A. Proof of Lemma 2

The proof of Lemma 2 relies on the following lemma:

**Lemma 7.** For positive measurable functions \( f_i \), we have:

\[
\left( \int \prod_{i=1}^{K} f_i(x) \, dx \right)^{\frac{1}{K}} \leq \prod_{i=1}^{K} \int f_i^K(x) \, dx.
\]  

(4)

**Proof.** By Cauchy-Schwarz inequality, Eq. (4) is correct for \( K = 2 \). We use mathematical induction to prove that it still holds for \( K > 2 \). Specifically, assume InEq. (4) is true for \( K - 1 \), considering case \( K \), by Holder’s inequality we have:

\[
\left( \int \prod_{i=1}^{K} f_i(x) \, dx \right)^{\frac{1}{K}} \leq \left( \int f_K^{K}(x) \, dx \right)^{\frac{1}{K}} \left( \int \prod_{i=1}^{K-1} f_i(x) \, dx \right)^{\frac{K-1}{K}} \leq \prod_{i=1}^{K} \int f_i^K(x) \, dx.
\]

Now we present the proof of Lemma 2.

**Proof.** With straight-forward computations, we have:

\[
\sum_{k=1}^{K} P_k(\psi = k) = \sum_{k=1}^{K} \int_{\psi=k} dP_k = \sum_{k=1}^{K} \int_{\psi=k} \frac{dP_k}{dP_1} dP_1.
\]

Because we need a bound for any measurable function \( \psi \), we need to construct the \( \psi \) that minimizes the last expression. Obviously, the last expression is minimized for \( \psi(x) = \arg \min_{k \leq K} \frac{dP_k}{dP_1}(x) \), so that

\[
\sum_{k=1}^{K} P_k(\psi = k) \geq \int dP_1 \min_{k} \frac{dP_k}{dP_1} = \int \min_{k} dP_k.
\]

For vector \( \vec{P} = \{dP_1, \cdots, dP_K\} \), define \( r_i(\vec{P}) \) be the \( i \)-th smallest value in \( \vec{P} \), we have

\[
\left( \int \prod_{k} dP_k \right)^{\frac{1}{K}} = \left( \int \prod_{k} r_k(\vec{P}) \right)^{\frac{1}{K}} \leq \prod_{k} \int r_k(\vec{P}) = \int \min_{k} dP_k \prod_{k \geq 2} \int r_k(\vec{P}).
\]

The above inequality is proven by Lemma 7. Note that \( \sum_k dP_k \leq K \), so \( \prod_{k \geq 2} \int r_k(\vec{P}) \leq \left( \frac{K-1}{K-1} \right)^{K-1} < e \). With Jensen’s inequality, it yields that:

\[
\sum_{k=1}^{K} P_k(\psi = k) \geq \frac{1}{e} \left( \int \prod_{k} dP_k \right)^{\frac{1}{K}}
\]

\[
= \frac{1}{e} \exp \left\{ K \log \left( \int \prod_{k} dP_k \right)^{\frac{1}{K}} \right\}
\]

\[
= \frac{1}{e} \exp \left\{ K \log \left( \int \prod_{k} \frac{dP_k}{dP_1} \right)^{\frac{1}{K}} dP_1 \right\}
\]

\[
\geq \frac{1}{e} \exp \left\{ \sum_{k} dP_1 \log \frac{dP_k}{dP_1} \right\}
\]

\[
= \frac{1}{e} \exp \left\{ -\sum_{k=2}^{K} KL(P_1, P_k) \right\}.
\]
B. Proof of Lemma 3

Proof. With Hoeffding’s inequality and InEq. (2) in the main text, we have

\[ P \left( i : \frac{1}{u_i} \sum_{i=1}^{u_i} Y_i^s - \mu_s > \beta(u_i, t_i) \right) \leq \sum_{i \geq 1} P \left( \frac{1}{u_i} \sum_{i=1}^{u_i} Y_i^s - \mu_s > \beta(u_i, t_i) \right) \]

\[ \leq \sum_{i \geq 1} \exp \left\{ -2u_i \beta(u_i, t_i)^2 \right\} \]

\[ \leq \sum_{i \geq 1} \exp \left\{ -2u_i \beta(u_i, t_i)^2 \right\} \leq \delta \frac{\sqrt{\log K}}{2K}. \]

Then, applying the union bound, we complete the proof.

C. Proof of Lemma 4

In this section, we are going to show that \( E[\gamma - T(t, \gamma)] \leq O(H_c(j) \log \gamma) \) to complete the proof of Lemma 4. The proof is the same as that of Theorem 1 in (Auer et al., 2002), except some constants.

Proof. For simplicity, let \( s_1 \) denote \( s^*(j) \). Considering \( s : \mu_s < \mu_{s_1} \), let \( ev(s, \gamma) \) be the event that \( s \) is pulled at line 22 by Alg. 1 at round \( \gamma \). Let \( T_s(t, \gamma) \) denote the number of pulls on arm \( s \) in line 22 Alg. 1 when Alg. 1 selects column \( j \) for the \( \gamma \)-th time. Then, \( E[T_s(t, \gamma)] \) can be bounded as follows:

\[ E[T_s(t, \gamma)] = \sum_{\gamma'=1}^{\gamma} \mathbb{I} \left[ \text{The algorithm pulls arm } s \text{ in line 22 when Alg. 1 selects column } j \text{ for the } \gamma' \text{-th round} \right] \]

\[ \leq l + \sum_{\gamma'=1}^{\gamma} \mathbb{I} \left[ T_s(t', \gamma' - 1) \geq l, ev(s, \gamma') \right] \]

\[ \leq l + \sum_{\gamma'=1}^{\gamma} \mathbb{I} \left[ T_s(t', \gamma' - 1) \geq l, \bar{\mu}_{s_1} + \frac{2 \log t}{T_s(t', \gamma' - 1)} \leq \bar{\mu}_s + \frac{2 \log \gamma'}{T_s(t', \gamma' - 1)} \right] \]

\[ \leq l + \sum_{\gamma'=1}^{\gamma} \mathbb{I} \left[ T_s(t', \gamma' - 1) \geq l, \min_{\gamma_1 \in [1, \gamma' - 1]} \bar{\mu}_{s_1} + \frac{2 \log \gamma_{\gamma_1}}{\gamma_1} \leq \max_{\gamma_2 \in [l, \gamma' - 1]} \bar{\mu}_s + \frac{2 \log \gamma_{\gamma_2}}{\gamma_2} \right] \]

\[ \leq l + \sum_{\gamma'=1}^{\gamma} \sum_{\gamma_1=1}^{\gamma-1} \sum_{\gamma_2=1}^{\gamma-1} \mathbb{I} \left[ \bar{\mu}_{s_1} + \frac{2 \log \gamma_{\gamma_1}}{\gamma_1} \leq \bar{\mu}_s + \frac{2 \log \gamma_{\gamma_2}}{\gamma_2} \right]. \]

Since \( \bar{\mu}_{s_1} + \frac{2 \log \gamma_{\gamma_1}}{\gamma_1} \leq \bar{\mu}_s + \frac{2 \log \gamma_{\gamma_2}}{\gamma_2} \), at least one of the following three events happens:

- \( \bar{\mu}_{s_1} \leq \mu_{s_1} - \frac{2 \log \gamma_{\gamma_1}}{\gamma_1} \);
- \( \bar{\mu}_{s} \leq \mu_{s} - \frac{2 \log \gamma_{\gamma_2}}{\gamma_2} \);
- \( \mu_{s_1} \leq \mu_{s} + 2 \frac{2 \log \gamma_{\gamma_2}}{\gamma_2} \).

By Hoeffding’s inequality, the probability that the first and the second events happen is at most \( 2\gamma' - 4 \), and for \( l \geq \frac{1}{(\mu_{s_1} - \mu_{s})^2} \log \gamma \), the third event will not happen. So we have \( E[T_s(t, \gamma)] = O \left( \frac{1}{(\mu_{s_1} - \mu_{s})^2} \log \gamma \right) \). Then, with straightforward computations we complete the proof. \( \square \)
D. Proof of Lemma 5

Proof. Suppose $s^* \neq \text{none}$. If at round $\gamma$, Alg. 1 does not select $R_{s^*[1]}$ or $C_{s^*[2]}$, then at least one of following events happens:

- $\exists s \notin \{\text{none}, s^*\}, s = NE(\overline{M_t})$;
- $NE(\overline{M_t}) = \text{none}$.

For the first part, it is easy to see that $s \neq s^*_s(s[1])$ or $s \neq s^*_c(s[2])$. So by Lemma 4, the size of this part is at most $2\left(\sum_i \Lambda(H_r(i)) + \sum_j \lambda(H_c(j))\right)$.

Similarly, if the second case happens, then $s^* \neq \arg \min_{s' \in \text{row}(s^*)} \bar{\mu}_{s'}$ or $s^* \neq \arg \max_{s' \in \text{col}(s^*)} \bar{\mu}_{s'}$. When the second case happens for the $\gamma$-th round, due to Alg.1, we selected $R_{s^*[1]}$ for at least $\lceil \gamma/m \rceil$ times and $C_{s^*[2]}$ for at least $\lceil \gamma/n \rceil$ times. So by Lemma 4, the size of the second part is at most $\left((m + n)(\Lambda(H_r(s^*[1])) + \Lambda(H_c(s^*[2])))\right)$. \hfill \Box

E. Proof of Lemma 6

This proof is the same as the proof of Theorem 6 in Kalyanakrishnan et al. (2012), except the statement and some constants.

Proof. Without loss of generality, consider a bandit model $v = \{s_1, ..., s_k\}$. Suppose $\mu_{s_1} > \mu_{s_{i+1}}$. At each round, we pull arms in $v$ as from line 21 to line 23 in Alg. 1. Let $H = \sum_{t=2}^{\infty} \frac{1}{(\bar{\mu}_{s_1} - \bar{\mu}_{s_2})^2}$. Let $s_1(\gamma) = \arg \max_{s \in v} \bar{\mu}_s$ after round $\gamma$, and $s_2(\gamma) = \arg \max_{s \in v \setminus s_1(\gamma)} U(s, |T_s(\tau)|, |\tau|)$. Denote $\tau(\gamma)$ be the set of time steps $t$ in the first $\gamma$ round. If after $\gamma$ round, we have $L(s_1(\gamma), |T_{s_1}(\tau)|, |\tau|) > U(s_2(\gamma), |T_{s_2}(\tau)|, |\tau|)$, we now show that $P_{\gamma} = O(H \log \frac{H}{\delta})$.

Now let us introduce some notations. Define $c = \frac{\mu_{s_1} + \mu_{s_2}}{2}$ and

$$\Delta_i := \begin{cases} \mu_{s_1} - \mu_{s_2} & i = 1, \\ \mu_{s_2} - \mu_{s_1} & i \geq 2 \end{cases}$$

During round $\gamma$, we partition the set of arms into three subsets:

- $\text{Above}^\gamma := \{s \in v : \bar{\mu}_s > \beta(T_s(\gamma), \gamma) > c\}$.
- $\text{Below}^\gamma := \{s \in v : \bar{\mu}_s + \beta(T_s(\gamma), \gamma) < c\}$.
- $\text{Middle}^\gamma := v \setminus (\text{Above}^\gamma \cup \text{Below}^\gamma)$.

And by Lemma 2 in Kalyanakrishnan et al. (2012), if we cannot identify $\arg \max_{s \in v} \mu(s)$ after round $\gamma$, then $s_1(\gamma) \in \text{Middle}^\gamma$ or $s_2(\gamma) \in \text{Middle}^\gamma$. And with similar computation to Lemma 4 in Kalyanakrishnan et al. (2012), if for all arm $s_i$, we pulled at least $4\left[\frac{1}{\Delta_i} \ln \frac{mn\Delta_i}{\delta}\right]$ times, then $\forall s, s \notin \text{Middle}^\gamma$ with probability at least $1 - \frac{3bH}{mn\Delta_i \delta}$. The rest is all the same as Theorem 6 in Kalyanakrishnan et al. (2012).

\hfill \Box

F. Proof of Theorem 5

Proof. With the same argument as Theorem 8 in Even-Dar et al. (2006), for $s_1, s_2$, suppose $\mu_{s_1} - \mu_{s_2} = \Delta > 0$. Then with probability at least $1 - \delta/(mn)$, after round $\gamma = O\left(\frac{\ln(mn/\delta \Delta)}{\Delta^2}\right)$, we have $\bar{\mu}_{s_1} - \bar{\mu}_{s_2} > 2\beta_1(\gamma)$. So by the union bound, the racing algorithm stops after finite time with probability at least $1 - \delta$.

Then we show that Alg. 2 is $\delta$-PAC.

According to Lemma 3, the probability that there is an arm $s$ such that $|\bar{\mu}_s - \mu_s| > \beta_1(\gamma)$ for some $\gamma$ is at most $\delta$. Suppose $\forall s, \gamma, |\bar{\mu}_s - \mu_s| > \beta_1(\gamma)$. 

\hfill \Box
Identify the Nash Equilibrium in Static Games with Random Payoffs

For the case $NE(M) = \text{none}$, consider arm $s$. If $s \notin \{s^*_r(s[1]), s^*_c(s[2])\}$, clearly, it will be eliminated when its confidence bounds are disjoint with the confidence bounds of $s^*_r(s[1])$ and $s^*_c(s[2])$. Otherwise, if $s \in \{s^*_r(s[1]), s^*_c(s[2])\}$, without loss of generality, suppose $s = s^*_r(s[1])$. Obviously, there is a sequence $S = \{s_1, s_2, \cdots, s_{2k}\}$ such that (1) $s = s_1$; (2) $s_{2i+1} = s^*_r(s_{2i+1}[1])$ and $s_{2i} = s^*_c(s_{2i}[2])$ for all $i$, where $s_j = s_{(j-1)%(2k)+1}$; (3) $s_{2i+1} \in \text{col}(s_{2i})$ and $s_{2i+2} \in \text{row}(s_{2i+1})$. And according to Alg.2, all the arms in $S$ will be eliminated when their confidence bounds are disjoint.

For the case $s^* = NE(M) \neq \text{none}$, obviously, it won’t be eliminated by our elimination rule, so Algorithm is $\delta$-PAC.

Up to now we have proven the two statements about Alg. 2 in Theorem 5.

\qed

G. Correctness of baseline

Here, We present the stopping and recommendation rules of our baseline algorithm in detail.

Stopping and recommendation rules: In each round, we pull all arms. Let $\beta_2(\gamma) = \sqrt{\log(5mn\gamma^2/4\delta)/\gamma}$. After the $t$-th round, if one of the following event happens, the algorithm stops:

- For some arm $s$, if for all $s' \in \text{row}(s) \setminus s$, $\exists \gamma_{s'}$, after $\gamma_{s'}$ rounds, $\bar{\mu}_s + 2\beta_2(\gamma_{s'}, \delta) \leq \bar{\mu}_{s'}$ and for all $s' \in \text{col}(s) \setminus s, \exists \gamma(s')$, after $\gamma_{s'}$ rounds, $\bar{\mu}_s \geq \bar{\mu}_{s'} + 2\beta_2(\gamma_{s'}, \delta)$, then the recommendation rule recommends $s$ as the NE.

- For all arm $s$, if $\exists s' \in \text{row}(s), \bar{\mu}_{s'} + \beta_2(t, \delta) \leq \bar{\mu}_s - \beta_2(t, \delta)$ or $\exists s' \in \text{col}(s), \bar{\mu}_{s'} - \beta_2(t, \delta) \geq \bar{\mu}_s + \beta_2(t, \delta)$, then the recommendation rule determines that the underlying game does not have a NE.

Obviously, this algorithm is $\delta$-PAC and the proof for this statement is the same as the proof for Theorem 5.