

---

# Optimal Best Arm Identification under Differential Privacy

---

Anonymous Author(s)

Affiliation

Address

email

## Abstract

1 Best Arm Identification (BAI) algorithms are deployed in data-sensitive applica-  
2 tions, such as adaptive clinical trials or user studies. Driven by the privacy concerns  
3 of these applications, we study the problem of fixed-confidence BAI under global  
4 Differential Privacy (DP) for Bernoulli distributions. While numerous asymptot-  
5 ically optimal BAI algorithms exist in the non-private setting, a significant gap  
6 remains between the best lower and upper bounds in the global DP setting. This  
7 work reduces this gap to a small multiplicative constant, for any privacy budget  $\epsilon$ .  
8 First, we provide a tighter lower bound on the expected sample complexity of any  
9  $\delta$ -correct and  $\epsilon$ -global DP strategy. Our lower bound replaces the Kullback–Leibler  
10 (KL) divergence in the transportation cost used by the non-private characteristic  
11 time with a new information-theoretic quantity that optimally trades off between  
12 the KL divergence and the Total Variation distance scaled by  $\epsilon$ . Second, we intro-  
13 duce a stopping rule based on these transportation costs and a private estimator  
14 of the means computed using an arm-dependent geometric batching. En route to  
15 proving the correctness of our stopping rule, we derive concentration results of  
16 independent interest for the Laplace distribution and for the sum of Bernoulli and  
17 Laplace distributions. Third, we propose a Top Two sampling rule based on these  
18 transportation costs. For any budget  $\epsilon$ , we show an asymptotic upper bound on  
19 its expected sample complexity that matches our lower bound to a multiplicative  
20 constant smaller than 8. Our algorithm outperforms existing  $\delta$ -correct and  $\epsilon$ -global  
21 DP BAI algorithms for different values of  $\epsilon$ .

## 22 1 Introduction

23 The stochastic Multi-Armed Bandit (MAB) is an interactive sequential decision-making model [18,  
24 59], introduced by William R. Thompson [81]. Thompson’s motivation for studying MABs is to  
25 design clinical trials that adapt treatment allocations on the fly as the medicines appear more or less  
26 effective. Specifically, in MABs, a learner interacts with  $K \in \mathbb{N}$  unknown probability distributions,  
27 referred to as *arms*. In clinical trials, the arms are the candidate medicines, while the observations  
28 are patient reactions, 1 if the patient is cured and 0 otherwise. The learner aims to identify the arm  
29 with the highest average efficiency, i.e., the medicine that cures most patients in expectation. Given  
30 a fixed error  $\delta \in (0, 1)$ , Best Arm Identification (BAI) [5, 47] algorithms in the fixed confidence  
31 setting [34, 36, 38] suggest a candidate answer that coincides with the optimal arm with probability  
32 more than  $1 - \delta$ , while using as few samples as possible.

33 BAI algorithms have been increasingly deployed in data-sensitive applications, such as adaptive  
34 clinical trials [81, 72, 8], pandemic mitigation [62], user studies [64], crowdsourcing [90], online  
35 advertisement [22], hyperparameter tuning [61], and communication networks [63], to name a few.  
36 Due to the adaptive nature of these procedures, critical data privacy concerns are raised [83], as

exemplified by the adaptive dose finding trial. For each new patient  $n$ , a physician chooses a dose level  $a_n \in [K] := \{1, \dots, K\}$  based on previous observations, and collects a binary observation measuring the effect of the selected dose on the patient. Crucially, the patients' reactions might reveal information regarding their health. Subsequently, these outcomes will guide the physician's decision for future patients. Eventually, the physician adaptively decides to stop the trial and recommends a dose  $\hat{a}_{\tau_\delta}$  after collecting  $\tau_\delta$  samples, referred to as *sample complexity*. Even if those outcomes are kept secret, the experimental findings and protocol are detailed thoroughly to the health authorities. This report contains the sequence of chosen dose levels  $(a_n)_{n \leq \tau_\delta}$  and the recommended dose level  $\hat{a}_{\tau_\delta}$ , both indirectly leaking information regarding the patients involved in the trial. This example underscores the need for privacy-preserving fixed-confidence BAI algorithms.

We adopt the Differential Privacy (DP) framework [30], which bounds the influence of any single data point. Given a privacy budget  $\epsilon$ , we consider the  $\epsilon$ -global DP constraint that assumes the existence of a trusted curator (e.g., the physician running the clinical trial), who observes the outcomes and ensures privacy when publishing these findings. While  $\epsilon$ -global DP is well-studied in regret minimization [68, 9, 13], its impact on fixed-confidence BAI is less understood [74, 54]. A significant gap remains between the existing lower and upper bounds [11, 12]. This paper reduces this gap to a small constant for any privacy budget  $\epsilon$ . Appendix C.1 contains a detailed literature review.

**Contributions.** Our contributions for fixed-confidence BAI under  $\epsilon$ -global DP are threefold.

**1. Lower bound under global DP.** We derive a novel information-theoretic lower bound on the expected sample complexity of any  $\delta$ -correct and  $\epsilon$ -global DP BAI algorithms (Theorem 2). Our lower bound replaces the Kullback-Leibler (KL) divergence in the transportation cost of the non-private characteristic time with an information-theoretic quantity  $d_\epsilon$  (Eq. (2)) that smoothly interpolates between the KL divergence and the Total Variation (TV) distance scaled by  $\epsilon$ .

**2. Private estimator and GLR stopping rule.** We introduce a private estimator using arm-dependent geometric batching without forgetting and a GLR stopping rule based on the  $d_\epsilon$  refined transportation costs. Its correctness (Theorem 5) required novel tails concentration results for Laplace distributions and the sum of Bernoulli and Laplace distributions, which could be of independent interest.

**3. Asymptotically optimal algorithm.** We propose a new Top Two sampling rule (DP-TT, Algorithm 1) based on the  $d_\epsilon$ -transportation costs suggested by our lower bound. We show that the asymptotic expected sample complexity of DP-TT matches our lower bound for any privacy budget  $\epsilon$  up to a constant smaller than 8 (Theorem 6). DP-TT outperforms all the other  $\delta$ -correct  $\epsilon$ -global DP BAI algorithms on all tested instances and all  $\epsilon$ .

## 2 Background: Best Arm Identification under Differential Privacy

In this section, we present the Best Arm Identification (BAI) under fixed confidence problem [38], introduce the Differential Privacy (DP) [31] constraint, and finally extend DP to BAI algorithms.

**BAI under Fixed Confidence.** A Bernoulli bandit instance  $\nu := (\nu_a)_{a \in [K]} \in \mathcal{F}^K$  is characterized by its means  $\mu := (\mu_a)_{a \in [K]} \in (0, 1)^K$ . The *best* (optimal) arm  $a^*$  is assumed to be unique, i.e.,  $a^*(\nu) = a^*(\mu) := \arg \max_{a \in [K]} \mu_a = \{a^*\}$ . Let  $\delta \in (0, 1)$  be the risk parameter. A fixed confidence BAI algorithm  $\pi$  specifies three rules that rely on previously observed samples and some exogenous randomness. The *sampling rule* determines the next arm to pull  $a_n \in [K]$  for which  $X_{n,a_n} \sim \nu_{a_n}$  is observed. The *recommendation rule* recommends a *candidate* arm  $\tilde{a} \in [K]$ . The *stopping rule* decides when to stop collecting additional samples and output the current candidate arm. The stopping time  $\tau_{\epsilon, \delta}$  is the *sample complexity*. Let  $\mathbb{P}_{\nu, \pi}$  and  $\mathbb{E}_{\nu, \pi}$  denote the probability and expectation taken over the randomness of the observations from  $\nu$  and the algorithm  $\pi$  (e.g., due to its privacy mechanism). A fixed-confidence BAI algorithm  $\pi$  is  $\delta$ -correct when  $\mathbb{P}_{\nu, \pi}(\tau_{\epsilon, \delta} < +\infty, \tilde{a} \notin a^*(\nu)) \leq \delta$  for all  $\nu \in \mathcal{F}^K$ .

**Differential Privacy (DP).** An algorithm satisfies the Differential Privacy constraint if the algorithm's outputs are "essentially" equally likely to occur, for any two input datasets that only differ in one individual's data. An adversary only observing the mechanism's output cannot distinguish whether any individual's data was included. A privacy budget  $\epsilon$  captures the closeness of the output distributions. Smaller  $\epsilon$  means stronger privacy.

**Definition 1** ( $\epsilon$ -DP [31]). A mechanism  $\mathcal{M}$  satisfies  $\epsilon$ -DP for a given  $\epsilon \geq 0$ , if, for all neighboring datasets  $D \sim D'$ , where  $D \sim D'$  if and only if  $d_{\text{Ham}}(D, D') := \sum_{t=1}^T \mathbb{1}\{D_t \neq D'_t\} \leq 1$ , i.e.,  $D$  and  $D'$  differ by at most one record, and for all sets of output  $\mathcal{O} \subseteq \text{Range}(\mathcal{M})$ ,  $\Pr[\mathcal{M}(D) \in \mathcal{O}] \leq e^\epsilon \Pr[\mathcal{M}(D') \in \mathcal{O}]$  where the probability space is over the coin flips of the mechanism  $\mathcal{M}$ .

To ensure  $\epsilon$ -DP, the Laplace mechanism [32, 30] adds calibrated Laplacian noise to the algorithm's output. Let  $\text{Lap}(b)$  be the Laplace distribution with mean/variance  $(0, 2b^2)$ .

**Theorem 1** (Laplace mechanism, Theorem 3.6 [30]). Let  $f : \mathcal{X} \rightarrow \mathbb{R}^d$  be an algorithm with sensitivity  $s(f) \triangleq \max_{\mathcal{D}, \mathcal{D}' \text{ s.t. } d_{\text{Ham}}(\mathcal{D}, \mathcal{D}')=1} \|f(\mathcal{D}) - f(\mathcal{D}')\|_1$ , where  $\|\cdot\|_1$  is the  $\ell_1$  norm. Let  $(Z_i)_{i \in [d]}$  be i.i.d. from  $\text{Lap}(s(f)/\epsilon)$ , then the noisy output  $f(\mathcal{D}) + (Z_i)_{i \in [d]}$  satisfies  $\epsilon$ -DP.

**DP for BAI.** In BAI algorithms, the private input is the observation dataset and the output is the recommended candidate arm  $\tilde{a}$  and the sequence of sampled actions  $(a_n)_{n < \tau_{\epsilon, \delta}}$  until stopping at  $\tau_{\epsilon, \delta}$ . Let  $R = \{r_1, \dots\}$  be a sequence of private observations. Given a fixed sequence of observations  $R$ , we denote by  $\Pr[\pi(R) = (T+1, \tilde{a}, (a_1, \dots, a_T))]$  the probability that the BAI algorithm  $\pi$  stops at step  $T+1$ , recommending action  $\tilde{a}$  and sampling actions  $(a_1, \dots, a_T)$  when interacting with  $R$ . The randomisation in this probability comes only from the BAI algorithm's sampling, recommendation and stopping rules, whereas the observations are fixed. Then, a BAI algorithm  $\pi$  is said to be  $\epsilon$ -global DP if, for every two neighboring sequences of observations  $R$  and  $R'$ , and for every possible stopping time, recommendation and sampled actions  $(T+1, \tilde{a}, (a_1, \dots, a_T))$ , we have that

$$\Pr[\pi(R) = (T+1, \tilde{a}, (a_1, \dots, a_T))] \leq e^\epsilon \Pr[\pi(R') = (T+1, \tilde{a}, (a_1, \dots, a_T))].$$

**Main Goal:** Design  $\epsilon$ -global DP  $\delta$ -correct BAI algorithms, with the smallest sample complexity  $\tau_{\epsilon, \delta}$ .

**Notation.** Let  $[x]_0^1 := \max\{0, \min\{1, x\}\}$  be the clipping operator to  $[0, 1]$ . Let  $\mathbb{1}(\cdot)$  be the indicator function. For two probability distributions  $\mathbb{P}$  and  $\mathbb{Q}$  on the measurable space  $(\Omega, \mathcal{G})$ , the Total Variation (TV) distance is  $\text{TV}(\mathbb{P} \parallel \mathbb{Q}) := \sup_{A \in \mathcal{G}} \{\mathbb{P}(A) - \mathbb{Q}(A)\}$  and the Kullback-Leibler (KL) divergence is  $\text{KL}(\mathbb{P} \parallel \mathbb{Q}) := \int \log\left(\frac{d\mathbb{P}}{d\mathbb{Q}}(\omega)\right) d\mathbb{P}(\omega)$ , when  $\mathbb{P} \ll \mathbb{Q}$ , and  $+\infty$  otherwise. The KL divergence and TV distance between two Bernoulli distributions with means  $(p, q) \in (0, 1)^2$  are the relative entropy denoted by  $\text{kl}$ , i.e.,  $\text{KL}(\text{Ber}(p) \parallel \text{Ber}(q)) = \text{kl}(p, q) := p \log(p/q) + (1-p) \log((1-p)/(1-q))$ , and the absolute mean difference, i.e.,  $\text{TV}(\text{Ber}(p) \parallel \text{Ber}(q)) = |p - q|$ . Let  $\Delta_K := \{w \in \mathbb{R}^K \mid w \geq 0, \sum_{a \in [K]} w_a = 1\}$  be the probability simplex of dimension  $K-1$ . For all  $a \in [K]$ , let  $N_{n,a} := \sum_{t \in [n-1]} \mathbb{1}(a_t = a)$  be the global pulling count of arm  $a$  before time  $n$ .

### 3 Lower Bound on the Expected Sample Complexity

In order to be  $\delta$ -correct, an algorithm has to be able to distinguish  $\nu$  from *alternative* instances with different best arms, i.e., an instance  $\kappa \in \text{Alt}(\nu) := \{\kappa \in \mathcal{F}^K \mid a^*(\kappa) \neq a^*(\nu)\}$ . On the other hand, being  $\epsilon$ -global DP forces an algorithm to have similar behaviour on similar instances. The tension between these two requirements yields the following problem-dependent lower bound on the expected sample complexity  $\mathbb{E}_{\nu\pi}[\tau_{\epsilon, \delta}]$  for any algorithm  $\pi$  on any instance  $\nu$ .

**Theorem 2.** Let  $(\epsilon, \delta) \in \mathbb{R}_+^* \times (0, 1)$ . For any algorithm  $\pi$  that is  $\delta$ -correct and  $\epsilon$ -global DP on  $\mathcal{F}^K$ ,

$$\mathbb{E}_{\nu\pi}[\tau_{\epsilon, \delta}] \geq T_\epsilon^*(\nu) \log(1/(3\delta))$$

for all  $\nu \in \mathcal{F}^K$  with unique best arm. The inverse of the characteristic time  $T_\epsilon^*(\nu)$  is defined as

$$T_\epsilon^*(\nu)^{-1} := \sup_{w \in \Delta_K} \inf_{\kappa \in \text{Alt}(\nu)} \sum_{a=1}^K w_a d_\epsilon(\nu_a, \kappa_a), \quad (1)$$

$$d_\epsilon(\nu_a, \kappa_a) := \inf_{\varphi_a \in \mathcal{F}} \{\text{KL}(\varphi_a \parallel \kappa_a) + \epsilon \cdot \text{TV}(\nu_a \parallel \varphi_a)\}. \quad (2)$$

**Comments.** (a) The characteristic time in the lower bound is the value of a two-player zero-sum game between a MIN player, who plays instances  $\kappa$  close of  $\nu$  in order to confuse the MAX player, who in order plays an arm allocation  $w \in \Delta_K$  to distinguish between  $\nu$  and  $\kappa$ .

(b) The crucial part in characteristic times similar to Eq. (1) is finding the “right” measure capturing the “distinguishability” between instances. In the non-private lower bounds, this is captured by the

KL divergence for parametric distributions [38] and by the Kinf (i.e., inf KL under mean constraint) for non-parametric distributions [2]. In the DP lower bounds of Azize et al. [11], it is captured by  $\min\{\text{KL}, \epsilon \text{TV}\}$ . In Theorem 2, it is captured by  $d_\epsilon$  (as in Eq. (2)) that smoothly interpolates between KL and TV. Azize et al. [13] recently introduced  $d_\epsilon$  for  $\epsilon$ -global DP regret minimization. Our results show that  $d_\epsilon$  also tightly captures the hardness of fixed-confidence BAI under  $\epsilon$ -global DP. Namely, our DP-TT algorithm achieves a matching upper bound when  $\delta \rightarrow 0$  (up to a constant smaller than 8), for all instances with distinct means and all values of  $\epsilon$ .

(c) Azize et al. [11, Theorem 2] provides a lower bound on the sample complexity of any  $\epsilon$ -global  $\delta$ -correct algorithm, where the inverse characteristic time is  $\sup_{w \in \Delta_K} \inf_{\kappa \in \text{Alt}(\nu)} \min\{\sum_{a=1}^K w_a \text{KL}(\nu_a \parallel \kappa_a), 6\epsilon \sum_{a=1}^K w_a \text{TV}(\nu_a \parallel \kappa_a)\}$ . The lower bound of Theorem 2 is strictly tighter than that of Theorem 2 in [11], for all instances  $\nu$  and values of  $\epsilon$ . The reason is that  $d_\epsilon(\mathbb{P}, \mathbb{Q}) \leq \min\{\text{KL}(\mathbb{P}, \mathbb{Q}), \epsilon \text{TV}(\mathbb{P}, \mathbb{Q})\}$  for any two distributions.

(d) The lower bound of Theorem 2 suggests the existence of two privacy regimes, depending on the value of  $\epsilon$  and the instance  $\nu$ . Specifically, when  $\epsilon$  is big,  $d_\epsilon$  reduces to the KL, and we retrieve the classic non-private lower bound. On the other hand, as  $\epsilon \rightarrow 0$ ,  $d_\epsilon$  reduces to  $\epsilon \text{TV}$ , and the characteristic time reduces to  $\frac{1}{\epsilon} T_{\text{TV}}^*(\nu) := \frac{1}{\epsilon} \left( \sup_{w \in \Delta_K} \inf_{\kappa \in \text{Alt}(\nu)} \sum_{a=1}^K w_a \text{TV}(\nu_a \parallel \kappa_a) \right)^{-1} = \frac{1}{\epsilon} \sum_{a=1}^k \frac{1}{\Delta_a}$ , where  $\Delta_a = \mu^* - \mu_a$  for  $a \neq a^*$  and  $\Delta_{a^*} = \min_{a \neq a^*} \Delta_a$ . This improves the high privacy regime lower bound of prior work by a factor 6. Also, the value of  $\epsilon$  at which the privacy regimes change can be tightly specified, which we quantify for Bernoulli instances in the following.

**Proof Sketch and Techniques.** The proof uses the standard reduction to hypothesis testing [38], using the data-processing inequality. The asymptotic techniques used by [13] for regret cannot be adapted for our non-asymptotic lower bound. Thus, new techniques are needed. The main technical novelty of the proof is a tighter quantification of the “similar” behaviour of a DP mechanism when applied to stochastic datasets. Specifically, let  $\mathcal{M}$  be an  $\epsilon$ -DP mechanism. Given two data-generating distributions  $\mathbb{P}$  and  $\mathbb{Q}$ , letting  $\mathbb{M}_{\mathbb{P}, \mathcal{M}}$  (resp.  $\mathbb{M}_{\mathbb{Q}, \mathcal{M}}$ ) be the marginal over outputs of the mechanism when the input dataset is generated through  $\mathbb{P}$  (resp.  $\mathbb{Q}$ ), then we show that

$$\text{KL}(\mathbb{M}_{\mathbb{P}, \mathcal{M}} \parallel \mathbb{M}_{\mathbb{Q}, \mathcal{M}}) \leq \inf_{\mathbb{L}} \left\{ \epsilon \inf_{\mathbb{C}_{\mathbb{P}, \mathbb{L}}} \left\{ \mathbb{E}_{D, D' \sim \mathbb{C}_{\mathbb{P}, \mathbb{L}}} [d_{\text{Ham}}(D, D')] \right\} + \text{KL}(\mathbb{L} \parallel \mathbb{Q}) \right\},$$

where the first infimum is over all distributions  $\mathbb{L}$  on the input space, and the second infimum is an optimal transport problem over all couplings between  $\mathbb{P}$  and  $\mathbb{L}$ , where the cost is the Hamming distance (introduced in Definition 1). This bound of general interest could be applied to get tighter lower bounds in any DP application using stochastic inputs. For product and bandit distributions, we solve the optimal transport using maximal couplings, where the Total Variation naturally appears, while keeping the first infimum unchanged, giving rise to the  $d_\epsilon$  quantity. Finally, plugging the new upper bound on the KL in the hypothesis reduction concludes the sample complexity lower bound proof. A detailed proof and discussion of all these claims is given in Appendix D.

**Properties of the Characteristic Time and Optimal Allocation.** The set  $w_\epsilon^*(\nu)$  of *optimal allocations* is the maximizer of the outer supremum on  $\Delta_K$  that defines  $T_\epsilon^*(\nu)^{-1}$  in Eq. (1). Theorem 3 gathers key properties satisfied by  $T_\epsilon^*(\nu)$  and  $w_\epsilon^*(\nu)$ , for Bernoulli distributions. See lemmas proven in Appendix G, i.e., Lemmas 22, 23, 24, 25, 35, 42 and 46.

**Theorem 3.** For all  $x \in [0, 1]$ , let us define  $g_\epsilon^-(x) := \frac{x e^\epsilon}{x(e^\epsilon - 1) + 1}$  and  $g_\epsilon^+(x) := 1 - g_\epsilon^-(1 - x) = (g_\epsilon^-)^{-1}(x)$ . For all  $(\lambda, \mu) \in \mathbb{R} \times [0, 1]$ , the signed divergences are defined as

$$\begin{aligned} d_\epsilon^+(\lambda, \mu) &:= \mathbb{1}(\mu > [\lambda]_0^1) \inf_{z \in [\lambda]_0^1, \mu]_0^1} \{ \text{kl}(z, \mu) + \epsilon(z - [\lambda]_0^1) \} \\ &= \begin{cases} 0 & \text{if } \mu \in [0, [\lambda]_0^1] \\ -\log(1 - \mu(1 - e^{-\epsilon})) - \epsilon[\lambda]_0^1 & \text{if } \mu \in (g_\epsilon^-([\lambda]_0^1), 1] \\ \text{kl}(\lambda, \mu) & \text{if } \lambda \in (0, 1) \text{ and } \mu \in ([\lambda]_0^1, g_\epsilon^-([\lambda]_0^1)) \end{cases}, \\ d_\epsilon^-(\lambda, \mu) &:= \mathbb{1}(\mu < [\lambda]_0^1) \inf_{z \in [\mu, [\lambda]_0^1]} \{ \text{kl}(z, \mu) + \epsilon([\lambda]_0^1 - z) \} = d_\epsilon^+(1 - \lambda, 1 - \mu). \end{aligned} \quad (3)$$

For  $(\mu, w) \in \mathbb{R}^K \times \mathbb{R}_+^K$ , the transportation cost of the pair of arms  $(a, b) \in [K]^2$  is defined as

$$W_{\epsilon, a, b}(\mu, w) := \mathbb{1}([\mu_a]_0^1 > [\mu_b]_0^1) \inf_{u \in [0, 1]} \{ w_a d_\epsilon^-(\mu_a, u) + w_b d_\epsilon^+(\mu_b, u) \}. \quad (4)$$

169 Let  $\nu \in \mathcal{F}^K$  having means  $\mu \in (0, 1)^K$  with unique best arm  $a^*$ . Then, we have

$$T_\epsilon^*(\nu)^{-1} = \max_{w \in \Delta_K} \min_{a \neq a^*} W_{\epsilon, a^*, a}(\mu, w) \quad \text{and} \quad T_\epsilon^*(\nu) \geq \sum_{a \in [K]} \Delta_{\epsilon, a}^{-1}. \quad (5)$$

170 where  $\Delta_{\epsilon, a^*} := \min_{a \neq a^*} d_\epsilon^-(\mu_{a^*}, \mu_a)$  and  $\Delta_{\epsilon, a} := d_\epsilon^+(\mu_a, \mu_{a^*})$  for all  $a \neq a^*$ . The optimal  
 171 allocation is unique, has dense support and ensures the equality of the transportation costs with  
 172  $T_\epsilon^*(\nu)^{-1}$  (i.e., information balance equation), namely  $w_\epsilon^*(\nu) = \{w_\epsilon^*\}$ ,  $\min_{a \in [K]} w_{\epsilon, a}^* > 0$  and  
 173  $W_{\epsilon, a^*, a}(\mu, w_\epsilon^*) = T_\epsilon^*(\nu)^{-1}$  for all  $a \neq a^*$ .

174 We use the notation  $d_\epsilon^\pm$  to refer to both  $(d_\epsilon^-, d_\epsilon^+)$  as in Eq. (3). Using signed divergences  $d_\epsilon^\pm$   
 175 instead of  $d_\epsilon$  as in Eq. (2) is a convention borrowed from non-parametric fixed-confidence BAI (i.e.,  
 176  $\mathcal{K}_{\text{inf}}^\pm$  [50]) that explicits the ordering between the mean parameters. Given  $(\kappa, \nu) \in \mathcal{F}^2$  with means  
 177  $(\lambda, \mu) \in (0, 1)^2$ , we have  $d_\epsilon(\kappa, \nu) = d_\epsilon^+(\lambda, \mu)$  when  $\mu > \lambda$ , and  $d_\epsilon(\kappa, \nu) = d_\epsilon^-(\lambda, \mu)$  otherwise  
 178 (Lemma 21). The signed divergences  $d_\epsilon^\pm$  and the transportation costs  $(W_{\epsilon, a, b})_{(a, b) \in [K]^2}$  satisfy all  
 179 the desired properties required to study BAI algorithms based on the empirical version of  $W_{\epsilon, a, b}$  (see  
 180 Lemmas 22, 23, 24 and 25, as well as Lemmas 34, 35, 36 and 37), e.g., symmetry, explicit formula,  
 181 monotonicity, strict convexity, etc. In Garivier and Kaufmann [38], the characteristic time and its  
 182 optimal allocation can be computed with a simpler optimisation problem. A simpler optimization  
 183 problem can also be solved to compute  $T_\epsilon^*(\nu)$  and  $w_\epsilon^*(\nu)$  explicitly (Lemma 46).

184 **Allocation Dependent Low Privacy Regime.** Let  $(\mu, w, a, b) \in (0, 1)^K \times \mathbb{R}_+^K \times [K]^2$  such that  
 185  $\mu_a > \mu_b$  and  $\min\{w_a, w_b\} > 0$ . The non-private Bernoulli transportation costs [38] are defined as

$$W_{a, b}(\mu, w) := w_a \text{kl}(\mu_a, \mu_{a, b}^w) + w_b \text{kl}(\mu_b, \mu_{a, b}^w) \quad \text{with} \quad \mu_{a, b}^w := \frac{w_a \mu_a + w_b \mu_b}{w_a + w_b}.$$

186 We provide an *allocation-dependent* low-privacy condition that depends on  $(\epsilon, \mu, w)$  (Lemma 44),  
 187 i.e.,  $W_{\epsilon, a, b}(\mu, w) = W_{a, b}(\mu, w)$  is implied by

$$\mu_a - \mu_b \leq (1 - e^{-\epsilon}) \min \left\{ (1 + w_a/w_b) \mu_a g_\epsilon^-(1 - \mu_a), (1 + w_b/w_a) (1 - \mu_b) g_\epsilon^-(\mu_b) \right\}. \quad (6)$$

188 Plugging  $w_\epsilon^*$  from Theorem 3 in Eq. (6) would give an implicit condition on  $(\epsilon, \mu)$  under which  
 189 the non-private characteristic time  $T^*(\nu)$  for Bernoulli distributions is recovered, i.e.,  $T_\epsilon^*(\nu) =$   
 190  $T^*(\nu)$ . A weaker (yet explicit) allocation-independent sufficient condition for  $T_\epsilon^*(\nu) = T^*(\nu)$  is  
 191  $\epsilon \geq \max_{a \neq a^*} \epsilon_{a^*, a}$  where  $\epsilon_{a, b} := \log \left( \frac{\mu_a(1 - \mu_b)}{\mu_b(1 - \mu_a)} \right)$ .

## 192 4 Generalized Likelihood Ratio Stopping Rule

193 Designing appropriate recommendation and stopping rules for the BAI problem can be framed as a  
 194 sequential hypothesis testing task with multiple hypotheses  $\{\mu_a = \max_{b \in [K]} \mu_b\}$ . One of the earliest  
 195 approaches to active hypothesis testing—where data collection is also optimized—was introduced  
 196 by Chernoff [25], who advocated for the use of Generalized Likelihood Ratio (GLR) tests for stopping  
 197 decisions. This methodology is also popular in the context of BAI [38]. Despite its relevance, fewer  
 198 works attempted to extend it for private sequential hypothesis testing, see, e.g., Zhang et al. [87]  
 199 under Rényi DP and Azize et al. [12] under  $\epsilon$ -local and  $\epsilon$ -global DP.

200 **Mean Estimator.** Three rules need to be specified to define a BAI algorithm: recommendation,  
 201 sampling, and stopping rules. An important remark in designing BAI algorithms is that the dependence  
 202 of these rules on the private input observation dataset comes solely through the sequence of mean  
 203 estimators. Thus, designing a sequence of mean estimators that satisfy DP is crucial when defining a  
 204  $\epsilon$ -global DP BAI algorithm. To estimate the sequence of means, defined in Lines 5-8 of Algorithm 1,  
 205 we rely on two ingredients: adaptive arm-dependent episodes with a geometric update grid and the  
 206 Laplace mechanism. We call this mechanism estimating the sequence of means the Geometric Private  
 207 Estimator, i.e.,  $\text{GPE}_\eta(\epsilon)$ . Most notably, we eliminate “observation forgetting” from  $\text{GPE}_\eta(\epsilon)$ , an  
 208 important design choice made in all past BAI algorithms [74, 11, 12]. Specifically, for some  $\eta > 0$   
 209 called the geometric grid parameter,  $\text{GPE}_\eta(\epsilon)$  estimates the noisy means in arm-dependent phases: a  
 210 phase changes when the counts of an arm has increased multiplicatively by  $1 + \eta$  (Line 5). Then,  
 211  $\text{GPE}_\eta(\epsilon)$  only updates the mean of the arm that changed phases, by accumulating the observations  
 212 collected from its last phase and adding Laplace noise (Line 7). Due to this accumulation step, we  
 213 *do not forget* the observations from past phases. Thus, each estimated noisy mean  $\tilde{\mu}_{n, a}$  in Line 7

contains  $\tilde{N}_{n,a}$  i.i.d. observations from  $\nu_a$  and  $k_{n,a} \approx \log_{1+\eta} \tilde{N}_{n,a}$  i.i.d. observations from  $\text{Lap}(1/\epsilon)$ . In contrast, using forgetting produces a noisy mean that contains fewer i.i.d. observations from  $\nu_a$  (e.g.  $\tilde{N}_{n,a}/2$  samples for forgetting with  $\eta = 1$ ), but only *one* Laplace noise. While removing forgetting allows us to keep more signal, i.e., more i.i.d samples from  $\nu_a$ , we need more noise, i.e., the cumulative sum of  $\text{Lap}(1/\epsilon)$ , which is logarithmic in the number of samples from  $\nu_a$ . Tighter concentration inequalities allow controlling the cumulative sum of Laplace noise. See below for a detailed discussion about our novel concentration results. As long as the number of samples from the  $\text{Lap}(1/\epsilon)$  is logarithmic in the number of samples from  $\nu_a$ , the effect of noise on the sample complexity is similar to having only *one* additional Laplace noise.

**Privacy Analysis.** By adaptive post-processing, the following lemma is proved naturally.

**Lemma 4.** *Any BAI algorithm using only  $\text{GPE}_\eta(\epsilon)$  to access observations is  $\epsilon$ -global DP on  $[0, 1]$ .*

**Proof Sketch.** The proof combines two steps. First, we show that the sequence of mean estimators produced by  $\text{GPE}_\eta(\epsilon)$  is  $\epsilon$ -DP. The crucial observation is that a change in one observation *only* affects the partial sum collected in just *one* arm-phase. By the Laplace mechanism, adding one  $\text{Lap}(1/\epsilon)$  to the partial sum is enough to make it  $\epsilon$ -DP. Then, by post-processing, the sequence of accumulated partial sums ( $\tilde{S}_{k_{n,a},a}$ ) and noisy means ( $\mu_{n,a}$ ) (Line 7) are also  $\epsilon$ -DP. The second step shows how to use the sequential nature of the process and adaptive post-processing to conclude that BAI algorithms using only  $\text{GPE}_\eta(\epsilon)$  are  $\epsilon$ -global DP. The detailed proof is in Appendix E.

**Recommendation Rule.** The recommendation rule  $\tilde{a}_n$  is defined as the arm with the highest clipped noisy empirical mean, i.e.,  $\tilde{a}_n \in \arg \max_{a \in [K]} [\tilde{\mu}_{n,a}]_0^1$  where ties are broken uniformly at random.

**GLR Stopping Rule.** The GLR stopping rule runs  $K$  sequential GLR tests in parallel, and stops as soon as one of these tests can reject the null hypothesis. When comparing the recommendation  $\tilde{a}_n$  with an alternative arm  $a$ , the GLR statistic is defined as the transportation cost  $W_{\epsilon, \tilde{a}_n, a}$  evaluated empirically at  $(\tilde{\mu}_n, \tilde{N}_n)$  (see Eq. (4)). Intuitively,  $W_{\epsilon, \tilde{a}_n, a}(\tilde{\mu}_n, \tilde{N}_n)$  represents the amount of empirical evidence to reject the hypothesis that arm  $a$  has a higher mean than  $\tilde{a}_n$ . One can stop and recommend  $\tilde{a}_n$  when all these statistics exceed a given stopping threshold. Given a privacy budget and risk  $(\epsilon, \delta) \in \mathbb{R}_+^* \times (0, 1)$  and a stopping threshold  $c : \mathbb{N} \times \mathbb{R}_+^* \times (0, 1) \rightarrow \mathbb{R}_+$ , we define

$$\tau_{\epsilon, \delta} = \inf \{ n \mid \forall a \neq \tilde{a}_n, W_{\epsilon, \tilde{a}_n, a}(\tilde{\mu}_n, \tilde{N}_n) > c(\tilde{N}_{n, \tilde{a}_n}, \epsilon, \delta) + c(\tilde{N}_{n, a}, \epsilon, \delta) \}. \quad (7)$$

Given its proximity to the characteristic time  $T_\epsilon^*(\nu)$ , see Eq. (5) (Theorem 3), the GLR stopping rule is a good candidate to match the lower bound, i.e., if one could sample arms according to  $w_\epsilon^*(\nu)$  and use the stopping threshold  $\log(1/\delta)$ . Unfortunately, this threshold is too good to be  $\delta$ -correct and  $w_\epsilon^*(\nu)$  should be estimated as it is unknown (Section 5).

**Calibration of the Stopping Threshold.** Regardless of the sampling rule, the stopping threshold should ensure  $\delta$ -correctness of the GLR stopping rule (Theorem 5).

**Theorem 5.** *Let  $(\epsilon, \delta, \eta) \in \mathbb{R}_+^* \times (0, 1) \times \mathbb{R}_+^*$ . Let  $s > 1$ ,  $\zeta$  be the Riemann  $\zeta$  function and  $\bar{W}_{-1}(x) = -W_{-1}(-e^{-x})$  for all  $x \geq 1$ , where  $W_{-1}$  is the negative branch of the Lambert  $W$  function, satisfying  $\bar{W}_{-1}(x) \approx x + \log x$  (Lemma 51). Given any sampling rule using the  $\text{GPE}_\eta(\epsilon)$ , using the GLR stopping rule as in Eq. (7) with the  $\text{GPE}_\eta(\epsilon)$  and the stopping threshold  $c(n, \epsilon, \delta) := c_1(n, \delta) + c_2(n, \epsilon)$  where*

$$\begin{aligned} c_1(n, \delta) &= \bar{W}_{-1}(\log(K\zeta(s)/\delta) + s \log(k_\eta(n))) + 3 - \log 2 - 3 + \log 2, \\ c_2(n, \epsilon) &= k_\eta(n) (\log(1 + 2\epsilon n/k_\eta(n)) + 1) \quad \text{with} \quad k_\eta(x) := 1 + \log_{1+\eta} x, \end{aligned} \quad (8)$$

yields a  $\delta$ -correct and  $\epsilon$ -global DP algorithm for all Bernoulli instances with a unique best arm.

The proof of Theorem 5 builds on novel concentration results of independent interest (Appendix F.2). Our explicit instance-independent upper bounds are pivotal to derive the stopping threshold in Eq. (8), which avoids the large instance-dependent constants used in the regret minimisation literature [13].

**Concentration Results.** First, we give tail bounds for the cumulative sum of i.i.d. Laplace observations (Lemma 15). We use Chernoff's method with the convex conjugate of the moment generating function of  $\text{Lap}(1/\epsilon)$ , hence improving on Azize et al. [13, Lemma 18] that approximates it. Second, we derive tail bounds for the sum between independent cumulative sums of  $t$  i.i.d. Bernoulli and  $n_t$  i.i.d. Laplace observations (Lemmas 17 and 18). They involve the *modified* signed divergences  $\tilde{d}_\epsilon^\pm$  that better capture the non-asymptotic tails behaviour, and are equivalent to  $d_\epsilon^\pm$  to an additive

term  $\Theta(\log(1 + 2\epsilon r_t)/r_t)$  where  $r_t := t/n_t$  (Lemma 29). Whenever  $r_t \rightarrow +\infty$ , we recover the same noise effect as adding only one  $\text{Lap}(1/\epsilon)$  observation. For  $x > 0$ , the exponential decrease of the probability of exceeding  $\mu + x$  (resp. being lower than  $\mu - x$ ) scales as  $t\tilde{d}_\epsilon^-(\mu + x, \mu, r_t)$  (resp.  $t\tilde{d}_\epsilon^+(\mu - x, \mu, r_t)$ ). The proof builds on fine-grained tail bounds of the sum of two independent random variables, i.e., we bound those probabilities by the maximal product between their respective survival functions (Lemma 9). While Azize et al. [13, Lemma 19] directly integrates their tail bounds, Lemma 11 can be used with any tail bounds. Third, we obtain time-uniform upper tail bounds for  $\tilde{N}_{n,a}\tilde{d}_\epsilon^\pm(\tilde{\mu}_{n,a}, \mu_a, \tilde{N}_{n,a}/k_{n,a})$  by exploiting the geometric-grid update of  $(\tilde{\mu}_n, \tilde{N}_n, k_n)$ .

**Threshold Scaling.** The threshold  $c_1$  in Eq. (8) ensures  $\delta$ -correctness of the *modified* GLR stopping rule, defined in Appendix F.1 with the *modified* transportation costs  $\tilde{W}_{\epsilon,a,b}$  and divergences  $\tilde{d}_\epsilon^\pm$  (Appendix F.1). Independent of  $\epsilon$ , it scales as  $\log(1/\delta) + \Theta(\log \log(1/\delta))$  when  $\delta \rightarrow 0$  and  $\Theta(\log \log(n))$  when  $n \rightarrow +\infty$ . The threshold  $c_2$  in Eq. (8) is an upper bound on  $\tilde{N}_{n,a}(d_\epsilon^\pm(\tilde{\mu}_{n,a}, \mu_a) - \tilde{d}_\epsilon^\pm(\tilde{\mu}_{n,a}, \mu_a, \tilde{N}_{n,a}/k_{n,a}))$  (Lemma 29) that scales as  $\Theta(\epsilon n)$  when  $\epsilon \rightarrow 0$  and as  $\Theta((\log n)^2)$  when  $n \rightarrow +\infty$ . Both  $c_1$  and  $c_2$  scales as  $\Theta(1/\log(1 + \eta))$  when  $\eta \rightarrow 0$ .

**Limitation.** As the threshold in Eq. (7) is the sum of per-arm thresholds, it scales as  $2 \log(1/\delta)$  when  $\delta \rightarrow 0$ , hence incurs a suboptimal factor 2 asymptotically. Obtaining a threshold in  $\log(1/\delta)$  is left for future work. It requires controlling the re-weighted sum of modified divergences  $\tilde{d}_\epsilon^\pm$ . Azize et al. [11, Theorem 4] has a suboptimal factor 2 for the same reason and incurs an additive factor  $\frac{1}{n\epsilon^2} \log(1/\delta)^2$  due to the separate control of the Laplace and the Bernoulli observations (based on sub-Gaussian concentration results). Azize et al. [12, Lemma 18] alleviates this factor 2 in their low privacy regime, yet it also pays  $\frac{1}{n\epsilon^2} \log(1/\delta)^2$ .

## 5 Top Two Sampling Rule

Equipped with a recommendation and stopping rules, we define a sampling rule using the  $\text{GPE}_\eta(\epsilon)$ . Within the fixed-confidence BAI literature, we adopt the Top Two approach [73, 71, 75, 50] that recently received increased scrutiny due to its good theoretical guarantees [49, 86, 52, 14], competitive empirical performance, and low computational cost. The Differentially Private Top Two (DP-TT) algorithm (Algorithm 1) uses the EB-TCI- $\beta$  sampling rule [50]. In Appendix I, we introduce the Track-and-Stop [38] and LUCB [55] sampling rules for fixed-confidence BAI under  $\epsilon$ -global DP.

After initialization, a Top Two sampling rule specifies four choices [48]: a leader arm  $B_n \in [K]$ , a challenger arm  $C_n \in [K] \setminus \{B_n\}$ , a target allocation  $\beta_n(B_n, C_n) \in [0, 1]$  and a mechanism to choose the next arm to sample from, i.e.,  $a_n \in \{B_n, C_n\}$  by using  $\beta_n(B_n, C_n)$ . The leader should select a good estimator of the best arm  $a^*$ . We use the empirical best (EB) leader that coincides with our recommendation rule, i.e.,  $B_n := \tilde{a}_n$ . The challenger should be a confusing alternative arm, for which the empirical evidence that the leader has a better mean is low. We use the TCI challenger [50] that penalizes oversampled challenger to foster implicit exploration, i.e.,  $C_n \in \arg \min_{a \neq B_n} \{W_{\epsilon, B_n, a}(\tilde{\mu}_n, N_n) + \log N_{n,a}\}$  where ties are broken uniformly at random. Crucially, we leverage our novel transportation costs  $(W_{\epsilon, B_n, a})_{a \neq B_n}$  featuring the signed divergences  $d_\epsilon^\pm$  that are evaluated empirically at  $(\tilde{\mu}_n, N_n)$ , see Eq. (3) and (4). The target should be chosen to balance the allocation between the leader and the challenger arms. Let  $\beta \in (0, 1)$ , e.g.,  $\beta = 1/2$ . We use a fixed  $\beta$ -design  $\beta_n(B_n, C_n) := \beta$ . The mechanism to choose the next arm to sample should enforce that this target is reached on average. We use  $K$  independent  $\beta$ -tracking procedures (one per leader), i.e.,  $a_n = B_n$  if  $N_{n, B_n}^{B_n} \leq \beta L_{n+1, B_n}$  and  $a_n = B_n$  otherwise, where  $N_{n,a}^a = \sum_{t \in [n-1]} \mathbb{1}((B_t, a_t) = (a, a))$  and  $L_{n,a} = \sum_{t \in [n-1]} \mathbb{1}(B_t = a)$ . Using Degenne et al. [29, Theorem 6] for each tracking procedure yields  $-1/2 \leq N_{n,a}^a - \beta L_{n,a} \leq 1$  for all  $a \in [K]$ .

**Computational and Memory Cost.** The  $\text{GPE}_\eta(\epsilon)$  sums the observations, and the recommendation and GLR stopping rules are updated when an arm is updated. Using the closed-form formula for  $W_{\epsilon,a,b}$  (Lemma 37), the per-iteration computational and global memory costs of DP-TT are  $\mathcal{O}(K)$ .

**Asymptotic Upper Bound on the Expected Sample Complexity.** Given a fixed target  $\beta$ , the empirical allocation of  $a^*$  converges towards  $\beta$ , that differs from  $w_{\epsilon, a^*}^*$ . At best, we can estimate the  $\beta$ -optimal allocation  $w_{\epsilon, \beta}^*(\nu)$ , i.e., maximizer of the inverse  $\beta$ -characteristic time  $T_{\epsilon, \beta}^*(\nu)^{-1}$

---

**Algorithm 1** Differentially Private Top Two (DP-TT) Algorithm.

---

1: **Input:** setting parameters  $(\epsilon, \delta) \in \mathbb{R}_+^* \times (0, 1)$ , algorithmic hyperparameters  $(\eta, \beta) \in \mathbb{R}_+^* \times (0, 1)$  and threshold  $c$ , e.g.,  $(\eta, \beta) = (1, 1/2)$  and  $c$  as in Eq. (8).  $(W_{\epsilon,a,b})_{(a,b) \in [K]}$  as in Eq. (4).  
2: **Output:** Stopping time  $\tau_{\epsilon,\delta}$ , recommendation  $\tilde{a}_{\tau_{\epsilon,\delta}}$  and pulling history  $(a_n)_{n < \tau_{\epsilon,\delta}}$ .  
3: **Initialization:** For all  $a \in [K]$ , pull arm  $a$ , observe  $X_{a,a} \sim \nu_a$  and draw  $Y_{1,a} \sim \text{Lap}(1/\epsilon)$ . Set  $n = K + 1$ . For all  $a \in [K]$ , set  $\tilde{S}_{n,a} = X_{a,a} + Y_{1,a}$ ,  $k_{n,a} = 1$ ,  $T_1(a) = n$ ,  $N_{n,a} = \tilde{N}_{n,a} = 1$ ,  $\tilde{\mu}_{n,a} = \tilde{S}_{n,a}/\tilde{N}_{n,a}$ ,  $L_{n,a} = 0$  and  $N_{n,a}^a = 0$ .  
4: **for**  $n \geq K + 1$  **do**  
5:     **if** there exists  $a \in [K]$  such that  $N_{n,a} \geq (1 + \eta)^{k_{n,a}}$  **then**  $\triangleright$  Per-arm geometric update grid  
6:         For this arm  $a$ , change phase  $k_{n,a} \leftarrow k_{n,a} + 1$ , and  $(T_{k_{n,a}}(a), \tilde{N}_{n,a}) = (n, N_{T_{k_{n,a}}(a),a})$ ;  
7:         Set  $\tilde{S}_{k_{n,a},a} = \sum_{t=T_{k_{n,a}-1}(a)}^{T_{k_{n,a}}(a)-1} X_{t,a} \mathbb{1}(a_t = a) + Y_{k_{n,a},a} + \tilde{S}_{k_{n,a}-1,a}$  with  $Y_{k_{n,a},a} \sim \text{Lap}(1/\epsilon)$ , and update the mean  $\tilde{\mu}_{n,a} = \tilde{S}_{k_{n,a},a}/\tilde{N}_{n,a}$ ;  
8:     **end if**  
9:     Set  $\tilde{a}_n \in \arg \max_{a \in [K]} [\tilde{\mu}_{n,a}]_0^1$ ;  $\triangleright$  Recommendation rule  
10:    **if**  $W_{\epsilon,\tilde{a}_n,a}(\tilde{\mu}_n, \tilde{N}_n) > \sum_{b \in \{\tilde{a}_n,a\}} c(\tilde{N}_{n,b}, \epsilon, \delta)$  for all  $a \neq \tilde{a}_n$  **then**  $\triangleright$  GLR stopping rule  
11:        **return**  $(n, \tilde{a}_n, (a_t)_{t < n})$ .  
12:    **end if**  
13:    Set  $B_n = \tilde{a}_n$  and  $C_n \in \arg \min_{a \neq B_n} \{W_{\epsilon,B_n,a}(\tilde{\mu}_n, N_n) + \log N_{n,a}\}$ ;  $\triangleright$  EB-TCI  
14:    Set  $a_n = B_n$  if  $N_{n,B_n}^{B_n} \leq \beta L_{n+1,B_n}$ , and  $a_n = C_n$  otherwise;  $\triangleright$   $\beta$ -tracking  
15:    Pull  $a_n$ , observe and store  $X_{n,a_n} \sim \nu_{a_n}$ ;  
16:    Update  $(N_{n+1,a_n}, L_{n+1,B_n}, N_{n+1,B_n}^{B_n}) = (N_{n,a_n}, L_{n,B_n}, N_{n,B_n}^{B_n}) + (1, 1, \mathbb{1}(B_n = a_n))$ ;  
17: **end for**

---

312 defined as in Eq. (5) with the constraint  $w_{a^*} = \beta$ . While being only nearly asymptotic optimal, i.e.,  
313  $T_\epsilon^*(\nu) = \min_{\beta \in (0,1)} T_{\epsilon,\beta}^*(\nu)$ , it satisfies  $T_{\epsilon,1/2}^*(\nu) \leq 2T_\epsilon^*(\nu)$  (Lemma 43).

314 DP-TT is  $\epsilon$ -global DP,  $\delta$ -correct and matches  $T_\epsilon^*(\nu)$  to a small constant, for any privacy budget  $\epsilon$ .  
315 The proof (Appendix H) builds on the unified analysis of Jourdan et al. [50] and relies heavily on the  
316 derived regularity properties for  $d_\epsilon^\pm$ ,  $(W_{\epsilon,a,b})_{(a,b)}$ ,  $T_{\epsilon,\beta}^*(\nu)$  and  $w_{\epsilon,\beta}^*(\nu)$  (Appendix G).

317 **Theorem 6.** Let  $(\epsilon, \delta, \eta, \beta) \in \mathbb{R}_+^* \times (0, 1) \times \mathbb{R}_+^* \times (0, 1)$  and  $c$  as in Eq. (8). The DP-TT algorithm is  
318  $\epsilon$ -global DP,  $\delta$ -correct and satisfies that, for any Bernoulli instance  $\nu$  with distinct means  $\mu \in (0, 1)^K$ ,

$$\limsup_{\delta \rightarrow 0} \frac{\mathbb{E}_{\nu\pi} [\tau_{\epsilon,\delta}]}{\log(1/\delta)} \leq 2(1 + \eta)T_{\epsilon,\beta}^*(\nu).$$

319 For  $(\eta, \beta) = (1, 1/2)$ , the asymptotic upper bound is  $4T_{\epsilon,1/2}^*(\nu) \leq 8T_\epsilon^*(\nu)$ . For any privacy budget  
320  $\epsilon$ , we reduced the gap between known lower and upper bounds for fixed-confidence BAI under  
321  $\epsilon$ -global DP to a constant lower than 8, hence closing the open problem in Azize et al. [12]. A  
322 discussion on how to improve this constant is deferred to Appendix C.2.

323 **Comparison with Azize et al. [11, 12].** AdaP-TT and AdaP-TT\* use the DAF( $\epsilon$ ) estimator, GLR-  
324 inspired recommendation/stopping rules and the TTUCB [49] sampling rule (i.e., UCB-TC- $\beta$  [48]),  
325 all based on arm-dependent doubling, forgetting and unclipped estimators. While AdaP-TT relies  
326 on the non-private *Gaussian* transportation costs, AdaP-TT\* accounts for a high privacy regime by  
327 clipping the mean gap, i.e.,  $(\mu_a - \mu_b)_+ \min\{3\epsilon, (\mu_a - \mu_b)_+\}$  instead of  $(\mu_a - \mu_b)^2$ . The AdaP-TT  
328 and AdaP-TT\* algorithms are  $\epsilon$ -global DP and  $\delta$ -correct. The sample complexity of AdaP-TT only  
329 matches the *high privacy* lower bound for instances where the means are of similar order. AdaP-TT\*  
330 improves on AdaP-TT by matching the *high privacy* lower bound for all instances with distinct means.  
331 However, both AdaP-TT and AdaP-TT\* fail to match the lower bound beyond the high-privacy  
332 regime, due to the use of non-adapted transportation costs. In contrast, DP-TT uses the  $d_\epsilon$ -inspired  
333 transportation costs, matching the lower bound up to a small constant for all values of  $\epsilon$ .

334 **Comparison with Sajed and Sheffet [74].** DP-SE [74] is an  $\epsilon$ -global DP version of the Successive  
335 Elimination algorithm introduced for the regret minimisation setting, modified by Azize et al. [11]  
336 into a  $\epsilon$ -global and  $\delta$ -correct BAI algorithm. Compared to DP-TT, DP-SE is less adaptive and



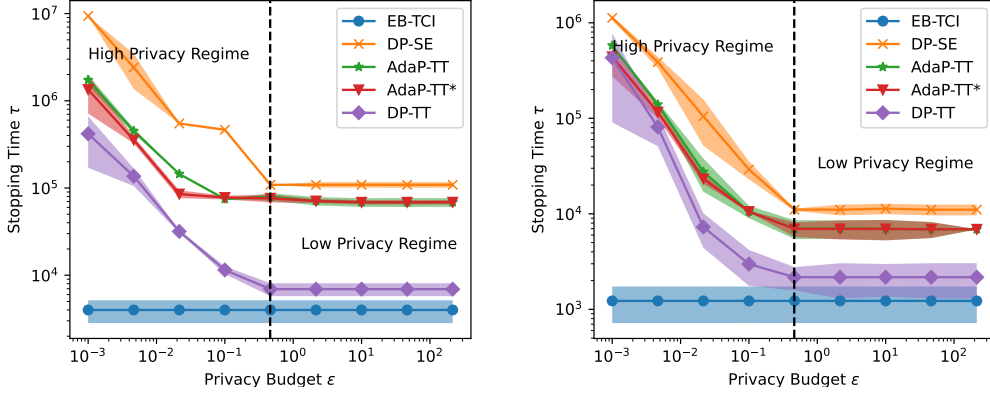


Figure 1: Empirical stopping time  $\tau_{\epsilon, \delta}$  (mean  $\pm 2$  std) for  $\delta = 10^{-2}$  with respect to the privacy budget  $\epsilon$  on Bernoulli instances (a)  $\mu_1$  and (b)  $\mu_2$ . The vertical line separates the two privacy regimes.

not anytime, since it relies on uniform sampling within each phase and the phase length depends explicitly on the risk  $\delta$ . The high probability upper bound on the sample complexity scales as  $\mathcal{O}(\sum_{a \neq a^*} (\Delta_a \min\{\epsilon, \Delta_a\})^{-1})$  with  $\Delta_a = \mu_{a^*} - \mu_a$ . This matches the lower bound when  $\epsilon \rightarrow 0$  (to a constant), but fails to recover the sample-complexity lower bound beyond this regime.

## 6 Experiments

The empirical performance of DP-TT with  $(\eta, \beta) = (1, 1/2)$  is compared to AdaP-TT [11], AdaP-TT\* [12], and DP-SE [74] on different Bernoulli instances for varying privacy budget. The first instance has means  $\mu_1 = (0.95, 0.9, 0.9, 0.9, 0.5)$  and the second instance has means  $\mu_2 = (0.75, 0.7, 0.7, 0.7, 0.7)$ . As a benchmark, we also compare to the non-private EB-TCI- $\beta$  algorithm with  $\beta = 1/2$ . For  $\delta = 10^{-2}$ , we run each algorithm 1000 times, and plot the averaged empirical stopping times in Figure 1. Additional experiments are available in Appendix J.

Figure 1 shows that DP-TT outperforms all the other  $\delta$ -correct and  $\epsilon$ -global DP BAI algorithms, for different values of  $\epsilon$  and in all the instances tested. The empirical performance of DP-TT demonstrates two regimes. A high-privacy regime, where the stopping time depends on the privacy budget  $\epsilon$ , and a low privacy regime, where the performance of DP-TT is independent of  $\epsilon$ , and requires twice the number of samples used by the non-private EB-TCI- $\beta$ .

## 7 Conclusion

Motivated by the privacy requirements of sensitive applications of BAI, we address the problem of fixed-confidence BAI under  $\epsilon$ -global DP. We narrow the gap between the lower and upper bounds on the expected sample complexity to a multiplicative constant smaller than 8, for all  $\epsilon$  values. Our novel lower bound incorporates  $d_\epsilon$  an information-theoretic quantity smoothly balancing KL divergence and TV distance, scaled by  $\epsilon$ . We design a private, arm-dependent geometric grid estimator *without forgetting* and a GLR stopping rule based on the  $d_\epsilon$ -transportation costs, whose correctness requires novel concentration results for Laplace and mixed distributions. Finally, we proposed a Top Two sampling rule that achieves an asymptotic upper bound matching our lower bound to a small constant.

We detailed research directions to further reduce the constant gap between the lower and upper bounds, by improving both the calibration of the stopping threshold and the analysis of the sampling rule. The most exciting direction for future work is to extend our results to other classes of distributions (e.g., Gaussian or bounded distributions), structured settings (e.g., linear or unimodal), or other identification problems (e.g., approximate BAI or Good Arm Identification). Another interesting research direction is to extend the proposed technique to other variants of pure DP (e.g.,  $(\epsilon, \delta)$ -DP or Rényi DP [67]) or other trust models (e.g., shuffle DP [26, 39]).

## References

- [1] Naman Agarwal and Karan Singh. The price of differential privacy for online learning. In *International Conference on Machine Learning*, pages 32–40. PMLR, 2017.
- [2] Shubhada Agrawal, Sandeep Juneja, and Peter W. Glynn. Optimal  $\delta$ -correct best-arm selection for heavy-tailed distributions. In *Algorithmic Learning Theory (ALT)*, 2020.
- [3] J-Y. Audibert and S. Bubeck. Regret Bounds and Minimax Policies under Partial Monitoring. *Journal of Machine Learning Research*, 2010.
- [4] J-Y. Audibert, S. Bubeck, and R. Munos. Best Arm Identification in Multi-armed Bandits. In *Conference on Learning Theory*, 2010.
- [5] Jean-Yves Audibert and Sébastien Bubeck. Regret bounds and minimax policies under partial monitoring. *The Journal of Machine Learning Research*, 11:2785–2836, 2010.
- [6] Peter Auer, Nicolo Cesa-Bianchi, and Paul Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 47(2-3):235–256, 2002.
- [7] Peter Auer, Chao-Kai Chiang, Ronald Ortner, and Madalina Drugan. Pareto front identification from stochastic bandit feedback. In *Artificial intelligence and statistics*, pages 939–947. PMLR, 2016.
- [8] Maryam Aziz, Emilie Kaufmann, and Marie-Karelle Riviere. On multi-armed bandit designs for dose-finding clinical trials. *The Journal of Machine Learning Research*, 22(1):686–723, 2021.
- [9] Achraf Azize and Debabrota Basu. When privacy meets partial information: A refined analysis of differentially private bandits. *Advances in Neural Information Processing Systems*, 35: 32199–32210, 2022.
- [10] Achraf Azize and Debabrota Basu. Concentrated differential privacy for bandits. In *2nd IEEE Conference on Secure and Trustworthy Machine Learning*, 2024.
- [11] Achraf Azize, Marc Jourdan, Aymen Al Marjani, and Debabrota Basu. On the complexity of differentially private best-arm identification with fixed confidence. *Thirty-Seventh Conference on Neural Information Processing Systems*, 2023.
- [12] Achraf Azize, Marc Jourdan, Aymen Al Marjani, and Debabrota Basu. Differentially private best-arm identification. *arXiv preprint arXiv:2406.06408*, 2024.
- [13] Achraf Azize, Yulian Wu, Junya Honda, Francesco Orabona, Shinji Ito, and Debabrota Basu. Optimal regret of bernoulli bandits under global differential privacy. *arXiv preprint arXiv:2505.05613*, 2025.
- [14] Agniv Bandyopadhyay, Sandeep Juneja, and Shubhada Agrawal. Optimal top-two method for best arm identification and fluid analysis. *Thirty-Eighth Conference on Neural Information Processing Systems*, 2024.
- [15] Debabrota Basu, Christos Dimitrakakis, and Aristide Tossou. Differential privacy for multi-armed bandits: What is it and what is its cost? *arXiv preprint arXiv:1905.12298*, 2019.
- [16] Claude Berge. *Topological Spaces: including a treatment of multi-valued functions, vector spaces, and convexity*. Courier Corporation, 1997.
- [17] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, 2013.
- [18] Sébastien Bubeck, Nicolo Cesa-Bianchi, et al. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends® in Machine Learning*, 5(1):1–122, 2012.
- [19] Alexandra Carpentier and Andrea Locatelli. Tight (lower) bounds for the fixed budget best arm identification bandit problem. In *Conference on Learning Theory*, pages 590–604. PMLR, 2016.

- 416 [20] T.-H. Hubert Chan, Elaine Shi, and Dawn Song. Private and continual release of statistics. *ACM*  
417 *Trans. Inf. Syst. Secur.*, 14(3), nov 2011. ISSN 1094-9224. doi: 10.1145/2043621.2043626.  
418 URL <https://doi.org/10.1145/2043621.2043626>.
- 419 [21] Lijie Chen, Anupam Gupta, Jian Li, Mingda Qiao, and Ruosong Wang. Nearly optimal  
420 sampling algorithms for combinatorial pure exploration. In *Proceedings of the 30th Conference*  
421 *on Learning Theory (COLT)*, 2017.
- 422 [22] Shouyuan Chen, Tian Lin, Irwin King, Michael R Lyu, and Wei Chen. Combinatorial pure  
423 exploration of multi-armed bandits. *Advances in neural information processing systems*, 27,  
424 2014.
- 425 [23] Wei Chen, Yajun Wang, and Yang Yuan. Combinatorial multi-armed bandit: General framework  
426 and applications. In *International Conference on Machine Learning*, 2013.
- 427 [24] Zhirui Chen, P. N. Karthik, Yeow Meng Chee, and Vincent Tan. Fixed-budget differentially  
428 private best arm identification. In *The Twelfth International Conference on Learning Representations*, 2024.  
429
- 430 [25] Herman Chernoff. Sequential Design of Experiments. *The Annals of Mathematical Statistics*,  
431 30:755–770, 1959.
- 432 [26] Albert Cheu. Differential privacy in the shuffle model: A survey of separations. *arXiv preprint*  
433 *arXiv:2107.11839*, 2021.
- 434 [27] Sayak Ray Chowdhury and Xingyu Zhou. Shuffle private linear contextual bandits. In *International*  
435 *Conference on Machine Learning*, pages 3984–4009. PMLR, 2022.
- 436 [28] R. Combes and A. Proutière. Unimodal bandits: Regret lower bounds and optimal algorithms.  
437 In *International Conference on Machine Learning (ICML)*, 2014.
- 438 [29] Rémy Degenne, Han Shao, and Wouter Koolen. Structure adaptive algorithms for stochastic  
439 bandits. In *International Conference on Machine Learning*, pages 2443–2452. PMLR, 2020.
- 440 [30] Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Foundations*  
441 *and Trends® in Theoretical Computer Science*, 9(3–4):211–407, 2014.
- 442 [31] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to  
443 sensitivity in private data analysis. In *Proceedings of the Third Conference on Theory of*  
444 *Cryptography, TCC’06*, pages 265–284, Berlin, Heidelberg, 2006. Springer-Verlag.
- 445 [32] Cynthia Dwork, Moni Naor, Toniann Pitassi, and Guy N Rothblum. Differential privacy under  
446 continual observation. In *ACM symposium on Theory of computing*, pages 715–724. ACM,  
447 2010.
- 448 [33] Eyal Even-Dar, Shie Mannor, and Yishay Mansour. PAC bounds for multi-armed bandit and  
449 Markov decision processes. In *Conference on Computational Learning Theory, COLT ’02*,  
450 page 255–270, Berlin, Heidelberg, 2002. Springer-Verlag. ISBN 354043836X.
- 451 [34] Eyal Even-Dar, Shie Mannor, Yishay Mansour, and Sridhar Mahadevan. Action elimination and  
452 stopping conditions for the multi-armed bandit and reinforcement learning problems. *Journal*  
453 *of machine learning research*, 7(6), 2006.
- 454 [35] S. Filippi, O. Cappé, A. Garivier, and C. Szepesvári. Parametric Bandits : The Generalized  
455 Linear case. In *Advances in Neural Information Processing Systems*, 2010.
- 456 [36] Victor Gabillon, Mohammad Ghavamzadeh, and Alessandro Lazaric. Best arm identification:  
457 A unified approach to fixed budget and fixed confidence. *Advances in Neural Information*  
458 *Processing Systems*, 25, 2012.
- 459 [37] Evrard Garcelon, Kamalika Chaudhuri, Vianney Perchet, and Matteo Pirota. Privacy amplification  
460 via shuffling for linear contextual bandits. In *International Conference on Algorithmic*  
461 *Learning Theory*, pages 381–407. PMLR, 2022.

- [38] Aurélien Garivier and Emilie Kaufmann. Optimal best arm identification with fixed confidence. In *Conference on Learning Theory*, pages 998–1027. PMLR, 2016.
- [39] Antonious M Girgis, Deepesh Data, Suhas Diggavi, Ananda Theertha Suresh, and Peter Kairouz. On the renyi differential privacy of the shuffle model. In *ACM SIGSAC Conference on Computer and Communications Security*, pages 2321–2341, 2021.
- [40] Yuxuan Han, Zhipeng Liang, Yang Wang, and Jiheng Zhang. Generalized linear bandits with local differential privacy. *Advances in Neural Information Processing Systems*, 34:26511–26522, 2021.
- [41] Osama Hanna, Antonious M Girgis, Christina Fragouli, and Suhas Diggavi. Differentially private stochastic linear bandits:(almost) for free. *IEEE Journal on Selected Areas in Information Theory*, 2024.
- [42] Junya Honda and Akimichi Takemura. Non-asymptotic analysis of a new bandit algorithm for semi-bounded rewards. *Journal of Machine Learning Research (JMLR)*, 16:3721–3756, 2015.
- [43] Bingshan Hu and Nidhi Hegde. Near-optimal thompson sampling-based algorithms for differentially private stochastic bandits. In *Uncertainty in Artificial Intelligence*, pages 844–852. PMLR, 2022.
- [44] Bingshan Hu, Zhiming Huang, and Nishant A. Mehta. Optimal algorithms for private online learning in a stochastic environment, 2021. URL <https://arxiv.org/abs/2102.07929>.
- [45] Bingshan Hu, Zhiming Huang, Tianyue H Zhang, Mathias Lécuyer, and Nidhi Hegde. Connecting thompson sampling and ucb: Towards more efficient trade-offs between privacy and regret. *arXiv preprint arXiv:2505.02383*, 2025.
- [46] K. Jamieson, S. Katariya, A. Deshpande, and R. Nowak. Sparse Dueling Bandits. In *Proceedings of the 18th Conference on Artificial Intelligence and Statistics*, 2015.
- [47] Kevin Jamieson and Robert Nowak. Best-arm identification algorithms for multi-armed bandits in the fixed confidence setting. In *Conference on Information Sciences and Systems (CISS)*, pages 1–6. IEEE, 2014.
- [48] Marc Jourdan. *Solving Pure Exploration Problems with the Top Two Approach*. PhD thesis, Université de Lille, 2024.
- [49] Marc Jourdan and Rémy Degenne. Non-asymptotic analysis of a ucb-based top two algorithm. *Advances in Neural Information Processing Systems*, 36, 2024.
- [50] Marc Jourdan, Rémy Degenne, Dorian Baudry, Rianne de Heide, and Emilie Kaufmann. Top two algorithms revisited. *Advances in Neural Information Processing Systems*, 35:26791–26803, 2022.
- [51] Marc Jourdan, Rémy Degenne, and Emilie Kaufmann. Dealing with unknown variances in best-arm identification. *International Conference on Algorithmic Learning Theory*, 2023.
- [52] Marc Jourdan, Rémy Degenne, and Emilie Kaufmann. An  $\varepsilon$ -best-arm identification algorithm for fixed-confidence and beyond. *Advances in Neural Information Processing Systems*, 36, 2024.
- [53] Kwang-Sung Jun, Lalit Jain, Houssam Nassif, and Blake Mason. Improved Confidence Bounds for the Linear Logistic Model and Applications to Bandits. In *International Conference on Machine Learning*, 2021.
- [54] Dionysios S Kalogerias, Kontantinos E Nikolakakis, Anand D Sarwate, and Or Sheffet. Quantile multi-armed bandits: Optimal best-arm identification and a differentially private scheme. *IEEE Journal on Selected Areas in Information Theory*, 2(2):534–548, 2021.
- [55] Shivaram Kalyanakrishnan, Ambuj Tewari, Peter Auer, and Peter Stone. Pac subset selection in stochastic multi-armed bandits. In *International Conference on Machine Learning*, volume 12, pages 655–662, 2012.

- [56] Julian Katz-Samuels and Clayton Scott. Top Feasible Arm Identification. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2019.
- [57] Tomas Kocak and Aurelien Garivier. Epsilon best arm identification in spectral bandits. In *International Joint Conference on Artificial Intelligence (IJCAI)*, pages 2636–2642, 2021.
- [58] Tze Leung Lai and Herbert Robbins. Asymptotically efficient adaptive allocation rules. *Advances in applied mathematics*, 6(1):4–22, 1985.
- [59] Tor Lattimore and Csaba Szepesvari. *Bandit algorithms*. Cambridge University Press, 2020.
- [60] Fengjiao Li, Xingyu Zhou, and Bo Ji. Differentially private linear bandits with partial distributed feedback. In *2022 20th International Symposium on Modeling and Optimization in Mobile, Ad hoc, and Wireless Networks (WiOpt)*, pages 41–48. IEEE, 2022.
- [61] Lisha Li, Kevin Jamieson, Giulia DeSalvo, Afshin Rostamizadeh, and Ameet Talwalkar. Hyperband: A novel bandit-based approach to hyperparameter optimization. *The Journal of Machine Learning Research*, 18(1):6765–6816, 2017.
- [62] Pieter JK Libin, Timothy Verstraeten, Diederik M Roijers, Jelena Grujic, Kristof Theys, Philippe Lemey, and Ann Nowe. Bayesian best-arm identification for selecting influenza mitigation strategies. In *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2018*, 2019.
- [63] Simon Lindstahl, Alexandre Proutiere, and Andreas Johnsson. Measurement-based admission control in sliced networks: A best arm identification approach. In *GLOBECOM 2022-2022 IEEE Global Communications Conference*, pages 1484–1490. IEEE, 2022.
- [64] David E Losada, David Elswiler, Morgan Harvey, and Christoph Trattner. A day at the races: using best arm identification algorithms to reduce the cost of information retrieval user studies. *Applied Intelligence*, 52(5):5617–5632, 2022.
- [65] S. Magureanu, R. Combes, and A. Proutiere. Lipschitz Bandits: Regret lower bounds and optimal algorithms. In *Proceedings on the 27th Conference On Learning Theory*, 2014.
- [66] Shie Mannor and John N Tsitsiklis. The sample complexity of exploration in the multi-armed bandit problem. *Journal of Machine Learning Research*, 5(Jun):623–648, 2004.
- [67] Ilya Mironov. Renyi differential privacy. In *2017 IEEE 30th computer security foundations symposium (CSF)*, pages 263–275. IEEE, 2017.
- [68] Nikita Mishra and Abhradeep Thakurta. (Nearly) optimal differentially private stochastic multi-arm bandits. In *Conference on Uncertainty in Artificial Intelligence*, 2015.
- [69] Seth Neel and Aaron Roth. Mitigating bias in adaptive data gathering via differential privacy. In *International Conference on Machine Learning*, pages 3720–3729. PMLR, 2018.
- [70] Nikola Pavlovic, Sudeep Salgia, and Qing Zhao. Differentially private kernelized contextual bandits. In *The 28th International Conference on Artificial Intelligence and Statistics*, 2025.
- [71] Chao Qin, Diego Klabjan, and Daniel Russo. Improving the expected improvement algorithm. *Advances in Neural Information Processing Systems*, 30, 2017.
- [72] H. Robbins. Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, 58(5):527–535, 1952.
- [73] Daniel Russo. Simple bayesian algorithms for best arm identification. In *Conference on Learning Theory*, pages 1417–1418. PMLR, 2016.
- [74] Touqir Sajed and Or Sheffet. An optimal private stochastic-mab algorithm based on optimal private stopping rule. In *International Conference on Machine Learning*, pages 5579–5588. PMLR, 2019.

- 553 [75] Xuedong Shang, Rianne Heide, Pierre Menard, Emilie Kaufmann, and Michal Valko. Fixed-  
554 confidence guarantees for bayesian best-arm identification. In *International Conference on*  
555 *Artificial Intelligence and Statistics*, pages 1823–1832. PMLR, 2020.
- 556 [76] Roshan Shariff and Or Sheffet. Differentially private contextual linear bandits. In *Advances in*  
557 *Neural Information Processing Systems*, pages 4296–4306, 2018.
- 558 [77] Marta Soare, Alessandro Lazaric, and Rémi Munos. Best-arm identification in linear bandits.  
559 *Advances in Neural Information Processing Systems*, 27, 2014.
- 560 [78] Rangarajan K Sundaram. *A first course in optimization theory*. Cambridge university press,  
561 1996.
- 562 [79] Jay Tenenbaum, Haim Kaplan, Yishay Mansour, and Uri Stemmer. Differentially private multi-  
563 armed bandits in the shuffle model. *Advances in Neural Information Processing Systems*, 34:  
564 24956–24967, 2021.
- 565 [80] Abhradeep Guha Thakurta and Adam Smith. (Nearly) optimal algorithms for private online  
566 learning in full-information and bandit settings. *Advances in Neural Information Processing*  
567 *Systems*, 26, 2013.
- 568 [81] William R Thompson. On the likelihood that one unknown probability exceeds another in view  
569 of the evidence of two samples. *Biometrika*, 25(3-4):285–294, 1933.
- 570 [82] Aristide CY Tossou and Christos Dimitrakakis. Achieving privacy in the adversarial multi-armed  
571 bandit. In *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.
- 572 [83] Katherine Tucker, Janice Branson, Maria Dilleen, Sally Hollis, Paul Loughlin, Mark J Nixon,  
573 and Zoë Williams. Protecting patient privacy when sharing patient-level data from clinical trials.  
574 *BMC medical research methodology*, 16(1):5–14, 2016.
- 575 [84] J. Ville. *Étude Critique de la Notion de Collectif*. Collection des monographies des probabilités.  
576 Gauthier-Villars, 1939.
- 577 [85] Siwei Wang and Jun Zhu. Optimal learning policies for differential privacy in multi-armed  
578 bandits. *Journal of Machine Learning Research*, 25(314):1–52, 2024.
- 579 [86] Wei You, Chao Qin, Zihao Wang, and Shuoguang Yang. Information-directed selection for  
580 top-two algorithms. In *Conference on Learning Theory*, pages 2850–2851. PMLR, 2023.
- 581 [87] Wanrong Zhang, Yajun Mei, and Rachel Cummings. Private sequential hypothesis testing for  
582 statisticians: Privacy, error rates, and sample size. In *International Conference on Artificial*  
583 *Intelligence and Statistics*, pages 11356–11373. PMLR, 2022.
- 584 [88] Yao Zhao, Connor James Stephens, Csaba Szepesvári, and Kwang-Sung Jun. Revisiting simple  
585 regret: Fast rates for returning a good arm. In *40th International Conference on Machine*  
586 *Learning (ICML)*, 2023.
- 587 [89] Kai Zheng, Tianle Cai, Weiran Huang, Zhenguo Li, and Liwei Wang. Locally differentially  
588 private (contextual) bandits learning. In *Advances in Neural Information Processing Systems*,  
589 volume 33, pages 12300–12310, 2020.
- 590 [90] Yuan Zhou, Xi Chen, and Jian Li. Optimal pac multiple arm identification with applications  
591 to crowdsourcing. In *International Conference on Machine Learning*, pages 217–225. PMLR,  
592 2014.

## NeurIPS Paper Checklist

### 1. Claims

Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

Answer: [\[Yes\]](#)

Justification: The claims made in the abstract and introduction reflect the paper's contributions and scope. Those claims are substantiated with details in the Sections 3, 4, 5 and 6 of the main paper. They are proven theoretically in Appendices D, F, G, H, I and J.

Guidelines:

- The answer NA means that the abstract and introduction do not include the claims made in the paper.
- The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
- The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
- It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

### 2. Limitations

Question: Does the paper discuss the limitations of the work performed by the authors?

Answer: [\[Yes\]](#)

Justification: The limitations of Theorems 5 and 6 are detailed in Sections 4 and Appendix C.2. The low computational and memory costs of the DP-TT algorithm is discussed in Sections 4 and 5. Our work addresses the  $\epsilon$ -global Differential Privacy constraint for fixed-confidence Best Arm Identification. Our algorithm can be deployed in data-sensitive applications, such as adaptive clinical trials or user studies.

Guidelines:

- The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
- The authors are encouraged to create a separate "Limitations" section in their paper.
- The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
- The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
- The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
- The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
- If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
- While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

### 3. Theory assumptions and proofs

Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: Theorem 2 is proven in Appendix D. Theorem 3 is based on numerous results proven in Appendix G. Lemma 4 is proven in Appendix E. Theorem 5 is proven in Appendix F. Theorem 6 is proven in Appendix H. All the other results mentioned in the main paper are proven theoretically in Appendices D, F, G, H and I.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

### 4. Experimental result reproducibility

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: Implementation details on our experiments are given in Section 6 and Appendix J. The DP-TT algorithm is summarized in Algorithm 1, and includes default choices of hyper-parameters. Our experiments are reproducible and the code is provided as supplementary material.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general, releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  - (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  - (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  - (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).



- (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

## 5. Open access to data and code

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [Yes]

Justification: The code for our experiments is provided as supplementary material, and we provide sufficient instructions to faithfully reproduce our experimental results. As we rely on synthetic experiments on Bernoulli distributions, there is no dataset per se.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so “No” is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

## 6. Experimental setting/details

Question: Does the paper specify all the training and test details (e.g., data splits, hyper-parameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [Yes]

Justification: The code for our experiments is provided as supplementary material. We provide sufficient implementation details in Section 6 and Appendix J. The DP-TT algorithm is summarized in Algorithm 1, and includes default choices of hyper-parameters.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

## 7. Experiment statistical significance

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [Yes]

Justification: The figures in Section 6 and Appendix J report the empirical means over 1000 runs for each algorithm. The 2-sigma error bars are based on the standard deviation of the means over those runs.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

## 8. Experiments compute resources

Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

Answer: [Yes]

Justification: Appendix J contains detailed information as regards the computational resources used for our experiments. The specific institutional cluster is omitted for the sake of anonymity. Overall, our experiments are relatively computationally inexpensive, and they could be run on a professional laptop.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
- The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
- The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

## 9. Code of ethics

Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics <https://neurips.cc/public/EthicsGuidelines>?

Answer: [Yes]

Justification: After reading it, we confirm that the research conducted in our paper complies with the NeurIPS Code of Ethics.

Guidelines:

- The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.

- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

## 10. Broader impacts

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [Yes]

Justification: Our work addresses the  $\epsilon$ -global Differential Privacy constraint for fixed-confidence Best Arm Identification. Our algorithm can be deployed in data-sensitive applications, such as adaptive clinical trials or user studies. Therefore, our paper can have a significant positive societal impact.

Guidelines:

- The answer NA means that there is no societal impact of the work performed.
- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

## 11. Safeguards

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification: To the best of our understanding, our paper doesn't necessitate safeguards. Our experiments use synthetic Bernoulli data and our algorithm satisfy the  $\epsilon$ -global Differential Privacy constraint.

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

## 12. Licenses for existing assets

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [NA]

Justification: Our paper does not rely on existing assets. Our experiments use synthetic Bernoulli data and our algorithm was implemented from scratch.

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

### 13. New assets

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [NA]

Justification: Our paper does not introduce new assets. The code for our experiments is provided as supplementary material, yet it does not constitute an asset per se.

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

### 14. Crowdsourcing and research with human subjects

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [NA]

Justification: Our paper does not involve crowdsourcing nor research with human subjects.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

909 **15. Institutional review board (IRB) approvals or equivalent for research with human**  
910 **subjects**

911 Question: Does the paper describe potential risks incurred by study participants, whether  
912 such risks were disclosed to the subjects, and whether Institutional Review Board (IRB)  
913 approvals (or an equivalent approval/review based on the requirements of your country or  
914 institution) were obtained?

915 Answer: [NA]

916 Justification: Our paper does not involve crowdsourcing nor research with human subjects.

917 Guidelines:

- 918 • The answer NA means that the paper does not involve crowdsourcing nor research with  
919 human subjects.
- 920 • Depending on the country in which research is conducted, IRB approval (or equivalent)  
921 may be required for any human subjects research. If you obtained IRB approval, you  
922 should clearly state this in the paper.
- 923 • We recognize that the procedures for this may vary significantly between institutions  
924 and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the  
925 guidelines for their institution.
- 926 • For initial submissions, do not include any information that would break anonymity (if  
927 applicable), such as the institution conducting the review.

928 **16. Declaration of LLM usage**

929 Question: Does the paper describe the usage of LLMs if it is an important, original, or  
930 non-standard component of the core methods in this research? Note that if the LLM is used  
931 only for writing, editing, or formatting purposes and does not impact the core methodology,  
932 scientific rigorousness, or originality of the research, declaration is not required.

933 Answer: [NA]

934 Justification: Our theoretical and empirical contributions do not involve LLMs.

935 Guidelines:

- 936 • The answer NA means that the core method development in this research does not  
937 involve LLMs as any important, original, or non-standard components.
- 938 • Please refer to our LLM policy (<https://neurips.cc/Conferences/2025/LLM>)  
939 for what should or should not be described.

## A Outline

The appendices are organized as follows:

- Notation are summarized in Appendix B.
- A detailed related work and a discussion on the limitations of Theorem 6 are given in Appendix C.
- The lower bound on the expected sample complexity under  $\epsilon$ -global DP (Theorem 2) is proven in Appendix D.
- The proof of Lemma 4 is given in Appendix E.
- The proof of our concentration results are detailed in Appendix F. In particular, this includes the proof of Theorem 5.
- Appendix G gathers key properties on the (resp. modified) divergence  $d_\epsilon^\pm$  (resp.  $\widetilde{d}_\epsilon^\pm$ ), the (resp. modified) transportation costs  $W_{\epsilon,a,b}$  (resp.  $\widetilde{W}_{\epsilon,a,b}$ ) and (resp.  $\beta$ -)characteristic times  $T_\epsilon^*(\nu)$  (resp.  $T_{\epsilon,\beta}^*(\nu)$ ) and their (resp.  $\beta$ -)optimal allocation  $w_\epsilon^*(\nu)$  (resp.  $w_{\epsilon,\beta}^*(\nu)$ ). In particular, this includes the proof of Theorem 3 based on Lemmas 42 and 46.
- The proof of the upper bound on the asymptotic expected sample complexity of DP-TT (Theorem 6) is given in Appendix H.
- In Appendix I, we propose variants of algorithms to tackle  $\epsilon$ -global DP BAI. We aim at providing several choices for the interested practitioners.
- Implementation details and additional experiments are presented in Appendix J.

Table 1: Notation for the setting.

Notation	Type	Description
$K$	$\mathbb{N}$	Number of arms
$\mathcal{F}$	$\subseteq \mathcal{P}([0, 1])$	Class of Bernoulli distributions
$\nu_a$	$\mathcal{F}$	Bernoulli distribution of arm $a \in [K]$
$\nu$	$\mathcal{F}^K$	Vector of Bernoulli distributions, $\nu := (\nu_a)_{a \in [K]}$
$\mu_a$	$(0, 1)$	Mean of arm $a \in [K]$
$\mu$	$(0, 1)^K$	Vector of means, $\mu := (\mu_a)_{a \in [K]}$
$a^*(\mu), a^*(\nu)$	$\subseteq [K]$	Set of best arms, $a^*(\nu) = a^*(\mu) := \arg \max_{a \in [K]} \mu_a$
$a^*$	$[K]$	Unique best arm, i.e., $a^*(\mu) = \{a^*\}$
$\epsilon$	$\mathbb{R}_+^*$	Privacy budget for $\epsilon$ -global DP
$\delta$	$(0, 1)$	Risk for $\delta$ -correctness
$\text{Alt}(\nu)$	$\subseteq \mathcal{F}^K$	Alternative instances with different best arms

## B Notation

We recall some commonly used notation: the set of integers  $[n] := \{1, \dots, n\}$ , the complement  $X^c$  and interior  $\overset{\circ}{X}$  of a set  $X$ , the indicator function  $\mathbb{1}(X)$  of an event, the probability  $\mathbb{P}_{\nu\pi}$  and the expectation  $\mathbb{E}_{\nu\pi}$  taken over the randomness of the observations from  $\nu$  and the algorithm  $\pi$ , Landau's notation  $o$ ,  $\mathcal{O}$ ,  $\Omega$  and  $\Theta$ , the  $(K - 1)$ -dimensional probability simplex  $\Delta_K := \left\{ w \in \mathbb{R}_+^K \mid w \geq 0, \sum_{i \in [K]} w_i = 1 \right\}$ . The functions  $[x]_0^1 := \max\{0, \min\{1, x\}\}$ ,  $k_\eta(x) := 1 + \log_{1+\eta} x$ ,  $\overline{W}_{-1}$  in Lemma 51,  $h$  in Eq. (31),  $r$  in Eq. (33),  $\zeta$  is the Riemann  $\zeta$  function. Moreover, we recall the definitions:  $d_\epsilon^\pm$  in Eq. (3),  $d_\epsilon$  in Eq. (2),  $\widetilde{d}_\epsilon^\pm$  in Eq. (32),  $W_{\epsilon,a,b}^\pm$  in Eq. (4),  $\widetilde{W}_{\epsilon,a,b}^\pm$  in Eq. (34),  $(T_\epsilon^*(\nu), T_{\epsilon,\beta}^*(\nu), w_\epsilon^*(\nu), w_{\epsilon,\beta}^*(\nu))$  in Eq. 35. While Table 1 gathers problem-specific notation, Table 2 groups notation for the algorithms.

Table 2: Notation for the algorithm.

Notation	Type	Description
$B_n$	$[K]$	(EB) Leader at time $n$
$C_n$	$[K]$	(TC) Challenger at time $n$
$a_n$	$[K]$	Arm sampled at time $n$
$X_{n,a_n}$	$\{0, 1\}$	Sample observed at the end of time $n$ , i.e. $X_{n,a_n} \sim \nu_{a_n}$
$Y_{k_{n,a},a}$	$\mathbb{R}$	Noisy perturbation drawn at the beginning of phase $k_{n,a}$ for arm $a$ , i.e. $Y_{k_{n,a},a} \sim \text{Lap}(1/\epsilon)$
$\mathcal{F}_n$		History before time $n$
$\tilde{a}_n$	$[K]$	Arm recommended before time $n$
$\tau_{\epsilon,\delta}$	$\mathbb{N}$	Sample complexity (stopping time)
$c(n, \epsilon, \delta)$	$\mathbb{N} \times \mathbb{R}_+^* \times (0, 1) \rightarrow \mathbb{R}_+^*$	Stopping threshold function
$c_1(n, \delta)$	$\mathbb{N} \times (0, 1) \rightarrow \mathbb{R}_+^*$	Stopping threshold function
$c_2(n, \epsilon)$	$\mathbb{N} \times \mathbb{R}_+^* \rightarrow \mathbb{R}_+^*$	Approximation threshold function
$N_{n,a}$	$\mathbb{N}$	Number of pulls of arm $a$ before time $n$
$k_{n,a}$	$\mathbb{N}$	Current phase of arm $a$ at time $n$
$T_k(a)$	$\mathbb{N}$	Time $n$ where the arm $a$ changes to phase $k$
$\tilde{S}_{k,a}$	$\mathbb{R}$	Private sum of observations for arm $a$ at phase $k$
$\tilde{N}_{n,a}$	$\mathbb{N}$	Number of pulls of arm $a$ at the beginning of phase $k_{n,a}$
$\tilde{\mu}_{n,a}$	$\mathbb{R}$	Private estimator of the empirical mean of arm $a$ at the beginning of phase $k_{n,a}$
$L_{n,a}$	$\mathbb{N}$	Counts of $B_t = a$ before time $n$
$N_{n,a}^a$	$\mathbb{N}$	Counts of $(B_t, a_t) = (a, a)$ before time $n$
$\beta$	$(0, 1)$	Fixed proportion

## C Related Work and Limitations

We provide a more detailed literature review in Appendix C.1, and discuss limitations of Theorem 6 in Appendix C.2.

### C.1 Related Work

**Structured Bandits.** While we consider unstructured bandits [6], numerous structural assumptions have been studied: linear bandits [77], generalized linear bandits [35] such as logistic bandits [53], combinatorial bandits [23], sparse bandits [46], spectral bandits [57], unimodal bandits [28], Lipschitz [65], partial monitoring [3], etc. Coping for the structural assumption while preserving  $\epsilon$ -global DP is an interesting direction for future works.

**Pure Exploration Problems.** While we consider only BAI [33], other pure exploration problems have been studied in the literature:  $\epsilon$ -BAI [66], thresholding bandits [19], Top- $k$  identification [56], Pareto set identification [7], best partition identification [21], etc. Extending our  $\epsilon$ -global DP results to answer these identification problems is an interesting research direction.

**Performance Metrics.** In pure exploration problems, the two major theoretical frameworks are the *fixed-confidence* setting [34, 47, 38], which is the focus of this paper, and the *fixed-budget* setting [4, 36]. In the fixed-budget setting, the objective is to minimize the probability of misidentifying a correct answer with a fixed number of samples  $T$ . Recent works have also considered the anytime setting, in which the agent aims at achieving a low probability of error at any deterministic time [88, 52]. Extending our findings to support  $\epsilon$ -global DP in the fixed-budget or the anytime setting is an interesting direction for future works, see e.g., Chen et al. [24].

**DP in bandits.** DP has been studied for multi-armed bandits under different bandit settings: finite-armed stochastic [68, 74, 89, 44, 9, 43, 10, 85, 45], adversarial [80, 1, 82], linear [41, 60, 10], contextual linear [76, 69, 89, 10], and kernel bandits [70], among others. Most of these works were for regret minimisation, but the problem has also been explored for best-arm identification, with fixed confidence [11, 12] and fixed budget [24]. The problem has also been studied under three different DP trust models: (a) global DP where the users trust the centralised decision maker [68, 76, 74, 9, 43],

995 (b) local DP where each user deploys a local perturbation mechanism to send a “noisy” version of  
 996 the rewards to the policy [15, 89, 40], and (c) shuffle DP where users still feed their data to a local  
 997 perturbation, but now they trust an intermediary to apply a uniformly random permutation on all  
 998 users’ data before sending to the central servers [79, 37, 27].

999 In the first papers on DP for bandits, the tree-based mechanism [32, 20] was used to compute the  
 1000 sum of rewards privately. However, this mechanism was proven to be sub-optimal, matching the  
 1001 lower bounds up to logarithmic factors. Then, forgetting was first proposed by [74] to get rid of  
 1002 the tree-based mechanism, then adapted to UCB in [44, 9]. Finally, if the KL is the divergence that  
 1003 controls the complexity of bandits without privacy [58, 38], then Azize and Basu [9] were the first to  
 1004 show that the TV controls the complexity of private bandits, in the high privacy regime.

1005 In this paper, *we focus on  $\epsilon$ -pure DP, under a global trust model, in stochastic finite-armed bandits,*  
 1006 *for best arm identification under fixed confidence.*

1007 **Gap in the literature.** This problem setting is first studied by Azize et al. [11], who proposed  
 1008 the first problem-dependent sample complexity lower bound, and introduced AdaP-TT, an  $\epsilon$ -global  
 1009 DP version of the Top Algorithm. However, the sample complexity upper bound of AdaP-TT only  
 1010 matches the lower bound in *the high privacy regime*  $\epsilon \rightarrow 0$ , and for instances where the means have  
 1011 similar order (see Condition 1 in [11] in the discussion after Theorem 5 in [11]).

1012 Azize et al. [12] proposes AdaP-TT\*, an improved version of AdaP-TT. The improvement is achieved  
 1013 by using a transport inspired by the sample complexity lower bound from [11]. Using the new  
 1014 transport, AdaP-TT\* gets rid of Condition 1 needed by AdaP-TT, and achieves the high privacy lower  
 1015 bound for all instances up to a multiplicative factor 48.

1016 However, both AdaP-TT and AdaP-TT\* do not match the lower bound, beyond the high privacy  
 1017 regime, i.e. for both the low privacy regime and transitional regimes.

## 1018 C.2 Limitations of Theorem 6

1019 Using adaptive targets  $\beta_n(B_n, C_n)$  in DP-TT could replace  $T_{\epsilon, \beta}^*(\nu)$  by  $T_\epsilon^*(\nu)$ . While we propose  
 1020 two adaptive choices of target based on IDS [86] or BOLD [14] (Appendix I), we leave their analysis  
 1021 for future work. The assumption that the means are distinct is used to prove sufficient exploration; it  
 1022 can be removed by using forced exploration or a fine-grained analysis [49, 52]. While it improves the  
 1023 asymptotic upper bound, choosing  $\eta$  too close to 0 negatively impacts the performance of DP-TT,  
 1024 due to the dependency in  $\mathcal{O}(1/\log(1 + \eta))$  of the stopping threshold. The suboptimal scaling in  
 1025  $2 \log(1/\delta)$  of the stopping threshold yields the factor 2.

## 1026 D Lower Bound

1027 Let  $\mathcal{M} : \mathcal{X}^n \rightarrow \mathcal{O}$  be an  $\epsilon$ -DP mechanism. For  $D \in \mathcal{X}^n$  an input dataset, we denote by  $\mathcal{M}_D$  the  
 1028 distribution over outputs, when the input is  $D$ , and  $\mathcal{M}_D(E)$  the probability of observing output  $E$   
 1029 when the input is  $D$ .

1030 Let  $\mathbb{P}$  and  $\mathbb{Q}$  be two data-generating distributions over  $\mathcal{X}^n$ . We denote by  $\mathbb{M}_{\mathbb{P}, \mathcal{M}}$  the marginal over  
 1031 outputs of the mechanism  $\mathcal{M}$  when the input dataset is generated through  $\mathbb{P}$ , i.e.

$$\mathbb{M}_{\mathbb{P}, \mathcal{M}}(A) := \int_{D \in \mathcal{X}^n} \mathcal{M}_D(A) \, d\mathbb{P}(D) , \quad (9)$$

1032 for any event  $A$  in the output space. We define similarly  $\mathbb{M}_{\mathbb{Q}, \mathcal{M}}$  the marginal over outputs of the  
 1033 mechanism  $\mathcal{M}$  when the input dataset is generated through  $\mathbb{Q}$ .

1034 The main question is to control the divergence  $\text{KL}(\mathbb{M}_{\mathbb{P}, \mathcal{M}} \parallel \mathbb{M}_{\mathbb{Q}, \mathcal{M}})$  when the mechanism  $\mathcal{M}$  satisfies  
 1035 DP. In general, for any mechanism  $\mathcal{M}$ , the data-processing inequality provides the following bound

$$\text{KL}(\mathbb{M}_{\mathbb{P}, \mathcal{M}} \parallel \mathbb{M}_{\mathbb{Q}, \mathcal{M}}) \leq \text{KL}(\mathbb{P} \parallel \mathbb{Q}) . \quad (10)$$

1036 Now, for  $\epsilon$ -DP mechanisms, we want to translate the DP constraint to a tight bound on the divergence  
 1037  $\text{KL}(\mathbb{M}_{\mathbb{P}, \mathcal{M}} \parallel \mathbb{M}_{\mathbb{Q}, \mathcal{M}})$ . To do so, let  $\mathbb{L}$  be any other distribution on  $\mathcal{X}^n$ . Let  $\mathbb{C}_{\mathbb{P}, \mathbb{L}}$  be a coupling of  
 1038  $(\mathbb{P}, \mathbb{L})$ , i.e., the marginals of  $\mathbb{C}_{\mathbb{P}, \mathbb{L}}$  are  $\mathbb{P}$  and  $\mathbb{L}$ . We can now rewrite our the marginals using the



1039 definition of couplings. For  $\mathbb{M}_{\mathbb{P}, \mathcal{M}}$ , we have

$$\mathbb{M}_{\mathbb{P}, \mathcal{M}}(A) := \int_{D \in \mathcal{X}^n} \mathcal{M}_D(A) \, d\mathbb{P}(D) = \int_{D, D' \in \mathcal{X}^n} \mathcal{M}_D(A) \, d\mathbb{C}_{\mathbb{P}, \mathbb{L}}(D, D'),$$

1040 and for  $\mathbb{Q}$  we get

$$\begin{aligned} \mathbb{M}_{\mathbb{Q}, \mathcal{M}}(A) &:= \int_{D' \in \mathcal{X}^n} \mathcal{M}_{D'}(A) \, d\mathbb{Q}(D') = \int_{D' \in \mathcal{X}^n} \mathcal{M}_{D'}(A) \frac{d\mathbb{Q}(D')}{d\mathbb{L}(D')} \, d\mathbb{L}(D') \\ &= \int_{D, D' \in \mathcal{X}^n} \mathcal{M}_{D'}(A) \frac{d\mathbb{Q}(D')}{d\mathbb{L}(D')} \, d\mathbb{C}_{\mathbb{P}, \mathbb{L}}(D, D'). \end{aligned}$$

1041 Using the data-processing inequality, we get

$$\begin{aligned} \text{KL}(\mathbb{M}_{\mathbb{P}, \mathcal{M}} \parallel \mathbb{M}_{\mathbb{Q}, \mathcal{M}}) &\leq \int_{D, D' \in \mathcal{X}^n} \int_{o \in \mathcal{O}} \log \left( \frac{\mathcal{M}_D(o)}{\mathcal{M}_{D'}(o) \frac{d\mathbb{Q}(D')}{d\mathbb{L}(D')}} \right) \mathcal{M}_D(o) \, do \, d\mathbb{C}_{\mathbb{P}, \mathbb{L}}(D, D') \\ &= \int_{D, D' \in \mathcal{X}^n} \left( \text{KL}(\mathcal{M}_D \parallel \mathcal{M}_{D'}) + \log \left( \frac{d\mathbb{L}(D')}{d\mathbb{Q}(D')} \right) \right) \, d\mathbb{C}_{\mathbb{P}, \mathbb{L}}(D, D') \\ &= \mathbb{E}_{D, D' \sim \mathbb{C}_{\mathbb{P}, \mathbb{L}}} [\text{KL}(\mathcal{M}_D \parallel \mathcal{M}_{D'})] + \text{KL}(\mathbb{L} \parallel \mathbb{Q}). \end{aligned}$$

1042 Since this is true for any coupling  $\mathbb{C}_{\mathbb{P}, \mathbb{L}}$  and any distribution  $\mathbb{L}$ , we get the final bound

$$\text{KL}(\mathbb{M}_{\mathbb{P}, \mathcal{M}} \parallel \mathbb{M}_{\mathbb{Q}, \mathcal{M}}) \leq \inf_{\mathbb{L} \in \mathcal{P}(\mathcal{X}^n)} \left\{ \inf_{\mathbb{C}_{\mathbb{P}, \mathbb{L}} \in \mathcal{C}(\mathbb{P}, \mathbb{L})} \left\{ \mathbb{E}_{D, D' \sim \mathbb{C}_{\mathbb{P}, \mathbb{L}}} [\text{KL}(\mathcal{M}_D \parallel \mathcal{M}_{D'})] \right\} + \text{KL}(\mathbb{L} \parallel \mathbb{Q}) \right\}$$

1043 where  $\mathcal{P}(\mathcal{X}^n)$  is the set of all distributions over  $\mathcal{X}^n$  and  $\mathcal{C}(\mathbb{P}, \mathbb{L})$  is the set of all couplings between  $\mathbb{P}$   
1044 and  $\mathbb{L}$ . Using that the  $\mathcal{M}$  is  $\epsilon$ -DP, we can use the simple bound  $\text{KL}(\mathcal{M}_D \parallel \mathcal{M}_{D'}) \leq \epsilon d_{\text{Ham}}(D, D')$   
1045 which gives

$$\text{KL}(\mathbb{M}_{\mathbb{P}, \mathcal{M}} \parallel \mathbb{M}_{\mathbb{Q}, \mathcal{M}}) \leq \inf_{\mathbb{L} \in \mathcal{P}(\mathcal{X}^n)} \left\{ \epsilon \inf_{\mathbb{C}_{\mathbb{P}, \mathbb{L}} \in \mathcal{C}(\mathbb{P}, \mathbb{L})} \left\{ \mathbb{E}_{D, D' \sim \mathbb{C}_{\mathbb{P}, \mathbb{L}}} [d_{\text{Ham}}(D, D')] \right\} + \text{KL}(\mathbb{L} \parallel \mathbb{Q}) \right\}. \quad (11)$$

## 1046 D.1 Product Distributions

1047 Suppose that  $\mathbb{P} := \bigotimes_{i=1}^n \mathbb{P}_i$  and  $\mathbb{Q} := \bigotimes_{i=1}^n \mathbb{Q}_i$  are product distributions. Consider the subset of  
1048 product distributions  $\mathbb{L} := \bigotimes_{i=1}^n \mathbb{L}_i$ , and the maximal coupling  $\mathbb{C}_{\infty}(\mathbb{P}, \mathbb{L}) := \prod_{i=1}^n \mathbb{C}_{\infty}(\mathbb{P}_i, \mathbb{L}_i)$ .  
1049 Plugging these in Equation (11), we get

$$\begin{aligned} \text{KL}(\mathbb{M}_{\mathbb{P}, \mathcal{M}} \parallel \mathbb{M}_{\mathbb{Q}, \mathcal{M}}) &\leq \inf_{\mathbb{L}_1, \dots, \mathbb{L}_n} \left\{ \epsilon \sum_{i=1}^n \mathbb{E}_{D_i, D'_i \sim \mathbb{C}_{\infty}(\mathbb{P}_i, \mathbb{L}_i)} [\mathbb{1}\{D_i \neq D'_i\}] + \sum_{i=1}^n \text{KL}(\mathbb{L}_i \parallel \mathbb{Q}_i) \right\} \\ &= \inf_{\mathbb{L}_1, \dots, \mathbb{L}_n} \left\{ \sum_{i=1}^n (\epsilon \text{TV}(\mathbb{P}_i \parallel \mathbb{L}_i) + \text{KL}(\mathbb{L}_i \parallel \mathbb{Q}_i)) \right\} \\ &= \sum_{i=1}^n \inf_{\mathbb{L}_i \in \mathcal{P}(\mathcal{X})} \{ \epsilon \text{TV}(\mathbb{P}_i \parallel \mathbb{L}_i) + \text{KL}(\mathbb{L}_i \parallel \mathbb{Q}_i) \} = \sum_{i=1}^n d_{\epsilon}(\mathbb{P}_i, \mathbb{Q}_i), \end{aligned}$$

1050 where

$$d_{\epsilon}(\mathbb{P}, \mathbb{Q}) := \inf_{\mathbb{L} \in \mathcal{P}(\mathcal{X})} \{ \epsilon \text{TV}(\mathbb{P} \parallel \mathbb{L}) + \text{KL}(\mathbb{L} \parallel \mathbb{Q}) \}. \quad (12)$$

## 1051 D.2 Sequential KL decomposition for bandits under DP

1052 In this section, we adapt the techniques from product distributions to bandit marginals.

1053 First, we introduce the bandit canonical model.

1054 **The bandit canonical model.** A stochastic bandit (or environment) is a collection of distributions  $\nu \triangleq$   
1055  $(P_a : a \in [K])$ , where  $[K]$  is the set of available  $K$  actions. The learner and the environment interact

---

**Algorithm 2** Bandit interaction between a policy and an environment

---

```

1: Input: A policy  $\pi$  and an environment  $\nu \triangleq (P_a : a \in [K])$ 
2: for  $t = 1, \dots$  do
3:   The policy samples an action  $a_t \sim \pi_t(\cdot \mid a_1, r_1, \dots, a_{t-1}, r_{t-1})$ 
4:   The policy observes a reward  $r_t \sim P_{a_t}$ 
5: end for
6: if Regret minimisation then
7:   The interaction ends after  $T$  steps
8: else FC-BAI
9:   The policy decides to stop the interaction at step  $\tau_{\epsilon, \delta}$  and recommends the final guess  $\hat{a}$ 
10: end if

```

---

1056 sequentially over  $T$  rounds. In each round  $t \in 1, \dots, T$ , the learner chooses an action  $a_t \in [K]$ ,  
 1057 which is fed to the environment. The environment then samples a reward  $r_t \in \mathbb{R}$  from distribution  
 1058  $P_{a_t}$  and reveals  $r_t$  to the learner. The interaction between the learner (or policy) and environment  
 1059 induces a probability measure on the sequence of outcomes  $H_T \triangleq (a_1, r_1, a_2, r_2, \dots, a_T, r_T)$ . In  
 1060 the following, we construct the probability space that carries these random variables.

1061 Let  $T \in \mathbb{N}^*$  be the horizon. Let  $\nu = (P_a : a \in [K])$  a bandit instance with  $K \in \mathbb{N}^*$  finite arms,  
 1062 and each  $P_a$  is a probability measure on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  with  $\mathfrak{B}$  being the Borel set. For each  $t \in [T]$ , let  
 1063  $\Omega_t = ([K] \times \mathbb{R})^t \subset \mathbb{R}^{2t}$  and  $\mathcal{F}_t = \mathfrak{B}(\Omega_t)$ . We first formalise the definition of a policy.

1064 **Definition 2** (The policy). *A policy  $\pi$  is a sequence  $(\pi_t)_{t=1}^T$ , where  $\pi_t$  is a probability kernel from*  
 1065  *$(\Omega_t, \mathcal{F}_t)$  to  $([K], 2^{[K]})$ . Since  $[K]$  is discrete, we adopt the convention that for  $a \in [K]$ ,*

$$\pi_t(a \mid a_1, r_1, \dots, a_{t-1}, r_{t-1}) = \pi_t(\{a\} \mid a_1, r_1, \dots, a_{t-1}, r_{t-1})$$

1066 We want to define a probability measure on  $(\Omega_T, \mathcal{F}_T)$  that respects our understanding of the sequential  
 1067 nature of the interaction between the learner and a stationary stochastic bandit. Specifically, the  
 1068 sequence of outcomes should satisfy the following two assumptions:

- 1069 (a) The conditional distribution of action  $a_t$  given  $a_1, r_1, \dots, a_{t-1}, r_{t-1}$  is  $\pi(a_t \mid H_{t-1})$  almost  
 1070 surely.
- 1071 (b) The conditional distribution of reward  $r_t$  given  $a_1, r_1, \dots, a_{t-1}, r_{t-1}, a_t$  is  $P_{a_t}$  almost  
 1072 surely.

1073 The probability measure on  $(\Omega_T, \mathcal{F}_T)$  depends on both the environment  $\nu$  and the policy  $\pi$ . To  
 1074 construct this probability, let  $\lambda$  be a  $\sigma$ -finite measure on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  for which  $P_a$  is absolutely  
 1075 continuous with respect to  $\lambda$  for all  $a \in [K]$ . Let  $p_a = dP_a/d\lambda$  be the Radon–Nikodym derivative of  
 1076  $P_a$  with respect to  $\lambda$ . Letting  $\rho$  be the counting measure with  $\rho(B) = |B|$ , the density  $p_{\nu\pi} : \Omega_T \rightarrow \mathbb{R}$   
 1077 can now be defined with respect to the product measure  $(\rho \times \lambda)^T$  by

$$p_{\nu\pi}(a_1, r_1, \dots, a_T, r_T) \triangleq \prod_{t=1}^T \pi_t(a_t \mid a_1, r_1, \dots, a_{t-1}, r_{t-1}) p_{a_t}(r_t)$$

1078 and  $\mathcal{P}_{\nu\pi}$  is defined as

$$\mathbb{P}_{\nu\pi}(B) \triangleq \int_B p_{\nu\pi}(\omega) (\rho \times \lambda)^T(d\omega) \quad \text{for all } B \in \mathcal{F}_T$$

1079 Hence  $(\Omega_T, \mathcal{F}_T, \mathbb{P}_{\nu\pi})$  is a probability space over histories induced by the interaction between  $\pi$  and  
 1080  $\nu$ . We define also a marginal distribution over the sequence of actions by

$$m_{\nu\pi}(a_1, \dots, a_T) \triangleq \int_{r_1, \dots, r_T} p_{\nu\pi}(a_1, r_1, \dots, a_T, r_T) dr_1 \dots dr_T,$$

1081 and for all  $C \in \mathcal{P}([K]^T)$ ,

$$\mathbb{M}_{\nu\pi}(C) \triangleq \sum_{(a_1, \dots, a_T) \in C} m_{\nu\pi}(a_1, a_2, \dots, a_T).$$

Finally,  $([K]^T, \mathcal{P}([K]^T), \mathbb{M}_{\nu\pi})$  is a probability space over sequence of actions produced when  $\pi$  interacts with  $\nu$  for  $T$  time-steps.

**The KL upper bound.** Now, we adapt the techniques for the bandit marginals. Let  $\nu = \{P_a, a \in [K]\}$  and  $\nu' = \{P'_a, a \in [K]\}$  be two bandit instances in  $\mathcal{F}^K$ . We recall that, when the policy  $\pi$  interacts with the bandit instance  $\nu$ , it induces a marginal distribution  $\mathbb{M}_{\nu\pi}$  over the sequence of actions. We define  $\mathbb{M}_{\nu'\pi}$  similarly.

The goal is to upper bound the quantity  $\text{KL}(\mathbb{M}_{\nu\pi} \parallel \mathbb{M}_{\nu'\pi})$ . The marginals  $\mathbb{M}_{\nu\pi}$  and  $\mathbb{M}_{\nu'\pi}$  in the sequential setting "look like" marginals generated by "product distributions". However, the hardness of the sequential setting lies in the fact that the data-generating distributions depend on the stochastic sequential actions chosen. Thus, the results of the previous section cannot be directly applied. To adapt the proof ideas of the previous section to the bandit case, we introduce the idea of a coupled bandit instance.

Let  $\nu'' = \{P''_a : a \in [K]\}$  be any "intermediary" bandit instance from  $\mathcal{F}^K$ . Define  $c_a$  as the maximal coupling between  $P_a$  and  $P''_a$ , i.e.,  $c_a := \mathbb{C}_\infty(P_a, P''_a)$ . Fix a policy  $\pi = \{\pi_t\}_{t=1}^T$ .

Here, we build a coupled environment  $\gamma$  of  $\nu$  and  $\nu''$ . The policy  $\pi$  interacts with the coupled environment  $\gamma$  up to a given time horizon  $T$  to produce an augmented history  $\{(a_t, r_t, r''_t)\}_{t=1}^T$ . The iterative steps of this interaction process are:

1. The probability of choosing an action  $a_t = a$  at time  $t$  is dictated only by the policy  $\pi_t$  and  $a_1, r_1, a_2, r_2, \dots, a_{t-1}, r_{t-1}$ , i.e. the policy ignores  $\{r''_s\}_{s=1}^{t-1}$ .
2. The distribution of rewards  $(r_t, r''_t)$  is  $c_{a_t}$  and is conditionally independent of the previous observed history  $\{(a_s, r_s, r''_s)\}_{s=1}^{t-1}$ .

This interaction is similar to the interaction process of policy  $\pi$  with the first bandit instance  $\nu$ , with the addition of sampling an extra  $r''_t$  from the coupling of  $P_{a_t}$  and  $P''_{a_t}$ .

The distribution of the augmented history induced by the interaction of  $\pi$  and the coupled environment can be defined as

$$p_{\gamma\pi}(a_1, r_1, r''_1, \dots, a_T, r_T, r''_T) := \prod_{t=1}^T \pi_t(a_t \mid a_1, r_1, \dots, a_{t-1}, r_{t-1}) c_{a_t}(r_t, r''_t).$$

To simplify the notation, let  $\mathbf{a} := (a_1, \dots, a_T)$ ,  $\mathbf{r} := (r_1, \dots, r_T)$ ,  $\mathbf{r}' := (r'_1, \dots, r'_T)$  and  $\mathbf{r}'' := (r''_1, \dots, r''_T)$ . Also, let  $c_a(\mathbf{r}, \mathbf{r}'') := \prod_{t=1}^T c_{a_t}(r_t, r''_t)$  and  $\pi(\mathbf{a} \mid \mathbf{r}) := \prod_{t=1}^T \pi_t(a_t \mid a_1, r_1, \dots, a_{t-1}, r_{t-1})$ . We put  $\mathbf{h} := (\mathbf{a}, \mathbf{r}, \mathbf{r}'')$ . With the new notation

$$p_{\gamma\pi}(\mathbf{a}, \mathbf{r}, \mathbf{r}'') := \pi(\mathbf{a} \mid \mathbf{r}) c_a(\mathbf{r}, \mathbf{r}'').$$

By the definition of the couplings, we have that  $m_{\nu\pi}$  is the marginal of  $p_{\gamma\pi}$  when integrated over  $(\mathbf{r}, \mathbf{r}'')$ , i.e.,

$$m_{\nu\pi}(\mathbf{a}) = \int_{\mathbf{r}, \mathbf{r}''} p_{\gamma\pi}(\mathbf{a}, \mathbf{r}, \mathbf{r}'') d\mathbf{r} d\mathbf{r}''.$$

Now, we define a new joint distribution  $q_{\gamma\pi}$ , inspired by the techniques used for product distributions:

$$q_{\gamma\pi}(\mathbf{a}, \mathbf{r}, \mathbf{r}'') := \pi(\mathbf{a} \mid \mathbf{r}'') \frac{p'_a(\mathbf{r}'')}{p''_a(\mathbf{r}'')} c_a(\mathbf{r}, \mathbf{r}''),$$

where  $p'_a(\mathbf{r}'') := \prod_{t=1}^T p'_{a_t}(r''_t)$ , and similarly,  $p''_a(\mathbf{r}'') := \prod_{t=1}^T p''_{a_t}(r''_t)$ .

First, observe that it is indeed a valid joint distribution, i.e.

$$\begin{aligned} \sum_{\mathbf{a}} \int_{\mathbf{r}, \mathbf{r}''} q_{\gamma\pi}(\mathbf{a}, \mathbf{r}, \mathbf{r}'') d\mathbf{r} d\mathbf{r}'' &= \sum_{\mathbf{a}} \int_{\mathbf{r}, \mathbf{r}''} \pi(\mathbf{a} \mid \mathbf{r}'') \frac{p'_a(\mathbf{r}'')}{p''_a(\mathbf{r}'')} c_a(\mathbf{r}, \mathbf{r}'') d\mathbf{r} d\mathbf{r}'' \\ &= \sum_{\mathbf{a}} \int_{\mathbf{r}''} \pi(\mathbf{a} \mid \mathbf{r}'') p'_a(\mathbf{r}'') d\mathbf{r}'' = \int_{\mathbf{r}''} p'_a(\mathbf{r}'') d\mathbf{r}'' = 1, \end{aligned}$$

and that  $m_{\nu'\pi}$  is the marginal of  $q_{\gamma\pi}$  when integrated over  $(\mathbf{r}, \mathbf{r}'')$ , i.e.,

$$\int_{\mathbf{r}, \mathbf{r}''} q_{\gamma\pi}(\mathbf{a}, \mathbf{r}, \mathbf{r}'') d\mathbf{r} d\mathbf{r}'' = \int_{\mathbf{r}, \mathbf{r}''} \pi(\mathbf{a} \mid \mathbf{r}'') \frac{p'_a(\mathbf{r}'')}{p''_a(\mathbf{r}'')} c_a(\mathbf{r}, \mathbf{r}'') d\mathbf{r} d\mathbf{r}''$$

$$= \int_{\mathbf{r}''} \pi(\mathbf{a} \mid \mathbf{r}'') p'_a(\mathbf{r}'') d\mathbf{r}'' = m_{\nu' \pi}(\mathbf{a}) .$$

1116 Using the data-processing inequality, we get

$$\text{KL}(\mathbb{M}_{\nu \pi} \parallel \mathbb{M}_{\nu' \pi}) \leq \text{KL}(p_{\gamma \pi} \parallel q_{\gamma \pi}) . \quad (13)$$

1117 Now, we compute

$$\begin{aligned} \text{KL}(p_{\gamma \pi} \parallel q_{\gamma \pi}) &\stackrel{(a)}{=} \mathbb{E}_{\mathbf{h}: (\mathbf{a}, \mathbf{r}, \mathbf{r}'') \sim p_{\gamma \pi}} \left[ \log \left( \frac{\pi(\mathbf{a} \mid \mathbf{r}) c_a(\mathbf{r}, \mathbf{r}'')}{\pi(\mathbf{a} \mid \mathbf{r}'') \frac{p'_a(\mathbf{r}'')}{p'_a(\mathbf{r}'')} c_a(\mathbf{r}, \mathbf{r}'')} \right) \right] \\ &\stackrel{(b)}{\leq} \mathbb{E}_{\mathbf{h}: (\mathbf{a}, \mathbf{r}, \mathbf{r}'') \sim p_{\gamma \pi}} \left[ \epsilon d_{\text{Ham}}(\mathbf{r}, \mathbf{r}'') + \log \left( \frac{p''_a(\mathbf{r}'')}{p'_a(\mathbf{r}'')} \right) \right] \\ &\stackrel{(c)}{=} \sum_{t=1}^T \mathbb{E}_{\mathbf{h} \sim p_{\gamma \pi}} \left[ \epsilon \mathbb{1}\{r_t \neq r''_t\} + \log \left( \frac{p''_{a_t}(r''_t)}{p'_{a_t}(r''_t)} \right) \right] \\ &\stackrel{(d)}{=} \sum_{t=1}^T \mathbb{E}_{\mathbf{h} \sim p_{\gamma \pi}} \left[ \mathbb{E}_{\mathbf{h} \sim p_{\gamma \pi}} [\epsilon \mathbb{1}\{r_t \neq r''_t\} + \log \left( \frac{p''_{a_t}(r''_t)}{p'_{a_t}(r''_t)} \right) \mid a_t] \right] \\ &\stackrel{(e)}{=} \sum_{t=1}^T \mathbb{E}_{\mathbf{h} \sim p_{\gamma \pi}} [\epsilon \text{TV}(p_{a_t} \parallel p''_{a_t}) + \text{KL}(p''_{a_t} \parallel p'_{a_t})] \\ &\stackrel{(f)}{=} \mathbb{E}_{\nu \pi} \left[ \sum_{t=1}^T \epsilon \text{TV}(p_{a_t} \parallel p''_{a_t}) + \text{KL}(p''_{a_t} \parallel p'_{a_t}) \right] . \end{aligned}$$

1118 where:

1119 (a) by the definition of the KL

(b) the group privacy property, applied to the  $\epsilon$ -global DP policy, we have

$$\pi(\mathbf{a} \mid \mathbf{r}) \leq e^{\epsilon d_{\text{Ham}}(\mathbf{r}, \mathbf{r}'')} \pi(\mathbf{a} \mid \mathbf{r}'')$$

1120 (c) by the definition of dham

1121 (d) by the towering property of conditional expectations

1122 (e) given  $a_t$ , we have  $r_t \sim p_{a_t}$ ,  $r'_t \sim p'_{a_t}$  and  $r'' \sim p''_{a_t}$

1123 (f) by linearity of the expectation, and the fact that the expression inside the expectation only depends  
1124 on the actions  $a_t$

1125 Since this is true for any “intermediary” bandit instance  $\nu'' \in \mathcal{F}^K$ , we take  $\nu''_*$  to be the environment  
1126 where the infimum of the  $d_\epsilon(P_a, P'_a)$  is attained for each arm  $a \in [K]$ . Specifically, let  $\nu''_* =$   
1127  $(p_a^*, a \in [K])$  where

$$p_a^* = \arg \min_{\mathbb{L} \in \mathcal{F}} \{\epsilon \text{TV}(p_a \parallel \mathbb{L}) + \text{KL}(\mathbb{L} \parallel p'_a)\}$$

1128 Plugging  $\nu''_*$  gives

$$\text{KL}(\mathbb{M}_{\nu \pi} \parallel \mathbb{M}_{\nu' \pi}) \leq \mathbb{E}_{\nu \pi} \left[ \sum_{t=1}^T d_\epsilon(p_{a_t}, p'_{a_t}) \right] \quad (14)$$

1129 Let  $N_{t,a} = \sum_{s < t} \mathbb{1}\{a_s = a\}$  be the counts of arm  $a$  before step  $t$ . Then, we can rewrite the bound  
1130 as

$$\text{KL}(\mathbb{M}_{\nu \pi} \parallel \mathbb{M}_{\nu' \pi}) \leq \sum_{a=1}^K \mathbb{E}_{\nu \pi} [N_{T+1,a}] d_\epsilon(p_a, p'_a) , \quad (15)$$

1131 **Stopping time version of the KL decomposition for BAI under DP.** Let  $\pi$  be an  $\epsilon$ -DP BAI strategy.  
1132 Let  $\nu$  and  $\lambda$  be two bandit instances. Denote by  $\mathbb{M}_{\nu \pi}$  the marginal distribution of the output of the

1133 BAI strategy when  $\pi$  interacts with  $\nu$ . By using Wald's lemma in the proof technique seen before for  
 1134 the canonical bandit setting under FC-BAI, we get that

$$\text{KL}(\mathbb{M}_{\nu\pi} \parallel \mathbb{M}_{\lambda\pi}) \leq \mathbb{E}_{\nu\pi} \left( \sum_{t=1}^{\tau_{\epsilon,\delta}} d_{\epsilon}(\nu_{a_t}, \lambda_{a_t}) \right) = \sum_{a=1}^K \mathbb{E}_{\nu\pi} [N_{\tau_{\epsilon,\delta}+1,a}] d_{\epsilon}(\nu_a, \lambda_a), \quad (16)$$

1135 where  $\tau$  is the stopping time.

### 1136 D.3 Sample Complexity Lower Bound Proof

**Theorem 2** (Sample complexity lower bound for BAI under  $\epsilon$ -DP). *Let  $(\epsilon, \delta) \in \mathbb{R}_+^* \times (0, 1)$ . For any algorithm  $\pi$  that is  $\delta$ -correct and  $\epsilon$ -global DP on  $\mathcal{F}^K$ ,*

$$\mathbb{E}_{\nu\pi}[\tau_{\epsilon,\delta}] \geq T_{\epsilon}^*(\nu) \log(1/(3\delta))$$

1137 for all  $\nu \in \mathcal{F}^K$  with unique best arm. The inverse of the characteristic time  $T_{\epsilon}^*(\nu)$  is defined as

$$T_{\epsilon}^*(\nu)^{-1} := \sup_{w \in \Delta_K} \inf_{\kappa \in \text{Alt}(\nu)} \sum_{a=1}^K w_a d_{\epsilon}(\nu_a, \kappa_a), \quad (17)$$

$$d_{\epsilon}(\nu_a, \kappa_a) := \inf_{\varphi_a \in \mathcal{F}} \{ \text{KL}(\varphi_a \parallel \kappa_a) + \epsilon \cdot \text{TV}(\nu_a \parallel \varphi_a) \}. \quad (18)$$

1138 *Proof.* Let  $\pi$  be an  $\epsilon$ -global DP  $\delta$ -correct BAI strategy. Let  $\nu$  be a bandit instance and  $\lambda \in \text{Alt}(\nu)$ .

1139 Let  $\mathbb{M}_{\nu\pi}$  denote the probability distribution of the output when the BAI strategy  $\pi$  interacts with  $\nu$ .

1140 For any alternative instance  $\lambda \in \text{Alt}(\nu)$ , the data-processing inequality gives that

$$\text{KL}(\mathbb{M}_{\nu\pi} \parallel \mathbb{M}_{\lambda,\pi}) \geq \text{kl}(\mathbb{M}_{\nu\pi}(\tilde{a} = a^*(\nu)), \mathbb{M}_{\lambda,\pi}(\tilde{a} = a^*(\nu))) \geq \text{kl}(1 - \delta, \delta), \quad (19)$$

1141 where the second inequality is because  $\pi$  is  $\delta$ -correct, i.e.,  $\mathbb{M}_{\nu\pi}(\tilde{a} = a^*(\nu)) \geq 1 - \delta$  and

1142  $\mathbb{M}_{\lambda,\pi}(\tilde{a} = a^*(\nu)) \leq \delta$ , and the monotonicity of the kl. Now, using the stopping time version

1143 of the KL decomposition for FC-BAI, we get that

$$\text{kl}(1 - \delta, \delta) \leq \text{KL}(\mathbb{M}_{\nu,\pi} \parallel \mathbb{M}_{\lambda,\pi}) \leq \sum_{a=1}^K \mathbb{E}_{\nu\pi} [N_{\tau_{\epsilon,\delta}+1,a}] d_{\epsilon}(\nu_a, \lambda_a).$$

1144 Since this is true for all  $\lambda \in \text{Alt}(\nu)$ , we get

$$\begin{aligned} \text{kl}(1 - \delta, \delta) &\leq \inf_{\lambda \in \text{Alt}(\nu)} \sum_{a=1}^K \mathbb{E}_{\nu\pi} [N_{\tau_{\epsilon,\delta}+1,a}] d_{\epsilon}(\nu_a, \lambda_a) \stackrel{(a)}{=} \mathbb{E}[\tau_{\epsilon,\delta}] \inf_{\lambda \in \text{Alt}(\nu)} \sum_{a=1}^K \frac{\mathbb{E}[N_{\tau_{\epsilon,\delta}+1,a}]}{\mathbb{E}[\tau_{\epsilon,\delta}]} d_{\epsilon}(\nu_a, \lambda_a) \\ &\stackrel{(b)}{\leq} \mathbb{E}[\tau_{\epsilon,\delta}] \left( \sup_{\omega \in \Delta_K} \inf_{\lambda \in \text{Alt}(\nu)} \sum_{a=1}^K \omega_a d_{\epsilon}(\nu_a, \lambda_a) \right). \end{aligned}$$

1145 (a) is due to the fact that  $\mathbb{E}[\tau_{\epsilon,\delta}]$  does not depend on  $\lambda$ . (b) is obtained by noting that the vector

1146  $(\omega_a)_{a \in [K]} \triangleq \left( \frac{\mathbb{E}_{\nu,\pi}[N_{\tau_{\epsilon,\delta}+1,a}]}{\mathbb{E}_{\nu,\pi}[\tau_{\epsilon,\delta}]} \right)_{a \in [K]}$  belongs to the simplex  $\Delta_K$ . The theorem follows by noting

1147 that for  $\delta \in (0, 1)$ ,  $\text{kl}(1 - \delta, \delta) \geq \log(1/3\delta)$ .  $\square$

## 1148 E Privacy Analysis

1149 In this section, we prove Lemma 4. First, we justify using a geometric grid for updating the means  
 1150 (Lemma 7). Second, we obtain Lemma 4 as a combination of Lemma 7 and the post-processing  
 1151 property of DP (Proposition 1).

### 1152 E.1 Releasing partial sums privately

1153 First, the following lemma justifies the use of geometric grids, and provides that the price of getting  
 1154 rid of forgetting is summing the Laplace noise from previous phases.

1155 **Lemma 7** (Privacy of our grid-based mean estimator). *Let  $T \in \{1, \dots\}$ ,  $\ell < T$  and  $t_1, \dots, t_\ell, t_{\ell+1}$*   
 1156 *be in  $[1, T]$  such that  $1 = t_1 < \dots < t_\ell < t_{\ell+1} - 1 = T$ .*

1157 *Let  $\mathcal{M}$  be the following mechanism:*

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{pmatrix} \xrightarrow{\mathcal{M}} \begin{pmatrix} (x_1 + \dots + x_{t_2-1}) + (Y_1) \\ (x_1 + \dots + x_{t_3-1}) + (Y_1 + Y_2) \\ \vdots \\ (x_1 + \dots + x_T) + (Y_1 + Y_2 + \dots + Y_{\ell-1}) \end{pmatrix}$$

1158 *where  $(Y_1, \dots, Y_\ell) \sim^{iid} \text{Lap}(1/\epsilon)$ .*

1159 *Then, for any  $\{x_1, \dots, x_T\} \in [0, 1]^T$ ,  $\mathcal{M}$  is  $\epsilon$ -DP.*

1160 *Proof.* First, consider the following mechanism, that only computes the partial sums:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{pmatrix} \rightarrow \begin{pmatrix} x_1 + \dots + x_{t_2-1} \\ x_{t_2} + \dots + x_{t_3-1} \\ \vdots \\ x_{t_{\ell-1}} + \dots + x_T \end{pmatrix}.$$

1161 Because  $x_t \in [0, 1]$ , the sensitivity of each partial sum is 1. Since each partial sum is computed  
 1162 over non-overlapping sequences, combining the Laplace mechanism (Theorem 1) with the parallel  
 1163 composition property of DP (Lemma 3) gives that the following mechanism:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{pmatrix} \xrightarrow{\mathcal{P}} \begin{pmatrix} x_1 + \dots + x_{t_2-1} + Y_1 \\ x_{t_2} + \dots + x_{t_3-1} + Y_2 \\ \vdots \\ x_{t_{\ell-1}} + \dots + x_T + Y_{\ell-1} \end{pmatrix}$$

1164 is  $\epsilon$ -DP, where  $(Y_1, \dots, Y_{\ell-1}) \sim^{iid} \text{Lap}(1/\epsilon)$ .

1165 Consider the post-processing function  $f : (x_1, \dots, x_{\ell-1}) \rightarrow (x_1, x_1 + x_2, \dots, x_1 + x_2 + \dots + x_{\ell-1})$ .  
 1166 Then, we have that that  $\mathcal{M} = f \circ \mathcal{P}$ . So, by the post-processing property of DP,  $\mathcal{M}$  is  $\epsilon$ -DP.  $\square$

1167 **Remark 1.** *Mechanism  $\mathcal{P}$ , defined in the proof of Lemma 7, is the fundamental mechanism used by*  
 1168 *all previous bandit algorithms [74, 9, 43, 12] to justify the use of forgetting. Our mechanism  $\mathcal{M}$  is*  
 1169 *just summing over the partial sums computed on each phase, and thus the price of having sums of*  
 1170  *$x_i$  that start from the beginning (i.e. do not forget) is that we have to sum now the noise from all*  
 1171 *previous phases too.*

## 1172 E.2 Proof of Lemma 4

1173 We are now ready to prove Lemma 4, i.e. that any BAI algorithm based solely on using  $\text{GPE}_\eta(\epsilon)$  to  
 1174 access observations is  $\epsilon$ -global DP on  $[0, 1]$ .

1175 *Proof.* Let  $\pi$  be a BAI algorithm using only  $\text{GPE}_\eta(\epsilon)$  to access observations. Let  $R = \{x_1, \dots\}$   
 1176 and  $R' = \{x'_1, \dots\}$  be two neighbouring sequences of private observations, i.e. there exists a  
 1177  $t^* \in \{1, \dots\}$  such that  $x_t = x'_t$  for all  $t \neq t^*$ , i.e. that  $R$  and  $R'$  only differ at  $t^*$ .

1178 Fix a stopping time, recommendation and sampled actions  $(T + 1, \tilde{a}, (a_1, \dots, a_T))$ , we want to show  
 1179 that

$$\Pr[\pi(R) = (T + 1, \tilde{a}, (a_1, \dots, a_T))] \leq e^\epsilon \Pr[\pi(R') = (T + 1, \tilde{a}, (a_1, \dots, a_T))].$$

1180 Step 1: Probability decompositions: First, let us denote by  $\tau, \tilde{A}$  and  $A_1, \dots, A_\tau$  the random variables  
 1181 of stopping, recommendation and sampled actions, when  $\pi$  interacts with  $R$ . Similarly, let  $\tau', \tilde{A}'$   
 1182 and  $A'_1, \dots, A'_\tau$  the random variables of stopping, recommendation and sampled actions, when  $\pi$   
 1183 interacts with  $R'$ .

1184 We have

$$\begin{aligned}\Pr[\pi(R) = (T+1, \tilde{a}, (a_1, \dots, a_T))] &= \Pr[\tau = T+1, \tilde{A} = \tilde{a}, A_1 = a_1, \dots, A_T = a_T] \\ \Pr[\pi(R') = (T+1, \tilde{a}, (a_1, \dots, a_T))] &= \Pr[\tau' = T+1, \tilde{A}' = \tilde{a}, A'_1 = a_1, \dots, A'_T = a_T]\end{aligned}$$

1185 Since for all  $t < t^*$ ,  $x_t = x'_t$ , the policy samples the same actions, up to step  $t^*$ , i.e.

$$\Pr[A_1 = a_1, \dots, A_{t^*} = a_{t^*}] = \Pr[A'_1 = a_1, \dots, A'_{t^*} = a_{t^*}]$$

1186 And thus

$$\begin{aligned}& \frac{\Pr[\pi(R) = (T+1, \tilde{a}, (a_1, \dots, a_T))]}{\Pr[\pi(R') = (T+1, \tilde{a}, (a_1, \dots, a_T))]} \\ &= \frac{\Pr[\tau = T+1, \tilde{A} = \tilde{a}, A_{t^*+1} = a_{t^*+1}, \dots, A_T = a_T \mid A_1 = a_1, \dots, A_{t^*} = a_{t^*}]}{\Pr[\tau' = T+1, \tilde{A}' = \tilde{a}, A'_{t^*+1} = a_{t^*+1}, \dots, A'_T = a_T \mid A'_1 = a_1, \dots, A'_{t^*} = a_{t^*}]}\end{aligned}$$

1187 Let us denote by  $t_1, \dots, t_\ell$  the time step corresponding to the beginning of the phases when  $\pi$  interacts  
1188 with  $R$ , and  $t'_1, \dots, t'_{\ell'}$  the the time step corresponding to the beginning of the phases  $\pi$  interacts with  
1189  $\mathbf{r}'$ .

1190 Also, let  $t_{k^*}$  be the beginning of the phase for which  $t^*$  belongs in  $R$  phases. Similarly, let  $t'_{k'_*}$  be the  
1191 beginning of the phase for which  $t^*$  belongs in  $R'$  phases.

1192 Since the actions  $a_1, \dots, a_T$  are fixed, and  $r_t = r'_t$  for  $t < t^*$ ,  $t^*$  falls in the same phase under both  
1193  $R$  and  $R'$ . Thus,  $t_{k^*} = t'_{k'_*}$  and  $k^* = k'_*$ .

1194 Step 2: Using the structure of  $\text{GPE}_\eta(\epsilon)$

1195 Let  $\tilde{S}_{k^*}^p = \sum_{s=t_{k^*}}^{t_{k^*+1}-1} x_s + Y_{k^*}$  be the noisy partial sum of observations collected at phase  $k^*$  for  
1196  $\mathbf{r}$ , where  $Y_{k^*} \sim \text{Lap}(1/\epsilon)$ . Similarly, let  $\tilde{S}'_{k^*}^p = \sum_{s=t_{k^*}}^{t_{k^*+1}-1} x'_s + Y'_{k^*}$  be the noisy partial sum of  
1197 observations collected at phase  $k^*$  for  $\mathbf{r}'$ , where  $Y'_{k^*} \sim \text{Lap}(1/\epsilon)$ . Using the structure of  $\text{GPE}_\eta(\epsilon)$ ,  
1198 we have that:

1199 (a) If the value of the noisy partial sums at phase  $k^*$  is exactly the same between the neighbouring  
1200  $R$  and  $R'$ , then the BAI algorithm  $\pi$  will sample the same sequence of actions from step  $t^*$  onward,  
1201 recommend the same final guess and stop at the same time, with the same probability under  $R$  and  
1202  $R'$ . Thus, for any  $s \in \mathbb{R}$ :

$$\begin{aligned}\Pr[\tau = T+1, \tilde{A} = \tilde{a}, A_{t^*+1} = a_{t^*+1}, \dots, A_T = a_T \mid A_1 = a_1, \dots, A_{t^*} = a_{t^*}, \tilde{S}_{k^*}^p = s] \\ = \Pr[\tau' = T+1, \tilde{A}' = \tilde{a}, A'_{t^*+1} = a_{t^*+1}, \dots, A'_T = a_T \mid A'_1 = a_1, \dots, A'_{t^*} = a_{t^*}, \tilde{S}'_{k^*}^p = s]\end{aligned}\tag{20}$$

1203 This is due to the fact that, in  $\text{GPE}_\eta(\epsilon)$ , the reward at step  $t^*$  only affects the statistic  $\tilde{S}_{k^*}^p$ , and nothing  
1204 else.

1205 (b) Since rewards are  $[0, 1]$ , using the Laplace mechanism, we have that

$$\Pr[\tilde{S}_{k^*}^p = s \mid A_1 = a_1, \dots, A_{t^*} = a_{t^*}] \leq e^\epsilon \Pr[\tilde{S}'_{k^*}^p = s \mid A_1 = a_1, \dots, A'_{t^*} = a_{t^*}]. \tag{21}$$

1206 Step 3: Combining Eq. 20 and Eq. 21, aka post-processing:

1207 We have

$$\begin{aligned}\Pr[\tau = T+1, \tilde{A} = \tilde{a}, A_{t^*+1} = a_{t^*+1}, \dots, A_T = a_T \mid A_1 = a_1, \dots, A_{t^*} = a_{t^*}] \\ = \int_{s \in \mathbb{R}} \Pr[\tau = T+1, \tilde{A} = \tilde{a}, A_{t^*+1} = a_{t^*+1}, \dots, A_T = a_T \mid A_1 = a_1, \dots, A_{t^*} = a_{t^*}, \tilde{S}_{k^*}^p = s] \\ \Pr[\tilde{S}_{k^*}^p = s \mid A_1 = a_1, \dots, A_{t^*} = a_{t^*}] \\ \leq \int_{s \in \mathbb{R}} e^\epsilon \Pr[\tau' = T+1, \tilde{A}' = \tilde{a}, A'_{t^*+1} = a_{t^*+1}, \dots, A'_T = a_T \mid A_1 = a_1, \dots, A'_{t^*} = a_{t^*}, \tilde{S}'_{k^*}^p = s] \\ \Pr[\tilde{S}'_{k^*}^p = s \mid A_1 = a_1, \dots, A'_{t^*} = a_{t^*}]\end{aligned}$$

$$= e^\epsilon \Pr[\tau' = T + 1, \tilde{A}' = \tilde{a}, A'_{t^*+1} = a_{t^*+1}, \dots, A'_T = a_T \mid A'_1 = a_1, \dots, A'_{t^*} = a_{t^*}].$$

1208 This concludes the proof:

$$\frac{\Pr[\pi(R) = (T + 1, \tilde{a}, (a_1, \dots, a_T))]}{\Pr[\pi(R') = (T + 1, \tilde{a}, (a_1, \dots, a_T))]} \leq e^\epsilon.$$

1209

□

### 1210 E.3 Recalling the post-processing and composition properties of DP

1211 **Proposition 1** (Post-processing [30]). *Let  $\mathcal{M}$  be a mechanism and  $f$  be an arbitrary randomised*  
 1212 *function defined on  $\mathcal{M}$ 's output. If  $\mathcal{M}$  is  $\epsilon$ -DP, then  $f \circ \mathcal{M}$  is  $\epsilon$ -DP.*

1213 The post-processing property ensures that any quantity constructed only from a private output is still  
 1214 private, with the same privacy budget. This is a consequence of the data processing inequality.

**Proposition 2** (Simple Composition). *Let  $\mathcal{M}^1, \dots, \mathcal{M}^k$  be  $k$  mechanisms. We define the mechanism*

$$\mathcal{G} : D \rightarrow \bigotimes_{i=1}^k \mathcal{M}_D^i$$

1215 *as the  $k$  composition of the mechanisms  $\mathcal{M}^1, \dots, \mathcal{M}^k$ .*

1216 *If each  $\mathcal{M}^i$  is  $\epsilon_i$ -DP, then  $\mathcal{G}$  is  $\sum_{i=1}^k \epsilon_i$ -DP.*

**Proposition 3** (Parallel Composition). *Let  $\mathcal{M}^1, \dots, \mathcal{M}^k$  be  $k$  mechanisms, such that  $k < n$ , where*  
 *$n$  is the size of the input dataset. Let  $t_1, \dots, t_k, t_{k+1}$  be indexes in  $[1, n]$  such that  $1 = t_1 < \dots <$   
 $t_k < t_{k+1} - 1 = n$ .*

*Let's define the following mechanism*

$$\mathcal{G} : \{x_1, \dots, x_n\} \rightarrow \bigotimes_{i=1}^k \mathcal{M}_{\{x_{t_i}, \dots, x_{t_{i+1}-1}\}}^i$$

1217  *$\mathcal{G}$  is the mechanism that we get by applying each  $\mathcal{M}^i$  to the  $i$ -th partition of the input dataset*  
 1218  *$\{x_1, \dots, x_n\}$  according to the indexes  $t_1 < \dots < t_k < t_{k+1}$ .*

1219 *If each  $\mathcal{M}^i$  is  $\epsilon$ -DP, then  $\mathcal{G}$  is  $\epsilon$ -DP.*

1220 In parallel composition, the  $k$  mechanisms are applied to different “non-overlapping” parts of the  
 1221 input dataset. If each mechanism is DP, then the parallel composition of the  $k$  mechanisms is DP, *with*  
 1222 *the same privacy budget*. This property will be the basis for designing private bandit algorithms.

## 1223 F Concentration Results

1224 In Appendix F, we detail the proof of all our concentration results. In Appendix F.1, we start by  
 1225 introducing a variant of GLR-based stopping rule using the modified transportation costs  $\widetilde{W}_{\epsilon, a, b}$   
 1226 (see Appendix G.2.1 for details) which are defined based on the modified divergences  $\widetilde{d}_\epsilon^\pm$  (see  
 1227 Appendix G.1.1 for details). The proof of Theorem 5 is given in Appendix F.2. In Appendix F.3, we  
 1228 show tail bounds for a sum between independent Bernoulli and Laplace observations that feature the  
 1229 product of the tail bounds of each process. We prove time-uniform and fixed-time tails concentration  
 1230 for Laplace distribution in Appendix F.4, and recall existing results for Bernoulli in Appendix F.5.  
 1231 In Appendix F.6, we provide tail bounds for a sum between independent Bernoulli and Laplace  
 1232 observations that feature the modified divergence  $\widetilde{d}_\epsilon$  defined in Eq. (32). In Appendix F.7, we give  
 1233 geometric grid time uniform tails concentration for the reweighted modified divergence.



## 1234 F.1 Modified GLR Stopping Rule

1235 The modified GLR stopping rule is defined as

$$\tau_{\epsilon, \delta}^{\text{MGLR}} = \inf \left\{ n \mid \forall a \neq \tilde{a}_n, \widetilde{W}_{\epsilon, \tilde{a}_n, a}(\tilde{\mu}_n, \tilde{N}_n) > \sum_{b \in \{\tilde{a}_n, a\}} \tilde{c}(k_{n,b}, \delta) \right\} \text{ with } \tilde{a}_n \in \arg \max_{a \in [K]} [\tilde{\mu}_{n,a}]_0^1, \quad (22)$$

1236 where  $(\tilde{\mu}_n, \tilde{N}_n)$  are the outputs of  $\text{GPE}_\eta(\epsilon)$ . The modified transportation costs  $(\widetilde{W}_{\epsilon, a, b})_{(a,b) \in [K]^2}$  are  
 1237 defined in Eq. (34), i.e., for all  $(\mu, w) \in \mathbb{R}^K \times \mathbb{R}_+^K$  and all  $(a, b) \in [K]^2$  such that  $a \neq b$ ,

$$\widetilde{W}_{\epsilon, a, b}(\mu, w) := \mathbf{1}([\mu_a]_0^1 > [\mu_b]_0^1) \inf_{u \in (0,1)} \left\{ w_a \tilde{d}_\epsilon^-(\mu_a, u, r(w_a)) + w_b \tilde{d}_\epsilon^+(\mu_b, u, r(w_b)) \right\},$$

1238 where  $r(x) := \frac{x}{1 + \log_{1+\eta} x}$  is defined in Eq. (33) for all  $x \geq 1$ . The modified divergence  $\tilde{d}_\epsilon^\pm$  are  
 1239 defined in Eq. (32), i.e., for all  $(\lambda, \mu, r) \in \mathbb{R} \times (0, 1) \times \mathbb{R}_+^*$ ,

$$\begin{aligned} \tilde{d}_\epsilon^-(\lambda, \mu, r) &:= \mathbf{1}(\mu < [\lambda]_0^1) \inf_{z \in (\mu, [\lambda]_0^1)} \left\{ \text{kl}(z, \mu) + \frac{1}{r} h(r\epsilon(\lambda - z)) \right\}, \\ \tilde{d}_\epsilon^+(\lambda, \mu, r) &:= \mathbf{1}(\mu > [\lambda]_0^1) \inf_{z \in ([\lambda]_0^1, \mu)} \left\{ \text{kl}(z, \mu) + \frac{1}{r} h(r\epsilon(z - \lambda)) \right\}, \end{aligned}$$

1240 where  $h(x) := \sqrt{1 + x^2} - 1 + \log\left(\frac{2}{x^2} (\sqrt{1 + x^2} - 1)\right)$  is defined in Eq. (31) for all  $x > 0$ .

1241 Lemma 8 gives a stopping threshold under which the modified GLR stopping rule is  $\delta$ -correct.

1242 **Lemma 8.** *Let  $\delta \in (0, 1)$  and  $\epsilon > 0$ . Let  $\eta > 0$ . Let  $s > 1$  and  $\zeta$  be the Riemann  $\zeta$  function.  
 1243 Let  $\overline{W}_{-1}(x) = -W_{-1}(-e^{-x})$  for all  $x \geq 1$ , where  $W_{-1}$  is the negative branch of the Lambert  
 1244  $W$  function. It satisfies  $\overline{W}_{-1}(x) \approx x + \log x$ , see Lemma 51. Given any sampling rule using the  
 1245  $\text{GPE}_\eta(\epsilon)$ , combining  $\text{GPE}_\eta(\epsilon)$  with the modified GLR stopping rule as in Eq. (22) with the stopping  
 1246 threshold*

$$\tilde{c}(k, \delta) = \overline{W}_{-1} \left( \log \left( \frac{K \zeta(s)}{\delta} \right) + s \log(k) + 3 - \log 2 \right) - 3 + \log 2. \quad (23)$$

1247 yields a  $\delta$ -correct and  $\epsilon$ -global DP algorithm for Bernoulli instances with unique best arm.

1248 *Proof.* Lemma 4 yields the  $\epsilon$ -global DP. Let  $\mathcal{E}_\delta = \mathcal{E}_{\delta, a^*, +} \cap \bigcap_{a \neq a^*} \mathcal{E}_{\delta, a, -}$  with

$$\begin{aligned} \mathcal{E}_{\delta, a^*, +} &= \left\{ \forall n \in \mathbb{N}, \tilde{N}_{n, a^*} \tilde{d}_\epsilon^+(\tilde{\mu}_{n, a^*}, \mu_{a^*}, \tilde{N}_{n, a^*}/k_{n, a^*}) \leq \tilde{c}(k_{n, a^*}, \delta) \right\}, \\ \mathcal{E}_{\delta, a, -} &= \left\{ \forall n \in \mathbb{N}, \tilde{N}_{n, a} \tilde{d}_\epsilon^-(\tilde{\mu}_{n, a}, \mu_a, \tilde{N}_{n, a}/k_{n, a}) \leq \tilde{c}(k_{n, a}, \delta) \right\}, \end{aligned}$$

1249 where  $(\tilde{\mu}_n, \tilde{N}_n, k_n)$  are given by  $\text{GPE}_\eta(\epsilon)$ ,  $\tilde{c}$  as in Eq. (23) and  $\tilde{d}_\epsilon^\pm$  as in Eq. (32).

1250 Using Lemmas 19 and 20, we have  $\mathbb{P}_{\nu\pi}(\mathcal{E}_{\delta, a, -}^c) \leq \delta/K$  for all  $a \neq a^*$ , and  $\mathbb{P}_{\nu\pi}(\mathcal{E}_{\delta, a^*, +}^c) \leq \delta/K$ .

1251 By union bound over  $a \in [K]$ , we obtain  $\mathbb{P}_{\nu\pi}(\mathcal{E}_\delta^c) \leq \delta$ .

1252 Let  $\tau_{\epsilon, \delta}^{\text{MGLR}}$  as in Eq. (22) and  $\tilde{a}_n \in \arg \max_{a \in [K]} [\tilde{\mu}_{n,a}]_0^1$ . Then, we directly have that

$$\begin{aligned} \mathbb{P}_{\nu\pi} \left( \tau_{\epsilon, \delta}^{\text{MGLR}} < +\infty, \tilde{a}_{\tau_{\epsilon, \delta}^{\text{MGLR}}} \neq a^* \right) &\leq \mathbb{P}_{\nu\pi}(\mathcal{E}_\delta^c) + \mathbb{P}_{\nu\pi} \left( \mathcal{E}_\delta \cap \{ \tau_{\epsilon, \delta}^{\text{MGLR}} < +\infty, \tilde{a}_{\tau_{\epsilon, \delta}^{\text{MGLR}}} \neq a^* \} \right) \\ &\leq \delta + \mathbb{P}_{\nu\pi} \left( \mathcal{E}_\delta \cap \{ \tau_{\epsilon, \delta}^{\text{MGLR}} < +\infty, \tilde{a}_{\tau_{\epsilon, \delta}^{\text{MGLR}}} \neq a^* \} \right). \end{aligned}$$

1253 Under  $\mathcal{E}_\delta \cap \{ \tau_{\epsilon, \delta}^{\text{MGLR}} < +\infty, \tilde{a}_{\tau_{\epsilon, \delta}^{\text{MGLR}}} \neq a^* \}$ , by definition of the stopping rule as in Eq. (7) and the

1254 stopping threshold in Eq. (8), we obtain that there exists  $a \neq a^*$  and  $n \in \mathbb{N}$  such that  $[\tilde{\mu}_{n,a}]_0^1 >$

1255  $[\tilde{\mu}_{n,a^*}]_0^1$  and

$$\sum_{b \in \{a, a^*\}} \tilde{c}(k_{n,b}, \delta) < \widetilde{W}_{\epsilon, a, a^*}(\tilde{\mu}_n, \tilde{N}_n)$$

$$\begin{aligned}
&= \inf_{u \in (0,1)} \left\{ \tilde{N}_{n,a} \tilde{d}_\epsilon^- (\tilde{\mu}_{n,a}, u, r(\tilde{N}_{n,a})) + \tilde{N}_{n,a^*} \tilde{d}_\epsilon^+ (\tilde{\mu}_{n,a^*}, u, r(\tilde{N}_{n,a^*})) \right\} \\
&= \inf_{(u_a, u_{a^*}) \in (0,1)^2, u_a \leq u_{a^*}} \left\{ \tilde{N}_{n,a} \tilde{d}_\epsilon^- (\tilde{\mu}_{n,a}, u_a, r(\tilde{N}_{n,a})) + \tilde{N}_{n,a^*} \tilde{d}_\epsilon^+ (\tilde{\mu}_{n,a^*}, u_{a^*}, r(\tilde{N}_{n,a^*})) \right\} \\
&\leq \tilde{N}_{n,a} \tilde{d}_\epsilon^- (\tilde{\mu}_{n,a}, \mu_a, r(\tilde{N}_{n,a})) + \tilde{N}_{n,a^*} \tilde{d}_\epsilon^+ (\tilde{\mu}_{n,a^*}, \mu_{a^*}, r(\tilde{N}_{n,a^*})) \\
&\leq \tilde{N}_{n,a} \tilde{d}_\epsilon^- (\tilde{\mu}_{n,a}, \mu_a, \tilde{N}_{n,a}/k_{n,a}) + \tilde{N}_{n,a^*} \tilde{d}_\epsilon^+ (\tilde{\mu}_{n,a^*}, \mu_{a^*}, \tilde{N}_{n,a^*}/k_{n,a^*}) \leq \sum_{b \in \{a, a^*\}} \tilde{c}(k_{n,b}, \delta),
\end{aligned}$$

1256 where we used the definition of  $\tilde{W}_{\epsilon, a, a^*}$  in Eq. (34) and Lemma 39 in the two equalities and  $\mu_{a^*} > \mu_a$   
1257 in the following inequality. The second to last inequality uses that  $r(\tilde{N}_{n,a}) \leq \tilde{N}_{n,a}/k_{n,a}$  for all  
1258  $a \in [K]$  by definition of  $(k_n, \tilde{N}_n)$ , i.e.,  $k_{n,a} \leq 1 + \log_{1+\eta} \tilde{N}_{n,a} \leq k_{n,a} + 1$ , and that  $r \mapsto \tilde{d}_\epsilon^\pm(\lambda, u, r)$   
1259 is non-decreasing, see Lemmas 30 and 31. The last inequality is obtained by the concentration event  
1260  $\mathcal{E}_\delta$ . Since this yields a contradiction, we obtain  $\mathcal{E}_\delta \cap \{\tau_{\epsilon, \delta}^{\text{MGLR}} < +\infty, \tilde{a}_{\tau_{\epsilon, \delta}^{\text{MGLR}}} \neq a^*\} = \emptyset$ . This  
1261 concludes the proof, i.e.,  $\mathbb{P}_{\nu\pi}(\tau_{\epsilon, \delta}^{\text{MGLR}} < +\infty, \tilde{a}_{\tau_{\epsilon, \delta}^{\text{MGLR}}} \neq a^*) \leq \delta$ .  $\square$

## 1262 F.2 Proof of Theorem 5

1263 Lemma 4 yields the  $\epsilon$ -global DP. The proof of  $\delta$ -correctness is the same as the one of Lemma 8  
1264 detailed above. In particular, we use the same concentration event  $\mathcal{E}_\delta = \mathcal{E}_{\delta, a^*, +} \cap \bigcap_{a \neq a^*} \mathcal{E}_{\delta, a, -}$  that  
1265 satisfies  $\mathbb{P}_{\nu\pi}(\mathcal{E}_\delta^c) \leq \delta$ .

1266 Under  $\mathcal{E}_\delta \cap \{\tau_{\epsilon, \delta} < +\infty, \tilde{a}_{\tau_{\epsilon, \delta}} \neq a^*\}$ , by definition of the GLR stopping rule as in Eq. (7)  
1267 and the stopping threshold in Eq. (8), we obtain that there exists  $a \neq a^*$  and  $n \in \mathbb{N}$  such that  
1268  $[\tilde{\mu}_{n,a}]_0^1 > [\tilde{\mu}_{n,a^*}]_0^1$ ,

$$\sum_{b \in \{a, a^*\}} \left( c_1(\tilde{N}_{n,b}, \delta) + c_2(\tilde{N}_{n,b}, \epsilon) \right) = \sum_{b \in \{a, a^*\}} c(k_{n,b}, \epsilon, \delta) < W_{\epsilon, a, a^*}(\tilde{\mu}_n, \tilde{N}_n).$$

1269 Then, we obtain

$$\begin{aligned}
W_{\epsilon, a, a^*}(\tilde{\mu}_n, \tilde{N}_n) &= \inf_{u \in [0,1]} \left\{ \tilde{N}_{n,a} d_\epsilon^- (\tilde{\mu}_{n,a}, u) + \tilde{N}_{n,a^*} d_\epsilon^+ (\tilde{\mu}_{n,a^*}, u) \right\} \\
&= \inf_{(u_a, u_{a^*}) \in [0,1]^2, u_a \leq u_{a^*}} \left\{ \tilde{N}_{n,a} d_\epsilon^- (\tilde{\mu}_{n,a}, u_a) + \tilde{N}_{n,a^*} d_\epsilon^+ (\tilde{\mu}_{n,a^*}, u_{a^*}) \right\} \\
&\leq \tilde{N}_{n,a} d_\epsilon^- (\tilde{\mu}_{n,a}, \mu_a) + \tilde{N}_{n,a^*} d_\epsilon^+ (\tilde{\mu}_{n,a^*}, \mu_{a^*}),
\end{aligned}$$

1270 where we used the definition of  $W_{\epsilon, a, a^*}$  in Eq. (4) and Lemma 34 in the two equalities, and  
1271  $(u_{a^*}, u_a) = (\mu_{a^*}, \mu_a) \in [0,1]^2$  that satisfies  $u_{a^*} > u_a$  in the following inequality.

1272 Using Lemma 38 and initialization yields  $\min\{r(\tilde{N}_{n,a^*}), r(\tilde{N}_{n,a})\} > 0$  by . When  $[\tilde{\mu}_{n,a}]_0^1 > \mu_a$   
1273 and  $\mu_{a^*} > [\tilde{\mu}_{n,a^*}]_0^1$ , Lemma 29 yields

$$\begin{aligned}
&\tilde{N}_{n,a^*} (d_\epsilon^+ (\tilde{\mu}_{n,a^*}, \mu_{a^*}) - \tilde{d}_\epsilon^+ (\tilde{\mu}_{n,a^*}, \mu_{a^*}, r(\tilde{N}_{n,a^*}))) \leq k_\eta(\tilde{N}_{n,a^*}) (\log(1 + 2\epsilon \frac{\tilde{N}_{n,a^*}}{k_\eta(\tilde{N}_{n,a^*})}) + 1) \\
&\tilde{N}_{n,a} (d_\epsilon^- (\tilde{\mu}_{n,a}, \mu_a) - \tilde{d}_\epsilon^- (\tilde{\mu}_{n,a}, \mu_a, r(\tilde{N}_{n,a}))) \leq k_\eta(\tilde{N}_{n,a}) \left( \log \left( 1 + 2\epsilon \frac{\tilde{N}_{n,a}}{k_\eta(\tilde{N}_{n,a})} \right) + 1 \right),
\end{aligned}$$

1274 where we used that  $(\mu_a, \mu_{a^*}) \in (0,1)^2$  and  $x/r(x) = 1 + \log_{1+\eta} x = k_\eta(x)$ . When  $[\tilde{\mu}_{n,a}]_0^1 \leq \mu_a$ , we  
1275 have  $d_\epsilon^- (\tilde{\mu}_{n,a}, \mu_a) = 0 = \tilde{d}_\epsilon^- (\tilde{\mu}_{n,a}, \mu_a, r(\tilde{N}_{n,a}))$ . When  $\mu_{a^*} \leq [\tilde{\mu}_{n,a^*}]_0^1$ , we have  $d_\epsilon^+ (\tilde{\mu}_{n,a^*}, \mu_{a^*}) =$   
1276  $0 = \tilde{d}_\epsilon^+ (\tilde{\mu}_{n,a^*}, \mu_{a^*}, r(\tilde{N}_{n,a^*}))$ . In either case, the above inequalities are still valid since the left hand  
1277 side is null and the right hand side is positive. Therefore, we have

$$\begin{aligned}
&\tilde{N}_{n,a} d_\epsilon^- (\tilde{\mu}_{n,a}, \mu_a) + \tilde{N}_{n,a^*} d_\epsilon^+ (\tilde{\mu}_{n,a^*}, \mu_{a^*}) \\
&\leq \tilde{N}_{n,a} \tilde{d}_\epsilon^- (\tilde{\mu}_{n,a}, \mu_a, r(\tilde{N}_{n,a})) + \tilde{N}_{n,a^*} \tilde{d}_\epsilon^+ (\tilde{\mu}_{n,a^*}, \mu_{a^*}, r(\tilde{N}_{n,a^*})) + \sum_{b \in \{a, a^*\}} c_2(\tilde{N}_{n,b}, \epsilon)
\end{aligned}$$

$$\leq \sum_{b \in \{a, a^*\}} \tilde{c}(k_{n,b}, \delta) + \sum_{b \in \{a, a^*\}} c_2(\tilde{N}_{n,b}, \epsilon) \leq \sum_{b \in \{a, a^*\}} \left( c_1(\tilde{N}_{n,b}, \delta) + c_2(\tilde{N}_{n,b}, \epsilon) \right),$$

1278 where the second inequality uses the proof of Lemma 8, and third leverages that

$$\tilde{c}(k_{n,a}, \delta) \leq \bar{W}_{-1} \left( \log \left( \frac{K\zeta(s)}{\delta} \right) + s \log(k_\eta(\tilde{N}_{n,a})) + 3 - \log 2 \right) - 3 + \log 2,$$

1279 by using that  $\bar{W}_{-1}$  is increasing (Lemma 51) and  $k_{n,a} \leq 1 + \log_{1+\eta} \tilde{N}_{n,a} = k_\eta(\tilde{N}_{n,a})$  for all  
1280  $a \in [K]$ , as well as  $r(x) = x/k_\eta(x)$ . Combining all the above inequalities, we have shown that

$$\sum_{b \in \{a, a^*\}} (c_1(\tilde{N}_{n,b}, \delta) + c_2(\tilde{N}_{n,b}, \epsilon)) < W_{\epsilon, a, a^*}(\tilde{\mu}_n, \tilde{N}_n) \leq \sum_{b \in \{a, a^*\}} (c_1(\tilde{N}_{n,b}, \delta) + c_2(\tilde{N}_{n,b}, \epsilon)).$$

1281 This yields a contradiction, hence we have  $\mathcal{E}_\delta \cap \{\tau_{\epsilon, \delta} < +\infty, \tilde{a}_{\tau_{\epsilon, \delta}} \neq a^*\} = \emptyset$ . This concludes the  
1282 proof, i.e.,  $\mathbb{P}_{\nu_\pi}(\tau_{\epsilon, \delta} < +\infty, \tilde{a}_{\tau_{\epsilon, \delta}} \neq a^*) \leq \delta$ .

### 1283 F.3 Fixed Time Tails Bounds for a Convolution of Probability Distributions

1284 We derive general upper and lower bounds on the upper and lower tails of the convolution (i.e., sum)  
1285 between two independent random variables (Lemma 9). We provide upper (Lemma 11) and lower  
1286 (Lemma 11) tail bounds for a sum (i.e., convolution) between independent Bernoulli and Laplace  
1287 i.i.d. observations for a fixed time. The bounds are expressed as a function of the infimum over a  
1288 bounded interval of a  $-\frac{1}{t} \log(\cdot)$  transform of the product between the (upper or lower) tail bounds of  
1289 each process. Therefore, in Lemmas 11 and 11, we can plug any bounds on the (upper or lower) tail  
1290 concentration of each process. While those bounds are standard for Bernoulli distribution (Lemma 16  
1291 in Appendix F.5), we propose new bounds for Laplace distribution (Lemma 15 in Appendix F.4).

1292 **Sketch of Proof of Lemma 9** The main difficulty when studying the sum of two random variables  
1293 lies in the fact that it involves the integral of the convolution of their probability measures. In all  
1294 generality, it is difficult to upper bound such a quantity. The main idea behind our proof technique  
1295 is to split the event of interest into a partition of carefully chosen events. Then, on each of those  
1296 smaller events, we derive a "tight" upper bound on the integral of the convolution of their probability  
1297 measures. It is reasonable to wonder how one could choose those events such that the upper bound  
1298 is easier to obtain. When the event is defined as the intersection of two independent events, then  
1299 we obtain a straightforward upper bound by the product of their respective probabilities. When the  
1300 event truly mixes the distributions, we need to use a smarter approach to control the integrated  
1301 function. First, we upper bound a sub-component of this function by a maximum of the product of  
1302 their respective probabilities (on a small interval that is defined by the smaller event). Second, after  
1303 this upper bound, the integrated function coincides with the hazard function, whose integral is the  
1304 cumulative hazard function. To conclude the proof, it only remains to merge together the different  
1305 upper bounds.

1306 To the best of our knowledge, the proof technique closest to ours is the one used to prove Lemma 64  
1307 in Jourdan et al. [50]. They control the probability that two random variables have an unexpected  
1308 empirical ranking as a function of the boundary crossing probabilities of each random variable.  
1309 While tackling a distinct problem, they adopt the same proof structure. They decompose the event  
1310 into carefully chosen events on which they can upper bound the integral of the convolution of their  
1311 probability distributions. The upper bounds are obtained similarly as ours, with fewer events to  
1312 consider.

1313 Lemma 9 gives upper and lower bounds on the upper and lower tails of the sum of two independent  
1314 random variables. This result is of independent interest.

1315 **Lemma 9.** *Let  $\theta$  and  $\lambda$  be two independent real random variables such that (i)  $\theta$  has bounded support*  
1316 *included in  $[\alpha, \beta]$  and mean  $\mu \in (\alpha, \beta)$  and (ii)  $\lambda$  has zero mean. Let*

$$\forall u \in [0, 1], \forall v \in (0, 1], \quad p(u, v) := u(1 - \log(u) + \log(v)).$$

1317 *Then, for all  $x > 0$ , we have*

$$\mathbb{P}(\theta + \lambda \geq \mu + x) \leq \mathbb{P}(\lambda \geq x) \mathbb{P}(\theta \in [\alpha, \mu]) + \mathbb{P}(\lambda \leq 0) \mathbb{P}(\theta \in [\min\{\beta, \mu + x\}, \beta])$$

$$\begin{aligned}
& + p \left( \sup_{z \in (\mu, \min\{\beta, \mu+x\})} \{\mathbb{P}(\theta \in [z, \beta])\mathbb{P}(\lambda \geq \mu+x-z)\}, \mathbb{P}(\theta \in [\mu, \beta])\mathbb{P}(\lambda \geq 0) \right), \\
\mathbb{P}(\theta + \lambda \geq \mu + x) & \geq \sup_{z \in (\mu, \min\{\beta, \mu+x\})} \{\mathbb{P}(\theta \in [z, \beta])\mathbb{P}(\lambda \geq \mu+x-z)\}, \\
\mathbb{P}(\theta + \lambda \leq \mu - x) & \leq \mathbb{P}(\lambda \leq -x)\mathbb{P}(\theta \in [\mu, \beta]) + \mathbb{P}(\lambda \geq 0)\mathbb{P}(\theta \in [\alpha, \max\{\alpha, \mu-x\}]) \\
& p \left( \sup_{z \in (\max\{\alpha, \mu-x\}, \mu)} \{\mathbb{P}(\theta \in [\alpha, z])\mathbb{P}(\lambda \leq \mu-x-z)\}, \mathbb{P}(\theta \in [\alpha, \mu])\mathbb{P}(\lambda \leq 0) \right), \\
\mathbb{P}(\theta + \lambda \leq \mu - x) & \geq \sup_{z \in (\max\{\alpha, \mu-x\}, \mu)} \{\mathbb{P}(\theta \in [\alpha, z])\mathbb{P}(\lambda \leq \mu-x-z)\}.
\end{aligned}$$

1318 **Proof. I. Upper Bound on Upper Tail.** We start by studying  $\mathbb{P}(\theta + \lambda \geq \mu + x)$  where  $x > 0$ . We can  
1319 suppose that there exists  $y_1 \in (\max\{x + \mu - \beta, 0\}, x)$  such that  $\mathbb{P}(\theta \geq x + \mu - y_1)\mathbb{P}(\lambda \geq y_1) > 0$ .  
1320 Otherwise, the probability of  $\{\theta + \lambda \geq \mu + x\}$  is 0, and both bounds are 0 as well. Let  $y_1$  be such a  
1321 value, and

$$y_3 \in [x, x + \mu) \quad \text{and} \quad y_2 \in (\min\{x + \mu - \beta, 0\}, 0].$$

1322 First, we note that  $-\log \mathbb{P}(\theta \geq x + \mu - y_1)$  and  $-\log \mathbb{P}(\lambda \geq y_1)$  are finite, since  $\mathbb{P}(\theta \geq x + \mu -$   
1323  $y_1)\mathbb{P}(\lambda \geq y_1) > 0$  implies that  $\min\{\mathbb{P}(\theta \geq x + \mu - y_1), \mathbb{P}(\lambda \geq y_1)\} > 0$ . Second, we note that  $y_2$   
1324 only exists when  $x + \mu < \beta$ , i.e.,  $(\min\{x + \mu - \beta, 0\}, 0] \neq \emptyset$ . In order to study the cases  $x + \mu < \beta$   
1325 and  $x + \mu \geq \beta$  simultaneously, we adopt the convention that the maximum of a positive quantity on  
1326 an empty set is defined as zero. Note that the situation  $x + \mu < \beta$  has more subcases.

1327 We partition the event  $\{\theta + \lambda \geq \mu + x\}$  into eight sets, namely

$$\begin{aligned}
\{\theta + \lambda \geq \mu + x, \theta \in [\alpha, \beta], \lambda \in \mathbb{R}\} & = \{\lambda \in (\max\{x + \mu - \beta, 0\}, y_1), \theta \in [x + \mu - \lambda, \beta]\} \\
& \cup \{\theta \in [x + \mu - y_1, \beta], \lambda \geq y_1\} \\
& \cup \{\theta \in (\mu, x + \mu - y_1), \lambda \geq x + \mu - \theta\} \\
& \cup \{\theta \in [x + \mu - y_3, \mu], \lambda \geq x + \mu - \theta\} \\
& \cup \{\theta \in [\alpha, x + \mu - y_3], \lambda \geq x + \mu\} \\
& \cup \{\lambda \in [y_3, x + \mu), \theta \in [x + \mu - \lambda, x + \mu - y_3)\} \\
& \cup \{\lambda \in [y_2, 0], \theta \in [x + \mu - \lambda, \beta]\} \\
& \cup \{\lambda \in [x + \mu - \theta, y_2), \theta \in [x + \mu - y_2, \beta]\}.
\end{aligned}$$

1328 First, it is direct to see that

$$\begin{aligned}
& \{\lambda \in [y_2, 0], \theta \in [x + \mu - \lambda, \beta]\} \cup \{\lambda \in [x + \mu - \theta, y_2), \theta \in [x + \mu - y_2, \beta]\} \\
& \subseteq \{\lambda \leq 0, \theta \in [\min\{\beta, \mu + x\}, \beta]\}, \\
& \{\theta \in [x + \mu - y_3, \mu], \lambda \geq x + \mu - \theta\} \cup \{\theta \in [\alpha, x + \mu - y_3], \lambda \geq x + \mu\} \\
& \cup \{\lambda \in [y_3, x + \mu), \theta \in [x + \mu - \lambda, x + \mu - y_3)\} \subseteq \{\lambda \geq x, \theta \in [\alpha, \mu]\}.
\end{aligned}$$

1329 By union bound, the probability of the union of those five events is upper bounded by the sum of the  
1330 probability of those two events, i.e.,  $\mathbb{P}(\lambda \geq x, \theta \in [\alpha, \mu]) + \mathbb{P}(\lambda \leq 0, \theta \in [\min\{\beta, \mu + x\}, \beta])$ .

1331 **A. Separate Conditions.** Those two events and one of the three remaining do not require to control  
1332  $(\theta, \lambda)$  simultaneously, as they separate the conditions on  $(\theta, \lambda)$ . Thanks to the independence of  
1333  $(\theta, \lambda)$ , the probability of those events can be simply upper bounded by the product of the respective  
1334 probability of those conditions. Therefore, we obtain

$$\begin{aligned}
& \mathbb{P}(\lambda \geq x, \theta \in [\alpha, \mu]) + \mathbb{P}(\lambda \leq 0, \theta \in [\min\{\beta, \mu + x\}, \beta]) + \mathbb{P}(\theta \in [x + \mu - y_1, \beta], \lambda \geq y_1) \\
& = \mathbb{P}(\lambda \geq x)\mathbb{P}(\theta \in [\alpha, \mu]) + \mathbb{P}(\lambda \leq 0)\mathbb{P}(\theta \in [\min\{\beta, \mu + x\}, \beta]) \\
& \quad + \mathbb{P}(\theta \in [x + \mu - y_1, \beta])\mathbb{P}(\lambda \geq y_1).
\end{aligned}$$

1335 **B. Mixed Conditions.** The two remaining events truly require to control  $(\theta, \lambda)$  simultaneously,  
1336 i.e., consider their convolution. The proof idea is the following: (1) we integrate one integral to  
1337 obtain one survival function, (2) we make appear the other survival function artificially, (3) we upper  
1338 bound the product of their survival functions on the whole set and (4) we integrate the remaining  
1339 hazard function, whose integral is the cumulative hazard function. Let  $dG$  and  $dF$  be the probability  
1340 measures of  $\theta$  and  $\lambda$  on  $\mathbb{R}$ .

1341 For all  $s \in (\max\{x + \mu - \beta, 0\}, y_1)$ , we have  $\mathbb{P}(\lambda \geq s) \geq \mathbb{P}(\lambda \geq y_1) > 0$ . Then, we obtain

$$\begin{aligned}
& \mathbb{P}(\lambda \in (\max\{x + \mu - \beta, 0\}, y_1), \theta \in [x + \mu - \lambda, \beta]) \\
&= \int_{s \in (\max\{x + \mu - \beta, 0\}, y_1)} \mathbb{P}(\theta \in [x + \mu - s, \beta]) dF(s) \\
&= \int_{s \in (\max\{x + \mu - \beta, 0\}, y_1)} \mathbb{P}(\lambda \geq s) \mathbb{P}(\theta \in [x + \mu - s, \beta]) \frac{1}{\mathbb{P}(\lambda \geq s)} dF(s) \\
&\leq \sup_{s \in (\max\{x + \mu - \beta, 0\}, y_1)} \{\mathbb{P}(\lambda \geq s) \mathbb{P}(\theta \in [x + \mu - s, \beta])\} \int_{s \in (\max\{x + \mu - \beta, 0\}, y_1)} \frac{1}{\mathbb{P}(\lambda \geq s)} dF(s) \\
&\leq \sup_{s \in (\max\{x + \mu - \beta, 0\}, y_1)} \{\mathbb{P}(\lambda \geq s) \mathbb{P}(\theta \in [x + \mu - s, \beta])\} (-\log(\mathbb{P}(\lambda \geq y_1)) + \log(\mathbb{P}(\lambda \geq 0))) ,
\end{aligned}$$

1342 where we used that  $\mathbb{P}(\lambda \geq \max\{x + \mu - \beta, 0\}) \leq \mathbb{P}(\lambda \geq 0)$ .

1343 For all  $z \in (\mu, x + \mu - y_1)$ , we have  $\mathbb{P}(\theta \in [z, \beta]) \geq \mathbb{P}(\theta \in [x + \mu - y_1, \beta]) > 0$ . Then, we obtain

$$\begin{aligned}
& \mathbb{P}(\theta \in (\mu, x + \mu - y_1), \lambda \geq x + \mu - \theta) \\
&= \int_{z \in (\mu, x + \mu - y_1)} \mathbb{P}(\lambda \geq x + \mu - z) dG(z) \\
&= \int_{z \in (\mu, x + \mu - y_1)} \mathbb{P}(\lambda \geq x + \mu - z) \mathbb{P}(\theta \in [z, \beta]) \frac{1}{\mathbb{P}(\theta \in [z, \beta])} dG(z) \\
&\leq \sup_{z \in (\mu, x + \mu - y_1)} \{\mathbb{P}(\lambda \geq x + \mu - z) \mathbb{P}(\theta \in [z, \beta])\} \int_{z \in (\mu, x + \mu - y_1)} \frac{1}{\mathbb{P}(\theta \in [z, \beta])} dG(z) \\
&\leq \sup_{s \in (y_1, x)} \{\mathbb{P}(\lambda \geq s) \mathbb{P}(\theta \in [x + \mu - s, \beta])\} \\
&\quad \cdot (-\log(\mathbb{P}(\theta \in [x + \mu - y_1, \beta])) + \log(\mathbb{P}(\theta \in [\mu, \beta]))) .
\end{aligned}$$

1344 **C. Combining Results.** Putting everything together, we have, for all  $y_1 \in (\max\{x + \mu - \beta, 0\}, x)$ ,

$$\begin{aligned}
& \mathbb{P}(\theta + \lambda \geq \mu + x) \leq \mathbb{P}(\lambda \geq x) \mathbb{P}(\theta \in [\alpha, \mu]) + \mathbb{P}(\lambda \leq 0) \mathbb{P}(\theta \in [\min\{\beta, \mu + x\}, \beta]) \\
&+ \mathbb{P}(\theta \in [x + \mu - y_1, \beta]) \mathbb{P}(\lambda \geq y_1) \\
&+ \sup_{s \in (\max\{x + \mu - \beta, 0\}, y_1)} \{\mathbb{P}(\lambda \geq s) \mathbb{P}(\theta \in [x + \mu - s, \beta])\} (-\log(\mathbb{P}(\lambda \geq y_1)) + \log(\mathbb{P}(\lambda \geq 0))) \\
&+ \sup_{s \in (y_1, x)} \{\mathbb{P}(\lambda \geq s) \mathbb{P}(\theta \in [x + \mu - s, \beta])\} (-\log(\mathbb{P}(\theta \in [x + \mu - y_1, \beta])) + \log(\mathbb{P}(\theta \in [\mu, \beta]))) \\
&\leq \mathbb{P}(\lambda \geq x) \mathbb{P}(\theta \in [\alpha, \mu]) + \mathbb{P}(\lambda \leq 0) \mathbb{P}(\theta \in [\min\{\beta, \mu + x\}, \beta]) \\
&+ \mathbb{P}(\theta \in [x + \mu - y_1, \beta]) \mathbb{P}(\lambda \geq y_1) \\
&+ \sup_{s \in (\max\{x + \mu - \beta, 0\}, x)} \{\mathbb{P}(\lambda \geq s) \mathbb{P}(\theta \in [x + \mu - s, \beta])\} \\
&\quad \cdot (-\log(\mathbb{P}(\lambda \geq y_1) \mathbb{P}(\theta \in [x + \mu - y_1, \beta])) + \log(\mathbb{P}(\theta \in [\mu, \beta]) \mathbb{P}(\lambda \geq 0))) ,
\end{aligned}$$

1345 where the second inequality is obtained by extending the two suprema to  $(\max\{x + \mu - \beta, 0\}, x)$ ,

1346 which is possible since multiplied by a positive value, and factorizing them together. Taking

$$y_1^* \in \arg \max_{s \in (\max\{x + \mu - \beta, 0\}, x)} \{\mathbb{P}(\lambda \geq s) \mathbb{P}(\theta \in [x + \mu - s, \beta])\} ,$$

1347 and the change of variable  $z = x + \mu - s$ , i.e.,

$$\begin{aligned}
& \sup_{s \in (\max\{x + \mu - \beta, 0\}, x)} \{\mathbb{P}(\lambda \geq s) \mathbb{P}(\theta \in [x + \mu - s, \beta])\} \\
&= \sup_{z \in (\mu, \min\{\beta, x + \mu\})} \{\mathbb{P}(\theta \in [z, \beta]) \mathbb{P}(\lambda \geq x + \mu - z)\} ,
\end{aligned}$$

1348 concludes the proof of the upper bound on the upper tail.

1349 **II. Lower Bound on Upper Tail.** Let  $z \in (\mu, \min\{\beta, \mu + x\})$ . Then, we have directly that

$$\{\theta \in [z, \beta], \lambda \geq \mu + x - z\} \subseteq \{\theta + \lambda \geq \mu + x\} .$$

1350 Using independence, we obtain

$$\mathbb{P}(\theta \in [z, \beta])\mathbb{P}(\lambda \geq \mu + x - z) = \mathbb{P}(\theta \in [z, \beta], \lambda \geq \mu + x - z) \leq \mathbb{P}(\theta + \lambda \geq \mu + x) .$$

1351 Taking the supremum over  $z \in (\mu, \min\{\beta, \mu + x\})$  on the left hand side concludes the proof of the  
1352 lower bound on the upper tail.

1353 **III. Upper/Lower Bound on Lower Tail.** The third and forth inequalities are a direct consequence  
1354 of the first and second inequalities applied to the two independent real random variables  $-\theta$  and  $-\lambda$   
1355 since (1)  $-\theta$  has bounded support included in  $[-\beta, -\alpha]$  and mean  $-\mu \in (-\beta, -\alpha)$ , and (2)  $-\lambda$  has  
1356 zero mean. Namely,

$$\begin{aligned} \mathbb{P}(\theta + \lambda \leq \mu - x) &= \mathbb{P}(-\theta - \lambda \geq -\mu + x) , \\ \mathbb{P}(\lambda \leq -x)\mathbb{P}(\theta \in [\mu, \beta]) &= \mathbb{P}(-\lambda \geq x)\mathbb{P}(-\theta \in [-\beta, -\mu]) , \\ \mathbb{P}(\lambda \geq 0)\mathbb{P}(\theta \in [\alpha, \max\{\alpha, \mu - x\}]) &= \mathbb{P}(-\lambda \leq 0)\mathbb{P}(-\theta \in [\min\{-\alpha, x - \mu\}, -\alpha]) , \\ \mathbb{P}(\theta \in [\alpha, \mu])\mathbb{P}(\lambda \leq 0) &= \mathbb{P}(-\theta \in [-\mu, -\alpha])\mathbb{P}(-\lambda \geq 0) , \\ \sup_{z \in (\max\{\alpha, \mu - x\}, \mu)} \{ \mathbb{P}(\theta \in [\alpha, z])\mathbb{P}(\lambda \leq \mu - x - z) \} \\ &= \sup_{\tilde{z} \in (-\mu, \min\{-\alpha, -\mu + x\})} \{ \mathbb{P}(-\theta \in [\tilde{z}, -\alpha])\mathbb{P}(-\lambda \geq -\mu + x - \tilde{z}) \} , \end{aligned}$$

1357 where we used the change of variable  $\tilde{z} = -z$ . □

1358 **Properties on the Rate Function** Lemma 10 gathers properties on the rate function  $f$  in Lemmas 11  
1359 and 11.

1360 **Lemma 10.** *Let us define*

$$\forall x \geq 0, \quad f(x) := (x + 3 - \log 2) \exp(-x) . \quad (24)$$

1361 *On  $\mathbb{R}_+$ , the function  $f$  is twice continuously differentiable, positive, decreasing and strictly convex. It*  
1362 *satisfies  $f(0) > 1$ ,  $\lim_{x \rightarrow +\infty} f(x) = 0$  and*

$$f(x) \leq \delta \quad \Longleftrightarrow \quad x \geq \overline{W}_{-1}(\log(1/\delta) + 3 - \log 2) - 3 + \log 2 ,$$

1363 *where  $\overline{W}_{-1}$  is defined in Lemma 51.*

1364 *Proof.* Direct manipulation yields  $f(0) = 3 - \log 2 > 1$ ,  $\lim_{x \rightarrow +\infty} f(x) = 0$ ,

$$\forall x \geq 0, \quad f'(x) = -(x + 2 - \log 2) \exp(-x) < 0 \quad \text{and} \quad f''(x) = (x + 1 - \log 2) \exp(-x) > 0 .$$

1365 Using that  $f(x) = e^{3 - \log 2} \exp(-h(x + 3 - \log 2))$  where  $h(x) = x - \log(x)$ , Lemma 51 yields

$$\begin{aligned} f(x) \leq \delta &\Longleftrightarrow h(x + 3 - \log 2) \geq \log(e^{3 - \log 2} / \delta) \\ &\Longleftrightarrow \overline{W}_{-1}(\log(1/\delta) + 3 - \log 2) - 3 + \log 2 \leq x . \end{aligned}$$

1366 □

1367 **Fixed Time Upper and Lower Tails Concentration** Lemma 11 gives an upper and lower tails  
1368 bound for a sum between independent Bernoulli and Laplace i.i.d. observations for a fixed time.

1369 **Lemma 11.** *Let  $\mu \in (0, 1)$  and  $\epsilon > 0$ . Let  $Z_t = \sum_{s \in [t]} X_s$  where  $X_s \sim \text{Ber}(\mu)$  are i.i.d.*  
1370 *observations. Let  $S_t = \sum_{s \in [n_t]} Y_s$  where  $Y_s \sim \text{Lap}(1/\epsilon)$  are i.i.d. observations where  $(n_t)_{t \in \mathbb{N}}$  be a*  
1371 *piece-wise constant increasing function from  $\mathbb{N}$  to  $\mathbb{N}$ . Let  $f$  as in Eq. (24). Then, for all  $t \in \mathbb{N}$  and all*  
1372  *$x > 0$ ,*

$$\begin{aligned} \mathbb{P}(Z_t + S_t \geq t(x + \mu)) &\leq f\left(t \inf_{z \in (\mu, \min\{1, x + \mu\})} \left\{ -\frac{1}{t} \log(\mathbb{P}(Z_t \geq tz)\mathbb{P}(S_t \geq t(x + \mu - z))) \right\}\right) \\ \mathbb{P}(Z_t + S_t \leq t(\mu - x)) &\leq f\left(t \inf_{z \in (\max\{0, \mu - x\}, \mu)} \left\{ -\frac{1}{t} \log(\mathbb{P}(Z_t \leq tz)\mathbb{P}(S_t \leq t(\mu - x - z))) \right\}\right) \end{aligned}$$

1373 *Proof.* Let  $t \in \mathbb{N}$  and  $x > 0$ . Then,  $Z_t$  and  $S_t$  are two independent real random variables such that  
 1374 (1)  $Z_t$  has bounded support included in  $[0, t]$  and mean  $t\mu \in (0, t)$  and (ii)  $S_t$  has zero mean. By  
 1375 symmetry of  $\text{Lap}(1/\epsilon)$  around 0, the cumulative sum of  $n_t$  observations (i.e.,  $S_t$ ) is also symmetric  
 1376 around 0. However,  $Z_t$  follows  $\text{Bin}(t, \mu)$  which can be skewed. Therefore, we have

$$\mathbb{P}(S_t \geq 0) = 1/2 = \mathbb{P}(S_t \leq 0) \quad \text{and} \quad \max\{\mathbb{P}(Z_t \in [t\mu, t]), \mathbb{P}(Z_t \in [0, t\mu])\} \leq 1, \\ \forall z \in [0, 1], \quad \mathbb{P}(Z_t \in [tz, t]) = \mathbb{P}(Z_t \geq tz) \quad \text{and} \quad \mathbb{P}(Z_t \in [0, tz]) = \mathbb{P}(Z_t \leq tz).$$

1377 Using that  $z \mapsto \mathbb{P}(Z_t \geq tz)$  is decreasing on  $(\mu, \min\{1, x + \mu\})$  and  $z \mapsto \mathbb{P}(S_t \geq t(\mu - x - z))$  is  
 1378 increasing on  $(\mu, \min\{1, x + \mu\})$ , we obtain

$$\max\{\mathbb{P}(Z_t \geq t \min\{1, x + \mu\})\mathbb{P}(S_t \leq 0), \mathbb{P}(S_t \geq tx)\mathbb{P}(Z_t \leq t\mu)\} \\ \leq \sup_{z \in (\mu, \min\{1, x + \mu\})} \{\mathbb{P}(Z_t \geq tz)\mathbb{P}(S_t \geq t(x + \mu - z))\}.$$

1379 Let us define  $g(x) := x(3 - \log(2) - \log(x))$ . Using Lemma 9 for  $(Z_t, S_t)$  and considering  $tx > 0$   
 1380 and  $z \in (\mu, \min\{\beta, \mu + x\})$  (i.e.,  $tz \in (t\mu, t \min\{\beta, \mu + x\})$ ), we obtain

$$\mathbb{P}(Z_t + S_t \geq t(x + \mu)) \leq g \left( \sup_{z \in (\mu, \min\{1, x + \mu\})} \{\mathbb{P}(Z_t \geq tz)\mathbb{P}(S_t \geq t(x + \mu - z))\} \right).$$

1381 Let  $f$  as in Eq. (24) of Lemma 10. Then, we have  $f(x) = g(\exp(-x))$ . This concludes the proof of  
 1382 the upper bound on the upper tail. The second result is obtained similarly based on Lemma 9 and the  
 1383 above results.  $\square$

#### 1384 **F.4 Tails Concentration of Cumulative Laplace Distributions**

1385 We derive time-uniform (Lemma 14) and fixed-time (Lemma 15) tails concentration for the cumulative  
 1386 sum of i.i.d. Laplace observations. Our proof technique is based on the Chernoff method and Ville's  
 1387 inequality as in Eq. (26). Therefore, we need to derive the convex conjugate of the moment  
 1388 generating function of a Laplace distribution (Lemma 12). While the time-uniform result requires  
 1389 using the peeling method, the proof of the fixed-time concentration is simpler. To use the peeling  
 1390 method, we need to control the deviation of the process on slices of time (Lemma 13).

1391 **Convex Conjugate of the Moment Generating Function of Laplace Distribution** Let  $\epsilon > 0$ .  
 1392 The moment generating function of the Laplace distribution  $\text{Lap}(1/\epsilon)$  is defined as

$$\forall \lambda \in (0, \epsilon), \quad \psi_{\text{Lap}, \epsilon}(\lambda) = \log \mathbb{E}_{X \sim \text{Lap}(1/\epsilon)} [\exp(\lambda X)] = -\log(1 - \lambda^2/\epsilon^2). \quad (25)$$

1393 Lemma 12 explicits the convex conjugate of  $\psi_{\text{Lap}, \epsilon}$  and its associated maximizer.

1394 **Lemma 12.** Let  $\psi_{\text{Lap}, \epsilon}$  as in Eq. (25). Let us define

$$\forall x > 0, \quad \psi_{\text{Lap}, \epsilon}^*(x) := \max_{\lambda \in (0, \epsilon)} \{\lambda x - \psi_{\text{Lap}, \epsilon}(\lambda)\} \quad \text{and} \quad \lambda(x) := \arg \max_{\lambda \in (0, \epsilon)} \{\lambda x - \psi_{\text{Lap}, \epsilon}(\lambda)\}.$$

1395 Then, for all  $x > 0$ , we have

$$\lambda(x) = \frac{1}{x} \left( \sqrt{1 + (x\epsilon)^2} - 1 \right) \in (0, \epsilon) \quad \text{and} \quad \psi_{\text{Lap}, \epsilon}^*(x) = h(\epsilon x) > 0.$$

1396 where  $h$  is defined in Eq. (31).

1397 *Proof.* Let  $f(\lambda) = \lambda x - \psi_{\text{Lap}, \epsilon}(\lambda)$  for all  $\lambda \in (0, \epsilon)$ . Direct manipulation yields that

$$\forall \lambda \in (0, \epsilon), \quad f'(\lambda) = x - \frac{2\lambda}{\epsilon^2 - \lambda^2} \quad \text{and} \quad f''(\lambda) = -2 \frac{\epsilon^2 + \lambda^2}{(\epsilon^2 - \lambda^2)^2} < 0.$$

1398 Moreover, for all  $\lambda \in (0, \epsilon)$ , we have

$$f'(\lambda) = 0 \quad \Longleftrightarrow \quad \lambda^2 + 2\lambda/x - \epsilon^2 = 0 \quad \Longleftrightarrow \quad \lambda = \frac{1}{x} \left( \sqrt{1 + (x\epsilon)^2} - 1 \right).$$

1399 We used that the second solution to the second order polynomial equation is negative, hence not in  
 1400  $(0, \epsilon)$ . Moreover, it is direct to see that  $\frac{1}{x} \left( \sqrt{1 + (x\epsilon)^2} - 1 \right) \in (0, \epsilon)$  since  $\sqrt{1 + x^2} - 1 \leq x$ , as it

1401 is equivalent to  $1 + x^2 \leq (x + 1)^2$  which is true when  $x > 0$ . Since  $f$  is strictly concave, the above  
 1402 computation gives its unique maximizer on  $(0, \epsilon)$ , namely we have  $\lambda(x) = \frac{1}{x} \left( \sqrt{1 + (x\epsilon)^2} - 1 \right)$ .  
 1403 Moreover, the convex conjugate of  $\psi_{\text{Lap}, \epsilon}$  is

$$\begin{aligned} \psi_{\text{Lap}, \epsilon}^*(x) &= f(\lambda(x)) = \sqrt{1 + (x\epsilon)^2} - 1 + \log \left( 1 - \frac{1}{(x\epsilon)^2} \left( \sqrt{1 + (x\epsilon)^2} - 1 \right)^2 \right) \\ &= \sqrt{1 + (x\epsilon)^2} - 1 + \log \left( \frac{2}{(x\epsilon)^2} \left( \sqrt{1 + (x\epsilon)^2} - 1 \right) \right). \end{aligned}$$

1404 This concludes the proof.  $\square$

1405 **Test Martingale for Cumulative Laplace Observations** Let  $\epsilon > 0$  and  $S_t = \sum_{s \in [t]} Y_s$  where  
 1406  $Y_s \sim \text{Lap}(1/\epsilon)$  are i.i.d. observations. Let us define

$$\forall \lambda \in (0, \epsilon), \quad M_t(\lambda) := \exp(\lambda S_t - t\psi_{\text{Lap}, \epsilon}(\lambda)).$$

1407 It is direct to see that  $M_0(\lambda) = 0$  almost surely and

$$\mathbb{E}[M_t(\lambda) \mid \mathcal{F}_{t-1}] = M_{t-1}(\lambda) \mathbb{E}_{X \sim \text{Lap}(1/\epsilon)}[\exp(\lambda X - \psi_{\text{Lap}, \epsilon}(\lambda))] = M_{t-1}(\lambda).$$

1408 Therefore,  $M_t(\lambda)$  is a test martingale. Using Ville's inequality [84] yields that

$$\forall \delta \in (0, 1), \forall \lambda \in (0, \epsilon), \quad \mathbb{P}(\exists t \in \mathbb{N}, \lambda S_t - t\psi_{\text{Lap}, \epsilon}(\lambda) \geq \log(1/\delta)) \leq \delta. \quad (26)$$

1409 **Time Uniform Tails Concentration** Lemma 13 controls the deviation of the process on slices of  
 1410 time.

1411 **Lemma 13.** *Let  $\epsilon > 0$  and  $S_t = \sum_{s \in [t]} Y_s$  where  $Y_s \sim \text{Lap}(1/\epsilon)$  are i.i.d. observations. Let  $N > 0$ .  
 1412 For all  $x > 0$ , there exists  $\lambda(x)$  such that for all  $t \geq N$ ,*

$$\{S_t \geq tx\} \subseteq \{\lambda(x)S_t - t\psi_{\text{Lap}, \epsilon}(\lambda(x)) \geq Nh(\epsilon x)\},$$

1413 where  $\lambda(x)$  as in Lemma 12 and  $h$  as in Eq. (31).

1414 *Proof.* Using Lemma 12, we obtain  $\lambda(x) \in (0, \epsilon)$  and  $\psi_{\text{Lap}, \epsilon}^*(x) = h(\epsilon x) > 0$ , hence  $t\psi_{\text{Lap}, \epsilon}^*(x) \geq$   
 1415  $N\psi_{\text{Lap}, \epsilon}^*(x)$  for  $t \geq N$ . Then, direct computations yield

$$\begin{aligned} S_t \geq tx &\implies \lambda(x)S_t - t\psi_{\text{Lap}, \epsilon}(\lambda(x)) \geq t(x\lambda(x) - \psi_{\text{Lap}, \epsilon}(\lambda(x))) = t\psi_{\text{Lap}, \epsilon}^*(x) \\ &\implies \lambda S_t - t\psi_{\text{Lap}, \epsilon}(\lambda) \geq N\psi_{\text{Lap}, \epsilon}^*(x) = Nh(\epsilon x). \end{aligned}$$

1416 This concludes the proof.  $\square$

1417 Lemma 14 gives time-uniform tails concentration for the cumulative sum of i.i.d. Laplace observations.  
 1418 It is obtained by applying Lemma 13 on slices of time with geometric growth rate.

1419 **Lemma 14.** *Let  $\delta \in (0, 1)$ . Let  $\gamma > 0$ ,  $s > 1$  and  $\zeta$  be the Riemann  $\zeta$  function. Let  $h^{-1}$  be the  
 1420 inverse of  $h$  defined as in Eq. (31), which is well-defined by Lemma 27. Let  $\epsilon > 0$  and  $S_t = \sum_{s \in [t]} Y_s$   
 1421 where  $Y_s \sim \text{Lap}(1/\epsilon)$  are i.i.d. observations. Then,*

$$\begin{aligned} \mathbb{P} \left( \exists t \in \mathbb{N}, S_t \geq \frac{t}{\epsilon} h^{-1} \left( \frac{1+\gamma}{t} \left( \log \left( \frac{\zeta(s)}{\delta} \right) + s \log(1 + \log_{1+\gamma} t) \right) \right) \right) &\leq \delta, \\ \mathbb{P} \left( \exists t \in \mathbb{N}, S_t \leq -\frac{t}{\epsilon} h^{-1} \left( \frac{1+\gamma}{t} \left( \log \left( \frac{\zeta(s)}{\delta} \right) + s \log(1 + \log_{1+\gamma} t) \right) \right) \right) &\leq \delta. \end{aligned}$$

1422 *Proof.* Let us define the geometric grid  $N_i = (1 + \gamma)^{i-1}$ , hence we have  $\mathbb{N} = \bigcup_{i \in \mathbb{N}} [N_i, N_{i+1})$ . For  
 1423 all  $i \in \mathbb{N}$ , let  $x_i(\delta) > 0$  to be defined later, and  $\lambda(x_i(\delta))$  as in Lemma 13. For all  $t \in \mathbb{N}$ , let  $g(t, \delta)$  to  
 1424 be defined later such that  $g(t, \delta) \geq x_i(\delta)$  for  $t \in [N_i, N_{i+1})$ . Using Lemma 13 with  $x_i(\delta) > 0$  and  
 1425  $g(t, \delta) \geq x_i(\delta)$  for  $t \in [N_i, N_{i+1})$ , a union bound yields that

$$\mathbb{P}(\exists t \in \mathbb{N}, S_t \geq tg(t, \delta))$$



$$\begin{aligned}
&\leq \sum_{i \in \mathbb{N}} \mathbb{P}(\exists t \in [N_i, N_{i+1}) : S_t \geq tx_i(\delta)) \\
&\leq \sum_{i \in \mathbb{N}} \mathbb{P}(\exists t \in [N_i, N_{i+1}) : \lambda(x_i(\delta))S_t - t\psi_{\text{Lap}, \epsilon}(\lambda(x_i(\delta))) \geq N_i h(\epsilon x_i(\delta))) \\
&\leq \sum_{i \in \mathbb{N}} e^{-N_i h(\epsilon x_i(\delta))},
\end{aligned}$$

1426 where the last inequality uses Ville's inequality as in Eq. (26) for all  $i \in \mathbb{N}$ . Let us define

$$\begin{aligned}
g(t, \delta) &= \frac{1}{\epsilon} h^{-1} \left( \frac{1+\gamma}{t} \left( \log \left( \frac{\zeta(s)}{\delta} \right) + s \log(1 + \log_{1+\gamma}(t)) \right) \right), \\
x_i(\delta) &= \frac{1}{\epsilon} h^{-1} \left( \frac{1}{N_i} \log \left( \frac{i^s \zeta(s)}{\delta} \right) \right).
\end{aligned}$$

1427 Using Lemma 27, we obtain that  $x_i(\delta) > 0$  and that  $h^{-1}$  is increasing on  $\mathbb{R}_+^*$ . Using  $t \in [N_i, N_{i+1})$   
1428 and  $i = 1 + \log_{1+\gamma} N_i$ , we obtain

$$\begin{aligned}
g(t, \delta) &\geq \frac{1}{\epsilon} h^{-1} \left( \frac{1}{N_i} \left( \log \left( \frac{\zeta(s)}{\delta} \right) + s \log(1 + \log_{1+\gamma}(t)) \right) \right) \\
&\geq \frac{1}{\epsilon} h^{-1} \left( \frac{1}{N_i} \log \left( \frac{i^s \zeta(s)}{\delta} \right) \right) = x_i(\delta).
\end{aligned}$$

1429 Therefore, we have

$$\mathbb{P}(\exists t \in \mathbb{N}, S_t \geq tg(t, \delta)) \leq \sum_{i \in \mathbb{N}} e^{-N_i h(\epsilon x_i(\delta))} \leq \frac{\delta}{\zeta(s)} \sum_{i \in \mathbb{N}} \frac{1}{i^s} = \delta.$$

1430 This concludes the proof of the first result.

1431 By symmetry of the  $\text{Lap}(1/\epsilon)$  around zero, the cumulative sum of i.i.d. observations is symmetric  
1432 around zero. Combining the first result with the symmetry around zero yields the second result.  $\square$

1433 **Fixed Time Tails Concentration** When the time is fixed and not random, there is no need to  
1434 consider slices of time and we can directly control the deviation of the process.

1435 **Lemma 15.** *Let  $\epsilon > 0$  and  $S_t = \sum_{s \in [t]} Y_s$  where  $Y_s \sim \text{Lap}(1/\epsilon)$  are i.i.d. observations. Let  $h$  as in*  
1436 *Eq. (31). Then,*

$$\begin{aligned}
\forall t \in \mathbb{N}, \forall x > 0, \quad \mathbb{P}(S_t \geq tx) &\leq \exp(-th(\epsilon x)), \\
\forall t \in \mathbb{N}, \forall x > 0, \quad \mathbb{P}(S_t \leq -tx) &\leq \exp(-th(\epsilon x)).
\end{aligned}$$

1437 *Proof.* The first result can be obtained with the same manipulation as in the proof of Lemma 14, i.e.,  
1438 combining Ville's inequality in Eq. (26) with Lemma 13 at  $N = t$ .

1439 By symmetry of the  $\text{Lap}(1/\epsilon)$  around zero, the cumulative sum of i.i.d. observations is symmetric  
1440 around zero. Combining the first result with the symmetry around zero yields the second result.  $\square$

## 1441 F.5 Fixed Time Tails Concentration of Cumulative Bernoulli Distributions

1442 The fixed time upper and lower tail concentration of cumulative Bernoulli distributions are well-  
1443 studied. Using the Chernoff method yields Lemma 16, whose proof is omitted since it is a classic  
1444 result.

1445 **Lemma 16** (Chernoff Tail Bound for Bernoulli Distributions [17]). *Let  $\mu \in (0, 1)$  and  $Z_t =$*   
1446  *$\sum_{s \in [t]} X_s$  where  $X_s \sim \text{Ber}(\mu)$  are i.i.d. observations. Then,*

$$\begin{aligned}
\forall t \in \mathbb{N}, \forall x \in (\mu, 1), \quad \mathbb{P}(Z_t \geq tx) &\leq \exp(-t\text{kl}(x, \mu)), \\
\forall t \in \mathbb{N}, \forall x \in (0, \mu), \quad \mathbb{P}(Z_t \leq tx) &\leq \exp(-t\text{kl}(x, \mu)).
\end{aligned}$$

1447 **F.6 Fixed Time Tails Concentration for a Convolution between Bernoulli and Laplace**  
1448 **Distributions**

1449 We provide upper (Lemma 17) and lower (Lemma 18) tail concentrations for a sum (i.e., convolution)  
1450 between independent Bernoulli and Laplace i.i.d. observations for a fixed time.

1451 **Fixed Time Upper Tail Concentration** Lemma 17 shows an upper tail concentration on the sum  
1452 (i.e., convolution) between independent Bernoulli and Laplace i.i.d. observations.

1453 **Lemma 17.** *Let  $\mu \in (0, 1)$  and  $Z_t = \sum_{s \in [t]} X_s$  where  $X_s \sim \text{Ber}(\mu)$  are i.i.d. observations. Let*  
1454  *$(n_t)_{t \in \mathbb{N}}$  be a piece-wise constant increasing function from  $\mathbb{N}$  to  $\mathbb{N}$ . Let  $\epsilon > 0$  and  $S_t = \sum_{s \in [n_t]} Y_s$*   
1455 *where  $Y_s \sim \text{Lap}(1/\epsilon)$  are i.i.d. observations. Then,*

$$\forall t \in \mathbb{N}, \forall x > 0, \quad \mathbb{P}(Z_t + S_t \geq t(\mu + x)) \leq f\left(t\tilde{d}_\epsilon^-(\mu + x, \mu, t/n_t)\right),$$

1456 where  $f$  is defined in Eq. (24) and  $\tilde{d}_\epsilon^-$  is defined in Eq. (32).

1457 *Proof.* Let  $t \in \mathbb{N}$  and  $x > 0$ . Combining Lemmas 15 and 16, we obtain, for all  $x > 0$  and all  
1458  $z \in (\mu, \min\{1, x + \mu\})$ ,

$$-\frac{1}{t} \log(\bar{G}_t(tz)\bar{F}_{n_t}(t(x + \mu - z))) \geq \text{kl}(z, \mu) + \frac{n_t}{t} h\left(\frac{t}{n_t} \epsilon(x + \mu - z)\right),$$

1459 where we used that  $x + \mu - z > 0$ . Taking the infimum on  $(\mu, \min\{1, x + \mu\})$  on both sides and  
1460 using that  $[x + \mu]_0^1 = \min\{1, x + \mu\} > \mu$ , we obtain

$$\inf_{z \in (\mu, \min\{1, x + \mu\})} \left\{ -\frac{1}{t} \log(\bar{G}_t(tz)\bar{F}_{n_t}(t(x + \mu - z))) \right\} \geq \tilde{d}_\epsilon^-(\mu + x, \mu, t/n_t),$$

1461 where  $\tilde{d}_\epsilon^-$  is defined in Eq. (32). Since  $f$  is decreasing on  $\mathbb{R}_+$  (Lemma 10), using Lemma 11 yields

$$\mathbb{P}(Z_t + S_t \geq t(\mu + x)) \leq f\left(t\tilde{d}_\epsilon^-(\mu + x, \mu, t/n_t)\right).$$

1462 which concludes the proof. □

1463 **Fixed Time Lower Tail Concentration** Lemma 18 shows a lower tail concentration on the sum  
1464 (i.e., convolution) between independent Bernoulli and Laplace i.i.d. observations.

1465 **Lemma 18.** *Let  $\mu \in (0, 1)$  and  $Z_t = \sum_{s \in [t]} X_s$  where  $X_s \sim \text{Ber}(\mu)$  are i.i.d. observations. Let*  
1466  *$(n_t)_{t \in \mathbb{N}}$  be a piece-wise constant increasing function from  $\mathbb{N}$  to  $\mathbb{N}$ . Let  $\epsilon > 0$  and  $S_t = \sum_{s \in [n_t]} Y_s$*   
1467 *where  $Y_s \sim \text{Lap}(1/\epsilon)$  are i.i.d. observations. Then,*

$$\forall t \in \mathbb{N}, \forall x > 0, \quad \mathbb{P}(Z_t + S_t \leq t(\mu - x)) \leq f\left(t\tilde{d}_\epsilon^+(\mu - x, \mu, t/n_t)\right),$$

1468 where  $f$  is defined in Eq. (24) and  $\tilde{d}_\epsilon^+$  is defined in Eq. (32).

1469 *Proof.* Let  $t \in \mathbb{N}$  and  $x > 0$ . Combining Lemmas 15 and 16, we obtain, for all  $x > 0$  and all  
1470  $z \in (\max\{0, \mu - x\}, \mu)$ ,

$$-\frac{1}{t} \log(G_t(tz)F_{n_t}(t(\mu - x - z))) \geq \text{kl}(z, \mu) + \frac{n_t}{t} h\left(\frac{t}{n_t} \epsilon(z + x - \mu)\right),$$

1471 where we used that  $\mu - x - z < 0$ . Taking the infimum on  $z \in (\max\{0, \mu - x\}, \mu)$  on both sides  
1472 and using that  $[\mu - x]_0^1 = \max\{0, \mu - x\} < \mu$ , we obtain

$$\inf_{z \in (\max\{0, \mu - x\}, \mu)} \left\{ -\frac{1}{t} \log(G_t(tz)F_{n_t}(t(\mu - x - z))) \right\} \geq \tilde{d}_\epsilon^+(\mu - x, \mu, t/n_t),$$

1473 where  $\tilde{d}_\epsilon^+$  is defined in Eq. (32). Since  $f$  is decreasing on  $\mathbb{R}_+$  (Lemma 10), using Lemma 11 yields

$$\mathbb{P}(Z_t + S_t \leq t(\mu - x)) \leq f\left(t\tilde{d}_\epsilon^+(\mu - x, \mu, t/n_t)\right),$$

1474 which concludes the proof. □

1475 **E.7 Geometric Grid Time Uniform Tails Concentration for a Convolution between Bernoulli**  
1476 **and Laplace Distributions**

1477 We provide upper (Lemma 19) and lower (Lemma 20) tail concentrations for a sum (i.e., convolution)  
1478 between independent Bernoulli and Laplace i.i.d. observations that holds time uniformly on a  
1479 geometric grid.

1480 **Geometric Grid Time Uniform Upper Tail Concentration** Lemma 19 gives a threshold ensuring  
1481 that a geometric grid time uniform upper tail concentration holds with probability at least  $1 - \delta$ .

1482 **Lemma 19.** *Let  $\delta \in (0, 1)$ . Let  $(\tilde{\mu}_n, \tilde{N}_n, k_n)$  are given by  $GPE_\eta(\epsilon)$ . Let  $s > 1$  and  $\zeta$  be the Riemann  
1483  $\zeta$  function. Let  $\overline{W}_{-1}(x) = -W_{-1}(-e^{-x})$  for all  $x \geq 1$ , where  $W_{-1}$  is the negative branch of the  
1484 Lambert  $W$  function. It satisfies  $\overline{W}_{-1}(x) \approx x + \log x$ , see Lemma 51. Let us define*

$$c(k, \delta) = \overline{W}_{-1}(\log(1/\delta) + s \log(k) + \log(\zeta(s)) + 3 - 2 \log 2) - 3 + 2 \log 2. \quad (27)$$

1485 For all  $a \in [K]$ , let us define

$$\mathcal{E}_{\delta, a, -} = \left\{ \forall n \in \mathbb{N}, \tilde{N}_{n, a} \tilde{d}_\epsilon^-(\tilde{\mu}_{n, a}, \mu_a, \tilde{N}_{n, a}/k_{n, a}) \leq c(k_{n, a}, \delta) \right\}, \quad (28)$$

1486 where  $\tilde{d}_\epsilon^-$  is defined in Eq. (32). Then, we have  $\mathbb{P}_{\nu\pi}(\mathcal{E}_{\delta, a, -}^c) \leq \delta$  for all  $a \in [K]$ .

1487 *Proof.* Let us define the geometric grid  $N_i = (1 + \eta)^{i-1}$ , hence we have  $\mathbb{N} = \bigcup_{i \in \mathbb{N}} [N_i, N_{i+1})$ . Let  
1488  $a \in [K]$ . If  $\tilde{N}_{n, a} \in [N_i, N_{i+1})$ , then we have  $\tilde{N}_{n, a} = \lceil N_i \rceil$  and  $k_{n, a} = i$ . By union bound, we  
1489 obtain

$$\begin{aligned} \mathbb{P}_{\nu\pi}(\mathcal{E}_{\delta, a, -}^c) &= \mathbb{P}_{\nu\pi} \left( \exists n \in \mathbb{N}, \tilde{N}_{n, a} \tilde{d}_\epsilon^-(\tilde{\mu}_{n, a}, \mu_a, \tilde{N}_{n, a}/k_{n, a}) \geq c(k_{n, a}, \delta) \right) \\ &\leq \sum_{i \in \mathbb{N}} \mathbb{P}_{\nu\pi} \left( \exists i \in \mathbb{N}, (\tilde{N}_{n, a}, k_{n, a}) = (\lceil N_i \rceil, i) \wedge \tilde{N}_{n, a} \tilde{d}_\epsilon^-(\tilde{\mu}_{n, a}, \mu_a, \tilde{N}_{n, a}/k_{n, a}) \geq c(k_{n, a}, \delta) \right) \\ &= \sum_{i \in \mathbb{N}} \mathbb{P} \left( \lceil N_i \rceil \tilde{d}_\epsilon^-((Z_{\lceil N_i \rceil} + S_i)/\lceil N_i \rceil, \mu_a, \lceil N_i \rceil/i) \geq c(i, \delta) \right), \end{aligned}$$

1490 where  $Z_{\lceil N_i \rceil}$  is the cumulative sum of  $\lceil N_i \rceil$  i.i.d. observations from  $\text{Ber}(\mu_a)$  and  $S_i$  is the cumulative  
1491 sum of  $i$  i.i.d. observations from  $\text{Lap}(1/\epsilon)$ .

1492 For all  $i \in \mathbb{N}$ , let  $x_i > 0$  be the unique solution of  $\lceil N_i \rceil \tilde{d}_\epsilon^-(x + \mu_a, \mu_a, \lceil N_i \rceil/i) = c(i, \delta)$ , which  
1493 exists by Lemma 32. Then, we obtain

$$\begin{aligned} &\mathbb{P} \left( \lceil N_i \rceil \tilde{d}_\epsilon^-((Z_{\lceil N_i \rceil} + S_i)/\lceil N_i \rceil, \mu_a, \lceil N_i \rceil/i) \geq c(i, \delta) \right) \\ &= \mathbb{P} \left( \tilde{d}_\epsilon^-((Z_{\lceil N_i \rceil} + S_i)/\lceil N_i \rceil, \mu_a, \lceil N_i \rceil/i) \geq \tilde{d}_\epsilon^-(x_i + \mu_a, \mu_a, \lceil N_i \rceil/i) \right) \\ &\leq \mathbb{P}(Z_{\lceil N_i \rceil} + S_i \geq \lceil N_i \rceil(x_i + \mu_a)) \leq f \left( \lceil N_i \rceil \tilde{d}_\epsilon^-(x_i + \mu_a, \mu_a, \lceil N_i \rceil/i) \right) = f(c(i, \delta)), \end{aligned}$$

1494 where  $f(x) := (x + 3 - \log 2) \exp(-x)$  for all  $x \geq 0$ . The first and the last equalities are obtained by  
1495 definition of  $x_i$ , i.e.,  $\lceil N_i \rceil \tilde{d}_\epsilon^-(x + \mu_a, \mu_a, \lceil N_i \rceil/i) = c(i, \delta)$ . The first inequality is obtained by using  
1496 Lemma 33, and the second inequality is obtained by using Lemma 17. Using Lemma 10 yields

$$f(x) \leq \delta \iff \overline{W}_{-1}(\log(1/\delta) + 3 - \log 2) - 3 + \log 2 \leq x.$$

1497 Taking

$$c(i, \delta) = \overline{W}_{-1}(\log(i^s \zeta(s)/\delta) + 3 - \log 2) - 3 + \log 2,$$

1498 we can conclude the proof since  $\mathbb{P}_{\nu\pi}(\mathcal{E}_{\delta, a, -}^c) \leq \sum_{i \in \mathbb{N}} f(c(i, \delta)) \leq \sum_{i \in \mathbb{N}} \frac{\delta}{\zeta(s)i^s} \leq \delta$ .  $\square$

1499 **Geometric Grid Time Uniform Lower Tail Concentration** Lemma 20 gives a threshold ensuring  
1500 that a geometric grid time uniform lower tail concentration holds with probability at least  $1 - \delta$ .

1501 **Lemma 20.** Let  $\delta \in (0, 1)$ . Let  $(\tilde{\mu}_n, \tilde{N}_n, k_n)$  are given by  $GPE_\eta(\epsilon)$ . Let  $c$  as in Eq. (27). For all  
 1502  $a \in [K]$ , let us define

$$\mathcal{E}_{\delta,a,+} = \left\{ \forall n \in \mathbb{N}, \tilde{N}_{n,a} \tilde{d}_\epsilon^+(\tilde{\mu}_{n,a}, \mu_a, \tilde{N}_{n,a}/k_{n,a}) \leq c(k_{n,a}, \delta) \right\}, \quad (29)$$

1503 where  $\tilde{d}_\epsilon^+$  is defined in Eq. (32). Then, we have  $\mathbb{P}_{\nu\pi}(\mathcal{E}_{\delta,a,+}^c) \leq \delta$  for all  $a \in [K]$ .

1504 *Proof.* Let us define the geometric grid  $N_i = (1 + \eta)^{i-1}$ , hence we have  $\mathbb{N} = \bigcup_{i \in \mathbb{N}} [N_i, N_{i+1})$ . Let  
 1505  $a \in [K]$ . If  $\tilde{N}_{n,a} \in [N_i, N_{i+1})$ , then we have  $\tilde{N}_{n,a} = \lceil N_i \rceil$  and  $k_{n,a} = i$ . By union bound, we  
 1506 obtain

$$\begin{aligned} \mathbb{P}_{\nu\pi}(\mathcal{E}_{\delta,a,+}^c) &= \mathbb{P}_{\nu\pi} \left( \exists n \in \mathbb{N}, \tilde{N}_{n,a} \tilde{d}_\epsilon^+(\tilde{\mu}_{n,a}, \mu_a, \tilde{N}_{n,a}/k_{n,a}) \geq c(k_{n,a}, \delta) \right) \\ &\leq \sum_{i \in \mathbb{N}} \mathbb{P}_{\nu\pi} \left( \exists i \in \mathbb{N}, (\tilde{N}_{n,a}, k_{n,a}) = (\lceil N_i \rceil, i) \wedge \tilde{N}_{n,a} \tilde{d}_\epsilon^+(\tilde{\mu}_{n,a}, \mu_a, \tilde{N}_{n,a}/k_{n,a}) \geq c(k_{n,a}, \delta) \right) \\ &= \sum_{i \in \mathbb{N}} \mathbb{P} \left( \lceil N_i \rceil \tilde{d}_\epsilon^+((Z_{\lceil N_i \rceil} + S_i)/\lceil N_i \rceil, \mu_a, \lceil N_i \rceil/i) \geq c(i, \delta) \right), \end{aligned}$$

1507 where  $Z_{\lceil N_i \rceil}$  is the cumulative sum of  $\lceil N_i \rceil$  i.i.d. observations from  $\text{Ber}(\mu_a)$  and  $S_i$  is the cumulative  
 1508 sum of  $i$  i.i.d. observations from  $\text{Lap}(1/\epsilon)$ .

1509 For all  $i \in \mathbb{N}$ , let  $x_i > 0$  be the unique solution of  $\lceil N_i \rceil \tilde{d}_\epsilon^+(\mu_a - x_i, \mu_a, \lceil N_i \rceil/i) = c(i, \delta)$ , which  
 1510 exists by Lemma 32. Then, we obtain

$$\begin{aligned} &\mathbb{P} \left( \lceil N_i \rceil \tilde{d}_\epsilon^+((Z_{\lceil N_i \rceil} + S_i)/\lceil N_i \rceil, \mu_a, \lceil N_i \rceil/i) \geq c(i, \delta) \right) \\ &= \mathbb{P} \left( \tilde{d}_\epsilon^+((Z_{\lceil N_i \rceil} + S_i)/\lceil N_i \rceil, \mu_a, \lceil N_i \rceil/i) \geq \tilde{d}_\epsilon^+(\mu_a - x_i, \mu_a, \lceil N_i \rceil/i) \right) \\ &\leq \mathbb{P}(Z_{\lceil N_i \rceil} + S_i \leq \lceil N_i \rceil(\mu_a - x_i)) \leq f \left( \lceil N_i \rceil \tilde{d}_\epsilon^+(\mu_a - x_i, \mu_a, \lceil N_i \rceil/i) \right) = f(c(i, \delta)) \leq \frac{\delta}{\zeta(s) i^s} \end{aligned}$$

1511 where  $f(x) := (x + 3 - \log 2) \exp(-x)$  for all  $x \geq 0$ . The first and the last equalities are obtained by  
 1512 definition of  $x_i$ , i.e.,  $\lceil N_i \rceil \tilde{d}_\epsilon^+(\mu_a - x_i, \mu_a, \lceil N_i \rceil/i) = c(i, \delta)$ . The first inequality is obtained by using  
 1513 Lemma 33, and the second inequality is obtained by using Lemma 18. The last inequality uses the  
 1514 same derivations based on Lemma 10 as in the proof of Lemma 19 by taking

$$c(i, \delta) = \overline{W}_{-1}(\log(i^s \zeta(s)/\delta) + 3 - \log 2) - 3 + \log 2.$$

1515 This concludes the proof since  $\mathbb{P}_{\nu\pi}(\mathcal{E}_{\delta,a,-}^c) \leq \sum_{i \in \mathbb{N}} \frac{\delta}{\zeta(s) i^s} \leq \delta$ .  $\square$

## 1516 G Divergence, Transportation Cost and Characteristic Time

1517 Appendix G is organized as follow. First, we derive regularity properties for the signed (modified)  
 1518 divergences  $d_\epsilon^\pm$  (Appendix G.1) and  $\tilde{d}_\epsilon^\pm$  (Appendix G.1.1). Second, we derive regularity properties  
 1519 the (modified) transportation costs  $W_{\epsilon,a,b}$  (Appendix G.2) and  $\tilde{W}_{\epsilon,a,b}$  (Appendix G.2.1) for a pair of  
 1520 arms  $(a, b)$ . Third, we study the characteristic time for  $\epsilon$ -global DP BAI (Appendix G.3).

### 1521 G.1 Signed Divergence

1522 Recall  $[x]_0^1 := \max\{0, \min\{1, x\}\}$  and

$$\forall (\lambda, \mu) \in (0, 1)^2, \quad \text{kl}(\lambda, \mu) := \lambda \log \left( \frac{\lambda}{\mu} \right) + (1 - \lambda) \log \left( \frac{1 - \lambda}{1 - \mu} \right)$$

1523 where  $\text{kl}$  is infinity when  $\{\mu, \lambda\} \cap \{0, 1\} \neq \emptyset$ . The signed divergences  $d_\epsilon^\pm$  are defined in Eq. (3), i.e.,

$$\begin{aligned} \forall (\lambda, \mu) \in \mathbb{R} \times [0, 1], \quad d_\epsilon^-(\lambda, \mu) &:= \mathbb{1}(\mu < [\lambda]_0^1) \inf_{z \in [\mu, [\lambda]_0^1]} \{ \text{kl}(z, \mu) + \epsilon([\lambda]_0^1 - z) \}, \\ d_\epsilon^+(\lambda, \mu) &:= \mathbb{1}(\mu > [\lambda]_0^1) \inf_{z \in [[\lambda]_0^1, \mu]} \{ \text{kl}(z, \mu) + \epsilon(z - [\lambda]_0^1) \}. \end{aligned}$$

1524 Lemma 21 relates  $d_\epsilon$  and  $d_\epsilon^\pm$ .

1525 **Lemma 21.** Let  $d_\epsilon^\pm$  and  $d_\epsilon$  as in Eq. (3) and (2). Let  $(\kappa, \nu) \in \mathcal{F}^2$  with means  $(\lambda, \mu) \in (0, 1)^2$ . Then,

$$d_\epsilon(\kappa, \nu) = \begin{cases} 0 & \text{if } \lambda = \mu \\ d_\epsilon^-(\lambda, \mu) & \text{if } \mu < \lambda \\ d_\epsilon^+(\lambda, \mu) & \text{if } \mu > \lambda \end{cases}.$$

1526 *Proof.* When  $\lambda = \mu$ , we have  $d_\epsilon(\kappa, \nu) = 0$  by taking  $\varphi = \nu$  and using the non-negativity of  $d_\epsilon$ .

1527 Let  $\varphi \in \mathcal{F}$  with mean  $z \in (0, 1)$ . When  $\mu < \lambda$ , we have

$$\begin{aligned} d_\epsilon(\kappa, \nu) &= \min \left\{ \inf_{z \in (0, \mu)} \{ \text{kl}(z, \mu) + \epsilon(\lambda - z) \}, \inf_{z \in [\mu, \lambda]} \{ \text{kl}(z, \mu) + \epsilon(\lambda - z) \}, \right. \\ &\quad \left. \inf_{z \in (\lambda, 1)} \{ \text{kl}(z, \mu) + \epsilon(z - \lambda) \} \right\} \\ &= \inf_{z \in [\mu, \lambda]} \{ \text{kl}(z, \mu) + \epsilon(\lambda - z) \} = d_\epsilon^-(\lambda, \mu), \end{aligned}$$

1528 where we partitioned  $(0, 1)$  and used that (1)  $z \mapsto \text{kl}(z, \mu) + \epsilon(z - \lambda)$  is increasing on  $(\lambda, 1)$ , hence  
1529 the infimum on this interval is achieved at  $\lambda$ , and (2)  $z \mapsto \text{kl}(z, \mu) + \epsilon(\lambda - z)$ , is decreasing on  $(0, \mu)$ ,  
1530 hence the infimum on this interval is achieved at  $\mu$ .

1531 When  $\mu > \lambda$ , we have

$$\begin{aligned} d_\epsilon(\kappa, \nu) &= \min \left\{ \inf_{z \in (0, \lambda)} \{ \text{kl}(z, \mu) + \epsilon(\lambda - z) \}, \inf_{z \in [\lambda, \mu]} \{ \text{kl}(z, \mu) + \epsilon(z - \lambda) \}, \right. \\ &\quad \left. \inf_{z \in (\mu, 1)} \{ \text{kl}(z, \mu) + \epsilon(z - \lambda) \} \right\} \\ &= \inf_{z \in [\lambda, \mu]} \{ \text{kl}(z, \mu) + \epsilon(z - \lambda) \} = d_\epsilon^+(\lambda, \mu), \end{aligned}$$

1532 where we partitioned  $(0, 1)$  and used that (1)  $z \mapsto \text{kl}(z, \mu) + \epsilon(z - \lambda)$  is increasing on  $(\mu, 1)$ , hence  
1533 the infimum on this interval is achieved at  $\mu$ , and (2)  $z \mapsto \text{kl}(z, \mu) + \epsilon(\lambda - z)$ , is decreasing on  $(0, \lambda)$ ,  
1534 hence the infimum on this interval is achieved at  $\lambda$ .  $\square$

1535 Lemma 22 shows a strong link between  $d_\epsilon^\pm$ . This symmetry property can be used to carry regularity  
1536 properties from  $d_\epsilon^+$  to  $d_\epsilon^-$ , and vice versa.

1537 **Lemma 22.** Let  $d_\epsilon^\pm$  as in Eq. 3. For all  $\mu \in [0, 1]$  and all  $\lambda \in \mathbb{R}$ ,

$$d_\epsilon^+(1 - \lambda, 1 - \mu) = d_\epsilon^-(\lambda, \mu) \quad \text{and} \quad d_\epsilon^-(1 - \lambda, 1 - \mu) = d_\epsilon^+(\lambda, \mu).$$

1538 *Proof.* By definitions and change of variable  $\tilde{z} = 1 - z$  and  $\text{kl}(1 - \tilde{z}, 1 - \mu) = \text{kl}(\tilde{z}, \mu)$ , we obtain

$$\begin{aligned} d_\epsilon^+(1 - \lambda, 1 - \mu) &= \mathbb{1}(\mu < [\lambda]_0^1) \inf_{z \in [1 - [\lambda]_0^1, 1 - \mu]} \{ \text{kl}(z, 1 - \mu) + \epsilon(\max\{0, \min\{1, \lambda\} - (1 - z)) \} \\ &= \mathbb{1}(\mu < [\lambda]_0^1) \inf_{\tilde{z} \in [\mu, [\lambda]_0^1]} \{ \text{kl}(1 - \tilde{z}, 1 - \mu) + \epsilon(\max\{0, \min\{1, \lambda\} - \tilde{z}) \} \\ &= \mathbb{1}(\mu < [\lambda]_0^1) \inf_{\tilde{z} \in [\mu, [\lambda]_0^1]} \{ \text{kl}(\tilde{z}, \mu) + \epsilon(\max\{0, \min\{1, \lambda\} - \tilde{z}) \} = d_\epsilon^-(\lambda, \mu). \end{aligned}$$

1539 The second equality is a consequence of the first.  $\square$

1540 Lemma 23 gathers regularity properties on the functions  $g_\epsilon^\pm$  that appear in the explicit solutions of  
1541  $d_\epsilon^\pm$ , as shown below. Intuitively, those functionals govern locally the separation between the low  
1542 privacy regime where  $d_\epsilon^\pm$  is equals to the kl and the high privacy regime where the divergence has to  
1543 be modified to account for the privacy budget  $\epsilon$ .

1544 **Lemma 23.** Let  $\epsilon > 0$ . Let  $g_\epsilon^\pm$  defined as

$$\forall x \in [0, 1], \quad g_\epsilon^+(x) := \frac{x}{x(1 - e^\epsilon) + e^\epsilon} \quad \text{and} \quad g_\epsilon^-(x) := \frac{x e^\epsilon}{x(e^\epsilon - 1) + 1}. \quad (30)$$

1545 On  $[0, 1]$ , the function  $g_\epsilon^+$  is twice continuously differentiable, increasing and strictly convex. It  
1546 satisfies  $g_\epsilon^+(0) = 0$ ,  $g_\epsilon^+(1) = 1$  and  $g_\epsilon^+(x) < x$  on  $(0, 1)$ . On  $[0, 1]$ , the function  $g_\epsilon^-$  is twice  
1547 continuously differentiable, increasing and strictly concave. It satisfies  $g_\epsilon^-(0) = 0$ ,  $g_\epsilon^-(1) = 1$  and  
1548  $g_\epsilon^-(x) > x$  on  $(0, 1)$ . For all  $x \in [0, 1]$ , we have  $g_\epsilon^+(g_\epsilon^-(x)) = x$  and  $g_\epsilon^-(1 - x) + g_\epsilon^+(x) = 1$ .  
1549 For all  $x \in [0, 1]$ , we have  $\lim_{\epsilon \rightarrow 0} g_\epsilon^+(x) = \lim_{\epsilon \rightarrow 0} g_\epsilon^-(x) = x$ ; it satisfies  $\lim_{\epsilon \rightarrow +\infty} g_\epsilon^-(x) = 1$  if  
1550  $x \neq 0$  and  $\lim_{\epsilon \rightarrow +\infty} g_\epsilon^+(x) = 0$  if  $x \neq 1$ .

1551 *Proof.* Using that  $e^\epsilon > 1$ , direct computations yield that, for all  $x \in [0, 1]$ ,

$$\begin{aligned} (g_\epsilon^+)'(x) &= \frac{e^\epsilon}{(x(1-e^\epsilon) + e^\epsilon)^2} > 0 \quad \text{and} \quad (g_\epsilon^+)''(x) = -2 \frac{e^\epsilon(1-e^\epsilon)}{(x(1-e^\epsilon) + e^\epsilon)^2} > 0, \\ (g_\epsilon^-)'(x) &= \frac{e^\epsilon}{(x(e^\epsilon - 1) + 1)^2} > 0 \quad \text{and} \quad (g_\epsilon^-)''(x) = -2 \frac{e^\epsilon(e^\epsilon - 1)}{(x(e^\epsilon - 1) + 1)^3} < 0. \end{aligned}$$

1552 Therefore,  $g_\epsilon^+$  is twice continuously differentiable, increasing and strictly convex on  $[0, 1]$  and  $g_\epsilon^-$   
 1553 is twice continuously differentiable, increasing and strictly concave on  $[0, 1]$ . It is direct to see  
 1554 that  $g_\epsilon^+(0) = g_\epsilon^-(0) = 0$  and  $g_\epsilon^+(1) = g_\epsilon^-(1) = 1$ . Since they are strictly convex and strictly  
 1555 concave, we obtain  $g_\epsilon^+(x) < x$  and  $g_\epsilon^-(x) > x$  for all  $x \in (0, 1)$ . It is direct to see that, for  
 1556 all  $x \in [0, 1]$ , we have  $g_\epsilon^+(g_\epsilon^-(x)) = x$  and  $1 - g_\epsilon^+(x) = g_\epsilon^-(1 - x)$ . It is direct to see that,  
 1557  $\lim_{\epsilon \rightarrow 0} g_\epsilon^+(x) = \lim_{\epsilon \rightarrow 0} g_\epsilon^-(x) = x$  for all  $x \in [0, 1]$ , and  $\lim_{\epsilon \rightarrow +\infty} g_\epsilon^+(x) = 0$  if  $x \neq 1$  and  
 1558  $\lim_{\epsilon \rightarrow +\infty} g_\epsilon^-(x) = 1$  if  $x \neq 0$ .  $\square$

1559 Lemma 24 gathers regularity properties of  $d_\epsilon^+$ . In particular, it gives a closed-form solution, which is  
 1560 a key property used in our implementation to reduce the computational cost.

1561 **Lemma 24.** Let  $d_\epsilon^+$  as in Eq. (3), and  $g_\epsilon^\pm$  as in Eq. (30). For all  $\mu \in [0, 1]$  and  $\lambda \in \mathbb{R}$ , we have

$$d_\epsilon^+(\lambda, \mu) = \begin{cases} 0 & \text{if } \mu \in [0, [\lambda]_0^1] \\ -\log(1 - \mu(1 - e^{-\epsilon})) - \epsilon[\lambda]_0^1 & \text{if } \mu \in (g_\epsilon^-([\lambda]_0^1), 1] \\ \text{kl}(\lambda, \mu) & \text{if } \lambda \in (0, 1) \wedge \mu \in ([\lambda]_0^1, g_\epsilon^-([\lambda]_0^1)] \end{cases}.$$

1562 The function  $(\lambda, \mu) \mapsto d_\epsilon^+(\lambda, \mu)$  is jointly continuous on  $\mathbb{R} \times [0, 1]$ . For all  $\mu \in [0, 1]$ , the function  
 1563  $\lambda \mapsto d_\epsilon^+(\lambda, \mu)$  is constant on  $(-\infty, 0]$  and on  $[1, +\infty)$ . Then,

$$\forall \lambda \in (0, 1), \forall \mu \in [0, 1], \quad d_\epsilon^+(\lambda, \mu) = \begin{cases} 0 & \text{if } \mu \in [0, \lambda] \\ \text{kl}(\lambda, \mu) & \mu \in (\lambda, g_\epsilon^-(\lambda)] \\ -\log(1 - \mu(1 - e^{-\epsilon})) - \epsilon\lambda & \text{if } \mu \in (g_\epsilon^-(\lambda), 1] \end{cases}.$$

1564 For all  $\mu \in [0, 1]$ , the function  $\lambda \mapsto d_\epsilon^+(\lambda, \mu)$  is continuously differentiable, positive, decreasing  
 1565 and convex on  $(0, \mu)$ ; it is affine with negative slope  $-\epsilon$  on  $(0, g_\epsilon^+(\mu))$  and twice continuously  
 1566 differentiable and strictly convex on  $(g_\epsilon^+(\mu), \mu)$ .

1567 For all  $\lambda \in (0, 1)$ , the function  $\mu \mapsto d_\epsilon^+(\lambda, \mu)$  is positive, three times differentiable with continuous  
 1568 first derivative, increasing and strictly convex on  $(\lambda, 1]$ ; its second derivative is discontinuous at  
 1569  $g_\epsilon^-(\lambda)$  with gap  $\frac{\partial^2 d_\epsilon^+}{\partial \mu^2}(\lambda, g_\epsilon^-(\lambda)) - \lim_{\mu \rightarrow g_\epsilon^-(\lambda)^+} \frac{\partial^2 d_\epsilon^+}{\partial \mu^2}(\lambda, \mu) > 0$ . Moreover, we have

$$\forall \mu \in (\lambda, 1], \quad \frac{\partial d_\epsilon^+}{\partial \mu}(\lambda, \mu) = \begin{cases} \frac{1-e^{-\epsilon}}{1-\mu(1-e^{-\epsilon})} & \text{if } \mu \in (g_\epsilon^-(\lambda), 1] \\ \frac{\mu-\lambda}{\mu(1-\mu)} & \text{if } \mu \in (\lambda, g_\epsilon^-(\lambda)] \end{cases}.$$

1570 The function  $d_\epsilon^+$  is jointly convex on  $(0, 1) \times [0, 1]$ .

1571 *Proof.* Recall that  $d_\epsilon^+(\lambda, \mu) = \mathbf{1}(\mu > [\lambda]_0^1) \inf_{z \in [[\lambda]_0^1, \mu]} f_\epsilon^+([\lambda]_0^1, \mu, z)$  where  $f_\epsilon^+(\lambda, \mu, z) =$   
 1572  $\text{kl}(z, \mu) + \epsilon(z - \lambda)$ . Direct computations yield that, for all  $z \in ([\lambda]_0^1, \mu)$ ,

$$\begin{aligned} \frac{\partial f_\epsilon^+}{\partial z}(\lambda, \mu, z) &= \log\left(\frac{z(1-\mu)}{(1-z)\mu}\right) + \epsilon \quad \text{and} \quad \frac{\partial f_\epsilon^+}{\partial z}(\lambda, \mu, z) = 0 \iff z = g_\epsilon^+(\mu), \\ \frac{\partial^2 f_\epsilon^+}{\partial z^2}(\lambda, \mu, z) &= \frac{1}{z(1-z)} > 0. \end{aligned}$$

1573 Therefore,  $z \rightarrow f_\epsilon^+(\lambda, \mu, z)$  is twice continuously differentiable, positive and strictly convex on  
 1574  $([\lambda]_0^1, \mu)$ . Moreover,  $z \rightarrow f_\epsilon^+(\lambda, \mu, z)$  is decreasing on  $([\lambda]_0^1, \max\{g_\epsilon^+(\mu), \lambda\})$  and increasing on  
 1575  $(\max\{g_\epsilon^+(\mu), \lambda\}, \mu)$ . Using Lemma 23, we obtain

$$\begin{aligned} f_\epsilon^+(\lambda, \mu, \lambda) &= \text{kl}(\lambda, \mu), \\ \text{kl}(g_\epsilon^+(\mu), \mu) &= -(g_\epsilon^+(\mu) + g_\epsilon^-(1 - \mu)) \log(\mu(1 - e^\epsilon) + e^\epsilon) + \epsilon g_\epsilon^-(1 - \mu) \end{aligned}$$

$$= -\log(1 - \mu(1 - e^{-\epsilon})) - \epsilon g_{\epsilon}^{+}(\mu),$$

$$f_{\epsilon}^{+}(\lambda, \mu, g_{\epsilon}^{+}(\mu)) = -\log(1 - \mu(1 - e^{-\epsilon})) - \epsilon \lambda.$$

1576 By definition of the indicator function, we have  $d_{\epsilon}^{+}(\lambda, \mu) = 0$  if  $\mu \in [0, [\lambda]_0^1]$ . When  $\lambda \leq 0$ , for all  
 1577  $\mu \in (0, 1)$ , we have

$$\forall \mu \in (0, 1), \quad d_{\epsilon}^{+}(\lambda, \mu) = f_{\epsilon}^{+}([\lambda]_0^1, \mu, g_{\epsilon}^{+}(\mu)) = -\log(1 - \mu(1 - e^{-\epsilon})) - \epsilon [\lambda]_0^1,$$

1578 by using the properties of  $z \rightarrow f_{\epsilon}^{+}(\lambda, \mu, z)$  on  $(0, 1) = (g_{\epsilon}^{-}([\lambda]_0^1), 1)$  by Lemma 23. This function  
 1579 can be extended by continuity to  $\mu = 0 = g_{\epsilon}^{-}([\lambda]_0^1)$  with value  $d_{\epsilon}^{+}(\lambda, 0) = 0$ . When  $\lambda \in (0, 1)$  and  
 1580  $\mu \in (g_{\epsilon}^{-}(\lambda), 1)$ , we have

$$\forall \mu \in (0, 1), \quad d_{\epsilon}^{+}(\lambda, \mu) = f_{\epsilon}^{+}([\lambda]_0^1, \mu, g_{\epsilon}^{+}(\mu)) = -\log(1 - \mu(1 - e^{-\epsilon})) - \epsilon [\lambda]_0^1,$$

1581 by using the properties of  $z \rightarrow f_{\epsilon}^{+}(\lambda, \mu, z)$  on  $(g_{\epsilon}^{-}(\lambda), 1) = (g_{\epsilon}^{-}([\lambda]_0^1), 1)$  by Lemma 23. This  
 1582 function can be extended by continuity to  $(\lambda, \mu) = (0, 0) = \lim_{\lambda \rightarrow 0^{+}}(\lambda, g_{\epsilon}^{-}([\lambda]_0^1))$  with value  
 1583  $d_{\epsilon}^{+}(0, 0) = 0$ . In both cases, this function can be extended by continuity to  $\mu = 1$  with value  
 1584  $d_{\epsilon}^{+}(\lambda, 1) = \epsilon(1 - [\lambda]_0^1)$ .

1585 When  $\lambda \in (0, 1)$ , i.e.,  $[\lambda]_0^1 = \lambda$ , and  $\mu \in (\lambda, g_{\epsilon}^{-}(\lambda)) \subseteq (0, 1)$  by Lemma 23, we have

$$d_{\epsilon}^{+}(\lambda, \mu) = f_{\epsilon}^{+}(\lambda, \mu, \lambda) = \text{kl}(\lambda, \mu).$$

1586 This function can be extended by continuity to  $\mu = \lambda$  with value  $d_{\epsilon}^{+}(\lambda, \lambda) = 0$  since  $\text{kl}(\lambda, \lambda) = 0$ .  
 1587 Using Lemma 23, this function can be extended by continuity to  $\mu = g_{\epsilon}^{-}(\lambda)$  (i.e.,  $\lambda = g_{\epsilon}^{+}(\mu)$ ) with  
 1588 value

$$d_{\epsilon}^{+}(\lambda, g_{\epsilon}^{-}(\lambda)) = \text{kl}(\lambda, g_{\epsilon}^{-}(\lambda)) = \text{kl}(g_{\epsilon}^{+}(\mu), \mu) = -\log(1 - \mu(1 - e^{-\epsilon})) - \epsilon g_{\epsilon}^{+}(\mu).$$

1589 Therefore, we have

$$\forall \lambda \in (0, 1), \forall \mu \in [\lambda, g_{\epsilon}^{-}(\lambda)], \quad d_{\epsilon}^{+}(\lambda, \mu) = \text{kl}(\lambda, \mu).$$

1590 Using that  $\lim_{\lambda \rightarrow 0^{+}}[\lambda, g_{\epsilon}^{-}(\lambda)] = \{0\}$ , this function can be extended by continuity to  $\lambda = 0$  with  
 1591 value 0. Using that  $\lim_{\lambda \rightarrow 1^{-}}[\lambda, g_{\epsilon}^{-}(\lambda)] = \{1\}$ , this function can be extended by continuity to  $\lambda = 1$   
 1592 with value  $0 = \lim_{(\mu, \lambda) \rightarrow 1^{-}} -\log(1 - \mu(1 - e^{-\epsilon})) - \epsilon [\lambda]_0^1$ .

1593 Putting all the continuity arguments together, we have shown that  $(\lambda, \mu) \rightarrow d_{\epsilon}^{+}(\lambda, \mu)$  is jointly  
 1594 continuous on  $\mathbb{R} \times [0, 1]$ . Moreover, it is direct to see that, for all  $\mu \in [0, 1]$ , the function  $\lambda \rightarrow d_{\epsilon}^{+}(\lambda, \mu)$   
 1595 is constant on  $(-\infty, 0]$  and on  $[1, +\infty)$ . Then,

$$\forall \lambda \in (0, 1), \forall \mu \in [0, 1], \quad d_{\epsilon}^{+}(\lambda, \mu) = \begin{cases} 0 & \text{if } \mu \in [0, \lambda] \\ \text{kl}(\lambda, \mu) & \mu \in (\lambda, g_{\epsilon}^{-}(\lambda)] \\ -\log(1 - \mu(1 - e^{-\epsilon})) - \epsilon \lambda & \text{if } \mu \in (g_{\epsilon}^{-}(\lambda), 1] \end{cases}.$$

1596 Let  $\mu \in [0, 1]$  and  $\lambda \in (0, \mu)$ . Using that  $\mu \in (g_{\epsilon}^{-}(\lambda), 1]$  if and only if  $\lambda \in (0, g_{\epsilon}^{+}(\mu))$ . For all  
 1597  $\mu \in [0, 1]$ , the function  $\lambda \rightarrow d_{\epsilon}^{+}(\lambda, \mu)$  is positive and affine with negative slope  $-\epsilon$  on  $(0, g_{\epsilon}^{+}(\mu))$ .  
 1598 Let  $\lambda \in (g_{\epsilon}^{+}(\mu), \mu)$ . Direct computation yields that

$$\frac{\partial d_{\epsilon}^{+}}{\partial \lambda}(\lambda, \mu) = \frac{\partial \text{kl}}{\partial \lambda}(\lambda, \mu) = \log\left(\frac{\lambda(1 - \mu)}{(1 - \lambda)\mu}\right) < 0,$$

$$\lim_{\lambda \rightarrow g_{\epsilon}^{+}(\mu)^{+}} \frac{\partial d_{\epsilon}^{+}}{\partial \lambda}(\lambda, \mu) = -\epsilon = \lim_{\lambda \rightarrow g_{\epsilon}^{+}(\mu)^{-}} \frac{\partial d_{\epsilon}^{+}}{\partial \lambda}(\lambda, \mu),$$

$$\frac{\partial^2 d_{\epsilon}^{+}}{\partial \lambda^2}(\lambda, \mu) = \frac{\partial^2 \text{kl}}{\partial \lambda^2}(\lambda, \mu) = \frac{1}{\lambda(1 - \lambda)} > 0.$$

1599 For all  $\mu \in [0, 1]$ , the function  $\lambda \rightarrow d_{\epsilon}^{+}(\lambda, \mu)$  is continuously differentiable, positive, decreasing and  
 1600 convex on  $(0, \mu)$ . For all  $\mu \in [0, 1]$ , the function  $\lambda \rightarrow d_{\epsilon}^{+}(\lambda, \mu)$  is twice continuously differentiable,  
 1601 positive and strictly convex on  $(g_{\epsilon}^{+}(\mu), \mu)$ . Combining the above results concludes the part of  
 1602  $\lambda \mapsto d_{\epsilon}^{+}(\lambda, \mu)$  on  $(0, \mu)$ .

1603 Let  $\lambda \in (0, 1)$ . Let  $a > 0$  and  $k \in \mathbb{N}$ . The  $k$ -th derivative of  $u(x) = a(1 - ax)^{-1}$  on  $[0, 1]$  is  
 1604  $u^{(k)}(x) = (k - 1)!a^{k+1}(1 - ax)^{-(k+1)}$ . Then,

$$\forall \mu \in (g_{\epsilon}^{-}(\lambda), 1], \forall k \in \mathbb{N}, \quad \frac{\partial^k d_{\epsilon}^{+}}{\partial \mu^k}(\lambda, \mu) = \frac{(1 - e^{-\epsilon})^k (k - 1)!}{(1 - \mu(1 - e^{-\epsilon}))^k} > 0,$$

$$\begin{aligned}\forall \mu \in (\lambda, g_\epsilon^-(\lambda)], \quad & \frac{\partial d_\epsilon^+}{\partial \mu}(\lambda, \mu) = \frac{\mu - \lambda}{\mu(1 - \mu)} > 0, \\ & \frac{\partial^2 d_\epsilon^+}{\partial \mu^2}(\lambda, \mu) = \frac{(\mu - \lambda)^2 + \lambda(1 - \lambda)}{\mu^2(1 - \mu)^2} > 0, \\ & \frac{\partial^3 d_\epsilon^+}{\partial \mu^3}(\lambda, \mu) > 0.\end{aligned}$$

1605 Direct computation yields

$$\begin{aligned}\lim_{\mu \rightarrow g_\epsilon^-(\lambda)} \frac{\mu - \lambda}{\mu(1 - \mu)} &= (1 - e^{-\epsilon})(1 + \lambda(e^\epsilon - 1)), \\ \lim_{\mu \rightarrow g_\epsilon^-(\lambda)} \frac{1 - e^{-\epsilon}}{1 - \mu(1 - e^{-\epsilon})} &= (1 - e^{-\epsilon})(1 + \lambda(e^\epsilon - 1)), \\ \lim_{\mu \rightarrow g_\epsilon^-(\lambda)} \left\{ \frac{(\mu - \lambda)^2 + \lambda(1 - \lambda)}{\mu^2(1 - \mu)^2} - \frac{(1 - e^{-\epsilon})^2}{(1 - \mu(1 - e^{-\epsilon}))^2} \right\} &= \frac{\lambda(1 - \lambda)}{g_\epsilon^-(\lambda)^2(1 - g_\epsilon^-(\lambda))^2} > 0.\end{aligned}$$

1606 For all  $\lambda \in (0, 1)$ , the function  $\mu \rightarrow d_\epsilon^+(\lambda, \mu)$  is positive, three times differentiable with continuous  
1607 first derivative and increasing on  $(\lambda, 1]$ . For all  $\lambda \in (0, 1)$ , the function  $\mu \rightarrow d_\epsilon^+(\lambda, \mu)$  is strictly  
1608 convex on  $(\lambda, g_\epsilon^-(\lambda)]$  and  $(g_\epsilon^-(\lambda), 1]$ . The second derivative is discontinuous at  $g_\epsilon^-(\lambda)$  with gap  
1609  $\frac{\partial^2 d_\epsilon^+}{\partial \mu^2}(\lambda, g_\epsilon^-(\lambda)) - \lim_{\mu \rightarrow g_\epsilon^-(\lambda)^+} \frac{\partial^2 d_\epsilon^+}{\partial \mu^2}(\lambda, \mu) > 0$ . Thanks to the continuity of the first derivative and  
1610 the sign of the second derivative, the function  $\mu \rightarrow d_\epsilon^+(\lambda, \mu)$  is strict convexity on  $(\lambda, 1]$ .

1611 Let  $(\mu_1, \mu_2) \in [0, 1]^2$  and  $(\lambda_1, \lambda_2) \in (0, 1)^2$ . On the convex set  $\mathcal{F}_0 = \{(\lambda, \mu) \in (0, 1) \times [0, 1] \mid \mu \in$   
1612  $[0, \lambda]\}$ , the function  $d_\epsilon^-$  is null hence jointly convex. Let  $((\mu_1, \lambda_1), (\mu_2, \lambda_2)) \in (((0, 1) \times [0, 1]) \setminus \mathcal{F}_0)^2$ .  
1613 Let  $(z_1, z_2) \in [\lambda_1, \mu_1] \times [\lambda_2, \mu_2]$  be the minimizers realizing  $d_\epsilon^+(\lambda_1, \mu_1)$  and  $d_\epsilon^+(\lambda_2, \mu_2)$ . Since it is  
1614 a convex set, we have  $(\alpha\lambda_1 + (1 - \alpha)\lambda_2, \alpha\mu_1 + (1 - \alpha)\mu_2) \in ((0, 1) \times [0, 1]) \setminus \mathcal{F}_0$  for all  $\alpha \in [0, 1]$ .  
1615 Moreover, we have  $\alpha z_1 + (1 - \alpha)z_2 \in [\alpha\lambda_1 + (1 - \alpha)\lambda_2, \alpha\mu_1 + (1 - \alpha)\mu_2]$  for all  $\alpha \in [0, 1]$ . Using  
1616 the definition of  $d_\epsilon^+$  as an infimum, we obtain

$$\begin{aligned}& d_\epsilon^+(\alpha\lambda_1 + (1 - \alpha)\lambda_2, \alpha\mu_1 + (1 - \alpha)\mu_2) \\ & \leq \text{kl}(\alpha z_1 + (1 - \alpha)z_2, \alpha\mu_1 + (1 - \alpha)\mu_2) + \epsilon(\alpha z_1 + (1 - \alpha)z_2 - (\alpha\lambda_1 + (1 - \alpha)\lambda_2)) \\ & \leq \alpha(\text{kl}(z_1, \mu_1) + \epsilon(z_1 - \lambda_1)) + (1 - \alpha)(\text{kl}(z_2, \mu_2) + \epsilon(z_2 - \lambda_2)) \\ & = \alpha d_\epsilon^+(\lambda_1, \mu_1) + (1 - \alpha)d_\epsilon^+(\lambda_2, \mu_2)\end{aligned}$$

1617 where the second inequality comes from the joint convexity of the Kullback-Leibler divergence.  
1618 Combining both results, we have shown that the function  $d_\epsilon^+$  is jointly convex on  $(0, 1) \times [0, 1]$ .  $\square$

1619 Lemma 25 gather regularity properties of  $d_\epsilon^-$ . In particular, it gives a closed-form solution, which is a  
1620 key property used in our implementation to reduce the computational cost.

1621 **Lemma 25.** Let  $d_\epsilon^-$  as in Eq. (3), and  $g_\epsilon^\pm$  as in Eq. (30). For all  $\mu \in [0, 1]$  and all  $\lambda \in \mathbb{R}$ , we have

$$d_\epsilon^-(\lambda, \mu) = \begin{cases} 0 & \text{if } \mu \in [[\lambda]_0^1, 1] \\ -\log(1 + \mu(e^\epsilon - 1)) + \epsilon[\lambda]_0^1 & \text{if } \mu \in [0, g_\epsilon^+([\lambda]_0^1)) \\ \text{kl}(\lambda, \mu) & \text{if } \lambda \in (0, 1) \text{ and } \mu \in [g_\epsilon^+([\lambda]_0^1), [\lambda]_0^1) \end{cases}.$$

1622 The function  $(\lambda, \mu) \mapsto d_\epsilon^-(\lambda, \mu)$  is jointly continuous on  $\mathbb{R} \times [0, 1]$ . For all  $\mu \in [0, 1]$ , the function  
1623  $\lambda \mapsto d_\epsilon^-(\lambda, \mu)$  is constant on  $(-\infty, 0]$  and on  $[1, +\infty)$ . Then,

$$\forall \lambda \in (0, 1), \forall \mu \in [0, 1], \quad d_\epsilon^-(\lambda, \mu) = \begin{cases} 0 & \text{if } \mu \in [\lambda, 1] \\ \text{kl}(\lambda, \mu) & \text{if } \mu \in [g_\epsilon^+(\lambda), \lambda) \\ -\log(1 + \mu(e^\epsilon - 1)) + \epsilon\lambda & \text{if } \mu \in [0, g_\epsilon^+(\lambda)) \end{cases}.$$

1624 For all  $\mu \in [0, 1]$ , the function  $\lambda \mapsto d_\epsilon^-(\lambda, \mu)$  is continuously differentiable, positive, increasing and  
1625 convex on  $(\mu, 1)$ ; it is affine with positive slope  $\epsilon$  on  $(g_\epsilon^-(\mu), 1)$  and twice continuously differentiable  
1626 and strictly convex on  $(\mu, g_\epsilon^-(\mu))$ .



1627 For all  $\lambda \in (0, 1)$ , the function  $\mu \mapsto d_\epsilon^-(\lambda, \mu)$  is positive, three times differentiable with continuous  
 1628 first derivative, decreasing and strictly convex on  $[0, \lambda)$ ; its second derivative is discontinuous at  
 1629  $g_\epsilon^+(\lambda)$  with gap  $\lim_{\mu \rightarrow g_\epsilon^+(\lambda)^-} \frac{\partial^2 d_\epsilon^-}{\partial \mu^2}(\lambda, \mu) - \frac{\partial^2 d_\epsilon^-}{\partial \mu^2}(\lambda, g_\epsilon^+(\lambda)) < 0$ . Moreover, we have

$$\forall \mu \in [0, \lambda), \quad \frac{\partial d_\epsilon^-}{\partial \mu}(\lambda, \mu) = \begin{cases} -\frac{e^\epsilon - 1}{1 + \mu(e^\epsilon - 1)} & \text{if } \mu \in [0, g_\epsilon^+(\lambda)) \\ -\frac{\lambda - \mu}{\mu(1 - \mu)} & \text{if } \mu \in [g_\epsilon^+(\lambda), \lambda) \end{cases}.$$

1630 The function  $d_\epsilon^-$  is jointly convex on  $(0, 1) \times [0, 1]$ .

1631 *Proof.* Using Lemmas 22 and 23, we have

$$\begin{aligned} d_\epsilon^-(\lambda, \mu) &= d_\epsilon^+(1 - \lambda, 1 - \mu) \quad \text{and} \quad g_\epsilon^+(\lambda) = 1 - g_\epsilon^-(1 - \lambda), \\ \frac{\partial d_\epsilon^-}{\partial \mu}(\lambda, \mu) &= -\frac{\partial d_\epsilon^+}{\partial \mu}(1 - \lambda, 1 - \mu) \quad \text{and} \quad \frac{\partial^2 d_\epsilon^-}{\partial \mu^2}(\lambda, \mu) = \frac{\partial^2 d_\epsilon^+}{\partial \mu^2}(1 - \lambda, 1 - \mu). \end{aligned}$$

1632 Moreover, we have  $\text{kl}(\lambda, \mu) = \text{kl}(1 - \lambda, 1 - \mu)$  and

$$-\log(1 + \mu(e^\epsilon - 1)) + \epsilon[\lambda]_0^1 = -\log(1 - (1 - \mu)(1 - e^{-\epsilon})) - \epsilon[1 - \lambda]_0^1.$$

1633 Combining the above with properties of  $d_\epsilon^+$  in Lemma 24 concludes the proof.  $\square$

### 1634 G.1.1 Modified Divergence

1635 Let us define

$$\forall x > 0, \quad h(x) := \sqrt{1 + x^2} - 1 + \log\left(\frac{2}{x^2}(\sqrt{1 + x^2} - 1)\right). \quad (31)$$

1636 For all  $(\lambda, \mu, r) \in \mathbb{R} \times (0, 1) \times \mathbb{R}_+^*$ , we define

$$\begin{aligned} \tilde{d}_\epsilon^-(\lambda, \mu, r) &:= \mathbb{1}(\mu < [\lambda]_0^1) \inf_{z \in (\mu, [\lambda]_0^1)} \left\{ \text{kl}(z, \mu) + \frac{1}{r} h(r\epsilon(\lambda - z)) \right\}, \\ \tilde{d}_\epsilon^+(\lambda, \mu, r) &:= \mathbb{1}(\mu > [\lambda]_0^1) \inf_{z \in ([\lambda]_0^1, \mu)} \left\{ \text{kl}(z, \mu) + \frac{1}{r} h(r\epsilon(z - \lambda)) \right\}. \end{aligned} \quad (32)$$

1637 Lemma 26 shows a strong link between  $\tilde{d}_\epsilon^\pm$ . This symmetry property can be used to carry regularity  
 1638 properties from  $\tilde{d}_\epsilon^+$  to  $\tilde{d}_\epsilon^-$ , and vice versa.

1639 **Lemma 26.** Let  $\tilde{d}_\epsilon^\pm$  as in Eq. (32). For all  $(\lambda, \mu) \in \mathbb{R} \times [0, 1]$ , we have

$$\tilde{d}_\epsilon^+(1 - \lambda, 1 - \mu, r) = \tilde{d}_\epsilon^-(\lambda, \mu, r) \quad \text{and} \quad \tilde{d}_\epsilon^-(1 - \lambda, 1 - \mu, r) = \tilde{d}_\epsilon^+(\lambda, \mu, r).$$

1640 *Proof.* Using the definitions, the change of variable  $\tilde{z} = 1 - z$  and  $\text{kl}(1 - \tilde{z}, 1 - \mu) = \text{kl}(\tilde{z}, \mu)$ , we  
 1641 obtain

$$\begin{aligned} \tilde{d}_\epsilon^+(1 - \lambda, 1 - \mu, r) &= \mathbb{1}(\mu < [\lambda]_0^1) \inf_{z \in [1 - [\lambda]_0^1, 1 - \mu]} \left\{ \text{kl}(z, 1 - \mu) + \frac{1}{r} h(r\epsilon(\lambda - (1 - z))) \right\} \\ &= \mathbb{1}(\mu < [\lambda]_0^1) \inf_{\tilde{z} \in [\mu, [\lambda]_0^1]} \left\{ \text{kl}(1 - \tilde{z}, 1 - \mu) + \frac{1}{r} h(r\epsilon(\lambda - \tilde{z})) \right\} \\ &= \mathbb{1}(\mu < [\lambda]_0^1) \inf_{\tilde{z} \in [\mu, [\lambda]_0^1]} \left\{ \text{kl}(\tilde{z}, \mu) + \frac{1}{r} h(r\epsilon(\lambda - \tilde{z})) \right\} = \tilde{d}_\epsilon^-(\lambda, \mu, r). \end{aligned}$$

1642 The second equality is a consequence of the first.  $\square$

1643 Lemma 27 gathers regularity properties of the function  $h$  defined in Eq. (31).

1644 **Lemma 27.** *Let  $h$  as in Eq. (31). Then,*

$$\forall x > 0, \quad h'(x) = \frac{x}{\sqrt{x^2+1}+1} > 0 \quad \text{and} \quad h''(x) = \frac{1}{1+x^2+\sqrt{1+x^2}} > 0.$$

1645 *On  $\mathbb{R}_+^*$ , the function  $h$  is twice continuously differentiable, increasing and strictly convex. Moreover,*  
 1646 *it satisfies*

$$h(x) =_{x \rightarrow 0} x^2/4 + \mathcal{O}(x^4) \quad \text{and} \quad h(x) =_{x \rightarrow +\infty} x - \mathcal{O}(\log(x)).$$

1647 *Proof.* For all  $x > 0$ ,  $h_1(x) = x + \log(x)$ ,  $h_2(x) = \sqrt{1+x^2} - 1$  and  $h_3(x) = \sqrt{1+x^2} - x$ . Then

$$h'_1(x) = 1 + \frac{1}{x}, \quad h'_2(x) = \frac{x}{\sqrt{1+x^2}} \quad \text{and} \quad h'_3(x) = \frac{x}{\sqrt{1+x^2}} - 1,$$

1648 Then, we have

$$\forall x > 0, \quad h(x) = h_1(h_2(x)) - 2\log(x) + \log 2.$$

1649 Therefore, we have

$$\begin{aligned} h'(x) &= h'_2(x)h'_1(h_2(x)) - \frac{2}{x} = \frac{x}{\sqrt{1+x^2}} \left(1 + \frac{1}{\sqrt{1+x^2}-1}\right) - \frac{2}{x} \\ &= \frac{x}{\sqrt{1+x^2}-1} - \frac{2}{x} = \sqrt{1 + \frac{1}{x^2}} - \frac{1}{x} = h_3(1/x). \end{aligned}$$

1650 Note that

$$\sqrt{1 + \frac{1}{x^2}} - \frac{1}{x} = \frac{x}{\sqrt{x^2+1}+1}.$$

1651 Moreover, we have w

$$h''(x) = -\frac{1}{x^2}h'_3(1/x) = -\frac{1}{x^2} \left( \frac{1/x}{\sqrt{1+(1/x)^2}} - 1 \right) = \frac{1}{1+x^2+\sqrt{1+x^2}}.$$

1652 By taking the limit, we have  $\lim_{x \rightarrow 0^+} h(x) = 0$ . Moreover, we see that  $\lim_{x \rightarrow 0^+} h'(x) = 0$  and  
 1653  $\lim_{x \rightarrow 0^+} h''(x) = 1/2$ . Therefore, one can conclude that  $h(x) =_{x \rightarrow 0} x^2/4 + \mathcal{O}(x^4)$  by Taylor  
 1654 expansion. The second result is obtained directly by limit.  $\square$

1655 Lemma 28 provides upper and lower bound on the function  $r \mapsto h(rx)/r$  involved in the definition  
 1656 of  $\tilde{d}_\epsilon^\pm$ .

1657 **Lemma 28.** *Let  $h$  as in Eq. (31). Let  $\kappa(r, x) = h(rx)/r - x$  for all  $r > 0$  and all  $x \in \mathbb{R}_+^*$ . Then,*  
 1658 *we have*

$$\forall r > 0, \quad \frac{\partial \kappa}{\partial r}(r, x) = \frac{rxh'(rx) - h(rx)}{r^2} = \log \left( \frac{1}{2}(\sqrt{1+(rx)^2} + 1) \right) > 0.$$

1659 *On  $\mathbb{R}_+^*$ , the function  $r \mapsto \kappa(r, x)$  is increasing. Moreover, we have*

$$\forall r > 0, \forall x \in \mathbb{R}_+, \quad 0 \leq r\kappa(r, x) + \log(1+2xr) + 1 \leq 1 + \log 4,$$

1660 *Proof.* Using Lemma 27 and the definition in Eq. (31), we obtain that

$$\forall x > 0, \quad xh'(x) - h(x) = -\log \left( \frac{2}{x^2} (\sqrt{1+x^2} - 1) \right) = \log \left( \frac{1}{2}(\sqrt{1+x^2} + 1) \right) > 0,$$

1661 where we used that  $\sqrt{1+x^2} + 1 > 2$  for the last inequality. Let us define

$$\forall x \in \mathbb{R}_+, \quad g_1(x) = \frac{2(1+2x)}{\sqrt{1+x^2}+1}.$$

1662 Then, we obtain  $g_1(0) = 1$ ,  $\lim_{x \rightarrow +\infty} g_1(x) = 4$  and

$$g'_1(x) = 2 \frac{2+2\sqrt{1+x^2}-x}{\sqrt{1+x^2}(\sqrt{1+x^2}+1)^2} > 2 \frac{2+x}{\sqrt{1+x^2}(\sqrt{1+x^2}+1)^2} > 0.$$

1663 Since  $g_1$  is strictly increasing on  $\mathbb{R}_+^*$ , we obtain  $\log g_1(x) \geq \log g_1(0) = 0$  and  $\log g_1(x) \leq \log 4$  for  
 1664 all  $x \in \mathbb{R}_+$ .

1665 By definition, we obtain

$$\begin{aligned} r\kappa(r, x) + \log(1 + 2xr) + 1 &= h(rx) - rx + 1 + \log(1 + 2xr) \\ &= \sqrt{1 + (rx)^2} - rx + \log\left(\frac{2(1 + 2xr)}{\sqrt{1 + r^2x^2} + 1}\right). \end{aligned}$$

1666 Using that  $0 \leq \sqrt{1 + x^2} - x \leq 1$  on  $\mathbb{R}_+$ , we obtain

$$\begin{aligned} r\kappa(r, x) + \log(1 + 2xr) + 1 &\geq \log\left(\frac{2(1 + 2xr)}{\sqrt{1 + r^2x^2} + 1}\right) = \log(g_1(rx)) \geq 0, \\ r\kappa(r, x) + \log(1 + 2xr) + 1 &\leq 1 + \log(g_1(rx)) \leq 1 + \log 4. \end{aligned}$$

1667 This concludes the proof. □

1668 Lemma 29 provides lower and upper bounds on the gap between  $\tilde{d}_\epsilon^\pm$  and  $d_\epsilon^\pm$ .

1669 **Lemma 29.** *Let  $d_\epsilon^\pm$  and  $\tilde{d}_\epsilon^\pm$  as in Eq. (3) and (32). For all  $(\lambda, \mu, r) \in \mathbb{R} \times (0, 1) \times \mathbb{R}_+^*$  such that*  
 1670  *$[\lambda]_0^1 < \mu$ . Then,*

$$d_\epsilon^+(\lambda, \mu) \leq \tilde{d}_\epsilon^+(\lambda, \mu, r) + \frac{\log(1 + 2\epsilon r) + 1}{r}.$$

1671 For all  $(\lambda, \mu, r) \in \mathbb{R} \times (0, 1) \times \mathbb{R}_+^*$  such that  $[\lambda]_0^1 > \mu$ . Then,

$$d_\epsilon^-(\lambda, \mu) \leq \tilde{d}_\epsilon^-(\lambda, \mu, r) + \frac{\log(1 + 2\epsilon r) + 1}{r}.$$

1672 For all  $(\lambda, \mu, r) \in [0, 1] \times (0, 1) \times \mathbb{R}_+^*$  such that  $\lambda < \mu$ . Then,

$$d_\epsilon^+(\lambda, \mu) \geq \tilde{d}_\epsilon^+(\lambda, \mu, r) - \frac{\log 4}{r}.$$

1673 For all  $(\mu, \lambda, r) \in [0, 1] \times \mathbb{R}_+^*$  such that  $\lambda > \mu$ . Then,

$$d_\epsilon^-(\lambda, \mu) \geq \tilde{d}_\epsilon^-(\lambda, \mu, r) - \frac{\log 4}{r}.$$

1674 *Proof.* Since  $\mu \in (0, 1)$ , we have  $[\lambda]_0^1 = \max\{0, \lambda\}$ . Therefore, we have  $z - \lambda \geq z - [\lambda]_0^1$  and  
 1675  $z - [\lambda]_0^1 \in (0, \mu - [\lambda]_0^1) \subset (0, 1)$  for all  $z \in ([\lambda]_0^1, \mu)$ . Using Lemmas 27 and 28 and  $\epsilon > 0$ , we  
 1676 obtain, for all  $r > 0$  and all  $z \in ([\lambda]_0^1, \mu)$ ,

$$\begin{aligned} \epsilon(z - [\lambda]_0^1) &\leq \frac{1}{r}h(r\epsilon(z - [\lambda]_0^1)) + \frac{\log(1 + 2\epsilon(z - [\lambda]_0^1)r) + 1}{r} \\ &\leq \frac{1}{r}h(r\epsilon(z - \lambda)) + \frac{\log(1 + 2\epsilon r) + 1}{r}. \end{aligned}$$

1677 Therefore, for all  $z \in ([\lambda]_0^1, \mu)$ , we obtain that

$$\text{kl}(z, \mu) + \epsilon(z - [\lambda]_0^1) \leq \text{kl}(z, \mu) + \frac{1}{r}h(r\epsilon(z - \lambda)) + \frac{\log(1 + 2\epsilon r) + 1}{r}.$$

1678 Taking the infimum over  $z \in ([\lambda]_0^1, \mu)$  on both sides of both inequalities and using that

$$\begin{aligned} d_\epsilon^+(\lambda, \mu) &= \inf_{z \in ([\lambda]_0^1, \mu)} \{\text{kl}(z, \mu) + \epsilon(z - [\lambda]_0^1)\} = \inf_{z \in ([\lambda]_0^1, \mu)} \{\text{kl}(z, \mu) + \epsilon(z - [\lambda]_0^1)\}, \\ \tilde{d}_\epsilon^+(\lambda, \mu, r) &= \inf_{z \in ([\lambda]_0^1, \mu)} \left\{ \text{kl}(z, \mu) + \frac{1}{r}h(r\epsilon(z - \lambda)) \right\}, \end{aligned}$$

1679 we obtain

$$d_\epsilon^+(\lambda, \mu) \leq \tilde{d}_\epsilon^+(\lambda, \mu, r) + \frac{\log(1 + 2\epsilon r) + 1}{r}.$$

1680 This concludes the proof of the first result. Using Lemmas 22 and 26 yields the second result.

1681 Suppose that  $\lambda \in [0, 1]$ , hence  $\lambda = [\lambda]_0^1$ . Using Lemmas 27 and 28 and  $\epsilon > 0$ , we obtain, for all  
1682  $r > 0$  and all  $z \in ([\lambda]_0^1, \mu)$ ,

$$\frac{1}{r}h(r\epsilon(z - \lambda)) \leq \epsilon(z - \lambda) + \frac{\log 4 - \log(1 + 2\epsilon(z - \lambda)r)}{r} \leq \epsilon(z - [\lambda]_0^1) + \frac{\log 4}{r}.$$

1683 Adding  $\text{kl}(z, \mu)$  on both sides and taking the infimum over  $z \in ([\lambda]_0^1, \mu)$  on both sides of both  
1684 inequalities yields the proof of third result. Using Lemmas 22 and 26 yields the forth result.  $\square$

1685 Lemma 30 gathers regularity properties on the modified divergences  $\tilde{d}_\epsilon^+$ . In particular, it gives a  
1686 closed-form solution based on an implicit solution of a fixed-point equation. This is a key property  
1687 used in our implementation to reduce the computational cost.

1688 **Lemma 30.** Let  $\tilde{d}_\epsilon^+$  as in Eq. (32), and  $g_\epsilon^\pm$  as in Eq. (30). For all  $\mu \in (0, 1)$ ,  $\lambda \in \mathbb{R}$  and  $r > 0$ , we  
1689 have

$$\begin{aligned} \tilde{d}_\epsilon^+(\lambda, \mu, r) &= \begin{cases} 0 & \text{if } \mu \in (0, [\lambda]_0^1) \\ \text{kl}(x_\epsilon^+(\lambda, \mu, r) + g_\epsilon^+(\mu), \mu) + \frac{1}{r}h(r\epsilon(x_\epsilon^+(\lambda, \mu, r) + g_\epsilon^+(\mu) - \lambda)) & \text{if } \mu \in ([\lambda]_0^1, 1) \end{cases}, \end{aligned}$$

1690 where  $x_\epsilon^+(\lambda, \mu, r) \in (\max\{0, \lambda - g_\epsilon^+(\mu)\}, \mu - g_\epsilon^+(\mu))$  is the unique solution for  $x \in (\max\{0, \lambda -$   
1691  $g_\epsilon^+(\mu)\}, \mu - g_\epsilon^+(\mu))$  of the equation

$$\log\left(1 + \frac{x}{g_\epsilon^+(\mu)(1 - x - g_\epsilon^+(\mu))}\right) + \epsilon\left(\frac{r\epsilon(x + g_\epsilon^+(\mu) - \lambda)}{\sqrt{(r\epsilon(x + g_\epsilon^+(\mu) - \lambda))^2 + 1 + 1}} - 1\right) = 0.$$

1692 For all  $(\mu, r) \in (0, 1) \times \mathbb{R}_+^*$ , the function  $\lambda \mapsto \tilde{d}_\epsilon^+(\lambda, \mu, r)$  is positive, twice continuously differ-  
1693 entiable, decreasing and strictly convex on  $(-\infty, \mu)$ ; it satisfies  $\lim_{\lambda \rightarrow \mu^-} \tilde{d}_\epsilon^+(\lambda, \mu, r) = 0$  and  
1694  $\lim_{\lambda \rightarrow -\infty} \tilde{d}_\epsilon^+(\lambda, \mu, r) = +\infty$ .

1695 For all  $(\lambda, r) \in \mathbb{R} \times \mathbb{R}_+^*$ , the function  $\mu \mapsto \tilde{d}_\epsilon^+(\lambda, \mu, r)$  is positive, twice continuously differentiable,  
1696 increasing and strictly convex on  $([\lambda]_0^1, 1)$ . Moreover, we have

$$\forall \mu \in ([\lambda]_0^1, 1), \quad \frac{\partial \tilde{d}_\epsilon^+}{\partial \mu}(\lambda, \mu, r) = \frac{\mu - g_\epsilon^+(\mu) - x_\epsilon^+(\lambda, \mu, r)}{\mu(1 - \mu)}.$$

1697 For all  $(\lambda, \mu) \in \mathbb{R} \times (0, 1)$  such that  $\mu \in (0, [\lambda]_0^1]$ , the function  $r \mapsto \tilde{d}_\epsilon^+(\lambda, \mu, r)$  is the zero function.  
1698 For all  $(\lambda, \mu) \in \mathbb{R} \times (0, 1)$  such that  $\mu \in ([\lambda]_0^1, 1)$ , the function  $r \mapsto \tilde{d}_\epsilon^+(\lambda, \mu, r)$  is positive,  
1699 continuously differentiable and increasing on  $\mathbb{R}_+$ .

1700 *Proof.* By definition of the indicator function, we have  $\tilde{d}_\epsilon^+(\lambda, \mu, r) = 0$  if  $\mu \in (0, [\lambda]_0^1]$ . Let  $(\lambda, \mu)$   
1701 such that  $\mu \notin (0, [\lambda]_0^1]$ , i.e.,  $([\lambda]_0^1, \mu)$  is non-empty. Since  $\mu \in (0, 1)$ , this implies that  $\lambda \in (-\infty, 1)$   
1702 necessarily, i.e.,  $[\lambda]_0^1 = \max\{0, \lambda\}$ .

1703 Recall that  $\tilde{d}_\epsilon^+(\lambda, \mu, r) = \mathbb{1}_{(\mu > [\lambda]_0^1)} \inf_{z \in ([\lambda]_0^1, \mu)} \tilde{f}_\epsilon^+(\lambda, \mu, r, z)$  where  $\tilde{f}_\epsilon^+(\lambda, \mu, r, z) = \text{kl}(z, \mu) +$   
1704  $\frac{1}{r}h(r\epsilon(z - \lambda))$ . Using Lemma 27, direct computations yield that, for all  $z \in ([\lambda]_0^1, \mu)$ ,

$$\begin{aligned} \frac{\partial \tilde{f}_\epsilon^+}{\partial z}(\lambda, \mu, r, z) &= \log\left(\frac{z(1 - \mu)}{(1 - z)\mu}\right) + \epsilon h'(r\epsilon(z - \lambda)) \\ &= \log\left(\frac{z(1 - \mu)}{(1 - z)\mu}\right) + \epsilon \frac{r\epsilon(z - \lambda)}{\sqrt{(r\epsilon(z - \lambda))^2 + 1 + 1}}, \\ \frac{\partial^2 \tilde{f}_\epsilon^+}{\partial z^2}(\lambda, \mu, r, z) &= \frac{1}{z(1 - z)} + r\epsilon^2 h''(r\epsilon(z - \lambda)) > 0. \end{aligned}$$

Therefore,  $z \rightarrow \tilde{f}_\epsilon^+(\lambda, \mu, r, z)$  is twice continuously differentiable, positive and strictly convex on  $([\lambda]_0^1, \mu)$ . Moreover, we have

$$\lim_{z \rightarrow \mu} \frac{\partial \tilde{f}_\epsilon^+}{\partial z}(\lambda, \mu, r, z) = \epsilon h'(r\epsilon(\mu - \lambda)) > 0,$$

$$\frac{\partial \tilde{f}_\epsilon^+}{\partial z}(\lambda, \mu, r, g_\epsilon^+(\mu)) = -\epsilon \left( 1 - \frac{r\epsilon(z - \lambda)}{\sqrt{(r\epsilon(z - \lambda))^2 + 1} + 1} \right) < 0,$$

$$\text{When } [\lambda]_0^1 > g_\epsilon^+(\mu), \quad \frac{\partial \tilde{f}_\epsilon^+}{\partial z}(\lambda, \mu, r, \lambda) = \log \left( \frac{\lambda(1 - \mu)}{(1 - \lambda)\mu} \right) < 0.$$

Note that  $\max\{[\lambda]_0^1, g_\epsilon^+(\mu)\} = \max\{\lambda, g_\epsilon^+(\mu)\}$  since  $\mu \in (0, 1)$ . Using that  $z \rightarrow \frac{\partial \tilde{f}_\epsilon^+}{\partial z}(\lambda, \mu, r, z)$  is continuously differentiable and increasing on  $([\lambda]_0^1, \mu)$ , with negative value at  $\max\{\lambda, g_\epsilon^+(\mu)\}$  and finite positive limit at  $\mu$ ,  $z \mapsto \tilde{f}_\epsilon^+(\lambda, \mu, r, z)$  admit a unique minimizer on  $(\max\{\lambda, g_\epsilon^+(\mu)\}, \mu)$ . Let  $\tilde{g}_\epsilon^+(\lambda, \mu, r) \in (\max\{\lambda, g_\epsilon^+(\mu)\}, \mu)$  be defined as this unique minimizer, defined implicitly as solution for  $z \in (\max\{\lambda, g_\epsilon^+(\mu)\}, \mu)$  of the equation

$$\log \left( \frac{z(1 - \mu)}{(1 - z)\mu} \right) + \epsilon \frac{r\epsilon(z - \lambda)}{\sqrt{(r\epsilon(z - \lambda))^2 + 1} + 1} = 0.$$

Then, we have  $\frac{\partial \tilde{f}_\epsilon^+}{\partial z}(\lambda, \mu, r, z) = 0$  if and only if  $z = \tilde{g}_\epsilon^+(\lambda, \mu, r)$ . Moreover,  $z \mapsto \tilde{f}_\epsilon^+(\lambda, \mu, r, z)$  is decreasing on  $([\lambda]_0^1, \tilde{g}_\epsilon^+(\lambda, \mu, r))$  and increasing on  $(\tilde{g}_\epsilon^+(\lambda, \mu, r), \mu)$ .

Let us define  $z = g_\epsilon^+(\mu) + x$  where  $x \in (\max\{0, \lambda - g_\epsilon^+(\mu)\}, \mu - g_\epsilon^+(\mu))$ . Then, we have

$$\begin{aligned} & \frac{\partial \tilde{f}_\epsilon^+}{\partial z}(\lambda, \mu, r, g_\epsilon^+(\mu) + x) \\ &= \log \left( 1 + \frac{x}{g_\epsilon^+(\mu)(1 - x - g_\epsilon^+(\mu))} \right) + \epsilon \left( \frac{r\epsilon(x + g_\epsilon^+(\mu) - \lambda)}{\sqrt{(r\epsilon(x + g_\epsilon^+(\mu) - \lambda))^2 + 1} + 1} - 1 \right) \end{aligned}$$

Therefore, we have  $\tilde{g}_\epsilon^+(\lambda, \mu, r) = g_\epsilon^+(\mu) + x_\epsilon^+(\lambda, \mu, r)$  where  $x_\epsilon^+(\lambda, \mu, r) \in (\max\{0, \lambda - g_\epsilon^+(\mu)\}, \mu - g_\epsilon^+(\mu))$  is the solution for  $x \in (\max\{0, \lambda - g_\epsilon^+(\mu)\}, \mu - g_\epsilon^+(\mu))$  of the equation

$$\log \left( 1 + \frac{x}{g_\epsilon^+(\mu)(1 - x - g_\epsilon^+(\mu))} \right) + \epsilon \left( \frac{r\epsilon(x + g_\epsilon^+(\mu) - \lambda)}{\sqrt{(r\epsilon(x + g_\epsilon^+(\mu) - \lambda))^2 + 1} + 1} - 1 \right) = 0.$$

When  $\lambda \in (0, 1)$  and  $\mu \rightarrow \lambda = [\lambda]_0^1$ , it is direct to see that  $\tilde{g}_\epsilon^+(\lambda, \mu, r) \rightarrow \lambda$ . Then, we have

$$\lim_{\mu \rightarrow \lambda^+} \tilde{d}_\epsilon^+(\lambda, \mu, r) = \lim_{\mu \rightarrow \lambda^+} \{\text{kl}(\tilde{g}_\epsilon^+(\lambda, \mu, r), \mu)\} + \frac{1}{r} \lim_{\tilde{g}_\epsilon^+(\lambda, \mu, r) \rightarrow \lambda^+} \{h(r\epsilon(\tilde{g}_\epsilon^+(\lambda, \mu, r) - \lambda))\} = 0.$$

Direct computation yields that, for  $z \in ([\lambda]_0^1, \mu) \subset (0, 1)$ ,

$$\begin{aligned} \frac{\partial \tilde{f}_\epsilon^+}{\partial \mu}(\lambda, \mu, r, z) &= \frac{\mu - z}{\mu(1 - \mu)} > 0, \\ \frac{\partial^2 \tilde{f}_\epsilon^+}{\partial \mu^2}(\lambda, \mu, r, z) &= \frac{(\mu - z)^2 + z(1 - z)}{\mu^2(1 - \mu)^2} > 0, \\ \frac{\partial^2 \tilde{f}_\epsilon^+}{\partial z^2}(\lambda, \mu, r, z) &= \frac{1}{z(1 - z)} + r\epsilon^2 h''(r\epsilon(z - \lambda)) > 0, \\ \frac{\partial^2 \tilde{f}_\epsilon^+}{\partial \mu \partial z}(\lambda, \mu, r, z) &= -\frac{1}{\mu(1 - \mu)} < 0. \end{aligned}$$

Since  $\frac{\partial \tilde{f}_\epsilon^+}{\partial z}(\lambda, \mu, r, g_\epsilon^+(\lambda, \mu, r)) = 0$ , the implicit function theorem yields that

$$\frac{\partial \tilde{g}_\epsilon^+}{\partial \mu}(\lambda, \mu, r) = -\frac{\frac{\partial^2 \tilde{f}_\epsilon^+}{\partial \mu \partial z}(\lambda, \mu, r, g_\epsilon^+(\lambda, \mu, r))}{\frac{\partial^2 \tilde{f}_\epsilon^+}{\partial z^2}(\lambda, \mu, r, g_\epsilon^+(\lambda, \mu, r))} > 0.$$

1720 Moreover, for  $\mu \in ([\lambda]_0^1, 1)$ ,

$$\begin{aligned}\frac{\partial \tilde{d}_\epsilon^+}{\partial \mu}(\lambda, \mu, r) &= \frac{\partial \tilde{f}_\epsilon^+}{\partial \mu}(\lambda, \mu, r, \tilde{g}_\epsilon^+(\lambda, \mu, r)) + \frac{\partial \tilde{g}_\epsilon^+}{\partial \mu}(\lambda, \mu, r) \frac{\partial \tilde{f}_\epsilon^+}{\partial z}(\lambda, \mu, r, \tilde{g}_\epsilon^+(\lambda, \mu, r)) \\ &= \frac{\partial \tilde{f}_\epsilon^+}{\partial \mu}(\lambda, \mu, r, \tilde{g}_\epsilon^+(\lambda, \mu, r)) = \frac{\mu - g_\epsilon^+(\mu) - x_\epsilon^+(\lambda, \mu, r)}{\mu(1 - \mu)} > 0, \\ \frac{\partial^2 \tilde{d}_\epsilon^+}{\partial \mu^2}(\lambda, \mu, r) &= \frac{\partial^2 \tilde{f}_\epsilon^+}{\partial \mu^2}(\lambda, \mu, r, \tilde{g}_\epsilon^+(\lambda, \mu, r)) \frac{\partial \tilde{g}_\epsilon^+}{\partial \mu}(\lambda, \mu, r) > 0.\end{aligned}$$

1721 Therefore, for all  $(\lambda, r) \in \mathbb{R} \times \mathbb{R}_+^*$ , the function  $\mu \mapsto \tilde{d}_\epsilon^+(\lambda, \mu, r)$  is positive, twice continuously  
1722 differentiable, increasing and strictly convex on  $([\lambda]_0^1, 1)$ .

1723 Let  $(\mu, r) \in (0, 1) \times \mathbb{R}_+^*$ . Direct computation yields that, for  $z \in ([\lambda]_0^1, \mu) \subset (0, 1)$ ,

$$\begin{aligned}\frac{\partial \tilde{f}_\epsilon^+}{\partial \lambda}(\lambda, \mu, r, z) &= -\epsilon h'(r\epsilon(z - \lambda)) < 0, \\ \frac{\partial^2 \tilde{f}_\epsilon^+}{\partial \lambda \partial z}(\lambda, \mu, r, z) &= -r\epsilon^2 h''(r\epsilon(z - \lambda)) < 0.\end{aligned}$$

1724 Since  $\frac{\partial \tilde{f}_\epsilon^+}{\partial z}(\lambda, \mu, r, g_\epsilon^+(\lambda, \mu, r)) = 0$ , the implicit function theorem yields that

$$\frac{\partial \tilde{g}_\epsilon^+}{\partial \lambda}(\lambda, \mu, r) = -\frac{\frac{\partial^2 \tilde{f}_\epsilon^+}{\partial \lambda \partial z}(\lambda, \mu, r, g_\epsilon^+(\lambda, \mu, r))}{\frac{\partial^2 \tilde{f}_\epsilon^+}{\partial z^2}(\lambda, \mu, r, g_\epsilon^+(\lambda, \mu, r))} = \frac{r\epsilon^2 h''(r\epsilon(g_\epsilon^+(\lambda, \mu, r) - \lambda))}{\frac{1}{z(1-z)} + r\epsilon^2 h''(r\epsilon(g_\epsilon^+(\lambda, \mu, r) - \lambda))} < 1.$$

1725 Direct computation yields that, for  $\lambda \in (-\infty, \mu)$ ,

$$\begin{aligned}\frac{\partial \tilde{d}_\epsilon^+}{\partial \lambda}(\lambda, \mu, r) &= \frac{\partial \tilde{f}_\epsilon^+}{\partial \lambda}(\lambda, \mu, r, \tilde{g}_\epsilon^+(\lambda, \mu, r)) + \frac{\partial \tilde{g}_\epsilon^+}{\partial \lambda}(\lambda, \mu, r) \frac{\partial \tilde{f}_\epsilon^+}{\partial z}(\lambda, \mu, r, \tilde{g}_\epsilon^+(\lambda, \mu, r)) \\ &= \frac{\partial \tilde{f}_\epsilon^+}{\partial \lambda}(\lambda, \mu, r, \tilde{g}_\epsilon^+(\lambda, \mu, r)) = -\epsilon h'(r\epsilon(\tilde{g}_\epsilon^+(\lambda, \mu, r) - \lambda)) < 0, \\ \frac{\partial^2 \tilde{d}_\epsilon^+}{\partial \lambda^2}(\lambda, \mu, r) &= r\epsilon^2 \left(1 - \frac{\partial \tilde{g}_\epsilon^+}{\partial \lambda}(\lambda, \mu, r)\right) h''(r\epsilon(\tilde{g}_\epsilon^+(\lambda, \mu, r) - \lambda)) > 0.\end{aligned}$$

1726 Therefore, for all  $(\mu, r) \in (0, 1) \times \mathbb{R}_+^*$ , the function  $\lambda \mapsto \tilde{d}_\epsilon^+(\lambda, \mu, r)$  is positive, twice continuously  
1727 differentiable, decreasing and strictly convex on  $(-\infty, \mu)$ . Similarly as above, it is direct to see that  
1728  $\lim_{\lambda \rightarrow \mu^-} \tilde{d}_\epsilon^+(\lambda, \mu, r) = 0$  and  $\lim_{\lambda \rightarrow -\infty} \tilde{d}_\epsilon^+(\lambda, \mu, r) = +\infty$ .

1729 Let  $(\lambda, \mu) \in \mathbb{R} \times (0, 1)$ . When  $\mu \in (0, [\lambda]_0^1]$ , we have  $\tilde{d}_\epsilon^+(\lambda, \mu, r) = 0$  for all  $r \in [1, +\infty)$ , hence  
1730  $r \mapsto \tilde{d}_\epsilon^+(\lambda, \mu, r)$  is non-decreasing. Let  $\kappa$  as in Lemma 28. Using Lemma 28, we have

$$\forall z > \lambda, \quad \frac{\partial \tilde{f}_\epsilon^+}{\partial r}(\lambda, \mu, r, z) = \frac{\partial \kappa}{\partial r}(r, \epsilon(z - \lambda)) > 0.$$

1731 When  $\mu \in ([\lambda]_0^1, 1)$ , we have  $\tilde{g}_\epsilon^+(\lambda, \mu, r) \in (\max\{\lambda, g_\epsilon^+(\mu)\}, \mu)$  and, for all  $r > 0$ ,

$$\begin{aligned}\frac{\partial \tilde{d}_\epsilon^+}{\partial r}(\lambda, \mu, r) &= \frac{\partial \tilde{f}_\epsilon^+}{\partial r}(\lambda, \mu, r, \tilde{g}_\epsilon^+(\lambda, \mu, r)) + \frac{\partial \tilde{g}_\epsilon^+}{\partial r}(\lambda, \mu, r) \frac{\partial \tilde{f}_\epsilon^+}{\partial z}(\lambda, \mu, r, \tilde{g}_\epsilon^+(\lambda, \mu, r)) \\ &= \frac{\partial \tilde{f}_\epsilon^+}{\partial r}(\lambda, \mu, r, \tilde{g}_\epsilon^+(\lambda, \mu, r)) > 0,\end{aligned}$$

1732 where we used that  $\tilde{g}_\epsilon^+(\lambda, \mu, r) > \lambda$ . This concludes the last part of the proof.  $\square$

1733 Lemma 31 gathers regularity properties on the modified divergences  $\tilde{d}_\epsilon^-$ . In particular, it gives a  
1734 closed-form solution based on an implicit solution of a fixed-point equation. This is a key property  
1735 used in our implementation to reduce the computational cost.

1736 **Lemma 31.** Let  $\tilde{d}_\epsilon^-$  as in Eq. (32), and  $g_\epsilon^\pm$  as in Eq. (30). For all  $\mu \in (0, 1)$ ,  $\lambda \in \mathbb{R}$  and  $r > 0$ , we  
 1737 have

$$\begin{aligned} & \tilde{d}_\epsilon^-(\lambda, \mu, r) \\ &= \begin{cases} 0 & \text{if } \mu \in [[\lambda]_0^1, 1) \\ \text{kl}(g_\epsilon^-(\mu) - x_\epsilon^-(\lambda, \mu, r), \mu) + \frac{1}{r}h(r\epsilon(x_\epsilon^-(\lambda, \mu, r) + \lambda - g_\epsilon^-(\mu))) & \text{if } \mu \in (0, [\lambda]_0^1) \end{cases}, \end{aligned}$$

1738 where  $x_\epsilon^-(\lambda, \mu, r) := x_\epsilon^+(1 - \lambda, 1 - \mu, r) \in (\max\{g_\epsilon^-(\mu) - \lambda, 0\}, g_\epsilon^-(\mu) - \mu)$  is the solution for  
 1739  $x \in (\max\{g_\epsilon^-(\mu) - \lambda, 0\}, g_\epsilon^-(\mu) - \mu)$  of the equation

$$\log \left( 1 + \frac{x}{(1 - g_\epsilon^-(\mu))(g_\epsilon^-(\mu) - x)} \right) + \epsilon \left( \frac{r\epsilon(x - g_\epsilon^-(\mu) + \lambda)}{\sqrt{(r\epsilon(x - g_\epsilon^-(\mu) + \lambda))^2 + 1 + 1}} - 1 \right) = 0.$$

1740 For all  $(\mu, r) \in (0, 1) \times \mathbb{R}_+^*$ , the function  $\lambda \mapsto \tilde{d}_\epsilon^-(\lambda, \mu, r)$  is positive, twice continuously dif-  
 1741 ferentiable, increasing and strictly convex on  $(\mu, +\infty)$ ; it satisfies  $\lim_{\lambda \rightarrow \mu^+} \tilde{d}_\epsilon^-(\lambda, \mu, r) = 0$  and  
 1742  $\lim_{\lambda \rightarrow +\infty} \tilde{d}_\epsilon^-(\lambda, \mu, r) = +\infty$ .

1743 For all  $(\lambda, r) \in \mathbb{R} \times \mathbb{R}_+^*$ , the function  $\mu \mapsto \tilde{d}_\epsilon^-(\lambda, \mu, r)$  is positive, twice continuously differentiable,  
 1744 decreasing and strictly convex on  $(0, [\lambda]_0^1)$ . Moreover, we have

$$\forall \mu \in (0, [\lambda]_0^1), \quad \frac{\partial \tilde{d}_\epsilon^-}{\partial \mu}(\lambda, \mu, r) = \frac{\mu - g_\epsilon^-(\mu) + x_\epsilon^-(\lambda, \mu, r)}{\mu(1 - \mu)}.$$

1745 For all  $(\lambda, \mu) \in \mathbb{R} \times (0, 1)$  such that  $\mu \in (0, [\lambda]_0^1]$ , the function  $r \mapsto \tilde{d}_\epsilon^-(\lambda, \mu, r)$  is the zero function.  
 1746 For all  $(\lambda, \mu) \in \mathbb{R} \times (0, 1)$  such that  $\mu \in ([\lambda]_0^1, 1)$ , the function  $r \mapsto \tilde{d}_\epsilon^-(\lambda, \mu, r)$  is positive,  
 1747 continuously differentiable and increasing on  $\mathbb{R}_+$ .

1748 *Proof.* Using Lemmas 26 and 23, we have

$$\begin{aligned} & \tilde{d}_\epsilon^-(\lambda, \mu, r) = \tilde{d}_\epsilon^+(1 - \lambda, 1 - \mu, r) \quad \text{and} \quad g_\epsilon^+(\lambda) = 1 - g_\epsilon^-(1 - \lambda), \\ & \frac{\partial \tilde{d}_\epsilon^-}{\partial \mu}(\lambda, \mu, r) = -\frac{\partial \tilde{d}_\epsilon^+}{\partial \mu}(1 - \lambda, 1 - \mu, r) \quad \text{and} \quad \frac{\partial^2 \tilde{d}_\epsilon^-}{\partial \mu^2}(\lambda, \mu, r) = \frac{\partial^2 \tilde{d}_\epsilon^+}{\partial \mu^2}(1 - \lambda, 1 - \mu, r). \end{aligned}$$

1749 Let  $x_\epsilon^+(1 - \lambda, 1 - \mu, r) \in (\max\{0, g_\epsilon^-(\mu) - \lambda\}, g_\epsilon^-(\mu) - \mu)$  be the unique solution for  $x \in$   
 1750  $(\max\{0, g_\epsilon^-(\mu) - \lambda\}, g_\epsilon^-(\mu) - \mu)$  of the equation

$$\log \left( 1 + \frac{x}{(1 - g_\epsilon^-(\mu))(g_\epsilon^-(\mu) - x)} \right) + \epsilon \left( \frac{r\epsilon(x - g_\epsilon^-(\mu) + \lambda)}{\sqrt{(r\epsilon(x - g_\epsilon^-(\mu) + \lambda))^2 + 1 + 1}} - 1 \right) = 0,$$

1751 where we used  $g_\epsilon^+(1 - \mu) = 1 - g_\epsilon^-(\mu)$  to simplify the formula given in Lemma 30. Therefore, we  
 1752 define  $x_\epsilon^-(\lambda, \mu, r) = x_\epsilon^+(1 - \lambda, 1 - \mu, r)$ . Then, we have

$$\begin{aligned} & \text{kl}(g_\epsilon^-(\mu) - x_\epsilon^-(\lambda, \mu, r), \mu) + \frac{1}{r}h(r\epsilon(x_\epsilon^-(\lambda, \mu, r) + \lambda - g_\epsilon^-(\mu))) = \\ & \text{kl}(x_\epsilon^+(1 - \lambda, 1 - \mu, r) + g_\epsilon^+(1 - \mu), 1 - \mu) + \frac{h(r\epsilon(x_\epsilon^+(1 - \lambda, 1 - \mu, r) + g_\epsilon^+(1 - \mu) - 1 + \lambda))}{r} \end{aligned}$$

1753 where we used that  $\text{kl}(g_\epsilon^-(\mu) - x_\epsilon^-(\lambda, \mu, r), \mu) = \text{kl}(1 - g_\epsilon^-(\mu) + x_\epsilon^-(\lambda, \mu, r), 1 - \mu)$ . Combining  
 1754 the above with the properties on  $\tilde{d}_\epsilon^+$  in Lemma 30 concludes the proof.  $\square$

1755 Lemma 32 shows that we can invert  $\tilde{d}_\epsilon^\pm$  with respect to their first argument, which is a key property  
 1756 used in Appendix F.

1757 **Lemma 32.** Let  $\tilde{d}_\epsilon^\pm$  as in Eq. (32). For all  $(\mu, r, c) \in (0, 1) \times \mathbb{R}_+^* \times \mathbb{R}_+^*$ , there exists  $x > 0$  such  
 1758 that  $\tilde{d}_\epsilon^+(\mu - x, \mu, r) = c$ . For all  $(\mu, r, c) \in (0, 1) \times \mathbb{R}_+^* \times \mathbb{R}_+^*$ , there exists  $x > 0$  such that  
 1759  $\tilde{d}_\epsilon^-(\mu + x, \mu, r) = c$ .

1760 *Proof.* Let us define  $f(x) = \tilde{d}_\epsilon^+(\mu - x, \mu, r)$  for all  $x > 0$ . Using Lemma 30, we know that  $f$  is  
 1761 continuous and increasing on  $\mathbb{R}_+^*$  and it satisfies  $\lim_{x \rightarrow 0^+} f(x) = 0$  and  $\lim_{x \rightarrow +\infty} f(x) = +\infty$ .  
 1762 Therefore, there exists a unique  $x > 0$  such that  $\tilde{d}_\epsilon^+(\mu - x, \mu, r) = c$ . Using Lemma 31, we can  
 1763 conclude similarly for  $\tilde{d}_\epsilon^-$ .  $\square$

1764 Lemma 32 shows that  $\tilde{d}_\epsilon^\pm$  is non-decreasing with respect to their first argument, which is a key  
 1765 property used in Appendix F.

1766 **Lemma 33.** *Let  $\tilde{d}_\epsilon^\pm$  as in Eq. (32). For all  $(\mu, r) \in (0, 1) \times \mathbb{R}_+^*$  and all  $(\lambda_1, \lambda_2) \in \mathbb{R} \times (-\infty, \mu)$ ,*

$$\tilde{d}_\epsilon^+(\lambda_1, \mu, r) \geq \tilde{d}_\epsilon^+(\lambda_2, \mu, r) \implies \lambda_1 \leq \lambda_2$$

1767 *For all  $(\mu, r) \in (0, 1) \times \mathbb{R}_+^*$  and all  $(\lambda_1, \lambda_2) \in \mathbb{R} \times (\mu, +\infty)$ ,*

$$\tilde{d}_\epsilon^-(\lambda_1, \mu, r) \geq \tilde{d}_\epsilon^-(\lambda_2, \mu, r) \implies \lambda_1 \geq \lambda_2$$

1768 *Proof.* Using Lemma 30, we know that  $\lambda \mapsto \tilde{d}_\epsilon^+(\lambda, \mu, r)$  is decreasing on  $(-\infty, \mu)$ . Let  $(\lambda_1, \lambda_2) \in$   
 1769  $\mathbb{R} \times (-\infty, \mu)$ . Then, we have

$$\lambda_1 > \lambda_2 \implies \tilde{d}_\epsilon^+(\lambda_1, \mu, r) < \tilde{d}_\epsilon^+(\lambda_2, \mu, r),$$

1770 which is equivalent to the statement of the lemma by contraposition. Using Lemma 31, we can  
 1771 conclude similarly for  $\tilde{d}_\epsilon^-$ .  $\square$

## 1772 G.2 Transportation Cost

1773 Recall that  $W_{\epsilon,a,b}$  is defined in Eq. (50), i.e., for all  $(\mu, w) \in \mathbb{R}^K \times \mathbb{R}_+^K$ ,

$$\forall (a, b) \in [K]^2, \quad W_{\epsilon,a,b}(\mu, w) := \mathbb{1}([\mu_a]_0^1 > [\mu_b]_0^1) \inf_{u \in [0,1]} \{w_a d_\epsilon^-(\mu_a, u) + w_b d_\epsilon^+(\mu_b, u)\},$$

1774 where  $d_\epsilon^\pm$  are defined in Eq. (3).

1775 Lemma 34 gathers regularity properties on the transportation costs.

1776 **Lemma 34.** *Let  $d_\epsilon^\pm$  as in Eq. (3). For all  $(\lambda, \mu) \in (0, 1)^2$  such that  $\lambda \geq \mu$  and  $w \in \mathbb{R}_+^2$ .*

1777 • *The function  $u \mapsto w_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)$  is strictly convex on  $[\mu, \lambda]$  when*  
 1778  *$\max\{w_1, w_2\} > 0$  and on  $[0, 1]$  when  $\min\{w_1, w_2\} > 0$ . Then,*

$$\inf_{u \in [0,1]} \{w_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)\} = \inf_{u \in [\mu, \lambda]} \{w_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)\}.$$

1779 • *The function  $(\lambda, \mu, w) \mapsto \inf_{u \in [0,1]} \{w_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)\}$  is continuous on  $(0, 1) \times$   
 1780  $(0, 1) \times \mathbb{R}_+^2$ .*

1781 • *If  $\max\{w_1, w_2\} > 0$ ,  $u_\star(\lambda, \mu, w) = \arg \min_{u \in [0,1]} \{w_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)\}$  is unique*  
 1782 *and continuous on  $(0, 1) \times (0, 1) \times \mathbb{R}_+^2$ .*

1783 • *If  $\min\{w_1, w_2\} > 0$  and  $\lambda > \mu$ ,  $u_\star(\lambda, \mu, w) \in (\mu, \lambda)$  and*  
 1784  *$\min\{d_\epsilon^-(\lambda, u_\star(\lambda, \mu, w)), d_\epsilon^+(\mu, u_\star(\lambda, \mu, w))\} > 0$ .*

1785 *Moreover,*

$$\inf_{u \in [0,1]} \{w_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)\} = \inf_{(u_1, u_2) \in [0,1]^2 : u_1 \leq u_2} \{w_1 d_\epsilon^-(\lambda, u_1) + w_2 d_\epsilon^+(\mu, u_2)\}.$$

1786 *Proof.* These results are obtained by leveraging Lemmas 24 and 25 at each step.

1787 For  $u \leq \mu$ , the function is equal to  $w_1 d_\epsilon^-(\lambda, u)$ , which is decreasing and strictly convex on  $[0, \lambda]$   
 1788 unless  $w_1 = 0$  since  $u \leq \mu \leq \lambda$ . Therefore, the minimum over that interval is attained at  $\mu$ . For  
 1789  $u \geq \lambda$ , the function is equal to  $w_2 d_\epsilon^+(\mu, u)$ , which is increasing and strictly convex on  $(\mu, 1]$  unless  
 1790  $w_2 = 0$  since  $u \geq \lambda \geq \mu$ . Therefore, the minimum over that interval is attained at  $\lambda$ . On the interval  
 1791  $(\mu, \lambda)$ , the function is equal to  $w_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)$ , hence it is the sum of two convex functions,



one of which is strictly convex. Furthermore, the function is continuous at  $\mu$  and  $\lambda$ . This concludes the first part of the proof.

As we have just shown, we can restrict the infimum to  $[\mu, \lambda]$ . We apply Berge's Maximum theorem [16, page 116]. Let

$$\begin{aligned}\phi(u, \lambda, \mu, w) &= -w_1 d_\epsilon^-(\lambda, u) - w_2 d_\epsilon^+(\mu, u), \\ \Gamma(\lambda, \mu, w) &= [\mu, \lambda], \\ M(\lambda, \mu, w) &= \max\{\phi(u, \lambda, \mu, w) \mid u \in \Gamma(\lambda, \mu, w)\}, \\ \Phi(\lambda, \mu, w) &= \arg \max\{\phi(u, \lambda, \mu, w) \mid u \in \Gamma(\lambda, \mu, w)\}.\end{aligned}$$

We verify the hypotheses of the theorem:

- $\phi$  is continuous on  $[\mu, \lambda] \times (0, 1) \times (0, 1) \times \mathbb{R}_+^2$ , by using the properties in Lemmas 24 and 25 since  $(\lambda, \mu) \in (0, 1)^2$ .
- $\Gamma$  is nonempty, compact-valued and continuous (since constant).

We obtain that  $M$  is continuous on  $(0, 1) \times (0, 1) \times \mathbb{R}_+^2$  and that  $\Phi$  is upper hemicontinuous. This concludes the second part of the proof.

When  $\max\{w_1, w_2\} > 0$ , we have just shown that  $\phi$  is a strictly concave function of  $u$ . Combining this with the fact that  $\Gamma$  is convex, we can argue as in [78, Theorem 9.17] to prove that  $\Phi$  is a single-valued upper hemicontinuous correspondence, hence a continuous function. This concludes the third part of the proof.

Suppose that  $\min\{w_1, w_2\} > 0$  and  $\lambda > \mu$ . Using Lemmas 24 and 25, the function  $u \mapsto w_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)$  is continuously differentiable on  $(\mu, \lambda)$  with derivative  $w_1 \frac{\partial d_\epsilon^-}{\partial u}(\lambda, u) + w_2 \frac{\partial d_\epsilon^+}{\partial u}(\mu, u)$  where

$$\begin{aligned}\forall u \in (\mu, 1], \quad \frac{\partial d_\epsilon^+}{\partial u}(\mu, u) &= \begin{cases} \frac{1-e^{-\epsilon}}{1-u(1-e^{-\epsilon})} & \text{if } u \in (g_\epsilon^-(\mu), 1] \\ \frac{u-\mu}{u(1-u)} & \text{if } u \in (\mu, g_\epsilon^-(\mu)] \end{cases}, \\ \forall u \in [0, \lambda), \quad \frac{\partial d_\epsilon^-}{\partial u}(\lambda, u) &= \begin{cases} -\frac{e^\epsilon-1}{1+u(e^\epsilon-1)} & \text{if } u \in [0, g_\epsilon^+(\lambda)) \\ -\frac{\lambda-u}{u(1-u)} & \text{if } u \in [g_\epsilon^+(\lambda), \lambda) \end{cases}.\end{aligned}$$

Since  $\frac{\partial d_\epsilon^-}{\partial u}(\lambda, u) \rightarrow_{u \rightarrow \lambda^-} 0$  and  $\frac{\partial d_\epsilon^+}{\partial u}(\mu, u) \rightarrow_{u \rightarrow \mu^+} 0$ , we obtain

$$\begin{aligned}\lim_{u \rightarrow \lambda^-} \left\{ w_1 \frac{\partial d_\epsilon^-}{\partial u}(\lambda, u) + w_2 \frac{\partial d_\epsilon^+}{\partial u}(\mu, u) \right\} &= w_2 \frac{\partial d_\epsilon^+}{\partial u}(\mu, \lambda) > 0, \\ \lim_{u \rightarrow \mu^+} \left\{ w_1 \frac{\partial d_\epsilon^-}{\partial u}(\lambda, u) + w_2 \frac{\partial d_\epsilon^+}{\partial u}(\mu, u) \right\} &= w_1 \frac{\partial d_\epsilon^-}{\partial u}(\lambda, \mu) < 0.\end{aligned}$$

Therefore, the infimum is attained inside the open interval. Using Lemmas 24 and 25, we can conclude the proof of the first part of the fourth property.

Using the strict convexity of  $u_1 \mapsto w_1 d_\epsilon^-(\lambda, u_1)$  and  $u_2 \mapsto w_2 d_\epsilon^+(\mu, u_2)$  on  $(\mu, \lambda)$ , we obtain that

$$\inf_{u \in (\mu, \lambda)} \{w_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)\} = \inf_{(u_1, u_2) : \mu < u_1 \leq u_2 < \lambda} \{w_1 d_\epsilon^-(\lambda, u_1) + w_2 d_\epsilon^+(\mu, u_2)\}.$$

Re-using the same arguments as above, we obtain that

$$\begin{aligned}& \inf_{(u_1, u_2) : \mu < u_1 \leq u_2 < \lambda} \{w_1 d_\epsilon^-(\lambda, u_1) + w_2 d_\epsilon^+(\mu, u_2)\} \\ &= \inf_{(u_1, u_2) \in [0, 1]^2 : u_1 \leq u_2} \{w_1 d_\epsilon^-(\lambda, u_1) + w_2 d_\epsilon^+(\mu, u_2)\}.\end{aligned}$$

This concludes the proof of the second part of the fourth property.  $\square$

Lemma 35 relates the transportation costs  $W_{\epsilon, a^*, a}$  with the transportation costs used in Eq. (1) to define the characteristic time. Crucially, this shows the equivalence with the definitions in Eq. (35).

1817 **Lemma 35.** Let  $W_{\epsilon, a^*, a}$  and  $d_\epsilon$  as in Eq. (4) and (2). Let  $\mu \in (0, 1)^K$  such that  $a^*(\mu) = \{a^*\}$ . Let  
 1818  $\text{Alt}(\mu) = \{\lambda \in (0, 1)^K \mid a^*(\lambda) \neq \{a^*\}\}$ . Then,

$$\forall w \in \Delta_K, \quad \inf_{\lambda \in \text{Alt}(\mu)} \sum_{a \in [K]} w_a d_\epsilon(\mu_a, \lambda_a) = \min_{a \neq a^*} W_{\epsilon, a^*, a}(\mu, w).$$

1819 *Proof.* It is direct to see that  $\text{Alt}(\mu) = \bigcup_{a \neq a^*} \mathcal{C}_a$  where  $\mathcal{C}_a = \{\lambda \in (0, 1)^K \mid \lambda_a \geq \lambda_{a^*}\}$ . Then,

$$\forall w \in \Delta_K, \quad \inf_{\lambda \in \text{Alt}(\mu)} \sum_{a \in [K]} w_a d_\epsilon(\mu_a, \lambda_a) = \min_{a \neq a^*} \inf_{\lambda \in \mathcal{C}_a} \sum_{c \in [K]} w_c d_\epsilon(\mu_c, \lambda_c).$$

1820 By non-negativity of  $d_\epsilon(\mu_a, \lambda_a)$  for all  $a \in [K]$ , we obtain

$$\begin{aligned} \inf_{\lambda \in \mathcal{C}_a} \sum_{c \in [K]} w_c d_\epsilon(\mu_c, \lambda_c) &= \inf_{\lambda \in \mathcal{C}_a} \sum_{c \in \{a, a^*\}} w_c d_\epsilon(\mu_c, \lambda_c) \\ &= \inf_{(\lambda_a, \lambda_{a^*}) \in (0, 1)^2, \lambda_a \geq \lambda_{a^*}} \sum_{c \in \{a, a^*\}} w_c d_\epsilon(\mu_c, \lambda_c), \end{aligned}$$

1821 where the two equalities are obtained by choosing  $\lambda(a) \in (0, 1)^K$  such that  $\lambda(a)_b = \mu_b$  for all  
 1822  $b \notin \{a, a^*\}$  with the two other coordinates choosen freely such that  $\lambda(a)_a \geq \lambda(a)_{a^*}$ . Using that  
 1823  $\mu_{a^*} > \mu_a$ , we can partition this set as follows

$$\begin{aligned} \mathcal{C}_{a, a^*} &= \{(\lambda_a, \lambda_{a^*}) \in (0, 1)^2 \mid \lambda_a \geq \lambda_{a^*}\} = \{(\lambda_a, \lambda_{a^*}) \in (0, \mu_a)^2 \mid \lambda_a \geq \lambda_{a^*}\} \\ &\quad \cup \{(\lambda_a, \lambda_{a^*}) \in [\mu_a, \mu_{a^*}] \times (0, \mu_a)\} \\ &\quad \cup \{(\lambda_a, \lambda_{a^*}) \in (\mu_{a^*}, 1)^2 \mid \lambda_a \geq \lambda_{a^*}\} \\ &\quad \cup \{(\lambda_a, \lambda_{a^*}) \in (\mu_{a^*}, 1) \times [\mu_a, \mu_{a^*}]\} \\ &\quad \cup \{(\lambda_a, \lambda_{a^*}) \in [\mu_a, \mu_{a^*}]^2 \mid \lambda_a \geq \lambda_{a^*}\}. \end{aligned}$$

1824 Using Lemma 21,  $\mu_{a^*} > \mu_a$  and Lemmas 24 and 25, we obtain

$$\begin{aligned} &\inf_{(\lambda_a, \lambda_{a^*}) \in (0, \mu_a)^2 \mid \lambda_a \geq \lambda_{a^*}} \{w_{a^*} d_\epsilon(\mu_{a^*}, \lambda_{a^*}) + w_a d_\epsilon(\mu_a, \lambda_a)\} \\ &= \inf_{(\lambda_a, \lambda_{a^*}) \in (0, \mu_a)^2 \mid \lambda_a \geq \lambda_{a^*}} \{w_{a^*} d_\epsilon^-(\mu_{a^*}, \lambda_{a^*}) + w_a d_\epsilon^-(\mu_a, \lambda_a)\} = w_{a^*} d_\epsilon^-(\mu_{a^*}, \mu_a), \\ &\inf_{(\lambda_a, \lambda_{a^*}) \in [\mu_a, \mu_{a^*}] \times (0, \mu_a)} \{w_{a^*} d_\epsilon(\mu_{a^*}, \lambda_{a^*}) + w_a d_\epsilon(\mu_a, \lambda_a)\} \\ &= \inf_{(\lambda_a, \lambda_{a^*}) \in [\mu_a, \mu_{a^*}] \times (0, \mu_a)} \{w_{a^*} d_\epsilon^-(\mu_{a^*}, \lambda_{a^*}) + w_a d_\epsilon^+(\mu_a, \lambda_a)\} = w_{a^*} d_\epsilon^-(\mu_{a^*}, \mu_a), \\ &\inf_{(\lambda_a, \lambda_{a^*}) \in (\mu_{a^*}, 1)^2 \mid \lambda_a \geq \lambda_{a^*}} \{w_{a^*} d_\epsilon(\mu_{a^*}, \lambda_{a^*}) + w_a d_\epsilon(\mu_a, \lambda_a)\} \\ &= \inf_{(\lambda_a, \lambda_{a^*}) \in (\mu_{a^*}, 1)^2 \mid \lambda_a \geq \lambda_{a^*}} \{w_{a^*} d_\epsilon^+(\mu_{a^*}, \lambda_{a^*}) + w_a d_\epsilon^+(\mu_a, \lambda_a)\} = w_a d_\epsilon^+(\mu_a, \mu_{a^*}), \\ &\inf_{(\lambda_a, \lambda_{a^*}) \in (\mu_{a^*}, 1) \times [\mu_a, \mu_{a^*}]} \{w_{a^*} d_\epsilon(\mu_{a^*}, \lambda_{a^*}) + w_a d_\epsilon(\mu_a, \lambda_a)\} \\ &= \inf_{(\lambda_a, \lambda_{a^*}) \in (\mu_{a^*}, 1) \times [\mu_a, \mu_{a^*}]} \{w_{a^*} d_\epsilon^-(\mu_{a^*}, \lambda_{a^*}) + w_a d_\epsilon^+(\mu_a, \lambda_a)\} = w_a d_\epsilon^+(\mu_a, \mu_{a^*}), \\ &\inf_{(\lambda_a, \lambda_{a^*}) \in [\mu_a, \mu_{a^*}]^2 \mid \lambda_a \geq \lambda_{a^*}} \{w_{a^*} d_\epsilon(\mu_{a^*}, \lambda_{a^*}) + w_a d_\epsilon(\mu_a, \lambda_a)\} \\ &= \inf_{(\lambda_a, \lambda_{a^*}) \in [\mu_a, \mu_{a^*}]^2 \mid \lambda_a \geq \lambda_{a^*}} \{w_{a^*} d_\epsilon^-(\mu_{a^*}, \lambda_{a^*}) + w_a d_\epsilon^+(\mu_a, \lambda_a)\}. \end{aligned}$$

1825 Therefore, we obtain

$$\begin{aligned} &\inf_{(\lambda_a, \lambda_{a^*}) \in \mathcal{C}_{a, a^*}} \{w_{a^*} d_\epsilon(\mu_{a^*}, \lambda_{a^*}) + w_a d_\epsilon(\mu_a, \lambda_a)\} \\ &= \inf_{(\lambda_a, \lambda_{a^*}) \in [\mu_a, \mu_{a^*}]^2 \mid \lambda_a \geq \lambda_{a^*}} \{w_{a^*} d_\epsilon^-(\mu_{a^*}, \lambda_{a^*}) + w_a d_\epsilon^+(\mu_a, \lambda_a)\} \\ &= \inf_{u \in [\mu_a, \mu_{a^*}]^2} \{w_{a^*} d_\epsilon^-(\mu_{a^*}, u) + w_a d_\epsilon^+(\mu_a, u)\} \\ &= \inf_{u \in [0, 1]} \{w_{a^*} d_\epsilon^-(\mu_{a^*}, u) + w_a d_\epsilon^+(\mu_a, u)\} = W_{\epsilon, a^*, a}(\mu, w), \end{aligned}$$

1826 where the second equality is obtained similarly as in Lemma 34 by leveraging the strict convexity  
 1827 of  $d_\epsilon^\pm$  in their second argument (see Lemmas 24 and 25). We used Lemma 34 and the definition of  
 1828  $W_{\epsilon,a^*,a}(\mu, w)$  for the last two equalities. This concludes the proof.  $\square$

1829 Lemma 36 gathers additional properties on the transportation costs.

1830 **Lemma 36.** *Let  $d_\epsilon^\pm$  as in Eq. (3).*

1831 • *Let  $(\lambda, \mu) \in (0, 1)^2$  such that  $\lambda > \mu$ . When  $w_2 > 0$ , the function  $w_1 \mapsto$   
 1832  $\min_{u \in [0,1]} \{w_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)\}$  is increasing on  $\mathbb{R}_+$ . When  $w_1 > 0$ , the function  
 1833  $w_2 \mapsto \min_{u \in [0,1]} \{w_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)\}$  is increasing on  $\mathbb{R}_+$ .*

1834 • *Let  $(\lambda, \mu) \in (0, 1)^2$  and  $\mu \in (0, 1)^K$ . The function  $w \mapsto$   
 1835  $\min_{u \in [0,1]} \{w_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)\}$  is concave on  $\mathbb{R}_+^2$ . The function  
 1836  $w \mapsto \min_{a \in [K] \setminus \{1\}} \min_{u \in [0,1]} \{w_1 d_\epsilon^-(\mu_1, u) + w_a d_\epsilon^+(\mu_a, u)\}$  is concave on  $\mathbb{R}_+^K$ .*

1837 *Proof.* Let  $w_2 > 0$  and  $w'_1 > w_1 \geq 0$ . Using Lemma 34, since  $w'_1 > 0$ , there exists  $u' \in [0, 1]$  with  
 1838  $d_\epsilon^-(\lambda, u') > 0$  such that

$$\begin{aligned} \min_{u \in [0,1]} \{w'_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)\} &= w'_1 d_\epsilon^-(\lambda, u') + w_2 d_\epsilon^+(\mu, u') \\ &> w_1 d_\epsilon^-(\lambda, u') + w_2 d_\epsilon^+(\mu, u') \\ &\geq \min_{u \in [0,1]} \{w_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)\}. \end{aligned}$$

1839 Let  $w_1 > 0$  and  $w'_2 > w_2 \geq 0$ . Then, we can show similarly by using Lemma 34 that

$$\min_{u \in [0,1]} \{w_1 d_\epsilon^-(\lambda, u) + w'_2 d_\epsilon^+(\mu, u)\} > \min_{u \in [0,1]} \{w_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)\}.$$

1840 This concludes the first part of the proof. The proof of the second part is direct since those functions  
 1841 are minimum of linear functions, hence concave.  $\square$

1842 Lemma 37 gives a closed-form solution for the transportation costs. This is a key property used in  
 1843 our implementation to reduce the computational cost.

1844 **Lemma 37.** *Let  $d_\epsilon^\pm$  and  $g_\epsilon^\pm$  as in Eq. (3) and (30). For all  $(a, c) \in \mathbb{R}_+^2$  and  $b \in \mathbb{R}$ , let  $r_{1,+}(a, b, c) :=$   
 1845  $\frac{\sqrt{b^2 + 4ac} - b}{2a}$ . For all  $(\lambda, \mu) \in (0, 1)^2$  and  $w \in \mathbb{R}_+^2$  such that  $\min\{w_1, w_2\} > 0$  and  $\lambda > \mu$ .*

1846 • *When (1)  $g_\epsilon^-(\mu) \geq \lambda$ , or (2)  $g_\epsilon^-(\mu) < \lambda$ ,  $g_\epsilon^+(\lambda) \leq g_\epsilon^-(\mu)$  and  $\frac{w_2 \mu + w_1 \lambda}{w_2 + w_1} \in [g_\epsilon^+(\lambda), g_\epsilon^-(\mu)]$ , we  
 1847 have  $u_\star(\lambda, \mu, w) = \frac{w_2 \mu + w_1 \lambda}{w_2 + w_1}$  and*

$$\min_{u \in [0,1]} \{w_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)\} = w_1 \text{kl}(\lambda, u_\star(\lambda, \mu, w)) + w_2 \text{kl}(\mu, u_\star(\lambda, \mu, w)).$$

1848 • *When (3)  $g_\epsilon^-(\mu) < \lambda$ ,  $g_\epsilon^+(\lambda) > g_\epsilon^-(\mu)$  and  $u_{3,\star}(w) \in [g_\epsilon^-(\mu), g_\epsilon^+(\lambda)]$  where*

$$u_{3,\star}(w) := \frac{w_1(e^\epsilon - 1) - w_2(1 - e^{-\epsilon})}{(w_2 + w_1)(1 - e^{-\epsilon})(e^\epsilon - 1)},$$

1849 *we have  $u_\star(\lambda, \mu, w) = u_{3,\star}(w)$  and*

$$\begin{aligned} &\min_{u \in [0,1]} \{w_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)\} \\ &= w_1 (-\log(1 + u_{3,\star}(w)(e^\epsilon - 1)) + \epsilon \lambda) + w_2 (-\log(1 - u_{3,\star}(w)(1 - e^{-\epsilon})) - \epsilon \mu). \end{aligned}$$

1850 • *When (4)  $g_\epsilon^-(\mu) < \lambda$ ,  $g_\epsilon^+(\lambda) \leq g_\epsilon^-(\mu)$  and  $\frac{w_2 \mu + w_1 \lambda}{w_2 + w_1} \in (\mu, g_\epsilon^+(\lambda))$ , or (5)  $g_\epsilon^-(\mu) < \lambda$ ,  $g_\epsilon^+(\lambda) >$   
 1851  $g_\epsilon^-(\mu)$  and  $u_{3,\star}(w) < g_\epsilon^-(\mu)$ , we have  $u_\star(\lambda, \mu, w) = u_{1,\star}(\mu, w)$  and*

$$\begin{aligned} &u_{1,\star}(\mu, w) := r_{1,+}((w_2 + w_1)(e^\epsilon - 1), (w_2 - (w_2 \mu + w_1)(e^\epsilon - 1)), w_2 \mu), \\ &\min_{u \in [0,1]} \{w_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)\} \end{aligned}$$

$$= w_1 (-\log(1 + u_{1,*}(\mu, w)(e^\epsilon - 1)) + \epsilon\lambda) + w_2 \text{kl}(\mu, u_{1,*}(\mu, w)) .$$

1852 • When (6)  $g_\epsilon^-(\mu) < \lambda$ ,  $g_\epsilon^+(\lambda) \leq g_\epsilon^-(\mu)$  and  $\frac{w_2\mu + w_1\lambda}{w_2 + w_1} \in (g_\epsilon^-(\mu), \lambda)$ , or (7)  $g_\epsilon^-(\mu) < \lambda$ ,  $g_\epsilon^+(\lambda) >$   
 1853  $g_\epsilon^-(\mu)$  and  $u_{3,*}(w) > g_\epsilon^+(\lambda)$ , we have  $u_*(\lambda, \mu, w) = u_{2,*}(\lambda, w)$  and

$$\begin{aligned} u_{2,*}(\lambda, w) &:= 1 - r_{1,+}((w_2 + w_1)(e^\epsilon - 1), (w_1 - (w_1(1 - \lambda) + w_2)(e^\epsilon - 1)), w_1(1 - \lambda)) , \\ &\min_{u \in [0,1]} \{w_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)\} \\ &= w_1 \text{kl}(\lambda, u_{2,*}(\lambda, w)) + w_2 (-\log(1 - u_{2,*}(\lambda, w)(1 - e^{-\epsilon})) - \epsilon\mu) . \end{aligned}$$

1854 *Proof.* Suppose that  $g_\epsilon^-(\mu) \geq \lambda$ . Using Lemma 23, we know that  $g_\epsilon^-(\mu) \geq \lambda$  if and only if  
 1855  $\mu \geq g_\epsilon^+(\lambda)$ . Therefore, for all  $u \in (\mu, \lambda)$ , we have

$$w_1 \frac{\partial d_\epsilon^-}{\partial u}(\lambda, u) + w_2 \frac{\partial d_\epsilon^+}{\partial u}(\mu, u) = -w_1 \frac{\lambda - u}{u(1 - u)} + w_2 \frac{u - \mu}{u(1 - u)} = \frac{(w_2 + w_1)u - (w_2\mu + w_1\lambda)}{u(1 - u)} .$$

1856 Therefore, we have

$$u_*(\lambda, \mu, w) = \frac{w_2\mu + w_1\lambda}{w_2 + w_1} \in (\mu, \lambda) .$$

1857 Suppose that  $g_\epsilon^-(\mu) < \lambda$ . Using Lemma 23, we know that  $g_\epsilon^-(\mu) < \lambda$  if and only if  $\mu < g_\epsilon^+(\lambda)$ .  
 1858 Using strict convexity of the function on  $(\mu, \lambda)$ , it is enough to exhibit one local minimum to obtain a  
 1859 global minimum on  $(\mu, \lambda)$ .

1860 Suppose that  $g_\epsilon^+(\lambda) \leq g_\epsilon^-(\mu)$ . Similarly as above, we obtain, for all  $u \in [g_\epsilon^+(\lambda), g_\epsilon^-(\mu)]$ ,

$$w_1 \frac{\partial d_\epsilon^-}{\partial u}(\lambda, u) + w_2 \frac{\partial d_\epsilon^+}{\partial u}(\mu, u) = \frac{(w_2 + w_1)u - (w_2\mu + w_1\lambda)}{u(1 - u)} .$$

1861 Suppose that  $\frac{w_2\mu + w_1\lambda}{w_2 + w_1} \in [g_\epsilon^+(\lambda), g_\epsilon^-(\mu)]$ . Then, we can conclude as above that

$$u_*(\lambda, \mu, w) = \frac{w_2\mu + w_1\lambda}{w_2 + w_1} \in [g_\epsilon^+(\lambda), g_\epsilon^-(\mu)] ,$$

1862 since it is a local minimum of a strictly convex function.

1863 Suppose that  $\frac{w_2\mu + w_1\lambda}{w_2 + w_1} < g_\epsilon^+(\lambda)$ . Since the gradient is positive on  $[g_\epsilon^+(\lambda), g_\epsilon^-(\mu)]$ , we know that  
 1864 the minimum on  $(\mu, \lambda)$  is achieved on  $(\mu, g_\epsilon^+(\lambda))$ , i.e.,  $u_*(\lambda, \mu, w) \in (\mu, g_\epsilon^+(\lambda))$ . Then, for all  
 1865  $u \in (\mu, g_\epsilon^+(\lambda))$ ,

$$w_1 \frac{\partial d_\epsilon^-}{\partial u}(\lambda, u) + w_2 \frac{\partial d_\epsilon^+}{\partial u}(\mu, u) = -w_1 \frac{e^\epsilon - 1}{1 + u(e^\epsilon - 1)} + w_2 \frac{u - \mu}{u(1 - u)} .$$

1866 Using Lemma 23, direct computation yields

$$\begin{aligned} w_1 \frac{\partial d_\epsilon^-}{\partial u}(\lambda, u) + w_2 \frac{\partial d_\epsilon^+}{\partial u}(\mu, u) > 0 &\iff \mu < u \left( 1 + \frac{w_1}{w_2} \left( 1 - \frac{g_\epsilon^-(u)}{u} \right) \right) , \\ \lim_{u \rightarrow \mu^+} u \left( 1 + \frac{w_1}{w_2} \left( 1 - \frac{g_\epsilon^-(u)}{u} \right) \right) &= \mu \left( 1 + \frac{w_1}{w_2} \left( 1 - \frac{g_\epsilon^-(\mu)}{\mu} \right) \right) < \mu , \\ \lim_{u \rightarrow g_\epsilon^+(\lambda)^-} u \left( 1 + \frac{w_1}{w_2} \left( 1 - \frac{g_\epsilon^-(u)}{u} \right) \right) &= g_\epsilon^+(\lambda) \left( 1 + \frac{w_1}{w_2} \left( 1 - \frac{\lambda}{g_\epsilon^+(\lambda)} \right) \right) > \mu , \end{aligned}$$

1867 where the second result uses that  $u < g_\epsilon^-(u)$  and the last result is obtained by continuity of the  
 1868 differentials (Lemmas 24 and 25) and the positivity on  $[g_\epsilon^+(\lambda), g_\epsilon^-(\mu)]$ . For all  $(a, c) \in \mathbb{R}_+^2$  and  
 1869  $b \in \mathbb{R}$ , we define  $r_{1,+}(a, b, c) = \frac{\sqrt{b^2 + 4ac} - b}{2a}$ . Therefore, we have

$$\begin{aligned} w_1 \frac{\partial d_\epsilon^-}{\partial u}(\lambda, u) + w_2 \frac{\partial d_\epsilon^+}{\partial u}(\mu, u) &= 0 \\ \iff (w_2 + w_1)(e^\epsilon - 1)u^2 + (w_2 - (w_2\mu + w_1)(e^\epsilon - 1))u - w_2\mu &= 0 \\ \iff u_*(\lambda, \mu, w) = r_{1,+}((w_2 + w_1)(e^\epsilon - 1), (w_2 - (w_2\mu + w_1)(e^\epsilon - 1)), w_2\mu) &\in (\mu, g_\epsilon^+(\lambda)) \end{aligned}$$

1870 where we used that  $u_*(\lambda, \mu, w) \in (\mu, g_\epsilon^+(\lambda))$  is unique for the last equivalence, and that the second  
1871 root of the second order polynomial equation is negative. Notice that  $u_*(\lambda, \mu, w)$  is independent of  $\lambda$ .

1872 Suppose that  $\frac{w_2\mu + w_1\lambda}{w_2 + w_1} > g_\epsilon^-(\mu)$ . Since the gradient is negative on  $[g_\epsilon^+(\lambda), g_\epsilon^-(\mu)]$ , we know that  
1873 the minimum on  $(\mu, \lambda)$  is achieved on  $(g_\epsilon^-(\mu), \lambda)$ , i.e.,  $u_*(\lambda, \mu, w) \in (g_\epsilon^-(\mu), \lambda)$ . Then, for all  
1874  $u \in (g_\epsilon^-(\mu), \lambda)$ ,

$$w_1 \frac{\partial d_\epsilon^-}{\partial u}(\lambda, u) + w_2 \frac{\partial d_\epsilon^+}{\partial u}(\mu, u) = -w_1 \frac{\lambda - u}{u(1 - u)} + w_2 \frac{1 - e^{-\epsilon}}{1 - u(1 - e^{-\epsilon})}.$$

1875 Using Lemma 23, direct computation yields

$$\begin{aligned} w_1 \frac{\partial d_\epsilon^-}{\partial u}(\lambda, u) + w_2 \frac{\partial d_\epsilon^+}{\partial u}(\mu, u) < 0 &\iff \lambda > u \left( 1 + \frac{w_2}{w_1} \left( 1 - \frac{g_\epsilon^+(u)}{u} \right) \right), \\ \lim_{u \rightarrow \lambda^-} u \left( 1 + \frac{w_2}{w_1} \left( 1 - \frac{g_\epsilon^+(u)}{u} \right) \right) &= \lambda \left( 1 + \frac{w_2}{w_1} \left( 1 - \frac{g_\epsilon^+(\lambda)}{\lambda} \right) \right) > \lambda, \\ \lim_{u \rightarrow g_\epsilon^-(\mu)^+} u \left( 1 + \frac{w_2}{w_1} \left( 1 - \frac{g_\epsilon^+(u)}{u} \right) \right) &= g_\epsilon^-(\mu) \left( 1 + \frac{w_2}{w_1} \left( 1 - \frac{\mu}{g_\epsilon^-(\mu)} \right) \right) < \lambda, \end{aligned}$$

1876 where the second result uses that  $u > g_\epsilon^+(u)$  and the last result is obtained by continuity of the  
1877 differentials (Lemmas 24 and 25) and the negativity on  $[g_\epsilon^+(\lambda), g_\epsilon^-(\mu)]$ .

1878 Using Lemma 22, we obtain

$$\arg \min_{u \in [0,1]} \{w_1 d_\epsilon^-(\lambda, u) + w_2 d_\epsilon^+(\mu, u)\} = 1 - \arg \min_{u \in [0,1]} \{w_1 d_\epsilon^+(1 - \lambda, u) + w_2 d_\epsilon^-(1 - \mu, u)\}.$$

1879 Using Lemma 23, we obtain

$$\begin{aligned} g_\epsilon^-(\mu) < \lambda &\iff g_\epsilon^-(1 - \lambda) < 1 - \mu, \\ \frac{w_2\mu + w_1\lambda}{w_2 + w_1} \in (g_\epsilon^-(\mu), \lambda) &\iff \frac{w_2(1 - \mu) + w_1(1 - \lambda)}{w_2 + w_1} \in (1 - \lambda, g_\epsilon^+(1 - \mu)). \end{aligned}$$

1880 Therefore, we can leverage the above case to obtain  $u_*(\lambda, \mu, w) = u_{2,*}(\lambda, w)$  where

$$u_{2,*}(\lambda, w) = 1 - r_{1,+}((w_2 + w_1)(e^\epsilon - 1), (w_1 - (w_1(1 - \lambda) + w_2)(e^\epsilon - 1)), w_1(1 - \lambda))$$

1881 Notice that  $u_*(\lambda, \mu, w)$  is independent of  $\mu$ .

1882 Suppose that  $g_\epsilon^-(\mu) < \lambda$  and  $g_\epsilon^+(\lambda) > g_\epsilon^-(\mu)$ . Similarly as above, we obtain, for all  $u \in$   
1883  $[g_\epsilon^-(\mu), g_\epsilon^+(\lambda)]$ ,

$$w_1 \frac{\partial d_\epsilon^-}{\partial u}(\lambda, u) + w_2 \frac{\partial d_\epsilon^+}{\partial u}(\mu, u) = -w_1 \frac{e^\epsilon - 1}{1 + u(e^\epsilon - 1)} + w_2 \frac{1 - e^{-\epsilon}}{1 - u(1 - e^{-\epsilon})}.$$

1884 Therefore, we obtain

$$\begin{aligned} w_1 \frac{\partial d_\epsilon^-}{\partial u}(\lambda, u) + w_2 \frac{\partial d_\epsilon^+}{\partial u}(\mu, u) &> 0 \\ \iff w_2(1 - e^{-\epsilon})(1 + u(e^\epsilon - 1)) - w_1(e^\epsilon - 1)(1 - u(1 - e^{-\epsilon})) &> 0 \\ \iff u > u_{3,*}(w) := \frac{w_1(e^\epsilon - 1) - w_2(1 - e^{-\epsilon})}{(w_2 + w_1)(1 - e^{-\epsilon})(e^\epsilon - 1)}. \end{aligned}$$

1885 Suppose that  $u_{3,*}(w) \in [g_\epsilon^-(\mu), g_\epsilon^+(\lambda)]$ . Then, we can conclude as above that

$$u_*(\lambda, \mu, w) = u_{3,*}(w) \in [g_\epsilon^-(\mu), g_\epsilon^+(\lambda)],$$

1886 since it is a local minimum of a strictly convex function. Notice that  $u_{3,*}(w)$  is independent of  $(\lambda, \mu)$ .

1887 Suppose that  $u_{3,*}(w) > g_\epsilon^+(\lambda)$ . Since the gradient is negative on  $[g_\epsilon^-(\mu), g_\epsilon^+(\lambda)]$ , we know that  
1888 the minimum on  $(\mu, \lambda)$  is achieved on  $(g_\epsilon^+(\lambda), \lambda)$ , i.e.,  $u_*(\lambda, \mu, w) \in (g_\epsilon^+(\lambda), \lambda)$ . Then, for all  
1889  $u \in (g_\epsilon^+(\lambda), \lambda)$ ,

$$w_1 \frac{\partial d_\epsilon^-}{\partial u}(\lambda, u) + w_2 \frac{\partial d_\epsilon^+}{\partial u}(\mu, u) = -w_1 \frac{\lambda - u}{u(1 - u)} + w_2 \frac{1 - e^{-\epsilon}}{1 - u(1 - e^{-\epsilon})}.$$

1890 This recovers the condition solved above. As we know that  $u_*(\lambda, \mu, w) \in (g_\epsilon^+(\lambda), \lambda)$ , we obtain  
 1891  $u_*(\lambda, \mu, w) = u_{2,*}(\lambda, w)$  where

$$u_{2,*}(\lambda, w) = 1 - r_{1,+}((w_2 + w_1)(e^\epsilon - 1), (w_1 - (w_1(1 - \lambda) + w_2)(e^\epsilon - 1)), w_1(1 - \lambda))$$

1892 Suppose that  $u_{3,*}(w) < g_\epsilon^-(\mu)$ . Since the gradient is positive on  $[g_\epsilon^-(\mu), g_\epsilon^+(\lambda)]$ , we know that  
 1893 the minimum on  $(\mu, \lambda)$  is achieved on  $(\mu, g_\epsilon^-(\mu))$ , i.e.,  $u_*(\lambda, \mu, w) \in (\mu, g_\epsilon^-(\mu))$ . Then, for all  
 1894  $u \in (\mu, g_\epsilon^-(\mu))$ ,

$$w_1 \frac{\partial d_\epsilon^-}{\partial u}(\lambda, u) + w_2 \frac{\partial d_\epsilon^+}{\partial u}(\mu, u) = -w_1 \frac{e^\epsilon - 1}{1 + u(e^\epsilon - 1)} + w_2 \frac{u - \mu}{u(1 - u)}.$$

1895 This recovers the condition solved above. As we know that  $u_*(\lambda, \mu, w) \in (\mu, g_\epsilon^-(\mu))$ , we obtain  
 1896  $u_*(\lambda, \mu, w) = u_{1,*}(\mu, w)$  where

$$u_{1,*}(\mu, w) = r_{1,+}((w_2 + w_1)(e^\epsilon - 1), (w_2 - (w_2\mu + w_1)(e^\epsilon - 1)), w_2\mu).$$

1897 This concludes the proof.

1898 □

### 1899 G.2.1 Modified Transportation Cost

1900 Let  $\eta > 0$  be the geometric parameter used for the geometric grid update of our private empirical  
 1901 mean estimator. Let us define

$$\forall x \geq 1, \quad r(x) := \frac{x}{1 + \log_{1+\eta} x}, \quad (33)$$

1902 which is increasing if and only if  $x > \frac{e}{1+\eta}$ . For all  $(\mu, w) \in \mathbb{R}^K \times \mathbb{R}_+^K$  and all  $(a, b) \in [K]^2$  such  
 1903 that  $a \neq b$ , we define

$$\widetilde{W}_{\epsilon,a,b}(\mu, w) := \mathbb{1}([\mu_a]_0^1 > [\mu_b]_0^1) \inf_{u \in (0,1)} \left\{ w_a \widetilde{d}_\epsilon^-(\mu_a, u, r(w_a)) + w_b \widetilde{d}_\epsilon^+(\mu_b, u, r(w_b)) \right\}, \quad (34)$$

1904 where  $\widetilde{d}_\epsilon^\pm$  are defined in Eq. (32).

1905 Lemma 38 gathers regularity properties of the function  $r$  defined in Eq. (33).

1906 **Lemma 38.** *Let  $r$  as in Eq. (33). Then,*

$$\begin{aligned} \forall x \geq 1, \quad r'(x) &= \frac{\log(x(1+\eta)/e)}{\log(1+\eta)(1 + \log_{1+\eta} x)^2}, \\ r''(x) &= -\frac{1}{x(\log(1+\eta))^2} \frac{\log((1+\eta)xe^{-2})}{(1 + \log_{1+\eta} x)^3}. \end{aligned}$$

1907 *On  $[1, +\infty)$ , the function  $r$  is twice continuously differentiable. It is decreasing on  $[1, e/(1+\eta))$   
 1908 and increasing on  $(e/(1+\eta), +\infty)$ ; its minimum is  $r(e/(1+\eta)) \in (0, 1)$ . It is strictly convex on  
 1909  $[1, e^2/(1+\eta))$  and strictly concave on  $(e^2/(1+\eta), +\infty)$ .*

1910 *Proof.* The proof is obtained by direct differentiation and manipulation. We have

$$\forall \eta > 0, \quad r(e/(1+\eta)) = \frac{e \log(1+\eta)}{1+\eta} \in (0, 1).$$

1911 □

1912 Lemma 39 shows that the modified transportation costs can be rewritten differently, which is a key  
 1913 property used in Appendix F.

1914 **Lemma 39.** *Let  $\widetilde{d}_\epsilon^\pm$  as in Eq. (32), and  $r$  as in Eq. (33). For all  $(\lambda, \mu) \in \mathbb{R}^2$  such that  $[\lambda]_0^1 > [\mu]_0^1$   
 1915 and  $(w_1, w_2) \in [1, +\infty)^2$ . Then,*

$$\begin{aligned} & \inf_{u \in (0,1)} \{w_1 \widetilde{d}_\epsilon^-(\lambda, u, r(w_1)) + w_2 \widetilde{d}_\epsilon^+(\mu, u, r(w_2))\} \\ &= \inf_{u \in ([\mu]_0^1, [\lambda]_0^1)} \{w_1 \widetilde{d}_\epsilon^-(\lambda, u, r(w_1)) + w_2 \widetilde{d}_\epsilon^+(\mu, u, r(w_2))\} \\ &= \inf_{(u_1, u_2) \in (0,1)^2: u_1 \leq u_2} \{w_1 \widetilde{d}_\epsilon^-(\lambda, u_1, r(w_1)) + w_2 \widetilde{d}_\epsilon^+(\mu, u_2, r(w_2))\}. \end{aligned}$$

1916 *Proof.* These results are obtained by leveraging Lemmas 30 and 31.

1917 Note that the condition  $[\lambda]_0^1 > [\mu]_0^1$  implies that  $\mu \in (-\infty, 1)$  and  $\lambda \in (0, +\infty)$ , i.e.,  $[\mu]_0^1 =$   
1918  $\max\{0, \mu\}$  and  $[\lambda]_0^1 = \min\{1, \lambda\}$ .

1919 Suppose that  $\mu \leq 0$  and  $\lambda \geq 1$ . Then, we have  $[\mu]_0^1 = 0$  and  $[\lambda]_0^1 = 1$ . Therefore, the first part of the  
1920 result holds by definition.

1921 Suppose that  $\mu \leq 0$  and  $\lambda \in (0, 1)$ . Then, we have  $[\mu]_0^1 = 0$  and  $[\lambda]_0^1 = \lambda$ . For  $u \in [\lambda, 1)$ , the  
1922 function is equal to  $w_2 \tilde{d}_\epsilon^+(\mu, u, r(w_2))$ , which is increasing and strictly convex on  $(0, 1)$ . There-  
1923 fore, the minimum over that interval is attained at  $\lambda$ . For  $u \in (0, \lambda)$ , the function is equal to  
1924  $w_1 \tilde{d}_\epsilon^-(\lambda, u, r(w_1)) + w_2 \tilde{d}_\epsilon^+(\mu, u, r(w_2))$ . Since it is the sum of two strictly convex function, the  
1925 minimum over that interval is achieved in  $(0, \lambda)$ . This concludes the proof of the first part of the  
1926 result for this case.

1927 Suppose that  $\mu \in (0, 1)$  and  $\lambda \geq 1$ . Then, we have  $[\mu]_0^1 = \mu$  and  $[\lambda]_0^1 = 1$ . For  $u \in (0, \mu]$ , the  
1928 function is equal to  $w_1 \tilde{d}_\epsilon^-(\lambda, u, r(w_1))$ , which is decreasing and strictly convex on  $(0, 1)$ . There-  
1929 fore, the minimum over that interval is attained at  $\mu$ . For  $u \in (\mu, 1)$ , the function is equal to  
1930  $w_1 \tilde{d}_\epsilon^-(\lambda, u, r(w_1)) + w_2 \tilde{d}_\epsilon^+(\mu, u, r(w_2))$ . Since it is the sum of two strictly convex function, the  
1931 minimum over that interval is achieved in  $(\mu, 1)$ . This concludes the proof of the first part of the  
1932 result for this case.

1933 Suppose that  $(\mu, \lambda) \in (0, 1)^2$ . Then, we have  $[\mu]_0^1 = \mu$  and  $[\lambda]_0^1 = \lambda$ . For  $u \in [\lambda, 1)$ , the function  
1934 is equal to  $w_2 \tilde{d}_\epsilon^+(\mu, u, r(w_2))$ , which is increasing and strictly convex on  $(0, 1)$ . Therefore, the  
1935 minimum over that interval is attained at  $\lambda$ . For  $u \in (0, \mu]$ , the function is equal to  $w_1 \tilde{d}_\epsilon^-(\lambda, u, r(w_1))$ ,  
1936 which is decreasing and strictly convex on  $(0, 1)$ . Therefore, the minimum over that interval is attained  
1937 at  $\mu$ . For  $u \in (\mu, \lambda)$ , the function is equal to  $w_1 \tilde{d}_\epsilon^-(\lambda, u, r(w_1)) + w_2 \tilde{d}_\epsilon^+(\mu, u, r(w_2))$ . Since it is  
1938 the sum of two strictly convex function, the minimum over that interval is achieved in  $(\mu, \lambda)$ . This  
1939 concludes the proof of the first part of the result for this case.

1940 In summary, we have shown that

$$\begin{aligned} & \inf_{u \in (0, 1)} \{w_1 \tilde{d}_\epsilon^-(\lambda, u, r(w_1)) + w_2 \tilde{d}_\epsilon^+(\mu, u, r(w_2))\} \\ &= \inf_{u \in ([\mu]_0^1, [\lambda]_0^1)} \{w_1 \tilde{d}_\epsilon^-(\lambda, u, r(w_1)) + w_2 \tilde{d}_\epsilon^+(\mu, u, r(w_2))\}. \end{aligned}$$

1941 Using the strict convexity of  $u_1 \mapsto w_1 \tilde{d}_\epsilon^-(\lambda, u_1, r(w_1))$  and  $u_2 \mapsto w_2 \tilde{d}_\epsilon^+(\mu, u_2, r(w_2))$  on  
1942  $([\mu]_0^1, [\lambda]_0^1)$ , we obtain that

$$\begin{aligned} & \inf_{u \in ([\mu]_0^1, [\lambda]_0^1)} \{w_1 \tilde{d}_\epsilon^-(\lambda, u, r(w_1)) + w_2 \tilde{d}_\epsilon^+(\mu, u, r(w_2))\} \\ &= \inf_{(u_1, u_2) : [\mu]_0^1 < u_1 \leq u_2 < [\lambda]_0^1} \{w_1 \tilde{d}_\epsilon^-(\lambda, u_1, r(w_1)) + w_2 \tilde{d}_\epsilon^+(\mu, u_2, r(w_2))\}. \end{aligned}$$

1943 Re-using the same arguments as above, we obtain that

$$\begin{aligned} & \inf_{(u_1, u_2) : [\mu]_0^1 < u_1 \leq u_2 < [\lambda]_0^1} \{w_1 \tilde{d}_\epsilon^-(\lambda, u_1, r(w_1)) + w_2 \tilde{d}_\epsilon^+(\mu, u_2, r(w_2))\} = \\ &= \inf_{(u_1, u_2) \in (0, 1)^2 : u_1 \leq u_2} \{w_1 \tilde{d}_\epsilon^-(\lambda, u_1, r(w_1)) + w_2 \tilde{d}_\epsilon^+(\mu, u_2, r(w_2))\}. \end{aligned}$$

1944 This concludes the proof.  $\square$

1945 Lemma 40 gives a closed-form solution for the modified transportation costs based on an implicit  
1946 solution of a fixed-point equation. This is a key property used in our implementation to reduce the  
1947 computational cost.

1948 **Lemma 40.** Let  $\tilde{d}_\epsilon^\pm$  as in Eq. (32),  $x_\epsilon^\pm$  as in Lemmas 30 and 31, and  $r$  as in Eq. (33). For all  
1949  $(\lambda, \mu) \in \mathbb{R}^2$  such that  $[\lambda]_0^1 > [\mu]_0^1$  and  $w \in [1, +\infty)^2$ . Then,

$$\begin{aligned} & \inf_{u \in (0, 1)} \{w_1 \tilde{d}_\epsilon^-(\lambda, u, r(w_1)) + w_2 \tilde{d}_\epsilon^+(\mu, u, r(w_2))\} \\ &= w_1 \tilde{d}_\epsilon^-(\lambda, u^*(\lambda, \mu, w), r(w_1)) + w_2 \tilde{d}_\epsilon^+(\mu, u^*(\lambda, \mu, w), r(w_2)), \end{aligned}$$

1950 where  $u^*(\lambda, \mu, w) \in ([\mu]_0^1, [\lambda]_0^1)$  is the unique solution for  $u \in ([\mu]_0^1, [\lambda]_0^1)$  of the equation

$$u(w_1 + w_2) - w_1 g_\epsilon^-(u) - w_2 g_\epsilon^+(u) + w_1 x_\epsilon^-(\lambda, u, r(w_1)) - w_2 x_\epsilon^+(\mu, u, r(w_2)) = 0.$$

1951 *Proof.* Using Lemma 39, we have

$$\begin{aligned} & \inf_{u \in (0,1)} \{w_1 \tilde{d}_\epsilon^-(\lambda, u, r(w_1)) + w_2 \tilde{d}_\epsilon^+(\mu, u, r(w_2))\} \\ &= \inf_{u \in ([\mu]_0^1, [\lambda]_0^1)} \{w_1 \tilde{d}_\epsilon^-(\lambda, u, r(w_1)) + w_2 \tilde{d}_\epsilon^+(\mu, u, r(w_2))\}. \end{aligned}$$

1952 Using Lemmas 31 and 30, we obtain

$$\begin{aligned} \forall u \in (0, [\lambda]_0^1), \quad & \frac{\partial \tilde{d}_\epsilon^-}{\partial u}(\lambda, u, r(w_1)) = \frac{u - g_\epsilon^-(u) + x_\epsilon^-(\lambda, u, r(w_1))}{u(1-u)}, \\ \forall u \in ([\mu]_0^1, 1), \quad & \frac{\partial \tilde{d}_\epsilon^+}{\partial u}(\mu, u, r(w_2)) = \frac{u - g_\epsilon^+(u) - x_\epsilon^+(\mu, u, r(w_2))}{u(1-u)}. \end{aligned}$$

1953 Therefore, for all  $u \in ([\mu]_0^1, [\lambda]_0^1)$ ,

$$\begin{aligned} & w_1 \frac{\partial \tilde{d}_\epsilon^-}{\partial u}(\lambda, u, r(w_1)) + w_2 \frac{\partial \tilde{d}_\epsilon^+}{\partial u}(\mu, u, r(w_2)) \\ &= \frac{w_1(u - g_\epsilon^-(u) + x_\epsilon^-(\lambda, u, r(w_1))) + w_2(u - g_\epsilon^+(u) - x_\epsilon^+(\mu, u, r(w_2)))}{u(1-u)} \\ &= \frac{u(w_1 + w_2) - (w_1 g_\epsilon^-(u) + w_2 g_\epsilon^+(u)) + w_1 x_\epsilon^-(\lambda, u, r(w_1)) - w_2 x_\epsilon^+(\mu, u, r(w_2))}{u(1-u)}. \end{aligned}$$

1954 For  $u \in ([\mu]_0^1, [\lambda]_0^1)$ , let us define

$$g_1(u) := u(w_1 + w_2) - w_1 g_\epsilon^-(u) - w_2 g_\epsilon^+(u) + w_1 x_\epsilon^-(\lambda, u, r(w_1)) - w_2 x_\epsilon^+(\mu, u, r(w_2)).$$

1955 Using the proof of Lemmas 31 and 30, we know that

$$\begin{aligned} \lim_{u \rightarrow [\mu]_0^1} \frac{\partial \tilde{d}_\epsilon^+}{\partial u}(\mu, u, r(w_2)) = 0 \quad \text{and} \quad \lim_{u \rightarrow [\lambda]_0^1} \frac{\partial \tilde{d}_\epsilon^-}{\partial u}(\lambda, u, r(w_1)) = 0, \\ \forall u \in (0, [\lambda]_0^1), \quad \frac{\partial \tilde{d}_\epsilon^-}{\partial u}(\lambda, u, r(w_1)) < 0 \quad \text{and} \quad \forall u \in ([\mu]_0^1, 1), \quad \frac{\partial \tilde{d}_\epsilon^+}{\partial u}(\mu, u, r(w_2)) > 0. \end{aligned}$$

1956 Combined with the strict convexity of  $\tilde{d}_\epsilon^\pm$  in their second argument, the equation  $g_1(u) = 0$  admits  
1957 a unique solution on  $([\mu]_0^1, [\lambda]_0^1)$ . Since  $u(1-u) > 0$ , we obtain the implicit equation defining  
1958  $u^*(\lambda, \mu, w)$  as above.  $\square$

### 1959 G.3 Characteristic Time

1960 Let  $\nu$  be a Bernoulli instance with means  $\mu \in (0, 1)^2$  and unique best arm  $a^* \in [K]$ , i.e.,  
1961  $\arg \max_{a \in [K]} \mu_a = \{a^*\}$ . For all  $\beta \in (0, 1)$ , we define

$$\begin{aligned} T_\epsilon^*(\nu)^{-1} &= \sup_{w \in \triangle_K} \min_{a \neq a^*} W_{\epsilon, a^*, b}(\mu, w) \quad \text{and} \quad w_\epsilon^*(\nu) = \arg \max_{w \in \triangle_K} \min_{a \neq a^*} W_{\epsilon, a^*, b}(\mu, w), \\ T_{\epsilon, \beta}^*(\nu)^{-1} &= \sup_{w \in \triangle_K, w_{a^*} = \beta} \min_{a \neq a^*} W_{\epsilon, a^*, b}(\mu, w) \quad \text{and} \quad w_{\epsilon, \beta}^*(\nu) = \arg \max_{w \in \triangle_K, w_{a^*} = \beta} \min_{a \neq a^*} W_{\epsilon, a^*, b}(\mu, w) \end{aligned} \quad (35)$$

1962 where  $W_{\epsilon, a, b}$  are defined in Eq. (4).

1963 Lemma 41 gathers regularity properties on the characteristic times and their optimal allocations.

1964 **Lemma 41.** Let  $W_{\epsilon, a, b}$  as in Eq. (4). Let  $(T_\epsilon^*, T_{\epsilon, \beta}^*)$  and  $(w_\epsilon^*, w_{\epsilon, \beta}^*)$  as in Eq. (35). The function  
1965  $(\mu, w) \mapsto \min_{a \neq a^*(\mu)} W_{\epsilon, a^*(\mu), a}(\mu, w)$  is continuous on  $(0, 1)^K \times \triangle_K$ . The functions  $\nu \mapsto T_\epsilon^*(\nu)^{-1}$   
1966 and  $\nu \mapsto T_{\epsilon, \beta}^*(\nu)^{-1}$  are continuous on  $\mathcal{F}^K$ . The correspondences  $\nu \mapsto w_\epsilon^*(\nu)$  and  $\nu \mapsto w_{\epsilon, \beta}^*(\nu)$   
1967 are upper hemicontinuous on  $\mathcal{F}^K$  with compact convex values.



1968 *Proof.* Let  $\mathcal{F}_a^K = \{\nu \in \mathcal{F}^K \mid a \in a^*(\nu)\}$ . Since  $\bigcup_{a \in [K]} \mathcal{F}_a^K = \mathcal{F}^K$ , it is enough to show the  
 1969 property for all  $\mathcal{F}_a^K$  for  $a \in [K]$ . Let  $a^* \in [K]$ .

1970 First, the function  $(w, \nu) \mapsto \min_{a \neq a^*} \inf_{u \in [0,1]} \{w_{a^*} d_\epsilon^-(\mu_{a^*}, u) + w_a d_\epsilon^+(\mu_{a^*}, u)\}$  is continuous on  
 1971  $\Delta_K \times \mathcal{F}^K$  by Lemma 34 and the fact that a minimum of continuous functions is continuous. It is  
 1972 concave in  $w$  by Lemma 36.

1973 The correspondence  $(w, \nu) \mapsto \Delta_K$  is nonempty compact-valued and continuous (since constant). By  
 1974 Berge's maximum theorem, we get that  $\nu \mapsto T_\epsilon^*(\nu)^{-1}$  is continuous on  $\mathcal{F}_a^K$  and that  $\nu \mapsto w_\epsilon^*(\nu)$  is  
 1975 upper hemicontinuous with compact values. By [78, Theorem 9.17], the concavity of the function  
 1976 being maximized implies that  $\nu \mapsto w_\epsilon^*(\nu)$  is convex-valued.

1977 The correspondence  $(w, \nu) \mapsto \Delta_K \cap \{w_{a^*} = \beta\}$  is nonempty compact-valued and continuous (since  
 1978 constant). By Berge's maximum theorem, we get that  $\nu \mapsto T_{\epsilon, \beta}^*(\nu)^{-1}$  is continuous on  $\mathcal{F}_a^K$  and that  
 1979  $w_{\beta}^*(\nu)$  is upper hemicontinuous with compact values. By [78, Theorem 9.17], the concavity of the  
 1980 function being maximized implies that  $\nu \mapsto w_{\epsilon, \beta}^*(\nu)$  is convex-valued.  $\square$

1981 Lemma 42 provides additional properties on the characteristic times and their optimal allocations.  
 1982 In particular, this results show that the  $(\beta)$ -optimal allocations is unique, has positive allocation for  
 1983 each arm and that the transportation costs are equal at equilibrium. Those properties are key in the  
 1984 analysis of a sampling rule.

1985 **Lemma 42.** Let  $W_{\epsilon, a, b}$  as in Eq. (4). Let  $(T_\epsilon^*, T_{\epsilon, \beta}^*)$  and  $(w_\epsilon^*, w_{\epsilon, \beta}^*)$  as in Eq. (35). Let  $\beta \in (0, 1)$   
 1986 and  $\nu \in \mathcal{F}^K$  such that  $a^*(\nu) = \{a^*\}$  is a singleton.

- 1987 •  $T_\epsilon^*(\nu)^{-1} > 0$  and  $T_{\epsilon, \beta}^*(\nu)^{-1} > 0$ .
- 1988 •  $\min_{a \in [K]} w_a^* > 0$  and  $\min_{a \in [K]} w_{\beta, a}^* > 0$  for all  $w^* \in w_\epsilon^*(\nu)$  and  $w_\beta^* \in w_{\epsilon, \beta}^*(\nu)$ .
- 1989 • the  $(\beta)$ -optimal allocations are unique and the transportation costs are all equals at equilibrium  
 $w_\epsilon^*(\nu) = \{w_\epsilon^*\}$  and  $\forall a \neq a^*, \inf_{u \in [0,1]} \{w_{\epsilon, a^*}^* d_\epsilon^-(\mu_{a^*}, u) + w_{\epsilon, a}^* d_\epsilon^+(\mu_{a^*}, u)\} = T_\epsilon^*(\nu)^{-1}$ ,  
 $w_{\epsilon, \beta}^*(\nu) = \{w_{\epsilon, \beta}^*\}$  and  $\forall a \neq a^*, \inf_{u \in [0,1]} \{w_{\epsilon, \beta, a^*}^* d_\epsilon^-(\mu_{a^*}, u) + w_{\epsilon, \beta, a}^* d_\epsilon^+(\mu_{a^*}, u)\} = T_{\epsilon, \beta}^*(\nu)^{-1}$

1990 *Proof.* Using the definition of the supremum with  $1_K/K \in \Delta_K$  and Lemma 34, we obtain

$$\begin{aligned} T_\epsilon^*(\nu)^{-1} &= \sup_{w \in \Delta_K} \min_{a \neq a^*} \inf_{u \in [0,1]} \{w_{a^*} d_\epsilon^-(\mu_{a^*}, u) + w_a d_\epsilon^+(\mu_{a^*}, u)\} \\ &\geq \frac{1}{K} \min_{a \neq a^*} \inf_{u \in [0,1]} \{d_\epsilon^-(\mu_{a^*}, u) + d_\epsilon^+(\mu_{a^*}, u)\} > 0, \end{aligned}$$

1991 where the last inequality strict uses Lemma 34 and  $\mu_a < \mu_{a^*}$  for all  $a \neq a^*$ . Similarly, we can prove  
 1992 that  $T_{\epsilon, \beta}^*(\nu)^{-1} > 0$ . This concludes the first part of the proof.

1993 We proceed towards contradiction. Suppose that there exists  $w^* \in w_\epsilon^*(\nu)$  and  $b$  with  $w_b^* = 0$ . Then,  
 1994 we will show  $T_\epsilon^*(\nu)^{-1} = 0$ , which is a contradiction with the above result. If  $b = a^*$  we have

$$T_\epsilon^*(\nu)^{-1} = \min_{a \neq a^*} \inf_{u \in [0,1]} w_a^* d_\epsilon^+(\mu_{a^*}, u) \leq \min_{a \neq a^*} w_a^* d_\epsilon^+(\mu_a, \mu_a) = 0.$$

1995 If  $b \neq a^*$ , we have

$$\begin{aligned} T_\epsilon^*(\nu)^{-1} &= \min_{a \neq a^*} \inf_{u \in [0,1]} \{w_{a^*}^* d_\epsilon^-(\mu_{a^*}, u) + w_a^* d_\epsilon^+(\mu_{a^*}, u)\} \\ &\leq \inf_{u \in [0,1]} \{w_{a^*}^* d_\epsilon^-(\mu_{a^*}, u) + w_b^* d_\epsilon^+(\mu_b, u)\} = \inf_{u \in [0,1]} w_{a^*}^* d_\epsilon^-(\mu_{a^*}, u) = 0. \end{aligned}$$

1996 A similar proof allows to show the result for  $w_{\epsilon, \beta}^*(\nu)$  by reasoning on  $T_{\epsilon, \beta}^*(\nu)^{-1}$ . This concludes the  
 1997 second part of the proof.

1998 For notational simplicity, we assume without loss of generality that  $a^* = 1$  is the best arm. At the  
 1999 optimal allocations, all  $w_a$  are positive. Let us define  $G_b(x) = \inf_{u \in [0,1]} \{d_\epsilon^-(\mu_1, u) + x d_\epsilon^+(\mu_b, u)\}$   
 2000 for all  $b \neq 1$ . Let  $w^* \in w_\epsilon^*(\nu)$ . Then, we have

$$T_\epsilon^*(\nu)^{-1} = \max_{w \in \Delta_K, w_1 > 0} w_1 \min_{b \neq 1} G_b \left( \frac{w_b}{w_1} \right) \quad \text{and} \quad w^* \in \arg \max_{w \in \Delta_K} w_1 \min_{b \neq 1} G_b \left( \frac{w_b}{w_1} \right).$$

2001 Introducing  $x_b^* = \frac{w_b^*}{w_1^*}$  for all  $b \neq 1$ , using that  $\sum_{b \in [K]} w_b^* = 1$ , one has

$$w_1^* = \frac{1}{1 + \sum_{c \neq 1} x_c^*} \quad \text{and} \quad \forall b \neq 1, w_b^* = \frac{x_b^*}{1 + \sum_{c \neq 1} x_c^*}.$$

2002 If  $x^*$  is unique, then so is  $w^*$ . Since it is optimal,  $\{x_b^*\}_{b=2}^K \in \mathbb{R}^{K-1}$  belongs to

$$\arg \max_{\{x_b\}_{b=2}^K \in \mathbb{R}^{K-1}} \frac{\min_{b \neq 1} G_b(x_b)}{1 + \sum_{c=2}^K x_c}. \quad (36)$$

2003 Let's show that all the  $G_b(x_b^*)$  have to be equal. Let  $\mathcal{O} =$   
 2004  $\{a \in [K] \setminus \{1\} \mid G_a(x_a^*) = \min_{b \neq 1} G_b(x_b^*)\}$  and  $\mathcal{A} = [K] \setminus (\{1\} \cup \mathcal{O})$ . Assume that  $\mathcal{A} \neq \emptyset$ . For  
 2005 all  $a \in \mathcal{A}$  and  $b \in \mathcal{O}$ , one has  $G_b(x_b^*) > G_a(x_a^*)$ . Using the continuity of the  $G_b$  functions and the  
 2006 fact that they are increasing (Lemma 36), there exists  $\epsilon > 0$  such that

$$\forall b \in \mathcal{A}, a \in \mathcal{O}, \quad G_b(x_b^* - \epsilon/|\mathcal{A}|) > G_a(x_a^* + \epsilon/|\mathcal{O}|) > G_a(x_a^*).$$

2007 We introduce  $\bar{x}_b = x_b^* - \epsilon/|\mathcal{A}|$  for all  $b \in \mathcal{A}$  and  $\bar{x}_a = x_a^* + \epsilon/|\mathcal{O}|$  for all  $a \in \mathcal{O}$ , hence  $\sum_{b=2}^K \bar{x}_b =$   
 2008  $\sum_{b=2}^K x_b^*$ . There exists  $a \in \mathcal{O}$  such that  $\min_{b \neq 1} G_b(\bar{x}_b) = G_a(x_a^* + \epsilon/|\mathcal{O}|)$ , hence

$$\frac{\min_{b \neq 1} G_b(\bar{x}_b)}{1 + \bar{x}_2 + \dots + \bar{x}_K} = \frac{G_a(x_a^* + \epsilon/|\mathcal{O}|)}{1 + x_2^* + \dots + x_K^*} > \frac{G_a(x_a^*)}{1 + x_2^* + \dots + x_K^*} = \frac{\min_{b \neq 1} G_b(x_b^*)}{1 + x_2^* + \dots + x_K^*}.$$

2009 This is a contradiction with the fact that  $x^*$  belongs to (36). Therefore, we have  $\mathcal{A} = \emptyset$ .

2010 We have proved that there is a unique value by  $y^* \in \mathbb{R}_+$ , such that for all  $b \neq 1$ ,  $G_b(x_b^*) = y^*$ . Now  
 2011 since  $G_b$  is increasing, this defines a unique value for  $x_b^*$ , equal to  $G_b^{-1}(y^*)$ .

2012 For  $y$  in the intersection of the ranges of all  $G_b$ , let  $x_b(y) = G_b^{-1}(y)$ . Then,  $y^*$  belongs to

$$\arg \max_{y \in [0, \min_{b \neq 1} \lim_{x \rightarrow \infty} G_b(x))} \frac{y}{1 + \sum_{b \neq 1} x_b(y)}. \quad (37)$$

2013 For  $\beta \in (0, 1)$ , the same results (and proof) hold for  $w_{\epsilon, \beta}^*(\nu)$  by noting that

$$T_{\epsilon, \beta}^*(\nu)^{-1} = \max_{w \in \Delta_K : w_1 = \beta} \beta \min_{b \neq 1} G_b(w_b/\beta).$$

2014 Let  $w_{\epsilon, \beta}^* \in w_{\epsilon, \beta}^*(\nu)$ , since we have equality at the equilibrium, we obtain  $\beta G_b(w_{\epsilon, \beta, b}^*/\beta) =$   
 2015  $T_{\epsilon, \beta}^*(\nu)^{-1}$  for all  $b \neq 1$ . Using the inverse mapping  $x_b$ , we obtain  $w_{\epsilon, \beta, b}^* = \beta x_b(T_{\epsilon, \beta}^*(\nu)^{-1}/\beta)$  for  
 2016 all  $b \neq 1$ . This concludes the third part of the proof.  $\square$

2017 Lemma 43 shows that an asymptotically 1/2-optimal algorithm has an asymptotic expected sample  
 2018 complexity which is at worse twice the asymptotic expected sample complexity of an asymptotically  
 2019 optimal algorithm. This result motivates the recommendation to the practitioner of using  $\beta = 1/2$   
 2020 when no prior information is available on the true instance  $\nu$ .

2021 **Lemma 43.** Let  $(T_\epsilon^*, T_{\epsilon, \beta}^*, w_\epsilon^*)$  as in Eq. (35). Let  $\beta \in (0, 1)$  and  $\nu \in \mathcal{F}^K$  such that  $a^*(\nu) = \{a^*\}$   
 2022 is a singleton. Then,

$$T_{\epsilon, 1/2}^*(\nu) \leq 2T_\epsilon^*(\nu) \quad \text{and} \quad \frac{T_\epsilon^*(\nu)^{-1}}{T_{\epsilon, \beta}^*(\nu)^{-1}} \leq \max \left\{ \frac{\beta^*}{\beta}, \frac{1 - \beta^*}{1 - \beta} \right\} \quad \text{with} \quad \beta^* = w_{\epsilon, a^*}^*.$$

2023 *Proof.* Define for each non-negative vector  $\psi \in \mathbb{R}_+^K$ ,

$$f(\psi) := \min_{a \neq a^*} \inf_{u \in [0, 1]} \{ \psi_{a^*} d_\epsilon^-(\mu_{a^*}, u) + \psi_a d_\epsilon^+(\mu_a, u) \}.$$

$T_\epsilon^*(\nu)^{-1}$  is the maximum of  $f(\psi)$  over probability vectors  $\psi \in \Delta_K$ . Here, we instead define  $f$  for all non-negative vectors, and proceed by varying the total budget of measurement effort available

$\sum_{a \in [K]} \psi_a$ . Using Lemma 36,  $f$  is non-decreasing in  $\psi_a$  for all  $a$ .  $f$  is homogeneous of degree 1. That is  $f(c\psi) = cf(\psi)$  for all  $c \geq 1$ . For each  $c_1, c_2 > 0$  define

$$g(c_1, c_2) = \max \left\{ f(\psi) \mid \psi \in \mathbb{R}_+^K, \psi_{a^*} = c_1, \sum_{a \neq a^*} \psi_a \leq c_2, \right\}.$$

The function  $g$  inherits key properties of  $f$ ; it is also non-decreasing and homogeneous of degree 1. We have

$$\begin{aligned} T_{\epsilon, \beta}^*(\nu)^{-1} &= \max \left\{ f(\psi) \mid \psi \in \mathbb{R}_+^K, \psi_{a^*} = \beta, \sum_{a \in [K]} \psi_a = 1 \right\} \\ &= \max \left\{ f(\psi) \mid \psi \in \mathbb{R}_+^K, \psi_{a^*} = \beta, \sum_{a \neq a^*} \psi_a \leq 1 - \beta \right\} = g(\beta, 1 - \beta), \end{aligned}$$

where the second equality uses that  $f$  is non-decreasing. Similarly,  $T_\epsilon^*(\nu)^{-1} = g(\beta^*, 1 - \beta^*)$  where  $\beta^* = w_{\epsilon, a^*}^*$ . Setting  $r := \max \left\{ \frac{\beta^*}{\beta}, \frac{1 - \beta^*}{1 - \beta} \right\}$  implies  $r\beta \geq \beta^*$  and  $r(1 - \beta) \geq 1 - \beta^*$ . Therefore

$$rT_{\epsilon, \beta}^*(\nu)^{-1} = rg(\beta, 1 - \beta) = g(r\beta, r(1 - \beta)) \geq g(\beta^*, 1 - \beta^*) = T_\epsilon^*(\nu)^{-1}.$$

2024 Taking  $\beta = \frac{1}{2}$ , yields that  $T_\epsilon^*(\nu)^{-1} \leq 2 \max\{\beta^*, 1 - \beta^*\} T_{1/2}^*(\nu)^{-1} \leq 2T_{1/2}^*(\nu)^{-1}$ .  $\square$

2025 Lemma 44 gives sufficient conditions on the means and allocations in order for the transportation  
2026 costs to be equals to the non-private transportation costs. Moreover, it gives sufficient conditions on  
2027 the means in order for this equality to hold irrespective of the considered allocation. Taken together,  
2028 this result allows to have fine and coarse understanding of the separation between the high privacy  
2029 regime and the low privacy regime for  $\epsilon$ -global DP BAI.

2030 **Lemma 44.** Let  $W_{\epsilon, a, b}$  as in Eq. (4). Let  $\mu \in (0, 1)^K$  such that  $a^* = \arg \max_{a \in [K]} \mu_a$  is unique.  
2031 Let  $w \in (\mathbb{R}_+^K)$ . Let  $\epsilon > 0$ . For all  $x \in (0, 1)$ , we define  $f_\epsilon(x) := (1 - x) \left( 1 - \frac{1}{1 + x(e^\epsilon - 1)} \right) =$   
2032  $(1 - x)g_\epsilon^-(x)(1 - e^{-\epsilon})$ . Let us define  $\mu_{a^*, a}^w := \frac{w_{a^*}\mu_{a^*} + w_a\mu_a}{w_{a^*} + w_a}$  for all  $a \neq a^*$ . For all  $a \neq a^*$ , we  
2033 have

$$\begin{aligned} \mu_{a^*} - \mu_a &\leq \min \left\{ \left( 1 + \frac{w_{a^*}}{w_a} \right) f_\epsilon(1 - \mu_{a^*}), \left( 1 + \frac{w_a}{w_{a^*}} \right) f_\epsilon(\mu_a) \right\} \\ \implies W_{\epsilon, a^*, a}(\mu, w) &= w_{a^*} \text{kl}(\mu_{a^*}, \mu_{a^*, a}^w) + w_a \text{kl}(\mu_a, \mu_{a^*, a}^w). \end{aligned}$$

2034 Moreover, we have

$$\max_{a^* \in [K], \mu \in (0, 1)^K, a^*(\mu) = \{a^*\}, w \in (\mathbb{R}_+^K)^K} \min \left\{ \left( 1 + \frac{w_{a^*}}{w_a} \right) f_\epsilon(1 - \mu_{a^*}), \left( 1 + \frac{w_a}{w_{a^*}} \right) f_\epsilon(\mu_a) \right\} \leq \epsilon/2$$

2035 and, for all  $a \neq a^*$ , we have

$$\begin{aligned} \epsilon &\geq \log \left( \frac{\mu_{a^*}(1 - \mu_a)}{\mu_a(1 - \mu_{a^*})} \right) = \frac{\partial \text{kl}}{\partial x_1}(\mu_{a^*}, \mu_a) = \frac{\partial \text{kl}}{\partial x_1}(\mu_a, \mu_{a^*}) \\ \implies \forall w \in (\mathbb{R}_+^K)^K, \quad W_{\epsilon, a^*, a}(\mu, w) &= w_{a^*} \text{kl}(\mu_{a^*}, \mu_{a^*, a}^w) + w_a \text{kl}(\mu_a, \mu_{a^*, a}^w). \end{aligned}$$

2036 *Proof.* Let us define  $f_\epsilon(x) = (1 - x) \left( 1 - \frac{1}{1 + x(e^\epsilon - 1)} \right)$  for all  $x \in (0, 1)$ . Then, we have

$$\begin{aligned} \frac{\mu_a(1 - \mu_a)(e^\epsilon - 1)}{1 + \mu_a(e^\epsilon - 1)} &= (1 - \mu_a) \left( 1 - \frac{1}{1 + \mu_a(e^\epsilon - 1)} \right) = f_\epsilon(\mu_a), \\ \frac{\mu_{a^*}(1 - \mu_{a^*})(e^\epsilon - 1)}{e^\epsilon - \mu_{a^*}(e^\epsilon - 1)} &= \mu_{a^*} \left( 1 - \frac{1}{1 + (1 - \mu_{a^*})(e^\epsilon - 1)} \right) = f_\epsilon(1 - \mu_{a^*}). \end{aligned}$$

2037 Using Lemma 23, direct manipulation yields that

$$f_\epsilon(1 - \mu_{a^*}) < \mu_{a^*} - \mu_a \iff g_\epsilon^+(\mu_{a^*}) > \mu_a \iff f_\epsilon(\mu_a) < \mu_{a^*} - \mu_a$$

$$\begin{aligned}
g_\epsilon^-(\mu_a) < \mu_{a^*,a}^w &\iff f_\epsilon(\mu_a) < \frac{w_{a^*}}{w_{a^*} + w_a}(\mu_{a^*} - \mu_a), \\
g_\epsilon^+(\mu_{a^*}) > \mu_{a^*,a}^w &\iff f_\epsilon(1 - \mu_{a^*}) < \frac{w_a}{w_{a^*} + w_a}(\mu_{a^*} - \mu_a).
\end{aligned}$$

2038 Using that  $\max \left\{ \frac{w_a}{w_{a^*} + w_a}, \frac{w_{a^*}}{w_{a^*} + w_a} \right\} \leq 1$ , we obtain that

$$\begin{aligned}
(g_\epsilon^-(\mu_a) < \mu_{a^*} \wedge g_\epsilon^-(\mu_a) < \mu_{a^*,a}^w) &\iff \left(1 + \frac{w_a}{w_{a^*}}\right) f_\epsilon(\mu_a) < \mu_{a^*} - \mu_a, \\
(g_\epsilon^-(\mu_a) < \mu_{a^*} \wedge g_\epsilon^+(\mu_{a^*}) > \mu_{a^*,a}^w) &\iff \left(1 + \frac{w_{a^*}}{w_a}\right) f_\epsilon(1 - \mu_{a^*}) < \mu_{a^*} - \mu_a, \\
(g_\epsilon^-(\mu_a) \geq \mu_{a^*} \vee (g_\epsilon^-(\mu_a) < \mu_{a^*} \wedge \mu_{a^*,a}^w \in [g_\epsilon^+(\mu_{a^*}), g_\epsilon^-(\mu_a)])) & \\
\iff (g_\epsilon^-(\mu_a) \geq \mu_{a^*} \vee \mu_{a^*,a}^w \in [g_\epsilon^+(\mu_{a^*}), g_\epsilon^-(\mu_a)]) & \\
\iff (\min\{f_\epsilon(\mu_a), f_\epsilon(1 - \mu_{a^*})\} \geq \mu_{a^*} - \mu_a & \\
\vee \mu_{a^*} - \mu_a \leq \min \left\{ \left(1 + \frac{w_{a^*}}{w_a}\right) f_\epsilon(1 - \mu_{a^*}), \left(1 + \frac{w_a}{w_{a^*}}\right) f_\epsilon(\mu_a) \right\} & \\
\iff \mu_{a^*} - \mu_a \leq \max \{ \min\{f_\epsilon(\mu_a), f_\epsilon(1 - \mu_{a^*})\}, & \\
\min \left\{ \left(1 + \frac{w_{a^*}}{w_a}\right) f_\epsilon(1 - \mu_{a^*}), \left(1 + \frac{w_a}{w_{a^*}}\right) f_\epsilon(\mu_a) \right\} \} & \\
\iff \mu_{a^*} - \mu_a \leq \min \left\{ \left(1 + \frac{w_{a^*}}{w_a}\right) f_\epsilon(1 - \mu_{a^*}), \left(1 + \frac{w_a}{w_{a^*}}\right) f_\epsilon(\mu_a) \right\}. &
\end{aligned}$$

2039 Combining those conditions with Lemma 37 concludes the first part of the proof.

2040 For all  $x \in (0, 1)$ , we have

$$\begin{aligned}
f'_\epsilon(x) &= \frac{(1-x)(e^\epsilon - 1) - x(e^\epsilon - 1)(1 + x(e^\epsilon - 1))}{(1 + x(e^\epsilon - 1))^2} = -(e^\epsilon - 1) \frac{x^2(e^\epsilon - 1) + 2x - 1}{(1 + x(e^\epsilon - 1))^2}, \\
f'_\epsilon(x) = 0 &\iff x = \frac{e^{\epsilon/2} - 1}{e^\epsilon - 1}, \\
f''_\epsilon(x) &= -2(e^\epsilon - 1) \frac{(1 + x(e^\epsilon - 1))^2 - (e^\epsilon - 1)(x^2(e^\epsilon - 1) + 2x - 1)}{(1 + x(e^\epsilon - 1))^3} \\
&= -\frac{2e^\epsilon(e^\epsilon - 1)}{(1 + x(e^\epsilon - 1))^3} \leq 0.
\end{aligned}$$

2041 As  $f_\epsilon$  is strictly concave, the maximum is achieved at  $\frac{e^{\epsilon/2}-1}{e^\epsilon-1}$  with value

$$\max_{x \in (0,1)} f_\epsilon(x) = f_\epsilon\left(\frac{e^{\epsilon/2}-1}{e^\epsilon-1}\right) = \frac{e^\epsilon - e^{\epsilon/2}}{e^\epsilon - 1} \left(1 - e^{-\epsilon/2}\right) = \frac{(e^{\epsilon/2}-1)^2}{e^\epsilon - 1}.$$

2042 Let  $\kappa_1(x) = x(e^x - 1) - 4(e^{x/2} - 1)^2$  for all  $x > 0$ . Then, we have

$$\frac{(e^{\epsilon/2}-1)^2}{e^\epsilon - 1} \leq \epsilon/4 \iff \kappa_1(\epsilon) \geq 0.$$

2043 Then, we have  $\kappa_1(0) = 0$  and

$$\kappa'_1(x) = 4e^{x/2} - 3e^x - 1 + xe^x \quad \text{and} \quad \kappa''_1(x) = e^x (2(e^{-x/2} - 1) + x).$$

2044 Using that  $e^{-x/2} - 1 \geq -x/2$ , we obtain  $\kappa''_1(x) \geq 0$ . Using that  $\kappa'_1(0) = 0$ , we obtain  $\kappa'_1(x) \geq 0$ .

2045 Using that  $\kappa_1(0) = 0$ , we obtain  $\kappa_1(x) \geq 0$ . Therefore, we have shown that

$$\forall \epsilon > 0, \quad \max_{x \in (0,1)} f_\epsilon(x) \leq \epsilon/4.$$

2046 Direct manipulation yields that

$$\min \left\{ \left(1 + \frac{w_{a^*}}{w_a}\right) f_\epsilon(1 - \mu_{a^*}), \left(1 + \frac{w_a}{w_{a^*}}\right) f_\epsilon(\mu_a) \right\}$$

$$\leq \left(1 + \min \left\{ \frac{w_{a^*}}{w_a}, \frac{w_a}{w_{a^*}} \right\}\right) \max_{x \in (0,1)} f_\epsilon(x) \leq \epsilon/2.$$

2047 Taking the supremum over  $w \in (\mathbb{R}_+^*)^K$ ,  $\mu \in (0,1)^K$  such that  $a^* = a^*(\mu)$  and over  $a^* \in [K]$   
 2048 concludes the second part of the proof.

2049 Let  $a \neq a^*$ . Direct manipulations yield that

$$\begin{aligned} \mu_{a^*} - \mu_a &\leq \min\{f_\epsilon(1 - \mu_{a^*}), f_\epsilon(\mu_a)\} \\ \implies \forall w \in (\mathbb{R}_+^*)^K, \quad \mu_{a^*} - \mu_a &\leq \min \left\{ \left(1 + \frac{w_{a^*}}{w_a}\right) f_\epsilon(1 - \mu_{a^*}), \left(1 + \frac{w_a}{w_{a^*}}\right) f_\epsilon(\mu_a) \right\} \\ \implies \forall w \in (\mathbb{R}_+^*)^K, \quad W_{\epsilon, a^*, a}(\mu, w) &= w_{a^*} \text{kl}(\mu_{a^*}, \mu_{a^*, a}^w) + w_a \text{kl}(\mu_a, \mu_{a^*, a}^w). \end{aligned}$$

2050 Recall that  $f_\epsilon(x) = (1-x) \left(1 - \frac{1}{1+x(e^\epsilon-1)}\right)$ . Then, we have directly that

$$f_\epsilon(x) \geq y \iff \frac{1-x-y}{1-x} \geq \frac{1}{1+x(e^\epsilon-1)} \iff \frac{(y+x)(1-x)}{x(1-x-y)} \leq e^\epsilon.$$

2051 Plugging this result, we obtain

$$\begin{aligned} \mu_{a^*} - \mu_a \leq \min\{f_\epsilon(1 - \mu_{a^*}), f_\epsilon(\mu_a)\} &\iff e^\epsilon \geq \frac{\mu_{a^*}(1 - \mu_a)}{\mu_a(1 - \mu_{a^*})} \\ &\iff \epsilon \geq \log \left( \frac{\mu_{a^*}(1 - \mu_a)}{\mu_a(1 - \mu_{a^*})} \right). \end{aligned}$$

2052 Recall that

$$\frac{\partial \text{kl}}{\partial x_1}(\mu_{a^*}, \mu_a) = \frac{\partial \text{kl}}{\partial x_1}(\mu_a, \mu_{a^*}) = \log \left( \frac{\mu_{a^*}(1 - \mu_a)}{\mu_a(1 - \mu_{a^*})} \right).$$

2053 This concludes the proof of the last part of the result.  $\square$

2054 Lemma 45 shows that our lower bound is larger (hence better) than the one derived in Azize et al.  
 2055 [12].

2056 **Lemma 45.** Let  $T_g^*(\nu, \epsilon)$  as in Theorem 13 in Azize et al. [12], and  $T_\epsilon^*(\nu)$  as in Eq. (35). Then, we  
 2057 have  $T_g^*(\nu, \epsilon) \leq T_\epsilon^*(\nu)$ .

2058 *Proof.* Let  $T_g^*(\nu, \epsilon)$  as in Theorem 13 in Azize et al. [12]. A sufficient condition to obtain  $T_g^*(\nu, \epsilon) \leq$   
 2059  $T_\epsilon^*(\nu)$  is to show that, for all  $\lambda \in \text{Alt}(\mu)$ , we have

$$\sum_{a \in [K]} w_a d_\epsilon(\mu_a, \lambda_a) \leq \min \left\{ \sum_{a \in [K]} w_a \text{kl}(\mu_a, \lambda_a), 6\epsilon \sum_{a \in [K]} w_a |\mu_a - \lambda_a| \right\},$$

2060 since we can conclude by taking the infimum over  $\lambda \in \text{Alt}(\mu)$  and the supremum over  $w \in \triangle_K$  on  
 2061 both sides of the inequalities. By definition of  $d_\epsilon$  and evaluation the function at  $z = \mu$  and  $z = \lambda$   
 2062 respectively, we obtain

$$d_\epsilon(\lambda, \mu) = \inf_{z \in (0,1)} \{\text{kl}(z, \mu) + \epsilon|\lambda - z|\} \leq \min \{\text{kl}(\lambda, \mu), \epsilon|\lambda - \mu|\}.$$

2063 By summing those inequalities over arms  $a \in [K]$ , we obtain

$$\begin{aligned} \sum_{a \in [K]} w_a d_\epsilon(\mu_a, \lambda_a) &\leq \sum_{a \in [K]} w_a \min \{\text{kl}(\mu_a, \lambda_a), \epsilon|\mu_a - \lambda_a|\} \\ &\leq \min \left\{ \sum_{a \in [K]} w_a \text{kl}(\mu_a, \lambda_a), \epsilon \sum_{a \in [K]} w_a |\mu_a - \lambda_a| \right\}. \end{aligned}$$

2064 Using that  $\sum_{a \in [K]} w_a |\mu_a - \lambda_a| \geq 0$  and  $6\epsilon \geq \epsilon$ , this concludes the proof.  $\square$

In Garivier and Kaufmann [38], the authors show how to rewrite the optimization problem underlying the characteristic time and its optimal allocation as a simpler optimization problem. Lemma 46 shows that similar properties holds for  $\epsilon$ -global DP BAI. In particular, it shows that computing the characteristic time  $T_\epsilon^*(\nu)$  and their optimal allocation  $w_\epsilon^*(\nu)$  can be done explicitly based on solving nested fixed-point equations. This result is key to implement computationally tractable Track-and-Stop algorithms. Additionally, Lemma 46 gives an explicit lower bound on the characteristic time  $T_\epsilon^*(\nu)$ .

**Lemma 46.** *Let  $d_\epsilon^\pm$  as in Eq. (3), and  $(T_\epsilon^*, w_\epsilon^*)$  as in Eq. (35). Let  $a \neq a^*$ . For  $x \in [0, +\infty)$ , let*

$$G_a(x) := \inf_{u \in [0,1]} \{d_\epsilon^-(\mu_{a^*}, u) + x d_\epsilon^+(\mu_a, u)\} \text{ and } u_a(x) := \arg \min_{u \in [\mu_a, \mu_{a^*}]} \{d_\epsilon^-(\mu_{a^*}, u) + x d_\epsilon^+(\mu_a, u)\}.$$

- The function  $G_a$  is an increasing and strictly concave one-to-one mapping from  $[0, +\infty)$  to  $[0, d_\epsilon^-(\mu_{a^*}, \mu_a))$ ; it satisfies that  $G_a(0) = 0$  and  $\lim_{x \rightarrow +\infty} G_a(x) = d_\epsilon^-(\mu_{a^*}, \mu_a)$ .
- The function  $u_a$  is a decreasing one-to-one mapping from  $[0, +\infty)$  to  $(\mu_a, \mu_{a^*}]$ ; it satisfies that  $u_a(0) = \mu_{a^*}$  and  $\lim_{x \rightarrow +\infty} u_a(x) = \mu_a$ .
- Let  $x_a(y)$  be defined as the unique solution of  $G_a(x) = y$  for all  $y \in [0, d_\epsilon^-(\mu_{a^*}, \mu_a))$ . The function  $x_a$  is an increasing and strictly convex one-to-one mapping from  $[0, d_\epsilon^-(\mu_{a^*}, \mu_a))$  to  $[0, +\infty)$ ; it satisfies that  $x_a(0) = 0$  and  $\lim_{y \rightarrow d_\epsilon^-(\mu_{a^*}, \mu_a)} x_a(y) = +\infty$ .

For all  $y \in [0, \min_{a \neq a^*} d_\epsilon^-(\mu_{a^*}, \mu_a))$ , let us define

$$G(y) := \frac{y}{1 + \sum_{a \neq a^*} x_a(y)} \quad \text{and} \quad F(y) := \sum_{a \neq a^*} \frac{d_\epsilon^-(\mu_{a^*}, u_a(x_a(y)))}{d_\epsilon^+(\mu_a, u_a(x_a(y)))}.$$

- The function  $F$  is an increasing one-to-one mapping from  $[0, \min_{a \neq a^*} d_\epsilon^-(\mu_{a^*}, \mu_a))$  to  $[0, +\infty)$ ; it satisfies that  $F(0) = 0$  and  $\lim_{y \rightarrow \min_{a \neq a^*} d_\epsilon^-(\mu_{a^*}, \mu_a)} F(y) = +\infty$ .
- On  $[0, \min_{a \neq a^*} d_\epsilon^-(\mu_{a^*}, \mu_a))$ , the function  $G$  is maximized at the unique  $y^*$  solution in  $[0, \min_{a \neq a^*} d_\epsilon^-(\mu_{a^*}, \mu_a))$  of the fixed-point equation  $F(y) = 1$ . Moreover, we have  $w_\epsilon^*(\nu)_{a^*} = w_\epsilon^*(\nu)_{a^*} x_a(y^*)$  for all  $a \neq a^*$ ,

$$w_\epsilon^*(\nu)_{a^*} = \frac{1}{1 + \sum_{a \neq a^*} x_a(y^*)} \quad \text{and} \quad T_\epsilon^*(\nu)^{-1} = \frac{y^*}{1 + \sum_{a \neq a^*} x_a(y^*)}.$$

- Moreover, we have

$$T_\epsilon^*(\nu) \geq \frac{1}{\min_{a \neq a^*} d_\epsilon^-(\mu_{a^*}, \mu_a)} + \sum_{a \neq a^*} \frac{1}{d_\epsilon^+(\mu_a, \mu_{a^*})}.$$

If  $\epsilon < \log\left(\frac{\mu_a(1-\mu_b)}{\mu_b(1-\mu_a)}\right)$ , we have  $d_\epsilon^+(\mu_a, \mu_{a^*}) = -\log(1 - \mu_{a^*}(1 - e^{-\epsilon})) - \epsilon\mu_a$  and  $d_\epsilon^-(\mu_{a^*}, \mu_a) = -\log(1 + \mu_a(e^\epsilon - 1)) + \epsilon\mu_{a^*}$ .

*Proof.* Using Lemma 36, we know that  $G_a$  is concave. Let  $u_a(x) \in \arg \min_{u \in [0,1]} \{d_\epsilon^-(\mu_{a^*}, u) + x d_\epsilon^+(\mu_a, u)\}$  for all  $x \in [0, +\infty)$ , whose explicit formula is given in Lemma 44. It is direct to see that  $G_a(0) = 0$  and  $u_a(0) = \mu_{a^*}$ . Using the optimality condition of  $u_a(x)$ , we obtain, for all  $x \in [0, +\infty)$ ,

$$\begin{aligned} G'_a(x) &= u'_a(x) \left( \frac{\partial d_\epsilon^-}{\partial u}(\mu_{a^*}, u_a(x)) + x \frac{\partial d_\epsilon^+}{\partial u}(\mu_a, u_a(x)) \right) + d_\epsilon^+(\mu_a, u_a(x)) \\ &= d_\epsilon^+(\mu_a, u_a(x)) > 0, \end{aligned}$$

where the last inequality is obtained by Lemma 34 and using that  $d_\epsilon^+(\mu_a, u_a(0)) = d_\epsilon^+(\mu_a, \mu_{a^*}) > 0$ . Therefore,  $G_a$  is an increasing one-to-one mapping from  $[0, +\infty)$  to  $[0, \lim_{x \rightarrow +\infty} G_a(x))$ .

Let  $\mu_{a^*,a}^x = \frac{\mu_{a^*} + x\mu_a}{1+x}$  for all  $x \in [0, +\infty)$ . It is easy to see that  $G_a(0) = 0$ ,  $u_a(0) = \mu_{a^*}$  and  $\lim_{x \rightarrow +\infty} \mu_{a^*,a}^x = \mu_a$ . Using Lemma 44, we obtain that

$$\lim_{x \rightarrow +\infty} \min \{(1 + 1/x) f_\epsilon(1 - \mu_{a^*}), (1 + x) f_\epsilon(\mu_a)\} = f_\epsilon(1 - \mu_{a^*}).$$

2097 When  $\mu_{a^*} - \mu_a \leq f_\epsilon(1 - \mu_{a^*})$ , we obtain

$$\begin{aligned} \lim_{x \rightarrow +\infty} G_a(x) &= \lim_{x \rightarrow +\infty} \{ \text{kl}(\mu_{a^*}, \mu_{a^*,a}^x) + x \text{kl}(\mu_a, \mu_{a^*,a}^x) \} \\ &= \text{kl}(\mu_{a^*}, \mu_a) + \lim_{x \rightarrow +\infty} \{ x \text{kl}(\mu_a, \mu_{a^*,a}^x) \} = \text{kl}(\mu_{a^*}, \mu_a), \end{aligned}$$

2098 where we used that

$$\begin{aligned} x \text{kl}(\mu_a, \mu_{a^*,a}^x) &= x \left( \mu_a \log \left( 1 - \frac{\mu_{a^*} - \mu_a}{\mu_a(1 - \mu_a)} \frac{1}{\frac{\mu_{a^*}}{\mu_a} + x} \right) + \log \left( 1 + \frac{\mu_{a^*} - \mu_a}{1 - \mu_a} \frac{1}{\frac{1 - \mu_{a^*}}{1 - \mu_a} + x} \right) \right), \\ \lim_{x \rightarrow +\infty} \{ x \text{kl}(\mu_a, \mu_{a^*,a}^x) \} &= \frac{(\mu_{a^*} - \mu_a)^2}{\mu_a(1 - \mu_a)^2} \lim_{x \rightarrow +\infty} \left\{ \frac{x}{\left( \frac{\mu_{a^*}}{\mu_a} + x \right) \left( \frac{1 - \mu_{a^*}}{1 - \mu_a} + x \right)} \right\} = 0, \end{aligned}$$

2099 where we used that  $\log(1 + x) =_{x \rightarrow 0} x + \mathcal{O}(x^2)$ . Using Lemma 25 and the proof of Lemma 44, we  
2100 know that  $\mu_{a^*} - \mu_a \leq f_\epsilon(1 - \mu_{a^*})$  if and only if  $\mu_a \in [g_\epsilon^+(\mu_{a^*}), \mu_{a^*})$ , hence we have  $\text{kl}(\mu_{a^*}, \mu_a) =$   
2101  $d_\epsilon^-(\mu_{a^*}, \mu_a)$ . This concludes the proof in the first case.

2102 When  $\mu_{a^*} - \mu_a > f_\epsilon(1 - \mu_{a^*})$ , we obtain

$$\lim_{x \rightarrow +\infty} G_a(x) = \lim_{x \rightarrow +\infty} \{ -\log(1 + u_{1,*}(\mu_a, x)(e^\epsilon - 1)) + \epsilon \mu_{a^*} + x \text{kl}(\mu_a, u_{1,*}(\mu_a, x)) \},$$

2103 where

$$\begin{aligned} u_{1,*}(\mu_a, x) &= \\ &= \frac{\sqrt{(x(1 - \mu_a(e^\epsilon - 1)) - (e^\epsilon - 1))^2 + 4(1 + x)(e^\epsilon - 1)x\mu_a - (x(1 - \mu_a(e^\epsilon - 1)) - (e^\epsilon - 1))}}{2(1 + x)(e^\epsilon - 1)}. \end{aligned}$$

2104 Direct manipulation yields that  $\lim_{x \rightarrow +\infty} u_{1,*}(\mu_a, x) = \mu_a$ , hence

$$\lim_{x \rightarrow +\infty} G_a(x) = -\log(1 + \mu_a(e^\epsilon - 1)) + \epsilon \mu_{a^*} + \lim_{x \rightarrow +\infty} \{ x \text{kl}(\mu_a, u_{1,*}(\mu_a, x)) \}.$$

2105 Let us denote  $v_{1,*}(\mu_a, x) = u_{1,*}(\mu_a, x) - \mu_a \geq 0$ , i.e.,  $\lim_{x \rightarrow +\infty} v_{1,*}(\mu_a, x) = 0$ . Direct manipula-  
2106 tion yields that

$$\begin{aligned} &v_{1,*}(\mu_a, x) \\ &= \frac{1 + \mu_a(e^\epsilon - 1)}{2(e^\epsilon - 1)} \left( 1 - \frac{1}{x + 1} \right) \\ &\quad \left( \sqrt{1 - \frac{2x(1 - \mu_a(e^\epsilon + 1))(e^\epsilon - 1) - (e^\epsilon - 1)^2}{x^2(1 + \mu_a(e^\epsilon - 1))^2}} - 1 + \frac{(e^\epsilon - 1)(1 - 2\mu_a)}{x(1 + \mu_a(e^\epsilon - 1))} \right) \\ &= \sqrt{1 - \frac{2x(1 - \mu_a(e^\epsilon + 1))(e^\epsilon - 1) - (e^\epsilon - 1)^2}{x^2(1 + \mu_a(e^\epsilon - 1))^2}} - 1 + \frac{(e^\epsilon - 1)(1 - 2\mu_a)}{x(1 + \mu_a(e^\epsilon - 1))} \\ &=_{x \rightarrow +\infty} \frac{(e^\epsilon - 1)(1 - 2\mu_a)}{x(1 + \mu_a(e^\epsilon - 1))} - \frac{2x(1 - \mu_a(e^\epsilon + 1))(e^\epsilon - 1) - (e^\epsilon - 1)^2}{2x^2(1 + \mu_a(e^\epsilon - 1))^2} + \mathcal{O}(1/x^2) \\ &=_{x \rightarrow +\infty} \frac{2x(e^\epsilon - 1)2(1 - \mu_a)\mu_a(e^\epsilon - 1) + (e^\epsilon - 1)^2}{2x^2(1 + \mu_a(e^\epsilon - 1))^2} + \mathcal{O}(1/x^2), \\ &\text{hence } v_{1,*}(\mu_a, x) =_{x \rightarrow +\infty} \frac{2(1 - \mu_a)\mu_a(e^\epsilon - 1)^2}{x(1 + \mu_a(e^\epsilon - 1))^2} + \mathcal{O}(1/x^2). \end{aligned}$$

2107 where we used that  $\sqrt{1 - x} - 1 =_{x \rightarrow 0} -x/2 + \mathcal{O}(x^2)$  to obtain the last result. Similarly as before,  
2108 we derive

$$\begin{aligned} x \text{kl}(\mu_a, u_{1,*}(\mu_a, x)) &= x \left( \mu_a \log \left( 1 - \frac{1}{\mu_a(1 - \mu_a)} \frac{v_{1,*}(\mu_a, x)}{1 + v_{1,*}(\mu_a, x)/\mu_a} \right) \right. \\ &\quad \left. + \log \left( 1 + \frac{1}{1 - \mu_a} \frac{v_{1,*}(\mu_a, x)}{1 - v_{1,*}(\mu_a, x)/(1 - \mu_a)} \right) \right) \end{aligned}$$

$$\begin{aligned} \lim_{x \rightarrow +\infty} \{x \text{kl}(\mu_a, \mu_{a^*}^x, a)\} &= \frac{1}{\mu_a(1 - \mu_a)^2} \lim_{x \rightarrow +\infty} \left\{ \frac{x v_{1,*}(\mu_a, x)^2}{\left(1 - \frac{v_{1,*}(\mu_a, x)}{1 - \mu_a}\right) \left(1 + \frac{v_{1,*}(\mu_a, x)}{\mu_a}\right)} \right\} \\ &= \frac{\lim_{x \rightarrow +\infty} x v_{1,*}(\mu_a, x)^2}{\mu_a(1 - \mu_a)^2} = 0, \end{aligned}$$

where we used that  $v_{1,*}(\mu_a, x) =_{x \rightarrow +\infty} \mathcal{O}(1/x)$  to conclude. Therefore, we have shown that  $\lim_{x \rightarrow +\infty} G_a(x) = -\log(1 + \mu_a(e^\epsilon - 1)) + \epsilon \mu_{a^*}$ . Using Lemma 25 and the proof of Lemma 44, we know that  $\mu_{a^*} - \mu_a > f_\epsilon(1 - \mu_{a^*})$  if and only if  $\mu_a \in [0, g_\epsilon^+(\mu_{a^*})]$ , hence we have  $-\log(1 + \mu_a(e^\epsilon - 1)) + \epsilon \mu_{a^*} = d_\epsilon^-(\mu_{a^*}, \mu_a)$ . This concludes the proof in the second case.

Therefore,  $G_a$  is a strictly increasing one-to-one mapping from  $[0, +\infty)$  to  $[0, d_\epsilon^-(\mu_{a^*}, \mu_a))$ . Using the implicit function theorem, we obtain

$$\forall x \in [0, +\infty), \quad u'_a(x) = -\frac{\frac{\partial d_\epsilon^+}{\partial u}(\mu_a, u_a(x))}{\frac{\partial^2 d_\epsilon^-}{\partial u^2}(\mu_{a^*}, u_a(x)) + x \frac{\partial^2 d_\epsilon^+}{\partial u^2}(\mu_a, u_a(x))} < 0,$$

where the strict inequality is obtained by using properties in Lemmas 24 and 25, since  $u_a(x) \in (\mu_a, \mu_{a^*})$  by Lemmas 34 and 24. Similarly, we obtain

$$\forall x \in [0, +\infty), \quad G''_a(x) = u'_a(x) \frac{\partial d_\epsilon^+}{\partial u}(\mu_a, u_a(x)) > 0,$$

Therefore, we have shown that  $G_a$  is strictly concave and that  $u_a$  is decreasing.

Let us define  $x_a(y)$  as the unique solution of  $G_a(x) = y$ , which is well-defined based on our above computations. Therefore, we have

$$y = d_\epsilon^-(\mu_{a^*}, u_a(x_a(y))) + x_a(y) d_\epsilon^+(\mu_a, u_a(x_a(y))).$$

Using the derivative of the inverse function, we obtain

$$\forall y \in [0, d_\epsilon^-(\mu_{a^*}, \mu_a)), \quad x'_a(y) = \frac{1}{G'_a(x_a(y))} = \frac{1}{d_\epsilon^+(\mu_a, u_a(x_a(y)))} > 0,$$

hence  $x_a$  is increasing on  $[0, d_\epsilon^-(\mu_{a^*}, \mu_a))$ . Moreover, we have

$$\forall y \in [0, d_\epsilon^-(\mu_{a^*}, \mu_a)), \quad x''_a(y) = -\frac{u'_a(x_a(y))}{d_\epsilon^+(\mu_a, u_a(x_a(y)))^3} \frac{\partial d_\epsilon^+}{\partial u}(\mu_a, u_a(x_a(y))) > 0,$$

hence  $x_a$  is strictly convex on  $[0, d_\epsilon^-(\mu_{a^*}, \mu_a))$ .

For all  $y \in [0, \min_{a \neq a^*} d_\epsilon^-(\mu_{a^*}, \mu_a))$ , let us define  $G(y) = \frac{y}{1 + \sum_{a \neq a^*} x_a(y)}$  and  $F(y) = \sum_{a \neq a^*} \frac{d_\epsilon^-(\mu_{a^*}, u_a(x_a(y)))}{d_\epsilon^+(\mu_a, u_a(x_a(y)))}$ . Using the above results, direct manipulations yield that, for all  $y \in [0, \min_{a \neq a^*} d_\epsilon^-(\mu_{a^*}, \mu_a))$ ,

$$\begin{aligned} G'(y) &= \frac{1 + \sum_{a \neq a^*} x_a(y) - y \sum_{a \neq a^*} x'_a(y)}{(1 + \sum_{a \neq a^*} x_a(y))^2} = \frac{1 + \sum_{a \neq a^*} x_a(y) - \sum_{a \neq a^*} \frac{y}{d_\epsilon^+(\mu_a, u_a(x_a(y)))}}{(1 + \sum_{a \neq a^*} x_a(y))^2} \\ &= \frac{1 - F(y)}{(1 + \sum_{a \neq a^*} x_a(y))^2}, \end{aligned}$$

hence we obtain that  $G'(y) = 0$  if and only if  $F(y) = 1$ . Using that  $x_a(0) = 0$ ,  $u_a(0) = \mu_{a^*}$  and  $d_\epsilon^-(\mu_{a^*}, \mu_{a^*}) = 0$ , we obtain that  $F(0) = 0$ .

Using that  $\lim_{y \rightarrow d_\epsilon^-(\mu_{a^*}, \mu_a)} x_a(y) = +\infty$ ,  $\lim_{x \rightarrow +\infty} u_a(x) = \mu_a$ ,  $d_\epsilon^-(\mu_{a^*}, \mu_a) > 0$  and  $d_\epsilon^+(\mu_a, \mu_a) = 0$ , we obtain that  $\lim_{y \rightarrow \min_{a \neq a^*} d_\epsilon^-(\mu_{a^*}, \mu_a)} F(y) = +\infty$ .

Let  $H(y) = \sum_{a \neq a^*} \frac{1}{d_\epsilon^+(\mu_a, u_a(x_a(y)))}$  for all  $y \in [0, \min_{a \neq a^*} d_\epsilon^-(\mu_{a^*}, \mu_a))$ . Then, we have

$$\sum_{a \neq a^*} x_a(y) = y H(y) - F(y), \quad \sum_{a \neq a^*} x'_a(y) = H(y),$$



$$\begin{aligned} \frac{\partial d_\epsilon^-}{\partial u}(\mu_{a^*}, u_a(x)) + x \frac{\partial d_\epsilon^+}{\partial u}(\mu_a, u_a(x)) &= 0, \\ d_\epsilon^-(\mu_{a^*}, u_a(x_a(y))) + x_a(y) d_\epsilon^+(\mu_a, u_a(x_a(y))) &= y. \end{aligned}$$

2131 Then, for all  $y \in [0, \min_{a \neq a^*} d_\epsilon^-(\mu_{a^*}, \mu_a))$ , we have

$$\begin{aligned} H'(y) &= - \sum_{a \neq a^*} \frac{u'_a(x_a(y)) x'_a(y)}{(d_\epsilon^+(\mu_a, u_a(x_a(y))))^2} \frac{\partial d_\epsilon^+}{\partial u}(\mu_a, u_a(x_a(y))) \\ &= - \sum_{a \neq a^*} \frac{u'_a(x_a(y))}{(d_\epsilon^+(\mu_a, u_a(x_a(y))))^3} \frac{\partial d_\epsilon^+}{\partial u}(\mu_a, u_a(x_a(y))), \\ F'(y) &= \sum_{a \neq a^*} \frac{u'_a(x_a(y)) x'_a(y)}{(d_\epsilon^+(\mu_a, u_a(x_a(y))))^2} \\ &\quad \left( d_\epsilon^+(\mu_a, u_a(x_a(y))) \frac{\partial d_\epsilon^-}{\partial u}(\mu_{a^*}, u_a(x_a(y))) - d_\epsilon^-(\mu_{a^*}, u_a(x_a(y))) \frac{\partial d_\epsilon^+}{\partial u}(\mu_a, u_a(x_a(y))) \right) \\ &= - \sum_{a \neq a^*} \frac{u'_a(x_a(y)) x'_a(y)}{(d_\epsilon^+(\mu_a, u_a(x_a(y))))^2} \frac{\partial d_\epsilon^+}{\partial u}(\mu_a, u_a(x_a(y))) \\ &\quad (x_a(y) d_\epsilon^+(\mu_a, u_a(x_a(y))) + d_\epsilon^-(\mu_{a^*}, u_a(x_a(y)))) \\ &= -y \sum_{a \neq a^*} \frac{u'_a(x_a(y)) x'_a(y)}{(d_\epsilon^+(\mu_a, u_a(x_a(y))))^2} \frac{\partial d_\epsilon^+}{\partial u}(\mu_a, u_a(x_a(y))) = y H'(y), \end{aligned}$$

2132 Therefore, showing that  $H$  is increasing is a sufficient condition to show that  $F$  is increasing. Using  
2133 the above results, we have, for all  $a \neq a^*$  and all  $y \in [0, \min_{a \neq a^*} d_\epsilon^-(\mu_{a^*}, \mu_a))$ , we have

$$\frac{1}{(d_\epsilon^+(\mu_a, u_a(x_a(y))))^3} \frac{\partial d_\epsilon^+}{\partial u}(\mu_a, u_a(x_a(y))) > 0 \quad \text{and} \quad u'_a(x_a(y)) < 0.$$

2134 Therefore,  $H$  is increasing as a summation of increasing function, hence  $F$  is increasing.

2135 Let  $y^*$  such that  $F(y^*) = 1$ . Reusing the above manipulation, we obtain

$$\begin{aligned} G''(y) &= - \frac{F'(y)(1 + \sum_{a \neq a^*} x_a(y)) + 2(1 - F(y)) \sum_{a \neq a^*} x'_a(y)}{(1 + \sum_{a \neq a^*} x_a(y))^3} \\ &= - \frac{y H'(y)(1 + y H(y) - F(y)) + 2(1 - F(y)) H(y)}{(1 + y H(y) - F(y))^3}, \\ G''(y^*) &= - \frac{H'(y^*)}{y^* H(y^*)^2} < 0, \end{aligned}$$

2136 Therefore,  $y^*$  is the unique maximum of  $G$ . We conclude this part of the proof by using the  
2137 intermediate results in the proof of Lemma 42.

2138 By strict convexity of  $x_a$  and using its properties proven above, we obtain

$$x_a(y) \geq x_a(0) + y x'_a(0) = \frac{y}{d_\epsilon^+(\mu_a, \mu_{a^*})}.$$

2139 Summing those inequalities, we obtain

$$\forall y \in [0, \min_{a \neq a^*} d_\epsilon^-(\mu_{a^*}, \mu_a)), \quad G(y) = \frac{y}{1 + \sum_{a \neq a^*} x_a(y)} \leq \frac{1}{\frac{1}{y} + \sum_{a \neq a^*} \frac{1}{d_\epsilon^+(\mu_a, \mu_{a^*})}}.$$

2140 Using that  $y \mapsto 1/(1/y + \alpha)$  is increasing for  $\alpha > 0$ , we obtain that

$$T_\epsilon^*(\nu)^{-1} = \max_{y \in [0, \min_{a \neq a^*} d_\epsilon^-(\mu_{a^*}, \mu_a))} G(y) \leq \frac{1}{\frac{1}{\min_{a \neq a^*} d_\epsilon^-(\mu_{a^*}, \mu_a)} + \sum_{a \neq a^*} \frac{1}{d_\epsilon^+(\mu_a, \mu_{a^*})}}.$$

2141 This concludes the proof of the second to last result. The last result is obtained by combining  
2142 Lemmas 25 and 24 and the derivation in the proof of Lemma 44.  $\square$

2143 Lemma 47 is a technical result used in the proof of sufficient exploration of our sampling rule.

2144 **Lemma 47.** *Let  $d_\epsilon^\pm$  as in Eq. (3). Let  $\mu \in (0, 1)^K$ . There exists  $\alpha > 0$  such that*

$$C_\mu := \min_{(a,b): \mu_a > \mu_b} \inf_{\lambda_a, \lambda_b: \max_{c \in \{a,b\}} |\mu_c - \lambda_c| \leq \alpha} \inf_{u \in [0,1]} \{d_\epsilon^-(\lambda_a, u) + d_\epsilon^+(\lambda_b, u)\} > 0. \quad (38)$$

2145 *Proof.* Using Lemma 34 for  $w_1 = w_2 = 1$ , the function  $\mu \mapsto \inf_{u \in [0,1]} \{d_\epsilon^-(\mu_a, u) + d_\epsilon^+(\mu_b, u)\}$  is  
 2146 continuous on  $\mathcal{F}^K$ . Since it has strictly positive values when  $\mu_a > \mu_b$  (Lemma 34), there exists  $\alpha$   
 2147 such that

$$\inf_{\lambda_a, \lambda_b: \max_{c \in \{a,b\}} |\mu_c - \lambda_c| \leq \alpha} \inf_{u \in [0,1]} \{d_\epsilon^-(\lambda_a, u) + d_\epsilon^+(\lambda_b, u)\} > 0.$$

2148 Further lower bounding by a finite number of strictly positive constants yields the result.  $\square$

2149 Lemma 47 is a technical result used in the proof of convergence towards the optimal allocation of our  
 2150 sampling rule.

2151 **Lemma 48.** *Let  $d_\epsilon^\pm$  as in Eq. (3). Let  $(\phi_1, \phi_2) \in (0, 1)^2$ . Let  $\mathcal{I}_a := \{\mu \in (0, 1)^K \mid a \in a^*(\mu)\}$  for  
 2152 all  $a \in [K]$ . For all  $a^* \in [K]$ , all  $\mu \in \mathcal{I}_{a^*}$ , all  $(a, b) \in ([K] \setminus \{a^*\})^2$  such that  $a \neq b$ , and all  
 2153  $\beta \in [0, 1]$ , define*

$$G_{a,b}(\mu, \beta) := \inf_{u \in [0,1]} \{\beta d_\epsilon^-(\mu_{a^*}, u) + \phi_1 d_\epsilon^+(\mu_a, u)\} - \inf_{u \in (0,1)} \{\beta d_\epsilon^-(\mu_{a^*}, u) + \phi_2 d_\epsilon^+(\mu_b, u)\}.$$

2154 *The function  $(\mu, \beta) \mapsto G_{a,b}(\mu, \beta)$  is continuous on  $(0, 1)^K \times [0, 1]$ . For all  $\xi > 0$ , the function*  
 2155  *$(\mu, \beta) \mapsto \inf_{\tilde{\beta}: |\beta - \tilde{\beta}| \leq \xi} G_{a,b}(\mu, \beta)$  is continuous on  $(0, 1)^K$ .*

2156 *Proof.* Since  $\bigcup_{a \in [K]} \mathcal{I}_a^K = (0, 1)^K$ , it is enough to show the property for all  $a \in [K]$ . Let  $a^* \in [K]$ ,  
 2157  $\mu \in \mathcal{I}_{a^*}$ ,  $(a, b) \in ([K] \setminus \{a^*\})^2$  such that  $a \neq b$ . As done in Lemma 41 by using Lemma 34,  
 2158 we obtain that the function  $(\mu, \beta) \mapsto G_{a,b}(\mu, \beta)$  is continuous on  $\mathcal{I}_{a^*} \times [0, 1]$  for all  $a^* \in [K]$ ,  
 2159 hence on  $(0, 1)^K \times [0, 1]$ . Let  $\Phi: \mu \mapsto \{\tilde{\beta}: |\beta - \tilde{\beta}| \leq \xi\}$ , it is a continuous (constant), compact  
 2160 valued and non-empty correspondence. Using the above continuity, Berge's theorem yields that  
 2161  $\mu \mapsto \inf_{\tilde{\beta}: |\beta - \tilde{\beta}| \leq \xi} G_{a,b}(\mu, \tilde{\beta})$  is continuous on  $(0, 1)^K$ .  $\square$

## 2162 H Asymptotic Upper Bound on the Expected Sample complexity

2163 Let  $\nu$  be a Bernoulli instance with means  $\mu \in (0, 1)^2$  and unique best arm  $a^* \in [K]$ , i.e.,  
 2164  $\arg \max_{a \in [K]} \mu_a = \{a^*\}$ . Let  $\beta \in (0, 1)$ . Let  $w_{\epsilon, \beta}^*(\nu) = \{w_{\epsilon, \beta}^*\}$  be the unique  $\beta$ -optimal al-  
 2165 location defined in Eq. (35), which satisfies  $\min_{a \in [K]} w_{\epsilon, \beta, a}^* > 0$  by Lemma 42. At equilibrium, we  
 2166 have equality of the transportation costs by Lemma 42, namely

$$\forall a \neq a^*, \quad W_{\epsilon, a^*, a}(\mu, w_{\epsilon, \beta}^*) = T_{\epsilon, \beta}^*(\nu)^{-1}, \quad (39)$$

2167 where  $W_{\epsilon, a, b}$  is defined in Eq. (4) and  $T_{\epsilon, \beta}^*$  is defined in Eq. (35).

2168 Let  $\gamma > 0$ . Let  $\omega \in \Delta_K$  be any allocation over arms such that  $\min_a \omega_a > 0$ . We denote by  $T_\gamma(\omega)$  the  
 2169 convergence time towards  $\omega$ , which is a random variable quantifying the number of samples required  
 2170 for the global empirical allocations  $N_n/(n-1)$  to be  $\gamma$ -close to  $\omega$  for any subsequent time, namely

$$T_\gamma(\omega) := \inf \left\{ T \geq 1 \mid \forall n \geq T, \left\| \frac{N_n}{n-1} - \omega \right\|_\infty \leq \gamma \right\}. \quad (40)$$

2171 The proof of Theorem 6 follows the same analysis as the unified analysis of Top Two algorithms, see,  
 2172 e.g., Jourdan et al. [50]. Appendix H is organised as follows. After recalling some technical results  
 2173 (Appendix H.1), we prove sufficient exploration of our sampling rule (Appendix H.2). Second, we  
 2174 prove that convergence time towards the  $\beta$ -optimal allocation of our sampling rule (Appendix H.3)  
 2175 has finite expectation. Finally, we conclude the proof of Theorems 6 (Appendix H.4).

## 2176 H.1 Technical Results from the Literature

2177 Lemma 49 relates the global counts  $(N_{n,a})_{a \in [K]}$  and the local counts  $(\tilde{N}_{n,a})_{a \in [K]}$ .

2178 **Lemma 49.** *Let  $\eta > 0$  be the geometric parameter used for the geometric grid update of our private*  
 2179 *empirical mean estimator. For all  $(a, k) \in [K] \times \mathbb{N}$  s.t.  $\mathbb{E}_{\nu\pi}[T_k(a)] < +\infty$ ,  $N_{T_k(a),a} = \tilde{N}_{k,a} =$*   
 2180  *$\lceil (1 + \eta)^{k-1} \rceil$ . For all  $a \in [K]$  and all  $n \in \mathbb{N}$ ,  $N_{n,a} \geq \tilde{N}_{n,a} \geq N_{n,a}/(1 + \eta)$ .*

2181 *Proof.* Let  $a \in [K]$ . After initialisation, we have  $k = 1$ ,  $T_1(a) = K + 1$  and  $N_{T_1(a),a} = 1$ .  
 2182 Using the definition of the phase switch, it is direct to see that  $N_{T_2(a),a} = \tilde{N}_{2,a} = \lceil 1 + \eta \rceil$  when  
 2183  $\mathbb{E}_{\nu\pi}[T_2(a)] < +\infty$ . Similarly, we obtain  $N_{T_k(a),a} = \tilde{N}_{k,a} = \lceil (1 + \eta)^{k-1} \rceil$  when  $\mathbb{E}_{\nu\pi}[T_k(a)] < +\infty$ .  
 2184 The last result is a direct consequence of the definition of the per-arm geometric update grid.  $\square$

2185 Lemma 50 controls the deviation  $N_{n,a}^a - \beta L_{n,a}$  enforced by the tracking procedure.

2186 **Lemma 50** (Lemma 2.2 in [49]). *For all  $n > K$  and all  $a \in [K]$ ,  $-1/2 \leq N_{n,a}^a - \beta L_{n,a} \leq 1$ .*

2187 Lemma 51 gathers properties on the  $\bar{W}_{-1}$  function used in the stopping threshold.

2188 **Lemma 51** ([51]). *Let  $\bar{W}_{-1}(x) := -W_{-1}(-e^{-x})$  for all  $x \geq 1$ , where  $W_{-1}$  is the negative branch*  
 2189 *of the Lambert  $W$  function. The function  $\bar{W}_{-1}$  is increasing on  $(1, +\infty)$  and strictly concave on*  
 2190  *$(1, +\infty)$ . In particular,  $\bar{W}_{-1}'(x) = \left(1 - \frac{1}{\bar{W}_{-1}(x)}\right)^{-1}$  for all  $x > 1$ . Then, for all  $y \geq 1$  and  $x \geq 1$ ,*

$$\bar{W}_{-1}(y) \leq x \iff y \leq x - \log(x).$$

2191 *Moreover, for all  $x > 1$ ,*

$$x + \log(x) \leq \bar{W}_{-1}(x) \leq x + \log(x) + \min \left\{ \frac{1}{2}, \frac{1}{\sqrt{x}} \right\}.$$

2192 Lemma 52 gives an upper bound on a time define implicit as a function of  $\bar{W}_{-1}$ , namely it is an  
 2193 inversion result.

2194 **Lemma 52** (Lemma 32 in [12]). *Let  $\bar{W}_{-1}$  defined in Lemma 51. Let  $A > 0$ ,  $B > 0$  such that*  
 2195  *$B/A + \log A > 1$  and  $C(A, B) = \sup \{x \mid x < A \log x + B\}$ . Then,  $C(A, B) < h_1(A, B)$  with*  
 2196  *$h_1(z, y) = z \bar{W}_{-1}(y/z + \log z)$ .*

2197 Lemma 53 shows that upon correction the supremum of sub-exponential random variables is also a  
 2198 sub-exponential random variable.

2199 **Lemma 53** (Lemma 72 in [50]). *Suppose that  $(X_n)_{n \geq 1}$  are sub-exponential random variables with*  
 2200 *constants  $(C_n)$ , such that  $c := \inf_n C_n > 0$ . Then  $\sup_n (X_n / \log(e + n))$  is sub-exponential.*

2201 Lemma 54 gives a coarse convergence rate of the private empirical estimators of the means towards  
 2202 their true means.

2203 **Lemma 54.** *There exist sub-exponential random variable  $W_\mu$  such that almost surely, for all  $a \in [K]$*   
 2204 *and all  $n$  such that  $\tilde{N}_{n,a} \geq 1$ ,*

$$\tilde{N}_{n,a} |\tilde{\mu}_{n,a} - \mu_a| \leq W_\mu \log(e + \tilde{N}_{n,a}).$$

2205 *In particular, any random variable which is polynomial in  $W_\mu$  has a finite expectation.*

2206 *Proof.* Let us define

$$W_\mu = \max_{a \in [K]} \sup_{n \in \mathbb{N}} \frac{\tilde{N}_{n,a} |\tilde{\mu}_{n,a} - \mu_a|}{\log(e + \tilde{N}_{n,a})}.$$

2207 Let  $a \in [K]$ . Let us define the geometric grid  $N_k = \lceil (1 + \eta)^{k-1} \rceil$  for all  $k \in \mathbb{N}$ , on which we  
 2208 effectively need to control the concentration. The maximum of a finite number of sub-exponential  
 2209 random variables is sub-exponential. Therefore, using the geometric update grid, it suffices to show  
 2210 that

$$\sup_{k \in \mathbb{N}} \frac{N_k |(Z_{N_k} + S_k)/N_k - \mu_a|}{\log(e + N_k)}$$

2211 is sub-exponential, where  $Z_{N_k}$  is the cumulative sum of  $N_k$  i.i.d. observations from  $\text{Ber}(\mu_a)$  and  $S_k$   
 2212 is the cumulative sum of  $k$  i.i.d. observations from  $\text{Lap}(1/\epsilon)$ .

2213 Using that  $Z_{N_k} - N_k\mu_a$  is sub-Gaussian and  $S_k$  is sub-exponential, for a fixed  $k$ ,  $|Z_{N_k} - N_k\mu_a + S_k|$   
 2214 is sub-exponential. Applying Lemma 53, we obtain that

$$\sup_{k \in \mathbb{N}} \frac{N_k |(Z_{N_k} + S_k)/N_k - \mu_a|}{\log(e + N_k)}$$

2215 is sub-exponential. We finally obtain that the maximum over the finitely many arms has the same  
 2216 property.  $\square$

## 2217 H.2 Sufficient Exploration

2218 The first step of in the generic analysis of Top Two algorithms [50] consists in showing sufficient  
 2219 exploration. The main idea is that, if there are still undersampled arms, either the leader or the  
 2220 challenger will be among them. Therefore, after a long enough time, no arm can still be undersampled.  
 2221 We emphasise that there are multiple ways to select the leader/challenger pair in order to ensure  
 2222 sufficient exploration. Therefore, other choices of leader/challenger pair would yield similar results.

2223 Given an arbitrary phase  $p \in \mathbb{N}$ , we define the sampled enough set, i.e., the arms having reached  
 2224 phase  $p$ , and the arm with highest mean in this set (when not empty) as

$$S_n^p = \{a \in [K] \mid N_{n,a} \geq (1 + \eta)^{p-1}\} \quad \text{and} \quad a_n^* = \arg \max_{a \in S_n^p} \mu_a. \quad (41)$$

2225 Since  $\min_{a \neq b} |\mu_a - \mu_b| > 0$ ,  $a_n^*$  is unique. Let  $p \in \mathbb{N}$  such that  $(p-1)/4 \in \mathbb{N}$ . We define the highly  
 2226 and the mildly under-sampled sets as

$$U_n^p := \{a \in [K] \mid N_{n,a} < (1 + \eta)^{(p-1)/2}\} \quad \text{and} \quad V_n^p := \{a \in [K] \mid N_{n,a} < (1 + \eta)^{3(p-1)/4}\}. \quad (42)$$

2227 Those arms have not reached phase  $(p-1)/2$  and phase  $3(p-1)/4$ , respectively.

2228 Lemma 55 shows that, when the leader is sampled enough, it is the arm with highest true mean among  
 2229 the sampled enough arms.

2230 **Lemma 55.** *Let  $S_n^p$  and  $a_n^*$  as in (41). There exists  $p_0$  with  $\mathbb{E}_{\nu^\pi}[\exp(\alpha p_0)] < +\infty$  for all  $\alpha > 0$   
 2231 such that if  $p \geq p_0$ , for all  $n$  such that  $S_n^p \neq \emptyset$ , we have*

- 2232 • For all  $a \in S_n^p$ , we have  $\tilde{\mu}_{n,a} \in (0, 1)$  and  $a_n^* = \arg \max_{a \in S_n^p} \tilde{\mu}_{n,a}$ .
- 2233 • If  $B_n \in S_n^p$ , then  $B_n = a_n^*$ .

2234 *Proof.* Let  $p_0$  to be specified later. Let  $p \geq p_0$ . Let  $n \in \mathbb{N}$  such that  $S_n^p \neq \emptyset$ , where  $S_n^p$  and  $a_n^*$  as  
 2235 in Equation (41). Since  $N_{n,a} \geq (1 + \eta)^{p-1}$  for all  $a \in S_n^p$ , we have  $\tilde{N}_{n,a} \geq (1 + \eta)^{p-1}$ . Using  
 2236 Lemma 54 and  $x \rightarrow \log(e + x)/x$  is decreasing, we obtain that

$$\begin{aligned} \tilde{\mu}_{n,a_n^*} &\geq \mu_{a_n^*} - W_\mu \frac{\log(e + (1 + \eta)^{p-1})}{(1 + \eta)^{p-1}}, \\ \forall a \in S_n^p \setminus \{a_n^*\}, \quad \tilde{\mu}_{n,a} &\leq \mu_a + W_\mu \frac{\log(e + (1 + \eta)^{p-1})}{(1 + \eta)^{p-1}}. \end{aligned}$$

2237 Let  $\bar{\Delta}_{\min} = \min_{a \neq b} |\mu_a - \mu_b|$  and  $\Delta_0 = \min_{a \in [K]} \min\{\mu_a, 1 - \mu_a\} > 0$ . By assumption on the  
 2238 considered instances, we know that  $\bar{\Delta}_{\min} > 0$ . Let  $p_1 = \lceil \log_{1+\eta}(X_1 - e) \rceil + 1$  with

$$\begin{aligned} X_1 &= \sup \{x > 1 \mid x \leq 4(\min\{\bar{\Delta}_{\min}, \Delta_0\})^{-1} W_\mu \log x + e\} \\ &\leq h_1(4(\min\{\bar{\Delta}_{\min}, \Delta_0\})^{-1} W_\mu, e), \end{aligned}$$

2239 where we used Lemma 52, and  $h_1$  defined therein. Then, for all  $p \in \mathbb{N}$  such that  $p \geq p_1 + 1$  and all  
 2240  $n \in \mathbb{N}$  such that  $S_n^p \neq \emptyset$ , we have

$$\forall a \in S_n^p, \quad \mu_a - \min\{\bar{\Delta}_{\min}, \Delta_0\}/4 \leq \tilde{\mu}_{n,a} \leq \mu_a + \min\{\bar{\Delta}_{\min}, \Delta_0\}/4.$$

2241 Therefore, we have  $\tilde{\mu}_{n,a} \in (0, 1)$  for all  $a \in S_n^p$ . Since  $\tilde{\mu}_{n,a_n^*} \geq \mu_{a_n^*} - \min\{\bar{\Delta}_{\min}, \Delta_0\}/4$  and  
 2242  $\tilde{\mu}_{n,a} \leq \mu_a + \min\{\bar{\Delta}_{\min}, \Delta_0\}/4$  for all  $a \in S_n^p \setminus \{a_n^*\}$ , we obtain  $a_n^* = \arg \max_{a \in S_n^p} \tilde{\mu}_{n,a}$  since  
 2243  $\arg \max_{a \in S_n^p} \tilde{\mu}_{n,a}$  is unique. The leader is defined as  $B_n = \arg \max_{a \in [K]} [\tilde{\mu}_{n,a}]_0^1$ . If  $B_n \in S_n^p$ , we  
 2244 obtain

$$B_n = \arg \max_{a \in S_n^p} [\tilde{\mu}_{n,a}]_0^1 = \arg \max_{a \in S_n^p} \tilde{\mu}_{n,a} = a_n^* .$$

2245 For all  $\alpha \in \mathbb{R}_+$ , we have  $\exp(\alpha p_1) \leq e^{3\alpha} (X_1 - e)^{\alpha/\log 2}$ , hence  $\mathbb{E}_{\nu\pi}[\exp(\alpha p_1)] < +\infty$  by using  
 2246 Lemma 54 and  $h_1(x, e) \sim_{x \rightarrow +\infty} x \log x$  to obtain that  $\exp(\alpha p_1)$  is at most polynomial in  $W_\mu$ .  
 2247 Taking  $p_0 = p_1$  concludes the proof.  $\square$

2248 Lemma 56 shows that the transportation costs between the sampled enough arms with largest true  
 2249 means and the other sampled enough arms are increasing fast enough.

2250 **Lemma 56.** *Let  $S_n^p$  as in Eq. (41). There exists  $p_1$  with  $\mathbb{E}_{\nu\pi}[\exp(\alpha p_1)] < +\infty$  for all  $\alpha > 0$  such  
 2251 that if  $p \geq p_1$ , for all  $n$  such that  $S_n^p \neq \emptyset$ , for all  $(a, b) \in (S_n^p)^2$  such that  $\mu_a > \mu_b$ , we have*

$$W_{\epsilon,a,b}(\tilde{\mu}_n, N_n) \geq (1 + \eta)^{p-1} C_\mu ,$$

2252 where  $C_\mu > 0$  is a problem dependent constant.

2253 *Proof.* Let  $p_0$  as in Lemma 55. Let  $p \geq p_0$ . Let  $n \in \mathbb{N}$  such that  $S_n^p \neq \emptyset$ , where  $S_n^p$  as in Eq. (41).  
 2254 Since  $N_{n,a} \geq (1 + \eta)^{p-1}$  for all  $a \in S_n^p$ , we have  $\tilde{N}_{n,a} \geq (1 + \eta)^{p-1}$  by using Lemma 49. Let  
 2255  $(a, b) \in (S_n^p)^2$  such that  $\mu_a > \mu_b$ . Using Lemma 47, there exists  $\alpha_\mu > 0$  such that

$$C_\mu = \min_{(a,b): \mu_a > \mu_b} \inf_{\lambda_a, \lambda_b: \max_{c \in \{a,b\}} |\mu_c - \lambda_c| \leq \alpha_\mu} \inf_{u \in [0,1]} \{d_\epsilon^-(\lambda_a, u) + d_\epsilon^+(\lambda_b, u)\} > 0 .$$

2256 Let  $\eta > 0$  s.t.  $\eta < \frac{1}{4} \min\{\bar{\Delta}_{\min}, \Delta_0, \alpha_\mu\}$  where  $\bar{\Delta}_{\min} = \min_{a \neq b} |\mu_a - \mu_b|$  and  $\Delta_0 =$   
 2257  $\min_{a \in [K]} \min\{\mu_a, 1 - \mu_a\}$ . Similarly as in the proof of Lemma 55, we can construct  $p_2$  with  
 2258  $\mathbb{E}_{\nu\pi}[\exp(\alpha p_2)] < +\infty$  for all  $\alpha > 0$  such that if  $p \geq p_2$ , for all  $n$  such that  $S_n^p \neq \emptyset$ , we have  
 2259  $|\tilde{\mu}_{n,a} - \mu_a| \leq \eta$  for all  $a \in S_n^p$ . Therefore, we have  $\tilde{\mu}_{n,a} = [\tilde{\mu}_{n,a}]_0^1$  and  $[\tilde{\mu}_{n,b}]_0^1 = \tilde{\mu}_{n,b}$ . Moreover,  
 2260 we have  $\tilde{\mu}_{n,a} \geq \mu_a - \eta > \mu_b + \eta \geq \tilde{\mu}_{n,b}$ . Then, we obtain

$$\begin{aligned} W_{\epsilon,a,b}(\tilde{\mu}_n, N_n) &= \inf_{u \in [0,1]} \{N_{n,a} d_\epsilon^-(\tilde{\mu}_{n,a}, u) + N_{n,b} d_\epsilon^+(\tilde{\mu}_{n,b}, u)\} \\ &\geq (1 + \eta)^{p-1} \inf_{u \in [0,1]} \{d_\epsilon^-(\tilde{\mu}_{n,a}, u) + d_\epsilon^+(\tilde{\mu}_{n,b}, u)\} \\ &\geq (1 + \eta)^{p-1} \inf_{\lambda_a, \lambda_b: \max_{c \in \{a,b\}} |\mu_c - \lambda_c| \leq \alpha_\mu} \inf_{u \in [0,1]} \{d_\epsilon^-(\lambda_a, u) + d_\epsilon^+(\lambda_b, u)\} \geq (1 + \eta)^{p-1} C_\mu . \end{aligned}$$

2261 This concludes the proof.  $\square$

2262 Lemma 57 shows that the transportation costs between sampled enough arms and undersampled arms  
 2263 are not increasing too fast.

2264 **Lemma 57.** *Let  $S_n^p$  be as in Eq. (41). There exists  $p_2$  with  $\mathbb{E}_{\nu\pi}[\exp(\alpha p_2)] < +\infty$  for all  $\alpha > 0$  such  
 2265 that if  $p \geq p_2$ , for all  $n$  such that  $S_n^p \neq \emptyset$ , For all  $p \geq p_2$  and all  $n$  such that  $S_n^p \neq \emptyset$ , for all  $a \in S_n^p$   
 2266 and  $b \notin S_n^p$ ,*

$$W_{\epsilon,a,b}(\tilde{\mu}_n, N_n) \leq (1 + \eta)^{p-1} D_\mu ,$$

2267 where  $D_\mu \in (0, +\infty)$  is a problem dependent constant.

2268 *Proof.* Let  $n \in \mathbb{N}$  such that  $S_n^p \neq \emptyset$ , where  $S_n^p$  as in Eq. (41). Since  $N_{n,a} \geq (1 + \eta)^{p-1}$  for all  
 2269  $a \in S_n^p$ , we have  $\tilde{N}_{n,a} \geq (1 + \eta)^{p-1}$  by using Lemma 49. Likewise,  $N_{n,b} < (1 + \eta)^{p-1}$  for all  
 2270  $b \notin S_n^p$ , we have  $\tilde{N}_{n,b} < (1 + \eta)^{p-1}$ . Let  $a \in S_n^p$  and  $b \notin S_n^p$ . Since the result is direct when  
 2271  $[\tilde{\mu}_{n,a}]_0^1 \leq [\tilde{\mu}_{n,b}]_0^1$ , we assume  $[\tilde{\mu}_{n,a}]_0^1 > [\tilde{\mu}_{n,b}]_0^1$  in the following.

2272 Let  $\eta > 0$  s.t.  $\eta < \frac{1}{4} \min\{\bar{\Delta}_{\min}, \Delta_0\}$  where  $\bar{\Delta}_{\min} = \min_{a \neq b} |\mu_a - \mu_b|$  and  $\Delta_0 =$   
 2273  $\min_{a \in [K]} \min\{\mu_a, 1 - \mu_a\} > 0$ . Similarly as in the proof of Lemma 55, we can construct  $p_2$

2274 with  $\mathbb{E}_{\nu\pi}[\exp(\alpha p_2)] < +\infty$  for all  $\alpha > 0$  such that if  $p \geq p_2$ , for all  $n$  such that  $S_n^p \neq \emptyset$ , we have  
 2275  $|\tilde{\mu}_{n,a} - \mu_a| \leq \eta$  for all  $a \in S_n^p$ . Let  $g_\epsilon^+(x) = \frac{x}{x(1-e^\epsilon)+e^\epsilon}$  as in Lemma 23. Using Lemma 23, for all  
 2276  $a \in S_n^p$ , we have

$$1 > \mu_a + \min\{\bar{\Delta}_{\min}, \Delta_0\}/4 \geq \tilde{\mu}_{n,a} > g_\epsilon^+(\tilde{\mu}_{n,a}) \geq g_\epsilon^+(\mu_a - \min\{\bar{\Delta}_{\min}, \Delta_0\}/4) > 0.$$

2277 Taking  $u = \tilde{\mu}_{n,a} \in [0, 1]$  and using that  $d_\epsilon^-(\tilde{\mu}_{n,a}, \tilde{\mu}_{n,a}) = 0$ , we obtain

$$\begin{aligned} W_{\epsilon,a,b}(\tilde{\mu}_n, N_n) &= \inf_{u \in [0,1]} \{N_{n,a}d_\epsilon^-(\tilde{\mu}_{n,a}, u) + N_{n,b}d_\epsilon^+(\tilde{\mu}_{n,b}, u)\} \\ &\leq N_{n,b}d_\epsilon^+(\tilde{\mu}_{n,b}, \tilde{\mu}_{n,a}) \leq (1+\eta)^{p-1}d_\epsilon^+(\tilde{\mu}_{n,b}, \tilde{\mu}_{n,a}), \end{aligned}$$

2278 where the last term is positive since  $\tilde{\mu}_{n,a} > [\tilde{\mu}_{n,b}]_0^1$  and  $\tilde{\mu}_{n,a} \in (0, 1)$  by Lemma 24.

2279 When  $\tilde{\mu}_{n,b} \leq 0$ , Lemma 24 yields that

$$d_\epsilon^+(\tilde{\mu}_{n,b}, \tilde{\mu}_{n,a}) = -\log(1 - \tilde{\mu}_{n,a}(1 - e^{-\epsilon})) \leq \epsilon,$$

2280 where we used that  $x \rightarrow -\log(1 - x(1 - e^{-\epsilon}))$  is increasing on  $(0, 1)$ . When  $\tilde{\mu}_{n,b} \in (0, g_\epsilon^+(\tilde{\mu}_{n,a}))$ ,

2281 Lemma 24 yields that

$$d_\epsilon^+(\tilde{\mu}_{n,b}, \tilde{\mu}_{n,a}) = -\log(1 - \tilde{\mu}_{n,a}(1 - e^{-\epsilon})) - \epsilon \tilde{\mu}_{n,b} \leq \epsilon.$$

2282 When  $\tilde{\mu}_{n,b} \in [g_\epsilon^+(\tilde{\mu}_{n,a}), \tilde{\mu}_{n,a})$ , Lemma 24 yields that

$$\begin{aligned} d_\epsilon^+(\tilde{\mu}_{n,b}, \tilde{\mu}_{n,a}) &= \text{kl}(\tilde{\mu}_{n,b}, \tilde{\mu}_{n,a}) \leq -\log \min\{\tilde{\mu}_{n,a}, 1 - \tilde{\mu}_{n,a}\} \\ &\leq -\log \min\{\mu_a - \min\{\bar{\Delta}_{\min}, \Delta_0\}/4, 1 - \mu_a - \min\{\bar{\Delta}_{\min}, \Delta_0\}/4\}, \end{aligned}$$

2283 where we used the classical result that  $\text{kl}(q, p) \leq -\log \min\{p, 1 - p\}$ . Let us define

$$D_\mu = \epsilon + \max_{a \in [K]} \{-\log \min\{\mu_a - \min\{\bar{\Delta}_{\min}, \Delta_0\}/4, 1 - \mu_a - \min\{\bar{\Delta}_{\min}, \Delta_0\}/4\}\}.$$

2284 Then, we have shown that  $d_\epsilon^+(\tilde{\mu}_{n,b}, \tilde{\mu}_{n,a}) \leq D_\mu$  where  $D_\mu \in (0, +\infty)$ . This yields the result.  $\square$

2285 Lemma 58 shows that the challenger is mildly undersampled if the leader is not mildly undersampled.

2286

2287 **Lemma 58.** Let  $V_n^p$  be as in Equation (42). There exists  $p_3$  with  $\mathbb{E}_{\nu\pi}[\exp(\alpha p_3)] < +\infty$  for all  
 2288  $\alpha > 0$  such that if  $p \geq p_3$ , for all  $n$  such that  $U_n^p \neq \emptyset$ ,  $B_n \notin V_n^p$  implies  $C_n \in V_n^p$ .

2289 *Proof.* Let  $p_3$  to be specified later. Let  $p \geq p_3$ . Let  $n \in \mathbb{N}$  such that  $U_n^p \neq \emptyset$  and  $V_n^p \neq [K]$ , where  
 2290  $U_n^p \subseteq V_n^p$  are defined in Eq. (42). Since the statement holds when  $B_n \in V_n^p$ , we suppose that  
 2291  $B_n \notin V_n^p$  in the following.

2292 Let  $p_0$  as in Lemma 55,  $p_1$  and  $C_\mu$  as in Lemma 56, and  $p_2$  and  $D_\mu$  as in Lemma 57. Let  $p_4 =$   
 2293  $\max\{2p_2 - 1, \frac{4}{3} \max\{p_0, p_1\} - 1/3\}$ , which satisfied that  $\mathbb{E}_{\nu\pi}[\exp(\alpha p_4)] < +\infty$  for all  $\alpha > 0$  by  
 2294 using Lemmas 55, 56 and 57. Then, for all  $p \geq p_4 = \max\{2p_2 - 1, \frac{4}{3} \max\{p_0, p_1\} - 1/3\}$  and all  $n$   
 2295 such that  $B_n \notin V_n^p$ , we have  $\tilde{\mu}_{n,a} \in (0, 1)$  for all  $a \notin V_n^p$ ,  $B_n = b_n^* := \arg \max_{a \notin V_n^p} \mu_a$ ,  $B_n \notin U_n^p$   
 2296 and

$$\forall b \notin \{b_n^*\} \cup V_n^p, \quad W_{\epsilon,b_n^*,b}(\tilde{\mu}_n, N_n) + \log N_{n,b} \geq (1+\eta)^{3(p-1)/4} C_\mu + \frac{3(p-1)}{4} \log(1+\eta),$$

$$\forall b \in U_n^p, \quad W_{\epsilon,b_n^*,b}(\tilde{\mu}_n, N_n) + \log N_{n,b} \leq (1+\eta)^{(p-1)/2} D_\mu + \frac{p-1}{2} \log(1+\eta),$$

2297 where we used Lemmas 55, 56 and 57. Direct manipulations yield that

$$\begin{aligned} (1+\eta)^{3(p-1)/4} C_\mu + \frac{3(p-1)}{4} \log(1+\eta) &\geq (1+\eta)^{(p-1)/2} D_\mu + \frac{p-1}{2} \log(1+\eta) \\ \iff p &\geq p_5 = 4\lceil \log_{1+\eta}(D_\mu/C_\mu) \rceil + 1, \end{aligned}$$

2298 where  $\mathbb{E}_{\nu\pi}[\exp(\alpha p_5)] < +\infty$  for all  $\alpha > 0$  since it is a deterministic constant. Let  $p_3 = \max\{p_4, p_5\}$   
 2299 which satisfies  $\mathbb{E}_{\nu\pi}[\exp(\alpha p_3)] < +\infty$  for all  $\alpha > 0$ . Then, we have shown that for all  $p \geq p_3$ , for  
 2300 all  $n$  such that  $B_n \notin V_n^p$ , we have  $B_n = b_n^*$  and

$$\min_{b \notin \{b_n^*\} \cup V_n^p} \{W_{\epsilon,b_n^*,b}(\tilde{\mu}_n, N_n) + \log N_{n,b}\} > \max_{b \in U_n^p} \{W_{\epsilon,b_n^*,b}(\tilde{\mu}_n, N_n) + \log N_{n,b}\}.$$

2301 By definition of the TC challenger, i.e.,  $C_n \in \arg \min_{b \neq B_n} \{W_{\epsilon,B_n,b}(\tilde{\mu}_n, N_n) + \log N_{n,b}\}$ , we  
 2302 obtain that  $C_n \in V_n^p$ . Otherwise, there would be a contradiction since we assumed  $U_n^p \neq \emptyset$ . This  
 2303 concludes the proof.  $\square$

2304 Lemma 59 shows that all the arms are sufficient explored for large enough  $n$ .

2305 **Lemma 59.** *There exists  $N_0$  with  $\mathbb{E}_{\nu\pi}[N_0] < +\infty$  such that, for all  $n \geq N_0$  and all  $a \in [K]$ ,*  
 2306  *$N_{n,a} \geq \sqrt{n/K}$ .*

2307 *Proof.* Let  $p_0$  and  $p_3$  as in Lemmas 55 and 58. Combining Lemmas 55 and 58 yields that, for all  $p \geq$   
 2308  $p_4 = \max\{p_3, 4p_0/3 - 1/3\}$  and all  $n$  such that  $U_n^p \neq \emptyset$ , we have  $B_n \in V_n^p$  or  $C_n \in V_n^p$ . We have  
 2309  $\mathbb{E}_{\nu\pi}[(1+\eta)^{p_2}] < +\infty$ . We have  $(1+\eta)^{p-1} \geq K(1+\eta)^{3(p-1)/4}$  for all  $p \geq p_5 = 4\lceil \log_{1+\eta} K \rceil + 1$ .  
 2310 Let  $p \geq \max\{p_5, p_4\}$ . For notational simplicity, we conduct the proof as if that  $k(1+\eta)^{p-1} \in \mathbb{N}$  for  
 2311 all  $k \in [K]$ . It is direct to adapt the proof by using the operator  $\lceil \cdot \rceil$ .

2312 Suppose towards contradiction that  $U_{K(1+\eta)^{p-1}}^p$  is not empty. Then, for any  $1 \leq t \leq K(1+\eta)^{p-1}$ ,  
 2313  $U_t^p$  and  $V_t^p$  are non empty as well. Using the pigeonhole principle, there exists some  $a \in [K]$  such  
 2314 that  $N_{(1+\eta)^{p-1},a} \geq (1+\eta)^{3(p-1)/4}$ . Thus, we have  $|V_{(1+\eta)^{p-1}}^p| \leq K - 1$ . Our goal is to show  
 2315 that  $|V_{2(1+\eta)^{p-1}}^p| \leq K - 2$ . A sufficient condition is that one arm in  $V_{(1+\eta)^{p-1}}^p$  is pulled at least  
 2316  $(1+\eta)^{3(p-1)/4}$  times between  $(1+\eta)^{p-1}$  and  $2(1+\eta)^{p-1} - 1$ .

2317 **Case 1.** Suppose there exists  $a \in V_{(1+\eta)^{p-1}}^p$  such that  $L_{2(1+\eta)^{p-1},a} - L_{(1+\eta)^{p-1},a} \geq$   
 2318  $\beta^{-1}((1+\eta)^{3(p-1)/4} + 3/2)$ . Using Lemma 50, we obtain

$$N_{2(1+\eta)^{p-1},a}^a - N_{(1+\eta)^{p-1},a}^a \geq \beta(L_{2(1+\eta)^{p-1},a} - L_{(1+\eta)^{p-1},a}) - 3/2 \geq (1+\eta)^{3(p-1)/4},$$

2319 hence  $a$  is sampled  $(1+\eta)^{3(p-1)/4}$  times between  $(1+\eta)^{p-1}$  and  $2(1+\eta)^{p-1} - 1$ .

2320 **Case 2.** Suppose that for all  $a \in V_{(1+\eta)^{p-1}}^p$ , we have  $L_{2(1+\eta)^{p-1},a} - L_{(1+\eta)^{p-1},a} <$   
 2321  $\beta^{-1}((1+\eta)^{3(p-1)/4} + 3/2)$ . Then,

$$\sum_{a \notin V_{(1+\eta)^{p-1}}^p} (L_{2(1+\eta)^{p-1},a} - L_{(1+\eta)^{p-1},a}) \geq (1+\eta)^{p-1} - K\beta^{-1}((1+\eta)^{3(p-1)/4} + 3/2).$$

2322 Using Lemma 50, we obtain

$$\left| \sum_{a \notin V_{(1+\eta)^{p-1}}^p} (N_{2(1+\eta)^{p-1},a}^a - N_{(1+\eta)^{p-1},a}^a) - \beta \sum_{a \notin V_{(1+\eta)^{p-1}}^p} (L_{2(1+\eta)^{p-1},a} - L_{(1+\eta)^{p-1},a}) \right| \leq 3(K-1)/2.$$

2323 Combining all the above, we obtain

$$\begin{aligned} & \sum_{a \notin V_{(1+\eta)^{p-1}}^p} (L_{2(1+\eta)^{p-1},a} - L_{(1+\eta)^{p-1},a}) - \sum_{a \notin V_{(1+\eta)^{p-1}}^p} (N_{2(1+\eta)^{p-1},a}^a - N_{(1+\eta)^{p-1},a}^a) \\ & \geq (1-\beta) \sum_{a \notin V_{(1+\eta)^{p-1}}^p} (L_{2(1+\eta)^{p-1},a} - L_{(1+\eta)^{p-1},a}) - 3(K-1)/2 \\ & \geq (1-\beta) \left( (1+\eta)^{p-1} - K\beta^{-1}((1+\eta)^{3(p-1)/4} + 3/2) \right) - 3(K-1)/2 \geq K(1+\eta)^{3(p-1)/4} \end{aligned}$$

2324 where the last inequality is obtained for  $p \geq p_6 + 1$  with

$$\begin{aligned} p_6 = \sup \left\{ p \in \mathbb{N} \mid (1-\beta) \left( (1+\eta)^{p-1} - K\beta^{-1}((1+\eta)^{3(p-1)/4} + 3/2) \right) - \frac{3}{2}(K-1) \right. \\ \left. < K(1+\eta)^{3(p-1)/4} \right\}. \end{aligned}$$

2325 The left hand side summation is exactly the number of times where an arm  $a \notin V_{(1+\eta)^{p-1}}^p$  was leader  
 2326 but wasn't sampled, hence we have shown that

$$\sum_{t=(1+\eta)^{p-1}}^{2(1+\eta)^{p-1}-1} \mathbb{1} \left( B_t \notin V_{(1+\eta)^{p-1}}^p, a_t = C_t \right) \geq K(1+\eta)^{3(p-1)/4}.$$

For any  $(1+\eta)^{p-1} \leq t \leq 2(1+\eta)^{p-1} - 1$ ,  $U_t^p$  is non-empty, hence we have  $B_t \notin V_{(1+\eta)^{p-1}}^p$  (hence  $B_t \notin V_t^p$ ) implies  $C_t \in V_t^p \subseteq V_{(1+\eta)^{p-1}}^p$ . Therefore, we have shown that

$$\sum_{t=(1+\eta)^{p-1}}^{2(1+\eta)^{p-1}-1} \mathbb{1}\left(a_t \in V_{(1+\eta)^{p-1}}^p\right) \geq \sum_{t=(1+\eta)^{p-1}}^{2(1+\eta)^{p-1}-1} \mathbb{1}\left(B_t \notin V_{(1+\eta)^{p-1}}^p, a_t = C_t\right) \geq K(1+\eta)^{3(p-1)/4}$$

Therefore, there is at least one arm in  $V_{(1+\eta)^{p-1}}^p$  that is sampled  $(1+\eta)^{3(p-1)/4}$  times between  $(1+\eta)^{p-1}$  and  $2(1+\eta)^{p-1} - 1$ .

In summary, we have shown  $|V_{2(1+\eta)^{p-1}}^p| \leq K - 2$  for all  $p \geq p_7 = \max\{p_6, p_4, p_5\}$ . By induction, for any  $1 \leq k \leq K$ , we have  $|V_{k(1+\eta)^{p-1}}^p| \leq K - k$ , and finally  $U_{K(1+\eta)^{p-1}}^p = \emptyset$  for all  $p \geq p_7$ . Defining  $N_0 = K(1+\eta)^{p_7-1}$ , we have  $\mathbb{E}_{\nu\pi}[N_0] < +\infty$  by using Lemmas 55 and 58 for  $p_4 = \max\{p_3, 4p_0/3 - 1/3\}$  and  $p_6$  and  $p_5$  are deterministic. For all  $n \geq N_0$ , we let  $(1+\eta)^{p-1} = \frac{n}{K}$ . Then, by applying the above, we have  $U_{K(1+\eta)^{p-1}}^p = U_n^{\log_{1+\eta}(n/K)+1}$  is empty, which shows that  $N_{n,a} \geq \sqrt{n/K}$  for all  $a \in [K]$ .  $\square$

### H.3 Convergence Towards $\beta$ -Optimal Allocation

The second step of in the generic analysis of Top Two algorithms [50] is to show the convergence of the empirical proportions towards the  $\beta$ -optimal allocation. First, we show that the leader coincides with the best arm. Hence, the tracking procedure will ensure that the empirical proportion of time we sample it is exactly  $\beta$ . Second, we show that a sub-optimal arm whose empirical proportion overshoots its  $\beta$ -optimal allocation will not be sampled next as challenger. Therefore, this ‘‘overshoots implies not sampled’’ mechanism will ensure the convergence towards the  $\beta$ -optimal allocation. We emphasise that there are multiple ways to select the leader/challenger pair in order to ensure convergence towards the  $\beta$ -optimal allocation. Therefore, other choices of leader/challenger pair would yield similar results. Note that our results heavily rely on having obtained sufficient exploration first.

Lemma 60 shows the leader and the candidate answer are equal to the best arm for large enough  $n$ .

**Lemma 60.** *Let  $N_0$  be as in Lemma 59. There exists  $N_1 \geq N_0$  with  $\mathbb{E}_{\nu\pi}[N_1] < +\infty$  such that, for all  $n \geq N_1$ , we have  $\tilde{\mu}_n \in (0, 1)^K$  and  $\tilde{a}_n = B_n = a^*$ .*

*Proof.* Let  $\Delta_{\min} = \min_{a \neq a^*} (\mu_{a^*} - \mu_a)$  and  $\Delta_0 = \min_{a \in [K]} \min\{\mu_a, 1 - \mu_a\} > 0$ . Using Lemma 54, we obtain, for all  $n \geq N_0$ ,

$$\begin{aligned} \tilde{\mu}_{n,a^*} &\geq \mu_{a^*} - W_\mu \frac{\log(e + \sqrt{n/K}/(1+\eta))}{\sqrt{n/K}/(1+\eta)} \\ \forall a \neq a^*, \quad \tilde{\mu}_{n,a} &\leq \mu_a + W_\mu \frac{\log(e + \sqrt{n/K}/(1+\eta))}{\sqrt{n/K}/(1+\eta)}, \end{aligned}$$

where we used that  $x \rightarrow \log(e+x)/x$  is decreasing and  $\tilde{N}_{n,a} \geq N_{n,a}/(1+\eta) \geq \sqrt{n/K}/(1+\eta)$ .

Let  $N_1 = \max\{N_0, \lceil K(1+\eta)^2 X_1^2 \rceil\}$  where

$$X_1 = \sup\{x > 1 \mid x \leq 4(\Delta_{\min}, \Delta_0)^{-1} W_\mu \log x + e\} \leq h_1(4(\Delta_{\min}, \Delta_0)^{-1} W_\mu, e),$$

where we used Lemma 52, and  $h_1$  defined therein. Using Lemmas 54 and 59, we obtain  $\mathbb{E}_{\nu\pi}[N_1] < +\infty$ . Then, we have  $0 < \mu_a - \Delta_0/4 \leq \tilde{\mu}_{n,a} \leq \mu_a + \Delta_0/4 < 1$  for all  $a \in [K]$ . Moreover, for all  $n \geq N_1$ , we have  $\tilde{\mu}_{n,a^*} \geq \mu_{a^*} - \Delta_{\min}/4$  and  $\tilde{\mu}_{n,a} \leq \mu_a + \Delta_{\min}/4$  for all  $a \neq a^*$ , hence

$$a^* = \arg \max_{a \in [K]} \tilde{\mu}_{n,a} = \arg \max_{a \in [K]} [\tilde{\mu}_{n,a}]_0^1 = \tilde{a}_n = B_n.$$

This concludes the proof.  $\square$

Lemma 61 shows that the pulling proportion of the best arm converges towards  $\beta$ . It is a direct consequence of Lemma 60 by using the same proof as Lemma 39 in Azize et al. [12], hence we omit the proof.



2361 **Lemma 61** (Lemma 39 in Azize et al. [12]). Let  $\gamma > 0$ , and  $N_1$  be as in Lemma 60. There exists a  
 2362 deterministic constant  $C_0 \geq 1$  such that, for all  $n \geq C_0 N_1$ , we have  $\left| \frac{N_{n,a^*}}{n-1} - \beta \right| \leq \gamma$ .

2363 Lemma 62 shows that if a sub-optimal arm overshoots its  $\beta$ -optimal allocation then it cannot be  
 2364 selected as challenger for large enough  $n$ .

2365 **Lemma 62.** Let  $\gamma \in (0, \gamma_\mu)$  where  $\gamma_\mu$  is a problem dependent constant. Let  $N_1$  and  $C_0$  be as in  
 2366 Lemma 60 and 61. There exists  $N_2 \geq C_0 N_1$  with  $\mathbb{E}_{\nu^\pi}[N_2] < +\infty$  such that, for all  $n \geq N_2$ ,

$$\exists a \neq a^*, \quad \frac{N_{n,a}}{n-1} \geq \gamma + \omega_{\epsilon,\beta,a}^* \implies C_n \neq a.$$

2367 *Proof.* Let  $\eta > 0$  and  $\gamma > 0$  be small enough, which we will specify below. Let  $\tilde{\gamma} \in (0, \gamma)$ . Let  $N_1$   
 2368 as in Lemma 60 and  $C_0$  as in Lemma 61 for  $\tilde{\gamma}$ . Let  $n \geq C_0 N_1$ . Therefore, we have  $\tilde{\mu}_n \in (0, 1)^K$  and  
 2369  $\tilde{a}_n = B_n = a^*$  and  $\left| \frac{N_{n,a^*}}{n-1} - \beta \right| \leq \tilde{\gamma}$ . Using the same proof as in Lemma 60, there exists  $N_3$  with  
 2370  $\mathbb{E}_{\nu^\pi}[N_3] < +\infty$  such that, for all  $n \geq N_3$ , we have  $\|\tilde{\mu}_n - \mu\|_\infty \leq \eta$ . Let  $n \geq \max\{C_0 N_1, N_3\}$ .

2371 Let  $a \neq a^*$  such that  $\frac{N_{n,a}}{n-1} \geq \omega_{\epsilon,\beta,a}^* + \gamma$ . Suppose towards contradiction that  $\frac{N_{n,b}}{n-1} > \omega_{\epsilon,\beta,b}^*$  for all  
 2372  $b \notin \{a^*, a\}$ . Then, for all  $n \geq C_0 N_1$ , we have

$$1 - \beta + \tilde{\gamma} \geq 1 - \frac{N_{n,a^*}}{n-1} = \sum_{b \neq a^*} \frac{N_{n,b}}{n-1} > \gamma + \sum_{b \neq a^*} \omega_{\epsilon,\beta,b}^* = 1 - \beta + \gamma,$$

2373 which yields a contradiction since  $\tilde{\gamma} < \gamma$ . Therefore, for all  $n \geq C_0 N_1$ , we have

$$\exists a \neq a^*, \quad \frac{N_{n,a}}{n-1} \geq \omega_{\epsilon,\beta,a}^* + \gamma \implies \exists b \notin \{a^*, a\}, \quad \frac{N_{n,b}}{n-1} \leq \omega_{\epsilon,\beta,b}^*.$$

2374 Let  $b \notin \{a^*, a\}$  such that  $\frac{N_{n,b}}{n-1} \leq \omega_{\epsilon,\beta,b}^*$ . By definition of the TC challenger, we obtain

$$\begin{aligned} C_n \neq a &\iff W_{\epsilon,a^*,a}(\tilde{\mu}_n, N_n) + \log N_{n,a} > W_{\epsilon,a^*,b}(\tilde{\mu}_n, N_n) + \log N_{n,b} \\ &\iff \frac{1}{n-1} (W_{\epsilon,a^*,a}(\tilde{\mu}_n, N_n) - W_{\epsilon,a^*,b}(\tilde{\mu}_n, N_n)) > \frac{1}{n-1} \log \frac{\omega_{\epsilon,\beta,b}^*}{\omega_{\epsilon,\beta,a}^* + \gamma} \\ &\iff \frac{1}{n-1} (W_{\epsilon,a^*,a}(\tilde{\mu}_n, N_n) - W_{\epsilon,a^*,b}(\tilde{\mu}_n, N_n)) > \frac{1}{n-1} \max_{a \neq b} \left| \log \frac{\omega_{\epsilon,\beta,b}^*}{\omega_{\epsilon,\beta,a}^*} \right|, \end{aligned}$$

2375 where we used the positivity of the  $\beta$ -optimal allocation (Lemma 42) to ensure that  
 2376  $\max_{a \neq b} \left| \log \frac{\omega_{\epsilon,\beta,b}^*}{\omega_{\epsilon,\beta,a}^*} \right| \in (0, +\infty)$ . Using that  $\tilde{\mu}_{n,a^*} > \max\{\tilde{\mu}_{n,a}, \tilde{\mu}_{n,b}\}$ , we obtain

$$\begin{aligned} &\frac{1}{n-1} (W_{\epsilon,a^*,a}(\tilde{\mu}_n, N_n) - W_{\epsilon,a^*,b}(\tilde{\mu}_n, N_n)) \\ &\geq \inf_{u \in [0,1]} \left\{ \frac{N_{n,a^*}}{n-1} d_\epsilon^-(\tilde{\mu}_{n,a^*}, u) + (\omega_{\epsilon,\beta,a}^* + \gamma) d_\epsilon^+(\tilde{\mu}_{n,a}, u) \right\} \\ &\quad - \inf_{u \in [0,1]} \left\{ \frac{N_{n,a^*}}{n-1} d_\epsilon^-(\tilde{\mu}_{n,a^*}, u) + \omega_{\epsilon,\beta,b}^* d_\epsilon^+(\tilde{\mu}_{n,b}, u) \right\} \\ &\geq \inf_{\tilde{\beta}: |\tilde{\beta} - \beta| \leq \tilde{\gamma}} G_{a,b}(\tilde{\mu}_n, \tilde{\beta}) \geq \inf_{\lambda: \|\lambda - \mu\|_\infty \leq \eta} \inf_{\tilde{\beta}: |\tilde{\beta} - \beta| \leq \tilde{\gamma}} G_{a,b}(\lambda, \tilde{\beta}), \end{aligned}$$

2377 where, for all  $(a, b) \in ([K] \setminus \{a^*\})^2$  such that  $a \neq b$ ,

$$\begin{aligned} G_{a,b}(\lambda, \tilde{\beta}) &= \inf_{u \in [0,1]} \left\{ \tilde{\beta} d_\epsilon^-(\lambda_{a^*}, u) + (\omega_{\epsilon,\beta,a}^* + \gamma) d_\epsilon^+(\lambda_a, u) \right\} \\ &\quad - \inf_{u \in [0,1]} \left\{ \tilde{\beta} d_\epsilon^-(\lambda_{a^*}, u) + \omega_{\epsilon,\beta,b}^* d_\epsilon^+(\lambda_b, u) \right\}. \end{aligned}$$

2378 Using the equality at equilibrium from (39) (see Lemma 42) and the fact that the transportation  
 2379 costs are increasing in their allocation argument (see Lemma 36), we obtain  $G_{a,b}(\mu, \beta) > 0$  for all  
 2380  $(a, b) \in ([K] \setminus \{a^*\})^2$  such that  $a \neq b$ , since

$$\inf_{u \in [0,1]} \left\{ \beta d_\epsilon^-(\mu_{a^*}, u) + (\omega_{\epsilon,\beta,a}^* + \gamma) d_\epsilon^+(\mu_a, u) \right\} > W_{\epsilon,a^*,a}(\mu, w_{\epsilon,\beta}^*) = W_{\epsilon,a^*,b}(\mu, w_{\epsilon,\beta}^*).$$

2381 By Lemma 48, the functions  $(\lambda, \tilde{\beta}) \rightarrow G_{a,b}(\lambda, \tilde{\beta})$  and  $\lambda \rightarrow \inf_{\tilde{\beta}: |\tilde{\beta} - \beta| \leq \tilde{\gamma}} G_{a,b}(\lambda, \tilde{\beta})$  are continuous.  
 2382 Therefore, there exists  $\eta_\mu$  and  $\gamma_\mu$  small enough such that

$$\inf_{\lambda: \|\lambda - \mu\|_\infty \leq \eta} \inf_{\tilde{\beta}: |\tilde{\beta} - \beta| \leq \tilde{\gamma}} G_{a,b}(\lambda, \tilde{\beta}) \geq G_{a,b}(\mu, \beta)/2 \geq \frac{1}{2} \min_{a \neq b, a \neq a^*, b \neq a^*} G_{a,b}(\mu, \beta) > 0,$$

2383 where the last strict inequality uses that the minimum of a finite number of positive constants  
 2384 is also positive. Considering such  $(\eta_\mu, \gamma_\mu)$  at the beginning of the proof and taking  $N_2 =$   
 2385  $\max\{C_0 N_1, N_3, \kappa_\mu\}$  where

$$\kappa_\mu = 2 + \frac{2 \max_{a \neq b} \left| \log \frac{\omega_{\epsilon, \beta, b}^*}{\omega_{\epsilon, \beta, a}^*} \right|}{\min_{a \neq b, a \neq a^*, b \neq a^*} G_{a,b}(\mu, \beta)} < +\infty,$$

2386 As it satisfies  $\mathbb{E}_{\nu\pi}[N_2] < +\infty$ , this concludes the proof.  $\square$

2387 Lemma 63 shows that the convergence time towards the  $\beta$ -optimal allocation has finite expectation.  
 2388 It is a direct consequence of Lemmas 60, 61 and 62 by using the same proof as Lemma 41 in Azize  
 2389 et al. [12], hence we omit the proof.

2390 **Lemma 63** (Lemma 41 in Azize et al. [12]). *Let  $\gamma \in (0, \gamma_\mu)$  where  $\gamma_\mu$  is a problem dependent*  
 2391 *constant, and  $T_\gamma(w)$  as in Eq. (40). Then, we have  $\mathbb{E}_{\nu\pi}[T_\gamma(\omega_{\epsilon, \beta}^*)] < +\infty$ .*

#### 2392 H.4 Asymptotic Upper Bound

2393 The final step of the generic analysis of Top Two algorithms [50] is to invert the GLR stopping rule in  
 2394 Eq. (7) by leveraging the convergence of the empirical proportions towards the  $\beta$ -optimal allocation.  
 2395 Provided this convergence is shown, the asymptotic upper bound on the expected sample complexity  
 2396 only depends on the dependence in  $\log(1/\delta)$  of the threshold that ensures  $\delta$ -correctness. Compared  
 2397 to the non-private GLR stopping rule, the GLR stopping rule in Eq. (7) pay an extra cost to ensure  
 2398 privacy.

2399 **Lemma 64.** *Let  $\epsilon > 0$ ,  $\eta > 0$  and  $(\delta, \beta) \in (0, 1)^2$ . Let  $T_{\epsilon, \beta}^*(\nu)$  as in Eq. (35) and  $\omega_{\epsilon, \beta}^*$  be its*  
 2400 *associated  $\beta$ -optimal allocation. Assume that there exists  $\gamma_\mu > 0$  such that  $\mathbb{E}_{\nu\pi}[T_\gamma(\omega_{\epsilon, \beta}^*)] < +\infty$*   
 2401 *for all  $\gamma \in (0, \gamma_\mu)$ , where  $T_\gamma(w)$  is defined in Eq. (40). Combining such a sampling rule, using the*  
 2402 *GPE $_\eta(\epsilon)$  update, with the GLR stopping rule as in Eq. (7) and the stopping threshold  $c$  as in Eq. (8)*  
 2403 *yields an  $\epsilon$ -global DP and  $\delta$ -correct algorithm which satisfies that, for all  $\nu$  with mean  $\mu$  such that*  
 2404  *$|a^*(\mu)| = 1$ ,*

$$\limsup_{\delta \rightarrow 0} \frac{\mathbb{E}_{\nu\pi}[\tau_{\epsilon, \delta}]}{\log(1/\delta)} \leq 2(1 + \eta) T_{\epsilon, \beta}^*(\nu).$$

2405 *Proof.* Lemma 4 yields the  $\epsilon$ -global DP. Theorem 5 yields the  $\delta$ -correctness.

2406 Let  $\zeta > 0$  and  $a^*$  be the unique best arm. Using the equality at equilibrium from (39) (see Lemma 42)  
 2407 and the continuity of  $(\mu, w) \mapsto \min_{a \neq a^*(\mu)} W_{\epsilon, a^*(\mu), a}(\mu, w)$  (see Lemma 41), there exists  $\gamma_\zeta > 0$   
 2408 such that  $\left\| \frac{N_n}{n-1} - \omega_{\epsilon, \beta}^* \right\|_\infty \leq \gamma_\zeta$  and  $\|\tilde{\mu}_n - \mu\|_\infty \leq \gamma_\zeta$  implies that

$$\forall a \neq a^*, \quad W_{\epsilon, a^*, a}(\tilde{\mu}_n, N_n/(n-1)) \geq \frac{(1 - \zeta)}{T_{\epsilon, \beta}^*(\nu)}.$$

2409 We choose such a  $\gamma_\zeta$ . Let  $\gamma_\mu > 0$  be such that for  $\mathbb{E}_{\nu\pi}[T_\gamma(\omega_{\epsilon, \beta}^*)] < +\infty$  for all  $\gamma \in (0, \gamma_\mu)$ , where  
 2410  $T_\gamma(w)$  is defined in Eq. (40). Let  $\gamma \in (0, \min\{\gamma_\mu, \gamma_\zeta, \min_{a \in [K]} \omega_{\epsilon, \beta, a}^*/4, \Delta_{\min}/4, \Delta_0/4\})$  where  
 2411  $\Delta_{\min} = \min_{a \neq a^*} (\mu_{a^*} - \mu_a)$  and  $\Delta_0 = \min_{a \in [K]} \min\{\mu_a, 1 - \mu_a\}$ . For all  $n \geq T_\gamma(\omega_{\epsilon, \beta}^*)$ , we have

$$\tilde{N}_{n, a} \geq N_{n, a}/(1 + \eta) \geq (n-1)(\omega_{\epsilon, \beta, a}^* - \gamma)/(1 + \eta) \geq (n-1) \frac{3}{4(1 + \eta)} \min_{a \in [K]} \omega_{\epsilon, \beta, a}^* > 0,$$

2412 where the last inequality used the positivity of the  $\beta$ -optimal allocation (Lemma 42). Since arms  
 2413 are sampled linearly, it is direct to construct  $N_3 \geq T_\gamma(\omega_{\epsilon, \beta}^*)$  with  $\mathbb{E}_{\nu\pi}[N_3] < +\infty$  such that  
 2414  $\|\tilde{\mu}_n - \mu\|_\infty \leq \gamma$  and  $\left\| \frac{N_n}{n-1} - \omega_{\epsilon, \beta}^* \right\|_\infty \leq \gamma$  (as well as  $\min_{a \in [K]} N_{n, a} > e$  trivially).

2415 Recall that  $c(n, \epsilon, \delta) = c_1(n, \delta) + c_2(n, \epsilon)$  where  $n \mapsto c_1(n, \delta)$  and  $n \mapsto c_2(n, \epsilon)$  are increasing (see  
 2416 Lemmas 51 and 38). Since  $\tilde{N}_{n,a} \leq N_{n,a} \leq n$ , we obtain

$$\sum_{b \in \{a^*, a\}} c(\tilde{N}_n, \epsilon, \delta) \leq 2(c_1(n, \delta) + c_2(n, \epsilon)) .$$

2417 Using Lemma 36 and  $\tilde{N}_{n,a} \geq N_{n,a}/(1 + \eta)$  for all  $a \in [K]$  (Lemma 49), we obtain

$$W_{\epsilon, a^*, a}(\tilde{\mu}_n, \tilde{N}_n) \geq \frac{n-1}{1+\eta} W_{\epsilon, a^*, a} \left( \tilde{\mu}_n, \frac{N_n}{n-1} \right) .$$

2418 Let  $\kappa \in (0, 1)$  and  $T > N_3/\kappa$ . For all  $n \in [\kappa T, T]$ , we have  $\tilde{a}_n = a^*$  and, for all  $a \neq a^*$ ,

$$\begin{aligned} & \tau_{\epsilon, \delta} > n \\ \implies & \exists a \neq a^*, \quad W_{\epsilon, a^*, a}(\tilde{\mu}_n, \tilde{N}_n) \leq \sum_{b \in \{a^*, a\}} c(\tilde{N}_n, \epsilon, \delta) \\ \implies & \exists a \neq a^*, \quad \frac{n-1}{1+\eta} W_{\epsilon, a^*, a} \left( \tilde{\mu}_n, \frac{N_n}{n-1} \right) \leq 2(c_1(n, \delta) + c_2(n, \epsilon)) \\ \implies & \exists a \neq a^*, \quad \frac{n-1}{1+\eta} \frac{(1-\zeta)}{T_{\epsilon, \beta}^*(\nu)} \leq 2c_1(T, \delta) + 2c_2(T, \epsilon) , \end{aligned}$$

2419 where we used that  $n \mapsto c_1(n, \delta)$  and  $n \mapsto c_2(n, \epsilon)$  are increasing and  $n \leq T$ . Therefore, we obtain

$$\begin{aligned} \min \{ \tau_{\epsilon, \delta}, T \} & \leq \kappa T + \sum_{n=\kappa T}^T \mathbb{1}(\tau_{\delta} > n) \\ & \leq \kappa T + \sum_{n=\kappa T}^T \mathbb{1} \left( \frac{n-1}{1+\eta} \frac{(1-\zeta)}{T_{\epsilon, \beta}^*(\nu)} \leq 2c_1(T, \delta) + 2c_2(T, \epsilon) \right) \\ & \leq \kappa T + 1 + \frac{2(1+\eta)T_{\epsilon, \beta}^*(\nu)}{1-\zeta} (c_1(T, \delta) + c_2(T, \epsilon)) . \end{aligned}$$

2420 Let  $T_\zeta(\delta)$  defined as

$$T_\zeta(\delta) := \inf \left\{ T \geq 1 \mid \frac{1}{1-\kappa} \left( 1 + \frac{2(1+\eta)T_{\epsilon, \beta}^*(\nu)}{1-\zeta} (c_1(T, \delta) + c_2(T, \epsilon)) \right) \leq T \right\} .$$

2421 Using Lemma 51, we know that  $\bar{W}_{-1}(x) =_{x \rightarrow \infty} x + \log x$ , hence we have  
 2422  $\limsup_{\delta \rightarrow 0} c_1(T, \delta)/\log(1/\delta) \leq 1$ . Since  $\lim_{\delta \rightarrow 0} c_2(T, \epsilon)/\log(1/\delta) = 0$ , we obtain  
 2423  $\limsup_{\delta \rightarrow 0} \frac{T_\zeta(\delta)}{\log(1/\delta)} \leq \frac{2(1+\eta)T_{\epsilon, \beta}^*(\nu)}{(1-\zeta)(1-\kappa)}$ . For every  $T \geq \max\{T_\zeta(\delta), N_3/\kappa\}$ , we have  $\tau_{\epsilon, \delta} \leq T$ ,  
 2424 hence  $\mathbb{E}_{\nu\pi}[\tau_{\epsilon, \delta}] \leq T_\zeta(\delta) + \mathbb{E}_{\nu\pi}[N_3]/\kappa < +\infty$ . Therefore, for all  $\zeta, \kappa > 0$ , we obtain

$$\limsup_{\delta \rightarrow 0} \frac{\mathbb{E}_{\nu\pi}[\tau_{\epsilon, \delta}]}{\log(1/\delta)} \leq \limsup_{\delta \rightarrow 0} \frac{T_\zeta(\delta)}{\log(1/\delta)} \leq \frac{2(1+\eta)T_{\epsilon, \beta}^*(\nu)}{(1-\zeta)(1-\kappa)} .$$

2425 Letting  $\zeta$  and  $\kappa$  go to zero concludes the proof.  $\square$

2426 **Proof of Theorem 6** The proof is obtained by combining Theorem 5 and Lemmas 4, 59, 63 and 64.

## 2427 I Variants of Algorithms

2428 In Appendix I, we propose several variants of the algorithmic components used in our algorithm. The  
 2429 objective is to give freedom of choice for the practitioners interested in solving  $\epsilon$ -global DP BAI.  
 2430 Given the rich literature on BAI, it is unreasonable to provide details for the  $\epsilon$ -global DP version of  
 2431 all the existing BAI algorithms. Therefore, we settle for a few instances that has received increased  
 2432 scrutiny in the BAI literature.

First, we adapt the Track-and-Stop sampling rule [38] to solve  $\epsilon$ -global DP BAI (Appendix I.1). This leverages the computational tractable procedure to compute the optimal allocation  $w_\epsilon^*$  derived in Lemma 46. Second, we explore some alternative choices of components of the Top Two sampling rule for  $\epsilon$ -global DP BAI (Appendix I.2). This includes adaptive choice of target for the leader, hence aiming at achieving  $T_\epsilon^*(\nu)$  instead of  $T_{\epsilon,\beta}^*(\nu)$ . Third, we adapt the LUCB sampling rule [55] for  $\epsilon$ -global DP BAI (Appendix I.3).

### I.1 Track-and-Stop Sampling Rule

The Track-and-Stop (TaS) sampling rule was introduced in the seminal paper [38]. At each time  $n$ , it solves the optimization problem defining the characteristic time for the current empirical estimator  $\tilde{\mu}_n$ . When  $\tilde{\mu}_n \in (0, 1)^K$ , we define  $\tilde{w}_n = w_\epsilon^*(\tilde{\nu}_n)$  where  $\tilde{\nu}_n$  is the Bernoulli instance with means  $\tilde{\mu}_n$ . When  $\tilde{\mu}_n \notin (0, 1)^K$ ,  $[\tilde{\mu}_n]_0^1$  corresponds to a degenerate Bernoulli instance, hence we define  $\tilde{w}_n = 1_K/K$ . Since  $\tilde{\mu}_n$  is updated on a per-arm geometric grid governed by  $\eta$ , the optimal allocation  $\tilde{w}_n$  is updated on the same per-arm geometric grid. Therefore, choosing a larger  $\eta$  yields lower computational cost of TaS at the cost of larger expected sampled complexity, i.e., asymptotic multiplicative factor  $1 + \eta$  due to the update grid.

Given the vector  $\tilde{w}_n \in \Delta_K$ , the next arm  $a_n$  to sample is obtained by using C-Tracking [38] with forced exploration in order to ensure that sufficient exploration holds. This is done here by projecting on  $\Delta_K^\epsilon = \{w \in [\epsilon, 1]^K \mid \sum_{a \in [K]} w_a = 1\}$  for a well chosen  $\epsilon \in (0, 1/K]$ . Let  $\tilde{w}_n^{\epsilon_n}$  be the  $\ell_\infty$  projection of  $\tilde{w}_n$  on  $\Delta_K^{\epsilon_n}$  with  $\epsilon_n = (K^2 + n)^{-1/2}/2$ . While we consider a projection that changes at each time  $n$  (due to  $\epsilon_n$ ),  $\tilde{w}_n^{\epsilon_n}$  could also be updated on a per-arm geometric grid, i.e., when  $\tilde{w}_n$  is updated itself. For all  $n \geq K + 1$ , the TaS sampling rule defines

$$a_n \in \arg \max_{a \in [K]} \left\{ \sum_{t \in [n]} \tilde{w}_{t,a}^{\epsilon_t} - N_{n,a} \right\}. \quad (43)$$

In summary, our proposed Track-and-Stop algorithm is defined as in DP-TT with the sole modification that Lines 13-14 are replaced by the sampling rule defined in Eq. (43).

**Optimal Allocation Oracle** In Lemma 46, we show that  $w_\epsilon^*(\nu)$  can be computed explicitly based on the unique fixed-point solution  $F_\mu(y) = 1$  for  $y \in [0, \min_{a \neq a^*(\mu)} d_\epsilon^-(\mu_{a^*(\mu)}, \mu_a))$ , where  $F_\mu$  is an increasing one-to-one mapping from  $[0, \min_{a \neq a^*(\mu)} d_\epsilon^-(\mu_{a^*(\mu)}, \mu_a))$  to  $[0, +\infty)$  defined as

$$F_\mu(y) = \sum_{a \neq a^*(\mu)} \frac{d_\epsilon^-(\mu_{a^*(\mu)}, u_a(x_a(y)))}{d_\epsilon^+(\mu_a, u_a(x_a(y)))}. \quad (44)$$

The definitions of  $u_a$  and  $x_a$  is deferred to Lemma 46,  $u_a$  is decreasing and  $x_a$  is increasing and strictly convex.

**Asymptotic Expected Sample Complexity** Combining the TaS sampling rule  $a_n$  as in Eq. (43) with the  $\text{GPE}_\eta(\epsilon)$  update and the GLR stopping rule as in Eq. (7) for the stopping threshold as in Eq. (8) yields a  $\delta$ -correct and  $\epsilon$ -global DP algorithm (see Lemma 4 and Theorem 5). Moreover, we conjecture that it satisfies that, for all  $\nu \in \mathcal{F}^K$  with unique best arm,

$$\limsup_{\delta \rightarrow 0} \frac{\mathbb{E}_{\nu\pi} [\tau_{\epsilon,\delta}]}{\log(1/\delta)} \leq 2(1 + \eta)T_\epsilon^*(\nu).$$

The multiplicative factor  $1 + \eta$  comes from the per-arm geometric update grid, and the factor 2 comes from the asymptotic scaling in  $2 \log(1/\delta)$  of the stopping threshold. Using Theorem 6 for  $\beta = 1/2$  and  $T_{\epsilon,1/2}^*(\nu) \leq 2T_\epsilon^*(\nu)$  (Lemma 43), proving this conjecture would only yield an asymptotic improvement by a factor of at most 2. However, this would come at the price of a significantly higher computational cost.

**Proof Sketch of Conjecture** While the detailed proof of this conjecture is beyond the scope of this work, an astute reader could notice that all the necessary steps were proven to derive Theorem 6 for DP-TT. At a higher level, it is intuitive that the asymptotic analysis of Track-and-Stop is simpler than the one of DP-TT.

First, the forced exploration is enforced algorithmically, hence an equivalent of Lemma 59 can be shown for the Track-and-Stop sampling rule. In contrast, the proof of sufficient exploration for DP-TT is more challenging and involves a subtle reasoning towards contradiction, see Appendix H.2 for more details.

Second, the convergence towards the optimal allocation is also enforced algorithmically. Thanks to the forced exploration and due to the continuity of  $\nu \mapsto w_\epsilon^*(\nu)$  (Lemma 41) and the convergence  $\tilde{\mu}_n \rightarrow_{n \rightarrow +\infty} \mu$ , the empirical optimal allocation  $\tilde{w}_n$  converges towards the true optimal allocation  $w_\epsilon^*(\nu)$ . Therefore, an equivalent of Lemma 63 can be shown for the Track-and-Stop sampling rule. In contrast, the proof of convergence towards  $\beta$ -optimal allocation for DP-TT is more challenging and leverage subtle regularity properties of the  $\beta$  characteristic time and its optimal allocation, e.g., the equality at equilibrium of all the transportations costs in Eq. (39), see Appendix H.3 for more details.

Third, the inversion of the GLR stopping rule can be done similarly as for DP-TT. The sole modification lies in using our derived regularity properties for  $w_\epsilon^*(\nu)$  instead of  $w_{\epsilon,\beta}^*(\nu)$ , e.g., the equality at equilibrium of all the transportations costs in Lemma 42. Therefore, an equivalent of Lemma 64 can be shown for the Track-and-Stop sampling rule with  $2(1 + \eta)T_\epsilon^*(\nu)$  instead of  $2(1 + \eta)T_{\epsilon,\beta}^*(\nu)$ , see Appendix H.4 for more details.

## I.2 Top Two Sampling Rule

As detailed in Chapter 2.2 in [48], a Top Two sampling rule is defined by four choices: a leader arm  $B_n \in [K]$ , a challenger arm  $C_n \in [K] \setminus \{B_n\}$ , a target  $\beta_n(B_n, C_n) \in [0, 1]$  and a mechanism to reach the target, i.e.,  $a_n \in \{B_n, C_n\}$  by using  $\beta_n(B_n, C_n)$ . For instance, the sampling rule in DP-TT uses the EB leader, the TCI challenger, a fixed target  $\beta \in (0, 1)$  and  $K$  independent  $\beta$ -tracking procedures (one per leader). We propose adaptive choice of target (Appendix I.2.1), as well as leader fostering implicit exploration (Appendix I.2.2).

### I.2.1 Adaptive Target

When the target is fixed to  $\beta$  beforehand, the Top Two sampling rule can achieve  $T_{\epsilon,\beta}^*(\nu)$  at best. We propose adaptive choices of the target inspired by the recent literature on asymptotically optimal Top Two algorithms [86, 14].

**BOLD Target** Given the EB-TCI leader/challenger pair  $(B_n, C_n)$  defined in DP-TT, we adapt the BOLD target from Bandyopadhyay et al. [14]. Let us define

$$u_{\epsilon,B_n,a}(\tilde{\mu}_n, N_n) = \arg \min_{u \in [0,1]} \{N_{n,B_n} d_\epsilon^-(\tilde{\mu}_{n,B_n}, u) + N_{n,a} d_\epsilon^+(\tilde{\mu}_{n,b}, u)\}, \quad (45)$$

whose closed-form solution is given in Lemma 44. Then, the deterministic BOLD target defines the next arm to pull as

$$a_n = B_n \quad \text{if} \quad \sum_{a \neq B_n} \frac{d_\epsilon^-(\tilde{\mu}_{n,B_n}, u_{\epsilon,B_n,a}(\tilde{\mu}_n, N_n))}{d_\epsilon^+(\tilde{\mu}_{n,a}, u_{\epsilon,B_n,a}(\tilde{\mu}_n, N_n))} > 1 \quad \text{and} \quad a_n = C_n \quad \text{otherwise.} \quad (46)$$

In summary, the sole modification in DP-TT is Line 14 that is replaced by the sampling rule defined in Eq. (46).

For any single-parameter exponential family of distributions, Bandyopadhyay et al. [14] shows that the BOLD target allows to reach asymptotic optimality. Forced exploration is added by Bandyopadhyay et al. [14] to ensure that sufficient exploration holds. Showing that the BOLD target can achieve asymptotic optimality without forced exploration, i.e., meaning that it ensures sufficient exploration on its own, is an open problem.

**IDS Target** Given the EB-TCI leader/challenger pair  $(B_n, C_n)$  defined in DP-TT, we adapt the IDS target from You et al. [86]. Namely, the randomized IDS target defines the next arm to pull from as

$$a_n = \begin{cases} B_n & \text{with proba } \beta_n(B_n, C_n) \\ C_n & \text{otherwise} \end{cases} \quad \text{where } \beta_n(B_n, C_n) = \frac{N_{n,B_n} d_\epsilon^-(\tilde{\mu}_{n,B_n}, u_{\epsilon,B_n,C_n}(\tilde{\mu}_n, N_n))}{W_{\epsilon,B_n,C_n}(\tilde{\mu}_n, N_n)}, \quad (47)$$

where  $u_{\epsilon, B_n, C_n}(\tilde{\mu}_n, N_n)$  is defined in Eq. (45). In summary, the sole modification in DP-TT is Line 14 that is replaced by the sampling rule defined in Eq. (47).

While we could use  $K(K-1)$  tracking procedures to select  $a_n \in \{B_n, C_n\}$ , we use randomization above for the sake of simplicity. For Gaussian distributions with known variance, You et al. [86] shows that the IDS target allows to reach asymptotic optimality. Showing that the IDS target can achieve optimality for other classes of distributions is an open problem.

**Asymptotic Expected Sample Complexity** Sampling  $a_n$  as in Eq. (46) or (47) for the EB-TCI leader/challenger pair  $(B_n, C_n)$  defined in DP-TT based on the  $\text{GPE}_\eta(\epsilon)$  update and the GLR stopping rule as in Eq. (7) for the stopping threshold as in Eq. (8) yields a  $\delta$ -correct and  $\epsilon$ -global DP algorithm (see Lemma 4 and Theorem 5).

While we conjecture that their asymptotic expected sample complexities  $\limsup_{\delta \rightarrow 0} \frac{\mathbb{E}_{\nu\pi}[\tau_{\epsilon, \delta}]}{\log(1/\delta)}$  are both upper bounded by  $2(1+\eta)T_\epsilon^*(\nu)$ , we emphasize that our analysis doesn't provide the necessary steps for this result to hold. This is an interesting research direction left for future work.

## I.2.2 Implicit Exploring Leaders and TC Challenger

The empirical best (EB) leader is a greedy choice of leader that doesn't foster implicit exploration. Without additional exploration mechanism, it can suffer from large empirical stopping time despite being enough for an asymptotic analysis, see [50]. This motivated the choice of the TCI challenger for DP-TT, since it fosters additional implicit exploration by penalizing over sampled challengers with the  $\log N_{n,a}$  term. We propose other choices of leaders that foster implicit exploration, and define the TC challenger that removes this penalization.

The UCB leader is defined as

$$B_n^{\text{UCB}} \in \arg \max_{a \in [K]} U_{n,a} \quad \text{where} \quad U_{n,a} = \max \left\{ u \in [0, 1] \mid N_{n,a} d_\epsilon^+([\tilde{\mu}_{n,a}]_0^1, u) \leq \log(n) \right\}. \quad (48)$$

By adding a bonus to the empirical mean, we are optimistic since we consider that the means are better than suggested by our observations.

The IMED leader builds on the IMED algorithm [42] is defined as

$$B_n^{\text{IMED}} \in \arg \min_{a \in [K]} \left\{ N_{n,a} d_\epsilon^+([\tilde{\mu}_{n,a}]_0^1, \tilde{\mu}_n^*) + \log N_{n,a} \right\} \quad \text{where} \quad \tilde{\mu}_n^* = \max_{a \in [K]} [\tilde{\mu}_{n,a}]_0^1. \quad (49)$$

The TC challenger is defined as

$$C_n^{\text{TC}} \in \arg \min_{a \neq B_n} W_{\epsilon, B_n, b}(\tilde{\mu}_n, N_n), \quad (50)$$

where  $W_{\epsilon, a, b}$  is defined as in Eq. (4).

In summary, the sole modification in DP-TT is Line 13 which can be replaced by choosing the leader as in Eq. (48) or Eq. (49), or choosing the challenger as in Eq. (50).

**Asymptotic Expected Sample Complexity** Choosing the leader as in Eq. (48) or Eq. (49) or the challenger as in Eq. (50) based on the  $\beta$ -tracking as in DP-TT, the  $\text{GPE}_\eta(\epsilon)$  update and the GLR stopping rule as in Eq. (7) for the stopping threshold as in Eq. (8) yields a  $\delta$ -correct and  $\epsilon$ -global DP algorithm (see Lemma 4 and Theorem 5). Moreover, we conjecture that it satisfies that, for all  $\nu \in \mathcal{F}^K$  with distinct means,

$$\limsup_{\delta \rightarrow 0} \frac{\mathbb{E}_{\nu\pi}[\tau_{\epsilon, \delta}]}{\log(1/\delta)} \leq 2(1+\eta)T_{\epsilon, \beta}^*(\nu).$$

While the detailed proof of this conjecture is beyond the scope of this work, an astute reader could notice that all the necessary steps were proven to derive Theorem 6 for DP-TT. When using the TC challenger as in Eq. (50), the proofs of Lemmas 58 and 62 can be readily adapted. When using the UCB leader as in Eq. (48) or the IMED leader as in Eq. (49), the proofs of Lemmas 55 and 60 could also be adapted.

### 2554 I.3 LUCB Sampling Rule

2555 While the Top Two terminology was introduced in Russo [73], the first sampling rule having a Top  
 2556 Two structure is the greedy sampling strategy in LUCB1 introduced by Kalyanakrishnan et al. [55].  
 2557 At each time  $n$ , it selects the EB leader  $B_n^{\text{EB}} = \tilde{a}_n$  and the UCB challenger defined as

$$C_n^{\text{UCB}} \in \arg \max_{a \neq B_n^{\text{EB}}} U_{n,a} \quad \text{where} \quad U_{n,a} \quad \text{as in Eq. (48)}. \quad (51)$$

2558 Then, it samples both  $B_n^{\text{EB}}$  and  $C_n^{\text{UCB}}$ . Instead of using the GLR stopping rule as in Eq.(7), LUCB1  
 2559 stops when the LCB (lower confidence bound) of the leader exceeds the UCB of the challenger, i.e.,

$$\tau_{\epsilon, \delta}^{\text{LUCB1}} = \inf \left\{ n \mid \tilde{L}_{n, B_n^{\text{EB}}} > U_{n, C_n^{\text{UCB}}} \right\}, \quad (52)$$

2560 where

$$\tilde{L}_{n,a} = \max \left\{ u \in [0, 1] \mid N_{n,a} d_{\epsilon}^{-}([\tilde{\mu}_{n,a}]_0^1, u) \leq \log(n) \right\}. \quad (53)$$

2561 In summary, the modifications in DP-TT are: (1) the sampling rule in Lines 13-15 is replaced by  
 2562 sampling both  $B_n^{\text{EB}}$  and  $C_n^{\text{UCB}}$ , and (2) the stopping rule in Line 10 is replaced by Eq. (52). While  
 2563 studying this algorithm is beyond the scope of this work, we emphasize that LUCB is known to not  
 2564 reach asymptotic ( $\beta$ -)optimality.

## 2565 J Implementation Details and Supplementary Experiments

2566 Appendix J is organized as follows. First, we provide additional detail on the implementation details  
 2567 for our algorithm (Appendix J.1). Second, we provide supplementary experiments to illustrate the  
 2568 good performance of our algorithm (Appendix J.2).

### 2569 J.1 Implementation Details

2570 We present additional experiments comparing the algorithms in different bandit instances with  
 2571 Bernoulli distributions, as defined by Sajed and Sheffet [74], namely

$$\begin{aligned} \mu_1 &= (0.95, 0.9, 0.9, 0.9, 0.5), & \mu_2 &= (0.75, 0.7, 0.7, 0.7, 0.7), \\ \mu_3 &= (0.1, 0.3, 0.5, 0.7, 0.9), & \mu_4 &= (0.75, 0.625, 0.5, 0.375, 0.25)\}, \\ \mu_5 &= (0.75, 0.53125, 0.375, 0.28125, 0.25), & \mu_6 &= (0.75, 0.71875, 0.625, 0.46875, 0.25)\}. \end{aligned}$$

2572 For each Bernoulli instance, we implement the algorithms with

$$\epsilon \in \{0.001, 0.005, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 10, 100, 125\}.$$

2573 The risk level is set at  $\delta = 0.01$ . We verify empirically that the algorithms are  $\delta$ -correct by running  
 2574 each algorithm 1000 times.

2575 We implement all the algorithms in Python (version 3.8) and on an 8 core 64-bits Intel i5@1.6 GHz  
 2576 CPU.

2577 **Remark 2.** *To implement the thresholds of AdaP-TT, AdaP-TT\* and DP-TT, we use empirical*  
 2578 *thresholds that we get by approximating the theoretical thresholds. The expressions of the empirical*  
 2579 *thresholds used can be found in the code in the supplementary material.*

### 2580 J.2 Supplementary Experiments

2581 Figure 2 confirms our experimental findings from Section 6. DP-TT outperforms all the other  
 2582  $\delta$ -correct and  $\epsilon$ -global DP BAI algorithms, for different values of  $\epsilon$  and in all the instances tested.  
 2583 The empirical performance of DP-TT demonstrates two regimes. A high-privacy regime, where the  
 2584 stopping time depends on the privacy budget  $\epsilon$ , and a low privacy regime, where the performance  
 2585 of DP-TT is independent of  $\epsilon$ , and requires twice the number of samples used by the non-private  
 2586 EB-TCI- $\beta$ .

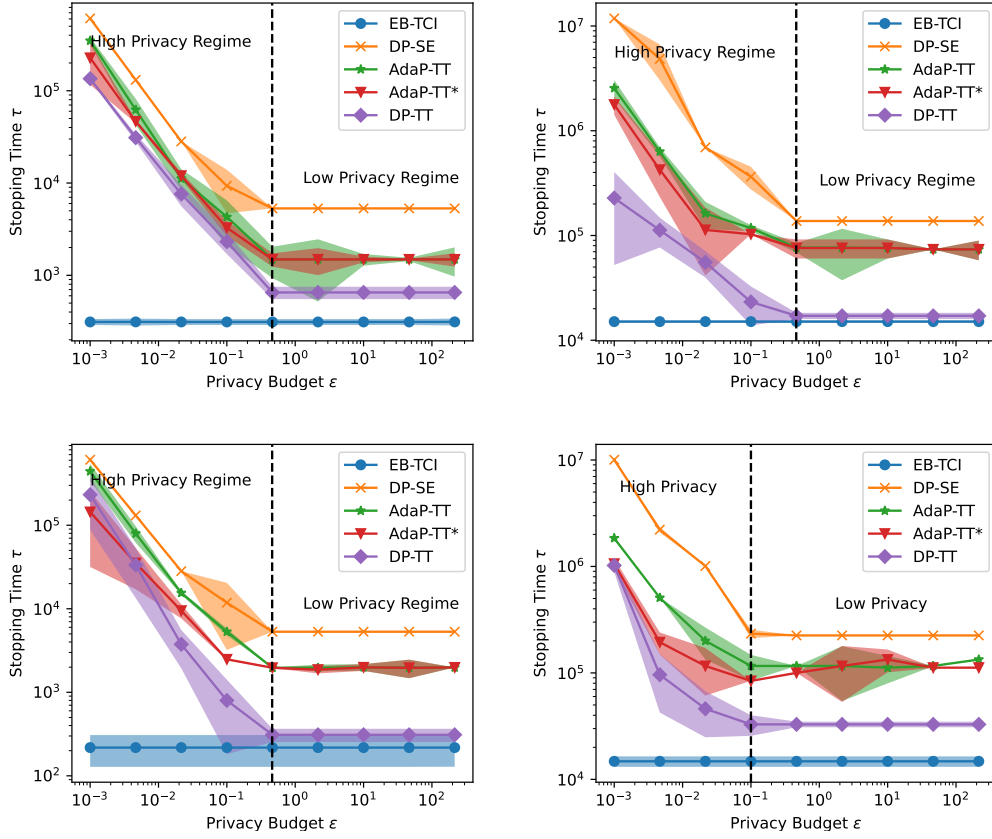


Figure 2: Empirical stopping time  $\tau_{\epsilon, \delta}$  (mean  $\pm 2$  std. over 1000 runs) for  $\delta = 10^{-2}$  with respect to the privacy budget  $\epsilon$  for  $\epsilon$ -global DP on Bernoulli instances  $\mu_3$ ,  $\mu_4$ ,  $\mu_5$  and  $\mu_6$  (top left to bottom right). The shaded vertical line separates the two privacy regimes.