Appendix

In order to prove Lemma 2, we need to introduce some auxiliary useful lemmas.

Lemma 5. *It holds* $\mathbb{P}[\mathcal{E}_1] \geq 1 - \delta$.

Proof. Let $s \geq H$ be an integer, and let $k \in \mathcal{K}$. By the Chernoff bound and the fact that the variables are all i.i.d., we have:

$$\begin{aligned}
\mathbb{P}\left[\sum_{i=1}^{s} \mathbb{I}\{B_{k,i} \geq p_k\} \leq \frac{s}{2}\mathbb{P}[B \geq p_k]\right] &\leq e^{-\frac{s\mathbb{P}[B \geq p_k]}{8}} \\
&= e^{-\frac{s\mathbb{P}[B\geq p_k]\log\left({KT^2}/{\delta}\right)}{8\log\left({KT^2}/{\delta}\right)}} \\
&= \left(\frac{\delta}{KT^2}\right)^{\frac{s\mathbb{P}[B\geq p_k]}{8\log\left({KT^2}/{\delta}\right)}} \\
&\leq \left(\frac{\delta}{KT^2}\right)^{\frac{T^{1/3}\mathbb{P}[B\geq p_k]}{8\log\left({KT^2}/{\delta}\right)}} \leq \frac{\delta}{KT^2}.
\end{aligned}$$

If $t \geq HK$, then $\mathcal{T}_{k,t} \geq H$ according to the definition of $\mathcal{T}_{k,t}$ and how our algorithm works. It follows that

$$\begin{aligned}
&\mathbb{P}\left[\bigcup_{t=HK}^{T}\bigcup_{k\in\mathcal{K}}\left\{\mathcal{Q}_{k,t} \leq \frac{\mathcal{T}_{k,t}}{2}\mathbb{P}[B \geq p_k]\right\}\right] \\
&\qquad = \mathbb{P}\left[\bigcup_{s=H}^{T}\bigcup_{t=HK}^{T}\bigcup_{k\in\mathcal{K}}\left\{\mathcal{Q}_{k,t} \leq \frac{\mathcal{T}_{k,t}}{2}\mathbb{P}[B \geq p_k]\right\} \cap \{\mathcal{T}_{k,t} = s\}\right] \\
&\qquad \leq \sum_{s=H}^{T}\sum_{t=HK}^{T}\sum_{k\in\mathcal{K}}\mathbb{P}\left[\left\{\mathcal{Q}_{k,t} \leq \frac{\mathcal{T}_{k,t}}{2}\mathbb{P}[B \geq p_k]\right\} \cap \{\mathcal{T}_{k,t} = s\}\right] \\
&\qquad = \sum_{s=H}^{T}\sum_{t=HK}^{T}\sum_{k\in\mathcal{K}}\mathbb{P}\left[\left\{\sum_{i=1}^{s}\mathbb{I}\{B_{k,i} \geq p_k\} \leq \frac{s}{2}\mathbb{P}[B \geq p_k]\right\} \cap \{\mathcal{T}_{k,t} = s\}\right] \\
&\qquad \leq \sum_{s=H}^{T}\sum_{t=HK}^{T}\sum_{k\in\mathcal{K}}\mathbb{P}\left[\sum_{i=1}^{s}\mathbb{I}\{B_{k,i} \geq p_k\} \leq \frac{s}{2}\mathbb{P}[B \geq p_k]\right] \\
&\qquad \leq \delta\ .
\end{aligned}$$

Hence,

$$\mathbb{P}\left[\bigcap_{t=HK}^{T}\bigcap_{k\in\mathcal{K}}\left\{\mathcal{Q}_{k,t} > \frac{\mathcal{T}_{k,t}}{2}\mathbb{P}[B \geq p_k]\right\}\right] \geq 1 - \delta\ ,$$

which concludes the proof. $\square$

Lemma 6. *It holds*

$$\mathbb{P}\left[\bigcap_{k\in\mathcal{K}}\left\{\left|\hat{G}_k - \int_0^{p_k}\mathbb{P}[S \leq \lambda]\,\mathrm{d}\lambda\right| \leq \sqrt{\frac{2\log(2/\delta)}{KN_k}} + \frac{1}{K}\right\}\right] \geq 1 - 2K\delta.$$

Proof. Under the event $\mathcal{E}_1$, for all $i \in [K]$ such that $\mathbb{P}[B \geq p_i] \geq 8T^{-\frac{1}{3}}\log({KT^2}/{\delta})$, we have:

$$\mathcal{Q}_{i,HK} \geq \frac{1}{2}\mathcal{T}_{i,HK}\mathbb{P}[B \geq p_i] = \frac{1}{2}H\mathbb{P}[B \geq p_i],$$

since $\mathcal{T}_{i,HK} = H$ for every $i \in [K]$, according how our algorithm works. This implies that, under the event $\mathcal{E}_1$, the following holds:

$$N_k = \min_{i \le k} \mathcal{Q}_{i,HK} \ge \frac{1}{2} \min_{i \le k} H\mathbb{P}[B \ge p_i] = \frac{1}{2} H\mathbb{P}[B \ge p_k] := n_k.$$

As a first step, we prove that, conditioned to the event $\{N_k = \ell\} \cap \mathcal{E}_1$, the estimates $\hat{G}_k$ concentrate around their expectation. Specifically, whenever $\ell \ge n_k$, the following holds:

$$\mathbb{P}\left[\left|\hat{G}_k - \mathbb{E}[\hat{G}_k]\right| \le 2\sqrt{\frac{\log(2/\delta)}{K\ell}} \mid \{N_k = \ell\} \cap \mathcal{E}_1\right] \ge 1 - \delta. \tag{4}$$

Indeed, by Azuma–Hoeffding inequality, we have:

$$\mathbb{P}\left[\left|\sum_{i=1}^{k}\sum_{j=1}^{\ell} \mathbb{I}\{S_{i,j} \le p_i\} - \sum_{i=1}^{k}\sum_{j=1}^{\ell} \mathbb{P}[S \le p_i]\right| \le \sqrt{2K\ell \log(2/\delta)}\right] \ge 1 - \delta \ .$$

Furthermore, noticing that

$$\left|\hat{G}_k - \mathbb{E}[\hat{G}_k]\right| = \frac{1}{K\ell}\left|\sum_{i=1}^{k}\sum_{j=1}^{\ell} \mathbb{I}\{S_{i,j} \le p_i\} - \sum_{i=1}^{k}\sum_{j=1}^{\ell} \mathbb{P}[S \le p_i]\right| ,$$

and that $\{N_k = \ell\} \cap \mathcal{E}_1$ is $\mathbb{P}$-independent from $S_1, S_2, \ldots$, we have:

$$\begin{aligned}\mathbb{P}\left[\left|\hat{G}_k - \mathbb{E}[\hat{G}_k]\right| \le \sqrt{\frac{2\log(2/\delta)}{K\ell}} \mid \{N_k = \ell\} \cap \mathcal{E}_1\right] &= \mathbb{P}\left[\left|\hat{G}_k - \mathbb{E}[\hat{G}_k]\right| \le \sqrt{\frac{2\log(2/\delta)}{K\ell}}\right] \\ &\ge 1 - \delta \ ,\end{aligned}$$

showing that Equation 4 holds. Therefore, we can prove that:

$$\begin{aligned}&\mathbb{P}\left[\left|\hat{G}_k - \mathbb{E}[\hat{G}_k]\right| \le \sqrt{\frac{2\log(2/\delta)}{KN_k}} \mid \mathcal{E}_1\right] \\ &\ge \sum_{\substack{\ell=0,\\ \ell \ge n_k}}^{H} \mathbb{P}\left[\left|\hat{G}_k - \mathbb{E}[\hat{G}_k]\right| \le \sqrt{\frac{2\log(2/\delta)}{K\ell}} \mid \{N_k = \ell\} \cap \mathcal{E}_1\right] \mathbb{P}\left[N_k = \ell \mid \mathcal{E}_1\right] \\ &\ge \sum_{\substack{\ell=0,\\ \ell \ge n_k}}^{H} (1-\delta)\mathbb{P}\left[N_k = \ell \mid \mathcal{E}_1\right] \\ &= 1 - \delta,\end{aligned}$$

where the first inequality holds because of the law of total probability, noticing that $\mathbb{P}[N_k = \ell \mid \mathcal{E}_1] = 0$ for all $\ell < n_k$, while the second inequality by eq. (4). Thanks to the i.i.d. hypothesis we have:

$$\begin{aligned}\mathbb{E}[\hat{G}_k] - \int_0^{p_k} \mathbb{P}[S \le \lambda]\,\mathrm{d}\lambda &= \mathbb{E}\left[\frac{1}{KN_k}\sum_{i=1}^{k}\sum_{j=1}^{N_k} \mathbb{I}\{S_{i,j} \le p_i\}\right] - \int_0^{p_k} \mathbb{P}[S \le \lambda]\,\mathrm{d}\lambda \\ &= \frac{1}{K}\sum_{i=1}^{k} \mathbb{P}[S \le p_i] - \int_0^{p_k} \mathbb{P}[S \le \lambda]\,\mathrm{d}\lambda \\ &= \sum_{i=1}^{k} \int_{\frac{i-1}{K}}^{\frac{i}{K}} \left(\mathbb{P}\left[S \le \frac{i}{K}\right] - \mathbb{P}[S \le \lambda]\right) \mathrm{d}\lambda =: (\star) \ ,\end{aligned}$$

Now, due to the fact that $\lambda \mapsto \mathbb{P}[S \le \lambda]$ is a non-decreasing function, we have that

$$0 \le (\star) \le \sum_{i=1}^{k} \int_{\frac{i-1}{K}}^{\frac{i}{K}} \left(\mathbb{P}\left[S \le \frac{i}{K}\right] - \mathbb{P}\left[S \le \frac{i-1}{K}\right]\right) \mathrm{d}\lambda$$

$$= \frac{1}{K}\sum_{i=1}^{k}\left(\mathbb{P}\left[S \le \frac{i}{K}\right] - \mathbb{P}\left[S \le \frac{i-1}{K}\right]\right) = \frac{\mathbb{P}[S \le p_k] - \mathbb{P}[S \le 0]}{K} \le \frac{1}{K}.$$

Thus, the following holds:

$$\mathbb{P}\left[\left|\hat{G}_k - \int_0^{p_k} \mathbb{P}[S \le \lambda]\, \mathrm{d}\lambda\right| \le \sqrt{\frac{2\log(2/\delta)}{KN_k}} + \frac{1}{K} \,\Big|\, \mathcal{E}_1\right] \ge 1-\delta.$$

Thus, thanks to Lemma 5, we have:

$$\begin{aligned}
&\mathbb{P}\left[\left|\hat{G}_k - \int_0^{p_k} \mathbb{P}[S \le \lambda]\, \mathrm{d}\lambda\right| \le \sqrt{\frac{2\log(2/\delta)}{KN_k}} + \frac{1}{K}\right] \\
&\ge \mathbb{P}\left[\left|\hat{G}_k - \int_0^{p_k} \mathbb{P}[S \le \lambda]\, \mathrm{d}\lambda\right| \le \sqrt{\frac{2\log(2/\delta)}{KN_k}} + \frac{1}{K} \,\Big|\, \mathcal{E}_1\right] \mathbb{P}[\mathcal{E}_1] \\
&\ge 1-2\delta.
\end{aligned}$$

Finally, taking a union bound over all possible sets of arms $\mathcal{K}$, we prove the lemma. □

Lemma 7. *It holds*

$$\mathbb{P}\left[\bigcap_{k\in\mathcal{K}}\left\{\left|\hat{F}_k - \int_{p_k}^{1} \mathbb{P}[B \ge \lambda]\, \mathrm{d}\lambda\right| \le \sqrt{\frac{2\log(2/\delta)}{KH}} + \frac{1}{K}\right\}\right] \ge 1-K\delta.$$

Proof. Thanks to the i.i.d. hypothesis and the fact that $\lambda \mapsto \mathbb{P}[B \ge \lambda]$ is a non-increasing function, with an argument analogous to that provided in the proof of Lemma 6, we have that

$$\left|\mathbb{E}[\hat{F}_k] - \int_{p_k}^{1} \mathbb{P}[B \ge \lambda]\, \mathrm{d}\lambda\right| \le \frac{1}{K}.$$

By Azuma–Hoeffding inequality, we have:

$$\left|\sum_{i=1}^{k}\sum_{j=1}^{H} \mathbb{I}\{B_{i,j} \ge p_i\} - \sum_{i=1}^{k}\sum_{j=1}^{H} \mathbb{P}\left[B \ge p_i\right]\right| \le \sqrt{2kH\log(2/\delta)} \le \sqrt{2KH\log(2/\delta)}$$

with probability $1-\delta$. Hence, to conclude, it is enough to notice that

$$\left|\hat{F}_k - \mathbb{E}[\hat{F}_k]\right| = \frac{1}{KH}\left|\sum_{i=1}^{k}\sum_{j=1}^{H} \mathbb{I}\{B_{i,j} \ge p_i\} - \sum_{i=1}^{k}\sum_{j=1}^{H} \mathbb{P}\left[B \ge p_i\right]\right|$$

and take a union bound over all possible sets of arms $[K]$. □

Lemma 8. *It holds* $\mathbb{P}[\mathcal{E}_3] \ge 1 - 2KT\delta$.

Proof. Fix $k \in \mathcal{K}$ and $t \ge HK$. Recall that, by definition,

$$\hat{\mu}_{k,t} := \frac{\sum_{\tau=1}^{\mathcal{Q}_{k,t}} \mathbb{I}\{S_{k,\tau} \le p_k\}}{\mathcal{Q}_{k,t}}.$$

Employing a union bound and the Hoeffding inequality, we have that:

$$\hat{\mu}_{k,t} - \sqrt{\frac{\log(2T/\delta)}{2\mathcal{Q}_{k,t}}} \le \mathbb{P}\left[S \le p_k\right] \le \hat{\mu}_{k,t} + \sqrt{\frac{\log(2T/\delta)}{2\mathcal{Q}_{k,t}}}$$

with probability at least $1-\delta$. Thus, taking a union bound, we have:

$$\mathbb{P}\left[\bigcap_{k\in[K]}\bigcap_{t=HK}^{T}\left\{\hat{\mu}_{k,t} - \sqrt{\frac{\log(2T/\delta)}{2\mathcal{Q}_{k,t}}} \le \mathbb{P}\left[S \le p_k\right] \le \hat{\mu}_{k,t} + \sqrt{\frac{\log(2T/\delta)}{2\mathcal{Q}_{k,t}}}\right\}\right] \ge 1-KT\delta.$$

With an analogous argument, it is possible to show that:

$$\mathbb{P}\left[\bigcap_{k\in[K]}\bigcap_{t=HK}^{T}\left\{\hat{\nu}_{k,t}-\sqrt{\frac{\log(2T/\delta)}{2\mathcal{T}_{k,t}}}\leq\mathbb{P}\left[B\geq p_k\right]\leq\hat{\nu}_{k,t}+\sqrt{\frac{\log(2T/\delta)}{2\mathcal{T}_{k,t}}}\right\}\right]\geq 1-KT\delta.$$

Thus, taking a union bound, we have that the lemma holds. $\square$

Lemma 9. *It holds* $\mathbb{P}[\mathcal{E}_4]\geq 1-\delta$

Proof. Let $\epsilon=8T^{-1/3}\log(KT^2/\delta)$ and $\hat{\nu}_k:=\hat{\nu}_{k,KH}$. If $k\notin\mathcal{K}$, then $\mathbb{P}[B\geq p_k]\leq\epsilon$. Therefore, we can employ the multiplicative Chernoff inequality as follows:

$$\mathbb{P}\big[\hat{\nu}_k\geq(1+c)\mathbb{P}[B\geq p_k]\big]\leq e^{-\frac{c^2H\mathbb{P}[B\geq p_k]}{2+c}},$$

with $c=\frac{\epsilon}{\mathbb{P}[B\geq p_k]}$. Thus, we get:

$$\begin{aligned}\mathbb{P}\big[\hat{\nu}_k\geq\mathbb{P}[B\geq p_k]+\epsilon\big]&\leq\exp\left(-\frac{\left(\frac{\epsilon}{\mathbb{P}[B\geq p_k]}\right)^2H\mathbb{P}[B\geq p_k]}{2+\frac{\epsilon}{\mathbb{P}[B\geq p_k]}}\right)\\&\leq\exp\left(-\frac{\frac{\epsilon}{\mathbb{P}[B\geq p_k]}}{2+\frac{\epsilon}{\mathbb{P}[B\geq p_k]}}\cdot\epsilon H\right)\\&\leq\left(\frac{\delta}{KT^2}\right)^{8/3}\leq\frac{\delta}{KT^2},\end{aligned}$$

since $x/(x+2)\geq 1/3$, for every $x\geq 1$. As a result, we have:

$$\mathbb{P}\big[\mathcal{Q}_{k,HK}\leq 2H\epsilon\big]=\mathbb{P}\big[\hat{\nu}_k\leq 2\epsilon\big]\geq\mathbb{P}\big[\hat{\nu}_k\leq\mathbb{P}[B\geq p_k]+\epsilon\big]\geq 1-\frac{\delta}{KT^2},$$

recalling that $\hat{\nu}_{k,HK}=\mathcal{Q}_{k,HK}/H$. Furthermore, we notice that:

$$2H\epsilon\leq 16HT^{-1/3}\log(KT^2/\delta)=16\frac{\lceil T^{1/3}\rceil}{T^{1/3}}\log(KT^2/\delta)\leq 32\log(KT^2/\delta).$$

Thus, by taking a union bound, we have:

$$\mathbb{P}\left[\bigcap_{k\notin\mathcal{K}}\left\{\mathcal{Q}_{k,HK}\leq 32\log(KT^2/\delta)\right\}\right]\geq 1-\frac{\delta}{T^2}\geq 1-\delta,$$

concluding the proof. $\square$

Lemma 2. *Let* $\mathcal{E}:=\mathcal{E}_1\cap\mathcal{E}_2\cap\mathcal{E}_3\cap\mathcal{E}_4$. *Then, we have:*

$$\mathbb{P}\left[\mathcal{E}\right]\geq 1-\mathcal{O}(KT\delta).$$

Proof. Thanks to Lemmas 5, 6, 7, 8, and 9, by taking a union bound we have:

$$\mathbb{P}\left[\mathcal{E}\right]\geq 1-(5KT+2)\delta=1-\mathcal{O}(KT\delta),$$

concluding the proof. $\square$

Lemma 3. *Let* $k\in\mathcal{K}$. *Then, for each* $t>HK$, *conditional on the event* $\mathcal{E}$, *we have:*

$$0\leq\left(\hat{\mu}_{k,t}+\sqrt{\frac{\log(2T/\delta)}{2\mathcal{Q}_{k,t}}}\right)\left(\hat{F}_k+\sqrt{\frac{2\log(2/\delta)}{HK}}+\frac{1}{K}\right)-\mathbb{P}[S\leq p_k]\int_{p_k}^{1}\mathbb{P}[B\geq\lambda]\,\mathrm{d}\lambda\leq\eta,$$

where $\eta:=C\log\left(\frac{T}{\delta}\right)\left(\frac{1}{T^{1/3}}+\frac{1}{\sqrt{\mathcal{T}_{k,t}}}\right)$ *and* $C>0$ *is an absolute constant.*

Proof. We first prove that:

$$\left(\hat{\mu}_{k,t} + \sqrt{\frac{\log(2T/\delta)}{2\mathcal{Q}_{k,t}}}\right)\left(\hat{F}_k + \sqrt{\frac{2\log(2/\delta)}{HK}} + \frac{1}{K}\right)$$

is the (optimistic) estimator of

$$\mathbb{P}[S \le p_k] \int_{p_k}^{1} \mathbb{P}[B \ge \lambda]\,\mathrm{d}\lambda.$$

To do so, we observe that under the event $\mathcal{E}$, we have:

$$\begin{aligned}
&\left(\hat{\mu}_{k,t} + \sqrt{\frac{\log(2T/\delta)}{2\mathcal{Q}_{k,t}}}\right)\left(\hat{F}_k + \sqrt{\frac{2\log(2/\delta)}{HK}} + \frac{1}{K}\right) \;-\; \mathbb{P}[S \le p_k] \int_{p_k}^{1} \mathbb{P}[B \ge \lambda]\,\mathrm{d}\lambda \\
&\quad\le \left(\mathbb{P}[S \le p_k] + \sqrt{\frac{2\log(2T/\delta)}{\mathcal{Q}_{k,t}}}\right)\left(\int_{p_k}^{1} \mathbb{P}[B \ge \lambda]\,\mathrm{d}\lambda + \sqrt{\frac{8\log(2/\delta)}{HK}} + \frac{2}{K}\right) \\
&\quad\quad - \mathbb{P}[S \le p_k] \int_{p_k}^{1} \mathbb{P}[B \ge \lambda]\,\mathrm{d}\lambda \\
&\quad\le \sqrt{\frac{2\log(2T/\delta)}{\mathcal{Q}_{k,t}}} \int_{p_k}^{1} \mathbb{P}[B \ge \lambda]\,\mathrm{d}\lambda + \mathbb{P}[S \le p_k]\sqrt{\frac{8\log(2/\delta)}{HK}} + \frac{4\log(2T/\delta)}{\sqrt{HK\mathcal{Q}_{k,t}}} \\
&\quad\quad + \frac{2}{K}\left(\mathbb{P}[S \le p_k] + \sqrt{\frac{2\log(2T/\delta)}{\mathcal{Q}_{k,t}}}\right) \\
&\quad\le \underbrace{\sqrt{\frac{2\log(2T/\delta)}{\mathcal{Q}_{k,t}}}\mathbb{P}[B \ge p_k]}_{(\star)} \quad + \quad \frac{20\log(2T/\delta)}{T^{1/3}}.
\end{aligned}$$

Now, notice that, since $k \in \mathcal{K}$ by assumption, we have $\mathcal{Q}_{k,t} \ge \frac{1}{2}\mathcal{T}_{k,t}\mathbb{P}\left[B \ge p_k\right]$ for all $t > HK$, under the event $\mathcal{E}$. Therefore, the following holds:

$$(\star) = \sqrt{\frac{2\log(2T/\delta)}{\mathcal{Q}_{k,t}}}\mathbb{P}\left[B \ge p_k\right] \le \sqrt{\frac{4\log(2T/\delta)}{\mathcal{T}_{k,t}\mathbb{P}\left[B \ge p_k\right]}}\mathbb{P}[B \ge p_k] \le 2\frac{\log(2T/\delta)}{\sqrt{\mathcal{T}_{k,t}}}.$$

Putting all together, we have:

$$\begin{aligned}
&\left(\hat{\mu}_{k,t} + \sqrt{\frac{\log(2T/\delta)}{2\mathcal{Q}_{k,t}}}\right)\left(\hat{F}_k + \sqrt{\frac{2\log(2/\delta)}{HK}} + \frac{1}{K}\right) - \mathbb{P}[S \le p_k] \int_{p_k}^{1} \mathbb{P}[B \ge \lambda]\,\mathrm{d}\lambda \\
&\qquad\qquad\qquad \le \log\left(\frac{T}{\delta}\right)\mathcal{O}\left(\sqrt{\mathcal{T}_{k,t}} + \frac{1}{T^{\frac{1}{3}}}\right).
\end{aligned}$$

Finally, we notice that:

$$\mathbb{P}[S \le p_k] \int_{p_k}^{1} \mathbb{P}[B \ge \lambda]\,\mathrm{d}\lambda \le \left(\hat{\mu}_{k,t} + \sqrt{\frac{\log(2T/\delta)}{2\mathcal{Q}_{k,t}}}\right)\left(\hat{F}_k + \sqrt{\frac{2\log(2/\delta)}{HK}} + \frac{1}{K}\right),$$

as a direct consequence of being under the clean event $\mathcal{E}$. This concludes the proof. □

Lemma 4. *Let $k \in \mathcal{K}$. Then, for each $t > HK$, conditional on the event $\mathcal{E}$, we have:*

$$0 \le \left(\hat{\nu}_{k,t} + \sqrt{\frac{\log(2T/\delta)}{2\mathcal{T}_{k,t}}}\right)\left(\hat{G}_k + \sqrt{\frac{2\log(2/\delta)}{KN_k}} + \frac{1}{K}\right) - \mathbb{P}[B \ge p_k] \int_{0}^{p_k} \mathbb{P}[S \le \lambda]\,\mathrm{d}\lambda \le \eta,$$

where $\eta \coloneqq C\log\left(\frac{T}{\delta}\right)\left(\frac{1}{T^{1/3}} + \frac{1}{\sqrt{\mathcal{T}_{k,t}}}\right)$ and $C > 0$ is an absolute constant.

Proof. Notice that the first inequality is trivially given by the fact that, under the event $\mathcal{E}$, the quantity

$$\left(\hat{\nu}_{k,t} + \sqrt{\frac{\log(2T/\delta)}{2\mathcal{T}_{k,t}}}\right)\left(\hat{G}_k + \sqrt{\frac{2\log(2/\delta)}{KN_k}} + \frac{1}{K}\right)$$

is an (optimistic) estimator of

$$\mathbb{P}[B \geq p_k]\int_0^{p_k}\mathbb{P}[S \leq \lambda]\,\mathrm{d}\lambda.$$

For the second inequality, we notice that, under the event $\mathcal{E}$:

$$\begin{aligned}
&\left(\hat{\nu}_{k,t} + \sqrt{\frac{\log(2T/\delta)}{2\mathcal{T}_{k,t}}}\right)\left(\hat{G}_k + \sqrt{\frac{2\log(2/\delta)}{KN_k}} + \frac{1}{K}\right) - \mathbb{P}[B \geq p_k]\int_0^{p_k}\mathbb{P}[S \leq \lambda]\,\mathrm{d}\lambda \\
&\overset{\mathcal{E}\subset\mathcal{E}_2\cap\mathcal{E}_3}{\leq} \left(\mathbb{P}[B \geq p_k] + \sqrt{\frac{2\log(2T/\delta)}{\mathcal{T}_{k,t}}}\right)\left(\int_0^{p_k}\mathbb{P}[S \leq \lambda]\,\mathrm{d}\lambda + \sqrt{\frac{8\log(2/\delta)}{KN_k}} + \frac{2}{K}\right) \\
&\qquad - \mathbb{P}[B \geq p_k]\int_0^{p_k}\mathbb{P}[S \leq \lambda]\,\mathrm{d}\lambda \\
&\overset{\mathcal{E}\subset\mathcal{E}_1}{\leq} \sqrt{\frac{2\log(2T/\delta)}{\mathcal{T}_{k,t}}}\int_0^{p_k}\mathbb{P}[S \leq \lambda]\,\mathrm{d}\lambda + 4\mathbb{P}[B \geq p_k]\sqrt{\frac{\log(2/\delta)}{KH\mathbb{P}[B \geq p_k]}} \\
&\qquad + \frac{4\sqrt{2}\log(2T/\delta)}{\sqrt{KH\mathbb{P}[B \geq p_k]}\sqrt{\mathcal{T}_{k,t}}} + \frac{2}{K}\left(\mathbb{P}[B \geq p_k] + \sqrt{\frac{2\log(2T/\delta)}{\mathcal{T}_{k,t}}}\right) \\
&\overset{k\in\mathcal{K}}{=} \log\left(\frac{T}{\delta}\right)\cdot\mathcal{O}\left(\sqrt{\mathcal{T}_{k,t}} + \frac{1}{T^{\frac{1}{3}}}\right),
\end{aligned}$$

concluding the proof. □

Theorem 1. *Algorithm 2 guarantees regret $R_T = \tilde{\mathcal{O}}(T^{2/3})$.*

Proof. We first notice that, by defining

$$\mathcal{R}_T \coloneqq \sum_{t=1}^T \mathrm{g}(p^\star) - \sum_{t=1}^T \mathrm{g}(P_t),$$

we have that $R_T = \mathbb{E}[\mathcal{R}_T]$ and

$$\begin{aligned}
R_T &= \mathbb{E}[\mathcal{R}_T\mathbb{I}_{\mathcal{E}}] + \mathbb{E}[\mathcal{R}_T\mathbb{I}_{\mathcal{E}^c}] \\
&\leq \mathbb{E}[\mathcal{R}_T\mathbb{I}_{\mathcal{E}}] + \mathbb{E}[T\mathbb{I}_{\mathcal{E}^c}] \\
&\leq \mathbb{E}[\mathcal{R}_T\mathbb{I}_{\mathcal{E}}] + 6KT^2\delta = \mathbb{E}[\mathcal{R}_T\mathbb{I}_{\mathcal{E}}] + \mathcal{O}(T^{1/3}).
\end{aligned}$$

It is then sufficient to control the magnitude of $\mathcal{R}_T$ under the clean event $\mathcal{E}$. Hence, from this point on, we assume we are under the clean event $\mathcal{E}$.

Let $k^\star \in \arg\max_{k\in[K]} \mathrm{g}(p_k)$.

First, notice that, if $\mathbb{P}[p_{k^\star} \leq B] \leq 64T^{-1/3}\log(KT^2/\delta)$, then the decomposition in eq. (1) leads to

$$\begin{aligned}
\mathrm{g}(p_{k^\star}) &\leq \mathbb{P}[S \leq p_k]\cdot(1-p)\cdot\mathbb{P}[B \geq p_k] + \mathbb{P}[B \geq p_k]\cdot p\cdot\mathbb{P}[S \leq p_k] \\
&\leq \mathbb{P}[B \geq p_k] \leq 64T^{-1/3}\log(KT^2/\delta) = \tilde{\mathcal{O}}(T^{-1/3}).
\end{aligned}$$

Thus, when $\mathbb{P}[p_{k^\star} \leq B] \leq 64T^{-1/3}\log(KT^2/\delta)$, we have, due to Lemma 1, that if we pay an additional term whose instantaneous regret is upper bounded by L/K, we can control $\mathcal{R}_T$ by comparing our performance against the performance of the best point in the grid $p_{k^\star}$. This leads to:

$$\mathcal{R}_T = T\cdot\tilde{\mathcal{O}}(T^{-1/3}) + \frac{TL}{K} = \tilde{\mathcal{O}}(T^{2/3}).$$

Hence, we are left to analyze what happens when $\mathbb{P}[p_{k^\star} \leq B] > 64T^{-1/3}\log(KT^2/\delta)$, which we assume being the case from this point on. First, since $\mathbb{P}[p_{k^\star} \leq B] > 64T^{-1/3}\log(KT^2/\delta)$, given that $\mathcal{E} \subset \mathcal{E}_1$, it follows that $k^\star \in K^\diamond$.

We now notice that for each $k \in K^\diamond$ we have that $\mathcal{Q}_{k,HK} > 32T^{-1/3}\log(KT^2/\delta)$ by definition. In the clean event $\mathcal{E}$, we have that $\mathcal{E}_4$ holds, and hence for each $h \notin \mathcal{K}$ we have that $\mathcal{Q}_{h,HK} \leq 32T^{-1/3}\log(KT^2/\delta)$. It follows that, in the clean event $\mathcal{E}$, $k \in K^\diamond$ implies $k \in \mathcal{K}$, *i.e.*, $K^\diamond \subset \mathcal{K}$.

Now, we recall that Lemma 3 and Lemma 4 imply that, under the event $\mathcal{E}$, for all $t > HK$ and $k \in \mathcal{K}$:

$$\mathrm{g}(p_k) \leq \textbf{\textit{UCB}}_{t,k} \leq \mathrm{g}(p_k) + \eta_{k,t}\ , \tag{5}$$

where $\eta_{k,t} \coloneqq \tilde{C}\log\left(\frac{T}{\delta}\right)\left(\frac{1}{T^{1/3}} + \frac{1}{\sqrt{\mathcal{T}_{k,t}}}\right)$ and $\tilde{C} > 0$ is a universal constant.

If, for ebery $p \in [0,1]$, we define the quantity

$$\Delta_p \coloneqq g(p_{k^\star}) - g(p)\ ,$$

given that, for each $t > HK$, if $k_t \in K^\diamond$ is the index of $p_k = P_t$, we have

$$\mathrm{g}(p_{k^\star}) = \max_{k \in K^\diamond} \mathrm{g}(p_k) \leq \max_{k \in K^\diamond} \textbf{\textit{UCB}}_{k,t} = \textbf{\textit{UCB}}_{k_t,t} \leq \mathrm{g}(P_t) + \eta_{k_t,t}\ ,$$

we also have

$$\Delta_{P_t} \leq \eta_{k_t,t}\ .$$

In addition, by Lemma 1 and the fact that the instantaneous regret is upper bounded by 1, we have:

$$\mathcal{R}_T \leq HK + \frac{LT}{K} + \sum_{t=HK+1}^{T} \Delta_{P_t}.$$

Now, we have

$$\begin{aligned}
\sum_{t=HK+1}^{T} \Delta_{P_t} &= \sum_{k \in K^\diamond} \sum_{t=HK+1}^{T} \Delta_{P_t} \mathbb{I}\{P_t = p_k\} \\
&\leq \sum_{k \in K^\diamond} \sum_{t=HK+1}^{T} \eta_{k,t} \mathbb{I}\{P_t = p_k\} \\
&\leq \sum_{k \in K^\diamond} \sum_{t=HK+1}^{T} \left(\tilde{C}\log\left(\frac{T}{\delta}\right)\left(\frac{1}{T^{1/3}} + \frac{1}{\sqrt{\mathcal{T}_{k,t}}}\right)\right) \mathbb{I}\{P_t = p_k\} \\
&\leq \tilde{C}\log\left(\frac{T}{\delta}\right) \sum_{k \in K^\diamond} \sum_{t=HK+1}^{T} \left(\frac{1}{T^{1/3}} + \frac{1}{\sqrt{\mathcal{T}_{k,t}}}\right) \mathbb{I}\{P_t = p_k\} \\
&\leq \tilde{C}\log\left(\frac{T}{\delta}\right) \left(\sum_{k \in K^\diamond} \sum_{t=HK+1}^{T} \frac{\mathbb{I}\{P_t = p_k\}}{\sqrt{\mathcal{T}_{k,t}}} + T^{2/3}\right) \\
&\leq \tilde{C}\log\left(\frac{T}{\delta}\right) \left(\sum_{k \in K^\diamond} 2\sqrt{\mathcal{T}_{k,T}} + T^{2/3}\right) \\
&\leq \tilde{C}\log\left(\frac{T}{\delta}\right) \left(2\sqrt{KT} + T^{2/3}\right) = 4\tilde{C}\log\left(\frac{T}{\delta}\right) T^{2/3}\ ,
\end{aligned}$$

where the last inequality is Jensen's inequality. Hence

$$\mathbb{E}[\mathcal{R}_T \mathbb{I}_{\mathcal{E}}] \leq \mathbb{E}\left[4\tilde{C}\log\left(\frac{T}{\delta}\right) T^{2/3} \mathbb{I}_{\mathcal{E}}\right] \leq 4\tilde{C}\log\left(\frac{T}{\delta}\right) T^{2/3} = \tilde{\mathcal{O}}(T^{2/3})\ ,$$

concluding the proof. □