Efficient Methods for Non-stationary Online Learning

Peng Zhao, Yan-Feng Xie, Lijun Zhang, Zhi-Hua Zhou
National Key Laboratory for Novel Software Technology,
Nanjing University, Nanjing, China
{zhaop, xieyf, zhanglj, zhouzh}@lamda.nju.edu.cn

Abstract

Non-stationary online learning has drawn much attention in recent years. In particular, *dynamic regret* and *adaptive regret* are proposed as two principled performance measures for online convex optimization in non-stationary environments. To optimize them, a two-layer online ensemble is usually deployed due to the inherent uncertainty of the non-stationarity, in which a group of base-learners are maintained and a meta-algorithm is employed to track the best one on the fly. However, the two-layer structure raises the concern about the computational complexity — those methods typically maintain $\mathcal{O}(\log T)$ base-learners simultaneously for a T-round online game and thus perform multiple projections onto the feasible domain per round, which becomes the computational bottleneck when the domain is complicated. In this paper, we present efficient methods for optimizing dynamic regret and adaptive regret, which reduce the number of projections per round from $\mathcal{O}(\log T)$ to 1. Moreover, our obtained algorithms require only one gradient query and one function evaluation at each round. Our technique hinges on the reduction mechanism developed in parameter-free online learning and requires non-trivial twists on non-stationary online methods. Empirical studies verify our theoretical findings.

1 Introduction

Classic online learning minimizes the static regret, which benchmarks the online learner's performance against the best fixed decision in hindsight. In many real-world online applications, however, the environments are non-stationary [Zhou, 2022] and static regret becomes less attractive since it would be unrealistic to assume the existence of a single decision behaved satisfactorily over time.

To address the limitation, in recent years, researchers have studied more strengthened performance measures to facilitate online algorithms with the capability of handling non-stationarity. In particular, dynamic regret [Zinkevich, 2003; Zhang et al., 2018a] and adaptive regret [Hazan and Seshadhri, 2009; Daniely et al., 2015] are proposed as two principled metrics to guide the algorithm design. We focus on the online convex optimization (OCO) setting [Hazan, 2016]. OCO can be deemed as a game between the learner and the environments. At each round $t \in [T]$, the learner submits her decision $\mathbf{x}_t \in \mathcal{X}$ from a convex feasible domain $\mathcal{X} \subseteq \mathbb{R}^d$ and simultaneously environments choose a convex function $f_t : \mathcal{X} \mapsto \mathbb{R}$, and subsequently the learner suffers an instantaneous loss $f_t(\mathbf{x}_t)$.

1.1 Dynamic Regret and Adaptive Regret

Dynamic regret is proposed by Zinkevich [2003] to compare the online learner's performance against a sequence of *any* feasible comparators $\mathbf{u}_1, \ldots, \mathbf{u}_T \in \mathcal{X}$. Formally, it is defined as

$$\text{D-REG}_T(\mathbf{u}_1, \ldots, \mathbf{u}_T) = \sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{u}_t). \tag{1}$$

36th Conference on Neural Information Processing Systems (NeurIPS 2022).

Dynamic regret minimization enables the learner to track changing comparators. A favorable dynamic regret bound should scale with a certain non-stationarity measure dependent on the comparators such as the path length $P_T = \sum_{t=2}\|\mathbf{u}_t - \mathbf{u}_{t-1}\|_2$. Notably, the classic static regret can be treated as a special case of dynamic regret by specifying the comparators as the best fixed decision in hindsight.

Adaptive regret is proposed by Hazan and Seshadhri [2009] and further strengthened by Daniely et al. [2015], which measures the regret over *any* interval $I = [r, s] \subseteq [T]$ with a length of $\tau = |I|$, i.e.,

$$\text{A-REG}_T(|I|) = \max_{[r,r+\tau-1]\subseteq[T]} \left\{ \sum_{t=r}^{r+\tau-1} f_t(\mathbf{x}_t) - \min_{\mathbf{u}\in\mathcal{X}} \sum_{t=r}^{r+\tau-1} f_t(\mathbf{u}) \right\}. \tag{2}$$

Since the minimizers of different intervals can be different, adaptive regret minimization also ensures the capability of competing with changing comparators. A desired adaptive regret bound should be as close as the minimax static regret of this interval. Algorithms with adaptive regret matching static regret of this interval up to logarithmic terms in T are referred to strongly adaptive [Daniely et al., 2015]. Moreover, adaptive regret includes static regret when choosing $I = [T]$.

It is worth noting that the relationship between dynamic regret and adaptive regret for OCO is generally unclear [Zhang, 2020, Section 5], even though a black-box reduction from dynamic regret to adaptive regret has been proven for the simper expert setting (i.e., online linear optimization over simplex) [Luo and Schapire, 2015, Theorem 4]. Hence, the two measures are separately developed and many algorithms have been proposed, including algorithms for dynamic regret [Zinkevich, 2003; Hall and Willett, 2013; Zhang et al., 2018a; Zhao et al., 2020, 2021b,a; Baby and Wang, 2021; Zhao et al., 2022a] and the ones for adaptive regret [Hazan and Seshadhri, 2009; Daniely et al., 2015; Jun et al., 2017; Zhang et al., 2018b, 2019]. Note that there are also studies [Zhang et al., 2020; Cutkosky, 2020] optimizing both measures simultaneously by an even strengthened metric $\sum_{t=r}^{s} f_t(\mathbf{x}_t) - \sum_{t=r}^{s} f_t(\mathbf{u}_t)$ over any interval $[r, s] \subseteq [T]$, hence called "interval dynamic regret".

1.2 Two-layer Structure and Projection Complexity Issue

The fundamental challenge of optimizing these two non-stationary regret measures is the uncertainty of the environmental non-stationarity. Concretely, to ensure the robustness to the unknown environments, dynamic regret aims to compete with *any* feasible comparator sequence, while adaptive regret examines the local performance over *any* intervals. The unknown comparators or unknown intervals bring considerable uncertainty to online optimization. To address the issue, a two-layer structure is usually deployed to optimize the measures, where a set of base-learners are maintained to handle the different possibilities of online environments and a meta-algorithm is employed to combine them all and track the unknown best one. Such a framework successfully achieves many state-of-the-art results, including the $\mathcal{O}(\sqrt{T(1+P_T)})$ dynamic regret [Zhang et al., 2018a] and the $\mathcal{O}(\sqrt{(F_T+P_T)(1+P_T)})$ small-loss dynamic regret for smooth functions [Zhao et al., 2020], where $P_T = \sum_{t=2}^{T}\|\mathbf{u}_t - \mathbf{u}_{t-1}\|_2$ is the path length and $F_T = \sum_{t=1}^{T} f_t(\mathbf{u}_t)$ is the cumulative loss of comparators; as well as the $\mathcal{O}(\sqrt{|I|\log T})$ adaptive regret [Jun et al., 2017] and the $\mathcal{O}(\sqrt{F_I \log F_I \log F_T})$ small-loss adaptive regret for smooth functions [Zhang et al., 2019] for any interval $I = [r, s] \subseteq [T]$, where $F_I = \min_{\mathbf{x}\in\mathcal{X}} \sum_{t=r}^{s} f_t(\mathbf{x})$ and $F_T = \min_{\mathbf{x}\in\mathcal{X}} \sum_{t=1}^{T} f_t(\mathbf{x})$. Besides, an $\mathcal{O}(\sqrt{|I|(\log T + P_I)})$ interval dynamic regret is also achieved by a two-layer (or even three-layer) structure [Zhang et al., 2020], where $P_I = \sum_{t=r}^{s}\|\mathbf{u}_t - \mathbf{u}_{t-1}\|_2$ is the path length over the interval.

The two-layer methods have demonstrated great effectiveness in tackling non-stationary online environments, whereas the gain is at the price of heavier computations than the methods for minimizing static regret. While it is believed that additional computations are necessary for more robustness, we are wondering whether it is possible to pay for a "minimal" computation overhead for adapting to the non-stationarity. To this end, we focus on the popular first-order online methods and aim to streamline unnecessary computations while retaining the same regret guarantees. Arguably, the most computationally expensive step of each round is the projection onto the convex feasible domain, namely, the projection operation $\Pi_{\mathcal{X}}[\mathbf{y}] = \arg\min_{\mathbf{x}\in\mathcal{X}}\|\mathbf{x} - \mathbf{y}\|_2$ for a convex set $\mathcal{X} \subseteq \mathbb{R}^d$. Typical two-layer non-stationary online algorithms require maintaining $N = \mathcal{O}(\log T)$ base-learners simultaneously to cover the possibility of unknown environments. Define the *projection complexity* of online methods as the number of projections onto the feasible domain per round. Then, those non-stationary methods suffer an $\mathcal{O}(\log T)$ projection complexity, whereas standard online methods for static regret minimization require only one projection per round such as online gradient descent [Zinkevich, 2003].

1.3 Our Contributions and Techniques

In this paper, we design a generic mechanism to reduce the projection complexity of many existing non-stationary methods from $\mathcal{O}(\log T)$ to 1 *without sacrificing the regret optimality*, hence matching the projection complexity of stationary methods. Our reduction is inspired by the recent advance in parameter-free online learning [Cutkosky and Orabona, 2018; Mhammedi et al., 2019]. The idea is simple: we reduce the original problem learned in the feasible domain $\mathcal{X}$ to an alternative one learned in a *surrogate domain* $\mathcal{Y} \supseteq \mathcal{X}$ such that the projection onto it is much cheaper, e.g., simply choosing $\mathcal{Y}$ as a properly scaled Euclidean ball; and moreover, a carefully designed *surrogate loss* is necessary for the alternative problem to retain the regret optimality. We reveal that a necessary condition for our reduction mechanism to deploy and reduce the projection complexity is that the non-stationary online algorithm shall *query the function gradient once and evaluate the function value once per round*. Several algorithms for the worst-case dynamic regret or adaptive regret already satisfy the requirements, so we can immediately deploy the reduction and obtain their efficient counterparts with the same regret guarantees and 1 projection complexity. However, many non-stationary algorithms, particularly those designed for small-loss bounds, do not satisfy the requirement. Hence, we require non-trivial efforts to make them compatible. Due to this, we have developed a series of algorithms that achieve worst-case/small-loss dynamic regret and adaptive regret with one projection per round (actually, with one gradient query and one function evaluation per round as well).

Despite that the reduction mechanism of this paper has been studied in parameter-free online learning, applying it to non-stationary online learning requires new ideas and non-trivial modifications. Here we highlight the technical innovation. The main challenge comes from the reduction condition mentioned earlier — as the surrogate loss involves the projection operation, our reduction requires the algorithm query one gradient and evaluate one function value at each round. However, many non-stationary algorithms do not satisfy the requirement, which is to be contrasted to the parameter-free algorithms such as MetaGrad [van Erven and Koolen, 2016; Mhammedi et al., 2019] that naturally satisfy the condition. For example, the SACS algorithm [Zhang et al., 2019] enjoys the best known small-loss adaptive regret, yet the method requires N gradient queries and $N+1$ function evaluations at each round, where $N = \mathcal{O}(\log T)$ is the number of base-learners. Thus, we have to dig into the algorithm and modify it to fit our reduction. First, we replace their meta-algorithm with Adapt-ML-Prod [Gaillard et al., 2014], an expert-tracking algorithm with a *second-order* regret with excess losses to accommodate the linearized loss that is used to ensure one gradient query per round. Second, we introduce a sequence of *time-varying* thresholds to adaptively determine the problem-dependent geometric covers in contrast to a fixed threshold used in their method. In particular, we register the cumulative loss of the final decisions rather than the base-learner's one to compare it with the changing thresholds, which renders the design of one function value evaluation per round and also turns out to be crucial for achieving an improved small-loss bound that can recover the best known worst-case adaptive adaptive regret (by contrast, SACS cannot obtain optimal worst-case adaptive regret). To summarize, our final algorithm only requires one projection/gradient query/function evaluation at each round, substantially improving the efficiency of SACS algorithm that requires N projections/gradient queries/function evaluations per round.

1.4 Assumptions

We list several standard assumptions used in OCO [Shalev-Shwartz, 2012; Hazan, 2016]. Notably, not all assumptions are always required. We will explicitly state the requirements in the theorem.

Assumption 1 (bounded gradient)**.** The norm of the gradients of online functions over the domain $\mathcal{X}$ is bounded by G, i.e., $\|\nabla f_t(\mathbf{x})\|_2 \le G$, for all $\mathbf{x} \in \mathcal{X}$ and $t \in [T]$.

Assumption 2 (bounded domain)**.** The domain $\mathcal{X} \subseteq \mathbb{R}^d$ contains the origin $\mathbf{0}$, and the diameter of the domain $\mathcal{X}$ is at most D, i.e., $\|\mathbf{x} - \mathbf{x}'\|_2 \le D$ for any $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$.

Assumption 3 (non-negativity and smoothness)**.** All the online functions are non-negative and L-smooth, i.e., for any $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$ and $t \in [T]$, $\|\nabla f_t(\mathbf{x}) - \nabla f_t(\mathbf{x}')\|_2 \le L\|\mathbf{x} - \mathbf{x}'\|_2$.

Organization. The rest is structured as follows. Section 2 presents the reduction mechanism and illustrates its application to dynamic regret minimization. Section 3 provides efficient methods for optimizing adaptive regret. Section 4 reports the experiments. Section 5 concludes the paper and makes discussions. All the proofs and omitted details for algorithms are deferred to the appendices.

2 The Reduction Mechanism and Dynamic Regret Minimization

We start from the dynamic regret minimization. First, we briefly review existing methods in Section 2.1, and then present our reduction mechanism and illustrate how to apply it to reducing the projection complexity of dynamic regret methods in Section 2.2.

2.1 A Brief Review of Dynamic Regret Minimization

Zhang et al. [2018a] propose a two-layer online algorithm called Ader with an $\mathcal{O}(\sqrt{T(1+P_T)})$ dynamic regret, which is proven to be minimax optimal for convex functions. Ader maintains a group of base-learners, each performing online gradient descent (OGD) [Zinkevich, 2003] with a customized step size specified by the pool $\mathcal{H} = \{\eta_1, \ldots, \eta_N\}$, and then uses a meta-algorithm to combine them all. Denoted by $\mathcal{B}_1, \ldots, \mathcal{B}_N$ the N base-learners. For each $i \in [N]$, $\mathcal{B}_i$ updates by

$$\mathbf{x}_{t+1,i} = \Pi_{\mathcal{X}}[\mathbf{x}_{t,i} - \eta_i \nabla f_t(\mathbf{x}_t)], \tag{3}$$

where $\eta_i \in \mathcal{H}$ is the associated step size and $\Pi_{\mathcal{X}}[\cdot]$ denotes the projection onto the feasible domain $\mathcal{X}$ with $\Pi_{\mathcal{X}}[\mathbf{y}] = \arg\min_{\mathbf{x}\in\mathcal{X}} \|\mathbf{y} - \mathbf{x}\|_2$. Notably, all the base-learners share the same gradient $\nabla f_t(\mathbf{x}_t)$ rather than using their individual one $\nabla f_t(\mathbf{x}_{t,i})$. This is because Ader optimizes the linearized loss $\ell_t(\mathbf{x}) = \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}\rangle$, which enjoys the benign property of $\nabla \ell_t(\mathbf{x}_{t,i}) = \nabla f_t(\mathbf{x}_t)$ for all $i \in [N]$.

Furthermore, the meta-algorithm evaluates each base-learner by $\ell_t(\mathbf{x}_{t,i}) = \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_{t,i}\rangle$ and updates the weight vector $\boldsymbol{p}_{t+1} \in \Delta_N$ by the Hedge algorithm [Freund and Schapire, 1997], namely,

$$p_{t+1,i} = \frac{p_{t,i} \exp(-\varepsilon \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_{t,i}\rangle)}{\sum_{j=1}^N p_{t,j} \exp(-\varepsilon \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_{t,j}\rangle)}, \quad \forall i \in [N], \tag{4}$$

where $\varepsilon > 0$ is the learning rate of the meta-algorithm. The final prediction is obtained by $\mathbf{x}_{t+1} = \sum_{i=1}^N p_{t+1,i} \mathbf{x}_{t+1,i}$. The learner submits the prediction $\mathbf{x}_{t+1}$ and then receives the loss $f_{t+1}(\mathbf{x}_{t+1})$ and the gradient $\nabla f_{t+1}(\mathbf{x}_{t+1})$ as the feedback of this round. Under a suitable configuration of the step size pool $\mathcal{H}$ with $N = \mathcal{O}(\log T)$ and learning rate $\varepsilon = \Theta(\sqrt{(\ln N)/T})$, Ader enjoys an $\mathcal{O}(\sqrt{T(1+P_T)})$ dynamic regret [Zhang et al., 2018a, Theorem 4].

For convex and smooth functions, Zhao et al. [2021b] demonstrate that a similar two-layer structure can attain an $\mathcal{O}(\sqrt{(F_T + P_T)(1+P_T)})$ small-loss dynamic regret under a suitable setting of the step size pool $\mathcal{H}$ and time-varying learning rates of meta-algorithm $\{\varepsilon_t\}_{t=1}^T$, where $F_T = \sum_{t=1}^T f_t(\mathbf{u}_t)$ is the cumulative loss of the comparators. This bound safeguards the minimax rate in the worst case, while can be much smaller than $\mathcal{O}(\sqrt{T(1+P_T)})$ bound in benign environments.

2.2 The Reduction Mechanism for Reducing Projection Complexity

As demonstrated in the update (3), all the base-learners require projecting the intermediate solution onto the domain $\mathcal{X}$ to ensure the feasibility. As a result, $\mathcal{O}(\log T)$ projections are required at each round, which is generally time-consuming particularly when the domain $\mathcal{X}$ is complicated.

We present a generic reduction mechanism for reducing the projection complexity and apply it to dynamic regret methods. Our reduction builds upon the seminal work [Cutkosky and Orabona, 2018] and a further refined result [Cutkosky, 2020], who propose a black-box reduction from constrained online learning to the unconstrained setting (or another constrained problem with a larger domain) .

Reduction mechanism. Given an algorithm for non-stationary online learning Algo whose projection complexity is $\mathcal{O}(\log T)$, our reduction mechanism builds on it to yield an algorithm Efficient-Algo with 1 projection onto $\mathcal{X}$ per round and retaining the same order of regret. The central idea is to replace expensive projections onto the original domain $\mathcal{X}$ with other much cheaper projections. To this end, we introduce a *surrogate domain* $\mathcal{Y}$ defined as the minimum Euclidean ball containing the feasible domain $\mathcal{X}$, i.e., $\mathcal{Y} = \{\mathbf{x} \mid \|\mathbf{x}\|_2 \le D\} \supseteq \mathcal{X}$. Then, the reduced algorithm Algo works on $\mathcal{Y}$ whose projection can be realized by a simple rescaling. More importantly, to avoid regret degeneration, it is necessary to carefully construct the surrogate loss $g_t : \mathcal{Y} \mapsto \mathbb{R}$ as

$$g_t(\mathbf{y}) = \langle \nabla f_t(\mathbf{x}_t), \mathbf{y}\rangle - \mathbb{1}_{\{\langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t\rangle < 0\}} \cdot \langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t\rangle \cdot S_{\mathcal{X}}(\mathbf{y}), \tag{5}$$

Algorithm 1 Efficient Algorithm for Minimizing Dynamic Regret

Input: step size pool $\mathcal{H} = \{\eta_1, \ldots, \eta_N\}$, learning rate of meta-algorithm ε_t (or simply a fixed ε).

1: Initialization: let $\mathbf{x}_1$ and $\{\mathbf{y}_{1,i}\}_{i=1}^N$ be any point in $\mathcal{X}$; $\forall i \in [N], p_{1,i} = 1/N$.

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

3: Receive the gradient information $\nabla f_t(\mathbf{x}_t)$.

4: Construct the surrogate loss $g_t : \mathcal{Y} \mapsto \mathbb{R}$ according to Eq. (5).

5: Compute the gradient $\nabla g_t(\mathbf{y}_t)$ according to Lemma 1.

6: For each $i \in [N]$, the base-learner $\mathcal{B}_i$ produces the local decision by

$$\widehat{\mathbf{y}}_{t+1,i} = \mathbf{y}_{t,i} - \eta_i \nabla g_t(\mathbf{y}_t), \;\; \mathbf{y}_{t+1,i} = \widehat{\mathbf{y}}_{t+1,i}\Big(\mathbb{1}_{\{\|\widehat{\mathbf{y}}_{t+1,i}\|_2 \le D\}} + \frac{D}{\|\widehat{\mathbf{y}}_{t+1,i}\|_2} \cdot \mathbb{1}_{\{\|\widehat{\mathbf{y}}_{t+1,i}\|_2 \ge D\}}\Big).$$

7: Meta-algorithm updates weight by $p_{t+1,i} \propto \exp(-\varepsilon_{t+1} \sum_{s=1}^t \langle \nabla g_s(\mathbf{y}_s), \mathbf{y}_{s,i} \rangle)$, $i \in [N]$.

8: Compute $\mathbf{y}_{t+1} = \sum_{i=1}^N p_{t+1,i} \mathbf{y}_{t+1,i}$.

9: Submit $\mathbf{x}_{t+1} = \Pi_{\mathcal{X}}[\mathbf{y}_{t+1}]$. ▷ The only step projects onto feasible domain $\mathcal{X}$ per round.

10: **end for**

where $S_{\mathcal{X}}(\mathbf{y}) = \inf_{\mathbf{x} \in \mathcal{X}} \|\mathbf{y} - \mathbf{x}\|_2$ is the distance function to $\mathcal{X}$ and $\mathbf{v}_t = (\mathbf{y}_t - \mathbf{x}_t)/\|\mathbf{y}_t - \mathbf{x}_t\|_2$ is the vector indicating the projection direction.

The main protocol of our reduction is presented as follows. The input includes original functions $\{f_t\}_{t=1}^T$, the feasible domain $\mathcal{X}$, and the reduced algorithm Algo.

for $t = 1, \ldots, T$ **do**

receive the gradient information $\nabla f_t(\mathbf{x}_t)$;

construct the surrogate loss $g_t : \mathcal{Y} \mapsto \mathbb{R}$ according to Eq. (5);

obtain the intermediate prediction $\mathbf{y}_{t+1} \leftarrow \mathsf{Algo}(g_t(\cdot), \mathbf{y}_t, \mathcal{Y})$;

submit the final prediction $\mathbf{x}_{t+1} = \Pi_{\mathcal{X}}[\mathbf{y}_{t+1}]$;

end for

Our reduction enjoys the regret safeness due to the following benign properties of surrogate loss.

Theorem 1 (Theorem 2 of Cutkosky [2020])**.** *The surrogate loss $g_t : \mathcal{Y} \mapsto \mathbb{R}$ defined in (5) is convex. Moreover, we have $\|\nabla g_t(\mathbf{y}_t)\|_2 \le \|\nabla f_t(\mathbf{x}_t)\|_2$ and for any $\mathbf{u}_t \in \mathcal{X}$*

$$\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{u}_t \rangle \le g_t(\mathbf{y}_t) - g_t(\mathbf{u}_t) \le \langle \nabla g_t(\mathbf{y}_t), \mathbf{y}_t - \mathbf{u}_t \rangle. \tag{6}$$

The theorem shows the convexity of the surrogate loss $g_t(\mathbf{y})$ and we thus have $f_t(\mathbf{x}_t) - f_t(\mathbf{u}_t) \le \langle \nabla g_t(\mathbf{y}_t), \mathbf{y}_t - \mathbf{u}_t \rangle$, which implies that it suffices to optimize the linearized upper bound, i.e., to optimize function $\ell_t(\mathbf{y}) = \langle \nabla g_t(\mathbf{y}_t), \mathbf{y} \rangle$. The following lemma specifies the gradient calculation.

Lemma 1. *For any $\mathbf{y} \in \mathcal{Y}$, $\nabla g_t(\mathbf{y}) = \nabla f_t(\mathbf{x}_t)$ when $\langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t \rangle \ge 0$; and $\nabla g_t(\mathbf{y}) = \nabla f_t(\mathbf{x}_t) - \langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t \rangle \cdot (\mathbf{y} - \Pi_{\mathcal{X}}[\mathbf{y}])/\|\mathbf{y} - \Pi_{\mathcal{X}}[\mathbf{y}]\|_2$ when $\langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t \rangle < 0$. Here $\mathbf{v}_t = (\mathbf{y}_t - \mathbf{x}_t)/\|\mathbf{y}_t - \mathbf{x}_t\|_2$. In particular, $\nabla g_t(\mathbf{y}_t) = \nabla f_t(\mathbf{x}_t) - \langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t \rangle \cdot \mathbf{v}_t$ when $\langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t \rangle < 0$.*

Reduction requirements. An important necessary condition for the reduction is to require the reduced algorithm satisfying *one gradient query* and *one function evaluation* at each round. Indeed, the reduction essentially updates according to the surrogate loss $\{g_t\}_{t=1}^T$. Note that the definition of surrogate loss involves the distance function $S_{\mathcal{X}}(\mathbf{y})$, see Eq. (5). Thus, each evaluation of $g_t(\mathbf{y})$ leads to one projection onto $\mathcal{X}$ due to the calculation of $S_{\mathcal{X}}(\mathbf{y})$. Similarly, each gradient query of $\nabla g_t(\mathbf{y})$ also contributes to one projection, see Lemma 1 for details. To summarize, we can use the reduction to ensure a 1 projection complexity, only when the reduced algorithm satisfies the requirements of one gradient query and one function evaluation per round. Below, we demonstrate the usage of our reduction mechanism for two methods of dynamic regret minimization that satisfy the conditions, including the worst-case method [Zhang et al., 2018a] and the small-loss method [Zhao et al., 2021b].

Application to dynamic regret minimization. Algorithm 1 summarizes the main procedures of our efficient methods for optimizing dynamic regret, which is an instance of the reduction mechanism by picking Algo as Ader [Zhang et al., 2018a]. More specifically, Lines 6 – 8 are essentially performing Ader algorithm using the surrogate loss $\{g_t\}_{t=1}^T$ over the surrogate domain $\mathcal{Y}$. Note that the base update in Line 6 is essentially performing OGD with projection onto $\mathcal{Y}$, a scaled Euclidean ball, and thus the projection admits a simple closed form. The overall algorithm requires projecting onto $\mathcal{X}$ only once per round, see Line 9. Our method provably retains the same dynamic regret.

Theorem 2. *Set the step size pool as* $\mathcal{H} = \{\eta_i = 2^{i-1}(D/G)\sqrt{5/(2T)} \mid i \in [N]\}$ *with* $N = \lceil 2^{-1}\log_2(1+2T/5)\rceil + 1$ *and the learning rate as* $\varepsilon = \sqrt{(\ln N)/(1+G^2D^2T)}$. *Under Assumptions 1 and 2, our algorithm requires one projection onto* $\mathcal{X}$ *per round and enjoys*

$$\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{u}_t) \leq \mathcal{O}\big(\sqrt{T(1+P_T)}\big). \tag{7}$$

For smooth and non-negative functions, the Sword++ algorithm [Zhao et al., 2021b] achieves an $\mathcal{O}(\sqrt{(F_T+P_T)(1+P_T)})$ small-loss dynamic regret, which requires one gradient and one function value per iteration.[1] However, notice that the surrogate loss $g_t(\cdot)$ in Eq. (5) is neither smooth nor non-negative, which hinders the application of our reduction to their method. Fortunately, owing to the benign property of $\|\nabla g_t(\mathbf{y}_t)\|_2 \leq \|\nabla f_t(\mathbf{x}_t)\|_2$ (see Theorem 1), we can still deploy the reduction via an improved analysis and obtain a projection-efficient algorithm with the same small-loss bound.

Theorem 3. *Set the step size pool as* $\mathcal{H} = \{\eta_i = 2^{i-1}\sqrt{5D^2/(1+8LGDT)} \mid i \in [N]\}$ *with* $N = \lceil 2^{-1}\log_2((5D^2+2D^2T)(1+8LGDT)/(5D^2))\rceil + 1$ *and the learning rate of the meta-algorithm as* $\varepsilon_t = \sqrt{(\ln N)/(1+D^2\sum_{s=1}^{t-1}\|\nabla g_s(\mathbf{y}_s)\|_2^2)}$. *Under Assumptions 1, 2, and 3, our algorithm requires one projection onto* $\mathcal{X}$ *per round and enjoys the following dynamic regret:*

$$\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{u}_t) \leq \mathcal{O}\big(\sqrt{(F_T+P_T)(1+P_T)}\big), \tag{8}$$

where $F_T = \sum_{t=1}^{T} f_t(\mathbf{u}_t)$ *is the cumulative loss of the comparators.*

3 Adaptive Regret Minimization

In this section, we present our efficient methods to minimize adaptive regret. First, we briefly review existing methods in Section 3.1, and then present our efficent methods to reducing the projection complexity of adaptive regret methods in Section 3.2.

3.1 A Brief Review of Adaptive Regret Minimization

Adaptive regret minimization ensures the online learner to be competitive with a fixed decision over every contiguous interval. For the worst-case bound, the best known result is the $\mathcal{O}(\sqrt{|I|\log T})$ adaptive regret bound achieved by the CBCE algorithm [Jun et al., 2017]. CBCE algorithm requires multiple gradients at each round. Wang et al. [2018] improve CBCE by using the linearized loss to make it requiring one gradient per iteration and retaining the same adaptive regret. Moreover, the improved CBCE algorithm only evaluates the function value once per iteration. Therefore, we can directly apply our reduction and obtain a projection-efficient variation with the same adaptive regret. More detailed elaborations can be found in Appendix C.1.

Now, we focus on the more challenging case of small-loss adaptive regret. The best known result is the $\mathcal{O}(\sqrt{F_I \log F_I \log F_T})$ bound for any interval $I = [r,s] \subseteq [T]$ obtained by the SACS algorithm [Zhang et al., 2019], where $F_I = \min_{\mathbf{x}\in\mathcal{X}} \sum_{t=r}^{s} f_t(\mathbf{x})$ and $F_T = \min_{\mathbf{x}\in\mathcal{X}} \sum_{t=1}^{T} f_t(\mathbf{x})$. However, SACS does not satisfy our reduction requirements, because it requires N gradient queries (i.e., $\nabla f_t(\mathbf{x}_{t,i})$ for $i \in [N]$) and $N+1$ function evaluations (i.e., $f_t(\mathbf{x}_{t,i})$ for $i \in [N]$, and $f_t(\mathbf{x}_t)$) at round $t \in [T]$, where N denotes the number of active base-learners and $\mathbf{x}_{t,i}$ denotes local decision returned by the i-th base-learner. To address so, we have to modify the algorithm to fit our purpose.

In the following, we first sketch the SACS algorithm and then present our modifications. In fact, to optimize the adaptive regret, an online algorithm usually consists of the three components:

[1]Sword++ algorithm is mainly proposed for gradient-variation dynamic regret, so there are advanced components (such as correction term and optimism) in algorithm design. It can be verified that their algorithm can be simplified by dropping the correction term and optimism when only small-loss bound is desired.

(i) base-algorithm: an online algorithm that can attain low (static) regret in a given interval;
(ii) scheduling: a set of intervals and each one is associated with a base-learner who aims to minimize the static regret over the interval (from starting time to ending time);
(iii) meta-algorithm: a combining algorithm that can track the best base-learner on the fly.

The specific configurations of the SACS algorithm is as follows. First, SACS uses scale-free online gradient descent (SOGD) [Orabona and Pál, 2018] as the base-algorithm, which ensures a small-loss regret in a given interval. Second, SACS employs AdaNormalHedge [Luo and Schapire, 2015] as the meta-algorithm, which supports the sleeping expert setup and also enjoys a small-loss regret. Finally, SACS designs a clever strategy of *problem-dependent* geometric covers to determine the set of intervals such that the number of active base-learners also depends on the small-loss quantity. As a result, SACS can achieve a fully problem-dependent adaptive regret of order $\mathcal{O}(\sqrt{F_I \log F_I \log F_T})$, scaling with the cumulative loss of comparators. However, SACS also suffers from an $\mathcal{O}(\log T)$ projection complexity in the worst case due to a two-layer structure; and moreover, it can be observed that SACS only attains an $\mathcal{O}(\sqrt{|I| \log |I| \log T})$ bound in the worst case, which exhibits an $\sqrt{\log |I|}$ gap compared with the best known result of $\mathcal{O}(\sqrt{|I| \log T})$ [Jun et al., 2017]. Below, we present an efficient algorithm for small-loss adaptive regret, which resolves the above two issues simultaneously.

3.2 Efficient Algorithms for Adaptive Regret

As multiple gradient queries and function evaluations are involved in all the three components of SACS, we have to make plenty of modifications to achieve an algorithm with small-loss adaptive regret yet requiring only one gradient query and function evaluation per round. With such an algorithm on hand, we can then deploy our reduction to achieve an efficient method with 1 projection complexity. Below we present the details. By the reduction mechanism, it is noticeable that we only need to consider the input online functions as surrogate loss $\{g_t\}_{t=1}^T$, where g_t is defined in Eq. (5).

Base-algorithm. We use SOGD with a *linearized* surrogate loss $\langle \nabla g_t(\mathbf{y}_t), \mathbf{y} \rangle$ over the surrogate domain $\mathcal{Y}$. Denote by A_t the set of active base-learners' indices, then the base-learner $\mathcal{B}_i$ updates by

$$\mathbf{y}_{t+1,i} = \Pi_{\mathcal{Y}}[\mathbf{y}_{t,i} - \eta_{t,i} \nabla g_t(\mathbf{y}_t)], \tag{9}$$

with $\eta_{t,i} = D / \sqrt{(\delta + \sum_{s=\tau_i}^t \|\nabla g_s(\mathbf{y}_s)\|_2^2)}$, where τ_i denotes the starting time of the base-learner $i \in A_t$. The projection onto $\mathcal{Y}$ can be easily calculated by a simple rescaling if needed. Notably, owing to the convexity of the surrogate loss g_t, we can use the *same* gradient $\nabla g_t(\mathbf{y}_t)$ for all the base-learners at each round, ensuring one gradient query of $\nabla f_t(\mathbf{x}_t)$ at each round.

Geometric Covers. The covers consist of a set of intervals that specify the alive time of base-learners. To achieve a small-loss adaptive regret, SACS [Zhang et al., 2019] employs a clever covering construction called problem-dependent geometric covers (PGC) — instead of initiating a base-learner at each round t like earlier algorithms [Daniely et al., 2015; Jun et al., 2017], SACS adds a new base-learner only when the cumulative loss exceeds a pre-defined *threshold*. As a result, the number of active base-learners relates to the small-loss quantity such that the overall algorithm achieves a fully problem-dependent adaptive regret. Notably, to determine the threshold, SACS monitors the cumulative loss of the latest base-learner $f_t(\mathbf{x}_{t,i^\dagger})$ with $i^\dagger$ being the latest base-learner's index, but clearly this will introduction an additional function evaluation beyond $f_t(\mathbf{x}_t)$ at each round.

To avoid the limitation, instead of using a fixed threshold to decide the initiations of base-learners, we design a sequence of *time-varying thresholds* to adaptively start a new base-learner according to amount of cumulative loss of *final decisions* (e.g., $f_t(\mathbf{x}_t)$), bypassing the requirement of additional function evaluation. This realizes the condition of one function evaluation per round. Also, the new design of thresholds mechanism is important to ensure that the overall small-loss bound can simultaneously recover the best known worst-case guarantee, which SACS fails to achieve [Zhang et al., 2019]. Let $C_1, C_2, C_3, \ldots$ denote the sequence of thresholds, and they will be determined by a threshold generating function $\mathcal{G}(\cdot) : \mathbb{N} \mapsto \mathbb{R}_+$ that will be specified later. Our problem-dependent geometric covers are set as follows. We initialize the setting by $s_1 = 1$. We set s_2 as the round when the cumulative loss of the overall algorithm (namely, $\sum_{s=1}^t f_s(\mathbf{x}_s)$) exceeds the threshold C_1 and then initialize a new instance of SOGD starting at this round. The process is repeated until the end of online game. We thus generate a sequence of points $s_1, s_2, \ldots$, referred to as the *markers*. See

Algorithm 2 Efficient Algorithm for Problem-dependent Adaptive Regret

Input: threshold generating function $\mathcal{G}(\cdot) : \mathbb{N} \mapsto \mathbb{R}_+$.

1: Initialize total intervals $m = 1$, marker $s_1 = 1$, threshold $C_1 = \mathcal{G}(1)$; let $\mathbf{x}_1$ be any point in $\mathcal{X}$.
2: **for** $t = 1$ **to** T **do**
3: Receive the gradient information $\nabla f_t(\mathbf{x}_t)$.
4: Construct the surrogate loss $g_t : \mathcal{Y} \mapsto \mathbb{R}$ according to Eq. (5).
5: Compute the (sub-)gradient $\nabla g_t(\mathbf{y}_t)$ according to Lemma 1.
6: Compute $L_t = L_{t-1} + f_t(\mathbf{x}_t)$
`% constructing Problem-dependent Geometric Covers(PGC)`
7: **if** $L_t > C_m$ **then**
8: Set $L_t = 0$, remove base-learners $\mathcal{B}_k$ whose ending point $e_k = m + 1$.
9: Set $m \leftarrow m + 1$, $s_m \leftarrow t$, $C_m = \mathcal{G}(m)$.
10: Initialize a new base-learner with ending point $e_m = j$ satisfying $[m, j-1] \in \mathcal{C}$, where $\mathcal{C} = \bigcup_{k \in \mathbb{N} \cup \{0\}} \mathcal{C}_k$ and $\mathcal{C}_k = \left\{[i \cdot 2^k, (i+1) \cdot 2^k - 1] \mid i \text{ is odd}\right\}$ for all $k \in \mathbb{N} \cup \{0\}$.
11: Set $\gamma_m = \ln(1 + 2m)$, $w_{t,m} = 1$, $\eta_{t,m} = \min\{1/2, \sqrt{\gamma_m}\}$ for the meta-algorithm.
12: **end if**
13: Send $\nabla g_t(\mathbf{y}_t)$ to all base-learners and obtain local predictions $\mathbf{y}_{t+1,i}$ for $i \in A_t$.
14: Meta-algorithm updates weight $\boldsymbol{p}_{t+1} \in \Delta_{|A_{t+1}|}$ according to Eq. (11), Eq. (12), and Eq. (13)
15: Compute $\mathbf{y}_{t+1} = \sum_{i \in A_{t+1}} p_{t+1,i} \mathbf{y}_{t+1,i}$.
16: Submit $\mathbf{x}_{t+1} = \Pi_{\mathcal{X}}[\mathbf{y}_{t+1}]$. ▷ The only projection onto feasible domain $\mathcal{X}$ per round.
17: **end for**

the condition in Line 7, registration of markers in Line 9, and the overall updates in Lines 7 – 11 of Algorithm 2. Those markers specify the starting time (and the ending time) of base-learners and thus construct the PGC as

$$\widetilde{\mathcal{C}} = \bigcup_{k \in \mathbb{N} \cup \{0\}} \widetilde{\mathcal{C}}_k, \text{ where } \widetilde{\mathcal{C}}_k = \left\{[s_{i \cdot 2^k}, s_{(i+1) \cdot 2^k} - 1] \mid i \text{ is odd}\right\} \text{ for all } k \in \mathbb{N} \cup \{0\}. \tag{10}$$

It is noteworthy to emphasize that PGC is constructed by the language of "marker", whose exact time stamp is *unknown* ahead of time but is only determined according to the learner's performance on the fly. Moreover, the notation $\mathcal{C}$ in Lines 10 of Algorithm 2 is defined based on the registered indexes of markers, and there is a one-one correspondence from the interval in $\mathcal{C}$ to that in PGC $\widetilde{\mathcal{C}}$. More concretely, an interval $[i \cdot 2^k, (i+1) \cdot 2^k - 1] \in \mathcal{C}$ will be mapped into the interval $[s_{i \cdot 2^k}, s_{(i+1) \cdot 2^k} - 1]$ in PGC, managing the alive time of base-learners in a geometric manner with respect to the subscripts.

Meta-algorithm. SACS uses the AdaNormalHedge [Luo and Schapire, 2015] as the meta-algorithm, however, this is not suitable for our propose. To ensure one projection per iteration, we cannot use multiple function values, i.e., $\{g_t(\mathbf{y}_{t,i})\}_{i=1}^N$, for meta-algorithm to evaluate the loss. Instead, we can only use the *linearized* loss value, namely, $\{\langle \nabla g_t(\mathbf{y}_t), \mathbf{y}_{t,i} \rangle\}_{i=1}^N$ in the weight update of meta-algorithm. The small-loss regret bound in the meta-algorithm of SACS crucially relies on the original function values, which is unfortunately inaccessible in our case. Technically, when fed with linearized loss, it is hard to establish a *squared* gradient-norm bound and then convert it to the small loss due to the *first-order* regret bound of AdaNormalHedge. Based on this crucial technical observation, we propose to use the Adapt-ML-Prod algorithm [Gaillard et al., 2014] as the meta-algorithm in our method. The key advantage is that it enjoys a *second-order* regret and also supports the sleeping expert setup. Adapt-ML-Prod maintains multiple learning rates $\boldsymbol{\eta}_{t+1}$ and an intermediate weight vector $\mathbf{w}_{t+1}$, which are updated by the following rule. For any active base-learner $i \in A_{t+1}$,

$$\eta_{t+1,i} = \min\left\{\frac{1}{2}, \sqrt{\frac{\gamma_i}{1 + \sum_{k=s_i}^t (\widehat{\ell}_k - \ell_{k,i})^2}}\right\}, w_{t+1,i} = \left(w_{t,i}\left(1 + \eta_{t,i}(\widehat{\ell}_t - \ell_{t,i})\right)\right)^{\frac{\eta_{t+1,i}}{\eta_{t,i}}}, \tag{11}$$

where $\gamma_i = \ln(1 + 2i)$ is a certain scaling factor and the feedback loss is constructed as for $i \in A_t$

$$\widehat{\ell}_t = \langle \nabla g_t(\mathbf{y}_t), \mathbf{y}_t \rangle / (2GD), \text{ and } \ell_{t,i} = \langle \nabla g_t(\mathbf{y}_t), \mathbf{y}_{t,i} \rangle / (2GD). \tag{12}$$

The final weight vector $\boldsymbol{p}_{t+1} \in \Delta_{|A_{t+1}|}$ is obtained by

$$p_{t+1,i} = \frac{w_{t+1,i} \cdot \eta_{t+1,i}}{\sum_{j \in A_{t+1}} w_{t+1,j} \cdot \eta_{t+1,j}}. \tag{13}$$

Notably, the meta update only uses one gradient at round t, namely, $\nabla g_t(\mathbf{y}_t)$.

Finally, we compute $\mathbf{y}_{t+1} = \sum_{i \in A_{t+1}} p_{t+1,i} \mathbf{y}_{t+1,i}$ as the overall prediction in the surrogate domain $\mathcal{Y}$ and calculate $\mathbf{x}_{t+1} = \Pi_{\mathcal{X}}[\mathbf{y}_{t+1}]$ to ensure the feasibility. This is the only projection onto $\mathcal{X}$ at each round. Algorithm 2 summarizes the main procedures of our efficient methods for small-loss adaptive regret. Albeit with a similar two-layer structure as SACS, our algorithm exhibits salient differences in base-algorithm, meta-algorithm, and geometric covers. As a benefit, we can successfully deploy our reduction mechanism and make the overall algorithm project onto the feasible domain $\mathcal{X}$ once per round, see Line 16. Our method retains the same small-loss adaptive regret as [Zhang et al., 2019].

Theorem 4. *Under Assumptions 1–3, setting the threshold generating function $\mathcal{G}(m) = \Theta(\log m)$ whose explicit form is in Eq. (46) of Appendix C, Algorithm 2 requires only one projection onto $\mathcal{X}$ per round and enjoys the small-loss adaptive regret:*

$$\sum_{t=r}^{s} f_t(\mathbf{x}_t) - \sum_{t=r}^{s} f_t(\mathbf{u}) \leq \mathcal{O}\Big(\min\big\{\sqrt{F_I \log F_I \log F_T}, \sqrt{|I| \log T}\big\}\Big) \tag{14}$$

for any interval $I = [r, s] \subseteq [T]$, where $F_I = \min_{\mathbf{x} \in \mathcal{X}} \sum_{t=r}^{s} f_t(\mathbf{x})$ and $F_T = \min_{\mathbf{x} \in \mathcal{X}} \sum_{t=1}^{T} f_t(\mathbf{x})$.

Remark 1. Note that the $\mathcal{O}(\sqrt{F_I \log F_I \log F_T})$ small-loss bound of Zhang et al. [2019] becomes $\mathcal{O}(\sqrt{|I| \log |I| \log T})$ in the worst case, looser than the $\mathcal{O}(\sqrt{|I| \log T})$ bound [Jun et al., 2017] by a factor of $\log |I|$. We show that this limitation can be actually avoided by the new design of thresholds mechanism and a refined analysis. More discussions can be found in Appendix C.3. Indeed, our result in (14) can *strictly* match the best known problem-independent result in the worst case.

4 Experiment

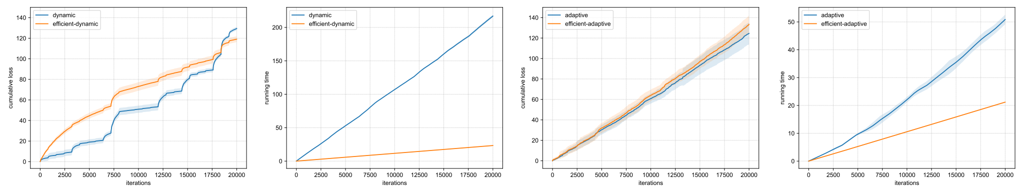
In this section, we provide empirical studies to evaluate our proposed methods.

General Setup. We conduct experiments on the synthetic data. We consider the following online regression problem. Let T denote the number of total rounds. At each round $t \in [T]$ the learner outputs the model parameter $\mathbf{w}_t \in \mathcal{W} \subseteq \mathbb{R}^d$ and simultaneously receives a data sample (x_t, y_t) with $x_t \in \mathcal{X} \subseteq \mathbb{R}^d$ being the feature and $y_t \in \mathbb{R}$ being the corresponding label.[2] The learner can then evaluate her model by the online loss $f_t(\mathbf{w}_t) = \frac{1}{2}(x_t^\top \mathbf{w}_t - y_t)^2$ which uses a square loss to evaluate the difference between the predictive value $x_t^\top \mathbf{w}_t$ and the ground-truth label y_t, and then use the feedback information to update the model. In the simulations, we set $T = 20000$, the domain diameter as $D = 6$, and the dimension of the domain as $d = 8$. The feasible domain $\mathcal{W}$ is set as an ellipsoid $\mathcal{W} = \big\{\mathbf{w} \in \mathbb{R}^d \mid \mathbf{w}^\top \mathbf{E} \mathbf{w} \leq \lambda_{\min}(\mathbf{E}) \cdot (D/2)^2\big\}$, where $\mathbf{E}$ is a certain diagonal matrix and $\lambda_{\min}(\mathbf{E})$ denotes its minimum eigenvalue. Then, a projection onto $\mathcal{W}$ requires solving a convex programming that is generally expensive. In the experiment, we use `scipy.optimize.NonlinearConstraint` to solve it to perform the projection onto the feasible domain.

To simulate the non-stationary online environments, we control the way to generate the date samples $\{(x_t, y_t)\}_{t=1}^T$. Specifically, for $t \in [T]$, the feature x_t is randomly sampled in an Euclidean ball with a diameter D same as the feasible domain of model parameters; and the corresponding label is set as $y_t = x_t^\top \mathbf{w}_t^* + \varepsilon_t$, where ε_t is the random noise drawn from $[0, 0.1]$ and $\mathbf{w}_t^*$ is the underlying ground-truth model from the feasible domain $\mathcal{W}$ generated according to a certain strategy specified below. For dynamic regret minimization, we simulate *piecewise-stationary* model drifts, as dynamic regret will be linear in T and thus vacuous when the model drift happens every round due to a linear path length measure. Concretely, we split the time horizon evenly into 25 stages and restrict the underlying model parameter $\mathbf{w}_t^*$ to be stationary within a stage. For adaptive regret minimization, we simulate *gradually evolving* model drifts, where the underlying model parameter $\mathbf{w}_{t+1}^*$ is generated based on the last-round model parameter $\mathbf{w}_t^*$ with an additional random walk in the feasible domain $\mathcal{W}$. The step size of random walk is set to be proportional to D/T to ensure a smooth model change.

Contenders. For both dynamic regret and adaptive regret minimization, we directly work on the small-loss online methods. We choose the Sword algorithm [Zhao et al., 2021b] as the contender of

[2] With a slight abuse of notations, we here use $\mathbf{w}$ to denote the model parameter and $\mathcal{W}$ to denote the feasible domain, while reserve the notations of x and $\mathcal{X}$ to denote the feature and feature space following the conventional notations of machine learning terminologies.



(a) dynamic regret (loss) (b) dynamic regret (time) (c) adaptive regret (loss) (d) adaptive regret (time)

Figure 1: Performance comparisons of existing methods and our methods (indicated by "efficient" prefix) in terms of cumulative loss and running time (in seconds). The first two figures plot the results of methods for dynamic regret minimization, while the latter ones are for adaptive regret.

our efficient method for dynamic regret (Algorithm 1) and choose the SACS algorithm [Zhang et al., 2019] as the contender of our efficient method for adaptive regret (Algorithm 2).

Results. We repeat the experiments for five times with different random seeds and report the results (mean and standard deviation) in Figure 1. We use a machine with a single CPU (Intel(R) Core(TM) i9-10900K CPU @ 3.70GHz) and 32GB main memory to conduct the experiments. We plot both cumulative loss and running time (in seconds) for all the methods. We first examine the performance of dynamic regret minimization, see Figure 1(a) for cumulative loss and see Figure 1(b) for running time. The empirical results show that our method has a comparable performance to Sword without much sacrifice of cumulative loss, while our method can achieve about 10 times speedup due to the improved projection complexity. Second, as shown in Figure 1(c) and Figure 1(d), a similar performance enhancement also appears in adaptive regret minimization, though the speedup is slightly smaller due to the fact that fewer learners are required to maintain for adaptive regret. To summarize, the empirical results show the effectiveness of our methods in retaining the regret performance and also the efficiency in terms of the running time due to the reduced projection complexity.

5 Conclusion

In this paper, we design a generic reduction mechanism that can reduce the projection complexity of two-layer methods for non-stationary online learning, hence approaching a clearer resolution of necessary computational overhead for robustness to non-stationarity. Building on the reduction mechanism, we develop a series of online algorithms for optimizing dynamic regret and adaptive regret. All the algorithms retain the best known regret guarantees, and more importantly, require one projection onto the feasible domain per iteration. It is further worth mentioning that, due to the requirement of our reduction, all our algorithms only need one gradient query and one function evaluation at each round as well, which can be appealing in situations with limited feedback.

Our reduction can also be applied to other settings to achieve light project complexity, for example, dynamic regret of OCO with memory [Zhao et al., 2022b], OCO with switching cost [Zhang et al., 2021], and related applications such as online non-stochastic control [Hazan et al., 2020]. Moreover, it is possible to derive similar efficient algorithms for minimizing the interval dynamic regret, an even stringent measure for non-stationary online convex optimization. There is one important open question left on another type of problem-dependent bound that scales with gradient variation [Chiang et al., 2012], which plays an important role in establishing fast convergence in zero-sum games [Syrgkanis et al., 2015; Zhang et al., 2022]. Although Zhao et al. [2021b] have devised a two-layer method that enjoys a gradient-variation dynamic regret and requires one gradient per iteration, it is quite challenging to incorporate the optimistic online learning into our reduction mechanism due to the constrained feasible domain and the complicated two-layer structure. Finally, it would be greatly important to further understand the minimal computational overhead in response to the robustness to non-stationarity, in particular, some information-theoretic arguments would be highly interesting.

Acknowledgements

This research was supported by NSFC (61921006, 62206125), JiangsuSF (BK20220776), National Postdoctoral Program for Innovative Talent, and Collaborative Innovation Center of Novel Software Technology and Industrialization.

References

P. Auer, N. Cesa-Bianchi, and C. Gentile. Adaptive and self-confident on-line learning algorithms. *Journal of Computer and System Sciences*, 64(1):48–75, 2002.

D. Baby and Y.-X. Wang. Optimal dynamic regret in exp-concave online learning. In *Proceedings of the 34th Conference on Learning Theory (COLT)*, pages 359–409, 2021.

C.-K. Chiang, T. Yang, C.-J. Lee, M. Mahdavi, C.-J. Lu, R. Jin, and S. Zhu. Online optimization with gradual variations. In *Proceedings of the 25th Conference On Learning Theory (COLT)*, pages 6.1–6.20, 2012.

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

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

A. Daniely, A. Gonen, and S. Shalev-Shwartz. Strongly adaptive online learning. In *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, pages 1405–1411, 2015.

Y. Freund and R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, 1997.

P. Gaillard, G. Stoltz, and T. van Erven. A second-order bound with excess losses. In *Proceedings of the 27th Conference on Learning Theory (COLT)*, pages 176–196, 2014.

E. C. Hall and R. Willett. Dynamical models and tracking regret in online convex programming. In *Proceedings of the 30th International Conference on Machine Learning (ICML)*, pages 579–587, 2013.

E. Hazan. Introduction to Online Convex Optimization. *Foundations and Trends in Optimization*, 2 (3-4):157–325, 2016.

E. Hazan and C. Seshadhri. Efficient learning algorithms for changing environments. In *Proceedings of the 26th International Conference on Machine Learning (ICML)*, pages 393–400, 2009.

E. Hazan, S. M. Kakade, and K. Singh. The nonstochastic control problem. In *Proceedings of the 31st International Conference on Algorithmic Learning Theory (ALT)*, pages 408–421, 2020.

K.-S. Jun, F. Orabona, S. Wright, and R. Willett. Improved strongly adaptive online learning using coin betting. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pages 943–951, 2017.

H. Luo and R. E. Schapire. Achieving all with no parameters: AdaNormalHedge. In *Proceedings of the 28th Annual Conference Computational Learning Theory (COLT)*, pages 1286–1304, 2015.

Z. Mhammedi, W. M. Koolen, and T. van Erven. Lipschitz adaptivity with multiple learning rates in online learning. In *Proceedings of the 32nd Annual Conference Computational Learning Theory (COLT)*, pages 2490–2511, 2019.

F. Orabona. A Modern Introduction to Online Learning. *ArXiv preprint*, 1912.13213, 2019.

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

R. Pogodin and T. Lattimore. On first-order bounds, variance and gap-dependent bounds for adversarial bandits. In *Proceedings of the 35th Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 894–904, 2019.

S. Shalev-Shwartz. Online Learning: Theory, Algorithms and Applications. *PhD Thesis*, 2007.

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

N. Srebro, K. Sridharan, and A. Tewari. Smoothness, low noise and fast rates. In *Advances in Neural Information Processing Systems 23 (NIPS)*, pages 2199–2207, 2010.

V. Syrgkanis, A. Agarwal, H. Luo, and R. E. Schapire. Fast convergence of regularized learning in games. In *Advances in Neural Information Processing Systems 28 (NIPS)*, pages 2989–2997, 2015.

T. van Erven and W. M. Koolen. Metagrad: Multiple learning rates in online learning. In *Advances in Neural Information Processing Systems 29 (NIPS)*, pages 3666–3674, 2016.

G. Wang, D. Zhao, and L. Zhang. Minimizing adaptive regret with one gradient per iteration. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 2762–2768, 2018.

L. Zhang. Online learning in changing environments. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 5178–5182, 2020. Early Career.

L. Zhang, S. Lu, and Z.-H. Zhou. Adaptive online learning in dynamic environments. In *Advances in Neural Information Processing Systems 31 (NeurIPS)*, pages 1330–1340, 2018a.

L. Zhang, T. Yang, R. Jin, and Z.-H. Zhou. Dynamic regret of strongly adaptive methods. In *Proceedings of the 35th International Conference on Machine Learning (ICML)*, pages 5877–5886, 2018b.

L. Zhang, T.-Y. Liu, and Z.-H. Zhou. Adaptive regret of convex and smooth functions. In *Proceedings of the 36th International Conference on Machine Learning (ICML)*, pages 7414–7423, 2019.

L. Zhang, S. Lu, and T. Yang. Minimizing dynamic regret and adaptive regret simultaneously. In *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics (AISTATS)*, pages 309–319, 2020.

L. Zhang, W. Jiang, S. Lu, and T. Yang. Revisiting smoothed online learning. In *Advances in Neural Information Processing Systems 34 (NeurIPS)*, pages 13599–13612, 2021.

M. Zhang, P. Zhao, H. Luo, and Z.-H. 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.

P. Zhao, Y.-J. Zhang, L. Zhang, and Z.-H. Zhou. Dynamic regret of convex and smooth functions. In *Advances in Neural Information Processing Systems 33 (NeurIPS)*, pages 12510–12520, 2020.

P. Zhao, G. Wang, L. Zhang, and Z.-H. Zhou. Bandit convex optimization in non-stationary environments. *Journal of Machine Learning Research*, 22(125):1–45, 2021a.

P. Zhao, Y.-J. Zhang, L. Zhang, and Z.-H. Zhou. Adaptivity and non-stationarity: Problem-dependent dynamic regret for online convex optimization. *ArXiv preprint*, arXiv:2112.14368, 2021b.

P. Zhao, L.-F. Li, and Z.-H. Zhou. Dynamic regret of online markov decision processes. In *Proceedings of the 39th International Conference on Machine Learning (ICML)*, pages 26865–26894, 2022a.

P. Zhao, Y.-X. Wang, and Z.-H. Zhou. Non-stationary online learning with memory and non-stochastic control. In *Proceedings of the 25th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pages 2101–2133, 2022b.

Z.-H. Zhou. Open-environment machine learning. *National Science Review*, 9(8):nwac123, 2022.

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

Checklist

The checklist follows the references. Please read the checklist guidelines carefully for information on how to answer these questions. For each question, change the default **[TODO]** to [Yes] , [No] , or [N/A] . You are strongly encouraged to include a **justification to your answer**, either by referencing the appropriate section of your paper or providing a brief inline description.

Please do not modify the questions and only use the provided macros for your answers. Note that the Checklist section does not count towards the page limit. In your paper, please delete this instructions block and only keep the Checklist section heading above along with the questions/answers below.

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes] See discussions in Section 5
   (c) Did you discuss any potential negative societal impacts of your work? [N/A]

(d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]

2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [Yes]
   (b) Did you include complete proofs of all theoretical results? [Yes]
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes]
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes]
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [Yes]
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [N/A]
   (b) Did you mention the license of the assets? [N/A]
   (c) Did you include any new assets either in the supplemental material or as a URL? [N/A]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]

A Omitted Details for Reduction Mechanism

In this section, we provide the proofs of Theorem 1 and Lemma 1.

A.1 Properties of Distance Function

Before presenting the proofs, we here collect two useful lemmas regarding the distance function used in the surrogate loss, which will be useful in the following proofs. The proofs of the two lemmas can be found in the seminal paper of Cutkosky and Orabona [2018].

Lemma 2 (Proposition 1 of Cutkosky and Orabona [2018])**.** *The distance function* $S_{\mathcal{X}}(\mathbf{y}) = \inf_{\mathbf{x}\in\mathcal{X}} \|\mathbf{y}-\mathbf{x}\|_2$ *is convex and* 1*-Lipschitz for any closed convex feasible domain* $\mathcal{X} \subseteq \mathbb{R}^d$.

Lemma 3 (Theorem 4 of Cutkosky and Orabona [2018])**.** *Let* $\mathcal{X} \subseteq \mathbb{R}^d$ *a closed convex set. Given* $\mathbf{y} \in \mathbb{R}^d$ *and* $\mathbf{y} \notin \mathcal{X}$. *Let* $\mathbf{x} = \Pi_{\mathcal{X}}[\mathbf{y}]$. *Then we have* $\{\frac{\mathbf{y}-\mathbf{x}}{\|\mathbf{y}-\mathbf{x}\|_2}\} = \partial S_{\mathcal{X}}(\mathbf{y})$.

A.2 Proof of Theorem 1

Theorem 1 is originally due to Cutkosky [2020], and for self-containedness we restate their proof using our notations.

Proof. When $\langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t\rangle \geq 0$, by the definition of the surrogate loss defined in Eq. (5), we have $g_t(\mathbf{y}) = \langle \nabla f_t(\mathbf{x}_t), \mathbf{y}\rangle$, which is linear in $\mathbf{y}$ and thus convex (in fact linear in y). It is clear that $\|\nabla g_t(\mathbf{y}_t)\|_2 = \|\nabla f_t(\mathbf{x}_t)\|_2$ and thus satisfies the claimed inequality of gradient norms in the statement. Moreover, the inequality (6) holds evidently due to the linear surrogate loss in this case.

Let us focus on the case when $\langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t\rangle < 0$. First, it can be verified that the surrogate loss $g_t(\mathbf{y}) = \langle \nabla f_t(\mathbf{x}_t), \mathbf{y}\rangle - \langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t\rangle \cdot S_{\mathcal{X}}(\mathbf{y})$ is convex due to the convexity of $S_{\mathcal{X}}(\mathbf{y})$ shown in Lemma 2 and the condition of $\langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t\rangle < 0$ in this case. Next, the gradient of $g_t(\cdot)$ at the $\mathbf{y}_t$ point can be calculated according to Lemma 1 as,

$$\nabla g_t(\mathbf{y}_t) = \nabla f_t(\mathbf{x}_t) - \langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t\rangle \cdot \mathbf{v}_t$$

where $\mathbf{v}_t = (\mathbf{y}_t - \mathbf{x}_t)/\|\mathbf{y}_t - \mathbf{x}_t\|_2$. Notice that $\|\mathbf{v}_t\|_2 = 1$ and $\nabla g_t(\mathbf{y}_t)$ is an orthogonal projection of $\nabla f_t(\mathbf{x}_t)$ onto the subspace perpendicular to the vector $\mathbf{v}_t$, so we have $\|\nabla g_t(\mathbf{y}_t)\|_2 \leq \|\nabla f_t(\mathbf{x}_t)\|_2$. Finally, we proceed to prove the inequality (6) in this case. Since the comparator $\mathbf{u}_t \in \mathcal{X}$ is in the feasible domain, we have $S_{\mathcal{X}}(\mathbf{u}_t) = \|\mathbf{u}_t - \mathbf{u}_t\|_2 = 0$ and get

$$\begin{aligned}
&\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{u}_t\rangle\\
&= \langle \nabla f_t(\mathbf{x}_t), \mathbf{y}_t\rangle + \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{y}_t\rangle - \langle \nabla f_t(\mathbf{x}_t), \mathbf{u}_t\rangle\\
&= \langle \nabla f_t(\mathbf{x}_t), \mathbf{y}_t\rangle - \langle \nabla f_t(\mathbf{x}_t), \frac{\mathbf{y}_t - \mathbf{x}_t}{\|\mathbf{y}_t - \mathbf{x}_t\|_2}\rangle \cdot \|\mathbf{y}_t - \mathbf{x}_t\|_2 - \langle \nabla f_t(\mathbf{x}_t), \mathbf{u}_t\rangle\\
&= \langle \nabla f_t(\mathbf{x}_t), \mathbf{y}_t\rangle - \langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t\rangle \cdot S_{\mathcal{X}}(\mathbf{y}_t) - \langle \nabla f_t(\mathbf{x}_t), \mathbf{u}_t\rangle + \langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t\rangle \cdot S_{\mathcal{X}}(\mathbf{u}_t)\\
&= g_t(\mathbf{y}_t) - g_t(\mathbf{u}_t)\\
&\leq \langle \nabla g_t(\mathbf{y}_t), \mathbf{y}_t - \mathbf{u}_t\rangle,
\end{aligned}$$

where the last inequality holds owing to the convexity of the surrogate loss proven earlier.

Combining the two cases finishes the proof. □

A.3 Proof of Lemma 1

Lemma 1 is originally due to Cutkosky and Orabona [2018], and for self-containedness we restate their proof using our notations.

Proof. With a slight abuse of notations, for simplicity we use the notation $\nabla g_t(\mathbf{y})$ to denote the (sub-)gradient of surrogate function $g_t(\cdot)$ at point $\mathbf{y}$, no matter whether the function is differentiable.

When $\langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t\rangle \geq 0$, the surrogate loss is $g_t(\mathbf{y}) = \langle \nabla f_t(\mathbf{x}_t), \mathbf{y}\rangle$ by definition in Eq. (5). Therefore, the gradient simply becomes $\nabla g_t(\mathbf{y}_t) = \nabla f_t(\mathbf{x}_t)$.

When $\langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t \rangle < 0$, the surrogate loss becomes $g_t(\mathbf{y}) = \langle \nabla f_t(\mathbf{x}_t), \mathbf{y} \rangle - \langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t \rangle \cdot S_{\mathcal{X}}(\mathbf{y})$ according to definition in Eq. (5). By Lemma 3, the gradient $\nabla g_t(\mathbf{y})$ can be calculated by

$$\nabla g_t(\mathbf{y}) = \nabla f_t(\mathbf{x}_t) - \langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t \rangle \cdot \frac{\mathbf{y} - \Pi_{\mathcal{X}}[\mathbf{y}]}{\|\mathbf{y} - \Pi_{\mathcal{X}}[\mathbf{y}]\|_2},$$

where the computation needs the projection onto domain $\mathcal{X}$. In particular, for $\mathbf{y}_t$, we have

$$\nabla g_t(\mathbf{y}_t) = \nabla f_t(\mathbf{x}_t) - \langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t \rangle \cdot \frac{\mathbf{y}_t - \mathbf{x}_t}{\|\mathbf{y}_t - \mathbf{x}_t\|_2} = \nabla f_t(\mathbf{x}_t) - \langle \nabla f_t(\mathbf{x}_t), \mathbf{v}_t \rangle \cdot \mathbf{v}_t.$$

This ends the proof. □

B Omitted Details for Dynamic Regret Minimization

In this section, we provide the proofs for the theorems presented in Section 2. Specifically, we first prove the worst-case bound (Theorem 2) and then work on the small-loss bound (Theorem 3).

B.1 Proof of Theorem 2

Proof. Notice that Zhang et al. [2018a] propose the improved Ader algorithm (see Algorithm 3 and Algorithm 4 in their paper), which uses the linearized loss as the input to make the online algorithm requiring one gradient and one function evaluation per iteration. So the algorithm satisfies the requirements of our reduction mechanism, and our algorithm can be regarded as the improved Ader equipped with the projection-efficient reduction. As a consequence, we can directly obtain the same dynamic regret guarantee and ensure 1 projection complexity at the same time, by following the same proof of the improved Ader as well as the reduction guarantee (Theorem 1). □

B.2 Proof of Theorem 3

Proof. By the reduction guarantee shown in Theorem 1, we have the following result that decomposes the dynamic regret into the two terms.

$$\begin{aligned}\sum_{t=1}^T f_t(\mathbf{x}_t) - \sum_{t=1}^T f_t(\mathbf{u}_t) &\leq \sum_{t=1}^T g_t(\mathbf{y}_t) - \sum_{t=1}^T g_t(\mathbf{u}_t) \leq \sum_{t=1}^T \langle \nabla g_t(\mathbf{y}_t), \mathbf{y}_t - \mathbf{u}_t \rangle \\ &= \underbrace{\sum_{t=1}^T \langle \nabla g_t(\mathbf{y}_t), \mathbf{y}_t - \mathbf{y}_{t,i} \rangle}_{\texttt{meta-regret}} + \underbrace{\sum_{t=1}^T \langle \nabla g_t(\mathbf{y}_t), \mathbf{y}_{t,i} - \mathbf{u}_t \rangle}_{\texttt{base-regret}}, \end{aligned} \tag{15}$$

where in (15) the first term is called *meta-regret* as it measures the regret overhead of the meta-algorithm to track the unknown best base-learner, and the second term is called the *base-regret* to denote the dynamic regret of the base-learner i. Note that the above decomposition holds for any base-learner index $i \in [N]$.

Upper bound of meta-regret. As the meta-algorithm can be regarded as a FTRL with time-varying learning rates and a negative entropy regularizer, we apply Lemma 10 to obtain an upper bound for the meta-regret. Indeed, by choosing $\ell_{t,i} = \langle \nabla g_t(\mathbf{y}_t), \mathbf{y}_{t,i} \rangle$ in Lemma 10, we can achieve that

$$\begin{aligned}\sum_{t=1}^T \langle \nabla g_t(\mathbf{y}_t), \mathbf{y}_t - \mathbf{y}_{t,i} \rangle &\leq 3\sqrt{\ln N \left(1 + \sum_{t=1}^T D^2 \|\nabla g_t(\mathbf{y}_t)\|_2^2\right)} + \frac{G^2 D^2 \sqrt{\ln N}}{2} \\ &\leq 3D\sqrt{\ln N \sum_{t=1}^T \|\nabla g_t(\mathbf{y}_t)\|_2^2} + \frac{(6 + G^2 D^2)\sqrt{\ln N}}{2} \\ &\leq 3D\sqrt{\ln N \sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t)\|_2^2} + \mathcal{O}(1)\end{aligned}$$

$$\leq 6D\sqrt{L\ln N\sum_{t=1}^{T} f_t(\mathbf{x}_t)} + \mathcal{O}(1), \tag{16}$$

where the first inequality holds because we have $\|\boldsymbol{\ell}_t\|_\infty^2 = \max_{i\in[N]}(\langle \nabla g_t(\mathbf{y}_t), \mathbf{y}_{t,i}\rangle)^2 \leq D^2\|\nabla g_t(\mathbf{y}_t)\|_2^2$ by Cauchy-Schwarz inequality and $\|\nabla g_t(\mathbf{y}_t)\|_2 \leq \|\nabla f_t(\mathbf{x}_t)\|_2 \leq G$ (see Theorem 1), the second inequality makes use of $\sqrt{a+b} \leq \sqrt{a}+\sqrt{b}$, the third inequality holds by the property of the surrogate loss (also via Theorem 1), and the last inequality is due to the self-bounding properties of smooth functions (see Lemma 12). Note that $\mathcal{O}(\ln N) = \mathcal{O}(\log\log T)$ can be treated as a constant following previous studies [Luo and Schapire, 2015; Gaillard et al., 2014]

Upper bound of base-regret. According to Lemma 7 and noticing that the comparator sequence $\mathbf{u}_1, \ldots, \mathbf{u}_T \in \mathcal{X} \subseteq \mathcal{Y}$ and the diameter of $\mathcal{Y}$ equals to $2D$ by definition, with slight modifications, we have the following dynamic regret bound.

$$\begin{aligned}
\sum_{t=1}^{T}\langle \nabla g_t(\mathbf{y}_t), \mathbf{y}_{t,i} - \mathbf{u}_t\rangle &\leq \frac{5D^2}{2\eta_i} + \frac{D}{\eta_i}\sum_{t=2}^{T}\|\mathbf{u}_t - \mathbf{u}_{t-1}\|_2 + \eta_i\sum_{t=1}^{T}\|\nabla g_t(\mathbf{y}_t)\|_2^2 \\
&\leq \frac{5D^2}{2\eta_i} + \frac{D}{\eta_i}\sum_{t=2}^{T}\|\mathbf{u}_t - \mathbf{u}_{t-1}\|_2 + \eta_i\sum_{t=1}^{T}\|\nabla f_t(\mathbf{x}_t)\|_2^2 \qquad (17) \\
&\leq \frac{5D^2}{2\eta_i} + \frac{D}{\eta_i}\sum_{t=2}^{T}\|\mathbf{u}_t - \mathbf{u}_{t-1}\|_2 + 4\eta_i L\sum_{t=1}^{T} f_t(\mathbf{x}_t), \qquad (18)
\end{aligned}$$

where the second inequality is due to the property of the surrogate loss (see Theorem 1) and the last one is due to the self-bounding property of smooth functions (see Lemma 12).

Note that the property of $\|\nabla g_t(\mathbf{y}_t)\|_2 \leq \|\nabla f_t(\mathbf{x}_t)\|_2$ (see Theorem 1) plays an important role in the above analysis. Although the surrogate functions $\{g_t\}_{t=1}^T$ are not guaranteed to be smooth and non-negative, we can upper bound its gradient norm by that defined over the original functions $\{g_t\}_{t=1}^T$, which are indeed smooth and non-negative. We thus can utilize the self-bounding properties to establish a small-loss bound for the meta-regret and base-regret.

Upper bound of dynamic regret. Plugging the above upper bounds of meta-regret and base-regret together, we achieve

$$\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{u}_t) \leq 6D\sqrt{L\ln N\sum_{t=1}^{T} f_t(\mathbf{x}_t)} + \frac{5D^2 + 2DP_T}{2\eta_i} + 4\eta_i L\sum_{t=1}^{T} f_t(\mathbf{x}_t) + \mathcal{O}(1), \tag{19}$$

which holds for any base-learner's index $i \in [N]$.

Next, we specify the base-algorithm $\mathcal{E}_i$ compared with. Indeed, we aim at choosing the one with step size closest to the (near-)optimal step size $\eta^* = \sqrt{\frac{5D^2+2DP_T}{1+8LF_T^{\mathbf{x}}}}$, where we denote by $F_T^{\mathbf{x}} = \sum_{t=1}^{T} f_t(\mathbf{x}_t)$ the cumulative loss of the decisions. By Assumption 1 and Assumption 2, we have $F_T^{\mathbf{x}} \in [0, GDT]$ and then the possible minimum optimal and maximum step size are

$$\eta_{\min} = \sqrt{\frac{5D^2}{1+8LGDT}}, \text{ and } \eta_{\max} = \sqrt{5D^2 + 2D^2 T}.$$

The construction of step size pool is by discretizing the interval $[\eta_{\min}, \eta_{\max}]$ with intervals with exponentially increasing length. The step size of each base-learner is designed to be monotonically increasing with respect to the index. Consequently, it is evident to verify that there exists an index $i^* \in [N]$ such that

$$\eta_{i^*} \leq \eta^* \leq \eta_{i^*+1} = 2\eta_{i^*}. \tag{20}$$

As the upper bounds of meta-regret and base-regret hold for any compared base-learner, we can choose the index as i^* in particular. Then the second and the third terms in the inequality (19) satisfy

$$\frac{5D^2 + 2DP_T}{2\eta_{i^*}} + 4\eta_{i^*} LF_T^{\mathbf{x}}$$

$$
\begin{aligned}
&\le \frac{5D^2 + 2DP_T}{\eta^*} + 4\eta^* L F_T^{\mathbf{x}} \\
&\le \sqrt{(5D^2 + 2DP_T)(1 + 8LF_T^{\mathbf{x}})} + \frac{1}{2}\sqrt{(5D^2 + 2DP_T)(1 + 8LF_T^{\mathbf{x}})} \\
&\le 3\sqrt{2(5D^2 + 2DP_T)(1 + LF_T^{\mathbf{x}})}. \qquad (21)
\end{aligned}
$$

Substituting inequality (21) into inequality (19), we have,

$$
\begin{aligned}
&\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{u}_t) \\
&\le 6D\sqrt{L \ln N F_T^{\mathbf{x}}} + 3\sqrt{2(5D^2 + 2DP_T)(1 + LF_T^{\mathbf{x}})} + \mathcal{O}(1) \\
&\le \left(6D\sqrt{L \ln N} + 3\sqrt{2L(5D^2 + 2DP_T)}\right)\sqrt{F_T^{\mathbf{x}}} + 3\sqrt{2(5D^2 + 2DP_T)} + \mathcal{O}(1) \\
&\le \mathcal{O}\big(\sqrt{(1 + P_T)(F_T + \sqrt{P_T} + \mathcal{O}(1))} + P_T + 1\big) \\
&= \mathcal{O}\big(\sqrt{(F_T + P_T)(1 + P_T)}\big),
\end{aligned}
$$

where the last inequality holds by Lemma 17. Hence, we complete the proof of Theorem 3. □

C Omitted Details for Adaptive Regret Minimization

In this section, we present omitted details for minimizing the worst-case and small-loss adaptive regret. First, in Appendix C.1 we describe the efficient algorithm for attaining the worst-case adaptive regret, which provably enjoys the same guarantee as the prior best known work and meanwhile requires one projection only per iteration. Next, we focus on the proof of the main theorem for small-loss adaptive regret, i.e., Theorem 4. To this end, Appendix C.2 provides three key lemmas, Appendix C.3 presents the proof of Theorem C.3, and Appendix C.4 – Appendix C.6 give the proofs of three key lemmas.

C.1 Results for the Worst-Case Adaptive Regret

In this section, we present results for the worst-case adaptive regret omitted in the main text. The best known result is the $\mathcal{O}(\sqrt{|I| \log T})$ worst-case adaptive regret attained by the CBCE algorithm [Jun et al., 2017], which is achieved by the coin-betting framework with the sleeping expert mechanism. However, CBCE requires multiple gradients at each round. Wang et al. [2018] improve CBCE by using the linearized loss to make the algorithm requiring one gradient per iteration and retaining the same adaptive regret. Moreover, the improved CBCE algorithm only evaluates the function value once per iteration, and the algorithm is presented in [Wang et al., 2018, Algorithm 4 and Algorithm 5] . Therefore, we can simply feed the surrogate loss g_t constructed in (5) to the improved CBCE of Wang et al. [2018], and the obtained algorithm can ensure the same order of adaptive regret and also require only one projection onto the feasible domain per iteration.

C.2 Key Lemmas

In this part, we present three key lemmas for proving the small-loss adaptive regret, namely, Theorem 4. We then prove Theorem 4 based on these lemmas in Appendix C.3. Finally, we present the proofs for those lemmas in the following several subsections.

The first lemma gives the upper bound of meta-regret of our efficient method for small-loss adaptive regret, which heavily relies on the structure of the problem-dependent geometric covers.

Lemma 4. *Under Assumptions 2 and 3, for any interval $I = [i, j] \in \widetilde{\mathcal{C}}$ in the geometric covers defined in Eq. (10) on which we suppose m-th base-learner is active, Algorithm 2 ensures*

$$
\sum_{\tau=i}^{t} \langle \nabla g_t(\mathbf{y}_t), \mathbf{y}_t - \mathbf{y}_{t,m} \rangle \le \mathcal{O}\left(\sqrt{\log(m) \sum_{\tau=i}^{t} f_\tau(\mathbf{x}_\tau)}\right),
$$

which holds for any time stamp $t \in [i, j]$.

Combining above lemma and the analysis of the base-regret upper bound, we can obtain the following adaptive regret for any interval of the problem-dependent geometric covers.

Lemma 5. *Under Assumptions 1, 2, and 3, for any interval $[i, j] \in \widetilde{\mathcal{C}}$ in the geometric covers defined in Eq. (10), on which we assume m-th expert-algorithm is active, Algorithm 2 ensures*

$$\sum_{\tau=i}^{t} f_\tau(\mathbf{x}_\tau) - \sum_{\tau=i}^{t} f_\tau(\mathbf{u}) \leq \mathcal{O}\left(\sqrt{\log(m) \sum_{\tau=i}^{t} f_\tau(\mathbf{u})}\right),$$

which holds for any time stamp $t \in [i, j]$ and any comparator $\mathbf{u} \in \mathcal{X}$.

The above two lemmas rely on the unknown variable of m, and the following lemma presents an upper bound for m in terms of the small-loss quantity F_t.

Lemma 6. *Under Assumptions 1, 2, and 3, for any interval $[i, j] \in \widetilde{\mathcal{C}}$ and any $t \in [i, j]$, the variable m specified in Lemma 4 and Lemma 5 can be bounded by*

$$m \leq \mathcal{O}\left(F_{[1,t]}\right). \tag{22}$$

This immediately implies that Algorithm 2 ensures

$$\sum_{\tau=i}^{t} f_\tau(\mathbf{x}_\tau) - \min_{\mathbf{u} \in \mathcal{X}} \sum_{\tau=i}^{t} f_\tau(\mathbf{u}) \leq \mathcal{O}\left(\sqrt{F_{[i,t]} \log F_{[1,t]}}\right),$$

where $F_{[a,b]} = \min_{\mathbf{u} \in \mathcal{X}} \sum_{\tau=a}^{b} f_\tau(\mathbf{u})$ denotes the cumulative loss of the comparator within the interval $[a, b] \subseteq [T]$.

C.3 Proof of Theorem 4

Proof. Recall that Theorem 4 exhibits an $\mathcal{O}\big(\min\{\sqrt{F_I \log F_I \log F_T}, \sqrt{|I| \log T}\}\big)$ adaptive regret for any interval $I = [r, s] \subseteq [T]$, where $F_I = \min_{\mathbf{x} \in \mathcal{X}} \sum_{t=r}^{s} f_t(\mathbf{x})$ and $F_T = \min_{\mathbf{x} \in \mathcal{X}} \sum_{t=1}^{T} f_t(\mathbf{x})$. The bound consists of two parts, including a small-loss bound of $\mathcal{O}(\sqrt{F_I \log F_I \log F_T})$ and a worst-case bound of $\mathcal{O}(\sqrt{|I| \log T})$. Below, we present the proofs of the two bounds respectively.

Before showing the proofs, we emphasize again that our result strictly improves the small-loss bound of Zhang et al. [2019], who give an $\mathcal{O}(\sqrt{F_I \log F_I \log F_T})$ bound that becomes $\mathcal{O}(\sqrt{|I| \log|I| \log T})$ in the worst case and thus is looser than the $\mathcal{O}(\sqrt{|I| \log T})$ problem-independent bound [Jun et al., 2017] by a factor of $\log|I|$. Our regret guarantee consists of the small-loss bound and another worst-case bound acting as a safety guarantee for the worst case. Indeed, in the worst-case situation, our bound becomes $\mathcal{O}(\sqrt{|I| \log T})$ and *strictly* match the best known worst-case result [Jun et al., 2017]. We note that our improvement is owing to a refined analysis in the proof as well as our careful algorithm design that only uses one function evaluations to adaptively determine the geometric covers, which is to be contrasted to SACS [Zhang et al., 2019] that uses the latest base-learner's decision to determine the covers.

Small-loss regret bound. Let s_p be the smallest marker that larger than r, and let s_q be the largest marker that is not large than s, then we have

$$s_{p-1} \leq r < s_p, \text{ and } s_q \leq s < s_{q+1}.$$

We bound the regret over the interval $[r, s_p - 1]$ as,

$$\sum_{t=r}^{s_p-1} f_t(\mathbf{x}_t) - \sum_{t=r}^{s_p-1} f_t(\mathbf{u}) \leq \sum_{t=r}^{s_p-1} f_t(\mathbf{x}_t) \leq \sum_{t=s_{p-1}}^{s_p-1} f_t(\mathbf{x}_t) \leq C_{p-1} + GD. \tag{23}$$

The last inequality is because of the construction rule of marker and the fact that $f_t(\mathbf{x}_t) \leq GD$ for any $t \in [T]$ by Assumptions 1 – 3.

By Lemma 11, we can find v consecutive intervals

$$I_1 = [s_{i_1}, s_{i_2} - 1],\ I_2 = [s_{i_2}, s_{i_3} - 1],\ \ldots, I_v = [s_{i_v}, s_{i_{v+1}} - 1] \in \widetilde{\mathcal{C}}, \tag{24}$$

such that

$$i_1 = p,\ i_v \leq q < i_{v+1}, \text{and } v \leq \lceil \log_2(q - p + 2) \rceil.$$

Notice that,

$$q < i_{v+1} \Rightarrow q + 1 \leq i_{v+1} \Rightarrow s_{q+1} - 1 \leq s_{i_{v+1}} - 1 \Rightarrow s \leq s_{i_{v+1}} - 1.$$

For the neat presentation, we define,

$$\alpha(t) = (27GD + 72D^2L)\ln\left(3 + \frac{8}{C_1}\sum_{\tau=1}^{t} f_\tau(\mathbf{u})\right) + 72D^2L\mu^2(t) + 9GD\mu(t) + 6D\sqrt{\delta} + 288D^2L,$$

$$\beta(t) = 4D\sqrt{L}\left(\sqrt{\ln\left(3 + \frac{8}{C_1}\sum_{\tau=1}^{t} f_\tau(\mathbf{u})\right)} + \frac{\mu(t)}{\sqrt{\ln(3 + 8(\sum_{\tau=1}^{t} f_\tau(\mathbf{u}))/C_1)}} + 2\right),$$

where we use the notations $\mu(t) = \ln(1+(1+\ln(1+t))/(2e)) = \mathcal{O}(\log\log t)$ that can be essentially regarded as a constant and $C_1 = \mathcal{G}(1)$, where $\mathcal{G}(\cdot)$ is the threshold generating function defined in (46).

For intervals I_1 to I_v, by Lemma 5,

$$\begin{aligned}
\sum_{t=s_p}^{s} f_t(\mathbf{x}_t) - \sum_{t=s_p}^{s} f_t(\mathbf{u}) &\leq \sum_{k=1}^{v-1}\left(\alpha(s) + \beta(s)\sqrt{F_{I_k}}\right) + \alpha(s) + \beta(s)\sqrt{F_{[s_{i_v}, s]}} \qquad (25)\\
&\leq v\alpha(s) + \beta(s)\sqrt{vF_{[s_p, s]}}\\
&\leq v\alpha(s) + \beta(s)\sqrt{vF_I}. \qquad (26)
\end{aligned}$$

Combining (23) and (26), the adaptive regret on any interval $i = [r, s]$ will be

$$\sum_{t=r}^{s} f_t(\mathbf{x}_t) - \sum_{t=r}^{s} f_t(\mathbf{u}) \leq v\alpha(s) + \beta(s)\sqrt{vF_I} + C_{p-1} + GD. \tag{27}$$

Next, we show that v and C_{p-1} are of order $\mathcal{O}(\log F_{[r,s]})$ and $\mathcal{O}(\log F_T)$ respectively. By the definition of the time-varying threshold (see the threshold generating function Eq. (46)) and Lemma 6, the threshold can be bounded as,

$$C_{p-1} \leq (54GD{+}168D^2L)\ln\left(3 + \frac{8}{C_1}F_{[1,r]}\right){+}168D^2L\mu^2(T){+}18GD\mu(T){+}6D\sqrt{\delta}{+}672D^2L,$$

which is of order $\mathcal{O}(\log F_T)$.

With the same argument as (48), it can be shown that between markers s_p and s_q, for any $\mathbf{u}' \in \mathcal{X}$,

$$\sum_{t=s_p}^{s_q-1} f_t(\mathbf{u}') \geq \frac{C_1}{4}(q - p),$$

which suggests

$$q - p \leq \frac{4}{C_1}\sum_{t=s_p}^{s_q-1} f_t(\mathbf{u}') \leq \frac{4}{C_1}\sum_{t=r}^{s} f_t(\mathbf{u}').$$

We thus have

$$v \leq \lceil \log_2(q - p + 2) \rceil \leq \left\lceil \log_2\left(\frac{4}{C_1}F_{[r,s]} + 2\right)\right\rceil = \mathcal{O}(\log F_{[r,s]}).$$

Combining the upper bounds of C_{p-1} and v as well as the adaptive regret bound in (27) yields

$$\sum_{t=r}^{s} f_t(\mathbf{x}_t) - \sum_{t=r}^{s} f_t(\mathbf{u}) \le \mathcal{O}(\log F_I \log F_T) + \mathcal{O}(\sqrt{F_I \log F_I \log F_T}) + \mathcal{O}(\log F_T) + \mathcal{O}(1)$$
$$= \mathcal{O}\left(\sqrt{F_I \log F_I \log F_T}\right),$$

where the last step is true as we follow the same convention in [Zhang et al., 2019] to treat the $\log F_I \log F_T$ as the non-leading term. Hence, we finish the proof of the small-loss adaptive regret.

Worst-case regret bound. The above proof aims at obtaining small-loss type regret bound, and one of the key steps is to use Cauchy-Schwarz inequality to bound (25), which results in an additional $\mathcal{O}(\sqrt{\log F_{[r,s]}})$ term. Next, we show that actually this extra term can be avoided by the new design thresholds mechanism and thus asymptotically achieve the same worst-case adaptive regret as the best known result [Jun et al., 2017].

From (43) in Lemma 5, we have that for any interval $I = [i, j]$ in problem-dependent covers defined in (10), the adaptive regret is at most

$$\sum_{t=i}^{j} f_t(\mathbf{x}_t) - \sum_{t=i}^{j} f_t(\mathbf{u}) \le \mathcal{O}\left(\sqrt{\log T \cdot F^{\mathbf{x}}_{[i,j]}} + \log T\right),$$

where we use the notation $F^{\mathbf{x}}_{[a,b]} = \sum_{t=a}^{b} f_t(\mathbf{x}_t)$ to denote the cumulative loss of the returned decisions within the interval $[a, b] \subseteq [T]$. Then, we can use Lemma 6 to upper bound $m \le \mathcal{O}(T)$ as only the worst-case behavior matters now.

Moreover, for the consecutive intervals defined in (24), we have the following facts:

$$i_{k+1} \le 2 \cdot i_k,\ \forall k \in [v],\ \text{and}\ |i_{l+1} - i_l| \le \frac{1}{2}|i_{l+2} - i_{l+1}|,\ \forall l \in [v-1].$$

The first relationship between consecutive foot-indexes of time markers will be used to show that thresholds will not grow too fast during an interval in the geometric covers, which can verified by the construction of cover defined in (10). The second inequality indicates that the times the cumulative loss of algorithm exceeds thresholds in an interval decreases exponentially from I_v to I_1, and this can be verified in the proof of [Zhang et al., 2019, Lemma 11].

For the interval I_k with $k \in [v-1]$ in (24), our algorithm's cumulative loss within the interval is upper bounded by

$$\sum_{t=s_{i_k}}^{s_{i_{k+1}}-1} f_t(\mathbf{x}_t) \le \left(\sum_{a=i_k}^{i_{k+1}-1} C_a\right) + GD|i_{k+1} - i_k| \le (GD + C_{i_{k+1}-1})|i_{k+1} - i_k|. \tag{28}$$

We then split a given interval $[r, s]$ into three parts to analyze, namely, the consecutive $v-1$ intervals I_1 to I_{v-1}, interval $[r, s_p - 1]$, and $[s_{i_v}, s]$, where notably the last two intervals are not fully covered by any interval in geometric covers. For intervals I_1 to I_{v-1}, we have

$$\begin{aligned}
\sum_{t=s_{i_1}}^{s_{i_v}-1} f_t(\mathbf{x}_t) - \sum_{t=s_{i_1}}^{s_{i_v}-1} f_t(\mathbf{u}) &\le \sum_{a=1}^{v-1} \mathcal{O}\left(\sqrt{\log T \cdot F^{\mathbf{x}}_{I_a}} + \log T\right) \\
&\le \sum_{a=1}^{v-1} \mathcal{O}\left(\sqrt{\log T \cdot C_{i_v-1} \cdot |i_{a+1} - i_a|} + \log T\right) \\
&\le \sum_{a=1}^{v-1} \mathcal{O}\left(\sqrt{\log T \cdot C_{i_v-1} \cdot \frac{|i_v - i_{v-1}|}{2^{v-1-a}}} + \log T\right) \\
&\le \mathcal{O}\left(v \log T + \sqrt{\log T \cdot C_{i_v-1}} \cdot \sum_{b=0}^{+\infty} \sqrt{\frac{|i_v - i_{v-1}|}{2^b}}\right) \\
&\le \mathcal{O}\left(v \log T + \sqrt{\log T \cdot C_{i_v-1} \cdot |i_v - i_{v-1}|}\right),
\end{aligned}$$

where the second inequality is due to the monotonically increasing property of thresholds, the third inequality is by (28), and the last inequality is by the summation of geometric sequence.

By the setting of time-varying thresholds as specified in Eq. (46), we know $C_{i_v-1} = \mathcal{O}(\log(i_v))$.

Furthermore, $|i_v - i_{v-1}|$ represents the number of markers that our algorithm generates during the interval I_{v-1}, so it can be upper bounded as

$$|i_v - i_{v-1}| \leq \mathcal{O}\left(\frac{GD|I|}{C_{i_{v-1}}}\right) = \mathcal{O}\left(\frac{GD|I|}{\log(i_{v-1})}\right),$$

because the total loss of the algorithm during $|I|$ is at most $GD|I|$ and we use the smallest threshold $C_{i_{v-1}}$ during the interval I_{v-1} to calculate the maximum number of possible markers. Due to the relationship of $2i_{v-1} \geq i_v$ by (28), we then get

$$\begin{aligned}
\sum_{t=s_{i_1}}^{s_{i_v}-1} f_t(\mathbf{x}_t) - \sum_{t=s_{i_1}}^{s_{i_v}-1} f_t(\mathbf{u}) &\leq \mathcal{O}\left(v \log T + \sqrt{\log T \cdot C_{i_v-1} \cdot |i_v - i_{v-1}|}\right) \\
&\leq \mathcal{O}\left(v \log T + \sqrt{\log T \cdot \log(i_v) \cdot \frac{|I|}{\log(i_{v-1})}}\right) \\
&\leq \mathcal{O}\left(v \log T + \sqrt{\log T \cdot |I|(1 + \frac{1}{\log i_{v-1}})}\right) \\
&\leq \mathcal{O}\left(\log |I| \log T + \sqrt{|I| \log T}\right),
\end{aligned}$$

where the last inequality is due to the fact that v is of order $\mathcal{O}(\log F_{[r,s]}) = \mathcal{O}(\log |I|)$. Remind that the variable v appears in our analysis by Lemma 11, which is independent of the worst-case analysis.

Now we proceed to upper bound the regret over intervals $[r, s_p - 1]$ and $[s_{i_v}, s]$. By a similar analysis used early, we have

$$\sum_{t=r}^{s_p-1} f_t(\mathbf{x}_t) - \sum_{t=r}^{s_p-1} f_t(\mathbf{u}) \leq C_{p-1} \leq \mathcal{O}(\log T),$$

and

$$\sum_{t=s_{i_v}}^{s} f_t(\mathbf{x}_t) - \sum_{t=s_{i_v}}^{s} f_t(\mathbf{u}) \leq \mathcal{O}\left(\log T + \sqrt{F_{[s_{i_v},s]} \log T}\right) \leq \mathcal{O}\left(\log T + \sqrt{|I| \log T}\right).$$

Combining everything together achieves

$$\begin{aligned}
&\sum_{t=r}^{s} f_t(\mathbf{x}_t) - \sum_{t=r}^{s} f_t(\mathbf{u}) \\
&\leq \mathcal{O}\left(\sqrt{|I| \log T} + \log |I| \log T\right) = \mathcal{O}\left(\sqrt{(|I| + \log T \cdot \log^2 |I|) \log T}\right) = \mathcal{O}(\sqrt{|I| \log T}).
\end{aligned}$$

The last step holds by considering the following cases.

- When the interval length is $|I| = \Theta(T^\alpha)$ with $\alpha \in (0, 1]$. Then,

$$\begin{aligned}
&\mathcal{O}\left(\sqrt{(|I| + \log T \cdot \log^2 |I|) \log T}\right) \\
&= \mathcal{O}\left(\sqrt{(T^\alpha + \alpha^2 \log^3 T) \log T}\right) \\
&= \mathcal{O}\left(\sqrt{T^\alpha \log T}\right) = \mathcal{O}(\sqrt{|I| \log T}).
\end{aligned}$$

- When the interval length is $|I| = \Theta(\log^\beta T)$, and note that $\beta \in [1, +\infty)$ as $|I| = \Omega(\log T)$ is the minimum order to ensure the adaptive regret to be non-trivial. Then,

$$\mathcal{O}\left(\sqrt{(|I| + \log T \cdot \log^2 |I|) \log T}\right)$$

$$= \mathcal{O}\left(\sqrt{\left(\log^{\beta} T + \beta^2 \log T \cdot (\log\log T)^2\right) \log T}\right)$$
$$= \mathcal{O}\left(\sqrt{(\log^{\beta} T + \beta^2 \log T) \log T}\right)$$
$$= \mathcal{O}\left(\sqrt{\log^{\beta} T \log T}\right) = \mathcal{O}(\sqrt{|I| \log T}),$$

where the second equality is true for $\beta > 1$ and otherwise we can treat $\mathcal{O}(\log\log T)$ as a constant following previous studies [Luo and Schapire, 2015; Gaillard et al., 2014].

Hence we finish the proof for the worst-case adaptive regret bound. Combining both small-loss bound and the worst-case safety guarantee, we complete the proof of Theorem 4.

Discussions about the analysis of worst-case bound. We point out that the same analysis for worst-case adaptive bound cannot be applied to SACS [Zhang et al., 2019] directly, because the new designed *time-varying thresholds* mechanism plays an important role. Indeed, SACS monitors the cumulative loss of the new-added base-learner and switches to the another one immediately after a new marker is registered, which causes intermittent performance monitoring for the remaining markers the base-learner may go through. Thus, we cannot employ a similar argument (like Eq. (28) in our analysis) to bound the cumulative loss by the summation of thresholds for adaptive regret of SACS, as the performance of base-learner on the whole active interval is absent.

Finally, we would like to emphasize the necessity of monitoring the final decisions' cumulative loss. To meet the one projection requirement, our base-algorithm is updated over the surrogate loss and the provided gradient information is about the final decision only. Thus the thresholds are set on the cumulative loss of final decisions to exploit the limited available information. □

C.4 Proof of Lemma 4

Proof. First we introduce some useful variables to help us prove the adaptivity of Adapt-ML-Prod under sleeping-expert setting. Similar to the proof technique proposed in [Daniely et al., 2015], for any interval $[i, j] \in \widetilde{\mathcal{C}}$ in the geometric covers defined in (10), on which we suppose m-th base-learner is active, we define the following pseudo-weight for the m-th base-learner,

$$\widetilde{w}_{\tau,m} = \begin{cases} 0 & \tau < i, \\ 1 & \tau = i, \\ \left(\widetilde{w}_{\tau-1,m}(1 + \eta_{\tau-1}(\widehat{\ell}_{\tau-1} - \ell_{\tau-1,m}))\right)^{\frac{\eta_{\tau,m}}{\eta_{\tau-1,m}}} & i < \tau \leq j+1, \\ \widetilde{w}_{j+1,m} & \tau > j+1. \end{cases}$$

In addition, we use $\widetilde{W}_t = \sum_{k \in [T]} \widetilde{w}_{t,k}$ to denote the summation of pseudo-weights for all possible base-learners up to time t. As for the problem-dependent geometric covers, in the worst case there are at most T base-learners generated, we use $[T]$ to denote the indexes for all the base-learners. Notice that the pseudo-weight $\widetilde{w}_t$ is defined as 0 for asleep base-learners till time t, so we can include all possible ones safely in the definition even though they are not generated in practical implementations of the algorithm.

In the following, we use the classic potential argument [Gaillard et al., 2014] by showing both lower bound and upper bound of $\ln \widetilde{W}_{t+1}$ to establish relationships between certain concerned quantities.

Lower bound of $\ln \widetilde{W}_{t+1}$. We claim that for $t \in [i, j]$ it holds that

$$\ln \widetilde{w}_{t+1,m} \geq \eta_{t+1,m} \sum_{\tau=i}^{t} (r_{\tau,m} - \eta_{\tau,m} r_{\tau,m}^2). \tag{29}$$

We prove the above inequality by induction on t. When $t = i$, by definition,

$$\ln \widetilde{w}_{i+1,m} = \frac{\eta_{i+1,m}}{\eta_{i,m}} \ln(1 + \eta_m r_{i,m}) \geq \frac{\eta_{i+1,m}}{\eta_{i,m}} \left(\eta_m r_{i,m} - \eta_m^2 r_{i,m}^2\right) = \eta_{i+1,m}(r_{i,m} - \eta_m r_{i,m}^2),$$

where the inequality is because of $\ln(1+x) \geq x - x^2, \forall x \geq -1/2$.

Suppose the statement holds for $\ln \widetilde{w}_{t,m}$, then we proceed to check the situation for $t+1$ round as follows. Indeed,

$$\begin{aligned}
\ln \widetilde{w}_{t+1,m} &= \frac{\eta_{t+1,m}}{\eta_{t,m}} \left(\ln \widetilde{w}_{t,m} + \ln\left(1+\eta_{t,m} r_{t,m}\right)\right) \\
&\geq \frac{\eta_{t+1,m}}{\eta_{t,m}} \left(\ln \widetilde{w}_{t,m} + \eta_{t,m} r_{t,m} - \eta_{t,m}^2 r_{t,m}^2\right) \\
&= \frac{\eta_{t+1,m}}{\eta_{t,m}} \ln \widetilde{w}_{t,m} + \eta_{t+1,m}\left(r_{t,m} - \eta_{t,m} r_{t,m}^2\right) \\
&\geq \frac{\eta_{t+1,m}}{\eta_{t,m}} \left(\eta_{t,m} \sum_{\tau=i}^{t-1} (r_{\tau,m} - \eta_{\tau,m} r_{\tau,m}^2)\right) + \eta_{t+1,m}\left(r_{t,m} - \eta_{t,m} r_{t,m}^2\right) \\
&= \eta_{t+1,m} \sum_{\tau=i}^{t} (r_{\tau,m} - \eta_{\tau,m} r_{\tau,m}^2). \qquad (30)
\end{aligned}$$

Then, as $\widetilde{w}_{t+1,m}$ is positive for any m-th base-learner, we have $\ln \widetilde{W}_{t+1} \geq \ln \widetilde{w}_{t+1,m}$. Combining (30) obtains the desired lower bound of $\ln \widetilde{W}_{t+1}$.

Upper bound of $\ln \widetilde{W}_{t+1}$. By the construction of the geometric covers as specified in Eq. (10), we know that there will be at most $2m$ base-learners initialized for the m-th base-learner active on interval $[i,j]$ till her end. This is because m-th base-learner is initialized when m-th marker is recorded, and she will expire before the moment when $2m$-th marker is recorded, as demonstrated by the construction of cover defined in Eq. (10). Owing to this property, we have $\widetilde{W}_{t+1} = \sum_{k\in[2m]} \widetilde{w}_{t+1,k}$ as others' pseudo-weight equals to 0 by definition. So we can upper bound $\widetilde{W}_{t+1}$ as,

$$\begin{aligned}
\widetilde{W}_{t+1} = \sum_{k\in[2m]} \widetilde{w}_{t+1,k} &= \sum_{k\in[2m]:i_k=t+1} \widetilde{w}_{t+1,k} + \sum_{k\in[2m]:i_k\leq t} \widetilde{w}_{t+1,k} \\
&= \mathbb{1}\{\text{new alg. at } t+1\} + \sum_{k\in[2m]:i_k\leq t} \widetilde{w}_{t+1,k}, \qquad (31)
\end{aligned}$$

where with a slight abuse of notations, we denote by $[i_k, j_k] \in \widetilde{\mathcal{C}}$ the active time for k-th base-learner.

For the second term in (31), we have

$$\begin{aligned}
\sum_{k:i_k\leq t} \widetilde{w}_{t+1,k} &= \sum_{k\in[2m]:t\in[i_k,j_k]} \widetilde{w}_{t+1,k} + \sum_{k\in[2m]:t>j_k} \widetilde{w}_{t+1,k} \\
&= \sum_{k\in[2m]:t\in[i_k,j_k]} \widetilde{w}_{t+1,k} + \sum_{k\in[2m]:t>j_k} \widetilde{w}_{t,k} \\
&\leq \sum_{k\in[2m]:t\in[i_k,j_k]} \widetilde{w}_{t,k}(1+\eta_{t,k} r_{t,k}) + \frac{1}{e}\left(\frac{\eta_{t,k}}{\eta_{t+1,k}} - 1\right) + \sum_{k\in[2m]:t>j_k} \widetilde{w}_{t,k} \\
&= \widetilde{W}_t + \underbrace{\sum_{k\in[2m]:t\in[i_k,j_k]} \eta_{t,k} \widetilde{w}_{t,k} r_{t,k}}_{=0} + \sum_{k\in[2m]:t\in[i_k,j_k]} \frac{1}{e}\left(\frac{\eta_{t,k}}{\eta_{t+1,k}} - 1\right), \qquad (32)
\end{aligned}$$

where the first equality holds by the definition of $\widetilde{w}_{t+1,k}$, the second inequality is by the updating rule of $\widetilde{w}_{t+1,k}$ and Lemma (18), and the second term in the last equality equals to 0 due to the weight update rule in (13) and the fact of $\widetilde{w}_{t,k} = w_{t,k}$ for any $t \in [i_k, j_k]$.

Combining (31), (32) and by induction, we obtain the following upper bound:

$$\widetilde{W}_{t+1} \leq 1 + 2m + \frac{1}{e} \sum_{k\in[2m]} \sum_{\tau=i_k}^{t\wedge j_k} \left(\frac{\eta_{\tau,k}}{\eta_{\tau+1,k}} - 1\right), \qquad (33)$$

where we denote $\alpha \wedge \beta = \min\{\alpha, \beta\}$.

We now turn to analyze the third term in (32). Indeed, Gaillard et al. [2014] have analyzed it under the static regret measure. For the sake of completeness, we present the proof with our notations. For any $k \in [2m]$, for any $\tau \in [i_k, t \wedge j_k]$, the relationship between $\eta_{\tau,k}$ and $\eta_{\tau+1,k}$ can be considered as three cases,

- $\eta_{\tau,k} = \eta_{\tau+1,k} = 1/2$,
- $\eta_{\tau+1,k} = \sqrt{\gamma_k/(1+\sum_{u=i_k}^{\tau} r_{u,k}^2)} < \eta_{\tau,k} = \frac{1}{2}$,
- $\eta_{\tau+1,k} \leq \eta_{\tau,k} < 1/2$.

In all cases, the ratio $\eta_{\tau,k}/\eta_{\tau+1,k} - 1$ is at most

$$
\begin{aligned}
\sum_{\tau=i_k}^{t\wedge j_k}\left(\frac{\eta_{\tau,k}}{\eta_{\tau+1,k}} - 1\right) &\leq \sum_{\tau=i_k}^{t\wedge j_k}\left(\sqrt{\frac{1+\sum_{u=i_k}^{\tau} r_{u,k}^2}{1+\sum_{u=i_k}^{\tau-1} r_{u,k}^2}} - 1\right)\\
&= \sum_{\tau=i_k}^{t\wedge j_k}\left(\sqrt{\frac{r_{\tau,k}^2}{1+\sum_{u=i_k}^{\tau-1} r_{u,k}^2} + 1} - 1\right)\\
&\leq \frac{1}{2}\sum_{\tau=i_k}^{t\wedge j_k}\frac{r_{\tau,k}^2}{1+\sum_{u=i_k}^{\tau-1} r_{u,k}^2}\\
&\leq \frac{1}{2}\left(1 + \ln\left(1 + \sum_{u=i_k}^{t\wedge j_k} r_{u,k}^2\right)\right) - \ln(1)\\
&\leq \frac{1}{2}\left(1 + \ln(1+t)\right), \qquad (34)
\end{aligned}
$$

where the second inequality uses $\sqrt{1+x} \leq 1 + x/2$ and the third inequality follows from Lemma 14 with the choice of $f(x) = 1/x$.

Substituting (34) into (33), we get

$$
\widetilde{W}_{t+1} \leq 1 + 2m + \frac{m}{e}\left(1 + \ln\left(1+t\right)\right) \leq (1+2m)\left(1 + \frac{1}{2e}\left(1 + \ln(1+t)\right)\right). \qquad (35)
$$

Further taking the logarithm over the above inequality gives the upper bound of $\ln \widetilde{W}_{t+1}$.

Upper bound of meta-regret. Now, we can lower bound and upper bound $\ln \widetilde{W}_{t+1}$ by (30) and (35), with arrangement, which yields the upper bound of scaled meta-regret. Concretely,

$$
\begin{aligned}
\sum_{\tau=i}^{t} r_{\tau,m} &\leq \sum_{\tau=i}^{t} \eta_{\tau,m} r_{\tau,m}^2 + \frac{\ln(1+2m) + \mu(t)}{\eta_{t+1,m}}\\
&\leq 2\sqrt{\gamma_i}\sqrt{1 + \sum_{\tau=i}^{t} r_{\tau,i}^2} + \frac{\ln(1+2m) + \mu(t)}{\eta_{t+1,m}} \qquad (36)\\
&\leq \frac{\ln(1+2m) + \mu(t) + 2\gamma_m}{\sqrt{\gamma_m}}\sqrt{1 + \sum_{\tau=i}^{t} r_{\tau,m}^2 + 2\ln(1+2m) + 4\gamma_m + 2\mu(t)}\\
&= \left(3\sqrt{\ln(1+2m)} + \frac{\mu(t)}{\sqrt{\ln(1+2m)}}\right)\sqrt{1 + \sum_{\tau=i}^{t} r_{\tau,m}^2 + 6\ln(1+2m) + 2\mu(t)}, \qquad (37)
\end{aligned}
$$

where we denote $\mu(t) = \ln(1 + (1 + \ln(1+t))/(2e))$.The second inequality is by Lemma 14 and choose $f(x) = 1/\sqrt{x}$. The last equality is by the choice of $\gamma_m = \ln(1+2m)$. As for the third inequality, there are two cases to be considered:

- when $\sqrt{1+\sum_{\tau=i}^{t} r_{\tau,m}^2} > 2\sqrt{\gamma_m}$, we have that (36) is at most

$$2\sqrt{\gamma_m}\sqrt{1+\sum_{\tau=i}^{t} r_{\tau,m}^2} + \frac{\ln(1+2m)+\mu(t)}{\sqrt{\gamma_m}}\sqrt{1+\sum_{\tau=i}^{t} r_{\tau,m}^2}.$$

- when $\sqrt{1+\sum_{\tau=i}^{t} r_{\tau,m}^2} \le 2\sqrt{\gamma_m}$, we have that $\eta_{t+1,m} = 1/2$ and (36) is at most

$$2\ln(1+2m) + 4\gamma_m + 2\mu(t).$$

Taking both cases into account implies the third inequality.

Finally, we end the proof by evaluating the meta-regret in terms of the surrogate loss.

$$\begin{aligned}
&\sum_{\tau=i}^{t}\langle \nabla g_t(\mathbf{y}_t), \mathbf{y}_t - \mathbf{y}_{t,m}\rangle \\
&= 2GD\cdot\sum_{\tau=i}^{t} r_{\tau,m} \\
&\le 2GD\left(3\sqrt{\ln(1+2m)} + \frac{\mu(t)}{\sqrt{\ln(1+2m)}}\right)\sqrt{1+\sum_{\tau=i}^{t} r_{\tau,m}^2} + 12GD\ln(1+2m) + 4GD\mu(t) \\
&\le \left(3\sqrt{\ln(1+2m)} + \frac{\mu(t)}{\sqrt{\ln(1+2m)}}\right)\sqrt{\sum_{\tau=i}^{t}\langle \nabla g_\tau(\mathbf{y}_\tau), \mathbf{y}_\tau - \mathbf{y}_{\tau,m}\rangle^2} + 18GD\ln(1+2m) + 6GD\mu(t) \\
&\le \left(3\sqrt{\ln(1+2m)} + \frac{\mu(t)}{\sqrt{\ln(1+2m)}}\right)\sqrt{\sum_{\tau=i}^{t} 4D^2\|\nabla g_\tau(\mathbf{y}_\tau)\|_2^2} + 18GD\ln(1+2m) + 6GD\mu(t) \\
&\le 2D\left(3\sqrt{\ln(1+2m)} + \frac{\mu(t)}{\sqrt{\ln(1+2m)}}\right)\sqrt{\sum_{\tau=i}^{t}\|\nabla f_\tau(\mathbf{x}_\tau)\|_2^2} + 18GD\ln(1+2m) + 6GD\mu(t) \qquad (38) \\
&\le 4D\left(3\sqrt{\ln(1+2m)} + \frac{\mu(t)}{\sqrt{\ln(1+2m)}}\right)\sqrt{L\sum_{\tau=i}^{t} f_t(\mathbf{x}_t)} + 18GD\ln(1+2m) + 6GD\mu(t), \qquad (39)
\end{aligned}$$

where the second inequality is true because $1 \le \sqrt{\ln(1+2m)} \le \ln(1+2m)$ holds for any $m \ge 1$, the third inequality is by Cauchy-Schwarz inequality, the forth inequality is by Theorem 1 and the last inequality is due to the self-bounded property of smooth functions (see Lemma 12). □

C.5 Proof of Lemma 5

Proof. Similar to the proof of dynamic regret (see Theorem 3), we start the proof by decomposing the interval regret into meta-regret and base-regret in terms of the surrogate loss by Theorem 1,

$$\begin{aligned}
\sum_{\tau=i}^{t} f_\tau(\mathbf{x}_\tau) - \sum_{t=i}^{t} f_\tau(\mathbf{u}) \le \sum_{\tau=i}^{t} g_\tau(\mathbf{x}_\tau) - \sum_{t=i}^{t} g_\tau(\mathbf{u}) &\le \sum_{\tau=i}^{t}\langle \nabla g_\tau(\mathbf{y}_\tau), \mathbf{y}_\tau - \mathbf{u}\rangle \\
&= \underbrace{\sum_{\tau=i}^{t}\langle \nabla g_\tau(\mathbf{y}_\tau), \mathbf{y}_\tau - \mathbf{y}_{\tau,m}\rangle}_{\texttt{meta-regret}} + \underbrace{\sum_{\tau=i}^{t}\langle \nabla g_\tau(\mathbf{y}_\tau), \mathbf{y}_{\tau,m} - \mathbf{u}\rangle}_{\texttt{base-regret}}, \qquad (40)
\end{aligned}$$

where our analysis will be performed by tracking the m-th base-learner, whose corresponding active interval is exactly the analyzed one. Our analysis is satisfied to any interval since there is always a base-learner active on it by the algorithm design.

Upper bound of base-regret. Since the base-algorithm (SOGD) guarantees anytime regret, direct application of Lemma 8 with the assumption of surrogate domain $\mathcal{Y}$ can upper bound the base-regret,

$$\sum_{\tau=i}^{t}\langle\nabla g_\tau(\mathbf{y}_\tau),\mathbf{y}_{\tau,m}-\mathbf{u}\rangle \leq 4D\sqrt{\delta+\sum_{\tau=i}^{t}\|\nabla g_\tau(\mathbf{y}_\tau)\|_2^2} \leq 8D\sqrt{L\sum_{\tau=i}^{t} f_\tau(\mathbf{x}_\tau)}+4D\sqrt{\delta}, \quad (41)$$

where we skip some steps for transforming $\|\nabla g_\tau(\mathbf{y}_\tau)\|_2^2$ into $4Lf_\tau(\mathbf{x}_\tau)$. The similar arguments can be found in the proof of Theorem 3.

Upper bound of meta-regret. By Lemma 4 (see Eq. (39) for a detailed form), we can upper bound the meta-regret as

$$\begin{aligned}&\sum_{\tau=i}^{t}\langle\nabla g_\tau(\mathbf{y}_\tau),\mathbf{y}_\tau-\mathbf{y}_{\tau,m}\rangle\\ &\leq 4D\left(3\sqrt{\ln(1+2m)}+\frac{\mu(t)}{\sqrt{\ln(1+2m)}}\right)\sqrt{L\sum_{\tau=i}^{t} f_\tau(\mathbf{x}_\tau)}+18GD\ln(1+2m)+6GD\mu(t),\end{aligned} \quad (42)$$

where we denote $\mu(t)=\ln(1+(1+\ln(1+t))/(2e))$, which is of order $\mathcal{O}(\log\log t)$ and can be treated as a constant.

Upper bound of interval regret. Substituting (41), (42) into (40) obtains the interval regret

$$\sum_{\tau=i}^{t} f_\tau(\mathbf{x}_\tau)-\sum_{\tau=i}^{t} f_\tau(\mathbf{u}) \leq \mathcal{O}\left(\sqrt{\log m\cdot\sum_{\tau=i}^{t} f_\tau(\mathbf{x}_\tau)}+\log m\right), \quad (43)$$

which is related to the returned decisions of our algorithm.

Further, by applying Lemma 17 to the preceding inequality, we can substitute $\sum_{\tau=i}^{t} f_\tau(\mathbf{x}_\tau)$ into decision-independent factor $\sum_{\tau=i}^{t} f_\tau(\mathbf{u})$,

$$\begin{aligned}&\sum_{\tau=i}^{t} f_\tau(\mathbf{x}_\tau)-\sum_{\tau=i}^{t} f_\tau(\mathbf{u})\\
&\leq 4D\sqrt{L}\left(\sqrt{\ln(1+2m)}+\frac{\mu(t)}{\sqrt{\ln(1+2m)}}+2\right)\sqrt{\sum_{\tau=i}^{t} f_\tau(\mathbf{u})+18GD\ln(1+2m)+6GD\mu(t)+4D\sqrt{\delta}}\\
&\qquad+18GD\ln(1+2m)+6GD\mu(t)+4D\sqrt{\delta}+16D^2L\left(\sqrt{\ln(1+2m)}+\frac{\mu(t)}{\sqrt{\ln(1+2m)}}+2\right)^2\\
&\leq 4D\sqrt{L}\left(\sqrt{\ln(1+2m)}+\frac{\mu(t)}{\sqrt{\ln(1+2m)}}+2\right)\sqrt{\sum_{\tau=i}^{t} f_\tau(\mathbf{u})}\\
&\qquad+27GD\ln(1+2m)+9GD\mu(t)+6D\sqrt{\delta}+24D^2L\left(\sqrt{\ln(1+2m)}+\frac{\mu(t)}{\sqrt{\ln(1+2m)}}+2\right)^2\\
&\leq 4D\sqrt{L}\left(\sqrt{\ln(1+2m)}+\frac{\mu(t)}{\sqrt{\ln(1+2m)}}+2\right)\sqrt{\sum_{\tau=i}^{t} f_\tau(\mathbf{u})}\\
&\qquad+(27GD+72D^2L)\ln(1+2m)+72D^2L\mu^2(t)+9GD\mu(t)+6D\sqrt{\delta}+288D^2L.\end{aligned} \quad (44)$$

The second inequality makes use of $\sqrt{a+b}\leq\sqrt{a}+\sqrt{b}$ and $\sqrt{ab}\leq(a^2+b^2)/2$. The last inequality holds by $(a+b+c)^2\leq 3(a^2+b^2+c^2)$.

Finally, with a slight abuse of notations, we show that actually the interval regret can be related to the best offline cumulative loss, i.e., the quantity $F_{[i,t]}=\min_{\mathbf{u}\in\mathcal{X}}\sum_{\tau=i}^{t} f_\tau(\mathbf{u})$. Notice that, the above arguments hold for any $\mathbf{u} \in \mathcal{X}$, so we can simply choose $\mathbf{u}^* \in \arg\min_{\mathbf{u}' \in \mathcal{X}} \sum_{\tau=i}^{t} f_\tau(\mathbf{u}')$ as the comparator in the above arguments, and for any $\mathring{\mathbf{u}} \in \mathcal{X}$,

$$\sum_{\tau=i}^{t} f_\tau(\mathbf{x}_\tau) - \sum_{\tau=i}^{t} f_\tau(\mathring{\mathbf{u}}) \leq \sum_{\tau=i}^{t} f_\tau(\mathbf{x}_\tau) - \min_{\mathbf{u} \in \mathcal{X}} \sum_{\tau=i}^{t} f_\tau(\mathbf{u}) \leq \mathcal{O}\left(\sqrt{\log m \cdot F_{[i,t]}}\right).$$

This ends the proof. □

C.6 Proof of Lemma 6

Proof. We claim that the cumulative loss of our algorithm within every interval in the geometric covers $\widetilde{\mathcal{C}}$ as defined in Eq. (10) can be upper bounded by the cumulative loss of any comparator and the threshold set for the interval.

To see this, we denote by $[i_k, j_k] \in \widetilde{\mathcal{C}}$ the active interval for the k-th base-learner. By Lemma 5, for any $t' \in [i_k, j_k]$ we have

$$\sum_{\tau=i_k}^{t'} f_\tau(\mathbf{x}_\tau) \leq 2 \sum_{\tau=i_k}^{t'} f_\tau(\mathbf{u}) + \frac{1}{2}\mathcal{G}(k), \tag{45}$$

where $\mathcal{G}(k)$ is the threshold generating function defined by

$$\mathcal{G}(k) = (54GD + 168D^2L)\ln(1+2k) + 168D^2L\mu^2(T) + 18GD\mu(T) + 6D\sqrt{\delta} + 672D^2L, \tag{46}$$

with $\mu(T) = \ln(1 + \frac{1+\ln(1+T)}{2e})$. This is true by applying $ab \leq a^2/4 + b^2$ and $(a+b+c)^2 \leq 3(a^2+b^2+c^2)$ to split $\sum f_t(\mathbf{u})$ outside the root in (44).

We denote s_k the k-th marker made by the algorithm and it is known that $i_k = s_k$ (but $s_{k+1} - 1 \leq j_k$ because the base-learner may survive several markers) by the cover mechanism. According to our algorithm design, we set the threshold for interval $[s_k, s_{k+1} - 1]$ as

$$C_k = \mathcal{G}(k).$$

By the construction rule of the markers, we have,

$$\sum_{\tau=s_k}^{s_{k+1}-1} f_\tau(\mathbf{x}_\tau) \geq C_k.$$

For the k-th base-learner, her ending time j_k should be equal or larger than $s_{k+1} - 1$, so we can use (45) to evaluate the lower bound of the cumulative loss of the comparator by setting $t' = s_{k+1} - 1$. Indeed, we have

$$\sum_{\tau=s_k}^{s_{k+1}-1} f_\tau(\mathbf{u}) \geq \frac{1}{2}\left(\sum_{\tau=s_k}^{s_{k+1}-1} f_\tau(\mathbf{x}_\tau) - \frac{1}{2}\mathcal{G}(k)\right) \geq \frac{1}{2}\left(C_k - \frac{1}{2}\mathcal{G}(k)\right) = \frac{1}{2}\left(C_k - \frac{1}{2}C_k\right) = \frac{1}{4}C_k. \tag{47}$$

It is worth emphasizing that, after making each marker, the cover will initialize a new base-learner and hence the evaluation as shown above can be made between every two consecutive markers thanks to the streaming initialized new learners.

Next, we proceed to upper bound m given in the lemma. As the m-th base-learner is active on interval $[i, j]$, we have the following result on the interval from marker s_1 to s_m,

$$\sum_{\tau=s_1}^{s_m-1} f_\tau(\mathbf{u}) \geq \frac{1}{4}\sum_{a=1}^{m-1} C_a \geq \frac{C_1}{4}(m-1).$$

The first inequality is because of (47). The second inequality holds since C_a is increasing with respect to its index, see the threshold generating function in Eq. (46).

Therefore, rearranging the above inequality shows that the quantity m satisfies the following inequality for any comparator $\mathbf{u} \in \mathcal{X}$,

$$m \leq 1 + \frac{4}{C_1}\sum_{\tau=s_1}^{s_m-1} f_\tau(\mathbf{u}) \leq 1 + \frac{4}{C_1}\sum_{\tau=1}^{t} f_\tau(\mathbf{u}),$$

where the last inequality makes use of the nonnegative assumption on loss function. In particular, we choose the comparator to be the best offline decision optimizing the cumulative loss within this interval and thus achieve

$$m \leq 1 + \frac{4}{C_1} \min_{\mathbf{u}\in\mathcal{X}} \sum_{\tau=1}^{t} f_\tau(\mathbf{u}) = \mathcal{O}(F_{[1,t]}), \tag{48}$$

where we denote $F_{[1,t]} = \min_{\mathbf{u}\in\mathcal{X}} \sum_{\tau=1}^{t} f_\tau(\mathbf{u})$.

Now combining Lemma 5 and (48), once given time t, we can upper bound the number of base-learners by the cumulative loss of comparators,

$$\sum_{\tau=i}^{t} f_\tau(\mathbf{x}_\tau) - \sum_{\tau=i}^{t} f_\tau(\mathbf{u}) \leq \mathcal{O}\left(\sqrt{\log(m) F_{[i,t]}}\right) = \mathcal{O}\left(\sqrt{\ln(F_{[1,t]}) \cdot F_{[i,t]}}\right),$$

which ends the proof of Lemma 6. □

D Useful Lemmas

This section collects some lemmas useful for the proofs.

D.1 OGD and Dynamic Regret

This part provides the dynamic regret of online gradient descent (OGD) [Zinkevich, 2003] and scale-free online gradient descent (SOGD) [Orabona and Pál, 2018] from the view of online mirror descent (OMD), which is a common and powerful online learning framework. Following the analysis in [Zhao et al., 2021b], we can directly obtain dynamic regret of OGD and SOGD [Orabona and Pál, 2018] in a unified view owing to the versatility of OMD.

Online mirror descent algorithm updates according to the following rule:

$$\mathbf{x}_{t+1} = \arg\min_{\mathbf{x}\in\mathcal{X}} \ \eta_t \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}\rangle + \mathcal{D}_\psi(\mathbf{x}, \mathbf{x}_t), \tag{49}$$

where $\eta_t > 0$ is the time-varying step size, $h_t(\cdot) : \mathbf{x} \mapsto \mathbb{R}$ is the convex loss function, and $\mathcal{D}_\psi(\cdot,\cdot)$ is the Bregman divergence induced by the regularizer function $\psi(\cdot)$ defined as $\mathcal{D}_\psi(\mathbf{x},\mathbf{y}) = \psi(\mathbf{x}) - \psi(\mathbf{y}) - \langle \nabla\psi(\mathbf{y}), \mathbf{x}-\mathbf{y}\rangle$. OMD following dynamic regret guarantee [Zhao et al., 2021b].

Theorem 5 (Theorem 1 of Zhao et al. [2021b])**.** *Suppose that the regularizer $\psi : \mathcal{X} \mapsto \mathbb{R}$ is 1-strongly convex with respect to the norm $\|\cdot\|$. The dynamic regret of Optimistic Mirror Descent (OMD) whose update rule specified in (49) is bounded as follows:*

$$\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{u}_t) \leq \sum_{t=1}^{T} \eta_t \|\nabla f_t(\mathbf{x}_t)\|_*^2 + \sum_{t=1}^{T} \frac{1}{\eta_t}\Big(\mathcal{D}_\psi(\mathbf{u}_t, \mathbf{x}_t) - \mathcal{D}_\psi(\mathbf{u}_t, \mathbf{x}_{t+1})\Big)$$
$$- \sum_{t=1}^{T} \frac{1}{\eta_t} \mathcal{D}_\psi(\mathbf{x}_{t+1}, \mathbf{x}_t),$$

which holds for any comparator sequence $\mathbf{u}_1, \ldots, \mathbf{u}_T \in \mathcal{X}$.

Choosing $\psi(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|_2^2$ will lead to the update form of online gradient descent used as base learners in our algorithm:

$$\mathbf{x}_{t+1} = \arg\min_{\mathbf{x}\in\mathcal{X}} \ \eta_t \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}\rangle + \frac{1}{2}\|\mathbf{x} - \mathbf{x}_t\|_2^2, \tag{50}$$

where the Bregman divergence becomes $\mathcal{D}_\psi(\mathbf{x}, \mathbf{x}_t) = \frac{1}{2}\|\mathbf{x} - \mathbf{x}_t\|_2^2$ w.r.t. the choice of regularizer.

We proceed to show the dynamic regret of online gradient descent (OGD),

Lemma 7. *Under Assumption 2, by choosing static step size $\eta_t = \eta > 0$, Online Gradient Descent defined in equation (50) satisfies:*

$$\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{u}_t) \leq \frac{7D^2}{4\eta} + \frac{D}{\eta} \sum_{t=2}^{T} \|\mathbf{u}_{t-1} - \mathbf{u}_2\|_2 + \eta \sum_{t=1}^{T} \|\nabla f_t(\mathbf{x}_t)\|_2^2$$

for any comparator sequence $\mathbf{u}_1, \ldots, \mathbf{u}_T \in \mathcal{X}$.

Proof. Applying Theorem 5 with the choices of $\psi(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|_2^2$ and fixed step size $\eta_t = \eta > 0$ gives:

$$\begin{aligned}
&\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{u}_t) \\
&\leq \frac{1}{2\eta}\sum_{t=1}^{T}\left(\|\mathbf{u}_t - \mathbf{x}_t\|_2^2 - \|\mathbf{u}_t - \mathbf{x}_{t+1}\|_2^2\right) + \eta\sum_{t=1}^{T}\|\nabla f_t(\mathbf{x}_t)\|_2^2 - \frac{1}{2\eta}\sum_{t=1}^{T}\|\mathbf{x}_t - \mathbf{x}_{t+1}\|_2^2 \\
&\leq \frac{1}{2\eta}\sum_{t=1}^{T}\left(\|\mathbf{x}_t\|_2^2 - \|\mathbf{x}_{t+1}\|_2^2\right) + \frac{1}{\eta}\sum_{t=1}^{T}\left(\mathbf{x}_{t+1} - \mathbf{x}_t\right)^\top \mathbf{u}_t + \eta\sum_{t=1}^{T}\|\nabla f_t(\mathbf{x}_t)\|_2^2 \\
&\leq \frac{1}{2\eta}\|\mathbf{x}_1\|_2^2 + \frac{1}{\eta}\left(\mathbf{x}_{T+1}^\top \mathbf{u}_T - \mathbf{x}_1^\top \mathbf{u}_1\right) + \frac{1}{\eta}\sum_{t=2}^{T}(\mathbf{u}_{t-1} - \mathbf{u}_t)^\top \mathbf{x}_t + \eta\sum_{t=1}^{T}\|\nabla f_t(\mathbf{x}_t)\|_2^2 \\
&\leq \frac{7D^2}{4\eta} + \frac{D}{\eta}\sum_{t=2}^{T}\|\mathbf{u}_{t-1} - \mathbf{u}_2\|_2 + \eta\sum_{t=1}^{T}\|\nabla f_t(\mathbf{x}_t)\|_2^2,
\end{aligned}$$

where the last inequality is due to:

$$\begin{aligned}
\|\mathbf{x}_1\|_2^2 = \|\mathbf{x}_1 - \mathbf{0}\|_2^2 &\leq D^2, \\
\mathbf{x}_{T+1}^\top \mathbf{u}_T \leq \|\mathbf{x}_{T+1}\|_2 \cdot \|\mathbf{u}_T\|_2 &= D^2, \\
-\mathbf{x}_1^\top \mathbf{u}_1 \leq \frac{1}{4}\|\mathbf{x}_1 - \mathbf{u}_1\|_2^2 &\leq \frac{1}{4}D^2, \\
(\mathbf{u}_{t-1} - \mathbf{u}_t)^\top \mathbf{x}_t \leq \|\mathbf{u}_{t-1} - \mathbf{u}_t\|_2 \cdot \|\mathbf{x}_t\|_2 &\leq D\|\mathbf{u}_{t-1} - \mathbf{u}_t\|_2.
\end{aligned}$$

□

D.2 Self-Confident Tuning

Orabona and Pál [2018] have analyzed the regret bound of SOGD. For completeness, we here provide the regret analysis under the OMD framework. Indeed, SOGD can be treated as OMD with a self-confident learning rate. Thus, we have the following lemma.

Lemma 8. *Under assumption 1 and 2, the OMD algorithm defined in equation (49) with the choices of regularizer $\psi(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|_2^2$ and the time-varying learning rate defined as*

$$\eta_t = \frac{D/2}{\sqrt{\delta + \sum_{\tau=1}^{t-1}\|\nabla f_t(\mathbf{x}_t)\|_2^2}}$$

for some $\delta > 0$, has the following guarantee:

$$\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{u}) \leq 2D \cdot \sqrt{\delta + \sum_{t=1}^{T}\|\nabla f_t(\mathbf{x}_t)\|_2^2},$$

where $\mathbf{u} \in \mathcal{X}$ can be any comparator.

Proof. We start the proof with the application of Theorem 5. For any fixed comparator $\mathbf{u} \in \mathcal{X}$,

$$\begin{aligned}
&\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{u}_t) \\
&\leq \sum_{t=1}^{T}\frac{1}{2\eta_t}\left(\|\mathbf{u} - \mathbf{x}_t\|_2^2 - \|\mathbf{u} - \mathbf{x}_{t+1}\|_2^2\right) + \sum_{t=1}^{T}\eta_t\|\nabla f_t(\mathbf{x}_t)\|_2^2 - \sum_{t=1}^{T}\frac{1}{2\eta_t}\|\mathbf{x}_t - \mathbf{x}_{t+1}\|_2^2 \\
&\leq \frac{1}{2\eta_1}\|\mathbf{u} - \mathbf{x}_1\|_2^2 + \sum_{t=2}^{T}\left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}}\right)\frac{\|\mathbf{u} - \mathbf{x}_t\|_2^2}{2} + \sum_{t=1}^{T}\eta_t\|\nabla f_t(\mathbf{x}_t)\|_2^2
\end{aligned}$$

$$\leq \frac{D^2}{2\eta_1} + \frac{D^2}{2} \sum_{t=2}^{T} \left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} \right) + \sum_{t=1}^{T} \eta_t \|\nabla f_t(\mathbf{x}_t)\|_2^2$$

$$= \frac{D^2}{2\eta_T} + \sum_{t=1}^{T} \eta_t \|\nabla f_t(\mathbf{x}_t)\|_2^2. \tag{51}$$

Applying Lemma 13 to the second term of (51) and by the definition of η_T, the regret bound becomes

$$\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{u}_t) \leq D \cdot \sqrt{\delta + \sum_{t=1}^{T} \|\nabla f_t(\mathbf{x}_t)\|_2^2} + D \left(\sqrt{\delta + \sum_{t=1}^{T} \|\nabla f_t(\mathbf{x}_t)\|_2^2} - \sqrt{\delta} \right)$$

$$\leq 2D \cdot \sqrt{\delta + \sum_{t=1}^{T} \|\nabla f_t(\mathbf{x}_t)\|_2^2},$$

which completes the proof. □

To bound the meta-regret of our dynamic methods, we introduce the FTRL lemma [Orabona, 2019, Corollary 7.8] under the time-varying learning rate.

Lemma 9 (FTRL Lemma)**.** *Suppose that the regularizer function $\psi : \mathcal{X} \mapsto \mathbb{R}$ is α-strongly convex with respect to the norm $\|\cdot\|$. Let f_t be a sequence of convex loss functions and $\psi_t(\mathbf{x}) = \frac{1}{\eta_t}(\psi(\mathbf{x}) - \min_{\mathbf{x}' \in \mathcal{X}} \psi(\mathbf{x}'))$, where $\eta_{t+1} \leq \eta_t$, $t = 1, \ldots, T$. Then the decision sequence $\mathbf{x}_t$ generated by*

$$\mathbf{x}_t = \operatorname*{arg\,min}_{\mathbf{x} \in \mathcal{X}} \left(\psi_t(\mathbf{x}) + \sum_{\tau=1}^{t-1} f_t(\mathbf{x}) \right),$$

satisfies the following regret upper bound for any $\mathbf{u} \in \mathcal{X}$,

$$\sum_{t=1}^{T} f_t(\mathbf{x}_t) - f_t(\mathbf{u}) \leq \frac{\psi(\mathbf{u}) - \min_{\mathbf{x} \in \mathcal{X}} \psi(\mathbf{x})}{\eta_{T+1}} + \frac{1}{2\alpha} \sum_{t=1}^{T} \eta_t \|\nabla f_t(\mathbf{x}_t)\|_*^2.$$

Based on the preceding lemma, we can derive the regret upper bound for Hedge algorithm with a self-confident learning rate.

Lemma 10. *Consider the prediction with expert advice setting with N experts and the linear loss $f_t(\mathbf{x}) = \langle \boldsymbol{\ell}_t, \mathbf{x} \rangle$, where $\ell_t \in \mathbb{R}^d$. Then the self-confident tuning Hedge, whose initial decision is $\boldsymbol{p}_1 = 1/N \cdot \mathbf{1}$ and update rules are*

$$p_{t+1,i} \propto \exp\left(\varepsilon_{t+1} \sum_{\tau=1}^{t} \ell_{\tau,i} \right) \text{ with } \varepsilon_{t+1} = \sqrt{\frac{\ln N}{1 + \sum_{\tau=1}^{t} \|\ell_\tau\|_\infty^2}}$$

ensures the following regret guarantee: for any $i \in [N]$

$$\sum_{t=1}^{T} \langle \boldsymbol{p}_t, \boldsymbol{\ell}_t \rangle - \sum_{t=1}^{T} \ell_{t,i} \leq 3\sqrt{\ln N \cdot \left(1 + \sum_{t=1}^{T} \|\boldsymbol{\ell}_t\|_\infty^2 \right)} + \frac{\sqrt{\ln N}}{2} \cdot \max_{t \in [T]} \|\boldsymbol{\ell}_t\|_\infty^2.$$

Proof. It is easy to verify that, the mentioned self-confident tuning Hedge can be treated as a special case of the time-varying FTRL algorithm by choosing $\psi(\boldsymbol{p}) = \sum_{s=1}^{N} p_s \ln p_s$, which is 1-strongly convex with respect to $\|\cdot\|_1$, and $\psi_t(\boldsymbol{p}) = \frac{1}{\varepsilon_t}\psi(\boldsymbol{p})$. Thus, by Lemma 9, we have

$$\sum_{t=1}^{T} \langle \boldsymbol{p}_t, \boldsymbol{\ell}_t \rangle - \sum_{t=1}^{T} \ell_{t,i} \leq \frac{\ln N}{\varepsilon_{T+1}} + \frac{1}{2} \sum_{t=1}^{T} \varepsilon_t \|\boldsymbol{\ell}_t\|_\infty^2$$

$$\leq \frac{\ln N}{\varepsilon_{T+1}} + \frac{\sqrt{\ln N}}{2} \cdot \left(4\sqrt{1 + \sum_{t=1}^{T} \|\boldsymbol{\ell}_t\|_\infty^2} + \max_{t \in [T]} \|\boldsymbol{\ell}_t\|_\infty^2 \right)$$

$$= 3\sqrt{\ln N \cdot \left(1 + \sum_{t=1}^{T} \|\boldsymbol{\ell}_t\|_\infty^2\right)} + \frac{\sqrt{\ln N}}{2} \cdot \max_{t\in[T]} \|\boldsymbol{\ell}_t\|_\infty^2,$$

where the first inequality chooses $\mathbf{u}$ as the one-hot vector with all entries being 0 except the i-th one as 1, and second inequality is by Lemma 15. ☐

D.3 Facts on Geometric Covers

Lemma 11 (Lemma 11 of Zhang et al. [2019]). *Let $[s_p, s_q] \subseteq [T]$ be an arbitrary interval that starts from a marker s_p and ends at another marker s_q. Then, we can find a sequence of consecutive intervals*

$$I_1 = [s_{i_1}, s_{i_2} - 1],\ I_2 = [s_{i_2}, s_{i_3} - 1],\ \ldots, I_v = [s_{i_v}, s_{i_{v+1}} - 1] \in \widetilde{\mathcal{C}}$$

such that

$$i_1 = p,\ i_v \le q < i_{v+1}, \text{and } v \le \lceil \log_2(q - p + 2) \rceil.$$

D.4 Technical Lemmas

In this section, we will present several technical lemmas used in our proof.

Lemma 12 (Lemma 3.1 of Srebro et al. [2010]). *For an L-smooth and nonnegative function $f : \mathcal{X} \mapsto \mathbb{R}_+$,*

$$\|\nabla f(\mathbf{x})\|_2 \le \sqrt{4Lf(\mathbf{x})},\ \forall \mathbf{x} \in \mathcal{X}.$$

Lemma 13 (Lemma 3.5 of Auer et al. [2002]). *Let $l_1, \ldots, l_T$ be non-negative real numbers. Then:*

$$\sum_{t=1}^{T} \frac{l_t}{\sqrt{\delta + \sum_{i=1}^{t} l_i}} \le 2\left(\sqrt{\delta + \sum_{t=1}^{T} l_t} - \sqrt{\delta}\right).$$

Lemma 14 (Lemma 14 of Gaillard et al. [2014]). *Let $a_0 > 0$ and $a_1, \ldots, a_m \in [0, 1]$ be real numbers and let $f : (0, +\infty) \mapsto [0, +\infty)$ be a non-increasing function. Then*

$$\sum_{i=1}^{m} a_i f(a_0 + \cdots + a_{i-1}) \le f(a_0) + \int_{a_0}^{a_0 + a_1 + \cdots + a_m} f(u) du.$$

Lemma 15 (Lemma 4.8 of Pogodin and Lattimore [2019]). *Let $a_1, a_2, \ldots, a_T$*

$$\sum_{t=1}^{T} \frac{a_t}{\sqrt{1 + \sum_{s=1}^{t-1} a_s}} \le 4\sqrt{1 + \sum_{t=1}^{T} a_t} + \max_{t\in[T]} a_t.$$

Lemma 16 (Lemma 5 of Shalev-Shwartz [2007]). *For any $x, y, a \in \mathbb{R}_+$ that satisfies $x - y \le \sqrt{ax}$,*

$$x - y \le \sqrt{ay} + a.$$

Based on Lemma 16, we can achieve the following result.

Lemma 17. *For any $x, y, a, b \in \mathbb{R}_+$ that satisfies $x - y \le \sqrt{ax} + b$,*

$$x - y \le \sqrt{ay + ab} + a + b.$$

Lemma 18 (Lemma 13 of Gaillard et al. [2014]). *For all $x > 0$ and all $\alpha \ge 1$, we have*

$$x \le x^\alpha + \frac{\alpha - 1}{e}.$$