Differentially Private Monotone Submodular Maximization Under Matroid and Knapsack Constraints

Omid Sadeghi
University of Washington

Maryam Fazel
University of Washington

Abstract

Numerous tasks in machine learning and artificial intelligence have been modeled as submodular maximization problems. These problems usually involve sensitive data about individuals, and in addition to maximizing the utility, privacy concerns should be considered. In this paper, we study the general framework of non-negative monotone submodular maximization subject to matroid or knapsack constraints in both offline and online settings. For the offline setting, we propose a differentially private $(1 - \frac{\kappa}{e})$-approximation algorithm, where $\kappa \in [0, 1]$ is the total curvature of the submodular set function, which improves upon prior works in terms of approximation guarantee and query complexity under the same privacy budget. In the online setting, we propose the first differentially private algorithm, and we specify the conditions under which the regret bound scales as $\mathcal{O}(\sqrt{T})$, i.e., privacy could be ensured while maintaining the same regret bound as the optimal regret guarantee in the non-private setting.

1 Introduction

A set function $F : 2^V \to \mathbb{R}$ over the ground set V is submodular if for all $j \in V$ and for all sets $A \subseteq B \subseteq V \setminus \{j\}$, the following holds:

$$F(A \cup \{j\}) - F(A) \geq F(B \cup \{j\}) - F(B).$$

The submodularity property of set functions has profound theoretical consequences and far-reaching applications. Submodular set functions play a significant role in combinatorial optimization as many well known combinatorial functions are indeed submodular. Cut functions of graphs and hypergraphs, rank functions of matroids and covering functions are a few examples of submodular set functions. Moreover, submodularity has been identified and utilized in applications such as viral marketing (Kempe et al., 2003), feature selection for classification (Krause and Guestrin, 2007), image segmentation (Kohli et al., 2008; Boykov and Jolly, 2001) and document summarization (Lin and Bilmes, 2011; Kirchhoff and Bilmes, 2014). The multilinear extension $f : [0, 1]^V \to \mathbb{R}$ of F is defined as (Calinescu et al., 2011a)

$$f(x) = \sum_{S \subset V} F(S) \prod_{i \in S} x_i \prod_{j \notin S} (1 - x_j) = \mathbb{E}_{S \sim x}[F(S)].$$

Multilinear extensions coupled with lossless rounding techniques are extensively used for maximizing the corresponding submodular set functions. In particular, for submodular maximization subject to matroid constraints, Calinescu et al. (2011a) and Chekuri et al. (2010) proposed the pipage rounding and swap rounding schemes respectively to round the fractional solution without losing in terms of the objective function. Kulik et al. (2009, 2013) provided lossless rounding techniques for knapsack constraints. It has been shown that multilinear extensions can be efficiently computed for a large class of submodular set functions, for example, weighted matroid rank function, set cover function, probabilistic coverage function and graph cut function.

In applications where the submodular function involves sensitive data about individuals, privacy concerns should be addressed as well as obtaining good approximation guarantees. For instance, consider the following feature selection problem (Krause and Guestrin, 2007; Mitrovic et al., 2017):

Example 1. *Let $D = \{(x_t, C_t)\}_{t=1}^T$ be a sensitive dataset consisting of a feature vector $x_t \in \mathbb{R}^m$ for each individual $t \in [T]$ along with a binary class label C_t. The goal is to select a small subset of the m features that provide a good classifier for the labels. In particular, determining the likeliness of an individual having a certain disease using a representative collection of his*

Proceedings of the 24th International Conference on Artificial Intelligence and Statistics (AISTATS) 2021, San Diego, California, USA. PMLR: Volume 130. Copyright 2021 by the author(s).

or her features (such as height, age and weight) could be cast in this framework.

In order to solve this problem, Krause and Guestrin (2007) proposed a non-private algorithm based on maximizing a submodular function capturing the mutual information between a subset of the features and the class label of interest. However, in this setting, along with obtaining the most relevant subset of features, it is crucial to ensure that the privacy of any individual included in the dataset is not compromised. See Mirzasoleiman et al. (2016) for more applications (such as personal data summarization) that could be cast as submodular problems under more general constraints in which privacy concerns should be addressed.

1.1 Preliminaries

Notations. We use $[T]$ to denote the set $\{1, 2, \ldots, T\}$. For a vector $x \in \mathbb{R}^m$, we use x_i to denote the i-th entry of x. The inner product of two vectors $x, y \in \mathbb{R}^m$ is denoted by either $\langle x, y\rangle$ or $x^T y$. Also, for two vectors $x, y \in \mathbb{R}^m$, we write $x \preceq y$ if $x_i \leq y_i$ holds for every $i \in [m]$. A set function $F : 2^V \to \mathbb{R}$ is called monotone if for all S, S' such that $S \subseteq S'$, $F(S) \leq F(S')$ holds. The dual norm $\|\cdot\|_*$ of a norm $\|\cdot\|$ is defined as $\|y\|_* = \max_{x:\|x\|\leq 1}\langle y, x\rangle$.

DR-submodular functions. (Sadeghi and Fazel, 2020) We say that a differentiable function $f : \mathcal{X} \to \mathbb{R}$, $\mathcal{X} \subset \mathbb{R}^m_+$ (i.e., *continuous* real variables), is DR-submodular if its gradient is an order-reversing mapping, i.e.,

$$x \succeq y \Rightarrow \nabla f(x) \preceq \nabla f(y).$$

A twice differentiable function f is DR-submodular if and only if its Hessian matrix $\nabla^2 f$ is entry-wise non-positive. It is noteworthy that although DR-submodularity and concavity are equivalent for the special case of $m = 1$, DR-submodular functions are generally non-concave. Nonetheless, an important consequence of DR-submodularity is concavity along non-negative directions (Bian et al., 2017; Calinescu et al., 2011b), i.e., for all x, y such that $x \preceq y$, we have $f(y) \leq f(x) + \langle\nabla f(x), y - x\rangle$.

For a DR-submodular function f, we say that f is L-smooth over non-negative directions with respect to the norm $\|\cdot\|$ if

$$f(y)-f(x) \geq \langle\nabla f(x), y-x\rangle - \frac{L}{2}\|y-x\|^2 \ \ \forall x, y; x \preceq y,$$

or equivalently, $\|\nabla f(y) - \nabla f(x)\|_* \leq L\|y - x\| \ \forall x, y; x \preceq y$ holds, where $\|\cdot\|_*$ is the dual norm of the norm $\|\cdot\|$. There are many functions which satisfy the DR-submodularity property (Sadeghi et al., 2020; Raut et al., 2020). In particular, multilinear extensions of submodular set functions are DR-submodular. The Hessian matrix of this class of functions has non-positive off-diagonal entries and all its diagonal entries are zero.

Total curvature of submodular set functions. For a monotone submodular set function $F : 2^V \to \mathbb{R}_+$, the total curvature κ_F is defined as follows:

$$\kappa_F = 1 - \min_{v\in V} \frac{F(V) - F(V \setminus \{v\})}{F(\{v\}) - F(\{\})}.$$

Since F is monotone, $\kappa_F \leq 1$ holds. Also, submodularity of F ensures that $\kappa_F \geq 0$. Therefore, we have $0 \leq \kappa_F \leq 1$.

Matroids and matroid polytopes. A matroid $\mathcal{M}$ is a pair $\mathcal{M} = (V, \mathcal{I})$, where V is a finite ground set and $\mathcal{I}$ is a collection of subsets of V called the independent sets, that satisfies the following properties: 1) $\emptyset \in \mathcal{I}$. 2) For $S' \subset S \subset V$, if $S \in \mathcal{I}$, then $S' \in \mathcal{I}$ holds. 3) For $S, S' \in \mathcal{I}$, if $|S| > |S'|$, there exists $v \in S \setminus S'$ such that $S' \cup \{v\} \in \mathcal{I}$. The matroid polytope corresponding to the matroid $\mathcal{M} = (V, \mathcal{I})$ is defined as

$$\begin{aligned} P(\mathcal{M}) &= \mathrm{conv}\{1_I : I \in \mathcal{I}\} \\ &= \{x \succeq 0 : \sum_{s\in S} x_s \leq r_{\mathcal{M}}(S),\ \forall S \subset V\}, \end{aligned}$$

where the rank function $r_{\mathcal{M}} : 2^V \to \mathbb{Z}_+$ is $r_{\mathcal{M}}(S) = \max\{|I| : I \subseteq S, I \in \mathcal{I}\}$ and conv denotes the convex hull. We define $\mathrm{rank}(\mathcal{M}) = r_{\mathcal{M}}(V)$ as the rank of the matroid $\mathcal{M}$.

Knapsack constraints and knapsack polytopes. Given a ground set V, a positive vector $c \in \mathbb{R}^{|V|}_{++}$ and a collection $\mathcal{I} = \{S \subseteq V : \sum_{s\in S} c_s \leq 1\}$ of subsets of V, $S \in \mathcal{I}$ is called a knapsack constraint. The natural continuous relaxation $\{x \in [0, 1]^{|V|} : c^T x \leq 1\}$ is the knapsack polytope corresponding to $\mathcal{I}$.

1.2 Related work

Non-private submodular maximization. Maximizing non-negative monotone submodular set functions under a certain constraint has been extensively studied in the literature in both offline and online settings. Consider the problem of maximizing the monotone submodular function $F(S)$ subject to a cardinality constraint $|S| \leq k$. For offline monotone submodular set function maximization subject to a cardinality constraint, Nemhauser et al. (1978) proposed a simple greedy algorithm that obtains the provably optimal approximation ratio of $1 - \frac{1}{e}$. At each round $i \in [k]$, the greedy algorithm constructs S_i from S_{i-1} by adding the element $v_i \in V \setminus S_{i-1}$ which maximizes the marginal gain $F(S_{i-1} \cup \{v_i\}) - F(S_{i-1})$. However, if $\mathcal{M} = (V, \mathcal{I})$ is a matroid, the greedy algorithm applied to the submodular maximization problem subject to the matroid constraint $S \in \mathcal{I}$ achieves a sub-optimal $\frac{1}{2}$ approximation ratio. Calinescu et al. (2011a) proposed the continuous greedy algorithm for this problem which achieves the optimal $1-\frac{1}{e}$ approximation ratio. The continuous greedy algorithm is applied to the multilinear extension of the submodular set function under the matroid polytope and is as follows:

$$\begin{array}{ll} dy/dt & = v_{\max}(y) \\ v_{\max}(y) & = \arg\max_{v\in P(\mathcal{M})}\langle v, \nabla f(y)\rangle \end{array},$$

where f is the multilinear extension of the submodular set function F and $P(\mathcal{M})$ is the matroid polytope corresponding to the matroid $\mathcal{M}$. In this algorithm, $y(1)=\int_0^1 v_{\max}(y(\tau))d\tau$ is the output. More recently, Sviridenko et al. (2017) introduced a modification of the continuous greedy algorithm with an approximation ratio of $1-\frac{\kappa}{e}$, where $\kappa\in[0,1]$ is the total curvature of the submodular set function, and proved that the derived approximation ratio is indeed optimal.
In the online setting, Chen et al. (2018) proposed an online variant of the continuous greedy algorithm, called the Meta Frank-Wolfe algorithm, with a provably optimal $\mathcal{O}(\sqrt{T})$ regret bound, where T is the length of the horizon.

Offline differentially private submodular maximization. Let D be a sensitive dataset associated to a monotone submodular set function $F_D: 2^V\rightarrow\mathbb{R}_+$. Offline submodular maximization in the context of differential privacy has been studied under two different settings:

- F_D is Δ-decomposable (Gupta et al.; Chaturvedi et al., 2020): In this setting, it is assumed that $D=(F_1,\dots,F_T)$, where for all $t\in[T]$, $F_t: 2^V\rightarrow[0,\Delta]$ is a *private* monotone submodular set function and $F_D(\cdot)=\frac{1}{T}\sum_{t=1}^T F_t(\cdot)$.
- F_D is Δ-sensitive (Mitrovic et al., 2017; Rafiey and Yoshida, 2020): For this framework, we have $D=(F_1,\dots,F_T)$, where for all $t\in[T]$, $F_t: 2^V\rightarrow\mathbb{R}_+$ is a *private* monotone submodular set function, however, the submodular objective function F_D depends on the dataset D in ways that could be much more complicated than simply averaging the private submodular functions $F_1,\dots,F_T$ (e.g., Example 1). Two datasets D and D' are neighboring ($D\sim D'$) if all but one of the T submodular functions in the datasets are equal. It is assumed that F_D is Δ-sensitive, i.e., $\Delta=\max_{D':D'\sim D}\max_{S\subseteq V}|F_D(S)-F_{D'}(S)|$ holds.

Note that if F_D is Δ-decomposable, the sensitivity parameter is bounded above by $\frac{\Delta}{T}$, implying F_D is $(\frac{\Delta}{T})$-sensitive. In this paper, we focus on the more general setting where F_D is Δ-sensitive, and we review the prior work in this setting below. For submodular maximization subject to cardinality or matroid constraints, Mitrovic et al. (2017) combined the greedy algorithm with the exponential mechanism of McSherry and Talwar (2007) for differential privacy as follows: At round $i\in[k]$ of the greedy algorithm, define a quality function q_i via $q_i(v,D)=F_D(S_{i-1}\cup\{v\})-F_D(S_{i-1})$, and select every $v\in V\setminus S_{i-1}$ with probability proportional to $\exp(\epsilon q(v,D)/2\lambda)$ where λ is the sensitivity of the quality function q, i.e., for all $v\in V$ and two neighboring datasets D and D', we have $|q(v,D)-q(v,D')|\leq\lambda$. Mitrovic et al. (2017) showed that this algorithm is ϵ-differentially private (see Section 2 for a formal definition of differential privacy) and obtained an expected utility bound of $(1-\frac{1}{e})\text{OPT}-\mathcal{O}(\frac{\Delta k^2\ln(|V|)}{\epsilon})$ and $\frac{1}{2}\text{OPT}-\mathcal{O}(\frac{\Delta(\text{rank}(\mathcal{M}))^2\ln(|V|)}{\epsilon})$ for submodular maximization subject to cardinality constraint $|S|\leq k$ and matroid constraint $\mathcal{M}=(V,\mathcal{I})$ respectively. However, the result of Mitrovic et al. (2017) for the setting with matroid constraints fails to achieve the optimal approximation ratio of $1-\frac{1}{e}$. More recently, for submodular maximization subject to a matroid constraint $\mathcal{M}=(V,\mathcal{I})$, Rafiey and Yoshida (2020) combined the discretized version of the continuous greedy algorithm with the exponential mechanism in the following way: Let C_ρ be a ρ-covering of the matroid polytope $P(\mathcal{M})$, i.e., for any $x\in P(\mathcal{M})$, there exists $y\in C_\rho$ such that $\|x-y\|_2\leq\rho$. At each round $k\in[K]$, where $K=\text{rank}(\mathcal{M})$ is the rank of the matroid, the algorithm samples $y_k\in C_\rho$ with probability proportional to $\exp(\epsilon\langle y_k,\nabla f_D(x_k)\rangle/2\Delta)$, and sets $x_{k+1}=x_k+\frac{1}{K}y_k$ (where $x_1=0$). Rafiey and Yoshida (2020) showed that the output of this algorithm (x_{K+1}) obtains the utility bound $(1-\frac{1}{e})\text{OPT}-\mathcal{O}(\sqrt{\epsilon}+\frac{\Delta(\text{rank}(\mathcal{M}))^7|V|\ln(|V|)}{\epsilon^3})$ with high probability while ensuring ϵ-differential privacy. Although the result in Rafiey and Yoshida (2020) has the optimal $1-\frac{1}{e}$ approximation ratio, it has several major drawbacks that are as follows:

1. The $\mathcal{O}(\sqrt{\epsilon})$ term in the approximation guarantee is unusual, i.e., if $\epsilon\rightarrow\infty$ (no differential privacy), the approximation guarantee is vacuous.
2. In order to ensure differential privacy, $\rho=\frac{\epsilon}{|V|}$ should hold and thus, $|C_\rho|=\mathcal{O}(|V|^{1+(\frac{\text{rank}(\mathcal{M})}{\epsilon})^2})$. Therefore, discretization of the matroid polytope and implementing the exponential mechanism over such a large domain requires $\mathcal{O}(\text{rank}(\mathcal{M})|V|^{1+(\frac{\text{rank}(\mathcal{M})}{\epsilon})^2})$ gradient evaluation of the multilinear extension, and is not computationally efficient (although Rafiey and Yoshida (2020) provided a second algorithm with improved query complexity, it is still computationally expensive).
3. The dependence of the additive factor $\mathcal{O}(\frac{\Delta(\text{rank}(\mathcal{M}))^7|V|\ln(|V|)}{\epsilon^3})$ in the approximation guarantee on ϵ and $\text{rank}(\mathcal{M})$ is not optimal.

Differentially private online learning. Online submodular maximization has not been studied under the context of differential privacy yet. Nonetheless, the problem of differentially private online learning is extensively studied for linear and convex objective functions (Dwork et al., 2010; Jain et al., 2012; Guha Thakurta and Smith, 2013; Agarwal and Singh, 2017). In particular, Agarwal and Singh (2017) considered an online linear optimization problem over T rounds where at each step $t \in [T]$, the algorithm first chooses a point $x_t \in \mathcal{X}$ in the convex and compact domain set $\mathcal{X}$ and subsequently, it observes the loss vector ℓ_t and incurs a loss of $\langle \ell_t, x_t \rangle$. Agarwal and Singh (2017) proposed an ϵ-differentially private modification of the well-known Follow the Regularized Leader (FTRL) scheme for linear objectives with a regret bound that scales as $\mathcal{O}(\sqrt{T}) + \tilde{\mathcal{O}}(\frac{1}{\epsilon})$. Therefore, if $\epsilon \geq \Omega(\frac{1}{\sqrt{T}})$, the regret incurred by the differentially private algorithm matches the optimal $\mathcal{O}(\sqrt{T})$ regret in the non-private setting, i.e., differential privacy could be ensured for free. We use this algorithm as a sub-routine in our proposed algorithm for differentially private online submodular maximization.

1.3 Contributions

In this paper, we study the general framework of monotone submodular maximization subject to matroid or knapsack constraints in both offline and online settings. Specifically, we make the following contributions (all missing proofs are provided in the Appendix):

- In Section 3.1, we propose the Differentially Private Continuous Greedy (DPCG) algorithm for offline monotone submodular maximization subject to matroid or knapsack constraints and we analyze its performance in both settings. The DPCG algorithm is ϵ-differentially private under both constraints. For matroid constraints, we obtain a utility bound of $(1 - \frac{1}{e})\text{OPT} - \mathcal{O}(\sqrt{\frac{\Delta(\text{rank}(\mathcal{M}))^3 |V| \ln(|V|)}{\epsilon}})$ with $\mathcal{O}(\sqrt{\frac{\epsilon.\text{rank}(\mathcal{M})}{|V| \ln(|V|)\Delta}})$ multilinear extension evaluations which is a significant improvement over the $(1 - \frac{1}{e})\text{OPT} - \mathcal{O}(\sqrt{\epsilon} + \frac{\Delta(\text{rank}(\mathcal{M}))^7 |V| \ln(|V|)}{\epsilon^3})$ bound in Rafiey and Yoshida (2020) with $\mathcal{O}(\text{rank}(\mathcal{M})|V|^{1+(\frac{\text{rank}(\mathcal{M})}{\epsilon})^2})$ multilinear extension evaluations. Also, we obtain the first approximation guarantee for knapsack constraint $\{S \subseteq V : \sum_{s \in S} c_s \leq 1\}$ which is $(1 - \frac{1}{e})\text{OPT} - \mathcal{O}(\sqrt{\frac{\Delta |V| \ln(|V|)}{(c_{\min})^3 \epsilon}})$, where $c_v \geq c_{\min}$, $\forall v \in V$.
- For submodular functions with bounded curvature, we propose a modification of the DPCG algorithm, called the κ-DPCG algorithm, in Section 3.2 which has a utility bound of $(1 - \frac{\kappa}{e})\text{OPT} - \mathcal{O}(\sqrt{\frac{\Delta(\text{rank}(\mathcal{M}))^3 |V| \ln(|V|)}{\epsilon}})$ and $(1 - \frac{\kappa}{e})\text{OPT} - \mathcal{O}(\sqrt{\frac{\Delta |V| \ln(|V|)}{(c_{\min})^3 \epsilon}})$ for matroid and knapsack constraints respectively, where $\kappa \in [0, 1]$ is the total curvature of the submodular set function. In other words, compared to the DPCG algorithm, the κ-DPCG algorithm maintains the same additive factor in its utility bound while its $1 - \frac{\kappa}{e}$ approximation ratio is strictly better than $1 - \frac{1}{e}$ for submodular functions with curvature $\kappa < 1$.
- In the online setting, we propose the first algorithm for (ϵ, δ)-differentially private (defined in Section 2) submodular maximization, namely the Differentially Private Meta Frank-Wolfe (DPMFW) algorithm, whose regret bound scales as $\mathcal{O}(\sqrt{T}) + \mathcal{O}(\frac{T^{1/4}\sqrt{\ln(1/\delta)}}{\epsilon})$. Therefore, if $\frac{\sqrt{\ln(1/\delta)}}{\epsilon} \leq \mathcal{O}(T^{1/4})$, the regret bound of the DPMFW algorithm matches the provably optimal $\mathcal{O}(\sqrt{T})$ bound in the non-private setting, i.e., privacy could be guaranteed for free.

Note that although all our proposed algorithms are applied to the multilinear extension of the discrete submodular objective function, we can couple the algorithms with the lossless rounding schemes of Calinescu et al. (2011a); Chekuri et al. (2010) for matroid constraints and Kulik et al. (2009, 2013) for knapsack constraints to obtain discrete solutions with similar guarantees for the original submodular set function.

2 Differential privacy

In this work, a sensitive dataset D consists of *private* non-negative monotone submodular set functions $F_1, \ldots, F_T : 2^V \to \mathbb{R}_+$. In the offline setting, the non-negative monotone submodular objective function $F_D : 2^V \to \mathbb{R}_+$ depends on $\{F_t\}_{t=1}^T$ and is given in advance. Two datasets D and D' are neighboring $(D \sim D')$ if all but one of the T submodular functions in the datasets are equal. We define the sensitivity of the submodular set function F_D via

$$\Delta = \max_{D' : D' \sim D} \max_{S \subseteq V} |F_D(S) - F_{D'}(S)|.$$

For the online setting, at each step $t \in [T]$ (where T is the length of the horizon), the private submodular function F_t arrives after committing to an action S_t leading to a utility $F_t(S_t)$ for the algorithm.

Definition 1. *(Dwork and Roth, 2014) For $\epsilon, \delta \in \mathbb{R}_+$, a randomized algorithm $\mathcal{A}$ is called (ϵ, δ)-differentially private if for any two neighboring datasets D and D' and any set of possible outcomes S, the following holds:*

$$\mathbb{P}[\mathcal{A}(D) \in S] \leq \exp(\epsilon)\mathbb{P}[\mathcal{A}(D') \in S] + \delta.$$

If $\delta = 0$, we say that $\mathcal{A}$ is ϵ-differentially private.

Theorem 1 (Basic composition theorem)**.** *(Dwork and Roth, 2014) Let $\mathcal{A}_k$ be an (ϵ_k, δ_k)-differentially private algorithm for $k \in [K]$. Then, if algorithm $\mathcal{A}$ is defined to be $\mathcal{A} = (\mathcal{A}_1, \dots, \mathcal{A}_K)$, $\mathcal{A}$ is $(\sum_{k=1}^{K} \epsilon_k, \sum_{k=1}^{K} \delta_k)$-differentially private.*

K-fold adaptive composition. Let $\{(\epsilon_k, \delta_k)\}_{k=1}^{K}$ be a sequence of privacy parameters and let $\mathcal{A}$ be an algorithm that works as follows on a dataset D: At each round $k \in [K]$, the algorithm chooses an (ϵ_k, δ_k)-differentially private algorithm $\mathcal{A}_k$ and releases the output of $\mathcal{A}_k$, where $\mathcal{A}_k$ depends on the output of the previous algorithms $\mathcal{A}_1, \dots, \mathcal{A}_{k-1}$ but not on the dataset D itself. The output of $\mathcal{A}$ is called the *K-fold adaptive composition* of (ϵ_k, δ_k)-differentially private algorithms $\mathcal{A}_k$. The following privacy guarantee holds for the composite algorithm $\mathcal{A}$.

Theorem 2 (Advanced composition theorem)**.** *(Dwork and Roth, 2014) Given target privacy parameters $0 < \epsilon' < 1$ and $\delta' > 0$, in order to ensure $(\epsilon', K\delta + \delta')$ cumulative privacy loss for the composite algorithm $\mathcal{A}$, it suffices that each algorithm is (ϵ, δ)-differentially private, where $\epsilon = \frac{\epsilon'}{2\sqrt{2K\ln(1/\delta')}}$.*

3 Differentially private offline submodular maximization

In this section, we first introduce the Differentially Private Continuous Greedy (DPCG) algorithm and analyze its approximation and privacy guarantees in Section 3.1. Then, in Section 3.2, we consider the setting where the submodular objective function has a bounded curvature, and we introduce the κ-DPCG algorithm with improved approximation ratio.

3.1 Differentially Private Continuous Greedy (DPCG) algorithm

We propose the Differentially Private Continuous Greedy (DPCG) algorithm in Algorithm 1. The algorithm performs K Frank-Wolfe iterations to obtain $\{v_k\}_{k=1}^{K}$ and outputs the average $x = \frac{1}{K}\sum_{k=1}^{K} v_k$. Note that the output is the average of K points in the convex constraint set P (the matroid or knapsack polytope) and hence, $x \in P$ holds. The privacy is ensured through adding noise sampled from the distribution $\mathcal{D}$ to the gradients $\{\nabla f(x^{(k)})\}_{k=1}^{K}$. We first provide three useful Lemmas below.

Algorithm 1 Differentially Private Continuous Greedy (DPCG) algorithm

```
Input: $K$, the constraint set $P$, the multilinear extension $f : [0,1]^{|V|} \to \mathbb{R}$ of the monotone submodular set function $F : 2^V \to \mathbb{R}$, and the noise distribution $\mathcal{D}$.
Initialization: $x^{(1)} = 0$.
for $k = 1, 2, \dots, K$ do
    Draw $Y^{(k)} \sim \mathcal{D}$.
    Set $v_k = \arg\max_{v \in P} \langle v, \nabla f(x^{(k)}) + Y^{(k)} \rangle$.
    Set $x^{(k+1)} = x^{(k)} + \frac{1}{K} v_k$.
end for
Output: $x = x^{(K+1)}$.
```

Lemma 1. *If the multilinear extension f is L-smooth with respect to $\|\cdot\|_1$ and the diameter of P is denoted by $R = \max_{x \in P} \|x\|_1$, Algorithm 1 outputs $x = x^{(K+1)}$ such that the following holds:*

$$\mathbb{E}[f(x)] \geq (1 - \frac{1}{e}) f(x^*) - G_{\mathcal{D}} - \frac{LR^2}{2K},$$

where expectation is taken with respect to the noise distribution $\mathcal{D}$ and $G_{\mathcal{D}} := \mathbb{E}_{Y \sim \mathcal{D}}\big[\max_{x \in P}\langle Y, x\rangle - \min_{x \in P}\langle Y, x\rangle\big]$ measures the width of P under $\mathcal{D}$.

Proof. For $k \in [K]$, we can write:

$$\begin{aligned}
f(x^{(k+1)}) - f(x^{(k)}) &\overset{(a)}{\geq} \frac{1}{K}\langle v_k, \nabla f(x^{(k)})\rangle - \frac{L}{2K^2}\|v_k\|_1^2 \\
&\overset{(b)}{\geq} \frac{1}{K}\langle x^*, \nabla f(x^{(k)})\rangle + \frac{1}{K}\langle x^* - v_k, Y^{(k)}\rangle - \frac{LR^2}{2K^2} \\
&\overset{(c)}{\geq} \frac{1}{K}\langle (x^* - x^{(k)}) \vee 0, \nabla f(x^{(k)})\rangle \\
&+ \frac{1}{K}\langle x^* - v_k, Y^{(k)}\rangle - \frac{LR^2}{2K^2} \\
&\overset{(d)}{\geq} \frac{1}{K}\big(f(x^* \vee x^{(k)}) - f(x^{(k)})\big) \\
&+ \frac{1}{K}\langle x^* - v_k, Y^{(k)}\rangle - \frac{LR^2}{2K^2} \\
&\overset{(e)}{\geq} \frac{1}{K}\big(f(x^*) - f(x^{(k)})\big) + \frac{1}{K}\langle x^* - v_k, Y^{(k)}\rangle - \frac{LR^2}{2K^2},
\end{aligned}$$

where (a) is due to L-smoothness of f, (b) follows from the update rule of the algorithm, (c) and (e) use the monotonocity of f and (d) exploits concavity of f along non-negative directions. Taking the expectation of the above inequality and rearranging the terms, we have:

$$\begin{aligned}
\mathbb{E}[f(x^{(k+1)})] - f(x^*) \geq (1 - \frac{1}{K})\big(\mathbb{E}[f(x^{(k)})] - f(x^*)\big) \\
- \frac{1}{K}G_{\mathcal{D}} - \frac{LR^2}{2K^2}.
\end{aligned}$$

Applying the inequality recursively for all $k \in [K]$, we obtain:

$$\begin{aligned}
\mathbb{E}[f(x^{(K+1)})] - f(x^*) \geq (1 - \frac{1}{K})^K \big(\mathbb{E}[f(\underbrace{x^{(1)}}_{=0})] - f(x^*)\big) \\
- G_{\mathcal{D}} - \frac{LR^2}{2K}.
\end{aligned}$$

Rearranging the terms and using the inequality $(1-\frac{1}{K})^K \leq \frac{1}{e}$, we obtain the desired result. □

Lemma 2. *L-smoothness parameter of the multilinear extension f of the submodular set function $F: 2^V \to \mathbb{R}$ is bounded as $L \leq m_F$, where $m_F = \max_{i \in V} F(\{i\})$. Moreover, for submodular maximization over the matroid polytope $P(\mathcal{M})$, $R \leq \mathrm{rank}(\mathcal{M})$ holds, and for submodular maximization subject to a knapsack constraint, we have $R \leq \frac{1}{c_{\min}}$.*

Assume that the submodular objective function is Δ-sensitive (defined in Section 2). The following lemma provides the performance guarantees of the DPCG algorithm under Laplace noise distribution.

Lemma 3. *If $\mathcal{D} = \mathrm{Lap}^{|V|}(\lambda)$, where $\mathrm{Lap}^{|V|}(\lambda)$ denotes a distribution over $\mathbb{R}^{|V|}$ such that each coordinate is drawn i.i.d. from the Laplace distribution with p.d.f. $f(z|\lambda) = \frac{1}{2\lambda}\exp(-\frac{|z|}{\lambda})$, $\forall z \in \mathbb{R}$, setting $\lambda = \frac{2K|V|\Delta}{\epsilon}$, the following holds in expectation:*

$$\mathbb{E}[f(x)] \geq (1-\frac{1}{e})f(x^*) - \frac{LR^2}{2K} - \mathcal{O}(\frac{RK|V|\ln(|V|)\Delta}{\epsilon}),$$

where $R = \mathrm{rank}(\mathcal{M})$ and $R = \frac{1}{c_{\min}}$ for matroid and knapsack constraints respectively. Also, with probability at least $1-\frac{1}{K}$, we have:

$$f(x) \geq (1-\frac{1}{e})f(x^*) - \frac{LR^2}{2K} - \mathcal{O}(\frac{RK|V|\ln(K|V|)\Delta}{\epsilon}).$$

Moreover, Algorithm 1 preserves ϵ-differential privacy.

Proof. In order to analyze the differential privacy of the proposed algorithm, let f_D and $f_{D'}$ be the multilinear extension of monotone submodular set functions F_D and $F_{D'}$ associated with neighboring datasets D and D'. Using the definition of multilinear extension, we can write:

$$\begin{aligned}
&\|\nabla f_D(x) - \nabla f_{D'}(x)\|_1 = \sum_{i=1}^{|V|} |\nabla_i f_D(x) - \nabla_i f_{D'}(x)| \\
&= \sum_{v \in V} |\mathbb{E}_{S \sim x}\big[F_D(S \cup \{v\}) - F_D(S \setminus \{v\}) \\
&- F_{D'}(S \cup \{v\}) + F_{D'}(S \setminus \{v\})\big]| \\
&\leq \sum_{v \in V} \mathbb{E}_{S \sim x}\big[|F_D(S \cup \{v\}) - F_{D'}(S \cup \{v\})| \\
&+ |F_{D'}(S \setminus \{v\}) - F_D(S \setminus \{v\})|\big] \\
&\leq 2|V|\Delta.
\end{aligned}$$

Let $\lambda = \frac{2K|V|\Delta}{\epsilon}$. We show that for each $k \in [K]$, $\nabla f(x^{(k)}) + Y^{(k)}$ is $(\frac{\epsilon}{K})$-differentially private. Considering the immunity of differential privacy to post-processing and using the basic composition theorem, we can conclude that the proposed algorithm is ϵ-differentially private. We have:

$$\begin{aligned}
&\frac{\mathbb{P}(\nabla f_D(x) + Y^{(k)} = z)}{\mathbb{P}(\nabla f_{D'}(x) + Y^{(k)} = z)} = \prod_{i=1}^{|V|} \frac{\exp\big(-\frac{\epsilon|z_i - \nabla_i f_D(x)|}{2K|V|\Delta}\big)}{\exp\big(-\frac{\epsilon|z_i - \nabla_i f_{D'}(x)|}{2K|V|\Delta}\big)} \\
&= \prod_{i=1}^{|V|} \exp\big(\frac{\epsilon(|z_i - \nabla_i f_{D'}(x)| - |z_i - \nabla_i f_D(x)|)}{2K|V|\Delta}\big) \\
&\leq \exp\big(\frac{\epsilon\|\nabla f_{D'}(x) - \nabla f_D(x)\|_1}{2K|V|\Delta}\big) \\
&\leq \exp(\frac{\epsilon}{K}).
\end{aligned}$$

Hence, $\nabla f(x^{(k)}) + Y^{(k)}$ is $(\frac{\epsilon}{K})$-differentially private. We now find an upper bound for $G_{\mathcal{D}}$. Using the definition of the matroid polytope $P(\mathcal{M})$, for all $x \in P(\mathcal{M})$, we have $\|x\|_1 \leq \mathrm{rank}(\mathcal{M})$, where $\mathrm{rank}(\mathcal{M})$ is the rank of the matroid constraint. Therefore, we have:

$$\begin{aligned}
\max_{x \in P(\mathcal{M})} \langle Y, x\rangle - \min_{x \in P(\mathcal{M})} \langle Y, x\rangle &\leq \|x\|_1\|Y\|_\infty + \|x\|_1\|Y\|_\infty \\
&= 2\|x\|_1\|Y\|_\infty \\
&\leq 2\mathrm{rank}(\mathcal{M})\|Y\|_\infty.
\end{aligned}$$

Thus, we have $G_{\mathcal{D}} \leq 2\mathrm{rank}(\mathcal{M})\mathbb{E}_{Y \sim \mathrm{Lap}^{|V|}(\lambda)}\|Y\|_\infty$. Note that in the case of knapsack constraint $c^T x \leq 1$, we have $c_{\min}\|x\|_1 \leq c^T x \leq 1$ and thus, $\|x\|_1 \leq \frac{1}{c_{\min}}$ holds. Therefore, we have $G_{\mathcal{D}} \leq \frac{2}{c_{\min}}\mathbb{E}_{Y \sim \mathrm{Lap}^{|V|}(\lambda)}\|Y\|_\infty$ under the knapsack constraint. For the Laplace random vector $Y \sim \mathrm{Lap}^{|V|}(\lambda)$, we have:

$$\mathbb{E}\|Y\|_\infty \leq \mathcal{O}\big(\lambda \ln(|V|)\big),$$

$$\mathbb{P}\big(\|Y\|_\infty \leq \sqrt{10}\lambda\ln(K|V|)\big) \geq 1 - \frac{1}{K^2}.$$

Therefore, using the union bound over $k \in [K]$, we can obtain the result as stated. □

Alternatively, we can use the Gaussian noise to ensure differential privacy. The analysis for this case is provided in the Appendix. Compared to the Laplace noise, the additive factor in the approximation guarantee using the Gaussian noise is smaller by an order of $\sqrt{|V|\ln(|V|)}$. However, this improved accuracy comes at the price of achieving (ϵ, δ)-differential privacy as opposed to ϵ-differential privacy using the Laplace noise. We can also use the advanced composition theorem to ensure differential privacy. In particular, for $0 < \epsilon < 1$ and $\delta > 0$, if $\nabla f(x^{(k)}) + Y^{(k)}$, $\forall k \in [K]$ is $(\frac{\epsilon}{2\sqrt{2K\ln(1/\delta)}})$-differentially private using the Laplace noise, we can use the advanced composition theorem for all $k \in [K]$ and ensure (ϵ, δ)-differential privacy of Algorithm 1. The result is summarized in the lemma below.

Lemma 4. *If $\mathcal{D} = \mathrm{Lap}^{|V|}(\lambda)$, setting $\lambda = \frac{4\sqrt{2K\ln(1/\delta)}|V|\Delta}{\epsilon}$, the following holds in expectation:*

$$\mathbb{E}[f(x)] \geq (1-\frac{1}{e})f(x^*) - \frac{LR^2}{2K} - \mathcal{O}(\frac{R\sqrt{K\ln(1/\delta)}|V|\ln(|V|)\Delta}{\epsilon}).$$

where $R = \mathrm{rank}(\mathcal{M})$ and $R = \frac{1}{c_{\min}}$ for matroid and knapsack constraints respectively. Also, with probability at least $1-\frac{1}{K}$, we have:

$$f(x) \geq (1-\frac{1}{e})f(x^*) - \frac{LR^2}{2K} - \mathcal{O}(\frac{R\sqrt{K\ln(1/\delta)}|V|\ln(K|V|)\Delta}{\epsilon}).$$

Moreover, Algorithm 1 preserves (ϵ,δ)-differential privacy.

Combining the result of Lemma 2 and 3, we provide the approximation and privacy guarantee of Algorithm 1 below.

Theorem 3. *Setting $K = \mathcal{O}(\sqrt{\frac{\epsilon\ \mathrm{rank}(\mathcal{M})}{|V|\ln(|V|)\Delta}})$, Algorithm 1 is ϵ-differentially private and has the following approximation guarantees for matroid and knapsack constraints respectively:*

$$\mathbb{E}[f(x)] \geq (1-\frac{1}{e})f(x^*) - \mathcal{O}(\sqrt{\frac{\Delta(\mathrm{rank}(\mathcal{M}))^3|V|\ln(|V|)}{\epsilon}}),$$

$$\mathbb{E}[f(x)] \geq (1-\frac{1}{e})f(x^*) - \mathcal{O}(\sqrt{\frac{\Delta|V|\ln(|V|)}{(c_{\min})^3\epsilon}}).$$

3.2 κ-Differentially Private Continuous Greedy (κ-DPCG) algorithm

For a general monotone continuous function $f : \mathbb{R}^m \to \mathbb{R}$, we define:

$$c_f = 1 - \min_{i\in[m]}\min_{x,y}\frac{\nabla_i f(y)}{\nabla_i f(x)}.$$

Note that $c_f \leq 1$ due to monotonocity of f and $c_f \geq 0$ by setting $x = y$ in the definition. If $f : [0,1]^V \to \mathbb{R}_+$ is the multilinear extension of a normalized monotone submodular set function $F : 2^V \to \mathbb{R}_+$, we can write:

$$\begin{aligned}
c_f &= 1 - \min_{i\in V}\min_{x,y}\frac{\nabla_i f(y)}{\nabla_i f(x)} \\
&\overset{(a)}{=} 1 - \min_{i\in V}\min_{y}\frac{\nabla_i f(y)}{\nabla_i f(0)} \\
&\overset{(b)}{=} 1 - \min_{i\in V}\min_{y}\frac{\mathbb{E}_{S\sim y}[F(S\cup\{i\}) - F(S\setminus\{i\})]}{F(\{i\}) - F(\{\})} \\
&\overset{(c)}{=} 1 - \min_{i\in V}\frac{F(V) - F(V\setminus\{i\})}{F(\{i\}) - F(\{\})} \\
&= \kappa_F,
\end{aligned}$$

Algorithm 2 κ-Differentially Private Continuous Greedy (κ-DPCG) algorithm

```
Input: $K$, $\lambda > 0$, the constraint set $P$, the multilinear extension $f : [0,1]^{|V|} \to \mathbb{R}$ of the monotone submodular set function $F : 2^V \to \mathbb{R}$, and the noise distribution $\mathcal{D}$.
Initialization: $x^{(1)} = 0$.
for $k = 1, 2, \dots, K$ do
  Draw $Y^{(k)} \sim \mathcal{D}$.
  Set $v_k = \arg\max_{v\in P:\ell^T v\geq\lambda}\langle v, \nabla g(x^{(k)}) + Y^{(k)}\rangle$.
  Set $x^{(k+1)} = x^{(k)} + \frac{1}{K}v_k$.
end for
Output: $x = x^{(K+1)}$.
```

where κ_F is the total curvature of the submodular set function F. (a) follows from DR-submodularity of f, (b) uses the definition of the multilinear extension and (c) is due to submodularity of F. Therefore, the parameter c_f extends the notion of curvature from submodular set functions to general monotone continuous functions and could be of independent interest. We propose the κ-DPCG algorithm in Algorithm 2. Let $\lambda = \ell^T x^*$ where $\ell_i = \min_x \nabla_i f(x) = F(V) - F(V\setminus\{i\})$ and $x^* = \mathbf{1}_{S^*}$ be the optimal point corresponding to the optimal solution $S^* \subseteq V$. Compared to the DPCG algorithm, the κ-DPCG algorithm is different in two important respects:

1. The DPCG algorithm is applied to the function $g(x) = f(x) - \ell^T x$. Note that the function g is normalized monotone DR-submodular as well.

2. The linear maximization step is performed over the intersection of the constraint set P and $\{x : \ell^T x \geq \lambda\}$.

Similar to Lemma 1, We provide the approximation guarantee of the κ-DPCG algorithm below.

Lemma 5. *If the multilinear extension f is L-smooth with respect to $\|\cdot\|_1$ and the diameter of P is denoted by $R = \max_{x\in P}\|x\|_1$, Algorithm 2 outputs $x = x^{(K+1)}$ such that the following holds:*

$$\mathbb{E}[f(x)] \geq (1-\frac{\kappa_F}{e})f(x^*) - G_{\mathcal{D}} - \frac{LR^2}{2K},$$

where $G_{\mathcal{D}} := \mathbb{E}_{Y\sim\mathcal{D}}\big[\max_{x\in P}\langle Y, x\rangle - \min_{x\in P}\langle Y, x\rangle\big]$ and κ_F is the total curvature of the submodular objective function F.

Therefore, all the analysis from Section 3.1 can be performed here as well and thus, the κ-DPCG algorithm maintains the same privacy and utility guarantees as the DPCG algorithm except for its improved approximation ratio $1 - \frac{\kappa_F}{e}$.

Note that although we have used $\lambda = \ell^T x^*$ in Algorithm 2, the optimal value x^* is generally unknown and therefore, we have to guess the value of λ in practice. Using the definition of ℓ and submodularity of F, we have $\ell_i \leq m_F \ \forall i \in V$. Therefore, $\lambda \leq |V| m_F$ holds. We discretize the interval $[0, m_F]$ with $\mathcal{O}(\frac{1}{\gamma})$ points of the form $i\gamma m_F$ for $0 \leq i \leq \frac{1}{\gamma}$, along with $\mathcal{O}(\frac{1}{\gamma}|V|\ln(|V|))$ points of the form $(1+\frac{\gamma}{|V|})^i m_F$ for $0 \leq i \leq \log_{1+\frac{\gamma}{|V|}} |V|$ to fill the interval $[m_F, |V| m_F]$. We then run Algorithm 2 using each point as a guess for λ and we output the best solution found, i.e., the solution with the highest utility. If $\lambda \in [0, m_F]$, we should have $\lambda \geq \hat{\lambda} \geq \lambda - \gamma m_F$ using one of the guesses $\hat{\lambda}$ in the interval $[0, m_F]$. Otherwise, if $\lambda \in [m_F, |V| m_F]$, consider the largest guess $\hat{\lambda}$ in the interval $[m_F, |V| m_F]$ satisfying $\lambda \geq \hat{\lambda}$. We have:

$$\lambda \geq \hat{\lambda} \geq \lambda(1+\frac{\gamma}{|V|})^{-1} \geq \lambda(1-\frac{\gamma}{|V|}) \geq \lambda - \gamma m_F,$$

where the last inequality uses $\lambda \leq |V| m_F$. Therefore, in practice, the approximation guarantee of the κ-DPCG algorithm has an additional γm_F error term which can be tuned through choosing γ.

4 Differentially private online submodular maximization

In this section, we study the following general protocol of online submodular maximization in the context of differential privacy: There is a fixed constraint set $(V, \mathcal{I})$ which could either be a matroid or a knapsack constraint. At each iteration $t \in [T]$, the online algorithm chooses $S_t \in \mathcal{I}$. Upon committing to this choice, a normalized monotone submodular set function F_t is revealed and the algorithm receives the payoff $F_t(S_t)$. The goal is to minimize the $(1-\frac{1}{e})$-regret defined below:

$$R_T = (1-\frac{1}{e}) \max_{S \in \mathcal{I}} \sum_{t=1}^{T} F_t(S) - \sum_{t=1}^{T} F_t(S_t),$$

where $1-\frac{1}{e}$ is the optimal polynomial time approximation ratio for offline monotone submodular maximization subject to matroid or knapsack constraints. We propose the Differentially Private Meta Frank-Wolfe (DPMFW) algorithm for online submodular maximization which exploits Algorithm 1 of Agarwal and Singh (2017) for differentially private online linear optimization as a sub-routine. The algorithm is presented in Algorithm 3. The DPMFW algorithm is applied to the multilinear extensions $\{f_t\}_{t=1}^T$ of the discrete monotone submodular utility functions $\{F_t\}_{t=1}^T$. At round $t \in [T]$, similar to the DPCG algorithm, the DPMFW algorithm outputs the average of K points $\{v_t^{(k)}\}_{k=1}^K$ in the constraint set and hence, $x_t \in P$ holds. However, since the utility function F_t remains unknown until the algorithm commits to a choice S_t, we instead run K instances $\{\mathcal{E}_k\}_{k=1}^K$ of the differentially private online linear optimization algorithm of Agarwal and Singh (2017) to mimic the K Frank-Wolfe updates of the DPCG algorithm. $\{\mathcal{E}_k\}_{k=1}^K$ combine the well-known Follow the Regularized Leader (FTRL) algorithm for online linear optimization with the Tree Based Aggregation Protocol (TBAP) of Dwork et al. (2010); Jain et al. (2012) for maintaining differentially private partial sums of linear utility vectors arriving online. See the Appendix for a more detailed presentation of Algorithm 1 of Agarwal and Singh (2017). We provide the regret bound and privacy guarantee of the DPMFW algorithm below.

Algorithm 3 Differentially Private Meta Frank-Wolfe (DPMFW) algorithm

```
**Input:** $K$, the constraint set $P$.
**Output:** $\{x_t\}_{t=1}^T$
Initialize $K$ instances $\{\mathcal{E}_k\}_{k=1}^K$ of the $(\frac{\epsilon}{2\sqrt{2K\ln(1/\delta)}})$-differentially private algorithm of Agarwal and Singh (2017) with noise distribution $\mathcal{D}$ and regularizer $g(x) = \sum_{i=1}^{|V|} x_i \ln(x_i)$ for online linear optimization over $P$.
for $t = 1, \ldots, T$ do
    $x_t^{(1)} = 0$.
    for $k = 1, 2, \ldots, K$ do
        Let $v_t^{(k)}$ be the output of $\mathcal{E}_k$ for round $t$.
        Set $x_t^{(k+1)} = x_t^{(k)} + \frac{1}{K} v_t^{(k)}$.
    end for
    Play $x_t = x_t^{(K+1)}$.
    for $k = 1, 2, \ldots, K$ do
        Feedback $\nabla f_t(x_t^{(k)})$ to $\mathcal{E}_k$ as the linear utility vector observed at round $t$.
    end for
end for
```

Theorem 4. *Let $0 < \epsilon < 1$ and $\delta > 0$. If $\mathcal{D} = \mathrm{Lap}^{|V|}(\lambda)$, where $\mathrm{Lap}^{|V|}(\lambda)$ is a distribution over $\mathbb{R}^{|V|}$ such that each coordinate is drawn i.i.d. from the Laplace distribution with p.d.f. $f(z|\lambda) = \frac{1}{2\lambda}\exp(-\frac{|z|}{\lambda}) \ \forall z \in \mathbb{R}$, setting $\lambda = \frac{2m_F|V|\ln T\sqrt{2K\ln(1/\delta)}}{\epsilon}$ and $K = \mathcal{O}(\sqrt{T})$, Algorithm 3 is (ϵ, δ)-differentially private and has the following expected regret bound for matroid and knapsack constraints respectively:*

$$\mathbb{E}[R_T] \leq \mathcal{O}(\mathrm{rank}(\mathcal{M})\sqrt{T\ln|V|}) + \tilde{\mathcal{O}}(\frac{(\mathrm{rank}(\mathcal{M}))^{3/2}|V|T^{1/4}\sqrt{\ln(1/\delta)}}{\epsilon}),$$

$$\mathbb{E}[R_T] \leq \mathcal{O}(\frac{\sqrt{T\ln|V|}}{c_{\min}}) + \tilde{\mathcal{O}}(\frac{|V|T^{1/4}\sqrt{\ln(1/\delta)}}{(c_{\min})^{3/2}\epsilon}).$$

The above theorem shows that if $\frac{\sqrt{\ln(1/\delta)}}{\epsilon} \leq \mathcal{O}(T^{1/4})$

holds, we can obtain a regret bound of $\mathcal{O}(\sqrt{T})$ which matches the optimal regret bound for online submodular maximization in the non-private setting.
We can alternatively use the Gaussian noise as the noise distribution $\mathcal{D}$. The analysis in this setting is provided in the Appendix.

5 Conclusion

In this paper, we studied the maximization of non-negative monotone submodular set functions, subject to matroid or knapsack constraints, in both offline and online settings. We proposed differentially private algorithms with optimal approximation ratios which are faster (i.e., less query complexity) and have improved accuracy compared to the prior works.

References

David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '03, page 137–146, New York, NY, USA, 2003. Association for Computing Machinery. ISBN 1581137370. doi: 10.1145/956750.956769. URL https://doi.org/10.1145/956750.956769.

Andreas Krause and Carlos Guestrin. Near-optimal observation selection using submodular functions. In *AAAI*, volume 7, pages 1650–1654, 2007.

Pushmeet Kohli, M Pawan Kumar, and Philip HS Torr. P^3 & beyond: Move making algorithms for solving higher order functions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(9): 1645–1656, 2008.

Yuri Y Boykov and M-P Jolly. Interactive graph cuts for optimal boundary & region segmentation of objects in nd images. In *Proceedings eighth IEEE international conference on computer vision. ICCV 2001*, volume 1, pages 105–112. IEEE, 2001.

Hui Lin and Jeff Bilmes. A class of submodular functions for document summarization. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 510–520, 2011.

Katrin Kirchhoff and Jeff Bilmes. Submodularity for data selection in machine translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 131–141, 2014.

Gruia Calinescu, Chandra Chekuri, Martin Pál, and Jan Vondrák. Maximizing a Monotone Submodular Function Subject to a Matroid Constraint. *SIAM Journal on Computing*, 40(6):1740–1766, January 2011a. ISSN 0097-5397, 1095-7111. doi: 10.1137/080733991.

C. Chekuri, J. Vondrak, and R. Zenklusen. Dependent randomized rounding via exchange properties of combinatorial structures. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pages 575–584, 2010.

Ariel Kulik, Hadas Shachnai, and Tami Tamir. Maximizing submodular set functions subject to multiple linear constraints. In *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '09, page 545–554, USA, 2009. Society for Industrial and Applied Mathematics.

Ariel Kulik, Hadas Shachnai, and Tami Tamir. Approximations for monotone and nonmonotone submodular maximization with knapsack constraints. *Math. Oper. Res.*, 38(4):729–739, November 2013. ISSN 0364-765X. doi: 10.1287/moor.2013.0592. URL https://doi.org/10.1287/moor.2013.0592.

Marko Mitrovic, Mark Bun, Andreas Krause, and Amin Karbasi. Differentially private submodular maximization: Data summarization in disguise. In Doina Precup and Yee Whye Teh, editors, *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 2478–2487, International Convention Centre, Sydney, Australia, 06–11 Aug 2017. PMLR. URL http://proceedings.mlr.press/v70/mitrovic17a.html.

Baharan Mirzasoleiman, Morteza Zadimoghaddam, and Amin Karbasi. Fast distributed submodular cover: Public-private data summarization. In *Advances in Neural Information Processing Systems*, pages 3594–3602, 2016.

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

Andrew An Bian, Baharan Mirzasoleiman, Joachim Buhmann, and Andreas Krause. Guaranteed Nonconvex Optimization: Submodular Maximization over Continuous Domains. In Aarti Singh and Jerry Zhu, editors, *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, volume 54 of *Proceedings of Machine Learning Research*, pages 111–120, Fort Lauderdale, FL, USA, 20–22 Apr 2017. PMLR. URL http://proceedings.mlr.press/v54/bian17a.html.

Gruia Calinescu, Chandra Chekuri, Martin Pál, and Jan Vondrák. Maximizing a monotone submodular function subject to a matroid constraint. *SIAM Journal on Computing*, 40(6):1740–1766, 2011b. doi: 10.

1137/080733991. URL https://doi.org/10.1137/080733991.

Omid Sadeghi, Prasanna Raut, and Maryam Fazel. A single recipe for online submodular maximization with adversarial or stochastic constraints. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 14712–14723. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper/2020/file/a8e5a72192378802318bf51063153729-Paper.pdf.

Prasanna Sanjay Raut, Omid Sadeghi, and Maryam Fazel. Online dr-submodular maximization with stochastic cumulative constraints. *arXiv preprint arXiv:2005.14708*, 2020.

George L Nemhauser, Laurence A Wolsey, and Marshall L Fisher. An analysis of approximations for maximizing submodular set functions—i. *Mathematical programming*, 14(1):265–294, 1978.

Maxim Sviridenko, Jan Vondrák, and Justin Ward. Optimal approximation for submodular and supermodular optimization with bounded curvature. *Mathematics of Operations Research*, 42(4):1197–1218, 2017.

Lin Chen, Hamed Hassani, and Amin Karbasi. Online continuous submodular maximization. volume 84 of *Proceedings of Machine Learning Research*, pages 1896–1905, Playa Blanca, Lanzarote, Canary Islands, 09–11 Apr 2018. PMLR. URL http://proceedings.mlr.press/v84/chen18f.html.

Anupam Gupta, Katrina Ligett, Frank McSherry, Aaron Roth, and Kunal Talwar. *Differentially Private Combinatorial Optimization*, pages 1106–1125. doi: 10.1137/1.9781611973075.90. URL https://epubs.siam.org/doi/abs/10.1137/1.9781611973075.90.

Anamay Chaturvedi, Huy Nguyen, and Lydia Zakynthinou. Differentially private decomposable submodular maximization. *arXiv preprint arXiv:2005.14717*, 2020.

Akbar Rafiey and Yuichi Yoshida. Fast and private submodular and k-submodular functions maximization with matroid constraints. In *Proceedings of the 37th International Conference on Machine Learning*, pages 6443–6453. 2020.

Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '07, page 94–103, USA, 2007. IEEE Computer Society. ISBN 0769530109. doi: 10.1109/FOCS.2007.41. URL https://doi.org/10.1109/FOCS.2007.41.

Cynthia Dwork, Moni Naor, Toniann Pitassi, and Guy N. Rothblum. Differential privacy under continual observation. In *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*, STOC '10, page 715–724, New York, NY, USA, 2010. Association for Computing Machinery. ISBN 9781450300506. doi: 10.1145/1806689.1806787. URL https://doi.org/10.1145/1806689.1806787.

Prateek Jain, Pravesh Kothari, and Abhradeep Thakurta. Differentially private online learning. volume 23 of *Proceedings of Machine Learning Research*, pages 24.1–24.34, Edinburgh, Scotland, 25–27 Jun 2012. JMLR Workshop and Conference Proceedings. URL http://proceedings.mlr.press/v23/jain12.html.

Abhradeep Guha Thakurta and Adam Smith. (nearly) optimal algorithms for private online learning in full-information and bandit settings. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 26*, pages 2733–2741. Curran Associates, Inc., 2013.

Naman Agarwal and Karan Singh. The price of differential privacy for online learning. volume 70 of *Proceedings of Machine Learning Research*, pages 32–40, International Convention Centre, Sydney, Australia, 06–11 Aug 2017. PMLR. URL http://proceedings.mlr.press/v70/agarwal17a.html.

Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3–4):211–407, August 2014. ISSN 1551-305X. doi: 10.1561/0400000042. URL https://doi.org/10.1561/0400000042.