Appendix

In this appendix, we give the detailed proofs of the main theorems and lemmas.

A.1 Proof of Lemma 1

Proof. We just analyze S_i for a fixed $i \in [K]$. Let the times of removing operation be J. Denote by $B = \alpha\mathcal{R}$, $\mathcal{J} = \{t_r, r \in [J]\}$, $T_r = \{t_{r-1}+1, \ldots, t_r\}$ and $t_0 = 0$. For any $t \in T_r$, if $\nabla_{t,i} \neq 0$, $\neg\mathrm{con}(a(i))$ and $b_{t,i} = 1$, then $(\boldsymbol{x}_t, y_t)$ will be added into S_i. For simplicity, we define a new notation $\nu_{t,i}$ as follows,

$$\nu_{t,i} = \mathbb{I}_{y_t f_{t,i}(\boldsymbol{x}_t)<1} \cdot \mathbb{I}_{\neg\mathrm{con}(a(i))} \cdot b_{t,i}.$$

At the end of the t_r-th round, the following equation can be derived,

$$|S_i| = |S_i(t_{r-1}+1)| + \sum_{t=t_{r-1}+1}^{t_r} \nu_{t,i} = \frac{B}{K},$$

where $|S_i(t_{r-1}+1)|$ is defined the initial size of S_i.

Let $s_r = t_{r-1}+1$. Assuming that there is no budget. We will present an expected bound on $\sum_{t=s_r}^{\bar{t}} \nu_{t,i}$ for any $\bar{t} > s_r$. In the first epoch, $s_1 = 1$ and $|S_i(s_1)| = 0$. Taking expectation w.r.t. $b_{t,i}$ gives

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=s_1}^{\bar{t}} \nu_{t,i}\right] =& \sum_{t=s_1}^{\bar{t}} \frac{\|\nabla_{t,i} - \hat{\nabla}_{t,i}\|^2_{\mathcal{H}_i} \cdot \mathbb{I}_{\nabla_{t,i}\neq 0}}{\|\nabla_{t,i} - \hat{\nabla}_{t,i}\|^2_{\mathcal{H}_i} + \|\hat{\nabla}_{t,i}\|^2_{\mathcal{H}_i}} \\
\leq& \frac{2}{k_1} \underbrace{\left(1 + \sum_{t=2}^{\bar{t}} \left\| y_t\kappa_i(\boldsymbol{x}_t, \cdot) - \frac{\sum_{(\boldsymbol{x},y)\in V_t} y\kappa_i(\boldsymbol{x}, \cdot)}{|V_t|}\right\|^2_{\mathcal{H}_i}\right)}_{\tilde{\mathcal{A}}_{[s_1,\bar{t}],\kappa_i}} \\
=& \frac{2}{k_1} \tilde{\mathcal{A}}_{[s_1,\bar{t}],\kappa_i},
\end{aligned}$$

where we use the fact $\kappa_i(\boldsymbol{x}_t, \boldsymbol{x}_t) \geq k_1$. Let t_1 be the minimal $\bar{t}$ such that

$$\frac{2}{k_1} \tilde{\mathcal{A}}_{[s_1,t_1],\kappa_i} \geq \frac{B}{K}. \tag{A1}$$

The first epoch will end at t_1 in expectation. We define $\tilde{\mathcal{A}}_{T_1,\kappa_i} := \tilde{\mathcal{A}}_{[s_1,t_1],\kappa_i}$.

Next we consider $r \geq 2$. It must be $|S_i(s_r)| = \frac{B}{2K}$. Similar to $r = 1$, we can obtain

$$\mathbb{E}\left[\sum_{t=s_r}^{\bar{t}} \nu_{t,i}\right] \leq \frac{2}{k_1} \underbrace{\sum_{t=s_r}^{\bar{t}} \left\| y_t\kappa_i(\boldsymbol{x}_t, \cdot) - \frac{\sum_{(\boldsymbol{x},y)\in V_t} y\kappa_i(\boldsymbol{x}, \cdot)}{|V_t|}\right\|^2_{\mathcal{H}_i}}_{\tilde{\mathcal{A}}_{[s_r,\bar{t}],\kappa_i}} = \frac{2}{k_1} \tilde{\mathcal{A}}_{[s_r,\bar{t}],\kappa_i}.$$

Let t_r be the minimal $\bar{t}$ such that

$$\frac{2}{k_1} \tilde{\mathcal{A}}_{[s_r,\bar{t}],\kappa_i} \geq \frac{B}{2K}, \tag{A2}$$

Let $\tilde{\mathcal{A}}_{T_r,\kappa_i} = \tilde{\mathcal{A}}_{[s_r,\bar{t}],\kappa_i}$. Combining (A1) and (A2), and summing over $r = 1, \ldots, J$ yields

$$\begin{aligned}
\frac{B}{K} + \frac{B(J-1)}{2K} \leq& \frac{2}{k_1} \tilde{\mathcal{A}}_{T_1,\kappa_i} + \sum_{r=2}^{J} \frac{2}{k_1} \tilde{\mathcal{A}}_{T_r,\kappa_i} \\
\leq& \frac{2}{k_1} \underbrace{\sum_{t=s_1}^{T} \left\| y_t\kappa_i(\boldsymbol{x}_t, \cdot) - \frac{\sum_{(\boldsymbol{x},y)\in V_t} y\kappa_i(\boldsymbol{x}, \cdot)}{|V_t|}\right\|^2_{\mathcal{H}_i}}_{\tilde{\mathcal{A}}_{T,\kappa_i}} \\
\leq& \frac{2}{k_1} \tilde{\mathcal{A}}_{T,\kappa_i}.
\end{aligned}$$