

On Maximizing Sums of Non-Monotone Submodular and Linear Functions

Benjamin Qi 

Massachusetts Institute of Technology, Cambridge, MA, USA

Abstract

We study the problem of **Regularized Unconstrained Submodular Maximization** (**RegularizedUSM**) as defined by [Bodek and Feldman '22]. In this problem, we are given query access to a non-negative submodular function $f: 2^{\mathcal{N}} \rightarrow \mathbb{R}_{\geq 0}$ and a linear function $\ell: 2^{\mathcal{N}} \rightarrow \mathbb{R}$ over the same ground set \mathcal{N} , and the objective is to output a set $T \subseteq \mathcal{N}$ approximately maximizing the sum $f(T) + \ell(T)$. Specifically, an algorithm is said to provide an (α, β) -approximation for **RegularizedUSM** if it outputs a set T such that $\mathbb{E}[f(T) + \ell(T)] \geq \max_{S \subseteq \mathcal{N}} [\alpha \cdot f(S) + \beta \cdot \ell(S)]$. We also study the setting where S and T are constrained to be independent in a given matroid, which we refer to as **Regularized Constrained Submodular Maximization** (**RegularizedCSM**).

The special case of **RegularizedCSM** with monotone f has been extensively studied [Sviridenko et al. '17, Feldman '18, Harshaw et al. '19]. On the other hand, we are aware of only one prior work that studies **RegularizedCSM** with non-monotone f [Lu et al. '21], and that work constrains ℓ to be non-positive. In this work, we provide improved (α, β) -approximation algorithms for both **RegularizedUSM** and **RegularizedCSM** with non-monotone f . In particular, we are the first to provide nontrivial (α, β) -approximations for **RegularizedCSM** where the sign of ℓ is unconstrained, and the α we obtain for **RegularizedUSM** improves over [Bodek and Feldman '22] for all $\beta \in (0, 1)$.

In addition to approximation algorithms, we provide improved inapproximability results for all of the aforementioned cases. In particular, we show that the α our algorithm obtains for **RegularizedCSM** with unconstrained ℓ is essentially tight for $\beta \geq \frac{e}{e+1}$. Using similar ideas, we are also able to show 0.478-inapproximability for maximizing a submodular function where S and T are subject to a cardinality constraint, improving a 0.491-inapproximability result due to [Oveis Gharan and Vondrak '10].

2012 ACM Subject Classification Mathematics of computing \rightarrow Combinatorial optimization; Theory of computation \rightarrow Approximation algorithms analysis

Keywords and phrases submodular maximization, regularization, continuous greedy, inapproximability

Digital Object Identifier 10.4230/LIPIcs.ISAAC.2022.41

Related Version *Full Version*: <https://arxiv.org/abs/2205.15874>

Supplementary Material *Software (Source Code)*: <https://github.com/bqi343/maximizing-sums> archived at `swh:1:dir:0dbefd7cb01ec27b7ec94585161ada5356de2906`

Acknowledgements I thank Tasuku Soma for mentoring me through the MIT Undergraduate Research Opportunities Program, as well as Moran Feldman for providing many helpful suggestions.

1 Introduction

Submodularity is a property satisfied by many fundamental set functions, including coverage functions, matroid rank functions, and directed cut functions. Optimization of submodular set functions has found a wealth of applications in machine learning, including the spread of influence in social networks [13], sensor placement [16], information gathering [15], document summarization [17, 26, 9], image segmentation [11], and multi-object tracking [23], among others (see Krause and Golovin [14] for a survey).



© Benjamin Qi;

licensed under Creative Commons License CC-BY 4.0

33rd International Symposium on Algorithms and Computation (ISAAC 2022).

Editors: Sang Won Bae and Heejin Park; Article No. 41; pp. 41:1–41:16

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Problems involving maximization of non-negative submodular functions can be classified as either *unconstrained* or *constrained*. In the unconstrained case, the objective is to return a set in the domain of the function approximately maximizing the function; we refer to this problem as *USM*. In the most commonly studied form of constrained submodular maximization, the returned set is subject to a “matroid constraint,” which means that the returned set is constrained to be independent in a given matroid. We refer to this form of constrained submodular maximization as *CSM*. The simplest nontrivial example of a matroid constraint is a *cardinality* constraint, which means that an upper bound is given on the allowed size of the returned set. Additionally, we refer to the special case where the function f we are maximizing is monotone as “monotone CSM.”

Approximation algorithms for both *USM* and *CSM* have been studied extensively. Say that an algorithm running in polynomial time with respect to the size of the ground set provides an α -approximation if it returns a set with expected value at least α times that of the optimum. For *USM*, a 0.5-approximation algorithm was provided by Buchbinder et al. [3]. For monotone *CSM*, a $(1 - e^{-1})$ -approximation was achieved by Nemhauser et al. [20] using a *greedy* algorithm for the special case of cardinality constraints and later generalized by Calinescu et al. [4] to matroid constraints using the *continuous greedy* algorithm. For general *CSM*, the *measured continuous greedy* algorithm of Feldman et al. [7] achieves an $e^{-1} > 0.367$ -approximation, and a subsequent algorithm due to Buchbinder and Feldman [2] achieves a 0.385-approximation.

To bound how far the algorithms from the previous paragraph are from optimal, corresponding inapproximability results have been shown. Say that a problem is α -inapproximable if no algorithm running in sub-exponential time with respect to the size of the ground set can provide an α -approximation. The first two approximation factors from the previous paragraph are in fact the best achievable; $(0.5 + \epsilon)$ -inapproximability and $(1 - e^{-1} + \epsilon)$ -inapproximability for any $\epsilon > 0$ were shown by Feige et al. [5] and Nemhauser and Wolsey [19], respectively, using ad hoc methods. On the other hand, the best achievable approximability for general *CSM* remains open; the best known inapproximability factor is 0.478 due to Oveis Gharan and Vondrak [8] using the *symmetry gap* technique of Vondrak [25]. This technique has the advantage of being able to succinctly reprove the inapproximability results of [19, 5] and many others.

In this work we study approximation algorithms for maximizing the sum of a non-negative submodular function f and a linear function ℓ . Sviridenko et al. [24] were the first to study algorithms for the sum $f + \ell$ in the case of f monotone, in order to provide improved approximation algorithms for monotone *CSM* with bounded *curvature*. Here, the curvature $c \in [0, 1]$ of a non-negative monotone submodular function g is roughly a measure of how far g is from being linear. They provided a $(1 - c/e - \epsilon)$ -approximation algorithm and a complementary $(1 - c/e + \epsilon)$ -inapproximability result. The idea of the algorithm is to decompose g into $f + \ell$ and show that an approximation factor of $1 - e^{-1}$ can be achieved with respect to f and an approximation factor of 1 can be achieved with respect to ℓ simultaneously. Formally, if \mathcal{I} is the independent set family of a matroid, the algorithm computes a set $T \in \mathcal{I}$ that satisfies

$$\mathbb{E}[g(T)] = \mathbb{E}[f(T) + \ell(T)] \geq \max_{S \in \mathcal{I}} [(1 - e^{-1} - \epsilon)f(S) + (1 - \epsilon)\ell(S)]$$

by first “guessing” the value of $\ell(S)$, and then running the continuous greedy algorithm. Subsequently, Feldman eliminated the need for the guessing step and the dependence on $\epsilon\ell(S)$ by introducing a *distorted objective* [6]. Many faster algorithms and practical applications for the case of f monotone have since been introduced [10, 12, 21]. Note that ℓ has several potential interpretations; while setting ℓ to be non-negative provides improved approximations

for monotone submodular functions with low curvature, setting ℓ to be non-positive allows it to serve as a *regularizer* or *soft constraint* that favors returning smaller sets as suggested by Harshaw et al. [10].

On the other hand, we know of only two prior works that study **RegularizedUSM** where f is not constrained to be monotone. Bodek and Feldman [1] were the first to consider the case where f is non-monotone and the sign of ℓ is unconstrained. They defined and studied the problem of **Regularized Unconstrained Submodular Maximization (RegularizedUSM)**:

► **Definition 1 (RegularizedUSM)**. *Given query access to a (not necessarily monotone) non-negative submodular function $f: 2^{\mathcal{N}} \rightarrow \mathbb{R}_{\geq 0}$ and a linear function $\ell: 2^{\mathcal{N}} \rightarrow \mathbb{R}$ over the same ground set \mathcal{N} , an algorithm is said to provide an (α, β) -approximation for **RegularizedUSM** if it outputs a set $T \subseteq \mathcal{N}$ such that $\mathbb{E}[f(T) + \ell(T)] \geq \max_{S \subseteq \mathcal{N}}[\alpha \cdot f(S) + \beta \cdot \ell(S)]$.*

The main approximation result of [1] is the first non-trivial approximation algorithm for **RegularizedUSM** with f non-monotone and the sign of ℓ unconstrained. Specifically, they used *non-oblivious local search* to provide $(\alpha(\beta) - \epsilon, \beta - \epsilon)$ -approximations for **RegularizedUSM** for all $\beta \in (0, 1]$, where $\alpha(\beta) \triangleq \beta(1 - \beta)/(1 + \beta)$ [1, Theorem 1.2]. They also proved inapproximability results for the cases of ℓ non-negative and ℓ non-positive using the symmetry gap technique of Vondrak [25]. In particular, they showed $(1 - e^{-\beta} + \epsilon, \beta)$ -inapproximability for monotone f and non-positive ℓ for all $\beta \geq 0$ [1, Theorem 1.1], essentially matching the $(1 - e^{-\beta} - \epsilon, \beta)$ -approximability provided by Lu et al.'s *distorted measured continuous greedy* algorithm [18] (note that (α, β) -inapproximability is defined in the same way as α -inapproximability).

In this work, we present improved approximability and inapproximability results for **RegularizedUSM** as well as the setting where S and T are subject to a matroid constraint, which we define analogously as **Regularized Constrained Submodular Maximization (RegularizedCSM)**:

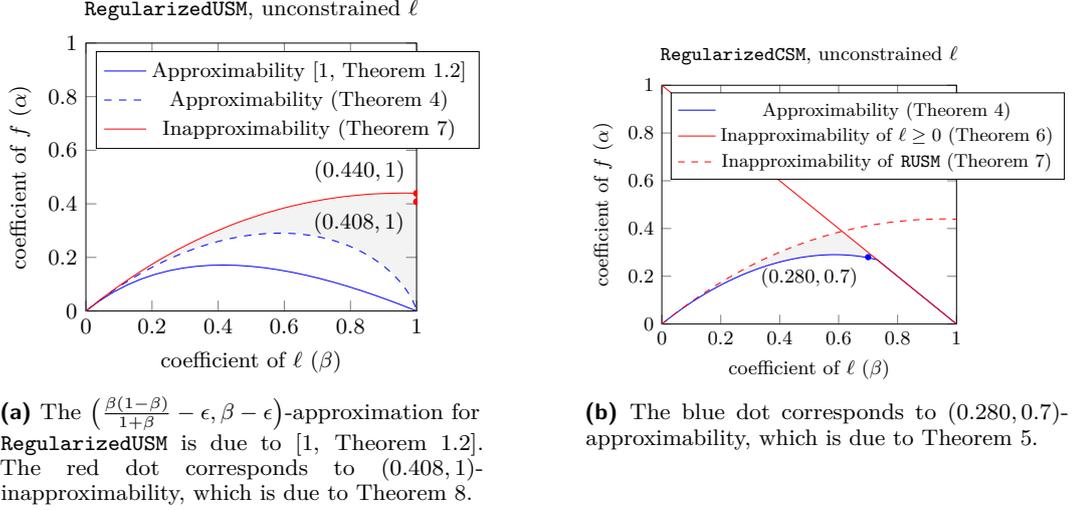
► **Definition 2 (RegularizedCSM)**. *Given query access to a (not necessarily monotone) non-negative submodular function $f: 2^{\mathcal{N}} \rightarrow \mathbb{R}_{\geq 0}$ and a linear function $\ell: 2^{\mathcal{N}} \rightarrow \mathbb{R}$ over the same ground set \mathcal{N} , as well as a matroid with family of independent sets denoted by \mathcal{I} also over the same ground set, an algorithm is said to provide an (α, β) -approximation for **RegularizedCSM** if it outputs a set $T \in \mathcal{I}$ such that $\mathbb{E}[f(T) + \ell(T)] \geq \max_{S \in \mathcal{I}}[\alpha \cdot f(S) + \beta \cdot \ell(S)]$.*

The only prior work considering **RegularizedCSM** for non-monotone f that we are aware of is that of Lu et al. [18], which as noted by [1] achieves $(\beta e^{-\beta} - \epsilon, \beta)$ -approximations for **RegularizedCSM** for all $\beta \in [0, 1]$, but only when ℓ is constrained to be non-positive.

Organization of the Paper. We present the definitions and notation used throughout this work in Section 2. Sections 3–5 form the bulk of our paper and are described below.

1.1 Our Contributions

In Section 3 we consider the inapproximability of **CSM**. Oveis Gharan and Vondrak [8] used a symmetry gap construction [25] to prove 0.491-inapproximability of **CSM** in the special case where the matroid constraint is a cardinality constraint. Our first result improves the inapproximability factor for a cardinality constraint to 0.478 using a modified construction, matching the factor of the current best inapproximability result for **CSM** in the general case (also due to [8]).



■ **Figure 1** Graphical presentation of results for Sections 4 and 5. Following the convention of Bodek and Feldman [1], the x and y axes represent the coefficients of ℓ and f , respectively. We use blue for approximation algorithms and red for inapproximability results, and the shaded area represents the gap between the best known approximation algorithms and inapproximability results.

► **Theorem 3.** *There exist instances of the problem $\max\{f(S) : S \subseteq \mathcal{N} \text{ and } |S| \leq w\}$ where f is non-negative submodular such that a 0.478 -approximation would require exponentially many value queries.*

In Section 4, we present the first nontrivial (α, β) -approximation algorithm for **RegularizedCSM** where the sign of ℓ is unconstrained. Furthermore, the α we obtain for **RegularizedUSM** improves over that of [1] for all $\beta \in (0, 1)$. The approximability and inapproximability results that follow are displayed in Figure 1.

► **Theorem 4.** *For all $t \geq 0$, there is a $\left(\frac{te^{-t}}{t+e^{-t}} - \epsilon, \frac{t}{t+e^{-t}}\right)$ -approximation algorithm for **RegularizedUSM**. This algorithm achieves the same approximation guarantee for **RegularizedCSM** when $t \leq 1$.*

For certain values of β , we can achieve greater α for **RegularizedCSM** than that guaranteed by Theorem 4. Because the improvement is marginal and we do not have a closed form, our following result addresses only the specific case of $\beta = 0.7$. Note that Theorem 4 guarantees a $(0.277, 0.7)$ -approximation (by setting $t \approx 0.925$).

► **Theorem 5** ([22, Theorem 8.3]). *There is a $(0.280, 0.7)$ -approximation algorithm for **RegularizedCSM**.*

The proof of this result is much more involved than Theorem 4. Thus, due to space constraints, we defer this proof to the full version of this paper [22].

Finally, in Section 5 we present three inapproximability results for $f + \ell$ sums. The first shows that Theorem 4 is essentially tight for **RegularizedCSM** when $\beta \geq \frac{e}{e+1}$.

► **Theorem 6** (Inapproximability of **RegularizedCSM** Near $\beta = 1$). *For any $0 \leq \beta \leq 1$, there exist instances of **RegularizedCSM** with non-negative ℓ such that a $(1 - \beta + \epsilon, \beta)$ -approximation would require exponentially many value queries.*

Our last two results show improved inapproximability for **RegularizedUSM**. The former generalizes Theorem 1.3 of [1] in order to show improved inapproximability for a range of β when ℓ is not necessarily constrained to be non-positive.

► **Theorem 7** (Inapproximability of **RegularizedUSM**). *There are instances of **RegularizedUSM** where $(\alpha(\beta), \beta)$ is inapproximable for any $(\alpha(\beta), \beta)$ in Table 1 (deferred to the appendix). In particular, $(0.440, 1)$ is inapproximable.*

The latter shows stronger inapproximability specifically for $\beta = 1$.

► **Theorem 8** (Inapproximability of **RegularizedUSM**, $\beta = 1$). *There are instances of **RegularizedUSM** where $(0.408, 1)$ is inapproximable.*

The best prior $(\alpha, 1)$ -inapproximability result for **RegularizedUSM** is $(0.478, 1)$ due to Theorem 1.3 of [1], matching the 0.478-inapproximability result for **CSM** due to Oveis Gharan and Vondrak [8]. We note that as Bodek and Feldman [1] show inapproximability specifically for the case of non-positive ℓ , it is not too surprising that we can show improved inapproximability for general ℓ (though in the full version of this paper, we also show slightly improved inapproximability for non-positive ℓ [22, Theorem 5.6]). The gap between the best approximability and inapproximability results for **RegularizedUSM** remains quite large; in fact, it remains unclear whether an $(\epsilon, 1)$ -approximation algorithm exists for any $\epsilon > 0$.

1.2 Our Techniques

To show approximability, the main techniques we use are the *measured continuous greedy* introduced by Feldman et al. [7] and used by [2, 18], as well as the *distorted objective* introduced by Feldman [6] and used by [18]. For Theorem 5 specifically, we additionally require the analysis of the 0.385-approximation algorithm for **CSM** due to Buchbinder et al. [2] and the “guessing step” used by Sviridenko et al. [24]. A more comprehensive overview of these techniques can be found in the full version of this paper [22, Section A.1].

To show inapproximability, the main technique we use is the symmetry gap of Vondrak [25], and most of our symmetry gap constructions are based on those of Oveis Gharan and Vondrak [8].

2 Preliminaries

Set Functions. Let $\mathcal{N} \triangleq \{u_1, u_2, \dots, u_n\}$ denote the *ground set*. A set function $f: 2^{\mathcal{N}} \rightarrow \mathbb{R}$ is said to be *submodular* if for every two sets $S, T \subseteq \mathcal{N}$, $f(S) + f(T) \geq f(S \cup T) + f(S \cap T)$. Equivalently, f is said to be submodular if it satisfies the property of “diminishing returns.” That is, for every two sets $S \subseteq T \subseteq \mathcal{N}$ and an element $u \in \mathcal{N} \setminus T$, $f(u|S) \geq f(u|T)$, where $f(u|S) \triangleq f(S \cup \{u\}) - f(S)$ is the *marginal value* of u with respect to S . We use $f(u)$ as shorthand for $f(\{u\})$. All submodular functions are implicitly assumed to be non-negative unless otherwise stated.

A set function f is said to be *monotone* if for every two sets $S \subseteq T \subseteq \mathcal{N}$, $f(S) \leq f(T)$. A set function ℓ is said to be *linear* if there exist values $\{\ell_u \in \mathbb{R} \mid u \in \mathcal{N}\}$ such that for every set $S \subseteq \mathcal{N}$, $\ell(S) = \sum_{u \in S} \ell_u$. When considering the sum of a non-negative submodular function f and a linear function ℓ whose sign is unconstrained, define $\ell_+(S) \triangleq \sum_{u \in S} \max(\ell_u, 0)$ and $\ell_-(S) \triangleq \sum_{u \in S} \min(\ell_u, 0)$ to be the components of ℓ with positive and negative sign, respectively.

Value Oracles. We make the standard assumption that an algorithm for the sum $f + \ell$ does not have direct access to the representation of f ; instead, it may obtain information about f only through a *value oracle*. Given any query set $S \subseteq \mathcal{N}$, a value oracle for f returns $f(S)$ in polynomial time. On the other hand, the coefficients of ℓ are directly provided to the algorithm.

Multilinear Extensions. All vectors of reals are in bold (e.g., \mathbf{x}). Given two vectors $\mathbf{x}, \mathbf{y} \in [0, 1]^{\mathcal{N}}$, we define $\mathbf{x} \circ \mathbf{y}$ to be the coordinate-wise multiplication of \mathbf{x} and \mathbf{y} . Given a set function $f: 2^{\mathcal{N}} \rightarrow \mathbb{R}$, its *multilinear extension* is the function $F: [0, 1]^{\mathcal{N}} \rightarrow \mathbb{R}$ defined by $F(\mathbf{x}) = \mathbb{E}[f(\mathbf{R}(\mathbf{x}))]$, where $\mathbf{R}(\mathbf{x})$ is a random subset of \mathcal{N} including every element $u \in \mathcal{N}$ with probability \mathbf{x}_u , independently. One can verify that F is a multilinear function of its arguments as well an extension of f in the sense that $F(\mathbf{1}_S) = f(S)$ for every set $S \subseteq \mathcal{N}$. Here, $\mathbf{1}_S$ is the vector with value 1 at each $u \in S$ and 0 at each $u \in \mathcal{N} \setminus S$, and is known as the *characteristic vector* of the set S .

Matroid Polytopes. A *matroid* \mathcal{M} may be specified by a pair of a ground set \mathcal{N} and a family \mathcal{I} of independent sets. The *matroid polytope* \mathcal{P} corresponding to \mathcal{M} is defined to be $\text{conv}(\{\mathbf{1}_S \mid S \in \mathcal{I}\})$, where conv denotes the convex hull. Due to the matroid axioms, \mathcal{P} is guaranteed to be *down-closed*; that is, $0 \leq \mathbf{x} \leq \mathbf{y}$ and $\mathbf{y} \in \mathcal{P}$ imply $\mathbf{x} \in \mathcal{P}$. It is also well-known that \mathcal{P} is *solvable*; that is, linear functions can be maximized over \mathcal{P} in polynomial time [4, Section 2.3]. For **CSM** and **RegularizedCSM**, we let OPT denote any set such that $OPT \in \mathcal{I}$ (equivalently, $\mathbf{1}_{OPT} \in \mathcal{P}$), while for **USM** and **RegularizedUSM**, we let OPT denote any subset of \mathcal{N} . For example, in the context of **CSM**, $\mathbb{E}[f(T)] \geq \alpha f(OPT)$ is equivalent to $\forall S \in \mathcal{I}, \mathbb{E}[f(T)] \geq \alpha f(S)$.

Miscellaneous. We let ϵ denote any positive real. Many of our algorithms are “almost” (α, β) approximations in the sense that they provide an $(\alpha - \epsilon, \beta)$ -approximation in $\text{poly}(n, \frac{1}{\epsilon})$ time for any $\epsilon > 0$. Similarly, some of our results show $(\alpha + \epsilon, \beta)$ -inapproximability for any $\epsilon > 0$.

3 Inapproximability of Maximization with Cardinality Constraint

In this section, we prove Theorem 3. First, we provide the relevant definitions about proving inapproximability using the symmetry gap technique from Vondrak [25].

► **Definition 9** (Symmetrization). Let \mathcal{G} be a group of permutations over \mathcal{N} . For $\mathbf{x} \in [0, 1]^{\mathcal{N}}$, define the “symmetrization of \mathbf{x} ” as $\bar{\mathbf{x}} = \mathbb{E}_{\sigma \in \mathcal{G}}[\sigma(\mathbf{x})]$, where $\sigma \in \mathcal{G}$ is uniformly random and $\sigma(\mathbf{x})$ denotes \mathbf{x} with coordinates permuted by σ .

► **Definition 10** (Symmetry Gap). Let $\max\{f(S) : S \in \mathcal{F} \subseteq 2^{\mathcal{N}}\}$ be strongly symmetric with respect to a group \mathcal{G} of permutations over \mathcal{N} , meaning that for all $\sigma \in \mathcal{G}$ and $S \subseteq 2^{\mathcal{N}}$, $f(S) = f(\sigma(S))$ and $S \in \mathcal{F} \Leftrightarrow S' \in \mathcal{F}$ whenever $\overline{\mathbf{1}_S} = \overline{\mathbf{1}_{S'}}$. Define $P(\mathcal{F}) = \text{conv}(\{\mathbf{1}_I : I \in \mathcal{F}\})$ to be the polytope associated with \mathcal{F} . Let $\mathbf{OPT} \triangleq \max_{\mathbf{x} \in P(\mathcal{F})} F(\mathbf{x})$ and $\overline{\mathbf{OPT}} \triangleq \max_{\mathbf{x} \in P(\mathcal{F})} F(\bar{\mathbf{x}})$. Then the symmetry gap of $\max\{f(S) : S \in \mathcal{F}\}$ is defined as $\gamma \triangleq \frac{\overline{\mathbf{OPT}}}{\mathbf{OPT}}$.

► **Lemma 11** (Inapproximability due to Symmetry Gap). Let $\max\{f(S) : S \in \mathcal{F}\}$ be an instance of non-negative submodular maximization, strongly symmetric with respect to \mathcal{G} , with symmetry gap γ . Let \mathcal{C} be the class of instances $\max\{\tilde{f}(S) : S \in \tilde{\mathcal{F}}\}$ where \tilde{f} is non-negative

submodular and $\tilde{\mathcal{F}}$ is a refinement of \mathcal{F} . Then for every $\epsilon > 0$, any (even randomized) $(1 + \epsilon)\gamma$ -approximation algorithm for the class \mathcal{C} would require exponentially many queries to the value oracle for $\tilde{f}(S)$.

The formal definition of refinement can be found in [25]. The important thing to note is that $\tilde{\mathcal{F}}$ satisfies the same properties as \mathcal{F} . In particular, $\tilde{\mathcal{F}}$ preserves cardinality and matroid independence constraints. Before proving Theorem 3, we start with a related lemma.

► **Lemma 12** (Inapproximability of Cardinality Constraint on Subset of Domain). *Let T be some subset of the ground set \mathcal{N} . There exist instances of the problem $\max\{f(S) : S \subseteq \mathcal{N} \wedge |S \cap T| \leq w\}$ such that a 0.478-approximation would require exponentially many value queries.*

Proof. It suffices to provide \mathcal{F} , f , and \mathcal{G} satisfying the definitions of Lemma 11 with symmetry gap $\gamma < 0.478$. The construction is identical to that of [8, Theorem E.2], except we omit $|S \cap \{a, b\}| \leq 1$ from the definition of \mathcal{F} . Specifically, we define $\mathcal{N} \triangleq \{a, b, a_{1\dots k}, b_{1\dots k}\}$ and

$$\mathcal{F} \triangleq \{S \mid S \subseteq \mathcal{N} \wedge |S \cap \{a_{1\dots k}, b_{1\dots k}\}| \leq 1\}$$

instead of:

$$\mathcal{F}_{orig} \triangleq \{S \mid S \subseteq \mathcal{N} \wedge |S \cap \{a, b\}| \leq 1 \wedge |S \cap \{a_{1\dots k}, b_{1\dots k}\}| \leq 1\}.$$

Recall that Theorem E.2 of [8] defines the submodular function f as the sum of the weighted cut functions of two directed hyperedges $(\{a_1, a_2, \dots, a_k\}, a)$, $(\{b_1, b_2, \dots, b_k\}, b)$ and the undirected edge (a, b) (see Figure 4 of [8] for an illustration). Specifically, the weighted cut function on the directed hyperedge $(\{a_1, a_2, \dots, a_k\}, a)$ contributes $\kappa \triangleq 0.3513$ to the value of $f(S)$ if $S \cap \{a_1, \dots, a_k\} \neq \emptyset$ and $a \notin S$, and 0 otherwise. The weighted cut function on the directed hyperedge $(\{b_1, b_2, \dots, b_k\}, b)$ is defined in the same way. Finally, the weighted cut function on the undirected edge (a, b) contributes $1 - \kappa$ if $|S \cap \{a, b\}| = 1$ and 0 otherwise. Thus, the multilinear extension of f is as follows:

$$\begin{aligned} F(\mathbf{x}_a, \mathbf{x}_b, \mathbf{x}_{a_{1\dots k}}, \mathbf{x}_{b_{1\dots k}}) &\triangleq (1 - \kappa)(\mathbf{x}_a(1 - \mathbf{x}_b) + \mathbf{x}_b(1 - \mathbf{x}_a)) \\ &+ \kappa \left[\left(1 - \prod_{i=1}^k (1 - \mathbf{x}_{a_i})\right) (1 - \mathbf{x}_a) + \left(1 - \prod_{i=1}^k (1 - \mathbf{x}_{b_i})\right) (1 - \mathbf{x}_b) \right]. \end{aligned}$$

As in [8, Lemma 5.4], we let \mathcal{G} be the group of permutations generated by $\{\sigma_1, \sigma_2\}$, where

$$\sigma_1(a) = b, \sigma_1(b) = a, \sigma_1(a_i) = b_i, \sigma_1(b_i) = a_i$$

swaps the two hyperedges, and

$$\sigma_2(a) = a, \sigma_2(b) = b, \sigma_2(a_i) = a_{i \pmod{k} + 1}, \sigma_2(b_i) = b_i$$

rotates the tail vertices of the first hyperedge. It is easy to check that (f, \mathcal{F}) are strongly symmetric with respect to both σ_1 and σ_2 , and that the symmetrization of \mathbf{x} is as follows:

$$\bar{\mathbf{x}} = \mathbb{E}_{\sigma \in \mathcal{G}} [\sigma(\mathbf{x})] = \begin{cases} \bar{\mathbf{x}}_a = \bar{\mathbf{x}}_b = \frac{\mathbf{x}_a + \mathbf{x}_b}{2} \\ \bar{\mathbf{x}}_{a_1} = \dots = \bar{\mathbf{x}}_{a_k} = \bar{\mathbf{x}}_{b_1} = \dots = \bar{\mathbf{x}}_{b_k} = \frac{\sum_{i=1}^k (\mathbf{x}_{a_i} + \mathbf{x}_{b_i})}{2k}. \end{cases}$$

Observe that

$$\mathbf{OPT} \geq \max_{S \in \mathcal{F}} f(S) \geq f(\{a, b_1\}) = (1 - \kappa) + \kappa = 1.$$

Defining $q \triangleq \frac{\mathbf{x}_a + \mathbf{x}_b}{2}$ and $p \triangleq \frac{\sum_{i=1}^k (\mathbf{x}_{a_i} + \mathbf{x}_{b_i})}{2}$, the maximum of F over all symmetric \mathbf{x} is thus:

$$\begin{aligned} \overline{\text{OPT}} &= \max_{\mathbf{x} \in P(\mathcal{F})} F(\bar{\mathbf{x}}) = \max_{\mathbf{x} \in P(\mathcal{F})} \overbrace{F(q, q, p/k, p/k, \dots, p/k)}^{2k \text{ times}} \triangleq \max_{\mathbf{x} \in P(\mathcal{F})} \hat{F}(q, p) \\ &= (1 - \kappa)2q(1 - q) + \kappa 2(1 - q)(1 - (1 - p/k)^k) \\ &\approx (1 - \kappa)2q(1 - q) + \kappa 2(1 - q)(1 - e^{-p}) \end{aligned}$$

where the approximate equality holds as $k \rightarrow \infty$. Now,

$$\overline{\text{OPT}} = \max_{\mathbf{x} \in P(\mathcal{F})} F(\bar{\mathbf{x}}) = \max_{p \leq 1/2} \hat{F}(q, p) = \max_{p, q \leq 1/2} \hat{F}(q, p) = \max_{\mathbf{x} \in P(\mathcal{F}_{orig})} F(\bar{\mathbf{x}}) < 0.478.$$

The third equality holds (i.e., adding the constraint $q \leq 1/2$ has no effect) since $\hat{F}(q, p) \leq \hat{F}(1 - q, p)$ for $q \in (1/2, 1]$, while the inequality holds due to the proof of [8, Theorem E.2] and may be verified using a numerical optimizer. So the symmetry gap $\gamma = \frac{\overline{\text{OPT}}}{\text{OPT}}$ is less than 0.478, as desired. \blacktriangleleft

Now, to show Theorem 3, all we need to do is to convert the cardinality constraint on $S \cap T$ in Lemma 12 into a cardinality constraint on all of S .

Proof of Theorem 3. Again, it suffices to provide \mathcal{F} , f , and \mathcal{G} satisfying the definitions of Lemma 11 with symmetry gap $\gamma < 0.478$. We start with the construction from Lemma 12, replace each element a_i and b_i with t copies $a_{i,1}, \dots, a_{i,t}$, and $b_{i,1}, \dots, b_{i,t}$ and set $w \triangleq t + 1$. Then we redefine f such that F is as follows:

$$\begin{aligned} F(\mathbf{x}_a, \mathbf{x}_b, \mathbf{x}_{a_{1 \dots k, 1 \dots t}}, \mathbf{x}_{b_{1 \dots k, 1 \dots t}}) &\triangleq (1 - \kappa)(\mathbf{x}_a(1 - \mathbf{x}_b) + \mathbf{x}_b(1 - \mathbf{x}_a)) \\ &+ \kappa \left[\left(1 - \prod_{i=1}^k \left(1 - \frac{\sum_{j=1}^t \mathbf{x}_{a_{i,j}}}{t} \right) \right) (1 - \mathbf{x}_a) + \left(1 - \prod_{i=1}^k \left(1 - \frac{\sum_{j=1}^t \mathbf{x}_{b_{i,j}}}{t} \right) \right) (1 - \mathbf{x}_b) \right]. \end{aligned}$$

Importantly, f remains non-negative submodular and symmetric, with the new symmetrization being as follows for an appropriate choice of \mathcal{G} :

$$\bar{\mathbf{x}} = \mathbb{E}_{\sigma \in \mathcal{G}} [\sigma(\mathbf{x})] = \begin{cases} \bar{\mathbf{x}}_a = \bar{\mathbf{x}}_b = \frac{\mathbf{x}_a + \mathbf{x}_b}{2} \\ \bar{\mathbf{x}}_{a_{1,1}} = \dots = \bar{\mathbf{x}}_{a_{k,t}} = \bar{\mathbf{x}}_{b_{1,1}} = \dots = \bar{\mathbf{x}}_{b_{k,t}} = \frac{\sum_{i=1}^k \sum_{j=1}^t (\mathbf{x}_{a_{i,j}} + \mathbf{x}_{b_{i,j}})}{2kt}. \end{cases}$$

For example, we may define \mathcal{G} to be the group generated by $\{\sigma_1, \sigma_2, \sigma_3\}$ where

- σ_1 swaps a with b and $a_{i,j}$ with $b_{i,j}$;
- σ_2 takes $a_{i,j}$ to $a_{i \pmod k + 1, j}$ and leaves all other vertices unchanged;
- σ_3 takes $a_{1,j}$ to $a_{1, j \pmod t + 1}$ and leaves all other vertices unchanged.

It can be verified that $F(\bar{\mathbf{x}})$ may be written in terms of the same function of two variables $\hat{F}(q, p)$ from Lemma 12. Let q be as defined above and redefine $p \triangleq \frac{\sum_{i=1}^k \sum_{j=1}^t (\mathbf{x}_{a_{i,j}} + \mathbf{x}_{b_{i,j}})}{2t}$, so that:

$$\begin{aligned} F(\bar{\mathbf{x}}) &\triangleq \overbrace{F(q, q, p/k, p/k, \dots, p/k)}^{2kt \text{ times}} \triangleq \hat{F}(q, p) \\ &= (1 - \kappa)2q(1 - q) + \kappa 2(1 - q)(1 - (1 - p/k)^k) \\ &\approx (1 - \kappa)2q(1 - q) + \kappa 2(1 - q)(1 - e^{-p}) \end{aligned}$$

where the approximate equality holds as $k \rightarrow \infty$, same as before.

To finish, we must show that symmetry gap of f with respect to \mathcal{F} remains less than 0.478 as $t \rightarrow \infty$. As in the proof of Lemma 12,

$$\mathbf{OPT} \geq \max_{S: |S| \leq t+1} f(S) \geq f(\{a, b_{11}, \dots, b_{1t}\}) = 1,$$

$$\overline{\mathbf{OPT}} = \max_{\mathbf{x}: |\mathbf{x}| \leq t+1} F(\mathbf{x}) = \max_{\mathbf{x}: |\mathbf{x}| \leq t+1} \hat{F}(q, p) \leq \max_{p \leq \frac{t+1}{2t}} \hat{F}(q, p) \approx \max_{p \leq 1/2} \hat{F}(q, p) < 0.478,$$

where the gap between the two sides of the approximate equality goes to 0 as $t \rightarrow \infty$ because \hat{F} is Lipschitz continuous and its domain is bounded. So the symmetry gap $\gamma = \frac{\overline{\mathbf{OPT}}}{\mathbf{OPT}}$ is again less than 0.478, as desired. \blacktriangleleft

4 Approximability for Unconstrained ℓ

In this section, we prove Theorem 4. We need two lemmas, the first of which is trivial.

► **Lemma 13.** *When ℓ is unconstrained, there exists a $(0, 1)$ -approximation algorithm for *RegularizedCSM*.*

Proof. As noted in the preliminaries, linear functions can be maximized over a matroid polytope \mathcal{P} in polynomial time. \blacktriangleleft

For the next lemma, we generalize the guarantee of the *distorted measured continuous greedy* of Lu et al. [18].

► **Lemma 14.** *For unconstrained ℓ and any $t_f \in [0, 1]$, there is a polynomial-time algorithm for *RegularizedCSM* that returns $T \in \mathcal{I}$ such that*

$$\mathbb{E}[f(T) + \ell(T)] \geq (t_f e^{-t_f} - o(1))f(\mathbf{OPT}) + (1 - e^{-t_f} - o(1))\ell_+(\mathbf{OPT}) + t_f \ell_-(\mathbf{OPT}). \quad (1)$$

When $t_f > 1$, the algorithm provides the same approximation guarantee but is allowed to return any $T \subseteq \mathcal{N}$.

Proof. It suffices to show that for any $\epsilon > 0$, with high probability, Algorithm 1 from [18] generates $\mathbf{y}(t_f) \in [0, 1]^{\mathcal{N}}$ such that $\mathbf{y}(t_f) \in t_f \cdot \mathcal{P}$ and

$$\begin{aligned} F(\mathbf{y}(t_f)) + L(\mathbf{y}(t_f)) &\geq t_f e^{-t_f} f(\mathbf{OPT}) + (1 - e^{-t_f})\ell_+(\mathbf{OPT}) \\ &\quad + t_f \ell_-(\mathbf{OPT}) - \mathcal{O}(\epsilon M) \end{aligned} \quad (2)$$

in $\text{poly}(n, 1/\epsilon)$ time, where $M \triangleq \max\{\max_{u \in \mathcal{N}} f(u|\emptyset), -\min_{u \in \mathcal{N}} f(u|\mathcal{N} - u)\} > 0$. How to use $\mathbf{y}(t_f)$ to generate a set T satisfying the conditions in the statement of this lemma is standard and is deferred to the appendix.

First we briefly review the *measured continuous greedy* algorithm introduced by Feldman et al. [7]. The idea is to continuously evolve a solution $\mathbf{y}(t)$ from time $t = 0$ to time $t = t_f$ such that $\mathbf{y}(t) \in (t \cdot \mathcal{P}) \cap ((1 - e^{-t}) \cdot [0, 1]^{\mathcal{N}})$. At all times, $\mathbf{y}'(t) = \mathbf{z}(t) \circ (\mathbf{1}_{\mathcal{N}} - \mathbf{y}(t))$, where $\mathbf{z}(t) \in \mathcal{P}$. To transform this continuous process into an algorithm running in finite time, it is necessary to discretize time into timesteps of size δ , where δ evenly divides t_f . Then $\mathbf{y}(t + \delta) \triangleq \mathbf{y}(t) + \delta \mathbf{z}(t) \circ (\mathbf{1}_{\mathcal{N}} - \mathbf{y}(t))$. How small δ needs to be to achieve the desired approximation factor is given by a polynomial in terms of n and ϵ .

Algorithm 1 of [18] combines measured continuous greedy with Feldman's *distorted objective* [6]. Specifically, algorithm 1 of [18] defines the objective at time t to be

$$\Phi(t) = (1 - \delta)^{(t_f - t)/\delta} F(\mathbf{y}(t)) + L(\mathbf{y}(t)) \approx e^{t - t_f} F(\mathbf{y}(t)) + L(\mathbf{y}(t))$$

41:10 On Maximizing Sums of Non-Monotone Submodular and Linear Functions

and chooses $\mathbf{z}(t)$ so that with high probability,

$$\Phi(t + \delta) - \Phi(t) \geq \delta [e^{-t_f} f(OPT) + \ell(OPT)] - \frac{\delta}{t_f} \cdot \mathcal{O}(\epsilon M),$$

assuming that ℓ is non-positive [18, Lemma 3.7]. Summing this inequality over all $\frac{t_f}{\delta}$ timesteps yields the desired result for non-positive ℓ .¹

We claim that when the sign of ℓ is unconstrained, the following generalization of [1, Lemma 3.7] holds:

$$\Phi(t + \delta) - \Phi(t) \geq \delta \left[e^{-t_f} f(OPT) + (1 - \delta)^{t/\delta} \ell_+(OPT) + \ell_-(OPT) \right] - \frac{\delta}{t_f} \cdot \mathcal{O}(\epsilon M), \quad (3)$$

To show this, the only part of the proof of [18, Lemma 3.7] that needs to change is the part where [18, Lemma 3.6] is invoked. Lemma 3.6 of [18] states that for non-positive ℓ ,

$$L(\mathbf{y}(t + \delta)) - L(\mathbf{y}(t)) = \delta L(\mathbf{z}(t) \circ (\mathbf{1}_{\mathcal{N}} - \mathbf{y}(t))) \geq \delta \langle \ell, \mathbf{z}(t) \rangle.$$

For unconstrained ℓ , we obtain the following inequality instead:

$$\begin{aligned} L(\mathbf{y}(t + \delta)) - L(\mathbf{y}(t)) &= \delta L(\mathbf{z}(t) \circ (\mathbf{1}_{\mathcal{N}} - \mathbf{y}(t))) \\ &= \delta (\langle \ell_+, \mathbf{z}(t) \circ (\mathbf{1}_{\mathcal{N}} - \mathbf{y}(t)) \rangle + \langle \ell_-, \mathbf{z}(t) \circ (\mathbf{1}_{\mathcal{N}} - \mathbf{y}(t)) \rangle) \\ &\geq \delta (\langle \ell_+, \mathbf{z}(t) \rangle \cdot (1 - \delta)^{t/\delta} + \langle \ell_-, \mathbf{z}(t) \rangle), \end{aligned} \quad (4)$$

where the last inequality follows from [18, Lemma 3.1], which states that $\mathbf{y}_u(t) \leq 1 - (1 - \delta)^{\frac{t}{\delta}}$ for all $u \in \mathcal{N}$. It is easy to verify that Equation (3) follows after substituting Equation (4) in place of [18, Lemma 3.6] in the proof of [18, Lemma 3.7].

To finish given Equation (3), we just need to check that $\sum_{i=0}^{t_f/\delta-1} \delta(1-\delta)^i \geq 1 - (1-\delta)^{t_f/\delta} \geq 1 - e^{-t_f}$. Thus, Equation (2) has been proven. \blacktriangleleft

Proof of Theorem 4. The algorithm is described below.

■ **Algorithm 1** Simple RegularizedCSM (t).

-
- 1 Let $T \leftarrow$ the set returned by running Lemma 14 for $t_f = t$ time.
 - 2 Let $T' \leftarrow$ the set returned by Lemma 13 (trivial approximation).
 - 3 if $f(T) + \ell(T) \geq \ell(T')$ then return T .
 - 4 else return T' .
-

Now we show that the desired approximation factor is achieved. Disregard the factors of $o(1)$ in Lemma 14; they can always be taken into account later at the cost of introducing the factor of ϵ . Next, add $t + e^{-t} - 1$ times the inequality of Lemma 13 to the inequality of Lemma 14.

$$\begin{aligned} (t + e^{-t})\mathbb{E}[\max(f(T) + \ell(T), \ell(T'))] &\geq \mathbb{E}[f(T) + \ell(T)] + (t + e^{-t} - 1)\mathbb{E}[\ell(T')] \\ &\geq te^{-t}f(OPT) + t(\ell_+(OPT) + \ell_-(OPT)) \\ &= te^{-t}f(OPT) + t\ell(OPT). \end{aligned}$$

¹ Actually, the authors of [18] only described their algorithm for the case of $t_f = 1$. Though as noted in [1], the proof of their algorithm easily generalizes to arbitrary t_f . Of course, the step size δ must be adjusted accordingly.

To finish, divide both sides by $t + e^{-t}$ and return the set out of T and T' that gives the higher value of $f + \ell$. Thus, we have the desired result after accounting for the factors of $o(1)$:

$$\mathbb{E}[\max(f(T) + \ell(T), \ell(T'))] \geq \left(\frac{te^{-t}}{t + e^{-t}} - \epsilon \right) f(OPT) + \frac{t}{t + e^{-t}} \ell(OPT). \quad \blacktriangleleft$$

5 Inapproximability for Unconstrained ℓ

First, we state a generalization of the symmetry gap technique to the sum $f + \ell$ that we'll need for all of these results in this section.

► **Definition 15.** We say that $\max_{S \in \mathcal{F}} [f(S) + \ell(S)]$ is strongly symmetric with respect to a group of permutations \mathcal{G} if $\ell(S) = \ell(\sigma(S))$ for all $\sigma \in \mathcal{G}$ and (f, \mathcal{F}) are strongly symmetric with respect to \mathcal{G} as defined in Definition 10.

► **Lemma 16 (Inapproximability of (α, β) Approximations).** Let $\max_{S \in \mathcal{F}} [f(S) + \ell(S)]$ be an instance of non-negative submodular maximization, strongly symmetric with respect to a group of permutations \mathcal{G} . For any two constants $\alpha, \beta \geq 0$, if

$$\max_{\mathbf{x} \in \mathcal{P}(\mathcal{F})} [F(\bar{\mathbf{x}}) + L(\bar{\mathbf{x}})] < \max_{S \in \mathcal{F}} [\alpha f(S) + \beta \ell(S)],$$

then no polynomial-time algorithm for *RegularizedCSM* can guarantee a (α, β) -approximation. The same inapproximability holds for *RegularizedUSM* by setting $\mathcal{F} = 2^{\mathcal{N}}$.

Proof. Theorem 3.1 of [1] proves this lemma only for the special case of $\mathcal{F} = 2^{\mathcal{N}}$ because the proof of [1, Lemma A.3] cites a special case of [25, Lemma 3.3] that only applies for $\mathcal{F} = 2^{\mathcal{N}}$. It suffices to modify the proof to cite the full version of [25, Lemma 3.3] instead. \blacktriangleleft

Now we are ready to prove Theorems 6–8. We note that the proofs of these results require the aid of a computer to verify. A link to the relevant code and a graph comparing these theorems to previous results are included in the appendix.

Proof of Theorem 6. Define $\alpha \triangleq 1 - \beta + \epsilon$. By Lemma 16, it suffices to construct a submodular function f satisfying

$$\max_{\mathbf{x} \in \mathcal{P}} [F(\bar{\mathbf{x}}) + L(\bar{\mathbf{x}})] < \max_{S \in \mathcal{I}} [\alpha \cdot f(S) + \ell(S)]. \quad (5)$$

We use the same f that Vondrak [25] uses for proving the inapproximability of maximization over matroid bases. Specifically, define $\mathcal{N} = \{a_1, \dots, a_k, b_1, \dots, b_k\}$ and let f correspond to the sum of directed cut functions of k disjoint arcs; that is, $f(S) \triangleq \sum_{i=1}^k [a_i \in S \text{ and } b_i \notin S]$. Its multilinear extension is $F(\mathbf{x}_{a_1 \dots a_k}, \mathbf{x}_{b_1 \dots b_k}) = \sum_{i=1}^k \mathbf{x}_{a_i} (1 - \mathbf{x}_{b_i})$. We define \mathcal{I} to consist of precisely the subsets of \mathcal{N} that contain at most one element from a_1, \dots, a_k and at most $k - 1$ elements from b_1, \dots, b_k , resulting in the following matroid independence polytope:

$$\mathcal{P} = \left\{ (\mathbf{x}_{a_i}, \mathbf{x}_{b_i}) \mid \sum_{i=1}^k \mathbf{x}_{a_i} \leq 1 \text{ and } \sum_{i=1}^k \mathbf{x}_{b_i} \leq k - 1 \right\}.$$

Finally, we define ℓ as $\ell(a_i) = 0, \ell(b_i) = \frac{1}{k}$. Then the RHS of Equation (5) is at least:

$$\max_{S \in \mathcal{I}} [\alpha f(S) + \beta \ell(S)] \geq (\alpha f + \beta \ell)(\{a_1, b_2, b_3, \dots, b_k\}) = \alpha + \beta \cdot \frac{k-1}{k},$$

41:12 On Maximizing Sums of Non-Monotone Submodular and Linear Functions

while the LHS of Equation (5) is:

$$\begin{aligned} \max_{\mathbf{x} \in \mathcal{P}} [F(\bar{\mathbf{x}}) + L(\bar{\mathbf{x}})] &= \max_{0 \leq p \leq 1/k, 0 \leq q \leq (k-1)/k} [kp(1-q) + q] \\ &= \max_{0 \leq p \leq 1/k, 0 \leq q \leq (k-1)/k} [q(1-kp) + kp] \\ &= \max_{0 \leq p \leq 1/k} [(1-kp) + kp] = 1, \end{aligned}$$

where the third equality follows because the expression is always maximized by setting $q = \frac{k-1}{k}$. For sufficiently large k we have

$$\alpha + \beta \cdot \frac{k-1}{k} \geq (\alpha + \beta) \frac{k-1}{k} = (1 + \epsilon) \cdot \frac{k-1}{k} > 1.$$

Equation (5) follows. ◀

In fact, Theorem 6 shows that Theorem 4 is essentially tight near $\beta = 1$ for **RegularizedCSM**. On the other hand, Theorem 6 cannot possibly apply to **RegularizedUSM** because Theorem 4 achieves $(1 - \beta + \epsilon, \beta)$ -approximations for β close to one.

► **Corollary 17** (Tight **RegularizedCSM** Near $\beta = 1$). *There is a $(1 - \beta - \epsilon, \beta)$ -approximation algorithm for **RegularizedCSM** for any $\frac{\epsilon}{e+1} \leq \beta < 1$, almost matching the bound of Theorem 6.*

Proof. Setting $t = 1$, the output of Theorem 4 is both a $(\frac{1}{e+1} - \epsilon, \frac{\epsilon}{e+1})$ -approximation and a $(0, 1)$ -approximation for **RegularizedCSM**. Therefore it is also an (α, β) -approximation for all (α, β) lying above the segment connecting $(\frac{1}{e+1} - \epsilon, \frac{\epsilon}{e+1})$ and $(0, 1)$. ◀

Proof of Theorem 7. Set f to be the same as defined in Lemma 12, and define $S \triangleq \{a, b_1\}$. For a fixed β , we can show (α, β) -inapproximability using Lemma 16 if it is possible to choose ℓ and κ such that:

$$\max_{\mathbf{x} \in [0,1]^{\mathcal{N}}} [F(\bar{\mathbf{x}}) + L(\bar{\mathbf{x}})] < \alpha f(\{a, b_1\}) + \beta \ell(\{a, b_1\}) = \alpha + \beta \ell(\{a, b_1\}),$$

which is equivalent to

$$\max_{\mathbf{x} \in [0,1]^{\mathcal{N}}} [F(\bar{\mathbf{x}}) + L(\bar{\mathbf{x}})] - \beta \ell(\{a, b_1\}) < \alpha.$$

For a fixed β , our goal is to choose ℓ and κ to minimize the LHS of the above inequality. Theorem 1.3 of [1] sets $\ell_a = \ell_b = 0$, and then chooses κ and $\ell_{a_1 \dots k} = \ell_{b_1 \dots k} \triangleq \ell_p$ in order to minimize the quantity

$$\begin{aligned} \max_{\mathbf{x} \in [0,1]^{\mathcal{N}}} [F(\bar{\mathbf{x}}) + L(\bar{\mathbf{x}})] - \beta \ell(\{a, b_1\}) &= \max_{\mathbf{x} \in [0,1]^{\mathcal{N}}} [F(\bar{\mathbf{x}}) + L(\bar{\mathbf{x}})] - \beta \ell_p \\ &\approx \max_{0 \leq q \leq 1, 0 \leq p} [(1 - \kappa)2q(1 - q) + \kappa 2(1 - q)(1 - e^{-p}) + 2p\ell_p] - \beta \ell_p \end{aligned}$$

However, allowing $\ell_a = \ell_b \triangleq \ell_q$ to be nonzero gives better bounds for all β . That is, our goal is to compute

$$\min_{0 \leq \kappa \leq 1, \ell_q, \ell_p} \left[\max_{0 \leq q \leq 1, 0 \leq p} [(1 - \kappa)2q(1 - q) + \kappa 2(1 - q)(1 - e^{-p}) + 2p\ell_p + 2q\ell_q] - \beta(\ell_p + \ell_q) \right].$$

We can approximate the optimal value by brute forcing over a range of (κ, ℓ_q, ℓ_p) . The best triples we found are displayed in Table 1. Allowing ℓ_q to be negative gives superior bounds for β near zero, while allowing ℓ_q to be positive gives superior bounds for β near one. In particular, for $\beta = 1$, taking $\kappa = 0.6022$, $\ell_p = -0.1819$, and $\ell_q = 0.2152$ gives $\alpha(\beta) \approx 0.4390 < 0.440$. ◀

We can do slightly better than Theorem 7 for β very close to one with a construction inspired by [1, Theorem 1.6].

Proof of Theorem 8. Again, we use Lemma 16. Let $\mathcal{N} \triangleq \{a_{1\dots k}, b_{1\dots k}\}$, and define f as the directed cut function of a generalized hyperedge $(a_{1\dots k}; b_{1\dots k})$; that is, the generalized hyperedge is said to be cut by S if S contains at least one of the tails of the hyperedge $(a_{1\dots k})$ but not all of the heads of the hyperedge $(b_{1\dots k})$:

$$f(S) \triangleq [S \cap \{a_{1\dots k}\} \neq \emptyset] \cdot [\{b_{1\dots k}\} \not\subseteq S].$$

Also define $\ell(a_i) = -0.2037$, $\ell(b_i) = 0.2037$, $p \triangleq \sum_{i=1}^k \mathbf{x}_{a_i}$ and $q \triangleq k - \sum_{i=1}^k \mathbf{x}_{b_i}$, and \mathcal{G} such that a_1, \dots, a_k and b_1, \dots, b_k are symmetric. Then as $k \rightarrow \infty$,

$$F(\bar{\mathbf{x}}) = \left(1 - (1 - p/k)^k\right) \left(1 - (1 - q/k)^k\right) \approx (1 - e^{-p})(1 - e^{-q}).$$

Now,

$$\begin{aligned} \max_{\mathbf{x} \in [0,1]^{\mathcal{N}}} [F(\bar{\mathbf{x}}) + L(\bar{\mathbf{x}})] &= \max_{p,q \geq 0} [(1 - e^{-p})(1 - e^{-q}) - 0.2037(p + q) + 0.2037k] \\ &= 0.2037k, \end{aligned}$$

where the last equality follows since the maximum is attained at $p = q = 0$, which may be verified using a numerical optimizer. On the other hand,

$$\max_S [\alpha f(S) + \ell(S)] \geq (f + \ell)(\{a_1, b_{1\dots k-1}\}) = \alpha + 0.2037(k - 2).$$

It follows from Lemma 16 that we have shown $(\alpha, 1)$ -inapproximability for any α satisfying

$$0.2037k < \alpha + 0.2037(k - 2) \implies \alpha > 0.4074. \quad \blacktriangleleft$$

References

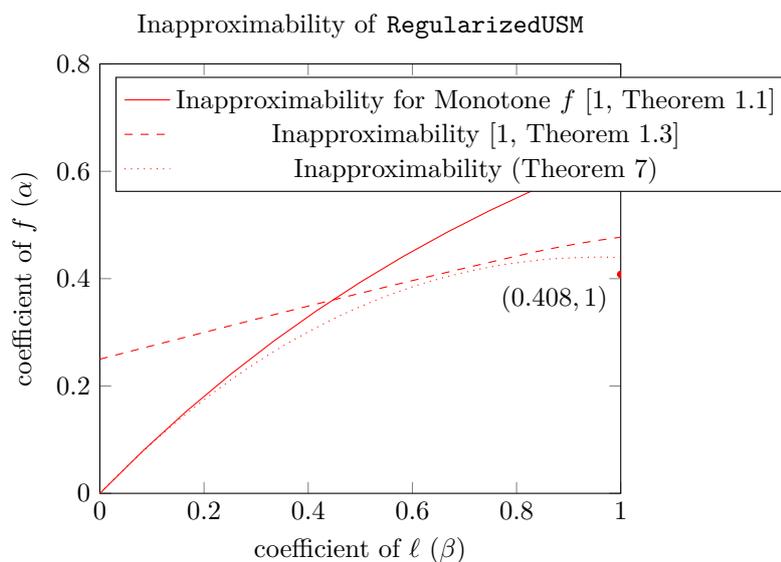
- 1 Kobi Bodek and Moran Feldman. Maximizing sums of non-monotone submodular and linear functions: Understanding the unconstrained case, 2022. To appear in ESA 2022. doi:10.48550/ARXIV.2204.03412.
- 2 Niv Buchbinder and Moran Feldman. Constrained submodular maximization via a nonsymmetric technique. *Mathematics of Operations Research*, 44(3):988–1005, 2019.
- 3 Niv Buchbinder, Moran Feldman, Joseph Naor, and Roy Schwartz. A tight linear time $(1/2)$ -approximation for unconstrained submodular maximization. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 649–658, 2012. doi:10.1109/FOCS.2012.73.
- 4 Gruia Calinescu, Chandra Chekuri, Martin Pal, and Jan Vondrák. Maximizing a monotone submodular function subject to a matroid constraint. *SIAM Journal on Computing*, 40(6):1740–1766, 2011.
- 5 Uriel Feige, Vahab S Mirrokni, and Jan Vondrák. Maximizing non-monotone submodular functions. *SIAM Journal on Computing*, 40(4):1133–1153, 2011.
- 6 Moran Feldman. Guess free maximization of submodular and linear sums. *Algorithmica*, 83(3):853–878, 2021.
- 7 Moran Feldman, Joseph Naor, and Roy Schwartz. A unified continuous greedy algorithm for submodular maximization. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, pages 570–579. IEEE, 2011.
- 8 Shayan Oveis Gharan and Jan Vondrák. Submodular maximization by simulated annealing. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, pages 1098–1116. SIAM, 2011.

- 9 Michael Gygli, Helmut Grabner, and Luc Van Gool. Video summarization by learning submodular mixtures of objectives. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3090–3098, 2015.
- 10 Chris Harshaw, Moran Feldman, Justin Ward, and Amin Karbasi. Submodular maximization beyond non-negativity: Guarantees, fast algorithms, and applications. In *International Conference on Machine Learning*, pages 2634–2643. PMLR, 2019.
- 11 Stefanie Jegelka and Jeff Bilmes. Submodularity beyond submodular energies: coupling edges in graph cuts. In *CVPR 2011*, pages 1897–1904. IEEE, 2011.
- 12 Ehsan Kazemi, Shervin Minaee, Moran Feldman, and Amin Karbasi. Regularized submodular maximization at scale. In *International Conference on Machine Learning*, pages 5356–5366. PMLR, 2021.
- 13 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*, pages 137–146, 2003.
- 14 Andreas Krause and Daniel Golovin. Submodular function maximization. *Tractability*, 3:71–104, 2014.
- 15 Andreas Krause and Carlos Guestrin. Submodularity and its applications in optimized information gathering. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(4):1–20, 2011.
- 16 Andreas Krause, Ajit Singh, and Carlos Guestrin. Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. *Journal of Machine Learning Research*, 9(2), 2008.
- 17 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.
- 18 Cheng Lu, Wenguo Yang, and Suixiang Gao. Regularized non-monotone submodular maximization, 2021. doi:10.48550/ARXIV.2103.10008.
- 19 George L Nemhauser and Laurence A Wolsey. Best algorithms for approximating the maximum of a submodular set function. *Mathematics of operations research*, 3(3):177–188, 1978.
- 20 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.
- 21 Sofia Maria Nikolakaki, Alina Ene, and Evimaria Terzi. An efficient framework for balancing submodularity and cost. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pages 1256–1266, 2021.
- 22 Benjamin Qi. On maximizing sums of non-monotone submodular and linear functions, 2022. doi:10.48550/ARXIV.2205.15874.
- 23 Jianbing Shen, Zhiyuan Liang, Jianhong Liu, Hanqiu Sun, Ling Shao, and Dacheng Tao. Multi-object tracking by submodular optimization. *IEEE transactions on cybernetics*, 49(6):1990–2001, 2018.
- 24 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.
- 25 Jan Vondrák. Symmetry and approximability of submodular maximization problems. *SIAM Journal on Computing*, 42(1):265–304, 2013.
- 26 Kai Wei, Yuzong Liu, Katrin Kirchhoff, and Jeff Bilmes. Using document summarization techniques for speech data subset selection. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 721–726, 2013.

A Appendix

■ **Table 1** Inapproximability of $(\alpha(\beta), \beta)$ -approximations for **RegularizedUSM** with unconstrained ℓ (Theorem 7).²

β	$\alpha(\beta)$	κ	ℓ_p	ℓ_q
0.1	0.0935	0.6708	-0.6064	-0.2644
0.2	0.1743	0.6598	-0.5370	-0.2102
0.3	0.2432	0.6577	-0.4809	-0.1504
0.4	0.3008	0.6519	-0.4261	-0.0940
0.5	0.3476	0.6465	-0.3778	-0.0407
0.6	0.3844	0.6390	-0.3309	0.0128
0.7	0.4114	0.6315	-0.2881	0.0655
0.8	0.4293	0.6226	-0.2498	0.1155
0.9	0.4383	0.6144	-0.2142	0.1668
1.0	0.4390	0.6022	-0.1819	0.2152



■ **Figure 2** Previous and new inapproximability results for **RegularizedUSM** with an unconstrained linear function ℓ . $(0.408, 1)$ -inapproximability is due to Theorem 8.

A.1 Omitted Proof

Proof of Lemma 14 (Omitted Details). We need to show that we can use an algorithm that outputs $\mathbf{y}(t_f)$ satisfying Equation (2) to generate $\mathbf{y} \in t_f \cdot \mathcal{P}$ such that

$$\mathbb{E}[F(\mathbf{y}) + L(\mathbf{y})] \geq (t_f e^{-t_f} - o(1))f(OPT) + (1 - e^{-t_f} - o(1))\ell_+(OPT) + t_f \ell_-(OPT), \quad (6)$$

² All values of $\alpha(\beta)$ are rounded to the nearest multiple of 10^{-4} . The numbers here are marginally better than those originally presented in the full paper [22, Table 4].

41:16 On Maximizing Sums of Non-Monotone Submodular and Linear Functions

where the RHS of Equation (6) is identical to that of Equation (1). Once we have \mathbf{y} satisfying Equation (6), we can use *pipage rounding* to round \mathbf{y} to an integral solution $T \in \mathcal{I}$ if $t_f \leq 1$ or $T \subseteq \mathcal{N}$ otherwise [25]. Specifically, given $\mathbf{y} \in \mathcal{P}$, pipage rounding generates $T \in \mathcal{I}$ such that $\mathbb{E}[\mathbf{1}_T] = \mathbf{y}$ and $\mathbb{E}[F(\mathbf{1}_T) + L(\mathbf{1}_T)] \geq F(\mathbf{y}) + L(\mathbf{y})$.

We note that $\mathbf{y}(t_f)$ satisfying Equation (2) with high probability does not necessarily satisfy Equation (6) if either:

1. $M = \omega(f(OPT) + \ell_+(OPT))$.
2. Equation (2) does not hold.

However, we claim that output of the following algorithm *does* satisfy the conditions of Lemma 14:

■ **Algorithm 2** Modified Distorted Measured Continuous Greedy (t).

- 1 Order the elements of \mathcal{N} such that $f(u_1) \leq f(u_2) \leq \dots \leq f(u_n)$.
 - 2 **for** $i = 0$ **to** n **do**
 - 3 Let $\mathbf{y}_i \leftarrow$ the result of running [18, Algorithm 1] on $\mathcal{N}_i = \{u_1, u_2, \dots, u_i\}$.
 - 4 Let $T_i \leftarrow$ the result of applying pipage rounding to \mathbf{y}_i .
 - 5 **return** the set T_i maximizing $f(T_i) + \ell(T_i)$.
-

It suffices to show that the two issues mentioned above are now resolved.

1. Assume

$$\emptyset \neq OPT \in \operatorname{argmax}_{OPT \subseteq \mathcal{N}} [t_f e^{-t_f} f(OPT) + (1 - e^{-t_f}) \ell_+(OPT) + t_f \ell_-(OPT)].$$

Defining $j \triangleq \max\{i \mid u_i \in OPT\}$, it follows that

$$t_f e^{-t_f} f(OPT) + (1 - e^{-t_f}) \ell_+(OPT) \geq t_f e^{-t_f} \max(f(u_j), f(\emptyset)) \geq t_f e^{-t_f} \frac{M_j}{j}, \quad (7)$$

where M_j is the value of M when restricted to \mathcal{N}_j . Thus, $M_j \leq n(f(OPT) + \ell_+(OPT))$. By choosing $\epsilon = o(\frac{1}{n})$ in [18, Algorithm 1], Equation (2) implies that \mathbf{y}_j satisfies Equation (6) with high probability.

2. Regardless of whether Equation (2) holds, the set T returned by Algorithm 2 always satisfies $f(T) + \ell(T) \geq f(T_0) + \ell(T_0) \geq 0$. Thus, if T satisfies Lemma 14 with high probability, it also satisfies Lemma 14 in expectation. ◀

A.2 Code

A Python notebook with code for all theorems in the full paper (including Theorems 5, 7, and 8 of this paper) can be found [here](#).