C}^\circ = \mathbb{R}_+(\Delta_d \times \{1\})$. Fix an arbitrary $a_1 \in \Delta_d$. This approachability problem admits the following oracle

$$\mathfrak{a}(z, \zeta) = \begin{cases} z/\zeta & \text{if } (z, \zeta) = (\zeta a, \zeta) \text{ for some } a \in \Delta_d \text{ and } \zeta > 0, \\ a_1 & \text{if } (z, \zeta) = (0, 0). \end{cases}$$

Then the regret with respect to a comparator $a \in \Delta_d$ can be expressed as

$$\sum_{t=1}^T \langle u_t, a - a_t\rangle = \left\langle \sum_{t=1}^T r_t, (a, 1)\right\rangle, \tag{9}$$

where, according to the definition of the outcome function, $r_t = (u_t, -\langle u_t, a_t\rangle)$. Therefore, any FTRL-based approachability algorithm from Section 4 defined by regularizers whose common domain generates $\mathcal{C}^\circ$, offers via Theorem 4.6 an upper bound on the above regret. Similarly, any OMD algorithm from Section 5 defined by a mirror map whose domain intersected with $\mathcal{C}^\circ$ generates $\mathcal{C}^\circ$ also offers some upper bound.

Although this reduction needs an additional dimension, it has the advantage (unlike the orthant reduction recalled in Section 6) of extending naturally to OLO with *arbitrary convex compact domains*. All the above indeed remains true if Δ_d is replaced by a convex compact set $\mathcal{A} \subset \mathbb{R}^d$. A variant of this extension[18] is used by Grand-Clément and Kroer (2024) to define the CBA and CBA$^+$ OLO algorithms, which are special cases of Blackwell's algorithm and the Greedy Blackwell algorithm, respectively.[19]

7.3. **Extension of Blackwell's reduction.** We now turn to unconstrained OLO for the remainder of this section. If the convex compact set $\mathcal{A}$ is replaced by $\mathbb{R}^d$, then the target set becomes $\mathcal{C} = \{0\} \times \mathbb{R}_-$ and thus $\mathcal{C}^\circ = \mathbb{R}^d \times \mathbb{R}_+$. The difficulty is that no oracle exists: for $w \in \mathbb{R}^d \setminus \{0\}$, $(w, 0)$ belongs to $\mathcal{C}^\circ$, and yet, for every $a \in \mathbb{R}^d$, taking $u = w$ gives

$$\langle \rho(a, w), (w, 0)\rangle = \langle (w, -\langle w, a\rangle), (w, 0)\rangle = \|w\|_2^2 > 0.$$

Nevertheless, restricting to the cone $\mathcal{K} = \mathbb{R}^d \times \mathbb{R}_+^* \cup \{(0, 0)\}$, which is dense in $\mathcal{C}^\circ$, it can immediately be seen that $\mathfrak{a}\colon \mathcal{K} \to \mathcal{A}$ defined as

$$\mathfrak{a}(w, \xi) = \begin{cases} w/\xi & \text{if } \xi > 0, \\ 0 & \text{if } (w, \xi) = (0, 0), \end{cases} \tag{10}$$

[18] Their conic embedding uses a bound on the action norms to scale the additional coordinate.

[19] Truncated conic lifting appears in Farina et al. (2022), who use optimistic FTRL with a logarithmic regularizer. Their lifted regret equals the positive part of the original regret, preventing cancellation between players' positive and negative regrets. Here we exploit arbitrary positive comparator rescaling on the whole cone (Lemma 3.2). Related truncated conic constructions yield logarithmic per-player regret in Soleymani et al. (2025) and constant per-player regret in Liu et al. (2026) and Abbadi et al. (2026). Constant per-player regret using multiplicative optimism and potential-based Regret Matching is obtained by Soleymani and Piliouras (2026). These logarithmic and constant rates concern prescribed simultaneous play in games, whereas our OLO guarantees hold against arbitrary payoff sequences.

is a *relaxed oracle*. Besides, the identity (9) between regret and cumulative outcomes remains valid. The FTRL bounds from Section 4 can therefore be applied to each comparator direction $(a, 1)$. Building on this key remark, the next two sections construct unconstrained OLO algorithms by carefully choosing regularizers whose subdifferential has domain in the cone $\mathcal{K} = \mathbb{R}^d \times \mathbb{R}_+^* \cup \{(0,0)\}$. The follow-up paper (Grand-Clément et al., 2026) extends this approach to work with non-Euclidean norms, arbitrary closed convex domains, and shows that it recovers coin-betting algorithms (Orabona and Pal, 2016) such as Krichevsky–Trofimov (Krichevsky and Trofimov, 1981), further demonstrating the relevance and unifying character of this approach.

8. Bounded payoff vectors

We here deal with unconstrained OLO with payoff vectors $(u_t)_{t\geqslant 1}$ such that $\|u_t\|_2 \leqslant 1$ for all $t \geqslant 1$. We introduce a regularizer on the cone $\mathcal{K} = \mathbb{R}^d \times \mathbb{R}_+^* \cup \{(0,0)\}$ which, combined with the relaxed oracle (10), yields a novel algorithm with several advantages over existing ones. First, it is simply an FTRL-based algorithm with a constant regularizer. Regarding regret bounds that explicitly depend on $\|a\|_2$ and $V_T \coloneqq \sum_{t=1}^T \|u_t\|_2^2$, it achieves

$$O\left(\|a\|_2 \log\left(1 + \|a\|_2\right) + \left(1 + \|a\|_2 \sqrt{\log(1 + \|a\|_2)}\right) \sqrt{V_T}\right), \tag{11}$$

which is a slight improvement over the existing bound discussed above. More importantly, the algorithm enjoys a novel *Blackwell-like adaptive guarantee*, which corresponds in some cases to a strong improvement, as will be demonstrated by Proposition 8.4 below. Furthermore, implementation is simple and cheap, as each iteration involves few elementary operations and an explicit 1-dimensional root-finding problem which can be quickly solved up to numerical precision by second-order methods. Implementation details are discussed in Section F.

Let $h\colon \mathbb{R}^d \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be defined as

$$h(w,\xi) = \begin{cases} \frac{3}{2}\left(\xi^2 \phi(\|w/\xi\|_2^2) + \xi\psi(\|w/\xi\|_2)\right) & \text{if } \xi > 0 \\ 0 & \text{if } (w,\xi) = (0,0) \\ +\infty & \text{otherwise,} \end{cases}$$

where $\phi(v) = 1 + v(1 + \log(1+v))$ and $\psi(v) = (1+v)\log(1+v) - v$. Only the above first term is 2-positively homogeneous but this will nevertheless give some interesting Blackwell-like adaptive character. The first proposition below establishes that it is indeed a regularizer. The quantity $h^*(R_T)$ appearing in Theorem 4.6 will be bounded from above by $\frac{1}{5}\sum_{t=1}^T \|u_t\|_2^2$ with the help of the key inequality in the second proposition, which bounds each increment. The proofs are deferred to Appendix E.

Proposition 8.1. *h is a regularizer with domain $\mathbb{R}^d \times \mathbb{R}_+^* \cup \{(0,0)\}$ and $h^*(0,0) = 0$.*

Proposition 8.2. *Let $u \in \mathbb{R}^d$ such that $\|u\|_2 \leqslant 1$. For $(z,\zeta) \in \mathbb{R}^d \times \mathbb{R}$, set $(w,\xi) \coloneqq \nabla h^*(z,\zeta)$ and $a \coloneqq \mathfrak{a}(w,\xi)$. Then,*

$$h^*(z + u, \zeta - \langle u, a\rangle) - h^*(z,\zeta) \leqslant \frac{1}{5}\|u\|_2^2.$$

The resulting algorithm is defined by $a_1 = 0$ and for $t \geqslant 1$,

$$R_t = \sum_{s=1}^t (u_s, -\langle u_s, a_s\rangle) \quad \text{and} \quad a_{t+1} = \mathfrak{a}\left(\nabla h^*\left(R_t\right)\right). \tag{12}$$

A practical description is given in Algorithm 3 in Appendix F.2.

Theorem 8.3. *Let $(u_t)_{t\geqslant 1}$ be a sequence in $\mathbb{R}^d$ such that $\|u_t\|_2 \leqslant 1$ $(t \geqslant 1)$. Algorithm (12) guarantees for all $a \in \mathbb{R}^d$ and $T \geqslant 1$,*

$$\sum_{t=1}^{T} \langle u_t, a - a_t\rangle \leqslant \frac{3}{2}\psi(\|a\|_2) + \sqrt{6\,\phi(\|a\|_2^2)\,h^*\left(R_T\right)}. \tag{13}$$

Consequently,

$$\sum_{t=1}^{T} \langle u_t, a - a_t\rangle \leqslant \frac{3}{2}\psi(\|a\|_2) + \sqrt{\frac{6}{5}}\sqrt{\phi(\|a\|_2^2)\sum_{t=1}^{T}\|u_t\|_2^2}. \tag{14}$$

Proof. Let $a \in \mathbb{R}^d$. Using (9) and applying general bound (i) from Theorem 4.6 with $x = (a, 1)$ (that indeed belongs to $\mathcal{C}^\circ = \mathbb{R}^d \times \mathbb{R}_+$) gives

$$\sum_{t=1}^{T} \langle u_t, a - a_t\rangle = \left\langle \sum_{t=1}^{T} r_t, (a, 1)\right\rangle \leqslant \inf_{\lambda>0} \frac{h(\lambda a, \lambda) + h^*(R_T)}{\lambda}.$$

Besides, $h(\lambda a, \lambda)$ is finite for all $\lambda > 0$ by definition of h, and the above infimum writes

$$\inf_{\lambda>0}\left\{\frac{3}{2}\psi(\|a\|_2) + \frac{3\lambda\phi(\|a\|_2^2)}{2} + \frac{h^*(R_T)}{\lambda}\right\}.$$

The minimization over $\lambda > 0$ balances the above last two terms and gives (13). Then using $h^*(0, 0) = 0$ and summing Proposition 8.2 gives $h^*(R_T) \leqslant \frac{1}{5}\sum_{t=1}^{T}\|u_t\|_2^2$, and (14) follows. □

Substituting the definitions of ψ and ϕ, explicit bound (14) can be seen to indeed be of the form (11), which for $\|a\|_2 \geqslant 2$, simplifies to

$$O\left(\|a\|_2 \log\|a\|_2 + \|a\|_2\sqrt{\log\|a\|_2 \sum_{t=1}^{T}\|u_t\|_2^2}\right).$$

At $a = 0$, bound (14) becomes $\sqrt{\frac{6}{5}\sum_{t=1}^{T}\|u_t\|_2^2}$.

Importantly, bound (13) retains $h^*(R_T)$, rather than replacing it by an upper bound in terms of the squared payoff norms. This corresponds to the inherited *Blackwell-like adaptive character*. As demonstrated by the following proposition, this gives in some cases a much tighter bound than the explicit one. The proof is given in Appendix E.6.

Proposition 8.4. *If $d \geqslant 2$, there exists a sequence $(u_t)_{t\geqslant 1}$ in $\mathbb{R}^d$ such that $\|u_t\|_2 \leqslant 1$ for all $t \geqslant 1$ and*

$$h^*(R_T) = O\left(\frac{\log^2 T}{T}\sum_{t=1}^{T}\|u_t\|_2^2\right)$$

as $T \to +\infty$.

9. Unbounded payoff vectors

Keeping the same notation, we turn to unconstrained OLO with arbitrary, possibly unbounded, payoff vectors. The algorithm below has several advantages over the highly ingenious but extremely complicated algorithm and analysis of Cutkosky and Mhammedi (2024). First, it is a simple FTRL algorithm with a time-dependent regularizer: one term is proportional to the magnitude of past

payoff vectors, while the other terms are constant. Regarding regret bounds that explicitly depend on $\|a\|_2$, $V_T \coloneqq \sum_{t=1}^T \|u_t\|_2^2$ and $G_T \coloneqq \max_{1\leqslant t\leqslant T} \|u_t\|_2$, it achieves

$$O\left(\|a\|_2^2 + G_T^2(1+\log(1+V_T/G_T)) + \left(1+\|a\|_2\sqrt{\log(1+\|a\|_2)}\right)\sqrt{V_T}\right), \tag{15}$$

where $G_T^2\log(1+V_T/G_T)$ is defined to be zero when $G_T = 0$. Ignoring logarithms, this bound is similar to that of Cutkosky and Mhammedi (2024). Beyond those dependencies, the algorithm inherits the *Blackwell-like adaptive character*, which in some cases strongly improves on the explicit bound, as shown by Proposition 9.4 below. Implementation is simple and cheap, similarly to the algorithm from Section 8, and is discussed in Appendix F.

Let $C_0 > 0$ and $C \geqslant 0$. Define $h_C\colon \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}\cup\{+\infty\}$ by

$$h_C(w,\xi) = \begin{cases} C(1+\xi\log\xi) + C_0\xi\|w/\xi\|_2^2 + \xi^2\phi\left(\|w/\xi\|_2^2\right) & \text{if } \xi > 0,\\ C & \text{if } (w,\xi) = (0,0),\\ +\infty & \text{otherwise,}\end{cases}$$

where $\phi(v) = 1 + v(1+\log(1+v))$. The following two propositions give the key properties of those regularizers and are proved in Appendices E.3 and E.4, respectively.

Proposition 9.1. *For all $C \geqslant 0$, h_C is a regularizer with domain $\mathbb{R}^d\times\mathbb{R}_+^*\cup\{(0,0)\}$ and $h_C^*(0,0) \leqslant 0$.*

Proposition 9.2. *Let $C_2 \geqslant C_1 \geqslant 0$, $u \in \mathbb{R}^d$ with $\|u\|_2^2 \leqslant C_0C_2/2$, and $(z,\zeta) \in \mathbb{R}^d\times\mathbb{R}$. Set $(w,\xi) = \nabla h_{C_1}^*(z,\zeta)$ and $a = \mathfrak{a}(w,\xi)$. Then*

$$h_{C_2}^*(z+u, \zeta - \langle u, a\rangle) - h_{C_1}^*(z,\zeta) \leqslant \frac{\|u\|_2^2}{2}.$$

The algorithm is defined as follows. Set $a_1 = 0$, $C_1 = 0$ and, for $t \geqslant 1$,

$$C_{t+1} = \max_{1\leqslant s\leqslant t}\frac{2\|u_s\|_2^2}{C_0}, \quad h_{t+1} = h_{C_{t+1}}, \quad R_t = \sum_{s=1}^t (u_s, -\langle u_s, a_s\rangle), \quad a_{t+1} = \mathfrak{a}(\nabla h_{t+1}^*(R_t)).$$

A practical description is given in Algorithm 5 in Appendix F.3.

Theorem 9.3. *Let $a \in \mathbb{R}^d$ and $T \geqslant 1$. If $G_T = 0$, the regret is zero. Otherwise,*

$$\sum_{t=1}^T \langle u_t, a - a_t\rangle \leqslant C_0\|a\|_2^2 + \frac{2G_T^2}{C_0}\left(1+\log\left(1+\frac{C_0h_{T+1}^*(R_T)}{2G_T^2}\right)\right) + 2\sqrt{\phi\left(\|a\|_2^2\right)\left(\frac{2G_T^2}{C_0} + h_{T+1}^*(R_T)\right)}.$$

Moreover, $h_{T+1}^*(R_T) \leqslant \frac{1}{2}\sum_{t=1}^T\|u_t\|_2^2$.

Proof. If $G_T = 0$, all the first T payoff vectors are zero, so the regret is zero. Assume that $G_T > 0$, so that $C_{T+1} > 0$. For $\xi > 0$, considering $(0,\xi)$ in the supremum that defines $h_{T+1}^*(R_T)$, we obtain

$$\begin{aligned} h_{T+1}^*(R_T) &= \sup_{(w,\xi)\in\mathbb{R}^d\times\mathbb{R}}\{\langle(w,\xi), R_T\rangle - h_{T+1}(w,\xi)\}\\ &\geqslant \langle(0,\xi), R_T\rangle - h_{T+1}(0,\xi)\\ &= \xi(R_T)_{d+1} - C_{T+1}(1+\xi\log\xi) - \xi^2\\ &> -C_{T+1}, \end{aligned} \tag{16}$$

where the last inequality holds for sufficiently small $\xi > 0$. This lower bound will be used below.

Let $a \in \mathbb{R}^d$. Using (9) and the general bound (i) from Theorem 4.6 applied with $x = (a, 1)$ gives for all $\lambda > 0$,

$$\begin{aligned}\sum_{t=1}^{T} \langle u_t, a - a_t \rangle = \left\langle \sum_{t=1}^{T} r_t, x \right\rangle &\leqslant \frac{h_{T+1}(\lambda a, \lambda) + h^*_{T+1}(R_T)}{\lambda} \\ &= C_0 \|a\|_2^2 + \frac{C_{T+1} + h^*_{T+1}(R_T)}{\lambda} + C_{T+1} \log \lambda + \lambda \phi\left(\|a\|_2^2\right).\end{aligned} \tag{17}$$

Choose for λ the unique positive solution of

$$\lambda^2 \phi\left(\|a\|_2^2\right) + C_{T+1}\lambda = C_{T+1} + h^*_{T+1}(R_T),$$

which exists thanks to (16). Then it can be verified that

$$\frac{C_{T+1} + h^*_{T+1}(R_T)}{\lambda} = C_{T+1} + \lambda \phi\left(\|a\|_2^2\right), \quad \lambda \leqslant 1 + \frac{h^*_{T+1}(R_T)}{C_{T+1}},$$

and

$$\lambda \phi\left(\|a\|_2^2\right) \leqslant \sqrt{\phi\left(\|a\|_2^2\right)\left(C_{T+1} + h^*_{T+1}(R_T)\right)}.$$

Substituting the above in (17) and using $C_{T+1} = 2G_T^2/C_0$ give the claimed regret bound.

Finally, $(C_t)_{t \geqslant 1}$ is nondecreasing and for all $t \geqslant 1$, $\|u_t\|_2 \leqslant \sqrt{C_0 C_{t+1}/2}$. Proposition 9.2 then applies and gives

$$h^*_{t+1}(R_t) - h^*_t(R_{t-1}) \leqslant \frac{\|u_t\|_2^2}{2}.$$

Summing the above and using $h^*_1(R_0) = h^*_{C_1}(0, 0) = 0$ proves $h^*_{T+1}(R_T) \leqslant \frac{1}{2}\sum_{t=1}^{T} \|u_t\|_2^2$, including when $G_T = 0$. □

For $C_0 = 1$, the above guarantee is of the form (15). Ignoring universal multiplicative constants and logarithms, it can be written

$$\tilde{O}\left(C_0 \|a\|_2^2 + \frac{G_T^2}{C_0} + \sqrt{\left(1 + \|a\|_2^2\right)\left(\frac{G_T^2}{C_0} + h^*_{T+1}(R_T)\right)}\right).$$

If $C_0 = 1$, and $\|a\|_2 \geqslant 1$, it further simplifies to

$$\tilde{O}\left(\|a\|_2^2 + G_T^2 + \|a\|_2 \sqrt{\left(h^*_{T+1}(R_T)\right)_+}\right).$$

Similarly to the algorithm in Section 8, $h^*_{T+1}(R_T)$ may be arbitrarily smaller than $\sum_{t=1}^{T} \|u_t\|_2^2$, almost up to a factor T, so the general bound above involving $h^*_{T+1}(R_T)$ illustrates the *Blackwell-like adaptive character*. The following proposition is proved in Appendix E.7.

Proposition 9.4. *Let $M > 0$. There exists a sequence $(u_t)_{t \geqslant 1}$ such that $\|u_t\|_2 \leqslant M$ for all $t \geqslant 1$ and*

$$h^*_{T+1}(R_T) = O\left(\frac{\log T}{T} \sum_{t=1}^{T} \|u_t\|_2^2\right)$$

as $T \to +\infty$.

10. Convex optimization

This section aims at demonstrating the relevance of applying an approachability-based unconstrained OLO algorithm to first-order optimization. We here restrict to a nonstochastic convex setting and to non-accelerated rates. We are nevertheless convinced that our approach could yield meaningful results beyond this scope.

We convert the unconstrained OLO for bounded payoff vectors defined in Section 8 using the normalized-gradient approach of Levy (2017) and Orabona (2023). The resulting algorithm will prove to be simultaneously adaptive to (i) the distance to the solution, (ii) Lipschitz continuity and (iii) (Hölder) smoothness, with no prior knowledge, no line-search and no parameter to tune. Obtaining such simultaneous adaptivity by applying the normalized-gradient conversion to parameter-free OLO algorithm was already suggested by Orabona (2023). Li and Lan (2026) also obtain *accelerated* rates, but their method for simultaneous adaptation to Lipschitz continuity and (Hölder) smoothness needs a prescribed target accuracy as input. Moshtaghifar et al. (2026) also obtain simultaneous distance and regularity adaptation on possibly unbounded domains, without line search or prescribed target accuracy. Ren et al. (2025) obtain accelerated (Hölder-)smooth rates without prescribing a target accuracy. Their method uses a line search, while their line-search-free variant assumes a bounded domain. Our algorithm will additionally inherit the Blackwell-like adaptive character.

Let $f\colon \mathbb{R}^d \to \mathbb{R}$ be a differentiable convex function[20] that admits a global minimizer $x_* \in \mathbb{R}^d$. We consider algorithm (12), denote $x_t = a_t$, and feed it with normalized gradients: for $t \geqslant 1$, if $\nabla f(x_t) = 0$, then x_t is a global minimizer, which we return. Otherwise, set

$$u_t = -\frac{\nabla f(x_t)}{\|\nabla f(x_t)\|_2}.$$

For every $T \geqslant 1$ before termination, define the output point

$$\bar{x}_T = \frac{\sum_{t=1}^T \|\nabla f(x_t)\|_2^{-1} x_t}{\sum_{t=1}^T \|\nabla f(x_t)\|_2^{-1}}.$$

As expected, the algorithm uses neither a Lipschitz constant nor a smoothness exponent or constant.

Theorem 10.1. *Let $T \geqslant 1$ be before termination. Then $h^*(R_T) \leqslant T/5$, and the following bounds hold.*

(i) If f is G-Lipschitz continuous for $\|\cdot\|_2$ for some $G > 0$, then

$$f(\bar{x}_T) - f(x_*) \leqslant \frac{G}{T}\left(\frac{3}{2}\psi\left(\|x_*\|_2\right) + \sqrt{6h^*(R_T)\phi\left(\|x_*\|_2^2\right)}\right),$$

where $\phi(v) = 1+v(1+\log(1+v))$, $\psi(v) = (1+v)\log(1+v)-v$, and $R_T = \sum_{t=1}^T (u_t, -\langle u_t, x_t\rangle)$. Consequently, ignoring universal multiplicative constants,

$$\begin{aligned} f(\bar{x}_T) - f(x_*) &= O\left(G\left(\frac{\|x_*\|_2 \log(1+\|x_*\|_2)}{T} + \frac{1+\|x_*\|_2\sqrt{\log(1+\|x_*\|_2)}}{\sqrt{T}}\right)\right) \\ &= \tilde{O}\left(G\left(\frac{\|x_*\|_2}{T} + \frac{1+\|x_*\|_2}{\sqrt{T}}\right)\right), \end{aligned}$$

where $\tilde{O}$ additionally ignores logarithms.

[20] We assume differentiability for simplicity, otherwise, subgradients can be used instead of gradients.

(ii) If, for some $\nu \in (0,1]$ and $L_\nu > 0$, the gradient of f is Hölder continuous for $\|\cdot\|_2$, meaning

$$\left\|\nabla f(x') - \nabla f(x)\right\|_2 \leqslant L_\nu \left\|x' - x\right\|_2^\nu, \quad x, x' \in \mathbb{R}^d,$$

then

$$f(\bar{x}_T) - f(x_*) \leqslant \left(\frac{1+\nu}{\nu}\right)^\nu \frac{L_\nu}{T^{1+\nu}} \left(\frac{3}{2}\psi\left(\|x_*\|_2\right) + \sqrt{6h^*(R_T)\phi\left(\|x_*\|_2^2\right)}\right)^{1+\nu}.$$

Consequently,

$$\begin{aligned} f(\bar{x}_T) - f(x_*) &= O\left(L_\nu \left(\frac{\|x_*\|_2 \log(1+\|x_*\|_2)}{T} + \frac{1+\|x_*\|_2\sqrt{\log(1+\|x_*\|_2)}}{\sqrt{T}}\right)^{1+\nu}\right) \\ &= \tilde{O}\left(L_\nu \left(\frac{\|x_*\|_2}{T} + \frac{1+\|x_*\|_2}{\sqrt{T}}\right)^{1+\nu}\right). \end{aligned}$$

Proof. By convexity of f and Theorem 8.3,

$$\begin{aligned} \sum_{t=1}^T \|\nabla f(x_t)\|_2^{-1} \left(f(x_t) - f(x_*)\right) &\leqslant \sum_{t=1}^T \langle u_t, x_* - x_t\rangle \\ &\leqslant \frac{3}{2}\psi(\|x_*\|_2) + \sqrt{6h^*(R_T)\phi(\|x_*\|_2^2)}. \end{aligned} \tag{18}$$

Hence, using Jensen's inequality,

$$f(\bar{x}_T) - f(x_*) \leqslant \frac{\frac{3}{2}\psi(\|x_*\|_2) + \sqrt{6h^*(R_T)\phi(\|x_*\|_2^2)}}{\sum_{t=1}^T \|\nabla f(x_t)\|_2^{-1}}. \tag{19}$$

If f is G-Lipschitz continuous for $\|\cdot\|_2$, the gradients have norm at most G and thus

$$f(\bar{x}_T) - f(x_*) \leqslant G\left(\frac{3\,\psi(\|x_*\|_2)}{2T} + \frac{\sqrt{6\,h^*(R_T)\,\phi(\|x_*\|_2^2)}}{T}\right).$$

By Proposition 8.2 and because the u_t are unit vectors,

$$h^*(R_T) \leqslant \frac{1}{5}\sum_{t=1}^T \|u_t\|_2^2 = \frac{T}{5}.$$

For every $v \geqslant 0$,

$$\psi(v) \leqslant v\log(1+v), \qquad \sqrt{\phi(v^2)} \leqslant 2\left(1 + v\sqrt{\log(1+v)}\right).$$

These inequalities give the asymptotic bounds in (i) for every $T \geqslant 1$ before termination.

By the Hölder continuity of ∇f assumed in (ii), the descent inequality from Nesterov (2015) gives,

$$f(x') \leqslant f(x) + \left\langle \nabla f(x), x' - x\right\rangle + \frac{L_\nu}{1+\nu}\left\|x' - x\right\|_2^{1+\nu}, \quad x, x' \in \mathbb{R}^d.$$

If $\nabla f(x) \neq 0$, consider

$$x' = x - \left(\frac{\|\nabla f(x)\|_2}{L_\nu}\right)^{1/\nu} \frac{\nabla f(x)}{\|\nabla f(x)\|_2}.$$

Since $f(x_*) \leqslant f(x')$, this yields

$$\|\nabla f(x)\|_2^{1+1/\nu} \leqslant \frac{1+\nu}{\nu} L_\nu^{1/\nu}\left(f(x) - f(x_*)\right).$$

Thus, by (18),

$$\sum_{t=1}^{T} \|\nabla f(x_t)\|_2^{1/\nu} \leqslant \frac{1+\nu}{\nu} L_\nu^{1/\nu} \sum_{t=1}^{T} \frac{f(x_t) - f(x_*)}{\|\nabla f(x_t)\|_2}$$
$$\leqslant \frac{1+\nu}{\nu} L_\nu^{1/\nu} \left(\frac{3}{2} \psi\left(\|x_*\|_2\right) + \sqrt{6h^*(R_T)\phi\left(\|x_*\|_2^2\right)} \right).$$

Since $v \mapsto v^{-\nu}$ is convex on $\mathbb{R}_+^*$, Jensen's inequality gives

$$\sum_{t=1}^{T} \|\nabla f(x_t)\|_2^{-1} \geqslant \frac{T^{1+\nu}}{\left(\sum_{t=1}^{T} \|\nabla f(x_t)\|_2^{1/\nu}\right)^{\nu}}.$$

Combining this with (19) proves the bound involving $h^*(R_T)$ in (ii). The estimates on ψ and ϕ above, together with $((1+\nu)/\nu)^{\nu} \leqslant 2$, then give the two asymptotic bounds. □

11. Fixed point iterations

We apply our approachability-based unconstrained OLO algorithm defined in Section 8 to solve fixed point problems and establish highly adaptive guarantees of a new kind.

In the context of fixed point iterations, the most favorable case is when the given map is a contraction, which allows for a geometric convergence to the unique solution by simply iterating the map (Banach, 1922). A useful relaxation is nonexpansiveness, meaning 1-Lipschitz continuity. Then, in a Euclidean space, assuming that a fixed point exists, the Krasnoselskii-Mann iteration (Mann, 1953; Krasnosel'skii, 1955) guarantees[21] a bound of order $R/\sqrt{T}$ on the fixed-point residual[22], where R is the distance of the initial point to the set of fixed points. Other methods (Lieder, 2021; Kim, 2021; Yoon and Ryu, 2021; Boţ and Nguyen, 2023; Boţ et al., 2025) achieve the *accelerated* R/T rate.

For a given map $F\colon \mathbb{R}^d \to \mathbb{R}^d$ and $L > 0$, an interesting remark is that maps F and $F_L := I + \frac{1}{L}(F - I)$ (where I denotes the identity map) have the same fixed points. Using map F_L instead of F may be beneficial: F_L may for instance be nonexpansive whereas F is not, and even if F is nonexpansive, the bound that corresponds to using F_L is linear in L. One may thus wish to use map F_L for the smallest value L such that F_L is nonexpansive, but the knowledge of such value may be unavailable. Kwon (2026) recently proposed an approach for converting OLO algorithms into fixed point iterations with nonexpansive maps. One highlight of that work is an algorithm that offers a guarantee similar to using map F_L for the lowest suitable value of L *without prior knowledge*, albeit with suboptimal dependence on the distance to the set of fixed points and with the slow $1/\sqrt{T}$ rate. This adaptivity to unknown scaling L can be seen as analogous to adaptivity to smoothness in convex optimization. The closest guarantee in the literature is that of Diakonikolas (2020, Algorithm 1 and Theorem 6) which has the advantage of achieving the optimal dependence on the distance to the set of fixed points, the fast $1/T$ rate, but on the other hand, only adapts to values of L that are larger than an initial guess, and thus misses the very favorable cases where the best suitable value of L is very small.

We convert the algorithm from Section 8 by adapting the approach from Kwon (2026) with normalized payoff vectors, analogously to Levy (2017) and Orabona (2023) and to Section 10. The fixed point iteration we obtain below is optimally adaptive (up to logarithms) to the distance to the set of fixed points. We also establish an adaptive guarantee analogous to adaptivity to Hölder

[21] See Vaisman (2005) and Cominetti et al. (2014) for the derivation of nonasymptotic bounds in the Euclidean/Hilbert setting.

[22] Meaning the distance between the output point and its image by the given map.

smoothness in convex optimization, and therefore stronger than adapting simply to scalar scalings $L > 0$ such that F_L is nonexpansive. This adaptivity is novel to the best of our knowledge. The guarantee additionally inherits the Blackwell-like adaptive character, although the rates are not *accelerated*.

The property of map F_L being nonexpansive is classically equivalent to associated map $G \coloneqq \frac{I-F}{2}$ being L-co-coercive (Bauschke and Combettes, 2017, Proposition 4.4), meaning that, for all $x, x' \in \mathbb{R}^d$,

$$\left\langle G(x') - G(x), x' - x \right\rangle \geqslant \frac{1}{L} \left\| G(x') - G(x) \right\|_2^2 .$$

The Baillon–Haddad theorem (Baillon and Haddad, 1977) establishes that the gradient of an L-smooth convex function satisfies this property[23]. We consider an extension of this property that is satisfied by gradients of Hölder-smooth convex functions.

Definition 11.1. Let $0 < \nu \leqslant 1$ and $L_\nu > 0$. A map $G\colon \mathbb{R}^d \to \mathbb{R}^d$ is (ν, L_ν)-*Hölder-co-coercive* if for all $x, x' \in \mathbb{R}^d$,

$$\left\langle G(x') - G(x), x' - x \right\rangle \geqslant L_\nu^{-1/\nu} \left\| G(x') - G(x) \right\|_2^{1+1/\nu} .$$

Relating to existing definitions, this property can be seen as inverse uniform monotonicity with power modulus (Sipoş, 2026; Oikonomidis et al., 2024). Note that $\nu = 1$ corresponds to regular co-coercivity.

Let $F\colon \mathbb{R}^d \to \mathbb{R}^d$ be a map that admits a fixed point $x_* \in \mathbb{R}^d$. We use algorithm (12), write $x_t = a_t$, and run it with payoff vectors:

$$u_t = \frac{F(x_t) - x_t}{\|F(x_t) - x_t\|_2}.$$

If $F(x_t) = x_t$, we return x_t and stop. As expected, the algorithm uses neither ν nor L_ν.

Theorem 11.2. *Assume that, for some $0 < \nu \leqslant 1$ and $L_\nu > 0$, $\frac{I-F}{2}$ is (ν, L_ν)-Hölder-co-coercive. Then, for every integer $T \geqslant 1$ before termination,*

$$\min_{1 \leqslant t \leqslant T} \|F(x_t) - x_t\|_2 \leqslant \frac{2L_\nu}{T^\nu} \left(\frac{3}{2} \psi\left(\|x_*\|_2 \right) + \sqrt{6h^*(R_T)\phi\left(\|x_*\|_2^2 \right)} \right)^\nu ,$$

where $R_T = \sum_{t=1}^T (u_t, -\langle u_t, x_t \rangle)$. Moreover, $h^*(R_T) \leqslant T/5$. Consequently, ignoring universal multiplicative constants,

$$\begin{aligned} \min_{1 \leqslant t \leqslant T} \|F(x_t) - x_t\|_2 &= O\left(L_\nu \left(\frac{\|x_*\|_2 \log(1 + \|x_*\|_2)}{T} + \frac{1 + \|x_*\|_2 \sqrt{\log(1 + \|x_*\|_2)}}{\sqrt{T}} \right)^\nu \right) \\ &= \tilde{O}\left(L_\nu \left(\frac{\|x_*\|_2}{T} + \frac{1 + \|x_*\|_2}{\sqrt{T}} \right)^\nu \right), \end{aligned}$$

where $\tilde{O}$ also ignores logarithms.

In particular, when $\nu = 1$, the bounds above apply under the equivalent assumption that $I + \frac{1}{L_1}(F - I)$ is nonexpansive.

Proof. Set $G = (I - F)/2$. Since $G(x_*) = 0$, Hölder-co-coercivity gives, for every $1 \leqslant t \leqslant T$,

$$\|F(x_t) - x_t\|_2^{1+1/\nu} \leqslant (2L_\nu)^{1/\nu} \langle F(x_t) - x_t, x_* - x_t \rangle .$$

[23] This motivates the convention of calling this property L-co-coercivity (as in "L-smoothness"), instead of $1/L$-co-coercivity as is sometimes proposed in the literature.

Since $F(x_t) \neq x_t$ before termination, we may divide this inequality by $\|F(x_t) - x_t\|_2$ and sum. Theorem 8.3 then gives

$$\begin{aligned}\sum_{t=1}^{T} \|F(x_t) - x_t\|_2^{1/\nu} &\leqslant (2L_\nu)^{1/\nu} \sum_{t=1}^{T} \langle u_t, x_* - x_t \rangle \\ &\leqslant (2L_\nu)^{1/\nu} \left(\frac{3}{2} \psi\left(\|x_*\|_2\right) + \sqrt{6h^*(R_T)\phi\left(\|x_*\|_2^2\right)} \right).\end{aligned}$$

The minimum of these T nonnegative terms is bounded by their average. Raising the resulting inequality to the power ν proves the first claim. By Proposition 8.2 and $h^*(0) = 0$,

$$h^*(R_T) \leqslant \frac{1}{5} \sum_{t=1}^{T} \|u_t\|_2^2 = \frac{T}{5}.$$

Substituting this estimate and the definitions of ϕ and ψ gives the stated bounds. □

12. Conclusion and perspectives

Inspired by empirical success reported in the literature, the initial goal of this work was to improve the theoretical understanding of the parameter-free and strong adaptive character of Blackwell's algorithms and related methods such as RM^+ and CBA^+.

In the context of approachability with convex cone targets, we introduced a generic approach for defining algorithms that share such favorable properties. It is based on an extended conversion scheme from OLO to approachability and the use of OLO algorithms defined on whole convex cones with the help of regularizers or mirror maps that are (partly) positively homogeneous of degree 2. This construction extends and unifies many known methods, such as Blackwell's algorithm, RM, RM^+, the Polynomially Weighted Average Forecaster, and its RM^+-like counterpart.

We then applied this approach to define algorithms for unconstrained OLO, thereby finding a new area of application for Blackwell's approachability. This is achieved by (i) considering action sets that are not compact, which is unusual in the topic of Blackwell's approachability, (ii) relaxing Blackwell's condition, (iii) extending Blackwell's reduction of OLO to approachability to work for unconstrained OLO, and (iv) carefully choosing regularizers on the unusual domain $\mathbb{R}^d \times \mathbb{R}_+^* \cup \{(0,0)\}$. The resulting algorithms compete with the state-of-the-art regarding explicit regret bounds, while being very simple, easy to implement and additionally inheriting the Blackwell-like adaptive character. The follow-up paper (Grand-Clément et al., 2026) extends this approach to work with non-Euclidean norms, arbitrary closed convex domains, and shows that it recovers coin-betting-based parameter-free regret minimizers such as Krichevsky–Trofimov, further demonstrating the unifying character of this approach.

These results provide a new approach for designing practical iterative methods with strong adaptive guarantees by applying Blackwell's approachability. Our first illustration in this spirit was the application to convex optimization, where the resulting algorithm offers nonaccelerated guarantees that adapt simultaneously to (i) the distance to the solution, (ii) Lipschitz continuity, and (iii) Hölder smoothness, with no prior knowledge of these quantities. Future work may develop this approach to achieve accelerated rates, per-coordinate adaptivity (similar to AdaGrad-Diagonal), adaptivity to noise, and guarantees for nonconvex objectives.

The other application to optimization dealt with fixed point iterations, where the resulting method achieves optimal adaptivity to (i) the distance to the set of fixed points and to (ii) regularity properties analogous to Hölder smoothness. These highly adaptive properties suggest that approachability-based methods could also be widely applied to the theory of monotone operators and splitting methods.

LLM usage disclosure

The core ideas of this work were laid down between 2023 and the beginning of 2026, before LLMs became sharp for mathematics. During Spring 2026, LLMs were used for exploring the analytic properties of numerous possible regularizers, before settling with those presented in Sections 8 and 9. LLMs were subsequently useful for proofreading and polishing.

Acknowledgments

The author is indebted to Vianney Perchet (resp. to Francesco Orabona) for introducing him to the topic of Blackwell's approachability (resp. of parameter-free regret minimization). Additionally, this work benefited from insightful discussions with Hector Denis, Ryan D'Orazio, Côme Fiegel, Julien Grand-Clément, Sylvain Sorin and Jérôme Taupin. This work was supported by the French Agence Nationale de la Recherche (ANR) under reference `ANR-21-CE40-0020` (CONVERGENCE project), and by a public grant as part of the *Investissement d'avenir project*, reference `ANR-11-LABX-0056-LMH`, LabEx LMH.

References

O. Abbadi, R. Laraki, and P. Mertikopoulos. Constant regret in general games via higher-order optimism. *arXiv:2609.04113*, 2026.

J. Abernethy, E. Hazan, and A. Rakhlin. Competing in the dark: An efficient algorithm for bandit linear optimization. In *Proceedings of the 21st Annual Conference on Learning Theory (COLT)*, pages 263–274, 2008.

J. Abernethy, P. L. Bartlett, and E. Hazan. Blackwell approachability and low-regret learning are equivalent. In *Proceedings of the 24th Annual Conference on Learning Theory*, volume 19 of *PMLR*, pages 27–46, 2011.

J.-B. Baillon and G. Haddad. Quelques propriétés des opérateurs angle-bornés et n-cycliquement monotones. *Israel J. Math.*, 26:137–150, 1977.

S. Banach. Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. *Fundam. Math.*, 3:133–181, 1922.

H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. Springer Science & Business Media, New York, 2nd edition, 2017.

A. Beck and M. Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Oper. Res. Lett.*, 31(3):167–175, 2003.

D. Blackwell. Controlled random walks. In *Proceedings of the International Congress of Mathematicians*, volume 3, pages 336–338, Amsterdam, 1956a.

D. Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific J. Math.*, 6(1):1–8, 1956b.

A. Blum and Y. Mansour. From external to internal regret. *J. Mach. Learn. Res.*, 8(6):1307–1324, 2007.

R. I. Boţ and D.-K. Nguyen. Fast Krasnosel'skiĭ–Mann algorithm with a convergence rate of the fixed point iteration of $o(1/k)$. *SIAM J. Numer. Anal.*, 61(6):2813–2843, 2023.

R. I. Boţ, E. R. Csetnek, and D.-K. Nguyen. Fast optimistic gradient descent ascent (OGDA) method in continuous and discrete time. *Found. Comput. Math.*, 25(1):163–222, 2025.

N. Brown and T. Sandholm. Superhuman AI for Heads-Up No-Limit Poker: Libratus beats top professionals. *Science*, 359(6374):418–424, 2018.

Y. Carmon and O. Hinder. Making SGD parameter-free. In *Proceedings of the 35th Conference on Learning Theory*, volume 178 of *PMLR*, pages 2360–2389, 2022.

N. Cesa-Bianchi and G. Lugosi. Potential-based algorithms in on-line prediction and game theory. *Mach. Learn.*, 51(3):239–261, 2003.

N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.

N. Cesa-Bianchi, Y. Freund, D. Haussler, D. P. Helmbold, R. E. Schapire, and M. K. Warmuth. How to use expert advice. *J. ACM*, 44(3):427–485, 1997.

N. Cesa-Bianchi, A. Conconi, and C. Gentile. On the generalization ability of on-line learning algorithms. *IEEE Trans. Inform. Theory*, 50(9):2050–2057, 2004.

D. Chakrabarti, J. Grand-Clément, and C. Kroer. Extensive-form game solving via Blackwell approachability on treeplexes. In *Advances in Neural Information Processing Systems 37*, 2024.

K. Chen, J. Langford, and F. Orabona. Better parameter-free stochastic optimization with ODE updates for coin-betting. *Proc. AAAI Conf. Artificial Intelligence*, 36(6):6239–6247, 2022.

E. Chzhen, C. Giraud, and G. Stoltz. A unified approach to fair online learning via Blackwell approachability. In *Advances in Neural Information Processing Systems 34*, 2021.

R. Cominetti, J. A. Soto, and J. Vaisman. On the rate of convergence of Krasnosel'skiĭ-Mann iterations and their connection with sums of Bernoullis. *Israel J. Math.*, 199(2):757–772, 2014.

A. Cutkosky. Anytime online-to-batch, optimism and acceleration. In *Proceedings of the International Conference on Machine Learning*, pages 1446–1454, 2019a.

A. Cutkosky. Artificial constraints and hints for unbounded online learning. In *Proceedings of the 32nd Conference on Learning Theory*, volume 99 of *PMLR*, pages 874–894, 2019b.

A. Cutkosky and K. Boahen. Online learning without prior information. In *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *PMLR*, pages 643–677, 2017.

A. Cutkosky and Z. Mhammedi. Fully unconstrained online learning. In *Advances in Neural Information Processing Systems 37*, 2024.

A. Cutkosky and F. Orabona. Black-box reductions for parameter-free online learning in Banach spaces. In *Proceedings of the 31st Conference on Learning Theory*, volume 75 of *PMLR*, pages 1493–1529, 2018.

A. Cutkosky, H. Mehta, and F. Orabona. Optimal stochastic non-smooth non-convex optimization through online-to-non-convex conversion. In *Proceedings of the International Conference on Machine Learning*, pages 6643–6670, 2023.

S. De Rooij, T. Van Erven, P. D. Grünwald, and W. M. Koolen. Follow the leader if you can, hedge if you must. *J. Mach. Learn. Res.*, 15(1):1281–1316, 2014.

A. Defazio and K. Mishchenko. Learning-rate-free learning by D-Adaptation. In *Proceedings of the International Conference on Machine Learning*, pages 7449–7479, 2023.

A. Defazio, X. Yang, H. Mehta, K. Mishchenko, A. Khaled, and A. Cutkosky. The road less scheduled. In *Advances in Neural Information Processing Systems 37*, 2024.

J. Diakonikolas. Halpern iteration for near-optimal and parameter-free monotone inclusion and strong solutions to variational inequalities. In *Proceedings of the 33rd Conference on Learning Theory*, volume 125 of *PMLR*, pages 1428–1451, 2020.

R. D'Orazio. Regret minimization with function approximation in extensive-form games. Master's thesis, University of Alberta, 2020.

R. D'Orazio and R. Huang. Optimistic and adaptive Lagrangian hedging. In *Proceedings of the AAAI Workshop on Reinforcement Learning and Games*, 2021.

J. Duchi, E. Hazan, and Y. Singer. Adaptive subgradient methods for online learning and stochastic optimization. *J. Mach. Learn. Res.*, 12:2121–2159, 2011.

E. Even-Dar, R. Kleinberg, S. Mannor, and Y. Mansour. Online learning for global cost functions. In *Proceedings of the 22nd Annual Conference on Learning Theory*, pages 316–326. Omnipress, 2009.

G. Farina, C. Kroer, and T. Sandholm. Faster game solving via predictive Blackwell approachability: Connecting Regret Matching and mirror descent. *Proc. AAAI Conf. Artificial Intelligence*, 35 (6):5363–5371, 2021.

G. Farina, I. Anagnostides, H. Luo, C.-W. Lee, C. Kroer, and T. Sandholm. Near-optimal no-regret learning dynamics for general convex games. In *Advances in Neural Information Processing Systems 35*, 2022.

G. E. Flaspohler, F. Orabona, J. Cohen, S. Mouatadid, M. Oprescu, P. Orenstein, and L. Mackey. Online learning with optimism and delay. In *Proceedings of the International Conference on Machine Learning*, pages 3363–3373, 2021.

D. J. Foster, S. Kale, M. Mohri, and K. Sridharan. Parameter-free online learning via model selection. In *Advances in Neural Information Processing Systems 30*, 2017.

D. J. Foster, A. Rakhlin, and K. Sridharan. Online learning: Sufficient statistics and the Burkholder method. In *Proceedings of the 31st Conference on Learning Theory*, volume 75 of *PMLR*, pages 3028–3064, 2018.

D. P. Foster. A proof of calibration via Blackwell's approachability theorem. *Games Econom. Behav.*, 29(1–2):73–78, 1999.

G. J. Gordon. *Approximate solutions to Markov decision processes*. PhD thesis, Carnegie Mellon University, 1999a.

G. J. Gordon. Regret bounds for prediction problems. In *Proceedings of the 12th Annual Conference on Computational Learning Theory*, pages 29–40. ACM, 1999b.

G. J. Gordon. No-regret algorithms for online convex programs. In *Advances in Neural Information Processing Systems 19*, 2007.

J. Grand-Clément and C. Kroer. Solving optimization problems with Blackwell approachability. *Math. Oper. Res.*, 49(2):697–728, 2024.

J. Grand-Clément, J. Kwon, and F. Orabona. Conic-lifting for learning-rate-free convex optimization. *arXiv:2610.39531*, 2026.

S. Hart and A. Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000.

S. Hart and A. Mas-Colell. A general class of adaptive strategies. *J. Econom. Theory*, 98(1):26–54, 2001.

E. Hazan. *Introduction to Online Convex Optimization*. Now Publishers, 2016.

E. Hazan and S. Kale. Extracting certainty from uncertainty: regret bounded by variation in costs. *Mach. Learn.*, 80(2–3):165–188, 2010.

M. Ivgi, O. Hinder, and Y. Carmon. DoG is SGD's best friend: A parameter-free dynamic step size schedule. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *PMLR*, pages 14465–14499, 2023.

A. Jacobsen and A. Cutkosky. Parameter-free mirror descent. In *Proceedings of the 35th Conference on Learning Theory*, volume 178 of *PMLR*, pages 4160–4211, 2022.

A. Juditsky, J. Kwon, and É. Moulines. Unifying mirror descent and dual averaging. *Math. Program.*, 199(1–2):793–830, 2023.

K.-S. Jun and F. Orabona. Parameter-free online convex optimization with sub-exponential noise. In *Proceedings of the 32nd Conference on Learning Theory*, volume 99 of *PMLR*, pages 1802–1823, 2019.

S. M. Kakade, S. Shalev-Shwartz, and A. Tewari. Regularization techniques for learning with matrices. *J. Mach. Learn. Res.*, 13(1):1865–1890, 2012.

A. Khaled, K. Mishchenko, and C. Jin. DoWG unleashed: An efficient universal parameter-free gradient descent method. In *Advances in Neural Information Processing Systems 36*, 2023.

D. Kim. Accelerated proximal point method for maximally monotone operators. *Math. Program.*, 190(1–2):57–87, 2021.

D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2015.

J. Kivinen and M. K. Warmuth. Exponentiated gradient versus gradient descent for linear predictors. *Inf. Comput.*, 132(1):1–63, 1997.

M. A. Krasnosel'skii. Two remarks on the method of successive approximations. *Uspekhi Matematicheskikh Nauk*, 10(1):123–127, 1955.

I. Kreisler, M. Ivgi, O. Hinder, and Y. Carmon. Accelerated parameter-free stochastic optimization. In *Proceedings of the 37th Conference on Learning Theory*, volume 247 of *PMLR*, pages 3257–3324, 2024.

R. Krichevsky and V. Trofimov. The performance of universal encoding. *IEEE Trans. Inform. Theory*, 27(2):199–207, 1981.

J. Kwon. *Mirror descent strategies for regret minimization and approachability*. PhD thesis, Université Pierre-et-Marie-Curie, 2016.

J. Kwon. Refined approachability algorithms and application to regret minimization with global costs. *J. Mach. Learn. Res.*, 22(200):1–38, 2021.

J. Kwon. A regret minimization approach to fixed-point iterations. *arXiv:2509.21653*, 2026.

J. Kwon, Y. Wan, and B. Ziliotto. Time-dependent Blackwell approachability and application to absorbing games. *arXiv:2303.04956*, 2024.

D. Lee, G. Noarov, M. Pai, and A. Roth. Online minimax multiobjective optimization: Multicalibeating and other applications. In *Advances in Neural Information Processing Systems 35*, 2022.

K. Y. Levy. Online to offline conversions, universality and adaptive minibatch sizes. In *Advances in Neural Information Processing Systems 30*, 2017.

T. Li and G. Lan. A simple uniformly optimal method without line search for convex optimization. *Math. Program.*, 219(1–2):219–256, 2026.

F. Lieder. On the convergence rate of the Halpern-iteration. *Optim. Lett.*, 15(2):405–418, 2021.

M. Liu, G. Farina, and A. Ozdaglar. Constant individual regret in general games. *arXiv:2608.31166*, 2026.

I. Loshchilov and F. Hutter. Decoupled weight decay regularization. In *Proceedings of the International Conference on Learning Representations*, 2019.

W. R. Mann. Mean value methods in iteration. *Proc. Amer. Math. Soc.*, 4(3):506–510, 1953.

S. Mannor and N. Shimkin. Regret minimization in repeated matrix games with variable stage duration. *Games Econom. Behav.*, 63(1):227–258, 2008.

H. B. McMahan and F. Orabona. Unconstrained online linear learning in Hilbert spaces: Minimax algorithms and normal approximations. In *Proceedings of the 27th Conference on Learning Theory*, volume 35 of *PMLR*, pages 1020–1039, 2014.

H. B. McMahan and M. Streeter. Adaptive bound optimization for online convex optimization. In *Proceedings of the 23rd Annual Conference on Learning Theory*, pages 244–256. ACM, 2010.

Z. Mhammedi and W. M. Koolen. Lipschitz and comparator-norm adaptivity in online learning. In *Proceedings of the 33rd Conference on Learning Theory*, volume 125 of *PMLR*, pages 2858–2887, 2020.

K. Mishchenko and A. Defazio. Prodigy: An expeditiously adaptive parameter-free learner. In *Proceedings of the International Conference on Machine Learning*, volume 235 of *PMLR*, pages 35779–35804, 2024.

M. Moravčík, M. Schmid, N. Burch, V. Lisý, D. Morrill, N. Bard, T. Davis, K. Waugh, M. Johanson, and M. Bowling. DeepStack: Expert-level artificial intelligence in Heads-Up No-Limit Poker. *Science*, 356(6337):508–513, 2017.

J. Moreau. Proximité et dualité dans un espace hilbertien. *Bull. Soc. Math. France*, 93:273–299, 1965.

M. Moshtaghifar, A. Rodomanov, D. Vankov, and S. Stich. DADA: Dual averaging with distance adaptation. In *Proceedings of the International Conference on Learning Representations*, 2026.

A. Nemirovski and D. B. Yudin. *Problem Complexity and Method Efficiency in Optimization*. Wiley-Interscience Series in Discrete Mathematics. John Wiley & Sons, Chichester, 1983.

Y. Nesterov. Primal-dual subgradient methods for convex problems. *Math. Program.*, 120(1): 221–259, 2009.

Y. Nesterov. Universal gradient methods for convex optimization problems. *Math. Program.*, 152 (1–2):381–404, 2015.

K. Oikonomidis, E. Laude, P. Latafat, A. Themelis, and P. Patrinos. Adaptive proximal gradient methods are universal without approximation. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *PMLR*, pages 38663–38682, 2024.

F. Orabona. Dimension-free exponentiated gradient. In *Advances in Neural Information Processing Systems 26*, 2013.

F. Orabona. Simultaneous model selection and optimization through parameter-free stochastic learning. In *Advances in Neural Information Processing Systems 27*, 2014.

F. Orabona. Normalized gradients for all. *arXiv:2308.05621*, 2023.

F. Orabona. Online learning: A modern introduction using convex optimization. *arXiv:1912.13213*, 2026. Version 10, June 2026; final version, to be published by Cambridge University Press, March 2027.

F. Orabona and D. Pal. Coin betting and parameter-free online learning. In *Advances in Neural Information Processing Systems 29*, 2016.

F. Orabona and T. Tommasi. Training deep networks without learning rates through coin betting. In *Advances in Neural Information Processing Systems 30*, 2017.

V. Perchet. Approachability, regret and calibration: Implications and equivalences. *J. Dyn. Games*, 1(2):181–254, 2014.

V. Perchet. Exponential weight approachability, applications to calibration and regret minimization. *Dyn. Games Appl.*, 5(1):136–153, 2015.

Y. Ren, H. Xu, and Q. Deng. Accelerated distance-adaptive methods for Hölder smooth and convex optimization. In *Advances in Neural Information Processing Systems 38*, 2025.

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

A. Rodomanov, A. Kavis, Y. Wu, K. Antonakopoulos, and V. Cevher. Universal gradient methods for stochastic convex optimization. *arXiv:2402.03210*, 2024.

E. K. Ryu and W. Yin. *Large-Scale Convex Optimization: Algorithms & Analyses via Monotone Operators*. Cambridge University Press, 2022.

S. Shalev-Shwartz. *Online learning: Theory, Algorithms, and Applications*. PhD thesis, The Hebrew University of Jerusalem, 2007.

S. Shalev-Shwartz. Online learning and online convex optimization. *Found. Trends Mach. Learn.*, 4(2):107–194, 2012.

S. Shalev-Shwartz and Y. Singer. A primal-dual perspective of online learning algorithms. *Mach. Learn.*, 69(2–3):115–142, 2007.

N. Shimkin. An online convex optimization approach to Blackwell's approachability. *J. Mach. Learn. Res.*, 17(1):4434–4456, 2016.

A. Sipoş. The computational content of super strongly nonexpansive mappings and uniformly monotone operators. *Isr. J. Math.*, 272(1):437–465, 2026.

A. Soleymani and G. Piliouras. Multiplicative optimism for constant regret in games. *arXiv:2609.21976*, 2026.

A. Soleymani, G. Piliouras, and G. Farina. Cautious optimism: A meta-algorithm for near-constant regret in general games. In *Proceedings of the 26th ACM Conference on Economics and Computation*, EC '25, page 870. ACM, 2025.

M. Streeter and H. B. McMahan. No-regret algorithms for unconstrained online convex optimization. In *Advances in Neural Information Processing Systems 25*, 2012.

O. Tammelin. Solving large imperfect information games using CFR+. *arXiv:1407.5042*, 2014.
O. Tammelin, N. Burch, M. Johanson, and M. Bowling. Solving Heads-Up Limit Texas Hold'em. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence*, 2015.
T. Tieleman and G. Hinton. Lecture 6.5-RMSProp: Divide the gradient by a running average of its recent magnitude. COURSERA: Neural Networks for Machine Learning, Lecture 6, 2012. Slide 29.
J. Vaisman. Convergencia fuerte del método de medias sucesivas para operadores lineales no-expansivos. Memoria de ingeniería civil matemática, Universidad de Chile, 2005.
T. van Erven, W. M. Koolen, S. de Rooij, and P. Grünwald. Adaptive Hedge. In *Advances in Neural Information Processing Systems 24*, 2011.
V. G. Vovk. Aggregating strategies. In *Proceedings of the 3rd Annual Workshop on Computational Learning Theory*, pages 371–383. Morgan Kaufmann, 1990.
M. K. Warmuth and A. K. Jagota. Continuous and discrete-time nonlinear gradient descent: Relative loss bounds and convergence. In *Electronic Proceedings of the 5th International Symposium on Artificial Intelligence and Mathematics*, 1998.
T. Yoon and E. K. Ryu. Accelerated algorithms for smooth convex-concave minimax problems with $O(1/k^2)$ rate on squared gradient norm. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *PMLR*, pages 12098–12109, 2021.
J. Zhang and A. Cutkosky. Parameter-free regret in high probability with heavy tails. In *Advances in Neural Information Processing Systems 35*, 2022.
Z. Zhang, H. Yang, A. Cutkosky, and I. C. Paschalidis. Improving adaptive online learning using refined discretization. In *Proceedings of the 35th International Conference on Algorithmic Learning Theory*, volume 237 of *PMLR*, pages 1208–1233, 2024.
Y. Zhong, Z. Zheng, M. C. Chou, and C.-P. Teo. Resource pooling and allocation policies to deliver differentiated service. *Manage. Sci.*, 64(4):1555–1573, 2018.
M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the Twentieth International Conference on Machine Learning*, 2003.
M. Zinkevich, M. Johanson, M. Bowling, and C. Piccione. Regret minimization in games with incomplete information. In *Advances in Neural Information Processing Systems 20*, 2007.

Appendix A. Convexity tools

Proposition A.1 (Fenchel–Young inequality)**.** *Let $f\colon \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a proper function. For all $x, y \in \mathbb{R}^d$, $\langle y, x\rangle \leqslant f(x) + f^*(y)$.*

Proposition A.2 (See, e.g., Rockafellar, 1970, Theorem 23.5)**.** *Let $f\colon \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous convex function and $x, y \in \mathbb{R}^d$. The following statements are equivalent:*

(i) $x \in \partial f^*(y)$,
(ii) $y \in \partial f(x)$,
(iii) $\langle y, x\rangle = f(x) + f^*(y)$,
(iv) $x \in \operatorname{Arg\,max}_{x' \in \mathbb{R}^d} \{\langle y, x'\rangle - f(x')\}$,
(v) $y \in \operatorname{Arg\,max}_{y' \in \mathbb{R}^d} \{\langle y', x\rangle - f^*(y')\}$.

Definition A.3 (Smoothness)**.** Let $f\colon \mathbb{R}^d \to \mathbb{R}$ be a differentiable function, $\|\cdot\|$ a norm in $\mathbb{R}^d$ and $L > 0$. f is *L-smooth* with respect to $\|\cdot\|$ if for all $x, x' \in \mathbb{R}^d$, $|D_f(x', x)| \leqslant \frac{L}{2}\|x' - x\|^2$, in other words,

$$\left|f(x') - f(x) - \left\langle \nabla f(x), x' - x\right\rangle\right| \leqslant \frac{L}{2}\left\|x' - x\right\|^2.$$

Definition A.4 (Strong convexity)**.** Let $f\colon \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$, $\|\cdot\|$ a norm in $\mathbb{R}^d$ and $K > 0$. f is *K-strongly convex* with respect to $\|\cdot\|$ if for all $x, x' \in \mathbb{R}^d$ and $\lambda \in [0,1]$,

$$f(\lambda x + (1-\lambda)x') \leqslant \lambda f(x) + (1-\lambda) f(x') - \frac{K\lambda(1-\lambda)}{2} \left\|x' - x\right\|^2 .$$

Proposition A.5 (Duality between strong convexity and smoothness, Kakade et al., 2012, Theorem 3)**.** *Let $f\colon \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a proper, lower semicontinuous convex function, $\|\cdot\|$ a norm in $\mathbb{R}^d$ and $K > 0$. Then f is K-strongly convex with respect to $\|\cdot\|$ if and only if f^* is finite and differentiable on $\mathbb{R}^d$ and $1/K$-smooth with respect to $\|\cdot\|_*$.*

Proposition A.6 (Shalev-Shwartz, 2007, Lemma 17)**.** *The squared Euclidean norm $h_2\colon x \mapsto \frac{1}{2}\|x\|_2^2$ is 1-strongly convex and 1-smooth with respect to $\|\cdot\|_2$. More generally, for $1 < p \leqslant 2$, the function $x \mapsto \frac{1}{2}\|x\|_p^2$ is $(p-1)$-strongly convex with respect to $\|\cdot\|_p$.*

Proposition A.7. *Let $f\colon \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a proper, lower semicontinuous, convex function that is positively homogeneous of degree 2. Then, $f, f^* \geqslant 0$ and $f(0) = f^*(0) = 0$.*

Proof. Choose $z \in \operatorname{dom} f$. By lower semicontinuity and homogeneity, $f(0) \leqslant \lim_{\lambda \to 0^+} \lambda^2 f(z) = 0$. Thus $f(0)$ is finite, and homogeneity at zero gives $f(0) = 0$. For $x \in \operatorname{dom} f$ and $0 < \lambda < 1$, convexity gives

$$\lambda^2 f(x) = f(\lambda x) \leqslant \lambda f(x),$$

so $f(x) \geqslant 0$. Finally, using $x = 0$ in the definition of the conjugate gives $f^*(y) \geqslant 0$ for every y, and $f^*(0) = -\inf f = 0$. □

Appendix B. Convex cones

For background on convex cones and polarity, see Rockafellar (1970, Section 14).

Definition B.1. A nonempty set $\mathcal{C} \subset \mathbb{R}^d$ is a *convex cone* if, for all $x, x' \in \mathcal{C}$ and $\lambda \geqslant 0$, $x + x' \in \mathcal{C}$ and $\lambda x \in \mathcal{C}$. It is a *closed convex cone* if it is also closed.

Remark B.2. Since a closed convex cone $\mathcal{C}$ is nonempty, closed, and convex, its Euclidean projection onto $\mathcal{C}$ is well-defined and denoted $\Pi_{\mathcal{C}}$.

Definition B.3. Let $\mathcal{A} \subset \mathbb{R}^d$. The *polar cone* of $\mathcal{A}$ is the set

$$\mathcal{A}^\circ = \left\{ y \in \mathbb{R}^d,\ \forall x \in \mathcal{A},\ \langle y, x \rangle \leqslant 0 \right\}.$$

Proposition B.4. *A polar cone is a closed convex cone.*

Proposition B.5. *If $\mathcal{C}$ is a closed convex cone, then $\mathcal{C}^{\circ\circ} = \mathcal{C}$.*

Proposition B.6. *The negative orthant $\mathbb{R}^d_-$ is a closed convex cone and its polar cone is the positive orthant $\mathbb{R}^d_+$.*

Theorem B.7 (Moreau's decomposition theorem, Moreau, 1965)**.** *Let $\mathcal{C} \subset \mathbb{R}^d$ be a closed convex cone. For all $x \in \mathbb{R}^d$, it holds that:*

$$x = \Pi_{\mathcal{C}}(x) + \Pi_{\mathcal{C}^\circ}(x),$$

where $\Pi_{\mathcal{C}}$ and $\Pi_{\mathcal{C}^\circ}$ denote the Euclidean projection onto $\mathcal{C}$ and $\mathcal{C}^\circ$ respectively.

Appendix C. Regularizers and FTRL

Proposition C.1 (ℓ^p regularization on a polar cone)**.** *Let $\mathcal{C} \subset \mathbb{R}^d$ be a closed convex cone, $1 < p < +\infty$, $q = p/(p-1)$, $H = \frac{1}{2}\|\cdot\|_p^2$ and $h = H + I_{\mathcal{C}^\circ}$. For $y \in \mathbb{R}^d$, denote by $\Pi_{\mathcal{C}}^{(q)}(y)$ the unique point of $\mathcal{C}$ minimizing $\|y - z\|_q$ over $z \in \mathcal{C}$. For $q = 2$, this is the Euclidean projection $\Pi_{\mathcal{C}}$.*

(i) h is a regularizer, positively homogeneous of degree 2, with domain $\mathcal{C}^\circ$ and $h(0) = h^*(0) = 0$. For every $y \in \mathbb{R}^d$,

$$h^*(y) = \frac{1}{2}\operatorname{dist}_q(y, \mathcal{C})^2, \qquad \nabla h^*(y) = \nabla H^*\Big(y - \Pi_{\mathcal{C}}^{(q)}(y)\Big),$$

where $H^* = \frac{1}{2}\|\cdot\|_q^2$.

(ii) If $1 < p \leqslant 2$, then h is $(p-1)$-strongly convex with respect to $\|\cdot\|_p$ and h^* is $1/(p-1)$-smooth with respect to $\|\cdot\|_q$.

(iii) If $\mathcal{C} = \mathbb{R}^d_-$, then

$$h^*(y) = \frac{1}{2}\|y_+\|_q^2, \qquad \nabla h^*(y) = \nabla H^*(y_+) = \|y_+\|_q^{2-q}\, y_+^{q-1}.$$

If $\mathcal{C} = \mathbb{R}^d_+$, then

$$h^*(y) = \frac{1}{2}\|(-y)_+\|_q^2, \qquad \nabla h^*(y) = -\|(-y)_+\|_q^{2-q}\,(-y)_+^{q-1}.$$

Powers of nonnegative vectors are taken component-wise. Each gradient is defined to be zero when its positive-part argument is zero.

Proof. Hölder's inequality gives $H^* = \frac{1}{2}\|\cdot\|_q^2$. The function h is proper, lower semicontinuous and strictly convex, and $0 \leqslant h^* \leqslant H^* < +\infty$, so it is a regularizer. Its domain, homogeneity and values at zero follow directly.

The projection $\Pi_{\mathcal{C}}^{(q)}(y)$ exists by coercivity and is unique by strict convexity of H^*. Set $x = \nabla H^*(y - \Pi_{\mathcal{C}}^{(q)}(y))$. First-order optimality and conicity give

$$x \in \mathcal{C}^\circ, \qquad \Big\langle x, \Pi_{\mathcal{C}}^{(q)}(y)\Big\rangle = 0.$$

Since $y - \Pi_{\mathcal{C}}^{(q)}(y) \in \partial H(x)$, these relations imply $y \in \partial h(x)$. Propositions A.2 and 4.3 therefore give $\nabla h^*(y) = x$ and

$$h^*(y) = \langle y, x\rangle - H(x) = H^*\Big(y - \Pi_{\mathcal{C}}^{(q)}(y)\Big).$$

Strong convexity follows from Proposition A.6 by restriction to $\mathcal{C}^\circ$, and smoothness follows from Proposition A.5. For the negative and positive orthants, the projection residuals are respectively y_+ and $-(-y)_+$. The stated formulas follow by differentiating H^*. □

Proposition C.2 (FTRL regret bound)**.** *Let $(h_t)_{t\geqslant 1}$ be regularizers with common domain $\mathcal{X}$ such that $h_{t+1} \geqslant h_t$ for every $t \geqslant 1$. For payoff vectors $u_t \in \mathbb{R}^d$, set $U_0 = 0$, $U_t = \sum_{s=1}^t u_s$ and $x_t = \nabla h_t^*(U_{t-1})$. Then, for every $x \in \mathcal{X}$ and $T \geqslant 1$,*

$$\sum_{t=1}^T \langle u_t, x - x_t\rangle \leqslant h_{T+1}(x) + h_1^*(0) + \sum_{t=1}^T D_{h_t^*}(U_t, U_{t-1}).$$

Proof. Since $h_{t+1} \geqslant h_t$ implies $h_{t+1}^* \leqslant h_t^*$, telescoping gives

$$\begin{aligned} h_{T+1}^*(U_T) - h_1^*(0) &= \sum_{t=1}^{T} \left(h_{t+1}^*(U_t) - h_t^*(U_{t-1}) \right) \\ &\leqslant \sum_{t=1}^{T} \left(\langle u_t, x_t \rangle + D_{h_t^*}(U_t, U_{t-1}) \right). \end{aligned}$$

The regret bound follows from the Fenchel–Young inequality $\langle U_T, x \rangle \leqslant h_{T+1}(x) + h_{T+1}^*(U_T)$. □

Appendix D. Mirror maps and OMD

Proposition D.1 (ℓ^p mirror maps)**.** *Let $1 < p < +\infty$, $q = p/(p-1)$ and $H = \frac{1}{2} \|\cdot\|_p^2$ on $\mathbb{R}^d$. Then H is a mirror map compatible with every nonempty closed convex set $\mathcal{X} \subset \mathbb{R}^d$, and $H^* = \frac{1}{2} \|\cdot\|_q^2$. Both H and H^* are positively homogeneous of degree 2, with $H(0) = H^*(0) = 0$. For nonzero $x, y \in \mathbb{R}^d$,*

$$\begin{aligned} \nabla H(x) &= \|x\|_p^{2-p} \left(\operatorname{sgn}(x_i) |x_i|^{p-1} \right)_{i=1}^d, \\ \nabla H^*(y) &= \|y\|_q^{2-q} \left(\operatorname{sgn}(y_i) |y_i|^{q-1} \right)_{i=1}^d, \end{aligned}$$

and $\nabla H(0) = \nabla H^*(0) = 0$. The maps ∇H and ∇H^* are inverses of one another. For $1 < p \leqslant 2$, H is $(p-1)$-strongly convex with respect to $\|\cdot\|_p$. For $2 \leqslant p < +\infty$, it is $(p-1)$-smooth with respect to $\|\cdot\|_p$.

Proof. The conjugate formula, strong convexity and smoothness follow from Proposition C.1 with $\mathcal{C} = \{0\}$, using exponent q for the smoothness assertion. The homogeneity and values at zero follow from the squared-norm formulas. Differentiation gives the two gradient formulas at nonzero inputs. Their continuous extensions at zero are the gradients there, since $H(x)$ and $H^*(x)$ are $O(\|x\|_2^2)$ as $x \to 0$. The formulas give $\nabla H^*(\nabla H(x)) = x$ and $\nabla H(\nabla H^*(y)) = y$ for every $x, y \in \mathbb{R}^d$. Thus ∇H is onto $\mathbb{R}^d$. Together with continuity, strict convexity and $\operatorname{dom} H = \mathbb{R}^d$, this verifies Definition 5.1 for every such $\mathcal{X}$. □

Proposition D.2. *Let $\mathcal{X} \subset \mathbb{R}^d$ be a nonempty closed convex set, H a mirror map compatible with $\mathcal{X}$, and $x_1 \in \mathcal{X} \cap \mathcal{D}_H$. For payoff vectors $r_t \in \mathbb{R}^d$, define*

$$x_{t+1} = \operatorname*{arg\,max}_{x \in \mathcal{X}} \left\{ \langle \nabla H(x_t) + r_t, x \rangle - H(x) \right\}.$$

Then, for every $x \in \mathcal{X} \cap \operatorname{dom} H$ and $T \geqslant 1$,

$$\begin{aligned} \sum_{t=1}^{T} \langle r_t, x - x_t \rangle &\leqslant D_H(x, x_1) - D_H(x, x_{T+1}) \\ &\quad + \sum_{t=1}^{T} D_{H^*} \left(\nabla H(x_t) + r_t, \nabla H(x_t) \right). \end{aligned}$$

Proof. The optimality condition for x_{t+1} and the definition of the Bregman divergence give

$$\begin{aligned} \langle r_t, x - x_t \rangle &\leqslant D_H(x, x_t) - D_H(x, x_{t+1}) \\ &\quad + \langle r_t, x_{t+1} - x_t \rangle - D_H(x_{t+1}, x_t). \end{aligned}$$

By the Fenchel–Young inequality (Proposition A.1) and the inverse-gradient identities in Proposition 5.2,

$$\begin{aligned}\langle r_t, x_{t+1} - x_t\rangle - D_H(x_{t+1}, x_t) &\leqslant H^*\left(\nabla H(x_t) + r_t\right) - H^*\left(\nabla H(x_t)\right) - \langle r_t, x_t\rangle \\ &= D_{H^*}\left(\nabla H(x_t) + r_t, \nabla H(x_t)\right).\end{aligned}$$

Summing over $t = 1, \dots, T$ proves the claim. □

Appendix E. Postponed proofs

E.1. **Proof of Proposition 6.1.** (i) Let $x \in (\mathbb{R}^d_-)^\circ = \mathbb{R}^d_+$. If $x = 0$, the dot product from Blackwell's condition is zero for any value of $\mathfrak{a}(x)$. If $x \neq 0$, because $\langle \mathbb{1}, x\rangle = \|x\|_1$, for all $u \in \mathbb{R}^d$,

$$\begin{aligned}\langle \rho\left(\mathfrak{a}(x), u\right), x\rangle &= \left\langle \rho\left(\frac{x}{\|x\|_1}, u\right), x\right\rangle = \left\langle u - \left\langle u, \frac{x}{\|x\|_1}\right\rangle \mathbb{1}, x\right\rangle \\ &= \langle u, x\rangle - \left\langle u, \frac{x}{\|x\|_1}\right\rangle \langle \mathbb{1}, x\rangle = \langle u, x\rangle - \langle u, x\rangle = 0.\end{aligned}$$

(ii) For all $a \in \Delta_d$,

$$\begin{aligned}\left\langle \sum_{t=1}^T r_t, a\right\rangle &= \left\langle \sum_{t=1}^T \rho(a_t, u_t), a\right\rangle = \sum_{t=1}^T \langle u_t - \langle u_t, a_t\rangle \mathbb{1}, a\rangle \\ &= \sum_{t=1}^T \left(\langle u_t, a\rangle - \langle u_t, a_t\rangle \langle \mathbb{1}, a\rangle\right) = \sum_{t=1}^T \langle u_t, a - a_t\rangle.\end{aligned}$$

E.2. **Proof of Proposition 8.1.** The domain of h is indeed $\mathbb{R}^d \times \mathbb{R}^*_+ \cup \{(0,0)\}$ by definition. For $\xi > 0$, we can write h as

$$h(w, \xi) = \frac{3}{2}\left(\|w\|_2^2 + \xi^2 + \|w\|_2^2 \log\left(1 + \frac{\|w\|_2^2}{\xi^2}\right) + \xi\psi\left(\|w/\xi\|_2\right)\right). \tag{20}$$

We prove that the last two terms are convex. The first two terms will then give strong (and thus strict) convexity.

For $v \geqslant 0$, set $f(v) = v\sqrt{\log(1+v^2)}$. This function is nonnegative and nondecreasing. For $v > 0$, differentiation gives

$$f''(v) = \frac{v\left((v^2+3)\log(1+v^2) - v^2\right)}{(1+v^2)^2\left(\log(1+v^2)\right)^{3/2}} > 0,$$

where positivity follows from $\log(1+v^2) \geqslant v^2/(1+v^2)$. Since f is continuous at zero, it is convex on $[0, +\infty)$. Therefore, $w \mapsto f(\|w\|_2)$ is convex, and so is the associated perspective function

$$(w, \xi) \longmapsto \xi f\left(\|w/\xi\|_2\right).$$

Moreover, this function is nonnegative on the half-space $\xi > 0$, so its square is convex as well. Thus,

$$\|w\|_2^2 \log\left(1 + \frac{\|w\|_2^2}{\xi^2}\right) = \left(\xi f\left(\|w/\xi\|_2\right)\right)^2$$

is convex on that half-space.

Similarly, for $v \geqslant 0$,

$$\psi'(v) = \log(1+v) \geqslant 0, \qquad \psi''(v) = \frac{1}{1+v} > 0.$$

Consequently, $w \mapsto \psi(\|w\|_2)$ is convex, and the associated perspective function $(w, \xi) \mapsto \xi\psi(\|w/\xi\|_2)$ is convex on $\xi > 0$.

Extending these two terms by zero at $(0, 0)$ and by $+\infty$ outside $\mathcal{K}$ preserves convexity. Indeed, the logarithmic term is nonnegative and positively homogeneous of degree 2, whereas the term involving ψ is positively homogeneous of degree 1, so the convexity inequality also holds on segments joining the origin to a point with $\xi > 0$. The last two terms in (20) are thus convex. h is then strictly convex because the sum of the first two terms is strictly convex.

We now turn to lower semicontinuity. The function h is continuous for $\xi > 0$. At the origin, lower semicontinuity follows from $h \geqslant 0 = h(0, 0)$. At any other point $(w_0, 0)$ of the boundary, we have $w_0 \neq 0$, and

$$\|w\|_2^2 \log\left(1 + \frac{\|w\|_2^2}{\xi^2}\right) \longrightarrow +\infty$$

as $(w, \xi) \to (w_0, 0)$ with $\xi > 0$. Since all terms in (20) are nonnegative, this proves lower semicontinuity there. At points with $\xi < 0$, the function is identically $+\infty$ in a neighborhood. Hence h is lower semicontinuous everywhere.

This proves that h is a pre-regularizer, and Proposition 4.2 ensures that strong convexity makes it a regularizer.

E.3. **Proof of Proposition 9.1.** The domain follows from the definition. For $\xi > 0$, we have

$$h_C(w, \xi) = \|w\|_2^2 + \xi^2 + C(1 + \xi \log \xi) + C_0 \frac{\|w\|_2^2}{\xi} + \|w\|_2^2 \log\left(1 + \frac{\|w\|_2^2}{\xi^2}\right).$$

We prove that the last three terms are convex. The squared Euclidean norm in the first two terms will then give strong convexity.

For $v \geqslant 0$, set $f(v) = v\sqrt{\log(1 + v^2)}$. This function is nonnegative and nondecreasing. For $v > 0$,

$$f''(v) = \frac{v\left((v^2 + 3)\log(1 + v^2) - v^2\right)}{(1 + v^2)^2 \left(\log(1 + v^2)\right)^{3/2}} > 0,$$

where we used $\log(1 + v^2) \geqslant v^2/(1 + v^2)$. Since f is continuous at zero, it is convex on $[0, +\infty)$. Thus $w \mapsto f\left(\|w\|_2\right)$ is convex, and its perspective

$$(w, \xi) \longmapsto \xi f\left(\|w/\xi\|_2\right)$$

is convex on $\xi > 0$. This function is nonnegative, so its square

$$\|w\|_2^2 \log\left(1 + \frac{\|w\|_2^2}{\xi^2}\right) = \left(\xi f\left(\|w/\xi\|_2\right)\right)^2$$

is convex as well.

The function $(w, \xi) \mapsto \|w\|_2^2/\xi$ is the perspective of the convex function $w \mapsto \|w\|_2^2$, and is therefore convex on $\xi > 0$. Moreover, the second derivative of $\xi \mapsto 1 + \xi \log \xi$ is $1/\xi > 0$.

We extend the logarithmic term and the quadratic perspective by zero at the origin, and the entropy term by its limit 1 there. These extensions preserve convexity. The logarithmic term and the quadratic perspective are nonnegative and positively homogeneous of degrees 2 and 1, respectively. For the entropy term, if $0 < t < 1$ and $\xi > 0$, then

$$1 + t\xi \log(t\xi) = 1 + t\xi \log \xi + t\xi \log t \leqslant 1 + t\xi \log \xi.$$

Thus the convexity inequality also holds on segments joining the origin to a point with $\xi > 0$. Extending by $+\infty$ outside the domain preserves convexity. Consequently, $h_C(w, \xi) - \|w\|_2^2 - \xi^2$ is convex, so h_C is 2-strongly convex for the Euclidean norm on $\mathbb{R}^{d+1}$.

We now prove lower semicontinuity. The function is continuous for $\xi > 0$. As $(w, \xi) \to (0, 0)$ with $\xi > 0$, the entropy term tends to C, and all the other terms are nonnegative. Hence

$$\liminf_{(w,\xi)\to(0,0)} h_C(w, \xi) \geqslant C = h_C(0, 0).$$

At a boundary point $(w_0, 0)$ with $w_0 \neq 0$,

$$C_0 \frac{\|w\|_2^2}{\xi} \longrightarrow +\infty$$

as $(w, \xi) \to (w_0, 0)$ with $\xi > 0$. The remaining terms are nonnegative, since $1 + \xi \log \xi \geqslant 1 - e^{-1} > 0$. At points with $\xi < 0$, the function is identically $+\infty$ in a neighborhood. This proves lower semicontinuity everywhere.

The function h_C is proper because it is finite at the origin. It is therefore a pre-regularizer, and its strong convexity and Proposition 4.2 make it a regularizer. Finally, $h_C \geqslant 0$, so

$$h_C^*(0, 0) = \sup_{(w,\xi)\in\mathbb{R}^d\times\mathbb{R}} \{-h_C(w, \xi)\} \leqslant 0.$$

E.4. **Proof of Proposition 9.2.** Set $\Phi(w, \xi) = \xi^2 \phi(\|w\|_2^2 / \xi^2)$ for $\xi > 0$ and $\Phi(0, 0) = 0$. We use the convention $0 \log 0 = 0$.

Suppose first that $\xi > 0$, so $a = w/\xi$ and $(z, \zeta) = \nabla h_{C_1}(w, \xi)$. Fenchel equality gives

$$\begin{aligned} &h_{C_2}^*(z + u, \zeta - \langle u, a\rangle) - h_{C_1}^*(z, \zeta) \\ &\quad = \sup_{(w',\xi')\in\operatorname{dom} h_{C_1}} \Big\{ \big\langle u, w' - \xi' a \big\rangle \\ &\qquad - D_{h_{C_1}}((w', \xi'), (w, \xi)) - (C_2 - C_1)(1 + \xi' \log \xi') \Big\}. \end{aligned}$$

The origin contributes at most zero. We bound the other candidates in two cases.

If $0 < \xi' \leqslant 2\xi$, we claim that

$$D_\Phi((w', \xi'), (w, \xi)) \geqslant \frac{\|w' - \xi' a\|_2^2}{2}.$$

Indeed, with the notation of Lemma E.2, we have $\Phi(w, \xi) = (\xi \widehat{\phi}(a))^2$, where $\nabla^2 \widehat{\phi}(a) \succeq 5 I_d / (6\widehat{\phi}(a))$. For $(\omega, \eta) \in \mathbb{R}^d \times \mathbb{R}$, differentiation gives

$$\begin{aligned} &\nabla^2 \Phi(w, \xi)[(\omega, \eta), (\omega, \eta)] \\ &\quad = 2 \left(\eta \widehat{\phi}(a) + \left\langle \nabla \widehat{\phi}(a), \omega - \eta a \right\rangle \right)^2 \\ &\qquad + 2\widehat{\phi}(a)(\omega - \eta a)^\top \nabla^2 \widehat{\phi}(a)(\omega - \eta a) \\ &\quad \geqslant \frac{5}{3} \|\omega - \eta a\|_2^2 . \end{aligned}$$

Integrating this inequality along the segment from (w, ξ) to (w', ξ') yields

$$\begin{aligned} D_\Phi((w', \xi'), (w, \xi)) &\geqslant \frac{5}{3} \left\|w' - \xi' a\right\|_2^2 \int_0^1 \frac{(1 - s)\xi^2}{((1 - s)\xi + s\xi')^2} \, ds \\ &\geqslant \frac{5}{3} \left\|w' - \xi' a\right\|_2^2 \int_0^1 \frac{1 - s}{(1 + s)^2} \, ds \\ &= \frac{5}{3}(1 - \log 2) \left\|w' - \xi' a\right\|_2^2 \geqslant \frac{\|w' - \xi' a\|_2^2}{2}. \end{aligned}$$

By convexity of the other terms and $1+\xi'\log\xi' \geqslant 0$, the corresponding expression in the supremum is at most

$$\langle u, w' - \xi' a\rangle - \frac{\|w' - \xi' a\|_2^2}{2} \leqslant \frac{\|u\|_2^2}{2}.$$

If $\xi' > 2\xi$, direct calculation and convexity of Φ give

$$\begin{aligned} &D_{h_{C_1}}((w',\xi'),(w,\xi)) + (C_2 - C_1)(1+\xi'\log\xi') \\ &\quad\geqslant \frac{C_0\|w' - \xi' a\|_2^2}{\xi'} + C_1\xi'\left(\log\frac{\xi'}{\xi} - 1 + \frac{\xi}{\xi'}\right) + (C_2 - C_1)\xi' \\ &\quad\geqslant \frac{C_0\|w' - \xi' a\|_2^2}{\xi'} + \frac{C_2\xi'}{6}, \end{aligned}$$

where we used $1 + s\log s \geqslant s$ and the fact that $r \mapsto \log r - 1 + 1/r$ increases for $r > 1$, with $\log 2 - 1/2 \geqslant 1/6$. Thus the expression in the supremum is at most

$$\begin{aligned} \langle u, w' - \xi' a\rangle - \frac{C_0\|w' - \xi' a\|_2^2}{\xi'} - \frac{C_2\xi'}{6} &\leqslant \frac{\|u\|_2^2\,\xi'}{4C_0} - \frac{C_2\xi'}{6} \\ &\leqslant \frac{C_2\xi'}{8} - \frac{C_2\xi'}{6} \leqslant 0. \end{aligned}$$

This proves the desired bound when $\xi > 0$.

It remains to consider $(w,\xi) = (0,0)$, in which case $a = 0$ and $(z,\zeta) \in \partial h_{C_1}(0,0)$. Applying the subgradient inequality at (sx', s) and dividing by $s > 0$ gives

$$\langle z, x'\rangle + \zeta \leqslant C_1\log s + C_0\left\|x'\right\|_2^2 + s\phi(\left\|x'\right\|_2^2).$$

Taking $s \to 0^+$ with $x' = 0$ shows that $C_1 = 0$, and taking the limit for arbitrary x' gives $\langle z, x'\rangle + \zeta \leqslant C_0\|x'\|_2^2$. Fenchel equality gives $h_0^*(z,\zeta) = 0$. For $\xi' > 0$, applying this inequality at $x' = w'/\xi'$ yields

$$\begin{aligned} \langle z + u, w'\rangle + \zeta\xi' - h_{C_2}(w',\xi') &\leqslant \langle u, w'\rangle - \Phi(w',\xi') - C_2(1+\xi'\log\xi') \\ &\leqslant \langle u, w'\rangle - \left\|w'\right\|_2^2 \leqslant \frac{\|u\|_2^2}{4}. \end{aligned}$$

The candidate origin contributes $-C_2 \leqslant 0$. Taking the supremum completes the proof.

E.5. **Proof of Proposition 8.2.** For $x \in \mathbb{R}^d$, set

$$\widehat{\phi}(x) = \sqrt{\phi\left(\|x\|_2^2\right)}, \qquad \widehat{\psi}(x) = \psi\left(\|x\|_2\right).$$

Thus, for $\xi > 0$,

$$h(\xi x, \xi) = \frac{3}{2}\left(\xi^2\widehat{\phi}(x)^2 + \xi\widehat{\psi}(x)\right).$$

The following estimate is the main ingredient of the proof.

Lemma E.1. *For all $x, x' \in \mathbb{R}^d$ and $\xi, \xi' > 0$,*

$$\begin{aligned} D_h\left((\xi' x', \xi'), (\xi x, \xi)\right) &\geqslant \frac{3}{2}\xi' D_{\widehat{\psi}}(x', x) \\ &\quad + \frac{5}{4}\xi'^2\left(\|x' - x\|_2 - \frac{3}{2}D_{\widehat{\psi}}(x', x)\right)_+^2. \end{aligned} \tag{21}$$

We first deduce the proposition from Lemma E.1, whose proof follows the two lemmas below. Suppose first that $\xi > 0$. Then $w = \xi a$ and $(z, \zeta) = \nabla h(w, \xi)$. Fenchel equality gives

$$\begin{aligned} &h^*\left(z + u, \zeta - \langle u, a\rangle\right) - h^*(z, \zeta) \\ &\quad = \sup_{(w', \xi') \in \operatorname{dom} h} \left(\langle u, w' - \xi' a\rangle - D_h\left((w', \xi'), (w, \xi)\right)\right). \end{aligned}$$

The origin $(w', \xi') = (0, 0)$ contributes $-D_h((0,0),(w,\xi)) \leqslant 0$. Every other point of $\operatorname{dom} h$ is of the form $(\xi' x', \xi')$ for some $\xi' > 0$ and $x' \in \mathbb{R}^d$. Since $\|u\|_2 \leqslant 1$ and $D_{\widehat{\psi}}(x', a) \geqslant 0$, Lemma E.1 yields

$$\begin{aligned} &\xi' \langle u, x' - a\rangle - D_h\left((\xi' x', \xi'), (\xi a, \xi)\right) \\ &\qquad \leqslant \max_{\beta \geqslant 0} \left(\|u\|_2 \beta - \frac{5}{4}\beta^2\right) = \frac{1}{5}\|u\|_2^2. \end{aligned} \tag{22}$$

Taking the supremum proves the bound when $\xi > 0$.

If $(w, \xi) = (0, 0)$, then $a = \mathfrak{a}(0,0) = 0$ by definition of the relaxed oracle. Fenchel equality gives $h^*(z, \zeta) = 0$ and $(z, \zeta) \in \partial h(0, 0)$. Thus, for every $\xi' > 0$ and $x' \in \mathbb{R}^d$,

$$\xi'\left(\langle z, x'\rangle + \zeta\right) \leqslant \frac{3}{2}\left(\xi'^2 \widehat{\phi}(x')^2 + \xi' \widehat{\psi}(x')\right).$$

Dividing by ξ' and taking the limit as $\xi' \to 0^+$ yields $\langle z, x'\rangle + \zeta \leqslant 3\widehat{\psi}(x')/2$. Consequently,

$$\begin{aligned} h^*(z + u, \zeta) - h^*(z, \zeta) &\leqslant \sup_{x' \in \mathbb{R}^d} \sup_{\xi' \geqslant 0} \left(\xi' \langle u, x'\rangle - \frac{3}{2}\xi'^2 \widehat{\phi}(x')^2\right) \\ &\leqslant \sup_{x' \in \mathbb{R}^d} \frac{\langle u, x'\rangle^2}{6\widehat{\phi}(x')^2} \leqslant \frac{\|u\|_2^2}{6} \leqslant \frac{\|u\|_2^2}{5}, \end{aligned}$$

where we used $\widehat{\phi}(x')^2 \geqslant \|x'\|_2^2$.

It remains to prove Lemma E.1. We first establish some lower estimates on $\widehat{\phi}$ and $\widehat{\psi}$.

Lemma E.2. *The functions $\widehat{\phi}$ and $\widehat{\psi}$ are twice continuously differentiable. For every $x \in \mathbb{R}^d$,*

$$\widehat{\phi}(x) \geqslant \frac{1 + \|x\|_2}{\sqrt{2}}, \qquad \nabla^2 \widehat{\phi}(x) \succeq \frac{5}{6\widehat{\phi}(x)} I_d, \qquad \nabla^2 \widehat{\psi}(x) \succeq \frac{1}{1 + \|x\|_2} I_d.$$

Moreover, for all $v, \varepsilon \geqslant 0$,

$$\sqrt{\phi\left((v + \varepsilon)^2\right)} \leqslant \sqrt{\phi\left(v^2\right)}\left(1 + \frac{\varepsilon}{1 + v}\right)^2.$$

Proof. Set $v = \|x\|_2$. The first bound follows from $\phi(v^2) \geqslant 1 + v^2 \geqslant (1 + v)^2/2$. Straightforward differentiation gives

$$\widehat{\phi}(x) \nabla^2 \widehat{\phi}(x) = \phi'(v^2) I_d + \left(2\phi''(v^2) - \frac{\phi'(v^2)^2}{\phi(v^2)}\right) x x^\top.$$

Since $\phi'(v^2) \geqslant 1$, it suffices to prove

$$\phi'(v^2) + 2v^2 \phi''(v^2) - \frac{v^2 \phi'(v^2)^2}{\phi(v^2)} \geqslant \frac{5}{6}.$$

Set $\lambda = 1 + \log(1 + v^2)$ and $\theta = 1/(1 + v^2)$. Rearranging gives

$$\begin{aligned} &\theta \phi(v^2)\left(\phi'(v^2) + 2v^2 \phi''(v^2) - \frac{v^2 \phi'(v^2)^2}{\phi(v^2)} - \frac{5}{6}\right) \\ &\qquad = \frac{\lambda}{6}\left(1 + 11\theta - 18\theta^2 + 12\theta^3\right) - 1 + \frac{31}{6}\theta - 4\theta^2 - \theta^3. \end{aligned} \tag{23}$$

The elementary bound $\log(1+v^2) \geqslant 2v^2/(v^2+2)$, also applied after rescaling, gives

$$\lambda \geqslant \begin{cases} \frac{13}{3} + \frac{2(1-32\theta)}{1+32\theta}, & \text{if } 0 < \theta \leqslant \dfrac{1}{32}, \\ 3 + \frac{2(1-8\theta)}{1+8\theta}, & \text{if } \dfrac{1}{32} < \theta \leqslant \dfrac{1}{8}, \\ \frac{3-\theta}{1+\theta}, & \text{if } \dfrac{1}{8} < \theta \leqslant 1. \end{cases}$$

The coefficient of λ in (23) is positive. Substitution and rearranging show that its right-hand side is nonnegative on each interval, proving the Hessian bound for $\widehat{\phi}$.

For $x \neq 0$, differentiating gives

$$\nabla^2 \widehat{\psi}(x) = \frac{1}{1+v}\frac{xx^\top}{v^2} + \frac{\log(1+v)}{v}\left(I_d - \frac{xx^\top}{v^2}\right) \succeq \frac{1}{1+v} I_d,$$

since $\log(1+v) \geqslant v/(1+v)$. The inequality also holds at zero, where $\nabla^2\widehat{\psi}(0) = I_d$.

Finally,

$$\frac{d}{dv}\log\sqrt{\phi(v^2)} = \frac{v\phi'(v^2)}{\phi(v^2)} \leqslant \frac{2}{1+v}.$$

The inequality follows by rearranging, using $\log(1+v^2) \leqslant 1$ for $0 \leqslant v \leqslant 1$ and $\log(1+v^2) \geqslant 2v^2/(v^2+2)$ for $v \geqslant 1$. Integrating between v and $v+\varepsilon$ proves the last estimate. □

The next lemma compares the logarithmic expressions used in the proof of Lemma E.1.

Lemma E.3. *For every $\alpha \geqslant 0$ and $\gamma \geqslant 1/2$,*

$$\gamma - \left(\sqrt{\gamma} - \frac{5}{6\sqrt{\gamma}}\left(\log(1+\alpha) - \frac{\alpha}{1+\alpha}\right)\right)_+^2 \geqslant \frac{5}{6}\left(\frac{3}{2}\log(1+\alpha) - \frac{\alpha}{2}\right)_+^2.$$

Proof. Set

$$\mu(\alpha) = \log(1+\alpha) - \frac{\alpha}{1+\alpha}, \qquad \nu(\alpha) = \frac{3}{2}\log(1+\alpha) - \frac{\alpha}{2}.$$

Straightforward differentiation gives

$$\mu'(\alpha) = \frac{\alpha}{(1+\alpha)^2}, \qquad \nu'(\alpha) = \frac{2-\alpha}{2(1+\alpha)}.$$

The left-hand side of the claimed inequality is nondecreasing in γ. Its value at $\gamma = 1/2$ is

$$\begin{cases} \frac{5}{3}\mu(\alpha) - \frac{25}{18}\mu(\alpha)^2, & \text{if } \mu(\alpha) \leqslant \dfrac{3}{5}, \\ \frac{1}{2}, & \text{if } \mu(\alpha) \geqslant \dfrac{3}{5}. \end{cases} \tag{24}$$

We use basic inequality

$$\log(1+\alpha) \leqslant \frac{2\alpha}{\alpha+2} + \frac{\alpha^3}{6(\alpha+2)(\alpha+1)}. \tag{25}$$

The desired comparison is immediate when $\nu(\alpha) \leqslant 0$. Suppose that $\nu(\alpha) > 0$. Since ν attains its maximum at 2, (25) gives $\nu(\alpha) \leqslant \nu(2) \leqslant 2/3$. Thus the desired comparison holds when $\mu(\alpha) \geqslant 3/5$, because $5\nu(\alpha)^2/6 \leqslant 10/27 < 1/2$.

It remains to consider $\mu(\alpha) \leqslant 3/5$. Set

$$\chi(\alpha) = 2\mu(\alpha) - \frac{5}{3}\mu(\alpha)^2 - \nu(\alpha)^2.$$

For $0 \leqslant \alpha \leqslant 2$, straightforward differentiation, (25), and rearranging give $\chi'(\alpha) \geqslant 0$. Hence $\chi(\alpha) \geqslant \chi(0) = 0$ on $[0, 2]$. For $\alpha \geqslant 2$ with $\mu(\alpha) \leqslant 3/5$ and $\nu(\alpha) \geqslant 0$,

$$\chi'(\alpha) = \left(2 - \frac{10}{3}\mu(\alpha)\right)\mu'(\alpha) - 2\nu(\alpha)\nu'(\alpha) \geqslant 0.$$

Since μ increases and ν decreases on $[2, +\infty)$, these two conditions hold on an interval starting at 2. Thus $\chi(\alpha) \geqslant \chi(2) \geqslant 0$ there as well, completing the comparison with (24). □

Proof of Lemma E.1. We first prove that, for all $x, x' \in \mathbb{R}^d$,

$$\widehat{\phi}(x')^2 - \left(\widehat{\phi}(x) + \left\langle \nabla\widehat{\phi}(x), x' - x \right\rangle\right)_+^2 \geqslant \frac{5}{6}\left(\|x' - x\|_2 - \frac{3}{2}D_{\widehat{\psi}}(x', x)\right)_+^2. \tag{26}$$

Indeed, the bound in Lemma E.1 follows from this estimate: straightforward differentiation gives

$$\frac{2}{3}D_h\left((\xi' x', \xi'), (\xi x, \xi)\right) = \xi'^2\widehat{\phi}(x')^2 + \xi^2\widehat{\phi}(x)^2 + \xi' D_{\widehat{\psi}}(x', x) \\ - 2\xi\xi'\widehat{\phi}(x)\left(\widehat{\phi}(x) + \left\langle \nabla\widehat{\phi}(x), x' - x \right\rangle\right).$$

The minimum of this quadratic expression over $\xi \geqslant 0$ is

$$\xi'^2\left(\widehat{\phi}(x')^2 - \left(\widehat{\phi}(x) + \left\langle \nabla\widehat{\phi}(x), x' - x \right\rangle\right)_+^2\right) + \xi' D_{\widehat{\psi}}(x', x).$$

To prove (26), we may assume that $x \neq x'$, since both sides vanish when $x = x'$. Set

$$\tau = \frac{\|x' - x\|_2}{1 + \|x'\|_2}.$$

For $0 \leqslant \omega \leqslant 1$, we have $\|x' + \omega(x - x')\|_2 \leqslant \|x'\|_2 + \omega\|x' - x\|_2$. Since ϕ is nondecreasing, Lemma E.2 therefore gives

$$\widehat{\phi}\left(x' + \omega(x - x')\right) \leqslant \widehat{\phi}(x')\left(1 + \omega\tau\right)^2.$$

Integrating twice along the segment from x' to x and using the Hessian estimate from Lemma E.2 yields

$$\begin{aligned} D_{\widehat{\phi}}(x', x) &= \int_0^1 \omega\left\langle x - x', \nabla^2\widehat{\phi}\left(x' + \omega(x - x')\right)(x - x')\right\rangle d\omega \\ &\geqslant \frac{5\|x' - x\|_2^2}{6\widehat{\phi}(x')}\int_0^1 \frac{\omega}{(1 + \omega\tau)^2}\, d\omega \\ &= \frac{5(1 + \|x'\|_2)^2}{6\widehat{\phi}(x')}\left(\log(1 + \tau) - \frac{\tau}{1 + \tau}\right). \end{aligned}$$

The same calculation with $\widehat{\psi}$ gives

$$D_{\widehat{\psi}}(x', x) \geqslant \frac{\|x' - x\|_2^2}{1 + \|x'\|_2}\int_0^1 \frac{\omega}{1 + \omega\tau}\, d\omega = (1 + \|x'\|_2)\left(\tau - \log(1 + \tau)\right).$$

In particular,

$$\|x' - x\|_2 - \frac{3}{2}D_{\widehat{\psi}}(x', x) \leqslant (1 + \|x'\|_2)\left(\frac{3}{2}\log(1 + \tau) - \frac{\tau}{2}\right). \tag{27}$$

By Lemma E.2, $\widehat{\phi}(x')^2/(1+\|x'\|_2)^2 \geqslant 1/2$. Applying Lemma E.3 with $\alpha = \tau$ and $\gamma = \widehat{\phi}(x')^2/(1+\|x'\|_2)^2$ to the bound on $D_{\widehat{\phi}}(x', x)$ gives

$$\frac{\widehat{\phi}(x')^2 - \left(\widehat{\phi}(x') - D_{\widehat{\phi}}(x', x)\right)_+^2}{(1+\|x'\|_2)^2} \geqslant \frac{5}{6}\left(\frac{3}{2}\log(1+\tau) - \frac{\tau}{2}\right)_+^2.$$

Taking positive parts in (27) and using

$$\widehat{\phi}(x') - D_{\widehat{\phi}}(x', x) = \widehat{\phi}(x) + \left\langle \nabla\widehat{\phi}(x), x' - x \right\rangle$$

proves (26).

□

E.6. **Proof of Proposition 8.4.** Let $e_1, e_2 \in \mathbb{R}^d$ be orthonormal, and set

$$u_t = \cos\left(2t^{1/3}\right) e_1 + \sin\left(2t^{1/3}\right) e_2, \qquad U_t = \sum_{s=1}^{t} u_s, \qquad p_t = \sqrt{6h^*(R_t)}.$$

Since $\sum_{t=1}^{T} \|u_t\|_2^2 = T$, it suffices to prove $p_T \sim \sqrt{2}\log T/3$. Set $F(b) = \sqrt{\phi(b^2)}$ and $G(b) = 3\psi(b)/2$ for $b \geqslant 0$. For $(z, \zeta) \in \mathbb{R}^d \times \mathbb{R}$, optimizing over ξ and the direction of w gives

$$\sqrt{6h^*(z,\zeta)} = \sup_{b \geqslant 0} \frac{(\|z\|_2 b + \zeta - G(b))_+}{F(b)}. \tag{28}$$

At a state with $\rho = \|z\|_2 > 0$ and $p = \sqrt{6h^*(z,\zeta)} > 0$, the action is bz/ρ, where $b > 0$ satisfies

$$\rho = G'(b) + pF'(b), \qquad \zeta = G(b) + pF(b) - \rho b. \tag{29}$$

Convexity and the oracle identity imply that h^* is nondecreasing along every update segment. Moreover, $p_1 > 0$ by (28), since $G(b) = O(b^2)$ near zero. We shall use $G'(b) = 3\log(1+b)/2$, $G''(b) = 3/(2(1+b))$, and

$$\frac{F(b)}{b} \sim F'(b) \sim \sqrt{2\log(1+b)}, \qquad F(b)F''(b) \longrightarrow 1, \tag{30}$$

which follow by differentiation as $b \to +\infty$. Lemma E.2 also gives $F(b)F''(b) \geqslant 5/6$ and $bF'(b)/F(b) \leqslant 2$ for all $b \geqslant 0$.

To estimate U_t, identify the plane spanned by e_1, e_2 with $\mathbb{C}$ and set $f(s) = \exp\left(2\mathrm{i}s^{1/3}\right)$. Since f'' is integrable on $[1, +\infty)$, the trapezoidal formula and the substitution $s = v^3$ give

$$e^{-2\mathrm{i}t^{1/3}} \sum_{s=1}^{t} f(s) = -\frac{3\mathrm{i}}{2}t^{2/3} + \frac{3}{2}t^{1/3} + O(1).$$

Consequently, setting $\rho_t = \|U_t\|_2$ and $q_t = \langle u_t, U_{t-1}/\rho_{t-1}\rangle$ for large t,

$$\rho_t = \frac{3}{2}t^{2/3} + O(1), \qquad q_t = t^{-1/3} + O(t^{-2/3}). \tag{31}$$

Fix a sufficiently large t, so that $q_t > 0$, and, for $0 \leqslant \lambda \leqslant 1$, consider

$$(z(\lambda), \zeta(\lambda)) = R_{t-1} + \lambda\left(u_t, -\langle u_t, a_t\rangle\right).$$

Write $\rho = \|z\|_2$, $p = \sqrt{6h^*(z,\zeta)}$, and let b be the norm of the action at this state. Primes on ρ, p, b denote differentiation in λ. All estimates below are uniform in λ. Since $\rho^2 = \rho_{t-1}^2 + 2\lambda\rho_{t-1}q_t + \lambda^2$,

$$\rho \sim \frac{3}{2}t^{2/3}, \qquad \rho' \sim t^{-1/3}, \qquad \rho'' \sim \frac{2}{3}t^{-2/3}. \tag{32}$$

The identities (29) can be differentiated by the implicit function theorem, since $F > 0$ and $G'' + pF'' > 0$. As $\zeta' = -b(0)q_t$, they give

$$p' = \frac{b\rho' - b(0)q_t}{F(b)},$$
$$p'' = \frac{b\rho''}{F(b)} + \frac{(\rho' - F'(b)p')^2}{F(b)\,(G''(b) + pF''(b))}.$$

Thus $p'(0) = 0$ and $p'' \geqslant 0$, so $p' \geqslant 0$. Since $q_t > 0$ and $bF'(b)/F(b) \leqslant 2$, the first identity implies $0 \leqslant F'(b)p' \leqslant 2\rho'$, hence $|\rho' - F'(b)p'| \leqslant \rho'$.

Set $L = 1 + \log(1+b)$. By (30), Lemma E.2, and continuity, there is a constant $C > 0$, enlarged below if necessary, such that

$$\frac{b}{F(b)} \leqslant \frac{C}{\sqrt{L}}, \qquad F'(b) \leqslant C\sqrt{L},$$
$$F(b)\left(G''(b) + pF''(b)\right) \geqslant C^{-1}\left(\sqrt{L} + p\right).$$

Also, (29) gives $L \leqslant 1 + 2\rho/3$ and $\rho \leqslant C\sqrt{L}\left(\sqrt{L} + p\right)$, hence

$$\frac{1}{\sqrt{L}} \leqslant C\left(\rho^{-1/2} + p/\rho\right), \qquad \sqrt{L} + p \geqslant C^{-1}\sqrt{\rho}.$$

Together with (32), they imply

$$0 \leqslant p'' \leqslant C\frac{\rho''}{\sqrt{L}} + C\frac{(\rho')^2}{\sqrt{L} + p} \leqslant \frac{C}{t} + \frac{Cp}{t^{4/3}}. \tag{33}$$

Integrating twice and using $p(\lambda) \leqslant p_t$ gives $p_t - p_{t-1} \leqslant C/t + Cp_t/t^{4/3}$. For sufficiently large t, this implies $p_t \leqslant (1 + Ct^{-4/3})p_{t-1} + C/t$. Iteration gives $p_T = O(\log T)$, because $\prod_{t \geqslant t_0}(1 + Ct^{-4/3}) < +\infty$.

Now $L = O(t^{2/3})$ and $pF'(b) = O(t^{1/3}\log t) = o(\rho)$, so (29) gives $\log(1+b) \sim t^{2/3}$. Also, (33) gives $p'' = O(t^{-1})$ and $p' = O(t^{-1})$. Thus $F'(b)p' = o(\rho')$ and $pF(b)F''(b) = o(F(b)G''(b))$. Using (30) and (32) in the formula for p'', its two terms are each equivalent to $\sqrt{2}/(3t)$. Consequently,

$$p_t - p_{t-1} = \int_0^1 (1-\lambda)p''(\lambda)\,d\lambda \sim \frac{\sqrt{2}}{3t}.$$

Summing yields $p_T \sim \sqrt{2}\log T/3$, and hence

$$h^*(R_T) \sim \frac{(\log T)^2}{27} = \frac{(\log T)^2}{27T}\sum_{t=1}^{T} \|u_t\|_2^2.$$

E.7. **Proof of Proposition 9.4.** Fix a unit vector $e \in \mathbb{R}^d$ and set $u_t = Me$ for every $t \geqslant 1$. Then $C_t = C := 2M^2/C_0$ for $t \geqslant 2$. For $t \geqslant 1$, symmetry and $C > 0$ give $\nabla h_C^*(R_t) = (\xi_{t+1}\alpha_{t+1}e, \xi_{t+1})$ with $\xi_{t+1}, \alpha_{t+1} > 0$, so $a_{t+1} = \alpha_{t+1}e$. Set $F_t = C + h_C^*(R_t)$. Since $h_{T+1}^*(R_T) = F_T - C$ and $\sum_{t=1}^T \|u_t\|_2^2 = M^2T$, it suffices to prove $0 < F_T \leqslant (M^2/8 + o(1))\log T$.

The first-order conditions, Fenchel equality, and Theorem 9.3 give

$$Mt = 2C_0\alpha_{t+1} + 2\xi_{t+1}\alpha_{t+1}\phi'(\alpha_{t+1}^2),$$
$$0 < F_t = C\xi_{t+1} + \xi_{t+1}^2\phi(\alpha_{t+1}^2) \leqslant C + \frac{M^2t}{2}.$$

Using $\alpha_{t+1} \leqslant Mt/(2C_0)$, $\phi(v) \geqslant v$, and $\phi'(v) = O(\log(2+v))$, we obtain

$$0 \leqslant Mt - 2C_0\alpha_{t+1} \leqslant 2\sqrt{F_t}\phi'(\alpha_{t+1}^2) = O(\sqrt{t}\log t).$$

Consequently,

$$\alpha_{t+1} \sim \frac{Mt}{2C_0}, \qquad \sqrt{\phi(\alpha_{t+1}^2)} \sim \frac{Mt\sqrt{2\log t}}{2C_0}, \qquad \xi_{t+1} = O(t^{-1/2}).$$

For $t \geqslant 2$, convexity and the oracle identity give $F_t \geqslant F_{t-1}$. Keeping the Bregman divergence of $C_0 \|w\|_2^2 / \xi$ in the Fenchel identity from the proof of Proposition 9.2, and using convexity of the other terms, gives

$$F_t - F_{t-1} \leqslant \xi_{t+1} \left(M(\alpha_{t+1} - \alpha_t) - C_0(\alpha_{t+1} - \alpha_t)^2 \right) \leqslant \frac{M^2 \xi_{t+1}}{4C_0}.$$

Since $F_t \geqslant F_1 > 0$ and $\xi_{t+1} \to 0$, we have $F_t / F_{t-1} \to 1$. Therefore,

$$\begin{aligned} \sqrt{F_t} - \sqrt{F_{t-1}} &\leqslant \frac{M^2}{4C_0} \frac{\sqrt{F_t}}{(\sqrt{F_t} + \sqrt{F_{t-1}})\sqrt{\phi(\alpha_{t+1}^2)}} \\ &= (1 + o(1)) \frac{M}{4\sqrt{2}\, t\sqrt{\log t}}. \end{aligned}$$

Summing and using $\sum_{t=2}^{T} (t\sqrt{\log t})^{-1} \sim 2\sqrt{\log T}$ yields

$$\sqrt{F_T} \leqslant \left(\frac{M}{2\sqrt{2}} + o(1) \right) \sqrt{\log T},$$

as required.

Appendix F. On implementation

We consider the relaxed oracle $\mathfrak{a}$ defined by (10).

Proposition F.1. *Let $h \colon \mathbb{R}^d \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be a regularizer of the form*

$$h(w, \xi) = \begin{cases} \xi^2 \Phi\left(\|w/\xi\|_2 \right) + \xi \Psi\left(\|w/\xi\|_2 \right) + \vartheta(\xi) & \text{if } \xi > 0, \\ \vartheta(0) & \text{if } (w, \xi) = (0, 0), \\ +\infty & \text{otherwise.} \end{cases}$$

Assume that Φ, Ψ are continuously differentiable on $\mathbb{R}_+$, that ϑ is continuous on $\mathbb{R}_+$ and continuously differentiable on $\mathbb{R}_+^*$, that $\Phi'(0) = \Psi'(0) = 0$, and that $\Phi'(\alpha) > 0$ for all $\alpha > 0$. Let $(z, \zeta) \in \mathbb{R}^d \times \mathbb{R}$ satisfy

$$z = 0 \quad \text{or} \quad h^*(z, \zeta) > -\vartheta(0). \tag{34}$$

(i) If $z = 0$, then $(\mathfrak{a} \circ \nabla h^*)(z, \zeta) = 0$.

(ii) If $z \neq 0$, there exists a unique $\alpha_* > 0$ such that $\Psi'(\alpha_*) < \|z\|_2$ and

$$\frac{2\Phi(\alpha_*)\left(\|z\|_2 - \Psi'(\alpha_*) \right)}{\Phi'(\alpha_*)} + \Psi(\alpha_*) + \vartheta'\left(\frac{\|z\|_2 - \Psi'(\alpha_*)}{\Phi'(\alpha_*)} \right) = \alpha_* \|z\|_2 + \zeta.$$

Moreover,

$$(\mathfrak{a} \circ \nabla h^*)(z, \zeta) = \alpha_* \frac{z}{\|z\|_2}.$$

Proof. Since h is a regularizer, the maximization problem defining $h^*(z, \zeta)$ has a unique maximizer $(w, \xi) = \nabla h^*(z, \zeta)$. If $z = 0$, the identity $h(w, \xi) = h(-w, \xi)$ and uniqueness force $w = 0$, proving (i).

Now assume that $z \neq 0$. The inequality $h^*(z,\zeta) > -\vartheta(0)$ excludes $(w,\xi) = (0,0)$ as a maximizer, so $\xi > 0$. For fixed ξ and $\|w\|_2$, the scalar product $\langle z, w\rangle$ is largest when $w = \|w\|_2 \, z/\|z\|_2$. Writing $w = \xi\alpha z/\|z\|_2$, the expression being maximized is

$$\xi\,(\alpha\,\|z\|_2 + \zeta - \Psi(\alpha)) - \xi^2\Phi(\alpha) - \vartheta(\xi).$$

Its derivative with respect to α at zero is $\xi\,\|z\|_2 > 0$, so the maximizer has $\alpha > 0$. Differentiating with respect to α and ξ there gives

$$\|z\|_2 = \xi\Phi'(\alpha) + \Psi'(\alpha), \qquad \alpha\,\|z\|_2 + \zeta = 2\xi\Phi(\alpha) + \Psi(\alpha) + \vartheta'(\xi).$$

Thus $\xi = (\|z\|_2 - \Psi'(\alpha))/\Phi'(\alpha) > 0$, and substitution gives the equation (ii). Conversely, every positive solution satisfying $\Psi'(\alpha) < \|z\|_2$ determines a point $(w,\xi) = (\xi\alpha z/\|z\|_2\,, \xi)$ at which the gradient of $(w,\xi) \mapsto \langle z, w\rangle + \zeta\xi - h(w,\xi)$ vanishes. By strict convexity of h, this point is its unique maximizer. This proves uniqueness of α_*, and using the definition of the relaxed oracle (10) gives the expression for $(\mathfrak{a} \circ \nabla h^*)(z,\zeta)$. □

F.1. **Solving the scalar equation.** Both algorithms use ROOT below. Its input interval satisfies $0 \leqslant \alpha_- < \alpha_+$ and contains the unique zero of F in its interior.

Algorithm 1 ROOT: safeguarded Newton iteration

Input: F, F', an interval $[\alpha_-, \alpha_+]$ containing the zero of F, and a tolerance $\texttt{tol} > 0$

$\texttt{width} \leftarrow \alpha_+ - \alpha_-$
$n \leftarrow 0$
$\alpha \leftarrow \alpha_- + (\alpha_+ - \alpha_-)/2$
while $(\alpha_+ - \alpha_-)/2 > \texttt{tol}(1 + \alpha_-)$ **do**
  if $F(\alpha) = 0$ **then**
    return α
  if $F(\alpha) > 0$ **then**
    $\alpha_- \leftarrow \alpha$
  else
    $\alpha_+ \leftarrow \alpha$
  if $\alpha_+ - \alpha_- \leqslant \texttt{width}/2$ **then**
    $\texttt{width} \leftarrow \alpha_+ - \alpha_-$
    $n \leftarrow 0$
  else
    $n \leftarrow n + 1$
  $\alpha_{\mathrm{N}} \leftarrow \alpha - F(\alpha)/F'(\alpha)$
  if $n < 2$ and $\alpha_- < \alpha_{\mathrm{N}} < \alpha_+$ **then**
    $\alpha \leftarrow \alpha_{\mathrm{N}}$
  else
    $\alpha \leftarrow \alpha_- + (\alpha_+ - \alpha_-)/2$
return $\alpha_- + (\alpha_+ - \alpha_-)/2$

For both functions F below, F is continuously differentiable and $F' < 0$ in the interior of the interval used by Algorithm 1. Indeed, substituting the displayed formulas for ξ' and using the convexity of $\alpha \mapsto \sqrt{\phi(\alpha^2)}$ from Lemma E.2 gives $F' < 0$. ROOT evaluates F and F' only strictly between the current interval endpoints α_- and α_+. It does not require their values at those endpoints. The counter n records the number of evaluations of F since the interval length was last reduced by at least a factor of two. After two evaluations without such a reduction, the next trial point is the arithmetic mean of the current interval endpoints. The stopping rule requires

the interval half-length to be at most $\texttt{tol}(1+\alpha_-)$. We use double precision, with $\texttt{tol} = 10^{-12}$ as a possible choice. Each online round uses $O(d)$ operations in addition to the scalar root-finding iterations, and $O(d)$ storage.

The derivatives needed in both cases are

$$\phi'(v) = 1 + \log(1+v) + \frac{v}{1+v}, \qquad \phi''(v) = \frac{v+2}{(1+v)^2}.$$

F.2. **Bounded payoff vectors.** For the updates in (12), convexity gives

$$h^*(R_t) \geqslant h^*(R_{t-1}) + \langle \nabla h^*(R_{t-1}), r_t \rangle = h^*(R_{t-1}).$$

At the first nonzero payoff vector, $\psi(0) = \psi'(0) = 0$ gives $h^*(R_t) > 0$. Since $\vartheta(0) = h(0,0) = 0$, condition (34) of Proposition F.1, namely $z = 0$ or $h^*(z,\zeta) > 0$, holds at every round.

For $z \neq 0$, consider

$$\begin{aligned}
\xi(\alpha) &= \frac{2\|z\|_2/3 - \log(1+\alpha)}{2\alpha\phi'(\alpha^2)}, \\
\xi'(\alpha) &= -\frac{(1+\alpha)^{-1} + \xi(\alpha)\left(2\phi'(\alpha^2) + 4\alpha^2\phi''(\alpha^2)\right)}{2\alpha\phi'(\alpha^2)}, \\
F(\alpha) &= 3\phi(\alpha^2)\xi(\alpha) + \frac{3}{2}\psi(\alpha) - \alpha\|z\|_2 - \zeta, \\
F'(\alpha) &= 3\alpha\phi'(\alpha^2)\xi(\alpha) + 3\phi(\alpha^2)\xi'(\alpha).
\end{aligned}$$

For this regularizer, $\Psi(\alpha) = 3\psi(\alpha)/2$, so $\Psi'(\alpha) = 3\log(1+\alpha)/2$. By Proposition F.1, F has a unique zero $\alpha_* > 0$ satisfying $\log(1+\alpha_*) < 2\|z\|_2/3$. Thus α_* belongs to $(0, \exp(2\|z\|_2/3) - 1)$, which is precisely the region where $\alpha > 0$ and $\xi(\alpha) > 0$.

To avoid cancellation for small α, evaluate $\psi(\alpha)$ with $\sum_{j=2}^{10} \frac{(-1)^j\alpha^j}{j(j-1)}$ when $\alpha \leqslant 10^{-2}$, and with $(1+\alpha)\log(1+\alpha) - \alpha$ otherwise.

The dedicated evaluations of $\log(1+v)$ and $\exp(v) - 1$ used in Algorithms 2 and 4 preserve accuracy near zero: forming $1+v$ first can round away a small v, and subtracting one after exponentiation can lose significant digits.

Algorithm 2 RootBounded: finding the positive scalar

Input: $z \neq 0$, ζ with $h^*(z,\zeta) > 0$, $\texttt{tol} > 0$, and F, F' defined above
In F, F', use $\mathrm{log1p}(v)$ for $\log(1+v)$
$\alpha_- \leftarrow 0$
$\alpha_+ \leftarrow \min\{1, \|z\|_2/3\}$
loop
  if $F(\alpha_+) = 0$ **then**
    return α_+
  if $F(\alpha_+) < 0$ **then**
    break
  $\alpha_- \leftarrow \alpha_+$
  if $\mathrm{log1p}(2\alpha_+) < 2\|z\|_2/3$ **then**
    $\alpha_+ \leftarrow 2\alpha_+$
  else
    $\alpha_+ \leftarrow \mathrm{expm1}(2\|z\|_2/3)$
    break
return $\textsc{Root}(F, F', \alpha_-, \alpha_+, \texttt{tol})$

Doubling finds a useful interval before calling ROOT. The exponential endpoint is formed only when it is at most twice the current α_+.

Algorithm 3 Practical implementation for bounded payoff vectors

Input: payoff vectors with $\|u_t\|_2 \leqslant 1$ and a tolerance $\texttt{tol} > 0$
$(z, \zeta) \leftarrow (0, 0)$
for $t = 1, 2, \ldots$ **do**
  if $z = 0$ **then**
    $a_t \leftarrow 0$
  else
    $\alpha \leftarrow \textsc{RootBounded}(z, \zeta, \texttt{tol})$
    $a_t \leftarrow \alpha z / \|z\|_2$
  Play a_t
  Observe u_t
  $(z, \zeta) \leftarrow (z + u_t, \zeta - \langle u_t, a_t \rangle)$

F.3. **Unbounded payoff vectors.** In the algorithm from Section 9, $C_t = 0$ implies that every preceding payoff vector is zero. Thus $z \neq 0$ implies $C_t > 0$. Testing the supremum defining $h^*_C(z, \zeta)$ at $(0, \xi)$ for sufficiently small $\xi > 0$ gives $h^*_C(z, \zeta) > -C$ whenever $C > 0$. Since $\vartheta(0) = C$, condition (34) of Proposition F.1, namely $z = 0$ or $h^*_C(z, \zeta) > -C$, holds at every round.

Whenever ROOTUNBOUNDED is called, $C = C_t > 0$. Set

$$\begin{aligned}
\xi(\alpha) &= \frac{\|z\|_2 - 2C_0\alpha}{2\alpha\phi'(\alpha^2)}, \\
\xi'(\alpha) &= -\frac{2C_0 + \xi(\alpha)\left(2\phi'(\alpha^2) + 4\alpha^2\phi''(\alpha^2)\right)}{2\alpha\phi'(\alpha^2)}, \\
F(\alpha) &= 2\phi(\alpha^2)\xi(\alpha) + C_0\alpha^2 + C(1 + \log\xi(\alpha)) - \alpha\|z\|_2 - \zeta, \\
F'(\alpha) &= 2\alpha\phi'(\alpha^2)\xi(\alpha) + \left(2\phi(\alpha^2) + C/\xi(\alpha)\right)\xi'(\alpha).
\end{aligned}$$

For this regularizer, $\Psi(\alpha) = C_0\alpha^2$, so $\Psi'(\alpha) = 2C_0\alpha$. By Proposition F.1, F has a unique zero $\alpha_* > 0$ satisfying $2C_0\alpha_* < \|z\|_2$. Thus α_* belongs to $(0, \|z\|_2/(2C_0))$, which is precisely the region where $\alpha > 0$ and $\xi(\alpha) > 0$.

To avoid taking the logarithm of a value of $\xi(\alpha)$ that may overflow or underflow, compute it from the factors in the definition of $\xi(\alpha)$:

$$\log\xi(\alpha) = \log\|z\|_2 + \log\left(1 - \frac{2C_0\alpha}{\|z\|_2}\right) - \log 2 - \log\alpha - \log\phi'(\alpha^2).$$

Algorithm 4 RootUnbounded: finding the positive scalar

Input: $z \neq 0$, ζ, $C > 0$, $C_0 > 0$, $\texttt{tol} > 0$, and F, F' defined above

In F, F', use $\mathrm{log1p}(v)$ for $\log(1+v)$
$\alpha_- \leftarrow 0$
$\alpha_+ \leftarrow \min\{1, \|z\|_2 / (4C_0)\}$
loop
  if $F(\alpha_+) = 0$ **then**
    return α_+
  if $F(\alpha_+) < 0$ **then**
    break
  $\alpha_- \leftarrow \alpha_+$
  if $2\alpha_+ < \|z\|_2 / (2C_0)$ **then**
    $\alpha_+ \leftarrow 2\alpha_+$
  else
    $\alpha_+ \leftarrow \|z\|_2 / (2C_0)$
    break
return $\textsc{Root}(F, F', \alpha_-, \alpha_+, \texttt{tol})$

Algorithm 5 Practical implementation for unbounded payoff vectors

Input: $C_0 > 0$, payoff vectors u_t, and a tolerance $\texttt{tol} > 0$

$(z, \zeta) \leftarrow (0, 0)$
$C_1 \leftarrow 0$
for $t = 1, 2, \ldots$ **do**
  if $z = 0$ **then**
    $a_t \leftarrow 0$
  else
    $\alpha \leftarrow \textsc{RootUnbounded}(z, \zeta, C_t, C_0, \texttt{tol})$
    $a_t \leftarrow \alpha z / \|z\|_2$
  Play a_t
  Observe u_t
  $(z, \zeta) \leftarrow (z + u_t, \zeta - \langle u_t, a_t \rangle)$
  $C_{t+1} \leftarrow \max\left\{C_t, 2\|u_t\|_2^2 / C_0\right\}$