A. Omitted Proofs

This appendix contains the proofs that have been omitted from Section 3.

A.1. Proof of Lemma 3.3

Proof. We first prove by induction that, for every $0 \leq i \leq k$,

$$\mathbb{E}[f(OPT_i)] \geq [1 - e^{-\gamma \cdot \sum_{j=i+1}^{k} j^{-1}}] \cdot f(OPT) \ . \tag{2}$$

For $i = k$, Inequality (2) follows from the non-negativity of f since

$$f(OPT_k) \geq 0 = [1 - e^{-\gamma \cdot 0}] \cdot f(OPT) = [1 - e^{-\gamma \cdot \sum_{j=k+1}^{k} j^{-1}}] \cdot f(OPT) \ .$$

Assume now that Inequality (2) holds for some $0 < i + 1 \leq k$, and let us prove it holds also for i. Observe that OPT_i is a uniformly random subset of OPT of size $k - i$, and OPT_{i+1} is a uniformly random subset of OPT of size $k - i - 1$. Thus, we can think of OPT_i as obtained from OPT_{i+1} by adding a uniformly random element of $OPT \setminus OPT_{i+1}$. Taking this point of view, we get, for every set $S \subseteq OPT$ of size $k - i - 1$,

$$\begin{aligned}\mathbb{E}[f(OPT_i) \mid OPT_{i+1} = S] &= f(S) + \frac{\sum_{u \in OPT \setminus S} f(u \mid S)}{|OPT \setminus S|} = f(S) + \frac{1}{i+1} \cdot \sum_{u \in OPT \setminus S} f(u \mid S) \\ &\geq f(S) + \frac{\gamma}{i+1} \cdot f(OPT \setminus S \mid S) = \left(1 - \frac{\gamma}{i+1}\right) \cdot f(S) + \frac{\gamma}{i+1} \cdot f(OPT) \ ,\end{aligned}$$

where the inequality holds by the γ-weak submodularity of f. Taking expectation over the set OPT_{i+1}, the last inequality becomes

$$\begin{aligned}\mathbb{E}[f(OPT_i)] &\geq \left(1 - \frac{\gamma}{i+1}\right) \cdot \mathbb{E}[f(OPT_{i+1})] + \frac{\gamma}{i+1} \cdot f(OPT) \\ &\geq \left(1 - \frac{\gamma}{i+1}\right) \cdot [1 - e^{-\gamma \cdot \sum_{j=i+2}^{k} j^{-1}}] \cdot f(OPT) + \frac{\gamma}{i+1} \cdot f(OPT) \\ &= f(OPT) - \left(1 - \frac{\gamma}{i+1}\right) \cdot e^{-\gamma \cdot \sum_{j=i+2}^{k} j^{-1}} \cdot f(OPT) \\ &\geq f(OPT) - e^{-\gamma \cdot \sum_{j=i+1}^{k} j^{-1}} \cdot f(OPT) = [1 - e^{-\gamma \cdot \sum_{j=i+1}^{k} j^{-1}}] \cdot f(OPT) \ ,\end{aligned}$$

where the second inequality follows by the induction hypothesis, and the last inequality follows by the inequality $1 - x \leq e^{-x}$ (which holds for every x). This completes the proof of Inequality (2). To see why the lemma follows from this inequality, we observe that

$$\sum_{j=i+1}^{k} \frac{1}{j} \geq \int_{i+1}^{k+1} \frac{dx}{x} = \ln x|_{i+1}^{k+1} = \ln\left(\frac{k+1}{i+1}\right) \ ,$$

which implies (by Inequality (2))

$$\mathbb{E}[f(OPT_i)] \geq [1 - e^{-\gamma \cdot \sum_{j=i+1}^{k} j^{-1}}] \cdot f(OPT) \geq \left[1 - \left(\frac{i+1}{k+1}\right)^{\gamma}\right] \cdot f(OPT) \ . \qquad \square$$

A.2. Proof of Observation 3.4

Proof. Fix $1 \leq i \leq k$, and let A_{i-1} be an arbitrary event fixing all the random decisions of Algorithm 1 up to iteration $i - 1$ (including). All the probabilities, expectations and random quantities in the first part of this proof are implicitly conditioned on A_{i-1}. The γ-weak submodularity of f implies

$$\sum_{u \in OPT_{i-1}} f(u \mid S_{i-1}) \geq \gamma \cdot f(OPT_{i-1} \mid S_{i-1}) \ .$$

Since OPT_{i-1} is one possible candidate to be M_i, we get

$$\sum_{u \in M_i} f(u \mid S_{i-1}) \geq \sum_{u \in OPT_{i-1}} f(u \mid S_{i-1}) \geq \gamma \cdot f(OPT_{i-1} \mid S_{i-1}) \enspace .$$

Algorithm 1 gets S_i by adding a uniformly random element $u_i \in M_i$ to the set S_{i-1}. Since the size of M_i is $k - i + 1$, this implies

$$\begin{aligned}\mathbb{E}[f(S_i)] = f(S_{i-1}) + \mathbb{E}[f(u_i \mid S_{i-1})] &= f(S_{i-1}) + \frac{1}{k-i+1} \cdot \sum_{u \in M_i} f(u \mid S_{i-1}) \\ &\geq f(S_{i-1}) + \frac{\gamma}{k-i+1} \cdot f(OPT_{i-1} \mid S_{i-1}) = f(S_{i-1}) + \gamma \cdot \frac{f(OPT_{i-1} \cup S_{i-1}) - f(S_{i-1})}{k-i+1} \enspace .\end{aligned}$$

Recall that the last inequality is implicitly conditioned on the event A_{i-1}. The observation now follows by taking the expectation of both sides of this inequality over all possible such events. □

A.3. Proof of Corollary 3.5

Proof. Note that

$$\begin{aligned}\mathbb{E}[f(S_i)] &\geq \mathbb{E}[f(S_{i-1})] + \gamma \cdot \frac{\mathbb{E}[f(OPT_{i-1} \cup S_{i-1})] - \mathbb{E}[f(S_{i-1})]}{k-i+1} \\ &\geq \mathbb{E}[f(S_{i-1})] + \gamma \cdot \frac{\mathbb{E}[f(OPT_{i-1})] - \mathbb{E}[f(S_{i-1})]}{k-i+1} \\ &\geq \mathbb{E}[f(S_{i-1})] + \gamma \cdot \frac{\{1 - [i/(k+1)]^\gamma\} \cdot f(OPT) - \mathbb{E}[f(S_{i-1})]}{k-i+1} \enspace ,\end{aligned}$$

where the first inequality follows by Observation 3.4, the second inequality holds by the monotonicity of f, and the last inequality follows by Lemma 3.3. □

A.4. Proof of Theorem 3.6

Proof. Let us first prove by induction that, for every $0 \leq i \leq k$,

$$\mathbb{E}[f(S_i)] \geq \gamma \cdot \frac{\sum_{j=1}^{i} \{1 - [j/(k+1)]^\gamma\} \cdot f(OPT) - i \cdot \mathbb{E}[f(S_k)]}{k} \enspace . \tag{3}$$

For $i = 0$, Inequality (3) follows from the non-negativity of f since

$$\mathbb{E}[f(S_0)] \geq 0 = \gamma \cdot \frac{0 \cdot f(OPT) - 0}{k} = \gamma \cdot \frac{\sum_{j=1}^{0} \{1 - [j/(k+1)]^\gamma\} \cdot f(OPT) - 0 \cdot \mathbb{E}[f(S_k)]}{k} \enspace .$$

Assume now that Inequality (3) holds for some $0 \leq i - 1 < k$, and let us prove that it holds for i as well. There are two cases to consider. If $\{1 - [i/(k+1)]^\gamma\} \cdot f(OPT) \leq \mathbb{E}[f(S_k)]$, then the monotonicity of f and the fact that S_{i-1} is a subset of S_i guarantee together that

$$\begin{aligned}\mathbb{E}[f(S_i)] \geq \mathbb{E}[f(S_{i-1})] &\geq \gamma \cdot \frac{\sum_{j=1}^{i-1} \{1 - [j/(k+1)]^\gamma\} \cdot f(OPT) - (i-1) \cdot \mathbb{E}[f(S_k)]}{k} \\ &\geq \gamma \cdot \frac{\sum_{j=1}^{i} \{1 - [j/(k+1)]^\gamma\} \cdot f(OPT) - i \cdot \mathbb{E}[f(S_k)]}{k} \enspace ,\end{aligned}$$

where the second inequality follows by the induction hypothesis. Thus, it remains to consider the case that $\{1 - [i/(k+1)]^\gamma\} \cdot f(OPT) \geq \mathbb{E}[f(S_k)]$. By Corollary 3.5, we get in this case

$$\begin{aligned}\mathbb{E}[f(S_i)] - \mathbb{E}[f(S_{i-1})] &\geq \gamma \cdot \frac{\{1 - [i/(k+1)]^\gamma\} \cdot f(OPT) - \mathbb{E}[f(S_{i-1})]}{k-i+1} \\ &\geq \gamma \cdot \frac{\{1 - [i/(k+1)]^\gamma\} \cdot f(OPT) - \mathbb{E}[f(S_k)]}{k-i+1} \\ &\geq \gamma \cdot \frac{\{1 - [i/(k+1)]^\gamma\} \cdot f(OPT) - \mathbb{E}[f(S_k)]}{k} \enspace ,\end{aligned}$$

where the second inequality follows by the monotonicity of f since S_{i-1} is a subset of S_k. Adding the last inequality to the induction hypothesis proves that Inequality (3) holds for i, and thus, completes the proof by induction of Inequality (3).

Let us now explain why the theorem follows from Inequality (3). Plugging $i = k$ into this inequality yields

$$\mathbb{E}[f(S_k)] \geq \gamma \cdot \frac{\sum_{j=1}^{k}\{1 - [j/(k+1)]^\gamma\} \cdot f(OPT) - k \cdot \mathbb{E}[f(S_k)]}{k} \ ,$$

which implies, by extracting $\mathbb{E}[f(S_k)]$,

$$\begin{aligned}
\mathbb{E}[f(S_k)] &\geq \frac{\gamma \cdot f(OPT)}{(1+\gamma)k} \cdot \sum_{j=1}^{k}\left[1 - \left(\frac{j}{k+1}\right)^\gamma\right] \geq \frac{\gamma \cdot f(OPT)}{(1+\gamma)k} \cdot \int_1^{k+1}\left[1 - \left(\frac{x}{k+1}\right)^\gamma\right] dx \\
&= \frac{\gamma \cdot f(OPT)}{(1+\gamma)k} \cdot \left[x - \frac{k+1}{1+\gamma} \cdot \left(\frac{x}{k+1}\right)^{1+\gamma}\right]_1^{k+1} \geq \frac{\gamma \cdot f(OPT)}{(1+\gamma)k} \cdot \left[k - \frac{k+1}{1+\gamma}\right] \\
&= \frac{\gamma \cdot f(OPT)}{(1+\gamma)k} \cdot \left[\frac{\gamma k}{1+\gamma} - \frac{1}{1+\gamma}\right] \geq \left[\left(\frac{\gamma}{1+\gamma}\right)^2 - \frac{1}{k}\right] \cdot f(OPT) \ . \qquad \square
\end{aligned}$$

B. Removal of the Low Order Term

Theorem 3.6 proves an approximation ratio guarantee for Algorithm 1 which is weaker than the approximation ratio guarantee of Theorem 1.1 by a low order term of $O(k^{-1})$. In this appendix we prove that this low order term can be dropped from the guarantee of Theorem 3.6, which yields Theorem 1.1. In fact, we even prove the following stronger theorem.

Theorem B.1. *Let $c(\gamma)$ be an arbitrary function of γ, and let $\varepsilon(k)$ be a function of k that approaches 0 as k increases. Then, if the approximation ratio of Algorithm 1 is at least $c(\gamma) - \varepsilon(k)$, then it is also at least $c(\gamma)$.*

Proof. Assume towards a contradiction that the theorem is wrong. This implies that the approximation ratio of Algorithm 1 is at least $c(\gamma) - \varepsilon(k)$, and yet there exists an instance I with a γ_I-weakly submodular objective function on which the approximation ratio of Algorithm 1 is $c' < c(\gamma_I)$. Since $\varepsilon(k)$ approaches 0 when k increases, we can find a value k' such that $c' < c(\gamma_I) - \varepsilon(k)$ for every $k \geq k'$. Note that we may assume, without loss of generality, that k' is a non-negative integer. Consider now the variant of Algorithm 1 given as Algorithm 2. Intuitively, this variant extends the input by introducing k' new elements which do not affect the objective function and can be used to extend every independent set of the matroid.

Algorithm 2 Residual Random Greedy for Matroids (Variant)$(f, \mathcal{M})$

Create k' new elements, and let $\mathcal{N}'$ denote the set of these new elements.
Extend the object function f to the ground set $\mathcal{N} \cup \mathcal{N}'$ by setting $f(S) = f(S \setminus \mathcal{N}')$ for every set S which includes new elements.
Extend the matroid $\mathcal{M}$ to the ground set $\mathcal{N} \cup \mathcal{N}'$ by defining that a set S which includes new elements is independent if and only if $S \setminus \mathcal{N}'$ is independent.
Initialize: $S_0 \leftarrow \varnothing$.
for $i \leftarrow 1, 2, \ldots, k + k'$ **do**
    Let M_i be a base of $\mathcal{M}/S_{i-1}$ maximizing $\sum_{u \in M_i} f(u \mid S_{i-1})$.
    Let u_i be a uniformly random element from M_i.
    $S_i \leftarrow S_{i-1} + u_i$.
end for
Return $S_{k+k'} \setminus \mathcal{N}'$.

Observe that the extension increases the rank of the matroid $\mathcal{M}$ to $k + k'$ and preserves the γ-weak submodularity of the objective function f. Thus, by our assumption on the approximation ratio of Algorithm 1, $S_{k+k'}$ must provide an approximation ratio of at least $c(\gamma) - \varepsilon(k + k')$ for the problem of maximizing the extended objective function f subject to the extended matroid $\mathcal{M}$. One can verify that, together with the properties of $\mathcal{N}'$, this implies that $S_{k+k'} \setminus \mathcal{N}'$ provides an approximation ratio of at least $c(\gamma) - \varepsilon(k + k')$ for the problem of maximizing the original objective function f subject to the original matroid $\mathcal{M}$. Hence, Algorithm 2 has an approximation ratio of at least $c(\gamma) - \varepsilon(k + k') > c(\gamma) - [c(\gamma_I) - c']$ in general, which implies that for the specific instance I the approximation ratio of Algorithm 2 is strictly better than c'.

The final step required for getting the contradiction that we seek is to observe that Algorithms 1 and 2 share an identical output distribution for every given instance. Before we explain why that observation is true, let us note that it indeed implies a contradiction because Algorithm 1 has an approximation ratio of c' for the instance I, while Algorithm 2 has a strictly better approximation ratio for this instance. Thus, it remains to explain why Algorithms 1 and 2 have identical output distributions, which is what we do in the rest of this paragraph. Note that Algorithm 2 must add all the elements of $\mathcal{N}'$ to its solution set at some point because every base of the extended matroid $\mathcal{M}$ contains all of $\mathcal{N}'$. This means that we can view Algorithm 2 as a variant of Algorithm 1 that has k' more rounds, but must waste k' of its rounds on adding the elements of $\mathcal{N}'$ which do not affect anything and are removed at the end anyhow. Hence, the two algorithms share an identical behavior if we disregard the extra k' rounds that Algorithm 2 wastes on adding elements of $\mathcal{N}'$. $\square$