Online Submodular Maximization via Online Convex Optimization

Tareq Si Salem[1], **Gözde Özcan**[1], **Iasonas Nikolaou**[2], **Evimaria Terzi**[2], **Stratis Ioannidis**[1]

[1]Northeastern University, Boston, MA, USA
[2]Boston University, Boston, MA, USA
[1]{sisalem, gozcan, ioannidis}@ece.neu.edu, [2]{nikolaou, evimaria}@bu.edu

Abstract

We study monotone submodular maximization under general matroid constraints in the online setting. We prove that online optimization of a large class of submodular functions, namely, weighted threshold potential functions, reduces to online convex optimization (OCO). This is precisely because functions in this class admit a concave relaxation; as a result, OCO policies, coupled with an appropriate rounding scheme, can be used to achieve sublinear regret in the combinatorial setting. We show that our reduction extends to many different versions of the online learning problem, including the dynamic regret, bandit, and optimistic-learning settings.

Introduction

In online submodular optimization (OSM) (Streeter, Golovin, and Krause 2009), submodular reward functions chosen by an adversary are revealed over several rounds. In each round, a decision maker first commits to a set satisfying, e.g., matroid constraints. Subsequently, the reward function is revealed and evaluated over this set. The objective is to minimize α-regret, i.e., the difference of the cumulative reward accrued from the one attained by a static set, selected by an α-approximation algorithm operating in hindsight. OSM has received considerable interest recently, both in the full information (Harvey, Liaw, and Soma 2020; Matsuoka, Ito, and Ohsaka 2021; Streeter, Golovin, and Krause 2009; Niazadeh et al. 2021) and bandit setting (Niazadeh et al. 2021; Matsuoka, Ito, and Ohsaka 2021; Wan et al. 2023), where only the reward values (rather than the entire functions) are revealed.

Online convex optimization (OCO) studies a similar online setting in which reward functions are concave, and decisions are selected from a compact convex set. First proposed by Zinkevich (2003), who showed that projected gradient ascent attains sublinear regret, OCO generalizes previous online problems like prediction with expert advice (Littlestone and Warmuth 1994), and has become widely influential in the learning community (Hazan 2016; Shalev-Shwartz 2012; McMahan 2017). Its success is evident from the multitude of OCO variants in literature: in *dynamic regret* OCO (Zinkevich 2003; Besbes, Gur, and Zeevi 2015; Jadbabaie et al. 2015; Mokhtari et al. 2016), the regret is evaluated w.r.t. an optimal comparator sequence instead of an optimal static decision in hindsight. *Optimistic* OCO (Rakhlin and Sridharan 2013; Dekel et al. 2017; Mohri and Yang 2016) takes advantage of benign sequences, in which reward functions are predictable: the decision maker attains tighter regret bounds when predictions are correct, falling back to the existing OCO regret guarantees when predictions are unavailable or inaccurate. *Bandit* OCO algorithms (Hazan and Levy 2014; Kleinberg 2004; Flaxman, Kalai, and McMahan 2005) study the aforementioned bandit setting, where again only rewards (i.e., function evaluations) are observed.

We make the following contributions:

- We provide a methodology for reducing OSM to OCO, when submodular functions selected by the adversary are bounded from above and below by concave relaxations and coupled with an opportune rounding. We prove that algorithms and regret guarantees in the OCO setting transfer to α-regret guarantees in the OSM setting, via a transformation that we introduce. Ratio α is determined by how well concave relaxations approximate the original submodular functions.
- We show that the above condition is satisfied by a wide class of submodular functions, namely, *weighted threshold potential* (WTP) functions. This class strictly generalizes weighted coverage functions (Karimi et al. 2017; Stobbe and Krause 2010), and includes many important applications, including influence maximization (Kempe, Kleinberg, and Tardos 2003), facility location (Krause and Golovin 2014; Frieze 1974), cache networks (Ioannidis and Yeh 2016; Li et al. 2021), similarity caching (Si Salem, Neglia, and Carra 2022), demand forecasting (Ito and Fujimaki 2016), and team formation (Li et al. 2018), to name a few.
- We show our reduction also extends to the dynamic regret and optimistic settings, reducing such full-information OSM settings to the respective OCO variants. To the best of our knowledge, our resulting algorithms are the first to come with guarantees for the dynamic regret and optimistic settings in the context of OSM with general matroid constraints. Finally, we also provide a different reduction for the bandit setting, again for all three (static, dynamic regret, and optimistic) variants; this reduction applies to general submodular functions, but is restricted to partition matroids.

Related Work

Offline Submodular Maximization and Relaxations of Submodular Functions. Continuous relaxations of submodular functions play a prominent role in submodular maximization. The so-called continuous greedy algorithm (Calinescu et al. 2011) maximizes the multilinear relaxation of a submodular objective over the convex hull of a matroid, using a variant of the Frank-Wolfe algorithm (Frank and Wolfe 1956). The fractional solution is then rounded via pipage (Ageev and Sviridenko 2004) or swap rounding (Chekuri, Vondrák, and Zenklusen 2010), which we also use. The multilinear relaxation is not convex but is continuous DR-submodular (Bach 2019; Bian et al. 2017), and continuous greedy comes with a $1 - \frac{1}{e}$ approximation guarantee. However, the multilinear relaxation is generally not tractable and is usually estimated via sampling. Ever since the seminal paper by Ageev and Sviridenko (Ageev and Sviridenko 2004), several works have exploited the existence of concave relaxations of weighted coverage functions (e.g., (Karimi et al. 2017; Ioannidis and Yeh 2016)), a strict subset of the threshold potential functions we consider here. For coverage functions, a version of our "sandwich" property (Asm. 2) follows directly by the Goemans & Williamson inequality (Goemans and Williamson 1994), which we also use. Both in the standard (Ageev and Sviridenko 2004) and stochastic offline (Karimi et al. 2017; Ioannidis and Yeh 2016) submodular maximization setting, in which the objective is randomized, exploiting concave relaxations of coverage functions yields significant computational dividends, as it eschews any sampling required for estimating the multilinear relaxation. We depart from both by considering a much broader class than coverage functions and studying the online/no-regret setting.

OSM via Regret Minimization. Several online algorithms have been proposed for maximizing general submodular functions (Niazadeh et al. 2021; Harvey, Liaw, and Soma 2020; Matsuoka, Ito, and Ohsaka 2021; Streeter, Golovin, and Krause 2009) under different matroid constraints. There has also been recent work (Chen, Hassani, and Karbasi 2018; Zhang et al. 2019, 2022) on the online maximization of continuous DR-submodular functions (Bach 2019). Proposed algorithms are applicable to our setting, because the multilinear relaxation is DR-submodular, and guarantees can be extended to matroid constraint sets again through rounding (Chekuri, Vondrák, and Zenklusen 2010), akin to the approach we follow here. Also pertinent is the work by Kakade, Kalai, and Ligett (2007): their proposed online algorithm operates over reward functions that can be decomposed as the weighted sum of finitely many (non-parametric) reference functions—termed Linearly Weighted Decomposable (LWD); the adversary selects only the weights. Applied to OSM, this is a more restrictive function class than the ones we study.

We compare these algorithms to our work in Table 1 in (Si Salem et al. 2023). In the full information setting, the OSM algorithm by Harvey, Liaw, and Soma (2020) has a slightly tighter α-regret than us, but also a much higher computational complexity. We attain the same or better regret than DR-S (Chen, Hassani, and Karbasi 2018; Zhang et al. 2019, 2022) and remaining algorithms that either operate on restricted constraint sets (Niazadeh et al. 2021; Matsuoka, Ito, and Ohsaka 2021; Streeter, Golovin, and Krause 2009) or on the much more restrictive LWD class (Kakade, Kalai, and Ligett 2007). Most importantly, our work generalizes to the dynamic and optimistic settings. With the exception of Matsuoka, Ito, and Ohsaka (2021), who study dynamic regret restricted to uniform matroids, our work provides the first guarantees for OSM in the dynamic and optimistic settings under general matroid constraints.

OSM in the Bandit Setting. Our reduction to OCO in the bandit setting extends the analysis by Wan et al. (2023), who provide a reduction to just FTRL in the static setting, under general submodular functions and partition matroid constraints. We generalize this to any OCO algorithm and to the dynamic and optimistic settings. Interestingly, Wan et al. (2023) conjecture that no sublinear regret algorithm exists for general submodular functions under general matroid constraints in the bandit setting. We compare to bounds attained by Wan et al. (2023) and other bandit algorithms for OSM (Niazadeh et al. 2021; Streeter, Golovin, and Krause 2009; Zhang et al. 2019; Kakade, Kalai, and Ligett 2007) in Table 4 in (Si Salem et al. 2023). Our main contribution is again the extension to the dynamic and optimistic settings.

Technical Preliminary

Submodularity and Matroids. Given a ground set $V = [n] \triangleq \{1, 2, \dots, n\}$, a set function $f : 2^V \to \mathbb{R}$ is *submodular* if $f(S \cup \{i, j\}) - f(S \cup \{j\}) \leq f(S \cup \{i\}) - f(S)$ for all $S \subseteq V$ and $i, j \in V \setminus S$ and *monotone* if $f(A) \leq f(B)$ for all $A \subseteq B \subseteq 2^V$. A *matroid* is a pair $\mathcal{M} = (V, \mathcal{I})$, where $\mathcal{I} \subseteq 2^V$, for which the following holds: (1) if $B \in \mathcal{I}$ and $A \subseteq B$, then $A \in \mathcal{I}$, (2) if $A, B \in \mathcal{I}$ and $|A| < |B|$, then there exists an $x \in B \setminus A$ s.t. $A \cup \{x\} \in \mathcal{I}$. The *rank* $r \in \mathbb{N}$ of $\mathcal{M}$ is the cardinality of the largest set in $\mathcal{I}$. With slight abuse of notation, we represent set functions $f : 2^V \to \mathbb{R}$ as functions over $\{0, 1\}^n$: given a set function f and an $\mathbf{x} \in \{0, 1\}^n$, we denote by $f(\mathbf{x})$ as the value $f(\operatorname{supp}(\mathbf{x}))$, where $\operatorname{supp}(\mathbf{x}) \triangleq \{i \in V : x_i \neq 0\} \subseteq V$ is the support of $\mathbf{x}$. Similarly, we treat matroids as subsets of $\{0, 1\}^n$.

Online Learning. In the general protocol of *online learning* (Cesa-Bianchi and Lugosi 2006), a decision-maker makes sequential decisions and incurs rewards as follows: at timeslot $t \in [T]$, where $T \in \mathbb{N}$ is the time horizon, the decision-maker first commits to a decision $\mathbf{x}_t \in \mathcal{X}$ from some set $\mathcal{X}$. Then, a reward function $f_t : \mathcal{X} \to \mathbb{R}_{\geq 0}$ is selected by an adversary from a set of functions $\mathcal{F}$ and revealed to the decision-maker, who accrues a reward $f_t(\mathbf{x}_t)$. The decision $\mathbf{x}_t$ is determined according to a (potentially random) mapping $\mathcal{P}_{\mathcal{X},t} : \mathcal{X}^{t-1} \times \mathcal{F}^{t-1} \to \mathcal{Y}$, i.e.,

$$\mathbf{x}_t = \mathcal{P}_{\mathcal{X},t}\left((\mathbf{x}_s)_{s=1}^{t-1}, (f_s)_{s=1}^{t-1}\right). \tag{1}$$

Let $\boldsymbol{\mathcal{P}}_{\mathcal{X}} = (\mathcal{P}_{\mathcal{X},t})_{t\in[T]}$ be the *online policy* of the decision-maker. We define $\operatorname{regret}_T(\boldsymbol{\mathcal{P}}_{\mathcal{X}})$, the *regret* at horizon T, as:

$$\sup_{(f_t)_{t=1}^T \in \mathcal{F}^T} \left\{ \max_{\mathbf{x}\in\mathcal{X}} \sum_{t=1}^{T} f_t(\mathbf{x}) - \mathbb{E}_{\boldsymbol{\mathcal{P}}_{\mathcal{X}}}\left[\sum_{t=1}^{T} f_t(\mathbf{x}_t)\right]\right\}. \tag{2}$$

We seek policies that attain a sublinear (i.e., $o(T)$) regret; intuitively, such policies perform on average as well as the

static optimum in hindsight. Note that the regret is defined w.r.t. the optimal *fixed* decision $\mathbf{x}$, i.e., the time-invariant decision $\mathbf{x} \in \mathcal{X}$ that would be optimal in hindsight, after the sequence $(f_t)_{t=1}^T$ is revealed. When selecting $\mathbf{x}_t$, the decision-maker has no information about the upcoming reward f_t. Finally, this is the *full-information* setting: at each timeslot, the decision maker observes the entire reward function $f_t(\,\cdot\,)$, rather than just the reward $f_t(\mathbf{x}_t) \in \mathbb{R}_{\geq 0}$.

Deviating from these assumptions is of both practical and theoretical interest. In the *dynamic regret* setting (Zinkevich 2003), the regret is measured w.r.t. a time-variant optimum, appropriately constrained so that changes from one timeslot to the next do not vary significantly. In *learning with optimism* (Rakhlin and Sridharan 2013), additional information is assumed to be available w.r.t. f_t, in the form of so-called predictions. In the *bandit* setting (Auer et al. 1995), the online policy $\boldsymbol{\mathcal{P}}_{\mathcal{X}}$ of the decision maker only has access to rewards $f_t(\mathbf{x}_t) \in \mathbb{R}_{\geq 0}$, as opposed to the entire reward function.

Online Convex Optimization. The *online convex optimization* (OCO) framework (Zinkevich 2003; Hazan 2016) follows the above online learning protocol (1), where (a) the decision space $\mathcal{X}$ is a convex set in $\mathbb{R}^n$, and (b) the set of reward functions $\mathcal{F}$ is a subset of concave functions over $\mathcal{X}$. Formally, OCO operates under the following assumption:

Assumption 1. *Set $\mathcal{X} \subset \mathbb{R}^n$ is convex and compact. The reward functions in $\mathcal{F}$ are all L-Lipschitz concave functions w.r.t. a norm $\|\cdot\|$ over $\mathcal{X}$, for some common $L \in \mathbb{R}_{>0}$.*

There is a rich literature on OCO policies (Hazan 2016); examples include Online Gradient Ascent (OGA) (Hazan 2016), Online Mirror Ascent (OMA) (Bubeck 2011), and Follow-The-Regularized-Leader (FTRL) (McMahan 2017). All three enjoy sublinear regret:

Theorem 1. *Under Asm. 1 OGA, OMA, and FTRL attain $O(\sqrt{T})$ regret.*

Details on all three algorithms and the regret they attain are in Appendix A of Si Salem et al. (2023). Most importantly, the OCO framework generalizes to the dynamic, learning-with-optimism, as well as bandit settings (see extentions section).

Weighted Threshold Potentials. A *threshold potential* (Stobbe and Krause 2010) $\Psi_{b,\mathbf{w},S} : \{0,1\}^n \to \mathbb{R}_{\geq 0}$, also known as a *budget-additive function* (Andelman and Mansour 2004; Dobzinski 2016; Buchfuhrer et al. 2010), is defined as:

$$\Psi_{b,\mathbf{w},S}(\mathbf{x}) \triangleq \min\left\{b, \textstyle\sum_{j\in S} x_j w_j\right\}, \text{ for } \mathbf{x} \in \{0,1\}^n, \quad (3)$$

where $b \in \mathbb{R}_{\geq 0} \cup \{\infty\}$ is a threshold, $S \subseteq V$ is a subset of $V = [n]$, and $\mathbf{w} = (w_j)_{j\in S} \in [0,b]^{|S|}$ is a weight vector bounded by b.[1] The linear combination of threshold potential functions defines the rich class of *weighted threshold potentials* (WTP) (Stobbe and Krause 2010), defined as:

$$f(\mathbf{x}) \triangleq \textstyle\sum_{\ell\in C} c_\ell \Psi_{b_\ell,\mathbf{w}_\ell,S_\ell}(\mathbf{x}), \quad \text{for } \mathbf{x} \in \{0,1\}^n, \quad (4)$$

where C is an arbitrary index set and $c_\ell \in \mathbb{R}_{\geq 0}$, for $\ell \in C$. WTP functions are submodular and monotone (see Appendix B of Si Salem et al. (2023)). We define the *degree* of a WTP function $\Delta_f = \max_{\ell\in C} |S_\ell|$ as the maximum number of variables that a threshold potential Ψ in f depends on.

We give several examples of WTP functions in Appendix B of Si Salem et al. (2023). In short, classic problems such as influence maximization (Kempe, Kleinberg, and Tardos 2003) and facility location (Krause and Golovin 2014; Frieze 1974), resource allocation problems like cache networks (Ioannidis and Yeh 2016; Li et al. 2021) and similarity caching (Si Salem, Neglia, and Carra 2022), as well as demand forecasting (Ito and Fujimaki 2016) and team formation (Li et al. 2018) can all be expressed using WTP functions. Overall, the WTP class is very broad: there exists a hierarchy among submodular functions, including weighted coverage functions (Karimi et al. 2017), weighted cardinality truncations (Dolhansky and Bilmes 2016), and sums of concave functions composed with non-negative modular functions (Stobbe and Krause 2010); all of them are strictly dominated by the WTP class (see Stobbe and Krause (2010), as well as Appendix B of Si Salem et al. (2023)).

Problem Formulation

We consider a combinatorial version of the online learning protocol defined in Eq. (1). In particular, we focus on the case where (a) the decision set is $\mathcal{X} \subseteq \{0,1\}^n$, i.e., the vectorized representation of subsets of $V = [n]$, and (b) the set $\mathcal{F}$ of reward functions comprises set functions over $\mathcal{X}$. Though some of our results (e.g., Thm. 2) pertain to this general combinatorial setting, we are particularly interested in the case where (a) $\mathcal{X}$ is a matroid, and (b) $\mathcal{F}$ is the WTP class, i.e., the set of functions whose form is given by Eq. (4).

Both in the general combinatorial setting and for WTP functions, evaluating the best fixed decision may be computationally intractable even in hindsight, i.e., when all reward functions were revealed. As is customary (see, e.g., (Krause and Golovin 2014)), instead of the regret in Eq. (2), we consider $\alpha\text{-regret}_T(\boldsymbol{\mathcal{P}}_{\mathcal{X}})$, the so-called α-regret, defined as:

$$\sup_{(f_t)_{t=1}^T \in \mathcal{F}^T} \left\{ \max_{\mathbf{x}\in\mathcal{X}} \alpha \sum_{t=1}^T f_t(\mathbf{x}) - \mathbb{E}_{\boldsymbol{\mathcal{P}}_{\mathcal{X}}}\Big[\sum_{t=1}^T f_t(\mathbf{x}_t)\Big]\right\}. \quad (5)$$

Intuitively, we compare the performance of the policy $P_{\mathcal{X}}$ w.r.t. the best polytime(n) α-approximation of the static offline optimum in hindsight. For example, the approximation ratio would be $\alpha = 1 - 1/e$ in the case of submodular set functions maximized over matroids.

Online Submodular Optimization via Online Convex Optimization

The Case of General Set Functions

First, we show how OCO can be leveraged to tackle online learning in the general combinatorial setting, i.e., when $\mathcal{X} \subseteq \{0,1\}^n$ and $\mathcal{F}$ comprises general functions defined over $\mathcal{X}$.

Rounding Augmented OCO Policy. We begin by stating a "sandwich" property that functions in $\mathcal{F}$ should satisfy, so that the reduction to OCO holds. To do so, we first need to introduce the notion of randomized rounding. Let $\mathcal{Y} \triangleq \operatorname{conv}(\mathcal{X})$ be the convex hull of $\mathcal{X}$. A *randomized rounding* is a random map $\Xi : \mathcal{Y} \to \mathcal{X}$, i.e., a map from a fractional $\mathbf{y} \in \mathcal{Y}$

[1] Assumption $w_j \leq b$, $j \in S$, is w.l.o.g., as replacing w_j with $\min\{w_j, b\}$ preserves values of f over $\{0,1\}^n$.

Algorithm 1: Rounding-Augmented OCO (RAOCO) policy

Require: OCO policy $\boldsymbol{\mathcal{P}}_{\mathcal{Y}}$, randomized rounding $\Xi : \mathcal{Y} \to \mathcal{X}$
for $t = 1, 2, \ldots, T$ **do**
  $\mathbf{y}_t \leftarrow \mathcal{P}_{\mathcal{Y},t}\left((\mathbf{y}_s)_{s=1}^{t-1}, (\tilde{f}_s)_{s=1}^{t-1}\right)$
  $\mathbf{x}_t \leftarrow \Xi(\mathbf{y}_t)$
  Receive reward $f_t(\mathbf{x}_t)$
  Reward function f_t is revealed
  Construct $\tilde{f}_t$ from f_t satisfying Asm. 2
end for

and, possibly, a source of randomness to an integral variable $\mathbf{x} \in \mathcal{X}$. We assume that the set $\mathcal{F}$ satisfies the following:

Assumption 2. *(Sandwich Property) There exists an $\alpha \in (0, 1]$, an $L \in \mathbb{R}_{>0}$, and a randomized rounding $\Xi : \mathcal{Y} \to \mathcal{X}$ such that, for every $f : \mathcal{X} \to \mathbb{R}_{\geq 0} \in \mathcal{F}$ there exists a L-Lipschitz concave function $\tilde{f} : \mathcal{Y} \to \mathbb{R}$ s.t.*

$$\tilde{f}(\mathbf{x}) \geq f(\mathbf{x}), \quad \textit{for all } \mathbf{x} \in \mathcal{X}, \textit{ and} \tag{6}$$

$$\mathbb{E}_{\Xi}\left[f(\Xi(\mathbf{y}))\right] \geq \alpha \cdot \tilde{f}(\mathbf{y}), \quad \textit{for all } \mathbf{y} \in \mathcal{Y}. \tag{7}$$

We refer to $\tilde{f}$ as the *concave relaxation* of f. Intuitively, Asm. 2 postulates the existence of such a concave relaxation $\tilde{f}$ that is not "far" from f: Eqs. (6) and (7) imply that $\tilde{f}$ bounds f both from above and below, up to the approximation factor α. Moreover, the upper bound (Eq. (6)) needs to only hold for integral values, while the lower bound (Eq. (7)) needs to only hold in expectation, under an appropriately-defined randomized rounding Ξ.

Armed with this assumption, we can convert any OCO policy $P_{\mathcal{Y}}$ operating over $\mathcal{Y} = \operatorname{conv}(\mathcal{X})$ to a *Randomized-rounding Augmented OCO* (RAOCO) policy, denoted by $\boldsymbol{\mathcal{P}}_{\mathcal{X}}$, operating over $\mathcal{X}$. This transformation (see Alg. 1) uses both the randomized-rounding Ξ, as well as the concave relaxations $(\tilde{f}_s)_{s=1}^{t-1}$ of the functions $(f_s)_{s=1}^{t-1}$ observed so far. At $t \in [T]$, the RAOCO policy amounts to:

$$\mathbf{y}_t = \mathcal{P}_{\mathcal{Y},t}\left((\mathbf{y}_s)_{s=1}^{t-1}, (\tilde{f}_s)_{s=1}^{t-1}\right), \tag{8a}$$

$$\mathbf{x}_t = \Xi(\mathbf{y}_t) \in \mathcal{X}. \tag{8b}$$

In short, the OCO policy $\boldsymbol{\mathcal{P}}_{\mathcal{Y}}$ is first used to generate a new fractional state $\mathbf{y}_t \in \mathcal{Y}$ by applying $\mathcal{P}_{\mathcal{Y},t}$ to the history of concave relaxations. Then, this fractional decision $\mathbf{y}_t$ is randomly mapped to an integral decision $\mathbf{x}_t \in \mathcal{X}$ according to the rounding scheme Ξ. Then, the reward $f(\mathbf{x}_t)$ is received and f_t is revealed, at which point a concave function $\tilde{f}_t$ is constructed from f_t and added to the history. Our first main result is the following:

Theorem 2. *Under Asm. 2, given an OCO policy $\boldsymbol{\mathcal{P}}_{\mathcal{Y}}$, the RAOCO policy $\boldsymbol{\mathcal{P}}_{\mathcal{X}}$ described by Alg. 1 satisfies $\alpha\text{-regret}_T\left(\boldsymbol{\mathcal{P}}_{\mathcal{X}}\right) \leq \alpha \cdot \text{regret}_T\left(\boldsymbol{\mathcal{P}}_{\mathcal{Y}}\right)$.*

The proof is in Appendix D of Si Salem et al. (2023). As a result, *any regret guarantee obtained by an OCO algorithm over $\mathcal{Y}$, immediately transfers to an α-regret for RAOCO, where α is determined by Asm. 2.* In particular, the decision set $\mathcal{Y}$ is closed, bounded, and convex by construction. Combined with the fact that concave relaxations $\tilde{f}$ are L-Lipschitz (by Asm. 2), Thms. 1 and 2 yield the following corollary:

Corollary 1. *Under Asm. 2, RAOCO policy $\boldsymbol{\mathcal{P}}_{\mathcal{X}}$ in Alg. 1 equipped with OGA, OMA, or FTRL as OCO policy $\boldsymbol{\mathcal{P}}_{\mathcal{Y}}$ has sublinear α-regret. That is, $\alpha\text{-regret}_T\left(\boldsymbol{\mathcal{P}}_{\mathcal{X}}\right) = \mathcal{O}\left(\sqrt{T}\right)$.*

To use this result, Asm. 2 should hold, and both the randomized rounding and the concave relaxations used in RAOCO should be poly-time: all are true for WTP functions optimized over matroid constraints, which we turn to next.

The Case of Weighted Threshold Potentials

We now consider the case where the decision set is a matroid, and reward functions belong to the class of WTP functions, defined by Eq. (4). We will show that, under appropriate definitions of a randomized rounding and concave relaxations, the class $\mathcal{F}$ satisfies Asm. 2 and, thus, online learning via RAOCO comes with the regret guarantees of Corollary 1. For $\mathcal{Y} = \operatorname{conv}(\mathcal{X})$, consider the map $f \mapsto \tilde{f}$ of WTP functions $f : \mathcal{X} \to \mathbb{R}$ to concave relaxations $\tilde{f} : \mathcal{Y} \to \mathbb{R}$ of the form:

$$\tilde{f}(\mathbf{y}) \triangleq f(\mathbf{y}) = \textstyle\sum_{\ell \in C} c_\ell \min\left\{b_\ell, \sum_{j \in S_\ell} y_j w_{\ell,j}\right\}, \tag{9}$$

for $\mathbf{y} \in \mathcal{Y}$. In other words, *the relaxation of f is itself*: it has the same functional form, allowing integral variables to become fractional.[2] This is clearly concave, as the minimum of affine functions is concave, and the positively weighted sum of concave functions is concave. Finally, all such functions are Lipschitz, with a parameter that depends on c_ℓ, b_ℓ, $\mathbf{w}_\ell$, $\ell \in C$. Let $\tilde{\mathcal{F}}$ be the image of $\mathcal{F}$ under the map (9). We make the following assumption, which is readily satisfied if, e.g., all constituent parameters (c_ℓ, b_ℓ, $\mathbf{w}_\ell$, $\ell \in C$) are uniformly bounded, or the set $\mathcal{F}$ is finite, etc.:

Assumption 3. *There exists an $L > 0$ such that all functions in $\tilde{\mathcal{F}}$ are L-Lipschitz.*

Next, we turn our attention to the randomized rounding Ξ. We can in fact characterize the property that Ξ must satisfy for Asm. 2 to hold for relaxations given by Eq. (9):

Definition 1. *A randomized rounding $\Xi : \mathcal{Y} \to \mathcal{X}$ is* negatively correlated *if, for $\mathbf{x} = \Xi(\mathbf{y}) \in \mathcal{X}$ (a) the coordinates of $\mathbf{x}$ are negatively correlated*[3] *and (b) $\mathbb{E}_{\Xi}[\mathbf{x}] = \mathbf{y}$.*

Our next result immediately implies that *any negatively correlated rounding can be used in RAOCO*:

Lemma 1. *Let $\Xi : \mathcal{Y} \to \mathcal{X}$ be a negatively correlated randomized rounding, and consider the concave relaxations $\tilde{f}$ constructed from $f \in \mathcal{F}$ via Eq. (9). Then, if Asm. 3 holds for some $L > 0$, the set $\mathcal{F}$ satisfies Asm. 2 with $\alpha = \left(1 - \frac{1}{\Delta}\right)^{\Delta}$ where $\Delta = \sup_{f \in \mathcal{F}} \Delta_f$.*

The proof is in Appendix E of Si Salem et al. (2023). As $\Delta \to \infty$, the approximation ratio α approaches $1 - 1/e$ from above, recovering the usual approximation guarantee. However, for finite Δ, we in fact obtain an improved approximation ratio; an example (quadratic submodular functions) is described in Appendix B of Si Salem et al. (2023). Finally, and

[2] In Appendix G.IV of Si Salem et al. (2023), we provide an example where the functional form of concave relaxations differs.

[3] A set of random variables $x_i \in \{0, 1\}, i \in [n]$, are negatively correlated if $\mathbb{E}\left[\prod_{i \in S} x_i\right] \leq \prod_{i \in S} \mathbb{E}[x_i]$ for all $S \subseteq [n]$.

most importantly, *a negatively-correlated randomized rounding can always be constructed if $\mathcal{X}$ is a matroid*. Chekuri, Vondrák, and Zenklusen (2010) provide two polynomial-time randomized rounding algorithms that satisfy this property:

Lemma 2. *(Chekuri, Vondrák, and Zenklusen (2010, Theorem 1.1.)) Given a matroid $\mathcal{X} \subset \{0,1\}^n$, let $\mathbf{y} \in \operatorname{conv}(\mathcal{X})$ and Ξ be either* swap rounding *or* randomized pipage rounding. *Then, Ξ is negatively correlated.*

We review swap rounding in Appendix F of Si Salem et al. (2023). Interestingly, the existence of a negatively correlated rounding is inherently linked to matroids: a negatively-correlated rounding exists *if and only if $\mathcal{X}$ is a matroid* (see Thm. I.1. in Chekuri, Vondrák, and Zenklusen (2010)). Lemma 1 thus implies that the reduction of RAOCO to OCO policies is also linked to matroids. Putting everything together, Lemmas 1–2 and Corollary 1 yield the following:

Theorem 3. *Let $\mathcal{X} \subseteq \{0,1\}^n$ be a matroid, and $\mathcal{F}$ be a subset of the WTP class for which Asm. 3 holds. Then, the RAOCO policy $P_{\mathcal{X}}$ defined by Alg. 1 using swap rounding or randomized pipage rounding as Ξ and OGA, OMA, or FTRL as OCO policy $\boldsymbol{\mathcal{P}}_{\mathcal{Y}}$, and concave relaxations in Eq.* (9) *has sublinear α-regret. In particular, $\alpha\text{-regret}(\boldsymbol{\mathcal{P}}_{\mathcal{X}}) = \mathcal{O}\left(\sqrt{T}\right)$.*

Note that, though all algorithms yield $O(\sqrt{T})$ regret, the dependence of constants on problem parameters (such as n and the matroid rank r), as reported in Table 1 in C of Si Salem et al. (2023) is optimized under OMA (see Appendix C of Si Salem et al. (2023)).

Computational Complexity

OCO Policy. OCO policies are polytime (see, e.g., (Hazan 2016)). Taking gradient-based OCO policies (e.g., OMA in Alg. 2 in Si Salem et al. (2023)), their computational complexity is dominated by a projection operation to the convex set $\mathcal{Y} = \operatorname{conv}(\mathcal{X})$. The exact time complexity of this projection depends on $\mathcal{Y}$, however, given a membership oracle that decides $\mathbf{y} \in \operatorname{conv}(\mathcal{X})$, the projection can be computed efficiently in polynomial time (Hazan 2016). Moreover, the projection problem is a convex problem that can be computed efficiently (e.g., iteratively to an arbitrary precision) and can also be computed in strongly polynomial time (Gupta, Goemans, and Jaillet 2016, Theorem 3).

Concave Relaxations and Randomized Rounding. Concave relaxations are linear in the representation of the function f, but in practice come "for free", once parameters in Eq. (4) are provided. Swap rounding over a general matroid is $\mathcal{O}\left(nr^2\right)$, where r is the rank of the matroid (Chekuri, Vondrák, and Zenklusen 2010). This dominates the remaining operations (including OCO), and thus determines the overall complexity of our algorithm. This $\mathcal{O}\left(nr^2\right)$ term assumes access to the decomposition of a fractional point $\mathbf{y} \in \mathcal{Y}$ to bases of the matroid. Carathéodory's theorem implies the existence of such decomposition of at most $n+1$ points in $\mathcal{X}$; moreover, there exists a decomposition algorithm (Cunningham 1984) for general matroids with running time $\mathcal{O}\left(n^6\right)$. However, in all practical cases we consider (including uniform and partition matroids) the complexity is significantly lower. More specifically, for partition matroids, swap rounding reduces to an algorithm with linear time complexity, namely, $\mathcal{O}(mn)$ for m partitions (Srinivasan 2001).

Extensions

Dynamic Setting. In the dynamic setting, the decision maker compares its performance to the best sequence of decisions $(\mathbf{y}_t^\star)_{t\in[T]}$ with a path length regularity condition (Zinkevich 2003; Cesa-Bianchi et al. 2012). I.e., let $\Lambda_{\mathcal{X}}(T, P_T) \triangleq \left\{(\mathbf{x}_t)_{t=1}^T \in \mathcal{X}^T : \sum_{t=1}^T \|\mathbf{x}_{t+1} - \mathbf{x}_t\| \le P_T\right\} \subset \mathcal{X}^T$ be the set of sequences of decisions in a set $\mathcal{X}$ with path length less than $P_T \in \mathbb{R}_{\ge 0}$ over a time horizon T. We define $\alpha\text{-regret}_{T,P_T}(\boldsymbol{\mathcal{P}}_{\mathcal{X}})$, the *dynamic α-regret*, as:

$$\sup_{(f_t)_{t=1}^T \in \mathcal{F}^T} \left\{ \max_{(\mathbf{x}_t^\star)_{t=1}^T \in \Lambda_{\mathcal{X}}(T,P_T)} \alpha \sum_{t=1}^T f_t(\mathbf{x}_t^\star) - \sum_{t=1}^T f_t(\mathbf{x}_t) \right\}$$

When $\alpha\text{-regret}_{T,P_T}(\boldsymbol{\mathcal{P}}_{\mathcal{X}})$ is sublinear in T, the policy attains average rewards that asymptotically compete with the optimal decisions *of bounded path length*, in hindsight.

Through our reduction to OCO, we can leverage OGA (Zinkevich 2003) or meta-learning algorithms over OGA (Zhao et al. 2020) to obtain dynamic regret guarantees in OSM. As an additional technical contribution, we provide the first sufficient and necessary conditions for OMA to admit a dynamic regret guarantee (see Appendix A of Si Salem et al. (2023)). This allows to extend a specific instance of OMA operating on the simplex, the so-called fixed-share algorithm (Cesa-Bianchi et al. 2012; Herbster and Warmuth 1998), to matroid polytopes. This yields a tighter regret guarantee than OGA (Zinkevich 2003; Zhao et al. 2020) (see Theorem 7 in Appendix A of Si Salem et al. (2023)). Putting everything together, we get:

Theorem 4. *Under Asm. 2, RAOCO policy $\boldsymbol{\mathcal{P}}_{\mathcal{X}}$ described by Alg. 1 equipped with an OMA policy in Appendix A of Si Salem et al. (2023) as OCO policy $\boldsymbol{\mathcal{P}}_{\mathcal{Y}}$ has sublinear dynamic α-regret, i.e., $\alpha\text{-regret}_{T,P_T}(\boldsymbol{\mathcal{P}}_{\mathcal{X}}) = \mathcal{O}\left(\sqrt{P_T T}\right)$.*

This follows from Thm. 2 and the dynamic regret guarantee for OMA in Appendix A of Si Salem et al. (2023).

Optimistic Setting. In the optimistic setting, the decision maker has access to additional information available in the form of predictions: a function $\pi_{t+1} : [0,1]^n \to \mathbb{R}_{\ge 0}$, serving as a prediction of the reward function $f_{t+1}(\mathbf{x})$ is made available before committing to a decision $\mathbf{x}_{t+1}$ at timeslot $t \in [T]$. The prediction π_t encodes prior information available to the decision maker at timeslot t.[4] Let $\mathbf{g}_t$ and $\mathbf{g}_t^\pi$ be supergradients of $\tilde{f}_t$ and π_t at point $\mathbf{y}_t$, respectively. We can define an optimistic OMA policy (see Alg. 3 in Appendix A of Si Salem et al. (2023)) that leverages both the prediction and the observed rewards. Applying again our reduction of OSM to this setting we get:

Theorem 5. *Under Asm. 2, RAOCO policy $\boldsymbol{\mathcal{P}}_{\mathcal{X}}$ in Alg. 1 equipped with OOMA in Appendix A of Si Salem et al. (2023) as policy $\boldsymbol{\mathcal{P}}_{\mathcal{Y}}$ yields $\alpha\text{-regret}_{T,P_T}(\boldsymbol{\mathcal{P}}_{\mathcal{X}}) = \mathcal{O}\left(\sqrt{P_T \sum_{t=1}^T \|\mathbf{g}_t - \mathbf{g}_t^\pi\|_\infty^2}\right)$.*

[4] Function π_t can extend over fractional values, e.g., be the multilinear relaxation of a set function.

Datasets & Constr.		$F^\star$	t	RAOCO-OGA $\bar{F}_{\mathcal{X}}/F^\star$	RAOCO-OMA $\bar{F}_{\mathcal{X}}/F^\star$	FSF* $\bar{F}_{\mathcal{X}}/F^\star$	TabularGreedy $\bar{F}_{\mathcal{X}}/F^\star$	Random $\bar{F}_{\mathcal{X}}/F^\star$
ZKC	unif.	0.234	33	0.902	**0.965**	0.839	0.833	0.642
			66	0.924	**0.967**	0.896	0.894	0.624
			99	0.945	**0.982**	0.933	0.931	0.622
	part.	0.83	33	0.994	**0.997**	×	0.985	0.953
			66	0.99	**0.994**		0.987	0.950
			99	0.993	**0.997**		0.995	0.953
Epinions	unif.	0.171	50	0.845	**0.853**	0.703	0.694	0.632
			100	0.865	**0.906**	0.776	0.768	0.615
			149	0.88	**0.925**	0.807	0.805	0.629
	part.	0.171	50	0.826	**0.861**	×	0.720	0.620
			100	0.854	**0.908**		0.786	0.619
			149	0.88	**0.927**		0.818	0.625
MovieLens	uniform	0.407	98	0.749	**0.792**	0.681	0.69	0.748
			196	**0.786**	0.781	0.713	0.676	0.7
			293	0.846	**0.866**	0.756	0.769	0.711
	part.	0.419	98	0.889	**0.948**	×	0.908	0.829
			196	0.872	**0.908**		0.902	0.814
			293	0.926	0.948		**0.964**	0.874
SynthTF	unif.	200	33	0.984	**0.987**	0.845	0.844	0.605
			66	0.994	**0.994**	0.868	0.886	0.603
			99	0.995	**0.998**	0.869	0.902	0.612
	part.	400	33	0.98	**0.983**	×	0.834	0.611
			66	0.990	**0.991**		0.852	0.607
			99	0.993	**0.994**		0.857	0.601

Table 1: Average cumulative reward $\bar{F}_{\mathcal{X}}$ ($t = T/3, 2T/3, T$), normalized by fractional optimal $F^\star$, of integral policies across different datasets and uniform (unif.) and partition (part.) matroid constraints. Optimal hyperparameters ($\eta, \gamma, c_{\mathrm{p}}$) and standard deviations are reported in the Appendix H of Si Salem et al. (2023) along with the value ranges explored. RAOCO combined with OGA or OMA outperforms competitors almost reaching a ratio of one, with the exception of MovieLens, where TabularGreedy does better. As Random also performs well on MovieLens, this indicates that the (static) offline optimal is quite poor for this reward sequence. By Property 2, fractional solutions strictly dominate the integral optimal, which implies that in all cases RAOCO outperformed the $1 - 1/e$ approximation, attaining an almost optimal value.

This theorem follows from Theorem 2 and the optimistic regret guarantee established for OMA policies in Theorem 7 in Appendix A of Si Salem et al. (2023). The optimistic regret guarantee in Theorem 5 shows that the regret of a policy can be reduced to 0 when the predictions are perfect, while providing $\mathcal{O}\left(\sqrt{T}\right)$ guarantee in Thm. 3 when the predictions are arbitrarily bad (with bounded gradients). To the best of our knowledge, ours is the first work to provide guarantees for optimistic OSM.

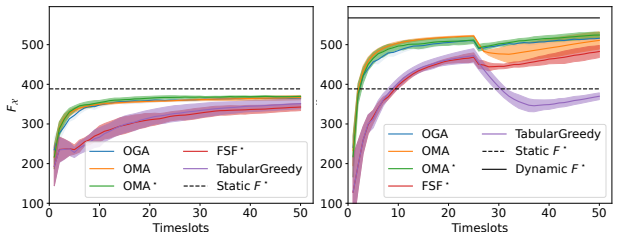
Bandit Setting. Recall that in the bandit setting the decision maker only has access to the reward $f_t(\mathbf{x}_t)$ after committing to a decision $\mathbf{x}_t \in \mathcal{X}$ at $t \in [T]$; i.e., the reward function is *not* revealed. Our reduction to OCO does not readily apply to the bandit setting; however, we show that the bandit algorithm by Wan et al. (2023) can be used to construct such a reduction. The main challenge is to estimate gradients of inputs in $\mathcal{Y}$ only from bandit feedback; this can be done via a perturbation method (see also Hazan and Levy (2014)). This approach, described in Appendix G of Si Salem et al. (2023), yields the following theorem:

Theorem 6. *Under bounded submodular monotone rewards and partition matroid constraint sets, LIRAOCO policy $\boldsymbol{\mathcal{P}}_{\mathcal{X}}$ in Alg. 5 in Appendix G of Si Salem et al. (2023) equipped with an OCO policy $\boldsymbol{\mathcal{P}}_{\mathcal{Y}_\delta}$ yields α-$\mathrm{regret}_{T,P_T,W}(\boldsymbol{\mathcal{P}}_{\mathcal{X}}) \leq W \cdot \mathrm{regret}_{T/W,P_T}(\boldsymbol{\mathcal{P}}_{\mathcal{Y}_\delta}) + \frac{T}{W} + 2\alpha\delta r^2 nT$, where δ, W, are tuneable parameters of the algorithm and $\mathrm{regret}_{T/W,P_T}(\boldsymbol{\mathcal{P}}_{\mathcal{Y}_\delta})$ is the regret of an OCO policy executed for T/W timeslots.*

Thm. 6 applies to general submodular functions, but is restricted to partition matroids. The theorem also extends to dynamic and optimistic settings: we provide the full description in Appendix G of Si Salem et al. (2023). Our analysis generalizes Wan et al. (2023) in that (a) we show that a reduction can be performed to any OCO algorithm, rather than just FTRL, as well as (b) in extending it to the dynamic and optimistic settings (see also Table 4 in Si Salem et al. (2023)).

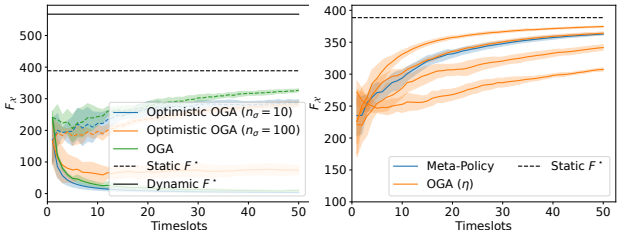
Experiments

Datasets and Problem Instances. We consider five different OSM problems: two influence maximization problems, over the ZKC (Zachary 1977) and Epinions (Richardson, Agrawal, and Domingos 2003) graphs, respectively, a facility location problem over the MovieLens dataset (Harper and Konstan 2015), a team formation, and a weighted cov-



(a) Stationary (b) Non-Stationary

Figure 1: Average cumulative reward $\bar{F}_{\mathcal{X}}$ of the different policies under SynthWC dataset under different setups: stationary in (a) and non-stationary in (b). Non-stationarity in (b) is applied by changing the objective at $t = 25$ (see Appendix H of Si Salem et al. (2023)). The area depicts the standard deviation over 5 runs.



(a) Optimistic Learning (b) Meta-Policies

Figure 2: Average cumulative reward $\bar{F}_{\mathcal{X}}$ of the different policies under SynthWC dataset under a non-stationary setup: the objective is changed at every timeslot. The algorithms Optimistic OGA and OGA are executed with different learning rates under different prediction accuracy (noise with std. dev. $n_\sigma \in \{10, 100\}$) in (a); the larger learning rate is depicted by a solid line. The meta-policy in (b) can learn the best configuration of OGA without tuning the learning rate. The area depicts the standard deviation over 5 runs.

erage problem over synthetic datasets. Reward functions are generated over a finite horizon and optimized online over both uniform and partition matroid constraints. Details are provided in Appendix H of Si Salem et al. (2023).

Algorithms. We implement the policy in Alg. 1 (RAOCO), coupled with OGA (RAOCO-OGA) and OMA (RAOCO-OMA) as OCO policies. As competitors, we also implement the *fixed share forecaster* (FSF⋆) policy by Matsuoka, Ito, and Ohsaka (2021), which only applies to uniform matroids, and the TabularGreedy policy by Streeter, Golovin, and Krause (2009), as well as a Random algorithm, which selects a decision u.a.r. from the bases of $\mathcal{X}$. Details and hyperparameters explored are reported in Appendix H of Si Salem et al. (2023). Our code is publicly available.[5]

Metrics. For each setting, we compute $F^\star = \max_{\mathbf{y}\in\mathcal{Y}} \frac{1}{T}\sum_{t=1}^{T} \tilde{f}(\mathbf{y})$, the value of the optimal fractional solution in hindsight, assuming rewards are replaced by their concave relaxations. Note that this involves solving a convex optimization problem, and overestimates the (intractable) offline optimal, i.e., $F^\star \geq \max_{\mathbf{x}\in\mathcal{X}} \frac{1}{T}\sum_{t=1}^{T} f(\mathbf{x})$. For each online policy, we compute both the integral and fractional average cumulative reward at different timeslots t, given by $\bar{F}_{\mathcal{X}} = \frac{1}{t}\sum_{s=1}^{t} f(\mathbf{x}_s)$, $\bar{F}_{\mathcal{Y}} = \frac{1}{t}\sum_{s=1}^{t} \tilde{f}(\mathbf{y}_s)$, respectively. We repeat each experiment for 5 different random seeds, and use this to report standard deviations.

OSM Policy Comparison. A comparison between the two versions of RAOCO (OGA and OMA) with the three competitors is shown in Table 1. We observe that both OGA and OMA significantly outperform competitors, with the exception of MovieLens, where TabularGreedy does better. OMA is almost always better than OGA. Most importantly, we significantly outperform both TabularGreedy and FSF⋆ w.r.t. running time, being 2.15–250 times faster (see Appendix H of Si Salem et al. (2023)).

Dynamic Regret and Optimistic Learning. Fig. 1 shows the performance of the different policies in a dynamic scenario. All the policies have similar performance in the stationary setting, however in the non-stationary setting only OGA, OMA, and FSF⋆ show robustness. We further experiment with an optimistic setting in Fig. 2 (a) under a non-stationary setting where we provide additional information about future rewards to the optimistic policies, which can leverage this information and yield better results (see Appendix H of Si Salem et al. (2023)).

Meta-Policies. Fig. 2 (b) shows the performance of the meta-policy that learns over different policies configured with different learning rates. We observe that the meta-policy quickly learns the best learning rate.

Conclusion

We provide a reduction of online maximization of a large class of submodular functions to online convex optimization and show that our reduction extends to many different versions of the online learning problem. There are many possible extensions. As our framework does not directly apply to general submodular functions, it would be interesting to derive a general method to construct the concave relaxations which are the building block of our reduction in Theorem 2. It is also meaningful to investigate the applicability of our reduction to monotone submodular functions with bounded curvature (Feldman 2021), and non-monotone submodular functions (Krause and Golovin 2014).

Acknowledgements

We gratefully acknowledge support from the NSF (grants 1750539, 2112471, 1813406, and 1908510).

References

Ageev, A. A.; and Sviridenko, M. I. 2004. Pipage rounding: A new method of constructing algorithms with proven performance guarantee. *Journal of Combinatorial Optimization*.

Andelman, N.; and Mansour, Y. 2004. Auctions with budget constraints. In *SWAT*.

[5] https://github.com/neu-spiral/OSMviaOCO

Auer, P.; Cesa-Bianchi, N.; Freund, Y.; and Schapire, R. E. 1995. Gambling in a rigged casino: The adversarial multi-armed bandit problem. In *FOCS*.

Bach, F. 2019. Submodular functions: from discrete to continuous domains. *Mathematical Programming*.

Besbes, O.; Gur, Y.; and Zeevi, A. 2015. Non-stationary stochastic optimization. *Operations research*, 63(5): 1227–1244.

Bian, A. A.; Mirzasoleiman, B.; Buhmann, J.; and Krause, A. 2017. Guaranteed non-convex optimization: Submodular maximization over continuous domains. In *AISTATS*.

Bubeck, S. 2011. Introduction to online optimization. *Lecture notes*.

Buchfuhrer, D.; Dughmi, S.; Fu, H.; Kleinberg, R.; Mossel, E.; Papadimitriou, C.; Schapira, M.; Singer, Y.; and Umans, C. 2010. Inapproximability for vcg-based combinatorial auctions. In *SODA*.

Calinescu, G.; Chekuri, C.; Pal, M.; and Vondrák, J. 2011. Maximizing a monotone submodular function subject to a matroid constraint. *SICOMP*.

Cesa-Bianchi, N.; Gaillard, P.; Lugosi, G.; and Stoltz, G. 2012. Mirror descent meets fixed share (and feels no regret). *NeurIPS*.

Cesa-Bianchi, N.; and Lugosi, G. 2006. *Prediction, Learning, and Games*. Cambridge University Press. ISBN 0521841089.

Chekuri, C.; Vondrák, J.; and Zenklusen, R. 2010. Dependent randomized rounding via exchange properties of combinatorial structures. In *FOCS*.

Chen, L.; Hassani, H.; and Karbasi, A. 2018. Online Continuous Submodular Maximization. In *AISTATS*.

Cunningham, W. H. 1984. Testing membership in matroid polyhedra. *Journal of Combinatorial Theory, Series B*.

Dekel, O.; Flajolet, A.; Haghtalab, N.; and Jaillet, P. 2017. Online learning with a hint. *NeurIPS*.

Dobzinski, S. 2016. Breaking the logarithmic barrier for truthful combinatorial auctions with submodular bidders. In *STOC*.

Dolhansky, B. W.; and Bilmes, J. A. 2016. Deep submodular functions: Definitions and learning. *NeurIPS*.

Feldman, M. 2021. Guess free maximization of submodular and linear sums. *Algorithmica*.

Flaxman, A. D.; Kalai, A. T.; and McMahan, H. B. 2005. Online Convex Optimization in the Bandit Setting: Gradient Descent without a Gradient. In *SODA*.

Frank, M.; and Wolfe, P. 1956. An algorithm for quadratic programming. *Naval research logistics quarterly*.

Frieze, A. M. 1974. A cost function property for plant location problems. *Mathematical Programming*.

Goemans, M. X.; and Williamson, D. P. 1994. New $\frac{3}{4}$-approximation algorithms for the maximum satisfiability problem. *SIDMA*.

Gupta, S.; Goemans, M.; and Jaillet, P. 2016. Solving combinatorial games using products, projections and lexicographically optimal bases. *arXiv preprint arXiv:1603.00522*.

Harper, F. M.; and Konstan, J. A. 2015. The MovieLens Datasets: History and Context. *TiiS*.

Harvey, N.; Liaw, C.; and Soma, T. 2020. Improved algorithms for online submodular maximization via first-order regret bounds. *NeurIPS*.

Hazan, E. 2016. Introduction to Online Convex Optimization. *Foundations and Trends® in Optimization*.

Hazan, E.; and Levy, K. 2014. Bandit convex optimization: Towards tight bounds. *NeurIPS*.

Herbster, M.; and Warmuth, M. K. 1998. Tracking the best expert. *Machine learning*.

Ioannidis, S.; and Yeh, E. 2016. Adaptive caching networks with optimality guarantees. *PER*.

Ito, S.; and Fujimaki, R. 2016. Large-scale price optimization via network flow. *NeurIPS*.

Jadbabaie, A.; Rakhlin, A.; Shahrampour, S.; and Sridharan, K. 2015. Online optimization: Competing with dynamic comparators. In *AISTATS*.

Kakade, S. M.; Kalai, A. T.; and Ligett, K. 2007. Playing games with approximation algorithms. In *STOC*.

Karimi, M.; Lucic, M.; Hassani, H.; and Krause, A. 2017. Stochastic Submodular Maximization: The Case of Coverage Functions. *NeurIPS*.

Kempe, D.; Kleinberg, J.; and Tardos, É. 2003. Maximizing the spread of influence through a social network. In *KDD*.

Kleinberg, R. 2004. Nearly tight bounds for the continuum-armed bandit problem. *NeurIPS*.

Krause, A.; and Golovin, D. 2014. Submodular Function Maximization. In *Tractability: Practical Approaches to Hard Problems*. Cambridge University Press.

Li, Y.; Cheng, M.; Fujii, K.; Hsieh, F.; and Hsieh, C.-J. 2018. Learning from group comparisons: exploiting higher order interactions. *NeurIPS*.

Li, Y.; Si Salem, T.; Neglia, G.; and Ioannidis, S. 2021. Online caching networks with adversarial guarantees. *POMACS*.

Littlestone, N.; and Warmuth, M. K. 1994. The Weighted Majority Algorithm. *Information and computation*.

Matsuoka, T.; Ito, S.; and Ohsaka, N. 2021. Tracking regret bounds for online submodular optimization. In *AISTATS*.

McMahan, H. B. 2017. A survey of algorithms and analysis for adaptive online learning. *JMLR*.

Mohri, M.; and Yang, S. 2016. Accelerating online convex optimization via adaptive prediction. In *AISTATS*.

Mokhtari, A.; Shahrampour, S.; Jadbabaie, A.; and Ribeiro, A. 2016. Online optimization in dynamic environments: Improved regret rates for strongly convex problems. In *CDC*.

Niazadeh, R.; Golrezaei, N.; Wang, J. R.; Susan, F.; and Badanidiyuru, A. 2021. Online learning via offline greedy algorithms: Applications in market design and optimization. In *EC*.

Rakhlin, A.; and Sridharan, K. 2013. Online learning with predictable sequences. In *COLT*.

Richardson, M.; Agrawal, R.; and Domingos, P. 2003. Trust Management for the Semantic Web. In *ISWC*.

Shalev-Shwartz, S. 2012. Online Learning and Online Convex Optimization. *Foundations and Trends® in Machine Learning*.

Si Salem, T.; Neglia, G.; and Carra, D. 2022. Ascent Similarity Caching With Approximate Indexes. *ToN*.

Si Salem, T.; Özcan, G.; Nikolaou, I.; Terzi, E.; and Ioannidis, S. 2023. Online Submodular Maximization via Online Convex Optimization. *arXiv preprint arXiv:2309.04339*.

Srinivasan, A. 2001. Distributions on level-sets with applications to approximation algorithms. In *FOCS*.

Stobbe, P.; and Krause, A. 2010. Efficient Minimization of Decomposable Submodular Functions. *NeurIPS*.

Streeter, M.; Golovin, D.; and Krause, A. 2009. Online Learning of Assignments. In *NeurIPS*.

Wan, Z.; Zhang, J.; Chen, W.; Sun, X.; and Zhang, Z. 2023. Bandit Multi-linear DR-Submodular Maximization and Its Applications on Adversarial Submodular Bandits. In *Proceedings of the 40th International Conference on Machine Learning (ICML)*.

Zachary, W. W. 1977. An information flow model for conflict and fission in small groups. *Journal of anthropological research*.

Zhang, M.; Chen, L.; Hassani, H.; and Karbasi, A. 2019. Online continuous submodular maximization: From full-information to bandit feedback. *NeurIPS*.

Zhang, Q.; Deng, Z.; Chen, Z.; Hu, H.; and Yang, Y. 2022. Stochastic continuous submodular maximization: Boosting via non-oblivious function. In *ICML*.

Zhao, P.; Zhang, Y.-J.; Zhang, L.; and Zhou, Z.-H. 2020. Dynamic regret of convex and smooth functions. *NeurIPS*.

Zinkevich, M. 2003. Online Convex Programming and Generalized Infinitesimal Gradient Ascent. In *ICML*. ISBN 1577351894.