

---

# Online Composite Optimization Between Stochastic and Adversarial Environments

---

Yibo Wang<sup>1,2</sup>, Sijia Chen<sup>1,2</sup>, Wei Jiang<sup>1</sup>, Wenhao Yang<sup>1,2</sup>, Yuanyu Wan<sup>3,1</sup>, Lijun Zhang<sup>1,2,\*</sup>

<sup>1</sup>National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China

<sup>2</sup>School of Artificial Intelligence, Nanjing University, Nanjing, China

<sup>3</sup>School of Software Technology, Zhejiang University, Ningbo, China

{wangyb, chensj, jiangw, yangwh, zhanglj}@lamda.nju.edu.cn, wanyy@zju.edu.cn

## Abstract

We study online composite optimization under the Stochastically Extended Adversarial (SEA) model. Specifically, each loss function consists of two parts: a fixed non-smooth and convex regularizer, and a time-varying function which can be chosen either stochastically, adversarially, or in a manner that interpolates between the two extremes. In this setting, we show that for smooth and convex time-varying functions, optimistic composite mirror descent (OptCMD) can obtain an  $\mathcal{O}(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})$  regret bound, where  $\sigma_{1:T}^2$  and  $\Sigma_{1:T}^2$  denote the cumulative stochastic variance and the cumulative adversarial variation of time-varying functions, respectively. For smooth and strongly convex time-varying functions, we establish an  $\mathcal{O}((\sigma_{\max}^2 + \Sigma_{\max}^2) \log(\sigma_{1:T}^2 + \Sigma_{1:T}^2))$  regret bound, where  $\sigma_{\max}^2$  and  $\Sigma_{\max}^2$  denote the maximal stochastic variance and the maximal adversarial variation, respectively. For smooth and exp-concave time-varying functions, we achieve an  $\mathcal{O}(d \log(\sigma_{1:T}^2 + \Sigma_{1:T}^2))$  bound where  $d$  denotes the dimensionality. Moreover, to deal with the unknown function type in practical problems, we propose a multi-level *universal* algorithm that is able to achieve the desirable bounds for three types of time-varying functions simultaneously. It should be noticed that all our findings match existing bounds for the SEA model without the regularizer, which implies that there is *no price* in regret bounds for the benefits gained from the regularizer.

## 1 Introduction

Online composite optimization has drawn considerable attention in recent years [Duchi and Singer, 2009, Ghadimi and Lan, 2012, Lei and Zhou, 2017, Scroccaro et al., 2023]. Formally, it can be viewed as an iterative game between a learner and the environment. In each round  $t$ , the learner makes a decision  $\mathbf{x}_t$  from a convex set  $\mathcal{X} \subseteq \mathbb{R}^d$  and then suffers a loss  $\phi_t(\mathbf{x}_t)$  in the form of

$$\phi_t(\mathbf{x}) = f_t(\mathbf{x}) + r(\mathbf{x}), \quad (1)$$

where  $f_t(\cdot) : \mathcal{X} \rightarrow \mathbb{R}$  denotes the time-varying function that is chosen by the environment, and  $r(\cdot) : \mathcal{X} \rightarrow \mathbb{R}$  denotes the fixed non-smooth and convex regularizer, such as the  $\ell_1$ -norm for sparse vectors and the trace norm for low-rank matrices [Langford et al., 2009, Flammarion and Bach, 2017, Zhang et al., 2019, Garber and Kaplan, 2019]. In the literature, online composite optimization is generally divided into two categories: stochastic composite optimization [Lan, 2012, 2016, Zhang et al., 2017, Lei and Tang, 2018, Lei et al., 2019, Kulunchakov and Mairal, 2019] and adversarial composite optimization [Xiao, 2009, Duchi et al., 2010, Mohri and Yang, 2016, Joulani et al., 2020, Yang et al., 2024c]. In the former one,  $f_t(\cdot)$  is assumed to be independent and identically distributed (i.i.d.) over time; in the latter one,  $f_t(\cdot)$  can be chosen arbitrarily or even adversarially. However,

---

\*Lijun Zhang is the corresponding author.

the environment in real-world scenarios is seldom purely stochastic or adversarial, but rather falls somewhere in between [Amir et al., 2020, Garber et al., 2020, Sherman et al., 2021, Zimmert and Seldin, 2021, Ito, 2021], and our understanding for the more common intermediate scenarios remains limited in online composite optimization.

Recently, Sachs et al. [2022] introduce the Stochastically Extended Adversarial (SEA) model as an intermediate setting in online optimization without regularizer, i.e.,  $r(\mathbf{x}) = 0$ . In the SEA model, loss functions are not restricted to fully i.i.d. or adversarial; instead, the environment selects them from any intermediate state between the two extreme settings. To reflect how stochastic or adversarial the environments are, they introduce the cumulative stochastic variance  $\sigma_{1:T}^2$  and the cumulative adversarial variation  $\Sigma_{1:T}^2$ , as shown below:

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

where  $\sigma_t^2 = \sup_{\mathbf{x} \in \mathcal{X}} \mathbb{E}_{f_t \sim \mathcal{D}_t} [\|\nabla f_t(\mathbf{x}) - \nabla F_t(\mathbf{x})\|_2^2]$  denotes the variance of gradients and  $F_t(\mathbf{x}) = \mathbb{E}_{f_t \sim \mathcal{D}_t} [f_t(\mathbf{x})]$  denotes the expected function for  $f_t(\cdot)$ . For the SEA model, two classical algorithms in optimistic online learning [Rakhlin and Sridharan, 2013]—optimistic Follow-The-Regularized Leader (FTRL) [Sachs et al., 2022] and optimistic Online Mirror Descent (OMD) [Chen et al., 2023]—have been proven ensuring sublinear regret bounds with respect to both  $\sigma_{1:T}^2$  and  $\Sigma_{1:T}^2$ .

In this paper, we extend the SEA model into online composite optimization, termed as composite SEA, to bridge the gap between stochastic and adversarial composite optimization. Specifically, in (1),  $r(\cdot)$  remains fixed over time and  $f_t(\cdot)$  is selected from a distribution  $\mathcal{D}_t$ , which is chosen by the environments either stochastically, adversarially, or in a manner that interpolates between the two extremes. The goal of composite SEA is to minimize the expected regret in terms of (1):

$$\mathbb{E} [\text{Regret}_T] = \mathbb{E} \left[ \sum_{t=1}^T [f_t(\mathbf{x}_t) + r(\mathbf{x}_t)] - \min_{\mathbf{x} \in \mathcal{X}} \sum_{t=1}^T [f_t(\mathbf{x}) + r(\mathbf{x})] \right], \quad (3)$$

which benchmarks the cumulative composite loss of the learner and that of the best fixed decision. To handle the composite setting, a natural impulse is to treat  $f_t(\cdot) + r(\cdot)$  as one function and directly apply existing methods for the SEA model, i.e., optimistic FTRL and optimistic OMD. However, such a straightforward application is unsuitable because (i) these methods heavily rely on the smoothness of loss functions, but due to the non-smooth component  $r(\cdot)$ , the summation function  $\phi_t(\cdot) = f_t(\cdot) + r(\cdot)$  loses this crucial property; (ii) directly applying these methods ignores the presence of  $r(\cdot)$  and thus fails to gain the benefits from the regularizer, e.g., the sparsity induced by the  $\ell_1$ -norm. Therefore, a natural question arises *whether it is possible to deal with both the intermediate nature of environments and the composite structure of loss functions concurrently*.

We affirmatively answer the above question by revisiting a variant of optimistic OMD [Scroccaro et al., 2023], named Optimistic Composite Mirror Descent (OptCMD). This method inherits the idea of optimistic online learning [Rakhlin and Sridharan, 2013], i.e., exploiting estimates on upcoming loss functions for decision updates, and can disentangle  $f_t(\cdot)$  and  $r(\cdot)$  during the optimization, thereby effectively leveraging distinct properties of each. However, OptCMD is originally designed for adversarial composite optimization, and a simple extension will lead to unsatisfactory bounds in composite SEA. In this paper, we reanalyze OptCMD and show that with suitable configurations, OptCMD attains  $\mathcal{O}(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})$ ,  $\mathcal{O}((\sigma_{\max}^2 + \Sigma_{\max}^2) \log(\sigma_{1:T}^2 + \Sigma_{1:T}^2))$  and  $\mathcal{O}(d \log(\sigma_{1:T}^2 + \Sigma_{1:T}^2))$  bounds for smooth and general convex, smooth and strongly convex, and smooth and exp-concave time-varying functions, respectively. Our findings generalize previous results in stochastic and adversarial composite optimization, and can reduce to them by specializing  $\sigma_{1:T}^2$  and  $\Sigma_{1:T}^2$ . Moreover, our results coincide with existing bounds for the SEA model [Sachs et al., 2022, Chen et al., 2023], which indicates that there is *no price* in regret bounds for the benefits from  $r(\cdot)$ .

One concern of OptCMD is the requirement of the function type in advance, which is often impractical in real-world problems and motivates us to design a *universal* algorithm that is agnostic to the prior knowledge about loss functions. We note that recently, Yan et al. [2023] have introduced a universal algorithm for the SEA model without regularizer, based on the meta-expert structure in online learning [van Erven and Koolen, 2016, Mhammedi et al., 2019, Wang et al., 2019, Zhang et al., 2022]. However, their method [Yan et al., 2023] cannot naturally support the composite setting, as it highly depends on the smoothness of the (summation) loss functions, which does not necessarily hold in the composite SEA model. In this paper, we propose a *novel* universal algorithm for composite SEA,

based on the observations that (i) the regularizer remains fixed over time and is accessible to the meta-algorithm from the beginning round; (ii) the meta-algorithm is able to obtain expert decisions for the current round before estimating their performance. With these two facts, we can utilize the information about  $r(\cdot)$  to track experts, ensuring the meta-regret will not deviate from that of the best one. Specifically, inspired by Yan et al. [2023], we employ a two-layer Multi-scale Multiplicative-weight with Correction (MsMwC) [Chen et al., 2021] as the meta-algorithm, and choose OptCMD as the expert-algorithm. To effectively estimate the expert performance, we first collect decisions of each expert for the current round, and then explicitly integrate them and expert performance on  $r(\cdot)$  into the losses and optimisms used in MsMwC. Theoretical analysis demonstrates that, with the proposed integration, our universal algorithm can achieve desirable bounds for three types of loss functions simultaneously. We summarize our contributions as shown below.

- We first introduce the composite SEA model, which serves as an intermediate setting between stochastic and adversarial composite optimization and can naturally adapt to two extreme settings;
- For the composite SEA model, we demonstrate that OptCMD can attain the regret bounds of  $\mathcal{O}(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})$ ,  $\mathcal{O}((\sigma_{\max}^2 + \Sigma_{\max}^2) \log(\sigma_{1:T}^2 + \Sigma_{1:T}^2))$  and  $\mathcal{O}(d \log(\sigma_{1:T}^2 + \Sigma_{1:T}^2))$  for smooth and general convex, smooth and strongly convex, and smooth and exp-concave time-varying functions, respectively. Owing to the versatility of  $\sigma_{1:T}^2$  and  $\Sigma_{1:T}^2$ , these bounds can *recover* previous results in stochastic and adversarial composite optimization;
- Moreover, to handle the unknown function type, we propose a *new* universal algorithm in composite SEA and show that it can concurrently achieve desirable bounds for three kinds of functions;
- Finally, we discuss implications of our theoretical findings for two common intermediate examples with composite loss functions, and derive favorable bounds specific to these practical scenarios.

## 2 Related work

In this section, we briefly review related work on adversarial and stochastic composite optimization, the SEA model and universal online learning.

**Adversarial composite optimization.** In this setting,  $f_t(\cdot)$  is assumed to be chosen by the environments arbitrarily or even adversarially, and the performance is measured by *regret* shown in brackets of (3). In the literature, the seminal work [Duchi and Singer, 2009] proposes the FOBOS method with  $\mathcal{O}(\sqrt{T})$  and  $\mathcal{O}(\lambda^{-1} \log T)$  regret bounds for general convex and  $\lambda$ -strongly convex  $f_t(\cdot)$ , respectively. Later, Xiao [2009] proposes the RDA method based on the primal-dual subgradient framework [Nesterov, 2009], and shows that RDA is able to generate sparser decisions than FOBOS when  $r(\mathbf{x}) = \mu \|\mathbf{x}\|_1$  for some  $\mu > 0$ . In the same time, another subsequent work [Duchi et al., 2010] introduces the COMID method under the mirror descent framework [Beck and Teboulle, 2003], and achieves the same bounds as FOBOS. Recently, for  $\alpha$ -exp-concave  $f_t(\cdot)$ , Yang et al. [2024c] establish an  $\mathcal{O}((d/\alpha) \log T)$  regret bound by proposing the ProxONS method.

Besides, there exist other powerful methods [Yang et al., 2014, Mohri and Yang, 2016, Joulani et al., 2020, Scroccaro et al., 2023] equipped with problem-dependent bounds. These bounds can safeguard the above results in the worst-case scenarios and become tighter when environments have special properties, such as smoothness. Among them, the most related work is by Scroccaro et al. [2023], who develop OptCMD based on the optimistic online learning framework [Rakhlin and Sridharan, 2013]. The key idea is to make an estimate for the upcoming loss function, which ensures tighter bounds when the estimate is accurate and still maintains the worst-case bound otherwise. By utilizing the smoothness of  $f_t(\cdot)$ , OptCMD achieves  $\mathcal{O}(\sqrt{V_T})$  and  $\mathcal{O}(\lambda^{-1} \log V_T)$  regret bounds for general convex and  $\lambda$ -strongly convex  $f_t(\cdot)$ , respectively, where  $V_T = \sum_{t=1}^T \max_{\mathbf{x} \in \mathcal{X}} \|\nabla f_t(\mathbf{x}) - \nabla f_{t-1}(\mathbf{x})\|_2^2$  denotes the gradient-variation and can be small when  $f_t(\cdot)$  gradually changes.

**Stochastic composite optimization.** In this setting,  $f_t(\cdot)$  is assumed to be i.i.d. sampled from a fixed distribution  $\mathcal{D}$ , and the goal is to minimize the composite objective:  $\min_{\mathbf{x} \in \mathcal{X}} \mathbb{E}_{f \sim \mathcal{D}} [f(\mathbf{x})] + r(\mathbf{x})$ . The performance is measured by *excess risk*, which compares the solution with the optimal one, i.e.,  $\mathbb{E}_{f \sim \mathcal{D}} [f(\mathbf{x}_T)] + r(\mathbf{x}_T) - \min_{\mathbf{x} \in \mathcal{X}} \{\mathbb{E}_{f \sim \mathcal{D}} [f(\mathbf{x})] + r(\mathbf{x})\}$ .

In the literature, there are a substantial body of methods designed for the stochastic setting [Ghadimi and Lan, 2012, Lan, 2012, 2016, Lei and Tang, 2018, Lei et al., 2019], and by utilizing the online-to-batch conversion technique [Cesa-Bianchi et al., 2004], existing methods for adversarial composite optimization [Xiao, 2009, Duchi et al., 2010, Yang et al., 2024c] can also be extended to the stochastic

scenarios. Specifically, Lan [2012] first establishes an  $\mathcal{O}(1/\sqrt{T})$  excess risk bound for general convex objective, and then Lan [2016] improves the rate to  $\mathcal{O}(1/T)$  for strongly convex objective. By utilizing the online-to-batch conversion, both RDA [Xiao, 2009] and COMID [Duchi et al., 2010] achieve  $\mathcal{O}(1/\sqrt{T})$  and  $\mathcal{O}(\log T/(\lambda T))$  excess risks for general convex and  $\lambda$ -strongly convex objective, respectively. When  $f(\cdot)$  is  $\alpha$ -exp-concave, ProxONS [Yang et al., 2024c] ensures an  $\mathcal{O}(d \log T/(\alpha T))$  excess bound.

**SEA model.** The SEA model is originally introduced by Sachs et al. [2022] as an intermediate setting in online optimization without regularizer. Moreover, they also propose two versatile quantities  $\sigma_{1:T}^2$  and  $\Sigma_{1:T}^2$  to reflect the stochastic and adversarial aspect of the environments, respectively. Theoretically, for the SEA model, Sachs et al. [2022] prove that optimistic FTRL enjoys the bounds of  $\mathcal{O}(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})$  and  $\mathcal{O}(\lambda^{-1}(\sigma_{\max}^2 + \Sigma_{\max}^2) \log T)$  for smooth and general convex, and smooth and  $\lambda$ -strongly convex loss functions, respectively. Later, Chen et al. [2023] demonstrate that optimistic OMD is able to attain the same bound for general convex losses and an improved  $\mathcal{O}(\lambda^{-1}(\sigma_{\max}^2 + \Sigma_{\max}^2) \log(\sigma_{1:T}^2 + \Sigma_{1:T}^2))$  bound for the strongly convex losses. Moreover, they also establish a new  $\mathcal{O}((d/\alpha) \log(\sigma_{1:T}^2 + \Sigma_{1:T}^2))$  bound for smooth and  $\alpha$ -exp-concave functions.

**Universal online learning.** The universal online learning is proposed to handle the uncertainty of the loss function types, when applying online algorithms to practical optimization problems. The center to universal algorithms is the powerful meta-expert structure [van Erven and Koolen, 2016, Mhammedi et al., 2019, van Erven et al., 2021], which is also widely-used in many other fields of online learning [Daniely et al., 2015, Jun et al., 2017a,b, Zhang et al., 2018, 2020, 2021, Cutkosky, 2020, Wan et al., 2021, 2022b, Wang et al., 2024, Wan et al., 2024a]. The key idea of meta-expert structure is to maintain multiple experts to process different types of loss functions and then, deploy a meta-algorithm to combine the decisions from experts [Wang et al., 2019, 2020, van Erven et al., 2021, Zhang et al., 2022, Yang et al., 2024a,b]. In the literature, the most related work is Yan et al. [2023], who propose a novel universal algorithm with a two-layer MSMWC [Chen et al., 2021] as the meta-algorithm. By maintaining multiple instances of optimistic OMD, their method can adapt to the SEA model and deliver the same bounds as those in Chen et al. [2023]. Recently, Yan et al. [2024] further improve this method by employing a simpler meta-algorithm while achieving optimal regret bounds for three types of functions. However, it should be noticed that these methods [Yan et al., 2023, 2024] heavily relies on the smoothness of the summation function (1), and thus cannot handle the composite structure in loss functions well, as previously discussed.

### 3 Preliminaries

In this section, we introduce some preliminaries, including standard assumptions and a brief review of OptCMD [Scroccaro et al., 2023].

#### 3.1 Assumptions

We first list the common assumptions in prior studies [Duchi et al., 2010, Chen et al., 2023, 2024].

**Assumption 1.** *The convex decision set  $\mathcal{X}$  belongs to an Euclidean ball with the diameter  $D$ .*

**Assumption 2.** *At each round  $t$ , the random function  $f_t(\cdot)$  is  $G$ -Lipschitz over  $\mathcal{X}$ , i.e.,*

$$\forall \mathbf{x}, \mathbf{y} \in \mathcal{X}, |f_t(\mathbf{x}) - f_t(\mathbf{y})| \leq G \|\mathbf{x} - \mathbf{y}\|_2.$$

**Assumption 3.** *The regularized function  $r(\cdot)$  is convex and bounded over  $\mathcal{X}$ , i.e.,*

$$\forall \mathbf{x}, \mathbf{y} \in \mathcal{X}, r(\mathbf{y}) \geq r(\mathbf{x}) + \nabla r(\mathbf{x})^\top (\mathbf{y} - \mathbf{x}) \text{ and } \forall \mathbf{x} \in \mathcal{X}, 0 \leq r(\mathbf{x}) \leq C.$$

**Assumption 4.** *At each round  $t$ , the expected function  $F_t(\cdot)$  is  $H$ -smooth over  $\mathcal{X}$ , i.e.,*

$$\forall \mathbf{x}, \mathbf{y} \in \mathcal{X}, \|\nabla F_t(\mathbf{x}) - \nabla F_t(\mathbf{y})\|_2 \leq H \|\mathbf{x} - \mathbf{y}\|_2.$$

**Assumption 5.** *At each round  $t$ , the expected function  $F_t(\cdot)$  is convex over  $\mathcal{X}$ , i.e.,*

$$\forall \mathbf{x}, \mathbf{y} \in \mathcal{X}, F_t(\mathbf{y}) \geq F_t(\mathbf{x}) + \nabla F_t(\mathbf{x})^\top (\mathbf{y} - \mathbf{x}).$$

**Assumption 6.** *At each round  $t$ , the expected function  $F_t(\cdot)$  is  $\lambda$ -strongly convex over  $\mathcal{X}$ , i.e.,*

$$\forall \mathbf{x}, \mathbf{y} \in \mathcal{X}, F_t(\mathbf{y}) \geq F_t(\mathbf{x}) + \nabla F_t(\mathbf{x})^\top (\mathbf{y} - \mathbf{x}) + (\lambda/2) \|\mathbf{x} - \mathbf{y}\|_2^2.$$

---

**Algorithm 1** Optimistic Composite Mirror Descent (OptCMD)

---

**Initialization:** Let  $\mathbf{x}_1 = \hat{\mathbf{x}}_1$  be an arbitrary point in  $\mathcal{X}$ .

- 1: **for**  $t = 1$  to  $T$  **do**
  - 2:   Submit  $\mathbf{x}_t$ , and the environment chooses a distribution  $\mathcal{D}_t$
  - 3:   Suffer the composite loss  $f_t(\mathbf{x}_t) + r(\mathbf{x}_t)$ , where  $f_t(\cdot)$  is sampled from  $\mathcal{D}_t$
  - 4:   Update  $\hat{\mathbf{x}}_{t+1}$  and  $\mathbf{x}_{t+1}$  according to (4) and (5), respectively
  - 5: **end for**
- 

**Assumption 7.** All the variance of the gradients and the adversarial variations are bounded, i.e.,

$$\forall t \in [T], \sigma_t^2 \leq \sigma_{\max}^2 \text{ and } \sup_{\mathbf{x} \in \mathcal{X}} \|\nabla F_t(\mathbf{x}) - \nabla F_{t-1}(\mathbf{x})\|_2^2 \leq \Sigma_{\max}^2.$$

**Assumption 8.** At each round  $t$ , the time-varying function  $f_t(\cdot)$  is  $\alpha$ -exp-concave over  $\mathcal{X}$ , i.e.,  $\forall \mathbf{x} \in \mathcal{X}$ ,  $\exp(-\alpha f_t(\mathbf{x}))$  is concave.

### 3.2 Optimistic composite mirror descent

OptCMD is an algorithmic realization of the powerful optimistic online learning [Rakhlin and Sridharan, 2013] in online composite optimization. Specifically, in each round  $t$ , the learner submits a decision  $\mathbf{x}_t$  and suffers a loss  $f_t(\mathbf{x}_t) + r(\mathbf{x}_t)$ . Then, the learner receives an optimism  $M_{t+1}$  that serves as an optimistic prediction for the gradient of subsequent function  $f_{t+1}(\cdot)$ . After that, the learner performs the following update steps:

$$\hat{\mathbf{x}}_{t+1} = \operatorname{argmin}_{\mathbf{x} \in \mathcal{X}} \{ \langle \nabla f_t(\mathbf{x}_t), \mathbf{x} \rangle + r(\mathbf{x}) + \mathcal{B}^{\mathcal{R}_t}(\mathbf{x}, \hat{\mathbf{x}}_t) \} \quad (4)$$

$$\mathbf{x}_{t+1} = \operatorname{argmin}_{\mathbf{x} \in \mathcal{X}} \{ \langle M_{t+1}, \mathbf{x} \rangle + r(\mathbf{x}) + \mathcal{B}^{\mathcal{R}_{t+1}}(\mathbf{x}, \hat{\mathbf{x}}_{t+1}) \} \quad (5)$$

where  $\mathcal{B}^{\mathcal{R}_t}(\mathbf{x}, \mathbf{y}) = \mathcal{R}_t(\mathbf{x}) - \mathcal{R}_t(\mathbf{y}) - \langle \nabla \mathcal{R}_t(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle$  denotes the Bregman divergence associated with a time-varying function  $\mathcal{R}_t$ , of which the specific form depends on the type of  $f_t(\cdot)$  and will be illuminated later. Note that in (4) and (5),  $r(\cdot)$  remains non-linearized, which allows the decisions  $\{\mathbf{x}_t\}_{t=1}^T$  to possess properties externally conferred by  $r(\cdot)$ , such as the sparsity.

**Remark.** Scroccaro et al. [2023] specialize the estimate  $M_{t+1} = \nabla \hat{f}_{t+1}(\hat{\mathbf{x}}_{t+1})$ , where  $\hat{f}_{t+1}(\cdot)$  denotes the function prediction of  $f_{t+1}(\cdot)$  and is generally set as  $f_t(\cdot)$  in practice. We highlight that this specification is unsuitable for the composite SEA model and inadvertently introduces a dependency issue during the analysis. More detailed discussions can be found in Sections 4.1 and 4.2.

## 4 OptCMD for composite SEA

In this section, we provide our results for general convex, strongly convex and exp-concave cases in the composite SEA model. Due to space limitations, all the proofs are deferred to Appendix B.

### 4.1 General convex case

Initially, we consider the case where the regularizer  $r(\cdot)$  is convex and non-smooth, and the expected functions  $F_t(\cdot)$  are general convex and smooth. We choose the following configuration for this case.

$$\mathcal{R}_t(\mathbf{x}) = \frac{1}{2\eta_t} \|\mathbf{x}\|_2^2, \eta_t = \min \left\{ \frac{D}{\sqrt{1 + \bar{V}_{t-1}}}, \frac{D}{\delta} \right\}, \text{ and } M_{t+1} = \nabla f_t(\mathbf{x}_t), \quad (6)$$

where  $\bar{V}_{t-1} = \sum_{s=1}^{t-1} \|\nabla f_s(\mathbf{x}_s) - \nabla f_{s-1}(\mathbf{x}_{s-1})\|_2^2$  and  $\delta > 0$  denotes the hyperparameter. With the configuration in (6), we establish the following regret bound for the general convex case.

**Theorem 1.** Under Assumptions 1, 2, 3, 4 and 5, Algorithm 1 ensures

$$\mathbb{E}[\text{Regret}_T] = \mathcal{O} \left( \sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2} \right)$$

with the configuration in (6) where  $\delta = 6\sqrt{2}HD$ .

**Remark.** For the fully adversarial environments where  $\Sigma_{1:T}^2 = V_T$  and  $\sigma_{1:T}^2 = 0$ , our bound degenerates to  $\mathcal{O}(\sqrt{V_T})$  matching the previous result of Scroccaro et al. [2023]. For the fully stochastic environments where  $\Sigma_{1:T}^2 = 0$  and  $\sigma_{1:T}^2 = \sigma^2 T$  with the stochastic variance  $\sigma^2$ , our result becomes  $\mathcal{O}(\sqrt{T})$  regret bound and further delivers an  $\mathcal{O}(1/\sqrt{T})$  excess risk bound by the online-to-batch conversion [Cesa-Bianchi et al., 2004], which coincides with the results of Lan [2012]. Furthermore, our finding also aligns with existing bounds for the SEA model without  $r(\cdot)$  [Sachs et al., 2022, Chen et al., 2023]. In other words, our method pays *no price* in regret bounds for handling the additional regularizer.

Beyond the configuration in (6), we also employ those adopted by Scroccaro et al. [2023]:

$$\mathcal{R}_t(\mathbf{x}) = \frac{1}{2\eta_t} \|\mathbf{x}\|_2^2, \quad \eta_t = \frac{1}{\sqrt{4H^2 + \bar{D}_{t-1}}}, \quad \text{and } M_{t+1} = \nabla f_t(\hat{\mathbf{x}}_{t+1}), \quad (7)$$

where  $\bar{D}_0 = 0$  and  $\bar{D}_{t-1} = \sum_{s=1}^{t-1} \|\nabla f_s(\hat{\mathbf{x}}_s) - \nabla f_{s-1}(\hat{\mathbf{x}}_s)\|_2^2$ . By setting the configuration in (7), we obtain the following bound.

**Theorem 2.** *Under Assumptions 1, 2, 3, 4 and 5, with configuration in (7), Algorithm 1 ensures*

$$\mathbb{E}[\text{Regret}_T] = \mathcal{O}\left(\sqrt{\tilde{\sigma}_{1:T}^2} + \sqrt{\Sigma_{1:T}^2}\right),$$

where  $\tilde{\sigma}_{1:T}^2$  with  $\tilde{\sigma}_t^2 = \mathbb{E}_{f_t \sim \mathcal{D}_t}[\sup_{\mathbf{x} \in \mathcal{X}} \|\nabla f_t(\mathbf{x}) - \nabla F_t(\mathbf{x})\|_2^2]$ .

**Remark.** This bound is less favorable than that in Theorem 1, as it scales with a new quantity  $\tilde{\sigma}_{1:T}^2$ , which also measures the stochasticity in composite SEA but is larger than  $\sigma_{1:T}^2$  because of the fact that  $\mathbb{E}_{f_t \sim \mathcal{D}_t}[\sup_{\mathbf{x} \in \mathcal{X}} \|\nabla f_t(\mathbf{x}) - \nabla F_t(\mathbf{x})\|_2^2] \geq \sup_{\mathbf{x} \in \mathcal{X}} \mathbb{E}_{f_t \sim \mathcal{D}_t}[\|\nabla f_t(\mathbf{x}) - \nabla F_t(\mathbf{x})\|_2^2]$ , where the inequality is due to the convexity of supremum operator. This discrepancy arises from the inappropriate choice of  $M_{t+1} = \nabla f_t(\hat{\mathbf{x}}_{t+1})$ , where  $\hat{\mathbf{x}}_{t+1}$  is generated based on  $\nabla f_t(\mathbf{x}_t)$  shown in (4), which implies that  $\hat{\mathbf{x}}_{t+1}$  has already incorporated partial information about  $f_t(\cdot)$ , leading to a dependence issue in analysis. To remove the dependency, we apply the supremum operator on  $\hat{\mathbf{x}}_t$  over  $\mathcal{X}$  (c.f. Lemma 4) but inevitably establish a reliance on  $\tilde{\sigma}_{1:T}^2$  in the final bound.

## 4.2 Strongly convex case

In this part, we focus on the case where the regularizer  $r(\cdot)$  is non-smooth and convex, and the expected functions  $F_t(\cdot)$  are smooth and  $\lambda$ -strongly convex. We set OptCMD with the configurations:

$$\mathcal{R}_t(\mathbf{x}) = \frac{1}{2\eta_t} \|\mathbf{x}\|_2^2, \quad \eta_t = \frac{2}{\delta + \lambda t}, \quad \text{and } M_{t+1} = \nabla f_t(\mathbf{x}_t), \quad (8)$$

where  $\delta > 0$  denotes the hyperparameter. The theoretical guarantee for this case is shown below.

**Theorem 3.** *Under Assumptions 1, 2, 3, 4, 6 and 7, Algorithm 1 ensures*

$$\mathbb{E}[\text{Regret}_T] = \mathcal{O}\left(\frac{1}{\lambda} (\sigma_{\max}^2 + \Sigma_{\max}^2) \log(\sigma_{1:T}^2 + \Sigma_{1:T}^2)\right).$$

with the configuration in (8) where  $\delta = 128H^2D^2$ .

**Remark.** Similar to Theorem 1, this bound reduces to  $\mathcal{O}(\lambda^{-1} \log V_T)$  matching that of Scroccaro et al. [2023] when environments become fully adversarial, and derives the same  $\mathcal{O}(\log T/(\lambda T))$  excess risk bound as that of Duchi and Singer [2009] and Xiao [2009] when environments become fully stochastic. Moreover, this bound also recovers that of Chen et al. [2023] for the SEA model.

Additionally, we further equip OptCMD with the original configuration of Scroccaro et al. [2023]:

$$\mathcal{R}_t(\mathbf{x}) = \frac{1}{2\eta_t} \|\mathbf{x}\|_2^2, \quad \eta_t = \frac{1}{4H^2 + (\lambda/2G^2)\bar{D}_{t-1}}, \quad \text{and } M_{t+1} = \nabla f_t(\hat{\mathbf{x}}_{t+1}), \quad (9)$$

and obtain the following theorem.

**Theorem 4.** *Under Assumptions 1, 2, 3, 4 and 6, with the configuration in (9), Algorithm 1 ensures*

$$\mathbb{E}[\text{Regret}_T] = \mathcal{O}\left(\frac{G^2}{\lambda} \log(\tilde{\sigma}_{1:T}^2 + \Sigma_{1:T}^2)\right).$$

**Remark.** This bound suffers two limitations. Firstly, it depends on the unfavorable  $\tilde{\sigma}_{1:T}^2$ , which is due to the improper estimation  $M_{t+1}$  in (9). Secondly, it scales with  $\mathcal{O}(G^2)$  due to the choice of  $\eta_t$  in (9), which is not tighter than  $\mathcal{O}(\sigma_{\max}^2 + \Sigma_{\max}^2)$  in Theorem 3.

---

**Algorithm 2** A Universal Strategy for Composite SEA
 

---

**Initialization:**  $\mathcal{M}_{\text{top}}$  with  $\eta^k = (C_0 \cdot 2^k)^{-1}$  and  $\hat{q}_1^k = (\eta^k)^2 / \sum_{k=1}^K (\eta^k)^2$ ; for each  $k \in [K]$ ,  $\mathcal{M}_{\text{mid}}^k$  with  $\eta^{k,i} = 2\eta^k$  and  $\hat{p}_1^{k,i} = 1/N$ , and experts  $\{E^{k,i}\}_{i \in [N]}$

- 1: **for**  $t = 1$  **to**  $T$  **do**
  - 2:   Receive  $\mathbf{x}_t^{k,i}$  from expert  $E^{k,i}$ , and update  $\mathbf{q}_t$  and  $\mathbf{p}_t^k$  according to (11)
  - 3:   Compute  $\mathbf{x}_t^k = \sum_i p_t^{k,i} \mathbf{x}_t^{k,i}$  and submit  $\mathbf{x}_t = \sum_k q_t^k \mathbf{x}_t^k$
  - 4:   Suffer the loss  $f_t(\mathbf{x}_t) + r(\mathbf{x}_t)$  and observe the gradient  $\mathbf{g}_t = \nabla f_t(\mathbf{x}_t)$
  - 5:   Update  $\hat{\mathbf{q}}_{t+1}$  and  $\hat{\mathbf{p}}_{t+1}^k$  according to (12)
  - 6:   Send the gradient  $\mathbf{g}_t$  and the regularizer  $r(\cdot)$  to experts
  - 7: **end for**
- 

### 4.3 Exp-concave case

We further investigate the exp-concave case where the regularizer  $r(\cdot)$  is non-smooth and convex, and the individual functions  $f_t(\cdot)$  are smooth and  $\alpha$ -exp-concave.

**Remark.** In this case, we require the exponential concavity on the individual function  $f_t(\cdot)$  instead of the expected one  $F_t(\cdot)$ . This is due to the technical demand in analysis, and similar assumptions are also adopted by Chen et al. [2023] and Yang et al. [2024c].

In this case, we set OptCMD with the following configuration:

$$\mathcal{R}_t(\mathbf{x}) = \frac{1}{2} \|\mathbf{x}\|_{H_t}^2, \quad H_t = \delta I + \frac{\beta G^2}{2} I + \frac{\beta}{2} \sum_{s=1}^{t-1} h_s, \quad \text{and } M_{t+1} = \nabla f_t(\mathbf{x}_t), \quad (10)$$

where  $\delta > 0$  denotes the hyperparameter,  $I$  denotes the  $d$ -dimension identity matrix,  $h_t = \nabla f_t(\mathbf{x}_t) \nabla f_t(\mathbf{x}_t)^\top$  and  $\beta = (1/2) \min\{1/(4GD), \alpha\}$ . The regret for this case is stated below.

**Theorem 5.** Under Assumptions 1, 2, 3, 4 and 8, Algorithm 1 ensures

$$\mathbb{E} [\text{Regret}_T] = \mathcal{O} \left( \frac{d}{\alpha} \log (\sigma_{1:T}^2 + \Sigma_{1:T}^2) \right).$$

with the configuration in (10) where  $\delta = 1$ .

**Remark.** For the fully adversarial setting, this bound degenerates to  $\mathcal{O}((d/\alpha) \log V_T)$ , which is the *first* problem-dependent bound for exp-concave time-varying functions in online composite optimization, and tighter than  $\mathcal{O}((d/\alpha) \log T)$  by Yang et al. [2024c] in benign environments. For the fully stochastic setting, our result reduces to  $\mathcal{O}((d/\alpha) \log T)$  regret bound and further implies the same excess risk of  $\mathcal{O}(d \log T / (\alpha T))$  as that of Yang et al. [2024c] through the online-to-batch conversion. Furthermore, it also matches existing results for the SEA model [Chen et al., 2023].

## 5 The universal strategy

Although we have established favorable theoretical results in the composite SEA model, achieving these guarantees requires prior knowledge of the function type to set appropriate configurations, e.g., (6) for the general convex case. This requirement is often impractical in real-world scenarios and motivates us to design a universal strategy. In the previous study of Yan et al. [2023], based on the meta-expert framework [van Erven and Koolen, 2016, Wang et al., 2020, Zhang et al., 2022], they propose a multi-layer universal algorithm, which can adapt to the SEA model without regularizer by choosing optimistic OMD [Chen et al., 2023] as the expert-algorithm. Unfortunately, their method [Yan et al., 2023] requires the smoothness of the (summation) loss functions, which does not necessarily hold for (1), so that it cannot be directly applied to the composite SEA model. In this paper, we propose a *novel* universal algorithm for composite SEA. For the expert-algorithm, we choose OptCMD due to its ability in handling three cases in composite SEA. For the meta-algorithm, we design a new method that is able to effectively manage the non-smooth component  $r(\cdot)$  in (1) and can thus track experts according to their performances in the composite SEA model.

Our method is based on two key observations. Firstly,  $r(\cdot)$  in (1) is fixed over time and available to the meta-algorithm from the beginning round. Secondly, the meta-algorithm can access expert

decisions for the current round before the performance estimation. Therefore, the information about  $r(\cdot)$  can be introduced into the expert tracking process, securing that the regret of meta-algorithm will not deviate from that of the best expert. Specifically, inspired by Yan et al. [2023], we employ a two-layer meta-algorithm, with each layer running an MsMwC algorithm [Chen et al., 2021] that maintains two weight sequences  $\{\mathbf{q}_t, \hat{\mathbf{q}}_t \in \Delta^N\}_{t \in [T]}$ , and a group of losses  $\mathbf{l}_t = (l_t^1, \dots, l_t^N)$  and optimism  $\mathbf{m}_t = (m_t^1, \dots, m_t^N)$  to measure the performance of  $N$  experts. To handle the composite loss, we explicitly incorporate  $r(\cdot)$  into both  $\mathbf{l}_t$  and  $\mathbf{m}_t$ . The incorporation serves two purposes: (i) it endows  $\mathbf{l}_t$  with a composite structure, facilitating the estimation of expert performance on composite losses; (ii) it utilizes the cancellation between  $\mathbf{l}_t$  and  $\mathbf{m}_t$  to eliminate the non-smooth component in loss functions, i.e.,  $r(\cdot)$ , so that the meta-algorithm is able to leverage the smoothness of  $f_t(\cdot)$ . In the following, we describe our universal algorithm, which is also summarized in Algorithm 2.

**Meta-algorithm.** Overall, our algorithm comprises two components: a two-layer meta-algorithm and a set of experts, together forming a three-layer structure. Specifically, on the top layer, we run one MsMwC named  $\mathcal{M}_{\text{top}}$ , which maintains the weights  $\mathbf{q} \in \Delta_K$  assigned to  $K = \mathcal{O}(\log T)$  MsMwCs, named  $\mathcal{M}_{\text{mid}}$ . On the middle layer, each  $\mathcal{M}_{\text{mid}}^k$  ( $k \in [K]$ ) maintains the weights  $\mathbf{p}^k \in \Delta_N$  assigned to  $N = \mathcal{O}(\log T)$  experts. At the round  $t$ , we receive the decision  $\mathbf{x}_t^{k,i}$  from each expert  $E^{k,i}$ , and update the weights  $\mathbf{q}_t$  and  $\mathbf{p}_t^k$  according to:

$$\mathbf{q}_t = \operatorname{argmin}_{\mathbf{q} \in \Delta^K} \{\langle \mathbf{m}_t, \mathbf{q} \rangle + \mathcal{B}^{\psi_1}(\mathbf{q}, \hat{\mathbf{q}}_t)\}, \quad \mathbf{p}_t^k = \operatorname{argmin}_{\mathbf{p} \in \Delta^N} \{\langle \mathbf{m}_t^k, \mathbf{p} \rangle + \mathcal{B}^{\psi_2}(\mathbf{p}, \hat{\mathbf{p}}_t^k)\}, \quad (11)$$

where  $\psi_1(\mathbf{q}) = \sum_k (q^k \ln q^k) / \eta^k$  and  $\psi_2(\mathbf{p}) = \sum_i (p^i \ln p^i) / \eta^{k,i}$  denote the negative entropy functions, and  $\mathbf{m}_t^k$  and  $\mathbf{m}_t$  denote the optimism used by  $\mathcal{M}_{\text{mid}}^k$  and  $\mathcal{M}_{\text{top}}$ , respectively. Then, we compute the weighted average decision  $\mathbf{x}_t^k = \sum_i p_t^{k,i} \mathbf{x}_t^{k,i}$  and submit  $\mathbf{x}_t = \sum_k q_t^k \mathbf{x}_t^k$ . Next, we suffer the loss  $f_t(\mathbf{x}_t) + r(\mathbf{x}_t)$  and observe the gradient  $\mathbf{g}_t = \nabla f_t(\mathbf{x}_t)$ . We update  $\hat{\mathbf{q}}_{t+1}$  and  $\hat{\mathbf{p}}_{t+1}^k$  by

$$\hat{\mathbf{q}}_{t+1} = \operatorname{argmin}_{\mathbf{q} \in \Delta^K} \{\langle \mathbf{l}_t + \mathbf{a}_t, \mathbf{q} \rangle + \mathcal{B}^{\psi_1}(\mathbf{q}, \hat{\mathbf{q}}_t)\}, \quad \hat{\mathbf{p}}_{t+1}^k = \operatorname{argmin}_{\mathbf{p} \in \Delta^N} \{\langle \mathbf{l}_t^k + \mathbf{b}_t^k, \mathbf{p} \rangle + \mathcal{B}^{\psi_2}(\mathbf{p}, \hat{\mathbf{p}}_t^k)\}, \quad (12)$$

where  $\mathbf{l}_t = (l_t^1, \dots, l_t^K)$  and  $\mathbf{l}_t^k = (l_t^{k,1}, \dots, l_t^{k,N})$  denote losses used by  $\mathcal{M}_{\text{top}}$  and  $\mathcal{M}_{\text{mid}}^k$ , respectively, and  $\mathbf{a}_t \in \mathbb{R}^K$  and  $\mathbf{b}_t^k \in \mathbb{R}^N$  denote the correction terms. At last, we send the gradient  $\mathbf{g}_t$  and the regularizer  $r(\cdot)$  to each expert. Now, we specify the loss and optimism in (11) and (12). For  $\mathcal{M}_{\text{top}}$ , we choose

$$l_t^k = \langle \mathbf{g}_t, \mathbf{x}_t^k \rangle + r(\mathbf{x}_t^k) + \gamma_1 \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2, \quad m_t^k = \langle \hat{\mathbf{m}}_t^k, \mathbf{p}_t^k \rangle + r(\mathbf{x}_t^k) + \gamma_1 \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2, \quad (13)$$

where  $\hat{\mathbf{m}}_t^k = (\hat{m}_t^{k,1}, \dots, \hat{m}_t^{k,N})$  with  $\hat{m}_t^{k,i} = \langle \mathbf{g}_t, \mathbf{x}_t \rangle - \langle \mathbf{g}_{t-1}, \mathbf{x}_{t-1} - \mathbf{x}_{t-1}^{k,i} \rangle$ . For  $\mathcal{M}_{\text{mid}}^k$ , we choose

$$l_t^{k,i} = \langle \mathbf{g}_t, \mathbf{x}_t^{k,i} \rangle + r(\mathbf{x}_t^{k,i}) + \gamma_2 \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2, \quad m_t^{k,i} = \hat{m}_t^{k,i} + r(\mathbf{x}_t^{k,i}) + \gamma_2 \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2. \quad (14)$$

In the above,  $\gamma_1, \gamma_2$  denote hyperparameters, and  $\gamma_1 \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2$  and  $\gamma_2 \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2$  are injected for cancellation during the analysis [Yan et al., 2023]. It should be noticed that we explicitly introduce  $r(\cdot)$  in (13) and (14). This is because experts are running over the composite losses and their performance assessment should similarly reflect the composite structure, ensuring an accurate measure for expert tracking. Moreover, the optimism  $\mathbf{m}_t$  with  $r(\cdot)$  helps eliminate the non-smooth component in  $\mathbf{l}_t$  and thereby enables the meta-algorithm to effectively utilize the smoothness of time-varying functions. To make it clearer, we take the general convex case as an example.

First, the expected regret can be decomposed into the meta-regret and the expert-regret, as below:

$$\begin{aligned} \mathbb{E} [\text{Regret}_T] &= \mathbb{E} \left[ \sum_{t=1}^T [f_t(\mathbf{x}_t) + r(\mathbf{x}_t)] - \min_{\mathbf{x} \in \mathcal{X}} \sum_{t=1}^T [f_t(\mathbf{x}) + r(\mathbf{x})] \right] \\ &\leq \underbrace{\mathbb{E} \left[ \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{x}_t - \mathbf{x}_t^{k^*, i^*} \rangle + r(\mathbf{x}_t) - r(\mathbf{x}_t^{k^*, i^*}) \right]}_{\text{:=meta-regret}} + \underbrace{\mathbb{E} \left[ \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{x}_t^{k^*, i^*} - \mathbf{x}^* \rangle + r(\mathbf{x}_t^{k^*, i^*}) - r(\mathbf{x}^*) \right]}_{\text{:=expert-regret}}, \end{aligned}$$

where  $\mathbf{x}_t^{k^*, i^*}$  denotes the decision made by the best expert  $E^{k^*, i^*}$ . For the expert-regret, we can directly apply the result in Theorem 1. Therefore, the key is how to bound the meta-regret. By the fact that  $r(\mathbf{x}_t) \leq \sum_k q_t^k r(\mathbf{x}_t^k)$  and  $r(\mathbf{x}_t^k) \leq \sum_i p_t^{k,i} r(\mathbf{x}_t^{k,i})$ , the meta-regret is bounded by

$$\text{meta-regret} \leq \sum_{t=1}^T \left[ \langle \mathbf{q}_t, \mathbf{l}_t \rangle - l_t^{k^*} \right] + \sum_{t=1}^T \left[ \langle \mathbf{p}_t^{k^*}, \mathbf{l}_t^{k^*} \rangle - l_t^{k^*, i^*} \right] + \mathcal{O}(1), \quad (15)$$



where the first and second terms denote the regrets for tracking experts of  $\mathcal{M}_{\text{top}}$  and  $\mathcal{M}_{\text{mid}}^{k^*}$ , respectively. While these two terms involve the losses  $l_t^k$  and  $l_t^{k^*,i}$  that incorporate the regularizer, they can be simultaneously controlled by a favorable bound of  $\mathcal{O}(\sum_{t \in [T]} (l_t^{k^*,i^*} - m_t^{k^*,i^*})^2)$ , which not only eliminates  $r(\cdot)$  by the cancellation between  $l_t^{k^*,i^*}$  and  $m_t^{k^*,i^*}$ , but also suffices to achieve the desirable bound of  $\tilde{\mathcal{O}}(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})$  for general convex case and  $\mathcal{O}(1)$  for other two cases. More details can be found in Appendix B.6.

**Experts.** Each expert is an instance of OptCMD that equips with certain configuration, i.e., (6), (8) or (10). To estimate the unknown curvature  $\lambda$  and  $\alpha$ , we utilize the discretization strategy by Zhang et al. [2022], constructing the candidate sets as

$$\mathcal{P}_{str} = \mathcal{P}_{exp} = \{1/T, 2/T, 2^2/T, \dots, 2^N/T\},$$

where  $N = \lceil \log_2 T \rceil$ . Based on the two sets, we create following three types of experts:

- Strongly convex experts, each of which is configured with (8) and a candidate  $\lambda^i \in \mathcal{P}_{str}$ ;
- Exp-concave experts, each of which is configured with (10) and a candidate  $\alpha^i \in \mathcal{P}_{exp}$ ;
- General convex expert, which is configured with (6).

As demonstrated by Zhang et al. [2022], these experts are sufficient to identify the best one that runs for the true function type with the most accurate curvature. For instance, in the strongly convex case, there must exist an expert  $E^i$  equipped with  $\lambda^i \in \mathcal{P}_{str}$  that satisfies  $\lambda^i \leq \lambda \leq 2\lambda^i$ . Moreover, instead of directly minimizing the original loss function (1), each expert runs over a *new* composite surrogate loss to avoid high-computational gradient query costs. Specifically, we choose  $h_t^c(\mathbf{x}) = \langle \mathbf{g}_t, \mathbf{x} \rangle + r(\mathbf{x})$ ,  $h_{t,i}^{sc}(\mathbf{x}) = \langle \mathbf{g}_t, \mathbf{x} \rangle + r(\mathbf{x}) + \lambda^i \|\mathbf{x} - \mathbf{x}_t\|_2^2/2$  and  $h_{t,i}^{exp}(\mathbf{x}) = \langle \mathbf{g}_t, \mathbf{x} \rangle + r(\mathbf{x}) + \beta^i \langle \mathbf{g}_t, \mathbf{x} - \mathbf{x}_t \rangle^2/2$  for general convex, strongly convex, and exp-concave experts, respectively. These surrogate losses not only inherit the composite structure and properties of  $f_t(\mathbf{x})$ , e.g., the strong convexity, but they also require only *one* gradient query for  $\mathbf{g}_t = \nabla f_t(\mathbf{x}_t)$  in each round. The theoretical guarantees of our universal algorithm are presented below.

**Theorem 6.** *Under Assumptions 1, 2, 3, 4 and 7, with proper hyperparameters, Algorithm 2 ensures the bounds of  $\tilde{\mathcal{O}}(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})$ ,  $\mathcal{O}((\sigma_{\max}^2 + \Sigma_{\max}^2) \log(\sigma_{1:T}^2 + \Sigma_{1:T}^2))$ , and  $\mathcal{O}(d \log(\sigma_{1:T}^2 + \Sigma_{1:T}^2))$  for general convex, strongly convex, and exp-concave time-varying functions, respectively.*

**Remark.** Theorem 6 indicates that Algorithm 2 can attain comparable regret bounds to those of OptCMD for the three cases, without knowing the function type and curvature in advance.

## 6 Implications

In the following, we discuss implications of our results for two common intermediate examples.

**Composite adversarially corrupted stochastic data.** We first focus on the composite adversarially corrupted stochastic model, of which a special case without regularizer has been widely investigated in learning with expert advice [Amir et al., 2020], bandits [Zimmert and Seldin, 2021, Ito, 2021], and online optimization [Sachs et al., 2022, Chen et al., 2023]. Specifically, each loss function consists of three terms:  $\phi_t(\cdot) = h_t(\cdot) + c_t(\cdot) + r(\cdot)$ , where  $h_t(\cdot)$  denotes the loss of i.i.d. data from a fixed distribution  $\mathcal{D}$  with the variance  $\sigma$ , and  $c_t(\cdot)$  denotes an adversarial perturbation measured by a parameter  $C_T > 0$ , satisfying  $\sum_{t \in [T]} \max_{\mathbf{x} \in \mathcal{X}} \|\nabla c_t(\mathbf{x})\|_2 \leq C_T$ , and  $r(\cdot)$  denotes the regularizer. In this model, the stochasticity and adversariality come from  $h_t(\cdot)$  and  $c_t(\cdot)$ , respectively. Therefore,  $\sigma_{1:T}^2$  and  $\Sigma_{1:T}^2$  can be formulated as

$$\sigma_{1:T}^2 = \sigma^2 T, \quad \Sigma_{1:T}^2 = \mathbb{E} \left[ \sum_{t=1}^T \sup_{\mathbf{x} \in \mathcal{X}} \|\nabla c_t(\mathbf{x}) - \nabla c_{t-1}(\mathbf{x})\|_2^2 \right] \leq 4GC_T. \quad (16)$$

With the above specification, our theoretical results can naturally be extended to this scenarios, delivering the following corollary.

**Corollary 1.** *With the specification (16) in the scenarios of composite adversarially corrupted stochastic data, we can obtain an  $\mathcal{O}(\sigma\sqrt{T} + \sqrt{C_T})$  bound for the general convex case by Theorem 1, an  $\mathcal{O}(\log(\sigma^2 T + C_T))$  bound for the strongly convex case by Theorem 3, and an  $\mathcal{O}(d \log(\sigma^2 T + C_T))$  bound for the exp-concave case by Theorem 5.*

**Adversarial online learning with limited resources.** We then consider the another common scenarios where loss functions arrive in batches and the available computing resources are insufficient to process them all [Bottou and Cun, 2003, Benczúr et al., 2018, Li et al., 2024, Zhou, 2024]. Specifically, in each round  $t$ , we receive a group of functions  $\mathcal{F}_t = \{f_t(\cdot, i)\}_{i \in [K_t]}$  with the size  $K_t$ , each of which is selected by the environments adversarially. We denote by  $F_t(\cdot) = K_t^{-1} \sum_{i \in [K_t]} f_t(\cdot, i)$  the average function. Due to limited computing resources, we can only sample a subset of functions for gradient estimation, and generate sparse decisions for efficient storage and inference. For this reason, our goal is to minimize  $\phi_t(\cdot) = h_t(\cdot) + r(\cdot)$ , where  $h_t(\cdot) = B_t^{-1} \sum_{i \in [B_t]} \hat{f}_t(\cdot, i)$  denotes the approximation of  $F_t(\cdot)$  by i.i.d. sampling  $B_t \in [K_t]$  functions  $\hat{f}_t(\cdot, i)$  from the group  $\mathcal{F}_t$ , and  $r(\cdot) = \|\cdot\|_1$  denotes the  $\ell_1$ -norm regularizer. In this scenarios, it can be verified that

$$\sigma_{1:T}^2 = \sum_{t=1}^T \sup_{\mathbf{x} \in \mathcal{X}} \mathbb{E}_{\hat{f}_t \sim \mathcal{F}_t} \left[ \left\| B_t^{-1} \sum_{i=1}^{B_t} \nabla \hat{f}_t(\cdot, i) - \nabla F_t(\mathbf{x}) \right\|_2^2 \right] \leq 4G^2 \sum_{t=1}^T B_t^{-1} \quad (17)$$

and  $\Sigma_{1:T}^2 = \mathbb{E}[\sum_{t=1}^T \sup_{\mathbf{x} \in \mathcal{X}} \|\nabla F_t(\mathbf{x}) - \nabla F_{t-1}(\mathbf{x})\|_2^2]$ . Applying the above specification into our theorems delivers the following corollary.

**Corollary 2.** *With the specification (17) in the scenarios of composite adversarially corrupted stochastic data, we can obtain an  $\mathcal{O}(\sqrt{\sum_{t \in [T]} B_t^{-1}} + \sqrt{\Sigma_{1:T}^2})$  bound for the general convex case by Theorem 1, an  $\mathcal{O}(\log(\sum_{t \in [T]} B_t^{-1} + \Sigma_{1:T}^2))$  bound for the strongly convex case by Theorem 3, and an  $\mathcal{O}(d \log(\sum_{t \in [T]} B_t^{-1} + \Sigma_{1:T}^2))$  bound for the exp-concave case by Theorem 5.*

## 7 Conclusion and future work

In this paper, we investigate the intermediate setting between stochastic and adversarial composite optimization, named composite SEA, and demonstrate that OptCMD is able to attain the regret bounds of  $\mathcal{O}(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2})$ ,  $\mathcal{O}((\sigma_{\max}^2 + \Sigma_{\max}^2) \log(\sigma_{1:T}^2 + \Sigma_{1:T}^2))$  and  $\mathcal{O}(d \log(\sigma_{1:T}^2 + \Sigma_{1:T}^2))$  for the general convex, strongly convex and exp-concave cases, respectively. To deal with the unknown function type in real-world problems, we further propose a *novel* universal algorithm in online composite optimization, and show that our universal algorithm is able to achieve the desirable bounds in the three cases, simultaneously. Due to the versatility of  $\sigma_{1:T}^2$  and  $\Sigma_{1:T}^2$ , all our theoretical findings can *recover* previous results in fully stochastic and adversarial composite optimization. Finally, we explore several practical intermediate scenarios to demonstrate the implications of our results. Additionally, we also conduct empirical studies in Appendix A to verify our theoretical results.

There are many valuable directions for future research. First, our methods still rely on the domain and gradient bounded assumptions. We notice that there have been several methods designed for unbounded cases [Orabona, 2014, Orabona and Pál, 2016, Cutkosky and Orabona, 2018, Jacobsen and Cutkosky, 2022], but all of them focus on the setting without the regularizer. Consequently, developing online algorithms with the unbounded domain and gradient for composite SEA remains an interesting research direction. Second, in the composite SEA model, we implicitly assume that the feedback (i.e., the loss function value and the gradient) is immediately revealed after making the decision, which, however, is not necessarily satisfied in practice [Joulani et al., 2013, Quanrud and Khashabi, 2015, Wan et al., 2022a,c, 2024b]. Therefore, it is also interesting to explore the composite SEA model with the delayed feedback. Thirdly, our proposed universal algorithm currently exhibits a three-layer structure, which presents several challenges in both analysis and practical implementation. Therefore, designing a simpler two-layer universal algorithm in composite SEA is another potential research direction in the future, which may require the advanced techniques in Yan et al. [2024].

## Acknowledgments and Disclosure of Funding

The authors would thank Yu-Hu Yan for helpful discussions, and the anonymous reviewers for their constructive suggestions. This work was partially supported by National Science and Technology Major Project (2022ZD0114801), and NSFC (U23A20382, 62306275).

## References

- I. Amir, I. Attias, T. Koren, R. Livni, and Y. Mansour. Prediction with corrupted expert advice. In *Advances in Neural Information Processing Systems 33*, pages 14315–14325, 2020.
- A. Beck and M. Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 31(3):167–175, may 2003.
- A. A. Benczúr, L. Kocsis, and R. Pálóvics. Online machine learning in big data streams. *ArXiv e-prints*, arXiv:1802.05872, 2018.
- L. Bottou and Y. L. Cun. Large scale online learning. In *Advances in Neural Information Processing Systems 16*, pages 217–224, 2003.
- N. Cesa-Bianchi, A. Conconi, and C. Gentile. On the generalization ability of on-line learning algorithms. *IEEE Transactions on Information Theory*, 50(9):2050–2057, 2004.
- C.-C. Chang and C.-J. Lin. Libsvm: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(27):1–27, 2011.
- L. Chen, H. Luo, and C.-Y. Wei. Impossible tuning made possible: A new expert algorithm and its applications. In *Proceedings of the 34th Conference on Learning Theory*, pages 1216–1259, 2021.
- S. Chen, W.-W. Tu, P. Zhao, and L. Zhang. Optimistic online mirror descent for bridging stochastic and adversarial online convex optimization. In *Proceedings of the 40th International Conference on Machine Learning*, pages 5002–5035, 2023.
- S. Chen, Y.-J. Zhang, W.-W. Tu, P. Zhao, and L. Zhang. Optimistic online mirror descent for bridging stochastic and adversarial online convex optimization. *Journal of Machine Learning Research*, 25(178):1–62, 2024.
- C.-K. Chiang, T. Yang, C.-J. Lee, M. Mahdavi, C.-J. Lu, R. Jin, and S. Zhu. Online optimization with gradual variations. In *Proceedings of the 25th Conference on Learning Theory*, pages 6.1–6.20, 2012.
- A. Cutkosky. Parameter-free, dynamic, and strongly-adaptive online learning. In *Proceedings of the 37th International Conference on Machine Learning*, pages 2250–2259, 2020.
- A. Cutkosky and F. Orabona. Black-box reductions for parameter-free online learning in banach spaces. In *Proceedings of the 31st Conference on Learning Theory*, pages 1493–1529, 2018.
- A. Daniely, A. Gonen, and S. Shalev-Shwartz. Strongly adaptive online learning. In *Proceedings of the 32nd International Conference on Machine Learning*, pages 1405–1411, 2015.
- J. Duchi and Y. Singer. Efficient online and batch learning using forward backward splitting. *Journal of Machine Learning Research*, 10:2899–2934, 2009.
- J. C. Duchi, S. Shalev-Shwartz, Y. Singer, and A. Tewari. Composite objective mirror descent. In *Proceedings of the 23rd Annual Conference on Learning Theory*, pages 14–26, 2010.
- N. Flammarion and F. Bach. Stochastic composite least-squares regression with convergence rate  $o(1/n)$ . In *Proceedings of the 30th Conference on Learning Theory*, pages 831–875, 2017.
- D. Garber and A. Kaplan. Fast stochastic algorithms for low-rank and nonsmooth matrix problems. In *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics*, pages 286–294, 2019.
- D. Garber, G. Korcia, and K. Levy. Online convex optimization in the random order model. In *Proceedings of the 37th International Conference on Machine Learning*, pages 3387–3396, 2020.
- S. Ghadimi and G. Lan. Optimal stochastic approximation algorithms for strongly convex stochastic composite optimization I: A generic algorithmic framework. *SIAM Journal on Optimization*, 22(4):1469–1492, 2012.
- E. Hazan, A. Agarwal, and S. Kale. Logarithmic regret algorithms for online convex optimization. *Foundations and Trends in Machine Learning*, 69(2):169–192, 2007.

- E. Hazan, T. Koren, and K. Y. Levy. Logistic regression: Tight bounds for stochastic and online optimization. In *Proceedings of the 27th Conference on Learning Theory*, pages 197–209, 2014.
- S. Ito. On optimal robustness to adversarial corruption in online decision problems. In *Advances in Neural Information Processing Systems 34*, pages 7409–7420, 2021.
- A. Jacobsen and A. Cutkosky. Parameter-free mirror descent. In *Proceedings of the 35th Conference on Learning Theory*, pages 4160–4211, 2022.
- P. Joulani, A. György, and C. Szepesvári. Online learning under delayed feedback. In *Proceedings of the 30th International Conference on Machine Learning*, pages 1453–1461, 2013.
- P. Joulani, A. György, and C. Szepesvári. A modular analysis of adaptive (non-)convex optimization: Optimism, composite objectives, variance reduction, and variational bounds. *Theoretical Computer Science*, 808:108–138, 2020.
- K.-S. Jun, F. Orabona, R. Willett, and S. Wright. Improved strongly adaptive online learning using coin betting. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, pages 943–951, 2017a.
- K.-S. Jun, F. Orabona, S. Wright, and R. Willett. Online learning for changing environments using coin betting. *Electronic Journal of Statistics*, 11(2):5282–5310, 2017b.
- A. Kulunchakov and J. Mairal. A generic acceleration framework for stochastic composite optimization. In *Advances in Neural Information Processing Systems 32*, pages 12556–12567, 2019.
- G. Lan. An optimal method for stochastic composite optimization. *Mathematical Programming*, 133:365–397, 2012.
- G. Lan. Gradient sliding for composite optimization. *Mathematical Programming*, 159:201–235, 2016.
- J. Langford, L. Li, and T. Zhang. Sparse online learning via truncated gradient. *Journal of Machine Learning Research*, 10:777–801, 2009.
- Y. Lei and K. Tang. Stochastic composite mirror descent: Optimal bounds with high probabilities. In *Advances in Neural Information Processing Systems 31*, pages 1519–1529, 2018.
- Y. Lei and D.-X. Zhou. Analysis of online composite mirror descent algorithm. *Neural Computation*, 29(3):825–860, 2017.
- Y. Lei, P. Yang, K. Tang, and D.-X. Zhou. Optimal stochastic and online learning with individual iterates. In *Advances in Neural Information Processing Systems 32*, pages 5415–5425, 2019.
- L. Li, T. Zhang, Z. Bu, S. Wang, H. He, J. Fu, Y. Wu, J. Bian, Y. Chen, and Y. Bengio. Map: Low-compute model merging with amortized pareto fronts via quadratic approximation. *ArXiv e-prints*, arXiv:2406.07529, 2024.
- Z. Mhammedi, W. M. Koolen, and T. van Erven. Lipschitz adaptivity with multiple learning rates in online learning. In *Proceedings of the 32nd Conference on Learning Theory*, pages 2490–2511, 2019.
- M. Mohri and S. Yang. Accelerating online convex optimization via adaptive prediction. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, pages 848–856, 2016.
- Y. Nesterov. Primal-dual subgradient methods for convex problems. *Mathematical Programming*, 120(1):261–283, 2009.
- F. Orabona. Simultaneous model selection and optimization through parameter-free stochastic learning. In *Advances in Neural Information Processing Systems 27*, pages 1116–1124, 2014.
- F. Orabona and D. Pál. Coin betting and parameter-free online learning. In *Advances in Neural Information Processing Systems 29*, pages 577–585, 2016.

- R. Pogodin and T. Lattimore. On first-order bounds, variance and gap-dependent bounds for adversarial bandits. In *Proceedings of the 35th Conference on Uncertainty in Artificial Intelligence*, pages 894–904, 2019.
- K. Quanrud and D. Khashabi. Online learning with adversarial delays. In *Advances in Neural Information Processing Systems 28*, pages 1270–1278, 2015.
- A. Rakhlin and K. Sridharan. Online learning with predictable sequences. In *Proceedings of the 26th Conference on Learning Theory*, pages 993–1019, 2013.
- S. Sachs, H. Hadiji, T. van Erven, and C. Guzmán. Between stochastic and adversarial online convex optimization: Improved regret bounds via smoothness. In *Advances in Neural Information Processing Systems 35*, pages 691–702, 2022.
- P. Z. Scroccaro, A. S. Kolarijani, and P. M. Esfahani. Adaptive composite online optimization: Predictions in static and dynamic environments. *IEEE Transactions on Automatic Control*, 68(5): 2906–2921, 2023.
- U. Sherman, T. Koren, and Y. Mansour. Optimal rates for random order online optimization. In *Advances in Neural Information Processing Systems 34*, pages 2097–2108, 2021.
- T. van Erven and W. M. Koolen. Metagrad: Multiple learning rates in online learning. In *Advances in Neural Information Processing Systems 29*, pages 3666–3674, 2016.
- T. van Erven, W. M. Koolen, and D. van der Hoeven. Metagrad: adaptation using multiple learning rates in online learning. *Journal of Machine Learning Research*, 22(161):1–61, 2021.
- Y. Wan, B. Xue, and L. Zhang. Projection-free online learning in dynamic environments. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence Advances*, pages 10067–10075, 2021.
- Y. Wan, W.-W. Tu, and L. Zhang. Online frank-wolfe with arbitrary delays. In *Advances in Neural Information Processing Systems 35*, pages 19703–19715, 2022a.
- Y. Wan, W.-W. Tu, and L. Zhang. Strongly adaptive online learning over partial intervals. *Science China Information Sciences*, 65(10):202101, 2022b.
- Y. Wan, Y. Wang, C. Yao, W.-W. Tu, and L. Zhang. Projection-free online learning with arbitrary delays. *ArXiv e-prints*, arXiv:2204.04964, 2022c.
- Y. Wan, C. Yao, M. Song, and L. Zhang. Non-stationary online convex optimization with arbitrary delays. In *Proceedings of the 41st International Conference on Machine Learning*, pages 49991–50011, 2024a.
- Y. Wan, C. Yao, M. Song, and L. Zhang. Improved regret for bandit convex optimization with delayed feedback. In *Advances in Neural Information Processing Systems 37*, 2024b.
- G. Wang, S. Lu, and L. Zhang. Adaptivity and optimality: A universal algorithm for online convex optimization. In *Proceedings of the 35th Conference on Uncertainty in Artificial Intelligence*, pages 659–668, 2019.
- G. Wang, S. Lu, Y. Hu, and L. Zhang. Adapting to smoothness: A more universal algorithm for online convex optimization. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence*, pages 6162–6169, 2020.
- Y. Wang, W. Yang, W. Jiang, S. Lu, B. Wang, H. Tang, Y. Wan, and L. Zhang. Non-stationary projection-free online learning with dynamic and adaptive regret guarantees. In *Proceedings of the 38th AAAI Conference on Artificial Intelligence*, pages 15671–15679, 2024.
- L. Xiao. Dual averaging method for regularized stochastic learning and online optimization. In *Advances in Neural Information Processing Systems 22*, pages 2116–2124, 2009.
- Y.-H. Yan, P. Zhao, and Z.-H. Zhou. Universal online learning with gradual variations: A multi-layer online ensemble approach. In *Advances in Neural Information Processing Systems 36*, pages 37682–37715, 2023.

- Y.-H. Yan, P. Zhao, and Z.-H. Zhou. A simple and optimal approach for universal online learning with gradient variations. In *Advances in Neural Information Processing Systems 37*, 2024.
- T. Yang, M. Mahdavi, R. Jin, and S. Zhu. Regret bounded by gradual variation for online convex optimization. *Machine Learning*, 95:183–223, 2014.
- W. Yang, W. Jiang, Y. Wang, P. Yang, Y. Hu, and L. Zhang. Small-loss adaptive regret for online convex optimization. In *Proceedings of the 41st International Conference on Machine Learning*, pages 56156–56195, 2024a.
- W. Yang, Y. Wang, P. Zhao, and L. Zhang. Universal online convex optimization with 1 projection per round. In *Advances in Neural Information Processing Systems 37*, 2024b.
- X. Yang, P. Tian, X. Cheng, Y. Wan, and M. Song. Regularized online exponentially concave optimization. *Neurocomputing*, 595:127789, 2024c.
- L. Zhang, S. Lu, and Z.-H. Zhou. Adaptive online learning in dynamic environments. In *Advances in Neural Information Processing Systems 31*, pages 1323–1333, 2018.
- L. Zhang, T. Yang, R. Jin, and Z.-H. Zhou. Relative error bound analysis for nuclear norm regularized matrix completion. *Journal of Machine Learning Research*, 20(97):1–22, 2019.
- L. Zhang, S. Lu, and T. Yang. Minimizing dynamic regret and adaptive regret simultaneously. In *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics*, pages 309–319, 2020.
- L. Zhang, G. Wang, W.-W. Tu, W. Jiang, and Z.-H. Zhou. Dual adaptivity: A universal algorithm for minimizing the adaptive regret of convex functions. In *Advances in Neural Information Processing Systems 34*, pages 24968–24980, 2021.
- L. Zhang, G. Wang, J. Yi, and T. Yang. A simple yet universal strategy for online convex optimization. In *Proceedings of the 39th International Conference on Machine Learning*, pages 26605–26623, 2022.
- W. Zhang, L. Zhang, Z. Jin, R. Jin, D. Cai, X. Li, R. Liang, and X. He. Sparse learning with stochastic composite optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6):1223–1236, 2017.
- Z.-H. Zhou. Learnability with time-sharing computational resource concerns. *National Science Review*, 11(10):nwae204, 2024.
- J. Zimmert and Y. Seldin. Tsallis-INF: An optimal algorithm for stochastic and adversarial bandits. *Journal of Machine Learning Research*, 22(28):1–49, 2021.
- M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th International Conference on Machine Learning*, pages 928–936, 2003.

## A Experiments

In this section, we conduct empirical studies to validate our theoretical results.

**Setup.** In this paper, we show that OptCMD with suitable configurations is able to achieve a series of favorable theoretical guarantees for the composite SEA model. Moreover, for the practical scenarios where the prior knowledge of loss functions is unavailable, we propose a novel universal strategy, called USC-SEA, which can still achieve the desired regret bounds for three cases in the composite SEA simultaneously. To verify our theoretical findings, we conduct experiments on the mushroom datasets from the LIBSVM repository [Chang and Lin, 2011], and consider the following online classification problem. Let  $T$  denote the number of the total rounds. At each round  $t \in [T]$ , the learner receives a sampled data  $(\mathbf{x}_t, y_t) \in \mathbb{R}^d \times \{-1, 1\}$  with  $d = 112$ . Then, the learner plays the decision  $\mathbf{w}_t$  from the ball  $\mathcal{X}$  with the diameter  $D = 20$ , and suffers a composite loss

$$\phi_t(\mathbf{w}_t; \mathbf{x}_t, y_t) = f_t(\mathbf{w}_t; \mathbf{x}_t, y_t) + \lambda r(\mathbf{w}_t),$$

where we set the hyper-parameter  $\lambda = 0.001$ . The dataset used in the experiments is considered to be sampled from an unknown distribution, possessing the inherent stochastic property. To simulate the stochastically extended adversarial environments, we perturb the dataset by randomly flipping the labels of 10% data as the adversarial corruptions.

We consider the following three types of loss functions.

- For the general convex case, we choose the smooth and convex cross-entropy function as the time-varying function:

$$f_t(\mathbf{w}_t; \mathbf{x}_t, y_t) = -y_t \log \sigma(\mathbf{x}_t^\top \mathbf{w}_t) - (1 - y_t) \log(1 - \sigma(\mathbf{x}_t^\top \mathbf{w}_t)),$$

where  $\sigma(\cdot)$  denotes the sigmoid function, and utilize the  $\ell_1$ -norm regularizer  $r(\mathbf{w}_t) = \|\mathbf{w}_t\|_1$ . Therefore, the composite function takes the form of:

$$\phi(\mathbf{w}_t; \mathbf{x}_t, y_t) = -y_t \log \sigma(\mathbf{x}_t^\top \mathbf{w}_t) - (1 - y_t) \log(1 - \sigma(\mathbf{x}_t^\top \mathbf{w}_t)) + \lambda \|\mathbf{w}_t\|_1.$$

- For the strongly convex case, we employ the cross-entropy functions with the  $\ell_2$ -norm regularizer as the time-varying function:

$$f_t(\mathbf{w}_t; \mathbf{x}_t, y_t) = -y_t \log \sigma(\mathbf{x}_t^\top \mathbf{w}_t) - (1 - y_t) \log(1 - \sigma(\mathbf{x}_t^\top \mathbf{w}_t)) + \delta \|\mathbf{w}_t\|_2^2,$$

which is  $2\delta$ -strongly convex, and still leverage the  $\ell_1$ -norm regularizer. Hence, the composite loss function is in the form of:

$$\phi(\mathbf{w}_t; \mathbf{x}_t, y_t) = -y_t \log \sigma(\mathbf{x}_t^\top \mathbf{w}_t) - (1 - y_t) \log(1 - \sigma(\mathbf{x}_t^\top \mathbf{w}_t)) + \delta \|\mathbf{w}_t\|_2^2 + \lambda \|\mathbf{w}_t\|_1$$

where we set the hyper-parameter  $\delta = 0.001$ .

- For the exp-concave case, we utilize the logistic function as the time-varying function:

$$f_t(\mathbf{w}_t; \mathbf{x}_t, y_t) = \log(1 + \exp(-y_t \mathbf{w}_t^\top \mathbf{x}_t))$$

which is exp-concave and smooth [Hazan et al., 2014], and still employ the  $\ell_1$ -norm regularizer. The composite loss function is shown below:

$$\phi(\mathbf{w}_t; \mathbf{x}_t, y_t) = \log(1 + \exp(-y_t \mathbf{w}_t^\top \mathbf{x}_t)) + \lambda \|\mathbf{w}_t\|_1.$$

**Contenders.** For the general convex and strongly convex cases, we compare our methods with OGD [Zinkevich, 2003], COMID [Duchi et al., 2010] and Optimistic-OMD [Chen et al., 2023]. For the exp-concave case, we choose ONS [Hazan et al., 2007], ProxONS [Yang et al., 2024c] and Optimistic-OMD [Chen et al., 2023] as the contenders. All parameters of each method are set according to their theoretical suggestions. For instance, in the general convex case, the learning rate is set as  $\eta = cT^{-1/2}$  in OGD, and  $\eta = cT^{-1/2}$  in COMID,  $\eta_t = D(c + \bar{V}_{t-1})^{-1/2}$  in Optimistic-OMD where  $c$  denotes the hyper-parameter selected from  $\{10^{-3}, 10^{-2}, \dots, 10^4\}$ .

**Results.** All experiments are repeated ten times, and we report the instantaneous loss, the cumulative loss and the average loss against the number of rounds in Figure 1 for the general convex case, Figure 2 for the strongly convex case and Figure 3 for the exp-concave case. From the experimental results, it is evident that adversarial corruptions cause considerable fluctuations in the instantaneous losses across different methods. Moreover, we also observe that, in the three cases of composite SEA, both OptCMD and USC-SEA suffer lower losses compared to baseline methods, demonstrating better performance. This phenomenon can be attributed to their ability to adapt to the composite SEA environment, and their explicit support for handling the non-smooth component  $r(\cdot)$ .

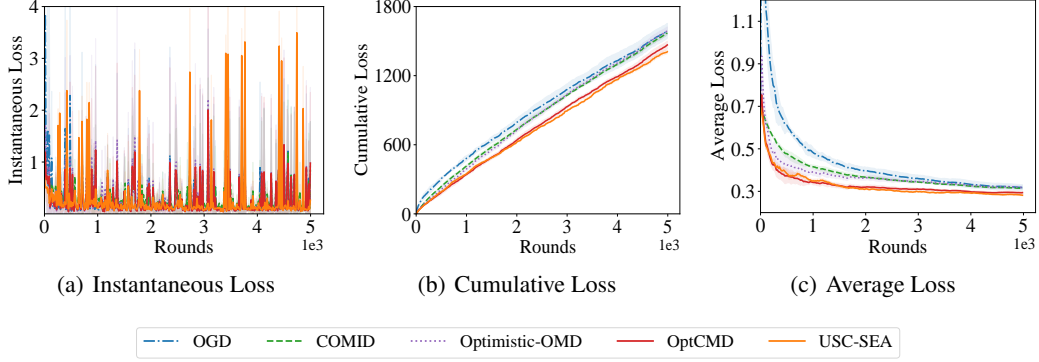


Figure 1: Experimental results for the general convex case.

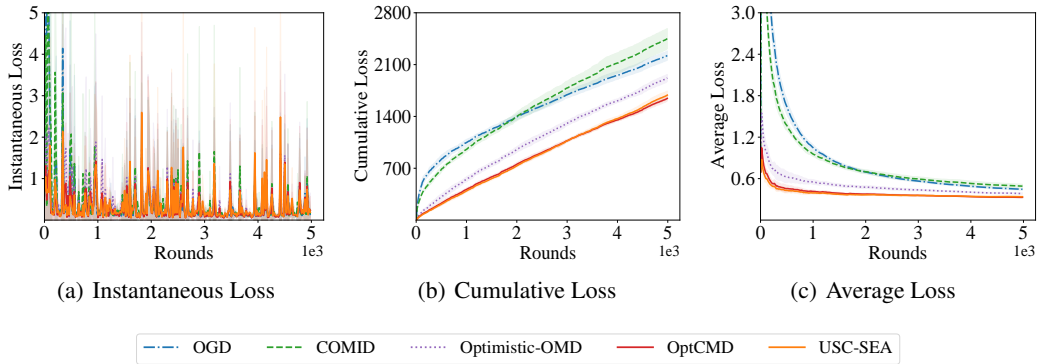


Figure 2: Experimental results for the strongly convex case.

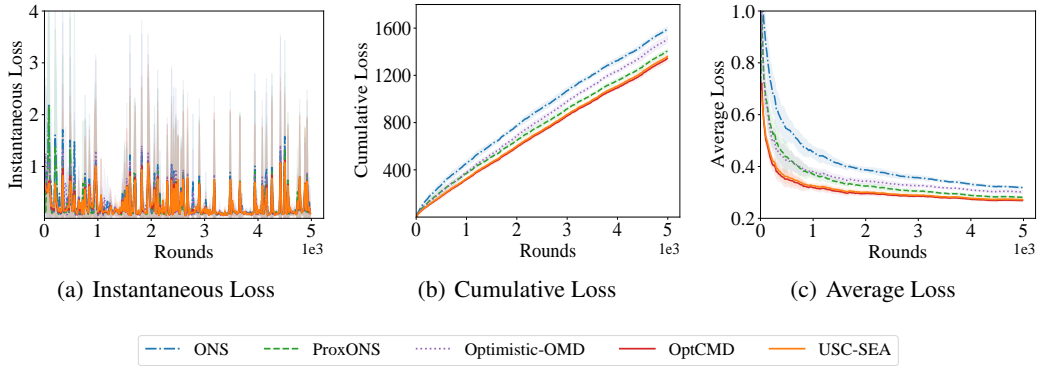


Figure 3: Experimental results for the exp-concave case.

## B Theoretical analysis

### B.1 Proof of Theorem 1

First, since the expected function  $F_t(\cdot)$  is convex in each round  $t$ , we could decompose the instantaneous regret as below:

$$\begin{aligned} \mathbb{E} [f_t(\mathbf{x}_t) + r(\mathbf{x}_t)] - [f_t(\mathbf{x}^*) + r(\mathbf{x}^*)] &= \mathbb{E} [F_t(\mathbf{x}_t) + r(\mathbf{x}_t)] - [F_t(\mathbf{x}^*) + r(\mathbf{x}^*)] \\ &\leq \mathbb{E} [\langle \nabla F_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*)] = \mathbb{E} [\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*)]. \end{aligned} \quad (18)$$



The inequality is due the convexity of  $F_t(\cdot)$  and the last equality is due to the interchangeability of differentiation and integration by Leibniz integral rule. In the following, we shed the light on the right side of (18).

Then, we introduce the following lemma for OptCMD.

**Lemma 1.** *Assume  $\mathcal{R}_t(\cdot)$  is an  $\gamma$ -strongly convex function with respect to the norm  $\|\cdot\|$  and denote by  $\|\cdot\|_*$  the dual norm. According to the updating rule in (4) and (5), for all  $\mathbf{x} \in \mathcal{X}$ , we have*

$$\begin{aligned} \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x} \rangle + r(\mathbf{x}_t) - r(\mathbf{x}) &\leq \gamma^{-1} \|M_t - \nabla f_t(\mathbf{x}_t)\|_*^2 + [\mathcal{B}^{\mathcal{R}_t}(\mathbf{x}, \hat{\mathbf{x}}_t) - \mathcal{B}^{\mathcal{R}_t}(\mathbf{x}, \hat{\mathbf{x}}_{t+1})] \\ &\quad - [\mathcal{B}^{\mathcal{R}_t}(\hat{\mathbf{x}}_{t+1}, \mathbf{x}_t) + \mathcal{B}^{\mathcal{R}_t}(\mathbf{x}_t, \hat{\mathbf{x}}_t)]. \end{aligned} \quad (19)$$

In the general convex setting, we choose  $\eta_t = \min\{D/\sqrt{1 + \bar{V}_{t-1}}, D/\delta\}$ , and  $\|\cdot\| = \frac{1}{\sqrt{\eta_t}} \|\cdot\|_2$ , and  $\|\cdot\|_* = \sqrt{\eta_t} \|\cdot\|_2$ . Therefore, the corresponding Bregman divergence becomes  $\mathcal{B}^{\mathcal{R}_t}(\mathbf{x}, \mathbf{y}) = \frac{1}{2\eta_t} \|\mathbf{x} - \mathbf{y}\|_2^2$  with respect to the  $\eta^{-1}$ -strongly convex function  $\mathcal{R}_t(\mathbf{x}) = \frac{1}{2\eta_t} \|\mathbf{x}\|_2^2$ . Now, we make use of Lemma 1 with  $M_t = \nabla f_{t-1}(\mathbf{x}_{t-1})$ , and obtain

$$\begin{aligned} \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) &\leq \eta_t \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\ &\quad + \frac{1}{2\eta_t} \{ \|\mathbf{x}^* - \hat{\mathbf{x}}_t\|_2^2 - \|\mathbf{x}^* - \hat{\mathbf{x}}_{t+1}\|_2^2 \} - \frac{1}{2\eta_t} \{ \|\hat{\mathbf{x}}_{t+1} - \mathbf{x}_t\|_2^2 + \|\mathbf{x}_t - \hat{\mathbf{x}}_t\|_2^2 \}. \end{aligned}$$

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

$$\begin{aligned} \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) &\leq \underbrace{\sum_{t=1}^T \eta_t \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2}_{\text{term (a)}} \\ &\quad + \underbrace{\sum_{t=1}^T \frac{1}{2\eta_t} \{ \|\mathbf{x}^* - \hat{\mathbf{x}}_t\|_2^2 - \|\mathbf{x}^* - \hat{\mathbf{x}}_{t+1}\|_2^2 \}}_{\text{term (b)}} - \underbrace{\sum_{t=1}^T \frac{1}{2\eta_t} \{ \|\hat{\mathbf{x}}_{t+1} - \mathbf{x}_t\|_2^2 + \|\mathbf{x}_t - \hat{\mathbf{x}}_t\|_2^2 \}}_{\text{term (c)}}. \end{aligned} \quad (20)$$

Now, we have decomposed the upper bound into three terms and will analyze them separately.

To bound the term (a), we introduce the following lemma.

**Lemma 2.** [Pogodin and Lattimore, 2019, Lemma 4.8] *Let  $l_1, \dots, l_T$  be non-negative real numbers with the  $l_t \in [0, B]$ . Then, we have*

$$\sum_{t=1}^T \frac{l_t}{\sqrt{1 + \sum_{s=1}^{t-1} l_s}} \leq 4 \sqrt{1 + \sum_{t=1}^T l_t} + B,$$

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

By utilizing Lemma 2 and the fact that  $\eta_t \leq D/\sqrt{1 + \bar{V}_{t-1}}$ , the term (a) can be upper bounded as follows:

$$\text{term (a)} \leq D \sum_{t=1}^T \frac{\|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2}{\sqrt{1 + \bar{V}_{t-1}}} \leq 4D\sqrt{1 + \bar{V}_T} + 4DG^2. \quad (21)$$

For the term (b), a simple calculation delivers

$$\begin{aligned} \text{term (b)} &= \frac{1}{2\eta_1} \|\mathbf{x}^* - \hat{\mathbf{x}}_1\|_2^2 + \frac{1}{2} \sum_{t=2}^T \left( \frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} \right) \|\mathbf{x}^* - \hat{\mathbf{x}}_t\|_2^2 - \frac{1}{2\eta_T} \|\mathbf{x}^* - \hat{\mathbf{x}}_{T+1}\|_2^2 \\ &\leq \frac{D^2}{2\eta_1} + \frac{D^2}{2} \sum_{t=2}^T \left( \frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} \right) = \frac{D^2}{2\eta_T} = \frac{D}{2} \sqrt{1 + \bar{V}_T} + \frac{D\delta}{2}. \end{aligned} \quad (22)$$

For the term (c), we utilize the fact that  $\eta_t \leq D/\delta$ , and thus obtain

$$\begin{aligned} \text{term (c)} &= \frac{1}{2} \sum_{t=2}^{T+1} \frac{1}{\eta_t} \|\hat{\mathbf{x}}_t - \mathbf{x}_{t-1}\|_2^2 + \frac{1}{2} \sum_{t=1}^T \frac{1}{\eta_t} \|\mathbf{x}_t - \hat{\mathbf{x}}_t\|_2^2 \\ &\geq \frac{\delta}{2D} \sum_{t=2}^T \{ \|\hat{\mathbf{x}}_t - \mathbf{x}_{t-1}\|_2^2 + \|\mathbf{x}_t - \hat{\mathbf{x}}_t\|_2^2 \} \geq \frac{\delta}{4D} \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2. \end{aligned} \quad (23)$$

The last step is due to the domain bounded assumption, i.e., Assumption 1. Combining (21), (22) and (23), we obtain that

$$\sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) \leq \frac{3D}{2} \sqrt{1 + \bar{V}_T} + \frac{D\delta}{2} - \frac{\delta}{4D} \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + 4DG^2. \quad (24)$$

To bound the term  $\bar{V}_T$ , we exploiting the following lemma.

**Lemma 3.** *Under Assumption 2 and 4, we have*

$$\begin{aligned} \bar{V}_T &\leq G^2 + 8 \sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 \\ &\quad + 4 \sum_{t=2}^T \|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 + 4H^2 \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2. \end{aligned} \quad (25)$$

Plugging (25) into (24) obtains

$$\begin{aligned} &\sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t) + \nabla r(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle \\ &\leq \frac{3(D+GD)}{2} + 4DG^2 + \frac{D\delta}{2} + 6HD \sqrt{\sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2} - \frac{\delta}{4D} \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \\ &\quad + 12D \sqrt{\sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2} + 6D \sqrt{\sum_{t=2}^T \|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2}. \end{aligned}$$

Then, we apply the AM-GM inequality, i.e.,  $6HD \sqrt{\sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2} \leq \frac{36H^2D^3}{\delta} + \frac{\delta}{4D} \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2$ , and obtain

$$\begin{aligned} \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t) + \nabla r(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle &\leq 2(D+GD) + 4DG^2 + \frac{D\delta}{2} + \frac{36H^2D^3}{\delta} \\ &\quad + 12D \sqrt{\sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2} + 6D \sqrt{\sum_{t=2}^T \|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2}. \end{aligned}$$

By setting  $\delta = 6\sqrt{2}HD$ , we have

$$\begin{aligned} \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t) + \nabla r(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle &\leq C_1 + 12D \sqrt{\sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2} \\ &\quad + 6D \sqrt{\sum_{t=2}^T \|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2}, \end{aligned}$$

where  $C_1 = 2(D + GD) + 4DG^2 + 6\sqrt{2}HD^2$ . Next, according to the definition of  $\sigma_{1:T}^2$  and  $\Sigma_{1:T}^2$ , the expected regret can be upper bound as follows:

$$\begin{aligned} \mathbb{E}[\text{Regret}_T] &\leq \mathbb{E} \left[ \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t) + \nabla r(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle \right] \\ &\leq C_1 + 12D\sqrt{\sigma_{1:T}^2} + 6D\sqrt{\Sigma_{1:T}^2} = \mathcal{O} \left( \sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2} \right). \end{aligned}$$

## B.2 Proof of Theorem 2

First, according to (18), we have

$$\mathbb{E} [[f_t(\mathbf{x}_t) + r(\mathbf{x}_t)] - [f_t(\mathbf{x}^*) + r(\mathbf{x}^*)]] \leq \mathbb{E} [\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*)]. \quad (26)$$

In this setting, we still choose  $\|\cdot\| = \frac{1}{\sqrt{\eta_t}} \|\cdot\|_2$ , and  $\|\cdot\|_* = \sqrt{\eta_t} \|\cdot\|_2$ . Hence, the function  $\mathcal{R}_t(\mathbf{x}) = \frac{1}{2\eta_t} \|\mathbf{x}\|_2^2$  is  $\eta_t^{-1}$ -strongly convex and the Bregman divergence becomes  $\mathcal{B}^{\mathcal{R}_t}(\mathbf{x}, \mathbf{y}) = \frac{1}{2\eta_t} \|\mathbf{x} - \mathbf{y}\|_2^2$ . Then, applying Lemma 1, we obtain

$$\begin{aligned} \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) &\leq \eta_t \|M_{t+1} - \nabla f_t(\mathbf{x}_t)\|_2^2 + \frac{1}{2\eta_t} \|\mathbf{x}^* - \hat{\mathbf{x}}_t\|_2^2 \\ &\quad - \frac{1}{2\eta_t} \|\mathbf{x}^* - \hat{\mathbf{x}}_{t+1}\|_2^2 - \frac{1}{2\eta_t} \|\hat{\mathbf{x}}_{t+1} - \mathbf{x}_t\|_2^2 - \frac{1}{2\eta_t} \|\mathbf{x}_t - \hat{\mathbf{x}}_t\|_2^2. \end{aligned}$$

Following the configuration in Scroccaro et al. [2023], we set  $M_t = \nabla f_{t-1}(\hat{\mathbf{x}}_t)$  and  $\eta_t = (4H^2 + \bar{D}_t)^{-1/2}$  where  $\bar{D}_t = \sum_{s=1}^t \|\nabla f_t(\hat{\mathbf{x}}_t) - \nabla f_{t-1}(\hat{\mathbf{x}}_t)\|_2^2$ . According to [Scroccaro et al., 2023, Theorem 2.5], we have

$$\sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) \leq \left(5 + \frac{3}{2}D^2\right) \sqrt{4H^2 + \bar{D}_T}.$$

To bound the term  $\bar{D}_T$ , we introduce the following lemma.

**Lemma 4.** *Under Assumption 2, we have*

$$\bar{D}_T \leq G^2 + 6 \sum_{t=1}^T \sup_{\mathbf{x} \in X} \|\nabla f_t(\mathbf{x}) - \nabla F_t(\mathbf{x})\|_2^2 + 4 \sum_{t=1}^T \sup_{\mathbf{x} \in X} \|\nabla F_t(\mathbf{x}) - \nabla F_{t-1}(\mathbf{x})\|_2^2.$$

By applying Lemma 4, we obtain

$$\begin{aligned} &\sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) \\ &\leq \left(5 + \frac{3}{2}D^2\right) \left( \sqrt{4H^2 + G^2} + 3 \sqrt{\sum_{t=1}^T \sup_{\mathbf{x} \in X} \|\nabla f_t(\mathbf{x}) - \nabla F_t(\mathbf{x})\|_2^2} \right. \\ &\quad \left. + 2 \sqrt{\sum_{t=1}^T \sup_{\mathbf{x} \in X} \|\nabla F_t(\mathbf{x}) - \nabla F_{t-1}(\mathbf{x})\|_2^2} \right). \end{aligned} \quad (27)$$

Substituting (27) into (26) and taking the expectation on both sides finishes the proof.

## B.3 Proof of Theorem 3

In the beginning, we state the parameter configuration under the strongly convex setting (i.e.,  $F_t(\cdot)$  is strongly convex and  $r(\cdot)$  is general convex). The learning rate is set as  $\eta_t = \frac{2}{\delta + \lambda t}$  and the norm is chosen as  $\|\cdot\| = \frac{1}{\sqrt{\eta_t}} \|\cdot\|_2$ , with the dual norm  $\|\cdot\|_* = \sqrt{\eta_t} \|\cdot\|_2$ . Therefore, the function becomes  $\mathcal{R}_t(\mathbf{x}) = \frac{1}{2\eta_t} \|\mathbf{x}\|_2^2$  and the corresponding Bregman divergence becomes  $\mathcal{B}^{\mathcal{R}_t}(\mathbf{x}, \mathbf{y}) = \frac{1}{2\eta_t} \|\mathbf{x} - \mathbf{y}\|_2^2$ .

Then, we make use of the strong convexity of  $F_t(\cdot)$ , and decompose the instantaneous regret as below:

$$\begin{aligned}
& \mathbb{E} [[f_t(\mathbf{x}_t) + r(\mathbf{x}_t)] - [f_t(\mathbf{x}^*) + r(\mathbf{x}^*)]] = \mathbb{E} [[F_t(\mathbf{x}_t) + r(\mathbf{x}_t)] - [F_t(\mathbf{x}^*) + r(\mathbf{x}^*)]] \\
& \leq \mathbb{E} \left[ \langle \nabla F_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) - \frac{\lambda}{2} \|\mathbf{x}^* - \mathbf{x}_t\|_2^2 \right] \\
& = \mathbb{E} \left[ \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) - \frac{\lambda}{2} \|\mathbf{x}^* - \mathbf{x}_t\|_2^2 \right].
\end{aligned} \tag{28}$$

Similar to the analysis in Theorem 1, we then focus on the the right side of (28). By utilizing Lemma 1, we have

$$\begin{aligned}
& \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) - \frac{\lambda}{2} \sum_{t=1}^T \|\mathbf{x}^* - \mathbf{x}_t\|_2^2 \leq \underbrace{\sum_{t=1}^T \eta_t \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2}_{\text{term (a)}} \\
& + \underbrace{\sum_{t=1}^T \left[ \frac{1}{2\eta_t} \{ \|\mathbf{x}^* - \hat{\mathbf{x}}_t\|_2^2 - \|\mathbf{x}^* - \hat{\mathbf{x}}_{t+1}\|_2^2 \} - \frac{\lambda}{2} \|\mathbf{x}^* - \mathbf{x}_t\|_2^2 \right]}_{\text{term (b)}} - \underbrace{\sum_{t=1}^T \frac{1}{2\eta_t} \{ \|\hat{\mathbf{x}}_{t+1} - \mathbf{x}_t\|_2^2 + \|\mathbf{x}_t - \hat{\mathbf{x}}_t\|_2^2 \}}_{\text{term (c)}}.
\end{aligned}$$

Next, we analyze the above three terms separately. For the term (a), we substitute  $\eta_t \leq \frac{2}{\lambda t}$  and obtain

$$\text{term (a)} \leq \sum_{t=1}^T \frac{2}{\lambda t} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2. \tag{29}$$

For the term (b), we exploit one result of Lemma 1 in (52), i.e.,

$$\|\mathbf{x}_t - \hat{\mathbf{x}}_{t+1}\|_2 \leq \eta_t \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2, \tag{30}$$

with  $\|\cdot\| = \frac{1}{\sqrt{\eta_t}} \|\cdot\|_2$  and  $\|\cdot\|_* = \sqrt{\eta_t} \|\cdot\|_2$ . By utilizing (30), we can upper bound the term (b) as following:

$$\begin{aligned}
\text{term (b)} & \leq \frac{D^2}{2\eta_1} + \sum_{t=2}^T \left( \frac{1}{2\eta_t} - \frac{1}{2\eta_{t-1}} \right) \|\mathbf{x}^* - \hat{\mathbf{x}}_t\|_2^2 - \frac{\lambda}{2} \sum_{t=1}^T \|\mathbf{x}^* - \mathbf{x}_t\|_2^2 \\
& \leq \frac{D^2\lambda}{4} + \frac{\lambda}{4} \sum_{t=1}^{T-1} [\|\mathbf{x}^* - \hat{\mathbf{x}}_{t+1}\|_2^2 - 2\|\mathbf{x}^* - \mathbf{x}_t\|_2^2] \\
& \leq \frac{D^2\lambda}{4} + \frac{\lambda}{2} \sum_{t=1}^{T-1} \|\mathbf{x}_t - \hat{\mathbf{x}}_{t+1}\|_2^2 \stackrel{(30)}{\leq} \frac{D^2\lambda}{4} + \frac{\lambda}{2} \sum_{t=1}^{T-1} \eta_t^2 \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\
& \leq \frac{D^2\lambda}{4} + \frac{\lambda\eta_1}{2} \sum_{t=1}^{T-1} \eta_t \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2 \leq \frac{D^2\lambda}{4} + \text{term (a)}.
\end{aligned} \tag{31}$$

For the term (c), we make use of the non-increasing property of  $\eta_t$  and  $\eta_t \leq 2/\delta$ , and obtain

$$\begin{aligned}
\text{term (c)} & \geq \sum_{t=2}^T \left\{ \frac{1}{2\eta_{t-1}} \|\hat{\mathbf{x}}_t - \mathbf{x}_{t-1}\|_2^2 + \frac{1}{2\eta_{t-1}} \|\mathbf{x}_t - \hat{\mathbf{x}}_t\|_2^2 \right\} \\
& \geq \sum_{t=2}^T \frac{1}{4\eta_{t-1}} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \geq \frac{\delta}{8} \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2.
\end{aligned} \tag{32}$$

Putting (29), (31) and (32) together, we obtain

$$\begin{aligned}
& \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) - \frac{\lambda}{2} \sum_{t=1}^T \|\mathbf{x}^* - \mathbf{x}_t\|_2^2 \\
& \leq \frac{D^2\lambda}{4} + \sum_{t=1}^T \frac{4}{\lambda t} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 - \frac{\delta}{8} \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2.
\end{aligned} \tag{33}$$

Now, we make use of a byproduct (55) from Lemma 3, i.e.,

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

and thus (33) becomes

$$\begin{aligned} &\mathbb{E} \left[ \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) - \frac{\lambda}{2} \sum_{t=1}^T \|\mathbf{x}^* - \mathbf{x}_t\|_2^2 \right] \\ &\leq \sum_{t=1}^T \frac{16}{\lambda t} (2\sigma_t^2 + \sup_{\mathbf{x} \in \mathcal{X}} \|\nabla F_t(\mathbf{x}) - \nabla F_{t-1}(\mathbf{x})\|_2^2) \\ &\quad + \sum_{t=1}^T \frac{16H^2}{\lambda t} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 - \frac{\delta}{8} \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + \frac{D^2\lambda}{4}. \end{aligned} \quad (34)$$

To bound the first two terms, we introduce the following lemma.

**Lemma 5.** [Yan et al., 2023, Lemma 9] For a sequence of  $\{a_t\}_{t=1}^T$  and  $b$ , where  $a_t, b > 0$  for any  $t \in [T]$ , denoting by  $a_{\max} \triangleq \max_t a_t$  and  $A \triangleq \lceil b \sum_{t=1}^T a_t \rceil$ , we have

$$\sum_{t=1}^T \frac{a_t}{bt} \leq \frac{a_{\max}}{b} (1 + \ln A) + \frac{1}{b^2}.$$

Let  $a_t = 2\sigma_t^2 + \sup_{\mathbf{x} \in \mathcal{X}} \|\nabla F_t(\mathbf{x}) - \nabla F_{t-1}(\mathbf{x})\|_2^2$ ,  $a_{\max} = 2\sigma_{\max}^2 + \Sigma_{\max}^2$ ,  $b = \lambda$  and  $A = \lceil \lambda(2\sigma_{1:T}^2 + \Sigma_{1:T}^2) \rceil$ . Then, applying Lemma 5, we obtain

$$\sum_{t=1}^T \frac{1}{\lambda t} (2\sigma_t^2 + \sup_{\mathbf{x} \in \mathcal{X}} \|\nabla F_t(\mathbf{x}) - \nabla F_{t-1}(\mathbf{x})\|_2^2) \leq \frac{2\sigma_{\max}^2 + \Sigma_{\max}^2}{\lambda} (1 + \ln(1 + \lambda(2\sigma_{1:T}^2 + \Sigma_{1:T}^2))) + \frac{1}{\lambda^2}. \quad (35)$$

Let  $a_t = \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2$ ,  $a_{\max} = D^2$ ,  $b = \lambda$  and  $A = \lceil \lambda \sum_{t=1}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \rceil$ . Applying Lemma 5, we obtain

$$\begin{aligned} \sum_{t=1}^T \frac{1}{\lambda t} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 &\leq \frac{D^2}{\lambda} \left( 1 + \ln(1 + \lambda \sum_{t=1}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2) \right) + \frac{1}{\lambda^2} \\ &\leq \frac{D^2}{\lambda} \left( 1 + \lambda \sum_{t=1}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \right) + \frac{1}{\lambda^2}, \end{aligned} \quad (36)$$

where the last step is due to  $\ln(1+x) \leq x$  for any  $x \geq 0$ .

Substituting (35) and (36) into (34), we have

$$\begin{aligned} &\mathbb{E} \left[ \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) - \frac{\lambda}{2} \sum_{t=1}^T \|\mathbf{x}^* - \mathbf{x}_t\|_2^2 \right] \\ &\leq \mathcal{O} \left( \frac{\sigma_{\max}^2 + \Sigma_{\max}^2}{\lambda} \ln(\sigma_{1:T}^2 + \Sigma_{1:T}^2) \right) + \left( 16H^2D^2 - \frac{\delta}{8} \right) \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + \mathcal{O}(1). \end{aligned}$$

Setting  $\delta = 128H^2D^2$  finishes the proof.

#### B.4 Proof of Theorem 4

Similar to (28) in Theorem 3, the instantaneous regret is upper bounded by

$$\mathbb{E} [f_t(\mathbf{x}_t) + r(\mathbf{x}_t)] - [f_t(\mathbf{x}^*) + r(\mathbf{x}^*)] = \mathbb{E} \left[ \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) - \frac{\lambda}{2} \|\mathbf{x}^* - \mathbf{x}_t\|_2^2 \right].$$

According to [Scroccaro et al., 2023, Theorem 2.9], OptCMD with the configuration in (9) ensures

$$\begin{aligned} & \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) - \frac{\lambda}{2} \|\mathbf{x}^* - \mathbf{x}_t\|_2^2 \\ & \leq 2HD^2 + \frac{G^2}{H} + \frac{4G^2}{\lambda} \log \left( 1 + \frac{\lambda}{4HG^2} \bar{D}_T \right). \end{aligned} \quad (37)$$

For the term  $\log \left( 1 + \frac{\lambda}{4HG^2} \bar{D}_T \right)$ , we apply Lemma 4 and obtain

$$\begin{aligned} & \log \left( 1 + \frac{\lambda}{4HG^2} \bar{D}_T \right) \leq \log \left( 1 + \frac{\lambda}{4H} \right) \\ & + \log \left( 1 + \frac{\lambda}{8HG^2} \left( 3 \sum_{t=1}^T \sup_{\mathbf{x} \in X} \|\nabla f_t(\mathbf{x}) - \nabla F_t(\mathbf{x})\|_2^2 + 2 \sum_{t=1}^T \sup_{\mathbf{x} \in X} \|\nabla F_t(\mathbf{x}) - \nabla F_{t-1}(\mathbf{x})\|_2^2 \right) \right). \end{aligned}$$

Taking expectation over both sides of the above inequality delivers

$$\mathbb{E} \left[ \log \left( 1 + \frac{\lambda}{4HG^2} \bar{D}_T \right) \right] \leq \log \left( 1 + \frac{\lambda}{4H} \right) + \log \left( 1 + \frac{\lambda}{8HG^2} (3\bar{\sigma}_{1:T}^2 + 2\Sigma_{1:T}^2) \right)$$

Plug the above result into (37) finishes the proof.

## B.5 Proof of Theorem 5

For the exp-concave setting (i.e.,  $f_t(\cdot)$  is exp-concave and  $r(\cdot)$  is general convex), our parameter configuration is follow Chiang et al. [2012]. To be precise, we define  $H_t = \delta I + \frac{\beta}{2} G^2 I + \frac{\beta}{2} \sum_{r=1}^{t-1} h_r$ , where  $h_r = \nabla f_r(\mathbf{x}_r) \nabla f_r(\mathbf{x}_r)^\top$ . The norm is set as  $\|\cdot\| = \|\cdot\|_{H_t}$  and  $\|\cdot\|_* = \|\cdot\|_{H_t^{-1}}$ , and the function becomes  $\mathcal{R}_t(\mathbf{x}) = \frac{1}{2} \|\mathbf{x}\|_{H_t}^2$ . With this  $\mathcal{R}_t(\mathbf{x})$ , the corresponding Bregman divergence becomes  $\mathcal{B}^{\mathcal{R}_t}(\mathbf{x}, \mathbf{y}) = \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_{H_t}^2$ .

First, we introduce a common property of exp-concave function, as shown below.

**Lemma 6.** [Hazan et al., 2007, Lemma 3] Under Assumption 1 and 2, if  $f(\cdot)$  is exp-concave over  $\mathcal{X}$ , we have

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{\beta}{2} \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle^2,$$

for  $\forall \mathbf{x}, \mathbf{y} \in \mathcal{X}$  and  $\beta = \frac{1}{2} \min\{\frac{1}{4GD}, \alpha\}$ .

Then, according to the exponential concavity of  $f_t(\cdot)$  shown in Lemma 6, we can decompose the instantaneous regret as below:

$$\begin{aligned} & \mathbb{E} [[f_t(\mathbf{x}_t) + r(\mathbf{x}_t)] - [f_t(\mathbf{x}^*) + r(\mathbf{x}^*)]] \\ & \leq \mathbb{E} \left[ \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) - \frac{\beta}{2} \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}^* - \mathbf{x}_t \rangle^2 \right]. \end{aligned}$$

Recall that we denote  $h_t = \nabla f_t(\mathbf{x}_t) \nabla f_t(\mathbf{x}_t)^\top$ . Hence, the last term in (38) can be rewrite as  $\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}^* - \mathbf{x}_t \rangle^2 = \|\mathbf{x}^* - \mathbf{x}_t\|_{h_t}^2$ .

Now, we focus on the right side of (38). By utilizing Lemma 1, we have

$$\begin{aligned} & \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) - \frac{\beta}{2} \sum_{t=1}^T \|\mathbf{x}^* - \mathbf{x}_t\|_{h_t}^2 \leq \underbrace{\sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_{H_t^{-1}}^2}_{\text{term (a)}} \\ & + \underbrace{\frac{1}{2} \sum_{t=1}^T [\|\mathbf{x}^* - \hat{\mathbf{x}}_t\|_{H_t}^2 - \|\mathbf{x}^* - \hat{\mathbf{x}}_{t+1}\|_{H_t}^2 - \beta \|\mathbf{x}^* - \mathbf{x}_t\|_{h_t}^2]}_{\text{term (b)}} - \underbrace{\frac{1}{2} \sum_{t=1}^T [\|\hat{\mathbf{x}}_{t+1} - \mathbf{x}_t\|_{H_t}^2 + \|\mathbf{x}_t - \hat{\mathbf{x}}_t\|_{H_t}^2]}_{\text{term (c)}}. \end{aligned}$$

The term (a) can be upper bounded as shown in the following lemma.

**Lemma 7.** *Let  $d$  be the dimension. Then, we have*

$$\text{term (a)} \leq \frac{8d}{\beta} \ln \left( 1 + \frac{\beta}{8\delta} \sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \right). \quad (38)$$

For the term (b), we exploit the fact that  $H_{t+1} - H_t = \frac{\beta}{2}h_t$  and obtain

$$\begin{aligned} \text{term (b)} &= \frac{1}{2} \left[ \|\mathbf{x}^* - \hat{\mathbf{x}}_1\|_{H_1}^2 - \|\mathbf{x}^* - \hat{\mathbf{x}}_{T+1}\|_{H_{T+1}}^2 \right] \\ &\quad + \frac{1}{2} \sum_{t=1}^T \left\{ \|\mathbf{x}^* - \hat{\mathbf{x}}_{t+1}\|_{H_{t+1}}^2 - \|\mathbf{x}^* - \hat{\mathbf{x}}_{t+1}\|_{H_t}^2 - \beta \|\mathbf{x}^* - \mathbf{x}_t\|_{h_t}^2 \right\} \\ &= \frac{1}{2} \left[ \|\mathbf{x}^* - \hat{\mathbf{x}}_1\|_{H_1}^2 - \|\mathbf{x}^* - \hat{\mathbf{x}}_{T+1}\|_{H_{T+1}}^2 \right] + \frac{\beta}{4} \sum_{t=1}^T \left\{ \|\mathbf{x}^* - \hat{\mathbf{x}}_{t+1}\|_{h_t}^2 - 2\|\mathbf{x}^* - \mathbf{x}_t\|_{h_t}^2 \right\} \\ &\stackrel{(1)}{\leq} \left( \frac{\delta}{2} + \frac{\beta}{4}G^2 \right) D^2 + \frac{\beta}{4} \sum_{t=1}^T \left\{ \|\mathbf{x}^* - \hat{\mathbf{x}}_{t+1}\|_{h_t}^2 - 2\|\mathbf{x}^* - \mathbf{x}_t\|_{h_t}^2 \right\}. \end{aligned} \quad (39)$$

To proceed the proof, we introduce the following lemma.

**Lemma 8.** *[Chiang et al., 2012, Proposition 1] For any  $y, z \in \mathbb{R}^N$  and any PSD  $H \in \mathbb{R}^{N \times N}$ , we have  $\|y + z\|_H^2 \leq 2\|y\|_H^2 + 2\|z\|_H^2$ .*

Applying Lemma 8 on (39) obtains:

$$\begin{aligned} \text{term (b)} &\leq \left( \frac{\delta}{2} + \frac{\beta}{4}G^2 \right) D^2 + \frac{\beta}{4} \sum_{t=1}^T \left\{ 2\|\mathbf{x}^* - \mathbf{x}_t\|_{h_t}^2 + 2\|\mathbf{x}_t - \hat{\mathbf{x}}_{t+1}\|_{h_t}^2 - 2\|\mathbf{x}^* - \mathbf{x}_t\|_{h_t}^2 \right\} \\ &\leq \left( \frac{\delta}{2} + \frac{\beta}{4}G^2 \right) D^2 + \frac{\beta}{2} \sum_{t=1}^T \|\mathbf{x}_t - \hat{\mathbf{x}}_{t+1}\|_{h_t}^2 \leq \left( \frac{1}{2} + \frac{\beta}{4}G^2 \right) D^2 + \sum_{t=1}^T \|\mathbf{x}_t - \hat{\mathbf{x}}_{t+1}\|_{H_t}^2 \\ &\leq \left( \frac{\delta}{2} + \frac{\beta}{4}G^2 \right) D^2 + \sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_{H_t^{-1}}, \end{aligned} \quad (40)$$

where the penultimate step is due to  $H_t \succeq \frac{\beta}{2}G^2I \succeq \frac{\beta}{2}h_t$ , and the last step is due to

$$\|\mathbf{x}_t - \hat{\mathbf{x}}_{t+1}\|_{H_t} \leq \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_{H_t^{-1}}, \quad (41)$$

which could be obtained from Lemma 1 with  $\|\cdot\| = \|\cdot\|_{H_t}$  and  $\|\cdot\|_* = \|\cdot\|_{H_t^{-1}}$ .

For the term (c), we have

$$\begin{aligned} \text{term (c)} &= \frac{1}{2} \sum_{t=2}^{T+1} \|\hat{\mathbf{x}}_t - \mathbf{x}_{t-1}\|_{H_{t-1}}^2 + \frac{1}{2} \sum_{t=1}^T \|\mathbf{x}_t - \hat{\mathbf{x}}_t\|_{H_t}^2 \\ &\geq \frac{\delta}{2} \sum_{t=2}^T \left\{ \|\hat{\mathbf{x}}_t - \mathbf{x}_{t-1}\|_2^2 + \|\mathbf{x}_t - \hat{\mathbf{x}}_t\|_2^2 \right\} \geq \frac{\delta}{4} \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2, \end{aligned} \quad (42)$$

where the first inequality is due to  $H_t \succeq H_{t-1} \succeq \delta I$ ,  $\|\hat{\mathbf{x}}_{T+1} - \mathbf{x}_T\|_{H_{T+1}}^2 \geq 0$  and  $\|\mathbf{x}_1 - \hat{\mathbf{x}}_1\|_{H_1}^2 \geq 0$ . Combining (38), (40) and (42) obtains

$$\begin{aligned} &\sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) - \frac{\beta}{2} \sum_{t=1}^T \|\mathbf{x}^* - \mathbf{x}_t\|_{h_t}^2 \\ &\leq \left( \frac{\delta}{2} + \frac{\beta}{4}G^2 \right) D^2 + \frac{24d}{\beta} \ln \left( 1 + \frac{\beta}{8\delta} \sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \right) - \frac{\delta}{4} \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2. \end{aligned}$$

Next, we leverage Lemma 3 and arrive at

$$\begin{aligned}
& \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) - \frac{\beta}{2} \sum_{t=1}^T \|\mathbf{x}^* - \mathbf{x}_t\|_{h_t}^2 \\
& \leq \left( \frac{\delta}{2} + \frac{\beta}{4} G^2 \right) D^2 - \frac{\delta}{4} \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + \frac{16d}{\beta} \ln \left( 1 + \frac{\beta}{8\delta} \left[ G^2 + 8 \sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 \right. \right. \\
& \quad \left. \left. + 4 \sum_{t=2}^T \|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 \right] + \frac{\beta H^2}{2\delta} \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \right) \\
& \leq \left( \frac{\delta}{2} + \frac{\beta}{4} G^2 \right) D^2 - \frac{\delta}{4} \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + \frac{16d}{\beta} \ln \left( 1 + \frac{\beta}{8\delta} \left[ G^2 + 8 \sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 \right. \right. \\
& \quad \left. \left. + 4 \sum_{t=2}^T \|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 \right] \right) + \frac{16d}{\beta} \ln \left( 1 + \frac{\beta H^2}{2\delta} \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \right), \tag{43}
\end{aligned}$$

where the last step is due to the inequality

$$\ln(1 + u + v) \leq \ln(1 + u) + \ln(1 + v), \quad \forall u, v \geq 0. \tag{44}$$

To simplify (43), we make use of the following lemma.

**Lemma 9.** [Chen et al., 2023, Lemma 7] Let  $A \geq 0$ ,  $a \geq 0$ ,  $b \geq 0$  and  $c > 0$ , we have

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

From Lemma 9, we have

$$\frac{16d}{\beta} \ln \left( \frac{\beta H^2}{2\delta} \sum_{t=1}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + 1 \right) - \frac{\delta}{4} \sum_{t=1}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \leq \frac{16d}{\beta} \ln(32dH^2 + 1), \tag{45}$$

where we set  $\delta = 1$  in the last step. Combining (43) and (45), we have

$$\begin{aligned}
& \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t) + \nabla r(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle - \frac{\beta}{2} \sum_{t=1}^T \|\mathbf{x}^* - \mathbf{x}_t\|_{h_t}^2 \\
& \leq \left( \frac{1}{2} + \frac{\beta}{4} G^2 \right) D^2 + \frac{16d}{\beta} \ln(32dH^2 + 1) + \frac{16d}{\beta} \ln \left( 1 + \frac{\beta}{8} \left[ G^2 + 8 \sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 \right. \right. \\
& \quad \left. \left. + 4 \sum_{t=2}^T \|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 \right] \right).
\end{aligned}$$

Taking the expectation over both sides delivers

$$\begin{aligned}
& \mathbb{E} \left[ \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t) + \nabla r(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle - \frac{\beta}{2} \sum_{t=1}^T \|\mathbf{x}^* - \mathbf{x}_t\|_{h_t}^2 \right] \\
& \leq \left( \frac{1}{2} + \frac{\beta}{4} G^2 \right) D^2 + \frac{16d}{\beta} \ln(32dH^2 + 1) + \frac{16d}{\beta} \ln \left( 1 + \frac{\beta}{8} G^2 + \beta \sigma_{1:T}^2 + \frac{\beta}{2} \Sigma_{1:T}^2 \right),
\end{aligned}$$

where the last step is due to the Jensen's inequality. Finally, we complete the proof by substitute the above result into (38).

## B.6 Proof of Theorem 6

In this part, we analyze three types of functions separately.



### B.6.1 For convex functions

First, since the expected function  $F_t(\cdot)$  is convex in each round  $t$ , the instantaneous dynamic regret could be upper bounded as below:

$$\begin{aligned} & \mathbb{E} [[f_t(\mathbf{x}_t) + r(\mathbf{x}_t)] - [f_t(\mathbf{x}^*) + r(\mathbf{x}^*)]] = \mathbb{E} [[F_t(\mathbf{x}_t) + r(\mathbf{x}_t)] - [F_t(\mathbf{x}^*) + r(\mathbf{x}^*)]] \\ & \leq \mathbb{E} [\langle \nabla F_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*)] = \mathbb{E} [\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*)]. \end{aligned}$$

Then, we decompose the regret into meta-regret and expert-regret:

$$\begin{aligned} \mathbb{E} [\text{Regret}_T] & \leq \mathbb{E} \left[ \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) \right] \\ & = \underbrace{\mathbb{E} \left[ \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}_t^{k,i} \rangle + r(\mathbf{x}_t) - r(\mathbf{x}_t^{k,i}) \right]}_{\text{meta-regret}} + \underbrace{\mathbb{E} \left[ \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t^{k,i} - \mathbf{x}^* \rangle + r(\mathbf{x}_t^{k,i}) - r(\mathbf{x}^*) \right]}_{\text{expert-regret}}. \end{aligned}$$

**Expert-regret analysis.** According to the above decomposition, we define the surrogate loss functions  $h_t^i(\mathbf{x}) = \langle \nabla f_t(\mathbf{x}_t), \mathbf{x} \rangle + r(\mathbf{x})$  for the  $i$ -th expert. According to (24), we can bound the expert-regret by

$$\begin{aligned} \text{expert-regret} & \leq \mathbb{E} \left[ \frac{3D}{2} \sqrt{1 + \bar{V}_T} + \frac{D\delta}{2} - \frac{\delta}{4D} \sum_{t=2}^T \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 + 4DG^2 \right] \\ & \leq \frac{3D}{2} (1 + G) + 4DG^2 + 12D \sqrt{\sigma_{1:T}^2} + 6D \sqrt{\Sigma_{1:T}^2} \\ & \quad + 10HD \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 - \frac{\delta}{4D} \sum_{t=2}^T \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2, \end{aligned}$$

where the last step is due to Lemma 3.

**Meta-regret analysis.** We first consider the instantaneous meta-regret:

$$\begin{aligned} & \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}_t^{k,i} \rangle + r(\mathbf{x}_t) - r(\mathbf{x}_t^{k,i}) \\ & = \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}_t^k \rangle + r(\mathbf{x}_t) - r(\mathbf{x}_t^k) + \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t^k - \mathbf{x}_t^{k,i} \rangle + r(\mathbf{x}_t^k) - r(\mathbf{x}_t^{k,i}) \\ & \leq \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}_t^k \rangle + \sum_{k=1}^K q_t^k r(\mathbf{x}_t^k) - r(\mathbf{x}_t^k) + \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t^k - \mathbf{x}_t^{k,i} \rangle + \sum_{i=1}^N p_t^{k,i} r(\mathbf{x}_t^{k,i}) - r(\mathbf{x}_t^{k,i}) \\ & \leq \sum_{k=1}^K q_t^k l_t^k - l_t^k + \sum_{i=1}^N p_t^{k,i} l_t^{k,i} - l_t^{k,i} - \gamma_1 \sum_{k=1}^K q_t^k \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2 + \gamma_1 \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2 \\ & \quad - \gamma_2 \sum_{i=1}^N p_t^{k,i} \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 + \gamma_2 \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 \\ & = \langle \mathbf{q}_t - \mathbf{e}_k, \mathbf{l}_t \rangle + \langle \mathbf{p}_t^k - \mathbf{e}_i, \mathbf{l}_t^k \rangle - \gamma_1 \sum_{k=1}^K q_t^k \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2 + \gamma_1 \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2 \\ & \quad - \gamma_2 \sum_{i=1}^N p_t^{k,i} \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 + \gamma_2 \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2, \end{aligned} \tag{46}$$

where the second step is due to Jensen's inequality, i.e.,  $r(\mathbf{x}_t) = r(\sum_{k=1}^K q_t^k \mathbf{x}_t^k) \leq \sum_{k=1}^K q_t^k r(\mathbf{x}_t^k)$  and  $r(\mathbf{x}_t^k) = r(\sum_{i=1}^N p_t^{k,i} \mathbf{x}_t^{k,i}) \leq \sum_{i=1}^N p_t^{k,i} r(\mathbf{x}_t^{k,i})$ , and the third step is due to the definition of  $l_t^k$  and  $l_t^{k,i}$  in (13) and (14). For brevity, we denote  $\mathbf{q}_t \triangleq (q_t^1, \dots, q_t^K)$ ,  $\mathbf{l}_t \triangleq (l_t^1, \dots, l_t^K)$ ,

$\mathbf{p}_t^k \triangleq (q_t^{k,1}, \dots, q_t^{k,N})$  and  $\mathbf{l}_t^k \triangleq (l_t^{k,1}, \dots, l_t^{k,N})$ . By the above results, we have

$$\begin{aligned} & \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}_t^{k,i} \rangle + r(\mathbf{x}_t) - r(\mathbf{x}_t^{k,i}) \\ & \leq \sum_{t=1}^T \langle \mathbf{q}_t - \mathbf{e}_k, \mathbf{l}_t \rangle + \langle \mathbf{p}_t^k - \mathbf{e}_i, \mathbf{l}_t^k \rangle - \gamma_1 \sum_{t=1}^T \sum_{k=1}^K q_t^k \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2 + \gamma_1 \sum_{t=1}^T \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2 \quad (47) \\ & \quad - \gamma_2 \sum_{t=1}^T \sum_{i=1}^N p_t^{k,i} \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 + \gamma_2 \sum_{t=1}^T \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2. \end{aligned}$$

Then, we introduce the following lemma, which is similar to Yan et al. [2023, Lemma 3] but with a different composite losses and optimisms in (13) and (14).

**Lemma 10.** *Let  $l_t^k, m_t^k$  be defined in (13), and  $l_t^{k,i}, m_t^{k,i}$  be defined in (14). Then, we have*

$$\begin{aligned} \sum_{t=1}^T \langle \mathbf{q}_t - \mathbf{e}_k, \mathbf{l}_t \rangle + \langle \mathbf{p}_t^k - \mathbf{e}_i, \mathbf{l}_t^k \rangle & \leq \frac{1}{\eta^k} \ln \frac{N}{3(C_0 \eta^k)^2} + 32\eta^k \sum_{t=1}^T \left( l_t^{k,i} - m_t^{k,i} \right)^2 \\ & \quad - \frac{C_0}{2} \sum_{t=2}^T \|\mathbf{q}_t - \mathbf{q}_{t-1}\|_1^2 - \frac{C_0}{16} \sum_{t=2}^T \|\mathbf{p}_t^k - \mathbf{p}_{t-1}^k\|_1^2. \end{aligned}$$

Note that

$$\begin{aligned} & \sum_{t=1}^T \left( l_t^{k,i} - m_t^{k,i} \right)^2 = \sum_{t=1}^T \left( \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}_t^{k,i} \rangle - \langle \nabla f_{t-1}(\mathbf{x}_{t-1}), \mathbf{x}_{t-1} - \mathbf{x}_{t-1}^{k,i} \rangle \right)^2 \\ & \leq 2 \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}_t^{k,i} \rangle^2 + 2 \sum_{t=1}^T \langle \nabla f_{t-1}(\mathbf{x}_{t-1}), \mathbf{x}_t - \mathbf{x}_t^{k,i} - \mathbf{x}_{t-1} + \mathbf{x}_{t-1}^{k,i} \rangle^2 \\ & \leq 2D^2 \bar{V}_T + 4G^2 \sum_{t=1}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + 4G^2 \sum_{t=1}^T \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2, \end{aligned}$$

where the second and third steps are due to the fact that  $(a+b)^2 \leq 2a^2 + 2b^2$  for any  $a, b \in \mathbb{R}$ . Taking the expectation on the both sides, we obtain

$$\begin{aligned} \mathbb{E} \left[ \sum_{t=1}^T \left( l_t^{k,i} - m_t^{k,i} \right)^2 \right] & \leq 2D^2 G^2 + 8D^2 (2\sigma_{1:T}^2 + \Sigma_{1:T}^2) + (8D^2 H^2 + 4G^2) \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \\ & \quad + 4G^2 \sum_{t=1}^T \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2, \end{aligned}$$

where the inequality is due to Lemma 3. Therefore, combining the above results with (47), Lemma 10 and Lemma 3 delivers

$$\begin{aligned} \text{meta-regret} & = \mathbb{E} \left[ \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}_t^{k,i} \rangle + r(\mathbf{x}_t) - r(\mathbf{x}_t^{k,i}) \right] \\ & \leq \frac{1}{\eta^k} \ln \frac{N}{3(C_0 \eta^k)^2} + 256\eta^k D^2 (2\sigma_{1:T}^2 + \Sigma_{1:T}^2) + \frac{64}{C_0} \left( 2D^2 H^2 + G^2 - \frac{C_0}{64} \gamma_1 \right) \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \\ & \quad + \left( \frac{64G^2}{C_0} - \gamma_2 \right) \sum_{t=1}^T \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 - \frac{C_0}{2} \sum_{t=2}^T \|\mathbf{q}_t - \mathbf{q}_{t-1}\|_1^2 - \frac{C_0}{16} \sum_{t=2}^T \|\mathbf{p}_t^k - \mathbf{p}_{t-1}^k\|_1^2 + \frac{32D^2 G^2}{C_0} \\ & \quad - \gamma_1 \sum_{t=2}^T \sum_{k=1}^K q_t^k \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2 - \gamma_2 \sum_{t=2}^T \sum_{i=1}^N p_t^{k,i} \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2. \end{aligned}$$

To bound the first two terms, we exploit the following lemma.

**Lemma 11.** [Yan et al., 2023, Lemma 7] For the step size pool  $\mathcal{H} = \{\eta_1, \dots, \eta_K\}$  with  $\eta_1 = \frac{1}{2C_0} \geq \dots \geq \eta_T = \frac{1}{2C_0 T}$ , if  $C_0 \geq \frac{\sqrt{X}}{2T}$ , there exists  $\eta^k \in \mathcal{H}$  satisfying

$$\frac{1}{\eta} \ln \frac{Y}{\eta^2} + \eta X \leq 2C_0 \ln(4YC_0^2) + 4\sqrt{X \ln(4XY)}.$$

Therefore, we have

$$\begin{aligned} & \frac{1}{\eta^k} \ln \frac{N}{3(C_0 \eta^k)^2} + 256\eta^k D^2 (2\sigma_{1:T}^2 + \Sigma_{1:T}^2) \\ & \leq 2C_0 \ln(2N) + 64D \sqrt{(2\sigma_{1:T}^2 + \Sigma_{1:T}^2) \ln \left( \frac{2^{10} D^2 N}{C_0^2} (2\sigma_{1:T}^2 + \Sigma_{1:T}^2) \right)} \\ & \leq 2C_0 \ln(2N) + 64D \sqrt{\ln \left( \frac{2^{10} D^2 N}{C_0^2} (2\sigma_{1:T}^2 + \Sigma_{1:T}^2) \right)} \left( \sqrt{2\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2} \right). \end{aligned}$$

Hence, by requiring  $C_0 \geq 1$ , the meta-regret is bounded as:

$$\begin{aligned} & \text{meta-regret} \\ & \leq \tilde{\mathcal{O}} \left( \sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2} \right) + \mathcal{O}(1) + 64 \left( 2D^2 H^2 + G^2 - \frac{1}{64} \gamma_1 \right) \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \\ & \quad + (64G^2 - \gamma_2) \sum_{t=1}^T \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 - \frac{C_0}{2} \sum_{t=2}^T \|\mathbf{q}_t - \mathbf{q}_{t-1}\|_1^2 - \frac{C_0}{16} \sum_{t=2}^T \|\mathbf{p}_t^k - \mathbf{p}_{t-1}^k\|_1^2 \\ & \quad - \gamma_1 \sum_{t=2}^T \sum_{k=1}^K q_t^k \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2 - \gamma_2 \sum_{t=2}^T \sum_{i=1}^N p_t^{k,i} \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2. \end{aligned}$$

Combining the expert-regret and meta-regret, we get

$$\begin{aligned} & \mathbb{E} [\text{Regret}_T] \\ & \leq \tilde{\mathcal{O}} \left( \sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2} \right) + \mathcal{O}(1) + (128D^2 H^2 + 64G^2 + 10DH - \gamma_1) \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \\ & \quad + \left( 64G^2 - \gamma_2 - \frac{\delta}{4D} \right) \sum_{t=1}^T \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 - \frac{C_0}{2} \sum_{t=2}^T \|\mathbf{q}_t - \mathbf{q}_{t-1}\|_1^2 - \frac{C_0}{16} \sum_{t=2}^T \|\mathbf{p}_t^k - \mathbf{p}_{t-1}^k\|_1^2 \\ & \leq \tilde{\mathcal{O}} \left( \sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2} \right) + \mathcal{O}(1) + (2C_1 - 2\gamma_1) \sum_{t=2}^T q_t^k \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2 \\ & \quad + \left( 2D^2 C_1 - 2D^2 \gamma_1 - \frac{C_0}{2} \right) \sum_{t=2}^T \|\mathbf{q}_t - \mathbf{q}_{t-1}\|_1^2 + \left( 64G^2 - \gamma_2 - \frac{\delta}{4D} \right) \sum_{t=1}^T \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2, \end{aligned}$$

where  $C_1 = 128D^2 H^2 + 64G^2 + 10DH$  and the last step is due to the following lemma.

**Lemma 12.** Let  $\mathbf{x}_t = \sum_{k=1}^K q_t^k \mathbf{x}_t^k$ . Then, we have

$$\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \leq 2 \sum_{k=1}^K q_t^k \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2 + 2D^2 \|\mathbf{q}_t - \mathbf{q}_{t-1}\|_1^2. \quad (48)$$

By setting  $C_0 \geq 1$ ,  $C_0 \geq 8D^2 C_1$ ,  $\gamma_1 \geq C_1$  and  $\delta \geq 256G^2 D + 4D\gamma_2$ , we finish the proof.

## B.6.2 For strongly convex functions

According to (28), we could decompose the regret as below:

$$\begin{aligned}
\mathbb{E} [\text{Regret}_T] &\leq \mathbb{E} \left[ \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) - \frac{\lambda}{2} \sum_{t=1}^T \|\mathbf{x}^* - \mathbf{x}_t\|_2^2 \right] \\
&\leq \mathbb{E} \left[ \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) - \frac{\lambda^i}{2} \sum_{t=1}^T \|\mathbf{x}^* - \mathbf{x}_t\|_2^2 \right] \\
&= \underbrace{\mathbb{E} \left[ \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}_t^{k,i} \rangle + r(\mathbf{x}_t) - r(\mathbf{x}_t^{k,i}) - \frac{\lambda^i}{2} \sum_{t=1}^T \|\mathbf{x}_t^{k,i} - \mathbf{x}_t\|_2^2 \right]}_{\text{meta-regret}} \\
&\quad + \underbrace{\mathbb{E} \left[ \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t^{k,i} - \mathbf{x}^* \rangle + r(\mathbf{x}_t^{k,i}) - r(\mathbf{x}^*) + \frac{\lambda^i}{2} \sum_{t=1}^T \|\mathbf{x}_t^{k,i} - \mathbf{x}_t\|_2^2 - \frac{\lambda^i}{2} \sum_{t=1}^T \|\mathbf{x}^* - \mathbf{x}_t\|_2^2 \right]}_{\text{expert-regret}},
\end{aligned}$$

where the first step is due to the fact that there exists an  $\lambda^i \in \mathcal{H}$  satisfying  $\lambda^i \leq \lambda \leq 2\lambda^i$ .

**Expert-regret analysis.** We define the surrogate loss functions  $h_{t,i}^{sc}(\mathbf{x}) = \langle \nabla f_t(\mathbf{x}_t), \mathbf{x} \rangle + r(\mathbf{x}) + \frac{\lambda^i}{2} \sum_{t=1}^T \|\mathbf{x} - \mathbf{x}_t\|_2^2$  for the  $i$ -th expert. According to Theorem 3, we can bound the expert-regret by

$$\begin{aligned}
&\leq \mathcal{O} \left( \frac{1}{\lambda} (\sigma_{\max}^2 + \Sigma_{\max}^2) \ln(\sigma_{1:T}^2 + \Sigma_{1:T}^2) \right) + 16D^2 (H^2 + 1)^2 \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \\
&\quad + \left( 8D^2 (H^2 + 1) - \frac{\delta}{8} \right) \sum_{t=2}^T \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 + \frac{1}{4} \delta D^2 + \frac{20G^2}{\lambda} + \frac{1}{\lambda^2} + \frac{\lambda D^2}{4}.
\end{aligned}$$

**Meta-regret analysis.** Similar to (46), the instantaneous meta-regret can be bounded by

$$\begin{aligned}
&\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}_t^{k,i} \rangle + r(\mathbf{x}_t) - r(\mathbf{x}_t^{k,i}) - \frac{\lambda^i}{2} \|\mathbf{x}_t^{k,i} - \mathbf{x}_t\|_2^2 \\
&\leq \langle \mathbf{q}_t - \mathbf{e}_k, \mathbf{l}_t \rangle + \langle \mathbf{p}_t^k - \mathbf{e}_i, \mathbf{l}_t^k \rangle - \gamma_1 \sum_{k=1}^K q_t^k \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2 + \gamma_1 \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2 \\
&\quad - \gamma_2 \sum_{i=1}^N p_t^{k,i} \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 + \gamma_2 \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 - \frac{\lambda^i}{2} \|\mathbf{x}_t^{k,i} - \mathbf{x}_t\|_2^2.
\end{aligned} \tag{49}$$

For the first two terms, we exploit Lemma 10 and obtain

$$\begin{aligned}
\sum_{t=1}^T \langle \mathbf{q}_t - \mathbf{e}_k, \mathbf{l}_t \rangle + \langle \mathbf{p}_t^k - \mathbf{e}_i, \mathbf{l}_t^k \rangle &\leq \frac{1}{\eta^k} \ln \frac{N}{3(C_0 \eta^k)^2} + 32\eta^k \sum_{t=1}^T (l_t^{k,i} - m_t^{k,i})^2 \\
&\quad - \frac{C_0}{2} \sum_{t=2}^T \|\mathbf{q}_t - \mathbf{q}_{t-1}\|_1^2 - \frac{C_0}{16} \sum_{t=2}^T \|\mathbf{p}_t^k - \mathbf{p}_{t-1}^k\|_1^2.
\end{aligned} \tag{50}$$

Note that

$$\begin{aligned}
\sum_{t=1}^T (l_t^{k,i} - m_t^{k,i})^2 &= \sum_{t=1}^T \left( \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}_t^{k,i} \rangle - \langle \nabla f_{t-1}(\mathbf{x}_{t-1}), \mathbf{x}_{t-1} - \mathbf{x}_{t-1}^{k,i} \rangle \right)^2 \\
&\leq 2 \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}_t^{k,i} \rangle^2 + 2 \sum_{t=1}^T \langle \nabla f_{t-1}(\mathbf{x}_{t-1}), \mathbf{x}_{t-1} - \mathbf{x}_{t-1}^{k,i} \rangle^2 \\
&\leq 4 \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}_t^{k,i} \rangle^2 + 2G^2 D^2 \leq 4G^2 \sum_{t=1}^T \|\mathbf{x}_t - \mathbf{x}_t^{k,i}\|_2^2 + 2G^2 D^2,
\end{aligned} \tag{51}$$

where the first step is due to the fact that  $(a + b)^2 \leq 2a^2 + 2b^2$  for any  $a, b \in \mathbb{R}$ . Substituting (50) and (51) into (49) obtains

$$\begin{aligned}
& \text{meta-regret} \\
& \leq \frac{1}{\eta^k} \ln \frac{N}{3(C_0 \eta^k)^2} + \left(128\eta^k G^2 - \frac{\lambda^i}{2}\right) \sum_{t=1}^T \left\| \mathbf{x}_t - \mathbf{x}_t^{k,i} \right\|^2 - \frac{C_0}{2} \sum_{t=2}^T \left\| \mathbf{q}_t - \mathbf{q}_{t-1} \right\|_1^2 \\
& \quad - \frac{C_0}{16} \sum_{t=2}^T \left\| \mathbf{p}_t^k - \mathbf{p}_{t-1}^k \right\|_1^2 - \gamma_1 \sum_{k=1}^K q_t^k \left\| \mathbf{x}_t^k - \mathbf{x}_{t-1}^k \right\|_2^2 + \gamma_1 \left\| \mathbf{x}_t^k - \mathbf{x}_{t-1}^k \right\|_2^2 \\
& \quad - \gamma_2 \sum_{i=1}^N p_t^{k,i} \left\| \mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i} \right\|_2^2 + \gamma_2 \left\| \mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i} \right\|_2^2 + 64\eta^k G^2 D^2 \\
& \leq \frac{1}{\eta^k} \ln \frac{N}{(\eta^k)^2} + \left(128\eta^k G^2 - \frac{\lambda^i}{2}\right) \sum_{t=1}^T \left\| \mathbf{x}_t - \mathbf{x}_t^{k,i} \right\|^2 - \frac{C_0}{2} \sum_{t=2}^T \left\| \mathbf{q}_t - \mathbf{q}_{t-1} \right\|_1^2 \\
& \quad + \left(2D^2\gamma_1 - \frac{C_0}{16}\right) \sum_{t=2}^T \left\| \mathbf{p}_t^k - \mathbf{p}_{t-1}^k \right\|_1^2 - \gamma_1 \sum_{k=1}^K q_t^k \left\| \mathbf{x}_t^k - \mathbf{x}_{t-1}^k \right\|_2^2 \\
& \quad + (2\gamma_1 - \gamma_2) \sum_{i=1}^N p_t^{k,i} \left\| \mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i} \right\|_2^2 + \gamma_2 \left\| \mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i} \right\|_2^2 + 64\eta^k G^2 D^2,
\end{aligned}$$

where the last step is due to Lemma 12 and  $C_0 \geq 1$ . To bound the first term in the above result, we employ the following lemma.

**Lemma 13.** [Yan et al., 2023, Lemma 8] For the step size pool  $\mathcal{H} = \{\eta_1, \dots, \eta_K\}$  with  $\eta_1 = \frac{1}{2C_0} \geq \dots \geq \eta_T = \frac{1}{2C_0 T}$ , if  $C_0 \geq \frac{1}{2\eta^* T}$  where  $\eta^*$  is the optimal step size, there exists  $\eta^k \in \mathcal{H}$  satisfying

$$\frac{1}{\eta} \ln \frac{Y}{\eta^2} \leq 2C_0 \ln(4YC_0^2) + \frac{1}{\eta^*} \ln \frac{4Y}{(\eta^*)^2}.$$

Applying the above lemma with  $\eta^* = \frac{\lambda^i}{256G^2}$  delivers

$$\begin{aligned}
\text{meta-regret} & \leq 2C_0 \ln(4NC_0^2) + \frac{512G^2}{\lambda} \ln \frac{2048NG^2}{\lambda^2} - \frac{C_0}{2} \sum_{t=2}^T \left\| \mathbf{q}_t - \mathbf{q}_{t-1} \right\|_1^2 \\
& \quad + \left(2D^2\gamma_1 - \frac{C_0}{16}\right) \sum_{t=2}^T \left\| \mathbf{p}_t^k - \mathbf{p}_{t-1}^k \right\|_1^2 - \gamma_1 \sum_{k=1}^K q_t^k \left\| \mathbf{x}_t^k - \mathbf{x}_{t-1}^k \right\|_2^2 \\
& \quad + (2\gamma_1 - \gamma_2) \sum_{i=1}^N p_t^{k,i} \left\| \mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i} \right\|_2^2 + \gamma_2 \left\| \mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i} \right\|_2^2 + \frac{\lambda D^2}{4}.
\end{aligned}$$

Then, we combine the expert-regret and meta-regret:

$$\begin{aligned}
\mathbb{E} [\text{Regret}_T] & \leq \mathcal{O} \left( \frac{1}{\lambda} (\sigma_{\max}^2 + \Sigma_{\max}^2) \ln(\sigma_{1:T}^2 + \Sigma_{1:T}^2) \right) \\
& \quad + \left(8D^2(H^2 + 1) - \frac{\delta}{8} + \gamma_2\right) \sum_{t=2}^T \left\| \mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i} \right\|^2 + \left(2D^2C_2 - \frac{C_0}{2}\right) \sum_{t=2}^T \left\| \mathbf{q}_t - \mathbf{q}_{t-1} \right\|_1^2 \\
& \quad + \left(2D^2\gamma_1 - \frac{C_0}{16}\right) \sum_{t=2}^T \left\| \mathbf{p}_t^k - \mathbf{p}_{t-1}^k \right\|_1^2 + (2C_2 - \gamma_1) \sum_{k=1}^K q_t^k \left\| \mathbf{x}_t^k - \mathbf{x}_{t-1}^k \right\|_2^2 \\
& \quad + (2\gamma_1 - \gamma_2) \sum_{i=1}^N p_t^{k,i} \left\| \mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i} \right\|_2^2 + \mathcal{O}(1).
\end{aligned}$$

where  $C_2 = 16D^2(H^2 + 1)^2$ . Setting  $C_0 \geq 4D^2C_2$ ,  $C_0 \geq 32D^2\gamma_1$ ,  $\delta \geq 64D^2(H^2 + 1) + 8\gamma_2$ ,  $\gamma_1 \geq 2C_2$  and  $\gamma_2 \geq \gamma_1$  completes the proof.

### B.6.3 For exp-concave functions

According to (38), we could decompose the regret as below:

$$\begin{aligned}
\mathbb{E} [\text{Regret}_T] &\leq \mathbb{E} \left[ \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) - \frac{\beta}{2} \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}^* - \mathbf{x}_t \rangle^2 \right] \\
&\leq \mathbb{E} \left[ \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}^* \rangle + r(\mathbf{x}_t) - r(\mathbf{x}^*) - \frac{\beta^i}{2} \sum_{t=1}^T \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}^* - \mathbf{x}_t \rangle^2 \right] \\
&= \underbrace{\mathbb{E} \left[ \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{x}_t - \mathbf{x}_t^{k,i} \rangle + r(\mathbf{x}_t) - r(\mathbf{x}_t^{k,i}) - \frac{\beta^i}{2} \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{x}_t^{k,i} - \mathbf{x}_t \rangle^2 \right]}_{\text{meta-regret}} \\
&\quad + \underbrace{\mathbb{E} \left[ \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{x}_t^{k,i} - \mathbf{x}^* \rangle + r(\mathbf{x}_t^{k,i}) - r(\mathbf{x}^*) + \frac{\beta^i}{2} \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{x}_t^{k,i} - \mathbf{x}_t \rangle^2 - \frac{\beta^i}{2} \sum_{t=1}^T \langle \mathbf{g}_t, \mathbf{x}^* - \mathbf{x}_t \rangle^2 \right]}_{\text{expert-regret}},
\end{aligned}$$

where  $\mathbf{g}_t = \nabla f_t(\mathbf{x}_t)$ ,  $\beta^i = \frac{1}{2} \min\{\frac{1}{4GD}, \alpha^i\}$ , and the first step is due to the fact that there exists an  $\alpha^i \in \mathcal{H}$  satisfying  $\alpha^i \leq \alpha \leq 2\alpha^i$ .

**Expert-regret analysis.** We define the surrogate loss functions  $h_{t,i}^{exp}(\mathbf{x}) = \langle \mathbf{g}_t, \mathbf{x} \rangle + r(\mathbf{x}) + \frac{\beta^i}{2} \langle \mathbf{g}_t, \mathbf{x} - \mathbf{x}_t \rangle^2$  for the  $i$ -th expert. According to Theorem 5, we can bound the expert-regret by

$$\begin{aligned}
&\text{expert-regret} \\
&\leq \mathcal{O} \left( \frac{d}{\alpha} \ln(\sigma_{1:T}^2 + \Sigma_{1:T}^2) \right) + \frac{24d}{\beta} \ln \left( 1 + \frac{C_3}{\delta} \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \right) \\
&\quad + \frac{24d}{\beta} \ln \left( 1 + \frac{\beta G^4}{\delta} \sum_{t=2}^T \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 \right) - \frac{\delta}{4} \sum_{t=2}^T \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 + \left( \frac{\delta}{2} + \frac{\beta}{4} G^2 \right) D^2 \\
&\leq \mathcal{O} \left( \frac{d}{\alpha} \ln(\sigma_{1:T}^2 + \Sigma_{1:T}^2) \right) + \frac{24dC_3}{\delta} \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + \left( \frac{24dG^4}{\delta} - \frac{\delta}{4} \right) \sum_{t=2}^T \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 \\
&\quad + \left( \frac{\delta}{2} + \frac{\beta}{4} G^2 \right) D^2,
\end{aligned}$$

where  $C_3 = 8H^2 + 64D^2G^2H^2 + 8G^4$ .

**Meta-regret analysis.** Similar to (46), the instantaneous meta-regret can be bounded by

$$\begin{aligned}
&\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}_t^{k,i} \rangle + r(\mathbf{x}_t) - r(\mathbf{x}_t^{k,i}) - \frac{\beta^i}{2} \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t^{k,i} - \mathbf{x}_t \rangle^2 \\
&\leq \langle \mathbf{q}_t - \mathbf{e}_k, \mathbf{l}_t \rangle + \langle \mathbf{p}_t^k - \mathbf{e}_i, \mathbf{l}_t^k \rangle - \gamma_1 \sum_{k=1}^K q_t^k \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2 + \gamma_1 \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2 \\
&\quad - \gamma_2 \sum_{i=1}^N p_t^{k,i} \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 + \gamma_2 \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 - \frac{\beta^i}{2} \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t^{k,i} - \mathbf{x}_t \rangle^2.
\end{aligned}$$

For the first two terms, we exploit Lemma 10 and obtain

$$\begin{aligned}
\sum_{t=1}^T \langle \mathbf{q}_t - \mathbf{e}_k, \mathbf{l}_t \rangle + \langle \mathbf{p}_t^k - \mathbf{e}_i, \mathbf{l}_t^k \rangle &\leq \frac{1}{\eta^k} \ln \frac{N}{3(C_0\eta^k)^2} + 32\eta^k \sum_{t=1}^T \left( l_t^{k,i} - m_t^{k,i} \right)^2 \\
&\quad - \frac{C_0}{2} \sum_{t=2}^T \|\mathbf{q}_t - \mathbf{q}_{t-1}\|_1^2 - \frac{C_0}{16} \sum_{t=2}^T \|\mathbf{p}_t^k - \mathbf{p}_{t-1}^k\|_1^2.
\end{aligned}$$

Note that

$$\begin{aligned}
& \sum_{t=1}^T \left( l_t^{k,i} - m_t^{k,i} \right)^2 = \sum_{t=1}^T \left( \left\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}_t^{k,i} \right\rangle - \left\langle \nabla f_{t-1}(\mathbf{x}_{t-1}), \mathbf{x}_{t-1} - \mathbf{x}_{t-1}^{k,i} \right\rangle \right)^2 \\
& \leq 2 \sum_{t=1}^T \left\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}_t^{k,i} \right\rangle^2 + 2 \sum_{t=1}^T \left\langle \nabla f_{t-1}(\mathbf{x}_{t-1}), \mathbf{x}_{t-1} - \mathbf{x}_{t-1}^{k,i} \right\rangle^2 \\
& \leq 4 \sum_{t=1}^T \left\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}_t^{k,i} \right\rangle^2 + 2G^2D^2.
\end{aligned}$$

Therefore, the meta-regret is bounded by

$$\begin{aligned}
& \text{meta-regret} \\
& \leq \frac{1}{\eta^k} \ln \frac{N}{3(C_0\eta^k)^2} + \left( 128\eta^k - \frac{\beta^i}{2} \right) \sum_{t=1}^T \left\langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x}_t^{k,i} \right\rangle^2 + 256G^2D^2\eta^k \\
& \quad - \frac{C_0}{2} \sum_{t=2}^T \|\mathbf{q}_t - \mathbf{q}_{t-1}\|_1^2 - \frac{C_0}{16} \sum_{t=2}^T \|\mathbf{p}_t^k - \mathbf{p}_{t-1}^k\|_1^2 - \gamma_1 \sum_{k=1}^K q_t^k \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2 + \gamma_1 \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2 \\
& \quad - \gamma_2 \sum_{i=1}^N p_t^{k,i} \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 + \gamma_2 \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2.
\end{aligned}$$

Applying Lemma 13 with  $\eta^* = \frac{\beta^i}{256}$  delivers

$$\begin{aligned}
\text{meta-regret} & \leq 2C_0 \ln(4NC_0^2) + \frac{512}{\beta} \ln \frac{2048N}{\beta^2} + G^2D^2\beta - \frac{C_0}{2} \sum_{t=2}^T \|\mathbf{q}_t - \mathbf{q}_{t-1}\|_1^2 \\
& \quad + \left( 2D^2\gamma_1 - \frac{C_0}{16} \right) \sum_{t=2}^T \|\mathbf{p}_t^k - \mathbf{p}_{t-1}^k\|_1^2 - \gamma_1 \sum_{k=1}^K q_t^k \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2 \\
& \quad + (2\gamma_1 - \gamma_2) \sum_{i=1}^N p_t^{k,i} \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 + \gamma_2 \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2.
\end{aligned}$$

Then, we combine the expert-regret and meta-regret:

$$\begin{aligned}
\mathbb{E}[\text{Regret}_T] & \leq \mathcal{O} \left( \frac{d}{\alpha} \ln(\sigma_{1:T}^2 + \Sigma_{1:T}^2) \right) + \left( \frac{48dDC_3}{\delta} - \frac{C_0}{2} \right) \sum_{t=2}^T \|\mathbf{q}_t - \mathbf{q}_{t-1}\|_1^2 \\
& \quad + \left( 2D^2\gamma_1 - \frac{C_0}{16} \right) \sum_{t=2}^T \|\mathbf{p}_t^k - \mathbf{p}_{t-1}^k\|_1^2 + \left( \frac{48dC_3}{\delta} - \gamma_1 \right) \sum_{k=1}^K q_t^k \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2 \\
& \quad + (2\gamma_1 - \gamma_2) \sum_{i=1}^N p_t^{k,i} \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 + \left( \frac{24dG^4}{\delta} - \frac{\delta}{4} + \gamma_2 \right) \|\mathbf{x}_t^{k,i} - \mathbf{x}_{t-1}^{k,i}\|_2^2 + \mathcal{O}(1).
\end{aligned}$$

where  $C_3 = 8H^2 + 64D^2G^2H^2 + 8G^4$ . Setting  $C_0 \geq 96dDC_3/\delta$ ,  $C_0 \geq 32D^2\gamma_1$ ,  $\delta \geq 4\sqrt{6d(G^4 + 4C_3)}$ ,  $\gamma_1 \geq 48dC_3/\delta$  and  $\gamma_2 \geq 2\gamma_1$  completes the proof.

## C Supporting lemmas

### C.1 Proof of Lemma 1

First, we introduce two lemmas that will be used in the following proof.

**Lemma 14.** [Scroccaro et al., 2023, Lemma 3.1] Let  $\mathcal{X}$  be a convex set,  $\varphi(\cdot) : \mathcal{X} \rightarrow \mathbb{R}$  be a convex function and  $\eta > 0$ . Define

$$\mathbf{u} = \arg \min_{\mathbf{x} \in \mathcal{X}} \{ \eta \varphi(\mathbf{x}) + \mathcal{B}^{\mathcal{R}}(\mathbf{x}, \mathbf{v}) \},$$

Then, for any  $\mathbf{z} \in \mathcal{X}$ , we have

$$\eta \langle \mathbf{g}(\mathbf{u}), \mathbf{u} - \mathbf{z} \rangle \leq \mathcal{B}^{\mathcal{R}}(\mathbf{z}, \mathbf{v}) - \mathcal{B}^{\mathcal{R}}(\mathbf{z}, \mathbf{u}) - \mathcal{B}^{\mathcal{R}}(\mathbf{u}, \mathbf{v})$$

where  $\mathbf{g}(\mathbf{u}) \in \partial \varphi(\mathbf{u})$ .

**Lemma 15.** [Scroccaro et al., 2023, Lemma 3.2] Let  $\mathcal{X}$  be a convex set,  $\mathcal{R} : \mathcal{X} \rightarrow \mathbb{R}$  be an  $\alpha$ -strongly convex function with respect to the norm  $\|\cdot\|$ . Define

$$\mathbf{u}_1 = \arg \min_{\mathbf{x}_1 \in \mathcal{X}} \{ \langle \mathbf{w}_1, \mathbf{x}_1 \rangle + r(\mathbf{x}_1) + \mathcal{B}^{\mathcal{R}}(\mathbf{x}_1, \mathbf{v}) \}$$

$$\mathbf{u}_2 = \arg \min_{\mathbf{x}_2 \in \mathcal{X}} \{ \langle \mathbf{w}_2, \mathbf{x}_2 \rangle + r(\mathbf{x}_2) + \mathcal{B}^{\mathcal{R}}(\mathbf{x}_2, \mathbf{v}) \}.$$

Then, we have

$$\|\mathbf{u}_1 - \mathbf{u}_2\| \leq \alpha^{-1} \|\mathbf{w}_1 - \mathbf{w}_2\|_*.$$

Then, by exploiting the convexity of  $r(\cdot)$ , we have

$$\begin{aligned} & \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x} \rangle + r(\mathbf{x}_t) - r(\mathbf{x}) = \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x} \rangle + r(\mathbf{x}_t) - r(\hat{\mathbf{x}}_{t+1}) + r(\hat{\mathbf{x}}_{t+1}) - r(\mathbf{x}) \\ & \leq \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t - \mathbf{x} \rangle + \langle \nabla r(\mathbf{x}_t), \mathbf{x}_t - \hat{\mathbf{x}}_{t+1} \rangle + \langle \nabla r(\hat{\mathbf{x}}_{t+1}), \hat{\mathbf{x}}_{t+1} - \mathbf{x} \rangle \\ & = \underbrace{\langle \nabla f_t(\mathbf{x}_t) - M_t, \mathbf{x}_t - \hat{\mathbf{x}}_{t+1} \rangle}_{\text{term (a)}} + \underbrace{\langle M_t + \nabla r(\mathbf{x}_t), \mathbf{x}_t - \hat{\mathbf{x}}_{t+1} \rangle}_{\text{term (b)}} + \underbrace{\langle \nabla f_t(\mathbf{x}_t) + \nabla r(\hat{\mathbf{x}}_{t+1}), \hat{\mathbf{x}}_{t+1} - \mathbf{x} \rangle}_{\text{term (c)}}. \end{aligned}$$

To upper bound the term (a), we exploit Lemma 15 and obtain

$$\text{term (a)} \leq \|\nabla f_t(\mathbf{x}_t) - M_t\|_* \|\mathbf{x}_t - \hat{\mathbf{x}}_{t+1}\| \leq \alpha^{-1} \|\nabla f_t(\mathbf{x}_t) - M_t\|_*^2. \quad (52)$$

Next, we apply Lemma 14 with the update (4) and (5), and obtain

$$\text{term (b)} \leq \mathcal{B}^{\mathcal{R}_t}(\hat{\mathbf{x}}_{t+1}, \hat{\mathbf{x}}_t) - \mathcal{B}^{\mathcal{R}_t}(\hat{\mathbf{x}}_{t+1}, \mathbf{x}_t) - \mathcal{B}^{\mathcal{R}_t}(\mathbf{x}_t, \hat{\mathbf{x}}_t) \quad (53)$$

$$\text{term (c)} \leq \mathcal{B}^{\mathcal{R}_t}(\mathbf{x}, \hat{\mathbf{x}}_t) - \mathcal{B}^{\mathcal{R}_t}(\mathbf{x}, \hat{\mathbf{x}}_{t+1}) - \mathcal{B}^{\mathcal{R}_t}(\hat{\mathbf{x}}_{t+1}, \hat{\mathbf{x}}_t) \quad (54)$$

Combining (52), (53) and (54) finishes the proof.

### C.2 Proof of Lemma 3

First, we consider the case  $t = 1$ . Under the Assumption 2, we obtain

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

In the case  $t \geq 2$ , according to the inequality  $\|\mathbf{a} + \mathbf{b}\|_2^2 \leq 2\|\mathbf{a}\|_2^2 + 2\|\mathbf{b}\|_2^2$ , we have

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

Under the Assumption 4, the above result becomes

$$\begin{aligned} & \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\ & \leq 4\|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 + 4H^2\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \\ & \quad + 4\|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 + 4\|\nabla F_{t-1}(\mathbf{x}_{t-1}) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2. \end{aligned} \quad (55)$$



Combining both cases, we reach at

$$\begin{aligned} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 &\leq G^2 + 4\|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 + 4\|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\ &\quad + 4H^2\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + 4\|\nabla F_{t-1}(\mathbf{x}_{t-1}) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2. \end{aligned}$$

Summing up both sides of the above inequality over  $t = 1, \dots, T$  obtains

$$\begin{aligned} \bar{V}_T &= G^2 + 4 \sum_{t=2}^T \|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 + 4 \sum_{t=2}^T \|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\ &\quad + 4H^2 \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + 4 \sum_{t=2}^T \|\nabla F_{t-1}(\mathbf{x}_{t-1}) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\ &= G^2 + 8 \sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - \nabla F_t(\mathbf{x}_t)\|_2^2 + 4 \sum_{t=2}^T \|\nabla F_t(\mathbf{x}_{t-1}) - \nabla F_{t-1}(\mathbf{x}_{t-1})\|_2^2 \\ &\quad + 4H^2 \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2, \end{aligned}$$

which completes the proof.

### C.3 Proof of Lemma 4

Similar to the analysis in Lemma 3, we first consider the case  $t = 1$ .

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

Then, for the case  $t \geq 2$ , we have

$$\begin{aligned} \|\nabla f_t(\hat{\mathbf{x}}_t) - \nabla f_{t-1}(\hat{\mathbf{x}}_t)\|_2^2 &\leq \sup_{\mathbf{x} \in \mathcal{X}} \|\nabla f_t(\mathbf{x}) - \nabla f_{t-1}(\mathbf{x})\|_2^2 \\ &\leq \sup_{\mathbf{x} \in \mathcal{X}} 2\|\nabla f_t(\mathbf{x}) - \nabla F_t(\mathbf{x})\|_2^2 + \sup_{\mathbf{x} \in \mathcal{X}} 2\|\nabla F_t(\mathbf{x}) - \nabla f_{t-1}(\mathbf{x})\|_2^2 \\ &\leq \sup_{\mathbf{x} \in \mathcal{X}} 2\|\nabla f_t(\mathbf{x}) - \nabla F_t(\mathbf{x})\|_2^2 + \sup_{\mathbf{x} \in \mathcal{X}} 2\|\nabla F_t(\mathbf{x}) - \nabla F_{t-1}(\mathbf{x})\|_2^2 + \sup_{\mathbf{x} \in \mathcal{X}} 2\|\nabla F_{t-1}(\mathbf{x}) - \nabla f_{t-1}(\mathbf{x})\|_2^2. \end{aligned}$$

Summing up both sides of the above inequality over  $t = 1, \dots, T$  completes the proof.

### C.4 Proof of Lemma 7

The analysis is based on Lemma 19 in Chiang et al. [2012]. Recall the definition  $H_t = I + \frac{\beta}{2}G^2\delta I + \frac{\beta}{2}\sum_{r=1}^{t-1} h_r$  with  $h_r = \nabla f_r(\mathbf{x}_r)\nabla f_r(\mathbf{x}_r)^\top$ . Then, we can proof that

$$H_t \succeq \delta I + \frac{\beta}{2}\sum_{r=1}^t h_r \succeq \delta I + \frac{\beta}{4}\sum_{r=1}^t (h_r + h_{r-1}). \quad (56)$$

The first step is due to Assumption 2, and the last step is due to  $\nabla f_0(\mathbf{x}_0)$  the all-0 vector. Next, according to the fact that

$$\begin{aligned} &\frac{1}{2}h_r + \frac{1}{2}h_{r-1} + \frac{1}{2}\nabla f_{r-1}(\mathbf{x}_{r-1})\nabla f_r(\mathbf{x}_r)^\top + \frac{1}{2}\nabla f_r(\mathbf{x}_r)\nabla f_{r-1}(\mathbf{x}_{r-1})^\top \\ &= \frac{1}{2}(\nabla f_r(\mathbf{x}_r) + \nabla f_{r-1}(\mathbf{x}_{r-1}))(\nabla f_r(\mathbf{x}_r) + \nabla f_{r-1}(\mathbf{x}_{r-1}))^\top \succeq 0, \end{aligned}$$

we have

$$h_r + h_{r-1} \succeq \frac{1}{2}(\nabla f_r(\mathbf{x}_r) - \nabla f_{r-1}(\mathbf{x}_{r-1}))(\nabla f_r(\mathbf{x}_r) - \nabla f_{r-1}(\mathbf{x}_{r-1}))^\top. \quad (57)$$

Substituting (57) into (56) delivers

$$H_t \succeq \delta I + \frac{\beta}{8}\sum_{r=1}^t (\nabla f_r(\mathbf{x}_r) - \nabla f_{r-1}(\mathbf{x}_{r-1}))(\nabla f_r(\mathbf{x}_r) - \nabla f_{r-1}(\mathbf{x}_{r-1}))^\top. \quad (58)$$

For brevity, we denote  $K_t = \delta I + \frac{\beta}{8} \sum_{r=1}^t (\nabla f_r(\mathbf{x}_r) - \nabla f_{r-1}(\mathbf{x}_{r-1})) (\nabla f_r(\mathbf{x}_r) - \nabla f_{r-1}(\mathbf{x}_{r-1}))^\top$ . From (58), we obtain that  $H_t^{-1} \preceq K_t^{-1}$ , which implies

$$\begin{aligned} \sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_{H_t^{-1}}^2 &\leq \sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_{K_t^{-1}}^2 \\ &= \frac{8}{\beta} \sum_{t=1}^T \left\| \sqrt{\frac{\beta}{8}} (\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})) \right\|_{K_t^{-1}}^2 \end{aligned} \quad (59)$$

Then, we introduce the following lemma:

**Lemma 16.** [Hazan et al., 2007, Lemma 11] Let  $\mathbf{u}_1, \dots, \mathbf{u}_T \in \mathbb{R}^d$  be any sequence vectors and  $\epsilon > 0$  be a positive real number. Denote  $V_t = \epsilon I + \sum_{i=1}^t \mathbf{u}_i \mathbf{u}_i^\top$ , then we have

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

By setting  $\epsilon = \delta$  and  $\mathbf{u}_t = \sqrt{\frac{\beta}{8}} (\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1}))$ , we have

$$\sum_{t=1}^T \left\| \sqrt{\frac{\beta}{8}} (\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})) \right\|_{K_t^{-1}}^2 \leq d \log \left( 1 + \frac{\beta}{8\delta} \sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 \right). \quad (60)$$

We complete the proof by combining (59) and (60).

## C.5 Proof of Lemma 10

Our proof starts from the following lemma for single-layer MsMwC.

**Lemma 17.** [Yan et al., 2023, Lemma 2] If  $\max_{t \in [T], i \in [N]} |l_t^i|, |m_t^i| \leq 1$ , then MsMwC enjoys

$$\begin{aligned} \sum_{t=1}^T \langle \mathbf{l}_t, \mathbf{p}_t \rangle - \sum_{t=1}^T l_t^i &\leq \frac{1}{\eta^i} \ln \frac{1}{\hat{p}_1^i} + \sum_{i=1}^N \frac{\hat{p}_1^i}{\eta^i} - 8 \sum_{t=1}^T \sum_{i=1}^N \eta^i p_t^i (l_t^i - m_t^i)^2 \\ &\quad + 16\eta^i \sum_{t=1}^T (l_t^i - m_t^i)^2 - \min_{k \in [K]} \frac{1}{4\eta^k} \sum_{t=2}^T \|\mathbf{p}_t - \mathbf{p}_{t-1}\|_1^2. \end{aligned}$$

It can be verified that under Assumption 1, 2 and 3, all our composite losses and optimisms defined in (13) and (14) are bounded

$$\begin{aligned} |l_t^k| &\leq GD + C + \gamma_1 D, \quad |m_t^k| \leq GD + C + \gamma_1 D \\ |l_t^{k,i}| &\leq GD + C + \gamma_2 D, \quad |m_t^{k,i}| \leq GD + C + \gamma_2 D, \end{aligned}$$

and we can rescale them to  $[-1, 1]$  with only constant multiplicative factors in the constant hyperparameter  $\gamma_1$  and  $\gamma_2$ . Therefore, applying Lemma 17 with the first layer (i.e., with  $l_t^k$  as the surrogate losses), we have

$$\begin{aligned} \sum_{t=1}^T \langle \mathbf{l}_t, \mathbf{q}_t \rangle - \sum_{t=1}^T l_t^k &\leq \frac{1}{\eta^k} \ln \frac{1}{\hat{q}_1^k} + \sum_{k=1}^K \frac{\hat{q}_1^k}{\eta^k} - 8 \sum_{t=1}^T \sum_{k=1}^K \eta^k q_t^k (l_t^k - m_t^k)^2 \\ &\quad + 16\eta^k \sum_{t=1}^T (l_t^k - m_t^k)^2 - \min_{k \in [K]} \frac{1}{4\eta^k} \sum_{t=2}^T \|\mathbf{q}_t - \mathbf{q}_{t-1}\|_1^2. \end{aligned} \quad (61)$$

For the first two terms, the initialization  $\hat{q}_1^k = (\eta^k)^2 / \sum_{k=1}^K (\eta^k)^2$  and  $\eta^k = 1/(C_0 2^k)$  imply

$$\frac{1}{\eta^k} \ln \frac{1}{\hat{q}_1^k} + \sum_{k=1}^K \frac{\hat{q}_1^k}{\eta^k} = \frac{1}{\eta^k} \ln \frac{\sum_{k=1}^K (\eta^k)^2}{(\eta^k)^2} + \sum_{k=1}^K \frac{\eta^k}{\sum_{k=1}^K (\eta^k)^2} \leq \frac{1}{\eta^k} \ln \frac{1}{3(C_0 \eta^k)^2} + 4C_0. \quad (62)$$

For the last term, due to  $\eta^k = 1/(C_0 2^k) \leq 2/C_0$ , we have

$$\min_{k \in [K]} \frac{1}{4\eta^k} \sum_{t=2}^T \|\mathbf{q}_t - \mathbf{q}_{t-1}\|_1^2 \geq \frac{C_0}{2} \sum_{t=2}^T \|\mathbf{q}_t - \mathbf{q}_{t-1}\|_1^2. \quad (63)$$

Substitute (62) and (63) into (61) delivers

$$\sum_{t=1}^T \langle \mathbf{l}_t, \mathbf{q}_t \rangle - \sum_{t=1}^T l_t^k \leq \frac{1}{\eta^k} \ln \frac{1}{3(C_0 \eta^k)^2} + 16\eta^k \sum_{t=1}^T (l_t^k - m_t^k)^2 - \frac{C_0}{2} \sum_{t=2}^T \|\mathbf{q}_t - \mathbf{q}_{t-1}\|_1^2. \quad (64)$$

Then, we apply Lemma 17 again with  $l_t^{k,i}$  and  $m_t^{k,i}$ .

$$\begin{aligned} \sum_{t=1}^T \langle \mathbf{l}_t^k, \mathbf{p}_t^k \rangle - \sum_{t=1}^T l_t^{k,i} &\leq \frac{1}{\eta^{k,i}} \ln \frac{1}{\hat{p}_1^{k,i}} + \sum_{i=1}^N \frac{\hat{p}_1^{k,i}}{\eta^{k,i}} - 8 \sum_{t=1}^T \sum_{i=1}^N \eta^{k,i} p_t^{k,i} (l_t^{k,i} - m_t^{k,i})^2 \\ &\quad + 16\eta^{k,i} \sum_{t=1}^T (l_t^{k,i} - m_t^{k,i})^2 - \min_{i \in [N]} \frac{1}{4\eta^{k,i}} \sum_{t=2}^T \|\mathbf{p}_t^k - \mathbf{p}_{t-1}^k\|_1^2. \end{aligned} \quad (65)$$

With the initialization  $\hat{p}_1^{k,i} = 1/N$  and  $\eta^{k,i} = 2\eta^k$ , we obtain

$$\frac{1}{\eta^{k,i}} \ln \frac{1}{\hat{p}_1^{k,i}} + \sum_{i=1}^N \frac{\hat{p}_1^{k,i}}{\eta^{k,i}} = \frac{\ln N + 1}{2\eta^k} \leq C_0 (\ln N + 1) = \mathcal{O}(1).$$

Therefore, (65) becomes

$$\begin{aligned} \sum_{t=1}^T \langle \mathbf{l}_t^k, \mathbf{p}_t^k \rangle - \sum_{t=1}^T l_t^{k,i} &\leq \frac{\ln N + 1}{2\eta^k} - 16\eta^k \sum_{t=1}^T \sum_{i=1}^N p_t^{k,i} (l_t^{k,i} - m_t^{k,i})^2 \\ &\quad + 32\eta^k \sum_{t=1}^T (l_t^{k,i} - m_t^{k,i})^2 - \frac{C_0}{16} \sum_{t=2}^T \|\mathbf{p}_t^k - \mathbf{p}_{t-1}^k\|_1^2. \end{aligned} \quad (66)$$

Combining (64) and (66), we obtain

$$\begin{aligned} &\sum_{t=1}^T \langle \mathbf{q}_t - \mathbf{e}_k, \mathbf{l}_t \rangle + \langle \mathbf{p}_t^k - \mathbf{e}_i, \mathbf{l}_t^k \rangle \\ &\leq \frac{1}{\eta^k} \ln \frac{1}{3(C_0 \eta^k)^2} + 32\eta^k \sum_{t=1}^T (l_t^{k,i} - m_t^{k,i})^2 - \frac{C_0}{16} \sum_{t=2}^T \|\mathbf{p}_t^k - \mathbf{p}_{t-1}^k\|_1^2 - \frac{C_0}{2} \sum_{t=2}^T \|\mathbf{q}_t - \mathbf{q}_{t-1}\|_1^2 \\ &\quad + 16\eta^k \sum_{t=1}^T \left( (l_t^k - m_t^k)^2 - \sum_{i=1}^N p_t^{k,i} (l_t^{k,i} - m_t^{k,i})^2 \right). \end{aligned} \quad (67)$$

According to (13) and (14), the last term is bounded by

$$\begin{aligned} &(l_t^k - m_t^k)^2 - \sum_{i=1}^N p_t^{k,i} (l_t^{k,i} - m_t^{k,i})^2 \\ &= \left( \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t^k \rangle - \langle \hat{\mathbf{m}}_t^k, \mathbf{p}_t^k \rangle \right)^2 - \sum_{i=1}^N p_t^{k,i} (l_t^{k,i} - m_t^{k,i})^2 \\ &= \langle \hat{\mathbf{l}}_t^k - \hat{\mathbf{m}}_t^k, \mathbf{p}_t^k \rangle^2 - \sum_{i=1}^N p_t^{k,i} (l_t^{k,i} - m_t^{k,i})^2 \leq 0 \end{aligned} \quad (68)$$

where  $\hat{\mathbf{l}}_t^k = \left( \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t^{k,1} \rangle, \dots, \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_t^{k,N} \rangle \right)$  and the last step is due to Cauchy-Schwarz inequality. Combining (67) and (68) completes the proof.

### C.6 Proof of Lemma 12

$$\begin{aligned}
\|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 &= \left\| \sum_{k=1}^K (q_t^k \mathbf{x}_t^k - q_{t-1}^k \mathbf{x}_{t-1}^k) \right\|_2^2 = \left\| \sum_{k=1}^K q_t^k (\mathbf{x}_t^k - \mathbf{x}_{t-1}^k) + \sum_{k=1}^K \mathbf{x}_{t-1}^k (q_t^k - q_{t-1}^k) \right\|_2^2 \\
&\leq 2 \left\| \sum_{k=1}^K q_t^k (\mathbf{x}_t^k - \mathbf{x}_{t-1}^k) \right\|_2^2 + 2 \left\| \sum_{k=1}^K \mathbf{x}_{t-1}^k (q_t^k - q_{t-1}^k) \right\|_2^2 \\
&\leq 2 \sum_{k=1}^K q_t^k \|\mathbf{x}_t^k - \mathbf{x}_{t-1}^k\|_2^2 + 2D^2 \|\mathbf{q}_t - \mathbf{q}_{t-1}\|_1^2,
\end{aligned}$$

where the second step is due to  $(a + b)^2 \leq 2a^2 + 2b^2$  and the last step is due to the triangle inequality.

## NeurIPS Paper Checklist

### 1. Claims

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

Answer: [Yes]

Justification: The contributions have been clearly claimed in the abstract and introduction.

Guidelines:

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

### 2. Limitations

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

Answer: [Yes]

Justification: The limitations have been discussed in Section 7.

Guidelines:

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

### 3. Theory Assumptions and Proofs

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

Answer: [Yes]

Justification: All the assumptions are listed in Section 3.1 and the complete proofs are provided in Appendix B.

Guidelines:

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

#### 4. Experimental Result Reproducibility

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

Answer: [Yes]

Justification: We have provided detailed descriptions for experiments in Section A that are sufficient to reproduce the experimental results.

Guidelines:

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

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

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

Answer: [No]

Justification: While the code is not included in the submission, the complete detailed descriptions of the experiments are provided in Section A.

Guidelines:

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

## 6. Experimental Setting/Details

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

Answer: [Yes]

Justification: We have provided detailed descriptions for experiments in Section A that are sufficient to reproduce the experimental results.

Guidelines:

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

## 7. Experiment Statistical Significance

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

Answer: [Yes]

Justification: All experiments are repeated ten times with different random seeds, and the mean and standard deviation are reported.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.

- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

## 8. Experiments Compute Resources

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

Answer: [No]

Justification: The experiments are conducted on a machine equipped with the Intel Xeon E5-2620 CPU and 32G memory.

Guidelines:

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

## 9. Code Of Ethics

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

Answer: [Yes]

Justification: The authors have diligently adhered to the NeurIPS Code of Ethics in all aspects of this research.

Guidelines:

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

## 10. Broader Impacts

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

Answer: [NA]

Justification: This paper is mainly theoretical and there is no societal impact of this work.

Guidelines:

- The answer NA means that there is no societal impact of the work performed.



- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

#### 11. Safeguards

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

Answer: [NA]

Justification: This paper is mainly theoretical and poses no such risks.

Guidelines:

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

#### 12. Licenses for existing assets

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

Answer: [NA]

Justification: This paper is mainly theoretical and does not use existing assets.

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.

- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, [paperswithcode.com/datasets](https://paperswithcode.com/datasets) has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

### 13. **New Assets**

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

Answer: [NA]

Justification: This paper does not release new assets.

Guidelines:

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

### 14. **Crowdsourcing and Research with Human Subjects**

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

Answer: [NA]

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

Guidelines:

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

### 15. **Institutional Review Board (IRB) Approvals or Equivalent for Research with Human Subjects**

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

Answer: [NA]

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

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.

- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.