Assume $\theta_1^* > \theta_2^* \geq \ldots \geq \theta_n^*$ (unique best-arm). $\Delta_i = \theta_1^* - \theta_i^*$

Assume $X_{i,s}$ is 1-sub-Gaussian $\left(\mathbb{E}[\exp(\lambda X_{i,s})] \leq e^{\lambda^2/2} \right)$

Input: n arms $\mathcal{X} = \{1, \ldots, n\}$, confidence level $\delta \in (0,1)$.
Let $\widehat{\mathcal{X}}_1 \leftarrow \mathcal{X}, \ell \leftarrow 1$
while $|\widehat{\mathcal{X}}_\ell| > 1$ **do**
  $\epsilon_\ell = 2^{-\ell}$
  Pull each arm in $\widehat{\mathcal{X}}_\ell$ exactly $\tau_\ell = \lceil 2\epsilon_\ell^{-2} \log(\frac{4\ell^2|\mathcal{X}|}{\delta}) \rceil$ times
  Compute the empirical mean of these rewards $\widehat{\theta}_{i,\ell}$ for all $i \in \widehat{\mathcal{X}}_\ell$
  $\widehat{\mathcal{X}}_{\ell+1} \leftarrow \widehat{\mathcal{X}}_\ell \setminus \{i \in \widehat{\mathcal{X}}_\ell : \max_{j \in \widehat{\mathcal{X}}_\ell} \widehat{\theta}_{j,\ell} - \widehat{\theta}_{i,\ell} > 2\epsilon_\ell\}$
  $\ell \leftarrow \ell + 1$
Output: $\widehat{\mathcal{X}}_{\ell+1}$ OR play single element of $\hat{\mathcal{X}}_l$ for remaining time T.

$\mathcal{E}_{i,\ell} = \{ |\hat{\theta}_{i,\ell} - \theta_i^*| \leq \varepsilon_\ell \}, \quad \mathcal{E} = \bigcap_{\ell=1}^{\infty} \bigcap_{i=1}^{n} \mathcal{E}_{i,\ell}$

$$\mathbb{P}(\mathcal{E}^c) = \mathbb{P}\left(\bigcup_{\ell=1}^{\infty} \bigcup_{i=1}^{n} \mathcal{E}_{i,\ell}^c \right)$$

$$\leq \sum_{\ell=1}^{\infty} \sum_{i=1}^{n} \mathbb{P}(\mathcal{E}_{i,\ell}^c) \leq \sum_{\ell=1}^{\infty} \sum_{i=1}^{n} \frac{\delta}{2\ell^2 n}$$

$$= \sum_{\ell=1}^{\infty} \frac{\delta}{2\ell^2} \leq \delta$$

$$\mathbb{P}(\mathcal{E}_{i,\ell}^c) = \mathbb{P}\left(|\hat{\theta}_{i,\ell} - \theta_i^*| > \varepsilon_\ell \right)$$

$$\leq \mathbb{P}\left(\hat{\theta}_{i,\ell} - \theta_i^* > \varepsilon_\ell \right) + \mathbb{P}\left(\hat{\theta}_{i,\ell} - \theta_i^* < -\varepsilon_\ell \right)$$

$$\leq 2 \exp\left(- \tau_\ell \varepsilon_\ell^2 / 2 \right)$$

$$\leq 2 \exp\left(- \cancel{\varepsilon_\ell^2 / 2 \cdot 2 \varepsilon_\ell^{-2}} \log\left(\frac{4\ell^2 |\mathcal{X}|}{\delta} \right) \right)$$

$$= 2 \cdot \frac{\delta}{4\ell^2 |\mathcal{X}|} = \frac{\delta}{2\ell^2 n}$$

Lemma: Assume $\max_i \Delta_i \le 4$. With prob $\ge 1-\delta$ we have $\{1\} \in \hat{\mathcal{X}}_\ell \; \forall \ell$ and $\max_{i \in \hat{\mathcal{X}}_\ell} \Delta_i \le 8 \varepsilon_\ell$.

Proof: Assume $\mathcal{E}$ holds since it hold w.p. $\ge 1-\delta$.

Fix ℓ and assume $\{1\} \in \hat{\mathcal{X}}_\ell$. (note $\{1\} \in \hat{\mathcal{X}}_1$)

For any $j \in \hat{\mathcal{X}}_\ell$ we have

$$\hat{\theta}_{j,\ell} - \hat{\theta}_{1,\ell} = \overbrace{(\hat{\theta}_{j,\ell} - \theta_j^*)}^{\le \varepsilon_\ell} - \overbrace{(\hat{\theta}_{1,\ell} - \theta_1^*)}^{\ge -\varepsilon_\ell} - \Delta_j$$

$$\overset{\mathcal{E}_j \cap \mathcal{E}_1}{\le} 2\varepsilon_\ell - \Delta_j$$

$$< 2\varepsilon_\ell$$

$$\Rightarrow \max_{j \in \hat{\mathcal{X}}_\ell} \hat{\theta}_{j,\ell} - \hat{\theta}_{1,\ell} \le 2\varepsilon_\ell \Rightarrow \{1\} \in \hat{\mathcal{X}}_{\ell+1}$$

(Second part) Fix i: $\Delta_i = \theta_1^* - \theta_i^* > 4\varepsilon_\ell = 8\varepsilon_{\ell+1}$

$$\max_{j \in \hat{\mathcal{X}}_\ell} \hat{\theta}_{j,\ell} - \hat{\theta}_{i,\ell} \ge \hat{\theta}_{1,\ell} - \hat{\theta}_{i,\ell}$$

$$= \underbrace{(\hat{\theta}_{1,\ell} - \theta_1^*)}_{\ge -\varepsilon_\ell} - \underbrace{(\hat{\theta}_{i,\ell} - \theta_i^*)}_{\le \varepsilon_\ell} + \Delta_i$$

$$\geq -2\varepsilon_\ell + \Delta_i$$

$$> 2\varepsilon_\ell$$

$\Rightarrow$ i is discarded $\Rightarrow i \notin \hat{\mathcal{X}}_{\ell+1}$

$$\Rightarrow \max_{i \in \hat{\mathcal{X}}_\ell} \Delta_i \leq 8\varepsilon_\ell$$

$\boxtimes$

$\boxed{\text{Theorem}}$ Assume $\max_i \Delta_i \leq 4$. Then w.p. $\geq 1-\delta$ $\hat{\mathcal{X}}_\ell = \{1\}$ once τ total measurements are taken

where $$\tau \leq c \sum_{i=2}^{n} \Delta_i^{-2} \log\left(n \log(\Delta_i^{-2})\right).$$

$(a \vee b) := \max(a, b)$

$\boxed{\text{Theorem}}$ Assume $\max_i \Delta_i \leq 4$. w.p. $\geq 1-\delta$

$$\sum_{i=2}^{n} \Delta_i T_i \leq \underbrace{\inf_{\nu \geq 0} \nu T + c \sum_{i=2}^{n} (\Delta_i \vee \nu)^{-1} \log\left(\frac{\log((\Delta \vee \nu)^{-1})|\mathcal{X}|}{\delta}\right)}.$$

$\boxed{\text{Corollary}}$ If $\delta = 1/T$ then

$$R_T = \sum_{i=2}^{n} \Delta_i \mathbb{E}[T_i] \leq \delta \cdot T + \quad 1 .$$

$$\leq \begin{cases} c \sum_{i=2}^{n} \Delta_i^{-1} \log\left(\log(\Delta_i^{-1}) |\mathcal{X}| T\right) & \nu = 0 \\ c\sqrt{nT\log(T)} & \nu \approx \sqrt{\frac{n}{T}} \end{cases}$$

After round ℓ, # meas taken $= \underbrace{\sum_{\ell} |\mathcal{X}_\ell| \, \tau_\ell}_{\approx n 2^{2\ell}}$

$\max_{i \in \hat{\mathcal{X}}_{\ell+1}} \leq 4 \varepsilon_\ell \approx 2^{-\ell}$

$\Rightarrow$ After T measurements

$\hat{i} := \arg\max_{i \in \hat{\mathcal{X}}_\ell} \hat{\theta}_{i,\ell}$ then $\theta^*_{\hat{i}} \geq \theta^*_1 - c\sqrt{\frac{n \log(T)}{T}}$