

A Proofs of lower bounds

A.1 Group-sparse dominance

The set $\Lambda(\mu)$ in the optimization problem can be decomposed into $\Lambda(\mu) = \Lambda_k(\mu) \sqcup \dots \sqcup \Lambda_K(\mu)$ where $\Lambda_k(\mu)$ is the set of alternative parameters in which arm k of category 1 is optimal. Indeed, as we know that $\lambda_1^1 = \mu_1^1 > 0$, the best category is known and the regret incurred by suboptimal categories is non-existent. Thus, asymptotically, we fall back into deriving a lower bound on the regret in one category, i.e. in the classic multi-armed bandit setting.

A.2 Strong dominance

Without loss of generality, we assume that we have $M = 2$ categories and category 2 has a unique worst arm. The condition in the optimization problem can be written as

$$\sum_{k=2}^K N_k^1 (\mu_k^1 - \lambda_k^2)^2 + \sum_{k=1}^K N_k^2 (\mu_k^2 - \lambda_k^2)^2 \geq 2, \forall \lambda \in \Lambda(\mu),$$

where $\Lambda(\mu) = \Lambda_2(\mu) \sqcup \dots \sqcup \Lambda_K(\mu) \sqcup \Lambda^2(\mu)$ where $\Lambda_k(\mu)$ is the event in which the best arm is mistaken by arm k in the category 1, i.e.,

$$\Lambda_k(\mu) = \{\mu_1^1 \times\} - \infty, \mu_1^1 [\times \dots \times] \mu_1^1, +\infty [\times \dots \times] - \infty, \mu_1^1 [\times] - \infty, \mu_1^1 [\times \dots \times] - \infty, \mu_1^1 [$$

and $\Lambda^2(\mu)$ is the event in which we mistake category 2 as the optimal category, i.e.,

$$\Lambda^2(\mu) = \{\mu_1^1 \times\} - \infty, \mu_1^1 [\times \dots \times] - \infty, \mu_1^1 [\times] \mu_1^1, +\infty [\times \dots \times] \mu_1^1, +\infty [.$$

On $\Lambda_k(\mu)$, the condition is equivalent to

$$N_k^1 (\mu_1^1 - \mu_k^2)^2 \geq 2,$$

and on $\Lambda^2(\mu)$,

$$\sum_{k=1}^K N_k^2 (\mu_1^1 - \mu_k^2)^2 \geq 2.$$

The minimization problem can thus be separated in two parts: the first part corresponds to finding the best arm in the optimal category and the second part to finding the optimal category.

For the first part, the solution is the same as in the multi-armed bandit setting and is given by $N_k^1 = \frac{2}{(\Delta_{1,k})^2}$.

For the second part, let us prove that the solution is given by $N_K^2 = \frac{2}{(\Delta_{2,K})^2}$ and $N_k^2 = 0$ for $k \neq K$. We have the following problem

$$\min_{N^2 \geq 0} \sum_{k=1}^K N_k^2 \Delta_{2,k} =: f(N^2) \quad \text{subject to} \quad \sum_{k=1}^K N_k^2 (\Delta_{2,k})^2 \geq 2.$$

On one side, we have

$$\min_{N \geq 0} f(N) \leq \min_{n \geq 0} f(0, \dots, 0, n) = f\left(0, \dots, 0, \frac{2}{(\Delta_{2,K})^2}\right) = \frac{2}{\Delta_{2,K}},$$

and on the other side, since $\Delta_{2,k} < \Delta_{2,K}$, we have

$$\sum_{k=1}^K N_k^2 \Delta_{2,k} > \frac{1}{\Delta_{2,K}} \sum_{k=1}^K N_k^2 (\Delta_{2,k})^2 \geq \frac{2}{\Delta_{2,K}}.$$

Hence the solution of the optimization problem in the suboptimal category and the lower bound on the regret follows.

A.3 First-order dominance

By simplifying the optimization problem, one obtains the two following conditions

$$\forall k \neq 1, N_k^1 (\Delta_{1,k})^2 \geq 2,$$

and $\forall k \in [K]$,

$$\begin{aligned} & \sum_{i=1}^{k-1} \left[\left(N_{i+1}^1 (\mu_{i+1}^1 - \tilde{\mu}_i)^2 + N_i^2 (\mu_i^2 - \tilde{\mu}_i)^2 \right) \mathbf{1} \{ \mu_i^2 < \mu_{i+1}^1 \} \right] + N_k^2 (\Delta_{2,k})^2 \\ & + \sum_{j=k+1}^K \left(N_j^1 (\mu_j^1 - \bar{\mu}_j)^2 + N_j^2 (\mu_j^2 - \bar{\mu}_j)^2 \right) \geq 2, \end{aligned}$$

where $\tilde{\mu}_i = \frac{N_{i+1}^1 \mu_{i+1}^1 + N_i^2 \mu_i^2}{N_{i+1}^1 + N_i^2}$ and $\bar{\mu}_j = \frac{N_j^1 \mu_j^1 + N_j^2 \mu_j^2}{N_j^1 + N_j^2}$.

Assuming the arms are intertwined, the first term in the above equation disappear since the condition in the indicator function is not verified. In the case of $M = 2$ categories and two arms per category $K = 2$, the following conditions are derived

$$N_2^1 \geq \frac{2}{(\Delta_{1,2})^2}, \quad N_2^2 \geq \frac{2}{(\Delta_{2,2})^2},$$

and

$$N_1^2 (\Delta_{2,1})^2 + N_2^1 (\mu_2^1 - \bar{\mu})^2 + N_2^2 (\mu_2^2 - \bar{\mu})^2 \geq 2,$$

where $\bar{\mu} = \frac{N_2^1 \mu_2^1 + N_2^2 \mu_2^2}{N_2^1 + N_2^2}$.

Since this is a minimization problem, it is clear that the regret is minimize on the lower bounds of N_2^1 and N_2^2 . Putting this two quantities in the last inequality, we obtain

$$N_1^2 \geq \frac{2}{(\Delta_{2,1})^2} \left[1 - \left(\left(\frac{\mu_2^1 - \bar{\mu}}{\Delta_{1,2}} \right)^2 + \left(\frac{\mu_2^2 - \bar{\mu}}{\Delta_{2,2}} \right)^2 \right) \right].$$

Developing $\bar{\mu}$, we have

$$\bar{\mu} = \frac{\frac{2\mu_2^1}{(\Delta_{1,2})^2} + \frac{2\mu_2^2}{(\Delta_{2,2})^2}}{\frac{2}{(\Delta_{1,2})^2} + \frac{2}{(\Delta_{2,2})^2}} = \frac{\mu_2^1 (\Delta_{2,2})^2 + \mu_2^2 (\Delta_{1,2})^2}{(\Delta_{1,2})^2 + (\Delta_{2,2})^2}.$$

Now developing $\frac{\mu_2^1 - \bar{\mu}}{\Delta_{1,2}}$, we get:

$$\frac{\mu_2^1 - \bar{\mu}}{\Delta_{1,2}} = \frac{\Delta_{1,2} (\mu_2^1 - \mu_2^2)}{(\Delta_{1,2})^2 + (\Delta_{2,2})^2} = \frac{\Delta_{1,2} \Delta_{2,2}^{1,2}}{(\Delta_{1,2})^2 + (\Delta_{2,2})^2}.$$

Similarly,

$$\frac{\mu_2^2 - \bar{\mu}}{\Delta_{2,2}} = -\frac{\Delta_{2,2} \Delta_{2,2}^{1,2}}{(\Delta_{1,2})^2 + (\Delta_{2,2})^2}.$$

Plugging this into the inequality on N_1^2 , we obtain

$$N_1^2 \geq \frac{2}{(\Delta_{2,1})^2} \left[1 - \frac{(\Delta_{2,2}^{1,2})^2}{(\Delta_{1,2})^2 + (\Delta_{2,2})^2} \right].$$

The result follows by the decomposition of the expected regret.

B Characterizations of dominance

B.1 Strong dominance

Let $(e_i)_i$ denotes the unit vectors. Taking $\mathbf{x} = e_k$ and $\mathbf{y} = e_l$ hands $\mu_k^1 \geq \mu_l^2$.

In the other direction, let $(\alpha, \beta) \in \Delta(K) \times \Delta(K)$. We have

$$\langle \alpha, \mu \rangle = \sum_{k=1}^K \alpha_k \mu_k = \sum_{k=1}^{K-1} \alpha_k \mu_k + \left(1 - \sum_{k=1}^{K-1} \alpha_k\right) \mu_K = \mu_K + \sum_{k=1}^{K-1} \alpha_k (\mu_k - \mu_K).$$

Now, using the previous equality, we obtain

$$\langle \alpha, \mu^1 \rangle - \langle \beta, \mu^2 \rangle = \sum_{k=1}^K \alpha_k \mu_k^1 - \sum_{k=1}^K \beta_k \mu_k^2 = (\mu_K^1 - \mu_1^2) + \sum_{k=1}^{K-1} \alpha_k (\mu_k^1 - \mu_K^1) + \sum_{k=2}^K \beta_k (\mu_1^2 - \mu_k^2) \geq 0.$$

B.2 First-order dominance

Taking $\mathbf{x} = e_k$ hands $\mu_k^1 \geq \mu_k^2$. In the other direction, let $\mathbf{x} \in \Delta(K)$. We have

$$\langle \mathbf{x}, \mu^1 - \mu^2 \rangle = \sum_{k=1}^K \mathbf{x}_k (\mu_k^1 - \mu_k^2) \geq 0.$$

C Regret upper bounds of CATSE

C.1 Group-sparse dominance

Consider the following clean event

$$\mathcal{E}_s = \left\{ \forall t \in [T], \forall k \in [K], |\widehat{\mu}_k^1(t) - \mu_k^1| \leq \sqrt{\frac{2 \log \frac{1}{\delta}}{N_k^1(t)}} \right\}.$$

Using union bounds over t and k , one obtains thanks to the subGaussian assumption that $\mathbb{P}(\mathcal{E}_s) \geq 2\delta KT$. In the following, we assume the clean event holds true. In the case in which only the optimal category is active, we get the regret of the UCB algorithm

$$R_T \leq \sum_{k=2}^K \frac{8 \log \frac{1}{\delta}}{\Delta_{1,k}}.$$

On the other hand, the set of active categories is empty if the optimal category is non active. That means that $\forall k \leq s, \widehat{\mu}_k^1(N_k^1(t)) < 2\sqrt{\frac{\log N_k^1(t)}{N_k^1(t)}}$ where s is the number of arms with positive expected reward. Let \mathcal{A}_s denote this event. The number of times it happen is bounded. Indeed, since

$$\mathcal{A}_s \subseteq \left\{ \widehat{\mu}_1^1(N_1^1(t)) < 2\sqrt{\frac{\log N_1^1(t)}{N_1^1(t)}} \right\} =: \mathcal{A}_1,$$

and

$$n \geq 3 + \frac{32}{(\mu_1^1)^2} \log \frac{16}{(\mu_1^1)^2} \Rightarrow 2\sqrt{\frac{\log n}{n}} - \mu_1^1 \leq -\frac{\mu_1^1}{2},$$

we have

$$\begin{aligned}
\mathbb{E} \left[\sum_{t=MK+1}^T \mathbf{1} \{ \mathcal{A}_s \} \right] &\leq \mathbb{E} \left[\sum_{t=MK+1}^T \mathbf{1} \{ \mathcal{A}_1 \} \right] \\
&\leq \left(3 + \frac{32}{(\mu_1^1)^2} \log \frac{16}{(\mu_1^1)^2} \right) + \sum_{u=1}^T \mathbb{P} \left(\widehat{\mu}_1^1(u) - \mu_1^1 < -\frac{\mu_k^1}{2} \right) \\
&\leq \left(3 + \frac{32}{(\mu_1^1)^2} \log \frac{16}{(\mu_1^1)^2} \right) + \sum_{u=1}^T \exp \left\{ -\frac{u}{8} (\mu_1^1)^2 \right\} \\
&\leq 3 + \frac{32}{(\mu_1^1)^2} \log \frac{16}{(\mu_1^1)^2} + \frac{8}{(\mu_1^1)^2}.
\end{aligned}$$

Finally, the set of active categories has more than one element if a sub-optimal category is active, i.e. $\exists m \neq 1, \exists k \in [K]; \widehat{\mu}_k^m(N_k^m(t)) \geq 2\sqrt{\frac{\log N_k^m(t)}{N_k^m(t)}}$. Let \mathcal{B} denote this event. The number of times it happen is also bounded. Indeed,

$$\begin{aligned}
\mathbb{E} \sum_{t=1}^T \mathbf{1} \{ \mathcal{B} \} &\leq \sum_{m,k} \sum_{u=1}^T \mathbb{P} \left(\widehat{\mu}_k^m(u) \geq 2\sqrt{\frac{\log u}{u}} \right) \\
&\leq \sum_{m,k} \sum_{u=1}^T \mathbb{P} \left(\widehat{\mu}_k^m(u) - \mu_k^m \geq 2\sqrt{\frac{\log u}{u}} \right) \\
&\leq \sum_{m,k} \sum_{u=1}^T \frac{1}{u^2} \leq (M-1)K \frac{\pi^2}{6}.
\end{aligned}$$

Combining the three inequalities, we conclude.

C.2 Strong dominance

Let \mathcal{E}_0 denote the clean event

$$\mathcal{E}_0 = \left\{ \forall t \in [T]; \forall m \in [M], \forall \mathbf{x} \in \mathbb{R}^K, \langle \mathbf{x}, \widehat{\mu}^m(t) - \mu^m \rangle \leq \|\mathbf{x}\|_2 \beta(t, \delta) \right\},$$

where $\beta(t, \delta) = \sqrt{\frac{2}{N^m(t)}} \left(K \log 2 + \log \frac{1}{\delta} \right)$.

Lemma 2. *With probability at least $1 - \delta$, the following holds uniformly overall all $\mathbf{x} \in \mathbb{R}^K$,*

$$\langle \mathbf{x}, \widehat{\mu}^m(p) - \mu^m \rangle \leq \|\mathbf{x}\|_2 \sqrt{\frac{2}{p}} \left(K \log 2 + \log \frac{1}{\delta} \right).$$

Proof. Fix $x \in \mathbb{R}$ and $\delta \in (0, 1)$ a confidence level. According to (Lattimore and Szepesvári, 2018), we have with probability at least $1 - \delta$,

$$\|\widehat{\mu}(t) - \mu\|_{V_t} \leq \sqrt{2 \left(K \log 2 + \log \frac{1}{\delta} \right)}.$$

If an agent pulls each arm sequentially, we are in the fixed design setting. In this case, (assuming t is a multiple of K), we have $V_t = N(t)\mathbf{I}_K$, i.e. it is a diagonal matrix and we conclude. \square

Using union bounds over the time and the categories, and using the definition of the confidence set, we obtain $\mathbb{P}(\mathcal{E}_0^c) \leq \delta MT$.

Suppose we are in the clean event and let $m \neq 1$ and t be the last time when we did not invoke the stopping rule, i.e. that the category m is still active. First remark that category 1 is never eliminated by category m on the clean event since $\min_k \mu_k^1 \geq \max_k \mu_k^m$. By Equation (1), this means that

$$\forall \mathbf{x} \in \Delta(K), \forall \mathbf{y} \in \Delta(K), \langle \mathbf{x}, \widehat{\mu}^1(t) \rangle - \langle \mathbf{y}, \widehat{\mu}^m(t) \rangle \leq (\|\mathbf{x}\|_2 + \|\mathbf{y}\|_2) \sqrt{\frac{2}{N(t)}} \left(\log \frac{1}{\delta} + K \log 2 \right),$$

where $N(t)$ denotes the number of times each category have been pulled. As we are in the clean event, we have

$$\forall \mathbf{x} \in \Delta(K), \forall \mathbf{y} \in \Delta(K), \langle \mathbf{x}, \mu^1 \rangle - \langle \mathbf{y}, \mu^m \rangle \leq 2(\|\mathbf{x}\|_2 + \|\mathbf{y}\|_2) \sqrt{\frac{2}{N(t)} \left(\log \frac{1}{\delta} + K \log 2 \right)}.$$

Inverting this equation, we obtain the following upper bound on $N(t)$

$$\forall \mathbf{x} \in \Delta(K), \forall \mathbf{y} \in \Delta(K), N(t) \leq 8 \left(K \log 2 + \log \frac{1}{\delta} \right) \left(\frac{\|\mathbf{x}\|_2 + \|\mathbf{y}\|_2}{\langle \mathbf{x}, \mu^1 \rangle - \langle \mathbf{y}, \mu^m \rangle} \right)^2.$$

The proof is conclude with the proof of the UCB algorithm [3].

C.3 First-order dominance

Lemma 3. *With probability at least $1 - \delta$,*

$$\|\widehat{\mu}_{\sigma_t^m}^m(t) - \mu^m\|_2 \leq \frac{1}{\sqrt{2t}} \left(\sqrt{K \log \frac{1}{\delta}} + \sqrt{1 + (K+1) \log K} \right),$$

where $\widehat{\mu}_{\sigma_t^m}^m(t)$ denotes the vector $\widehat{\mu}^m(t)$ ordered in decreasing order.

Proof. The McDiarmid inequality gives the following

$$\mathbb{P} \left\{ \|\widehat{\mu}_{\sigma_t^m}^m(t) - \mu^m\| \geq \mathbb{E} \|\widehat{\mu}_{\sigma_t^m}^m(t) - \mu^m\| + \varepsilon \right\} \leq \exp(-2t\varepsilon^2/K)$$

Now we just has to bound $\mathbb{E} \|\widehat{\mu}_{\sigma_t^m}^m(t) - \mu^m\|_2$. If Y_1, \dots, Y_N are σ^2 sub-Gaussian, then

$$\mathbb{P} \left\{ \max_{i=1, \dots, N} Y_i \geq \varepsilon \right\} \leq N \exp \left(-\frac{\varepsilon^2}{2\sigma^2} \right).$$

This give, by a careful integration, that

$$\mathbb{E} \left(\max_{i=1, \dots, N} Y_i \right)^2 \leq 2\sigma^2(\log(N) + 1).$$

In our case, we have $\sigma^2 = \frac{1}{4t}$. Using that the expectation of the k^{th} maximum of N random variables is smaller than the expectation of the maximum of $N - (k - 1)$ random variables [11], we obtain

$$\mathbb{E} \|\widehat{\mu}_{\sigma_t^m}^m(t) - \mu^m\|_2^2 \leq \frac{1}{2t} \sum_{k=1}^K (1 + \log(K - (k - 1))) = \frac{1}{2t} (K + \log K!) \leq \frac{1 + (K + 1) \log K}{2t},$$

where the last inequality comes from the Stirling formulae. The result follows. \square

Let define the clean event

$$\mathcal{E}_1 = \left\{ \forall t \in [T], \forall m \in [m], \|\widehat{\mu}_{\sigma_t^m}^m(t) - \mu^m\|_2 \leq \frac{1}{\sqrt{2t}} \left(\sqrt{K \log \frac{1}{\delta}} + \sqrt{1 + (K+1) \log K} \right) \right\}$$

By the lemma and with union bounds over t and m , we have $\mathbb{P}(\mathcal{E}_1^c) \leq \delta MT$. Let $m \neq 1$ and t be the last time we pulled category m .

By Equation (2), we have

$$\forall \mathbf{x} \in \Delta(K), \langle \mathbf{x}, \widehat{\mu}_{\sigma_t^1}^1(t) - \widehat{\mu}_{\sigma_t^m}^m(t) \rangle \leq 2\|\mathbf{x}\|_2 \gamma(t, \delta).$$

Moreover, notice that after t samples

$$\begin{aligned} \forall \mathbf{x} \in \Delta(K), \frac{1}{\|\mathbf{x}\|_2} \left| \langle \mathbf{x}, \widehat{\mu}_{\sigma_t^1}^1(t) - \widehat{\mu}_{\sigma_t^m}^m(t) \rangle - \langle \mathbf{x}, \mu^1 - \mu^m \rangle \right| &\leq \|\widehat{\mu}_{\sigma_t^1}^1(t) - \mu^1\|_2 + \|\widehat{\mu}_{\sigma_t^m}^m(t) - \mu^m\|_2 \\ &\leq 2\gamma(t, \delta), \end{aligned}$$

where the last inequality holds true with probability at least $1 - \delta MT$. Combining the two inequalities, one obtains with probability at least $1 - \delta MT$,

$$\begin{aligned} N^m(t) &\leq \frac{8}{\|\mu^1 - \mu^m\|_2^2} \left(\sqrt{K \log \frac{1}{\delta}} + \sqrt{1 + (K+1) \log K} \right)^2 \\ &\leq \frac{16}{\|\mu^1 - \mu^m\|_2^2} \left(K \log \frac{1}{\delta} + K \log K + \log K + 1 \right) \end{aligned}$$

where in the last inequality we used the Cauchy–Schwarz inequality. Hence the result.