



Analyzing Residual Random Greedy for monotone submodular maximization

Kristóf Bérczi^a, Karthekeyan Chandrasekaran^{b,*}, Tamás Király^a, Aditya Pillai^{c,1}

^a MTA-ELTE Egerváry Research Group, Department of Operations Research, Eötvös Loránd University, Hungary

^b University of Illinois, Urbana-Champaign, United States of America

^c Georgia Institute of Technology, Atlanta, United States of America

ARTICLE INFO

Article history:

Received 15 February 2021

Received in revised form 31 January 2022

Accepted 17 October 2022

Available online 20 October 2022

Communicated by Marek Chrobak

Keywords:

Submodular maximization

Analysis of algorithms

Randomized Greedy

Matroid constraints

ABSTRACT

Residual Random Greedy (RRGREEDY) is a natural randomized version of the greedy algorithm for submodular maximization. It was introduced to address non-monotone submodular maximization [1] and plays an important role in the deterministic algorithm for monotone submodular maximization that beats the $(1/2)$ -factor barrier [2]. In this work, we analyze RRGREEDY for monotone submodular functions along two fronts: (1) For matroid constrained maximization of monotone submodular functions with bounded curvature α , we show that RRGREEDY achieves a $(1/(1+\alpha))$ -approximation in the worst-case (i.e., irrespective of the randomness in the algorithm). In particular, this implies that it achieves a $(1/2)$ -approximation in the worst-case (not just in expectation). (2) We generalize RRGREEDY to k matroid intersection constraints and show that the generalization achieves a $(1/(k+1))$ -approximation in expectation relative to the optimum value of the Lovász relaxation over the intersection of k matroid polytopes. Our results suggest that RRGREEDY is at least as good as GREEDY for matroid and matroid intersection constraints.

© 2022 Elsevier B.V. All rights reserved.

1. Introduction

A set function $f: 2^{\mathcal{N}} \rightarrow \mathbb{R}_+$ is submodular if $f(A) + f(B) \geq f(A \cap B) + f(A \cup B)$ for every pair of subsets $A, B \subseteq \mathcal{N}$, and is monotone if $f(B) \geq f(A)$ for every pair of subsets $A \subseteq B \subseteq \mathcal{N}$. Submodular functions arise in several areas including combinatorial optimization, game theory, and economics. The problem of maximizing a monotone submodular function subject to matroid intersection constraints has been extensively studied in the approximation literature. Here, we are given a monotone sub-

modular function $f: 2^{\mathcal{N}} \rightarrow \mathbb{R}_+$ by a valuation oracle and matroids $\mathcal{M}_j = (\mathcal{N}, \mathcal{I}_j)$ for $j = 1, \dots, k$ by independent set oracles, and the goal is to find a common independent set $I \in \cap_{j=1}^k \mathcal{I}_j$ so as to maximize $f(I)$. Throughout this work, we will be interested in non-negative monotone submodular functions. For subsets $U, A \subseteq \mathcal{N}$, we use $f(U|A) := f(U \cup A) - f(A)$ to denote the marginal gain in function value.

For single matroid constraint, Fisher, Nemhauser, and Wolsey [3] showed that the greedy algorithm achieves a $(1/2)$ -approximation. The greedy algorithm, denoted GREEDY, builds a candidate set A by starting with $A := \emptyset$ and repeatedly adding an element u into A such that $u + A \in \mathcal{I}$ and the marginal gain $f(u|A)$ is maximized. It is known that no polynomial-time algorithm can yield $(1 - 1/e + \epsilon)$ -approximation for any constant $\epsilon > 0$ [4]. In a breakthrough work, Călinescu, Chekuri, Pál, and Vondrák [5] designed a randomized $(1 - 1/e)$ -approximation. Very

* Corresponding author.

E-mail addresses: berkri@cs.elte.hu (K. Bérczi), karthe@illinois.edu (K. Chandrasekaran), tkiraly@cs.elte.hu (T. Király), apillai32@gatech.edu (A. Pillai).

¹ Work done while the author was a student at the University of Illinois, Urbana-Champaign.

```

RRGREEDY ( $\mathcal{M}, f$ )
Input: A matroid  $\mathcal{M} = (\mathcal{N}, \mathcal{I})$  and
a monotone submodular function  $f : 2^{\mathcal{N}} \rightarrow \mathbb{R}_+$ 
Initialize  $A \leftarrow \emptyset$ 
While there exists  $u \in \mathcal{N} - A$  such that  $A + u \in \mathcal{I}$ :
  Let  $M$  be a subset of  $\mathcal{N} - A$  that maximizes
   $\sum_{u \in M} f(u|A)$  such that  $M \cup A \in \mathcal{I}$ 
  Let  $u$  be a uniformly random element from  $M$ 
   $A \leftarrow A + u$ 
Return  $A$ 

```

Fig. 1. RRGREEDY for matroid constraint.

recently, Buchbinder, Feldman, and Garg [2] designed the first deterministic algorithm that broke the $(1/2)$ -factor barrier. One of the main ingredients of their improved approximation is a randomized version of the greedy algorithm, known as RRGREEDY (short for Residual Random Greedy). As part of their proof, they showed that RRGREEDY achieves $(1/2)$ -approximation in expectation. In this work, we investigate RRGREEDY along two fronts: (1) What is the approximation factor of RRGREEDY for monotone submodular functions with bounded curvature, and (2) Is there a generalization of RRGREEDY for matroid intersection with good approximation guarantee?

RRGREEDY was introduced by Buchbinder, Feldman, Naor, and Schwartz [1] to address non-monotone submodular maximization as a natural way to randomize the greedy strategy. Instead of choosing an element that maximizes the marginal gain with respect to the current set, it considers a set of good choices to add and randomly chooses one among them. We state RRGREEDY for matroid constraint in Fig. 1.

It is easy to see that RRGREEDY achieves the optimum value for modular functions while previous works have shown that it achieves $(1/2)$ -approximation in expectation for monotone submodular functions [2,1]. Is there a smooth interpolation of its approximation factor based on *distance to modularity* of the submodular function? Conforti and Cornuéjols [6] introduced the notion of curvature as a measure of distance to modularity of a set function. The curvature of $f : 2^{\mathcal{N}} \rightarrow \mathbb{R}_+$ is defined as $\alpha := 1 - \min_{u \in \mathcal{N}: f(u) > 0} (f(\mathcal{N}) - f(\mathcal{N} - u)) / f(u)$. We note that $\alpha = 0$ if and only if f is modular and $0 \leq \alpha \leq 1$ for monotone submodular functions. Conforti and Cornuéjols showed that GREEDY achieves $(1/(1 + \alpha))$ -approximation for matroid constraint. Their result gives a unified explanation for the approximation factor of GREEDY for both modular as well as monotone submodular functions. Given this status of GREEDY, can we analyze the approximation factor of RRGREEDY for bounded curvature functions? Is it at least as good as GREEDY? We address this in Theorem 1.1. Our theorem shows a surprisingly strong result about RRGREEDY: it is at least as good as GREEDY in the worst-case and not just in expectation (i.e., for all random choices).

Next, we turn to k matroid intersection constraints where $k \geq 2$. Fisher, Nemhauser, and Wolsey [3] considered the natural extension of GREEDY, where the next element u is chosen to be the one with maximum $f(u|A)$ subject to $u + A$ being independent in *all* matroids. We will denote this extension also as GREEDY. They showed that

GREEDY achieves $(1/(k + 1))$ -approximation. We note that for $k = 1$, the expected approximation factor of RRGREEDY and the approximation factor of GREEDY coincide. While GREEDY has a natural generalization for $k \geq 2$ matroid constraints, it is apriori unclear if there is a natural generalization of RRGREEDY to $k \geq 2$ matroid constraints and even if so, would its approximation ratio be comparable to that of GREEDY? In this work, we present a generalization of RRGREEDY for matroid intersection constraints and show that the approximation factor of the generalization is indeed at least as good as GREEDY.

Although we do not improve the approximation factors of the respective problems, our results present evidence suggesting that RRGREEDY might be a viable algorithm for improving the approximation factors. We show surprising results which were not known for RRGREEDY. Our results raise several interesting open questions about RRGREEDY that we summarize in the conclusion section.

1.1. Our contributions

Our first result is about the approximation factor of RRGREEDY for maximizing a monotone submodular function with bounded curvature subject to matroid constraint. We recall that RRGREEDY achieves an optimum for $\alpha = 0$ and a $(1/2)$ -approximation in expectation for $\alpha = 1$. We show the following:

Theorem 1.1. RRGREEDY achieves a $(1/(1 + \alpha))$ -approximation for all random choices in the algorithm.

The theorem highlights two important features of RRGREEDY. Firstly, it shows that the approximation ratio of RRGREEDY is at least as good as GREEDY even for bounded curvature functions. Secondly, for arbitrary submodular functions (i.e., when curvature $\alpha = 1$), RRGREEDY achieves a $(1/2)$ -approximation in the worst-case and not only just in expectation as was known before [2,1]. If we consider the execution of a randomized algorithm as a branching tree where the algorithm's decision can be viewed as follows a particular branch based on the outcome of a random coin toss, then our result says that every branch of the execution tree would still lead to a $(1/2)$ -approximation. This also tells us that the greedy algorithm is forgiving in terms of mistakes: even if we make certain kind of mistakes in greedy, we would still get a $(1/2)$ -approximation! In particular, instead of picking the element u with maximum $f(u|A)$, we could have chosen an arbitrary element $u \in M$, where M is a set in the contracted matroid $\mathcal{M}/A$ with maximum sum of marginals $\sum_{u \in M} f(u|A)$, and it would still lead to a $(1/2)$ -approximation. We note that this result was previously known only for partition matroids [7].

Next, we generalize RRGREEDY for k matroid intersection constraints. For the case of $k = 2$, a natural generalization is to define the set M to be a common independent set in $\mathcal{M}_1/A$ and $\mathcal{M}_2/A$ that maximizes $\sum_{u \in M} f(u|A)$. We recall that the problem of finding a maximum weight common independent set in two matroids is solvable in polynomial time [8]. This leads to the algorithm given in Fig. 2. We show that this algorithm

```

2MATROIDINTERSECTIONRRGREEDY ( $\mathcal{M}_1, \mathcal{M}_2, f$ )
Input: Matroids  $\mathcal{M}_b = (\mathcal{N}, \mathcal{I}_b)$  for  $b = 1, 2$  and
a monotone submodular function  $f : 2^{\mathcal{N}} \rightarrow \mathbb{R}_+$ 
Initialize  $A \leftarrow \emptyset$ 
While there exists  $u \in \mathcal{N} - A$  such that  $A + u \in \mathcal{I}_1 \cap \mathcal{I}_2$ :
  Let  $M$  be a subset of  $\mathcal{N} - A$  that maximizes
   $\sum_{u \in M} f(u|A)$  such that  $M \cup A \in \mathcal{I}_1 \cap \mathcal{I}_2$ 
  Let  $u$  be a uniformly random element from  $M$ 
   $A \leftarrow A + u$ 
Return  $A$ 

```

Fig. 2. RRGREEDY for 2 matroid intersection constraints.

(i.e., 2MATROIDINTERSECTIONRRGREEDY) achieves a $(1/3)$ -approximation.

Theorem 1.2. For an arbitrary common independent set T and the set A output by 2MATROIDINTERSECTIONRRGREEDY, we have that

$$\mathbb{E}[f(A)] \geq \frac{f(T)}{3}.$$

Remark 1.3. A natural question is whether Theorem 1.2 is tight, i.e., does 2MATROIDINTERSECTIONRRGREEDY achieve an approximation factor that is better than $1/3$, perhaps $1/2$? Based on computer experiments, we were able to show the following: There exist matroids $\mathcal{M}_1$ and $\mathcal{M}_2$ and a monotone submodular function f such that the best possible approximation factor achieved for *all* random choices in the execution of 2MATROIDINTERSECTIONRRGREEDY is at most $13/33$. This example is described in the appendix.

Remark 1.4. Based on the result in Theorem 1.1, one might conjecture that the statement of Theorem 1.2 may also hold in the worst case (and not just in expectation). However, we were able to rule out this conjecture with a computer-generated example.

Next, we generalize RRGREEDY for k matroid intersection constraints where $k \geq 3$. We note that the algorithm in Fig. 2 does not generalize to $k \geq 3$ since the problem of finding a maximum weight common independent set in three matroids is NP-hard. Instead, we take an alternative perspective of RRGREEDY. We view RRGREEDY for $k = 1$ as optimizing over the matroid polytope to obtain a solution $x \in [0, 1]^{\mathcal{N}}$ and then sampling an element u with probability proportional to $x(u)$. We recall that extreme point solutions x to the matroid polytope are integral and hence, this naturally leads to the choice $M = \text{support}(x)$. This perspective leads to a natural generalization of RRGREEDY for $k \geq 2$ matroid intersection constraints: optimize over the intersection of k matroid polytopes to obtain a point z and then sample a point u with probability proportional to $z(u)$ to add to the current set A . For a fixed set A that is independent in all k matroids, we consider the following LP whose feasible region is the intersection of the k matroid polytopes (where r_j is the rank function of the j 'th matroid):

$$\max \sum_{u \in \mathcal{N}} z(u) f(u|A) \quad (\text{LP}(A))$$

```

MATROIDINTERSECTIONRRGREEDY ( $\mathcal{M}_1, \dots, \mathcal{M}_k, f$ )
Input: Matroids  $\mathcal{M}_j = (\mathcal{N}, \mathcal{I}_j)$  for  $j \in [k]$  and
a monotone submodular function  $f : 2^{\mathcal{N}} \rightarrow \mathbb{R}_+$ 
Initialize  $A \leftarrow \emptyset$ 
While there exists  $u \in \mathcal{N} - A$  such that  $A + u \in \cap_{j=1}^k \mathcal{I}_j$ :
  Let  $z$  be an optimum solution to LP( $A$ )
  Let  $u$  be an element in  $\mathcal{N} - A$  sampled with
  probability  $z(u) / \sum_{v \in \mathcal{N} - A} z(v)$ 
   $A \leftarrow A + u$ 
Return  $A$ 

```

Fig. 3. RRGREEDY for k matroid intersection constraints.

$$\sum_{u \in S - A} z(u) \leq r_j(S \cup A) - |A| \quad \forall S \subseteq \mathcal{N}, j \in [k]$$

$$z(u) \geq 0 \quad \forall u \in \mathcal{N}$$

$$z(u) = 1 \quad \forall u \in A.$$

We use this LP to generalize RRGREEDY for k -matroid intersection as shown in Fig. 3.

In order to analyze MATROIDINTERSECTIONRRGREEDY, we compare the objective value of the set returned by the algorithm with the optimum objective value of the Lovász relaxation. For a set $S \subseteq \mathcal{N}$, let $\chi^S \in \{0, 1\}^{\mathcal{N}}$ denote the indicator vector of S . For a function $f : 2^{\mathcal{N}} \rightarrow \mathbb{R}_+$, the convex closure $f^- : [0, 1]^{\mathcal{N}} \rightarrow \mathbb{R}$ is defined as $f^-(y) :=$

$$\min \left\{ \sum_{S \subseteq \mathcal{N}} \lambda_S f(S) : \sum_{S \subseteq \mathcal{N}} \lambda_S \chi^S = y, \lambda_S \geq 0 \quad \forall S \subseteq \mathcal{N} \right\}. \quad (1)$$

We note that the definition we use here is not the same as the original definition of the Lovász relaxation. It is well known that the convex closure of a function f is equivalent to the Lovász relaxation in the cube $[0, 1]^{\mathcal{N}}$ when f is submodular [9].

We show the following result for the algorithm MATROIDINTERSECTIONRRGREEDY.

Theorem 1.5. Let A be the set returned by MATROIDINTERSECTIONRRGREEDY and let t be any feasible point for LP($\emptyset$). Then,

$$\mathbb{E}[f(A)] \geq \left(\frac{1}{k+1} \right) f^-(t).$$

The theorem shows that the expected approximation factor of RRGREEDY for k matroid intersection is at least that of GREEDY. Moreover, our approximation guarantee is relative to the optimum value of the Lovász relaxation over the k matroid intersection polytope. Thus, it could be viewed as bounding the integrality gap. We note that our techniques underlying the proof of Theorem 1.5 can also be extended to show that the approximation factor of GREEDY is also the same factor relative to the optimum value of the Lovász relaxation over the k matroid intersection polytope.

1.2. Related work

As mentioned earlier, Conforti and Cornuéjols [6] introduced the notion of curvature as a measure of distance to modularity of a monotone submodular function.

They showed that GREEDY for matroid constraint achieves a $(1/(1+\alpha))$ -approximation, where α is the curvature of the monotone submodular function. For uniform matroid, they showed that GREEDY achieves a $(1/\alpha)(1-e^{-\alpha})$ -approximation. This result also recovers the $(1-e^{-1})$ -approximation guarantee for the case of $\alpha = 1$ which was shown by Fisher, Nemhauser, and Wolsey [3]. Recent works have extended the results for uniform matroids beyond submodular functions [10]. Vondrák introduced a slightly different definition of curvature, denoted by c [11], and showed that the continuous greedy algorithm of [5] achieves a $(1/c)(1-e^{-c})$ -approximation for any matroid. Subsequently, Sviridenko, Vondrák, and Ward showed a different algorithm that achieves a $(1-\alpha/e)$ -approximation for any matroid [12].

For $k \geq 2$ matroid intersection constraints, Fisher, Nemhauser, and Wolsey showed that GREEDY achieves a $(1/(k+1))$ -approximation. Conforti and Cornuéjols showed that it achieves a $(1/(k+\alpha))$ -approximation, where α is the curvature of the monotone submodular function. Lee, Sviridenko, and Vondrák [13] gave a polynomial-time local-search algorithm that achieves a $(1/k+\epsilon)$ -approximation for $k \geq 2$ matroid intersection constraints for arbitrary curvature (for any constant ϵ). We observed via an additional lemma from their main theorem that their algorithm also achieves a $(1/(k-1+\alpha+\epsilon))$ -approximation for monotone submodular functions with curvature α .

For $k \geq 3$ matroid intersection constraints, the case where the submodular function is in fact a modular function (i.e., a linear function) is also of special interest. In particular, if the modular function is the cardinality function, then Lau, Ravi, and Singh [14] showed an iterative rounding algorithm that achieves a $(1/(k-1))$ -approximation relative to the LP-optimum value. For arbitrary modular functions, Lee, Sviridenko, and Vondrák [13] showed that their local search algorithm achieves a $(1/(k-1+\epsilon))$ -approximation. A well-known open question for modular functions is whether the integrality gap of the LP is at least $(1/(k-1))$. In a recent work, Linhares, Olver, Swamy, and Zenklusen [15] showed that the integrality gap for $k = 3$ is indeed $1/2$.

1.3. Preliminaries

Let $\mathcal{N}$ be a finite set of size n and $f: 2^{\mathcal{N}} \rightarrow \mathbb{R}_+$ be a set function. We use $+$ and $-$ as a shorthand for set union and difference, respectively. We repeat the definitions of submodularity and monotonicity for the sake of completeness. Let $f(U|S)$ denote $f(U+S) - f(S)$ where $U, S \subseteq 2^{\mathcal{N}}$. A set function $f: 2^{\mathcal{N}} \rightarrow \mathbb{R}_+$ is *submodular* if and only if for all $S \subseteq T \subseteq \mathcal{N}$ and $u \in \mathcal{N} - T$, we have $f(u|S) \geq f(u|T)$. Additionally, the function f is *monotone* if $f(S) \leq f(T)$ for all $S \subseteq T \subseteq \mathcal{N}$. A submodular function is *modular* if there are weights $w: \mathcal{N} \rightarrow \mathbb{R}_{\geq 0}$ defined on every element of the ground set such that $f(S) = \sum_{s \in S} w(s)$ for all $S \subseteq \mathcal{N}$. Throughout the paper, we assume that f is a monotone, submodular function satisfying $f(\emptyset) = 0$.² We

note that this immediately implies that $f(X) \geq 0$ for every $X \subseteq S$. We may assume that $f(\emptyset) = 0$ without loss of generality since otherwise we may work with the normalized function $g: 2^{\mathcal{N}} \rightarrow \mathbb{R}$ defined by $g(S) := f(S) - f(\emptyset)$ for all $S \subseteq \mathcal{N}$ which is non-negative, submodular, and monotone. It is clear that subtracting $f(\emptyset)$ from the function does not change the optimum solution of the problem, so a set which is a c -approximation with respect to the function g (for $c \leq 1$) is still a c -approximation with respect to f since $g(S) \geq c \cdot g(OPT)$ implies that $f(S) \geq c \cdot f(OPT) + (1-c)f(\emptyset) \geq c \cdot f(OPT)$.

Matroids were introduced by Whitney [16] and Nakasawa [17]. A matroid $\mathcal{M}$ is a pair $(\mathcal{N}, \mathcal{I})$ where $\mathcal{N}$ is a ground set and $\mathcal{I} \subseteq 2^{\mathcal{N}}$ is the family of *independent sets* satisfying the following, so-called *independence axioms*: (I1) $\emptyset \in \mathcal{I}$, (I2) $X \subseteq Y \in \mathcal{I} \Rightarrow X \in \mathcal{I}$, and (I3) $X, Y \in \mathcal{I}$, $|X| < |Y| \Rightarrow \exists e \in Y - X$ s.t. $X + e \in \mathcal{I}$. A maximal independent set is called a *base*. The independent set axioms imply that all bases have the same size called the *rank* of the matroid. The rank function $r: 2^{\mathcal{N}} \rightarrow \mathbb{Z}_+$ of a matroid $\mathcal{M} = (\mathcal{N}, \mathcal{I})$ is defined as $r(S) := \max\{|X| \mid X \subseteq S, X \in \mathcal{I}\}$. Contracting a matroid $\mathcal{M}$ by an independent set $S \in \mathcal{I}$ is denoted by $\mathcal{M}/S$: here, a set $I \subseteq \mathcal{N} - S$ is an independent set of $\mathcal{M}/S$ if $I \cup S$ is independent in $\mathcal{M}$. It is easy to verify that $\mathcal{M}/S$ is a matroid with rank function $r': 2^{\mathcal{N}-S} \rightarrow \mathbb{Z}_+$ given by $r'(X) = r(X \cup S) - |S|$. We will need the following well known result about matroids (see e.g. [18, Theorem 5.3.4]).

Lemma 1.6. *Let A, B be independent sets in a matroid $\mathcal{M}$ such that $|A| = |B|$. Then, there exists a bijection $h: A \rightarrow B$ such that $h(a) = a$ for all $a \in A \cap B$ and the set $B - h(a) + a$ is an independent set for all $a \in A$.*

2. Analyzing RRGREEDY for bounded curvature functions under matroid constraint

In this section we analyze the approximation factor of RRGREEDY for submodular functions with curvature α under matroid constraint. We show that the approximation factor of RRGREEDY is at least that of GREEDY, i.e., it achieves a $(\frac{1}{1+\alpha})$ -approximation. Furthermore, we show that the approximation ratio holds in the worst-case: any set output by RRGREEDY is a $(\frac{1}{1+\alpha})$ -approximation. We need the following well known fact about matroids.

Lemma 2.1. *Let P, Q be bases of $\mathcal{M}$. For all $p \in P - Q$, there exists $q \in Q - P$ such that $P - p + q$ and $Q - q + p$ are bases of $\mathcal{M}$.*

Let T be an arbitrary base in the matroid and let A be the set output by RRGREEDY. We will now use Lemma 2.1 to get a bijection from $A - T$ to $T - A$ that satisfies a non-increasing condition. This bijection is our key insight to prove Theorem 1.1.

Lemma 2.2. *Let T be an arbitrary base in the matroid. In the execution of RRGREEDY, let u_i be the element chosen from the set M_i in the i 'th iteration of the while loop, let A_i be the set after i iterations of the while loop, m be the number of iterations of the while loop and let A be the set output by RRGREEDY. Then, there exists a mapping $\phi: T - A \rightarrow A - T$ such that*

² Functions satisfying these three properties are sometimes called *polymatroid functions* in the literature.

1. for every $w \in T - A$ we have $\phi(w) = u_i$ for some $i \in [m]$ with $f(w|A) \leq f(u_i|A_{i-1})$, and
2. ϕ is a bijection.

Proof. For every $0 \leq i \leq m$, we will define a set $T_i \subseteq T$ that will be a base of $\mathcal{M}/A_i$. Let $T_0 := T$. The set T_i will be defined recursively based on the choices of the algorithm. We may assume that the set M_i is a base of $\mathcal{M}/A_{i-1}$ since f is monotone, so we get $|M_i| = |T_{i-1}|$ for all $i \in [m]$. We define a bijection $h: A \rightarrow T$ such that for all $u_i \in A$

1. $f(u_i|A_{i-1}) \geq f(h(u_i)|A_{i-1})$,
2. $h(u_i) = u_i$ if $u_i \in T_{i-1}$, and
3. $T_{i-1} + u_i - h(u_i)$ is a base of $\mathcal{M}/A_{i-1}$.

We now define $h(u_i)$ assuming we have A_{i-1} , M_i , and T_{i-1} . If $u_i \in M_i \cap T_{i-1}$, we set $h(u_i) := u_i$ and all properties are satisfied (here $T_{i-1} + u_i - h(u_i) = T_{i-1}$). If $u_i \in M_i - T_{i-1}$, we set $h(u_i) := t$ where $t \in T_{i-1} - A_{i-1}$ is acquired from Lemma 2.1 by considering the matroid $\mathcal{M}/A_{i-1}$ and bases $P := M_i$, $Q := T_{i-1}$. With this choice, we now prove that the above-mentioned properties are satisfied. The first property is satisfied because $M_i + t - u_i$ is a base of $\mathcal{M}/A_{i-1}$ and if $f(u_i|A_{i-1}) < f(t|A_{i-1})$, then it would imply that $\sum_{u \in M_i + t - u_i} f(u|A_{i-1}) > \sum_{u \in M_i} f(u|A_{i-1})$, contradicting the optimality of M_i . The third property is satisfied by Lemma 2.1.

We set $T_i := T_{i-1} - h(u_i)$. Consequently, T_i is a base of $\mathcal{M}/A_i$ by the third property and the fact that $A_i = A_{i-1} + u_i$.

Using the bijection h , we now construct the required bijection ϕ between $T - A$ and $A - T$. Let $w \in T - A$. Then, there exists a unique $u_{j_0} \in A$ where $u_{j_0} = h^{-1}(w)$ since h is a bijection. By submodularity and the first property of h , we have

$$f(w|A) \leq f(w|A_{j_0-1}) \leq f(u_{j_0}|A_{j_0-1}).$$

If $u_{j_0} \notin T$, then we define $\phi(w) := u_{j_0}$. If $u_{j_0} \in T$, then it must be the case that $u_{j_0} \notin T_{j_0-1}$: suppose for contradiction $u_{j_0} \in T_{j_0-1}$, then by the second property of the bijection h , we have $h(u_{j_0}) = u_{j_0} \in A$, but $h(u_{j_0}) = w \notin A$, a contradiction. Therefore, we have $u_{j_0} \notin T_{j_0-1}$. Since h is a bijection, there exists a unique index $j_1 < j_0$ such that $h(u_{j_1}) = u_{j_0}$. Once again, by submodularity and the first property of h , we have

$$f(u_{j_0}|A_{j_0-1}) \leq f(u_{j_0}|A_{j_1-1}) \leq f(u_{j_1}|A_{j_1-1}).$$

If $u_{j_1} \notin T$, then we define $\phi(w) := u_{j_1}$. Otherwise, if $u_{j_1} \in T$, then it must be the case that $u_{j_1} \notin T_{j_1-1}$ by the second property of the bijection h . Since h is a bijection, there exists a unique index $j_2 < j_1$ such that $h(u_{j_2}) = u_{j_1}$. Once again, by submodularity and the first property of h , we have

$$f(u_{j_1}|A_{j_1-1}) \leq f(u_{j_1}|A_{j_2-1}) \leq f(u_{j_2}|A_{j_2-1}).$$

If $u_{j_2} \notin T$, then we define $\phi(w) := u_{j_2}$. Otherwise, we continue to trace back by the same argument as above until we arrive at an element $u_{j_r} \notin T$ and define $\phi(w) := u_{j_r}$.

We note that the tracing back procedure has to terminate since the indices are strictly decreasing: $j_0 > j_1 > j_2 > \dots > j_r$. Thus, for every $w \in T - A$, we have a sequence of indices $j_0 > j_1 > j_2 > \dots > j_r$ such that

1. $w, u_{j_0}, \dots, u_{j_r-1} \in T$ and $u_{j_r} \in A \setminus T$,
2. $f(w|A) \leq f(u_{j_0}|A_{j_0-1}) \leq \dots \leq f(u_{j_r}|A_{j_r-1})$,
3. $h(u_{j_i}) = u_{j_{i-1}}$ for every $i = r, \dots, 1$ and $h(u_{j_0}) = w$.

By these properties, we have $u_{j_r} \in A - T$ and moreover, $f(w|A) \leq f(u_{j_r}|A_{j_r-1})$.

We now prove that ϕ is indeed a bijection. Since $|T - A| = |A - T|$, it suffices to show that ϕ is a one-to-one mapping. Suppose for contradiction, we have two distinct elements $w, w' \in T - A$ such that $\phi(w) = \phi(w')$. Let $(j_0, j_1, \dots, j_r)$ and $(i_0, i_1, \dots, i_p)$ be the sequence of indices that are encountered while defining $\phi(w)$ and $\phi(w')$ respectively. Since $w \neq w'$ and h is a mapping it follows that $u_{j_0} \neq u_{i_0}$. Since $\phi(w) = \phi(w')$, we have $u_{j_r} = u_{i_p}$. It follows that there exists a pair of indices j_ℓ and i_m such that $u_{j_\ell} \neq u_{i_m}$ but $u_{j_{\ell+1}} = u_{i_{m+1}}$. However, we know that $h(u_{j_{\ell+1}}) = u_{j_\ell}$ and $h(u_{i_{m+1}}) = u_{i_m}$. This contradicts the fact that h is a mapping. $\square$

Remark 2.3. A natural question that arises from our proof of Lemma 2.2 is whether we could have simply defined ϕ to be h^{-1} restricted to $A - T$. For this to work we would need $h(A \cap T) = A \cap T$; however, this condition fails since the second property of h holds only for elements $u_i \in T_{i-1}$ but not necessarily for elements $u_i \in T$. Consider the following scenario. Suppose for some $u_i \in A - T$ we have $h(u_i) = w$. It may be the case that for some $j > i$, we have that $u_j = w$ (i.e., the algorithm picks w in an iteration $j > i$) and now we cannot set $h(w) = w$ since h is a bijection and we have already set $h(u_i) = w$.

We need the following lemma from [6].

Lemma 2.4. (Conforti and Cornuéjols [6]) Given $T \subseteq \mathcal{N}$ and an ordered set $A = \{u_1, u_2, \dots, u_t\}$, define $A_i := \{u_1, u_2, \dots, u_i\}$ for $1 \leq i \leq t$. Then, $f(T)$ is at most

$$\alpha \sum_{u_i \in A-T} f(u_i|A_{i-1}) + \sum_{u_i \in A \cap T} f(u_i|A_{i-1}) + \sum_{w \in T-A} f(w|A).$$

We now restate and prove Theorem 1.1.

Theorem 1.1. RRGREEDY achieves a $(1/(1+\alpha))$ -approximation for all random choices in the algorithm.

Proof. By construction, RRGREEDY outputs a base in the matroid. We will show that for every base T and for every base A output by RRGREEDY, we have that $f(A) \geq \frac{f(T)}{1+\alpha}$. Let u_i be the element chosen in the i 'th iteration of the while loop and let A_i be the set A after i iterations of the while loop ($A_0 = \emptyset$). By Lemma 2.2, we have that

$$\sum_{w \in T-A} f(w|A) \leq \sum_{u_i \in A-T} f(u_i|A_{i-1}). \quad (2)$$

Using Lemma 2.4, we have that $f(T)$

$$\begin{aligned}
&\leq \alpha \sum_{u_i \in A-T} f(u_i|A_{i-1}) + \sum_{u_i \in A \cap T} f(u_i|A_{i-1}) + \sum_{w \in T-A} f(w|A) \\
&\leq \alpha \sum_{u_i \in A-T} f(u_i|A_{i-1}) + \sum_{u_i \in A \cap T} f(u_i|A_{i-1}) \\
&\quad + \sum_{u_i \in A-T} f(u_i|A_{i-1}) \quad (\text{By inequality (2)}) \\
&= \alpha \sum_{u_i \in A-T} f(u_i|A_{i-1}) + f(A) \leq (1 + \alpha)f(A). \quad \square
\end{aligned}$$

3. RRGreedy for k -matroid intersection

In this section, we consider the problem of maximizing a monotone submodular function subject to matroid intersection constraints. Here, we are given a monotone submodular function $f : 2^{\mathcal{N}} \rightarrow \mathbb{R}_+$ and matroids $\mathcal{M}_j = (\mathcal{N}, \mathcal{I}_j)$ for $j = 1, \dots, k$ for $k \geq 2$ and the goal is to find a common independent set $I \in \cap_{j=1}^k \mathcal{I}_j$ that maximizes $f(I)$. We note that, in contrast to Theorem 1.1, the approximation factors of the algorithms in this section hold only in expectation and not in the worst-case. We will analyze 2MATROIDINTERSECTIONRRGREEDY and MATROIDINTERSECTIONRRGREEDY (both of which were presented in Section 1.1) and prove Theorems 1.2 and 1.5.

For ease of understanding, we begin by analyzing the approximation factor of 2MATROIDINTERSECTIONRRGREEDY and prove Theorem 1.2 in Section 3.1. Next, we follow the ideas underlying this analysis to bound the approximation factor of MATROIDINTERSECTIONRRGREEDY, i.e., the generalization of RRGREEDY for arbitrary k , and prove Theorem 1.5. We need the following proposition.

Proposition 3.1. *For all independent sets P, Q in a matroid $\mathcal{M} = (\mathcal{N}, \mathcal{I})$ there exists a function $h : P \rightarrow Q \cup \{d\}$ where d is a dummy element such that*

1. $Q - h(a) + a \in \mathcal{I}$ for all $a \in P$ and
2. h is injective when restricted to $h^{-1}(Q)$, i.e., for all $p_1, p_2 \in P$ if $h(p_1), h(p_2) \in Q$, then $h(p_1) \neq h(p_2)$.

Proof. First we consider the case when $|P| \leq |Q|$. By the exchange property of matroids, there exists $V \subseteq Q - P$ such that $|V \cup P| = |Q|$ and $V \cup P \in \mathcal{I}$. By Lemma 1.6, there exists a bijection $h : P \cup V \rightarrow Q$. The bijection maps elements in V to themselves, so elements of P map to unique elements of Q and both properties are satisfied. Next, we consider the case when $|Q| < |P|$. There exists $V \subseteq P - Q$ such that $V \cup Q \in \mathcal{I}$ and $|V \cup Q| = |P|$. This implies that $|P - V| = |Q|$, so we can use Lemma 1.6 to get $h : P - V \rightarrow Q$. All elements in $v \in V$ have the property that $Q + v \in \mathcal{I}$, so we define $h(v) := d$ for all $v \in V$. Such a mapping h satisfies both properties. $\square$

3.1. Generalization of RRGreedy for $k = 2$

In this section, we analyze the approximation factor of 2MATROIDINTERSECTIONRRGREEDY which was presented in Section 1.1. We restate and prove Theorem 1.2

Theorem 1.2. *For an arbitrary common independent set T and the set A output by 2MATROIDINTERSECTIONRRGREEDY, we have that*

$$\mathbb{E}[f(A)] \geq \frac{f(T)}{3}.$$

Proof. In the execution of 2MATROIDINTERSECTIONRRGREEDY, let M_i be the set M and u_i be the element chosen in the i 'th iteration of the while loop. Let A_i be the set A after i iterations of the while loop. We now define sets T_i which will be useful in the analysis. Let $T_0 := T$. The sets T_i will be defined inductively so that $A_i \cup T_i \in \mathcal{I}_1 \cap \mathcal{I}_2$. Assuming we have A_{i-1} , M_i , and T_{i-1} , we use Proposition 3.1 with $P = M_i$, $Q = T_{i-1}$, and $\mathcal{M} = \mathcal{M}_b/A_{i-1}$ to obtain functions $h_b : M_i \rightarrow T_{i-1} \cup \{d\}$, where d is a dummy element, such that for $b \in \{1, 2\}$,

1. $T_{i-1} - h_b(u) + u \in \mathcal{I}_b/A_{i-1}$ for every $u \in M_i$ and
2. h_b is injective when restricted to $h_b^{-1}(T_{i-1})$.

With such functions h_1 and h_2 , we define $T_i := T_{i-1} - h_1(u_i) - h_2(u_i)$. We note that the first property of h_b implies that $A_i \cup T_i \in \mathcal{I}_1 \cap \mathcal{I}_2$.

To prove the theorem, we will show that for every iteration i , we have that

$$\begin{aligned}
&2\mathbb{E}[f(A_i)] + \mathbb{E}[f(A_i \cup T_i)] \\
&\geq 2\mathbb{E}[f(A_{i-1})] + \mathbb{E}[f(A_{i-1} \cup T_{i-1})]. \quad (3)
\end{aligned}$$

This would imply the theorem as follows: let m be the number of iterations executed by the algorithm. Then,

$$\begin{aligned}
3\mathbb{E}[f(A_m)] &= 2\mathbb{E}[f(A_m)] + \mathbb{E}[f(A_m \cup T_m)] \\
&\geq 2\mathbb{E}[f(A_0)] + \mathbb{E}[f(A_0 \cup T_0)] \\
&= f(T).
\end{aligned}$$

We note that $T_m = \emptyset$ because the algorithm outputs a maximal common independent set.

We now prove inequality (3). We first fix the choices made by the algorithm in the first $i - 1$ iterations. Consequently, the sets A_{i-1} , M_i , and T_{i-1} are fixed. From the second property of h_b , we obtain that

$$\mathbb{E}_{u_i}[f(h_b(u_i)|A_{i-1})] = \frac{\sum_{u \in M_i} f(h_b(u)|A_{i-1})}{|M_i|} \quad (4)$$

$$\leq \frac{\sum_{w \in T_{i-1}} f(w|A_{i-1})}{|M_i|} \quad (5)$$

for $b = 1, 2$. The inequality follows because the sum in the LHS contains unique elements of T_{i-1} . If d appears in the sum, we can think of $f(d|A_{i-1}) = 0$ since f is not defined on d . We have, $\mathbb{E}_{u_i}[f(A_i)] - f(A_{i-1})$

$$\begin{aligned}
&= \mathbb{E}_{u_i}[f(u_i|A_{i-1})] \\
&= \frac{\sum_{u \in M_i} f(u|A_{i-1})}{|M_i|} \\
&\geq \frac{\sum_{w \in T_{i-1}} f(w|A_{i-1})}{|M_i|} \quad (6)
\end{aligned}$$

$$\geq \frac{\mathbb{E}_{u_i}[f(h_1(u_i)|A_{i-1})] + \mathbb{E}_{u_i}[f(h_2(u_i)|A_{i-1})]}{2} \quad (7)$$

$$\geq \frac{\mathbb{E}_{u_i}[f(h_1(u_i)|A_{i-1} \cup T_i)] + \mathbb{E}_{u_i}[f(h_2(u_i)|A_{i-1} \cup T_i)]}{2} \quad (8)$$

$$\geq \frac{\mathbb{E}_{u_i}[f(\{h_1(u_i), h_2(u_i)\}|A_{i-1} \cup T_i)]}{2} \quad (9)$$

$$= \frac{f(T_{i-1} \cup A_{i-1}) - \mathbb{E}_{u_i}[f(A_{i-1} \cup T_i)]}{2} \quad (10)$$

$$\geq \frac{f(T_{i-1} \cup A_{i-1}) - \mathbb{E}_{u_i}[f(A_i \cup T_i)]}{2}.$$

Inequality (6) is by the choice of the algorithm since $A_{i-1} \cup T_{i-1} \in \mathcal{I}_1 \cap \mathcal{I}_2$ and hence, T_{i-1} is a valid candidate for M_i . Inequality (7) follows by (5). Inequalities (8) and (9) follow by submodularity and inequality (10) follows by monotonicity. Thus, for any fixed choice of A_{i-1} we have, $2\mathbb{E}_{u_i}[f(A_i)] + \mathbb{E}_{u_i}[f(A_i \cup T_i)] \geq 2f(A_{i-1}) + f(A_{i-1} \cup T_{i-1})$. So this inequality also holds in expectation when A_{i-1} is unfixed. This shows that $2\mathbb{E}[f(A_i)] + \mathbb{E}[f(A_i \cup T_i)]$ is monotone in i . $\square$

3.2. Generalization of RRGreedy for arbitrary k

In this section, we analyze the approximation factor of MATROIDINTERSECTIONRRGREEDY and prove Theorem 1.5. We note that the linear program $\text{LP}(A)$ presented in Section 1.1 has exponentially many constraints, but it can be solved in polynomial-time [19]. We first prove certain useful properties of the Lovász extension in Section 3.2.1. We prove Theorem 1.1 in Section 3.2.2.

3.2.1. Properties of the Lovász extension

We recall that the Lovász extension of $f : 2^{\mathcal{N}} \rightarrow \mathbb{R}_+$ is denoted by f^- . For ease of notation, let $f^-(x + A) := f^-(x + \chi^A)$ for all $A \subseteq \mathcal{N}$ and $x \in \mathbb{R}_+^{\mathcal{N}}$. For a set $S \subseteq \mathcal{N}$, we define $f^-(x|S) := g^-(x)$ where $g : 2^{\mathcal{N}} \rightarrow \mathbb{R}_+$ is the submodular function defined by $g(P) = f(P|S)$.

There are two further definitions of the Lovász extension that are equivalent to (1) (see [9]). We state these two definitions now. Let $y \in [0, 1]^{\mathcal{N}}$. Let $s_1, \dots, s_n$ be an ordering of the elements of $\mathcal{N}$ such that $y(s_1) \geq y(s_2) \geq \dots \geq y(s_n)$ and let $S_i := \{s_1, \dots, s_i\}$ for $i = 1, \dots, n$ and $S_0 = \emptyset$. Then, we have that $f^-(y)$

$$= \sum_{j=1}^n (y(s_j) - y(s_{j+1})) f(S_j) = \sum_{j=1}^n y(s_j) f(s_j | S_{j-1}). \quad (11)$$

Next, consider the base polyhedron $B(f) := \{x \in \mathbb{R}_+^{\mathcal{N}} \mid x(S) \leq f(S) \text{ for every } S \subseteq \mathcal{N}, x(\mathcal{N}) = f(\mathcal{N})\}$. Although the non-negativity of x is stated explicitly in the definition of $B(f)$, this also follows by the monotonicity of the function f , since $x(s) = x(\mathcal{N}) - x(\mathcal{N} - s) \geq f(\mathcal{N}) - f(\mathcal{N} - s) \geq 0$. For a vector $y \in \mathbb{R}_+^{\mathcal{N}}$, we have

$$f^-(y) = \max\{y^T x : x \in B(f)\}. \quad (12)$$

The next two lemmas summarize the main properties of f^- that will be used in the analysis.

Lemma 3.2. Let $y \in [0, 1]^{\mathcal{N}}$ and $A \subseteq \mathcal{N}$ such that $A \cap \text{supp}(y) = \emptyset$. Then

$$f^-(y|A) = f^-(y + A) - f(A).$$

Proof. Since $y \in [0, 1]^{\mathcal{N}}$ and $A \cap \text{supp}(y) = \emptyset$, we can order the elements of $\mathcal{N}$ in non-increasing order such that the elements of A appear before the elements of $\text{supp}(y)$. Let $i_1, \dots, i_j$ be the ordering of elements of $\text{supp}(y)$ and $S_p = \{s_{i_k} \mid 1 \leq k \leq p\}$. Then, by the chain definition (11), $f^-(y + A) = f(A) + \sum_{k=1}^j y_{i_k} f(s_{i_k} | A \cup S_{k-1}) = f(A) + f^-(y|A)$. $\square$

The next lemma follows from the definition of the Lovász extension given in (12) (i.e., $f^-(y)$ is the optimum value of $\max y^T x$ over $B(f)$).³

Lemma 3.3. Let $y, z \in \mathbb{R}_+^{\mathcal{N}}$. We have the following properties for f^- :

1. *Monotonicity:* If $y \leq z$, then $f^-(y) \leq f^-(z)$.
2. *Subadditivity:* $f^-(y) + f^-(z) \geq f^-(y + z)$.

Proof. 1. As every $x \in B(f)$ has non-negative coordinates and $y \leq z$, the optimum value of $\max\{y^T x : x \in B(f)\}$ does not decrease when the objective changes from y to z .
2. As $f^-(y)$ is defined as the maximum of a linear function over the base polyhedron $B(f)$, it is convex, and the statement follows. $\square$

3.2.2. Proof of Theorem 1.5

We now prove Theorem 1.5. Our proof strategy is similar to the one for two matroid intersection—however, instead of using intermediate sets T_i , we will now use intermediate vectors t_i .

Let z_i be the vector z and u_i be the element chosen in the i 'th iteration of the while loop. Let A_i be the set A after i iterations of the while loop. For notational convenience, let $x_i := z_i - \chi^{A_{i-1}}$ for each iteration i . We note that $x_i \geq 0$ due to the constraints in $\text{LP}(A_{i-1})$ and moreover, x_i is in the matroid polytope of $\mathcal{M}_b/A_{i-1}$ for all $b \in [k]$. We now define vectors t_i which will be useful in the analysis. Let $t_0 := t$ be a feasible point for $\text{LP}(\emptyset)$. We will define vectors t_i inductively so that $\chi^{A_i} + t_i$ is in the matroid polytope of $\mathcal{M}_b$ for every $b \in [k]$. Assuming we have A_{i-1} , x_i , and t_{i-1} , we now define t_i .

Let $b \in [k]$. Since x_i and t_{i-1} are in the matroid polytope of $\mathcal{M}_b/A_{i-1}$, they can be written as convex combinations of indicator vectors of independent sets in $\mathcal{M}_b/A_{i-1}$. Let $x_i = \sum_{J \in J_b} \lambda_J \chi^J$ and $t_{i-1} = \sum_{T \in T_b} \alpha_T \chi^T$ where J_b and T_b are collections of independent sets in $\mathcal{M}_b/A_{i-1}$, $\lambda_J, \alpha_T \geq 0$ for all $J \in J_b, T \in T_b$, and $\sum_{J \in J_b} \lambda_J = \sum_{T \in T_b} \alpha_T = 1$. Let $J \in J_b, T \in T_b$. By Proposition 3.1, there exists a function $h_b^{J,T} : J \rightarrow T \cup \{d\}$ where d is a dummy element such that

³ Although it is not needed in the present work, definitions (11) and (12) can be generalized to vectors in $\mathbb{R}^{\mathcal{N}}$. Using this generalization, both statements in Lemma 3.3 hold for every $y, z \in \mathbb{R}^{\mathcal{N}}$.

1. $T - h_b^{J,T}(u) + u \in \mathcal{I}_b/A_{i-1}$ for all $u \in J$ and
2. $h_b^{J,T}$ restricted to $(h_b^{J,T})^{-1}(T)$ is injective.

Let $u \in \text{supp}(x_i)$. Then, the point

$$\begin{aligned} r_T &:= \sum_{J \in J_b: u \in J} \left(\frac{\lambda_J}{\sum_{S \in J_b: u \in S} \lambda_S} \right) \chi^{T+u-h_b^{J,T}(u)} \\ &= \chi^{T+u} - \sum_{J \in J_b: u \in J} \left(\frac{\lambda_J}{x_i(u)} \right) \chi^{h_b^{J,T}(u)} \end{aligned}$$

is in the matroid polytope of $\mathcal{M}_b/A_{i-1}$ because it is a convex combination of characteristic vectors of independent sets in $\mathcal{M}_b/A_{i-1}$ (we use the convention that χ^d is the all zeroes vector for the dummy element d). Consequently, the convex combination of the r_T vectors for $T \in T_b$ given by

$$\sum_{T \in T_b} \alpha_T \cdot r_T = t_{i-1} + \chi^u - \sum_{T \in T_b} \alpha_T \sum_{J \in J_b: u \in J} \left(\frac{\lambda_J}{x_i(u)} \right) \chi^{h_b^{J,T}(u)}$$

is in the matroid polytope of $\mathcal{M}_b/A_{i-1}$. Now for each $u \in \text{supp}(x_i)$, define $y_b^u := \sum_{T \in T_b} \alpha_T \sum_{J \in J_b: u \in J} \left(\frac{\lambda_J}{x_i(u)} \right) \chi^{h_b^{J,T}(u)}$ and for each $a \in \mathcal{N} - A_{i-1}$, define $y^u(a) := \min \left\{ t_{i-1}(a), \sum_{b=1}^k y_b^u(a) \right\}$. We recall that u_i is the element in $\text{supp}(x_i)$ chosen by the algorithm in the i 'th iteration. We set $t_i := t_{i-1} - y^{u_i}$.

Claim 3.4. *The point $\chi^{A_i} + t_i$ is feasible for $\text{LP}(A_i)$.*

Proof. By the definition of y^{u_i} we get $0 \leq t_i \leq t_{i-1} - y_b^{u_i}$ for all $b \in [k]$. We have shown that $t_{i-1} - y_b^{u_i}$ is in the matroid polytope of $\mathcal{M}_b/A_i$ for all $b \in [k]$, so $\chi^{A_i} + t_{i-1} - y_b^{u_i}$ is feasible for $\text{LP}(A_i)$. Consequently, the point t_i is also in the matroid polytope of $\mathcal{M}_b/A_i$ for all $b \in [k]$, and $\chi^{A_i} + t_i$ is feasible for $\text{LP}(A_i)$. $\square$

We now lower bound the marginal improvement of u_i on A_{i-1} with an expression that represents the function value of $y_b^{u_i}$.

Claim 3.5. *For a fixed choice of A_{i-1} and for all $b \in [k]$, we have*

$$\mathbb{E}_{u_i}[f(u_i|A_{i-1})] \geq \mathbb{E}_{u_i}[f^-(y_b^{u_i}|A_{i-1})].$$

Proof. We first show the following inequality:

$$\begin{aligned} &\mathbb{E}_{u_i}[f(u_i|A_{i-1})] \\ &\geq \mathbb{E}_{u_i} \left[\sum_{T \in T_b} \alpha_T \sum_{J \in J_b: u_i \in J} \left(\frac{\lambda_J}{x_i(u_i)} \right) f(h_b^{J,T}(u_i)|A_{i-1}) \right]. \end{aligned} \quad (13)$$

Expanding the RHS using the probability distribution on u_i shows that $\left(\sum_{v \in \mathcal{N} - A_{i-1}} x_i(v) \right) \text{RHS}$

$$\begin{aligned} &= \sum_{u \in \mathcal{N} - A_{i-1}} \left(\sum_{T \in T_b} \alpha_T \sum_{J \in J_b: u \in J} \lambda_J f(h_b^{J,T}(u)|A_{i-1}) \right) \\ &= \sum_{T \in T_b, J \in J_b} \alpha_T \lambda_J \left(\sum_{u \in J} f(h_b^{J,T}(u)|A_{i-1}) \right) \\ &\leq \sum_{T \in T_b, J \in J_b} \alpha_T \lambda_J \left(\sum_{w \in T} f(w|A_{i-1}) \right) \quad (14) \\ &= \sum_{T \in T_b} \alpha_T \left(\sum_{w \in T} f(w|A_{i-1}) \right) \\ &= \sum_{w \in \text{supp}(t_{i-1})} f(w|A_{i-1}) \sum_{T \in T_b: w \in T} \alpha_T \\ &= \sum_{w \in \text{supp}(t_{i-1})} f(w|A_{i-1}) \cdot t_{i-1}(w). \end{aligned}$$

Inequality (14) follows by the injectivity of $h_b^{J,T}$. Since $\chi^{A_{i-1}} + t_{i-1}$ is feasible for $\text{LP}(A_{i-1})$ and $\chi^{A_{i-1}} + x_i$ is an optimal solution for $\text{LP}(A_{i-1})$, the previous expression is at most

$$\begin{aligned} &\sum_{u \in \mathcal{N} - A_{i-1}} f(u|A_{i-1}) \cdot x_i(u) = \mathbb{E}_{u_i}[f(u_i|A_{i-1})] \sum_{u \in \mathcal{N} - A_{i-1}} x_i(u) \\ &= \left(\sum_{u \in \mathcal{N} - A_{i-1}} x_i(u) \right) \text{LHS}, \end{aligned}$$

thus proving inequality (13).

Let $u \in \text{supp}(x_i)$. Then, we have that

$$\begin{aligned} &\sum_{T \in T_b} \alpha_T \sum_{J \in J_b: u \in J} \left(\frac{\lambda_J}{x_i(u)} \right) f(h_b^{J,T}(u)|A_{i-1}) \\ &= \left(\sum_{T \in T_b} \alpha_T \sum_{J \in J_b: u \in J} \left(\frac{\lambda_J}{x_i(u)} \right) f(h_b^{J,T}(u) + A_{i-1}) \right) \\ &\quad - f(A_{i-1}). \end{aligned}$$

We observe that the point y_b^u is a convex combination of indicator vectors of sets $\{h_b^{J,T}(u)\}$ for $T \in T_b$ and $J \in J_b$. By definition of Lovász extension, this implies that

$$\begin{aligned} &\sum_{T \in T_b} \alpha_T \sum_{J \in J_b: u \in J} \left(\frac{\lambda_J}{x_i(u)} \right) f(h_b^{J,T}(u) + A_{i-1}) \\ &\geq f^-(y_b^u + A_{i-1}). \end{aligned}$$

By Lemma 3.2, we have that $f^-(y_b^u) - f(A_{i-1}) = f^-(y_b^u|A_{i-1})$. Hence, using inequality (13), we have that $\mathbb{E}_{u_i}[f(u_i|A_{i-1})]$

$$\geq \mathbb{E}_{u_i}[f^-(y_b^{u_i} + A_{i-1})] - f(A_{i-1}) = \mathbb{E}_{u_i}[f^-(y_b^{u_i}|A_{i-1})].$$

$\square$

Now we have all the ingredients needed to bound the approximation factor.

Theorem 1.5. *Let A be the set returned by MATROIDINTERSECTIONRRGREEDY and let t be any feasible point for $LP(\emptyset)$. Then,*

$$\mathbb{E}[f(A)] \geq \left(\frac{1}{k+1}\right) f^-(t).$$

Proof. Consider a fixed choice of A_{i-1} . By Claim 3.5, we have that

$$k\mathbb{E}_{u_i}[f(u_i|A_{i-1})] \geq \sum_{b=1}^k \mathbb{E}_{u_i}[f^-(y_b^{u_i}|A_{i-1})]. \quad (15)$$

Using Lemma 3.3 (subadditivity and monotonicity of f^-), we get that

$$\mathbb{E}_{u_i} \left[\sum_{b=1}^k f^-(y_b^{u_i}|A_{i-1}) \right] \geq \mathbb{E}_{u_i} \left[f^-\left(\sum_{b=1}^k y_b^{u_i}|A_{i-1}\right) \right] \quad (16)$$

$$\geq \mathbb{E}_{u_i}[f^-(y^{u_i}|A_{i-1})]. \quad (17)$$

Now by Lemma 3.3 (subadditivity) we have, $\mathbb{E}_{u_i}[f^-(y^{u_i}|A_{i-1})] \geq \mathbb{E}_{u_i}[f^-(y^{u_i} + t_i|A_{i-1})] - \mathbb{E}_{u_i}[f^-(t_i|A_{i-1})] = f^-(t_{i-1}|A_{i-1}) - \mathbb{E}_{u_i}[f^-(t_i|A_{i-1})]$. We note that $\text{supp}(t_{i-1}) \cap A_{i-1} = \emptyset$ and $\text{supp}(t_i) \cap A_{i-1} = \emptyset$. By Lemma 3.2, we have that

$$\begin{aligned} f^-(t_{i-1}|A_{i-1}) - \mathbb{E}_{u_i}[f^-(t_i|A_{i-1})] \\ = f^-(t_{i-1} + A_{i-1}) - \mathbb{E}_{u_i}[f^-(t_i + A_{i-1})]. \end{aligned}$$

By monotonicity, we get

$$\begin{aligned} f^-(t_{i-1} + A_{i-1}) - \mathbb{E}_{u_i}[f^-(t_i + A_{i-1})] \\ \geq f^-(t_{i-1} + A_{i-1}) - \mathbb{E}_{u_i}[f^-(t_i + A_i)]. \end{aligned}$$

Thus, we obtain that $k\mathbb{E}_{u_i}[f(u_i|A_{i-1})] \geq f^-(t_{i-1} + A_{i-1}) - \mathbb{E}_{u_i}[f^-(t_i + A_i)]$. Consequently, for any fixed choice of A_{i-1} , we have

$$\begin{aligned} k\mathbb{E}_{u_i}[f(A_i)] + \mathbb{E}_{u_i}[f^-(t_i + A_i)] \\ \geq kf(A_{i-1}) + f^-(t_{i-1} + A_{i-1}). \end{aligned}$$

So, this inequality also holds when we take expectation over A_{i-1} . Thus, we have $k\mathbb{E}[f(A_i)] + \mathbb{E}[f^-(t_i + A_i)] \geq k\mathbb{E}[f(A_{i-1})] + \mathbb{E}[f^-(t_{i-1} + A_{i-1})]$ which implies that $k\mathbb{E}[f(A)] + \mathbb{E}[f^-(A)] = (k+1)\mathbb{E}[f(A)] \geq f^-(t_0) = f^-(t)$. $\square$

Remark 3.6. We note that the above analysis can also be extended to show that GREEDY outputs a common independent set A with $f(A) \geq (1/(k+1))f^-(t)$ for any point t that is feasible to $LP(\emptyset)$. This provides a stronger guarantee for the approximation factor of GREEDY as it is relative to the optimum value of the Lovász relaxation over the intersection of matroid polytopes.

Corollary 3.6.1. *If f is modular, then $\mathbb{E}[f(A)] \geq \frac{f^-(t)}{k}$, where A is the set returned by MATROIDINTERSECTIONRRGREEDY and t is any point that is feasible for $LP(\emptyset)$.*

Proof. For f being modular, the function f^- is simply the natural extension of linear functions to fractional domains, so we have that $f^-(x) = \sum_{i=1}^n f(s_i)x_i$. In the proof of Theorem 1.5, we showed that $k\mathbb{E}[f(u_i|A_{i-1})] \geq \mathbb{E}[f^-(y^{u_i}|A_{i-1})]$ (using inequalities (15) and (17) and taking expectation over A_{i-1}). Since f is modular this can be simplified to $k\mathbb{E}[f(u_i)] \geq \mathbb{E}[f^-(y^{u_i})]$. Now we take the sum over all u_i and get,

$$\begin{aligned} k\mathbb{E}[f(A)] &= k\mathbb{E} \left[\sum_{u_i \in A} f(u_i) \right] \geq \mathbb{E} \left[\sum_{u_i \in A} f^-(y^{u_i}) \right] \\ &= f^-(t). \quad \square \end{aligned}$$

Remark 3.7. In Corollary 3.6.1, since we use f^- and allow t to be fractional, the $(1/k)$ -approximation is relative to the optimum of the LP relaxation of maximizing a modular function subject to a k -matroid intersection constraint. This again matches the guarantee of Greedy.

4. Conclusions and open problems

In this work, we analyzed the approximation factor of RRGREEDY for monotone submodular maximization subject to matroid and matroid intersection constraints. Our work raises many interesting open questions. Firstly, we showed that RRGREEDY achieves $(1/2)$ -approximation in the worst-case (not just in expectation) for matroid constraint. An intriguing open question here is whether RRGREEDY beats the $(1/2)$ -factor in expectation. Indeed, it beats the $(1/2)$ -factor in expectation for partition matroid [7]. We conjecture that it beats the $(1/2)$ -factor in expectation for any matroid. If true this would imply a better deterministic approximation factor for the algorithm in [2] as their analysis uses the fact that the approximation factor of RRGREEDY is $1/2$.

Another intriguing direction is to understand whether the approximation factor of RRGREEDY is at least as good as GREEDY even for specific matroids. In particular, for the uniform matroid (i.e., for cardinality constraint), we recall that GREEDY achieves a $(1 - 1/e)$ -approximation [3] and $(1/\alpha)(1 - e^{-\alpha})$ for functions with curvature α [6]. Does RRGREEDY also achieve these same guarantees in expectation? This is indeed known for the case of $\alpha = 1$ [1].

We note that 2MATROIDINTERSECTIONRRGREEDY achieves an optimum when the function is modular (i.e., $\alpha = 0$) and achieves a $(1/3)$ -approximation in expectation for arbitrary monotone submodular functions (i.e., when $\alpha = 1$). A natural open question here is whether MATROIDINTERSECTIONRRGREEDY achieves a $(1/(k-1+2\alpha))$ -approximation for maximizing monotone submodular functions with curvature α subject to k matroid intersection constraints. We note that this conjecture if true, would also give an algorithm to obtain $(1/(k-1))$ -approximation for maximizing a linear function subject to k matroid intersection constraints.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Karthekeyan Chandrasekaran was supported by NSF grants - CCF-1814613 and CCF-1907937. Kristóf Bérczi was supported by the János Bolyai Research Fellowship. Kristóf Bérczi and Tamás Király were supported by the Lendület Programme of the Hungarian Academy of Sciences—grant number LP2021-1/2021 and by the Hungarian National Research, Development and Innovation Office – NKFIH, grant numbers FK128673, K120254, and TKP2020-NKA-06.

Appendix A

A.1. Upper bound on 2MATROIDINTERSECTIONRRGREEDY approximation ratio

We now describe a computer generated example that upper bounds the expected approximation ratio of 2MATROIDINTERSECTIONRRGREEDY by 13/33. To generate the example, we fix two matroids and fix the choices made by the algorithm. The choices of the algorithm include the set M_i and the element u_i that is chosen in each iteration i . For fixed choices of the algorithm, our goal is to find a monotone submodular function that is consistent with the choices of the algorithm and that will minimize the approximation ratio of the algorithm. The choices of the algorithm are determined by the function values of subsets of $\mathcal{N}$, so we create decision variables $f(S)$ (that will correspond to the function value $f(S)$) for all $S \subseteq \mathcal{N}$ and encode the decisions of the algorithm with linear constraints. We also ensure that the values of the variables correspond to a monotone submodular function by imposing linear constraints. Since all constraints are linear, we can write a linear program (LP) that minimizes the approximation ratio of the algorithm.

We now describe the choice of two matroids that we use to generate our example. The common independent sets of two matroids will correspond to matchings in a bipartite graph. The bipartite graph that we use is a cycle with $2t$ edges for some parameter t . The elements of the ground set $\mathcal{N} := [2t]$ represent the edges of the cycle in order.

Next, we describe the choices of the algorithm and an optimal solution. We will set up the example so that 2MATROIDINTERSECTIONRRGREEDY will always output the matching $A = \{j \in \mathcal{N} | j \bmod 2 = 1\}$ irrespective of the random choices made by the algorithm. We will also set up the example so that the matching $B = \{j \in \mathcal{N} | j \bmod 2 = 0\}$ is the optimal solution.

Let $\mathcal{C}$ denote the set of maximal matchings in the graph and for any matching F , let $\mathcal{C}(F)$ denote the set of maximal matchings that contain F . Maximal matchings are used to reduce the number of constraints in the LP. The example is generated by solving the following LP that has

a variable representing the function value for every subset of $\mathcal{N}$.

$$\begin{aligned} & \text{minimize } f(A) \\ & \sum_{u \in A - M'} f(u|M') \geq \sum_{u \in S - M'} f(u|M') \end{aligned} \quad (18)$$

$$\forall M' \subseteq A, \forall S \in \mathcal{C}(M')$$

$$f(B) \geq f(X) \forall X \in \mathcal{C} \quad (19)$$

$$f(B) = 1 \quad (20)$$

$$f(u|S) \geq 0 \forall S \subseteq \mathcal{N}, \forall u \in \mathcal{N} - S \quad (21)$$

$$f(X + x_1) + f(X + x_2) \geq f(X + x_1 + x_2) + f(X)$$

$$\forall X \subseteq \mathcal{N}, \forall x_1, x_2 \in \mathcal{N} - X \quad (22)$$

$$f(\emptyset) = 0 \quad (23)$$

Constraint (19) ensures optimality of B . Constraints (21) and (22) ensure that the function values correspond to a monotone submodular function. Constraints (20) and (23) are for the purpose of normalization. We now show using Constraint (18) that 2MATROIDINTERSECTIONRRGREEDY will indeed return A .

Claim A.1. For all $M' \subseteq A$, the set $S = A - M'$ is an optimum solution to the following:

$$\max \left\{ \sum_{u \in S} f(u|M') | S \cap M' = \emptyset, S \cup M' \text{ is a matching} \right\}.$$

Proof. Let F be a matching such that $F \cup M'$ is also matching. Let F' be the set acquired by adding edges to F so that it is a maximal matching. Hence, $F' \in \mathcal{C}(M')$. By constraint (18) and by monotonicity of f , we have that

$$\sum_{u \in A - M'} f(u|M') \geq \sum_{u \in F'} f(u|M') \geq \sum_{u \in F} f(u|M'). \quad \square$$

Claim A.2. 2MATROIDINTERSECTIONRRGREEDY will always output A .

Proof. We prove by induction on i that 2MATROIDINTERSECTIONRRGREEDY will always choose elements of A in iterations $1, \dots, i$. Let $i = 1$ and $M' = \emptyset$. Now Claim A.1 implies that $A = \arg \max \{ \sum_{u \in S} f(u) | S \text{ is a matching} \}$, so the algorithm can choose an element of A in the first iteration. Now assume that $A_{i-1} \subseteq A$ is the set of elements chosen by the algorithm in the first $i - 1$ iterations. The set $A - A_{i-1}$ is a valid candidate for M_i and Claim A.1 implies $A - A_{i-1} = \arg \max \{ \sum_{u \in S - A_{i-1}} f(u|A_{i-1}) | S \cup A_{i-1} \text{ is a matching} \}$. So, the algorithm can choose M_i to be $A - A_{i-1}$ and the algorithm chooses an element of A in iteration i . Hence, we have $A_i \subseteq A$. $\square$

Lemma A.3. The approximation ratio of 2MATROIDINTERSECTIONRRGREEDY is upper bounded by 13/33.

Proof. Solving the LP for $t = 6$ generates a solution where $f(A) = 13/33$. $\square$

References

- [1] N. Buchbinder, M. Feldman, J.S. Naor, R. Schwartz, Submodular maximization with cardinality constraints, in: *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA*, 2014, pp. 1433–1452.
- [2] N. Buchbinder, M. Feldman, M. Garg, Deterministic $(1/2 + \epsilon)$ -approximation for submodular maximization over a matroid, in: *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA*, 2019, pp. 241–254.
- [3] G. Nemhauser, L. Wolsey, M. Fisher, An analysis of approximations for maximizing submodular set functions - I, *Math. Program.* 14 (1978) 265–294.
- [4] G. Nemhauser, L. Wolsey, Best algorithms for approximating the maximum of a submodular set function, *Math. Oper. Res.* 3 (1978) 177–188.
- [5] G. Călinescu, C. Chekuri, M. Pál, J. Vondrák, Maximizing a monotone submodular function subject to a matroid constraint, *SIAM J. Comput.* 40 (2011) 1740–1766.
- [6] M. Conforti, G. Cornuéjols, Submodular set functions, matroids and the greedy algorithm: tight worst-case bounds and some generalizations of the rado-edmonds theorem, *Discrete Appl. Math.* 7 (1984) 251–274.
- [7] N. Buchbinder, M. Feldman, Y. Filmus, M. Garg, Online submodular maximization: beating $1/2$ made simple, in: *Integer Programming and Combinatorial Optimization, IPCO*, 2019, pp. 101–114.
- [8] E.L. Lawler, Matroid intersection algorithms, *Math. Program.* 9 (1975) 31–56.
- [9] L. Lovász, Submodular functions and convexity, in: *Mathematical Programming the State of the Art*, Springer, 1983, pp. 235–257.
- [10] A.A. Bian, J.M. Buhmann, A. Krause, S. Tschischek, Guarantees for greedy maximization of non-submodular functions with applications, in: *Proceedings of the 34th International Conference on Machine Learning*, vol. 70, ICML, 2017, pp. 498–507.
- [11] J. Vondrák, Submodularity and curvature: the optimal algorithm, in: *RIMS Kokyuroku Bessatsu*, vol. B23, RIMS, 2010, pp. 253–266.
- [12] M. Sviridenko, J. Vondrák, J. Ward, Optimal approximation for submodular and supermodular optimization with bounded curvature, in: *Proceedings of the 26th Annual ACM-SIAM Symposium on Discrete Algorithms, SODA*, 2015, pp. 1134–1148.
- [13] J. Lee, M. Sviridenko, J. Vondrák, Submodular maximization over multiple matroids via generalized exchange properties, *Math. Oper. Res.* 35 (2010) 795–806.
- [14] L.C. Lau, R. Ravi, M. Singh, *Iterative Methods in Combinatorial Optimization*, Cambridge University Press, 2011.
- [15] A. Linhares, N. Olver, C. Swamy, R. Zenklusen, Approximate multi-matroid intersection via iterative refinement, in: *Integer Programming and Combinatorial Optimization, IPCO*, 2019, pp. 299–312.
- [16] H. Whitney, On the abstract properties of linear dependence, in: *Hasler Whitney Collected Papers*, Springer, 1992, pp. 147–171.
- [17] H. Nishimura, S. Kuroda, A. Lost Mathematician, Takeo Nakasawa, *The Forgotten Father of Matroid Theory*, Springer Science & Business Media, 2009.
- [18] A. Frank, *Connections in Combinatorial Optimization*, vol. 38, OUP, Oxford, 2011.
- [19] W.H. Cunningham, Testing membership in matroid polyhedra, *J. Comb. Theory, Ser. B* 36 (1984) 161–188.