A Detailed Procedure of Simulation Based on Real-World Data

In this appendix we give the detailed procedure of the experiment presented in Section 7.2. We use the Yahoo! Webscope dataset R6A, which consists of more than 45 million user visits to the Yahoo! Today module collected over 10 days in May 2009. The log describes the interaction (view/click) of each user with one randomly chosen article out of 271 articles. It was originally used as an unbiased evaluation benchmark for the LB in explore-exploration setting (Li et al., 2010). The dataset is made of features describes each user u and each article a, both are expressed in 6 dimension feature vectors, accompanied with a binary outcome (clicked/not clicked). We use article-user interaction feature $z_{a,u} \in \mathbb{R}^{36}$, which is expressed by a Kronecker product of a feature vector of article a and that of u. Chu et al. (2009) present a detailed description of the dataset, features and the collection methodology.

In our setting, we use the subset of the dataset which is collected on the one day (May 1st). We first conduct the regularized linear regression on whether the target is clicked ($r_t = 1$) or not clicked ($r_t = -1$). Here, the regularize term is set as 0.01. Let θ^* be the learned parameter, which we regard as the "true" parameter in the simulation. We consider the LB with K arms, the features of which are sampled from the dataset. We limit the the case of $\Delta_i \geq 0.05$ for all arms i in order to make the problem not too hard. The reward r_t at the t-th round is given by

$$r_t = \begin{cases} 1 & \left(\text{w.p. } \frac{1+x_{a_t}^\top \theta^*}{2}\right) \\ -1 & (\text{otherwise}) \end{cases},$$

where x_{a_t} is the feature of the arm selected at the tth round. Although it does not always the case, $x^\top \theta^*$ is happended to be bounded in $[-1, 1]$ for all feature x in the dataset, therefore $(1 + x_{a_t}^\top \theta^*)/2$ is always valid for probability. Furthermore, since $x_{a_t}^\top \theta^* \in [-1, 1]$, the noise variable ε_t is bounded as $\varepsilon_t \in [-2, 2]$, which is known as 2-sub-Gaussian. We run LinGapE on this setting, where the parameter is fixed as $\varepsilon = 0$, $\delta = 0.05$, and $\lambda = 1$, in comparison with $\mathcal{XY}$-static allocation, where the estimation is given by the regularized least-squares estimator with $\lambda = 0.01$.

B Algorithm with Problem Complexity Independent of K

The problem complexity of Algorithm 1 is shown in (13) and can be $\mathcal{O}(K)$ in the worst-case. This is problematic when $K \gg d$. In this section, we describe the trick that makes problem complexity completely independent of K.

The idea is to restrict the arms to be pulled. We first choose $K' = \mathcal{O}(d)$ arms, denoted as $\mathcal{B} = \{b_1, b_2, \ldots, b_{K'}\} \subset [K]$, and force the agent to select arms from $\mathcal{B}$. In other word, the arm selection strategy in (12) would be

$$a_{t+1} = \mathop{\arg\min}_{a \in \mathcal{B}:\, p_a^*(i_t,j_t) > 0} T_a(t)/\tilde{p}_a^*(i_t, j_t), \tag{19}$$

where $\tilde{p}_a^*(i_t, j_t)$ is defined as

$$\tilde{p}_k^*(i_t, j_t) = \begin{cases} \frac{|\tilde{w}_k^*(i_t,j_t)|}{\sum_{k=1}^K |\tilde{w}_k^*(i_t,j_t)|} & (k \in \mathcal{B}) \\ 0 & (k \notin \mathcal{B}) \end{cases}$$

$$\begin{aligned} \tilde{\mathbf{w}}^*(i_t, j_t) = & \mathop{\arg\min}_{\mathbf{w} \in \mathbb{R}^d} \|\mathbf{w}\|_1 \\ & \text{s.t.} \quad x_{i_t} - x_{j_t} = \sum_{k=1}^K w_k x_k \\ & \qquad\quad w_k = 0 \quad (\forall k \notin \mathcal{B}). \end{aligned} \tag{20}$$

This modification does not affect the proof given in Appendix D, hence the bounds in Theorems 2 and 3 remain true with new problem complexity:

$$\tilde{H}_\varepsilon = \sum_{k=1}^{K'} \max_{i,j \in [K]} \frac{\tilde{p}_{b_k}^*(i,j)\tilde{\rho}(i,j)}{\max\left(\varepsilon, \frac{\varepsilon+\Delta_i}{3}, \frac{\varepsilon+\Delta_j}{3}\right)^2},$$

where $\tilde{\rho}$ is

$$\tilde{\rho}(i,j) = \|\tilde{\mathbf{w}}^*(i,j)\|_1.$$

As showed in the following lemma, this problem complexity $\tilde{H}_\varepsilon$ is independent of the number of arms K.

Lemma 1. *Let $\tilde{X}$ be the matrix*

$$\tilde{X} = \left[x_{b_1}, x_{b_2}, \ldots, x_{b_{K'}}\right]^\top \in \mathbb{R}^{K' \times d},$$

and σ be the smallest eigenvalue of $\tilde{X}^\top \tilde{X}$. Then, problem complexity $\tilde{H}_\varepsilon$ is bounded as follows.

$$\tilde{H}_\varepsilon \leq \sum_{k=1}^{K'} \max_{i,j \in [K]} \frac{2L\sqrt{\frac{K'}{\sigma}}}{\max\left(\varepsilon, \frac{\varepsilon+\Delta_i}{3}, \frac{\varepsilon+\Delta_j}{3}\right)^2} = \mathcal{O}(d\sqrt{d}).$$

This lemma states that minimizing $\sqrt{K'/\sigma}$ is the key to reduce the problem complexity $\tilde{H}$. Thus we can improve the sample complexity by choosing a set of arms $\mathcal{B}$ with small size $K' = |\mathcal{B}|$ and large eigenvalues σ.

Proof of Lemma 1. Due to the constraint in (20), we have

$$x_i - x_j = \tilde{X}\tilde{\mathbf{w}}^*(i,j).$$

Using this, we can derive the upper bound of $\tilde{\rho}(i,j)$ as follows.

$$\begin{aligned}
\tilde{\rho}(i,j) &= \sum_{k=1}^{K} |\tilde{w}_k^*(i,j)| \\
&= \sqrt{\left(\sum_{k=1}^{K} |\tilde{w}_k^*(i,j)|\right)^2} \\
&\le \sqrt{K'\left(\sum_{k=1}^{K} |\tilde{w}_k^*(i,j)|^2\right)} \\
&= \sqrt{K'(\tilde{\mathbf{w}}^*(i,j))^\top \tilde{\mathbf{w}}^*(i,j)} \\
&\le \sqrt{K'(\tilde{\mathbf{w}}^*(i,j))^\top \tilde{X}^\top \tilde{X}\tilde{\mathbf{w}}^*(i,j)} \\
&\quad \times \max_{\mathbf{w}\in\mathbb{R}^d} \sqrt{\frac{\mathbf{w}^\top \mathbf{w}}{\mathbf{w}^\top \tilde{X}^\top \tilde{X}\mathbf{w}}} \\
&= \sqrt{\frac{K'}{\sigma}(\tilde{\mathbf{w}}^*(i,j))^\top \tilde{X}^\top \tilde{X}\tilde{\mathbf{w}}^*(i,j)} \\
&\le \sqrt{\frac{K'}{\sigma}} \|x_i - x_j\|_2 \\
&\le 2L\sqrt{\frac{K'}{\sigma}}.
\end{aligned}$$

Considering that $\tilde{p}^*(i,j) \le 1$ and $K' = \mathcal{O}(d)$, we have

$$\tilde{H}_\varepsilon \le \sum_{k=1}^{K'} \max_{i,j\in[K]} \frac{2L\sqrt{\frac{K'}{\sigma}}}{\max\left(\varepsilon, \frac{\varepsilon+\Delta_i}{3}, \frac{\varepsilon+\Delta_j}{3}\right)^2} = \mathcal{O}(d\sqrt{d}),$$

which does not depend on K. □

C Derivation of Ratio $p_k^*(i,j)$

In this appendix, we present the derivation of $p_k^*(i,j)$ defined in (10) and the proof of Lemma 2, which bounds the matrix norm when the arm selection strategy based on the ratio $p_k^*(i,j)$.

The original problem of reducing the interval of confidence bound for given $y \in \mathcal{Y} = \{x - x' | x, x' \in \mathcal{X}\}$ is to obtain

$$\arg\min_{\mathbf{x}_n} \|y\|_{(A^\lambda_{\mathbf{x}_n})^{-1}}$$

in the limit of $n \to \infty$. Since we choose features from the finite set $\mathcal{X}$ in the LB, the problem becomes

$$\min_{C_i\in\mathbb{N}\cup\{0\}} y^\top \left(\frac{\lambda}{n}I + \sum_{i=1}^{K} \frac{C_i}{n} x_i x_i^\top\right)^{-1} y \quad \text{s.t.} \sum_{i=1}^{K} C_i = n. \tag{21}$$

where the C_i represents the number of times that the arm $i \in [K]$ is pulled before the n-th round.

We first conduct the continuous relaxation, which turns the optimization problem (21) into

$$\min_{p_i\ge 0} y^\top \left(\frac{\lambda}{n}I + \sum_{i=1}^{K} p_i x_i x_i^\top\right)^{-1} y \quad \text{s.t.} \sum_{i=1}^{K} p_i = 1,$$

where p_i corresponds to the ratio C_i/n. Although this relaxed problem can be solved by convex optimization, it is not suited for the LB setting because the solution depends on the sample size n. Therefore, we consider the asymptotic case, where the sample size n goes to infinity.

It is proved (Yu et al., 2006, Thm. 3.2) that the continuous relaxed problem is equivalent to

$$\begin{aligned}
&\min_{p_i, w_i} \left\| y - \sum_{i=1}^{K} w_i x_i \right\|^2 + \frac{\lambda}{n}\sum_{i=1}^{K} \frac{w_i^2}{p_i} \\
&\text{s.t.} \sum_{i=1}^{K} p_i = 1,\, p_i \ge 0,\, p_i, w_i \in \mathbb{R}.
\end{aligned} \tag{22}$$

Since we consider $y \in \mathcal{Y}$, there always exists w_i such that $y = \sum_{i=1}^{K} w_i x_i$. Then, $\{w_i\}$ such that $\|y - \sum_{i=1}^{K} w_i x_i\| > 0$ cannot be the optimal solution for sufficiently small λ/n and thus the optimal solution has to satisfy $\|y - \sum_{i=1}^{K} w_i x_i\| = 0$. Therefore, the asymptotic case of (22) corresponds to the problem

$$\begin{aligned}
&\min_{p_i, w_i} \sum_{i=1} \frac{w_i^2}{p_i} \\
&\text{s.t.}\, y = \sum_{i=1}^{K} w_i x_i \\
&\sum_{i=1}^{K} p_i = 1,\, p_i \ge 0,\, w_i \in \mathbb{R},
\end{aligned} \tag{23}$$

the KKT condition of which yields the definition of p^* and $\mathbf{w}^*$ in (10) and (11), respectively. Hence, $\rho(i,j)$, the optimal value of (11), is the optimal value of (23) as well.

If we employ the arm selection strategy in (12) based on p^* in (10), we can bound the matrix norm $\|x_i - x_j\|_{A_t^{-1}}$ as described in the following lemma.

Lemma 2. *Recall that $\rho(i,j)$ and $p_k^*(i,j)$ are defined in (15) and (10), respectively. Let $T_i(t)$ be the number of times that the arm i is pulled before the t-th round. Then, the matrix norm $\|x_i - x_j\|_{A_t^{-1}}$ is bounded by*

$$\|x_i - x_j\|_{A_t^{-1}} \le \sqrt{\frac{\rho(i,j)}{T_{i,j}(t)}},$$

where

$$T_{i,j}(t) = \min_{\substack{k\in[K]:\\ p^*_k(i,j)>0}} T_k(t)/p^*_k(i,j).$$

This lemma is proved by the following lemma.

Lemma 3. *Let A be a positive definite matrix in $\mathbb{R}^{d\times d}$ and x, y be vectors in $\mathbb{R}^d$. Then, for any constant $\alpha > 0$,*

$$y^\top (A + \alpha x x^\top)^{-1} y \le y^\top A^{-1} y$$

holds.

Proof. By Sherman-Morrison formula (Sherman and Morrison, 1950) we have,

$$\begin{aligned} y^\top (A + \alpha x x^\top)^{-1} y &= y^\top \left(A^{-1} - \frac{\alpha A^{-1} x x^\top A^{-1}}{1 + \alpha x^T A^{-1} x} \right) y \\ &= y^\top A^{-1} y - y^\top \frac{\alpha A^{-1} x x^\top A^{-1}}{1 + \alpha x^T A^{-1} x} y \\ &\le y^\top A^{-1} y. \end{aligned}$$

The last inequality follows from the fact that A^{-1} is positive definite. □

Using Lemma 3, we can prove Lemma 2 as follows.

Proof of Lemma 2. By the definition of A_t, we have

$$A_t = \lambda I + \sum_{k=1}^K T_k(t) x_k x_k^\top.$$

Then, for

$$\tilde{A}_t = \lambda I + \sum_{k=1}^K p^*_k(i,j) T_{i,j}(t) x_k x_k^\top,$$

we have

$$\|x_i - x_j\|_{A_t^{-1}} \le \|x_i - x_j\|_{\tilde{A}_t^{-1}}$$

from Lemma 3 and the fact

$$T_k(t) \le p^*_k(i,j) T_{i,j}(t),$$

which can be inferred from the definition of $T_t(i,j)$. Therefore, proving

$$\|x_i - x_j\|^2_{\tilde{A}_t^{-1}} \le \frac{\rho(i,j)}{T_{i,j}(t)}$$

completes the proof of the lemma.

For convenience, we write $x_i - x_j$ as y. The KKT condition of (23) implies that $w^*_k(i,j)$ and $p^*_k(i,j)$ satisfy

$$\begin{aligned} w^*_k(i,j) &= \frac{1}{2} p^*_k(i,j) x_k^\top \gamma \\ y &= \frac{1}{2} \sum_{k=1}^K p^*_k(i,j) x_k x_k^\top \gamma, \end{aligned}$$

where $\gamma \in \mathbb{R}^d$ corresponds to the Lagrange multiplier. Therefore, the optimal value $\rho(i,j)$ can be written as

$$\rho(i,j) = \sum_{i=1}^K \frac{{w^*}^2_k(i,j)}{p^*_k(i,j)} = \frac{1}{4} \gamma^\top \left(\sum_{k=1}^K p^*_k(i,j) x_k x_k^\top \right) \gamma.$$

Now, let B be denoted as

$$B = \left(\sum_{k=1}^K p^*_k(i,j) x_k x_k^\top \right).$$

Then, since $y = \frac{1}{2} B\gamma$, we have

$$\begin{aligned} y^\top \tilde{A}_t^{-1} y - \frac{\rho(y)}{T_{i,j}(t)} &= \frac{1}{4} \gamma^\top B^\top \tilde{A}_t^{-1} B \gamma - \frac{1}{4T_{i,j}(t)} \gamma^\top B \gamma \\ &= \frac{1}{4} \gamma^\top \left(B^\top - \frac{\tilde{A}_t}{T_{i,j}(t)} \right) \tilde{A}_t^{-1} B \gamma \\ &= -\frac{1}{4} \gamma^\top \frac{\lambda}{T_{i,j}(t)} \tilde{A}_t^{-1} B \gamma \\ &\le 0. \end{aligned}$$

The inequality follows from the fact that both of $\tilde{A}_t^{-1}$ and B are positive semi-definite matrices. □

D Proofs of Theorems

In this appendix, we give the proofs of Theorems 1, 2, and 3, which are the main theoretical contribution of this paper. In the proof, we assume that the event $\mathcal{E}$ defined as

$$\mathcal{E} = \{\forall t > 0, \, \forall i, j \in [K], \, |\Delta(i,j) - \hat{\Delta}_t(i,j)| \le \beta_t(i,j)\}$$

occurs, where $\Delta(i,j) = (x_i - x_j)^\top \theta$ is the gap of expected rewards between arms i and j. The following lemma states that this assumption holds with high probability.

Lemma 4. *Event $\mathcal{E}$ holds w.p. at least $1 - \delta$.*

Combining Prop. 2 and union bounds proves this lemma.

D.1 Proof of Theorem 1

Let τ be the stopping round of LinGapE. If $\Delta(a^*, \hat{a}^*) > \varepsilon$ holds, that is the returned arm $\hat{a}^*$ is worse than the best arm a^* by ε, then we have

$$\Delta(a^*, \hat{a}^*) > \varepsilon \ge B(\tau) \ge \hat{\Delta}_\tau(a^*, \hat{a}^*) + \beta_\tau(a^*, \hat{a}^*).$$

The second inequality holds for stopping condition $B(\tau) \le \varepsilon$ and the last follows from the definition of $B(\tau)$ (Line 5 in Algorithm 2). From this inequality, we can see that $\Delta(a^*, \hat{a}^*) > \varepsilon$ means that event $\mathcal{E}$

does not occur. Thus, the probability that LinGapE returns such arms is

$$\mathbb{P}[\Delta(a^*, \hat{a}_\tau) > \varepsilon] \le \mathbb{P}[\bar{\mathcal{E}}] = 1 - \mathbb{P}[\mathcal{E}] \le \delta,$$

where $\bar{\mathcal{E}}$ represents the complement of the event $\mathcal{E}$. The last inequality follows from Lemma 4. Therefore, we can conclude that the returned arm satisfies the condition (1). □

D.2 Proofs of Theorems 2 and 3

We prove Theorems 2 and 3 by combining Lemma 2 with following lemmas.

Lemma 5. *Under event $\mathcal{E}$, $B(t)$ is bounded as follows. If i_t or j_t is the best arm, then*

$$B(t) \le \min(0, -(\Delta_{i_t} \vee \Delta_{j_t}) + \beta_t(i_t, j_t)) + \beta_t(i_t, j_t).$$

Otherwise, we have

$$B(t) \le \min(0, -(\Delta_{i_t} \vee \Delta_{j_t}) + 2\beta_t(i_t, j_t)) + \beta_t(i_t, j_t),$$

where $a \vee b = \max(a, b)$.

Proof. First, we consider the case where either arm i_t or j_t is the best arm a^*. Since arm i_t is the estimated best arm (Line 3 in Algorithm 2), we have

$$\hat{\Delta}_t(j_t, i_t) = (x_{j_t} - x_{i_t})^\top \theta_t^\lambda \le 0. \tag{24}$$

Thus, $B(t)$ is bounded by

$$B(t) = \hat{\Delta}(j_t, i_t) + \beta_t(i_t, j_t) \le \beta_t(i_t, j_t). \tag{25}$$

Therefore, it is sufficient to show

$$B(t) \le -(\Delta_{i_t} \vee \Delta_{j_t}) + 2\beta_t(i_t, j_t). \tag{26}$$

If $i_t = a^*$, then

$$(\Delta_{i_t} \vee \Delta_{j_t}) = \Delta_{j_t} \tag{27}$$

follows from the definition of Δ_a in (14). In this case, $B(t)$ is bounded as

$$\begin{aligned} B(t) &\overset{\text{(a)}}{=} \hat{\Delta}_t(j_t, i_t) + \beta_t(i_t, j_t) \\ &\overset{\text{(b)}}{\le} \Delta(j_t, i_t) + 2\beta_t(i_t, j_t) \\ &\overset{\text{(c)}}{=} -\Delta_{j_t} + 2\beta_t(i_t, j_t) \\ &\overset{\text{(d)}}{=} -(\Delta_{i_t} \vee \Delta_{j_t}) + 2\beta_t(i_t, j_t), \end{aligned}$$

where (a), (b), (c) and (d) follow from the definition of $B(t)$, event $\mathcal{E}$, definition of Δ_a and (14), respectively.

On the other hand, in the case where $j_t = a^*$, we have

$$(\Delta_{i_t} \vee \Delta_{j_t}) = \Delta_{i_t}. \tag{28}$$

In this case, the upper bound of $B(t)$ is derived as

$$\begin{aligned} B(t) &\overset{\text{(a)}}{\le} \beta_t(i_t, j_t) \\ &\overset{\text{(b)}}{\le} -\hat{\Delta}_t(j_t, i_t) + \beta_t(i_t, j_t) \\ &\overset{\text{(c)}}{\le} -\Delta(j_t, i_t) + 2\beta_t(i_t, j_t) \\ &\overset{\text{(d)}}{=} -(\Delta_{i_t} \vee \Delta_{j_t}) + 2\beta_t(i_t, j_t), \end{aligned}$$

where (a), (b), (c) and (d) follow from (25), (24), event $\mathcal{E}$, and (28), respectively.

Therefore, in both cases, (26) holds, which completes the proof of the first inequality in Lemma 5.

Next, we prove the second inequality, which holds when neither $i_t \neq a^*$ nor $j_t \neq a^*$. Again, with (25), it is sufficient to prove

$$B(t) \le -(\Delta_{i_t} \vee \Delta_{j_t}) + 3\beta_t(i_t, j_t). \tag{29}$$

Since $j_t \neq a^*$,

$$\hat{\Delta}_t(a^*, i_t) + \beta_t(a^*, i_t) \le \hat{\Delta}_t(j_t, i_t) + \beta_t(j_t, i_t). \tag{30}$$

follows from the definition of j_t (Line 4 in Algorithm 2). Thus, we have

$$\begin{aligned} \beta_t(i_t, j_t) &\overset{\text{(a)}}{\ge} \hat{\Delta}_t(j_t, i_t) + \beta_t(j_t, i_t) \\ &\overset{\text{(b)}}{\ge} \hat{\Delta}_t(a^*, i_t) + \beta_t(a^*, i_t) \\ &\overset{\text{(c)}}{\ge} \Delta(a^*, i_t), \end{aligned} \tag{31}$$

where (a), (b) and (c) follow from (24), (30), event $\mathcal{E}$, respectively. By using (31) and event $\mathcal{E}$, we have

$$\begin{aligned} B(t) &= \hat{\Delta}_t(j_t, i_t) + \beta_t(i_t, j_t) \\ &\le \Delta(j_t, i_t) + 2\beta_t(i_t, j_t) \\ &= \Delta(j_t, a^*) + \Delta(a^*, i_t) + 2\beta_t(i_t, j_t) \\ &\le -\Delta_{j_t} + 3\beta_t(i_t, j_t). \end{aligned} \tag{32}$$

Moreover, from (25) and (31), we obtain

$$B(t) \le 2\beta_t(i_t, j_t) \le -\Delta_{i_t} + 3\beta_t(i_t, j_t). \tag{33}$$

Combining (32) and (33) yields (29), which was what we wanted. □

Based on Lemmas 2 and 5, we can derive the following lemma, which is the essential part of the proofs of Theorems 2 and 3.

Lemma 6. *Let τ be the stopping time of LinGapE when a_t is determined by* (12)*. Then, statement*

$$\tau \le H_\varepsilon C_\tau^2 + K \tag{34}$$

holds with probability at least $1-\delta$, where C_n is defined as (3)*.*

Proof. From Lemma 4, it suffices to show the (34) holds in the case where event $\mathcal{E}$ occurs. First we derive the upper bound of $T_k(\tau)$. Let $\tilde{t} \le \tau$ be the last round that arm k is pulled. Then,

$$\begin{aligned}\min(0, -\Delta_k + 2\beta_{\tilde{t}-1}(i_{\tilde{t}-1}, j_{\tilde{t}-1})) + \beta_{\tilde{t}-1}(i_{\tilde{t}-1}, j_{\tilde{t}-1}))\\ \ge B(\tilde{t}-1) \ge \varepsilon\end{aligned}$$

follows from Lemma 5 and the fact that stopping condition is not satisfied at the $\tilde{t}$-th round. Applying Lemma 2 yields

$$T_{i_{\tilde{t}-1}, j_{\tilde{t}-1}}(\tilde{t}-1) \le \frac{\rho(i_{\tilde{t}-1}, j_{\tilde{t}-1})}{\max\left(\varepsilon, \frac{\varepsilon+\Delta_{i_{\tilde{t}-1}}}{3}, \frac{\varepsilon+\Delta_{j_{\tilde{t}-1}}}{3}\right)^2} C_{\tilde{t}-1}^2,$$

where C_t is defined in (3). Now, since arm k is pulled at $\tilde{t}$-th round,

$$T_k(\tilde{t}-1) = p_k^*(i_{\tilde{t}-1}, j_{\tilde{t}-1}) T_{i_{\tilde{t}-1}, j_{\tilde{t}-1}}(\tilde{t}-1)$$

holds by definition. Therefore, $T_k(\tau)$ can be bounded as

$$\begin{aligned}T_k(\tau) &= T_k(\tilde{t}-1) + 1\\ &= p_k^*(i_{\tilde{t}-1}, j_{\tilde{t}-1}) T_{i_{\tilde{t}-1}, j_{\tilde{t}-1}}(\tilde{t}-1) + 1\\ &\le \max_{i,j \in [K]} p_k^*(i,j) T_{i,j}(\tilde{t}-1) + 1\\ &\le \frac{p_k^*(i_{\tilde{t}-1}, j_{\tilde{t}-1}) \rho(i_{\tilde{t}-1}, j_{\tilde{t}-1})}{\max\left(\varepsilon, \frac{\varepsilon+\Delta_{i_{\tilde{t}-1}}}{3}, \frac{\varepsilon+\Delta_{j_{\tilde{t}-1}}}{3}\right)^2} C_{\tilde{t}-1}^2 + 1\\ &\le \max_{i,j \in [K]} \frac{p_k^*(i,j)\rho(i,j)}{\max\left(\varepsilon, \frac{\varepsilon+\Delta_i}{3}, \frac{\varepsilon+\Delta_j}{3}\right)^2} C_\tau^2 + 1.\end{aligned}$$

Since $\sum_{k=1}^K T_k(\tau) = \tau$, summing up the upper bound of $T_k(t)$ above yields

$$\tau \le H_\varepsilon C_\tau^2 + K.$$

□

Now, we can complete the proofs by bounding C_τ in (34) by the following proposition.

Proposition 3. *(Abbasi-Yadkori et al., 2011, Lemma 10) Let the maximum l_2 norm of features be denoted as L. Then,* $\det(A_n^\lambda)$ *is bounded as*

$$\det(A_n^\lambda) \le (\lambda + nL^2/d)^d.$$

Proof of Theorem 2. From Proposition 3, we have

$$\begin{aligned}C_\tau &= R\sqrt{2\log\frac{K^2 \det(A_t)^{\frac{1}{2}} \det(\lambda I)^{-\frac{1}{2}}}{\delta}} + \lambda^{\frac{1}{2}} S\\ &\le R\sqrt{2\log\frac{K^2}{\delta} + d\log\left(1 + \frac{\tau L^2}{\lambda d}\right)} + \lambda^{\frac{1}{2}} S\\ &\le 2R\sqrt{2\log\frac{K^2}{\delta} + d\log\left(1 + \frac{\tau L^2}{\lambda d}\right)}.\end{aligned}$$

The second inequality follows from condition $\lambda \le \frac{2R^2}{S^2}\log\frac{K^2}{\delta}$. Therefore, using Lemma 6, we have

$$\begin{aligned}\tau &\le H_\varepsilon C_\tau^2 + K\\ &\le 4H_\varepsilon R^2\left(2\log\frac{K^2}{\delta} + d\log\left(1 + \frac{\tau L^2}{\lambda d}\right)\right) + K.\end{aligned}$$

Let τ' a parameter satisfying

$$\tau = 4H_\varepsilon R^2\left(2\log\frac{K^2}{\delta} + d\log\left(1 + \frac{\tau' L^2}{\lambda d}\right)\right) + K. \tag{35}$$

Then, $\tau' \le \tau$ holds.

For N defined as

$$N = 8H_\varepsilon R^2 \log\frac{K^2}{\delta} + K,$$

we have

$$\begin{aligned}\tau' &\le \tau\\ &= 4H_\varepsilon R^2 d\log\left(1 + \frac{\tau' L^2}{\lambda d}\right) + N\\ &\le 4H_\varepsilon R^2\sqrt{dL^2\tau'/\lambda} + N.\end{aligned}$$

By solving this inequality, we obtain

$$\begin{aligned}\sqrt{\tau'} &\le 4H_\varepsilon R^2\sqrt{dL^2/\lambda} + \sqrt{16H_\varepsilon^2 R^4 dL^2\tau'/\lambda + N^2}\\ &\le 2\sqrt{16H_\varepsilon^2 R^4 dL^2/\lambda + N^2}.\end{aligned}$$

Let M be the right hand side of the inequality:

$$M = 2\sqrt{16H_\varepsilon^2 R^4 dL^2/\lambda + N^2}.$$

Then, using this upper bound of τ' in (35) yields

$$\tau \le K + 8H_\varepsilon R^2\log\frac{K^2}{\delta} + C(H_\varepsilon, \delta),$$

where $C(H_\varepsilon, \delta)$ is denoted as

$$\begin{aligned}C(H_\varepsilon, \delta) &= 4H_\varepsilon R^2 d\log\left(1 + \frac{M^2L^2}{\lambda d}\right) \qquad (36)\\ &= \mathcal{O}\left(H_\varepsilon\log\left(H_\varepsilon\log\frac{1}{\delta}\right)\right)\end{aligned}$$

□

Proof of Theorem 3. By Prop. 3, we have

$$C_\tau \le R\sqrt{2\log\frac{K^2}{\delta} + d\log\left(1 + \frac{\tau L^2}{\lambda d}\right)} + \lambda^{\frac{1}{2}} S.$$

Using the fact $(a+b)^2 \le 2(a^2+b^2)$ and $(1+\frac{1}{x})^x \le e$, we have

$$\begin{aligned}\tau &\le H_\varepsilon C_\tau^2 + K\\ &\le 2H_\varepsilon\left(2R^2\log\frac{K^2}{\delta} + \frac{\tau R^2L^2}{\lambda} + \lambda S^2\right) + K,\end{aligned}$$

from Lemma 6. Therefore, we can conclude that if $\lambda > 4H_\varepsilon R^2L^2$, then

$$\begin{aligned}\tau &\le \left(1 - \frac{2H_\varepsilon R^2L^2}{\lambda}\right)^{-1}\left(4H_\varepsilon R^2\log\frac{K^2}{\delta} + C'\right)\\ &\le 2\left(4H_\varepsilon R^2\log\frac{K^2}{\delta} + C'\right),\end{aligned}$$

where $C' = 2H_\varepsilon\lambda S^2 + K$. □

D.3 Proof of Theorem 4

In this appendix we give the proof of Theorem 4. This follows straightforwardly from the definition of problem complexity H_ε in (13) and the ratio $p_k^*(i,j)$ in (10).

Proof of Theorem 4. First, we bound the $\rho(i,j)$, which is the optimal value of

$$\begin{aligned}\min_{p_k,w_k}\quad & \sum_{k=1}\frac{w_k^2}{p_k}\\ \text{s.t.}\quad & x_i - x_j = \sum_{k=1}^K w_k x_k\\ & \sum_{i=1}^K p_k = 1,\ p_k \ge 0,\ p_k, w_k \in \mathbb{R}. \qquad (37)\end{aligned}$$

Now, since $x_i - x_j = (x_i - x_{a^*}) + (x_{a^*} - x_j)$, p'_k and w'_k defined as

$$\begin{aligned}p'_k &= \frac{p_k^*(i,a^*) + p_k^*(a^*,j)}{2},\\ w'_k &= w_k^*(i,a^*) + w_k^*(a^*,j)\end{aligned}$$

satisfy the condition of (37). Therefore, we have

$$\begin{aligned}\rho(y(i,j)) &\le \sum_{k=1}\frac{(w'_k)^2}{p'_k}\\ &= 2\sum_{k=1}\frac{(w_k^*(i,a^*) + w_k^*(a^*,j))^2}{p_k^*(i,a^*) + p_k^*(a^*,j)}\\ &\le 4\sum_{k=1}\frac{(w_k^*(i,a^*))^2 + (w_k^*(a^*,j))^2}{p_k^*(i,a^*) + p_k^*(a^*,j)}\\ &\le 4\sum_{k=1}\frac{(w_k^*(i,a^*))^2}{p_k^*(i,a^*)} + \frac{(w_k^*(a^*,j))^2}{p_k^*(a^*,j)}\\ &= 4\rho(i,a^*) + 4\rho(a^*,j).\end{aligned}$$

Using this upper bound, we can bound the problem complexity H_0 as follows. Let i_k^* and j_k^* be defined as

$$(i_k^*, j_k^*) = \arg\max_{i,j\in[K]}\frac{p_k^*(i,j)\rho(i,j)}{\max\left(\Delta_i^2, \Delta_j^2\right)}.$$

and we have

$$\begin{aligned}H_0 &= 9\sum_{k=1}^K \max_{i,j\in[K]}\frac{p_k^*(i,j)\rho(i,j)}{\max\left(\Delta_i^2,\Delta_j^2\right)}\\ &= 9\sum_{k=1}^K\frac{p_k^*(i_k^*,j_k^*)\rho(i_k^*,j_k^*)}{\max\left(\Delta_{i_k^*}^2,\Delta_{j_k^*}^2\right)}\\ &\le 36\sum_{k=1}^K p_k^*(i_k^*,j_k^*)\frac{\rho(i_k^*,a^*) + \rho(a^*,j_k^*))}{\max\left(\Delta_{i_k^*}^2,\Delta_{j_k^*}^2\right)}\\ &\le 36\sum_{k=1}^K p_k^*(i_k^*,j_k^*)\left(\frac{\rho(i_k^*,a^*)}{\Delta_{i_k^*}^2} + \frac{\rho(a^*,j_k^*))}{\Delta_{j_k^*}^2}\right)\\ &\le 72H'_{\mathrm{oracle}}.\end{aligned}$$

The last inequality holds from $\sum_{k=1}^K p_k^*(i,j) = 1$ for all $i,j\in[K]$. Now, it is sufficient to prove

$$KH_{\mathrm{oracle}} \ge H'_{\mathrm{oracle}}.$$

This can be derived as follows. By definition, we have

$$\begin{aligned}H_{\mathrm{oracle}} &= \lim_{n\to\infty}\min_{\mathbf{x}_n}\max_{i\in[K]\setminus\{a^*\}}\frac{n\|x_{a^*} - x_i\|^2_{A^{-1}_{\mathbf{x}_n}}}{\Delta_i^2}\\ &\ge \max_{i\in[K]\setminus\{a^*\}}\lim_{n\to\infty}\min_{\mathbf{x}_n}\frac{n\|x_{a^*} - x_i\|^2_{A^{-1}_{\mathbf{x}_n}}}{\Delta_i^2}.\end{aligned}$$

As discussed in Appendix C, $\rho(a^*,i)$ is the optimal value of (23) and is also equal to the limit of the optimal value of (21) as $n\to\infty$ for $y = x_{a^*} - x_i$, that is,

$$\begin{aligned}&\lim_{n\to\infty}\min_{\mathbf{x}_n} n\|x_{a^*} - x_i\|^2_{A^{-1}_{\mathbf{x}_n}}\\ &\quad= \lim_{n\to\infty}\min_{\substack{C_k\in\mathbb{N}\cup\{0\}:\\ \sum_{k=1}^K C_k = n}} y^\top\left(\frac{\lambda}{n}I + \sum_{k=1}^K\frac{C_k}{n}x_k^\top x_k\right)y\\ &\quad= \rho(a^*,i).\end{aligned}$$

Therefore, we have

$$\begin{aligned}H_{\mathrm{oracle}} &\ge \max_{i\in[K]\setminus\{a^*\}}\frac{\rho(a^*,i)}{\Delta_i^2}\\ &\ge \frac{1}{K}H'_{\mathrm{oracle}}.\end{aligned}$$

□