Appendix

A Proof for Theorem 1

A.1 Notations

We start by defining some notations. For each time $t$, we define a random permutation $(a_1^{*t}, \ldots, a_K^{*t})$ of $A^*$ based on $A_t$ as follows: for any $k = 1, \ldots, K$, if $a_k^t \in A^*$, then we set $a_k^{*t} = a_k^t$. The remaining optimal items are positioned arbitrarily. Notice that under this random permutation, we have:

$$\bar{w}(a_k^{*t}) \geq \bar{w}(a_k^t) \quad \text{and} \quad U_t(a_k^t) \geq U_t(a_k^{*t}) \quad \forall k = 1, \ldots, K$$

Moreover, we use $\mathcal{H}_t$ to denote the “history” (rigorously speaking, $\sigma$-algebra) by the end of time $t$. Then both $A_t = (a_1^t, \ldots, a_K^t)$ and the permutation $(a_1^{*t}, \ldots, a_K^{*t})$ of $A^*$ are $\mathcal{H}_{t-1}$-adaptive. In other words, they are conditionally deterministic at the beginning of time $t$. To simplify the notation, in this paper, we use $E_t[\cdot]$ to denote $E[\cdot | \mathcal{H}_{t-1}]$ when appropriate.

When appropriate, we also use $\langle \cdot, \cdot \rangle$ to denote the inner product of two vectors. Specifically, for two vectors $u$ and $v$ with the same dimension, we use $\langle u, v \rangle$ to denote $u^Tv$.

A.2 Regret Decomposition

We first prove the following technical lemma:

Lemma 1. For any $B = (b_1, \ldots, b_K) \in \mathbb{R}^K$ and $C = (c_1, \ldots, c_K) \in \mathbb{R}^K$, we have

$$\prod_{k=1}^K b_k - \prod_{k=1}^K c_k = \sum_{k=1}^K \left( \prod_{i=1}^{k-1} b_i \right) \times [b_k - c_k] \times \left( \prod_{j=k+1}^K c_j \right).$$

Proof. Notice that

$$\sum_{k=1}^K \left( \prod_{i=1}^{k-1} b_i \right) \times [b_k - c_k] \times \left( \prod_{j=k+1}^K c_j \right)
= \sum_{k=1}^K \left\{ \left[ \prod_{i=1}^{k-1} b_i \right] \times \left[ \prod_{j=k+1}^K c_j \right] \right. - \left[ \prod_{i=1}^{k-1} b_i \right] \times \left[ \prod_{j=k}^K c_j \right] \}
= \prod_{k=1}^K b_k - \prod_{k=1}^K c_k.$$

Thus we have

$$R(A_t, w_t) = f(A^*, w_t) - f(A_t, w_t)
= \prod_{k=1}^K (1 - w_t(a_k^t)) - \prod_{k=1}^K (1 - w_t(a_k^{*t}))
\overset{(a)}{=} \sum_{k=1}^K \left[ \prod_{i=1}^{k-1} (1 - w_t(a_i^t)) \right] [w_t(a_k^{*t}) - w_t(a_k^t)] \left[ \prod_{j=k+1}^K (1 - w_t(a_j^{*t})) \right]
\overset{(b)}{=} \sum_{k=1}^K \left[ \prod_{i=1}^{k-1} (1 - w_t(a_i^t)) \right] [\bar{w}(a_k^{*t}) - \bar{w}(a_k^t)],$$

(5)

where equality (a) is based on Lemma 1 and inequality (b) is based on the fact that $\prod_{j=k+1}^K (1 - w_t(a_j^{*t})) \leq 1$. Recall that $A^t$ and the permutation $(a_1^{*t}, \ldots, a_K^{*t})$ of $A^*$ are deterministic conditioning on $\mathcal{H}_{t-1}$, and $a_k^{*t} \neq a_i^t$ for all $i < k$, thus we have

$$E_t[R(A_t, w_t)] \leq E_t \left[ \sum_{k=1}^K \left[ \prod_{i=1}^{k-1} (1 - w_t(a_i^t)) \right] [\bar{w}(a_k^{*t}) - \bar{w}(a_k^t)] \right]
= \sum_{k=1}^K E_t \left[ \prod_{i=1}^{k-1} (1 - w_t(a_i^t)) \right] E_t [\bar{w}(a_k^{*t}) - \bar{w}(a_k^t)]
= \sum_{k=1}^K E_t \left[ \prod_{i=1}^{k-1} (1 - w_t(a_i^t)) \right] [\bar{w}(a_k^{*t}) - \bar{w}(a_k^t)].$$
For any \( t \leq n \) and any \( e \in E \), we define event 
\[
G_{t,k} = \{ \text{item } a_k \text{ is examined in episode } t \},
\]
notice that \( \mathbb{1}\{G_{t,k}\} = \prod_{i=1}^{k-1} (1 - w_i(a_i^t)) \). Thus, we have 
\[
E_t[R_t] \leq \sum_{k=1}^{K} \mathbb{E}_t[\mathbb{1}\{G_{t,k}\} \left[ \bar{w}(a_{k}^{*,t}) - \bar{w}(a_{k}^t) \right]].
\]
Hence, from the tower property, we have 
\[
R(n) \leq \mathbb{E} \left[ \sum_{t=1}^{n} \sum_{k=1}^{K} \mathbb{1}\{G_{t,k}\} \left[ \bar{w}(a_{k}^{*,t}) - \bar{w}(a_{k}^t) \right] \right].
\]
(6)

We further define event \( \mathcal{E} \) as 
\[
\mathcal{E} = \left\{ \langle x_e, \bar{\theta}_{t-1} - \theta^* \rangle \leq c \sqrt{x_e^T M_{t-1}^{-1} x_e}, \forall e \in E, \forall t \leq n \right\},
\]
and \( \bar{\mathcal{E}} \) as the complement of \( \mathcal{E} \). Then we have 
\[
R(n) \overset{(a)}{\leq} P(\bar{\mathcal{E}}) \mathbb{E} \left[ \sum_{t=1}^{n} \sum_{k=1}^{K} \mathbb{1}\{G_{t,k}\} \left[ \bar{w}(a_{k}^{*,t}) - \bar{w}(a_{k}^t) \right] | \mathcal{E} \right] 
+ P(\mathcal{E}) \mathbb{E} \left[ \sum_{t=1}^{n} \sum_{k=1}^{K} \mathbb{1}\{G_{t,k}\} \left[ \bar{w}(a_{k}^{*,t}) - \bar{w}(a_{k}^t) \right] | \bar{\mathcal{E}} \right] 
\overset{(b)}{\leq} \mathbb{E} \left[ \sum_{t=1}^{n} \sum_{k=1}^{K} \mathbb{1}\{G_{t,k}\} \left[ \bar{w}(a_{k}^{*,t}) - \bar{w}(a_{k}^t) \right] | \mathcal{E} \right] + nKP(\bar{\mathcal{E}}),
\]
(8)

where inequality (a) is based on the law of total probability, and the inequality (b) is based on the naive bounds (1) \( P(\mathcal{E}) \leq 1 \) and (2) \( \mathbb{1}\{G_{t,k}\} \left[ \bar{w}(a_{k}^{*,t}) - \bar{w}(a_{k}^t) \right] \leq 1 \). Notice that from the definition of event \( \mathcal{E} \), we have
\[
\bar{w}(e) = \langle x_e, \bar{\theta}_{t-1} - \theta^* \rangle \leq (x_e, \bar{\theta}_{t-1} - \theta^*) + c \sqrt{x_e^T M_{t-1}^{-1} x_e}, \forall e \in E, \forall t \leq n.
\]
under event \( \mathcal{E} \). Moreover, since \( \bar{w}(e) \leq 1 \) by definition, we have \( \bar{w}(e) \leq U_{t}(e) \) for all \( e \in E \) and all \( t \leq n \) under event \( \mathcal{E} \). Hence under event \( \mathcal{E} \), we have 
\[
\bar{w}(a_k^{t}) \leq \bar{w}(a_k^t) \leq U_{t}(a_k^{*,t}) \leq U_t(a_k^t) \leq \langle x_{a_k^t}, \bar{\theta}_{t-1} \rangle + c \sqrt{x_{a_k^t}^T M_{t-1}^{-1} x_{a_k^t}}, \forall t \leq n.
\]

Thus we have 
\[
\bar{w}(a_k^{*,t}) - \bar{w}(a_k^t) \overset{(a)}{\leq} \langle x_{a_k^t}, \bar{\theta}_{t-1} - \theta^* \rangle + c \sqrt{x_{a_k^t}^T M_{t-1}^{-1} x_{a_k^t}} 
\overset{(b)}{\leq} 2c \sqrt{x_{a_k^t}^T M_{t-1}^{-1} x_{a_k^t}},
\]
where inequality (a) follows from the fact that \( \bar{w}(a_k^{*,t}) \leq \langle x_{a_k^t}, \bar{\theta}_{t-1} \rangle + c \sqrt{x_{a_k^t}^T M_{t-1}^{-1} x_{a_k^t}} \) and inequality (b) follows from the fact that \( \langle x_{a_k^t}, \bar{\theta}_{t-1} - \theta^* \rangle \leq c \sqrt{x_{a_k^t}^T M_{t-1}^{-1} x_{a_k^t}} \) under event \( \mathcal{E} \). Thus, we have 
\[
R(n) \leq 2c \mathbb{E} \left[ \sum_{t=1}^{n} \sum_{k=1}^{K} \mathbb{1}\{G_{t,k}\} \sqrt{x_{a_k^t}^T M_{t-1}^{-1} x_{a_k^t}} \left| \mathcal{E} \right\} \right] + nKP(\bar{\mathcal{E}}).
\]

Define \( K_t = \min\{C_t, K\} \), notice that 
\[
\sum_{k=1}^{K} \mathbb{1}\{G_{t,k}\} \sqrt{x_{a_k^t}^T M_{t-1}^{-1} x_{a_k^t}} = \sum_{k=1}^{K_t} \sqrt{x_{a_k^t}^T M_{t-1}^{-1} x_{a_k^t}}.
\]

Thus, we have 
\[
R(n) \leq 2c \mathbb{E} \left[ \sum_{t=1}^{n} \sum_{k=1}^{K_t} \sqrt{x_{a_k^t}^T M_{t-1}^{-1} x_{a_k^t}} \left| \mathcal{E} \right\} \right] + nKP(\bar{\mathcal{E}}).
\]
(9)

In the next two subsections, we will provide a worst-case bound on \( \sum_{t=1}^{n} \sum_{k=1}^{K_t} \sqrt{x_{a_k^t}^T M_{t-1}^{-1} x_{a_k^t}} \) and a bound on \( P(\bar{\mathcal{E}}) \).
A.3 Worst-Case Bound on $\sum_{t=1}^{n} \sum_{k=1}^{K_t} \sqrt{x_{a_k}^T M_{t-1}^{-1} x_{a_k}}$

Lemma 2. $\sum_{t=1}^{n} \sum_{k=1}^{K_t} \sqrt{x_{a_k}^T M_{t-1}^{-1} x_{a_k}} \leq K \sqrt{\frac{dn \log(1 + \frac{dn}{d})}{\log(1 + \frac{1}{d^{\sigma/2}})}}$.

Proof. To simplify the exposition, we define $z_{t,k} = \sqrt{x_{a_k}^T M_{t-1}^{-1} x_{a_k}}$ for all $(t, k)$ s.t. $k \leq K_t$. Recall that

$$M_t = M_{t-1} + \frac{1}{\sigma^2} \sum_{k=1}^{K_t} x_{a_k} x_{a_k}^T.$$ 

Thus, for all $(t, k)$ s.t. $k \leq K_t$, we have that

$$\det[M_t] \geq \det[M_{t-1}] \left( 1 + \frac{1}{\sigma^2} x_{a_k} x_{a_k}^T \right) \geq \det[M_{t-1}] \left( 1 + \frac{1}{\sigma^2} x_{a_k} x_{a_k}^T \right)^{-1}.$$ 

Since $\det[M_t] \geq \det[M_{t-1}]$ and $K_t \leq K$, we have

$$(\det[M_t])^K \geq (\det[M_{t-1}])^K \prod_{k=1}^{K_t} \left( 1 + \frac{z_{t,k}^2}{\sigma^2} \right).$$ 

So we have

$$\left( \det[M_n] \right)^K \geq (\det[M_0])^K \prod_{t=1}^{n} \prod_{k=1}^{K_t} \left( 1 + \frac{z_{t,k}^2}{\sigma^2} \right) = \prod_{t=1}^{n} \prod_{k=1}^{K_t} \left( 1 + \frac{z_{t,k}^2}{\sigma^2} \right),$$ 

since $M_0 = I$. On the other hand, we have that

$$\text{trace}(M_n) = \text{trace} \left( I + \frac{1}{\sigma^2} \sum_{t=1}^{n} \sum_{k=1}^{K_t} x_{a_k} x_{a_k}^T \right) = d + \frac{1}{\sigma^2} \sum_{t=1}^{n} \sum_{k=1}^{K_t} \|x_{a_k}\|^2 \leq d + \frac{nK}{\sigma^2},$$ 

where the last inequality follows from the fact that $\|x_{a_k}\|^2 \leq 1$ and $K_t \leq K$. From the trace-determinant inequality, we have $\frac{1}{d} \text{trace}(M_n) \geq (\det(M_n))^{\frac{1}{d}}$, thus we have

$$\left[ 1 + \frac{nK}{d\sigma^2} \right]^{\frac{dK}{\sigma^2}} \geq \left[ \frac{1}{d} \text{trace}(M_n) \right]^{\frac{dK}{\sigma^2}} \geq (\det(M_n))^K \geq \prod_{t=1}^{n} \prod_{k=1}^{K_t} \left( 1 + \frac{z_{t,k}^2}{\sigma^2} \right).$$ 

Taking the logarithm, we have

$$dK \log \left[ 1 + \frac{nK}{d\sigma^2} \right] \geq \sum_{t=1}^{n} \sum_{k=1}^{K_t} \log \left( 1 + \frac{z_{t,k}^2}{\sigma^2} \right).$$

Notice that $z_{t,k} = \sqrt{x_{a_k}^T M_{t-1}^{-1} x_{a_k}} \leq \sqrt{x_{a_k}^T M_0^{-1} x_{a_k}} = \|x_{a_k}\|_2 \leq 1$, thus we have $z_{t,k} \leq \frac{\log^2(1 + \frac{1}{\sigma^2})}{\log^2(1 + \frac{1}{\sigma^2})}$. Hence we have

$$\sum_{t=1}^{n} \sum_{k=1}^{K_t} z_{t,k}^2 \leq \frac{1}{\log(1 + \frac{1}{\sigma^2})} \sum_{t=1}^{n} \sum_{k=1}^{K_t} \log \left( 1 + \frac{z_{t,k}^2}{\sigma^2} \right) \leq \frac{dK \log \left[ 1 + \frac{nK}{d\sigma^2} \right]}{\log(1 + \frac{1}{\sigma^2})}.$$ 

Notice that for any $y \in [0, 1]$, we have $y \leq \frac{\log(1 + \frac{1}{\sigma^2})}{\log(1 + \frac{1}{\sigma^2})} = h(y)$. To see it, notice that $h(y)$ is a strictly concave function, and $h(0) = 0$ and $h(1) = 1$. 

\[5\]
Finally, from Cauchy-Schwarz inequality, we have that
\[
\sum_{t=1}^{n} \sum_{k=1}^{K_t} z_{t,k} \leq \sqrt{nK} \left( \sum_{t=1}^{n} \sum_{k=1}^{K_t} z_{t,k}^2 \right) \leq K \sqrt{\frac{dn \log(1 + nK)}{\log(1 + 1/\delta)}}.
\]

### A.4 Bound on \( P(\mathcal{E}) \)

**Lemma 3.** For any \( \sigma > 0 \), any \( \delta \in (0, 1) \), and any
\[
e \geq \frac{1}{\sigma} \sqrt{d \log \left( 1 + \frac{nK}{d\sigma^2} \right) + 2 \log \left( \frac{1}{\delta} \right)} + \| \theta^* \|_2,
\]
we have \( P(\mathcal{E}) \leq \delta \).

**Proof.** We start by defining some useful notations. For any \( t = 1, 2, \ldots \), any \( k = 1, 2, \ldots, K_t \), we define

\[
\eta_{t,k} = w_t(a_k^t) - \bar{w}(a_k^t).
\]

One key observation is that \( \eta_{t,k} \)'s form a Martingale difference sequence (MDS). Moreover, since \( \eta_{t,k} \)'s are bounded in \([-1, 1]\) and hence they are conditionally sub-Gaussian with constant \( R = 1 \). We further define that

\[
V_t = \sigma^2 M_t = \sigma^2 I + \sum_{t=1}^{n} \sum_{k=1}^{K_t} x_{a_k^t} x_{a_k^t}^T,
\]

\[
S_t = \sum_{t=1}^{n} \sum_{k=1}^{K_t} x_{a_k^t} \eta_{t,k} = B_t - \sum_{t=1}^{n} \sum_{k=1}^{K_t} x_{a_k^t} \bar{w}(a_k^t) = B_t - \left[ \sum_{t=1}^{n} \sum_{k=1}^{K_t} x_{a_k^t} x_{a_k^t}^T \right] \theta^*.
\]

As we will see later, we define \( V_t \) and \( S_t \) to use the “self normalized bound” developed in [1] (see Algorithm 1 of [1]). Notice that

\[
M_t \tilde{\theta}_t = \frac{1}{\sigma^2} B_t = \frac{1}{\sigma^2} S_t + \frac{1}{\sigma^2} \left( \sum_{t=1}^{n} \sum_{k=1}^{K_t} x_{a_k^t} x_{a_k^t}^T \right) \theta^* = \frac{1}{\sigma^2} S_t + \left[ M_t - I \right] \theta^*,
\]

where the last equality is based on the definition of \( M_t \). Hence we have

\[
\tilde{\theta}_t - \theta^* = M_t^{-1} \left[ \frac{1}{\sigma^2} S_t - \theta^* \right].
\]

Thus, for any \( e \in E \), we have

\[
\langle x_e, \tilde{\theta}_t - \theta^* \rangle = \left\| x_e^T M_t^{-1} \left[ \frac{1}{\sigma^2} S_t - \theta^* \right] \right\| \leq \left\| x_e \right\|_{M_t^{-1}} \left\| \frac{1}{\sigma^2} S_t - \theta^* \right\|_{M_t^{-1}}
\]

\[
\leq \left\| x_e \right\|_{M_t^{-1}} \left[ \frac{1}{\sigma^2} S_t \right]_{M_t^{-1}} \left\| \theta^* \right\|_{M_t^{-1}},
\]

where the first inequality follows from the Cauchy-Schwarz inequality and the second inequality follows from the triangle inequality. Notice that \( \left\| \theta^* \right\|_{M_t^{-1}} \leq \left\| \theta^* \right\|_{M_0^{-1}} = \left\| \theta^* \right\|_2 \), and \( \frac{1}{\sigma^2} S_t \)_{M_t^{-1}} = \frac{1}{\sigma^2} S_{\tilde{V}^{-1}} \) (since \( M_t^{-1} = \sigma^2 V_t^{-1} \)), so we have

\[
\left\| x_e \right\|_{M_t^{-1}} \left[ \frac{1}{\sigma^2} S_t \right]_{M_t^{-1}} \left\| \theta^* \right\|_2 \].
\]

Notice that the above inequality always holds. We now provide a high-probability bound on \( \left\| S_t \right\|_{V_t^{-1}} \) based on “self normalized bound” proposed in [1]. From Theorem 1 of [1], we know that for any \( \delta \in (0, 1) \), with probability at least \( 1 - \delta \), we have

\[
\left\| S_t \right\|_{V_t^{-1}} \leq \sqrt{2 \log \left( \frac{\det(V_t)^{1/2} \det(V_0)^{-1/2}}{\delta} \right)} \quad \forall t = 0, 1, \ldots
\]

\( ^{6} \text{Notice that the notion of “time” is indexed by the pair (t, k), and follows the lexicographical order.} \)
Notice that \( \det(V_0) = \det(\sigma^2 I) = \sigma^{2d} \). Moreover, from the trace-determinant inequality, we have
\[
[\det(V_t)]^{1/d} \leq \frac{\text{trace}(V_t)}{d} = \sigma^2 + \frac{1}{d} \sum_{\tau=1}^{t} \sum_{k=1}^{K} \|x_{a_k}^\tau\|^2 \leq \sigma^2 + \frac{tK}{d} \leq \sigma^2 + \frac{nK}{d},
\]
where the second inequality follows from the assumption that \( \|x_{a_k}^\tau\|_2 \leq 1 \) and \( K \tau \leq K \), and the last inequality follows from \( t \leq n \). Thus, with probability at least \( 1 - \delta \), we have
\[
\|S_t\|_{V_t^{-1}} \leq \sqrt{d \log \left(1 + \frac{nK}{d\sigma^2}\right) + 2 \log \left(\frac{1}{\delta}\right)} \quad \forall t = 0, 1, \ldots, n - 1.
\]
That is, with probability at least \( 1 - \delta \), we have
\[
|\langle x_e, \bar{\theta}_t - \theta^* \rangle| \leq \|x_e\|_{M_t^{-1}} \left[ \frac{1}{\sigma} \sqrt{d \log \left(1 + \frac{nK}{d\sigma^2}\right) + 2 \log \left(\frac{1}{\delta}\right) + \|\theta^*\|_2} \right]
\]
for all \( t = 0, 1, \ldots, n - 1 \) and \( \forall e \in E \). Recall that by definition of event \( \mathcal{E} \), the above inequality implies that, if
\[
c \geq \frac{1}{\sigma} \sqrt{d \log \left(1 + \frac{nK}{d\sigma^2}\right) + 2 \log \left(\frac{1}{\delta}\right) + \|\theta^*\|_2},
\]
then \( P(\mathcal{E}) \geq 1 - \delta \). That is, \( P(\bar{\mathcal{E}}) \leq \delta \). ■

### A.5 Conclude the Proof

Putting it together, for any \( \sigma > 0 \), any \( \delta \in (0, 1) \), and any
\[
c \geq \frac{1}{\sigma} \sqrt{d \log \left(1 + \frac{nK}{d\sigma^2}\right) + 2 \log \left(\frac{1}{\delta}\right) + \|\theta^*\|_2},
\]
we have that
\[
R(n) \leq 2cE \left[ \sum_{t=1}^{n} \sum_{k=1}^{K_t} x_{a_k}^t x_{a_k}^t M_t^{-1} x_{a_k}^t \mathcal{E} \right] + nKP(\bar{\mathcal{E}})
\]
\[
\leq 2cK \sqrt{dn \log \left(1 + \frac{nK}{d\sigma^2}\right) \log \left(1 + \frac{1}{\sigma^2}\right) + nK}.
\]
(12)

Choose \( \delta = \frac{1}{nK} \), we have the following result: for any \( \sigma > 0 \) and any
\[
c \geq \frac{1}{\sigma} \sqrt{d \log \left(1 + \frac{nK}{d\sigma^2}\right) + 2 \log (nK) + \|\theta^*\|_2},
\]
we have
\[
R(n) \leq 2cK \sqrt{dn \log \left(1 + \frac{nK}{d\sigma^2}\right) \log \left(1 + \frac{1}{\sigma^2}\right) + 1}.
\]