Parameter-free Algorithms for the Stochastically Extended Adversarial Model

Anonymous Author(s)
Affiliation
Address
email

Abstract

We develop the first parameter-free algorithms for the Stochastically Extended Adversarial (SEA) model, a framework that bridges adversarial and stochastic online convex optimization. Existing approaches for the SEA model require prior knowledge of problem-specific parameters, such as the diameter of the domain D and the Lipschitz constant of the loss functions G, which limits their practical applicability. Addressing this, we develop parameter-free methods by leveraging the Optimistic Online Newton Step (OONS) algorithm to eliminate the need for these parameters. We first establish a comparator-adaptive algorithm for the scenario with unknown domain diameter but known Lipschitz constant, achieving an expected regret bound of $\widetilde{O}\big(\|u\|_2^2 + \|u\|_2(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})\big)$, where u is the comparator vector and $\sigma_{1:T}^2$ and $\Sigma_{1:T}^2$ represent the cumulative stochastic variance and cumulative adversarial variation, respectively. We then extend this to the more general setting where both D and G are unknown, attaining the comparator- and Lipschitz-adaptive algorithm. Notably, the regret bound exhibits the same dependence on $\sigma_{1:T}^2$ and $\Sigma_{1:T}^2$, demonstrating the efficacy of our proposed methods even when both parameters are unknown in the SEA model.

1 Introduction

We focus on online convex optimization (OCO) [1, 2, 3], a broad framework for sequential decision-making. In each round $t \in [T]$, a learner chooses a point x_t from a convex set $\mathcal{X} \subseteq \mathbb{R}^d$. The environment then discloses a convex function $f_t : \mathcal{X} \to \mathbb{R}$, after which the learner incurs a loss $f_t(x_t)$ and updates their decision. The standard way to show the performance is via the *regret*, the total loss relative to a comparator $u \in \mathcal{X}$, defined as $\mathfrak{R}_T(u) = \sum_{t=1}^T f_t(x_t) - \sum_{t=1}^T f_t(u)$.

For convex problems, the regret can be bounded by $O(\sqrt{T})$ [4], which is known to be minimax optimal [5]. OCO encompasses two primary frameworks: adversarial OCO [4, 6], which aims to minimize regret against arbitrarily chosen loss functions, and stochastic OCO (SCO) [6, 7], which minimizes excess risk under i.i.d. losses. While both frameworks are well-studied, real-world scenarios typically fall between these theoretical extremes of purely adversarial or stochastic settings. The Stochastically Extended Adversarial (SEA) model proposed in [8] bridges the gap between traditional adversarial and stochastic frameworks in OCO. This hybrid approach serves as an intermediate formulation that captures aspects of both adversarial OCO and SCO settings.

Optimal performance in OCO, SCO, and SEA models typically relies on careful step-size tuning, which requires prior knowledge of problem parameters such as the diameter of the decision set and Lipschitz constants. However, these parameters are often unknown in practice, motivating the development of *parameter-free* algorithms that achieve comparable regret without requiring such oracle information. Specifically, parameter-free algorithms include *comparator-adaptive* algorithms (unknown diameter D) and *Lipschitz-adaptive* algorithms (unknown Lipschitz constant G). A related

Table 1: Comparison of the regret bounds of existing results and our proposed algorithms.

Algorithm	Free of D	Free of G	Bound on Expected Regret $\mathbb{E}[\mathfrak{R}_T(u)]$
OFTLR, OMD (Sachs et al. [8], Chen et al. [12])	✗	✗	$O\left(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2}\right)$
OONS (Theorem 3.2)	✗	✗	$\widetilde{O}\left(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2}\right)$
CA-OONS (Theorem 4.1)	✓	✗	$\widetilde{O}\left(\|u\|_2^2 + \|u\|_2(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})\right)$
CLA-OONS (Theorem 4.4)	✓	✓	$\widetilde{O}\left(\|u\|_2^2(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2}) + \|u\|_2^4 + \sqrt{\sigma_{1:T} + \mathfrak{G}_{1:T}}\right)$

challenge arises when the decision set $\mathcal{X}$ is unbounded, allowing adversaries to induce arbitrarily large losses for linear functions. Traditional methods often circumvent this by assuming bounded domains, where $\sup_{x,y\in\mathcal{X}} \|x-y\|_2 \le D$. Consequently, developing OCO algorithms that remain effective under both unknown parameters and unbounded domains is significantly more challenging than in classical settings [9, 10, 11].

To address these challenges, we propose "parameter-free" algorithms for the SEA model, accommodating potentially unbounded decision sets. Using the Optimistic Online Newton Step as our base algorithm, we systematically relax assumptions: first tackling the case of an unknown domain diameter D (potentially infinite) with a known Lipschitz constant G, and then extending to the more complex scenario where both D and G are unknown. In the SEA model, at each time step t, the learner selects a distribution $\mathcal{D}_t$ over functions and incurs a loss $f_t(x_t)$, where f_t is sampled from $\mathcal{D}_t$. The *expected gradient* is denoted as $\nabla F_t(x) = \mathbb{E}_{f_t\sim\mathcal{D}_t}[\nabla f_t(x)]$.

Main Contributions. Our main results and contributions are summarized as follows.

(1) We begin by introducing the Optimistic Online Newton Step (OONS) as our foundational algorithm. The OONS algorithm is inspired by [10]; however, we incorporate an adaptive step-size η_t rather than a fixed step-size throughout the learning process. When the parameters D and G are known, we demonstrate that OONS achieves an expected regret bound of $\widetilde{O}(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})$, matching the state-of-the-art results in terms of dependence on the cumulative stochastic variance $\sigma_{1:T}^2$ and the cumulative adversarial variation $\Sigma_{1:T}^2$ [8, 12]. This establishes a solid foundation for our subsequent extensions to parameter-free algorithms.

(2) We introduce the first parameter-free (comparator-adaptive) algorithm for the SEA model that remains effective when the domain diameter D is unknown, provided the Lipschitz constant G is known. This is achieved through a meta-framework wherein each base learner operates within a distinct bounded domain, complemented by the Multi-scale Multiplicative-Weight with Correction (MsMwC) algorithm [10] for the meta-algorithm's weight updates. This construction yields an expected regret bound of $\widetilde{O}(\|u\|_2^2 + \|u\|_2(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2}))$, where the bound scales with the ℓ_2-norm of the comparator u without requiring prior knowledge of the domain diameter D.

(3) We further consider a setting in which *both* the domain diameter D and the Lipschitz constant G are unknown. By devising appropriate update rules for the estimation of the domain diameter, we establish an expected regret bound of $\widetilde{O}(\|u\|_2^2(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2}) + \|u\|_2^4 + \sqrt{\sigma_{1:T} + \mathfrak{G}_{1:T}})$ where $\sigma_{1:T}$ captures the deviation of the stochastic gradients (excluding squared norms), and $\mathfrak{G}_{1:T}$ denotes the sum of the maximum expected gradients over the sequence.

A summary of our results and the best existing results are included in Table 1. Due to space limitations, we hide the Lipschitz constant G in the $\widetilde{O}(\cdot)$-notation in the regret bound of CLA-OONS algorithm.

1.1 Related Work

The SEA model [8] is motivated by foundational insights from the *gradual-variation online learning*. The study of gradual variation can be traced back to the works of [13] and [14], and it has gained significant traction in recent years [15, 16, 17, 18, 19]. Notably, the SEA model has emerged as a practical application of the gradual variation principle [16, 18, 19]. Furthermore, this model serves as a bridge between adversarial OCO and SCO. This intermediate framework is comprehensively understood in the context of expert prediction [20, 21] and the bandit setting [22, 23].

Parameter-free online learning has emerged as a fundamental advancement in machine learning, offering solutions to the critical challenge of parameter tuning in practice. In the baseline scenario, when both the diameter parameter D and the gradient bound G are known, algorithms leveraging Follow the Regularized Leader or Mirror Descent principles achieve the minimax optimal regret bound of $\mathfrak{R}_T(u) \le O(GD\sqrt{T})$ [2]. The field has subsequently progressed to address more practical scenarios where complete parameter knowledge is unavailable. Notably, in the Lipschitz-adaptive setting, it is still possible to attain the same optimal regret bound, differing only by constant factors [24, 25]. Xie et al. [18] extended these principles to gradient-variation online learning.

In the comparator-adaptive setting, the online learning problem becomes substantially more challenging due to the unknown comparator's magnitude, which could cause the algorithm's predictions to significantly deviate from the optimal solution, leading to a large regret. For this challenging scenario, a key result has been established as $\mathfrak{R}_T(u) \le \widetilde{O}(\|u\|_2 G\sqrt{T})$ [26, 24, 9, 27]. For scenarios where both parameters D and G are unknown, significant progress has been made recently. Cutkosky [25] developed an algorithm with $\mathfrak{R}_T(u) \le \widetilde{O}(G\|u\|_2^3 + \|u\|_2 G\sqrt{T})$, while Mhammedi & Koolen [28] achieve $\mathfrak{R}_T(u) \le \widetilde{O}(G\|u\|_2^3 + G\sqrt{\max_{t\le T}(\sum_{s=1}^{t}\|g_s\|_2/\max_{s\le t}\|g_s\|_2)})$. An alternative approach by [29] presented the regret bound $\mathfrak{R}_T(u) \le \widetilde{O}(\|u\|_2^2 G\sqrt{T})$. More recent advances including [11] achieve the regret $\mathfrak{R}_T(u) \le \widetilde{O}(G\|u\|_2\sqrt{T} + L\|u\|_2^2\sqrt{T})$ under the condition that subgradients satisfy $\|g_t\|_2 \le G + L\|x_t\|_2$. Cutkosky & Mhammedi [30] further improve it to $\widetilde{O}(G\|u\|_2\sqrt{T} + \|u\|_2^2 + G^2)$.

Besides parameter-free algorithms for OCO, [31] and [32] studied the parameter-free stochastic gradient descent (SGD) algorithms. Khaled et al. (2024) [33] introduced the concept of "tuning-free" algorithms, which achieve performance comparable to optimally-tuned SGD within polylogarithmic factors, requiring only approximate estimates of the relevant problem parameters.

Although this series of works on parameter-free algorithms in OCO provides valuable insights, these approaches cannot be directly applied to attain the optimal regret bounds for the SEA model without prior knowledge of parameters. This limitation stems from the fact that the desired bounds for the SEA model should be expressed in terms of the variance-like quantities $\sigma_{1:T}^2$ and $\Sigma_{1:T}^2$, rather than the time horizon T. While Sachs et al. [8] have attempted to address this issue by proposing an algorithm that adapts to an unknown strong convexity parameter, their step-size search range still depends on both D and G, thereby restricting its fully parameter-free adaptivity.

2 Problem Setup and Preliminaries

In this section, we formulate the problem setup of the Stochastically Extended Adversarial (SEA) model, present the existing results, and discuss the key challenges.

2.1 Problem Setup of the SEA Model

In iteration $t \in [T]$, the learner selects a decision x_t from a convex feasible domain $\mathcal{X} \subseteq \mathbb{R}^d$, and nature chooses a distribution $\mathcal{D}_t$ from a set of distributions over functions. Then, the learner suffers a loss $f_t(x_t)$, where f_t is a random function sampled from the distribution $\mathcal{D}_t$. The distributions are allowed to vary over time, and by choosing them appropriately, the SEA model reduces to the adversarial OCO, SCO, or other intermediate settings. Additionally, for each $t \in [T]$, the *(conditional) expected function* is defined as $F_t(x) = \mathbb{E}_{f_t\sim\mathcal{D}_t}[f_t(x)]$ and the *expected gradient* is defined as $\nabla F_t(x) = \mathbb{E}_{f_t\sim\mathcal{D}_t}[\nabla f_t(x)]$. We define $\mathfrak{G}_t := \sup_{x\in\mathcal{X}} \|\nabla F_t(x)\|_2$ to be the largest norm of the expected gradient, and use the shorthand $\mathfrak{G}_{1:T}$ to denote the sum $\sum_{t=1}^{T}\mathfrak{G}_t$.

Due to the randomness in the online decision-making process, our goal in the SEA model is to bound the *expected regret* with respect to the randomness in the loss functions f_t drawn from the distribution $\mathcal{D}_t$ against any fixed comparator $u \in \mathcal{X}$, defined as $\mathbb{E}[\mathfrak{R}_T(u)] \triangleq \mathbb{E}[\sum_{t=1}^{T} f_t(x_t) - \sum_{t=1}^{T} f_t(u)]$. To capture the characteristics of the SEA model, we introduce the following quantities. For each $t \in [T]$, define the *(conditional) variance of the gradients* and *cumulative stochastic variance* respectively as

$$\sigma_t^2 = \sup_{x\in\mathcal{X}} \mathbb{E}_{f_t\sim\mathcal{D}_t}\left[\|\nabla f_t(x) - \nabla F_t(x)\|_2^2\right], \quad \sigma_{1:T}^2 = \mathbb{E}\left[\sum_{t=1}^{T}\sigma_t^2\right], \tag{1}$$

which reflect the stochasticity of the online process. Additionally, we introduce the concepts of stochastic gradient deviation and cumulative gradient deviation to characterize the stochastic variation of gradients, without the squared norm. The stochastic gradient deviation is defined as $\sigma_t = \sup_{x\in\mathcal{X}} \mathbb{E}_{f_t\sim\mathcal{D}_t}\left[\|\nabla f_t(x) - \nabla F_t(x)\|_2\right]$, and the *cumulative gradient deviation* is defined as $\sigma_{1:T} = \mathbb{E}\big[\sum_{t=1}^{T}\sigma_t\big]$. The *cumulative adversarial variation* is defined as

$$\Sigma_{1:T}^2 = \mathbb{E}\left[\sum_{t=1}^{T}\sup_{x\in\mathcal{X}}\|\nabla F_t(x) - \nabla F_{t-1}(x)\|_2^2\right],$$

where $\nabla F_0(x) = 0$, reflecting the adversarial difficulty. This work aims to provide expected regret bounds that depend on problem-dependent quantities such as $\sigma_{1:T}^2$, $\Sigma_{1:T}^2$, and $\mathfrak{G}_{1:T}$ instead of T.

Below, we present several assumptions. Note that our results do not rely on *all* of these assumptions; rather, specific assumptions are required for each result, which will be explicitly stated in the theorem.

Assumption 2.1 (Boundedness of gradient norms)**.** The gradient norms of all loss functions are bounded by G, i.e., $\max_{t\in[T]}\max_{x\in\mathcal{X}}\|\nabla f_t(x)\|_2 \le G$.

Assumption 2.2 (Boundedness of domain)**.** The diameter of the convex set $\mathcal{X}$ (the feasible domain) is bounded by D i.e., $\sup_{x,y\in\mathcal{X}}\|x-y\|_2 \le D$.

Assumption 2.3 (Smoothness)**.** For all $t\in[T]$, the expected function F_t is L-smooth over $\mathcal{X}$, i.e., $\|\nabla F_t(x) - \nabla F_t(y)\|_2 \le L\|x-y\|_2$ for all $x, y\in\mathcal{X}$.

Assumption 2.4 (Convexity)**.** For all $t\in[T]$, the expected function F_t is convex on $\mathcal{X}$.

Notations. Given a positive definite matrix A, the norm induced by A is $\|x\|_A = \sqrt{x^\top Ax}$. Δ_N denotes the $(N-1)$-dimensional simplex. Let $\psi : \mathcal{X}\to\mathbb{R}$ be a continuously differentiable and strictly convex function, the associated Bregman divergence is defined as $D_\psi(x,y) \coloneqq \psi(x) - \psi(y) - \langle\nabla\psi(y), x-y\rangle$. The notation $O(\cdot)$ hides constants and $\widetilde{O}(\cdot)$ additionally hides polylog factors.

2.2 Existing Results for the SEA Model

Bounded domain and gradient norm. Sachs et al. [8] established a regret bound for the SEA model using both Optimistic Follow-The-Regularized-Leader (OFTRL) and Optimistic Mirror Descent (OMD), given by $\mathbb{E}[\mathfrak{R}_T(u)] = O(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})$, achieved by setting the step-size as $\eta_t = \frac{D^2}{\sum_{s=1}^{t-1}\min\{\frac{\eta_s}{2}\|g_s - m_s\|_2^2, D\|g_s - m_s\|_2\}}$, where $g_t = \nabla f_t(x_t)$ and $m_t = g_{t-1}$. Similarly, Chen et al. [12] derived the same bound by Optimistic Online Mirror Descent (OMD) with the step-size $\eta_t = \frac{D}{\sqrt{\delta + 4G^2 + \bar{V}_{t-1}}}$, where $\bar{V}_{t-1} = \sum_{s=1}^{t-1}\|g_s - m_s\|_2^2$ and $\delta > 0$.

In all of the above settings, the optimal step-size η_t is dependent on the parameters D (the diameter of decision set $\mathcal{X}$) and G (Lipschitz constant), so there has been a natural motivation to develop algorithms that achieve similar regret bounds *without* knowing such parameters *a priori*. We term such algorithms as *"parameter-free"* algorithms for the SEA model.

Parameter-free algorithm for the SEA model. Theorem 5 in [27] demonstrates that the parameter-free mirror descent algorithm can be extended to enjoy a gradient-variation regret of $\mathfrak{R}_T(u) \le \widetilde{O}(\|u\|_2\sqrt{\sum_{t=1}^{T}\|\nabla f_t(x_t) - \nabla f_{t-1}(x_t)\|_2^2})$. In fact, this can directly yield an expected regret bound for the SEA model scaling with $\widetilde{\sigma}_{1:T}^2 \coloneqq \mathbb{E}\left[\sum_{t=1}^{T}\mathbb{E}_{f_t\sim\mathcal{D}_t}\left[\sup_x\|\nabla f_t(x) - \nabla F_t(x)\|_2^2\right]\right]$, i.e.,

$$\mathbb{E}[\mathfrak{R}_T(u)] \le \widetilde{O}\left(\|u\|_2\Big(\sqrt{\widetilde{\sigma}_{1:T}^2} + \sqrt{\Sigma_{1:T}^2}\Big)\right). \tag{2}$$

Akin to $\sigma_{1:T}^2$, $\widetilde{\sigma}_{1:T}^2$ defined in [12] also captures the stochastic nature of the SEA model. Furthermore, in the worst case, the bound in (2) reduces to $\widetilde{O}(\|u\|_2\sqrt{T})$, matching the best available problem-independent bound. The outer expectation in the definition of $\widetilde{\sigma}_{1:T}^2$ accounts for the randomness in the choice of the distribution $\mathcal{D}_t$ at each step. Refer to Appendix A.1 for a self-contained proof of (2).

Key Challenge. However, we emphasize that our goal is to obtain regret bounds scaling with $\sigma_{1:T}^2$, as defined in (1). As pointed out in previous work on the SEA model [12, Remark 9], $\sigma_{1:T}^2$

Algorithm 1 Optimistic Online Newton Step (OONS)

Input: learning rate $\eta_t > 0$, $x'_1 = 0$.

for $t = 1, \ldots, T$ **do**
  Receive optimistic prediction m_t and range hint z_t.
  Update $x_t = \arg\min_{x \in \mathcal{X}} \{\langle x, m_t \rangle + D_{\psi_t}(x, x'_t)\}$ where $\psi_t(x) = \frac{1}{2}\|x\|^2_{A_t}$ and $A_t = 4z_1^2 I + \sum_{s=1}^{t-1} \eta_s(\nabla_s - m_s)(\nabla_s - m_s)^\top + 4\eta_t z_t^2 I$.
  Receive $g_t = \nabla f_t(x_t)$ and construct $\nabla_t = g_t + 32\eta_t \langle x_t, g_t - m_t \rangle (g_t - m_t)$.
  Update $x'_{t+1} = \arg\min_{x \in \mathcal{X}} \{\langle x, \nabla_t \rangle + D_{\psi_t}(x, x'_t)\}$.
end for

is more favorable than $\widetilde{\sigma}^2_{1:T}$. First, from a mathematical perspective, the latter is generally larger due to the convexity of the supremum operator. Second, from an algorithmic perspective, an algorithm with a regret bound involving $\widetilde{\sigma}^2_{1:T}$ typically involves an *implicit* update, typically requires an *implicit* update, which operates on the original function and is significantly more costly than standard first-order methods (see Remark 10 in [12]). Third, achieving regret bounds scaling with $\sigma^2_{1:T}$ typically requires leveraging the *Regret Bounded by Variation in Utilities (RVU)* property [34], which captures the regret to be bounded not only by the gradient variations but also an additional negative stability term. Formally, an algorithm satisfies the RVU property if its regret upper bound is in the form of $\sum_{t=1}^T \langle x^* - x_t, u_t \rangle \le \alpha + \beta \sum_{t=1}^T \|u_t - u_{t-1}\|^2 - \gamma \sum_{t=1}^T \|x_t - x_{t-1}\|^2$, for some constants $\alpha, \beta, \gamma > 0$. This structure enables finer control over the regret by explicitly analyzing trajectory stability, establishing profound connections to game theory [34] and accelerations in smooth optimization [35]. Consequently, the key challenge lies in how to achieve this preferred $\sigma^2_{1:T}$-scaling *without* knowledge of D and G for unbounded domains.

3 Optimistic Online Newton Step (ONS) for the SEA Model

In this section, different from the Optimistic follow-the-regularized-leader (OFTRL) [8] and Optimistic mirror descent (OMD) [12], we first introduce the Optimistic Online Newton Step (OONS) algorithm as the base algorithm for the "parameter-free" algorithms to be introduced later. This algorithm is summarized in Algorithm 1.

The ONS algorithm [6] is an iterative algorithm that adaptively updates a second-order (Hessian-based) model of the loss, allowing more efficient gradient-based updates and improved regret bounds. OONS also maintains two sequences $\{x_t\}_{t=1}^T$ and $\{x'_t\}_{t=1}^T$ like OMD and OFTRL, which is achieved by introducing the optimistic prediction m_t. Chen et al. [10] also considered combining their Multi-scale Multiplicative-weight with Correction (MsMwC) algorithm with this variant of the ONS algorithm. However, the step-size η is fixed in their algorithm and the MsMwC algorithm is applied to learn the optimal $\eta_\star$. Different from it, in OONS, we consider adaptive step-sizes η_t.

Theorem 3.1. *Suppose that $\|g_t - m_t\|_2 \le z_t$, z_t is non-decreasing in t, $64\eta_t D z_T \le 1, \forall t \in [T]$, and η_t is non-increasing in t. Then,* OONS *guarantees that*

$$\mathfrak{R}_T(u) \le O\left(\frac{r \ln(T\eta_1 z_T/z_1)}{\eta_T} + z_1^2\|u\|_2^2 + D(z_T - z_1) + \sum_{t=1}^T \eta_t \langle u, g_t - m_t \rangle^2 - z_1^2 \sum_{t=2}^T \|x_t - x_{t-1}\|_2^2\right) \tag{3}$$

where r is the rank of $\sum_{t=1}^T (g_t - m_t)(g_t - m_t)^\top$.

Next, we verify that OONS also works for the case with known parameters D, G, and we can also obtain a similar regret bound as [8] and [12]. The regret bound of OONS for the SEA model with known parameters D and G is presented below. We specify the adaptive step-size for all $t \in [T]$ as

$$\eta_t = \min\left\{\frac{1}{64 D z_T}, \frac{1}{D\sqrt{\sum_{s=1}^{t-1}\|g_s - m_s\|_2^2}}\right\}. \tag{4}$$

Since G is known, we have $\|g_t - m_t\|_2 \le z_t = 2G, \forall t \in [T]$ and η_t is defined in terms of z_T here.

Theorem 3.2. *Under Assumptions 2.1, 2.2, 2.3, and 2.4,* OONS *with step-size η_t given in* (4)*, $m_t = \nabla f_{t-1}(x_{t-1})$ and $z_t = 2G$ for all $t \in [T]$ ensures $\mathbb{E}[\mathfrak{R}_T(u)] = \widetilde{O}(\sqrt{\sigma^2_{1:T}} + \sqrt{\Sigma^2_{1:T}})$.*

Remark 3.3. Theorem 3.2 achieves the same (up to logarithmic terms) dependence on $\sigma^2_{1:T}$ and $\Sigma^2_{1:T}$ as in [8] and [12]. The primary reason to use OONS as the base algorithm instead of OMD [12] is

Algorithm 2 Comparator-adaptive algorithm for the SEA model (CA-OONS)

Input: Lipschitz constant G.

1: **for** $t = 1, \ldots, T$ **do**
2: Create $N = \lceil \log T \rceil$ base-learners. Each base-learner $j \in [N]$ runs OONS with stepsize η_t^j.
3: Each base-learner j provides x_t^j.
4: Run Algorithm 3 to obtain $w_t \in \Delta_N$.
5: The final decision is $x_t = \sum_{j=1}^N w_{t,j} x_t^j$.
6: **end for**

Algorithm 3 Meta Algorithm

Input: Additional expert set $\mathcal{S}$ defined in (7).
Initialization: $p_1' \in \Delta_{\mathcal{S}}$ such that $p_{1,k}' \propto \beta_k^2$ for all $k \in \mathcal{S}$.

1: **for** $t = 1, \ldots, T$ **do**
2: Construct $h_t \in \mathbb{R}^N$ with $h_t^j = \langle \nabla f_{t-1}(x_{t-1}^j), x_t^j \rangle$.
3: Each expert $k \in \mathcal{S}$ runs MsMwC with step-size $\beta_{t,j}^k$ and plays $w_t^k \in \Delta_N$.
4: Receive w_t^k for each $k \in \mathcal{S}$ and compute $H_t^k = \langle w_t^k, h_t \rangle$.
5: Compute $p_t = \arg\min_{p \in \Delta_{\mathcal{S}}} \langle p, H_t \rangle + D_\phi(p, p_t')$.
6: Play $w_t = \sum_{k \in \mathcal{S}} p_{t,k} w_t^k \in \Delta_N$.
7: Receive $\ell_t \in \mathbb{R}^N$. Define $L_t^k = \langle w_t^k, \ell_t \rangle$ and $b_t^k = 32\beta_k (L_t^k - H_t^k)$.
8: Compute $p_{t+1}' = \arg\min_{p \in \Delta_{\mathcal{S}}} \langle p, L_t + b_t \rangle + D_\phi(p, p_t')$.
9: **end for**

that the final regret bound for OMD typically depends on $D\sqrt{\sum_{t=1}^T \|g_t - m_t\|_2^2}$. In scenarios when D is unknown or potentially infinite, like in Section 4.2, this might lead to $O(T)$ regret bounds. By contrast, OONS leverages adaptive second-order information, which helps remove (or substantially reduce) explicit dependence on D. In (3), the only term relevant to D is $D(z_T - z_1)$, which solely depends on the starting and ending points, z_1 and z_T.

4 Parameter-free Algorithms for the SEA Model

In this section, we develop parameter-free algorithms for the SEA model, building on OONS which we use as the base algorithm. Moreover, we allow the decision set $\mathcal{X}$ to be potentially unbounded throughout this section, i.e. $D = \infty$ in Assumption 2.2. In Section 4.1, we present a *comparator-adaptive* algorithm for the SEA model for unknown D but known G, and then, we develop the *comparator- and Lipschitz-adaptive* algorithm where both D and G are unknown in Section 4.2.

4.1 Comparator-adaptive algorithm

We now propose the Comparator-Adaptive Optimistic Online Newton Step (CA-OONS) algorithm for the unknown D (potentially infinite) but known G case by using a meta-base algorithm framework. Recent works in [10] and [11] addressed the challenges associated with unbounded domains by developing a base-learner framework. Building on this philosophy, we propose CA-OONS(Algorithm 2) where we adopt the MsMwC–Master algorithm [10] as the meta algorithm.

The algorithm uses N base-learners. For *any* base-learner $i \in [N]$, the regret can be decomposed as

$$\mathfrak{R}_T(u) = \sum_{t=1}^T f_t(x_t) - \sum_{t=1}^T f_t(u) = \underbrace{\sum_{t=1}^T f_t(x_t) - \sum_{t=1}^T f_t(x_t^i)}_{\text{Meta Regret}} + \underbrace{\sum_{t=1}^T f_t(x_t^i) - \sum_{t=1}^T f_t(u)}_{\text{Base Regret}} .$$

Let $\mathcal{A}_i$ denote the base algorithm for the i-th base-learner. We denote $\mathfrak{R}_T^{\mathcal{A}_i}(u) = \sum_{t=1}^T f_t(x_t^i) - \sum_{t=1}^T f_t(u)$ as the base regret by taking $\mathcal{A}_i$ as the base algorithm. Moreover, the final decision x_t is a weighted-average of all the base-learners' decisions: $x_t = \sum_{j=1}^N w_{t,j} x_t^j$ with $w_t \in \Delta_N$. As such,

$$\mathfrak{R}_T(u) \le \mathfrak{R}_T^{\mathcal{A}_i}(u) + \sum_{t=1}^T \langle \ell_t, w_t - w_\star^i \rangle, \tag{5}$$

Table 2: Three-layer hierarchy of CA-OONS

Layer	Algorithm	Loss	Optimism	Decision	Output
Top Meta	MsMwC	$(L_t^k)_{k\in\mathcal{S}}$	$(H_t^k)_{k\in\mathcal{S}}$	$p_t \in \Delta_{\mathcal{S}}$	$w_t = \sum_k p_{t,k} w_t^k$
Middle Meta	MsMwC	$\ell_t \in \mathbb{R}^N$	$h_t \in \mathbb{R}^N$	$(w_t^k)_{k\in\mathcal{S}} \in \Delta_N$	w_t^k
Base	OONS	$\nabla f_t(x_t^j)$	$\nabla f_{t-1}(x_{t-1}^j)$	$(x_t^j)_{j\in N} \in \mathcal{X}_j$	x_t^j

where $\ell_t \in \mathbb{R}^N$ with $\ell_t^j = \langle \nabla f_t(x_t^j), x_t^j\rangle$ and $w_\star^i$ is a vector in Δ_N whose j-th component is $(w_\star^i)_j = 1$ if $j = i$ and 0 otherwise. Refer to Appendix C.2 for the proof of (5).

We first consider the base algorithm. Specifically, for each base-learner $j \in [N]$, we impose a constraint that it operates within $\mathcal{X}_j = \{x : \|x\|_2 \le D_j \wedge x \in \mathcal{X}\}$, where $D_j = 2^j$. Then, we define $g_t^j = \nabla f_t(x_t^j)$ and $m_t^j = \nabla f_{t-1}(x_{t-1}^j)$. Each base-learner $j \in [N]$ runs OONS with step-size

$$\eta_t^j = \min\left\{\frac{1}{64 D_j z_T}, \frac{1}{D_j\sqrt{\sum_{s=1}^{t-1}\|g_s^j - m_s^j\|_2^2}}\right\}, \tag{6}$$

which depends on D_j instead of D in OONS. Hence, each base-learner j can update x_t^j via OONS with step-size η_t^j. Since the final decision is $x_t = \sum_{j=1}^N w_{t,j} x_t^j$, we need to adopt a meta-algorithm to learn the weight parameter $w_t \in \Delta_N$.

As mentioned above, we introduce a constraint that each base-learner $j \in [N]$ operates within a D_j-bounded domain. We can consider this as a "multi-scale" base-learner problem [36, 9, 10] where each base-learner j has a different loss range such that $|\ell_t^j| \le GD_j$ since $\ell_t^j = \langle \nabla f_t(x_t^j), x_t^j\rangle$. We choose the Multi-scale Multiplicative-weight with Correction (MsMwC)–Master algorithm (Algorithm 2 in [10]) as the meta-algorithm to learn w_t, which is implemented based on the MsMwC [10]. Details of MsMwC are presented in Appendix C.1.Specifically, we define a new expert set

$$\mathcal{S} = \{k \in \mathbb{Z} : \exists j \in [N], GD_j \le 2^{k-2} \le GD_j\sqrt{T}\}. \tag{7}$$

For all $k \in \mathcal{S}$, the step-size of the MsMwC–Master algorithm is set to $\beta_k = \frac{1}{32\cdot 2^k}$. Each expert $k \in \mathcal{S}$ runs the MsMwC algorithm with w_1' being uniform over $\mathcal{Z}(k)$, where $\mathcal{Z}_k = \{j \in [N] : GD_j \le 2^{k-2}\}$. Moreover, each base MsMwC algorithm only works in the subset $\mathcal{Z}(k)$, i.e., $w_t \in \Delta_N$ with $w_{t,j} = 0$ for all $j \notin \mathcal{Z}(k)$. We can view CA-OONS (Algorithm 2) as a three-layer structure, where the meta-algorithm (Algorithm 3) itself consists of two layers, which we refer to as *meta top* and *meta middle*. Also, the base layer is OONS algorithm. For clarity, we summarize the notations of the three-layer hierarchy for CA-OONS in Table 2.

In the following, we provide the expected regret guarantee for CA-OONS.

Theorem 4.1. *Let D be unknown (potentially infinite). Under Assumptions 2.1, 2.3 and 2.4,* CA-OONS *provides the following regret*

$$\mathbb{E}[\mathfrak{R}_T(u)] = \widetilde{O}\Big(\|u\|_2^2 + \|u\|_2\big(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2}\big)\Big). \tag{8}$$

This regret guarantee is referred to as "comparator-adaptive" because it depends directly on the norm of the comparator, $\|u\|_2$, rather than explicitly relying on the diameter of the decision set, D. Notably, when considering the constrained decision set with a diameter D, our regret bound immediately recovers the result $\mathbb{E}[\mathfrak{R}_T(u)] = \widetilde{O}(D^2 + D(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2}))$ established in [8, 12].

One limitation of our regret bound (8) is that when particularizing to adversarial OCO, it achieves only an $\widetilde{O}(\|u\|_2^2 + \|u\|_2\sqrt{T})$ worst-case regret bound, which falls short of the best-known $\widetilde{O}(\|u\|_2\sqrt{T})$ regret bound [24, 27, 28]. However, in the following two remarks, we will justify the $\|u\|_2$-dependence in the gradient-variation regret and emphasize the fundamental challenge of achieving adaptivity from the gradient-variation bound (for smooth functions) to the worst-case bound (for the non-smooth case) when the decision set of online learning is *unconstrained*.

Remark 4.2 (Dependency on $\|u\|_2^2$). Recent studies have established the connection between gradient-variation online learning and accelerated offline optimization through advanced online-to-batch conversions [35]. Specifically, let $d_0 = \|x_0 - x_*\|_2$ denote the distance of an initial point x_0 to the optimum x_*. For an L-smooth function, gradient-variation online algorithms using first-order information correspond to an accelerated convergence rate of $O(Ld_0^2/T^2)$ via *weighted* online-to-batch conversions [37]. For a G-Lipschitz function, the problem-independent regret bounds translate to an $O(Gd_0/\sqrt{T})$ rate through the standard conversion [1]. In this context, we hypothesize that the $\|u\|_2^2$ term may be unavoidable in gradient-variation regret for unconstrained online learning, paralleling how the d_0^2 term also appears in the accelerated rate of unconstrained offline optimization.

Remark 4.3 (Adaptivity between gradient-variation bound and worst-case bound). We argue that achieving *adaptivity* between the gradient-variation bound and the problem-independent worst-case bound in unconstrained online learning may be as challenging as achieving *universality* in offline optimization over unconstrained domains, where the method must adapt to both smooth and Lipschitz functions. To the best of our knowledge, the best-known universal method for offline unconstrained optimization is by [38], which combines UNIXGRAD [37] with the DOG step size [31]. Nonetheless, this method is complex and still relies on a predefined range of parameters, highlighting both the difficulty of the problem and the fact that it remains only partially solved. Consequently, designing a *single* online learning algorithm that adaptively bridges the gradient-variation regret bound for smooth functions and the worst-case bound for Lipschitz functions is non-trivial, which could provide new insights into universal offline optimization methods. We leave this for future work.

4.2 Comparator- and Lipschitz-adaptive Algorithm

The algorithm in the previous subsection requires *prior* knowledge of the Lipschitz constant G. Due to practical limitations, as such knowledge may not be available in real applications. A comparator- and Lipschitz-adaptive algorithm would instead *adapt* to an unknown Lipschitz constant G.

A simple approach to handling the unknown gradient norms, proposed by [25], relies on a gradient-clipping reduction. The key idea is to design an algorithm $\mathcal{A}$ that achieves appropriate regret when given prescient "hints" $h_t \geq \|g_t\|_2$ at the start of round t. Since such hints are impractical (as g_t is not observed beforehand), we instead approximate them using a clipped gradient, inspired by [25]. We start with an initial guess B_0 on the range of $\max_t \|g_t - m_t\|_2$, where $g_t = \nabla f_t(x_t)$. We define $B_t = \max_{0\leq s\leq t} \|g_s - m_s\|_2$ as the predicted error range up to iteration t. The truncated gradient is then defined as $\widetilde{g}_t = m_t + \frac{B_{t-1}}{B_t}(g_t - m_t)$. The truncated gradient satisfies $\|\widetilde{g}_t - m_t\|_2 \leq B_{t-1}$, allowing the learner to assume that the range of predicted error in iteration t is known at the start.

Next, we initialize the decision set diameter guess as $D_1 = 1$. For each iteration $t \in [T]$, we first play x_t and receive $g_t = \nabla f_t(x_t)$. To update D_t, we consider the condition $D_t < \sqrt{\sum_{s=1}^t \frac{\|g_s\|_2}{\max\{1, \max_{k\leq s}\|g_k\|_2\}}}$. If this condition holds, we update D_{t+1} using the doubling trick. This ensures that we need to update D_t a maximum of $M = O(\log T)$ times. We divide the total T iterations into disjoint subsets of M iterations. If the "doubling" occurs at the t-th iteration, we update $t_a \leftarrow t$ and reset $x'_{t+1} = 0$ and the matrix A_{t+1} in OONS as follows

$$A_{t+1} = 4z_{t_a+1}^2 I + \sum_{s=t_a+1}^{t} \eta_s(\nabla_s - m_s)(\nabla_s - m_s)^\top + 4\eta_{t+1}z_{t+1}^2 I. \tag{9}$$

Then, we feed $\widetilde{g}_t$ to OONS running in the D_{t+1}-bounded domain $\mathcal{X}_{t+1} = \{x : \|x\|_2 \leq D_{t+1} \wedge x \in \mathcal{X}\}$ and obtain x_{t+1}. We summarize the ideas in Algorithm 4 and term it as Comparator and Lipschitz-Adaptive Optimistic Online Newton Step (or CLA-OONS) algorithm.

Theorem 4.4. *Let both D (potentially infinite) and G be unknown. Under Assumption 2.1 (but G is unknown), 2.3 and 2.4, the proposed* FPF-OONS *algorithm satisfies*

$$\mathbb{E}[\mathfrak{R}_T(u)] \leq \widetilde{O}\Big(\|u\|_2^2(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2}) + G^2\|u\|_2^2 + \|u\|_2^4 + G\|u\|_2^3 + G^2\sqrt{\sigma_{1:T} + \mathfrak{G}_{1:T}}\Big),$$

where $\sigma_{1:T}$ captures the stochastic gradient deviation (without the squared norm) and $\mathfrak{G}_{1:T}$ denotes the sum of maximum expected gradients.

Remark 4.5 (Discussion and challenges). In Theorem 4.4, the regret includes $\|u\|_2^2(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})$. Ideally, we aim to achieve a dependence of $\widetilde{O}(\|u\|_2)$, consistent with [25] and [39]. However, achieving this within the SEA framework presents significant challenges. As mentioned in Section 2, obtaining regret bounds that scale with $\sigma_{1:T}^2$ in the SEA framework is difficult. These challenges are

Algorithm 4 Comparator and Lipschitz-Adaptive (or CLA-OONS) for the SEA model

Input: Initial scale B_0.
Initialize: $D_1 = 1$.

1: **for** $t = 1, \ldots, T$ **do**
2: Run OONS in D_t-bounded domain and obtain x_t. Play x_t and receive $g_t = \nabla f_t(x_t)$.
3: Construct $\widetilde{g}_t = m_t + \frac{B_{t-1}}{B_t}(g_t - m_t)$, where $B_t = \max_{0 \le s \le t} \|g_s - m_s\|_2$.
4: **if** $D_t < \sqrt{\sum_{s=1}^{t} \frac{\|g_s\|_2}{\max\{1, \max_{k \le s} \|g_k\|_2\}}}$ **then**
5: Update $D_{t+1} = 2\sqrt{\sum_{s=1}^{t} \frac{\|g_s\|_2}{\max\{1, \max_{k \le s} \|g_k\|_2\}}}$ and reset A_{t+1} as (9) and $x'_{t+1} = 0$.
6: **end if**
7: Feed $\widetilde{g}_t$ to OONS running in the D_{t+1}-bounded domain and get x_{t+1}, where $z_{t+1} = B_t$.
8: **end for**

compounded in the comparator and Lipschitz-adaptive setting. Below, we outline some of the main technical challenges associated with achieving the desired bound of $\widetilde{O}(\|u\|_2(\sqrt{\sigma^2_{1:T}} + \sqrt{\Sigma^2_{1:T}}))$.

As stated in Section 2, the methods such as those proposed by [25, 27, 11] cannot be applied to obtain the expected regret bound in terms of $\sigma^2_{1:T}$. The work [40] also looks promising; however, it remains unclear to us whether the approach proposed in the paper can be directly extended to the SEA framework. Their results, presented in Theorems 3 and 5 of [40], are not Lipschitz-adaptive. Specifically, they operate under the assumptions that $\|g_t\|_2 \le 1$ and $\|m_t\|_2 \le 1$.

One could also use a large number of base-learners to achieve a regret of $\widetilde{O}(\|u\|_2\sqrt{r\sum_t \|g_t - m_t\|_2^2} + \|u\|_2^3)$, similar to [10]. However, this approach presents a subtle yet significant challenge. Following a similar analysis [10], we get the following decomposition: $\sum_t \langle g_t, x_t - u\rangle = \sum_t \langle g_t, x_t - x_t^{k_*}\rangle + \sum_t \langle g_t, x_t^{k_*} - u\rangle$. By leveraging Theorem 23 in [10], we can write $\sum_t \langle g_t, x_t^{k_*} - u\rangle$ as $\sum_t \langle g_t, x_t^{k_*} - u\rangle \le \widetilde{O}(\|u\|\sqrt{r\sum_t \|g_t - m_t\|_2^2} - \sum_t \|x_t^{k_*} - x_{t-1}^{k_*}\|_2^2)$. Observe that expressing the first term, $\sqrt{r\sum_t \|g_t - m_t\|_2^2}$, in terms of $\sigma_{1:T}$ and $\Sigma_{1:T}$ introduces additional terms involving $\sum_t \|x_t - x_{t-1}\|_2^2$ (See Lemma 2 in [12]). The only way to address this term is through the negative term $-\sum_t \|x_t^{k_*} - x_{t-1}^{k_*}\|_2^2$, which becomes tricky. This challenge is reminiscent of the problem encountered by [17]. Their solution, as outlined in Equation (17) of [17], relies on the bounded domain assumption, which is not applicable in our setting. Consequently, this limitation prevents us from improving the $\|u\|_2^2$ dependence in the leading term by using additional base-learners.

The bound in Theorem 4.4 also includes an additive term involving $\sqrt{\mathfrak{G}_{1:T}}$, which reflects the sum of the maximum expected gradients' norms over T rounds, and arises because the domain is potentially unbounded. Note that this term does not have a $\|u\|$ dependence. Hence, the comparator having a large norm in an unbounded setting (potentially dependent on T) does not affect its growth. In the worst case, $\sqrt{\mathfrak{G}_{1:T}} = O(\sqrt{T})$, which underscores that this additive term does not have a significant adverse effect on the regret as $\sqrt{\sigma^2_{1:T}}$ and $\sqrt{\Sigma^2_{1:T}}$ also scale as $\sqrt{T}$ [8].

5 Conclusions and Future Work

This paper presents novel parameter-free algorithms for the SEA model, addressing critical challenges in online optimization where traditional approaches require prior knowledge of parameters such as the diameter of the domain D and the Lipschitz constant of the loss functions G. Our proposed algorithms: CA-OONS and CLA-OONS are designed to operate effectively even when D and G are unknown, demonstrating their adaptability and practicality.

There are several avenues for future research. First, we would like to improve the regret's dependence on $\|u\|_2$ when both D and G are unknown. Another promising direction is to reduce the number of gradient queries in CA-OONS from $O(\log T)$ to $O(1)$, thus enhancing its efficiency. An intriguing question in the comparator-adaptive setting is whether it is possible to design a single, simple online algorithm that simultaneously achieves two types of bounds: $\widetilde{O}(\|u\|_2^2 + \|u\|_2(\sqrt{\sigma^2_{1:T}} + \sqrt{\Sigma^2_{1:T}}))$ and $\widetilde{O}(\|u\|_2 G\sqrt{T})$. As discussed in Remark 4.3, it remains an open challenge to construct an adaptive parameter-free online algorithm that can interpolate between these bounds.

References

[1] Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.

[2] Elad Hazan et al. Introduction to online convex optimization. *Foundations and Trends® in Optimization*, 2(3-4):157–325, 2016.

[3] Francesco Orabona. A modern introduction to online learning. *arXiv preprint arXiv:1912.13213*, 2019.

[4] Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th international conference on machine learning (icml-03)*, pages 928–936, 2003.

[5] Jacob Abernethy, Peter L Bartlett, Alexander Rakhlin, and Ambuj Tewari. Optimal strategies and minimax lower bounds for online convex games. In *Proceedings of the 21st annual conference on learning theory*, pages 414–424, 2008.

[6] Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2):169–192, 2007.

[7] Hao Yu, Michael Neely, and Xiaohan Wei. Online convex optimization with stochastic constraints. *Advances in Neural Information Processing Systems*, 30, 2017.

[8] Sarah Sachs, Hedi Hadiji, Tim van Erven, and Cristobal Guzman. Accelerated rates between stochastic and adversarial online convex optimization. *arXiv preprint arXiv:2303.03272*, 2023.

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

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

[11] Andrew Jacobsen and Ashok Cutkosky. Unconstrained online learning with unbounded losses. In *International Conference on Machine Learning*, pages 14590–14630. PMLR, 2023.

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

[13] Elad Hazan and Satyen Kale. Extracting certainty from uncertainty: Regret bounded by variation in costs. *Machine learning*, 80:165–188, 2010.

[14] Chao-Kai Chiang, Tianbao Yang, Chia-Jung Lee, Mehrdad Mahdavi, Chi-Jen Lu, Rong Jin, and Shenghuo Zhu. Online optimization with gradual variations. In *Conference on Learning Theory*, pages 6–1. JMLR Workshop and Conference Proceedings, 2012.

[15] Peng Zhao, Yu-Jie Zhang, Lijun Zhang, and Zhi-Hua Zhou. Dynamic regret of convex and smooth functions. *Advances in Neural Information Processing Systems*, 33:12510–12520, 2020.

[16] Yu-Hu Yan, Peng Zhao, and Zhi-Hua Zhou. Universal online learning with gradient variations: A multi-layer online ensemble approach. *Advances in Neural Information Processing Systems*, 36:37682–37715, 2023.

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

[18] Yan-Feng Xie, Peng Zhao, and Zhi-Hua Zhou. Gradient-variation online learning under generalized smoothness. *arXiv preprint arXiv:2408.09074*, 2024.

[19] Yu-Hu Yan, Peng Zhao, and Zhi-Hua Zhou. A simple and optimal approach for universal online learning with gradient variations. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024.

[20] Idan Amir, Idan Attias, Tomer Koren, Yishay Mansour, and Roi Livni. Prediction with corrupted expert advice. *Advances in Neural Information Processing Systems*, 33:14315–14325, 2020.

[21] Shinji Ito. On optimal robustness to adversarial corruption in online decision problems. *Advances in Neural Information Processing Systems*, 34:7409–7420, 2021.

[22] Julian Zimmert and Yevgeny Seldin. An optimal algorithm for stochastic and adversarial bandits. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 467–475. PMLR, 2019.

[23] Chung-Wei Lee, Haipeng Luo, Chen-Yu Wei, Mengxiao Zhang, and Xiaojin Zhang. Achieving near instance-optimality and minimax-optimality in stochastic and adversarial linear bandits simultaneously. In *International Conference on Machine Learning*, pages 6142–6151. PMLR, 2021.

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

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

[26] Dylan J Foster, Alexander Rakhlin, and Karthik Sridharan. Adaptive online learning. *Advances in Neural Information Processing Systems*, 28, 2015.

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

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

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

[30] Ashok Cutkosky and Zakaria Mhammedi. Fully unconstrained online learning. *arXiv preprint arXiv:2405.20540*, 2024.

[31] Maor Ivgi, Oliver Hinder, and Yair Carmon. Dog is sgd's best friend: A parameter-free dynamic step size schedule. In *International Conference on Machine Learning*, pages 14465–14499. PMLR, 2023.

[32] Ahmed Khaled, Konstantin Mishchenko, and Chi Jin. Dowg unleashed: An efficient universal parameter-free gradient descent method. *Advances in Neural Information Processing Systems*, 36:6748–6769, 2023.

[33] Ahmed Khaled and Chi Jin. Tuning-free stochastic optimization. *arXiv preprint arXiv:2402.07793*, 2024.

[34] Vasilis Syrgkanis, Alekh Agarwal, Haipeng Luo, and Robert E Schapire. Fast convergence of regularized learning in games. *Advances in Neural Information Processing Systems*, 28, 2015.

[35] Peng Zhao. Lecture Notes for Advanced Optimization, 2025. Lecture 9. Optimism for Acceleration.

[36] Sebastien Bubeck, Nikhil R Devanur, Zhiyi Huang, and Rad Niazadeh. Online auctions and multi-scale online learning. In *Proceedings of the 2017 ACM Conference on Economics and Computation*, pages 497–514, 2017.

[37] Ali Kavis, Kfir Y Levy, Francis Bach, and Volkan Cevher. UniXGrad: A universal, adaptive algorithm with optimal guarantees for constrained optimization. *Advances in neural information processing systems*, 32, 2019.

[38] Itai Kreisler, Maor Ivgi, Oliver Hinder, and Yair Carmon. Accelerated parameter-free stochastic optimization. In *The Thirty Seventh Annual Conference on Learning Theory*, pages 3257–3324. PMLR, 2024.

[39] H Brendan McMahan and Matthew Streeter. Adaptive bound optimization for online convex optimization. *arXiv preprint arXiv:1002.4908*, 2010.

[40] Ashok Cutkosky. Combining online learning guarantees. In *Conference on Learning Theory*, pages 895–913. PMLR, 2019.

A Omitted Details of Section 2

A.1 Proof of Equation (2)

Proof. From Theorem 5 in [27], we have

$$\mathfrak{R}_T(u) \le \widetilde{O}\left(\|u\|_2\sqrt{\sum_{t=1}^{T}\|\nabla f_t(x_t) - \nabla f_{t-1}(x_t)\|_2^2}\right)$$
$$\le \widetilde{O}\left(\|u\|_2\sqrt{\sum_{t=1}^{T}\sup_{x\in\mathcal{X}}\|\nabla f_t(x) - \nabla f_{t-1}(x)\|_2^2}\right)$$

From Lemma 8 in [12], we also have

$$\sum_{t=1}^{T}\sup_{x\in\mathcal{X}}\|\nabla f_t(x) - \nabla f_{t-1}(x)\|_2^2$$
$$\le G^2 + 6\sum_{t=1}^{T}\sup_{x\in\mathcal{X}}\|\nabla f_t(x) - \nabla F_t(x)\|_2^2 + 4\sum_{t=1}^{T}\sup_{x\in\mathcal{X}}\|\nabla F_t(x) - \nabla F_{t-1}(x)\|_2^2.$$

Taking expectations with Jensen's inequality and the definition of $\widetilde{\sigma}_{1:T}^2$ and $\Sigma_{1:T}^2$, we obtain

$$\mathbb{E}[\mathfrak{R}_T(u)] \le \widetilde{O}\left(\|u\|_2(\sqrt{\widetilde{\sigma}_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})\right). \qquad \square$$

B Omitted Details of Section 3

B.1 Auxiliary Lemmas

Lemma B.1 (Bregman Proximal Inequality)**.** *The Bregman Proximal update in the form of* $x_{t+1} = \arg\min_{x\in\mathcal{X}}\{\langle x, g_t\rangle + D_\psi(x, x_t)\}$ *satisfies*

$$\langle g_t, x_{t+1} - u\rangle \le D_\psi(u, x_t) - D_\psi(u, x_{t+1}) - D_\psi(x_t, x_{t+1}). \tag{10}$$

Proof. By the first-order optimality condition at x_{t+1}, for any $u \in \mathcal{X}$, we have

$$\langle g_t + \nabla\psi(x_{t+1}) - \nabla\psi(x_t), u - x_{t+1}\rangle \ge 0,$$

On the RHS of (10), we expand each term by the definition of Bregman divergence

$$D_\psi(u, x_t) - D_\psi(u, x_{t+1}) - D_\psi(x_t, x_{t+1}) = \langle\nabla\psi(x_{t+1}) - \nabla\psi(x_t), u - x_{t+1}\rangle.$$

Hence, the proof is finished by rearranging the terms. $\square$

Lemma B.2. *Let* $x_t = \arg\min_{x\in\mathcal{X}}\{\langle x, m_t\rangle + D_{\psi_t}(x, x'_t)\}$ *and* $x'_{t+1} = \arg\min_{x\in\mathcal{X}}\{\langle x, g_t\rangle + D_{\psi_t}(x, x'_t)\}$*. Then, it holds for any* u *in* $\mathcal{X}$

$$\sum_{t=1}^{T}\langle x_t - u, g_t\rangle \le \sum_{t=1}^{T}\langle x_t - x'_{t+1}, g_t - m_t\rangle + D_{\psi_t}(u, x'_t) - D_{\psi_t}(u, x'_{t+1})$$
$$- D_{\psi_t}(x_t, x'_{t+1}) - D_{\psi_t}(x_t, x'_t).$$

Proof. We have

$$\langle x_t - u, g_t\rangle = \langle x_t - x'_{t+1}, m_t\rangle + \langle x'_{t+1} - u, g_t\rangle + \langle x_t - x'_{t+1}, g_t - m_t\rangle. \tag{11}$$

We apply Lemma B.1 twice, i.e., $\langle a - u, f\rangle \le D_\psi(u, b) - D_\psi(u, a) - D_\psi(a, b)$ since $a = \arg\min_{x\in\mathcal{X}}\langle x, f\rangle + D_\psi(x, b)$. Then, we have

$$\left\langle x_t - x'_{t+1}, m_t\right\rangle \le D_{\psi_t}(x'_{t+1}, x'_t) - D_{\psi_t}(x'_{t+1}, x_t) - D_{\psi_t}(x_t, x'_t),$$
$$\left\langle x'_{t+1} - u, g_t\right\rangle \le D_{\psi_t}(u, x'_t) - D_{\psi_t}(x'_{t+1}, u) - D_{\psi_t}(x'_t, x'_{t+1}).$$

Substitute these two back to (11) and sum over T providing the desired result. $\square$

Lemma B.3. *In* OONS*, we have* $0 \le \langle x_t - x'_{t+1}, \nabla_t - m_t\rangle \le 2\|\nabla_t - m_t\|^2_{A_t^{-1}}$ *and* $\sum_{t=1}^T \langle x_t - x'_{t+1}, \nabla_t - m_t\rangle \le O\Big(\frac{r\ln(T\eta_1 z_T/z_1)}{\eta_T}\Big)$.

Proof. We define

$$F_{m_t}(x) = \langle x, m_t\rangle + D_{\psi_t}(x, x'_t), \qquad F_{\nabla_t}(x) = \langle x, \nabla_t\rangle + D_{\psi_t}(x, x'_t)$$

By OONS, we have

$$x_t = \arg\min_{x\in\mathcal{X}} F_{m_t}(x), \qquad x'_{t+1} = \arg\min_{x\in\mathcal{X}} F_{\nabla_t}(x).$$

Since $\nabla^2 D_{\psi_t} = A_t$ and by the first-order optimality at x'_{t+1}, we have

$$F_{\nabla_t}(x_t) - F_{\nabla_t}(x'_{t+1}) \ge \frac{1}{2}\|x_t - x'_{t+1}\|^2_{A_t}.$$

Also, we can write $F_{\nabla_t}(x_t) - F_{\nabla_t}(x'_{t+1})$ as

$$F_{\nabla_t}(x_t) - F_{\nabla_t}(x'_{t+1}) = \langle x_t - x'_{t+1}, \nabla_t\rangle + D_{\psi_t}(x_t, x'_t) - D_{\psi_t}(x'_{t+1}, x'_t).$$

Then,

$$\begin{aligned}\frac{1}{2}\|x_t - x'_{t+1}\|^2_{A_t} &\le \langle x_t - x'_{t+1}, \nabla_t - m_t\rangle + F_{m_t}(x_t) - F_{m_t}(x'_{t+1}),\\ &\le \langle x_t - x'_{t+1}, \nabla_t - m_t\rangle,\end{aligned}$$

where the second inequality comes from $x_t = \arg\min F_{m_t}(x)$. Therefore, we have

$$\langle x_t - x'_{t+1}, \nabla_t - m_t\rangle \le 2\|\nabla_t - m_t\|^2_{A_t^{-1}}.$$

Since x'_{t+1} minimizes $F_{m_t}(x)$ and x_t minimizes $F_{\nabla_t}(x)$, we have

$$\begin{aligned}0 \le F_{\nabla_t}(x_t) - F_{\nabla_t}(x'_{t+1}) &= \langle x_t - x'_{t+1}, \nabla_t\rangle + D_{\psi_t}(x_t, x'_t) - D_{\psi_t}(x'_{t+1}, x'_t),\\ &= \langle x_t - x'_{t+1}, \nabla_t - m_t\rangle + F_{m_t}(x_t) - F_{m_t}(x'_{t+1})\\ &\le \langle x_t - x'_{t+1}, \nabla_t - m_t\rangle.\end{aligned}$$

By the definition of $\nabla_t = g_t + 32\eta_t\langle x_t, g_t - m_t\rangle(g_t - m_t)$, we have

$$\begin{aligned}\|\nabla_t - m_t\|_2 &= \|g_t - m_t + 32\eta_t\langle x_t, g_t - m_t\rangle(g_t - m_t)\|_2\\ &\le \|g_t - m_t\|_2 + 32\eta_t D\|g_t - m_t\|_2^2 \le \frac{3}{2}\|g_t - m_t\|_2. \end{aligned} \tag{12}$$

Next, we define

$$\bar{A}_t = 4z_1^2 \cdot I + \sum_{s=1}^t \eta_s(\nabla_s - m_s)(\nabla_s - m_s)^\top.$$

Hence, $A_t \succeq \bar{A}_t$ since $\|\nabla_t - m_t\|_2^2 \le 4\|g_t - m_t\|_2^2 \le 4z_t^2$. Also, we have

$$(\nabla_t - m_t)(\nabla_t - m_t)^\top = \frac{1}{\eta_t}[\eta_t(\nabla_t - m_t)(\nabla_t - m_t)^\top] = \frac{1}{\eta_t}(\bar{A}_t - \bar{A}_{t-1})$$

Then,

$$\begin{aligned}\sum_{t=1}^T \|\nabla_t - m_t\|^2_{A_t^{-1}} &\le \sum_{t=1}^T \|\nabla_t - m_t\|^2_{\bar{A}_t^{-1}},\\ &= \sum_{t=1}^T \operatorname{trace}\Big((\nabla_t - m_t)(\nabla_t - m_t)^\top \bar{A}_t^{-1}\Big),\\ &\le \sum_{t=1}^T \frac{1}{\eta_t}\operatorname{trace}\Big(\bar{A}_t^{-1}(\bar{A}_t - \bar{A}_{t-1})\Big),\\ &\le \sum_{t=1}^T \frac{1}{\eta_t}(\ln|\bar{A}_t| - \ln|\bar{A}_{t-1}|),\\ &\le \frac{1}{\eta_T}\ln\frac{|\bar{A}_T|}{|\bar{A}_0|}.\end{aligned}$$

For $|\bar{A}_T|$:

$$|\bar{A}_T| \leq |4z_1^2 I + \sum_{t=1}^{T} \eta_t (\nabla_t - m_t)(\nabla_t - m_t)^\top|$$

$$\ln |\bar{A}_T| \leq O\Big(r \ln(1 + \sum_{t=1}^{T} \frac{\eta_t}{4z_1^2} \|\nabla_t - m_t\|_2^2)\Big)$$

$$\leq O\Big(r \ln(1 + \frac{4z_T^2}{4z_1^2} \sum_{t=1}^{T} \eta_t)\Big)$$

$$\leq O\Big(r \ln(1 + \frac{\eta_1 z_T^2}{z_1^2} T)\Big),$$

where r is the rank of $\sum_{t=1}^{T} (\nabla_t - m_t)(\nabla_t - m_t)^\top$.

Therefore, we have

$$\sum_{t=1}^{T} \langle x_t - x'_{t+1}, \nabla_t - m_t \rangle \leq O\Big(\frac{r \ln(T\eta_1 z_T / z_1)}{\eta_T}\Big).$$

□

Lemma B.4. *Let $s_t, \forall t \in [T]$ be non-negative. Then,*

$$\sum_{t=1}^{T} \frac{s_t}{\sqrt{\sum_{j=1}^{t} s_j}} \leq 2\sqrt{\sum_{t=1}^{T} s_t}.$$

Proof. Let $S_t = \sum_{j=1}^{t} s_j$. Then,

$$\sum_{t=1}^{T} \frac{s_t}{\sqrt{S_t}} \leq \sum_{t=1}^{T} \int_{S_{t-1}}^{S_t} \frac{1}{\sqrt{x}} dx = \int_0^{S_T} \frac{1}{\sqrt{x}} dx = 2\sqrt{S_T}.$$

□

Lemma B.5 (Theorem 5 in [8], Lemma 3 in [12])**.** *Under Assumptions 2.1 and 2.3, we have*

$$\sum_{t=1}^{T} \|\nabla f_t(x_t) - \nabla f_{t-1}(x_{t-1})\|_2^2 \leq G^2 + 4\sum_{t=2}^{T} \|\nabla F_t(x_{t-1}) - \nabla F_{t-1}(x_{t-1})\|_2^2 + 8\sum_{t=1}^{T} \|\nabla f_t(x_t) - \nabla F_t(x_t)\|_2^2 + 4L^2 \sum_{t=2}^{T} \|x_t - x_{t-1}\|_2^2.$$

B.2 Proof of Theorem 3.1

Proof. By Lemma B.1, we have

$$\begin{aligned}
&\sum_{t=1}^{T} \langle x_t - u, \nabla_t \rangle \\
&\leq \sum_{t=1}^{T} \langle x_t - x'_{t+1}, \nabla_t - m_t \rangle + D_{\psi_t}(u, x'_t) - D_{\psi_t}(u, x'_{t+1}) - D_{\psi_t}(x_t, x'_{t+1}) - D_{\psi_t}(x_t, x'_t), \\
&\leq \sum_{t=1}^{T} \langle x_t - x'_{t+1}, \nabla_t - m_t \rangle + D_{\psi_1}(u, x'_1) + \sum_{t=1}^{T-1} D_{\psi_{t+1}}(u, x'_{t+1}) - D_{\psi_t}(u, x'_{t+1}) \\
&\quad - \sum_{t=1}^{T} (D_{\psi_t}(x_t, x'_{t+1}) + D_{\psi_t}(x_t, x'_t))
\end{aligned}$$

Term $D_{\psi_1}(u, x'_1)$: Since the initialization of $x'_1 = 0$ and $A_1 = O(z_1^2 I)$, we have

$$D_{\psi_1}(u, x'_1) = \frac{1}{2}\|u\|_{A_1}^2 \leq O(z_1^2\|u\|_2^2).$$

Term $\sum_{t=1}^{T-1} D_{\psi_{t+1}}(u, x'_{t+1}) - D_{\psi_t}(u, x'_{t+1})$: First, we have

$$A_{t+1} - A_t = \eta_t(\nabla_t - m_t)(\nabla_t - m_t)^\top + 4\eta_t(z_{t+1}^2 - z_t^2)I.$$

Also,

$$\begin{aligned}
D_{\psi_{t+1}}(u, x'_{t+1}) - D_{\psi_t}(u, x'_{t+1}) &= \frac{1}{2}(u - x'_{t+1})^\top (A_{t+1} - A_t)(u - x'_{t+1}) \\
&= \frac{1}{2}(u - x'_{t+1})^\top (\eta_t(\nabla_t - m_t)(\nabla_t - m_t)^\top + 4\eta_t(z_{t+1}^2 - z_t^2)I)(u - x'_{t+1}) \\
&= \frac{\eta_t}{2}\left\langle u - x'_{t+1}, \nabla_t - m_t\right\rangle^2 + 2\eta_t(z_{t+1}^2 - z_t^2)\|u - x'_{t+1}\|_2^2.
\end{aligned}$$

Then, we have

$$\begin{aligned}
&\sum_{t=1}^{T-1} D_{\psi_{t+1}}(u, x'_{t+1}) - D_{\psi_t}(u, x'_{t+1}) \\
&\leq \sum_{t=1}^{T-1} \frac{\eta_t}{2}\langle u - x'_{t+1}, \nabla_t - m_t\rangle^2 + O(\sum_{t=1}^{T-1} \eta_t D^2(z_{t+1}^2 - z_t^2)), \\
&\leq \sum_{t=1}^{T-1} \frac{\eta_t}{2}\langle u - x'_{t+1}, \nabla_t - m_t\rangle^2 + O(\sum_{t=1}^{T-1} D(z_{t+1}^2 - z_t^2)/z_T), \qquad \text{(since } \eta_t \leq \frac{1}{64Dz_T}\text{)} \\
&\leq \sum_{t=1}^{T-1} \frac{\eta_t}{2}\langle u - x'_{t+1}, \nabla_t - m_t\rangle^2 + O(D(z_T - z_1)), \\
&\leq \sum_{t=1}^{T-1} \eta_t\langle u - x_t, \nabla_t - m_t\rangle^2 + \sum_{t=1}^{T-1} \eta_t\langle x_t - x'_{t+1}, \nabla_t - m_t\rangle^2 + O(D(z_T - z_1)), \\
&\leq \sum_{t=1}^{T-1} \eta_t\langle u - x_t, \nabla_t - m_t\rangle^2 + \sum_{t=1}^{T-1} \eta_t\langle x_t - x'_{t+1}, \nabla_t - m_t\rangle\langle x_t - x'_{t+1}, \nabla_t - m_t\rangle + O(D(z_T - z_1)), \\
&\leq \sum_{t=1}^{T-1} \eta_t\langle u - x_t, \nabla_t - m_t\rangle^2 + \sum_{t=1}^{T-1} 2\eta_t D z_t\langle x_t - x'_{t+1}, \nabla_t - m_t\rangle + O(D(z_T - z_1)), \qquad \text{(Using Equation 12)} \\
&\leq \sum_{t=1}^{T-1} \eta_t\langle u - x_t, \nabla_t - m_t\rangle^2 + O\Big(\frac{r\ln(T\eta_1 z_T/z_1)}{\eta_T} + D(z_T - z_1)\Big), \qquad \text{(Using Lemma B.3)} \\
&\leq \sum_{t=1}^{T-1} 2\eta_t\langle u, \nabla_t - m_t\rangle^2 + 2\eta_t\langle x_t, \nabla_t - m_t\rangle^2 + O\Big(\frac{r\ln(T\eta_1 z_T/z_1)}{\eta_T} + D(z_T - z_1)\Big), \\
&\leq \sum_{t=1}^{T-1} 8\eta_t\langle u, g_t - m_t\rangle^2 + 8\eta_t\langle x_t, g_t - m_t\rangle^2 + O\Big(\frac{r\ln(T\eta_1 z_T/z_1)}{\eta_T} + D(z_T - z_1)\Big). \qquad \text{(since } 32\eta_t\left\langle x_t, g_t - m_t\right\rangle \leq \frac{1}{2}\text{)}
\end{aligned}$$

Term $\sum_{t=1}^{T}(D_{\psi_t}(x_t, x'_{t+1}) + D_{\psi_t}(x_t, x'_t))$:

$$\begin{aligned}
\sum_{t=1}^{T}(D_{\psi_t}(x_t, x'_{t+1}) + D_{\psi_t}(x_t, x'_t)) &= \sum_{t=1}^{T} \frac{1}{2}(\|x_t - x'_t\|_{A_t}^2 + \|x'_{t+1} - x_t\|_{A_t}^2) \\
&\geq \frac{1}{2}\sum_{t=2}^{T} \|x_t - x'_t\|_{A_{t-1}}^2 + \frac{1}{2}\sum_{t=2}^{T+1} \|x_{t-1} - x'_t\|_{A_{t-1}}^2 \\
&\geq \frac{4z_1^2}{4}\sum_{t=2}^{T} \|x_t - x_{t-1}\|_2^2 = z_1^2\sum_{t=2}^{T} \|x_t - x_{t-1}\|_2^2,
\end{aligned}$$

where the last inequality comes from $A_t \succeq A_{t-1} \succeq 4z_1^2 I$.

Since $c_t(x)$ is convex in x, we have

$$\begin{aligned}&\sum_{t=1}^{T} c_t(x_t) - c_t(u)\\ &= \sum_{t=1}^{T}\langle x_t - u, g_t\rangle + 16\eta_t\langle x_t, g_t - m_t\rangle^2 - 16\eta_t\langle u, g_t - m_t\rangle^2 \le \sum_{t=1}^{T}\langle \nabla_t, x_t - u\rangle.\end{aligned}$$

Therefore, the final regret bound here is

$$\begin{aligned}\mathfrak{R}_T(u) &\le \sum_{t=1}^{T}\langle x_t - u, g_t\rangle\\ &\le O\Big(\frac{r\ln(T\eta_1 z_T/z_1)}{\eta_T} + z_1^2\|u\|_2^2 + D(z_T - z_1) + \sum_{t=1}^{T}\eta_t\langle u, g_t - m_t\rangle^2 - z_1^2\sum_{t=2}^{T}\|x_t - x_{t-1}\|_2^2\Big).\end{aligned}$$

□

B.3 Proof of Theorem 3.2

Proof. In this case, we have $\|g_t - m_t\|_2 \le 2G, \forall t \in [T]$. Then, we set $z_t = 2G, \forall t \in [T]$ and step-size η_t as

$$\eta_t = \min\Big\{\frac{1}{64Dz_T}, \frac{1}{D\sqrt{\sum_{s=1}^{t-1}\|g_s - m_s\|_2^2}}\Big\} \quad \text{for all} \quad t \in [T].$$

By substituting η_t and z_t into (3), we have

$$\mathfrak{R}_T(u) \le \widetilde{O}\Big(D\sqrt{\sum_{t=1}^{T-1}\|g_t - m_t\|_2^2} + G^2D^2 + \sum_{t=1}^{T}\eta_t\left\langle u, g_t - m_t\right\rangle^2 - G^2\sum_{t=2}^{T}\|x_t - x_{t-1}\|_2^2\Big).$$

By Lemma B.4, we have

$$\sum_{t=1}^{T}\eta_t\left\langle u, g_t - m_t\right\rangle^2 \le 2D\sqrt{\sum_{t=1}^{T}\|g_t - m_t\|_2^2}$$

Then,

$$\mathfrak{R}_T(u) \le \widetilde{O}\Big(D\sqrt{\sum_{t=1}^{T}\|g_t - m_t\|_2^2} - G^2\|x_t - x_{t-1}\|_2^2\Big).$$

By Lemma B.5 and the definition of g_t and m_t, we have

$$\begin{aligned}&D\sqrt{\sum_{t=1}^{T}\|g_t - m_t\|_2^2} - G^2\sum_{t=2}^{T}\|x_t - x_{t-1}\|_2^2\\ &\le DG + 2D\sqrt{\sum_{t=2}^{T}\|\nabla F_t(x_{t-1}) - \nabla F_{t-1}(x_{t-1})\|_2^2} + 2\sqrt{2}D\sqrt{\sum_{t=1}^{T}\|\nabla f_t(x_t) - \nabla F_t(x_t)\|_2^2}\\ &\quad + 2LD\sqrt{\sum_{t=2}^{T}\|x_t - x_{t-1}\|_2^2} - G^2\sum_{t=2}^{T}\|x_t - x_{t-1}\|_2^2\\ &\le DG + \frac{D^2L^2}{G^2} + 2D\sqrt{\sum_{t=2}^{T}\|\nabla F_t(x_{t-1}) - \nabla F_{t-1}(x_{t-1})\|_2^2} + 2\sqrt{2}D\sqrt{\sum_{t=1}^{T}\|\nabla f_t(x_t) - \nabla F_t(x_t)\|_2^2}\end{aligned}$$

Therefore, we have

$$\mathbb{E}[\mathfrak{R}_T(u)] \le \widetilde{O}(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})$$

$\square$

C Omitted Details of Section 4.1

C.1 Multi-scale Multiplicative-weight with Correction (MsMwC)

We rephrase the MsMwC algorithm [10] as the following Algorithm 5.

Algorithm 5 Multi-scale Multiplicative-weight with Correction (MsMwC)

Input: $w'_1 \in \Delta_N$.

1: **for** $t = 1, \ldots, T$ **do**
2: Receive the prediction $h_t \in \mathbb{R}^N$.
3: Compute $w_t = \arg\min_{w \in \Delta_N} \langle w, h_t \rangle + D_\phi(w, w'_t)$, where $\phi_t(w) = \sum_{j=1}^N \frac{w_j}{\beta_{t,j}} \ln w_j$.
4: Play w_t, receive ℓ_t and construct correction term $a_t \in \mathbb{R}^N$ with $a_{t,j} = 32\beta_{t,j}(\ell_t^j - m_t^j)^2$.
5: Compute $w'_{t+1} = \arg\min_{w \in \Delta_N} \langle w, \ell_t + a_t \rangle + D_\phi(w, w'_t)$.
6: **end for**

C.2 Proof of equation (5)

The final decision x_t is a weighted-average of base-learners' decisions: $x_t = \sum_{j=1}^N w_{t,j} x_t^j$. Then,

$$\begin{aligned}
\mathfrak{R}_T(u) &= \mathfrak{R}_T^{\mathcal{A}_i}(u) + \sum_{t=1}^T f_t(x_t) - \sum_{t=1}^T f_t(x_t^i) \\
&= \mathfrak{R}_T^{\mathcal{A}_i}(u) + \sum_{t=1}^T f_t(\sum_{j \in [N]} w_{t,j} x_t^j) - \sum_{t=1}^T f_t(x_t^i) \\
&\le \mathfrak{R}_T^{\mathcal{A}_i}(u) + \sum_{t=1}^T \sum_{j \in [N]} w_{t,j} f_t(x_t^j) - \sum_{t=1}^T f_t(x_t^i) \\
&= \mathfrak{R}_T^{\mathcal{A}_i}(u) + \sum_{t=1}^T \sum_{j \in [N]} f_t(x_t^j)[w_{t,j} - \mathbb{1}(j = i)] \\
&= \mathfrak{R}_T^{\mathcal{A}_i}(u) + \sum_{t=1}^T \sum_{j \in [N]} (f_t(x_t^j) - f_t(0))[w_{t,j} - \mathbb{1}(j = i)] \\
&\le \mathfrak{R}_T^{\mathcal{A}_i}(u) + \sum_{t=1}^T \sum_{j \in [N]} \langle \nabla f_t(x_t^j), x_t^j \rangle [w_{t,j} - \mathbb{1}(j = i)] \\
&= \mathfrak{R}_T^{\mathcal{A}_i}(u) + \sum_{t=1}^T \langle \ell_t, w_t - w_\star^i \rangle,
\end{aligned}$$

where $\ell_t \in \mathbb{R}^N$ with $\ell_t^j = \langle \nabla f_t(x_t^j), x_t^j \rangle$ and $w_\star^i$ is a vector in Δ_N whose j-th component is $(w_\star^i)_j = 1$ if $j = i$ and 0 otherwise.

C.3 Auxiliary Lemma

Lemma C.1. *(Theorem 6 in [10]) Suppose for all $t \in [T]$ and $j \in [N]$, $|\ell_t^j| \le GD_j$ and $|h_t^j| \le GD_j$, where $\ell_t^j = \langle \nabla f_t(x_t^j), x_t^j \rangle$ and $h_t^j = \langle \nabla f_{t-1}(x_{t-1}^j), x_t^j \rangle$. Define $\Gamma_j = \ln(\frac{NTD_j}{D_1})$ and the set $\mathcal{E}$ as*

$$\mathcal{E} = \Big\{(\beta_k, \mathcal{G}_k) : \forall k \in \mathcal{S}, \beta_k = \frac{1}{32 \cdot 2^k}\Big\},$$

where $\mathcal{G}_k$ is the MsMwC algorithm with w_1' being uniform over $\mathcal{Z}(k)$, $\mathcal{S} = \{k \in \mathbb{Z} : \exists j \in [N], GD_j \le 2^{k-2} \le GD_j\sqrt{T}\}$ and $\mathcal{Z}_k = \{j \in [N] : GD_j \le 2^{k-2}\}$. We have the following regret bound

$$\sum_{t=1}^{T}\langle \ell_t, w_t - w_\star^i\rangle \le O\Big(D_i\Gamma_i + \sqrt{\Gamma_i\sum_{t=1}^{T}(\ell_t^i - h_t^i)^2}\Big).$$

Proof. The regret $\sum_{t=1}^{T}\langle \ell_t, w_t - w_\star^i\rangle$ can also be decomposed as

$$\begin{aligned}\sum_{t=1}^{T}\langle \ell_t, w_t - w_\star^i\rangle &= \sum_{t=1}^{T}\left\langle \ell_t, w_t^{k_\star} - w_\star^i\right\rangle + \sum_{t=1}^{T}\left\langle \ell_t, w_t - w_t^{k_\star}\right\rangle \\ &= \sum_{t=1}^{T}\left\langle \ell_t, w_t^{k_\star} - w_\star^i\right\rangle + \sum_{t=1}^{T}\left\langle L_t, p_t - e_{k_\star}\right\rangle,\end{aligned}$$

where $e_{k_\star}$ is the $k_\star$-th standard basis vector and the second equality is from the definition of L_t.

For any $i \in [N]$, there exists a $k_\star$ such that $\eta_{k_\star} \le \min\left\{\frac{1}{128GD_i}, \sqrt{\frac{\Gamma_i}{\sum_{t=1}^{T}(\ell_t^i - h_t^i)^2}}\right\} \le 2\eta_{k_\star}$. By Lemma 1 and Theorem 4 in [10], we have

$$\sum_{t=1}^{T}\langle \ell_t, w_t - w_\star^i\rangle \le O\Big(D_i\Gamma_i + \sqrt{\Gamma_i\sum_{t=1}^{T}(\ell_t^i - h_t^i)^2}\Big).$$

□

C.4 Proof of Theorem 4.1

Proof. We begin with considering the first and second cases that $\|u\|_2 \le D_1$ and $\|u\|_2 \le D_i \le 2\|u\|_2$.

Here, we define $g_t^j = \nabla f_t(x_t^j)$ and $m_t^j = \nabla f_{t-1}(x_{t-1}^j)$ for all $t \in [T]$ and $j \in [N]$. For the meta regret, we define $\ell_t \in \mathbb{R}^N$ with $\ell_t^j = \langle \nabla f_t(x_t^j), x_t^j\rangle$ and $h_t \in \mathbb{R}^N$ with $h_t^j = \langle \nabla f_{t-1}(x_{t-1}^j), x_t^j\rangle$. Then, $|\ell_t^j| \le GD_j$ and $|h_t^j| \le GD_j$ for all $t \in [T]$ and $j \in [N]$. By applying Lemma C.1, we have

$$\sum_{t=1}^{T}\langle \ell_t, w_t - w_\star^i\rangle \le O\Big(D_i\Gamma_i + \sqrt{\Gamma_i\sum_{t=1}^{T}(\ell_t^i - h_t^i)^2}\Big).$$

By the definition of ℓ_t^i and h_t^i, we have

$$\begin{aligned}\sum_{t=1}^{T}(\ell_t^i - h_t^i)^2 &\le \sum_{t=1}^{T}\langle \nabla f_t(x_t^i) - \nabla f_{t-1}(x_{t-1}^i), x_t^i\rangle^2 \\ &\le D_i^2\sum_{t=1}^{T}\|\nabla f_t(x_t^i) - \nabla f_{t-1}(x_{t-1}^i)\|_2^2 = D_i^2\sum_{t=1}^{T}\|g_t^i - m_t^i\|_2^2.\end{aligned}$$

Hence,

$$\sum_{t=1}^{T}\langle \ell_t, w_t - w_\star^i\rangle \le O\Big(D_i\Gamma_i + D_i\sqrt{\Gamma_i\sum_{t=1}^{T}\|g_t^i - m_t^i\|_2^2}\Big). \tag{13}$$

Then, we investigate the expert regret part. Here, we set the step-size for the expert i as

$$\eta_t^i = \min\Big\{\frac{1}{64D_iz_T}, \frac{1}{D_i\sqrt{\sum_{s=1}^{t-1}\|g_s^i - m_s^i\|_2^2}}\Big\}. \tag{14}$$

By substituting the step-size specified in (14) to (3), we have

$$\mathfrak{R}_T^{\mathcal{A}_i}(u) \le \widetilde{O}\Big(D_i\sqrt{\sum_{t=1}^T \|g_t^i - m_t^i\|_2^2 + G^2\|u\|_2^2} + \frac{\|u\|_2^2}{D_i}\sqrt{\sum_{t=1}^T \|g_t^i - m_t^i\|_2^2 - G^2\|x_t^i - x_{t-1}^i\|_2^2}\Big)$$
$$\le \widetilde{O}\Big(D_i^2 + D_i\sqrt{\sum_{t=1}^T \|g_t^i - m_t^i\|_2^2 - G^2\|x_t^i - x_{t-1}^i\|_2^2}\Big). \tag{15}$$

Therefore, by combining (13) and (15), we have

$$\mathfrak{R}_T(u) \le \widetilde{O}\Big(D_i^2 + D_i\sqrt{\sum_{t=1}^T \|g_t^i - m_t^i\|_2^2} - G^2\sum_{t=2}^T \|x_t^i - x_{t-1}^i\|_2^2\Big)$$

Then, by applying Lemma B.5, we have

$$D_i\sqrt{\sum_{t=1}^T \|g_t^i - m_t^i\|_2^2} - G^2\sum_{t=2}^T \|x_t^i - x_{t-1}^i\|_2^2$$
$$\le GD_i + \frac{D_i^2L^2}{G^2} + 2D_i\sqrt{\sum_{t=2}^T \|\nabla F_t(x_{t-1}^i) - \nabla F_{t-1}(x_{t-1}^i)\|_2^2} + 2\sqrt{2}D_i\sqrt{\sum_{t=1}^T \|\nabla f_t(x_t^i) - \nabla F_t(x_t^i)\|_2^2}.$$

Since the expert i runs OONS within the set $\mathcal{X}_i = \{x : \|x\|_2 \le D_i\} \subseteq \mathcal{X}$, we have

$$\sup_{x\in\mathcal{X}_i} \mathbb{E}_{f_t\sim\mathcal{D}_t}\left[\|\nabla f_t(x) - \nabla F_t(x)\|_2^2\right] \le \sup_{x\in\mathcal{X}} \mathbb{E}_{f_t\sim\mathcal{D}_t}\left[\|\nabla f_t(x) - \nabla F_t(x)\|_2^2\right]$$
$$\sup_{x\in\mathcal{X}_i} \|\nabla F_t(x) - \nabla F_{t-1}(x)\|_2^2 \le \sup_{x\in\mathcal{X}} \|\nabla F_t(x) - \nabla F_{t-1}(x)\|_2^2.$$

Hence,

$$\mathbb{E}[\mathfrak{R}_T(u)] \le \widetilde{O}\Big(D_i^2 + \frac{D_i^2L^2}{G^2} + D_i\sqrt{\sigma_{1:T}^2} + D_i\sqrt{\Sigma_{1:T}^2}\Big). \tag{16}$$

Case 1: $\|u\|_2 \le D_1$. We take $i = 1$ and substitute D_i with D_1 into (16). Hence, we have

$$\mathbb{E}[\mathfrak{R}_T(u)] \le \widetilde{O}\Big(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2}\Big).$$

Case 2: In this case, let i be the smallest integer such that $\|u\|_2 \le D_i = 2^i$. We have $\|u\|_2 \le D_i \le 2\|u\|_2$ since $D_{i+1} = 2D_i$. Then, we substitute D_i with $2\|u\|_2$ into the regret bound (16). Then,

$$\mathbb{E}[\mathfrak{R}_T(u)] \le \widetilde{O}\Big(\|u\|_2^2 + \|u\|_2(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})\Big).$$

Case 3: $\|u\|_2 > D_{\max}$.

Next, we consider the case when $\|u\|_2 > D_{\max} = 2^N$. Then, we have

$$\sum_{t=1}^T f_t(x_t) - f_t(u) \le \sum_{t=1}^T \langle \nabla f_t(x_t), x_t - u\rangle$$
$$\le 2GT\|u\|_2.$$

We take $N = \lceil \log T\rceil$, then $T \le \|u\|_2$. Therefore,

$$\sum_{t=1}^T f_t(x_t) - f_t(u) \le 2G\|u\|_2^2.$$

By combining these two cases above, the desirable regret bound is achieved. □

D Omitted Details of Section 4.2

In this section, we also denote $g_t = \nabla f_t(x_t)$ and $m_t = \nabla f_{t-1}(x_{t-1})$. We first note that

$$\max_{t \le T} D_t < \sqrt{\sum_{s=1}^{t} \frac{\|g_s\|_2}{\max\{1, \max_{k \le s} \|g_k\|_2\}}} \le \sqrt{T}\,.$$

Thus, we need to update D_t at most $O(\log T)$ times. Let M be the number of total updates in D_t, where $M = \mathcal{O}(\log T)$. We split the T iterations into M intervals I_m with $m \in [M]$, where the last iteration of I_m (denoted by t_m) either equals to T or $D_{t_m+1} \neq D_{t_m}$.

D.1 Proof of Theorem 4.4

Proof. We have

$$\begin{aligned}
\mathfrak{R}_T(u) &= \sum_{t=1}^{T} f_t(x_t) - \sum_{t=1}^{T} f_t(u) \le \sum_{t=1}^{T} \langle g_t, x_t - u\rangle \\
&= \sum_{m=1}^{M} \underbrace{\sum_{t \in I_m} \langle g_t, x_t - u_t\rangle}_{T_m} + \underbrace{\sum_{t=1}^{T} \langle g_t, u_t - u\rangle}_{T_{\text{extra}}},
\end{aligned}$$

where we define $u_t = \min\{1, \frac{D_t}{\|u\|_2}\} u$.

We first consider T_m as

$$\sum_{t \in I_m} \langle g_t, x_t - u_t\rangle = \sum_{t \in I_m} \langle \widetilde{g}_t, x_t - u_t\rangle + \sum_{t \in I_m} \langle g_t - \widetilde{g}_t, x_t - u_t\rangle\,.$$

Note that iteration t within interval I_m, i.e., $t \in I_m$, the domain has a bounded diameter D_t. When $t \in I_m$, we take

$$\eta_t = \min\{\frac{1}{64 D_t z_t}, \frac{1}{\sqrt{\sum_{s=t_1}^{t-1} \|\widetilde{g}_s - m_s\|_2^2}}\},$$

where t_1 is first index in I_m. Also, we denote the last index in I_m as t_m, respectively. From the Line 5 or 8 in FPF-OONS we need to reset $x'_{t_1} = 0$ and A_t for all $t \in I_m$ at iteration t_1 as follows

$$A_t = 4 z_{t_1}^2 I + \sum_{s=t_1}^{t-1} \eta_s (\nabla_s - m_s)(\nabla_s - m_s)^\top + 4\eta_t z_t^2 I.$$

Similar to the proof of Theorem 3.1, we have

$$\begin{aligned}
&\sum_{t=t_1}^{t_m} \langle x_t - u_t, \nabla_t\rangle \\
&\le \sum_{t=t_1}^{t_m} \langle x_t - x'_{t+1}, \nabla_t - m_t\rangle + D_{\psi_{t_1}}(u, x'_{t_1}) + \sum_{t=t_1}^{t_m-1} D_{\psi_{t+1}}(u_t, x'_{t+1}) - D_{\psi_t}(u_t, x'_{t+1}) \\
&\quad - \sum_{t=t_1}^{t_m} (D_{\psi_t}(x_t, x'_{t+1}) + D_{\psi_t}(x_t, x'_t)),
\end{aligned}$$

where $\nabla_t = \widetilde{g}_t + 32\eta_t \langle x_t, \widetilde{g}_t - m_t\rangle (\widetilde{g}_t - m_t)$.

We first consider the term $\sum_{t=t_1}^{t_m} \langle x_t - x'_{t+1}, \nabla_t - m_t\rangle$. By Lemma B.3, we have $0 \le \langle x_t - x'_{t+1}, \nabla_t - m_t\rangle \le 2\|\nabla_t - m_t\|^2_{A_t^{-1}}$.

Also, we have

$$\begin{aligned}\|\widetilde{g}_t - m_t\|_2 &= \|m_t + \frac{B_{t-1}}{B_t}(g_t - m_t) - m_t\|_2 \\ &\le \|g_t - m_t\|_2\end{aligned}$$

By the definition of $\nabla_t = \widetilde{g}_t + 32\eta_t\langle x_t, \widetilde{g}_t - m_t\rangle(\widetilde{g}_t - m_t)$ and $\|\widetilde{g}_t - m_t\|_2 \le \|g_t - m_t\|_2$, we have

$$\begin{aligned}\|\nabla_t - m_t\|_2 &= \|\widetilde{g}_t - m_t + 32\eta_t\langle x_t, \widetilde{g}_t - m_t\rangle(\widetilde{g}_t - m_t)\|_2 \\ &\le \|\widetilde{g}_t - m_t\|_2 + 32\eta_t D_t\|\widetilde{g}_t - m_t\|_2^2 \\ &\le \|\widetilde{g}_t - m_t\|_2 + 32\eta_t D_t z_t\|\widetilde{g}_t - m_t\|_2 \\ &\le \frac{3}{2}\|\widetilde{g}_t - m_t\|_2 \le \frac{3}{2}\|g_t - m_t\|_2.\end{aligned}$$

For $t \in I_m$, we also redefine

$$\bar{A}_t = 4z_{t_1}^2 \cdot I + \sum_{s=t_1}^{t}\eta_s(\nabla_s - m_s)(\nabla_s - m_s)^\top.$$

Hence, $A_t \succeq \bar{A}_t$ since $\|\nabla_t - m_t\|_2^2 \le 4\|\widetilde{g}_t - m_t\|_2^2 \le 4z_t^2$. Also, for $t \in [t_1+1, t_m]$, we have

$$(\nabla_t - m_t)(\nabla_t - m_t)^\top = \frac{1}{\eta_t}[\eta_t(\nabla_t - m_t)(\nabla_t - m_t)^\top] = \frac{1}{\eta_t}(\bar{A}_t - \bar{A}_{t-1})$$

Then,

$$\begin{aligned}\sum_{t=t_1}^{t_m}\|\nabla_t - m_t\|_{A_t^{-1}}^2 &= \|\nabla_{t_1} - m_{t_1}\|_{A_{t_1}^{-1}}^2 + \sum_{t=t_1+1}^{t_m}\|\nabla_t - m_t\|_{A_t^{-1}}^2 \\ &= \frac{1}{4z_{t_1}^2}\|\nabla_{t_1} - m_{t_1}\|_2^2 + \sum_{t=t_1+1}^{t_m}\|\nabla_t - m_t\|_{A_t^{-1}}^2 \\ &\le 1 + \sum_{t=t_1+1}^{t_m}\|\nabla_t - m_t\|_{A_t^{-1}}^2 \qquad (\text{since } \|\nabla_{t_1} - m_{t_1}\|_2^2 \le 4z_{t_1}^2) \\ &\le \sum_{t=t_1+1}^{t_m}\|\nabla_t - m_t\|_{\bar{A}_t^{-1}}^2 + 1 \\ &\le \sum_{t=t_1+1}^{t_m}\frac{1}{\eta_t}(\ln|\bar{A}_t| - \ln|\bar{A}_{t-1}|) + 1 \\ &\le \frac{1}{\eta_{t_m}}\ln\frac{|\bar{A}_{t_m}|}{|\bar{A}_{t_1}|} + 1.\end{aligned}$$

For $|\bar{A}_{t_m}|$:

$$|\bar{A}_{t_m}| \le |4z_{t_1}^2 I + \sum_{t=t_1}^{t_m}\eta_t(\nabla_t - m_t)(\nabla_t - m_t)^\top|.$$

$$\begin{aligned}\ln|\bar{A}_{t_m}| &\le O\Big(r\ln(1 + \sum_{t=t_1}^{t_m}\frac{\eta_t}{4z_{t_1}^2}\|\nabla_t - m_t\|_2^2)\Big) \\ &\le O\Big(r\ln(1 + \frac{4z_{t_m}^2}{4z_{t_1}^2}\sum_{t=t_1}^{t_m}\eta_t)\Big) \\ &\le O\Big(r\ln(1 + \frac{\eta_{t_1}z_{t_m}^2}{z_{t_1}^2}T)\Big).\end{aligned}$$

Therefore, we have

$$\sum_{t=t_1}^{t_m}\langle x_t - x'_{t+1}, \nabla_t - m_t\rangle \le O\Big(\frac{r\ln(T\eta_{t_1}z_{t_m}/z_{t_1})}{\eta_{t_m}}\Big).$$

Term $D_{\psi_{t_1}}(u, x'_{t_1})$:

$$D_{\psi_{t_1}}(u_t, x'_{t_1}) = \frac{1}{2}\|u_t\|^2_{A_{t_1}} \le O(z_{t_1}^2\|u_t\|_2^2).$$

Term $\sum_{t=t_1}^{t_m-1} D_{\psi_{t+1}}(u_t, x'_{t+1}) - D_{\psi_t}(u_t, x'_{t+1})$:

By the definition of $u_t = \min\{1, \frac{D_t}{\|u\|_2}\}u$, we have

$$\|u_t\|_2 = \min\{\|u\|_2, D_t\}.$$

Then, we have

$$\begin{aligned}
&\sum_{t=t_1}^{t_m-1} D_{\psi_{t+1}}(u_t, x'_{t+1}) - D_{\psi_t}(u_t, x'_{t+1})\\
&\le \sum_{t=t_1}^{t_m-1}\frac{\eta_t}{2}\left\langle u_t - x'_{t+1}, \nabla_t - m_t\right\rangle^2 + O(\sum_{t=t_1}^{t_m-1}\eta_t\|u_t - x'_{t+1}\|_2^2(z_{t+1}^2 - z_t^2))\\
&\le \sum_{t=t_1}^{t_m-1}\frac{\eta_t}{2}\left\langle u_t - x'_{t+1}, \nabla_t - m_t\right\rangle^2 + O(\sum_{t=t_1}^{t_m-1}\eta_t D_t^2(z_{t+1}^2 - z_t^2)) \qquad \text{(since } \|u_t\|_2 \le \|D_t\|_2\text{)}\\
&\le \sum_{t=t_1}^{t_m-1}\frac{\eta_t}{2}\left\langle u_t - x'_{t+1}, \nabla_t - m_t\right\rangle^2 + O(\sum_{t=t_1}^{t_m-1} D_t(z_{t+1}^2 - z_t^2)/z_t) \qquad \text{(since } \eta_t \le \frac{1}{64D_tz_t}\text{)}\\
&\le \sum_{t=t_1}^{t_m-1}\frac{\eta_t}{2}\left\langle u_t - x'_{t+1}, \nabla_t - m_t\right\rangle^2 + O(D_{t_m}z_{t_m}^2) \qquad \text{(since } z_t \ge z_1 = 1\text{)}\\
&\le \sum_{t=t_1}^{t_m-1} 8\eta_t\langle u_t, \widetilde{g}_t - m_t\rangle^2 + 8\eta_t\langle x_t, \widetilde{g}_t - m_t\rangle^2 + O\Big(\frac{r\ln(T\eta_{t_1}z_{t_m}/z_{t_1})}{\eta_{t_m}} + D_{t_m}z_{t_m}^2\Big).
\end{aligned}$$

Term $\sum_{t=t_1}^{t_m}(D_{\psi_t}(x_t, x'_{t+1}) + D_{\psi_t}(x_t, x'_t))$:

$$\begin{aligned}
\sum_{t=t_1}^{t_m}(D_{\psi_t}(x_t, x'_{t+1}) + D_{\psi_t}(x_t, x'_t)) &= \sum_{t=t_1}^{t_m}\frac{1}{2}(\|x_t - x'_t\|^2_{A_t} + \|x'_{t+1} - x_t\|^2_{A_t})\\
&\ge \frac{1}{2}\sum_{t=t_1+1}^{t_m}\|x_t - x'_t\|^2_{A_{t-1}} + \frac{1}{2}\sum_{t=t_1+1}^{t_m+1}\|x_{t-1} - x'_t\|^2_{A_{t-1}}\\
&\ge \frac{4z_{t_1}^2}{4}\sum_{t=t_1+1}^{t_m}\|x_t - x_{t-1}\|_2^2 = z_1^2\sum_{t=t_1+1}^{t_m}\|x_t - x_{t-1}\|_2^2,
\end{aligned}$$

where the last inequality comes from $A_t \succeq A_{t-1} \succeq 4z_{t_1}^2 I$ when $t \in [t_1 + 1, t_m]$.

Then, we have

$$\begin{aligned}
&\sum_{t\in I_m}\langle \widetilde{g}_t, x_t - u_t\rangle\\
&\le O\Big(\frac{r\ln(T\eta_{t_1}z_{t_m}/z_{t_1})}{\eta_{t_m}} + z_{t_1}^2\|u_t\|_2^2 + D_tz_{t_m}^2 + \sum_{t=t_1}^{t_m-1}\eta_t\langle u_t, \widetilde{g}_t - m_t\rangle^2 - z_{t_1}^2\sum_{t_1+1}^{t_m}\|x_t - x_{t-1}\|_2^2\Big)\\
&\le \widetilde{O}\Big(\sqrt{\sum_{t\in I_m}\|\widetilde{g}_t - m_t\|_2^2} + z_{t_1}^2\|u\|_2^2 + z_{t_m}^2 D_t + \|u\|_2^2\sqrt{\sum_{t\in I_m}\|\widetilde{g}_t - m_t\|_2^2} - z_{t_1}^2\sum_{t_1+1}^{t_m}\|x_t - x_{t-1}\|_2^2\Big).
\end{aligned}$$

Furthermore, by $\|\widetilde{g}_t - m_t\|_2 \le \|g_t - m_t\|_2$, we have

$$\begin{aligned}\sum_{t \in I_m} \langle g_t - \widetilde{g}_t, x_t - u_t \rangle \le (D_t + \|u\|_2) \sum_{t \in I_m} \|g_t - \widetilde{g}_t\| \le (D_t + \|u\|_2) \sum_{t \in I_m} \frac{B_t - B_{t-1}}{B_t} \|g_t - m_t\|_2 \\ \le 2(D_t + \|u\|_2) G.\end{aligned}$$

At the last iteration T, we have

$$\begin{aligned} D_T &< \sqrt{\sum_{t=1}^{T} \frac{\|g_t\|_2}{\max\{1, \max_{k \le t} \|g_k\|_2\}}} \\ &\le \sqrt{\sum_{t=1}^{T} \|g_t\|_2} \\ &\le \sqrt{\sum_{t=1}^{T} \|g_t - \mathbb{E}[g_t]\|_2 + \|\nabla F_t(x_t)\|_2} \\ &= \sqrt{\sum_{t=1}^{T} \|\nabla f_t(x_t) - \nabla F_t(x_t)\|_2 + \|\nabla F_t(x_t)\|_2}. \end{aligned}$$

Therefore, we have

$$\begin{aligned} &\sum_{m=1}^{M} \sum_{t \in I_m} \langle g_t, x_t - u_t \rangle \le \widetilde{O}\Big(\|u\|_2^2 \sqrt{\sum_{t \in [T]} \|g_t - m_t\|_2^2} + G^2 D_T + G^2 \|u\|_2^2 - z_1^2 \sum_{t=2}^{T} \|x_t - x_{t-1}\|_2^2 \Big) \\ &\le \widetilde{O}\Big(\|u\|_2^2 \sqrt{\sum_{t \in [T]} \|g_t - m_t\|_2^2} + G^2 D_T + G^2 \|u\|_2^2 - \sum_{t=2}^{T} \|x_t - x_{t-1}\|_2^2 \Big). \qquad \text{(since } z_1 = B_0 = 1\text{)} \end{aligned}$$

Now, we bound T_{extra}. We observe that u_t is either u or $\frac{D_t}{\|u\|_2} u$. When $u_t \neq u$, $\|u\|_2 \ge D_t > \sqrt{\sum_{s=1}^{t} \frac{\|g_s\|_2}{\max_{k \le s} \|g_k\|_2}}$. Once $u_t = u$, it stays there. Let t^* be the last round when $u_t \neq u$.

$$\begin{aligned} \sum_{t=1}^{T} \langle g_t, u_t - u \rangle = \sum_{t=1}^{t^*} \langle g_t, u_t - u \rangle &\le \sum_{t=1}^{t^*-1} \langle g_t, u_t - u \rangle + 2\|u\|_2 G \\ &\le 2\|u\|_2 \sum_{t=1}^{t^*-1} \|g_t\|_2 + 2\|u\|_2 G \\ &\le 2\|u\|_2 G \sum_{t=1}^{t^*-1} \frac{\|g_t\|_2}{\max_{k \le t^*-1} \|g_k\|_2} + 2\|u\|_2 G \\ &\le 2\|u\|_2^3 G + 2\|u\|_2 G. \end{aligned}$$

Therefore, we have

$$\mathfrak{R}_T(u) \le \widetilde{O}\Big(\|u\|_2^2 \sqrt{\sum_{t \in [T]} \|g_t - m_t\|_2^2} + G^2 D_T + G^2 \|u\|_2^2 + G\|u\|_2^3 - \sum_{t=2}^{T} \|x_t - x_{t-1}\|_2^2 \Big)$$

By Lemma B.5, we have

$$
\begin{aligned}
&\|u\|_2^2\sqrt{\sum_{t=1}^{T}\|g_t-m_t\|_2^2}-\sum_{t=2}^{T}\|x_t-x_{t-1}\|_2^2 \\
&\leq G^2\|u\|_2^2+2\|u\|_2^2\sqrt{\sum_{t=2}^{T}\|\nabla F_t(x_{t-1})-\nabla F_{t-1}(x_{t-1})\|_2^2} \\
&\quad+2\sqrt{2}\|u\|_2^2\sqrt{\sum_{t=1}^{T}\|\nabla f_t(x_t)-\nabla F_t(x_t)\|_2^2}+2L\|u\|_2^2\sqrt{\sum_{t=2}^{T}\|x_t-x_{t-1}\|_2^2}-\sum_{t=2}^{T}\|x_t-x_{t-1}\|_2^2 \\
&\leq G^2\|u\|_2^2+L^2\|u\|_2^4+2\|u\|_2^2\sqrt{\sum_{t=2}^{T}\|\nabla F_t(x_{t-1})-\nabla F_{t-1}(x_{t-1})\|_2^2} \\
&\quad+2\sqrt{2}\|u\|_2^2\sqrt{\sum_{t=1}^{T}\|\nabla f_t(x_t)-\nabla F_t(x_t)\|_2^2}
\end{aligned}
$$

Therefore, we can conclude that

$$
\mathbb{E}[\mathfrak{R}_T(u)] \leq \widetilde{O}\Big(\|u\|_2^2(\sqrt{\sigma_{1:T}^2}+\sqrt{\Sigma_{1:T}^2})+G^2\sqrt{\sigma_{1:T}+\mathfrak{G}_{1:T}}+G^2\|u\|_2^2+\|u\|_2^4+G\|u\|_2^3\Big),
$$

where $\sigma_{1:T}$ captures the stochastic gradient deviation (without the squared norm) and $\mathfrak{G}_{1:T}$ denotes the sum of maximum expected gradients. $\square$

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction accurately reflect the paper's contributions and scope.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: The limitations of our regret bound are discussed in Remarks 4.2, 4.3, and 4.5.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

Justification: The problem setting is clearly introduced in Section 2 and all theoretical results are accompanied by proofs.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [NA]

   Justification: This paper is a fully theoretical paper without an experimental section.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [NA]

Justification: This paper is a fully theoretical paper without an experimental section.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so “No” is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyper-parameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [NA]

   Justification: This paper is a fully theoretical paper without an experimental section.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [NA]

   Justification: This paper is a fully theoretical paper without an experimental section.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
   - The assumptions made should be given (e.g., Normally distributed errors).
   - It should be clear whether the error bar is the standard deviation or the standard error of the mean.

- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [NA]

   Justification: This paper is a fully theoretical paper without an experimental section.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The authors confirm with the NeurIPS Code of Ethics.

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [NA]

    Justification: The work is mainly theoretical, thus there is no identifiable societal impact.

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
    - The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to

generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.

- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification: The work is mainly theoretical.

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [NA]

Justification: The paper does not use existing assets.

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [NA]

Justification: The paper does not release new assets.

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

    Answer: [NA]

Justification: The core method development in this research does not involve LLMs as any important, original, or non-standard components.

Guidelines:

- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy (`https://neurips.cc/Conferences/2025/LLM`) for what should or should not be described.