# A ADDITIONAL PROOFS

## A NP-hardness

Proof. Let $SK(U, S, P, W, q)$ be an instance of the Set Union Knapsack Problem [15], where $U = \{u_1, \ldots u_n\}$ is a set of items, $S = \{S_1, \ldots S_m\}$ is a set of subsets ($S_i \subseteq U$), $P : S \to \mathbb{R}_+$ is a subset profit function, $w : U \to \mathbb{R}_+$ is an item weight function, and $q \in \mathbb{R}_+$ is the budget. For a subset $\mathcal{A} \subseteq S$, the weighted union of set $\mathcal{A}$ is $W(\mathcal{A}) = \sum_{e \in \cup_{t \in \mathcal{A}} S_t} w_e$ and $P(\mathcal{A}) = \sum_{t \in \mathcal{A}} p_t$. The problem is to find a subset $\mathcal{A}^* \subseteq S$ such that $W(\mathcal{A}^*) \leq q$ and $P(\mathcal{A}^*)$ is maximized. SK is NP-hard to approximate within a constant factor [3]. We reduce a version of $SK$ with equal profits and weights (also NP-hard) to the Mbed problem. We define a corresponding Mbed problem instance via constructing a graph $\Gamma$ as follows.

For each $S_i \in S$ and $u_j \in U$ we create nodes $x_i$ and $y_j$ respectively. We also add a node $v$ with a large connected component $L$ of size $l$ only with positive edges attached to it. The node $v$ has negative edges with every node $x_i$, $\forall i \in [m]$ and every node $y_j$, $\forall j \in [n]$. Additionally, if $u_j \in S_i$, a negative edge $(x_i, y_j)$ will be added to the edge set $E$.

In Mbed, the number of edges to be removed is the budget, $b = q$. The candidate set, $\mathbb{C} = \{(v, y_j) | \forall j \in [n]\}$. Note that initial largest connected balanced component is $\{v \cup L\} \cup \{y_j \forall j \in [n]\}$ if $l > m+1$ (assuming $n > m$). Our claim is that, for any solution $\mathcal{A}$ of an instance of $SK$ there is a corresponding solution set of edges, $B$ (where $|B| = b$) in the graph $\Gamma$ of the Mbed version, such that $f(B) = P(\mathcal{A}) + n + l + 1$ if $B = \{(v, y) | y \in \mathcal{A}\}$ are removed.

In the new balanced graph, we aim to build two partitions ($W_1$ and $W_2$) as follows. One partition $W_1$ consists of $\{v \cup L\}$ initially. Our goal is to delete edges from $\mathbb{C}$ and add the nodes $y_j$'s in $W_1$. If $(v, y_{j'})$ for any $j'$ does not get deleted then it would be in $W_2$. If there is any node $x_i$ that is connected with only nodes in $\mathcal{A}$ beside being connected with $v$, then removing all the edges in $B$ would put the node $x_i$ in $W_2$. Thus removing edges in $\mathcal{A}$ would put $P(\mathcal{A})$ nodes in $W_2$. Thus, $f(B) = P(\mathcal{A}) + n + l + 1$.

□

## B Proportionally Submodular

Lemma A.1. *The objective function $f$ is not proportionally submodular[31]. In other words, there exists $S, T \in E$ for some graph $H$ such that $|T|f(S) + |S|f(T) < |S \cap T|f(S \cup T) + |S \cup T|f(S \cap T)$.*

Proof. Consider a balanced subgraph of $H$, $S(H)$ has a partition $V_1$ and $V_2$. A node $v$ is outside $S(H)$ and it is connected to $V_1$ with positive edges $e_1$ and $e_2$, $V_2$ with another positive edge $e_3$. Thus the node $v$ cannot be the part of $S(H)$. Consider an edge $e_4$ inside $V_1$ which can be removed without making the graph disconnected. Let us assume $S = \{e_1, e_4\}, T = \{e_2, e_4\}$. Then, $f(\{e_1, e_4\}) = 0$ and $f(\{e_2, e_4\}) = 0$, since even after removing any of these edges it is not possible to add the node $v$ to $S(H)$. Note that $f(S \cap T) = f(\{e_4\}) = 0$. However, $f(S \cup T) = f(\{e_1, e_2, e_4\}) = 1$ since the node $v$ can be added. Substituting these values, we get $|T|f(S) + |S|f(T) < |S \cap T|f(S \cup T) + |S \cup T|f(S \cap T)$. □

## C Proof of Lemma 9

Proof. We prove this by induction on the number of edges, $b$. Let us denote $B_k \subseteq B$ as $\{e_1, \cdots, e_k\}$. We construct $B$ by only considering peripheral edges $e_{k+1}$ such that, for all $k \leq b$: $(e_{k+1}, e') \in cep(H_{B_k}, x)$, for some node $x$ and edge $e'$.

**Base case** ($b = 1$)**:** $f(\{e_1\}) \leq f(\{e_1\})$. Also, $\alpha(\{e_1\}) = 0$.

**Inductive hypothesis (IH):** Suppose the equation holds for $b = k$, i.e., $f(B_k) \leq \sum_{i=1}^{k} f(\{e_i\}) + (C^* + 1)\alpha(B_k)$.

**Inductive step** ($b = k + 1$)**:** We present different cases for $e_{k+1}$. Note that we have $(e_{k+1}, e') \in cep(H_{B_k}, x)$ for some $x, e'$.

**Case 1:** $(e_{k+1}, e') \in cep(H, x)$ and $\left|cep\left(H_{e_{k+1}}, x\right)\right| = 0$, i.e., after deleting $e_{k+1}$, $x$ moves into the balanced subgraph. Then, we must also have $\left|cep\left(H_{B_{k+1}}, x\right)\right| = 0$. Hence, $f(B_k \cup \{e_{k+1}\}) - f(B_k) = f(\{e_{k+1}\})$ and the inequality holds.

**Case 2:** Either (1) $(e_{k+1}, e') \in cep(H, x)$ and $\left|cep\left(H_{e_{k+1}}, x\right)\right| > 0$ or (2) $(e_{k+1}, e') \notin cep(H, x)$.

Thus, by Observation 2, we have $f(\{e_{k+1}\}) = 0$.

**Case 2a:** Suppose $\left|cep\left(H_{B_{k+1}}, x\right)\right| = 0$. Then by definition of $\alpha, C^*$, we have $\alpha(B_{k+1}) = \alpha(B_k) + 1$, and $f(\{B_k \cup e_{k+1}\}) - f(B_k) \leq C^* + 1$.

Substituting this, we get $f(B_{k+1}) \leq (C^* + 1) + \sum_{i=1}^{k} f(\{e_i\}) + (C^* + 1)\alpha(B_k) = \sum_{i=1}^{k+1} f(\{e_i\}) + (C^* + 1)\alpha(B_{k+1})$.

**Case 2b:** In other cases, $f(B_k \cup \{e_{k+1}\}) - f(B_k) = f(\{e_{k+1}\}) = 0$.

This exhausts our cases and the claim is true $\forall b, b > 0$. □

## D Construction for the tight lower bound in Thm. 2

One can construct a graph $H$ and the sets $Q, R$ where equality holds. In particular, let $R$ be of an arbitrary size $b$. Consider $H_Q$ to have the MBS partition as $V_1, V_2$ each of size $\frac{\Delta(H_Q)}{2}$. Nodes of type 1 (Obs. 2) are attached to these each with the sole connected component of size $\frac{\Delta(H_Q)-2}{2}$. Let these nodes have 3 such connections (thus, removing two will help - any two such that our "connected assumption" holds are in the set $R$). We have another node of type 1 such that only two such connections are connected and one of these is in $R$ and the connected component $C$ to it is of size 0. This completes the set $R$. Thus, $\sum_{e \in R} [f(Q \cup \{e\}) - f(Q) = 1$ and $f(Q \cup R) - f(Q) = 1 + \left(\frac{\Delta(H_Q)-2}{2} + 1\right)\frac{b-1}{2}$.