Optimistic Online Mirror Descent for Bridging Stochastic and Adversarial Online Convex Optimization

Sijia Chen [1 2] **Wei-Wei Tu** [1 3 4] **Peng Zhao** [1] **Lijun Zhang** [1 2]

Abstract

Stochastically Extended Adversarial (SEA) model is introduced by Sachs et al. (2022) as an interpolation between stochastic and adversarial online convex optimization. Under the smoothness condition, they demonstrate that the expected regret of optimistic follow-the-regularized-leader (FTRL) depends on the *cumulative stochastic variance* $\sigma^2_{1:T}$ and the *cumulative adversarial variation* $\Sigma^2_{1:T}$ for convex functions. They also provide a slightly weaker bound based on the *maximal stochastic variance* $\sigma^2_{\max}$ and the *maximal adversarial variation* $\Sigma^2_{\max}$ for strongly convex functions. Inspired by their work, we investigate the theoretical guarantees of optimistic online mirror descent (OMD) for the SEA model. For convex and smooth functions, we obtain the *same* $\mathcal{O}(\sqrt{\sigma^2_{1:T}} + \sqrt{\Sigma^2_{1:T}})$ regret bound, without the convexity requirement of individual functions. For strongly convex and smooth functions, we establish an $\mathcal{O}(\min\{\log(\sigma^2_{1:T} + \Sigma^2_{1:T}), (\sigma^2_{\max} + \Sigma^2_{\max})\log T\})$ bound, *better* than their $\mathcal{O}((\sigma^2_{\max} + \Sigma^2_{\max})\log T)$ result. For exp-concave and smooth functions, we achieve a *new* $\mathcal{O}(d\log(\sigma^2_{1:T} + \Sigma^2_{1:T}))$ bound. Owing to the OMD framework, we further establish *dynamic regret* for convex and smooth functions, which is more favorable in non-stationary online scenarios.

1. Introduction

Online convex optimization (OCO) is a fundamental framework for online learning and has been applied in a variety of real-world applications such as spam filtering and portfolio management (Hazan, 2016). OCO problems can be mainly divided into two categories: adversarial online convex optimization (adversarial OCO) (Zinkevich, 2003; Hazan et al., 2007) and stochastic online convex optimization (SCO) (Nemirovski et al., 2009; Hazan & Kale, 2011; Lan, 2012). Adversarial OCO assumes that the loss functions are chosen arbitrarily or adversarially and the goal is to minimize the regret. SCO assumes that the loss functions are independently and identically distributed (i.i.d.), and the goal is to minimize the excess risk. Although the two models have been extensively studied (Shalev-Shwartz et al., 2009; Shapiro et al., 2014; Hazan, 2016; Orabona, 2019), in real scenarios the nature is not always completely adversarial or stochastic, but often lies somewhere in between.

Recently, Sachs et al. (2022) introduce the Stochastically Extended Adversarial (SEA) model, in which the nature chooses distribution $\mathcal{D}_t$ and the learner suffers a loss $f_t(\mathbf{x}_t)$ where $f_t \sim \mathcal{D}_t$ for each round t. The distributions are allowed to vary over time, and by choosing them appropriately, SEA reduces to adversarial OCO, SCO, or other intermediate settings. To analyze the performance, they propose to use two quantities — *cumulative stochastic variance* $\sigma^2_{1:T}$ and *cumulative adversarial variation* $\Sigma^2_{1:T}$ — to bound the expected regret, which measure how stochastic or adversarial the distributions are. For convex and smooth functions, Sachs et al. (2022) prove that optimistic follow-the-regularized-leader (FTRL) enjoys an $\mathcal{O}(\sqrt{\sigma^2_{1:T}} + \sqrt{\Sigma^2_{1:T}})$ expected regret bound, from which we can derive an excess risk bound for SCO (Lan, 2012) and a gradient-variation regret bound for adversarial OCO (Chiang et al., 2012) (see details in Section 3). However, for the strongly convex case, Sachs et al. (2022) only establish a weak bound of $\mathcal{O}((\sigma^2_{\max} + \Sigma^2_{\max})\log T)$, which depends on the *maximal* stochastic variance $\sigma^2_{\max}$ and the *maximal* adversarial variation $\Sigma^2_{\max}$ instead of the cumulative counterparts.

Optimistic FTRL belongs to optimistic online learning algorithms (Rakhlin & Sridharan, 2013), which aim to exploit prior knowledge during the online process. Besides optimistic FTRL, optimistic online mirror descent (OMD) is another popular framework, and in fact, the gradient-variation bound of Chiang et al. (2012) is derived from optimistic OMD. Given the encouraging results of optimistic

[1]National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China [2]Peng Cheng Laboratory, Shenzhen 518055, China [3]4Paradigm, Beijing, China [4]ChaLearn, USA. Correspondence to: Peng Zhao <zhaop@lamda.nju.edu.cn>, Lijun Zhang <zhanglj@lamda.nju.edu.cn>.

Proceedings of the 40^{th} International Conference on Machine Learning, Honolulu, Hawaii, USA. PMLR 202, 2023.

FTRL (Sachs et al., 2022), it is natural to ask what are the theoretical guarantees of optimistic OMD for SEA, and we provide answers below.

- For convex and smooth functions, optimistic OMD enjoys an $\mathcal{O}(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})$ expected regret bound, which is the same as that of Sachs et al. (2022) but without the convexity condition of individual functions.
- For strongly convex and smooth functions, optimistic OMD achieves an $\mathcal{O}(\min\{\log(\sigma_{1:T}^2 + \Sigma_{1:T}^2), (\sigma_{\max}^2 + \Sigma_{\max}^2)\log T\})$ expected regret bound, which is better than the $\mathcal{O}((\sigma_{\max}^2 + \Sigma_{\max}^2)\log T)$ bound of Sachs et al. (2022) for optimistic FTRL when $\sigma_{1:T}^2$ and $\Sigma_{1:T}^2$ are small, and is no worse than theirs in any case.
- For exp-concave and smooth functions, our work establishes a new $\mathcal{O}(d\log(\sigma_{1:T}^2 + \Sigma_{1:T}^2))$ bound for optimistic OMD, where d denotes the dimensionality.
- Our better results for optimistic OMD are due to more careful analyses and do not suggest that optimistic OMD is inherently superior to optimistic FTRL for regret minimization. Indeed, we present a different analysis of optimistic FTRL when encountering convex functions, which eliminates the convexity condition of individual functions; similarly, we also provide new analyses for strongly convex functions and exp-concave functions respectively, both obtaining the same expected regret bounds as optimistic OMD.

Based on our theoretical guarantees, we apply optimistic OMD to many intermediate examples between adversarial OCO and SCO, leading to *better* results for strongly convex functions and *new* results for exp-concave functions, hence deepening our understanding of the intermediate scenarios.

Furthermore, owing to the OMD framework, we can generalize our results to *dynamic regret*, a strengthened measure comparing the online performance with a sequence of *changing* comparators (Zinkevich, 2003). For convex and smooth functions, we obtain an $\mathcal{O}(P_T + \sqrt{1+P_T}(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2}))$ expected dynamic regret bound for SEA, where P_T denotes the path length. The bound is new and immediately recovers the $\mathcal{O}(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})$ expected regret bound by setting $P_T = 0$, and reduces to the $\mathcal{O}(\sqrt{T(1+P_T)})$ dynamic regret of Zhang et al. (2018) in the adversarial setting. We regard the support of dynamic regret as an advantage of optimistic OMD over optimistic FTRL. To the best of our knowledge, even the $\mathcal{O}(\sqrt{T(1+P_T)})$ dynamic regret has not been established for FTRL-style methods in online convex optimization.

2. Related Work

In this section, we review the related work in adversarial OCO and SCO, as well as studies on intermediate states.

2.1. Adversarial Online Convex Optimization

Adversarial OCO can be seen as a repeated game between the online learner and the nature (or called environments). In round $t \in [T]$, the online learner chooses a decision $\mathbf{x}_t$ from the convex feasible set $\mathcal{X} \subseteq \mathbb{R}^d$, and suffers a convex loss $f_t(\mathbf{x}_t)$ which may be adversarially selected by the nature. In adversarial OCO, we aim to minimize the *regret*:

$$\mathbf{Reg}_T = \sum_{t=1}^{T} f_t(\mathbf{x}_t) - \min_{\mathbf{x}\in\mathcal{X}} \sum_{t=1}^{T} f_t(\mathbf{x}),$$

which measures the cumulative loss difference between the learner and the best decision in hindsight (Orabona, 2019). For the adversarial OCO, an $\mathcal{O}(\sqrt{T})$ regret bound is achieved by online gradient descent (OGD) with step size $\eta_t = \mathcal{O}(1/\sqrt{t})$ (Zinkevich, 2003). For λ-strongly convex functions, an $\mathcal{O}(\frac{1}{\lambda}\log T)$ bound is achieved by changing the step size η_t of OGD to $\mathcal{O}(1/[\lambda t])$ (Shalev-Shwartz, 2007). For α-exp-concave functions, online Newton step (ONS) (Hazan et al., 2007) obtains an $\mathcal{O}(\frac{d}{\alpha}\log T)$ bound. Those results are minimax optimal (Ordentlich & Cover, 1998; Abernethy et al., 2008) and can not be improved in general.

Furthermore, various algorithms are proposed to achieve *problem-dependent* regret guarantees, which can safeguard the minimax rates in the worst case and become better when problems satisfy benign properties such as smoothness (Srebro et al., 2010; Chiang et al., 2012; Orabona et al., 2012), sparsity (Duchi et al., 2011), or other structural properties (Yang et al., 2014; Kingma & Ba, 2015; Défossez et al., 2022). Among them, Chiang et al. (2012) demonstrate that the regret of smooth functions can be upper bounded by the gradient-variation defined as

$$V_T = \sum_{t=2}^{T} \sup_{\mathbf{x}\in\mathcal{X}} \|\nabla f_t(\mathbf{x}) - \nabla f_{t-1}(\mathbf{x})\|_2^2. \tag{1}$$

Based on OMD (Nemirovski & Yudin, 1983), they prove $\mathcal{O}(\sqrt{V_T})$ regret and $\mathcal{O}(\frac{d}{\alpha}\log V_T)$ regret for convex functions and α-exp-concave functions respectively, under the smoothness condition. Zhang et al. (2022) further extend the result to λ-strongly convex and smooth functions, and obtain an $\mathcal{O}(\frac{1}{\lambda}\log V_T)$ bound. These bounds can be tighter than previous results when the loss functions change slowly such that the gradient variation V_T is small.

Later, Rakhlin & Sridharan (2013) introduce the paradigm of optimistic online learning that aims to take advantage of prior knowledge about the loss functions. In each round, the learner can obtain a prediction of the next loss, which is exploited to yield tighter bounds when the predictions are accurate and maintain the worst-case regret bound otherwise. To this end, they develop two frameworks: optimistic FTRL and optimistic OMD, where the latter generalizes the algorithm of Chiang et al. (2012).

2.2. Stochastic Online Convex Optimization

SCO assumes the loss functions to be i.i.d. and aims to minimize a convex objective in an expectation form, that is, $\min_{\mathbf{x}\in\mathcal{X}} F(\mathbf{x}) = \mathbb{E}_{f\sim\mathcal{D}}[f(\mathbf{x})]$, where $f(\cdot)$ is sampled from a fixed distribution $\mathcal{D}$. The performance is measured by the *excess risk* of the solution point over the optimum

$$F(\mathbf{x}_T) - \min_{\mathbf{x}\in\mathcal{X}} F(\mathbf{x}).$$

For Lipschitz and convex functions, stochastic gradient descent (SGD) achieves an $\mathcal{O}(1/\sqrt{T})$ risk bound. When the functions are equipped with additional properties, faster rates can be obtained. For smooth functions, SGD can reach an $\mathcal{O}(1/T + \sqrt{F_*/T})$ rate, where $F_* = \min_{\mathbf{x}\in\mathcal{X}} F(\mathbf{x})$, which will be tighter than $\mathcal{O}(1/\sqrt{T})$ when F_* is small (Srebro et al., 2010). For strongly convex functions, Hazan & Kale (2011) establish an $\mathcal{O}(1/[\lambda T])$ risk bound through a variant of SGD. When the functions are α-exp-concave, online Newton step (ONS) enjoys an $\mathcal{O}(d\log T/[\alpha T])$ rate (Hazan et al., 2007; Mahdavi et al., 2015). Furthermore, when strong convexity and smoothness hold at the same time, accelerated stochastic approximation (AC-SA) achieves an $\mathcal{O}(1/T)$ rate with a smaller constant (Ghadimi & Lan, 2012). Even faster results can be attained with strengthened conditions and advanced algorithms (Johnson & Zhang, 2013; Zhang et al., 2013; Zhang & Zhou, 2019).

2.3. Intermediate setting

In recent years, intermediate settings between adversarial OCO and SCO have drawn attention in prediction with expert advice (PEA) (Amir et al., 2020) and bandit problems (Zimmert & Seldin, 2021). Amir et al. (2020) study the stochastic regime with adversarial corruptions in PEA and obtain an $\mathcal{O}(\log N/\Delta + C_T)$ bound, where N is the number of experts, Δ is the suboptimality gap and $C_T \geq 0$ is the corruption level. Zimmert & Seldin (2021) focus on the adversarial regime with a self-bounding constraint and establish an $\mathcal{O}(N\log T/\Delta + \sqrt{C_T N \log T/\Delta})$ bound for bandit problems. Ito (2021) further demonstrates an expected regret bound of $\mathcal{O}(\log N/\Delta + \sqrt{C_T \log N/\Delta})$ on the above setting. However, as mentioned by Ito (2021), we know very little about the intermediate setting in OCO with the recent exception of Sachs et al. (2022).

3. Problem Setup and Existing Results

Problem Setup. Stochastically Extended Adversarial (SEA) model is introduced by Sachs et al. (2022) as an intermediate problem setup between adversarial OCO and SCO. In round $t \in [T]$, the learner selects a decision $\mathbf{x}_t$ from a convex feasible domain $\mathcal{X} \subseteq \mathbb{R}^d$, and the nature chooses a distribution $\mathcal{D}_t$. Then, the learner suffers a loss $f_t(\mathbf{x}_t)$, where the function f_t is sampled from the distribution $\mathcal{D}_t$. Due to the randomness in online process, our goal is to bound the *expected* regret against a comparator $\mathbf{x} \in \mathcal{X}$,

$$\mathbb{E}[\mathbf{Reg}_T] \triangleq \mathbb{E}\left[\sum_{t=1}^T f_t(\mathbf{x}_t) - \sum_{t=1}^T f_t(\mathbf{x})\right]. \tag{2}$$

Sachs et al. (2022) introduce the following quantities to capture the characteristics of SEA. For each $t \in [T]$, define the (conditional) expected function as

$$F_t(\mathbf{x}) = \mathbb{E}_{f_t\sim\mathcal{D}_t}[f_t(\mathbf{x})]$$

and the (conditional) variance of gradients as

$$\sigma_t^2 = \max_{\mathbf{x}\in\mathcal{X}} \mathbb{E}_{f_t\sim\mathcal{D}_t}\left[\|\nabla f_t(\mathbf{x}) - \nabla F_t(\mathbf{x})\|_2^2\right]. \tag{3}$$

Notice that both $F_t(\mathbf{x})$ and σ_t^2 can be random variables due to the randomness of distribution $\mathcal{D}_t$. The cumulative stochastic variance is defined as

$$\sigma_{1:T}^2 = \mathbb{E}\left[\sum_{t=1}^T \sigma_t^2\right], \tag{4}$$

which reflects the stochastic aspect of the online process. Moreover, the cumulative adversarial variation is defined as

$$\Sigma_{1:T}^2 = \mathbb{E}\left[\sum_{t=1}^T \sup_{\mathbf{x}\in\mathcal{X}} \|\nabla F_t(\mathbf{x}) - \nabla F_{t-1}(\mathbf{x})\|_2^2\right], \tag{5}$$

where $\nabla F_0(\mathbf{x}) = 0$, reflecting the adversarial difficulty.[1]

Besides, the following standard assumptions for online convex optimization are required (Hazan, 2016).

Assumption 1 (boundness of the gradient norm)**.** The gradients of all the random functions are bounded by G, i.e. for all $t \in [T]$, we have $\max_{\mathbf{x}\in\mathcal{X}} \|\nabla f_t(\mathbf{x})\|_2 \leq G$.

Assumption 2 (diameter of the domain)**.** The domain $\mathcal{X}$ contains the origin $\mathbf{0}$, and the diameter of $\mathcal{X}$ is bounded by D, i.e., for all $\mathbf{x}, \mathbf{y} \in \mathcal{X}$, we have $\|\mathbf{x} - \mathbf{y}\|_2 \leq D$.

Existing Results. With the condition of smoothness, Sachs et al. (2022) establish a series of results for the SEA model, including convex functions and strongly convex functions. In the case of convex and smooth functions, they prove an $\mathcal{O}(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})$ regret bound of optimistic FTRL. Note that they require the individual functions $\{f_t\}_{t=1}^T$ to be convex, which is relatively strict. When facing the adversarial setting, we have $\sigma_t^2 = 0$ for all t and $\Sigma_{1:T}^2$ is equivalent to V_T, so the bound implies $\mathcal{O}(\sqrt{V_T})$ regret matching the result of Chiang et al. (2012), which also reduces to the $\mathcal{O}(\sqrt{T})$ bound in the worst case (Zinkevich, 2003). In the SCO setting, we have $\Sigma_{1:T}^2 = 0$ and $\sigma_t = \sigma$ for all t,

[1] If the nature is oblivious, then both $F_t(\mathbf{x})$ and σ_t^2 are deterministic and we can remove the expectation in (4) and (5).

where σ denotes the variance of stochastic gradients. Then they obtain $\mathcal{O}(\sigma\sqrt{T})$ regret, leading to an $\mathcal{O}(\sigma/\sqrt{T})$ excess risk bound through the standard online-to-batch conversion (Cesa-Bianchi et al., 2004).

To investigate the strongly convex case, they make an additional assumption about the *maximum value* of stochastic variance and adversarial variation.

Assumption 3 (maximal stochastic variance and adversarial variation)**.** All the variance of the gradients are at most $\sigma_{\max}^2$, and all the adversarial variations are upper bounded by $\Sigma_{\max}^2$, that is, for all $t \in [T]$, it holds that $\sigma_t^2 \leq \sigma_{\max}^2$ and $\sup_{\mathbf{x}\in\mathcal{X}} \|\nabla F_t(\mathbf{x}) - \nabla F_{t-1}(\mathbf{x})\|_2^2 \leq \Sigma_{\max}^2$.

Based on Assumption 3, Sachs et al. (2022) prove an $\mathcal{O}((\sigma_{\max}^2 + \Sigma_{\max}^2)\log T)$ expected regret bound of optimistic FTRL for λ-strongly convex and smooth functions. Considering the adversarial setting, we have $\sigma_{\max}^2 = 0$ and $\Sigma_{\max}^2 \leq 4G^2$, so their bound implies an $\mathcal{O}(\log T)$ regret bound. We note that unlike in the convex and smooth case, their expected regret bound fails to recover the $\mathcal{O}(\log V_T)$ gradient-variation bound (Zhang et al., 2022). In the SCO setting, we have $\Sigma_{\max}^2 = 0$ and $\sigma_{\max}^2 = \sigma^2$. Therefore, their result brings an $\mathcal{O}([\sigma^2 \log T]/T)$ excess risk bound through online-to-batch conversion.

4. Our Results

We first introduce optimistic OMD, our main algorithmic framework, and then present its theoretical guarantees for SEA, as well as new results of optimistic FTRL. Finally, we investigate a new performance measure—dynamic regret, which is more suitable for non-stationary environments.

4.1. Algorithm

Optimistic OMD is a versatile and powerful framework for online learning (Rakhlin & Sridharan, 2013). During the learning process, it maintains two sequences $\{\mathbf{x}_t\}_{t=1}^T$ and $\{\widehat{\mathbf{x}}_t\}_{t=1}^T$. At round $t \in [T]$, the learner first submits the decision $\mathbf{x}_t$ and observes the random function $f_t(\cdot)$. Then, an optimistic vector $M_{t+1} \in \mathbb{R}^d$ is received that encodes certain prior knowledge of the (unknown) function $f_{t+1}(\cdot)$, and the algorithm updates the decision by

$$\widehat{\mathbf{x}}_{t+1} = \operatorname*{argmin}_{\mathbf{x}\in\mathcal{X}} \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}\rangle + B_{\psi_t}(\mathbf{x}, \widehat{\mathbf{x}}_t), \tag{6}$$

$$\mathbf{x}_{t+1} = \operatorname*{argmin}_{\mathbf{x}\in\mathcal{X}} \langle M_{t+1}, \mathbf{x}\rangle + B_{\psi_{t+1}}(\mathbf{x}, \widehat{\mathbf{x}}_{t+1}), \tag{7}$$

where $B_\psi(\mathbf{x},\mathbf{y}) = \psi(\mathbf{x}) - \psi(\mathbf{y}) - \langle \nabla\psi(\mathbf{y}), \mathbf{x}-\mathbf{y}\rangle$ denotes the Bregman divergence induced by a differentiable convex function $\psi : \mathcal{X} \mapsto \mathbb{R}$ (or usually called regularizer). We allow the regularizer to be time-varying, and the specific choice of $\psi_t(\cdot)$ depends on the type of online functions and will be determined later.

Algorithm 1 Optimistic OMD

1: Set $\mathbf{x}_1 = \widehat{\mathbf{x}}_1$ to be any point in $\mathcal{X}$
2: **for** $t = 1, \ldots, T$ **do**
3: Predict $\mathbf{x}_t$ and the nature selects a distribution $\mathcal{D}_t$
4: Receive $f_t(\cdot)$, which is sampled from $\mathcal{D}_t$
5: Update the two solutions according to (6) and (7)
6: **end for**

To exploit smoothness, we simply set the optimism as the last-round gradient, that is, $M_{t+1} = \nabla f_t(\mathbf{x}_t)$ (Chiang et al., 2012). We set $\mathbf{x}_1 = \widehat{\mathbf{x}}_1$ to be an arbitrary point in $\mathcal{X}$. The overall procedures are summarized in Algorithm 1.

Remark 1. If we drop the expectation operation, the measure (2) becomes the standard regret. Thus, a straightforward way is to plug in existing regret bounds of optimistic OMD (Chiang et al., 2012; Rakhlin & Sridharan, 2013), and then simplify the expectation. However, as elaborated by Sachs et al. (2022, Remark 4), this only yields very loose bounds. So, we need to dig into the analysis and examine the effect of expectations in the intermediate steps. $\triangleleft$

In the following, we consider three different instantiations of Algorithm 1 and present their theoretical guarantees.

4.2. Convex and Smooth Functions

In this part, we consider the case that the *expected function* is convex and smooth, as stated below.

Assumption 4 (smoothness of expected function)**.** For all $t \in [T]$, the expected function $F_t(\cdot)$ is L-smooth over $\mathcal{X}$, i.e., $\|\nabla F_t(\mathbf{x}) - \nabla F_t(\mathbf{y})\|_2 \leq L\|\mathbf{x}-\mathbf{y}\|_2,\ \forall \mathbf{x},\mathbf{y}\in\mathcal{X}$.

Assumption 5 (convexity of expected function)**.** For all $t \in [T]$, the expected function $F_t(\cdot)$ is convex over $\mathcal{X}$.

Remark 2. Sachs et al. (2022) require the random function $f_t(\cdot)$ to be convex (see A1 of their paper), whereas we only require the expected function $F_t(\cdot)$ to be convex, which is a much weaker condition. In fact, this relaxation is important and was studied in many works of stochastic optimization (Shalev-Shwartz, 2016; Hu et al., 2017; Ahn et al., 2020).

Indeed, we identify that the expectation in (2) allows us to avoid the convexity assumption of the individual function. Specifically, we have

$$\begin{aligned}&\mathbb{E}\big[f_t(\mathbf{x}_t) - f_t(\mathbf{x})\big] = \mathbb{E}\big[F_t(\mathbf{x}_t) - F_t(\mathbf{x})\big]\\ &\leq \mathbb{E}\big[\langle \nabla F_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}\rangle\big] = \mathbb{E}\big[\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}\rangle\big],\end{aligned} \tag{8}$$

where the inequality follows from the convexity of $F_t(\cdot)$, and the last step is because one can interchange differentiation and integration by Leibniz integral rule. As a result, we only need to bound the expected regret in terms of the linearized function, i.e., $\sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}\rangle$. $\triangleleft$

For convex and smooth functions, we set the Euclidean regularizer $\psi_t(\mathbf{x}) = \frac{1}{2\eta_t}\|\mathbf{x}\|_2^2$ with

$$\eta_t = \frac{D}{\sqrt{\delta + 4G^2 + \bar{V}_{t-1}}} \leq \frac{D}{\sqrt{\delta}}, \tag{9}$$

where $\bar{V}_{t-1} = \sum_{s=1}^{t-1} \|\nabla f_s(\mathbf{x}_s) - \nabla f_{s-1}(\mathbf{x}_{s-1})\|_2^2$, $\nabla f_0(\mathbf{x}_0) = 0$ and $\delta > 0$ is a parameter to be specified later. Then, the updates in (6) and (7) become

$$\widehat{\mathbf{x}}_{t+1} = \Pi_{\mathcal{X}}\left[\widehat{\mathbf{x}}_t - \eta_t \nabla f_t(\mathbf{x}_t)\right], \tag{10}$$

$$\mathbf{x}_{t+1} = \Pi_{\mathcal{X}}\left[\widehat{\mathbf{x}}_{t+1} - \eta_{t+1} \nabla f_t(\mathbf{x}_t)\right], \tag{11}$$

where $\Pi_{\mathcal{X}}[\cdot]$ denotes the Euclidean projection onto the feasible domain $\mathcal{X}$. The algorithm essentially performs gradient descent twice in each round, and the step size is determined adaptively, in a similar spirit with self-confident tuning (Auer et al., 2002). We do not need to apply the doubling trick used in previous works (Chiang et al., 2012; Rakhlin & Sridharan, 2013; Jadbabaie et al., 2015).

Below, we provide the theoretical guarantee of optimistic OMD for SEA with convex and smooth functions.

Theorem 1. *Under Assumptions 1, 2, 4 and 5, we have*

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^T f_t(\mathbf{x}_t) - \sum_{t=1}^T f_t(\mathbf{x})\right] \\
&\leq 5\sqrt{10}D^2 L + \frac{5\sqrt{5}DG}{2} + 5\sqrt{2}D\sqrt{\sigma_{1:T}^2} + 5D\sqrt{\Sigma_{1:T}^2} \\
&= \mathcal{O}\left(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2}\right),
\end{aligned}$$

where we set $\delta = 10D^2L^2$ in the step size (9).

Remark 3. Theorem 1 exhibits the same bound as Sachs et al. (2022), but under weaker assumptions since we only need the convexity of expected functions instead of individual functions. The regret bound is optimal according to the lower bound of Sachs et al. (2022, Theorem 6). ◁

We further improve the analysis of optimistic FTRL for SEA and demonstrate that even *without* convexity of individual functions, optimistic FTRL can achieve the *same* guarantee (Sachs et al., 2022) by feeding the algorithm with the linearized surrogate loss $\{\langle \nabla f_t(\mathbf{x}_t), \cdot\rangle\}_{t=1}^T$ instead of the original loss $\{f_t(\cdot)\}_{t=1}^T$.

Theorem 2. *Under Assumptions 1, 2, 4, and 5 (without assuming convexity of individual functions), with appropriate setup (see details in Appendix A.1), the expected regret of optimistic FTRL is at most $\mathcal{O}\big(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2}\big)$.*

4.3. Strongly Convex and Smooth Functions

In this part, we examine the case when *expected functions* are strongly convex and smooth.

Assumption 6 (strong convexity of expected function)**.** For $t \in [T]$, the expected function $F_t(\cdot)$ is λ-strongly convex over $\mathcal{X}$.

We still employ optimistic OMD (Algorithm 1) with regularizer $\psi_t(\mathbf{x}) = \frac{1}{2\eta_t}\|\mathbf{x}\|_2^2$ and set the step size as

$$\eta_t = \frac{2}{\lambda t} \leq \frac{2}{\lambda}. \tag{12}$$

It is worth mentioning that this setting of step size is *new* and much simpler than that in the earlier study of gradient-variation bound for strongly convex and smooth functions (Zhang et al., 2022), which uses a self-confident step size. To summarize, the updating rules take the same form as (10) and (11) with the step size in (12).

Then we pose the expected regret of optimistic OMD for SEA with strongly convex and smooth functions.

Theorem 3. *Under Assumptions 1, 2, 3, 4 and 6, we obtain*

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^T f_t(\mathbf{x}_t) - \sum_{t=1}^T f_t(\mathbf{x})\right] \\
&\leq \min\Bigg\{\frac{16G^2}{\lambda}\ln\left(8\sigma_{1:T}^2 + 4\Sigma_{1:T}^2 + G^2 + 1\right) \\
&\quad + \frac{16G^2}{\lambda}\ln\left(\frac{512G^2L^2}{\lambda^2} + 1\right) + \frac{16G^2+4}{\lambda} + \frac{\lambda D^2}{4}, \\
&\quad \frac{1}{\lambda}\left(32\sigma_{\max}^2 + 16\Sigma_{\max}^2\right)(\ln T + 1) + \frac{16L^2D^2 + 4G^2}{\lambda} \\
&\quad + \frac{16L^2D^2}{\lambda}\ln\left(1 + 8\sqrt{2}\frac{L}{\lambda}\right) + \frac{\lambda D^2}{4}\Bigg\} \\
&= \mathcal{O}\Big(\min\Big\{\frac{1}{\lambda}\log(\sigma_{1:T}^2 + \Sigma_{1:T}^2), \frac{1}{\lambda}(\sigma_{\max}^2 + \Sigma_{\max}^2)\log T\Big\}\Big).
\end{aligned}$$

Remark 4. Compare with the $\mathcal{O}(\frac{1}{\lambda}(\sigma_{\max}^2 + \Sigma_{\max}^2)\log T)$ bound of Sachs et al. (2022), our result demonstrates advantages on benign problems with small cumulative quantities $\sigma_{1:T}^2$ and $\Sigma_{1:T}^2$. When the cumulative quantities $\sigma_{1:T}^2$ and $\Sigma_{1:T}^2$ are small, the maximal quantities $\sigma_{\max}^2$ and $\Sigma_{\max}^2$ can still be large, making the bound of Sachs et al. (2022) loose. For instance, consider the adversarial setting where $\sigma_{1:T}^2 = \sigma_{\max}^2 = 0$, and online functions only change *once* such that $\Sigma_{1:T}^2 = \Sigma_{\max}^2 = \mathcal{O}(1)$. In this case, Theorem 3 gives an $\mathcal{O}(1)$ bound, while Sachs et al. (2022) only have an $\mathcal{O}(\log T)$ guarantee. Moreover, our additional bound $\mathcal{O}(\frac{1}{\lambda}\log(\sigma_{1:T}^2 + \Sigma_{1:T}^2))$ can reduce an $\mathcal{O}(\frac{1}{\lambda}\log V_T)$ gradient-variation bound in the adversarial OCO setting, whereas Sachs et al. (2022)'s bound cannot. ◁

Remark 5. Our new upper bound in Theorem 3 does not contradict with the $\Omega(\frac{1}{\lambda}(\sigma_{\max}^2 + \Sigma_{\max}^2)\log T)$ lower bound of Sachs et al. (2022, Theorem 8), because their lower bound focuses on the worst-case behavior while our result is better only in certain cases. ◁

Similar to Theorem 3, we demonstrate that in the strongly convex case, optimistic FTRL can also attain the *same* guarantee for SEA as our optimistic OMD does.

Theorem 4. *Under Assumptions 1, 2, 3, 4 and 6, with appropriate setup (see details in Appendix A.2), the expected regret of optimistic FTRL is at most* $\mathcal{O}(\min\{\frac{1}{\lambda}\log(\sigma_{1:T}^2+\Sigma_{1:T}^2), \frac{1}{\lambda}(\sigma_{\max}^2+\Sigma_{\max}^2)\log T\})$.

4.4. Exp-concave and Smooth Functions

We further investigate SEA with exp-concave and smooth functions. Note that Sachs et al. (2022) only study convex and strongly convex functions, without considering the exp-concave functions, so our result in this part is *new*.

Assumption 7 (exponential concavity of individual function)**.** For $t \in [T]$, the individual function $f_t(\cdot)$ is α-exp-concave over $\mathcal{X}$.

Remark 6. Note that Assumption 7 is about the random function $f_t(\cdot)$ rather than the expected function $F_t(\cdot)$. This is due to the need of using the exponential concavity of random functions in the regret analysis. We note that in the studies of stochastic exp-concave optimization, it is common to assume the random function to be exp-concave (Mahdavi et al., 2015; Koren & Levy, 2015) $\triangleleft$

Following Chiang et al. (2012), we set the regularizer $\psi_t(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|_{H_t}^2$, where the matrix H_t is defined as

$$H_t = I + \frac{\beta}{2}G^2 I + \frac{\beta}{2}\sum_{s=1}^{t-1}\nabla f_s(\mathbf{x}_s)\nabla f_s(\mathbf{x}_s)^\top, \tag{13}$$

where I is the d-dimensional identity matrix and $\beta = \frac{1}{2}\min\{\frac{1}{4GD}, \alpha\}$. Then, the updating rules of optimistic OMD in (6) and (7) become

$$\widehat{\mathbf{x}}_{t+1} = \operatorname*{argmin}_{\mathbf{x}\in\mathcal{X}} \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}\rangle + \frac{1}{2}\|\mathbf{x}-\widehat{\mathbf{x}}_t\|_{H_t}^2,$$
$$\mathbf{x}_{t+1} = \operatorname*{argmin}_{\mathbf{x}\in\mathcal{X}} \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}\rangle + \frac{1}{2}\|\mathbf{x}-\widehat{\mathbf{x}}_{t+1}\|_{H_{t+1}}^2.$$

For exp-concave and smooth functions, we can prove the following bound of optimistic OMD for the SEA model.

Theorem 5. *Under Assumptions 1, 2, 4 and 7, we obtain*

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^T f_t(\mathbf{x}_t) - \sum_{t=1}^T f_t(\mathbf{x})\right] \\
\le{}& \frac{16d}{\beta}\ln\left(\frac{\beta}{d}\sigma_{1:T}^2 + \frac{\beta}{2d}\Sigma_{1:T}^2 + \frac{\beta}{8d}G^2 + 1\right) \\
&+ \frac{16d}{\beta}\ln\left(32L^2+1\right) + D^2\left(1+\frac{\beta}{2}G^2\right) \\
={}& \mathcal{O}\Big(\frac{d}{\alpha}\log(\sigma_{1:T}^2+\Sigma_{1:T}^2)\Big),
\end{aligned}$$

where $\beta = \frac{1}{2}\min\{\frac{1}{4GD}, \alpha\}$, *and d is the dimensionality.*

Remark 7. This is the *first* regret bound of exp-concave and smooth functions under the SEA model. Due to the difference in the analysis, we are unable to achieve an $\mathcal{O}(\frac{d}{\alpha}(\sigma_{\max}^2+\Sigma_{\max}^2)\log T)$ regret bound like in the strongly convex case (Theorem 3). We will further investigate this possibility in the future. $\triangleleft$

Similarly, we obtain the same guarantee by optimistic FTRL in the exp-concave case.

Theorem 6. *Under Assumptions 1, 2, 4 and 7 with appropriate setup (see details in Appendix A.3), the expected regret of optimistic FTRL is at most* $\mathcal{O}(\frac{d}{\alpha}\log(\sigma_{1:T}^2+\Sigma_{1:T}^2))$.

4.5. Extension to Dynamic Regret Minimization

Sections 4.2–4.4 show that with appropriate step size and regularizer, optimistic OMD can achieve favorable guarantees for the SEA model.

Note that all those results minimize the measure (2), i.e., $\mathbb{E}[\sum_{t=1}^T f_t(\mathbf{x}_t) - \sum_{t=1}^T f_t(\mathbf{x})]$, where the comparator is fixed. Thus, it is usually called *static regret*. In this part, we further consider a more strengthened measure called *dynamic regret* (Zinkevich, 2003), defined as

$$\mathbf{Reg}_T^{\mathbf{d}} = \mathbb{E}\left[\sum_{t=1}^T f_t(\mathbf{x}_t) - \sum_{t=1}^T f_t(\mathbf{u}_t)\right].$$

where the comparators $\mathbf{u}_1, \ldots, \mathbf{u}_T \in \mathcal{X}$ are allowed to change over time. Therefore, this measure is more attractive in non-stationary online learning (Zhao et al., 2021). Notably, the static regret can be treated as its special case with a fixed comparator, i.e., $\mathbf{u}_1 = \ldots = \mathbf{u}_T = \mathbf{u}$, and thus it is much more general.

To optimize the dynamic regret, following the recent studies of non-stationary online learning (Zhang et al., 2018; Zhao et al., 2020), we develop a two-layer approach based on the optimistic OMD framework, which consists of a meta-learner running over a group of base-learners. The full procedure is summarized in Algorithm 2. Specifically, we maintain a pool for candidate step sizes $\mathcal{H} = \{\eta_i = c \cdot 2^i | i \in [N]\}$, where N is the number of base-learners of order $\mathcal{O}(\log T)$ and c is some small constant given later. We denote by $\mathcal{B}_i$ the i-th base-learner for $i \in [N]$. At round $t \in [T]$, the online learner obtains the decision $\mathbf{x}_t$ by aggregating local base decisions via the meta-learner, namely, $\mathbf{x}_t = \sum_{i=1}^N p_{t,i}\mathbf{x}_{t,i}$, where $\mathbf{x}_{t,i}$ is the decision returned by the base-learner $\mathcal{B}_i$ for $i \in [N]$ and $\boldsymbol{p}_t \in \Delta_N$ is the weight vector returned by the meta-algorithm. The nature then chooses a distribution $\mathcal{D}_t$ and the random function $f_t(\cdot)$ is sampled from $\mathcal{D}_t$. Subsequently, the online learner suffers the loss $f_t(\mathbf{x}_t)$ and observes the gradient $\nabla f_t(\mathbf{x}_t)$.

For the base-learner $\mathcal{B}_i$, in each round t, she obtains her local decision $\mathbf{x}_{t+1,i}$ by instantiating the optimistic OMD

Algorithm 2 Dynamic Regret Minimization of SEA Model

Input: step size pool $\mathcal{H} = \{\eta_1, \ldots, \eta_N\}$, learning rate of meta-algorithm $\varepsilon_t > 0$, correction coefficient $\lambda > 0$

1: Initialization: $\mathbf{x}_1 = \widehat{\mathbf{x}}_1 \in \mathcal{X}$, $\boldsymbol{p}_1 = \frac{1}{N} \cdot \mathbf{1}_N$
2: **for** $t = 1$ **to** T **do**
3: Receive $\mathbf{x}_{t,i}$ from base-learner $\mathcal{B}_i$ for $i \in [N]$
4: Submit the decision $\mathbf{x}_t = \sum_{i=1}^N p_{t,i}\mathbf{x}_{t,i}$
5: Observe the online function $f_t : \mathcal{X} \mapsto \mathbb{R}$ sampled from the underlying distribution $\mathcal{D}_t$ and suffer the loss $f_t(\mathbf{x}_t)$
6: Base-learner $\mathcal{B}_i$ updates the local decision by optimistic OMD, that is, $\widehat{\mathbf{x}}_{t+1,i} = \Pi_{\mathcal{X}}\left[\widehat{\mathbf{x}}_{t,i} - \eta_i \nabla f_t(\mathbf{x}_t)\right]$ and $\mathbf{x}_{t+1,i} = \Pi_{\mathcal{X}}\left[\widehat{\mathbf{x}}_{t+1,i} - \eta_i \nabla f_t(\mathbf{x}_t)\right], \forall i \in [N]$
7: Construct the feedback loss $\boldsymbol{\ell}_t \in \mathbb{R}^N$ with $\ell_{t,i} = \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_{t,i}\rangle + \lambda\|\mathbf{x}_{t,i} - \mathbf{x}_{t-1,i}\|_2$ for $i \in [N]$
8: Construct the optimism $\boldsymbol{m}_{t+1} \in \mathbb{R}^N$ with $m_{t+1,i} = \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_{t+1,i}\rangle + \lambda\|\mathbf{x}_{t+1,i} - \mathbf{x}_{t,i}\|_2$ for $i \in [N]$
9: Update the weight $\boldsymbol{p}_{t+1} \in \Delta_N$ by optimistic Hedge, that is, $p_{t+1,i} \propto \exp\left(-\varepsilon_t(\sum_{s=1}^t \ell_{s,i} + m_{t+1,i})\right)$
10: **end for**

algorithm (see Algorithm 1) with $\psi(\mathbf{x}) = \frac{1}{2\eta_i}\|\mathbf{x}\|_2^2$ and $M_{t+1} = \nabla f_t(\mathbf{x}_t)$ over the linearized surrogate loss $g_t(\mathbf{x}) = \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}\rangle$, where $\eta_i \in \mathcal{H}$ is the step size associated with the i-th base-learner. Since $\nabla g_t(\mathbf{x}_{t,i}) = \nabla f_t(\mathbf{x}_t)$, the updating rules of $\mathcal{B}_i$ are demonstrated as

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

The meta-learner updates the weight vector $\boldsymbol{p}_{t+1} \in \Delta_N$ by Optimistic Hedge (Syrgkanis et al., 2015) with a time-varying learning rate ε_t, that is,

$$p_{t+1,i} \propto \exp\left(-\varepsilon_t\Big(\sum_{s=1}^t \ell_{s,i} + m_{t+1,i}\Big)\right),$$

where the loss $\boldsymbol{\ell}_t \in \mathbb{R}^N$ is $\ell_{t,i} = \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_{t,i}\rangle + \lambda\|\mathbf{x}_{t,i} - \mathbf{x}_{t-1,i}\|_2^2$ for $t \geq 2$ and $\ell_{1,i} = \langle \nabla f_1(\mathbf{x}_1), \mathbf{x}_{1,i}\rangle$; the optimism $\boldsymbol{m}_{t+1} \in \mathbb{R}^N$ is constructed as $m_{t+1,i} = \langle M_{t+1}, \mathbf{x}_{t+1,i}\rangle + \lambda\|\mathbf{x}_{t+1,i} - \mathbf{x}_{t,i}\|_2^2$ with $M_{t+1} = \nabla f_t(\mathbf{x}_t)$ for $t \geq 2$ and $M_1 = \mathbf{0}$; $\lambda \geq 0$ being the coefficient of the correction terms; and we set $\mathbf{x}_{0,i} = \mathbf{0}$ for $i \in [N]$.

Remark 8. Our algorithm design and regret analysis follow the collaborative online ensemble framework proposed by Zhao et al. (2021), where the correction term $\lambda\|\mathbf{x}_{t,i} - \mathbf{x}_{t-1,i}\|_2^2$ in the meta-algorithm (both feedback loss and optimism) plays an important role. Technically, in this two-layer structure, to cancel the additional positive term $\sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2$ appearing in the derivation of $\sigma_{1:T}^2$ and $\Sigma_{1:T}^2$, one needs to ensure an effective collaboration between the meta and base layers. This involves simultaneously exploiting negative terms of the regret upper bounds in both the base and meta layers as well as leveraging additional negative terms introduced by the above correction term. ◁

Remark 9. After the submission of our paper, Sachs et al. (2022) released an updated version (Sachs et al., 2023), where they also utilized optimistic OMD to achieve the same dynamic regret as our approach. However, there is a significant difference between their method and ours. They employed an optimism design with $m_{t,i} = \langle \nabla f_{t-1}(\bar{\mathbf{x}}_t), \mathbf{x}_{t,i}\rangle$, based on the work of Zhao et al. (2020), where $\bar{\mathbf{x}}_t = \sum_{i=1}^N p_{t-1,i}\mathbf{x}_{t,i}$. But this design may introduce a dependence issue in SEA because $\bar{\mathbf{x}}_t$ depends on $f_{t-1}(\cdot)$. ◁

Below, we provide the dynamic regret upper bound.

Theorem 7. *Under Assumptions 1, 2, 4 and 5, setting the step size pool $\mathcal{H} = \{\eta_1, \ldots, \eta_N\}$ with $\eta_i = \min\{1/(8L), \sqrt{(D^2/(8G^2T)) \cdot 2^{i-1}}\}$ and $N = \lceil 2^{-1}\log_2(G^2T/(8L^2D^2))\rceil + 1$, and setting the learning rate of meta-algorithm as $\varepsilon_t = \min\{1/(8D^2L), \sqrt{(\ln N)/(D^2\bar{V}_t)}\}$ for all $t \in [T]$, Algorithm 2 ensures*

$$\mathbf{Reg}_T^{\mathbf{d}} \leq \mathcal{O}\Big(P_T + \sqrt{1+P_T}\big(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2}\big)\Big),$$

which holds for any comparator sequence $\mathbf{u}_1, \ldots, \mathbf{u}_T \in \mathcal{X}$. In above, $\bar{V}_t = \sum_{s=2}^t \|\nabla f_s(\mathbf{x}_s) - \nabla f_{s-1}(\mathbf{x}_{s-1})\|_2^2$ with $\nabla f_0(\mathbf{x}_0)$ defined as $\mathbf{0}$, and $P_T = \sum_{t=2}^T \|\mathbf{u}_t - \mathbf{u}_{t-1}\|_2$ denotes the path length of comparators.

Remark 10. As mentioned, the static regret studied in earlier sections is a special case of dynamic regret with a fixed comparator. As a consequence, Theorem 7 directly implies an $\mathcal{O}(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})$ static regret bound by noticing that $P_T = 0$ when comparing to a fixed benchmark, which recovers the result in Theorem 1. Moreover, Theorem 7 also recovers the $\mathcal{O}(\sqrt{(1+P_T+V_T)(1+P_T)})$ gradient-variation bound of Zhao et al. (2020) for the adversarial setting and the minimax optimal $\mathcal{O}(\sqrt{(1+P_T)T})$ bound of Zhang et al. (2018) since $\sigma_{1:T}^2 = 0$ and $\Sigma_{1:T}^2 = V_T \leq 4G^2T$ in this case. ◁

Remark 11. We focus on the convex and smooth case, while for the strongly convex and exp-concave cases, current understandings of their dynamic regret are still far from complete (Baby & Wang, 2021). In particular, how to realize optimistic online learning in strongly convex/exp-concave dynamic regret minimization remains open. ◁

Finally, we note that to the best of our knowledge, optimistic FTRL even has not been able to achieve the minimax optimal bound of Zhang et al. (2018). In fact, FTRL is more like a lazy update (Hazan, 2016), which seems unable to track a sequence of changing comparators. We found that Jacobsen & Cutkosky (2022) have given some preliminary results (in Theorem 2 and Theorem 3 of their work): *all the parameter-free FTRL-based algorithms we are aware of cannot achieve a dynamic regret bound better than $\mathcal{O}(P_T\sqrt{T})$.* Although this conclusion cannot cover all the cases of FTRL-based algorithms on dynamic regret, it has at least shown that FTRL-based algorithms do have some limitations in dynamic regret minimization.

5. Implications

First, we demonstrate how our results can be applied to recover the regret bound for adversarial data and the excess risk bound for stochastic data. Then, we discuss the implications for other intermediate examples.

We begin by listing two points followed in all the examples.

- For convex and smooth functions, we obtain the same $\mathcal{O}(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})$ bound as Sachs et al. (2022), so we will not repeat the analysis below unless necessary. We emphasize, however, that our result eliminates the assumption for convexity in individual functions, which is required in their work.
- For strongly convex and smooth functions, we give an $\mathcal{O}(\min\{\frac{1}{\lambda}\log(\sigma_{1:T}^2+\Sigma_{1:T}^2), \frac{1}{\lambda}(\sigma_{\max}^2+\Sigma_{\max}^2)\log T\})$ bound. The $\mathcal{O}(\frac{1}{\lambda}(\sigma_{\max}^2+\Sigma_{\max}^2)\log T)$ bound has been fully discussed by Sachs et al. (2022), so we will focus on the newly obtained $\mathcal{O}(\frac{1}{\lambda}\log(\sigma_{1:T}^2 + \Sigma_{1:T}^2))$ bound.

5.1. Fully Adversarial Data

For fully adversarial data, we have $\sigma_{1:T}^2 = 0$ as $\sigma_t^2 = 0$ for $t \in [T]$, and $\Sigma_{1:T}^2$ is equivalent to V_T. In this case, the new $\mathcal{O}(\frac{1}{\lambda}\log(\sigma_{1:T}^2 + \Sigma_{1:T}^2))$ bound in Theorem 3 guarantees an $\mathcal{O}(\frac{1}{\lambda}\log V_T)$ regret bound for λ-strongly convex and smooth functions, recovering the gradient-variation bound of Zhang et al. (2022). By contrast, the result of Sachs et al. (2022) can only recover the $\mathcal{O}(\frac{1}{\lambda}\log T)$ worst-case bound. Furthermore, for α-exp-concave functions, our new result (Theorem 5) implies an $\mathcal{O}(\frac{d}{\alpha}\log V_T)$ regret bound for OCO, recovering the result of Chiang et al. (2012).

5.2. Fully Stochastic Data

For fully stochastic data, the loss functions are i.i.d., so we have $\Sigma_{1:T}^2 = 0$ and $\sigma_t = \sigma, \forall t \in [T]$. Then for λ-strongly convex functions, Theorem 3 implies the same $\mathcal{O}(\log T/[\lambda T])$ excess risk bound as Sachs et al. (2022). Besides, Theorem 5 further delivers a new $\mathcal{O}(d\log T/[\alpha T])$ bound for α-exp-concave functions. These results match the well-known bounds in SCO (Hazan et al., 2007; Mahdavi et al., 2015) through online-to-batch conversion.

5.3. Adversarially Corrupted Stochastic Data

In the adversarially corrupted stochastic model, the loss function consists of two parts:

$$f_t(\cdot) = h_t(\cdot) + c_t(\cdot),$$

where $h_t(\cdot)$ is the loss of i.i.d. data sampled from a fixed distribution $\mathcal{D}$, and $c_t(\cdot)$ is a smooth adversarial perturbation satisfying that $\sum_{t=1}^T \max_{\mathbf{x}\in\mathcal{X}} \|\nabla c_t(\mathbf{x})\| \le C_T$, where $C_T \ge 0$ is a parameter called the *corruption level*.

Ito (2021) studies this model in expert and bandit problems, proposing a bound consisting of regret of i.i.d. data and an $\sqrt{C_T}$ term measuring the corrupted performance. Sachs et al. (2022) achieve a similar $\mathcal{O}(\sigma\sqrt{T} + \sqrt{C_T})$ bound in OCO problems under convexity and smoothness conditions, and raise an open question about how to extend the results to strongly convex losses. We resolve the problem by applying Theorem 3 of optimistic OMD to this model.

Corollary 1. *In the adversarially corrupted stochastic model, our $\mathcal{O}(\frac{1}{\lambda}\log(\sigma_{1:T}^2 + \Sigma_{1:T}^2))$ bound in Theorem 3 implies an $\mathcal{O}(\frac{1}{\lambda}\log(\sigma^2 T + C_T))$ bound for λ-strongly convex functions; and the result in Theorem 5 implies an $\mathcal{O}(\frac{d}{\alpha}\log(\sigma^2 T + C_T))$ bound for α-exp-concave functions.*

The proof of Corollary 1 is in Appendix B.1. We successfully extend results of Ito (2021) not only to strongly convex functions, but also to exp-concave functions.

5.4. Random Order Model

Random Order Model (ROM) (Garber et al., 2020; Sherman et al., 2021) relaxes the adversarial setting in standard adversarial OCO, in which the nature is allowed to choose the set of loss functions even with complete knowledge of the algorithm. However, the nature has no right to choose the order of loss functions, which will be arranged in uniformly random order instead.

Same as Sachs et al. (2022), let $\bar{\nabla}_T(\mathbf{x}) \triangleq \frac{1}{T}\sum_{s=1}^T \nabla f_s(\mathbf{x})$. Then we define $\sigma_1^2 = \max_{\mathbf{x}\in\mathcal{X}} \frac{1}{T}\sum_{t=1}^T \|\nabla f_t(\mathbf{x}) - \bar{\nabla}_T(\mathbf{x})\|_2^2$ and $\widetilde{\sigma}_1^2 = \frac{1}{T}\sum_{t=1}^T \max_{\mathbf{x}\in\mathcal{X}} \|\nabla f_t(\mathbf{x}) - \bar{\nabla}_T(\mathbf{x})\|_2^2$. Note that $\widetilde{\sigma}_1^2$ is a relaxation of σ_1^2 and the logarithm of $\widetilde{\sigma}_1^2/\sigma_1^2$ will not be large in reasonable scenarios. Sachs et al. (2022) establish an $\mathcal{O}(\sigma_1\sqrt{\log(\widetilde{\sigma}_1/\sigma_1)T})$ bound but require the convexity of individual functions, and they ask whether σ-dependent regret bounds can be realized under weaker assumptions on convexity of expected functions like Sherman et al. (2021). In Corollary 2, we give an affirmative answer based on Theorem 1 and obtain the results with weak assumptions. The proof is in Appendix B.2.

Corollary 2. *For convex expected functions, ROM enjoys an $\mathcal{O}(\sigma_1\sqrt{\log(\widetilde{\sigma}_1/\sigma_1)T})$ bound by Theorem 1.*

For λ-strongly convex expected functions, our new bound in Theorem 3 leads to an $\mathcal{O}(\frac{1}{\lambda}\log(T\sigma_1^2\log(\widetilde{\sigma}_1^2/\sigma_1^2)))$ bound, which is more stronger than the $\mathcal{O}(\frac{1}{\lambda}\sigma_1^2\log T)$ bound of Sachs et al. (2022) when σ_1^2 is not too small. Meanwhile, the best-of-both-worlds guarantee in Theorem 3 safeguards that our final bound is never worse than theirs. Besides, for α-exp-concave functions, we establish a new

$\mathcal{O}(\frac{d}{\alpha}\log(T\sigma_1^2\log(\widetilde{\sigma}_1^2/\sigma_1^2)))$ bound from Theorem 5, but the curvature assumption is imposed over individual functions. Thus an open question is whether a similar σ-dependent bound can be obtained under weaker assumptions.

5.5. Slow Distribution Shift

We consider a simple problem instance of online learning with slow distribution shifts, in which the underlying distributions selected by the nature in every two adjacent rounds are close on average. Formally, we suppose that $(1/T)\sum_{t=1}^{T}\sup_{\mathbf{x}\in\mathcal{X}}\|\nabla F_t(\mathbf{x})-\nabla F_{t-1}(\mathbf{x})\|_2^2 \le \varepsilon$, where ε is a constant. So we can get that $\Sigma_{1:T}^2 \le T\varepsilon$. For λ-strongly convex functions, our Theorem 3 realizes an $\mathcal{O}(\frac{1}{\lambda}\log(\sigma_{1:T}^2+\varepsilon T))$ regret bound, which is tighter than the $\mathcal{O}(\frac{1}{\lambda}(\sigma_{\max}^2\log T+\varepsilon T))$ bound of Sachs et al. (2022) for a large range of ε. We can further extend the analysis to the α-exp-concave functions and obtain an $\mathcal{O}(\frac{d}{\alpha}\log(\sigma_{1:T}^2+\varepsilon T))$ regret bound from Theorem 5.

5.6. Online Learning with Limited Resources

In real-world online learning applications, functions often arrive not individually but rather in groups. Let K_t denote the number of functions coming in round t and $f_t(\cdot,i)$ denote the i-th function. Denote by $F_t(\cdot)\triangleq\frac{1}{K_t}\sum_{i=1}^{K_t}f_t(\cdot,i)$ the average of all functions.

We consider the scenarios with limited computing resources such that gradient estimation can only be achieved by sampling a portion of the functions, leading to gradient variance. Assume that at each time t we sample $1\le B_t\le K_t$ functions, where the i-th function is expressed as $\widehat{f}_t(\cdot,i)$. We can then estimate $F_t(\cdot)$ by $f_t(\cdot)\triangleq\frac{1}{B_t}\sum_{i=1}^{B_t}\widehat{f}_t(\cdot,i)$, and further we have an upper bound for σ_t^2 as follows.

$$
\begin{aligned}
\sigma_t^2 &= \max_{\mathbf{x}\in\mathcal{X}}\mathbb{E}\left[\left\|\frac{1}{B_t}\sum_{i=1}^{B_t}\nabla\widehat{f}_t(\mathbf{x},i)-\nabla F_t(\mathbf{x})\right\|_2^2\right]\\
&=\frac{1}{B_t^2}\max_{\mathbf{x}\in\mathcal{X}}\Bigg(\sum_{i=1}^{B_t}\mathbb{E}\left[\left\|\nabla\widehat{f}_t(\mathbf{x},i)-\nabla F_t(\mathbf{x})\right\|_2^2\right]\\
&\quad+\mathbb{E}\Bigg[\sum_{i\neq j}\Big\langle\mathbb{E}\left[\nabla\widehat{f}_t(\mathbf{x},i)-\nabla F_t(\mathbf{x})\right],\\
&\quad\mathbb{E}\left[\nabla\widehat{f}_t(\mathbf{x},j)-\nabla F_t(\mathbf{x})\right]\Big\rangle\Bigg]\Bigg)\\
&=\frac{1}{B_t^2}\max_{\mathbf{x}\in\mathcal{X}}\left(\sum_{i=1}^{B_t}\mathbb{E}\left[\left\|\nabla\widehat{f}_t(\mathbf{x},i)-\nabla F_t(\mathbf{x})\right\|_2^2\right]\right)\\
&\le\frac{4G^2}{B_t},
\end{aligned}
$$

where we use the fact that $\nabla\widehat{f}_t(\mathbf{x},i)$ and $\nabla\widehat{f}_t(\mathbf{x},j)$ are independent when $i\neq j$, and the fact that $\mathbb{E}[\nabla\widehat{f}_t(\mathbf{x},i)-\nabla F_t(\mathbf{x})]=0$. The last inequality is due to Assumption 1. As a result, we have $\sigma_{1:T}^2=\mathbb{E}[\sum_{t=1}^{T}\sigma_t^2]\le 4G^2\sum_{t=1}^{T}\frac{1}{B_t}$ and obtain the following corollary by substituting it into Theorem 1, Theorem 3, and Theorem 5, respectively.

Corollary 3. *In online learning with limited resources, we can obtain an $\mathcal{O}(2G\sqrt{\sum_{t=1}^{T}\frac{1}{B_t}}+\sqrt{\Sigma_{1:T}^2})$ bound for convex functions by Theorem 1; and our $\mathcal{O}(\frac{1}{\lambda}\log(\sigma_{1:T}^2+\Sigma_{1:T}^2))$ bound in Theorem 3 implies an $\mathcal{O}(\frac{1}{\lambda}\log(4G^2\sum_{t=1}^{T}\frac{1}{B_t}+\Sigma_{1:T}^2))$ bound for λ-strongly convex functions; and Theorem 5 leads to an $\mathcal{O}\big(\frac{d}{\alpha}\log(4G^2\sum_{t=1}^{T}\frac{1}{B_t}+\Sigma_{1:T}^2)\big)$ bound for α-exp-concave functions.*

When the number of sampled functions increases, the estimated gradient will gradually approach the real gradient, and the gradient variance will be close to 0. It is noteworthy that the ratio B_t/K_t can be viewed as the *data throughput* determined by the available computing resources (Zhou, 2023). Corollary 3 demonstrates the impact of data throughput on the learning performance.

6. Conclusion and Future Work

In this paper, we investigate the Stochastically Extended Adversarial (SEA) model of Sachs et al. (2022) and propose a different solution via the optimistic OMD framework. Our results yield the *same* regret bound for convex and smooth functions under weaker assumptions and a *better* regret bound for strongly convex and smooth functions; moreover, we establish the *first* regret bound for exp-concave and smooth functions. For all three cases, we further improve analyses of optimistic FTRL, proving equal regret bounds with optimistic OMD for the SEA model. Furthermore, we study the SEA model under *dynamic regret* and propose a new two-layer algorithm based on optimistic OMD, which obtains the *first* dynamic regret guarantee for the SEA model. Lastly, we explore implications for intermediate learning scenarios, leading to various new results.

Although our algorithms for various functions can be unified using the optimistic OMD framework, they still necessitate distinct configurations for parameters such as step sizes and regularizers. Consequently, it becomes crucial to conceive and develop more adaptive online algorithms that eliminate the need for pre-set parameters. Exploring this area of research and designing such algorithms will be an important focus in future studies.

Acknowledgements

This work was supported by the National Key R&D Program of China (2022ZD0114801), NSFC (62122037, 61976112), and the major key project of PCL (PCL2021A12). The authors would thank Yu-Jie Zhang for helpful discussions.

References

Abernethy, J., Bartlett, P. L., Rakhlin, A., and Tewari, A. Optimal strategies and minimax lower bounds for online convex games. In *Proceedings of the 21st Annual Conference on Learning Theory*, pp. 415–423, 2008.

Ahn, K., Yun, C., and Sra, S. Sgd with shuffling: optimal rates without component convexity and large epoch requirements. *Advances in Neural Information Processing Systems 33 (NeurIPS)*, pp. 17526–17535, 2020.

Amir, I., Attias, I., Koren, T., Livni, R., and Mansour, Y. Prediction with corrupted expert advice. In *Advances in Neural Information Processing Systems 33 (NeurIPS)*, pp. 14315–14325, 2020.

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

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

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

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

Défossez, A., Bottou, L., Bach, F., and Usunier, N. A simple convergence proof of adam and adagrad. *Transactions on Machine Learning Research*, pp. to appear, 2022.

Duchi, J., Hazan, E., and Singer, Y. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12(7):2121–2159, 2011.

Garber, D., Korcia, G., and Levy, K. Online convex optimization in the random order model. In *Proceedings of the 37th International Conference on Machine Learning (ICML)*, pp. 3387–3396, 2020.

Ghadimi, S. and Lan, G. Optimal stochastic approximation algorithms for strongly convex stochastic composite optimization I: A generic algorithmic framework. *Siam Journal on Optimization*, 22(4):1469–1492, 2012.

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

Hazan, E. and Kale, S. Beyond the regret minimization barrier: an optimal algorithm for stochastic strongly-convex optimization. In *Proceedings of the 24th Annual Conference on Learning Theory (COLT)*, pp. 421–436, 2011.

Hazan, E., Agarwal, A., and Kale, S. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192, 2007.

Hu, B., Seiler, P., and Rantzer, A. A unified analysis of stochastic optimization methods using jump system theory and quadratic constraints. In *Proceedings of the 30th Conference on Learning Theory (COLT)*, pp. 1157–1189, 2017.

Ito, S. On optimal robustness to adversarial corruption in online decision problems. In *Advances in Neural Information Processing Systems 34 (NeurIPS)*, pp. 7409–7420, 2021.

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

Jadbabaie, A., Rakhlin, A., Shahrampour, S., and Sridharan, K. Online optimization: Competing with dynamic comparators. In *Proceedings of the 18th International Conference on Artificial Intelligence and Statistics*, pp. 398–406, 2015.

Johnson, R. and Zhang, T. Accelerating stochastic gradient descent using predictive variance reduction. In *Advances in Neural Information Processing Systems 26 (NIPS)*, pp. 315–323, 2013.

Kingma, D. P. and Ba, J. L. Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, 2015.

Koren, T. and Levy, K. Fast rates for exp-concave empirical risk minimization. In *Advances in Neural Information Processing Systems 28 (NIPS)*, pp. 1477–1485, 2015.

Lan, G. An optimal method for stochastic composite optimization. *Mathematical Programming*, 133(1):365–397, 2012.

Luo, H., Agarwal, A., Cesa-Bianchi, N., and Langford, J. Efficient second order online learning by sketching. In *Advances in Neural Information Processing Systems 29 (NIPS)*, pp. 902–910, 2016.

Mahdavi, M., Zhang, L., and Jin, R. Lower and upper bounds on the generalization of stochastic exponentially concave optimization. In *Proceedings of the 28th Annual Conference on Learning Theory (COLT)*, pp. 1305–1320, 2015.

Nemirovski, A. Prox-method with rate of convergence $O(1/t)$ for variational inequalities with lipschitz continuous monotone operators and smooth convex-concave saddle point problems. *SIAM Journal on Optimization*, 15(1):229–251, 2005.

Nemirovski, A. and Yudin, D. B. *Problem Complexity and Method Efficiency in Optimization*. John Wiley & Sons Ltd, 1983.

Nemirovski, A., Juditsky, A., Lan, G., and Shapiro, A. Robust stochastic approximation approach to stochastic programming. *SIAM Journal on Optimization*, 19(4):1574–1609, 2009.

Orabona, F. A modern introduction to online learning. *ArXiv preprint*, arXiv:1912.13213, 2019.

Orabona, F., Cesa-Bianchi, N., and Gentile, C. Beyond logarithmic bounds in online learning. In *Proceedings of the 15th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pp. 823–831, 2012.

Ordentlich, E. and Cover, T. M. The cost of achieving the best portfolio in hindsight. *Mathematics of Operations Research*, 23(4):960–982, 1998.

Rakhlin, A. and Sridharan, K. Online learning with predictable sequences. In *Proceedings of the 26th Conference on Learning Theory (COLT)*, pp. 993–1019, 2013.

Sachs, S., Hadiji, H., van Erven, T., and Guzmán, C. A. Between stochastic and adversarial online convex optimization: Improved regret bounds via smoothness. In *Advances in Neural Information Processing Systems 35 (NeurIPS)*, pp. 691–702, 2022.

Sachs, S., Hadiji, H., van Erven, T., and Guzman, C. Accelerated rates between stochastic and adversarial online convex optimization. *ArXiv preprint*, arXiv:2303.03272, 2023.

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

Shalev-Shwartz, S. SDCA without duality, regularization, and individual convexity. In *Proceedings of the 33th International Conference on Machine Learning (ICML)*, pp. 747–754, 2016.

Shalev-Shwartz, S., Shamir, O., Srebro, N., and Sridharan, K. Stochastic convex optimization. In *Proceedings of the 22nd Annual Conference on Learning Theory (COLT)*, pp. 5, 2009.

Shapiro, A., Dentcheva, D., and Ruszczyński, A. *Lectures on Stochastic Programming: Modeling and Theory*. SIAM, 2014.

Sherman, U., Koren, T., and Mansour., Y. Optimal rates for random order online optimization. In *Advances in Neural Information Processing Systems 34 (NeurIPS)*, pp. 2097–2108, 2021.

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

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

Yang, T., Mahdavi, M., Jin, R., and Zhu, S. Regret bounded by gradual variation for online convex optimization. *Machine Learning*, 95(2):183–223, 2014.

Zhang, L. and Zhou, Z.-H. Stochastic approximation of smooth and strongly convex functions: Beyond the $O(1/T)$ convergence rate. In *Proceedings of the 32nd Annual Conference on Learning Theory (COLT)*, pp. 3160–3179, 2019.

Zhang, L., Mahdavi, M., and Jin, R. Linear convergence with condition number independent access of full gradients. In *Advance in Neural Information Processing Systems 26 (NIPS)*, pp. 980–988, 2013.

Zhang, L., Lu, S., and Zhou, Z.-H. Adaptive online learning in dynamic environments. In *Advances in Neural Information Processing Systems 31 (NeurIPS)*, pp. 1323–1333, 2018.

Zhang, L., Wang, G., Yi, J., and Yang, T. A simple yet universal strategy for online convex optimization. In *Proceedings of the 39th International Conference on Machine Learning (ICML)*, pp. 26605–26623, 2022.

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

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

Zhou, Z.-H. Stream efficient learning. *ArXiv preprint*, arXiv:2305.02217, 2023.

Zimmert, J. and Seldin, Y. Tsallis-INF: An optimal algorithm for stochastic and adversarial bandits. *Journal of Machine Learning Research*, 22(28):1–49, 2021.

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

A. Omitted Details for Section 4

In this section, we present the omitted details for Section 4. We first provide the proofs of Theorem 1–Theorem 6 in Appendix A.1–Appendix A.3, and then supplement the omitted details of Section 4.5 for the dynamic regret minimization in Appendix A.4. We also collect some useful lemmas in Appendix A.5.

A.1. Convex and Smooth Functions

Proof of Theorem 1. Notice that our algorithm performs optimistic OMD on the random functions $\{f_1, \ldots, f_T\}$. So we can use Lemma 1 (variant of Bregman proximal inequality) by noting that the $\psi_t(\mathbf{x}) = \frac{1}{2\eta_t}\|\mathbf{x}\|_2^2$ is $\frac{1}{\eta_t}$-strongly convex with respect to $\|\cdot\|_2$ and obtain

$$\begin{aligned}\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}\rangle \le{}& \frac{1}{2\eta_t}\|\mathbf{x} - \widehat{\mathbf{x}}_t\|_2^2 - \frac{1}{2\eta_t}\|\mathbf{x} - \widehat{\mathbf{x}}_{t+1}\|_2^2 \\ &+ \eta_t\|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 - \frac{1}{2\eta_t}\Big(\|\mathbf{x}_t - \widehat{\mathbf{x}}_t\|_2^2 + \|\widehat{\mathbf{x}}_{t+1} - \mathbf{x}_t\|_2^2\Big).\end{aligned}$$

Summing the above inequality over $t = 1, \ldots, T$, we have

$$\begin{aligned}&\sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}\rangle \\ \le{}& \underbrace{\sum_{t=1}^T \left(\frac{1}{2\eta_t}\|\mathbf{x} - \widehat{\mathbf{x}}_t\|_2^2 - \frac{1}{2\eta_t}\|\mathbf{x} - \widehat{\mathbf{x}}_{t+1}\|_2^2\right)}_{\texttt{term (a)}} + \underbrace{\sum_{t=1}^T \eta_t\|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2}_{\texttt{term (b)}} \\ &\underbrace{- \sum_{t=1}^T \frac{1}{2\eta_t}\left(\|\mathbf{x}_t - \widehat{\mathbf{x}}_t\|_2^2 + \|\widehat{\mathbf{x}}_{t+1} - \mathbf{x}_t\|_2^2\right)}_{\texttt{term (c)}}. \end{aligned} \tag{14}$$

In the following, we will bound the three terms on the right hand respectively. Before that, according to that $\eta_t = D/\sqrt{\delta + 4G^2 + \bar{V}_{t-1}}$ where $\bar{V}_{t-1} = \sum_{s=1}^{t-1}\|\nabla f_s(\mathbf{x}_s) - \nabla f_{s-1}(\mathbf{x}_{s-1})\|_2^2$, we derive that $\eta_t \le D/\sqrt{\delta + \bar{V}_t}$ by Assumption 1 (boundness of the gradient norm).

To bound term (a), we notice that $\eta_t \le \eta_{t-1}$. Based on Assumption 2 (domain boundedness),

$$\begin{aligned}\texttt{term (a)} ={}& \frac{1}{2\eta_1}\|\mathbf{x} - \widehat{\mathbf{x}}_1\|_2^2 + \frac{1}{2}\sum_{t=2}^T \left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}}\right)\|\mathbf{x} - \widehat{\mathbf{x}}_t\|_2^2 - \frac{1}{2\eta_T}\|\mathbf{x} - \widehat{\mathbf{x}}_{T+1}\|_2^2 \\ \le{}& \frac{1}{2\eta_1}D^2 + \frac{1}{2}\sum_{t=2}^T \left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}}\right) D^2 = \frac{D^2}{2\eta_T} = \frac{D}{2}\sqrt{\delta + 4G^2 + \bar{V}_{T-1}}.\end{aligned}$$

By applying Lemma 8 (self-confident tuning), we can bound term (b) as

$$\texttt{term (b)} \le \sum_{t=1}^T \frac{D}{\sqrt{\delta + \bar{V}_t}}\|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \le 2D\sqrt{\delta + \bar{V}_T}.$$

To bound term (c), we use the fact that $\eta_t \le \frac{D}{\sqrt{\delta}}$:

$$\begin{aligned}\texttt{term (c)} ={}& \sum_{t=1}^T \frac{1}{2\eta_t}\left(\|\mathbf{x}_t - \widehat{\mathbf{x}}_t\|_2^2 + \|\widehat{\mathbf{x}}_{t+1} - \mathbf{x}_t\|_2^2\right) \ge \frac{\sqrt{\delta}}{2D}\sum_{t=1}^T \left(\|\mathbf{x}_t - \widehat{\mathbf{x}}_t\|_2^2 + \|\widehat{\mathbf{x}}_{t+1} - \mathbf{x}_t\|_2^2\right) \\ \ge{}& \frac{\sqrt{\delta}}{2D}\sum_{t=2}^T \left(\|\mathbf{x}_t - \widehat{\mathbf{x}}_t\|_2^2 + \|\widehat{\mathbf{x}}_t - \mathbf{x}_{t-1}\|_2^2\right) \ge \frac{\sqrt{\delta}}{4D}\sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2.\end{aligned}$$

Then we substitute the three bounds above into (14)

$$
\begin{aligned}
\sum_{t=1}^{T}\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}\rangle &\le \frac{D}{2}\sqrt{\delta + 4G^2 + \bar{V}_{T-1}} + 2D\sqrt{\delta + \bar{V}_T} - \frac{\sqrt{\delta}}{4D}\sum_{t=2}^{T}\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \\
&\le \frac{5D}{2}\sqrt{\delta + 4G^2 + \bar{V}_{T-1}} - \frac{\sqrt{\delta}}{4D}\sum_{t=2}^{T}\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2.
\end{aligned}
$$

Applying Lemma 5 (boundness of cumulative norm of gradient difference), we have

$$
\begin{aligned}
&\sum_{t=1}^{T}\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}\rangle \\
&\le \frac{5D}{2}\sqrt{\delta + 5G^2} + 5\sqrt{2}D\sqrt{\sum_{t=1}^{T}\|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2} + 5DL\sqrt{\sum_{t=2}^{T}\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2} \\
&\quad + 5D\sqrt{\sum_{t=2}^{T}\|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2} - \frac{\sqrt{\delta}}{4D}\sum_{t=2}^{T}\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \\
&\le \frac{5D}{2}\sqrt{\delta + 5G^2} + \frac{25D^3L^2}{\sqrt{\delta}} + 5\sqrt{2}D\sqrt{\sum_{t=1}^{T}\|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2} \\
&\quad + 5D\sqrt{\sum_{t=2}^{T}\|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2},
\end{aligned} \tag{15}
$$

where in the last step we make use of the AM-GM inequality

$$
5DL\sqrt{\sum_{t=2}^{T}\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2} \le \frac{25D^3L^2}{\sqrt{\delta}} + \frac{\sqrt{\delta}}{4D}\sum_{t=2}^{T}\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2.
$$

Taking expectations over (15) and applying Jensen's inequality lead to

$$
\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^{T}\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}\rangle\right] &\le \frac{5D}{2}\sqrt{\delta} + \frac{25D^3L^2}{\sqrt{\delta}} + \frac{5\sqrt{5}DG}{2} + 5\sqrt{2}D\sqrt{\sigma_{1:T}^2} + 5D\sqrt{\Sigma_{1:T}^2} \\
&= 5\sqrt{10}D^2L + \frac{5\sqrt{5}DG}{2} + 5\sqrt{2}D\sqrt{\sigma_{1:T}^2} + 5D\sqrt{\Sigma_{1:T}^2} \\
&= \mathcal{O}(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2}),
\end{aligned}
$$

where we set $\delta = 10D^2L^2$ and recall the definitions of $\sigma_{1:T}^2$ and $\Sigma_{1:T}^2$ in (4) and (5), as restated below:

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

Then we complete the proof with the fact that the above expectation upper bounds the expected regret in (8). □

Proof of Theorem 2. We introduce the procedure of optimistic FTRL for convex and smooth functions first. At each step t, we define a new surrogate loss: $\ell_t(\mathbf{x}) = \langle \nabla f_t(\mathbf{x}_t), \mathbf{x} - \mathbf{x}_t\rangle$. Note that we do not directly use the original function $f_t(\cdot)$ to update the decision point in the convex case as in Sachs et al. (2022), this is to remove the requirement on the convexity of individual functions (which is required by Sachs et al. (2022)). The decision $\mathbf{x}_t$ is updated by deploying optimistic FTRL over the linearized loss:

$$
\mathbf{x}_t = \operatorname*{argmin}_{\mathbf{x}\in\mathcal{X}}\sum_{s=1}^{t-1}\ell_s(\mathbf{x}) + \langle M_t, \mathbf{x}\rangle + \frac{1}{\eta_t}\|\mathbf{x}\|_2^2,
$$

where $\mathbf{x}_0$ is an arbitrary point in $\mathcal{X}$, and the optimistic vector $M_t = \nabla f_{t-1}(\mathbf{x}_{t-1})$ (we set $M_1 = \nabla f_0(\mathbf{x}_0) = 0$). And the step size η_t is designed as $\eta_t = D^2/(\delta + \sum_{s=1}^{t-1} \eta_s \|\nabla f_s(\mathbf{x}_s) - f_{s-1}(\mathbf{x}_{s-1})\|_2^2)$ with δ to be defined latter, which is non-increasing for $t \in [T]$.

Obviously, we can easily obtain that

$$\mathbb{E}\left[\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{x})\right] \le \mathbb{E}\left[\sum_{t=1}^{T} \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}\rangle\right] \le \mathbb{E}\left[\sum_{t=1}^{T} \ell_t(\mathbf{x}_t) - \sum_{t=1}^{T} \ell_t(\mathbf{x})\right].$$

As a result, we only need to consider the regret of the surrogate loss $\ell_t(\cdot)$. The following proof is similar with Sachs et al. (2022). To exploit Lemma 11 (standard analysis of optimistic FTRL), we map the G_t term in Lemma 11 to $\frac{1}{\eta_t}\|\mathbf{x}\|_2^2 + \sum_{s=1}^{t-1} \ell_s(\mathbf{x})$ and map the $\widetilde{\mathbf{g}}_t$ term to M_t. Note that G_t is $\frac{2}{\eta_t}$-strongly convex and ℓ_t is convex, we have

$$\begin{aligned}
\sum_{t=1}^{T} \ell_t(\mathbf{x}_t) - \sum_{t=1}^{T} \ell_t(\mathbf{x}) &\le \frac{D^2}{\eta_T} + \sum_{t=1}^{T}\left(\langle \nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1}), \mathbf{x}_t - \mathbf{x}_{t+1}\rangle - \frac{1}{\eta_t}\|\mathbf{x}_t - \mathbf{x}_{t+1}\|_2^2\right) \\
&= \frac{D^2}{\eta_T} + \sum_{t=1}^{T}\left(\frac{\eta_t}{2}\|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 - \frac{1}{2\eta_t}\|\mathbf{x}_t - \mathbf{x}_{t+1}\|_2^2\right) \\
&\le \delta + \frac{3}{2}\sum_{t=1}^{T} \eta_t \|\nabla f_t(\mathbf{x}_t) - f_{t-1}(\mathbf{x}_{t-1})\|_2^2 - \frac{\delta}{2D^2}\sum_{t=1}^{T}\|\mathbf{x}_t - \mathbf{x}_{t+1}\|_2^2, \qquad (16)
\end{aligned}$$

where we use the fact that $\langle a, b\rangle \le \|a\|_* \|b\| \le \frac{1}{2c}\|a\|_*^2 + \frac{c}{2}\|b\|^2$ in the second inequality ($\|\cdot\|_*$ denotes the dual norm of $\|\cdot\|$), based on the Hölder's inequality.

To bound the second term above, we directly use the following inequality from Sachs et al. (2022, proof of Theorem 5)

$$\sum_{t=1}^{T} \eta_t \|\nabla f_t(\mathbf{x}_t) - f_{t-1}(\mathbf{x}_{t-1})\|^2 \le D\sqrt{2\bar{V}_T} + \frac{4D^2G^2}{\delta}.$$

In this way, we substitute the above bound into the origin regret (16) and obtain

$$\sum_{t=1}^{T} \ell_t(\mathbf{x}_t) - \sum_{t=1}^{T} \ell_t(\mathbf{x}) \le \frac{3\sqrt{2}}{2} D\sqrt{\bar{V}_T} + \frac{6D^2G^2}{\delta} + \delta - \frac{\delta}{2D^2}\sum_{t=1}^{T}\|\mathbf{x}_t - \mathbf{x}_{t+1}\|_2^2.$$

By applying Lemma 5 (boundness of cumulative norm of the gradient difference), we have

$$\begin{aligned}
&\sum_{t=1}^{T} \ell_t(\mathbf{x}_t) - \sum_{t=1}^{T} \ell_t(\mathbf{x}) \\
&\le 6D\sqrt{\sum_{t=1}^{T}\|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2} + 3\sqrt{2}D\sqrt{\sum_{t=2}^{T}\|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2} \\
&\quad + 3\sqrt{2}DL\sqrt{\sum_{t=2}^{T}\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2} - \frac{\delta}{2D^2}\sum_{t=1}^{T}\|\mathbf{x}_t - \mathbf{x}_{t+1}\|_2^2 + \frac{6D^2G^2}{\delta} + \delta + \frac{3\sqrt{2}}{2}DG \\
&\le 6D\sqrt{\sum_{t=1}^{T}\|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2} + 3\sqrt{2}D\sqrt{\sum_{t=2}^{T}\|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2} \\
&\quad + \frac{9D^4L^2}{\delta} + \frac{6D^2G^2}{\delta} + \delta + \frac{3\sqrt{2}}{2}DG, \qquad (17)
\end{aligned}$$

where we use the following inequality in the last step:

$$3\sqrt{2}DL\sqrt{\sum_{t=2}^{T}\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2} \le \frac{9D^4L^2}{\delta} + \frac{\delta}{2D^2}\sum_{t=1}^{T}\|\mathbf{x}_t - \mathbf{x}_{t+1}\|_2^2,$$

which cancels out the negative term in (17) with the second term.

Then, we take expectations over (17) with the help of definitions of $\sigma^2_{1:T}$ and $\Sigma^2_{1:T}$, and use Jensen's inequality. Given that the expected regret of surrogate loss functions upper bounds the expected regret of original functions, we get the final result and complete the proof:

$$\begin{aligned}\mathbb{E}\left[\sum_{t=1}^T f_t(\mathbf{x}_t) - \sum_{t=1}^T f_t(\mathbf{x})\right] &\le \mathbb{E}\left[\sum_{t=1}^T \ell_t(\mathbf{x}_t) - \sum_{t=1}^T \ell_t(\mathbf{x})\right] \\ &\le 6D\sqrt{\sigma^2_{1:T}} + 3\sqrt{2}D\sqrt{\Sigma^2_{1:T}} + \frac{9D^4L + 6D^2G^2}{\delta} + \delta + \frac{3\sqrt{2}}{2}DG \\ &= 6D\sqrt{\sigma^2_{1:T}} + 3\sqrt{2}D\sqrt{\Sigma^2_{1:T}} + 2\sqrt{9D^4L + 6D^2G^2} + \frac{3\sqrt{2}}{2}DG \\ &= \mathcal{O}(\sqrt{\sigma^2_{1:T}} + \sqrt{\Sigma^2_{1:T}}),\end{aligned}$$

where we set $\delta = \sqrt{9D^4L + 6D^2G^2}$. $\square$

A.2. Strongly Convex and Smooth Functions

Proof of Theorem 3. For the case where we operate optimistic OMD on λ-strongly convex expected functions (see Assumption 6), we have $F_t(\mathbf{x}_t) - F_t(\mathbf{x}) \le \langle \nabla F_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}\rangle - \frac{\lambda}{2}\|\mathbf{x} - \mathbf{x}_t\|_2^2$. Then in view of the definition that $F_t(\mathbf{x}) = \mathbb{E}_{f_t\sim\mathcal{D}_t}[f_t(\mathbf{x})]$, we bound the expected regret as

$$\begin{aligned}\mathbb{E}\left[\sum_{t=1}^T f_t(\mathbf{x}_t) - \sum_{t=1}^T f_t(\mathbf{x})\right] &= \mathbb{E}\left[\sum_{t=1}^T F_t(\mathbf{x}_t) - \sum_{t=1}^T F_t(\mathbf{x})\right] \\ \le \mathbb{E}\left[\sum_{t=1}^T \left(\langle \nabla F_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}\rangle - \frac{\lambda}{2}\|\mathbf{x} - \mathbf{x}_t\|_2^2\right)\right] &= \mathbb{E}\left[\sum_{t=1}^T \left(\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}\rangle - \frac{\lambda}{2}\|\mathbf{x} - \mathbf{x}_t\|_2^2\right)\right].\end{aligned} \tag{18}$$

Since we use the same $\psi_t(\mathbf{x})$ as Theorem 1, we can reuse the regret approximation in (14):

$$\begin{aligned}&\sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}\rangle - \frac{\lambda}{2}\sum_{t=1}^T \|\mathbf{x} - \mathbf{x}_t\|_2^2 \\ \le\ &\underbrace{\sum_{t=1}^T \left(\frac{1}{2\eta_t}\|\mathbf{x} - \widehat{\mathbf{x}}_t\|_2^2 - \frac{1}{2\eta_t}\|\mathbf{x} - \widehat{\mathbf{x}}_{t+1}\|_2^2\right) - \frac{\lambda}{2}\sum_{t=1}^T \|\mathbf{x} - \mathbf{x}_t\|_2^2}_{\texttt{term (a)}} \\ &+ \underbrace{\sum_{t=1}^T \eta_t\|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2}_{\texttt{term (b)}} - \underbrace{\sum_{t=1}^T \frac{1}{2\eta_t}\left(\|\mathbf{x}_t - \widehat{\mathbf{x}}_t\|_2^2 + \|\widehat{\mathbf{x}}_{t+1} - \mathbf{x}_t\|_2^2\right)}_{\texttt{term (c)}}.\end{aligned} \tag{19}$$

Because the regret bound in Theorem 3 consists of two parts, we first prove the $\mathcal{O}\big(\frac{1}{\lambda}\log(\sigma^2_{1:T} + \Sigma^2_{1:T})\big)$ bound and then prove the $\mathcal{O}\big(\frac{1}{\lambda}(\sigma^2_{\max} + \Sigma^2_{\max})\log T\big)$ bound.

The $\mathcal{O}\big(\frac{1}{\lambda}\log(\sigma^2_{1:T} + \Sigma^2_{1:T})\big)$ bound. First, we upper bound term (a), term (b), and term (c) respectively.

From Lemma 3 (stability lemma) we get that $\|\widehat{\mathbf{x}}_{t+1} - \mathbf{x}_t\|_2 \le \eta_t\|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2$. Then based on Assumption 2 (diameter of the domain) and the step size $\eta = \frac{2}{\lambda t}$ (see (12)), we can bound term (a) as

$$\begin{aligned}\texttt{term (a)} &\le \frac{1}{2\eta_1}D^2 + \frac{1}{2}\sum_{t=2}^T\left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}}\right)\|\mathbf{x} - \widehat{\mathbf{x}}_t\|_2^2 - \frac{\lambda}{2}\sum_{t=1}^T \|\mathbf{x} - \mathbf{x}_t\|_2^2 \\ &\le \frac{\lambda D^2}{4} + \frac{\lambda}{4}\sum_{t=1}^{T-1}\left(\|\mathbf{x} - \widehat{\mathbf{x}}_{t+1}\|_2^2 - 2\|\mathbf{x} - \mathbf{x}_t\|_2^2\right) \le \frac{\lambda D^2}{4} + \frac{\lambda}{2}\sum_{t=1}^{T-1}\|\widehat{\mathbf{x}}_{t+1} - \mathbf{x}_t\|_2^2\end{aligned}$$

$$\leq \frac{\lambda D^2}{4} + \frac{\lambda \eta_1}{2} \sum_{t=1}^{T-1} \eta_t \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \leq \frac{\lambda D^2}{4} + \texttt{term (b)},$$

where in the penultimate step, we use the fact that η_t is non-increasing. From the above derivation, we observe that the upper bound of term (a) depends on term (b). To bound term (b), we apply Lemma 6 to term (b):

$$\texttt{term (b)} = 2\sum_{t=1}^{T} \frac{1}{\lambda t} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \leq \frac{8G^2}{\lambda} \ln\left(\bar{V}_T + 1\right) + \frac{8G^2 + 2}{\lambda}.$$

Similar to the proof of Theorem 1, we substitute $\eta_t = \frac{2}{\lambda t}$ into term (c) and bound it as

$$\texttt{term (c)} = \sum_{t=1}^{T} \frac{1}{2\eta_t} \left(\|\mathbf{x}_t - \widehat{\mathbf{x}}_t\|_2^2 + \|\widehat{\mathbf{x}}_{t+1} - \mathbf{x}_t\|_2^2\right) \geq \frac{\lambda}{8} \sum_{t=2}^{T} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2.$$

Replacing term (a), term (b), and term (c) with their corresponding upper bounds in (19), we have

$$\begin{aligned}
&\sum_{t=1}^{T} \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x} \rangle - \frac{\lambda}{2} \sum_{t=1}^{T} \|\mathbf{x} - \mathbf{x}_t\|_2^2 \\
\leq {} & \frac{\lambda D^2}{4} + \frac{16G^2}{\lambda} \ln\left(\bar{V}_T + 1\right) + \frac{16G^2 + 4}{\lambda} - \frac{\lambda}{8} \sum_{t=2}^{T} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2.
\end{aligned}$$

Then, applying Lemma 5 (boundness of cumulative norm of gradient difference), we can upper bound the $\ln\left(\bar{V}_T + 1\right)$ term and get that

$$\begin{aligned}
&\sum_{t=1}^{T} \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x} \rangle - \frac{\lambda}{2} \sum_{t=1}^{T} \|\mathbf{x} - \mathbf{x}_t\|_2^2 \\
\leq {} & \frac{16G^2}{\lambda} \ln\left(8\sum_{t=1}^{T} \|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 + 4\sum_{t=2}^{T} \|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 + G^2 + 1\right) \\
& + \frac{16G^2}{\lambda} \ln\left(4L^2 \sum_{t=2}^{T} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + 1\right) - \frac{\lambda}{8} \sum_{t=2}^{T} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + \frac{16G^2 + 4}{\lambda} + \frac{\lambda D^2}{4},
\end{aligned} \tag{20}$$

where we use the inequality below

$$\ln(1 + u + v) \leq \ln(1 + u) + \ln(1 + v), \ \forall u, v \geq 0. \tag{21}$$

To simplify the last line of the above bound, we use Lemma 7 and obtain

$$\frac{16G^2}{\lambda} \ln\left(4L^2 \sum_{t=2}^{T} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + 1\right) - \frac{\lambda}{8} \sum_{t=2}^{T} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \leq \frac{16G^2}{\lambda} \ln\left(\frac{512G^2L^2}{\lambda^2} + 1\right).$$

Under this simplification, we can reduce the regret bound to the following form:

$$\begin{aligned}
&\sum_{t=1}^{T} \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x} \rangle - \frac{\lambda}{2} \sum_{t=1}^{T} \|\mathbf{x} - \mathbf{x}_t\|_2^2 \\
\leq {} & \frac{16G^2}{\lambda} \ln\left(8\sum_{t=1}^{T} \|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 + 4\sum_{t=2}^{T} \|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 + G^2 + 1\right) \\
& + \frac{16G^2}{\lambda} \ln\left(\frac{512G^2L^2}{\lambda^2} + 1\right) + \frac{16G^2 + 4}{\lambda} + \frac{\lambda D^2}{4}.
\end{aligned} \tag{22}$$

To get the bound of expected regret, taking expectations over the above regret bound, and applying Jensen's inequality, we have

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T}\langle\nabla f_t(\mathbf{x}_t),\mathbf{x}_t-\mathbf{x}\rangle-\frac{\lambda}{2}\sum_{t=1}^{T}\|\mathbf{x}-\mathbf{x}_t\|_2^2\right]\\
\leq{}& \frac{16G^2}{\lambda}\ln\left(8\sigma_{1:T}^2+4\Sigma_{1:T}^2+G^2+1\right)+\frac{16G^2}{\lambda}\ln\left(\frac{512G^2L^2}{\lambda^2}+1\right)+\frac{16G^2+4}{\lambda}+\frac{\lambda D^2}{4}. \qquad (23)
\end{aligned}$$

Thus we have proven the $\mathcal{O}\big(\frac{1}{\lambda}\log(\sigma_{1:T}^2+\Sigma_{1:T}^2)\big)$ bound.

The $\mathcal{O}\big(\frac{1}{\lambda}(\sigma_{\max}^2+\Sigma_{\max}^2)\log T\big)$ bound. With the techniques in Sachs et al. (2022), we can get another bound. Indeed, their analysis is for optimistic FTRL, while ours is for optimistic OMD, but we show that our algorithm can also enjoy that kind of guarantee.

For term (a) and term (b), we have already derived that

$$\texttt{term (b)}=2\sum_{t=1}^{T}\frac{1}{\lambda t}\|\nabla f_t(\mathbf{x}_t)-\nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \quad\text{and}\quad \texttt{term (a)}\leq\frac{\lambda D^2}{4}+\texttt{term (b)}.$$

For term (c), we give a new bound, which follows Sachs et al. (2022):

$$\texttt{term (c)}\geq\sum_{t=2}^{T}\left(\frac{1}{2\eta_t}\|\mathbf{x}_t-\widehat{\mathbf{x}}_t\|_2^2+\frac{1}{2\eta_{t-1}}\|\widehat{\mathbf{x}}_t-\mathbf{x}_{t-1}\|_2^2\right)\geq\sum_{t=2}^{T}\frac{1}{4\eta_{t-1}}\|\mathbf{x}_t-\mathbf{x}_{t-1}\|_2^2,$$

where we make use of the fact that η_t is non-increasing.

Integrating the new upper bounds of term (a), term (b) and term (c) into (19) together with $\eta_t=\frac{2}{\lambda t}$, we have

$$\begin{aligned}
&\sum_{t=1}^{T}\langle\nabla f_t(\mathbf{x}_t),\mathbf{x}_t-\mathbf{x}\rangle-\frac{\lambda}{2}\sum_{t=1}^{T}\|\mathbf{x}-\mathbf{x}_t\|_2^2\\
\leq{}&\frac{\lambda D^2}{4}+4\sum_{t=1}^{T}\frac{1}{\lambda t}\|\nabla f_t(\mathbf{x}_t)-\nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2-\sum_{t=2}^{T}\frac{\lambda(t-1)}{8}\|\mathbf{x}_t-\mathbf{x}_{t-1}\|_2^2.
\end{aligned}$$

Applying Lemma 4 (boundness of the norm of gradient difference) to the $\|\nabla f_t(\mathbf{x}_t)-\nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2$ term, we can get that

$$\begin{aligned}
&\sum_{t=1}^{T}\langle\nabla f_t(\mathbf{x}_t),\mathbf{x}_t-\mathbf{x}\rangle-\frac{\lambda}{2}\sum_{t=1}^{T}\|\mathbf{x}-\mathbf{x}_t\|_2^2\\
\leq{}&\frac{4G^2}{\lambda}+4\sum_{t=2}^{T}\frac{1}{\lambda t}\big(4\|\nabla f_t(\mathbf{x}_t)-\nabla F_t(\mathbf{x}_t)\|_2^2+4\|\nabla F_t(\mathbf{x}_{t-1})-\nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2\\
&+4\|\nabla F_{t-1}(\mathbf{x}_{t-1})-\nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2\big)+\sum_{t=2}^{T}\left(\frac{16L^2}{\lambda t}-\frac{\lambda(t-1)}{8}\right)\|\mathbf{x}_t-\mathbf{x}_{t-1}\|_2^2+\frac{\lambda D^2}{4}.
\end{aligned}$$

The inequality can be simplified by

$$\begin{aligned}
&\sum_{t=2}^{T}\frac{4}{\lambda t}\|\nabla F_t(\mathbf{x}_t)-\nabla f_t(\mathbf{x}_t)\|_2^2+\sum_{t=2}^{T}\frac{4}{\lambda t}\|\nabla F_{t-1}(\mathbf{x}_{t-1})-\nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2\\
\leq{}&\sum_{t=2}^{T}\frac{4}{\lambda t}\|\nabla F_t(\mathbf{x}_t)-\nabla f_t(\mathbf{x}_t)\|_2^2+\sum_{t=2}^{T}\frac{4}{\lambda(t-1)}\|\nabla F_{t-1}(\mathbf{x}_{t-1})-\nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2\\
\leq{}&\sum_{t=1}^{T}\frac{8}{\lambda t}\|\nabla F_t(\mathbf{x}_t)-\nabla f_t(\mathbf{x}_t)\|_2^2. \qquad (24)
\end{aligned}$$

As a result, it turns out that

$$\begin{aligned}
&\sum_{t=1}^{T}\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}\rangle - \frac{\lambda}{2}\sum_{t=1}^{T}\|\mathbf{x}-\mathbf{x}_t\|_2^2 \\
\overset{(24)}{\leq}& \frac{4G^2}{\lambda} + \sum_{t=1}^{T}\frac{32}{\lambda t}\|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 + \sum_{t=2}^{T}\frac{16}{\lambda t}\|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\
&+ \sum_{t=2}^{T}\left(\frac{16L^2}{\lambda t} - \frac{\lambda(t-1)}{8}\right)\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + \frac{\lambda D^2}{4} \\
\leq& \frac{4G^2}{\lambda} + \sum_{t=1}^{T}\frac{32}{\lambda t}\|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 + \sum_{t=2}^{T}\frac{16}{\lambda t}\|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\
&+ \sum_{t=1}^{T-1}\left(\frac{16L^2}{\lambda t} - \frac{\lambda t}{8}\right)\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + \frac{\lambda D^2}{4}.
\end{aligned}$$

Following Sachs et al. (2022), we define $\kappa = \frac{L}{\lambda}$. Then for $t \geq 8\sqrt{2}\kappa$, we have $\frac{16L^2}{\lambda t} - \frac{\lambda t}{8} \leq 0$. Using Assumption 2 (diameter of the domain), the fourth term above is bounded as

$$\begin{aligned}
&\sum_{t=1}^{T-1}\left(\frac{16L^2}{\lambda t} - \frac{\lambda t}{8}\right)\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \leq \sum_{t=1}^{\lceil 8\sqrt{2}\kappa\rceil}\left(\frac{16L^2}{\lambda t} - \frac{\lambda t}{8}\right)D^2 \leq \frac{16L^2D^2}{\lambda}\sum_{t=1}^{\lceil 8\sqrt{2}\kappa\rceil}\frac{1}{t} \\
\leq& \frac{16L^2D^2}{\lambda}\left(1 + \int_{t=1}^{\lceil 8\sqrt{2}\kappa\rceil}\frac{1}{t}dt\right) = \frac{16L^2D^2}{\lambda}\ln\left(1 + 8\sqrt{2}\frac{L}{\lambda}\right) + \frac{16L^2D^2}{\lambda}.
\end{aligned}$$

Combining the above two formulas, we get the bound:

$$\begin{aligned}
&\sum_{t=1}^{T}\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}\rangle - \frac{\lambda}{2}\sum_{t=1}^{T}\|\mathbf{x}-\mathbf{x}_t\|_2^2 \\
\leq& \frac{4G^2}{\lambda} + \sum_{t=1}^{T}\frac{32}{\lambda t}\|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 + \sum_{t=2}^{T}\frac{16}{\lambda t}\|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\
&+ \frac{16L^2D^2}{\lambda}\ln\left(1 + 8\sqrt{2}\frac{L}{\lambda}\right) + \frac{16L^2D^2}{\lambda} + \frac{\lambda D^2}{4}.
\end{aligned}$$

Under Assumption 3 (maximal stochastic variance and maximal adversarial variation), we take expectations over the above bound and get that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T}\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}, \rangle - \frac{\lambda}{2}\sum_{t=1}^{T}\|\mathbf{x}-\mathbf{x}_t\|_2^2\right] \\
=& \frac{32}{\lambda}\sum_{t=1}^{T}\frac{1}{t}\mathbb{E}\left[\|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2\right] + \frac{16}{\lambda}\sum_{t=2}^{T}\frac{1}{t}\mathbb{E}\left[\|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2\right] \\
&+ \frac{16L^2D^2}{\lambda}\ln\left(1 + 8\sqrt{2}\frac{L}{\lambda}\right) + \frac{16L^2D^2 + 4G^2}{\lambda} + \frac{\lambda D^2}{4} \\
\leq& \frac{32}{\lambda}\sum_{t=1}^{T}\frac{1}{t}\sigma_t^2 + \frac{16}{\lambda}\sum_{t=2}^{T}\frac{1}{t}\sup_{\mathbf{x}\in\mathcal{X}}\|\nabla F_t(\mathbf{x}) - \nabla F_{t-1}(\mathbf{x})\|_2^2 + \frac{16L^2D^2}{\lambda}\ln\left(1 + 8\sqrt{2}\frac{L}{\lambda}\right) \\
&+ \frac{16L^2D^2 + 4G^2}{\lambda} + \frac{\lambda D^2}{4} \\
\leq& \frac{1}{\lambda}\left(32\sigma_{\max}^2 + 16\Sigma_{\max}^2\right)(\ln T + 1) + \frac{16L^2D^2}{\lambda}\ln\left(1 + 8\sqrt{2}\frac{L}{\lambda}\right) + \frac{16L^2D^2 + 4G^2}{\lambda} + \frac{\lambda D^2}{4},
\end{aligned} \tag{25}$$

where the last inequality uses the fact that $\sum_{t=1}^T \frac{1}{t} \le \ln T + 1$. Hence, we have proven the $\mathcal{O}\big(\frac{1}{\lambda}(\sigma_{\max}^2 + \Sigma_{\max}^2)\log T\big)$ bound.

To summarize, combining the above two upper bounds for smooth and strongly convex functions by optimistic OMD (see (23) and (25)) and considering the upper bound on expected regret we have proven in (18), we finally achieve the following guarantee:

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^T f_t(\mathbf{x}_t) - \sum_{t=1}^T f_t(\mathbf{x})\right] \\
\le\ & \min\left\{\frac{16G^2}{\lambda}\ln\left(8\sigma_{1:T}^2 + 4\Sigma_{1:T}^2 + G^2 + 1\right) + \frac{16G^2}{\lambda}\ln\left(\frac{512G^2L^2}{\lambda^2}+1\right) + \frac{16G^2+4}{\lambda} + \frac{\lambda D^2}{4},\right.\\
&\left.\frac{1}{\lambda}\left(32\sigma_{\max}^2 + 16\Sigma_{\max}^2\right)(\ln T + 1) + \frac{16L^2D^2 + 4G^2}{\lambda} + \frac{16L^2D^2}{\lambda}\ln\left(1 + 8\sqrt{2}\frac{L}{\lambda}\right) + \frac{\lambda D^2}{4}\right\}\\
=\ & \mathcal{O}\Big(\min\Big\{\frac{1}{\lambda}\log(\sigma_{1:T}^2 + \Sigma_{1:T}^2), \frac{1}{\lambda}(\sigma_{\max}^2 + \Sigma_{\max}^2)\log T\Big\}\Big).
\end{aligned}
$$

□

Proof of Theorem 4. We first present the procedure of optimistic FTRL for λ-strongly convex and smooth functions (Sachs et al., 2022). In each round t, we define a new surrogate loss: $\ell_t(\mathbf{x}) = \langle \nabla f_t(\mathbf{x}_t), \mathbf{x} - \mathbf{x}_t\rangle + \frac{\lambda}{2}\|\mathbf{x} - \mathbf{x}_t\|_2^2$. And the decision $\mathbf{x}_{t+1}$ is determined by

$$
\mathbf{x}_{t+1} = \operatorname*{argmin}_{\mathbf{x}\in\mathcal{X}} \frac{\lambda}{2}\|\mathbf{x} - \mathbf{x}_0\|_2^2 + \sum_{s=1}^t \ell_s(\mathbf{x}) + \langle M_{t+1}, \mathbf{x}\rangle,
$$

where $\mathbf{x}_0$ is an arbitrary point in $\mathcal{X}$, and the optimistic vector $M_{t+1} = \nabla f_t(\mathbf{x}_t)$. In the beginning, we set $M_1 = \nabla f_0(\mathbf{x}_0) = 0$ and thus $\mathbf{x}_1 = \mathbf{x}_0$. Compared with the original algorithm of Sachs et al. (2022), we insert an additional $\frac{\lambda}{2}\|\mathbf{x} - \mathbf{x}_0\|_2^2$ term in the updating rule above, and in this way, the objective function in the t-th round is λt-strongly convex, which facilitates the subsequent analysis.

According to (18), it is easy to verify that

$$
\mathbb{E}\left[\sum_{t=1}^T f_t(\mathbf{x}_t) - \sum_{t=1}^T f_t(\mathbf{x})\right] \le \mathbb{E}\left[\sum_{t=1}^T \ell_t(\mathbf{x}_t) - \sum_{t=1}^T \ell_t(\mathbf{x})\right]. \tag{26}
$$

Thus, we can focus on the regret of the surrogate loss $\ell_t(\cdot)$. From Lemma 11 (standard analysis of optimistic FTRL), since $\frac{\lambda}{2}\|\mathbf{x} - \mathbf{x}_0\|_2^2 + \sum_{s=1}^{t-1}\ell_s(\mathbf{x})$ is λt-strongly convex, we obtain

$$
\begin{aligned}
&\sum_{t=1}^T \ell_t(\mathbf{x}_t) - \sum_{t=1}^T \ell_t(\mathbf{x})\\
\le\ & \frac{\lambda}{2}\|\mathbf{x} - \mathbf{x}_0\|_2^2 + \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1}), \mathbf{x}_t - \mathbf{x}_{t+1}\rangle - \sum_{t=1}^T \frac{\lambda t}{2}\|\mathbf{x}_t - \mathbf{x}_{t+1}\|_2^2\\
\le\ & \frac{\lambda D^2}{2} + \sum_{t=1}^T \frac{1}{\lambda t}\|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 - \sum_{t=1}^T \frac{\lambda t}{4}\|\mathbf{x}_t - \mathbf{x}_{t+1}\|_2^2\\
\le\ & \frac{\lambda D^2}{2} + \sum_{t=1}^T \frac{1}{\lambda t}\|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 - \frac{\lambda}{4}\sum_{t=1}^T \|\mathbf{x}_t - \mathbf{x}_{t+1}\|_2^2,
\end{aligned}\tag{27}
$$

where we again use the inequality of $\langle a, b\rangle \le \|a\|_*\|b\| \le \frac{1}{2c}\|a\|_*^2 + \frac{c}{2}\|b\|^2$ in the second step.

Because there are indeed two different upper bounds for strongly convex and smooth functions by optimistic FTRL in Theorem 4, we will prove the two bounds respectively.

The $\mathcal{O}\big(\frac{1}{\lambda}\log(\sigma_{1:T}^2+\Sigma_{1:T}^2)\big)$ bound. Compared with the proof of Sachs et al. (2022), the main difference here is that we propose a novel way to bound the second term in the last line of (27), which makes use of Lemma 6, in this way we obtain that

$$\sum_{t=1}^T \ell_t(\mathbf{x}_t) - \sum_{t=1}^T \ell_t(\mathbf{x}) \le \frac{4G^2}{\lambda}\ln\left(\bar{V}_T+1\right) - \frac{\lambda}{4}\sum_{t=1}^T \|\mathbf{x}_t-\mathbf{x}_{t+1}\|_2^2 + \frac{4G^2+1}{\lambda} + \frac{\lambda D^2}{2}.$$

Then, we just simplify the above bound in the same way as we do in the proof of Theorem 3 (see (20) $\sim$ (22)). That means, by applying Lemma 5 (boundness of cumulative norm of gradient difference) and the inequality (21), we have

$$\begin{aligned}
&\sum_{t=1}^T \ell_t(\mathbf{x}_t) - \sum_{t=1}^T \ell_t(\mathbf{x})\\
\le{}& \frac{4G^2}{\lambda}\ln\left(8\sum_{t=1}^T\|\nabla f_t(\mathbf{x}_t)-\nabla F_t(\mathbf{x}_t)\|_2^2 + 4\sum_{t=2}^T \|\nabla F_t(\mathbf{x}_{t-1})-\nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 + G^2+1\right)\\
&+\frac{4G^2+1}{\lambda}+\frac{\lambda D^2}{2}+\frac{4G^2}{\lambda}\ln\left(4L^2\sum_{t=2}^T\|\mathbf{x}_t-\mathbf{x}_{t-1}\|_2^2+1\right) - \frac{\lambda}{4}\sum_{t=1}^T\|\mathbf{x}_t-\mathbf{x}_{t+1}\|_2^2.
\end{aligned} \tag{28}$$

And by exploiting Lemma 7, we have

$$\frac{4G^2}{\lambda}\ln\left(4L^2\sum_{t=2}^T\|\mathbf{x}_t-\mathbf{x}_{t-1}\|_2^2+1\right) - \frac{\lambda}{4}\sum_{t=1}^T\|\mathbf{x}_t-\mathbf{x}_{t+1}\|_2^2 \le \frac{4G^2}{\lambda}\ln\left(1+\frac{64G^2L^2}{\lambda^2}\right).$$

Substituting the above inequality into (28), we arrive at

$$\begin{aligned}
&\sum_{t=1}^T \ell_t(\mathbf{x}_t) - \sum_{t=1}^T \ell_t(\mathbf{x})\\
\le{}& \frac{4G^2}{\lambda}\ln\left(8\sum_{t=1}^T\|\nabla f_t(\mathbf{x}_t)-\nabla F_t(\mathbf{x}_t)\|_2^2 + 4\sum_{t=2}^T \|\nabla F_t(\mathbf{x}_{t-1})-\nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 + G^2+1\right)\\
&+\frac{4G^2}{\lambda}\ln\left(1+\frac{64G^2L^2}{\lambda^2}\right)+\frac{4G^2+1}{\lambda}+\frac{\lambda D^2}{2}.
\end{aligned}$$

Taking expectations over the above bound, and applying Jensen's inequality, we have

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^T \ell_t(\mathbf{x}_t) - \sum_{t=1}^T \ell_t(\mathbf{x})\right]\\
\le{}& \frac{4G^2}{\lambda}\ln\left(8\sigma_{1:T}^2+4\Sigma_{1:T}^2+G^2+1\right)+\frac{4G^2}{\lambda}\ln\left(1+\frac{64G^2L^2}{\lambda^2}\right)+\frac{4G^2+1}{\lambda}+\frac{\lambda D^2}{2},
\end{aligned} \tag{29}$$

which completes the proof for the $\mathcal{O}\big(\frac{1}{\lambda}\log(\sigma_{1:T}^2+\Sigma_{1:T}^2)\big)$ bound.

The $\mathcal{O}\big(\frac{1}{\lambda}(\sigma_{\max}^2+\Sigma_{\max}^2)\log T\big)$ bound. For the completeness of the work, we also show the proof of optimistic FTRL on strongly convex functions in Sachs et al. (2022) below. Return to the standard analysis of optimistic FTRL (see (27)), which derives that

$$\sum_{t=1}^T \ell_t(\mathbf{x}_t) - \sum_{t=1}^T \ell_t(\mathbf{x}) \le \frac{\lambda D^2}{2} + \sum_{t=1}^T \frac{1}{\lambda t}\|\nabla f_t(\mathbf{x}_t)-\nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 - \sum_{t=1}^T\frac{\lambda t}{4}\|\mathbf{x}_t-\mathbf{x}_{t+1}\|_2^2.$$

Here they directly use Lemma 4 (boundness of the norm of gradient difference) to bound the above formula:

$$\sum_{t=1}^T \ell_t(\mathbf{x}_t) - \sum_{t=1}^T \ell_t(\mathbf{x})$$

$$
\begin{aligned}
&\le \frac{G^2}{\lambda} + \sum_{t=2}^{T} \frac{1}{\lambda t} \left(4\|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 + 4\|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2\right. \\
&\quad \left. + 4\|\nabla F_{t-1}(\mathbf{x}_{t-1}) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2\right) + \sum_{t=1}^{T} \left(\frac{4L^2}{\lambda(t+1)} - \frac{\lambda t}{4}\right) \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + \frac{\lambda D^2}{2} \\
&\le \frac{G^2}{\lambda} + \sum_{t=1}^{T} \frac{8}{\lambda t}\|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 + \sum_{t=2}^{T} \frac{4}{\lambda t}\|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\
&\quad + \sum_{t=1}^{T} \left(\frac{4L^2}{\lambda t} - \frac{\lambda t}{4}\right) \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + \frac{\lambda D^2}{2}.
\end{aligned} \tag{30}
$$

The above formula reuses the simplification techniques in (24).

Still defining $\kappa = \frac{L}{\lambda}$, then for $t \ge 16\kappa$, there is $\frac{4L^2}{\lambda t} - \frac{\lambda t}{4} \le 0$. For this reason, it turns out that

$$
\begin{aligned}
&\sum_{t=1}^{T} \left(\frac{4L^2}{\lambda t} - \frac{\lambda t}{4}\right) \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \le \sum_{t=1}^{\lceil 16\kappa \rceil} \left(\frac{4L^2}{\lambda t} - \frac{\lambda t}{4}\right) D^2 \le \frac{4L^2 D^2}{\lambda} \sum_{t=1}^{\lceil 16\kappa \rceil} \frac{1}{t} \\
&\le \frac{4L^2 D^2}{\lambda} \left(1 + \int_{t=1}^{\lceil 16\kappa \rceil} \frac{1}{t}\right) = \frac{4L^2 D^2}{\lambda} \ln\left(1 + 16\frac{L}{\lambda}\right) + \frac{4L^2 D^2}{\lambda}.
\end{aligned}
$$

Then substitute the above inequality into (30), and take expectations over it under Assumption 3 (maximal stochastic variance and maximal adversarial variation):

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T} \ell_t(\mathbf{x}_t) - \sum_{t=1}^{T} \ell_t(\mathbf{x})\right] \\
=& \frac{8}{\lambda} \sum_{t=1}^{T} \frac{1}{t} \mathbb{E}\left[\|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2\right] + \frac{4}{\lambda} \sum_{t=2}^{T} \frac{1}{t} \mathbb{E}\left[\|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2\right] \\
&+ \frac{4L^2 D^2}{\lambda} \ln\left(1 + 16\frac{L}{\lambda}\right) + \frac{4L^2 D^2 + G^2}{\lambda} + \frac{\lambda D^2}{2} \\
\le& \frac{8}{\lambda} \sum_{t=1}^{T} \frac{1}{t} \sigma_t^2 + \frac{4}{\lambda} \sum_{t=2}^{T} \frac{1}{t} \sup_{\mathbf{x} \in \mathcal{X}} \|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 + \frac{4L^2 D^2}{\lambda} \ln\left(1 + 16\frac{L}{\lambda}\right) \\
&+ \frac{4L^2 D^2 + G^2}{\lambda} + \frac{\lambda D^2}{2} \\
\le& \frac{1}{\lambda} \left(8\sigma_{\max}^2 + 4\Sigma_{\max}^2\right) (\ln T + 1) + \frac{4L^2 D^2}{\lambda} \ln\left(1 + 16\frac{L}{\lambda}\right) + \frac{4L^2 D^2 + G^2}{\lambda} + \frac{\lambda D^2}{2},
\end{aligned} \tag{31}
$$

where the last inequality uses the fact that $\sum_{t=1}^{T} \frac{1}{t} \le \ln T + 1$. So the proof for the $\mathcal{O}\big(\frac{1}{\lambda}(\sigma_{\max}^2 + \Sigma_{\max}^2) \log T\big)$ bound is finished.

To summarize, we complete the proof by combining the two upper bounds in (29) and (31):

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{x})\right] \\
\le& \min\Big\{\frac{4G^2}{\lambda} \ln\left(8\sigma_{1:T}^2 + 4\Sigma_{1:T}^2 + G^2 + 1\right) + \frac{4G^2}{\lambda} \ln\left(1 + \frac{64G^2L^2}{\lambda^2}\right) + \frac{4G^2 + 1}{\lambda} + \frac{\lambda D^2}{2} \\
&\frac{1}{\lambda} \left(8\sigma_{\max}^2 + 4\Sigma_{\max}^2\right) (\ln T + 1) + \frac{4L^2 D^2}{\lambda} \ln\left(1 + 16\frac{L}{\lambda}\right) + \frac{4L^2 D^2 + G^2}{\lambda} + \frac{\lambda D^2}{2}\Big\} \\
\le& \mathcal{O}\Big(\min\Big\{\frac{1}{\lambda} \log(\sigma_{1:T}^2 + \Sigma_{1:T}^2), \frac{1}{\lambda}(\sigma_{\max}^2 + \Sigma_{\max}^2) \log T\Big\}\Big).
\end{aligned}
$$

□

A.3. Exp-Concave and Smooth Functions

Proof of Theorem 5. As stated in Assumption 7 (exponential concavity of individual function), we operate optimistic OMD on α-exp-concave individual functions in Theorem 5. Accordingly, we have $f_t(\mathbf{x}_t) - f_t(\mathbf{x}) \leq \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}\rangle - \frac{\beta}{2}\|\mathbf{x} - \mathbf{x}_t\|_{h_t}^2$, where $\beta = \frac{1}{2}\min\left\{\frac{1}{4GD}, \alpha\right\}$, and $h_t = \nabla f_t(\mathbf{x}_t)\nabla f_t(\mathbf{x}_t)^\top$. Therefore, we can take advantage of the above formula to get tighter regret bounds as follows

$$\mathbb{E}\left[\sum_{t=1}^T f_t(\mathbf{x}_t) - \sum_{t=1}^T f_t(\mathbf{x})\right] \leq \mathbb{E}\left[\sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}\rangle - \frac{\beta}{2}\sum_{t=1}^T \|\mathbf{x} - \mathbf{x}_t\|_{h_t}^2\right]. \tag{32}$$

Clearly, $\psi_t(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|_{H_t}^2$ is a 1-strongly convex function with respect to $\|\cdot\|_{H_t}$, and $\|\cdot\|_{H_t}^{-1}$ is the dual norm of it. Thus we can get the following formula from Lemma 1 (variant of Bregman proximal inequality) with a sum operation:

$$\begin{aligned}
&\sum_{t=1}^T \langle \mathbf{x}_t - \mathbf{x}, \nabla f_t(\mathbf{x}_t)\rangle - \frac{\beta}{2}\sum_{t=1}^T \|\mathbf{x} - \mathbf{x}_t\|_{h_t}^2 \\
\leq\ & \underbrace{\sum_{t=1}^T \left(\frac{1}{2}\|\mathbf{x} - \widehat{\mathbf{x}}_t\|_{H_t}^2 - \frac{1}{2}\|\mathbf{x} - \widehat{\mathbf{x}}_{t+1}\|_{H_t}^2\right) - \frac{\beta}{2}\sum_{t=1}^T \|\mathbf{x} - \mathbf{x}_t\|_{h_t}^2}_{\texttt{term (a)}} \\
&+ \underbrace{\sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_{H_t^{-1}}^2}_{\texttt{term (b)}} - \underbrace{\sum_{t=1}^T \frac{1}{2}\left(\|\mathbf{x}_t - \widehat{\mathbf{x}}_t\|_{H_t}^2 + \|\widehat{\mathbf{x}}_{t+1} - \mathbf{x}_t\|_{H_t}^2\right)}_{\texttt{term (c)}}.
\end{aligned} \tag{33}$$

Then, we discuss the upper bounds of term (a), term (b) and term (c), respectively. According to Chiang et al. (2012, Proof of Lemma 14), we write term (a) as

$$\begin{aligned}
\texttt{term (a)} = &\frac{1}{2}\|\mathbf{x} - \widehat{\mathbf{x}}_1\|_{H_1}^2 - \frac{1}{2}\|\mathbf{x} - \widehat{\mathbf{x}}_{T+1}\|_{H_{T+1}}^2 + \frac{1}{2}\sum_{t=1}^T \left(\|\mathbf{x} - \widehat{\mathbf{x}}_{t+1}\|_{H_{t+1}}^2 - \|\mathbf{x} - \widehat{\mathbf{x}}_{t+1}\|_{H_t}^2\right) \\
&- \frac{\beta}{2}\sum_{t=1}^T \|\mathbf{x} - \mathbf{x}_t\|_{h_t}^2.
\end{aligned}$$

Based on Assumption 2 (diameter of the domain), with the definition that $H_t = I + \frac{\beta}{2}G^2 I + \frac{\beta}{2}\sum_{\tau=1}^{t-1}\nabla f_\tau(\mathbf{x}_\tau)\nabla f_\tau(\mathbf{x}_\tau)^\top$ (see (13)) and $h_t = \nabla f_t(\mathbf{x}_t)\nabla f_t(\mathbf{x}_t)^\top$, we have $\|\mathbf{x} - \widehat{\mathbf{x}}_1\|_{H_1}^2 \leq D^2\left(1 + \frac{\beta}{2}G^2\right)$ and $H_{t+1} - H_t = \frac{\beta}{2}h_t$. Thus we can simplify term (a) to

$$\begin{aligned}
\texttt{term (a)} &\leq D^2\left(1 + \frac{\beta}{2}G^2\right) + \frac{\beta}{4}\sum_{t=1}^T \|\mathbf{x} - \widehat{\mathbf{x}}_{t+1}\|_{h_t}^2 - \frac{\beta}{2}\sum_{t=1}^T \|\mathbf{x} - \mathbf{x}_t\|_{h_t}^2 \\
&\leq D^2\left(1 + \frac{\beta}{2}G^2\right) + \frac{\beta}{2}\sum_{t=1}^T \|\mathbf{x}_t - \widehat{\mathbf{x}}_{t+1}\|_{h_t}^2 \leq D^2\left(1 + \frac{\beta}{2}G^2\right) + \sum_{t=1}^T \|\mathbf{x}_t - \widehat{\mathbf{x}}_{t+1}\|_{H_t}^2 \\
&\leq D^2\left(1 + \frac{\beta}{2}G^2\right) + \sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_{H_t^{-1}}^2 = D^2\left(1 + \frac{\beta}{2}G^2\right) + \texttt{term (b)},
\end{aligned}$$

where the third inequality is based on the fact that $H_t \succeq \frac{\beta}{2}G^2 I \succeq \frac{\beta}{2}h_t$ and the 4th inequality uses Lemma 3 (stability lemma).

Obviously, the upper bound of term (b) determines that of term (a). Hence we move to bound term (b). Due to the definition of H_t and Assumption 1 (boundness of the gradient norm), there is $G^2 I \succeq \nabla f_t(\mathbf{x}_t)\nabla f_t(\mathbf{x}_t)^\top$ for every t. In addition, we know $\nabla f_0(\mathbf{x}_0) = 0$, so we have

$$H_t \succeq I + \frac{\beta}{4}\sum_{\tau=1}^t \left(\nabla f_\tau(\mathbf{x}_\tau)\nabla f_\tau(\mathbf{x}_\tau)^\top + \nabla f_{\tau-1}(\mathbf{x}_{\tau-1})\nabla f_{\tau-1}(\mathbf{x}_{\tau-1})^\top\right). \tag{34}$$

As well as Chiang et al. (2012), we claim that

$$\nabla f_\tau(\mathbf{x}_\tau)\nabla f_\tau(\mathbf{x}_\tau)^\top + \nabla f_{\tau-1}(\mathbf{x}_{\tau-1})\nabla f_{\tau-1}(\mathbf{x}_{\tau-1})^\top \\ \succeq \frac{1}{2}\left(\nabla f_\tau(\mathbf{x}_\tau) - \nabla f_{\tau-1}(\mathbf{x}_{\tau-1})\right)\left(\nabla f_\tau(\mathbf{x}_\tau) - \nabla f_{\tau-1}(\mathbf{x}_{\tau-1})\right)^\top. \tag{35}$$

The above inequality comes from subtracting the RHS of it from the left and getting that $\frac{1}{2}\left(\nabla f_\tau(\mathbf{x}_\tau) + \nabla f_{\tau-1}(\mathbf{x}_{\tau-1})\right)\left(\nabla f_\tau(\mathbf{x}_\tau) + \nabla f_{\tau-1}(\mathbf{x}_{\tau-1})\right)^\top \succeq 0$.

To this end, we substituting (35) into (34) and obtain

$$H_t \overset{(35)}{\succeq} I + \frac{\beta}{8}\sum_{\tau=1}^{t}\left(\nabla f_\tau(\mathbf{x}_\tau) - \nabla f_{\tau-1}(\mathbf{x}_{\tau-1})\right)\left(\nabla f_\tau(\mathbf{x}_\tau) - \nabla f_{\tau-1}(\mathbf{x}_{\tau-1})\right)^\top.$$

Let $P_t = I + \frac{\beta}{8}\sum_{\tau=1}^{t}\left(\nabla f_\tau(\mathbf{x}_\tau) - \nabla f_{\tau-1}(\mathbf{x}_{\tau-1})\right)\left(\nabla f_\tau(\mathbf{x}_\tau) - \nabla f_{\tau-1}(\mathbf{x}_{\tau-1})\right)^\top$, we have

$$\texttt{term (b)} \le \sum_{t=1}^{T}\|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_{P_t^{-1}}^2 = \frac{8}{\beta}\sum_{t=1}^{T}\left\|\sqrt{\frac{\beta}{8}}\left(\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\right)\right\|_{P_t^{-1}}^2.$$

As a result, applying Lemma 10 with $\mathbf{u}_t = \sqrt{\frac{\beta}{8}}\left(\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\right)$ and $\varepsilon = 1$, we get

$$\texttt{term (b)} \le \frac{8d}{\beta}\ln\left(\frac{\beta}{8d}\bar{V}_T + 1\right).$$

Then, derived from the fact that $H_t \succeq H_{t-1} \succeq I$, we can bound term (c) as

$$\begin{aligned}\texttt{term (c)} &= \frac{1}{2}\sum_{t=1}^{T}\|\mathbf{x}_t - \widehat{\mathbf{x}}_t\|_{H_t}^2 + \frac{1}{2}\sum_{t=2}^{T+1}\|\mathbf{x}_{t-1} - \widehat{\mathbf{x}}_t\|_{H_{t-1}}^2 \\ &\ge \frac{1}{2}\sum_{t=2}^{T}\|\mathbf{x}_t - \widehat{\mathbf{x}}_t\|_{H_{t-1}}^2 + \frac{1}{2}\sum_{t=2}^{T}\|\mathbf{x}_{t-1} - \widehat{\mathbf{x}}_t\|_{H_{t-1}}^2 \ge \frac{1}{4}\sum_{t=2}^{T}\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2.\end{aligned}$$

Combining the above bounds of term (a), term (b) and term (c), we can get

$$\begin{aligned}&\sum_{t=1}^{T}\langle \mathbf{x}_t - \mathbf{x}, \nabla f_t(\mathbf{x}_t)\rangle - \frac{\beta}{2}\sum_{t=1}^{T}\|\mathbf{x} - \mathbf{x}_t\|_{h_t}^2 \\ \le{}& \frac{16d}{\beta}\ln\left(\frac{\beta}{8d}\bar{V}_T + 1\right) + D^2\left(1 + \frac{\beta}{2}G^2\right) - \frac{1}{4}\sum_{t=2}^{T}\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2.\end{aligned}$$

Then exploiting Lemma 5 (boundness of cumulative norm of gradient difference) with the inequality $\ln(1+u+v) \le \ln(1+u) + \ln(1+v)$, we have

$$\begin{aligned}&\sum_{t=1}^{T}\langle \mathbf{x}_t - \mathbf{x}, \nabla f_t(\mathbf{x}_t)\rangle - \frac{\beta}{2}\sum_{t=1}^{T}\|\mathbf{x} - \mathbf{x}_t\|_{h_t}^2 \\ \le{}& \frac{16d}{\beta}\ln\left(\frac{\beta}{d}\sum_{t=1}^{T}\|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 + \frac{\beta}{2d}\sum_{t=2}^{T}\|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 + \frac{\beta}{8d}G^2 + 1\right) \\ &+ \frac{16d}{\beta}\ln\left(\frac{\beta L^2}{2d}\sum_{t=2}^{T}\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + 1\right) + D^2\left(1 + \frac{\beta}{2}G^2\right) - \frac{1}{4}\sum_{t=2}^{T}\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2,\end{aligned} \tag{36}$$

Similarly, we still use Lemma 7 to simplify the above formula:

$$\frac{16d}{\beta}\ln\left(\frac{\beta L^2}{2d}\sum_{t=2}^{T}\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + 1\right) - \frac{1}{4}\sum_{t=2}^{T}\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \le \frac{16d}{\beta}\ln\left(32L^2 + 1\right). \tag{37}$$

Through combining (36) and (37), we obtain that

$$\begin{aligned}
&\sum_{t=1}^{T}\langle \mathbf{x}_t - \mathbf{x}, \nabla f_t(\mathbf{x}_t)\rangle - \frac{\beta}{2}\sum_{t=1}^{T}\|\mathbf{x}-\mathbf{x}_t\|_{h_t}^2 \\
\le& \frac{16d}{\beta}\ln\left(\frac{\beta}{d}\sum_{t=1}^{T}\|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 + \frac{\beta}{2d}\sum_{t=2}^{T}\|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 + \frac{\beta}{8d}G^2 + 1\right) \\
&+ \frac{16d}{\beta}\ln\left(32L^2+1\right) + D^2\left(1+\frac{\beta}{2}G^2\right).
\end{aligned}$$

Taking the expectation, and making use of Jensen's inequality, the above bound becomes

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T}\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}\rangle - \frac{\beta}{2}\sum_{t=1}^{T}\|\mathbf{x}-\mathbf{x}_t\|_{h_t}^2\right] \\
\le& \frac{16d}{\beta}\ln\left(\frac{\beta}{d}\sigma_{1:T}^2 + \frac{\beta}{2d}\Sigma_{1:T}^2 + \frac{\beta}{8d}G^2 + 1\right) + \frac{16d}{\beta}\ln\left(32L^2+1\right) + D^2\left(1+\frac{\beta}{2}G^2\right) \\
=& \mathcal{O}\Big(\frac{d}{\alpha}\log(\sigma_{1:T}^2 + \Sigma_{1:T}^2)\Big)
\end{aligned}$$

We finish the proof by integrating the above inequality with the proven result of expected regret for exp-concave losses (see (32)). □

Proof of Theorem 6. We demonstrate the new procedure of optimistic FTRL for α-exp-concave and smooth functions. We design a surrogate loss for each round t: $\ell_t(\mathbf{x}) = \langle \nabla f_t(\mathbf{x}_t), \mathbf{x} - \mathbf{x}_t\rangle + \frac{\beta}{2}\|\mathbf{x}-\mathbf{x}_t\|_{h_t}^2$, where $\beta = \frac{1}{2}\min\left\{\frac{1}{4GD}, \alpha\right\}$, and $h_t = \nabla f_t(\mathbf{x}_t)\nabla f_t(\mathbf{x}_t)^\top$. And we use the following updating rule:

$$\mathbf{x}_{t+1} = \operatorname*{argmin}_{\mathbf{x}\in\mathcal{X}} \frac{1}{2}(1+\beta G^2)\|\mathbf{x}\|_2^2 + \sum_{s=1}^{t}\ell_s(\mathbf{x}) + \langle M_{t+1}, \mathbf{x}\rangle,$$

where $\mathbf{x}_0$ is an arbitrary point in $\mathcal{X}$, and $M_{t+1} = \nabla f_t(\mathbf{x}_t)$. Furthermore, we set $M_1 = \nabla f_0(\mathbf{x}_0) = 0$. From (32), we can easily derive that

$$\mathbb{E}\left[\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{x})\right] \le \mathbb{E}\left[\sum_{t=1}^{T}\ell_t(\mathbf{x}_t) - \sum_{t=1}^{T}\ell_t(\mathbf{x})\right]. \tag{38}$$

So in the following, we concentrate on the regret of surrogate losses. Denoting by $H_t = I + \beta G^2 I + \beta\sum_{s=1}^{t-1} h_s$ (where I is the $d\times d$ identity matrix) and $G_t(\mathbf{x}) = \frac{1}{2}(1+\beta G^2)\|\mathbf{x}\|_2^2 + \sum_{s=1}^{t-1}\ell_s(\mathbf{x})$, we have that $G_t(\mathbf{x})$ is 1-strongly convex w.r.t. $\|\cdot\|_{H_t}$. Hence, using Lemma 11 (standard analysis of optimistic FTRL), we immediately get the following guarantee

$$\begin{aligned}
&\sum_{t=1}^{T}\ell_t(\mathbf{x}_t) - \sum_{t=1}^{T}\ell_t(\mathbf{x}) \\
\le& \frac{1+\beta G^2}{2}\|\mathbf{x}\|_2^2 + \sum_{t=1}^{T}\left(\langle \nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1}), \mathbf{x}_t - \mathbf{x}_{t+1}\rangle - \frac{1}{2}\|\mathbf{x}_t - \mathbf{x}_{t+1}\|_{H_t}^2\right) \\
\le& \frac{(1+\beta G^2)D^2}{2} + \underbrace{\sum_{t=1}^{T}\|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_{H_t^{-1}}^2}_{\texttt{term (a)}} - \underbrace{\frac{1}{4}\sum_{t=1}^{T}\|\mathbf{x}_t - \mathbf{x}_{t+1}\|_{H_t}^2}_{\texttt{term (b)}},
\end{aligned} \tag{39}$$

where we denote the dual norm of $\|\cdot\|_{H_t}$ by $\|\cdot\|_{H_t^{-1}}$, and use Assumption 2 (diameter of the domain) and $\langle a, b\rangle \le \|a\|_*\|b\| \le \frac{1}{2c}\|a\|_*^2 + \frac{c}{2}\|b\|^2$ in the second inequality.

To bound term (a) in (39), we begin with the fact that

$$\begin{aligned} H_t &\succeq I + \beta \sum_{s=1}^{t} \nabla f_s(\mathbf{x}_s) \nabla f_s(\mathbf{x}_s)^\top \\ &\succeq I + \frac{\beta}{2} \sum_{s=1}^{t} \left(\nabla f_s(\mathbf{x}_s) \nabla f_s(\mathbf{x}_s)^\top + \nabla f_{s-1}(\mathbf{x}_{s-1}) \nabla f_{s-1}(\mathbf{x}_{s-1})^\top \right), \end{aligned} \tag{40}$$

where the first inequality is due to Assumption 1 (boundness of the gradient norm) and the second inequality comes from the definition that $\nabla f_0(\mathbf{x}_0) = 0$. We substitute (35) in the proof of Theorem 5 into (40) and obtain that

$$H_t \succeq I + \frac{\beta}{4} \sum_{s=1}^{t} \left(\nabla f_s(\mathbf{x}_s) - \nabla f_{s-1}(\mathbf{x}_{s-1}) \right) \left(\nabla f_s(\mathbf{x}_s) - \nabla f_{s-1}(\mathbf{x}_{s-1}) \right)^\top.$$

Let $P_t = I + \frac{\beta}{4} \sum_{s=1}^{t} \left(\nabla f_s(\mathbf{x}_s) - \nabla f_{s-1}(\mathbf{x}_{s-1}) \right) \left(\nabla f_s(\mathbf{x}_s) - \nabla f_{s-1}(\mathbf{x}_{s-1}) \right)^\top$ so that $H_t \succeq P_t$, then we can bound term (a) in (39) as

$$\texttt{term (a)} \leq \sum_{t=1}^{T} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_{P_t^{-1}}^2 = \frac{4}{\beta} \sum_{t=1}^{T} \left\| \sqrt{\frac{\beta}{4}} (\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})) \right\|_{P_t^{-1}}^2.$$

By applying Lemma 10 with $\mathbf{u}_t = \sqrt{\frac{\beta}{4}} (\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1}))$ and $\varepsilon = 1$, we get that

$$\texttt{term (a)} \leq \frac{4d}{\beta} \ln \left(\frac{\beta}{4d} \bar{V}_T + 1 \right).$$

Then we move to term (b). Since $H_t = I + \beta G^2 I + \beta \sum_{s=1}^{t-1} h_s \succeq I$, we can derive that

$$\texttt{term (b)} = \frac{1}{4} \sum_{t=1}^{T} \|\mathbf{x}_t - \mathbf{x}_{t+1}\|_{H_t}^2 \geq \frac{1}{4} \sum_{t=1}^{T} \|\mathbf{x}_t - \mathbf{x}_{t+1}\|_I^2 = \frac{1}{4} \sum_{t=1}^{T} \|\mathbf{x}_t - \mathbf{x}_{t+1}\|_2^2.$$

As a result, we bound the guarantee in (39) by substituting the bounds of term (a) and term (b):

$$\sum_{t=1}^{T} \ell_t(\mathbf{x}_t) - \sum_{t=1}^{T} \ell_t(\mathbf{x}) \leq \frac{(1 + \beta G^2) D^2}{2} + \frac{4d}{\beta} \ln \left(\frac{\beta}{4d} \bar{V}_T + 1 \right) - \frac{1}{4} \sum_{t=1}^{T} \|\mathbf{x}_t - \mathbf{x}_{t+1}\|_2^2.$$

Through Lemma 5 (boundness of cumulative norm of gradient difference) together with the inequality of $\ln(1 + u + v) \leq \ln(1 + u) + \ln(1 + v)$, we get that

$$\begin{aligned} &\sum_{t=1}^{T} \ell_t(\mathbf{x}_t) - \sum_{t=1}^{T} \ell_t(\mathbf{x}) \\ \leq &\frac{4d}{\beta} \ln \left(\frac{2\beta}{d} \sum_{t=1}^{T} \|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 + \frac{\beta}{d} \sum_{t=2}^{T} \|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 + \frac{\beta}{4d} G^2 + 1 \right) \\ &+ \frac{(1 + \beta G^2) D^2}{2} + \frac{4d}{\beta} \ln \left(\frac{\beta L^2}{d} \sum_{t=2}^{T} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + 1 \right) - \frac{1}{4} \sum_{t=1}^{T} \|\mathbf{x}_t - \mathbf{x}_{t+1}\|_2^2. \end{aligned} \tag{41}$$

By applying Lemma 7, the last two terms in (41) are bounded by

$$\frac{4d}{\beta} \ln \left(\frac{\beta L^2}{d} \sum_{t=2}^{T} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + 1 \right) - \frac{1}{4} \sum_{t=1}^{T} \|\mathbf{x}_t - \mathbf{x}_{t+1}\|_2^2 \leq \frac{4d}{\beta} \ln(16L^2 + 1). \tag{42}$$

Combining (41) with (42), we arrive at

$$\begin{aligned}
&\sum_{t=1}^{T}\ell_t(\mathbf{x}_t)-\sum_{t=1}^{T}\ell_t(\mathbf{x})\\
&\leq \frac{4d}{\beta}\ln\left(\frac{2\beta}{d}\sum_{t=1}^{T}\|\nabla f_t(\mathbf{x}_t)-\nabla F_t(\mathbf{x}_t)\|_2^2+\frac{\beta}{d}\sum_{t=2}^{T}\|\nabla F_t(\mathbf{x}_{t-1})-\nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2+\frac{\beta}{4d}G^2+1\right)\\
&\quad+\frac{(1+\beta G^2)D^2}{2}+\frac{4d}{\beta}\ln(16L^2+1).
\end{aligned}$$

Then we compute the expected regret by taking the expectation over the above regret with the help of Jensen's inequality and the derived result in (38):

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T}f_t(\mathbf{x}_t)-\sum_{t=1}^{T}f_t(\mathbf{x})\right]\leq\mathbb{E}\left[\sum_{t=1}^{T}\ell_t(\mathbf{x}_t)-\sum_{t=1}^{T}\ell_t(\mathbf{x})\right]\\
&\leq\frac{4d}{\beta}\ln\left(\frac{2\beta}{d}\sigma_{1:T}^2+\frac{\beta}{d}\Sigma_{1:T}^2+\frac{\beta}{4d}G^2+1\right)+\frac{(1+\beta G^2)D^2}{2}+\frac{4d}{\beta}\ln(16L^2+1)\\
&=\mathcal{O}\Big(\frac{d}{\alpha}\log(\sigma_{1:T}^2+\Sigma_{1:T}^2)\Big).
\end{aligned}$$

□

A.4. Proof of Theorem 7

In this part, we present the proof of Theorem 7. Since our algorithmic design is based on the collaborative online ensemble framework proposed by Zhao et al. (2021), we first introduce the following general theorem (Zhao et al., 2021, Theorem 5) and provide the proof for our theorem based on it.

Theorem 8 (Theorem 5 of Zhao et al. (2021).). *Under Assumption 1 (boundness of the gradient norm) and Assumption 2 (diameter of the domain), setting the step size pool $\mathcal{H}$ as*

$$\mathcal{H}=\left\{\eta_i=\min\left\{\bar{\eta},\sqrt{\frac{D^2}{8G^2T}\cdot 2^{i-1}}\right\}|i\in[N]\right\},\tag{43}$$

where $N=\lceil 2^{-1}\log_2((8G^2T\bar{\eta}^2)/D^2)\rceil+1$, and setting the meta learning rate as

$$\varepsilon_t=\min\left\{\bar{\varepsilon},\sqrt{\frac{\ln N}{D^2\sum_{s=1}^{t}\|\nabla f_t(\mathbf{x}_t)-f_{t-1}(\mathbf{x}_{t-1})\|_2^2}}\right\},$$

Algorithm 2 enjoys the following dynamic regret guarantee:

$$\begin{aligned}
&\sum_{t=1}^{T}\langle\nabla f_t(\mathbf{x}_t),\mathbf{x}_t-\mathbf{u}_t\rangle\\
&\leq 5\sqrt{D^2\ln N\bar{V}_T}+2\sqrt{(D^2+2DP_T)\bar{V}_T}+\frac{\ln N}{\bar{\varepsilon}}+8\bar{\varepsilon}D^2G^2\\
&\quad+\frac{D^2+2DP_T}{\bar{\eta}}+\left(\lambda-\frac{1}{4\bar{\eta}}\right)\sum_{t=2}^{T}\|\mathbf{x}_{t,i}-\mathbf{x}_{t-1,i}\|_2^2-\frac{1}{4\bar{\varepsilon}}\sum_{t=2}^{T}\|\boldsymbol{p}_t-\boldsymbol{p}_{t-1}\|_1^2-\lambda\sum_{t=2}^{T}\sum_{i=1}^{N}p_{t,i}\|\mathbf{x}_{t,i}-\mathbf{x}_{t-1,i}\|_2^2.
\end{aligned}$$

In the above, $\bar{V}_T=\sum_{t=1}^{T}\|\nabla f_t(\mathbf{x}_t)-\nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2$ is the adaptivity term measuring the quality of optimistic gradient vectors $\{M_t=f_{t-1}(\mathbf{x}_{t-1})\}_{t=1}^{T}$, and $P_T=\sum_{t=2}^{T}\|\mathbf{u}_{t-1}-\mathbf{u}_t\|_2$ is the path length of comparators.

In the following, we prove Theorem 7 based on Theorem 8.

Proof. Noting that $\bar{V}_T = \sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2$ in Theorem 8, thus by applying Lemma 5 (Boundness of cumulative norm of gradient difference) and the inequality $\sqrt{a+b} \le \sqrt{a} + \sqrt{b} (a, b > 0)$, we can bound the first and second term in the regret of Theorem 8 as

$$
\begin{aligned}
&5\sqrt{D^2 \ln N \bar{V}_T} + 2\sqrt{(D^2 + 2DP_T)\bar{V}_T} \\
= &\left(5\sqrt{D^2 \ln N} + 2\sqrt{(D^2 + 2DP_T)}\right)\sqrt{\bar{V}_T} \\
\le &\, G\left(5\sqrt{D^2 \ln N} + 2\sqrt{(D^2 + 2DP_T)}\right) + \left(5\sqrt{D^2 \ln N} + 2\sqrt{(D^2 + 2DP_T)}\right)\sqrt{4L^2 \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2} \\
&+ 2\sqrt{2}\left(5\sqrt{D^2 \ln N} + 2\sqrt{(D^2 + 2DP_T)}\right)\sqrt{\sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2} \\
&+ 2\left(5\sqrt{D^2 \ln N} + 2\sqrt{(D^2 + 2DP_T)}\right)\sqrt{\sum_{t=2}^T \|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2}.
\end{aligned} \tag{44}
$$

To eliminate the relevant terms of $\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2$, we first prove that

$$
\begin{aligned}
\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 &= \left\|\sum_{i=1}^N p_{t,i}\mathbf{x}_{t,i} - \sum_{i=1}^N p_{t-1,i}\mathbf{x}_{t-1,i}\right\|_2^2 \\
&\le 2\left\|\sum_{i=1}^N p_{t,i}\mathbf{x}_{t,i} - \sum_{i=1}^N p_{t,i}\mathbf{x}_{t-1,i}\right\|_2^2 + 2\left\|\sum_{i=1}^N p_{t,i}\mathbf{x}_{t-1,i} - \sum_{i=1}^N p_{t-1,i}\mathbf{x}_{t-1,i}\right\|_2^2 \\
&\le 2\left(\sum_{i=1}^N p_{t,i}\|\mathbf{x}_{t,i} - \mathbf{x}_{t-1,i}\|_2\right)^2 + 2\left(\sum_{i=1}^N |p_{t,i} - p_{t-1,i}|\|\mathbf{x}_{t-1,i}\|_2\right)^2 \\
&\le 2\sum_{i=1}^N p_{t,i}\|\mathbf{x}_{t,i} - \mathbf{x}_{t-1,i}\|_2^2 + 2D^2\|\boldsymbol{p}_t - \boldsymbol{p}_{t-1}\|_1^2.
\end{aligned}
$$

Thus we can get

$$
\sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 = 2\sum_{t=2}^T\sum_{i=1}^N p_{t,i}\|\mathbf{x}_{t,i} - \mathbf{x}_{t-1,i}\|_2^2 + 2D^2\sum_{t=2}^T \|\boldsymbol{p}_t - \boldsymbol{p}_{t-1}\|_1^2.
$$

Then we can use the above inequality and the AM-GM inequality to bound the second term in (44):

$$
\begin{aligned}
&\left(5\sqrt{D^2 \ln N} + 2\sqrt{(D^2 + 2DP_T)}\right)\sqrt{4L^2 \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2} \\
\le &\, 5\sqrt{D^2 \ln N\left(8L^2\sum_{t=2}^T\sum_{i=1}^N p_{t,i}\|\mathbf{x}_{t,i} - \mathbf{x}_{t-1,i}\|_2^2 + 8L^2D^2\sum_{t=2}^T \|\boldsymbol{p}_t - \boldsymbol{p}_{t-1}\|_1^2\right)} \\
&+ 2\sqrt{(D^2 + 2DP_T)\left(8L^2\sum_{t=2}^T\sum_{i=1}^N p_{t,i}\|\mathbf{x}_{t,i} - \mathbf{x}_{t-1,i}\|_2^2 + 8L^2D^2\sum_{t=2}^T \|\boldsymbol{p}_t - \boldsymbol{p}_{t-1}\|_1^2\right)} \\
\le &\, \frac{25\ln N}{4\bar{\varepsilon}} + 8\bar{\varepsilon}D^2L^2\sum_{t=2}^T\sum_{i=1}^N p_{t,i}\|\mathbf{x}_{t,i} - \mathbf{x}_{t-1,i}\|_2^2 + 8\bar{\varepsilon}L^2D^4\sum_{t=2}^T \|\boldsymbol{p}_t - \boldsymbol{p}_{t-1}\|_1^2 \\
&+ \frac{D^2 + 2DP_T}{\bar{\eta}} + 8\bar{\eta}L^2\sum_{t=2}^T\sum_{i=1}^N p_{t,i}\|\mathbf{x}_{t,i} - \mathbf{x}_{t-1,i}\|_2^2 + 8\bar{\eta}L^2D^2\sum_{t=2}^T \|\boldsymbol{p}_t - \boldsymbol{p}_{t-1}\|_1^2
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{25\ln N}{4\bar{\varepsilon}} + \frac{D^2+2DP_T}{\bar{\eta}} + \left(8\bar{\varepsilon}D^2L^2 + 8\bar{\eta}L^2\right)\sum_{t=2}^{T}\sum_{i=1}^{N} p_{t,i}\|\mathbf{x}_{t,i}-\mathbf{x}_{t-1,i}\|_2^2 \\
&\quad + \left(8\bar{\varepsilon}L^2D^4 + 8\bar{\eta}L^2D^2\right)\sum_{t=2}^{T}\|\boldsymbol{p}_t-\boldsymbol{p}_{t-1}\|_1^2.
\end{aligned}
$$

Combining (44) and the above formula with the regret in Theorem 8, we have

$$
\begin{aligned}
&\sum_{t=1}^{T}\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{u}_t\rangle \\
&\le G\left(5\sqrt{D^2\ln N} + 2\sqrt{(D^2+2DP_T)}\right) + 2\sqrt{2}\left(5\sqrt{D^2\ln N} + 2\sqrt{(D^2+2DP_T)}\right)\sqrt{\sum_{t=1}^{T}\|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2} \\
&\quad + 2\left(5\sqrt{D^2\ln N} + 2\sqrt{(D^2+2DP_T)}\right)\sqrt{\sum_{t=2}^{T}\|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2} + \frac{29\ln N}{4\bar{\varepsilon}} \\
&\quad + 8\bar{\varepsilon}D^2G^2 + \frac{2D^2+4DP_T}{\bar{\eta}} + \left(\lambda - \frac{1}{4\bar{\eta}}\right)\sum_{t=2}^{T}\|\mathbf{x}_{t,i}-\mathbf{x}_{t-1,i}\|_2^2 \\
&\quad + \left(8\bar{\varepsilon}L^2D^4 + 8\bar{\eta}L^2D^2 - \frac{1}{4\bar{\varepsilon}}\right)\sum_{t=2}^{T}\|\boldsymbol{p}_t-\boldsymbol{p}_{t-1}\|_1^2 + \left(8\bar{\varepsilon}D^2L^2 + 8\bar{\eta}L^2 - \lambda\right)\sum_{t=2}^{T}\sum_{i=1}^{N} p_{t,i}\|\mathbf{x}_{t,i}-\mathbf{x}_{t-1,i}\|_2^2.
\end{aligned}
$$

By setting $\lambda = 2L$, $\bar{\eta} = \frac{1}{8L}$ and $\bar{\varepsilon} = \frac{1}{8D^2L}$, we can drop the last three non-positive terms and take expectations over the above formula to get that

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T}\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{u}_t\rangle\right] \\
&\le G\left(5\sqrt{D^2\ln N} + 2\sqrt{(D^2+2DP_T)}\right) + \left(5\sqrt{D^2\ln N} + 2\sqrt{(D^2+2DP_T)}\right)\left(2\sqrt{2}\sqrt{\sigma_{1:T}^2} + 2\sqrt{\Sigma_{1:T}^2}\right) \\
&\quad + (58\ln N + 16)D^2L + 32DLP_T + \frac{1}{L}G^2 \\
&= \mathcal{O}\left(P_T + \sqrt{(1+P_T)}\left(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2}\right)\right),
\end{aligned}
$$

which completes the proof. □

A.5. Useful Lemmas

Lemma 1 (variant of Bregman proximal inequality)**.** Assume $R_t(\cdot)$ is a α-strongly convex function with respect to $\|\cdot\|$, and denote by $\|\cdot\|_*$ the dual norm. Based on the updating rules of optimistic OMD in (6) and (7), for all $\mathbf{x} \in \mathcal{X}$ and $t \in [T]$, we have

$$
\begin{aligned}
\langle \mathbf{x}_t - \mathbf{x}, \nabla f_t(\mathbf{x}_t)\rangle \le\, & B_{R_t}(\mathbf{x}, \widehat{\mathbf{x}}_t) - B_{R_t}(\mathbf{x}, \widehat{\mathbf{x}}_{t+1}) \\
& + \frac{1}{\alpha}\|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_*^2 - \frac{\alpha}{2}\left(\|\mathbf{x}_t - \widehat{\mathbf{x}}_t\|^2 + \|\widehat{\mathbf{x}}_{t+1} - \mathbf{x}_t\|^2\right),
\end{aligned}
$$

where $\nabla f_0(\mathbf{x}_0) = 0$.

Proof. The above lemma can be extracted from previous studies (Nemirovski, 2005; Chiang et al., 2012). Here we provide its proof with the following lemma (Nemirovski, 2005, Lemma 3.1). Let us review the updating rules of optimistic OMD in (6) and (7):

$$
\mathbf{x}_t = \operatorname*{argmin}_{\mathbf{x}\in\mathcal{X}} \langle \nabla f_{t-1}(\mathbf{x}_{t-1}), \mathbf{x}\rangle + B_{\psi_t}(\mathbf{x}, \widehat{\mathbf{x}}_t), \tag{45}
$$

$$\widehat{\mathbf{x}}_{t+1} = \operatorname*{argmin}_{\mathbf{x}\in\mathcal{X}} \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}\rangle + B_{\psi_t}(\mathbf{x}, \widehat{\mathbf{x}}_t). \tag{46}$$

Notice that the updating rule (45) of $\mathbf{x}_t$ holds even for $t = 1$, since we define $\nabla f_0(\mathbf{x}_0) = 0$ and

$$\mathbf{x}_1 = \widehat{\mathbf{x}}_1 = \operatorname*{argmin}_{\mathbf{x}\in\mathcal{X}} B_{\psi_1}(\mathbf{x}, \widehat{\mathbf{x}}_1).$$

Then, based on the following mappings

$$\mathcal{Z} = \mathcal{E} = \mathcal{U} \leftarrow \mathcal{X},\ \mathbf{w} \leftarrow \mathbf{x}_t,\ \gamma \leftarrow 1,\ \boldsymbol{\xi} \leftarrow \nabla f_{t-1}(\mathbf{x}_{t-1}),\ \omega \leftarrow R_t,\ \mathbf{z}_- \leftarrow \widehat{\mathbf{x}}_t,$$
$$\mathbf{z}_+ \leftarrow \widehat{\mathbf{x}}_{t+1},\ \boldsymbol{\eta} \leftarrow \nabla f_t(\mathbf{x}_t),$$

we apply Lemma 2 to (45) and (46), and obtain Lemma 1. ☐

Lemma 2 (Lemma 3.1 of Nemirovski (2005))**.** Let $\mathcal{Z}$ be a convex compact set in Euclidean space $\mathcal{E}$ with inner product $\langle\cdot,\cdot\rangle$, let $\|\cdot\|$ be a norm on $\mathcal{E}$ and $\|\cdot\|_*$ be its dual norm, and let $\omega(\mathbf{z}) : \mathcal{Z} \mapsto \mathbb{R}$ be a α-strongly convex function with respect to $\|\cdot\|$, and $B_\omega(\mathbf{z}, \mathbf{w})$ be the Bregman distance associated with ω. Let $\mathcal{U}$ be a convex and closed subset of $\mathcal{Z}$, and let $\mathbf{z}_- \in \mathcal{Z}$, let $\boldsymbol{\xi}, \boldsymbol{\eta} \in \mathcal{E}$, and let $\gamma > 0$. Consider the points

$$\mathbf{w} = \operatorname*{argmin}_{\mathbf{y}\in\mathcal{U}} \big[\langle \gamma\boldsymbol{\xi}, \mathbf{y}\rangle + B_\omega(\mathbf{y}, \mathbf{z}_-)\big],$$
$$\mathbf{z}_+ = \operatorname*{argmin}_{\mathbf{y}\in\mathcal{U}} \big[\langle \gamma\boldsymbol{\eta}, \mathbf{y}\rangle + B_\omega(\mathbf{y}, \mathbf{z}_-)\big].$$

Then for all $\mathbf{z} \in \mathcal{U}$, one has

$$\langle \mathbf{w} - \mathbf{z}, \gamma\boldsymbol{\eta}\rangle \le B_\omega(\mathbf{z}, \mathbf{z}_-) - B_\omega(\mathbf{z}, \mathbf{z}_+) + \frac{\gamma^2}{\alpha}\|\boldsymbol{\eta} - \boldsymbol{\xi}\|_*^2 - \frac{\alpha}{2}\big(\|\mathbf{w} - \mathbf{z}_-\|^2 + \|\mathbf{z}_+ - \mathbf{w}\|^2\big)$$

and

$$\|\mathbf{w} - \mathbf{z}_+\| \le \alpha^{-1}\gamma\|\boldsymbol{\xi} - \boldsymbol{\eta}\|_*.$$

Lemma 3 (stability lemma)**.** By the same argument with Lemma 1, we have

$$\|\widehat{\mathbf{x}}_{t+1} - \mathbf{x}_t\| \le \alpha^{-1}\|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_*.$$

Proof. This conclusion is also proved by Lemma 2, which can be obtained from the proof of Lemma 1, and will not be repeated here. ☐

Lemma 4 (Boundness of the norm of gradient difference (Sachs et al. (2022), Analysis of Theorem 5))**.** Under Assumptions 4 and 1, we have

$$\begin{aligned}\|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \le& 4\|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 + 4\|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\ &+ 4L^2\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + 4\|\nabla F_{t-1}(\mathbf{x}_{t-1}) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2,\end{aligned}$$

where $\|\nabla f_1(\mathbf{x}_1) - \nabla f_0(\mathbf{x}_0)\|_2^2 = \|\nabla f_1(\mathbf{x}_1)\|_2^2 \le G^2$.

Proof. For $t \ge 2$, from Jensen's inequality and Assumption 4 (smoothness of expected function), we have

$$\begin{aligned}&\|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\ =&16\bigg\|\frac{1}{4}\big[\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\big] + \frac{1}{4}\big[\nabla F_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_{t-1})\big] \\ &+\frac{1}{4}\big[\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\big] + \frac{1}{4}\big[\nabla F_{t-1}(\mathbf{x}_{t-1}) - \nabla f_{t-1}(\mathbf{x}_{t-1})\big]\bigg\|_2^2 \\ \le&4\|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 + 4\|\nabla F_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_{t-1})\|_2^2 \\ &+ 4\|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 + 4\|\nabla F_{t-1}(\mathbf{x}_{t-1}) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\ \le&4\|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 + 4L^2\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \\ &+ 4\|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 + 4\|\nabla F_{t-1}(\mathbf{x}_{t-1}) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2.\end{aligned} \tag{47}$$

For $t = 1$, from Assumption 1 (boundness of the gradient norm), we have

$$\|\nabla f_1(\mathbf{x}_1) - \nabla f_0(\mathbf{x}_0)\|_2^2 = \|\nabla f_1(\mathbf{x}_1)\|_2^2 \leq G^2. \tag{48}$$

□

Lemma 5 (Boundness of cumulative norm of gradient difference (Sachs et al. (2022), Analysis of Theorem 5))**.** Under Assumptions 4 and 1, we have

$$\begin{aligned}\sum_{t=1}^{T} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \leq & G^2 + 4L^2 \sum_{t=2}^{T} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \\ & + 8 \sum_{t=1}^{T} \|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 + 4 \sum_{t=2}^{T} \|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2.\end{aligned} \tag{49}$$

Proof. It is easy to verify the above lemma by substituting (47) and (48) in Lemma 4 into $\sum_{t=1}^{T} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2$ and simplifying the result. □

Lemma 6. Under Assumption 1 and 6, we have

$$\sum_{t=1}^{T} \frac{1}{\lambda t} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \leq \frac{4G^2}{\lambda} \ln \left(\sum_{t=1}^{T} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 + 1 \right) + \frac{4G^2 + 1}{\lambda}.$$

Proof. Define

$$\alpha = \left\lceil \sum_{t=1}^{T} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \right\rceil.$$

Since α is obtained by rounding up the term $\sum_{t=1}^{T} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2$, we have

$$\begin{aligned}& \sum_{t=1}^{T} \frac{1}{\lambda t} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\ = & \sum_{t=1}^{\alpha} \frac{1}{\lambda t} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 + \sum_{t=\alpha+1}^{T} \frac{1}{\lambda t} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\ \leq & \frac{4G^2}{\lambda} \sum_{t=1}^{\alpha} \frac{1}{t} + \frac{1}{\lambda(\alpha+1)} \sum_{t=\alpha+1}^{T} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\ \leq & \frac{4G^2}{\lambda} \left(1 + \int_{t=1}^{\alpha} \frac{1}{t} dt\right) + \frac{1}{\lambda} \leq \frac{4G^2}{\lambda} (\ln \alpha + 1) + \frac{1}{\lambda} \\ \leq & \frac{4G^2}{\lambda} \ln \left(\sum_{t=1}^{T} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 + 1 \right) + \frac{4G^2 + 1}{\lambda}.\end{aligned}$$

□

Lemma 7. Let A_T be a non-negative term, a, b be non-negative constants and c be a positive constant, then we have

$$a \ln(bA_T + 1) - cA_T \leq a \ln \left(\frac{ab}{c} + 1 \right).$$

Proof. We use the following inequality to prove the lemma:

$$\ln p \leq \frac{p}{q} + \ln q + -1, \forall p > 0, \ q > 0.$$

By setting $p = bA_T + 1$ and $q = \frac{ab}{c} + 1$, we obtain

$$\begin{aligned} a\ln(bA_T+1) - cA_T &\le a\left(\frac{bA_T+1}{ab/c+1} + \ln\left(\frac{ab}{c}+1\right) - 1\right) - cA_T \\ &= c\left(\frac{ab}{ab+c} - 1\right)A_T + \left(\frac{1}{ab/c+1} - 1\right)a + a\ln\left(\frac{ab}{c}+1\right) \\ &\le a\ln\left(\frac{ab}{c}+1\right). \end{aligned}$$

□

B. Omitted Details for Section 5

In this section, we present the missing proofs of corollaries in Section 5, including proof of Corollary 1 for the adversarially corrupted stochastic model in Appendix B.1 and proof of Corollary 2 for the ROM model in Appendix B.2.

B.1. Proof of Corollary 1

Proof. We have introduced in Section 5.3 that the loss functions in adversarially corrupted stochastic model satisfy that

$$f_t(\mathbf{x}) = h_t(\mathbf{x}) + c_t(\mathbf{x}) \tag{50}$$

for all $t \in [T]$, where $h_t(\cdot)$ is sampled from a fixed distribution every iteration and $\sum_{t=1}^T \max_{\mathbf{x}\in\mathcal{X}} \|\nabla c_t(\mathbf{x})\| \le C_T$ ($C_T > 0$ is a corruption level parameter). Then according to the definition of $F_t(\mathbf{x})$, we have

$$F_t(\mathbf{x}) = \mathbb{E}_{f_t\sim\mathcal{D}_t}[f_t(\mathbf{x})] = \mathbb{E}_{h_t\sim\mathcal{D}}[h_t(\mathbf{x}_t) + c_t(\mathbf{x})] = \mathbb{E}_{h_t\sim\mathcal{D}}[h_t(\mathbf{x}_t)] + c_t(\mathbf{x}). \tag{51}$$

Since $h_t(\cdot)$ is i.i.d, the expectations of $h_t(\cdot)$ for each t are the same. Refer to Sachs et al. (2022), we have

$$\begin{aligned} \|\nabla F_t(\mathbf{x}) - \nabla F_{t-1}(\mathbf{x})\|_2^2 \le 2G\|\nabla F_t(\mathbf{x}) - \nabla F_{t-1}(\mathbf{x})\|_2 &\overset{(51)}{=} 2G\|\nabla c_t(\mathbf{x}) - \nabla c_{t-1}(\mathbf{x})\| \\ &\le 2G(\|\nabla c_t(\mathbf{x})\| + \|\nabla c_{t-1}(\mathbf{x})\|). \end{aligned}$$

Because $\Sigma_{1:T}^2 = \mathbb{E}\left[\sum_{t=2}^T \sup_{\mathbf{x}\in\mathcal{X}} \|\nabla F_t(\mathbf{x}) - \nabla F_{t-1}(\mathbf{x})\|_2^2\right]$ (see (5)), we can substitute the above formula into it and get that

$$\Sigma_{1:T}^2 \le \sum_{t=2}^T \sup_{\mathbf{x}\in\mathcal{X}} 2G(\|\nabla c_t(\mathbf{x})\| + \|\nabla c_{t-1}(\mathbf{x})\|) \le 4GC_T. \tag{52}$$

Besides, due to that

$$\sigma_t^2 = \max_{\mathbf{x}\in\mathcal{X}} \mathbb{E}_{f_t\sim\mathcal{D}_t}[\|\nabla f_t(\mathbf{x}) - \nabla F_t(\mathbf{x})\|_2^2] \overset{(50),(51)}{=} \max_{\mathbf{x}\in\mathcal{X}} \mathbb{E}_{h_t\sim\mathcal{D}_t}[\|\nabla h_t(\mathbf{x}) - \nabla \mathbb{E}_{h_t\sim\mathcal{D}}[h_t(\mathbf{x}_t)]\|_2^2] = \sigma,$$

where σ denotes the variance of the stochastic gradients, we have

$$\sigma_{1:T}^2 = \mathbb{E}\left[\sum_{t=1}^T \sigma_t^2\right] = \sigma T. \tag{53}$$

We complete the proof by integrating the bound of $\sigma_{1:T}^2$ and $\Sigma_{1:T}^2$ (see (53), (52)) with the regret bounds of optimistic OMD in Theorem 3 and Theorem 5. □

B.2. Proof of Corollary 2

Proof. The difference between ROM and i.i.d. stochastic model is that ROM samples a loss from the loss set without replacement in each round, while i.i.d. stochastic model samples independently and uniformly with replacement in each round. However, following Sachs et al. (2022), we can bound the variance of ROM with respect to $\mathcal{D}_t$ for each t by the variance σ_1^2 of the first round, which can also be regarded as the variance of the i.i.d. model for every round. Specifically, for $\forall \mathbf{x} \in \mathcal{X}$ and every $t \in [T]$, we have

$$\mathbb{E}_{f_t \sim \mathcal{D}_t}\left[\|\nabla f_t(\mathbf{x}) - \nabla F_t(\mathbf{x})\|_2^2\right] \leq \mathbb{E}_{f_t \sim \mathcal{D}_t}\left[\|\nabla f_t(\mathbf{x}) - \nabla F_1(\mathbf{x})\|_2^2\right]. \tag{54}$$

Since ROM samples losses without replacement, let set Γ_t represent the index set of losses that can be selected in the tth round, thus $\Gamma_1 = [T]$, then we have

$$\begin{aligned}\mathbb{E}_{f_t \sim \mathcal{D}_t}\left[\|\nabla f_t(\mathbf{x}) - \nabla F_1(\mathbf{x})\|_2^2\right] &= \frac{1}{T-(t-1)} \sum_{i \in \Gamma_t} \|\nabla f_i(\mathbf{x}) - \nabla F_1(\mathbf{x})\|_2^2 \\ &\leq \frac{1}{T-(t-1)} \sum_{i \in \Gamma_1} \|\nabla f_i(\mathbf{x}) - \nabla F_1(\mathbf{x})\|_2^2 \leq \frac{T}{T-(t-1)} \sigma_1^2.\end{aligned}$$

So combining (54) with the above inequality, we get that

$$\mathbb{E}_{f_t \sim \mathcal{D}_t}\left[\|\nabla f_t(\mathbf{x}) - \nabla F_t(\mathbf{x})\|_2^2\right] \leq \frac{T}{T-(t-1)} \sigma_1^2, \ \forall \mathbf{x} \in \mathcal{X}, \, t \in [T]. \tag{55}$$

Besides, from (54), we can also get that

$$\begin{aligned}\mathbb{E}\left[\sigma_t^2\right] &\leq \mathbb{E}\left[\max_{\mathbf{x} \in \mathcal{X}} \mathbb{E}_{f_t \sim \mathcal{D}_t}\left[\|\nabla f_t(\mathbf{x}) - \nabla F_1(\mathbf{x})\|_2^2\right]\right] \\ &\leq \mathbb{E}\left[\mathbb{E}_{f_t \sim \mathcal{D}_t}\left[\max_{\mathbf{x} \in \mathcal{X}} \|\nabla f_t(\mathbf{x}) - \nabla F_1(\mathbf{x})\|_2^2\right]\right] = \widetilde{\sigma}_1^2,\end{aligned} \tag{56}$$

where we review the definition that $\widetilde{\sigma}_1^2 = \frac{1}{T}\sum_{t=1}^T \max_{\mathbf{x} \in \mathcal{X}} \|\nabla f_t(\mathbf{x}) - \bar{\nabla}_T(\mathbf{x})\|_2^2$. Then, we use a technique in Sachs et al. (2022) by introduce a variable $\tau \in [T]$, which help us upper bound $\sigma_{1:T}^2$ as

$$\begin{aligned}\sigma_{1:T}^2 = \mathbb{E}\left[\sum_{t=1}^T \sigma_t^2\right] &\leq \mathbb{E}\left[\sum_{t=1}^\tau \sigma_t^2\right] + \mathbb{E}\left[\sum_{t=\tau+1}^T \sigma_t^2\right] \\ &\overset{(55),(56)}{\leq} \sum_{t=1}^\tau \frac{T}{T-(t-1)} \sigma_1^2 + (T-\tau)\widetilde{\sigma}_1^2 \\ &\leq \sum_{n=T-(\tau-1)}^T \frac{1}{n} T\sigma_1^2 + (T-\tau)\widetilde{\sigma}_1^2 \\ &\leq \left(\int_{t=T-(\tau-1)}^T \frac{1}{t} dt + \frac{1}{T-(\tau-1)}\right) T\sigma_1^2 + (T-\tau)\widetilde{\sigma}_1^2 \\ &\leq \left(1 + \log \frac{T}{T-(\tau-1)}\right) T\sigma_1^2 + (T-\tau)\widetilde{\sigma}_1^2.\end{aligned}$$

If $T\sigma_1^2/\widetilde{\sigma}_1^2 > 2$, we set $\tau = T - \lfloor T\sigma_1^2/\widetilde{\sigma}_1^2 \rfloor$, then we have

$$\begin{aligned}\sigma_{1:T}^2 &\leq \left(1 + \log \frac{T}{T-(\tau-1)}\right) T\sigma_1^2 + (T-\tau)\widetilde{\sigma}_1^2 \\ &= \left(1 + \log \frac{T}{\lfloor T\sigma_1^2/\widetilde{\sigma}_1^2 \rfloor + 1}\right) T\sigma_1^2 + \lfloor T\sigma_1^2/\widetilde{\sigma}_1^2 \rfloor \widetilde{\sigma}_1^2 \\ &\leq \left(1 + \log \frac{T}{\lfloor T\sigma_1^2/\widetilde{\sigma}_1^2 \rfloor}\right) T\sigma_1^2 + T\sigma_1^2 \\ &\leq \left(1 + \log \frac{1}{\sigma_1^2/\widetilde{\sigma}_1^2 - 1/T}\right) T\sigma_1^2 + T\sigma_1^2\end{aligned}$$

$$\leq \left(1 + \log \frac{2\widetilde{\sigma}_1^2}{\sigma_1^2}\right) T\sigma_1^2 + T\sigma_1^2$$
$$\leq T\sigma_1^2 \log\left(\frac{2e^2\widetilde{\sigma}_1^2}{\sigma_1^2}\right).$$

Otherwise, if $T\sigma_1^2/\widetilde{\sigma}_1^2 \leq 2$, we set $\tau = T$, then we can get the regret bound of $\mathcal{O}(T\sigma_1^2(1+\log T))$. Since we have

$$\mathcal{O}(T\sigma_1^2(1+\log T)) \leq \mathcal{O}(T\sigma_1^2(1+\log(2\widetilde{\sigma}_1^2/\sigma_1^2))) \leq \mathcal{O}(T\sigma_1^2\log(2e^2\widetilde{\sigma}_1^2/\sigma_1^2)),$$

then the final bound of $\sigma_{1:T}^2$ is of order $\mathcal{O}\left(T\sigma_1^2 \log\left(\frac{2e^2\widetilde{\sigma}_1^2}{\sigma_1^2}\right)\right)$.

Next, we try to bound $\Sigma_{1:T}^2$. We suppose that $k_t = \Gamma_t \backslash \Gamma_{t+1}$ represents the loss selected in round t, then we have

$$\begin{aligned}
\|\nabla F_t(\mathbf{x}) - \nabla F_{t-1}(\mathbf{x})\|_2^2 &= \left\| \frac{1}{T-(t-1)} \sum_{i\in\Gamma_t} \nabla f_i(\mathbf{x}) - \frac{1}{T-(t-2)} \sum_{i\in\Gamma_{t-1}} \nabla f_i(\mathbf{x}) \right\|_2^2 \\
&= \left\| \frac{(T-t+2)-(T-t+1)}{(T-t+1)(T-t+2)} \sum_{i\in\Gamma_t} \nabla f_i(\mathbf{x}) - \frac{1}{T-t+2} \nabla f_{k_{t-1}}(\mathbf{x}) \right\|_2^2 \\
&\leq \frac{2}{(T-t+2)^2} \left\| \frac{1}{T-t+1} \sum_{i\in\Gamma_t} \nabla f_i(\mathbf{x}) \right\|_2^2 + \frac{2}{(T-t+2)^2} \|\nabla f_{k_{t-1}}(\mathbf{x})\|_2^2 \\
&\leq \frac{4G^2}{(T-t+2)^2},
\end{aligned}$$

where the last inequality is derived from Assumption 1 (boundness of the gradient norm).

Summing the above inequality over $t = 1, ..., T$, and taking the expectation give

$$\Sigma_{1:T}^2 = \mathbb{E}\left[\sum_{t=1}^T \sup_{\mathbf{x}\in\mathcal{X}} \|\nabla F_t(\mathbf{x}) - \nabla F_{t-1}(\mathbf{x})\|_2^2\right] \leq \sum_{t=1}^T \frac{4G^2}{(T-t+2)^2} \leq 8G^2.$$

Finally, we substitute the bound of $\sigma_{1:T}^2$ and $\Sigma_{1:T}$ into Theorem 1, which is for convex and smooth functions, and complete the proof. □

C. Technical Lemmas

Lemma 8 (self-confident tuning (Auer et al., 2002, Lemma 3.5))**.** Let $l_1, \ldots, l_T$ and δ be non-negative real numbers. Then

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

where we define $0/\sqrt{0} = 0$ for simplicity.

Lemma 9 (Lemma 12 of Hazan et al. (2007))**.** Let $A \succeq B \succ 0$ be positive definite matrices. Then $\langle A^{-1}, A - B\rangle \leq \ln \frac{|A|}{|B|}$, where $|A|$ denotes the determinant of matrix A.

Lemma 10. Let $\mathbf{u}_t \in \mathbb{R}^d$ $(t = 1, ..., T)$, be a sequence of vectors. Define $S_t = \sum_{\tau=1}^t \mathbf{u}_\tau \mathbf{u}_\tau^\top + \varepsilon I$, where $\varepsilon > 0$. Then

$$\sum_{t=1}^T \mathbf{u}_t^\top S_t^{-1} \mathbf{u}_t \leq d \ln\left(1 + \frac{\sum_{t=1}^T \|\mathbf{u}_t\|_2^2}{d\varepsilon}\right).$$

Proof. Using Lemma 9, we have $\langle A^{-1}, A - B\rangle \leq \ln \frac{|A|}{|B|}$ for any two positive definite matrices $A \succeq B \succ 0$. Then, following the argument of Luo et al. (2016, Theorem 2), we have

$$\begin{aligned}
&\sum_{t=1}^{T} \mathbf{u}_t^\top S_t^{-1} \mathbf{u}_t = \sum_{t=1}^{T} \langle S_t^{-1}, \mathbf{u}_t \mathbf{u}_t^\top \rangle = \sum_{t=1}^{T} \langle S_t^{-1}, S_t - S_{t-1} \rangle \\
\leq & \sum_{t=1}^{T} \ln \frac{|S_t|}{|S_{t-1}|} = \ln \frac{|S_T|}{|S_0|} = \sum_{i=1}^{d} \ln \left(1 + \frac{\lambda_i(\sum_{t=1}^{T} \mathbf{u}_t \mathbf{u}_t^\top)}{\varepsilon} \right) \\
= & d \sum_{i=1}^{d} \frac{1}{d} \ln \left(1 + \frac{\lambda_i(\sum_{t=1}^{T} \mathbf{u}_t \mathbf{u}_t^\top)}{\varepsilon} \right) \\
\leq & d \ln \left(1 + \frac{\sum_{i=1}^{d} \lambda_i(\sum_{t=1}^{T} \mathbf{u}_t \mathbf{u}_t^\top)}{d\varepsilon} \right) = d \ln \left(1 + \frac{\sum_{t=1}^{T} \|\mathbf{u}_t\|_2^2}{d\varepsilon} \right),
\end{aligned}$$

where the last inequality is due to Jensen's inequality. □

Lemma 11 (standard analysis of optimistic FTRL (Orabona, 2019, Theorem 7.35))**.** Let $V \in \mathbb{R}^d$ be convex, closed, and non-empty. Denote by $G_t(\mathbf{x}) = \Psi_t(\mathbf{x}) + \sum_{s=1}^{t-1} \ell_s(\mathbf{x})$. Assume for $t = 1, \cdots, T$ that G_t is proper and λ_t-strongly convex w.r.t. $\|\cdot\|$, ℓ_t and $\widetilde{\ell}_t$ proper and convex ($\widetilde{\ell}_t$ is the predicted next loss), and int dom $G_t \cap V \neq \{\}$. Also, assume that $\partial \ell_t(\mathbf{x}_t)$ and $\partial \widetilde{\ell}_t(\mathbf{x}_t)$ are non-empty. Then there exists $\widetilde{\mathbf{g}}_t \in \partial \widetilde{\ell}_t(\mathbf{x}_t)$ for $t = 1, \cdots, T$ such that

$$\begin{aligned}
&\sum_{t=1}^{T} \ell_t(\mathbf{x}_t) - \sum_{t=1}^{T} \ell_t(\mathbf{x}) \\
\leq\ & \Psi_{T+1}(\mathbf{x}) - \Psi_1(\mathbf{x}_1) + \sum_{t=1}^{T} \left(\langle \mathbf{g}_t - \widetilde{\mathbf{g}}_t, \mathbf{x}_t - \mathbf{x}_{t+1} \rangle - \frac{\lambda_t}{2} \|\mathbf{x}_t - \mathbf{x}_{t+1}\|^2 + \Psi_t(\mathbf{x}_{t+1}) - \Psi_{t+1}(\mathbf{x}_{t+1}) \right)
\end{aligned}$$

for all $\mathbf{g}_t \in \partial \ell_t(\mathbf{x}_t)$.

Lemma 12 (Lemma 13 of Zhao et al. (2021))**.** Let $a_1, a_2, \cdots, a_T, b$ and $\bar{c}$ be non-negative real numbers and $a_t \in [0, B]$ for any $t \in [T]$. Let the step size be

$$c_t = \min \left\{ \bar{c}, \sqrt{\frac{b}{\sum_{s=1}^{t} a_s}} \right\} \text{ and } c_0 = \bar{c}.$$

Then, we have

$$\sum_{t=1}^{T} c_{t-1} a_t \leq 2\bar{c}B + 4\sqrt{b \sum_{t=1}^{T} a_t}.$$