A Upper Bound

Lemma 1. *Let $\overline{\mathcal{E}}$ be defined as in the proof of Theorem 1. Then*

$$P(\overline{\mathcal{E}}) \leq \frac{2(K + L)}{n}\,.$$

Proof. Let $\mathcal{E}_\ell = \mathcal{E}_\ell^{\text{U}} \cap \mathcal{E}_\ell^{\text{V}}$. Then from the definition of $\overline{\mathcal{E}}$,

$$\overline{\mathcal{E}} = \overline{\mathcal{E}_0} \cup (\overline{\mathcal{E}_1} \cap \mathcal{E}_0) \cup \ldots \cup (\overline{\mathcal{E}_{n-1}} \cap \mathcal{E}_{n-2} \cap \ldots \cap \mathcal{E}_0)\,,$$

and from the definition of $\mathcal{E}_\ell$,

$$\overline{\mathcal{E}_\ell} \cap \mathcal{E}_{\ell-1} \cap \ldots \cap \mathcal{E}_0 = (\overline{\mathcal{E}_\ell^{\text{U}}} \cap \mathcal{E}_{\ell-1} \cap \ldots \cap \mathcal{E}_0) \cup (\overline{\mathcal{E}_\ell^{\text{V}}} \cap \mathcal{E}_{\ell-1} \cap \ldots \cap \mathcal{E}_0)\,.$$

It follows that the probability of event $\overline{\mathcal{E}}$ is bounded as

$$\begin{aligned} P(\overline{\mathcal{E}}) &\leq \sum_{\ell=0}^{n-1} P(\overline{\mathcal{E}_\ell^{\text{U}}},\, \mathcal{E}_0^{\text{U}},\, \ldots,\, \mathcal{E}_{\ell-1}^{\text{U}},\, \mathcal{E}_0^{\text{V}},\, \ldots,\, \mathcal{E}_{\ell-1}^{\text{V}}) + \sum_{\ell=0}^{n-1} P(\overline{\mathcal{E}_\ell^{\text{V}}},\, \mathcal{E}_0^{\text{U}},\, \ldots,\, \mathcal{E}_{\ell-1}^{\text{U}},\, \mathcal{E}_0^{\text{V}},\, \ldots,\, \mathcal{E}_{\ell-1}^{\text{V}}) \\ &\leq \sum_{\ell=0}^{n-1} P(\overline{\mathcal{E}_\ell^{\text{U}}},\, \mathcal{E}_0^{\text{V}},\, \ldots,\, \mathcal{E}_{\ell-1}^{\text{V}}) + \sum_{\ell=0}^{n-1} P(\overline{\mathcal{E}_\ell^{\text{V}}},\, \mathcal{E}_0^{\text{U}},\, \ldots,\, \mathcal{E}_{\ell-1}^{\text{U}})\,. \end{aligned}$$

From the definition of $\overline{\mathcal{E}_\ell^{\text{U}}}$, it follows that

$$\begin{aligned} P(\overline{\mathcal{E}_\ell^{\text{U}}},\, \mathcal{E}_0^{\text{V}},\, \ldots,\, \mathcal{E}_{\ell-1}^{\text{V}}) \leq{} & P(\exists i \in \mathbf{I}_\ell \text{ s.t. } \bar{\mathbf{u}}_\ell(i) \notin [\mathbf{L}_\ell^{\text{U}}(i), \mathbf{U}_\ell^{\text{U}}(i)]) + {} \\ & P(\exists i \in \mathbf{I}_\ell \text{ s.t. } \bar{\mathbf{u}}_\ell(i) < \mu\bar{u}(i),\, \mathcal{E}_0^{\text{V}},\, \ldots,\, \mathcal{E}_{\ell-1}^{\text{V}})\,. \end{aligned}$$

Now we bound the probability of the above two events. The probability $P(\overline{\mathcal{E}_\ell^{\text{V}}},\, \mathcal{E}_0^{\text{U}},\, \ldots,\, \mathcal{E}_{\ell-1}^{\text{U}})$ can be bounded similarly and we omit this proof.

Event 1: $\exists i \in \mathbf{I}_\ell$ s.t. $\bar{\mathbf{u}}_\ell(i) \notin [\mathbf{L}_\ell^{\text{U}}(i), \mathbf{U}_\ell^{\text{U}}(i)]$

Fix any $i \in \mathbf{I}_\ell$. Let $\mathbf{c}_k$ be the k-th observation of row i in the row exploration stage of `Rank1Elim` and $\ell(k)$ be the index of that stage. Then

$$\left(\mathbf{c}_k - \bar{u}(i) \sum_{j=1}^{L} \frac{\bar{v}(\mathbf{h}_{\ell(k)}^{\text{V}}(j))}{L}\right)_{k=1}^{n}$$

is a martingale difference sequence with respect to history $\mathbf{h}_0^{\text{V}}, \ldots, \mathbf{h}_{\ell(k)}^{\text{V}}$ in step k. This follows from the observation that

$$\mathbb{E}\left[\mathbf{c}_k \,\middle|\, \mathbf{h}_0^{\text{V}}, \ldots, \mathbf{h}_{\ell(k)}^{\text{V}}\right] = \bar{u}(i) \sum_{j=1}^{L} \frac{\bar{v}(\mathbf{h}_{\ell(k)}^{\text{V}}(j))}{L}\,,$$

because column $j \in [L]$ in stage $\ell(k)$ is chosen randomly and then mapped to at least as rewarding column $\mathbf{h}_{\ell(k)}^{\text{V}}(j)$. By the definition of our sequence and from the Azuma-Hoeffding inequality (Remark 2.2.1 of Raginsky and Sason [26]),

$$\begin{aligned} P(\bar{\mathbf{u}}_\ell(i) \notin [\mathbf{L}_\ell^{\text{U}}(i), \mathbf{U}_\ell^{\text{U}}(i)]) &= P\left(\left|\frac{1}{n_\ell} \sum_{j=1}^{L} \mathbf{C}_\ell^{\text{U}}(i, j) - \bar{\mathbf{u}}_\ell(i)\right| > \sqrt{\frac{\log n}{n_\ell}}\right) \\ &= P\left(\left|\sum_{k=1}^{n_\ell}\left[\mathbf{c}_k - \bar{u}(i) \sum_{j=1}^{L} \frac{\bar{v}(\mathbf{h}_{\ell(k)}^{\text{V}}(j))}{L}\right]\right| > \sqrt{n_\ell \log n}\right) \\ &\leq 2\exp[-2\log n] \\ &= 2n^{-2} \end{aligned}$$

for any stage ℓ. By the union bound,

$$P(\exists i \in \mathbf{I}_\ell \text{ s.t. } \bar{\mathbf{u}}_\ell(i) \notin [\mathbf{L}^{\text{U}}_\ell(i), \mathbf{U}^{\text{U}}_\ell(i)]) \leq 2Kn^{-2}$$

for any stage ℓ.

Event 2: $\exists i \in \mathbf{I}_\ell$ s.t. $\bar{\mathbf{u}}_\ell(i) < \mu \bar{u}(i),\ \mathcal{E}^{\text{V}}_0,\ \ldots,\ \mathcal{E}^{\text{V}}_{\ell-1}$

We claim that this event cannot happen. Fix any $i \in \mathbf{I}_\ell$. When $\ell = 0$, we get that $\bar{\mathbf{u}}_0(i) = \bar{u}(i)(1/L)\sum_{j=1}^L \bar{v}(j) \geq \mu \bar{u}(i)$ from the definitions of $\bar{\mathbf{u}}_0(i)$ and μ, and event 2 obviously does not happen. When $\ell > 0$ and events $\mathcal{E}^{\text{V}}_0, \ldots, \mathcal{E}^{\text{V}}_{\ell-1}$ happen, any eliminated column j up to stage ℓ is substituted with column j' such that $\bar{v}(j') \geq \bar{v}(j)$, by the design of `Rank1Elim`. From this fact and the definition of $\bar{\mathbf{u}}_\ell(i)$, $\bar{\mathbf{u}}_\ell(i) \geq \mu \bar{u}(i)$. Therefore, event 2 does not happen when $\ell > 0$.

Total probability

Finally, we sum all probabilities up and get that

$$P(\overline{\mathcal{E}}) \leq n\left(\frac{2K}{n^2}\right) + n\left(\frac{2L}{n^2}\right) \leq \frac{2(K+L)}{n}\,.$$

This concludes our proof. ■

Lemma 2. *Let event $\mathcal{E}$ happen and m be the first stage where $\tilde{\Delta}_m < \mu \Delta^{\text{U}}_i / 2$. Then row i is guaranteed to be eliminated by the end of stage m. Moreover, let m be the first stage where $\tilde{\Delta}_m < \mu \Delta^{\text{V}}_j / 2$. Then column j is guaranteed to be eliminated by the end of stage m.*

Proof. We only prove the first claim. The other claim is proved analogously.

Before we start, note that by the design of `Rank1Elim` and from the definition of m,

$$\tilde{\Delta}_m = 2^{-m} < \frac{\mu \Delta^{\text{U}}_i}{2} \leq 2^{-(m-1)} = \tilde{\Delta}_{m-1}\,. \tag{6}$$

By the design of our confidence intervals,

$$\begin{aligned}
\frac{1}{n_m}\sum_{j=1}^K \mathbf{C}^{\text{U}}_m(i,j) + \sqrt{\frac{\log n}{n_m}} &\overset{\text{(a)}}{\leq} \bar{\mathbf{u}}_m(i) + 2\sqrt{\frac{\log n}{n_m}} \\
&= \bar{\mathbf{u}}_m(i) + 4\sqrt{\frac{\log n}{n_m}} - 2\sqrt{\frac{\log n}{n_m}} \\
&\overset{\text{(b)}}{\leq} \bar{\mathbf{u}}_m(i) + 2\tilde{\Delta}_m - 2\sqrt{\frac{\log n}{n_m}} \\
&\overset{\text{(c)}}{\leq} \bar{\mathbf{u}}_m(i) + \mu \Delta^{\text{U}}_i - 2\sqrt{\frac{\log n}{n_m}} \\
&= \bar{\mathbf{u}}_m(i^*) + \mu \Delta^{\text{U}}_i - [\bar{\mathbf{u}}_m(i^*) - \bar{\mathbf{u}}_m(i)] - 2\sqrt{\frac{\log n}{n_m}}\,,
\end{aligned}$$

where inequality (a) is from $\mathbf{L}^{\text{U}}_m(i) \leq \bar{\mathbf{u}}_m(i)$, inequality (b) is from $n_m \geq 4\tilde{\Delta}_m^{-2} \log n$, and inequality (c) is by (6). Now note that

$$\bar{\mathbf{u}}_m(i^*) - \bar{\mathbf{u}}_m(i) = q(\bar{u}(i^*) - \bar{u}(i)) \geq \mu \Delta^{\text{U}}_i$$

for some $q \in [0,1]$. The equality holds because $\bar{\mathbf{u}}_m(i^*)$ and $\bar{\mathbf{u}}_m(i)$ are estimated from the same sets of random columns. The inequality follows from the fact that events $\mathcal{E}^{\text{V}}_0, \ldots, \mathcal{E}^{\text{V}}_{m-1}$ happen. The events imply that any eliminated column j up to stage m is substituted with column j' such that $\bar{v}(j') \geq \bar{v}(j)$, and thus $q \geq \mu$. From the above inequality, we get that

$$\bar{\mathbf{u}}_m(i^*) + \mu \Delta^{\text{U}}_i - [\bar{\mathbf{u}}_m(i^*) - \bar{\mathbf{u}}_m(i)] - 2\sqrt{\frac{\log n}{n_m}} \leq \bar{\mathbf{u}}_m(i^*) - 2\sqrt{\frac{\log n}{n_m}}\,.$$

Finally,

$$\bar{\mathbf{u}}_m(i^*) - 2\sqrt{\frac{\log n}{n_m}} \overset{\text{(a)}}{\leq} \frac{1}{n_m}\sum_{j=1}^{K}\mathbf{C}^{\text{U}}_m(i^*, j) - \sqrt{\frac{\log n}{n_m}}$$
$$\overset{\text{(b)}}{\leq} \frac{1}{n_m}\sum_{j=1}^{K}\mathbf{C}^{\text{U}}_m(\mathbf{i}_m, j) - \sqrt{\frac{\log n}{n_m}}\,,$$

where inequality (a) follows from $\bar{\mathbf{u}}_m(i^*) \leq \mathbf{U}^{\text{U}}_m(i^*)$ and inequality (b) follows from $\mathbf{L}^{\text{U}}_m(i^*) \leq \mathbf{L}^{\text{U}}_m(\mathbf{i}_m)$, since $i^* \in \mathbf{I}_m$ and $\mathbf{i}_m = \arg\max_{i \in \mathbf{I}_m} \mathbf{L}^{\text{U}}_m(i)$. Now we chain all inequalities and get our final claim. ■

Lemma 3. *The expected cumulative regret due to exploring any row $i \in [K]$ and any column $j \in [L]$ is bounded as*

$$\mathbb{E}\left[\sum_{\ell=0}^{n-1}\mathbb{E}\left[\mathbf{R}^{\text{U}}_\ell(i) \,|\, \mathcal{H}_\ell\right]\mathbb{1}\{\mathcal{F}_\ell\}\right] \leq \frac{384}{\mu^2\bar{\Delta}^{\text{U}}_i}\log n + 1\,,$$
$$\mathbb{E}\left[\sum_{\ell=0}^{n-1}\mathbb{E}\left[\mathbf{R}^{\text{V}}_\ell(j) \,|\, \mathcal{H}_\ell\right]\mathbb{1}\{\mathcal{F}_\ell\}\right] \leq \frac{384}{\mu^2\bar{\Delta}^{\text{V}}_j}\log n + 1\,.$$

Proof. We only prove the first claim. The other claim is proved analogously. This proof has two parts. In the first part, we assume that row i is suboptimal, $\Delta^{\text{U}}_i > 0$. In the second part, we assume that row i is optimal, $\Delta^{\text{U}}_i = 0$.

Row i is suboptimal

Let row i be suboptimal and m be the first stage where $\tilde{\Delta}_m < \mu\Delta^{\text{U}}_i/2$. Then row i is guaranteed to be eliminated by the end of stage m (Lemma 2), and thus

$$\mathbb{E}\left[\sum_{\ell=0}^{n-1}\mathbb{E}\left[\mathbf{R}^{\text{U}}_\ell(i) \,|\, \mathcal{H}_\ell\right]\mathbb{1}\{\mathcal{F}_\ell\}\right] \leq \mathbb{E}\left[\sum_{\ell=0}^{m}\mathbb{E}\left[\mathbf{R}^{\text{U}}_\ell(i) \,|\, \mathcal{H}_\ell\right]\mathbb{1}\{\mathcal{F}_\ell\}\right].$$

By Lemma 4, the expected regret of choosing row i in stage ℓ can be bounded from above as

$$\mathbb{E}\left[\mathbf{R}^{\text{U}}_\ell(i) \,|\, \mathcal{H}_\ell\right]\mathbb{1}\{\mathcal{F}_\ell\} \leq (\Delta^{\text{U}}_i + \max_{j \in \mathbf{J}_\ell}\Delta^{\text{V}}_j)(n_\ell - n_{\ell-1})\,,$$

where $\max_{j \in \mathbf{J}_\ell}\Delta^{\text{V}}_j$ is the maximum column gap in stage ℓ, n_ℓ is the number of steps by the end of stage ℓ, and $n_{-1} = 0$. From the definition of $\mathcal{F}_\ell$ and $\tilde{\Delta}_\ell$, if column j is not eliminated before stage ℓ, we have that

$$\Delta^{\text{V}}_j \leq \frac{2\tilde{\Delta}_{\ell-1}}{\mu} = \frac{2 \cdot 2^{m-\ell+1}\tilde{\Delta}_m}{\mu} < 2^{m-\ell+1}\Delta^{\text{U}}_i\,.$$

From the above inequalities and the definition of n_ℓ, it follows that

$$\begin{aligned}
\mathbb{E}\left[\sum_{\ell=0}^{m}\mathbb{E}\left[\mathbf{R}^{\text{U}}_\ell(i) \,|\, \mathcal{H}_\ell\right]\mathbb{1}\{\mathcal{F}_\ell\}\right] &\leq \sum_{\ell=0}^{m}(\Delta^{\text{U}}_i + \max_{j \in \mathbf{J}_\ell}\Delta^{\text{V}}_j)(n_\ell - n_{\ell-1}) \\
&\leq \sum_{\ell=0}^{m}(\Delta^{\text{U}}_i + 2^{m-\ell+1}\Delta^{\text{U}}_i)(n_\ell - n_{\ell-1}) \\
&\leq \Delta^{\text{U}}_i\left(n_m + \sum_{\ell=0}^{m}2^{m-\ell+1}n_\ell\right) \\
&\leq \Delta^{\text{U}}_i\left(2^{2m+2}\log n + 1 + \sum_{\ell=0}^{m}2^{m-\ell+1}(2^{2\ell+2}\log n + 1)\right) \\
&= \Delta^{\text{U}}_i\left(2^{2m+2}\log n + 1 + \sum_{\ell=0}^{m}2^{m+\ell+3}\log n + \sum_{\ell=0}^{m}2^{m-\ell+1}\right) \\
&\leq \Delta^{\text{U}}_i(5 \cdot 2^{2m+2}\log n + 2^{m+2}) + 1 \\
&\leq 6 \cdot 2^4 \cdot 2^{2m-2}\Delta^{\text{U}}_i\log n + 1\,,
\end{aligned}$$

where the last inequality follows from $\log n \geq 1$ for $n \geq 3$. From the definition of $\tilde{\Delta}_{m-1}$ in (6), we have that

$$2^{m-1} = \frac{1}{\tilde{\Delta}_{m-1}} \leq \frac{2}{\mu \Delta_i^{\text{U}}}\,.$$

Now we chain all above inequalities and get that

$$\mathbb{E}\left[\sum_{\ell=0}^{n-1} \mathbb{E}\left[\mathbf{R}_\ell^{\text{U}}(i) \,|\, \mathcal{H}_\ell\right] \mathbb{1}\{\mathcal{F}_\ell\}\right] \leq 6 \cdot 2^4 \cdot 2^{2m-2} \Delta_i^{\text{U}} \log n + 1 \leq \frac{384}{\mu^2 \Delta_i^{\text{U}}} \log n + 1\,.$$

This concludes the first part of our proof.

Row i is optimal

Let row i be optimal and m be the first stage where $\tilde{\Delta}_m < \mu \Delta_{\min}^{\text{V}}/2$. Then similarly to the first part of the analysis,

$$\mathbb{E}\left[\sum_{\ell=0}^{n-1} \mathbb{E}\left[\mathbf{R}_\ell^{\text{U}}(i) \,|\, \mathcal{H}_\ell\right] \mathbb{1}\{\mathcal{F}_\ell\}\right] \leq \sum_{\ell=0}^{m} (\max_{j \in \mathbf{J}_\ell} \Delta_j^{\text{V}})(n_\ell - n_{\ell-1}) \leq \frac{384}{\mu^2 \Delta_{\min}^{\text{V}}} \log n + 1\,.$$

This concludes our proof. ■

Lemma 4. *Let $\mathbf{u} \sim P_{\text{U}}$ and $\mathbf{v} \sim P_{\text{V}}$ be drawn independently. Then the expected regret of choosing any row $i \in [K]$ and column $j \in [L]$ is bounded from above as*

$$\mathbb{E}\left[\mathbf{u}(i^*)\mathbf{v}(j^*) - \mathbf{u}(i)\mathbf{v}(j)\right] \leq \Delta_i^{\text{U}} + \Delta_j^{\text{V}}\,.$$

Proof. Note that for any $x, y, x^*, y^* \in [0, 1]$,

$$x^*y^* - xy = x^*y^* - xy^* + xy^* - xy = y^*(x^* - x) + x(y^* - y) \leq (x^* - x) + (y^* - y)\,.$$

By the independence of the entries of $\mathbf{u}$ and $\mathbf{v}$, and from the above inequality,

$$\mathbb{E}\left[\mathbf{u}(i^*)\mathbf{v}(j^*) - \mathbf{u}(i)\mathbf{v}(j)\right] = \bar{u}(i^*)\bar{v}(j^*) - \bar{u}(i)\bar{v}(j) \leq (\bar{u}(i^*) - \bar{u}(i)) + (\bar{v}(j^*) - \bar{v}(j))\,.$$

This concludes our proof. ■

B Lower Bound

In this section we present the missing details of the proof of Theorem 2. Recall that we need to bound from below the value of $f(\bar{u}, \bar{v})$ where

$$\begin{aligned} f(\bar{u}, \bar{v}) = & \inf_{c \in [0,\infty)^{K \times L}} \sum_{i=1}^{K} \sum_{j=1}^{L} (\bar{u}(i^*)\bar{v}(j^*) - \bar{u}(i)\bar{v}(j)) c_{i,j} \\ & \text{s.t. } \forall (\bar{u}', \bar{v}') \in B(\bar{u}, \bar{v}) : \\ & \qquad \sum_{i=1}^{K} \sum_{j=1}^{L} d(\bar{u}(i)\bar{v}(j), \bar{u}'(i)\bar{v}'(j)) c_{i,j} \geq 1 \end{aligned}$$

and

$$B(\bar{u}, \bar{v}) = \{(\bar{u}', \bar{v}') \in [0,1]^K \times [0,1]^L : \bar{u}(i^*) = \bar{u}'(i^*),\, \bar{v}(j^*) = \bar{v}'(j^*),\, w^*(\bar{u}, \bar{v}) < w^*(\bar{u}', \bar{v}')\}\,.$$

Without loss of generality, we assume that the optimal action in the original model $(\bar{u}, \bar{v})$ is $(i^*, j^*) = (1, 1)$. Moreover, we consider a class of *identifiable* bandit models, meaning that we assume that

$$\forall (i, i', j, j') \in [0,1]^{2K} \times [0,1]^{2L}, \quad (i, j) \neq (i', j') \implies 0 < d(\bar{u}(i)\bar{v}(j), \bar{u}(i')\bar{v}(j')) < +\infty.$$

This implies in particular that $\bar{u}(i^*)\bar{v}(j^*)$ must be less than 1. An intuitive justification of this assumption is the following. Remark that for the Bernoulli problem we consider here, if the mean of the best arm is exactly 1, the rewards from optimal pulls are always 1 so that the empirical average is always exactly 1 and as we cap the UCBs to 1, the optimal arm is always a candidate to the next pull, which leads to constant regret. Also note that by our assumption, the optimal action is unique. To get a lower bound, we consider the same optimization problem as above, but replace B with its subset. Clearly, this can only decrease the optimal value.

Concretely, we consider only those models in $B(\bar{u}, \bar{v})$ where only one parameter changes at a time. Let

$$B_{\text{U}}(\bar{u}, \bar{v}) = \{(\bar{u}', \bar{v}) : \bar{u}' \in [0,1]^K, \exists i_0 \in \{2, \dots, K\}, \epsilon \in [0,1] \text{ s.t. } [\forall i \neq i_0 : \bar{u}'(i) = \bar{u}(i)] \text{ and } \bar{u}'(i_0) = \bar{u}(1) + \epsilon\}\,,$$
$$B_{\text{V}}(\bar{u}, \bar{v}) = \{(\bar{u}, \bar{v}') : \bar{v}' \in [0,1]^L, \exists j_0 \in \{2, \dots, L\}, \epsilon \in [0,1] \text{ s.t. } [\forall j \neq j_0 : \bar{v}'(j) = \bar{v}(j)] \text{ and } \bar{v}'(j_0) = \bar{v}(1) + \epsilon\}\,.$$

Let $f'(\bar{u}, \bar{v})$ be the optimal value of the above optimization problem when $B(\bar{u}, \bar{v})$ is replaced by $B_{\text{U}}(\bar{u}, \bar{v}) \cup B_{\text{V}}(\bar{u}, \bar{v}) \subset B(\bar{u}, \bar{v})$. Now suppose that $(\bar{u}', \bar{v}') \in B_{\text{U}}(\bar{u}, \bar{v})$ and $i_0 = 2$. Then, for any $i \neq 2$ and $j \in [L]$, $d(\bar{u}(i)\bar{v}(j), \bar{u}'(i)\bar{v}'(j)) = 0$; and for $i = 2$ and any $j \in [L]$, $d(\bar{u}(i)\bar{v}(j), \bar{u}'(i)\bar{v}'(j)) = d(\bar{u}(2)\bar{v}(j), (\bar{u}(1) + \epsilon)\bar{v}(j))$. Hence,

$$\sum_{i=1}^{K} \sum_{j=1}^{L} d(\bar{u}(i)\bar{v}(j), \bar{u}'(i)\bar{v}'(j)) = \sum_{j=1}^{L} d(\bar{u}(2)\bar{v}(j), (\bar{u}(1) + \epsilon)\bar{v}(j))\,.$$

Reasoning similarly for $B_{\text{V}}(\bar{u}, \bar{v})$, we see that $f'(\bar{u}, \bar{v})$ satisfies

$$\begin{aligned} f'(\bar{u}, \bar{v}) = & \inf_{c \in [0,\infty)^{K \times L}} \sum_{i=1}^{K} \sum_{j=1}^{L} (\bar{u}(i^*)\bar{v}(j^*) - \bar{u}(i)\bar{v}(j)) c_{i,j} \\ & \text{s.t.} \quad \forall \epsilon_{\text{V}} \in (0, 1 - \bar{v}(1)], \epsilon_{\text{U}} \in (0, 1 - \bar{u}(1)] \\ & \qquad \forall j \neq 1, \sum_{i=1}^{K} d(\bar{u}(i)\bar{v}(j), \bar{u}(i)(\bar{v}(1) + \epsilon_{\text{V}})) c_{i,j} \geq 1 \\ & \qquad \forall i \neq 1, \sum_{j=1}^{L} d(\bar{u}(i)\bar{v}(j), (\bar{u}(1) + \epsilon_{\text{U}})\bar{v}(j)) c_{i,j} \geq 1. \end{aligned}$$

Clearly, the smaller the coefficients of $c_{i,j}$ in the constraints, the tighter the constraints. We obtain the smallest coefficients when $\epsilon_{\text{V}}, \epsilon_{\text{U}} \to 0$. By continuity, we get

$$\begin{aligned} f'(\bar{u}, \bar{v}) = & \inf_{c \in [0,\infty)^{K \times L}} \sum_{i=1}^{K} \sum_{j=1}^{L} (\bar{u}(i^*)\bar{v}(j^*) - \bar{u}(i)\bar{v}(j)) c_{i,j} \\ & \text{s.t.} \quad \forall j \neq 1, \sum_{i=1}^{K} d(\bar{u}(i)\bar{v}(j), \bar{u}(i)\bar{v}(1)) c_{i,j} \geq 1 \\ & \qquad \forall i \neq 1, \sum_{j=1}^{L} d(\bar{u}(i)\bar{v}(j), \bar{u}(1)\bar{v}(j)) c_{i,j} \geq 1. \end{aligned}$$

Let

$$c_{i,j} = \begin{cases} 1/d(\bar{u}(i)\bar{v}(1), \bar{u}(1)\bar{v}(1))\,, & j = 1 \text{ and } i > 1\,; \\ 1/d(\bar{u}(1)\bar{v}(j), \bar{u}(1)\bar{v}(1))\,, & i = 1 \text{ and } j > 1\,; \\ 0\,, & \text{otherwise.} \end{cases}$$

We claim that $(c_{i,j})$ is an optimal solution for the problem defining f'.

First, we show that $(c_{i,j})$ is feasible. Let $i \neq 1$. Then $\sum_{j=1}^{L} d(\bar{u}(i)\bar{v}(j), \bar{u}(1)\bar{v}(j)) c_{i,j} = d(\bar{u}(i)\bar{v}(1), \bar{u}(1)\bar{v}(1)) c_{i,1} = 1$. Similarly, we can verify the other constraint, too, showing that $(c_{i,j})$ is indeed feasible.

Now, it remains to show that the proposed solution is indeed optimal. We prove this by contradiction, following the ideas of [6]. We suppose that there exists a solution c of the optimization problem such that $c_{i_0,j_0} > 0$ for $i_0 \neq 1$ and $j_0 \neq 1$. Then, we prove that it is possible to find another feasible solution c' but with an objective lower than that obtained with c, contradicting the assumption of optimality of c.

We define c' as follows, redistributing the mass of c_{i_0,j_0} on the first row and the first column:

$$c'_{i,j} = \begin{cases} 0\,, & i = i_0 \text{ and } j = j_0\,; \\ c_{i_0,1} + c_{i_0,j_0}\dfrac{d(\bar{u}(i_0)\bar{v}(j_0), \bar{u}(1)\bar{v}(j_0))}{d(\bar{u}(i_0)\bar{v}(1), \bar{u}(1)\bar{v}(1))}\,, & i = i_0 \text{ and } j = 1\,; \\ \\ c_{1,j_0} + c_{i_0,j_0}\dfrac{d(\bar{u}(i_0)\bar{v}(j_0), \bar{u}(i_0)\bar{v}(1))}{d(\bar{u}(1)\bar{v}(j_0), \bar{u}(1)\bar{v}(1))}\,, & i = 1 \text{ and } j = j_0\,; \\ c_{i,j}\,, & \text{otherwise.} \end{cases}$$

It is easily verified that if c satisfies the constraints, then so does c' because the missing mass of c_{i_0,j_0} is simply redistributed on $c'_{i_0,1}$ and c'_{1,j_0}. For example, for $i = i_0$ we have

$$\begin{aligned} &\sum_{j=1}^{L} d(\bar{u}(i_0)\bar{v}(j), \bar{u}(1)\bar{v}(j))c'_{i_0,j} - \sum_{j=1}^{L} d(\bar{u}(i_0)\bar{v}(j), \bar{u}(1)\bar{v}(j))c_{i_0,j} \\ &= d(\bar{u}(i_0)\bar{v}(1), \bar{u}(1)\bar{v}(1))c_{i_0,j_0}\frac{d(\bar{u}(i_0)\bar{v}(j_0), \bar{u}(1)\bar{v}(j_0))}{d(\bar{u}(i_0)\bar{v}(1), \bar{u}(1)\bar{v}(1))} - c_{i_0,j_0}d(\bar{u}(i_0)\bar{v}(j_0), \bar{u}(1)\bar{v}(j_0)) \\ &= 0 \end{aligned}$$

while for $i \notin \{1, i_0\}$, $c'_{i,j} = c_{i,j}$, so $\sum_{j=1}^{L} d(\bar{u}(i)\bar{v}(j), \bar{u}(1)\bar{v}(j))c'_{i,j} = \sum_{j=1}^{L} d(\bar{u}(i)\bar{v}(j), \bar{u}(1)\bar{v}(j))c_{i,j}$.

Now, we prove that the objective function is lower for c' than for c by showing that the difference between them is negative:

$$\begin{aligned} \Delta &\doteq \sum_{i=1}^{K}\sum_{j=1}^{L}(\bar{u}(1)\bar{v}(1) - \bar{u}(i)\bar{v}(j))c'_{i,j} - \sum_{i=1}^{K}\sum_{j=1}^{L}(\bar{u}(1)\bar{v}(1) - \bar{u}(i)\bar{v}(j))c_{i,j} \\ &= \quad c_{i_0,j_0}\,(\bar{u}(1)\bar{v}(1) - \bar{u}(i_0)\bar{v}(1))\frac{d(\bar{u}(i_0)\bar{v}(j_0), \bar{u}(1)\bar{v}(j_0))}{d(\bar{u}(i_0)\bar{v}(1), \bar{u}(1)\bar{v}(1))} \\ &\quad + c_{i_0,j_0}(\bar{u}(1)\bar{v}(1) - \bar{u}(1)\bar{v}(j_0))\frac{d(\bar{u}(i_0)\bar{v}(j_0), \bar{u}(i_0)\bar{v}(1))}{d(\bar{u}(1)\bar{v}(j_0), \bar{u}(1)\bar{v}(1))} \\ &\quad - c_{i_0,j_0}(\bar{u}(1)\bar{v}(1) - \bar{u}(i_0)\bar{v}(j_0)) \\ &= \quad c_{i_0,j_0}\Big\{(\bar{u}(1) - \bar{u}(i_0))\bar{v}(1)\frac{d(\bar{u}(i_0)\bar{v}(j_0), \bar{u}(1)\bar{v}(j_0))}{d(\bar{u}(i_0)\bar{v}(1), \bar{u}(1)\bar{v}(1))} \\ &\qquad + (\bar{v}(1) - \bar{v}(j_0))\bar{u}(1)\frac{d(\bar{u}(i_0)\bar{v}(j_0), \bar{u}(i_0)\bar{v}(1))}{d(\bar{u}(1)\bar{v}(j_0), \bar{u}(1)\bar{v}(1))} \\ &\qquad - (\bar{u}(1)\bar{v}(1) - \bar{u}(i_0)\bar{v}(j_0))\Big\} \end{aligned}$$

Writing

$$\bar{u}(1)\bar{v}(1) - \bar{u}(i_0)\bar{v}(j_0) = (\bar{u}(1) - \bar{u}(i_0))\bar{v}(j_0) + (\bar{v}(1) - \bar{v}(j_0))\bar{u}(1)$$

we get

$$\begin{aligned} \Delta = c_{i_0,j_0}(\bar{u}(1) - \bar{u}(i_0))\left(\bar{v}(1)\frac{d(\bar{u}(i_0)\bar{v}(j_0), \bar{u}(1)\bar{v}(j_0))}{d(\bar{u}(i_0)\bar{v}(1), \bar{u}(1)\bar{v}(1))} - \bar{v}(j_0)\right) \\ + c_{i_0,j_0}(\bar{v}(1) - \bar{v}(j_0))\left(\bar{u}(1)\frac{d(\bar{u}(i_0)\bar{v}(j_0), \bar{u}(i_0)\bar{v}(1))}{d(\bar{u}(1)\bar{v}(j_0), \bar{u}(1)\bar{v}(1))} - \bar{u}(1)\right). \end{aligned}$$

To finish the proof, it suffices to prove that both terms of the above sum are negative. First, $\bar{u}(1) - \bar{u}(i_0), \bar{v}(1) - \bar{v}(j_0), c_{i_0,j_0} > 0$, hence it remains to consider the terms involving the ratios of KL divergences. Note that both ratios take the form $\frac{d(\alpha p, \alpha q)}{d(p,q)}$ with $\alpha < 1$, but one must be compared to $\alpha < 1$ while the other can simply be compared to 1. For the first such term, showing the negativity of the difference is equivalent to showing that for $\alpha = \bar{v}(j_0)/\bar{v}(1) < 1$,

$$\frac{d(\alpha\bar{u}(i_0)\bar{v}(1), \alpha\bar{u}(1)\bar{v}(1))}{d(\bar{u}(i_0)\bar{v}(1), \bar{u}(1)\bar{v}(1))} < \alpha.$$

Lemma 5 below shows that for fixed $(p,q) \in (0,1)^2$, $f : \alpha \mapsto d(\alpha p, \alpha q)$ is convex, which proves the above inequality. For the second term, it remains to see whether the ratio of the KL divergences is below one. Lemma 5 proven below shows that the function $\alpha \mapsto d(\alpha p, \alpha q)$ is increasing on $(0,1)$, showing that

$$\frac{d(\bar{u}(i_0)\bar{v}(j_0), \bar{u}(i_0)\bar{v}(1))}{d(\bar{u}(1)\bar{v}(j_0), \bar{u}(1)\bar{v}(1))} < 1\,.$$

Thus, the proof is finished once we prove Lemma 5.

Lemma 5. *Let p, q be any fixed real numbers in $(0,1)$. The function $f : \alpha \mapsto d(\alpha p, \alpha q)$ is convex and increasing on $(0,1)$. As a consequence, for any $\alpha < 1$, $d(\alpha p, \alpha q) < d(p,q)$.*

Proof. We first re-parametrize our problem into polar coordinates (r, θ) :

$$\begin{cases} p &= r\cos\theta \\ q &= r\sin\theta \end{cases}$$

In order to prove the statement of the lemma, it now suffices to prove that $f_\theta : r \mapsto d(r\sin\theta, r\cos\theta)$ is increasing. We have

$$f_\theta(r) = r\cos\theta\log\left(\frac{\cos\theta}{\sin\theta}\right) + (1 - r\cos\theta)\log\left(\frac{1-r\cos\theta}{1-r\sin\theta}\right)$$

which can be differentiated along r for a fixed θ :

$$f'_\theta(r) = cos\theta\log\left(\frac{1-r\sin\theta}{1-r\cos\theta}\right) + \frac{\sin\theta-\cos\theta}{1-r\sin\theta} + cos\theta\log\left(\frac{\cos\theta}{\sin\theta}\right).$$

Now, we can differentiate again along r and after some calculations we obtain

$$f''_\theta(r) = \frac{(\sin\theta-\cos\theta)^2}{(1-r\sin\theta)^2(1-r\cos\theta)} > 0$$

which proves that the function f_θ is convex. It remains to prove that $f'_\theta(0) \geq 0$ for any $\theta \in (0, \pi/2)$. We rewrite $f'_\theta(0)$ as a function of θ :

$$\begin{aligned} f'_\theta(0) &= \cos\theta\log\left(\frac{\cos\theta}{\sin\theta}\right) + \sin\theta - \cos\theta \\ &:= \phi(\theta) \end{aligned}$$

Let us assume that there exists $\theta_0 \in (0, \pi/2)$ such that $\phi(\theta_0) < 0$. Then, in this direction $f'_\theta(0) < 0$ and as $f_\theta(0) = 0$ for any $\theta \in (0, \pi/2)$, it means that there exists $r_0 > 0$ such that $f_{\theta_0}(r_0) < 0$. Yet, $f_{\theta_0}(r_0) = d(r_0\cos\theta_0, r_0\sin\theta_0) > 0$ because of the positivity of the KL divergence.

So by contradiction, we proved that for all $\theta \in (0, \pi/2)$, $f'_\theta(0) = \phi(\theta) \geq 0$ and by convexity f_θ is non-negative and non-decreasing on $[0, +\infty)$.

∎

B.1 Gaussian payoffs

The lower bound naturally extends to other classes of distributions, such as Gaussians. For illustration here we show the lower bound for this case. We still assume that the means are in $[0,1]$, as before. We also assume that all payoffs have a common variance $\sigma^2 > 0$. Recall that the Kullback-Leibler divergence between two distributions with fixed variance σ^2 is $d(p,q) = (p-q)^2/(2\sigma^2)$. Then, the proof of Theorem 2 can be repeated with minor differences (in particular, the proof of the analogue of Lemma 5 becomes trivial) and we get the following result:

Theorem 3. *For any $(\bar{u}, \bar{v}) \in [0,1]^K \times [0,1]^L$ with a unique optimal action and any uniformly efficient algorithm $\mathcal{A}$ whose regret is $R(n)$, assuming Gaussian row and column rewards with common variance σ^2,*

$$\liminf_{n\to\infty} \frac{R(n)}{\log(n)} \geq \frac{2\sigma^2}{\bar{v}(j^*)} \sum_{i \in [K]\setminus\{i^*\}} \frac{1}{\Delta^{\mathrm{U}}_i} + \frac{2\sigma^2}{\bar{u}(i^*)} \sum_{j \in [L]\setminus\{j^*\}} \frac{1}{\Delta^{\mathrm{V}}_j}\,.$$