Supplementary Materials: Tracking Regret Bounds for Online Submodular Optimization

Tatsuya Matsuoka
NEC Corporation
ta.matsuoka@nec.com

Shinji Ito
NEC Corporation
i-shinji@nec.com

Naoto Ohsaka
NEC Corporation
ohsaka@nec.com

A DESCRIPTION OF ALGORITHM FSF* MENTIONED IN COROLLARY 1

The description of the algorithm FSF* is given in Algorithm 4. We initialize parameters J, γ, β, $w_1 \in \mathbb{R}^J_{>0}$, $\gamma^{(j)}$, and $w_1^{(j)} \in \mathbb{R}^m_{>0}$ for $j = 1, 2, \ldots, J$ as stated in Step 1 in Algorithm 4. We run J copies of fixed share forecaster algorithms (Herbster and Warmuth, 1998) with different value of parameter $\gamma^{(j)}$. For each round $t = 1, 2, \ldots, T$, first we normalize $w_t \in \mathbb{R}^J_{>0}$ to calculate q_t, where w_{tj} represents the weight of the j-th copy, and normalize $w_t^{(j)} \in \mathbb{R}^m_{>0}$ to calculate $p_t^{(j)}$, which represents, for each $j \in [J]$, the weights of each action $i \in [m]$. Then, we compute and output $p_t \in \mathbb{R}^m_{>0}$, the sum of the vectors $p_t^{(j)}$ weighted by q_{tj} for $j = 1, 2, \ldots, J$. After outputting p_t, the algorithm receives feedback ℓ_{ti} for each $i \in [m]$, which represents a loss of choosing i in round t.

After receiving feedback ℓ_t, we update the weights for the next round. We use different value of parameter $\gamma^{(j)}$ for each $j \in [J]$. For each $j \in [J]$, we calculate $w_{t+1}^{(j)} \in \mathbb{R}^m_{>0}$, as stated in (6) and (7). We calculate $w_{t+1} \in \mathbb{R}^J_{>0}$ by multiplicative weight update with parameter γ and with loss for j-th action defined to be $\ell_t^\top p_t^{(j)}$, i.e., $w_{t+1,j}$ is calculated as in Step 9 in Algorithm 4 for each $j \in [J]$.

Algorithm 4 FSF*

Require: The number T of rounds and the number m of actions.

1: Set $J = \lceil \log T \rceil$, $\gamma = \sqrt{\frac{\log J}{T}}$, $\beta = \frac{1}{T}$ and initialize $w_t = (w_{t1}, w_{t2}, \ldots, w_{tJ})^\top$ by $w_{1j} = 1$ for $j = 1, 2, \ldots, J$. For $j = 1, 2, \ldots, J$, set $\gamma^{(j)} = \sqrt{\frac{\log(mT)}{2^{j-1}}}$ and initialize $w_t^{(j)} = (w_{t1}^{(j)}, w_{t2}^{(j)}, \ldots, w_{tm}^{(j)})^\top$ by $w_{1i}^{(j)} = 1$ for $i = 1, 2, \ldots, m$.
2: **for** $t = 1, 2, \ldots, T$ **do**
3:   Set $q_t = \frac{w_t}{\|w_t\|_1}$ and $p_t^{(j)} = \frac{w_t^{(j)}}{\|w_t^{(j)}\|_1}$ for $j = 1, 2, \ldots, J$.
4:   [**t-th output**] Compute $p_t = \sum_{j=1}^J q_{tj} p_t^{(j)}$ and output p_t.
5:   [**t-th input**] Get feedback of $\ell_t = (\ell_{t1}, \ell_{t2}, \ldots, \ell_{tm})^\top$.
6:   **for** $j = 1, 2, \ldots, J$ **do**
7:     Compute $v_{ti}^{(j)} = w_{ti}^{(j)} \exp(-\gamma^{(j)} \ell_{ti})$ for $i = 1, 2, \ldots, m$.
8:     Update $w_t^{(j)}$ by $w_{t+1,i}^{(j)} = \beta \frac{W_t^{(j)}}{m} + (1 - \beta) v_{ti}^{(j)}$ for $i = 1, 2, \ldots, m$ where $W_t^{(j)} = v_{t1}^{(j)} + \cdots + v_{tm}^{(j)}$.
9:     Update w_{tj} by $w_{t+1,j} = w_{tj} \exp(-\gamma \ell_t^\top p_t^{(j)})$.
10:   **end for**
11: **end for**

B OMITTED PROOFS

B.1 Proof of Corollary 1

If $m=1$, then $\sum_{t=1}^{T}\left(\ell_t^\top p_t - \ell_{ti_t^*}\right) = 0$ holds. If $T=1$ and $m \geq 2$, since $\ell_t \in [-1,1]^m$, $\sum_{t=1}^{T}\left(\ell_t^\top p_t - \ell_{ti_t^*}\right) \leq 2$. On the other hand, we have

$$8\sqrt{T((P'+1)\log(mT)+\log(1+\log T))} = 8\sqrt{(P'+1)\log 2} > 4.$$

Thus, we suppose $m \geq 2$ and $T \geq 2$. We have the following inequality (see, e.g., Cesa-Bianchi and Lugosi (2006)):

$$\sum_{t=1}^{T}(\ell_t^\top p_t - \ell_t^\top p_t^{(j)}) \leq \gamma T + \frac{1}{\gamma}\log J = 2\sqrt{T\log J}. \tag{16}$$

From Theorem 1,

$$\begin{aligned}\sum_{t=1}^{T}(\ell_t^\top p_t^{(j)} - \ell_{ti_t^*}) &\leq \gamma^{(j)}T + \frac{1}{\gamma^{(j)}}\left((2P'+1)\log m + \log\frac{1}{\beta^{P'}(1-\beta)^{T-P'-1}}\right)\\ &\leq \gamma^{(j)}T + \frac{2}{\gamma^{(j)}}(P'+1)\log(mT) \end{aligned}\tag{17}$$

holds for each $j \in [J]$, where we utilize

$$\begin{aligned}\log\frac{1}{\beta^{P'}(1-\beta)^{T-P'-1}} &= P'\log T + (T-P'-1)\log\left(1+\frac{1}{T-1}\right)\\ &\leq P'\log T + \frac{T-P'-1}{T-1}\\ &\leq P'\log T + 1\end{aligned}$$

and $\log 2 > \frac{1}{2}$. By the definition of $\{\gamma^{(j)}\}_{j=1}^{J}$, there exists $j \in [J]$ such that

$$\frac{\gamma^{(j)}}{2} \leq \sqrt{\frac{(P'+1)\log(mT)}{T}} \leq \gamma^{(j)} \left(\iff \sqrt{\frac{(P'+1)\log(mT)}{T}} \leq \gamma^{(j)} \leq 2\sqrt{\frac{(P'+1)\log(mT)}{T}}\right)$$

holds. For such j, by (17) we have

$$\sum_{t=1}^{T}(\ell_t^\top p_t^{(j)} - \ell_{ti_t^*}) \leq 4\sqrt{T(P'+1)\log(mT)}. \tag{18}$$

Therefore, by (16) and (18),

$$\begin{aligned}\sum_{t=1}^{T}(\ell_t^\top p_t - \ell_{ti_t^*}) &\leq 2\sqrt{T\log J} + 4\sqrt{T(P'+1)\log(mT)}\\ &= 2\left(\sqrt{T\log J} + \sqrt{4T(P'+1)\log(mT)}\right)\\ &\leq 2\left(2\sqrt{T\log(1+\log T)} + \sqrt{4T(P'+1)\log(mT)}\right)\\ &\leq 8\sqrt{T\left((P'+1)\log(mT)+\log(1+\log T)\right)}\end{aligned}$$

holds, where the second inequality holds since for $T \geq 2$,

$$\log J = \log(\lceil \log_2 T\rceil) \leq \log(1+\log_2 T) \leq 2\log(1+\log T)$$

holds.

B.2 Proof of Lemma 2

The procedure given by (10) can be expressed by

$$y_{t+1} = x_t - \eta g_t, \quad x_{t+1} \in \operatorname*{argmin}_{x \in [0,1]^n} \|x - y_{t+1}\|_2^2. \tag{19}$$

We have

$$\|x_{t+1} - x_t^*\|_2^2 \le \|y_{t+1} - x_t^*\|_2^2 = \|x_t - x_t^* - \eta g_t\|_2^2 = \|x_t - x_t^*\|_2^2 + \eta^2 \|g_t\|_2^2 - 2\eta g_t^\top (x_t - x_t^*),$$

where the inequality follows from the second part of (19) and the generalized Pythagorean theorem, and the first equality follows from the first part of (19). By dividing both sides with 2η, we obtain

$$g_t^\top (x_t - x_t^*) \le \frac{\eta}{2}\|g_t\|_2^2 + \frac{1}{2\eta}\left(\|x_t - x_t^*\|_2^2 - \|x_{t+1} - x_t^*\|_2^2\right).$$

By taking the sum for $t \in [T]$, we obtain

$$\begin{aligned}
\sum_{t=1}^T g_t^\top x_t - \sum_{t=1}^T g_t^\top x_t^* &\le \sum_{t=1}^T \left(\frac{\eta}{2}\|g_t\|_2^2 + \frac{1}{2\eta}\left(\|x_t - x_t^*\|_2^2 - \|x_{t+1} - x_t^*\|_2^2\right)\right) \\
&\le \frac{\eta}{2}\sum_{t=1}^T \|g_t\|_2^2 + \frac{\|x_1 - x_1^*\|_2^2}{2\eta} + \frac{1}{2\eta}\sum_{t=1}^{T-1}\left(\|x_{t+1} - x_{t+1}^*\|_2^2 - \|x_{t+1} - x_t^*\|_2^2\right) \\
&\le \frac{\eta}{2}\sum_{t=1}^T \|g_t\|_2^2 + \frac{n}{2\eta} + \frac{1}{2\eta}\sum_{t=1}^{T-1}(2x_{t+1} - x_{t+1}^* - x_t^*)^\top (x_t^* - x_{t+1}^*) \\
&\le \frac{\eta}{2}\sum_{t=1}^T \|g_t\|_2^2 + \frac{n}{2\eta} + \frac{1}{\eta}\sum_{t=1}^{T-1}\|x_t^* - x_{t+1}^*\|_1,
\end{aligned}$$

where the last inequality follows from $\|2x_{t+1} - x_{t+1}^* - x_t^*\|_\infty \le 2$.

B.3 Proof of Lemma 3

Define δ_s by

$$\delta_s = \sum_{t=1}^T \left(f_t(X_t^*) - f_t(X_{ts})\right) \tag{20}$$

for $s = 0, 1, \ldots, k$. Then, for an arbitrary fixed $s \in \{0, 1, \ldots, k-1\}$, we have

$$\begin{aligned}
\delta_s &= \sum_{t=1}^T \left(f_t(X_t^*) - f_t(X_{ts})\right) \\
&\le \sum_{t=1}^T \sum_{i \in X_t^*} \left(f_t(X_{ts} \cup \{i_j^*\}) - f_t(X_{ts})\right) \\
&= \sum_{t=1}^T \left(- \sum_{i \in X_t^*} \ell_{ti}^{(s+1)}\right) \\
&= -k \sum_{t=1}^T \ell_{ti_{t,s+1}}^{(s+1)} + B_T^{(s+1)} \\
&= k\left(\delta_s - \delta_{s+1}\right) + B_T^{(s+1)},
\end{aligned}$$

where the inequality follows from submodularity of f_t, the second equality follows from the definition (12), (13) of $\ell_{ti}^{(s+1)}$, the third equality follows from the definition (14) of $B_T^{(s+1)}$, and the fourth equality follows from (12), (13), and (20).

Thus,

$$\delta_{s+1} \leq \left(1 - \frac{1}{k}\right)\delta_s + \frac{1}{k}B_T^{(s+1)}$$

holds for each $s \in \{0, 1, \ldots, k\}$, and hence, we have

$$\delta_{s+1} \leq \left(1 - \frac{1}{k}\right)^{s+1}\delta_0 + \frac{1}{k}\sum_{j=1}^{s+1}\left(1 - \frac{1}{k}\right)^{s+1-j} B_T^{(j)}.$$

Therefore,

$$\delta_k - \left(1 - \frac{1}{k}\right)^k \delta_0 \leq \frac{1}{k}\sum_{s=1}^{k}\left(1 - \frac{1}{k}\right)^{k-s} B_T^{(s)} \tag{21}$$

holds. On the other hand, we have

$$\begin{aligned}
\delta_k - \left(1 - \frac{1}{k}\right)^k \delta_0 &= \left(1 - \left(1 - \frac{1}{k}\right)^k\right)\sum_{t=1}^{T} f_t(X_t^*) - \sum_{t=1}^{T} f_t(X_{tk}) + \left(1 - \frac{1}{k}\right)^k \sum_{t=1}^{T} f_t(X_{t0}) \\
&\geq \left(1 - \frac{1}{\mathrm{e}}\right)\sum_{t=1}^{T} f_t(X_t^*) - \sum_{t=1}^{T} f_t(X_t) \\
&= R_T^{(1-1/\mathrm{e})}(\{X_t^*\}_{t=1}^T), \qquad (22)
\end{aligned}$$

where the first equality follows from the definition (20) of δ_s, the inequality follows from the nonnegativity of f_t and $\left(1 - \frac{1}{k}\right)^k \leq \frac{1}{\mathrm{e}}$, and the second equality follows from the definition (2) of R_T^α. By (21) and (22), we obtain

$$R_T^{(1-1/\mathrm{e})}(\{X_t^*\}_{t=1}^T) \leq \frac{1}{k}\sum_{s=1}^{k}\left(1 - \frac{1}{k}\right)^{k-s} B_T^{(s)}.$$

B.4 Proof of Lemma 4

For proving Lemma 4, we first state Lemmas 5 and 6.

Lemma 5 (Lemma 2.1 of Buchbinder et al. (2015))**.** *For each $t \in [T]$ and for each $s \in [n]$, $a_{ts} + b_{ts} \geq 0$ holds.*

Lemma 6. *It holds that*

$$\mathbb{E}\left[R_T^{(1/2)}(\{X_t^*\}_{t=1}^T)\right] \leq \mathbb{E}\left[\sum_{t=1}^{T}\left(\frac{1}{2}\sum_{s \in X_t^*}(1 - q_t^{(s)})a_{ts} + \frac{1}{2}\sum_{s \in [n]\setminus X_t^*} q_t^{(s)} b_{ts} - \frac{1}{4}\sum_{s=1}^{n}(q_{ts}a_{ts} + (1 - q_{ts})b_{ts})\right)\right].$$

Proof of Lemma 6. Let $Z_{ts} = (X_t^* \cup X_{ts}) \cap Y_{ts}$. Then,

$$f_t(X_t^*) = \sum_{s=1}^{n}\left(f_t(Z_{t,s-1}) - f_t(Z_{ts})\right) + f_t(X_t) \tag{23}$$

holds.

- Suppose that $s \in X_t^*$. If $X_{ts} = X_{t,s-1} \cup \{s\}$ (with probability $q_t^{(s)}$), we have

  $$f_t(Z_{t,s-1}) = f_t(Z_{ts}).$$

  Otherwise (with probability $1 - q_t^{(s)}$), by submodularity,

  $$f_t(Z_{t,s-1}) - f_t(Z_{ts}) \leq f_t(X_{t,s-1} \cup \{s\}) - f_t(X_{t,s-1}) = a_{ts}$$

  holds since $Z_{t,s-1} = (X_{t,s-1} \cup \{s\}) \cup Z_{ts}$ and $X_{t,s-1} = (X_{t,s-1} \cup \{s\}) \cap Z_{ts}$. Thus, we obtain

  $$\mathbb{E}\left[f_t(Z_{t,s-1}) - f_t(Z_{ts})\right] \leq \mathbb{E}\left[(1 - q_t^{(s)})a_{ts}\right]. \tag{24}$$

- Suppose that $s \notin X_t^*$. If $X_{ts} = X_{t,s-1} \cup \{s\}$ (with probability $q_t^{(s)}$), by submodularity,

$$f_t(Z_{t,s-1}) - f_t(Z_{ts}) \leq f_t(Y_{t,s-1} \setminus \{s\}) - f_t(Y_{t,s-1}) = b_{ts}.$$

holds since $Z_{t,s-1} = (Y_{t,s-1} \setminus \{s\}) \cup Z_{ts}$ and $Y_{t,s-1} = (Y_{t,s-1} \setminus \{s\}) \cap Z_{ts}$. Otherwise (with probability $1 - q_t^{(s)}$), we have

$$f_t(Z_{t,s-1}) = f_t(Z_{ts}).$$

Thus, we obtain

$$\mathbb{E}\left[f_t(Z_{t,s-1}) - f_t(Z_{ts})\right] \leq \mathbb{E}\left[q_t^{(s)} b_{ts}\right]. \tag{25}$$

Therefore, by combining (23), (24), and (25), we obtain

$$\mathbb{E}\left[f_t(X_t^*) - f_t(X_t)\right] \leq \mathbb{E}\left[\sum_{s \in X_t^*} (1 - q_t^{(s)}) a_{ts} + \sum_{s \in [n] \setminus X_t^*} q_t^{(s)} b_{ts}\right]. \tag{26}$$

Here, $f_t(X_t)$ can be decomposed as

$$f_t(X_t) = f_t(X_{tn}) = \sum_{s=1}^{n} \left(f_t(X_{ts}) - f_t(X_{t,s-1})\right) + f_t(X_{t0}). \tag{27}$$

From Step 6 of Algorithm 3, we have

$$\mathbb{E}\left[f_t(X_{ts}) - f_t(X_{t,s-1})\right] = \mathbb{E}\left[q_t^{(s)} (f_t(X_{t,s-1} \cup \{i\}) - f_t(X_{t,s-1}))\right] = \mathbb{E}\left[q_t^{(s)} a_{ts}\right].$$

By the above equation and (27), we have

$$\mathbb{E}\left[f_t(X_t) - f_t(X_{t0})\right] = \mathbb{E}\left[\sum_{s=1}^{n} q_t^{(s)} a_{ts}\right]. \tag{28}$$

Similarly,

$$\mathbb{E}\left[f_t(Y_t) - f_t(Y_{t0})\right] = \mathbb{E}\left[\sum_{s=1}^{n} (1 - q_t^{(s)}) b_{ts}\right] \tag{29}$$

holds. By (26), (28), and (29), we obtain

$$\begin{aligned}
\mathbb{E}\left[R_T^{(1/2)}(\{X_t^*\}_{t=1}^T)\right] &= \mathbb{E}\left[\frac{1}{2}\sum_{t=1}^{T} f_t(X_t^*) - \sum_{t=1}^{T} f_t(X_t)\right] \\
&\leq \frac{1}{2}\mathbb{E}\left[f_t(X_t^*) - f_t(X_t)\right] - \frac{1}{4}\mathbb{E}\left[f_t(X_t) - f_t(X_{t0})\right] - \frac{1}{4}\mathbb{E}\left[f_t(Y_t) - f_t(Y_{t0})\right] \\
&\leq \mathbb{E}\left[\sum_{t=1}^{T}\left(\frac{1}{2}\sum_{s \in X_t^*}(1 - q_t^{(s)}) a_{ts} + \frac{1}{2}\sum_{s \in [n] \setminus X_t^*} q_t^{(s)} b_{ts} - \frac{1}{4}\sum_{s=1}^{n}(q_{ts} a_{ts} + (1 - q_{ts}) b_{ts})\right)\right].
\end{aligned}$$

□

It holds from the definition of $\ell_t^{(s)}$ that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T}\left(\frac{1}{2}\sum_{s \in X_t^*}(1 - q_t^{(s)}) a_{ts} + \frac{1}{2}\sum_{s \in [n] \setminus X_t^*} q_t^{(s)} b_{ts} - \frac{1}{4}\sum_{s=1}^{n}(q_{ts} a_{ts} + (1 - q_{ts}) b_{ts})\right)\right] \\
&= \frac{1}{2}\sum_{s=1}^{n}\mathbb{E}\left[\sum_{t \in [T]} \ell_t^{(s)\top} p_t^{(s)} - \sum_{t \in [T]: s \in X_t^*} \ell_{t1}^{(s)} - \sum_{t \in [T]: s \notin X_t^*} \ell_{t2}^{(s)}\right] - \mathbb{E}\left[\sum_{s=1}^{n}\sum_{t=1}^{T}\left(\frac{1}{2}\ell_t^{(s)\top} p_t^{(s)} + \frac{1}{4}(q_t^{(s)} a_{ts} + (1 - q_t^{(s)}) b_{ts})\right)\right].
\end{aligned} \tag{30}$$

Since $p_{t1}^{(s)} = \frac{1}{2}(4q_t^{(s)} - 1)$ and $p_{t2}^{(s)} = \frac{1}{2}(3 - 4q_t^{(s)})$, by Step 5 of Algorithm 3, it holds that

$$\frac{1}{2}\ell_t^{(s)\top} p_t^{(s)} + \frac{1}{4}(q_t^{(s)} a_{ts} + (1 - q_t^{(s)}) b_{ts}) = \frac{1}{4}\left(-(1 - q_t^{(s)}) a_{ts}(4q_t^{(s)} - 1) - q_t^{(s)} b_{ts}(3 - 4q_t^{(s)}) + q_t^{(s)} a_{ts} + (1 - q_t^{(s)}) b_{ts}\right)$$
$$= \frac{a_{ts} + b_{ts}}{4}(2q_t^{(s)} - 1)^2 \geq 0.$$

By the above inequality, (30), and Lemma 6, we obtain Lemma 4.

B.5 Proof of Theorem 5

To prove Theorem 5, we use the following lemma.

Lemma 7. *Let $X = B_1 + B_2 + \cdots + B_m$, where each $B_i \in \{-1, 1\}$ follows a Bernoulli distribution of parameter $1/2$, independent and identically distributed for $i \in [m]$. We then have $\mathbb{E}[|X|] \geq \sqrt{m/3}$.*

Proof of Lemma 7. By Hölder's inequality ($\mathbb{E}[|A||B|] \leq (\mathbb{E}[|A|^p])^{1/p}(\mathbb{E}[|B|^q])^{1/q}$ ($p > 0$, $q > 0$, $1/p + 1/q = 1$)) with $A = |X|^{4/3}$, $B = |X|^{2/3}$, $p = 3$, and $q = 3/2$,

$$\mathbb{E}[|X|^2] = \mathbb{E}[AB] = \mathbb{E}[|A||B|] \leq (\mathbb{E}[|A|^3])^{1/3}(\mathbb{E}[|B|^{3/2}])^{2/3} = (\mathbb{E}[|X|^4])^{1/3}(\mathbb{E}[|X|])^{2/3}$$

holds, which implies

$$\mathbb{E}[|X|] \geq \frac{(\mathbb{E}[|X|^2])^{3/2}}{(\mathbb{E}[|X|^4])^{1/2}}. \tag{31}$$

Then, we have

$$\mathbb{E}[|X|^2] = \mathbb{E}[X^2] = \mathbb{E}\left[\sum_{i=1}^m B_i^2 + \sum_{i=1}^m \sum_{j \in [m] \setminus \{i\}} B_i B_j\right] = \mathbb{E}\left[\sum_{i=1}^m B_i^2\right] = m, \tag{32}$$

where the third equality follows from $\mathbb{E}[B_j \mid B_i] = 0$ for $j \neq i$ and the fourth equality follows from $B_i \in \{-1, 1\}$ for each $i \in [m]$. Similarly, by utilizing $\mathbb{E}[B_j \mid B_i] = 0$ for $j \neq i$ and $B_i \in \{-1, 1\}$ for each $i \in [m]$, we obtain

$$\mathbb{E}[|X|^4] = \mathbb{E}[X^4] = \mathbb{E}\left[\sum_{i=1}^m B_i^4 + \frac{\binom{4}{2}}{2}\sum_{i=1}^m \sum_{j \in [m] \setminus \{i\}} B_i^2 B_j^2\right] = m + 3m(m-1) \leq 3m^2. \tag{33}$$

Therefore, by (31), (32), and (33), we have

$$\mathbb{E}[|X|] \geq \frac{m^{3/2}}{(3m^2)^{1/2}} = \sqrt{\frac{m}{3}}.$$

$\square$

Let $B_t \in \{-1, 1\}$ be a random variable which depends on a Bernoulli distribution of parameter $1/2$, independent and identically distributed for $t \in [T]$. Let $n' = \min\{n + P, T\}$, $T' = \lfloor T/n' \rfloor$, and $n'' = \min\{n', n\}$. For each $i \in [n'']$, define $f_t \colon 2^{[n]} \to [0, 1]$ by

$$f_t(X) = \begin{cases} (1 - B_t)/2 & (i \in X) \\ (1 + B_t)/2 & (i \notin X) \end{cases}$$

for $t = (i-1)T' + 1, (i-1)T' + 2, \ldots, iT'$. For $t = n''T' + 1, n''T' + 2, \ldots, n'T'$, let

$$f_t(X) = \begin{cases} (1 - B_t)/2 & (1 \in X) \\ (1 + B_t)/2 & (1 \notin X) \end{cases}.$$

Define $f_t(X) = 0$ for $t > n'T'$. Let $\{X_t\}_{t=1}^T$ denote a (random) output sequence of an arbitrary fixed algorithm. Then, X_t is independent of f_t since X_t depends only on $f_1, f_2, \ldots, f_{t-1}$ which are independent of f_t. By combining this fact and $\mathbb{E}[f_t(X)] = 1/2$ for all $X \subseteq [n]$, we obtain

$$\mathbb{E}[f_t(X_t)] = 1/2 \tag{34}$$

for $t \in [n'T']$, where the expectation is taken for X_t and B_t. Define $X^* \subseteq [n]$ so that

$$i \in X^* \iff \sum_{t=(i-1)T'+1}^{iT} B_t \geq 0$$

for all $i \in [n'']$.

Set $X_t^* = X^*$ for $t \in [n''T']$. For each $j \in [P]$, define

$$X^{(j)} = \begin{cases} X^* \cup \{1\} & \left(\sum_{t=(n''+j-1)T'+1}^{(n''+j)T'} B_t \geq 0\right) \\ X^* \setminus \{1\} & \text{(otherwise)} \end{cases}$$

and set $X_t^* = X^{(j)}$ for all $t = (n''+j-1)T'+1, (n''+j-1)T'+2, \ldots, (n''+j)T'$. Set $X_t^* = X_{(n''+P)T'}$ for $t > (n''+P)T'$. Then, we have

$$\sum_{t=1}^{T-1} |X_t^* \triangle X_{t+1}^*| \leq P$$

and

$$\sum_{t=1}^{T} f_t(X_t^*) = \sum_{j=1}^{n'} \sum_{t=(j-1)T'+1}^{jT'} f_t(X_t^*) = \sum_{j=1}^{n'} \left(\frac{T'}{2} - \frac{1}{2} \sum_{t=(j-1)T'+1}^{jT'} |B_t| \right). \tag{35}$$

Then, by combining (34) and (35), we obtain

$$\begin{aligned} \mathbb{E}[R_T(\{X_t^*\}_{t=1}^T)] &= \frac{1}{2} \sum_{j=1}^{n'} \mathbb{E}\left[\sum_{t=(j-1)T'+1}^{jT'} |B_t| \right] \geq \frac{1}{2} n' \sqrt{\frac{T'}{3}} = \frac{n'}{2\sqrt{3}} \sqrt{\left\lfloor \frac{T}{n'} \right\rfloor} \\ &\geq \frac{n'}{2\sqrt{3}} \sqrt{\frac{T}{2n'}} \geq \frac{1}{2\sqrt{6}} \sqrt{n'T} = \frac{1}{2\sqrt{6}} \sqrt{T \min\{T, n+P\}}, \end{aligned}$$

where the expectation is taken over the randomness of $\{B_t\}_{t=1}^T$ and $\{X_t\}_{t=1}^T$. Therefore, there exists a realization of $\{B_t\}_{t=1}^T$ for which this inequality holds.

References

N. Buchbinder, M. Feldman, J. Seffi, and R. Schwartz. A tight linear time (1/2)-approximation for unconstrained submodular maximization. *SIAM Journal on Computing*, 44(5):1384–1402, 2015.

N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.

M. Herbster and M. K. Warmuth. Tracking the best expert. *Machine Learning*, 32(2):151–178, 1998.