Proof Notes

The objective submodular function f is non-negative, non-monotone and normalized with $f(\emptyset) = 0$ by default.

1 Deterministic + Multi-Solution + k-System

Algorithm 1: DeterministicMultiSolutionKSystem

Input: the number of candidates $\ell \in \mathbb{Z}_{\geq 2}$, independence system $(\mathcal{N}, \mathcal{I})$

Output: A feasible set S^*

foreach $i \in [\ell]$ **do** $S_i \leftarrow \emptyset$;

while true do

  foreach $i \in [\ell]$ **do** $\mathcal{M}_i \leftarrow \{e \in \mathcal{N} \backslash \big(\bigcup_{j \in [\ell]} S_j\big) : S_i \cup \{e\} \in \mathcal{I}\}$;

  $(j, u) \leftarrow \text{argmax}_{i \in [\ell], e \in \mathcal{M}_i} f(e|S_i)$; // *u could be emptyset.*

  if $f(u|S_j) \leq 0$ **then break**;

  $S_j \leftarrow S_j \cup \{u\}$;

$S^* \leftarrow \text{argmax}_{X \in \{S_1, \ldots, S_\ell\}} f(X)$;

return S^*

Definition. Let $S = \cup_{t=1}^{\ell} S_t$. We can write S as $\{v_1, v_2, \ldots, v_{|S|}\}$, such that v_p is added into S by the algorithm before v_q for any $1 \leq p < q \leq |S|$. With this ordered list, given any $e = v_p \in S$, we define

$$\text{Pre}(e, S_i) = \{v_1, \ldots, v_{p-1}\} \cap S_i, \quad \forall i \in [\ell]$$

That is, $\text{Pre}(e, S_i)$ denotes the set of elements in S_i that are added by Algorithm 1 before adding e to S. Furthermore, we define

$$O_{ij}^{+} = \{e \in O \cap S_i : \text{Pre}(e, S_j) \cup \{e\} \in \mathcal{I}\},$$
$$O_{ij}^{-} = \{e \in O \cap S_i : \text{Pre}(e, S_j) \cup \{e\} \notin \mathcal{I}\},$$
$$Q_i = \{e \in O \backslash S : S_i \cup \{e\} \notin \mathcal{I}\}$$

Note that all the three kinds of sets defined above are subsets of optimal set O. We also define the marginal gain of any $e \in S$ as

$$\delta(e) = \sum_{i=1}^{\ell} f(e|\text{Pre}(e, S_i)) \mathbb{I}_{S_i}(e)$$

Lemma 1.1. For any $i \in [\ell]$, there exists a function $\pi_i : R_i \to S_i$ s.t.

1. $\forall e \in R_i$ we have $\mathrm{Pre}(\pi_i(e), S_i) \cup \{e\} \in \mathcal{I}$.
2. $\forall e \in O \cap S_i$ we have $\pi_i(e) = e$.
3. $|\pi_i^{-1}(y)| \le k$ for any $y \in S_i$.

where

$$R_i = (O \cap S_i) \cup \Big(\bigcup_{j \in [\ell] \backslash \{i\}} O^-_{ji} \Big) \cup Q_i$$

Proof. We just prove the case $i = 1$. Suppose that the elements in S_1 are $\{u_1, \dots, u_s\}$ (listed according to the order that they added into S_1). Let $L_s = R_1$. We execute the following iterations from $j = s$ to $j = 0$. At the beginning of the j-th iteration, we compute a set $A_j = \{x \in L_j \backslash \{u_1, \dots, u_{j-1}\} : \{u_1, \dots, u_{j-1}, x\} \in \mathcal{I}\}$. If $|A_j| \le k$, then we set $D_j = A_j$. If $|A_j| > k$ and $u_j \in O \cap S_1$ (so $u_j \in A_j$), then we pick a subset $D_j \subseteq A_j$ satisfying $|D_j| = k$ and $u_j \in D_j$. If $|A_j| > k$ and $u_j \notin O \cap S_1$, then we pick a subset $D_j \subseteq A_j$ satisfying $|D_j| = k$. After that, we set $\pi_1(e) = u_j$ for every $e \in D_j$ and let $L_{j-1} = L_j \backslash D_j$, then enter the $(j-1)$-th itertation.

From the above process, it can be easily seen that conditions 1-3 in the lemma are satisfied. So we only need to prove that each $e \in R_1$ is mapped to an element in S_1, which is equivalent to prove $L_0 = \emptyset$ as each $e \in L_s \backslash L_0$ is mapped to an element in S_1 according to the above process. In the following we will prove $L_j \le kj$ for $0 \le j \le s$ by induction.

When $j = s$, consider the set $M = S_1 \cup \Big(\bigcup_{j \in [\ell] \backslash \{1\}} O^-_{j1} \Big) \cup Q_1$. Cleary, each element $e \in Q_1$ satisfies $S_1 \cup \{e\} \notin \mathcal{I}$ according to the definition of Q_1. Besides, we must have $S_1 \cup \{e\} \notin \mathcal{I}$ for each $e \in \bigcup_{j \in [\ell] \backslash \{1\}} O^-_{j1}$, because otherwise there exists $e \in \bigcup_{j \in [\ell] \backslash \{1\}} O^-_{j1}$ satisfying $\mathrm{Pre}(e, S_1) \cup \{e\} \in \mathcal{I}$, contradicting the definition of $\bigcup_{j \in [\ell] \backslash \{1\}} O^-_{j1}$. Therefore, we know that S_1 is a base of M. As $R_1 \in \mathcal{I}$ and $R_1 \subseteq M$, we get $|L_s| = |R_1| \le p|S_1| = ps$ according to the definition of p-set system.

Now suppose that $|L_j| \le kj$ for certain $j \le s$. If $|A_j| > k$, then we have $|D_j| = k$ and hence $|L_{j-1}| = |L_j| - k \le k(j-1)$. If $|A_j| \le k$, then we know that there does not exist $e \in L_{j-1} \backslash \{u_1, \dots, u_{j-1}\}$ s.t. $\{u_1, \dots, u_{j-1}\} \cup \{e\} \in \mathcal{I}$ due to the process for constructing π_1. Now consider the set $M' = \{u_1, \dots, u_{j-1}\} \cup L_{j-1}$, we know that $\{u_1, \dots, u_{j-1}\}$ is a base of M' and $L_{j-1} \in \mathcal{I}$, which implies $|L_{j-1}| \le k(j-1)$ according to the definition of k-set system.

The above reason proves $L_j \le pj$ for all $0 \le j \le s$ by induction, so we get $L_0 = \emptyset$ and hence lemma follows. □

Lemma 1.2. For any $i, j \in [\ell]$, $i \ne j$, the Algorithm 1 satisfies:

$$f(O^+_{ij} | S_j) \le \sum_{e \in O^+_{ij}} \delta(\pi_i(e)), \quad f(O^-_{ij} | S_j) \le \sum_{e \in O^-_{ij}} \delta(\pi_j(e)), \quad f(Q_i | S_i) \le \sum_{e \in Q_i} \delta(\pi_i(e))$$

Proof. Obviously we have

$$f(O^+_{ij} | S_j) \le \sum_{e \in O^+_{ij} \backslash S_j} f(e | S_j) \overset{\text{(a)}}{\le} \sum_{e \in O^+_{ij}} f(e | \mathrm{Pre}(e, S_j)) \overset{\text{(b)}}{\le} \sum_{e \in O^+_{ij}} \delta(\pi_i(e))$$

where (a) is because of the submodularity and (b) is the greedy choice. For $e \in O^-_{ij}$, we claim that e has not been inserted into S_i yet when SimultaneousGreedy inserts $\pi_j(e)$ into S_j. If not, we have $\mathrm{Pre}(e, S_j) \subseteq \mathrm{Pre}(\pi_j(e), S_j)$. By hereditary property of independence system, this means

$\text{Pre}(e, S_j) \cup \{e\} \in \mathcal{I}$, which contradicts the definition of O^-_{ij}. Similar as above, we can get

$$f(O^-_{ij}|S_j) \leq \sum_{e \in O^-_{ij} \setminus S_j} f(e|S_j) \leq \sum_{e \in O^-_{ij}} f(e|\text{Pre}(\pi_j(e), S_j)) \leq \sum_{e \in O^-_{ij}} \delta(\pi_j(e))$$

For the third inequality, we do the same thing

$$f(Q_i|S_i) \leq \sum_{e \in Q_i \setminus S_i} f(e|S_i) \leq \sum_{e \in Q_i} f(e|\text{Pre}(\pi_i(e), S_i)) \leq \sum_{e \in Q_i} \delta(\pi_i(e))$$

which completes our proof. $\square$

Proposition 1. The Algorithm 1 returns a solution S^* with $\frac{\ell-1}{\ell(k+\ell-1)}$ approximation ratio.

Proof. Applying our definition and Lemma 1.2, by submodularity and the non-negative greedy choice, we have

$$\begin{aligned} f(O \cup S_i) - f(S_i) &\leq \sum_{j \in [\ell] \setminus \{i\}} f(O \cap S_j|S_i) + f(Q_i|S_i) + f(O \setminus (S \cup Q_i)|S_i) \\ &\leq \sum_{j \in [\ell] \setminus \{i\}} \Big(f(O^+_{ji}|S_i) + f(O^-_{ji}|S_i) \Big) + f(Q_i|S_i) \\ &\leq \sum_{j \in [\ell] \setminus \{i\}} \left(\sum_{e \in O^+_{ji}} \delta(\pi_j(e)) + \sum_{e \in O^-_{ji}} \delta(\pi_i(e)) \right) + \sum_{e \in Q_i} \delta(\pi_i(e)) \end{aligned}$$

Now we add the ℓ inequalities together

$$\begin{aligned} &\sum_{i \in [\ell]} \Big(f(O \cup S_i) - f(S_i) \Big) \\ \leq &\sum_{i \in [\ell]} \left(\sum_{j \in [\ell] \setminus \{i\}} \left(\sum_{e \in O^+_{ji}} \delta(\pi_j(e)) + \sum_{e \in O^-_{ji}} \delta(\pi_i(e)) \right) + \sum_{e \in Q_i} \delta(\pi_i(e)) \right) \\ \overset{\text{(c)}}{=} &\sum_{i \in [\ell]} \left\{ \left(\sum_{e \in O^+_{ij_i}} \delta(\pi_i(e)) + \sum_{j \in [\ell] \setminus \{i\}} \sum_{e \in O^-_{ji}} \delta(\pi_i(e)) + \sum_{e \in Q_i} \delta(\pi_i(e)) \right) + \sum_{j \in [\ell] \setminus \{i, j_i\}} \sum_{e \in O^+_{ij}} \delta(\pi_i(e)) \right\} \\ \leq &\sum_{i \in [\ell]} \left(\sum_{e \in R_i} \delta(\pi_i(e)) + \sum_{j \in [\ell] \setminus \{i, j_i\}} \sum_{e \in S_i} \delta(\pi_i(e)) \right) \\ \leq &\sum_{i \in [\ell]} \left(\sum_{y \in S_i} |\pi_i^{-1}(y)| \delta(y) + \sum_{j \in [\ell] \setminus \{i, j_i\}} \sum_{e \in S_i} \delta(e) \right) \\ \leq &(k + \ell - 2) \sum_{i \in [\ell]} f(S_i) \end{aligned}$$

where for each $i \in [\ell]$, j_i is an arbitratry fixed number chosen from $[\ell] \setminus \{i\}$. And the equality (c) is using the order change of the double summation as following

$$\begin{aligned} \sum_{i \in [\ell]} \sum_{j \in [\ell] \setminus \{i\}} \sum_{e \in O^+_{ji}} \delta(\pi_j(e)) &= \sum_{j \in [\ell]} \sum_{i \in [\ell] \setminus \{j\}} \sum_{e \in O^+_{ji}} \delta(\pi_j(e)) \\ &= \sum_{i \in [\ell]} \sum_{j \in [\ell] \setminus \{i\}} \sum_{e \in O^+_{ij}} \delta(\pi_i(e)) \\ &= \sum_{i \in [\ell]} \left(\sum_{e \in O^+_{ij_i}} \delta(\pi_i(e)) + \sum_{j \in [\ell] \setminus \{i, j_i\}} \sum_{e \in O^+_{ij}} \delta(\pi_i(e)) \right) \end{aligned}$$

For the other side, by induction we have

$$\sum_{i\in[\ell]} f(O\cup S_i) \geq (\ell-1)f(O)$$

Finally we get

$$(\ell-1)f(O) \leq \sum_{i\in[\ell]} f(O\cup S_i) \leq (k+\ell-1)\sum_{i\in[\ell]} f(S_i) \leq \ell(k+\ell-1)f(S^*)$$

which means the approximation ratio is $\frac{\ell-1}{\ell(k+\ell-1)}$. □

We now turn to consider the best choice for ℓ under the fixed k case. Assume that the approximation ratio is $h(\ell) = \frac{\ell(k+\ell-1)}{\ell-1}$, then let

$$\frac{\mathrm{d}h}{\mathrm{d}\ell} = \frac{\mathrm{d}}{\mathrm{d}\ell}\Big(\ell+k+\frac{k}{\ell-1}\Big) = 1-\frac{k}{(\ell-1)^2} = 0 \Rightarrow \ell_0 = 1+\sqrt{k}$$

So we have $h_{\min} = h(\ell_0) = (1+\sqrt{k})^2$. Since $\sqrt{k}$ may take non-integer value, we could choose $\ell_1 = 1+\lceil\sqrt{k}\rceil$. And then

$$h(\ell_1) = \frac{(1+\lceil\sqrt{k}\rceil)(k+\lceil\sqrt{k}\rceil)}{\lceil\sqrt{k}\rceil} = k+\lceil\sqrt{k}\rceil+\frac{k}{\lceil\sqrt{k}\rceil}+1 \leq k+2\sqrt{k}+2$$

2 Randomized + One-Solution + k-System

Algorithm 2: RandomizedOneSolutionKSystem

Input: probability $p\in(0,1)$, independence system $(\mathcal{N},\mathcal{I})$

Output: A feasible set S

```
1 S ← ∅;
2 while true do
3     M = {e ∈ N : S ∪ {e} ∈ I};
4     u ← argmax_{e∈M} f(e|S); // u could be emptyset.
5     if f(u|S) ≤ 0 then break;
6     S ← S ∪ {u} with probability p;
7     N = N\{u};
8 return S
```

Definition. $U = \{u_1,\ldots,u_\ell\}$ is the sequence according to the order of them being selected as argmax in algorithm except the last turn, and $S_1,\ldots,S_\ell$ is the sequence of results attained by the algorithm after these turns. Let

$$O_{\leq\ell} := \{e\in O : \tau(e)\leq\ell\}$$
$$O_{>\ell} := \{e\in O : \tau(e)>\ell \wedge f(e|S_\ell)>0\}$$

where $\tau(u_i)=i$, $\tau(e)=\infty, \forall e\notin U$. Let $S_u = S_{\tau(u)-1}$ for $u\in U$.

Sketch of Derivation. We first prove a result for the $O_{>\ell}$ part. For any $u \in O_{>\ell}$, define $\pi : O_{>\ell} \to \{1, \dots, \ell\}$ as

$$\pi(u) := \mathrm{argmax}_i \{S_{i-1} \cup \{u\} \in \mathcal{I}\}$$

So $\pi^{-1}(i) \neq \emptyset$ means $u_i \in S_\ell$. Now we assume $S_\ell = \{v_1, \dots, v_s\}$ according to the order of adding elements. Since for any $i \in [\ell]$, we have $\{v_1, \dots, v_i\}$ is the base of $\{v_1, \dots, v_i\} \cup \bigcup_{j\in[i]} \pi^{-1}(\tau(v_j))$, we have

$$\sum_{j\in[i]} |\pi^{-1}(\tau(v_j))| \leq ik$$

So let $k_i = \frac{1}{i}\sum_{j\in[i]} |\pi^{-1}(\tau(v_j))| \leq k$, we have

$$\begin{aligned}
\sum_{u\in O_{>\ell}} f(u|S_\ell) \leq & \sum_{i=1}^{s} |\pi^{-1}(\tau(v_i))| f(v_i|S_{v_i}) \\
= & \sum_{i=1}^{s} \big(ik_i - (i-1)k_{i-1}\big) f(v_i|S_{v_i}) \\
= & sk_s f(v_s|S_{v_s}) + \sum_{i=1}^{s-1} i\Big(f(v_i|S_{v_i}) - f(v_{i+1}|S_{v_{i+1}})\Big)k_i \\
\overset{\text{(d)}}{\leq} & skf(v_s|S_{v_s}) + k\sum_{i=1}^{s-1} i\Big(f(v_i|S_{v_i}) - f(v_{i+1}|S_{v_{i+1}})\Big) \\
= & k\sum_{u\in S_\ell} f(u|S_u)
\end{aligned}$$

where (d) is because of the non-negativeness of $f(v_i|S_{v_i}) - f(v_{i+1}|S_{v_{i+1}})$.

Now we turn to consider the other part. For $\forall u \in U$ define r.v. as

$$R_u = \begin{cases} f(u|S_u), & \text{if } u \in S_\ell \\ 0, & \text{else} \end{cases}, \quad X_u = \begin{cases} 1, & \text{if } u \in O_{\leq\ell}\backslash S_\ell \\ 0, & \text{else} \end{cases}$$

We will prove a claim that, for $\forall u \in U$ we have

$$\mathbb{E}[X_u f(u|S_u)] \leq \frac{1-p}{p}\mathbb{E}[R_u]$$

It's easy to see $\mathbb{E}[R_u] = pf(u|S_u)$. Note that if $u \in O$ and is not discarded by algorithm, then $u \notin O_{\leq\ell}\backslash S_\ell$ hence $X_u = 0$. And if it is discarded, then $X_u = 1$ and $\mathbb{E}[X_u f(u|S_u)] = (1-p)f(u|S_u)$. And for $u \notin O$ case we have $\mathbb{E}[X_u f(u|S_u)] = 0$.

Combine things together, we have

$$\begin{aligned}
\mathbb{E}\big[f(S_\ell \cup O) - f(S_\ell)\big] \leq & \mathbb{E}\Big[\sum_{e\in O_{\leq\ell}\backslash S_\ell} f(e|S_\ell) + \sum_{e\in O_{>\ell}} f(e|S_\ell)\Big] \\
\leq & \mathbb{E}\Big[\sum_{u\in U} X_u f(u|S_u) + k\sum_{u\in S_\ell} f(u|S_u)\Big] \\
\leq & \frac{1-p}{p}\mathbb{E}\Big[\sum_{u\in U} R_u\Big] + k\mathbb{E}\Big[\sum_{u\in S_\ell} f(u|S_u)\Big] \\
= & \Big(k - 1 + \frac{1}{p}\Big)\mathbb{E}[f(S_\ell)]
\end{aligned}$$

So we get

Proposition 2. Algorithm 2 will output a set S with approximation ratio $\frac{1-p}{k+\frac{1}{p}}$ in expectation.

By taking $p = \frac{1}{1+\sqrt{1+k}}$, we could achieve the best approximation ratio $(1+\sqrt{1+k})^2$.

3 Randomized + Multi-Solution + k-System

Algorithm 3: RandomizedMultiSolutionKSystem

Input: probability $p \in (0,1)$, the number of candidates ℓ, independence system $(\mathcal{N}, \mathcal{I})$

Output: A feasible set S^*

foreach $i \in [\ell]$ **do** $S_i \leftarrow \emptyset$;

while true do

  foreach $i \in [\ell]$ **do** $\mathcal{M}_i \leftarrow \{e \in \mathcal{N} \backslash \left(\bigcup_{j\in[\ell]} S_j\right) : S_i \cup \{e\} \in \mathcal{I}\}$;

  $(j, u) \leftarrow \mathrm{argmax}_{i\in[\ell], e\in\mathcal{M}_i} f(e|S_i)$; // *$u$ could be emptyset.*

  if $f(u|S_j) \leq 0$ **then break**;

  $S_j \leftarrow S_j \cup \{u\}$ with probability p;

  $\mathcal{N} = \mathcal{N} \backslash \{u\}$;

$S^* \leftarrow \mathrm{argmax}_{X\in\{S_1,\ldots,S_\ell\}} f(X)$;

return S^*

To extend the randomized algorithm using multi-solution method, the derivation way is just similar as we did in deterministic case (Algorithm 1). Every set S_i holds an ordered argmax list U_i. Let $U = \cup_{i\in[\ell]} U_i = \{u_1, \ldots, u_{|U|}\}$ according to the order, by seperating the elements discarded or not discarded by the algorithm we could define

$$
\begin{aligned}
&\mathrm{Pre}(u_j, S_i) = \{u_1, \ldots, u_{j-1}\} \cap S_i, \ \forall i \in [\ell] \\
&O_{ij}^+ = \{e \in O \cap S_i : \mathrm{Pre}(e, S_j) \cup \{e\} \in \mathcal{I}\}, \\
&O_{ij}^- = \{e \in O \cap S_i : \mathrm{Pre}(e, S_j) \cup \{e\} \notin \mathcal{I}\}, \\
&T_{ij}^+ = \{e \in O \cap (U_i \backslash S_i) : \mathrm{Pre}(e, S_j) \cup \{e\} \in \mathcal{I}\}, \\
&T_{ij}^- = \{e \in O \cap (U_i \backslash S_i) : \mathrm{Pre}(e, S_j) \cup \{e\} \notin \mathcal{I}\}, \\
&Q_i = \{e \in O \backslash U : S_i \cup \{e\} \notin \mathcal{I}\}
\end{aligned}
$$

Then

Lemma 3.1. For any $i \in [\ell]$, there exists a function $\pi_i : R_i \to S_i$ s.t.

1. $\forall e \in R_i$ we have $\mathrm{Pre}(\pi_i(e), S_i) \cup \{e\} \in \mathcal{I}$.
2. $\forall e \in O \cap S_i$ we have $\pi_i(e) = e$.
3. $|\pi_i^{-1}(y)| \leq k$ for any $y \in S_i$.

where

$$R_i = (O \cap S_i) \cup \Big(\bigcup_{j\in[\ell]\backslash\{i\}} O_{ji}^-\Big) \cup \Big(\bigcup_{j\in[\ell]\backslash\{i\}} T_{ji}^-\Big) \cup Q_i$$

Lemma 3.2. For any $i, j \in [\ell]$, $i \neq j$, the Algorithm 3 satisfies:

$$f(O_{ij}^{+}|S_j) \leq \sum_{e \in O_{ij}^{+}} \delta(\pi_i(e)), \quad f(O_{ij}^{-}|S_j) \leq \sum_{e \in O_{ij}^{-}} \delta(\pi_j(e)),$$
$$f(T_{ij}^{-}|S_j) \leq \sum_{e \in T_{ij}^{-}} \delta(\pi_j(e)), \quad f(Q_i|S_i) \leq \sum_{e \in Q_i} \delta(\pi_i(e))$$

Now we consider

$$\begin{aligned}
f(O \cup S_i) - f(S_i) &\leq \sum_{j \in [\ell] \setminus \{i\}} f(O \cap S_j|S_i) + \sum_{j \in [\ell]} f(O \cap (U_j \setminus S_j)|S_i) + f(Q_i|S_i) + f(O \setminus (S \cup U \cup Q_i)|S_i) \\
&\leq \sum_{j \in [\ell] \setminus \{i\}} \Big(f(O_{ji}^{+}|S_i) + f(O_{ji}^{-}|S_i) \Big) + \sum_{j \in [\ell]} \Big(f(T_{ji}^{+}|S_i) + f(T_{ji}^{-}|S_i) \Big) + f(Q_i|S_i) \\
&\leq \sum_{j \in [\ell] \setminus \{i\}} \Bigg(\sum_{e \in O_{ji}^{+}} \delta(\pi_j(e)) + \sum_{e \in O_{ji}^{-}} \delta(\pi_i(e)) + \sum_{e \in T_{ji}^{-}} \delta(\pi_i(e)) \Bigg) \\
&\quad + \sum_{j \in [\ell]} f(T_{ji}^{+}|S_i) + \sum_{e \in Q_i} \delta(\pi_i(e))
\end{aligned}$$

So

$$\sum_{i \in [\ell]} \Big(f(O \cup S_i) - f(S_i) \Big) \leq \sum_{i,j \in [\ell]} f(T_{ji}^{+}|S_i) + \sum_{i \in [\ell]} \Bigg(\sum_{e \in R_i} \delta(\pi_i(e)) + \sum_{j \in [\ell] \setminus \{i, j_i\}} \sum_{e \in O_{ij}^{+}} \delta(\pi_i(e)) \Bigg)$$

The second part is what we have met in Algorithm 1, the first part could be inducted according to the way in Algorithm 2. So we have the similar inequality

$$\mathbb{E}[X_{i,u} f(u|(S_i)_u)] \leq \frac{1-p}{p} \mathbb{E}[R_{i,u}], \ \forall e \in U_i, \ i \in [\ell]$$

where

$$R_{i,u} = \begin{cases} f(u|S_u), & \text{if } u \in S_j \\ 0, & \text{else} \end{cases}, \quad X_{i,u} = \begin{cases} 1, & \text{if } u \in O \cap (U_j \setminus S_j) \\ 0, & \text{else} \end{cases}$$

Finally we could get

$$(\ell - p) f(O) \leq \mathbb{E}\Big[\sum_{i \in [\ell]} (f(S_i \cup O)) \Big] \leq \ell(k + \ell - 1) f(S^*) + \frac{1-p}{p} \ell^2 f(S^*)$$

Proposition 3: The Algorithm 3 returns a solution S^* with approximation ratio $\frac{\ell - p}{\ell(k - 1 + \frac{\ell}{p})}$.

Since the constrained optimization problem

$$\begin{aligned}
&\textbf{Maximize} \quad \frac{\ell - p}{\ell(k - 1 + \frac{\ell}{p})} \text{ where } k \in \mathbb{Z}_{\geq 1} \text{ is a const} \\
&\textbf{Subject to} \quad p \in (0, 1), \ell \in \mathbb{Z}_{\geq 2},
\end{aligned}$$

is decided by ℓ/p. For simplicity, we choose the smallest $\ell = 2$ and then $p = \frac{2}{1+\sqrt{k}}$ we could have the best ratio $(1 + \sqrt{k})^2$.

4 Randomized + Two-Solution + k-System + Multi-Knapsack

Now we incorporate knapsack constraints with k-system under our randomized multi-solution framework. In section 3 we have found out that we only need to take $\ell = 2$. So this section we will just study the two-solution case.

4.1 FANTOM Framework (Discarded)

The conditional expectation differs from the two termination condition. We can not follow the same way as FANTOM did.

4.2 A New Proposed Way (Discarded)

See Algorithm 4.

In k-system, the greedy strategy needs us to use marginal gain but the threshold needs the density expression.

Algorithm 4: RandomizedTwinGreedywithKnapsacks

```
Input: ϵ, p ∈ (0,1), γ_max > γ_min > 0, independence system (N, I)
Output: A feasible set S*
S_1, S_2 ← ∅, γ ← γ_max, t ← 1, flag ← true, u* = argmax_{e∈N ∧ {e}∈I} f(e)
while (γ ≥ γ_min) ∧ flag do
    A_1 ← {e ∈ N\(S_1 ∪ S_2) : S_1 + e ∈ I};
    A_2 ← {e ∈ N\(S_1 ∪ S_2) : S_2 + e ∈ I};
    if A_1 ∪ A_2 = ∅ then break;
    repeat
        (j, u) ← argmax_{(i,e)∈{1,2}×A_i} f(e|S_i) / Σ_{r=1}^{h} c_r(e);
        if f(u|S_j) ≤ 0 then
            N = N\(A_1 ∪ A_2);
            break;
        else if f(u|S_j) ≥ γ then
            if ⋀_{r∈[h]} c_r(S_j + u) ≤ 1 then
                u_t ← u, t ← t + 1;
                S_j ← S_j + u_t with prob. p;
                N ← N − u_t;
                A_1 ← {e ∈ A_1\{u_t} : S_1 + e ∈ I};
                A_2 ← {e ∈ A_2\{u_t} : S_2 + e ∈ I};
            else
                flag ← false;
                break;
        else
            A_j ← A_j − u;
    until A_1 ∪ A_2 = ∅;
    γ ← (1 − ϵ)γ
return S* ← argmax_{S∈{S_1, S_2, {u*}}} f(S)
```

5 Randomized + Multi-Solution + One-Knapsack

Algorithm 5: RandomizedMultiSolutionwithOneKnapsack

Input: probability $p \in (0,1)$, the number of candidates ℓ, ground set $\mathcal{N}$, budget B

Output: A feasible set S^*

$j \leftarrow 0$; **foreach** $i \in [\ell]$ **do** $S_i \leftarrow \emptyset$;
while true do
  foreach $i \in [\ell]$ **do**
    $\mathcal{M}_i \leftarrow \{e \in \mathcal{N} \setminus \{u_1, \ldots, u_j\} : c(S_i \cup \{e\}) \leq B\}$;
    $\mathcal{M}_i^* \leftarrow \{e \in \mathcal{N} \setminus S_i : c(S_i \cup \{e\}) \leq B\}$;
  foreach $i \in [\ell]$ **do**
    $u_{i,j}^* \leftarrow \operatorname{argmax}_{e \in \mathcal{M}_i^*} f(e|S_i)$;
    if $u_{i,j}^* \neq$ **NULL** $\wedge f(u_{i,j}^*|S_i) > 0$ **then**
      $S_{i,j}^* \leftarrow S_i \cup \{u_{i,j}^*\}$;
    else
      $S_{i,j}^* \leftarrow S_i$;
  $(t, u_{j+1}) \leftarrow \operatorname{argmax}_{i \in [\ell], e \in \mathcal{M}_i} \frac{f(e|S_i)}{c(e)}$;
  if $u_{j+1} \neq$ **NULL** $\wedge f(u_{j+1}|S_t) > 0$ **then**
    $S_t \leftarrow S_t \cup \{u_{j+1}\}$ with prob. p;
    $j \leftarrow j + 1$;
  else
    $h \leftarrow j$;
    break;
$S^* \leftarrow \operatorname{argmax}_{X \in \{S_{i,j}^*\}_{i \in [\ell], j \in [h] \cup \{0\}}} f(X)$;
return S^*

Definition. $S_1, \ldots, S_n$ are n candidate solutions when algorithm returns. $U = \{u_1, \ldots, u_\ell\}$ are all argmax elements in algorithm. Let

$$\begin{cases} U_i = \{u_{i,1}, \ldots, u_{i,\ell_i}\} \text{ is the argmax sequence for } S_i, \\ S_{i,j} \text{ is } S_i\text{'s first } j \text{ rounds result, } j \in [\ell_i] \end{cases}, \quad \tau_i(e) = \begin{cases} j, & \text{if } e = u_{i,j} \\ \infty, & \text{if } e \notin U_i \end{cases}$$

for each $i \in [n]$. Also we set $\tau(e) = \begin{cases} j, & \text{if } e = u_j \\ \infty, & \text{if } e \notin U \end{cases}$ and $S_{i,u} = S_{i,j-1}$, if $u = u_{i,j}$. Define

$$T_i = \min\left\{ j \;\middle|\; [0 \leq j \leq \ell_i - 1 \wedge c(S_{i,j}) + c(u_{i,j+1}) > c(O \setminus \{o_m\})] \vee (j = \ell_i) \right\}$$

$$O_{\leq T_i} = \{v \mid v \in O \wedge \tau_i(v) \leq T_i\}$$

$$O_{>T_i}^r = \{v \mid v \in O \wedge \tau_i(v) > T_i \wedge \tau(v) > \tau(u_{i,T_i}) \wedge f(v|S_{i,T_i}) > 0\}$$

$$O_{>T_i}^l = \{v \mid v \in O \wedge \tau_i(v) > T_i \wedge \tau(v) \leq \tau(u_{i,T_i}) \wedge f(v|S_{i,T_i}) > 0\}$$

Lemma 5.1. When $\ell_i > T_i$, there exists a mapping Ψ_i satisfying the following properties. For each $u \in O^r_{>T_i} \setminus \cup \{o_m\}$, $\Psi_i(u)$ is a set of 2-tuples s.t. each tuple $(v, \lambda_{i,v}(u)) \in \Psi_i(u)$ satisfies $v \in [S_{i,T_i} \setminus (O \setminus \{o_m\})] \cup \{u_{i,T_i+1}\}$ and $0 < \lambda_{i,v}(u) \leq \min\{c(u), c(v)\}$. Moreover, we have

$$\forall u \in O^r_{>T_i} \setminus \cup \{o_m\} : \sum_{(v,\lambda_{i,v}(u)) \in \Psi_i(u)} \lambda_{i,v}(u) = c(u)$$

$$\forall v \in [S_{i,T_i} \setminus (O \setminus \{o_m\})] \cup \{u_{i,T_i+1}\} : \sum_{u: \exists (v,\lambda_{i,v}(u)) \in \Psi_i(u)} \lambda_{i,v}(u) \leq c(v)$$

Lemma 5.2. For each $u \in V$, define $X_{i,u} = 1$ if $u \in O_{\leq T_i} \setminus (S_{i,T_i} \cup \{o_m\})$ or $u \in S_{i,T_i} \setminus (O \setminus \{o_m\})$, and we define $X_{i,u} = 0$ for any $u \in V$ that does not satisfy these conditions. Based on this definition, we have

$$f(S_{i,T_i} \cup O) \leq f(S_{i,T_i} \cup \{o_m\}) + \sum_{u \in V} X_{i,u} f(u|S_{i,u}) + f(S^*_{i,T_i}|S_{i,T_i})$$

Proof. Notice that

$$\begin{aligned}
&f(S_{i,T_i} \cup O) - f(S_{i,T_i} \cup \{o_m\}) \leq \sum_{O \setminus (S_{i,T_i} \cup \{o_m\})} f(u|S_{i,T_i}) \\
&\leq \sum_{u \in O^l_{>T_i} \setminus \{o_m\}} f(u|S_{i,T_i}) + \sum_{u \in O^r_{>T_i} \setminus \{o_m\}} f(u|S_{i,T_i}) + \sum_{O_{\leq T_i} \setminus (S_{i,T_i} \cup \{o_m\})} f(u|S_{i,T_i})
\end{aligned}$$

If $\ell_i = T_i$ (which implies $c(S_{i,\ell_i}) \leq c(O \setminus \{o_m\})$), then we must have $O^r_{>T_i} = \emptyset$, because otherwise there exists an element in $O^r_{>T_i}$ that can be added into S_{i,T_i} without violating the budget constraint, contradicting $\ell_i = T_i$. So we have

$$\begin{aligned}
&f(S_{i,T_i} \cup O) - f(S_{i,T_i} \cup \{o_m\}) \leq \sum_{O_{\leq T_i} \setminus (S_{i,T_i} \cup \{o_m\})} f(u|S_{i,T_i}) \\
&\leq \sum_{O_{\leq T_i} \setminus (S_{i,T_i} \cup \{o_m\})} f(u|S_{i,T_i}) + \sum_{S_{i,T_i} \setminus (O \setminus \{o_m\})} f(u|S_{i,u}) + f(S^*_{i,T_i}|S_{i,T_i}) \\
&= \sum_{u \in V} X_{i,u} f(u|S_{i,u}) + f(S^*_{i,T_i}|S_{i,T_i})
\end{aligned}$$

Now we consider the case $\ell_i > T_i$. In this case, we can use Lemma to get

$$
\begin{aligned}
\sum_{u \in O^r_{>T_i}\backslash\{o_m\}} f(u|S_{i,T_i}) &= \sum_{u \in O^r_{>T_i}\backslash\{o_m\}} \frac{f(u|S_{i,T_i})}{c(u)} \cdot c(u) \\
&= \sum_{u \in O^r_{>T_i}\backslash\{o_m\}} \frac{f(u|S_{i,T_i})}{c(u)} \sum_{(v,\lambda_{i,v}(u)) \in \Psi_i(u)} \lambda_{i,v}(u) \\
&\leq \sum_{u \in O^r_{>T_i}\backslash\{o_m\}} \sum_{(v,\lambda_{i,v}(u)) \in \Psi_i(u)} \frac{f(v|S_{i,v})}{c(v)} \cdot \lambda_{i,v}(u) \\
&= \sum_{v \in [S_{i,T_i}\backslash(O\backslash\{o_m\})] \cup \{u_{i,T_i+1}\}} \sum_{u: \exists (v,\lambda_{i,v}(u)) \in \Psi_i(u)} \frac{f(v|S_{i,v})}{c(v)} \cdot \lambda_{i,v}(u) \\
&\leq \sum_{v \in [S_{i,T_i}\backslash(O\backslash\{o_m\})] \cup \{u_{i,T_i+1}\}} \frac{f(v|S_{i,v})}{c(v)} \cdot c(v) \\
&\leq f(u^*_{i,T_i}|S_{i,T_i}) + \sum_{v \in S_{i,T_i}\backslash(O\backslash\{o_m\})} f(v|S_{i,v}) \\
&= f(S^*_{i,T_i}|S_{i,T_i}) + \sum_{v \in S_{i,T_i}\backslash(O\backslash\{o_m\})} f(v|S_{i,v})
\end{aligned}
$$

Lemma 5.3. We have

$$\mathbb{E}[f(S_{i,T_i})] = \mathbb{E}\Big[\sum_{u \in V} X_{i,u} f(u|S_{i,u})\Big]$$

Proof. Define $R_{i,u}$ a random variable denoting the marginal contribution of u for $f(S_{i,T_i})$, i.e., $R_{i,u} = f(u|S_{i,u})$ if $u \in S_{i,T_i}$ and $R_{i,u} = 0$ otherwise. So we have $f(S_{i,T_i}) = \sum_{u \in V} R_{i,u}$ and hence we just prove

$$\forall u \in V : \mathbb{E}[R_{i,u}] = \mathbb{E}[X_{i,u} f(u|S_{i,u})]$$

For any $u \in V$, let $\mathcal{E}_{i,u}$ be an arbitratry event specifying all the random choices of algorithm up

1. until the moment that u is considered by Line 8 with index i, and $c(S_{i,u} \cup \{u\}) \leq c(O\backslash\{o_m\})$.
2. until the moment that u is considered by Line 8 with index i, and $c(S_{i,u} \cup \{u\}) > c(O\backslash\{o_m\})$.
3. until the moment that u is considered by Line 8 but the index is not i.
4. if u is never considered by Line 8.

For case 2, 3 and 4, by definition we have

$$\mathbb{E}[R_{i,u}|\mathcal{E}_{i,u}] = \mathbb{E}[X_{i,u} f(u|S_{i,u})|\mathcal{E}_{i,u}] = 0$$

For case 1, $\mathbb{E}[R_{i,u}|\mathcal{E}_{i,u}] = \frac{1}{2} f(u|S_{i,u})$. Now we consider

(1) $u \in O\backslash\{o_m\}$.
In this case, we must have $u \notin S_{i,T_i}\backslash(O\backslash\{o_m\})$. If u is not discarded by S_i, then $u \in S_{i,T_i}$ and $X_{i,u} = 0$. If u is discarded, then we have $u \in O_{\leq T_i}\backslash(S_{i,T_i} \cup \{o_m\})$ and hence $X_{i,u} = 1$. Therefore, we get $\mathbb{E}[X_{i,u} f(u|S_{i,u})|\mathcal{E}_{i,u}] = \frac{1}{2} f(u|S_{i,u})$.

(2) $u \notin O\backslash\{o_m\}$.

Algorithm 6: STREAMINGRANDOMIZEDKSYSTEM

Input: $p \in (0,1)$, $\tau_1 > \cdots > \tau_\ell$, streaming elements of k-system $(\mathcal{N}, \mathcal{I})$

Output: A feasible set S^*

1 $S_j \leftarrow \emptyset,\ j = 1, \ldots, \ell$;

2 Take a new pass over the stream;

3 **while** *there is an incoming element* u **do**

4   **if** $\{u\} \notin \mathcal{I}$ **then continue**;

5     $i \leftarrow 1$;

6   **while** $i \leq \ell$ **do**

7   **if** $f(u| \cup_{j=1}^{i} S_j) \geq \tau_i$ **then**

8   **if** $S_i + u \in \mathcal{I}$ **then**

9                 $S_i \leftarrow S_i + u$ with prob. p;

10   **break**

11         $i \leftarrow i + 1$;

12 **return** $S^* = \text{argmax}_{X \in \{S_1, \ldots, S_\ell\}} f(X)$

6 Streaming + Randomized + k-System

Definition. Let O be an optimal solution (i.e. $O = \text{argmax}_{S \in \mathcal{I}} f(S)$), ρ be the size of largest independent set (i.e. $\rho = \max_{S \in \mathcal{I}} |S|$), $\tau_1 = \max_{u \in \mathcal{N}, \{u\} \in \mathcal{I}} f(\{u\})$ and $\tau_i = \tau_1 / q^{i-1},\ i = 2, \ldots, \ell$, where $q > 1$ is a constant to be determined later. Define

$$
\begin{aligned}
&S^{<}(e):\ \text{the set of elements in } S \text{ the moment that algoritm meets element } e\\
&O_i = \{e \in O:\ \text{Considered by line 8 with threshold } \tau_i\}\\
&O_{i1} = \{e \in O_i:\ \text{Do not satisfy the condition of line 8}\}\\
&O_{i2} = \{e \in O_i:\ \text{Satisfy the condition of line 8 but discarded by line 9}\}\\
&O_{>\ell} = O \backslash (\cup_{i=1}^{\ell} O_i)
\end{aligned}
$$

For the first set S_1, due to the submodularity we have

$$
\begin{aligned}
f(O_1 \cup S_1) - f(S_1) &\leq \sum_{e \in O_1 \backslash S_1} f(e|S_1)\\
&= \sum_{e \in O_{11}} f(e|S_1) + \sum_{e \in O_{12}} f(e|S_1)
\end{aligned} \tag{1}
$$

Here we just consider the O_{11} part and leave the other part O_{12} later to be solved. Since S_1 is a base of $O_{11} \cup S_1$ and $O_{11} \in \mathcal{I}$, we have $|O_{11}| \leq k|S_1|$. And for each $e \in S_1$ we have $\tau_1 \leq f(e|S_1^{<}(e)) \leq f(e) \leq \tau_1$, so $f(e|S_1^{<}(e)) = \tau_1$. So we have

$$\sum_{e \in O_{11}} f(e|S_1) \leq \sum_{e \in O_{11}} f(e) < |O_{11}|\tau_1 \leq k|S_1|\tau_1 = kf(S_1)$$

For the second one, we have

$$
\begin{aligned}
f(O_2 \cup S_1 \cup S_2) - f(S_1 \cup S_2) \le & \sum_{e \in O_2 \setminus (S_1 \cup S_2)} f(e|S_1 \cup S_2) \\
= & \sum_{e \in O_{21}} f(e|S_1 \cup S_2) + \sum_{e \in O_{22}} f(e|S_1 \cup S_2)
\end{aligned}
$$

Also, we just consider the O_{21} part. Since S_2 is a base of $O_{21} \cup S_2$, we have $|O_{21}| \le k|S_2|$. And for each $e \in O_2$, we have $\tau_2 \le f(e|S_1 \cup S_2) \le f(e|S_1) < \tau_1$. So we have

$$
\begin{aligned}
\sum_{e \in O_{21}} f(e|S_1 \cup S_2) \le & \sum_{e \in O_{21}} f(e|S_1) < \tau_1 |O_{21}| \le k\tau_1 |S_2| \\
\le & k\frac{\tau_1}{\tau_2} \sum_{e \in S_2} f(e|S_1 \cup S_2) \le qk \sum_{e \in S_2} f(e|S_2^{<}(e)) = qkf(S_2)
\end{aligned}
$$

So we can do the same thing to get the t-th ($t = 2, \ldots, \ell$) equality as

$$
\begin{aligned}
f\big(O_t \cup \cup_{i=1}^t S_i\big) - f\big(\cup_{i=1}^t S_i\big) \le & \sum_{e \in O_{t1}} f\big(e|\cup_{i=1}^t S_i\big) + \sum_{e \in O_{t2}} f\big(e|\cup_{i=1}^t S_i\big) \\
\le & qkf(S_t) + \sum_{e \in O_{t2}} f\big(e|\cup_{i=1}^t S_i\big)
\end{aligned}
\tag{2}
$$

For the set $O_{>\ell}$, we have

$$
f\big(O_{>\ell}|\cup_{i=1}^\ell S_i\big) \le \sum_{e \in O_{>\ell}} f\big(e|\cup_{i=1}^\ell S_i\big) < |O_{>\ell}|\frac{\tau_1}{q^{\ell-1}} \le \rho\tau_1 q^{1-\ell} \tag{3}
$$

Now add the $\ell + 1$ inequalities (1), (2) and (3) together, we will get

$$
\begin{aligned}
& \sum_{t=1}^{\ell} \Big(f\big(O_t \cup \cup_{i=1}^t S_i\big) - f\big(\cup_{i=1}^t S_i\big) \Big) + f\big(O_{>\ell}|\cup_{i=1}^\ell S_i\big) \\
\le & \sum_{t=1}^{\ell} \Bigg(\sum_{e \in O_{t1}} f\big(e|\cup_{i=1}^t S_i\big) + \sum_{e \in O_{t2}} f\big(e|\cup_{i=1}^t S_i\big) \Bigg) + f\big(O_{>\ell}|\cup_{i=1}^\ell S_i\big) \\
\le & kf(S_1) + qk\sum_{i=2}^{\ell} f(S_i) + \sum_{t=1}^{\ell}\sum_{e \in O_{t2}} f\big(e|\cup_{i=1}^t S_i\big) + \rho\tau_1 q^{1-\ell}
\end{aligned}
\tag{4}
$$

By submodularity, the left-hand side is

$$
\begin{aligned}
& \sum_{t=1}^{\ell} \Big(f\big(O_t \cup \cup_{i=1}^t S_i\big) - f\big(\cup_{i=1}^t S_i\big) \Big) + f\big(O_{>\ell}|\cup_{i=1}^\ell S_i\big) \\
\ge & f\big(\cup_{i=1}^2 (O_i \cup S_i)\big) - f\big(\cup_{i=1}^2 S_i\big) + \sum_{t=3}^{\ell} \Big(f\big(O_t \cup \cup_{i=1}^t S_i\big) - f\big(\cup_{i=1}^t S_i\big) \Big) + f\big(O_{>\ell}|\cup_{i=1}^\ell S_i\big) \\
\ge & f\big(\cup_{i=1}^3 (O_i \cup S_i)\big) - f\big(\cup_{i=1}^3 S_i\big) + \sum_{t=4}^{\ell} \Big(f\big(O_t \cup \cup_{i=1}^t S_i\big) - f\big(\cup_{i=1}^t S_i\big) \Big) + f\big(O_{>\ell}|\cup_{i=1}^\ell S_i\big) \\
\ge & \cdots \ge f\big(\cup_{i=1}^\ell (O_i \cup S_i)\big) - f\big(\cup_{i=1}^\ell S_i\big) + f\big(O_{>\ell}|\cup_{i=1}^\ell S_i\big) \\
\ge & f\big(O \cup \cup_{i=1}^\ell S_i\big) - f\big(\cup_{i=1}^\ell S_i\big)
\end{aligned}
\tag{5}
$$

Now we consider the third part of the right-hand side of (4). We will prove the following lemma.

Lemma. $\mathbb{E}\Big[\sum_{t=1}^{\ell}\sum_{e\in O_{t2}} f\big(e|S_t\big)\Big] \le \frac{1-p}{p}\mathbb{E}\Big[\sum_{i=1}^{\ell} f(S_i)\Big]$

Proof. For each $e\in\mathcal{N}$, we define random variables

$$R_e = \begin{cases}\delta(e)=\sum_{i=1}^{\ell}\mathbb{I}_{S_i}(e)f(e|S_i^{<}(e)), & \text{if } e\in\cup_{i=1}^{\ell}S_i\\ 0, & \text{else}\end{cases},\quad X_e=\begin{cases}1, & \text{if } e\in\cup_{i=1}^{\ell}O_{i2}\\ 0, & \text{else}\end{cases}$$

where $\mathbb{I}_S$ is the characteristic function for set S. So we have $\sum_{i=1}^{\ell} f(S_i)=\sum_{e\in\mathcal{N}} R_e$ and hence we only need to prove

$$\mathbb{E}\Big[X_e\sum_{i=1}^{\ell}\mathbb{I}_{O_{i2}}(e)f(e|S_i^{<}(e))\Big] \le \frac{1-p}{p}\mathbb{E}[R_e],\ \forall e\in\mathcal{N}$$

due to the linearity of expectation. Note that $\mathbb{E}[R_e]=p\delta(e)$ because of the conditional probability. Thus, we only need to prove LHS $\le (1-p)\delta(e)$, as done by the following discussions:

(1) $e\in O$:

In this case, if e is not discarded by line 9, then we have $e\notin\cup_{i=1}^{\ell}O_{i2}$ and hence $X_e=0$. If e is discarded, then we have $e\in\cup_{i=1}^{\ell}O_{i2}$ and hence $X_e=1$. Therefore, we get LHS $=(1-p)\delta(e)$.

(2) $e\notin O$:

In this case, we must have $X_e=0$. Therefore, we get LHS $=0$.

So the lemma follows from the above analysis.

Now we take expectation on the both sides of inequality (4) applying the results of (5) and the lemma, we have

$$\begin{aligned}\mathbb{E}\Big[f\big(O\cup\cup_{i=1}^{\ell}S_i\big)\Big] \le &\mathbb{E}\Big[f\big(\cup_{i=1}^{\ell}S_i\big)+kf(S_1)+qk\sum_{i=2}^{\ell}f(S_i)\Big]+\mathbb{E}\Big[\sum_{t=1}^{\ell}\sum_{e\in O_{t2}}f(e|S_t)\Big]+\rho\tau_1 q^{1-\ell}\\ \le&\mathbb{E}\Big[\sum_{i=1}^{\ell}f(S_i)+kf(S_1)+qk\sum_{i=2}^{\ell}f(S_i)\Big]+\frac{1-p}{p}\mathbb{E}\Big[\sum_{i=1}^{\ell}f(S_i)\Big]+\rho\tau_1 q^{1-\ell}\\ \le&\Big((1+q(\ell-1))k+\frac{\ell}{p}\Big)\mathbb{E}[f(S^*)]+\rho\tau_1 q^{1-\ell}\end{aligned}$$

Since LHS $\ge (1-p)f(O)$ and $\tau_1\le f(O)$, we finally get

$$\mathbb{E}[f(S^*)] \ge \frac{1-p-\rho q^{1-\ell}}{(1+q(\ell-1))k+\frac{\ell}{p}}f(O)$$