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 by a binary outcome (clicked/not clicked). We use article-user interaction feature $z_{a,u} \in \mathbb{R}^{6}$, 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 $\theta^*$ 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 $\Delta_{t} \geq 0.05$ 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

$$r_{t} = \begin{cases} 
1 & \text{w.p. } \frac{1+x_{a,t}^{T} \theta^*}{2} \\
-1 & \text{(otherwise)}
\end{cases},$$

where $x_{a,t}$ is the feature of the arm selected at the $t$th round. Although it does not always the case, $x^{T} \theta^*$ is happened to be bounded in $[-1,1]$ for all feature $x$ in the dataset, therefore $(1 + x_{a,t}^{T} \theta^*)/2$ is always valid for probability. Furthermore, since $x_{a,t}^{T} \theta^* \in [-1,1]$, the noise variable $\varepsilon_t$ 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 $X^{T}$ 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 $O(K)$ 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 $K' = O(d)$ arms, denoted as $B = \{b_1, b_2, \ldots, b_{K'}\} \subset [K]$, and force the agent to select arms from $B$. In other word, the arm selection strategy in (12) would be

$$a_{t+1} = \underset{a \in B: p_{t}^{*}(i_t, j_t) > 0}{\arg\min} T_{a}(t)/\hat{p}_{a}^{*}(i_t, j_t), \quad (19)$$

where $\hat{p}_{a}^{*}(i_t, j_t)$ is defined as

$$\hat{p}_{a}^{*}(i_t, j_t) = \arg\min_{w \in \mathbb{R}^{d}} \|w\|_1 \quad \begin{cases} 
\frac{|w_{k}^{*}(i_t, j_t)|}{\sum_{k=1}^{K} |w_{k}^{*}(i_t, j_t)|} & (k \in B) \\
0 & (k \notin B)
\end{cases},

\text{s.t. } x_{i_t} - x_{j_t} = \sum_{k=1}^{K} w_{k} x_{k} \\
w_{k} = 0 \quad (\forall k \notin B). \quad (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:

$$\hat{H}_{\varepsilon} = \sum_{k=1}^{K'} \max_{i,j \in [K]} \frac{\hat{p}_{a}^{*}(i, j) \hat{\rho}(i, j)}{\max \left( \varepsilon, \varepsilon + \Delta_{i}, \varepsilon + \Delta_{j} \right)^{2}} \left( \frac{\sqrt{\varepsilon}}{\sigma} \right)^2,$$

where $\hat{\rho}$ is

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

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

Lemma 1. Let $X$ be the matrix

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

and $\sigma$ be the smallest eigenvalue of $X^{T} X$. Then, problem complexity $\hat{H}_{\varepsilon}$ is bounded as follows:

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

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


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

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

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

\[
\begin{align*}
\tilde{p}(i, j) &= \sum_{k=1}^{K} |\tilde{w}_k^+(i, j)| \\
&= \sqrt{\left( \sum_{k=1}^{K} |\tilde{w}_k^+(i, j)| \right)^2} \\
&\leq \sqrt{K'} \left( \sum_{k=1}^{K} |\tilde{w}_k^+(i, j)|^2 \right) \\
&= \sqrt{K'} \left( \tilde{w}^+(i, j) \right)^\top \tilde{w}^+(i, j) \\
&\leq \sqrt{K'} \left( \tilde{w}^+(i, j) \right)^\top \tilde{X}^\top \tilde{X} \tilde{w}^+(i, j) \\
&\leq \sqrt{K'} \left( \tilde{X}^\top \tilde{X} \right) \frac{2L}{\sigma} \\
&\leq 2L \sqrt{\frac{K'}{\sigma}}.
\end{align*}
\]

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

\[
\tilde{H}_\varepsilon \leq \sum_{k=1}^{K'} \text{max}_{y \in [K]} \frac{2L}{\sigma} \left( \frac{K'}{\sigma} \right)^2 = O(d \sqrt{d}),
\]

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 \( \gamma \in \mathcal{X} = \{ x - x' \mid x, x' \in \mathcal{X} \} \) is to obtain

\[
\arg \min_{x_n} \| y \| (A_{x_n})^{-1}.
\]

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

\[
\begin{align*}
\min_{C_i \in \mathbb{N} \cup \{ 0 \}} \quad & y^\top \left( \frac{\lambda}{n} I + \sum_{i=1}^{K} \frac{C_i}{n} x_i x_i^\top \right)^{-1} y \
\text{s.t.} \quad & \sum_{i=1}^{K} C_i = n.
\end{align*}
\]

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

\[
\begin{align*}
\min_{p_i \geq 0} \quad & y^\top \left( \frac{\lambda}{n} I + \sum_{i=1}^{K} p_i x_i x_i^\top \right)^{-1} y \
\text{s.t.} \quad & \sum_{i=1}^{K} p_i = 1, \quad \text{where } p_i \text{ corresponds to the ratio } C_i/n.
\end{align*}
\]

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

\[
\begin{align*}
\min_{p_i, w_i} \quad & \| y - \sum_{i=1}^{K} w_i x_i \| \quad \text{such that } \| y - \sum_{i=1}^{K} w_i x_i \| > 0 \\
\text{s.t.} \quad & \sum_{i=1}^{K} p_i = 1, \quad p_i \geq 0, \quad w_i \in \mathbb{R},
\end{align*}
\]

(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 \( \lambda/n \) 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

\[
\begin{align*}
\min_{p_i, w_i} \quad & \sum_{i=1}^{K} p_i w_i^2 \\
\text{s.t.} \quad & y = \sum_{i=1}^{K} w_i x_i
\end{align*}
\]

(23)

the KKT condition of which yields the definition of \( p^* \) and \( w^* \) 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 \( \| x_i - x_j \|_{A^{-1}} \) 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 \( \| x_i - x_j \|_{A^{-1}} \) is bounded by

\[
\| x_i - x_j \|_{A^{-1}} \leq \sqrt{\frac{\rho(i, j)}{T_{i,j}(t)}}.
\]
where \( T_{i,j}(t) = \min_{k \in [K]: p_k^*(i,j) > 0} T_k(t)/p_k^*(i,j) \).

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 \( \alpha > 0 \),

\[
y^\top (A + \alpha xx^\top)^{-1} y \leq y^\top A^{-1} y
\]

holds.

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

\[
y^\top (A + \alpha xx^\top)^{-1} y = y^\top \left( A^{-1} - \frac{\alpha A^{-1} xx^\top A^{-1}}{1 + \alpha xx^\top A^{-1} x} \right) y
\]

\[
= y^\top A^{-1} y - y^\top \frac{\alpha A^{-1} xx^\top A^{-1}}{1 + \alpha xx^\top A^{-1} x} y
\]

\[
\leq y^\top A^{-1} y.
\]

The last inequality follows from the fact that \( A^{-1} \) is positive definite. \( \square \)

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

**Proof of Lemma 2.** By the definition of \( A_t \), we have

\[ A_t = \lambda I + \sum_{k=1}^{K} T_k(t)x_kx_k^\top. \]

Then, for

\[ \tilde{A}_t = \lambda I + \sum_{k=1}^{K} p_k^*(i,j)T_{i,j}(t)x_kx_k^\top, \]

we have

\[ \|x_i - x_j\|_{\tilde{A}_t^{-1}} \leq \|x_i - x_j\|_{A_t^{-1}} \]

from Lemma 3 and the fact

\[ T_k(t) \leq p_k^*(i,j)T_{i,j}(t), \]

which can be inferred from the definition of \( T_{i,j}(t) \). Therefore, proving

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

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

\[
w_k^*(i,j) = \frac{1}{2} p_k^*(i,j)x_k^\top \gamma
\]

\[
y = \frac{1}{2} \sum_{k=1}^{K} p_k^*(i,j)x_kx_k^\top \gamma,
\]

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

\[
\rho(i,j) = \frac{1}{4} \gamma^\top \left( \sum_{k=1}^{K} p_k^*(i,j)x_kx_k^\top \right) \gamma.
\]

Now, let \( B \) be denoted as

\[
B = \left( \sum_{k=1}^{K} p_k^*(i,j)x_kx_k^\top \right).
\]

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

\[
y^\top \tilde{A}_t^{-1} y - \rho(y) = \frac{1}{4} \gamma^\top B^\top \tilde{A}_t^{-1} B\gamma - \frac{1}{4T_{i,j}(t)} \gamma^\top B\gamma
\]

\[
= \frac{1}{4} \gamma^\top \left( B^\top - \frac{\tilde{A}_t}{T_{i,j}(t)} \right) \tilde{A}_t^{-1} B\gamma
\]

\[
= -\frac{1}{4} \gamma^\top \frac{\lambda}{T_{i,j}(t)} \tilde{A}_t^{-1} B\gamma
\]

\[ \leq 0. \]

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

### 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

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

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 \( \tau \) 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 \( \varepsilon \), then we have

\[ \Delta(a^*, \hat{a}^*) > \varepsilon \geq B(\tau) \geq \hat{\Delta}_\tau(a^*, \hat{a}^*) + \beta_\tau(a^*, \hat{a}^*). \]

The second inequality holds for stopping condition \( B(\tau) \leq \varepsilon \) 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 \( \mathcal{E} \)
On the other hand, in the case where (a), (b), (c) and (d) follow from the definition of $\Delta_a$. Thus, the probability that LinGapE returns such arms is

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

where $\hat{E}$ 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

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

Otherwise, we have

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

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

**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

$$\hat{\Delta}_t(j_t, i_t) = (x_{j_t} - x_{i_t})^\top \theta^*_t \leq 0. \quad (24)$$

Thus, $B(t)$ is bounded by

$$B(t) = \hat{\Delta}_t(j_t, i_t) + \beta_t(i_t, j_t) \leq \beta_t(i_t, j_t). \quad (25)$$

Therefore, it is sufficient to show

$$B(t) \leq -(\Delta_{i_t} \lor \Delta_{j_t}) + 2\beta_t(i_t, j_t). \quad (26)$$

If $i_t = a^*$, then

$$(\Delta_{i_t} \lor \Delta_{j_t}) = \Delta_{j_t} \quad (27)$$

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

$$B(t) \overset{(a)}{=} \hat{\Delta}_t(j_t, i_t) + \beta_t(i_t, j_t) \overset{(b)}{=} \Delta_t(j_t, i_t) + 2\beta_t(i_t, j_t) \overset{(c)}{=} -\Delta_{j_t} + 2\beta_t(i_t, j_t) \overset{(d)}{=} -(\Delta_{i_t} \lor \Delta_{j_t}) + 2\beta_t(i_t, j_t),$$

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

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

$$(\Delta_{i_t} \lor \Delta_{j_t}) = \Delta_{i_t}. \quad (28)$$

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

$$B(t) \overset{(a)}{=} \beta_t(i_t, j_t) \overset{(b)}{=} -\hat{\Delta}_t(j_t, i_t) + \beta_t(i_t, j_t) \overset{(c)}{=} -\Delta(j_t, i_t) + 2\beta_t(i_t, j_t) \overset{(d)}{=} -(\Delta_{i_t} \lor \Delta_{j_t}) + 2\beta_t(i_t, j_t),$$

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

$$B(t) \leq -(\Delta_{i_t} \lor \Delta_{j_t}) + 3\beta_t(i_t, j_t). \quad (29)$$

Since $j_t \neq a^*$,

$$\hat{\Delta}_t(a^*, i_t) + \beta_t(a^*, i_t) \leq \Delta_t(j_t, i_t) + \beta_t(j_t, i_t). \quad (30)$$

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

$$\beta_t(i_t, j_t) \overset{(a)}{=} \hat{\Delta}_t(j_t, i_t) + \beta_t(j_t, i_t) \overset{(b)}{=} \hat{\Delta}_t(a^*, i_t) + \beta_t(a^*, i_t) \overset{(c)}{=} \Delta(a^*, i_t). \quad (31)$$

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

$$B(t) = \hat{\Delta}_t(j_t, i_t) + \beta_t(i_t, j_t) \leq \Delta(j_t, i_t) + 2\beta_t(i_t, j_t) = \Delta(j_t, a^*) + \Delta(a^*, i_t) + 2\beta_t(i_t, j_t) \leq -\Delta_{j_t} + 3\beta_t(i_t, j_t). \quad (32)$$

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

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

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 $\tau$ be the stopping time of LinGapE when $a_t$ is determined by (12). Then, statement

$$\tau \leq H_n C^2_t + K \quad (34)$$

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 $E = \mathcal{E}$ occurs. First we derive the upper bound that arm $k$ is pulled. Then,
\[
\min(0, -\Delta_k + 2\beta_{i-1}(i_{i-1}, j_{i-1})) + \beta_{i-1}(i_{i-1}, j_{i-1}) \\
\geq B(i - 1) \geq \varepsilon
\]
follows from Lemma 5 and the fact that stopping condition is not satisfied at the $\hat{t}$-th round. Applying Lemma 2 yields
\[
T_{i_{i-1}, j_{i-1}}(\hat{t} - 1) \leq \frac{\rho(i_{i-1}, j_{i-1})}{\max(\varepsilon, \frac{e^{+\Delta_k}}{3}, \frac{e^{+\Delta_{i-1}}}{3})^{2}C_{t-1}},
\]
where $C_t$ is defined in (3). Now, since arm $k$ is pulled at $t$-th round,
\[
T_{k}(\hat{t} - 1) = p_k(i_{i-1}, j_{i-1})T_{i_{i-1}, j_{i-1}}(\hat{t} - 1)
\]
holds by definition. Therefore, $T_k(\tau)$ can be bounded as
\[
T_k(\tau) = T_k(\hat{t} - 1) + 1 \\
= p_k(i_{i-1}, j_{i-1})T_{i_{i-1}, j_{i-1}}(\hat{t} - 1) + 1 \\
\leq \max_{i,j \in [K]} p_k(i, j)T_{i_{i}, j_{i}}(\hat{t} - 1) + 1 \\
\leq \frac{p_k(i_{i-1}, j_{i-1})p(i_{i-1}, j_{i-1})}{\max(\varepsilon, \frac{e^{+\Delta_k}}{3}, \frac{e^{+\Delta_{i-1}}}{3})^{2}C_{t-1}} + 1 \\
\leq \max_{i,j \in [K]} \frac{p_k(i, j)p(i, j)}{\max(\varepsilon, \frac{e^{+\Delta_k}}{3}, \frac{e^{+\Delta_{i-1}}}{3})^{2}C_{t}^{2} + 1.
\]
Since $\sum_{k=1}^{K} T_k(\tau) = \tau$, summing up the upper bound of $T_k(t)$ above yields
\[
\tau \leq H_{\varepsilon}C_{t}^{2} + K.
\]

Now, we can complete the proofs by bounding $C_{t}$ 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
\[
\det(A_{n}^{\lambda}) \leq (\lambda + nL^2/d)^{d}.
\]

**Proof of Theorem 2.** From Proposition 3, we have
\[
C_{t} = R \sqrt{2 \log \frac{K^2 \det(A_{n}^{\lambda})}{\delta} \frac{\lambda^{1/2} \det(\lambda^{1/2})^{-1/2}}{\delta}} + \lambda^{1/2} S \\
\leq R \sqrt{2 \log \frac{K^2}{\delta} + d \log \left(1 + \frac{\tau L^2}{\lambda d}\right)} + \lambda^{1/2} S \\
\leq 2R \sqrt{2 \log \frac{K^2}{\delta} + d \log \left(1 + \frac{\tau L^2}{\lambda d}\right)}.
\]

The second inequality follows from condition $\lambda \leq \frac{2K^2}{d^{2}} \log \frac{K^2}{\delta}$. Therefore, using Lemma 6, we have
\[
\tau \leq H_{\varepsilon}C_{t}^{2} + K \\
\leq 4H_{\varepsilon}R^2 \left(2 \log \frac{K^2}{\delta} + d \log \left(1 + \frac{\tau L^2}{\lambda d}\right)\right) + K.
\]

Let $\tau'$ a parameter satisfying
\[
\tau = 4H_{\varepsilon}R^2 \left(2 \log \frac{K^2}{\delta} + d \log \left(1 + \frac{\tau L^2}{\lambda d}\right)\right) + K.
\]

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

For $N$ defined as
\[
N = 8H_{\varepsilon}R^2 \log \frac{K^2}{\delta} + K,
\]
we have
\[
\tau' \leq \tau \\
= 4H_{\varepsilon}R^2d \log \left(1 + \frac{\tau' L^2}{\lambda d}\right) + N \\
\leq 4H_{\varepsilon}R^2 \sqrt{dL^2\tau'/\lambda} + N.
\]

By solving this inequality, we obtain
\[
\sqrt{\tau'} \leq 4H_{\varepsilon}R^2 \sqrt{dL^2/\lambda} + \sqrt{16H_{\varepsilon}^2R^4d^2L^2\tau'/\lambda + N^2} \\
\leq 2\sqrt{16H_{\varepsilon}^2R^4d^2L^2/\lambda + N^2}.
\]

Let $M$ be the right hand side of the inequality:
\[
M = 2\sqrt{16H_{\varepsilon}^2R^4d^2L^2/\lambda + N^2}.
\]

Then, using this upper bound of $\tau'$ in (35) yields
\[
\tau \leq K + 8H_{\varepsilon}R^2 \log \frac{K^2}{\delta} + C(H_{\varepsilon}, \delta),
\]
where $C(H_{\varepsilon}, \delta)$ is denoted as
\[
C(H_{\varepsilon}, \delta) = 4H_{\varepsilon}R^2d \log \left(1 + \frac{M^2L^2}{\lambda d}\right) = \mathcal{O}\left(\frac{H_{\varepsilon}^2 \log \frac{1}{\delta}}{\delta}\right)
\]

**Proof of Theorem 3.** By Prop. 3, we have
\[
C_{t} \leq R \sqrt{2 \log \frac{K^2}{\delta} + d \log \left(1 + \frac{\tau L^2}{\lambda d}\right)} + \lambda^{1/2} S.
\]
Using the fact \((a+b)^2 \leq 2(a^2 + b^2)\) and \((1 + \frac{1}{n})^n \leq e\), we have

\[
\tau \leq H_x C_x^2 + K \\
\leq 2H_x \left(2R^2 \log \frac{K^2}{\delta} + \frac{\tau R^2 L^2}{\lambda} + \lambda S^2\right) + K,
\]
from Lemma 6. Therefore, we can conclude that if \(\lambda > 4H_x R^2 L^2\), then

\[
\tau \leq \left(1 - \frac{2H_x R^2 L^2}{\lambda}\right)^{-1} \left(4H_x R^2 \log \frac{K^2}{\delta} + C'\right) \\
\leq 2 \left(4H_x R^2 \log \frac{K^2}{\delta} + C'\right),
\]
where \(C' = 2H_x \lambda S^2 + K\). \(\square\)

### 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 \(H_x\) 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

\[
\min_{p_k, w_k} \sum_{k=1}^{K} w_k p_k \\
\text{s.t. } x_i - x_j = \sum_{k=1}^{K} w_k x_k \\
\sum_{i=1}^{K} p_k = 1, p_k \geq 0, p_k, w_k \in \mathbb{R}. \quad (37)
\]

Now, since \(x_i - x_j = (x_i - x_{a^*}) + (x_{a^*} - x_j)\), \(p_k'\) and \(w_k'\) defined as

\[
p_k' = \frac{p_k(i,a^*) + p_k'(a^*,j)}{2}, \quad w_k' = w_k(i,a^*) + w_k'(a^*,j)
\]
satisfy the condition of (37). Therefore, we have

\[
\rho(y(i,j)) \leq \sum_{k=1}^{K} \frac{(w_k')^2}{p_k'} \\
= 2 \sum_{k=1}^{K} \left(\frac{w_k(i,a^*) + w_k'(a^*,j)}{2}\right)^2 \\
\leq 4 \sum_{k=1}^{K} \left(p_k(i,a^*) + p_k'(a^*,j)\right)^2 \\
\leq 4 \sum_{k=1}^{K} \left(p_k(i,a^*)\right)^2 + \left(w_k'(a^*,j)\right)^2 \\
\leq 4 \rho(i,a^*) + 4 \rho(a^*,j).
\]

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

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

and we have

\[
H_0 = 9 \sum_{k=1}^{K} \max_{i,j \in [K]} p_k(i,j)\rho(i,j) \\
\max \left(\Delta_i^2, \Delta_j^2\right) \\
= 9 \sum_{k=1}^{K} \frac{\rho(i_k^*,a^*) + \rho(a^*,j_k^*)}{\Delta_i^2, \Delta_j^2} \\
\leq 36 \sum_{k=1}^{K} \frac{\rho(i_k^*,a^*), \rho(a^*,j_k^*)}{\Delta_i^2, \Delta_j^2} \\
\leq 36 \sum_{k=1}^{K} \frac{\rho(i_k^*,a^*)}{\Delta_i^2} + \frac{\rho(a^*,j_k^*)}{\Delta_j^2} \\
\leq 72H_{oracle}'.
\]

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

\[
KH_{oracle} \geq H_{oracle}'.
\]

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

\[
H_{oracle} = \lim_{n \to \infty} \min_{\mathbf{x}_n} \max_{i \in [K] \setminus \{a^*\}} \frac{n\|x_{a^*} - x_i\|_{A_{a^*}}^2}{\Delta_i^2} \\
\geq \lim_{n \to \infty} \min_{i \in [K] \setminus \{a^*\}} \frac{n\|x_{a^*} - x_i\|_{A_{a^*}}^2}{\Delta_i^2}.
\]

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,

\[
\lim_{n \to \infty} \min_{\mathbf{x}_n} \frac{n\|x_{a^*} - x_i\|_{A_{a^*}}^2}{\Delta_i^2} \\
= \lim_{n \to \infty} \min_{\mathbf{x}_n} \frac{y^\top \left(\frac{\lambda}{n} I + \sum_{k=1}^{K} \frac{C_k}{n} x_k x_k^\top\right) y}{\Delta_i^2} \\
= \rho(a^*,i).
\]

Therefore, we have

\[
H_{oracle} \geq \min_{i \in [K] \setminus \{a^*\}} \frac{\rho(a^*,i)}{\Delta_i^2} \\
\geq \frac{1}{K} H_{oracle}'. \quad \square
\]