Proceedings of Machine Learning Research vol 237:1–26, 2024 35th International Conference on Algorithmic Learning Theory

Improving Adaptive Online Learning Using Refined Discretization

Zhiyu Zhang ZHIYUZ@SEAS.HARVARD.EDU
Harvard University

Heng Yang HANKYANG@SEAS.HARVARD.EDU
Harvard University

Ashok Cutkosky ASHOK@CUTKOSKY.COM
Boston University

Ioannis Ch. Paschalidis YANNISP@BU.EDU
Boston University

Editors: Claire Vernade and Daniel Hsu

Abstract

We study unconstrained Online Linear Optimization with Lipschitz losses. The goal is to simultaneously achieve (i) second order gradient adaptivity; and (ii) comparator norm adaptivity also known as "parameter freeness" in the literature. Existing regret bounds (Cutkosky and Orabona, 2018; Mhammedi and Koolen, 2020; Jacobsen and Cutkosky, 2022) have the suboptimal $O(\sqrt{V_T \log V_T})$ dependence on the gradient variance V_T, while the present work improves it to the optimal rate $O(\sqrt{V_T})$ using a novel continuous-time-inspired algorithm, without any impractical doubling trick. This result can be extended to the setting with unknown Lipschitz constant, eliminating the range ratio problem from prior works (Mhammedi and Koolen, 2020).

Concretely, we first show that the aimed simultaneous adaptivity can be achieved fairly easily in a continuous time analogue of the problem, where the environment is modeled by an arbitrary continuous semimartingale. Then, our key innovation is a new discretization argument that preserves such adaptivity in the discrete time adversarial setting. This refines a non-gradient-adaptive discretization argument from (Harvey et al., 2023), both algorithmically and analytically, which could be of independent interest.[1]

Keywords: adaptive online learning, scale-free online learning, continuous time method

1. Introduction

We study unconstrained *Online Linear Optimization* (OLO) with Lipschitz losses,[2] which is a repeated game between us (the learner) and an adversarial environment denoted by Env. In each (the t-th) round, with a mutually known Lipschitz constant G:

1. We make a decision $x_t \in \mathbb{R}^d$ based on the observations before the t-th round.
2. The environment Env reveals a loss gradient $g_t \in \mathbb{R}^d$ dependent on our decision history $x_1, \ldots, x_t$, which satisfies the Lipschitz condition with respect to the Euclidean norm, $\|g_t\| \leq G$.
3. We suffer the linear loss $\langle g_t, x_t \rangle$.

1. Future versions available at `https://arxiv.org/abs/2309.16044`.
2. In a black-box manner, solutions of this problem can also solve bounded domain *Online Convex Optimization* (OCO) with Lipschitz losses, as shown in (Orabona, 2023, Section 2.3) and (Cutkosky, 2020, Section 4).

The game ends after T rounds, and then, our total loss is compared to that of an arbitrary fixed decision $u \in \mathbb{R}^d$. Without knowing the time horizon T, the environment Env and the comparator u, the goal is to guarantee low *regret*, defined as

$$\mathrm{Regret}_T(Env, u) := \sum_{t=1}^{T} \langle g_t, x_t - u \rangle .$$

In a nutshell, the present work uses a novel and practical strategy to achieve the tightest known regret upper bound (even including a near-optimal leading constant) that depends simultaneously on the loss gradients $g_1, \ldots, g_T$ and the comparator u.

To be concrete, let us survey a bit more of the context. Existing research on regret minimization started from the minimax regime: under the additional assumption of $\|u\| \leq D$, it has been long known that *Online Gradient Descent* (OGD) (Zinkevich, 2003) guarantees the optimal upper bound on the *worst case regret*, $\sup_{Env; \|u\| \leq D} \mathrm{Regret}_T(Env, u) \leq O\left(DG\sqrt{T}\right)$. Refining such worst case optimality by instance optimality, improvements have been achieved under the notion of *adaptive online learning*, with *gradient adaptivity* and *comparator adaptivity* being the two prominent types.

- Gradient adaptivity aims at bounding $\sup_{\|u\| \leq D} \mathrm{Regret}_T(Env, u)$ by a function of the observed gradient sequence $g_1, \ldots, g_T$. Using learning rates dependent on past observations, OGD can achieve the optimal second order gradient adaptive bound (McMahan and Streeter, 2010; Duchi et al., 2011)
$$\sup_{\|u\| \leq D} \mathrm{Regret}_T(Env, u) \leq O\left(D\sqrt{V_T}\right), \tag{1}$$
where $V_T := \sum_{t=1}^{T} \|g_t\|^2$ is the (uncentered) *gradient variance*. This has been a hallmark of practical online learning algorithms, popularized by the massive success of ADAGRAD (Duchi et al., 2011).

- Comparator adaptivity aims at bounding $\sup_{Env} \mathrm{Regret}_T(Env, u)$ by a function of the comparator u. Without imposing the extra bounded-u assumption, one could use a dual space framework to achieve the optimal bound (McMahan and Orabona, 2014; Zhang et al., 2022a)
$$\sup_{Env} \mathrm{Regret}_T(Env, u) \leq O\left(\|u\| G\sqrt{T \log \|u\|}\right). \tag{2}$$
Due to the absence of learning rates, such algorithms are also called "parameter-free" (Orabona and Pál, 2016) in the literature. They have exhibited the potential to reduce hyperparameter tuning in the modern deep learning workflow (Orabona and Tommasi, 2017; Cutkosky et al., 2023).

While both types of adaptivity are well-studied separately, achieving them simultaneously is an active research direction, which we call *simultaneous adaptivity*. A series of works (Cutkosky and Orabona, 2018; Mhammedi and Koolen, 2020; Jacobsen and Cutkosky, 2022) proposed drastically different approaches to obtain simultaneously adaptive regret bounds like[3]

$$\mathrm{Regret}_T(Env, u) \leq O\left(\|u\| \sqrt{V_T \log(G^{-2} \|u\| V_T)}\right), \tag{3}$$

3. Omitting an additive $O\left(\|u\| G \log(G^{-2} \|u\| V_T)\right)$ term for clarity. Essentially, it means the order of V_T is considered "more important" than the order of $\|u\|$, which fits into the convention of the field.

whose dependence on the gradient variance V_T alone is $O\left(\sqrt{V_T \log V_T}\right)$ rather than the standard optimal rate $O\left(\sqrt{V_T}\right)$ from Eq.(1). Roughly speaking (with subtleties explained in Appendix A), Eq.(3) can be improved to the optimal rate $O\left(\|u\| \sqrt{V_T \log \|u\|}\right)$ through the classical *doubling trick* – restarting the algorithm with a doubling "confidence hyperparameter" whenever the observed gradient variance exceeds a doubling threshold. However, such a restarting scheme is notoriously impractical (also explained in Appendix A), thus doing so will violate the key practical considerations that motivated adaptive online learning in the first place. The first goal of this paper, on the quantitative side, is to achieve the optimal $O\left(\|u\| \sqrt{V_T \log \|u\|}\right)$ bound without the doubling trick.

To this end, we will take a detour through the *continuous time* (CT), first solving a CT analogue of the problem, and then converting the solution back to *discrete time* (DT). Quantitatively, our goal above can be seen as the gradient adaptive refinement of (Zhang et al., 2022a) – without considering gradient adaptivity, the latter showed that an algorithm designed in CT natively achieves the optimal comparator adaptive bound, Eq.(2), while earlier algorithms designed in DT relied on the doubling trick. Broadly speaking, such a result exemplifies a higher level observation: while various benefits of the CT approach have been demonstrated in online learning before (Kapralov and Panigrahy, 2011; Drenska and Kohn, 2020; Kobzar et al., 2020; Zhang et al., 2022b; Harvey et al., 2023), it remains somewhat unsatisfactory that no existing work (to the best of our knowledge) used it to obtain DT gradient adaptive regret bounds, even though the CT analogue of gradient adaptivity is often natural[4] and fairly standard to achieve (Freund, 2009; Harvey et al., 2023). In other words, one would expect the CT approach to make gradient adaptivity easier as well, but such a benefit has not been demonstrated in the literature.

The key reason of this limitation appears to be the crudity of existing *discretization arguments*, i.e., the modification applied to a CT algorithm and its analysis to make them work well in DT. The state-of-the-art technique, due to (Harvey et al., 2023), replaces the continuous derivative in potential-based CT algorithms[5] by the discrete derivative, and consequently, the standard *Itô's formula* in the CT regret analysis by the *discrete Itô's formula*. Applying the discrete derivative amounts to implicitly assuming the worst case gradient magnitude ($\|g_t\| = G$), therefore any gradient adaptivity in CT is lost in DT by construction. The second goal of this paper, on the technical side, is to propose a refined discretization argument that preserves such gradient adaptivity.

1.1. Contribution

We first show that in a CT analogue of our OLO problem, simultaneous adaptivity is actually easy to obtain by combining Itô's formula and the *Backward Heat Equation* (BHE). Building on this intuition, our main result is a new DT algorithm achieving the following regret bound without the doubling trick. With an arbitrary hyperparameter $\varepsilon > 0$, in the asymptotic regime of large $\|u\|$ and V_T,

$$\text{Regret}_T(Env, u) \leq \varepsilon \cdot O\left(\sqrt{V_T}\right) + \|u\| \cdot O\left(\sqrt{V_T \log(\|u\| \varepsilon^{-1})} \vee G \log(\|u\| \varepsilon^{-1})\right).$$

This is the first simultaneously adaptive regret bound matching the optimal $O(\sqrt{V_T})$ rate (with respect to V_T alone), improving a series of prior works (Cutkosky and Orabona, 2018; Mhammedi

4. Example: in finance, the gradient variance is analogous to the *price volatility*. This is ubiquitous in the continuous time modeling of financial instruments, such as the *geometric Brownian motion*.
5. A perhaps better-known name is *Follow the Regularized Leader* (FTRL) (Abernethy et al., 2008).

and Koolen, 2020; Jacobsen and Cutkosky, 2022). Furthermore, given any hyperparameter $\alpha > \frac{1}{2}$, the multiplying constant on the leading order term $\|u\| \sqrt{V_T \log(\|u\| \varepsilon^{-1})}$ is $\sqrt{4\alpha}$, almost[6] matching the $\sqrt{2}$ lower bound from (Zhang et al., 2022a). This is the first result characterizing the leading constant optimality in simultaneously adaptive online learning.

In addition, we generalize the above to the setting without a known Lipschitz constant G, making the algorithm and its regret bound *scale-free*. Since the hyperparameter ε is truly "unitless" in our algorithm design, there is no need to estimate the scale of the loss gradients ($\max_{t\in[1:T]} \|g_t\|$) at the beginning, which eliminates the *range ratio problem* from existing solutions (Mhammedi and Koolen, 2020; Jacobsen and Cutkosky, 2022).[7] This keeps the algebra simple, while also avoiding the standard range-ratio-induced penalties in the scale-free regret bound.

Technically, our key innovation is a new gradient adaptive discretization argument, refining the non-adaptive one from (Harvey et al., 2023). The essential idea is connecting the CT algorithm and its DT analogue via a *change of variables*, which allows using the *exact* BHE from CT to simplify the complicated algebra in DT. For our specific problem of simultaneous adaptivity, this procedure is arguably easier and more intuitive than existing approaches (Cutkosky and Orabona, 2018; Mhammedi and Koolen, 2020; Jacobsen and Cutkosky, 2022) that tackle DT directly.

1.2. Notation

Let $C^{1,2}(\mathcal{X})$ be the class of bivariate functions on an open set $\mathcal{X}$, continuously differentiable in their first argument and twice continuously differentiable in their second argument. For any $\Phi \in C^{1,2}(\mathcal{X})$, let $\partial_1\Phi$ and $\partial_2\Phi$ be its first order partial derivatives with respect to the first and the second argument of Φ. Similarly, $\partial_{11}\Phi$, $\partial_{12}\Phi$ and $\partial_{22}\Phi$ denote the second order partial derivatives ($\partial_{12}\Phi = \partial_{21}\Phi$). For all x and u, define $\Phi^*_x(z) := \sup_y[zy - \Phi(x,y)]$, i.e., the Fenchel conjugate of Φ with respect to its second argument.

We define the *imaginary error function* as $\mathrm{erfi}(x) = \int_0^x \exp(u^2)du$; this is scaled by $\sqrt{\pi}/2$ from the conventional definition, thus can also be queried from standard software packages like SCIPY and JAX. Let erfi^{-1} be its inverse function.

$\Pi_{\mathcal{X}}(x)$ is the Euclidean projection of x onto a closed convex set $\mathcal{X}$. $\log$ represents natural logarithm when the base is omitted. Throughout this paper, $\|\cdot\|$ denotes the Euclidean norm.

2. Related work

To streamline the exposition, a detailed discussion of existing works on simultaneously adaptive online learning is deferred to Appendix A.

Continuous time approach Our result fits into an emerging direction of online learning: exploiting the synergy between CT and DT algorithms. Concrete benefits in DT, including better bounds and simpler analyses, have been demonstrated in various settings of minimax online learning (Bayraktar et al., 2020; Drenska and Kohn, 2020; Kobzar et al., 2020; Wang and Kohn, 2022; Harvey et al., 2023) and adaptive online learning (Kapralov and Panigrahy, 2011; Daniely and Mansour, 2019; Portella

6. There is an additive factor on V_T proportional to $(\alpha - \frac{1}{2})^{-1}G^2$. For any fixed $\alpha > \frac{1}{2}$ this additive factor is negligible in the large-V_T regime, but we cannot pick $\alpha = \frac{1}{2}$.

7. Clarification: existing works (Mhammedi and Koolen, 2020; Jacobsen and Cutkosky, 2022) proposed various techniques to mitigate the range ratio problem, while our approach does not encounter this problem by default. See Section 5 for a detailed explanation.

et al., 2022; Zhang et al., 2022a,b). Conversely, such a synergy can benefit traditional "model-based" CT decision making as well; for example, Abernethy et al. (2012, 2013) showed that the celebrated *Black-Scholes model* (Black and Scholes, 1973) for option pricing can be derived as the scaling limit of a DT adversarial online learning model, which provides a strong justification of its validity.

Most of these works established the synergy via potential-based algorithms. Roughly speaking,[8] the decision at time t has the form $\phi_t'(S_t)$, which denotes the derivative of a potential function ϕ_t evaluated at a "sufficient statistic" S_t that summarizes the history. In the CT regime, the crucial simplicity is that a suitable ϕ_t satisfies a *Partial Differential Equation* (PDE), thus finding it can be a tractable task. The tricky step is to properly convert this CT algorithm to DT and quantify their performance discrepancy, which we call the discretization argument.

The most natural idea is to apply the CT algorithm to DT as is, and characterize the performance discrepancy using Taylor's theorem (Abernethy et al., 2013; Kobzar et al., 2020). However, this approach requires a terminal condition at a fixed time horizon T, which is missing from many common settings of adaptive online learning. Harvey et al. (2023) proposed a particularly strong and elegant alternative: replacing the standard derivative $\phi_t'(S_t)$ in the CT algorithm by the discrete derivative, e.g., $\frac{1}{2G}[\phi_t(S_t+G)-\phi_t(S_t-G)]$. Then, the performance discrepancy between CT and DT can be characterized by a DT analogue of the PDE, which can be analyzed in a principled manner. Such an analysis has been adopted in several recent works (Greenstreet et al., 2022; Portella et al., 2022; Zhang et al., 2022a,b), but the downside is that any gradient adaptive upgrade on the CT algorithm (not hard to obtain, as shown in Section 3) is lost in DT by construction. The present work addresses this limitation.

3. Warm up: Adaptivity in continuous time

To begin with, we study a one dimensional continuous time analogue of the unconstrained OLO problem, in order to demonstrate the inherent simplicity of simultaneous adaptivity. The restriction to 1D is justified by a well-known polar-decomposition technique from (Cutkosky and Orabona, 2018), which will be made concrete in Section 4.

Technically, much of this section is standard: the critical use of Itô's formula in CT online learning was pioneered by Freund (2009) and greatly streamlined by Harvey et al. (Harvey et al., 2023, Appendix B). Our treatment of simultaneous adaptivity will follow Harvey et al.'s argument and a classical loss-regret duality from (McMahan and Orabona, 2014). Nonetheless, it embodies the key intuition, thus paves the way for our main contribution on the discretization argument.

3.1. Setting

Unlike the DT adversarial setting universally recognized as a repeated game, the definition of a reasonable CT analogue has been elusive. First, following (Harvey et al., 2023), we model the combined actions of the CT environment, i.e., the CT analogue of the gradient sum $\sum_{i=1}^{t} g_i$, as an arbitrary *continuous semimartingale* denoted by S_t ($t \in \mathbb{R}_{\geq 0}$),[9] with $S_0 = 0$ – examples include

8. For clarity, we do not account for gradient adaptivity in this discussion. Essentially, the idea of second order gradient adaptivity is replacing t by the running gradient variance $V_t = \sum_{i=1}^{t} \|g_i\|^2$.

9. In DT, the gradient sum $\sum_{i=1}^{t} g_i$ is a classical quantity for dual space online learning algorithms, sometimes called the "sufficient statistic". In the CT analogue, we essentially assume the sufficient statistic evolves as a stochastic process with a very general law. Our treatment has a slight difference from (Harvey et al., 2023): the latter models $|S_t|$ rather than S_t.

the *Brownian motion* and the *Itô process*.[10] Such an assumption is motivated by the analysis: semimartingales form the largest class of integrators with respect to which the Itô integral can be defined; "continuous" means that the sample paths are continuous almost surely – this avoids the jump-correction terms in the Itô's formula.

Next, for any continuous semimartingale S, one can define its *quadratic variation process*, denoted by $[S]_t$. Intuitively, $[S]_t$ is the CT analogue of the 1D gradient variance $V_t = \sum_{i=1}^t g_i^2$. For the standard Brownian motion, $[S]_t = t$. For the Itô's process, $[S]_t = \int_0^t \sigma_s^2 \,\mathrm{d}s$, where σ is the diffusion coefficient of S (note that σ is itself a stochastic process).

After characterizing the CT environment, let us turn to the CT learner. We consider the potential framework. The learner fixes a potential function $\phi \in C^{1,2}(\mathcal{X})$ at the beginning, where $\mathcal{X}$ is an open set containing $\mathbb{R}_{\geq 0} \times \mathbb{R}$. The learner's decision against the environment S is $\partial_2\phi([S]_t, -S_t)$, which is a continuous process – this mirrors the standard FTRL family in DT (Orabona, 2023, Chapter 7). Furthermore, the "adversarialness" of the DT setting is analogous to the fact that the law of the CT environment S can depend on the learner's potential function ϕ.

With the above, the learner's total loss is the Itô integral $\int_0^T \partial_2\phi([S]_t, -S_t)\,\mathrm{d}S_t$ (Revuz and Yor, 2013, Definition IV.2.9), and any fixed comparator u induces the total loss uS_T. The goal of this CT problem is thus choosing ϕ to minimize the *continuous time regret*,

$$\mathrm{Regret}_T^{\mathrm{CT}}(Env, u) := \left(\int_0^T \partial_2\phi([S]_t, -S_t)\,\mathrm{d}S_t\right) - uS_T.$$

3.2. Analysis

The crucial simplicity of the CT setting can be seen in the following theorem. This is new to the literature, but just a combination of steps in (Harvey et al., 2023, Theorem B.2) and (Orabona, 2023, Theorem 9.6).

Theorem 1 *If $\phi \in C^{1,2}(\mathcal{X})$ satisfies the Backward Heat Equation (BHE) $\partial_1\phi + \frac{1}{2}\partial_{22}\phi = 0$, then for all $T \in \mathbb{R}_{\geq 0}$ and $u \in \mathbb{R}$, almost surely,*

$$\mathrm{Regret}_T^{\mathrm{CT}}(Env, u) \leq \phi(0,0) + \phi^*_{[S]_T}(u).$$

Here we follow the notation from Section 1.2: $\phi^*_\cdot(\cdot)$ is the Fenchel conjugate of ϕ with respect to its second argument.

Let us include the proof for completeness, which also highlights the ideal type of analysis that the DT regime should also follow. The central component is the Itô's formula, i.e., the stochastic analogue of the chain rule. The specific version below combines two results from (Revuz and Yor, 2013): Proposition IV.1.18, and Remark 1 after Theorem IV.3.3.

Lemma 2 (Itô's formula) *If $f \in C^{1,2}(\mathcal{X})$ and X is a continuous semimartingale, then for all $T \in \mathbb{R}_{\geq 0}$, almost surely,*

$$\begin{aligned} &f\left([X]_T, X_T\right) - f\left(0, X_0\right) \\ &\qquad = \int_0^T \partial_2 f\left([X]_t, X_t\right)\mathrm{d}X_t + \int_0^T \left[\partial_1 f\left([X]_t, X_t\right) + \frac{1}{2}\partial_{22} f\left([X]_t, X_t\right)\right]\mathrm{d}[X]_t. \end{aligned}$$

10. Itô's process is a general form of diffusion defined by a differential equation $\mathrm{d}S_t = \sigma_t\,\mathrm{d}B_t + \mu_t\,\mathrm{d}t$, where B is the standard Brownian motion. Here, the *diffusion coefficient* σ and the *drift coefficient* μ are both stochastic processes. Rigorous definitions of relevant stochastic process concepts can be found in (Revuz and Yor, 2013).

In the language of the potential framework, the Itô's formula is a potential verification argument (on f), but in the strongest form: an equality. The regret bound can then be proved with ease.

Proof [Proof of Theorem 1] Applying Lemma 2 with $f \leftarrow -\phi$ and $X \leftarrow -S$, and further using the BHE $\partial_1\phi + \frac{1}{2}\partial_{22}\phi = 0$ to eliminate the integral with respect to $[S]$, we have

$$\begin{aligned}\mathrm{Regret}_T^{\mathrm{CT}}(Env, u) &= \phi\left(0,0\right) - uS_T - \phi\left([S]_T, -S_T\right)\\ &\leq \phi\left(0,0\right) + \sup_{y\in\mathbb{R}}\left[uy - \phi\left([S]_T, y\right)\right].\end{aligned}$$

The proof is complete by plugging in the definition of the Fenchel conjugate $\phi^*_{[S]_T}(u)$. ■

It remains to pick a specific potential function ϕ. Similar to (Zhang et al., 2022a; Harvey et al., 2023), with arbitrary constants $\varepsilon, \delta > 0$, we define

$$\phi^{\mathrm{CT}}(x, y) = \varepsilon\sqrt{x+\delta}\left(2\int_0^{\frac{y}{\sqrt{2(x+\delta)}}} \mathrm{erfi}(u)du - 1\right), \tag{4}$$

which satisfies the BHE. Also, the shift δ on the first argument x ensures $\phi^{\mathrm{CT}} \in C^{1,2}(\mathbb{R}_{>-\delta} \times \mathbb{R})$. Plugging in the Fenchel conjugate computation from (Zhang et al., 2022a, Theorem 4), Theorem 1 becomes

$$\mathrm{Regret}_T^{\mathrm{CT}}(Env, u) \leq \varepsilon\sqrt{[S]_T + \delta} + |u|\sqrt{2\left([S]_T + \delta\right)}\left[\sqrt{\log\left(1 + \frac{|u|}{\sqrt{2}\varepsilon}\right)} + 1\right]. \tag{5}$$

With $\varepsilon = 1$, $\mathrm{Regret}_T^{\mathrm{CT}}(Env, u) = O\left(|u|\sqrt{[S]_T \log|u|}\right)$, which is the desirable CT simultaneously adaptive bound, analogous to the $O\left(|u|\sqrt{V_T \log|u|}\right)$ bound we aim for in DT. Furthermore, the leading constant $\sqrt{2}$ in Eq.(5) matches the optimal leading constant in the DT setting (Zhang et al., 2022a).

To conclude, the key takeaway is that in CT, one can use the Itô's formula and the Backward Heat Equation to achieve an "ideal form" of simultaneous adaptivity fairly easily. In some sense, they capture the important problem structure, which suggests that their DT analogues could potentially improve and simplify prior works (Cutkosky and Orabona, 2018; Mhammedi and Koolen, 2020; Jacobsen and Cutkosky, 2022) that do not exploit such structures. Next, we make this intuition concrete.

4. Main result: Refined discretization

In this section, we consider a DT setting that slightly generalizes the one at the beginning of this paper. Let us assume the Lipschitz constant G is unknown, but at the beginning of each (the t-th) round we have access to a *hint* h_t which satisfies $h_t \geq h_{t-1}$ and $\|g_t\| \leq h_t$ (initially, assume $h_0 = h_1$ and $\|g_1\| > 0$ w.l.o.g.). Such a setting is motivated by (Cutkosky, 2019), where designing a full Lipschitzness-adaptive algorithm can be reduced to solving this OLO problem with hints; details will be discussed in Section 5. For clarity, one may think of $h_t = G$ when G is known.

Our solution centers around a 1D potential function Φ_t defined by a *change of variables*. With hyperparameters $\varepsilon, \alpha, k_t > 0$ and $z_t > k_t h_t$, define

$$\Phi_t(V, S) := \phi(V + z_t + k_t S, S), \tag{6}$$

where

$$\phi(x,y) := \varepsilon\sqrt{\alpha x}\left(2\int_0^{\frac{y}{\sqrt{4\alpha x}}} \mathrm{erfi}(u)du - 1\right).$$

Here, ϕ is a generalization of the CT potential (4), satisfying $\partial_1\phi + \alpha\partial_{22}\phi = 0$, the generalized Backward Heat Equation with constant α. It can be verified that $\phi \in C^{1,2}(\mathbb{R}_{>0}\times\mathbb{R})$.

Intuitively, Φ_t is the potential function we apply in the t-th round, therefore k_t and z_t should be functions of the hint h_t. By a simple dimensional analysis, $z_t \propto h_t^2$, $k_t \propto h_t$, while ε and α are real numbers. Also, the definition of $\Phi_t(V,S)$ is only valid when S is larger than a threshold, since the first argument of ϕ can only be positive – the choice of z_t and k_t will ensure that all the "interesting" values of S are above this threshold.

4.1. Algorithm

Similar to many other comparator adaptive OLO algorithms, our algorithm has a two-level hierarchical structure. On the high level is the meta algorithm from (Cutkosky and Orabona, 2018, Section 3), decomposing the OLO problem on $\mathbb{R}^d$ into two independent subtasks: learning the direction and the magnitude of the comparator. Direction learning is handled by gradient adaptive OGD on a unit norm ball. Magnitude learning is handled by the novel 1D base algorithm employing the potential function Φ_t, followed by a constraint-imposing technique (Cutkosky, 2020, Section 4) which restricts its output from $\mathbb{R}$ to $\mathbb{R}_{\geq 0}$.

Concretely, we present the important 1D base algorithm as Algorithm 1, while the meta algorithm is presented as Algorithm 2. Since the base algorithm is updated using the surrogate feedback provided by the meta algorithm, we denote these surrogate algorithmic quantities with tilde.

Algorithm 1 1D base algorithm

Require: The potential function Φ_t defined in Eq.(6). Hints $h_1, h_2, \ldots \in \mathbb{R}_{>0}$ satisfying $h_t \geq h_{t-1}$. Surrogate loss gradients $\tilde{l}_1, \tilde{l}_2, \ldots$ satisfying $\tilde{l}_t \in [-h_t, h_t]$ and $\sum_{i=1}^t \tilde{l}_i \leq h_t$ for all t.

1: Initialize $\tilde{V}_0 = 0$, $\tilde{S}_0 = 0$.
2: **for** $t = 1, 2, \ldots$ **do**
3: Observe the hint h_t and use it to define k_t and z_t in the potential function Φ_t.
4: Predict $\tilde{y}_t = \partial_2\Phi_t(\tilde{V}_{t-1}, \tilde{S}_{t-1})$.
5: Receive the surrogate loss gradient $\tilde{l}_t$.
6: Let $\tilde{V}_t = \tilde{V}_{t-1} + \tilde{l}_t^2$, and $\tilde{S}_t = \tilde{S}_{t-1} - \tilde{l}_t$.
7: **end for**

Notice that Algorithm 1 requires a somewhat nonstandard condition, $\sum_{i=1}^t \tilde{l}_i \leq h_t$ for all t. This is to ensure that the update $\tilde{y}_t = \partial_2\Phi_t(\tilde{V}_{t-1}, \tilde{S}_{t-1})$ is well-defined: $\tilde{S}_{t-1} = -\sum_{i=1}^t \tilde{l}_i \geq -h_t$, therefore with $z_t > k_t h_t$, we always have $\tilde{V}_{t-1} + z_t + k_t\tilde{S}_{t-1} > 0$, which complies with the positivity requirement on the first argument of ϕ. The following lemma, proved in Appendix D, shows that the surrogate losses defined by the meta algorithm indeed satisfy this requirement, thus the entire algorithm procedure is well-posed.

Lemma 3 (Well-posedness) *The surrogate loss $\tilde{l}_t$ defined in Algorithm 2 satisfies $\sum_{i=1}^t \tilde{l}_i \leq h_t$ for all t.*

Algorithm 2 Meta algorithm on $\mathbb{R}^d$.

1: Define $\mathcal{A}_{1d}$ as a copy of Algorithm 1. Define $\mathcal{A}_{\mathsf{B}}$ as OGD on the d-dimensional unit L_2 norm ball, with adaptive learning rate $\eta_t = \sqrt{2/\sum_{i=1}^t \|g_i\|^2}$. The initialization of $\mathcal{A}_{\mathsf{B}}$ is arbitrary.
2: **for** $t = 1, 2, \dots$ **do**
3: Query $\mathcal{A}_{1d}$ for its prediction $\tilde{y}_t \in \mathbb{R}$. Let $y_t = \Pi_{\mathbb{R}_+}(\tilde{y}_t)$.
4: Query $\mathcal{A}_{\mathsf{B}}$ for its prediction $w_t \in \mathbb{R}^d$; $\|w_t\| \le 1$.
5: Predict $x_t = w_t y_t$, receive the loss gradient $g_t \in \mathbb{R}^d$.
6: Send g_t as the surrogate loss gradient to $\mathcal{A}_{\mathsf{B}}$.
7: Define $l_t = \langle g_t, w_t \rangle$, and

$$\tilde{l}_t = \begin{cases} l_t, & l_t \tilde{y}_t \ge l_t y_t, \\ 0, & \text{else.} \end{cases}$$

8: Send $\tilde{l}_t$ as the surrogate loss gradient to $\mathcal{A}_{1d}$.
9: **end for**

4.2. Analysis

Turning to the analysis, our key innovation is the following lemma.

Lemma 4 (Key lemma: one step potential bound) *Let $\varepsilon > 0$, $\alpha > \frac{1}{2}$, and for all t, $k_t = 2h_t$ and $z_t = \frac{12\alpha+4}{2\alpha-1} h_t^2$. Then, the 1D potential functions satisfy*

$$\Phi_t(V + c^2, S + c) - \Phi_{t-1}(V, S) - c\partial_2 \Phi_t(V, S) \le 0,$$

for all $V \ge 0$, $S \ge -h_{t-1}$ and $c \in [-h_t - \min(S, 0), h_t]$.

This is a potential verification argument, serving a similar purpose as the Itô's formula in the CT analysis (Lemma 2). The condition on c simply means we require $c \in [-h_t, h_t]$ and $S + c \ge -h_t$. With the lemma, one can take a telescopic sum with $c = -\tilde{l}_t$, which returns a cumulative loss upper bound of the 1D base algorithm (Algorithm 1): $\sum_{t=1}^T \tilde{l}_t \tilde{y}_t \le \Phi_0(0,0) - \Phi_T\left(\tilde{V}_T, \tilde{S}_T\right)$. Then, similar to what we did in CT, the regret bound of Algorithm 1 follows from the loss-regret duality (McMahan and Orabona, 2014) and a Fenchel conjugate computation (Lemma 12).

Now let us sketch the proof of Lemma 4.

Proof sketch of Lemma 4 The proof is structured into three steps.

1. Proving $\Phi_t(V, S) \le \Phi_{t-1}(V, S)$.

   This is due to $\partial_1 \phi(x, y) \le 0$. After that, we have

   $$\Phi_t(V + c^2, S + c) - \Phi_{t-1}(V, S) - c\partial_2 \Phi_t(V, S) \le \Phi_t(V + c^2, S + c) - \Phi_t(V, S) - c\partial_2 \Phi_t(V, S),$$

   and it suffices to show RHS ≤ 0. Since all the subscripts are t now, let us simply drop this subscript from Φ, z and k, and write G in place of h_t.

2. Convert checking Φ to checking ϕ, using the change of variables.

   Let us define

   $$f_{V,S}(c) = \Phi(V + c^2, S + c) - \Phi(V, S) - c\partial_2 \Phi(V, S).$$

The task now is showing $f_{V,S}(c) \le 0$. Taking the derivatives with respect to c,

$$f'_{V,S}(c) = 2c\partial_1\Phi(V + c^2, S + c) + \partial_2\Phi(V + c^2, S + c) - \partial_2\Phi(V, S),$$

$$\begin{aligned} f''_{V,S}(c) &= 2\partial_1\Phi + 4c^2\partial_{11}\Phi + 4c\partial_{12}\Phi + \partial_{22}\Phi\Big|_{(V+c^2,S+c)} \\ &\le 2\partial_1\Phi + 4G^2\partial_{11}\Phi + 4G\,|\partial_{12}\Phi| + \partial_{22}\Phi\Big|_{(V+c^2,S+c)}, \end{aligned}$$

where the final subscript means all the involved partial derivatives are evaluated at $(V + c^2, S + c)$. Notice that $f_{V,S}(0) = f'_{V,S}(0) = 0$. Therefore, to prove $f_{V,S}(c) \le 0$, it suffices to show $f''_{V,S}(c) \le 0$ for all considered values of V, S and c.

Crucially, due to the change of variables, partial derivatives of Φ can be easily rewritten using partial derivatives of ϕ (Appendix B). Plugging that in, it suffices to verify the following two cases.

Case 1. If $\partial_{12}\Phi(V + c^2, S + c) \le 0$, then

$$2\partial_1\phi + (k - 2G)^2\partial_{11}\phi + 2(k - 2G)\partial_{12}\phi + \partial_{22}\phi\Big|_{(V+c^2+z+k(S+c),S+c)} \le 0.$$

Case 2. If $\partial_{12}\Phi(V + c^2, S + c) > 0$, then

$$2\partial_1\phi + (k + 2G)^2\partial_{11}\phi + 2(k + 2G)\partial_{12}\phi + \partial_{22}\phi\Big|_{(V+c^2+z+k(S+c),S+c)} \le 0.$$

3. Controlling $\partial_{11}\phi$ and $\partial_{12}\phi$ by picking k and z, and applying the BHE.

   Closely examining the above two inequalities on ϕ, one could notice a striking similarity with the BHE $\partial_1\phi + \frac{1}{2}\partial_{22}\phi = 0$, which the ideal CT potential ϕ^{CT} was *designed* to satisfy, cf., Eq.(4). In particular, if one could drop the two annoying terms, $\partial_{11}\phi$ and $\partial_{12}\phi$, then ϕ^{CT} already fits into the above two cases with *equality*. Essentially, the exhibited similarity is due to the fact that both ϕ (in DT) and ϕ^{CT} (in CT) have their "time variable" (i.e., their first argument) growing according to the quadratic variation of the environment. Therefore, it appears to be an important problem structure, rather than a coincidence that happens to work in our favor.

   With that, our idea is to pick k and z such that the annoying residual terms are upper bounded by a small constant multiplying $\partial_{22}\phi$. Then, we can still use the BHE $\partial_1\phi + \alpha\partial_{22}\phi = 0$, but with a different diffusivity constant $\alpha > \frac{1}{2}$, to control the LHS of the above two cases. Eventually, this will only cost us a slightly suboptimal leading constant in the regret bound ($\sqrt{4 \cdot \frac{1}{2}} = \sqrt{2} \Rightarrow \sqrt{4\alpha}$).

   Formally, to this end, notice that $k = 2G$ trivially satisfies Case 1. Case 2 is a bit more involved, but from the full expressions of $\partial_{11}\phi$, $\partial_{12}\phi$ and $\partial_{22}\phi$, it is not hard to see that a large enough $z \ge \frac{12\alpha+4}{2\alpha-1}G^2$ suffices. This completes the proof. ■

With Lemma 4 above, it is fairly straightforward to obtain the regret bound of the 1D base algorithm (Lemma 13), as we sketched earlier. Then, since the meta algorithm is simply the combination of two existing black-box reductions, its regret bound follows from (Cutkosky and Orabona, 2018, Theorem 2) and (Cutkosky, 2020, Theorem 2). This returns the following theorem as the main result of this paper.

Theorem 5 (Main result) *With $\varepsilon > 0$, $\alpha > \frac{1}{2}$, $k_t = 2h_t$ and $z_t = \frac{12\alpha+4}{2\alpha-1}h_t^2$, Algorithm 2 guarantees for all $T \in \mathbb{N}_+$ and $u \in \mathbb{R}^d$,*

$$\mathrm{Regret}_T(Env, u) \leq \varepsilon\sqrt{\alpha\left(V_T + z_T + k_T\bar{S}\right)} + \|u\|\left(\bar{S} + 2\sqrt{2V_T}\right),$$

where

$$\bar{S} = 4\alpha k_T\left(1 + \sqrt{\log(2\|u\|\varepsilon^{-1} + 1)}\right)^2 + \sqrt{4\alpha\left(V_T + z_T\right)}\left(1 + \sqrt{\log(2\|u\|\varepsilon^{-1} + 1)}\right).$$

Theorem 5 contains the precise regret bound without any big-Oh. Nonetheless, one could use asymptotic orders to make it more interpretable (see Appendix D for the derivation).

$$\mathrm{Regret}_T(Env, u) \leq \varepsilon\left(\sqrt{\alpha\left(V_T + \frac{12\alpha+4}{2\alpha-1}h_T^2\right)} + 4\alpha h_T\right) + \|u\|\, O\left(\sqrt{V_T\log(\|u\|\varepsilon^{-1})} \vee h_T\log(\|u\|\varepsilon^{-1})\right),$$

which is simultaneously valid in two regimes: (i) $\|u\| \gg \varepsilon$ and $V_T \gg h_T^2$; and (ii) $u = 0$. In comparator adaptive online learning, regret bounds of this form are said to characterize the *loss-regret tradeoff* (Zhang et al., 2022a): with a small ε, one could ensure that the cumulative loss $\mathrm{Regret}_T(Env, 0)$ is low, while only sacrificing a $\sqrt{\log(\varepsilon^{-1})}$ penalty on the leading term of the regret bound. We also note that the additive logarithmic "residual" term $h_T\log(\|u\|\varepsilon^{-1})$ is standard in simultaneously adaptive regret bounds, and not removable in some sense (Cutkosky, 2018, Section 5.5.1).

The key strength of this bound is that, the dependence on V_T alone is $O(\sqrt{V_T})$, matching the optimal gradient adaptive bound achieved by OGD, and improving prior works on simultaneous adaptivity (Cutkosky and Orabona, 2018; Mhammedi and Koolen, 2020; Jacobsen and Cutkosky, 2022). Furthermore, in the regime of large $\|u\|$ and V_T, the leading order term is $\sqrt{4\alpha V_T\log(\|u\|\varepsilon^{-1})}$, where the multiplying constant $\sqrt{4\alpha}$ almost matches the $\sqrt{2}$ lower bound from (Zhang et al., 2022a).

5. Extension: Unknown Lipschitz constant without hints

Finally, we discuss the full extension of our main result to the setting with unknown G, removing the need of hints. This follows from a reduction to our Section 4, developed by (Cutkosky, 2019; Mhammedi and Koolen, 2020). Let us define $G_t := \max_{i\leq t}\|g_t\|$, and w.l.o.g., $G := G_T$.

The essential idea is the following. Without knowing G, we use the hint h_t as a guess of the "running Lipschitz constant" G_t, before observing $\|g_t\|$. Naturally, $h_t = G_{t-1}$ makes a reasonable guess, but there is always some chance of "surprise", where $\|g_t\| > G_{t-1}$, and the analysis from Section 4 breaks. To fix this issue, instead of feeding the algorithm the true gradient g_t, one could feed its *clipped version*, $\bar{g}_t = g_tG_{t-1}/G_t$. Now, $h_t = G_{t-1}$ is always a valid hint for the clipped gradient $\bar{g}_t$, therefore our main result (Theorem 5) can be applied in a black-box manner.

Ultimately, we care about the regret evaluated on the true gradients $g_{1:T}$, rather than the clipped gradients $\bar{g}_{1:T}$. Their difference is related to the magnitude of the predictions $\|x_t\|$, thus one could use the standard constraint-imposing technique (Cutkosky, 2020, Theorem 2) once again to control it. In combination, this yields the following lemma (Mhammedi and Koolen, 2020, Corollary 3).

Lemma 6 *If we denote the simultaneously adaptive regret bound in our Theorem 5 as a function $R(u, V_T, h_T)$, then there exists an algorithm taking ours as a black-box subroutine, guaranteeing*

$$\mathrm{Regret}_T(Env, u) \le R(u, V_T, G) + G\|u\|^3 + G\sqrt{\max_{t\le T} B_t} + G\|u\|,$$

where $B_t := \sum_{i=1}^t \|g_i\|/G_t$.

Notably, this algorithm needs neither the knowledge of G, nor any other oracle knowledge like hints. Moreover, no restarting (e.g., the doubling trick) is required. In the special case with bounded domain ($\|x_t\|, \|u\| \le D$), a simplified variant of this procedure guarantees an even smaller bound, $\mathrm{Regret}_T(Env, u) \le R(u, V_T, G) + 2DG$.

The strength of our main result can be demonstrated in this general setting as well.

- First, the algorithm obtained from Section 4 and Lemma 6 is *scale-free* (Orabona and Pál, 2018), in the sense that if all the loss gradients are scaled by $c \in \mathbb{R}_{>0}$, then the prediction x_t of the algorithm remains unchanged, and the above G-adaptive regret bound is scaled by exactly c. This is a favorable property in practice, also satisfied by prior works (Mhammedi and Koolen, 2020; Jacobsen and Cutkosky, 2022) which we quantitatively improve.

- Second, we eliminate the *range ratio* problem in prior works (Mhammedi and Koolen, 2020; Jacobsen and Cutkosky, 2022). For OLO with hints, existing analogues of our Theorem 5 have the shape like $\mathrm{Regret}_T(Env, u) \le O\left(\|u\|\sqrt{V_T \log(\|u\| V_T/h_1^2)}\right)$, where the range ratio V_T/h_1^2 can be arbitrarily large despite being inside the log. Using a restarting trick (Mhammedi and Koolen, 2020) or a somewhat more complicated "soft-thresholding" on the prediction x_t (Jacobsen and Cutkosky, 2022), one could replace this range ratio by $\mathrm{poly}(T)$. However, this sacrifices either the practicality of the algorithm or its analytical simplicity, and in both cases, one is left with a suboptimal $\sqrt{\log T}$ multiplicative factor on the leading term of the regret bound.

  Essentially, such a range ratio problem originated in (Mhammedi and Koolen, 2020) due to the existence of "unit" in their confidence hyperparameter ε. Analogous to (McMahan and Orabona, 2014; Orabona and Pál, 2016), the paper applies the potential function

  $$\Phi_t(V, S) = \frac{\varepsilon}{\sqrt{V}} \exp\left(\frac{S^2}{2V + 2h_t|S|}\right),$$

  where ε has the unit of G^2 due to a dimensional analysis. If G is known, then one could pick $\varepsilon \propto G^2$, leading to the regret bound Eq.(3). Without knowing G, since ε needs to be determined at the beginning, the feasible choice becomes $\varepsilon \propto h_1^2$ – this replaces G^2 in Eq.(3) by h_1^2, causing the range ratio problem.

  Our algorithm has an important difference. The confidence hyperparameter ε in our potential function Eq.(6) is "unitless", therefore when selecting it at the beginning, we do not need a guess of G. This eliminates the range ratio problem, since there are no V_T/h_1^2 or G/h_1 terms in our regret bound at all. Neither restarting nor soft-thresholding is needed. The takeaway is that, such a range ratio problem appears to be an analytical artifact due to certain unit inconsistency (and ultimately, the suboptimal loss-regret tradeoff (Zhang et al., 2022a)), which can be eliminated by a better design of the potential function.

6. Conclusion

The present work studies how to achieve simultaneous gradient and comparator adaptivity in OLO with Lipschitz losses. A new continuous-time-inspired algorithm is proposed, improving the $O(\sqrt{V_T \log V_T})$ regret bound from prior works (with respect to the gradient variance V_T alone) to the optimal rate $O(\sqrt{V_T})$. The crucial technique is a new discretization argument that preserves gradient adaptivity from CT to DT, improving an already powerful, but non-gradient-adaptive one from the literature. This could be of broader applicability, and a natural step forward is exploiting the benefits of this technique in other online learning problems of interest. Finally, the extension to the setting with unknown Lipschitz constant is discussed, where our algorithm is made scale-free.

References

Jacob Abernethy, Elad E Hazan, and Alexander Rakhlin. Competing in the dark: An efficient algorithm for bandit linear optimization. In *Conference on Learning Theory*, pages 263–273, 2008.

Jacob Abernethy, Rafael M Frongillo, and Andre Wibisono. Minimax option pricing meets black-scholes in the limit. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 1029–1040, 2012.

Jacob Abernethy, Peter L Bartlett, Rafael Frongillo, and Andre Wibisono. How to hedge an option against an adversary: Black-scholes pricing is minimax optimal. *Advances in Neural Information Processing Systems*, 26, 2013.

Erhan Bayraktar, Ibrahim Ekren, and Xin Zhang. Finite-time 4-expert prediction problem. *Communications in Partial Differential Equations*, 45(7):714–757, 2020.

Lilian Besson and Emilie Kaufmann. What doubling tricks can and can't do for multi-armed bandits. *arXiv preprint arXiv:1803.06971*, 2018.

Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of political economy*, 81(3):637–654, 1973.

Liyu Chen, Haipeng Luo, and Chen-Yu Wei. Impossible tuning made possible: A new expert algorithm and its applications. In *Conference on Learning Theory*, pages 1216–1259. PMLR, 2021.

Ashok Cutkosky. *Algorithms and Lower Bounds for Parameter-free Online Learning*. Stanford University, 2018.

Ashok Cutkosky. Artificial constraints and hints for unbounded online learning. In *Conference on Learning Theory*, pages 874–894. PMLR, 2019.

Ashok Cutkosky. Parameter-free, dynamic, and strongly-adaptive online learning. In *International Conference on Machine Learning*, pages 2250–2259. PMLR, 2020.

Ashok Cutkosky and Francesco Orabona. Black-box reductions for parameter-free online learning in banach spaces. In *Conference On Learning Theory*, pages 1493–1529. PMLR, 2018.

Ashok Cutkosky, Aaron Defazio, and Harsh Mehta. Mechanic: A learning rate tuner. *arXiv preprint arXiv:2306.00144*, 2023.

Amit Daniely and Yishay Mansour. Competitive ratio vs regret minimization: achieving the best of both worlds. In *Algorithmic Learning Theory*, pages 333–368. PMLR, 2019.

Nadejda Drenska and Robert V Kohn. Prediction with expert advice: A PDE perspective. *Journal of Nonlinear Science*, 30(1):137–173, 2020.

John Duchi, Elad Hazan, and Yoram Singer. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12(7), 2011.

Yoav Freund. A method for hedging in continuous time. *arXiv preprint arXiv:0904.3356*, 2009.

Laura Greenstreet, Nicholas JA Harvey, and Victor Sanches Portella. Efficient and optimal fixed-time regret with two experts. In *International Conference on Algorithmic Learning Theory*, pages 436–464. PMLR, 2022.

Nicholas JA Harvey, Christopher Liaw, Edwin Perkins, and Sikander Randhawa. Optimal anytime regret with two experts. *Mathematical Statistics and Learning*, 6(1):87–142, 2023.

Andrew Jacobsen and Ashok Cutkosky. Parameter-free mirror descent. In *Conference on Learning Theory*, pages 4160–4211. PMLR, 2022.

Michael Kapralov and Rina Panigrahy. Prediction strategies without loss. *Advances in Neural Information Processing Systems*, 24, 2011.

Vladimir A Kobzar, Robert V Kohn, and Zhilei Wang. New potential-based bounds for prediction with expert advice. In *Conference on Learning Theory*, pages 2370–2405. PMLR, 2020.

H Brendan McMahan and Francesco Orabona. Unconstrained online linear learning in hilbert spaces: Minimax algorithms and normal approximations. In *Conference on Learning Theory*, pages 1020–1039. PMLR, 2014.

H Brendan McMahan and Matthew Streeter. Adaptive bound optimization for online convex optimization. In *Conference on Learning Theory*, 2010.

Zakaria Mhammedi and Wouter M Koolen. Lipschitz and comparator-norm adaptivity in online learning. In *Conference on Learning Theory*, pages 2858–2887. PMLR, 2020.

Francesco Orabona. A modern introduction to online learning. *arXiv preprint arXiv:1912.13213*, 2023.

Francesco Orabona and Dávid Pál. Coin betting and parameter-free online learning. *Advances in Neural Information Processing Systems*, 29, 2016.

Francesco Orabona and Dávid Pál. Scale-free online learning. *Theoretical Computer Science*, 716: 50–69, 2018.

Francesco Orabona and Tatiana Tommasi. Training deep networks without learning rates through coin betting. *Advances in Neural Information Processing Systems*, 30:2160–2170, 2017.

Victor Sanches Portella, Christopher Liaw, and Nicholas JA Harvey. Continuous prediction with experts' advice. *arXiv preprint arXiv:2206.00236*, 2022.

Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293. Springer Science & Business Media, 2013.

Matthew Streeter and Brendan Mcmahan. No-regret algorithms for unconstrained online convex optimization. *Advances in Neural Information Processing Systems*, 25, 2012.

Zhilei Wang and Robert V Kohn. A new approach to drifting games, based on asymptotically optimal potentials. *arXiv preprint arXiv:2207.11405*, 2022.

Zhiyu Zhang, Ashok Cutkosky, and Ioannis Paschalidis. PDE-based optimal strategy for unconstrained online learning. In *International Conference on Machine Learning*, pages 26085–26115. PMLR, 2022a.

Zhiyu Zhang, Ashok Cutkosky, and Yannis Paschalidis. Optimal comparator adaptive online learning with switching cost. *Advances in Neural Information Processing Systems*, 35:23936–23950, 2022b.

Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *International Conference on Machine Learning*, pages 928–936, 2003.

Appendix A. More related work

Simultaneous adaptivity Within the two types of adaptive online learning, comparator adaptivity is often considered more challenging than gradient adaptivity. Therefore, existing works on the simultaneous adaptivity problem (Cutkosky and Orabona, 2018; Mhammedi and Koolen, 2020; Jacobsen and Cutkosky, 2022) are all built on the core algorithmic frameworks of comparator adaptivity, with various "gradient adaptive modifications". Let us take a closer look.

- The state-of-the-art result for comparator adaptivity alone is due to (Zhang et al., 2022a), which achieved not only the order-optimal regret bound Eq.(2), but also the optimal leading constant $\sqrt{2}$, improving a series of earlier works (Streeter and Mcmahan, 2012; McMahan and Orabona, 2014; Orabona and Pál, 2016). The key technique is a CT analysis of dual space OLO algorithms (specifically, the *potential framework*), combining the *potential verification* argument from (Harvey et al., 2023) and the *loss-regret duality* from (McMahan and Orabona, 2014).
- Simultaneous gradient and comparator adaptivity was achieved in (Cutkosky and Orabona, 2018) for the first time, cf., Eq.(3). It relied on a refined version of a *coin betting* algorithm (Orabona and Pál, 2016), with the betting fractions selected separately by online learning. It has been shown that coin betting is a variant of the potential framework with a suboptimal loss-regret tradeoff (Zhang et al., 2022a), which is inherently tied to suboptimal logarithmic factors in the regret bound.
- Mhammedi and Koolen (2020) designed different potential-based algorithms achieving Eq.(3). Unique to this work is a novel computer-aided analysis: a suitable potential function candidate (inspired by (McMahan and Orabona, 2014; Orabona and Pál, 2016)) was first guessed, and then verified using MATHEMATICA. It is possible that the same approach could verify better guesses

that achieve our goal. The difficulty is that, the success would be subject to "good intuitions" combined with the capability of the software, which is not well-understood yet.

- Most recently, Jacobsen and Cutkosky (2022) achieved Eq.(3) (modulo $\log(\log)$ factors) using a variant of *Online Mirror Descent* (OMD). This is a primal space approach fundamentally different from all the dual space approaches discussed above, therefore several related questions remain open. For example, it is unclear if results like (Zhang et al., 2022a) (on comparator adaptivity alone, based on a CT analysis) can also be achieved using this primal space approach.

  During our revision, we also notice that the algorithm from (Jacobsen and Cutkosky, 2022) with a different hyperparameter setting (not reported there) might achieve the $O(\sqrt{V_T})$ regret bound as well, without the extra $\log(V_T)$ factor. We are currently verifying such an observation with the authors, and the outcome will be incorporated into the arXiv version of this work.

- Besides the above, one can also achieve simultaneous adaptivity using a powerful aggregation approach (Chen et al., 2021). This approach is formulated for a much more general setting, but on an unbounded domain, it requires increased computation, i.e., $O(\sqrt{T})$ per round, making it somewhat incomparable to the other baselines surveyed above.

In light of these works, the present paper aims to use a better CT workflow than (Zhang et al., 2022a) to improve the results of (Cutkosky and Orabona, 2018; Mhammedi and Koolen, 2020; Jacobsen and Cutkosky, 2022). There are also techniques relaxing the assumption of knowing the Lipschitz constant G (Cutkosky, 2019; Mhammedi and Koolen, 2020). We study this setting as an extension of our main result.

Existing bound + doubling trick We mention in Section 1 that the doubling trick can *roughly* turn an existing regret bound Eq.(3) (Cutkosky and Orabona, 2018; Mhammedi and Koolen, 2020; Jacobsen and Cutkosky, 2022) to the form we aim for. Here is an explanation.

Concretely, using $\hat{O}$ to hide $\log(\log)$ factors, (Jacobsen and Cutkosky, 2022, Theorem 1) achieves

$$\mathrm{Regret}_T(u) \le \hat{O}\left(\varepsilon G + \|u\| \left[\sqrt{V_T \log\left(\frac{\|u\|\sqrt{V_T}}{\varepsilon G} + 1\right)} \vee G \log\left(\frac{\|u\|\sqrt{V_T}}{\varepsilon G} + 1\right)\right]\right).$$

With a known V_T budget, setting $\varepsilon = O(\sqrt{V_T}/G)$ yields the desirable bound $\hat{O}(\|u\|\sqrt{V_T \log \|u\|})$, up to a "morally secondary term" $\|u\| G \log \|u\|$. Applying the standard doubling trick can relax the known-V_T assumption. The only small price to pay is that the secondary term is multiplied by $O(\log T)$.

Apart from that, (Cutkosky and Orabona, 2018; Mhammedi and Koolen, 2020) cannot be applied similarly due to a tuning issue. Even knowing V_T, the parameter ε there cannot be set accordingly to achieve the $O(\|u\|\sqrt{V_T \log \|u\|})$ regret bound we aim for.

Weakness of doubling trick Section 1 also claims that the doubling trick is impractical. To justify this claim, we first refer the readers to an empirical study: (Besson and Kaufmann, 2018, Section 5) evaluated several versions of the doubling trick in the bandit setting. In almost all cases, the combination of the fixed-T algorithm and the doubling trick performs considerably worse than the "intrinsically anytime" algorithm with weaker theoretical guarantees; and especially, such a performance gap widens with each restart. There is a theoretical support for this: the doubling trick incurs an extra multiplicative constant, which is at least 3 (Besson and Kaufmann, 2018, Section 5).

It means for our problem, the combination of existing regret bounds and the doubling trick cannot achieve (or get close to) the optimal leading constant.

In addition, we would argue that the doubling trick is aesthetically unsatisfactory. Restarting wastes data and causes large "jumps" in the decision sequence, which are often undesirable. Therefore, although the doubling trick can be theoretically sufficient (e.g., making fixed-learning-rate OGD anytime), more streamlined solutions are typically favored (e.g., time-varying learning rates).

Appendix B. Basics of the potential function

For the two potential functions defined in Section 4, we compute their derivatives as follows. Starting from ϕ,

$$\partial_1\phi(x,y) = -\frac{\varepsilon\sqrt{\alpha}}{2\sqrt{x}}\exp\left(\frac{y^2}{4\alpha x}\right),$$
$$\partial_2\phi(x,y) = \varepsilon\mathrm{erfi}\left(\frac{y}{\sqrt{4\alpha x}}\right),$$
$$\partial_{11}\phi(x,y) = \frac{\varepsilon\sqrt{\alpha}}{4x^{3/2}}\left(\frac{y^2}{2\alpha x}+1\right)\exp\left(\frac{y^2}{4\alpha x}\right),$$
$$\partial_{12}\phi(x,y) = -\frac{\varepsilon y}{4\sqrt{\alpha}x^{3/2}}\exp\left(\frac{y^2}{4\alpha x}\right),$$
$$\partial_{22}\phi(x,y) = \frac{\varepsilon}{2\sqrt{\alpha x}}\exp\left(\frac{y^2}{4\alpha x}\right).$$

Due to the change of variables, the derivatives of Φ_t can be concisely represented by the derivatives of ϕ.

$$\partial_1\Phi_t(V,S) = \partial_1\phi(V+z_t+k_tS,S),$$
$$\partial_2\Phi_t(V,S) = k_t\partial_1\phi(V+z_t+k_tS,S)+\partial_2\phi(V+z_t+k_tS,S),$$
$$\partial_{11}\Phi_t(V,S) = \partial_{11}\phi(V+z_t+k_tS,S),$$
$$\partial_{12}\Phi_t(V,S) = k_t\partial_{11}\phi(V+z_t+k_tS,S)+\partial_{12}\phi(V+z_t+k_tS,S),$$
$$\partial_{22}\Phi_t(V,S) = k_t^2\partial_{11}\phi(V+z_t+k_tS,S)+2k_t\partial_{12}\phi(V+z_t+k_tS,S)+\partial_{22}\phi(V+z_t+k_tS,S).$$

Next, we present two simple lemmas on the potential function Φ_t. The first lemma shows Φ_t is convex in its second argument. The second lemma shows the negativity of $\partial_2\Phi_t(V,S)$ when $S\leq 0$.

Lemma 7 (Convexity) *If $\varepsilon,\alpha,k_t>0$ and $z_t>k_th_t$, then the potential function $\Phi_t(V,S)$ satisfies $\partial_{22}\Phi_t(V,S)>0$ for all $V\geq 0$ and $S\geq -h_t$.*

Proof [Proof of Lemma 7] Let us drop all the subscript t and let $G=h_t$. Define the shorthands $x=V+z+kS$ and $y=S$. For all $V\geq 0$ and $S\geq -G$, we have $x>0$, therefore

$$\begin{aligned}\partial_{22}\Phi(V,S) &= k^2\partial_{11}\phi(x,y)+2k\partial_{12}\phi(x,y)+\partial_{22}\phi(x,y)\\ &= \frac{\varepsilon\sqrt{\alpha}}{4x^{3/2}}\exp\left(\frac{y^2}{4\alpha x}\right)\left(\frac{k^2y^2}{2\alpha x}+k^2-\frac{2ky}{\alpha}+\frac{2x}{\alpha}\right)\\ &= \frac{\varepsilon\sqrt{\alpha}}{4x^{3/2}}\exp\left(\frac{y^2}{4\alpha x}\right)\left(\frac{k^2y^2}{2\alpha x}+k^2+\frac{2(V+z)}{\alpha}\right).\end{aligned}$$

The RHS is strictly positive. ■

Lemma 8 (The sign of prediction) *If $\varepsilon, \alpha, k_t > 0$ and $z_t > k_t h_t$, then the potential function $\Phi_t(V,S)$ satisfies $\partial_2 \Phi_t(V,S) < 0$ for all $V \geq 0$ and $-h_t \leq S \leq 0$.*

Proof [Proof of Lemma 8] Same as before, drop all the subscript t. Let us check $\partial_2 \Phi(V,0) < 0$. Indeed,

$$\partial_2 \Phi(V,0) = k\partial_1 \phi(V+z,0) + \partial_2 \phi(V+z,0) = -\frac{\varepsilon k\sqrt{\alpha}}{2\sqrt{V+z}} < 0.$$

Moreover, $\Phi(V,S)$ is convex with respect to its second argument, due to Lemma 7. ■

The final two basic lemmas concern the property of the erfi function.

Lemma 9 *For all $x \geq 1$,* $\operatorname{erfi}(x) \geq \exp(x^2)/2x$.

Proof [Proof of Lemma 9] Let $f(x) = \operatorname{erfi}(x) - \exp(x^2)/2x$. $f(1) = \operatorname{erfi}(1) - e/2 > 0$. For all $x \geq 1$,

$$f'(x) = \frac{1}{2x^2}\exp(x^2) > 0,$$

which means $f(x) > 0$ for all $x \geq 1$. ■

Lemma 10 *For all $x \geq 0$,* $\operatorname{erfi}^{-1}(x) \leq 1 + \sqrt{\log(x+1)}$.

Proof [Proof of Lemma 10] We first show $\operatorname{erfi}(x) \geq \exp(x^2 - x) - 1$ for all $x \geq 0$. Let $f(x) = \operatorname{erfi}(x) - \exp(x^2 - x) + 1$, then $f(0) = 0$,

$$f'(x) = \exp(x^2 - x) \cdot [\exp(x) - (2x-1)]\,.$$

It is easy to verify $\exp(x) \geq (2x-1)$ for all $x \geq 0$.

Next, for any $y \geq 0$ let $x = \operatorname{erfi}^{-1}(y)$, which is also nonnegative. From the first step,

$$y = \operatorname{erfi}(x) \geq \exp\left[\left(x - \frac{1}{2}\right)^2 - \frac{1}{4}\right] - 1,$$

therefore

$$x \leq \frac{1}{2} + \sqrt{\frac{1}{4} + \log(y+1)} \leq 1 + \sqrt{\log(y+1)}.$$

Substituting $x = \operatorname{erfi}^{-1}(y)$ completes the proof. ■

Appendix C. Analysis of the base algorithm

The key step of our analysis is the following potential verification argument.

Lemma 11 (Key lemma: one step potential bound) *Let $\varepsilon > 0$, $\alpha > \frac{1}{2}$, and for all t, $k_t = 2h_t$ and $z_t = \frac{12\alpha+4}{2\alpha-1}h_t^2$. Then, the 1D potential functions satisfy*

$$\Phi_t(V + c^2, S + c) - \Phi_{t-1}(V, S) - c\partial_2\Phi_t(V, S) \leq 0,$$

for all $V \geq 0$, $S \geq -h_{t-1}$ and $c \in [-h_t - \min(S, 0), h_t]$.

Proof [Proof of Lemma 4] The first, preparatory step is to show $\Phi_t(V, S) \leq \Phi_{t-1}(V, S)$ for all $V \geq 0$ and $S \geq -h_{t-1}$. To see this, notice that

$$\Phi_t(V, S) = \phi(V + z_t + k_t S, S),$$

$$\Phi_{t-1}(V, S) = \phi(V + z_{t-1} + k_{t-1} S, S).$$

From Appendix B, $\partial_1\phi(x, y) \leq 0$ for all $x > 0$ and $y \in \mathbb{R}$. Furthermore, compare the first argument on the above right hand sides,

$$\begin{aligned} V + z_t + k_t S &= V + h_t\left(\frac{12\alpha+4}{2\alpha-1}h_t + 2S\right) \\ &\geq V + h_{t-1}\left(\frac{12\alpha+4}{2\alpha-1}h_{t-1} + 2S\right) \qquad (S \geq -h_{t-1} \text{ and } h_t \geq h_{t-1}) \\ &= V + z_{t-1} + k_{t-1}S. \end{aligned}$$

Concluding this argument, we have shown $\Phi_t(V, S) \leq \Phi_{t-1}(V, S)$. It means to prove the present lemma, it suffices to show

$$\Phi_t(V + c^2, S + c) - \Phi_t(V, S) - c\partial_2\Phi_t(V, S) \leq 0,$$

for all $V \geq 0$, $S \geq -h_t$ and $c \in [-h_t - \min(S, 0), h_t]$. Since all the subscripts are the same t, let us drop them completely to simplify the exposition. Also, let us denote $h_t = G$, which hopefully makes the "unit" clearer.

Now, let us view our remaining objective as a function of c,

$$f_{V,S}(c) := \Phi(V + c^2, S + c) - \Phi(V, S) - c\partial_2\Phi(V, S).$$

Taking the derivatives,

$$f'_{V,S}(c) = 2c\partial_1\Phi(V + c^2, S + c) + \partial_2\Phi(V + c^2, S + c) - \partial_2\Phi(V, S),$$

$$\begin{aligned} f''_{V,S}(c) &= 2\partial_1\Phi(V + c^2, S + c) + 4c^2\partial_{11}\Phi(V + c^2, S + c) + 4c\partial_{12}\Phi(V + c^2, S + c) \\ &\qquad + \partial_{22}\Phi(V + c^2, S + c) \\ &\leq 2\partial_1\Phi(V + c^2, S + c) + 4G^2\partial_{11}\Phi(V + c^2, S + c) + 4G\left|\partial_{12}\Phi(V + c^2, S + c)\right| \\ &\qquad + \partial_{22}\Phi(V + c^2, S + c). \end{aligned} \tag{7}$$

Note that $f_{V,S}(0) = f'_{V,S}(0) = 0$. Therefore, to prove $f_{V,S}(c) \le 0$, it suffices to show $f''_{V,S}(c) \le 0$ for all considered values of V, S and c. Also notice that the RHS of Eq.(7) has a striking similarity to the Backward Heat Equation – in fact, after a change of variable, the resulting expressions, namely Eq.(8) and Eq.(9) below, will resemble the standard BHE on ϕ ($\partial_1\phi + \frac{1}{2}\partial_{22}\phi = 0$) plus "lower order terms". The main goal of this proof is to control such lower order terms by properly choosing k and z.

Concretely, due to the absolute value in Eq.(7), we will analyze two cases. Technically, the first case is harder, therefore we pick k to simplify its analysis. The second case is relatively easier to handle.

Case 1: $\partial_{12}\Phi(V + c^2, S + c) \le 0$. Substituting the derivatives of Φ by the derivatives of ϕ, we have

$$f''_{V,S}(c) \le 2\partial_1\phi + (k - 2G)^2\partial_{11}\phi + 2(k - 2G)\partial_{12}\phi + \partial_{22}\phi\Big|_{(V+c^2+z+k(S+c),S+c)}. \tag{8}$$

The RHS means we evaluate all the derivative functions at $(V + c^2 + z + k(S + c), S + c)$. Plugging in our specific choice of k and α,

$$\begin{aligned} f''_{V,S}(c) &\le 2\partial_1\phi + \partial_{22}\phi\Big|_{(V+c^2+z+2(S+c),S+c)} && (k = 2G) \\ &\le 2\left(\partial_1\phi + \alpha\partial_{22}\phi\right)\Big|_{(V+c^2+z+2(S+c),S+c)} && (\alpha > \tfrac{1}{2} \text{ and Lemma 7}) \\ &= 0. && (\phi \text{ satisfies the BHE with constant } \alpha) \end{aligned}$$

Case 2: $\partial_{12}\Phi(V + c^2, S + c) \ge 0$. Similar to the first case,

$$f''_{V,S}(c) \le 2\partial_1\phi + (k + 2G)^2\partial_{11}\phi + 2(k + 2G)\partial_{12}\phi + \partial_{22}\phi\Big|_{(V+c^2+z+k(S+c),S+c)}. \tag{9}$$

Consider the k-dependent terms in Eq.(9), $(k + 2G)^2\partial_{11}\phi + 2(k + 2G)\partial_{12}\phi$. Our goal is to upper bound it by $(2\alpha - 1)\partial_{22}\phi$, such that in total, the RHS of Eq.(9) becomes $2(\partial_1\phi + \alpha\partial_{22}\phi)$, which equals 0 due to the BHE.

Plugging in the derivatives of ϕ from Appendix B, for all inputs (x, y),

$$\begin{aligned} &(k + 2G)^2\partial_{11}\phi + 2(k + 2G)\partial_{12}\phi - (2\alpha - 1)\partial_{22}\phi\Big|_{(x,y)} \\ = \;& \frac{\varepsilon}{4\sqrt{\alpha}x^{3/2}} \exp\left(\frac{y^2}{4\alpha x}\right)\left[(k + 2G)^2\left(\frac{y^2}{2x} + \alpha\right) - 2(k + 2G)y - 2x(2\alpha - 1)\right] \\ = \;& \frac{\varepsilon}{4\sqrt{\alpha}x^{3/2}} \exp\left(\frac{y^2}{4\alpha x}\right)\left[16G^2\left(\frac{y^2}{2x} + \alpha\right) - 8Gy - 2x(2\alpha - 1)\right] \end{aligned}$$

We aim to show the bracket on the RHS is negative at $x = V + c^2 + z + k(S + c)$ and $y = S + c$, where $k = 2G$. This amounts to showing

$$\lozenge := \frac{4G^2(S + c)^2}{V + c^2 + z + 2G(S + c)} + 8\alpha G^2 - (4\alpha G + 2G)(S + c) - (2\alpha - 1)(V + c^2 + z) \le 0.$$

The idea is that we can pick a large enough z to make it hold. Concretely,

- If $S+c>0$, then since $\alpha>\frac{1}{2}$,

$$\begin{aligned}\Diamond &\leq \frac{4G^2(S+c)^2}{2G(S+c)}+8\alpha G^2-(4\alpha G+2G)(S+c)-(2\alpha-1)z\\ &= -4\alpha G(S+c)+8\alpha G^2-(2\alpha-1)z\\ &\leq 8\alpha G^2-(2\alpha-1)z.\end{aligned}$$

  and it suffices to pick
$$z\geq \frac{8\alpha}{2\alpha-1}G^2.$$

- If $S+c\leq 0$, then since we require $c\geq -G-\min(S,0)$, we also have $S+c\geq \min(S,0)+c\geq -G$. As long as $z\geq 4G^2$,

$$\begin{aligned}\Diamond &\leq \frac{4G^4}{z-2G^2}+12\alpha G^2+2G^2-(2\alpha-1)z\\ &\leq 12\alpha G^2+4G^2-(2\alpha-1)z.\end{aligned}$$

  It suffices to pick
$$z\geq \frac{12\alpha+4}{2\alpha-1}G^2.$$

In summary, $z=\frac{12\alpha+4}{2\alpha-1}G^2$ ensures $\Diamond\leq 0$. Due to the BHE on ϕ,

$$f''_{V,S}(c)\leq 2(\partial_1\phi+\alpha\partial_{22}\phi)\Big|_{(V+c^2+z+2G(S+c),S+c)}=0.$$

Combining the two cases completes the proof. ■

The following lemma characterizes the Fenchel conjugate of Φ_t (with respect to its second argument).

Lemma 12 (Conjugate) *With $\varepsilon,\alpha,k_t>0$ and $z_t>k_th_t$, for all $u\geq 0$,*

$$\sup_{S\in[-h_t,\infty)}[uS-\Phi_t(V,S)]\leq u\bar{S}+\varepsilon\sqrt{\alpha(V+z_t+k_t\bar{S})},$$

where

$$\bar{S}=4\alpha k_t\left(1+\sqrt{\log(2u\varepsilon^{-1}+1)}\right)^2+\sqrt{4\alpha(V+z_t)}\left(1+\sqrt{\log(2u\varepsilon^{-1}+1)}\right).$$

Proof [Proof of Lemma 12] Throughout this proof, we drop all the subscript t and write G in place of h_t.

We first show that the supremum over S in the Fenchel conjugate is attainable by some $S^*\in[0,\infty)$. To this end, define the function $f(S):=uS-\Phi(V,S)$. f is continuous, with $f'(S)=u-\partial_2\Phi(V,S)$. Moreover, due to Lemma 7, f is concave on $[-G,\infty)$. The existence of $S^*\in[0,\infty)$ then follows from analyzing the boundary.

- For all $S \in [-G, 0]$, we have $f'(S) \geq 0$. The reason is $u \geq 0$, and $\partial_2\Phi(V, S) \leq 0$ due to Lemma 8.
- For sufficiently large S, we aim to show $f'(S) < 0$. Let us begin by writing down $\partial_2\Phi(V, S)$, from Appendix B.

$$\partial_2\Phi(V,S) = \varepsilon \mathrm{erfi}\left(\frac{S}{\sqrt{4\alpha(V+z+kS)}}\right) - \frac{\varepsilon k\sqrt{\alpha}}{2\sqrt{V+z+kS}}\exp\left(\frac{S^2}{4\alpha(V+z+kS)}\right).$$

Now consider large S that satisfies $S \geq \sqrt{4\alpha(V+z+kS)}$. Due to an estimate of the erfi function (Lemma 9),

$$\mathrm{erfi}\left(\frac{S}{\sqrt{4\alpha(V+z+kS)}}\right) \geq \frac{\sqrt{\alpha(V+z+kS)}}{S}\exp\left(\frac{S^2}{4\alpha(V+z+kS)}\right).$$

Moreover,

$$\frac{\sqrt{\alpha(V+z+kS)}}{S} - \frac{k\sqrt{\alpha}}{\sqrt{V+z+kS}} = \frac{\sqrt{\alpha}(V+z)}{S\sqrt{V+z+kS}} \geq 0.$$

Therefore,

$$\begin{aligned}\partial_2\Phi(V,S) &= \left[\frac{\varepsilon}{2}\mathrm{erfi}\left(\frac{S}{\sqrt{4\alpha(V+z+kS)}}\right) - \frac{\varepsilon k\sqrt{\alpha}}{2\sqrt{V+z+kS}}\exp\left(\frac{S^2}{4\alpha(V+z+kS)}\right)\right] \\ &\qquad + \frac{\varepsilon}{2}\mathrm{erfi}\left(\frac{S}{\sqrt{4\alpha(V+z+kS)}}\right) \\ &\geq \frac{\varepsilon}{2}\mathrm{erfi}\left(\frac{S}{\sqrt{4\alpha(V+z+kS)}}\right). \end{aligned} \tag{10}$$

For sufficiently large S, we have RHS $> u$, hence $f'(S) < 0$.

Summarizing the above, we have shown that there exists $S^* \in [0, \infty)$ such that

$$\Phi_V^*(u) := \sup_{S\in[-G,\infty)} uS - \Phi(V,S) = uS^* - \Phi(V,S^*).$$

Moreover, S^* should satisfy the first order optimality condition $f'(S^*) = 0$, i.e., $u = \partial_2\Phi(V, S^*)$. Our goal next is to upper bound S^* by a function of u. Again, we analyze two cases.

Case 1. If S^* satisfies $S^* \leq \sqrt{4\alpha(V+z+kS^*)}$, then by taking the square on both sides and regrouping the terms, we have $(S^*)^2 - 4\alpha kS^* - 4\alpha(V+z) < 0$. Solving this quadratic inequality,

$$\begin{aligned} S^* &\leq \frac{1}{2}\left[4\alpha k + \sqrt{(4\alpha k)^2 + 16\alpha(V+z)}\right] \\ &= 2\alpha k + \sqrt{4\alpha^2k^2 + 4\alpha(V+z)} \\ &\leq 4\alpha k + \sqrt{4\alpha(V+z)}. \end{aligned}$$

Case 2. If S^* satisfies $S^* \geq \sqrt{4\alpha(V+z+kS^*)}$, then same as the earlier analysis in the present proof, Eq.(10), we have

$$u \geq \frac{\varepsilon}{2}\mathrm{erfi}\left(\frac{S^*}{\sqrt{4\alpha(V+z+kS^*)}}\right).$$

For conciseness, define the notation $p = \mathrm{erfi}^{-1}(2u\varepsilon^{-1})$. Then, $(S^*)^2 - 4\alpha kp^2S^* - 4\alpha p^2(V+z) \leq 0$. Solving the quadratic inequality,

$$\begin{aligned}
S^* &\leq \frac{1}{2}\left[4\alpha kp^2 + \sqrt{(4\alpha kp^2)^2 + 16\alpha p^2(V+z)}\right] \\
&\leq 2\alpha kp^2 + \sqrt{4\alpha^2k^2p^4 + 4\alpha p^2(V+z)} \\
&\leq 4\alpha kp^2 + \sqrt{4\alpha(V+z)}p.
\end{aligned}$$

Now we can combine the above two cases. Specifically, $p \leq 1 + \sqrt{\log(2u\varepsilon^{-1}+1)}$ due to Lemma 10. Therefore,

$$S^* \leq 4\alpha k\left(1+\sqrt{\log(2u\varepsilon^{-1}+1)}\right)^2 + \sqrt{4\alpha(V+z)}\left(1+\sqrt{\log(2u\varepsilon^{-1}+1)}\right).$$

Define the RHS as $\bar{S}$. Then, from the definition of the Fenchel conjugate,

$$\begin{aligned}
\sup_{S\in[-G,\infty)}[uS - \Phi(V,S)] &= uS^* - \Phi(V,S^*) \\
&= uS^* - \varepsilon\sqrt{\alpha(V+z+kS^*)}\left[2\int_0^{\frac{S^*}{\sqrt{4\alpha(V+z+kS^*)}}}\mathrm{erfi}(u)du - 1\right] \\
&\leq u\bar{S} + \varepsilon\sqrt{\alpha(V+z+k\bar{S})}.
\end{aligned}$$

Plugging in $\bar{S}$ completes the proof. ■

Combining the previous two lemmas, the following lemma characterizes the regret of the base algorithm.

Lemma 13 (Regret of the base algorithm) *With $\varepsilon > 0$, $\alpha > \frac{1}{2}$, $k_t = 2h_t$ and $z_t = \frac{12\alpha+4}{2\alpha-1}h_t^2$, Algorithm 1 guarantees for all $T \in \mathbb{N}_+$ and $\tilde{u} \geq 0$,*

$$\sum_{t=1}^T \tilde{l}_t(\tilde{y}_t - \tilde{u}) \leq \varepsilon\sqrt{\alpha\left(\tilde{V}_T + z_T + k_T\bar{S}\right)} + \tilde{u}\bar{S},$$

where

$$\bar{S} = 4\alpha k_T\left(1+\sqrt{\log(2\tilde{u}\varepsilon^{-1}+1)}\right)^2 + \sqrt{4\alpha\left(\tilde{V}_T+z_T\right)}\left(1+\sqrt{\log(2\tilde{u}\varepsilon^{-1}+1)}\right).$$

Proof [Proof of Lemma 13] First, we can obtain a cumulative loss upper bound by simply summing the one-step potential verification bound (Lemma 4). Letting $c = -\tilde{l}_t$, $V = \tilde{V}_{t-1}$ and $S = \tilde{S}_{t-1}$ in Lemma 4,

$$\tilde{l}_t\tilde{y}_t \leq \Phi_{t-1}\left(\tilde{V}_{t-1}, \tilde{S}_{t-1}\right) - \Phi_t\left(\tilde{V}_t, \tilde{S}_t\right),$$

$$\sum_{t=1}^{T} \tilde{l}_t \tilde{y}_t \leq \Phi_0(0,0) - \Phi_T\left(\tilde{V}_T, \tilde{S}_T\right).$$

Then, the proof follows from a loss-regret duality and the Fenchel conjugate computation from Lemma 12.

$$\begin{aligned}
\sum_{t=1}^{T} \tilde{l}_t(\tilde{y}_t - \tilde{u}) &\leq \tilde{S}_T \tilde{u} + \Phi_0(0,0) - \Phi_T\left(\tilde{V}_T, \tilde{S}_T\right) \\
&\leq \Phi_0(0,0) + \sup_{S \in [-h_T, \infty)} \left[S\tilde{u} - \Phi_T\left(\tilde{V}_T, S\right)\right] \\
&\leq -\varepsilon\sqrt{\alpha \cdot \frac{12\alpha + 4}{2\alpha - 1} h_0^2} + \varepsilon\sqrt{\alpha\left(\tilde{V}_T + z_T + k_T \bar{S}\right)} + \tilde{u}\bar{S} \\
&\leq \varepsilon\sqrt{\alpha\left(\tilde{V}_T + z_T + k_T \bar{S}\right)} + \tilde{u}\bar{S}.
\end{aligned}$$

Plugging in $\bar{S}$ from Lemma 12 completes the proof. ∎

Appendix D. Analysis of the meta algorithm

First, we prove Lemma 3, which certifies the well-posedness of our algorithm.

Lemma 14 (Well-posedness) *The surrogate loss $\tilde{l}_t$ defined in Algorithm 2 satisfies $\sum_{i=1}^{t} \tilde{l}_i \leq h_t$ for all t.*

Proof [Proof of Lemma 3] First, notice that $|\tilde{l}_t| \leq |l_t| = |\langle g_t, w_t \rangle| \leq h_t$.

Next, we prove by induction. Consider $-\sum_{i=1}^{t-1} \tilde{l}_i$, which is defined as $\tilde{S}_{t-1}$ in the base algorithm (Algorithm 1). Suppose $\tilde{S}_{t-1} \geq -h_{t-1}$, which trivially holds at $t = 1$. Let us analyze two cases.

- If $\tilde{S}_{t-1} \geq 0$, then $-\sum_{i=1}^{t} \tilde{l}_i = \tilde{S}_{t-1} - \tilde{l}_t \geq \tilde{S}_{t-1} - |\tilde{l}_t| \geq -h_t$.
- If $-h_{t-1} \leq \tilde{S}_{t-1} < 0$, then due to Lemma 8, the prediction $\tilde{y}_t$ of the base algorithm satisfies $\tilde{y}_t < 0$. The meta algorithm projects it to $y_t = 0$. Then, due to our definition of $\tilde{l}_t$ in the meta algorithm,
$$\tilde{l}_t = \begin{cases} l_t, & l_t \leq 0, \\ 0, & \text{else}, \end{cases}$$
which is non-positive. Therefore, $-\sum_{i=1}^{t} \tilde{l}_i = \tilde{S}_{t-1} - \tilde{l}_t \geq \tilde{S}_{t-1} \geq -h_{t-1} \geq -h_t$.

An induction completes the proof. ∎

Next, we prove our main result.

Theorem 5 (Main result) *With $\varepsilon > 0$, $\alpha > \frac{1}{2}$, $k_t = 2h_t$ and $z_t = \frac{12\alpha+4}{2\alpha-1} h_t^2$, Algorithm 2 guarantees for all $T \in \mathbb{N}_+$ and $u \in \mathbb{R}^d$,*

$$\text{Regret}_T(Env, u) \leq \varepsilon\sqrt{\alpha\left(V_T + z_T + k_T \bar{S}\right)} + \|u\|\left(\bar{S} + 2\sqrt{2V_T}\right),$$

where

$$\bar{S} = 4\alpha k_T\left(1 + \sqrt{\log(2\|u\|\varepsilon^{-1} + 1)}\right)^2 + \sqrt{4\alpha\left(V_T + z_T\right)}\left(1 + \sqrt{\log(2\|u\|\varepsilon^{-1} + 1)}\right).$$

Proof [Proof of Theorem 5] Since the meta algorithm simply applies two existing black-box reductions (Cutkosky and Orabona, 2018; Cutkosky, 2020), the proof is straightforward given Lemma 13, the regret bound of the 1D base algorithm. First, due to a polar decomposition theorem (Cutkosky and Orabona, 2018, Theorem 2), the regret can be decomposed into the regret of $\mathcal{A}_B$ with respect to $u/\|u\|$, plus the regret of y_t with respect to $\|u\|$. Then, the latter is upper-bounded by the regret of $\tilde{y}_t$ evaluated on the surrogate losses $\tilde{l}_t$ – this is because our definition of y_t and $\tilde{l}_t$ follows the procedure of (Cutkosky, 2020, Theorem 2), where a convex constraint can be added to an unconstrained algorithm without changing its regret bound. In summary, we have

$$\begin{aligned}\mathrm{Regret}_T(Env, u) &= \sum_{t=1}^{T} l_t(y_t - \|u\|) + \|u\| \sum_{t=1}^{T} \langle g_t, w_t - u/\|u\|\rangle \\ &\leq \sum_{t=1}^{T} \tilde{l}_t(\tilde{y}_t - \|u\|) + \|u\| \sum_{t=1}^{T} \langle g_t, w_t - u/\|u\|\rangle\,.\end{aligned}$$

The two regret terms on the RHS represent the regret bound of $\mathcal{A}_{1d}$ and $\mathcal{A}_{\mathsf{B}}$, respectively.

Now, the first term follows from Lemma 13, where

$$\tilde{V}_T = \sum_{t=1}^{T} \tilde{l}_t^2 \leq \sum_{t=1}^{T} l_t^2 = \sum_{t=1}^{T} \langle g_t, w_t\rangle^2 \leq V_T.$$

As for the regret of $\mathcal{A}_{\mathsf{B}}$, due to (Orabona, 2023, Theorem 4.14),

$$\sum_{t=1}^{T} \langle g_t, w_t - u/\|u\|\rangle \leq 2\sqrt{2V_T}.$$

Combining these two components completes the proof. ∎

Finally, let us use asymptotic orders to make this bound a bit more interpretable. Consider the regime of large $\|u\|$ and V_T, i.e., $\|u\| \gg \varepsilon$ and $V_T \gg h_T^2$. We preserve the dependence on ε^{-1}, as ε is a "confidence hyperparameter" that can get arbitrarily close to 0. In contrast, α is a moderate constant slightly larger than $\frac{1}{2}$, therefore we subsume it by the big-Oh.

Using $\log(1+x) \leq x$, we can crudely bound $\bar{S}$ in Theorem 5 by

$$\begin{aligned}\bar{S} &\leq 8\alpha h_T \left(1 + \sqrt{2\|u\|\varepsilon^{-1}}\right)^2 + \sqrt{4\alpha(V_T + z_T)}\left(1 + \sqrt{2\|u\|\varepsilon^{-1}}\right) \\ &= 8\alpha h_T + \sqrt{4\alpha(V_T + z_T)} + o\left(\|u\|\varepsilon^{-1}\sqrt{V_T}\right).\end{aligned}$$

Plugging this crude bound of $\bar{S}$ into the first term of the regret bound, we have

$$\begin{aligned}\mathrm{Regret}_T(Env, u) &\leq \varepsilon\sqrt{\alpha\left(V_T + z_T + 16\alpha h_T^2 + 2h_T\sqrt{4\alpha(V_T + z_T)}\right)} + \varepsilon\sqrt{o\left(\|u\|\varepsilon^{-1}\sqrt{V_T}\right)} \\ &\quad + \|u\|\left(\bar{S} + 2\sqrt{2V_T}\right) \\ &\leq \varepsilon\sqrt{\alpha}\left(\sqrt{V_T + z_T} + 4\sqrt{\alpha}h_T\right) + o\left(\|u\| V_T^{1/4}\right) + \|u\|\left(\bar{S} + 2\sqrt{2V_T}\right).\end{aligned}$$

Next, notice that $\bar{S} = O\left(\sqrt{V_T \log(\|u\| \varepsilon^{-1})} \vee h_T \log(\|u\| \varepsilon^{-1})\right)$. Using it to replace the remaining $\bar{S}$ above,

$$\begin{aligned}\mathrm{Regret}_T(Env, u) \le \varepsilon\left(\sqrt{\alpha(V_T + z_T)} + 4\alpha h_T\right) + o\left(\|u\| V_T^{1/4}\right) \\ + \|u\| O\left(\sqrt{V_T \log(\|u\| \varepsilon^{-1})} \vee h_T \log(\|u\| \varepsilon^{-1})\right).\end{aligned}$$

The second term can be assimilated into the third term. The result becomes

$$\begin{aligned}\mathrm{Regret}_T(Env, u) \le \varepsilon\left(\sqrt{\alpha\left(V_T + z_T\right)} + 4\alpha h_T\right) \\ + \|u\| O\left(\sqrt{V_T \log(\|u\| \varepsilon^{-1})} \vee h_T \log(\|u\| \varepsilon^{-1})\right). \quad (11)\end{aligned}$$

Note that this bound is not only valid for large $\|u\|$, but also valid when $u = 0$ (this can be directly verified from Theorem 5). Therefore, it characterizes the loss-regret tradeoff.