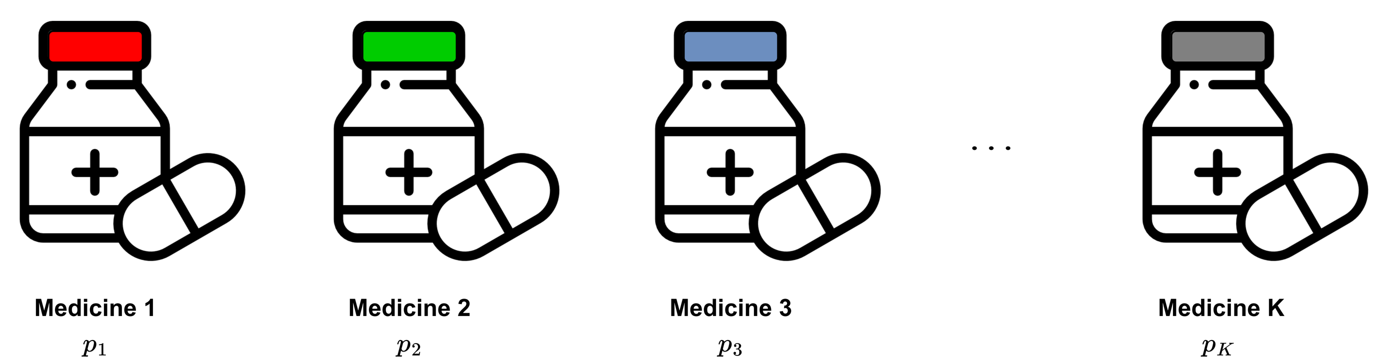
On the Complexity of Differentially Private Best-Arm Identification with Fixed Confidence

Achraf Azize, Marc Jourdan, Aymen Al Marjani, Debabrota Basu
Univ. Lille, Inria, CNRS, Centrale Lille, UMR 9189 CRIStAL, F-59000 Lille, France.

FC-BAI with ϵ-global Differential Privacy

Setting: Clinical trials with K candidate medicines



Goal: Find the medicine with the highest mean $a^\star \triangleq \arg\max_{a\in[K]} p_a$.

Constraint: Protect the privacy of the patients. *A patient's reaction to a medicine can reveal sensitive information about health conditions.*

Interaction Protocol: For the t-th patient in the study:

1. The doctor π chooses a Medicine $a_t \in \{1, \dots, K\}$
2. The doctor observes a reward $r_t \in \{0,1\}$ such that $r_t \sim$ Bernoulli(p_{a_t})

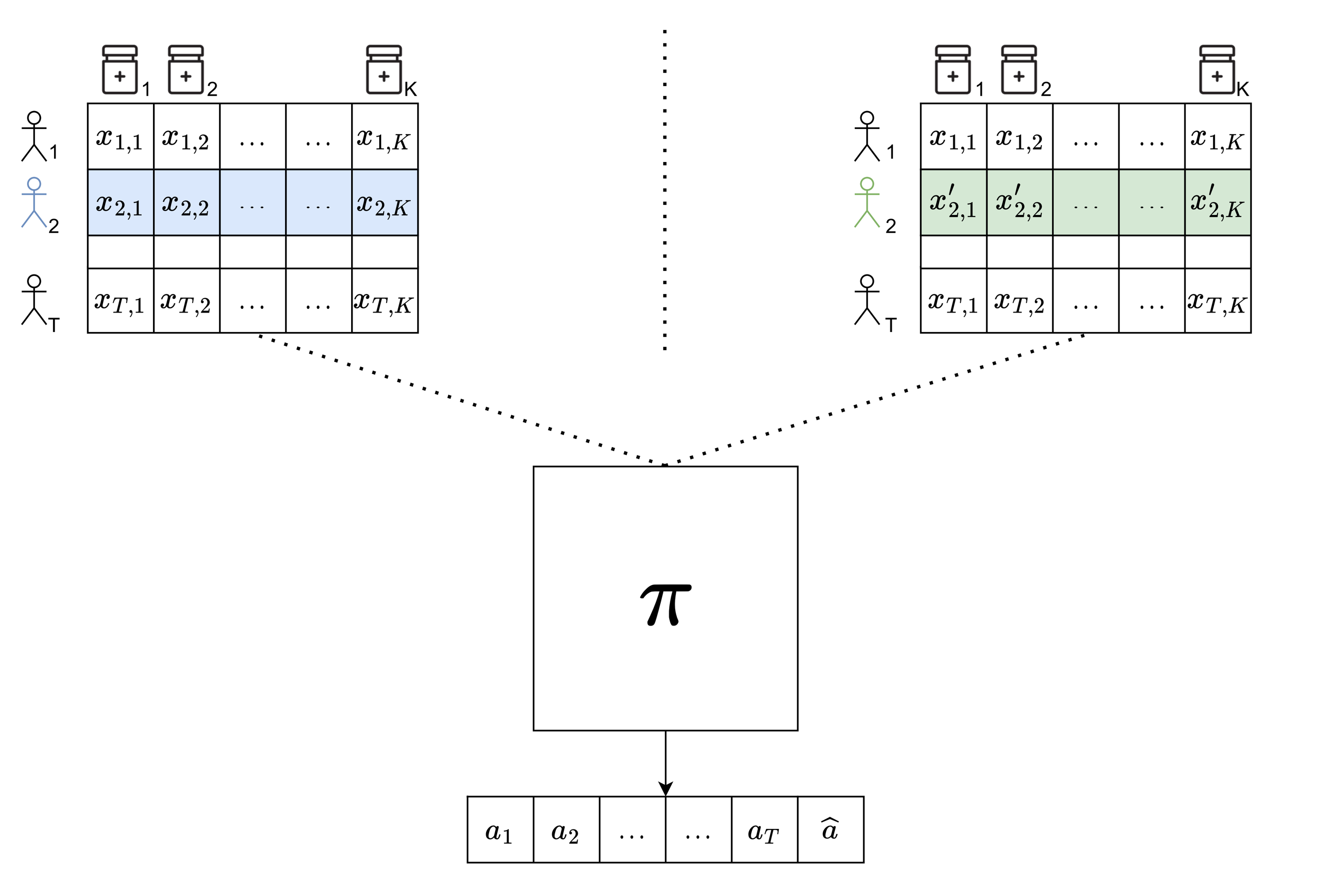
Stop the interaction at time τ and Recommend a final guess $\widehat{a} \in [K]$

Correctness: A BAI strategy π is δ-**correct** for a class $\mathcal{M}$, if for every instance $\boldsymbol{\nu} \in \mathcal{M}$,

$$\mathbb{P}_{\boldsymbol{\nu},\pi}(\tau < \infty, \widehat{a} = a^\star(\boldsymbol{\nu})) \geq 1-\delta.$$

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 $\widehat{a}$,

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



Contributions

1. Lower bound on the sample complexity of δ-correct ϵ-global DP BAI strategies.
2. Algorithm design: an ϵ-global DP variant of Top Two algorithms named AdaP-TT
3. Analysis of AdaP-TT: Enjoys both theoretical near-optimality and good experimental performance.

Algorithm Design

Main Ingredients:

1. Per-arm doubling (Line 5).
2. Forgetting (Line 8).
3. Adding Laplace noise (Line 9).

Algorithm 1 AdaP-TT

1: **Input:** $\beta \in (0,1)$, risk $\delta \in (0,1)$, privacy budget ϵ, thresholds $c_{\epsilon,k_1,k_2} : \mathbb{N}^2 \times (0,1) \to \mathbb{R}^+$
2: **Output:** Recommendation $\hat{a}$ and Stopping time τ satisfying ϵ-global DP
3: **Initialization:** $\forall a \in [K]$, pull arm a, set $k_a = 1, T_1(a) = K+1, L_{n,a} = 0, N_{n,a} = 1, n = K+1$.
4: **for** $n > K$ **do**
5: **if** there exists $a \in [K]$ such that $N_{n,a} \geq 2N_{T_{k_a}(a),a}$ **then**
6: Change phase $k_a \leftarrow k_a + 1$ for this arm a
7: Set $T_{k_a}(a) = n$ and $\tilde{N}_{k_a,a} = N_{T_{k_a}(a),a} - N_{T_{k_a-1}(a),a}$
8: Set $\hat{\mu}_{k_a,a} = \tilde{N}_{k_a,a}^{-1} \sum_{s=T_{k_a-1}(a)}^{T_{k_a}(a)-1} X_s \mathbb{1}\{I_s = a\}$
9: Set $\tilde{\mu}_{k_a,a} = \hat{\mu}_{k_a,a} + Y_{k_a,a}$ where $Y_{k_a,a} \sim$ Lap$((\epsilon \tilde{N}_{k_a,a})^{-1})$
10: **end if**
11: Set $\hat{a}_n = \arg\max_{b\in[K]} \tilde{\mu}_{k_b,b}$
12: **if** $\frac{(\tilde{\mu}_{k_{\hat{a}_n},\hat{a}_n} - \tilde{\mu}_{k_b,b})^2}{1/\tilde{N}_{k_{\hat{a}_n},\hat{a}_n} + 1/\tilde{N}_{k_b,b}} \geq 2c_{\epsilon,k_{\hat{a}_n},k_b}(\tilde{N}_{k_{\hat{a}_n},\hat{a}_n}, \tilde{N}_{k_b,b}, \delta)\ \forall b \neq \hat{a}_n$ **then**
13: **return** $(\hat{a}_n, n)$
14: **end if**
15: Set $B_n = \arg\max_{a\in[K]} \{\tilde{\mu}_{k_a,a} + \sqrt{k_a/\tilde{N}_{k_a,a}} + k_a/(\epsilon\tilde{N}_{k_a,a})\}$
16: Set $C_n = \arg\min_{a\neq B_n} \frac{\tilde{\mu}_{k_{B_n},B_n} - \tilde{\mu}_{k_a,a}}{\sqrt{1/N_{n,B_n} + 1/N_{n,a}}}$
17: Set $I_n = B_n$ if $N_{n,B_n}^{B_n} \leq \beta L_{n+1,B_n}$, else $I_n = C_n$
18: Pull I_n and observe $X_n \sim \nu_{I_n}$
19: Set $N_{n+1,I_n} \leftarrow N_{n,I_n} + 1$, $N_{n+1,I_n}^{B_n} \leftarrow N_{n,I_n}^{B_n} + 1$ and $L_{n+1,B_n} \leftarrow L_{n,B_n} + 1$. Set $n \leftarrow n+1$
20: **end for**

Privacy analysis: For rewards in $[0,1]$, AdaP-TT is ϵ-global DP. A change in one user *only affects* the empirical mean at one episode of an arm, which is made private using the Laplace Mechanism.

Correctness: AdaP-TT is δ-correct for thresholds which verify

$$\tilde{c}_{\epsilon,k_1,k_2}(n,m,\delta) \approx \log(1/\delta) + (1/n + 1/m)\log(1/\delta)^2/\epsilon^2 .$$

Upper bound on expected sample complexity: For Bernoulli instances verifying that $\exists C \geq 1$ such that $\Delta_{\max}/\Delta_{\min} \leq C$, AdaP-TT is ϵ-global DP, δ-correct and satisfies

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

Comparison to DP-SE: DP-SE has two drawbacks:

1. DP-SE is less adaptive than AdaP-TT, i.e. in a phase, DP-SE continues to sample arms that might already be known to be bad.
2. AdaP-TT is anytime, i.e. its sampling does not depend on the risk δ.

Sample complexity lower bound

The lower bound: Let $\delta \in (0,1)$ and $\epsilon > 0$. For any δ-correct ϵ-global DP BAI strategy, we have that

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

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

Simplification: $T^\star_{\mathrm{KL}}(\boldsymbol{\nu}) \approx \sum_a \frac{1}{(\mu_{a^\star}-\mu_a)^2}$ and $T^\star_{\mathrm{TV}}(\boldsymbol{\nu}) \approx \sum_a \frac{1}{\mu_{a^\star}-\mu_a}$

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

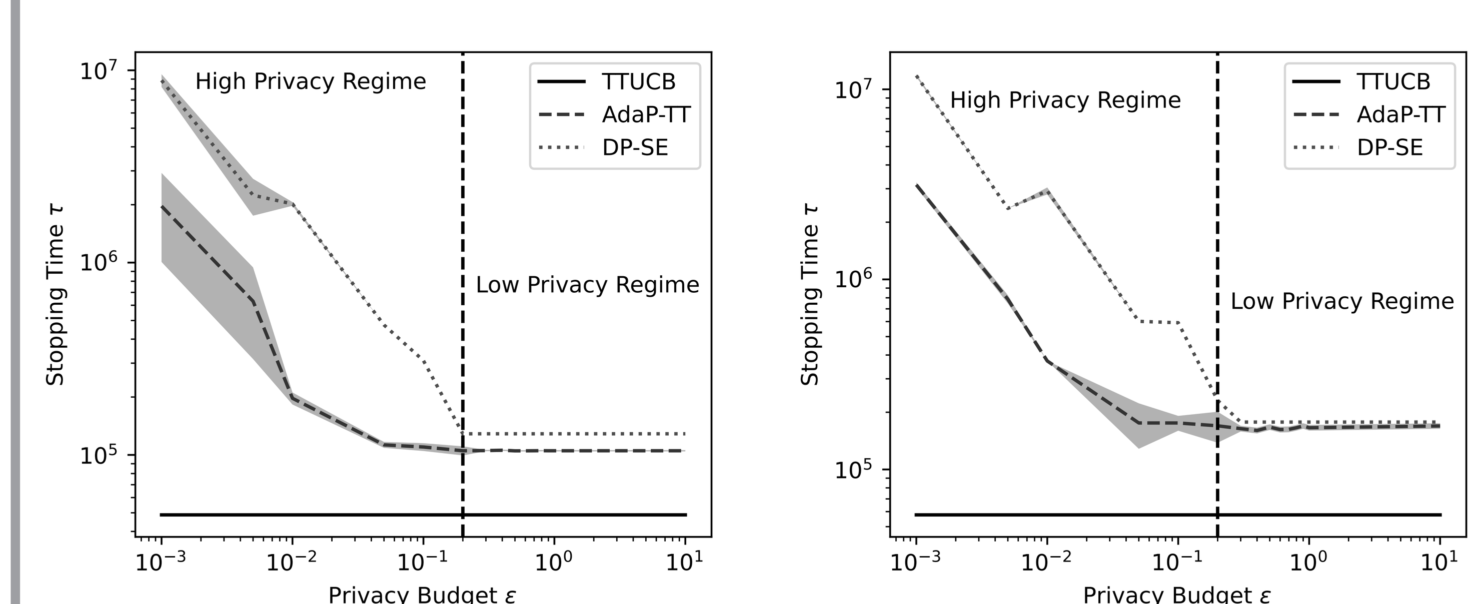
- *Low-privacy regime*: when $\epsilon > \frac{T^\star_{\mathrm{TV}}(\boldsymbol{\nu})}{6T^\star_{\mathrm{KL}}(\boldsymbol{\nu})}$, **privacy can be achieved for free.**
- *High-privacy regime*: when $\epsilon < \frac{T^\star_{\mathrm{TV}}(\boldsymbol{\nu})}{6T^\star_{\mathrm{KL}}(\boldsymbol{\nu})}$, the ϵ-global constraint requires more samples than non-private ones.

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

DP and Total Variation: Stochastic Group Privacy.

- d and d' differ in 1 sample $\to \exp(\epsilon)$
- d and d' differ in k samples $\to \exp(k\epsilon)$
- Sample $d \sim \otimes^n P$ and $d' \sim \otimes^n Q \to \exp(nTV(P,Q)\epsilon)$
- Sample $d \sim \otimes_{i=1}^n P_i$ and $d' \sim \otimes_{i=1}^n Q_i \to \exp(\sum_{i=1}^n TV(P_i,Q_i)\epsilon)$

Experimental analysis



1. AdaP-TT outperforms DP-SE.
2. The performance of AdaP-TT has two regimes: a high-privacy regime (for $\epsilon < 0.2$) and a low privacy regime (for $\epsilon > 0.2$).

Future work

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