Inequalities on Submodular Functions via Term Rewriting

Yuval Filmus[a,1,*]

[a] *University of Toronto*

Abstract

We devise a method for proving inequalities on submodular functions, with a term rewriting flavour. Our method comprises of the following steps: (1) start with a linear combination X of the values of the function; (2) define a set of simplification rules; (3) conclude that $X \geq Y$, where Y is a linear combination of a small number of terms which cannot be simplified further; (4) calculate the coefficients of Y by evaluating X and Y on functions on which the inequality is tight.

The crucial third step is non-constructive, since it uses compactness of the dual cone of submodular functions. Its proof uses the classical uncrossing technique with a quadratic potential function. We prove several inequalities using our method, and use them to tightly analyze the performance of two natural (but non-optimal) algorithms for submodular maximization, the random set algorithm and local search.

Keywords: inequalities, submodularity, local search

1. Introduction

This paper concerns submodular functions [1, 2, 3]. These are functions $f\colon 2^{\mathcal{U}} \to \mathbb{R}$ for some ground set $\mathcal{U}$ (we say that f is a set-function on $\mathcal{U}$), satisfying the submodular inequality:

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B), \quad A, B \subseteq \mathcal{U}. \tag{1}$$

We will be interested in inequalities satisfied by submodular functions. A simple example is the inequality

$$\mathop{\mathbb{E}}_{A \in \binom{\mathcal{U}}{k}} f(A) \geq \frac{k}{n} f(\mathcal{U}) + \left(1 - \frac{k}{n}\right) f(\emptyset), \tag{2}$$

where n is the cardinality of $\mathcal{U}$, and $\binom{\mathcal{U}}{k}$ consists of all subsets of $\mathcal{U}$ of size k. (The symbol $\mathbb{E}$ denotes expectation.)

*Corresponding author.
Email address: `yuvalf@cs.toronto.edu` (Yuval Filmus)
[1]Supported by NSERC.

Inequality (2) can be proved in many ways. Here is how our method goes about proving it. The starting point is the left-hand side. Each time we apply an instance of inequality (1), we replace one term by another term which is simpler in some precise sense (see §2). The only terms that cannot be simplified further are of the form $\alpha f(\mathcal{U}) + \beta f(\emptyset)$, and so we conclude that for some α, β,

$$\mathop{\mathbb{E}}_{A \in \binom{\mathcal{U}}{k}} f(A) \geq \alpha f(\mathcal{U}) + \beta f(\emptyset). \tag{3}$$

In order to compute α, β, we consider modular functions, which are functions in which (1) holds with equality. If f is a modular function, then (3) must also hold with equality. By picking appropriate modular functions, we can evaluate the coefficients α, β.

Inequality (2) is a particularly simple example, since we allowed the proof to use arbitrary instances of the submodular inequality (1). Other cases necessitate limiting the instances that are used.

Organization. We present the method, together with a few examples, in §2. We interpret the example inequalities in the context of algorithms for submodular maximization in §3.

Notation. For $n \in \mathbb{N}$, we use $[n] = \{1, \ldots, n\}$. We denote set difference by $\setminus$. We denote symmetric difference by $A \triangle B = (A \setminus B) \cup (B \setminus A)$. For a set S, the complete permutation group on S is denoted $\mathrm{Sym}(S)$. The transposition exchanging i and j is denoted (i, j). For a set S and integer k, $\binom{S}{k}$ stands for the family of all subsets of S of cardinality k. The expression $\mathop{\mathbb{E}}_{A \in \mathcal{I}}$ denotes expectation over a random variable A distributed uniformly over $\mathcal{I}$.

Acknowledgements. The author would like to thank several anonymous referees for bringing to his attention several references and for many helpful suggestions.

2. Method

In order to describe our method in full generality, we need to make several definitions (§2.1). Then we state and prove the method (§2.2). Our proof uses the classical technique of uncrossing with a quadratic potential function (see for example Schrijver [4, Chapter 49]). Finally, we give several examples (§2.3), including a proof of (2). We also discuss some similar results due to Vondrák [5] concerning the multilinear relaxation (§2.4).

2.1. Setup

The data for our method consists of the following:

- A ground set $\mathcal{U}$.
- A permutation group $\mathcal{G}$ acting on $\mathcal{U}$.
- A symmetric relation $\circledast$ (read: star) on $2^{\mathcal{U}}$ satisfying the following property: if $A \circledast B$ then $A \not\subseteq B$ and $B \not\subseteq A$ (if this holds, we say that A, B form an *antichain*).

- A subset $\mathcal{I} \subseteq 2^{\mathcal{U}}$ satisfying the following property: if $A, B \in \mathcal{I}$ and $A \circledast B$ then $A \cup B, A \cap B \in \mathcal{I}$.

We extend the action of $\mathcal{G}$ to subsets of $\mathcal{U}$: for $A \subseteq \mathcal{U}$ and $\pi \in \mathcal{G}$, we define $A^\pi = \{x^\pi : x \in A\}$, where x^π is the action of π applied to x.

Given the foregoing, we define the cone $\mathcal{C}(\mathcal{I}, \mathcal{G}) \subseteq \mathbb{R}^{\mathcal{I}}$ as the set of all non-negative vectors $X = (X_A)_{A \in \mathcal{I}}$ which are symmetric under the action of $\mathcal{G}$, that is, $X_A = X_{A^\pi}$ for all $A \in \mathcal{I}$ and $\pi \in \mathcal{G}$. The *magnitude* of a vector $X \in \mathcal{C}(\mathcal{I}, \mathcal{G})$ is $\sum_A X_A$. If $X \in \mathcal{C}(\mathcal{I}, \mathcal{G})$ and f is a set-function on $\mathcal{U}$, then $X|_f \in \mathbb{R}$ is defined by

$$X|_f = \sum_{A \in \mathcal{I}} X_A f(A).$$

For $A \in \mathcal{I}$, the vector $\chi^A \in \mathcal{C}(\mathcal{I}, \mathcal{G})$ has a single non-zero coordinate $\chi^A_A = 1$.

We say that a set $A \subseteq \mathcal{U}$ is $(\mathcal{G}, \circledast)$-*stable* if there does *not* exist $\pi \in \mathcal{G}$ such that $A \circledast A^\pi$.

We say that a set-function f on $\mathcal{U}$ is $\circledast$-*modular* if whenever $A \circledast B$, we have $f(A) + f(B) = f(A \cup B) + f(A \cap B)$.

2.2. Method

Our method consists of the following proposition.

Proposition 1. *Let $\mathcal{U}, \mathcal{G}, \circledast, \mathcal{I}$ be data for the method, as described in §2.1. For every $X \in \mathcal{C}(\mathcal{I}, \mathcal{G})$ there exists $Y \in \mathcal{C}(\mathcal{I}, \mathcal{G})$ of the same magnitude as X such that*

(a) The support of Y consists of $(\mathcal{G}, \circledast)$-stable sets.

(b) If f is submodular then $X|_f \geq Y|_f$.

(c) If f is $\circledast$-modular then $X|_f = Y|_f$.

The proof of the proposition will require the following lemma.

Lemma 2. *Let $\mathcal{U}, \mathcal{G}, \circledast, \mathcal{I}$ be data for the method. Define a mapping $\sigma \colon \mathcal{C}(\mathcal{I}, \mathcal{G}) \to \mathbb{R}$ by*

$$\sigma(X) = \sum_{A \in \mathcal{I}} X_A |A|^2.$$

Suppose A, B are an antichain, $Y \in \mathcal{C}(\mathcal{I}, \mathcal{G})$ and $t > 0$. Define

$$Z = Y + t(\chi^{A \cup B} + \chi^{A \cap B} - \chi^A - \chi^B).$$

If $Z \in \mathcal{C}(\mathcal{I}, \mathcal{G})$ then $\sigma(Z) > \sigma(Y)$.

Proof. Let $|A \setminus B| = a$, $|B \setminus A| = b$ and $|A \cap B| = c$. We have

$$\sigma(Z) - \sigma(Y) = t[(a + b + c)^2 + c^2 - (a + c)^2 - (b + c)^2] = 2tab > 0,$$

using the fact that A, B are an antichain. $\square$

Proof of Proposition 1. Consider the following subset of $\mathcal{C}(\mathcal{I}, \mathcal{G})$:

$$\mathcal{D} = \left\{ X + \sum_{\substack{A,B \in \mathcal{I} \\ A \circledast B}} t_{A,B}(\chi^{A \cup B} + \chi^{A \cap B} - \chi^A - \chi^B) : t_{A,B} \geq 0 \right\} \cap \mathcal{C}(\mathcal{I}, \mathcal{G}).$$

Intuitively, $\mathcal{D}$ results from iteratively applying to X the uncrossing operation which is the subject of Lemma 2.

Since the magnitude of all elements in $\mathcal{D}$ is the same, $\mathcal{D}$ is compact. Therefore σ attains its maximum at some point $Y \in \mathcal{D}$. We claim that the support of Y consists of sets which are $(\mathcal{G}, \circledast)$-stable.

Suppose to the contrary that $Y_A = t \neq 0$ for some A which is not $(\mathcal{G}, \circledast)$-stable. Thus $A \circledast B$ for some set $B = A^\pi$, where $\pi \in \mathcal{G}$. Since $Y \in \mathcal{C}(\mathcal{I}, \mathcal{G})$, it is symmetric with respect to $\mathcal{G}$, and so

$$Z = Y + t \mathop{\mathbb{E}}_{\pi \in \mathcal{G}} (\chi^{A^\pi \cup B^\pi} + \chi^{A^\pi \cap B^\pi} - \chi^{A^\pi} - \chi^{B^\pi}) \in \mathcal{D}.$$

Note that $(A \cup B)^\pi = A^\pi \cup B^\pi$ and $(A \cap B)^\pi = A^\pi \cap B^\pi$. Lemma 2 then shows that $\sigma(Z) > \sigma(Y)$, contradicting our choice of Y.

The other two properties are immediate from the definition of $\mathcal{D}$. $\square$

It is instructive at this point to give an example showing the non-constructive nature of the proof. For a ground set $\mathcal{U}$, let

$$F(k) = \mathop{\mathbb{E}}_{A \in \binom{\mathcal{U}}{k}} f(A).$$

This allows us to rewrite inequality (2) as

$$F(k) \geq \frac{k}{n} F(n) + \left(1 - \frac{k}{n}\right) F(0). \tag{4}$$

For any $A \in \binom{\mathcal{U}}{k-1}$ and any $x, y \in \mathcal{U} \setminus A$ such that $x \neq y$, submodularity implies that

$$f(A \cup \{x\}) + f(A \cup \{y\}) \geq f(A \cup \{x, y\}) + f(A).$$

Averaging over all possible choices of A, x, y, we deduce

$$2F(k) \geq F(k+1) + F(k-1).$$

Here is one attempt at proving inequality (4) for $n = 3$ and $k = 2$:

$$\begin{aligned} F(2) &\geq \frac{1}{2} F(3) + \frac{1}{2} F(1) \\ &\geq \frac{1}{2} F(3) + \frac{1}{4} F(2) + \frac{1}{4} F(0) \\ &\geq \frac{5}{8} F(3) + \frac{1}{8} F(1) + \frac{1}{4} F(0) \\ &\geq \frac{5}{8} F(3) + \frac{1}{16} F(2) + \frac{5}{16} F(0) \\ &\geq \cdots \end{aligned}$$

Continuing this way, after infinitely many steps we reach the desired inequality

$$F(2) \geq \frac{2}{3}F(3) + \frac{1}{3}F(0).$$

Our method guarantees that this approach always terminates, and also implies that the inequality reached this way has a finitary proof.

2.3. Examples

We illustrate our method by proving four different inequalities. When describing data for the method, we omit $\mathcal{U}$ when it is clear from the context, and X is always chosen so that the left-hand side of the inequality we are trying to prove is $X|_f$. The simplest is (2).

Lemma 3. *Let f be a submodular function on $\mathcal{U}$, where $|\mathcal{U}| = n$. For $0 \leq k \leq n$,*

$$\mathop{\mathbb{E}}_{A \in \binom{\mathcal{U}}{k}} f(A) \geq \frac{k}{n} f(\mathcal{U}) + \left(1 - \frac{k}{n}\right) f(\emptyset).$$

Proof. We apply the method with the following data:

- $\mathcal{G} = \mathrm{Sym}(\mathcal{U})$.
- $A \circledast B$ if A, B is an antichain.
- $\mathcal{I} = 2^{\mathcal{U}}$.

Proposition 1 gives us a vector Y supported on $(\mathcal{G}, \circledast)$-stable sets. Clearly the sets $\mathcal{U}, \emptyset$ are $(\mathcal{G}, \circledast)$-stable. Any other set A is not $(\mathcal{G}, \circledast)$-stable since $A \circledast A^{(x,y)}$ for any two elements $x \in A$ and $y \notin A$. Hence

$$Y|_f = \alpha f(\mathcal{U}) + (1 - \alpha) f(\emptyset)$$

for some $\alpha \in \mathbb{R}$ (recall that X and Y have the same magnitude). It is easy to check that the function $f(A) = |A|$ is modular and so $\circledast$-modular. Simple calculation gives $X|_f = k$ and $Y|_f = \alpha n$, hence $\alpha = k/n$. $\square$

The following example requires $\mathcal{G}$ to be smaller than the complete group, and also requires $\circledast$ to be defined in a non-trivial way. This is always the case when the inequality is tight for some function which is not modular.

Lemma 4. *Let f be a submodular function on $\mathcal{U}$. Suppose $\mathcal{U}$ is partitioned into two non-empty parts P_1, P_2 of sizes n_1, n_2. For $0 \leq k_1 \leq n_1$ and $0 \leq k_2 \leq n_2$,*

$$\mathop{\mathbb{E}}_{\substack{A \subseteq \mathcal{U} \\ |A \cap P_1| = k_1 \\ |A \cap P_2| = k_2}} f(A) \geq \frac{k_1 k_2}{n_1 n_2} f(\mathcal{U}) + \frac{k_1(n_2 - k_2)}{n_1 n_2} f(P_1) + \frac{(n_1 - k_1)k_2}{n_1 n_2} f(P_2) + \frac{(n_1 - k_1)(n_2 - k_2)}{n_1 n_2} f(\emptyset).$$

Proof. We apply the method with the following data:

- $\mathcal{G} = \mathrm{Sym}(P_1) \times \mathrm{Sym}(P_2)$.
- $A \circledast B$ if (1) A, B is an antichain and (2) either $A \cap P_1 = B \cap P_1$ or $A \cap P_2 = B \cap P_2$.
- $\mathcal{I} = 2^{\mathcal{U}}$.

Proposition 1 gives us a vector Y supported on $(\mathcal{G}, \circledast)$-stable sets. Clearly the sets $\mathcal{U}, P_1, P_2, \emptyset$ are $(\mathcal{G}, \circledast)$-stable. Any other set A is not $(\mathcal{G}, \circledast)$-stable since $A \circledast A^{(x,y)}$ for any two elements $x \in A$ and $y \notin A$ belonging to the same part. Hence

$$Y|_f = \alpha f(\mathcal{U}) + \beta f(P_1) + \gamma f(P_2) + (1 - \alpha - \beta - \gamma) f(\emptyset)$$

for some $\alpha, \beta, \gamma \in \mathbb{R}$.

Consider the function

$$f(A) = |A \cap P_1| \cdot |A \cap P_2|.$$

It is easy to check that f is $\circledast$-modular, though not modular (consider $A = P_1$ and $B = P_2$). We have

$$k_1 k_2 = X|_f = Y|_f = \alpha n_1 n_2.$$

Hence $\alpha = k_1 k_2 / n_1 n_2$.

Next, it is easy to check that the function $g(A) = |A \cap P_1|$ is modular and so $\circledast$-modular. Simple calculation gives

$$k_1 = X|_g = Y|_g = (\alpha + \beta) n_1.$$

Hence $\beta = \alpha - k_1/n_1$. Similarly, the function $h(A) = |A \cap P_2|$ implies $\gamma = \alpha - k_2/n_2$. $\square$

The next example also requires $\mathcal{G}$ and $\circledast$ to be defined in a non-trivial way. The definition of $\circledast$ is more complicated this time.

Lemma 5. *Let f be a submodular function on $\mathcal{U}$, where $|\mathcal{U}| = n$. Suppose $\mathcal{U}$ is partitioned into two non-empty parts P_1, P_2. For $1 \le k \le n-1$,*

$$\mathop{\mathbb{E}}_{A \in \binom{\mathcal{U}}{k}} f(A) \ge \frac{k(k-1)}{n(n-1)} f(\mathcal{U}) + \frac{k(n-k)}{n(n-1)} (f(P_1) + f(P_2)) + \frac{(n-k)(n-k-1)}{n(n-1)} f(\emptyset).$$

Proof. We apply the method with the following data:

- $\mathcal{G} = \mathrm{Sym}(P_1) \times \mathrm{Sym}(P_2)$.
- $A \circledast B$ if (1) A, B is an antichain and (2) for all $p_1 \in P_1, p_2 \in P_2$ it is *not* the case that $A \cap \{p_1, p_2\} = \{p_1\}$ and $B \cap \{p_1, p_2\} = \{p_2\}$ or vice versa.
- $\mathcal{I} = 2^{\mathcal{U}}$.

Proposition 1 gives us a vector Y supported on $(\mathcal{G}, \circledast)$-stable sets. The sets $\mathcal{U}, P_1, P_2, \emptyset$ are clearly $(\mathcal{G}, \circledast)$-stable. We claim that these are the only ones. Indeed, consider an arbitrary subset $Q = Q_1 \cup Q_2$ of $\mathcal{U}$ where $Q_1 \subseteq P_1$, $Q_2 \subseteq P_2$, and (without loss of generality) $Q_1 \notin \{\emptyset, P_1\}$. Then $Q \circledast Q^{(x,y)}$ for any two elements $x \in Q_1$ and $y \in P_1 \setminus Q_1$, since both sets contain the same elements of P_2. We deduce that for some $\alpha, \beta, \gamma \in \mathbb{R}$,

$$Y|_f = \alpha f(\mathcal{U}) + \beta f(P_1) + \gamma f(P_2) + (1 - \alpha - \beta - \gamma) f(\emptyset).$$

Pick $x_1 \in P_1, x_2 \in P_2$, and define

$$f(A) = \begin{cases} 1 & \text{if } x_1 \in A \text{ and } x_2 \notin A, \\ 0 & \text{otherwise.} \end{cases}$$

Simple case analysis shows that f is $\circledast$-modular. Indeed, if $f(A) = f(B) = 1$ then $f(A \cup B) = f(A \cap B) = 1$, and similarly if $f(A) = f(B) = 0$ then $f(A \cup B) = f(A \cap B) = 0$. When $f(A) = 1$ and $f(B) = 0$, we could have $f(A \cup B) = f(A \cap B) = 0$ only if $A \cap \{x_1, x_2\} = \{x_1\}$ and $B \cap \{x_1, x_2\} = \{x_2\}$, which would contradict $A \circledast B$. Thus f is $\circledast$-modular. Simple calculation shows that

$$\beta = Y|_f = X|_f = \frac{\binom{n-2}{k-1}}{\binom{n}{k}} = \frac{k(n-k)}{n(n-1)}.$$

A similar argument shows that $\gamma = \beta$. To determine α, consider the modular function $g(A) = |A|$. Simple calculation shows that

$$k = X|_g = Y|_g = \alpha n + \beta |P_1| + \gamma |P_2| = \left(\alpha + \frac{k(n-k)}{n(n-1)} \right) n,$$

from which it is easy to calculate α. $\square$

The final example requires us to define $\mathcal{I}$ in a non-trivial way.

Lemma 6. *Let f be a submodular function on $\mathcal{U} = \mathcal{S} \cup \mathcal{O}$, where $\mathcal{S} = \{S_1, \ldots, S_n\}$ and $\mathcal{O} = \{O_1, \ldots, O_n\}$ are disjoint sets of size n. For $1 \leq k \leq n-1$,*

$$\mathop{\mathbb{E}}_{I \in \binom{[n]}{k}} f(\{O_i : i \in I\} \cup \{S_i : i \notin I\}) \geq$$
$$\frac{k(n-k)}{n(n-1)} f(\mathcal{S} \cup \mathcal{O}) + \frac{k(k-1)}{n(n-1)} f(\mathcal{O}) + \frac{(n-k)(n-k-1)}{n(n-1)} f(\mathcal{S}) + \frac{k(n-k)}{n(n-1)} f(\emptyset).$$

Proof. We apply the method with the following data:

- $\mathcal{G} = \mathrm{Sym}([n])$ which acts on $\mathcal{U}$ by permuting the indices.
- $A \circledast B$ if (1) A, B is an antichain, (2) for all $i \neq j$ it is *not* the case that $A \cap \{S_i, O_j\} = \{S_i\}$ and $B \cap \{S_i, O_j\} = \{O_j\}$ or vice versa, and (3) $A \cup B, A \cap B \in \mathcal{I}$ (defined next). Note that (2) can be falsified only if $A \triangle B$ contains both S_i and O_j for some $i \neq j$.

- $A \in \mathcal{I}$ if either (1) for every $i \in [n]$, $|A \cap \{S_i, O_i\}| \le 1$, or (2) for every $i \in [n]$, $|A \cap \{S_i, O_i\}| \ge 1$.

Proposition 1 gives us a vector Y supported on $(\mathcal{G}, \circledast)$-stable sets, each in addition belonging to $\mathcal{I}$. The sets $\mathcal{U}, \mathcal{O}, \mathcal{S}, \emptyset$ clearly satisfy these requirements, and we claim that these are the only ones. Indeed, consider an arbitrary $A \in \mathcal{I}$ which is $(\mathcal{G}, \circledast)$-stable. We consider three cases, corresponding to the cases in the definition of $\mathcal{I}$ and their combination.

Case 1: For all $i \in [n]$, $|A \cap \{S_i, O_i\}| \le 1$, and for some $j \in [n]$, $S_j, O_j \notin A$. If $A \neq \emptyset$ then (without loss of generality) $S_i \in A$ for some $i \in [n]$. Let $B = A^{(i,j)}$, so that $S_j \in B$ and $S_i, O_i \notin B$. Clearly both $A \cup B$ and $A \cap B$ satisfy case 1 of the definition of $\mathcal{I}$. Furthermore, $A \triangle B = \{S_i, S_j\}$, hence item 2 of the definition of $\circledast$ is satisfied. We conclude that $A \circledast B$ and so A is not $(\mathcal{G}, \circledast)$-stable.

Case 2: For all $i \in [n]$, $|A \cap \{S_i, O_i\}| \ge 1$, and for some $j \in [n]$, $S_j, O_j \in A$. If $A \neq \mathcal{U}$ then (without loss of generality) $S_i \notin A$ for some $i \in [n]$. Let $B = A^{(i,j)}$, so that $S_j \notin B$ and $S_i, O_i \in B$. Clearly both $A \cup B$ and $A \cap B$ satisfy case 2 of the definition of $\mathcal{I}$. Furthermore, $A \triangle B = \{S_i, S_j\}$, hence item 2 of the definition of $\circledast$ is satisfied. We conclude that $A \circledast B$ and so A is not $(\mathcal{G}, \circledast)$-stable.

Case 3: For all $i \in [n]$, $|A \cap \{S_i, O_i\}| = 1$. If $A \notin \mathcal{S}, \mathcal{O}$ then $S_i, O_j \in A$ for some $i \neq j$. Let $B = A^{(i,j)}$, so that $S_j, O_i \in B$. Clearly $A \cup B$ satisfies case 2 of the definition of $\mathcal{I}$, and $A \cap B$ satisfies case 1 of the definition. Furthermore, $A \triangle B = \{S_i, S_j, O_i, O_j\}$, hence item 2 of the definition of $\circledast$ can be falsified only for i, j, and that is not the case by inspection. We conclude that $A \circledast B$ and so A is not $(\mathcal{G}, \circledast)$-stable.

It follows that for some $\alpha, \beta, \gamma \in \mathbb{R}$,

$$Y|_f = \alpha f(\mathcal{U}) + \beta f(\mathcal{O}) + \gamma f(\mathcal{S}) + (1 - \alpha - \beta) f(\emptyset).$$

Define

$$f(A) = \begin{cases} 1 & \text{if } O_1 \in A \text{ and } S_2 \notin A, \\ 0 & \text{otherwise.} \end{cases}$$

As in the proof of Lemma 5, item 2 of the definition of $\circledast$ guarantees that f is $\circledast$-modular. Simple calculation shows that

$$\beta = Y|_f = X|_f = \frac{\binom{n-2}{k-2}}{\binom{n}{k}} = \frac{k(k-1)}{n(n-1)}.$$

Similarly we can calculate the value of γ. To determine α, consider the modular function $g(A) = |A|$. Simple calculation shows that

$$n = X|_g = Y|_g = 2\alpha n + \beta n + \gamma n = (2\alpha + \beta + \gamma)\, n.$$

Therefore $2\alpha + \beta + \gamma = 1$, and we can compute α. $\square$

2.4. Relation to the multilinear relaxation

Results very similar to Lemma 3, Lemma 4 and Lemma 5 appear in Vondrák [5], in the context of analyzing the multilinear relaxation. Given a set function f on a ground set $\mathcal{U}$, the multilinear relaxation $F\colon [0,1]^{\mathcal{U}} \to \mathbb{R}$ is defined by $F(X) = \mathbb{E}_A f(A)$, where each element $i \in \mathcal{U}$ is in A with probability X_i independently. The related Lovász extension $\tilde{F}(X)\colon [0,1]^{\mathcal{U}} \to \mathbb{R}$ is defined by

$$\tilde{F}(X) = \mathop{\mathbb{E}}_{\lambda \in [0,1]} f(\{i \in \mathcal{U} : X_i > \lambda\}).$$

Vondrák proves two results. His Lemma A.4 shows that for every submodular function f and vector $X \in [0,1]^{\mathcal{U}}$,

$$F(X) \geq \tilde{F}(X).$$

His Lemma A.5 shows that for every partition $\mathcal{U} = P_1 \cup P_2$, submodular function f and vector $X \in [0,1]^{\mathcal{U}}$,

$$F(X) \geq \mathop{\mathbb{E}}_{\lambda_1, \lambda_2 \in [0,1]} f(\{i \in P_1 : X_i > \lambda_1\} \cup \{i \in P_2 : X_i > \lambda_2\}).$$

These two results can be used to prove analogs of our lemmas where hard cardinality constraints of the form $|A| = k$ are replaced by choosing each element to be in A with probability k/n independently, where $n = |\mathcal{U}|$. We denote the distribution of such a set by $\mathrm{Bin}(\mathcal{U}, k/n)$.

For the analog of Lemma 3, let f be a submodular function on $\mathcal{U}$, where $|\mathcal{U}| = n$, and let k satisfy $0 \leq k \leq n$. Take X to be the constant k/n vector. Vondrák's Lemma A.4 shows that

$$\mathop{\mathbb{E}}_{A \sim \mathrm{Bin}(\mathcal{U}, k/n)} f(A) \geq \frac{k}{n} f(\mathcal{U}) + \left(1 - \frac{k}{n}\right) f(\emptyset).$$

For the analog of Lemma 4, let f be a submodular function on $\mathcal{U}$, where $|\mathcal{U}| = n$, let P_1, P_2 be a partition of $\mathcal{U}$ into sets of size n_1, n_2, respectively, and let k_1, k_2 satisfy $0 \leq k_1 \leq n_1$, $0 \leq k_2 \leq n_2$. Take X to be k_1/n_1 over P_1 and k_2/n_2 over P_2. Vondrák's Lemma A.5 shows that

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\substack{A_1 \sim \mathrm{Bin}(P_1, k_1/n_1) \\ A_2 \sim \mathrm{Bin}(P_2, k_2/n_2)}} f(A_1 \cup A_2) \\
&\qquad \geq \frac{k_1 k_2}{n_1 n_2} f(\mathcal{U}) + \frac{k_1(n_2 - k_2)}{n_1 n_2} f(P_1) + \frac{(n_1 - k_1)k_2}{n_1 n_2} f(P_2) + \frac{(n_1 - k_1)(n_2 - k_2)}{n_1 n_2} f(\emptyset).
\end{aligned}$$

For the analog of Lemma 5, let f be a submodular function on $\mathcal{U}$, where $|\mathcal{U}| = n$, let P_1, P_2 be a partition of $\mathcal{U}$, and let k satisfy $1 \leq k \leq n - 1$. Take X to be the constant k/n vector. Vondrák's Lemma A.5 shows that

$$\mathop{\mathbb{E}}_{A \sim \mathrm{Bin}(\mathcal{U}, k/n)} f(A) \geq \frac{k^2}{n^2} f(\mathcal{U}) + \frac{k(n-k)}{n^2} (f(P_1) + f(P_2)) + \frac{(n-k)^2}{n^2} f(\emptyset).$$

Note the slightly different coefficients in this case.

3. Applications

We consider several different algorithms for maximizing a given submodular function f defined on some ground set $\mathcal{U}$. We assume throughout that f attains only non-negative values.

The functions we consider are either unconstrained, symmetric or monotone:

- A *symmetric* function satisfies $f(A) = f(\mathcal{U} \setminus A)$.
- A *monotone* function satisfies $f(A) \leq f(B)$ whenever $A \subseteq B$.

The algorithms we consider attempt to maximize f over some subset $\mathcal{D} \subseteq 2^{\mathcal{U}}$. We consider two different sets of constraints:

1. Uniform constraint: $\mathcal{D} = \binom{\mathcal{U}}{k}$ for some k.
2. Partition constraint: The ground set $\mathcal{U}$ is partitioned into k parts $P_1, \ldots, P_k$, and $\mathcal{D}$ consists of those sets containing exactly one element from each part.

We will denote by $\mathcal{O}$ a set maximizing f under the constraint $\mathcal{O} \in \mathcal{D}$.

Comments. (1) Both types of constraints we consider are matroid constraints: the domain $\mathcal{D}$ consists of the bases of a uniform matroid or a partition matroid, respectively. Our results under the partition constraint actually hold even when $\mathcal{D}$ is an SBO matroid[2] of rank k. When $k = 1$, a theorem of Brualdi [7] implies that the results hold for any matroid.

(2) The algorithms we consider are natural but naive, and are superseded by superior algorithms. However, we believe that our analysis has merit. The first algorithm we consider, the random set algorithm, is a variant of an algorithm considered by Feige et al. [8], which is optimal for maximizing symmetric submodular functions over $\mathcal{D} = 2^{\mathcal{U}}$. The second algorithm we consider, the local search algorithm, is not only natural but also leads to an optimal algorithm for maximizing monotone submodular functions over an arbitrary matroid constraint [9]. Local search algorithms are also used for non-monotone submodular maximization. Whereas in some situations the approximation ratio improves by a constant when larger local changes are considered (see for example Lee et al. [10]), in our case the improvement is only sub-constant. A short survey of the best known algorithms appears in Appendix A.

3.1. Random set algorithm

The random set algorithm for maximizing a submodular function over the uniform constraint $\mathcal{D} = \binom{\mathcal{U}}{k}$ consists of choosing a uniformly random set from $\mathcal{D}$. We say that the algorithm has *approximation ratio* θ if $\mathbb{E}\, f(\mathcal{S}) \geq \theta f(\mathcal{O})$, where $\mathcal{S}$ is a uniformly random set chosen from $\mathcal{D}$.

---

[2] A matroid is *strongly base orderable* (SBO) if for any two bases B_1, B_2 there is a bijection $\rho\colon B_1 \to B_2$ such that $(B_1 \setminus X) \cup \rho(X)$ is a base for all $X \subseteq B_1$; see Brualdi [6].

Theorem 7. *Let f be a non-negative submodular function on $\mathcal{U}$, $n = |\mathcal{U}|$ and $1 \le k \le n-1$.*

(a) The random set algorithm has the approximation ratio

$$\frac{k(n-k)}{n(n-1)}.$$

(b) If f is symmetric, then the algorithm has the approximation ratio

$$\frac{2k(n-k)}{n(n-1)}.$$

(c) If f is monotone, then the algorithm has the approximation ratio

$$\frac{k}{n}.$$

Moreover, the bounds are tight for the random set algorithm in the worst case.

Proof. We first prove the stated approximation ratios, and then show that they are tight. For an unconstrained f, the approximation ratio follows from substituting $P_1 = \mathcal{O}, P_2 = \mathcal{U} \setminus \mathcal{O}$ in Lemma 5. For symmetric f, the same application yields a stronger inequality using $f(P_1) = f(P_2)$. Finally, for monotone f, the approximation ratio follows from Lemma 3 using $f(\mathcal{U}) \ge f(\mathcal{O})$.

Next, we exhibit in each of the three cases functions for which $\mathbb{E}\, f(\mathcal{S}) = \theta f(\mathcal{O})$, where θ is the relevant approximation ratio. Let $x, y \in \mathcal{U}$ be two different elements. For an unconstrained f, such a function is given by

$$f_1(A) = \begin{cases} 1 & \text{if } x \in A \text{ and } y \notin A, \\ 0 & \text{otherwise.} \end{cases}$$

For symmetric f, such a function is given by

$$f_2(A) = \begin{cases} 1 & \text{if } |A \cap \{x, y\}| = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Finally, for monotone f we can take the function $f_3(A) = |A|$. $\square$

Feige et al. [8] prove a similar result when the random set is chosen by putting each element in with probability $1/2$ independently. They show that the approximation ratio of this algorithm for the constraint $\mathcal{D} = 2^{\mathcal{U}}$ is $1/2$ when the function is symmetric and $1/4$ otherwise. Theorem 7 suggests a slightly superior algorithm for this constraint: choose a random set of size $\lfloor n/2 \rfloor$.

3.2. Local search

Local search [11] attempts to find the maximum by starting at an arbitrary set in $\mathcal{D}$, and making small changes that increase the value of f. The algorithm has a parameter $k \in \mathbb{N}$, and returns a set $\mathcal{S} \in \mathcal{D}$ satisfying

$$f(\mathcal{S}) \geq f(\mathcal{S} \setminus A \cup B)$$

for all $A \subseteq \mathcal{S}, B \subseteq \mathcal{U} \setminus \mathcal{S}$ such that $|A| = |B| \leq k$ and $\mathcal{S} \setminus A \cup B \in \mathcal{D}$. Such a set is known as a *local optimum*. We say that the algorithm has *approximation ratio* θ if $f(\mathcal{S}) \geq \theta f(\mathcal{O})$ whenever $\mathcal{S}$ is a local optimum. This algorithm is also known as k-local search.

We consider separately uniform constraints and partition constraints. In both cases, we analyze both unconstrained f and monotone f, but not symmetric f.

Theorem 8. *Let f be a non-negative submodular function on $\mathcal{U}$ and $1 \leq k \leq n \leq |\mathcal{U}|$. Set $\mathcal{D} = \binom{\mathcal{U}}{n}$.*

(a) Local search with parameter k has the approximation ratio

$$\frac{k}{2n-k}.$$

(b) If f is monotone, then local search has the approximation ratio

$$\frac{n}{2n-k}.$$

Proof. Write $\mathcal{O} = \{O_1, \ldots, O_n\}$ and $\mathcal{S} = \{S_1, \ldots, S_n\}$. Suppose first that $\mathcal{O}$ and $\mathcal{S}$ are disjoint. The fact that $\mathcal{S}$ is a local optimum implies that

$$f(\mathcal{S}) \geq \mathop{\mathbb{E}}_{I,J \in \binom{[n]}{k}} f(\{S_i : i \notin I\} \cup \{O_j : j \in J\}).$$

The stated approximation ratios now follow directly from Lemma 4, in the case of monotone f using $f(\mathcal{S} \cup \mathcal{O}) \geq f(\mathcal{O})$; we use $P_1 = \mathcal{S}$, $P_2 = \mathcal{O}$, $n_1 = n_2 = n$, $k_1 = n - k$ and $k_2 = k$.

If $\mathcal{O}$ and $\mathcal{S}$ are not disjoint, let $\mathcal{C} = \mathcal{O} \cap \mathcal{S}$ and $c = |\mathcal{C}|$. If $n - c \leq k$ then local optimality implies $f(\mathcal{S}) \geq f(\mathcal{O})$, so assume $n - c > k$. Define a function f' on $\mathcal{U}' = \mathcal{U} \setminus \mathcal{C}$ by $f'(A) = f(\mathcal{C} \cup A)$. Set $\mathcal{D}' = \binom{\mathcal{U}'}{n'}$, where $n' = n - c$. The following properties are easy to verify: f' is submodular, if f is monotone then so is f', $\mathcal{O}' = \mathcal{O} \setminus \mathcal{C}$ is a set maximizing f' over $\mathcal{D}'$, and $\mathcal{S}' = \mathcal{S} \setminus \mathcal{C}$ is a local optimum for local search with parameter k. For unconstrained f, since $\mathcal{O}'$ and $\mathcal{S}'$ are disjoint,

$$\frac{f(\mathcal{S})}{f(\mathcal{O})} = \frac{f'(\mathcal{S}')}{f'(\mathcal{O}')} \geq \frac{k}{2n'-k} > \frac{k}{2n-k}.$$

So the stated approximation ratio holds even when $\mathcal{O}$ and $\mathcal{S}$ are not disjoint. A similar argument works when f is monotone:

$$\frac{f(\mathcal{S})}{f(\mathcal{O})} = \frac{f'(\mathcal{S}')}{f'(\mathcal{O}')} \geq \frac{n'}{2n'-k} = \frac{n-c}{2n-k-2c} > \frac{n}{2n-k}. \qquad \square$$

Theorem 9. *Let f be a non-negative submodular function on $\mathcal{U}$ and $1 \leq k \leq n \leq |\mathcal{U}|$. Suppose $\mathcal{U}$ is partitioned into n non-empty parts $P_1, \ldots, P_n$, and let $\mathcal{D}$ consist of all sets intersecting each P_i at exactly one point.*

(a) Local search with parameter k has the approximation ratio

$$\frac{k-1}{2n-k-1}.$$

(b) If f is monotone, then local search has the approximation ratio

$$\frac{n-1}{2n-k-1}.$$

Proof. Write $\mathcal{O} = \{O_1, \ldots, O_n\}$ and $\mathcal{S} = \{S_1, \ldots, S_n\}$, where $O_i, S_i \in P_i$. Suppose first that $\mathcal{O}$ and $\mathcal{S}$ are disjoint. The fact that $\mathcal{S}$ is a local optimum implies that

$$f(\mathcal{S}) \geq \mathop{\mathbb{E}}_{I \in \binom{[n]}{k}} f(\{S_i : i \notin I\} \cup \{O_i : i \in I\}).$$

The stated approximation ratios now follow directly from Lemma 6, in the case of monotone f using $f(\mathcal{S} \cup \mathcal{O}) \geq f(\mathcal{O})$. When $\mathcal{O}$ and $\mathcal{S}$ are not disjoint, use the argument in the proof of the preceding theorem. ☐

The results obtained in Theorem 8 and Theorem 9 are tight. In order to show that, we present a repertory of well-known submodular functions.

Set coverage functions. A set coverage function $f \colon 2^{\mathcal{U}} \to \mathbb{R}$ is given by a universe $\mathcal{V}$, a weight function $w \colon \mathcal{V} \to \mathbb{R}_{\geq 0}$, and for each $x \in \mathcal{U}$, a set $S(x) \subseteq \mathcal{V}$. The function itself is defined by

$$f(A) = \sum_{a \in \bigcup_{x \in A} S(x)} w(a).$$

In other words, $f(A)$ is the total weight of all elements covered by the sets in A. A set coverage function is always monotone and submodular.

Directed cut functions. A directed cut function is given by a directed graph on $\mathcal{U}$ with non-negative weights on the edges $w(a, b)$; edges not in the graph have zero weight. The function is given by

$$f(A) = \sum_{\substack{a \in A, \\ b \notin A}} w(a, b).$$

In other words, $f(A)$ is the total weight of all edges crossing the cut $(A, \overline{A})$. A directed cut function is always submodular, but not necessarily monotone.

Set coverage functions will be used in the monotone case, while directed cut functions will be used in the unconstrained case.

Theorem 10. *The approximation ratios given in Theorem 8 and Theorem 9 are tight.*

Proof. Let k, n be given. In each case, we will describe a function defined on $\mathcal{U} = \mathcal{S} \cup \mathcal{O} = \{S_i : i \in [n]\} \cup \{O_i : i \in [n]\}$. For Theorem 9 we will consider the partition $P_i = \{S_i, O_i\}$. In order to show that $\mathcal{S}$ is a local optimum, we will consider the hybrids

$$\mathcal{C}_t = \{S_1, \ldots, S_{n-t}, O_{n-t+1}, \ldots, O_n\}$$

for $t \leq k$.

Theorem 8, monotone case. Consider the following set coverage function. The universe consists of elements $x(i,j)$, $i,j \in [n]$ (possibly equal) of weight 1, and elements $y(i)$, $i \in [n]$ of weight $n-k$. The sets are given by

$$S_i = \{x(i,j) : j \in [n]\}, \quad O_i = \{x(j,i) : j \in [n]\} \cup \{y(i)\}.$$

We have $f(\mathcal{S}) = n^2$ and $f(\mathcal{O}) = n^2 + n(n-k)$, so $f(\mathcal{S})/f(\mathcal{O}) = n/(2n-k)$, as claimed.

In order to see that $\mathcal{S}$ in a local optimum, consider the hybrid $\mathcal{C}_t$. An element $x(i,j)$ is covered unless $i \in \{n-t+1, \ldots, n\}$ and $j \in \{1, \ldots, n-t\}$, for a total of $n^2 - t(n-t)$ covered elements of this form. Also, t elements of the form $y(i)$ are covered. Hence

$$f(\mathcal{C}_t) = n^2 - t(n-t) + t(n-k) = n^2 - (k-t)t \leq f(\mathcal{S}).$$

Theorem 9, monotone case. We slightly modify the previous example by setting the weight of the elements $x(i,i)$ to zero. We then have $f(\mathcal{S}) = n(n-1)$ and $f(\mathcal{O}) = n(n-1) + n(n-k)$, so $f(\mathcal{S})/f(\mathcal{O}) = (n-1)/(2n-k-1)$, as claimed. The hybrid $\mathcal{C}_t$ now satisfies

$$f(\mathcal{C}_t) = n(n-1) - t(n-t) + t(n-k) = n(n-1) - (k-t)t \leq f(\mathcal{S}).$$

Theorem 8, unconstrained case. Consider the following directed cut function. For $i,j \in [n]$ (possibly equal), there is an edge (S_i, O_j) of weight k, and an edge (O_j, S_i) of weight $2n-k$. Clearly $f(\mathcal{S})/f(\mathcal{O}) = k/(2n-k)$, as claimed. The hybrid $\mathcal{C}_t$ satisfies

$$f(\mathcal{C}) = (n-t)^2 k + t^2(2n-k) = n^2 k - 2n(k-t)t \leq n^2 k = f(\mathcal{S}).$$

Theorem 9, unconstrained case. We slightly modify the previous example, by the setting the weight of the edges (S_i, O_i) and (O_i, S_i) to zero, by reducing the weight of the edges (S_i, O_j) to $k-1$, and by reducing the weight of the edges (O_i, S_j) to $2n-k-1$ (where $i \neq j$). Clearly $f(\mathcal{S})/f(\mathcal{O}) = (k-1)/(2n-k-1)$. The hybrid $\mathcal{C}_t$ now satisfies

$$\begin{aligned} f(\mathcal{C}) &= (n-t)(n-t-1)(k-1) + t(t-1)(2n-k-1) \\ &= n(n-1)(k-1) - 2(n-1)(k-t)t \leq n(n-1)(k-1) = f(\mathcal{S}). \end{aligned}$$

□

Theorem 8(b) is proved in Fisher et al. [11]. Theorem 9(b) is proved in Filmus and Ward [12] for the case of coverage functions. The reader is invited to compare the proof there to the proof of the underlying Lemma 6 here. Extensions to optimization over the intersection of several matroids are considered by Fisher et al. [13] and by Lee et al. [14].

[1] J. Edmonds, Submodular functions, matroids, and certain polyhedra, in: Combinatorial structures and their applications, Gordon and Breach, New York, 1970, pp. 69–87.

[2] L. Lovász, Submodular functions and convexity, in: Mathematical Programming — The State of The Art, Springer, Berlin, 1983, pp. 235–257.

[3] S. Fujishige, Submodular Functions and Optimization, volume 58 of *Annals of Discrete Mathematics*, Elsevier, second edition, 2005.

[4] A. Schrijver, Combinatorial Optimization: Polyhedra and Efficiency, Springer, 2003.

[5] J. Vondrák, Symmetry and approximability of submodular maximization problems, in: Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2009).

[6] R. A. Brualdi, Induced matroids, Proceedings of the American Mathematical Society 29 (1971) 213–221.

[7] R. A. Brualdi, Comments on bases in dependence structure, Bulletin of the Australian Mathematical Society 1 (1969) 161–167.

[8] U. Feige, V. S. Mirrokni, J. Vondrak, Maximizing non-monotone submodular functions, SIAM Journal on Computing 40 (2011) 1133–1153.

[9] Y. Filmus, J. Ward, A tight combinatorial algorithm for submodular maximization subject to a matroid constraint, in: Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS 2012).

[10] J. Lee, V. Mirrokni, V. Nagarajan, M. Sviridenko, Non-monotone submodular maximization under matroid and knapsack constraints, SIAM Journal on Discrete Mathematics 23 (2010) 2053–2078.

[11] M. L. Fisher, G. L. Nemhauser, L. A. Wolsey, An analysis of approximations for maximizing submodular set functions I, Mathematical Programming 14 (1978) 265–294.

[12] Y. Filmus, J. Ward, The power of local search: Maximum coverage over a matroid, in: Proceedings of the 29th annual Symposium on Theoretical Aspects of Computer Science (STACS 2012), pp. 601–612.

[13] M. L. Fisher, G. L. Nemhauser, L. A. Wolsey, An analysis of approximations for maximizing submodular set functions—II, in: Polyhedral Combinatorics, Springer Berlin Heidelberg, 1978, pp. 73–87.

[14] J. Lee, M. Sviridenko, J. Vondrák, Matroid matching: the power of local search, in: Proceedings of the 42nd ACM Symposium on Theory of Computing (STOC 2010), pp. 369–378.

[15] G. Cornuejols, M. L. Fisher, G. L. Nemhauser, Location of bank accounts to optimize float: An analytic study of exact and approximate algorithms, Management Science 23 (1977) pp. 789–810.

[16] D. S. Hochbaum, A. Pathria, Analysis of the greedy approach in covering problems, Naval Research Quarterly 45 (1998) 615–627.

[17] P. R. Goundan, A. S. Schulz, Revisiting the greedy approach to submodular set function maximization, 2009. Preprint.

[18] U. Feige, A threshold of ln n for approximating set cover, Journal of the ACM 45 (1998) 634–652.

[19] G. Calinescu, C. Chekuri, M. Pál, J. Vondrák, Maximizing a monotone submodular function subject to a matroid constraint, SIAM Journal on Computing 40 (2011) 1740–1766.

[20] N. Buchbinder, M. Feldman, J. S. Naor, R. Schwartz, A tight linear time (1/2)-approximation for unconstrained submodular maximization, in: Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS 2012).

Appendix A. Algorithms for submodular maximization

In §3 we analyze two algorithms: the random set algorithm and local search. In order to complete the picture, we describe the best known polynomial-time algorithms for the problems we consider.

Monotone functions. Monotone submodular maximization over a uniform constraint is a generalization of the classical NP-complete problem *maximum coverage*. The greedy algorithm [15, 11, 16, 17] produces the approximation ratio $1-1/e$ for monotone functions, which was shown to be optimal by Feige [18] unless P=NP.

Monotone submodular maximization over a partition constraint is a generalization of another classical NP-complete problem, *maximum satisfiability* (MAX-SAT). The function f in question is a set coverage function (see §3.2) in which the parts correspond to variables, and the two elements in each part represent the set of clauses satisfied by a given assignment to the variable. Both the continuous greedy algorithm [19] and non-oblivious local search [9] produce the optimal approximation ratio $1-1/e$ for monotone functions over matroid constraints, which generalize partition constraints.

In contrast, the best algorithm we consider in §3 (local search) produces an approximation ratio which is roughly $1/2$ in any variant which runs in polynomial time.

Non-monotone functions. Non-monotone submodular maximization is often considered for $\mathcal{D} = 2^{\mathcal{U}}$, and this is a generalization of the classical NP-complete problem *maximum cut* (MAX-CUT). Feige et al. [8] show that the random set algorithm produces a $1/2$ approximation for symmetric functions, which is optimal in the value oracle model: every algorithm that queries the function as a black box and has an approximation ratio better than $1/2$ must query the function in exponentially many places. The recent double greedy algorithm of Feldman et al. [20] produces a $1/2$ approximation for an arbitrary submodular function.

Submodular maximization over bases of a matroid, a constraint generalizing the partition constraint, is considered by Vondrák [5] and Lee et al. [10]. Lee et al. give a $1/3 - \epsilon$ approximation algorithm for symmetric functions, and Vondrák gives a $1/4-\epsilon$ approximation algorithm for arbitrary submodular functions, which works under some mild constraints on the matroid (it has to contain at least two disjoint bases). For the case of a partition constraint, this condition holds if each part contains at least two different elements. Vondrák also shows that in the value oracle model, the best achievable approximation ratio is at most $1/2$.

In contrast, the local search algorithm has an $o(1)$ approximation ratio in any variant which runs in polynomial time, even for a uniform constraint.

For completeness, we also mention that submodular maximization is often considered over the independent sets of a matroid. For a uniform constraint, that means optimization over all sets of cardinality at most k instead of exactly k. For a partition constraint, that means optimization over all sets containing at most one element from each part.