# Supplementary Material for
# 'Thompson Sampling for Unsupervised Sequential Selection'

## Appendix A. Useful results needed to prove regret bounds of USS-TS

We use the following results in our proofs.

**Fact 2** (Chernoff bound for Bernoulli distributed random variables)**.** *Let $X_1, \ldots, X_n$ be i.i.d. Bernoulli distributed random variables. Let $\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^n X_i$ and $\mu = \mathbb{E}[X_i]$. Then, for any $\varepsilon \in (0, 1-\mu)$,*

$$\mathbb{P}\{\hat{\mu}_n \geq \mu + \varepsilon\} \leq \exp(-d(\mu+\varepsilon, \mu)n),$$

*and, for any $\varepsilon \in (0, \mu)$,*

$$\mathbb{P}\{\hat{\mu}_n \leq \mu - \varepsilon\} \leq \exp(-d(\mu-\varepsilon, \mu)n),$$

*where $d(x,\mu) = x\log\left(\frac{x}{\mu}\right) + (1-x)\log\left(\frac{1-x}{1-\mu}\right)$.*

See Section 10.1 of Chapter 10 of book 'Bandit Algorithms' (Lattimore and Szepesvári, 2020) for proof.

**Fact 3** (Pinsker's Inequality for Bernoulli distributed random variables)**.** *For $p, q \in (0,1)$, the KL divergence between two Bernoulli distributions is bounded as:*

$$d(p,q) \geq 2(p-q)^2.$$

**Fact 4.** *Let $x > 0$ and $D > 0$. Then, for any $a \in (0,1)$,*

$$\frac{1}{\exp^{Dx} - 1} \leq \begin{cases} \frac{\exp^{-Dx}}{1-a} & (x \geq \ln(1/a)/D) \\ \frac{1}{Dx} & (x < \ln(1/a)/D). \end{cases}$$

*Further, we have,*

$$\sum_{x=1}^{n} \frac{1}{\exp^{Dx} - 1} \leq \Theta\left(\frac{1}{D^2} + \frac{1}{D}\right).$$

*Proof.* Using $\exp^y \geq y+1$ (by Taylor Series expansion), we have $\frac{1}{\exp^{Dx}-1} \leq \frac{1}{Dx}$ as $\exp^{Dx} - 1 \geq Dx$. We can re-write, $\frac{1}{\exp^{Dx}-1} = \frac{\exp^{-Dx}}{1-\exp^{-Dx}}$. Since $\exp^{-Dx}$ is strictly decreasing function for all $Dx > 0$, it is easy to check that $\exp^{-Dx} \leq a$ holds for any $x \geq \ln(1/a)/D$ and $a \in (0,1)$. Hence, $\frac{\exp^{-Dx}}{1-\exp^{-Dx}} \leq \frac{\exp^{-Dx}}{1-a}$ for all $x \geq \ln(1/a)/D$.

Now we will prove the second part,

$$\sum_{x=1}^{n} \frac{1}{\exp^{Dx}-1} \leq \frac{\ln(1/a)}{D^2} + \sum_{x \geq \ln(1/a)/D}^{n} \frac{\exp^{-Dx}}{1-a}$$

$$
\begin{aligned}
&\leq \frac{\ln(1/a)}{D^2} + \frac{1}{(1-a)} \int_{x=0}^{\infty} \exp^{-Dx} dx \\
&= \frac{\ln(1/a)}{D^2} + \frac{1}{(1-a)} \left( \frac{\exp^{-Dx}}{-D} \right) \Big|_{x=0}^{\infty} \\
&= \frac{\ln(1/a)}{D^2} + \frac{1}{(1-a)} \left( 0 - \frac{\exp^{0}}{-D} \right) \\
&= \frac{\ln(1/a)}{D^2} + \frac{1}{(1-a)D}
\end{aligned}
$$

$$
\implies \sum_{x=1}^{n} \frac{1}{\exp^{Dx} - 1} \leq \Theta\left( \frac{1}{D^2} + \frac{1}{D} \right). \qquad \square
$$

**Fact 5.** *Let* $\varepsilon \in (0,1)$ *and* $0 < x < y < z < 1$. *If* $d(y,z) = d(x,z)/(1+\varepsilon)$ *then*

$$
y - x \geq \frac{\varepsilon}{1+\varepsilon} \cdot \frac{d(x,z)}{\ln\left(\frac{z(1-x)}{x(1-z)}\right)}.
$$

*Proof.* By definition

$$
\begin{aligned}
d(p,q) &= p \ln \frac{p}{q} + (1-p) \ln \left( \frac{1-p}{1-q} \right) \\
&= \ln \left( \left( \frac{p}{q} \right)^{p} \left( \frac{1-p}{1-q} \right)^{1-p} \right) \\
&= \ln \left( \left( \frac{q(1-p)}{p(1-q)} \right)^{-p} \right) + \ln \left( \frac{1-p}{1-q} \right) \\
\implies d(p,q) &= -p \ln \left( \frac{q(1-p)}{p(1-q)} \right) + \ln \left( \frac{1-p}{1-q} \right).
\end{aligned}
$$

Set $l(p,q) = \ln\left(\frac{q(1-p)}{p(1-q)}\right)$. Note that $l(p,\cdot)$ is a strictly decreasing function of $p$ and positive for all $p < q$. We can re-arrange above equation as

$$
p \cdot l(p,q) = -d(p,q) + \ln \left( \frac{1-p}{1-q} \right).
$$

Using above equation, we have

$$
y \cdot l(y,z) - x \cdot l(x,z) = -d(y,z) + \ln \left( \frac{1-y}{1-z} \right) + d(x,z) - \ln \left( \frac{1-x}{1-z} \right).
$$

Using $d(y,z) = d(x,z)/(1+\varepsilon)$,

$$
y \cdot l(y,z) - x \cdot l(x,z) = \frac{\varepsilon}{1+\varepsilon} d(x,z) + \ln \left( \frac{1-y}{1-x} \right).
$$

After adding $y(l(x,z) - l(y,z))$ both side, we have

$$
(y-x) l(x,z) = \frac{\varepsilon}{1+\varepsilon} d(x,z) + \ln \left( \frac{1-y}{1-x} \right) + y(l(x,z) - l(y,z)).
$$

Using $l(x,z) = \ln\left(\frac{z(1-x)}{x(1-z)}\right)$ and $l(y,z) = \ln\left(\frac{z(1-y)}{y(1-z)}\right)$

$$
\begin{aligned}
&= \frac{\varepsilon}{1+\varepsilon} d(x,z) + \ln\left(\frac{1-y}{1-x}\right) + y \ln\left(\frac{y(1-x)}{x(1-y)}\right) \\
&= \frac{\varepsilon}{1+\varepsilon} d(x,z) + \ln\left(\left(\frac{y(1-x)}{x(1-y)}\right)^{y} \cdot \frac{1-y}{1-x}\right) \\
&= \frac{\varepsilon}{1+\varepsilon} d(x,z) + \ln\left(\left(\frac{y}{x}\right)^{y}\left(\frac{1-y}{1-x}\right)^{1-y}\right) \\
&= \frac{\varepsilon}{1+\varepsilon} d(x,z) + d(y,x)
\end{aligned}
$$

As $d(p,q) \geq 0$ and dividing both side by $l(x,z)$,

$$
\implies y - x \geq \frac{\varepsilon}{1+\varepsilon} \cdot \frac{d(x,z)}{l(x,z)}.
$$

Substituting value of $l(x,z)$ in the above equation, we get

$$
y - x \geq \frac{\varepsilon}{1+\varepsilon} \cdot \frac{d(x,z)}{\ln\left(\frac{z(1-x)}{x(1-z)}\right)}.
$$

□

## Appendix B. Leftover proofs from Section 4

**Lemma 4.** *Let $P \in \mathcal{P}_{\mathrm{WD}}$ and satisfies the transitivity property. If $s$ be the number of times the sub-optimal arm $j$ is selected by USS-TS then, for any $j < i^\star$,*

$$
\sum_{t=1}^{T} \mathbb{P}\left\{I_t = j, j < i^\star\right\} \leq \frac{24}{\xi_j^2} + \sum_{s \geq 8/\xi_j} \Theta\left(\exp^{-s\xi_j^2/2} + \frac{\exp^{-sd(p_{i^\star j} - \xi_j, p_{i^\star j})}}{(s+1)\xi_j^2} + \frac{1}{\exp^{s\xi_j^2/4} - 1}\right).
$$

*Proof.* Applying Lemma 3 and properties of conditional expectations, we have

$$
\sum_{t=1}^{T} \mathbb{P}\left\{I_t = j, j < i^\star\right\} = \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{P}\left\{I_t = j, j < i^\star | \mathcal{H}_t\right\}\right].
$$

As $q_{j,t}$ is fixed given $\mathcal{H}_t$,

$$
\begin{aligned}
\implies \sum_{t=1}^{T} \mathbb{P}\left\{I_t = j, j < i^\star\right\} &\leq \sum_{t=1}^{T} \mathbb{E}\left[\frac{(1-q_{j,t})}{q_{j,t}} \mathbb{P}\left\{I_t \geq i^\star | \mathcal{H}_t\right\}\right] \\
&\leq \sum_{t=1}^{T} \mathbb{E}\left[\mathbb{E}\left[\frac{(1-q_{j,t})}{q_{j,t}} \mathbb{1}_{\{I_t \geq i^\star\}} | \mathcal{H}_t\right]\right].
\end{aligned}
$$

Using law of iterated expectations,

$$
\implies \sum_{t=1}^{T} \mathbb{P}\left\{I_t = j, j < i^\star\right\} \leq \sum_{t=1}^{T} \mathbb{E}\left[\frac{(1-q_{j,t})}{q_{j,t}} \mathbb{1}_{\{I_t \geq i^\star\}}\right]. \tag{14}
$$

Let $s_m$ denote the time step at which the output of arm $i^\star$ is observed for the $m^{th}$ time for $m \geq 1$, and let $s_0 = 0$. For $j < i^\star$, whenever the output from arm $i^\star$ is observed then the output from arm $j$ is also observed due to the cascade structure. Note that $q_{j,t} = \mathbb{P}\left\{\tilde{p}_{i^\star j}^{(t)} > p_{i^\star j} - \xi_j | \mathcal{H}_t\right\}$ changes only when the distribution of $\tilde{p}_{i^\star j}^{(t)}$ changes, that is, only on the time step when the feedback from arms $i^\star$ and $j$ are observed. It only happens when selected arm $I_t \geq i^\star$. Hence, $q_{j,t}$ is the same at all time steps $t \in \{s_m + 1, \ldots, s_{m+1}\}$ for every $m$. Using this fact, we can decompose the right hand side term in Eq. (14) as follows,

$$\begin{aligned}\sum_{t=1}^{T} \mathbb{E}\left[\frac{(1-q_{j,t})}{q_{j,t}} \mathbb{1}_{\{I_t \geq i^\star\}}\right] &= \sum_{m=0}^{T-1} \mathbb{E}\left[\frac{(1-q_{j,s_m+1})}{q_{j,s_m+1}} \sum_{t=s_m+1}^{s_{m+1}} \mathbb{1}_{\{I_t \geq i^\star\}}\right] \\ &\leq \sum_{m=0}^{T-1} \mathbb{E}\left[\frac{(1-q_{j,s_m+1})}{q_{j,s_m+1}}\right] \\ &= \sum_{k=0}^{T-1} \mathbb{E}\left[\frac{1}{q_{j,s_m+1}} - 1\right].\end{aligned}$$

Using above bound in Eq. (14), we get

$$\sum_{t=1}^{T} \mathbb{P}\left\{I_t = j, j < i^\star\right\} \leq \sum_{m=0}^{T-1} \mathbb{E}\left[\frac{1}{q_{j,s_m+1}} - 1\right].$$

Substituting the bound from Lemma 2 with $\mu = p_{i^\star j}, x = p_{i^\star j} - \xi_j, \Delta(x) = \xi_j$, and $q_n(x) = q_{j,s_m}$, we obtain the following bound,

$$\sum_{t=1}^{T} \mathbb{P}\left\{I_t = j, j < i^\star\right\} \leq \frac{24}{\xi_j^2} + \sum_{s \geq 8/\xi_j} \Theta\left(\exp^{-s\xi_j^2/2} + \frac{\exp^{-sd(p_{i^\star j} - \xi_j, p_{i^\star j})}}{(s+1)\xi_j^2} + \frac{1}{\exp^{s\xi_j^2/4} - 1}\right). \quad \square$$

**Lemma 6.** *For any $x_j > p_{i^\star j}$,*

$$\sum_{t=1}^{T} \mathbb{P}\left\{\hat{p}_{i^\star j}^{(t)} > x_j\right\} \leq \frac{1}{d(x_j, p_{i^\star j})}.$$

*Proof.* Let $s_m$ denote the time step at which the outputs of arm $i^\star$ and $j$ is observed for the $m^{th}$ time for $m \geq 1$, and let $s_0 = 0$. Note that probability $\mathbb{P}\left\{\hat{p}_{i^\star j}^{(t)} > x_j\right\}$ changes when the outputs from both arm $i^\star$ and $j$ are observed. Hence, we have

$$\begin{aligned}\sum_{t=1}^{T} \mathbb{P}\left\{\hat{p}_{i^\star j}^{(t)} > x_j\right\} &\leq \sum_{m=0}^{T-1} \mathbb{P}\left\{\hat{p}_{i^\star j}(s_{m+1}) > x_j\right\} \\ &= \sum_{m=0}^{T-1} \mathbb{P}\left\{\hat{p}_{i^\star j}(s_{m+1}) - p_{i^\star j} > x_j - p_{i^\star j}\right\} \\ &\leq \sum_{m=0}^{T-1} \exp^{-kd(p_{i^\star j} + x_j - p_{i^\star j}, p_{i^\star j})} \qquad \text{(using Fact 2)}\end{aligned}$$

$$= \sum_{m=0}^{T-1} \exp^{-kd(x_j, p_{i^\star j})}.$$

Using $\sum_{s \geq 0} \exp^{-sa} \leq 1/a$, we get

$$\sum_{t=1}^{T} \mathbb{P}\left\{\hat{p}_{i^\star j}^{(t)} > x_j\right\} \leq \frac{1}{d(x_j, p_{i^\star j})}. \qquad \square$$

**Lemma 7.** *Let $P \in \mathcal{P}_{\text{WD}}$. For any $\varepsilon > 0$ and $j > i^\star$,*

$$\sum_{t=1}^{T} \mathbb{P}\left\{j \succ_t i^\star, j > i^\star\right\} \leq (1+\varepsilon)\frac{\ln T}{d(p_{i^\star j}, p_{i^\star j} + \xi_j)} + O\left(\frac{1}{\varepsilon^2}\right).$$

*Proof.* Let $p_{i^\star j} < x_j < y_j < p_{i^\star j} + \xi_j$ for any $j > i^\star$. Than,

$$\begin{aligned}
\sum_{t=1}^{T} \mathbb{P}\left\{j \succ_t i^\star, j > i^\star\right\} &= \sum_{t=1}^{T} \mathbb{P}\left\{\tilde{p}_{i^\star j}^{(t)} > p_{i^\star j} + \xi_j\right\} \\
&\leq \sum_{t=1}^{T} \mathbb{P}\left\{\tilde{p}_{i^\star j}^{(t)} > y_j\right\} \\
&\leq \sum_{t=1}^{T} \mathbb{P}\left\{\hat{p}_{i^\star j}^{(t)} \leq x_j, \tilde{p}_{i^\star j}^{(t)} > y_j\right\} + \sum_{t=1}^{T} \mathbb{P}\left\{\hat{p}_{i^\star j}^{(t)} > x_j\right\}.
\end{aligned}$$

Using Lemma 6 and Lemma 5, we have

$$\sum_{t=1}^{T} \mathbb{P}\left\{j \succ_t i^\star, j > i^\star\right\} \leq \frac{\ln T}{d(x_j, y_j)} + 1 + \frac{1}{d(x_j, p_{i^\star j})}.$$

For $\varepsilon \in (0,1)$, we set $x_j \in (p_{i^\star j}, p_{i^\star j} + \xi_j)$ such that $d(x_j, p_{i^\star j} + \xi_j) = d(p_{i^\star j}, p_{i^\star j} + \xi_j)/(1+\varepsilon)$, and set $y_j \in (x_j, p_{i^\star j} + \xi_j)$ such that $d(x_j, y_j) = d(x_j, p_{i^\star j} + \xi_j)/(1+\varepsilon) = d(p_{i^\star j}, p_{i^\star j} + \xi_j)/(1+\varepsilon)^2$. Then this gives

$$\frac{\ln(T)}{d(x_j, y_j)} = (1+\varepsilon)^2 \frac{\ln(T)}{d(p_{i^\star j}, p_{i^\star j} + \xi_j)}.$$

Using Fact 5, if $\varepsilon \in (0,1)$, $x_j \in (p_{i^\star j}, p_{i^\star j} + \xi_j)$, and $d(x_j, p_{i^\star j} + \xi_j) = d(p_{i^\star j}, p_{i^\star j} + \xi_j)/(1+\varepsilon)$ then

$$x_j - p_{i^\star j} \geq \frac{\varepsilon}{1+\varepsilon} . \frac{d(p_{i^\star j}, p_{i^\star j} + \xi_j)}{\ln\left(\frac{(p_{i^\star j} + \xi_j)(1 - p_{i^\star j})}{p_{i^\star j}(1 - p_{i^\star j} - \xi_j)}\right)}.$$

Using Pinsker's Inequality (Fact 3), $1/d(x_j, p_{i^\star j}) \leq 1/2(x_j - p_{i^\star j})^2 = O(1/\varepsilon^2)$ where big-Oh is hiding functions of the $p_{i^\star j}$ and $\xi_j$,

$$\sum_{t=1}^{T} \mathbb{P}\left\{j \succ_t i^\star, j > i^\star\right\} \leq (1+\varepsilon)^2 \frac{\ln(T)}{d(p_{i^\star j}, p_{i^\star j} + \xi_j)} + O\left(\frac{1}{\varepsilon^2}\right)$$

$$\leq (1+3\varepsilon)\frac{\ln(T)}{d(p_{i^\star j}, p_{i^\star j}+\xi_j)} + O\left(\frac{1}{\varepsilon^2}\right)$$
$$\leq (1+\varepsilon')\frac{\ln(T)}{d(p_{i^\star j}, p_{i^\star j}+\xi_j)} + O\left(\frac{1}{\varepsilon'^2}\right),$$

where $\varepsilon' = 3\varepsilon$ and the big-Oh above hides $p_{i^\star j}$ and $\xi_j$ in addition to the absolute constants. Replacing $\varepsilon$ by $\varepsilon'$ completes the proof. □

**Theorem 1** (Problem Dependent Bound)**.** *Let $P \in \mathcal{P}_{\text{WD}}$ and satisfies the transitivity property. If $\varepsilon > 0$ then, the expected regret of USS-TS in $T$ rounds is bounded by*

$$\mathfrak{R}_T \leq \sum_{j>i^\star} \frac{(1+\varepsilon)\ln T}{d(p_{i^\star j}, p_{i^\star j}+\xi_j)}\Delta_j + O\left(\frac{K-i^\star}{\varepsilon^2}\right),$$

*Proof.* Let $M_j(T)$ is the number of times arm $j$ is selected by USS-TS. Than, the regret is

$$\mathfrak{R}_T = \sum_{j\in[K]} \mathbb{E}\left[M_j(T)\right]\Delta_j = \sum_{j\in[K]} \mathbb{E}\left[\sum_{t=1}^{T} \mathbb{1}_{\{I_t=j\}}\right]\Delta_j$$
$$= \sum_{j\in[K]}\sum_{t=1}^{T} \mathbb{E}\left[\mathbb{1}_{\{I_t=j\}}\right]\Delta_j = \sum_{j\in[K]}\sum_{t=1}^{T} \mathbb{P}\left\{I_t=j\right\}\Delta_j$$
$$= \sum_{j\in[K]}\sum_{t=1}^{T} \mathbb{P}\left\{I_t=j, j\neq i^\star\right\}\Delta_j$$
$$\implies \mathfrak{R}_T = \sum_{j<i^\star}\sum_{t=1}^{T} \mathbb{P}\left\{I_t=j, j< i^\star\right\}\Delta_j + \sum_{j>i^\star}\sum_{t=1}^{T} \mathbb{P}\left\{I_t=j, j> i^\star\right\}\Delta_j \tag{15}$$

First, we bound the first of term of summation. From Lemma 4, we have

$$\sum_{t=1}^{T} \mathbb{P}\left\{I_t=j, j<i^\star\right\} \leq \frac{24}{\xi_j^2} + \sum_{s\geq 8/\xi_j} \Theta\left(\exp^{-s\xi_j^2/2} + \frac{\exp^{-sd(p_{i^\star j}-\xi_j, p_{i^\star j})}}{(s+1)\xi_j^2} + \frac{1}{\exp^{s\xi_j^2/4}-1}\right).$$

Using $\sum_{s\geq 0}\exp^{-sa} \leq 1/a$, $d(p_{i^\star j}-\xi_j, p_{i^\star j}) \leq 2\xi_j^2$ (Fact 3), and Fact 4, we have

$$\sum_{t=1}^{T} \mathbb{P}\left\{I_t=j, j<i^\star\right\} \leq \frac{24}{\xi_j^2} + \Theta\left(\frac{1}{\xi_j^2} + \frac{1}{\xi_j^4} + \left(\frac{1}{\xi_j^4} + \frac{1}{\xi_j^2}\right)\right) \leq O(1). \tag{16}$$

If arm $I_t > i^\star$ is selected then there exists at least one arm $k_1 > i^\star$ which must be preferred over $i^\star$. If the index of arm $k_1$ is smaller than the selected arm, then there must be an arm $k_2 > k_1$, which must be preferred over $k_1$. By transitivity property, arm $k_2$ is also preferred over $i^\star$. If the index of arm $k_2$ is still smaller of the selected arm, we can repeat the same argument. Eventually, we can find an arm $k'$ whose index is larger than the selected arm, and it is preferred over arm $k_i, \ldots, k_1, i^\star$. Note that the selected arm must be preferred over $k'$; hence the selected arm is also preferred over $i^\star$. We can write it as follows:

$$\sum_{t=1}^{T} \mathbb{P}\left\{I_t=j, j>i^\star\right\}\Delta_j = \sum_{t=1}^{T} \mathbb{P}\left\{I_t=j, j>i^\star, k'\succ_t k, k\succ_t i^\star, k'>j, k>i^\star\right\}\Delta_j$$

$$
\begin{aligned}
&= \sum_{t=1}^{T} \mathbb{P}\left\{I_t = j, j > i^\star, k' \succ_t i^\star, k' > j\right\} \Delta_j \quad \text{(Definition 5)} \\
&= \sum_{t=1}^{T} \mathbb{P}\left\{j \succ_t k, \forall k > j, j > i^\star, k' \succ_t i^\star, k' > j\right\} \Delta_j \text{ (Lemma 1)} \\
&= \sum_{t=1}^{T} \mathbb{P}\left\{j \succ_t k, \forall k > j, j > i^\star, j \succ_t i^\star\right\} \Delta_j \quad \text{(Definition 5)}
\end{aligned}
$$

$$
\implies \sum_{t=1}^{T} \mathbb{P}\left\{I_t = j, j > i^\star\right\} \Delta_j \leq \sum_{t=1}^{T} \mathbb{P}\left\{j \succ_t i^\star, j > i^\star\right\} \Delta_j. \tag{17}
$$

Using Lemma 7 to upper bound $\sum_{t=1}^{T} \mathbb{P}\left\{j \succ_t i^\star, j > i^\star\right\} \Delta_j$ and with Eq. (16), we get

$$
\mathfrak{R}_T \leq O(1) + \sum_{j > i^\star} \left( (1+\varepsilon) \frac{\ln(T)}{d(p_{i^\star j}, p_{i^\star j} + \xi_j)} + O\left(\frac{1}{\varepsilon^2}\right) \right) \Delta_j
$$

$$
\implies \mathfrak{R}_T \leq \sum_{j > i^\star} \frac{(1+\varepsilon)\ln(T)}{d(p_{i^\star j}, p_{i^\star j} + \xi_j)} \Delta_j + O\left(\frac{K - i^\star}{\varepsilon^2}\right). \qquad \square
$$

**Theorem 2** (Problem Independent Bound)**.** *Let $P \in \mathcal{P}_{\mathrm{WD}}$ and satisfies the transitivity property. Then the expected regret of USS-TS in $T$ rounds*

- *for any instance in $\mathcal{P}_{\mathrm{SD}}$ is bounded as*

$$
\mathfrak{R}_T \leq O\left(\sqrt{KT \ln T}\right).
$$

- *for any instance in $\mathcal{P}_{\mathrm{WD}}$ is bounded as*

$$
\mathfrak{R}_T \leq O\left((K \ln T)^{1/3} T^{2/3}\right).
$$

*Proof.* Let $M_j(T)$ is the number of times arm $j$ preferred over the optimal arm in $T$ rounds. From Lemma 4, for any $j < i^\star$, we have

$$
\begin{aligned}
\mathbb{E}\left[M_j(T)\right] &= \sum_{t=1}^{T} \mathbb{P}\left\{I_t = j, j < i^\star\right\} \\
&\leq \frac{24}{\xi_j^2} + \sum_{s \geq 8/\xi_j} \Theta\left( \exp^{-s\xi_j^2/2} + \frac{\exp^{-sd(p_{i^\star j} - \xi_j, p_{i^\star j})}}{(s+1)\xi_j^2} + \frac{1}{\exp^{s\xi_j^2/4} - 1} \right).
\end{aligned}
$$

It is east to show that $\frac{\exp^{-sd(p_{i^\star j} - \xi_j, p_{i^\star j})}}{(s+1)\xi_j^2} \leq \frac{1}{(s+1)\xi_j^2}$ and $\exp^{s\xi_j^2/4} - 1 \geq s\xi_j^2/4$ (as $\exp^y \geq y + 1$),

$$
\mathbb{E}\left[M_j(T)\right] \leq \frac{24}{\xi_j^2} + \sum_{s \geq 8/\xi_j} \Theta\left( \frac{1}{\xi_j^2} + \frac{1}{(s+1)\xi_j^2} + \frac{4}{s\xi_j^2} \right).
$$

By using $\sum_{s\geq 0}\exp^{-sa} \leq 1/a$ and $\sum_{s=1}^{T}(1/s) = \log T$,

$$\mathbb{E}\left[M_j(T)\right] \leq \frac{24}{\xi_j^2} + \Theta\left(\frac{1}{\xi_j^2} + \frac{\ln T}{\xi_j^2}\right) \implies \mathbb{E}\left[M_j(T)\right] \leq O\left(\frac{\ln T}{\xi_j^2}\right). \tag{18}$$

For any $j > i^\star$, using Lemma 5 and Lemma 6 with Eq. (17), we have

$$\mathbb{E}\left[M_j(T)\right] = \sum_{t=1}^{T}\mathbb{P}\left\{I_t = j, j > i^\star\right\} \leq \sum_{t=1}^{T}\mathbb{P}\left\{j \succ_t i^\star, j > i^\star\right\} \leq \frac{\ln T}{d(x_j, y_j)} + 1 + \frac{1}{d(x_j, p_{i^\star j})}.$$

By setting $x_j = p_{i^\star j} + \frac{\xi_j}{3}$ and $y_j = p_{i^\star j} + \frac{2\xi_j}{3}$, we have $d(x_j, y_j) \geq \frac{2\xi_j^2}{9}$ and $d(x_j, p_{i^\star j}) \geq \frac{2\xi_j^2}{9}$ (using Fact 3).

$$\begin{aligned}\mathbb{E}\left[M_j(T)\right] &\leq \frac{9\ln T}{2\xi_j^2} + 1 + \frac{9}{2\xi_j^2} \\ \implies \mathbb{E}\left[M_j(T)\right] &\leq O\left(\frac{\ln T}{\xi_j^2}\right).\end{aligned} \tag{19}$$

The regret of USS-TS is given by

$$\mathfrak{R}_T = \sum_{j\neq i^\star}\mathbb{E}\left[M_j(T)\right]\Delta_j = \sum_{j<i^\star}\mathbb{E}\left[M_j(T)\right]\Delta_j + \sum_{j>i^\star}\mathbb{E}\left[M_j(T)\right]\Delta_j$$

Recall $\Delta_j = C_j + \gamma_j - (C_{i^\star} + \gamma_{i^\star})$ and for any two arms $i$ and $j$, $0 \leq p_{ij} - (\gamma_j - \gamma_{i^\star}) \leq \beta$. By using Eq. (8a) for $j < i^\star$, we have $\Delta_j = \xi_j - (p_{i^\star j} - (\gamma_{i^\star} - \gamma_j)) \implies \Delta_j \leq \xi_j$, and using Eq. (8b) for $j > i^\star$, we have $\Delta_j = \xi_j + (p_{i^\star j} - (\gamma_{i^\star} - \gamma_j)) \implies \Delta_j \leq \xi_j + \beta$. Replacing $\Delta_j$,

$$\Rightarrow \mathfrak{R}_T \leq \sum_{j<i^\star}\mathbb{E}\left[M_j(T)\right]\xi_j + \sum_{j>i^\star}\mathbb{E}\left[M_j(T)\right](\xi_j + \beta).$$

Let $0 < \xi' < 1$. Then $\mathfrak{R}_T$ can be written as:

$$\begin{aligned}\mathfrak{R}_T \leq &\sum_{\substack{\xi'>\xi_j\\ j<i^\star}}\mathbb{E}\left[M_j(T)\right]\xi_j + \sum_{\substack{\xi'<\xi_j\\ j<i^\star}}\mathbb{E}\left[M_j(T)\right]\xi_j \\ &+ \sum_{\substack{\xi'>\xi_j\\ j>i^\star}}\mathbb{E}\left[M_j(T)\right](\xi_j + \beta) + \sum_{\substack{\xi'<\xi_j\\ j>i^\star}}\mathbb{E}\left[M_j(T)\right](\xi_j + \beta).\end{aligned}$$

Using $\sum\limits_{\xi'>\xi_j}\mathbb{E}\left[M_j(T)\right] \leq T$ for any $j$ such that $\xi' > \xi_j$,

$$\mathfrak{R}_T \leq T\xi' + \sum_{\substack{\xi'<\xi_j\\ j<i^\star}}\mathbb{E}\left[M_j(T)\right]\xi_j + \sum_{\substack{\xi'<\xi_j\\ j>i^\star}}\mathbb{E}\left[M_j(T)\right](\xi_j + \beta).$$

Substituting the value of $\mathfrak{R}_T$ from Eq. (18) and Eq. (19),

$$\begin{aligned}
\mathfrak{R}_T &\le T\xi' + \sum_{\substack{\xi'<\xi_j\\ j<i^\star}} O\left(\frac{\xi_j \ln T}{\xi_j^2}\right) + \sum_{\substack{\xi'<\xi_j\\ j>i^\star}} O\left(\frac{(\xi_j+\beta)\ln T}{\xi_j^2}\right) \\
&\le T\xi' + \sum_{\substack{\xi'<\xi_j\\ j<i^\star}} O\left(\frac{\ln T}{\xi_j}\right) + \sum_{\substack{\xi'<\xi_j\\ j>i^\star}} O\left(\frac{\ln T}{\xi_j} + \frac{\beta \ln T}{\xi_j^2}\right) \\
&\le T\xi' + O\left(\frac{K\ln T}{\xi'}\right) + O\left(\frac{K\ln T}{\xi'} + \frac{\beta K \ln T}{\xi'^2}\right) \\
&= T\xi' + O\left(K\ln T\left(\frac{1}{\xi'} + \frac{\beta}{\xi'^2}\right)\right)
\end{aligned}$$

Let there exist a variable $\alpha$ such that $O\left(K\ln T\left(\frac{1}{\xi'} + \frac{\beta}{\xi'^2}\right)\right) \le \alpha K \ln T\left(\frac{1}{\xi'} + \frac{\beta}{\xi'^2}\right)$,

$$\implies \mathfrak{R}_T \le= T\xi' + \alpha K\ln T\left(\frac{1}{\xi'} + \frac{\beta}{\xi'^2}\right). \tag{20}$$

Consider $\mathcal{P}_{\mathrm{WD}}$ class of problems. As $\xi' < 1$ and $\beta \le 2$ (as arms in the cascade may not be ordered by their error-rates, it is possible that $\gamma_i < \gamma_j$), we have $\left(\frac{1}{\xi'} + \frac{\beta}{\xi'^2}\right) \le \frac{\beta+1}{\xi'^2} \le \frac{3}{\xi'^2}$,

$$\mathfrak{R}_T \le= T\xi' + \frac{3\alpha K \ln T}{\xi'^2}.$$

Choose $\xi' = \left(\frac{6\alpha K\ln T}{T}\right)^{1/3}$ which maximize above upper bound and we get,

$$\begin{aligned}
\mathfrak{R}_T &\le (6\alpha K\ln T)^{1/3} T^{2/3} + \frac{(6\alpha K \ln T)^{1/3}}{2} T^{2/3} \\
\implies \mathfrak{R}_T &\le 2\left(6\alpha K\ln T\right)^{1/3} T^{2/3} = O\left((K\ln T)^{1/3} T^{2/3}\right)
\end{aligned}$$

It completes our proof for the case when any problem instance belongs to $\mathcal{P}_{\mathrm{WD}}$.

Now we consider any problem instance $\theta \in \mathcal{P}_{\mathrm{SD}}$. For any $\theta \in \mathcal{P}_{\mathrm{SD}} \Rightarrow \forall j \in [K],\ p_{ij} = \gamma_i - \gamma_j \implies \beta = 0$ (Setting $\mathbb{P}\left\{Y^i = Y, Y^j \neq Y\right\} = 0$ for $j > i$ in Proposition 3 of Hanawal et al. (2017)). We can rewrite Eq. (20) as

$$\mathfrak{R}_T \le T\xi' + \frac{\alpha K \ln T}{\xi'}.$$

Choose $\xi' = \left(\frac{\alpha K\ln T}{T}\right)^{1/2}$ which maximize above upper bound and we get,

$$\implies \mathfrak{R}_T \le 2\left(\alpha K T\ln T\right)^{1/2} = O\left(\sqrt{KT\ln T}\right)$$

This complete proof for second part of Theorem 2. □