Gradient-Variation Regret Bounds for Unconstrained Online Learning

Yuheng Zhao* ZHAOYH@LAMDA.NJU.EDU.CN
State Key Laboratory for Novel Software Technology, Nanjing University
School of Artificial Intelligence, Nanjing University

Andrew Jacobsen* CONTACT@ANDREW-JACOBSEN.COM
Università degli Studi di Milano & Politecnico di Milano

Nicolò Cesa-Bianchi NICOLO.CESA-BIANCHI@UNIMI.IT
Università degli Studi di Milano & Politecnico di Milano

Peng Zhao† ZHAOP@LAMDA.NJU.EDU.CN
State Key Laboratory for Novel Software Technology, Nanjing University
School of Artificial Intelligence, Nanjing University

Editors: Steve Hanneke and Tor Lattimore

Abstract

We develop parameter-free algorithms for unconstrained online learning with regret guarantees that scale with the gradient variation $V_T(u) = \sum_{t=2}^T \|\nabla f_t(u) - \nabla f_{t-1}(u)\|^2$. For L-smooth convex losses, we provide fully-adaptive algorithms achieving regret of $\widetilde{O}(\|u\| \sqrt{V_T(u)} + L\|u\|^2 + G^4)$ without requiring prior knowledge of comparator norm $\|u\|$, Lipschitz constant G, or smoothness L. The update in each round can be computed efficiently via a closed-form expression. Our results extend to dynamic regret and find immediate implications for the stochastically-extended adversarial (SEA) model, which significantly improves upon the previous best-known result (Wang et al., 2025).

1. Introduction

Online learning (Cesa-Bianchi and Lugosi, 2006; Orabona, 2019) is a fundamental paradigm in machine learning for modeling and analyzing sequential prediction and decision-making problems. An online learning process is formalized as an interaction between a learner and the environment. At iteration $t \in [T]$, the learner chooses a decision w_t from a feasible domain $\mathcal{W} \subseteq \mathbb{R}^d$, after which the environment reveals a loss function $f_t : \mathcal{W} \to \mathbb{R}$, and the learner incurs a loss $f_t(w_t)$. A general performance metric is the *dynamic regret* (Zinkevich, 2003; Zhang et al., 2018), which evaluates the cumulative loss against a sequence of comparators:

$$\textsc{Reg}_T(u_{1:T}) \triangleq \sum_{t=1}^T f_t(w_t) - \sum_{t=1}^T f_t(u_t), \tag{1}$$

where the comparators $u_{1:T} \triangleq (u_1, \ldots, u_T)$ in $\mathcal{W}$ are unknown, and their variability is typically measured by the path length $P_T(u_{1:T}) \triangleq \sum_{t=2}^T \|u_t - u_{t-1}\|$. When restricting to a fixed comparator $u \in \mathcal{W}$, dynamic regret reduces to the standard notion of static regret, denoted by $\textsc{Reg}_T(u)$.

* Equal Contribution.
† Corresponding Author.

1.1. Gradient-Variation Adaptivity

A rich theory has been developed for both static and dynamic regret minimization in the last decades (Zinkevich, 2003; Shalev-Shwartz, 2012; Orabona, 2019; Hazan, 2022). Notably, recent studies suggest the importance of *problem-dependent adaptivity* (de Rooij et al., 2014; Foster et al., 2015; Roughgarden, 2021; Zhao et al., 2024), which aims to achieve tighter bounds for benign problem instances while preserving minimax optimality in the worst case. Among various measures of problem difficulty, a key quantity is the *gradient variation* (Chiang et al., 2012; Yang et al., 2014), defined as

$$V_T^+ \triangleq \sum_{t=2}^{T} \sup_{w \in \mathcal{W}} \|\nabla f_t(w) - \nabla f_{t-1}(w)\|^2 , \tag{2}$$

which captures how the problem evolves over time in terms of the function gradients. When the feasible domain is bounded (i.e., $\|x - y\| \le D$ for all $x, y \in \mathcal{W}$) and the online functions are convex and L-smooth (i.e., $\|\nabla f_t(x) - \nabla f_t(y)\| \le L \|x - y\|$ for all $x, y \in \mathcal{W}$), it is known that optimal $O\big(D\sqrt{V_T^+} + LD^2\big)$ static regret (Chiang et al., 2012) and $O\big(D\sqrt{V_T^+(1 + P_T)} + L(D^2 + DP_T)\big)$ dynamic regret (Zhao et al., 2020) can be achieved. Gradient-variation-based online learning has attracted growing interest in recent years. In particular, Zhao et al. (2020) introduced gradient variation into dynamic regret minimization and proposed novel techniques that have inspired many subsequent works (Zhang et al., 2022; Sachs et al., 2022; Xie et al., 2024; Mhaisen and Iosifidis, 2025; Zhao et al., 2025c; Wang et al., 2025; Zhao et al., 2025b; Yu et al., 2026). Gradient-variation adaptivity has been revealed to have tight connections to a broad class of optimization problems. For example, this adaptivity is shown to be crucial for achieving fast convergence rates in minimax optimization/games (Syrgkanis et al., 2015; Zhang et al., 2022), as well as for attaining acceleration in offline smooth convex optimization (Cutkosky, 2019a; Zhao et al., 2025c). Moreover, recent work demonstrates that controlling gradient-variation regret is essential for obtaining adaptive guarantees under the *Stochastically Extended Adversarial* (SEA) model, which interpolates between adversarial online optimization and stochastic convex optimization (Sachs et al., 2022; Chen et al., 2024).

1.2. Parameter-Free Online Learning

Most existing gradient-variation online learning algorithms rely on the assumption of a bounded feasible domain. In many practical scenarios, however, the domain is naturally *unbounded*, making it infeasible to impose an *a priori* diameter upper bound D on the comparator norm $\|u\|$. This limitation motivates the study of *parameter-free online learning* (Chaudhuri et al., 2009; Mcmahan and Streeter, 2012; Orabona, 2013; McMahan and Orabona, 2014; Orabona and Pál, 2016; Jacobsen and Cutkosky, 2022; Cutkosky and Mhammedi, 2024), which aims to design algorithms that do not require such problem-dependent quantities as inputs.

A central requirement of parameter-free algorithms is being *comparator-adaptive*: the ability to achieve regret bounds that scale favorably with the unknown comparator norm $\|u\|$. When the Lipschitz constant G of loss functions is further unknown,[1] a parameter-free online algorithm must also be *Lipschitz-adaptive*, meaning it attains the desired regret without prior knowledge of G. Algorithms that satisfy both comparator-adaptivity and Lipschitz-adaptivity are sometimes referred to as *fully-adaptive* methods (Cutkosky and Mhammedi, 2024). The best known fully-

1. The Lipschitz constant G is used to denote the empirical gradient norm $\max_t \|\nabla f_t(w_t)\|$, when there's no ambiguity.

adaptive result is achieved by Cutkosky and Mhammedi (2024), who attained a static regret of order $\text{REG}_T(u) \le \widetilde{O}\big(\|u\| G\sqrt{T} + \|u\|^2 + G^2\big)$ with $\widetilde{O}(\cdot)$ omitting poly-logarithmic factors.

For gradient-variation regret over unbounded domains, two previous works are most relevant. Jacobsen and Cutkosky (2022) proposes a mirror descent-based algorithm that obtains comparator-adaptive gradient-variation regret, but requires full-information feedback $f_t(\cdot)$ since it applies an *implicit* update (Campolongo and Orabona, 2020) in its optimistic step. Wang et al. (2025) study the stochastically-extended adversarial (SEA) model—a closely related setting where gradient-variation online learning plays a central role—and attain a comparator-adaptive regret bound of $\widetilde{O}\big(\|u\| \sqrt{V_T^+} + \|u\|^2\big)$ when the Lipschitz constant is known, using only first-order feedback. However, their algorithm relies on a two-layer meta-base structure that maintains $O(\log^2 T)$ base learners, which is significantly more expensive than the $O(d)$ per-round computation used to obtain gradient-variation bounds in the bounded domain setting. Moreover, when further targeting Lipschitz adaptivity, their method suffers a significant deterioration of the leading term to $\|u\|^2 \sqrt{V_T^+}$, giving a sub-optimal dependence on the comparator norm. As such, a natural open question arises:

Is it possible to achieve gradient-variation regret in a fully adaptive, parameter-free manner over unbounded domains with an efficient algorithm?

1.3. Our Contributions

In this paper, we provide an affirmative answer by developing the *first* fully-adaptive algorithm for gradient-variation online learning in the unconstrained domain, requiring no prior knowledge of the comparator norm $\|u\|$, the Lipschitz constant G, or the smoothness parameter L. Notably, our algorithm is efficient, with a closed-form update that can be computed in $O(d)$ time per round.

To begin, we clarify the definition of gradient variation in the context of unbounded domains. The original definition of V_T^+ in Eq. (2) does not work well in this context, as the $\mathbb{R}^d$-domain may easily cause it to scale as $O(G^2T)$, reducing back to a non-adaptive worst-case dependence. Instead, we introduce a more appropriate definition that supports arbitrary comparators $u_{1:T} \in \mathbb{R}^d$ as:

$$V_T(u_{1:T}) \triangleq \sum_{t=2}^{T} \|\nabla f_t(u_{t-1}) - \nabla f_{t-1}(u_{t-1})\|^2, \tag{3}$$

which quantifies the gradient variation between consecutive functions on the sequence. The time-varying comparators primarily serve to accommodate dynamic regret. For static regret over a bounded domain $\mathcal{W}$, it captures the original definition in Eq. (2) since $V_T(u) \le V_T^+$ for any $u \in \mathcal{W}$.

We provide both comparator-adaptive and fully-adaptive gradient-variation regret bounds. A summary of our contributions and a comparison to prior works can be found in Table 1.

Comparator-adaptive regret. We first propose an *optimistic-to-gradient-variation reduction*, which transforms the challenge of achieving comparator or fully-adaptive gradient-variation regret bounds into the problem of attaining standard optimistic regret in online learning. This reduction leverages the negative Bregman divergence terms that naturally arises in regret linearization. As a warm-up, we show that by instantiating this reduction with the existing optimistic comparator-adaptive algorithm from Jacobsen and Cutkosky (2022), we obtain a *comparator-adaptive* gradient-variation bound of $\widetilde{O}(\|u\| \sqrt{V_T(u)} + L\|u\|^2 + G\|u\|)$ with *efficient* closed-form updates requiring only $O(d)$ computational cost per round, significantly improving the efficiency of the best-known prior result (Wang et al., 2025), which maintains a meta-base ensemble structure requiring $O(\log^2 T)$ computation on each round.

Table 1: Comparison of parameter-free gradient-adaptive regret, where $g_t \in \partial f_t(w_t)$ denotes the gradient feedback on round t. "Efficiency" denotes the order of per-round computational cost.

Setting	Reference	Regret in $\widetilde{O}(\cdot)$-notation	Efficiency
Comparator Adaptive	Jacobsen and Cutkosky (2022)	$\|u\| \sqrt{\sum_{t=2}^T \|g_t - g_{t-1}\|^2} + G\|u\|$	$O(d)$
	Wang et al. (2025)	$\|u\| \sqrt{V_T^+} + L^2 \|u\|^2 / G^2 + G^2 \|u\|^2$	$O(d \log^2 T)$
	Ours, Theorem 2	$\|u\| \sqrt{V_T(u)} + L\|u\|^2 + G\|u\|$	$O(d)$
Fully Adaptive	Cutkosky and Mhammedi (2024)	$\|u\| \sqrt{\sum_{t=1}^T \|g_t\|^2} + \gamma \|u\|^2 / \epsilon + \epsilon G^2 / \gamma$	$O(d)$
	Wang et al. (2025)	$\|u\|^2 \sqrt{V_T^+} + L^2 \|u\|^4 + G\|u\|^3 + G^2 \|u\|^2 + G^2 \sqrt{\sum_{t=1}^T \|g_t\|}$	$O(d)$
	Ours, Theorem 3	$\|u\| \sqrt{V_T(u)} + L\|u\|^2 + \gamma \|u\|^2 / \epsilon + \epsilon G^2 / \gamma$	$O(d + \log T)$
	Ours, Theorem 4	$\|u\| \sqrt{V_T(u)} + L\|u\|^2 + \gamma \|u\|^2 / \epsilon + \epsilon G^4 / \gamma^3$	$O(d)$

Fully-adaptive regret. We then design fully-adaptive methods for gradient-variation regret that require no prior knowledge of $\|u\|$ or G, which is the main focus of this work and constitutes the key technical contributions of the paper. We provide two algorithms. *(i)* Starting from the comparator-adaptive optimistic result of Jacobsen and Cutkosky (2022), we propose a simple algorithm to incorporate Lipschitz adaptivity, which is accomplished by a *virtual clipping technique* over the optimistic gap in the regularizer and adding a quadratic penalty to form a hybrid regularizer. This yields a fully-adaptive gradient-variation bound $\widetilde{O}(\|u\| \sqrt{V_T(u)} + L\|u\|^2 + G^2)$, strictly improving the prior best-known result of Wang et al. (2025) whose leading term is $\|u\|^2 \sqrt{V_T^+}$. However, its update lacks a closed-form expression and may incur a computational cost of $O(d + \log T)$ per round. *(ii)* We extend the fully-adaptive non-optimistic algorithm of Cutkosky and Mhammedi (2024) by equipping it with optimistic guarantees, which is achieved by a *refined optimistic reduction* (Cutkosky, 2019c). The resulting algorithm admits closed-form updates in $O(d)$ time per round and achieves regret $\widetilde{O}(\|u\| \sqrt{V_T(u)} + L\|u\|^2 + G^4)$, though suffers a slightly larger lower-order dependence of G^4 instead of G^2. A comparison of the results can be found in second half of Table 1.

Dynamic regret and the SEA model. We extend our fully-adaptive algorithm to optimize *dynamic regret* in Eq. (1). By combining it with a one-dimensional reduction (Cutkosky and Orabona, 2018) and an anytime Lipschitz-adaptive algorithm over the unit ball, we obtain a dynamic regret bound with leading terms as $\widetilde{O}(\sqrt{(M^2 + MP_T)V_T(u_{1:T})} + L(M^2 + MP_T) + GP_T)$, where $M = \max_t \|u_t\|$, using $O(d \log t)$ time per iteration t. We further apply this to the *Stochastically Extended Adversarial* (SEA) model (Sachs et al., 2022), an intermediate setting between adversarial and stochastic convex optimization. Our gradient-variation dynamic regret bound translates to the SEA setting by replacing $V_T(u_{1:T})$ with the sum of stochastic variance and adversarial variation. This yields the *first* dynamic regret guarantee for the SEA model in the unconstrained setting and substantially improves upon Wang et al. (2025), who only provide static regret and suffer a quadratic dependency $\|u\|^2 \sqrt{V_T^+}$, where V_T^+ can be much larger than our problem-dependent $V_T(u_{1:T})$.

Notations. Let $[N] = \{1, \ldots, N\}$ denote the set of integers up to $N \geq 1$. We define $\log_+(\cdot) \triangleq \max\{1, \log(\cdot)\}$. For an indexed collection $\{a_t\}_{t\geq 1}$, we also write $(a_t)_t$ when it is clear from the context. The Bregman divergence with respect to a differentiable convex function ψ is $\mathcal{D}_\psi(x, y) = \psi(x) - \psi(y) - \langle \nabla\psi(y), x - y\rangle$. Unless otherwise specified, $\|\cdot\|$ denotes the ℓ_2-norm. Finally, the $\widetilde{O}(\cdot)$-notation suppresses logarithmic factors, and we treat $\log\log T$ as a constant.

Organization. The rest of the paper is organized as follows. Section 2 provides the results for comparator-adaptive gradient-variation bounds. In Section 3, we propose two fully-adaptive algorithms and their analysis. Section 4 presents the applications to dynamic regret and the SEA model. Section 5 concludes the paper. All the proofs are deferred to the appendix.

2. Warm Up: Comparator-Adaptive Gradient-Variation Bounds via Optimism

In this section, we first establish a simple black-box reduction that obtains comparator-dependent gradient variation bounds from standard optimistic OLO bounds. We then present as an immediate application a comparator-adaptive gradient variation bound using Jacobsen and Cutkosky (2022).

2.1. Gradient-Variation Black-Box Reduction

Instead of building gradient-variation bounds from scratch, we provide a simple black-box reduction that leverages the existing results in *optimistic* online learning (Jacobsen and Cutkosky, 2022). In addition to receiving standard gradient feedback $g_t \in \partial f_t(w_t)$, an optimistic online algorithm receives an optimistic "hint" vector h_t at the beginning of each round, which can be leveraged to achieve improved guarantees when the hints accurately predict the next gradient, $h_t \approx g_t$. The following theorem shows that any online algorithm achieving a standard optimistic regret bound (i.e., regret scaling with $\sqrt{\sum_t \|g_t - h_t\|^2}$) can also deliver a gradient-variation regret bound when the losses are L-smooth by setting the optimistic hint as $h_t = g_{t-1}$.

Theorem 1 (Optimistic-to-Gradient-Variation Reduction) *Let $\mathcal{U}$ be a class of sequences in $\mathbb{R}^d$, and let $\mathcal{A}$ be an online learning algorithm that receives $\{g_t\}_{t=1}^T$ as gradients and takes $\{h_t\}_{t=1}^T$ as optimistic hints. Suppose $\mathcal{A}$ guarantees that for any sequence $u_{1:T} \in \mathcal{U}$ that*

$$\sum_{t=1}^{T} \langle g_t, w_t - u_t \rangle \le A_T(u_{1:T}) + B_T(u_{1:T}) \sqrt{\sum_{t=1}^{T} \|g_t - h_t\|^2}. \tag{4}$$

Then for any sequence of G-Lipschitz, L-smooth convex functions $\{f_t\}_{t=1}^T$, by setting $g_t = \nabla f_t(w_t)$ and $h_t = g_{t-1}$, $\mathcal{A}$ achieves

$$\text{REG}_T(u_{1:T}) \le A_T(u_{1:T}) + 4LB_T(u_{1:T})^2 + 2B_T(u_{1:T}) \sqrt{G^2 + V_T(u_{1:T}) + L^2 P_T^{\|\cdot\|^2}(u_{1:T})}, \tag{5}$$

where $P_T^{\|\cdot\|^2}(u_{1:T}) \triangleq \sum_{t=2}^T \|u_t - u_{t-1}\|^2$ is the squared path length.

The proof is provided in Appendix A.1, where the key idea is to appropriately decompose the $\|g_t - g_{t-1}\|^2$ term and leverage a negative Bregman divergence term that naturally arises from regret linearization (Yan et al., 2024). While the theorem is framed in terms of dynamic regret, it also applies to static regret by considering the class of fixed sequences $u_1 = \cdots = u_T$ in $\mathbb{R}^d$. Similarly, the theorem can also be applied to constrained settings, so long as smoothness is still defined over the entire space (or at least over a slightly augmented space; see, e.g., Yan et al. (2024, Appendix A)).

We also remark that the result above simultaneously implies *small-loss* bounds, scaling with $F_T(u_{1:T}) \triangleq \sum_{t=1}^T (f_t(u_t) - \inf_{w \in \mathbb{R}^d} f_t(w))$. Indeed, in Appendix A.1 we show that the obtained bounds more generally scale with $O\big(\sqrt{\min\{LF_T(u_{1:T}), V_T(u_{1:T})\}}\big)$. Thus, each of our results in what follows simultaneously achieve small-loss bounds. Throughout the paper we focus our discussion on the gradient variation bounds for simplicity.

2.2. Implication to Comparator-Adaptive Gradient-Variation Regret

When the Lipschitz constant G is known *a priori*, we can immediately apply Theorem 1 to achieve comparator-adaptive gradient variation bounds using existing algorithms. Indeed, Jacobsen and Cutkosky (2022) showed that optimistic Follow-The-Regularized-Leader (FTRL) with a carefully-designed regularizer can achieve comparator-adaptive optimistic regret.[2] Briefly, it updates by

$$w_t = \arg\min_{w \in \mathbb{R}^d} \left\langle h_t + \sum_{s=1}^{t-1} g_s, w \right\rangle + \psi_t^{\text{PF}}(w, \alpha_t), \quad \psi_t^{\text{PF}}(w, \alpha) = \int_0^{\|w\|} \min_{\eta \le 1/G} \left[\frac{\log(x/\alpha+1)}{\eta} + \eta \bar{V}_t \right] \mathrm{d}x. \tag{6}$$

Here, $\psi_t^{\text{PF}}(w, \alpha_t)$ is a *parameter-free*[3] regularizer, $\bar{V}_t \propto \sum_{s=1}^{t-1} \|g_s - h_s\|^2$ is the empirical gradient variation, and $(\alpha_t)_t$ is a non-increasing sequence. Intuitively, this regularizer applies weaker regularization compared to the typical quadratic regularizer $\frac{1}{\eta}\|w\|^2$, adaptively balancing the trade-off between the $\|u\|$-term and the empirical gradient variation $\bar{V}_T$. This ensures that $\psi_T^{\text{PF}}(u) \le \widetilde{O}(\|u\| \sqrt{\bar{V}_T})$ without any explicit tuning based on $\|u\|$. With an analysis similar to that of Jacobsen and Cutkosky (2022), this algorithm guarantees:

$$\text{REG}_T(u) \le \widetilde{O}\left(\|u\| \sqrt{\sum_{t=1}^T \|g_t - h_t\|^2} + G\|u\| + \epsilon G \right). \tag{7}$$

A detailed specification of the algorithm can be found in Algorithm 4 of Appendix A.3. Combined with the optimistic-to-gradient-variation reduction in Theorem 1, we obtain the following comparator-adaptive gradient-variation regret, whose proof is provided in Appendix A.4.

Theorem 2 (Comparator-Adaptive Gradient-Variation Regret) *For any $u \in \mathbb{R}^d$, Algorithm 4 (in Appendix A.3) with $h_t = g_{t-1}$ guarantees*

$$\text{REG}_T(u) \le \widetilde{O}\left(\|u\| \sqrt{V_T(u) \left(\log_+ \frac{\|u\|\sqrt{T}}{\epsilon} \right)} + L\|u\|^2 + G\|u\| + \epsilon G \right).$$

where we keep the log factors of the dominant term in the $\widetilde{O}(\cdot)$-notation for clarity. Moreover, the algorithm admits an efficient closed-form update with $O(d)$ time per iteration.

Importantly, albeit with a seemingly complex form, the update Eq. (6) admits an efficient closed-form update formula, as shown in Algorithm 4, requiring only $O(d)$ time per round. In contrast, the prior best-known result (Wang et al., 2025) requires $O(d \log^2 T)$ computation per-round and is significantly less efficient than ours. This inefficiency arises from their use of an online ensemble that maintains $O(\log^2 T)$ base learners: they derive a range for the comparator norm $\|u\|$ to prevent the bound from being vacuous, discretize this range, and run a base learner in a bounded domain for each candidate diameter. A meta-algorithm is then used to combine the outputs of base learners.

2. Jacobsen and Cutkosky (2022) primarily focus on a mirror-descent-based formulation for dynamic regret minimization. Here, since we focus on static regret, we provide an analysis based on FTRL which significantly streamlines their arguments.
3. The regularizers associated with comparator-adaptive guarantees are also sometimes referred to as *linearithmic* regularizers, due to their log-linear form (Orabona and Pál, 2021).

3. Fully-adaptive Gradient-Variation Bounds

Now we target fully-adaptive gradient-variation regret, where neither the comparator norm $\|u\|$ nor the Lipschitz constant G is known in advance. Leveraging the optimistic-to-gradient-variation reduction, it suffices to design a *fully-adaptive optimistic* algorithm. There are two possible approaches:

(i) Begin with the comparator-adaptive optimistic algorithm (Jacobsen and Cutkosky, 2022) and enhance it with Lipschitz adaptivity, as described in Section 3.1.

(ii) Alternatively, start from the fully-adaptive non-optimistic algorithm of Cutkosky and Mhammedi (2024) and extend it to its optimistic counterpart, as presented in Section 3.2.

We will discuss the challenges inherent in these extensions when using existing techniques and introduce new ideas to overcome them. The algorithm in Section 3.1 is conceptually simple and enjoys optimal regret bounds, but it lacks a closed-form update and requires a line search to implement. Alternatively, the approach in Section 3.2 extends the algorithm of Cutkosky and Mhammedi (2024) using a new optimistic reduction, leading to an algorithm having an efficient closed-form update while maintaining nearly the same regret bounds, with only slight deterioration in horizon-independent lower-order terms.

3.1. A Simple Algorithm via Virtual Clipping

When the Lipschitz constant $G \geq \max_{t\in[T]} \|g_t\|$ is unknown, it is natural to employ the observable maximum gradient norm $\widehat{G}_t = \max_{s<t} \|g_s\|$ as a guess of G. This motivates the *gradient clipping* technique from Cutkosky (2019b), which is now a standard approach to obtain Lipschitz-adaptivity. In particular, Cutkosky (2019b) proposes feeding the online algorithm with *clipped* gradients $\widetilde{g}_t \triangleq g_t \min\{1, \widehat{G}_t / \|g_t\|\}$, which are instead bounded by the *known* Lipschitz constant $\widehat{G}_t$. Then this clipping reduces the problem to regret against a sequence of clipped gradients in a *black-box* manner:

$$\sum_{t=1}^{T} \langle g_t, w_t - u\rangle \leq G\left(\|u\| + \max_{t\in[T]} \|w_t\|\right) + \sum_{t=1}^{T} \langle \widetilde{g}_t, w_t - u\rangle . \tag{8}$$

However, in unbounded domains the term $\max_t \|w_t\|$ is typically difficult to control, leading to an additional *cubic* penalty $G\|u\|^3$ in prior works (Cutkosky, 2019b; Mhammedi and Koolen, 2020).

Our key insight is that instead of applying a black-box clipping argument, feeding the *true* gradients g_t to the algorithm while including an additional quadratic regularizer allows one to apply a *virtual* clipping argument which incurs only an $O(\|u\|^2)$ overhead. Crucially, the $\|u\|^3$ vs. $\|u\|^2$ dependency is a *key distinction* as the latter matches the $L\|u\|^2$ term in the optimal gradient-variation regret (Chiang et al., 2012), and in the context of our gradient-variation bounds from the previous section, a $\|u\|^2$ overhead is negligible because we already expect to incur an $O(\|u\|^2)$ term from Theorem 1. Moreover, the $\|u\|^2$ penalty only dominates the worst-case bound $O(\|u\|^2 + \|u\|\sqrt{T})$ when that bound is vacuous (i.e., $\|u\| = \Omega(\sqrt{T})$). By contrast, a $\|u\|^3$ penalty can dominate even in non-vacuous regimes, so it cannot be treated as a lower-order term.

Virtual clipping via quadratic regularization. To illustrate the crux of the virtual clipping argument, consider a standard FTRL update, $w_{t+1} = \arg\min_{w\in\mathbb{R}^d} \left\langle \sum_{s=1}^{t} g_s, w\right\rangle + \Psi_{t+1}(w)$ where

Algorithm 1: Simple Fully-adaptive Optimistic Algorithm

Input: $\epsilon > 0, \gamma > 0$

Initialize: $w_1 = \mathbf{0}$, $h_1 = \mathbf{0}$, $\widehat{M}_1 = \gamma$, $\beta = \frac{\gamma}{\epsilon}$

Define: Regularizer $\psi^{\mathrm{PF}}(w; \alpha, V, M) = 3\int_0^{\|w\|} \min_{\eta \le 1/M} \left[\frac{\log(x/\alpha+1)}{\eta} + \eta V\right] \mathrm{d}x$

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

Play w_t, receive gradient $g_t = \nabla f_t(w_t)$ and optimistic hint h_{t+1}

Define:

$$\Delta_t = g_t - h_t, \qquad \widehat{\Delta}_t = \Delta_t \min\left\{1, \frac{\widehat{M}_t}{\|\Delta_t\|}\right\}, \qquad \widehat{M}_{t+1} = \max\left\{\widehat{M}_t, \|\Delta_t\|\right\}$$

$$B_{t+1} = 4 + \sum_{i=1}^t \frac{\|\widehat{\Delta}_i\|^2}{\widehat{M}_i^2}, \qquad \bar{V}_{t+1} = 4\widehat{M}_{t+1}^2 + \sum_{i=1}^t \|\widehat{\Delta}_i\|^2, \qquad \alpha_{t+1} = \frac{\epsilon}{\sqrt{B_{t+1}}\log^2(B_{t+1})}$$

Choose regularizer $\Psi_{t+1}(w) = \psi^{\mathrm{PF}}\left(w; \alpha_{t+1}, \bar{V}_{t+1}, \widehat{M}_{t+1}\right) + \frac{\beta}{2}\|w\|^2$

Update $w_{t+1} = \arg\min_{w \in \mathbb{R}^d} \left\langle h_{t+1} + \sum_{s=1}^t g_s, w\right\rangle + \Psi_{t+1}(w)$

end

Ψ_{t+1} is an arbitrary convex regularizer. Then using the standard regret analysis (Orabona, 2019, Theorem 7.1), it can be shown that if $(\Psi_t)_t$ is a non-decreasing sequence of regularizers, we have

$$\sum_{t=1}^T \langle g_t, w_t - u\rangle \le \Psi_{T+1}(u) + \sum_{t=1}^T \Big(\langle g_t, w_t - w_{t+1}\rangle - \mathcal{D}_{\Psi_t}(w_{t+1}, w_t)\Big) =: \Psi_{T+1}(u) + \sum_{t=1}^T \delta_t. \quad (9)$$

Inside the stability term $\sum_{t=1}^T \delta_t$, if we replace g_t with their clipped quantities $\widetilde{g}_t$, then

$$\sum_{t=1}^T \delta_t \le \frac{G^2}{2\beta} + \sum_{t=1}^T \Big(\langle \widetilde{g}_t, w_t - w_{t+1}\rangle + \frac{\beta}{2}\|w_t - w_{t+1}\|^2 - \mathcal{D}_{\Psi_t}(w_{t+1}, w_t)\Big), \qquad (10)$$

where we have used Fenchel-Young inequality to bound $\|g_t - \widetilde{g}_t\|\,\|w_t - w_{t+1}\| \le \frac{1}{2\beta}\|g_t - \widetilde{g}_t\|^2 + \frac{\beta}{2}\|w_t - w_{t+1}\|^2$ and used an argument similar to Cutkosky (2019a) to bound $\sum_{t=1}^T \|g_t - \widetilde{g}_t\|^2 = O(G^2)$. Now the differences $\frac{\beta}{2}\|w_t - w_{t+1}\|^2$ can be cancelled completely by simply including a matching quadratic penalty in Ψ_t to form a hybrid regularizer: $\Psi_t(w) = \psi_t^{\mathrm{PF}}(w) + \frac{\beta}{2}\|w\|^2$.

The above illustrates why we refer to our approach as *virtual* clipping. Unlike Cutkosky (2019b), our approach passes the algorithm the true subgradients g_t rather than the clipped $\widetilde{g}_t$, yet by including a quadratic penalty in Ψ_t we are able to replace the g_t appearing in the stability term with the clipped quantity *in the analysis*, allowing us to control this term using standard comparator-adaptive regularizers, such as the one discussed in the previous section, defined in terms of the clipped gradients $\widetilde{g}_t$.

Notably, the discussion above easily extends to optimistic updates by replacing g_t with $\Delta_t \triangleq g_t - h_t$ and $\widetilde{g}_t$ with $\widehat{\Delta}_t \triangleq \Delta_t \min\{1, \max_{s<t}\|\Delta_s\|/\|\Delta_t\|\}$. The resulting algorithm is shown in Algorithm 1, and its regret guarantee is presented in Theorem 3, with proof in Appendix B.1.

Theorem 3 *For any $u \in \mathbb{R}^d$, Algorithm 1 with $h_t = g_{t-1}$ guarantees*

$$\mathrm{REG}_T(u) \le \widetilde{O}\left(\|u\|\sqrt{V_T(u)\left(\log_+ \frac{\|u\|\sqrt{T}}{\epsilon}\right)} + \left(L + \frac{\gamma}{\epsilon}\right)\|u\|^2 + \gamma\|u\| + \epsilon G\left(\frac{G}{\gamma} + 1\right) + \epsilon\gamma\right).$$

Ideally, we would set ϵ in terms of $\|u\|$ and γ in terms of G, so that these low-order terms would correspond to the appropriate units of $G\|u\|$. Nonetheless, observe that we may set these parameters naively (e.g, $\epsilon = \gamma = 1$) without significantly impacting the main terms in the bound.

Theorem 3 strictly improves over the previous best-known bound for fully-adaptive gradient-variation regret $\widetilde{O}\big(\|u\|^2\sqrt{V_T^+} + L^2\|u\|^4 + G\|u\|^3 + G^2\|u\|^2 + G^2\sqrt{\sum_{t=1}^T \|g_t\|}\big)$ from Wang et al. (2025), improving both the leading terms and constant penalties. Moreover, Algorithm 1 is relatively *simple* compared to alternative approaches and analyses (Cutkosky and Mhammedi, 2024) because it lets us immediately apply existing parameter-free regularization strategies—all we had to do was add an additional quadratic regularizer.

Computational efficiency. Although Algorithm 1 is conceptually simple, computing the update requires careful implementation to ensure computational efficiency. In each round the direction of w_{t+1} is straightforward to obtain, but its magnitude must be determined by solving an equation that, in general, admits no closed-form solution and requires a line search. Fortunately, our algorithm guarantees that $\|w_{t+1}\|$ remains close to $\|w_t\|$, so one can efficiently search for $\|w_{t+1}\|$ within a small interval around $\|w_t\|$ to achieve a sufficiently small approximation error. The following proposition states useful constraints for computing w_{t+1}, with full details deferred to Appendix B.2.

Proposition 1 *Algorithm 1 guarantees that, for any $t \in [T]$, the direction of w_{t+1} is determined by*

$$\frac{w_{t+1}}{\|w_{t+1}\|} = -\frac{h_{t+1} + \sum_{s=1}^t g_s}{\left\|h_{t+1} + \sum_{s=1}^t g_s\right\|}.$$

Moreover, denoting $\alpha := \alpha_{t+1}$, $V := \bar{V}_{t+1}$, and $M := \widehat{M}_{t+1}$ the magnitude of w_{t+1} satisfies

$$\big|\,\|w_{t+1}\| - \|w_t\|\,\big| \leq \frac{\|g_t - h_t + h_{t+1}\|}{\beta}, \quad \textit{and}$$

$$\left\|h_{t+1} + \sum_{s=1}^t g_s\right\| = \begin{cases} 6\sqrt{V\ln(\|w_{t+1}\|/\alpha + 1)} + \beta\|w_{t+1}\|, & \textit{if } \sqrt{\frac{\ln(\|w_{t+1}\|/\alpha+1)}{V}} \leq \frac{1}{M}, \\ 3\left(M\ln(\|w_{t+1}\|/\alpha + 1) + \frac{V}{M}\right) + \beta\|w_{t+1}\|, & \textit{otherwise.} \end{cases}$$

Proposition 1 implies that, at each round, the direction of w_{t+1} is determined exactly, and the norm of w_{t+1} lies in an interval of length at most $O(G/\beta)$. This interval can be efficiently searched via binary search to obtain an ϵ-approximate solution. For example, to achieve up to $\epsilon = O(1/T)$ accuracy, it suffices to use $O(\log T)$ iterations to check the above 1-D equation per round, leading to an overall $O(d + \log T)$ per-round time complexity, which actually improves upon the $O(d\log^2 T)$ complexity of the ensemble method of Wang et al. (2025), while also avoiding their assumption of a *known* Lipschitz constant. Hence, our approach improves the state-of-the-art both in terms of regret and computational overhead. Given the conceptual simplicity of the algorithm and its analysis, we believe this approach will be of independent interest for Lipschitz adaptivity in online learning.

3.2. An Efficient Algorithm via Optimistic Reduction

In this part, we present an efficient fully-adaptive optimistic algorithm by extending the algorithm of Cutkosky and Mhammedi (2024) using a refined optimistic reduction.

Algorithm 2: Efficient Fully-adaptive Optimistic Algorithm

Input: $\epsilon > 0, \gamma > 0$

Initialize: Instantiate two instances of the algorithm of Cutkosky and Mhammedi (2024), $\mathcal{A}_x(\epsilon, \gamma)$ applied on $\mathbb{R}^d$, and $\mathcal{A}_y(\epsilon/\gamma, \gamma^2)$ applied on $\mathbb{R}$.

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

  Receive optimistic hint h_t

  Get $x_t \in \mathbb{R}^d$ from $\mathcal{A}_x$ and $y_t \in \mathbb{R}$ from $\mathcal{A}_y$

  Play $w_t = x_t - y_t h_t$ and observe $g_t = \nabla f_t(w_t)$

  Pass g_t to $\mathcal{A}_x$ as the t^{th} gradient

  Pass $-\langle g_t, h_t\rangle$ to $\mathcal{A}_y$ as the t^{th} gradient

end

Challenge in extending Cutkosky and Mhammedi (2024) to optimistic updates. Cutkosky and Mhammedi (2024) proposed an efficient algorithm that is free of both G and $\|u\|$, ensuring an $\widetilde{O}\big(\|u\| \sqrt{\sum_t \|g_t\|^2 + \|u\|^2 + G^2}\big)$ regret. Unfortunately, their approach requires a very delicate analysis involving low-level interactions between the gradient clipping reduction (Cutkosky, 2019b) and an internal application of a constraint-set reduction due to Cutkosky and Orabona (2018), making it difficult to extend beyond the original scope.

In fact, the extension to optimistic updates is particularly problematic, as the constraint-set reduction adds an additional term to the feedback that can ruin the optimistic guarantee, as detailed by Cutkosky (2019b) and subsequently observed in several follow-up related works (Cutkosky, 2019b; Bhaskara et al., 2020, 2021; Zhao et al., 2025a). Indeed, the constraint-set reduction of Cutkosky and Orabona (2018) modifies the learner's feedback at time t to $\widetilde{\ell}_t \triangleq g_t + \|g_t\| \nabla S(w_t)$, where $\nabla S(w_t) \in \partial \|w_t - \Pi_{\mathcal{W}}(w_t)\|$ and satisfies $\|\nabla S(w_t)\| \le 1$.[4] In the context of an optimistic update, this additional term can dominate the desired optimistic dependency, since $\sum_{t=1}^T \|\widetilde{\ell}_t - h_t\|^2 = O(\sum_{t=1}^T \|g_t - h_t\|^2 + \|g_t\|^2)$. This *incompatibility between optimism and the constraint-set reduction* in turn makes it highly non-trivial to extend the approach of Cutkosky and Mhammedi (2024) to an optimistic guarantee in any *white-box manner*.

Refined optimistic reduction. To address this issue, we avoid this incompatibility by using a refined version of the optimistic reduction from Cutkosky (2019c), which directly converts the regret to an optimistic form in a *black-box manner*. In this way, the aforementioned constraints are applied internally within the base algorithms, rather than externally on top of them, allowing us to retain the efficiency benefits of Cutkosky and Mhammedi (2024) without ruining the optimistic guarantees. However, this reduction is typically applied under the assumption that $\|h_t\| \le 1$ for all t, which is not suitable for our Lipschitz adaptive setting. The following proposition, proven in Appendix B.4, provides a refinement of the reduction which accounts for a time-varying Lipschitz constant.

Proposition 2 (Refined Optimistic Reduction) *Let $\mathcal{A}_x$ and $\mathcal{A}_y$ be online algorithms defined on $\mathcal{W}_x = \mathbb{R}^d$ and $\mathcal{W}_y = \mathbb{R}$ respectively. Suppose the following conditions hold:*

- *For each $z \in \{x, y\}$, $\mathcal{A}_z$ guarantees $\text{REG}_T^{\mathcal{A}_z}(u) \le A_T^{\mathcal{A}_z}(u) + B_T^{\mathcal{A}_z}(u)\sqrt{\sum_{t=1}^T \|g_t\|^2}$ for any $u \in \mathcal{W}_z$ and $\{g_t\}_{t=1}^T$ in $\mathcal{W}_z$, where $A_T^{\mathcal{A}_z}$ and $B_T^{\mathcal{A}_z}$ are non-negative functions.*

4. Refinements of the constraint-set reduction exist which slightly improve constant factors and take a slightly different surrogate penalties, but these still suffer the same incompatibility with optimism described above (Cutkosky, 2020).

- $A_T^{\mathcal{A}_y}$ *is non-decreasing and* $B_T^{\mathcal{A}_y}(\mathring{y}) \le |\mathring{y}|\, \lambda_T(\mathring{y})$ *for some non-decreasing function* $\lambda_T(\mathring{y}) \ge 1$.

Then for any $u \in \mathbb{R}^d$*, Algorithm 5 (in Appendix B.4) enjoys the following optimistic bound:*

$$\sum_{t=1}^{T} \langle g_t, w_t - u \rangle \le A_T^{\mathcal{A}_x}(u) + A_T^{\mathcal{A}_y}(\mathring{y}) + 2B_T^{\mathcal{A}_x}(u) \sqrt{H_T^2 \lambda_T(\mathring{y})^2 + \frac{1}{2}\left[\sum_{t=1}^{T} \|g_t - h_t\|^2 - \|h_t\|^2\right]_+},$$

where $H_T = \max_{t\in[T]} \|h_t\|$, $\mathring{y} = B_T^{\mathcal{A}_x}(u)/H_T$, *and* $[x]_+ \triangleq \max\{x, 0\}$.

Applying this refined optimistic reduction with $\mathcal{A}_x$ and $\mathcal{A}_y$ being two instances of the algorithm of Cutkosky and Mhammedi (2024) (with regret guarantee restated in Lemma 11) leads to the full algorithm summarized in Algorithm 2. We then have the following fully-adaptive regret guarantee with proof in Appendix B.3.

Theorem 4 *For any* $u \in \mathbb{R}^d$*, Algorithm 2 with hints* $h_t = g_{t-1}$ *guarantees* $\textsc{Reg}_T(u)$ *bounded by*

$$\widetilde{O}\left(\|u\| \sqrt{V_T(u) \left(\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}\right)} + \left(L + \tfrac{\gamma}{\epsilon} + \tfrac{\gamma^3}{\epsilon G^2}\right) \|u\|^2 + \epsilon G\left(\tfrac{G^3}{\gamma^3} + \tfrac{G}{\gamma}\right) + \epsilon\gamma\right).$$

Moreover, the algorithm admits an efficient closed-form update with $O(d)$ *time per iteration.*

Algorithm 2 is the *first* efficient and fully-adaptive optimistic algorithm for unconstrained settings. Compared to Theorem 3, it maintains the same favorable leading term related to gradient variation. However, the lower-order term slightly deteriorates from G^2 to G^4, because the algorithm $\mathcal{A}_y$ receives feedback $\langle g_t, h_t \rangle$, and hence has an effective Lipschitz constant of G^2 when setting $h_t = g_{t-1}$. This then leads to the G^4 penalty when applying the fully-adaptive guarantee in Cutkosky and Mhammedi (2024). We suspect that this is an artifact of the analysis, though it is currently unclear how to further refine the black-box reduction to avoid this penalty while still allowing the proper cancellations in the analysis. We leave this as an important direction for future work.

4. Applications

In this section, we extend our parameter-free gradient variation results to two important applications: dynamic regret minimization and the stochastically-extended adversarial (SEA) setting.

4.1. Dynamic Regret

Previous works on gradient-variation dynamic regret minimization have primarily focused on constrained domains (Zhao et al., 2020, 2024). In this section we provide comparator-adaptive and fully-adaptive algorithms achieving dynamic gradient variation bounds in *unconstrained* settings.

Comparator-adaptive dynamic regret. When the Lipschitz constant is known, we can simply apply an optimistic extension of the comparator-adaptive dynamic regret algorithm of Jacobsen and Cutkosky (2022) (e.g., adding an optimistic step, as in their Algorithm 3, to the base learner in their Algorithm 2), followed by Theorem 1 to get comparator-adaptive gradient-variation dynamic regret.

Theorem 5 *For any sequence $u_1, \dots, u_T$ in $\mathbb{R}^d$, the optimistic extension of Jacobsen and Cutkosky (2022, Algorithm 2) with $h_t = g_{t-1}$ guarantees*

$$\textsc{Reg}_T(u_{1:T}) \leq \widetilde{O}\left(\sqrt{(M^2 + MP_T)V_T(u_{1:T})\left(\log_+ \tfrac{MT}{\epsilon}\right)} + LMP_T + LM^2 + \epsilon G\right),$$

where $M \triangleq \max_{t\in[T]} \|u_t\|$.

While this result is a simple extension of existing results when applied with our reduction in Theorem 1, we note that this is the first instance of an explicit gradient-variation bound for dynamic regret in unconstrained domains, and hence fills an important gap in the literature.

Fully-adaptive dynamic regret. With an unknown Lipschitz constant, directly extending our previous fully-adaptive results to dynamic regret is challenging. For our first approach in Section 3.1, including a quadratic regularizer in unconstrained domains significantly complicates dynamic regret guarantees (Jacobsen and Cutkosky, 2023), and our black-box approach in Section 3.2 also does not naturally extend to dynamic regret, since Cutkosky and Mhammedi (2024) provides only *static* regret bounds based on FTRL, and extending their analysis to dynamic regret is highly non-trivial.

Instead, we can utilize a black-box reduction from Cutkosky and Orabona (2018), which decomposes the regret into an unconstrained *static regret* problem — wherein we can apply our results from the previous section — and a dynamic regret problem *on the unit ball*, where existing algorithms such as SWORD (Zhao et al., 2020) and SWORD++ (Zhao et al., 2024) can be deployed. Specifically, the decision w_t is decomposed as $w_t = y_t x_t$, with one online algorithm producing the magnitude $y_t \in \mathbb{R}$ and another producing the direction x_t. Applied to dynamic regret, this reduction guarantees the following decomposition (Jacobsen and Cutkosky, 2022, Appendix J):

$$\textsc{Reg}_T(u_{1:T}) \leq \textsc{Reg}_T^{\text{1d}}(M) + M\textsc{Reg}_T^{\mathcal{B}}(u_{1:T}/M),$$

where $\textsc{Reg}_T^{\text{1d}}(M)$ is the 1D static regret with comparator $M \triangleq \max_{t\in[T]} \|u_t\|$, and $\textsc{Reg}_T^{\mathcal{B}}(u_{1:T}/M)$ is the dynamic regret of the direction learner on the unit ball w.r.t the re-scaled comparator sequence.

To obtain gradient-variation guarantees, we further incorporate the optimistic scheme into this reduction and summarize our algorithm in Algorithm 3. For the 1D algorithm $\mathcal{A}_{\text{1d}}$, we employ our fully-adaptive optimistic Algorithm 2. For the bounded domain algorithm $\mathcal{A}_{\mathcal{B}}$, we use a standard online ensemble method similar to Zhao et al. (2024), with an enhancement that applies time-varying step sizes for base learners to achieve Lipschitz adaptivity, and the doubling trick to guide the instantiation of new base learners. This online ensemble method is summarized in Algorithm 6 in Appendix C.2, with the Lipschitz-adaptive and anytime theoretical guarantee in Theorem 12.

Finally, we provide our fully-adaptive and anytime dynamic regret guarantee in Theorem 6, with a more detailed version and the proof in Appendix C.1.

Theorem 6 *For any sequence $u_1, \dots, u_T \in \mathbb{R}^d$, Algorithm 3 guarantees $\textsc{Reg}_T(u_{1:T})$ bounded by*

$$\widetilde{O}\Bigg(\sqrt{(M^2 + MP_T)\min\left\{V_T(u_{1:T}), LF_T(u_{1:T})\right\}\left(\log_+ \tfrac{MG\sqrt{T}}{\epsilon\gamma}\right)} \\ + (LM + G + \gamma)P_T + \left(L + \tfrac{\gamma}{\epsilon} + \tfrac{\gamma^3}{\epsilon G^2}\right)M^2 + \gamma M + \epsilon G\left(\tfrac{G^3}{\gamma^3} + \tfrac{G}{\gamma} + 1\right) + \epsilon\gamma\Bigg),$$

where $M \triangleq \max_{t\in[T]} \|u_t\|$ and $F_T(u_{1:T}) \triangleq \sum_{t=1}^T (f_t(u_t) - \inf_{w\in\mathbb{R}^d} f_t(w))$. Moreover, the algorithm runs in $O(d \log t)$ time on iteration t.

Algorithm 3: Fully-adaptive Gradient-Variation Dynamic Regret Minimization

Input: $\epsilon > 0, \gamma > 0$

Initialize: $g_0 = \mathbf{0}$. Instantiate Algorithm 2 as $\mathcal{A}_{1\text{d}}(\epsilon, \gamma)$ acting on $\mathbb{R}$, and instantiate Algorithm 6 as $\mathcal{A}_{\mathcal{B}}(D := 1, \gamma)$ acting on the unit ball $\mathcal{B} = \{w \in \mathbb{R}^d : \|w\| \leq 1\}$

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

  Set optimistic hint $h_t = g_{t-1}$

  Pass h_t to $\mathcal{A}_{\mathcal{B}}$ as the t^{th} hint, and get x_t from $\mathcal{A}_{\mathcal{B}}$

  Pass $\langle h_t, x_t \rangle$ to $\mathcal{A}_{1\text{d}}$ as the t^{th} hint, and get y_t from $\mathcal{A}_{1\text{d}}$

  Play $w_t = y_t x_t$ and observe $g_t = \nabla f_t(w_t)$

  Pass g_t to $\mathcal{A}_{\mathcal{B}}$ as the t^{th} gradient, and pass $\langle g_t, x_t \rangle$ to $\mathcal{A}_{1\text{d}}$ as the t^{th} gradient

end

Theorem 6 provides the first fully-adaptive gradient-variation dynamic regret in the unconstrained setting. It matches the best-known bound in the constrained setting (Zhao et al., 2024) up to logarithmic factors. By the optimistic-to-gradient-variation reduction, this theorem simultaneously achieves a small-loss bound that can also be generally applied to static regret. Due to gradient-variation adaptivity, this result also readily applies to the stochastically extended adversarial model (Sachs et al., 2022), yielding the *first* unconstrained dynamic regret in this setting, formally discussed below.

4.2. Stochastically-extended Adversarial Model

The stochastically-extended adversarial (SEA) model (Sachs et al., 2022) is an intermediate setting between adversarial OCO and Stochastic Convex Optimization (SCO). The key difference from standard OCO is that the losses f_t are sampled from a distribution $\mathfrak{D}_t$ chosen by the environment on each round. This setting seamlessly interpolates between (potentially non-stationary) SCO when $\mathfrak{D}_t$ is chosen obliviously to the learners decisions—naturally modelling the non-stationary found in many real-world applications such as the Online Label Shift problem (Bai et al., 2022; Qian et al., 2023; Baby et al., 2023)—and the adversarial setting when $\mathfrak{D}_t = \delta_{\bar{f}_t}$ for arbitrary $\bar{f}_t$.

For the SEA model, the natural performance measure is the expected dynamic regret, i.e.,

$$\mathbb{E}[\text{REG}_T(u_{1:T})] \triangleq \mathbb{E}\left[\sum_{t=1}^{T} f_t(w_t) - \sum_{t=1}^{T} f_t(u_t)\right],$$

against a sequence of comparators $u_{1:T}$ that capture the non-stationarity of the environment, where the expectation is taken over the randomness of the sampled functions f_t. We assume that the comparator sequence is oblivious, as elaborated later in Remark 1.

Letting $F_t(w) := \mathbb{E}_{f_t \sim \mathfrak{D}_t}[f_t(w)]$, Sachs et al. (2022) introduced two quantities characterizing the difficulty of the environment: the stochastic variance $\sigma_t^2 \triangleq \sup_{w \in \mathcal{W}} \mathbb{E}_{f_t \sim \mathfrak{D}_t}[\|\nabla f_t(w) - \nabla F_t(w)\|^2]$, and the adversarial variation $\Sigma_t^2 \triangleq \sup_{w \in \mathcal{W}} \|\nabla F_t(w) - \nabla F_{t-1}(w)\|^2$. We consider the following comparator-adaptive generalizations based on an arbitrary $u \in \mathbb{R}^d$:

$$\sigma_t^2(u) \triangleq \mathbb{E}_{f_t \sim \mathfrak{D}_t}\left[\|\nabla f_t(u) - \nabla F_t(u)\|^2 \,\middle|\, \mathcal{F}_{t-1}\right], \text{ and } \Sigma_t^2(u) \triangleq \|\nabla F_t(u) - \nabla F_{t-1}(u)\|^2, \quad (11)$$

where we denote $\mathcal{F}_{t-1}$ the sigma field generated up to the beginning of round t. The following theorem then provides fully-adaptive dynamic regret guarantees for the SEA model, with proof provided in Appendix C.3.

Theorem 7 *In the SEA model, for any oblivious comparator sequence* $u_1, \ldots, u_T \in \mathbb{R}^d$ *such that* $\max_t \|u_t\| \le M$ *with* $M > 0$*, Algorithm 3 guarantees* $\mathbb{E}\left[\mathrm{REG}_T(u_{1:T})\right]$ *bounded by*

$$\widetilde{O}\Bigg(\sqrt{(M^2 + MP_T)\left(\sigma^2_{1:t}(u_{1:T}) + \Sigma^2_{1:T}(u_{1:T})\right)\left(\log_+ \tfrac{MG\sqrt{T}}{\epsilon\gamma}\right)} \\ + (LM + G + \gamma)P_T + \left(L + \tfrac{\gamma}{\epsilon} + \tfrac{\gamma^3}{\epsilon G^2}\right)M^2 + \gamma M + \epsilon G\left(\tfrac{G^3}{\gamma^3} + \tfrac{G}{\gamma} + 1\right) + \epsilon\gamma\Bigg).$$

where $\sigma^2_{1:T}(u_{1:T}) \triangleq \sum_{t=2}^T (\sigma_t^2(u_{t-1}) + \sigma^2_{t-1}(u_{t-1}))$*, and* $\Sigma^2_{1:T}(u_{1:T}) \triangleq \sum_{t=2}^T \Sigma_t^2(u_{t-1})$*. Moreover, the algorithm runs in* $O(d \log t)$ *time on iteration* t*.*

Remark 1 *We assume that the comparator sequence* $u_{1:T}$ *is oblivious, ensuring* $\mathbb{E}[P_T \cdot V_T(u_{1:T})] = P_T \cdot \mathbb{E}[V_T(u_{1:T})]$*. This assumption is reasonable in many real-world applications such as the online label shift problem (Bai et al., 2022; Qian et al., 2023; Chen et al., 2024)—a classification task in non-stationary environments where the label distribution changes over time (e.g., species monitoring, where the learner does not affect the underlying environment dynamics). In such settings, the learner competes against optimal parameters* $u_t = \arg\min_w F_t(w)$*, which (we can assume) is determined by the underlying environment and is independent of the learner's predictions.*

The closely-related work of Wang et al. (2025, Theorem 4.5) obtained a fully-adaptive static regret of $\widetilde{O}\big(\|u\|^2 \sqrt{\sigma^2_{1:T}(w_{1:T}) + \Sigma^2_{1:T}(w_{1:T})} + L^2\|u\|^4 + G\|u\|^3 + G^2\|u\|^2 + G^2\sqrt{\sum_{t=1}^T \|g_t\|}\big)$ for the SEA model, suffering from a quadratic dependency on the comparator norm in the leading term and a large penalty of $\|u\|^4$. In contrast, our Theorem 7 implies a static regret bound $\widetilde{O}(\|u\|\sqrt{\sigma^2_{1:T}(u) + \Sigma^2_{1:T}(u)} + L\|u\|^2 + G^4)$, which has significantly better $\|u\|$-dependencies.

It is worth noting that while our Theorem 7 introduces a lower-order G^4 penalty that does not appear in Wang et al. (2025), this penalty is not present in our algorithm from Section 3.1, which when applied in the SEA setting obtains a strict improvement over their result while being slightly more efficient. Moreover, our result is *problem-dependent*, with $\sigma^2_{1:T}(u) + \Sigma^2_{1:T}(u)$, whereas Wang et al. (2025) obtained *algorithm-dependent* $\sigma^2_{1:T}(w_{1:T}) + \Sigma^2_{1:T}(w_{1:T})$, with the algorithm's trajectory $w_{1:T}$. For a problem-dependent guarantee, one has to upper bound the algorithm-dependent quantity by the worst-case $\sum_{t=1}^T(\sigma_t^2 + \Sigma_t^2)$, which can easily become $O(T)$ in an unbounded domain.

5. Conclusion

In this work, we investigate parameter-free gradient-variation online learning. We provided a simple black-box reduction that converts general optimistic regret bounds into gradient-variation bounds, and leverage this reduction to develop two fully-adaptive optimistic algorithms achieving gradient-variation regret bounds in unconstrained settings. As a direct application of our approach, we obtain novel gradient-variation guarantees for dynamic regret and for the SEA model, where our results significantly improve the state-of-the-art.

We anticipate several promising directions for future work. We expect that our fully-adaptive gradient-variation guarantees will find natural applications in areas such as minimax games and accelerated (universal) optimization. We also believe developing a more refined optimistic reduction to achieve an efficient fully-adaptive optimistic regret bound without the G^4 penalty artifact would also be valuable in general. Finally, it would be interesting to explore more sophisticated solvers to compute the simple fully-adaptive algorithm from Section 3.1.

Acknowledgments

Peng Zhao and Yuheng Zhao were supported by NSFC (62361146852, 62576164), the Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China (No. JYB2025XDXM118). Nicolò Cesa-Bianchi and Andrew Jacobsen acknowledge the financial support from the EU Horizon CL4-2022-HUMAN-02 research and innovation action under grant agreement 101120237, project ELIAS (European Lighthouse of AI for Sustainability). This research was also supported by the "111 Center" (No. B26023).

References

Dheeraj Baby, Saurabh Garg, Tzu-Ching Yen, Sivaraman Balakrishnan, Zachary Chase Lipton, and Yu-Xiang Wang. Online label shift: Optimal dynamic regret meets practical algorithms. In *Advances in Neural Information Processing Systems 36 (NeurIPS)*, pages 65703–65742, 2023.

Yong Bai, Yu-Jie Zhang, Peng Zhao, Masashi Sugiyama, and Zhi-Hua Zhou. Adapting to online label shift with provable guarantees. In *Advances in Neural Information Processing Systems 35 (NeurIPS)*, pages 29960–29974, 2022.

Aditya Bhaskara, Ashok Cutkosky, Ravi Kumar, and Manish Purohit. Online learning with imperfect hints. In *Proceedings of the 37th International Conference on Machine Learning (ICML)*, pages 822–831, 2020.

Aditya Bhaskara, Ashok Cutkosky, Ravi Kumar, and Manish Purohit. Logarithmic regret from sublinear hints. In *Advances in Neural Information Processing Systems 34 (NeurIPS)*, pages 10148–10201, 2021.

Nicolò Campolongo and Francesco Orabona. Temporal variability in implicit online learning. In *Advances in Neural Information Processing Systems 33 (NeurIPS)*, pages 12377–12387, 2020.

Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.

Kamalika Chaudhuri, Yoav Freund, and Daniel J Hsu. A parameter-free hedging algorithm. In *Advances in Neural Information Processing Systems 22 (NIPS)*, pages 297–305, 2009.

Sijia Chen, Yu-Jie Zhang, Wei-Wei Tu, Peng Zhao, and Lijun Zhang. Optimistic online mirror descent for bridging stochastic and adversarial online convex optimization. *Journal of Machine Learning Research*, 25(178):1–62, 2024.

Chao-Kai Chiang, Tianbao Yang, Chia-Jung Lee, Mehrdad Mahdavi, Chi-Jen Lu, Rong Jin, and Shenghuo Zhu. Online optimization with gradual variations. In *Proceedings of the 25th Conference on Learning Theory (COLT)*, pages 6.1–6.20, 2012.

Ashok Cutkosky. Anytime online-to-batch, optimism and acceleration. In *Proceedings of the 36th International Conference on Machine Learning (ICML)*, pages 1446–1454, 2019a.

Ashok Cutkosky. Artificial constraints and hints for unbounded online learning. In *Proceedings of the 32nd Conference on Learning Theory (COLT)*, pages 874–894, 2019b.

Ashok Cutkosky. Combining online learning guarantees. In *Proceedings of the 32nd Conference on Learning Theory (COLT)*, pages 895–913, 2019c.

Ashok Cutkosky. Parameter-free, dynamic, and strongly-adaptive online learning. In *Proceedings of the 37th International Conference on Machine Learning (ICML)*, pages 2250–2259, 2020.

Ashok Cutkosky and Zakaria Mhammedi. Fully unconstrained online learning. In *Advances in Neural Information Processing Systems 37 (NeurIPS)*, pages 10148–10201, 2024.

Ashok Cutkosky and Francesco Orabona. Black-box reductions for parameter-free online learning in Banach spaces. In *Proceedings of the 31st Conference on Learning Theory (COLT)*, pages 1493–1529, 2018.

Steven de Rooij, Tim van Erven, Peter D. Grünwald, and Wouter M. Koolen. Follow the leader if you can, hedge if you must. *Journal of Machine Learning Research*, 15(1):1281–1316, 2014.

Dylan J. Foster, Alexander Rakhlin, and Karthik Sridharan. Adaptive online learning. In *Advances in Neural Information Processing Systems 28 (NIPS)*, pages 3375–3383, 2015.

Elad Hazan. *Introduction to Online Convex Optimization*. MIT Press, 2nd edition, 2022.

Andrew Jacobsen and Ashok Cutkosky. Parameter-free mirror descent. In *Proceedings of the 35th Conference on Learning Theory (COLT)*, pages 4160–4211, 2022.

Andrew Jacobsen and Ashok Cutkosky. Unconstrained online learning with unbounded losses. In *Proceedings of the 40th International Conference on Machine Learning (ICML)*, pages 14590–14630, 2023.

Pooria Joulani, András György, and Csaba Szepesvári. A modular analysis of adaptive (non-)convex optimization: Optimism, composite objectives, variance reduction, and variational bounds. *Theoretical Computer Science*, 808:108–138, 2020.

Kfir Levy. Online to offline conversions, universality and adaptive minibatch sizes. In *Advances in Neural Information Processing Systems 30 (NIPS)*, pages 1613–1622, 2017.

Brendan Mcmahan and Matthew Streeter. No-regret algorithms for unconstrained online convex optimization. In *Advances in Neural Information Processing Systems 25 (NIPS)*, pages 2402–2410, 2012.

H. Brendan McMahan and Francesco Orabona. Unconstrained online linear learning in Hilbert spaces: Minimax algorithms and normal approximations. In *Proceedings of the 27th Conference on Learning Theory (COLT)*, pages 1020–1039, 2014.

Naram Mhaisen and George Iosifidis. On the dynamic regret of following the regularized leader: Optimism with history pruning. In *Proceedings of the 42nd International Conference on Machine Learning (ICML)*, pages 43990–44016, 2025.

Zakaria Mhammedi and Wouter M. Koolen. Lipschitz and comparator-norm adaptivity in online learning. In *Proceedings of 33rd Conference on Learning Theory (COLT)*, pages 2858–2887, 2020.

Yurii Nesterov. *Lectures on Convex Optimization*, volume 137. Springer, 2018.

Francesco Orabona. Dimension-free exponentiated gradient. In *Advances in Neural Information Processing Systems 26 (NIPS)*, pages 1806–1814, 2013.

Francesco Orabona. A Modern Introduction to Online Learning. *ArXiv preprint*, arxiv:1912.13213, 2019.

Francesco Orabona and Dávid Pál. Coin betting and parameter-free online learning. In *Advances in Neural Information Processing Systems 29 (NIPS)*, pages 577–585, 2016.

Francesco Orabona and Dávid Pál. Parameter-free stochastic optimization of variationally coherent functions. *ArXiv preprint*, arxiv:2102.00236, 2021.

Yu-Yang Qian, Yong Bai, Zhen-Yu Zhang, Peng Zhao, and Zhi-Hua Zhou. Handling new class in online label shift. In *Proceedings of the 23rd IEEE International Conference on Data Mining (ICDM)*, pages 1283–1288, 2023.

Tim Roughgarden. *Beyond the Worst-Case Analysis of Algorithms*. Cambridge University Press, 2021.

Sarah Sachs, Hedi Hadiji, Tim van Erven, and Cristóbal A Guzmán. Between stochastic and adversarial online convex optimization: Improved regret bounds via smoothness. In *Advances in Neural Information Processing Systems 35 (NeurIPS)*, pages 691–702, 2022.

Shai Shalev-Shwartz. Online Learning and Online Convex Optimization. *Foundations and Trends in Machine Learning*, 4(2):107–194, 2012.

Vasilis Syrgkanis, Alekh Agarwal, Haipeng Luo, and Robert E. Schapire. Fast convergence of regularized learning in games. In *Advances in Neural Information Processing Systems 28 (NIPS)*, pages 2989–2997, 2015.

Shuche Wang, Adarsh Barik, Peng Zhao, and Vincent Y. F. Tan. Parameter-free algorithms for the stochastically extended adversarial model. In *Advances in Neural Information Processing Systems 38 (NeurIPS)*, pages 64068–64102, 2025.

Yan-Feng Xie, Peng Zhao, and Zhi-Hua Zhou. Gradient-variation online learning under generalized smoothness. In *Advances in Neural Information Processing Systems 37 (NeurIPS)*, pages 37865–37899, 2024.

Yu-Hu Yan, Peng Zhao, and Zhi-Hua Zhou. A simple and optimal approach for universal online learning with gradient variations. In *Advances in Neural Information Processing Systems 37 (NeurIPS)*, pages 11132–11163, 2024.

Tianbao Yang, Mehrdad Mahdavi, Rong Jin, and Shenghuo Zhu. Regret bounded by gradual variation for online convex optimization. *Machine Learning*, 95(2):183–223, 2014.

Hang Yu, Yu-Hu Yan, and Peng Zhao. Improved dimension dependence for bandit convex optimization with gradient variations. In *Proceedings of the 43rd International Conference on Machine Learning (ICML)*, page to appear, 2026.

Lijun Zhang, Shiyin Lu, and Zhi-Hua Zhou. Adaptive online learning in dynamic environments. In *Advances in Neural Information Processing Systems 31 (NeurIPS)*, pages 1330–1340, 2018.

Mengxiao Zhang, Peng Zhao, Haipeng Luo, and Zhi-Hua Zhou. No-regret learning in time-varying zero-sum games. In *Proceedings of the 39th International Conference on Machine Learning (ICML)*, pages 26772–26808, 2022.

Peng Zhao, Yu-Jie Zhang, Lijun Zhang, and Zhi-Hua Zhou. Dynamic regret of convex and smooth functions. In *Advances in Neural Information Processing Systems 33 (NeurIPS)*, pages 12510–12520, 2020.

Peng Zhao, Yu-Jie Zhang, Lijun Zhang, and Zhi-Hua Zhou. Adaptivity and non-stationarity: Problem-dependent dynamic regret for online convex optimization. *Journal of Machine Learning Research*, 25(98):1–52, 2024.

Peng Zhao, Yan-Feng Xie, Lijun Zhang, and Zhi-Hua Zhou. Efficient methods for non-stationary online learning. *Journal of Machine Learning Research*, 26(208):1–66, 2025a.

Peng Zhao, Yu-Hu Yan, Hang Yu, and Zhi-Hua Zhou. Adaptivity and universality: Problem-dependent universal regret for online convex optimization. *ArXiv preprint*, arxiv:2511.19937, 2025b.

Yuheng Zhao, Yu-Hu Yan, Kfir Yehuda Levy, and Peng Zhao. Gradient-variation online adaptivity for accelerated optimization with Hölder smoothness. In *Advances in Neural Information Processing Systems 38 (NeurIPS)*, pages 3749–3779, 2025c.

Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th International Conference on Machine Learning (ICML)*, pages 928–936, 2003.

Appendix A. Proofs for Section 2

A.1. Proof of Theorem 1

In this section we provide proof of the optimistic-to-gradient-variation reduction. The full version of the theorem is re-stated below, which additionally includes the *small-loss* bounds.

Theorem 1 (Full Version) *Let $\mathcal{U}$ be a class of sequences in $\mathbb{R}^d$, and let $\mathcal{A}$ be an online learning algorithm that receives $\{g_t\}_{t=1}^T$ as gradients and takes $\{h_t\}_{t=1}^T$ as optimistic hints. Suppose $\mathcal{A}$ guarantees that for any sequence $u_{1:T} \in \mathcal{U}$ that*

$$\sum_{t=1}^{T} \langle g_t, w_t - u_t \rangle \le A_T(u_{1:T}) + B_T(u_{1:T}) \sqrt{\sum_{t=1}^{T} \|g_t - h_t\|^2}.$$

Then for any sequence of G-Lipschitz, L-smooth convex functions $\{f_t\}_{t=1}^T$, by setting $g_t = \nabla f_t(w_t)$ and $h_t = g_{t-1}$, $\mathcal{A}$ achieves

$$\begin{aligned}\textsc{Reg}_T(u_{1:T}) \le\ & A_T(u_{1:T}) + 4LB_T(u_{1:T})^2 \\ & + 2B_T(u_{1:T})\sqrt{\min\left\{G^2 + V_T(u_{1:T}) + L^2 P_T^{\|\cdot\|^2}(u_{1:T}), 4LF_T(u_{1:T})\right\}},\end{aligned}$$

where $P_T^{\|\cdot\|^2}(u_{1:T}) \triangleq \sum_{t=2}^T \|u_t - u_{t-1}\|^2$ is the squared path length, and $F_T(u_{1:T}) \triangleq \sum_{t=1}^T (f_t(u_t) - \inf_{w\in\mathbb{R}^d} f_t(w))$ is the small loss.

Proof By the definition of Bregman divergence $\mathcal{D}_f(x, y) = f(x) - f(y) - \langle \nabla f(y), x - y\rangle$, we have

$$\begin{aligned}\textsc{Reg}_T(u_{1:T}) &= \sum_{t=1}^{T} f_t(w_t) - f_t(u_t) = \sum_{t=1}^{T} \langle \nabla f_t(w_t), w_t - u_t \rangle - \mathcal{D}_{f_t}(u_t, w_t) \\ &= \textsc{Reg}_T^{\mathcal{A}}(u_{1:T}) - \sum_{t=1}^{T} \mathcal{D}_{f_t}(u_t, w_t) \\ &\le A_T(u_{1:T}) + B_T(u_{1:T})\sqrt{\sum_{t=1}^{T} \|\nabla f_t(w_t) - \nabla f_{t-1}(w_{t-1})\|^2} - \sum_{t=1}^{T} \mathcal{D}_{f_t}(u_t, w_t).\end{aligned}$$

Now apply Lemma 15 to get

$$\begin{aligned}\sum_{t=1}^{T} \|\nabla f_t(w_t) - \nabla f_{t-1}(w_{t-1})\|^2 \le\ & \min\left\{\|\nabla f_1(w_1)\|^2 + 4V_T(u_{1:T}) + 4L^2 P_T^{\|\cdot\|^2}(u_{1:T}), 16LF_T(u_{1:T})\right\} \\ & + 16L\sum_{t=1}^{T} \mathcal{D}_{f_t}(u_t, w_t).\end{aligned}$$

where we define

$$V_T(u_{1:T}) \triangleq \sum_{t=2}^{T} \|\nabla f_t(u_{t-1}) - \nabla f_{t-1}(u_{t-1})\|^2,$$

$$F_T(u_{1:T}) \triangleq \sum_{t=1}^{T}(f_t(u_t) - \inf_{w\in\mathbb{R}^d} f_t(w)), \qquad P_T^{\|\cdot\|^2}(u_{1:T}) \triangleq \sum_{t=2}^{T} \|u_t - u_{t-1}\|^2 .$$

Hence, using $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$ yields

$$\begin{aligned}
\text{REG}_T(u_{1:T}) \le A_T(u_{1:T}) &+ 2B_T(u_{1:T})\sqrt{\min\left\{G^2 + V_T(u_{1:T}) + L^2 P_T^{\|\cdot\|^2}(u_{1:T}), 4LF_T(u_{1:T})\right\}} \\
&+ 4B_T(u_{1:T})\sqrt{L\sum_{t=1}^{T}\mathcal{D}_{f_t}(u_t, w_t)} - \sum_{t=1}^{T}\mathcal{D}_{f_t}(u_t, w_t) \\
\le A_T(u_{1:T}) &+ 4LB_T(u_{1:T})^2 \\
&+ 2B_T(u_{1:T})\sqrt{\min\left\{G^2 + V_T(u_{1:T}) + L^2 P_T^{\|\cdot\|^2}(u_{1:T}), 4LF_T(u_{1:T})\right\}},
\end{aligned}$$

where the last line uses $ax - bx^2 \le a^2/4b$. ■

A.2. Parameter-free Regularizer from Jacobsen and Cutkosky (2022)

In this part, we introduce a key component of our parameter-free algorithms, which is what we call a "parameter-free regularizer" from Jacobsen and Cutkosky (2022). The following lemma provides the theoretical guarantees.

Lemma 8 *Let $g_1, \dots, g_T$ be an arbitrary sequence of vectors. Suppose $0 < M_1 \le \dots \le M_T$ is non-decreasing magnitude hint sequence that $\|g_t\| \le M_t$ for all t, and let $\alpha_1 \ge \dots \ge \alpha_T$ be a non-increasing sequence. Set $\bar{V}_t = 4M_t^2 + \sum_{i=1}^{t-1}\|g_i\|^2$ and define:*

$$\psi_t(w) = 3\int_0^{\|w\|} \min_{\eta \le 1/M_t}\left[\frac{\log(x/\alpha_t + 1)}{\eta} + \eta\bar{V}_t\right]\mathrm{d}x,$$

then for any sequence $w_1, \dots, w_{T+1} \in \mathbb{R}^d$:

$$\sum_{t=1}^{T}\langle g_t, w_t - w_{t+1}\rangle - \mathcal{D}_{\psi_t}(w_{t+1}, w_t) - (\psi_{t+1} - \psi_t)(w_{t+1}) \le \sum_{t=1}^{T}\frac{2\alpha_t\|g_t\|^2}{\sqrt{\bar{V}_t}}.$$

Moreover, for any $u \in \mathbb{R}^d$:

$$\psi_{T+1}(u) \le 6\|u\|\max\left\{\sqrt{\bar{V}_{T+1}\log(\|u\|/\alpha_{T+1} + 1)}, M_{T+1}\log(\|u\|/\alpha_{T+1} + 1)\right\}.$$

Proof This lemma follows from the proof of Jacobsen and Cutkosky (2022, Theorem 6). ■

Algorithm 4: Optimistic FTRL/Centered Mirror Descent

Input: Lipschitz bound G, value $\epsilon > 0$

Define: parameter-free regularizer $\psi(w;\bar{V},\alpha) = 3\int_0^{\|w\|} \min_{\eta\le 1/(2G)} \left[\frac{\log(x/\alpha+1)}{\eta} + \eta\bar{V}\right] \mathrm{d}x$

Initialize: $V_1 = 4(2G)^2$, $w_1 = \mathbf{0}$, $\theta_1 = \mathbf{0}$, $B_1 = 4$, $h_1 = \mathbf{0}$

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

Play w_t, receive subgradient g_t and hint h_{t+1}

Set $\bar{V}_{t+1} = \bar{V}_t + \|g_t - h_t\|^2$, $B_{t+1} = B_t + \frac{\|g_t-h_t\|^2}{(2G)^2}$, and $\alpha_{t+1} = \frac{\epsilon}{\sqrt{B_{t+1}}\log^2(B_{t+1})}$

Set $\theta_{t+1} = \theta_t - g_t$ and $\widetilde{\theta}_{t+1} = \theta_{t+1} - h_{t+1}$

Update

$$w_{t+1} = \arg\min_{w\in\mathbb{R}^d} \left\langle h_{t+1} + \sum_{s=1}^{t} g_s, w \right\rangle + \psi(w;\bar{V}_{t+1},\alpha_{t+1}) \tag{12}$$

$$= \frac{\widetilde{\theta}_{t+1}}{\|\widetilde{\theta}_{t+1}\|}\alpha_{t+1} \begin{cases} \exp\left(\frac{\|\widetilde{\theta}_{t+1}\|^2}{36\bar{V}_{t+1}}\right) - 1 & \text{if } \|\widetilde{\theta}_{t+1}\| \le \frac{6\bar{V}_{t+1}}{(2G)} \\ \exp\left(\frac{\|\widetilde{\theta}_{t+1}\|}{3(2G)} - \frac{6\bar{V}_{t+1}}{(2G)}\right) - 1 & \text{otherwise} \end{cases} \tag{13}$$

end

A.3. Optimistic FTRL with Parameter-free Regularizer

In this section, we provide a comparator-adaptive optimistic algorithm, Algorithm 4, that combines optimistic FTRL with the parameter-free regularizer in Jacobsen and Cutkosky (2022). This is a slight simplification of Jacobsen and Cutkosky (2022, Algorithm 3) which used a somewhat complicated formulation based on mirror descent in order to develop dynamic regret guarantees; this is more general than we need since we focus primarily on static regret. Instead, we provide an FTRL-based formulation which significantly simplifies the exposition and analysis, and is likely much easier to follow for most readers familiar with online learning than the centered mirror descent argument of Jacobsen and Cutkosky (2022).

The algorithm is shown in Algorithm 4, which also clearly demonstrates that the algorithm can be implemented in $O(d)$ per-round computation. The closed-form shown in the pseudocode follows from the same arguments as Jacobsen and Cutkosky (2022, Theorem 1), which shows how to compute the equivalent non-optimistic update from the first-order optimality condition, so we omit the details here for brevity.

Theorem 9 (Comparator-adaptive Optimistic Algorithm) *Assume that the hints $h_t \in \mathbb{R}^d$ satisfy $\|h_t\| \le G$ for all t. For any $u \in \mathbb{R}^d$, Algorithm 4 guarantees*

$$\sum_{t=1}^{T} \langle g_t, w_t - u\rangle \le O\left(\|u\|\sqrt{\sum_{t=1}^{T}\|g_t - h_t\|^2\left(\log_+ \frac{\|u\|\sqrt{T}}{\epsilon}\right)} + \epsilon G + G\|u\|\left(\log_+ \frac{\|u\|\sqrt{T}}{\epsilon}\right).\right).$$

Moreover, the algorithm admits an efficient closed-form update with $O(d)$ time per iteration.

Proof For all t, let $\psi_t^{\text{PF}}(w) \triangleq \psi(w; \bar{V}_t, \alpha_t)$ as defined in Algorithm 4. Applying a typical regret bound for optimistic FTRL (e.g., see (Orabona, 2019, Theorem 7.36)), we have

$$\sum_{t=1}^{T} \langle g_t, w_t - u \rangle \leq \psi_{T+1}^{\text{PF}}(u) + \sum_{t=1}^{T} \underbrace{\left(\langle g_t - h_t, w_t - w_{t+1} \rangle - \mathcal{D}_{\psi_t^{\text{PF}}}(w_{t+1}, w_t) - (\psi_{t+1}^{\text{PF}} - \psi_t^{\text{PF}})(w_{t+1}) \right)}_{\delta_t}.$$

Then apply Lemma 8 with $M_t = 2G \geq \|g_t - h_t\|$, $\bar{V}_t = 4M_t^2 + \sum_{s=1}^{t-1} \|g_s - h_s\|^2$, and with any non-increasing sequence $\alpha_1 \geq \ldots \geq \alpha_T > 0$,

$$\sum_{t=1}^{T} \delta_t \leq \sum_{t=1}^{T} \frac{2\alpha_t \|g_t - h_t\|^2}{\sqrt{\bar{V}_T}},$$

apply Lemma 18 that defines $\alpha_t = \frac{\epsilon}{\sqrt{B_t} \log^2(B_t)}$ where $B_t = 4 + \sum_{s=1}^{t-1} \|g_s - h_s\|^2 / (2G)^2$, then

$$\sum_{t=1}^{T} \delta_t \leq 16\epsilon G.$$

And Lemma 8 also gives us

$$\begin{aligned} \psi_{T+1}^{\text{PF}}(u) &\leq O\left(\|u\| \max\left\{ \sqrt{\bar{V}_{T+1} \log(\|u\| / \alpha_{T+1} + 1)}, G \log(\|u\| / \alpha_{T+1} + 1) \right\} \right) \\ &\leq O\left(\|u\| \max\left\{ \sqrt{\bar{V}_{T+1} \log(\|u\| \sqrt{T}/\epsilon + 1)}, G \log(\|u\| \sqrt{T}/\epsilon + 1) \right\} \right). \end{aligned}$$

Plugging the previous two displays in above yields the stated regret guarantee. The per-iteration efficiency can be easily seen from the updates in Eq. (13). ■

A.4. Proof of Theorem 2

In this section we provide the guarantee for our comparator-adaptive gradient variation bound. The result is re-stated below for convenience.

Theorem 2 (Comparator-Adaptive Gradient-Variation Regret) *For any $u \in \mathbb{R}^d$, Algorithm 4 (in Appendix A.3) with $h_t = g_{t-1}$ guarantees*

$$\text{REG}_T(u) \leq \widetilde{O}\left(\|u\| \sqrt{V_T(u) \left(\log_+ \frac{\|u\|\sqrt{T}}{\epsilon} \right)} + L \|u\|^2 + G \|u\| + \epsilon G \right).$$

where we keep the log factors of the dominant term in the $\widetilde{O}(\cdot)$-notation for clarity. Moreover, the algorithm admits an efficient closed-form update with $O(d)$ time per iteration.

Proof Applying Theorem 9 that guarantees

$$\sum_{t=1}^{T} \langle g_t, w_t - u \rangle \leq O\left(\|u\| \sqrt{\bar{V}_T \left(\log_+ \frac{\|u\| \sqrt{T}}{\epsilon} \right)} + \epsilon G + G \|u\| \left(\log_+ \frac{\|u\| \sqrt{T}}{\epsilon} \right) \right),$$

where $\bar{V}_T = \sum_{t=1}^T \|g_t - h_t\|^2$. Hence, Algorithm 4 satisfies the conditions of Theorem 1 with

$$A_T(u) = O\left(\epsilon G + G\|u\|\left(\log_+ \frac{\|u\|\sqrt{T}}{\epsilon}\right)\right), \quad B_T(u) = O\left(\|u\|\sqrt{\left(\log_+ \frac{\|u\|\sqrt{T}}{\epsilon}\right)}\right).$$

Plugging these into Theorem 1 yields the desired result. ■

Appendix B. Proofs for Section 3

B.1. Proof of Theorem 3

Here we re-state the full version for Theorem 3 of the main text, which additionally includes the derived optimistic regret bound.

Theorem 3 (Full Version) *Assume that the optimistic hints h_t satisfy $\|h_t\| \le G$ for all t. For any $u \in \mathbb{R}^d$, Algorithm 1 guarantees*

$$\text{REG}_T(u) \le \widetilde{O}\left(\|u\|\sqrt{\sum_{t=1}^T \|g_t - h_t\|^2 \left(\log_+ \tfrac{\|u\|\sqrt{T}}{\epsilon}\right)} + \epsilon\gamma + \epsilon G + \gamma\|u\| + \tfrac{\gamma}{\epsilon}\|u\|^2 + \tfrac{\epsilon}{\gamma}G^2\right).$$

Moreover, by setting hints $h_t = g_{t-1}$, for any $u \in \mathbb{R}^d$ it holds that

$$\text{REG}_T(u) \le \widetilde{O}\left(\|u\|\sqrt{V_T(u)\left(\log_+ \tfrac{\|u\|\sqrt{T}}{\epsilon}\right)} + L\|u\|^2 + \epsilon\gamma + \epsilon G + \gamma\|u\| + \tfrac{\gamma}{\epsilon}\|u\|^2 + \tfrac{\epsilon}{\gamma}G^2\right).$$

Proof Consider the standard optimistic FTRL bound (setting $h_1 = h_{T+1} = \mathbf{0}$):

$$\sum_{t=1}^T \langle g_t, w_t - u\rangle \le \Psi_{T+1}(u) + \sum_{t=1}^T \underbrace{\langle g_t - h_t, w_t - w_{t+1}\rangle - \mathcal{D}_{\Psi_t}(w_{t+1}, w_t) - (\Psi_{t+1} - \Psi_t)(w_{t+1})}_{\delta_t}.$$

We proceed by bounding the terms $\sum_{t=1}^T \delta_t$. We begin by replacing $\Delta_t = g_t - h_t$ with a *clipped* version $\widehat{\Delta}_t = \Delta_t \cdot \min\left\{1, \frac{\widehat{M}_t}{\|\Delta_t\|}\right\}$, where $\widehat{M}_t = \max_{s<t}\|\Delta_t\|$. To do this, let $\Psi_t(w) = \psi_t(w) + \frac{\beta}{2}\|w\|^2$, where $\psi_t(w) \triangleq \psi^{\text{PF}}(w; \alpha_t, \bar{V}_t, \widehat{M}_t)$ and $\beta > 0$, then we have:

$$\begin{aligned}
\sum_{t=1}^T \delta_t &= \sum_{t=1}^T \langle \Delta_t, w_t - w_{t+1}\rangle - \mathcal{D}_{\Psi_t}(w_{t+1}, w_t) - (\Psi_{t+1} - \Psi_t)(w_{t+1}) \\
&= \sum_{t=1}^T \left\langle \widehat{\Delta}_t, w_t - w_{t+1}\right\rangle - \mathcal{D}_{\psi_t}(w_{t+1}, w_t) - (\psi_{t+1} - \psi_t)(w_{t+1}) \\
&\quad + \sum_{t=1}^T \left\langle \Delta_t - \widehat{\Delta}_t, w_t - w_{t+1}\right\rangle - \frac{\beta}{2}\|w_t - w_{t+1}\|^2
\end{aligned}$$

$$\leq \sum_{t=1}^{T} \left\langle \widehat{\Delta}_t, w_t - w_{t+1} \right\rangle - \mathcal{D}_{\psi_t}(w_{t+1}, w_t) - (\psi_{t+1} - \psi_t)(w_{t+1}) + \frac{1}{2\beta} \sum_{t=1}^{T} \left\| \Delta_t - \widehat{\Delta}_t \right\|^2 .$$

By definition of the clipped optimistic gap $\widehat{\Delta}_t$, we have $\|\Delta_t - \widehat{\Delta}_t\| = \widehat{M}_{t+1} - \widehat{M}_t$, hence:

$$\sum_{t=1}^{T} \left\| \Delta_t - \widehat{\Delta}_t \right\|^2 = \sum_{t=1}^{T} \left(\widehat{M}_{t+1} - \widehat{M}_t \right)^2 \leq \left(\sum_{t=1}^{T} \left(\widehat{M}_{t+1} - \widehat{M}_t \right) \right)^2 = \left(\widehat{M}_{T+1} - \widehat{M}_1 \right)^2 .$$

Therefore, we conclude that:

$$\sum_{t=1}^{T} \langle g_t, w_t - u \rangle \leq \psi_{T+1}(u) + \sum_{t=1}^{T} \underbrace{\left\langle \widehat{\Delta}_t, w_t - w_{t+1} \right\rangle - \mathcal{D}_{\psi_t}(w_{t+1}, w_t) - (\psi_{t+1} - \psi_t)(w_{t+1})}_{=:\widehat{\delta}_t}$$
$$+ \frac{\beta}{2} \|u\|^2 + \frac{1}{2\beta} \left(\widehat{M}_{T+1} - \widehat{M}_1 \right)^2 .$$

To bound $\sum_{t=1}^{T} \widehat{\delta}_t$, we apply Lemma 8 by defining $\bar{V}_t = 4\widehat{M}_t^2 + \sum_{s=1}^{t-1} \|\widehat{\Delta}_s\|^2$, let $\alpha_1 \geq \ldots \geq \alpha_T$ be a non-increasing sequence, and define $\psi_t(w) = 3 \int_0^{\|w\|} \min_{\eta \leq 1/\widehat{M}_t} \left[\frac{\log(x/\alpha_t + 1)}{\eta} + \eta \bar{V}_t \right] \mathrm{d}x$, then

$$\sum_{t=1}^{T} \widehat{\delta}_t \leq \sum_{t=1}^{T} \frac{2\alpha_t \left\| \widehat{\Delta}_t \right\|^2}{\sqrt{\bar{V}_t}} .$$

We conclude that:

$$\begin{aligned}
\sum_{t=1}^{T} \langle g_t, w_t - u \rangle &\leq \psi_{T+1}(u) + \sum_{t=1}^{T} \widehat{\delta}_t + \frac{\beta}{2} \|u\|^2 + \frac{1}{2\beta} \left(\widehat{M}_{T+1} - \widehat{M}_1 \right)^2 \\
&\leq \psi_{T+1}(u) + \sum_{t=1}^{T} \frac{2\alpha_t \left\| \widehat{\Delta}_t \right\|^2}{\sqrt{\bar{V}_t}} + \frac{\beta}{2} \|u\|^2 + \frac{1}{2\beta} \left(\widehat{M}_{T+1} - \widehat{M}_1 \right)^2 \\
&\leq 6 \|u\| \max \left\{ \sqrt{\bar{V}_{T+1} \log(\|u\| / \alpha_{T+1} + 1)}, \widehat{M}_{T+1} \log(\|u\| / \alpha_{T+1} + 1) \right\} \\
&\quad + \sum_{t=1}^{T} \frac{2\alpha_t \left\| \widehat{\Delta}_t \right\|^2}{\sqrt{\bar{V}_t}} + \frac{\beta}{2} \|u\|^2 + \frac{1}{2\beta} \left(\widehat{M}_{T+1} - \widehat{M}_1 \right)^2 ,
\end{aligned}$$

where the last inequality uses Lemma 8. Finally, to apply Lemma 18, we define $\alpha_t = \frac{\epsilon}{\sqrt{B_t} \log^2(B_t)}$ with $B_t = 4 + \sum_{i=1}^{t-1} \|\widehat{\Delta}_i\|^2 / \widehat{M}_i^2$, which yields

$$\sum_{t=1}^{T} \langle g_t, w_t - u \rangle \leq 6 \|u\| \max \left\{ \sqrt{\bar{V}_{T+1} \log \left(\frac{\|u\| \sqrt{B_{T+1}} \log^2(B_{T+1})}{\epsilon} + 1 \right)}, \right.$$
$$\left. \widehat{M}_{T+1} \log \left(\frac{\|u\| \sqrt{B_{T+1}} \log^2(B_{T+1})}{\epsilon} + 1 \right) \right\}$$

$$+ 8\epsilon\widehat{M}_T + \frac{\beta}{2}\|u\|^2 + \frac{1}{2\beta}\left(\widehat{M}_{T+1} - \widehat{M}_1\right)^2.$$

Finally, the stated bound follows by bounding $B_t \le 3 + t$ (since $\|\widehat{\Delta}_t\| \le \widehat{M}_t$ by definition), and that $0 < \widehat{M}_{T+1} \le \max\{\gamma, 2G\} \le \gamma + 2G$, then $(\widehat{M}_{T+1} - \widehat{M}_1)^2 = (\widehat{M}_{T+1} - \gamma)^2 \le 4G^2$.

The second regret bound is by applying Theorem 1, for which we have

$$A_T(u) = O\left(\|u\|(\gamma + G)\left(\log_+ \frac{\|u\|\sqrt{T}}{\epsilon}\right) + \epsilon(\gamma + G) + \frac{\gamma}{\epsilon}\|u\|^2 + \frac{\epsilon}{\gamma}G^2\right),$$

$$B_T(u) = O\left(\|u\|\sqrt{\log_+\left(\frac{\|u\|\sqrt{T}}{\epsilon}\right)}\right).$$

Plugging these into Theorem 1 yields the desired result. ■

B.2. Efficiency of the Virtual Clipping Algorithm

In this section we provide the detailed discussion of the computational efficiency of the virtual clipping algorithm, Algorithm 1, proposed in Section 3.1. We first recall Proposition 1, re-stated below for convenience, which we prove later in this section.

Proposition 1 *Algorithm 1 guarantees that, for any $t \in [T]$, the direction of w_{t+1} is determined by*

$$\frac{w_{t+1}}{\|w_{t+1}\|} = -\frac{h_{t+1} + \sum_{s=1}^t g_s}{\left\|h_{t+1} + \sum_{s=1}^t g_s\right\|}.$$

Moreover, denoting $\alpha := \alpha_{t+1}$, $V := \bar{V}_{t+1}$, and $M := \widehat{M}_{t+1}$ the magnitude of w_{t+1} satisfies

$$\big|\, \|w_{t+1}\| - \|w_t\| \,\big| \le \frac{\|g_t - h_t + h_{t+1}\|}{\beta}, \quad \textit{and}$$

$$\left\|h_{t+1} + \sum_{s=1}^t g_s\right\| = \begin{cases} 6\sqrt{V\ln(\|w_{t+1}\|/\alpha + 1)} + \beta\|w_{t+1}\|, & \textit{if } \sqrt{\frac{\ln(\|w_{t+1}\|/\alpha+1)}{V}} \le \frac{1}{M}, \\ 3\left(M\ln(\|w_{t+1}\|/\alpha + 1) + \frac{V}{M}\right) + \beta\|w_{t+1}\|, & \textit{otherwise}. \end{cases}$$

Using Proposition 1 and denoting $\theta_t := \|g_t - h_t + h_{t+1}\|$, we have that the norm of w_{t+1} is no more than a factor of θ_t/β away from the norm of w_t. This means that we could naively apply a binary search over a bracket of $[\|w_t\| - \theta_t/\beta, \|w_t\| + \theta_t/\beta]$ to compute $\|w_{t+1}\|$ up to ϵ accuracy using $O(\log(2\theta_t/\beta\epsilon))$ iterations. Thus, we can solve $\|w_{t+1}\|$ up to $\epsilon = \frac{2\theta_t}{\beta T^k} = O(1/T^k)$ accuracy for any $k \ge 1$ while still matching the efficiency of Wang et al. (2025), which would already be sufficient for most applications. In practice, this naive estimate above could potentially be significantly reduced by applying a more sophisticated solver than the simple binary search suggested above (e.g., by applying Newton's method).

Proof [Proof of Proposition 1] We first present a direct derivation of the update for Algorithm 1. Denote $\widehat{\psi}(x; \alpha, V, M) = 3\int_0^x \min_{\eta \le 1/M}\left[\frac{\ln(x/\alpha+1)}{\eta} + \eta V\right] dx$ and $\psi(w; \alpha, V, M) = \widehat{\psi}(\|w\|; \alpha, V, M)$. On round t, the algorithm chooses a hybrid regularizer $\Psi_{t+1}(w) = \psi^{\text{PF}}\left(w; \alpha_{t+1}, \bar{V}_{t+1}, \widehat{M}_{t+1}\right) + \frac{\beta}{2}\|w\|^2$ (we use simplified notations of α, V, M in the following) and updates using $w_{t+1} =$

$\arg\min_{w\in\mathbb{R}^d} \left\langle h_{t+1} + \sum_{s=1}^{t} g_s, w\right\rangle + \Psi_{t+1}(w)$. By the first-order optimality condition of w_{t+1}, we have

$$\mathbf{0} = h_{t+1} + \sum_{s=1}^{t} g_s + \nabla\psi^{\text{PF}}\big(w_{t+1}; \alpha, V, M\big) + \beta w_{t+1}$$

$$= \begin{cases} h_{t+1} + \sum_{s=1}^{t} g_s + 6\frac{w_{t+1}}{\|w_{t+1}\|}\sqrt{V\ln(\|w_{t+1}\|/\alpha+1)} + \beta w_{t+1}, & \text{if } \sqrt{\frac{\ln(\|w_{t+1}\|/\alpha+1)}{V}} \le \frac{1}{M}, \\ h_{t+1} + \sum_{s=1}^{t} g_s + 3\frac{w_{t+1}}{\|w_{t+1}\|}\left(M\ln(\|w_{t+1}\|/\alpha+1) + \frac{V}{M}\right) + \beta w_{t+1}, & \text{otherwise.} \end{cases}$$

This immediately implies that w_{t+1} is in the direction of $-h_{t+1} - \sum_{s=1}^{t} g_s$. Taking the norm on both sides of the above equation, we have

$$\left\|h_{t+1} + \sum_{s=1}^{t} g_s\right\| = \begin{cases} 6\sqrt{V\ln(\|w_{t+1}\|/\alpha+1)} + \beta\|w_{t+1}\|, & \text{if } \sqrt{\frac{\ln(\|w_{t+1}\|/\alpha+1)}{V}} \le \frac{1}{M}, \\ 3\left(M\ln(\|w_{t+1}\|/\alpha+1) + \frac{V}{M}\right) + \beta\|w_{t+1}\|, & \text{otherwise.} \end{cases}$$

This equation has no closed-form solution for $\|w_{t+1}\|$ in general, but an approximate solution can be found via binary search over a small bracket around $\|w_t\|$, as shown in Proposition 10 applied with $\widehat{\psi}_t(\|w\|) = \widehat{\psi}_t^{\text{PF}}(\|w\|; \alpha_t, \bar{V}_t, \widehat{M}_t)$, which can easily be seen to satisfy the required conditions $\widehat{\psi}_t'(x) \ge 0$ and $\psi_t'(x) \le \psi_{t+1}'(x)$ for non-increasing sequence $(\alpha_t)_t$.

∎

Proposition 10 *Let $\widehat{\psi}_1, \widehat{\psi}_2, \ldots$ be an sequence of convex regularizers satisfying $\widehat{\psi}_t'(x) \ge 0$ and $\widehat{\psi}_t'(x) \le \widehat{\psi}_{t+1}'(x)$ for any t and any $x \ge 0$. For each t define $\psi_t(w) = \widehat{\psi}_t(\|w\|)$ and $\Psi_t(w) = \psi_t(w) + \frac{\beta}{2}\|w\|^2$ for $\beta > 0$, and suppose that on each round we play*

$$w_t = \arg\min_{w\in\mathbb{R}^d} \left\langle h_t + \sum_{s=1}^{t-1} g_s, w\right\rangle + \Psi_t(w).$$

Then for any t, it holds that

$$|\|w_{t+1}\| - \|w_t\|| \le \frac{\|g_t - h_t + h_{t+1}\|}{\beta}.$$

Proof Observe that for any t, we have via the first-order optimality condition for the optimistic FTRL update that

$$w_{t+1} = \arg\min_{w\in\mathbb{R}^d} \left\langle h_{t+1} + \sum_{s=1}^{t} g_s, w\right\rangle + \Psi_{t+1}(w)$$

$$\Longrightarrow \quad \nabla\Psi_{t+1}(w_{t+1}) = -h_{t+1} - \sum_{s=1}^{t} g_s,$$

and moreover, since this holds for any t, we also have $\nabla\Psi_t(w_t) = -h_t - \sum_{s=1}^{t-1} g_s$, so the previous display can be written as

$$\nabla\Psi_{t+1}(w_{t+1}) = \nabla\Psi_t(w_t) - (g_t - h_t + h_{t+1}).$$

Therefore, by denoting $\Psi_t(w) = \widehat{\Psi}_t(\|w\|)$ and $\psi_t(w) = \widehat{\psi}_t(\|w\|)$, we have

$$\begin{aligned}\widehat{\Psi}'_{t+1}(\|w_{t+1}\|) &= \widehat{\psi}_{t+1}{}'(\|w_{t+1}\|) + \beta \|w_{t+1}\| \\ &= \|\nabla\Psi_t(w_t) - (g_t - h_t + h_{t+1})\| \le \|\nabla\Psi_t(w_t)\| + \|g_t - h_t + h_{t+1}\| \\ &= \widehat{\Psi}'_t(\|w_t\|) + \|g_t - h_t + h_{t+1}\| \le \widehat{\Psi}'_{t+1}(\|w_t\|) + \|g_t - h_t + h_{t+1}\|,\end{aligned}$$

where we've used the facts that $\widehat{\Psi}'_t(x) \ge 0$ and $\widehat{\Psi}'_t(x) \le \widehat{\Psi}'_{t+1}(x)$ for all $x \ge 0$. From this, we have in particular that

$$\left|\widehat{\Psi}'_{t+1}(\|w_{t+1}\|) - \widehat{\Psi}'_{t+1}(\|w_t\|)\right| \le \|g_t - h_t + h_{t+1}\| := \theta_t.$$

We can use this to derive a bound on how far $\|w_{t+1}\|$ is from $\|w_t\|$ by studying the inverse of $\widehat{\Psi}'_{t+1}$, since $\|w\| = (\widehat{\Psi}'_{t+1})^{-1}(\widehat{\Psi}'_{t+1}(\|w\|))$.

Now, since $\widehat{\psi}_{t+1}$ is convex, we have $\widehat{\psi}_{t+1}{}''(x) \ge 0$ for all x and therefore, $\widehat{\Psi}''_{t+1}(x) = \widehat{\psi}_{t+1}{}''(x) + \beta \ge \beta > 0$ for all x. Therefore, $\widehat{\Psi}'_{t+1}(x)$ is β-strongly monotone, which implies that it has a unique inverse $(\widehat{\Psi}'_{t+1})^{-1}$, and moreover, that $(\widehat{\Psi}'_{t+1})^{-1}$ is $1/\beta$-Lipschitz. Hence,

$$\begin{aligned}|\|w_{t+1}\| - \|w_t\|| &= \left|(\widehat{\Psi}'_{t+1})^{-1}\left(\widehat{\Psi}'_{t+1}(\|w_{t+1}\|)\right) - (\widehat{\Psi}'_{t+1})^{-1}\left(\widehat{\Psi}'_{t+1}(\|w_t\|)\right)\right| \\ &\le \frac{\left|\widehat{\Psi}'_{t+1}(\|w_{t+1}\|) - \widehat{\Psi}'_{t+1}(\|w_t\|)\right|}{\beta} \le \frac{\|g_t - h_t + h_{t+1}\|}{\beta} = \frac{\theta_t}{\beta}.\end{aligned}$$

■

B.3. Proof of Theorem 4

To make the paper self-contained, we first restate Cutkosky and Mhammedi (2024, Theorem 1) as a lemma here. To be consistent with this paper, we replace their notation γ with γ'/ϵ', replace h_1 with γ', and replace notation ϵ with ϵ'.

Lemma 11 (Theorem 1 of Cutkosky and Mhammedi (2024)) *There exists an online learning algorithm $\mathcal{A}$ with any inputs $\epsilon' > 0$ and $\gamma' > 0$, that runs in $O(d)$ time per iteration, such that for any sequence $g_1, \ldots, g_T$, and for any $u \in \mathbb{R}^d$, guarantees*

$$\begin{aligned}\sum_{t=1}^T \langle g_t, w_t - u\rangle \le O\Bigg(&\|u\|\sqrt{\sum_{t=1}^T \|g_t\|^2 \left(\log_+ \tfrac{G\|u\|\sqrt{T}}{\epsilon'\gamma'}\right)} + G\|u\|\left(\log_+ \tfrac{G\|u\|\sqrt{T}}{\epsilon'\gamma'}\right) \\ &+ \tfrac{\epsilon' G^2}{\gamma'}\left(\log_+ \tfrac{G}{\gamma'}\right) + \tfrac{\gamma'\|u\|^2}{\epsilon'}\left(\log_+ \tfrac{\|u\|}{\epsilon'}\right) + \epsilon' G + \epsilon'\gamma'\Bigg),\end{aligned}$$

where $G = \max\{\gamma', \max_{t\in[T]} \|g_t\|\}$.

We restate Theorem 4 here with more detailed logarithmic factors.

Theorem 4 (Full Version) *For any $u \in \mathbb{R}^d$, Algorithm 2 guarantees*

$$\begin{aligned}\mathrm{REG}_T(u) \le O\Bigg(&\epsilon G + \epsilon\gamma + \frac{\epsilon G^2}{\gamma}\left(\log_+ \frac{G}{\gamma}\right) + \frac{\gamma \|u\|^2}{\epsilon}\left(\log_+ \frac{\|u\|}{\epsilon}\right) + \frac{\epsilon G^4}{\gamma^3}\left(\log_+ \frac{G}{\gamma}\right) \\ &+ \|u\| G\left(\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}\right)^{\frac{3}{2}} + \frac{\gamma^3 \|u\|^2}{\epsilon G^2}\left(\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}\right)\left(\log_+ \frac{\gamma\|u\|}{\epsilon G}\right) \\ &+ \|u\| \sqrt{V_T(u)\left(\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}\right)} + L\|u\|^2\left(\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}\right)\Bigg).\end{aligned}$$

Moreover, the algorithm admits an efficient closed-form update with $O(d)$ time per iteration.

Proof We will apply our black-box reduction (Proposition 2) with the algorithm of Cutkosky and Mhammedi (2024, Theorem 1), provided in Lemma 11 above, for both $\mathcal{A}_x$ and $\mathcal{A}_y$. By Lemma 11, the algorithm $\mathcal{A}_x$ with inputs $\epsilon' = \epsilon$ and $\gamma' = \gamma$ satisfies the conditions of Proposition 2 with

$$\begin{aligned}A_T^{\mathcal{A}_x}(u) = O\Bigg(&\epsilon G + \epsilon\gamma + \frac{\epsilon G^2}{\gamma}\left(\log_+ \frac{G}{\gamma}\right) \\ &+ \frac{\gamma\|u\|^2}{\epsilon}\left(\log_+ \frac{\|u\|}{\epsilon}\right) + \|u\| G\left(\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}\right)\Bigg), \quad (14)\\ B_T^{\mathcal{A}_x}(u) = O&\left(\|u\|\sqrt{\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}}\right), \quad (15)\end{aligned}$$

where G is the Lipschitz constant. Likewise, for $\mathcal{A}_y$, given any $\epsilon' > 0$ and $\gamma' > 0$ (to be specified later) we have:

$$\begin{aligned}A_T^{\mathcal{A}_y}(\mathring{y}) = O\Bigg(&\epsilon' GH + \epsilon'\gamma' + \frac{\epsilon' G^2 H^2}{\gamma'}\left(\log_+ \frac{GH}{\gamma'}\right) \\ &+ \frac{\gamma' |\mathring{y}|^2}{\epsilon'}\left(\log_+ \frac{|\mathring{y}|}{\epsilon'}\right) + |\mathring{y}| GH\left(\log_+ \frac{|\mathring{y}| GH\sqrt{T}}{\epsilon'\gamma'}\right)\Bigg), \\ B_T^{\mathcal{A}_y}(\mathring{y}) = O&\left(|\mathring{y}|\sqrt{\log_+ \frac{|\mathring{y}| GH\sqrt{T}}{\epsilon'\gamma'}}\right) = O\left(|\mathring{y}|\lambda_T(\mathring{y})\right),\end{aligned}$$

where $H \triangleq \max_{t\in[T]} \|h_t\|$ and $\lambda_T(\mathring{y}) \triangleq 1 + \sqrt{\log_+\left(|\mathring{y}| GH\sqrt{T}/(\epsilon'\gamma')\right)}$. Applying Proposition 2 we have

$$\sum_{t=1}^T \langle g_t, w_t - u\rangle \le A_T^{\mathcal{A}_x}(u) + A_T^{\mathcal{A}_y}(\mathring{y}) + 2B_T^{\mathcal{A}_x}(u)\sqrt{H^2\lambda_T(\mathring{y})^2 + \frac{1}{2}\sum_{t=1}^T \|g_t - h_t\|^2}. \quad (16)$$

where $\mathring{y} = B_T^{\mathcal{A}_x}(u)/H$. Expanding $\mathring{y}$ in $A_T^{\mathcal{A}_y}(\mathring{y})$ and setting $\epsilon' = \epsilon/\gamma$, and $\gamma' = \gamma^2$, we have

$$\begin{aligned}
A_T^{\mathcal{A}_y}(\mathring{y}) &= A_T^{\mathcal{A}_y}(B_T^{\mathcal{A}_x}(u)/H) \\
&= O\Bigg(\epsilon' GH + \epsilon'\gamma' + \frac{\epsilon' G^2 H^2}{\gamma'}\left(\log_+ \frac{GH}{\gamma'}\right) \\
&\qquad + \frac{\gamma' B_T^{\mathcal{A}_x}(u)^2}{\epsilon' H^2}\left(\log_+ \frac{B_T^{\mathcal{A}_x}(u)}{\epsilon' H}\right) + B_T^{\mathcal{A}_x}(u) G \left(\log_+ \frac{B_T^{\mathcal{A}_x}(u) G\sqrt{T}}{\epsilon'\gamma'}\right)\Bigg) \\
&= O\Bigg(\epsilon' GH + \epsilon'\gamma' + \frac{\epsilon' G^2 H^2}{\gamma'}\left(\log_+ \frac{GH}{\gamma'}\right) \\
&\qquad + \frac{\gamma' \|u\|^2}{\epsilon' H^2}\left(\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}\right)\left(\log_+ \frac{\|u\|}{\epsilon' H}\right) \\
&\qquad + \|u\| G \sqrt{\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}}\left(\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon'\gamma'}\right)\Bigg) \\
&= O\Bigg(\frac{\epsilon GH}{\gamma} + \epsilon\gamma + \frac{\epsilon G^2 H^2}{\gamma^3}\left(\log_+ \frac{GH}{\gamma^2}\right) \\
&\qquad + \frac{\gamma^3 \|u\|^2}{\epsilon H^2}\left(\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}\right)\left(\log_+ \frac{\gamma\|u\|}{\epsilon H}\right) \\
&\qquad + \|u\| G \left(\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}\right)^{\frac{3}{2}}\Bigg),
\end{aligned} \tag{17}$$

and likewise, expanding $\mathring{y}$ in $\lambda_T(\mathring{y})$ yields

$$\lambda_T(\mathring{y})^2 = \lambda_T(B_T^{\mathcal{A}_x}(u)/H)^2 = 1 + \log_+ \frac{B_T^{\mathcal{A}_x}(u) G\sqrt{T}}{\epsilon\gamma} = O\left(\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}\right). \tag{18}$$

Plugging Equations (14), (15), (17) and (18) back into Equation (16), we have

$$\begin{aligned}
&\sum_{t=1}^{T} \langle g_t, w_t - u\rangle \le A_T^{\mathcal{A}_x}(u) + A_T^{\mathcal{A}_y}(\mathring{y}) + 2B_T^{\mathcal{A}_x}(u)\sqrt{H^2\lambda_T(\mathring{y})^2 + \frac{1}{2}\sum_{t=1}^{T}\|g_t - h_t\|^2} \\
&\le O\Bigg(\epsilon G + \epsilon\gamma + \frac{\epsilon GH}{\gamma} + \frac{\epsilon G^2}{\gamma}\left(\log_+ \frac{G}{\gamma}\right) + \frac{\gamma\|u\|^2}{\epsilon}\left(\log_+ \frac{\|u\|}{\epsilon}\right) + \frac{\epsilon G^2 H^2}{\gamma^3}\left(\log_+ \frac{GH}{\gamma^2}\right) \\
&\qquad + \|u\| H\left(\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}\right) + \frac{\gamma^3\|u\|^2}{\epsilon H^2}\left(\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}\right)\left(\log_+ \frac{\gamma\|u\|}{\epsilon H}\right) \\
&\qquad + \|u\| G\left(\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}\right)^{\frac{3}{2}} + \|u\|\sqrt{\bar{V}_T\left(\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}\right)}\Bigg),
\end{aligned}$$

Algorithm 5: Optimistic Reduction (Cutkosky, 2019c)

Input: Online Learning algorithms $\mathcal{A}_x$ defined on $\mathbb{R}^d$ and $\mathcal{A}_y$ defined on $\mathbb{R}$.
for $t = 1 : T$ **do**
  Observe $h_t \in \mathbb{R}^d$
  Get $x_t \in \mathbb{R}^d$ from $\mathcal{A}_x$ and $y_t \in \mathbb{R}$ from $\mathcal{A}_y$
  Play $w_t = x_t - y_t h_t$ and observe $g_t = \nabla f_t(w_t)$
  Pass g_t to $\mathcal{A}_x$ as the t^{th} subgradient
  Pass $-\langle g_t, h_t\rangle$ to $\mathcal{A}_y$ as the t^{th} subgradient
end

where we define $\bar{V}_T = \sum_{t=1}^T \|g_t - h_t\|^2$. Finally by setting $h_t = g_{t-1}$ and hence $H = G$, then apply Theorem 1, we conclude that:

$$
\begin{aligned}
&\sum_{t=1}^T f_t(w_t) - f_t(u) = \sum_{t=1}^T \langle g_t, w_t - u\rangle - \mathcal{D}_{f_t}(u, w_t) \\
&\leq O\Bigg(\epsilon G + \epsilon\gamma + \frac{\epsilon G^2}{\gamma}\left(\log_+ \frac{G}{\gamma}\right) + \frac{\gamma \|u\|^2}{\epsilon}\left(\log_+ \frac{\|u\|}{\epsilon}\right) + \frac{\epsilon G^4}{\gamma^3}\left(\log_+ \frac{G}{\gamma}\right) \\
&\qquad + \|u\| G \left(\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}\right)^{\frac{3}{2}} + \frac{\gamma^3 \|u\|^2}{\epsilon G^2}\left(\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}\right)\left(\log_+ \frac{\gamma \|u\|}{\epsilon G}\right) \\
&\qquad + \|u\| \sqrt{V_T(u)\left(\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}\right)} + L\|u\|^2\left(\log_+ \frac{\|u\| G\sqrt{T}}{\epsilon\gamma}\right)\Bigg),
\end{aligned}
$$

where $V_T(u) = \sum_{t=2}^T \|\nabla f_t(u) - \nabla f_{t-1}(u)\|^2$. ■

B.4. Proof of Proposition 2

We provide the Optimistic Reduction (Cutkosky, 2019c) in Algorithm 5 for self-containment, and give the following proof for the refined reduction in Proposition 2. The result is re-stated below for convenience.

Proposition 2 (Refined Optimistic Reduction) *Let $\mathcal{A}_x$ and $\mathcal{A}_y$ be online algorithms defined on $\mathcal{W}_x = \mathbb{R}^d$ and $\mathcal{W}_y = \mathbb{R}$ respectively. Suppose the following conditions hold:*

- *For each $z \in \{x, y\}$, $\mathcal{A}_z$ guarantees $\text{REG}_T^{\mathcal{A}_z}(u) \leq A_T^{\mathcal{A}_z}(u) + B_T^{\mathcal{A}_z}(u)\sqrt{\sum_{t=1}^T \|g_t\|^2}$ for any $u \in \mathcal{W}_z$ and $\{g_t\}_{t=1}^T$ in $\mathcal{W}_z$, where $A_T^{\mathcal{A}_z}$ and $B_T^{\mathcal{A}_z}$ are non-negative functions.*
- *$A_T^{\mathcal{A}_y}$ is non-decreasing and $B_T^{\mathcal{A}_y}(\mathring{y}) \leq |\mathring{y}|\,\lambda_T(\mathring{y})$ for some non-decreasing function $\lambda_T(\mathring{y}) \geq 1$.*

Then for any $u \in \mathbb{R}^d$, Algorithm 5 (in Appendix B.4) enjoys the following optimistic bound:

$$
\sum_{t=1}^T \langle g_t, w_t - u\rangle \leq A_T^{\mathcal{A}_x}(u) + A_T^{\mathcal{A}_y}(\mathring{y}) + 2B_T^{\mathcal{A}_x}(u)\sqrt{H_T^2 \lambda_T(\mathring{y})^2 + \frac{1}{2}\left[\sum_{t=1}^T \|g_t - h_t\|^2 - \|h_t\|^2\right]_+},
$$

where $H_T = \max_{t\in[T]} \|h_t\|$, $\mathring{y} = B_T^{\mathcal{A}_x}(u)/H_T$, *and* $[x]_+ \triangleq \max\{x, 0\}$.

Proof For $w_t = x_t - y_t h_t$ and $g_t = \nabla f_t(w_t)$, we have

$$\begin{aligned}
\sum_{t=1}^{T} \langle g_t, w_t - u\rangle &= \sum_{t=1}^{T} \langle g_t, x_t - u\rangle + \sum_{t=1}^{T} \langle -g_t, h_t\rangle y_t \\
&= \text{REG}_T^{\mathcal{A}_x}(u) + \sum_{t=1}^{T} \left(\langle -g_t, h_t\rangle y_t - \langle -g_t, h_t\rangle \mathring{y}\right) + \mathring{y} \sum_{t=1}^{T} \langle -g_t, h_t\rangle \\
&= \text{REG}_T^{\mathcal{A}_x}(u) + \text{REG}_T^{\mathcal{A}_y}(\mathring{y}) + \mathring{y} \sum_{t=1}^{T} \langle -g_t, h_t\rangle,
\end{aligned}$$

and moreover, using the fact that $-\langle g_t, h_t\rangle = \frac{1}{2}\|g_t - h_t\|^2 - \frac{1}{2}\|h_t\|^2 - \frac{1}{2}\|g_t\|^2$, we have

$$\sum_{t=1}^{T} \langle g_t, w_t - u\rangle \leq \text{REG}_T^{\mathcal{A}_x}(u) + \text{REG}_T^{\mathcal{A}_y}(\mathring{y}) + \frac{\mathring{y}}{2}\left(\bar{V}_T - \sum_{t=1}^{T}\|g_t\|^2\right),$$

where we've defined $\bar{V}_T := \left[\sum_{t=1}^{T} \|g_t - h_t\|^2 - \|h_t\|^2\right]_+$ and $[x]_+ = \max\{x, 0\}$. Applying the regret guarantees of $\mathcal{A}_x$ and $\mathcal{A}_y$, we have

$$\begin{aligned}
\sum_{t=1}^{T} \langle g_t, w_t - u\rangle &\leq A_T^{\mathcal{A}_x}(u) + B_T^{\mathcal{A}_x}(u)\sqrt{\sum_{t=1}^{T}\|g_t\|^2} + A_T^{\mathcal{A}_y}(\mathring{y}) + B_T^{\mathcal{A}_y}(\mathring{y})\sqrt{\sum_{t=1}^{T}\langle g_t, h_t\rangle^2} \\
&\quad + \frac{\mathring{y}}{2}\left(\bar{V}_T - \sum_{t=1}^{T}\|g_t\|^2\right) \\
&= A_T^{\mathcal{A}_x}(u) + A_T^{\mathcal{A}_y}(\mathring{y}) + \frac{\mathring{y}}{2}\bar{V}_T \\
&\quad + \left[B_T^{\mathcal{A}_x}(u) + B_T^{\mathcal{A}_y}(\mathring{y})\max_{t\in[T]}\|h_t\|\right]\sqrt{\sum_{t=1}^{T}\|g_t\|^2} - \frac{\mathring{y}}{2}\sum_{t=1}^{T}\|g_t\|^2 \\
&\leq A_T^{\mathcal{A}_x}(u) + A_T^{\mathcal{A}_y}(\mathring{y}) + \frac{\mathring{y}}{2}\bar{V}_T \\
&\quad + \sup_{X\geq 0}\left\{\left[B_T^{\mathcal{A}_x}(u) + B_T^{\mathcal{A}_y}(\mathring{y})\max_{t\in[T]}\|h_t\|\right]X - \frac{\mathring{y}}{2}X^2\right\} \\
&\leq A_T^{\mathcal{A}_x}(u) + A_T^{\mathcal{A}_y}(\mathring{y}) + \frac{\mathring{y}}{2}\bar{V}_T + \frac{\left[B_T^{\mathcal{A}_x}(u) + B_T^{\mathcal{A}_y}(\mathring{y})H_T\right]^2}{2\mathring{y}} \\
&\leq A_T^{\mathcal{A}_x}(u) + A_T^{\mathcal{A}_y}(\mathring{y}) + \frac{\mathring{y}}{2}\bar{V}_T + \frac{B_T^{\mathcal{A}_x}(u)^2 + B_T^{\mathcal{A}_y}(\mathring{y})^2 H_T^2}{\mathring{y}},
\end{aligned}$$

where we've used $aX - bX^2 \leq a^2/4b$ and defined $H_T = \max_t \|h_t\|$. For $B_T^{\mathcal{A}_y}(\mathring{y}) \leq \mathring{y}\lambda_T(\mathring{y})$, we have

$$\sum_{t=1}^{T} \langle g_t, w_t - u\rangle \leq A_T^{\mathcal{A}_x}(u) + A_T^{\mathcal{A}_y}(\mathring{y}) + \mathring{y}\left[H_T^2\lambda_T(\mathring{y})^2 + \frac{1}{2}\bar{V}_T\right] + \frac{B_T^{\mathcal{A}_x}(u)^2}{\mathring{y}}.$$

Since this holds for any $\mathring{y} \in \mathbb{R}_+$, we may take $\mathring{y} = B_T^{\mathcal{A}_x}(u)/\sqrt{H_T^2\lambda_T^2(B_T^{\mathcal{A}_x}(u)/H_T) + \frac{1}{2}\bar{V}_T} \leq B_T^{\mathcal{A}_x}(u)/H_T$, and so for monotonically increasing $A_T^{\mathcal{A}_y}(\mathring{y})$ and $\lambda_T(\mathring{y})$, we have

$$\sum_{t=1}^{T} \langle g_t, w_t - u \rangle \leq A_T^{\mathcal{A}_x}(u) + A_T^{\mathcal{A}_y}\left(B_T^{\mathcal{A}_x}(u)/H_T\right) + 2B_T^{\mathcal{A}_x}(u)\sqrt{H_T^2\lambda_T\left(B_T^{\mathcal{A}_x}(u)/H_T\right)^2 + \frac{1}{2}\bar{V}_T}.$$

The stated result defines $\mathring{y} = B_T^{\mathcal{A}_x}(u)/H_T$ for brevity. ■

Appendix C. Proofs for Section 4

C.1. Proof of Theorem 6

We re-state Theorem 6 here with more detailed logarithmic factors.

Theorem 6 (Full Version) *For any sequence $u_1, \ldots, u_T \in \mathbb{R}^d$, Algorithm 3 guarantees*

$$\begin{aligned}\text{REG}_T(u_{1:T}) \leq O\Bigg(&\sqrt{(M^2 + MP_T)\, V_T(u_{1:T})\left(\log_+ \tfrac{MG\sqrt{T}}{\epsilon\gamma}\right)} + L(M^2 + MP_T)\left(\log_+ \tfrac{MG\sqrt{T}}{\epsilon\gamma}\right) \\ &+ G\sqrt{M^2 + MP_T} + P_T G + \gamma P_T \\ &+ \epsilon G + \gamma M + \epsilon\gamma + \tfrac{\epsilon G^2}{\gamma}\left(\log_+ \tfrac{G}{\gamma}\right) + \tfrac{\gamma M^2}{\epsilon}\left(\log_+ \tfrac{M}{\epsilon}\right) + \tfrac{\epsilon G^4}{\gamma^3}\left(\log_+ \tfrac{G}{\gamma}\right) \\ &+ MG\left(\log_+ \tfrac{MG\sqrt{T}}{\epsilon\gamma}\right)^{\frac{3}{2}} + MG\left(\log_+ \tfrac{GT}{\gamma}\right) \\ &+ \tfrac{\gamma^3 M^2}{\epsilon G^2}\left(\log_+ \tfrac{MG\sqrt{T}}{\epsilon\gamma}\right)\left(\log_+ \tfrac{\gamma M}{\epsilon G}\right)\Bigg),\end{aligned}$$

where $M \triangleq \max_{t\in[T]} \|u_t\|$. Moreover, the algorithm runs in $O(d\log t)$ time on iteration t.

Proof Following Cutkosky and Orabona (2018, Theorem 2) and Jacobsen and Cutkosky (2022, Lemma 10), the linearized dynamic regret is bounded as

$$\begin{aligned}\sum_{t=1}^{T} \langle g_t, w_t - u_t \rangle &= \sum_{t=1}^{T} \langle g_t, x_t \rangle y_t - \langle g_t, u_t \rangle \\ &= \sum_{t=1}^{T} \langle g_t, x_t \rangle y_t - \langle g_t, x_t \rangle M + \sum_{t=1}^{T} \langle g_t, x_t \rangle M - \langle g_t, u_t \rangle \\ &= \text{REG}_T^{1\text{d}}(M) + M\sum_{t=1}^{T}\left\langle g_t, x_t - \frac{u_t}{M}\right\rangle \\ &= \text{REG}_T^{1\text{d}}(M) + M\text{REG}_T^{\mathcal{B}}(u_{1:T}/M),\end{aligned}$$

where we define $M \triangleq \max_{t\in[T]} \|u_t\|$, $\text{REG}_T^{1\text{d}}(M)$ denotes the static regret of an algorithm $\mathcal{A}_{1\text{d}}$ which receives optimistic hint $\langle h_t, x_t \rangle$, then chooses $y_t \in \mathbb{R}$ and suffers losses $y \mapsto \langle g_t, x_t \rangle y$, while $\text{REG}_T^{\mathcal{B}}(u_{1:T}/M)$ denotes the dynamic regret of an algorithm $\mathcal{A}_{\mathcal{B}}$ receiving optimism h_t and choosing $x_t \in \mathcal{B} = \{x \in \mathbb{R}^d : \|x\| \leq 1\}$ against losses $x \mapsto \langle g_t, x \rangle$. Moreover, $\|h_t\| \leq G$ since we set $h_t = g_{t-1}$.

For $\mathcal{A}^{\mathcal{B}}$, we apply Theorem 12 which ensures the following guarantee:

$$\text{REG}_T^{\mathcal{B}}(u_{1:T}/M) \leq O\Bigg(\sqrt{(1 + P_T/M) \sum_{t=1}^{T} \|g_t - h_t\|^2} + G\sqrt{P_T/M} \\ + GP_T/M + G \log_+ \left(\tfrac{GT}{\gamma}\right) + \gamma(1 + P_T/M)\Bigg).$$

For $\mathcal{A}^{1\text{d}}$, we apply proof in Theorem 4 with optimism $\langle h_t, x_t \rangle$ and gradient $\langle g_t, x_t \rangle$, hence:

$$\begin{aligned} \text{REG}_T^{1\text{d}}(M) &= \sum_{t=1}^{T} \langle g_t, x_t \rangle (y_t - M) \\ &\leq O\left(M\sqrt{\sum_{t=1}^{T} \|\langle g_t - h_t, x_t \rangle\|^2 \left(\log_+ \frac{MG\sqrt{T}}{\epsilon\gamma}\right)} + \text{CONST}\right) \\ &\leq O\left(M\sqrt{\sum_{t=1}^{T} \|g_t - h_t\|^2 \left(\log_+ \frac{MG\sqrt{T}}{\epsilon\gamma}\right)} + \text{CONST}\right), \end{aligned}$$

where we use $\|x_t\| \leq 1$ and denote constants by CONST:

$$\text{CONST} = \epsilon G + \epsilon\gamma + \frac{\epsilon G^2}{\gamma}\left(\log_+ \frac{G}{\gamma}\right) + \frac{\gamma M^2}{\epsilon}\left(\log_+ \frac{M}{\epsilon}\right) + \frac{\epsilon G^4}{\gamma^3}\left(\log_+ \frac{G}{\gamma}\right) \\ + MG\left(\log_+ \frac{MG\sqrt{T}}{\epsilon\gamma}\right)^{\frac{3}{2}} + \frac{\gamma^3 M^2}{\epsilon G^2}\left(\log_+ \frac{MG\sqrt{T}}{\epsilon\gamma}\right)\left(\log_+ \frac{\gamma M}{\epsilon G}\right).$$

Combining everything together and applying Theorem 1, we have:

$$\text{REG}_T(u_{1:T}) \leq O\Bigg(\sqrt{(M^2 + MP_T) V_T(u_{1:T}) \left(\log_+ \frac{MG\sqrt{T}}{\epsilon\gamma}\right)} \\ + L(M^2 + MP_T)\left(\log_+ \frac{MG\sqrt{T}}{\epsilon\gamma}\right) + G\sqrt{M^2 + MP_T} \\ + P_T G + MG\left(\log_+ \tfrac{GT}{\gamma}\right) + \gamma(D + P_T) + \text{CONST}\Bigg).$$

Finally, the computational efficiency is from Algorithm 2 and Algorithm 6. ∎

Algorithm 6: Anytime Dynamic Regret Minimization

Input: Domain diameter D, $\gamma > 0$.
Initialize: Base learner number $n = 1$, $h_1 = \mathbf{0}$, $w_{1,1} = \mathbf{0}$, $p_1 = 1$, index set $\mathcal{I}_1 = \emptyset$.
for $t = 1$ **to** T **do**
  Play $w_t = \sum_{i=1}^n p_{t,i} w_{t,i}$, then observe g_t and hint h_{t+1}
  Update $\mathcal{I}_n \leftarrow \mathcal{I}_n \cup \{t\}$, define $\bar{V}_{\mathcal{I}_n} = \sum_{s \in \mathcal{I}_n} \|g_s - h_s\|^2$ and $\bar{V}_t = \sum_{s=1}^t \|g_s - h_s\|^2$
  for $i = 1$ **to** n **do**
    Update $w_{t+1,i} = \Pi_{\mathcal{W}}[w_{t,i} - \eta_{t+1,i}(g_t - h_t + h_{t+1})]$ with $\eta_{t+1,i} = \frac{D 2^n}{\sqrt{\bar{V}_t}}$
  end
  if $\bar{V}_{\mathcal{I}_n}/\gamma^2 > 2^n$ **then**
    $n \leftarrow n + 1$
    Initialize $w_{t+1,n} = \mathbf{0}$, $p_{t+1} = \frac{1}{n}\mathbf{1}_n$, $\mathcal{I}_n = \emptyset$
  end
  else
    Define $\ell_t \in \mathbb{R}^n$ where $\ell_{t,i} = \langle g_t, w_{t,i} \rangle$
    Define $m_{t+1} \in \mathbb{R}^n$ where $m_{t+1,i} = \langle h_{t+1}, w_{t+1,i} \rangle$
    Update $p_{t+1} \in \Delta_n$ by $p_{t+1,i} \propto \exp\left(-\varepsilon_{t+1}\left(\sum_{s \in \mathcal{I}_n} \ell_s + m_{t+1}\right)\right)$ with $\varepsilon_{t+1} = \frac{1}{D\sqrt{\bar{V}_{\mathcal{I}_n}}}$
  end
end

C.2. Anytime and Lipschitz-Adaptive Dynamic Regret Algorithm

Theorem 12 *Assume that the domain $\mathcal{W}$ is bounded by D. For any sequence $u_1, \ldots, u_T$ in $\mathcal{W}$, Algorithm 6 guarantees*

$$\sum_{t=1}^T \langle g_t, w_t - u_t \rangle \le O\left(\sqrt{(D^2 + DP_T)\,\bar{V}_T} + \sqrt{DP_T}\widehat{M} + P_T\widehat{M} + D\widehat{M}\left(\log_+ \frac{\bar{V}_T}{\gamma^2}\right) + \gamma(D + P_T)\right),$$

where $\bar{V}_T = \sum_{t=1}^T \|g_t - h_t\|^2$, $P_T = \sum_{t=2}^T \|u_t - u_{t-1}\|$ and $\widehat{M} = \max_{t \in [T]} \|g_t - h_t\|$. Moreover, the algorithm runs in $O(d \log t)$ time on iteration t.

Proof Denote by N the value of n at the beginning of round T. We partition the interval $[T]$ into N segments $\mathcal{I}_1, \ldots, \mathcal{I}_N$, and denote by $\mathcal{I}_n = [s_n, e_n]$, where $s_1 = 1, e_N = T$. Note that at the interval $\mathcal{I}_n$, there are exactly n base learners. Specifically, at round t when n is the current number of base learners, if we find $\bar{V}_{\mathcal{I}_n}/\gamma^2 > 2^n$, then we define $e_n = t$ and $s_{n+1} = t + 1$, and add one base learner in the next round $t + 1$. By definition, we have $\bar{V}_{\mathcal{I}_n} = \sum_{t \in \mathcal{I}_n} \|g_t - h_t\|^2$ for all $n \in [N]$. Moreover, the algorithm ensures that for all $n \in [N - 1]$:

$$\gamma^2 2^n < \bar{V}_{\mathcal{I}_n} = \sum_{t=s_n}^{e_n - 1} \|g_t - h_t\|^2 + \|g_{e_n} - h_{e_n}\|^2 \le \gamma^2 2^n + \widehat{M}^2,$$

where we define $\widehat{M} = \max_{t\in[T]} \|g_t - h_t\|$. We also have, without loss of generality, by assuming $N \geq 2$,

$$\gamma^2 2^{N-1} \leq \gamma^2(2^N - 2) = \gamma^2 \sum_{n=1}^{N-1} 2^n < \sum_{n=1}^{N-1} \bar{V}_{\mathcal{I}_n} \leq \sum_{t=1}^{T} \|g_t - h_t\|^2 = \bar{V}_T,$$

which implies $N \leq 1 + \log_2(\bar{V}_T/\gamma^2)$.
Define $i_\star \in \mathbb{N}$ that $2^{i_\star} D \leq \sqrt{5D^2 + 12DP_T} < 2^{i_\star+1} D$. Then decompose the linearized regret by:

$$\sum_{t=1}^{T} \langle g_t, w_t - u_t \rangle = \underbrace{\sum_{n=1}^{N} \sum_{t\in\mathcal{I}_n} \left\langle g_t, w_t - w_{t,\min\{n,i_\star\}} \right\rangle}_{\text{META-REG}} + \underbrace{\sum_{n=1}^{N} \sum_{t\in\mathcal{I}_n} \left\langle g_t, w_{t,\min\{n,i_\star\}} - u_t \right\rangle}_{\text{BASE-REG}}.$$

Consider the base regret BASE-REG. For $n < i_\star$ and $t \in \mathcal{I}_n$, the n-th base learner starts from $w_{s_n,n}$ and performs one-step variant of Optimistic OGD as in Lemma 13, applying which we have:

$$\begin{aligned}
\sum_{t\in\mathcal{I}_n} \langle g_t, w_{t,n} - u_t \rangle &\leq \frac{D^2 + 3DP_{\mathcal{I}_n}}{2\eta_{e_n+1,n}} + \sum_{t\in\mathcal{I}_n} \eta_{t+1,n} \|g_t - h_t\|^2 \\
&= \left(\frac{D^2 + 3DP_{\mathcal{I}_n}}{2 \cdot 2^n D} + 2 \cdot 2^n D \right) \sqrt{\bar{V}_{\mathcal{I}_n}} \\
&\leq \left(\frac{D^2 + 3DP_{\mathcal{I}_n}}{2 \cdot 2^n D} + 2 \cdot 2^n D \right) \left(\gamma\sqrt{2^n} + \widehat{M} \right) \\
&\leq \frac{\gamma(D + 3P_T)}{2\sqrt{2^n}} + \frac{(D + 3P_T)\widehat{M}}{2 \cdot 2^n} + 2 \cdot 2^n D \left(\gamma\sqrt{2^n} + \widehat{M} \right),
\end{aligned}$$

where in the first line we define $P_{\mathcal{I}_n} = \sum_{t=s_n}^{e_n-1} \|u_t - u_{t+1}\|$, in the second line we use the definition $\eta_{t+1,n} = \frac{2^n D}{\sqrt{\bar{V}_t}}$ and the inequality $\sum_{t=1}^{T} \frac{a_t}{\sqrt{\sum_{s=1}^{t} a_s}} \leq 2\sqrt{\sum_{t=1}^{T} a_t}$ for any positive sequence $\{a_t\}_{t=1}^T$, then use the condition $\bar{V}_{\mathcal{I}_n} \leq \gamma^2 2^n + \widehat{M}^2$ in the third line. Then

$$\begin{aligned}
\sum_{n=1}^{\min\{N,i_\star-1\}} \sum_{t\in\mathcal{I}_n} \langle g_t, w_{t,n} - u_t \rangle &\leq O\left((D + P_T)(\gamma + \widehat{M}) + 4 \sum_{n=1}^{\min\{N,i_\star-1\}} 2^n D \left(\gamma\sqrt{2^n} + \widehat{M} \right) \right) \\
&\leq O\Bigg((D + P_T)(\gamma + \widehat{M}) + 4\sqrt{5D^2 + 12DP_T} \sum_{n=1}^{\min\{N,i_\star-1\}} \gamma\sqrt{2^n} \\
&\qquad + 4\widehat{M} \sum_{t=1}^{\min\{N,i_\star-1\}} 2^n D \Bigg) \\
&\leq O\left((D + P_T)(\gamma + \widehat{M}) + \sqrt{(D^2 + DP_T)\,\bar{V}_T} + \sqrt{D^2 + DP_T}\widehat{M} \right),
\end{aligned}$$

where we use $\sum_{i=1}^{\infty} \frac{1}{\sqrt{2^i}} = O(1)$ and $\sum_{i=1}^{\infty} \frac{1}{2^i} = O(1)$ in the first line, and for $n < i_\star$ use $2^n D < 2^{i_\star} D \leq \sqrt{5D^2 + 12DP_T}$ in the second and third line, use $\sum_{n=1}^{N} \gamma\sqrt{2^n} \leq O(\gamma\sqrt{2^N}) \leq O(\sqrt{\bar{V}_T})$

and $\sum_{n=1}^{i_\star} 2^n D \leq O(2^{i_\star} D) \leq O(\sqrt{D^2 + DP_T})$ in the third line. Without loss of generality, if $N \geq i_\star$, then:

$$\begin{aligned}
\sum_{n=i_\star}^{N} \sum_{t \in \mathcal{I}_n} \left\langle g_t, w_{t,\min\{n,i_\star\}} - u_t \right\rangle &= \sum_{n=i_\star}^{N} \sum_{t \in \mathcal{I}_n} \langle g_t, w_{i_\star} - u_t \rangle = \sum_{t=s_{i_\star}}^{T} \langle g_t, w_{i_\star} - u_t \rangle \\
&\leq \frac{5D^2 + 12DP_T}{8\eta_{T+1,i_\star}} + \sum_{t=s_{i_\star}}^{T} 2\eta_{t+1,i_\star} \|g_t - h_t\|^2 \\
&\leq \left(\frac{5D^2 + 12DP_T}{8 \cdot 2^{i_\star} D} + 4 \cdot 2^{i_\star} D \right) \sqrt{\bar{V}_T} = O\left(\sqrt{(D^2 + DP_T)\, \bar{V}_T} \right),
\end{aligned}$$

where in the last line we use $2^{i_\star} D \leq \sqrt{5D^2 + 12DP_T} < 2^{i_\star+1} D$. Hence we bound base regret by

$$\textsc{Base-Reg} \leq O\left(\sqrt{(D^2 + DP_T)\, \bar{V}_T} + (D + P_T)(\gamma + \widehat{M}) + \sqrt{D^2 + DP_T}\widehat{M} \right).$$

For meta regret, at the interval $\mathcal{I}_n = [s_n, e_n]$, we restart an Optimistic Hedge algorithm as in Lemma 19 in domain $\Delta_{n-1} = \left\{ a \in \mathbb{R}^n_{\geq 0} : \sum_{i=1} a_i = 1 \right\}$, with $\ell_{t,i} = \langle g_t, w_{t,i} \rangle$, $m_{t+1,i} = \langle h_{t+1}, w_{t,i} \rangle$, and $\varepsilon_{t+1} = \frac{1}{D\sqrt{\sum_{s=s_n}^{t} \|g_s - h_s\|^2}}$, then:

$$\begin{aligned}
\sum_{t \in \mathcal{I}_n} \left\langle g_t, w_t - w_{t,\min\{n,i_\star\}} \right\rangle &= \sum_{t \in \mathcal{I}_n} \left\langle \ell_t, p_t - e_{\min\{n,i_\star\}} \right\rangle \\
&\leq \frac{\ln n}{\varepsilon_{e_n}} + \sum_{t \in \mathcal{I}_n} \langle \ell_t - m_t, p_t - p_{t+1} \rangle - \sum_{t \in \mathcal{I}_n} \frac{1}{2\varepsilon_{t-1}} \|p_t - p_{t+1}\|_1^2 \\
&\leq \frac{\ln n}{\varepsilon_{e_n}} + \sum_{t \in \mathcal{I}_n} \|\ell_t - m_t\|_\infty \|p_t - p_{t+1}\|_1 - \sum_{t \in \mathcal{I}_n} \frac{1}{2\varepsilon_{t+1}} \|p_t - p_{t+1}\|_1^2 \\
&\quad + \sum_{t \in \mathcal{I}_n} \left(\frac{1}{2\varepsilon_{t+1}} - \frac{1}{2\varepsilon_t} \right) \|p_t - p_{t+1}\|_1^2 \\
&\quad + \sum_{t \in \mathcal{I}_n} \left(\frac{1}{2\varepsilon_t} - \frac{1}{2\varepsilon_{t-1}} \right) \|p_t - p_{t+1}\|_1^2 \\
&\leq \frac{1 + \ln n}{\varepsilon_{e_n+1}} + \frac{1}{2} \sum_{t \in \mathcal{I}_n} \varepsilon_{t+1} \|\ell_t - m_t\|_\infty^2 \\
&\leq \frac{1 + \ln n}{\varepsilon_{e_n+1}} + \frac{D^2}{2} \sum_{t \in \mathcal{I}_n} \varepsilon_{t+1} \|g_t - h_t\|^2 \\
&\leq (2 + \ln n)\, D \sqrt{\bar{V}_{\mathcal{I}_n}}.
\end{aligned}$$

Moreover, when $n \in [N-1]$, we have $\bar{V}_{\mathcal{I}_n} \leq \gamma^2 2^n + \widehat{M}^2$ by definition. Then we have:

$$\textsc{Meta-Reg} = \sum_{n=1}^{N} \sum_{t \in \mathcal{I}_n} \left\langle g_t, w_t - w_{t,\min\{n,i_\star\}} \right\rangle \leq (2 + \ln N) D \left(\sum_{n=1}^{N-1} \sqrt{\bar{V}_{\mathcal{I}_n}} + \sqrt{\bar{V}_{\mathcal{I}_N}} \right)$$

$$\leq (2+\ln N)D\left(\sum_{n=1}^{N-1}\left(\gamma\sqrt{2^n}+\widehat{M}\right)+\sqrt{\bar{V}_T}\right)=O\left(D\sqrt{\bar{V}_T}+D\widehat{M}\left(\log_+\frac{\bar{V}_T}{\gamma^2}\right)\right),$$

where we use $N \leq 1+\log_2(\bar{V}_T/\gamma^2)$. Combining meta-regret and base-regret, we have:

$$\sum_{t=1}^{T}\langle g_t, w_t-u_t\rangle \leq O\left(\sqrt{(D^2+DP_T)\,\bar{V}_T}+\sqrt{DP_T\widehat{M}}+P_T\widehat{M}+D\widehat{M}\left(\log_+\frac{\bar{V}_T}{\gamma^2}\right)+\gamma(D+P_T)\right).$$

■

We provide for completeness the dynamic regret of optimistic OGD with a time-varying step-size. The result follows via a mild generalization of the usual mirror-descent argument.

Lemma 13 (Dynamic Regret of One-step Optimistic OGD) *Assume that the domain $\mathcal{W}$ is bounded with diameter $D = \sup_{x,y\in\mathcal{W}}\|x-y\|$. The one-step optimistic OGD (Joulani et al., 2020) that starts at $w_1 \in \mathcal{W}$ and updates using $w_{t+1} = \arg\min_{w\in\mathcal{W}}\langle g_t - h_t + h_{t+1}, w\rangle + \frac{1}{2\eta_{t+1}}\|w-w_t\|^2$ with non-increasing sequence $\eta_t > 0$ and $h_1 = \mathbf{0}$, guarantees*

$$\sum_{t=1}^{T}\langle g_t, w_t-u_t\rangle \leq \frac{D^2+3DP_T}{2\eta_{T+1}}+\sum_{t=1}^{T}\eta_{t+1}\|g_t-h_t\|^2$$

for any sequence $u_1,\ldots,u_T$ in $\mathcal{W}$, where $P_T = \sum_{t=2}^{T}\|u_t-u_{t-1}\|$ is the path-length.

Proof Let $\psi_t(w) = \frac{1}{2\eta_t}\|w\|^2$ for all t, and observe that $\mathcal{D}_{\psi_t}(x,y) = \frac{1}{2\eta_t}\|x-y\|^2$. As usual, we seek to apply the first-order optimality condition $w_{t+1} = \arg\min_{w\in\mathcal{W}}\langle g_t-h_t+h_{t+1}, w\rangle + \mathcal{D}_{\psi_{t+1}}(w,w_t)$, so we begin by exposing terms of the form $\langle g_t-h_t+h_{t+1}, w_{t+1}-u_{t+1}\rangle$:

$$\begin{aligned}\sum_{t=1}^{T}\langle g_t, w_t-u_t\rangle &= \sum_{t=1}^{T}\langle g_t-h_t, w_t-u_t\rangle + \sum_{t=1}^{T}\langle h_t, w_t-u_t\rangle\\ &= \sum_{t=1}^{T}\langle g_t-h_t, w_{t+1}-u_{t+1}\rangle + \sum_{t=1}^{T}\langle h_t, w_t-u_t\rangle\\ &\quad + \sum_{t=1}^{T}\langle g_t-h_t, w_t-w_{t+1}\rangle + \langle g_t-h_t, u_{t+1}-u_t\rangle\\ &= \underbrace{\sum_{t=1}^{T}\langle g_t-h_t+h_{t+1}, w_{t+1}-u_{t+1}\rangle}_{\text{(A)}}\\ &\quad + \underbrace{\sum_{t=1}^{T}\langle h_t, w_t-u_t\rangle - \langle h_{t+1}, w_{t+1}-u_{t+1}\rangle}_{\text{(B)}}\\ &\quad + \sum_{t=1}^{T}\langle g_t-h_t, w_t-w_{t+1}\rangle + \langle g_t-h_t, u_{t+1}-u_t\rangle\end{aligned}$$

$$\tag{19}$$

where we've defined an arbitrary $u_{T+1} \in \mathcal{W}$, which we may set to u_T without loss of generality. Now by the first-order optimality condition $w_{t+1} = \arg\min_{w\in\mathcal{W}} \langle g_t - h_t + h_{t+1}, w\rangle + \mathcal{D}_{\psi_{t+1}}(w, w_t)$, we have that for any $w \in \mathcal{W}$ that

$$\langle g_t - h_t + h_{t+1} + \nabla\psi_{t+1}(w_{t+1}) - \nabla\psi_{t+1}(w_t), w_{t+1} - w\rangle \leq 0,$$

hence, we can bound

$$\begin{aligned}
\text{Ⓐ} &= \sum_{t=1}^{T} \langle g_t - h_t + h_{t+1}, w_{t+1} - u_{t+1}\rangle \\
&\leq \sum_{t=1}^{T} \langle \nabla\psi_{t+1}(w_t) - \nabla\psi_{t+1}(w_{t+1}), w_{t+1} - u_{t+1}\rangle \\
&= \sum_{t=1}^{T} \mathcal{D}_{\psi_{t+1}}(u_{t+1}, w_t) - \mathcal{D}_{\psi_{t+1}}(u_{t+1}, w_{t+1}) - \mathcal{D}_{\psi_{t+1}}(w_{t+1}, w_t)
\end{aligned}$$

where the last line uses the three-point relation for Bregman divergences, $\langle\nabla f(x) - \nabla f(x'), x' - u\rangle = \mathcal{D}_f(u, x) - \mathcal{D}_f(u, x') - \mathcal{D}_f(x', x)$. Meanwhile, the terms Ⓑ telescope to zero:

$$\begin{aligned}
\text{Ⓑ} &= \sum_{t=1}^{T} \langle h_t, w_t - u_t\rangle - \langle h_{t+1}, w_{t+1} - u_{t+1}\rangle \\
&= \langle h_1, w_1 - u_1\rangle - \langle h_{T+1}, w_{T+1} - u_{T+1}\rangle = 0
\end{aligned}$$

where we've used $h_1 = \mathbf{0}$ and observed that we may set $h_{T+1} = \mathbf{0}$ without loss of generality, since the regret does not depend on h_{T+1}. Plugging the previous two displays back into Eq. (19) and re-arranging terms, we have

$$\begin{aligned}
\sum_{t=1}^{T} \langle g_t, w_t - u_t\rangle &= \underbrace{\sum_{t=1}^{T} \mathcal{D}_{\psi_{t+1}}(u_{t+1}, w_t) - \mathcal{D}_{\psi_{t+1}}(u_{t+1}, w_{t+1})}_{\Omega_T} \\
&\quad + \sum_{t=1}^{T} \langle g_t - h_t, w_t - w_{t+1}\rangle - \mathcal{D}_{\psi_{t+1}}(w_{t+1}, w_t) + \langle g_t - h_t, u_{t+1} - u_t\rangle \\
&\leq \Omega_T + \sum_{t=1}^{T} \eta_{t+1} \|g_t - h_t\|^2 + \sum_{t=1}^{T} \frac{1}{2\eta_{t+1}} \|u_t - u_{t+1}\|^2 \qquad (20) \\
&\leq \Omega_T + \sum_{t=1}^{T} \eta_{t+1} \|g_t - h_t\|^2 + \frac{DP_T}{2\eta_{T+1}},
\end{aligned}$$

where the second-to-last line applies Fenchel-young inequality twice to bound $\langle g_t - h_t, \Delta\rangle \leq \frac{\eta_{t+1}}{2} \|g_t - h_t\|^2 + \frac{1}{2\eta_{t+1}} \|\Delta\|^2$ for $\Delta = w_t - w_{t+1}$ and for $\Delta = u_{t+1} - u_t$, while the last line bounds $\sum_{t=2}^{T} \frac{1}{2\eta_{t+1}} \|u_t - u_{t+1}\|^2 \leq \frac{DP_T}{2\eta_{T+1}}$ for non-increasing sequence $(\eta_t)_t$ and $u_{T+1} = u_T$. Finally, the terms Ω_T can be bound as

$$\begin{aligned}
\Omega_T &= \sum_{t=1}^{T} \mathcal{D}_{\psi_{t+1}}(u_{t+1}, w_t) - \mathcal{D}_{\psi_{t+1}}(u_{t+1}, w_{t+1}) \\
&\leq \sum_{t=1}^{T} \mathcal{D}_{\psi_t}(u_{t+1}, w_t) - \mathcal{D}_{\psi_{t+1}}(u_{t+1}, w_{t+1}) + \sum_{t=1}^{T} \mathcal{D}_{\psi_{t+1}-\psi_t}(u_{t+1}, w_t) \\
&\leq \sum_{t=1}^{T} \mathcal{D}_{\psi_t}(u_{t+1}, w_t) - \mathcal{D}_{\psi_{t+1}}(u_{t+1}, w_{t+1}) + D^2 \sum_{t=1}^{T} \left(\frac{1}{2\eta_{t+1}} - \frac{1}{2\eta_t} \right) \\
&\leq \frac{D^2}{2\eta_1} + \sum_{t=2}^{T} \left(\mathcal{D}_{\psi_t}(u_{t+1}, w_t) - \mathcal{D}_{\psi_t}(u_t, w_t) \right) + D^2 \left(\frac{1}{2\eta_{T+1}} - \frac{1}{2\eta_1} \right) \\
&\leq \frac{D^2}{2\eta_{T+1}} + \sum_{t=2}^{T} \frac{1}{\eta_t} \langle u_{t+1} - w_t, u_{t+1} - u_t \rangle \leq \frac{D^2}{2\eta_{T+1}} + \frac{DP_T}{\eta_{T+1}},
\end{aligned}$$

where the last line follows via $\frac{1}{2} \|x - y\|^2 - \frac{1}{2} \|x - z\|^2 \leq \langle x - y, z - y \rangle$ via convexity of $w \mapsto \frac{1}{2} \|w\|^2$. Plugging this back into the previous display yields

$$\sum_{t=1}^{T} \langle g_t, w_t - u_t \rangle \leq \frac{D^2 + 3DP_T}{2\eta_{T+1}} + \sum_{t=1}^{T} \eta_{t+1} \|g_t - h_t\|^2$$

■

We also provide a slightly different version of this result which exposes an additional negative term (the so-called *RVU property* (Syrgkanis et al., 2015)). We were unable to find an explicit statement of this result in the literature (though it is easily derived from, e.g., Zhao et al. (2025b, Theorem 1)), so we provide it here for posterity.

Corollary 14 *Under the same assumptions as Lemma 13, for any sequence $u_1, \ldots, u_T$ it holds that*

$$R_T(u_1, \ldots, u_T) \leq \frac{D^2 + 3DP_T}{2\eta_{T+1}} + \sum_{t=1}^{T} \frac{3\eta_{t+1}}{2} \|g_t - h_t\|^2 - \frac{1}{4\eta_{t+1}} \|w_t - w_{t+1}\|^2 .$$

Proof The result follows via the same steps as the proof of Lemma 13, but in line Eq. (20) applies Fenchel-Young inequality as $\langle g_t - h_t, w_t - w_{t+1} \rangle \leq \frac{\rho}{2} \|g_t - h_t\|^2 + \frac{1}{2\rho} \|w_t - w_{t+1}\|^2$ with $\rho = 2\eta_{t+1}$, so that

$$\begin{aligned}
&\langle g_t - h_t, w_t - w_{t+1} \rangle - \mathcal{D}_{\psi_{t+1}}(w_{t+1}, w_t) \\
&\quad \leq \eta_{t+1} \|g_t - h_t\|^2 + \frac{1}{4\eta_{t+1}} \|w_t - w_{t+1}\|^2 - \frac{1}{2\eta_{t+1}} \|w_{t+1} - w_t\|^2 \\
&\quad = \eta_{t+1} \|g_t - h_t\|^2 - \frac{1}{4\eta_{t+1}} \|w_t - w_{t+1}\|^2 ,
\end{aligned}$$

which leads to the stated result by following the same steps thereafter. ■

C.3. Proof of Theorem 7

In this section we prove our main result for the SEA model. The theorem is re-stated below for convenience.

Theorem 7 *In the SEA model, for any oblivious comparator sequence* $u_1, \ldots, u_T \in \mathbb{R}^d$ *such that* $\max_t \|u_t\| \le M$ *with* $M > 0$*, Algorithm 3 guarantees* $\mathbb{E}\left[\mathrm{REG}_T(u_{1:T})\right]$ *bounded by*

$$\widetilde{O}\Bigg(\sqrt{(M^2 + MP_T)\left(\sigma^2_{1:t}(u_{1:T}) + \Sigma^2_{1:T}(u_{1:T})\right)\left(\log_+ \tfrac{MG\sqrt{T}}{\epsilon\gamma}\right)} \\ + (LM + G + \gamma)P_T + \left(L + \tfrac{\gamma}{\epsilon} + \tfrac{\gamma^3}{\epsilon G^2}\right)M^2 + \gamma M + \epsilon G\left(\tfrac{G^3}{\gamma^3} + \tfrac{G}{\gamma} + 1\right) + \epsilon\gamma\Bigg).$$

where $\sigma^2_{1:T}(u_{1:T}) \triangleq \sum_{t=2}^T (\sigma_t^2(u_{t-1}) + \sigma^2_{t-1}(u_{t-1}))$*, and* $\Sigma^2_{1:T}(u_{1:T}) \triangleq \sum_{t=2}^T \Sigma_t^2(u_{t-1})$*. Moreover, the algorithm runs in* $O(d \log t)$ *time on iteration* t*.*

Proof The gradient variations can be decomposed as:

$$\begin{aligned} \|\nabla f_t(u_{t-1}) - \nabla f_{t-1}(u_{t-1})\|^2 &= \|\nabla f_t(u_{t-1}) - \nabla F_t(u_{t-1}) + \nabla F_t(u_{t-1}) - \nabla F_{t-1}(u_{t-1}) \\ &\quad + \nabla F_{t-1}(u_{t-1}) - \nabla f_{t-1}(u_{t-1})\|^2 \\ &\le 3\|\nabla F_t(u_{t-1}) - \nabla F_{t-1}(u_{t-1})\|^2 + 3\|\nabla f_t(u_{t-1}) - \nabla F_t(u_{t-1})\|^2 \\ &\quad + 3\|\nabla f_{t-1}(u_{t-1}) - \nabla F_{t-1}(u_{t-1})\|^2 . \end{aligned}$$

By definition (Eq. (11)) we have:

$$\mathbb{E}\left[\|\nabla f_t(u_{t-1}) - \nabla f_{t-1}(u_{t-1})\|^2\right] \le 3\Sigma_t^2(u_{t-1}) + 3\sigma_t^2(u_{t-1}) + 3\sigma^2_{t-1}(u_{t-1}).$$

Therefore, applying Theorem 6 and taking expectations, an application of Jensen's inequality yields

$$\begin{aligned} &\mathbb{E}\left[\sqrt{(M^2 + MP_T)V_T(u_{1:T})\left(\log_+ \frac{MG\sqrt{T}}{\epsilon\gamma}\right)}\right] \\ &\le \sqrt{\mathbb{E}\left[(M^2 + MP_T)V_T(u_{1:T})\left(\log_+ \frac{MG\sqrt{T}}{\epsilon\gamma}\right)\right]} \\ &= \sqrt{(M^2 + MP_T)\left(\log_+ \frac{MG\sqrt{T}}{\epsilon\gamma}\right)\mathbb{E}\left[V_T(u_{1:T})\right]} \\ &\le \sqrt{(M^2 + MP_T)\left(\log_+ \frac{MG\sqrt{T}}{\epsilon\gamma}\right)(3\Sigma^2_{1:T} + 3\sigma^2_{1:t})}, \end{aligned}$$

where we define $\Sigma^2_{1:T} \triangleq \sum_{t=2}^T \Sigma_t^2(u_{t-1})$, and $\sigma^2_{1:t} \triangleq \sum_{t=2}^T (\sigma_t^2(u_{t-1}) + \sigma^2_{t-1}(u_{t-1}))$. ■

Appendix D. Supporting Lemmas

In this section, we collect various supporting lemmas that are used throughout. The results in this section are mostly used to simplify calculations or restate existing results for completeness.

The following lemma provides a formula that relates the empirical gradient variation to $V_T(u_{1:T})$.

Lemma 15 *Let $f_1, \ldots, f_T$ be an arbitrary sequence of L-smooth functions. Then*

$$\sum_{t=1}^{T} \|\nabla f_t(w_t) - \nabla f_{t-1}(w_{t-1})\|^2 \leq 16L \sum_{t=1}^{T} \mathcal{D}_{f_t}(u_t, w_t)$$
$$+ \min\left\{\|\nabla f_1(w_1)\|^2 + 4V_T(u_{1:T}) + 4L^2 P_T^{\|\cdot\|^2}(u_{1:T}), 16LF_T(u_{1:T})\right\}$$

where we denote $V_T(u_{1:T}) \triangleq \sum_{t=2}^{T} \|\nabla f_t(u_{t-1}) - \nabla f_{t-1}(u_{t-1})\|^2$, $F_T(u_{1:T}) \triangleq \sum_{t=1}^{T}(f_t(u_t) - \inf_{w\in\mathbb{R}^d} f_t(w))$, and $P_T^{\|\cdot\|^2}(u_{1:T}) \triangleq \sum_{t=2}^{T} \|u_t - u_{t-1}\|^2$,

Proof For brevity denote $g_t(w) = \nabla f_t(w)$ and define $f_0(\cdot) \triangleq 0$.

Bounding via Gradient Variation. For all $t \geq 2$,

$$\begin{aligned}\|g_t(w_t) - g_{t-1}(w_{t-1})\| &= \|g_t(w_t) - g_t(u_t) + g_t(u_t) - g_t(u_{t-1}) \\ &\quad + g_t(u_{t-1}) - g_{t-1}(u_{t-1}) + g_{t-1}(u_{t-1}) - g_{t-1}(w_{t-1})\| \\ &\leq \|g_t(w_t) - g_t(u_t)\| + \|g_{t-1}(w_{t-1}) - g_{t-1}(u_{t-1})\| \\ &\quad + \|g_t(u_t) - g_t(u_{t-1})\| + \|g_t(u_{t-1}) - g_{t-1}(u_{t-1})\|.\end{aligned}$$

Then using Cauchy-Schwarz inequality $(\sum_{i=1}^{n} a_i)^2 \leq \left(\sqrt{n\sum_{i=1}^{n} a_i^2}\right)^2 = n\sum_{i=1}^{n} a_i^2$, we can bound $\sum_{t=1}^{T} \|g_t(w_t) - g_{t-1}(w_{t-1})\|^2$ by

$$\begin{aligned}&\sum_{t=1}^{T} \|g_t(w_t) - g_{t-1}(w_{t-1})\|^2 \\ &\leq \|g_1(w_1)\|^2 + 4\sum_{t=2}^{T}\Bigg[\|g_t(w_t) - g_t(u_t)\|^2 + \|g_{t-1}(w_{t-1}) - g_{t-1}(u_{t-1})\|^2 \\ &\qquad\qquad + \|g_t(u_t) - g_t(u_{t-1})\|^2 + \|g_t(u_{t-1}) - g_{t-1}(u_{t-1})\|^2\Bigg] \\ &\leq \|g_1(w_1)\|^2 + 16L\sum_{t=1}^{T}\mathcal{D}_{f_t}(u_t, w_t) + 4\sum_{t=2}^{T}\|g_t(u_t) - g_t(u_{t-1})\|^2 \\ &\quad + 4\sum_{t=2}^{T}\|g_t(u_{t-1}) - g_{t-1}(u_{t-1})\|^2 \\ &\leq \|g_1(w_1)\|^2 + 16L\sum_{t=1}^{T}\mathcal{D}_{f_t}(u_t, w_t) + 4L^2\sum_{t=2}^{T}\|u_t - u_{t-1}\|^2 + 4V_T(u_{1:T})\end{aligned}$$

$$= \|g_1(w_1)\|^2 + 16L \sum_{t=1}^{T} \mathcal{D}_{f_t}(u_t, w_t) + 4L^2 P_T^{\|\cdot\|^2}(u_{1:T}) + 4V_T(u_{1:T}). \tag{21}$$

where the second inequality uses $\|\nabla f_t(x) - \nabla f_t(y)\|^2 \leq 2L\mathcal{D}_{f_t}(y, x)$ for L-smooth functions (Lemma 16), the last inequality uses $\|\nabla f_t(x) - \nabla f_t(y)\| \leq L\|x - y\|$, and we've defined the $V_T(u_{1:T}) \triangleq \sum_{t=2}^{T} \|\nabla f_t(u_{t-1}) - \nabla f_{t-1}(u_{t-1})\|^2$ and $P_T^{\|\cdot\|^2}(u_{1:T}) \triangleq \sum_{t=2}^{T} \|u_t - u_{t-1}\|^2$

Bounding via Comparator Loss. Again, for all $t \geq 2$:

$$\begin{aligned}\|g_t(w_t) - g_{t-1}(w_{t-1})\| &= \|g_t(w_t) - g_t(u_t) + g_t(u_t) - g_{t-1}(u_{t-1}) + g_{t-1}(u_{t-1}) - g_{t-1}(w_{t-1})\| \\ &\leq \|g_t(w_t) - g_t(u_t)\| + \|g_{t-1}(w_{t-1}) - g_{t-1}(u_{t-1})\| + \|g_t(u_t)\| + \|g_{t-1}(u_{t-1})\|.\end{aligned}$$

And for $t = 1$, we bound $\|g_1(w_1)\| \leq \|g_1(w_1) - g_1(u_1)\| + \|g_1(u_1)\|$. Hence, we can bound

$$\begin{aligned}\sum_{t=1}^{T} \|g_t(w_t) - g_{t-1}(w_{t-1})\|^2 &\leq \|g_1(w_1)\|^2 + 4\sum_{t=1}^{T}\left[\|g_t(u_t)\|^2 + \|g_{t-1}(u_{t-1})\|^2\right] \\ &\quad + 4\sum_{t=2}^{T}\left[\|g_t(w_t) - g_t(u_t)\|^2 + \|g_{t-1}(w_{t-1}) - g_{t-1}(u_{t-1})\|^2\right] \\ &\leq 8\sum_{t=1}^{T}\|g_t(u_t)\|^2 + 16L\sum_{t=1}^{T}\mathcal{D}_{f_t}(u_t, w_t) \\ &\leq 16L\sum_{t=1}^{T}\left(f_t(u_t) - \inf_{w\in\mathbb{R}^d} f_t(w)\right) + 16L\sum_{t=1}^{T}\mathcal{D}_{f_t}(u_t, w_t) \\ &= 16LF_T(u_{1:T}) + 16L\sum_{t=1}^{T}\mathcal{D}_{f_t}(u_t, w_t),\end{aligned} \tag{22}$$

where the second inequality uses $\|\nabla f_t(x) - \nabla f_t(y)\|^2 \leq 2L\mathcal{D}_{f_t}(y, x)$ for L-smooth functions (Lemma 16). The last inequality uses the self-bounding property of smooth functions (Lemma 17), and the last line defines $F_T(u_{1:T}) \triangleq \sum_{t=1}^{T}(f_t(u_t) - \inf_{w\in\mathbb{R}^d} f_t(w))$. Combining Eq. (21) and Eq. (22) concludes the proof. ■

The following lemma provides a crucial bound on the difference between the gradients of smooth functions in terms of the Bregman divergence.

Lemma 16 (Theorem 2.1.5 of Nesterov (2018)) *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a L-smooth function, then for all $x, y \in \mathbb{R}^d$,*

$$\|\nabla f(x) - \nabla f(y)\|^2 \leq 2L\mathcal{D}_f(x, y).$$

The following lemma provides the well-known self-bounding property of smooth functions (see, e.g., Levy (2017)).

Lemma 17 *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a L-smooth function. Then for all $x \in \mathbb{R}^d$,*

$$\|\nabla f(x)\|^2 \leq 2L\left(f(x) - \inf_{z\in\mathbb{R}^d} f(z)\right).$$

Proof By smoothness, we have, for all $x \in \mathbb{R}^d$:

$$f(x - \frac{1}{L}\nabla f(x)) \le f(x) + \left\langle \nabla f(x), -\frac{1}{L}\nabla f(x) \right\rangle + \frac{1}{2L} \|\nabla f(x)\|^2 = f(x) - \frac{1}{2L} \|\nabla f(x)\|^2 .$$

Rearranging, we have

$$\|\nabla f(x)\|^2 \le 2L(f(x) - f(x - \frac{1}{L}\nabla f(x))) \le 2L(f(x) - \inf_{z\in\mathbb{R}^d} f(z)).$$

■

Finally, we borrow the following lemmas from existing works, included here for completeness.

Lemma 18 (Adapted from Lemma 21 of Cutkosky and Mhammedi (2024)) *Suppose* $g_1, \ldots, g_T$ *and* $0 < h_1 \le h_2 \le \ldots \le h_T$ *are such that* $\|g_t\| \le h_t$ *for all* t. *Define* $\bar{V}_t = 4h_t^2 + \sum_{i=1}^{t-1} \|g_i\|^2$, $B_t = 4 + \sum_{i=1}^{t-1} \|g_i\|^2/h_i^2$, *and* $\alpha_t = \frac{\epsilon}{\sqrt{B_t}\log^2(B_t)}$, *then:*

$$\sum_{t=1}^{T} \frac{\alpha_t \|g_t\|^2}{\sqrt{\bar{V}_t}} \le 4\epsilon h_T.$$

Lemma 19 (Theorem 7.36 of Orabona (2019)) *The Optimistic Hedge algorithm, which starts from* $p_1 = \frac{1}{N}\mathbf{1}_N$ *and updates* $p_{t+1} \in \Delta_N$ *by* $p_{t+1,i} \propto \exp(-\varepsilon_{t+1}(\sum_{s=1}^{t} \ell_t + m_{t+1}))$ *with a time-varying and non-increasing learning rate* $\varepsilon_t > 0$, *ensures that for any expert* $i \in [N]$*:*

$$\sum_{t=1}^{T} \langle \ell_t, p_t - e_i \rangle \le \frac{\ln N}{\varepsilon_T} + \sum_{t=1}^{T} \langle \ell_t - m_t, p_t - p_{t+1} \rangle - \sum_{t=1}^{T} \frac{1}{2\varepsilon_{t-1}} \|p_t - p_{t+1}\|_1^2 .$$