A. Missing Proofs from Section 4

A.1. Proof of Lemma 1

Proof. The proof is constructive and is inspired by (Calinescu et al., 2011; Han et al., 2020). For clarity, we provide a procedure to construct $\sigma_i(\cdot)$, as shown by Algorithm 3. Suppose that the elements in S_i are $\{z_1, \cdots, z_q\}$ (listed according to the order that they are added into S_i). Algorithm 3 finds a series of sets $J_0 \subseteq J_1 \subseteq \cdots \subseteq J_q = Q_i$ such that all the elements in $M_t = J_t \setminus J_{t-1}$ is mapped to z_t by $\sigma_i(\cdot)$ for any $t \in \{1, 2, \cdots, q\}$. From Algorithm 3, it can be easily seen that $\sigma_i(\cdot)$ satisfies the conditions required by the lemma. The only problem left is to prove that all the elements in Q_i is mapped by $\sigma_i(\cdot)$, i.e., to prove $J_0 = \emptyset$. Indeed, we can prove a stronger result $\forall t \in \{0, 1, \cdots, q\} : |J_t| \leq kt$ by induction:

- When $t = q$, we will prove $|J_q| \leq kq$ by showing that S_i is a base of $Q_i \cup S_i$. It is obvious that each element $u \in O_i^-$ satisfies $S_i \cup \{u\} \notin \mathcal{I}$ according to the definition of O_i^-. Moreover, for any element $u \in \cup_{j \in [\ell] \setminus \{i\}} (O_j^{i-} \cup \widehat{O}_j^{i-})$, we must have $S_i \cup \{u\} \notin \mathcal{I}$, because otherwise we have $S_i^<(u) \cup \{u\} \in \mathcal{I}$ due to $S_i^<(u) \subseteq S_i$ and the down-closed property of independence systems, contradicting the definition of O_j^{i-} and $\widehat{O}_j^{i-}$. These reasoning implies that S_i is a base of $Q_i \cup S_i$. Note that $Q_i \subseteq O$. So we can get $|J_q| = |Q_i| \leq k|S_i| = kq$ according to the definition of k-systems.
- Suppose that $|J_t| \leq kt$ holds, we will prove $|J_{t-1}| \leq k(t-1)$. If the set C_t determined in Line 3 of Algorithm 3 has a cardinality larger than k, then we have $|M_t| = k$ according to Algorithm 3 and hence $|J_{t-1}| = |J_t| - k \leq k(t-1)$. If $|C_t| \leq k$, then $\{z_1, \cdots, z_{t-1}\}$ must be a base of $\{z_1, \cdots, z_{t-1}\} \cup J_{t-1}$, because there does not exist $u \in J_{t-1} \setminus \{z_1, \cdots, z_{t-1}\}$ such that $\{z_1, \cdots, z_{t-1}\} \cup \{u\} \in \mathcal{I}$ according to Algorithm 3. So we also have $|J_{t-1}| \leq k(t-1)$ according to $J_{t-1} \in \mathcal{I}$ and the definition of k-systems.

From the above reasoning we know $J_0 = \emptyset$. So the lemma follows. □

Algorithm 3 CONSTRUCTING THE MAPPING $\sigma_i(\cdot)$

Initialize: Denote the elements in S_i as $\{z_1, \cdots, z_q\}$, where elements are listed according to the order that they are added into S_i; $J_q \leftarrow Q_i$

```
1:  for t = q to 0 do
2:    C_t ← {e ∈ J_t \ {z_1, ···, z_{t-1}} : {z_1, ···, z_{t-1}, e} ∈ I}
3:    if |C_t| ≤ k then
4:      M_t ← C_t
5:    end if
6:    if |C_t| > k then
7:      if z_t ∈ C_t then
8:        Find a subset M_t ⊆ C_t satisfying |M_t| = k and z_t ∈ M_t
9:      else
10:       Find a subset M_t ⊆ C_t satisfying |M_t| = k
11:     end if
12:   end if
13:   Let σ_i(z) = z_t for all z ∈ M_t; J_{t-1} ← J_t \ M_t
14: end for
```

A.2. Proof of Lemma 2

Proof. We first prove Eqn. (3). According to the definitions of O_j^{i+} and $\widehat{O}_j^{i+}$, any element $u \in O_j^{i+} \cup \widehat{O}_j^{i+}$ can also be added into S_i without violating the feasibility of $\mathcal{I}$ when u is inserted into S_j. Therefore, according to the greedy rule of RANDOMMULTIGREEDY and the submodularity of $f(\cdot)$, we must have

$$\forall u \in O_j^{i+} \cup \widehat{O}_j^{i+} : f(u \mid S_i) \leq f(u \mid S_i^<(u))) \leq f(u \mid S_j^<(u))) = \delta(u) \tag{11}$$

Now we prove Eqn. (4). Recall that $Q_i = \cup_{j \in [\ell] \setminus \{i\}} (O_j^{i-} \cup \widehat{O}_j^{i-}) \cup (O \cap S_i) \cup O_i^-$. According to Lemma 1, any element $u \in O_j^{i-} \cup \widehat{O}_j^{i-} (j \neq i)$ can be added into $S_i^<(\pi_i(u))$ without violating the feasibility of $\mathcal{I}$. Moreover, u must have not been considered by the algorithm at the moment that $\pi_i(u)$ is added into S_i, because otherwise we have $S_i^{<}(u) \subseteq S_i^{<}(\pi_i(u))$ and hence $S_i^{<}(u) \cup \{u\} \in \mathcal{I}$ due to the definition of independence systems, which contradicts the definitions of O_j^{i-} and $\widehat{O}_j^{i-}$. Therefore, according to the greedy rule of RANDOMMULTIGREEDY and submodularity, we can get

$$\forall u \in O_j^{i-} \cup \widehat{O}_j^{i-} : f(u \mid S_i) \leq f(u \mid S_i^{<}(u))) \leq f(u \mid S_i^{<}(\pi_i(u))) \leq f(\pi_i(u) \mid S_i^{<}(\pi_i(u))) = \delta(\pi_i(u)) \tag{12}$$

By similar reasoning, we can also prove $\forall u \in O_i^- : f(u \mid S_i) \leq f(u \mid S_i^{<}(\pi_i(u))) \leq \delta(\pi_i(u))$. Finally, $f(u \mid S_i) \leq \delta(\pi_i(u))$ trivially holds for all $u \in O \cap S_i$ as $\pi_i(u) = u$ due to Lemma 1. So the lemma follows. □

A.3. Proof of Lemma 3

As the proof of Lemma 3 is a bit involved, we first introduce Lemma 9, and then use Lemma 9 to prove Lemma 3.

Lemma 9. *We have*

$$\sum_{i\in[\ell]} \left(\sum_{j\in[\ell]\setminus\{i\}} \left(\sum_{u\in O_j^{i+}} \delta(u) + \sum_{u\in O_j^{i-}\cup\widehat{O}_j^{i-}} \delta(\pi_i(u)) \right) + \sum_{u\in O_i^-} \delta(\pi_i(u)) \right) \leq \ell(k+\ell-2) f(S^*) \tag{13}$$

Proof. For any $i \in [\ell]$, let $\lambda(i) = (i \mod \ell) + 1$. So we have

$$\sum_{i\in[\ell]} \sum_{u\in O_{\lambda(i)}^{i+}} \delta(u) = \sum_{j\in[\ell]} \sum_{u\in O_j^{\lambda^{-1}(j)+}} \delta(u) \leq \sum_{j\in[\ell]} \sum_{u\in O\cap S_j} \delta(u) = \sum_{i\in[\ell]} \sum_{u\in O\cap S_i} \delta(u), \tag{14}$$

where the inequality is due to $O_j^{\lambda^{-1}(j)+} \subseteq O \cap S_j$ and $\forall u \in S_j : \delta(u) > 0$. So we can get

$$\begin{aligned}
\sum_{i\in[\ell]} \sum_{j\in[\ell]\setminus\{i\}} \sum_{u\in O_j^{i+}} \delta(u) &= \sum_{i\in[\ell]} \left(\sum_{j\in[\ell]\setminus\{i,\lambda(i)\}} \sum_{u\in O_j^{i+}} \delta(u) + \sum_{u\in O_{\lambda(i)}^{i+}} \delta(u) \right) \\
&\leq \sum_{i\in[\ell]} \sum_{j\in[\ell]\setminus\{i,\lambda(i)\}} \sum_{u\in O\cap S_j} \delta(u) + \sum_{i\in[\ell]} \sum_{u\in O\cap S_i} \delta(u) \qquad (15) \\
&\leq \ell(\ell-2) f(S^*) + \sum_{i\in[\ell]} \sum_{u\in O\cap S_i} \delta(u) \qquad (16)
\end{aligned}$$

where we leverage Eqn. (14) to derive Eqn. (15), and Eqn. (16) is due to $\sum_{u\in O\cap S_j} \delta(u) \leq \sum_{u\in S_j} \delta(u) \leq f(S_j) \leq f(S^*)$. Moreover, we can get

$$\begin{aligned}
&\sum_{i\in[\ell]} \left(\sum_{j\in[\ell]\setminus\{i\}} \left(\sum_{u\in O_j^{i+}} \delta(u) + \sum_{u\in O_j^{i-}\cup\widehat{O}_j^{i-}} \delta(\pi_i(u)) \right) + \sum_{u\in O_i^-} \delta(\pi_i(u)) \right) \\
=\ & \sum_{i\in[\ell]} \sum_{j\in[\ell]\setminus\{i\}} \sum_{u\in O_j^{i+}} \delta(u) + \sum_{i\in[\ell]} \left(\sum_{j\in[\ell]\setminus\{i\}} \sum_{u\in O_j^{i-}\cup\widehat{O}_j^{i-}} \delta(\pi_i(u)) + \sum_{u\in O_i^-} \delta(\pi_i(u)) \right) \\
\leq\ & \ell(\ell-2) f(S^*) + \sum_{i\in[\ell]} \left(\sum_{u\in O\cap S_i} \delta(u) + \sum_{j\in[\ell]\setminus\{i\}} \sum_{u\in O_j^{i-}\cup\widehat{O}_j^{i-}} \delta(\pi_i(u)) + \sum_{u\in O_i^-} \delta(\pi_i(u)) \right) \qquad (17) \\
=\ & \ell(\ell-2) f(S^*) + \sum_{i\in[\ell]} \sum_{u\in Q_i} \delta(\pi_i(u)) \\
\leq\ & \ell(\ell-2) f(S^*) + k \sum_{i\in[\ell]} \sum_{u\in S_i} \delta(u) \qquad (18) \\
\leq\ & \ell(\ell-2) f(S^*) + k \sum_{i\in[\ell]} f(S_i) \leq \ell(k+\ell-2) f(S^*) \qquad (19)
\end{aligned}$$

where $Q_i = \cup_{j\in[\ell]\setminus\{i\}}(O_j^{i-} \cup \widehat{O}_j^{i-}) \cup (O \cap S_i) \cup O_i^-$ is defined in Lemma 1; Eqn. (17) is due to Eqn. (16); and Eqn. (18) is due to Lemma 1. So the lemma follows. □

Now we provide the proof of Lemma 3:

Proof. Let $G_i = [\cup_{i\in[\ell]\setminus\{i\}}(O_j^{i+} \cup O_j^{i-} \cup \widehat{O}_j^{i+} \cup \widehat{O}_j^{i-})] \cup O_i^- \cup [O \cap D_i]$ for all $i \in [\ell]$. It is not hard to see that $G_i \subseteq O \setminus S_i$ and $\forall u \in O \setminus (S_i \cup G_i) : f(u \mid S_i) \leq 0$. Therefore, we can get

$$\begin{aligned}
&\sum_{i\in[\ell]}\Big(f(O\cup S_i) - f(S_i)\Big)\\
\leq\ &\sum_{i\in[\ell]}\Bigg(\sum_{j\in[\ell]\setminus\{i\}}\Bigg(\sum_{u\in O_j^{i+}\cup\widehat{O}_j^{i+}} f(u\mid S_i) + \sum_{u\in O_j^{i-}\cup\widehat{O}_j^{i-}} f(u\mid S_i)\Bigg) + \sum_{u\in O_i^-} f(u\mid S_i) + \sum_{u\in O\cap D_i} f(u\mid S_i)\Bigg) &(20)\\
\leq\ &\sum_{i\in[\ell]}\Bigg(\sum_{j\in[\ell]\setminus\{i\}}\Bigg(\sum_{u\in O_j^{i+}\cup\widehat{O}_j^{i+}} \delta(u) + \sum_{u\in O_j^{i-}\cup\widehat{O}_j^{i-}} \delta(\pi_i(u))\Bigg) + \sum_{u\in O_i^-} \delta(\pi_i(u)) + \sum_{u\in O\cap D_i} \delta(u)\Bigg) &(21)\\
=\ &\sum_{i\in[\ell]}\Bigg(\sum_{j\in[\ell]\setminus\{i\}}\Bigg(\sum_{u\in O_j^{i+}} \delta(u) + \sum_{u\in O_j^{i-}\cup\widehat{O}_j^{i-}} \delta(\pi_i(u))\Bigg) + \sum_{u\in O_i^-} \delta(\pi_i(u))\Bigg)\\
&+\sum_{i\in[\ell]}\Bigg(\sum_{j\in[\ell]\setminus\{i\}}\sum_{u\in\widehat{O}_j^{i+}} \delta(u) + \sum_{u\in O\cap D_i}\delta(u)\Bigg)\\
\leq\ &\ell(k+\ell-2)f(S^*) + \sum_{i\in[\ell]}\Bigg(\sum_{j\in[\ell]\setminus\{i\}}\sum_{u\in\widehat{O}_j^{i+}} \delta(u) + \sum_{u\in O\cap D_i}\delta(u)\Bigg) &(22)
\end{aligned}$$

where Eqn. (20) is due to submodularity of $f(\cdot)$; Eqn. (21) is due to Lemma 2 and submodularity; and Eqn. (22) is due to Lemma 9. Moreover, we can get

$$\begin{aligned}
&\sum_{i\in[\ell]}\Bigg(\sum_{j\in[\ell]\setminus\{i\}}\sum_{u\in\widehat{O}_j^{i+}} \delta(u) + \sum_{u\in O\cap D_i}\delta(u)\Bigg) \leq \sum_{i\in[\ell]}\Bigg(\sum_{j\in[\ell]\setminus\{i\}}\sum_{u\in O\cap D_j} \delta(u) + \sum_{u\in O\cap D_i}\delta(u)\Bigg) &(23)\\
=\ &\sum_{i\in[\ell]}\sum_{j\in[\ell]}\sum_{u\in O\cap D_j}\delta(u) = \ell\sum_{u\in\mathcal{N}} X_u\cdot\delta(u) &(24)
\end{aligned}$$

where Eqn. (23) is due to $\widehat{O}_j^{i+} \subseteq O \cap D_j$ and $\forall u \in D_j : \delta(u) > 0$. Combining Eqn. (22) and Eqn. (24) finishes the proof of Lemma 3. □

A.4. Proof of Lemma 4

We first quote the following lemma presented in (Buchbinder et al., 2014):

Lemma 10. *(Buchbinder et al., 2014) Given a ground set $\mathcal{N}$ and any non-negative submodular function $g(\cdot)$ defined on $2^{\mathcal{N}}$, we have $\mathbb{E}[g(Y)] \geq (1-p)g(\emptyset)$ if Y is a random subset of $\mathcal{N}$ such that each element in $\mathcal{N}$ appears in Y with probability of at most p (not necessarily independently).*

With the above lemma, Lemma 4 can be proved as follows:

Proof. We first prove Eqn. (6). Note that $S_1, S_2, \cdots, S_\ell$ are disjoint sets. Using submodularity, we have

$$\begin{aligned}
&\sum_{i=1}^{\ell} f(S_i\cup O) \geq f(O) + f(S_1\cup S_2\cup O) + \sum_{i=3}^{\ell} f(S_i\cup O)\\
\geq\ &2f(O) + f(S_1\cup S_2\cup S_3\cup O) + \sum_{i=4}^{\ell} f(S_i\cup O) \geq \cdots \geq (\ell-1)f(O) + f(\cup_{i=1}^{\ell} S_i\cup O) &(25)
\end{aligned}$$

Let $g : 2^{\mathcal{N}} \mapsto \mathbb{R}_{\geq 0}$ be a non-negative submodular function defined as: $\forall S \subseteq \mathcal{N} : g(S) = f(S \cup O)$. As each element in $\mathcal{N}$ appears in $\cup_{i=1}^{\ell} S_i$ with probability of no more than p, We can use Lemma 10 to get

$$\mathbb{E}[f(\cup_{i=1}^{\ell} S_i \cup O)] = \mathbb{E}[g(\cup_{i=1}^{\ell} S_i)] \geq (1-p)g(\emptyset) = (1-p)f(O) \tag{26}$$

Combining Eqn. (25) and Eqn. (26) finishes the proof of Eqn. (6).

Next, we prove Eqn. (7). For any $u \in \mathcal{N}$, let $Y_u = 1$ if $u \in \cup_{i=1}^{\ell} S_i$ and $Y_u = 0$ otherwise; let $\mathcal{E}_u$ be an arbitrary event denoting all the random choices of RANDOMMULTIGREEDY up until the time that u is considered to be added into a candidate solution, or denoting all the randomness of RANDOMMULTIGREEDY if u is never considered. Note that we have $\sum_{u \in \mathcal{N}} Y_u \cdot \delta(u) \leq \sum_{i=1}^{\ell} f(S_i)$. Therefore, by the law of total probability, we only need to prove

$$\forall u \in \mathcal{N} : \frac{1-p}{p} \mathbb{E}[Y_u \cdot \delta(u) \mid \mathcal{E}_u] \geq \mathbb{E}[X_u \cdot \delta(u) \mid \mathcal{E}_u] \tag{27}$$

for any event $\mathcal{E}_u$ defined above. Note that we have $X_u = 0$ and hence Eqn. (27) clearly holds if $u \notin O$ or u is never considered by the algorithm. Otherwise we have $\mathbb{E}[Y_u \cdot \delta(u) \mid \mathcal{E}_u] = p \cdot \delta(u)$ and $\mathbb{E}[X_u \cdot \delta(u) \mid \mathcal{E}_u] = (1-p) \cdot \delta(u)$ due to the reason that u is accepted with probability of p and discarded with probability of $1-p$. Combining all these results completes the proof of Eqn. (7). □

A.5. Proof of Theorem 2

For clarity, we first provide the detailed design of the accelerated version of RANDOMMULTIGREEDY, as shown by Algorithm 5. In the t-th iteration, Algorithm 5 calls a procedure CHOOSE to greedily find an candidate element v_i for S_i satisfying $f(v_i \mid S_i) > 0$ and $S_i \cup \{v_i\} \in \mathcal{I}$ for each $i \in [\ell]$. The CHOOSE procedure also returns an index i_t same to that in Algorithm 1. After that, Algorithm 5 runs similarly as Algorithm 1, i.e., it inserts v_{i_t} into S_{i_t} with probability p, and then enters the $(t+1)$-th iteration. Note that the elements $v_1, \cdots, v_\ell$ and v_{i_t} found in the t-th iteration are also used to call CHOOSE in the $(t+1)$-th iteration, so that CHOOSE need not to identify a new v_i for all $i \in [\ell] : v_i \neq v_{i_t}$ (as S_i does not change for these i's) and hence time efficiency can be improved. Finally, Algorithm 5 returns the optimal set among $S_1, \cdots, S_\ell$ and S_0, where S_0 is the singleton set with the maximum utility.

Next, we provide a brief description on the CHOOSE procedure. As explained in Sec. 4.1, CHOOSE maintains ℓ sets $A_1, A_2, \cdots, A_\ell$ such that v_i can be selected from A_i. At the first time that CHOOSE is called, CHOOSE assigns each element $u \in A_i$ a weight $w_i(u) = f(u \mid \emptyset)$ and an integer $\tau_i(u)$ indicating how many times $w_i(u)$ has been updated (Lines 3–7). Afterwards, CHOOSE runs as that described in Sec. 4.1 and finds v_i for each $i \in [\ell]$. Finally, CHOOSE identifies v_{i^*} from $\{v_i : i \in [\ell]\}$ which has the maximum marginal gain, and it also removes v_{i^*} from all $A_i : i \in [\ell]$ because v_{i^*} will used as v_{i_t} by Algorithm 5.

Note that Algorithm 5 differs from Algorithm 1 in two points: (1) the element u_t found in the t-th iteration is only an $(\frac{1}{1+\epsilon})$-approximate solution; (2) there are elements removed from A_i due to "too many updates". Based on this observation, we can slightly modify the proofs for Algorithm 1 to prove Theorem 2, as presented below:

Proof. Let L_i denote the set of all elements removed from A_i due to Line 25 of Algorithm 4. We can slightly modify Definition 5 to re-define the sets $O_j^{i+}, O_j^{i-}, \widehat{O}_j^{i+}, \widehat{O}_j^{i-}, O_i^-$ as follows:

$$O_j^{i+} = \{u \in O \cap S_j : S_i^<(u) \cup \{u\} \in \mathcal{I}\} \setminus L_i;$$
$$O_j^{i-} = \{u \in O \cap S_j : S_i^<(u) \cup \{u\} \notin \mathcal{I}\} \setminus L_i;$$
$$\widehat{O}_j^{i+} = \{u \in O \cap D_j : S_i^<(u) \cup \{u\} \in \mathcal{I}\} \setminus L_i;$$
$$\widehat{O}_j^{i-} = \{u \in O \cap D_j : S_i^<(u) \cup \{u\} \notin \mathcal{I}\} \setminus L_i;$$
$$O_i^- = \{u \in O \setminus U : S_i \cup \{u\} \notin \mathcal{I} \wedge f(u \mid S_i) > 0\} \setminus L_i;$$

With this new definition, it can be easily verified that each element u in $O_j^{i+} \cup O_j^{i-} \cup \widehat{O}_j^{i+} \cup \widehat{O}_j^{i-}$ is still a candidate considered for S_i in the CHOOSE procedure when the algorithm tries to insert u into S_j. Therefore, according to the greedy rule of RANDOMMULTIGREEDY and the $(1+\epsilon)^{-1}$-approximation ratio of CHOOSE, we can use similar reasoning as that

Algorithm 4 CHOOSE$(S_1, S_2, \cdots, S_\ell, v_1, \cdots, v_\ell, v^*)$

if $\cup_{i=1}^{\ell} S_i = \emptyset$ **then**
  Let $A_i \leftarrow \{u \in \mathcal{N} : \{u\} \in \mathcal{I} \wedge f(u \mid \emptyset) > 0\}$ for all $i \in [\ell]$;
  for all $i \in [\ell]$ **do**
    Let $w_i(u) \leftarrow f(u \mid \emptyset)$ and $\tau_i(u) \leftarrow 0$ for all $u \in A_i$;
    Store A_i as a priority list according to the non-increasing order of $w_i(u) : u \in A_i$ for all $i \in [\ell]$;
    Let $v_i \leftarrow \arg\max_{u \in A_i} w_i(u)$;
  end for
else
  $C \leftarrow [\ell] \backslash \{j \in [\ell] : (v_j \neq v^*) \vee (v_j = \mathrm{NULL})\}$
  for all $i \in C$ **do**
    Let $v_i \leftarrow \mathrm{NULL}$ and remove all elements in A_i with non-positive weights;
    while $A_i \neq \emptyset$ **do**
      pop out the top element u from A_i;
      if $f(u \mid S_i)$ has been computed **then**
        $v_i \leftarrow u$; **exit while**;
      end if
      if $S_i \cup \{u\} \notin \mathcal{I}$ **then**
        continue;
      end if
      $old \leftarrow w_i(u)$; $\tau_i(u) \leftarrow \tau_i(u) + 1$;
      Compute $f(u \mid S_i)$ and let $w_i(u) \leftarrow f(u \mid S_i)$;
      if $w_i(u) \geq \frac{old}{1+\epsilon}$ **then**
        $v_i \leftarrow u$; **exit while**;
      else
        if $\tau_i(u) \leq \lceil \log_{1+\epsilon} \frac{\ell r}{\epsilon} \rceil$ **then**
          re-insert u into A_i and resort the elements in A_i;
        end if
      end if
    end while
  end for
end if
Let $i^* \leftarrow \arg\max_{i \in [\ell] : v_i \neq \mathrm{NULL}} f(v_i \mid S_i)$ and remove v_{i^*} from A_i for all $i \in [\ell]$
Output: $v_1, v_2, \cdots, v_\ell, i^*$

for Lemma 2 to prove

$$\forall u \in O_j^{i+} \cup \widehat{O}_j^{i+} : f(u \mid S_i) \leq (1+\epsilon)\delta(u); \tag{28}$$

$$\forall u \in \cup_{j \in [\ell] \setminus \{i\}} (O_j^{i-} \cup \widehat{O}_j^{i-}) \cup (O \cap S_i) \cup O_i^- : f(u \mid S_i) \leq (1+\epsilon)\delta(\pi_i(u)); \tag{29}$$

With the above results, we can use similar reasoning as that in Lemma 3 to prove:

$$\frac{1}{1+\epsilon} \sum_{i \in [\ell]} f(O \mid S_i) \leq \ell(k+\ell-2) f(S^*) + \ell \sum_{u \in \mathcal{N}} X_u \cdot \delta(u) + \sum_{i \in [\ell]} \sum_{u \in L_i \cap O} f(u \mid S_i) \tag{30}$$

Moreover, we have

$$\sum_{u \in L_i \cap O} f(u \mid S_i) \leq \sum_{u \in L_i \cap O} f(u \mid \emptyset)(1+\epsilon)^{-\lceil \log_{1+\epsilon} \frac{\ell r}{\epsilon} \rceil} \leq \sum_{u \in L_i \cap O} \frac{\epsilon}{\ell r} f(u) \leq \epsilon f(S^*)/\ell \tag{31}$$

where the first inequality is due the reason that the weight of each element $u \in L_i$ have been updated in CHOOSE procedure for more than $\lceil \log_{1+\epsilon} \frac{\ell r}{\epsilon} \rceil$ times and it diminishes by a factor of $\frac{1}{1+\epsilon}$ for each update. Combining Eqn. (30), Eqn. (31) and Lemma 4, we can prove

$$f(O) \leq \left[(1+\epsilon) \frac{\ell(k + \frac{\ell}{p} - 1)}{\ell - p} - \frac{(\ell-1)\epsilon - \epsilon^2}{\ell - p} \right] \mathbb{E}[f(S^*)] \tag{32}$$

Algorithm 5 RANDOMMULTIGREEDY(ℓ, p) /*with acceleration*/

Initialize: $\forall i \in [\ell]: \ S_i \leftarrow \emptyset; v_i \leftarrow \mathrm{NULL}; \ t \leftarrow 1; u_0 \leftarrow \mathrm{NULL};$

repeat
  $(v_1, v_2, \cdots, v_\ell, i_t) \leftarrow$ CHOOSE$(S_1, \cdots, S_\ell, v_1, \cdots, v_\ell, u_{t-1})$
  if $\exists j \in [\ell] : v_j \neq \mathrm{NULL}$ **then**
    $u_t \leftarrow v_{i_t}$;
    With probability p **do** $S_{i_t} \leftarrow S_{i_t} \cup \{u_t\}$
    $t \leftarrow t + 1$
  end if
until $(\forall i \in [\ell] : v_i = \mathrm{NULL})$
$u^* \leftarrow \arg\max_{u \in \mathcal{N} \wedge \{u\} \in \mathcal{I}} f(u); \ S_0 \leftarrow \{u^*\}$
$S^* \leftarrow \arg\max_{S \in \{S_0, S_1, S_2, \cdots, S_\ell\}} f(S); \ T \leftarrow t - 1$
Output: S^*, T

Therefore, the approximation ratio of the accelerated RANDOMMULTIGREEDY algorithm is at most $(1+\epsilon)(1+\sqrt{k})^2$ when $\ell = 2, p = \frac{2}{1+\sqrt{k}}$ (for a randomized algorithm), or at most $(1+\epsilon)(k + \sqrt{k} + \lceil\sqrt{k}\rceil + 1)$ when $\ell = \lceil\sqrt{k}\rceil + 1, p = 1$ (for a deterministic algorithm). Finally, it can be seen that the CHOOSE procedure incurs at most $\mathcal{O}(\log_{1+\epsilon} \frac{\ell r}{\epsilon})$ value and independence oracle queries for each element in each $A_i : i \in [\ell]$. So the total time complexity of the accelerated RANDOMMULTIGREEDY algorithm is at most $\mathcal{O}(\ell n \log_{1+\epsilon} \frac{\ell r}{\epsilon}) = \mathcal{O}(\frac{\ell n}{\epsilon} \log \frac{\ell r}{\epsilon})$, which completes the proof. □

B. Missing Proofs from Section 5

B.1. Proof of Lemma 5

Proof. Given any element set $Y \subseteq \mathcal{N}$ and any realization ϕ, let $g(Y, \phi) := f(Y \cup \mathcal{N}(\pi_{\mathrm{opt}}, \phi), \phi)$. It is easy to verify that the non-negative function $g(\cdot, \phi)$ is submodular. Thus, given a fixed realization ϕ, by Lemma 10, we know that

$$\mathbb{E}_{\pi_\mathcal{A}}[g(\mathcal{N}(\pi_\mathcal{A}, \phi), \phi)] \geq (1-p) g(\emptyset, \phi) \tag{33}$$

Therefore, we have

$$f_{\mathrm{avg}}(\pi_{\mathrm{opt}} @ \pi_\mathcal{A}) = \mathbb{E}_\Phi[\mathbb{E}_{\pi_\mathcal{A}}[g(\mathcal{N}(\pi_\mathcal{A}, \Phi), \Phi)]] \geq \mathbb{E}_\Phi[(1-p) g(\emptyset, \Phi)] = (1-p) f_{\mathrm{avg}}(\pi_{\mathrm{opt}}), \tag{34}$$

which completes the proof. □

B.2. Proof of Lemma 6

Proof. We first give an equivalent expression of the expected utility by a function of conditional expected marginal gains. Given a deterministic policy π and a realization ϕ, for each $u \in \mathcal{N}$, let $Y_u(\phi)$ be a boolean random variable such that $Y_u(\phi) = 1$ if $u \in \mathcal{N}(\pi, \phi)$ and $Y_u(\phi) = 0$ otherwise. Further, denote by $\psi_u^\pi(\phi)$ the partial realization observed by π right before considering u under realization ϕ, and denote by Ψ_u^π a random partial realization right before considering u by π. We also use $Y_u(\psi_u^\pi(\phi))$ to represent $Y_u(\phi)$, since the partial realization $\psi_u^\pi(\phi)$ suffices to determine whether u is added to the solution under realization ϕ. Thus,

$$\begin{aligned}
& \mathbb{E}_\Phi[f(\mathcal{N}(\pi, \Phi), \Phi)] \\
= \ & \mathbb{E}_\Phi\Big[\sum_{u \in \mathcal{N}} \Big(Y_u(\Phi) \cdot \big(f(\mathrm{dom}(\psi_u^\pi(\Phi)) \cup \{u\}, \Phi) - f(\mathrm{dom}(\psi_u^\pi(\Phi)), \Phi)\big)\Big)\Big] \\
= \ & \sum_{u \in \mathcal{N}} \mathbb{E}_{\Psi_u^\pi}\Big[\mathbb{E}_\Phi\Big[Y_u(\Phi) \cdot \big(f(\mathrm{dom}(\Psi_u^\pi) \cup \{u\}, \Phi) - f(\mathrm{dom}(\Psi_u^\pi), \Phi)\big) \ \Big| \ \Phi \sim \Psi_u^\pi\Big]\Big] \\
= \ & \sum_{u \in \mathcal{N}} \mathbb{E}_{\Psi_u^\pi}\Big[Y_u(\Psi_u^\pi) \cdot \Delta(u \mid \Psi_u^\pi)\Big] = \sum_{u \in \mathcal{N}} \mathbb{E}_\Phi\Big[\mathbb{E}_{\Psi_u^\pi}\Big[Y_u(\Psi_u^\pi) \cdot \Delta(u \mid \Psi_u^\pi) \ \Big| \ \Phi \sim \Psi_u^\pi\Big]\Big] \\
= \ & \sum_{u \in \mathcal{N}} \mathbb{E}_\Phi\Big[Y_u(\Phi) \cdot \Delta(u \mid \psi_u^\pi(\Phi))\Big] = \mathbb{E}_\Phi\Big[\sum_{u \in \mathcal{N}(\pi, \Phi)} \Delta(u \mid \psi_u^\pi(\Phi))\Big].
\end{aligned} \tag{35}$$

Denote by $\psi(\pi_{\mathcal{A}}, \phi)$ the observed partial realization at the end of $\pi_{\mathcal{A}}$ under realization ϕ. Then, similar to the above analysis, we have

$$\begin{aligned} f_{\mathrm{avg}}(\pi_{\mathcal{A}}@\pi_{\mathrm{opt}}) &= \mathbb{E}_{\Phi,\pi_{\mathcal{A}}@\pi_{\mathrm{opt}}}[f(\mathcal{N}(\pi_{\mathcal{A}}@\pi_{\mathrm{opt}}, \Phi), \Phi)] \\ &= \mathbb{E}_{\pi_{\mathcal{A}}@\pi_{\mathrm{opt}}}\Big[\sum_{u\in\mathcal{N}(\pi_{\mathcal{A}},\Phi)} \Delta(u \mid \psi_u(\Phi)) + \sum_{u\in\mathcal{N}(\pi_{\mathrm{opt}},\Phi)\setminus\mathcal{N}(\pi_{\mathcal{A}},\Phi)} \Delta(u \mid \psi(\pi_{\mathcal{A}}, \Phi) \cup \psi_u^{\pi_{\mathrm{opt}}}(\Phi))\Big] \\ &= f_{\mathrm{avg}}(\pi_{\mathcal{A}}) + \mathbb{E}_{\pi_{\mathcal{A}}@\pi_{\mathrm{opt}}}\Big[\sum_{u\in\mathcal{N}(\pi_{\mathrm{opt}},\Phi)\setminus\mathcal{N}(\pi_{\mathcal{A}},\Phi)} \Delta(u \mid \psi(\pi_{\mathcal{A}}, \Phi) \cup \psi_u^{\pi_{\mathrm{opt}}}(\Phi))\Big] \\ &\le f_{\mathrm{avg}}(\pi_{\mathcal{A}}) + \mathbb{E}_{\pi_{\mathcal{A}}}\Big[\sum_{u\in\mathcal{N}(\pi_{\mathrm{opt}},\Phi)\setminus\mathcal{N}(\pi_{\mathcal{A}},\Phi)} \Delta(u \mid \psi_u(\Phi))\Big], \end{aligned}$$

where the inequality is due to adaptive submodularity and $\psi_u(\Phi) \subseteq \psi(\pi_{\mathcal{A}}, \Phi) \subseteq \psi(\pi_{\mathcal{A}}, \Phi) \cup \psi_u^{\pi_{\mathrm{opt}}}(\Phi)$. □

B.3. Proof of Lemma 7

Proof. Since $f_{\mathrm{avg}}(\pi_{\mathcal{A}}) = \mathbb{E}_{\pi_{\mathcal{A}}}\Big[\mathbb{E}_{\Phi}\Big[\sum_{u\in\mathcal{N}(\pi_{\mathcal{A}},\Phi)} \Delta(u \mid \psi_u(\Phi))\Big]\Big]$, it suffices to prove

$$\sum\nolimits_{u\in O_1(\phi)} \Delta(u \mid \psi_u(\phi)) \le k \cdot \sum\nolimits_{u\in\mathcal{N}(\pi_{\mathcal{A}},\phi)} \Delta(u \mid \psi_u(\phi)) \tag{36}$$

for any given realization $\phi \in Z^{\mathcal{N}}$ and fixed randomness of $\pi_{\mathcal{A}}$. Given a realization ϕ, let $\hat{u}_i$ be the i-th element selected by $\pi_{\mathcal{A}}$ and let $\hat{S}_i$ be the first i elements picked, i.e., $\hat{S}_i = \{\hat{u}_1, \ldots, \hat{u}_i\}$, for $i = 1, 2, \ldots, h$ where $h := |\mathcal{N}(\pi_{\mathcal{A}}, \phi)|$. Suppose that there exists a partition $O_{1,1}, O_{1,2}, \ldots, O_{1,h}$ of $O_1(\phi)$ such that for all $i = 1, 2, \ldots, h$,

$$\sum\nolimits_{u\in O_{1,i}} \Delta(u \mid \psi_u(\phi)) \le k \cdot \Delta(\hat{u}_i \mid \psi_{\hat{u}_i}(\phi)), \tag{37}$$

then Eqn. (36) must hold due to

$$\sum_{u\in O_1(\phi)} \Delta(u \mid \psi_u(\phi)) = \sum_{i=1}^{h} \sum_{u\in O_{1,i}} \Delta(u \mid \psi_u(\phi)) \le k \cdot \sum_{i=1}^{h} \Delta(\hat{u}_i \mid \psi_{\hat{u}_i}(\phi)) = k \cdot \sum_{u\in\mathcal{N}(\pi_{\mathcal{A}},\phi)} \Delta(u \mid \psi_u(\phi)). \tag{38}$$

Therefore, we just need to show the existence of such a desired partition of O_1, as proved below.

We use the following iterative algorithm to find the partition, which is inspired by (Calinescu et al., 2011). Define $\mathcal{N}_h := O_1(\phi)$. For $i = h, h-1, \ldots, 2$, let $B_i := \{u \in \mathcal{N}_i \mid \hat{S}_{i-1} \cup \{u\} \in \mathcal{I}\}$. If $|B_i| \le k$, set $O_{1,i} = B_i$. Otherwise, pick an arbitrary $O_{1,i} \subseteq B_i$ with $|O_{1,i}| = k$. Then, set $\mathcal{N}_{i-1} = \mathcal{N}_i \setminus O_{1,i}$. Finally, set $O_{1,1} = \mathcal{N}_1$. Clearly, $|O_{1,i}| \le k$ for $i = 2, \ldots, h$. We further show that $|O_{1,1}| \le k$. We prove it by contradiction and assume $|O_{1,1}| > k$. If $|B_2| \le k$, then we have $\hat{S}_1 \cup \{u\} \notin \mathcal{I}$ for every $u \in \mathcal{N}_1$ according to the above process. So $\hat{S}_1$ is a base of $\hat{S}_1 \cup \mathcal{N}_1$, which implies that $|\mathcal{N}_1| \le k \cdot |\hat{S}_1|$, contradicting the assumption that $|\mathcal{N}_1| = |O_{1,1}| > k$. Consequently, it must hold that $|B_2| > k$ and hence $|O_{1,2}| = k$ and $|\mathcal{N}_2| > 2k$. Using a similar argument, we can recursively get that $|B_i| > k$ and hence $|O_{1,i}| = k$ and $|\mathcal{N}_i| > ik$ for any $i = 3, \ldots, h$, e.g., $|\mathcal{N}_h| > hk$. However, as $\hat{S}_h$ is a base of $\hat{S}_h \cup O_1(\phi)$, we should have $|\mathcal{N}_h| = |O_1(\phi)| \le hk$, which shows a contradiction. Therefore, we can conclude that $|O_{1,i}| \le k$ for all $i = 1, 2, \ldots, h$.

According to the partition $O_{1,i} : i \in [h]$ constructed above, it is obvious that for every $u \in O_{1,i}$, $\hat{S}_{i-1} \cup \{u\} \in \mathcal{I}$. This implies that for every $u \in O_{1,i}$, u cannot be considered before $\hat{u}_i$ is added by $\pi_{\mathcal{A}}$, i.e., $\psi_{\hat{u}_i}(\phi) \subseteq \psi_u(\phi)$. Meanwhile, due to the greedy rule of AdaptRandomGreedy, it follows that $\Delta(\hat{u}_i \mid \psi_{\hat{u}_i}(\phi)) \ge \Delta(u \mid \psi_{\hat{u}_i}(\phi))$ for each $u \in O_{1,i}$. Hence,

$$\sum_{u\in O_{1,i}} \Delta(u \mid \psi_u(\phi)) \le \sum_{u\in O_{1,i}} \Delta(u \mid \psi_{\hat{u}_i}(\phi)) \le \sum_{u\in O_{1,i}} \Delta(\hat{u}_i \mid \psi_{\hat{u}_i}(\phi)) \le k \cdot \Delta(\hat{u}_i \mid \psi_{\hat{u}_i}(\phi)) \tag{39}$$

holds for any $i \in [h]$. Combining the above results completes the proof. □

B.4. Proof of Lemma 8

Proof. Again, since $f_{\mathrm{avg}}(\pi_{\mathcal{A}}) = \mathbb{E}_{\pi_{\mathcal{A}}}\Big[\mathbb{E}_{\Phi}\Big[\sum_{u\in\mathcal{N}(\pi_{\mathcal{A}},\Phi)} \Delta(u \mid \psi_u(\Phi))\Big]\Big]$, we only need to prove that, for any $\phi \in Z^{\mathcal{N}}$,

$$\mathbb{E}_{\pi_{\mathcal{A}}}\Big[\sum\nolimits_{u\in O_2(\phi)} \Delta(u \mid \psi_u(\phi))\Big] \le \frac{1-p}{p} \cdot \mathbb{E}_{\pi_{\mathcal{A}}}\Big[\sum\nolimits_{u\in\mathcal{N}(\pi_{\mathcal{A}},\phi)} \Delta(u \mid \psi_u(\phi))\Big]. \tag{40}$$

Given a realization $\phi \in Z^{\mathcal{N}}$, for each $u \in \mathcal{N}$, let X_u be a random variable such that $X_u = 1$ if $u \in O_2(\phi)$ and $X_u = 0$ otherwise. So we have

$$\sum_{u \in O_2(\phi)} \Delta(u \mid \psi_u(\phi)) = \sum_{u \in \mathcal{N}} \big(X_u \cdot \Delta(u \mid \psi_u(\phi))\big). \tag{41}$$

Similarly, for each $u \in \mathcal{N}$, let Y_u be a random variable such that $Y_u = 1$ if $u \in \mathcal{N}(\pi_{\mathcal{A}}, \phi)$ and $Y_u = 0$ otherwise. Thus,

$$\sum_{u \in \mathcal{N}(\pi_{\mathcal{A}}, \phi)} \Delta(u \mid \psi_u(\phi)) = \sum_{u \in \mathcal{N}} \big(Y_u \cdot \Delta(u \mid \psi_u(\phi))\big). \tag{42}$$

Therefore, it is sufficient to prove:

$$\forall u \in \mathcal{N}: \ \mathbb{E}_{\pi_{\mathcal{A}}}\big[X_u \cdot \Delta(u \mid \psi_u(\phi))\big] \leq \frac{1-p}{p} \cdot \mathbb{E}_{\pi_{\mathcal{A}}}\big[Y_u \cdot \Delta(u \mid \psi_u(\phi))\big] \tag{43}$$

Observe that, for any given $u \in \mathcal{N}$, if $\Delta(u \mid \psi_u(\phi)) \leq 0$ or $\mathrm{dom}(\psi_u(\phi)) \cup \{u\} \notin \mathcal{I}$, then we have $u \notin \mathcal{N}(\pi_{\mathcal{A}}, \phi)$ and $u \notin O_2(\phi)$ by definition, which indicates $X_u = Y_u = 0$. Consider the event that $\Delta(u \mid \psi_u(\phi)) > 0$ and $\mathrm{dom}(\psi_u(\phi)) \cup \{u\} \in \mathcal{I}$, and denote such an event as $\mathcal{E}_u$. Since $\Pr[u \in \mathcal{N}(\pi_{\mathcal{A}}, \phi) \mid \mathcal{E}_u] = p$, it is trivial to see that

$$\mathbb{E}_{\pi_{\mathcal{A}}}\big[Y_u \cdot \Delta(u \mid \psi_u(\phi))\big] = p \cdot \mathbb{E}_{\psi_u(\phi)}[\Delta(u \mid \psi_u(\phi)) \mid \mathcal{E}_u] \cdot \Pr[\mathcal{E}_u], \tag{44}$$

where the expectation is taken over the randomness of $\psi_u(\phi)$ (i.e., $\psi_u(\phi) \sim \mathcal{E}_u$) due to the internal randomness of algorithm. On the other hand, if $u \in O(\phi)$, then we have $\Pr[u \in O_2(\phi) \mid \mathcal{E}_u] = 1 - p$ as u is discarded with probability of $1 - p$, while we also have $\Pr[u \in O_2(\phi) \mid \mathcal{E}_u] = 0$ if $u \notin O(\phi)$. Thus, we know $\Pr[u \in O_2(\phi) \mid \mathcal{E}_u] \leq (1 - p)$ and hence we can immediately get

$$\mathbb{E}_{\pi_{\mathcal{A}}}\big[X_u \cdot \Delta(u \mid \psi_u)\big] \leq (1-p) \cdot \mathbb{E}_{\psi_u(\phi)}[\Delta(u \mid \psi_u(\phi)) \mid \mathcal{E}_u] \cdot \Pr[\mathcal{E}_u]. \tag{45}$$

The lemma then follows by combining all the above reasoning. □

B.5. Proof of Theorem 3

Proof. According to Lemmas 6–8, we have

$$\begin{aligned}
f_{\mathrm{avg}}(\pi_{\mathcal{A}} @ \pi_{\mathrm{opt}}) - f_{\mathrm{avg}}(\pi_{\mathcal{A}}) &\leq \mathbb{E}_{\pi_{\mathcal{A}}, \Phi}\Big[\sum_{u \in \mathcal{N}(\pi_{\mathrm{opt}}, \Phi) \setminus \mathcal{N}(\pi_{\mathcal{A}}, \Phi)} \Delta(u \mid \psi_u(\Phi))\Big] \\
&\leq \mathbb{E}_{\pi_{\mathcal{A}}, \Phi}\Big[\sum_{u \in O_1(\Phi)} \Delta(u \mid \psi_u(\Phi)) + \sum_{u \in O_2(\Phi)} \Delta(u \mid \psi_u(\Phi))\Big] \\
&\leq \Big(k + \frac{1-p}{p}\Big) \cdot f_{\mathrm{avg}}(\pi_{\mathcal{A}})
\end{aligned}$$

where the second inequality is due to the definition of $O_3(\Phi)$, i.e., $\Delta(u \mid \psi_u(\Phi)) \leq 0$ for every $u \in O_3(\Phi)$. Combining the above result with Lemma 5 gives

$$f(\pi_{\mathrm{opt}}) \leq \frac{kp + 1}{p(1-p)} \cdot f_{\mathrm{avg}}(\pi_{\mathcal{A}}). \tag{46}$$

Moreover, $\frac{kp+1}{p(1-p)}$ achieves its minimum value of $(1+\sqrt{k+1})^2$ at $p = (1+\sqrt{k+1})^{-1}$. Finally, the $\mathcal{O}(nr)$ time complexity is evident, as the algorithm incurs $\mathcal{O}(n)$ oracle queries for each selected element. □