We can ... $H_1(\mu^{[a]}) \leq H_1(\mu)$

s.t.

$$\max\left(e_T(\mu_*),\ e_T(\mu^{[a]})\right) \geq \frac{1}{4} e^{-\frac{2T}{H_1(\mu)}}$$

<u>Proof</u>: <u>Claim</u>: $\exists\, a \in \{2, \ldots, k\}$ s.t.

14/03/24

$$E_\mu[N_a(T)] \leq \frac{T}{H_1 \Delta_a^2}$$

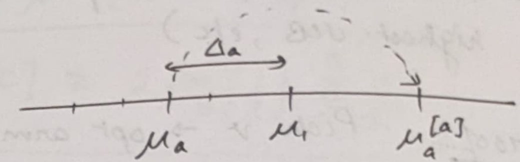
<u>Reason</u>: If $E_\mu[N_a(T)] > \dfrac{T}{H_1 \Delta_a^2}$ for all $a \in \{2, \ldots, k\}$

then $\sum_{a>1} E[N_a(T)] > \cancel{\frac{T}{H_1}} \left(\cancel{\sum \frac{1}{\Delta_a^2}}\right)$

$\Rightarrow$ <u>Contradiction</u>

$\rightarrow$ Pick such a, define $\mu^{[a]}$ as;

$$\mu_i^{[a]} = \begin{cases} \mu_i & i \neq a \\ \mu_a + 2\Delta_a & i = a \end{cases}$$



Easy to see : $H_1(\mu^{[a]}) < H_1(\mu)$

Since $\mu_a^{[a]}$ is now the new opt arm the sub opt gap increases for all arm (except μ_a & μ_1)

By B-H, Divergence decomposition,

$$\mathbb{P}_\mu(A) + \mathbb{P}_{\mu^{[a]}}(A^c) \geq \frac{1}{2} e^{-D(\mu_a, \mu_a^{[a]}) \cdot E_\mu[N_a(T)]}$$

$$\geq \frac{1}{2} e^{-\frac{(2\Delta_a)^2}{2} \cdot \frac{T}{H_1 \Delta_a^2}}$$

$$= \frac{1}{2} e^{-2T/H_1}$$

$A = \{a_T = a\} \rightarrow$ Bad for μ
and A^c bad for $\mu^{[a]}$

$e_T(\mu) \geq \mathbb{P}_\mu(A)$

$e_T(\mu^{[a]}) = \mathbb{P}_{\mu^{[a]}}(A^c)$

$\Rightarrow \quad e_T(\mu) + e_T(\mu^{[a]}) \geq \frac{1}{2} e^{-2T/H_1}$

$\Rightarrow \quad \max(e_T(\mu), e_T(\mu^{[a]})) \geq \frac{1}{4} e^{-2T/H_1}$

$\therefore \alpha(\text{any algo}) \leq \frac{2}{H_1} \leq \frac{4}{H_2}$ (Using Lemma $\frac{H_2}{2} \leq H_1$)

And $\alpha(SR) \geq \frac{1}{8(\overline{\log} K) H_2}$

Thus, SR is already near-perfect.

<u>Theorem</u>: Consider algo s.t. for any instance ν, for any $a \neq 1$

$$\mathbb{E}[N_a(T)] \leq C_a(\nu) \log(T)$$

For this algo, prob. of error cannot decay exponentially in T $\forall \nu$.

(Basically, for any algo, which gives log regret like UCB for all instances of ν, we cannot get exponential decay (for BAI) for all instances)

→ This is irrespective of how we declare the best arm (highest mean, highest UCB, etc)

Proof: Pick ν → opt arm 1
Pick $a \neq 1$,

Define $\nu^{[a]}$ s.t. only arm a is perturbed; a is opt. in $\nu^{[a]}$

By B-H, divergence decomp,

$$\mathbb{P}_\nu(A) + \mathbb{P}_{\nu^{[a]}}(A^c) \geq \frac{1}{2} e^{-D(\nu_a, \nu_a^{[a]}) E_\nu[N_a(T)]}$$

As before, $(A = \{a_T = a\})$

⋮

$$e_T(\nu) + e_T(\nu^{[a]}) \geq \frac{1}{2} e^{-D(\nu_a, \nu_a^{[a]}) E_\nu[N_a(T)]}$$

$$\geq \frac{1}{2} e^{-DC\log T}$$

$$= \frac{1}{2} T^{-DC}$$

Powerlaw decay slower than exponentials.

⇒ Both $e_T(\nu)$ and $e_T(\nu^{[a]})$ cannot decay exponentially in T.

FIXED CONFIDENCE BAI

Given: Error threshold $\delta \in (0,1)$ $\{$Prob. error $\leq \delta\}$

Algo keeps sampling until at stopping time τ, it stops, and reports output $\hat{a}_\tau$

Formally, we need, $\mathbb{P}(\tau < \infty, \hat{a}_\tau \neq 1) \leq \delta$ —— ①

prob. that we stop & doesn't give right answer less than δ

Algo, that satisfies ① are called Sound / δ-PC
(prob. correct)

→ With prob. $1-\delta$, we give right ans (after stopping) or we don't stop. (for sound algo)

→ If we prove that we stop with prob. 1 then we give right answer with prob. $1-\delta$.

Goal: Minimize $\mathbb{E}[\tau]$, or high prob. upper bound on τ.

<u>Theorem</u> : Consider 1-Gaussian instance μ (arm 1 opt).

For any δ-sound algo, $\boxed{E[\tau] \geq 2 \log\left(\frac{1}{\delta}\right) \sum_{1}^{k} \frac{1}{\Delta_i^2}}$

<u>Convention</u> : $\Delta_1 = \Delta_2$

or $\Delta_i = \begin{cases} \mu_1 - \mu_2 & i = 1 \\ \mu_1 - \mu_i & i \neq 1 \end{cases}$

* This is the first instance where we <u>need</u> that only one arm is optimal. If there are even 2 opt arms $E[\tau]$ blows to ∞. Earlier, we used only 1 optical arm only for cosmetic reasons (not needed). Here, the algo cannot differentiate whether there are 2 opt arms or the other arm is arbit. close to the first.