Online Min-max Optimization: Nonconvexity, Nonstationarity, and Dynamic Regret

Yu Huang  
Tsinghua University

Yuan Cheng  
University of Science and Technology of China

Yingbin Liang  
The Ohio State University

Longbo Huang  
Tsinghua University

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 dynamic 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 [7] 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 [17, 36], online bilevel optimization [39], online federated learning [8], etc. Recently, the online min-max (i.e., saddle point) problem has gained considerable interest due to its broad applications in game theory [34, 43], multi-agent reinforcement learning [5, 42], online robust learning [3, 13], to name a few.

On the theoretical side, a line of works have explored provably efficient algorithms for online min-max optimization. Specifically, [6, 12, 19, 44] considered the zero-sum matrix games where the online objective function takes a bilinear form. [33, 34] studied a more general online min-max problem, where the objective is strongly-convex and strongly-concave. [30] 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 [28, 34, 44], 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., [33] 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 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 [15] 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 $O(\frac{T}{w^2})$ and defer the details to Appx. C.
We verify our theoretical results and demonstrate the effectiveness of our algorithm through several empirical experiments on real-world datasets and defer the details to Appx. D. To our best knowledge, this is the first study on online nonconvex min-max optimization with theoretical characterization of the regret performance.

2. Problem Setup

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

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

where $f_t : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$ is generally nonconvex in $x$ but concave in $y$ 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 $x$-learner and $y$-learner are tasked with predicting $x_t$ and $y_t$ 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 $f_t(x_t, y_t)$.

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

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

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

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

These assumptions are standard in literature of online [16] and min-max optimization [25, 26].

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 [14, 15]. Specifically, for each $t$, we first define the smoothed functions of $f_t$ over a sliding-window of size $w$ as:

$$F_{t,w}(x, y) \overset{def}{=} \frac{1}{w} \sum_{i=0}^{w-1} f_{t-i}(x, y).$$

(2)

For notation convenience, we treat $f_0(x, y)$ as 0 for all $t < 0$. Moreover, since the averaging preserves strongly-convexity, which implies $F_{t,w}$ is strongly-concave in $y$, the maximization problem $\max_{y \in \mathcal{Y}} F_{t,w}(x, y)$ can be solved efficiently. Then, we can naturally define the following function:

$$\Phi_{t,w}(x) \overset{def}{=} \max_{y \in \mathcal{Y}} F_{t,w}(x, y).$$

(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 $\{f_t\}_{t \in [T]}$ satisfy Assumption 1-3, with $\Phi_{t,w}(\cdot)$ defined in eq. (3). The $w$-local Nash Equilibrium (NE)-Regret is defined as:

$$\mathfrak{R}_{w-NE}(T) \overset{def}{=} \sum_{t=1}^{T} \|\nabla \Phi_{t,w}(x_t)\|^2.$$

(4)

$\nabla \Phi_{t,w}$ is well-defined since $\Phi_{t,w}$ is differentiable for nonconvex-strongly-concave min-max problem [25]. We next justify the above notion of the local NE-regret from three aspects1.

---

1. Our local NE-regret naturally is a dynamic regret. Detailed explanation please refer to Appx. A.1
Why norm of gradient as metric? In online convex-concave min-max optimization, it is standard to consider the Nash Equilibrium (NE)-Regret [33] metric, defined as: \( |\sum_{t=1}^T f_t(x_t, y_t) - \min_{x \in \mathbb{R}^m} \max_{y \in \mathbb{Y}} \sum_{t=1}^T f_t(x, y)| \). However, the above metric of NE-regret is inapproparate 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 global optimum of \( \min_{x \in \mathbb{R}^m} \max_{y \in \mathbb{Y}} f_t(x, y) \) in hindsight. Clearly, the problem is equivalent to \( \min_{x \in \mathbb{R}^m} \Phi(x) \), where \( \Phi(x) = \max_{y \in \mathbb{Y}} f_t(x, y) \) is generally nonconvex, and hence finding global minimum for \( \Phi(x) \) is NP hard. A common surogate for the global minimum of \( \Phi \) in the offline nonconvex-strongly-concave min-max literature is the notion of \( \epsilon \)-stationary point \([25, 26]\) for differentiable \( \Phi \), i.e., there exists some iterate \( x_t \) for which \( \|\nabla \Phi(x_t)\| < \epsilon \). If \( \epsilon = 0 \), then \( x_t \) 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 window averaging is two-fold: (i) \( F_{t,w} \) and \( \Phi_{t,w} \) represent 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 nonconvex stochastic optimization literature \([4]\) and can reduce the variation caused by noises. Such smoothed notion is also a common practice in the field of online nonconvex optimization \([2, 14, 15, 45]\). (ii) The average performance itself is also a typical notion people are interested in real applications. Suppose a decision maker in 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 performance metrics.

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 \([34, 44]\) or local \([14]\) 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 \( y^*_{t,w}(x) = \max_{y \in \mathbb{Y}} F_{t,w}(x, y) \).

The sliding-window variation in \( x \) is defined as:

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

Moreover, the sliding-window variation in \( y \) is defined as:

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

**Remark 3** Clearly, \( V_{x,w}[T] \) and \( V_{y,w}[T] \) 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 \( y^*_{t,w} = y^*_{t-1,w} \), 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 \( V_{x,w}[T] = V_{y,w}[T] = 0 \) in the above case.

\[\text{\textsuperscript{2}}\] If we view \( \mathbb{Y} \) to be singleton, the local NE-regret degenerates to local regret proposed in \([15]\).
Algorithm 1 Time-Smoothed Online Gradient Descent Ascent (SODA)

**Input:** window size $w \geq 1$, stepsizes $(\eta_x, \eta_y)$, tolerance $\delta > 0$

**Initialization:** $(x_1, y_1)$

1: for $t = 1$ to $T$ do
2: Predict $(x_t, y_t)$. Observe the cost function $f_t : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$
3: Set $(x_{t+1}, y_{t+1}) \leftarrow (x_t, y_t)$
4: repeat
5: $x_{t+1} \leftarrow x_t - \eta_x \nabla_x F_{t,w}(x_{t+1}, y_{t+1})$
6: $y_{t+1} \leftarrow \mathcal{P}_Y (y_t + \eta_y \nabla_y F_{t,w}(x_{t+1}, y_{t+1}))$
7: until $\frac{\delta^2}{\eta_y^2} \| y_{t+1} - \mathcal{P}_Y (y_t + \eta_y \nabla_y F_{t,w}(x_{t+1}, y_{t+1})) \|^2 + \| \nabla_x F_{t,w}(x_{t+1}, y_{t+1}) \|^2 \leq \frac{\delta^2}{2w^2}$
8: end for

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.

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 $x$ with the stepsize $\eta_x$ and gradient ascent over the variable $y$ with the stepsize $\eta_y$ on function $F_{t,w}(x, y)$ 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 $x_{t+1}$ with small local regret at round $t$, i.e., $\| \nabla \Phi_{t,w}(x_{t+1}) \|^2$ 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 and Lewis [11] to translate conditions on $\nabla \Phi_{t,w}(x_{t+1})$ into restrictions on tractable $\nabla F_{t,w}$. Specifically, we prove that $\| \nabla \Phi_{t,w}(x_{t+1}) \|^2$ is upper bounded by the left-hand side of inequality in Algorithm 1 line 7 (see Theorem 8). 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 [25, 26], which are only guaranteed to visit an $\epsilon$-stationary point within a certain number of iterations, i.e., where the return $\bar{x}$ is uniformly drawn from previous iterations. Crucially, we will establish the total iteration bound (see Theorem 5) 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 4 (Local NE regret minimization)** Let $\kappa = \ell/\mu$ denote the condition number. Under Assumptions 1-3, and letting the stepsizes be chosen as $\eta_x = \Theta (1/\kappa^3 \ell)$ and $\eta_y = \Theta (1/\ell)$, then Algorithm 1 enjoys the following local NE-regret bound:

$$R_{w-NE}(T) = \sum_{t=1}^{T} \| \nabla \Phi_{t,w}(x_t) \|^2 \leq \frac{3}{w^4}(T \delta^2 + \frac{(\kappa w)^2}{(w-1)^2} V_{y,w}[T] + V_{x,w}[T]).$$

**Theorem 5 (Iteration bound)** Let $\tau$ denote the total number of iterations incurred by Algorithm 1. Then $\tau$ can be upper bounded as:

$$\tau \leq \frac{480 e^2 \ell M w T}{\delta^4} + 256 \kappa^2 T \mu + \frac{256 D^2 \kappa^3 \ell w^2}{\delta^2} + 512 \frac{w^2 \kappa^5}{(w-1)^3 \ell^2} V_{y,w}[T].$$

Theorems 4 and 5 together reveal trade-offs between the impact of size $w$ on the regret and computational complexity, where larger $w$ leads to smaller regret bound but incurs more gradient calls.

**Robustness of SODA.** Our results in Theorems 4 and 5 are expressed in terms of variation measures $V_{x,w}[T]$ and $V_{y,w}[T]$ of the environment introduced in Section 3.2. If we make the more restrictive assumption similar to that in [15] 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 $O(T\frac{w}{\kappa^2})$ local NE-regret with $O(Tw)$ iterations since $V_{x,w}[T]$ and $V_{y,w}[T]$ 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 Theorem 3, for the scenario that $f_t$ is periodic with period $w$, $V_{x,w}[T] = V_{y,w}[T] = 0$.

**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 [14, 15]. Consequently, the adversarial example that incurs the local regret of $\Omega(T\frac{w}{\kappa^2})$ constructed in [14] can also serve as a worst case example for our online nonconvex min-max setting. Moreover, under the same assumption made in [15] (which is more restrictive than our assumption here), we achieve a robust regret upper bound of $O(T\frac{w}{\kappa^2})$ (as discussed in the previous paragraph), which matches the worst-case lower bound, indicating that our bound Theorem 4 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 $V_{x,w}[T] = V_{y,w}[T] = 0$ will disappear from our results. Therefore, an immediate implication from our theorems is that GDA is guaranteed to find $\epsilon$-stationary point with $O(\kappa^3 \epsilon^{-2})$ iteration complexity. The best known complexity bound for GDA in offline min-max optimization is $O(\kappa^2 \epsilon^{-2})$ [25]. However, as we discussed in Section 4.1, SODA aims to output $x$ with last-iterate type guarantee, which is a stronger notion than that considered in [25], where GDA are only guaranteed to visit an $\epsilon$-stationary point within a certain number of iterations. Thus, these results are not directly comparable.

5. 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. Furthermore, we conduct experiments with real-world data to validate the theoretical statements and show its superiority in practice.
References


Appendix

A Missing Discussions and Related Work 12
A.1 Why capturing the dynamic nature? 12
A.2 Related Work 13

B Missing Proof of Section 4 13
B.1 Technical Lemma 13
B.2 Local Regret: Proof of Theorem 4 15
B.3 Oracle Queries: Proof of Theorem 5 16
B.4 Proof of Theorem 5 18

C SODA with Stochastic First-order Oracle 20
C.1 Proposed Algorithm 20
C.2 Theoretical Guarantees 21
C.3 Missing Proof of Appx. C.2 22

D Experiments 34
D.1 Implementation Details 34
D.2 Results 35

Notation \([T] \triangleq \{1, \ldots , T\}\). We use bold lower-case letters to denote vectors as in \(\mathbf{x}, \mathbf{y}\), and denote its \(\ell_2\)-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 \(\Phi(\cdot) : \mathbb{R}^m \rightarrow \mathbb{R}\), we let \(\nabla \Phi(\mathbf{x})\) denote the gradient of \(\Phi\) at \(\mathbf{x}\). For a function \(f(\cdot, \cdot) : \mathbb{R}^m \times \mathcal{Y} \rightarrow \mathbb{R}\) of two variables, \(\nabla_x f(\mathbf{x}, \mathbf{y})\) (or \(\nabla_y f(\mathbf{x}, \mathbf{y})\)) 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 \(\nabla f(\mathbf{x}, \mathbf{y}) = (\nabla_x f(\mathbf{x}, \mathbf{y}), \nabla_y f(\mathbf{x}, \mathbf{y}))\). Finally, we use the notation \(O(\cdot)\) and \(\Omega(\cdot)\) to hide constant factors which are independent of problem parameters.

Appendix A. Missing Discussions and Related Work

A.1. 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 [44] 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.
A.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 $A_t$, where the objective function is given by $f_t(x, y) = x^\top A_t y_t$. Several works, for example, [6, 12, 19, 44] proposed and analyzed algorithms with respect to different notions of regret. For more general objective functions, [33, 34] studied the case where the loss function $f_t$ is strongly-convex-strongly-concave. Very recently, [30] 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, [1, 38] studied the performance of “follow the perturbed leader” (FTPL) algorithm [21], and their regrets are all static regret. Further, [2, 14, 15] considered online nonconvex problems under non-stationary environments, and utilized sliding windows method with window size $w$. They proposed different notions of dynamic regrets and algorithms, and achieved an order of $O(T^{2\gamma})$ according regret notions. Additionally, [18] studied online nonconvex optimization with imperfect feedback. Except first-order optimization, [18, 35] considered zeroth-order online nonconvex optimization and [23] 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 [22, 31, 32, 40], 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 [25] and references therein for a good collection of such studies). Another approach, manifesting in the recent works of [27] and [25] considers less complicated single-loop structures. Specifically, the two-timescale GDA analyzed in [25] 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.

**Appendix B. Missing Proof of Section 4**

**B.1. Technical Lemma**

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

**Lemma 6** $\Phi_{t,w}(\cdot)$ is $(\ell + \kappa \ell)$-smooth with $\nabla \Phi_{t,w}(\cdot) = \nabla_x F_{t,w}(\cdot, y^{*}_{t,w}(\cdot))$. Also, $y_{t,w}^{*}(\cdot)$ is $\kappa$-Lipschitz.
By the definition of $F$, summing up Equation (9) with plug Equations (12) and (13) into Equation (11), then we have

$$
\|y_{t-1,w}^*(x) - y_{t,w}^*(x)\| \leq \frac{\|\nabla y_f_{t,w}(x, y_{t,w}^*(x)) - \nabla y_f_{t-w}(x, y_{t-1,w}^*(x))\|}{\mu(w-1)}.
$$

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

$$
(y - y_{t,w}^*(x))^T \nabla y_f_{t,w}(x, y_{t,w}^*(x)) \leq 0, \forall y \in \mathcal{Y}, \tag{9}
$$

$$
(y - y_{t-1,w}^*(x))^T \nabla y_f_{t-1,w}(x, y_{t-1,w}^*(x)) \leq 0, \forall y \in \mathcal{Y}. \tag{10}
$$

Summing up Equation (9) with $y = y_{t-1,w}^*(x)$ and Equation (10) with $y = y_{t,w}^*(x)$ yields that

$$
(y_{t-1,w}^*(x) - y_{t,w}^*(x))^T (\nabla y_f_{t,w}(x, y_{t,w}^*(x)) - \nabla y_f_{t-1,w}(x, y_{t-1,w}^*(x))) \leq 0. \tag{11}
$$

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

$$
\nabla y_f_{t,w}(x, y_{t,w}^*(x)) - \nabla y_f_{t-1,w}(x, y_{t-1,w}^*(x)) \tag{12}
$$

$$
= \frac{1}{w} \sum_{i=0}^{w-1} \nabla y_f_{t-i}(x, y_{t,w}^*(x)) - \frac{1}{w} \sum_{i=0}^{w-1} \nabla y_f_{t-i-1}(x, y_{t-1,w}^*(x))
$$

$$
= \frac{1}{w} \{ \nabla y_f_{t,w}(x, y_{t,w}^*(x)) - \nabla y_f_{t-w}(x, y_{t-1,w}^*(x)) \}
$$

$$
+ \frac{1}{w} \sum_{i=1}^{w-1} \{ \nabla y_f_{t-i}(x, y_{t,w}^*(x)) - \nabla y_f_{t-i}(x, y_{t-1,w}^*(x)) \}.
$$

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

$$
(y_{t-1,w}^*(x) - y_{t,w}^*(x))^T \{ \nabla y_f_{t-i}(x, y_{t-1,w}^*(x)) - \nabla y_f_{t-i}(x, y_{t,w}^*(x)) \}
$$

$$
+ \mu \| (y_{t-1,w}^*(x) - y_{t,w}^*(x)) \|^2 \leq 0. \tag{13}
$$

Plug Equations (12) and (13) into Equation (11), then we have

$$
(y_{t-1,w}^*(x) - y_{t,w}^*(x))^T \frac{1}{w} \{ \nabla y_f_{t,w}(x, y_{t,w}^*(x)) - \nabla y_f_{t-w}(x, y_{t-1,w}^*(x)) \}
$$

$$
+ \frac{w-1}{w} \mu \| (y_{t-1,w}^*(x) - y_{t,w}^*(x)) \|^2 \leq 0.
$$
As a result
\[
\frac{w-1}{w} \mu \| y_{t-1,w}^*(x) - y_{t,w}^*(x) \|^2 \\
\leq - (y_{t-1,w}^*(x) - y_{t,w}^*(x))^\top \frac{1}{w} \{ \nabla_y f_{t,w}(x, y_{t,w}^*(x)) - \nabla_y f_{t-1,w}(x, y_{t-1,w}^*(x)) \} \\
\leq \frac{1}{w} \| y_{t-1,w}^*(x) - y_{t,w}^*(x) \| \| \nabla_y f_{t,w}(x, y_{t,w}^*(x)) - \nabla_y f_{t-1,w}(x, y_{t-1,w}^*(x)) \| ,
\]
where the last inequality follows from Cauchy-Schwartz inequality.

Finally, by some algebra manipulation, we finish the proof as following
\[
\| y_{t-1,w}^*(x) - y_{t,w}^*(x) \| \leq \frac{\| \nabla_y f_{t,w}(x, y_{t,w}^*(x)) - \nabla_y f_{t-1,w}(x, y_{t-1,w}^*(x)) \|}{\mu(w - 1)}.
\]

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 8** Given a pair \( (x, y) \in \mathbb{R}^m \times \mathcal{Y} \), for \( t \in [T] \) and \( w > 0 \), it holds that
\[
\| \nabla \Phi_{t,w}(x) \|^2 \leq \frac{2\kappa^2}{\eta_y^2} \| y - \mathcal{P}_\mathcal{Y}(y + \eta_y \nabla_y F_{t,w}(x, y)) \|^2 \\
+ 2\| \nabla_x F_{t,w}(x, y) \|^2
\]

**Proof** By Cauchy-Schwartz inequality, we have
\[
\| \nabla \Phi_{t,w}(x) \|^2 \leq \| \nabla \Phi_{t,w}(x) - \nabla_x F_{t,w}(x, y) \|^2 + 2\| \nabla_x F_{t,w}(x, y) \|^2 \\
\leq 2\| y_{t,w}^*(x) - y \|^2 + 2\| \nabla_x F_{t,w}(x, y) \|^2
\]
where the last inequality holds by combining Theorem 6 and the fact that \( F_{t,w}(x, \cdot) \) is \( \ell \)-smooth. Since \( F_{t,w}(x, \cdot) \) is \( \mu \)-strongly-concave over \( \mathcal{Y} \), from the global error bound condition in [11], we obtain
\[
\| y_{t,w}^*(x) - y \|^2 \leq \frac{\kappa^2}{\eta_y^2 \ell^2} \| \nabla \Phi_{t,w}(x) - \nabla_x F_{t,w}(x, y) \|^2
\]
Thus, we complete the proof.

**B.2. Local Regret: Proof of Theorem 4**

**Proof** [Proof of Theorem 4] Recall the definition of \( \Phi_{t,w} \) and notice that
\[
\Phi_{t,w}(x) = \max_{y \in \mathcal{Y}} \frac{1}{w} \sum_{i=t-w+1}^{t} f_i(x, y) = \max_{y \in \mathcal{Y}} \left[ F_{t-1,w}(x, y) + \frac{1}{w} (f_t(x, y) - f_{t-w}(x, y)) \right]
\]
Then
\[ \| \nabla \Phi_{t,w}(x_t) \|^2 = \| \nabla_x F_{t,w}(x_t, y_{t,w}^*(x_t)) \|^2 \]
\[ = \| \nabla_x F_{t-1,w}(x_t, y_{t-1,w}^*(x_t)) + \nabla_x F_{t-1,w}(x_t, y_{t-1,w}^*(x_t)) - \nabla_x F_{t-1,w}(x_t, y_{t-1,w}^*(x_t)) \]
\[ + \frac{1}{w} \left( \nabla_x f_t(x_t, y_{t,w}^*(x_t)) - \nabla_x f_{t-w}(x_t, y_{t,w}^*(x_t)) \right) \|^2 \]
\[ \leq 3 \| \nabla \Phi_{t-1,w}(x_t) \|^2 + \frac{3\kappa^2}{(w - 1)^2} \| \nabla y f_t(x_t, y_{t,w}^*(x_t)) - \nabla y f_{t-w}(x_t, y_{t,w}^*(x_t)) \|^2 \]
\[ + \frac{3}{w^2} \| \nabla_x f_t(x_t, y_{t,w}^*(x_t)) - \nabla_x f_{t-w}(x_t, y_{t,w}^*(x_t)) \|^2 , \] (14)
where the second term in last inequality follows from that
\[ \| \nabla_x F_{t-1,w}(x_t, y_{t-1,w}^*(x_t)) - \nabla_x F_{t-1,w}(x_t, y_{t-1,w}^*(x_t)) \| \]
\[ \leq \ell \| y_{t,w}^*(x_t) - y_{t-1,w}^*(x_t) \| \]
\[ \leq \kappa \| \nabla y f_t(x, y_{t,w}^*(x)) - \nabla y f_{t-w}(x, y_{t,w}^*(x)) \| \]
\[ (w - 1) \]
where \((a)\) is implied by Theorem 7.

Moreover, for the first term in Equation (14), by Theorem 8, and the stop condition, we obtain
\[ \| \nabla \Phi_{t-1,w}(x_t) \|^2 \leq \frac{\delta^2}{w^2} \]

Summing over \( t = 1, \cdots, T \), and combining the definition of variation measures \( V_{x,w} \) and \( V_{y,w} \), then we have
\[ \mathcal{R}_{w-N}(T) = \sum_{t=1}^{T} \| \Phi_{t,w}(x_t) \|^2 \leq \frac{3}{w^2} (T\delta^2 + \frac{(\kappa w)^2}{(w - 1)^2} V_{y,w}[T] + V_{x,w}[T]) \]

B.3. Oracle Queries: Proof of Theorem 5

Denote the sequence generated in the inner loop at time \( t \in [T] \) by
\[ x_t^0 = x_t \quad x_t^{k+1} \leftarrow x_t^k - \eta_x \nabla_x F_{t,w}(x_t^k, y_t^k) \]
\[ y_t^0 = y_t \quad y_t^{k+1} \leftarrow \mathcal{P}_y \left( y_t^k + \eta_y \nabla_y F_{t,w}(x_t^k, y_t^k) \right) \]

Let \( \tau_t \) be the number of times the gradient update is executed at the \( t \)-th iteration. Note that \( x_t^{\tau_t} = x_{t+1} \) and \( y_t^{\tau_t} = y_{t+1} \).
B.3.1. Supporting Lemmas

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

**Lemma 9** Denote $\tau_t$ the total iteration of inner loop at step $t$, for $0 \leq k \leq \tau_t - 1$

$$\Phi_{t,w}(x_{t}^{k+1}) \leq \Phi_{t,w}(x_{t}^{k}) - \left( \frac{\eta_x}{2} - \eta_x^2\kappa\ell \right) \left\| \nabla_x F_{t,w}(x_{t}^{k}, y_{t}^{k}) \right\|^2 + \frac{\eta_x^2\ell^2}{2}\delta_{t,w}^{k} \tag{15}$$

**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

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

Plugging $x^+ - x = -\eta_x \nabla_x F_{t,w}(x, y)$ yields that

$$\Phi_{t,w}(x^+) \leq \Phi_{t,w}(x) - \eta_x \left\| \nabla_x F_{t,w}(x, y) \right\|^2 + \eta_x^2\kappa\ell \left\| \nabla_x F_{t,w}(x, y) \right\|^2$$

$$+ \eta_x \left( \nabla_x F_{t,w}(x, y) - \nabla \Phi_{t,w}(x) \right)^\top \nabla_x F_{t,w}(x, y)$$

By Young’s inequality, we have

$$\left( \nabla_x F_{t,w}(x, y) - \nabla \Phi_{t,w}(x) \right)^\top \nabla_x F_{t,w}(x, y) \leq \frac{\left\| \nabla_x F_{t,w}(x, y) - \nabla \Phi_{t,w}(x) \right\|^2 + \left\| \nabla_x F_{t,w}(x, y) \right\|^2}{2}$$

Since $\nabla \Phi_{t,w}(x) = \nabla_x F_{t,w}(x, y_{t,w}^k(x))$, we have

$$\left\| \nabla_x F_{t,w}(x, y) - \nabla \Phi_{t,w}(x) \right\|^2 \leq \ell^2 \left\| y - y_{t,w}^k(x) \right\|^2$$

Putting these pieces together, we obtain

$$\Phi_{t,w}(x^+) \leq \Phi_{t,w}(x) - \left( \frac{\eta_x}{2} - \eta_x^2\kappa\ell \right) \left\| \nabla_x F_{t,w}(x, y) \right\|^2$$

$$+ \frac{\eta_x^2\ell^2}{2} \left\| y - y_{t,w}^k(x) \right\|^2$$

**Lemma 10** For any $t, k \geq 0$, the following statement holds true,

$$\left\| y_{t}^{k+1} - y_{t}^{k} \right\|^2 \leq (4 - \frac{2}{\kappa})\delta_{t,w}^{k} \tag{16}$$

**Proof** By Young’s inequality, we have

$$\left\| y_{t}^{k+1} - y_{t}^{k} \right\|^2 \leq 2\left\| y_{t}^{k+1} - y_{t,w}^k(x_k^t) \right\|^2 + 2\left\| y_{t,w}^k(x_k^t) - y_{t}^{k} \right\|^2$$

$$\leq \left( 2\left( 1 - \frac{1}{\kappa} \right) + 2 \right) \delta_{t,w}^{k} = (4 - \frac{2}{\kappa})\delta_{t,w}^{k}.$$
Lemma 11 Let \( \delta_{t,w}^k = \|y_{t,w}^* (x_t^k) - y_t^k \|^2 \), the following statement holds true,

\[
\delta_{t,w}^k \leq \left( 1 - \frac{1}{2\kappa} \right) \delta_{t,w}^{k-1} + 2\kappa^3 \eta_x^2 \| \nabla_x F_{t,w}(x_t^{k-1}, y_t^{k-1}) \|^2
\]

Proof Since \( f_t(x, \cdot) \) is \( \mu \)-strongly concave and \( \eta_y = 1/\ell \), we have

\[
\|y_{t,w}^* (x_t^{k-1}) - y_t^k \|^2 \leq (1 - \frac{1}{\kappa}) \delta_{t,w}^{k-1}
\]

By Young’s inequality, we have

\[
\delta_{t,w}^k \leq \left( 1 + \frac{1}{2(\kappa - 1)} \right) \|y_{t,w}^* (x_t^{k-1}) - y_t^k \|^2 + (1 + 2(\kappa - 1)) \|y_{t,w}^* (x_t^k) - y_{t,w}^* (x_t^{k-1}) \|^2
\]

\[
\leq \left( \frac{2\kappa - 1}{2\kappa - 2} \right) \|y_{t,w}^* (x_t^{k-1}) - y_t^k \|^2 + 2\kappa \|y_{t,w}^* (x_t^k) - y_{t,w}^* (x_t^{k-1}) \|^2
\]

\[
\leq \left( 1 - \frac{1}{2\kappa} \right) \delta_{t,w}^{k-1} + 2\kappa \|y_{t,w}^* (x_t^k) - y_{t,w}^* (x_t^{k-1}) \|^2
\]

(17)

Since \( y_{t,w}^* (\cdot) \) is \( \kappa \)-Lipschitz, we have

\[
\|y_{t,w}^* (x_t^k) - y_{t,w}^* (x_t^{k-1}) \|^2 \leq 2\kappa^2 \|x_t^k - x_t^{k-1} \|^2 = 2\kappa^2 \eta_x^2 \| \nabla_x F_{t,w}(x_t^{k-1}, y_t^{k-1}) \|^2.
\]

Thus, plug into eq. (17)

\[
\delta_{t,w}^k \leq \left( 1 - \frac{1}{2\kappa} \right) \delta_{t,w}^{k-1} + 2\kappa^3 \eta_x^2 \| \nabla_x F_{t,w}(x_t^{k-1}, y_t^{k-1}) \|^2
\]

B.4. Proof of Theorem 5

Proof [Proof of Theorem 5]

Denote \( \gamma = 1 - \frac{1}{2\kappa} \), from Theorem 11 and using telescoping we have

\[
\delta_{t,w}^k \leq \gamma^{k} \delta_{t,w}^0 + 2\kappa^3 \eta_x^2 \left( \sum_{j=0}^{k-1} \gamma^{k-1-j} \| \nabla_x F_{t,w}(x_j^0, y_j^0) \|^2 \right)
\]

(18)

Specially, for \( t > 1 \),

\[
\delta_{t,w}^0 = \|y_t^0 - y_{t,w}^* (x_t^0) \|^2
\]

\[
\leq 2\|y_{t-1}^{\tau_{t-1}} - y_{t-1,w}^* (x_{t-1}^{\tau_{t-1}}) \|^2 + 2\|y_{t-1,w}^* (x_{t-1}^{\tau_{t-1}}) - y_{t,w}^* (x_{t-1}^{\tau_{t-1}}) \|^2
\]

\[
\leq \frac{\delta^2}{\ell^2 u^2} + \frac{2}{\mu^2 (w - 1)^2} \| \nabla y f_t(x_{t-1}^{\tau_{t-1}}, y_{t,w}^* (x_{t-1}^{\tau_{t-1}})) - \nabla y f_{t-1,w}(x_{t-1}^{\tau_{t-1}}, y_{t-1,w}^* (x_{t-1}^{\tau_{t-1}})) \|^2
\]
Then plug Equation (18) into Equations (15) and (16) from Theorems 9 and 10, and sum over outer loop number.

\[
\left(\eta_x - \eta_x^2 \kappa \ell - 2 \kappa^2 \eta_x^2 \ell^2 \right) \sum_{j=0}^{\tau_t-1} \left\| \nabla_x F_{t,w} \left(x_t^j, y_t^j \right) \right\|^2 \leq \Phi_{t,w}(x_t) - \Phi_{t,w}(x_{t+1}) + \kappa \eta_x \ell^2 \delta_{t,w}^0
\]

\[
\sum_{j=0}^{\tau_t-1} \left\| y_t^{k+1} - y_t^k \right\|^2 \leq (8 \kappa - 4) \delta_{t,w}^0 + (16 - \frac{8}{\kappa}) \kappa^2 \eta_x^2 \sum_{j=0}^{\tau_t-1} \left\| \nabla_x F_{t,w} \left(x_t^j, y_t^j \right) \right\|^2
\]

Letting \( \eta_x = \frac{1}{8 \kappa \ell^2} \), we have

\[
\sum_{j=0}^{\tau_t-1} \left\| \nabla_x F_{t,w} \left(x_t^j, y_t^j \right) \right\|^2 \leq \frac{8}{\eta_x} \left( \Phi_{t,w}(x_t) - \Phi_{t,w}(x_{t+1}) \right) + 8 \kappa \ell^2 \delta_{t,w}^0
\]  
(19)

\[
\sum_{j=0}^{\tau_t-1} (\kappa \ell)^2 \left\| y_t^{j+1} - y_t^j \right\|^2 \leq (8 \kappa - 4)(\kappa \ell)^2 \delta_{t,w}^0 + \frac{1}{4} \sum_{j=0}^{\tau_t-1} \left\| \nabla_x F_{t,w} \left(x_t^j, y_t^j \right) \right\|^2
\]  
(20)

Therefore add Equation (19) \( \times \frac{\eta_x}{5} \) and Equation (20) \( \times \frac{\eta_x}{10} \) we have

\[
\frac{\eta_x}{10} \sum_{j=0}^{\tau_t-1} \left[ \left\| \nabla_x F_{t,w} \left(x_t^j, y_t^j \right) \right\|^2 + (\kappa \ell)^2 \left\| y_t^{j+1} - y_t^j \right\|^2 \right] \leq \left( \Phi_{t,w}(x_t) - \Phi_{t,w}(x_{t+1}) \right) + \frac{8 \ell}{5} \delta_{t,w}^0.
\]  
(21)

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

\[
\Phi_{T,w}(x_T) = \sum_{t=1}^{T} \left( \Phi_{t,w}(x_t) - \Phi_{t-1,w}(x_{t-1}) \right)
\]

\[
= \sum_{t=1}^{T} \left( \Phi_{t,w}(x_t) - \Phi_{t-1,w}(x_{t-1}) \right) + \sum_{t=2}^{T} \left( \Phi_{t-1,w}(x_t) - \Phi_{t-1,w}(x_{t-1}) \right)
\]

\[
= \frac{1}{w} \sum_{t=1}^{T} \left( F_{t-1,w}(x_t, y_{t-1,w}^*(x_t)) - F_{t-1,w}(x_t, y_{t-1,w}^*(x_t)) \right)
\]

\[
+ \frac{1}{w} \sum_{t=1}^{T} \left( F_{t-1,w}(x_t, y_{t-1,w}^*(x_t)) - f_{t-1,w}(x_t, y_{t-1,w}^*(x_t)) \right) + \frac{1}{w} \sum_{t=2}^{T} \left( \Phi_{t-1,w}(x_t) - \Phi_{t-1,w}(x_{t-1}) \right)
\]

\[
\leq \frac{1}{w} \sum_{t=1}^{T} \left( F_{t-1,w}(x_t, y_{t-1,w}^*(x_t)) - f_{t-1,w}(x_t, y_{t-1,w}^*(x_t)) \right) + \frac{1}{w} \sum_{t=2}^{T} \left( \Phi_{t-1,w}(x_t) - \Phi_{t-1,w}(x_{t-1}) \right),
\]

where \((i)\) follows from that \( y_{t-1,w}^*(x_t) \) is the maximizer of \( F_{t-1,w}(x_t, \cdot) \).

By some algebra, we have

\[
\sum_{t=1}^{T} \Phi_{t,w}(x_t) - \Phi_{t,w}(x_{t+1}) \leq \frac{1}{w} \sum_{t=1}^{T} \left( f_t(x_t, y_{t,w}^*(x_t)) - f_{t-1,w}(x_t, y_{t-1,w}^*(x_t)) \right) - \Phi_{T+1,w}(x_{T+1}).
\]
Sum Equation (21) over $t$, we have

\[
\frac{\eta_k}{10} \times \frac{\delta^2}{2w^2} \tau = \frac{\eta_k \delta^2 \tau}{20w^2} \leq \sum_{t=1}^{T} \sum_{j=0}^{T-1} \left[ \left\| \nabla_x F_{t,w}(x^j_t, y^j_t) \right\|^2 + (\kappa \ell)^2 \| y^{j+1}_t - y^j_t \|^2 \right]
\]

\[
\leq \sum_{t=1}^{T} (\Phi_{t,w}(x_t) - \Phi_{t,w}(x_{t+1})) + \frac{8\ell}{5} \sum_{t=1}^{T} \delta^0_{t,w}
\]

\[
\leq \frac{1}{w} \sum_{t=1}^{T} \left( f_t(x_t, y^*_t, w(x_t)) - f_{t-w}(x_t, y^*_{t-w}(x_t)) \right) - \Phi_{T+1,w}(x_{T+1})
\]

\[
+ \frac{8T \delta^2}{5\ell w^2} + \frac{16\ell}{5\mu^2(w-1)^2} V_{y,w}[T] + \frac{8\ell D^2}{5}
\]

\[
\leq \frac{2MT}{w} + M + \frac{8T \delta^2}{5\ell w^2} + \frac{16\ell}{5\mu^2(w-1)^2} V_{y,w}[T] + \frac{8\ell D^2}{5}.
\]

Hence

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

Appendix C. 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 [14, 15, 29]. Specifically, the formal SFO definition is as follows.

**Definition 12** (Stochastic first-order oracle) A stochastic first-order oracle (SFO) is a function $S_\sigma$ such that, given a point $(x, y) \in \mathbb{R}^m \times \mathcal{Y}$, a random seed $\zeta$, and a smooth function $h : \mathbb{R}^m \times \mathcal{Y} \to \mathbb{R}$ satisfies:

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

C.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
Consider the following algorithm:

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

**Input:** window size $w \geq 1$, step sizes $(\eta_x, \eta_y)$, tolerance $\delta > 0$

**Initialization:** $(x_1, y_1)$

1. for $t = 1$ to $T$
   1. Cost function $f_t : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$ is updated;
   2. Sample $\nabla f_t(x_t, y_t) \leftarrow S_{f/w}(x_t, y_t; \zeta, f_t)$
   3. Set $\nabla F_{t,w}(x_t, y_t) = \nabla F_{t-1,w}(x_t, y_t) + \frac{1}{w} (\nabla f_t(x_t, y_t) - \nabla f_t(x_t, y_t))$
   4. Set $x_t^0 = x_t, y_t^0 = y_t, G^0_{x,t} = \nabla F_{t,w}(x_t, y_t), G^0_{y,t} = \nabla F_{t,w}(x_t, y_t), k = 0$
   5. while $\frac{2w^2}{\eta_x^2} \left\| y_t^k - P_{Y} \left( y_t^k + \eta_y G^k_{y,t} \right) \right\|^2 + \left\| G^k_{x,t} \right\|^2 > \delta^2 / 3w^2$
   6. $x_t^{k+1} \leftarrow x_t^k - \eta_x G^k_{x,t}$
   7. $y_t^{k+1} \leftarrow P_{Y} \left( y_t^k + \eta_y G^k_{y,t} \right)$
   8. Sample $\tilde{\nabla} f_i(x_t^{k+1}, y_t^{k+1}) \leftarrow S_{\tilde{\nabla}}(x_t^{k+1}, y_t^{k+1}; \zeta, f_i)$ for $i = t - w + 1, \ldots, t$
   9. Set $G^{k+1} := (G^{k+1}_{t,x}, G^{k+1}_{t,y}) = \frac{1}{w} \sum_{i=t-w+1}^{t} \tilde{\nabla} f_i(x_t^{k+1}, y_t^{k+1})$
   10. $k \leftarrow k + 1$
   11. end while
   12. $x_{t+1} = x_t^k, y_{t+1} = y_t^k, \text{ and } \nabla F_{t,w}(x_{t+1}, y_{t+1}) = G^k_t$
   13. end for

In different coefficients in stop criterion compared to SODA. (ii) The stopping criterion in SODA-SFO only ensures that $\left\| \nabla \Phi_{t,w}(x_{t+1}) \right\|^2$ is bounded by the 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.

### C.2. Theoretical Guarantees

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

**Theorem 13 (Finite iteration with SFO)** Let $\kappa = \ell / \mu$ denote the condition number, and let the step sizes be chosen as $\eta_x = \Theta \left( 1 / (k^3 \ell) \right)$ and $\eta_y = \Theta (1 / \ell)$. Under Assumptions 1-3, for any $\epsilon \in [T]$, if $\delta, w$ and $\sigma$ satisfy that $\delta^2 = O(\kappa^4 \ell^2 \sigma^2 / w)$, then $\tau_i$ and $\tau$ is finite with high probability. Specially, when $K \in \mathbb{R}$ is large enough, $P(\tau_i > K) = O(1/K)$.

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 $w, T, V_{x,w}[T], V_{y,w}[T]$.

**Theorem 14 (Expectation local NE-regret with SFO)** Under the setting of Theorem 13, SODA-SFO enjoys the following expectation local NE regret bound:

$$
E \left[ R_{w-N_E}(T) \right] \leq \frac{T}{w^2} \left( 3\delta^2 + \frac{(18\kappa^2 + 9)\sigma^2}{w} \right) + \frac{3\kappa^2}{(w-1)^2} V_{y,w}[T] + \frac{3}{w} V_{x,w}[T].
$$
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 5. In the stochastic online nonconvex min-max setting, we further provide an upper bound on the complexity of SFO calls as in Theorem 16. Furthermore, Theorem 16 can provide deterministic guarantees rather than high probability guarantees because Assumption 4 controls the variation of noise in an absolute and deterministic manner.

**Remark 15** We remark here that Theorems 13 and 14 do not require Assumption 4, and Theorem 13 provide the finite iteration guarantee with high probability and Theorem 14 provides an upper bound for expectation regret. With Assumption 4, which is slightly stronger than the assumptions in Theorem 12, we are able to provide the following deterministic bound on iterations and the number of SFO calls as in Theorem 16. Furthermore, Theorem 16 can provide deterministic guarantees rather than high probability guarantees because Assumption 4 controls the variation of noise in an absolute and deterministic manner.

**Theorem 16 (Iterations and SFO calls bounds)** Under the setting of Theorem 13 and Assumption 4, and suppose that \( \sigma^2 > 540\kappa^4\sigma^2 \). Then the total number of iterations satisfies

\[
\tau \leq \frac{1}{\eta_k} \left( \frac{2MT \ell \chi^2}{3\chi^2} + \frac{\ell \mu^2 \chi^2}{3\mu^2(\chi - 1)^2} V_{Y_\ell}[T] + \chi^2 M + \frac{5\ell D^2 \chi^2}{92} \right). 
\]

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

The above results also provide a robust guarantee for SODA-SFO, where SODA-SFO achieves an expectation regret of \( O(\frac{T \chi}{w}) \) with at most \( O(T \chi^2) \) iterations and hence \( O(T \chi^2) \) calls of SFO, as long as \( V_{Y_\ell}[T] \) and \( V_{Y_\ell}[T] \) scale with \( O(T) \). Following the similar discussions from Theorem 3 and Section 4.2, such condition can hold with relaxed assumptions depending on nonstationarity.

Specially, if the variance of SFO defined in Theorem 12 is zero, then SFO reduces to perfect first order feedback. Hence, as discussed in Section 4.2, the adversarial example provided by [15] is also applicable to the stochastic setting, and thus indicates that the expectation regret \( O(\frac{T \chi}{w}) \) 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 \( V_{Y_\ell}[T] \) will disappear in our analysis, and our theorems recover the results in [14].

**C.3. Missing Proof of Appx. C.2**

In this section, we first make some notation clearly here, \( G_t^{k+1} = \frac{1}{w} \sum_{i=t-w+1}^{t} \hat{\nabla} f_i(x_t^{k+1}, y_t^{k+1}) = \hat{\nabla} F_t,w(x_t^{k+1}, y_t^{k+1}) \) and \( \hat{\nabla} F_t,w(x_t^{k+1}, y_t^{k+1}) \) denotes \( \hat{\nabla} F_t,w(x_t^{k+1}, y_t^{k+1}) \). And for simplification, we denote \( y_t^{k} = y_t^{\tau_i} \) for any \( k \geq \tau_i \).

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

**Definition 17 (Filtration)** For any \( t \geq 1 \), we denote filtration \( \mathcal{F}_t \) to be the \( \sigma \)-fields that corresponds to the randomness of all gradient feedback up to stage \( t - 1 \) and the decision of \( f_i \) at stage \( t \). In
particular, \( F_t \) includes \( f_t, x_t \) and \( \bar{\nabla} F_{t-1,w}(x_t, y_t) \), but doesn't include \( \bar{\nabla} f_t(x_t, y_t), \bar{\nabla} F_{t,w}(x_t, y_t) \). For any \( t \geq 1, k \geq 1 \), we denote filtration \( F^k_t \) to be the \( \sigma \)-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, \( F^k_t \) includes \( f_t, x_t^k, y_t^k, \bar{\nabla} F_{t,w}(x_t, y_t), \{ \bar{\nabla} f_i(x_t, y_t) \}_{t-w+1} \) and \( G^{k-1}_t \), but doesn’t include \( G^k_t, \{ \bar{\nabla} f_i(x_t^k, y_t^k) \}_{t-w+1} \).

C.3.1. Supporting Lemmas

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

**Lemma 18** Denote \( \tau_t \) the total iteration of inner loop at stage \( t \) and \( \delta_{t,w}^k = \| y_{t,w}^*(x_t^k) - y_t^k \|^2 \), for \( 0 \leq k \leq \tau_t - 1 

\[
\Phi_{t,w}(x_t^{k+1}) \leq \Phi_{t,w}(x_t^k) - \left( \frac{\eta x}{2} - \eta_x^2 \kappa \ell \right) \| \bar{\nabla}_x F_{t,w}(x,y) \|^2 + \eta x^2 \delta_{t,w}^k \\
+ \| \bar{\nabla}_x F_{t,w}(x,y) - \bar{\nabla}_x F_{t,w}(x, y) \|^2
\]

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

\[
\Phi_{t,w}(x^+) - \Phi_{t,w}(x) - (x^+ - x)^\top \nabla \Phi_{t,w}(x) \leq \kappa \ell \| x^+ - x \|^2.
\]

Set \( x^+ = x_t^{k+1}, x = x_t^k \), we have \( x^+ - x = x_t^{k+1} - x_t^k = -\eta x \bar{\nabla}_x F_{t,w}(x_t^k, y_t^k) \), which yields that

\[
\Phi_{t,w}(x_t^{k+1}) \leq \Phi_{t,w}(x_t^k) - \eta x \| \bar{\nabla}_x F_{t,w}(x_t^k, y_t^k) \|^2 + \eta x^2 \kappa \ell \| \bar{\nabla}_x F_{t,w}(x_t^k, y_t^k) \|^2 \\
+ \eta x (\bar{\nabla}_x F_{t,w}(x_t^k, y_t^k) - \nabla \Phi_{t,w}(x_t^k))^\top \bar{\nabla}_x F_{t,w}(x_t^k, y_t^k).
\]

By Young’s inequality, we have

\[
\left( \bar{\nabla}_x F_{t,w}(x_t^k, y_t^k) - \nabla \Phi_{t,w}(x_t^k) \right)^\top \bar{\nabla}_x F_{t,w}(x_t^k, y_t^k) \\
\leq \frac{\| \bar{\nabla}_x F_{t,w}(x_t^k, y_t^k) - \nabla \Phi_{t,w}(x_t^k) \|^2}{2} + \frac{\| \bar{\nabla}_x F_{t,w}(x_t^k, y_t^k) \|^2}{2} \\
\leq 2\| \bar{\nabla}_x F_{t,w}(x_t^k, y_t^k) - \nabla \Phi_{t,w}(x_t^k) \|^2 + 2\| \bar{\nabla}_x F_{t,w}(x_t^k, y_t^k) - \nabla \Phi_{t,w}(x_t^k) \|^2 \\
+ \frac{\| \bar{\nabla}_x F_{t,w}(x_t^k, y_t^k) \|^2}{2}.
\]

Since \( \nabla \Phi_{t,w}(x_t^k) = \nabla \Phi_{t,w}(x_t^k, y_t^k(x_t^k)) \), we have

\[
\| \nabla \Phi_{t,w}(x_t^k, y_t^k) - \nabla \Phi_{t,w}(x_t^k) \|^2 \leq \ell \| y_t^k - y_{t,w}(x_t^k) \|^2.
\]

Putting Equations (22) to (24) together, we obtain

\[
\Phi_{t,w}(x_t^{k+1}) \leq \Phi_{t,w}(x_t^k) - \left( \frac{\eta x}{2} - \eta_x^2 \kappa \ell \right) \| \bar{\nabla}_x F_{t,w}(x_t^k, y_t^k) \|^2 \\
+ \eta x^2 \| y_t^k - y_{t,w}(x_t^k) \|^2 + \| \bar{\nabla}_x F_{t,w}(x_t^k, y_t^k) - \nabla \Phi_{t,w}(x_t^k, y_t^k) \|^2.
\]
The next lemma characterizes the descent property of distance to the maximizer $y^*_{t,w}$.

**Lemma 19**  Let $\delta_{t,w}^k = \|y^*_{t,w}(x^k_t) - y^k_t\|^2$, the following statement holds true,

$$
\delta_{t,w}^k \leq \left(1 - \frac{1}{4K}\right)\delta_{t,w}^{k-1} + 8K^3\eta^2_k\|\nabla_x F_{t,w}(x^{k-1}_t, y^{k-1}_t)\|^2 + \frac{2K}{\ell^2}\left\|\nabla_x F_{t,w}(x^{k-1}_t, y^{k-1}_t) - \tilde{\nabla}_x F_{t,w}(x^{k-1}_t, y^{k-1}_t)\right\|^2.
$$

**Proof** Since $f(x, \cdot)$ is $\mu$-strongly concave and $\eta = 1/\ell$, we have

$$
\|y^*_{t,w}(x^{k-1}_t) - y^k_t\|^2
\leq \left\|y^*_{t,w}(x^{k-1}_t) - \mathcal{P}_y(y^{k-1}_t + \eta y\nabla_x F_{t,w}(x^{k-1}_t, y^{k-1}_t))\right\|^2
+ \|y^*_{t,w}(x^{k-1}_t) - \mathcal{P}_y(y^{k-1}_t + \eta y\nabla_x F_{t,w}(x^{k-1}_t, y^{k-1}_t))\|^2
+ \left(1 + \frac{1}{2(\kappa - 1)}\right)\|y^*_{t,w}(x^{k-1}_t) - \mathcal{P}_y(y^{k-1}_t + \eta y\nabla_x F_{t,w}(x^{k-1}_t, y^{k-1}_t))\|^2
+ (1 + 2(\kappa - 1))\|\mathcal{P}_y(y^{k-1}_t + \eta y\nabla_x F_{t,w}(x^{k-1}_t, y^{k-1}_t)) - \mathcal{P}_y(y^{k-1}_t + \eta y\nabla_x F_{t,w}(x^{k-1}_t, y^{k-1}_t))\|^2
\leq (1 - \frac{1}{2\kappa})\delta_{t,w}^{k-1} + \frac{2K}{\ell^2}\left\|\nabla_x F_{t,w}(x^{k-1}_t, y^{k-1}_t) - \tilde{\nabla}_x F_{t,w}(x^{k-1}_t, y^{k-1}_t)\right\|^2.
$$

By Young’s inequality, we have

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

Since $y^*_{t,w}(\cdot)$ is $\kappa$-Lipschitz, we have

$$
\|y^*_{t,w}(x^k_t) - y^*_{t,w}(x^{k-1}_t)\|^2 \leq 2\kappa^2\|x^k_t - x^{k-1}_t\|^2 = 2\kappa^2\eta^2_k\|\nabla_x F_{t,w}(x^{k-1}_t, y^{k-1}_t)\|^2
$$
Thus, plug into
\[
\delta_{t,w}^k \leq \left(1 - \frac{1}{4\kappa}\right)\delta_{t,w}^{k-1} + 8\kappa^3\eta_k^2\|\nabla_x F_{t,w}(x_t^{k-1}, y_t^{k-1})\|^2 \\
+ \frac{2\kappa}{\ell^2} \left\| \nabla_x F_{t,w}(x_t^{k-1}, y_t^{k-1}) - \nabla_x F_{t,w}(x_t^{k-1}, y_t^{k-1}) \right\|^2.
\]

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

**Lemma 20** For any \(t, k \geq 0\), the following statement holds true,
\[
\|y_t^{k+1} - y_t^k\|^2 \leq (4 - \frac{1}{\kappa})\delta_{t,w}^k + \frac{4\kappa}{\ell^2} \left\| \nabla_x F_{t,w}(x_t^k, y_t^k) - \nabla_x F_{t,w}(x_t^k, y_t^k) \right\|^2.
\]

**Proof** By Young’s inequality, we have
\[
\|y_t^{k+1} - y_t^k\|^2 \leq 2\|y_t^{k+1} - y_{t,w}^k(x_t^k)\|^2 + 2\|y_{t,w}^k(x_t^k) - y_t^k\|^2 \\
\leq (2(1 - \frac{1}{2\kappa}) + 2)\delta_{t,w}^k + \frac{4\kappa}{\ell^2} \left\| \nabla_x F_{t,w}(x_t^k, y_t^k) - \nabla_x F_{t,w}(x_t^k, y_t^k) \right\|^2 \\
\leq (4 - \frac{1}{\kappa})\delta_{t,w}^k + \frac{4\kappa}{\ell^2} \left\| \nabla_x F_{t,w}(x_t^k, y_t^k) - \nabla_x F_{t,w}(x_t^k, y_t^k) \right\|^2,
\]
where (i) follows from Equation (25).

**C.3.2. Finite Iteration: Proof of Theorem 13**

**Proof** [Proof of Theorem 13] From Theorem 18
\[
\delta_{t,w}^k \leq \left(1 - \frac{1}{4\kappa}\right)\delta_{t,w}^{k-1} + 8\kappa^3\eta_k^2\|\nabla_x F_{t,w}(x_t^{k-1}, y_t^{k-1})\|^2 \\
+ \frac{2\kappa}{\ell^2} \left\| \nabla_x F_{t,w}(x_t^{k-1}, y_t^{k-1}) - \nabla_x F_{t,w}(x_t^{k-1}, y_t^{k-1}) \right\|^2.
\]

Denote \(\gamma = 1 - \frac{1}{4\kappa}\), Given \(\gamma^{k-1}\) we have
\[
\delta_{t,w}^k \leq \gamma^k \delta_{t,w}^0 + 8\kappa^3\eta_k^2 \left(\sum_{j=0}^{k-1} \gamma^{k-1-j} \left\| \nabla_x F_{t,w}(x_t^j, y_t^j) \right\|^2\right) \\
+ \frac{2\kappa}{\ell^2} \left(\sum_{j=0}^{k-1} \gamma^{k-1-j} \left\| \nabla_x F_{t,w}(x_t^j, y_t^j) - \nabla_x F_{t,w}(x_t^j, y_t^j) \right\|^2\right) \\
\leq \gamma^k D^2 + \frac{32\kappa^4\eta_k^2\delta^2}{3\ell^2} + \frac{2\kappa}{\ell^2} \left(\sum_{j=0}^{k-1} \gamma^{k-1-j} \left\| \nabla_x F_{t,w}(x_t^j, y_t^j) - \nabla_x F_{t,w}(x_t^j, y_t^j) \right\|^2\right),
\]
where the first term of (i) follows from that \(\gamma\) 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 \( j \in [k - 1] \),
\[
E \left\| \nabla_{x} F_{t,w}(x_{t}^{j}, y_{t}^{j}) - \nabla_{x} F_{t,w}(x_{t}^{j}, y_{t}^{j}) \right\|^{2} = \left( E \right)_{F_{t}^{j}} \left[ E \left( \left\| \nabla_{x} F_{t,w}(x_{t}^{j}, y_{t}^{j}) - \nabla_{x} F_{t,w}(x_{t}^{j}, y_{t}^{j}) \right\|^{2} | F_{t}^{j} \right) \right] = \frac{1}{w^{2}} \sum_{i=t-\ell}^{t-1} \left\| \nabla_{x} f_{i}(x_{t}^{j}, y_{t}^{j}) - \nabla_{x} f_{i}(x_{t}^{j}, y_{t}^{j}) \right\|^{2} = \frac{\sigma_{i}^{2}}{w^{2}},
\]
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 \( F_{t}^{j} \).

Thus take expectation over two sides of Equation (26), we have
\[
E \left[ \delta_{t,w}^{k} \right] \leq \gamma^{k} D^{2} + \frac{32 \kappa^{2} \eta \delta^{2}}{3 \omega^{2}} + \frac{8 \kappa^{2} \sigma^{2}}{\ell^{2} \omega^{3}}.
\]
Then by Theorem 19
\[
\Phi_{t,w} \left( x_{t}^{k} \right) - \Phi_{t,w} \left( x_{t}^{k+1} \right) \geq \left( \frac{\eta x}{2} - \eta x \kappa \ell \right) \left\| \nabla_{x} F_{t,w} \left( x_{t}^{k}, y_{t}^{k} \right) \right\|^{2} - \eta x \delta_{t,w}^{k} - \left\| \nabla_{x} F_{t,w} \left( x_{t}^{k}, y_{t}^{k} \right) - \nabla_{x} F_{t,w} \left( x_{t}^{k}, y_{t}^{k} \right) \right\|^{2} \geq \frac{15 \eta x \delta_{t,w}^{k}}{32} + \frac{\sigma_{i}^{2}}{w^{2}}. \tag{29}
\]
By Theorem 20
\[
\frac{15}{4} \kappa^{2} \ell^{2} \eta x \delta_{t,w}^{k} + \frac{15}{4} \left\| \nabla_{x} F_{t,w}(x_{t}^{k}, y_{t}^{k}) - \nabla_{x} F_{t,w}(x_{t}^{k}, y_{t}^{k}) \right\|^{2} \geq \frac{15 \eta x}{32} \times 2 \kappa^{2} \ell^{2} \left\| y_{t}^{k+1} - y_{t}^{k} \right\|^{2}. \tag{30}
\]
Sum Equation (29) and Equation (30), we have
\[
\Phi_{t,w} \left( x_{t}^{k} \right) - \Phi_{t,w} \left( x_{t}^{k+1} \right) \geq \frac{15 \eta x}{32} \times \left( 2 \kappa^{2} \ell^{2} \left\| y_{t}^{k+1} - y_{t}^{k} \right\|^{2} + \left\| \nabla_{x} F_{t,w} \left( x_{t}^{k}, y_{t}^{k} \right) \right\|^{2} \right) - \eta x \delta_{t,w}^{k} - \left\| \nabla_{x} F_{t,w} \left( x_{t}^{k}, y_{t}^{k} \right) - \nabla_{x} F_{t,w} \left( x_{t}^{k}, y_{t}^{k} \right) \right\|^{2}
\]
Rearranging the term, we have
\[ \Phi_{t,w} \left( x_t^k \right) - \Phi_{t,w} \left( x_t^{k+1} \right) \]
\[ \geq \frac{15\eta_k}{32} \left( 2\kappa^2\ell^2 \| y_t^{k+1} - y_t^k \|^2 + \| \nabla_x F_{t,w} \left( x_t^k, y_t^k \right) \|^2 \right) \]
\[ - 5\kappa^2\ell^2 \eta_k \delta_t,w \left( \frac{15}{4\ell} + 1 \right) \| \nabla_x F_{t,w} \left( x_t^k, y_t^k \right) - \nabla_x F_{t,w} \left( x_t^k, y_t^k \right) \|^2 \]
Equation (31)
\[ \Phi_{t,w} \left( x_t^k \right) - \Phi_{t,w} \left( x_t^{k+1} \right) \geq \frac{5\eta_k}{32w^2} - \frac{5\kappa^2\ell^2}{\eta_k} \left( \frac{35\kappa^4\eta_k^2 \delta^2}{3w^2} + \frac{9\kappa^2 \sigma^2}{\ell^2 w^3} \right) - \left( \frac{15}{4\ell} + 1 \right) \frac{\sigma^2}{w^3} \]
Because \( \gamma = 1 - \frac{1}{\ell} \leq 1 \), there exist a constant \( \tilde{K} \) such that \( \gamma \tilde{K} D^2 \leq \max \left\{ \frac{32\kappa^4\eta_k^2}{3w^2}, \frac{8\kappa^2 \sigma^2}{\ell^2 w^3} \right\} \). Thus for \( k \geq \tilde{K} \), we have
\[ \Phi_{t,w} \left( x_t^k \right) - \Phi_{t,w} \left( x_t^{k+1} \right) \geq \frac{25\eta_k \delta^2}{256w^2} - \frac{45\kappa^4 \eta_k \sigma^2}{w^3} - \left( \frac{15}{4\ell} + 1 \right) \frac{\sigma^2}{w^3} \]
when \( \delta^2 > \frac{2304\kappa^4 \sigma^2}{4w^2} + \frac{256(4\ell+1) \sigma^2}{256w^2} \), we set \( \alpha = \frac{25\eta_k \delta^2}{256w^2} - \frac{45\kappa^4 \eta_k \sigma^2}{w^3} - \left( \frac{15}{4\ell} + 1 \right) \frac{\sigma^2}{w^3} > 0 \). Then for \( K \geq \tilde{K} \), we have
\[ 2M \geq \mathbb{E} \left[ \Phi_{t,w} \left( x_t^K \right) - \Phi_{t,w} \left( x_t^{K+1} \right) \right] \]
\[ = \mathbb{E} \left[ \sum_{k=K}^{K} \left( \Phi_{t,w} \left( x_t^k \right) - \Phi_{t,w} \left( x_t^{k+1} \right) \right) \right] \]
\[ = \sum_{k=K}^{K} \left( \mathbb{E} \left[ \Phi_{t,w} \left( x_t^k \right) - \Phi_{t,w} \left( x_t^{k+1} \right) \left| \tau_t \geq k + 1 \right. \right] \mathbb{P} \left( \tau_t \geq k + 1 \right) + 0 \cdot \mathbb{P} \left( \tau_t < k + 1 \right) \right) \]
\[ \geq \alpha \sum_{k=K}^{K} \mathbb{P} \left( \tau_t \geq k + 1 \right) \]
\[ \geq \alpha \sum_{k=K}^{K} \mathbb{P} \left( \tau_t > K \right) = \alpha \left( K - \tilde{K} \right) \mathbb{P} \left( \tau_t > K \right), \]
where the third equation follows from the Optional Stopping Theorem. Consequently, we have \( \tau_t \) 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. \( \blacksquare \)
C.3.3. LOCAL REGRET: PROOF OF THEOREM 14

Proof [Proof of Theorem 14] Following from Equation (14), we have

\[ \| \nabla \Theta_{t,w}(x_t) \|^2 = \| \nabla_x F_{t,w}(x_t, y^*_t(w)(x_t)) \|^2 \]
\[ \leq 3 \| \nabla \Theta_{t-1,w}(x_t) \|^2 + \frac{3 \kappa^2}{(w-1)^2} \| \nabla_y f_t(x_t, y^*_t(w)(x_t)) - \nabla_y f_{t-w}(x_t, y^*_{t-1,w}(x_t)) \|^2 \]
\[ + \frac{3}{w} \| \nabla_x f_t(x_t, y^*_t(w)(x_t)) - \nabla_x f_{t-w}(x_t, y^*_{t-1,w}(x_t)) \|^2. \]

For the first term

\[ \| \nabla \Theta_{t-1,w}(x_t) \|^2 \]
\[ = \| \nabla \Theta_{t-1,w}(x^{t-1}_t) \|^2 \]
\[ \leq 3 \| \nabla \Theta_{t-1,w}(x^{t-1}_t) - \nabla_x F_{t-1,w}(x^{t-1}_t, y^{t-1}_t) \|^2 \]
\[ + 3 \| \nabla_x F_{t-1,w}(x^{t-1}_t, y^{t-1}_t) - \nabla \Theta_{t-1,w}(x^{t-1}_t, y^{t-1}_t) \|^2 \]
\[ + 3 \| \nabla \Theta_{t-1,w}(x^{t-1}_t, y^{t-1}_t) \|^2. \]

Consider \( \| y^*_{t-1}(x^{t-1}_t) - y^{t-1}_t \|^2 \)
\[ \| y^*_{t-1}(x^{t-1}_t) - y^{t-1}_t \|^2 \]
\[ \leq \kappa^2 \cdot \frac{1}{\eta \bar{\nu}^2} \| y^{t-1}_t - \mathcal{P}_Y(y_t + \eta y \nabla_y F_{t-1,w}(x^{t-1}_t, y^{t-1}_t)) \|^2 \]
\[ \leq 2 \kappa^2 \cdot \frac{1}{\eta \bar{\nu}^2} \| y^{t-1}_t - \mathcal{P}_Y(y_t + \eta y \nabla_y F_{t-1,w}(x^{t-1}_t, y^{t-1}_t)) \|^2 \]
\[ + 2 \kappa^2 \cdot \frac{1}{\eta \bar{\nu}^2} \| \mathcal{P}_Y(y_t + \eta y \nabla_y F_{t-1,w}(x^{t-1}_t, y^{t-1}_t)) - \mathcal{P}_Y(y_t + \eta y \nabla_y F_{t-1,w}(x^{t-1}_t, y^{t-1}_t)) \|^2 \]
\[ \leq 2 \kappa^2 \cdot \frac{1}{\eta \bar{\nu}^2} \| y^{t-1}_t - \mathcal{P}_Y(y_t + \eta y \nabla_y F_{t-1,w}(x^{t-1}_t, y^{t-1}_t)) \|^2 \]
\[ + 2 \kappa^2 \cdot \frac{1}{\bar{\nu}^2} \| \nabla_y F_{t-1,w}(x^{t-1}_t, y^{t-1}_t) - \nabla_y F_{t-1,w}(x^{t-1}_t, y^{t-1}_t) \|^2, \]

where (i) follows from the global error bound condition in [9] and (ii) follows from the project operator is a contraction.
Then

\[
\| \nabla \Phi_{t-1,w}(x_t) \|^2 \leq 6C \cdot \frac{1}{\eta y} \| y_{t-1} - P(y_t + \eta y \nabla_y F_{t-1,w}(x_{t-1}, y_{t-1})) \|^2 + 3 \| \nabla_x F_{t-1,w}(x_{t-1}, y_{t-1}) \|^2 \\
+ 6C \| \nabla_y F_{t-1,w}(x_{t-1}, y_{t-1}) - \nabla_y F_{t-1,w}(x_{t-1}, y_{t-1}) \|^2 \\
+ 3 \| \nabla_x F_{t-1,w}(x_{t-1}, y_{t-1}) - \nabla_x F_{t-1,w}(x_{t-1}, y_{t-1}) \|^2 \\
\leq \frac{\delta^2}{w^2} + 6C \| \nabla_y F_{t-1,w}(x_{t-1}, y_{t-1}) - \nabla_y F_{t-1,w}(x_{t-1}, y_{t-1}) \|^2 \\
+ 3 \| \nabla_x F_{t-1,w}(x_{t-1}, y_{t-1}) - \nabla_x F_{t-1,w}(x_{t-1}, y_{t-1}) \|^2,
\]

where (i) follows from the stopping condition of inner loop and \( \eta y = 1/\ell \).

Plug Equation (33) into Equation (32) and sum over \( t \), we have

\[
R_w(T) = \sum_{t=1}^{T} \| \nabla \Phi_{t,w}(x_t) \|^2 \\
\leq \sum_{t=1}^{T} \left\{ \frac{3\delta^2}{w^2} + 18C \| \nabla_y F_{t-1,w}(x_{t-1}, y_{t-1}) - \nabla_y F_{t-1,w}(x_{t-1}, y_{t-1}) \|^2 \\
+ 9 \| \nabla_x F_{t-1,w}(x_{t-1}, y_{t-1}) - \nabla_x F_{t-1,w}(x_{t-1}, y_{t-1}) \|^2 \\
+ \frac{3\kappa^2}{(w-1)^2} \| \nabla_y f_t(x_{t-1}, y_{t-1}, x_{t-1}) - \nabla_y f_t(x_{t-1}, y_{t-1}, x_{t-1}) \|^2 \\
+ \frac{3}{w^2} \| \nabla_x f_t(x_{t-1}, y_{t-1}, x_{t-1}) - \nabla_x f_t(x_{t-1}, y_{t-1}, x_{t-1}) \|^2 \right\} \\
= \frac{3T\delta^2}{w^2} + \frac{3\kappa^2}{(w-1)^2} V_{y,w}[T] + \frac{3}{w^2} V_{x,w}[T] \\
+ \sum_{t=1}^{T} \left\{ 18C \| \nabla_y F_{t-1,w}(x_{t-1}, y_{t-1}) - \nabla_y F_{t-1,w}(x_{t-1}, y_{t-1}) \|^2 \\
+ 9 \| \nabla_x F_{t-1,w}(x_{t-1}, y_{t-1}) - \nabla_x F_{t-1,w}(x_{t-1}, y_{t-1}) \|^2 \right\}.
\]
Notice that for any $t \in [T]$, 
\[
\mathbb{E}\left\| \nabla_y F_{t-1,w}(x_{t-1}^{T-1}, y_{t-1}^{T-1}) - \nabla_y F_{t-1,w}(x_{t-1}^{T-1}, y_{t-1}^{T-1}) \right\|^2
\]
\[
\overset{(i)}{=} \mathbb{E}\sum_{t=1}^{T-1} \mathbb{E}\left[ \left\| \nabla_y f_t(x_{t-1}^{T-1}, y_{t-1}^{T-1}) - \nabla_y f_t(x_{t-1}^{T-1}, y_{t-1}^{T-1}) \right\|^2 \mathbb{F}_{t-1} \right]
\]
\[
\overset{(ii)}{=} \mathbb{E}\sum_{t=1}^{T-1} \mathbb{E}\left[ \left\| \nabla_y f_t(x_{t-1}^{T-1}, y_{t-1}^{T-1}) - \nabla_y f_t(x_{t-1}^{T-1}, y_{t-1}^{T-1}) \right\|^2 \mathbb{F}_{t-1} \right]
\]
\[
\overset{(iii)}{=} \mathbb{E}\mathbb{E}\sum_{t=1}^{T-1} \left[ \frac{1}{w^2} \cdot \frac{w \cdot \sigma^2}{w^2} \right] = \frac{\sigma^2}{w^3},
\]
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
\[
\mathbb{E}\left\| \nabla_x F_{t-1,w}(x_{t-1}^{T-1}, y_{t-1}^{T-1}) - \nabla_x F_{t-1,w}(x_{t-1}^{T-1}, y_{t-1}^{T-1}) \right\|^2 = \frac{\sigma^2}{w^3}.
\]
Plug Equations (35) and (36) into Equation (34), we have
\[
\mathbb{E}[\mathcal{R}_w(T)] = \sum_{t=1}^{T} \mathbb{E}\left[ \|\nabla \Phi_{t,w}(x_t)\|^2 \right]
\]
\[
\leq 3T\delta^2 + \frac{3\kappa^2}{(w - 1)^2} \mathbb{V}_{y,w}[T] + \frac{3}{w^2} \mathbb{V}_{x,w}[T] + \frac{(18\kappa^2 + 9)T\sigma^2}{w^3}.
\]

C.3.4. Iteration and SFO Calls Bound: Proof of Theorem 16

**Proof** [Proof of Theorem 16] From Theorem 18
\[
\delta_{t,w}^k \leq \left( 1 - \frac{1}{4\kappa} \right) \delta_{t,w}^{k-1} + 8\kappa^3 \eta_x^2 \left\| \nabla_x F_{t,w}(x_t^{k-1}, y_t^{k-1}) \right\|^2
\]
\[
+ \frac{2\kappa}{\ell^2} \left\| \nabla_x F_{t,w}(x_t^{k-1}, y_t^{k-1}) - \nabla_x F_{t,w}(x_t^{k-1}, y_t^{k-1}) \right\|^2
\]
\[
+ \frac{2\kappa}{\ell^2} \left( \sum_{j=0}^{k-1} \gamma^{k-1-j} \left\| \nabla_x F_{t,w}(x_t^{j}, y_t^{j}) - \nabla_x F_{t,w}(x_t^{j}, y_t^{j}) \right\|^2 \right)
\]
\[
\overset{(37)}{=} \left( 1 - \frac{1}{4\kappa} \right) \delta_{t,w}^{k-1} + 8\kappa^3 \eta_x^2 \left\| \nabla_x F_{t,w}(x_t^{k-1}, y_t^{k-1}) \right\|^2
\]
\[
+ \frac{2\kappa}{\ell^2} \left( \sum_{j=0}^{k-1} \gamma^{k-1-j} \left\| \nabla_x F_{t,w}(x_t^{j}, y_t^{j}) - \nabla_x F_{t,w}(x_t^{j}, y_t^{j}) \right\|^2 \right)
\]

30
Then by Theorem 19
\[
\Phi_{t,w}(x_t^{k+1}) \leq \Phi_{t,w}(x_t^k) - \left(\frac{\eta}{2} - \eta^2 \kappa \ell\right) \left\| \nabla x F_{t,w}(x_t^k, y_t^k) \right\|^2 + \eta \kappa \ell^2 \delta_{t,w}^0 + \left\| \nabla x F_{t,w}(x_t^k, y_t^k) - \nabla x F_{t,w}(x_t^{k-1}, y_t^{k-1}) \right\|^2 \tag{38}
\]

Then plugging Equation (37) into Equation (38) and summing up them over \( k = 0, \ldots, \tau_t - 1 \), we have
\[
\Phi_{t,w}(x_t)^\tau \leq \Phi_{t,w}(x_t^0) - \left(\frac{\eta}{2} - \eta^2 \kappa \ell - 32 \kappa^4 \eta^2 \ell^2\right) \sum_{k=0}^{\tau_t-1} \left\| \nabla x F_{t,w}(x_t^k, y_t^k) \right\|^2 + 4\kappa \eta \ell^2 \delta_{t,w}^0 + \left(8 \eta^2 + 1\right) \sum_{k=0}^{\tau_t-1} \left\| \nabla x F_{t,w}(x_t^k, y_t^k) - \nabla x F_{t,w}(x_t^{k-1}, y_t^{k-1}) \right\|^2
\]

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
\[
\left(\frac{\eta}{2} - \eta^2 \kappa \ell - 32 \kappa^4 \eta^2 \ell^2\right) \sum_{k=0}^{\tau_t-1} \left\| \nabla x F_{t,w}(x_t^k, y_t^k) \right\|^2 \leq \Phi_{t,w}(x_t) - \Phi_{t,w}(x_{t+1}) + 4\kappa \eta \ell^2 \delta_{t,w}^0 + \left(8 \eta^2 + 1\right) \sum_{k=0}^{\tau_t-1} \left\| \nabla x F_{t,w}(x_t^k, y_t^k) - \nabla x F_{t,w}(x_t^{k-1}, y_t^{k-1}) \right\|^2.
\]

By Theorem 20
\[
\left\| y_t^{k+1} - y_t^k \right\|^2 \leq \left(4 - \frac{1}{\kappa} \right) \delta_{t,w}^k + \frac{4\kappa}{\ell^2} \left\| \nabla x F_{t,w}(x_t^k, y_t^k) - \nabla x F_{t,w}(x_t^{k-1}, y_t^{k-1}) \right\|^2.
\]
Then
\[
\sum_{k=0}^{\tau-1} \| y_{t+1}^k - y_t^k \|^2 \leq \left( 16\kappa - 4 \right) \delta_{t,w}^0 + \frac{128\kappa^2 \eta_x^2 \sum_{k=0}^{\tau-1} \left\| \nabla_{x} F_{t,w} \left( x_t^k, y_t^k \right) \right\|^2}{\ell^2} \\
+ \frac{36\kappa^2}{\ell^2} \sum_{k=0}^{\tau-1} \left\| \nabla_{x} F_{t,w} \left( x_t^k, y_t^k \right) - \nabla_{x} F_{t,w} \left( x_t^k, y_t^k \right) \right\|^2
\]

Notice that \( \delta_{0,w}^0 \leq D^2 \) and for any \( t \geq 2 \)
\[
\delta_{t,w}^0 = \left\| y_t^0 - y_{t,w}^0 (x_t^0) \right\|^2 \\
\leq 2 \left\| y_{t-1}^0 - y_{t-1,w}^0 (x_{t-1}^0) \right\|^2 + 2 \left\| y_{t-1,w}^0 (x_{t-1}^0) - y_{t,w}^0 (x_{t-1}^0) \right\|^2 \\
\leq 2\kappa^2 \left\| y_{t-1,w}^0 - P_y \left( y_{t-1,w}^0 + \eta_y G_{y,t-1} \right) \right\|^2 \\
+ \frac{2}{\mu^2 (w-1)^2} \left\| \nabla_y f_t \left( x_{t-1,w}^0, y_{t,w}^0 \left( x_{t}^0 \right) \right) - \nabla_y f_{t-w} \left( x_{t-1,w}^0, y_{t-1,w}^0 \left( x_{t-1}^0 \right) \right) \right\|^2 \\
\leq \frac{\delta^2}{4\ell^2 w^2} + \frac{2}{\mu^2 (w-1)^2} \left\| \nabla_y f_t \left( x_{t-1,w}^0, y_{t,w}^0 \left( x_{t}^0 \right) \right) - \nabla_y f_{t-w} \left( x_{t-1,w}^0, y_{t-1,w}^0 \left( x_{t-1}^0 \right) \right) \right\|^2
\]

Letting \( \eta_x = \frac{1}{32\kappa^2 \tau} \), we have
\[
\sum_{k=0}^{\tau-1} \left\| \nabla_{x} F_{t,w} \left( x_t^k, y_t^k \right) \right\|^2 \leq \frac{16}{7\eta_x} \left( \Phi_{t,w} \left( x_t^k \right) - \Phi_{t,w} \left( x_{t+1}^k \right) \right) + \frac{64\kappa^2 \delta_{t,w}^0}{7} \\
+ \frac{640}{7\eta_x} \left( \sum_{k=0}^{\tau-1} \left\| \nabla_{x} F_{t,w} \left( x_t^k, y_t^k \right) - \nabla_{x} F_{t,w} \left( x_t^k, y_t^k \right) \right\|^2 \right) \tag{39}
\]
\[
\sum_{k=0}^{\tau-1} 2(\kappa \ell)^2 \left\| y_{t+1}^k - y_t^k \right\|^2 \leq \frac{32(\kappa \ell)^2}{9} \delta_{t,w}^0 + \frac{1}{8} \sum_{k=0}^{\tau-1} \left\| \nabla_{x} F_{t,w} \left( x_t^k, y_t^k \right) \right\|^2 \\
+ \frac{72\kappa^4}{9} \left( \sum_{k=0}^{\tau-1} \left\| \nabla_{x} F_{t,w} \left( x_t^k, y_t^k \right) - \nabla_{x} F_{t,w} \left( x_t^k, y_t^k \right) \right\|^2 \right) \tag{40}
\]

Therefore add Equation (39) \times \frac{9}{8} and Equation (40) \times \frac{9}{8}, we have
\[
\frac{\eta_x}{9} \sum_{k=0}^{\tau-1} \left[ \left\| \nabla_{x} F_{t,w} \left( x_t^k, y_t^k \right) \right\|^2 + 2(\kappa \ell)^2 \left\| y_{t+1}^k - y_t^k \right\|^2 \right] \\
\leq \frac{2}{7} \left( \Phi_{t,w} \left( x_t \right) - \Phi_{t,w} \left( x_{t+1} \right) \right) + \frac{43\eta_x \kappa^3 \ell^2}{9} \delta_{t,w}^0 \\
+ \frac{20\eta_x \kappa^4}{9} \left( \sum_{k=0}^{\tau-1} \left\| \nabla_{x} F_{t,w} \left( x_t^k, y_t^k \right) - \nabla_{x} F_{t,w} \left( x_t^k, y_t^k \right) \right\|^2 \right) \tag{41}
\]
Denote $\Phi_{0,w}(x) = 0$, we notice that

$$
\Phi_{T,w}(x_T) = \sum_{t=1}^{T} (\Phi_{t,w}(x_t) - \Phi_{t-1,w}(x_{t-1}))
$$

$$
= \sum_{t=1}^{T} (\Phi_{t,w}(x_t) - \Phi_{t-1,w}(x_{t-1})) + \sum_{t=2}^{T} (\Phi_{t-1,w}(x_t) - \Phi_{t-1,w}(x_{t-1}))
$$

$$
= \frac{1}{w} \sum_{t=1}^{T} \left( F_{t-1,w}(x_t, y_{t,w}^*) - F_{t-1,w}(x_t, y_{t-1,w}^*) \right)
$$

$$
+ \frac{1}{w} \sum_{t=1}^{T} \left( f_t(x_t, y_{t,w}^*) - f_{t-w}(x_t, y_{t-w}^*) \right) + \sum_{t=2}^{T} \left( \Phi_{t-1,w}(x_t) - \Phi_{t-1,w}(x_{t-1}) \right)
$$

$$
\leq \frac{1}{w} \sum_{t=1}^{T} \left( f_t(x_t, y_{t,w}^*) - f_{t-w}(x_t, y_{t-w}^*) \right) + \sum_{t=2}^{T} \left( \Phi_{t-1,w}(x_t) - \Phi_{t-1,w}(x_{t-1}) \right),
$$

where \((i)\) follows from that $y_{t-1,w}^*(x_t)$ is the maximizer of $F_{t-1,w}(x_t, \cdot)$.

By some algebra, we have

$$
\sum_{t=1}^{T} \Phi_{t,w}(x_t) - (\Phi_{t,w}(x_{t+1}) \leq \frac{1}{w} \sum_{t=1}^{T} \left( f_t(x_t, y_{t,w}^*) - f_{t-w}(x_t, y_{t-w}^*) \right) - \Phi_{T+1,w}(x_{T+1}).
$$

Sum Equation (41) over $t$ and take expectation, we have

$$
\left( \frac{\delta^2}{2T^2} - 20k^4 \frac{\sigma^2}{w^2} \right) \eta_x \tau
$$

$$
\leq \sum_{t=1}^{T} \eta_x T \sum_{k=0}^{\tau_1 - 1} \left[ \left\| \nabla F_{t,w}(x_t^k, y_t^k) \right\|^2 + 2(\kappa \ell)^2 \|y_t^{k+1} - y_t^k\|^2 \right]
$$

$$
- 20\eta_x k^4 \sum_{t=1}^{T} \left( \sum_{k=0}^{\tau_1 - 1} \left\| \nabla F_{t,w}(x_t^k, y_t^k) - \nabla F_{t,w}(x_t^k, y_t^k) \right\|^2 \right)
$$

$$
\leq \sum_{t=1}^{T} \left( \Phi_{t,w}(x_t) - \Phi_{t,w}(x_{t+1}) \right) + \frac{43\eta_x k^3 \ell^2}{9} \sum_{t=1}^{T} \delta_{t,w}
$$

$$
\leq \frac{2}{w} \sum_{t=1}^{T} \left( f_t(x_t, y_{t,w}^*) - f_{t-w}(x_t, y_{t-w}^*) \right) - \Phi_{T+1,w}(x_{T+1}) + \frac{43\eta_x k^3 \ell^2}{9} \left\{ \frac{(T-1)\delta^2}{4\ell^2 w^2} \right\}
$$

$$
+ \frac{2}{w} \sum_{t=2}^{T} \left\| \nabla f_t(x_{t-1,w}^{\tau_1-1}, y_{t-1,w}^*) - \nabla f_{t-w}(x_{t-1,w}^{\tau_1-1}, y_{t-1,w}(x_{t-1,w}^{\tau_1-1})) \right\|^2 + D^2
$$

$$
\leq \frac{2TM}{w} + M + \frac{43T\eta_x k^3 \ell^2}{36w^2} + \frac{86\eta_x k^3 \ell^2}{9\mu^2 (w - 1)^2} V_{y,w}[T] + \frac{43\eta_x k^3 \ell^2 D^2}{9},
$$

where the first inequality follows from Assumption 4.
Thus

$$\tau \leq \frac{1}{\eta_x} \frac{2MTw + \frac{3\delta^2 T}{64t} + \frac{3\mu^2 \ell w^2}{3} V_{y,w}[T] + w^2 M + \frac{5TD^2 w^2}{32}}{(\delta^2 - 20\kappa^4 \sigma^2)}$$

Appendix D. Experiments

In this section, we evaluate the efficiency of the proposed SODA algorithm and verify the theoretical results through numerical simulations. We consider the min-max problem of training an empirical Wasserstein robustness model (WRM) [37], which has the following form:

$$\min_{\mathbf{x}} \max_{y \in \mathcal{Y}, i \in \mathcal{N}} \mathcal{L}(\mathbf{x}, y; \mathcal{D}) \triangleq \frac{1}{N} \sum_{(x_i, y_i) \in \mathcal{D}} \left[ \ell(h_{\mathbf{x}}(y_i), z_i) - \gamma \|\xi - y_i\|^2 \right],$$

where $\ell$ is the cross-entropy loss function, $N$ is the number of training samples, $\mathbf{x}$ is the network parameter, $(\xi_i, z_i) \in \mathcal{D}$ corresponds to the $i$-th data sample and label, respectively, and $y_i$ is the adversarial sample corresponding to $\xi_i$. Denote $\{y_i\}_{i=1}^N$ as $\mathbf{y}$. We simulate the online WRM model as follows. We randomly split the given dataset into $T$ pieces $\{\mathcal{D}_t\}_{t=1}^T$, and the learner sequentially receives $\mathcal{D}_t$. At each round $t$, $f_t(\mathbf{x}, \mathbf{y}) = \mathcal{L}(\mathbf{x}, \mathbf{y}; \mathcal{D}_t)$.

D.1. Implementation Details

The real-world datasets we consider are MNIST [10] and Fashion-MNIST [41], each containing 60k samples. We choose $T = 100$ for the online setting. The network architecture mainly follows [37], which consists of three convolution blocks with filters of size $8 \times 8, 6 \times 6$ and $5 \times 5$ respectively activated by ELU function, then followed by a fully connected layer and softmax output. Furthermore, we set the adversarial perturbation $\gamma \in \{0.4, 1.3\}$, which is consistent with [37].

For the results presented in Figure 1 to investigate the effect of window size, the two-timescale learning rates are set to $\eta_x = 5 \times 10^{-3}$, $\eta_y = 0.01$ for MNIST and $\eta_x = 10^{-4}$, $\eta_y = 5 \times 10^{-3}$ for Fashion-MNIST. Moreover, considering experiments in Figure 2, which compared SODA and onlineGDmax, for onlineGDmax, the learning rates are set to $\eta_x = \eta_y = 5 \times 10^{-3}$ for MNIST and $\eta_x = \eta_y = 10^{-3}$ for Fashion-MNIST; And for SODA, the learning rates are set to $\eta_x = 10^{-3}$; $\eta_y = 5 \times 10^{-3}$ for MNIST and $\eta_x = 10^{-4}$, $\eta_y = 5 \times 10^{-3}$ for Fashion-MNIST. The window size is 10.

Metrics. Since we do not have access to the first-order oracle of $\nabla \Phi_{t,w}$ in practice, two alternative performance metrics are considered, which capture the essence of the online setting and are consistent with the definition of our local NE-regret. The first metric is the stronger notion we utilize in the stop criterion, which provides an upper bound for $\|\nabla \Phi_{t,w}(\mathbf{x}_t)\|^2$. Observing that the projected gradient of $\mathbf{y}$ does not change significantly in experiments, we only compute $\|\nabla_x \nabla \Phi_{t,w}(\mathbf{x}_t, \mathbf{y}_t)\|^2$ and report the average $R_{\text{avg}} \triangleq \frac{1}{t} \sum_{j=1}^t \|\nabla_x \nabla \Phi_{j,w}(\mathbf{x}_j, \mathbf{y}_j)\|^2$ of these at each round $t$, which serves as an approximation of $\frac{1}{t} \|w_{t-\text{NE}}(t)$. The second metric is the average accuracy, where we evaluate the test accuracy of output $(\mathbf{x}_t, \mathbf{y}_t)$ from the last round on the newly coming $\mathcal{D}_t$ and report the average from round 1 to $t$.

---

3. Note that we can choose sufficiently large $\gamma > 0$ to make the maximization part be strongly-concave.
ONLINE MIN-MAX OPTIMIZATION

Figure 1: Performance of SODA with different window size. Average regret $R_{avg}$ vs. round.

Figure 2: Comparison of SODA and onlineGDmax. Number of gradient calls vs. average accuracy.

D.2. Results

**Algorithm 3 OnlineGDmax**

**Input:** window size $w \geq 1$, stepsizes $(\eta_x, \eta_y)$, tolerance $\delta > 0$, ascent step $K$

**Initialization:** $(x_1, y_1)$

1: for $t = 1$ to $T$

2: Predict $(x_t, y_t)$. Observe the cost function $f_t : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$

3: Set $(x_{t+1}, y_{t+1}) \leftarrow (x_t, y_t)$

4: repeat

5: $x_{t+1} \leftarrow x_{t+1} - \eta_x \nabla_x F_{t,w}(x_{t+1}, y_{t+1})$

6: $k = 0$

7: while $k < K$

8: $y_{t+1} \leftarrow P_Y (y_{t+1} + \eta_y \nabla_y F_{t,w}(x_{t+1}, y_{t+1}))$

9: $k = k + 1$

10: end while

11: until $\frac{\eta_x^2}{2} \|y_{t+1} - P_Y (y_{t+1} + \eta_y \nabla_y F_{t,w}(x_{t+1}, y_{t+1}))\|^2 + \|\nabla_x F_{t,w}(x_{t+1}, y_{t+1})\|^2 \leq \frac{\delta^2}{2w^2}$

12: end for

The effect of window size $w$. In Figure 1, we plot $R_{avg}$ of SODA on MNIST and Fashion-MNIST with different $w$. It can be observed that as $w$ increases from 2 to 10, the local regret becomes smaller, which verifies the bound in Theorem 4 and justifies the usage of large window size.

Comparison of SODA and baseline algorithm. To further investigate the performance of SODA, we conduct experiments to compare it with a baseline algorithm. Note that to our best knowledge, there has been no existing formal studies on the performance of any developed algorithm for online nonconvex min-max problems. Here, we consider a baseline algorithm, which is a natural extension of the well-known offline min-max method GDmax [20] to the online framework,
named *onlineGDmax*. Specifically, onlineGDmax replaces the inner-loop procedure of SODA by the nested-loop GDmax, i.e., at each iteration in the inner loop of round $t$, onlineGDmax will firstly maximize the function by multi-step gradient ascent for $y$, which is 10 steps in our setting, then perform one-step GD for $x$. Typically, the stepsizes for GDmax are chosen to be equal, i.e. $\eta_x, \eta_y$ \[37\]. OnlineGDmax is summarized in Algorithm 3. As shown in Figure 2, to achieve similar accuracy, onlineGDmax requires significantly larger number of gradient calls than SODA, which demonstrates the efficiency of our algorithm.