Partial Monotonicity for Submodular Maximization with a Knapsack Constraint

Tong Cheng
Nanyang Technological University
tong030@e.ntu.edu.sg

Xueyan Tang
Nanyang Technological University
asxytang@ntu.edu.sg

Abstract

Submodular maximization has become increasingly important in the fields of machine learning and data mining. For general submodular maximization without monotonicity, many previous analyses provide poor approximation guarantees, especially for submodular functions that are approximately monotone. To address this issue, the research community has proposed a continuous metric called the monotonicity ratio for submodular functions. The monotonicity ratio has been studied for submodular maximization under no constraint, a cardinality constraint, and a matroid constraint. However, the implications of using the monotonicity ratio for submodular maximization with a knapsack constraint remain unclear. Although a knapsack constraint can be regarded as a continuous extension of the cardinality constraint with non-uniform costs, the gap in analysis between these two constraints is substantial. In this paper, we analytically show that many previously proposed algorithms for monotone submodular maximization with a knapsack constraint can achieve improved approximation guarantees under partial monotonicity with a simple modification: enforcing positive marginal gain. In addition, we evaluate our proposed algorithms for two machine learning applications of movie recommendation and influence-and-exploit marketing, showing that our algorithms could achieve better empirical performance than state-of-the-art algorithms under partial monotonicity.

1 INTRODUCTION

Submodular maximization has emerged over the past two decades as an indispensable tool for tackling discrete optimization problems in machine learning. Submodularity captures the property of diminishing returns in set functions, which is ubiquitous in a wide range of applications such as maximum entropy sampling (Shewry and Wynn, 1987; Fampa and Lee, 2022), maximum facility location (Cornuejols et al., 1977; Ageev and Sviridenko, 1999) and influence maximization in social networks (Chen et al., 2009; Lei et al., 2015; Li et al., 2018). Submodular maximization is NP-hard (Feige, 1998). Hence, a substantial body of research has focused on the development and analysis of approximation algorithms for submodular maximization. In the context of constrained submodular maximization, the cardinality, matroid, and knapsack constraints represent three of the most commonly studied formulations. Among these, the knapsack-constrained setting is generally considered more challenging, due to the variability and heterogeneity of the associated cost functions. This complexity makes the optimization process more difficult compared to the relatively uniform structures of cardinality and matroid constraints (Amanatidis et al., 2020). In this paper, we study the problem of non-monotone submodular maximization with a knapsack constraint.

Many prior studies have focused on monotone submodular maximization with a knapsack constraint, yielding a range of approximation ratios (Wolsey, 1982a; Khuller et al., 1999; Yaroslavtsev et al., 2020; Tang et al., 2021; Feldman et al., 2023). Beyond the monotone setting, the problem of non-monotone submodular maximization with a knapsack constraint is also of practical significance, including applications such as computing core values in supermodular games (Schulz and Uhan, 2007) and optimizing marketing strategies for revenue maximization in social networks (Hartline et al., 2008). Hence, substantial research efforts have been devoted to non-monotone submodular maximiza-

---

Proceedings of the 29th International Conference on Artificial Intelligence and Statistics (AISTATS) 2026, Tangier, Morocco. PMLR: Volume 300.

tion with a knapsack constraint (Mirzasoleiman et al., 2016; Amanatidis et al., 2020, 2021; Cui et al., 2023).

However, there exists a significant gap of the approximation ratios between the monotone case (Wolsey, 1982a; Khuller et al., 1999; Yaroslavtsev et al., 2020; Tang et al., 2021; Feldman et al., 2023) and the non-monotone case (Mirzasoleiman et al., 2016; Amanatidis et al., 2020, 2021; Cui et al., 2023). Specifically, the approximation ratios for monotone submodular maximization with a knapsack constraint (e.g., 0.402, 0.5, 0.632) are substantially larger than those for the non-monotone counterpart (e.g., 0.167, 0.171, 0.25). Intuitively, for submodular functions that are nearly monotone, one would expect the approximation ratio to degrade smoothly rather than abruptly. Relying solely on results from strictly monotone or fully non-monotone settings overlooks this continuous variation.

Instead of treating monotonicity as a binary property, a continuous metric called monotonicity ratio, proposed in (Mualem and Feldman, 2022), quantifies how close a submodular function is to being monotone. This concept is a useful tool for enhancing the approximation analysis. Mualem and Feldman (2022) leveraged the monotonicity ratio to analyze three submodular maximization problems: unconstrained submodular maximization, submodular maximization with a cardinality constraint, and submodular maximization with a matroid constraint. While Mualem and Feldman (2022) derives convex-combination approximation ratios for unconstrained, cardinality, and matroid settings, it leaves the knapsack setting unresolved. We close this gap by developing a monotonicity-ratio analysis of greedy-based algorithms for submodular maximization under a knapsack constraint. Our approximation ratio is continuous with the monotonicity ratio and hence bridge the classical monotone and fully non-monotone cases.

However, the proof techniques introduced in (Mualem and Feldman, 2022) do not directly extend from submodular maximization with a cardinality constraint or matroid constraint into submodular maximization with a knapsack constraint. In the cardinality or matroid constraint setting, the greedy algorithms select elements by maximizing the marginal gain. Most analyses lower bound the marginal gain based on the intermediate solution set by some multiplier times the optimal function value. By telescoping these inequalities over all intermediate solution sets, we could obtain the approximation ratio of the greedy algorithm. However, in the knapsack constraint setting, the greedy algorithms select elements by maximizing the density instead of marginal gain. The analyses can only lower bound the density based on the intermediate solution sets by some multiplier times the optimal value. Since the cost of selected elements in the solution set are different, telescoping these inequalities of densities over all intermediate solution sets does not recover the function value of the final solution and hence fails to give an approximation ratio of the greedy algorithm. This discrepancy necessitates more specialized analysis to derive an approximation ratio under a knapsack constraint setting.

Our contributions. We introduce two novel techniques designed to overcome these knapsack-specific analytical barriers: 1) Positive-marginal Filter (PF): A technique that identifies and excludes elements with negative marginal gains. This ensures the monotone growth of the objective value throughout the greedy process 2) Residual Lift (RL): A technique that constructs a residual set function for the analysis. We prove that for any non-negative, m-monotone submodular function, the residual set function has monotonicity ratio of $m' = m + \frac{m^2 - m}{\lambda - m}$ for any tunable parameter $\lambda \geq 1$. With these two techniques, we prove a PF+RL growth template theorem as follows:

Theorem 1.1. *For any PF-respecting algorithm with an output S_{out} and a seed set Y (e.g., $\emptyset$, the best singleton, or the best two-set), it holds that*

$$f(S_{\text{out}}) \geq A(m, \lambda) f(\text{OPT}) + B(m, \lambda) f(Y).$$

(The same as Theorem 3.1).

Instantiating the PF+RL growth template theorem with five PF-respecting greedy algorithms: Positive Modified Greedy (PMG), Positive Greedy+Max (PG+Max), Two Set Enumeration Positive Greedy (2EPG), One Set Enumeration Positive Greedy+Max (1EPG+Max), and Sample Greedy (SG), we instantiate $A(m, \lambda)$ and $B(m, \lambda)$ as a corollary of Theorem 1.1, yielding a continuous approximation curve as a function of the monotonicity ratio m. In particular, the values of the approximation ratios when the monotonicity ratio $m = 1$ are consistent with the classical monotone results from (Khuller et al., 1999; Yaroslavtsev et al., 2020; Feldman et al., 2023; Tang et al., 2021; Amanatidis et al., 2020).

To facilitate a direct comparison between methods, Table 1 reports conservative, closed-form expressions for the approximation ratios. We provide tighter approximation ratios with respect to monotonicity ratio in Appendix C. These sharper analytical results are further visualized in Figure 1.

Guided by our theoretical results, we outline a deployment guideline. From Figure 1, one may select the pointwise maximum of the five approximation ratios for each $m \in [0, 1]$. When $m \leq 0.2$, SG is preferable, as it remains non-trivial as m goes to 0. For $0.2 < m \leq 0.6$, PG+Max offers a parameter-free $m/2$

Algorithm	Approximation Ratio
PMG	$m(1-1/e)/2$
PG+Max	$m/2$
2EPG	$\max\{m(1-1/e)/2, \text{goes to } m(1-1/e)\}$
1EPG+Max	$\max\{m/2, \text{approximates to } m(1-1/e)\}$
SG	$(m+1)/6$ $(m \geq 1/5)$; otherwise 0.1716

Table 1: Conservative, closed-form approximation ratios for the five PF-respecting instantiations. The tighter (and slightly more intricate) m-dependent curves are plotted in Figure 1 and stated precisely in Appendix C.

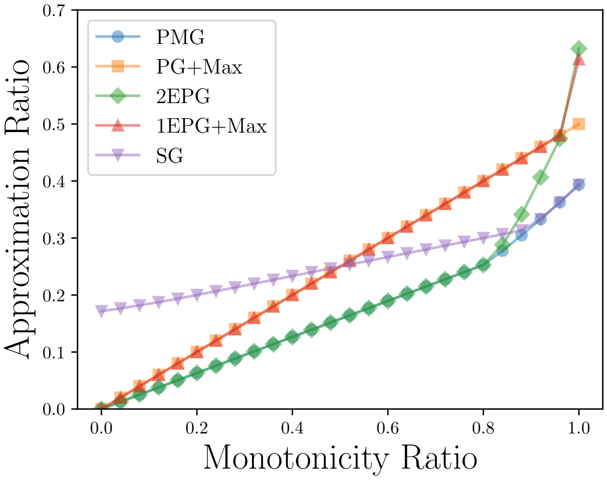

Figure 1: Approximation ratios with respect to monotonicity ratio m for the five PF-respecting instantiations.

approximation ratio with reduced query complexity. For $m > 0.6$, a light enumeration step activates a seeding dividend: 2EPG or 1EPG+Max typically achieves the maximum approximation ratio. When enumeration cost is prohibitive, PMG and SG have lower query complexity and hence are favorable. Specifically, when the monotonicity ratio m is large PMG is generally preferable whereas when m is small SG is advantageous.

Beyond analyzing approximation ratios using PF+RL growth template theorem, another contribution of this paper is a new inapproximability framework that is asymmetry-aware and well-suited to knapsack constraint analysis. For cardinality constraint and matroid constraint, the feasible sets exhibit substantial permutation symmetry, making the symmetry-gap framework in (Vondrák, 2009) applicable. Mualem and Feldman (2022) extends the symmetry gap framework to partial monotonicity setting to derive inapproximability results for these two constraints. By contrast, a knapsack constraint breaks permutation symmetry since heterogeneous item costs make the feasible family variant and sets of the same cardinality are neither uniformly feasible nor share the same function value under permutations. Hence, the symmetry gap framework used in (Mualem and Feldman, 2022) does not transfer directly to the knapsack constraint case.

Using our asymmetry-aware inapproximability framework, we obtain the following theorem:

Theorem 1.2. *For any $\varepsilon > 0$, no polynomial-time algorithm achieves an approximation ratio larger than $\rho_{\text{card}}(m) - \frac{1-m}{512} + \varepsilon$ where $\rho_{\text{card}}(m)$ is the inapproximability result for the cardinality case from (Mualem and Feldman, 2022) (The same as Theorem 3.2).*

The proof of this theorem follows an information-theoretic indistinguishability framework derived from (Le Cam, 1986; Yao, 1977) by constructing two distributions over problem instances that are hard to distinguish from oracle transcripts and apply Yao's minimax principle. These two distributions $\mathcal{D}_0$ and $\mathcal{D}_1$ are statistically close under oracle access, so no polynomial-time algorithm can reliably distinguish them with $o(1)$ advantage. These two distributions exhibit a gap in optimal value aligned with $\rho_{\text{card}}(m)$. The key departure from (Mualem and Feldman, 2022) is that without symmetry under a knapsack constraint, we pad heavy items with carefully chosen costs and introduce light items so that the knapsack feasibility region encodes the hidden bit distinguishing $\mathcal{D}_0$ from $\mathcal{D}_1$. This yields a gap-preserving reduction from the cardinality lower bound to knapsack without assuming invariance under permutations of feasible sets.

We admit that the quantitative downgrade $(1-m)/512$ is not significant. But our major contribution here is the analytical framework: an asymmetry-aware indistinguishability method for knapsack that does not rely on the symmetric feasible-set structure used in (Mualem and Feldman, 2022).

2 PRELIMINARIES

Let V be a ground set of elements with size n. A set function $f : 2^V \to \mathbb{R}_{\geq 0}$ is a submodular function if for any subsets $A, B \subseteq V$, it holds that $f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$. Denote $f(e \mid A) := f(A \cup \{e\}) - f(A)$ as the *marginal gain* of element e over a base set $A \subseteq V$. Equivalently, if a set function f is submodular, then for any subsets $A \subseteq B \subseteq V$, it holds that $f(e \mid A) \geq f(e \mid B)$ for any $e \in V \setminus B$, i.e., the marginal gain of an element over a base set never increases as more elements are added into the base set.

The submodular maximization with a knapsack constraint problem is formulated as: $\max_{S \subseteq V : c(S) \leq b} f(S)$, where $c : V \to \mathbb{R}_{>0}$, $c(S) := \sum_{e \in S} c(e)$ is the cost function for the knapsack constraint and $b \in \mathbb{R}_{>0}$ is the budget. We call the ratio between the marginal gain $f(e \mid S)$ and cost $c(e)$ of element e as *density* $\rho(e \mid S) := f(e \mid S)/c(e)$. Let OPT denote an optimal solution of this problem.

A submodular function f is *monotone* if for any subsets $A \subseteq B \subseteq V$, it holds that $f(A) \leq f(B)$. To characterize the degree of monotonicity for non-monotone submodular functions, we adopt the concept of the monotonicity ratio from (Mualem and Feldman, 2022). Specifically, for a non-negative submodular function f, the monotonicity ratio is defined as $m := \min_{A \subseteq B \subseteq V} f(B)/f(A)$, where the ratio of function values $f(B)/f(A)$ is defined as 1 if $f(A) = 0$. A submodular function f is *m-monotone* if its monotonicity ratio is m.

3 THEORETICAL RESULTS

In this section, we first present the PF+RL growth template theorem and then instantiate the template with five PF-respecting algorithms, deriving explicit m-dependent approximation ratios. Finally, we provide an improved inapproximability results that extend the results of (Mualem and Feldman, 2022) to the knapsack setting.

Unlike a cardinality or matroid constraint where the greedy algorithm advances in uniform rank increments, the algorithm under a knapsack constraint advances in non-uniform cost steps and may terminate early due to a blocking item that does not fit the residual budget. This breaks the telescoping and averaging arguments that give a convex-combination of approximation ratios as in (Mualem and Feldman, 2022).

To guarantee a monotone growth solution sequence under partial monotonicity, we forbid negative-marginal moves and analyze a residual function that lifts the monotonicity ratio. Here we introduce two techniques.

- **Positive-marginal Filter (PF).** A greedy algorithm is *PF-respecting* if it only adds elements with non-negative marginal gains. PF rejects any elements with negative marginal gains into the solution set to prevent the objective function value from decreasing.
- **Residual Lift (RL).** Fix a parameter $\lambda \geq 1$. For a base set X, define the residual function $g_Y(S) := f(S \cup Y) - \frac{m}{\lambda} f(Y)$. When f is non-negative m-monotone submodular, the residual function $g_Y(S)$ is m'-monotone submodular where $m' = m'(m, \lambda) := m + \frac{m^2 - m}{\lambda - m}$ (See more details in Appendix B). Intuitively, RL provides an analytical tool for greedy algorithm starting from a seed set.

Collectively, PF guarantees monotone progress of the partial solution and RL provides a residual objective with an enhanced monotonicity ratio $m'(m, \lambda)$. Combining these two techniques, we obtain the PF+RL growth template theorem. For the five PF-respecting algorithms, the PF+RL growth template theorem applies to yield instance-specific approximation ratios.

3.1 PF+RL Growth Template

Denote $S_0 \subseteq S_1 \subseteq \cdots \subseteq S_{\text{out}}$ as the solution set sequence generated by a PF-respecting algorithm with an initial seed set $S_0 = Y$ (e.g., $Y = \emptyset$, the best singleton, or the best two-set). This algorithm halts when either no item fits the residual budget or the marginal gains of the remaining elements are negative.

Theorem 3.1. *For any PF-respecting algorithm with an output S_{out} and a seed set Y (e.g., $\emptyset$, the best singleton, or the best two-set), it holds that*

$$f(S_{\text{out}}) \geq A(m, \lambda) f(\text{OPT}) + B(m, \lambda) f(Y). \qquad (1)$$

Proof sketch. The PF technique ensures that the objective value of the solution remains monotonically increasing throughout the greedy selection process. At each step, the density of the chosen item lower-bounds a residual averaging of $f(\text{OPT})$, which is standard in the analysis of greedy algorithms for submodular maximization with a knapsack constraint. RL provides a lifted monotonicity m' so that the residual contributions from $\text{OPT} \setminus S_i$ do not cancel when non-uniform cost steps and early termination occur. When the seed set is not empty($|Y| \in \{1, 2\}$), a seeding dividend $B(m, \lambda) f(Y)$ arises. Summing the per-step inequalities up to the iteration round where the budget violation occurs and projecting back from g_Y to f yields (1). The full proof is given in Appendix C.1. ☐

Theorem 3.1 decomposes the worst-case approximation ratio into sum of a universal term $A(m, \lambda) f(\text{OPT})$ and an explicit seeding dividend $B(m, \lambda) f(Y)$. Crucially, the coefficients depend only on the specification of greedy algorithm and the seeding scheme (e.g., $|Y| \in \{0, 1, 2\}$), making the template a useful tool for the analyses of PF-respecting greedy algorithms.

Many previous works have developed constant approximation ratios for monotone submodular maximization with a knapsack constraint (Wolsey, 1982b,a; Khuller et al., 1999; Sviridenko, 2004; Yaroslavtsev et al., 2020; Kulik et al., 2021; Feldman et al., 2023; Tang et al., 2021; Balkanski et al., 2021). However, the analyses used in these works can fail for non-monotone objectives because negative marginal gains and early blocking break the telescoping and averaging arguments.

Our PF+RL growth template theorem resolves these failure modes by excluding elements with negative marginal gain and supplies a lifted monotonicity ratio for the residual objective.

We therefore instantiate the template with five PF-respecting greedy algorithms with a seed set Y: Positive Modified Greedy (PMG), Positive Greedy+Max (PG+Max), Two Set Enumeration Positive Greedy (2EPG), One Set Enumeration Positive Greedy+Max (1EPG+Max), and Sample Greedy (SG). For each procedure, the PF+RL growth template theorem yields a m-dependent approximation ratio as a short corollary of Theorem 3.1, which recovers the classical monotone constants at $m = 1$ and remains informative for $m < 1$.

3.2 Positive Modified Greedy (PMG)

Algorithm 1 Positive Greedy

```
procedure PositiveGreedy($V, X$)
    $S \leftarrow X$, $\bar{V} \leftarrow V$;
    while $\bar{V} \neq \emptyset$ do
        $u \leftarrow \arg\max_{v \in \bar{V}} f(v \mid S)/c(v)$;
        if $f(e \mid S) \geq 0$ then
            $S \leftarrow S \cup \{u\}$;
        else
            break;
        $\bar{V} \leftarrow \{v \in \bar{V} \setminus S : c(S) + c(v) \leq b\}$;
    return $S$;
```

Algorithm 2 Positive Modified Greedy

```
$v^* \leftarrow \arg\max_{v \in V} f(v)$;
$S \leftarrow$ PositiveGreedy($V, \emptyset$);
return best among $S$ and $\{v^*\}$;
```

Corollary 3.1.1. *Algorithm 2 is PF-respecting and instantiates Theorem 3.1 with seed* $Y = \emptyset$. *There exists a closed-form* $h_1^\lambda(m)$ *such that for every* $\lambda \geq 1$ *it holds that* $f(S_{\text{out}}) \geq \max\{m(1 - e^{-1})/2, h_1^\lambda(m)\} \cdot f(\text{OPT})$.

At $m = 1$, PMG recovers the approximation ratio from the analyses of the modified greedy algorithm under a knapsack constraint. For more details of $h_1^\lambda(m)$, please refer to Appendix C.2 and Appendix D.

3.3 Positive Greedy+Max (PG+Max)

Corollary 3.1.2. *Algorithm 3 is PF-respecting and satisfies* $f(S_{\text{out}}) \geq (m/2) \cdot f(\text{OPT})$ *as an instantiation of Theorem 3.1 with seed* $Y = \emptyset$.

The approximation ratio $m/2$ interpolates smoothly from 0 at $m = 0$ to $1/2$ at $m = 1$, recovering the approximation ratio under the monotone case. For proof details, please refer to Appendix C.3 and Appendix E.

Algorithm 3 Positive Greedy+Max

```
return PG+Max+Initial($V, \emptyset$)
procedure PG+Max+Initial($V, X$)
    $S \leftarrow X$, $T \leftarrow X$, $\bar{V} \leftarrow V$;
    while $\bar{V} \neq \emptyset$ do
        $v' \leftarrow \arg\max_{v \in \bar{V}} f(v \mid S)$;
        if $f(v' \mid S) < 0$ then
            break ;
        if $f(T) < f(S \cup \{v'\})$ then
            $T \leftarrow S \cup \{v'\}$;
        $u \leftarrow \arg\max_{v \in \bar{V}} f(v \mid S)/c(v)$;
        $S \leftarrow S \cup \{u\}$;
        $\bar{V} \leftarrow \{v \in V \setminus S : c(S) + c(v) \leq b\}$;
    return $T$;
```

3.4 Two Set Enumeration Positive Greedy (2EPG)

Algorithm 4 Two Set Enumeration Positive Greedy

```
$v^* \leftarrow \arg\max_{v \in V} f(v)$;
$S \leftarrow \emptyset$;
for every $U \subseteq V$ with $|U| = 2$ and $c(U) \leq b$ do
    $\bar{V} \leftarrow \{v \in V \setminus U, c(U) + c(v) \leq b\}$;
    $S' \leftarrow$ PositiveGreedy($\bar{V}, U$)
    if $f(S') > f(S)$ then
        $S \leftarrow S'$;
return best among $S$ and $\{v^*\}$
```

Let $Y^\star \in \arg\max\{f(Y) : Y \subseteq \text{OPT}, |Y| = 2\}$ be the best two-element subset of OPT.

Corollary 3.1.3. *Algorithm 4 is PF-respecting and instantiates Theorem 3.1 with seed* $Y = Y^\star$. *There exists a sign-sensitive correction term* $t_2(m, m', \lambda)$ *such that for every* $\lambda \geq 1$,

$$f(S_{\text{out}}) \geq \alpha m' \cdot f(\text{OPT}) + t_2(m, m', \lambda) \cdot f(Y^\star), \quad (2)$$

where $\alpha := 1 - e^{-1}$ *and* $m' = m + \frac{m^2 - m}{\lambda - m}$. *Eliminating the unknown ratio* $f(Y^\star)/f(\text{OPT})$ *in the worst case yields a closed-form* $h_2^\lambda(m) \in [0, \alpha]$ *such that for every* $\lambda \geq 1$, *it holds that* $f(S_{\text{out}}) \geq \max\{h_2^\lambda(m), m\alpha/2, h_1^\lambda(m)\} \cdot f(\text{OPT})$.

At $m = 1$, the sign-sensitive correction term t_2 is nonnegative and (2) reduces to the classical two-set enumeration guarantee $1 - 1/e$ (Feldman et al., 2023). For more details of $h_2^\lambda(m)$, please refer to Appendix C.4 and Appendix F.

Algorithm 5 One Set Enumeration Positive Greedy+Max

```
$S \leftarrow \emptyset$;
for every $U \subseteq V$ with $|U| = 1$ do
    $\bar{V} \leftarrow \{v \in V \setminus U : c(U) + c(v) \leq b\}$;
    $S' \leftarrow$ PG+Max+Initial$(\bar{V}, U)$
    if $f(S') > f(S)$ then
        $S \leftarrow S'$;
return $S$;
```

3.5 One Set Enumeration Positive Greedy+Max (1EPG+Max)

Let $w^\star \in \arg\max\{f(\{w\}) : w \in \text{OPT}\}$ be the best singleton solution within OPT.

Corollary 3.1.4. *Algorithm 5 is PF-respecting and instantiates Theorem 3.1 with seed* $Y = \{w^\star\}$ *such that for every* $\lambda \geq 1$,

$$f(S_{\text{out}}) \geq \frac{m'}{2} \cdot f(\text{OPT}) + \frac{mm'}{2\lambda} \cdot f(\{w^\star\}). \quad (3)$$

Worst-case elimination of $f(\{w^\star\})/f(\text{OPT})$ *yields a closed-form* $h_3^\lambda(m)$ *such that for every* $\lambda \geq 1$, *it holds that* $f(S_{\text{out}}) \geq \max\{h_3^\lambda(m), m/2\} \cdot f(\text{OPT})$.

At $m = 1$, the seeding dividend is non-negative and (3) reduces to the standard one-set enumeration greedy+max guarantee. For more details of $h_3^\lambda(m)$, please refer to Appendix C.5 and Appendix G.

By far, the approximation ratios go to 0 as the monotonicity ratio m goes to 0, which is a weakness of these theoretical results. Next, we consider a randomized algorithm which has a non-trivial approximation ratio when the monotonicity ratio m goes to 0.

3.6 Sample Greedy (SG)

Corollary 3.1.5. *Algorithm 6 is PF-respecting and satisfies for every* $\lambda \geq 1$, *it holds that* $f(S_{\text{out}}) \geq \max\{h_4(m), h_1^\lambda(m)\} \cdot f(\text{OPT})$, *as an instantiation Theorem 3.1 with* $Y = \emptyset$ *where* $h_1^\lambda(m)$ *is defined in Corollary 3.1.1.*

Noticeably, SG stays non-trivial as the monotonicity ratio m goes to 0 and its approximation ratio is larger than those of other deterministic algorithms when the function is far from monotone. As m increases, deterministic PF-respecting algorithms overtake it. For more details of $h_4^\lambda(m)$, please refer to Appendix C.6 and Appendix H.

Algorithm 6 Sample Greedy

```
Input: $\delta \in (0, 1/5)$
$T \leftarrow \lfloor 1/(5\delta) \rfloor$
$S_1 \leftarrow$ SampleGreedyWithP$(1/2)$
$S_2 \leftarrow$ SampleGreedyWithP$(1)$
for $i = 0$ to $T$ do
    $\tilde{m} \leftarrow i\delta$
    $p \leftarrow \frac{-\tilde{m} - \sqrt{(\tilde{m}-2)(\tilde{m}-1)} + 1}{\tilde{m}-1}$
    $S_i \leftarrow$ SampleGreedyWithP$(p)$
return best among $S_1$, $S_2$ and $S_i$ for $0 \leq i \leq T$
procedure SampleGreedyWithP$(p)$
    $v^* \leftarrow \arg\max_{v \in V} f(v)$;
    $S \leftarrow \emptyset$, $\bar{V} \leftarrow V$;
    while $\bar{V} \neq \emptyset$ do
        $u \leftarrow \arg\max_{v \in \bar{V}} f(v \mid S)/c(v)$;
        if $f(u \mid S) < 0$ then
            break;
        $r \sim$ Bernoulli$(p)$;
        if $r = 1$ then
            $S \leftarrow S \cup \{u\}$;
        $\bar{V} \leftarrow \bar{V} \setminus \{u\}$
        $\bar{V} \leftarrow \{v \in \bar{V} \setminus S : c(S) + c(v) \leq b\}$;
    return best among $S$ and $\{v^*\}$;
```

3.7 Inapproximability Results

Theorem 3.2. *For any constant* $\varepsilon > 0$, *no polynomial time algorithm can obtain an approximation ratio of* $\rho_{\text{card}}(m) - \frac{1-m}{512} + \varepsilon$ *for the problem of maximizing a non-negative* m*-monotone submodular function with a knapsack constraint where* $\rho_{\text{card}}(m)$ *is the inapproximability result for the cardinality case from (Mualem and Feldman, 2022).*

Proof sketch. Starting from the objective f from §4.3 of (Mualem and Feldman, 2022), we construct two distributions $\mathcal{D}_0$ and $\mathcal{D}_1$ over submodular maximization with a knapsack constraint problem instances that share the same partition of ground set and multiset of item costs but differ by a hidden alignment between costs and a planted block of types over partition. Then, it follows that $\mathbb{E}_{\mathcal{D}_0}[f(\text{OPT})] \leq \rho_{\text{card}}(m)\mathbb{E}_{\mathcal{D}_1}[f(\text{OPT})] + o(1)$.

Let $\Phi(x) := \mathbb{E}[f(S) \mid x]$ given a type counts vector $x = (x_1, \cdots, x_T)$ with $x_t := |S \cap V_t|$ where $V_t (1 \leq t \leq T)$ is a partition of the ground set V. The value of Φ over two distributions $\mathcal{D}_0$ and $\mathcal{D}_1$ differs at most $(1-m)/512$ relative to expected optimal value. Since the total variation between two distributions $\mathcal{D}_0$ and $\mathcal{D}_1$ are infinitesimal TV $= o(1)$, it follows that any algorithm achieving an approximation ratio of no less than $\rho_{\text{card}}(m) - \frac{1-m}{512} + \varepsilon$ on both distributions would distinguish the two distributions with constant advantage, which is a contradiction via the information-theoretic indistinguishability framework derived from (Le Cam, 1986; Yao, 1977). Therefore, no polynomial-time algorithm can obtain an approximation ratio of $\rho_{\text{card}}(m) - \frac{1-m}{512} + \varepsilon$ for any constant $\varepsilon > 0$. The full proof is given in Appendix I. $\square$

4 EXPERIMENTS

We evaluate the five PF-respecting algorithms across two distinct applications: movie recommendation (Mirzasoleiman et al., 2016) and influence-and-exploit marketing (Hartline et al., 2008).

We compare our algorithms against two state-of-the-art algorithms from (Pham et al., 2023), including the Deterministic Linear Query algorithm, which achieves an approximation ratio of 1/6, and the Randomized Linear Query algorithm, which achieves an approximation ratio of 1/4 for non-monotone submodular maximization with a knapsack constraint.

Our experiments are conducted on a 36-core Linux server equipped with an Intel Core i9-10980XE CPU @ 3.00GHz and 125GB of RAM. In the implementation, we apply the lazy update technique (Minoux, 1978) to accelerate the greedy-based algorithms.

Since the optimal function value is generally inaccessible in submodular maximization, we compute a data-dependent upper bound on the optimal value for each problem instance. We define the ratio between the output function value and this upper bound as the *empirical approximation ratio*. Specifically, we sort the elements in ground set V by density $\rho(e_i)$ in decreasing order as $\{e_1, e_2, \ldots, e_n\}$ where $\rho(e_i) \geq \rho(e_{i+1})$ for $1 \leq i \leq n-1$. The data-dependent upper bound is defined as $\Lambda := \sum_{i=1}^{r-1} f(e_i) + \big(b - \sum_{i=1}^{r-1} c(e_i)\big)\rho(e_r)$, where r is the smallest index such that $\sum_{i=1}^{r} c(e_i) \geq b$. It is easy to verify that $\Lambda \geq f(\text{OPT})$ by submodularity. Hence, $f(S)/\Lambda$ gives an empirical approximation ratio (Tang et al., 2021).

4.1 Movie Recommendation

To effectively promote movies in a movie recommendation system, the system must generate a recommendation list for each user, aiming to maximize user satisfaction within the user's budget for purchasing movies. Following the setting in (Mirzasoleiman et al., 2016), we consider a movie recommendation system with a ground set V consisting of n movies. Each movie $u \in V$ is represented by a user-rating vector $\mathbf{x}_u \in \mathbb{R}^d$, where d is the total number of active users in the system. The similarity score $s_{u,v}$ between any two movies u and v is computed based on their user-rating vectors using methods such as cosine similarity or Euclidean distance. To maximize revenue, the price set by the content provider is proportional to the movie's popularity. Consequently, the cost (price) of each movie is defined as the norm of its rating vector, i.e., $c(u) := |\mathbf{x}_u|$. In the real-world scenario, each user has an expenditure budget b. Therefore, the recommendation list $S \subseteq V$ generated by the movie recommendation system must not exceed the user's budget b, i.e., $c(S) \leq b$. Given the similarity scores for all movie pairs, to measure the quality of a recommendation list S, we use the following objective function: $f(S) = \sum_{u \in V} \sum_{v \in S} s_{u,v} - \beta \sum_{u \in S} \sum_{v \in S} s_{u,v}$ where $\beta \in [0, 1]$ is a parameter. The first term measures the coverage of the recommendation list, while the second term penalizes the similarity within the recommendation list, with the penalty controlled by the parameter β. The goal of the movie recommendation task is to generate a recommendation list S that maximizes the objective function $f(S)$ with the knapsack constraint defined by the user's budget b, i.e., $c(S) \leq b$.

Lemma 4.1. *(Mualem and Feldman, 2022) The set function f is non-negative and submodular. If $\beta \in [0, 1/2]$, f is monotone; otherwise, it has a monotonicity ratio of $2(1-\beta)$.*

In our experiments, we use the MovieLens dataset (Harper and Konstan, 2016), which includes 10,000 movies, to construct the cosine similarity score between movie pairs.

4.2 Influence-and-Exploit Marketing

With the rise of social media, businesses have found numerous opportunities within social networks, particularly in product advertising. Influence maximization is one of the most common applications of submodular maximization on social networks (Li et al., 2018; Chen et al., 2009). Recently, there has been significant interest in monetizing social networks by assigning value (potential revenue) to social media users (Tanase, 2024). Within a social media network, a purchase by one user influences others in the network to make similar purchases. For sellers on social media, it can be advantageous in the long term to offer product trials to influential users (online celebrities), as these individuals can spread information about the product to a wider audience, potentially increasing revenue. This strategy is known as influence-and-exploit marketing.

We use the setting from (Hartline et al., 2008), which considers a social network with a seller and multiple potential buyers, denoted by the ground set V. For any pair of buyers i and j in the network, the weight w_{ij} on the undirected edge (i, j) quantifies the influence between them. A buyer's decision to purchase an item is influenced by the set of other buyers who already own it. Specifically, the potential revenue generated by buyer i depends on the extent to which he has been influenced by his neighbors who own the prod-

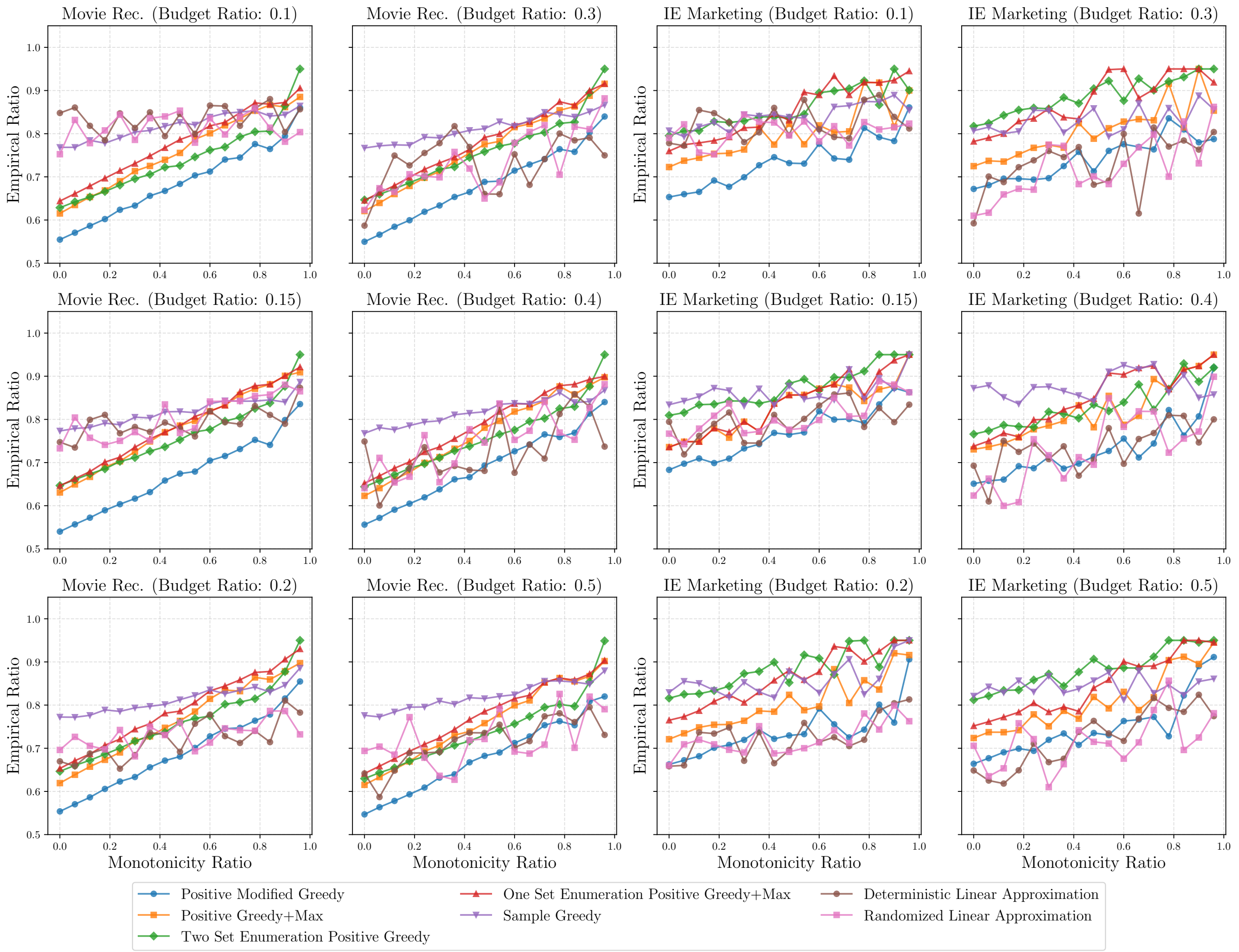

Figure 2: The evaluation results for movie recommendation (Movie Rec.) and influence-and-exploit marketing (IE Marketing) with respect to monotonicity ratios under six different budget ratios.

uct and disseminate information about it across the network. Let $v_i(S)$ represent the value (potential revenue) of buyer i with respect to a selected set S of influential users.

Under the Concave Graph Model, each user $i \in V$ is associated with a non-negative, monotone, concave function f_i. The value of buyer i is then defined as $v_i(S) := f_i(\sum_{j \in S} w_{ij})$ for $i \notin S$ where $w_{ij} = 0$ if there is no edge (i, j) in the network.

Additionally, there is an advertising cost c for each user $i \in V$, and the total cost of advertising to the selected influential users S must not exceed the budget b of the seller, i.e., $c(S) \leq b$. Given a set of influential users S, the total revenue function is the aggregate value of all buyers except those in S, expressed as $f(S) := \sum_{i \in V \setminus S} v_i(S)$. This means that the revenue comes from all buyers except the influential users in S, who receive free products.

In our experiments, we introduce a relaxation on the total revenue function. Suppose the product, such as a high-quality entertainment video, could be offered partially to users. Instead of giving full products to influential users for free trials, the seller could provide partial products with a percentage denoted as β. Then, the total revenue function is defined as $f(S) := \sum_{i \in V} v_i(S) - \beta \sum_{i \in S} v_i(S)$. In the experiments, we set $f_i(x) = x$. The objective of the influence-and-exploit task is to select a set of influential users S by offering them free partial products to maximize the potential revenue $f(S)$ generated from all buyers in the social network, while adhering to the seller's advertisement budget b, i.e., $c(S) \leq b$.

Lemma 4.2. *The objective function f for the influence-and-exploit marketing problem is 1) non-negative $2(1-\beta)$-monotone submodular when $\beta \in [1/2, 1]$; 2) non-negative monotone submodular when $\beta \in [0, 1/2]$.*

Please refer to Appendix J for the proof. In our experiments, the weights are generated on a graph based on top 5,000 communities of Youtube (Yang and Leskovec, 2012) using a uniform distribution in the range $[0, 1]$. The cost of a vertex is the total weight of the edges incident to it.

4.3 Evaluation Results

Define the *budget ratio* b' as the ratio between the budget b of a knapsack constraint and the total cost of elements in ground set V, i.e., $b' := b / \sum_{e \in V} c(e)$. Since the cost scales of the two aforementioned applications are different, we conduct these two experiments with the same budget ratios instead of budget values.

For each application, we construct problem instances with different monotonicity ratios by specifying the parameter β and different budget ratios. The evaluation results are illustrated in Figure 2, where each column corresponds to a distinct budget ratio and each row represents a distinct application. When the monotonicity ratio is relatively low, Sample Greedy, Deterministic Linear Approximation and Randomized Linear Approximation achieve competitive empirical approximation ratio compared to other algorithms in most cases. However, as the monotonicity ratio increases, the performance advantage of these three algorithms gradually diminishes. This trend aligns with our theoretical results, as shown in Figure 1.

5 DISCUSSION

One can combine the analysis of the measured continuous greedy algorithm (MCG) for m-monotone submodular maximization in (Mualem and Feldman, 2022) with the contention-resolution rounding schemes (CRS) in (Chekuri et al., 2014). Specifically, MCG produces a fractional solution $\vec{x}$ over a down-closed polytope $P \subseteq [0, 1]^V$ whose value of multilinear extension $F(\cdot)$ of the submodular function f satisfies $F(\vec{x}) \geq \alpha(m) f(\text{OPT})$ where $\alpha(m) := m(1 - e^{-1}) + (1 - m)e^{-1}$. Since the polytope induced by a knapsack constraint is down-closed, this result could be applied to the m-monotone submodular maximization with a knapsack constraint problem. Using a $(1 - \varepsilon, 1 - \varepsilon)$-balanced CRS for constant-dimensional knapsack polytope according to (Chekuri et al., 2014), we can round the fractional solution $\vec{x}$ into a discrete solution S such that $f(S) \geq (1 - O(\varepsilon))\alpha(m) f(\text{OPT})$. Hence, we obtain an approximation ratio arbitrarily close to $\alpha(m)$ for the m-monotone submodular maximization with a knapsack constraint problem. We call the combination of the MCG algorithm and the $(1 - \varepsilon, 1 - \varepsilon)$-balanced CRS as the MCG+CRS algorithm. The approximation ratio of the MCG+CRS algorithm is stronger than the approximation ratio curves of the PF-respecting greedy algorithms in Figure 1. However, this does not make our work trivial. In this paper, we focus on simple combinatorial algorithms with relatively low query complexity instead of continuous methods. The MCG+CRS algorithm has much higher query complexity than the PF-respecting algorithms due to (i) multilinear-extension evaluation (often via sampling) during the continuous phase, and (ii) knapsack rounding procedures that may involve partial enumeration of large items. In contrast, our PF+RL growth template theorem yields PF-respecting greedy algorithms with at most $O(n^4)$ query complexity, and standard thresholding techniques can further accelerate them to at most $O(n^3)$ query complexity. Empirically, these PF-respecting greedy algorithms are easy to implement and often competitive in partially monotone submodular maximization with a knapsack constraint problem instances, even though their worst-case approximation ratios are weaker than $\alpha(m)$. Hence, continuous algorithms are usually less desirable in practice despite their strong approximation ratios and the PF-respecting algorithms are more practical in real-world applications.

Recent work develops deterministic frameworks for extended multilinear optimization that can substantially reduce or eliminate Monte Carlo evaluation overhead for certain constraint families like matroids (Buchbinder and Feldman, 2025). Incorporating these deterministic frameworks could potentially lower the query complexity of the MCG+CRS algorithm. However, these deterministic frameworks are not directly tailored to the knapsack polytopes. Adapting these deterministic frameworks to the MCG+CRS algorithm for m-monotone submodular maximization with a knapsack constraint is still technically nontrivial. Since this method is beyond the scope of this paper, we view it as a promising future direction to design approximate algorithms that simultaneously achieve an approximation ratio arbitrarily close to $\alpha(m)$ with at most $O(n^3)$ query complexity.

6 CONCLUSIONS

In this work, we propose a novel PF+RL growth template theorem to close the gap between monotone and non-monotone submodular maximization with a knapsack constraint by deriving approximation ratios as a function of the monotonicity ratio m. Enforcing positive-marginal filter yields five PF-respecting greedy algorithms with provable m-dependent approximations that extend over prior constant approximation ratios. Experiments on two applications indicate the informativeness of our theoretical results.

Acknowledgements

We are grateful to the anonymous reviewers for their constructive feedback and suggestions that improved this paper. This research is supported by the Ministry of Education, Singapore, under its AcRF Tier 2 (Award MOE-T2EP20122-0007).

References

A. A. Ageev and M. I. Sviridenko. An 0.828–approximation algorithm for the uncapacitated facility location problem. *Discrete Appl. Math.*, 93 (2–3):149–156, jul 1999. ISSN 0166-218X. doi: 10.1016/S0166-218X(99)00103-1. URL `https://doi.org/10.1016/S0166-218X(99)00103-1`.

Georgios Amanatidis, Federico Fusco, Philip Lazos, Stefano Leonardi, and Rebecca Reiffenhauser. Fast adaptive non-monotone submodular maximization subject to a knapsack constraint. *J. Artif. Intell. Res.*, 74:661–690, 2020.

Georgios Amanatidis, Federico Fusco, Philip Lazos, Stefano Leonardi, Alberto Marchetti-Spaccamela, and Rebecca Reiffenhäuser. Submodular maximization subject to a knapsack constraint: Combinatorial algorithms with near-optimal adaptive complexity. In Marina Meila and Tong Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 231–242. PMLR, 18–24 Jul 2021.

Eric Balkanski, Sharon Qian, and Yaron Singer. Instance Specific Approximations for Submodular Maximization. In *Proceedings of the 38th International Conference on Machine Learning*, pages 609–618. PMLR, July 2021. URL `https://proceedings.mlr.press/v139/balkanski21a.html`. ISSN: 2640-3498.

Niv Buchbinder and Moran Feldman. Extending the extension: Deterministic algorithm for non-monotone submodular maximization. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, STOC '25, page 1130–1141, New York, NY, USA, 2025. Association for Computing Machinery. ISBN 9798400715105. doi: 10.1145/3717823.3718120. URL `https://doi.org/10.1145/3717823.3718120`.

Chandra Chekuri, Jan Vondrák, and Rico Zenklusen. Submodular function maximization via the multilinear relaxation and contention resolution schemes. *SIAM Journal on Computing*, 43(6):1831–1879, 2014. doi: 10.1137/110839655. URL `https://doi.org/10.1137/110839655`.

Wei Chen, Yajun Wang, and Siyu Yang. Efficient influence maximization in social networks. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 199–208, 2009.

Gerard Cornuejols, Marshall L. Fisher, and George L. Nemhauser. Exceptional paper—location of bank accounts to optimize float: An analytic study of exact and approximate algorithms. *Manage. Sci.*, 23(8):789–810, apr 1977. ISSN 0025-1909. doi: 10.1287/mnsc.23.8.789. URL `https://doi.org/10.1287/mnsc.23.8.789`.

Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, 2 edition, 2006.

Shuang Cui, Kai Han, Jing Tang, He Huang, Xueying Li, and Aakas Zhiyuli. Practical parallel algorithms for submodular maximization subject to a knapsack constraint with nearly optimal adaptivity. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(6):7261–7269, Jun. 2023. doi: 10.1609/aaai.v37i6.25885. URL `https://ojs.aaai.org/index.php/AAAI/article/view/25885`.

Marcia Fampa and Jon Lee. An outer-approximation algorithm for maximum-entropy sampling. In Ivana Ljubić, Francisco Barahona, Santanu S. Dey, and A. Ridha Mahjoub, editors, *Combinatorial Optimization*, pages 130–142, Cham, 2022. Springer International Publishing. ISBN 978-3-031-18530-4.

Uriel Feige. A threshold of ln n for approximating set cover. *J. ACM*, 45(4):634–652, jul 1998. ISSN 0004-5411. doi: 10.1145/285055.285059. URL `https://doi.org/10.1145/285055.285059`.

Moran Feldman, Zeev Nutov, and Elad Shoham. Practical budgeted submodular maximization. *Algorithmica*, 85(5):1332–1371, May 2023. ISSN 1432-0541. doi: 10.1007/s00453-022-01071-2. URL `https://doi.org/10.1007/s00453-022-01071-2`.

F. Maxwell Harper and Joseph A. Konstan. The MovieLens Datasets: History and Context. *ACM Transactions on Interactive Intelligent Systems*, 5 (4):1–19, January 2016. ISSN 2160-6455, 2160-6463. doi: 10.1145/2827872. URL `https://dl.acm.org/doi/10.1145/2827872`.

Jason Hartline, Vahab Mirrokni, and Mukund Sundararajan. Optimal marketing strategies over social networks. In *Proceedings of the 17th International Conference on World Wide Web*, WWW '08, page 189–198, New York, NY, USA, 2008. Association for Computing Machinery. ISBN 9781605580852. doi: 10.1145/1367497.1367524. URL `https://doi.org/10.1145/1367497.1367524`.

Samir Khuller, Anna Moss, and Joseph Naor. The budgeted maximum coverage problem. *Information*

Processing Letters, 70(1):39–45, 1999. ISSN 0020-0190. doi: https://doi.org/10.1016/S0020-0190(99)00031-9. URL `https://www.sciencedirect.com/science/article/pii/S0020019099000319`.

Ariel Kulik, Roy Schwartz, and Hadas Shachnai. A refined analysis of submodular greedy. *Operations Research Letters*, 49(4):507–514, 2021. ISSN 0167-6377. doi: https://doi.org/10.1016/j.orl.2021.04.006. URL `https://www.sciencedirect.com/science/article/pii/S0167637721000584`.

Lucien M. Le Cam. *Asymptotic Methods in Statistical Decision Theory.* Springer-Verlag, 1986.

Siyu Lei, Silviu Maniu, Luyi Mo, Reynold Cheng, and Pierre Senellart. Online influence maximization. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, pages 645–654, 2015.

Yuchen Li, Ju Fan, Yanhao Wang, and Kian-Lee Tan. Influence maximization on social graphs: A survey. *IEEE Transactions on Knowledge and Data Engineering*, 30(10):1852–1872, 2018.

Michel Minoux. Accelerated greedy algorithms for maximizing submodular set functions. In J. Stoer, editor, *Optimization Techniques*, volume 7, pages 234–243. Springer-Verlag, Berlin/Heidelberg, 1978. ISBN 978-3-540-08708-3. doi: 10.1007/BFb0006528. URL `http://link.springer.com/10.1007/BFb0006528`. Series Title: Lecture Notes in Control and Information Sciences.

Baharan Mirzasoleiman, Ashwinkumar Badanidiyuru, and Amin Karbasi. Fast constrained submodular maximization: Personalized data summarization. In Maria Florina Balcan and Kilian Q. Weinberger, editors, *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 1358–1367, New York, New York, USA, 20–22 Jun 2016. PMLR. URL `https://proceedings.mlr.press/v48/mirzasoleiman16.html`.

Loay Mualem and Moran Feldman. Using partial monotonicity in submodular maximization. In *Proceedings of the 36th International Conference on Neural Information Processing Systems*, NIPS '22, Red Hook, NY, USA, 2022. Curran Associates Inc. ISBN 9781713871088.

Canh V. Pham, Tan D. Tran, Dung T. K. Ha, and My T. Thai. Linear query approximation algorithms for non-monotone submodular maximization under knapsack constraint. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, IJCAI '23, Macao, P.R.China, 2023. ISBN 978-1-956792-03-4. doi: 10.24963/ijcai.2023/459. URL `https://doi.org/10.24963/ijcai.2023/459`.

Andreas S. Schulz and Nelson A. Uhan. Encouraging cooperation in sharing supermodular costs. In Moses Charikar, Klaus Jansen, Omer Reingold, and José D. P. Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 271–285, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg. ISBN 978-3-540-74208-1.

Michael C Shewry and Henry P Wynn. Maximum entropy sampling. *Journal of applied statistics*, 14(2):165–170, 1987.

Maxim Sviridenko. A note on maximizing a submodular set function subject to a knapsack constraint. *Operations Research Letters*, 32(1):41–43, January 2004. ISSN 01676377. doi: 10.1016/S0167-6377(03)00062-2. URL `https://linkinghub.elsevier.com/retrieve/pii/S0167637703000622`.

George Cosmin Tanase. Personal brand management through social networks in the virtual age of the 21st century individualism. *Romanian Distribution Committee Magazine*, 15(2):28–33, 2024.

Jing Tang, Xueyan Tang, Andrew Lim, Kai Han, Chongshou Li, and Junsong Yuan. Revisiting Modified Greedy Algorithm for Monotone Submodular Maximization with a Knapsack Constraint. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 5(1):1–22, February 2021. ISSN 2476-1249. doi: 10.1145/3447386. URL `https://dl.acm.org/doi/10.1145/3447386`.

Jan Vondrák. Symmetry and approximability of submodular maximization problems. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 651–670, 2009. doi: 10.1109/FOCS.2009.24.

L. A. Wolsey. An analysis of the greedy algorithm for the submodular set covering problem. *Combinatorica*, 2(4):385–393, Dec 1982a. ISSN 1439-6912. doi: 10.1007/BF02579435. URL `https://doi.org/10.1007/BF02579435`.

Laurence A. Wolsey. Maximising real-valued submodular functions: Primal and dual heuristics for location problems. *Mathematics of Operations Research*, 7(3):410–425, 1982b. ISSN 0364765X, 15265471. URL `http://www.jstor.org/stable/3689607`.

Jaewon Yang and Jure Leskovec. Defining and evaluating network communities based on groundtruth. In *Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics*, MDS '12, New York, NY, USA, 2012. Association for Computing Machinery. ISBN 9781450315463. doi: 10.

1145/2350190.2350193. URL `https://doi.org/10.1145/2350190.2350193`.

Andrew Chi-Chin Yao. Probabilistic computations: Toward a unified measure of complexity. In *18th Annual Symposium on Foundations of Computer Science (sfcs 1977)*, pages 222–227, 1977. doi: 10.1109/SFCS.1977.24.

Grigory Yaroslavtsev, Samson Zhou, and Dmitrii Avdiukhin. "bring your own greedy"+max: Near-optimal 1/2-approximations for submodular knapsack. In Silvia Chiappa and Roberto Calandra, editors, *The 23rd International Conference on Artificial Intelligence and Statistics, AISTATS 2020, 26-28 August 2020, Online [Palermo, Sicily, Italy]*, volume 108 of *Proceedings of Machine Learning Research*, pages 3263–3274. PMLR, 2020. URL `http://proceedings.mlr.press/v108/yaroslavtsev20a.html`.

Checklist

1. For all models and algorithms presented, check if you include:
   (a) A clear description of the mathematical setting, assumptions, algorithm, and/or model. [Yes]
   (b) An analysis of the properties and complexity (time, space, sample size) of any algorithm. [Yes]
   (c) (Optional) Anonymized source code, with specification of all dependencies, including external libraries. [Yes]

2. For any theoretical claim, check if you include:
   (a) Statements of the full set of assumptions of all theoretical results. [Yes]
   (b) Complete proofs of all theoretical results. [Yes]
   (c) Clear explanations of any assumptions. [Yes]

3. For all figures and tables that present empirical results, check if you include:
   (a) The code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL). [Yes]
   (b) All the training details (e.g., data splits, hyperparameters, how they were chosen). [Not Applicable]
   (c) A clear definition of the specific measure or statistics and error bars (e.g., with respect to the random seed after running experiments multiple times). [Yes]
   (d) A description of the computing infrastructure used. (e.g., type of GPUs, internal cluster, or cloud provider). [Yes]

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets, check if you include:
   (a) Citations of the creator If your work uses existing assets. [Yes]
   (b) The license information of the assets, if applicable. [Not Applicable]
   (c) New assets either in the supplemental material or as a URL, if applicable. [Not Applicable]
   (d) Information about consent from data providers/curators. [Not Applicable]
   (e) Discussion of sensible content if applicable, e.g., personally identifiable information or offensive content. [Not Applicable]

5. If you used crowdsourcing or conducted research with human subjects, check if you include:
   (a) The full text of instructions given to participants and screenshots. [Not Applicable]
   (b) Descriptions of potential participant risks, with links to Institutional Review Board (IRB) approvals if applicable. [Not]Applicable
   (c) The estimated hourly wage paid to participants and the total amount spent on participant compensation. [Not Applicable]

Supplementary Materials

A Discussions on Technical Distinctions

Our work extends the applicability of five well-known algorithms from the classical setting of monotone submodular maximization with a knapsack constraint to a broader setting in which the objective function exhibits partial monotonicity by enforcing positive marginal gain. Although our analytical framework retains the overall structure of prior work, the lack of a monotonicity assumption introduces new difficulties in the analysis, causing portions of the original arguments to break down. To overcome these difficulties, we substitute the monotonicity-based inequality with monotonicity-ratio-based inequality, discuss different cases determined by the value of monotonicity ratio and analyze the behavior of selected elements with negative marginal gains. Depending on the context, we either obtain generalized versions of prior results or develop entirely new arguments. Additionally, we introduce a dummy element to guarantee the budget violation condition always holds (this condition holds under monotone case but may not hold under non-monotone case) and leverage case-based reasoning to handle complicated scenarios caused by different values of monotonicity ratio. These techniques collectively form a unified analytical framework for submodular maximization with a knapsack constraint under partial monotonicity.

B Useful Lemmas

In the following, we derive the monotonicity ratio of a special set function. If $f(S)$ is a non-negative monotone submodular function, its residual function $g_Y(S) := f(Y \cup S) - f(Y)$ given a base set $Y \subseteq V$ is also non-negative, monotone and submodular. However, if f is m-monotone where $m < 1$, we cannot derive a meaningful closed-form expression for the monotonicity ratio by directly incorporating the monotonicity ratio m into the residual function $g_Y(S) := f(Y \cup S) - mf(Y)$. To address this issue, we introduce a hyper-parameter $\lambda \geq 1$ (whose value will be determined later) and focus on the residual function $g_Y(S) := f(Y \cup S) - \frac{m}{\lambda} f(Y)$ depending on hyper-parameter λ.

Lemma B.1. *Let f be a non-negative, m-monotone submodular function and let $\lambda \geq 1$ be a hyper-parameter (to be determined later). For any set $Y \subseteq V$, define the residual function $g_Y(S) := f(S \cup Y) - \frac{m}{\lambda} f(Y)$. Then g_Y is non-negative, m'-monotone and submodular where*

$$m' := \begin{cases} 1, & \text{if } \lambda = m = 1, \\ m + \dfrac{m^2 - m}{\lambda - m}, & \text{otherwise.} \end{cases}$$

Proof. If $\lambda = m = 1$, then $g_Y(S) = f(S \cup Y) - f(Y)$. Since f is monotone ($m = 1$), it follows that $g_Y(S) \geq 0$ and g_Y is also monotone. Submodularity of f implies that g_Y is submodular.

In the remaining proof, assume $\lambda \neq m$ and $m \neq 1$.

Since function f is m-monotone, function $g_Y(S)$ is non-negative since $f(S \cup Y) \geq mf(Y) \geq \frac{m}{\lambda} f(Y)$ for any subsets $S, Y \subseteq V$.

Given any two subsets $S \subseteq T \subseteq V$, let $s := \frac{f(T \cup Y)}{f(S \cup Y)}$ and $t := \frac{f(S \cup Y)}{f(Y)}$, it follows that

$$\frac{g_Y(T)}{g_Y(S)} = \frac{f(T \cup Y) - \frac{m}{\lambda} f(Y)}{f(S \cup Y) - \frac{m}{\lambda} f(Y)} = \frac{stf(Y) - \frac{m}{\lambda} f(Y)}{tf(Y) - \frac{m}{\lambda} f(Y)} = \frac{st - \frac{m}{\lambda}}{t - \frac{m}{\lambda}}.$$

By the definition of monotonicity ratio, it holds that $s \geq m$, $t \geq m$ and $st \geq m$.

Case 1: If $t \geq 1$, since $s \geq m$, it holds that

$$\frac{g_Y(T)}{g_Y(S)} = \frac{st - \frac{m}{\lambda}}{t - \frac{m}{\lambda}} \geq \frac{mt - \frac{m}{\lambda}}{t - \frac{m}{\lambda}} \geq \frac{m - \frac{m}{\lambda}}{1 - \frac{m}{\lambda}}.$$

Case 2: If $t \leq 1$, it holds that

$$\frac{g_Y(T)}{g_Y(S)} = \frac{st - \frac{m}{\lambda}}{t - \frac{m}{\lambda}} \geq \frac{m - \frac{m}{\lambda}}{t - \frac{m}{\lambda}} \geq \frac{m - \frac{m}{\lambda}}{1 - \frac{m}{\lambda}}.$$

Define $m' := \frac{m - \frac{m}{\lambda}}{1 - \frac{m}{\lambda}} = m + \frac{m^2 - m}{\lambda - m}$. Combining the results of Case 1 and Case 2, we conclude that for any two subsets $S \subseteq T \subseteq V$, it holds that $g_Y(T) \geq m' g_Y(S)$.

Given any two subsets $S \subseteq T \subseteq V$ and element $x \in V \setminus T$, it follows that

$$\begin{aligned} g_Y(x \mid S) - g_Y(x \mid T) &= g_Y(S \cup \{x\}) - g_Y(S) - g_Y(T \cup \{x\}) + g_Y(T) \\ &= f(S \cup \{x\} \cup Y) - f(S \cup Y) - f(T \cup \{x\} \cup Y) + f(T \cup Y) \\ &= f(x \mid S \cup Y) - f(x \mid T \cup Y) \geq 0 \end{aligned}$$

where the last inequality holds due to the submodularity of function f. Hence, g_Y is submodular. □

Lemma B.2. *(Corollary 2.2 in (Mualem and Feldman, 2022)) Let f be a non-negative m-monotone submodular function. For any deterministic set $O \subseteq V$ and random set $D \subseteq V$, it holds that*

$$\mathbb{E}[f(O \cup D)] \geq \left(1 - (1 - m) \max_{u \in V} \mathbb{P}[u \in D]\right) f(O).$$

C Details of Theoretical Results

For simplicity, we denote $\alpha := 1 - e^{-1}$, $m' := m + \dfrac{m^2 - m}{\lambda - m}$ and use the hyper-parameter $\lambda \geq 1$ defined in Lemma B.1 in our analysis. λ trades off inherited monotonicity from the base set vs. slack to avoid over-penalizing densities. m' increases with λ and equals 0 when $\lambda = 1$.

C.1 Proof of Theorem 3.1

Proof. For the fixed seed Y and parameter $\lambda \geq 1$, consider the residual $g(S) := g_Y(S) = f(S \cup Y) - \frac{m}{\lambda} f(Y)$ where $S \subseteq V \setminus Y$. By Lemma B.1, g is nonnegative, submodular and m'-monotone with

$$m' = \begin{cases} 1, & \lambda = m = 1, \\ m + \frac{m^2 - m}{\lambda - m}, & \text{otherwise}, \end{cases}$$

and m' is non-decreasing in m for each fixed λ. In particular, since g is monotone it follows that any PF-respecting growth rule can be analyzed on g using standard monotone knapsack arguments.

Let OPT_g denote an optimal solution for the residual problem instance $(V \setminus Y, c, b - c(Y))$ with objective g. Let $S^{\mathcal{R}}_{\text{out}}$ denote some candidate solution among some greedy algorithm $\mathcal{R}$ (e.g., positive modified greedy, positive greedy+max, two set enumeration positive greedy, one set enumeration positive greedy+max and sample greedy) by setting objective function as g and initial solution as Y . Since g is m'-monotone and submodular, there exists some constants $\alpha_{\mathcal{R}} \geq 0$ and $\beta_{\mathcal{R}} \geq 0$ such that

$$g\big(S^{\mathcal{R}}_{\text{out}}\big) \geq \alpha_{\mathcal{R}} \cdot g(\text{OPT}_g) + \beta_{\mathcal{R}} \cdot g(Y). \tag{4}$$

For concrete theoretical results in Appendix C, we have (i) $\alpha_{\mathcal{R}} = \alpha/2$ with $\alpha = 1 - 1/e$ and $\beta_{\mathcal{R}} = 0$ for PMG; (ii) $\alpha_{\mathcal{R}} = 1/2$, $\beta_{\mathcal{R}} = 0$ for PG+Max; (iv) $\alpha_{\mathcal{R}} = 1 - 1/e$, $\beta_{\mathcal{R}} > 0$ for 2EPG; (iii) $\alpha_{\mathcal{R}}$ equals the constant in the analysis of one-set enumeration algorithm for monotone case (Feldman et al., 2023) and $\beta_{\mathcal{R}} > 0$ due to explicit seed comparison for 1EPG+Max. Intuitively, $\alpha_{\mathcal{R}}$ is similar to the approximation ratio derived for monotone submodular maximization with a knapsack constraint and $\beta_{\mathcal{R}}$ captures explicit seed dividends for the enumeration-based algorithms ($|Y| \in \{1,2\}$).

By the definition of g and Inequality (4), it follows that $f\left(S_{\text{out}}^{\mathcal{R}} \cup Y\right) - \frac{m}{\lambda} f(Y) \geq \alpha_{\mathcal{R}} \left(f(\text{OPT}_g \cup Y) - \frac{m}{\lambda} f(Y)\right) + \beta_{\mathcal{R}} \left(f(Y) - \frac{m}{\lambda} f(Y)\right) = \alpha_{\mathcal{R}} f(\text{OPT}_g \cup Y) + (\beta_{\mathcal{R}} - \alpha_{\mathcal{R}}) \left(1 - \frac{m}{\lambda}\right) f(Y)$. Let $B_{\mathcal{R}}(m,\lambda) := (\beta_{\mathcal{R}} - \alpha_{\mathcal{R}}) \left(1 - \frac{m}{\lambda}\right)$. Then we have

$$f\left(S_{\text{out}}^{\mathcal{R}} \cup Y\right) \geq \alpha_{\mathcal{R}} f(\text{OPT}_g \cup Y) + B_{\mathcal{R}}(m,\lambda) f(Y). \tag{5}$$

For $|Y| \in \{0,1,2\}$, there exists at least one subset Y such that $Y \subseteq \text{OPT}$ since 1) $\emptyset \subseteq \text{OPT}$ when $|Y| = 0$; 2) enumerating any subset with size $|Y| \in \{1,2\}$ guarantees the existence. For simplicity, we still use Y to denote the seed subset such that $Y \subseteq \text{OPT}$. Hence, it follows that $f(\text{OPT}_g \cup Y) \geq f(\text{OPT} \cup Y) \geq m f(\text{OPT})$.

Let S_{out} be the final output of some PF-respecting algorithm. Since the PF-respecting algorithm returns the best among its candidates, it holds that

$$f(S_{\text{out}}) \geq f\left(S_{\text{out}}^{\mathcal{R}} \cup Y\right). \tag{6}$$

Combining Inequality (5) and Inequality 6, we obtain

$$f(S_{\text{out}}) \geq f\left(S_{\text{out}}^{\mathcal{R}} \cup Y\right) \geq \underbrace{\alpha_{\mathcal{R}} m}_{=:A(m,\lambda)} f(\text{OPT}) + \underbrace{B_{\mathcal{R}}(m,\lambda)}_{=:B(m,\lambda)} f(Y).$$

In the technical details of Appendix D,E ,F, G, H, the closed form details are given per instantiation. ☐

Leveraging Theorem 3.1, we obtain closed form approximation ratios with respect to monotonicity ratio per instantiation in the remaining part of Appendix C.

C.2 Positive Modified Greedy

Theorem C.1. *Positive Modified Greedy (Algorithm 2) achieves an approximation ratio of* $\max\{m\alpha/2, \max_{\lambda \geq 1}\{h_1^{\lambda}(m)\}\}$ *for maximizing a non-negative m-monotone submodular function with a knapsack constraint where*

$$h_1^{\lambda}(m) := \min\left\{\frac{m(\lambda - 1)}{\lambda + m(\lambda - 1) - m}, \frac{m^2}{\lambda}\left(1 - \frac{1}{\sqrt{e}}\right)\right\}.$$

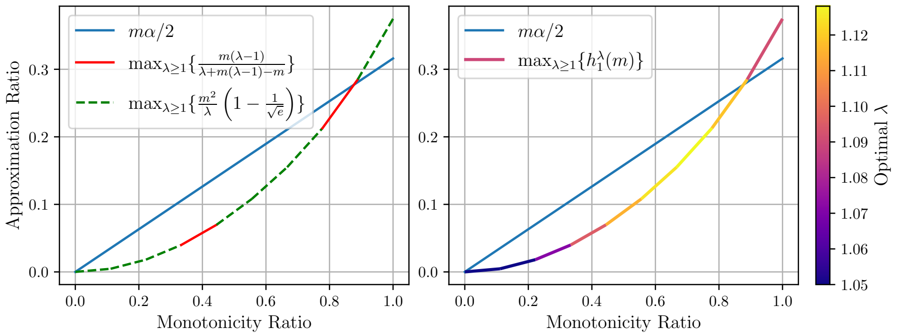

Figure 3: **Comparison of Approximation Ratios and Optimal λ.** This figure compares the approximation ratio $m\alpha/2$ (blue solid line) and $\max_{\lambda\geq 1}\{h_1^{\lambda}(m)\}$ (which is plotted as a segmented line) across different monotonicity ratios (m). The left panel highlights the two components of $\max_{\lambda\geq 1}\{h_1^{\lambda}(m)\}$, while the right panel highlights the optimal λ values.

We illustrate in Figure 3 the approximation ratio with respect to the monotonicity ratio in Theorem C.1. As $h_1^{\lambda}(m)$ is defined as the minimum of two expressions, we use two line styles to distinguish the intervals where each term is selected by the min operator. The left panel depicts $\max_{\lambda\geq 1}\{h_1^{\lambda}(m)\}$, where the red solid line corresponds to the interval where $\frac{m(\lambda-1)}{\lambda+m(\lambda-1)-m}$ is smaller, while the green dashed line corresponds to the interval where $\frac{m^2}{\lambda}\left(1-\frac{1}{\sqrt{e}}\right)$ is smaller. This distinction highlights the necessity to include both expressions within $h_1^\lambda(m)$. The right panel depicts the optimal λ value that maximizes $h_1^\lambda(m)$ for each monotonicity ratio m. Since the optimal λ varies with m, the inclusion of the max operator over λ is important to capture the best approximation ratio across all monotonicity ratios. From both panels, we observe that the two terms: $m\alpha/2$ and $\max_{\lambda\geq 1}\{h_1^\lambda(m)\}$ does not dominate each other. For relatively small values of m, $m\alpha/2$ yields a higher approximation ratio. For relatively large values of m, $\max_{\lambda\geq 1}\{h_1^\lambda(m)\}$ becomes dominant. Therefore, combining these two terms using max operator is necessary for the final approximation ratio of the Positive Modified Greedy algorithm.

C.3 Positive Greedy+Max

Theorem C.2. *Positive Greedy+Max (Algorithm 3) achieves an approximation ratio of $m/2$ for maximizing a non-negative m-monotone submodular function with a knapsack constraint.*

C.4 Two Set Enumeration Positive Greedy

Theorem C.3. *Let* $t_2(m,m',\lambda) := \frac{m}{\lambda} - \frac{3}{2}\alpha + \frac{\alpha m(1-m')}{\lambda}$ *and* $\alpha_0 := \frac{25\alpha^2-4\alpha+4}{24\alpha^2} \approx 1.195$. *Two Set Enumeration Positive Greedy (Algorithm 4) achieves an approximation ratio of* $\max_{\lambda\geq 1}\{h_2^\lambda(m)\}$ *for maximizing a non-negative m-monotone submodular function with a knapsack constraint where*

- *Case 1:*

$$h_2^\lambda(m) := \begin{cases} \alpha m', & m \in [m_1, m_2] \\ \alpha m' + t_2(m,m',\lambda), & \textit{otherwise} \end{cases}$$

$$m_1 := \frac{\lambda\left(5\alpha+2-\sqrt{-24\alpha^2\lambda+25\alpha^2-4\alpha+4}\right)}{4(\alpha\lambda+1)} \textit{ and } m_2 := \frac{\lambda\left(5\alpha+2+\sqrt{-24\alpha^2\lambda+25\alpha^2-4\alpha+4}\right)}{4(\alpha\lambda+1)} \textit{ if } 1 \leq \lambda \leq \alpha_0;$$

- *Case 2:* $h_2^\lambda(m) := \alpha m' + t_2(m,m',\lambda)$ *if* $\lambda > \alpha_0$.

The Two Set Enumeration Positive Greedy algorithm not only runs the greedy procedure starting from all possible initial solution sets of size two (line 4 of Algorithm 4), but also identifies the single element with the maximum function value (line 1 of Algorithm 4). Therefore, its approximation ratio is at least as good as that of the Positive Modified Greedy algorithm. Consequently, the approximation ratio of the Two Set Enumeration Positive Greedy algorithm can be improved to $\max\{\max_{\lambda\geq 1}\{h_2^\lambda(m)\}, m\alpha/2, \max_{\lambda\geq 1}\{h_1^\lambda(m)\}\}$.

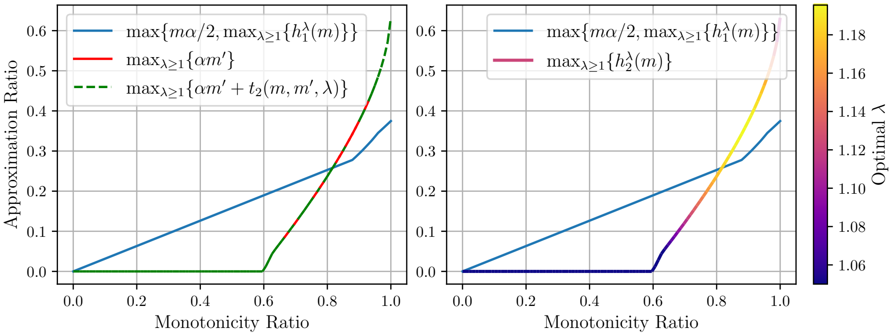

Figure 4: **Comparison of Approximation Ratios and Optimal λ.** This figure compares the approximation ratio $\max\{m\alpha/2, \max_{\lambda\geq 1}\{h_1^\lambda(m)\}\}$ (blue solid line) and $\max_{\lambda\geq 1}\{h_2^\lambda(m)\}$ (which is plotted as a segmented line) across different monotonicity ratios (m). The left panel highlights the two cases of $\max_{\lambda\geq 1}\{h_2^\lambda(m)\}$, while the right panel highlights the optimal λ values.

We illustrate the approximation ratio $\max\{\max_{\lambda\geq 1}\{h_2^\lambda(m)\}, m\alpha/2, \max_{\lambda\geq 1}\{h_1^\lambda(m)\}\}$ in Figure 4. The left panel depicts $\max_{\lambda\geq 1}\{h_2^\lambda(m)\}$, where the red solid line corresponds to the interval where $\max_{\lambda\geq 1}\{h_2^\lambda(m)\} = \max_{\lambda\geq 1}\{\alpha m'\}$ and the green dashed line corresponds to the interval where $\max_{\lambda\geq 1}\{h_2^\lambda(m)\} = \max_{\lambda\geq 1}\{\alpha m' + t_2(m,m',\lambda)\}$. This distinction highlights the necessity to include both components within $h_2^\lambda(m)$. The right panel depicts the optimal λ values that maximizes $h_2^\lambda(m)$ for each monotonicity ratio m. Since the optimal λ varies with m, the inclusion of the max operator over λ is essential to capture the best approximation ratio across all monotonicity ratios. From both panels, we observe that for relatively small values of m, $\max\{m\alpha/2, \max_{\lambda\geq 1}\{h_1^\lambda(m)\}\}$ yields a higher approximation ratio. For relatively large values of m, $\max_{\lambda\geq 1}\{\alpha m'\}$ and $\max_{\lambda\geq 1}\{\alpha m' + t_2(m, m', \lambda)\}$ becomes dominant over $\max\{m\alpha/2, \max_{\lambda\geq 1}\{h_1^\lambda(m)\}\}$ intermittently. Therefore, combining these terms using max operator is necessary for the final approximation ratio of the Two Set Enumeration Positive Greedy algorithm.

C.5 One Set Enumeration Positive Greedy+Max

Theorem C.4. *One Set Enumeration Positive Greedy+Max (Algorithm 5) achieves an approximation ratio of* $\max_{\lambda\geq 1}\{h_3^\lambda(m)\}$ *for maximizing a non-negative m-monotone submodular function with a knapsack constraint where*

$$h_3^\lambda(m) := \min\left\{\frac{m'}{2} + \frac{m(2-m')}{8\lambda}, m'\alpha\right\}.$$

Since the inner loop of the One Set Enumeration Positive Greedy+Max algorithm is exactly the Positive Greedy+Max algorithm, the approximation ratio of One Set Enumeration Positive Greedy+Max is at least as good as that of Positive Greedy+Max. Consequently, the approximation ratio of the One Set Enumeration Positive Greedy+Max algorithm can be enhanced as $\max\{\max_{\lambda\geq 1}\{h_3^\lambda(m)\}, m/2\}$.

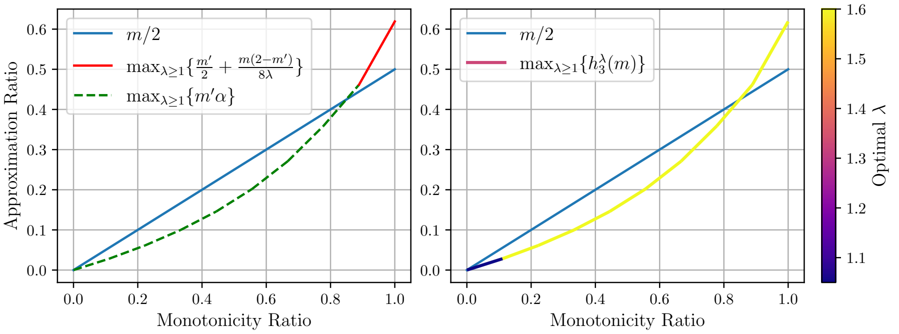

Figure 5: **Comparison of Approximation Ratios and Optimal λ.** This figure compares the approximation ratio $m/2$ (blue solid line) and $\max_{\lambda\geq 1}\{h_3^\lambda(m)\}$ (which is plotted as a segmented line) across different monotonicity ratios (m). The left panel highlights the two components of $\max_{\lambda\geq 1}\{h_3^\lambda(m)\}$, while the right panel highlights the optimal λ values.

We illustrate the approximation ratio $\max\{\max_{\lambda\geq 1}\{h_3^\lambda(m)\}, m/2\}$ in Figure 5. The left panel depicts $\max_{\lambda\geq 1}\{h_3^\lambda(m)\}$, where the red solid line corresponds to the interval where $\max_{\lambda\geq 1}\{h_3^\lambda(m)\} = \max_{\lambda\geq 1}\{\frac{m'}{2} + \frac{m(2-m')}{8\lambda}\}$ and the green dashed line corresponds to the interval where $\max_{\lambda\geq 1}\{h_3^\lambda\ (m)\} = \max_{\lambda\geq 1}\{m'\alpha\}$. This distinction highlights the necessity to include both components within $h_3^\lambda(m)$. The right panel illustrates the optimal λ values that maximizes $h_3^\lambda(m)$ for each monotonicity ratio m. Since the optimal λ varies with m, the inclusion of the max operator over λ is essential to capture the best approximation ratio across all monotonicity ratios. From both panels, we observe that for relatively small values of m, $m/2$ yields a higher approximation ratio. For relatively large values of m, $\max_{\lambda\geq 1}\{h_3^\lambda(m)\}$ becomes dominant over $m/2$. Therefore, combining these two terms using max operator is necessary for the final approximation ratio of the One Set Enumeration Positive Greedy+Max algorithm.

C.6 Sample Greedy

Theorem C.5. *Let* $t_4(m) := \sqrt{(m-2)(m-1)} + m$ *and* $\delta \in (0, 1/5)$ *be a small positive number. Sample Greedy (Algorithm 6) achieves an approximation ratio of* $h_4(m)$ *for maximizing a non-negative m-monotone submodular* *function with a knapsack constraint where*

$$h_4(m) := \begin{cases} (m+1)/6, & m \in [1/5, 1] \\ -\dfrac{(t_4(m)-2)\,(t_4(m)-1)}{t_4(m)-m} - O(\delta), & \textit{otherwise.} \end{cases}$$

When setting $p = 1$, Sample Greedy Wi-th P reduces to Positive Modified Greedy. Hence, the approximation ratio of the Sample Greedy algorithm is at least as good as that of the Positive Modified Greedy algorithm. Consequently the approximation ratio of the Sample Greedy algorithm can be enhanced as $\max\{h_4(m), m\alpha/2, \max_{\lambda\geq 1}\{h_1^\lambda(m)\}\}$. Furthermore, it is easy to verify that $\frac{m+1}{6} > \frac{m\alpha}{2}$ for all $m \in [0, 1]$. Hence, we could further simplify the approximation ratio of Sample Greedy as $\max\{h_4(m), \max_{\lambda\geq 1}\{h_1^\lambda(m)\}\}$.

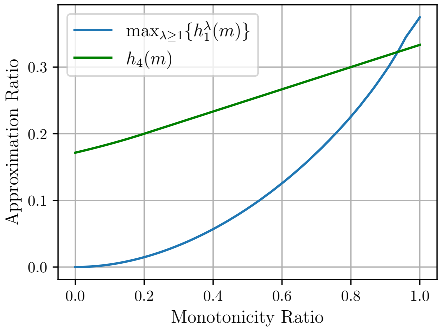

Figure 6: **Comparison of Approximation Ratios.** This figure compares the approximation ratio $\max_{\lambda\geq 1}\{h_1^\lambda(m)\}$ (blue solid line) and $h_4(m)$ (green solid line) across different monotonicity ratios (m).

We illustrate the approximation ratio $\max\{h_4\ (m), \max_{\lambda\geq 1}\{h_1^\lambda(m)\}\}$ in Figure 6. When m is relatively small, $h_4(m)$ is larger than $\max_{\lambda\geq 1}\{h_1^\lambda(m)\}$. When m is relatively large, $\max_{\lambda\geq 1}\{h_1^\lambda(m)\}$ is larger than $h_4(m)$. This indicates the necessity of including these two terms in the final approximation ratio.

D Analysis of Positive Modified Greedy

In this section, we give analysis of approximation ratio for Positive Modified Greedy (Algorithm 2).

Let f be a non-negative m-monotone submodular function with ground set V. Let OPT be the optimal solution set with maximum function value, i.e., $f(\mathrm{OPT}) = \max_{S\subseteq V, c(S)\leq b} f(S)$. Positive Modified Greedy (Algorithm 2) makes one call to Positive Greedy (Algorithm 1) to generate the greedy solution sets $S_i (0 \leq i \leq l)$. Let u_i be the element considered at each step with maximum density and added into the current solution set S_{i-1}, i.e. $S_i = S_{i-1} \cup \{u_i\}$. Let $v^* := \arg\max_{v\in V} f(v)$ and $S_g := \arg\max_{S\in\{S_l, \{v^*\}\}} f(S)$ be the greedy solution set with maximum function value.

Let o_1 be the first element in OPT that is considered by the greedy heuristic but not added to the greedy solution set S_q due to budget violation, i.e., $c(S_q) + c(o_1) > b$ where $1 \leq q \leq l$. If o_1 does not exist, we split into two cases. On the one hand, if $\mathrm{OPT} \subseteq S_l$, it follows that $f(S_g) \geq f(S_l) = f(\mathrm{OPT})$, which is consistent with Theorem C.1. On the other hand, if $\mathrm{OPT} \not\subseteq S_l$, i.e., $\mathrm{OPT} \setminus S_l \neq \emptyset$, it holds that $\forall e \in \mathrm{OPT} \setminus S_l$ s.t. $c(e) + c(S_l) \leq b$ and $f(e \mid S_l) < 0$. By submodularity and definition of monotonicity ratio, it follows that $mf(\mathrm{OPT}) \leq f(\mathrm{OPT} \cup S_l) \leq f(S_l) + \sum_{e\in\mathrm{OPT}\setminus S_l} f(e \mid S_l) < f(S_l) \leq f(S_g)$, which is also consistent with Theorem C.1. Hence, in the remaining part of this section, we assume that o_1 exists.

Define the greedy solution sets $S'_i := S_i (0 \leq i \leq q)$ and $S'_{q+1} := S_i \cup \{o_1\}$ where $u'_i := S_i \setminus S_{i-1} = u_i$ for $1 \leq i \leq q$ and $u'_{q+1} := o_1$. Let $\alpha_i := \frac{b-c(u'_i)}{b-mc(u'_i)}$ for $0 \leq i \leq q+1$.

Lemma D.1. *For* $1 \leq i \leq q+1$*, it holds that*

$$f(S'_i) - \alpha_i f(S'_{i-1}) \geq \frac{mc(u'_i)}{b}[f(\mathrm{OPT}) - \alpha_i f(S'_{i-1})].$$

Proof. According to the definition of m-monotonicity and submodularity, it holds

$$mf(\mathrm{OPT}) - f(S'_{i-1}) \leq f(\mathrm{OPT} \cup S'_{i-1}) - f(S'_{i-1}) \leq \sum_{e \in \mathrm{OPT} \setminus S'_{i-1}} f(e \mid S'_{i-1}) \leq b \cdot \rho(u'_i \mid S'_{i-1}).$$

Since $f(u'_i \mid S'_{i-1}) = f(S'_i) - f(S'_{i-1})$, we have

$$mf(\mathrm{OPT}) - f(S'_{i-1}) \leq b \cdot \frac{f(S'_i) - f(S'_{i-1})}{c(u'_i)}.$$

The result is obtained by rearranging the inequality. $\square$

Lemma D.2. *For $1 \leq i \leq q+1$, it holds that*

$$f(S'_i) \geq \sum_{j=1}^{i} \prod_{k=j+1}^{i} \gamma_k (1-\beta_j) f(\mathrm{OPT})$$

where $\beta_i := 1 - \frac{m \cdot c(u'_i)}{b}$ and $\gamma_i := \alpha_i \cdot \beta_i = 1 - \frac{c(u'_i)}{b}$.

Proof. According to Lemma D.1, it holds

$$\begin{aligned} f(S'_i) &= \alpha_i f(S'_{i-1}) + f(S'_i) - \alpha_i f(S'_{i-1}) \geq \alpha_i f(S'_{i-1}) + \frac{mc(u'_i)}{b}[f(\mathrm{OPT}) - \alpha_i f(S'_{i-1})] \\ &= \alpha_i \left(1 - \frac{mc(u'_i)}{b}\right) f(S'_{i-1}) + \frac{mc(u'_i)}{b} f(\mathrm{OPT}) = \gamma_i \cdot f(S'_{i-1}) + (1-\beta_i) \cdot f(\mathrm{OPT}). \end{aligned}$$

by definitions of β_i and γ_i.

This inequality could be used recursively until $f(S'_1)$, i.e.

$$\begin{aligned} f(S'_i) &\geq \gamma_i \cdot f(S'_{i-1}) + (1-\beta_i) \cdot f(\mathrm{OPT}) \\ f(S'_{i-1}) &\geq \gamma_{i-1} \cdot f(S'_{i-2}) + (1-\beta_{i-1}) \cdot f(\mathrm{OPT}) \\ &\vdots \\ f(S'_1) &\geq \gamma_1 \cdot f(S'_0) + (1-\beta_1) \cdot f(\mathrm{OPT}) \\ &= (1-\beta_1) \cdot f(\mathrm{OPT}). \end{aligned}$$

Combining these inequalities together by multiplying $\prod_{k=j+1}^{i} \gamma_i$ for each $f(S'_j)$ $1 \leq j \leq i-1$, we have

$$\begin{aligned} f(S'_i) &\geq \left[(1-\beta_i) + \gamma_i \cdot (1-\beta_{i-1}) + \gamma_i \gamma_{i-1} \cdot (1-\beta_{i-2}) + \cdots + \prod_{k=2}^{i} \gamma_k (1-\beta_1)\right] \cdot f(\mathrm{OPT}) \\ &= \sum_{j=1}^{i} \prod_{k=j+1}^{i} \gamma_k (1-\beta_j) f(\mathrm{OPT}). \end{aligned}$$

$\square$

Theorem D.3. *Positive Modified Greedy (Algorithm 2) achieves an approximation factor of $m(1-e^{-1})/2$ for maximizing a non-negative m-monotone submodular function with a knapsack constraint.*

Proof. Using Lemma D.2 by setting $i = q+1$, we have

$$f(S'_{q+1}) \geq \sum_{j=1}^{q+1} \prod_{k=j+1}^{q+1} \gamma_k (1-\beta_j) f(\mathrm{OPT}).$$

Let $\phi_w := \prod_{k=w}^{q+1} \gamma_k$ for $w \in [1, q+1]$ and $\phi_{q+2} = 1$. Then it holds that $\phi_w \cdot \gamma_{w-1} = \phi_{w-1}$ for $2 \leq w \leq q+1$ ($\phi_{i+1} = 1$) and $\{\phi_w\}$ is a non-decreasing sequence.

Using results above, it holds that

$$\begin{aligned}\sum_{j=1}^{q+1} \prod_{k=j+1}^{q+1} \gamma_k (1-\beta_j) &= \sum_{j=1}^{q+1} \phi_{j+1}(1-\beta_j) = \sum_{j=1}^{q+1} \phi_{j+1} \cdot \frac{mc(u_j)}{b} \\ &= \sum_{j=1}^{q+1} \phi_{j+1} \cdot m(1-\gamma_j) = m \cdot \left(\sum_{j=1}^{q+1} \phi_{j+1} - \sum_{j=1}^{q+1} \phi_j \right) = m \cdot (1-\phi_1).\end{aligned}$$

According to the definition of q, it follows that $D := \sum_{k=1}^{q+1} c(u_k) = c(S_{q+1}) > b$. For term ϕ_1, it follows that

$$1-\phi_1 = 1 - \prod_{k=1}^{q+1} \gamma_k = 1 - \prod_{k=1}^{q+1} \left(1 - \frac{c(u_k)}{b}\right) \geq 1 - \prod_{k=1}^{q+1} \left(1 - \frac{c(u_k)}{D}\right) > 1 - \left(1 - \frac{1}{q+1}\right)^{q+1} \geq 1 - e^{-1}$$

where the first inequality holds since for $x_1, \cdots, x_n \in \mathbb{R}^+$ such that $\sum_{i=1}^n x_i = D$, the multivariate function $1 - \prod_{i=1}^n (1 - \frac{x_i}{D})$ achieves its minimum when $x_1 = \cdots = x_n = D/n$.

Hence, we have

$$f(S'_{q+1}) \geq \sum_{j=1}^{q+1} \prod_{k=j+1}^{q+1} \gamma_k (1-\beta_j) f(\mathrm{OPT}) = m(1-\phi_1) f(\mathrm{OPT}) \geq m\left(1-e^{-1}\right) f(\mathrm{OPT}).$$

By submodularity, it follows that

$$f(S'_q) + \max_{e \in V} f(e) \geq f(S'_q) + f(o_1) \geq f(S'_q) + f(o_1 \mid S'_q) \geq m\left(1-e^{-1}\right) f(\mathrm{OPT}). \tag{7}$$

Since Positive Greedy constructs solution set by selecting elements with positive marginal gain, it holds that $f(S_g) \geq f(S'_q)$. By definition of S_g, it holds that $f(S_g) \geq \max_{e \in V} f(e)$. Combining these two inequalities and Inequaity (7), we obtain that

$$f(S_g) \geq 1/2 \left(f(S'_q) + \max_{e \in V} f(e) \right) \geq \frac{m}{2}\left(1-e^{-1}\right) f(\mathrm{OPT}).$$

□

By far, we have obtained one part of the approximation ratio ($m\alpha/2$) in Theorem C.1. Next we aim to prove another part of approximation ratio denoted as $h_1^{\lambda}(m)$.

Lemma D.4. *For any two subsets $S \subseteq T \subseteq V$, it follows that*

$$\max_{t \in T} \frac{f(t \mid S)}{c(t)} \geq \frac{f(T \mid S)}{c(T)}.$$

Proof. Denote $T \setminus S := \{t_1, \cdots, t_n\}$. By submodularity, it follows that

$$\begin{aligned} f(T \mid S) &= \sum_{i=1}^n f(t_i \mid S \cup \{t_1, \cdots, t_{i-1}\}) = \sum_{i=1}^n c(t_i) \cdot \frac{f(t_i \mid S \cup \{t_1, \cdots, t_{i-1}\})}{c(t_i)} \\ &\leq \sum_{i=1}^n c(t_i) \cdot \frac{f(t_i \mid S)}{c(t_i)} \leq \max_{t \in T} \frac{f(t \mid S)}{c(t)} \cdot \sum_{i=1}^n c(t_i) = c(T) \cdot \max_{t \in T} \frac{f(t \mid S)}{c(t)}. \end{aligned}$$

□

Lemma D.5. *Given any element set T, it holds that*

$$f(S_l) \geq m\left(1 - \frac{c(T)}{b}\right) f(T).$$

Proof. If $T \subseteq S_l$, we have $f(S_l) \geq mf(T)$ due to m-monotonicity of submodular function f.

If $T \not\subseteq S_l$, i.e., $T \setminus S_l \neq \emptyset$. Let S_k be the element set constructed by the greedy heuristic when the first element from T is considered but not added to S_k due to budget constraint (meaning that all elements in T can be added into previous solution set). If such element does not exist due to negative marginal gain, it means that $\forall e \in T \setminus S_l$, s.t. $c(e) + c(S_l) \leq b$ and $f(e \mid S_l) < 0$. Under this condition, it follows that

$$mf(T) \leq f(S_l \cup T) \leq f(S_l) + \sum_{e \in T \setminus S_l} f(e \mid S_l) < f(S_l),$$

which also proves the conclusion.

By submodularity and greedy heuristic, we have

$$\rho(u_1 \mid S_0) \geq \rho(u_2 \mid S_1) \geq \cdots \geq \rho(u_t \mid S_{k-1}) \geq \max_{v \in T'} \rho(v \mid S_k) \geq \frac{f(T' \mid S_k)}{c(T')} \tag{8}$$

where $T' := T \setminus S_k$ and the last inequality holds according to Lemma D.4.

According to definition of m-monotonicity, we have

$$mf(T) \leq f(T' \cup S_k) = f(S_k) + f(T' \mid S_k). \tag{9}$$

Using Inequality (8), it holds that

$$f(S_k) = \sum_{i=1}^{k} f(u_i \mid S_{i-1}) = \sum_{i=1}^{k} c(u_i') \cdot \rho(u_i \mid S_{i-1}) \geq \sum_{i=1}^{k} c(u_i') \cdot \frac{f(T' \mid S_k)}{c(T')} = c(S_k) \cdot \frac{f(T' \mid S_k)}{c(T')} \tag{10}$$

Let u_k denote the element in T' abandoned by S_k due to budget violation, it holds that

$$c(S_k) + c(T') \geq c(S_k) + c(u_k) > b. \tag{11}$$

Combining Inequalities (9) (10) (11), we have

$$f(S_l) \geq f(S_k) \geq m\left(1 - \frac{c(T)}{b}\right) f(T).$$

□

Let $OPT_1 := \mathrm{OPT} \setminus (S_q \cup \{o_1\})$.

Lemma D.6. *For the final solution set S_l, it holds that*

$$f(S_l) \geq \frac{m}{\lambda} f(S_q) + m'\left(1 - \frac{c(S_q)}{b - c(S_q)}\right)\left[f(OPT_1 \cup S_q) - \frac{m}{\lambda} f(S_q)\right]$$

where $m' = m + \frac{m^2 - m}{\lambda - m}$ is defined in Lemma B.1.

Proof. According to Lemma B.1, it holds that $g_{S_q}(S) := f(S \cup S_q) - \frac{m}{\lambda} f(S_q)$ is a non-negative m'-monotone submodular function. Using Lemma D.5, we have

$$\begin{aligned} f(S_l) &= f(S_l \cup S_q) - \frac{m}{\lambda} f(S_q) + \frac{m}{\lambda} f(S_q) = \frac{m}{\lambda} f(S_q) + g_{S_q}(S_l \setminus S_q) \qquad (12) \\ &\geq \frac{m}{\lambda} f(S_q) + m'\left(1 - \frac{c(OPT_1)}{b - c(S_q)}\right) g_{S_q}(OPT_1) \qquad (13) \end{aligned}$$

where $S_q \subseteq S_l$.

Furthermore, by definition of S_q and o_1, it holds that if $c(S_q)+c(o_1) > b$ and $OPT_1 = \mathrm{OPT}\setminus(S_q \cup \{o_1\})$. Hence, it follows that $c(o_1)+c(OPT_1) \le c(\mathrm{OPT}) \le b$, implying that $c(OPT_1) < c(S_q)$ since $c(S_q)+c(o_1) > b$. By plugging the result into Inequality (13), we have

$$\begin{aligned} f(S_l) &\ge \frac{m}{\lambda} f(S_q) + m'\left(1-\frac{c(OPT_1)}{b-c(S_q)}\right) g_{S_q}(OPT_1) \\ &\ge \frac{m}{\lambda} f(S_q) + m'\left(1-\frac{c(S_q)}{b-c(S_q)}\right)\cdot\left[f(OPT_1\cup S_q) - \frac{m}{\lambda} f(S_q)\right]. \end{aligned}$$

□

For Positive Greedy, an element u picked due to maximum density can be dumped because budget violation $c(u)+c(S_i) > b$ given the current solution set S_i. Define A_i as the abandoned element set at time when S_i is constructed from S_{i-1}.

Lemma D.7. *For any greedy solution set S_i and element set T such that $T \cap A_i = \emptyset$, it follows that*

$$f(S_i) \ge m\left[1-\exp\left(-\frac{c(S_i)}{c(T)}\right)\right] f(T).$$

Proof. If $f(S_i) \ge f(T)$, the lemma holds since $1-e^{-x} \in [0,1)$ when $x \in [0,\infty)$ and $m \in [0,1]$.

If $f(S_i) < f(T)$, we have

$$mf(T) \le f(T\cup S_i) \le f(S_i) + \sum_{v\in T\setminus S_i} f(v \mid S_i) = f(S_i) + \sum_{v\in T\setminus S_i} c(v)\cdot\frac{f(v\mid S_i)}{c(v)}$$

where the first inequality holds due to definition of monotonicity ratio and the second inequality holds by submodularity. According to the greedy heuristic (choosing element with maximum density), for any $j \le i$ and $v \in T\setminus S_j$, we have

$$\frac{f(u_{j+1}\mid S_j)}{c(u_{j+1})} \ge \frac{f(v\mid S_j)}{c(v)}.$$

Combining this result with previous inequality, it follows

$$\begin{aligned} mf(T) &\le f(S_i) + \sum_{v\in T\setminus S_i} c(v)\cdot\frac{f(v\mid S_i)}{c(v)} \\ &\le f(S_i) + \frac{f(u_{i+1}\mid S_i)}{c(u_{i+1})}\cdot\sum_{v\in T\setminus S_i} c(v) \\ &\le f(S_i) + \frac{f(u_{i+1}\mid S_i)}{c(u_{i+1})}\cdot c(T) \\ &= f(S_i) + \frac{f(S_{i+1})-f(S_i)}{c(u_{i+1})}\cdot c(T) \\ &= \left(1-\frac{c(T)}{c(u_{i+1})}\right)\cdot f(S_i) + \frac{c(T)}{c(u_{i+1})}\cdot f(S_{i+1}). \end{aligned}$$

which could be rewritten as

$$mf(T) - f(S_{i+1}) \le \left(1-\frac{c(u_{i+1})}{c(T)}\right)[mf(T)-f(S_i)]. \tag{14}$$

Using inequality $1-x \le e^{-x}$, we have $1-\frac{c(u_{i+1})}{c(T)} \le \exp\left(-\frac{c(u_{i+1})}{c(T)}\right)$. Apply it into Inequality (14), it follows that

$$mf(T) - f(S_{i+1}) \le \left(1-\frac{c(u_{i+1})}{c(T)}\right)\cdot[mf(T)-f(S_i)] \le \exp\left(-\frac{c(u_{i+1})}{c(T)}\right)\cdot[mf(T)-f(S_i)]. \tag{15}$$

By applying Inequality (15) recursively, it follows that

$$mf(T)-f(S_i) \leq \exp\left(-\frac{c(u_i)}{c(T)}\right) \cdot [mf(T) - f(S_{i-1})] \leq \exp\left(-\frac{\sum_{j=1}^{i} c(u_j)}{c(T)}\right) \cdot [mf(T) - f(S_0)],$$

which could be simplified as

$$f(S_i) \geq m\left[1 - \exp\left(-\frac{c(S_i)}{c(T)}\right)\right] f(T).$$

□

Lemma D.8. *For the greedy solution set S_q, it holds that*

$$f(S_q) \geq m\left[1 - \exp\left(-\frac{c(S_q)}{b}\right)\right] f(\mathrm{OPT}).$$

Proof. According to the definition of S_q, it follows that $\mathrm{OPT} \cap A_q = \emptyset$.

Using Lemma D.7, we have

$$f(S_q) \geq m\left[1 - \exp\left(-\frac{c(S_q)}{c(\mathrm{OPT})}\right)\right] f(\mathrm{OPT}).$$

Using the fact that $c(\mathrm{OPT}) \leq b$ where b is the budget of problem instance, we have

$$f(S_q) \geq m\left[1 - \exp\left(-\frac{c(S_q)}{c(\mathrm{OPT})}\right)\right] f(\mathrm{OPT}) \geq m\left[1 - \exp\left(-\frac{c(S_q)}{b}\right)\right] f(\mathrm{OPT}).$$

□

Lemma D.9. *Let $\alpha = \frac{f(S_g)}{f(\mathrm{OPT})}$, $x_1 = \frac{f(S_q)}{f(\mathrm{OPT})}$, $x_2 = \frac{f(OPT_1|S_q)}{f(\mathrm{OPT})}$ and $x_3 = \frac{c(S_q)}{b}$. It follows that $f(S_g) \geq \alpha^* f(\mathrm{OPT})$ where α^* is the minimum of the following mathematical programming with respect to α, x_1, x_2, x_3*

$$\begin{aligned}
\min\ & \alpha \\
s.t.\ & \alpha \geq x_1 && (16)\\
& \alpha \geq 1 - x_1 - x_2 && (17)\\
& \alpha \geq \frac{m}{\lambda}x_1 + m'\left(1 - \frac{x_3}{1-x_3}\right)\left[x_2 + \left(1 - \frac{m}{\lambda}\right)x_1\right] && (18)\\
& x_1 \geq m(1 - e^{-x_3}) && (19)\\
& \alpha, x_1, x_3 \in [0,1] \quad x_2 \in [-1,1].
\end{aligned}$$

Proof. Since $f(S_l) \geq f(S_q)$ by positive marginal gain condition, it follows that $\alpha \geq x_1$, which corresponds to constraint (16).

Since $f(S_g) \geq \max_{e\in V} f(e) \geq f(o_1) \geq f(o_1 \mid S_q)$, it follows that $\alpha = \frac{f(S_g)}{f(\mathrm{OPT})} \geq \frac{f(o_1|S_q)}{f(\mathrm{OPT})}$. Plus the fact that $f(S_q) + f(OPT_1 \mid S_q) + f(o_1 \mid S_q) \geq f(\mathrm{OPT})$ by definition of $OPT_1 = \mathrm{OPT} \setminus (S_q \cup \{o_1\})$ and submodularity, we have $\alpha \geq 1 - x_1 - x_2$, which corresponds to constraint (17).

Using Lemma D.8, we have

$$f(S_q) \geq m\left[1 - \exp\left(-\frac{c(S_q)}{c(\mathrm{OPT})}\right)\right] f(\mathrm{OPT}) \geq m\left[1 - \exp\left(-\frac{c(S_q)}{b}\right)\right] f(\mathrm{OPT}).$$

Dividing both sides with $f(\mathrm{OPT})$, we have $x_1 \geq m(1 - e^{-x_3})$, which corresponds to constraint (19).

Since $c(S_q) + c(o_1) > b$, using Lemma D.6, we have

$$\begin{aligned}
f(S_l) &\geq \frac{m}{\lambda} f(S_q) + m'\left(1 - \frac{c(S_q)}{b - c(S_q)}\right)\left[f(OPT_1 \cup S_q) - \frac{m}{\lambda} \cdot f(S_q)\right] \\
&= \frac{m}{\lambda} f(S_q) + m'\left(1 - \frac{c(S_q)}{b - c(S_q)}\right)\left[f(OPT_1 \mid S_q) + \left(1 - \frac{m}{\lambda}\right) f(S_q)\right].
\end{aligned}$$

Dividing each side with $f(\mathrm{OPT})$, we have

$$\alpha \geq \frac{m}{\lambda} x_1 + m' \left(1 - \frac{x_3}{1 - x_3}\right) \left[x_2 + \left(1 - \frac{m}{\lambda}\right) x_1\right],$$

which corresponds to constraint (18).

Due to the non-negativity of function f and cost function c, it follows that $\alpha, x_1, x_3 \geq 0$. By definition of OPT, it follows that $\alpha, x_1 \leq 1$. Since $c(S_q) \leq b$, it holds that $x_3 \leq 1$. By submodularity, it holds that

$$x_2 = \frac{f(OPT_1 \mid S_q)}{f(\mathrm{OPT})} \leq \frac{f(\mathrm{OPT} \setminus \{o_1\})}{f(\mathrm{OPT})} \leq 1.$$

By the definition of monotonicity ratio, it holds that

$$x_2 = \frac{f(OPT_1 \mid S_q)}{f(\mathrm{OPT})} \geq \frac{(m-1) f(S_q)}{f(\mathrm{OPT})} \geq m - 1 \geq -1.$$

Hence, $x_2 \in [-1, 1]$.

□

Lemma D.10. *Define* $d_4(x_3) = -m^2 m'^2 - 2m^2 m' + m^2 + x_3^2(-4m^2 m'^2 + m^2) + x_3 \cdot \left(4m^2 m'^2 - 2m^2\right) + \left(-\lambda m' + 2m^2 m'\right) e^{x_3}$ *where* $x_3 \in [0, 1/2]$, m *is monotonicity ratio and* m' *is defined in Lemma B.1.*

1. *If there exists only one zero* x_3^* *such that* $d_4(x_3^*) = 0$, *it holds that* $d_4(x_3) \geq 0$ *when* $x_3 \in [0, x_3^*]$ *and* $d_4(x_3) < 0$ *when* $x_3 \in [x_3^*, 1/2]$.
2. *Otherwise, either* $d_4(x_3) \geq 0$ *or* $d_4(x_3) \leq 0$ *for* $x_3 \in [0, 1/2]$.

Proof. Define the coefficients of x_3 and e^{x_3} in $d_4(x_3)$ as follows:

$$\begin{aligned} a &= -4m^2 m'^2 + m^2 = m^2(1 - 4m'^2), \\ b &= 4m^2 m'^2 - 2m^2 = 2m^2(2m'^2 - 1), \\ c &= -m^2 m'^2 - 2m^2 m' + m^2 = m^2(-m'^2 - 2m' + 1), \\ d &= -\lambda m' + 2m^2 m' = m'(2m^2 - \lambda). \end{aligned}$$

Using these notations, we have $d_4(x) = ax^2 + bx + c + de^x$ and $d_4'(x) = 2ax + b + de^x$.

If there exists only one zero x^* such that $d_4(x^*) = 0$, we need to prove that $d_4'(x) \leq 0$, meaning $d_4(x)$ decreases over $x \in [0, 1/2]$. Hence, it follows that $d_4(x_3) \geq 0$ when $x_3 \in [0, x_3^*]$ and $d_4(x_3) < 0$ when $x_3 \in [x_3^*, 1/2]$, which is consistent with the first part of the claim.

According to definition of zero x^*, it holds $d_4(x^*) = 0$, which is equivalent to

$$a(x^*)^2 + bx^* + c = -de^{x^*}. \tag{20}$$

The symmetric axis of left hand side of Equation (20) is

$$x = -\frac{b}{2a} = -\frac{2m^2(2m'^2 - 1)}{2m^2(1 - 4m'^2)} = \frac{2m'^2 - 1}{4m'^2 - 1}.$$

By spliting the range of m'^2 into three cases, we observe that

1. If $m'^2 \leq \frac{1}{4}$, it follows that $a \geq 0$ and $1 - 4m'^2 \leq 1 - 2m'^2$, i.e., symmetric axis $x = \frac{2m'^2 - 1}{4m'^2 - 1} \geq 1$. Hence, $ax^2 + bx + c$ decreases over $x \in [0, \frac{1}{2}]$.
2. If $m'^2 \in (\frac{1}{4}, \frac{1}{2}]$, it follows that $a < 0$ and $4m'^2 - 1 > 0 \geq 2m'^2 - 1$, i.e., symmetric axis $x = \frac{2m'^2 - 1}{4m'^2 - 1} \leq 0$. Hence, $ax^2 + bx + c$ decreases over $x \in [0, \frac{1}{2}]$.

3. If $m'^2 \in (\frac{1}{2}, 1]$, it follows that $a < 0$ and $4m'^2 - 1 > 2m'^2 - 1 > 0$ and $4m'^2 - 1 > 4m'^2 - 2$, i.e., symmetric axis $x = \frac{2m'^2-1}{4m'^2-1} \in (0, \frac{1}{2})$. Hence, $ax^2 + bx + c$ first increases and then decreases over $x \in [0, \frac{1}{2}]$.

By spliting the range of m^2 into two cases, we observe that

1. If $-d = m'(\lambda - 2m^2) \geq 0$, i.e., $m^2 \leq \frac{\lambda}{2}$, $-de^x$ increases from $-d$ to $-d\sqrt{e}$.
2. If $-d = m'(\lambda - 2m^2) < 0$, i.e., $m^2 > \frac{\lambda}{2}$, $-de^x$ decreases from $-d$ to $-d\sqrt{e}$.

Range of m'^2 Range of m^2	$[0, \frac{1}{4}]$	$(\frac{1}{4}, \frac{1}{2}]$	$(\frac{1}{2}, 1]$
$m^2 \leq \lambda/2$	A1	A2	A3
$m^2 > \lambda/2$	B1	B2	B3

Table 2: Six cases classified by ranges of m'^2 and m^2.

Hence, by specifying the range of m'^2 and m^2, we split the analysis into six cases as described in Table 2. Next we discuss the cases A1-A3 and B1-B3.

If A1 holds, $a \geq 0$, $d \leq 0$, $m'^2 \in [0, \frac{1}{4}]$, $m^2 \in [0, \frac{\lambda}{2}]$, $x \in [0, \frac{1}{2}]$, it follows that

$$d_4'(x) = 2ax + b + de^x, \quad d_4''(x) = 2a + de^x.$$

Since $m = \frac{\lambda m'}{\lambda - 1 + m'}$, it holds that $m \in [0, \frac{\lambda}{2\lambda - 1}]$. Since $d_4''(x)$ decreases over $x \in [0, \frac{1}{2}]$, it holds that $\min_{x \in [0, \frac{1}{2}]} d_4''(x) = d_4''(\frac{1}{2})$ and $\max_{x \in [0, \frac{1}{2}]} d_4''(x) = d_4''(0)$. We split into 3 cases:

1. $d_4''(0) = 2a + d \geq 0$ and $d_4''(\frac{1}{2}) = 2a + d\sqrt{e} \leq 0$, then we aim to prove that $d_4'(x) \leq d_4'(x^\sharp) \leq 0$ where $d_4''(x^\sharp) = 0$.
2. $d_4''(0) = 2a + d < 0$, then we aim to prove that $d_4'(x) \leq d_4'(0) \leq 0$.
3. $d_4''(\frac{1}{2}) = 2a + d\sqrt{e} > 0$, then we aim to prove that $d_4'(x) \leq d_4'(\frac{1}{2}) \leq 0$.

Case 1: Since $d_4''(x^\sharp) = 0$, it holds that $e^{x^\sharp} = -\frac{2a}{d}$ where $-\frac{2a}{d} \in [1, \sqrt{e}]$, i.e., $\ln\left(-\frac{2a}{d}\right) \in [0, \frac{1}{2}]$. Since $e^{x^\sharp} = -\frac{2a}{d}$, it holds that

$$d_4'(x^\sharp) = 2a \ln\left(-\frac{2a}{d}\right) - 2a + b \leq b - a = m^2(8m'^2 - 3) \leq -m^2 \leq 0.$$

Case 2: Since $2a + d < 0$, it holds that

$$d_4'(0) = b + d < b + d - (2a + d) = b - 2a \leq b - a \leq 0$$

where the last inequality holds by case 1 above and $a \geq 0$.

Case 3: Since $2a + d\sqrt{e} > 0$, it follows that

$$d_4'(\frac{1}{2}) = a + b + \sqrt{e}d = -m^2 + m'\sqrt{e}\left(-\lambda + 2m^2\right) \leq -m^2 \leq 0$$

where the second last inequality holds since $m^2 \leq \frac{\lambda}{2}$.

If B1 holds, since $m = \frac{\lambda m'}{\lambda - 1 + m'}$, it holds that $m \in [0, \frac{\lambda}{2\lambda - 1}]$ when $m'^2 \leq \frac{1}{4}$. There exists $\lambda_0 \approx 1.309$ such that when $\lambda < \lambda_0$, it holds that $(\frac{\lambda}{2\lambda - 1})^2 > \frac{\lambda}{2}$, hence $m^2 \in (\frac{\lambda}{2}, (\frac{\lambda}{2\lambda - 1})^2]$. Since $a \geq 0$, $d > 0$, $m'^2 \leq \frac{1}{4}$ and $m^2 > \frac{\lambda}{2}$, we need to prove that

$$\begin{aligned} d_4'(x) \leq d_4'(\frac{1}{2}) &= a + b + \sqrt{e}d = -m^2 + m'\sqrt{e}\left(-\lambda + 2m^2\right) \\ &\leq (\sqrt{e} - 1)m^2 - \frac{\sqrt{e}}{2}\lambda < \frac{\sqrt{e}}{2}(m^2 - \lambda) < 0. \end{aligned}$$

where the first inequality holds since $m^2 \geq \frac{\lambda}{2}$.

If A2 holds, $a < 0$, $d \leq 0$, $m'^2 \in (\frac{1}{4}, \frac{1}{2}]$, $m^2 \in [0, \frac{\lambda}{2}]$, $x \in [0, \frac{1}{2}]$, it follows that 1)

$$d_4'(x) = 2ax + b + de^x \leq b + d = 4m^2m'^2 + m'\left(-\lambda + 2m^2\right) - 2m^2 \leq \frac{\sqrt{2}}{2}(2m^2 - \lambda) < 0$$

if $\frac{\lambda - 2m^2}{8m^2} \leq \frac{1}{2}$, i.e., $m^2 \geq \frac{\lambda}{6}$. 2)

$$d_4'(x) = 2ax + b + de^x \leq b + d = 4m^2m'^2 + m'\left(-\lambda + 2m^2\right) - 2m^2 \leq -\frac{\lambda}{2} < 0$$

if $\frac{\lambda - 2m^2}{8m^2} \geq \frac{1}{\sqrt{2}}$, i.e., $m^2 \leq \frac{\lambda}{4\sqrt{2}+2}$. 3) if $m^2 \in (\frac{\lambda}{4\sqrt{2}+2}, \frac{\lambda}{6})$, $d_4'(x) = 2ax + b + de^x \leq b + d = 4m^2m'^2 + m'\left(-\lambda + 2m^2\right) - 2m^2 \leq \min_{m' \in \{\frac{1}{2}, \frac{1}{\sqrt{2}}\}} 4m^2m'^2 + m'\left(-\lambda + 2m^2\right) - 2m^2 \leq 0$.

If B2 holds, since $m = \frac{\lambda m'}{\lambda - 1 + m'}$ and $m^2 \geq \frac{\lambda}{2}$, it holds that $m \in (\frac{\lambda}{2\lambda - 1}, \frac{\lambda}{\sqrt{2}\lambda + 1 - \sqrt{2}}]$ when $m'^2 \in (\frac{1}{4}, \frac{1}{2}]$. There exists $\lambda_0 \approx 1.5297$ such that when $\lambda < \lambda_0$, it holds that $(\frac{\lambda}{\sqrt{2}\lambda + 1 - \sqrt{2}})^2 > \frac{\lambda}{2}$. Hence, when $\lambda \in (1, \lambda_0]$, $m^2 \in ((\frac{\lambda}{2\lambda - 1})^2, (\frac{\lambda}{\sqrt{2}\lambda + 1 - \sqrt{2}})^2]$; when $\lambda \in (\lambda_0, +\infty)$, $m^2 \in (\frac{\lambda}{2}, (\frac{\lambda}{\sqrt{2}\lambda + 1 - \sqrt{2}})^2]$.

Since $a < 0$, $d > 0$, $m'^2 \in (\frac{1}{4}, \frac{1}{2}]$ and $m^2 > \frac{\lambda}{2}$, we need to prove that $d_4'(x) \leq 0$.

We prove there exists m_0' such that when $m' \in (m_0', \frac{1}{\sqrt{2}}]$, $d_4(x) < 0$. Hence, zero x^* does not exist anymore. In this case, if zero x^* exists, it must have $m' \leq m_0'$.

If there exists $x^\circ \in [0, \frac{1}{2}]$ such that $d_4'(x^\circ) = 2ax^\circ + b + de^{x^\circ} = 0$, it follows that

$$d_4(x^\circ) = a(x^\circ)^2 + (b - 2a)(x^\circ) + c - b$$

whose symmetric axis is $x = -\frac{b}{2a} + 1 = \frac{2(3m'^2 - 1)}{4m'^2 - 1}$.

We need to prove that $d_4'(x) \geq 0$ when $x \in [0, x^\circ]$ and $d_4'(x) < 0$ when $x \in (x^\circ, \frac{1}{2}]$. Hence, $d_4(x) \leq d_4(x^\circ)$.

When $m'^2 \in (\frac{25}{64}, \frac{1}{2}]$, it follows that

$$\begin{aligned} d_4''(x) = 2a + de^x \leq 2a + \sqrt{e}d = &\frac{\lambda m'}{(\lambda + m' - 1)^2} \cdot \left(-\lambda^2\sqrt{e} - 8\lambda m'^3 + 2\lambda\sqrt{e}\right. \\ &\left.+ m'^2 \cdot \left(2\lambda\sqrt{e} - \sqrt{e}\right) + m'\left(-2\lambda\sqrt{e} + 2\lambda + 2\sqrt{e}\right) - \sqrt{e}\right) \\ &\leq -\lambda^2\sqrt{e} - \frac{45\lambda}{64} + \frac{49\lambda\sqrt{e}}{32} - \frac{9\sqrt{e}}{64} < 0 \end{aligned}$$

where the first inequality holds since $d > 0$ and de^x increases over $x \in [0, \frac{1}{2}]$, the second inequality holds since $-\lambda^2\sqrt{e} - 8\lambda x^3 + 2\lambda\sqrt{e} + x^2 \cdot (2\lambda\sqrt{e} - \sqrt{e}) + x\left(-2\lambda\sqrt{e} + 2\lambda + 2\sqrt{e}\right) - \sqrt{e}$ decreases when $x \in (\frac{1}{2}, \frac{1}{\sqrt{2}}]$ and the third inequality holds since $-\lambda^2\sqrt{e} - \frac{45\lambda}{64} + \frac{49\lambda\sqrt{e}}{32} - \frac{9\sqrt{e}}{64}$ decreases over $\lambda \geq 1$ and $\left(-\lambda^2\sqrt{e} - \frac{45\lambda}{64} + \frac{49\lambda\sqrt{e}}{32} - \frac{9\sqrt{e}}{64}\right)_{\lambda=1} < 0$. Specifically, let $f(x) := -\lambda^2\sqrt{e} - 8\lambda x^3 + 2\lambda\sqrt{e} + x^2 \cdot (2\lambda\sqrt{e} - \sqrt{e}) + x\left(-2\lambda\sqrt{e} + 2\lambda + 2\sqrt{e}\right) - \sqrt{e}$. It follows that $f'(x) = -24\lambda x^2 - 2\lambda\sqrt{e} + 2\lambda + 2x\left(2\lambda\sqrt{e} - \sqrt{e}\right) + 2\sqrt{e}$, whose symmetric axis $x = \frac{(2\lambda - 1)\sqrt{e}}{24\lambda} < \frac{\sqrt{e}}{12} < \frac{1}{2}$. Hence, $f'(x) < f'(\frac{1}{2}) = -4\lambda + \sqrt{e} < 0$.

Moreover, $m'^2 \in (\frac{3}{8}, \frac{1}{2}]$, $\max_{x \in [0, \frac{1}{2}]} d_4(x^\circ) = d_4(\frac{1}{2})$ over $x^\circ \in [0, \frac{1}{2}]$. Hence, it holds that

$$d_4(x^\circ) = a(x^\circ)^2 + (b - 2a)(x^\circ) + c - b \leq -\frac{m^2 \cdot (8m' - 5)}{4} \leq 0.$$

Hence, we know that when $m'^2 > \frac{25}{64}$, $d_4(x^\circ) < 0$.

In the following analysis, it holds that $m'^2 \in (\frac{1}{4}, \frac{25}{64}]$. By definition, it follows that

$$d_4(x) = ax^2 + bx + c + de^x, \quad d_4'(x) = 2ax + b + de^x, \quad d_4''(x) = 2a + de^x.$$

Since ax^2+bx+c decreases over $x \in [0, \frac{1}{2}]$, $-de^x < 0$ decreases over $x \in [0, \frac{1}{2}]$ and zero x^* exists, it holds that $d_4(\frac{1}{2}) < 0$ and $d_4(0) = c + \sqrt{e}d > 0$.

Since $d_4''(x)$ increases over $x \in [0, \frac{1}{2}]$, it holds that $\min_{x\in[0,\frac{1}{2}]} d_4''(x) = d_4''(0)$ and $\max_{x\in[0,\frac{1}{2}]} d_4''(x) = d_4''(\frac{1}{2})$. We split into three cases:

1. $d_4''(0) = 2a + d \le 0$ and $d_4''(\frac{1}{2}) = 2a + d\sqrt{e} \ge 0$, then we aim to prove that $d_4'(x) \le \min\{d_4'(0), d_4'(\frac{1}{2})\} \le 0$.
2. $d_4''(0) = 2a + d > 0$, then we aim to prove that $d_4'(x) \le d_4'(\frac{1}{2}) \le 0$.
3. $d_4''(\frac{1}{2}) = 2a + d\sqrt{e} < 0$, then we aim to prove that $d_4'(x) \le d_4'(0) \le 0$.

Case 1: Note that case 2 also holds without condition $2a+d > 0$. Hence, $d_4'(x) \le \min\{d_4'(0), d_4'(\frac{1}{2})\} \le d_4'(\frac{1}{2}) < 0$.

Case 2: Since $2a + d > 0$, it holds that

$$d_4'(\frac{1}{2}) = a + b + \sqrt{e}d < a + b + \sqrt{e}d - d_4(\frac{1}{2}) = \frac{\lambda^2 m'^2 \cdot (8m' - 5)}{4\left(\lambda + m' - 1\right)^2} \le 0$$

where the first inequality holds since $d_4(\frac{1}{2}) < 0$.

Case 3: Since $2a + d\sqrt{e} < 0$, it follows that

$$\begin{aligned} d_4'(0) = b + d < b + d - \frac{1}{\sqrt{e}}(2a + d\sqrt{e}) &= \frac{2\lambda^2 m'^2 \cdot \left(2m'^2\sqrt{e} + 4m'^2 - \sqrt{e} - 2\right)}{\left(\lambda + m' - 1\right)^2 \sqrt{e}} \\ &\le \frac{2\lambda^2 m'^2}{\left(\lambda + m' - 1\right)^2 \sqrt{e}} \cdot \left(\frac{25(2\sqrt{e}+4)}{64} - \sqrt{e} - 2\right) < 0. \end{aligned}$$

where the first inequality holds since $2a + d\sqrt{e} < 0$.

If A3 and B3 holds, zero x^* does not exist. If 3 holds ($m'^2 \in (\frac{1}{2}, 1]$), $ax^2 + bx + c \le \frac{m^2 m'\left(-8m'^2+m'+2\right)}{4m'^2-1}$ when $x = \frac{2m'^2-1}{4m'^2-1}$. To prove it contradicts the existence of zero x^*, we only need to verify that

1) $\frac{m^2 m'\left(-8m'^2+m'+2\right)}{4m'^2-1} < -d$ when $m'^2 \le \frac{\lambda}{2}$;

2) $\frac{m^2 m'\left(-8m'^2+m'+2\right)}{4m'^2-1} < -d\sqrt{e}$ when $m'^2 > \frac{\lambda}{2}$.

Case 1: by rearranging inequality, we need to prove $\frac{m'\left(-4\lambda m'^2+\lambda+m^2 m'\right)}{4m'^2-1} < 0$. It is easy to verify that $-4\lambda m'^2 + \lambda + m^2 m' \le 0$ since

$$\begin{aligned} -4\lambda m'^2 + \lambda + m^2 m' &= (1 - 4m'^2)\lambda + m^2 m' \\ &<^{\lambda=1} 1 - 4m'^2 + m^2 m' \\ &\le 1 - 4m'^2 + m^2 m' \\ &\le 1 - 4m'^2 + m' < 0 \end{aligned}$$

Plus the fact that $4m'^2 - 1 > 0$, it holds that $\frac{m'\left(-4\lambda m'^2+\lambda+m^2 m'\right)}{4m'^2-1} < 0$.

Case 2: in this case, we need to guarantee that $\lambda \le 2$ such that m'^2 is well-defined. By rearranging inequality, we need to prove

$$\frac{m' t(m,\lambda)}{4m'^2 - 1} < 0$$

where $t(m,\lambda) := m^2\left(-8m'^2 + m' + 2\right) - \left(\lambda - 2m^2\right)\left(4m'^2 - 1\right)\sqrt{e}$. It is easy to verify that $t(m,\lambda) = \lambda\left(-4m'^2\sqrt{e} + \sqrt{e}\right) - 8m^2m'^2 + 8m^2m'^2\sqrt{e} + m^2m' - 2m^2\sqrt{e} + 2m^2 < 0$ since

$$\begin{aligned}&\lambda\left(-4m'^2\sqrt{e}+\sqrt{e}\right)-8m^2m'^2+8m^2m'^2\sqrt{e}+m^2m'-2m^2\sqrt{e}+2m^2\\&\qquad<m^2\left(-8m'^2+8m'^2\sqrt{e}+m'-2\sqrt{e}+2\right)-4m'^2\sqrt{e}+\sqrt{e}\leq-5+3\sqrt{e}\approx-0.0538<0\end{aligned}$$

where the first inequality holds due to negativity of coefficient $\sqrt{e}(1-4m'^2)$ of λ, the second inequality holds since $8(\sqrt{e}-1)m'^2+m'+2(1-\sqrt{e})$ increases over $m'^2\in(\frac{\lambda}{2},1]$ and $8(\sqrt{e}-1)m'^2+m'+2(1-\sqrt{e})\geq 4(\sqrt{e}-1)\lambda+2(1-\sqrt{e})>0$.

If zero x^* does not exist, it means that $f_1(x)=ax^2+bx+c$ and and $f_2(x)=-de^x$ does not have intersection over interval $x\in[0,\frac{1}{2}]$. Hence, $d_4(x)\geq 0$ or $d_4(x)\leq 0$ for $x\in[0,\frac{1}{2}]$.

We need to prove that for $x\in[0,\frac{1}{2}]$, the number of zeros of $d_4(x)$ is no more than 1.

$$d_4(x^*)=0\Longleftrightarrow a(x^*)^2+bx^*+c=-de^{x^*}$$

where we denote left hand side as LHS(x) and right hand side as RHS(x). Like Case 1, we split the analysis into six cases as described in Table 2.

If A1 or A2 hold, the number of zeros is at most 1.

If A3 holds, when $m'^2\in(\frac{1}{2},1])$, $ax^2+bx+c\leq\frac{m^2m'\left(-8m'^2+m'+2\right)}{4m'^2-1}\leq\frac{\lambda\left(-4+\sqrt{2}\right)}{2\cdot\left(2\lambda-2+\sqrt{2}\right)}<0$ and $-de^x>0$. Hence, the number of zeros is 0.

If B1 or B2 hold, we consider three cases:

1) if $\text{LHS}(0)=c\geq-d=\text{RHS}(0)$ and $\text{LHS}(\frac{1}{2})=\frac{a}{4}+\frac{b}{2}+c\leq-d\sqrt{e}=\text{RHS}(\frac{1}{2})$, the number of zeros is 1.
2) if $\text{LHS}(0)=c\geq-d=\text{RHS}(0)$ and $\text{LHS}(\frac{1}{2})=\frac{a}{4}+\frac{b}{2}+c\geq-d\sqrt{e}=\text{RHS}(\frac{1}{2})$, the number of zeros is 0. In fact, we have proved that under 1 and 5 (or 2 and 5), $d_4'(x)<0$, meaning that $\text{LHS}'(x)\leq\text{RHS}'(x)$. Plus the conditions $\text{LHS}(0)\geq\text{RHS}(0)$ and $\text{LHS}(\frac{1}{2})\geq\text{RHS}(\frac{1}{2})$, it is easy to verify that the number of zeros is 0.
3) if $\text{LHS}(0)=c<-d=\text{RHS}(0)$, it follows that $d_4(x)<0$, i.e., the number of zeros is 0.

Hence, if B1 or B2 hold, the number of zeros is at most 1.

If B3 holds, we have proved previously that zero x^* does not exist.

In conclusion, we prove that if there exists only one zero x_3^* such that $d_4(x_3^*)=0$, it holds that $d_4(x_3)\geq 0$ when $x_3\in[0,x_3^*]$ and $d_4(x_3)<0$ when $x_3\in[x_3^*,\frac{1}{2}]$. Otherwise, either $d_4(x_3)\geq 0$ or $d_4(x_3)\leq 0$ for $x_3\in[0,\frac{1}{2}]$.

□

Theorem D.11. *Positive Modified Greedy (Algorithm 2) achieves an approximation ratio of*

$$\min\left\{\frac{m\left(\lambda-1\right)}{\lambda+m\left(\lambda-1\right)-m},\frac{m^2}{\lambda}\left(1-\frac{1}{\sqrt{e}}\right)\right\}$$

for maximizing a non-negative m-monotone submodular function subject to a knapsack constraint for any $\lambda\geq 1$.

Proof. In the proof, we analyze the mathematical programming defined in Lemma D.9.

If $x_3>1/2$, using constraint (16) and constraint (19), it follows that

$$\alpha\geq x_1\geq m(1-e^{-x_3})\geq m\left(1-\frac{1}{\sqrt{e}}\right)\geq\frac{m^2}{\lambda}\left(1-\frac{1}{\sqrt{e}}\right)$$

since $m\in[0,1]$ and $1-e^{-x}$ is increasing with respect to x. Hence, the theorem holds if $x_3>1/2$.

In the remaining proof, we assume $x_3\leq 1/2$. By computing

$$m'(1-2x_3)\times(17)+(1-x_3)\times(18)+\frac{m\left((2m'-1)x_3-m'+1\right)}{\lambda}\times(19),\tag{21}$$

we have $d_1/\lambda \geq 0$ where $d_1 := \alpha\left(-2\lambda m' x_3 + \lambda m' - \lambda x_3 + \lambda\right) + 2\lambda m' x_3 - \lambda m' - 2m^2 m' x_3 + 2m^2 m' x_3 e^{-x_3} + m^2 m' - m^2 m' e^{-x_3} + m^2 x_3 - m^2 x_3 e^{-x_3} - m^2 + m^2 e^{-x_3}$. (Note that it is easy to verify that $-2m' x_3 + m' = m'(1-2x_3) \geq 0$ and $1 - x_3 > 0$. If $m' \geq \frac{1}{2}$, it follows that $(2m'-1)x_3 - m' + 1 \geq 1 - m' \geq 0$. If $m' < \frac{1}{2}$, it follows that $(2m'-1)x_3 - m' + 1 \geq = \frac{1-m'}{2} > 0$. Hence, the coefficients used in (21) do not change directions of inequalities.)

Consider the coefficient of α in d_1, by substituting $m' := m + \frac{m^2 - m}{\lambda - m}$, it follows that

$$-2\lambda m' x_3 + \lambda m' - \lambda x_3 + \lambda = \frac{(-2\lambda^2 m - \lambda^2 + 3\lambda m)x_3 + \lambda^2 m + \lambda^2 - 2\lambda m}{\lambda - m}.$$

Consider the coefficient of x_3 as a quadratic form with respect to λ. Since $-2m - 1 < 0$ and symmetric axis $\lambda = \frac{3m}{4m+2} < 1$, it follows that $-2\lambda^2 m - \lambda^2 + 3\lambda m$ decreases when $\lambda \geq 1$. Hence, $-2\lambda^2 m - \lambda^2 + 3\lambda m <^{\lambda=1} -2m - 1 + 3m = m - 1 \leq 0$. Plus the fact that $\lambda - m > 0$, it follows that $-2\lambda m' x_3 + \lambda m' - \lambda x_3 + \lambda \geq \lambda/2 > 0$ since $-2\lambda m' x_3 + \lambda m' - \lambda x_3 + \lambda$ decreases over $x_3 \in [0, \frac{1}{2}]$.

Hence, we obtain inequality $\alpha \geq d_2(x_3)$ where

$$d_2(x_3) := \frac{\text{Denom}(d_2)}{-2\lambda m' x_3 + \lambda m' - \lambda x_3 + \lambda}$$

and $\text{Denom}(d_2) := -2\lambda m' x_3 + \lambda m' + 2m^2 m' x_3 - 2m^2 m' x_3 e^{-x_3} - m^2 m' + m^2 m' e^{-x_3} - m^2 x_3 + m^2 x_3 e^{-x_3} + m^2 - m^2 e^{-x_3}$. Computing the first derivative of $d_2(x_3)$ with respect to variable x_3, we have

$$d_2'(x_3) = \frac{d_4(x_3)}{\lambda e^{x_3} d_3(x_3)}.$$

where $d_4(x_3) := -\lambda m' e^{x_3} - 4m^2 m'^2 x_3^2 + 4m^2 m'^2 x_3 - m^2 m'^2 + 2m^2 m' e^{x_3} - 2m^2 m' + m^2 x_3^2 - 2m^2 x_3 + m^2$ and $d_3(x_3) := 4m'^2 x_3^2 - 4m'^2 x_3 + m'^2 + 4m' x_3^2 - 6m' x_3 + 2m' + x_3^2 - 2x_3 + 1$.

Recollecting term $d_3(x_3)$ with respect to x_3, we obtain that

$$d_3(x_3) = m'^2 + 2m' + 1 + x_3^2 \cdot \left(4m'^2 + 4m' + 1\right) + x_3\left(-4m'^2 - 6m' - 2\right).$$

Let $a := 4m'^2 + 4m' + 1$, $b := -4m'^2 - 6m' - 2$, $c := m'^2 + 2m' + 1$, it is easy to verify that $b^2 - 4ac = 0$. Plus the fact that $a > 0$, it holds that $d_3(x_3)$ is non-negative.

Hence, it follows that $\lambda e^{x_3} d_3(x_3) \geq 0$, which is exactly the denominator of $d_2'(x_3)$.

Next we focus on the property of numerator of $d_2'(x_3)$. By collecting coefficients of x_3 and e^{x_3}, we could re-write $d_4(x_3)$ as $d_4(x_3) = -m^2 m'^2 - 2m^2 m' + m^2 + x_3^2(-4m^2 m'^2 + m^2) + x_3 \cdot \left(4m^2 m'^2 - 2m^2\right) + \left(-\lambda m' + 2m^2 m'\right) e^{x_3}$.

By Lemma D.10, one of the following three cases hold: 1) $d_4(x_3) \geq 0$ when $x_3 \in [0, x_3^*]$ and $d_4(x_3) < 0$ when $x_3 \in [x_3^*, 1/2]$ for some zero $x^* \in [0, 1/2]$; 2) $d_4(x_3) \geq 0$ for $x_3 \in [0, 1/2]$; 3) $d_4(x_3) \leq 0$ for $x_3 \in [0, 1/2]$. Since $d_2'(x_3) = d_4(x_3)/(\lambda e^{x_3} d_3(x_3))$ and $\lambda e^{x_3} d_3(x_3) \geq 0$, it follows that $d_2(x_3)$ cannot achieve its minimum at middle point on interval $x_3 \in [0, \frac{1}{2}]$. Hence, $\alpha \geq \min\{d_2(0), d_2(\frac{1}{2})\} = \min\left\{\frac{m(\lambda-1)}{\lambda + m(\lambda-1) - m}, \frac{m^2}{\lambda}\left(1 - \frac{1}{\sqrt{e}}\right)\right\}$.

□

Theorem C.1. *Positive Modified Greedy (Algorithm 2) achieves an approximation ratio of* $\max\{m\alpha/2, \max_{\lambda \geq 1}\{h_1^\lambda(m)\}\}$ *for maximizing a non-negative m-monotone submodular function with a knapsack constraint where*

$$h_1^\lambda(m) := \min\left\{\frac{m(\lambda-1)}{\lambda + m(\lambda-1) - m}, \frac{m^2}{\lambda}\left(1 - \frac{1}{\sqrt{e}}\right)\right\}.$$

Proof. We obtain this result by combining Theorem D.3 and Theorem D.11. □

E Analysis of Positive Greedy+Max

In this section, we give detailed analysis of approximation ratio for Positive Greedy+Max (Algorithm 3).

Lemma E.1. *For the greedy solution set* S_{i^*}*, it holds that* $g_1(c(S_{i^*})) + (b - c(o_1))g'(c(S_{i^*})) \geq mf(\mathrm{OPT})$.

Proof. By definition of g_1, we have $g_1(c(S_{i^*})) = f(S_{i^*} \cup \{o_1\})$. By the definition of monotonicity ratio, submodularity and greedy heuristic, it follows that

$$\begin{aligned} mf(\mathrm{OPT}) &\leq f(S_{i^*} \cup \mathrm{OPT}) = f(S_{i^*} \cup \{o_1\} \cup \mathrm{OPT} \setminus \{o_1\}) \\ &= f(S_{i^*} \cup \{o_1\}) + f(\mathrm{OPT} \setminus (S_{i^*} \cup \{o_1\}) \mid S_{i^*} \cup \{o_1\}) \\ &\leq g_1(c(S_{i^*})) + \sum_{e \in \mathrm{OPT} \setminus (S_{i^*} \cup \{o_1\})} f(e \mid S_{i^*} \cup \{o_1\}) \\ &= g_1(c(S_{i^*})) + \sum_{e \in \mathrm{OPT} \setminus (S_{i^*} \cup \{o_1\})} c(e)\rho(e \mid S_{i^*} \cup \{o_1\}). \end{aligned}$$

Since $c(S_{i^*}) \leq b - c(o_1)$ and $o_1 = \max_{e \in \mathrm{OPT}} c(e)$, it follows that $c(S_{i^*}) + c(e) \leq b$ for any $e \in \mathrm{OPT} \setminus (S_{i^*} \cup \{o_1\})$. By submodularity and greedy heuristic, it holds that

$$\max_{e \in \mathrm{OPT} \setminus (S_{i^*} \cup \{o_1\})} \rho(e \mid S_{i^*} \cup \{o_1\}) \leq \max_{e \in \mathrm{OPT} \setminus (S_{i^*} \cup \{o_1\})} \rho(e \mid S_{i^*}) \leq g'(c(S_{i^*})).$$

Hence, we have

$$\begin{aligned} mf(\mathrm{OPT}) &\leq g_1(c(S_{i^*})) + \sum_{e \in \mathrm{OPT} \setminus (S_{i^*} \cup \{o_1\})} c(e)\rho(e \mid S_{i^*} \cup \{o_1\}) \\ &\leq g_1(c(S_{i^*})) + \sum_{e \in \mathrm{OPT} \setminus (S_{i^*} \cup \{o_1\})} c(e)g'(c(S_{i^*})) \\ &= g_1(c(S_{i^*})) + g'(c(S_{i^*}))c(\mathrm{OPT} \setminus (S_{i^*} \cup \{o_1\})) \\ &\leq g_1(c(S_{i^*})) + g'(c(S_{i^*}))(b - c(o_1)). \end{aligned}$$

□

Proof of Theorem C.2. Let T be the output of the Positive Greedy+Max. For each greedy solution set S_i ($1 \leq i \leq l$), its function value is no larger than the augmented solution of S_{i-1}. Since T is the augmented solution with largest function value, it holds that $f(T) \geq f(S_i)$ for any $1 \leq i \leq l$. For the greedy solution set S_{i^*}, it holds that $f(T) \geq f(S_{i^*} \cup \{o_1\})$.

If $g_1(c(S_{i^*})) \geq mf(\mathrm{OPT})/2$, we have $m/2$-approximation guarantee since $f(T) \geq g_1(c(S_{i^*})) = f(S_{i^*} \cup \{o_1\})$.

In the remaining proof, we assume $g_1(c(S_{i^*})) < mf(\mathrm{OPT})/2$. By Lemma E.1, it holds that

$$g'(c(S_{i^*})) \geq \frac{mf(\mathrm{OPT}) - g_1(c(S_{i^*}))}{b - c(o_1)} \tag{22}$$

by rearrangement. By greedy heuristic and submodularity, $g'(x)$ is non-increasing. Hence, we have

$$g(x) = \int_{t=0}^{x} g'(t)dt \geq \int_{t=0}^{x} g'(x)dt = g'(x)x.$$

Let $x = c(S_{i^*})$, we have $g(c(S_{i^*})) \geq g'(c(S_{i^*})) \cdot c(S_{i^*})$. Combining with Inequality (22), it follows that

$$g(c(S_{i^*})) \geq \frac{mf(\mathrm{OPT}) - g_1(c(S_{i^*}))}{b - c(o_1)} \cdot c(S_{i^*}). \tag{23}$$

Since $c(S_{i^*+1}) > b - c(o_1)$, it holds that $c(S_{i^*+1}) - c(S_{i^*}) > b - c(o_1) - c(S_{i^*})$. Hence, we have

$$\begin{aligned} g(c(S_{i^*+1})) &> g(c(S_{i^*})) + (b - c(o_1) - c(S_{i^*}))\frac{g(c(S_{i^*+1})) - g(c(S_{i^*}))}{c(S_{i^*+1}) - c(S_{i^*})} \\ &= g(c(S_{i^*})) + (b - c(o_1) - c(S_{i^*}))g'(c(S_{i^*})) \\ &\geq \frac{mf(\mathrm{OPT}) - g_1(c(S_{i^*}))}{b - c(o_1)} \cdot c(S_{i^*}) + (b - c(o_1) - c(S_{i^*})) \cdot \frac{mf(\mathrm{OPT}) - g_1(c(S_{i^*}))}{b - c(o_1)} \\ &= mf(\mathrm{OPT}) - g_1(c(S_{i^*})) > \frac{mf(\mathrm{OPT})}{2} \end{aligned}$$

where the first inequality holds since $c(S_{i^*+1}) - c(S_{i^*}) > b - c(o_1) - c(S_{i^*})$, the first equality holds since g' is the right derivative, the second inequality holds due to Inequality (23) and Inequality (22), the third inequality holds by assumption $g_1(c(S_{i^*})) < mf(\mathrm{OPT})/2$. Since $f(T) \geq g(c(S_{i^*+1})) = f(S_{i^*+1})$, it holds that $f(T) \geq mf(\mathrm{OPT})/2$.

In conclusion, Positive Greedy+Max (Algorithm 3) gives an approximation ratio of $m/2$ to the non-negative m-monotone submodular maximization with a knapsack constraint. ☐

F Analysis of Two Set Enumeration Positive Greedy

In this section, we give analysis of approximation ratio for Two Set Enumeration Positive Greedy (Algorithm 4).

Let f be a non-negative m-monotone submodular function with ground set V where m is monotonicity ratio. Let OPT be the optimal solution set with maximum function value, i.e., $f(\mathrm{OPT}) = \max_{S\subseteq V, c(S)\leq b} f(S)$. Without loss of generality, we assume $f(\emptyset) = 0$. Given an initial set U, Two Set Enumeration Positive Greedy makes one call to Positive Greedy algorithm (Algorithm 1) to obtain a partial solution set S_l with cardinality l and cost $c(S_l) \leq b$. Let $S_i (0 \leq i \leq l)$ be the partial solution set generated by Positive Greedy algorithm. It holds that $S_0 = U$. Let u_i be the element considered at each step with maximum density and added into the current solution set S_{i-1}, i.e. $S_i = S_{i-1} \cup \{u_i\}$.

Define the continuous extension of greedy solution sequence as $g(x)$ where $x \in [0, b]$ as $g(x) := f(S_i) + \frac{x - c(S_i)}{c(u_{i+1})} \cdot f(u_{i+1} \mid S_i)$ where i is the largest solution index satisfying $c(S_i) \leq x$. It is easy to verify that $g(c(S_i)) = f(S_i)$.

Lemma F.1. *For any two subsets $R \subseteq T \subseteq V$ and $0 \leq i \leq l$, if $c(T) \leq b$ and $b - c(S_i) \geq \max_{s\in T\setminus(R\cup S_i)} c(s)$, it follows that*

$$\begin{aligned} (b - c(R)) \cdot \frac{f(u_{i+1} \mid S_i)}{c(u_{i+1})} &\geq mf(T) - f(S_i \cup R). && (24) \\ &\geq mf(T) - g(c(S_i)) - \sum_{s\in R\setminus S_i} f(s) && (25) \\ &\geq mf(T) - g(c(S_i)) - \sum_{s\in R} f(s). && (26) \end{aligned}$$

Proof. Define $R_i := S_i \cup R$. Since $T \cap R_i = T \cap (S_i \cup R) = (T \cap S_i) \cup (T \cap R) = (T \cap S_i) \cup R$, it follows that $c(T \cap R_i) \geq c(R)$.

By the definition of monotonicity ratio, submodularity and greedy heuristic, it follows that

$$\begin{aligned} mf(T) - f(R_i) \leq f(T \cup R_i) - f(R_i) &\leq \sum_{s\in T\setminus R_i} f(s \mid R_i) \leq \sum_{s\in T\setminus R_i} f(s \mid S_i) \\ &= \sum_{s\in T\setminus R_i} c(s) \cdot \frac{f(s \mid S_i)}{c(s)} \leq \frac{f(u_{i+1} \mid S_i)}{c(u_{i+1})} \cdot \sum_{s\in T\setminus R_i} c(s) \\ &= [c(T) - c(T \cap R_i)] \cdot \frac{f(u_{i+1} \mid S_i)}{c(u_{i+1})} \leq (b - c(R)) \cdot \frac{f(u_{i+1} \mid S_i)}{c(u_{i+1})} \end{aligned}$$

where the first inequality holds by the definition of monotonicity ratio, the second and third inequalities hold due to submodularity, the fourth inequality holds due to greedy heuristic and condition: $b - c(S_i) \geq \max_{s\in T\setminus(R\cup S_i)} c(s)$, the last inequality holds since $c(T) \leq b$ and $c(T \cap R_i) \geq c(R)$. Hence, Inequality (24) holds.

Furthermore, due to the non-negativity and the submodularity of f and the definition of g, we have

$$f(S_i \cup R) \leq f(S_i) + \sum_{u \in R \setminus S_i} f(u) = g(c(S_i)) + \sum_{u \in R \setminus S_i} f(u) \leq g(c(S_i)) + \sum_{u \in R} f(u).$$

Hence, Inequality (25) and Inequality (26) hold.

☐

Lemma F.2. *Given a subset $T \subseteq V$ such that $c(T) \leq b$, let $k (1 \leq k \leq l)$ be the smallest index such that there exists an element $w \in T$ satisfying $c(w) > b - c(S_k)$.*

1. *If k exists, $f(S_l) \geq (1 - e^{-1})[mf(T) - f(w)]$;*
2. *Otherwise, $f(S_l) \geq mf(T)$.*

Proof. Firstly, we consider the case when k exists. If $f(S_l) \geq mf(T) - f(w)$, the lemma already holds. In the following, we assume that $f(S_l) < mf(T) - f(w)$. Hence, it holds that $mf(T) - f(S_i) - f(w) > 0$ for $0 \leq i \leq l$.

According to Lemma F.1, by choosing $R := \{w\}$, we have $R \subseteq T$ since $w \in T$ and $b - c(S_i) \geq \max_{s \in T \setminus (\{w\} \cup S_i)} c(s)$ for $0 \leq i \leq k-1$ since w is the budget violation element with the smallest index k. Using Inequality (24), we have

$$\frac{f(u_{i+1} \mid S_i)}{c(u_{i+1})} \geq \frac{mf(T) - f(S_i) - f(w)}{b - c(w)}, \quad 0 \leq i \leq k-1.$$

By rearranging the inequality and applying $1 - x \leq e^{-x}$, it follows that

$$\begin{aligned} mf(T) - f(S_{i+1}) - f(w) &\leq \left(1 - \frac{c(u_{i+1})}{b - c(w)}\right) [mf(T) - f(S_i) - f(w)] \\ &\leq \exp\left(-\frac{c(u_{i+1})}{b - c(w)}\right) [mf(T) - f(S_i) - f(w)]. \end{aligned}$$

Recursively applying this inequality for $i \in [0, k-1]$, it holds that

$$\begin{aligned} mf(T) - f(S_k) - f(w) &\leq \exp\left(\sum_{j=0}^{k-1} -\frac{c(u_{j+1})}{b - c(w)}\right) [mf(T) - f(S_0) - f(w)] \\ &= \exp\left(-\frac{c(S_k)}{b - c(w)}\right) [mf(T) - f(S_0) - f(w)] \\ &\leq e^{-1}[mf(T) - f(S_0) - f(w)] \leq e^{-1}[mf(T) - f(w)] \end{aligned}$$

where the last inequality holds due to $c(S_k) + c(w) > 1$ and the non-negativity of f. Hence, we have $f(S_k) \geq (1 - e^{-1})[mf(T) - f(w)]$.

Secondly, if k does not exist, it follows that $\forall 1 \leq i \leq l$, $w \in T$, s.t. $c(w) + c(S_i) \leq b$. Consider the final solution set S_l. If $T \subseteq S_l$, it follows that $f(S_l) \geq mf(T)$ by the definition of monotonicity ratio. Otherwise, for any element $w \in T \setminus S_l$, it holds that $f(w \mid S_l) \leq 0$ since w could still be added to S_l but it is not selected by Positive Greedy algorithm. Denote set $T \setminus S_l := \{w_1, w_2, \cdots, w_p\}$. It follows that

$$f(S_l \cup T) = f(S_l) + \sum_{i=1}^{p} f(w_i \mid S_l \cup \{w_1, \cdots, w_{i-1}\}) \leq f(S_l) + \sum_{i=1}^{p} f(w_i \mid S_l) \leq f(S_l)$$

where the first inequality holds due to submodularity of function f and the second inequality holds due to the condition $f(w_i \mid S_l) \leq 0$. Hence, it holds that $f(S_l) \geq f(S_l \cup T) \geq mf(T)$. ☐

Lemma F.3. *Let $\bar{t}_2(m, \lambda) = -\frac{\alpha m^3}{\lambda^2 - \lambda m} + \frac{\alpha m^2}{\lambda^2 - \lambda m} - \frac{3\alpha}{2} - \frac{\alpha m^2}{\lambda} + \frac{\alpha m}{\lambda} + \frac{m}{\lambda}$ where $\alpha = 1 - e^{-1}$. When $\lambda \in (1, \frac{25\alpha^2 - 4\alpha + 4}{24\alpha^2}]$, there exists two roots $m_1, m_2 \in [0, 1]$ and $m_1 \leq m_2$, i.e.,*

$$m_1 := \frac{\lambda\left(5\alpha + 2 - \sqrt{-24\alpha^2\lambda + 25\alpha^2 - 4\alpha + 4}\right)}{4\left(\alpha\lambda + 1\right)},$$

$$m_2 := \frac{\lambda\left(5\alpha + 2 + \sqrt{-24\alpha^2\lambda + 25\alpha^2 - 4\alpha + 4}\right)}{4\left(\alpha\lambda + 1\right)},$$

such that $\bar{t}_2(m_1, \lambda) = \bar{t}_2(m_2, \lambda) = 0$. *Furthermore,* $\bar{t}_2(m, \lambda) \geq 0$ *when* $m \in [m_1, m_2]$*; otherwise,* $\bar{t}_2(m, \lambda) < 0$.

Proof. By multiplying $(\lambda^2 - \lambda m)$ on $\bar{t}_2(m, m', \lambda)$, it follows that

$$\bar{t}_2(m, m', \lambda) \cdot (\lambda^2 - \lambda m) = (-\alpha\lambda - 1)\, m^2 + \left(\frac{5\alpha\lambda}{2} + \lambda\right) m - \frac{3\alpha\lambda^2}{2}, \tag{27}$$

indicating that $\bar{t}_2(m_1, \lambda) = \bar{t}_2(m_2, \lambda) = 0$. In order to guarantee the existence of these two roots, we obtain an upper bound of parameter λ:

$$-24\alpha^2\lambda + 25\alpha^2 - 4\alpha + 4 \geq 0 \iff \lambda \leq \frac{25\alpha^2 - 4\alpha + 4}{24\alpha^2} \approx 1.195.$$

Define

$$g_2(x) = -\frac{\alpha x^3}{\lambda^2 - \lambda x} + \frac{\alpha x^2}{\lambda^2 - \lambda x} - \frac{3\alpha}{2} - \frac{\alpha x^2}{\lambda} + \frac{\alpha x}{\lambda} + \frac{x}{\lambda}.$$

Taking the first derivative of $g_2(x)$, we obtain

$$g_2'(x) = -\frac{\alpha\lambda x^3}{(\lambda^2 - \lambda x)^2} + \frac{\alpha\lambda x^2}{(\lambda^2 - \lambda x)^2} - \frac{3\alpha x^2}{\lambda^2 - \lambda x} + \frac{2\alpha x}{\lambda^2 - \lambda x} - \frac{2\alpha x}{\lambda} + \frac{\alpha}{\lambda} + \frac{1}{\lambda}.$$

Since

$$g_2''(x) = -\frac{2\alpha\lambda^2 m^3}{(\lambda^2 - \lambda m)^3} + \frac{2\alpha\lambda^2 m^2}{(\lambda^2 - \lambda m)^3} - \frac{6\alpha\lambda m^2}{(\lambda^2 - \lambda m)^2} + \frac{4\alpha\lambda m}{(\lambda^2 - \lambda m)^2} - \frac{6\alpha m}{\lambda^2 - \lambda m} + \frac{2\alpha}{\lambda^2 - \lambda m} - \frac{2\alpha}{\lambda},$$

and

$$(\lambda^2 - \lambda m)^3 \cdot g_2''(x) = 2\alpha\lambda^4 \cdot (1 - \lambda) < 0,$$

it follows that $g_2'(x)$ decreases on interval $x \in [0, 1]$.

We assert that $g_2'(1) = \frac{-\alpha\lambda + \lambda - 1}{\lambda(\lambda - 1)}$ increases on interval $\lambda \in (1, 1.195]$. Let $u(\lambda) := \frac{-\alpha\lambda + \lambda - 1}{\lambda(\lambda - 1)}$, it follows that $u'(\lambda) = \frac{(\sqrt{\alpha}\lambda - (\lambda - 1))(\sqrt{\alpha}\lambda + (\lambda - 1))}{\lambda^2(\lambda - 1)^2}$. Since $\lambda \leq 1.195$, it follows that $\sqrt{\alpha}\lambda - (\lambda - 1) \geq 0.755 > 0$, indicating that $u'(\lambda) > 0$. Hence, $g_2'(1)$ increases on interval $\lambda \in (1, 1.195]$.

Since $g_2'(0) = \frac{\alpha}{\lambda} + \frac{1}{\lambda} > 0$ and $g_2'(1) = \frac{-\alpha\lambda + \lambda - 1}{\lambda(\lambda - 1)} < 0$,

it follows that $g_2'(x)$ has a root x_0. Hence, $g_2'(x) \geq 0$ on interval $x \in [0, x_0]$ and $g_2'(x) \leq 0$ on interval $x \in [x_0, 1]$, meaning that $g_2(x)$ firstly increases and then decreases, i.e., $\max_{x \in [0,1]} g_2(x) = g_2(x_0)$. Since $g_2(m) = \bar{t}_2(m, m', \lambda)$, it follows that $\bar{t}_2(m, m', \lambda) \geq 0$ when $m \in [m_1, m_2]$; otherwise, $\bar{t}_2(m, m', \lambda) < 0$.

□

Theorem C.3. *Let* $t_2(m, m', \lambda) := \frac{m}{\lambda} - \frac{3}{2}\alpha + \frac{\alpha m(1 - m')}{\lambda}$ *and* $\alpha_0 := \frac{25\alpha^2 - 4\alpha + 4}{24\alpha^2} \approx 1.195$. *Two Set Enumeration Positive Greedy (Algorithm 4) achieves an approximation ratio of* $\max_{\lambda \geq 1}\{h_2^{\lambda}(m)\}$ *for maximizing a non-negative m-monotone submodular function with a knapsack constraint where*

- *Case 1:*

$$h_2^{\lambda}(m) := \begin{cases} \alpha m', & m \in [m_1, m_2] \\ \alpha m' + t_2(m, m', \lambda), & \text{otherwise} \end{cases}$$

$m_1 := \frac{\lambda\left(5\alpha + 2 - \sqrt{-24\alpha^2\lambda + 25\alpha^2 - 4\alpha + 4}\right)}{4(\alpha\lambda + 1)}$ *and* $m_2 := \frac{\lambda\left(5\alpha + 2 + \sqrt{-24\alpha^2\lambda + 25\alpha^2 - 4\alpha + 4}\right)}{4(\alpha\lambda + 1)}$ *if* $1 \leq \lambda \leq \alpha_0$;

- *Case 2:* $h_2^{\lambda}(m) := \alpha m' + t_2(m, m', \lambda)$ *if* $\lambda > \alpha_0$.

Proof. Let $Y := \{u_1, u_2\} = \arg\max_{Y' \subseteq \mathrm{OPT}, |Y'|=2} f(Y')$. Define a problem instance as $\mathcal{I} := \{V, f, c, b\}$. Using Y as initial solution set, we run Positive Greedy algorithm on a residual problem instance $\mathcal{I}_1 := \{V \setminus Y, f_1, c, b - c(Y)\}$ where $f_1(S) = f(S \cup Y) - f(Y)$. However, since monotonicity ratio of f_1 does not have explicit form, we introduce another equivalent residual problem instance only for purpose of analysis. Specifically, define residual instance for analysis as $\mathcal{I}_2 := \{V \setminus Y, f_2, c, b - c(Y)\}$ where $f_2(S) := f(S \cup Y) - \frac{m}{\lambda} \cdot f(Y)$. Given any partial solution set S_i, it follows that $\frac{f(u_{i+1}|S_i)}{c(u_{i+1})} = \frac{f_1(u_{i+1}|S_i)}{c(u_{i+1})} = \frac{f_2(u_{i+1}|S_i)}{c(u_{i+1})}$. Hence, as long as the initial solution set is determined, the generated partial solution set is identical under Positive Greedy algorithm among problem instances $\{\mathcal{I}, \mathcal{I}_1, \mathcal{I}_2\}$. Since monotonicity ratio m cannot be efficiently computed, we cannot optimize problem instance $\mathcal{I}_2$ in our algorithm. Hence, problem instance $\mathcal{I}_2$ is introduced for the purpose of analysis.

For residual instance $\mathcal{I}_2$, denote its partial solution set as S_i' $(0 \le i \le l')$ given the initial set Y. Let k' be the smallest integer $1 \le k' \le l'$ such that there exists an element $w' \in \mathrm{OPT} \setminus Y$ satisfying $c(w') > b - c(Y) - c(S'_{k'})$. In other words, k' is the smallest index for budget violation condition. If k' does not exist, we could use the second case of Lemma F.2 to obtain approximation ratio.

$\mathrm{OPT} \setminus Y$ is an optimal solution for the residual instance $\mathcal{I}_2$. In fact, for objective $f_2(S)$ of the residual instance $\mathcal{I}_2$, we have

$$\arg\max_{c(S) \le b - c(Y)} f_2(S) = \arg\max_{c(S) \le b - c(Y)} \Big(f(S \cup Y) - \frac{m}{\lambda} f(Y) \Big) = \arg\max_{c(S) + c(Y) \le b} f(S \cup Y) = \mathrm{OPT} \setminus Y.$$

Case 1: If k' exists, it holds that $f_2(S'_{l'}) \ge \alpha[m' f_2(\mathrm{OPT} \setminus Y) - f_2(w')]$ by Lemma F.2. According to definition of Y, it follows that $f(\{u_1, w'\}) \le f(Y)$ and $f(\{u_2, w'\}) \le f(Y)$. Using these two inequalities, it follows that

$$\begin{aligned} f(Y) &\ge f(\{u_1, w'\}) + f(\{u_2, w'\}) - f(Y) \\ &= f(w' \mid u_1) + f(w' \mid u_2) + f(u_1) + f(u_2) - f(Y) \\ &\ge 2f(w' \mid Y) + f(u_1 \mid u_2) + f(u_2) - f(Y) \\ &= 2f(w' \mid Y) \end{aligned}$$

where the last two inequalities hold due to submodularity of f. Hence,

$$\begin{aligned} f(Y \cup S'_{l'}) &= \frac{m}{\lambda} f(Y) + f_2(S'_{l'}) \\ &\ge \frac{m}{\lambda} f(Y) + \alpha[m' f_2(\mathrm{OPT} \setminus Y) - f_2(w')] \\ &= \frac{m}{\lambda} f(Y) + \alpha m' [f(\mathrm{OPT}) - \frac{m}{\lambda} f(Y)] - \alpha [f(Y \cup \{w'\}) - \frac{m}{\lambda} f(Y)] \\ &= \alpha m' f(\mathrm{OPT}) + \left[\frac{m}{\lambda} - \alpha + \frac{\alpha m (1 - m')}{\lambda} \right] f(Y) - \alpha f(w' \mid Y) \\ &\ge \alpha m' f(\mathrm{OPT}) + \left[\frac{m}{\lambda} - \frac{3}{2}\alpha + \frac{\alpha m (1 - m')}{\lambda} \right] f(Y) \end{aligned} \tag{28}$$

where the first inequality holds since $f_2(S'_{l'}) \ge \alpha[m' f_2(\mathrm{OPT} \setminus Y) - f_2(w')]$ and the second inequality holds due to the fact that $f(Y) \ge 2f(w' \mid Y)$.

Next we investigate the function

$$t_2(m, m', \lambda) := \frac{m}{\lambda} - \frac{3}{2}\alpha + \frac{\alpha m (1 - m')}{\lambda}.$$

Using definition of $m' := m + \frac{m^2 - m}{\lambda - m}$, we could simplify $t_2(m, m', \lambda)$ into

$$t_2(m, m', \lambda) = \bar{t}_2(m, \lambda) = -\frac{\alpha m^3}{\lambda^2 - \lambda m} + \frac{\alpha m^2}{\lambda^2 - \lambda m} - \frac{3\alpha}{2} - \frac{\alpha m^2}{\lambda} + \frac{\alpha m}{\lambda} + \frac{m}{\lambda}.$$

By Lemma F.3, when $m \in [m_1, m_2]$, it holds that $\bar{t}_2(m, \lambda) \ge 0$. Using Inequality (28), it holds that

$$f(Y \cup S'_{l'}) \ge \alpha m' f(\mathrm{OPT}) + t_2(m, m', \lambda) \cdot f(Y) \ge \alpha m' f(\mathrm{OPT})$$

when $m \in [m_1, m_2]$. Otherwise, if $m \in [0, m_1) \cup (m_2, 1]$, it follows that $t_2(m, m', \lambda) \leq 0$. Using the fact that $f(Y) \leq f(\text{OPT})$, it follows that

$$f(Y \cup S'_{l'}) \geq \alpha m' f(\text{OPT}) + t_2(m, m', \lambda) \cdot f(Y) \geq (\alpha m' + t_2(m, m', \lambda)) \cdot f(\text{OPT}).$$

Case 2: If k' does not exist, since $\text{OPT} \setminus Y$ is an optimal solution for the residual instance $\mathcal{I}_2$, Lemma F.2 guarantees that $f_2(S'_{l'}) \geq m' f_2(\text{OPT} \setminus Y)$. Hence, it follows that

$$\begin{aligned}
f(Y \cup S'_{l'}) &= f(Y \cup S'_{l'}) - \frac{m}{\lambda} f(Y) + \frac{m}{\lambda} f(Y) \\
&= \frac{m}{\lambda} f(Y) + f_2(S'_{l'}) \\
&\geq \frac{m}{\lambda} f(Y) + m' f_2(\text{OPT} \setminus Y) \\
&= \frac{m}{\lambda} f(Y) + m'[f(\text{OPT}) - \frac{m}{\lambda} f(Y)] \\
&= m' f(\text{OPT}) + \frac{m}{\lambda}(1 - m') f(Y) \\
&\geq m' f(\text{OPT})
\end{aligned}$$

where the last inequality follows from the non-negativity of f and $1 - m'$.

□

G Analysis of One Set Enumeration Positive Greedy+Max

In this section, we give analysis of approximation ratio for One Set Enumeration Positive Greedy+Ma x (Algorithm 5).

Let f be a non-negative m-monotone submodular function with ground set V where m is monotonicity ratio. Let OPT be the optimal solution set with maximum function value, i.e., $f(\text{OPT}) = \max_{S \subseteq V, c(S) \leq b} f(S)$. Let r be an element of OPT with maximum cost, i.e., $r = \arg\max_{e \in \text{OPT}} c(e)$ where c is the cost function. Without loss of generality, we assume $f(\emptyset) = 0$, $f(\text{OPT}) = 1$ and $b = 1$. Given an initial set U, One Set Enumeration Positive Greedy+Max makes one call to Positive Greedy+Max algorithm (Algorithm 3). It constructs two sequences of partial solutions S_i and T_i for $0 \leq i \leq l$. It holds that $S_0 = U$. Let u_i be the element considered at each step with maximum density and added into the current solution set S_{i-1}, i.e., $S_i = S_{i-1} \cup \{u_i\}$. Let v_i be the selected element at each step with maximum marginal gain, i.e., $T_i = S_{i-1} \cup \{v_i\}$. We assume $T_l := T_{l-1}$ since in the last iteration round, partial solution set S_l is constructed and then the loop ends without constructing corresponding T_l. However, due to optimality of augmentation, it holds that $f(T_l) = f(T_{l-1}) \geq f(S_l)$.

Define the continuous extension of greedy solution sequence as $g(x)$ where $x \in [0, 1]$ as $g(x) := f(S_i) + \frac{x - c(S_i)}{c(u_{i+1})} \cdot f(u_{i+1} \mid S_i)$ where i is the largest solution index satisfying $c(S_i) \leq x$. It is easy to verify that $g(c(S_i)) = f(S_i)$. Let $g'(x)$ be the right derivative of $g(x)$ at x.

Lemma G.1. *Given a non-negative, m-monotone and submodular function f, for $t \in [0, 1 - c(r)]$, it follows that*

$$g'(t) \geq \frac{m - f(S_{i^*} \cup \{r\})}{1 - c(r)} \geq \frac{m - g(t) - f(r)}{1 - c(r)}$$

where i^* is the largest index satisfying $c(S_{i^*}) \leq t$.

Proof. For $t \in [0, 1 - c(r)]$, it follows that $1 - c(S_{i^*}) \geq 1 - t \geq c(r) = \max_{e \in \text{OPT}} c(e) \geq \max_{e \in \text{OPT} \setminus (\{r\} \cup S_{i^*})} c(e)$ since $t \geq c(S_{i^*})$.

According to Lemma F.1, by choosing $T = \text{OPT}$, $R = \{r\}$ and $b = 1$, it follows that

$$\begin{aligned}
g'(t) &= \frac{f(u_{i^*+1} \mid S_{i^*})}{c(u_{i^*+1})} \geq \frac{m f(\text{OPT}) - f(S_{i^*} \cup \{r\})}{1 - c(r)} = \frac{m - f(S_{i^*} \cup \{r\})}{1 - c(r)} \\
&\geq \frac{m f(\text{OPT}) - g(c(S_{i^*})) - f(r)}{1 - c(r)} = \frac{m - f(S_{i^*}) - f(r)}{1 - c(r)} \geq \frac{m - g(t) - f(r)}{1 - c(r)}.
\end{aligned}$$

since $\{r\} \subseteq \mathrm{OPT}$, $c(\mathrm{OPT}) \leq 1$ and $1 - c(S_{i^*}) \geq \max_{s \in \mathrm{OPT} \setminus (\{r\} \cup S_{i^*})} c(s)$ and $g(t) \geq f(S_{i^*})$.

□

Lemma G.2. *(Lemma 9 in (Feldman et al., 2023)) Given monotonicity ratio $m \in [0, 1]$, for every $y \in [0, 1/2]$, there is a unique value $z(y) \in [y, 1/2]$, satisfying*

$$\frac{y}{z(y)} - 1 = \ln\left(\frac{z(y)}{1-y}\right).$$

Moreover, $z(y)$ is a non-decreasing function of y and $z(y) \geq z(0) = 1/e$ and $z(y) \leq z(1/2) = 1/2$.

Proof. Here we only prove $z(1/2) = 1/2$. For other details, please refer to Lemma 9 in (Feldman et al., 2023).

Plugging $y = 1/2$ into the definition, it follows that

$$\frac{1}{2z(1/2)} - 1 = \ln\left(2z\left(1/2\right)\right).$$

It is easy to verify that $z(1/2) = 1/2$ satisfies this equation. Since $1/(2x) - 1$ decreases and $\ln(2x)$ increases with respect to variable x, the solution for equation $1/(2x) - 1 = \ln(2x)$ is unique, i.e., $1/2$. □

Lemma G.3. *If $f(r) \leq f(\mathrm{OPT})/2 = 1/2$, Positive Greedy+Max (Algorithm 3) achieves an approximation ratio of $m\left(1 - z\left(\frac{f(r)}{f(\mathrm{OPT})}\right)\right)$.*

Proof. Let $z_0 := z\left(\frac{f(r)}{f(\mathrm{OPT})}\right)$. Considering the function z from Lemma G.2, its domain is $[0, 1/2]$. Since $f(r) \leq f(\mathrm{OPT})/2$, it follows that z_0 is well defined.

Next we prove $f(T_l) \geq m\left(1 - z_0\right)$ by contradiction. Without loss of generality, we assume that $f(T_l) < m(1 - z_0)$. Otherwise, the approximation ratio is at least $m(1 - z_0)$.

Using Lemma G.1, for $t \in [0, 1 - c(r)]$ it holds that

$$g'(t) \geq \frac{m - f(S_{i^*} \cup \{r\})}{1 - c(r)} \geq \frac{m z_0}{1 - c(r)} \tag{29}$$

where the second inequality holds since $f(S_{i^*} \cup \{r\}) \leq f(T_l) \leq m(1 - z_0)$.

By Lemma G.1, it follows that

$$g'(t) \geq \frac{m - g(t) - f(r)}{1 - c(r)}.$$

By solving this differential inequality using integral over right derivative, we obtain a solution

$$g(t) \geq m - f(r) - C \cdot \exp\left(-\frac{t}{1 - c(r)}\right) \tag{30}$$

where C is some constant.

By setting constant $C := m(1 - f(r)) - (1 - m)e \cdot f(r)$, it follows that

$$\begin{aligned} -C \exp\left(-\frac{t}{1-c(r)}\right) &= -m(1 - f(r)) \exp\left(-\frac{t}{1-c(r)}\right) + (1 - m) f(r) \exp\left(1 - \frac{t}{1-c(r)}\right) \\ &\geq -m(1 - f(r)) \exp\left(-\frac{t}{1-c(r)}\right) + (1 - m) f(r) \end{aligned} \tag{31}$$

where the last inequality holds since $1 - t/(1 - c(r)) \geq 0$.

Combining Inequality (30) and Inequality (31), it follows that

$$g(t) \geq m(1 - f(r)) \left[1 - \exp\left(-\frac{t}{1 - c(r)}\right)\right].$$

Let $t_r := -[1 - c(r)] \cdot \ln\left(\frac{z_0}{1-f(r)}\right)$. By plugging $y = f(r) \in [0, 1/2]$ into the definition of z_0, we have

$$\ln\left(\frac{z_0}{1-f(r)}\right) = \frac{f(r)}{z_0} - 1 \in [2f(r) - 1, 0] \subseteq [-1, 0].$$

Using this result, we obtain that $t_r \in [0, 1 - c(r)]$.

Plugging t_r into this solution, we have

$$g(t_r) \geq m(1 - f(r)) \left[1 - \exp\left(-\frac{t_r}{1 - c(r)}\right)\right] = m(1 - f(r)) \left(1 - \frac{z_0}{1 - f(r)}\right).$$

Since

$$\frac{t_r z}{1 - c(r)} = -z \ln\left(\frac{z}{1 - f(r)}\right) = -z\left(\frac{f(r)}{z} - 1\right) = z - f(r)$$

where the second last equality is derived by the definition of $z := z(f(r))$, it holds that

$$\begin{aligned}
g(1 - c(r)) &= g(t_r) + \int_{t_r}^{1-c(r)} g'(t)dt \\
&\geq g(t_r) + (1 - c(r) - t_r) \cdot \frac{mz}{1 - c(r)} \\
&\geq m(1 - f(r)) \cdot \left(1 - \frac{z}{1 - f(r)}\right) + (1 - c(r) - t_r) \cdot \frac{mz}{1 - c(r)} \\
&= m(1 - f(r)) \cdot \left(1 - \frac{z}{1 - f(r)}\right) + mz - m(z - f(r)) \\
&= m(1 - f(r)) \cdot \left(1 - \frac{z}{1 - f(r)}\right) + mf(r) \\
&= m(1 - z)
\end{aligned}$$

where the first inequality holds due to Inequality (29) and the last inequality holds due to the fact that $mz - m(1-z) - m(-x+z) + m - x - \frac{z(m-x)}{1-x} = \frac{x(m-1)(x+z-1)}{1-x} \geq 0$ when $x \in [0, \frac{1}{2}]$ and $z \in [\frac{1}{e}, \frac{1}{2}]$. This result contradicts the assumption $f(T_l) < m(1-z)$ since $f(T_l) \geq f(S_l) \geq g(1-c(r)) \geq m(1-z)$ due to the monotonicity of partial solution sequence. Hence, it holds that $f(T_l) \geq m(1-z)$.

□

Using $\{w\}$ as the initial solution set, define the corresponding residual instance $\mathcal{I}' := \{V \setminus \{w\}, \bar{f}, c, b - c(w)\}$ where $\bar{f}(S) := f(S \cup \{w\}) - \frac{m}{\lambda} f(w)$. Let T_l' be the output of Positive Greedy+Max algorithm on residual instance $\mathcal{I}'$.

Lemma G.4. *If $f(w) \geq 1/4$, One Set Enumeration Greedy+Max (Algorithm 5) achieves an approximation ratio of $\frac{m'}{2} + \frac{m(2-m')}{8\lambda}$.*

Proof. According to Theorem C.2, Positive Greedy+Max algorithm achieves an approximation ratio of $m/2$. Thus, it follows that

$$\bar{f}(T_l') \geq \frac{m'}{2} \bar{f}(\mathrm{OPT}') = \frac{m'}{2} \bar{f}(\mathrm{OPT} \setminus \{w\}) = \frac{m'}{2}(f(\mathrm{OPT}) - \frac{m}{\lambda} f(w)),$$

which is equivalent to

$$f(T_l' \cup \{w\}) \geq \frac{m'}{2} + \frac{m(2-m')}{2\lambda} f(w) \geq \frac{m'}{2} + \frac{m(2-m')}{8\lambda}.$$

Since the function value of the output of One Set Enumeration Positive Greedy+Max (Algorithm 5) is no worse than $T_l' \cup \{w\}$, the approximation ratio is at least $\frac{m'}{2} + \frac{m(2-m')}{2\lambda} f(w)$ if $f(w) \geq 1/4$. □

Lemma G.5. *If $f(w) \leq 1/4$, One Set Enumeration Positive Greedy+Max (Algorithm 5) achieves an approximation ratio of $m'(1-e^{-1})$.*

Proof. By definition of r', w and submodularity of f, it holds that

$$\bar{f}(r') = f(\{r', w\}) - \frac{m}{\lambda} f(w) \leq f(r') + f(w) - \frac{m}{\lambda} f(w) \leq \Big(2 - \frac{m}{\lambda}\Big) f(w).$$

Hence, it follows that

$$\frac{\bar{f}(r')}{\bar{f}(\mathrm{OPT} \setminus \{w\})} \leq \frac{(2-\frac{m}{\lambda})f(w)}{f(\mathrm{OPT}) - \frac{m}{\lambda}f(w)} = \frac{(2-\frac{m}{\lambda})f(w)}{1-\frac{m}{\lambda}f(w)} \leq \frac{2-\frac{m}{\lambda}}{4-\frac{m}{\lambda}} \leq \frac{1}{2}$$

since function $f(x) = \frac{2-x}{4-x}$ decreases on interval $[0,1]$, $f(w) \in [0,1]$ and $\frac{m}{\lambda} \in [0,1)$.

Using Lemma G.3 by substituting $f(\mathrm{OPT})$ with $\bar{f}(\mathrm{OPT} \setminus \{w\})$ and $f(r)$ with $\bar{f}(r')$, we have

$$\begin{aligned} \frac{\bar{f}(T_l')}{\bar{f}(\mathrm{OPT} \setminus \{w\})} &= \frac{f(T_l' \cup \{w\}) - \frac{m}{\lambda}f(w)}{1-\frac{m}{\lambda}f(w)} \\ &\geq m'\left(1 - z\Big(\frac{\bar{f}(r')}{\bar{f}(\mathrm{OPT} \setminus \{w\})}\Big)\right) \geq m'\left(1 - z\Big(\frac{(2-\frac{m}{\lambda})f(w)}{1-\frac{m}{\lambda}f(w)}\Big)\right) \end{aligned}$$

where the last inequality holds due to the monotonicity of function z.

Using this result, we have

$$\begin{aligned} f(T_l' \cup \{w\}) &\geq \Big(1 - \frac{m}{\lambda}f(w)\Big)m' - \Big(1 - \frac{m}{\lambda}f(w)\Big) \cdot m' \cdot z\Big(\frac{(2-\frac{m}{\lambda})f(w)}{1-\frac{m}{\lambda}f(w)}\Big) + \frac{m}{\lambda}f(w) \\ &= m' + (1-m')\frac{m}{\lambda}f(w) - \Big(1 - \frac{m}{\lambda}f(w)\Big) \cdot m' \cdot z\Big(\frac{(2-\frac{m}{\lambda})f(w)}{1-\frac{m}{\lambda}f(w)}\Big). \end{aligned} \tag{32}$$

Let $x := \frac{(2-\frac{m}{\lambda})f(w)}{1-\frac{m}{\lambda}f(w)}$, it follows that $f(w) = \frac{\lambda x}{2\lambda - m + mx}$. Substituting $f(w)$ into the right hand side of Inequality (32), we have $f(T_l' \cup \{w\}) \geq q(x)$ where

$$\begin{aligned} q(x) &= m' + (1-m')\frac{m}{\lambda}\frac{\lambda x}{2\lambda - m + mx} - (1 - \frac{m}{\lambda}\frac{\lambda x}{2\lambda - m + mx}) \cdot m' \cdot z(x) \\ &= m' + (1-m')\frac{mx}{2\lambda - m + mx} - (1 - \frac{mx}{2\lambda - m + mx}) \cdot m' \cdot z(x) \\ &= 1 + (1 - \frac{mx}{2\lambda - m + mx}) \cdot (m' - 1 - m'z(x)) \\ &= 1 + \frac{2\lambda - m}{2\lambda - m + mx} \cdot (m' - 1 - m'z(x)). \end{aligned}$$

Since $0 \leq f(w) \leq \frac{1}{4}$ and $0 \leq \frac{m}{\lambda} < 1$, it follows that $x \in [0, \frac{2-\frac{m}{\lambda}}{4-\frac{m}{\lambda}}] \subseteq [0, \frac{1}{2})$, which is consistent with the domain of $z(y)$ where $y \in [0, \frac{1}{2}]$. Analyzing the lower bound of $q(x)$, we have

$$q(x) = 1 + \frac{2\lambda - m}{2\lambda - m + mx} \cdot (m' - 1 - m'z(x)) \geq 1 + \frac{2\lambda - m}{2\lambda - m + mx}(m'(1-e^{-1}) - 1) \geq m'(1-e^{-1})$$

where the first inequality holds since $\frac{2\lambda - m}{2\lambda - m + mx} > 0$ and $z(x) \geq e^{-1}$, the second inequality holds since $m'(1 - e^{-1}) - 1) < 0$ and $x \in [0, \frac{1}{2})$. Hence, it holds that $f(T_l' \cup \{w\}) \geq m'(1-e^{-1})$.

□

Combining Lemma G.4 and Lemma G.5, we have the following theorem:

Theorem C.4. *One Set Enumeration Positive Greedy+Max (Algorithm 5) achieves an approximation ratio of $\max_{\lambda \geq 1}\{h_3^\lambda(m)\}$ for maximizing a non-negative m-monotone submodular function with a knapsack constraint where*

$$h_3^\lambda(m) := \min\left\{\frac{m'}{2} + \frac{m(2-m')}{8\lambda}, m'\alpha\right\}.$$

H Analysis of Sample Greedy

Amanatidis et al. (2020) proposed Sample Greedy With P (line 9 in Algorithm 6) for non-monotone submodular maximization with a knapsack constraint, which has an approximation ratio of $(3+2\sqrt{2})^{-1}$.

Let f be a non-negative m-monotone submodular function with ground set V where m is monotonicity ratio. Let OPT be the optimal solution set with maximum function value, i.e., $f(\mathrm{OPT}) = \max_{S\subseteq V, c(S)\le b} f(S)$. Let p be the hyper-parameter of Sample Greedy With P, solution S and $v^* = \arg\max_{v\in V} f(v)$ be the solution sets constructed by Sample Greedy With P algorithm and $S_g := \arg\max_{T\in\{S,v^*\}} f(T)$ be the output of Sample Greedy With P algorithm.

Lemma H.1. *(Inequality (4) in (Amanatidis et al., 2020)) Let S be the solution set constructed by while loop by Sample Greedy With P (Procedure defined in Algorithm 6) and p be the hyper-parameter. It follows that*

$$\mathbb{E}[f(\mathrm{OPT}\cup S)] \le \max\{2, 1/p\}\cdot \mathbb{E}[f(S)] + f(v^*).$$

Theorem C.5. *Let $t_4(m) := \sqrt{(m-2)(m-1)} + m$ and $\delta \in (0, 1/5)$ be a small positive number. Sample Greedy (Algorithm 6) achieves an approximation ratio of $h_4(m)$ for maximizing a non-negative m-monotone submodular function with a knapsack constraint where*

$$h_4(m) := \begin{cases} (m+1)/6, & m\in[1/5,1] \\ -\dfrac{(t_4(m)-2)(t_4(m)-1)}{t_4(m)-m} - O(\delta), & \textit{otherwise.} \end{cases}$$

Proof. By setting $O = \mathrm{OPT}$ and $D = S$ in Lemma B.2, it follows that $\mathbb{E}[f(\mathrm{OPT}\cup S)] \ge (1-(1-m)\cdot p)\cdot f(\mathrm{OPT})$ since $\max_{u\in V}\Pr[u\in S]\le p$. By Lemma H.1, it follows that

$$\begin{aligned}(1-(1-m)\cdot p)\cdot f(\mathrm{OPT}) &\le \mathbb{E}[f(\mathrm{OPT}\cup S)] \\ &\le \max\{2,1/p\}\cdot\mathbb{E}[f(S)] + f(v^*) \\ &\le \max\{2,1/p\}\cdot\mathbb{E}[f(S_g)] + \mathbb{E}[f(S_g)],\end{aligned}$$

which is equivalent to

$$\mathbb{E}[f(S_g)] \ge \frac{1-(1-m)\cdot p}{\max\{2,1/p\}+1}\cdot f(\mathrm{OPT}).$$

Case 1: If $p\le 1/2$, $\max\{2,1/p\} = 1/p$. Then

$$\frac{1-(1-m)\cdot p}{\max\{2,1/p\}+1} = \frac{p(1-(1-m)\cdot p)}{1+p}.$$

Let $d_1(p) := \frac{p(1-p(1-m))}{1+p}$ Considering its derivative $d_1'(p) = \frac{p^2\cdot(m-1)+p(2m-2)+1}{(p+1)^2}$. whose numerator is defined as $d_2(p) := p^2(m-1) + p(2m-2) + 1$. Since $p\in[0,1/2]$ and the symmetric axis is $p=-1$, it follows that $d_2(p)$ decreases over $p\in[0,1/2]$ for all $m\in[0,1]$. Hence, it follows that $d_2(p)\in[d_2(1/2), d_2(0)] = [(5m-1)/4, 1]$. Here we observe that the monotonicity of $d_1(p)$ depends on the value of m.

When $m\ge 1/5$, $d_1(p)$ increases over $p\in[0,1/2]$, i.e., $\max d_1(p) = d_1(1/2) = \frac{m+1}{6}$. When $m<1/5$, $d_1(p)$ firstly increases and then decreases over $p\in[0,1/2]$. Denote $p_0 := \frac{-m-\sqrt{(m-2)(m-1)}+1}{m-1}$. Since $d_2(p_0)=0$, $d_1(p)$ achieves its maximum value when $p=p_0$. Hence, $\max d_1(p) = d_1(p_0) = -\frac{\left(m+\sqrt{(m-2)(m-1)}-2\right)\left(m+\sqrt{(m-2)(m-1)}-1\right)}{\sqrt{(m-2)(m-1)}} = -\frac{(t_4(m)-2)(t_4(m)-1)}{t_4(m)-m}$.

Hence, when $p\in[0,1/2]$, we choose $p=1/2$ when $m\ge 1/5$, in which case it holds that

$$\mathbb{E}[f(S_g)] \ge d_1(1/2)f(\mathrm{OPT}) = \frac{m+1}{6}f(\mathrm{OPT}).$$

Additionally, we choose $p=p_0$ when $m\le 1/5$. However, since the monotonicity ratio m is inaccessible, we cannot apply exact value of p_0 to the Sample Greedy With P algorithm. Instead, we construct an arithmetic sequence in $m \in [0, 1/5]$ with interval $\delta \in (0, 1/5)$. For each guess $\tilde{m}$ from the arithmetic sequence, we apply $p_0(\tilde{m})$ into Sample Greedy algorithm. As δ is small enough, the approximation ratio we obtain is not too far way from that we obain by applying exact value of $p_0(m)$. If fact, for any $m \in [0, 1/5]$, there exist $\tilde{m} \in [0, 1/5]$ such that $|\tilde{m} - m| \leq \delta$. Given $p_0(\tilde{m}) = \frac{-\tilde{m}-\sqrt{(\tilde{m}-2)(\tilde{m}-1)}+1}{\tilde{m}-1}$, it follows that

$$|p_0(\tilde{m}) - p_0| \leq \delta \sup_{m \in [0,1/5]} |p_0'(m)| \leq C\delta$$

for some constant C. Denote $x_1 := \min\{p_0(\tilde{m}), p_0\}$ and $x_2 := \max\{p_0(\tilde{m}), p_0\}$, it follows that

$$|d_1(p_0(\tilde{m})) - d_1(p_0)| \leq |p_0(\tilde{m}) - p_0| \sup_{x \in [x_1, x_2]} |d_1'(x)| \leq C\delta \sup_{x \in [x_1, x_2]} |d_1'(x)| \leq C_1\delta$$

for some constant C_1. Hence, we have

$$\mathbb{E}[f(S_g)] \geq d_1(p_0(\tilde{m}))f(\mathrm{OPT}) \geq (d_1(p_0) - O(\delta))f(\mathrm{OPT}).$$

Case 2: If $p \geq 1/2$, $\max\{2, 1/p\} = 2$. Then

$$\frac{1 - (1-m) \cdot p}{\max\{2, 1/p\} + 1} = \frac{1 - (1-m) \cdot p}{3} \leq \frac{m+1}{6},$$

which is the same result as Case 1.

Concluding Case 1 and Case 2, by calling Sample Greedy With P by setting $p = 1/2$ and $p = p_0(\tilde{m})$, the Sample Greedy algorithm achieves an approximation ratio of $h_4(m)$.

□

I Proof of Theorem 3.2

In this section, we prove Theorem 3.2, which we expand here into full version.

Theorem I.1. *For any constant $\varepsilon > 0$, no polynomial time algorithm can obtain an approximation ratio of*

$$\left(\min_{\alpha \in [0,1]} \frac{\max_{x \in [0,1]}\{\alpha(mx^2 + 2x - 2x^2) + 2(1-\alpha)(1 - e^{x-1})(1 - (1-m)x)\}}{\max\{1, 2(1-\alpha)\}}\right) - \frac{1-m}{512} + \varepsilon$$

for the problem of maximizing a non-negative m-monotone submodular function with a knapsack constraint.

For simplicity, denote

$$\rho_{\mathrm{card}}(m) := \min_{\alpha \in [0,1]} \frac{\max_{x \in [0,1]}\{\alpha(mx^2 + 2x - 2x^2) + 2(1-\alpha)(1 - e^{x-1})(1 - (1-m)x)\}}{\max\{1, 2(1-\alpha)\}}.$$

We prove the theorem by firstly constructing a problem instance under a knapsack constraint. Let V denote the ground set with size n such that $V = \{a, b\} \cup \{a_i, b_i \mid i \in [r]\}$ where $r = (n-2)/2$. The objective $f : 2^V \to \mathbb{R}_{\geq 0}$ is exactly the function from Section 4.3 in (Mualem and Feldman, 2022), i.e., $f(S) := \alpha f_1(S) + (1-\alpha)(f_2(S) + f_3(S))$, where

$$\begin{aligned}
f_1(S) &= m \cdot \mathbf{1}\{S \cap \{a, b\} \neq \emptyset\} + (1-m) \cdot (|S \cap \{a, b\}| \bmod 2), \\
f_2(S) &= \mathbf{1}\{S \cap \{a_i \mid i \in [r]\} \neq \emptyset\} \cdot (1 - (1-m) \cdot \mathbf{1}\{a \in S\}), \\
f_3(S) &= \mathbf{1}\{S \cap \{b_i \mid i \in [r]\} \neq \emptyset\} \cdot (1 - (1-m) \cdot \mathbf{1}\{b \in S\}).
\end{aligned}$$

We consider maximizing the objective $f(S)$ with a knapsack constraint $c(S) \leq b$ where c is a nonnegative modular cost function and b is a real-valued budget. We fix an arbitrary constant $c \geq 1$ and analyze algorithms making at most n^c value-oracle queries.

Partition V into T hidden types $V = \bigsqcup_{t=1}^{T} V_t$ with integer sizes $k_t := |V_t|$ satisfying $\sum_t k_t = n$ and $k_t \in \{\lfloor n/T \rfloor, \lceil n/T \rceil\}$. Assign each type a pairwise-distinct cost label $\tilde{c}_t > 0$ (e.g., $\tilde{c}_t := 2^t$), and set each item's cost by $c(v) := \tilde{c}_{\sigma(t)}$ for $v \in V_t$ where σ is a permutation of $[T]$. For any subset $S \subseteq V$, its type–count vector is $x(S) = (x_1, \ldots, x_T)$ with $x_t := |S \cap V_t|$. Let $\mathcal{P} \subseteq [T]$ be a hidden planted block of type indices of size $\lfloor T/2 \rfloor$.

Define two reference mixes $b^{(0)}, b^{(1)} \in \mathbb{Z}_{\geq 0}^{T}$ for balanced-shifted and planted-shifted problem instance distributions and let the common budget be $b := \sum_{t=1}^{T} \tilde{c}_t b_t^{(0)} = \sum_{t=1}^{T} \tilde{c}_t b_t^{(1)}$. We define two problem instance distributions over all problem instances denoted by (f, c, b) that share the same f and the same multiset of item costs as follows:

- $\mathcal{D}_0$ (balanced reference): draw the hidden partition $\{V_t\}$ and $\mathcal{P}$ uniformly. Draw σ uniformly at random and set $c(v) = \tilde{c}_{\sigma(t)}$ for $v \in V_t$. Use budget b as above ;
- $\mathcal{D}_1$ (planted reference): same draw of $\{V_t\}$ and $\mathcal{P}$. Choose σ so that more low $\tilde{c}_t$ are mapped to $t \in \mathcal{P}$ and more high $\tilde{c}_t$ to $t \notin \mathcal{P}$, while keeping the multiset $\{c(v)\}$ identical and the budget b equal to the same expression with $b^{(1)}$.

In both distributions, the objective f and the multiset of item costs $\{c(v)\}$ have the same law. They differ only in the hidden correlation between costs and the planted block $\mathcal{P}$. Let $\Phi(x)$ denote the mean value as a function of type counts $x = (x_1, \ldots, x_T)$ with $x_t := |S \cap V_t|$, i.e., $\Phi(x) := \mathbb{E}[f(S) \mid x]$.

Lemma I.2. *There exist two constants $C, c > 0$ independent of n such that for any fixed subset $S \subseteq V$ it holds that*

$$\mathbb{P}\Big(|f(S) - \Phi(x(S))| > C\sqrt{n \log n} \Big) \leq n^{-c}.$$

Proof. By construction, f is obtained by aggregating a finite number of independent, bounded primitive random variables whose expectations depend only on the type-count vector $x(S)$. Let $\{Z_j(S)\}_{j=1}^{M}$ denote these primitives after the standard normalization that ensures $\Phi(x^\star) = 1$ (i.e., each raw indicator is scaled by a factor $\beta = \Theta(1/n)$ so the overall value is $\Theta(1)$). Then it follows that $f(S) = \sum_{j=1}^{M} Z_j(S)$ and $\Phi(x(S)) = \mathbb{E}\big[f(S)\big] = \sum_{j=1}^{M} \mathbb{E}[Z_j(S)]$. Since $|Z_j(S) - \mathbb{E}Z_j(S)| \leq \beta = \Theta(1/n)$, it holds that each $Z_j(S)$ is independent and bounded. Moreover, since $M = \Theta(n^2)$ (a constant number of components, each summing over $O(n^2)$ pairs), it follows that

$$\mathrm{Var}\,\big(f(S)\big) = \sum_{j=1}^{M} \mathrm{Var}(Z_j(S)) \leq \sum_{j=1}^{M} \mathbb{E}\big[(Z_j(S) - \mathbb{E}Z_j(S))^2\big] \leq M \cdot \beta^2 = \Theta(n).$$

Applying Bernstein's inequality to the centered sum $f(S) - \Phi(x(S))$ with per-summand bound $b := \beta = \Theta(1/n)$ and variance proxy $\sigma^2 := \Theta(n)$, we have

$$\mathbb{P}\big(|f(S) - \Phi(x(S))| > t\big) \leq \; 2\exp\!\left(-\frac{t^2}{2\sigma^2 + \frac{2}{3}bt}\right)$$

for any $t > 0$.

Choose $t = C\sqrt{n \log n}$ with C large enough. Since $bt = \Theta\big((1/n) \cdot \sqrt{n \log n}\big) = \Theta(\sqrt{\log n / n}) \ll \sigma^2 = \Theta(n)$, the denominator is $2\sigma^2(1 + o(1)) = \Theta(n)$, and the exponent is $-\Theta\big((n \log n)/n\big) = -\Theta(\log n)$. Thus $\mathbb{P}(|f(S) - \Phi(x(S))| > C\sqrt{n \log n}) \leq n^{-c}$ for some constant $c > 0$. ☐

By Lemma I.2, it holds that $f(S) = \Phi(x(S)) \pm \tilde{O}(\sqrt{n})$ with high probability.

Lemma I.3. *For the knapsack instance distributions $\mathcal{D}_0, \mathcal{D}_1$ defined above, it holds that $\mathbb{E}_{\mathcal{D}_0}[f(\mathrm{OPT})] \leq \rho_{\mathrm{card}}(m) \cdot \mathbb{E}_{\mathcal{D}_1}[f(\mathrm{OPT})] + o(1)$.*

Proof. We first prove how the budget b pins the feasible type counts under $\mathcal{D}_0$ and guarantees feasibility of the planted mix under $\mathcal{D}_1$.

Fix a tiny parameter $\zeta \in (0, 1/10)$ and set the type-cost labels $\tilde{c}_t := (1 + \zeta)2^t$ for $t \in [T]$. Let $b^{(0)}, b^{(1)} \in \mathbb{Z}_{\geq 0}^{T}$ be the reference mixes and define $b := \sum_{t=1}^{T} \tilde{c}_t$ and $b_t^{(0)} = \sum_{t=1}^{T} \tilde{c}_t b_t^{(1)}$. Consider any $x \in \mathbb{Z}_{\geq 0}^{T}$ and let $\Delta := x - b^{(0)}$. If $\Delta \neq 0$ and $t^\star$ is the largest index with $\Delta_{t^\star} > 0$, then $\sum_{t=1}^{T} \tilde{c}_t \Delta_t \geq \tilde{c}_{t^\star} - \sum_{t < t^\star} \tilde{c}_t |\Delta_t| \geq (1 + \zeta)2^{t^\star} - \sum_{t < t^\star}(1 +$

$\zeta)2^t|\Delta_t|$. Since $|\Delta_t| \le b_t^{(0)} \le |V_t| \le n$ and $\sum_{t<t^\star} 2^t \le 2^{t^\star}-1$, we obtain $\sum_{t=1}^{T} \tilde{c}_t \Delta_t \ge (1+\zeta)2^{t^\star} - (1+\zeta)(2^{t^\star}-1) \ge \zeta 2^{t^\star} - (1+\zeta) \ge 1$ for all large $t^\star$ (and hence large n), because $\zeta 2^{t^\star}$ dominates the additive constant. Therefore,

$$\sum_{t=1}^{T} \tilde{c}_t x_t > \sum_{t=1}^{T} \tilde{c}_t b_t^{(0)} = b \qquad \text{whenever } x_{t^\star} > b_{t^\star}^{(0)} \text{ for the maximal } t^\star. \tag{33}$$

In particular, under the cost assignment of $\mathcal{D}_0$ which has uniform random permutation σ of the labels $\tilde{c}_t$ across types, any feasible x must satisfy the coordinate-wise inequalities

$$x_t \le b_t^{(0)} \qquad \forall t \in [T], \tag{34}$$

because otherwise (33) is violated and the budget is already tight at $b^{(0)}$ along the most expensive active coordinate.

By (34), the $\mathcal{D}_0$ optimum is bounded by $f(\mathrm{OPT}) \le \max\{\Phi(x) : 0 \le x_t \le b_t^{(0)} \forall t\} + \tilde{O}(\sqrt{n})$, where we inserted a $\tilde{O}(\sqrt{n})$ concentration slack (Lemma I.2) to pass from f to Φ. Let $\bar{x}^{(0)}$ be the maximizer of this box-constrained mean-field problem. Since the objective is invariant under permutations of types that preserve $\mathcal{P}$, averaging $\bar{x}^{(0)}$ over those permutations does not decrease Φ by convexity of expectation. Hence there exists an optimal solution that is symmetric within the planted and non-planted blocks and satisfies the knapsack bound with equality by the cost separation. Therefore, we have

$$\mathbb{E}_{\mathcal{D}_0}[f(\mathrm{OPT})] \le \mathbb{E}\big[\Phi(x^{\mathrm{sym}})\big] + o(1) = \rho_{\mathrm{card}}(m) \cdot \mathbb{E}\big[\Phi(x^\star)\big] + o(1),$$

where x^{sym} denotes the best symmetric mix, $x^\star$ the planted true maximizer, and $\Phi(x^{\mathrm{sym}}) = \rho_{\mathrm{card}}(m)\Phi(x^\star)$.

Under $\mathcal{D}_1$, the permutation σ aligns smaller costs with types in $\mathcal{P}$ and larger costs with types outside $\mathcal{P}$, and b is defined so that $b^{(1)}$ saturates the budget. By construction, $x^\star$ is coordinate-wise $\le b^{(1)}$ and thus feasible. Hence $\mathbb{E}_{\mathcal{D}_1}[f(\mathrm{OPT})] \ge \mathbb{E}\big[\Phi(x^\star)\big]$. Combining the two inequalities yields the claim. □

Fix a deterministic algorithm that uses at most $q \le n^c$ value queries. Let $T := \lceil n^{2c+3} \rceil$ and draw the hidden type partition $V = \bigsqcup_{t=1}^{T} V_t$ by distributing the n items so that $k_t := |V_t| \in \{0,1\}$ and $\sum_t k_t = n$ (hence exactly n types are nonempty). Let $\mathcal{P} \subseteq [T]$ be a uniformly random planted block of size $\lfloor T/2 \rfloor$.

Lemma I.4. *For each queried set $S^{(i)}$, write its mix $x^{(i)} \in \{0,1\}^T$ with $x_t^{(i)} = |S^{(i)} \cap V_t| \in \{0,1\}$. Then, with probability $1 - o(1)$ over the random partition and $\mathcal{P}$, we have*

$$\frac{\sum_{t\in\mathcal{P}} x_t^{(i)} - \sum_{t\notin\mathcal{P}} x_t^{(i)}}{\sum_{t=1}^{T} x_t^{(i)}} \le \gamma = 1/8$$

for all $i \in [q]$. Equivalently, along mix directions, $\|\Pi(x^{(i)} - x^\star)\|_2 \ge \gamma\Theta(n)$.

Proof. Condition on the algorithm's adaptivity, $S^{(i)}$ may depend on past answers. However, given the partition into singletons V_t the indicator vector $x^{(i)}$ is a $\{0,1\}$ vector with exactly $|S^{(i)}|$ ones placed on the chosen type indices.

Let $\mathbf{1}_{\mathcal{P}} \in \{0,1\}^T$ be the indicator of the planted block. For fixed $x^{(i)}$ and random $\mathcal{P}$ (uniform of size $\lfloor T/2 \rfloor$), we have $\langle x^{(i)}, \mathbf{1}_{\mathcal{P}} \rangle \sim \mathrm{Hypergeom}\big(T, \lfloor T/2 \rfloor, |S^{(i)}|\big)$, implying that its mean is $|S^{(i)}|/2$. By Hoeffding's bound for sampling without replacement, it follows that $\mathbb{P}\big[\big|\langle x^{(i)}, \mathbf{1}_{\mathcal{P}} \rangle - |S^{(i)}|/2\big| \ge \lambda\big] \le 2\exp\Big(-\frac{2\lambda^2}{|S^{(i)}|}\Big)$ for any $\lambda > 0$. By setting λ as $\lambda_i := \sqrt{\frac{|S^{(i)}|}{2}\log(4qn^2)}$, it follows that $\mathbb{P}\big[\big|\langle x^{(i)}, \mathbf{1}_{\mathcal{P}} \rangle - |S^{(i)}|/2\big| \ge \lambda_i\big] \le \frac{1}{2qn^2}$.

A union bound over $i \in [q]$ yields $\forall i \le q :\ \frac{\big|\langle x^{(i)}, \mathbf{1}_{\mathcal{P}} \rangle - |S^{(i)}|/2\big|}{|S^{(i)}|} \le \sqrt{\frac{\log(4qn^2)}{2|S^{(i)}|}} \le \sqrt{\frac{\log(4n^{c+2})}{2n}} \le \frac{1}{10}$ with probability at least $1 - 1/(2n^2)$ for all large n as $|S^{(i)}| \le n$. Hence, for all i, we have $\frac{\sum_{t\in\mathcal{P}} x_t^{(i)} - \sum_{t\notin\mathcal{P}} x_t^{(i)}}{\sum_t x_t^{(i)}} = \frac{2\langle x^{(i)}, \mathbf{1}_{\mathcal{P}} \rangle - |S^{(i)}|}{|S^{(i)}|} \le \frac{2\lambda_i}{|S^{(i)}|} \le \frac{1}{5} < \frac{1}{8}$ for all sufficiently large n.

Finally, the projection-distance claim follows because the normalized planted direction $u := \frac{1}{\sqrt{n}}(\mathbf{1}_{\mathcal{P}} - \mathbf{1}_{[T]\setminus\mathcal{P}})$ has constant correlation with $x^\star$ while $\langle x^{(i)}, u \rangle$ is bounded by $\gamma\|x^{(i)}\|_1/\sqrt{n}$. By Pythagoras the component of $x^{(i)} - x^\star$ orthogonal to budget direction has ℓ_2-norm $\Omega(n)$. □

Definition I.5 (Total Variation). *For two probability measures P, Q on a common measurable space, their total variation (TV) distance is defined as*

$$\mathrm{TV}(P,Q) := \sup_A |P(A) - Q(A)| = \frac{1}{2}\int |dP - dQ|.$$

A query is a subset $S \subseteq V$ into the value oracle f and a corresponding answer is the function value $f(S)$ of the subset S. We apply the concept of total variation to the distributions of oracle transcripts, which is defined as random vectors consisting of queries and answers. By Pinsker's inequality, it holds that $\mathrm{TV}(P,Q) \le \sqrt{\frac{1}{2}\mathrm{KL}(P\|Q)}$, where KL is Kullback-Leibler divergence (Cover and Thomas, 2006).

Lemma I.6. *Fix any deterministic algorithm issuing at most $q \le n^c$ value queries. Let Tr_b be its oracle transcript distribution on instances drawn from $\mathcal{D}_b$. Then $\mathrm{TV}(\mathsf{Tr}_0, \mathsf{Tr}_1) = o(1)$.*

Proof. Let $S^{(1)}, \ldots, S^{(q)}$ be the random queries and $A_i := f(S^{(i)})$ the answers. We bound $\mathrm{KL}(\mathsf{Tr}_0\|\mathsf{Tr}_1)$ by Pinsker's inequality.

By Lemma I.2, for each fixed S, it holds that $A_i = \Phi(x(S)) + Z_i$ with $\mathbb{P}(|Z_i| > C\sqrt{n\log n}) \le n^{-10}$. In particular, A_i is sub-Gaussian with variance proxy $\sigma^2 = \Theta(n)$ (primitive indicators are $O(n^2)$ many, each scaled by $O(1/n)$).

Let $\mu_i^{(b)} := \mathbb{E}_{\mathcal{D}_b}[A_i \mid \mathcal{H}_{i-1}]$, where $\mathcal{H}_{i-1}$ is the transcript history up to $i-1$. The only difference between $\mathcal{D}_0$ and $\mathcal{D}_1$ is the hidden cost-type alignment. By Lemma I.4, the queried mixes $x^{(i)}$ have normalized planted correlation at most $\gamma = 1/8$. Since Φ is 1-Lipschitz in that correlation after the standard normalization $\Phi(x^\star) = \Theta(1)$ since the coefficients of the type interaction terms are $O(1)$, we get $\Delta_i := |\mu_i^{(1)} - \mu_i^{(0)}| \le C_1$ for an absolute constant C_1 independent of n.

Conditioned on $\mathcal{H}_{i-1}$, the one-step KL between A_i under $\mathcal{D}_0$ vs. $\mathcal{D}_1$ is upper bounded by the Gaussian proxy $\mathrm{KL}_i \le C_2\Delta_i^2/\sigma^2 \le C_3/n$. By the chain rule for KL divergence (adaptivity only decreases the sum of conditional KLs), it follows that $\mathrm{KL}(\mathsf{Tr}_0\|\mathsf{Tr}_1) \le \sum_{i=1}^q \mathbb{E}[\mathrm{KL}_i] \le q \cdot C_3/n = O(n^{c-1}) = o(1)$. Pinsker's inequality gives $\mathrm{TV}(\mathsf{Tr}_0, \mathsf{Tr}_1) \le \sqrt{\frac{1}{2}\mathrm{KL}(\mathsf{Tr}_0\|\mathsf{Tr}_1)} = o(1)$. ☐

Lemma I.7. *There exists a constant $c_0 \in (0,1)$ such that for all feasible mixes x it holds that $\Phi(x) \le \Phi(x^\star) - (1-m)c_0\big\|\Pi(x-x^\star)\big\|_2^2$, where $x^\star$ is a planted maximizer of Φ, and Π is the orthogonal projector onto the subspace of mix-preserving directions. i.e., orthogonal to the budget direction.*

Proof. After symmetrization within the planted and non-planted meta-types, Φ is a smooth quadratic function of x, with Hessian $H := \nabla^2\Phi$ that is block-constant on $\mathcal{P}\times\mathcal{P}$, $\mathcal{P}\times\mathcal{P}^c$, $\mathcal{P}^c\times\mathcal{P}$, and $\mathcal{P}^c\times\mathcal{P}^c$. The budget direction w which is proportional to the all-ones vector weighted by costs is an eigenvector corresponding to the knapsack modular part. Along w, Φ is affine once budget is fixed.

Restrict H to the orthogonal complement of w. In this mix-preserving subspace, the quadratic form coincides with the cross-type interaction term whose coefficients are controlled by the partial-monotonicity parameter m. The imbalance between positive and negative expected marginals is $\Theta(1-m)$. A standard block-matrix eigenvalue bound (Gershgorin or explicit diagonalization of a 2×2 reduced form after averaging within blocks) yields that the smallest eigenvalue of $-H$ on that subspace is at least $c_0(1-m)$ for an absolute $c_0 \in (0,1)$. Hence, for any v orthogonal to w, it follows that $\Phi(x^\star + v) \le \Phi(x^\star) - c_0(1-m)\|v\|_2^2$.

Writing $x = x^\star + u + v$ with $u \parallel w$, $v = \Pi(x - x^\star)$, and using that $\Phi(x^\star + u) \le \Phi(x^\star)$ by optimality of $x^\star$ under the fixed budget, the lemma holds. ☐

Lemma I.8. *It holds that $\Phi(x) \le \Phi(x^\star) - (1-m)/512$.*

Proof. By Lemma I.7, it holds that $\Phi(x) \le \Phi(x^\star) - (1-m)c_0\|\Pi(x-x^\star)\|_2^2$. By Lemma I.4, for all transcript queries we have $\|\Pi(x-x^\star)\|_2 \ge \gamma\Theta(n)$. After the standard normalization of primitive weights that enforces $\Phi(x^\star) = \Theta(1)$, the scale factor turns $\Theta(n^2)$ in the norm square into an $O(1)$ constant. Absorbing slack into $\mu(m) \ge (1-m)/4$ yields the displayed bound. Hence, there exists a constant $c_0 \in (0,1)$ and $\gamma \in (0, 1/8]$ such that it holds that $\Phi(x) \le \Phi(x^\star) - \mu(m)\gamma^2 c_0$ with $\mu(m) \ge \frac{1-m}{4}$ for any queried mix x with $\|\Pi(x-x^\star)\|_2 \ge \gamma\Theta(n)$.

Choosing $\gamma = 1/8$ and $c_0 \geq 1/2$, we have $\Phi(x) \leq \Phi(x^\star) - \mu(m)\gamma^2 c_0 \leq \Phi(x^\star) - (1-m)/512$.

□

Proof of Theorem I.1. If there exists a deterministic algorithm $\mathcal{A}$ making at most $q \leq n^c$ value-oracle queries that achieves an approximation ratio of $\rho_{\text{card}}(m) - \frac{1-m}{512} + \varepsilon$ for some constant $\varepsilon > 0$. We evaluate $\mathcal{A}$ on the distributions $\mathcal{D}_0$ and $\mathcal{D}_1$ constructed in Appendix I.

By Lemma I.3, it follows that $\mathbb{E}_{\mathcal{D}_0}[f(\text{OPT})] \leq \rho_{\text{card}}(m)\mathbb{E}_{\mathcal{D}_1}[f(\text{OPT})] + o(1)$. Under distribution $\mathcal{D}_1$, let $\{S^{(i)}\}_{i=1}^q$ be the sequence of queries. By Lemma I.4, the algorithm's queries are statistically "blind" to the planted alignment $\mathcal{P}$, ensuring that the type-count vectors $x^{(i)}$ satisfy $\|\Pi(x^{(i)} - x^\star)\|_2 \geq \gamma\Theta(n)$ with high probability. By Lemma I.8, it holds that $\Phi(x^{(i)}) \leq \Phi(x^\star) - \frac{1-m}{512}$. By the concentration bound in Lemma I.2, the actual oracle answers $f(S^{(i)})$ satisfy $f(S^{(i)}) = \Phi(x^{(i)}) \pm \tilde{O}(\sqrt{n})$. Thus, in $\mathcal{D}_1$, the algorithm is constrained to values strictly below $\mathbb{E}_{\mathcal{D}_1}[f(\text{OPT})] - \frac{1-m}{512}$.

Define the success event E as the event that the algorithm returns a set S with value, i.e., $f(S) \geq \left(\rho_{\text{card}}(m) - \frac{1-m}{512} + \varepsilon\right)\mathbb{E}_{\mathcal{D}_1}[f(\text{OPT})]$. Under $\mathcal{D}_0$, since $f(\text{OPT}) \approx \rho_{\text{card}}(m)\mathbb{E}_{\mathcal{D}_1}[f(\text{OPT})]$, the algorithm can satisfy this condition by finding a local optimum. However, under $\mathcal{D}_1$, the curvature gap ensures that E occurs with only negligible probability, as the required value is unattainable for queries distant from $x^\star$. By Lemma I.6, the total variation distance between the oracle transcripts is $\text{TV}(\mathsf{Tr}_0, \mathsf{Tr}_1) = o(1)$. This implies that $|\mathbb{P}_{\mathcal{D}_0}(E) - \mathbb{P}_{\mathcal{D}_1}(E)| \leq o(1)$. If $\mathcal{A}$ could achieve an approximation ratio of $\rho_{\text{card}}(m) - \frac{1-m}{512} + \varepsilon$, it would distinguish $\mathcal{D}_0$ from $\mathcal{D}_1$ with constant advantage, contradicting the indistinguishability framework of (Le Cam, 1986; Yao, 1977). This concludes the proof. □

Comparison with the hardness result for cardinality constraint. Our hardness result departs from the strong-symmetry paradigm in the cardinality constraint setting. Specifically, for a strongly symmetric instance with symmetry gap $\rho_{\text{card}}(m)$, no polynomial-time algorithm achieves better than $(1+o(1))\rho_{\text{card}}(m)$-approximation Mualem and Feldman (2022). However, a knapsack constraint with arbitrary costs destroys strong symmetry because item costs can be chosen pairwise distinct across hidden types, so permutations preserving feasibility act only within types. To handle this non-symmetric regime, we replace symmetry-gap arguments with an information-theoretic indistinguishability framework derived from (Le Cam, 1986; Yao, 1977): we construct two carefully matched distributions of instances $\mathcal{D}_0, \mathcal{D}_1$ that share the same objective f and the same multiset of costs, yet differ in an unobservable planted alignment between costs and types. We then bound the total-variation distance between oracle transcripts via KL+Pinsker, and convert the induced, algorithm-enforced separation in type mixes into a quantitative value loss using a curvature argument in the type-mix space. This toolkit applies whenever global symmetry is absent but one can (i) hide planted structure and (ii) certify uniform curvature of the mean-field potential Φ.

Within this framework we establish a strict separation from the cardinality barrier for every $m < 1$. Specializing to the objective f, we prove that non-monotone submodular maximization under a single knapsack with arbitrary weights is strictly harder than the cardinality-constrained version, i.e., no polynomial time algorithm achieves an approximation ratio of $\rho_{\text{card}}(m) - (1-m)/512 + \varepsilon$ for any $\varepsilon > 0$.

J Proof of Lemma 4.2

Lemma 4.2. *The objective function f for the influence-and-exploit marketing problem is 1) non-negative $2(1-\beta)$-monotone submodular when $\beta \in [1/2, 1]$; 2) non-negative monotone submodular when $\beta \in [0, 1/2]$.*

Proof. Since the weights are non-negative for any $S \subseteq V$ and $\beta \in [0, 1]$, it holds that f is non-negative.

For any $S \subseteq T \subseteq V$ and $k \in V \setminus T$, it follows that

$$\begin{aligned} f(k \mid S) &= \sum_{i \in V} w_{ik} - \beta(\sum_{i \in S} w_{ik} + \sum_{j \in S} w_{kj} + w_{kk}) - \beta w_{kk} \\ &= \sum_{i \in V} w_{ik} - 2\beta(\sum_{i \in S} w_{ik} + w_{kk}) \\ &\geq \sum_{i \in V} w_{ik} - 2\beta(\sum_{i \in T} w_{ik} + w_{kk}) \\ &= f(k \mid T) \end{aligned}$$

where the first inequality holds since $S \subseteq T$ and weights are non-negative. Hence, f is submodular.

When $\beta \in [0, \frac{1}{2}]$, we consider the marginal gain $f(k \mid S)$ where $k \in V \setminus S$ and $S \subseteq V$:

$$\begin{aligned} f(k \mid S) &= \sum_{i \in V} w_{ik} - 2\beta(\sum_{i \in S} w_{ik} + w_{kk}) \\ &\geq \sum_{i \in V} w_{ik} - (\sum_{i \in S} w_{ik} + w_{kk}) \\ &= \sum_{i \in V} w_{ik} - \sum_{i \in S \cup \{k\}} w_{ik} \geq 0 \end{aligned}$$

where the first inequality holds since $\beta \leq \frac{1}{2}$. Hence, f is monotone.

When $\beta \in [\frac{1}{2}, 1]$, we consider the ratio between $f(S)$ and $f(T)$ where $S \subseteq T \subseteq V$.

Since

$$\begin{aligned} &\sum_{i \in V} \sum_{j \in T \setminus S} w_{ij} - \sum_{i \in S} \sum_{j \in T \setminus S} w_{ij} - \beta \sum_{i \in T \setminus S} \sum_{j \in T \setminus S} w_{ij} + (\beta - 1) \sum_{i \in T \setminus S} w_{ii} \\ &= \sum_{i \in V \setminus S} \sum_{j \in T \setminus S} w_{ij} - \sum_{i \in T \setminus S} w_{ii} - \beta \sum_{i \in T \setminus S} \sum_{j \in T \setminus (S \cup \{i\})} w_{ij} \\ &\geq \sum_{i \in T \setminus S} \sum_{j \in T \setminus (S \cup \{i\})} w_{ij} - \beta \sum_{i \in T \setminus S} \sum_{j \in T \setminus (S \cup \{i\})} w_{ij} \geq 0, \end{aligned}$$

it follows that

$$\sum_{i \in V} \sum_{j \in T \setminus S} w_{ij} - \beta \sum_{i \in T \setminus S} \sum_{j \in T \setminus S} w_{ij} - \beta \sum_{i \in T \setminus S} w_{ii} \geq \sum_{i \in S} \sum_{j \in T \setminus S} w_{ij} + (1 - 2\beta) \sum_{i \in T \setminus S} w_{ii}. \tag{35}$$

Since $2\beta - 1 \geq 0$, it holds that

$$f(S) = (2 - 2\beta) f(S) + (2\beta - 1) f(S) \geq (2 - 2\beta) f(S) + (2\beta - 1) \sum_{i \in V \setminus S} \sum_{j \in S \cup \{i\}} w_{ij},$$

where the first inequality holds since $\beta \leq 1$ and the second inequality holds since $S \subseteq T$. This could be rewritten as

$$(1 - 2\beta) \sum_{i \in V \setminus S} \sum_{j \in S \cup \{i\}} w_{ij} \geq (1 - 2\beta) f(S). \tag{36}$$

Using Inequality (35) and Inequality (36), it holds that

$$\begin{aligned} f(T)-f(S) &= \sum_{i\in V}\sum_{j\in T\setminus S} w_{ij} - \beta(2\sum_{i\in S}\sum_{j\in T\setminus S} w_{ij} + \sum_{i\in T\setminus S}\sum_{j\in T\setminus S} w_{ij} + \sum_{i\in T\setminus S} w_{ii}) \\ &\geq \sum_{i\in S}\sum_{j\in T\setminus S} w_{ij} + (1-2\beta)\sum_{i\in T\setminus S} w_{ii} - 2\beta\sum_{i\in S}\sum_{j\in T\setminus S} w_{ij} \\ &= (1-2\beta)(\sum_{i\in S}\sum_{j\in T\setminus S} w_{ij} + \sum_{j\in T\setminus S} w_{ij}) \\ &\geq (1-2\beta)f(S), \end{aligned}$$

which is equivalent to $f(T) \geq 2(1-\beta)f(S)$. Hence, the monotonicity ratio of f when $\beta \in [\frac{1}{2}, 1]$ is $2(1-\beta)$. $\square$

K Additional Experiments

Additionally, we summarize the average and standard deviation of approximation ratios in Table 3. We observe that when the budget ratio is set at 0.1 and 0.15, the average empirical approximation ratios of the Deterministic Linear Approximation and Randomized Linear Approximation algorithms are larger than other algorithms. However, as the budget ratio increases to $0.2, 0.3, 0.4, 0.5$, One Set Enumeration Positive Greedy+Max and Sample Greedy demonstrate superior average empirical approximation ratios. This phenomenon could be attributed to the property of the thresholding technique used in Deterministic Linear Approximation and Randomized Linear Approximation algorithms. When the budget ratio becomes larger, the thresholding technique could include more elements since the budget is larger. However, in the non-monotone case, aggressively including more elements could sabotage the function value of final solution set.

Budget Ratio	Algorithm	Average	Standard Deviation
0.10	Positive Modified Greedy	0.6867	0.0837
	Positive Greedy+Max	0.7591	0.0850
	Two Set Enumeration Positive Greedy	0.7422	0.0821
	One Set Enumeration Positive Greedy+Max	0.7786	0.0793
	Sample Greedy	0.8179	0.0304
	Deterministic Linear Approximation	0.8421	0.0323
	Randomized Linear Approximation	0.8179	0.0327
0.15	Positive Modified Greedy	0.6706	0.0831
	Positive Greedy+Max	0.7786	0.0877
	Two Set Enumeration Positive Greedy	0.7608	0.0793
	One Set Enumeration Positive Greedy+Max	0.7853	0.0845
	Sample Greedy	0.8181	0.0297
	Deterministic Linear Approximation	0.8168	0.0365
	Randomized Linear Approximation	0.8299	0.0410
0.20	Positive Modified Greedy	0.6890	0.0854
	Positive Greedy+Max	0.7647	0.0868
	Two Set Enumeration Positive Greedy	0.7611	0.0789
	One Set Enumeration Positive Greedy+Max	0.7905	0.0827
	Sample Greedy	0.8131	0.0303
	Deterministic Linear Approximation	0.7399	0.0470
	Randomized Linear Approximation	0.7471	0.0310
0.30	Positive Modified Greedy	0.6823	0.0817
	Positive Greedy+Max	0.7678	0.0873
	Two Set Enumeration Positive Greedy	0.7620	0.0802
	One Set Enumeration Positive Greedy+Max	0.7815	0.0817
	Sample Greedy	0.8126	0.0306
	Deterministic Linear Approximation	0.7365	0.0570
	Randomized Linear Approximation	0.7439	0.0715
0.40	Positive Modified Greedy	0.6897	0.0829
	Positive Greedy+Max	0.7691	0.0867
	Two Set Enumeration Positive Greedy	0.7591	0.0794
	One Set Enumeration Positive Greedy+Max	0.7880	0.0802
	Sample Greedy	0.8181	0.0296
	Deterministic Linear Approximation	0.7352	0.0682
	Randomized Linear Approximation	0.7601	0.0719
0.50	Positive Modified Greedy	0.6788	0.0824
	Positive Greedy+Max	0.7603	0.0874
	Two Set Enumeration Positive Greedy	0.7409	0.0801
	One Set Enumeration Positive Greedy+Max	0.7754	0.0794
	Sample Greedy	0.8204	0.0303
	Deterministic Linear Approximation	0.7239	0.0564
	Randomized Linear Approximation	0.7331	0.0654

Table 3: The average and standard deviation of empirical approximation ratios for different algorithms over different monotonicity ratios under six budget ratios.