A Useful theorems

Here are some basic facts from the literature that we will use:

Theorem A.1 ((Audibert et al., 2009), speicalized to Berounlli random variables)**.** *Consider N independently and identically distributed Bernoulli random variables $Y_1, \ldots, Y_N \in \{0,1\}$, which have the common mean $m = \mathbb{E}[Y_1]$. In addition, consider their sample mean $\hat{\xi}$ and their sample variance $\hat{V}$:*

$$\hat{\xi} = \frac{1}{N}\sum_{i=1}^{N} Y_i, \quad \hat{V} = \frac{1}{N}\sum_{i=1}^{N}(Y_i - \hat{\xi})^2 = \hat{\xi}(1-\hat{\xi}).$$

For any $\delta \in (0,1)$, the following inequality holds:

$$\Pr\left(\left|\hat{\xi} - m\right| \le \sqrt{\frac{2\hat{V}\log(1/\delta)}{N}} + \frac{3\log(1/\delta)}{N}\right) \ge 1 - 3\delta.$$

Theorem A.2 ((Abramowitz and Stegun, 1964))**.** *Let $Z \sim \mathcal{N}(\mu, \sigma^2)$. For any $z \ge 0$, the following inequalities hold:*

$$\frac{1}{4\sqrt{\pi}}\exp\left(-\frac{7z^2}{2}\right) \le \Pr\left(|Z-\mu| > z\sigma\right) \le \frac{1}{2}\exp\left(-\frac{z^2}{2}\right).$$

B Proofs of main results

In this Section, we provide proofs of Lemmas 4.2, 4.4, as well as Lemmas 4.1, B.1.

B.1 Proof of Lemma 4.1

Lemma 4.1. *For each $t \in [T]$, $H_t \in \mathcal{E}_{\hat{\mu},t}$, we have*

$$\Pr\left[\mathcal{E}_{\hat{\mu},t}\right] \ge 1 - \frac{3L}{(t+1)^3}, \quad \Pr\left[\mathcal{E}_{\theta,t} | H_t\right] \ge 1 - \frac{1}{2(t+1)^2}.$$

Proof. **Bounding probability of event $\mathcal{E}_{\hat{\mu},t}$** : We first consider a fixed non-negative integer N and a fixed item i. Let $Y_1, \ldots, Y_N$ be i.i.d. Bernoulli random variables, with the common mean $w(i)$. Denote $\hat{\xi}_N = \sum_{i=1}^{N} Y_i/N$ as the sample mean, and $\hat{V}_N = \hat{\xi}_N(1-\hat{\xi}_N)$ as the empirical variance. By applying Theorem A.1 with $\delta = 1/(t+1)^4$, we have

$$\Pr\left(\left|\hat{\xi}_N - w(i)\right| > \sqrt{\frac{8\hat{V}_N\log(t+1)}{N}} + \frac{12\log(t+1)}{N}\right) \le \frac{3}{(t+1)^4}. \tag{B.1}$$

By an abuse of notation, let $\hat{\xi}_N = 0$ if $N = 0$. Inequality (B.1) implies the following concentration bound when N is non-negative:

$$\Pr\left(\left|\hat{\xi}_N - w(i)\right| > \sqrt{\frac{16\hat{V}_N\log(t+1)}{N+1}} + \frac{24\log(t+1)}{N+1}\right) \le \frac{3}{(t+1)^4}. \tag{B.2}$$

Subsequently, we can establish the concentration property of $\hat{\mu}_t(i)$ by a union bound of $N_t(i)$ over $\{0, 1, \ldots, t-1\}$:

$$\begin{aligned}
&\Pr\left(|\hat{\mu}_t(i) - w(i)| > \sqrt{\frac{16\hat{\nu}_t(i)\log(t+1)}{N_t(i)+1}} + \frac{24\log(t+1)}{N_t(i)+1}\right) \\
&\le \Pr\left(\left|\hat{\xi}_N - w(i)\right| > \sqrt{\frac{16\hat{V}_N\log(t+1)}{N+1}} + \frac{24\log(t+1)}{N+1} \text{ for some } N \in \{0, 1, \ldots t-1\}\right) \\
&\le \frac{3}{(t+1)^3}.
\end{aligned}$$

Finally, taking union bound over all items $i \in L$, we know that event $\mathcal{E}_{\hat{\mu},t}$ holds true with probability at least $1 - 3L/(t+1)^3$.

Bounding probability of event $\mathcal{E}_{\theta,t}$, conditioned on event $\mathcal{E}_{\hat{\mu},t}$: Consider an observation trajectory H_t satisfying event $\mathcal{E}_{\hat{\mu},t}$. By the definition of the Thompson sample $\theta_t(i)$ (see Line 7 in Algorithm 1), we have

$$\begin{aligned}
&\Pr\left(|\theta_t(i) - \hat{\mu}_t(i)| > h_t(i) \text{ for all } i \in L | H_{\hat{\mu},t}\right) \\
&= \Pr_{Z_t}\left(\left|Z_t \cdot \max\left\{\sqrt{\frac{\hat{\nu}_t(i)\log(t+1)}{N_t(i)+1}}, \frac{\log(t+1)}{N_t(i)+1}\right\}\right| > \right. \\
&\qquad\qquad \left. \sqrt{\log(t+1)}\left[\sqrt{\frac{16\hat{\nu}_t(i)\log(t+1)}{N_t(i)+1}} + \frac{24\log(t+1)}{N_t(i)+1}\right] \text{ for all } i \in [L] \;\middle|\; \hat{\mu}_t(i), N_t(i)\right) \\
&\le \Pr\left(\left|Z_t \cdot \max\left\{\sqrt{\frac{\hat{\nu}_t(i)\log(t+1)}{N_t(i)+1}}, \frac{\log(t+1)}{N_t(i)+1}\right\}\right| > \right. \\
&\qquad\qquad \left. \sqrt{16\log(t+1)}\max\left\{\sqrt{\frac{\hat{\nu}_t(i)\log(t+1)}{N_t(i)+1}}, \frac{\log(t+1)}{N_t(i)+1}\right\} \text{ for all } i \in [L] \;\middle|\; \hat{\mu}_t(i), N_t(i)\right) \\
&\le \frac{1}{2}\exp\left[-8\log(t+1)\right] \le \frac{1}{2(t+1)^3}.
\end{aligned} \tag{B.3}$$

The inequality in (B.3) is by the concentration property of a Gaussian random variable, see Theorem A.2. Altogether, the lemma is proved. □

B.2 Proof of Lemma 4.2

Proof. To start, we denote the shorthand $\theta_t(i)^+ = \max\{\theta_t(i), 0\}$. We demonstrate that, if events $\mathcal{E}_{\hat{\mu},t}, \mathcal{E}_{\theta,t}$ and inequality (4.2) hold, then for all $\bar{S} = (\bar{i}_1, \ldots, \bar{i}_K) \in \bar{\mathcal{S}}_t$ we have:

$$\sum_{k=1}^{K}\left[\prod_{j=1}^{k-1}(1 - w(\bar{i}_j))\right] \cdot \theta_t(\bar{i}_k)^+ \overset{(\ddagger)}{<} \sum_{k=1}^{K}\left[\prod_{j=1}^{k-1}(1 - w(j))\right] \cdot \theta_t(k)^+ \overset{(\dagger)}{\le} \sum_{k=1}^{K}\left[\prod_{j=1}^{k-1}(1 - w(i_j^t))\right] \cdot \theta_t(i_k^t)^+, \tag{B.4}$$

where we recall that $S_t = (i_1^t, \ldots, i_K^t)$ in an optimal arm for $\boldsymbol{\theta}_t$, and $\theta_t(i_1^t) \ge \theta_t(i_2^t) \ge \ldots \ge \theta_t(i_K^t) \ge \max_{i \in [L] \setminus \{i_1^t, \ldots, i_K^t\}} \theta_t(i)$. The inequalities in (B.4) clearly implies that $S_t \in \mathcal{S}_t$. To justifies these inequalities, we proceed as follows:

Showing (†): This inequality is true even without requiring events $\mathcal{E}_{\hat{\mu},t}, \mathcal{E}_{\theta,t}$ and inequality (4.2) to be true. Indeed, we argue the following:

$$\begin{aligned}
\sum_{k=1}^{K}\left[\prod_{j=1}^{k-1}(1 - w(j))\right] \cdot \theta_t(k)^+ &\le \sum_{k=1}^{K}\left[\prod_{j=1}^{k-1}(1 - w(j))\right] \cdot \theta_t(i_k^t)^+ &\quad (\text{B.5}) \\
&\le \sum_{k=1}^{K}\left[\prod_{j=1}^{k-1}(1 - w(i_j^t))\right] \cdot \theta_t(i_k^t)^+. &\quad (\text{B.6})
\end{aligned}$$

To justify inequality (B.5), consider function $f : \pi_K(L) \to \mathbb{R}$ defined as

$$f\left((i_k)_{k=1}^K\right) := \sum_{k=1}^{K}\left[\prod_{j=1}^{k-1}(1 - w(j))\right] \cdot \theta_t(i_k)^+.$$

We assert that $S_t \in \mathrm{argmax}_{S \in \pi_K(L)} f(S)$. The assertion can be justified by the following two properties. First, by the choice of S_t, we know that $\theta_t(i_1^t)^+ \ge \theta_t(i_2^t)^+ \ge \ldots \ge \theta_t(i_K^t)^+ \ge \max_{i \in [L] \setminus S_t} \theta_t(i)^+$. Second, the linear coefficients in the function f are monotonic and non-negative, in the sense that $1 \ge 1 - w(1) \ge (1 - w(1))(1 - w(2)) \ge \ldots \ge \prod_{k=1}^{K-1}(1 - w(k)) \ge 0$. Altogether, we have $f(S_t) \ge f(S^*)$, hence inequality (B.5) is shown.

Next, inequality (B.6) clearly holds, since for each $k \in [K]$ we know that $\theta_t(i^t_k)^+ \geq 0$, and $\prod_{j=1}^{k-1}(1 - w(j)) \leq \prod_{j=1}^{k-1}(1 - w(i^t_j))$. The latter is due to the fact that $1 \geq w(1) \geq w(2) \geq \ldots \geq w(K) \geq \max_{i\in[L]\setminus[K]} w(i)$. Altogether, inequality (†) is established.

Showing (‡): The demonstration crucially hinges on events $\mathcal{E}_{\hat{\mu},t}, \mathcal{E}_{\theta,t}$ and inequality (4.2) being held true. For any $\bar{S} = (\bar{i}_1, \ldots, \bar{i}_K) \in \bar{\mathcal{S}}_t$, we have

$$\begin{aligned}
\sum_{k=1}^{K}\left[\prod_{j=1}^{k-1}(1 - w(\bar{i}_j))\right]\theta_t(\bar{i}_k)^+ &\leq \sum_{k=1}^{K}\left[\prod_{j=1}^{k-1}(1 - w(\bar{i}_j))\right](w(\bar{i}_k) + g_t(\bar{i}_k) + h_t(\bar{i}_k)) &\text{(B.7)}\\
&< \left\{\sum_{k=1}^{K}\left[\prod_{j=1}^{k-1}(1 - w(\bar{i}_j))\right]w(\bar{i}_k)\right\} + r(S^*|\boldsymbol{w}) - r(\bar{S}|\boldsymbol{w}) &\text{(B.8)}\\
&= r(S^*|\boldsymbol{w}) = \sum_{k=1}^{K}\left[\prod_{j=1}^{k-1}(1 - w(j))\right]w(k)\\
&\leq \sum_{k=1}^{K}\left[\prod_{j=1}^{k-1}(1 - w(j))\right]\theta_t(k) \leq \sum_{k=1}^{K}\left[\prod_{j=1}^{k-1}(1 - w(j))\right]\theta_t(k)^+. &\text{(B.9)}
\end{aligned}$$

Inequality (B.7) is by the assumption that events $\mathcal{E}_{\hat{\mu},t}, \mathcal{E}_{\theta,t}$ are true, which means that for all $i \in [L]$ we have $\theta_t(i)^+ \leq \mu(i) + g_t(i) + h_t(i)$. Inequality (B.8) is by the fact that $S \in \bar{\mathcal{S}}_t$. Inequality (B.9) is by our assumption that inequality (4.2) holds.

Altogether, the inequalities (†, ‡) in (B.4) are shown, and the Lemma is established. □

B.3 Proof of Lemma 4.4

Lemma 4.4. *Let c be an absolute constant such that Lemma 4.3 holds true. Consider a time step t that satisfies $c - 1/(t+1)^3 > 0$. Conditional on an arbitrary but fixed historical observation $H_t \in \mathcal{H}_{\hat{\mu},t}$, we have*

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}_t}[r(S^*|\boldsymbol{w}) - r(S_t|\boldsymbol{w})|H_t]\\
&\quad\leq \left(1 + \frac{4}{c}\right)\mathbb{E}_{\boldsymbol{\theta}_t}\left[F(S_t, t) \;\middle|\; H_t\right] + \frac{L}{2(t+1)^2}.
\end{aligned}$$

The proof of Lemma 4.4 crucially uses the following lemma on the expression of the difference in expected reward between two arms:

Lemma B.1. *[Implied by Zong et al. (2016)] Let $S = (i_1, \ldots, i_K)$, $S' = (i'_1, \ldots, i'_K)$ be two arbitrary ordered K-subsets of $[L]$. For any $\boldsymbol{w}, \boldsymbol{w}' \in \mathbb{R}^L$, the following equalities holds:*

$$\begin{aligned}
r(S|\boldsymbol{w}) - r(S'|\boldsymbol{w}') &= \sum_{k=1}^{K}\left[\prod_{j=1}^{k-1}(1 - w(i_j))\right]\cdot(w(i_k) - w'(i'_k))\cdot\left[\prod_{j=k+1}^{K}(1 - w'(i'_j))\right]\\
&= \sum_{k=1}^{K}\left[\prod_{j=1}^{k-1}(1 - w'(i'_j))\right]\cdot(w(i_k) - w'(i'_k))\cdot\left[\prod_{j=k+1}^{K}(1 - w(i_j))\right].
\end{aligned}$$

While Lemma B.1 is folklore in the cascading bandit literature, we provide a proof in Appendix B.4 for the sake of completeness. Now, we proceed to the proof of Lemma 4.4:

Proof. In the proof, we always condition to the historical observation H_t stated in the Lemma. To proceed with the analysis, we define $\tilde{S}_t = (\tilde{i}^t_1, \ldots, \tilde{i}^t_K) \in \mathcal{S}_t$ as an ordered K-subset that satisfies the following minimization criterion:

$$\sum_{k=1}^{K}\left[\prod_{j=1}^{k-1}(1 - w(\tilde{i}_j))\right](g_t(\tilde{i}_j) + h_t(\tilde{i}_j)) = \min_{S=(i_1,\ldots,i_K)\in\mathcal{S}_t}\sum_{k=1}^{K}\left[\prod_{j=1}^{k-1}(1 - w(i_j))\right](g_t(i_j) + h_t(i_j)). \quad \text{(B.10)}$$

We emphasize that both $\tilde{S}_t$ and the left hand side of (B.10) are deterministic in the current discussion, where we condition on H_t. To establish tight bounds on the regret, we consider the truncated version, $\tilde{\boldsymbol{\theta}}_t \in [0,1]^L$, of the Thompson sample $\boldsymbol{\theta}_t$. For each $i \in L$, define

$$\tilde{\theta}_t(i) = \min\{1, \max\{0, \theta_t(i)\}\}.$$

The truncated version $\tilde{\theta}_t(i)$ serves as a correction of $\theta_t(i)$, in the sense that the Thompson sample $\theta_t(i)$, which serves as a Bayesian estimate of click probability $w(i)$, should lie in $[0,1]$. It is important to observe the following two properties hold under the truncated Thompson sample $\tilde{\boldsymbol{\theta}}_t$:

Property 1 Our pulled arm S_t is still optimal under the truncated estimate $\tilde{\boldsymbol{\theta}}_t$, i.e.

$$S_t \in \text{argmax}_{S\in\pi_K(L)}\ r(S|\tilde{\boldsymbol{\theta}}_t).$$

Indeed, the truncated Thompson sample can be sorted in a descending order in the same way as for the original Thompson sample[1], i.e. $\tilde{\theta}_t(i_1^t) \geq \tilde{\theta}_t(i_2^t) \geq \ldots \geq \tilde{\theta}_t(i_K^t) \geq \max_{i\in[L]\setminus\{i_1^t,\ldots,i_K^t\}} \tilde{\theta}_t(i)$. The optimality of S_t thus follows.

Property 2 For any t, i, if it holds that $|\theta_t(i) - w(i)| \leq g_t(i) + h_t(i)$, then it also holds that $|\tilde{\theta}_t(i) - w(i)| \leq g_t(i) + h_t(i)$. Indeed, we know that $|\tilde{\theta}_t(i) - w(i)| \leq |\theta_t(i) - w(i)|$.

Now, we use the ordered K-subset $\tilde{S}_t$ and the truncated Thompson sample $\tilde{\boldsymbol{\theta}}_t$ to decompose the conditionally expected round t regret as follows:

$$\begin{aligned}
r(S^*|\boldsymbol{w}) - r(S_t|\boldsymbol{w}) &= \Big[r(S^*|\boldsymbol{w}) - r(\tilde{S}_t|\boldsymbol{w})\Big] + \Big[r(\tilde{S}_t|\boldsymbol{w}) - r(S_t|\boldsymbol{w})\Big] \\
&\leq \Big[r(S^*|\boldsymbol{w}) - r(\tilde{S}_t|\boldsymbol{w})\Big] + \Big[r(\tilde{S}_t|\boldsymbol{w}) - r(S_t|\boldsymbol{w})\Big]\, 1(\mathcal{E}_{\theta,t}) + 1(\neg\mathcal{E}_{\theta,t}) \\
&\leq \underbrace{\Big[r(S^*|\boldsymbol{w}) - r(\tilde{S}_t|\boldsymbol{w})\Big]}_{(\diamondsuit)} + \underbrace{\Big[r(\tilde{S}_t|\tilde{\boldsymbol{\theta}}_t) - r(S_t|\tilde{\boldsymbol{\theta}}_t)\Big]}_{(\clubsuit)} \\
&\quad + \underbrace{\Big[r(S_t|\tilde{\boldsymbol{\theta}}_t) - r(S_t|\boldsymbol{w})\Big]\, 1(\mathcal{E}_{\theta,t})}_{(\heartsuit)} + \underbrace{\Big[r(\tilde{S}_t|\boldsymbol{w}) - r(\tilde{S}_t|\tilde{\boldsymbol{\theta}}_t)\Big]\, 1(\mathcal{E}_{\theta,t})}_{(\spadesuit)} + 1(\neg\mathcal{E}_{\theta,t}).
\end{aligned} \tag{B.11}$$

We bound $(\diamondsuit, \clubsuit, \heartsuit, \spadesuit)$ from above as follows:

Bounding $(\diamondsuit)$: By the assumption that $\tilde{S}_t = (\tilde{i}_1^t, \ldots, \tilde{i}_K^t) \in \mathcal{S}_t$, with certainty we have

$$(\diamondsuit) \leq \sum_{k=1}^{K} \left[\prod_{j=1}^{k-1} (1 - w(\tilde{i}_j^t))\right] (g_t(\tilde{i}_j^t) + h_t(\tilde{i}_j^t)). \tag{B.12}$$

Bounding $(\clubsuit)$: By **Property 1** of the truncated Thompson sample $\tilde{\boldsymbol{\theta}}_t$, we know that $r(S_t|\tilde{\boldsymbol{\theta}}_t) = \max_{S\in\pi_K(L)}\ r(S|\tilde{\boldsymbol{\theta}}_t) \geq r(\tilde{S}_t|\tilde{\boldsymbol{\theta}}_t)$. Therefore, with certainty we have

$$(\clubsuit) \leq 0. \tag{B.13}$$

[1]Recall that that $\theta_t(i_1^t) \geq \theta_t(i_2^t) \geq \ldots \geq \theta_t(i_K^t) \geq \max_{i\in[L]\setminus\{i_k^t\}_{k=1}^K} \theta_t(i)$ for the original Thompson sample $\boldsymbol{\theta}_t$.

Bounding ($\heartsuit$)**:** . We bound the term as follows:

$$
\begin{aligned}
&\mathsf{1}(\mathcal{E}_{\theta,t}) \left[r(S_t|\tilde{\boldsymbol{\theta}}_t) - r(S_t|\boldsymbol{w}) \right] \\
&= \mathsf{1}(\mathcal{E}_{\theta,t}) \sum_{k=1}^{K} \left[\prod_{j=1}^{k-1} (1 - w(i_j^t)) \right] \cdot (\tilde{\theta}_t(i_k^t) - w(i_k^t)) \cdot \left[\prod_{j=k+1}^{K} (1 - \tilde{\theta}_t(i_j^t)) \right] && \text{(B.14)} \\
&\leq \mathsf{1}(\mathcal{E}_{\theta,t}) \sum_{k=1}^{K} \left[\prod_{j=1}^{k-1} (1 - w(i_j^t)) \right] \cdot \left| \tilde{\theta}_t(i_k^t) - w(i_k^t) \right| \cdot \left[\prod_{j=k+1}^{K} \left| 1 - \tilde{\theta}_t(i_j^t) \right| \right] \\
&\leq \mathsf{1}(\mathcal{E}_{\theta,t}) \sum_{k=1}^{K} \left[\prod_{j=1}^{k-1} (1 - w(i_j^t)) \right] \cdot \left[g_t(i_k^t) + h_t(i_k^t) \right] && \text{(B.15)} \\
&\leq \sum_{k=1}^{K} \left[\prod_{j=1}^{k-1} (1 - w(i_j^t)) \right] \cdot \left[g_t(i_k^t) + h_t(i_k^t) \right]. && \text{(B.16)}
\end{aligned}
$$

Equality (B.14) is by applying the second equality in Lemma B.1, with $S = S' = S_t$, as well as $\boldsymbol{w}' \leftarrow \boldsymbol{w}$, $\boldsymbol{w} \leftarrow \boldsymbol{\theta}_t$. Inequality (B.15) is by the following two facts: (1) By the definition of the truncated Thompson sample $\tilde{\boldsymbol{\theta}}$, we know that $\left|1 - \tilde{\theta}_t(i)\right| \leq 1$ for all $i \in [L]$; (2) By assuming event $\mathcal{E}_{\theta,t}$ and conditioning on H_t where event $\mathcal{E}_{\hat{\mu},t}$ holds true, **Property 2** implies that that $|\tilde{\theta}_t(i) - w(i)| \leq g_t(i) + h_t(i)$ for all i.

Bounding ($\spadesuit$)**:** The analysis is similar to the analysis on ($\heartsuit$):

$$
\begin{aligned}
&\mathsf{1}(\mathcal{E}_{\theta,t}) \left[r(\tilde{S}_t|\boldsymbol{w}) - r(\tilde{S}_t|\tilde{\boldsymbol{\theta}}_t) \right] \\
&= \mathsf{1}(\mathcal{E}_{\theta,t}) \sum_{k=1}^{K} \left[\prod_{j=1}^{k-1} (1 - w(\tilde{i}_j^t)) \right] \cdot (w(\tilde{i}_k^t) - \tilde{\theta}_t(\tilde{i}_k^t)) \cdot \left[\prod_{j=k+1}^{K} (1 - \tilde{\theta}_t(\tilde{i}_j^t)) \right] && \text{(B.17)} \\
&\leq \mathsf{1}(\mathcal{E}_{\theta,t}) \sum_{k=1}^{K} \left[\prod_{j=1}^{k-1} (1 - w(\tilde{i}_j^t)) \right] \cdot \left[g_t(\tilde{i}_k^t) + h_t(\tilde{i}_k^t) \right] && \text{(B.18)} \\
&\leq \sum_{k=1}^{K} \left[\prod_{j=1}^{k-1} (1 - w(\tilde{i}_j^t)) \right] \cdot \left[g_t(\tilde{i}_k^t) + h_t(\tilde{i}_k^t) \right]. && \text{(B.19)}
\end{aligned}
$$

Equality (B.17) is by applying the first equality in Lemma B.1, with $S = S' = \tilde{S}_t$, and $\boldsymbol{w} \leftarrow \boldsymbol{w}$, $\boldsymbol{w}' \leftarrow \boldsymbol{\theta}_t$. Inequality (B.18) follows the same logic as inequality (B.15).

Altogether, collating the bounds (B.12, B.13, B.16, B.19) for ($\diamondsuit, \clubsuit, \heartsuit, \spadesuit$) respectively, we bound (B.11) from above (conditioned on H_t) as follows:

$$
\begin{aligned}
r(S^*|\boldsymbol{w}) - r(S_t|\boldsymbol{w}) \leq 2 \sum_{k=1}^{K} \left[\prod_{j=1}^{k-1} (1 - w(\tilde{i}_j^t)) \right] (g_t(\tilde{i}_j^t) + h_t(\tilde{i}_j^t)) \\
+ \sum_{k=1}^{K} \left[\prod_{j=1}^{k-1} (1 - w(i_j^t)) \right] \cdot \left[g_t(i_k^t) + h_t(i_k^t) \right] + \mathsf{1}(\neg \mathcal{E}_{\theta,t}). \qquad \text{(B.20)}
\end{aligned}
$$

Now, observe that

$$
\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}_t} \left[\sum_{k=1}^{K} \left[\prod_{j=1}^{k-1} (1 - w(i_j^t)) \right] (g_t(i_j^t) + h_t(i_j^t)) \;\middle|\; H_t \right] \\
&\geq \mathbb{E}_{\boldsymbol{\theta}_t} \left[\sum_{k=1}^{K} \left[\prod_{j=1}^{k-1} (1 - w(i_j^t)) \right] (g_t(i_j^t) + h_t(i_j^t)) \;\middle|\; H_t, S_t \in \mathcal{S}_t \right] \Pr_{\boldsymbol{\theta}_t} \left[S_t \in \mathcal{S}_t \middle| H_t \right]
\end{aligned}
$$

$$\geq \left\{ \sum_{k=1}^{K} \left[\prod_{j=1}^{k-1} (1 - w(\tilde{i}_j)) \right] (g_t(\tilde{i}_j) + h_t(\tilde{i}_j)) \right\} \cdot \left(c - \frac{1}{2(t+1)^3} \right), \tag{B.21}$$

where we recall that $f(\lambda, t)$ is the probability lower bound defined in Equation (4.7). Thus, taking conditional expectation $\mathbb{E}_{\boldsymbol{\theta}_t}[\cdot|H_t]$ on both sides in inequality (B.20) gives

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\theta}_t}[R(S^*|\boldsymbol{w}) - R(S_t|\boldsymbol{w})|H_t] \\
&\leq \left(1 + \frac{2}{c - \frac{1}{2(t+1)^3}}\right) \mathbb{E}_{\boldsymbol{\theta}_t} \left[\sum_{k=1}^{K} \left[\prod_{j=1}^{k-1} (1 - w(i_j^t)) \right] \cdot \left[g_t(i_k^t) + h_t(i_k^t) \right] \;\middle|\; H_t \right] + \mathbb{E}_{\boldsymbol{\theta}_t}[1(\neg \mathcal{E}_{\theta,t})|H_t].
\end{aligned}$$

Finally, the Lemma is proved by the assumption that $c > 1/(t+1)^3$, and noting from Lemma 4.1 that $\mathbb{E}_{\boldsymbol{\theta}_t}[1(\neg \mathcal{E}_{\theta,t})|H_t] \leq 1/(2(t+1)^3)$. □

B.4 Proof of Lemma B.1

Lemma B.1. *[Implied by Zong et al. (2016)] Let $S = (i_1, \ldots, i_K)$, $S' = (i'_1, \ldots, i'_K)$ be two arbitrary ordered K-subsets of $[L]$. For any $\boldsymbol{w}, \boldsymbol{w}' \in \mathbb{R}^L$, the following equalities holds:*

$$\begin{aligned}
r(S|\boldsymbol{w}) - r(S'|\boldsymbol{w}') &= \sum_{k=1}^{K} \left[\prod_{j=1}^{k-1} (1 - w(i_j)) \right] \cdot (w(i_k) - w'(i'_k)) \cdot \left[\prod_{j=k+1}^{K} (1 - w'(i'_j)) \right] \\
&= \sum_{k=1}^{K} \left[\prod_{j=1}^{k-1} (1 - w'(i'_j)) \right] \cdot (w(i_k) - w'(i'_k)) \cdot \left[\prod_{j=k+1}^{K} (1 - w(i_j)) \right].
\end{aligned}$$

Proof. Observe that

$$\begin{aligned}
&\sum_{k=1}^{K} \left[\prod_{j=1}^{k-1} (1 - w(i_j)) \right] \cdot (w(i_k) - w'(i'_k)) \cdot \left[\prod_{j=k+1}^{K} (1 - w'(i'_j)) \right] \\
&= \sum_{k=1}^{K} \left\{ \left[\prod_{j=1}^{k-1} (1 - w(i_j)) \right] \cdot \left[\prod_{j=k}^{K} (1 - w'(i'_j)) \right] - \left[\prod_{j=1}^{k} (1 - w(i_j)) \right] \cdot \left[\prod_{j=k+1}^{K} (1 - w'(i'_j)) \right] \right\} \\
&= \prod_{k=1}^{K} (1 - w'(i'_k)) - \prod_{k=1}^{K} (1 - w(i_k)) = R(S|\boldsymbol{w}) - R(S'|\boldsymbol{w}'),
\end{aligned}$$

and also that (actually we can also see this by a symmetry argument)

$$\begin{aligned}
&\sum_{k=1}^{K} \left[\prod_{j=1}^{k-1} (1 - w'(i'_j)) \right] \cdot (w(i_k) - w'(i'_k)) \cdot \left[\prod_{j=k+1}^{K} (1 - w(i_j)) \right] \\
&= \sum_{k=1}^{K} \left\{ \left[\prod_{j=1}^{k} (1 - w'(i'_j)) \right] \cdot \left[\prod_{j=k+1}^{K} (1 - w(i_j)) \right] - \left[\prod_{j=1}^{k-1} (1 - w'(i'_j)) \right] \cdot \left[\prod_{j=k}^{K} (1 - w(i_j)) \right] \right\} \\
&= \prod_{k=1}^{K} (1 - w'(i'_k)) - \prod_{k=1}^{K} (1 - w(i_k)) = R(S|\boldsymbol{w}) - R(S'|\boldsymbol{w}').
\end{aligned}$$

This completes the proof. □

C Additional Numerical Results

We set $L \in \{16, 32, 64, 128, 256\}$, $K \in \{2, 4, 8\}$ and $\Delta \in \{0.15, 0.075\}$. This results in 30 parameter settings. We record all the results in Table 3. Here we can see that our algorithm clearly outperforms the other two algorithms when L is large, Δ is small. This superiority manifests itself in the expected regret and the time complexity.

Table 3: The performances of TS-Cascade, CascadeKL-UCB and CascadeUCB1 under 30 different settings. For each algorithm, the first column shows the mean and the standard deviation of $\mathrm{Reg}(T)$ and the second column shows the average running time in seconds. For each problem setting, the algorithm with smallest average $\mathrm{Reg}(T)$ or shortest running time is marked in bold.

L	K	Δ	TS-Cascade		CascadeKL-UCB		CascadeUCB1	
16	2	0.15	377.07 ± 11.67	3.16	**359.35 ± 26.42**	54.3	1277.42 ± 25.88	**2.82**
16	4	0.15	294.55 ± 15.08	3.03	**265.9 ± 20.36**	54.48	990.51 ± 31.72	**2.84**
16	8	0.15	**138.85 ± 9.81**	3.51	148.36 ± 12.35	55.5	555.83 ± 14.41	**3.17**
16	2	0.075	**691.6 ± 58.39**	2.98	736.08 ± 56.36	54.11	2028.56 ± 71.56	**2.94**
16	4	0.075	546.46 ± 40.78	3.15	**526.93 ± 52.76**	54.41	1485.14 ± 58.43	**2.85**
16	8	0.075	**252.74 ± 20.52**	3.44	261.76 ± 33.86	54.24	713.43 ± 46.93	**2.9**
32	2	0.15	**738.19 ± 19.23**	3.41	764.42 ± 48.57	105.4	2711.44 ± 58.41	**2.98**
32	4	0.15	**612.36 ± 10.66**	3.55	619.68 ± 34.56	105.56	2237.77 ± 43.7	**3.02**
32	8	0.15	**381.8 ± 13.19**	3.68	419.39 ± 19.59	105.64	1526.97 ± 24.48	**3.14**
32	2	0.075	**1159 ± 63.43**	**3.49**	1583.33 ± 104.04	106.62	4217.87 ± 129.08	3.95
32	4	0.075	**1062.9 ± 80.06**	**3.55**	1208.06 ± 59.25	106.08	3301.44 ± 85.43	3.84
32	8	0.075	**631.45 ± 51.51**	**3.58**	718.65 ± 32.27	106.51	1890.06 ± 47.8	3.97
64	2	0.15	**1400.97 ± 45.61**	4.62	1555.42 ± 44.88	208.48	5408.46 ± 83.34	**4.13**
64	4	0.15	**1194.26 ± 21.69**	5.47	1283.29 ± 49.22	208.3	4609.41 ± 84.2	**4.17**
64	8	0.15	**812.1 ± 29.36**	**4.73**	937.02 ± 30.52	208.03	3307.08 ± 43.78	4.74
64	2	0.075	**1810.43 ± 126.74**	4.74	3169.17 ± 156.98	207.31	7599.58 ± 199.99	**4.24**
64	4	0.075	**1730.13 ± 128.09**	**4.88**	2512.28 ± 106.85	208.08	6437.43 ± 239.96	5.04
64	8	0.075	**1175.07 ± 46.91**	**4.7**	1565.76 ± 72.98	208.34	3962.35 ± 87.61	4.77
128	2	0.15	**2520.03 ± 74.04**	5.06	3114.73 ± 74.62	416.05	10677.3 ± 193.72	**4.18**
128	4	0.15	**2216.26 ± 50.54**	4.71	2602.08 ± 51.29	413.77	9163.15 ± 126.39	**4.52**
128	8	0.15	**1591.75 ± 32.73**	5.39	1916.45 ± 61.9	414.58	6589.88 ± 67.56	**4.77**
128	2	0.075	**2784.44 ± 185.08**	5.36	6160.86 ± 300.48	414.45	11055.68 ± 156.27	**5.17**
128	4	0.075	**2837.25 ± 239.41**	4.76	5004.45 ± 188.68	412.55	11516.47 ± 227.48	**4.7**
128	8	0.075	**2004.58 ± 122.26**	4.87	3084.67 ± 105.78	413.6	7432.14 ± 129.24	**4.61**
256	2	0.15	**4386.43 ± 315.68**	**8.05**	6255.14 ± 131.46	817.17	19088.19 ± 318.55	9.37
256	4	0.15	**3998.61 ± 107.35**	**6.95**	5209.96 ± 80.16	820.48	17287.79 ± 221.64	8.64
256	8	0.15	**2934.38 ± 53.36**	**7.47**	3786.36 ± 66.26	818.43	12519.56 ± 125.97	7.81
256	2	0.075	**4128.96 ± 400.88**	8.35	10426.63 ± 249.33	816.52	12191.23 ± 39.69	**7.22**
256	4	0.075	**4376.73 ± 373.99**	**7.49**	9389.72 ± 251.5	818.07	15748.08 ± 131.08	7.56
256	8	0.075	**3258.24 ± 238.91**	**7.24**	6019.24 ± 145.95	820	12417.86 ± 160.53	7.83

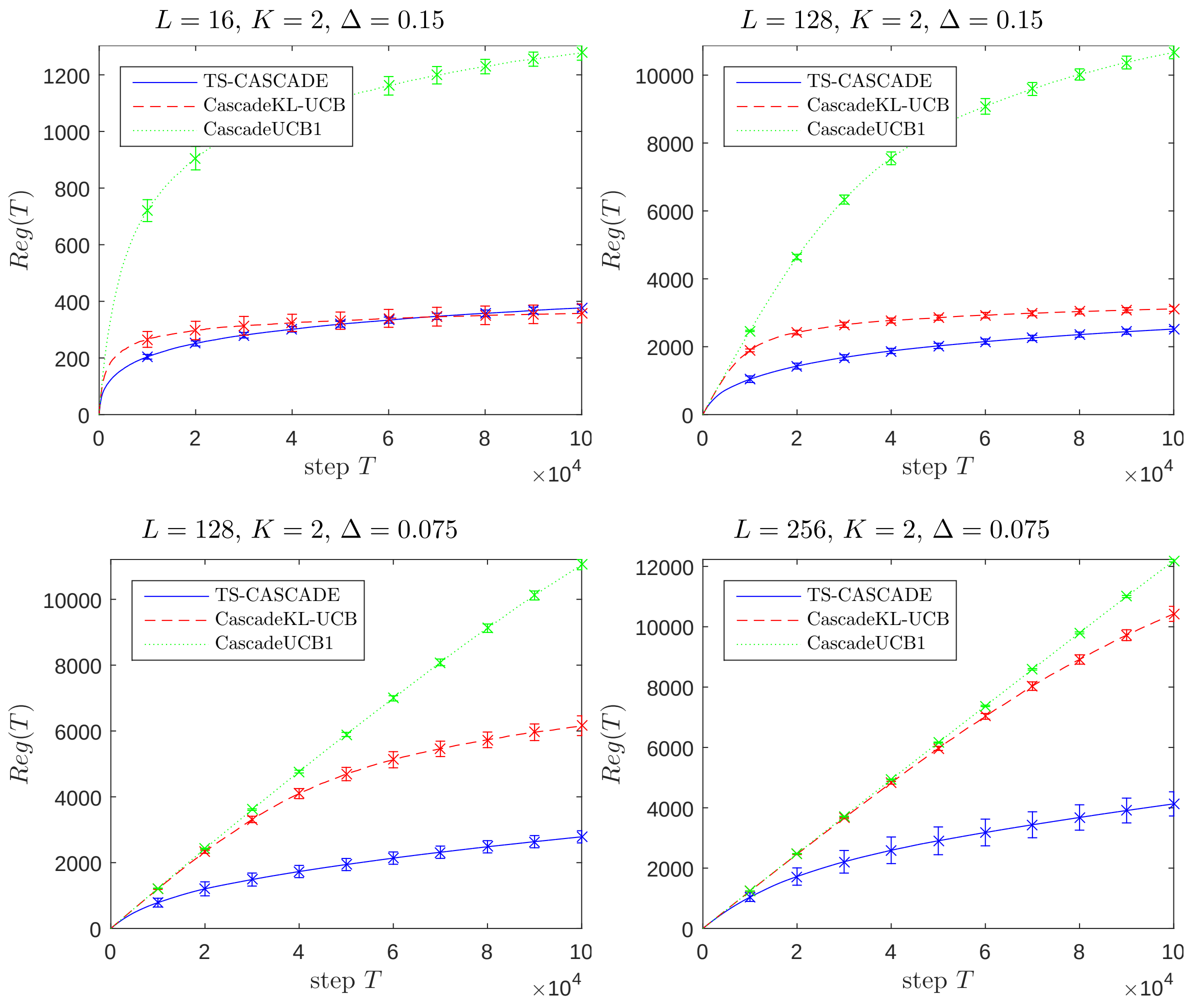

Figure 2: The T-step regret $\mathrm{Reg}(T)$ of TS-Cascade, CascadeKL-UCB and CascadeUCB1 under 4 different parameter settings. Each line indicates the average $\mathrm{Reg}(T)$ of an algorithm and the length of each vertical error bar above and below each data point is the standard deviation.

To better understand the evolution of T-step regret $\mathrm{Reg}(T)$, we present four more plots in Figure 2. First of all, our algorithm clearly beats CascadeUCB1 in all simulations. Even though our algorithm sometimes requires slightly more time to run than CascadeUCB1, there is a significant improvement of TS-Cascade over CascadeUCB1. Comparing TS-Cascade to CascadeKL-UCB, we notice that when $L = 16$, $K = 2$, $\Delta = 0.15$, CascadeKL-UCB slightly outperforms our algorithm but requires more than ten times the computational time. Besides, when $L = 128$, $K = 2$, $\Delta = 0.15$, our algorithm outperforms CascadeKL-UCB in terms of regret and computational time. For the other two settings, our algorithm generates a T-step regret smaller than half of that of CascadeKL-UCB, which confirms the superiority of our algorithm when the ground set is large.