UNIFORM WRAPPERS: BRIDGING CONCAVE TO QUADRATIZABLE FUNCTIONS IN ONLINE OPTIMIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

This paper presents novel contributions to the field of online optimization, particularly focusing on the adaptation of algorithms from concave optimization to more challenging classes of functions. Key contributions include the introduction of uniform wrappers, establishing a vital link between upper-quadratizable functions and algorithmic conversions. Through this framework, the paper demonstrates superior regret guarantees for various classes of up-concave functions under zeroth-order feedback. Furthermore, the paper extends zeroth-order online algorithms to bandit feedback counterparts and offline counterparts, achieving a notable improvement in regret/sample complexity compared to existing approaches.

1 INTRODUCTION

The optimization of continuous DR-submodular functions has become increasingly prominent in recent years. This form of optimization represents an important subset of non-convex optimization problems at the forefront of machine learning and statistics. These challenges have numerous real-world applications like revenue maximization, mean-field inference, and recommendation systems, among others (Bian et al., 2019; Hassani et al., 2017; Mitra et al., 2021; Djolonga & Krause, 2014; Ito & Fujimaki, 2016; Gu et al., 2023; Li et al., 2023).

A natural staring point to for DR-submodular maximization is to start from a convex optimization algorithm and adapt it to the setting of DR-submodular functions. Online Convex Optimization (OCO) is extensively utilized across various fields due to its numerous practical applications and robust theoretical underpinnings. The tools from the area of online convex optimization have been applied to many online non-concave optimization algorithms, e.g., to converge to stationary points in online non-concave optimization (Yang et al., 2018), or algorithms with approximation guarantees for DR-submodular optimization (Chen et al., 2018; Niazadeh et al., 2020; Zhang et al., 2022; Pedramfar et al., 2023).

In this paper, we focus on a large class of functions, namely the class of quadratizable functions, first introduced in (Pedramfar & Aggarwal, 2024a). Quadratizable functions includes special subclasses of non-convex/non-concave functions where the offline constrained optimization problem is NP-hard to solve but we can find an α-approximation of the optimal value in polynomial time. Indeed, it is shown that this class of online upper quadratizable optimization includes up-concave optimization (a generalization of DR-submodular and concave optimization) in the following cases: (i) monotone γ-weakly μ-strongly DR-submodular functions with curvature c over general convex sets, (ii) monotone γ-weakly DR-submodular functions over convex sets containing the origin, and (iii) non-monotone DR-submodular optimization over general convex sets.

Even though the tools from OCO have proven effective in more challenging classes, much of past work along these lines involve taking inspiration from OCO and manually designing new algorithms and analyzing them specific to each problem setting. This raises the following question

When and how can we adapt algorithms from the (simpler) setup of online convex optimization into algorithms for online optimization over more general classes of functions?

In this paper, we try to provide partial solutions to this question for adapting OCO algorithms to algorithms for online quadratizable optimization. The notion of quadratizability is built upon a generalization of the defining condition $f(\mathbf{x}) - f(\mathbf{y}) \geq \langle \nabla f(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle$ of convex functions. This similarity with convex functions is a starting point which allows us to define a class of meta-algorithm called "uniform wrappers". Uniform wrappers provide a straightforward way to convert OCO algorithms into algorithms that can handle quadratizable functions. We also develop a guideline to convert the existing proofs for regret bounds of the base algorithms in the convex setting into regret bounds of the new algorithms over the quadratizable functions.

We note that, for a specific class of algorithms, this question was partly addressed in (Pedramfar & Aggarwal, 2024a). Specifically, as we will discuss in Appendix B, their result can be formulated as a special case of ours, where they assume that the starting algorithm is a first order online algorithm with semi-bandit feedback that obtains sub-linear regret against fully adaptive adversaries. This condition is too restrictive to allow for adapting many of the ideas in OCO literature. In this paper we take a step further and can handle broader classes of algorithms, including the more challenging setting of zeroth order feedback.

As an application of our framework, we propose a variant of a bandit convex optimization algorithm that was introduced in (Saha & Tewari, 2011) as the base algorithm, namely Zeroth Order Follow the Regularized Leader (ZO-FTRL) and demonstrate how it can be converted using uniform wrappers (denoted by $\mathcal{W}$) to obtain 3 algorithms for function classes (i)-(iii) mentioned above. See Tables 1 and 2 for details. Note that ZO-FTRL and $\mathcal{W}(\text{ZO-FTRL})$ are zeroth order, but they are not bandit feedback algorithms. We also extend the results to those with bandit feedback, as well as derive sample complexity guarantees for the offline algorithm.

The main contributions in this work include:

1. We develop a general framework for converting algorithms and their regret guarantees from online convex concave optimization to online quadratizable optimization. Conversion of the algorithm could be applied to any online optimization algorithm, and the conversion of the proof is described using a general guideline.
2. Our framework obtains or matches the state of the art algorithm in all online optimization settings considered. (See Table 1) Note that our framework also recovers all known results for non-stationary DR-submodular maximization. (See Remark 6 and Table 3 in (Pedramfar & Aggarwal, 2024a))
3. Except for deterministic first order feedback and the special case of γ-weakly non-monotone functions with $\gamma < 1$, our framework obtains or matches the state of the art algorithm in all online optimization settings considered. (See Table 2)
4. We obtain superior regret guarantees for several classes of weakly DR-submodular functions under zeroth order feedback, specifically (i) monotone γ-weakly μ-strongly DR-submodular functions with curvature c over general convex sets, (ii) monotone γ-weakly DR-submodular functions over convex sets containing the origin, and (iii) non-monotone DR-submodular optimization over general convex sets. (See Table 1 and Theorem 6)
5. Those results can be extended to the bandit setting yielding improved results for bandit feedback. (See Table 1 and Theorem 7)
6. The results for zeroth order online algorithms can be specialized to offline algorithms, resulting in three new algorithms with a sample complexity of $1/\epsilon^3$ in different settings, which is significantly better than the state of art $1/\epsilon^4$. (See Table 2 and Theorem 8)

To simplify the notation and statements, we define regret for maximization problems and focus on concave maximization and DR-submodular maximization.

2 Background and Notation

For a set $\mathcal{D} \subseteq \mathbb{R}^d$, we define its *affine hull* $\operatorname{aff}(\mathcal{D})$ to be the set of $\alpha\mathbf{x} + (1-\alpha)\mathbf{y}$ for all $\mathbf{x}, \mathbf{y}$ in $\mathcal{K}$ and $\alpha \in \mathbb{R}$. The *relative interior* of $\mathcal{D}$ is defined as $\operatorname{relint}(\mathcal{D}) := \{\mathbf{x} \in \mathcal{D} \mid \exists r > 0, \mathbb{B}_r(\mathbf{x}) \cap \operatorname{aff}(\mathcal{D}) \subseteq \mathcal{D}\}$. All convex functions are continuous on any point in the relative interior of their domains. In this work, we will only focus on continuous functions. If $\mathbf{x} \in \operatorname{relint}(\mathcal{K})$ and f is convex and is

Table 1: Online up-concave maximization

F	Set	Feedback			Reference	Appx.	# of queries	$\log_T(\alpha\text{-regret})$
Monotone	$0 \in \mathcal{K}$	∇F	Full Information	stoch.	(Zhang et al., 2022) (*)	$1-e^{-\gamma}$	1	$1/2$
					(Pedramfar et al., 2024a)	$1-e^{-1}$	$T^\theta(\theta \in [0,1/2])$	$2/3-\theta/3$
					(Pedramfar & Aggarwal, 2024a) (*)	$1-e^{-\gamma}$	1	$1/2$
			Semi-bandit	stoch.	(Pedramfar et al., 2024a)	$1-e^{-1}$	-	$3/4$
					(Pedramfar & Aggarwal, 2024a) (*)	$1-e^{-\gamma}$	-	$2/3$
		F	Full Information	det.	(Pedramfar & Aggarwal, 2024a) (*)	$1-e^{-1}$	2	$1/2$
				stoch. †	Theorem 6	$1-e^{-1}$	1	$2/3$
				stoch.	(Pedramfar et al., 2024a)	$1-e^{-1}$	$T^\theta(\theta \in [0,1/4])$	$4/5-\theta/5$
					(Pedramfar & Aggarwal, 2024a) (*)	$1-e^{-\gamma}$	1	$3/4$
			Bandit	det.	(Wan et al., 2023) (*)	$1-e^{-1}$	-	$3/4$
					(Zhang et al., 2024) (*)	$1-e^{-\gamma}$	-	$4/5$
				stoch.†	Theorem 7	$1-e^{-1}$	-	$3/4$
				stoch.	(Pedramfar et al., 2024a)	$1-e^{-1}$	-	$5/6$
					(Pedramfar & Aggarwal, 2024a) (*)	$1-e^{-\gamma}$	-	$4/5$
	general	∇F	Full Information	stoch.	(Pedramfar et al., 2024a)	$1/2$	$T^\theta(\theta \in [0,1/2])$	$2/3-\theta/3$
			Semi-bandit	stoch.	(Chen et al., 2018) (*)	$\gamma^2/(1+\gamma^2)$	-	$1/2$
					(Pedramfar et al., 2024a)	$1/2$	-	$3/4$
					(Pedramfar & Aggarwal, 2024a) (*)	$\gamma^2/(1+c\gamma^2)$	-	$1/2$
		F	Full Information	det.	(Pedramfar & Aggarwal, 2024a) (*)	$\gamma^2/(1+c\gamma^2)$	2	$1/2$
				stoch.	(Pedramfar et al., 2024a)	$1/2$	$T^\theta(\theta \in [0,1/4])$	$4/5-\theta/5$
					Theorem 6	$\gamma^2/(1+c\gamma^2)$	1	$2/3$
			Bandit	stoch.	(Pedramfar et al., 2024a)	$1/2$	-	$5/6$
					(Pedramfar & Aggarwal, 2024a) (*)	$\gamma^2/(1+c\gamma^2)$	-	$3/4$
					Theorem 7	$\gamma^2/(1+c\gamma^2)$	-	$3/4$
Non-Monotone	general	∇F	Full Information	stoch.	(Pedramfar et al., 2024a)	$(1-h)/4$	$T^\theta(\theta \in [0,1/2])$	$2/3-\theta/3$
					(Zhang et al., 2024) (*)	$(1-h)/4$	1	$1/2$
					(Pedramfar & Aggarwal, 2024a) (*)	$(1-h)/4$	1	$1/2$
			Semi-bandit	stoch.	(Pedramfar et al., 2024a)	$(1-h)/4$	-	$3/4$
					(Pedramfar & Aggarwal, 2024a) (*)	$(1-h)/4$	-	$2/3$
		F	Full Information	det.	(Pedramfar & Aggarwal, 2024a) (*)	$(1-h)/4$	2	$1/2$
				stoch. †	Theorem 6	$(1-h)/4$	1	$2/3$
				stoch.	(Pedramfar et al., 2024a)	$(1-h)/4$	$T^\theta(\theta \in [0,1/4])$	$4/5-\theta/5$
					(Pedramfar & Aggarwal, 2024a) (*)	$(1-h)/4$	1	$3/4$
			Bandit	det.	(Zhang et al., 2024) (*)	$(1-h)/4$	-	$4/5$
				stoch. †	Theorem 7	$(1-h)/4$	-	$3/4$
				stoch.	(Pedramfar et al., 2024a)	$(1-h)/4$	-	$5/6$
					(Pedramfar & Aggarwal, 2024a) (*)	$(1-h)/4$	-	$4/5$

This table compares different static regret results for the online up-concave maximization. The logarithmic terms in regret are ignored. Here $h := \min_{\mathbf{z} \in \mathcal{K}} \|\mathbf{z}\|_\infty$. **Rows marked with (*) are results in the literature that are special cases of the results stated here and therefore fit within the framework described in this paper.** The rows describing results with stochastic feedback that are marked with † assume that the random query oracle is contained with a cone, as detailed in Theorem 6.

differentiable at $\mathbf{x}$, then we have $f(\mathbf{y}) - f(\mathbf{x}) \geq \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle$, for all $\mathbf{y} \in \mathcal{K}$. More generally, given $\mu \geq 0$, we say a vector $\mathbf{o} \in \mathbb{R}^d$ is a *μ-subgradient* of f at $\mathbf{x}$ if $f(\mathbf{y}) - f(\mathbf{x}) \geq \langle \mathbf{o}, \mathbf{y} - \mathbf{x} \rangle + \frac{\mu}{2}\|\mathbf{y} - \mathbf{x}\|^2$. for all $\mathbf{y} \in \mathcal{K}$. Given a convex set $\mathcal{K}$, a function $f : \mathcal{K} \to \mathbb{R}$ is μ-strongly convex if and only if it has a μ-subgradient at all points $\mathbf{x} \in \mathcal{K}$. A function $F : \mathcal{D} \to \mathbb{R}^+$ is *G-Lipschitz continuous* if for all $\mathbf{x}, \mathbf{y} \in \mathcal{D}$, $\|F(\mathbf{x}) - F(\mathbf{y})\| \leq G\|\mathbf{x} - \mathbf{y}\|$. A differentiable function $F : \mathcal{D} \to \mathbb{R}^+$ is *L-smooth* if for all $\mathbf{x}, \mathbf{y} \in \mathcal{D}$, $\|\nabla F(\mathbf{x}) - \nabla F(\mathbf{y})\| \leq L\|\mathbf{x} - \mathbf{y}\|$. Given a continuous monotone function $f : \mathcal{K} \to \mathbb{R}$, its *curvature* is defined as the smallest number $c \in [0, 1]$ such that $f(\mathbf{y}+\mathbf{z}) - f(\mathbf{y}) \geq (1-c)(f(\mathbf{x}+\mathbf{z}) - f(\mathbf{x}))$, for all $\mathbf{x}, \mathbf{y} \in \mathcal{K}$ and $\mathbf{z} \geq 0$ such that $\mathbf{x}+\mathbf{z}, \mathbf{y}+\mathbf{z} \in \mathcal{K}$. We define the curvature of a function class $\mathbf{F}$ as the supremum of the curvature of functions in $\mathbf{F}$.

We say $\tilde{\nabla} f : \mathcal{K} \to \mathbb{R}^d$ is a *μ-strongly γ-weakly up-super-gradient* of f if for all $\mathbf{x} \leq \mathbf{y}$ in $\mathcal{K}$, we have $\gamma(\langle \tilde{\nabla} f(\mathbf{y}), \mathbf{y} - \mathbf{x} \rangle + \frac{\mu}{2}\|\mathbf{y} - \mathbf{x}\|^2) \leq f(\mathbf{y}) - f(\mathbf{x}) \leq \frac{1}{\gamma}(\langle \tilde{\nabla} f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle - \frac{\mu}{2}\|\mathbf{y} - \mathbf{x}\|^2)$. Then we say f is μ-strongly γ-weakly up-concave if it is continuous and it has a μ-strongly γ-weakly up-super-gradient. When $\gamma = 1$ and the above inequality holds for all $\mathbf{x}, \mathbf{y} \in \mathcal{K}$, we say f is μ-strongly concave. A differentiable function $f : \mathcal{K} \to \mathbb{R}$ is called *continuous DR-submodular* if for all $\mathbf{x} \leq \mathbf{y}$, we have $\nabla f(\mathbf{x}) \geq \nabla f(\mathbf{y})$. More generally, we say f is *γ-weakly continuous DR-submodular* if for all $\mathbf{x} \leq \mathbf{y}$, we have $\nabla f(\mathbf{x}) \geq \gamma \nabla f(\mathbf{y})$. It follows that any γ-weakly continuous DR-submodular functions is γ-weakly up-concave.

3 PROBLEM SETUP

Online optimization problems can be formalized as a repeated game between an agent and an adversary. The game lasts for T rounds on a convex domain $\mathcal{K}$ where T and $\mathcal{K}$ are known to both players. In the t-th round, the agent chooses an action $\mathbf{x}_t$ from an action set $\mathcal{K} \subseteq \mathbb{R}^d$, then the adversary chooses a loss function $f_t \in \mathbf{F}$ and a query oracle for the function f_t. Then, for $1 \leq i \leq k_t$, the agent chooses a points $\mathbf{y}_{t,i}$ and receives the output of the query oracle.

To be more precise, an agent consists of a tuple $(\Omega^{\mathcal{A}}, \mathcal{A}^{\text{action}}, \mathcal{A}^{\text{query}})$, where $\Omega^{\mathcal{A}}$ is a probability space that captures all the randomness of $\mathcal{A}$. We assume that, before the first action, the agent samples $\omega \in \Omega$. The next element in the tuple, $\mathcal{A}^{\text{action}} = (\mathcal{A}_1^{\text{action}}, \cdots, \mathcal{A}_T^{\text{action}})$ is a sequence of functions such that $\mathcal{A}_t$ that maps the history $\Omega^{\mathcal{A}} \times \mathcal{K}^{t-1} \times \prod_{s=1}^{t-1} (\mathcal{K} \times \mathcal{O})^{k_s}$ to $\mathbf{x}_t \in \mathcal{K}$ where we use $\mathcal{O}$ to denote range of the query oracle. The last element in the tuple, $\mathcal{A}^{\text{query}}$, is the query policy. For each $1 \leq t \leq T$ and $1 \leq i \leq k_t$, $\mathcal{A}_{t,i}^{\text{query}} : \Omega^{\mathcal{A}} \times \mathcal{K}^t \times \prod_{s=1}^{t-1} (\mathcal{K} \times \mathcal{O})^{k_s} \times (\mathcal{K} \times \mathcal{O})^{i-1}$ is a function that, given previous actions and observations, either selects a point $\mathbf{y}_t^i \in \mathcal{K}$, i.e., query, or signals that the query policy at this time-step is terminated. We may drop ω as one of the inputs of the above functions when there is no ambiguity. We say the agent query function is *trivial* if $k_t = 1$ and $\mathbf{y}_{t,1} = \mathbf{x}_t$ for all $1 \leq t \leq T$. In this case, we simplify the notation and use the notation $\mathcal{A} = \mathcal{A}^{\text{action}} = (\mathcal{A}_1, \cdots, \mathcal{A}_T)$ to denote the agent action functions and assume that the domain of $\mathcal{A}_t$ is $\Omega^{\mathcal{A}} \times (\mathcal{K} \times \mathcal{O})^{t-1}$.

A query oracle is a function that provides the observation to the agent. Formally, a query oracle for a function f is a map $\mathcal{Q}$ defined on $\mathcal{K}$ such that for each $\mathbf{x} \in \mathcal{K}$, the $\mathcal{Q}(\mathbf{x})$ is a random variable taking value in the observation space $\mathcal{O}$. The query oracle is called a *stochastic value oracle* or *stochastic zeroth order oracle* if $\mathcal{O} = \mathbb{R}$ and $f(\mathbf{x}) = \mathbb{E}[\mathcal{Q}(\mathbf{x})]$. Similarly, it is called a *stochastic up-super-gradient oracle* or *stochastic first order oracle* if $\mathcal{O} = \mathbb{R}^d$ and $\mathbb{E}[\mathcal{Q}(\mathbf{x})]$ is a up-super-gradient of f at $\mathbf{x}$. In all cases, if the random variable takes a single value with probability one, we refer to it as a *deterministic* oracle. Note that, given a function, there is at most a single deterministic gradient oracle, but there may be many deterministic up-super-gradient oracles. We will use ∇ to denote the deterministic gradient oracle. We say an oracle is bounded by B if its output is always within the Euclidean ball of radius B centered at the origin. We say the agent takes *semi-bandit feedback* if the oracle is first-order and the agent query function is trivial. Similarly, it takes *bandit feedback* if the oracle is zeroth-order and the agent query function is trivial[1]. If the agent query function is non-trivial, then we say the agent requires *full-information feedback*.

An adversary Adv is a set such that each element $\mathcal{B} \in \mathrm{Adv}$, referred to as a *realized adversary*, is a sequence $(\mathcal{B}_1, \cdots, \mathcal{B}_T)$ of functions where each $\mathcal{B}_t$ maps a tuple $(\mathbf{x}_1, \cdots, \mathbf{x}_t) \in \mathcal{K}^t$ to a tuple $(f_t, \mathcal{Q}_t)$ where $f_t \in \mathbf{F}$ and $\mathcal{Q}_t$ is a query oracle for f_t. We say an adversary Adv is *oblivious* if for any realization $\mathcal{B} = (\mathcal{B}_1, \cdots, \mathcal{B}_T)$, all functions $\mathcal{B}_t$ are constant, i.e., they are independent of $(\mathbf{x}_1, \cdots, \mathbf{x}_t)$. In this case, a realized adversary may be simply represented by a sequence of functions $(f_1, \cdots, f_T) \in \mathbf{F}^T$ and a sequence of query oracles $(\mathcal{Q}_1, \cdots, \mathcal{Q}_T)$ for these functions. We say an adversary is a *weakly adaptive* adversary if each function $\mathcal{B}_t$ described above does not depend on $\mathbf{x}_t$ and therefore may be represented as a map defined on $\mathcal{K}^{t-1}$. In this work we also consider adversaries that are *fully adaptive*, i.e., adversaries with no restriction. Clearly any oblivious adversary is a weakly adaptive adversary and any weakly adaptive adversary is a fully adaptive adversary. Given a function class $\mathbf{F}$ and $i \in \{0, 1\}$, we use $\mathrm{Adv}_i^{\mathrm{f}}(\mathbf{F})$ to denote the set of all possible realized adversaries with deterministic i-th order oracles. If the oracle is instead stochastic and bounded by B, we use $\mathrm{Adv}_i^{\mathrm{f}}(\mathbf{F}, B)$ to denote such an adversary. Finally, we use $\mathrm{Adv}_i^{\mathrm{o}}(\mathbf{F})$ and $\mathrm{Adv}_i^{\mathrm{o}}(\mathbf{F}, B)$ to denote all oblivious realized adversaries with i-th order deterministic and stochastic oracles, respectively.

In order to handle different notions of regret with the same approach, for an agent $\mathcal{A}$, adversary Adv, compact set $\mathcal{U} \subseteq \mathcal{K}^T$, approximation coefficient $0 < \alpha \leq 1$ and $1 \leq a \leq b \leq T$, we define *regret* as

$$\mathcal{R}_{\alpha,\mathrm{Adv}}^{\mathcal{A}}(\mathcal{U})[a, b] := \sup_{\mathcal{B} \in \mathrm{Adv}} \mathbb{E}\left[\alpha \max_{\mathbf{u}=(\mathbf{u}_1, \cdots, \mathbf{u}_T) \in \mathcal{U}} \sum_{t=a}^{b} f_t(\mathbf{u}_t) - \sum_{t=a}^{b} f_t(\mathbf{x}_t)\right],$$

where the expectation in the definition of the regret is over the randomness of the algorithm and the query oracle. We use the notation $\mathcal{R}_{\alpha,\mathcal{B}}^{\mathcal{A}}(\mathcal{U})[a, b] := \mathcal{R}_{\alpha,\mathrm{Adv}}^{\mathcal{A}}(\mathcal{U})[a, b]$ when $\mathrm{Adv} = \{\mathcal{B}\}$ is a singleton. We may drop α when it is equal to 1. When $\alpha < 1$, we often assume that the functions are non-negative. *Static adversarial regret* or simply *adversarial regret* corresponds to $a = 1$, $b = T$ and $\mathcal{U} = \mathcal{K}_\star^T := \{(\mathbf{x}, \cdots, \mathbf{x}) \mid \mathbf{x} \in \mathcal{K}\}$. When $a = 1$, $b = T$ and $\mathcal{U}$ contains only a single element then it is referred to as the *dynamic regret* (Zinkevich, 2003; Zhang et al., 2018). *Adaptive regret*, is defined as $\max_{1 \leq a \leq b \leq T} \mathcal{R}_{\alpha,\mathrm{Adv}}^{\mathcal{A}}(\mathcal{K}_\star^T)[a, b]$ (Hazan & Seshadhri, 2009). We drop a, b and $\mathcal{U}$ when the statement is independent of their value or their value is clear from the context.

[1]This is a slight generalization of the common use of the term bandit feedback. Usually, bandit feedback refers to the case where the oracle is a *deterministic* zeroth-order oracle and the agent query function is trivial.

4 UNIFORM WRAPPERS

We next introduce a class of meta-algorithms that will be a central element of our proposed framework for adapting algorithms. At a high level, the meta-algorithms we consider wrap around the base algorithm and translate each action and feedback signal between the base algorithm and the adversary. The qualifier "uniform" highlights that the translations are one-to-one and independent of time.

Definition 1. Given a function class $\mathbf{F}$ and a family of query oracles $\mathcal{Q}$ over $\mathbf{F}$, we say a *uniform wrapper* $\mathcal{W} = (\mathcal{W}^{\text{action}}, \mathcal{W}^{\text{function}}, \mathcal{W}^{\text{query}})$ is a tuple of maps where $\mathcal{W}^{\text{action}} : \mathcal{K} \to \mathcal{K}$, $\mathcal{W}^{\text{function}} : \mathbf{F} \to \mathbf{H}$ for a function class $\mathbf{H}$ and for any $f \in \mathbf{F}$ and any query oracle $\mathcal{Q}_f \in \mathcal{Q}$, $\mathcal{W}^{\text{query}}(\mathcal{Q}_f)$ is a query oracle for $\mathcal{W}^{\text{function}}(f) \in \mathbf{H}$. Given an adversary Adv choosing functions in $\mathbf{F}$ and query oracles in $\mathcal{Q}$, we define $\mathcal{W}(\mathrm{Adv})$ to be the adversary over $\mathbf{H}$ where the selected function and query by the adversary are transformed according to $\mathcal{W}^{\text{function}}$ and $\mathcal{W}^{\text{query}}$. We say $\mathcal{W} = \mathrm{Id}$ if all the maps in $\mathcal{W}$ are identity.

In Section 7 we will discuss several examples of uniform wrappers for up-concave optimization. We drop the superscripts and use $\mathcal{W}(\mathbf{x})$, $\mathcal{W}(f)$ and $\mathcal{W}(\mathcal{Q}_f)$ to denote $\mathcal{W}^{\text{action}}(\mathbf{x})$, $\mathcal{W}^{\text{function}}(f)$ and $\mathcal{W}^{\text{query}}(\mathcal{Q}_f)$, respectively, when there is no ambiguity.

Meta-algorithm 1 details the pseudo-code for $\mathcal{W}(\mathcal{A})$ for a uniform wrapper $\mathcal{W}$ and an online optimization algorithm $\mathcal{A}$. Note that, when $\mathcal{W} = \mathrm{Id}$, the meta-algorithm also reduces to the identity meta-algorithm and we see that $\mathcal{W}(\mathcal{A}) = \mathcal{A}$. Note that in the special case where $\mathcal{A}$ is an online algorithm with semi-bandit feedback, Meta-algorithm 1 reduces to Algorithm 1 in (Pedramfar & Aggarwal, 2024a).

Meta-algorithm 1: Application of a uniform wrapper to the base algorithm - $\mathcal{W}(\mathcal{A})$

Input : horizon T, algorithm $\mathcal{A}$, uniform wrapper $\mathcal{W}$
for $t = 1, 2, \ldots, T$ **do**
  Play $\mathcal{W}^{\text{action}}(\mathbf{x}_t)$ where $\mathbf{x}_t$ is the action chosen by $\mathcal{A}^{\text{action}}$
  The adversary selects f_t and a query oracle $\mathcal{Q}_t$ for f_t
  for *i starting from 1, while $\mathcal{A}^{query}$ is not terminated for this time-step* **do**
    Let $\mathbf{y}_{t,i}$ be the query chosen by $\mathcal{A}^{\text{query}}$
    Return $\mathbf{o}_{t,i} = \mathcal{W}^{\text{query}}(\mathcal{Q}_t)(\mathbf{y}_{t,i})$ as the output of the query oracle to $\mathcal{A}^{\text{query}}$
  end
end

In this paper, we will design uniform wrappers that could convert algorithms for concave optimization into algorithms for more general class of functions that contains many DR-submodular functions. Specifically, we consider upper-quadratizable/linearizable functions which we will discuss in the following section.

5 LINEARIZABLE AND QUADRATIZABLE FUNCTIONS CLASSES

We next define an important function class significantly generalizes concavity but preserves enough structure that will enable us to obtain improved regret bounds for various problems.

Definition 2 ((Pedramfar & Aggarwal, 2024a))**.** Let $\mathcal{K} \subseteq \mathbb{R}^d$ be a convex set, $\mathbf{F}$ be a function class over $\mathcal{K}$. We say the function class $\mathbf{F}$ is *upper quadratizable* if there are maps $\mathfrak{g} : \mathbf{F} \times \mathcal{K} \to \mathbb{R}^d$ and $h : \mathcal{K} \to \mathcal{K}$ and constants $\mu \geq 0$, $0 < \alpha \leq 1$ and $\beta > 0$ such that

$$\alpha f(\mathbf{y}) - f(h(\mathbf{x})) \leq \beta \left(\langle \mathfrak{g}(f, \mathbf{x}), \mathbf{y} - \mathbf{x} \rangle - \frac{\mu}{2} \|\mathbf{y} - \mathbf{x}\|^2 \right). \tag{1}$$

As a special case, when $\mu = 0$, we say $\mathbf{F}$ is *upper linearizable*. By setting $\mathfrak{g}(f, \mathbf{x}) = \nabla f(\mathbf{x})$, $h = \mathrm{Id}_{\mathcal{K}}$ and $\alpha = \beta = 1$, we see that the notion of upper linearizability generalizes concavity and upper quadratizability generalizes strong concavity. It was shown in (Pedramfar & Aggarwal, 2024a) that several classes of DR-submodular (and up-concave) functions are upper quadratizable. (see Lemmas 1, 2 and 3) A similar notion of *lower-quadratizable/linearizable* may be similarly defined for minimization problems such as convex minimization [2].

[2] We say $\mathbf{F}$ is lower quadratizable if $\alpha f(\mathbf{y}) - f(h(\mathbf{x})) \geq \beta \left(\langle \mathfrak{g}(f, \mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{\mu}{2} \|\mathbf{y} - \mathbf{x}\|^2 \right)$. This generalizes the notion of convexity and strong convexity.

Definition 3. We say $\mathbf{F}$ is *upper quadratizable with a uniform wrapper* $\mathcal{W}$ if $\mathcal{W}(\mathbf{F})$ is defined and differentiable over $\mathcal{K}$ and, for all $f \in \mathbf{F}$, we have

$$\alpha f(\mathbf{y}) - f(\mathcal{W}(\mathbf{x})) \leq \beta \left(\langle \nabla \mathcal{W}(f)(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle - \frac{\mu}{2} \|\mathbf{y} - \mathbf{x}\|^2 \right). \tag{2}$$

Note that a uniform wrapper is not uniquely determined by h and $\mathfrak{g}$ in the definition of upper quadratizable functions as it also needs to describe transformations of query oracles. The special case with $\alpha = \beta = 1$, $\mathcal{W} = \mathrm{Id}$ reduces to the definition of (strong) concavity. In Section 7, we will construct uniform wrappers for several classes of upper quadratizable functions.

6 When Is Concave Optimization Enough?

As can be seen in Meta-algorithm 1, we may apply a uniform wrapper $\mathcal{W}$ to any online optimization algorithm $\mathcal{A}$. However, even if the original algorithm has a sublinear regret over concave functions and $\mathbf{F}$ is a function class that is upper quadratizable with $\mathcal{W}$, this does not guarantee that the resulting algorithm $\mathcal{W}(\mathcal{A})$ has a sublinear regret over $\mathbf{F}$. In this section we discuss how we might convert the proofs of the regret bound for $\mathcal{A}$ over concave functions into a proof of a similar regret bound for $\mathcal{W}(\mathcal{A})$ over $\mathbf{F}$. We will refer to algorithms $\mathcal{A}$ where the regret bounds could be be converted as *wrappable* algorithms.

The core idea for converting proof for concave optimization into proofs for upper-quadratizable optimization can be informally summarized in a few steps:

(0) Sometimes, if the algorithm $\mathcal{A}$ is the result of application of a meta-algorithm to another algorithm $\mathcal{B}$, e.g. $\mathcal{A} = \mathrm{SFTT}(\mathcal{B})$ (the meta-algorithm SFTT converts algorithms that require full-information feedback to ones that work with (semi)-bandit feedback; see Appendix J), we may need to consider the base algorithm instead. For example, in the example of SFTT, we might want to consider $\mathrm{SFTT}(\mathcal{W}(\mathcal{B}))$ instead of $\mathcal{W}(\mathrm{SFTT}(\mathcal{B})) = \mathcal{W}(\mathcal{A})$.

(1) Rewrite the parts of proof (after possibly adapting the algorithm) of the original regret bound without assuming that the function class in concave, in order to isolate the use on concavity in the proof. In this step, we hope to obtain a result that would only require a single use of an inequality of the type $f(\mathbf{y}) - f(\mathbf{x}) \leq \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle - \frac{\mu}{2}\|\mathbf{y} - \mathbf{x}\|^2$ to complete the proof for the concave case. See Theorems 1 (as an example of a family of zeroth order results) and 9 (as an example of a family of first order results) for examples of this step.

(2) Verify that the results of the previous step could be adapted to upper-quadratizable setting. See the proof of Theorems 2 and 10 for examples of this step.

In the following subsection, we discuss a version of Follow The Regularized Leader (FTRL) algorithm for concave optimization and adapt it to fit the guidelines discussed above. As another application of the guideline, we refer to Appendix B for a discussion of applying this guideline to recover some previous results in the literature, including all the results in Tables 1 and 2 that are marked with (*).

6.1 Follow The Regularized Leader

Follow The Regularized Leader is a popular online optimization algorithm. When applied to a sequence of vectors $\{\mathbf{g}_t\}_{t=1}^T$ in $\mathcal{K}$, FTRL outputs a sequence of points $\{\mathbf{x}_t\}_{t=1}^T$, where

$$\mathbf{x}_1 = \operatorname*{argmin}_{\mathbf{x} \in \mathcal{K}} \Phi(\mathbf{x}), \quad \mathbf{x}_{t+1} = \operatorname*{argmin}_{\mathbf{x} \in \mathcal{K}} \eta \sum_{s=1}^{t} \langle -\mathbf{g}_s, \mathbf{x} \rangle + \Phi(\mathbf{x}). \tag{3}$$

Here $\Phi(\mathbf{x})$ is an arbitrary regularizer and η is a parameter. In this paper, we use a self-concordant barrier of $\mathcal{K}$ as the regularizer of FTRL. Self-concordant barriers were first proposed in the convex optimization literature, with (Abernethy et al., 2008) the first use in bandit feedback setting. We refer to Appendix E for an overview of the main ideas present in FTRL, including the definition of self-concordant barrier Φ, the Minkowski set $\mathcal{K}_{\gamma,\mathbf{x}_1}$, and Σ-smoothing of function f to obtain f^{Σ}.

Here we propose a FTRL variant for zeroth-order feedback, based on (Saha & Tewari, 2011), which will be a key base algorithm for our framework. See Algorithm 2 for pseudo-code.

The following theorems demonstrate how to apply the guideline described in the beginning of Section 6 to the results of (Saha & Tewari, 2011). The first step is to analyze the proof and modify the base algorithm so that we could obtain a result that is valid for non-convex functions and would only require a single use of an inequality similar to $f(\mathbf{y}) - f(\mathbf{x}) \leq \langle \nabla f(\mathbf{x}), \mathbf{y}-\mathbf{x} \rangle$ to obtain a regret bound for concave case. By a small modification in the original algorithm, we get ZO-FTRL which differs from the original in that it is no longer a bandit algorithm. While the agent plays $\mathbf{x}_t$ it queries the oracle at $\mathbf{x}_t + \delta \Sigma_t \mathbf{v}_t \neq \mathbf{x}_t$. This modification allows us to obtain the following result.

Algorithm 2: Zeroth Order Follow The Regularized Leader - ZO-FTRL

Input : Horizon T, smoothing radius δ, learning rate η, ν-self-concordant barrier Φ

$\mathbf{x}_1 \leftarrow \operatorname{argmin}_{\mathbf{x} \in \mathcal{K}} \Phi(\mathbf{x})$

for $t = 1, 2, \dots, T$ **do**

- Play $\mathbf{x}_t$
- The adversary selects f_t and reveals a zeroth-order query oracle $\mathcal{Q}_t$ for f_t
- $\Sigma_t \leftarrow \left(\nabla^2 \Phi(\mathbf{x}_t) \right)^{-1/2}$
- Draw $\mathbf{v}_t$ uniformly from $\mathbb{S}^{d-1}$
- $y_t \leftarrow$ a sample of Q_t at $\mathbf{x}_t + \delta \Sigma_t \mathbf{v}_t$
- $\mathbf{o}_t \leftarrow \frac{d}{\delta} y_t \Sigma_t^{-1} \mathbf{v}_t$
- $\mathbf{x}_{t+1} \leftarrow \operatorname*{argmin}_{\mathbf{x} \in \mathcal{K}} \sum_{s=1}^t -\eta \langle \mathbf{o}_t, \mathbf{x} \rangle + \Phi(\mathbf{x})$

end

Theorem 1. *Let $\mathbf{F}$ be an M_1-Lipschitz M_2-smooth function class that is bounded by M_0 and let $B_0 \geq M_0$. Also let $\mathcal{B} \in \mathrm{Adv}_0^o(\mathbf{F}, B_0)$ be a realized adversary that returns $f_1, \cdots, f_T$, let $\mathbf{u}_* \in \operatorname{argmax}_{\mathbf{u} \in \mathcal{K}} \sum_{t=1}^T f_t(\mathbf{u})$ and $\hat{\mathbf{u}}_* \in \operatorname{argmin}_{\mathbf{x} \in \mathcal{K}_{\gamma, \mathbf{x}_1}} \|\mathbf{u}_* - \mathbf{x}\|$ where $\gamma = T^{-1}$. Then, when running Algorithm 2 against $\mathcal{B}$, we have*

$$\sum_{t=1}^T \mathbb{E}\left[f_t(\mathbf{u}_*) - f_t(\mathbf{x}_t) \right] - O(\delta^2 T) \leq \sum_{t=1}^T \mathbb{E}\left[f_t^{\delta \Sigma_t}(\hat{\mathbf{u}}_*) - f_t^{\delta \Sigma_t}(\mathbf{x}_t) \right],$$

$$\text{and} \qquad \sum_{t=1}^T \mathbb{E}\left[\langle \nabla f_t^{\delta \Sigma_t}(\mathbf{x}_t), \hat{\mathbf{u}}_* - \mathbf{x}_t \rangle \right] \leq O\left(\eta \delta^{-2} T + \eta^{-1} \log T \right).$$

See Appendix G for the proof. Note that if f is concave, then we use use Lemma 4 to see that the right hand side of the first inequality is bounded by the left hand side of the second inequality and obtain the regret bound for the concave case. See Appendix H for the proof.

Theorem 2. *Let $\mathbf{F}$ be an M_1-Lipschitz M_2-smooth function class over $\mathcal{K}$ that is upper-linearizable with $0 < \alpha \leq 1$, $\beta \geq 0$ and a zeroth-order uniform wrapper $\mathcal{W}$. Also assume that $\mathcal{W}^{\text{action}}$ is M_1'-Lipschitz and M_2'-smooth. If Adv is a zeroth order oblivious adversary over $\mathbf{F}$ such that for for any $f \in \mathbf{F}$ and any query oracle $\mathcal{Q}_f$ returned by Adv, $\mathcal{W}(\mathcal{Q}_f)$ is a stochastic zeroth order query oracle for $\mathcal{W}(f)$ that is bounded by B_0, then*

$$\mathcal{R}_{\alpha, \mathrm{Adv}}^{\mathcal{W}(\text{ZO-FTRL})} = O\left(\eta \delta^{-2} T + \eta^{-1} \log T + \delta^2 T \right),$$

In particular, by setting $\eta = T^{-2/3}$ and $\delta = T^{-1/6}$, we see that $\mathcal{R}_{\alpha, \mathrm{Adv}}^{\mathcal{W}(\text{ZO-FTRL})} = \tilde{O}(T^{2/3})$.

7 UNIFORM WRAPPERS FOR UP-CONCAVE OPTIMIZATION

In this section, we study three classes of up-concave functions and show that they are upper-quadratizable with appropriate uniform wrappers. By identifying appropriate uniform wrappers, Theorem 2 immediately implies $\tilde{O}(T^{2/3})$ α-regret using UNIFORMWRAPPER with ZO-FTRL as a base algorithm along with the respective uniform wrapper.

7.1 MONOTONE μ-STRONGLY γ-WEAKLY UP-CONCAVE FUNCTIONS WITH BOUNDED CURVATURE ($\mathbf{F}^{\mathrm{M}}$)

For differentiable DR-submodular functions, the following lemma is proved for the case $\gamma = 1$ in (Fazel & Sadeghi, 2023) and for the case $\mu = 0$ in (Hassani et al., 2017). The general form we use here is proved in Lemma 1 in (Pedramfar & Aggarwal, 2024a).

Lemma 1. *Let $f : [0,1]^d \to \mathbb{R}$ be a non-negative monotone μ-strongly γ-weakly up-concave function with curvature bounded by c. Then, for all $\mathbf{x}, \mathbf{y} \in [0,1]^d$, we have*

$$\frac{\gamma^2}{1+c\gamma^2} f(\mathbf{y}) - f(\mathbf{x}) \leq \frac{\gamma}{1+c\gamma^2}\big(\langle \tilde{\nabla} f(\mathbf{x}), \mathbf{y}-\mathbf{x}\rangle - \frac{\mu}{2}\|\mathbf{y}-\mathbf{x}\|^2\big),$$

where $\tilde{\nabla} f$ is an up-super-gradient for f.

Lemma 1, together with Definition 1 of uniform wrappers, immediately imply the following.

Theorem 3. *Let $\mathbf{F}^{\mathrm{M}}$ be the class of functions over $\mathcal{K}$ where every $f \in \mathbf{F}^{\mathrm{M}}$ may be extended to a non-negative differentiable monotone μ-strongly γ-weakly up-concave function with curvature bounded by c defined over $[0,1]^d$. Then $\mathbf{F}^{\mathrm{M}}$ is upper-quadratizable with uniform wrapper $\mathcal{W}^{\mathrm{M}} = \mathrm{Id}$.*

If $\mathcal{A}$ is a wrappable algorithm for online optimization with sublinear regret bound of $O(T^\beta)$ for some $\beta < 1$ over concave functions, then the above theorem shows that by directly applying $\mathcal{A}$ to monotone DR-submodular functions, we get $\frac{\gamma}{1+c\gamma^2}$-regret bound of $O(T^\beta)$. As a special case, when $\mathcal{A}$ is one of the wrappable algorithm described in Theorem 10, using the above theorem recovers Theorem 2 in (Pedramfar & Aggarwal, 2024a) which itself is a generalization of Theorem 2 in (Chen et al., 2018) and Theorem 3 in (Fazel & Sadeghi, 2023).

7.2 Monotone γ-weakly up-concave functions over convex sets containing the origin ($\mathbf{F}^{\mathrm{M0}}$)

For differentiable monotone DR-submodular functions, the following lemma is proved in (Zhang et al., 2022). The general form here is proved in Lemma 2 in (Pedramfar & Aggarwal, 2024a).

Lemma 2. *Let $f : [0,1]^d \to \mathbb{R}$ be a non-negative monotone γ-weakly up-concave differentiable function and let $F : [0,1]^d \to \mathbb{R}$ be the function defined by $F(\mathbf{x}) := \int_0^1 \frac{\gamma e^{\gamma(z-1)}}{(1-e^{-\gamma})z}(f(z*\mathbf{x}) - f(\mathbf{0}))dz$. Then F is differentiable and*

$$(1-e^{-\gamma})f(\mathbf{y}) - f(\mathbf{x}) \leq \frac{1-e^{-\gamma}}{\gamma}\langle \nabla F(\mathbf{x}), \mathbf{y}-\mathbf{x}\rangle.$$

Let the random variable $\mathcal{Z}^{\mathrm{M0}} \in [0,1]$ be defined by the law $\forall z \in [0,1]$, $\mathbb{P}(\mathcal{Z}^{\mathrm{M0}} \leq z) = \int_0^z \frac{\gamma e^{\gamma(u-1)}}{1-e^{-\gamma}} du$. Then we have $\mathbb{E}_{z\sim\mathcal{Z}^{\mathrm{M0}}}\left[z^{-1}(f(z*\mathbf{x}) - f(\mathbf{0}))\right] = F(\mathbf{x})$. Moreover, for $i \geq 1$, if f is i times differentiable then we also have $\mathbb{E}_{z\sim\mathcal{Z}^{\mathrm{M0}}}\left[z^{i-1}\nabla^i f(z*\mathbf{x})\right] = \nabla^i F(\mathbf{x})$.

Definition 4. Let $\mathcal{K} \subseteq [0,1]^d$ be a convex set containing the origin and, for any $i \geq 0$, let $\mathbf{F}_i^{\mathrm{M0}}$ be the class of functions over $\mathcal{K}$ that are $\max\{i,1\}$ times differentiable and where every $f \in \mathbf{F}_i^{\mathrm{M0}}$ may be extended to a non-negative monotone γ-weakly up-concave function defined over $[0,1]^d$. We also assume that $f(\mathbf{0}) = 0$ for all $f \in \mathbf{F}_0^{\mathrm{M0}}$. We define $\mathcal{W}_i^{\mathrm{M0}} := ((\mathcal{W}_i^{\mathrm{M0}})^{\text{action}}, (\mathcal{W}_i^{\mathrm{M0}})^{\text{function}}, (\mathcal{W}_i^{\mathrm{M0}})^{\text{query}})$ to be the uniform wrapper with (i) $(\mathcal{W}_i^{\mathrm{M0}})^{\text{action}} := \mathrm{Id}_{\mathcal{K}}$; (ii) for any $f \in \mathbf{F}_i^{\mathrm{M0}}$, $(\mathcal{W}_i^{\mathrm{M0}})^{\text{function}}(f) := \mathbf{x} \mapsto \mathbb{E}_{z\sim\mathcal{Z}^{\mathrm{M0}}}\left[z^{-1}(f(z*\mathbf{x}) - f(\mathbf{0}))\right] : \mathcal{K} \to \mathbb{R}$; and (iii) for any $f \in \mathbf{F}_i^{\mathrm{M0}}$ and any i-th order query oracle $\mathcal{Q}_f$ for f, we have $(\mathcal{W}_i^{\mathrm{M0}})^{\text{query}}(\mathcal{Q}_f)(\mathbf{x}) := z^{i-1} * Q_f(z*\mathbf{x})$, where z is sampled according to $\mathbb{P}(\mathcal{Z}^{\mathrm{M0}} \leq z)$.

Theorem 4. *For any $i \geq 0$, the function class $\mathbf{F}_i^{\mathrm{M0}}$ defined above is upper-linearizable with the uniform wrapper $\mathcal{W}_i^{\mathrm{M0}}$.*

Remark 1. The meta-algorithm $\mathcal{A} \mapsto \mathrm{OMBQ}(\mathcal{A}, \mathrm{BQM0}, \mathrm{Id})$, described in (Pedramfar & Aggarwal, 2024a), is identical to $\mathcal{A} \mapsto \mathcal{W}_1^{\mathrm{M0}}(\mathcal{A})$. In other words, the results of Theorem 3 in (Pedramfar & Aggarwal, 2024a) are about the first order uniform wrapper $\mathcal{W}_1^{\mathrm{M0}}$. Here we consider a more general case where we are not necessarily limited to first order.

7.3 Non-monotone up-concave functions over general convex sets ($\mathbf{F}^{\mathrm{NM}}$)

For differentiable monotone DR-submodular functions, the following lemma is proved in (Zhang et al., 2024). The general form we use is proven in Lemma 3 in (Pedramfar & Aggarwal, 2024a).

Lemma 3. *Let $f : [0,1]^d \to \mathbb{R}$ be a non-negative continuous up-concave differentiable function and let $\underline{\mathbf{x}} \in \mathcal{K}$. Define $F : [0,1]^d \to \mathbb{R}$ as the function $F(\mathbf{x}) := \int_0^1 \frac{2}{3z(1-\frac{z}{2})^3}(f(\frac{z}{2} * (\mathbf{x} - \underline{\mathbf{x}}) + \underline{\mathbf{x}}) - f(\underline{\mathbf{x}}))dz$, then F is differentiable and we have*

$$\frac{1 - \|\underline{\mathbf{x}}\|_\infty}{4} f(\mathbf{y}) - f\left(\frac{\mathbf{x} + \underline{\mathbf{x}}}{2}\right) \leq \frac{3}{8}\langle \nabla F(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle.$$

Let the random variable $\mathcal{Z}^{\mathrm{NM}} \in [0,1]$ be defined by the law $\forall z \in [0,1], \quad \mathbb{P}(\mathcal{Z}^{\mathrm{NM}} \leq z) = \int_0^z \frac{1}{3(1-\frac{u}{2})^3} du$. Then we have $\mathbb{E}_{z \sim \mathcal{Z}^{\mathrm{NM}}}[(\frac{z}{2})^{-1} * (f(\frac{z}{2} * (\mathbf{x} - \underline{\mathbf{x}}) + \underline{\mathbf{x}}) - f(\underline{\mathbf{x}}))] = F(\mathbf{x})$. Moreover, if $i \geq 1$ and f is i times differentiable, then $\mathbb{E}_{z \sim \mathcal{Z}^{\mathrm{NM}}}[(\frac{z}{2})^{i-1} * \nabla^i f(\frac{z}{2} * (\mathbf{x} - \underline{\mathbf{x}}) + \underline{\mathbf{x}})] = \nabla^i F(\mathbf{x})$.

Definition 5. Let $\mathcal{K} \subseteq [0,1]^d$ be a convex set and, for any $i \geq 0$, let $\mathbf{F}_i^{\mathrm{NM}}$ be the class of functions over $\mathcal{K}$ where every $f \in \mathbf{F}_i^{\mathrm{NM}}$ may be extended to a non-negative up-concave function defined over $[0,1]^d$. We also assume that $\mathbf{F}_i^{\mathrm{NM}}$ is $\max\{i, 1\}$ times differentiable for all $i \geq 0$ and, for some known constant $c \geq 0$ and all $f \in \mathbf{F}_0^{\mathrm{NM}}$, $f(\underline{\mathbf{x}}) = c$. For $i \geq 0$, we define $\mathcal{W}_i^{\mathrm{NM}} = ((\mathcal{W}_i^{\mathrm{NM}})^{\text{action}}, (\mathcal{W}_i^{\mathrm{NM}})^{\text{function}}, (\mathcal{W}_i^{\mathrm{NM}})^{\text{query}})$ to be the uniform wrapper with
(i) $(\mathcal{W}_i^{\mathrm{NM}})^{\text{action}} := \mathbf{x} \mapsto \frac{\mathbf{x} + \underline{\mathbf{x}}}{2} : \mathcal{K} \to \mathcal{K}$; (ii) for any $f \in \mathbf{F}_i^{\mathrm{NM}}$,
$(\mathcal{W}_i^{\mathrm{NM}})^{\text{function}}(f) := \mathbf{x} \mapsto \mathbb{E}_{z \sim \mathcal{Z}^{\mathrm{NM}}}[(\frac{z}{2})^{-1} * (f(\frac{z}{2} * (\mathbf{x} - \underline{\mathbf{x}}) + \underline{\mathbf{x}}) - f(\underline{\mathbf{x}}))] : \mathcal{K} \to \mathbb{R}$; and
(iii) for any $f \in \mathbf{F}_i^{\mathrm{NM}}$ and any i-th order query oracle $\mathcal{Q}_f$ for f,

$$(\mathcal{W}_i^{\mathrm{NM}})^{\text{query}}(\mathcal{Q}_f)(\mathbf{x}) := \begin{cases} \left(\frac{z}{2}\right)^{i-1} * Q_f\left(\frac{z}{2} * (\mathbf{x} - \underline{\mathbf{x}}) + \underline{\mathbf{x}}\right) & \text{if } i \geq 1 \\ \left(\frac{z}{2}\right)^{-1} * \left(Q_f\left(\frac{z}{2} * (\mathbf{x} - \underline{\mathbf{x}}) + \underline{\mathbf{x}}\right) - c\right) & \text{if } i = 0 \end{cases}$$

where z is sampled according to $\mathbb{P}(\mathcal{Z}^{\mathrm{NM}} \leq z)$.

Theorem 5. *For any $i \geq 0$, the function class $\mathbf{F}_i^{\mathrm{NM}}$ defined above is upper-linearizable with the uniform wrapper $\mathcal{W}_i^{\mathrm{NM}}$.*

Remark 2. The meta-algorithm $\mathcal{A} \mapsto \mathrm{OMBQ}(\mathcal{A}, \mathrm{BQN}, \mathbf{x} \mapsto \frac{\mathbf{x} + \underline{\mathbf{x}}}{2})$, described in (Pedramfar & Aggarwal, 2024a), is identical to $\mathcal{A} \mapsto \mathcal{W}_1^{\mathrm{NM}}(\mathcal{A})$. In other words, the results of Theorem 4 in (Pedramfar & Aggarwal, 2024a) are about the first order uniform wrapper $\mathcal{W}_1^{\mathrm{NM}}$. Here we consider a more general case where we are not necessarily limited to first order.

8 Applications

We next discuss some specific online and offline non-convex/non-concave optimization problems for which we can use our new framework to derive improved regret and sample complexity bounds respectively by applying uniform wrappers proposed in Section 7 to the zeroth order feedback OCO base algorithm ZO-FTRL (Algorithm 2). We note that we can also apply our framework to other base algorithms to recover many existing results in the literature. (See Appendix B for more details).

We start with a definition. For $\mathbf{x} \in \mathcal{K}$ and $C > 0$, we say a zeroth order query oracle $\mathcal{Q}_f$ is contained in a $(\mathbf{x}, C)$ cone if we have $|Q_f(\mathbf{z}) - f(\mathbf{x})| \leq C\|\mathbf{z} - \mathbf{x}\|$ for all $\mathbf{z} \in \mathcal{K}$. In other words, the randomness of the query oracle approaches to zero at least linearly as we approach the point $\mathbf{x}$. We use the notation $\mathrm{Adv}_0^o(\mathbf{F}, \mathrm{Cone}(\mathbf{x}, C))$ to denote the oblivious adversary over $\mathbf{F}$ with query oracles that are contained within this cone. Note that $\mathcal{Q}_f \in \mathrm{Adv}_0^o(\mathbf{F}, \mathrm{Cone}(\mathbf{x}, C))$ is equivalent to $\mathcal{W}_0^{\mathrm{NM}}(\mathcal{Q}_f)$ being bounded. See condition (iii) of Definition 5 for details. If $\mathcal{Q}_f$ does not belong to a cone as described above, we can see that the term $\left(\frac{z}{2}\right)^{-1}$ causes $\mathcal{W}_0^{\mathrm{NM}}(\mathcal{Q}_f)$ to blow up. Similarly, in the special case when $\underline{x} = \mathbf{0}$ and $f(\mathbf{0}) = 0$, it is also equivalent to $\mathcal{W}_0^{\mathrm{M0}}(\mathcal{Q}_f)$ being bounded.

We begin by showing $\tilde{O}(T^{2/3})$ α-regret bounds for online optimization problems for the three function classes discussed in Section 7 under zeroth order feedback. See Appendix I for the proof.

Theorem 6. *Let $\mathbf{F}_0^{\mathrm{M}}$, $\mathbf{F}_0^{\mathrm{M0}}$ and $\mathbf{F}_0^{\mathrm{NM}}$ denote the function classes described in Lemmas 1, 2 and 3 respectively and let α^{M}, α^{M0} and α^{NM} be the values of α. If the function classes are M_1-Lipschitz and M_2-smooth, then for any $C > 0$ and $B_0 \geq M_0 = \max_{\mathbf{x} \in \mathcal{K}} f(\mathbf{x})$, the following are $\tilde{O}(T^{2/3})$:*

$$\mathcal{R}^{\mathcal{W}_0^{\mathrm{M}}(\text{ZO-FTRL})}_{\alpha^{\mathrm{M}}, \mathrm{Adv}_0^o(\mathbf{F}_0^{\mathrm{M}}, B_0)}, \quad \mathcal{R}^{\mathcal{W}_0^{\mathrm{M0}}(\text{ZO-FTRL})}_{\alpha^{\mathrm{M0}}, \mathrm{Adv}_0^o(\mathbf{F}_0^{\mathrm{M0}}, \mathrm{Cone}(\mathbf{0}, C))}, \quad \mathcal{R}^{\mathcal{W}_0^{\mathrm{NM}}(\text{ZO-FTRL})}_{\alpha^{\mathrm{NM}}, \mathrm{Adv}_0^o(\mathbf{F}_0^{\mathrm{NM}}, \mathrm{Cone}(\underline{\mathbf{x}}, C))}.$$

Remark 3. For each function class, the SOTA for noisy zeroth order feedback achieved $\tilde{O}(T^{3/4})$ α-regret bounds while we achieve $\tilde{O}(T^{2/3})$. For the special case of exact zeroth order feedback, the SOTA is $\tilde{O}(\sqrt{T})$. All the SOTA algorithms mentioned are special cases of our framework.

We next show $\tilde{O}(T^{3/4})$ α-regret bounds for online optimization problems for the three function classes discussed in Section 7 under bandit feedback. For full information zeroth order algorithms, the query location may differ from the action taken. Here we convert them into bandit algorithms using the meta-algorithm Stochastic Full-information To Trivial query (SFTT) from (Pedramfar & Aggarwal, 2024a) (see Appendix J for details). The proof is in Appendix K.

Theorem 7. *Under the assumptions of Theorem 6, the following are $\tilde{O}(T^{3/4})$:*

$$\mathcal{R}^{\mathrm{SFTT}(\mathcal{W}_0^{\mathrm{M}}(\text{ZO-FTRL}))}_{\alpha^{\mathrm{M}},\mathrm{Adv}_0^o(\mathbf{F}_0^{\mathrm{M}},B_0)}, \quad \mathcal{R}^{\mathrm{SFTT}(\mathcal{W}_0^{\mathrm{M0}}(\text{ZO-FTRL}))}_{\alpha^{\mathrm{M0}},\mathrm{Adv}_0^o(\mathbf{F}_0^{\mathrm{M0}},\mathrm{Cone}(\mathbf{0},C))}, \quad \mathcal{R}^{\mathrm{SFTT}(\mathcal{W}_0^{\mathrm{NM}}(\text{ZO-FTRL}))}_{\alpha^{\mathrm{NM}},\mathrm{Adv}_0^o(\mathbf{F}_0^{\mathrm{NM}},\mathrm{Cone}(\underline{\mathbf{x}},C))},$$

where SFTT *is Algorithm 4 in (Pedramfar & Aggarwal, 2024a) with $L = T^{1/4}$.*

Remark 4. Note that Algorithm 3 in (Wan et al., 2023) is in fact $\mathrm{SFTT}(\mathcal{W}_0^{\mathrm{M0}}(\text{ZO-FTRL}))$. However, our analysis simplifies the proof and generalizes the result to allow for stochastic feedback.

Remark 5. For the class $\mathbf{F}^{\mathrm{NM}}$ of non-monotone up-concave functions over general convex sets, our $\tilde{O}(T^{3/4})$ bound beats the SOTA $\tilde{O}(T^{4/5})$ bounds for exact and for noisy bandit feedback. For the class $\mathbf{F}^{\mathrm{M0}}$ of monotone γ-weakly up-concave functions over convex sets containing the origin, our $\tilde{O}(T^{3/4})$ bound beats the SOTA $\tilde{O}(T^{4/5})$ bound for noisy bandit feedback and matches the bound for exact bandit feedback. For the third class $\mathbf{F}^{\mathrm{M}}$ of monotone μ-strongly γ-weakly up-concave functions with bounded curvature, our results match the SOTA. All of the SOTA algorithms mentioned here are special cases of our framework.

Conversions of online algorithms to offline are referred to online-to-batch techniques and are well-known in the literature (See (Shalev-Shwartz, 2012)). A simple approach is to simply run the online algorithm and if the actions chosen by the algorithm are $\mathbf{x}_1, \cdots, \mathbf{x}_T$, return $\mathbf{x}_t$ for $1 \leq t \leq T$ with probability $1/T$. We use OTB to denote the meta-algorithm that uses this approach to convert online algorithms to offline algorithms.

We next show that using OTB conversion (on top of $\mathcal{W}(\text{ZO-FTRL})$), we obtain $\tilde{O}(1/\epsilon^3)$ sample complexity for finding an α-approximate solution in each function class under a noisy value oracle model, beating the SOTA $\tilde{O}(1/\epsilon^4)$ sample complexity. The proof is in Appendix L

Theorem 8. *Under the assumptions of Theorem 6, the following is true.*

(i) If the stochastic query oracle is bounded by B_0, then the sample complexity of the offline algorithm $\mathrm{OTB}(\mathcal{W}_0^{\mathrm{M}}(\text{ZO-FTRL}))$ *over* $\mathbf{F}_0^{\mathrm{M}}$ *is* $\tilde{O}(\epsilon^{-3})$.

(ii) If the stochastic query oracle is contained in the cone $\mathrm{Cone}(\mathbf{0}, C)$*, then the sample complexity of the offline algorithm* $\mathrm{OTB}(\mathcal{W}_0^{\mathrm{M0}}(\text{ZO-FTRL}))$ *over* $\mathbf{F}_0^{\mathrm{M0}}$ *is* $\tilde{O}(\epsilon^{-3})$.

(iii) If the stochastic query oracle is contained in the cone $\mathrm{Cone}(\underline{\mathbf{x}}, C)$*, then the sample complexity of the offline algorithm* $\mathrm{OTB}(\mathcal{W}_0^{\mathrm{NM}}(\text{ZO-FTRL}))$ *over* $\mathbf{F}_0^{\mathrm{NM}}$ *is* $\tilde{O}(\epsilon^{-3})$.

REFERENCES

Jacob Abernethy, Elad E Hazan, and Alexander Rakhlin. Competing in the dark: An efficient algorithm for bandit linear optimization. In *21st Annual Conference on Learning Theory, COLT 2008*, pp. 263–273, 2008.

An Bian, Kfir Levy, Andreas Krause, and Joachim M Buhmann. Continuous DR-submodular maximization: Structure and algorithms. In *Advances in Neural Information Processing Systems*, 2017a.

Andrew An Bian, Baharan Mirzasoleiman, Joachim Buhmann, and Andreas Krause. Guaranteed Non-convex Optimization: Submodular Maximization over Continuous Domains. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, April 2017b.

Yatao Bian, Joachim Buhmann, and Andreas Krause. Optimal continuous DR-submodular maximization and applications to provable mean field inference. In *Proceedings of the 36th International Conference on Machine Learning*, June 2019.

Gruia Calinescu, Chandra Chekuri, Martin Pál, and Jan Vondrák. Maximizing a monotone submodular function subject to a matroid constraint. *SIAM Journal on Computing*, 40(6):1740–1766, 2011.

Lin Chen, Hamed Hassani, and Amin Karbasi. Online continuous submodular maximization. In *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*, April 2018.

Lin Chen, Mingrui Zhang, and Amin Karbasi. Projection-free bandit convex optimization. In *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, pp. 2047–2056. PMLR, 2019.

Shengminjie Chen, Donglei Du, Wenguo Yang, Dachuan Xu, and Suixiang Gao. Continuous non-monotone DR-submodular maximization with down-closed convex constraint. *arXiv preprint arXiv:2307.09616*, July 2023.

Josip Djolonga and Andreas Krause. From map to marginals: Variational inference in Bayesian submodular models. *Advances in Neural Information Processing Systems*, 2014.

Maryam Fazel and Omid Sadeghi. Fast first-order methods for monotone strongly dr-submodular maximization. In *SIAM Conference on Applied and Computational Discrete Algorithms (ACDA23)*, 2023.

Yuval Filmus and Justin Ward. A tight combinatorial algorithm for submodular maximization subject to a matroid constraint. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pp. 659–668, 2012.

Dan Garber and Ben Kretzu. New projection-free algorithms for online convex optimization with adaptive regret guarantees. In *Proceedings of Thirty Fifth Conference on Learning Theory*, pp. 2326–2359. PMLR, 2022.

Shuyang Gu, Chuangen Gao, Jun Huang, and Weili Wu. Profit maximization in social networks and non-monotone DR-submodular maximization. *Theoretical Computer Science*, 957:113847, 2023.

Hamed Hassani, Mahdi Soltanolkotabi, and Amin Karbasi. Gradient methods for submodular maximization. In *Advances in Neural Information Processing Systems*, 2017.

Hamed Hassani, Amin Karbasi, Aryan Mokhtari, and Zebang Shen. Stochastic conditional gradient++: (non)convex minimization and continuous submodular maximization. *SIAM Journal on Optimization*, 30(4):3315–3344, 2020.

Elad Hazan and C. Seshadhri. Efficient learning algorithms for changing environments. In *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML '09, pp. 393–400. Association for Computing Machinery, 2009.

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

Shinji Ito and Ryohei Fujimaki. Large-scale price optimization via network flow. *Advances in Neural Information Processing Systems*, 2016.

M. Kirszbraun. Über die zusammenziehende und lipschitzsche transformationen. *Fundamenta Mathematicae*, 22:77–108, 1934.

Duksang Lee, Nam Ho-Nguyen, and Dabeen Lee. Non-smooth, hölder-smooth, and robust submodular maximization. *arXiv preprint arXiv:2210.06061*, 2023.

Yuanyuan Li, Yuezhou Liu, Lili Su, Edmund Yeh, and Stratis Ioannidis. Experimental design networks: A paradigm for serving heterogeneous learners under networking constraints. *IEEE/ACM Transactions on Networking*, 2023.

Yucheng Liao, Yuanyu Wan, Chang Yao, and Mingli Song. Improved Projection-free Online Continuous Submodular Maximization. *arXiv preprint arXiv:2305.18442*, May 2023.

Siddharth Mitra, Moran Feldman, and Amin Karbasi. Submodular+ concave. *Advances in Neural Information Processing Systems*, 2021.

Aryan Mokhtari, Hamed Hassani, and Amin Karbasi. Stochastic conditional gradient methods: From convex minimization to submodular maximization. *The Journal of Machine Learning Research*, 21(1):4232–4280, 2020.

Loay Mualem and Moran Feldman. Resolving the approximability of offline and online non-monotone DR-submodular maximization over general convex sets. In *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, April 2023.

Yurii Nesterov and Arkadii Nemirovskii. *Interior-point polynomial algorithms in convex programming*. SIAM, 1994.

Rad Niazadeh, Tim Roughgarden, and Joshua R Wang. Optimal algorithms for continuous non-monotone submodular and DR-submodular maximization. *The Journal of Machine Learning Research*, 21(1):4937–4967, 2020.

Rad Niazadeh, Negin Golrezaei, Joshua Wang, Fransisca Susan, and Ashwinkumar Badanidiyuru. Online learning via offline greedy algorithms: Applications in market design and optimization. *Management Science*, 69(7):3797–3817, July 2023.

Mohammad Pedramfar and Vaneet Aggarwal. From linear to linearizable optimization: A novel framework with applications to stationary and non-stationary DR-submodular optimization. *arXiv preprint arXiv:2405.00065*, 2024a.

Mohammad Pedramfar and Vaneet Aggarwal. A generalized approach to online convex optimization. *arXiv preprint arXiv:2402.08621*, 2024b.

Mohammad Pedramfar, Christopher John Quinn, and Vaneet Aggarwal. A unified approach for maximizing continuous DR-submodular functions. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023.

Mohammad Pedramfar, Yididiya Y. Nadew, Christopher John Quinn, and Vaneet Aggarwal. Unified projection-free algorithms for adversarial DR-submodular optimization. In *The Twelfth International Conference on Learning Representations*, 2024a.

Mohammad Pedramfar, Christopher Quinn, and Vaneet Aggarwal. A unified approach for maximizing continuous γ-weakly DR-submodular functions. *optimization-online preprint optimization-online:25915*, 2024b.

Ankan Saha and Ambuj Tewari. Improved regret guarantees for online smooth convex optimization with bandit feedback. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pp. 636–642. JMLR Workshop and Conference Proceedings, 2011.

Shai Shalev-Shwartz. Online learning and online convex optimization. *Foundations and Trends® in Machine Learning*, 4(2):107–194, 2012.

Zongqi Wan, Jialin Zhang, Wei Chen, Xiaoming Sun, and Zhijie Zhang. Bandit multi-linear dr-submodular maximization and its applications on adversarial submodular bandits. In *International Conference on Machine Learning*, 2023.

Hassler Whitney. Analytic extensions of differentiable functions defined in closed sets. *Transactions of the American Mathematical Society*, 36(1):63–89, 1934.

Bryan Wilder. Equilibrium computation and robust optimization in zero sum games with submodular structure. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

Lin Yang, Lei Deng, Mohammad H Hajiesmaili, Cheng Tan, and Wing Shing Wong. An optimal algorithm for online non-convex learning. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 2(2):1–25, 2018.

Lijun Zhang, Shiyin Lu, and Zhi-Hua Zhou. Adaptive online learning in dynamic environments. In *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.

Mingrui Zhang, Lin Chen, Hamed Hassani, and Amin Karbasi. Online continuous submodular maximization: From full-information to bandit feedback. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

Qixin Zhang, Zengde Deng, Zaiyi Chen, Haoyuan Hu, and Yu Yang. Stochastic continuous submodular maximization: Boosting via non-oblivious function. In *Proceedings of the 39th International Conference on Machine Learning*, 2022.

Qixin Zhang, Zengde Deng, Zaiyi Chen, Kuangqi Zhou, Haoyuan Hu, and Yu Yang. Online learning for non-monotone DR-submodular maximization: From full information to bandit feedback. In *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, April 2023.

Qixin Zhang, Zongqi Wan, Zengde Deng, Zaiyi Chen, Xiaoming Sun, Jialin Zhang, and Yu Yang. Boosting gradient ascent for continuous DR-submodular maximization. *arXiv preprint arXiv:2401.08330*, 2024.

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

Table 2: Offline up-concave maximization

F	Set	Feedback		Reference	Appx.	Complexity
Monotone	$0 \in \mathcal{K}$	∇F	stoch.	(Mokhtari et al., 2020)	$1-e^{-\gamma}$	$O(1/\epsilon^3)$
				(Hassani et al., 2020) (*)	$1-e^{-\gamma}$	$O(1/\epsilon^2)$
				(Zhang et al., 2022) (*)	$1-e^{-\gamma}$	$O(1/\epsilon^2)$
				(Pedramfar & Aggarwal, 2024a) (*)	$1-e^{-\gamma}$	$O(1/\epsilon^2)$
		F	det.	(Pedramfar et al., 2024b)	$1-e^{-\gamma}$	$O(1/\epsilon^3)$
				(Pedramfar & Aggarwal, 2024a) (*)	$1-e^{-\gamma}$	$O(1/\epsilon^2)$
			stoch. †	Theorem 8	$1-e^{-\gamma}$	$\tilde{O}(1/\epsilon^3)$
			stoch.	(Pedramfar et al., 2024b)	$1-e^{-\gamma}$	$O(1/\epsilon^5)$
				(Pedramfar & Aggarwal, 2024a) (*)	$1-e^{-\gamma}$	$O(1/\epsilon^4)$
	general	∇F	stoch.	(Hassani et al., 2017) (*)	$\gamma^2/(1+\gamma^2)$	$O(1/\epsilon^2)$
				(Pedramfar et al., 2024b)	$\gamma^2/(1+\gamma^2)$	$\tilde{O}(1/\epsilon^3)$
				(Pedramfar & Aggarwal, 2024a) (*)	$\gamma^2/(1+c\gamma^2)$	$O(1/\epsilon^2)$
		F	det.	(Pedramfar et al., 2023)	$\gamma^2/(1+\gamma^2)$	$\tilde{O}(1/\epsilon^3)$
				(Pedramfar & Aggarwal, 2024a) (*)	$\gamma^2/(1+c\gamma^2)$	$O(1/\epsilon^2)$
			stoch.	(Pedramfar et al., 2023)	$\gamma^2/(1+\gamma^2)$	$\tilde{O}(1/\epsilon^5)$
				(Pedramfar & Aggarwal, 2024a) (*)	$\gamma^2/(1+c\gamma^2)$	$O(1/\epsilon^4)$
				Theorem 8	$\gamma^2/(1+c\gamma^2)$	$\tilde{O}(1/\epsilon^3)$
Non-Monotone	general	∇F	stoch.	(Pedramfar et al., 2024b)	$\frac{\gamma(1-\gamma h)}{\gamma'-1}\left(\frac{1}{2}-\frac{1}{2^{\gamma'}}\right)$	$O(1/\epsilon^3)$
				(Zhang et al., 2024) (*)	$(1-h)/4$	$O(1/\epsilon^2)$
				(Pedramfar & Aggarwal, 2024a) (*)	$(1-h)/4$	$O(1/\epsilon^2)$
		F	det.	(Pedramfar et al., 2024b)	$\frac{\gamma(1-\gamma h)}{\gamma'-1}\left(\frac{1}{2}-\frac{1}{2^{\gamma'}}\right)$	$O(1/\epsilon^3)$
				(Pedramfar & Aggarwal, 2024a) (*)	$(1-h)/4$	$O(1/\epsilon^2)$
			stoch. †	Theorem 8	$(1-h)/4$	$\tilde{O}(1/\epsilon^3)$
			stoch.	(Pedramfar et al., 2024b)	$\frac{\gamma(1-\gamma h)}{\gamma'-1}\left(\frac{1}{2}-\frac{1}{2^{\gamma'}}\right)$	$O(1/\epsilon^5)$
				(Pedramfar & Aggarwal, 2024a) (*)	$(1-h)/4$	$O(1/\epsilon^4)$

This table compares the different results for the number of oracle calls (complexity) *within the constraint set* for up-concave maximization. We refer to (Pedramfar et al., 2024b) for a more comprehensive table that includes results for deterministic first order feedback. Here $h := \min_{\mathbf{z}\in\mathcal{K}} \|\mathbf{z}\|_\infty$ and $\gamma' := \gamma + 1/\gamma$. **Rows marked with (*) are results in the literature that fit within the framework described in this paper.** The rows describing results with stochastic feedback that are marked with † assume that the random query oracle is contained with a cone, as detailed in Theorem 6.

A Additional Related Works

DR-submodular maximization Two of the main methods for continuous DR-submodular maximization are *Frank-Wolfe type methods* and *Boosting based methods*. This division is based on how the approximation coefficient appears in the proof.

In Frank-Wolfe type algorithms, the approximation coefficient appears by specific choices of the Frank-Wolfe update rules. (See Lemma 8 in (Pedramfar et al., 2024a)) The specific choices of the update rules for different settings have been proposed in (Bian et al., 2017b;a; Mualem & Feldman, 2023; Pedramfar et al., 2023; Chen et al., 2023). The momentum technique of (Mokhtari et al., 2020) has been used to convert algorithms designed for deterministic feedback to stochastic feedback setting. (Hassani et al., 2020) proposed a Frank-Wolfe variant with access to a stochastic gradient oracle with *known distribution*. Frank-Wolfe type algorithms been adapted to the online setting using Meta-Frank-Wolfe (Chen et al., 2018; 2019) or using Blackwell approachablity (Niazadeh et al., 2023). Later (Zhang et al., 2019) used a Meta-Frank-Wolfe with random permutation technique to obtain full-information results that only require a single query per function and also bandit results. This was extended to another settings by (Zhang et al., 2023) and generalized to many different settings with improved regret bounds by (Pedramfar et al., 2024a).

Some techniques construct an alternative function such that maximization of this function results in approximate maximization of the original function. Given this definition, we may consider the result of (Hassani et al., 2017; Chen et al., 2018; Fazel & Sadeghi, 2023) as the first boosting based results. However, in the case of monotone DR-submodular functions over general convex sets, the alternative function is identical to the original function. The term boosting in this context was first used in (Zhang et al., 2022), based on ideas presented in (Filmus & Ward, 2012; Mitra et al., 2021), for monotone functions over convex sets containing the origin. This idea has been used later in (Wan et al., 2023; Liao et al., 2023) in bandit and projection-free full-information settings. Finally, in (Zhang et al., 2024) a boosting based method was introduced for non-monotone functions over general convex sets.

Up-concave maximization Not all continuous DR-submodular functions are concave and not all concave functions are continuous DR-submodular. (Mitra et al., 2021) considers functions that are the sum of a concave and a continuous DR-submodular function. It is well-known that continuous DR-submodular functions are concave along positive directions (Calinescu et al., 2011; Bian et al., 2017b). Based on this idea, (Wilder, 2018) defined an up-concave function as a function that is concave along positive directions. Up-concave maximization has been considered in the offline setting before, e.g. in (Lee et al., 2023) and (Pedramfar & Aggarwal, 2024a). In this work, we focus on up-concave maximization which is a generalization of DR-submodular maximization.

B RECOVERING PREVIOUS RESULTS IN THE LITERATURE

As mentioned in Remark 4, Algorithm 3 in (Wan et al., 2023) is in fact $\operatorname{SFTT}(\mathcal{W}_0^{\mathrm{M0}}(\text{ZO-FTRL}))$ and therefore their result fits within our framework. The way the remaining results in the tables that are marked with (*) is discussed in the following.

We demonstrate how to apply the guideline described in the beginning of Section 6 to Theorem 2 in (Pedramfar & Aggarwal, 2024b). This allows us to obtain a generalized version of Theorems 1 in (Pedramfar & Aggarwal, 2024a). As we will discuss below, this will allow us to recover all the remaining results in Tables 1 and 2 that are marked with (*) and all the results of (Pedramfar & Aggarwal, 2024a). Note that the results of (Pedramfar & Aggarwal, 2024a) in non-stationary setting are not discussed in this paper, but they are also recovered.

We start with some definitions. Given a function class $\mathbf{F}$, we use the notation $\mathbf{F}_{\mu,\mathbf{g}}$ to denote the class of functions $q(\mathbf{y}) := \langle \mathfrak{g}(f,\mathbf{x}), \mathbf{y}-\mathbf{x}\rangle - \frac{\mu}{2}\|\mathbf{y}-\mathbf{x}\|^2 : \mathcal{K} \to \mathbb{R}$, for all $f \in \mathbf{F}$ and $\mathbf{x} \in \mathcal{K}$. This is the class of quadratic (or linear, when $\mu = 0$) functions that form the upper bound in Equation 1. Similarly, for any $B_1 > 0$, we use the notation $\mathbf{Q}_\mu[B_1]$ to denote the class of functions $q(\mathbf{y}) := \langle \mathbf{o}, \mathbf{y}-\mathbf{x}\rangle - \frac{\mu}{2}\|\mathbf{y}-\mathbf{x}\|^2 : \mathcal{K} \to \mathbb{R}$, for all $\mathbf{x} \in \mathcal{K}$ and $\mathbf{o} \in \mathbb{B}_{B_1}(\mathbf{0})$. In the following theorems, we will obtain results that allow us to reduce the problem of online optimization over $\mathbf{F}$ to the problem of online optimization over the quadratic (or linear) function class $\mathbf{F}_{\mu,\mathbf{g}}$.

Theorem 9. *Let $\mathcal{A}$ be algorithm for online optimization with semi-bandit feedback. Also let $\mathbf{F}$ be a differentiable function class over $\mathcal{K}$ and $\mu \geq 0$.Then the following are true.*

- *If query oracles in* Adv *are deterministic gradient oracles, then we have*

$$\sup_{\mathcal{B}\in\mathrm{Adv}} \mathbb{E}\left[\max_{\mathbf{u}\in\mathcal{U}} \sum_{t=a}^{b} \left(\langle \nabla f_t(\mathbf{x}_t), \mathbf{u}_t - \mathbf{x}_t\rangle - \frac{\mu}{2}\|\mathbf{u}_t - \mathbf{x}_t\|\right)\right] \leq \mathcal{R}^{\mathcal{A}}_{1,\mathrm{Adv}_1^f(\mathbf{F}_{\mu,\nabla})}.$$

- *On the other hand, if $\mathbf{F}$ is M_1-Lipschitz and query oracles in* Adv *are stochastic gradient oracles that are bounded by $B_1 \geq M_1$, then we have*

$$\sup_{\mathcal{B}\in\mathrm{Adv}} \max_{\mathbf{u}\in\mathcal{U}} \mathbb{E}\left[\sum_{t=a}^{b} \left(\langle \nabla f_t(\mathbf{x}_t), \mathbf{u}_t - \mathbf{x}_t\rangle - \frac{\mu}{2}\|\mathbf{u}_t - \mathbf{x}_t\|\right)\right] \leq \mathcal{R}^{\mathcal{A}}_{1,\mathrm{Adv}_1^f(\mathbf{Q}_\mu[B_1])}.$$

See Appendix C for proof. Note that if f_t are μ-strongly concave, then this result reduces to Theorem 2 in (Pedramfar & Aggarwal, 2024b). Next, we follow step (2) in the guideline to obtain the following result.

Theorem 10. *Let $\mathbf{F}$ be function class over $\mathcal{K}$ that is upper-quadratizable with $\mu \geq 0$, $0 < \alpha \leq 1$ and $\beta \geq 0$ and a first-order uniform wrapper $\mathcal{W}$.*

- *If $\mathcal{W}(\nabla) = \nabla$, i.e., it maps deterministic gradient oracles into deterministic gradient oracles, then we have $\mathcal{R}^{\mathcal{W}(\mathcal{A})}_{\alpha,\mathrm{Adv}_1^f(\mathbf{F})} \leq \beta \mathcal{R}^{\mathcal{A}}_{1,\mathrm{Adv}_1^f(\mathbf{F}_{\mu,\nabla})}$.*
- *If, for any $f \in \mathbf{F}$ and any query oracle $\mathcal{Q}_f$ bounded by B_1, $\mathcal{W}(\mathcal{Q}_f)$ is a stochastic query oracle for $\mathcal{W}(f)$ that is bounded by B_1', then we have $\mathcal{R}^{\mathcal{W}(\mathcal{A})}_{\alpha,\mathrm{Adv}_1^o(\mathbf{F},B_1)} \leq \beta \mathcal{R}^{\mathcal{A}}_{1,\mathrm{Adv}_1^f(\mathbf{Q}_\mu[B_1'])}$.*

See Appendix D for proof. In this theorem, by using the uniform wrappers described in Section 7, in the special case of $i = 1$, we recover Theorems 2, 3 and 4 in (Pedramfar & Aggarwal, 2024a). (See Remarks 1 and 2) In other words, we recover all meta-algorithms in (Pedramfar & Aggarwal, 2024a) that are used to convert concave optimization algorithms into up-concave optimization algorithms.

Remark 6. By applying these uniform wrappers to base algorithms SO-OGA ((Garber & Kretzu, 2022)) or IA ((Zhang et al., 2018)), we recover all the results of (Pedramfar & Aggarwal, 2024a). In particular, we also recover the results for non-stationary regret described in Table 3 in (Pedramfar & Aggarwal, 2024a).

C PROOF OF THEOREM 9

Proof.

Deterministic oracle:

For any realization $\mathcal{B} = (\mathcal{B}_1, \cdots, \mathcal{B}_T) \in \mathrm{Adv} \subseteq \mathrm{Adv}_1^{\mathrm{f}}(\mathbf{F})$, we define $\mathcal{B}'_t(\mathbf{x}_1, \cdots, \mathbf{x}_t)$ to be the tuple (q_t, ∇) where

$$\mathcal{B}'_t(\mathbf{x}_1, \cdots, \mathbf{x}_t) := q_t := \mathbf{y} \mapsto \langle \nabla f_t(\mathbf{x}_t), \mathbf{y} - \mathbf{x}_t \rangle - \frac{\mu}{2}\|\mathbf{y} - \mathbf{x}_t\|^2,$$

and $\mathcal{B}' = (\mathcal{B}'_1, \cdots, \mathcal{B}'_T)$. Note that each $\mathcal{B}'_t$ is a deterministic function of $\mathbf{x}_1, \cdots, \mathbf{x}_t$ and therefore $\mathcal{B}' \in \mathrm{Adv}_1^{\mathrm{f}}(\mathbf{F}_{\mu,\nabla})$. Since the algorithm uses semi-bandit feedback, the sequence of random vectors $(\mathbf{x}_1, \cdots, \mathbf{x}_T)$ chosen by $\mathcal{A}$ is identical between the game with $\mathcal{B}$ and $\mathcal{B}'$. Hence

$$\begin{aligned}
\sup_{\mathcal{B} \in \mathrm{Adv}} \mathbb{E}&\left[\max_{\mathbf{u} \in \mathcal{U}} \sum_{t=a}^{b} \left(\langle \nabla f_t(\mathbf{x}_t), \mathbf{u}_t - \mathbf{x}_t \rangle - \frac{\mu}{2}\|\mathbf{u}_t - \mathbf{x}_t\|^2 \right)\right] \\
&= \sup_{\mathcal{B} \in \mathrm{Adv}} \mathbb{E}\left[\max_{\mathbf{u} \in \mathcal{U}} \left(\sum_{t=a}^{b} q_t(\mathbf{u}_t) - \sum_{t=a}^{b} q_t(\mathbf{x}_t) \right)\right] \\
&\leq \sup_{\mathcal{B}' \in \mathrm{Adv}_1^{\mathrm{f}}(\mathbf{F}_{\mu,\nabla})} \mathcal{R}^{\mathcal{A}}_{1,\mathcal{B}'} = \mathcal{R}^{\mathcal{A}}_{1,\mathrm{Adv}_1^{\mathrm{f}}(\mathbf{F}_{\mu,\nabla})}.
\end{aligned}$$

Stochastic oracle:

Let $\Omega^{\mathcal{Q}} = \Omega_1^{\mathcal{Q}} \times \cdots \times \Omega_T^{\mathcal{Q}}$ capture all sources of randomness in the query oracles of $\mathrm{Adv}_1^{\mathrm{o}}(\mathbf{F}, B_1)$, i.e., for any choice of $\theta \in \Omega^{\mathcal{Q}}$, the query oracle is deterministic. Hence for any $\theta \in \Omega^{\mathcal{Q}}$ and realized adversary $\mathcal{B} \in \mathrm{Adv} \subseteq \mathrm{Adv}_1^{\mathrm{f}}(\mathbf{F}, B_1)$, we may consider $\mathcal{B}_\theta$ as an object similar to an adversary with a deterministic oracle. However, note that $\mathcal{B}_\theta$ does not satisfy the unbiasedness condition of the oracle, i.e., the returned value of the oracle is not necessarily the gradient of the function at that point. Recall that $\mathcal{B}_t$ maps a tuple $(\mathbf{x}_1, \cdots, \mathbf{x}_t)$ to a tuple of f_t and a stochastic query oracle for f_t. We will use $\mathbb{E}_{\Omega^{\mathcal{Q}}}$ to denote the expectation with respect to the randomness of query oracle and $\mathbb{E}_{\Omega_t^{\mathcal{Q}}}[\cdot] := \mathbb{E}_{\Omega^{\mathcal{Q}}}[\cdot | f_t, \mathbf{x}_t]$ to denote the expectation conditioned on the action of the agent and the adversary. Similarly, let $\mathbb{E}_{\Omega^{\mathcal{A}}}$ denote the expectation with respect to the randomness of the agent. Let $\mathbf{o}_t$ be the random variable denoting the output of $\mathcal{Q}$ at time-step t and let

$$\bar{\mathbf{o}}_t := \mathbb{E}[\mathbf{o}_t \mid f_t, \mathbf{x}_t] = \mathbb{E}_{\Omega_t^{\mathcal{Q}}}[\mathbf{o}_t] = \nabla f_t(\mathbf{x}_t).$$

Similar to the deterministic case, for any realization $\mathcal{B} = (\mathcal{B}_1, \cdots, \mathcal{B}_T) \in \mathrm{Adv}$ and any $\theta \in \Omega^{\mathcal{Q}}$, we define $\mathcal{B}'_{\theta,t}(\mathbf{x}_1, \cdots, \mathbf{x}_t)$ to be the pair (q_t, ∇) where

$$q_t := \mathbf{y} \mapsto \langle \mathbf{o}_t, \mathbf{y} - \mathbf{x}_t \rangle - \frac{\mu}{2}\|\mathbf{y} - \mathbf{x}_t\|^2.$$

We also define $\mathcal{B}'_\theta := (\mathcal{B}'_{\theta,1}, \cdots, \mathcal{B}'_{\theta,T})$. Note that a specific choice of θ is necessary to make sure that the function returned by $\mathcal{B}'_{\theta,t}$ is a deterministic function of $\mathbf{x}_1, \cdots, \mathbf{x}_t$ and not a random variable and therefore $\mathcal{B}'_\theta$ belongs to $\mathrm{Adv}_1^{\mathrm{f}}(\mathbf{F}_\mu[B_1])$.

Since the algorithm uses (semi-)bandit feedback, given a specific value of θ, the sequence of random vectors $(\mathbf{x}_1, \cdots, \mathbf{x}_T)$ chosen by $\mathcal{A}$ is identical between the game with $\mathcal{B}_\theta$ and $\mathcal{B}'_\theta$. Therefore, for any $\mathbf{u} \in \mathcal{U}$, we have

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=a}^{b}\left(\langle \nabla f_t(\mathbf{x}_t), \mathbf{u}_t - \mathbf{x}_t\rangle - \frac{\mu}{2}\|\mathbf{u}_t - \mathbf{x}_t\|^2\right)\right] \\
&\qquad = \mathbb{E}\left[\sum_{t=a}^{b}\left(\langle \mathbb{E}\left[\mathbf{o}_t \mid f_t, \mathbf{x}_t\right], \mathbf{u}_t - \mathbf{x}_t\rangle - \frac{\mu}{2}\|\mathbf{u}_t - \mathbf{x}_t\|^2\right)\right] \\
&\qquad = \mathbb{E}\left[\sum_{t=a}^{b}\left(\mathbb{E}\left[\langle \mathbf{o}_t, \mathbf{u}_t - \mathbf{x}_t\rangle - \frac{\mu}{2}\|\mathbf{u}_t - \mathbf{x}_t\|^2 \mid f_t, \mathbf{x}_t\right]\right)\right] \\
&\qquad = \mathbb{E}\left[\sum_{t=a}^{b}\left(\mathbb{E}\left[q_t(\mathbf{u}_t) - q_t(\mathbf{x}_t) \mid f_t, \mathbf{x}_t\right]\right)\right] \\
&\qquad = \mathbb{E}\left[\sum_{t=a}^{b}\left(q_t(\mathbf{u}_t) - q_t(\mathbf{x}_t)\right)\right].
\end{aligned}$$

Hence we have

$$\begin{aligned}
\max_{\mathbf{u}\in\mathcal{U}} \mathbb{E}\left[\sum_{t=a}^{b}\left(\langle \nabla f_t(\mathbf{x}_t), \mathbf{u}_t - \mathbf{x}_t\rangle - \frac{\mu}{2}\|\mathbf{u}_t - \mathbf{x}_t\|\right)\right] &= \max_{\mathbf{u}\in\mathcal{U}} \mathbb{E}\left[\sum_{t=a}^{b}\left(q_t(\mathbf{u}_t) - q_t(\mathbf{x}_t)\right)\right] \\
&\le \mathbb{E}\left[\max_{\mathbf{u}=(\mathbf{u}_1,\cdots,\mathbf{u}_T)\in\mathcal{U}} \sum_{t=a}^{b}\left(q_t(\mathbf{u}_t) - q_t(\mathbf{x}_t)\right)\right] \\
&= \mathcal{R}^{\mathcal{A}}_{\mathcal{B}'_\theta}(\mathcal{U})[a,b]
\end{aligned}$$

where the inequality follows from Jensen's inequality. Therefore

$$\begin{aligned}
&\sup_{\mathcal{B}\in\mathrm{Adv}} \max_{\mathbf{u}\in\mathcal{U}} \mathbb{E}\left[\sum_{t=a}^{b}\left(\langle \nabla f_t(\mathbf{x}_t), \mathbf{u}_t - \mathbf{x}_t\rangle - \frac{\mu}{2}\|\mathbf{u}_t - \mathbf{x}_t\|\right)\right] \\
&\qquad \le \sup_{\mathcal{B}\in\mathrm{Adv}, \theta\in\Omega^{\mathcal{Q}}} \mathcal{R}^{\mathcal{A}}_{\mathcal{B}'_\theta} \\
&\qquad \le \sup_{\mathcal{B}'\in\mathrm{Adv}^{\mathrm{f}}_1(\mathbf{F}_\mu[B_1])} \mathcal{R}^{\mathcal{A}}_{\mathcal{B}'} \\
&\qquad = \mathcal{R}^{\mathcal{A}}_{\mathrm{Adv}^{\mathrm{f}}_1(\mathbf{F}_\mu[B_1])}
\end{aligned}$$

□

D PROOF OF THEOREM 10

Proof.

(i):

We have

$$\begin{aligned}
\mathcal{R}^{\mathcal{W}(\mathcal{A})}_{\alpha,\mathrm{Adv}^{\mathrm{f}}_1(\mathbf{F})} &= \sup_{\mathcal{B}\in\mathrm{Adv}^{\mathrm{f}}_1(\mathbf{F})} \mathbb{E}\left[\max_{\mathbf{u}=(\mathbf{u}_1,\cdots,\mathbf{u}_T)\in\mathcal{U}} \sum_{t=a}^{b}\left(\alpha f_t(\mathbf{u}_t) - f_t(\mathcal{W}(\mathbf{x}_t))\right)\right] \\
&\le \sup_{\mathcal{B}\in\mathrm{Adv}^{\mathrm{f}}_1(\mathbf{F})} \max_{\mathbf{u}=(\mathbf{u}_1,\cdots,\mathbf{u}_T)\in\mathcal{U}} \mathbb{E}\left[\sum_{t=a}^{b}\beta\left(\langle \nabla \mathcal{W}(f_t)(\mathbf{x}_t), \mathbf{u}_t - \mathbf{x}_t\rangle - \frac{\mu}{2}\|\mathbf{u}_t - \mathbf{x}_t\|\right)\right] \\
&= \beta\mathcal{R}^{\mathcal{A}}_{1,\mathrm{Adv}^{\mathrm{f}}_1(\mathbf{H}_{\mu,\nabla})}.
\end{aligned}$$

(ii):

Since Adv is oblivious, the sequence of functions $(f_1, \cdots, f_T)$ is not random and we have

$$
\begin{aligned}
\mathcal{R}^{\mathcal{W}(\mathcal{A})}_{\alpha,\mathrm{Adv}^{\mathrm{o}}_1(\mathbf{F},B_1)} &= \sup_{\mathcal{B}\in\mathrm{Adv}^{\mathrm{o}}_1(\mathbf{F},B_1)} \mathbb{E}\left[\max_{\mathbf{u}=(\mathbf{u}_1,\cdots,\mathbf{u}_T)\in\mathcal{U}} \sum_{t=a}^{b} \left(\alpha f_t(\mathbf{u}_t) - f_t(\mathcal{W}(\mathbf{x}_t))\right)\right] \\
&= \sup_{\mathcal{B}\in\mathrm{Adv}^{\mathrm{o}}_1(\mathbf{F},B_1)} \max_{\mathbf{u}=(\mathbf{u}_1,\cdots,\mathbf{u}_T)\in\mathcal{U}} \mathbb{E}\left[\sum_{t=a}^{b} \left(\alpha f_t(\mathbf{u}_t) - f_t(\mathcal{W}(\mathbf{x}_t))\right)\right] \\
&\leq \sup_{\mathcal{B}\in\mathrm{Adv}^{\mathrm{o}}_1(\mathbf{F},B_1)} \max_{\mathbf{u}=(\mathbf{u}_1,\cdots,\mathbf{u}_T)\in\mathcal{U}} \mathbb{E}\left[\sum_{t=a}^{b} \beta\left(\langle \nabla\mathcal{W}(f_t)(\mathbf{x}_t), \mathbf{u}_t - \mathbf{x}_t\rangle - \frac{\mu}{2}\|\mathbf{u}_t - \mathbf{x}_t\|\right)\right] \\
&= \beta\mathcal{R}^{\mathcal{A}}_{1,\mathrm{Adv}^{\mathrm{f}}_1(\mathbf{Q}_\mu[B_1'])}.
\end{aligned}
$$

□

E Follow The Regularized Leader

We start by defining the notion of self-concordant barrier.

Definition 6 ((Hazan et al., 2016))**.** Let $\mathcal{K} \in \mathbb{R}^d$ be a convex set with non empty interior $\mathrm{int}(\mathcal{K})$. We call a function $\Phi : \mathrm{int}(\mathcal{K}) \longrightarrow \mathbb{R}$ a *ν-self-concordant barrier* of $\mathcal{K}$ if:

(i) Φ is three-times continuously differentiable, convex, and tends to infinity along any sequence of points approaching the boundary of $\mathcal{K}$;

(ii) For every $\mathbf{h} \in \mathbb{R}^d$ and $\mathbf{x} \in \mathrm{int}(\mathcal{K})$, we have:

$$|\nabla^3\Phi(\mathbf{x})[\mathbf{h},\mathbf{h},\mathbf{h}]| \leq 2(\nabla^2\Phi(\mathbf{x})[\mathbf{h},\mathbf{h}])^{3/2}, \quad |\nabla\Phi(x)[\mathbf{h}]| \leq \nu^{1/2}(\nabla^2\Phi(\mathbf{x})[\mathbf{h},\mathbf{h}])^{1/2}$$

where the third-order differential is defined as $\nabla^3\Phi(\mathbf{x})[\mathbf{h},\mathbf{h},\mathbf{h}] := \frac{\partial^3}{\partial t_1 \partial t_2 \partial t_3}\Phi(x + t_1\mathbf{h} + t_2\mathbf{h} + t_3\mathbf{h})|_{t_1=t_2=t_3=0}$.

Next we define the notion of local norm and dual norm with respect to a self-concordant barrier.

Definition 7. For every $x \in \mathrm{int}(\mathcal{K})$, the Hessian of the self-concordant barrier induces a *local* norm, denoted as $\|\cdot\|_{\Phi,x}$, and a *dual* norm, denoted as $\|\cdot\|_{\Phi,\mathbf{x},*}$, where for any $\mathbf{v} \in \mathbb{R}^d$,

$$\|\mathbf{v}\|_{\Phi,\mathbf{x}} = \sqrt{\mathbf{v}^T\nabla^2\Phi(\mathbf{x})\mathbf{v}}, \qquad \|\mathbf{v}\|_{\Phi,\mathbf{x},*} = \sqrt{\mathbf{v}^T(\nabla^2\Phi(\mathbf{x}))^{-1}\mathbf{v}}.$$

An important result for FTRL is the following theorem which was proved in (Abernethy et al., 2008). It shows that if we set the regularizer to be a self-concordant barrier of $\mathcal{K}$ and the algorithm can access the unbiased estimator of $\mathbf{g}_t$, then the regret of the generated solution sequence $\{x_t\}_{t=1}^T$ can be bounded in terms of the local norm of the estimator.

Theorem 11 ((Abernethy et al., 2008))**.** *Let $\mathcal{K} \subseteq \mathbb{R}^d$ be a convex set, $\Phi(x)$ be a self-concordant barrier on $\mathcal{K}$, $\{\mathbf{g}_t\}_{t=1}^T$ be a sequence of random vectors in $\mathbb{R}^d$. Then running FTRL (described in Equation 3) on a vector sequence $\{\mathbf{g}_t\}_{t=1}^T$ in $\mathbb{R}^d$ with $\Phi(x)$ as the regularizer will produce a sequence of point $\{\mathbf{x}_t\}_{t=1}^T$ in $\mathcal{K}$ where*

$$\sum_{t=1}^{T}\langle \mathbf{g}_t, \mathbf{y} - \mathbf{x}_t\rangle \leq \eta\sum_{t=1}^{T}\|\mathbf{g}_t\|^2_{\Phi,\mathbf{x}_t,*} + \frac{\Phi(\mathbf{y}) - \Phi(\mathbf{x}_1)}{\eta},$$

for any $\mathbf{y} \in \mathcal{K}$.

The ellipsoid gradient estimator was proposed in (Abernethy et al., 2008), where the authors use it along with Theorem 11 to design an $\widetilde{O}(\sqrt{T})$ regret algorithm for bandit linear optimization. For a continuous function but possibly non-smooth $f : \mathbb{R}^d \to \mathbb{R}$ and an invertible matrix $\Sigma \in \mathbb{R}^{d\times d}$, we define the Σ-smoothed version of f.

Definition 8. For function $f(\mathbf{x}) : \mathbb{R}^d \to \mathbb{R}$ and invertible matrix $\Sigma \in \mathbb{R}^{d\times d}$, we call $f^{\Sigma}(\mathbf{x})$ a Σ-*smoothed version* of $f(\mathbf{x})$, where $f^{\Sigma}(\mathbf{x}) = \mathbb{E}_{\mathbf{v}\sim\mathbb{B}^d}\left[f(\mathbf{x}+\Sigma\mathbf{v})\right]$. Here $\mathbf{v}\sim\mathbb{B}^d$ means that $\mathbf{v}$ is sampled from the unit ball $\mathbb{B}^d$ uniformly at random.

There is a surprising fact that there is an unbiased estimator of $\nabla f^{\Sigma}(\mathbf{x})$ for any $\mathbf{x}$, and the estimator uses only a single query to the value oracle of f.

Lemma 4 ((Abernethy et al., 2008)). *Let $\Sigma \in \mathcal{R}^{d\times d}$ be an invertible matrix, $f(\mathbf{x}) : \mathbb{R}^d \to \mathbb{R}$ be an arbitrary function. Then $\nabla f^{\Sigma}(\mathbf{x}) = d\mathbb{E}_{\mathbf{v}\sim\mathbb{S}^{d-1}}\left[f(\mathbf{x}+\Sigma\mathbf{v})\Sigma^{-1}\mathbf{v}\right]$. Here $\mathbf{v}\sim\mathbb{S}^{d-1}$ means that $\mathbf{v}$ is sampled from the $(d-1)$-dimensional unit sphere $\mathbb{S}^{d-1}$ uniformly at random.*

If f is a linear function, $f^{\Sigma}(\mathbf{x}) = f(\mathbf{x})$, so Lemma 4 provides a one-sample unbiased estimator of the gradient of the linear function. The ellipsoid gradient estimator is usually used along with FTRL with a self-concordant regularizer Φ of $\mathcal{K}$. When the invertible matrix Σ is set to be $(\nabla^2\Phi(\mathbf{x}))^{-1/2}$ and $\mathbf{x}\in \mathrm{int}(\mathcal{K})$, the sampled action $\mathbf{x}+\Sigma\mathbf{v}$ is located in the surface of a so-called **Dikin ellipsoid** centered at $\mathbf{x}$, i.e. $\{\mathbf{x}' \mid \|\mathbf{x}'-\mathbf{x}\|_{\Phi,\mathbf{x}} \le 1\}$. The fact that Dikin ellipsoid is entirely contained in $\mathcal{K}$ allows us to define f^{Σ} at $\mathbf{x}$.

We finish this section with quick overview of the concept of the Minkowski function, the Minkowski set and some of their useful properties.

Definition 9. Let $\mathcal{K}$ be a compact convex set, the Minkowski function $\pi_{\mathbf{x}} : \mathcal{K} \to \mathbb{R}$ parameterized by a pole $\mathbf{x}\in\mathrm{int}(\mathcal{K})$ is defined as $\pi_{\mathbf{x}}(\mathbf{y}) := \inf\{t \ge 0 \mid x + t^{-1}(y-x) \in \mathcal{K}\}$. Given $\delta\in\mathbb{R}^+$ and $\mathbf{x}_1 \in \mathrm{int}(\mathcal{K})$, we define the Minkowski set

$$\mathcal{K}_{\gamma,\mathbf{x}_1} := \{\mathbf{x}\in\mathcal{K} \mid \pi_{\mathbf{x}_1}(\mathbf{x}) \le (1+\gamma)^{-1}\}.$$

Lemma 5 ((Abernethy et al., 2008)). *Let $\mathcal{K}$ be a compact convex set, $\mathbf{x}\in\mathrm{int}(\mathcal{K})$ with diameter D, $\mathbf{u}_*\in\mathcal{K}$ and $\hat{\mathbf{u}}_* := \mathrm{argmin}_{\mathbf{z}\in\mathcal{K}_{\gamma,\mathbf{x}}}\|\mathbf{z}-\mathbf{u}_*\|$ be the projection of $\mathbf{u}_*$ onto the Minkowski set $\mathcal{K}_{\gamma,\mathbf{x}}$, then*

$$\|\mathbf{u}_* - \hat{\mathbf{u}}_*\| \le \gamma D.$$

The following lemma provides an upper bound of the difference between the function value of a self-concordant barrier at two different points.

Lemma 6 ((Nesterov & Nemirovskii, 1994)). *Let Φ be a ν-self-concordant barrier over a compact convex set $\mathcal{K}$, then for all $\mathbf{x},\mathbf{y}\in\mathrm{int}(\mathcal{K})$:*

$$\Phi(\mathbf{y}) - \Phi(\mathbf{x}) \le \nu \log\frac{1}{1-\pi_{\mathbf{x}}(\mathbf{y})}.$$

F Technical Lemmas

This section provides some technical lemmas that will be used in the proofs later.

Lemma 7. *Let $\mathcal{K}$ be a compact set and let $f : \mathcal{K} \to \mathbb{R}^d$ be an M_2-smooth function. Then f may be extended to an M_2-smooth function $\tilde{f} : \mathbb{R}^d \to \mathbb{R}$.*

Proof. The function ∇F is an M_2-Lipschitz function defined on $\mathcal{K}$. Therefore, according to Kirszbraun theorem (Kirszbraun, 1934) it may be extended to a function $g : \mathbb{R}^d \to \mathbb{R}^d$ that is M_2-Lipschitz. Now the result follows directly from Whitney's extension theorem (Whitney, 1934). □

Parts (i)-(iii) of the following lemma are well-known in the literature. (See Lemma A.5 in (Wan et al., 2023) for a proof). Here we provide a proof for part (iv).

Lemma 8. *Following properties hold for $\mathbf{\Sigma}$-smoothed version of a function $f(\mathbf{x})$ for an invertible matrix Σ.*

(i) If $f(\mathbf{x})$ is a monotone function, then so is $f^{\Sigma}(\mathbf{x})$.

(ii) If $f(\mathbf{x})$ is M_1-Lipschitz, then so is $f^{\Sigma}(\mathbf{x})$.

(iii) If $f(\mathbf{x})$ is M_2-smooth, then so is $f^{\Sigma}(\mathbf{x})$.

(iv) If f is upper-quadratizable with a uniform wrapper $\mathcal{W}$ and α, β and μ, then we have

$$\alpha f^{\Sigma}(\mathbf{y}) - (f \circ \mathcal{W})^{\Sigma}(\mathbf{x}) \leq \beta \left(\left\langle \nabla (\mathcal{W}(f))^{\Sigma}(\mathbf{x}), \mathbf{y} - \mathbf{x} \right\rangle - \frac{\mu}{2} \|\mathbf{y} - \mathbf{x}\|^2 \right).$$

Proof. We have

$$\begin{aligned}
\alpha f^{\Sigma}(\mathbf{y}) - (f \circ \mathcal{W})^{\Sigma}(\mathbf{x}) &= \mathbb{E}_{\mathbf{v} \sim \mathbb{B}^d} \left[\alpha f(\mathbf{y} + \Sigma \mathbf{v}) - f(\mathcal{W}(\mathbf{x} + \Sigma \mathbf{v})) \right] \\
&\leq \mathbb{E}_{\mathbf{v} \sim \mathbb{B}^d} \left[\beta \left(\langle \nabla \mathcal{W}(f)(\mathbf{x} + \Sigma \mathbf{v}), \mathbf{y} - \mathbf{x} \rangle - \frac{\mu}{2} \|\mathbf{y} - \mathbf{x}\|^2 \right) \right] \\
&= \beta \left(\left\langle \mathbb{E}_{\mathbf{v} \sim \mathbb{B}^d} \left[\nabla \mathcal{W}(f)(\mathbf{x} + \Sigma \mathbf{v}) \right], \mathbf{y} - \mathbf{x} \right\rangle - \frac{\mu}{2} \|\mathbf{y} - \mathbf{x}\|^2 \right) \\
&= \beta \left(\left\langle \nabla \mathbb{E}_{\mathbf{v} \sim \mathbb{B}^d} \left[\mathcal{W}(f)(\mathbf{x} + \Sigma \mathbf{v}) \right], \mathbf{y} - \mathbf{x} \right\rangle - \frac{\mu}{2} \|\mathbf{y} - \mathbf{x}\|^2 \right) \\
&= \beta \left(\left\langle \nabla (\mathcal{W}(f))^{\Sigma}(\mathbf{x}), \mathbf{y} - \mathbf{x} \right\rangle - \frac{\mu}{2} \|\mathbf{y} - \mathbf{x}\|^2 \right).
\end{aligned}$$

□

Lemma 9. *If $f : \mathcal{K} \to \mathbb{R}$ is M_1-Lipschitz and M_2-smooth and $g : \mathcal{K} \to \mathcal{K}$ is M_1'-Lipschitz and M_2'-smooth, then $f \circ g$ is M_1''-Lipschitz and M_2''-smooth where $M_1'' := M_1 M_1'$ and $M_2'' := M_1 M_2' + M_2 {M_1'}^2$.*

Proof. We have

$$\|D(f \circ g)(\mathbf{x})\| = \|Df(g(\mathbf{x})) \cdot Dg(\mathbf{x})\| \leq M_1 M_1',$$

and therefore for all $\mathbf{x}, \mathbf{y} \in \mathcal{K}$, we have

$$\begin{aligned}
\|D(f \circ g)(\mathbf{x}) - D(f \circ g)(\mathbf{y})\| &= \|Df(g(\mathbf{x})) \cdot Dg(\mathbf{x}) - Df(g(\mathbf{y})) \cdot Dg(\mathbf{y})\| \\
&\leq \|Df(g(\mathbf{x})) \cdot Dg(\mathbf{x}) - Df(g(\mathbf{x})) \cdot Dg(\mathbf{y})\| \\
&\quad + \|Df(g(\mathbf{x})) \cdot Dg(\mathbf{y}) - Df(g(\mathbf{y})) \cdot Dg(\mathbf{y})\| \\
&= \|Df(g(\mathbf{x}))\| \|Dg(\mathbf{x}) - Dg(\mathbf{y})\| \\
&\quad + \|Df(g(\mathbf{x})) - Df(g(\mathbf{y}))\| \|Dg(\mathbf{y})\| \\
&\leq M_1 M_2' \|\mathbf{x} - \mathbf{y}\| + M_2 M_1' \|g(\mathbf{x}) - g(\mathbf{y})\| \\
&\leq (M_1 M_2' + M_2 {M_1'}^2) \|\mathbf{x} - \mathbf{y}\|.
\end{aligned}$$

□

G PROOF OF THEOREM 1

Proof of Theorem 1. We have

$$\begin{aligned}
&\sum_{t=1}^{T} \mathbb{E} \left[f_t(\mathbf{u}_*) - f_t(\mathbf{x}_t) \right] \\
&\quad = \sum_{t=1}^{T} \mathbb{E} \left[f_t^{\delta \Sigma_t}(\hat{\mathbf{u}}_*) - f_t^{\delta \Sigma_t}(\mathbf{x}_t) \right] + \underbrace{\sum_{t=1}^{T} \mathbb{E} \left[f_t^{\delta \Sigma_t}(\mathbf{u}_*) - f_t^{\delta \Sigma_t}(\hat{\mathbf{u}}_*) \right]}_{(A)} \\
&\quad + \underbrace{\sum_{t=1}^{T} \mathbb{E} \left[f_t(\mathbf{u}_*) - f_t^{\delta \Sigma_t}(\mathbf{u}_*) \right]}_{(B)} + \underbrace{\sum_{t=1}^{T} \mathbb{E} \left[f_t^{\delta \Sigma_t}(\mathbf{x}_t) - f_t(\mathbf{x}_t) \right]}_{(C)}
\end{aligned} \tag{4}$$

Note that, for the terms above to be well-defined, we need to be able to define $f_t^{\delta \Sigma_t}$ over $\mathcal{K}$ which requires computing f_t over a set that is slightly larger than $\mathcal{K}$. Using Lemma 7, we assume that all functions f_t are well-defined and M_2-smooth over $\mathbb{R}^d$.

Bounding (A): Since $f_t(\mathbf{x})$ is M_1-Lipschitz continuous, $f_t^{\delta\Sigma_t}$ is also M_1-Lipschitz continuous by Lemma 8. Since $\|\hat{\mathbf{u}}_* - \mathbf{u}_*\| \leq \gamma D$ by Lemma 5,

$$\begin{aligned}\sum_{t=1}^{T} \mathbb{E}\left[f_t^{\delta\Sigma_t}(\mathbf{u}_*) - f_t^{\delta\Sigma_t}(\hat{\mathbf{u}}_*)\right] &\leq \sum_{t=1}^{T} \mathbb{E}\left[\left|f_t^{\delta\Sigma_t}(\hat{\mathbf{u}}_*) - f_t^{\delta\Sigma_t}(\mathbf{u}_*)\right|\right] \\ &\leq \sum_{t=1}^{T} M_1\gamma D = M_1\gamma DT.\end{aligned} \tag{5}$$

Bounding (B): Since $f_t(\mathbf{x})$ is M_2-smooth, by Lemma 8, $f_t^{\delta\Sigma_t}$ is M_2-smooth. Thus,

$$\begin{aligned}f_t(\mathbf{u}_*) - f_t^{\delta\Sigma_t}(\mathbf{u}_*) &= \mathbb{E}_{\mathbf{v}\sim\mathbb{B}^d}\left[f_t(\mathbf{u}_*) - f_t(\mathbf{u}_* + \delta\Sigma_t\mathbf{v})\right] \\ &\leq \mathbb{E}_{\mathbf{v}\sim\mathbb{B}^d}\left[-\langle\nabla f_t(\mathbf{u}_*), \delta\Sigma_t\mathbf{v}\rangle + \frac{M_2}{2}\|\delta\Sigma_t\mathbf{v}\|^2\right] \\ &= \mathbb{E}_{\mathbf{v}\sim\mathbb{B}^d}\left[-\langle\nabla f_t(\mathbf{u}_*), \delta\Sigma_t\mathbf{v}\rangle\right] + \mathbb{E}_{\mathbf{v}\sim\mathbb{B}^d}\left[\frac{M_2}{2}\|\delta\Sigma_t\mathbf{v}\|^2\right] \\ &= \mathbb{E}_{\mathbf{v}\sim\mathbb{B}^d}\left[\frac{M_2}{2}\|\delta\Sigma_t\mathbf{v}\|^2\right] \\ &\leq \frac{M_2\delta^2 D^2}{2}.\end{aligned}$$

Note that in the last inequality, we used the fact that the Dikin ellipsoid centered at $\mathbf{x}_t$ is contained in $\mathcal{K}$ which implies that $\mathbf{x}_t + \Sigma_t\mathbf{v} \in \mathcal{K}$ and therefore $\|\Sigma_t\mathbf{v}\| \leq D$. It follows that,

$$\sum_{t=1}^{T} \mathbb{E}\left[f_t(\hat{\mathbf{u}}_*) - f_t^{\delta\Sigma_t}(\hat{\mathbf{u}}_*)\right] \leq \frac{M_2\delta^2 D^2 T}{2}. \tag{6}$$

Bounding (C): Similarly,

$$\begin{aligned}f_t^{\delta\Sigma_t}(\mathbf{x}_t) - f_t(\mathbf{x}_t) &= \mathbb{E}_{\mathbf{v}\sim\mathbb{B}^d}\left[f_t(\mathbf{x}_t + \delta\Sigma_t\mathbf{v}) - f_t(\mathbf{x}_t)\right] \\ &\leq \mathbb{E}_{\mathbf{v}\sim\mathbb{B}^d}\left[-\left\langle\nabla f_t(\mathbf{x}_t), \delta\Sigma_t\mathbf{v}\right\rangle + \frac{M_2}{2}\|\delta\Sigma_t\mathbf{v}\|^2\right] \leq \frac{M_2\delta^2 D^2}{2}.\end{aligned}$$

Therefore,

$$\sum_{t=1}^{T} \mathbb{E}\left[f_t^{\delta\Sigma_t}(\mathbf{x}_t) - f_t(\mathbf{x}_t)\right] \leq \frac{M_2\delta^2 D^2 T}{2} \tag{7}$$

Putting 5,6,7 in 4, we see that

$$\begin{aligned}\sum_{t=1}^{T} \mathbb{E}\left[f_t(\mathbf{u}_*) - f_t(\mathbf{x}_t)\right] \leq \sum_{t=1}^{T} \mathbb{E}\left[f_t^{\delta\Sigma_t}(\hat{\mathbf{u}}_*) - f_t^{\delta\Sigma_t}(\mathbf{x}_t)\right] \\ + \alpha M_1\gamma DT + \frac{M_2\delta^2 D^2 T}{2} + \frac{M_2\delta^2 D^2 T}{2},\end{aligned}$$

which completes the proof of the first claim.

To prove the second claim, we first use Lemma 4, with $\Sigma = \delta\Sigma_t$, to see that $\mathbb{E}\left[\mathbf{o}_t \mid \mathbf{x}_t\right] = \nabla f_t^{\delta\Sigma_t}(\mathbf{x}_t)$. On the other hand, since $\mathcal{Q}_t$ is bounded by B_0, we have

$$\|\mathbf{o}_t\|_{\mathbf{x}_t,*}^2 = \left\|\frac{d}{\delta}y_t\Sigma_t^{-1}\mathbf{v}_t\right\|_{\mathbf{x}_t,*}^2 = \frac{d^2}{\delta^2}|y_t|^2\mathbf{v}_t^T\Sigma_t^{-1}\left(\nabla^2\Phi(\mathbf{x}_t)\right)^{-1}\Sigma_t^{-1}\mathbf{v}_t \leq \frac{d^2}{\delta^2}B_0^2\|\mathbf{v}_t\|^2 \leq \frac{d^2B_0^2}{\delta^2}$$

Hence, using Theorem 11 with $\mathbf{g}_t = \mathbf{o}_t$ and $\mathbf{y} = \hat{\mathbf{u}}_*$, we see that

$$\begin{aligned}\sum_{t=1}^{T} \mathbb{E}\left[\langle \nabla f_t^{\delta\Sigma_q}(\mathbf{x}_t), \hat{\mathbf{u}}_* - \mathbf{x}_t \rangle\right] &= \sum_{t=1}^{T} \mathbb{E}\left[\langle \mathbb{E}\left[\mathbf{o}_t \mid \mathbf{x}_t\right], \hat{\mathbf{u}}_* - \mathbf{x}_t \rangle\right] \\ &= \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{E}\left[\langle \mathbf{o}_t, \hat{\mathbf{u}}_* - \mathbf{x}_t \rangle \mid \mathbf{x}_t\right]\right] \\ &= \mathbb{E}\left[\sum_{t=1}^{T} \langle \mathbf{o}_t, \hat{\mathbf{u}}_* - \mathbf{x}_t \rangle\right] \\ &\le \mathbb{E}\left[\eta \sum_{t=1}^{T} \|\mathbf{o}_t\|_{\Phi,\mathbf{x}_t,*}^2 + \frac{\Phi(\hat{\mathbf{u}}_*) - \Phi(\mathbf{x}_1)}{\eta}\right] \\ &\le \eta \sum_{t=1}^{T} \frac{d^2 B_0^2}{\delta^2} + \frac{\Phi(\hat{\mathbf{u}}_*) - \Phi(\mathbf{x}_1)}{\eta} \\ &\le \frac{\eta d^2 B_0^2 T}{\delta^2} + \frac{\nu \log(\frac{1}{1-(1+\gamma)^{-1}})}{\eta},\end{aligned}$$

where we used Lemma 6 in the last inequality. □

H Proof of Theorem 2

Proof. Let $\mathcal{B} \in \mathrm{Adv}$ be a realized adversary and let $f_1, \cdots, f_T$ be the sequence of functions selected by $\mathcal{B}$. Also let $\mathbf{u}_* \in \operatorname{argmax}_{\mathbf{u}\in\mathcal{K}} \sum_{t=1}^{T} f_t(\mathbf{u})$ and $\hat{\mathbf{u}}_* \in \operatorname{argmin}_{\mathbf{x}\in\mathcal{K}_{\gamma,\mathbf{x}_1}} \|\mathbf{u}_* - \mathbf{x}\|$ where $\gamma = T^{-1}$. We have

$$\begin{aligned}\mathcal{R}_{\alpha,\mathcal{B}}^{\mathcal{W}(\text{ZO-FTRL})} &= \sum_{t=1}^{T} \mathbb{E}\left[\alpha f_t(\mathbf{u}_*) - f_t(\mathcal{W}(\mathbf{x}_t))\right] \\ &= \sum_{t=1}^{T} \mathbb{E}\left[\alpha f_t^{\delta\Sigma_t}(\hat{\mathbf{u}}_*) - (f_t \circ \mathcal{W})^{\delta\Sigma_t}(\mathbf{x}_t)\right] + \alpha \underbrace{\sum_{t=1}^{T} \mathbb{E}\left[f_t^{\delta\Sigma_t}(\mathbf{u}_*) - f_t^{\delta\Sigma_t}(\hat{\mathbf{u}}_*)\right]}_{(A)} \\ &\quad + \alpha \underbrace{\sum_{t=1}^{T} \mathbb{E}\left[f_t(\mathbf{u}_*) - f_t^{\delta\Sigma_t}(\mathbf{u}_*)\right]}_{(B)} + \underbrace{\sum_{t=1}^{T} \mathbb{E}\left[(f_t \circ \mathcal{W})^{\delta\Sigma_t}(\mathbf{x}_t) - f_t(\mathcal{W}(\mathbf{x}_t))\right]}_{(C)}\end{aligned}$$

As in the proof of Theorem 1, we use Lemma 7 to extend all functions f_t to M_2-smooth functions over $\mathbb{R}^d$ and we bound the terms (A) and (B) by $M_1 \gamma D T$ and $\frac{M_2 \delta^2 D^2 T}{2}$, respectively. To bound (C), we first use Lemma 9 to see that $f_t \circ \mathcal{W}$ is M_2''-smooth, where $M_2'' = M_1 M_2' + M_2 {M_1'}^2$. Hence, we see that

$$\begin{aligned}(f_t \circ \mathcal{W})^{\delta\Sigma_t}(\mathbf{x}_t) - f_t(\mathcal{W}(\mathbf{x}_t)) &= \mathbb{E}_{\mathbf{v}\sim\mathbb{B}^d}\left[f_t(\mathcal{W}(\mathbf{x}_t + \delta\Sigma_t \mathbf{v})) - f_t(\mathcal{W}(\mathbf{x}_t))\right] \\ &\le \mathbb{E}_{\mathbf{v}\sim\mathbb{B}^d}\left[-\left\langle \nabla f_t(\mathcal{W}(\mathbf{x}_t)), \delta\Sigma_t \mathbf{v}\right\rangle + \frac{M_2''}{2}\|\delta\Sigma_t \mathbf{v}\|^2\right] \le \frac{M_2'' \delta^2 D^2}{2}.\end{aligned}$$

Therefore,

$$\sum_{t=1}^{T} \mathbb{E}\left[(f_t \circ \mathcal{W})^{\delta\Sigma_t}(\mathbf{x}_t) - f_t(\mathcal{W}(\mathbf{x}_t))\right] \le \frac{M_2'' \delta^2 D^2 T}{2}$$

Putting the bounds for (A), (B) and (C) together, we see that

$$\begin{aligned}\mathcal{R}_{\alpha,\mathcal{B}}^{\mathcal{W}(\text{ZO-FTRL})} &= \sum_{t=1}^{T} \mathbb{E}\left[\alpha f_t(\mathbf{u}_*) - f_t(\mathcal{W}(\mathbf{x}_t))\right] \\ &\le \sum_{t=1}^{T} \mathbb{E}\left[\alpha f_t^{\delta\Sigma_t}(\hat{\mathbf{u}}_*) - (f_t \circ \mathcal{W})^{\delta\Sigma_t}(\mathbf{x}_t)\right] + \alpha M_1 \gamma D T + \frac{(\alpha M_2 + M_2'')\delta^2 D^2 T}{2} \\ &\le \sum_{t=1}^{T} \mathbb{E}\left[\beta\langle \nabla(\mathcal{W}(f_t))^{\delta\Sigma_t}(\mathbf{x}_t), \hat{\mathbf{u}}_* - \mathbf{x}_t\rangle\right] + \alpha M_1 \gamma D T + \frac{(\alpha M_2 + M_2'')\delta^2 D^2 T}{2},\end{aligned} \tag{8}$$

where the second inequality follows from Lemma 8.

To bound the remaining term, we use an argument similar to the one used in the proof of Theorem 1 again. Using Lemma 4 with $\Sigma = \delta\Sigma_t$ and the fact that y_t is an unbiased sample of $\mathcal{W}(f_t)$ at $\mathbf{x}_t + \delta\Sigma_t\mathbf{v}_t$, we see that $\mathbb{E}\left[\mathbf{o}_t \mid \mathbf{x}_t\right] = \nabla(\mathcal{W}(f_t))^{\delta\Sigma_t}(\mathbf{x}_t)$. On the other hand, since $\mathcal{W}(\mathcal{Q}_t)$ is bounded by B_0, we have $|y_t| \le B_0$, which implies that

$$\|\mathbf{o}_t\|_{\mathbf{x}_t,*}^2 = \left\|\frac{d}{\delta} y_t \Sigma_t^{-1}\mathbf{v}_t\right\|_{\mathbf{x}_t,*}^2 = \frac{d^2}{\delta^2}|y_t|^2 \mathbf{v}_t^T \Sigma_t^{-1}\left(\nabla^2\Phi(\mathbf{x}_t)\right)^{-1}\Sigma_t^{-1}\mathbf{v}_t \le \frac{d^2}{\delta^2}B_0^2\|\mathbf{v}_t\|^2 \le \frac{d^2 B_0^2}{\delta^2}.$$

Hence, using Theorem 11 with $\mathbf{g}_t = \mathbf{o}_t$ and $\mathbf{y} = \hat{\mathbf{u}}_*$, we see that

$$\begin{aligned}\sum_{t=1}^{T} \mathbb{E}\left[\beta\langle \nabla(\mathcal{W}(f_t))^{\delta\Sigma_t}(\mathbf{x}_t), \hat{\mathbf{u}}_* - \mathbf{x}_t\rangle\right] &= \beta \sum_{t=1}^{T} \mathbb{E}\left[\langle \mathbb{E}\left[\mathbf{o}_t \mid \mathbf{x}_t\right], \hat{\mathbf{u}}_* - \mathbf{x}_t\rangle\right] \\ &= \beta \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{E}\left[\langle \mathbf{o}_t, \hat{\mathbf{u}}_* - \mathbf{x}_t\rangle \mid \mathbf{x}_t\right]\right] \\ &= \beta\mathbb{E}\left[\sum_{t=1}^{T}\langle \mathbf{o}_t, \hat{\mathbf{u}}_* - \mathbf{x}_t\rangle\right] \\ &\le \beta\mathbb{E}\left[\eta\sum_{t=1}^{T}\|\mathbf{o}_t\|_{\Phi,\mathbf{x}_t,*}^2 + \frac{\Phi(\hat{\mathbf{u}}_*) - \Phi(\mathbf{x}_1)}{\eta}\right] \\ &\le \beta\eta\sum_{t=1}^{T}\frac{d^2 B_0^2}{\delta^2} + \beta\frac{\Phi(\hat{\mathbf{u}}_*) - \Phi(\mathbf{x}_1)}{\eta} \\ &\le \frac{\beta\eta d^2 B_0^2 T}{\delta^2} + \frac{\beta\nu\log(\frac{1}{1-(1+\gamma)^{-1}})}{\eta},\end{aligned}$$

where we used Lemma 6 in the last inequality. Plugging this into Equation 8 and using $M_2'' = M_1 M_2' + M_2 {M_1'}^2$ and $\gamma = T^{-1}$, we see that

$$\begin{aligned}\mathcal{R}_{\alpha,\mathcal{B}}^{\mathcal{W}(\text{ZO-FTRL})} &\le \frac{\beta\eta d^2 B_0^2 T}{\delta^2} + \frac{\beta\nu\log(\frac{1}{1-(1+\gamma)^{-1}})}{\eta} \\ &\qquad + \alpha M_1\gamma D T + \frac{\left(\alpha M_2 + M_1 M_2' + M_2 {M_1'}^2\right)\delta^2 D^2 T}{2} \\ &= O\left(\eta\delta^{-2}T + \eta^{-1}\log T + \delta^2 T\right).\end{aligned}$$

□

I Proof of Theorem 6

Proof. Note that in all three cases, $\mathcal{W}^{\text{action}}$ is 1-Lipschitz and 0-smooth. Now the result for the first case follows immediately from the fact that $\mathcal{W}^{\text{M}} = \text{Id}$. Also note that for any zeroth order query oracle Q_f for a function $f \in \mathbf{F}^{\mathrm{M0}}$ and any $\mathbf{y} \in \mathcal{K}$

$$|\mathcal{W}^{\mathrm{M0}}(\mathcal{Q}_f)(\mathbf{y})| = |z^{-1}\mathcal{Q}_f(z * \mathbf{y})| \leq z^{-1} \cdot C\|z * \mathbf{y}\| = \|\mathbf{y}\| \leq D.$$

Thus the query oracle $\mathcal{W}(\mathcal{Q}_f)$ is bounded by D and the assumptions of Theorem 2 are satisfied. The proof of boundedness of $\mathcal{W}^{\mathrm{NM}}(\mathcal{Q}_f)$ for any $f \in \mathbf{F}^{\mathrm{NM}}$ is similar. □

J Stochastic Full-information To Trivial query - SFTT

In this section, we discuss the SFTT meta-algorithm (Algorithm 4 in (Pedramfar & Aggarwal, 2024a)) which converts algorithms that require full-information feedback into algorithms that have a trivial query oracle. In particular, it converts algorithms require zeroth-order full-information feedback into bandit algorithms.

We say a function class $\mathbf{F}$ is closed under convex combination if for any $f_1, \cdots, f_k \in \mathbf{F}$ and any $\delta_1, \cdots, \delta_k \geq 0$ with $\sum_i \delta_i = 1$, we have $\sum_i \delta_i f_i \in \mathbf{F}$.

Theorem 12 (Theorem 7 and Remark 1 and Corollary 6 in (Pedramfar & Aggarwal, 2024a))**.** *Let $\mathcal{A}$ be an online optimization algorithm with full-information feedback and with K queries at each time-step where $\mathcal{A}^{query}$ does not depend on the observations in the current round and $\mathcal{A}' = \mathrm{SFTT}(\mathcal{A})$. Then, for any M_1-Lipschitz function class $\mathbf{F}$ that is closed under convex combination and any $B_1 \geq M_1$, $0 < \alpha \leq 1$ and $1 \leq a \leq b \leq T$, let $a' = \lfloor (a-1)/L \rfloor + 1$, $b' = \lceil b/L \rceil$, $D = \mathrm{diam}(\mathcal{K})$ and let $\{T\}$ and $\{T/L\}$ denote the horizon of the adversary. If we also have $\mathcal{R}^{\mathcal{A}'}_{\alpha,\mathrm{Adv}^o_i(\mathbf{F},B)}(\mathcal{K}^T_\star)[a,b] = O(BT^\eta)$, $K = O(1)$ and $L = O(T^{\frac{1-\eta}{2-\eta}})$, then*

$$\mathcal{R}^{\mathcal{A}'}_{\alpha,\mathrm{Adv}^o_i(\mathbf{F},B)}(\mathcal{K}^T_\star)[a,b] = O\left(BT^{\frac{1}{2-\eta}}\right).$$

More generally, the above result holds even if the query oracles are not bounded. Specifically, what we require is that the set of query oracles to be closed under convex combinations.

Algorithm 3: Stochastic Full-information To Trivial query - $\mathrm{SFTT}(\mathcal{A})$

Input : base algorithm $\mathcal{A}$, horizon T, block size $L > K$.
for $q = 1, 2, \dots, T/L$ **do**
  Let $\hat{\mathbf{x}}_q$ be the action chosen by $\mathcal{A}^{\text{action}}$
  Let $(\hat{\mathbf{y}}^i_q)^K_{i=1}$ be the queries selected by $\mathcal{A}^{\text{query}}$
  Let $(t_{q,1}, \dots, t_{q,L})$ be a random permutation of $\{(q-1)L + 1, \dots, qL\}$
  for $t = (q-1)L + 1, \dots, qL$ **do**
    if $t = t_{q,i}$ *for some* $1 \leq i \leq K$ **then**
      Play the action $\mathbf{x}_t = \hat{\mathbf{y}}^i_q$
      Return the observation to the query oracle as the response to the i-th query
    else
      Play the action $\mathbf{x}_t = \hat{\mathbf{x}}_q$
    end
  end
end

K Proof of Theorem 7

Proof. All three class of functions considered are closed under convex combination. Therefore we may directly apply Theorems 6 and 12 to obtain this result for the first case.

For any sequence of functions $f_1, \cdots, f_k$ and query oracles $\mathcal{Q}_1, \cdots, \mathcal{Q}_k$ for these functions that are contained within a cone $\mathrm{Cone}(\mathbf{0}, C)$ and non-negative numbers $\delta_1, \cdots, \delta_k$ such that $\sum_i \delta_i = 0$, the query oracle $\overline{Q}$ that uses $\mathcal{Q}_i$ with probability δ_i is trivially a query oracle for $\sum_i \delta_i f_i$ that is also contained within this cone. Therefore, we may apply Theorem 12 to obtain this result for the second case as well. The proof of the last case is similar. □

L Proof of Theorem 8

First we state the following simple result about OTB.

Theorem 13 (Theorem 8 in (Pedramfar & Aggarwal, 2024a)). *If $\mathcal{A}$ is an online algorithm that queries no more than $K = T^\theta$ times per time-step and obtains an α-regret bound of $O(T^\delta)$, then the sample complexity of $\operatorname{OTB}(\mathcal{A})$ is $\Omega(\epsilon^{-\frac{1+\theta}{1-\delta}})$.*

Proof of Theorem 8. This is an immediate corollary of Theorem 6 and the guarantees for the OTB meta-algorithm stated in Theorem 13. □