Online Learning for Non-monotone DR-Submodular Maximization: From Full Information to Bandit Feedback

Qixin Zhang
City University of Hong Kong

Zengde Deng[†]
Cainiao Network

Zaiyi Chen
Cainiao Network

Kuangqi Zhou
National University of Singapore

Haoyuan Hu
Cainiao Network

Yu Yang[†]
City University of Hong Kong

Abstract

In this paper, we revisit the online non-monotone continuous DR-submodular maximization problem over a down-closed convex set, which finds wide real-world applications in the domain of machine learning, economics, and operations research. At first, we present the **Meta-MFW** algorithm achieving a $1/e$-regret of $O(\sqrt{T})$ at the cost of $T^{3/2}$ stochastic gradient evaluations per round. As far as we know, **Meta-MFW** is the first algorithm to obtain $1/e$-regret of $O(\sqrt{T})$ for the online non-monotone continuous DR-submodular maximization problem over a down-closed convex set. Furthermore, in sharp contrast with **ODC** algorithm (Thang & Srivastav, 2021), **Meta-MFW** relies on the simple online linear oracle without discretization, lifting, or rounding operations. Considering the practical restrictions, we then propose the **Mono-MFW** algorithm, which reduces the per-function stochastic gradient evaluations from $T^{3/2}$ to 1 and achieves a $1/e$-regret bound of $O(T^{4/5})$. Next, we extend **Mono-MFW** to the bandit setting and propose the **Bandit-MFW** algorithm which attains a $1/e$-regret bound of $O(T^{8/9})$. To the best of our knowledge, **Mono-MFW** and **Bandit-MFW** are the first sublinear-regret algorithms to explore the one-shot and bandit setting for online non-monotone continuous DR-submodular maximization problem over a down-closed convex set, respectively. Finally, we conduct numerical experiments on both synthetic and real-world datasets to verify the effectiveness of our methods.

[†]Corresponding author. Proceedings of the 26[th] International Conference on Artificial Intelligence and Statistics (AISTATS) 2023, Valencia, Spain. PMLR: Volume 206.

1 Introduction

Continuous DR-submodular maximization draws wide attention since it mathematically depicts the diminishing return phenomenon in continuous domains. Numerous real-world applications in machine learning, operations research, and other related areas, such as non-definite quadratic programming (Ito & Fujimaki, 2016), revenue maximization (Soma & Yoshida, 2017; Bian et al., 2020), viral marketing (Kempe et al., 2003; Yang et al., 2016), determinantal point processes (Kulesza et al., 2012; Mitra et al., 2021), to name a few, could be modeled throughout the notion of continuous DR-submodularity.

In recent years, the prominent paradigm of online optimization (Zinkevich, 2003; Hazan et al., 2016) has led to spectacular successes in modelling the imperfect and complicated environment. In this framework, at each step, the online algorithm first chooses an action from a predefined set of feasible actions; then the adversary reveals the utility function. The objective of the online algorithm is to minimize the gap between the accumulative reward and that of the best fixed policy in hindsight.

Previously, a large body of algorithms (Bian et al., 2020; Mokhtari et al., 2020) with approximation guarantees rely on the monotone assumption of continuous DR-submodular functions. However, many real-world problems, such as the general DR-submodular quadratic programming (Ito & Fujimaki, 2016) and revenue maximization (Soma & Yoshida, 2017), are instances of non-monotone DR-submodular maximization. Motivated by these real applications, in this paper we focus on the problem of online non-monotone continuous DR-submodular maximization over a down-closed convex set under different feedbacks, i.e., full-information and bandit feedback.

Recently, based on a special online non-convex oracle, Thang & Srivastav (2021) presented the first online algorithm (**ODC**) for non-monotone continuous DR-submodular maximization over a down-closed convex set. **ODC** achieves a $1/e$-regret of $O(T^{3/4})$ where T is the horizon. Notably, the non-convex oracles of **ODC** need to discretize the original constrained domain and lift the n-dimensional subroutine problem into a solvable linear programming in a higher $(M \times n)$-dimensional space where $M = (T/n)^{1/4}$, which will incur a heavy computation burden when T is large. Moreover, the rounding operation (Mirrokni et al., 2017) in the online non-convex oracle assumes the knowledge of the vertices of the down-closed convex set, which is infeasible in many real applications. In this paper, we propose a new method to overcome these drawbacks. Motivated via the measured continuous greedy (Feldman et al., 2011), we first present the Meta-Measured Frank-Wolfe (**Meta-MFW**) algorithm, which achieves a faster $1/e$-regret of $O(T^{1/2})$ with only a simple online linear oracle.

Note that **ODC** and **Meta-MFW** require inquiring $T^{3/4}$ and $T^{3/2}$ stochastic gradient evaluations at each round, respectively. Therefore, when T is large, the huge amount of gradient estimates at each round makes both algorithms computationally prohibitive. In many scenarios, the stochastic gradient is 1) time-consuming to acquire, for instance, in the influence maximization task (Kempe et al., 2003; Yang et al., 2016), we need to generate enormous samples on large-scale social graphs to estimate the gradient, or 2) impossible to compute, e.g., black-box attacks and optimization (Chen et al., 2017; Ilyas et al., 2018; Chen et al., 2020). Considering these practical limitations, we also want to extend our proposed **Meta-MFW** into both one-shot and bandit feedback scenarios. As for the one-shot setting, we merge the blocking procedures (Zhang et al., 2019) into **Meta-MFW** to present **Mono-MFW** algorithm which yields a result with a $1/e$-regret of $O(T^{4/5})$ and reduces the number of per-function stochastic gradient evaluations from $T^{3/2}$ (or $T^{3/4}$) to 1. Finally, in the bandit feedback where the only observable information is the reward we receive, a new algorithm **Bandit-MFW** is proposed with the exploration-exploitation policy (Zhang et al., 2019) and achieves a $1/e$-regret of $O(T^{8/9})$.

To be specific, we make the following contributions:

1. We first develop a new algorithm **Meta-MFW** for online non-monotone continuous DR-submodular maximization problem over a down-closed convex set, which only relies on the simple online linear oracle without discretization, lifting, or rounding operations. Moreover, in sharp contrast with **ODC** (Thang & Srivastav, 2021), **Meta-MFW** achieves a faster $1/e$-regret of $O(T^{1/2})$ at the cost of $T^{3/2}$ stochastic gradient evaluations for each reward function. It is worth mentioning that the $1/e$-regret of $O(T^{1/2})$ result not only achieves the best-known approximation guarantee for the offline problem (Bian et al., 2017a) but also matches the optimal $O(\sqrt{T})$ regret (Hazan et al., 2016). Meanwhile, like **ODC** algorithm, our proposed **Meta-MFW** also can achieve $1/e$-regret of $O(T^{3/4})$ with $T^{3/4}$ per-function stochastic gradient evaluations.
2. Considering the practical restrictions, we then present the one-shot algorithm **Mono-MFW** equipped with blocking procedures (Zhang et al., 2019), which achieves a $1/e$-regret of $O(T^{4/5})$ and reduces the stochastic gradient evaluations from $T^{3/2}$ or $T^{3/4}$ to 1 at each round. Next, in the bandit setting, we propose the **Bandit-MFW** algorithm achieving a $1/e$-regret of $O(T^{8/9})$ by only inquiring one-point function value for each reward function. To the best of our knowledge, **Mono-MFW** and **Bandit-MFW** are the first sublinear-regret algorithm to explore the one-shot and bandit settings for online non-monotone continuous DR-submodular maximization problem over a down-closed convex set, respectively.
3. Finally, we empirically evaluate our proposed methods on both synthetic and real-world datasets. Numerical experiments demonstrate the superior performance of our proposed algorithms.

1.1 Related Work

Continuous DR-submodular maximization problem has been extensively investigated as it admits efficient approximate maximization routines. In this section, we provide a summary about these known results.

Monotone Setting: In the deterministic setting, Bian et al. (2017b) first proposed a variant of Frank-Wolfe achieving $(1 - 1/e)OPT - \epsilon$ after $O(1/\epsilon)$ iterations where OPT is the optimal objective value. When a stochastic gradient oracle is available, Hassani et al. (2017) proved that the stochastic gradient ascent guarantees $(1/2)OPT - \epsilon$ after $O(1/\epsilon^2)$ iterations. Next, Mokhtari et al. (2018) proposed the stochastic continuous greedy algorithm, which achieves a $(1-1/e)$-approximation after $O(1/\epsilon^3)$ iterations. Then, an accelerated stochastic continuous greedy algorithm is presented in Hassani et al. (2020), which guarantees a $(1-1/e)OPT-\epsilon$ after $O(1/\epsilon^2)$ iterations. As for the online settings, Chen et al. (2018b) first investigated the online gradient ascent with a $(1/2)$-regret of $O(\sqrt{T})$. Then, inspired by the meta actions (Streeter & Golovin, 2008), Chen et al. (2018b) proposed the Meta-Frank-Wolfe algorithm with a $(1 - 1/e)$-regret bound of $O(\sqrt{T})$ under the deterministic setting. With an unbiased gradient oracle, then Chen et al. (2018a) proposed a variant of the Meta-Frank-Wolfe algorithm having a $(1 - 1/e)$-regret bound of $O(T^{1/2})$ and requiring $T^{3/2}$ stochastic gradient queries for each function. In order to reduce the number of gradient evaluations, Zhang et al. (2019) presented Mono-Frank-Wolfe taking the blocking procedure, which achieves a $(1 - 1/e)$-regret bound of $O(T^{4/5})$ with only one stochastic gradient evaluation at each round. Leveraging this one-shot algorithm, Zhang et al. (2019) also presented a bandit algorithm Bandit-Frank-Wolfe achieving $(1 - 1/e)$-regret bound of $O(T^{8/9})$. Next, Sadeghi & Fazel (2020) and Sadeghi et al. (2020) extended the Meta-Frank-Wolfe algorithm to the online continuous DR-submodular maximization problem with long-term or stochastic constraints. Recently, based on a novel auxiliary function, Zhang et al. (2022) have presented a variant of gradient ascent improving the approximation ratio of the standard gradient ascent (Hassani et al., 2017; Chen et al., 2018b) from $1/2$ to $1-1/e$ in both offline and online settings.

Non-Monotone Setting: Without the monotone property, maximizing the continuous DR-submodular function becomes much harder. Under the down-closed convex constraint, Bian et al. (2017a) proposed the deterministic Two-Phase Frank-Wolfe and Non-monotone Frank-Wolfe with $1/4$-approximation and $1/e$-approximation guarantee, respectively. When only an unbiased estimate of gradient is available, Hassani et al. (2020) improved the Non-monotone Frank-Wolfe by variance reduction technique, which yields a result with $1/e$-approximation guarantee. Moreover, inspired by the Double Greedy (Buchbinder et al., 2015; Buchbinder & Feldman, 2018) for discrete unconstrained submodular set maximization, Niazadeh et al. (2018) and Bian et al. (2019) proposed a similar $1/2$-approximation algorithms for unconstrained continuous DR-submodular maximization. Note that Vondrák (2013) pointed that any algorithm with a constant-factor approximation for maximizing a non-monotone DR-submodular function over a non-down-closed convex set would require exponentially many value queries and the approximation guarantee of $1/2$ is tight for unconstrained DR-submodular maximization. Thang & Srivastav (2021) is the first work to explore the sublinear-regret online algorithm for the non-monotone continuous DR-submodular maximization problems over a down-closed convex set.

We present a comparison between this work and previous studies in Table 1.

2 Preliminaries

Notation: In this paper, a lower boldface denotes a vector with suitable dimension and an uppercase boldface for a matrix. For each vector $\mathbf{x}$, the i-th element of $\mathbf{x}$ is denoted as $(\mathbf{x})_i$. Specially, $\mathbf{0}$ and $\mathbf{1}$ represent the vector whose elements are all zero or one, respectively. For any positive integer number K, the symbol $[K]$ denotes the set $\{1,\dots,K\}$. Moreover, the symbol $\odot$ and $\oslash$ denote coordinate-wise multiplication and coordinate-wise division, respectively. For instance, given two vector $\mathbf{x}$ and $\mathbf{y}$, if $\mathbf{y}>\mathbf{0}$, the i-th element of vector $\mathbf{x}\oslash\mathbf{y}$ is $\frac{(\mathbf{x})_i}{(\mathbf{y})_i}$. The product $\langle\mathbf{x},\mathbf{y}\rangle=\sum_i(\mathbf{x})_i(\mathbf{y})_i$ and the norm $\|\mathbf{x}\|=\sqrt{\langle\mathbf{x},\mathbf{x}\rangle}$. We say the domain $\mathcal{C}\subseteq[0,1]^n$ is down-closed, if there exist a lower vector $\underline{\mathbf{u}}\in\mathcal{C}$ such that 1) $\mathbf{y}\geq\underline{\mathbf{u}}$ for any $\mathbf{y}\in\mathcal{C}$; 2) $\mathbf{x}\in\mathcal{C}$ if there exists a vector $\mathbf{y}\in\mathcal{C}$ satisfying $\underline{\mathbf{u}}\leq\mathbf{x}\leq\mathbf{y}$. Additionally, the radius $r(\mathcal{C})=\max_{\mathbf{x}\in\mathcal{C}}\|\mathbf{x}\|$ and the diameter $\mathrm{diam}(\mathcal{C})=\max_{\mathbf{x},\mathbf{y}\in\mathcal{C}}\|\mathbf{x}-\mathbf{y}\|$.

Table 1: Comparison of regrets for online non-monotone continuous DR-submodular function maximization over a down-closed convex set with stochastic gradient oracles. '**# Grad**' means the number of stochastic gradient evaluations at each round; '**Oracle**' indicates which type of online oracle used in algorithms; '**Feedback**' indicates full-information or bandit feedback scenario. '**ODC**' refers to the algorithm 2 in (Thang & Srivastav, 2021). '**Meta-MFW**', '**Mono-MFW**', and '**Bandit-MFW**' are our proposed algorithms. Note that all these algorithms can achieve $1/e$ approximation ratio.

Method	Regret	# Grad	Oracle	Feedback
ODC	$O(T^{3/4})$	$T^{3/4}$	nonconvex	full
Meta-MFW	$O(T^{3/4})$	$T^{3/4}$	linear	full
Meta-MFW	$O(T^{1/2})$	$T^{3/2}$	linear	full
Mono-MFW	$O(T^{4/5})$	1	linear	full
Bandit-MFW	$O(T^{8/9})$	0	linear	bandit

DR-Submodularity: A differentiable function $f:[0,1]^n\rightarrow\mathbb{R}_+$ is DR-submodular iff $\nabla f(\mathbf{x})\leq\nabla f(\mathbf{y})$ when $\mathbf{x}\geq\mathbf{y}$ (Bian et al., 2020).

Smoothness: A differentiable function f is called L_0-*smooth* if for any $\mathbf{x},\mathbf{y}\in[0,1]^n$, $\|\nabla f(\mathbf{x})-\nabla f(\mathbf{y})\|\leq L_0\|\mathbf{x}-\mathbf{y}\|$.

Problem Settings and α-regret: In this paper, we revisit the online non-monotone continuous DR-submodular maximization problem over a down-closed convex set $\mathcal{C}$. For a T-*round* game, after the learner chooses an action $\mathbf{x}_t\in\mathcal{C}$ at each round, the adversary reveals a DR-submodular function $f_t:[0,1]^n\rightarrow\mathbb{R}_+$ and feeds back the reward $f_t(\mathbf{x}_t)$ to the learner. The goal is to design efficient algorithms such that the gap between the accumulative reward and that of the best fixed policy in hindsight with scale parameter α, i.e., $\mathcal{R}_\alpha(T)=\alpha\max_{\mathbf{x}\in\mathcal{C}}\sum_{t=1}^{T}f_t(\mathbf{x})-\sum_{t=1}^{T}f_t(\mathbf{x}_t)$, is sublinear in horizon T. That is, $\lim_{T\rightarrow\infty}\mathcal{R}_\alpha(T)/T=0$. In this paper, we consider $\alpha=1/e$.

3 Algorithms and Main Results

3.1 Online Non-monotone Continuous DR-submodular maximization

In previous literature (Bian et al., 2017a; Mokhtari et al., 2018; Chen et al., 2018a,b), the monotone assumption of continuous DR-submodular function plays a core role in deriving the efficient competitive ratio. Therefore, we can not directly apply these algorithms to the non-monotone settings, due to lack of theoretical guarantee.

In this subsection, we will present a new online algorithm (Algorithm 1) for non-monotone continuous DR-submodular maximization over a down-closed convex set, which is inspired by the measured continuous greedy (Feldman et al., 2011; Mitra et al., 2021) for offline continuous DR-submodular maximization and the meta-action framework (Streeter & Golovin, 2008; Chen et al., 2018a) which utilizes the online linear optimization oracles (Hazan et al., 2016). Note that an online linear optimization oracle is an instance of the off-the-shelf online linear maximization algorithm that sequentially maximizes linear objectives.

In sharp contrast with the Meta-Frank-Wolfe (Chen et al., 2018a) for online monotone continuous DR-submodular maximization, in our Algorithm 1 we adopt a different update rule (line 6) and a novel feedback (line 11). Given a series of update directions $\mathbf{v}_t^{(k)} \in \mathcal{C}, \forall k \in [K]$ and initial point $\mathbf{x}_t^{(0)} = \mathbf{0}$, we consider

$$\mathbf{x}_t^{(k)} = \mathbf{x}_t^{(k-1)} + \frac{1}{K}\mathbf{v}_t^{(k)} \odot (\mathbf{1} - \mathbf{x}_t^{(k-1)}), \qquad (1)$$

where we re-weight the i-th element of $\mathbf{v}_t^{(k)}$ by $(\mathbf{1}-\mathbf{x}_t^{(k-1)})_i$ at each round and push the iteration point $\mathbf{x}_t^{(k-1)}$ along the weighted update direction $\mathbf{v}_t^{(k)} \odot (\mathbf{1}-\mathbf{x}_t^{(k-1)})$ with step size $\frac{1}{K}$. Due to the update rule of Equation (1), then Algorithm 1 feeds back the weighted gradient estimate $\mathbf{g}_t^{(k)} \odot (\mathbf{1} - \mathbf{x}_t^{(k)})$ for the linear oracle $\mathcal{E}^{(k)}$, where we view the vector $\mathbf{g}_t^{(k)}$ as an estimate for $\nabla f_t(\mathbf{x}_t^{(k)})$. Our update rule guarantees that $\mathbf{x}_t^{(k)} \in \mathcal{C}$ (proof in Appendix B).

Next, we demonstrate how the K different linear oracles work. Each linear oracle $\mathcal{E}^{(k)}$ in Algorithm 1 tries to online maximize the cumulative linear reward function $\sum_{t=1}^{T}\langle \mathbf{g}_t^{(k)} \odot (\mathbf{1} - \mathbf{x}_t^{(k)}), \cdot\rangle$. Precisely, after $\mathcal{E}^{(k)}$ commits to the action $\mathbf{v}_t^{(k)} \in \mathcal{C}$ at t-th round, Algorithm 1 feeds back the vector $\mathbf{g}_t^{(k)} \odot (\mathbf{1} - \mathbf{x}_t^{(k)})$ and the reward $\langle \mathbf{g}_t^{(k)} \odot (\mathbf{1} - \mathbf{x}_t^{(k)}), \mathbf{v}_t^{(k)}\rangle$ to the oracle $\mathcal{E}^{(k)}$; then the oracle $\mathcal{E}^{(k)}$ updates the action via some well-known strategies such as the online gradient ascent or regularized-follow-the-leader (Hazan et al., 2016). Taking the online gradient ascent as an example, the oracle $\mathcal{E}^{(k)}$ will choose the next action $\mathbf{v}_{t+1}^{(k)} = \arg\min_{\mathbf{v}\in\mathcal{C}} \|\mathbf{v} - (\mathbf{v}_t^{(k)} + \frac{1}{\sqrt{T}}\mathbf{g}_t^{(k)} \odot (\mathbf{1} - \mathbf{x}_t^{(k)}))\|$. Predictably, compared with the complicated online non-convex oracle of **ODC** (See online vee learning algorithm in (Thang & Srivastav, 2021)), the online linear oracle in the **Meta-MFW**, without discretization, lifting, or rounding operations, is simpler and more efficient.

We then make some assumptions for the regret analysis of Algorithm 1.

Assumption 1.

(i) The domain $\mathcal{C} \subseteq [0,1]^n$ is a down-closed convex set including the original point $\mathbf{0}$, where n is the dimensional parameter.

Algorithm 1 Meta-Measured Frank-Wolfe (**Meta-MFW**)

1: **Input:** K online linear maximization oracles over $\mathcal{C}$, i.e, $\mathcal{E}^{(1)}, \dots, \mathcal{E}^{(K)}$, η_k, $\mathbf{g}_t^{(0)} = \mathbf{x}_t^{(0)} = \mathbf{0}$.
2: **Output:** $\mathbf{y}_1, \dots, \mathbf{y}_T$.
3: **for** $t = 1, \dots, T$ **do**
4:   **for** $k = 1, \dots, K$ **do**
5:     Receive $\mathbf{v}_t^{(k)}$ which is the output of oracle $\mathcal{E}^{(k)}$.
6:     $\mathbf{x}_t^{(k)} = \mathbf{x}_t^{(k-1)} + \frac{1}{K}\mathbf{v}_t^{(k)} \odot (\mathbf{1} - \mathbf{x}_t^{(k-1)})$.
7:   **end for**
8:   Play $\mathbf{y}_t = \mathbf{x}_t^{(K)}$ for f_t to get reward $f_t(\mathbf{y}_t)$ and observe the stochastic gradient information of f_t.
9:   **for** $k = 1, \dots, K$ **do**
10:     $\mathbf{g}_t^{(k)} = (1 - \eta_k)\mathbf{g}_t^{(k-1)} + \eta_k\widetilde{\nabla} f_t(\mathbf{x}_t^{(k)})$ where $\mathbb{E}(\widetilde{\nabla} f_t(\mathbf{x}_t^{(k)})|\mathbf{x}_t^{(k)}) = \nabla f_t(\mathbf{x}_t^{(k)})$.
11:     Feed back $\langle \mathbf{g}_t^{(k)} \odot (\mathbf{1} - \mathbf{x}_t^{(k)}), \mathbf{v}_t^{(k)}\rangle$ as the payoff to be received by oracle $\mathcal{E}^{(k)}$.
12:   **end for**
13: **end for**

(ii) Each $f_t : [0,1]^n \to \mathbb{R}_+$ is a differentiable, DR-submodular function with smoothness parameter L_0.

(iii) For any linear maximization oracle $\mathcal{E}^{(k)}$, the regret at horizon t is at most $M_0\sqrt{t}$, where M_0 is a parameter.

Assumption 2. *For any $t \in [T]$ and $\mathbf{x} \in [0,1]^n$, there exists a stochastic gradient oracle $\widetilde{\nabla} f_t(\mathbf{x})$ with $\mathbb{E}(\widetilde{\nabla} f_t(\mathbf{x})|\mathbf{x}) = \nabla f_t(\mathbf{x})$ and $\mathbb{E}(\|\nabla f_t(\mathbf{x}) - \widetilde{\nabla} f_t(\mathbf{x})\|^2) \le \sigma^2$.*

Theorem 1. *[Proof in Appendix C] Under Assumption 1 and 2, if we set $\eta_k = \frac{2}{(k+3)^{2/3}}$ for any $k \in [K]$, we could verify that Algorithm 1 achieves:*

$$\frac{1}{e}\sum_{t=1}^{T} f_t(\mathbf{x}^*) - \sum_{t=1}^{T}\mathbb{E}(f_t(\mathbf{y}_t)) \le M_0\sqrt{T} + L_0 r^2(\mathcal{C})\frac{T}{2K} + \frac{\text{diam}(\mathcal{C})}{2}(3N_0+1)\frac{T}{K^{1/3}},$$

where $N_0 = \max\{4^{2/3}\max_{t\in[T]}\|\nabla f_t(\mathbf{x}_t^{(1)})\|^2, 4\sigma^2 + 6(L_0 r(\mathcal{C}))^2\}$ and $\mathbf{x}^* = \arg\max_{\mathbf{x}\in\mathcal{C}}\sum_{t=1}^{T} f_t(\mathbf{x})$.

Remark 1. *According to Theorem 1, if we set $K = T^{3/2}$, **Meta-MFW** yields the first result to achieve a $1/e$-regret of $O(\sqrt{T})$, which is faster than the previous outcome of **ODC** (Thang & Srivastav, 2021). It is worth mentioning that the $1/e$-regret of $O(\sqrt{T})$ not only achieves the best-known guarantee for the offline problem, but also matches the optimal $O(\sqrt{T})$ regret of online convex optimization (Hazan et al., 2016).*

Remark 2. *Meanwhile, when $K = T^{3/4}$, **Meta-MFW** achieves a $1/e$-regret of $O(T^{3/4})$, which has the same approximation ratio and regret as **ODC** (Thang & Srivastav, 2021). Although the oracle number $K = T^{3/4}$ of **Meta-MFW** is the same as **ODC**, **Meta-MFW** is more time-efficient than **ODC** since we adopt the simple online linear* *oracles while **ODC** utilizes complicated online non-convex oracles with discretization, lifting, and rounding operations.*

3.2 One-shot Online Non-monotone Continuous DR-submodular maximization

In many real-world scenarios, it could be time-consuming or even impossible to compute the stochastic gradient, e.g., influence maximization (Yang et al., 2016) as well as black-box attacks (Ito & Fujimaki, 2016). Thus, our Algorithm 1, which needs to inquire K gradient estimates for each reward function f_t (See line 10 in **Meta-MFW**), seems to be restrictive for many applications. To tackle the practical challenges, we hope to extend our proposed **Meta-MFW** into one-shot or bandit settings, where we only are permitted to inquire an unbiased gradient or one-point function value for each f_t, respectively. At first, we investigate the one-shot non-monotone DR-submodular maximization in this subsection.

We begin by reviewing the fairly known blocking technique in online learning (Zhang et al., 2019; Hazan et al., 2016). Specifically, we divide the T reward functions $f_1, \dots, f_T$ into Q blocks of the same size K, where $T = QK$, i.e., the q-th block includes the K different functions $f_{(q-1)K+1}, \dots, f_{qK}$. We also define the average function in the q-th block as $\bar{f}_q = \sum_{t=(q-1)K+1}^{qK} f_t/K$. To reduce the number of per-function stochastic gradient evaluations, the key idea is to view each $\bar{f}_q$ as a virtual reward function, such that the original T-round online optimization can be transferred into a new Q-round game. In this new Q-round game, at the q-th step, the algorithm first chooses an action $\mathbf{x}_q \in \mathcal{C}$, then the adversary reveals the reward $\bar{f}_q(\mathbf{x}_q)$ for the algorithm.

Since each $\bar{f}_q$ is also continuous DR-submodular, we could directly adopt Algorithm 1 to tackle the new Q-round game, which also requires inquiring K unbiased gradient estimates for each $\bar{f}_q$. Note that, in q-th block, there exist K different stochastic gradient oracles $\{\widetilde{\nabla} f_{(q-1)K+1}, \dots, \widetilde{\nabla} f_{qK}\}$. Moreover, for each random permutation $\{t_q^{(1)}, \dots, t_q^{(K)}\}$ of the indices $\{(q-1)K+1, \dots, qK\}$, it could be verified that the $\mathbb{E}(f_{t_q^{(k)}}(\mathbf{x})|\mathbf{x}) = \bar{f}_q(\mathbf{x})$ and $\mathbb{E}(\widetilde{\nabla} f_{t_q^{(k)}}(\mathbf{x})|\mathbf{x}) = \nabla \bar{f}_q(\mathbf{x})$. As a result, we can construct unbiased gradient estimates of $\bar{f}_q$ at K different points via the K existing oracles $\{\widetilde{\nabla} f_{(q-1)K+1}, \dots, \widetilde{\nabla} f_{qK}\}$, and each oracle inquires only one gradient evaluation. In this manner, we successfully reduce the number of per-function gradient evaluations from K to 1. Motivated via this high-level idea, we present a one-shot variant in Algorithm 2 (**Mono-MFW**). Note that in the q-th block, we play the same point $\mathbf{y}_t = \mathbf{x}_q^{(K)}$ for each objective function in $\{f_{(q-1)K+1}, \dots, f_{qK}\}$. We provide the regret analysis of Algorithm 2 in Theorem 2.

Theorem 2 (Proof in Appendix D). *Under Assumption 1-2 and $\max_{\mathbf{x}\in\mathcal{C}} \|\nabla f_t(\mathbf{x})\| \le G$ for any $t \in [T]$, if we set $\eta_k = \frac{2}{(k+3)^{2/3}}$, when $1 \le k \le \frac{K}{2}+1$, and $\eta_k = \frac{1.5}{(K-k+2)^{2/3}}$, when $\frac{K}{2}+2 \le k \le K$, then Algorithm 2 achieves:*

$$\frac{1}{e}\sum_{t=1}^{T} f_t(\mathbf{x}^*) - \sum_{t=1}^{T} \mathbb{E}(f_t(\mathbf{y}_t))$$
$$\le 2\mathrm{diam}(\mathcal{C})(N_1+1)QK^{2/3} + \frac{L_0 r^2(\mathcal{C})}{2}Q + M_0\sqrt{Q}K,$$

where $\mathbf{x}^* = \arg\max_{\mathbf{x}\in\mathcal{C}} \sum_{t=1}^{T} f_t(\mathbf{x})$ and $N_1 = \max\{5^{2/3}G^2, 8(\sigma^2+G^2) + 32(2G+L_0 r(\mathcal{C}))^2, 4.5(\sigma^2+G^2) + 7(2G+L_0 r(\mathcal{C}))^2/3\}$.

Algorithm 2 Mono-Measured Frank-Wolfe (**Mono-MFW**)

Input: K online linear maximization oracles over $\mathcal{C}$, i.e., $\mathcal{E}^{(1)}, \dots, \mathcal{E}^{(K)}$, $Q = \frac{T}{K}$, η_k, $\mathbf{g}_q^{(0)} = \mathbf{x}_q^{(0)} = \mathbf{0}$.
Output: $\mathbf{y}_1, \dots, \mathbf{y}_T$.
for $q = 1, \dots, Q$ **do**
  for $k = 1, \dots, K$ **do**
    Receive the update direction $\mathbf{v}_q^{(k)}$ which is the output of oracle $\mathcal{E}^{(k)}$.
    $\mathbf{x}_q^{(k)} = \mathbf{x}_q^{(k-1)} + \frac{1}{K}\mathbf{v}_q^{(k)} \odot (\mathbf{1} - \mathbf{x}_q^{(k-1)})$.
  end for
  Generate a random permutation $\{t_q^{(1)}, \dots, t_q^{(K)}\}$ for $\{(q-1)K+1, \dots, qK\}$.
  for $t = (q-1)K+1, \dots, qK$ **do**
    Play $\mathbf{y}_t = \mathbf{x}_q^{(K)}$ to get reward $f_t(\mathbf{y}_t)$ and observe the stochastic gradient information of f_t.
  end for
  for $k = 1, \dots, K$ **do**
    $\mathbf{g}_q^{(k)} = (1-\eta_k)\mathbf{g}_q^{(k-1)} + \eta_k \widetilde{\nabla} f_{t_q^{(k)}}(\mathbf{x}_q^{(k)})$.
    Feed back $\langle (\mathbf{1} - \mathbf{x}_q^{(k)}) \odot \mathbf{g}_q^{(k)}, \mathbf{v}_q^{(k)} \rangle$ as the payoff to be received by oracle $\mathcal{E}^{(k)}$.
  end for
end for

Remark 3. *According to Theorem 2, if we set $K = T^{3/5}$ and $Q = T^{2/5}$, the **Mono-MFW** achieves a $1/e$-regret of $O(T^{4/5})$. To the best of our knowledge, this is the first result with sublinear regret for one-shot online non-monotone DR-submodular maximization over a down-closed convex set.*

3.3 Bandit Online Non-monotone Continuous DR-submodular maximization

In this subsection, we turn to the bandit setting for online non-monotone continuous DR-submodular maximization. To begin, we review the one-point estimator (Flaxman et al., 2005), which is of great importance to our proposed bandit algorithm.

3.3.1 One-point Estimator

For any function $f : [0,1]^n \to \mathbb{R}_+$, define the δ-smooth version of f as $\hat{f}_\delta(\mathbf{x}) = \mathbb{E}_{\mathbf{v}\sim B^n}(f(\mathbf{x}+\delta\mathbf{v}))$ where $\mathbf{v} \sim$

B^n represents that the vector $\mathbf{v}$ is uniformly sampled from the n-dimensional unit ball B^n. If $\|\nabla f(\mathbf{x})\| \le G$, we have $|f(\mathbf{x}) - \hat{f}_\delta(\mathbf{x})| \le G\delta$. Thus, $\hat{f}_\delta$ can be viewed as an approximation of f, when δ is small. Roughly speaking, we can approximately maximize f via the maximizer of $\hat{f}_\delta$. Note that if f is continuous DR-submodular and L_0-smooth, so is $\hat{f}_\delta$. Moreover, according to Flaxman et al. (2005), $\nabla \hat{f}_\delta(\mathbf{x}) = \frac{n}{\delta}\mathbb{E}_{\mathbf{v}\sim S^{n-1}}(f(\mathbf{x}+\delta\mathbf{v})\mathbf{v})$ where $\mathbf{v} \sim S^{n-1}$ implies that the vector $\mathbf{v}$ is uniformly sampled from the unit sphere S^{n-1}, which sheds light on the possibility of estimating the gradient of $\hat{f}_\delta(\mathbf{x})$ via the function value at a random point $\mathbf{x}+\delta\mathbf{v}$.

However, we cannot use this estimate method directly. The point $\mathbf{x}+\delta\mathbf{v}$ may fall outside of the constraint set $\mathcal{C}$, when $\mathbf{x}$ is close to the boundary of $\mathcal{C}$. To tackle this challenge, we introduce the concept of δ-interior. We say that a subset $\mathcal{C}^{'}$ is a δ-interior of $\mathcal{C}$, if the ball $B(\mathbf{x},\delta)$ centered at $\mathbf{x}$ with radius δ, is included in $\mathcal{C}$ for any $\mathbf{x}\in\mathcal{C}^{'}$. As a result, for every point $\mathbf{x}\in\mathcal{C}^{'}$, $\mathbf{x}+\delta\mathbf{v}$ is included in $\mathcal{C}$, which enables us to use the one-point estimator. Recently, for a down-closed convex set $\mathcal{C}$, Zhang et al. (2019) provided a method to construct a δ-interior down-closed convex set $\mathcal{C}^{'}$. Next, we show this outcome in Lemma 1.

Lemma 1 (Zhang et al. (2019)). *Under Assumption 1, if there exists a positive number r such that $rB^n_{\ge 0}\subseteq\mathcal{C}$ where $B^n_{\ge 0} = B^n\cap\mathbb{R}^n_+$, and $\delta < \frac{r}{\sqrt{n}+1}$, the set $\mathcal{C}^{'} = (1-\alpha)\mathcal{C}+\delta\mathbf{1}$ is a down-closed convex δ-interior of $\mathcal{C}$ with $\sup_{\mathbf{x}\in\mathcal{C},\mathbf{y}\in\mathcal{C}^{'}}\|\mathbf{x}-\mathbf{y}\| \le ((\sqrt{n}+1)\frac{r(\mathcal{C})}{r}+\sqrt{n})\delta$, where $\alpha = \frac{(\sqrt{n}+1)\delta}{r}$.*

3.3.2 Bandit Measured Frank-Wolfe

To design an efficient algorithm in the bandit setting, a simple idea is to replace the stochastic gradient in Algorithm 2 with the one-point estimator and run it on the δ-interior $\mathcal{C}^{'}$. However, we cannot take this simple policy directly. In Algorithm 2, for each t in the q-th block, we play $\mathbf{x}_q^{(K)}$ for f_t, but we may require inquiring the gradient at a different point $\mathbf{x}_q^{(k)}$. Therefore, we could not construct the one-point gradient estimate at point $\mathbf{x}_q^{(k)}$ via the reward $f_t(\mathbf{x}_q^{(K)})$, when $k\neq K$.

To circumvent this drawback, we take the exploration-exploitation trade-off strategy in Zhang et al. (2019). Specifically, we divide the T reward functions into Q blocks of size L, where $T = LQ$. Then, we cut each block into two phases (i.e., exploration and exploitation). Taking the q-th block as an example, in the exploration phase, we select K random reward functions to play the $\mathbf{x}_q^{(k)}+\delta\mathbf{u}_q^{(k)}$ which provide the one-point gradient estimators. Then, in the exploitation phase, we commit to the point $\mathbf{x}_q^{(K)}$ for the remaining $(L-K)$ reward functions. Combining Algorithm 2 with this strategy, we present Algorithm 3 (**Bandit-MFW**). Moreover, we make an additional assumption and provide the regret bound of Algorithm 3.

Algorithm 3 Bandit-Measured Frank-Wolfe (**Bandit-MFW**)

Input: δ, r, $\alpha = \frac{(\sqrt{n}+1)\delta}{r}$, δ-interior down-closed convex set $\mathcal{C}^{'} = (1-\alpha)\mathcal{C}+\delta\mathbf{1}$, K online linear maximization oracles on $\mathcal{C}^{'}$, i.e., $\mathcal{E}^{(1)},\dots,\mathcal{E}^{(K)}$, L, $Q=\frac{T}{L}$, η_k, $\mathbf{g}_q^{(0)}=\mathbf{0}$, $\mathbf{x}_q^{(0)}=\delta\mathbf{1}$.
Output: $\mathbf{y}_1,\dots,\mathbf{y}_T$.
for $q=1,\dots,Q$ **do**
  for $k=1,\dots,K$ **do**
    Receive $\mathbf{v}_q^{(k)}$ which is the output of oracle $\mathcal{E}^{(k)}$.
    $\tilde{\mathbf{v}}_q^{(k)} = (\mathbf{v}_q^{(k)}-\delta\mathbf{1})\oslash(\mathbf{1}-\delta\mathbf{1})$.
    $\mathbf{x}_q^{(k)} = \mathbf{x}_q^{(k-1)}+\frac{1}{K}\tilde{\mathbf{v}}_q^{(k)}\odot(\mathbf{1}-\mathbf{x}_q^{(k-1)})$.
  end for
  Generate a random permutation $\{t_q^{(1)},\dots,t_q^{(L)}\}$ for $\{(q-1)L+1,\dots,qL\}$.
  for $t=(q-1)L+1,\dots,qL$ **do**
    if $t\in\{t_q^{(1)},\dots,t_q^{(K)}\}$ **then**
      Play $\mathbf{y}_t = \mathbf{x}_q^{(k)}+\delta\mathbf{u}_q^{(k)}$ for f_t, where $\mathbf{u}_q^{(k)}\sim S^{n-1}$. ▷ Exploration
    end if
    if $t\in\{(q-1)L+1,\dots,qL\}\setminus\{t_q^{(1)},\dots,t_q^{(K)}\}$ **then**
      Play $\mathbf{y}_t = \mathbf{x}_q^{(K)}$ for f_t. ▷ Exploitation
    end if
  end for
  for $k=1,\dots,K$ **do**
    $\mathbf{g}_q^{(k)} = (1-\eta_k)\mathbf{g}_q^{(k-1)}+\eta_k\frac{n}{\delta}f_{t_q^{(k)}}(\mathbf{x}_q^{(k)}+\delta\mathbf{u}_q^{(k)})\mathbf{u}_q^{(k)}$.
    $\tilde{\mathbf{x}}_q^{(k)} = (\mathbf{x}_q^{(k)}-\delta\mathbf{1})\oslash(\mathbf{1}-\delta\mathbf{1})$.
    Feed back $\langle(\mathbf{1}-\tilde{\mathbf{x}}_q^{(k)})\odot\mathbf{g}_q^{(k)},\mathbf{v}_q^{(k)}\rangle$ as the payoff to be received by oracle $\mathcal{E}^{(k)}$.
  end for
end for

Assumption 3.

(i) There exists a positive number r such that $rB^n_{\ge 0}\subseteq\mathcal{C}$ where $B^n_{\ge 0} = B^n\cup\mathbb{R}^n_+$.

(ii) For each $t\in[T]$, $\sup_{x\in\mathcal{C}} f_t(\mathbf{x})\le M_1$.

Theorem 3 (Proof in Appendix E). *Under Assumption 1, Assumption 3, and $\max_{\mathbf{x}\in\mathcal{C}}\|\nabla f_t(\mathbf{x})\|\le G$ for any $t\in[T]$, if we set $\eta_k = \frac{2}{(k+3)^{2/3}}$ for $k\in[K]$, then Algorithm 3 achieves:*

$$\frac{1}{e}\sum_{t=1}^{T} f_t(\mathbf{x}^*) - \sum_{t=1}^{T}\mathbb{E}(f_t(\mathbf{y}_t)) \le C_1\frac{LQ}{K} + M_0L\sqrt{Q} + \frac{C_2LQ}{2\delta K^{1/3}} + \frac{C_3\delta LQ}{2K^{1/3}} + 2M_1KQ + C_4T\delta,$$

where $\mathbf{x}^* = \arg\max_{\mathbf{x}\in\mathcal{C}}\sum_{t=1}^{T} f_t(\mathbf{x})$, $C_1 = \frac{L_0 r^2(\mathcal{C})}{2}$,

$C_2 = (8n^2M_1^2 + 1)\text{diam}(\mathcal{C})$, $C_3 = \max\{3^{2/3}G^2, 8G^2 + 3(4.5L_0r(\mathcal{C}) + 3G)^2/2\}\text{diam}(\mathcal{C})$ *and* $C_4 = ((\sqrt{n}+1)\frac{r(\mathcal{C})}{r} + \sqrt{n} + 2)G$.

Remark 4. *According to Theorem 3, if we set* $L = T^{7/9}$, $Q = T^{2/9}$, $K = T^{2/3}$, *and* $\delta = \frac{r}{(\sqrt{n}+2)T^{1/9}}$, ***Bandit-MFW*** *achieves a* $1/e$*-regret of* $O(T^{8/9})$ *. As far as we know, this is the first sublinear-regret online algorithm for continuous non-monotone DR-submodular maximization with bandit feedback.*

4 Empirical Evaluation

In this section, we compare the performance of the following algorithms in Python 3.9.7 with CVX optimization tool (Grant & Boyd, 2014) on a MacBook Pro with M1 chip with 16 GB RAM:

Meta-Measured Frank-Wolfe (β-Meta): In Algorithm 1, we set $K = T^{\beta}$ and $\eta_k = \frac{2}{(k+3)^{2/3}}$ for any $k \in [K]$. In the experiments, we consider $\beta = \frac{3}{4}$ or $\beta = \frac{3}{2}$.
Mono-Measured Frank-Wolfe (Mono): In Algorithm 2, we set $K = T^{3/5}$ and $Q = T^{2/5}$. Simultaneously, $\eta_k = \frac{2}{(k+3)^{2/3}}$ for any $1 \le k \le \frac{K}{2} + 1$ and $\eta_k = \frac{1.5}{(K-k+2)^{2/3}}$ for any $\frac{K}{2} + 2 \le k \le K$.
Bandit-Measured Frank-Wolfe (Bandit): In Algorithm 3, we set $L = T^{7/9}$, $Q = T^{2/9}$, $K = T^{2/3}$, $\delta = \frac{r}{(\sqrt{n}+2)T^{1/9}}$ as well as $\eta_k = \frac{2}{(k+3)^{2/3}}$ for any $k \in [K]$.
Online algorithm for down-closed convex sets (ODC): We consider Algorithm 2 in (Thang & Srivastav, 2021) where $L = T^{3/4}$ and $\rho_l = \frac{2}{(l+3)^{2/3}}$ for all $1 \le l \le L$.

4.1 Non-Convex/Non-Concave Quadratic Programming

We consider the quadratic objective $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{H}\mathbf{x} + \mathbf{h}^T\mathbf{x} + c$ and constraints $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}_+^n | \mathbf{A}\mathbf{x} \le \mathbf{b}, \mathbf{0} \le \mathbf{x} \le \mathbf{u}, \mathbf{A} \in \mathbb{R}_+^{m\times n}, \mathbf{b} \in \mathbb{R}_+^m\}$. Following Bian et al. (2017a,b); Chen et al. (2018b), we choose the matrix $\boldsymbol{H} \in \mathbb{R}^{n\times n}$ to be a randomly generated symmetric matrix with entries H_{ij} uniformly distributed in $[-10, 0]$, and the matrix $\boldsymbol{A}$ to be a random matrix with entries uniformly distributed in $[0, 1]$. It can be verified that f is a continuous DR-submodular function and $\mathcal{C}$ is down-closed. We set $\mathbf{b} = \mathbf{u} = \mathbf{1}$. Meanwhile, we set $\mathbf{h} = -0.1 * \mathbf{H}^T\mathbf{u}$, which ensures the non-monotone property. To make f non-negative, we choose $c = -0.5 * \sum_{i,j} H_{ij}$. We consider the Gaussian noise for gradient, i.e., $(\widetilde{\nabla} f_t(\boldsymbol{x}))_i = (\nabla f_t(\boldsymbol{x}))_i + \delta\mathcal{N}(0,1)$ for any $i \in [n]$ and $\mathbf{x} \in [0,1]^n$, where we set $\delta = 0.1$ in the experiments.

In our simulations, we first generate $T = 200$ reward functions $f_1, \dots, f_T$ with associated matrices $\mathbf{H}_1, \dots, \mathbf{H}_T$. Next, we run the well-studied offline algorithms (Bian et al., 2017a; Mitra et al., 2021) to produce an effective solution $\mathbf{x}_t^*$ that is a $(1/e)$-approximation to the optimum of the objective $\sum_{m=1}^{t} f_m$ for each $t \in [T]$. Then, under different n and m, we present the trend of the ratio between regret and horizon, namely, $(\sum_{m=1}^{t} f_m(\mathbf{x}_t^*) - \sum_{m=1}^{t} f_m(\mathbf{y}_m))/t$ in Figure 1(a)-1(c). Simultaneously, we report the 200-round running time in Table 2.

Table 2: Running time (in seconds)

Method \ (n,m)	$(25,15)$	$(40,20)$	$(50,50)$
ODC	14.14	26.38	45.08
3/2-Meta	471.99	609.50	895.97
3/4-Meta	8.75	11.54	16.61
Mono	0.16	0.21	0.31
Bandit	0.11	0.14	0.21

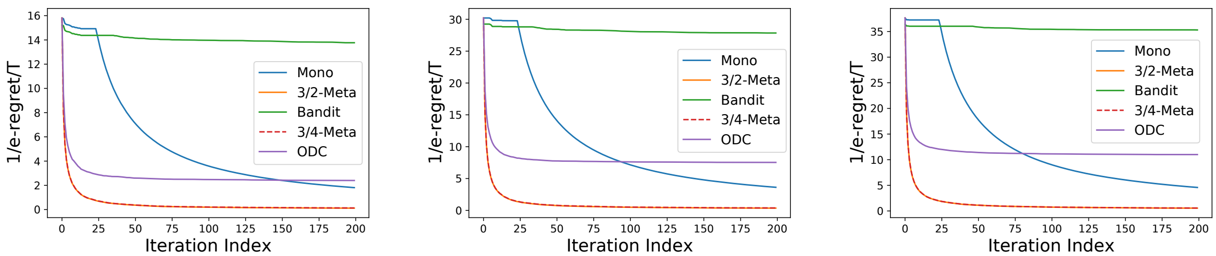
Table 3: Running time (in seconds)

Method	CA-HepPH	CA-GrQc	CA-HepTH
ODC	51.70	94.21	161.27
3/2-Meta	1302.50	1818.33	3156.41
3/4-Meta	24.97	35.26	59.56
Mono	0.52	0.64	1.08
Bandit	0.24	0.33	0.68

As shown in Figure 1, our proposed Meta-MFW with $\beta = 3/2$ and $3/4$ (i.e., 3/2-Meta and 3/4-Meta) achieve lower regret in contrast with ODC (Thang & Srivastav, 2021). Interestingly, the regret curves of both 3/2-Meta and 3/4-Meta are nearly the same in Figure 1. When the iteration index increases, Mono (Algorithm 2) also outperforms ODC in all three settings. Moreover, according to Table 2, 3/4-Meta and Mono effectively save running time compared with ODC. For example, when $n = 50, m = 50$, we spend 16.61 and 0.31 seconds in running 3/4-Meta and Mono, respectively, while the ODC takes 45.08 seconds. It is worth mentioning that although the bandit algorithm (Algorithm 3) with only one-point reward information exhibits the lowest convergence rate among all algorithms, it has the least running time as demonstrated in Table 2.

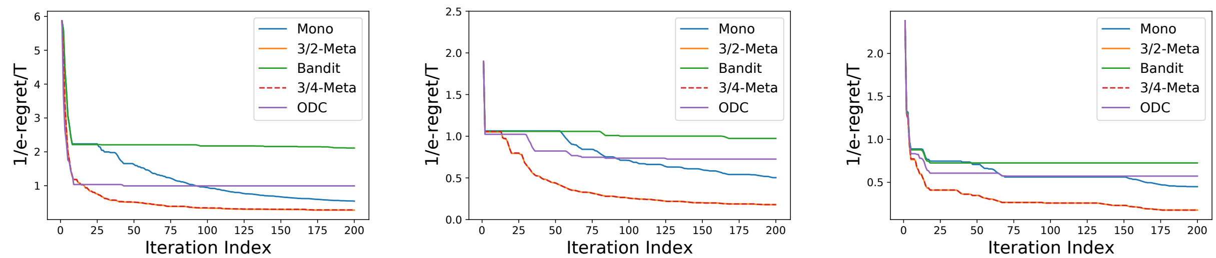
4.2 Revenue Maximization

In this application, we consider revenue maximization on an undirected social network $G = (V, W)$ where V is the set of nodes, and $w_{ij} \in W$ represents the weight of the edge between node i and node j. If we invest x proportion of the budget B on a user (node) $i \in V$, the user becomes an advocate of some product with probability $1 - (1-p)^{xB}$, where $p \in (0, 1)$ is a parameter. Intuitively, for investing a unit cost to user (node) i, we have an extra chance that the user i becomes an advocate with probability p. Let $S \subseteq V$ be a random set of users who advocate the product. Follow-



(a) $n = 25, m = 15$ (b) $n = 40, m = 20$ (c) $n = 50, m = 50$

Figure 1: We test the performance of the **3/2-Meta**, **3/4-Meta**, **Mono**, **Bandit**, and **ODC** in the simluated continuous DR-submodular quadratic programming under different dimension n and number of linear constraints m.



(a) CA-HepPH (b) CA-GrQc (c) CA-HepTH

Figure 2: We test the performance of the **3/2-Meta**, **3/4-Meta**, **Mono**, **Bandit**, and **ODC** in revenue maximization on social network **CA-HepPH**, **CA-GrQc** and **CA-HepTH**.

ing Thang & Srivastav (2021), the revenue with respect to S is defined as $\sum_{i\in S}\sum_{j\in V\setminus S} w_{ij}$. Let $f : [0, 1]^{|V|} \rightarrow \mathbb{R}_+$ be the expected revenue obtained in this model, that is $f(\mathbf{x}) = \sum_i \sum_{j\neq i} w_{ij}(1-(1-p)^{(\mathbf{x})_i B})(1-p)^{(\mathbf{x})_j B}$. It has been shown that f is a non-monotone continuous submodular function (Niazadeh et al., 2018; Soma & Yoshida, 2017).

In our experiments, we first sample three subgraphs from social networks (Leskovec et al., 2007) to simulate the online revenue maximization, i.e., a part of collaboration network of Arxiv High Energy Physics (CA-HepPH) with 316 edges and 56 vertices, a part of collaboration network of Arxiv General Relativity (CA-GrQc) with 316 edges and 81 vertices and a part of collaboration network of Arxiv High Energy Physics Theory (CA-HepTH) with 658 edges and 106 vertices. At each round $t \in [T]$, we randomly select 20 vertices $V_t \subseteq V$ and construct W_t with edge-weight $w_{ij}^t = 100$ if $i, j \in V_t$ and edge (i, j) exists in the network. Otherwise, $w_{ij}^t = 0$. As a result, the reward function $f_t(\mathbf{x}) = \sum_i \sum_{j\neq i} w_{ij}^t(1-(1-p)^{(\mathbf{x})_i B})(1-p)^{(\mathbf{x})_j B}$. We also impose a down-closed convex constraint as $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}_+^n | \mathbf{A}\mathbf{x} \leq \mathbf{b}, \sum_i(\mathbf{x})_i \leq 1, \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}\}$ where $\boldsymbol{A}$ is a random matrix with entries uniformly distributed in $[0, 1]$. We set $p = 0.002$, $m = 25$ as well as $B = 5$. Similarly, we consider the Gaussian noise for gradient with $\delta = 0.1$. Then, we report the trend of the ratio between regret and time horizon in Figure 2(a)-2(c) and running time in Table 3.

As shown in Figure 2, our proposed **Meta-MFW** with $\beta = 3/2$ and $3/4$ (i.e., 3/2-Meta and 3/4-Meta) have nearly the same curves and outperform the ODC (Thang & Srivastav, 2021). Similarly, compared to the **ODC**, the **Mono-MFW** (Algorithm 2) achieves lower regret in all three real-world social networks, when T is large. Moreover, according to Table 3, our proposed 3/4-Meta and Mono take less running time than the **ODC** algorithm. Note that the bandit algorithm (Algorithm 3) exhibits the lowest convergence rate among all algorithms with the fastest running time.

5 Conclusion

In this paper, we design three online no-regret algorithms for non-monotone continuous DR-submodular maximization over a down-closed convex set. The first one, **Meta-MFW**, attains a $1/e$-regret bound of $O(\sqrt{T})$ while requiring inquiring the $T^{3/2}$ amounts of gradient evaluations for each reward function. The second one, **Mono-MFW**, reduces the number of per-function gradient evaluations from $T^{3/2}$ to 1, and achieves a $1/e$-regret bound of $O(T^{4/5})$. Finally, we present the **Bandit-MFW** algorithm, which is the first bandit algorithm for online continuous non-monotone DR-submodular maximization over a down-closed convex set and achieves a $1/e$-regret bound of $O(T^{8/9})$. Numerical experiments demonstrate the superior performance of our algorithms.

Acknowledgements

Zhang and Yang's research is supported in part by the Hong Kong Research Grants Council (ECS 21214720), City University of Hong Kong (Project 9610465) and Alibaba Group through Alibaba Innovative Research (AIR) Program.

References

Auer, P., Cesa-Bianchi, N., Freund, Y., and Schapire, R. E. Gambling in a rigged casino: The adversarial multi-armed bandit problem. In *Proceedings of IEEE 36th annual foundations of computer science*, pp. 322–331. IEEE, 1995.

Bian, A., Levy, K., Krause, A., and Buhmann, J. M. Continuous dr-submodular maximization: Structure and algorithms. *Advances in Neural Information Processing Systems*, 30, 2017a.

Bian, A. A., Mirzasoleiman, B., Buhmann, J., and Krause, A. Guaranteed non-convex optimization: Submodular maximization over continuous domains. In *Artificial Intelligence and Statistics*, pp. 111–120. PMLR, 2017b.

Bian, Y., Buhmann, J., and Krause, A. Optimal continuous dr-submodular maximization and applications to provable mean field inference. In *International Conference on Machine Learning*, pp. 644–653. PMLR, 2019.

Bian, Y., Buhmann, J. M., and Krause, A. Continuous submodular function maximization. *arXiv preprint arXiv:2006.13474*, 2020.

Buchbinder, N. and Feldman, M. Deterministic algorithms for submodular maximization problems. *ACM Transactions on Algorithms (TALG)*, 14(3):1–20, 2018.

Buchbinder, N., Feldman, M., Seffi, J., and Schwartz, R. A tight linear time (1/2)-approximation for unconstrained submodular maximization. *SIAM Journal on Computing*, 44(5):1384–1402, 2015.

Chen, L., Harshaw, C., Hassani, H., and Karbasi, A. Projection-free online optimization with stochastic gradient: From convexity to submodularity. In *International Conference on Machine Learning*, pp. 814–823. PMLR, 2018a.

Chen, L., Hassani, H., and Karbasi, A. Online continuous submodular maximization. In *International Conference on Artificial Intelligence and Statistics*, pp. 1896–1905. PMLR, 2018b.

Chen, L., Zhang, M., Hassani, H., and Karbasi, A. Black box submodular maximization: Discrete and continuous settings. In *International Conference on Artificial Intelligence and Statistics*, pp. 1058–1070. PMLR, 2020.

Chen, P.-Y., Zhang, H., Sharma, Y., Yi, J., and Hsieh, C.-J. Zoo: Zeroth order optimization based black-box attacks to deep neural networks without training substitute models. In *Proceedings of the 10th ACM workshop on artificial intelligence and security*, pp. 15–26, 2017.

Feldman, M., Naor, J., and Schwartz, R. A unified continuous greedy algorithm for submodular maximization. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, pp. 570–579. IEEE, 2011.

Flaxman, A. D., Kalai, A. T., and McMahan, H. B. Online convex optimization in the bandit setting: gradient descent without a gradient. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pp. 385–394, 2005.

Grant, M. and Boyd, S. Cvx: Matlab software for disciplined convex programming, version 2.1, 2014.

Hassani, H., Soltanolkotabi, M., and Karbasi, A. Gradient methods for submodular maximization. In *Advances in Neural Information Processing Systems*, pp. 5841–5851, 2017.

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

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

Ilyas, A., Engstrom, L., Athalye, A., and Lin, J. Black-box adversarial attacks with limited queries and information. In *International Conference on Machine Learning*, pp. 2137–2146. PMLR, 2018.

Ito, S. and Fujimaki, R. Large-scale price optimization via network flow. *Advances in Neural Information Processing Systems*, 29, 2016.

Kempe, D., Kleinberg, J., and Tardos, É. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 137–146, 2003.

Kulesza, A., Taskar, B., et al. Determinantal point processes for machine learning. *Foundations and Trends® in Machine Learning*, 5(2–3):123–286, 2012.

Leskovec, J., Kleinberg, J., and Faloutsos, C. Graph evolution: Densification and shrinking diameters. *ACM transactions on Knowledge Discovery from Data (TKDD)*, 1 (1):2–es, 2007.

Mirrokni, V., Leme, R. P., Vladu, A., and Wong, S. C.-w. Tight bounds for approximate carathéodory and beyond. In *International Conference on Machine Learning*, pp. 2440–2448. PMLR, 2017.

Mitra, S., Feldman, M., and Karbasi, A. Submodular+ concave. In *Advances in Neural Information Processing Systems*, 2021.

Mokhtari, A., Hassani, H., and Karbasi, A. Conditional gradient method for stochastic submodular maximization: Closing the gap. In *International Conference on Artificial Intelligence and Statistics*, pp. 1886–1895. PMLR, 2018.

Mokhtari, A., Hassani, H., and Karbasi, A. Stochastic conditional gradient methods: From convex minimization to submodular maximization. *Journal of Machine Learning Research*, 2020.

Niazadeh, R., Roughgarden, T., and Wang, J. Optimal algorithms for continuous non-monotone submodular and dr-submodular maximization. *Advances in Neural Information Processing Systems*, 31, 2018.

Sadeghi, O. and Fazel, M. Online continuous dr-submodular maximization with long-term budget constraints. In *International Conference on Artificial Intelligence and Statistics*, pp. 4410–4419. PMLR, 2020.

Sadeghi, O., Raut, P., and Fazel, M. A single recipe for online submodular maximization with adversarial or stochastic constraints. *Advances in Neural Information Processing Systems*, 33:14712–14723, 2020.

Soma, T. and Yoshida, Y. Non-monotone dr-submodular function maximization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, 2017.

Streeter, M. and Golovin, D. An online algorithm for maximizing submodular functions. In *Advances in Neural Information Processing Systems*, pp. 1577–1584, 2008.

Thang, N. K. and Srivastav, A. Online non-monotone dr-submodular maximization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 9868–9876, 2021.

Vondrák, J. Symmetry and approximability of submodular maximization problems. *SIAM Journal on Computing*, 42(1):265–304, 2013.

Yang, Y., Mao, X., Pei, J., and He, X. Continuous influence maximization: What discounts should we offer to social network users? In *Proceedings of the 2016 International Conference on Management of Data*, pp. 727–741, 2016.

Zhang, M., Chen, L., Hassani, H., and Karbasi, A. Online continuous submodular maximization: From full-information to bandit feedback. In *Advances in Neural Information Processing Systems*, pp. 9206–9217, 2019.

Zhang, Q., Deng, Z., Chen, Z., Hu, H., and Yang, Y. Stochastic continuous submodular maximization: Boosting via non-oblivious function. In *International Conference on Machine Learning*, pp. 26116–26134. PMLR, 2022.

Zinkevich, M. Online convex programming and generalized infinitesimal gradient ascent. In *International Conference on Machine Learning*, pp. 928–936, 2003.

Algorithm 4 Online Algorithm using the doubling trick

Input:down-closed domain $\mathcal{C}$.
Output: $\mathbf{y}_1, \dots, \mathbf{y}_T$.
for $m = 0, 1, 2, \dots$ **do**
  Run base algorithm (e.g., Algorithm 1-3) with horizon 2^m from the 2^m-th iteration to the $(2^{m+1}-1)$-th iteration.
  We select the results $\mathbf{y}_{2^m}, \dots, \mathbf{y}_{2^{m+1}-1}$ of the previous base algorithm for the objectives $f_{2^m}, \dots, f_{2^{m+1}-1}$.
end for

A The Doubling Trick and Error of Soma & Yoshida (2017)

The authors would like to thank the anonymous reviewers for their helpful comments. In this section, we highlight two points during the peer review process

A minor drawback of our proposed algorithms (i.e., Algorithm 1-3) is that it requires the knowledge of the horizon T . This problem can be easily tackled via the doubling trick while preserving the order of the regret bound. The doubling trick was first proposed in Auer et al. (1995) and its key idea is to repeat the base algorithm with a doubling horizon. Algorithm 4 shows a online framework with the doubling trick.

Soma & Yoshida (2017) incorrectly assumed that the revenue maximization objective function is DR-submodular. Actually, it is continuous submodular (not DR-submodular), considering a counterexample $f(x_1, x_2) = (1 - q^{x_1})q^{x_2}$ where $q = 1 - p \in (0, 1)$ and $\frac{\partial^2 f}{\partial x_2^2} = (1 - q^{x_1})q^{x_2}(\log(q))^2 > 0$. Although our proposed algorithms achieve superior empirical performance in revenue maximization problems, it's unsure whether our methods can guarantee the same approximation ratio for the general continuous submodular maximization over down-closed sets.

B Variance Reduction Techniques

Our algorithms rely on the well-studied variance reduction techniques in (Chen et al., 2018a; Zhang et al., 2019; Mokhtari et al., 2020). Next, we demonstrate some results about variance reduction in the following lemmas.

Lemma 2 (Chen et al. (2018a); Mokhtari et al. (2020)). *Let $\{\mathbf{a}_t\}_{t=0}^K$ be a sequence of points in $\mathbb{R}^n$ such that $\|\mathbf{a}_t - \mathbf{a}_{t-1}\| \le \frac{G}{t+s}$ for all $1 \le t \le K$ with fixed constant $G \ge 0$ and $s \ge 3$. Let $\{\widetilde{\mathbf{a}}_t\}_{t=0}^K$ be a sequence of random variables such that $\mathbb{E}(\widetilde{\mathbf{a}}_t|\mathcal{F}_{t-1}) = \mathbf{a}_t$ and $\mathbb{E}(\|\widetilde{\mathbf{a}}_t - \mathbf{a}_t\|^2|\mathcal{F}_{t-1}) \le \sigma^2$ for every $t \ge 0$, where $\mathcal{F}_{t-1}$ is the σ-field generated by $\{\widetilde{\mathbf{a}}_k\}_{k=0}^{t-1}$ and $\mathcal{F}_0 = \emptyset$. Let $\{\mathbf{d}_t\}_{t=0}^K$ be a sequence of random variables where $\mathbf{d}_0$ is fixed and subsequent $\mathbf{d}_t$ are obtained by $\mathbf{d}_t = (1 - \eta_t)\mathbf{d}_{t-1} + \eta_t\widetilde{\mathbf{a}}_t$. If we set $\eta_t = \frac{2}{(t+s)^{2/3}}$, we have*

$$\mathbb{E}(\|\mathbf{d}_t - \mathbf{a}_t\|^2) \le \frac{N}{(t+s+1)^{2/3}}, \tag{2}$$

where $N = \max\{\|\mathbf{a}_0 - \mathbf{d}_0\|^2(s+1)^{2/3}, 4\sigma^2 + 3G^2/2\}$.

Lemma 3 (Zhang et al. (2019)). *Let $\{a_t\}_{t=0}^K$ be a sequence of points in $\mathbb{R}^n$ such that $\|\mathbf{a}_t - \mathbf{a}_{t-1}\| \le \frac{G}{K+2-t}$ for all $1 \le t \le K$ with fixed constant $G \ge 0$. Let $\{\widetilde{\mathbf{a}}_t\}_{t=0}^K$ be a sequence of random variables such that $\mathbb{E}(\widetilde{\mathbf{a}}_t|\mathcal{F}_{t-1}) = \mathbf{a}_t$ and $\mathbb{E}(\|\widetilde{\mathbf{a}}_t - \mathbf{a}_t\|^2|\mathcal{F}_{t-1}) \le \sigma^2$ for every $t \ge 0$, where $\mathcal{F}_{t-1}$ is the σ-field generated by $\{\widetilde{\mathbf{a}}_k\}_{k=0}^{t-1}$ and $\mathcal{F}_0 = \emptyset$. Let $\{\mathbf{d}_t\}_{t=0}^K$ be a sequence of random variables where $\mathbf{d}_0$ is fixed and subsequent $\mathbf{d}_t$ are obtained by $\mathbf{d}_t = (1 - \eta_t)\mathbf{d}_{t-1} + \eta_t\widetilde{\mathbf{a}}_t$. If we set $\eta_t = \frac{2}{(t+3)^{2/3}}$, when $1 \le t \le \frac{K}{2} + 1$, and when $\frac{K}{2} + 2 \le t \le K$, $\eta_t = \frac{1.5}{(K-t+2)^{2/3}}$, we have*

$$\mathbb{E}(\|\mathbf{d}_t - \mathbf{a}_t\|^2) \le \begin{cases} \dfrac{N}{(t+4)^{2/3}} & 1 \le t \le \dfrac{K}{2} + 1 \\ \dfrac{N}{(K-t+1)^{2/3}} & \dfrac{K}{2} + 2 \le t \le K \end{cases} \tag{3}$$

where $N = \max\{5^{2/3}\|\mathbf{a}_0 - \mathbf{d}_0\|^2, 4\sigma^2 + 32G^2, 2.25\sigma^2 + 7G^2/3\}$.

C Proofs in Section 3.1

C.1 The Properties of New Update Rule

In our Algorithm 1, we take a novel update rule (Equation (1) in Section 3.1). Before going into the detail, we first demonstrate some important properties of this new update rule. Next, we use new symbols to retell this update rule: Given a series of update directions $\mathbf{d}_k \in \mathcal{C}, \forall k \in [K]$ and initial point $\mathbf{y}_0 = \mathbf{0}$, we consider the following update rule, i.e.,

$$\mathbf{y}_k = \mathbf{y}_{k-1} + \frac{1}{K}\mathbf{d}_k \odot (\mathbf{1} - \mathbf{y}_{k-1}). \tag{4}$$

A prompt benefit of this rule is shown in the following lemma.

Lemma 4. *When $\mathcal{C} \subseteq [0,1]^n$ is down-closed convex set and $\mathbf{0} \in \mathcal{C}$, then $\mathbf{y}_k \in \mathcal{C}$ for any $k \in [K]$.*

Proof. First, we prove that $\mathbf{y}_k \leq \mathbf{1}$ for any $k \leq K$. By induction, we know $\mathbf{y}_0 = \mathbf{0}$. If we assume $\mathbf{y}_{k-1} \leq \mathbf{1}$, then

$$\begin{aligned}
\mathbf{y}_k &= \mathbf{y}_{k-1} + \frac{1}{K}\mathbf{d}_k \odot (\mathbf{1} - \mathbf{y}_{k-1}) \\
&= \frac{1}{K}\mathbf{d}_k + \mathbf{y}_{k-1} \odot (\mathbf{1} - \frac{1}{K}\mathbf{d}_k) \\
&\leq \frac{1}{K}\mathbf{d}_k + \mathbf{1} - \frac{1}{K}\mathbf{d}_k \\
&= \mathbf{1}.
\end{aligned}$$

As a result, $\mathbf{y}_k \leq \mathbf{1}$ for any $k \leq K$. Next, we verify that $\mathbf{y}_k \in \mathcal{C}$. According to Equation (4), we could conclude that $\mathbf{y}_K = \frac{1}{K}\sum_{k=1}^{K} \mathbf{d}_k \odot (\mathbf{1} - \mathbf{y}_{k-1})$. Due to convexity and each $\mathbf{d}_k \in \mathcal{C}$, we know $\frac{1}{K}\sum_{k=1}^{K} \mathbf{d}_k \in \mathcal{C}$. Also, we know that $\mathbf{0} \leq \mathbf{y}_1 \leq \mathbf{y}_2 \leq \dots \mathbf{y}_K \leq \frac{1}{K}\sum_{k=1}^{K} \mathbf{d}_k$ ($\mathbf{y}_k \leq \mathbf{1}$) so that $\mathbf{y}_k \in \mathcal{C}$ for any $k \in [K]$(the down-closed property). □

Moreover, we could derive a upper bound about every element of $\mathbf{y}_k$, i.e.,

Lemma 5. *For $i \in [n]$ and $k \in [K]$, we have $(\mathbf{y}_k)_i \leq 1 - (1 - \frac{1}{K})^k$.*

Proof. From Equation (4), we have

$$\begin{aligned}
(\mathbf{y}_k)_i &= (\mathbf{y}_{k-1})_i + \frac{1}{K}(\mathbf{d}_k \odot (\mathbf{1} - \mathbf{y}_{k-1}))_i \\
&= (\mathbf{y}_{k-1})_i + \frac{1}{K}(\mathbf{d}_k)_i * (1 - (\mathbf{y}_{k-1})_i) \\
&\leq (\mathbf{y}_{k-1})_i + \frac{1}{K}(1 - (\mathbf{y}_{k-1})_i) \\
&= (1 - \frac{1}{K})(\mathbf{y}_{k-1})_i + \frac{1}{K},
\end{aligned} \tag{5}$$

where the inequality follows from $(\mathbf{d}_k)_i \leq 1$ and $(\mathbf{y}_{k-1})_i \leq 1$.

First, we have $(\mathbf{y}_0)_i = 0 \leq 0$. If $(\mathbf{y}_k)_i \leq 1 - (1 - \frac{1}{K})^k$, we have

$$\begin{aligned}
(\mathbf{y}_k)_i &\leq (1 - \frac{1}{K})(\mathbf{y}_{k-1})_i + \frac{1}{K} \\
&\leq (1 - \frac{1}{K})(1 - (1 - \frac{1}{K})^k) + \frac{1}{K} \\
&= 1 - (1 - \frac{1}{K})^{k+1}.
\end{aligned} \tag{6}$$

Therefore, we have $(\mathbf{y}_k)_i \leq 1 - (1 - \frac{1}{K})^k$ by induction. □

Next, for any continuous DR-submodular function $f : [0,1]^n \to \mathbb{R}_+$, we show the relationship between $f(\mathbf{z})$ and $f(\mathbf{x})$ when the vector $\mathbf{z}$ take a similar form of the update rule (Equation (4)), namely, $\mathbf{z} = \mathbf{y} + (\mathbf{1} - \mathbf{y}) \odot \mathbf{x}$ where $\mathbf{x}, \mathbf{y} \in [0,1]^n$. Noticeably, $\mathbf{z} \geq \mathbf{x}$.

Lemma 6. *For any continuous DR-submodular function* $f : [0,1]^n \to \mathbb{R}_+$*, when* $\mathbf{z} = \mathbf{y} + (\mathbf{1}-\mathbf{y}) \odot \mathbf{x}$ *where* $\mathbf{x}, \mathbf{y} \in [0,1]^n$*, we have*

$$f(\mathbf{z}) \geq (1 - \|\mathbf{y}\|_\infty) f(\mathbf{x}).$$

Proof. First, we set $g(z) = f(\mathbf{x} + z(\mathbf{1}-\mathbf{x}) \odot \mathbf{y})$. Moreover, we know $\mathbf{x} + \frac{1}{\|\mathbf{y}\|_\infty}(\mathbf{1}-\mathbf{x}) \odot \mathbf{y} \in [0,1]^n$. According to (Bian et al., 2020; Thang & Srivastav, 2021), we know continuous DR-submodular function f is concave along the any positive direction. Therefore, g is a concave function in the interval $[0, \frac{1}{\|\mathbf{y}\|_\infty}]$. As a result, we have

$$\begin{aligned} f(\mathbf{y} + (\mathbf{1}-\mathbf{y}) \odot \mathbf{x}) &= f(\mathbf{x} + (\mathbf{1}-\mathbf{x}) \odot \mathbf{y}) \\ &= g(1) \\ &= g(\|\mathbf{y}\|_\infty * \frac{1}{\|\mathbf{y}\|_\infty} + (1 - \|\mathbf{y}\|_\infty) * 0) \\ &\geq (1 - \|\mathbf{y}\|_\infty) g(0) + \|\mathbf{y}\|_\infty g(\frac{1}{\|\mathbf{y}\|_\infty}) \\ &\geq (1 - \|\mathbf{y}\|_\infty) g(0) \\ &= (1 - \|\mathbf{y}\|_\infty) f(\mathbf{x}), \end{aligned} \tag{7}$$

where the first inequality comes from the concave property of g; the second from $g(\frac{1}{\|\mathbf{y}\|_\infty}) \geq 0$. ☐

Thus, according to Equation (4) and Lemma 5-6, we have $f(\mathbf{y}_k + (\mathbf{1}-\mathbf{y}_k) \odot \mathbf{y}^*) \geq (1 - \|\mathbf{y}_k\|_\infty) f(\mathbf{y}^*) \geq (1 - \frac{1}{K})^k f(\mathbf{y}^*)$ where $\mathbf{y}^* = \arg\max_{\mathbf{y} \in \mathcal{C}} f(\mathbf{y})$, which sheds light on the possibility to derive a constant-factor approximation for maximizing a non-monotone DR-submodular function for our proposed algorithms.

C.2 Proof of Theorem 1

First, we present a frequently used lemma.

Lemma 7. *For any continuous DR-submodular function* $f : [0,1]^n \to \mathbb{R}_+$ *with smoothness parameter* L_0*, if* $\mathbf{x}_k = \mathbf{x}_{k-1} + \frac{1}{K}\mathbf{v}_k \odot (\mathbf{1} - \mathbf{x}_{k-1})$ *for any* $0 \leq k \leq K$*, then we have, for* $\forall \mathbf{d} \in \mathbb{R}^n$ *and* $\forall \mathbf{y} \in [0,1]^n$*,*

$$\begin{aligned} f(\mathbf{x}_k) \geq &(1 - \frac{1}{K}) f(\mathbf{x}_{k-1}) + \frac{1}{K} f(\mathbf{x}_{k-1} + (\mathbf{1} - \mathbf{x}_{k-1}) \odot \mathbf{y}) + \frac{1}{K}\langle (\mathbf{1} - \mathbf{x}_{k-1}) \odot \mathbf{d}, \mathbf{v}_k - \mathbf{y} \rangle \\ &+ \frac{1}{K}\langle (\mathbf{v}_k - \mathbf{y}) \odot (\mathbf{1} - \mathbf{x}_{k-1}), \nabla f(\mathbf{x}_{k-1}) - \mathbf{d} \rangle - \frac{L_0}{2}\|\mathbf{x}_k - \mathbf{x}_{k-1}\|^2. \end{aligned} \tag{8}$$

Proof. According to the L_0-smooth condition, we have

$$\begin{aligned} f(\mathbf{x}_k) - f(\mathbf{x}_{k-1}) &\geq \langle \mathbf{x}_k - \mathbf{x}_{k-1}, \nabla f(\mathbf{x}_{k-1}) \rangle - \frac{L_0}{2}\|\mathbf{x}_k - \mathbf{x}_{k-1}\|^2 \\ &= \frac{1}{K}\langle \mathbf{v}_k \odot (\mathbf{1} - \mathbf{x}_{k-1}), \nabla f(\mathbf{x}_{k-1}) \rangle - \frac{L_0}{2}\|\mathbf{x}_k - \mathbf{x}_{k-1}\|^2. \end{aligned} \tag{9}$$

Then,

$$\begin{aligned} &\langle \mathbf{v}_k \odot (\mathbf{1} - \mathbf{x}_{k-1}), \nabla f(\mathbf{x}_{k-1}) \rangle \\ =&\langle \mathbf{v}_k \odot (\mathbf{1} - \mathbf{x}_{k-1}), \mathbf{d} \rangle + \langle \mathbf{v}_k \odot (\mathbf{1} - \mathbf{x}_{k-1}), \nabla f(\mathbf{x}_{k-1}) - \mathbf{d} \rangle \\ =&\langle (\mathbf{1} - \mathbf{x}_{k-1}) \odot \mathbf{d}, \mathbf{v}_k \rangle + \langle \mathbf{v}_k \odot (\mathbf{1} - \mathbf{x}_{k-1}), \nabla f(\mathbf{x}_{k-1}) - \mathbf{d} \rangle \\ =&\langle (\mathbf{1} - \mathbf{x}_{k-1}) \odot \mathbf{d}, \mathbf{y} \rangle + \langle (\mathbf{1} - \mathbf{x}_{k-1}) \odot \mathbf{d}, \mathbf{v}_k - \mathbf{y} \rangle + \langle \mathbf{v}_k \odot (\mathbf{1} - \mathbf{x}_{k-1}), \nabla f(\mathbf{x}_{k-1}) - \mathbf{d} \rangle \\ =&\langle (\mathbf{1} - \mathbf{x}_{k-1}) \odot \nabla f(\mathbf{x}_{k-1}), \mathbf{y} \rangle + \langle (\mathbf{1} - \mathbf{x}_{k-1}) \odot \mathbf{d}, \mathbf{v}_k - \mathbf{y} \rangle + \langle (\mathbf{v}_k - \mathbf{y}) \odot (\mathbf{1} - \mathbf{x}_{k-1}), \nabla f(\mathbf{x}_{k-1}) - \mathbf{d} \rangle \\ =&\langle \nabla f(\mathbf{x}_{k-1}), (\mathbf{1} - \mathbf{x}_{k-1}) \odot \mathbf{y} \rangle + \langle (\mathbf{1} - \mathbf{x}_{k-1}) \odot \mathbf{d}, \mathbf{v}_k - \mathbf{y} \rangle + \langle (\mathbf{v}_k - \mathbf{y}) \odot (\mathbf{1} - \mathbf{x}_{k-1}), \nabla f(\mathbf{x}_{k-1}) - \mathbf{d} \rangle. \end{aligned} \tag{10}$$

For DR-submodular function f, we also have

$$\langle \nabla f(\mathbf{x}_{k-1}), (\mathbf{1} - \mathbf{x}_{k-1}) \odot \mathbf{y} \rangle \geq f(\mathbf{x}_{k-1} + (\mathbf{1} - \mathbf{x}_{k-1}) \odot \mathbf{y}) - f(\mathbf{x}_{k-1}), \tag{11}$$

because f is concave along the direction $(\mathbf{1}-\mathbf{x}_{k-1})\odot\mathbf{y}$ and $\mathbf{x}_{k-1}+(\mathbf{1}-\mathbf{x}_{k-1})\odot\mathbf{y}\in[0,1]^n$. Finally, we have

$$\begin{aligned}f(\mathbf{x}_k)\geq&(1-\frac{1}{K})f(\mathbf{x}_{k-1})+\frac{1}{K}f(\mathbf{x}_{k-1}+(\mathbf{1}-x_{k-1})\odot\mathbf{y})+\frac{1}{K}\langle(\mathbf{1}-\mathbf{x}_{k-1})\odot\mathbf{d},\mathbf{v}_k-\mathbf{y}\rangle\\&+\frac{1}{K}\langle(\mathbf{v}_k-\mathbf{y})\odot(\mathbf{1}-\mathbf{x}_{k-1}),\nabla f(\mathbf{x}_{k-1})-\mathbf{d}\rangle-\frac{L_0}{2}\|\mathbf{x}_k-\mathbf{x}_{k-1}\|^2.\end{aligned}\tag{12}$$

□

Then, we show how $\mathbf{g}_t^{(k)}$ (See Line 10 in Algorithm 1) approximates the gradient $\nabla f_t(\mathbf{x}_t^{(k)})$.

Lemma 8. *Under Assumption 1 and Assumption 2, if we set $\eta_k=\frac{2}{(k+3)^{2/3}}$ for any $k\in[K]$, then we have, for any fixed $t\in[T]$,*

$$\mathbb{E}(\|\mathbf{g}_t^{(k)}-\nabla f_t(\mathbf{x}_t^{k})\|^2)\leq\frac{N_0}{(k+4)^{2/3}},\tag{13}$$

where $N_0=\max\{4^{2/3}\max_{t\in[T]}\|\nabla f_t(\mathbf{x}_t^{(1)})\|^2,4\sigma^2+6(L_0r(\mathcal{C}))^2\}$.

Proof. According to Algorithm 1, $\mathbf{g}_t^{(k)}=(1-\eta_k)\mathbf{g}_t^{(k-1)}+\eta_k\widetilde{\nabla}f_t(\mathbf{x}_t^{(k)})$ where $\mathbb{E}(\widetilde{\nabla}f_t(\mathbf{x}_t^{(k)})|\mathbf{x}_t^{(k)})=\nabla f_t(\mathbf{x}_t^{(k)})$. We first derive that

$$\|\nabla f_t(\mathbf{x}_t^{(k)})-\nabla f_t(\mathbf{x}_t^{(k-1)})\|\leq\frac{L_0}{K}\|\mathbf{v}_t^{(k)}\|\leq\frac{2L_0r(\mathcal{C})}{k+3},\tag{14}$$

where the first inequality follows from the L_0-smoothness of f_t Therefore, if we set the $\widetilde{\mathbf{a}}_t$ in Lemma 2 as $\widetilde{\nabla}f_t(\mathbf{x}_t^{(k)})$, we have

$$\mathbb{E}(\|\mathbf{g}_t^{(k)}-\nabla f_t(\mathbf{x}_t^{(k)})\|^2)\leq\frac{N_0}{(k+4)^{2/3}}.\tag{15}$$

□

Now, we present the proof of Theorem 1.

Proof. If we set $(f,\mathbf{x}_k,\mathbf{x}_{k-1},\mathbf{v}_k,\mathbf{d},\mathbf{y})$ in Lemma 7 as $(f_t,\mathbf{x}_t^{(k+1)},\mathbf{x}_t^{(k)},\mathbf{v}_t^{(k+1)},\mathbf{g}_t^{(k)},\mathbf{x}^*)$ in the Algorithm 1 where $\mathbf{x}^*=\arg\max_{\mathbf{x}\in\mathcal{C}}\sum_{t=1}^{T}f_t(\mathbf{x})$, we have, for any $k\in[K]$,

$$\begin{aligned}&f_t(\mathbf{x}_t^{(k+1)})\\\geq&(1-\frac{1}{K})f_t(\mathbf{x}_t^{(k)})+\frac{1}{K}f_t(\mathbf{x}_t^{(k)}+(\mathbf{1}-\mathbf{x}_t^{(k)})\odot\mathbf{x}^*)+\frac{1}{K}\langle(\mathbf{1}-\mathbf{x}^{(k)})\odot\mathbf{g}_t^{(k)},\mathbf{v}_t^{(k+1)}-\mathbf{x}^*\rangle\\&+\frac{1}{K}\langle(\mathbf{v}_t^{(k+1)}-\mathbf{x}^*)\odot(\mathbf{1}-\mathbf{x}_t^{(k)}),\nabla f(\mathbf{x}_{k-1})-\mathbf{g}_t^{(k)}\rangle-\frac{L_0}{2}\|\mathbf{x}_t^{(k+1)}-\mathbf{x}_t^{(k)}\|^2\\\geq&(1-\frac{1}{K})f_t(\mathbf{x}_t^{(k)})+\frac{1}{K}(1-\frac{1}{K})^kf_t(\mathbf{x}^*)+\frac{1}{K}\langle(\mathbf{1}-\mathbf{x}^{(k)})\odot\mathbf{g}_t^{(k)},\mathbf{v}_t^{(k+1)}-\mathbf{x}^*\rangle\\&+\frac{1}{K}\langle(\mathbf{v}_t^{(k+1)}-\mathbf{x}^*)\odot(\mathbf{1}-\mathbf{x}_t^{(k)}),\nabla f(\mathbf{x}_{k-1})-\mathbf{g}_t^{(k)}\rangle-\frac{L_0}{2}\|\mathbf{x}_t^{(k+1)}-\mathbf{x}_t^{(k)}\|^2,\end{aligned}\tag{16}$$

where the second inequality comes from Lemma 5 and Lemma 6. Therefore, by iteration, we have

$$\begin{aligned}
&f_t(\mathbf{x}_t^{(K)})\\
\geq&(1-\frac{1}{K})f_t(\mathbf{x}_t^{(K-1)})+\frac{1}{K}(1-\frac{1}{K})^{K-1}f_t(\mathbf{x}^*)+\frac{1}{K}\langle(\mathbf{1}-\mathbf{x}^{(K-1)})\odot\mathbf{g}_t^{(K-1)},\mathbf{v}_t^{(K)}-\mathbf{x}^*\rangle\\
&+\frac{1}{K}\langle(\mathbf{v}_t^{(K)}-\mathbf{x}^*)\odot(\mathbf{1}-\mathbf{x}_t^{(K-1)}),\nabla f(\mathbf{x}_t^{(K-1)})-\mathbf{g}_t^{(K-1)}\rangle-\frac{L_0r^2(\mathcal{C})}{2K^2}\\
\geq&\dots\\
\geq&(1-\frac{1}{K})f_t(\mathbf{x}_t^{(0)})+(1-\frac{1}{K})^{K-1}f_t(\mathbf{x}^*)+\frac{1}{K}\sum_{m=0}^{K-1}(1-\frac{1}{K})^{K-1-m}\langle(\mathbf{1}-\mathbf{x}^{(m)})\odot\mathbf{g}_t^{(m)},\mathbf{v}_t^{(m+1)}-\mathbf{x}^*\rangle\\
&+\frac{1}{K}\sum_{m=0}^{K-1}(1-\frac{1}{K})^{K-1-m}\langle(\mathbf{v}_t^{(m+1)}-\mathbf{x}^*)\odot(\mathbf{1}-\mathbf{x}_t^{(m)}),\nabla f(\mathbf{x}_t^{(m)})-\mathbf{g}_t^{(m)}\rangle-\frac{L_0r^2(\mathcal{C})}{2K}\\
\geq&(1-\frac{1}{K})f_t(\mathbf{x}_t^{(0)})+\frac{1}{e}f_t(\mathbf{x}^*)+\frac{1}{K}\sum_{m=0}^{K-1}(1-\frac{1}{K})^{K-1-m}\langle(\mathbf{1}-\mathbf{x}^{(m)})\odot\mathbf{g}_t^{(m)},\mathbf{v}_t^{(m+1)}-\mathbf{x}^*\rangle\\
&+\frac{1}{K}\sum_{m=0}^{K-1}(1-\frac{1}{K})^{K-1-m}\langle(\mathbf{v}_t^{(m+1)}-\mathbf{x}^*)\odot(\mathbf{1}-\mathbf{x}_t^{(m)}),\nabla f(\mathbf{x}_t^{(m)})-\mathbf{g}_t^{(m)}\rangle-\frac{L_0r^2(\mathcal{C})}{2K},
\end{aligned}\tag{17}$$

where the final inequality comes from $(1-\frac{1}{K})^{K-1}\geq\frac{1}{e}$ for any $K\geq 2$.

Finally,

$$\begin{aligned}
&\sum_{t=1}^{T}\mathbb{E}(f_t(\mathbf{x}_t^{(K)}))\\
\geq&\frac{1}{e}\sum_{t=1}^{T}f_t(\mathbf{x}^*)+\frac{1}{K}\sum_{t=1}^{T}\sum_{m=0}^{K-1}(1-\frac{1}{K})^{K-1-m}\mathbb{E}(\langle(\mathbf{1}-\mathbf{x}^{(m)})\odot\mathbf{g}_t^{(m)},\mathbf{v}_t^{(m+1)}-\mathbf{x}^*\rangle)\\
&+\frac{1}{K}\sum_{t=1}^{T}\sum_{m=0}^{K-1}(1-\frac{1}{K})^{K-1-m}\mathbb{E}\langle(\mathbf{v}_t^{(m+1)}-\mathbf{x}^*)\odot(\mathbf{1}-\mathbf{x}_t^{(m)}),\nabla f(\mathbf{x}_t^{(m)})-\mathbf{g}_t^{(m)}\rangle)-\frac{L_0Tr^2(\mathcal{C})}{2K}\\
=&\frac{1}{e}\sum_{t=1}^{T}f_t(\mathbf{x}^*)+\frac{1}{K}\sum_{m=0}^{K-1}(1-\frac{1}{K})^{K-1-m}\sum_{t=1}^{T}\mathbb{E}(\langle(\mathbf{1}-\mathbf{x}^{(m)})\odot\mathbf{g}_t^{(m)},\mathbf{v}_t^{(m+1)}-\mathbf{x}^*\rangle)\\
&+\frac{1}{K}\sum_{t=1}^{T}\sum_{m=0}^{K-1}(1-\frac{1}{K})^{K-1-m}\mathbb{E}(\langle(\mathbf{v}_t^{(m+1)}-\mathbf{x}^*)\odot(\mathbf{1}-\mathbf{x}_t^{(m)}),\nabla f(\mathbf{x}_t^{(m)})-\mathbf{g}_t^{(m)}\rangle)-\frac{L_0Tr^2(\mathcal{C})}{2K}\\
\geq&\frac{1}{e}\sum_{t=1}^{T}f_t(\mathbf{x}^*)-\frac{1}{K}\sum_{m=0}^{K-1}(1-\frac{1}{K})^{K-1-m}M_0\sqrt{T}-\frac{L_0Tr^2(\mathcal{C})}{2K}\\
&-\frac{1}{2K}\sum_{t=1}^{T}\sum_{m=0}^{K-1}(\mathrm{diam}^2(\mathcal{C})b+\mathbb{E}(\frac{\|\nabla f(\mathbf{x}_t^{(m)})-\mathbf{g}_t^{(m)}\|^2}{b}))\\
\geq&\frac{1}{e}\sum_{t=1}^{T}f_t(\mathbf{x}^*)-M_0\sqrt{T}-\frac{L_0Tr^2(\mathcal{C})}{2K}-\mathrm{diam}(\mathcal{C})(\frac{3}{2}N_0+\frac{1}{2})\frac{T}{K^{1/3}},
\end{aligned}\tag{18}$$

where the second inequality follows from $\langle(\mathbf{v}_t^{(m+1)}-\mathbf{x}^*)\odot(\mathbf{1}-\mathbf{x}_t^{(m)}),\nabla f(\mathbf{x}_t^{(m)})-\mathbf{g}_t^{(m)}\rangle\leq\frac{1}{2}(\mathrm{diam}^2(\mathcal{C})b+\frac{\|\nabla f(\mathbf{x}_t^{(m)})-\mathbf{g}_t^{(m)}\|^2}{b})$ for any $b>0$ and $\sum_{t=1}^{T}\mathbb{E}(\langle(\mathbf{1}-\mathbf{x}^{(m)})\odot\mathbf{g}_t^{(m)},\mathbf{v}_t^{(m+1)}-\mathbf{x}^*\rangle)\leq M_0\sqrt{T}$ from Assumption 1; and the last inequality comes from $\sum_{m=0}^{K-1}(1-\frac{1}{K})^{K-1-m}\leq K$, $\sum_{m=0}^{K-1}\mathbb{E}(\|\nabla f(\mathbf{x}_t^{(m)})-\mathbf{g}_t^{(m)}\|^2)\leq\sum_{m=0}^{K-1}\frac{N_0}{(m+4)^{2/3}}\leq\int_{t=0}^{K}\frac{N_0}{t^{2/3}}\leq 3N_0K^{1/3}$ and setting $b=\frac{1}{\mathrm{diam}(\mathcal{C})K^{1/3}}$. □

D Proofs in Section 3.2

In this section, we begin by deriving the upper bound of $\mathbf{x}_q^{(k)}$ in Algorithm 2. First, like the Equation (4), we also take a similar rule to update the $\mathbf{x}_q^{(k)}$. As a result, we have:

Lemma 9. *For $i\in[n]$ and $q\in[Q]$, we have $(\mathbf{x}_q^{(k)})_i\le 1-(1-\frac{1}{K})^k$.*

Before going into the detail, we define the average function of the remaining $(K-k)$ functions as $\bar{f}_{q,k}=\frac{\sum_{m=k+1}^{K}f_{t_q^{(m)}}}{K-k}$ for any $0\le k\le K-1$. Also, we use $\mathcal{F}_{q,k}$ to denote the σ-field generated via $t_q^{(1)},\dots,t_q^{(k)}$. As a result, according to Lemma 3 in variance reduction section, we show how $\mathbf{g}_q^{(k)}$ (See Line 13 in Algorithm 2) approximates the gradient $\nabla\bar{f}_{q,k-1}(\mathbf{x}_q^k)$, i.e.,

Lemma 10. *Under Assumption 1-2 and $\|\nabla f_t(\mathbf{x})\|\le G$, if we set $\eta_k=\frac{2}{(k+3)^{2/3}}$, when $1\le k\le\frac{K}{2}+1$, and $\eta_k=\frac{1.5}{(K-k+2)^{2/3}}$, when $\frac{K}{2}+2\le k\le K$, we have, for any fixed $q\in[Q]$,*

$$\mathbb{E}(\|\mathbf{g}_q^{(k)}-\nabla\bar{f}_{q,k-1}(\mathbf{x}_q^k)\|^2)\le\begin{cases}\dfrac{N_1}{(k+4)^{2/3}}, & 1\le k\le\dfrac{K}{2}+1\\ \dfrac{N_1}{(K-k+1)^{2/3}}, & \dfrac{K}{2}+2\le k\le K\end{cases}\tag{19}$$

where $N_1=\max\{5^{2/3}G^2, 8(\sigma^2+G^2)+32(2G+L_0r(\mathcal{C}))^2, 4.5(\sigma^2+G^2)+7(2G+L_0r(\mathcal{C}))^2/3\}$.

Proof. From Algorithm 2, $\mathbf{g}_q^{(k)}=(1-\eta_k)\mathbf{g}_q^{(k-1)}+\eta_k\widetilde{\nabla}f_{t_q^{(k)}}(\mathbf{x}_q^{(k)})$. As we know, $\mathbb{E}(\widetilde{\nabla}f_{t_q^{(k)}}(\mathbf{x}_q^{(k)})|\mathcal{F}_{q,k-1})=\nabla\bar{f}_{q,k-1}(\mathbf{x}_q^{(k)})$. Also, we have

$$\begin{aligned}&\nabla\bar{f}_{q,k-1}(\mathbf{x}_q^{(k)})-\nabla\bar{f}_{q,k-2}(\mathbf{x}_q^{(k-1)})\\ =&\frac{\sum_{m=k}^{K}\nabla f_{t_q^{(m)}}(\mathbf{x}_q^{(k)})}{K-k+1}-\frac{\sum_{m=k-1}^{K}\nabla f_{t_q^{(m)}}(\mathbf{x}_q^{(k-1)})}{K-k+2}\\ =&\frac{\sum_{m=k}^{K}(\nabla f_{t_q^{(m)}}(\mathbf{x}_q^{(k)})-\nabla f_{t_q^{(m)}}(\mathbf{x}_q^{(k-1)}))}{K-k+2}+\frac{\sum_{m=k}^{K}\nabla f_{t_q^{(m)}}(\mathbf{x}_q^{(k)})}{(K-k+1)(K-k+2)}-\frac{\nabla f_{t_q^{(k-1)}}(\mathbf{x}_q^{(k-1)})}{K-k+2}.\end{aligned}\tag{20}$$

Thus,

$$\begin{aligned}&\|\nabla\bar{f}_{q,k-1}(\mathbf{x}_q^{(k)})-\nabla\bar{f}_{q,k-2}(\mathbf{x}_q^{(k-1)})\|\\ \le&\|\frac{\sum_{m=k}^{K}(\nabla f_{t_q^{(m)}}(\mathbf{x}_q^{(k)})-\nabla f_{t_q^{(m)}}(\mathbf{x}_q^{(k-1)}))}{K-k+2}\|+\|\frac{\sum_{m=k}^{K}\nabla f_{t_q^{(m)}}(\mathbf{x}_q^{(k)})}{(K-k+1)(K-k+2)}\|+\|\frac{\nabla f_{t_q^{(k-1)}}(\mathbf{x}_q^{(k-1)})}{K-k+2}\|\\ \le&\frac{(K-k+1)L_0r(\mathcal{C})}{K(K-k+2)}+\frac{G}{K-k+2}+\frac{G}{K-k+2}\\ \le&\frac{L_0r(\mathcal{C})+2G}{K-k+2}.\end{aligned}\tag{21}$$

Moreover,

$$\begin{aligned}&\mathbb{E}(\|\widetilde{\nabla}f_{t_q^{(k)}}(\mathbf{x}_q^{(k)})-\nabla\bar{f}_{q,k-1}(\mathbf{x}_q^{(k)})\|^2|\mathcal{F}_{q,k-1})\\ \le&2(\mathbb{E}(\|\widetilde{\nabla}f_{t_q^{(k)}}(\mathbf{x}_q^{(k)})-\nabla f_{t_q^{(k)}}(\mathbf{x}_q^{(k)})\|^2|\mathcal{F}_{q,k-1})+\mathbb{E}(\|\nabla f_{t_q^{(k)}}(\mathbf{x}_q^{(k)})-\nabla\bar{f}_{q,k-1}(\mathbf{x}_q^{(k)})\|^2|\mathcal{F}_{q,k-1}))\\ =&2(\mathbb{E}(\|\widetilde{\nabla}f_{t_q^{(k)}}(\mathbf{x}_q^{(k)})-\nabla f_{t_q^{(k)}}(\mathbf{x}_q^{(k)})\|^2|\mathcal{F}_{q,k-1})+Var(\nabla f_{t_q^{(k)}}(\mathbf{x}_q^{(k)})|\mathcal{F}_{q,k-1}))\\ \le&2(\sigma^2+G^2),\end{aligned}\tag{22}$$

where $Var(\nabla f_{t_q^{(k)}}(\mathbf{x}_q^{(k)})|\mathcal{F}_{q,k-1})=\mathbb{E}(\|\nabla f_{t_q^{(k)}}(\mathbf{x}_q^{(k)})-\nabla\bar{f}_{q,k-1}(\mathbf{x}_q^{(k)})\|^2|\mathcal{F}_{q,k-1})$.

According to Lemma 3 where we set $\widetilde{\mathbf{a}}_k = \widetilde{\nabla} f_{t_q^{(k)}}(\mathbf{x}_q^{(k)})$, we have

$$\mathbb{E}(\|\mathbf{g}_q^{(k)} - \nabla \bar{f}_{q,k-1}(\mathbf{x}_q^k)\|^2) \leq \begin{cases} \frac{N_1}{(k+4)^{2/3}} & 1 \leq k \leq \frac{K}{2}+1 \\ \frac{N_1}{(K-k+1)^{2/3}} & \frac{K}{2}+2 \leq k \leq K \end{cases} \tag{23}$$

where $N_1 = \max\{5^{2/3}G^2, 8(\sigma^2+G^2)+32(2G+L_0 r(\mathcal{C}))^2, 4.5(\sigma^2+G^2)+7(2G+L_0 r(\mathcal{C}))^2/3\}$. □

Now, we prove Theorem 2.

Proof. Note that $\bar{f}_{q,k-1}$ is continuous DR-submodular and L_0-smooth. Thus, if we set $(f, \mathbf{x}_k, \mathbf{x}_{k-1}, \mathbf{v}_k, \mathbf{d}, \mathbf{y})$ in Lemma 7 as $(\bar{f}_{q,k-1}, \mathbf{x}_q^{(k+1)}, \mathbf{x}_q^{(k)}, \mathbf{v}_q^{(k+1)}, \mathbf{g}_q^{(k)}, \mathbf{x}^*)$ in the Algorithm 2 where $\mathbf{x}^* = \arg\max_{\mathbf{x}\in\mathcal{C}} \sum_{t=1}^{T} f_t(\mathbf{x})$, we have, for any $k \in [K]$,

$$\begin{aligned}
&\bar{f}_{q,k-1}(\mathbf{x}_q^{(k+1)}) \\
\geq&(1-\frac{1}{K})\bar{f}_{q,k-1}(\mathbf{x}_q^{(k)}) + \frac{1}{K}\bar{f}_{q,k-1}(\mathbf{x}_q^{(k)} + (\mathbf{1}-\mathbf{x}_q^{(k)})\odot\mathbf{x}^*) + \frac{1}{K}\langle(\mathbf{1}-\mathbf{x}_q^{(k)})\odot\mathbf{g}_q^{(k)}, \mathbf{v}_q^{(k+1)} - \mathbf{x}^*\rangle \\
&+ \frac{1}{K}\langle(\mathbf{v}_q^{(k+1)} - \mathbf{x}^*)\odot(\mathbf{1}-\mathbf{x}_q^{(k)}), \nabla\bar{f}_{q,k-1}(\mathbf{x}_q^k) - \mathbf{g}_q^{(k)}\rangle - \frac{L_0}{2}\|\mathbf{x}_q^{(k+1)} - \mathbf{x}_q^{(k)}\|^2 \\
\geq&(1-\frac{1}{K})\bar{f}_{q,k-1}(\mathbf{x}_q^{(k)}) + \frac{1}{K}(1-\frac{1}{K})^k\bar{f}_{q,k-1}(\mathbf{x}^*) + \frac{1}{K}\langle(\mathbf{1}-\mathbf{x}_q^{(k)})\odot\mathbf{g}_q^{(k)}, \mathbf{v}_q^{(k+1)} - \mathbf{x}^*\rangle \\
&+ \frac{1}{K}\langle(\mathbf{v}_q^{(k+1)} - \mathbf{x}^*)\odot(\mathbf{1}-\mathbf{x}_q^{(k)}), \nabla\bar{f}_{q,k-1}(\mathbf{x}_q^k) - \mathbf{g}_q^{(k)}\rangle - \frac{L_0}{2}\|\mathbf{x}_q^{(k+1)} - \mathbf{x}_q^{(k)}\|^2,
\end{aligned} \tag{24}$$

where the second inequality comes from Lemma 9 and Lemma 6. Therefore, by iteration, we have

$$\begin{aligned}
&\mathbb{E}(\bar{f}_q(\mathbf{x}_q^{(K)})) \\
=&\mathbb{E}(\bar{f}_{q,K-2}(\mathbf{x}_q^{(K)})) \\
\geq&(1-\frac{1}{K})\mathbb{E}(\bar{f}_{q,K-2}(\mathbf{x}_q^{(K-1)})) + \frac{1}{K}(1-\frac{1}{K})^{K-1}\mathbb{E}(\bar{f}_{q,K-2}(\mathbf{x}^*)) + \frac{1}{K}\mathbb{E}(\langle(\mathbf{1}-\mathbf{x}_q^{(K-1)})\odot\mathbf{g}_q^{(K-1)}, \mathbf{v}_q^{(K)} - \mathbf{x}^*\rangle) \\
&+ \frac{1}{K}\mathbb{E}(\langle(\mathbf{v}_q^{(K)} - \mathbf{x}^*)\odot(\mathbf{1}-\mathbf{x}_q^{(K-1)}), \nabla\bar{f}_{q,K-2}(\mathbf{x}_q^{K-1}) - \mathbf{g}_q^{(K-1)}\rangle) - \frac{L_0 r^2(\mathcal{C})}{2K^2} \\
=&(1-\frac{1}{K})\mathbb{E}(\bar{f}_{q,K-3}(\mathbf{x}_q^{(K-1)})) + \frac{1}{K}(1-\frac{1}{K})^{K-1}\bar{f}_q(\mathbf{x}^*) + \frac{1}{K}\mathbb{E}(\langle(\mathbf{1}-\mathbf{x}_q^{(K-1)})\odot\mathbf{g}_q^{(K-1)}, \mathbf{v}_q^{(K)} - \mathbf{x}^*\rangle) \\
&+ \frac{1}{K}\mathbb{E}(\langle(\mathbf{v}_q^{(K)} - \mathbf{x}^*)\odot(\mathbf{1}-\mathbf{x}_q^{(K-1)}), \nabla\bar{f}_{q,K-2}(\mathbf{x}_q^{K-1}) - \mathbf{g}_q^{(K-1)}\rangle) - \frac{L_0 r^2(\mathcal{C})}{2K^2} \\
\geq&\cdots \\
\geq&(1-\frac{1}{K})^{K-1}f_q(\mathbf{x}^*) + \frac{1}{K}\sum_{m=1}^{K-1}(1-\frac{1}{K})^{K-1-m}\mathbb{E}(\langle(\mathbf{1}-\mathbf{x}_q^{(m)})\odot\mathbf{g}_q^{(m)}, \mathbf{v}_q^{(m+1)} - \mathbf{x}^*\rangle) \\
&+ \frac{1}{K}\sum_{m=1}^{K-1}(1-\frac{1}{K})^{K-1-m}\mathbb{E}(\langle(\mathbf{v}_q^{(m+1)} - \mathbf{x}^*)\odot(\mathbf{1}-\mathbf{x}_q^{(m)}), \nabla\bar{f}_{q,m-1}(\mathbf{x}_q^m) - \mathbf{g}_q^{(m)}\rangle) - \frac{L_0 r^2(\mathcal{C})}{2K}.
\end{aligned} \tag{25}$$

Finally,

$$
\begin{aligned}
&\sum_{q=1}^{Q}\mathbb{E}(\bar{f}_q(\mathbf{x}_q^{(K)}))\\
\geq&(1-\frac{1}{K})^{K-1}\sum_{q=1}^{Q}f_q(\mathbf{x}^*)+\frac{1}{K}\sum_{q=1}^{Q}\sum_{m=1}^{K-1}(1-\frac{1}{K})^{K-1-m}\mathbb{E}(\langle(\mathbf{1}-\mathbf{x}_q^{(m)})\odot\mathbf{g}_q^{(m)},\mathbf{v}_q^{(m+1)}-\mathbf{x}^*\rangle)\\
&+\frac{1}{K}\sum_{q=1}^{Q}\sum_{m=1}^{K-1}(1-\frac{1}{K})^{K-1-m}\mathbb{E}(\langle(\mathbf{v}_q^{(m+1)}-\mathbf{x}^*)\odot(\mathbf{1}-\mathbf{x}_q^{(m)}),\nabla\bar{f}_{q,m-1}(\mathbf{x}_q^m)-\mathbf{g}_q^{(m)}\rangle)-\frac{L_0Qr^2(\mathcal{C})}{2K}\\
\geq&\frac{1}{e}\sum_{q=1}^{Q}f_q(\mathbf{x}^*)+\frac{1}{K}\sum_{m=1}^{K-1}(1-\frac{1}{K})^{K-1-m}\sum_{q=1}^{Q}\mathbb{E}(\langle(\mathbf{1}-\mathbf{x}_q^{(m)})\odot\mathbf{g}_q^{(m)},\mathbf{v}_q^{(m+1)}-\mathbf{x}^*\rangle)\\
&-\frac{1}{2K}\sum_{q=1}^{Q}\sum_{m=1}^{K-1}\mathbb{E}(b_m*\mathrm{diam}^2(\mathcal{C})+\frac{\|\nabla\bar{f}_{q,m-1}(\mathbf{x}_q^m)-\mathbf{g}_q^{(m)}\|^2}{b_m})-\frac{L_0Qr^2(\mathcal{C})}{2K}\\
\geq&\frac{1}{e}\sum_{q=1}^{Q}f_q(\mathbf{x}^*)-M_0\sqrt{Q}-\frac{L_0Qr^2(\mathcal{C})}{2K}-\frac{1}{2K}\sum_{q=1}^{Q}\sum_{m=1}^{K-1}\mathbb{E}(b_m*\mathrm{diam}^2(\mathcal{C})+\frac{\|\nabla\bar{f}_{q,m-1}(\mathbf{x}_q^m)-\mathbf{g}_q^{(m)}\|^2}{b_m}),
\end{aligned}
\tag{26}
$$

where the second inequality comes from $(1-\frac{1}{K})^{K-1}\geq\frac{1}{e}$ and $\langle(\mathbf{v}_q^{(m+1)}-\mathbf{x}^*)\odot(\mathbf{1}-\mathbf{x}_q^{(m)}),\nabla\bar{f}_{q,m-1}(\mathbf{x}_q^m)-\mathbf{g}_q^{(m)}\rangle\leq\frac{1}{2}(b_m*\mathrm{diam}^2(\mathcal{C})+\frac{\|\nabla\bar{f}_{q,m-1}(\mathbf{x}_q^m)-\mathbf{g}_q^{(m)}\|^2}{b_m})$ for any positive constant $b_m>0$; the third comes from $\sum_{q=1}^{Q}\mathbb{E}(\langle(\mathbf{1}-\mathbf{x}_q^{(m)})\odot\mathbf{g}_q^{(m)},\mathbf{v}_q^{(K)}-\mathbf{x}^*\rangle)\leq M_0\sqrt{Q}$.

If we consider $b_m=\frac{1}{\mathrm{diam}(\mathcal{C})(m+4)^{1/3}}$ when $1\leq m\leq\frac{K}{2}+1$ and $b_m=\frac{1}{\mathrm{diam}(\mathcal{C})(K-m+1)^{1/3}}$ when $\frac{K}{2}+2\leq m\leq K$, then we have

$$
\begin{aligned}
&\sum_{m=1}^{K-1}\mathrm{diam}^2(\mathcal{C})b_m\leq\sum_{m=1}^{K/2+1}\frac{\mathrm{diam}(\mathcal{C})}{(m+4)^{1/3}}+\sum_{m=K/2+2}^{K}\frac{\mathrm{diam}(\mathcal{C})}{(K-m+1)^{1/3}}\leq2\mathrm{diam}(\mathcal{C})K^{2/3},\\
&\sum_{m=1}^{K-1}\mathbb{E}(\frac{\|\nabla\bar{f}_{q,m-1}(\mathbf{x}_q^m)-\mathbf{g}_q^{(m)}\|^2}{b_m})\leq\sum_{m=1}^{K/2+1}\frac{\mathrm{diam}(\mathcal{C})N_1}{(m+4)^{1/3}}+\sum_{m=K/2+2}^{K}\frac{N_1}{(K-m+1)^{1/3}}\leq2N_1\mathrm{diam}(\mathcal{C})K^{2/3},
\end{aligned}
\tag{27}
$$

where the second inequality comes from Lemma 10 and $N_1=\max\{5^{2/3}G^2,8(\sigma^2+G^2)+32(2G+L_0r(\mathcal{C}))^2,4.5(\sigma^2+G^2)+7(2G+L_0r(\mathcal{C}))^2/3\}$.

As a result,

$$
\begin{aligned}
&\frac{1}{e}\sum_{t=1}^{T}f_t(\mathbf{x}^*)-\sum_{t=1}^{T}\mathbb{E}(f_t(\mathbf{y}_t))\\
=&K(\frac{1}{e}\sum_{q=1}^{Q}\bar{f}_q(\mathbf{x}^*)-\sum_{q=1}^{Q}\mathbb{E}(\bar{f}_q(\mathbf{x}_q^{(K)})))\\
\leq&2\mathrm{diam}(\mathcal{C})(N_1+1)QK^{2/3}+\frac{L_0r^2(\mathcal{C})}{2}Q+M_0\sqrt{Q}K.
\end{aligned}
\tag{28}
$$

□

E Proofs in Section 3.3

To begin, we review the properties of smoothed function.

Lemma 11 (Zhang et al. (2019); Chen et al. (2020))**.** *If $f:[0,1]^n\to\mathbb{R}_+$ is continuous DR-submodular, G-Lipschitz, and L_0-smooth, then so is $\hat{f}_\delta$ where $\hat{f}_\delta(\mathbf{x})=\mathbb{E}_{\mathbf{v}\sim B^n}(f(\mathbf{x}+\delta\mathbf{v}))$ and we have $|\hat{f}_\delta(\mathbf{x})-f(\mathbf{x})|\leq G\delta$ for all $\mathbf{x}\in[0,1]^n$.*

In this section, we begin by examining the sequence of iterates $\mathbf{x}_q^{(0)},\mathbf{x}_q^{(1)},\dots,\mathbf{x}_q^{(K)}$ in Algorithm 3. First, we derive the upper bound of $\tilde{\mathbf{x}}_q^{(k)}$ where $\tilde{\mathbf{x}}_q^{(k)}=(\mathbf{x}_q^{(k)}-\delta\mathbf{1})\oslash(\mathbf{1}-\delta\mathbf{1})$.

Lemma 12. *For $i \in [n]$ and $q \in [Q]$, we have $(\tilde{\mathbf{x}}_q^{(k)})_i \leq 1 - (1 - \frac{1}{K})^k$ where $\tilde{\mathbf{x}}_q^{(k)} = (\mathbf{x}_q^{(k)} - \delta\mathbf{1}) \oslash (\mathbf{1} - \delta\mathbf{1})$.*

Proof. From Algorithm 3 (See Line 7), we have

$$\mathbf{x}_q^{(k)} = \mathbf{x}_q^{(k-1)} + \frac{1}{K}\tilde{\mathbf{v}}_q^{(k)} \odot (\mathbf{1} - \mathbf{x}_q^{(k-1)}). \tag{29}$$

Therefore, we have

$$\tilde{\mathbf{x}}_q^{(k)} = \tilde{\mathbf{x}}_q^{(k-1)} + \frac{1}{K}\tilde{v}_q^{(k)} \odot (\mathbf{1} - \tilde{\mathbf{x}}_q^{(k-1)}). \tag{30}$$

Finally, due to $0 \leq \tilde{v}_q^{(k)} \leq \mathbf{1}$, we obtain

$$\tilde{\mathbf{x}}_q^{(k)} \leq \tilde{\mathbf{x}}_q^{(k-1)} + \frac{1}{K}(\mathbf{1} - \tilde{\mathbf{x}}_q^{(k-1)}). \tag{31}$$

According to Lemma 5, we get the result. □

Next, we define some notations frequently used in this section. For any f_t, we denote its δ-smoothed approximation as $\hat{f}_{t,\delta}(\mathbf{x}) = \mathbb{E}_{\mathbf{v}\sim B^n}(f_t(\mathbf{x} + \delta\mathbf{v}))$. Then, the average function for q-block is denoted as $\bar{F}_q(\mathbf{x}) = \frac{\sum_{m=(q-1)L+1}^{qL} \hat{f}_{m,\delta}((\mathbf{1}-\delta\mathbf{1})\odot\mathbf{x}+\delta\mathbf{1})}{L}$. Also, the average function of remaining $(L-l)$ rewards is $\bar{F}_{q,l}(\mathbf{x}) = \frac{\sum_{m=l+1}^{L} \hat{f}_{t_q^{(m)},\delta}(((\mathbf{1}-\delta\mathbf{1})\odot\mathbf{x}+\delta\mathbf{1}))}{L-l}$ where $0 \leq l \leq L-1$.

In the following part, we assume $\mathbf{x}^* = \arg\max_{\mathbf{x}\in\mathcal{C}} f_t(\mathbf{x})$ and $\mathbf{x}_\delta^* = \arg\max_{\mathbf{x}\in\mathcal{C}'} f_t(\mathbf{x})$. Then, we could conclude that

Lemma 13. *Under Assumption 1 and Assumption 3 , if $\|\nabla f_t(\mathbf{x})\| \leq G$, then*

$$\begin{aligned}
&\sum_{t=1}^{T}\frac{1}{e}f_t(\mathbf{x}^*) - \sum_{t=1}^{T} f_t(\mathbf{y}_t) \\
\leq& L\sum_{q=1}^{Q}\frac{1}{e}\bar{F}_q(\tilde{\mathbf{x}}_\delta^*) - L\sum_{q=1}^{Q}\bar{F}_q(\tilde{\mathbf{x}}_q^{(K)}) + 2M_1KQ + \left((\sqrt{n}+1)\frac{r(\mathcal{C})}{r} + \sqrt{n} + 2\right)TG\delta,
\end{aligned} \tag{32}$$

where $\tilde{\mathbf{x}}_\delta^* = (\mathbf{x}_\delta^* - \delta\mathbf{1}) \oslash (\mathbf{1} - \delta\mathbf{1})$ and $\tilde{\mathbf{x}}_q^{(K)} = (\mathbf{x}_q^{(K)} - \delta\mathbf{1}) \oslash (\mathbf{1} - \delta\mathbf{1})$.

Proof. We denote the $\mathbf{x}'$ as the projection of $\mathbf{x}^*$ on the $\mathcal{C}'$, i.e., $\mathbf{x}' = \arg\min_{x\in\mathcal{C}'} \|\mathbf{x} - \mathbf{x}^*\|$, we could conclude that

$$\begin{aligned}
\sum_{t=1}^{T}\frac{1}{e}f_t(\mathbf{x}^*) - \sum_{t=1}^{T} f_t(\mathbf{y}_t) =& \sum_{t=1}^{T}\frac{1}{e}f_t(\mathbf{x}^*) - \sum_{t=1}^{T}\frac{1}{e}f_t(\mathbf{x}_\delta^*) + \sum_{t=1}^{T}\frac{1}{e}f_t(\mathbf{x}_\delta^*) - \sum_{t=1}^{T}\frac{1}{e}\hat{f}_{t,\delta}(\mathbf{x}_\delta^*) \\
&+ \sum_{t=1}^{T}\frac{1}{e}\hat{f}_{t,\delta}(\mathbf{x}_\delta^*) - \sum_{t=1}^{T}\hat{f}_{t,\delta}(\mathbf{y}_t) + \sum_{t=1}^{T}\hat{f}_{t,\delta}(\mathbf{y}_t) - \sum_{t=1}^{T} f_t(\mathbf{y}_t).
\end{aligned} \tag{33}$$

First, $|\hat{f}_{t,\delta}(\mathbf{y}_t) - f_t(\mathbf{y}_t)| \leq G\delta$ and $|\hat{f}_{t,\delta}(\mathbf{x}_\delta^*) - f_t(\mathbf{x}_\delta^*)| \leq G\delta$. Then,

$$\begin{aligned}
&\sum_{t=1}^{T} f_t(\mathbf{x}^*) - \sum_{t=1}^{T} f_t(\mathbf{x}_\delta^*) \\
\leq& \sum_{t=1}^{T} f_t(\mathbf{x}^*) - \sum_{t=1}^{T} f_t(\mathbf{x}') \\
\leq& TG\|\mathbf{x}^* - \mathbf{x}'\| \\
\leq& \left((\sqrt{n}+1)\frac{r(\mathcal{C})}{r} + \sqrt{n}\right)TG\delta,
\end{aligned} \tag{34}$$

where the first inequality comes from the definition of $\mathbf{x}_\delta^*$ and $\mathbf{x}' \in \mathcal{C}'$; the second follows from the lipschitz of f_t; the final from Lemma 1 in Zhang et al. (2019).

Finally, if setting $\tilde{\mathbf{x}}_{\delta}^{*}=(\mathbf{x}_{\delta}^{*}-\delta\mathbf{1})\oslash(\mathbf{1}-\delta\mathbf{1})$ and $\tilde{\mathbf{x}}_{q}^{(K)}=(\mathbf{x}_{q}^{(K)}-\delta\mathbf{1})\oslash(\mathbf{1}-\delta\mathbf{1})$,

$$\begin{aligned}&\sum_{t=1}^{T}\frac{1}{e}\hat{f}_{t,\delta}(\mathbf{x}_{\delta}^{*})-\sum_{t=1}^{T}\hat{f}_{t,\delta}(\mathbf{y}_{t})\\=&L\sum_{q=1}^{Q}\frac{1}{e}\bar{F}_{q}(\tilde{\mathbf{x}}_{\delta}^{*})-L\sum_{q=1}^{Q}\bar{F}_{q}(\tilde{\mathbf{x}}_{q}^{(K)})+\sum_{q=1}^{Q}\sum_{k=1}^{K}(\hat{f}_{t_{q}^{(k)}}(\mathbf{x}_{q}^{(K)})-\hat{f}_{t_{q}^{(k)}}(\mathbf{y}_{t_{q}^{(k)}}))\\\leq&L\sum_{q=1}^{Q}\frac{1}{e}\bar{F}_{q}(\tilde{\mathbf{x}}_{\delta}^{*})-L\sum_{q=1}^{Q}\bar{F}_{q}(\tilde{\mathbf{x}}_{q}^{(K)})+2M_{1}KQ,\end{aligned}\tag{35}$$

where the inequality comes from $|\hat{f}_{t_{q}^{k}}(\mathbf{x}_{q}^{(K)})-\hat{f}_{t_{q}^{k}}(\mathbf{y}_{t_{q}^{k}})|\leq 2M_{1}$. Therefore,

$$\begin{aligned}&\sum_{t=1}^{T}\frac{1}{e}f_{t}(\mathbf{x}^{*})-\sum_{t=1}^{T}f_{t}(\mathbf{y}_{t})\\\leq&L\sum_{q=1}^{Q}\frac{1}{e}\bar{F}_{q}(\tilde{\mathbf{x}}_{\delta}^{*})-L\sum_{q=1}^{Q}\bar{F}_{q}(\tilde{\mathbf{x}}_{q}^{(K)})+2M_{1}KQ+\left((\sqrt{n}+1)\frac{r(\mathcal{C})}{r}+\sqrt{n}+2\right)TG\delta.\end{aligned}\tag{36}$$

□

Next, we demonstrate how $(\mathbf{1}-\delta\mathbf{1})\odot\mathbf{g}_{q}^{(k)}$(See Line 19 in Algorithm 3) approximates $\nabla\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_{q}^{(k)})$ where $\tilde{\mathbf{x}}_{q}^{(k)}=(\mathbf{x}_{q}^{(k)}-\delta\mathbf{1})\oslash(\mathbf{1}-\delta\mathbf{1})$.

Lemma 14. *Under Assumption 1 and Assumption 3, if $\|\nabla f_{t}(\mathbf{x})\|\leq G$, $L\geq 2K$ and $\eta_{k}=\frac{2}{(k+3)^{2/3}}$ for $k\in[K]$, we have, for any fixed $q\in[Q]$,*

$$\mathbb{E}(\|(\mathbf{1}-\delta\mathbf{1})\odot\mathbf{g}_{q}^{(k)}-\nabla\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_{q}^{(k)})\|^{2})\leq\frac{N_{2}}{(k+4)^{2/3}},\tag{37}$$

where $N_{2}=\max\{3^{2/3}G^{2},8(\frac{n^{2}M_{1}^{2}}{\delta^{2}}+G^{2})+3(4.5L_{0}r(\mathcal{C})+3G)^{2}/2\}$.

Proof. Similarly, we use $\mathcal{F}_{q,k}$ to denote the σ-field generated via $t_{q}^{(1)},\dots,t_{q}^{(k)}$. From Algorithm 3, $\mathbf{g}_{q}^{(k)}=(1-\eta_{k})\mathbf{g}_{q}^{(k-1)}+\eta_{k}\frac{n}{\delta}f_{t_{q}^{(k)}}(\mathbf{x}_{q}^{(k)}+\delta\mathbf{u}_{q}^{(k)})\mathbf{u}_{q}^{(k)}$. Also, we have $\frac{n}{\delta}\mathbb{E}(f_{t_{q}^{(k)}}(\mathbf{x}_{q}^{(k)}+\delta\mathbf{u}_{q}^{(k)})(\mathbf{1}-\delta\mathbf{1})\odot\mathbf{u}_{q}^{(k)}|\mathcal{F}_{q,k-1})=\nabla\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_{q}^{(k)})$. Next, we prove that

$$\begin{aligned}&\nabla\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_{q}^{(k)})-\nabla\bar{F}_{q,k-2}(\tilde{\mathbf{x}}_{q}^{(k-1)})\\=&\frac{\sum_{m=k}^{L}(1-\delta)\nabla\hat{f}_{t_{q}^{(m)},\delta}(\mathbf{x}_{q}^{(k)})}{L-k+1}-\frac{\sum_{m=k-1}^{L}(1-\delta)\nabla\hat{f}_{t_{q}^{(m)},\delta}(\mathbf{x}_{q}^{(k-1)})}{L-k+2}\\=&\frac{\sum_{m=k}^{L}(1-\delta)(\nabla\hat{f}_{t_{q}^{(m)},\delta}(\mathbf{x}_{q}^{(k)})-\nabla\hat{f}_{t_{q}^{(m)},\delta}(\mathbf{x}_{q}^{(k-1)}))}{L-k+2}+\frac{\sum_{m=k}^{L}(1-\delta)\nabla\hat{f}_{t_{q}^{(m)},\delta}(\mathbf{x}_{q}^{(k)})}{(L-k+1)(L-k+2)}\\&-\frac{(1-\delta)\nabla\hat{f}_{t_{q}^{(k-1)},\delta}(\mathbf{x}_{q}^{(k-1)})}{L-k+2}.\end{aligned}\tag{38}$$

Thus, when $L\geq 2K$,

$$\begin{aligned}&\|\nabla\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_{q}^{(k)})-\nabla\bar{F}_{q,k-2}(\tilde{\mathbf{x}}_{q}^{(k-1)})\|\\\leq&\|\frac{\sum_{m=k}^{L}(\nabla\hat{f}_{t_{q}^{(m)},\delta}(\mathbf{x}_{q}^{(k)})-\nabla\hat{f}_{t_{q}^{(m)},\delta}(\mathbf{x}_{q}^{(k-1)}))}{L-k+2}\|+\|\frac{\sum_{m=k}^{L}\nabla\hat{f}_{t_{q}^{(m)},\delta}(\mathbf{x}_{q}^{(k)})}{(L-k+1)(L-k+2)}\|+\|\frac{\nabla\hat{f}_{t_{q}^{(k-1)},\delta}(\mathbf{x}_{q}^{(k-1)})}{L-k+2}\|\\\leq&\frac{(L-k+1)L_{0}r(\mathcal{C})}{K(L-k+2)}+\frac{G}{L-k+2}+\frac{G}{L-k+2}\\\leq&\frac{4.5L_{0}r(\mathcal{C})+3G}{k+3},\end{aligned}\tag{39}$$

where the final inequality follows from $L-k+2\geq 2K-k+2\geq k+2$. Moreover,

$$\begin{aligned}
&\mathbb{E}(\|\frac{n}{\delta}f_{t_q^{(k)}}(\mathbf{x}_q^{(k)}+\delta\mathbf{u}_q^{(k)})(\mathbf{1}-\delta\mathbf{1})\odot\mathbf{u}_q^{(k)}-\nabla\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_q^{(k)})\|^2|\mathcal{F}_{q,k-1})\\
\leq&2\mathbb{E}(\|\frac{n}{\delta}f_{t_q^{(k)}}(\mathbf{x}_q^{(k)}+\delta\mathbf{u}_q^{(k)})(\mathbf{1}-\delta\mathbf{1})\odot\mathbf{u}_q^{(k)}-(1-\delta)\mathbf{1}\odot\nabla\hat{f}_{t_q^{(k)},\delta}(\mathbf{x}_q^{(k)})\|^2|\mathcal{F}_{q,k-1})\\
&+2\mathbb{E}((1-\delta)\mathbf{1}\odot\nabla\hat{f}_{t_q^{(k)},\delta}(\mathbf{x}_q^{(k)})-\nabla\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_q^{(k)})\|^2|\mathcal{F}_{q,k-1})\\
=&2\mathbb{E}(\|\frac{n}{\delta}f_{t_q^{(k)}}(\mathbf{x}_q^{(k)}+\delta\mathbf{u}_q^{(k)})(\mathbf{1}-\delta\mathbf{1})\odot\mathbf{u}_q^{(k)}-(1-\delta)\mathbf{1}\odot\nabla\hat{f}_{t_q^{(k)},\delta}(\mathbf{x}_q^{(k)})\|^2|\mathcal{F}_{q,k-1})\\
&+2Var((1-\delta)\mathbf{1}\odot\nabla\hat{f}_{t_q^{(k)},\delta}(\mathbf{x}_q^{(k)})|\mathcal{F}_{q,k-1}))\\
\leq&2\left(\frac{n^2M_1^2}{\delta^2}+G^2\right).
\end{aligned}\tag{40}$$

According to Lemma 3 where we set $\widetilde{\mathbf{a}}_k=\frac{n}{\delta}f_{t_q^{(k)}}(\mathbf{x}_q^{(k)}+\delta\mathbf{u}_q^{(k)})(\mathbf{1}-\delta\mathbf{1})\odot\mathbf{u}_q^{(k)}$, we have

$$\mathbb{E}(\|(\mathbf{1}-\delta\mathbf{1})\odot\mathbf{g}_q^{(k)}-\nabla F_{q,k-1}(\tilde{\mathbf{x}}_q^{(k)})\|^2)\leq\frac{N_2}{(k+4)^{2/3}},\tag{41}$$

where $N_2=\max\{3^{2/3}G^2,8(\frac{n^2M_1^2}{\delta^2}+G^2)+3(4.5L_0r(\mathcal{C})+3G)^2/2\}$. □

Lemma 15. *Under Assumption 1 and Assumption 3, if $\|\widetilde{\nabla}f_t(\mathbf{x})\|\leq G$ and $L\geq 2K$, we could conclude that*

$$\sum_{q=1}^{Q}\frac{1}{e}\bar{F}_q(\mathbf{x}_\delta^*)-\sum_{q=1}^{Q}\mathbb{E}(\bar{F}_q(\tilde{\mathbf{x}}_q^{(K)}))\leq\frac{L_0Qr^2(\mathcal{C})}{2K}+M_0\sqrt{Q}+\frac{\mathrm{diam}(\mathcal{C})Q}{2\delta K^{1/3}}+\frac{\mathrm{diam}(\mathcal{C})N_2\delta Q}{2K^{1/3}},\tag{42}$$

where $N_2=\max\{3^{2/3}G^2,8(\frac{n^2M_1^2}{\delta^2}+G^2)+3(4.5L_0r(\mathcal{C})+3G)^2/2\}$.

Proof. If we set $(f,\mathbf{x}_k,\mathbf{x}_{k-1},\mathbf{v}_k)$ in Lemma 7 as $(\bar{F}_{q,k-1},\tilde{\mathbf{x}}_q^{(k+1)},\tilde{\mathbf{x}}_q^{(k)},\mathbf{v}_q^{(k+1)})$ in the Algorithm 3, we have

$$\begin{aligned}
&\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_q^{(k+1)})\\
\geq&(1-\frac{1}{K})\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_q^{(k)})+\frac{1}{K}\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_q^{(k)}+(\mathbf{1}-\tilde{\mathbf{x}}_q^{(k)})\odot\mathbf{y})-\frac{L_0r(\mathcal{C})^2}{2K^2}\\
&+\frac{1}{K}\langle(\mathbf{1}-\tilde{\mathbf{x}}_q^{(k)})\odot d,\tilde{\mathbf{v}}_q^{(k+1)}-\mathbf{y}\rangle+\frac{1}{K}\langle(\tilde{\mathbf{v}}_q^{(k+1)}-\tilde{\mathbf{x}}_\delta^*)\odot(\mathbf{1}-\tilde{\mathbf{x}}_q^{(k)}),\nabla\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_q^k)-\mathbf{d}\rangle\\
\geq&(1-\frac{1}{K})\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_q^{(k)})+\frac{1}{K}(1-\frac{1}{K})^k\bar{F}_{q,k-1}(\mathbf{y})+\frac{1}{K}\langle(\mathbf{1}-\tilde{\mathbf{x}}_q^{(k)})\odot\mathbf{d},\tilde{\mathbf{v}}_q^{(k+1)}-\tilde{\mathbf{x}}_\delta^*\rangle\\
&+\frac{1}{K}\langle(\tilde{\mathbf{v}}_q^{(k+1)}-\mathbf{y})\odot(\mathbf{1}-\tilde{\mathbf{x}}_q^{(k)}),\nabla\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_q^k)-\mathbf{d}\rangle-\frac{L_0r(\mathcal{C})^2}{2K^2},
\end{aligned}\tag{43}$$

where the second inequality comes from Lemma 12 and Lemma 6.

If we set $\mathbf{d}=(1-\delta)\mathbf{g}_q^{(k)}$ and $\mathbf{y}=\tilde{\mathbf{x}}_\delta^*$, we have

$$\begin{aligned}
&\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_q^{(k+1)})\\
\geq&(1-\frac{1}{K})\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_q^{(k)})+\frac{1}{K}(1-\frac{1}{K})^k\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_\delta^*)+\frac{1}{K}\langle(\mathbf{1}-\tilde{\mathbf{x}}_q^{(k)})\odot((1-\delta)\mathbf{g}_q^{(k)}),\tilde{\mathbf{v}}_q^{(k+1)}-\tilde{\mathbf{x}}_\delta^*\rangle\\
&+\frac{1}{K}\langle(\tilde{\mathbf{v}}_q^{(k+1)}-\tilde{\mathbf{x}}_\delta^*)\odot(\mathbf{1}-\tilde{\mathbf{x}}_q^{(k)}),\nabla\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_q^k)-(1-\delta)\mathbf{g}_q^{(k)}\rangle-\frac{L_0r^2(\mathcal{C})}{2K^2}\\
=&(1-\frac{1}{K})\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_q^{(k)})+\frac{1}{K}(1-\frac{1}{K})^k\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_\delta^*)+\frac{1}{K}\langle(\mathbf{1}-\tilde{\mathbf{x}}_q^{(k)})\odot\mathbf{g}_q^{(k)},(1-\delta)(\tilde{\mathbf{v}}_q^{(k+1)}-\tilde{\mathbf{x}}_\delta^*)\rangle\\
&+\frac{1}{K}\langle(\tilde{\mathbf{v}}_q^{(k+1)}-\tilde{\mathbf{x}}_\delta^*)\odot(\mathbf{1}-\tilde{\mathbf{x}}_q^{(k)}),\nabla\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_q^k)-(1-\delta)\mathbf{g}_q^{(k)}\rangle-\frac{L_0r^2(\mathcal{C})}{2K^2}\\
=&(1-\frac{1}{K})\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_q^{(k)})+\frac{1}{K}(1-\frac{1}{K})^k\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_\delta^*)+\frac{1}{K}\langle(\mathbf{1}-\tilde{\mathbf{x}}_q^{(k)})\odot\mathbf{g}_q^{(k)},\mathbf{v}_q^{(k+1)}-\mathbf{x}_\delta^*)\rangle\\
&+\frac{1}{K}\langle(\tilde{\mathbf{v}}_q^{(k+1)}-\tilde{\mathbf{x}}_\delta^*)\odot(\mathbf{1}-\tilde{\mathbf{x}}_q^{(k)}),\nabla\bar{F}_{q,k-1}(\tilde{\mathbf{x}}_q^k)-(1-\delta)\mathbf{g}_q^{(k)}\rangle-\frac{L_0r^2(\mathcal{C})}{2K^2}.
\end{aligned}\tag{44}$$

Therefore, by iteration, we have

$$
\begin{aligned}
&\mathbb{E}(\bar{F}_q(\tilde{\mathbf{x}}_q^{(K)}))\\
=&\mathbb{E}(\bar{F}_{q,K-2}(\tilde{\mathbf{x}}_q^{(K)}))\\
\geq&(1-\frac{1}{K})\mathbb{E}(\bar{F}_{q,K-2}(\tilde{\mathbf{x}}_q^{(K-1)}))+\frac{1}{K}(1-\frac{1}{K})^{K-1}\mathbb{E}(\bar{F}_{q,K-2}(\tilde{\mathbf{x}}_\delta^{*}))+\frac{1}{K}\mathbb{E}(\langle(\mathbf{1}-\tilde{\mathbf{x}}_q^{(K-1)})\odot\mathbf{g}_q^{(K-1)},\mathbf{v}_q^{(K)}-\mathbf{x}_\delta^{*}\rangle)\\
&+\frac{1}{K}\mathbb{E}(\langle(\tilde{\mathbf{v}}_q^{(K)}-\tilde{\mathbf{x}}_\delta^{*})\odot(\mathbf{1}-\tilde{\mathbf{x}}_q^{(K-1)}),\nabla\bar{F}_{q,K-2}(\tilde{\mathbf{x}}_q^{K-1})-(1-\delta)\mathbf{g}_q^{(K-1)}\rangle)-\frac{L_0r^2(\mathcal{C})}{2K^2}\\
=&(1-\frac{1}{K})\mathbb{E}(\bar{F}_{q,K-3}(\tilde{\mathbf{x}}_q^{(K-1)}))+\frac{1}{K}(1-\frac{1}{K})^{K-1}\bar{F}_q(\tilde{\mathbf{x}}_\delta^{*})+\frac{1}{K}\mathbb{E}(\langle(\mathbf{1}-\tilde{\mathbf{x}}_q^{(K-1)})\odot\mathbf{g}_q^{(K-1)},\mathbf{v}_q^{(K)}-\mathbf{x}_\delta^{*}\rangle)\\
&+\frac{1}{K}\mathbb{E}(\langle(\tilde{\mathbf{v}}_q^{(K)}-\tilde{\mathbf{x}}_\delta^{*})\odot(\mathbf{1}-\tilde{\mathbf{x}}_q^{(K-1)}),\nabla\bar{F}_{q,K-2}(\tilde{\mathbf{x}}_q^{K-1})-(1-\delta)\mathbf{g}_q^{(K-1)}\rangle)-\frac{L_0r^2(\mathcal{C})}{2K^2}\\
\geq&\dots\\
\geq&(1-\frac{1}{K})^{K-1}\bar{F}_q(\mathbf{x}_\delta^{*})+\frac{1}{K}\sum_{m=1}^{K-1}(1-\frac{1}{K})^{K-1-m}\mathbb{E}(\langle(\mathbf{1}-\tilde{\mathbf{x}}_q^{(m)})\odot\mathbf{g}_q^{(m)},\mathbf{v}_q^{(m+1)}-\mathbf{x}_\delta^{*}\rangle)\\
&+\frac{1}{K}\sum_{m=1}^{K-1}(1-\frac{1}{K})^{K-1-m}\mathbb{E}(\langle(\tilde{\mathbf{v}}_q^{(m+1)}-\tilde{\mathbf{x}}_\delta^{*})\odot(\mathbf{1}-\tilde{\mathbf{x}}_q^{(m)}),\nabla\bar{F}_{q,m-1}(\tilde{\mathbf{x}}_q^{m})-(1-\delta)\mathbf{g}_q^{(m)}\rangle)-\frac{L_0r^2(\mathcal{C})}{2K}.
\end{aligned}
\tag{45}
$$

Then,

$$
\begin{aligned}
&\sum_{q=1}^{Q}\mathbb{E}(\bar{F}_q(\tilde{\mathbf{x}}_q^{(K)}))\\
\geq&(1-\frac{1}{K})^{K-1}\sum_{q=1}^{Q}\bar{F}_q(\mathbf{x}_\delta^{*})+\frac{1}{K}\sum_{m=1}^{K-1}(1-\frac{1}{K})^{K-1-m}\sum_{q=1}^{Q}\mathbb{E}(\langle(\mathbf{1}-\tilde{\mathbf{x}}_q^{(m)})\odot\mathbf{g}_q^{(m)},\mathbf{v}_q^{(m+1)}-\mathbf{x}_\delta^{*}\rangle)\\
&+\frac{1}{K}\sum_{q=1}^{Q}\sum_{m=1}^{K-1}(1-\frac{1}{K})^{K-1-m}\mathbb{E}(\langle(\tilde{\mathbf{v}}_q^{(m+1)}-\tilde{\mathbf{x}}_\delta^{*})\odot(\mathbf{1}-\tilde{\mathbf{x}}_q^{(m)}),\nabla\bar{F}_{q,m-1}(\tilde{\mathbf{x}}_q^{m})-(1-\delta)\mathbf{g}_q^{(m)}\rangle)-\frac{L_0Qr^2(\mathcal{C})}{2K}
\end{aligned}
\tag{46}
$$

First, for any $m\leq K$, $\sum_{q=1}^{Q}\mathbb{E}(\langle(\mathbf{1}-\tilde{\mathbf{x}}_q^{(m)})\odot\mathbf{g}_q^{(m)},\mathbf{x}_\delta^{*}-\mathbf{v}_q^{(m+1)}\rangle)\leq M_0\sqrt{Q}$. Next,

$$
\begin{aligned}
&\frac{1}{K}\sum_{m=1}^{K-1}\mathbb{E}(\langle(\tilde{\mathbf{v}}_q^{(m+1)}-\tilde{\mathbf{x}}_\delta^{*})\odot(\mathbf{1}-\tilde{\mathbf{x}}_q^{(m)}),\nabla\bar{F}_{q,m-1}(\tilde{\mathbf{x}}_q^{m})-(1-\delta)\mathbf{1}\odot\mathbf{g}_q^{(m)}\rangle)\\
\geq&-\frac{1}{2K}\sum_{m=1}^{K-1}\mathbb{E}(\mathrm{diam}^2(\mathcal{C})b+\|\nabla\bar{F}_{q,m-1}(\tilde{\mathbf{x}}_q^{m})-(1-\delta)\mathbf{1}\odot\mathbf{g}_q^{(m)}\|^2/b)\\
\geq&-\frac{\mathrm{diam}(\mathcal{C})}{2\delta K^{1/3}}-\frac{\mathrm{diam}(\mathcal{C})N_2\delta}{2K^{1/3}},
\end{aligned}
\tag{47}
$$

where the first inequality comes from the Cauchy inequality; the second follows from Lemma 14 and $b=\frac{1}{\mathrm{diam}(\mathcal{C})K^{1/3}\delta}$. Finally, due to $(1-\frac{1}{K})^{K-1}\geq\frac{1}{e}$, we have

$$
\sum_{q=1}^{Q}\frac{1}{e}\bar{F}_q(\mathbf{x}_\delta^{*})-\sum_{q=1}^{Q}\mathbb{E}(\bar{F}_q(\tilde{\mathbf{x}}_q^{(K)}))\leq\frac{L_0Qr^2(\mathcal{C})}{2K}+M_0\sqrt{Q}+\frac{\mathrm{diam}(\mathcal{C})Q}{2\delta K^{1/3}}+\frac{\mathrm{diam}(\mathcal{C})N_2\delta Q}{2K^{1/3}}.
\tag{48}
$$

□

Next, we prove Theorem 3.

Proof. Finally, according to Lemma 13 and Lemma 15, we have

$$
\begin{aligned}
&\sum_{t=1}^{T}\frac{1}{e}f_t(\mathbf{x}^*)-\sum_{t=1}^{T}\mathbb{E}(f_t(\mathbf{y}_t))\\
\le&\frac{L_0r^2(\mathcal{C})}{2}\frac{LQ}{K}+M_0L\sqrt{Q}+\frac{\mathrm{diam}(\mathcal{C})LQ}{2\delta K^{1/3}}+\frac{\mathrm{diam}(\mathcal{C})N_2\delta LQ}{2K^{1/3}}+2M_1KQ\\
&+\left((\sqrt{n}+1)\frac{r(\mathcal{C})}{r}+\sqrt{n}+2\right)TG\delta\\
\le&C_1\frac{LQ}{K}+M_0L\sqrt{Q}+\frac{C_2LQ}{2\delta K^{1/3}}+\frac{C_3\delta LQ}{2K^{1/3}}+2M_1KQ+C_4T\delta,
\end{aligned}
$$

where the first inequality follows from the $N_2\le\max\{3^{2/3}G^2,8G^2+3(4.5L_0r(\mathcal{C})+3G)^2/2\}+8\frac{n^2M_1^2}{\delta^2}$ and in the second inequality, we set $C_1=\frac{L_0r^2(\mathcal{C})}{2}$, $C_2=(8n^2M_1^2+1)\mathrm{diam}(\mathcal{C})$, $C_3=\max\{3^{2/3}G^2,8G^2+3(4.5L_0r(\mathcal{C})+3G)^2/2\}\mathrm{diam}(\mathcal{C})$ and $C_4=((\sqrt{n}+1)\frac{r(\mathcal{C})}{r}+\sqrt{n}+2)G$. □