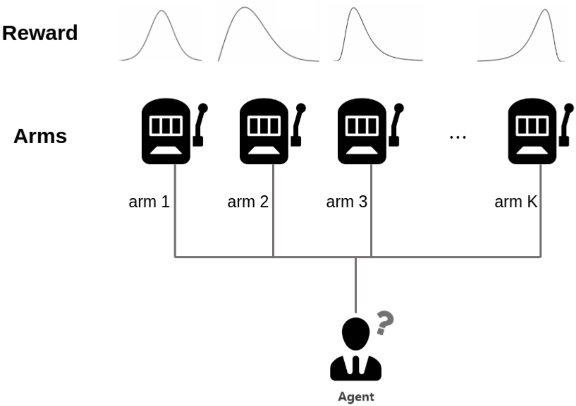
On the Complexity of Differentially Private Best-Arm Identification with Fixed Confidence

Achraf Azize, **Marc Jourdan**, **Aymen Al Marjani** and **Debabrota Basu**

Univ. Lille, CNRS, Inria, Centrale Lille, UMR 9189 CRIStAL, F-59000 Lille, France

Best Arm identification



Objective: Identify the arm with the highest mean $a^\star \triangleq \operatorname{argmax}_{a \in [K]} \mu_a$

Privacy Concern: Rewards may reveal sensitive information about individuals

Clinical trials

Objective: Find the most effective medicine between K candidates

Clinical trials

Objective: Find the most effective medicine between K candidates

Reward: $r = 1$ if the patient is cured, $r = 0$ if not cured

Clinical trials

Objective: Find the most effective medicine between K candidates

Reward: $r = 1$ if the patient is cured, $r = 0$ if not cured

For each round $t = 1, \ldots,$

- A new patient p_t arrives
- The agent chooses a medicine $a_t \in [K]$ based on the history $\mathcal{H}_{t-1} \triangleq \{a_1, r_1, \ldots, a_{t-1}, r_{t-1}\}$
- The agent observes the reaction r_t of patient p_t to medicine a_t
- If the agent decides to stop:
  - The agent proposes a guess $\hat{a}$ of $a^\star$
  - Stop

Clinical trials

Objective: Find the most effective medicine between K candidates

Reward: $r = 1$ if the patient is cured, $r = 0$ if not cured

For each round $t = 1, \ldots,$

- A new patient p_t arrives
- The agent chooses a medicine $a_t \in [K]$ based on the history $\mathcal{H}_{t-1} \triangleq \{a_1, r_1, \ldots, a_{t-1}, r_{t-1}\}$
- The agent observes the reaction r_t of patient p_t to medicine a_t
- If the agent decides to stop:
  - The agent proposes a guess $\hat{a}$ of $a^\star$
  - Stop

Privacy: A patient's reaction to a medicine can reveal sensitive information about their health conditions

Differential Privacy Background

Intuition: Indistinguishability from the mass

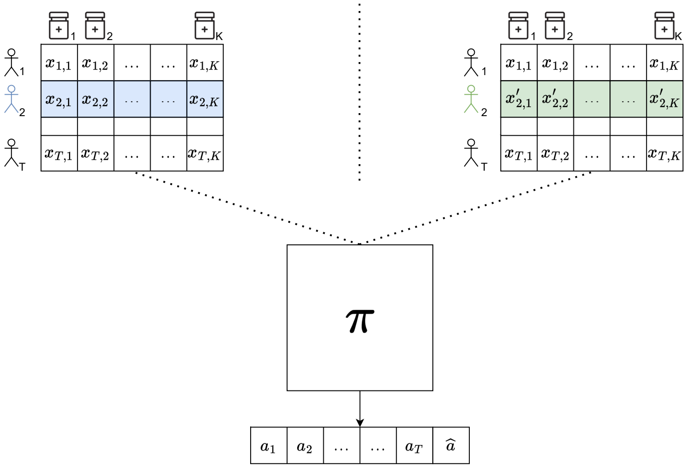
Definition: [Dwork and Roth, 2014] A randomised algorithm $\mathcal{A}$ satisfies ϵ-DP if for any two neighbouring datasets d and d' that differ only in one row, i.e $d \sim d'$, and for all sets of output $\mathcal{O} \subseteq \mathrm{Range}(\mathcal{A})$,

$$\Pr[\mathcal{A}(d) \in \mathcal{O}] \leq e^{\epsilon} \Pr\left[\mathcal{A}\left(d'\right) \in \mathcal{O}\right]$$

ϵ-global DP BAI

Definition: π satisfies ϵ-**global DP**, if $\forall T \geq 1$, $\forall \underline{\mathbf{d}}^T \sim \underline{\mathbf{d}'}^T, \forall \underline{a}^T$ and $\widehat{a}$,

$$\pi(\underline{a}^T, \widehat{a}, T \mid \underline{\mathbf{d}}^T) \leq e^{\epsilon} \pi(\underline{a}^T, \widehat{a}, T \mid \underline{\mathbf{d}'}^T).$$



Main Question and Contributions

Main Question: What is the cost of ϵ-global DP in BAI?

Contributions:

- We provide a lower bound on the sample complexity of any δ-correct ϵ-global DP BAI strategy
- We design a near-optimal algorithm matching the sample complexity lower bound, up to multiplicative constants

Lower Bound

Theorem: For any δ-correct ϵ-global DP BAI strategy, we have that

$$\mathbb{E}_{\boldsymbol{\nu}}[\tau] \geq \max\left(T^\star_{\mathrm{KL}}(\boldsymbol{\nu}), \frac{1}{6\epsilon} T^\star_{\mathrm{TV}}(\boldsymbol{\nu})\right) \log(1/3\delta),$$

where $\left(T^\star_{\mathbf{d}}(\boldsymbol{\nu})\right)^{-1} \triangleq \sup_{\omega \in \Sigma_K} \inf_{\boldsymbol{\lambda} \in \mathrm{Alt}(\boldsymbol{\nu})} \sum_{a=1}^{K} \omega_a \mathbf{d}(\nu_a, \lambda_a)$,
and $\mathbf{d}$ is either KL or TV.

Lower Bound

Theorem: For any δ-correct ϵ-global DP BAI strategy, we have that

$$\mathbb{E}_{\boldsymbol{\nu}}[\tau] \geq \max\left(T^\star_{\mathrm{KL}}(\boldsymbol{\nu}), \frac{1}{6\epsilon} T^\star_{\mathrm{TV}}(\boldsymbol{\nu}) \right) \log(1/3\delta),$$

where $\left(T^\star_{\mathbf{d}}(\boldsymbol{\nu})\right)^{-1} \triangleq \sup_{\omega \in \Sigma_K} \inf_{\boldsymbol{\lambda} \in \mathrm{Alt}(\boldsymbol{\nu})} \sum_{a=1}^{K} \omega_a \mathbf{d}(\nu_a, \lambda_a)$,
and $\mathbf{d}$ is either KL or TV.

Properties: Let $\Delta_a \triangleq \mu_1 - \mu_a$ be the gap between the means.
$T^\star_{\mathrm{KL}}(\boldsymbol{\nu}) \approx \sum_a \frac{1}{\Delta_a^2}$ [Garivier and Kaufmann, 2016] and $T^\star_{\mathrm{TV}}(\boldsymbol{\nu}) \approx \sum_a \frac{1}{\Delta_a}$.

Pinsker: $T^\star_{\mathrm{TV}}(\boldsymbol{\nu}) \geq \sqrt{2 T^\star_{\mathrm{KL}}(\boldsymbol{\nu})}$.

Lower Bound

Discussion

$$\mathbb{E}_{\boldsymbol{\nu}}[\tau] \geq \max\left(T^{\star}_{\mathrm{KL}}(\boldsymbol{\nu}), \frac{1}{6\epsilon} T^{\star}_{\mathrm{TV}}(\boldsymbol{\nu}) \right) \log(1/3\delta)$$

Two hardness regimes depending on ϵ and the environment $\boldsymbol{\nu}$:

- *Low-privacy regime*: When $\epsilon > \frac{T^{\star}_{\mathrm{TV}}(\boldsymbol{\nu})}{6 T^{\star}_{\mathrm{KL}}(\boldsymbol{\nu})}$, the lower bound retrieves the non-private $T^{\star}_{\mathrm{KL}}(\boldsymbol{\nu})$ lower bound and **privacy can be achieved for free**.

- *High-privacy regime:* When $\epsilon < \frac{T^{\star}_{\mathrm{TV}}(\boldsymbol{\nu})}{6 T^{\star}_{\mathrm{KL}}(\boldsymbol{\nu})}$, the lower bound becomes $\frac{1}{6\epsilon} T^{\star}_{\mathrm{TV}}(\boldsymbol{\nu})$ and ϵ-global DP δ-BAI requires more samples than non-private ones.

Algorithm Design

Top Two Algorithm

Algorithm Design

Top Two Algorithm

Intuition: The Top Two sampling rule consist of:

- Choosing a **leader** $B_n \in [K]$
- Choosing a **challenger** $C_n \in [K] \setminus \{B_n\}$
- Sampling B_n with probability β, else sampling C_n

Algorithm Design

Top Two Algorithm

Intuition: The Top Two sampling rule consist of:

- Choosing a **leader** $B_n \in [K]$
- Choosing a **challenger** $C_n \in [K] \setminus \{B_n\}$
- Sampling B_n with probability β, else sampling C_n

The recommendation rule: recommend the empirical best-arm

$$\widehat{a}_n = \operatorname*{argmax}_{a \in [K]} \widehat{\mu}_{n,a}$$

Algorithm Design

Top Two Algorithm

Intuition: The Top Two sampling rule consist of:

- Choosing a **leader** $B_n \in [K]$
- Choosing a **challenger** $C_n \in [K] \setminus \{B_n\}$
- Sampling B_n with probability β, else sampling C_n

The recommendation rule: recommend the empirical best-arm

$$\widehat{a}_n = \operatorname*{argmax}_{a \in [K]} \widehat{\mu}_{n,a}$$

The stopping rule is based on a calibrated GLR

$$\tau_\delta = \inf\{n \mid \min_{j \neq \widehat{a}_n} W_n(\widehat{a}_n, j) > c(n, \delta)\} ,$$

where $c(n, \delta)$ is a calibrated threshold and $W_n(i, j)$ is the empirical transportation cost between arms (i, j).

Algorithm Design

Private Top Two

To make the Top Two algorithm satisfy ϵ-global DP, we

Algorithm Design

Private Top Two

To make the Top Two algorithm satisfy ϵ-global DP, we

- Estimate the sequence of empirical means $(\widehat{\mu}_{a,n})$ privately, i.e. $(\tilde{\mu}_{a,n}) = (\widehat{\mu}_{a,n}) + \frac{1}{\epsilon} Lap$, using
  - Per-arm doubling
  - Forgetting
  - Adding calibrated Laplace noise

Algorithm Design

Private Top Two

To make the Top Two algorithm satisfy ϵ-global DP, we

- Estimate the sequence of empirical means $(\widehat{\mu}_{a,n})$ privately, i.e. $(\tilde{\mu}_{a,n}) = (\widehat{\mu}_{a,n}) + \frac{1}{\epsilon} Lap$, using
  - Per-arm doubling
  - Forgetting
  - Adding calibrated Laplace noise

- Count for the noise in:
  - The sampling rule: leader and challenger based on the private $(\tilde{\mu}_{a,n})$
  - The recommendation rule: Recommend $\widehat{a}_n = \operatorname{argmax}_{a \in [K]} \tilde{\mu}_{n,a}$
  - The stopping rule: re-calibrate the GLR threshold $\tilde{c}(n, \delta) = c(n, \delta) + \frac{1}{\epsilon} c_2(n, \delta)$

Algorithm Design

Privacy and sample complexity

Theorem: For Bernoulli instances verifying that $\exists C \geq 1$ such that $\Delta_{\text{max}}/\Delta_{\text{min}} \leq C$ and $\beta = 1/2$, AdaP-TT is ϵ-global DP, δ-correct and satisfies

$$\limsup_{\delta \to 0} \frac{\mathbb{E}_{\mu}[\tau_\delta]}{\log(1/\delta)} \leq c \ \max\left\{ T^\star_{\mathrm{KL}}(\boldsymbol{\mu}), C \frac{T^\star_{\mathrm{TV}}(\boldsymbol{\mu})}{\epsilon} \right\}.$$

where c is a universal constant.

☞ Matches the lower bound up to constants

Experimental Analysis

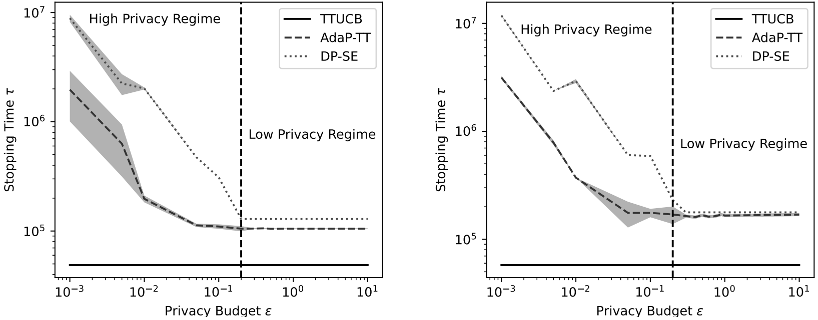


Figure: Evolution of the stopping time τ of AdaP-TT, DP-SE, and TTUCB with respect to the privacy budget ϵ for $\delta = 10^{-2}$ on two Bernoulli instances. The shaded vertical line separates the two privacy regimes. AdaP-TT outperforms DP-SE.

Conclusion and Future Work

Conclusion: We derive sample complexity lower bounds and matching upper bounds for BAI with ϵ-global DP.

Future Work:

- Close the multiplicative gap between the lower and upper bounds.
- Extend the analysis to other DP settings, like (ϵ, δ)-DP and Rényi-DP.
- Extend the analysis to other trust models, like local DP and shuffle DP.

Thank you for your interest in the paper

Come see us at the poster session!

Bibliography I

Dwork, C. and Roth, A. (2014). The algorithmic foundations of differential privacy. Foundations and Trends® in Theoretical Computer Science, 9(3–4):211–407.

Garivier, A. and Kaufmann, E. (2016). Optimal best arm identification with fixed confidence. In Conference on Learning Theory, pages 998–1027. PMLR.