Online Min-max Problems: Nonconvexity and Saddle Point

Anonymous authors
Paper under double-blind review

Abstract

Online min-max optimization has recently gained considerable interest due to its rich applications to game theory, multi-agent reinforcement learning, online robust learning, etc. Theoretical understanding in this field has been mainly focused on convex-concave settings. Online min-max optimization with nonconvex geometries, which captures various online deep learning problems, has yet been studied so far. In this paper, we make the first effort and investigate online nonconvex-strongly-concave min-max optimization in the nonstationary environment. We first introduce a natural notion of local Nash equilibrium (NE)-regret, and then propose a novel algorithm coined SODA to achieve the optimal regret. We further generalize our study to the setting with stochastic first-order feedback, and show that a variation of SODA can also achieve the same optimal regret in expectation. Our theoretical results and the superior performance of the proposed method are further validated by empirical experiments. To our best knowledge, this is the first exploration of efficient online nonconvex min-max optimization.

1 Introduction

Online optimization (Cesa-Bianchi & Lugosi, 2006) is a powerful paradigm for modeling many applications that require decision making based on information available sequentially. Specially, at each time instance, an online player needs to make a decision based on the history information, and then receives a feedback (which can be a possibly adversarial and nonstationary reward or loss value) that may be used in the future. There have been extensive studies in this field for various scenarios, such as online convex optimization (Shalev-Shwartz, 2012; Hazan et al., 2016), online bilevel optimization (Tarzanagh & Balzano, 2022), online federated learning (Chen et al., 2020), etc. Recently, the online min-max (i.e., saddle point) problem has gained considerable interest due to its broad applications in game theory (Roy et al., 2019; Zhang et al., 2022a), multi-agent reinforcement learning (Buşoniu et al., 2010; Zhang et al., 2021), online robust learning (Gabrel et al., 2014; Ben-Tal et al., 2015), to name a few.

On the theoretical side, a line of works have explored provably efficient algorithms for online min-max optimization. Specifically, Cardoso et al. (2019); Fiez et al. (2021); Immorlica et al. (2019); Zhang et al. (2022b) considered the zero-sum matrix games where the online objective function takes a bilinear form. Rivera et al. (2018); Roy et al. (2019) studied a more general online min-max problem, where the objective is strongly-convex and strongly-concave. Noarov et al. (2021) focused on multi-objective online min-max games, where the reward is convex-concave in each coordinate.

Despite many efforts so far, existing literature on online min-max optimization has mainly focused on online convex-concave problems and did not take **nonconvexity** into consideration. However, in practice, nonconvexity occurs very often in online min-max problems, particularly those that apply deep neural networks (DNNs) for decision making. For instance, in the time-varying two-player zero-sum stochastic games (Mertens & Neyman, 1981; Roy et al., 2019; Zhang et al., 2022b), where the payoffs change with time, the policies are modeled by DNNs with strong regularization, and hence the online objective function is nonconvex-strongly-concave.

Motivated by the aforementioned practical problems, the goal of this paper is to take the first step towards exploring the **online nonconvex-strongly-concave min-max** problem with dynamic (and hence nonstationary) loss functions. Due to the nonconvexity and nonstationarity nature of the problem, two new challenges arise as we explain below.

First, *how to define an appropriate notion of regret for the nonstationary environment under the online nonconvex setting?* The standard notion of Nash Equilibrium (NE)-regret, e.g., (Rivera et al., 2018) for online convex-concave problems, which quantifies the difference between the cumulative loss of players and the min-max value of the cumulative payoff loss, is highly unreasonable for nonconvex-concave setting, since the min-max comparator is intractable for a nonconvex-concave function. Hence, new surrogate for regret is in demand.

Second, *with a desirable notion of regret, how to design efficient algorithms?* A natural strategy to handle the nonstationarity is that at each round, the decision maker first learns a good enough decision based on history losses and then applies it to the adversarial loss of current round. Two key difficulties will arise during this process. First, how to identify a good decision? In nonconvex min-max problems, a good decision usually refers to a stationary point. The standard definition of a stationary point involves an optimization oracle, which is unknown to the decision maker. Thus the decision maker needs to find a surrogate to identify a near stationary point at each round. Second, when applying the decision based on history information to the adversarial loss, mismatch errors arise due to the variability of the environment, which motivates the need for nonstationarity measures.

1.1 Our contributions

In this paper, we handle the aforementioned challenges by introducing a new regret measure and developing efficient algorithms for online nonconvex min-max problem with optimal regret guarantees. The main contributions are highlighted below.

- We first introduce a novel notion of dynamic regret for online nonconvex-strongly-concave min-max problem, called **local Nash equilibrium (NE)-regret**, which jointly captures the nonconvexity, nonstationarity, and min-max nature of our problem.

- Based on the regret notion, we propose an efficient online min-max optimization algorithm, named time-Smoothed Online gradient Descent Ascent (SODA). The main idea underlying SODA is to output a near-stationary point at each round by performing two-timescale gradient descent ascent and utilizing a specially designed stop criterion component.

- We show that the local NE-regret of SODA scales as $O(\frac{T}{w^2})$ with a iteration complexity of $O(Tw)$, which matches the $\Omega(\frac{T}{w^2})$ regret lower bound and the order of iteration complexity of $O(Tw)$ provided in Hazan et al. (2017a) for online minimization (where we set the maximization to be over a singleton). Thus, SODA achieves the optimal performance for online nonconvex-strongly-concave min-max optimization.

- We further generalize our study to the setting with stochastic first-order feedback and show that a variation of SODA can also achieve a regret of $O(\frac{T}{w^2})$.

To our best knowledge, this is the first study on online nonconvex min-max optimization with theoretical characterization of the regret performance.

1.2 Related Work

Online min-max optimization. Recently, online min-max optimization, also known as online saddle-point game, has emerged as an interesting optimization framework, and has been studied under various settings. More specifically, the zero-sum matrix game considers the special case that the function is bilinear with a payoff matrix $\mathbf{A}_t$, where the objective function is given by $f_t(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\top \mathbf{A}_t \mathbf{y}_t$. Several works, for example, Cardoso et al. (2019); Fiez et al. (2021); Immorlica et al. (2019); Zhang et al. (2022b) proposed and analyzed algorithms with respect to different notions of regret. For more general objective functions, Rivera et al. (2018); Roy et al. (2019) studied the case where the loss function f_t is strongly-convex-strongly-concave. Very recently, Noarov et al. (2021) formulated a general multi-objective framework, where the goal is to minimize the maximum coordinate of the cumulative vector-valued loss with convex-concave function in every coordinate. We emphasize that all of the above studies did not consider nonconvexity in their objective functions, which is the focus of our study here.

Online nonconvex optimization. As online nonconvex optimization is an active research area, various works have taken different approaches to handle the nonconvexity. Assuming access to an offline nonconvex optimization oracle to approximate minimizers of perturbed nonconvex functions, Suggala & Netrapalli (2020); Agarwal et al. (2019) studied the performance of "follow the perturbed leader" (FTPL) algorithm (Kalai & Vempala, 2005), and their regrets are all static regret. Further, Hazan et al. (2017a); Hallak et al. (2021); Aydore et al. (2019) considered online nonconvex problems under nonstationary environments, and utilized sliding windows method with window size w. They proposed different notions of dynamic regrets and algorithms, and achieved an order of $O(\frac{T}{w^2})$ according regret notions. Additionally, Héliou et al. (2020) studied online nonconvex optimization with imperfect feedback. Except first-order optimization, Héliou et al. (2020); Roy et al. (2022) considered zeroth-order online nonconvex optimization and Lesage-Landry et al. (2020) studied second-order online nonconvex optimization.

Offline min-max optimization. There is a rich literature that studies a diverse set of algorithms for min-max optimization with nonconvexity in the offline setting. We next describe only those studies highly relevant to our study here. One celebrated approach is the nested-loop type algorithm (Rafique et al., 2021; Nouiehed et al., 2019; Thekumparampil et al., 2019; Kong & Monteiro, 2021), where the outer loop can be treated as an inexact gradient descent on a nonconvex function while the inner loop aims to find an approximate solution to the maximization problem (see Lin et al. (2020a) and references therein for a good collection of such studies). Another approach, manifesting in the recent works of Lu et al. (2020) and Lin et al. (2020a) considers less complicated single-loop structures. Specifically, the two-timescale GDA analyzed in Lin et al. (2020a) is closest to the implementation at each round of our proposed SODA method. But it is not straightforward to generalize the design to the online setting, and our analysis of the new local NE-regret for online optimization is also very different from such a offline min-max problem.

1.3 Notations

$[T] \triangleq \{1, \dots, T\}$. We use bold lower-case letters to denote vectors as in $\mathbf{x}, \mathbf{y}$, and denote its ℓ_2-norm as $\|\cdot\|$. We use calligraphic upper case letters to denote sets as in $\mathcal{Y}$, and use the notation $\mathcal{P}_{\mathcal{Y}}$ to denote projections onto the set. For a differentiable function $\Phi(\cdot) : \mathbb{R}^m \to \mathbb{R}$, we let $\nabla\Phi(\mathbf{x})$ denote the gradient of Φ at $\mathbf{x}$. For a function $f(\cdot,\cdot) : \mathbb{R}^m \times \mathcal{Y} \to \mathbb{R}$ of two variables, $\nabla_{\mathbf{x}} f(\mathbf{x}, \mathbf{y})$ (or $\nabla_{\mathbf{y}} f(\mathbf{x}, \mathbf{y})$) denotes the partial gradient of f with respect to the first variable (or the second variable) at the point $(\mathbf{x}, \mathbf{y})$. We also use $\nabla f(\mathbf{x}, \mathbf{y})$ to denote the full gradient at $(\mathbf{x}, \mathbf{y})$ where $\nabla f(\mathbf{x}, \mathbf{y}) = (\nabla_{\mathbf{x}} f(\mathbf{x}, \mathbf{y}), \nabla_{\mathbf{y}} f(\mathbf{x}, \mathbf{y}))$. Finally, we use the notation $O(\cdot)$ and $\Omega(\cdot)$ to hide constant factors which are independent of problem parameters.

2 Problem Setup

We consider solving the following online min-max (i.e., saddle-point) problem:

$$\min_{\mathbf{x}\in\mathbb{R}^m} \max_{\mathbf{y}\in\mathcal{Y}} f_t(\mathbf{x}, \mathbf{y}), \quad t \in [T] \tag{1}$$

where $f_t : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ is generally *nonconvex* in $\mathbf{x}$ but *concave* in $\mathbf{y}$ and where $\mathcal{Y}$ is a convex set. At each round $t \in [T]$, the environment first incurs a loss function f_t. Without knowing the knowledge of f_t, the $\mathbf{x}$-learner and $\mathbf{y}$-learner are tasked with predicting $\mathbf{x}_t$ and $\mathbf{y}_t$ respectively to solve eq. (1) based on loss functions up to round $t-1$, i.e., $\{f_i\}_{i=1}^{t-1}$. The learners then observe the function $f_t(\cdot)$ and suffer a loss of $f_t(\mathbf{x}_t, \mathbf{y}_t)$.

The following regularity assumptions for f_t are made throughout the entire paper:

Assumption 1 (Smoothness). *f_t is ℓ-smooth $\forall t \in [T]$, i.e., $\forall (\mathbf{x}, \mathbf{y}), (\mathbf{x}', \mathbf{y}')$, it holds that $\|\nabla f_t(\mathbf{x}, \mathbf{y}) - \nabla f_t(\mathbf{x}', \mathbf{y}')\| \le \ell \|(\mathbf{x}, \mathbf{y}) - (\mathbf{x}', \mathbf{y}')\|$.*

Assumption 2 (Strong Concavity). *The function $f_t(\mathbf{x}, \cdot)$ is μ-strongly concave $\forall t \in [T]$, i.e., given $\mathbf{x} \in \mathbb{R}^m$, $\forall \mathbf{y}, \mathbf{y}'$, it holds that $f_t(\mathbf{x}, \mathbf{y}) \leq f_t(\mathbf{x}, \mathbf{y}') + \langle \nabla_{\mathbf{y}} f_t(\mathbf{x}, \mathbf{y}'), \mathbf{y} - \mathbf{y}' \rangle - \frac{\mu}{2} \|\mathbf{y} - \mathbf{y}'\|^2$.*

Assumption 3 (Boundness). *The set $\mathcal{Y}$ is a convex and bounded set with diameter $D \geq 0$. There exists $M > 0$, s.t. $|f_t(\mathbf{x}, \mathbf{y})| \leq M$, $\forall t \in [T], \mathbf{x} \in \mathbb{R}^m, \mathbf{y} \in \mathcal{Y}$.*

The above assumptions are standard in the literature of online learning (Hazan et al., 2017b) and min-max optimization (Lin et al., 2020a;b).

When the loss f_t is fixed for all t, our framework specializes to the standard nonconvex-strongly-concave min-max optimization (Lin et al., 2020a;b). Putting into the context of online min-max optimization, our formulation is similar to those in Roy et al. (2019); Rivera et al. (2018); Zhang et al. (2022b), where they studied only the case where f_t is convex-concave. However, their standard regret minimization and equilibrium computation will be computationally infeasible for general nonconvex-strongly-concave losses.

3 How to Measure the Performance?

3.1 Local Nash Equilibrium (NE)-Regret

We introduce a new definition of a local regret that suits online nonconvex-strongly-concave min-max problems. Our new metric is motivated by the online nonconvex optimization literature; see for example Hazan et al. (2017a); Hallak et al. (2021). Specifically, for each t, we first define the smoothed functions of f_t over a sliding-window of size w as:

$$F_{t,w}(\mathbf{x}, \mathbf{y}) \overset{\text{def}}{=} \tfrac{1}{w} \textstyle\sum_{i=0}^{w-1} f_{t-i}(\mathbf{x}, \mathbf{y}). \tag{2}$$

For notation convenience, we treat $f_t(\mathbf{x}, \mathbf{y})$ as 0 for all $t < 0$. Moreover, since the averaging preserves strongly-convexity, which implies $F_{t,w}$ is strongly-concave in $\mathbf{y}$, the maximization problem $\max_{\mathbf{y} \in \mathcal{Y}} F_{t,w}(\mathbf{x}, \mathbf{y})$ can be solved efficiently. Then, we can naturally define the following function:

$$\Phi_{t,w}(\mathbf{x}) \overset{\text{def}}{=} \max\nolimits_{\mathbf{y} \in \mathcal{Y}} F_{t,w}(\mathbf{x}, \mathbf{y}). \tag{3}$$

The overall goal of online min-max optimization can be viewed as online minimization over the above defined $\Phi_{t,w}(\cdot)$ function.Thus, we define the following regret metric with respect to $\Phi_{t,w}(\cdot)$.

Definition 1 (Local Nash Equilibrium (NE)-Regret). *Let f_t be a sequence of functions satisfying Assumption 1-3, with $\Phi_{t,w}(\cdot)$ defined in eq. (3). The w-local Nash Equilibrium (NE)-Regret is defined as:*

$$\mathfrak{R}_{w-NE}(T) \overset{def}{=} \textstyle\sum_{t=1}^{T} \|\nabla \Phi_{t,w}(\mathbf{x}_t)\|^2. \tag{4}$$

$\nabla \Phi_{t,w}$ is well-defined since $\Phi_{t,w}$ is differentiable for nonconvex-strongly-concave min-max problem (Lin et al., 2020a). We next justify the above notion of the local NE-regret from three aspects.

Why norm of gradient as metric? In online *convex-concave* min-max optimization, it is standard to consider the *Nash Equilibrium (NE)-Regret* (Rivera et al., 2018) metric, defined as:

$$|\textstyle\sum_{t=1}^{T} f_t(\mathbf{x}_t, \mathbf{y}_t) - \min_{\mathbf{x} \in \mathbb{R}^m} \max_{\mathbf{y} \in \mathcal{Y}} \sum_{t=1}^{T} f_t(\mathbf{x}, \mathbf{y})|.$$

However, the above metric of NE-regret is inappropriate and face a major issue in the online *nonconvex-concave* formulation. The core challenge is that, even in the offline case ($T = 1$), it is hard to efficiently find the global optimum of $\min_{\mathbf{x} \in \mathbb{R}^m} \max_{\mathbf{y} \in \mathcal{Y}} f_1(\mathbf{x}, \mathbf{y})$ in the hindsight. Clearly, the problem is equivalent to $\min_{\mathbf{x} \in \mathbb{R}^m} \Phi(\mathbf{x})$, where $\Phi(\cdot) = \max_{\mathbf{y} \in \mathcal{Y}} f_1(\cdot, \mathbf{y})$ is generally nonconvex, and hence finding the global minimum for $\Phi(\mathbf{x})$ is NP hard. A common surrogate for the global minimum of Φ in the offline nonconvex-strongly-concave min-max literature is the notion of ϵ-stationary point (Lin et al., 2020a;b) for differentiable Φ, i.e., there exists some iterate $\mathbf{x}_t$ for which $\|\nabla \Phi(\mathbf{x}_t)\|^2 < \epsilon$. If $\epsilon = 0$, then $\mathbf{x}_t$ is a stationary point. Therefore, it is reasonable to leverage such a norm of gradient as the optimality criterion from the offline nonconvex min-max analysis.

Why sliding-window averaging? The motivation behind the window averaging is two-fold: (*i*) $F_{t,w}$ and $\Phi_{t,w}$ represent the average performance during the window, which is widely adopted to handle noises and fluctuations when the environment and the loss function f_t encounter mild perturbations and variations. For instance, when each loss function f_t is an unbiased noisy realization of some f, the expected gradient norm of a randomly selected update inside the window is a standard measure in the nonconvex stochastic optimization literature (Bottou et al., 2018) and can reduce the variation caused by noises. Such smoothed notion is also a common practice in the field of online nonconvex optimization[1] (Hazan et al., 2017a; Hallak et al., 2021; Aydore et al., 2019; Zhuang et al., 2020). (*ii*) The average performance itself is also a typical notion that people are interested in real-world applications. Suppose a decision maker in a time-varying environment (with loss functions f_t) has only finite term memory w. Then she naturally wishes to find the best decision based on the entire finite term memory and will choose the average loss function $F_{t,w}$ and $\Phi_{t,w}$ as the performance metrics. As another example, if the environment varies in a periodic manner, such an average performance metric during a whole period is naturally adopted in time series forecasting problems.

Why capturing the dynamic nature? It is desirable that the regret can capture how well the players adapt their actions to the best decision at *each round* if the environment is nonstationary and changes over time. In the well-studied online convex-concave setting, the notion of dynamic regret (Zhang et al., 2022b) is defined for this purpose, since its definition of $|\sum_{t=1}^{T} f_t(\mathbf{x}_t, \mathbf{y}_t) - \sum_{t=1}^{T} \min_{\mathbf{x}\in\mathbb{R}^m} \max_{\mathbf{y}\in\mathcal{Y}} f_t(\mathbf{x}, \mathbf{y})|$ evaluates the gap to the min-max comparator at *each round* instead of the min-max solution of the sum of functions over all rounds. For the nonconvex min-max setting, the best min-max comparator at *each round* can be set as the stationary point of the window function $\Phi_{t,w}(\cdot)$, which has zero gradient. Hence, our local regret in eq. (4) can be interpreted as evaluating the gap between $\|\nabla\Phi_{t,w}(\mathbf{x}_t)\|^2$ and its comparator (which equals zero gradient) at *each round*, and thus implicitly captures the player's adaption to the dynamic setting.

3.2 Variability of Environment

Intuitively, if the environment (and hence the loss function f_t) changes drastically over time, it will be hard to obtain meaningful guarantees efficiently. To handle this problem, dynamic (Roy et al., 2019; Zhang et al., 2022b) or local (Hallak et al., 2021) regret serves as better performance metrics to take the changing environment into consideration. Such notions typically rely on certain nonstationarity measures of the environment in order to reflect how the system dynamics affects the performance. Therefore, in this subsection, we introduce such measures of variation for loss functions, which will be crucial in our analysis and capture the nonstationarity of our online min-max settings.

Definition 2 (Variation of Sliding-window)**.** *Let us denote $\mathbf{y}^*_{t,w}(\mathbf{x}) = \arg\max_{\mathbf{y}\in\mathcal{Y}} F_{t,w}(\mathbf{x}, \mathbf{y})$. The sliding-window variation in $\mathbf{x}$ is defined as:*

$$V_{\mathbf{x},w}[T] := \sum_{t=1}^{T} \sup_{\mathbf{x}\in\mathbb{R}^m} \|\nabla_{\mathbf{x}} f_t\left(\mathbf{x}, \mathbf{y}^*_{t,w}(\mathbf{x})\right) - \nabla_{\mathbf{x}} f_{t-w}\left(\mathbf{x}, \mathbf{y}^*_{t,w}(\mathbf{x})\right)\|^2. \tag{5}$$

Moreover, the sliding-window variation in $\mathbf{y}$ is defined as:

$$V_{\mathbf{y},w}[T] := \sum_{t=1}^{T} \sup_{\mathbf{x}\in\mathbb{R}^m} \|\nabla_{\mathbf{y}} f_t\left(\mathbf{x}, \mathbf{y}^*_{t,w}(\mathbf{x})\right) - \nabla_{\mathbf{y}} f_{t-w}\left(\mathbf{x}, \mathbf{y}^*_{t-1,w}(\mathbf{x})\right)\|^2. \tag{6}$$

Remark 1. Clearly, $V_{\mathbf{x},w}[T]$ and $V_{\mathbf{y},w}[T]$ are $O(T)$ if the gradients of f_t are bounded and can be zero in the offline setting, i.e., $T = 1$. A key observation is that if the loss function encounters a periodic shift with certain period length of w^*, i.e., $f_{t+w^*} = f_t$, then for $w = w^*$, $f_t = f_{t-w}$ and $\mathbf{y}^*_{t,w} = \mathbf{y}^*_{t-1,w}$, which is implied by the fact that $F_{t+1,w} = F_{t,w}$. As a consequence, for the well-tuned w, the sliding-window variations could be considerably small compared to T, especially $V_{\mathbf{x},w}[T] = V_{\mathbf{y},w}[T] = 0$ in the above case.

4 SODA: Time-Smoothed Online Gradient Descent Ascent

In this section, we present our proposed method, named time-Smoothed Online gradient Descent Ascent (SODA), for online nonconvex-strongly-concave problem, and we show that our approach is capable of efficiently achieving a favorable local NE-regret bound.

[1]If we view $\mathcal{Y}$ to be singleton, the local NE-regret degenerates to local regret proposed in Hazan et al. (2017a).

Algorithm 1 Time-**S**moothed **O**nline Gradient **D**escent **A**scent (SODA)

Input: window size $w \geq 1$, stepsizes $(\eta_{\mathbf{x}}, \eta_{\mathbf{y}})$, tolerance $\delta > 0$
Initialization: $(\mathbf{x}_1, \mathbf{y}_1)$
1: **for** $t = 1$ to T **do**
2: Predict $(\mathbf{x}_t, \mathbf{y}_t)$. Observe the cost function $f_t : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$
3: Set $(\mathbf{x}_{t+1}, \mathbf{y}_{t+1}) \leftarrow (\mathbf{x}_t, \mathbf{y}_t)$
4: **repeat**
5: $\mathbf{x}_{t+1} \leftarrow \mathbf{x}_{t+1} - \eta_{\mathbf{x}} \nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}_{t+1}, \mathbf{y}_{t+1})$
6: $\mathbf{y}_{t+1} \leftarrow \mathcal{P}_{\mathcal{Y}}(\mathbf{y}_{t+1} + \eta_{\mathbf{y}} \nabla_{\mathbf{y}} F_{t,w}(\mathbf{x}_{t+1}, \mathbf{y}_{t+1}))$
7: **until** $\frac{\kappa^2}{\eta_{\mathbf{y}}^2} \|\mathbf{y}_{t+1} - \mathcal{P}_{\mathcal{Y}}(\mathbf{y}_{t+1} + \eta_{\mathbf{y}} \nabla_{\mathbf{y}} F_{t,w}(\mathbf{x}_{t+1}, \mathbf{y}_{t+1}))\|^2 + \|\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}_{t+1}, \mathbf{y}_{t+1})\|^2 \leq \frac{\delta^2}{2w^2}$
8: **end for**

4.1 Proposed Algorithm

At the high-level, our algorithm plays following the-leader iterates, aiming to find a suitable approximating stationary point at each round using two-timescale gradient descent ascent (GDA). At each round t, SODA performs gradient descent over the variable $\mathbf{x}$ with the stepsize $\eta_{\mathbf{x}}$ and gradient ascent over the variable $\mathbf{y}$ with the stepsize $\eta_{\mathbf{y}}$ on function $F_{t,w}(\mathbf{x}, \mathbf{y})$ until the stop condition is satisfied. Then, SODA observes the loss function f_{t+1} to be used in the next round. The pseudocode of SODA is summarized in Algorithm 1.

Discussions about stopping criterion. Due to the online nature, the design of the stopping condition is to guarantee that the learner outputs a good $\mathbf{x}_{t+1}$ with small local regret at round t, i.e., $\|\nabla \Phi_{t,w}(\mathbf{x}_{t+1})\|^2$ is small enough. However, we do not have direct access to the first order oracle of $\Phi_{t,w}$. To circumvent this issue, we adopt the global error bound condition from the seminal paper Drusvyatskiy & Lewis (2018) to translate conditions on $\nabla \Phi_{t,w}(\mathbf{x}_{t+1})$ into restrictions on tractable $\nabla F_{t,w}$. Specifically, we prove that $\|\nabla \Phi_{t,w}(\mathbf{x}_{t+1})\|^2$ is upper bounded by the left-hand side of inequality in Algorithm 1 line 7 (see Lemma A.3). Therefore, alternatively we can utilize the accessible information of $\nabla F_{t,w}$ to terminate the inner loop iterations at time t.

Last-iterate guarantee. At each round t, the stop condition will be triggered only when the local regret of last iteration is small enough. Such a *last-iterate* type guarantee is different by nature from existing offline nonconvex-strongly-concave min-max results (Lin et al., 2020b;a), which are only guaranteed to visit an ϵ-stationary point within a certain number of iterations, i.e., where the return $\bar{\mathbf{x}}$ is uniformly drawn from previous iterations. Crucially, we will establish the total iteration bound (see Theorem 2) in the next subsection, which indicates that such last-iterate type outputs can be obtained efficiently. Furthermore, since the stopping criterion leads to stronger guarantee, our result is incomparable with former offline iteration complexity in the special case that $T = 1$.

4.2 Theoretical Guarantees

In this subsection, we provide the regret and computational complexity guarantees of our algorithm under local NE-regret and highlight several connections with the existing results from offline min-max optimization and online nonconvex problem.

Theorem 1 (Local NE regret minimization)**.** *Let $\kappa = \ell/\mu$ denote the condition number. Under Assumptions 1-3, and letting the stepsizes be chosen as $\eta_{\mathbf{x}} = \Theta\left(1/\kappa^3\ell\right)$ and $\eta_{\mathbf{y}} = \Theta(1/\ell)$, then Algorithm 1 enjoys the following local NE-regret bound:*

$$\Re_{w-NE}(T) = \textstyle\sum_{t=1}^{T} \|\nabla \Phi_{t,w}(\mathbf{x}_t)\|^2 \leq \frac{3}{w^2}(T\delta^2 + \frac{(\kappa w)^2}{(w-1)^2} V_{\mathbf{y},w}[T] + V_{\mathbf{x},w}[T]).$$

Theorem 2 (Iteration bound)**.** *Let τ denote the total number of iterations incurred by Algorithm 1. Then τ can be upper bounded as:*

$$\tau \leq \frac{480\kappa^3 \ell M w T}{\delta^2} + 256\frac{\kappa^2 T}{\mu} + \frac{256 D^2 \kappa^3 \ell^2 w^2}{\delta^2} + 512\frac{w^2 \kappa^5}{(w-1)^2 \delta^2} V_{\mathbf{y},w}[T].$$

Furthermore, the number of first-order gradient calls is bounded by $O(w\tau)$.

Theorems 1 and 2 together reveal the trade-off between the impact of sliding-window size w on the regret and the computational complexity, where larger w leads to smaller regret bound but incurs more gradient calls.

Robustness of SODA. Our results in Theorems 1 and 2 are expressed in terms of variation measures $V_{\mathbf{x},w}[T]$ and $V_{\mathbf{y},w}[T]$ of the environment introduced in Section 3.2. If we make the more restrictive assumption similar to that in Hazan et al. (2017a) that the gradient of f_t is bounded, the above theorems provide a robust guarantee for SODA; namely, no matter how the environment changes at each round, SODA always ensures $O(\frac{T}{w^2})$ local NE-regret with $O(Tw)$ iterations since $V_{\mathbf{x},w}[T]$ and $V_{\mathbf{y},w}[T]$ are $O(T)$ by definitions. Therefore, the regret can be made sublinear in T if w is selected accordingly. Interestingly, following SODA, the local NE-regret can achieve the same order without the bounded gradient assumption depending on the nonstationarity. Particularly, as we discussed in Remark 1, for the scenario that f_t is periodic with period w, $V_{\mathbf{x},w}[T] = V_{\mathbf{y},w}[T] = 0$.

Optimality of regret bound. Note that the basic online nonconvex minimization problem can be viewed as a special case of our online nonconvex min-max problem, if $f_t(x, y)$ takes values independent of y. In such a degenerate case, our local NE-regret is equivalent to the local regret analyzed in Hazan et al. (2017a); Hallak et al. (2021). Consequently, the adversarial example that incurs the local regret of $\Omega(\frac{T}{w^2})$ constructed in Hallak et al. (2021) can also serve as a worst case example for our online noncovex min-max setting. Moreover, under the same assumption made in Hazan et al. (2017a) (which is more restrictive than our assumption here), we achieve a robust regret upper bound of $O(\frac{T}{w^2})$ (as discussed in the previous paragraph), which matches the worst-case lower bound, indicating that our bound Theorem 1 for online nonconvex min-max problem is optimal.

Comparison to offline min-max optimization. When the environment is fixed, i.e. $f_t \equiv f$ or $T = 1$ with $w = 1$, our problem specializes to offline min-max optimization and $V_{\mathbf{x},w}[T] = V_{\mathbf{y},w}[T] = 0$ will disappear from our results. Therefore, an immediate implication from our theorems is that GDA is guaranteed to find ϵ-stationary point with $O(\kappa^3\epsilon^{-2})$ iteration complexity. The best known complexity bound for GDA in offline min-max optimization is $O(\kappa^2\epsilon^{-2})$ (Lin et al., 2020a). However, as we discussed in Section 4.1, SODA aims to output $\mathbf{x}$ with last-iterate type guarantee, which is a stronger notion than that considered in Lin et al. (2020a), where GDA are only guaranteed to visit an ϵ-stationary point within a certain number of iterations. Thus, these results are not directly comparable.

5 SODA with Stochastic First-order Oracle

In this section, we extend our online min-max framework to an online stochastic version. This setting is motivated by the fact that, in real world application, such as training a neural network, an oracle with access to the gradient of loss function is hard to reach. Instead, a stochastic first-order oracle (SFO) is used to approximate the ground truth gradient. Similar settings have been studied in Nemirovski et al. (2009); Hazan et al. (2017a); Hallak et al. (2021). Specifically, the formal SFO definition is as follows.

Definition 3 (Stochastic first-order oracle)**.** *A stochastic first-order oracle (SFO) is a function $\mathcal{S}_\sigma$ such that, given a point $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^m \times \mathcal{Y}$, a random seed ζ, and a smooth function $h : \mathbb{R}^m \times \mathcal{Y} \to \mathbb{R}$ satisfies:*

- $\mathcal{S}_\sigma(\mathbf{x}, \mathbf{y}; \zeta, h)$ *is an unbiased estimate of* $\nabla h(\mathbf{x}, \mathbf{y})$: $\mathbb{E}\left(\mathcal{S}(\mathbf{x}, \mathbf{y}; \zeta, h) - \nabla h(\mathbf{x}, \mathbf{y})\right) = 0$;
- $\mathcal{S}_\sigma(\mathbf{x}, \mathbf{y}; \zeta, h)$ *has variance bounded by* $\sigma^2 > 0$: $\mathbb{E}\left(\|\mathcal{S}(\mathbf{x}, \mathbf{y}; \zeta, h) - \nabla h(\mathbf{x}, \mathbf{y})\|^2\right) \leq \sigma^2$.

5.1 Proposed Algorithm

With the above definition of SFO, we introduce the stochastic version of Algorithm 1, named SODA-SFO (see Algorithm 2). Similarly, SODA-SFO also follows the-leader iterates using two-time scale GDA. Taking the noise brought by SFO into consideration, nested loops and special stopping criterion (in line 6 in Algorithm 2) are modified accordingly. Specially, (i) SFO results in different coefficients in stop criterion compared to SODA. (ii) The stopping criterion in SODA-SFO only ensures that $\|\nabla \Phi_{t,w}(\mathbf{x}_{t+1})\|^2$ is bounded by the

Algorithm 2 SODA with Stochastic First-order Oracle (SODA-SFO)

Input: window size $w \geq 1$, stepsizes $(\eta_{\mathbf{x}}, \eta_{\mathbf{y}})$, tolerance $\delta > 0$

Initialization: $(\mathbf{x}_1, \mathbf{y}_1)$

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

  Cost function $f_t : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ is updated;

  Sample $\tilde{\nabla} f_t(\mathbf{x}_t, \mathbf{y}_t) \leftarrow \mathcal{S}_{\sigma/w}(\mathbf{x}_t, \mathbf{y}_t; \zeta, f_t)$

  Set $\tilde{\nabla} F_{t,w}(\mathbf{x}_t, \mathbf{y}_t) = \tilde{\nabla} F_{t-1,w}(\mathbf{x}_t, \mathbf{y}_t) + \frac{1}{w}(\tilde{\nabla} f_{t-w}(\mathbf{x}_t, \mathbf{y}_t) - \tilde{\nabla} f_t(\mathbf{x}_t, \mathbf{y}_t))$

  Set $\mathbf{x}_t^0 = \mathbf{x}_t$, $\mathbf{y}_t^0 = \mathbf{y}_t$, $G_{\mathbf{x},t}^0 = \tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t, \mathbf{y}_t)$, $G_{\mathbf{y},t}^0 = \tilde{\nabla}_{\mathbf{y}} F_{t,w}(\mathbf{x}_t, \mathbf{y}_t)$, $k = 0$

  while $\frac{2\kappa^2}{\eta_{\mathbf{y}}^2}\|\mathbf{y}_t^k - \mathcal{P}_{\mathcal{Y}}\left(\mathbf{y}_t^k + \eta_{\mathbf{y}} G_{\mathbf{y},t}^k\right)\|^2 + \|G_{\mathbf{x},t}^k\|^2 > \delta^2/3w^2$ **do**

    $\mathbf{x}_t^{k+1} \leftarrow \mathbf{x}_t^k - \eta_{\mathbf{x}} G_{\mathbf{x},t}^k$

    $\mathbf{y}_t^{k+1} \leftarrow \mathcal{P}_{\mathcal{Y}}\left(\mathbf{y}_t^k + \eta_{\mathbf{y}} G_{\mathbf{y},t}^k\right)$

    Sample $\tilde{\nabla} f_i(\mathbf{x}_t^{k+1}, \mathbf{y}_t^{k+1}) \leftarrow \mathcal{S}_{\frac{\sigma}{w}}(\mathbf{x}_t^{k+1}, \mathbf{y}_t^{k+1}; \zeta, f_i)$ for $i = t-w+1, \cdots, t$;

    Set $G_t^{k+1} := (G_{t,\mathbf{x}}^{k+1}, G_{t,\mathbf{y}}^{k+1}) = \frac{1}{w}\sum_{i=t-w+1}^{t} \tilde{\nabla} f_i(\mathbf{x}_t^{k+1}, \mathbf{y}_t^{k+1})$

    $k \leftarrow k+1$

  end while

  $\mathbf{x}_{t+1} = \mathbf{x}_t^k$, $\mathbf{y}_{t+1} = \mathbf{y}_t^k$, and $\tilde{\nabla} F_{t,w}(\mathbf{x}_{t+1}, \mathbf{y}_{t+1}) = G_t^k$

end for

threshold plus the variation of SFO. But the variation here does not play an important role, since sliding windows serve variance reduction purpose to reduce the variation in the final expectation regret.

5.2 Theoretical Guarantees

Denote τ_t as the number of iterations of inner loop at round t and thus $\tau = \sum_{t=1}^{T} \tau_t$. We first establish that for each round t, the inner loop terminates with finite iterations τ_t provided that δ is not too small (recall that δ is the tolerance for stopping criterion), which justifies that SODA-SFO is computation tractable.

Theorem 3 (Finite iteration with SFO). *Let $\kappa = \ell/\mu$ denote the condition number, and let the stepsizes be chosen as $\eta_{\mathbf{x}} = \Theta\left(1/\kappa^3\ell\right)$ and $\eta_{\mathbf{y}} = \Theta(1/\ell)$. Under Assumptions 1-3, for any $t \in [T]$, if δ, w and σ satisfy that $\delta^2 = O(\frac{\kappa^4\ell^2\sigma^2}{w})$, then τ_t and τ is finite with high probability. Specially, when $K \in \mathbb{R}$ is large enough, $\mathbb{P}(\tau_t > K) = O(1/K)$.*

With the finite step stopping guarantee on hand, we next characterize the performance of SODA-SFO with expectation local NE-regret formally in terms of $w, T, V_{\mathbf{x},w}[T], V_{\mathbf{y},w}[T]$.

Theorem 4 (Expectation local NE-regret with SFO). *Under the setting of Theorem 3, SODA-SFO enjoys the following expectation local NE regret bound:*

$$\mathbb{E}\left[\Re_{w-NE}(T)\right] \leq \frac{T}{w^2}\left(3\delta^2 + \frac{(18\kappa^2+9)\sigma^2}{w}\right) + \frac{3\kappa^2}{(w-1)^2}V_{\mathbf{y},w}[T] + \frac{3}{w^2}V_{\mathbf{x},w}[T].$$

Beyond finite stopping and the regret bound, people may wonder whether the inner loop is meaningful if the per-stage calls of SFOs increase greatly, and are also interested in the total complexity of SFO calls. To address such an issue, we further provide an upper bound on the complexity of SFO calls similar to Theorem 2. In the stochastic online nonconvex min-max setting, we further need the following widely adopted assumption (Li & Orabona, 2019; Hallak et al., 2021) to control the noise caused by SFO calls.

Assumption 4. *Given any point $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^m \times \mathcal{Y}$, random seed ζ, and smooth function h: $\mathbb{R}^m \times \mathcal{Y}$, the SFO defined in Definition 3 satisfys that $\|\mathcal{S}(\mathbf{x}, \mathbf{y}; \zeta, h) - \nabla h(\mathbf{x}, \mathbf{y})\|^2 \leq \sigma^2$.*

Remark 2. We remark here that Theorems 3 and 4 do not require Assumption 4, and Theorem 3 provide the finite iteration guarantee with high probability and Theorem 4 provides an upper bound for expectation regret. With Assumption 4, which is slightly stronger than the assumptions in Definition 3, we are able to provide the following deterministic bound on iterations and the number of SFO calls as in Theorem 5. Furthermore, Theorem 5 can provide deterministic guarantees rather than high probability guarantees because Assumption 4 controls the variation of noise in an absolute and deterministic manner.

Theorem 5 (Iterations and SFO calls bounds)**.** *Under the setting of Theorem 3 and Assumption 4, and suppose that* $\delta^2 > 540\kappa^4\sigma^2$*. Then the total number of iterations satisfies*

$$\tau \leq \frac{1}{\eta_{\mathbf{x}}} \frac{2MTw + \frac{3\delta^2 T}{64\ell} + \frac{\ell w^2}{3\mu^2(w-1)^2} V_{\mathbf{y},w}[T] + w^2 M + \frac{5\ell D^2 w^2}{32}}{\left(\frac{\delta^2}{27} - 20\kappa^4\sigma^2\right)}.$$

Furthermore, the number of SFO calls is bounded by $O(w\tau)$*.*

The above results also provide a robust guarantee for SODA-SFO, where SODA-SFO achieves a expectation regret of $O(\frac{T}{w^2})$ with at most $O(Tw)$ iterations and hence $O(Tw^2)$ calls of SFO, as long as $V_{\mathbf{x},w}[T]$ and $V_{\mathbf{y},w}[T]$ scale with $O(T)$. Following the similar discussions from Remark 1 and Section 4.2, such condition can hold with relaxed assumptions depending on nonstationarity.

Specially, if the variance of SFO defined in Definition 3 is zero, then SFO reduces to perfect first order feedback. Hence, as discussed in Section 4.2, the adversarial example provided by Hazan et al. (2017a) is also applicable to the stochastic setting, and thus indicates that the expectation regret $O(\frac{T}{w^2})$ reaches optimality. If the set $\mathcal{Y}$ is a singleton, online nonconvex min-max problem with SFO reduces to the online nonconvex problem with SFO. In this case, the term consisting of $V_{\mathbf{y},w}[T]$ will disappear in our analysis, and our theorems recover the results in Hallak et al. (2021).

6 Conclusions

This paper provides the first analysis for the online nonconvex-concave min-max optimization problem. We introduced a novel notion of local Nash Equilibrium regret to capture the nonconvexity and nonstationary of the environment. We developed and analyzed algorithms SODA and its stochastic version with respect to the proposed notions of regret, establishing favorable regret and complexity guarantees.

References

Naman Agarwal, Alon Gonen, and Elad Hazan. Learning in non-convex games with an optimization oracle. In *Conference on Learning Theory*, pp. 18–29. PMLR, 2019.

Sergul Aydore, Tianhao Zhu, and Dean P Foster. Dynamic local regret for non-convex online forecasting. *Advances in neural information processing systems*, 32, 2019.

Aharon Ben-Tal, Elad Hazan, Tomer Koren, and Shie Mannor. Oracle-based robust optimization via online learning. *Operations Research*, 63(3):628–638, 2015.

Léon Bottou, Frank E Curtis, and Jorge Nocedal. Optimization methods for large-scale machine learning. *Siam Review*, 60(2):223–311, 2018.

Lucian Buşoniu, Robert Babuška, and Bart De Schutter. Multi-agent reinforcement learning: An overview. *Innovations in multi-agent systems and applications-1*, pp. 183–221, 2010.

Adrian Rivera Cardoso, Jacob D. Abernethy, He Wang, and Huan Xu. Competing against nash equilibria in adversarially changing zero-sum games. In *ICML*, 2019.

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

Yujing Chen, Yue Ning, Martin Slawski, and Huzefa Rangwala. Asynchronous online federated learning for edge devices with non-iid data. In *2020 IEEE International Conference on Big Data (Big Data)*, pp. 15–24. IEEE, 2020.

Damek Davis and Dmitriy Drusvyatskiy. Stochastic model-based minimization of weakly convex functions. *SIAM J. Optim.*, 29(1):207–239, 2019. doi: 10.1137/18M1178244. URL https://doi.org/10.1137/18M1178244.

Dmitriy Drusvyatskiy and Adrian S Lewis. Error bounds, quadratic growth, and linear convergence of proximal methods. *Mathematics of Operations Research*, 43(3):919–948, 2018.

Tanner Fiez, Ryann Sim, Stratis Skoulakis, Georgios Piliouras, and Lillian Ratliff. Online learning in periodic zero-sum games. *Advances in Neural Information Processing Systems*, 34:10313–10325, 2021.

Virginie Gabrel, Cécile Murat, and Aurélie Thiele. Recent advances in robust optimization: An overview. *European journal of operational research*, 235(3):471–483, 2014.

Nadav Hallak, Panayotis Mertikopoulos, and Volkan Cevher. Regret minimization in stochastic non-convex learning via a proximal-gradient approach. In *International Conference on Machine Learning*, pp. 4008–4017. PMLR, 2021.

Elad Hazan, Karan Singh, and Cyril Zhang. Efficient regret minimization in non-convex games. In *International Conference on Machine Learning*, pp. 1433–1441. PMLR, 2017a.

Elad Hazan, Karan Singh, and Cyril Zhang. Efficient regret minimization in non-convex games. In Doina Precup and Yee Whye Teh (eds.), *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 1433–1441. PMLR, 06–11 Aug 2017b. URL https://proceedings.mlr.press/v70/hazan17a.html.

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

Amélie Héliou, Matthieu Martin, Panayotis Mertikopoulos, and Thibaud Rahier. Online non-convex optimization with imperfect feedback. *Advances in Neural Information Processing Systems*, 33:17224–17235, 2020.

Nicole Immorlica, Karthik Abinav Sankararaman, Robert Schapire, and Aleksandrs Slivkins. Adversarial bandits with knapsacks. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 202–219. IEEE, 2019.

Adam Kalai and Santosh Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3):291–307, 2005.

Weiwei Kong and Renato DC Monteiro. An accelerated inexact proximal point method for solving nonconvex-concave min-max problems. *SIAM Journal on Optimization*, 31(4):2558–2585, 2021.

Antoine Lesage-Landry, Joshua A Taylor, and Iman Shames. Second-order online nonconvex optimization. *IEEE Transactions on Automatic Control*, 66(10):4866–4872, 2020.

Xiaoyu Li and Francesco Orabona. On the convergence of stochastic gradient descent with adaptive stepsizes. In Kamalika Chaudhuri and Masashi Sugiyama (eds.), *The 22nd International Conference on Artificial Intelligence and Statistics, AISTATS 2019, 16-18 April 2019, Naha, Okinawa, Japan*, volume 89 of *Proceedings of Machine Learning Research*, pp. 983–992. PMLR, 2019. URL http://proceedings.mlr.press/v89/li19c.html.

Tianyi Lin, Chi Jin, and Michael Jordan. On gradient descent ascent for nonconvex-concave minimax problems. In Hal Daumé III and Aarti Singh (eds.), *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 6083–6093. PMLR, 13–18 Jul 2020a. URL https://proceedings.mlr.press/v119/lin20a.html.

Tianyi Lin, Chi Jin, and Michael I Jordan. Near-optimal algorithms for minimax optimization. In *Conference on Learning Theory*, pp. 2738–2779. PMLR, 2020b.

Songtao Lu, Ioannis Tsaknakis, Mingyi Hong, and Yongxin Chen. Hybrid block successive approximation for one-sided non-convex min-max problems: algorithms and applications. *IEEE Transactions on Signal Processing*, 68:3676–3691, 2020.

J-F Mertens and Abraham Neyman. Stochastic games. *International Journal of Game Theory*, 10(2):53–66, 1981.

Arkadi Nemirovski, Anatoli B. Juditsky, Guanghui Lan, and Alexander Shapiro. Robust stochastic approximation approach to stochastic programming. *SIAM J. Optim.*, 19(4):1574–1609, 2009. doi: 10.1137/070704277. URL https://doi.org/10.1137/070704277.

Georgy Noarov, Mallesh Pai, and Aaron Roth. Online multiobjective minimax optimization and applications. *arXiv preprint arXiv:2108.03837*, 2021.

Maher Nouiehed, Maziar Sanjabi, Tianjian Huang, Jason D Lee, and Meisam Razaviyayn. Solving a class of non-convex min-max games using iterative first order methods. *Advances in Neural Information Processing Systems*, 32, 2019.

Hassan Rafique, Mingrui Liu, Qihang Lin, and Tianbao Yang. Weakly-convex–concave min–max optimization: provable algorithms and applications in machine learning. *Optimization Methods and Software*, pp. 1–35, 2021.

A. Pallares Rivera, He Wang, and Huan Xu. The online saddle point problem and online convex optimization with knapsacks. *arXiv: Machine Learning*, 2018.

Abhishek Roy, Yifang Chen, Krishnakumar Balasubramanian, and Prasant Mohapatra. Online and bandit algorithms for nonstationary stochastic saddle-point optimization. *ArXiv*, abs/1912.01698, 2019.

Abhishek Roy, Krishnakumar Balasubramanian, Saeed Ghadimi, and Prasant Mohapatra. Stochastic zeroth-order optimization under nonstationarity and nonconvexity. *Journal of Machine Learning Research*, 23 (64):1–47, 2022.

Shai Shalev-Shwartz. Online learning and online convex optimization. *Found. Trends Mach. Learn.*, 4(2): 107–194, 2012. doi: 10.1561/2200000018. URL https://doi.org/10.1561/2200000018.

Arun Sai Suggala and Praneeth Netrapalli. Online non-convex learning: Following the perturbed leader is optimal. In *Algorithmic Learning Theory*, pp. 845–861. PMLR, 2020.

Davoud Ataee Tarzanagh and Laura Balzano. Online bilevel optimization: Regret analysis of online alternating gradient methods. *arXiv preprint arXiv:2207.02829*, 2022.

Kiran K Thekumparampil, Prateek Jain, Praneeth Netrapalli, and Sewoong Oh. Efficient algorithms for smooth minimax optimization. *Advances in Neural Information Processing Systems*, 32, 2019.

Kaiqing Zhang, Zhuoran Yang, and Tamer Başar. Multi-agent reinforcement learning: A selective overview of theories and algorithms. *Handbook of Reinforcement Learning and Control*, pp. 321–384, 2021.

Mengxiao Zhang, Peng Zhao, Haipeng Luo, and Zhi-Hua Zhou. No-regret learning in time-varying zero-sum games. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvári, Gang Niu, and Sivan Sabato (eds.), *International Conference on Machine Learning, ICML 2022, 17-23 July 2022, Baltimore, Maryland, USA*, volume 162 of *Proceedings of Machine Learning Research*, pp. 26772–26808. PMLR, 2022a. URL https://proceedings.mlr.press/v162/zhang22an.html.

Mengxiao Zhang, Peng Zhao, Haipeng Luo, and Zhi-Hua Zhou. No-regret learning in time-varying zero-sum games. In *ICML*, 2022b.

Zhenxun Zhuang, Yunlong Wang, Kezi Yu, and Songtao Lu. No-regret non-convex online meta-learning. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3942–3946. IEEE, 2020.

Appendices

A Missing Proof of Section 4

A.1 Technical Lemma

Recall that $\Phi_{t,w}(\mathbf{x}) = \max_{\mathbf{y}\in\mathcal{Y}} F_{t,w}(\mathbf{x},\mathbf{y})$ and $\mathbf{y}^*_{t,w}(\mathbf{x}) = \arg\max_{\mathbf{y}\in\mathcal{Y}} F_{t,w}(\mathbf{x},\mathbf{y})$. In this section, we first present some technical lemmas to characterize the structure of the function $\Phi_{t,w}$ and $\mathbf{y}^*_{t,w}$ in the nonconvex-strongly-concave setting, which will be essential throughout the analysis.

Lemma A.1. *$\Phi_{t,w}(\cdot)$ is $(\ell+\kappa\ell)$-smooth with $\nabla\Phi_{t,w}(\cdot) = \nabla_{\mathbf{x}} F_{t,w}\left(\cdot, \mathbf{y}^\star_{t,w}(\cdot)\right)$. Also, $\mathbf{y}^\star_{t,w}(\cdot)$ is κ-Lipschitz.*

Proof. Since the averaging maintains the strongly-nonconcavity and smoothness, i.e. $F_{t,w}(\mathbf{x},\mathbf{y})$ is still μ-strongly-convave in $\mathbf{y}$ and ℓ-smooth. Thus, the proof directly follow Lemma 4.3 in Lin et al. (2020a) and we omit the details. □

Furthermore, we derive the following lemma to provide the smoothness property of $\mathbf{y}^*_{t,w}(\mathbf{x})$ with respect to t. In other words, given any fixed $\mathbf{x}$, the movement of $\mathbf{y}^*_{t,w}(\mathbf{x})$ when t changes can be controlled by the variability of environment of the sliding window.

Lemma A.2. *For any $\mathbf{x}\in\mathbb{R}^m$, $t\in[T]$, it holds that*

$$\left\|\mathbf{y}^*_{t-1,w}(\mathbf{x}) - \mathbf{y}^*_{t,w}(\mathbf{x})\right\| \le \frac{\left\|\nabla_{\mathbf{y}} f_{t,w}(\mathbf{x},\mathbf{y}^*_{t,w}(\mathbf{x})) - \nabla_{\mathbf{y}} f_{t-w}(\mathbf{x},\mathbf{y}^*_{t-1,w}(\mathbf{x}))\right\|}{\mu(w-1)}.$$

Proof. By the optimality of $\mathbf{y}^*_{t,w}(\mathbf{x})$ and $\mathbf{y}^*_{t-1,w}(\mathbf{x})$, for $\forall\mathbf{x}$, we have

$$(\mathbf{y}-\mathbf{y}^*_{t,w}(\mathbf{x}))^\top \nabla_{\mathbf{y}} F_{t,w}(\mathbf{x},\mathbf{y}^*_{t,w}(\mathbf{x})) \le 0, \forall \mathbf{y}\in\mathcal{Y}, \tag{7}$$

$$(\mathbf{y}-\mathbf{y}^*_{t-1,w}(\mathbf{x}))^\top \nabla_{\mathbf{y}} F_{t-1,w}(\mathbf{x},\mathbf{y}^*_{t-1,w}(\mathbf{x})) \le 0, \forall \mathbf{y}\in\mathcal{Y}. \tag{8}$$

Summing up Equation (7) with $\mathbf{y} = \mathbf{y}^*_{t-1,w}(\mathbf{x})$ and Equation (8) with $\mathbf{y} = \mathbf{y}^*_{t,w}(\mathbf{x})$ yields that

$$(\mathbf{y}^*_{t-1,w}(\mathbf{x}) - \mathbf{y}^*_{t,w}(\mathbf{x}))^\top (\nabla_{\mathbf{y}} F_{t,w}(\mathbf{x},\mathbf{y}^*_{t,w}(\mathbf{x})) - \nabla_{\mathbf{y}} F_{t-1,w}(\mathbf{x},\mathbf{y}^*_{t-1,w}(\mathbf{x}))) \le 0. \tag{9}$$

By the definition of $F_{t,w}(\mathbf{x},\mathbf{y})$, we have

$$\begin{aligned}
&\nabla_{\mathbf{y}} F_{t,w}(\mathbf{x},\mathbf{y}^*_{t,w}(\mathbf{x})) - \nabla_{\mathbf{y}} F_{t-1,w}(\mathbf{x},\mathbf{y}^*_{t-1,w}(\mathbf{x}))\\
&= \frac{1}{w}\sum_{i=0}^{w-1} \nabla_{\mathbf{y}} f_{t-i}(\mathbf{x},\mathbf{y}^*_{t,w}(\mathbf{x})) - \frac{1}{w}\sum_{i=0}^{w-1} \nabla_{\mathbf{y}} f_{t-i-1}(\mathbf{x},\mathbf{y}^*_{t-1,w}(\mathbf{x}))\\
&= \frac{1}{w}\left\{\nabla_{\mathbf{y}} f_{t,w}(\mathbf{x},\mathbf{y}^*_{t,w}(\mathbf{x})) - \nabla_{\mathbf{y}} f_{t-w}(\mathbf{x},\mathbf{y}^*_{t-1,w}(\mathbf{x}))\right\}\\
&+ \frac{1}{w}\sum_{i=1}^{w-1}\left\{\nabla_{\mathbf{y}} f_{t-i}(\mathbf{x},\mathbf{y}^*_{t,w}(\mathbf{x})) - \nabla_{\mathbf{y}} f_{t-i}(\mathbf{x},\mathbf{y}^*_{t-1,w}(\mathbf{x}))\right\}. \qquad (10)
\end{aligned}$$

Since for any t and fixed $\mathbf{x}$, the $f_t(\mathbf{x},\cdot)$ is μ-strongly-concave, we have

$$\begin{aligned}
&(\mathbf{y}^*_{t-1,w}(\mathbf{x}) - \mathbf{y}^*_{t,w}(\mathbf{x}))^\top \left\{\nabla_{\mathbf{y}} f_{t-i}(\mathbf{x},\mathbf{y}^*_{t-1,w}(\mathbf{x})) - \nabla_{\mathbf{y}} f_{t-i}(\mathbf{x},\mathbf{y}^*_{t,w}(\mathbf{x}))\right\}\\
&+ \mu\left\|(\mathbf{y}^*_{t-1,w}(\mathbf{x}) - \mathbf{y}^*_{t,w}(\mathbf{x}))\right\|^2 \le 0. \qquad (11)
\end{aligned}$$

Plug Equations (10) and (11) into Equation (9), then we have

$$\begin{aligned}
&(\mathbf{y}^*_{t-1,w}(\mathbf{x}) - \mathbf{y}^*_{t,w}(\mathbf{x}))^\top \frac{1}{w}\left\{\nabla_{\mathbf{y}} f_{t,w}(\mathbf{x},\mathbf{y}^*_{t,w}(\mathbf{x})) - \nabla_{\mathbf{y}} f_{t-w}(\mathbf{x},\mathbf{y}^*_{t-1,w}(\mathbf{x}))\right\}\\
&+ \frac{w-1}{w}\mu\left\|(\mathbf{y}^*_{t-1,w}(\mathbf{x}) - \mathbf{y}^*_{t,w}(\mathbf{x}))\right\|^2 \le 0.
\end{aligned}$$

As a result

$$\begin{aligned}
&\frac{w-1}{w}\mu\left\|\mathbf{y}^*_{t-1,w}(\mathbf{x}) - \mathbf{y}^*_{t,w}(\mathbf{x})\right\|^2\\
&\le -(\mathbf{y}^*_{t-1,w}(\mathbf{x}) - \mathbf{y}^*_{t,w}(\mathbf{x}))^\top \frac{1}{w}\left\{\nabla_{\mathbf{y}} f_{t,w}(\mathbf{x},\mathbf{y}^*_{t,w}(\mathbf{x})) - \nabla_{\mathbf{y}} f_{t-w}(\mathbf{x},\mathbf{y}^*_{t-1,w}(\mathbf{x}))\right\}\\
&\le \frac{1}{w}\left\|\mathbf{y}^*_{t-1,w}(\mathbf{x}) - \mathbf{y}^*_{t,w}(\mathbf{x})\right\|\left\|\nabla_{\mathbf{y}} f_{t,w}(\mathbf{x},\mathbf{y}^*_{t,w}(\mathbf{x})) - \nabla_{\mathbf{y}} f_{t-w}(\mathbf{x},\mathbf{y}^*_{t-1,w}(\mathbf{x}))\right\|,
\end{aligned}$$

where the last inequality follows from Cauchy-Schwartz inequality.

Finally, by some algebra manipulation, we finish the proof as following

$$\left\|\mathbf{y}^*_{t-1,w}(\mathbf{x}) - \mathbf{y}^*_{t,w}(\mathbf{x})\right\| \le \frac{\left\|\nabla_{\mathbf{y}} f_{t,w}(\mathbf{x},\mathbf{y}^*_{t,w}(\mathbf{x})) - \nabla_{\mathbf{y}} f_{t-w}(\mathbf{x},\mathbf{y}^*_{t-1,w}(\mathbf{x}))\right\|}{\mu(w-1)}.$$

□

The next lemma provides an upper bound for the gradient norm of $\nabla\Phi_{t,w}$ in term of notions about $\nabla F_{t,w}$, which justifies our design of stop conditions.

Lemma A.3. *Given a pair $(\mathbf{x},\mathbf{y}) \in \mathbb{R}^m \times \mathcal{Y}$, for $t \in [T]$ and $w > 0$, it holds that*

$$\begin{aligned}
\|\nabla\Phi_{t,w}(\mathbf{x})\|^2 \le &\frac{2\kappa^2}{\eta_{\mathbf{y}}^2}\|\mathbf{y} - \mathcal{P}_{\mathcal{Y}}\left(\mathbf{y} + \eta_{\mathbf{y}}\nabla_{\mathbf{y}} F_{t,w}\left(\mathbf{x},\mathbf{y}\right)\right)\|^2\\
&+ 2\|\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x},\mathbf{y})\|^2
\end{aligned}$$

Proof. By Cauchy-Schwartz inequality, we have

$$\begin{aligned}
\|\nabla\Phi_{t,w}(\mathbf{x})\|^2 \le &2\|\nabla\Phi_{t,w}(\mathbf{x}) - \nabla_{\mathbf{x}} F_{t,w}(\mathbf{x},\mathbf{y})\|^2 + 2\|\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x},\mathbf{y})\|^2\\
\le &2\ell^2\|\mathbf{y}^*_{t,w}(\mathbf{x}) - \mathbf{y}\|^2 + 2\|\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x},\mathbf{y})\|^2
\end{aligned}$$

where the last inequality holds by combining Lemma A.1 and the fact that $F_{t,w}$ is ℓ-smooth. Since $F_{t,w}(\mathbf{x},\cdot)$ is μ-strongly-concave over $\mathcal{Y}$, from the global error bound condition in Drusvyatskiy & Lewis (2018), we obtain

$$\|\mathbf{y}^*_{t,w}(\mathbf{x}) - \mathbf{y}\|^2 \le \frac{\kappa^2}{\eta_{\mathbf{y}}^2\ell^2}\|\nabla\Phi_{t,w}(\mathbf{x}) - \nabla_{\mathbf{x}} F_{t,w}(\mathbf{x},\mathbf{y})\|^2$$

Thus, we complete the proof. □

A.2 Local Regret: Proof of Theorem 1

Proof of Theorem 1. Recall the definition of $\Phi_{t,w}$ and notice that

$$\Phi_{t,w}(\mathbf{x}) = \max_{\mathbf{y}\in\mathcal{Y}} \frac{1}{w} \sum_{i=t-w+1}^{t} f_i(\mathbf{x},\mathbf{y}) = \max_{\mathbf{y}\in\mathcal{Y}} \left[F_{t-1,w}(\mathbf{x},\mathbf{y}) + \frac{1}{w}(f_t(\mathbf{x},\mathbf{y}) - f_{t-w}(\mathbf{x},\mathbf{y}))\right]$$

Then

$$\begin{aligned}
&\left\|\nabla\Phi_{t,w}\left(\mathbf{x}_t\right)\right\|^2 = \left\|\nabla_{\mathbf{x}}F_{t,w}(\mathbf{x}_t,\mathbf{y}^*_{t,w}(\mathbf{x}_t))\right\|^2 \\
&= \Big\|\nabla_{\mathbf{x}}F_{t-1,w}(\mathbf{x}_t,\mathbf{y}^*_{t-1,w}(\mathbf{x}_t)) + \nabla_{\mathbf{x}}F_{t-1,w}(\mathbf{x}_t,\mathbf{y}^*_{t,w}(\mathbf{x}_t)) - \nabla_{\mathbf{x}}F_{t-1,w}(\mathbf{x}_t,\mathbf{y}^*_{t-1,w}(\mathbf{x}_t)) \\
&+\frac{1}{w}\left(\nabla_{\mathbf{x}}f_t\left(\mathbf{x}_t,\mathbf{y}^*_{t,w}(\mathbf{x}_t)\right) - \nabla_{\mathbf{x}}f_{t-w}\left(\mathbf{x}_t,\mathbf{y}^*_{t,w}(\mathbf{x}_t)\right)\right)\Big\|^2 \\
&\le 3\left\|\nabla\Phi_{t-1,w}\left(\mathbf{x}_t\right)\right\|^2 + \frac{3\kappa^2}{(w-1)^2}\|\nabla_{\mathbf{y}}f_t\left(\mathbf{x}_t,\mathbf{y}^*_{t,w}(\mathbf{x}_t)\right) - \nabla_{\mathbf{y}}f_{t-w}\left(\mathbf{x}_t,\mathbf{y}^*_{t-1,w}(\mathbf{x}_t)\right)\|^2 \\
&+\frac{3}{w^2}\|\nabla_{\mathbf{x}}f_t\left(\mathbf{x}_t,\mathbf{y}^*_{t,w}(\mathbf{x}_t)\right) - \nabla_{\mathbf{x}}f_{t-w}\left(\mathbf{x}_t,\mathbf{y}^*_{t,w}(\mathbf{x}_t)\right)\|^2,
\end{aligned} \tag{12}$$

where the second term in last inequality follows from that

$$\begin{aligned}
\|\nabla_{\mathbf{x}}F_{t-1,w}(\mathbf{x}_t,\mathbf{y}^*_{t,w}(\mathbf{x}_t)) - \nabla_{\mathbf{x}}F_{t-1,w}(\mathbf{x}_t,\mathbf{y}^*_{t-1,w}(\mathbf{x}_t))\| \\
&\le \ell\|\mathbf{y}^*_{t,w}(\mathbf{x}_t) - \mathbf{y}^*_{t-1,w}(\mathbf{x}_t)\| \\
&\overset{(a)}{\le} \frac{\kappa\left\|\nabla_{\mathbf{y}}f_{t,w}(\mathbf{x},\mathbf{y}^*_{t,w}(\mathbf{x})) - \nabla_{\mathbf{y}}f_{t-w}(\mathbf{x},\mathbf{y}^*_{t-1,w}(\mathbf{x}))\right\|}{(w-1)}
\end{aligned}$$

where (a) is implied by Lemma A.2.

Moreover, for the first term in Equation (12), by Lemma A.3, and the stop condition, we obtain

$$\left\|\nabla\Phi_{t-1,w}\left(\mathbf{x}_t\right)\right\|^2 \le \frac{\delta^2}{w^2}$$

Summing over $t=1,\cdots,T$, and combining the definition of variation measures $V_{\mathbf{x},w}$ and $V_{\mathbf{y},w}$, then we have

$$\Re_{w-NE}(T) = \sum_{t=1}^{T}\|\Phi_{t,w}(\mathbf{x}_t)\|^2 \le \frac{3}{w^2}(T\delta^2 + \frac{(\kappa w)^2}{(w-1)^2}V_{\mathbf{y},w}[T] + V_{\mathbf{x},w}[T])$$

□

A.3 Oracle Queries: Proof of Theorem 2

Denote the sequence generated in the inner loop at time $t\in[T]$ by

$$\begin{aligned}
\mathbf{x}_t^0 = \mathbf{x}_t \quad & \mathbf{x}_t^{k+1} \leftarrow \mathbf{x}_t^k - \eta_{\mathbf{x}}\nabla_{\mathbf{x}}F_{t,w}\left(\mathbf{x}_t^k,\mathbf{y}_t^k\right) \\
\mathbf{y}_t^0 = \mathbf{y}_t \quad & \mathbf{y}_t^{k+1} \leftarrow \mathcal{P}_{\mathcal{Y}}\left(\mathbf{y}_t^k + \eta_{\mathbf{y}}\nabla_{\mathbf{y}}F_{t,w}\left(\mathbf{x}_t^k,\mathbf{y}_t^k\right)\right)
\end{aligned}$$

Let τ_t be the number of times the gradient update is executed at the t-th iteration. Note that $\mathbf{x}_t^{\tau} = \mathbf{x}_{t+1}$ and $\mathbf{y}_t^{\tau} = \mathbf{y}_{t+1}$.

A.3.1 Supporting Lemmas

We present three key lemmas which are important step descent lemmas. In this section, we focus on a crucial quantity, $\delta_{t,w}^k = \left\|\mathbf{y}^\star_{t,w}\left(\mathbf{x}_t^k\right) - \mathbf{y}_t^k\right\|^2$, which are useful for the subsequent analysis. Throughout our analysis, we choose $\eta_{\mathbf{x}} = \frac{1}{8\kappa^3\ell}$ and $\eta_{\mathbf{y}} = \frac{1}{\ell}$.

Lemma A.4. *Denote τ_t the total iteration of inner loop at step t, for $0 \leq k \leq \tau_t - 1$*

$$\Phi_{t,w}\left(\mathbf{x}_t^{k+1}\right) \leq \Phi_{t,w}\left(\mathbf{x}_t^k\right) - \left(\frac{\eta_{\mathbf{x}}}{2} - \eta_{\mathbf{x}}^2\kappa\ell\right)\left\|\nabla_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right)\right\|^2 + \frac{\eta_{\mathbf{x}}\ell^2}{2}\delta_{t,w}^k \tag{13}$$

Proof. Since $\Phi_{t,w}$ is $(\ell + \kappa\ell)$-smooth and $\ell + \kappa\ell \leq 2\kappa\ell$, for any $x, x^+ \in \mathbb{R}^m$, we have

$$\Phi_{t,w}\left(\mathbf{x}^+\right) - \Phi_{t,w}(\mathbf{x}) - \left(\mathbf{x}^+ - \mathbf{x}\right)^\top \nabla\Phi_{t,w}(\mathbf{x}) \leq \kappa\ell\left\|\mathbf{x}^+ - \mathbf{x}\right\|^2$$

Plugging $\mathbf{x}^+ - \mathbf{x} = -\eta_{\mathbf{x}}\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}, \mathbf{y})$ yields that

$$\begin{aligned}\Phi_{t,w}\left(\mathbf{x}^+\right) \leq &\Phi_{t,w}(\mathbf{x}) - \eta_{\mathbf{x}}\left\|\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}, \mathbf{y})\right\|^2 + \eta_{\mathbf{x}}^2\kappa\ell\left\|\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}, \mathbf{y})\right\|^2 \\ &+ \eta_{\mathbf{x}}\left(\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}, \mathbf{y}) - \nabla\Phi_{t,w}(\mathbf{x})\right)^\top \nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}, \mathbf{y})\end{aligned}$$

By Young's inequality, we have

$$\begin{aligned}&\left(\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}, \mathbf{y}) - \nabla\Phi_{t,w}(\mathbf{x})\right)^\top \nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}, \mathbf{y}) \\ &\leq \frac{\|\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}, \mathbf{y}) - \nabla\Phi_{t,w}(\mathbf{x})\|^2 + \|\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}, \mathbf{y})\|^2}{2}\end{aligned}$$

Since $\nabla\Phi_{t,w}(\mathbf{x}) = \nabla_{\mathbf{x}} F_{t,w}\left(\mathbf{x}, \mathbf{y}_{t,w}^*(\mathbf{x})\right)$, we have

$$\|\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}, \mathbf{y}) - \nabla\Phi_{t,w}(\mathbf{x})\|^2 \leq \ell^2\|\mathbf{y} - \mathbf{y}_{t,w}^*(\mathbf{x})\|^2$$

Putting these pieces together, we obtain

$$\begin{aligned}\Phi_{t,w}\left(\mathbf{x}^+\right) \leq &\Phi_{t,w}(\mathbf{x}) - \left(\frac{\eta_{\mathbf{x}}}{2} - \eta_{\mathbf{x}}^2\kappa\ell\right)\left\|\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}, \mathbf{y})\right\|^2 \\ &+ \frac{\eta_{\mathbf{x}}\ell^2}{2}\|\mathbf{y} - \mathbf{y}_{t,w}^*(\mathbf{x})\|^2\end{aligned}$$

□

Lemma A.5. *For any $t, k \geq 0$, the following statement holds true,*

$$\|\mathbf{y}_t^{k+1} - \mathbf{y}_t^k\|^2 \leq \left(4 - \frac{2}{\kappa}\right)\delta_{t,w}^k. \tag{14}$$

Proof. By Young's inequality, we have

$$\begin{aligned}\|\mathbf{y}_t^{k+1} - \mathbf{y}_t^k\|^2 &\leq 2\|\mathbf{y}_t^{k+1} - \mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^k\right)\|^2 + 2\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^k\right) - \mathbf{y}_t^k\|^2 \\ &\leq \left(2\left(1 - \frac{1}{\kappa}\right) + 2\right)\delta_{t,w}^k = \left(4 - \frac{2}{\kappa}\right)\delta_{t,w}^k.\end{aligned}$$

□

Lemma A.6. *Let $\delta_{t,w}^k = \left\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^k\right) - \mathbf{y}_t^k\right\|^2$, the following statement holds true,*

$$\delta_{t,w}^k \leq \left(1 - \frac{1}{2\kappa}\right)\delta_{t,w}^{k-1} + 2\kappa^3\eta_{\mathbf{x}}^2\|\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^{k-1}, \mathbf{y}_t^{k-1})\|^2$$

Proof. Since $f_t(\mathbf{x}, \cdot)$ is μ-strongly concave and $\eta_{\mathbf{y}} = 1/\ell$, we have

$$\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k-1}\right) - \mathbf{y}_t^k\|^2 \leq \left(1 - \frac{1}{\kappa}\right)\delta_{t,w}^{k-1}$$

By Young's inequality, we have

$$\begin{aligned}\delta_{t,w}^k &\leq \left(1+\frac{1}{2(\kappa-1)}\right)\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k-1}\right)-\mathbf{y}_t^k\|^2+(1+2(\kappa-1))\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^k\right)-\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k-1}\right)\|^2\\ &\leq \left(\frac{2\kappa-1}{2\kappa-2}\right)\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k-1}\right)-\mathbf{y}_t^k\|^2+2\kappa\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^k\right)-\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k-1}\right)\|^2\\ &\leq \left(1-\frac{1}{2\kappa}\right)\delta_{t,w}^{k-1}+2\kappa\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^k\right)-\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k-1}\right)\|^2 \qquad (15)\end{aligned}$$

Since $\mathbf{y}_{t,w}^\star(\cdot)$ is κ-Lipschitz, we have

$$\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^k\right)-\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k-1}\right)\|^2\leq 2\kappa^2\|\mathbf{x}_t^k-\mathbf{x}_t^{k-1}\|^2=2\kappa^2\eta_{\mathbf{x}}^2\|\nabla_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1})\|^2.$$

Thus, plug into eq. (15)

$$\delta_{t,w}^k\leq\left(1-\frac{1}{2\kappa}\right)\delta_{t,w}^{k-1}+2\kappa^3\eta_{\mathbf{x}}^2\|\nabla_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1})\|^2$$

□

A.4 Proof of Theorem 2

Proof of Theorem 2. Denote $\gamma=1-\frac{1}{2\kappa}$, from Lemma A.6 and using telescoping we have

$$\delta_{t,w}^k\leq\gamma^k\delta_{t,w}^0+2\kappa^3\eta_{\mathbf{x}}^2\left(\sum_{j=0}^{k-1}\gamma^{k-1-j}\left\|\nabla_{\mathbf{x}}F_{t,w}\left(\mathbf{x}_t^j,\mathbf{y}_t^j\right)\right\|^2\right) \qquad (16)$$

Specially, for $t>1$,

$$\begin{aligned}\delta_{t,w}^0 &= \|\mathbf{y}_t^0-\mathbf{y}_{t,w}^*(\mathbf{x}_t^0)\|^2\\ &\leq 2\|\mathbf{y}_{t-1}^{\tau_{t-1}}-\mathbf{y}_{t-1,w}^*(\mathbf{x}_{t-1}^{\tau_{t-1}})\|^2+2\|\mathbf{y}_{t-1,w}^*(\mathbf{x}_{t-1}^{\tau_{t-1}})-\mathbf{y}_{t,w}^*(\mathbf{x}_{t-1}^{\tau_{t-1}})\|^2\\ &\leq \frac{\delta^2}{\ell^2w^2}+\frac{2}{\mu^2(w-1)^2}\|\nabla_{\mathbf{y}}f_t(\mathbf{x}_{t-1}^{\tau_{t-1}},\mathbf{y}_{t,w}^*(\mathbf{x}_{t-1}^{\tau_{t-1}}))-\nabla_{\mathbf{y}}f_{t-w}(\mathbf{x}_{t-1}^{\tau_{t-1}},\mathbf{y}_{t-1,w}^*(\mathbf{x}_{t-1}^{\tau_{t-1}}))\|^2\end{aligned}$$

Then plug Equation (16) into Equations (13) and (14) from Lemmas A.4 and A.5, and sum over outer loop number.

$$(\frac{\eta_{\mathbf{x}}}{2}-\eta_{\mathbf{x}}^2\kappa\ell-2\kappa^4\eta_{\mathbf{x}}^3\ell^2)\sum_{j=0}^{\tau_t-1}\left\|\nabla_{\mathbf{x}}F_{t,w}\left(\mathbf{x}_t^j,\mathbf{y}_t^j\right)\right\|^2\leq\Phi_{t,w}\left(\mathbf{x}_t\right)-\Phi_{t,w}\left(\mathbf{x}_{t+1}\right)+\kappa\eta_{\mathbf{x}}\ell^2\delta_{t,w}^0$$

$$\sum_{j=0}^{\tau_t-1}\|\mathbf{y}_t^{k+1}-\mathbf{y}_t^k\|^2\leq(8\kappa-4)\delta_{t,w}^0+(16-\frac{8}{\kappa})\kappa^4\eta_{\mathbf{x}}^2\sum_{j=0}^{\tau_t-1}\left\|\nabla_{\mathbf{x}}F_{t,w}\left(\mathbf{x}_t^j,\mathbf{y}_t^j\right)\right\|^2$$

Letting $\eta_{\mathbf{x}}=\frac{1}{8\kappa^3\ell}$, we have

$$\sum_{j=0}^{\tau_t-1}\left\|\nabla_{\mathbf{x}}F_{t,w}\left(\mathbf{x}_t^j,\mathbf{y}_t^j\right)\right\|^2\leq\frac{8}{\eta_{\mathbf{x}}}(\Phi_{t,w}\left(\mathbf{x}_t\right)-\Phi_{t,w}\left(\mathbf{x}_{t+1}\right))+8\kappa\ell^2\delta_{t,w}^0 \qquad (17)$$

$$\sum_{j=0}^{\tau_t-1}(\kappa\ell)^2\|\mathbf{y}_t^{j+1}-\mathbf{y}_t^j\|^2\leq(8\kappa-4)(\kappa\ell)^2\delta_{t,w}^0+\frac{1}{4}\sum_{j=0}^{\tau_t-1}\left\|\nabla_{\mathbf{x}}F_{t,w}\left(\mathbf{x}_t^j,\mathbf{y}_t^j\right)\right\|^2 \qquad (18)$$

Therefore add Equation (17) $\times \frac{\eta_{\mathbf{x}}}{8}$ and Equation (18) $\times \frac{\eta_{\mathbf{x}}}{10}$ we have

$$\frac{\eta_{\mathbf{x}}}{10}\sum_{j=0}^{\tau_t-1}\left[\left\|\nabla_{\mathbf{x}}F_{t,w}\left(\mathbf{x}_t^j,\mathbf{y}_t^j\right)\right\|^2+(\kappa\ell)^2\|\mathbf{y}_t^{j+1}-\mathbf{y}_t^j\|^2\right]\le\left(\Phi_{t,w}\left(\mathbf{x}_t\right)-\Phi_{t,w}\left(\mathbf{x}_{t+1}\right)\right)+\frac{8\ell}{5}\delta_{t,w}^0. \tag{19}$$

Denote $\Phi_{0,w}(\mathbf{x})=0$, we notice that

$$\begin{aligned}
\Phi_{T,w}(\mathbf{x}_T) &= \sum_{t=1}^{T}(\Phi_{t,w}(\mathbf{x}_t)-\Phi_{t-1,w}(\mathbf{x}_{t-1})) \\
&= \sum_{t=1}^{T}(\Phi_{t,w}(\mathbf{x}_t)-\Phi_{t-1,w}(\mathbf{x}_t))+\sum_{t=2}^{T}(\Phi_{t-1,w}(\mathbf{x}_t)-\Phi_{t-1,w}(\mathbf{x}_{t-1})) \\
&= \frac{1}{w}\sum_{t=1}^{T}\left(F_{t-1,w}(\mathbf{x}_t,\mathbf{y}_{t,w}^*(\mathbf{x}_t))-F_{t-1,w}(\mathbf{x}_t,\mathbf{y}_{t-1,w}^*(\mathbf{x}_t))\right) \\
&+\frac{1}{w}\sum_{t=1}^{T}\left(f_t(\mathbf{x}_t,\mathbf{y}_{t,w}^*(\mathbf{x}_t))-f_{t-w}(\mathbf{x}_t,\mathbf{y}_{t,w}^*(\mathbf{x}_t))\right)+\sum_{t=2}^{T}(\Phi_{t-1,w}(\mathbf{x}_t)-\Phi_{t-1,w}(\mathbf{x}_{t-1})) \\
&\overset{(i)}{\le} \frac{1}{w}\sum_{t=1}^{T}\left(f_t(\mathbf{x}_t,\mathbf{y}_{t,w}^*(\mathbf{x}_t))-f_{t-w}(\mathbf{x}_t,\mathbf{y}_{t,w}^*(\mathbf{x}_t))\right)+\sum_{t=2}^{T}(\Phi_{t-1,w}(\mathbf{x}_t)-\Phi_{t-1,w}(\mathbf{x}_{t-1})),
\end{aligned}$$

where (i) follows from that $\mathbf{y}_{t-1,w}^*(\mathbf{x}_t)$ is the maximizer of $F_{t-1,w}(\mathbf{x}_t,\cdot)$.

By some algebra, we have

$$\sum_{t=1}^{T}\Phi_{t,w}(\mathbf{x}_t))-(\Phi_{t,w}(\mathbf{x}_{t+1})\le\frac{1}{w}\sum_{t=1}^{T}\left(f_t(\mathbf{x}_t,\mathbf{y}_{t,w}^*(\mathbf{x}_t))-f_{t-w}(\mathbf{x}_t,\mathbf{y}_{t,w}^*(\mathbf{x}_t))\right)-\Phi_{T+1,w}(\mathbf{x}_{T+1}).$$

Sum Equation (19) over t, we have

$$\begin{aligned}
&\frac{\eta_{\mathbf{x}}}{10}\times\frac{\delta^2}{2w^2}\tau=\frac{\eta_{\mathbf{x}}\delta^2\tau}{20w^2} \\
&\overset{(i)}{\le}\frac{\eta_{\mathbf{x}}}{10}\sum_{t=1}^{T}\sum_{j=0}^{\tau_t-1}\left[\left\|\nabla_{\mathbf{x}}F_{t,w}\left(\mathbf{x}_t^j,\mathbf{y}_t^j\right)\right\|^2+(\kappa\ell)^2\|\mathbf{y}_t^{j+1}-\mathbf{y}_t^j\|^2\right] \\
&\le\sum_{t=1}^{T}(\Phi_{t,w}\left(\mathbf{x}_t\right)-\Phi_{t,w}\left(\mathbf{x}_{t+1}\right))+\frac{8\ell}{5}\sum_{t=1}^{T}\delta_{t,w}^0 \\
&\le\frac{1}{w}\sum_{t=1}^{T}\left(f_t(\mathbf{x}_t,\mathbf{y}_{t,w}^*(\mathbf{x}_t))-f_{t-w}(\mathbf{x}_t,\mathbf{y}_{t,w}^*(\mathbf{x}_t))\right)-\Phi_{T+1,w}(\mathbf{x}_{T+1}) \\
&+\frac{8T\delta^2}{5\ell w^2}+\frac{16\ell}{5\mu^2(w-1)^2}V_{\mathbf{y},w}[T]+\frac{8\ell D^2}{5} \\
&\le\frac{2MT}{w}+M+\frac{8T\delta^2}{5\ell w^2}+\frac{16\ell}{5\mu^2(w-1)^2}V_{\mathbf{y},w}[T]+\frac{8\ell D^2}{5}.
\end{aligned}$$

Hence

$$\tau\le\frac{480\kappa^3\ell MwT}{\delta^2}+256\frac{\kappa^2T}{\mu}+256\frac{w^2\kappa^3\ell^2}{\delta^2}+512\frac{w^2\kappa^5}{(w-1)^2\delta^2}V_{\mathbf{y},w}[T]+\frac{256D^2\kappa^3\ell^2w^2}{\delta^2}.$$

□

B Missing Proof of Section 5

We first make some notation in Algorithm 2 clearly here, $G_t^{k+1} = \frac{1}{w}\sum_{i=t-w+1}^{t}\tilde{\nabla} f_i(\mathbf{x}_t^{k+1}, \mathbf{y}_t^{k+1}) = \tilde{\nabla} F_{t,w}(\mathbf{x}_t^{k+1}, \mathbf{y}_t^{k+1}) = \left(\tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^{k+1}, \mathbf{y}_t^{k+1}), \tilde{\nabla}_{\mathbf{y}} F_{t,w}(\mathbf{x}_t^{k+1}, \mathbf{y}_t^{k+1})\right)$. And for casimplification, we denote $\mathbf{y}_t^k = \mathbf{y}_t^{\tau_t}$ for any $k \geq \tau_t$.

Before our theoretical analysis of Algorithm 2 and proof of Section 5, we define the filtration in Algorithm 2 formally to describe clearly what is known and what is unknown at certain stage.

Definition 4 (Filtration)**.** *For any $t \geq 1$, we denote filtration $\mathcal{F}_t$ to be the σ-fields that corresponds to the randomness of all gradient feedback up to stage $t-1$ and the decision of f_t at stage t. In particular, $\mathcal{F}_t$ includes $f_t, \mathbf{x}_t$ and $\tilde{\nabla} F_{t-1,w}(\mathbf{x}_t, \mathbf{y}_t)$, but doesn't include $\tilde{\nabla} f_t(\mathbf{x}_t, \mathbf{y}_t), \tilde{\nabla} F_{t,w}(\mathbf{x}_t, \mathbf{y}_t)$.*
For any $t \geq 1$, $k \geq 1$, we denote filtration $\mathcal{F}_t^k$ to be the σ-fields that corresponds to the randomness of all gradient feedback up to the k-th iteration in line 6 at stage t in Algorithm 2. In particular, $\mathcal{F}_t^k$ includes $f_t, \mathbf{x}_t^k, \mathbf{y}_t^k, \tilde{\nabla} F_{t,w}(\mathbf{x}_t, \mathbf{y}_t), \{\tilde{\nabla} f_i(\mathbf{x}_t^{k-1}, \mathbf{y}_t^{k-1})\}_{i=t-w+1}^{t}$ and G_t^{k-1}, but doesn't include $G_t^k, \{\tilde{\nabla} f_i(\mathbf{x}_t^k, \mathbf{y}_t^k)\}_{i=t-w+1}^{t}$.

B.1 Finite Iteration: Proof of Theorem 3

B.1.1 Supporting Lemmas

Generally speaking, the lemmas in this section extends lemmas in Appendix A.3.1 to noisy setting. We first provide a descend lemma for $\Phi_{t,w}(\mathbf{x})$ in each iteration of inner loop.

Lemma B.1. *Denote τ_t the total iteration of inner loop at stage t and $\delta_{t,w}^k = \left\|\mathbf{y}_{t,w}^{\star}\left(\mathbf{x}_t^k\right) - \mathbf{y}_t^k\right\|^2$, for $0 \leq k \leq \tau_t - 1$*

$$\begin{aligned}\Phi_{t,w}\left(\mathbf{x}_t^{k+1}\right) \leq & \Phi_{t,w}\left(\mathbf{x}_t^k\right) - \left(\frac{\eta_{\mathbf{x}}}{2} - \eta_{\mathbf{x}}^2\kappa\ell\right)\left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}, \mathbf{y}\right)\right\|^2 + \eta_{\mathbf{x}}\ell^2\delta_{t,w}^k \\ & + \|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}, \mathbf{y}\right) - \nabla_{\mathbf{x}} F_{t,w}\left(\mathbf{x}, \mathbf{y}\right)\|^2\end{aligned}$$

Proof. Since $\Phi_{t,w}$ is $(\ell + \kappa\ell)$-smooth, for any $x, x^+ \in \mathbb{R}^m$, we have

$$\Phi_{t,w}\left(\mathbf{x}^+\right) - \Phi_{t,w}\left(\mathbf{x}\right) - \left(\mathbf{x}^+ - \mathbf{x}\right)^\top \nabla\Phi_{t,w}\left(\mathbf{x}\right) \leq \kappa\ell\left\|\mathbf{x}^+ - \mathbf{x}\right\|^2.$$

Set $\mathbf{x}^+ = \mathbf{x}_t^{k+1}, \mathbf{x} = \mathbf{x}_t^k$, we have $\mathbf{x}^+ - \mathbf{x} = \mathbf{x}_t^{k+1} - \mathbf{x}_t^k = -\eta_{\mathbf{x}}\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right)$, which yeilds that

$$\begin{aligned}\Phi_{t,w}\left(\mathbf{x}_t^{k+1}\right) \leq & \Phi_{t,w}\left(\mathbf{x}_t^k\right) - \eta_{\mathbf{x}}\left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right)\right\|^2 + \eta_{\mathbf{x}}^2\kappa\ell\left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right)\right\|^2 \\ & + \eta_{\mathbf{x}}\left(\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right) - \nabla\Phi_{t,w}\left(\mathbf{x}_t^k\right)\right)^\top \tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right).\end{aligned} \tag{20}$$

By Young's inequality, we have

$$\begin{aligned}& \left(\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right) - \nabla\Phi_{t,w}\left(\mathbf{x}_t^k\right)\right)^\top \tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right) \\ \leq & \frac{\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right) - \nabla\Phi_{t,w}\left(\mathbf{x}_t^k\right)\|^2 + \|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right)\|^2}{2} \\ \leq & \frac{2\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right) - \nabla_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right)\|^2 + 2\|\nabla_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right) - \nabla\Phi_{t,w}\left(\mathbf{x}_t^k\right)\|^2}{2} \\ + & \frac{\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right)\|^2}{2}.\end{aligned} \tag{21}$$

Since $\nabla\Phi_{t,w}\left(\mathbf{x}_t^k\right) = \nabla_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_{t,w}^*(\mathbf{x}_t^k)\right)$, we have

$$\|\nabla_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right) - \nabla\Phi_{t,w}\left(\mathbf{x}_t^k\right)\|^2 \leq \ell^2\|\mathbf{y}_t^k - \mathbf{y}_{t,w}^*(\mathbf{x}_t^k)\|^2. \tag{22}$$

Putting Equations (20) to (22) together, we obtain

$$\Phi_{t,w}\left(\mathbf{x}_t^{k+1}\right) \leq \Phi_{t,w}\left(\mathbf{x}_t^{k}\right) - \left(\frac{\eta_{\mathbf{x}}}{2} - \eta_{\mathbf{x}}^2\kappa\ell\right)\left\|\tilde{\nabla}_{\mathbf{x}}F_{t,w}\left(\mathbf{x}_t^k,\mathbf{y}_t^k\right)\right\|^2 + \eta_{\mathbf{x}}\ell^2\|\mathbf{y}_t^k - \mathbf{y}_{t,w}^*(\mathbf{x}_t^k)\|^2 + \|\tilde{\nabla}_{\mathbf{x}}F_{t,w}\left(\mathbf{x}_t^k,\mathbf{y}_t^k\right) - \nabla_{\mathbf{x}}F_{t,w}\left(\mathbf{x}_t^k,\mathbf{y}_t^k\right)\|^2.$$

□

The next lemma characterizes the descent property of distance to the maximizer $\mathbf{y}_{t,w}^*$.

Lemma B.2. *Let $\delta_{t,w}^k = \left\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^k\right) - \mathbf{y}_t^k\right\|^2$, the following statement holds true,*

$$\delta_{t,w}^k \leq \left(1 - \frac{1}{4\kappa}\right)\delta_{t,w}^{k-1} + 8\kappa^3\eta_{\mathbf{x}}^2\|\tilde{\nabla}_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1})\|^2 + \frac{2\kappa}{\ell^2}\left\|\nabla_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1}) - \tilde{\nabla}_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1})\right\|^2.$$

Proof. Since $f(\mathbf{x},\cdot)$ is μ-strongly concave and $\eta_{\mathbf{y}} = 1/\ell$, we have

$$\begin{aligned}
&\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k-1}\right) - \mathbf{y}_t^k\|^2 \\
&= \left\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k-1}\right) - \mathcal{P}_{\mathcal{Y}}\left(\mathbf{y}_t^{k-1} + \eta_{\mathbf{y}}\tilde{\nabla}_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1})\right)\right\|^2 \\
&= \left\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k-1}\right) - \mathcal{P}_{\mathcal{Y}}\left(\mathbf{y}_t^{k-1} + \eta_{\mathbf{y}}\nabla_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1})\right)\right. \\
&\left.+\mathcal{P}_{\mathcal{Y}}\left(\mathbf{y}_t^{k-1} + \eta_{\mathbf{y}}\nabla_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1})\right) - \mathcal{P}_{\mathcal{Y}}\left(\mathbf{y}_t^{k-1} + \eta_{\mathbf{y}}\tilde{\nabla}_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1})\right)\right\|^2 \\
&\leq \left(1 + \frac{1}{2(\kappa-1)}\right)\left\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k-1}\right) - \mathcal{P}_{\mathcal{Y}}\left(\mathbf{y}_t^{k-1} + \eta_{\mathbf{y}}\nabla_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1})\right)\right\|^2 \\
&+(1+2(\kappa-1))\left\|\mathcal{P}_{\mathcal{Y}}\left(\mathbf{y}_t^{k-1} + \eta_{\mathbf{y}}\nabla_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1})\right) - \mathcal{P}_{\mathcal{Y}}\left(\mathbf{y}_t^{k-1} + \eta_{\mathbf{y}}\tilde{\nabla}_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1})\right)\right\|^2 \\
&\leq \left(1 - \frac{1}{2\kappa}\right)\delta_{t,w}^{k-1} + \frac{2\kappa-1}{\ell^2}\left\|\nabla_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1}) - \tilde{\nabla}_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1})\right\|^2.
\end{aligned} \tag{23}$$

By Young's inequality, we have

$$\begin{aligned}
\delta_{t,w}^k &\leq \left(1 + \frac{1}{2(2\kappa-1)}\right)\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k-1}\right) - \mathbf{y}_t^k\|^2 \\
&+ (1 + 2(2\kappa-1))\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k}\right) - \mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k-1}\right)\|^2 \\
&\leq \left(\frac{4\kappa-1}{2(2\kappa-1)}\right)\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k-1}\right) - \mathbf{y}_t^k\|^2 + 4\kappa\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k}\right) - \mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k-1}\right)\|^2 \\
&\leq \left(1 - \frac{1}{4\kappa}\right)\delta_{t,w}^{k-1} + 4\kappa\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k}\right) - \mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k-1}\right)\|^2 \\
&+ \frac{2\kappa}{\ell^2}\left\|\nabla_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1}) - \tilde{\nabla}_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1})\right\|^2
\end{aligned}$$

Since $\mathbf{y}_{t,w}^\star(\cdot)$ is κ-Lipschitz, we have

$$\begin{aligned}
\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k}\right) - \mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^{k-1}\right)\|^2 &\leq 2\kappa^2\|\mathbf{x}_t^k - \mathbf{x}_t^{k-1}\|^2 \\
&= 2\kappa^2\eta_{\mathbf{x}}^2\|\tilde{\nabla}_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1})\|^2
\end{aligned}$$

Thus, plug into

$$\delta_{t,w}^k \leq \left(1 - \frac{1}{4\kappa}\right)\delta_{t,w}^{k-1} + 8\kappa^3\eta_{\mathbf{x}}^2\|\tilde{\nabla}_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1})\|^2 + \frac{2\kappa}{\ell^2}\left\|\nabla_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1}) - \tilde{\nabla}_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^{k-1},\mathbf{y}_t^{k-1})\right\|^2.$$

□

The next lemma shows that updates over $\mathbf{y}$ can be controlled by $\delta_{t,w}^k$ plus a noisy term.

Lemma B.3. *For any $t, k \geq 0$, the following statement holds true,*

$$\|\mathbf{y}_t^{k+1} - \mathbf{y}_t^k\|^2 \leq (4 - \frac{1}{\kappa})\delta_{t,w}^k + \frac{4\kappa}{\ell^2}\left\|\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^k, \mathbf{y}_t^k) - \tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^k, \mathbf{y}_t^k)\right\|^2.$$

Proof. By Young's inequality, we have

$$\begin{aligned}
\|\mathbf{y}_t^{k+1} - \mathbf{y}_t^k\|^2 &\leq 2\|\mathbf{y}_t^{k+1} - \mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^k\right)\|^2 + 2\|\mathbf{y}_{t,w}^\star\left(\mathbf{x}_t^k\right) - \mathbf{y}_t^k\|^2 \\
&\overset{(i)}{\leq} \left(2(1 - \frac{1}{2\kappa}) + 2\right)\delta_{t,w}^k + \frac{4\kappa}{\ell^2}\left\|\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^k, \mathbf{y}_t^k) - \tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^k, \mathbf{y}_t^k)\right\|^2 \\
&\leq (4 - \frac{1}{\kappa})\delta_{t,w}^k + \frac{4\kappa}{\ell^2}\left\|\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^k, \mathbf{y}_t^k) - \tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^k, \mathbf{y}_t^k)\right\|^2,
\end{aligned}$$

where (i) follows from Equation (23). □

B.1.2 Proof of Theorem 3

Proof. From Lemma B.1

$$\begin{aligned}
\delta_{t,w}^k \leq& \left(1 - \frac{1}{4\kappa}\right)\delta_{t,w}^{k-1} + 8\kappa^3\eta_{\mathbf{x}}^2\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^{k-1}, \mathbf{y}_t^{k-1})\|^2 \\
&+ \frac{2\kappa}{\ell^2}\left\|\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^{k-1}, \mathbf{y}_t^{k-1}) - \tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^{k-1}, \mathbf{y}_t^{k-1})\right\|^2
\end{aligned}$$

Denote $\gamma = 1 - \frac{1}{4\kappa}$, Given $\mathcal{F}_t^{k-1}$ we have

$$\begin{aligned}
\delta_{t,w}^k &\leq \gamma^k \delta_{t,w}^0 + 8\kappa^3\eta_{\mathbf{x}}^2\left(\sum_{j=0}^{k-1}\gamma^{k-1-j}\left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^j, \mathbf{y}_t^j\right)\right\|^2\right) \\
&+ \frac{2\kappa}{\ell^2}\left(\sum_{j=0}^{k-1}\gamma^{k-1-j}\left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^j, \mathbf{y}_t^j) - \nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^j, \mathbf{y}_t^j)\right\|^2\right) \\
&\overset{(i)}{\leq} \gamma^k D^2 + \frac{32\kappa^4\eta_{\mathbf{x}}^2\delta^2}{3w^2} + \frac{2\kappa}{\ell^2}\left(\sum_{j=0}^{k-1}\gamma^{k-1-j}\left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^j, \mathbf{y}_t^j) - \nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^j, \mathbf{y}_t^j)\right\|^2\right),
\end{aligned} \tag{24}$$

where the first term of (i) follows from that $\mathcal{Y}$ is bounded with D, and the second term of (i) follows from the stopping criterion of Algorithm 2 and $\sum_{j=0}^{k-1}\gamma^{k-1-j} \leq 4\kappa$.

Notice that for any fixed t, k and $j \in [k-1]$,

$$\begin{aligned}
&\mathbb{E}\left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^j, \mathbf{y}_t^j) - \nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^j, \mathbf{y}_t^j)\right\|^2 \\
&\overset{(i)}{=} \mathbb{E}_{\mathcal{F}_t^j}\left[\mathbb{E}\left[\left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^j, \mathbf{y}_t^j) - \nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^j, \mathbf{y}_t^j)\right\|^2 \middle| \mathcal{F}_t^j\right]\right] \\
&= \mathbb{E}_{\mathcal{F}_t^j}\left[\frac{1}{w^2}\mathbb{E}\left[\left\|\sum_{i=t-w}^{t-1}\left\{\tilde{\nabla}_{\mathbf{x}} f_i(\mathbf{x}_t^j, \mathbf{y}_t^j) - \nabla_{\mathbf{x}} f_i(\mathbf{x}_t^j, \mathbf{y}_t^j)\right\}\right\|^2 \middle| \mathcal{F}_t^j\right]\right] \\
&\overset{(ii)}{=} \mathbb{E}_{\mathcal{F}_t^j}\left[\frac{1}{w^2}\sum_{i=t-w}^{t-1}\mathbb{E}\left[\left\|\tilde{\nabla}_{\mathbf{x}} f_i(\mathbf{x}_t^j, \mathbf{y}_t^j) - \nabla_{\mathbf{x}} f_i(\mathbf{x}_t^j, \mathbf{y}_t^j)\right\|^2 \middle| \mathcal{F}_t^j\right]\right] \\
&\overset{(iii)}{=} \mathbb{E}_{\mathcal{F}_t^j}\left[\frac{1}{w^2}\cdot w \cdot \frac{\sigma^2}{w^2}\right] = \frac{\sigma^2}{w^3},
\end{aligned} \tag{25}$$

where (i) follows from the property of conditional expectation, (ii) follows from that the SFO calls in line 9 of Algorithm 2 is independent and (iii) follows from definition of SFO and filtration $\mathcal{F}_t^j$.

Thus take expectation over two sides of Equation (24), we have

$$\mathbb{E}\left[\delta_{t,w}^k\right] \leq \gamma^k D^2 + \frac{32\kappa^4\eta_{\mathbf{x}}^2\delta^2}{3w^2} + \frac{8\kappa^2\sigma^2}{\ell^2 w^3}. \tag{26}$$

Then by Lemma B.2

$$\begin{aligned}
&\Phi_{t,w}\left(\mathbf{x}_t^k\right) - \Phi_{t,w}\left(\mathbf{x}_t^{k+1}\right) \\
&\quad\geq \left(\frac{\eta_{\mathbf{x}}}{2} - \eta_{\mathbf{x}}^2\kappa\ell\right)\left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right)\right\|^2 \\
&\quad\quad - \eta_{\mathbf{x}}\ell^2\delta_{t,w}^k - \|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right) - \nabla_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right)\|^2 \\
&\quad\geq \frac{15\eta_{\mathbf{x}}}{32}\left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right)\right\|^2 \\
&\quad\quad - \eta_{\mathbf{x}}\ell^2\delta_{t,w}^k - \|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right) - \nabla_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right)\|^2
\end{aligned} \tag{27}$$

By Lemma B.3

$$\frac{15}{4}\kappa^2\ell^2\eta_{\mathbf{x}}\delta_{t,w}^k + \frac{15}{4\ell}\left\|\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^k, \mathbf{y}_t^k) - \tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^k, \mathbf{y}_t^k)\right\|^2 \geq \frac{15\eta_{\mathbf{x}}}{32} \times 2\kappa^2\ell^2\|\mathbf{y}_t^{k+1} - \mathbf{y}_t^k\|^2. \tag{28}$$

Sum Equation (27) and Equation (28), we have

$$\begin{aligned}
&\Phi_{t,w}\left(\mathbf{x}_t^k\right) - \Phi_{t,w}\left(\mathbf{x}_t^{k+1}\right) + \frac{15}{4}\kappa^2\ell^2\eta_{\mathbf{x}}\delta_{t,w}^k + \frac{15}{4\ell}\left\|\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^k, \mathbf{y}_t^k) - \tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^k, \mathbf{y}_t^k)\right\|^2 \\
&\geq \frac{15\eta_{\mathbf{x}}}{32} \times \left(2\kappa^2\ell^2\|\mathbf{y}_t^{k+1} - \mathbf{y}_t^k\|^2 + \left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right)\right\|^2\right) \\
&\quad - \eta_{\mathbf{x}}\ell^2\delta_{t,w}^k - \|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right) - \nabla_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right)\|^2
\end{aligned}$$

Rearranging the term, we have

$$\begin{aligned}
&\Phi_{t,w}\left(\mathbf{x}_t^k\right) - \Phi_{t,w}\left(\mathbf{x}_t^{k+1}\right) \\
&\quad\geq \frac{15\eta_{\mathbf{x}}}{32} \times \left(2\kappa^2\ell^2\|\mathbf{y}_t^{k+1} - \mathbf{y}_t^k\|^2 + \left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right)\right\|^2\right) \\
&\quad\quad - 5\kappa^2\ell^2\eta_{\mathbf{x}}\delta_{t,w}^k - \left(\frac{15}{4\ell} + 1\right)\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right) - \nabla_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right)\|^2
\end{aligned} \tag{29}$$

Take expectation over both sides of Equation (29), plug into Equation (26) and follow from the similar step of Equation (25), we have

$$\begin{aligned}
&\mathbb{E}\left[\Phi_{t,w}\left(\mathbf{x}_t^k\right) - \Phi_{t,w}\left(\mathbf{x}_t^{k+1}\right)\right] \\
&\geq \frac{5\eta_{\mathbf{x}}\delta^2}{32w^2} - 5\kappa^2\ell^2\eta_{\mathbf{x}}\left(\gamma^k D^2 + \frac{32\kappa^4\eta_{\mathbf{x}}^2\delta^2}{3w^2} + \frac{8\kappa^2\sigma^2}{\ell^2 w^3}\right) - \left(\frac{15}{4\ell} + 1\right)\frac{\sigma^2}{w^3}.
\end{aligned}$$

Because $\gamma = 1 - \frac{1}{4\kappa} \leq 1$, there exist a constant $\tilde{K}$ such that $\gamma^{\tilde{K}} D^2 \leq \max\left\{\frac{32\kappa^4\eta_{\mathbf{x}}^2}{3w^2}, \frac{8\kappa^2\sigma^2}{\ell^2 w^3}\right\}$. Thus for $k \geq \tilde{K}$, we have

$$\begin{aligned}
&\mathbb{E}\left[\Phi_{t,w}\left(\mathbf{x}_t^k\right) - \Phi_{t,w}\left(\mathbf{x}_t^{k+1}\right)\right] \\
&\quad\geq \frac{5\eta_{\mathbf{x}}\delta^2}{32w^2} - 5\kappa^2\ell^2\eta_{\mathbf{x}}\left(\frac{35\kappa^4\eta_{\mathbf{x}}^2\delta^2}{3w^2} + \frac{9\kappa^2\sigma^2}{\ell^2 w^3}\right) - \left(\frac{15}{4\ell} + 1\right)\frac{\sigma^2}{w^3} \\
&\quad\geq \frac{25\eta_{\mathbf{x}}\delta^2}{256w^2} - \frac{45\kappa^4\eta_{\mathbf{x}}\sigma^2}{w^3} - \left(\frac{15}{4\ell} + 1\right)\frac{\sigma^2}{w^3}.
\end{aligned}$$

when $\delta^2 > \frac{2304\kappa^4\sigma^2}{5w} + \frac{256(4\ell+1)\sigma^2}{25\eta_{\mathbf{x}} w}$, we set $\alpha = \frac{25\eta_{\mathbf{x}}\delta^2}{256w^2} - \frac{45\kappa^4\eta_{\mathbf{x}}\sigma^2}{w^3} - \left(\frac{15}{4\ell}+1\right)\frac{\sigma^2}{w^3} > 0$. Then for $K \geq \tilde{K}$, we have

$$\begin{aligned}
2M &\geq \mathbb{E}\left[\Phi_{t,w}\left(\mathbf{x}_t^{\tilde{K}}\right) - \Phi_{t,w}\left(\mathbf{x}_t^{K+1}\right)\right] \\
&= \mathbb{E}\left[\sum_{k=\tilde{K}}^{K}\left(\Phi_{t,w}\left(\mathbf{x}_t^{k}\right) - \Phi_{t,w}\left(\mathbf{x}_t^{k+1}\right)\right)\right] \\
&= \sum_{k=\tilde{K}}^{K}\left(\mathbb{E}\left[\left(\Phi_{t,w}\left(\mathbf{x}_t^{k}\right) - \Phi_{t,w}\left(\mathbf{x}_t^{k+1}\right)\middle|\tau_t \geq k+1\right)\right]\mathbb{P}\left(\tau_t \geq k+1\right) + 0\cdot\mathbb{P}\left(\tau_t < k+1\right)\right) \\
&\geq \alpha\sum_{k=\tilde{K}}^{K}\mathbb{P}\left(\tau_t \geq k+1\right) \\
&\geq \alpha\sum_{k=\tilde{K}}^{K}\mathbb{P}\left(\tau_t > K\right) = \alpha\left(K-\tilde{K}\right)\mathbb{P}\left(\tau_t > K\right),
\end{aligned}$$

where the third equation follows from the Optional Stopping Theorem. Consequently, we have τ_t is finite in probability, which implies that $\tau = \sum_{t=1}^{T}\tau_t$ is finite in probability since it is the finite sum of finite variables in probability. □

B.2 Local Regret: Proof of Theorem 4

Proof of Theorem 4. Following from Equation (12), we have

$$\begin{aligned}
&\left\|\nabla\Phi_{t,w}\left(\mathbf{x}_t\right)\right\|^2 = \left\|\nabla_{\mathbf{x}}F_{t,w}(\mathbf{x}_t, \mathbf{y}^*_{t,w}(\mathbf{x}_t))\right\|^2 \\
&\leq 3\left\|\nabla\Phi_{t-1,w}\left(\mathbf{x}_t\right)\right\|^2 + \frac{3\kappa^2}{(w-1)^2}\|\nabla_{\mathbf{y}}f_t\left(\mathbf{x}_t, \mathbf{y}^*_{t,w}(\mathbf{x}_t)\right) - \nabla_{\mathbf{y}}f_{t-w}\left(\mathbf{x}_t, \mathbf{y}^*_{t-1,w}(\mathbf{x}_t)\right)\|^2 \\
&+ \frac{3}{w^2}\|\nabla_{\mathbf{x}}f_t\left(\mathbf{x}_t, \mathbf{y}^*_{t,w}(\mathbf{x}_t)\right) - \nabla_{\mathbf{x}}f_{t-w}\left(\mathbf{x}_t, \mathbf{y}^*_{t,w}(\mathbf{x}_t)\right)\|^2.
\end{aligned} \tag{30}$$

For the first term

$$\begin{aligned}
&\left\|\nabla\Phi_{t-1,w}(\mathbf{x}_t)\right\|^2 \\
&= \left\|\nabla\Phi_{t-1,w}(\mathbf{x}_{t-1}^{\tau_{t-1}})\right\|^2 \\
&\leq 3\left\|\nabla\Phi_{t-1,w}(\mathbf{x}_{t-1}^{\tau_{t-1}}) - \nabla_{\mathbf{x}}F_{t-1,w}(\mathbf{x}_{t-1}^{\tau_{t-1}}, \mathbf{y}_{t-1}^{\tau_{t-1}})\right\|^2 \\
&+ 3\left\|\nabla_{\mathbf{x}}F_{t-1,w}(\mathbf{x}_{t-1}^{\tau_{t-1}}, \mathbf{y}_{t-1}^{\tau_{t-1}}) - \tilde{\nabla}_{\mathbf{x}}F_{t-1,w}(\mathbf{x}_{t-1}^{\tau_{t-1}}, \mathbf{y}_{t-1}^{\tau_{t-1}})\right\|^2 + 3\left\|\tilde{\nabla}_{\mathbf{x}}F_{t-1,w}(\mathbf{x}_{t-1}^{\tau_{t-1}}, \mathbf{y}_{t-1}^{\tau_{t-1}})\right\|^2 \\
&\leq 3\ell^2\left\|\mathbf{y}^\star_{t-1}(\mathbf{x}_{t-1}^{\tau_{t-1}}) - \mathbf{y}_{t-1}^{\tau_{t-1}}\right\|^2 + 3\left\|\tilde{\nabla}_{\mathbf{x}}F_{t-1,w}(\mathbf{x}_{t-1}^{\tau_{t-1}}, \mathbf{y}_{t-1}^{\tau_{t-1}})\right\|^2 \\
&+ 3\left\|\nabla_{\mathbf{x}}F_{t-1,w}(\mathbf{x}_{t-1}^{\tau_{t-1}}, \mathbf{y}_{t-1}^{\tau_{t-1}}) - \tilde{\nabla}_{\mathbf{x}}F_{t-1,w}(\mathbf{x}_{t-1}^{\tau_{t-1}}, \mathbf{y}_{t-1}^{\tau_{t-1}})\right\|^2.
\end{aligned}$$

Consider $\left\|\mathbf{y}^{\star}_{t-1}(\mathbf{x}^{\tau_{t-1}}_{t-1}) - \mathbf{y}^{\tau_{t-1}}_{t-1}\right\|^2$

$$\begin{aligned}
&\left\|\mathbf{y}^{\star}_{t-1}(\mathbf{x}^{\tau_{t-1}}_{t-1}) - \mathbf{y}^{\tau_{t-1}}_{t-1}\right\|^2 \\
&\overset{(i)}{\leq} \kappa^2 \cdot \frac{1}{\eta_{\mathbf{y}}^2 \ell^2} \left\|\mathbf{y}^{\tau_{t-1}}_{t-1} - \mathcal{P}_{\mathcal{Y}}\left(\mathbf{y}_t + \eta_{\mathbf{y}} \nabla_{\mathbf{y}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1})\right)\right\|^2 \\
&\leq 2\kappa^2 \cdot \frac{1}{\eta_{\mathbf{y}}^2 \ell^2} \left\|\mathbf{y}^{\tau_{t-1}}_{t-1} - \mathcal{P}_{\mathcal{Y}}\left(\mathbf{y}_t + \eta_{\mathbf{y}} \tilde{\nabla}_{\mathbf{y}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1})\right)\right\|^2 \\
&+ 2\kappa^2 \cdot \frac{1}{\eta_{\mathbf{y}}^2 \ell^2} \left\|\mathcal{P}_{\mathcal{Y}}\left(\mathbf{y}_t + \eta_{\mathbf{y}} \nabla_{\mathbf{y}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1})\right) - \mathcal{P}_{\mathcal{Y}}\left(\mathbf{y}_t + \eta_{\mathbf{y}} \tilde{\nabla}_{\mathbf{y}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1})\right)\right\|^2 \\
&\overset{(ii)}{\leq} 2\kappa^2 \cdot \frac{1}{\eta_{\mathbf{y}}^2 \ell^2} \left\|\mathbf{y}^{\tau_{t-1}}_{t-1} - \mathcal{P}_{\mathcal{Y}}\left(\mathbf{y}_t + \eta_{\mathbf{y}} \tilde{\nabla}_{\mathbf{y}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1})\right)\right\|^2 \\
&+ 2\kappa^2 \cdot \frac{1}{\ell^2} \left\|\nabla_{\mathbf{y}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1}) - \tilde{\nabla}_{\mathbf{y}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1})\right\|^2,
\end{aligned}$$

where (i) follows from the global error bound condition in Davis & Drusvyatskiy (2019) and (ii) follows from the project operator is a contraction.

Then

$$\begin{aligned}
&\left\|\nabla \Phi_{t-1,w}(\mathbf{x}_t)\right\|^2 \\
&\leq 6\kappa^2 \cdot \frac{1}{\eta_{\mathbf{y}}^2} \left\|\mathbf{y}^{\tau_{t-1}}_{t-1} - \mathcal{P}_{\mathcal{Y}}\left(\mathbf{y}_t + \eta_{\mathbf{y}} \tilde{\nabla}_{\mathbf{y}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1})\right)\right\|^2 + 3\left\|\tilde{\nabla}_{\mathbf{x}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1})\right\|^2 \\
&+ 6\kappa^2 \left\|\nabla_{\mathbf{y}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1}) - \tilde{\nabla}_{\mathbf{y}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1})\right\|^2 \\
&+ 3\left\|\nabla_{\mathbf{x}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1}) - \tilde{\nabla}_{\mathbf{x}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1})\right\|^2 \\
&\overset{(i)}{\leq} \frac{\delta^2}{w^2} + 6\kappa^2 \left\|\nabla_{\mathbf{y}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1}) - \tilde{\nabla}_{\mathbf{y}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1})\right\|^2 \\
&+ 3\left\|\nabla_{\mathbf{x}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1}) - \tilde{\nabla}_{\mathbf{x}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1})\right\|^2,
\end{aligned} \tag{31}$$

where (i) follows from the stopping condition of inner loop and $\eta_{\mathbf{y}} = 1/\ell$.

Plug Equation (31) into Equation (30) and sum over t, we have

$$\begin{aligned}
R_w(T) &= \sum_{t=1}^{T} \left\|\nabla \Phi_{t,w}(\mathbf{x}_t)\right\|^2 \\
&\leq \sum_{t=1}^{T} \left\{ \frac{3\delta^2}{w^2} + 18\kappa^2 \left\|\nabla_{\mathbf{y}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1}) - \tilde{\nabla}_{\mathbf{y}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1})\right\|^2 \right. \\
&+ 9\left\|\nabla_{\mathbf{x}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1}) - \tilde{\nabla}_{\mathbf{x}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1})\right\|^2 \\
&+ \frac{3\kappa^2}{(w-1)^2} \| \nabla_{\mathbf{y}} f_t\left(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{*}_{t,w}(\mathbf{x}^{\tau_{t-1}}_{t-1})\right) - \nabla_{\mathbf{y}} f_{t-w}\left(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{*}_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1})\right) \|^2 \\
&\left. + \frac{3}{w^2} \left\|\nabla_{\mathbf{x}} f_t\left(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{*}_{t,w}(\mathbf{x}^{\tau_{t-1}}_{t-1})\right) - \nabla_{\mathbf{x}} f_{t-w}\left(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{*}_{t,w}(\mathbf{x}^{\tau_{t-1}}_{t-1})\right)\right\|^2 \right\} \\
&= \frac{3T\delta^2}{w^2} + \frac{3\kappa^2}{(w-1)^2} V_{\mathbf{y},w}[T] + \frac{3}{w^2} V_{\mathbf{x},w}[T] \\
&+ \sum_{t=1}^{T} \left\{ 18\kappa^2 \left\|\nabla_{\mathbf{y}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1}) - \tilde{\nabla}_{\mathbf{y}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1})\right\|^2 \right. \\
&\left. + 9\left\|\nabla_{\mathbf{x}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1}) - \tilde{\nabla}_{\mathbf{x}} F_{t-1,w}(\mathbf{x}^{\tau_{t-1}}_{t-1}, \mathbf{y}^{\tau_{t-1}}_{t-1})\right\|^2 \right\}.
\end{aligned} \tag{32}$$

Notice that for any $t \in [T]$,

$$\begin{aligned}
&\mathbb{E}\left\|\nabla_{\mathbf{y}} F_{t-1,w}(\mathbf{x}_{t-1}^{\tau_{t-1}}, \mathbf{y}_{t-1}^{\tau_{t-1}}) - \tilde{\nabla}_{\mathbf{y}} F_{t-1,w}(\mathbf{x}_{t-1}^{\tau_{t-1}}, \mathbf{y}_{t-1}^{\tau_{t-1}})\right\|^2 \\
&\overset{(i)}{=} \mathbb{E}_{\mathcal{F}_{t-1}^{\tau_{t-1}}}\left[\mathbb{E}\left[\left\|\nabla_{\mathbf{y}} F_{t-1,w}(\mathbf{x}_{t-1}^{\tau_{t-1}}, \mathbf{y}_{t-1}^{\tau_{t-1}}) - \tilde{\nabla}_{\mathbf{y}} F_{t-1,w}(\mathbf{x}_{t-1}^{\tau_{t-1}}, \mathbf{y}_{t-1}^{\tau_{t-1}})\right\|^2 \middle| \mathcal{F}_{t-1}^{\tau_{t-1}}\right]\right] \\
&= \mathbb{E}_{\mathcal{F}_{t-1}^{\tau_{t-1}}}\left[\frac{1}{w^2}\mathbb{E}\left[\left\|\sum_{i=t-w}^{t-1}\{\nabla_{\mathbf{y}} f_i(\mathbf{x}_{t-1}^{\tau_{t-1}}, \mathbf{y}_{t-1}^{\tau_{t-1}}) - \tilde{\nabla}_{\mathbf{y}} f_i(\mathbf{x}_{t-1}^{\tau_{t-1}}, \mathbf{y}_{t-1}^{\tau_{t-1}})\}\right\|^2 \middle| \mathcal{F}_{t-1}^{\tau_{t-1}}\right]\right] \\
&\overset{(ii)}{=} \mathbb{E}_{\mathcal{F}_{t-1}^{\tau_{t-1}}}\left[\frac{1}{w^2}\sum_{i=t-w}^{t-1}\mathbb{E}\left[\left\|\nabla_{\mathbf{y}} f_i(\mathbf{x}_{t-1}^{\tau_{t-1}}, \mathbf{y}_{t-1}^{\tau_{t-1}}) - \tilde{\nabla}_{\mathbf{y}} f_i(\mathbf{x}_{t-1}^{\tau_{t-1}}, \mathbf{y}_{t-1}^{\tau_{t-1}})\right\|^2 \middle| \mathcal{F}_{t-1}^{\tau_{t-1}}\right]\right] \\
&\overset{(iii)}{=} \mathbb{E}_{\mathcal{F}_{t-1}^{\tau_{t-1}}}\left[\frac{1}{w^2}\cdot w \cdot \frac{\sigma^2}{w^2}\right] = \frac{\sigma^2}{w^3},
\end{aligned} \tag{33}$$

where (i) follows from the property of conditional expectation, (ii) follows from that the SFO calls in line 9 of Algorithm 2 is independent and (iii) follows from definition of SFO.

Similarly, for any t, we have

$$\mathbb{E}\left\|\nabla_{\mathbf{x}} F_{t-1,w}(\mathbf{x}_{t-1}^{\tau_{t-1}}, \mathbf{y}_{t-1}^{\tau_{t-1}}) - \tilde{\nabla}_{\mathbf{x}} F_{t-1,w}(\mathbf{x}_{t-1}^{\tau_{t-1}}, \mathbf{y}_{t-1}^{\tau_{t-1}})\right\|^2 = \frac{\sigma^2}{w^3}. \tag{34}$$

Plug Equations (33) and (34) into Equation (32), we have

$$\begin{aligned}
\mathbb{E}\left[\Re_w(T)\right] &= \sum_{t=1}^{T}\mathbb{E}\left[\|\nabla \Phi_{t,w}(\mathbf{x}_t)\|^2\right] \\
&\leq \frac{3T\delta^2}{w^2} + \frac{3\kappa^2}{(w-1)^2}V_{\mathbf{y},w}[T] + \frac{3}{w^2}V_{\mathbf{x},w}[T] + \frac{\left(18\kappa^2+9\right)T\sigma^2}{w^3}.
\end{aligned}$$

□

B.3 Iteration and SFO Calls Bound: Proof of Theorem 5

Proof of Theorem 5. From Lemma B.1

$$\begin{aligned}
\delta_{t,w}^{k} \leq& \left(1 - \frac{1}{4\kappa}\right)\delta_{t,w}^{k-1} + 8\kappa^3\eta_{\mathbf{x}}^2\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^{k-1}, \mathbf{y}_t^{k-1})\|^2 \\
&+ \frac{2\kappa}{\ell^2}\left\|\nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^{k-1}, \mathbf{y}_t^{k-1}) - \tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^{k-1}, \mathbf{y}_t^{k-1})\right\|^2
\end{aligned}$$

Denote $\gamma = 1 - \frac{1}{4\kappa}$, Given $\mathcal{F}_t$ we have

$$\begin{aligned}
\delta_{t,w}^{k} \leq \gamma^k\delta_{t,w}^{0} &+ 8\kappa^3\eta_{\mathbf{x}}^2\left(\sum_{j=0}^{k-1}\gamma^{k-1-j}\left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^{j}, \mathbf{y}_t^{j}\right)\right\|^2\right) \\
&+ \frac{2\kappa}{\ell^2}\left(\sum_{j=0}^{k-1}\gamma^{k-1-j}\left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^{j}, \mathbf{y}_t^{j}) - \nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^{j}, \mathbf{y}_t^{j})\right\|^2\right)
\end{aligned} \tag{35}$$

Then by Lemma B.2

$$\begin{aligned}
\Phi_{t,w}\left(\mathbf{x}_t^{k+1}\right) \leq& \Phi_{t,w}\left(\mathbf{x}_t^{k}\right) - \left(\frac{\eta_{\mathbf{x}}}{2} - \eta_{\mathbf{x}}^2\kappa\ell\right)\left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^{k}, \mathbf{y}_t^{k}\right)\right\|^2 \\
&+ \eta_{\mathbf{x}}\ell^2\delta_{t,w}^{k} + \|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^{k}, \mathbf{y}_t^{k}\right) - \nabla_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^{k}, \mathbf{y}_t^{k}\right)\|^2
\end{aligned} \tag{36}$$

Then plugging Equation (35) into Equation (36) and summing up them over $k=0,\ldots,\tau_t-1$, we have

$$
\begin{aligned}
\Phi_{t,w}(\mathbf{x}_t^{\tau_t}) &\leq \Phi_{t,w}(\mathbf{x}_t^0) - (\frac{\eta_{\mathbf{x}}}{2} - \eta_{\mathbf{x}}^2\kappa\ell)\sum_{k=0}^{\tau_t-1}\left\|\tilde{\nabla}_{\mathbf{x}}F_{t,w}\left(\mathbf{x}_t^k,\mathbf{y}_t^k\right)\right\|^2 + \eta_{\mathbf{x}}\ell^2\sum_{k=0}^{\tau_t-1}\gamma^k\delta_{t,w}^0 \\
&+ 8\kappa^3\eta_{\mathrm{x}}^3\ell^2\sum_{k=0}^{\tau_t-1}\left(\sum_{j=0}^{k-1}\gamma^{k-1-j}\left\|\tilde{\nabla}_{\mathbf{x}}F_{t,w}\left(\mathbf{x}_t^j,\mathbf{y}_t^j\right)\right\|^2\right) \\
&+ 2\eta_{\mathbf{x}}\kappa\sum_{k=0}^{\tau_t-1}\left(\sum_{j=0}^{k-1}\gamma^{k-1-j}\left\|\tilde{\nabla}_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^j,\mathbf{y}_t^j) - \nabla_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^j,\mathbf{y}_t^j)\right\|^2\right) \\
&+ \sum_{k=0}^{\tau_t-1}\|\tilde{\nabla}_{\mathbf{x}}F_{t,w}\left(\mathbf{x}_t^k,\mathbf{y}_t^k\right) - \nabla_{\mathbf{x}}F_{t,w}\left(\mathbf{x}_t^k,\mathbf{y}_t^k\right)\|^2 \\
&\leq \Phi_{t,w}(\mathbf{x}_t^0) - (\frac{\eta_{\mathbf{x}}}{2} - \eta_{\mathbf{x}}^2\kappa\ell - 32\kappa^4\eta_{\mathbf{x}}^3\ell^2)\sum_{k=0}^{\tau_t-1}\left\|\tilde{\nabla}_{\mathbf{x}}F_{t,w}\left(\mathbf{x}_t^k,\mathbf{y}_t^k\right)\right\|^2 \\
&+ 4\kappa\eta_{\mathbf{x}}\ell^2\delta_{t,w}^0 + \left(8\kappa^2\eta_{\mathbf{x}}+1\right)\left(\sum_{k=0}^{\tau_t-1}\left\|\tilde{\nabla}_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^k,\mathbf{y}_t^k) - \nabla_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^k,\mathbf{y}_t^k)\right\|^2\right),
\end{aligned}
$$

where the last inequality follows from that $\sum_{k=0}^{\tau_t-1}\gamma^k = \frac{1-\gamma^{\tau_t}}{1-\gamma} \leq 4\kappa$ and changing the order of summation over j and k.

Rearranging the terms, we have

$$
\begin{aligned}
&(\frac{\eta_{\mathbf{x}}}{2} - \eta_{\mathbf{x}}^2\kappa\ell - 32\kappa^4\eta_{\mathbf{x}}^3\ell^2)\sum_{k=0}^{\tau_t-1}\left\|\tilde{\nabla}_{\mathbf{x}}F_{t,w}\left(\mathbf{x}_t^k,\mathbf{y}_t^k\right)\right\|^2 \\
&\quad\leq \Phi_{t,w}\left(\mathbf{x}_t\right) - \Phi_{t,w}\left(\mathbf{x}_{t+1}\right) + 4\kappa\eta_{\mathbf{x}}\ell^2\delta_{t,w}^0 \\
&\quad+ \left(8\kappa^2\eta_{\mathbf{x}}+1\right)\left(\sum_{k=0}^{\tau_t-1}\left\|\tilde{\nabla}_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^k,\mathbf{y}_t^k) - \nabla_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^k,\mathbf{y}_t^k)\right\|^2\right).
\end{aligned}
$$

By Lemma B.3

$$
\|\mathbf{y}_t^{k+1} - \mathbf{y}_t^k\|^2 \leq (4 - \frac{1}{\kappa})\delta_{t,w}^k + \frac{4\kappa}{\ell^2}\left\|\nabla_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^k,\mathbf{y}_t^k) - \tilde{\nabla}_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^k,\mathbf{y}_t^k)\right\|^2.
$$

Then

$$
\begin{aligned}
\sum_{k=0}^{\tau_t-1}\|\mathbf{y}_t^{k+1} - \mathbf{y}_t^k\|^2 &\leq (16\kappa-4)\delta_{t,w}^0 + 128\kappa^4\eta_{\mathbf{x}}^2\sum_{k=0}^{\tau_t-1}\left\|\tilde{\nabla}_{\mathbf{x}}F_{t,w}\left(\mathbf{x}_t^k,\mathbf{y}_t^k\right)\right\|^2 \\
&\quad+ \frac{36\kappa^2}{\ell^2}\sum_{k=0}^{\tau_t-1}\left\|\tilde{\nabla}_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^k,\mathbf{y}_t^k) - \nabla_{\mathbf{x}}F_{t,w}(\mathbf{x}_t^k,\mathbf{y}_t^k)\right\|^2
\end{aligned}
$$

Notice that $\delta_{0,w}^0 \leq D^2$ and for any $t \geq 2$

$$
\begin{aligned}
\delta_{t,w}^0 &= \|\mathbf{y}_t^0 - \mathbf{y}_{t,w}^*(\mathbf{x}_t^0)\|^2 \\
&\leq 2\|\mathbf{y}_{t-1}^{\tau_{t-1}} - \mathbf{y}_{t-1,w}^*(\mathbf{x}_{t-1}^{\tau_{t-1}})\|^2 + 2\|\mathbf{y}_{t-1,w}^*(\mathbf{x}_{t-1}^{\tau_{t-1}}) - \mathbf{y}_{t,w}^*(\mathbf{x}_{t-1}^{\tau_{t-1}})\|^2 \\
&\leq 2\kappa^2\|\mathbf{y}_{t-1}^{\tau_{t-1}} - \mathcal{P}_{\mathcal{Y}}\left(\mathbf{y}_{t-1}^{\tau_{t-1}} + \eta_{\mathbf{y}}G_{\mathbf{y},t-1}^{\tau_{t-1}}\right)\|^2 \\
&+ \frac{2}{\mu^2(w-1)^2}\|\nabla_{\mathbf{y}}f_t(\mathbf{x}_{t-1}^{\tau_{t-1}},\mathbf{y}_{t,w}^*(\mathbf{x}_{t-1}^{\tau_{t-1}})) - \nabla_{\mathbf{y}}f_{t-w}(\mathbf{x}_{t-1}^{\tau_{t-1}},\mathbf{y}_{t-1,w}^*(\mathbf{x}_{t-1}^{\tau_{t-1}}))\|^2 \\
&\leq \frac{\delta^2}{4\ell^2w^2} + \frac{2}{\mu^2(w-1)^2}\|\nabla_{\mathbf{y}}f_t(\mathbf{x}_{t-1}^{\tau_{t-1}},\mathbf{y}_{t,w}^*(\mathbf{x}_{t-1}^{\tau_{t-1}})) - \nabla_{\mathbf{y}}f_{t-w}(\mathbf{x}_{t-1}^{\tau_{t-1}},\mathbf{y}_{t-1,w}^*(\mathbf{x}_{t-1}^{\tau_{t-1}}))\|^2
\end{aligned}
$$

Letting $\eta_{\mathbf{x}} = \frac{1}{32\kappa^3\ell}$, we have

$$\begin{aligned}\sum_{k=0}^{\tau_t-1} \left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right)\right\|^2 \leq& \frac{16}{7\eta_{\mathbf{x}}}(\Phi_{t,w}\left(\mathbf{x}_t^k\right) - \Phi_{t,w}\left(\mathbf{x}_{t+1}^k\right)) + \frac{64\kappa\ell^2\delta_{t,w}^0}{7} \\ &+ \frac{640}{7\eta_{\mathbf{x}}}\left(\sum_{k=0}^{\tau_t-1} \left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^k, \mathbf{y}_t^k) - \nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^k, \mathbf{y}_t^k)\right\|^2\right) \end{aligned} \tag{37}$$

$$\begin{aligned}\sum_{k=0}^{\tau_t-1} 2(\kappa\ell)^2\|\mathbf{y}_t^{k+1} - \mathbf{y}_t^k\|^2 \leq& 32(\kappa\ell)^2\delta_{t,w}^0 + \frac{1}{8}\sum_{k=0}^{\tau_t-1} \left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right)\right\|^2 \\ &+ 72\kappa^4\left(\sum_{k=0}^{\tau_t-1} \left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^k, \mathbf{y}_t^k) - \nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^k, \mathbf{y}_t^k)\right\|^2\right). \end{aligned} \tag{38}$$

Therefore add Equation (37) $\times \frac{\eta_{\mathbf{x}}}{8}$ and Equation (38) $\times \frac{\eta_{\mathbf{x}}}{9}$, we have

$$\begin{aligned} &\frac{\eta_{\mathbf{x}}}{9}\sum_{k=0}^{\tau_t-1}\left[\left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right)\right\|^2 + 2(\kappa\ell)^2\|\mathbf{y}_t^{k+1} - \mathbf{y}_t^k\|^2\right] \\ &\leq \frac{2}{7}(\Phi_{t,w}\left(\mathbf{x}_t\right) - \Phi_{t,w}\left(\mathbf{x}_{t+1}\right)) + \frac{43\eta_{\mathbf{x}}\kappa^3\ell^2}{9}\delta_{t,w}^0 \\ &+ 20\eta_{\mathbf{x}}\kappa^4\left(\sum_{k=0}^{\tau_t-1} \left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^k, \mathbf{y}_t^k) - \nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^k, \mathbf{y}_t^k)\right\|^2\right). \end{aligned} \tag{39}$$

Denote $\Phi_{0,w}(\mathbf{x}) = 0$, we notice that

$$\begin{aligned} &\Phi_{T,w}(\mathbf{x}_T) \\ &= \sum_{t=1}^{T}(\Phi_{t,w}(\mathbf{x}_t) - \Phi_{t-1,w}(\mathbf{x}_{t-1})) \\ &= \sum_{t=1}^{T}(\Phi_{t,w}(\mathbf{x}_t) - \Phi_{t-1,w}(\mathbf{x}_t)) + \sum_{t=2}^{T}(\Phi_{t-1,w}(\mathbf{x}_t) - \Phi_{t-1,w}(\mathbf{x}_{t-1})) \\ &= \frac{1}{w}\sum_{t=1}^{T}\left(F_{t-1,w}(\mathbf{x}_t, \mathbf{y}_{t,w}^*(\mathbf{x}_t)) - F_{t-1,w}(\mathbf{x}_t, \mathbf{y}_{t-1,w}^*(\mathbf{x}_t))\right) \\ &+ \frac{1}{w}\sum_{t=1}^{T}\left(f_t(\mathbf{x}_t, \mathbf{y}_{t,w}^*(\mathbf{x}_t)) - f_{t-w}(\mathbf{x}_t, \mathbf{y}_{t,w}^*(\mathbf{x}_t))\right) + \sum_{t=2}^{T}(\Phi_{t-1,w}(\mathbf{x}_t) - \Phi_{t-1,w}(\mathbf{x}_{t-1})) \\ &\overset{(i)}{\leq} \frac{1}{w}\sum_{t=1}^{T}\left(f_t(\mathbf{x}_t, \mathbf{y}_{t,w}^*(\mathbf{x}_t)) - f_{t-w}(\mathbf{x}_t, \mathbf{y}_{t,w}^*(\mathbf{x}_t))\right) + \sum_{t=2}^{T}(\Phi_{t-1,w}(\mathbf{x}_t) - \Phi_{t-1,w}(\mathbf{x}_{t-1})), \end{aligned}$$

where (i) follows from that $\mathbf{y}_{t-1,w}^*(\mathbf{x}_t)$ is the maximizer of $F_{t-1,w}(\mathbf{x}_t, \cdot)$.

By some algebra, we have

$$\sum_{t=1}^{T}\Phi_{t,w}(\mathbf{x}_t)) - (\Phi_{t,w}(\mathbf{x}_{t+1}) \leq \frac{1}{w}\sum_{t=1}^{T}\left(f_t(\mathbf{x}_t, \mathbf{y}_{t,w}^*(\mathbf{x}_t)) - f_{t-w}(\mathbf{x}_t, \mathbf{y}_{t,w}^*(\mathbf{x}_t))\right) - \Phi_{T+1,w}(\mathbf{x}_{T+1}).$$

Sum Equation (39) over t and take expectation, we have

$$
\begin{aligned}
&\left(\frac{\delta^2}{27w^2} - 20\kappa^4 \frac{\sigma^2}{w^2}\right)\eta_{\mathbf{x}}\tau \\
&\leq \sum_{t=1}^{T} \frac{\eta_{\mathbf{x}}}{9} \sum_{k=0}^{\tau_t-1} \left[\left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}\left(\mathbf{x}_t^k, \mathbf{y}_t^k\right)\right\|^2 + 2(\kappa\ell)^2 \|\mathbf{y}_t^{k+1} - \mathbf{y}_t^k\|^2\right] \\
&- 20\eta_{\mathbf{x}}\kappa^4 \sum_{t=1}^{T}\left(\sum_{k=0}^{\tau_t-1}\left\|\tilde{\nabla}_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^k, \mathbf{y}_t^k) - \nabla_{\mathbf{x}} F_{t,w}(\mathbf{x}_t^k, \mathbf{y}_t^k)\right\|^2\right) \\
&\leq \sum_{t=1}^{T}(\Phi_{t,w}(\mathbf{x}_t) - \Phi_{t,w}(\mathbf{x}_{t+1})) + \frac{43\eta_{\mathbf{x}}\kappa^3\ell^2}{9}\sum_{t=1}^{T}\delta_{t,w}^0 \\
&\leq \frac{1}{w}\sum_{t=1}^{T}\left(f_t(\mathbf{x}_t, \mathbf{y}_{t,w}^*(\mathbf{x}_t)) - f_{t-w}(\mathbf{x}_t, \mathbf{y}_{t,w}^*(\mathbf{x}_t))\right) - \Phi_{T+1,w}(\mathbf{x}_{T+1}) + \frac{43\eta_{\mathbf{x}}\kappa^3\ell^2}{9}\left\{\frac{(T-1)\delta^2}{4\ell^2w^2}\right. \\
&\left. + \frac{2}{\mu^2(w-1)^2}\sum_{t=2}^{T}\|\nabla_{\mathbf{y}} f_t(\mathbf{x}_{t-1}^{\tau_{t-1}}, \mathbf{y}_{t,w}^*(\mathbf{x}_{t-1}^{\tau_{t-1}})) - \nabla_{\mathbf{y}} f_{t-w}(\mathbf{x}_{t-1}^{\tau_{t-1}}, \mathbf{y}_{t-1,w}^*(\mathbf{x}_{t-1}^{\tau_{t-1}}))\|^2 + D^2\right\} \\
&\leq \frac{2TM}{w} + M + \frac{43T\eta_{\mathbf{x}}\kappa^3\delta^2}{36w^2} + \frac{86\eta_{\mathbf{x}}\kappa^3\ell^2}{9\mu^2(w-1)^2}V_{\mathbf{y},w}[T] + \frac{43\eta_{\mathbf{x}}\kappa^3\ell^2D^2}{9},
\end{aligned}
$$

where the first inequality follows from Assumption 4.

Thus

$$
\tau \leq \frac{1}{\eta_{\mathbf{x}}}\frac{2MTw + \frac{3\delta^2T}{64\ell} + \frac{\ell w^2}{3\mu^2(w-1)^2}V_{\mathbf{y},w}[T] + w^2M + \frac{5\ell D^2w^2}{32}}{\left(\frac{\delta^2}{27} - 20\kappa^4\sigma^2\right)}
$$

$\square$