Optim Lett (2017) 11:655–665
DOI 10.1007/s11590-016-1039-z

ORIGINAL PAPER

New performance guarantees for the greedy maximization of submodular set functions

Jussi Laitila[1] · **Atte Moilanen**[2]

Received: 3 June 2015 / Accepted: 2 May 2016 / Published online: 10 May 2016

Abstract We present new tight performance guarantees for the greedy maximization of monotone submodular set functions. Our main result first provides a performance guarantee in terms of the overlap of the optimal and greedy solutions. As a consequence we improve performance guarantees of Nemhauser et al. (Math Program 14:265–294, 1978) and Conforti and Cornuéjols (Discr Appl Math 7:251–274, 1984) for maximization over subsets, which are at least half the size of the problem domain. As a further application, we obtain a new tight approximation guarantee in terms of the cardinality of the problem domain.

Keywords Approximation · Cardinality · Convex optimization · Greedy algorithm · Maximization · Steepest ascent

1 Introduction

Let X be a finite set, $X = \{x_1, \dots, x_n\}$, and let T be an integer such that $0 < T \leq n$. We consider the cardinality-constrained maximization problem

$$\max\{f(S) \colon |S| = T, S \subset X\}, \tag{1}$$

✉ Jussi Laitila
jussi.laitila@helsinki.fi

1 Department of Mathematics and Statistics, University of Helsinki, P.O. Box 68, 00014 Helsinki, Finland

2 Department of Biosciences, University of Helsinki, P.O. Box 65, 00014 Helsinki, Finland

where $f: 2^X \to \mathbb{R}_+$ is a submodular set function. Recall that f is submodular if

$$f(S) + f(R) \geq f(S \cup R) + f(S \cap R) \tag{2}$$

for all $S, R \subset X$; see, e.g., [1]. We further assume that f is nondecreasing; $f(S) \leq f(R)$ for all $S \subset R$, and, without loss of generality, that $f(\emptyset) = 0$. We consider the following well-known greedy algorithm for solving problem (1):

Algorithm A

Step 0: Set $S_0 = \emptyset$. Go to Step 1.
Step t $(1 \leq t \leq T)$: Select any $x_t \in S_{t-1}$ such that

$$f(S_{t-1} \cup \{x_t\}) = \max\{f(S_{t-1} \cup \{x\}) : x \in X \backslash S_{t-i}\}.$$

Set $S_t = S_{t-1} \cup \{x_t\}$. Go to step $t+1$.
Step $T+1$: Set $S_{gr} = S_T$. Stop.

Algorithm A has been extensively studied in the literature. It is well known [2,3], that it finds an optimal solution when f is an additive set function, i.e., when (2) holds with an equality for all $S, R \subset X$. Nemhauser et al. [1] (see also [4,5]) gave the following performance guarantee for Algorithm A for nonadditive functions f:

$$\frac{f(S_{gr})}{f(S_{opt})} \geq 1 - \left(1 - \frac{1}{T}\right)^T =: G_{NWF}(T), \tag{3}$$

where S_{opt} is an optimal solution to problem (1). Conforti and Cornuéjols [2] improved (3) to

$$\frac{f(S_{gr})}{f(S_{opt})} \geq \frac{1}{\alpha}\left(1 - \left(1 - \frac{\alpha}{T}\right)^T\right) =: G_{CC}(T, \alpha), \tag{4}$$

for $\alpha \in (0, 1]$, where $\alpha \in (0, 1]$ is the total curvature

$$\alpha = \max\left\{1 - \frac{f(X) - f(X\backslash\{x\})}{f(\{x\}) - f(\emptyset)} : x \in X, f(\{x\}) \neq f(\emptyset)\right\}.$$

It is known that $\alpha \in (0, 1]$ if and only if f is nonadditive [2]. We also clearly have $G_{NWT}(T) = G_{CC}(T, 1)$ and $G_{CC}(T, \alpha) \to 1$ as $\alpha \to 0^+$. The above performance guarantees further satisfy the estimates

$$G_{CC}(T, \alpha) \geq \max\left\{G_{NWF}(T), \frac{1 - e^{-\alpha}}{\alpha}\right\} \geq 1 - e^{-1},$$

for all α and T. The guarantees (3) and (4) are tight for suitable choices of parameters T and α. For example, for all $\alpha \in (0, 1]$ and $T \geq 1$ there is a problem of the type (1) and the corresponding greedy solution S_{gr} such that $f(S_{gr}) = G_{CC}(T, \alpha) f(S_{opt})$ [2].

Submodular optimization has played a central role in operations research and combinatorial optimization [6]. By now it has applications in various fields, including computer science [7], economics [8] and, more recently, ecology [9–11]. Problem (1) and the above performance guarantees have been extended to various other settings and problem structures, related to, for example, matroid [2,12] and knapsack [13,14] constraints, continuous algorithms [15,16], nonmonotone functions [17], nonsubmodular functions [18] and supermodular minimization [19,20].

To authors' knowledge, previously presented performance guarantees either do not depend on T or n, or, like (3) and (4), they are decreasing in T. However, when $T = n$, it is clear that $S_{opt} = S_{gr}$, so the greedy algorithm returns the optimal solution. This suggests that any performance guarantee should in fact be improving when T approaches and is close enough to n. We show that this is indeed the case. More generally, we show that increasing degree of overlap $m = |S_{opt} \cap S_{gr}|$ between the sets S_{opt} and S_{gr} improves the approximation guarantees. While in applications the overlap m may not be known, we can give this quantity a useful lower bound. In fact, when $T > n/2$, we have $m \geq 2T - n > 0$. Our results thus have particular relevance for optimization problems where the maximum is sought over subsets of cardinality larger than $n/2$.

Let

$$G(T, \alpha, m) = \frac{1}{\alpha}\left(1 - \left(1 - \frac{\alpha m}{T}\right)\left(1 - \frac{\alpha}{T}\right)^{T-m}\right)$$

and $\widetilde{G}(T, \alpha, n) = G(T, \alpha, \max\{0, 2T - n\})$. Our main result is the following.

Theorem 1 *Let* $\alpha \in (0, 1]$, *let* $1 \leq T \leq n$ *and let* S_{opt} *and* S_{gr} *be an optimal, respectively a greedy, solution to problem* (1) *and let* $m = |S_{opt} \cap S_{gr}|$. *Then*

$$\frac{f(S_{gr})}{f(S_{opt})} \geq G(T, \alpha, m) \geq \widetilde{G}(T, \alpha, n). \tag{5}$$

Moreover, these bounds are tight in the following sense: for every $\alpha \in (0, 1]$ *and numbers* n *and* T *such that* $1 \leq T \leq n$, *there is a problem of the type* (1) *and its greedy solution* S_{gr} *such that* $\max\{0, 2T - n\} = |S_{opt} \cap S_{gr}|$ *and*

$$\frac{f(S_{gr})}{f(S_{opt})} = \widetilde{G}(T, \alpha, n).$$

We postpone the proof of Theorem 1 to Sect. 2.

Remark 1 Theorem 1 strictly improves (4) and provides further examples of cases where the performance guarantee equals one. Indeed, for all T and n such that $T > n/2$, we have the strict inequality

$$\widetilde{G}(T, \alpha, n) > G_{CC}(T, \alpha).$$

For $T = n$, we get $\widetilde{G}(n, \alpha, n) = 1$. Note that, by (4), $\lim_{\alpha \to 0^+} \widetilde{G}(T, \alpha, n) = 1$. Moreover, in the case $m = T$, we again get $G(T, \alpha, T) = 1$. Note also that $G(T, \alpha, m)$ is decreasing in α, so (5) can be substituted by a weaker but simpler approximation guarantee

$$\frac{f(S_{gr})}{f(S_{opt})} \geq 1 - \left(1 - \frac{m}{T}\right)\left(1 - \frac{1}{T}\right)^{T-m}.$$

Using Theorem 1, one can derive other new performance guarantees for the greedy algorithm. As an example of independent interest, we present the following performance guarantee in terms of n only.

Corollary 1 *Let $\alpha \in (0, 1]$, $1 \leq T \leq n$, and let S_{opt} and S_{gr} be an optimal, respectively a greedy, solution to problem (1). Then*

$$\frac{f(S_{gr})}{f(S_{opt})} \geq \frac{1}{\alpha}\left(1 - \left(1 - \frac{\alpha}{\lfloor \frac{n}{2} \rfloor}\right)^{\lfloor \frac{n}{2} \rfloor}\right) \geq \frac{1}{\alpha}\left(1 - \left(1 - \frac{2\alpha}{n}\right)^{n/2}\right), \tag{6}$$

where $\lfloor x \rfloor$ denotes the largest integer not greater than x. The left-hand estimate is tight in the following sense: for every $\alpha \in (0, 1]$ and $n \geq 2$, there is a problem of the type (1) and its greedy solution S_{gr} such that

$$\frac{f(S_{gr})}{f(S_{opt})} = \frac{1}{\alpha}\left(1 - \left(1 - \frac{\alpha}{\lfloor \frac{n}{2} \rfloor}\right)^{\lfloor \frac{n}{2} \rfloor}\right).$$

Proof If n is an odd integer, it is easy to check that the minimum of $\widetilde{G}(T, \alpha, n)$ over all integers T with $0 \leq T \leq n$ is $\widetilde{G}((n-1)/2, \alpha, n)$. Moreover, when treated as a continuous function of T, $\widetilde{G}(T, \alpha, n)$ attains its minimum at $T = n/2$. Together with Theorem 1 this yields (6). Tightness of the left-hand inequality in (6) follows from Theorem 1 with the choice $T = \lfloor \frac{n}{2} \rfloor$.

2 Proof of Theorem 1

In this section we present a proof of Theorem 1. We first prove (5). Note that the right-hand inequality in (5) follows directly from $m = |S_{opt} \cap S_{gr}| \geq \max\{0, 2T - n\}$ and the fact that $G(T, \alpha, m)$ is increasing in m.

We next prove the left-hand inequality in (5). We may assume that $0 < m < T$. Indeed, if $m = T$, then $S_{gr} = S_{opt}$ and the claim is trivial. If $m = 0$, the claim follows from (4).

Let $S_0 = \emptyset$ and $S_t = \{y_1, \ldots, y_t\} \subset X$ be the successive sets chosen by the greedy algorithm for $t = 1, \ldots, T$, so that $S_0 \subset S_1 \subset \cdots \subset S_T$. Let

$$a_t = \frac{f(S_t) - f(S_{t-1})}{f(S_{opt})},$$

for $t = 1, \ldots, T$. Because f is nondecreasing, each a_t is nonnegative and

$$\frac{f(S_{gr})}{f(S_{opt})} = \sum_{i=1}^{T} a_i.$$

Let $1 \le j_1 \le \cdots \le j_m \le T$ denote the indices for which $S_{gr} \cap S_{opt} = \{y_{j_1}, \ldots y_{j_m}\}$. Denote $j_0 = 0$ and $j_{m+1} = T$. We will use the following lemma from [2].

Lemma 1 ([2, Lemma 5.1]) *We have*

$$\alpha \sum_{\{i:\ y_i \in S_{t-1} \setminus S_{opt}\}} a_i + \sum_{\{i:\ y_i \in S_{t-1} \cap S_{opt}\}} a_i + (T - |S_{t-1} \cap S_{opt}|) a_t \ge 1,$$

for $t = 1, \ldots, T$.

Using Lemma 1, we get

$$\frac{f(S_{gr})}{f(S_{opt})} \ge B(J), \tag{7}$$

where $J = \{j_1, \ldots, j_m\}$ and, for $U \subset \{1, \ldots, T\}$, $B(U)$ denotes the minimum of the linear program

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{T} b_i \\ \text{s.t.} \quad & \alpha \sum_{i \in V_{t-1} \setminus U} b_i + \sum_{i \in U \cap V_{t-1}} b_i + (T - |U \cap V_{t-1}|) b_t \ge 1, \quad t = 1, \ldots, T \\ & b_t \ge 0, \quad t = 1, \ldots, T, \end{aligned} \tag{8}$$

where $V_t = \{1, \ldots, t\}$. The following lemma refines [2, Lemma 5.2].

Lemma 2 $B(J) \ge B(\{T - m + 1, T - m + 2, \ldots, T\})$.

Proof Fix $1 \le r \le m$ and consider $q = j_r \in J$. We first show that $b_q \le b_{q+1}$ for some optimal solution to (8) with $U = J$. To this end, assume that this does not hold for some optimal solution $b = (b_1, \ldots, b_T)$. Then $\varepsilon := b_q - b_{q+1} > 0$. The constraints q and $q + 1$ are

$$\alpha \sum_{i \in V_{q-1} \setminus J} b_i + \sum_{i \in J \cap V_{q-1}} b_i + (T - r + 1) b_q \ge 1;$$

$$\alpha \sum_{i \in V_q \setminus J} b_i + \sum_{i \in J \cap V_q} b_i + (T - r) b_{q+1} \ge 1.$$

Because $V_q \backslash J = V_{q-1} \backslash J$ and $J \cap V_q = (J \cap V_{q-1}) \cup \{q\}$, the constraint $q+1$ is equivalent to

$$\alpha \sum_{i \in V_{q-1} \backslash J} b_i + \sum_{i \in J \cap V_{q-1}} b_i + b_q + (T-r) b_{q+1} \geq 1.$$

Therefore

$$\alpha \sum_{i \in V_{q-1} \backslash J} b_i + \sum_{i \in J \cap V_{q-1}} b_i + (T-r+1) b_q \geq 1 + \varepsilon (T-r) > 1,$$

which shows that the constraint q is not tight. Form a new solution $b' = (b'_1, \ldots, b'_T)$ by setting $b'_i = b_i$ for $1 \leq i \leq q-1$, $b'_q = b_q - \varepsilon(T-r)/(T-r+1)$ and $b'_i = b_i + \varepsilon/(T-r+1)$ for $q+1 \leq i \leq T$. It is easy to check that b' is feasible. Moreover,

$$b'_q - b'_{q+1} = b_q - \frac{\varepsilon(T-r)}{T-r+1} - b_{q+1} - \frac{\varepsilon}{T-r+1} = 0$$

and

$$\sum_{i=1}^{T} b'_i = \sum_{i=1}^{T} b_i + \frac{\varepsilon(T-q)}{T-r+1} - \frac{\varepsilon(T-r)}{T-r+1} \leq \sum_{i=1}^{T} b_i,$$

because $r \leq q$. Hence b' is an optimal solution with $b'_q \leq b'_{q+1}$.

Assume next that $q = j_r \in J$ and $q+1 \notin J$ for some r. Let $b = (b_1, \ldots, b_T)$ be a feasible solution to (8) with $U = J$, so that

$$\alpha \sum_{i \in V_{t-1} \backslash J} b_i + \sum_{i \in J \cap V_{t-1}} b_i + (T - |J \cap V_{t-1}|) b_t \geq 1, \tag{9}$$

for $1 \leq t \leq T$. Assume also that $b_q \leq b_{q+1}$. Let $J' = \{j_1, \ldots, j_{r-1}, q+1, j_{r+1}, \ldots, j_m\}$. We will show that b is a feasible solution to (8) with $U = J'$. Consider first $1 \leq t \leq q$. Then $V_{t-1} \backslash J' = V_{t-1} \backslash J$ and $J' \cap V_{t-1} = J \cap V_{t-1}$, so

$$\alpha \sum_{i \in V_{t-1} \backslash J'} b_i + \sum_{i \in J' \cap V_{t-1}} b_i + (T - |J' \cap V_{t-1}|) b_t \geq 1,$$

by (9). Consider next $t = q + 1$. Then $V_{t-1}\backslash J' = (V_{t-1}\backslash J) \cup \{q\}$ and $J' \cap V_{t-1} = (J \cap V_{t-1})\backslash\{q\}$. By (9) and using $b_q \leq b_{q+1}$, we get

$$\begin{aligned} &\alpha \sum_{i \in V_q \backslash J'} b_i + \sum_{i \in J' \cap V_q} b_i + (T - |J' \cap V_q|) b_{q+1} \\ &= \alpha \sum_{i \in V_q \backslash J} b_i + \alpha b_q + \sum_{i \in J \cap V_q} b_i - b_q + (T - |J \cap V_q| + 1) b_{q+1} \\ &\geq 1 + \alpha b_q - b_q + b_{q+1} \geq 1. \end{aligned}$$

Finally, consider $t = q + k$ for $k \geq 2$. Then $V_{t-1}\backslash J' = ((V_{t-1}\backslash J) \cup \{q\})\backslash\{q+1\}$ and $J' \cap V_{t-1} = ((J \cap V_{t-1})\backslash\{q\}) \cup \{q+1\}$. By (9) and using $b_q \leq b_{q+1}$, we get similarly as above

$$\begin{aligned} &\alpha \sum_{i \in V_{q+k-1} \backslash J'} b_i + \sum_{i \in J' \cap V_{q+k-1}} b_i + (T - |J' \cap V_{q+k-1}|) b_{q+k} \\ &\geq 1 + (b_{q+1} - b_q)(1 - \alpha) \geq 1. \end{aligned}$$

This shows that b is a feasible solution to (8) with $U = J'$.
By combining the above results, we get

$$B(J) \geq B(J').$$

The proof of Lemma 2 is completed by repeating this argument sufficiently many times. □

Lemma 2 and (7) now imply

$$\frac{f(S_{gr})}{f(S_{opt})} \geq B(\{T - m + 1, T - m + 2, \ldots, T\}).$$

By the weak duality theorem, we get

$$\frac{f(S_{gr})}{f(S_{opt})} \geq \sum_{i=1}^{T} c_i^*, \tag{10}$$

where $c^* = (c_1^*, \ldots, c_T^*)$ is an optimal solution to the dual problem of (8):

$$\text{maximize} \quad \sum_{i=1}^{T} c_i \tag{11}$$

$$\text{s.t.} \quad Tc_t + \alpha \sum_{i=t+1}^{T} c_i \leq 1, \qquad 1 \leq t \leq T - m \tag{12}$$

$$(2T - m + 1 - t)c_t + \sum_{i=t+1}^{T} c_i \leq 1, \qquad T - m + 1 \leq t \leq T \tag{13}$$

$$c_i \geq 0, \qquad i = 1, \ldots, T. \tag{14}$$

Define the vector $c = (c_1, \ldots, c_T)$ by

$$c_t = \begin{cases} \frac{1}{T}\left(1 - \frac{\alpha m}{T}\right)\left(1 - \frac{\alpha}{T}\right)^{T-m-t}, & 1 \leq t \leq T - m, \\ \frac{T-m}{(2T-m+1-t)(2T-m-t)}, & T - m + 1 \leq t \leq T. \end{cases}$$

We will need the following two straightforward indentities:

$$\sum_{i=s}^{T-m} c_i = \frac{1}{\alpha}\left(1 - \frac{\alpha m}{T}\right)\left(1 - \left(1 - \frac{\alpha}{T}\right)^{T-m-s+1}\right), \qquad 1 \leq s \leq T - m; \tag{15}$$

$$\sum_{i=k}^{T} c_i = \frac{T - k + 1}{2T - m - k + 1}, \qquad T - m + 1 \leq k \leq T + 1. \tag{16}$$

Lemma 3 *The vector c is a feasible solution to problem* (11).

Proof Consider first $1 \leq t \leq T - m - 1$. By (15) and (16),

$$\sum_{i=s}^{T} c_i = \sum_{i=s}^{T-m} c_i + \sum_{i=T-m+1}^{T} c_i = \frac{1}{\alpha}\left(1 - \frac{\alpha m}{T}\right)\left(1 - \left(1 - \frac{\alpha}{T}\right)^{T-m-s+1}\right) + \frac{m}{T},$$

for $1 \leq s \leq T - m$. Hence

$$\begin{aligned} Tc_t + \alpha \sum_{i=t+1}^{T} c_i &= \left(1 - \frac{\alpha m}{T}\right)\left(1 - \frac{\alpha}{T}\right)^{T-m-t} \\ &\quad + \left(1 - \frac{\alpha m}{T}\right)\left(1 - \left(1 - \frac{\alpha}{T}\right)^{T-m-t}\right) + \frac{\alpha m}{T} \\ &= 1, \end{aligned}$$

so c_t satisfies the constraint (12).

By (16),

$$Tc_{T-m} + \alpha \sum_{i=T-m+1}^{T} c_i = \left(1 - \frac{\alpha m}{T}\right) + \frac{\alpha m}{T} = 1,$$

so c_{T-m} also satisfies the constraint (12).

For $T - m + 1 \leq t \leq T$, we get from (16) that

$$(2T - m + 1 - t)c_t + \sum_{i=t+1}^{T} c_i = 1,$$

so c_t satisfies the constraint (13).

Finally, it is clear from the definition that each c_t satisfies the constraint (14). This completes the proof of Lemma 3. □

Lemma 3 and (10) imply

$$\frac{f(S_{gr})}{f(S_{opt})} \geq \sum_{i=1}^{T} c_i.$$

Moreover, by (15) and (16),

$$\sum_{i=1}^{T} c_i = \frac{1}{\alpha}\left(1 - \frac{\alpha m}{T}\right)\left(1 - \left(1 - \frac{\alpha}{T}\right)^{T-m}\right) + \frac{m}{T} = G(T, \alpha, m),$$

which yields the desired estimate

$$\frac{f(S_{gr})}{f(S_{opt})} \geq G(T, \alpha, m)$$

and completes the proof of (5).

We next show the tightness of $\widetilde{G}(T, \alpha, n)$ by modifying the proof of [2, Theorem 5.4]. Let $1 \leq T < n$ be any positive numbers. Pick any number $1 \leq r \leq n/2$, let $X = \{a_1, \ldots, a_r, b_1, \ldots, b_{n-r}\}$ and let $f\colon 2^X \to \mathbb{R}_+$ be the set function

$$f(\{a_{i_1}, \ldots, a_{i_s}, b_{j_1}, \ldots, b_{j_u}\}) = u + \left(1 - \frac{\alpha u}{T}\right)\sum_{k=1}^{s}\left(1 - \frac{\alpha}{T}\right)^{i_k - 1},$$

defined for all subsets $\{a_{i_1}, \ldots, a_{i_s}, b_{j_1}, \ldots, b_{j_u}\} \subset X$. Then $f(\emptyset) = 0$. For any $S = \{a_{i_1}, \ldots, a_{i_s}, b_{j_1}, \ldots, b_{j_u}\} \subsetneq X$, where $s < r$ and $u \leq n - r$, and $a_i \in X \backslash S$, we have

$$f(S \cup \{a_i\}) - f(S) = \left(1 - \frac{\alpha u}{T}\right)\left(1 - \frac{\alpha}{T}\right)^{i-1} \geq 0.$$

For any $S = \{a_{i_1}, \ldots, a_{i_s}, b_{j_1}, \ldots, b_{j_u}\} \subsetneq X$, where $s \leq r$ and $u < n - r$, and $b_j \in X \backslash S$, we have

$$f(S \cup \{b_j\}) - f(S) = 1 - \frac{\alpha}{T} \sum_{k=1}^{s} \left(1 - \frac{\alpha}{T}\right)^{i_k - 1} \geq 0.$$

By recalling that a set function $g\colon 2^X \to \mathbb{R}_+$ is submodular if and only if

$$g(S \cup \{x\}) - g(S) \geq g(R \cup \{x\}) - g(R),$$

for all $S \subset R \subsetneq X$ and $x \in X \backslash R$ (e.g., [1]), these inequalities show that f is submodular and nondecreasing. Moreover,

$$\max \left\{1 - \frac{f(X) - f(X \backslash \{x\})}{f(\{x\})} : x \in X,\ f(\{x\}) \neq 0\right\}$$
$$= 1 - \frac{f(X) - f(X \backslash \{a_i\})}{f(\{a_i\})} = \alpha,$$

for any $1 \leq i \leq r$, so f has total curvature α.

Consider next the case where $T > n/2$. Set $r = n - T$, so that $r < n/2 < T$ and $n - r = T$. It is easy to verify that $S_{opt} = \{b_1, \ldots, b_T\}$ is an optimal solution to problem (1) with $f(S_{opt}) = T$. Since $f(\{a_1\}) = f(\{b_j\}) = 1$, for any $1 \leq j \leq T$, the greedy algorithm can choose the element a_1 at the first iteration. Assume next that the greedy algorithm has chosen $S_{t-1} = \{a_1, \ldots, a_{t-1}\}$ for some $t \leq n - T$. Using the fact

$$\sum_{k=1}^{l} \left(1 - \frac{\alpha}{T}\right)^{k-1} = \frac{T}{\alpha} \left(1 - \left(1 - \frac{\alpha}{T}\right)^{l}\right)$$

it is easy to see that

$$f(S_{t-1} \cup \{a_t\}) = f(S_{t-1} \cup \{b_j\}) = \sum_{i=1}^{t} \left(1 - \frac{\alpha}{T}\right)^{i-1},$$

so the greedy algorithm can choose a_t at the tth iteration. We therefore can have $S_{gr} = \{a_1, \ldots a_{n-T}, b_1, \ldots, b_{2T-n}\}$. This solution has the value

$$f(S_{gr}) = \frac{T}{\alpha} \left(1 - \left(1 - \frac{\alpha m}{T}\right) \left(1 - \frac{\alpha}{T}\right)^{n-T}\right).$$

The claim follows because $m = |S_{opt} \cap S_{gr}| = 2T - n$, whence we obtain $n - T = T - m$.

The proof of case $T \leq n/2$ is easier, so we omit its proof.

Acknowledgements J.L. and A.M were supported by the ERC-StG Grant 260393. A.M. was supported by the Academy of Finland Centre of Excellence programme 2012–2017, Grant 250444, and the Finnish Natural Heritage Services (Metsähallitus).

References

1. Nemhauser, G.L., Wolsey, L.A., Fisher, M.L.: An analysis of approximations for maximizing submodular set functions I. Math. Program. **14**, 265–294 (1978)
2. Conforti, M., Cornuéjols, G.: Submodular set functions, matroids and the greedy algorithm: tight worst-case bounds and some generalizations of the Rado–Edmonds theorem. Discr. Appl. Math. **7**, 251–274 (1984)
3. Edmonds, J.: Matroids and the greedy algorithm. Math. Program. **1**, 127–136 (1971)
4. Cornuéjols, G., Fisher, M.L., Nemhauser, G.L.: Location of bank accounts to optimize float: an analytic study of exact and approximate algorithms. Manag. Sci. **23**, 789–810 (1977)
5. Nemhauser, G.L., Wolsey, L.A.: Best algorithms for approximating the maximum of a submodular set function. Math. Oper. Res. **3**, 177–188 (1978)
6. Goundan, P.R., Schulz, A.S.: Revisiting the greedy approach to submodular set function maximization. Working Paper, Massachusetts Institute of Technology (2007). http://www.optimization-online.org/DB_HTML/2007/08/1740.html
7. Krause, A., Golovin, D.: Submodular function maximization. In: Bordeaux, L., Hamadi, Y., Kohli, P. (eds.) Tractability: Practical Approaches to Hard Problems, pp. 71–104. Cambridge University Press, Cambridge (2014)
8. Topkis, D.M.: Supermodularity and Complementarity. Princeton University Press, Princeton (1998)
9. Bordewich, M., Semple, C.: Budgeted nature reserve selection with diversity feature loss and arbitrary split systems. J. Math. Biol. **64**, 69–85 (2012)
10. Golovin, D., Krause, A., Gardner, B., Converse, S.J., Morey, S.: Dynamic resource allocation in conservation planning. In: Proceeding of the 25th AAAI Conference on Artificial Intelligence, pp. 1331–1336 (2011)
11. Moilanen, A.: Landscape Zonation, benefit functions and target-based planning: unifying reserve selection strategies. Biol. Conserv. **134**, 571–579 (2007)
12. Fisher, M.L., Nemhauser, G.L., Wolsey, L.A.: An analysis of approximations for maximizing submodular set functions II. Math. Program. Study **8**, 73–87 (1978)
13. Kulik, A., Shachnai, H., Tamir, T.: Approximations for monotone and non-monotone submodular maximization with knapsack constraints. Math. Oper. Res. **38**, 729–739 (2013)
14. Sviridenko, M.: A note on maximizing a submodular set function subject to a knapsack constraint. Oper. Res. Lett. **32**, 41–43 (2004)
15. Calinescu, G., Chekuri, C., Pál, M., Vondrák, J.: Maximizing a submodular set function subject to a matroid constraint. SIAM J. Comput. **40**, 1740–1766 (2011)
16. Vondrák, J.: Submodularity and curvature: the optimal algorithm. RIMS Kôkyûroku Bessatsu **B23**, 253–266 (2010)
17. Feige, U., Mirrokni, V.S., Vondrák, J.: Maximizing non-monotone submodular functions. SIAM J. Comput. **40**, 1133–1153 (2011)
18. Wang, Z., Moran, B., Wang, X., Pan, Q.: Approximation for maximizing monotone non-decreasing set functions with a greedy method. J. Comb. Optim. **31**, 29–43 (2016)
19. Il'ev, V.: An approximation guarantee of the greedy descent algorithm for minimizing a supermodular set function. Discr. Appl. Math. **114**, 131–146 (2001)
20. Il'ev, V., Linker, N.: Performance guarantees of a greedy algorithm for minimizing a supermodular set function on comatroid. Eur. J. Oper. Res. **171**, 648–660 (2006)