A. Proofs of Main Theorems

A.1. Proof of Theorem 2

Let $\mathbf{R}_t = R(\mathbf{A}_t, \mathbf{w}_t)$ be the regret of the learning algorithm at time t, where $\mathbf{A}_t$ is the recommended list at time t and $\mathbf{w}_t$ are the weights of items at time t. Let $\mathcal{E}_t = \left\{\exists e \in E \text{ s.t. } \left|\bar{w}(e) - \hat{\mathbf{w}}_{\mathbf{T}_{t-1}(e)}(e)\right| \geq c_{t-1,\mathbf{T}_{t-1}(e)}\right\}$ be the event that $\bar{w}(e)$ is not in the high-probability confidence interval around $\hat{\mathbf{w}}_{\mathbf{T}_{t-1}(e)}(e)$ for some e at time t; and let $\overline{\mathcal{E}}_t$ be the complement of $\mathcal{E}_t$, $\bar{w}(e)$ is in the high-probability confidence interval around $\hat{\mathbf{w}}_{\mathbf{T}_{t-1}(e)}(e)$ for all e at time t. Then we can decompose the regret of `CascadeUCB1` as:

$$R(n) = \mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\{\mathcal{E}_t\}\,\mathbf{R}_t\right] + \mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\{\overline{\mathcal{E}}_t\}\,\mathbf{R}_t\right]. \tag{10}$$

Now we bound both terms in the above regret decomposition.

The first term in (10) is small because all of our confidence intervals hold with high probability. In particular, Hoeffding's inequality (Boucheron et al., 2013, Theorem 2.8) yields that for any e, s, and t:

$$P(|\bar{w}(e) - \hat{\mathbf{w}}_s(e)| \geq c_{t,s}) \leq 2\exp[-3\log t]\,,$$

and therefore:

$$\mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\{\mathcal{E}_t\}\right] \leq \sum_{e \in E}\sum_{t=1}^{n}\sum_{s=1}^{t} P(|\bar{w}(e) - \hat{\mathbf{w}}_s(e)| \geq c_{t,s}) \leq 2\sum_{e \in E}\sum_{t=1}^{n}\sum_{s=1}^{t} \exp[-3\log t] \leq 2\sum_{e \in E}\sum_{t=1}^{n} t^{-2} \leq \frac{\pi^2}{3}L\,.$$

Since $\mathbf{R}_t \leq 1$, $\mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\{\mathcal{E}_t\}\,\mathbf{R}_t\right] \leq \frac{\pi^2}{3}L$.

Recall that $\mathbb{E}_t[\cdot] = \mathbb{E}[\cdot \,|\, \mathcal{H}_t]$, where $\mathcal{H}_t$ is the history of the learning agent up to choosing $\mathbf{A}_t$, the first $t-1$ observations and t actions (4). Based on this definition, we rewrite the second term in (10) as:

$$\mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\{\overline{\mathcal{E}}_t\}\,\mathbf{R}_t\right] \overset{\text{(a)}}{=} \sum_{t=1}^{n} \mathbb{E}\left[\mathbb{1}\{\overline{\mathcal{E}}_t\}\,\mathbb{E}_t[\mathbf{R}_t]\right] \overset{\text{(b)}}{\leq} \sum_{e=K+1}^{L} \mathbb{E}\left[\sum_{e^*=1}^{K}\sum_{t=1}^{n} \Delta_{e,e^*}\mathbb{1}\{\overline{\mathcal{E}}_t, G_{e,e^*,t}\}\right],$$

where equality (a) is due to the tower rule and that $\mathbb{1}\{\overline{\mathcal{E}}_t\}$ is only a function of $\mathcal{H}_t$, and inequality (b) is due to the upper bound in Theorem 1.

Now we bound $\sum_{e^*=1}^{K}\sum_{t=1}^{n} \Delta_{e,e^*}\mathbb{1}\{\overline{\mathcal{E}}_t, G_{e,e^*,t}\}$ for any suboptimal item e. Select any optimal item e^*. When event $\overline{\mathcal{E}}_t$ happens, $\left|\bar{w}(e) - \hat{\mathbf{w}}_{\mathbf{T}_{t-1}(e)}(e)\right| < c_{t-1,\mathbf{T}_{t-1}(e)}$. Moreover, when event $G_{e,e^*,t}$ happens, $\mathbf{U}_t(e) \geq \mathbf{U}_t(e^*)$ by Theorem 1. Therefore, when both $G_{e,e^*,t}$ and $\overline{\mathcal{E}}_t$ happen:

$$\bar{w}(e) + 2c_{t-1,\mathbf{T}_{t-1}(e)} \geq \mathbf{U}_t(e) \geq \mathbf{U}_t(e^*) \geq \bar{w}(e^*)\,,$$

which implies:

$$2c_{t-1,\mathbf{T}_{t-1}(e)} \geq \Delta_{e,e^*}\,.$$

Together with $c_{n,\mathbf{T}_{t-1}(e)} \geq c_{t-1,\mathbf{T}_{t-1}(e)}$, this implies $\mathbf{T}_{t-1}(e) \leq \tau_{e,e^*}$, where $\tau_{e,e^*} = \frac{6}{\Delta_{e,e^*}^2}\log n$. Therefore:

$$\sum_{e^*=1}^{K}\sum_{t=1}^{n} \Delta_{e,e^*}\mathbb{1}\{\overline{\mathcal{E}}_t, G_{e,e^*,t}\} \leq \sum_{e^*=1}^{K} \Delta_{e,e^*}\sum_{t=1}^{n} \mathbb{1}\{\mathbf{T}_{t-1}(e) \leq \tau_{e,e^*},\ G_{e,e^*,t}\}\,. \tag{11}$$

Let:

$$\mathbf{M}_{e,e^*} = \sum_{t=1}^{n} \mathbb{1}\{\mathbf{T}_{t-1}(e) \leq \tau_{e,e^*},\ G_{e,e^*,t}\}$$

be the inner sum in (11). Now note that (i) the counter $\mathbf{T}_{t-1}(e)$ of item e increases by one when the event $G_{e,e^*,t}$ happens for any optimal item e^*, (ii) the event $G_{e,e^*,t}$ happens for at most one optimal e^* at any time t; and (iii) $\tau_{e,1} \leq \ldots \leq \tau_{e,K}$.

Based on these facts, it follows that $\mathbf{M}_{e,e^*} \leq \tau_{e,e^*}$, and moreover $\sum_{e^*=1}^{K} \mathbf{M}_{e,e^*} \leq \tau_{e,K}$. Therefore, the right-hand side of (11) can be bounded from above by:

$$\max \left\{ \sum_{e^*=1}^{K} \Delta_{e,e^*} m_{e,e^*} : 0 \leq m_{e,e^*} \leq \tau_{e,e^*}, \ \sum_{e^*=1}^{K} m_{e,e^*} \leq \tau_{e,K} \right\} .$$

Since the gaps are decreasing, $\Delta_{e,1} \geq \ldots \geq \Delta_{e,K}$, the solution to the above problem is $m_{e,1}^* = \tau_{e,1}$, $m_{e,2}^* = \tau_{e,2} - \tau_{e,1}$, $\ldots$, $m_{e,K}^* = \tau_{e,K} - \tau_{e,K-1}$. Therefore, the value of (11) is bounded from above by:

$$\left[\Delta_{e,1} \frac{1}{\Delta_{e,1}^2} + \sum_{e^*=2}^{K} \Delta_{e,e^*} \left(\frac{1}{\Delta_{e,e^*}^2} - \frac{1}{\Delta_{e,e^*-1}^2} \right) \right] 6 \log n \,.$$

By Lemma 3 of Kveton et al. (2014a), the above term is bounded by $\frac{12}{\Delta_{e,K}} \log n$. Finally, we chain all inequalities and sum over all suboptimal items e.

A.2. Proof of Theorem 3

Let $\mathbf{R}_t = R(\mathbf{A}_t, \mathbf{w}_t)$ be the regret of the learning algorithm at time t, where $\mathbf{A}_t$ is the recommended list at time t and $\mathbf{w}_t$ are the weights of items at time t. Let $\mathcal{E}_t = \{\exists 1 \leq e \leq K \text{ s.t. } \bar{w}(e) > \mathbf{U}_t(e)\}$ be the event that the attraction probability of at least one optimal item is above its upper confidence bound at time t. Let $\overline{\mathcal{E}}_t$ be the complement of event $\mathcal{E}_t$. Then we can decompose the regret of `CascadeKL-UCB` as:

$$R(n) = \mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\{\mathcal{E}_t\} \, \mathbf{R}_t \right] + \mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\{\overline{\mathcal{E}}_t\} \, \mathbf{R}_t \right] . \tag{12}$$

By Theorems 2 and 10 of Garivier & Cappe (2011), thanks to the choice of the upper confidence bound $\mathbf{U}_t$, the first term in (12) is bounded as $\mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\{\mathcal{E}_t\} \, \mathbf{R}_t\right] \leq 7K \log\log n$. As in the proof of Theorem 2, we rewrite the second term as:

$$\mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\{\overline{\mathcal{E}}_t\} \, \mathbf{R}_t \right] = \sum_{t=1}^{n} \mathbb{E}\left[\mathbb{1}\{\overline{\mathcal{E}}_t\} \, \mathbb{E}_t\left[\mathbf{R}_t\right] \right] \leq \sum_{e=K+1}^{L} \mathbb{E}\left[\sum_{e^*=1}^{K} \sum_{t=1}^{n} \Delta_{e,e^*} \mathbb{1}\{\overline{\mathcal{E}}_t, G_{e,e^*,t}\} \right] .$$

Now note that for any suboptimal item e and $\tau_{e,e^*} > 0$:

$$\begin{aligned} \mathbb{E}\left[\sum_{e^*=1}^{K} \sum_{t=1}^{n} \Delta_{e,e^*} \mathbb{1}\{\overline{\mathcal{E}}_t, G_{e,e^*,t}\} \right] \leq {} & \mathbb{E}\left[\sum_{e^*=1}^{K} \sum_{t=1}^{n} \Delta_{e,e^*} \mathbb{1}\{\mathbf{T}_{t-1}(e) \leq \tau_{e,e^*}, \ G_{e,e^*,t}\} \right] + \\ & \sum_{e^*=1}^{K} \Delta_{e,e^*} \mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\{\mathbf{T}_{t-1}(e) > \tau_{e,e^*}, \ \overline{\mathcal{E}}_t, \ G_{e,e^*,t}\} \right] . \end{aligned} \tag{13}$$

Let:

$$\tau_{e,e^*} = \frac{1+\varepsilon}{D_{\mathrm{KL}}(\bar{w}(e) \,\|\, \bar{w}(e^*))} (\log n + 3 \log\log n) \,.$$

Then by the same argument as in Theorem 2 and Lemma 8 of Garivier & Cappe (2011):

$$\mathbb{E}\left[\sum_{t=1}^{n} \mathbb{1}\{\mathbf{T}_{t-1}(e) > \tau_{e,e^*}, \ \overline{\mathcal{E}}_t, \ G_{e,e^*,t}\} \right] \leq \frac{C_2(\varepsilon)}{n^{\beta(\varepsilon)}}$$

holds for any suboptimal e and optimal e^*. So the second term in (13) is bounded from above by $K \frac{C_2(\varepsilon)}{n^{\beta(\varepsilon)}}$. Now we bound the first term in (13). By the same argument as in the proof of Theorem 2:

$$\begin{aligned} & \sum_{e^*=1}^{K} \sum_{t=1}^{n} \Delta_{e,e^*} \mathbb{1}\{\mathbf{T}_{t-1}(e) \leq \tau_{e,e^*}, \ G_{e,e^*,t}\} \leq \\ & \quad \left[\frac{\Delta_{e,1}}{D_{\mathrm{KL}}(\bar{w}(e) \,\|\, \bar{w}(1))} + \sum_{e^*=2}^{K} \Delta_{e,e^*} \left(\frac{1}{D_{\mathrm{KL}}(\bar{w}(e) \,\|\, \bar{w}(e^*))} - \frac{1}{D_{\mathrm{KL}}(\bar{w}(e) \,\|\, \bar{w}(e^*-1))} \right) \right] (1+\varepsilon)(\log n + 3 \log\log n) \end{aligned}$$

holds for any suboptimal item e. By Lemma 2, the leading constant is bounded as:

$$\frac{\Delta_{e,1}}{D_{\mathrm{KL}}(\bar{w}(e) \,\|\, \bar{w}(1))} + \sum_{e^*=2}^{K} \Delta_{e,e^*} \left(\frac{1}{D_{\mathrm{KL}}(\bar{w}(e) \,\|\, \bar{w}(e^*))} - \frac{1}{D_{\mathrm{KL}}(\bar{w}(e) \,\|\, \bar{w}(e^*-1))}\right) \leq \frac{\Delta_{e,K}(1 + \log(1/\Delta_{e,K}))}{D_{\mathrm{KL}}(\bar{w}(e) \,\|\, \bar{w}(K))} .$$

Finally, we chain all inequalities and sum over all suboptimal items e.

B. Technical Lemmas

Lemma 1. *Let $A = (a_1, \dots, a_K)$ and $B = (b_1, \dots, b_K)$ be any two lists of K items from $\Pi_K(E)$ such that $a_i = b_j$ only if $i = j$. Let $\mathbf{w} \sim P$ in Assumption 1. Then:*

$$\mathbb{E}\left[\prod_{k=1}^{K} \mathbf{w}(a_k) - \prod_{k=1}^{K} \mathbf{w}(b_k)\right] = \sum_{k=1}^{K} \mathbb{E}\left[\prod_{i=1}^{k-1} \mathbf{w}(a_i)\right] \mathbb{E}\left[\mathbf{w}(a_k) - \mathbf{w}(b_k)\right] \left(\prod_{j=k+1}^{K} \mathbb{E}\left[\mathbf{w}(b_j)\right]\right) .$$

Proof. First, we prove that:

$$\prod_{k=1}^{K} w(a_k) - \prod_{k=1}^{K} w(b_k) = \sum_{k=1}^{K} \left(\prod_{i=1}^{k-1} w(a_i)\right) (w(a_k) - w(b_k)) \left(\prod_{j=k+1}^{K} w(b_j)\right)$$

holds for any $w \in \{0,1\}^L$. The proof is by induction on K. The claim holds obviously for $K = 1$. Now suppose that the claim holds for any $A, B \in \Pi_{K-1}(E)$. Let $A, B \in \Pi_K(E)$. Then:

$$\begin{aligned}
\prod_{k=1}^{K} w(a_k) - \prod_{k=1}^{K} w(b_k) &= \prod_{k=1}^{K} w(a_k) - w(b_K) \prod_{k=1}^{K-1} w(a_k) + w(b_K) \prod_{k=1}^{K-1} w(a_k) - \prod_{k=1}^{K} w(b_k) \\
&= (w(a_K) - w(b_K)) \prod_{k=1}^{K-1} w(a_k) + w(b_K) \left[\prod_{k=1}^{K-1} w(a_k) - \prod_{k=1}^{K-1} w(b_k)\right] \\
&= (w(a_K) - w(b_K)) \prod_{k=1}^{K-1} w(a_k) + \sum_{k=1}^{K-1} \left(\prod_{i=1}^{k-1} w(a_i)\right) (w(a_k) - w(b_k)) \left(\prod_{j=k+1}^{K} w(b_j)\right) \\
&= \sum_{k=1}^{K} \left(\prod_{i=1}^{k-1} w(a_i)\right) (w(a_k) - w(b_k)) \left(\prod_{j=k+1}^{K} w(b_j)\right) .
\end{aligned}$$

The third equality is by our induction hypothesis. Finally, note that $\mathbf{w}$ is drawn from a factored distribution. Therefore, we can decompose the expectation of the product as a product of expectations, and our claim follows. ■

Lemma 2. *Let $p_1 \geq \ldots \geq p_K > p$ be $K + 1$ probabilities and $\Delta_k = p_k - p$ for $1 \leq k \leq K$. Then:*

$$\frac{\Delta_1}{D_{\mathrm{KL}}(p \,\|\, p_1)} + \sum_{k=2}^{K} \Delta_k \left(\frac{1}{D_{\mathrm{KL}}(p \,\|\, p_k)} - \frac{1}{D_{\mathrm{KL}}(p \,\|\, p_{k-1})}\right) \leq \frac{\Delta_K(1 + \log(1/\Delta_K))}{D_{\mathrm{KL}}(p \,\|\, p_K)} .$$

Proof. First, we note that:

$$\frac{\Delta_1}{D_{\mathrm{KL}}(p \,\|\, p_1)} + \sum_{k=2}^{K} \Delta_k \left(\frac{1}{D_{\mathrm{KL}}(p \,\|\, p_k)} - \frac{1}{D_{\mathrm{KL}}(p \,\|\, p_{k-1})}\right) = \sum_{k=1}^{K-1} \frac{\Delta_k - \Delta_{k+1}}{D_{\mathrm{KL}}(p \,\|\, p_k)} + \frac{\Delta_K}{D_{\mathrm{KL}}(p \,\|\, p_K)} .$$

The summation over k can be bounded from above by a definite integral:

$$\sum_{k=1}^{K-1} \frac{\Delta_k - \Delta_{k+1}}{D_{\mathrm{KL}}(p \,\|\, p_k)} = \sum_{k=1}^{K-1} \frac{\Delta_k - \Delta_{k+1}}{D_{\mathrm{KL}}(p \,\|\, p + \Delta_k)} \leq \int_{\Delta_K}^{\Delta_1} \frac{1}{D_{\mathrm{KL}}(p \,\|\, p + x)} \,\mathrm{d}x \leq \int_{\Delta_K}^{1} \frac{1}{D_{\mathrm{KL}}(p \,\|\, p + x)} \,\mathrm{d}x ,$$

where the first inequality follows from the fact that $1/D_{\mathrm{KL}}(p \,\|\, p + x)$ decreases on $x \geq 0$. To the best of our knowledge, the integral of $1/D_{\mathrm{KL}}(p \,\|\, p + x)$ over x does not have a simple analytic solution. Therefore, we integrate an upper bound on $1/D_{\mathrm{KL}}(p \,\|\, p + x)$ which does. In particular, note that for any $x \geq \Delta_K$:

$$D_{\mathrm{KL}}(p \,\|\, p + x) \geq \frac{D_{\mathrm{KL}}(p \,\|\, p + \Delta_K)}{\Delta_K} x = \frac{D_{\mathrm{KL}}(p \,\|\, p_K)}{\Delta_K} x$$

because $D_{\mathrm{KL}}(p \,\|\, p + x)$ is convex, increasing in $x \geq 0$, and its minimum is attained at $x = 0$. Therefore:

$$\int_{\Delta_K}^{1} \frac{1}{D_{\mathrm{KL}}(p \,\|\, p + x)} \, \mathrm{d}x \leq \frac{\Delta_K}{D_{\mathrm{KL}}(p \,\|\, p_K)} \int_{\Delta_K}^{1} \frac{1}{x} \, \mathrm{d}x = \frac{\Delta_K}{D_{\mathrm{KL}}(p \,\|\, p_K)} \log(1/\Delta_K) \,.$$

Finally, we chain all inequalities and get the final result. ■