Appendices

A Detailed Proofs of Theorem 1

A.1 Proof of Lemma 1

Proof of Lemma 1. Define that $R(s) = \sum_{t=1}^{s} \max_{k\in\mathcal{K}} X_{k,t} - X_{A_t,t}$, then we have that $\mathcal{R}(T) = \mathbb{E}[R(T)]$.

$$\begin{aligned}\mathcal{R}(T) =&\mathbb{E}[R(T)]\\ =&\mathbb{E}[R(T)\mathbb{I}\{\tau_1 \le T\}] + \mathbb{E}[R(T)\mathbb{I}\{\tau_1 > T\}]\\ \le& T\cdot\mathbb{P}(\tau_1 \le T) + \mathbb{E}[R(T)\mathbb{I}\{\tau_1 > T\}].\end{aligned} \tag{10}$$

Define $N_k(t)$ as the number of times arm k has been selected by the Algorithm 2 in the first t steps, i.e., $N_k(t) = \sum_{i=1}^{t}\mathbb{I}(A_i = k)$. No false alarm is raised and we do not restart the UCB algorithm if the evnet $\{\tau_1 > T\}$ happens. Therefore, we have the following equation:

$$\mathbb{E}[R(T)\mathbb{I}\{\tau_1 > T\}] = \sum_{\Delta_k^{(1)}>0} \Delta_k^{(1)} \cdot \mathbb{E}[N_k(T)\mathbb{I}\{\tau_1 > T\}].$$

Thus, it remains to show an upper bound for $\mathbb{E}[N_k(T)\mathbb{I}\{\tau_1 > T\}]$. By the definition of Algorithm 2, we have that for any $k \in \mathcal{K}$

$$\begin{aligned}&N_k(T)\mathbb{I}\{\tau_1 > T\}\\ =&\sum_{t=1}^{T}\mathbb{I}\{A_t = k, \tau_1 > T, N_k(t) < l\}\\ +&\sum_{t=1}^{T}\mathbb{I}\{A_t = k, \tau_1 > T, N_k(t) \ge l\}\\ \le& l + \sum_{t=1}^{T}\mathbb{I}\{t \bmod \lfloor K/\gamma\rfloor = k, N_k(t) \ge l\}\\ +&\sum_{t=1}^{T}\mathbb{I}\{k = \underset{\tilde{k}\in\mathcal{K}}{argmax}\,\mathrm{UCB}_{\tilde{k}}, N_k(t) \ge l\}\\ \le& l + \lceil T\gamma/K\rceil + \sum_{t=1}^{T}\mathbb{I}\{k = \underset{\tilde{k}\in\mathcal{K}}{argmax}\,\mathrm{UCB}_{\tilde{k}}, N_k(t) \ge l\},\end{aligned} \tag{11}$$

where the first inequality is due to the fact that if the event $\{A_t = k, \tau_1 > T\}$ happens, then we do not restart the UCB algorithm before time T and the selection of the kth arm is based on either the uniform sampling or the largest UCB index in a stochastic bandit setting. Setting $l = \lceil 8\log T/(\Delta_k^{(1)})^2\rceil$ and following the same argument as in the proof of Theorem 1 of [Auer et al., 2002a], we have that

$$\mathbb{E}[N_k(T)\mathbb{I}\{\tau_1 > T\}] \le \frac{T\gamma}{K} + \frac{8\log T}{(\Delta_k^{(1)})^2} + 1 + \frac{\pi^2}{3} + K.$$

Summing over $k \in \mathcal{K}$ we prove the result.

□

A.2 Proof of Lemma 2

Proof of Lemma 2. Define $\tau_{k,1}$ as the first detection time of the kth arm. Then, $\tau_1 = \min_{k\in\mathcal{K}}\{\tau_{k,1}\}$ since Algorithm 2 is designed to reinitialize the UCB algorithm if a change is detected on any of the K arms. Using the union bound, we have that

$$\mathbb{P}(\tau_1 \leq T) \leq \sum_{k=1}^{K} \mathbb{P}(\tau_{k,1} \leq T).$$

Define that for any $k \in \mathcal{K}$ and $t \geq w$

$$S_{k,t} = \left| \sum_{i=t-w/2+1}^{i=t} Z_{k,i} - \sum_{i=t-w+1}^{t-w/2} Z_{k,i} \right|. \tag{12}$$

Then, for any $k \in \mathcal{K}$, $\tau_{k,1}$ is given by

$$\tau_{k,1} = \inf\{t \geq w : S_{k,t} > b\}$$

Let $\mathbb{Z}^+$ be the set of all positive integers. Define that for any $0 \leq j \leq w-1$ the stopping times

$$\tau_{k,1}^{(j)} = \inf\left\{t = j + nw, n \in \mathbb{Z}^+ : S_{k,t} > b\right\}.$$

We have that $\tau_{k,1} = \min\{\tau_1^{(0)}, \ldots, \tau_1^{(w-1)}\}$. Note that under the stationary environment, for any $0 \leq j \leq w-1$, $\tau_{k,1}^{(j)}$ is a random variable with the geometric distribution

$$\mathbb{P}(\tau_{k,1}^{(j)} = nw + j) = p(1-p)^{n-1},$$

where $p = \mathbb{P}(S_{k,w} > b)$. Therefore, considering union bound we have that for any $k \in \mathcal{K}$

$$\mathbb{P}(\tau_{k,1} \leq T) \leq w\left(1 - (1-p)^{\lfloor T/w \rfloor}\right).$$

The remaining task is to find an upper bound for p. Note that for any $k \in \mathcal{K}$, $S_{k,w}$ is a random variable with zero mean. We have by the McDiarmid's inequality and the union bound that

$$p \leq 2 \cdot \exp\left(-\frac{2b^2}{w}\right).$$

Combining the above analysis we conclude the result. □

A.3 Proof of Lemma 3

Proof of Lemma 3. Assume that $\delta_{\tilde{k}}^{(1)} \geq 2b/w + c$ for some $\tilde{k} \in \mathcal{K}$. Since the uniformly sampling scheme (step 2-4 of Algorithm 2) guarantees that in any time interval with length larger than $L/2$ each arm is sampled at least $w/2$ times, conditioning on $\{\tau_1 > \nu_1\}$, we have that

$$\begin{aligned}
&\mathbb{P}(\nu_1 < \tau_1 \leq \nu_1 + L/2 \mid \tau_1 > \nu_1) \\
\geq &\mathbb{P}\left(S_{\tilde{k},w} > b\right) \\
\geq &1 - 2\exp\left(-\frac{(w|\delta_{\tilde{k}}^{(1)}|/2 - b)^2}{w}\right) \\
\geq &1 - 2\exp\left(-\frac{wc^2}{4}\right),
\end{aligned} \tag{13}$$

where $S_{k,t}$ is defined in (12) and we use McDiarmid's inequality in the second inequality. □

A.4 Proof of Lemma 4

Proof of Lemma 4. First, define that $N = \lceil b/\delta_{\tilde{k}}^{(1)} \rceil \cdot \lceil K/\gamma \rceil$, we obtain a simple upper bound for the EDD as follows.

$$\begin{aligned}
&\mathbb{E}[\tau_1 - \nu_1 \mid \nu_1 < \tau_1 \leq \nu_1 + L/2] \\
=&\sum_{i=1}^{L/2} \mathbb{P}(\tau_1 \geq \nu_1 + i \mid \nu_1 < \tau_1 \leq \nu_1 + L/2) \\
\leq& N + \sum_{i=N}^{L/2} \left(\mathbb{P}(\tau_1 \geq \nu_1 + i \mid \nu_1 < \tau_1 \leq \nu_1 + L/2) \right).
\end{aligned}$$

Since the uniformly sampling scheme guarantees that we have at least $i/\lceil K/\gamma \rceil$ samples from each arm within i time steps, we use McDiarmid's inequality and Lemma 3 to have that

$$\begin{aligned}
&\sum_{i=N}^{L/2} \left(\mathbb{P}(\tau_1 \geq \nu_1 + i \mid \nu_1 < \tau_1 \leq \nu_1 + L/2) \right) \\
=&\sum_{i=N}^{L/2} \frac{\mathbb{P}(\nu_1 + i \leq \tau_1 \leq \nu_1 + L/2 \mid \tau_1 > \nu_1)}{\mathbb{P}(\nu_1 \leq \tau_1 \leq \nu_1 + L/2 \mid \tau_1 > \nu_1)} \\
\leq& \frac{1}{1 - 2\exp\left(-wc^2/4\right)} \cdot \sum_{i=N}^{L/2} 2\exp\left(- \frac{(i/\lceil (K/\gamma) \rceil \delta_{\tilde{k}}^{(1)} - b)^2}{w} \right) \\
\leq& \frac{\lceil K/\gamma \rceil}{1 - 2\exp\left(-wc^2/4\right)} \cdot \sum_{j=\lceil b/\delta_{\tilde{k}}^{(1)} \rceil}^{w/2} 2\exp\left(- \frac{(j\delta_{\tilde{k}}^{(1)} - b)^2}{w} \right).
\end{aligned}$$

Define $q = \lceil (w/2) \cdot \delta_{\tilde{k}}^{(1)} \rceil - b$ and we have $q > 1$ from the assumption that $\delta_{\tilde{k}}^{(1)} > 2b/w + c$. Combining the above analysis, we have that

$$\begin{aligned}
&\left(1 - 2\exp\left(-wc^2/4\right)\right) \cdot \mathbb{E}[\tau_1 - \nu_1 \mid \nu_1 < \tau_1 \leq \nu_1 + L/2] \\
\leq& N + \lceil K/\gamma \rceil \cdot \sum_{j=\lceil b/\delta_{\tilde{k}}^{(1)} \rceil}^{w/2} 2\exp\left(- \frac{(j\delta_{\tilde{k}}^{(1)} - b)^2}{w} \right) \\
\leq& N + 2\lceil K/\gamma \rceil \cdot \left(1 + \int_1^q \exp\left(-\frac{l^2}{w} \right) dl \right) \\
\leq& N + 2\lceil K/\gamma \rceil \cdot \\
&\left[1 + \sqrt{w} \left(1 - \frac{1}{\sqrt{w}} + \int_1^{q/\sqrt{w}} \exp(-u^2) du \right) \right] \\
\leq& N + 2\lceil K/\gamma \rceil \cdot \left(\sqrt{w} + \sqrt{w} \int_1^{q/\sqrt{w}} u \exp(-u^2) du \right) \\
\leq& \left(\lceil b/\delta_{\tilde{k}}^{(1)} \rceil + 3\sqrt{w} \right) \cdot \lceil K/\gamma \rceil,
\end{aligned}$$

where we transform l into $u = l/\sqrt{w}$ in the third inequality and we use the fact that $\exp(-u^2) \leq u \exp(-u^2), u \geq 1$ in the fourth inequality. On the other hand, by the definition of the conditioning event we also have that

$$\left(1 - 2\exp\left(-wc^2/4\right)\right) \cdot \mathbb{E}[\tau_1 - \nu_1 \mid \nu_1 < \tau_1 \leq \nu_1 + L/2] \leq L/2.$$

Combining the above analysis we conclude the result. □

A.5 Proof of Theorem 1

Proof of Theorem 1. Recall that $L = w\lceil K/\gamma \rceil$. Algorithm 2 guarantees that in any time interval with length larger than L each arm is sampled at least w times. Define events $F_i = \{\tau_i > \nu_i\}, 1 \le i \le M-1$. Define events $D_i = \{\tau_i \le \nu_i + L/2\}, 1 \le i \le M-2$ and event $D_{M-1} = \{\tau_{M-1} \le T\}$. Therefore, the event $F_i D_i$ is the good event where the ith change can be detected correctly and efficiently. Define that $R(s) = \sum_{t=1}^{s} \max_{k\in\mathcal{K}} X_{k,t} - X_{A_t,t}$, then we have that $\mathcal{R}(T) = \mathbb{E}[R(T)]$. Equipped with the sequence of good events, we have that

$$\begin{aligned}\mathcal{R}(T) = \mathbb{E}[R(T)] \le & \mathbb{E}[R(T)\mathbb{I}\{F_1\}] + T \cdot (1 - \mathbb{P}(F_1)) \\ \le & \mathbb{E}[R(\nu_1)\mathbb{I}\{F_1\}] + \mathbb{E}[R(T) - R(\nu_1)] + 1 \\ \le & \tilde{C}_1 + \gamma\nu_1 + \mathbb{E}[R(T) - R(\nu_1)] + 1.\end{aligned}$$

Above, the second inequality is due to Lemma 2 that $\mathbb{P}(F_1) \ge 1 - 1/T$ provided that $b = [w \log(2KT^2)/2]^{1/2}$ and the third inequality is due to the bound in the end of Lemma 1, which is the bound for the UCB algorithm in a stochastic bandit setting.

The next step is to bound $\mathbb{E}[R(T) - R(\nu_1)]$. Using the law of total expectation, we have that

$$\begin{aligned}&\mathbb{E}[R(T) - R(\nu_1)] \\ \le & \mathbb{E}[R(T) - R(\nu_1) \mid F_1 D_1] + T \cdot (1 - \mathbb{P}(F_1 D_1)) \\ \le & \mathbb{E}[R(T) - R(\nu_1) \mid F_1 D_1] + 2,\end{aligned}$$

where the last inequality is due to Lemma 3 that we have $\mathbb{P}(D_1 \mid F_1) \ge 1 - 1/T$ provided that $c = 2\sqrt{\log(2T)/w}$ and the fact that $\mathbb{P}(F_1 D_1) = \mathbb{P}(D_1 \mid F_1) \cdot \mathbb{P}(F_1)$ for any probability measure $\mathbb{P}$.

Therefore, the remaining task is to bound $\mathbb{E}[R(T) - R(\nu_1) \mid F_1 D_1]$. Denote $\tilde{\mathbb{E}}$ as the expectation according to the piecewise-stationary bandit starts from the second segment. Further splitting the regret, we have that

$$\begin{aligned}&\mathbb{E}[R(T) - R(\nu_1) \mid F_1 D_1] \\ \le & \mathbb{E}[R(T) - R(\tau_1) \mid F_1 D_1] + \mathbb{E}[R(\tau_1) - R(\nu_1) \mid F_1 D_1] \\ \le & \tilde{\mathbb{E}}[R(T - \nu_1)] + \mathbb{E}[\tau_1 - \nu_1 \mid F_1 D_1] \\ \le & \tilde{\mathbb{E}}[R(T - \nu_1)] + \min(L/2, (\lceil b/\delta^{(1)} \rceil + 3\sqrt{w}) \cdot \lceil K/\gamma \rceil) / (1 - 1/T)\end{aligned}$$

where the second inequality is due to the renewal property given that the whole algorithm restarts in the time interval between ν_1 and $\nu_1 + L/2$ and the last inequality is due to Lemma 4 by setting $c = 2\sqrt{\log(2T)/w}$.

Combining the above analysis, we bound the regret in a recursive manner as follows (assuming $T \ge 2$):

$$\begin{aligned}\mathbb{E}[R(T)] \le & \tilde{\mathbb{E}}[R(T - \nu_1)] + \tilde{C}_1 \\ & + \gamma\nu_1 + 2\min(L/2, (\lceil b/\delta^{(1)} \rceil + 3\sqrt{w}) \cdot \lceil K/\gamma \rceil) + 3.\end{aligned}$$

The recursive manner means that we can apply the same method to bound $\tilde{\mathbb{E}}[R(T - \nu_1)]$, by conditioning on the event $D_2 F_2$. Repeating this procedure $M - 1$ times, we obtain that

$$\begin{aligned}\mathbb{E}[R(T)] \le & \textstyle\sum_{i=1}^{M} \tilde{C}_i + \gamma T \\ & + \textstyle\sum_{i=1}^{M-1} \frac{2K \cdot \min(\frac{w}{2}, \lceil \frac{b}{\delta^{(i)}} \rceil + 3\sqrt{w})}{\gamma} + 3M.\end{aligned}$$

□