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 $\varepsilon \in (0, 1 - \mu)$,
\[
\mathbb{P}\{\hat{\mu}_n \geq \mu + \varepsilon\} \leq \exp\left(-d(\mu + \varepsilon, \mu)n\right),
\]
and, for any $\varepsilon \in (0, \mu)$,
\[
\mathbb{P}\{\hat{\mu}_n \leq \mu - \varepsilon\} \leq \exp\left(-d(\mu - \varepsilon, \mu)n\right),
\]
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 $a \in (0, 1)$,
\[
\frac{1}{\exp^{Dx} - 1} = \begin{cases} \frac{\exp^{-Dx}}{1-a} & (x \geq \ln (1/a)/D) \\ \frac{1}{Dx} & (x < \ln (1/a)/D) \end{cases}
\]
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,
\[
\sum_{x=1}^{n} \frac{1}{\exp^{Dx} - 1} \leq \frac{\ln(1/a)}{D^2} + \sum_{x \geq \ln(1/a)/D}^{n} \frac{\exp^{-Dx}}{1-a}
\]
\[
\begin{align*}
\leq & \ln(1/a) + \frac{1}{(1-a)} \int_{x=0}^{\infty} \exp^{-Dx} \, dx \\
= & \frac{\ln(1/a)}{D^2} + \frac{1}{(1-a)} \left( \exp^{-Dx} \right)_{x=0}^{\infty} \\
= & \frac{\ln(1/a)}{D^2} + \frac{1}{(1-a)} \left( 0 - \exp^{0} \right) \\
= & \frac{\ln(1/a)}{D^2} + \frac{1}{(1-a)D} \\
\implies & \sum_{x=1}^{n} \frac{1}{\exp^{Dx} - 1} \leq \Theta \left( \frac{1}{D^2} + \frac{1}{D} \right). \\
\end{align*}
\]

**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

\[
d(p, q) = p \ln \frac{p}{q} + (1-p) \ln \left( \frac{1-p}{1-q} \right)
\]

\[
= \ln \left( \left( \frac{p}{q} \right)^p \left( \frac{1-p}{1-q} \right)^{1-p} \right)
\]

\[
= \ln \left( \left( \frac{q(1-p)}{p(1-q)} \right)^{-p} \right) + \ln \left( \frac{1-p}{1-q} \right)
\]

\[
\implies d(p, q) = -p \ln \left( \frac{q(1-p)}{p(1-q)} \right) + \ln \left( \frac{1-p}{1-q} \right).
\]

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

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

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

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

Appendix B. Leftover proofs from Section 4

**Lemma 4.** Let \( P \in \mathcal{P}_{WD} \) 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 \( j < i^* \),

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

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

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

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

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

Using law of iterated expectations,

\[
\implies \sum_{t=1}^{T} \mathbb{P} \{ I_t = j, j < i^* \} \leq \sum_{t=1}^{T} \mathbb{E} \left[ \frac{(1 - q_{j,t})}{q_{j,t}} 1_{\{I_t \geq i^*\}} \right].
\]
Let $s_m$ denote the time step at which the output of arm $i^*$ is observed for the $m^{th}$ time for $m \geq 1$, and let $s_0 = 0$. For $j < i^*$, whenever the output from arm $i^*$ is observed then the output from arm $j$ is also observed due to the cascade structure. Note that $q_{j,t} = P\left\{ \hat{p}_{i^*j}^{(t)} > p_{i^*j} - \xi_j | \mathcal{H}_t \right\}$ changes only when the distribution of $\hat{p}_{i^*j}^{(t)}$ changes, that is, only on the time step when the feedback from arms $i^*$ and $j$ are observed. It only happens when selected arm $I_t \geq i^*$. 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,

\[
\sum_{t=1}^{T} \mathbb{E}\left[ \frac{1 - q_{j,t}}{q_{j,t}} \mathbbm{1}_{\{I_t \geq i^*\}} \right] = \sum_{m=0}^{T-1} \mathbb{E}\left[ \frac{1 - q_{j,s_{m+1}}}{q_{j,s_{m+1}}} \sum_{t=s_{m+1}}^{s_{m+1}} \mathbbm{1}_{\{I_t \geq i^*\}} \right] \\
\leq \sum_{m=0}^{T-1} \mathbb{E}\left[ \frac{1 - q_{j,s_{m+1}}}{q_{j,s_{m+1}}} \right] \\
= \sum_{k=0}^{T-1} \mathbb{E}\left[ \frac{1}{q_{j,s_{m+1}}} - 1 \right].
\]

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

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

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

\[
\sum_{i=1}^{T} P\{I_t = j, j < i^*\} \leq \frac{24}{\xi_j^2} \sum_{s \geq 8/\xi_j} \Theta\left( \exp^{-s \xi_j^2/2} + \frac{\exp^{-s d(p_{i^*j} - \xi_j, p_{i^*j})}}{(s + 1) \xi_j^2} + \frac{1}{\exp^{s \xi_j^2/4} - 1} \right).
\]

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

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

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

\[
\sum_{t=1}^{T} P\left\{ \hat{p}_{i^*j}^{(t)} > x_j \right\} \leq \sum_{m=0}^{T-1} P\left\{ \hat{p}_{i^*j}(s_{m+1}) > x_j \right\} \\
= \sum_{m=0}^{T-1} P\left\{ \hat{p}_{i^*j}(s_{m+1}) - p_{i^*j} > x_j - p_{i^*j} \right\} \\
\leq \sum_{m=0}^{T-1} \exp^{-kd(p_{i^*j} + x_j - p_{i^*j}, p_{i^*j})} \quad \text{(using Fact 2)}
\]
Using $\sum_{a \geq 0} \exp^{-sa} \leq 1/a$, we get
\[
\sum_{t=1}^{T} \mathbb{P} \left\{ \hat{p}_{i,*}^{(t)} > x_{j} \right\} \leq \frac{1}{d(x_{j}, p_{i,*})}.
\]

**Lemma 7.** Let $P \in \mathcal{P}_{WD}$. For any $\varepsilon > 0$ and $j > i^*$,
\[
\sum_{t=1}^{T} \mathbb{P} \left\{ j > i^*, j > i^* \right\} \leq (1 + \varepsilon) \frac{\ln T}{d(p_{i,*}, p_{i,*} + \xi_j)} + O \left( \frac{1}{\varepsilon^2} \right).
\]

**Proof.** Let $p_{i,*} < x_j < y_j < p_{i,*} + \xi_j$ for any $j > i^*$. Then,
\[
\sum_{t=1}^{T} \mathbb{P} \left\{ j > i^*, j > i^* \right\} = \sum_{t=1}^{T} \mathbb{P} \left\{ \hat{p}_{i,*}^{(t)} > p_{i,*} + \xi_j \right\}
\]
\[
\leq \sum_{t=1}^{T} \mathbb{P} \left\{ \hat{p}_{i,*}^{(t)} > y_j \right\}
\]
\[
\leq \sum_{t=1}^{T} \mathbb{P} \left\{ \hat{p}_{i,*}^{(t)} \leq x_j, \hat{p}_{i,*}^{(t)} > y_j \right\} + \sum_{t=1}^{T} \mathbb{P} \left\{ \hat{p}_{i,*}^{(t)} > x_j \right\}.
\]

Using Lemma 6 and Lemma 5, we have
\[
\sum_{t=1}^{T} \mathbb{P} \left\{ j > i^*, j > i^* \right\} \leq \frac{\ln T}{d(x_j, y_j)} + 1 + \frac{1}{d(x_j, p_{i,*})}.
\]

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

Using Fact 5, if $\varepsilon \in (0, 1)$, $x_j \in (p_{i,*}, p_{i,*} + \xi_j)$, and $d(x_j, p_{i,*} + \xi_j) = d(p_{i,*}, p_{i,*} + \xi_j)/(1 + \varepsilon)$ then
\[
x_j - p_{i,*} \geq \frac{\varepsilon}{1 + \varepsilon} \frac{d(p_{i,*}, p_{i,*} + \xi_j)}{\ln \left( \frac{p_{i,*} + \xi_j}{p_{i,*} (1 + \varepsilon)} \right)}.
\]

Using Pinsker’s Inequality (Fact 3), $1/d(x_j, p_{i,*}) \leq 1/(x_j - p_{i,*})^2 = O(1/\varepsilon^2)$ where big-Oh is hiding functions of the $p_{i,*}$ and $\xi_j$,
\[
\sum_{t=1}^{T} \mathbb{P} \left\{ j > i^*, j > i^* \right\} \leq (1 + \varepsilon)^2 \frac{\ln(T)}{d(p_{i,*}, p_{i,*} + \xi_j)} + O \left( \frac{1}{\varepsilon^2} \right)
\]
Theorem 1 (Problem Dependent Bound). Let $P \in \mathcal{P}_{WD}$ and satisfies the transitivity property. If $\varepsilon > 0$ then, the expected regret of USS-TS in $T$ rounds is bounded by

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

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

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

$$R_T = \sum_{j \in [K]} \mathbb{E} [M_j(T)] \Delta_j = \sum_{j \in [K]} \mathbb{E} \left[ \sum_{t=1}^T \mathbbm{1}_{\{I_t = j\}} \right] \Delta_j$$

$$= \sum_{j \in [K]} \sum_{t=1}^T \mathbb{E} \left[ \mathbbm{1}_{\{I_t = j\}} \right] \Delta_j = \sum_{j \in [K]} \sum_{t=1}^T \mathbb{P} \left\{ I_t = j \right\} \Delta_j$$

$$= \sum_{j \in [K]} \sum_{t=1}^T \mathbb{P} \left\{ I_t = j, j \neq i^* \right\} \Delta_j$$

$$\implies R_T = \sum_{j < i^*} \sum_{t=1}^T \mathbb{P} \left\{ I_t = j, j < i^* \right\} \Delta_j + \sum_{j > i^*} \sum_{t=1}^T \mathbb{P} \left\{ I_t = j, j > i^* \right\} \Delta_j$$

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

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

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

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

If arm $I_t > i^*$ is selected then there exists at least one arm $k_1 > i^*$ which must be preferred over $i^*$, since the same property, arm $k_2 > k_1$, which must be preferred over $k_1$, By transivity property, arm $k_2$ is also preferred over $i^*$. 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 over $k_1, \ldots, k_1, i^*$. Note that the selected arm must be preferred over $k'$; hence the selected arm is also preferred over $i^*$. We can write it as follows:

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

6
\[
\sum_{t=1}^{T} P\{I_t = j, j > i^*, k' > t, k' > j\} \Delta_j \quad \text{(Definition 5)}
\]
\[
= \sum_{t=1}^{T} P\{j > t, k, \forall k > j, j > i^*, k' > t, k' > j\} \Delta_j \quad \text{(Lemma 1)}
\]
\[
= \sum_{t=1}^{T} P\{j > t, k, \forall k > j, j > i^*, j > t, i^*\} \Delta_j \quad \text{(Definition 5)}
\]
\[
\Rightarrow \sum_{t=1}^{T} P\{I_t = j, j > i^*\} \Delta_j \leq \sum_{t=1}^{T} P\{j > t, i^*\} \Delta_j.
\] (17)

Using Lemma 7 to upper bound \[
\sum_{t=1}^{T} P\{j > t, i^*, j > i^*\} \Delta_j \quad \text{and with Eq. (16), we get}
\]
\[
\mathcal{R}_T \leq O(1) + \sum_{j > i^*} (1 + \varepsilon) \frac{\ln(T)}{d(p_{i^*j}, p_{i^*j} + \xi_j)} + O\left(\frac{1}{\varepsilon^2}\right) \Delta_j
\]
\[
\Rightarrow \mathcal{R}_T \leq \sum_{j > i^*} \frac{(1 + \varepsilon) \ln(T)}{d(p_{i^*j}, p_{i^*j} + \xi_j)} \Delta_j + O\left(\frac{K - i^*}{\varepsilon^2}\right).
\]

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

- for any instance in \(\mathcal{P}_{SD}\) is bounded as
  \[
  \mathcal{R}_T \leq O\left(\sqrt{KT \ln T}\right).
  \]
- for any instance in \(\mathcal{P}_{WD}\) is bounded as
  \[
  \mathcal{R}_T \leq O\left((K \ln T)^{1/3} T^{2/3}\right).
  \]

**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 \(j < i^*\), we have

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

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

\[
\mathbb{E}[M_j(T)] \leq \frac{24}{\xi_j^2} + \sum_{s \geq 8/\xi_j} \Theta\left(\frac{1}{\xi_j^2} + \frac{1}{(s+1)\xi_j^2} + \frac{4}{s \xi_j^2}\right).
\]
By using $\sum_{s \geq 0} \exp^{-sa} \leq 1/a$ and $\sum_{s=1}^{T} (1/s) = \log T$,

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

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

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

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

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

The regret of USS-TS is given by

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

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

$$
\implies \mathcal{R}_T \leq \sum_{j < i^*} \mathbb{E}[M_j(T)] \xi_j + \sum_{j > i^*} \mathbb{E}[M_j(T)] (\xi_j + \beta).
$$

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

$$
\mathcal{R}_T \leq \sum_{\xi' > \xi_j \ j < i^*} \mathbb{E}[M_j(T)] \xi_j + \sum_{\xi' < \xi_j \ j < i^*} \mathbb{E}[M_j(T)] \xi_j + \sum_{\xi' > \xi_j \ j > i^*} \mathbb{E}[M_j(T)] (\xi_j + \beta) + \sum_{\xi' < \xi_j \ j > i^*} \mathbb{E}[M_j(T)] (\xi_j + \beta).
$$

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

$$
\mathcal{R}_T \leq T \xi' + \sum_{\xi' < \xi_j \ j < i^*} \mathbb{E}[M_j(T)] \xi_j + \sum_{\xi' < \xi_j \ j > i^*} \mathbb{E}[M_j(T)] (\xi_j + \beta).
$$
Substituting the value of $\mathcal{R}_T$ from Eq. (18) and Eq. (19),

\[ \mathcal{R}_T \leq T\xi' + \sum_{\xi' < \xi_j} O\left(\frac{\xi_j \ln T}{\xi_j^2}\right) + \sum_{\xi' \geq \xi_j} O\left(\frac{(\xi_j + \beta) \ln T}{\xi_j^2}\right) \]

\[ \leq T\xi' + \sum_{\xi' < \xi_j} O\left(\frac{\ln T}{\xi_j}\right) + \sum_{\xi' < \xi_j} O\left(\frac{\ln T}{\xi_j} + \frac{\beta \ln T}{\xi_j^2}\right) \]

\[ \leq T\xi' + O\left(\frac{K \ln T}{\xi'}\right) + O\left(\frac{K \ln T}{\xi'} + \frac{\beta K \ln T}{\xi'^2}\right) \]

\[ = T\xi' + O\left(K \ln T \left(\frac{1}{\xi'} + \frac{\beta}{\xi'^2}\right)\right) \]

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

\[ \implies \mathcal{R}_T \leq T\xi' + \alpha K \ln T \left(\frac{1}{\xi'} + \frac{\beta}{\xi'^2}\right). \] \hspace{1cm} (20)

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

\[ \mathcal{R}_T \leq T\xi' + \frac{3\alpha K \ln T}{\xi'^2}. \]

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

\[ \mathcal{R}_T \leq (6\alpha K \ln T)^{1/3} T^{2/3} + \frac{(6\alpha K \ln T)^{1/3} T^{2/3}}{2} \]

\[ \implies \mathcal{R}_T \leq 2 \left(6\alpha K \ln T\right)^{1/3} T^{2/3} = O\left((K \ln T)^{1/3} T^{2/3}\right) \]

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

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

\[ \mathcal{R}_T \leq T\xi' + \frac{\alpha K \ln T}{\xi'}. \]

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

\[ \implies \mathcal{R}_T \leq 2 \left(\alpha KT \ln T\right)^{1/2} = O\left(T \ln T\right) \]

This complete proof for second part of Theorem 2.