A Analysis of Greedy Power Density

Denote OPT as an optimal solution for the monotone submodular maximization problem under a knapsack constraint. Let l be the size of the final solution set constructed by Greedy Power Density and S_i be the partial solution set constructed by Greedy Power Density after selecting i elements for $0 \le i \le l$. We define l' as the largest index such that, up to the point when the partial solution set $S_{l'}$ is constructed, no element from the optimal set OPT has been selected (i.e., chosen by maximizing power density) but not added to the solution set due to the budget constraint. Let $u_i := S_i \setminus S_{i-1}$. Hence, for all $1 \le i \le l'$, it holds that

$$\frac{f(u_i \mid S_{i-1})}{c(u_i)^\theta} \ge \max_{e \in OPT \setminus S_{i-1}} \frac{f(e \mid S_{i-1})}{c(e)^\theta}. \tag{1}$$

Let $C_i := \sum_{e \in OPT \setminus S_{i-1}} c(e)^\theta$ and $D := \sum_{i=1}^{l'} c(u_i)^\theta$.

Lemma A.1. *For $i = 1, \cdots, l'$, it holds that*

$$f(S_i) - f(S_{i-1}) \ge \frac{c(u_i)^\theta}{C_i}[f(OPT) - f(S_{i-1})].$$

Proof. Since f is monotone and submodular, we have

$$\begin{aligned} f(OPT) &\le f(OPT \cup S_{i-1}) \\ &\le f(S_{i-1}) + \sum_{e \in OPT \setminus S_{i-1}} f(e \mid S_{i-1}) \\ &= f(S_{i-1}) + \sum_{e \in OPT \setminus S_{i-1}} c(e)^\theta \cdot \frac{f(e \mid S_{i-1})}{c(e)^\theta} \\ &\le f(S_{i-1}) + \frac{f(u_i \mid S_{i-1})}{c(u_i)^\theta} \cdot C_i \\ &= f(S_{i-1}) + \frac{f(S_i) - f(S_{i-1})}{c(u_i)^\theta} \cdot C_i, \end{aligned}$$

where the third inequality holds by Inequality (1). The lemma follows by rearranging the terms. □

Lemma A.2. *Let $\{a_i\}_{i=0}^m$ be a sequence with $a_0 = 0$, $\{b_i\}_{i=1}^m$ be a sequence with $0 \le b_i \le 1$ and $c > 0$ be a constant satisfying*

$$a_i - a_{i-1} \ge b_i(c - a_{i-1}). \tag{2}$$

Then for every $i \in \{1, \dots, m\}$, it holds that

$$a_i \ge \left[1 - \prod_{j=1}^{i}(1 - b_j)\right] c. \tag{3}$$

Proof. We proceed by induction on i. From (2) with $i = 1$, it holds that $a_1 \ge b_1 c$, which matches (3) for $i = 1$. Assume (3) holds for $i - 1$. Then from (2), it holds that $a_i \ge (1 - b_i)a_{i-1} + b_i c$. Applying the induction hypothesis to a_{i-1}, it holds that $a_i \ge (1 - b_i)\left[1 - \prod_{j=1}^{i-1}(1 - b_j)\right] c + b_i c$. Factoring out c and simplifying gives $a_i \ge \left[1 - \prod_{j=1}^{i}(1 - b_j)\right] c$, which completes the induction. □

Lemma A.3. *If $0 \le b_i \le 1$ for all $1 \le i \le n$, it holds that $1 - \prod_{i=1}^n (1 - b_i) \ge 1 - e^{-\sum_{i=1}^n b_i}$.*

Proof. For $x \in [0, 1]$, it holds that $\ln(1 - x) \le -x$. Hence, it follows that

$$\prod_{i=1}^n (1 - b_i) = \exp\left(\sum_{i=1}^n \ln(1 - b_i)\right) \le \exp\left(-\sum_{i=1}^n -b_i\right),$$

which proves the lemma. □

Definition A.1 (Satisfiable Bounds). *We call a pair of bounds (A, B) (possibly dependent on the cost ratio r and budget b) is satisfiable if and only if*

(1) $A \ge C_i$ for all $1 \le i \le l'$;
(2) $c(u_i)^\theta / A \le 1$ for all u_i $(1 \le i \le l')$;
(3) $B \le D$.

Lemma A.4. *Given a satisfiable bounds (A, B), it holds that*

$$\frac{f(S_{l'})}{f(OPT)} \ge 1 - \exp\left(-\frac{B}{A}\right). \tag{4}$$

In particular, if $\theta = 0$, it holds that

$$\frac{f(S_{l'})}{f(OPT)} \ge 1 - \left(1 - \frac{1}{A}\right)^B. \tag{5}$$

Proof. Since $C_i \le A$ for all $1 \le i \le l'$, it holds that

$$\begin{aligned} f(S_i) - f(S_{i-1}) &\ge \frac{c(u_i)^\theta}{C_i}[f(OPT) - f(S_{i-1})] \\ &\ge \frac{c(u_i)^\theta}{A}[f(OPT) - f(S_{i-1})] \end{aligned}$$

for $i = 1, \cdots, l'$ by Lemma A.1.

Using Lemma A.2 by setting $a_i = f(S_i)$ with $a_0 = f(S_0) = 0$, $m = l'$, $b_i = \frac{c(u_i)^\theta}{A} \in [0, 1]$ and $c = f(OPT) > 0$, we could obtain that

$$f(S_i) \ge \left[1 - \prod_{j=1}^{i}\left(1 - \frac{c(u_j)^\theta}{A}\right)\right] f(OPT)$$

for $i = 1, \cdots, l'$. Consider the solution set $S_{l'}$, it holds that

$$f(S_{l'}) \ge \left[1 - \prod_{j=1}^{l'}\left(1 - \frac{c(u_j)^\theta}{A}\right)\right] f(OPT).$$

Using Lemma A.3 by setting $b_j := c(u_j)^\theta / A$, we have

$$1 - \prod_{j=1}^{l'}\left(1 - \frac{c(u_j)^\theta}{A}\right) \ge 1 - \exp\left(-\frac{D}{A}\right).$$

Furthermore, since $D \ge B$, it follows that $1 - \exp\left(-\frac{D}{A}\right) \ge 1 - \exp\left(-\frac{B}{A}\right)$. Hence, we have

$$\frac{f(S_{l'})}{f(OPT)} \ge 1 - \exp\left(-\frac{B}{A}\right).$$

In particular, if $\theta = 0$, it holds that $D = l' \ge B$ and

$$1 - \prod_{j=1}^{l'}\left(1 - \frac{c(u_j)^\theta}{A}\right) = 1 - \left(1 - \frac{1}{A}\right)^{l'} \ge 1 - \left(1 - \frac{1}{A}\right)^B.$$

□

Lemma A.5. *It holds that $C_i \le \min\{\lfloor b \rfloor r^\theta, \lfloor b \rfloor^{1-\theta} b^\theta\}$ for all $1 \le i \le l'$.*

Proof. For each $e \in OPT$, it holds that $1 \le c(e) \le r$. Then it follows that $C_i \le |OPT \setminus S_{i-1}| \cdot r^\theta \le \lfloor b \rfloor r^\theta$. Since $\theta \in [0,1]$ and OPT is a feasible solution, by Hölder's inequality, it holds that

$$\sum_{e \in OPT \setminus S_{i-1}} c(e)^\theta \le |OPT \setminus S_{i-1}|^{1-\theta} \left(\sum_{e \in OPT \setminus S_{i-1}} c(e) \right)^\theta \le \lfloor b \rfloor^{1-\theta} b^\theta.$$

Together with the earlier bound $C_i \le \lfloor b \rfloor r^\theta$, it holds that $C_i \le \min\{\lfloor b \rfloor r^\theta, \lfloor b \rfloor^{1-\theta} b^\theta\}$. □

Define

$$\alpha_1(r,b) := \begin{cases} b-r, & b < r\lfloor b \rfloor, \\ \lfloor b \rfloor, & \text{otherwise.} \end{cases} \quad (6)$$

Lemma A.6. *For the solution set $S_{l'}$, it holds that $c(S_{l'}) \ge \alpha_1(r,b)$ where $\alpha_1(r,b)$ is defined as Equation (6).*

Proof. We assume that the budget b is strictly less than the cost $c(V)$ of the ground set V for monotone submodular maximization under a knapsack constraint.

We always have $c(S_{l'}) \ge b - r$. If $l' < l$, for the solution set $S_{l'}$, there exists an element $e^* \in OPT \setminus S_{l'}$ such that $c(e^*) + c(S_{l'}) > b$, implying that $c(S_{l'}) > b - r$ since $c(e^*) \le r$. If $l' = l$, we prove it by contradiction. If $c(S_l) < b - r$, there exists $e^* \in V \setminus S_l$ such that $c(e^*) + c(S_l) \le r + c(S_l) \le b$, which contradicts the fact that S_l is the output of the Greedy algorithm. Hence, it holds that $c(S_l) \ge b - r$.

If $b \ge r\lfloor b \rfloor$, we assert that $|S_{l'}| \ge \lfloor b \rfloor$. In fact, it holds that $\lfloor b \rfloor \le b/r \le b$, implying that $\lfloor b/r \rfloor = \lfloor b \rfloor$. Hence, we have $c(S_{l'}) = \sum_{e \in S_{l'}} c(e) \ge |S_{l'}| \ge \lfloor b/r \rfloor = \lfloor b \rfloor$ because we can choose at least $\lfloor b/r \rfloor = \lfloor b \rfloor$ elements without discarding any element due to budget violation. □

Lemma A.7. *It holds that $l' \ge \lfloor b/r \rfloor$.*

Proof. By the definition of cost ratio r, for any element $e \in V$, it holds that $c(e) \le r$. If $l' < \lfloor b/r \rfloor$, it holds that

$$c(S_{l'}) + c(e) \le (\lfloor b/r \rfloor - 1)r + r \le b \quad (7)$$

for any element $e \in V \setminus S_{l'}$.

By the definition of l', there exists an element $e^* \in OPT \setminus S_{l'}$ such that $e^* = \arg\max_{e \in V \setminus S_{l'}} \frac{f(e|S_{l'})}{c(e)^\theta}$ and $c(S_{l'}) + c(e^*) > b$, which contradicts Inequality (7).

Hence, $l' \ge \lfloor b/r \rfloor$. □

Lemma A.8. *It holds that $D \ge \max\{\lfloor b/r \rfloor, \alpha_1(r,b)/r^{1-\theta}\}$.*

Proof. For any element $e \in V$ and $\theta \in [0,1]$, it holds that $c(e)^\theta \ge 1$. Hence, it follows that $D = \sum_{i=1}^{l'} c(u_i)^\theta \ge l'$. Since $l' \ge \lfloor b/r \rfloor$ (Lemma A.7), it holds that $D \ge \lfloor b/r \rfloor$.

By the definition of cost ratio r, it holds that $c(e) \le r$, implying that $c(e)^\theta \ge c(e)/r^{1-\theta}$. Hence, it follows that $D = \sum_{i=1}^{l'} c(u_i)^\theta \ge \frac{c(S_{l'})}{r^{1-\theta}}$. Since $c(S_{l'}) \ge \alpha_1(r,b)$ (Lemma A.6), it holds that $D \ge \frac{\alpha_1(r,b)}{r^{1-\theta}}$.

Combining the above bounds gives

$$D \ge \max\{\lfloor b/r \rfloor, \alpha_1(r,b)/r^{1-\theta}\}.$$

□

In the following, we specify these lower and upper bounds to derive approximation guarantees.

Let $A_1 := \lfloor b \rfloor r^\theta$ and $A_2 := \lfloor b \rfloor^{1-\theta} b^\theta$. From Lemma A.5, it holds that $C_i \le \min\{A_1, A_2\}$. Let $B_1 := \lfloor b/r \rfloor$, $B_2 := \frac{\alpha_1(r,b)}{r^{1-\theta}}$. From Lemma A.8, it holds that $D \ge \max\{B_1, B_2\}$.

Then we could conclude the approximation ratio of Greedy Power Density as the following theorem from Lemma A.4:

Theorem A.9. *[Theorem 2.2 in the main body] For any $\theta \in [0,1]$, Greedy Power Density achieves an approximation ratio of $1 - e^{-\gamma(\theta)}$ where $\gamma(\theta) := \frac{\max\{B_1,B_2\}}{\min\{A_1,A_2\}}$. In particular, it achieves $1 - \left(1 - \frac{1}{\min\{A_1,A_2\}}\right)^{\max\{B_1,B_2\}}$-approximation if $\theta = 0$.*

B Corollaries of Greedy Power Density

Lemma B.1. *When $\theta = 0$, Greedy Power Density achieves an approximation ratio of $1 - (1 - 1/\lfloor b \rfloor)^{\max\{\lfloor b/r \rfloor, \alpha_1(r,b)/r\}}$.*

Proof. By setting $\theta = 0$ in Theorem A.9, the lemma follows. □

Lemma B.2. *When (r,b) satisfies either (i) $r \le b/\lfloor b \rfloor$; or (ii) $r > b/\lfloor b \rfloor$, $2 \le \lfloor b/r \rfloor \le 5$ and $\lfloor b \rfloor \le \lfloor b/r \rfloor + 1$, the approximation ratio of Greedy Power Density achieves a $(5/9)$-approximation ratio when $\theta = 0$.*

Proof. By Lemma B.1, it holds that Greedy Power Density with $\theta = 0$ achieves an approximation ratio of $1 - (1 - 1/\lfloor b \rfloor)^{\max\{\lfloor b/r \rfloor, \alpha_1(r,b)/r\}} \ge 1 - (1 - 1/\lfloor b \rfloor)^{\lfloor b/r \rfloor}$.

Case 1: if $r \le b/\lfloor b \rfloor$, it holds that $b/r \ge \lfloor b \rfloor$, implying that $\lfloor b \rfloor = \lfloor b/r \rfloor$ since $1 \le r \le b$. Hence, we have $1 - (1 - 1/\lfloor b \rfloor)^{\lfloor b/r \rfloor} \ge 1 - e^{-1} > 5/9$.

Case 2: $r > b/\lfloor b \rfloor$. Let $k := \lfloor b/r \rfloor$ and $m = \lfloor b \rfloor$. The condition $1 - (1 - 1/m)^k \ge 5/9$ is equivalent to $m \le \frac{1}{1-(9/4)^{-1/k}}$. Evaluating this inequality for $k = 2, 3, 4, 5$ gives the integer thresholds $k = 2 : m \le 3, k = 3 : m \le 4, k = 4 : m \le 5, k = 5 : m \le 6$. Combining with these results, we could conclude that $\lfloor b \rfloor \le \lfloor b/r \rfloor + 1$. (For $k \ge 6$, we have $k = 6 : m \le 7$ and $k = 7 : m \le 9$, implying that when $k \ge 7$, we could relax the bound of $\lfloor b \rfloor$ into $\lfloor b/r \rfloor + 2$. Since we only focus on $2 \le k \le 5$, we stop the discussion for $k \ge 6$.) □

Lemma B.3. *When $\theta = 1$ and $r > b/\lfloor b \rfloor$, Greedy Power Density achieves an approximation ratio of $1 - e^{-\max\{\lfloor b/r \rfloor, b-r\}/b}$.*

Proof. By setting $\theta = 1$ in Theorem A.9, the lemma follows. □

Lemma B.4. *When $r > b/\lfloor b \rfloor$, the approximation ratio of Greedy Power Density when $\theta = 1$ achieves $5/9$ approximation ratio if $b/r \ge (1 - \ln(9/4))^{-1} \approx 5.2890$.*

Proof. By Lemma B.3, it holds that when $r > b/\lfloor b \rfloor$, Greedy Power Density when $\theta = 1$ achieves an approximation ratio of $1 - e^{-\max\{\lfloor b/r \rfloor, b-r\}/b} \ge 1 - e^{-1+r/b}$.

It is easy to verify that $1 - e^{-1+r/b} \ge \frac{5}{9}$ is equivalent to $b/r \ge (1 - \ln(9/4))^{-1}$. □

Concluding Lemma B.2 and Lemma B.4, we have identified three regions in the (r, b)-plane:

$$A := \{(r, b) \in S : r \leq b/\lfloor b \rfloor\},$$
$$B := \{(r, b) \in S : b/r \geq (1 - \ln(9/4))^{-1}\},$$
$$C := \{(r, b) \in S : \lfloor b/r \rfloor \in \{2, 3, 4, 5\},\ \lfloor b \rfloor \leq \lfloor b/r \rfloor + 1\}$$

where Greedy with Power Density could achieve a $(5/9)$-approximation ratio and proves Lemma 2.4.