Best Arm Identification with Arm Erasures

P. N. Karthik
Assistant Professor,
Department of Artificial Intelligence

Applied Mathematics Day
IIT Hyderabad

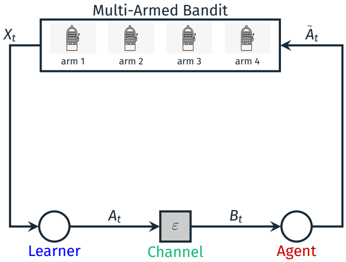
25 July 2024

Joint Work With

Kota Srinivas Reddy
ksreddy@ee.iitm.ac.in
sites.google.com/view/srinivas-reddy-kota

Vincent Y. F. Tan
vtan@nus.edu.sg
vyftan.github.io

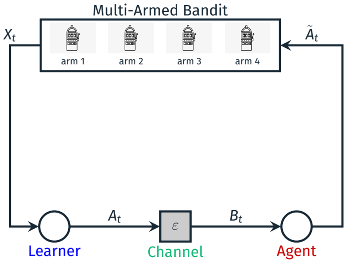
Preliminaries



Model Inspired from [Hanna et al., 2023]

- Two parties: Learner and Agent

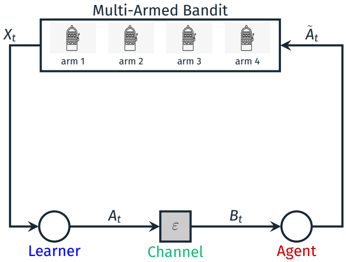
Preliminaries



Model Inspired from [Hanna et al., 2023]

- Two parties: Learner and Agent
- Channel with erasure probability $\varepsilon \in (0, 1)$

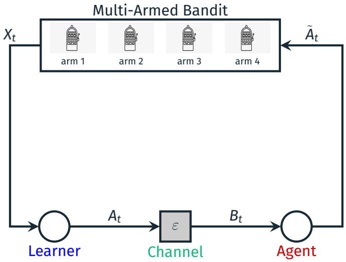
Preliminaries



Model Inspired from [Hanna et al., 2023]

- Two parties: Learner and Agent
- Channel with erasure probability $\varepsilon \in (0,1)$
- K-armed bandit with arm means $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_K]^\top$

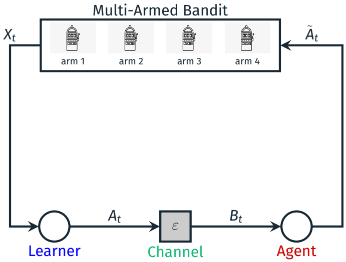
Preliminaries



Model Inspired from [Hanna et al., 2023]

- Two parties: Learner and Agent
- Channel with erasure probability $\varepsilon \in (0, 1)$
- K-armed bandit with arm means $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_K]^\top$
- Let $a^\star(\boldsymbol{\mu}) = \arg\max_a \mu_a$

Preliminaries

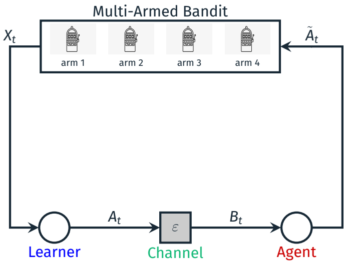


Model Inspired from [Hanna et al., 2023]

- Two parties: Learner and Agent
- Channel with erasure probability $\varepsilon \in (0, 1)$
- K-armed bandit with arm means $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_K]^\top$
- Let $a^\star(\boldsymbol{\mu}) = \arg\max_a \mu_a$
- Learner's goal:

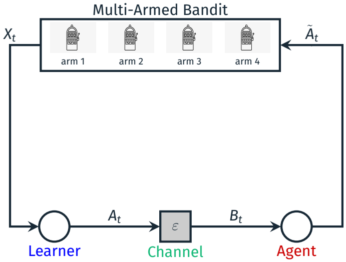
  Find $a^\star(\boldsymbol{\mu})$ quickly and accurately

PRELIMINARIES



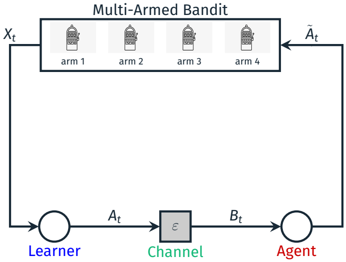
- A_t: learner's transmitted arm at time t

Preliminaries



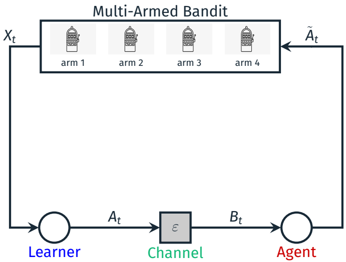
- A_t: learner's transmitted arm at time t
- $B_t \in \{A_t, \mathrm{null}\}$

Preliminaries



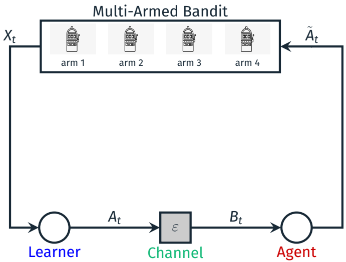
- A_t: learner's transmitted arm at time t
- $B_t \in \{A_t, \mathrm{null}\}$
  - $B_t = A_t$ if no erasure at time t

Preliminaries



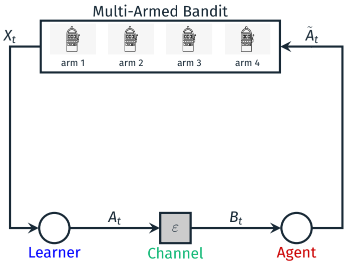
- A_t: learner's transmitted arm at time t
- $B_t \in \{A_t, \mathrm{null}\}$
  - $B_t = A_t$ if no erasure at time t
  - $B_t = \mathrm{null}$ if erasure at time t

Preliminaries



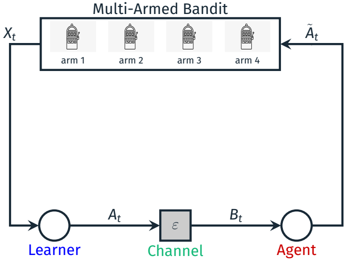
- A_t: learner's transmitted arm at time t
- $B_t \in \{A_t, \text{null}\}$
  - $B_t = A_t$ if no erasure at time t
  - $B_t = \text{null}$ if erasure at time t
  - $\mathbb{P}(B_t = \text{null}) = \varepsilon = 1 - \mathbb{P}(B_t = A_t)$

PRELIMINARIES



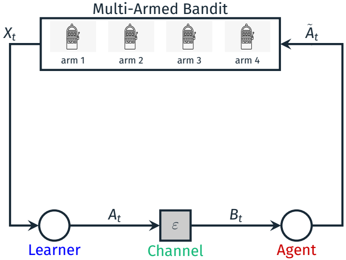
- A_t: learner's transmitted arm at time t
- $B_t \in \{A_t, \text{null}\}$
  - $B_t = A_t$ if no erasure at time t
  - $B_t = \text{null}$ if erasure at time t
  - $\mathbb{P}(B_t = \text{null}) = \varepsilon = 1 - \mathbb{P}(B_t = A_t)$
- No erasure feedback; learner cannot see B_t

Preliminaries



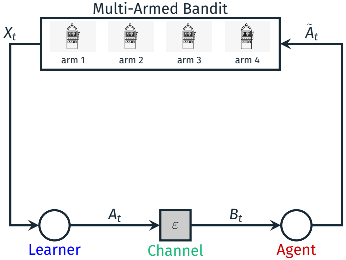
- A_t: learner's transmitted arm at time t
- $B_t \in \{A_t, \mathrm{null}\}$
  - $B_t = A_t$ if no erasure at time t
  - $B_t = \mathrm{null}$ if erasure at time t
  - $\mathbb{P}(B_t = \mathrm{null}) = \varepsilon = 1 - \mathbb{P}(B_t = A_t)$
- No erasure feedback; learner cannot see B_t
- $\tilde{A}_t$: arm pulled by agent at time t

Preliminaries



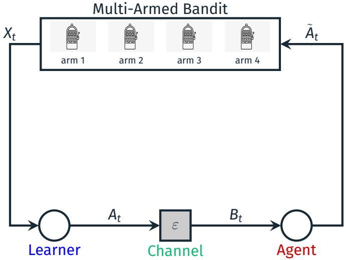
- A_t: learner's transmitted arm at time t
- $B_t \in \{A_t, \mathrm{null}\}$
  - $B_t = A_t$ if no erasure at time t
  - $B_t = \mathrm{null}$ if erasure at time t
  - $\mathbb{P}(B_t = \mathrm{null}) = \varepsilon = 1 - \mathbb{P}(B_t = A_t)$
- No erasure feedback; learner cannot see B_t
- $\tilde{A}_t$: arm pulled by agent at time t
- X_t: reward from arm $\tilde{A}_t$, revealed to learner

Preliminaries



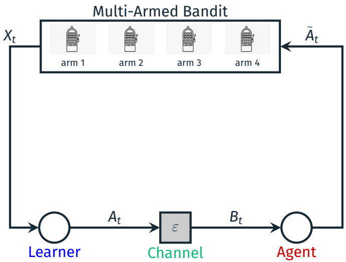
- A_t: learner's transmitted arm at time t
- $B_t \in \{A_t, \mathrm{null}\}$
  - $B_t = A_t$ if no erasure at time t
  - $B_t = \mathrm{null}$ if erasure at time t
  - $\mathbb{P}(B_t = \mathrm{null}) = \varepsilon = 1 - \mathbb{P}(B_t = A_t)$
- No erasure feedback; learner cannot see B_t
- $\tilde{A}_t$: arm pulled by agent at time t
- X_t: reward from arm $\tilde{A}_t$, revealed to learner
- ε known to learner

Preliminaries



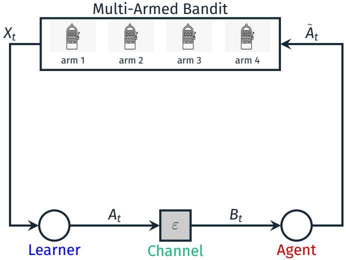
- A_t: learner's transmitted arm at time t
- $B_t \in \{A_t, \text{null}\}$
  - $B_t = A_t$ if no erasure at time t
  - $B_t = \text{null}$ if erasure at time t
  - $\mathbb{P}(B_t = \text{null}) = \varepsilon = 1 - \mathbb{P}(B_t = A_t)$
- No erasure feedback; learner cannot see B_t
- $\tilde{A}_t$: arm pulled by agent at time t
- X_t: reward from arm $\tilde{A}_t$, revealed to learner
- ε known to learner
- Agent's strategy known to learner. E.g.,

Preliminaries



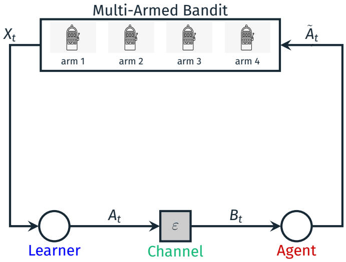
- A_t: learner's transmitted arm at time t
- $B_t \in \{A_t, \mathrm{null}\}$
  - $B_t = A_t$ if no erasure at time t
  - $B_t = \mathrm{null}$ if erasure at time t
  - $\mathbb{P}(B_t = \mathrm{null}) = \varepsilon = 1 - \mathbb{P}(B_t = A_t)$
- No erasure feedback; learner cannot see B_t
- $\tilde{A}_t$: arm pulled by agent at time t
- X_t: reward from arm $\tilde{A}_t$, revealed to learner
- ε known to learner
- Agent's strategy known to learner. E.g.,
  - $\tilde{A}_t \sim \mathrm{Unif}([K])$ under erasures

Preliminaries



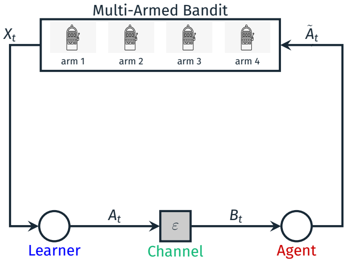
- A_t: learner's transmitted arm at time t
- $B_t \in \{A_t, \mathrm{null}\}$
  - $B_t = A_t$ if no erasure at time t
  - $B_t = \mathrm{null}$ if erasure at time t
  - $\mathbb{P}(B_t = \mathrm{null}) = \varepsilon = 1 - \mathbb{P}(B_t = A_t)$
- No erasure feedback; learner cannot see B_t
- $\tilde{A}_t$: arm pulled by agent at time t
- X_t: reward from arm $\tilde{A}_t$, revealed to learner
- ε known to learner
- Agent's strategy known to learner. E.g.,
  - $\tilde{A}_t \sim \mathrm{Unif}([K])$ under erasures
  - $\tilde{A}_t = \tilde{A}_{t-1}$ under erasures

Preliminaries



Find $a^\star(\boldsymbol{\mu})$ quickly and accurately

Preliminaries



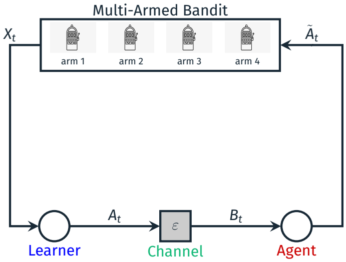
Find $a^\star(\boldsymbol{\mu})$ quickly and accurately

Algo	Stopping time	Output
π	τ	$\widehat{a}$

- Stop in finite time:

$$\mathbb{P}_{\boldsymbol{\mu}}^{\pi}(\tau < +\infty) = 1, \quad \forall \boldsymbol{\mu}.$$

Preliminaries



Find $a^\star(\boldsymbol{\mu})$ quickly and accurately

Algo	Stopping time	Output
π	τ	$\widehat{a}$

- Stop in finite time:

$$\mathbb{P}_{\boldsymbol{\mu}}^{\pi}(\tau < +\infty) = 1, \quad \forall \boldsymbol{\mu}.$$

- Given $\delta \in (0,1)$,

$$\mathbb{P}_{\boldsymbol{\mu}}^{\pi}(\widehat{a} = a^\star(\boldsymbol{\mu})) \geq 1 - \delta, \quad \forall \boldsymbol{\mu}.$$

Preliminaries

$$\Pi(\delta) = \left\{ \pi \text{ satisfying } (1), (2) \right\}$$

Growth rate of $\inf_{\pi \in \Pi(\delta)} \mathbb{E}_{\boldsymbol{\mu}}^{\pi}[\tau]$ = ?

Find $a^\star(\boldsymbol{\mu})$ quickly and accurately

Algo	Stopping time	Output
π	τ	$\widehat{a}$

- Stop in finite time:

$$\mathbb{P}_{\boldsymbol{\mu}}^{\pi}(\tau < +\infty) = 1, \quad \forall \boldsymbol{\mu}. \tag{1}$$

- Given $\delta \in (0, 1)$,

$$\mathbb{P}_{\boldsymbol{\mu}}^{\pi}(\widehat{a} = a^\star(\boldsymbol{\mu})) \geq 1 - \delta, \quad \forall \boldsymbol{\mu}. \tag{2}$$

Our Results at a Glance

<table>
<tr><td rowspan="2">Uniform</td><td colspan="2">Results</td></tr>
<tr><td>Converse: $\forall \pi \in \Pi(\delta)$,
$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \geq \log\left(\frac{1}{4\delta}\right) \cdot \left\{\sum_{a=1}^{K} \frac{1}{\Delta_{\varepsilon,a}^2/2}\right\}$$</td><td>Algorithm (SEUnif): w.p. $\geq 1-\delta$,
$$\tau_{\text{SEUnif}} \leq \sum_{a=1}^{K} \left[1 + \frac{102}{\Delta_{\varepsilon,a}^2} \log\left(\frac{64\sqrt{\frac{8K}{\delta}}}{\Delta_{\varepsilon,a}^2}\right)\right]$$</td></tr>
<tr><td rowspan="2">Previous</td><td colspan="2">Results</td></tr>
<tr><td colspan="2">Algorithm (MSEA): w.p. $\geq 1-\delta$,
$$\tau_{\text{MSEA}} \leq \frac{1}{1-\alpha} \sum_{a=1}^{K} \left[1 + \frac{102}{\Delta_a^2} \log\left(\frac{64\sqrt{\frac{8K}{\delta}}}{\Delta_a^2}\right)\right] + O\left(\sqrt{\log\left(\frac{1}{\delta}\right)}\right)$$</td></tr>
</table>

Uniform Sampling By Agent Under Erasures

Shifting and Scaling of Arm Means

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\mu}}[X_t|A_t = a] \\
&= (1-\varepsilon)\,\mathbb{E}_{\boldsymbol{\mu}}[X_t|A_t = a,\ B_t = A_t] \\
&\qquad\qquad + \varepsilon\,\mathbb{E}_{\boldsymbol{\mu}}[X_t|A_t = a,\ B_t = \mathrm{null}] \\
&= (1-\varepsilon)\,\mu_a + \frac{\varepsilon}{K}\sum_{a'=1}^{K}\mu_{a'} \\
&= \underbrace{(1-\varepsilon)\,\mu_a + \varepsilon\,\bar{\mu}}_{\mu_a^{\varepsilon}}
\end{aligned}$$

$$\Delta_a = \mu_1 - \mu_a$$
$$\Delta_1 = \mu_1 - \mu_2$$

$$\Delta_{\varepsilon,a} = \mu_1^{\varepsilon} - \mu_a^{\varepsilon}$$
$$\Delta_{\varepsilon,1} = \mu_1^{\varepsilon} - \mu_2^{\varepsilon}$$

$$\Delta_{\varepsilon,a} = (1-\varepsilon)\,\Delta_a$$

Lower Bound

Transportation Lemma [Kaufmann et al., 2016, Lemma 19]

For all $\delta \in (0,1)$, $\pi \in \Pi(\delta)$, and $\boldsymbol{\lambda}$ such that $a^\star(\boldsymbol{\lambda}) \neq a^\star(\boldsymbol{\mu})$,

$$\sum_{a=1}^{K} \mathbb{E}_{\boldsymbol{\mu}}^{\pi}\left[N_a(\tau)\right] \frac{(\mu_a^\varepsilon - \lambda_a^\varepsilon)^2}{2} \geq \log\left(\frac{1}{4\delta}\right), \qquad N_a(\tau) = \sum_{t=0}^{\tau} \mathbf{1}\{A_t = a\}$$

Lower Bound

Transportation Lemma [Kaufmann et al., 2016, Lemma 19]

For all $\delta \in (0,1)$, $\pi \in \Pi(\delta)$, and $\boldsymbol{\lambda}$ such that $a^\star(\boldsymbol{\lambda}) \neq a^\star(\boldsymbol{\mu})$,

$$\sum_{a=1}^{K} \mathbb{E}_{\boldsymbol{\mu}}^{\pi}\left[N_a(\tau)\right] \frac{(\mu_a^\varepsilon - \lambda_a^\varepsilon)^2}{2} \geq \log\left(\frac{1}{4\delta}\right), \qquad N_a(\tau) = \sum_{t=0}^{\tau} \mathbf{1}\{A_t = a\}$$

$\forall a \neq 1:$

$$\boldsymbol{\mu} = [\mu_1,\ \mu_2,\ \mu_a,\ \ldots, \mu_K]^\top$$

$$\boldsymbol{\lambda} = [\mu_1,\ \mu_2,\ \underbrace{\mu_1 + \gamma}_{\text{arm } a},\ \ldots, \mu_K]^\top$$

$$\mathbb{E}_{\boldsymbol{\mu}}^{\pi}[N_a(\tau)] \geq \frac{\log \frac{1}{4\,\delta}}{(\Delta_{\varepsilon,a} - \gamma)^2/2}$$

$a = 1:$

$$\boldsymbol{\mu} = [\mu_1,\ \mu_2,\ \mu_a,\ \ldots, \mu_K]^\top$$

$$\boldsymbol{\lambda} = [\underbrace{\mu_2 - \gamma}_{\text{arm } 1},\ \mu_2, \mu_a,\ \ldots, \mu_K]^\top$$

$$\mathbb{E}_{\boldsymbol{\mu}}^{\pi}[N_1(\tau)] \geq \frac{\log \frac{1}{4\,\delta}}{(\Delta_{\varepsilon,1} + \gamma)^2/2}$$

Converse: Lower Bound

Proposition: Lower Bound

For all $\delta \in (0,1)$ and $\pi \in \Pi(\delta)$,

$$\mathbb{E}^{\pi}_{\boldsymbol{\mu}}[\tau] \geq \log\left(\frac{1}{4\delta}\right) \cdot \left\{\sum_{a=1}^{K} \frac{1}{\Delta^2_{\varepsilon,a}/2}\right\}$$

$\forall a \neq 1:$

$$\boldsymbol{\mu} = [\mu_1,\ \mu_2,\ \mu_a,\ \ldots, \mu_K]^\top$$

$$\boldsymbol{\lambda} = [\mu_1,\ \mu_2,\ \underbrace{\mu_1 + \gamma,}_{\text{arm } a}\ \ldots, \mu_K]^\top$$

$$\mathbb{E}^{\pi}_{\boldsymbol{\mu}}[N_a(\tau)] \geq \frac{\log \frac{1}{4\,\delta}}{(\Delta_{\varepsilon,a} - \gamma)^2/2}$$

$a = 1:$

$$\boldsymbol{\mu} = [\mu_1,\ \mu_2,\ \mu_a,\ \ldots, \mu_K]^\top$$

$$\boldsymbol{\lambda} = [\underbrace{\mu_2 - \gamma,}_{\text{arm } 1}\ \mu_2, \mu_a,\ \ldots, \mu_K]^\top$$

$$\mathbb{E}^{\pi}_{\boldsymbol{\mu}}[N_1(\tau)] \geq \frac{\log \frac{1}{4\,\delta}}{(\Delta_{\varepsilon,1} + \gamma)^2/2}$$

Achievability: Successive Elimination (SEUnif)

Algorithm 1 SEUnif

Input: $K \in \mathbb{N}, \delta \in (0,1)$
Output: $\widehat{a}_{\text{SEUnif}} \in [K]$ (best arm)

Initialization: $n = 0, S = [K], t_a(n) = 0$ for all a

```
while |S| > 1 do
    n ← n + 1
    Pull each arm a ∈ S once.
    Set t_a(n) ← t_a(n − 1) + 1. Update μ̂_a(n), UCB_a(n)
    and LCB_a(n) for all a ∈ S.
    if ∃a′ ∈ S such that UCB_a(n) < LCB_{a′}(n) then
        S ← S \ {a}
    end if
    if |S| = 1 then
        â_SEUnif ← a ∈ S
    end if
end while
return â_SEUnif.
```

$$\mathrm{UCB}_a(n) := \widehat{\mu}_a(n) + \alpha_\delta(t_a(n)),$$
$$\mathrm{LCB}_a(n) := \widehat{\mu}_a(n) - \alpha_\delta(t_a(n)),$$
$$\alpha_\delta(x) := \sqrt{\frac{2\log(8Kx^2/\delta)}{x}}$$

Theorem

- SEUnif $\in \Pi(\delta)$
- With probability $\geq 1-\delta$,

$$\tau_{\text{SEUnif}} \leq \sum_{a=1}^{K}\left[1 + \frac{102}{\Delta_{\varepsilon,a}^2}\log\left(\frac{64\sqrt{\frac{8K}{\delta}}}{\Delta_{\varepsilon,a}^2}\right)\right]$$

Previous Arm Sampling By Agent Under Erasures

$$
\begin{aligned}
\mathbb{E}_{\boldsymbol{\mu}}[X_t | A_{0:t}, X_{0:t-1}] &= \mathbb{E}_{\boldsymbol{\mu}}[X_t \, \mathbf{1}\{B_t = A_t\} | A_{0:t}, X_{0:t-1}] + \mathbb{E}_{\boldsymbol{\mu}}[X_t \, \mathbf{1}\{B_t = \text{null}\} | A_{0:t}, X_{0:t-1}] \\
&= (1-\varepsilon)\, \mathbb{E}_{\boldsymbol{\mu}}[X_t | \tilde{A}_t = A_t] + \varepsilon \, \mathbb{E}_{\boldsymbol{\mu}}[X_t | \tilde{A}_{t-1}] \\
&= (1-\varepsilon)\, \mathbb{E}_{\boldsymbol{\mu}}[X_t | \tilde{A}_t = A_t] + \varepsilon \left[(1-\varepsilon)\, \mathbb{E}_{\boldsymbol{\mu}}[X_t | \tilde{A}_{t-1} = A_{t-1}] + \varepsilon \, \mathbb{E}_{\boldsymbol{\mu}}[X_t | \tilde{A}_{t-2}] \right] \\
&\;\vdots \\
&= (1-\varepsilon) \sum_{s=0}^{t} \varepsilon^s \, \mathbb{E}_{\boldsymbol{\mu}}[X_t | \tilde{A}_{t-s} = A_{t-s}] + \frac{\varepsilon^{t+1}}{K} \sum_{a'=1}^{K} \mathbb{E}_{\boldsymbol{\mu}}[X_t | \tilde{A}_0 = a'] \\
&\overset{(a)}{=} (1-\varepsilon) \sum_{s=0}^{t} \varepsilon^s \, \mu_{A_{t-s}} + \frac{\varepsilon^{t+1}}{K} \sum_{a'=1}^{K} \mu_{a'} \\
&= (1-\varepsilon) \sum_{u=0}^{t} \varepsilon^{t-u} \, \mu_{A_u} + \frac{\varepsilon^{t+1}}{K} \sum_{a'=1}^{K} \mu_{a'} \\
&= \mu_{A_{0:t}, \varepsilon}
\end{aligned}
$$

Achievability: Modified Successive Elimination Algorithm (MSEA)

Algorithm 2 MSEA

Input: $K \in \mathbb{N}, \delta \in (0,1), \alpha \in (0,1)$
Output: $\widehat{a}_{\text{MSEA}} \in [K]$ (best arm)

Initialization: $n = 0, t = 0, S = [K], t_a(n) = 0$ for all a

```
while |S| > 1 do
    n ← n + 1
    Pull each active arm a ∈ S for a total of t_n times, and ignore
    the first ⌊α t_n⌋ pulls and the associated rewards.
    For all a ∈ [K], set t_a(n) ← t_a(n − 1) + ⌈(1 − α) t_n⌉.
    Also, update μ̂_a(n), UCB_a(n) and LCB_a(n) based on
    the last ⌈(1 − α) t_n⌉ rewards seen from arm a.
    if ∃a′ ∈ S such that UCB_a(n) < LCB_a′(n) then
        S ← S \ {a}.
    end if
    if |S| = 1 then
        â_MSEA ← a ∈ S
    end if
end while
return â_MSEA.
```

$$T := \max\left\{ \left\lceil \frac{\log\left(\frac{2K}{\delta}+1\right)}{\alpha \log(1/\varepsilon)} \right\rceil, \ 1 \right\}, \quad t_n := nT$$

Neglect first α fraction of arm transmissions and associated rewards

Theorem

- MSEA $\in \Pi(\delta)$
- With probability $\geq 1-\delta$,

$$\tau_{\text{MSEA}} \leq \sum_{a=1}^{K} \beta_a,$$

where $\beta_a = \beta_a := \frac{T'_a}{1-\alpha} + \sqrt{\frac{2\,T\,T'_a}{1-\alpha}} + T$, and

$$T'_a := 1 + \frac{102}{\Delta_a^2} \log\left(\frac{64\sqrt{\frac{8K}{\delta}}}{\Delta_a^2} \right)$$

Which Scheme for Faster Identification?

Uniform Sampling under Erasures

- SEUnif $\in \Pi(\delta)$
- With probability $\geq 1-\delta$,

$$\tau_{\text{SEUnif}} \leq \sum_{a=1}^{K} \left[1 + \frac{102}{\Delta_{\varepsilon,a}^2} \log \left(\frac{64\sqrt{\frac{8K}{\delta}}}{\Delta_{\varepsilon,a}^2} \right) \right]$$

Previous Arm Sampling under Erasures

- MSEA $\in \Pi(\delta)$
- With probability $\geq 1-\delta$,

$$\tau_{\text{MSEA}} \leq \sum_{a=1}^{K} \beta_a,$$

where $\beta_a = \beta_a := \frac{T'_a}{1-\alpha} + \sqrt{\frac{2\,T\,T'_a}{1-\alpha}} + T$, and

$$T'_a := 1 + \frac{102}{\Delta_a^2} \log \left(\frac{64\sqrt{\frac{8K}{\delta}}}{\Delta_a^2} \right)$$

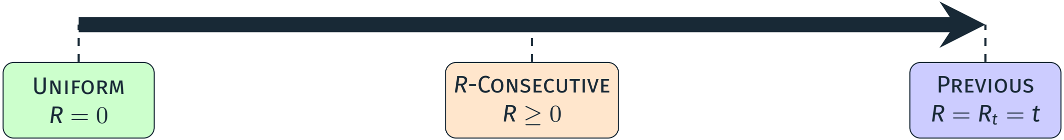
Handling R-Consecutive Erasures

R or less Consecutive Erasures – Agent Pulls Previous Arm

$(R+1)^{\text{st}}$ Erasure – Agent Pulls Arm Uniformly

UNIFORM	$\mathbb{E}_{\boldsymbol{\mu}}^{\pi}[X_t \| A_{0:t}, X_{0:t-1}] = (1-\varepsilon)\,\mu_{A_t} + \varepsilon\,\bar{\mu}$
PREVIOUS	$\mathbb{E}_{\boldsymbol{\mu}}^{\pi}[X_t \| A_{0:t}, X_{0:t-1}] = (1-\varepsilon)\left[\mu_{A_t} + \varepsilon\,\mu_{A_{t-1}} + \ldots + \varepsilon^{t}\,\mu_{A_0}\right] + \varepsilon^{t+1}\,\bar{\mu}$
R-CONSECUTIVE	$\mathbb{E}_{\boldsymbol{\mu}}^{\pi}[X_t \| A_{0:t}, X_{0:t-1}] = (1-\varepsilon)\left[\mu_{A_t} + \varepsilon\,\mu_{A_{t-1}} + \ldots + \varepsilon^{R}\,\mu_{A_{t-R}}\right] + \varepsilon^{R+1}\,\bar{\mu}$

Uniform	$\mathbb{E}^{\pi}_{\boldsymbol{\mu}}[X_t \mid A_{0:t}, X_{0:t-1}] = (1-\varepsilon)\,\mu_{A_t} + \varepsilon\,\bar{\mu}$
Previous	$\mathbb{E}^{\pi}_{\boldsymbol{\mu}}[X_t \mid A_{0:t}, X_{0:t-1}] = (1-\varepsilon)\left[\mu_{A_t} + \varepsilon\,\mu_{A_{t-1}} + \ldots + \varepsilon^{t}\,\mu_{A_0}\right] + \varepsilon^{t+1}\,\bar{\mu}$
R-Consecutive	$\mathbb{E}^{\pi}_{\boldsymbol{\mu}}[X_t \mid A_{0:t}, X_{0:t-1}] = (1-\varepsilon)\left[\mu_{A_t} + \varepsilon\,\mu_{A_{t-1}} + \ldots + \varepsilon^{R}\,\mu_{A_{t-R}}\right] + \varepsilon^{R+1}\,\bar{\mu}$



Lower Bound

Uniform

For all $\delta \in (0,1)$, $\pi \in \Pi(\delta)$, and $\boldsymbol{\lambda}$ such that $a^\star(\boldsymbol{\lambda}) \neq a^\star(\boldsymbol{\mu})$,

$$\sum_{a=1}^{K} \mathbb{E}_{\boldsymbol{\mu}}^{\pi}[N_a(\tau)] \, \frac{(\mu_a^\varepsilon - \lambda_a^\varepsilon)^2}{2} \geq \log\left(\frac{1}{4\delta}\right), \qquad N_a(\tau) = \sum_{t=0}^{\tau} \mathbf{1}\{A_t = a\}$$

$\forall a \neq 1:$

$$\boldsymbol{\mu} = [\mu_1,\ \mu_2,\ \mu_a,\ \ldots, \mu_K]^\top$$

$$\boldsymbol{\lambda} = [\mu_1,\ \mu_2,\ \underbrace{\mu_1 + \gamma}_{\text{arm } a},\ \ldots, \mu_K]^\top$$

$$\mathbb{E}_{\boldsymbol{\mu}}^{\pi}[N_a(\tau)] \geq \frac{\log \frac{1}{4\,\delta}}{(\Delta_{\varepsilon,a} - \gamma)^2/2}$$

$a = 1:$

$$\boldsymbol{\mu} = [\mu_1,\ \mu_2,\ \mu_a,\ \ldots, \mu_K]^\top$$

$$\boldsymbol{\lambda} = [\underbrace{\mu_2 - \gamma}_{\text{arm } 1},\ \mu_2, \mu_a,\ \ldots, \mu_K]^\top$$

$$\mathbb{E}_{\boldsymbol{\mu}}^{\pi}[N_1(\tau)] \geq \frac{\log \frac{1}{4\,\delta}}{(\Delta_{\varepsilon,1} + \gamma)^2/2}$$

A Tighter Lower Bound

Uniform

For all $\delta \in (0,1)$ and $\pi \in \Pi(\delta)$,

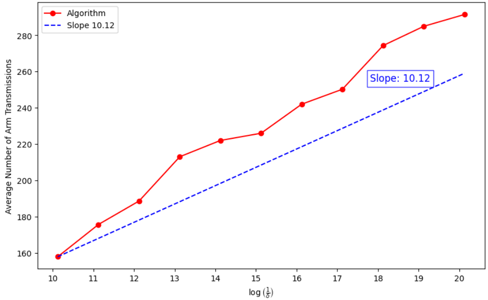
$$\inf_{\boldsymbol{\lambda}:\, a^\star(\boldsymbol{\lambda}) \neq a^\star(\boldsymbol{\mu})} \sum_{a=1}^{K} \mathbb{E}_{\boldsymbol{\mu}}^{\pi}\left[N_a(\tau)\right] \frac{(\mu_a^\varepsilon - \lambda_a^\varepsilon)^2}{2} \geq \log\left(\frac{1}{4\delta}\right)$$

$$(\mathbb{E}_{\boldsymbol{\mu}}^{\pi}[\tau] + 1) \times \inf_{\boldsymbol{\lambda}:\, a^\star(\boldsymbol{\lambda}) \neq a^\star(\boldsymbol{\mu})} \sum_{a=1}^{K} \frac{\mathbb{E}_{\boldsymbol{\mu}}^{\pi}\left[N_a(\tau)\right]}{\mathbb{E}_{\boldsymbol{\mu}}^{\pi}[\tau] + 1} \frac{(\mu_a^\varepsilon - \lambda_a^\varepsilon)^2}{2} \geq \log\left(\frac{1}{4\delta}\right)$$

$$(\mathbb{E}_{\boldsymbol{\mu}}^{\pi}[\tau] + 1) \times \underbrace{\left\{ \sup_{\mathrm{w} \in \Sigma_K} \inf_{\boldsymbol{\lambda}:\, a^\star(\boldsymbol{\lambda}) \neq a^\star(\boldsymbol{\mu})} \sum_{a=1}^{K} w_a \frac{(\mu_a^\varepsilon - \lambda_a^\varepsilon)^2}{2} \right\}}_{T^\star_{\varepsilon,\mathrm{unif}}(\boldsymbol{\mu})} \geq \log\left(\frac{1}{4\delta}\right)$$

Complexity under Uniform Scheme

$$\inf_{\pi \in \Pi(\delta)} \frac{\mathbb{E}_{\boldsymbol{\mu}}^{\pi}[\tau]}{\log\left(\frac{1}{\delta}\right)} \gtrapprox \frac{1}{T^{\star}_{\varepsilon,\mathrm{unif}}(\boldsymbol{\mu})}$$



Complexity under Uniform Scheme

$$\inf_{\pi \in \Pi(\delta)} \frac{\mathbb{E}_{\boldsymbol{\mu}}^{\pi}[\tau]}{\log\left(\frac{1}{\delta}\right)} \approx \frac{1}{T^{\star}_{\varepsilon,\mathrm{unif}}(\boldsymbol{\mu})}, \qquad T^{\star}_{\varepsilon,\mathrm{unif}}(\boldsymbol{\mu}) = \sup_{\mathbf{w} \in \Sigma_K} \inf_{\boldsymbol{\lambda}:\, a^{\star}(\boldsymbol{\lambda}) \neq a^{\star}(\boldsymbol{\mu})} \sum_{a=1}^{K} w_a \frac{(\mu_a^{\varepsilon} - \lambda_a^{\varepsilon})^2}{2}$$

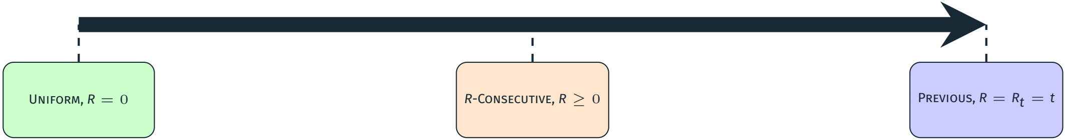
Complexity under Uniform Scheme

$$\inf_{\pi \in \Pi(\delta)} \frac{\mathbb{E}_{\boldsymbol{\mu}}^{\pi}[\tau]}{\log\left(\frac{1}{\delta}\right)} \approx \frac{1}{T^{\star}_{\varepsilon,\mathrm{unif}}(\boldsymbol{\mu})}, \qquad T^{\star}_{\varepsilon,\mathrm{unif}}(\boldsymbol{\mu}) = \sup_{\mathrm{w} \in \Sigma_K} \inf_{\boldsymbol{\lambda}:\, a^{\star}(\boldsymbol{\lambda}) \neq a^{\star}(\boldsymbol{\mu})} \sum_{a=1}^{K} w_a \, \frac{(\mu_a^{\varepsilon} - \lambda_a^{\varepsilon})^2}{2}$$

Complexity when Handling R-Consecutive Erasures

$$\inf_{\pi \in \Pi(\delta)} \frac{\mathbb{E}_{\boldsymbol{\mu}}^{\pi}[\tau]}{\log\left(\frac{1}{\delta}\right)} \approx \frac{1}{T^{\star}_{\varepsilon,R}(\boldsymbol{\mu})}, \quad T^{\star}_{\varepsilon,R}(\boldsymbol{\mu}) := \sup_{\mathrm{w} \in \Sigma([K]^{R+1})} \inf_{\boldsymbol{\lambda}:\, a^{\star}(\boldsymbol{\lambda}) \neq a^{\star}(\boldsymbol{\mu})} \sum_{\mathbf{a} \in [K]^R} \sum_{b=1}^{K} w(\mathbf{a}, b) \, \frac{(\mu_{\mathbf{a},b,\varepsilon} - \lambda_{\mathbf{a},b,\varepsilon})^2}{2}$$

$$\mu_{\mathbf{a},b,\varepsilon} := (1-\varepsilon)\left[\varepsilon^{R}\,\mu_{a_1} + \varepsilon^{R-1}\,\mu_{a_2} + \cdots + \varepsilon\,\mu_{a_R} + \mu_b\right] + \frac{\varepsilon^{R+1}}{K} \sum_{a'=1}^{K} \mu_{a'}$$



References

Hanna, O. A., Karakas, M., Yang, L. F., and Fragouli, C. (2023).
Multi-arm bandits over action erasure channels.
In *2023 IEEE International Symposium on Information Theory (ISIT)*, pages 1312–1317. IEEE.

Kaufmann, E., Cappé, O., and Garivier, A. (2016).
On the complexity of best-arm identification in multi-armed bandit models.
The Journal of Machine Learning Research, 17(1):1–42.