Bandit Pareto Set Identification: the Fixed Budget Setting

Cyrille Kone[1] **Emilie Kaufmann**[1] **Laura Richert**[2]
[1] Univ. Lille, Inria, CNRS, Centrale Lille, UMR 9198-CRIStAL, F-59000 Lille, France
[2] Univ. Bordeaux, Inserm, Inria, BPH, U1219, Sistm, F-33000 Bordeaux, France

Abstract

We study a multi-objective pure exploration problem in a multi-armed bandit model. Each arm is associated to an unknown multivariate distribution and the goal is to identify the distributions whose mean is not uniformly worse than that of another distribution: the Pareto optimal set. We propose and analyze the first algorithms for the *fixed budget* Pareto Set Identification task. We propose Empirical Gap Elimination, a family of algorithms combining a careful estimation of the "hardness to classify" each arm in or out of the Pareto set with a generic elimination scheme. We prove that two particular instances, EGE-SR and EGE-SH, have a probability of error that decays exponentially fast with the budget, with an exponent supported by an information theoretic lower-bound. We complement these findings with an empirical study using real-world and synthetic datasets, which showcase the good performance of our algorithms.

1 INTRODUCTION

The multi-armed bandit problem has been extensively studied in the literature, predominantly as a single-objective stochastic optimization problem. In this framework, an agent sequentially collects samples from a set of K (unknown) probability distributions, called arms. Two common goals are either to maximize the sum of observations (regret minimization, (Lattimore and CSzepesvri, 2020)) or to identify the arm with the largest expected value (best arm identification, Even-Dar et al. (2002); Audibert and Bubeck (2010)). In many real-world problems, however, there are multiple (possibly conflicting) objectives to optimize simultaneously and there might not exist a unique arm that maximizes all these objectives. This leads to a scenario in which the observations are multi-dimensional, and there may be several *Pareto optimal* arms. An arm belongs to the Pareto optimal set $\mathcal{S}^\star$ if it is not dominated by any other arm (in the sense that the mean values of all the objectives would be larger for the latter). While some line of work aims at minimizing the *Pareto regret* by strategically selecting arms from the Pareto set along the agent's trajectory (Drugan and Nowe, 2013), we are interested in the pure exploration counterpart of this problem, first studied by Auer et al. (2016), in which the goal is to identify (relaxations of) the set $\mathcal{S}^\star$, as quickly and accurately as possible. This multi-objective exploration problem has a wide range of applications including adaptive clinical trials (with multiple indicators of efficacy), software and hardware design (with multiple criteria such as speed and energy consumption), or A/B/n testing for recommender systems (with different metrics of users engagement).

Notably, prior work (Auer et al., 2016; Ararat and Tekin, 2023; Kone et al., 2023) has studied the Pareto Set Identification (PSI) problem under the *fixed-confidence setting*: given a risk $\delta \in (0,1)$, the agent stops at a random time τ and recommends a set $\widehat{S}_\tau$ such that $\mathbb{P}(\widehat{S}_\tau \neq \mathcal{S}^\star) \leq \delta$, while minimizing its expected stopping time $\mathbb{E}[\tau]$. However, to the best of our knowledge, no algorithms have been proposed for the dual *fixed-budget setting* in which the agent should output a guess for the Pareto set after a given budget T while minimizing its probability of error $\mathbb{P}(\widehat{S}_T \neq \mathcal{S}^\star)$. For some applications, the fixed-budget setting can be viewed as more practical: e.g., in clinical trials the maximal number of patients is typically defined in advance (due to some financial constraints), and the stopping time used in the fixed-confidence formulation could be prohibitively large.

We fill this gap by proposing the first algorithm(s) for Pareto Set Identification in the fixed-budget setting. Our approach may be viewed as an extension of

Proceedings of the 27$^{\text{th}}$ International Conference on Artificial Intelligence and Statistics (AISTATS) 2024, Valencia, Spain. PMLR: Volume 238.

Successive Rejects (SR) (Audibert and Bubeck, 2010) and Sequential Halving (SH) (Karnin et al., 2013), two state-of-the-art algorithms for fixed-budget best arm identification (i.e. PSI in the uni-dimensional setting), that are based on arm eliminations. We propose a novel arm elimination criterion based on the computation of an *Empirical Gap*. This quantity can be viewed as a measure of the hardness to classify an arm as Pareto optimal or sub-optimal and is inspired by the gaps featured in the analysis of fixed-confidence algorithms (Auer et al., 2016; Kone et al., 2023). While SR and SH successively eliminate arms with smallest empirical means, in Empirical Gap Elimination (EGE), the arms with the largest empirical gaps are eliminated and appropriately classified as optimal or sub-optimal.

We prove an upper bound on the error probability of a generic version of EGE using any arm allocations satisfying some budget constraints (Theorem 1). We explicit our bound for two instances using the same arm allocations as SR and SH respectively, called EGE-SR and EGE-SH, showing that their error probability decrease exponentially fast with the budget (Corollary 1.1). We derive an information-theoretic lower bound showing that these algorithms are essentially optimal in the worse case (Theorem 2). On the practical side, we report in Section 6 some experimental results showing that EGE-SR/SH outperform uniform sampling and are much more robust than APE-FB, an adaptation of an existing fixed-confidence algorithm that requires some prior knowledge of the problem complexity as input (discussed in Appendix D). Finally, we propose in Section 5, EGE-SR-k, a variant of EGE-SR that tackles a relaxation of PSI first considered by Kone et al. (2023) in which one should identify at most k Pareto optimal arms, with a provably reduced probability of error.

Related work The Best Arm Identification problem (BAI) is one of the most studied pure exploration problems in the bandit literature. The fixed-confidence formulation of BAI is by now well understood as there are efficient algorithms with guaranteed bounds on their sample complexity, $\mathbb{E}[\tau]$ (e.g., (Kalyanakrishnan et al., 2012; Gabillon et al., 2012; Jamieson et al., 2014)) and some algorithms that attain the minimal sample complexity on any (parametric) problem instance in the asymptotic regime $\delta \to 0$ (e.g., (Garivier and Kaufmann, 2016; You et al., 2023)). In contrast, the theoretical understanding of fixed-budget BAI is more elusive. While the error probability of any reasonable strategy (e.g. uniform sampling) decays exponentially fast with the budget T, Degenne (2023) shows that there is no optimal exponent that can be attained by a single algorithm on all bandit instances. Still, the Successive Rejects and Sequential Halving algorithms mentioned above have their error probability scaling in $\exp(-T/\log(K)H(\nu))$ for some complexity quantity $H(\nu)$ depending on the bandit instance ν and Carpentier and Locatelli (2016) proved a (worst-case) lower bound showing that there exists instances ν in which this rate is un-improvable (up to constant factors in the exponent) when the budget T is large.

In recent years, there has been a growing interest in pure exploration problems in a multi-dimensional setting. Adaptive algorithms for Pareto Set Identification were first proposed in the literature on black-box multi-objective optimization. Besides several evolutionary heuristics (e.g., Deb et al. (2002); Knowles (2006)), the works of Zuluaga et al. (2013, 2016) have studied PSI under a Gaussian Process modeling assumption, obtaining some error bounds that decay polynomially in T and depend on some notion of information gain that is specific to the GP setting. In the bandit literature, the first fixed-confidence PSI algorithm proposed by Auer et al. (2016) is based on uniform sampling and an accept/reject mechanism. Kone et al. (2023) then proposed a more adaptive LUCB-like algorithm for PSI and some relaxations introduced therein. Ararat and Tekin (2023) studied an extension of PSI to any $\mathbb{R}^D$ partial order defined by an ordering cone, still in the fixed confidence setting.

Alongside PSI, other multi-dimensional pure exploration problems have been investigated such as *feasible arm identification* (Katz-Samuels and Scott, 2018) in which the goal is to identify the arms which belong to a given polyhedron $\mathcal{P} \subset \mathbb{R}^D$ or *top feasible arm identification* (Katz-Samuels and Scott, 2019) in which one seeks a feasible arm that further maximizes some linear combinations of the different objectives. While the former work considers the fixed-budget setting and also proposes some counterpart of Successive Rejects, the latter considers the fixed-confidence setting. In the fixed-budget setting, a special case was recently considered by Faizal and Nair (2022) in a 2-dimensional setting in which the goal is to find the arm maximizing one attribute under the constraint that the second attribute is larger than a given threshold. They propose a SR-like algorithm that is similar in spirit to our Empirical Gap Elimination, but with a very different notion of empirical gap.

2 SETTING

We formalize the fixed budget Pareto Set Identification problem and introduce some notation.

We are given K distributions (or arms) $\nu_1, \dots, \nu_K$ over $\mathbb{R}^D$ with means (resp.) $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_K$. Let $\nu := (\nu_1, \dots, \nu_K)$ be the bandit instance and $\boldsymbol{\Theta} := (\boldsymbol{\theta}_1 \dots \boldsymbol{\theta}_K)^\intercal$. We use boldfaced symbols for $\mathbb{R}^D$ elements. As in prior work on multi-objective pure exploration, we assume that the marginal distributions of each arm are all σ-subgaussian. We recall that a random variable X is σ-subgaussian if $\forall\ u \in \mathbb{R}$, $\mathbb{E}[\exp(u(X - \mathbb{E}[X]))] \leq \exp(u^2\sigma^2/2)$.

Definition 1. *Given two arms $i, j \in [K]$, i is weakly (Pareto) dominated by j (denoted by $\boldsymbol{\theta}_i \leq \boldsymbol{\theta}_j$) if for any $d \in \{1, \dots, D\}$, $\theta_i^d \leq \theta_j^d$. The arm i is (Pareto) dominated by j ($\boldsymbol{\theta}_i \preceq \boldsymbol{\theta}_j$ or $i \preceq j$) if i is weakly dominated by j and there exists $d \in \{1, \dots, D\}$ such that $\theta_i^d < \theta_j^d$. The arm i is strictly (Pareto) dominated by j ($\boldsymbol{\theta}_i \prec \boldsymbol{\theta}_j$ or $i \prec j$) if for any $d \in \{1, \dots, D\}$, $\theta_i^d < \theta_j^d$.*

The Pareto set $\mathcal{S}^\star(\nu)$ is

$$\mathcal{S}^\star(\nu) := \{i \in [K] \text{ s.t } \nexists j \in [K] : \boldsymbol{\theta}_i \prec \boldsymbol{\theta}_j\},$$

and will be denoted by $\mathcal{S}^\star$ when ν is clear from the context. Any arm $a \in \mathcal{S}^\star$ will be called (Pareto) *optimal* and an arm $a \notin \mathcal{S}^\star$ is called *sub-optimal*.

At each time $t = 1, 2, \dots, T$, the agent chooses an arm a_t and observes an independent draw $\mathbf{X}_t \sim \nu_{a_t}$ with $\mathbb{E}(\mathbf{X}_{a_t}) = \boldsymbol{\theta}_{a_t}$. We denote by $\mathbb{P}_\nu$ the law of the stochastic process $\{\mathbf{X}_t\}_{t\geq 1}$ and by $\mathbb{E}_\nu$, the expectation under $\mathbb{P}_\nu$. Let $\mathcal{F}_t := \sigma(a_1, \mathbf{X}_1, \dots, a_t, \mathbf{X}_t, a_{t+1})$ be the σ-algebra representing the history of the process up to time t and $\mathcal{F} := \{\mathcal{F}_t\}_{t\geq 1}$, the filtration of the process. The agent's strategy is adaptive in the sense that a_t is $\mathcal{F}_{t-1}$ measurable. At time T, the agent outputs a guess $\widehat{S}_T$ (which is $\mathcal{F}_T$ measurable) for $\mathcal{S}^\star$ which should minimize the error probability $e_T(\nu) := \mathbb{P}_\nu(\widehat{S}_T \neq \mathcal{S}^\star(\nu))$.

To characterize the difficulty of the problem, we introduce below some quantities that allow to measure how much an arm is dominated. For any arms i, j, we let

$$\begin{aligned} \mathrm{m}(i,j) &:= \min_d \left[\theta_j^d - \theta_i^d\right], \\ \mathrm{M}(i,j) &:= \max_d \left[\theta_i^d - \theta_j^d\right]. \end{aligned}$$

If $\boldsymbol{\theta}_i \not\prec \boldsymbol{\theta}_j$, $\mathrm{M}(i,j)$ is the smallest uniform increase of j that makes it dominate i. If $\boldsymbol{\theta}_i \prec \boldsymbol{\theta}_j$ then $\mathrm{m}(i,j)$ is the smallest increase of any component of i which makes it non-dominated by j. Auer et al. (2016) proposed an algorithm in the fixed-confidence setting whose sample complexity is characterized by some sub-optimality gaps Δ_i's defined as follows. For a sub-optimal arm $i \notin \mathcal{S}^\star$,

$$\Delta_i := \Delta_i^\star := \max_{j\in\mathcal{S}^\star} \mathrm{m}(i,j), \tag{1}$$

which is the smallest quantity that should be added component-wise to $\boldsymbol{\theta}_i$ to make i appear Pareto optimal w.r.t $\{\boldsymbol{\theta}_k : k \in [K] \setminus \{i\}\}$. For an optimal arm $i \in \mathcal{S}^\star$,

$$\Delta_i := \min(\delta_i^+, \delta_i^-) \tag{2}$$

where

$$\begin{aligned} \delta_i^+ &:= \min_{j\in\mathcal{S}^\star\setminus\{i\}} \min(\mathrm{M}(i,j), \mathrm{M}(j,i)), \\ \delta_i^- &:= \min_{j\in[K]\setminus\mathcal{S}^\star} [(\mathrm{M}(j,i))^+ + \Delta_j], \end{aligned}$$

with the convention $\min_\emptyset = +\infty$. δ_i^+ accounts for how much i is close to dominate (or to be dominated by) another optimal arm while δ_i^- translates in a sense the smallest "margin" from an optimal arm i to the sub-optimal arms. We illustrate these quantities in Appendix C. Auer et al. (2016) show that the difficulty (i.e. near-optimal sample complexity) of fixed-confidence PSI is characterized by the complexity term

$$H(\nu) := \sum_{a=1}^K \frac{1}{\Delta_a^2}.$$

In this work, we show that it is also a relevant complexity measure for the fixed-budget setting.

3 ALGORITHMS

In this section, we introduce the family of Empirical Gap Elimination (EGE) algorithms.

3.1 Empirical Gap Elimination

An Empirical Gap Elimination algorithm uses a round-based structure. The algorithm is parameterized by a number of rounds $R > 0$, an arm schedule vector $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_R, \lambda_{R+1}) \in [K]^{R+1}$ satisfying $\lambda_r > \lambda_{r+1}$ where λ_r indicates how many arms are active in round r, and an allocation vector $\boldsymbol{t} = (t_1, \dots, t_R) \in [T]^R$, where t_r is the number of samples gathered from each active arm in round r. Given the maximal budget T, these vectors should further satisfy the following:

$$\lambda_1 = K \quad \text{and} \quad \lambda_{R+1} \in \{0, 1\} \tag{3}$$

$$\text{and} \quad \sum_{r=1}^R \lambda_r t_r \leq T, \tag{4}$$

ideally with an equality.

In each round r, the algorithm maintains a set of active arms, denoted by A_r, that is of size λ_r, and collects t_r new samples from each arm in A_r. The total number of samples from each arm in A_r gathered in the first r rounds[1] is therefore

$$n_r = \sum_{s=1}^r t_s.$$

[1]We emphasize that EGE algorithms do not discard samples between rounds

For each $a \in A_r$, the empirical estimate of $\boldsymbol{\theta}_a$ at the end of round r is denoted by $\widehat{\boldsymbol{\theta}}_a(r) := \frac{1}{n_r}\sum_{s=1}^{n_r} \mathbf{X}_{a,s}$ where $\mathbf{X}_{a,s}$ denote the s-th observation drawn *i.i.d* from distribution ν_a. These estimates are used to carefully decide which arm to explore in the next round, based on an appropriate notion of *empirical gap*.

These gaps are empirical variants of the (fixed-confidence) gaps introduced in the previous section. We first propose a rewriting of these quantities to remove the explicit dependency on $\mathcal{S}^\star(\nu)$.

Lemma 1. *For any arm $i \in [K]$,*

$$\Delta_i = \begin{cases} \Delta_i^\star = \max_{j\in[K]} \mathrm{m}(i,j) & \text{if } i \notin \mathcal{S}^\star \\ \delta_i^\star & \text{if } i \in \mathcal{S}^\star \end{cases},$$

where $\delta_i^\star := \min_{j\neq i}[\mathrm{M}(i,j) \wedge (\mathrm{M}(j,i)^+ + (\Delta_j^\star)^+)]$.

We introduce the empirical quantities

$$\mathrm{m}(i,j;r) := \min_d \left[\widehat{\theta}_j^d(r) - \widehat{\theta}_i^d(r)\right],$$

$$\mathrm{M}(i,j;r) := \max_d \left[(\widehat{\theta}_i^d(r) - \widehat{\theta}_j^d(r))\right].$$

and the empirical Pareto set

$$\begin{aligned} S_r &:= \{i \in A_r : \nexists j \in A_r : \widehat{\boldsymbol{\theta}}_i(r) \prec \widehat{\boldsymbol{\theta}}_j(r)\}, \\ &= \{i \in A_r : \forall j \in A_r\backslash\{i\}, \mathrm{M}(i,j;r) > 0\}. \end{aligned}$$

Finally, we define for any arm $i \in A_r$

$$\widehat{\Delta}_{i,r}^\star := \max_{j\in A_r\backslash\{i\}} \mathrm{m}(i,j;r),$$

$$\widehat{\delta}_{i,r}^\star := \min_{j\in A_r\backslash\{i\}} [\mathrm{M}(i,j;r) \wedge (\mathrm{M}(j,i;r)^+ + (\widehat{\Delta}_{i,r}^\star)^+)]$$

and for any $i \in A_r$

$$\widehat{\Delta}_{i,r} := \begin{cases} \widehat{\Delta}_{i,r}^\star & \text{if } i \in A_r\backslash S_r, \\ \widehat{\delta}_{i,r}^\star & \text{else.} \end{cases} \tag{5}$$

At the end of round r, the algorithm sorts the arms by increasing order of their empirical gaps: $\widehat{\Delta}_{(1),r} \leq \dots \leq \widehat{\Delta}_{(\lambda_r),r}$ and we define $A_{r+1} = \{(1), \dots, (\lambda_{r+1})\}$. In case of ties, i.e. if $\widehat{\Delta}_{(\lambda_{r+1}),r} = \widehat{\Delta}_{(\lambda_{r+1}+m),r}$ for some m, we first add arms in S_r to A_{r+1}. We emphasize that this tie-breaking rule is crucial in our analysis. Arms in $A_{r+1}\backslash A_r$ are further classified as optimal (and added to the set B_{r+1}) or sub-optimal (and added to the set D_{r+1}) based on whether or not they belong to S_r. The output of the algorithm is the set $B_{R+1} \cup A_{R+1}$.

3.2 Particular Instances

Arm elimination algorithms have been proposed for different (mostly uni-dimensional) fixed-budget identification tasks (Audibert and Bubeck, 2010; Bubeck et al., 2013; Karnin et al., 2013; Katz-Samuels and Scott, 2018) with different elimination rules, that could also be rewritten featuring some (simpler) gaps. In these works, two different types of arm schedule and sampling allocations have been mostly investigated:

- In Successive Rejects (SR), one arm is de-activated in each round, ie. $R = K-1$ and $\lambda_r^{\mathrm{SR}} = K - r + 1$ for all $r \leq K$. The sampling allocation proposed by Audibert and Bubeck (2010) satisfies $t_r^{\mathrm{SR}} = n_r^{\mathrm{SR}} - n_{r-1}^{\mathrm{SR}}$ with $n_r^{\mathrm{SR}} = \left\lceil \frac{1}{\overline{\log}(K)} \frac{T-K}{K+1-r} \right\rceil$ where $\overline{\log}(K) := 2^{-1} + \sum_{i=2}^K i^{-1}$ and $n_0^{\mathrm{SR}} = 0$.
- In Sequential Halving (SH), one half of the active set is de-activated in each round, that is $R = \lceil \log_2(K) \rceil$ and for all $r \in \{1, \dots, \lceil \log_2(K) \rceil\}$, $\lambda_{r+1}^{\mathrm{SH}} := \lceil \lambda_r^{\mathrm{SH}}/2 \rceil$ (we easily very that $\lambda_{R+1}^{\mathrm{SH}} = 1$). The sampling allocation proposed by Karnin et al. (2013) is uniformly spread across rounds, that is $t_r^{\mathrm{SH}} := \left\lfloor \frac{T}{|A_r|\lceil \log_2(K) \rceil} \right\rfloor$.

We refer to EGE-SR (resp. EGE-SH) as the instances of EGE using the same allocation as SR (resp. SH). In Appendix F we propose a third instantiation using the geometric allocation of Karpov and Zhang (2022).

Remark 1. *For $D = 1$, the PSI problem coincides with BAI and EGE-SR (resp. EGE-SH) coincides with SR (resp. SH[2]). Indeed, in that case $S_r = \{\widehat{a}_r\}$ reduces to the empirical best arm, $\widehat{\Delta}_{i,r} = \widehat{\theta}_{\widehat{a}_r,r} - \widehat{\theta}_{i,r}$ for $i \neq \hat{a}_r$, $\widehat{\Delta}_{\widehat{a}_r,r} = \min_{i\in A_r\backslash\{\widehat{a}_r\}} \widehat{\Delta}_{i,r}$. Then, our tie-breaking rule ensures that no arm is accepted before the last round and at each round r, A_{r+1} is defined as the λ_{r+1} arms in A_r with the largest empirical means and the final survival arm is recommended as optimal.*

Algorithm 1: Empirical Gap Elimination (EGE)

Result: Pareto set
Data: parameters $R, \boldsymbol{t}, \boldsymbol{\lambda}$
initialize: Set $A_1 = \{1, \dots, K\}, B_1 := \emptyset$, $D_1 = \emptyset$
for $r = 1, 2, \dots, R$ **do**
  Collect t_r samples from each arm $a \in A_r$
  Compute S_r the empirical Pareto set
  Let A_{r+1} be the set of λ_{r+1} arms in A_r with the smallest empirical gaps $\widehat{\Delta}_{a,r}$
  `// ties broken in favor of arms in` S_r
  $B_{r+1} \leftarrow B_r \cup \{S_r \cap (A_r\backslash A_{r+1})\}$;
  $D_{r+1} \leftarrow D_r \cup \{(A_r\backslash A_{r+1})\backslash S_r\}$;
return : $\widehat{S}_T = B_{R+1} \cup A_{R+1}$

---

[2] Besides the fact that the original SH algorithm for BAI discards samples collected in previous rounds to compute the empirical means.

3.3 Alternative Approach

Another idea to tackle fixed-budget PSI is to adapt an existing fixed-confidence algorithm to that setting. The APE algorithm of Kone et al. (2023) takes $\delta \in (0,1)$ as input and uses a stopping time τ_δ such that

$$\mathbb{P}\left(\mathcal{S}^\star = S(\tau_\delta) \text{ and } \tau_\delta \le CH(\nu)f(\delta)\right) \ge 1-\delta,$$

where $S(\tau_\delta)$ is the empirical Pareto set at time τ_δ and f is a function of δ and C is a constant. The idea is then to choose δ so as $CH(\nu)f(\delta) \le T$, that is tune δ w.r.t $H(\nu)$ and T. This is roughly the approach used by UGapEb (Gabillon et al., 2012) in BAI. In Appendix D, we analyze APE-FB, a fixed-budget version of APE that takes as input a parameter $a \ge 0$ and we prove an upper-bound on its probability of error when $a \le \frac{25}{36}\frac{T-K}{H(\nu)}$. However, such an "oracle" tuning is not very satisfying, as assuming $H(\nu)$ to be known is quite unrealistic in practice. This is why in the sequel we focus on presenting the analysis of EGE which does not require any prior knowledge about ν.

4 THEORETICAL GUARANTEES

We first propose an analysis of EGE for a generic number of rounds R, arm schedule $\boldsymbol{\lambda}$ and sampling allocation $\boldsymbol{t}$ satisfying (3) and (4). It features the quantity

$$\widetilde{T}^{R,\boldsymbol{t},\boldsymbol{\lambda}}(\nu) := \min_{r\in[R]} \left(\sum_{s=1}^{r} t_s\right) \Delta^2_{(\lambda_{r+1}+1)},$$

in which the dependency in ν is captured in the gaps.

Theorem 1. *Let ν be a bandit with marginally σ-subgaussian arms. Then Empirical Gap Elimination with parameters $R, \boldsymbol{\lambda}$ and $\boldsymbol{t}$ satisfies*

$$e_T^{EGE}(\nu) \le 2(K-1)|\mathcal{S}^\star| RD \exp\left(-\frac{\widetilde{T}^{R,\boldsymbol{t},\boldsymbol{\lambda}}(\nu)}{144\sigma^2}\right).$$

This result shows that the probability of failure of EGE decreases exponentially fast with $\widetilde{T}^{R,\boldsymbol{t},\boldsymbol{\lambda}}(\nu)$. In Appendix A.2, we further show that for both EGE-SR and EGE-SH, $\widetilde{T}^{R,\boldsymbol{t},\boldsymbol{\lambda}}(\nu)$ is of order $T/(H_2(\nu)\log(K))$ with

$$H_2(\nu) := \max_{i\in[K]} i\Delta_{(i)}^{-2},$$

where $(\cdot)$ is a permutation such that $\Delta_{(1)} \le \cdots \le \Delta_{(K)}$. More precisely, we obtain the following.

Corollary 1.1. *Let $T \ge K$ and ν be a bandit with σ-subgaussian marginals. Then* EGE-SR *satisfies*

$$e_T^{SR}(\nu) \le 2(K-1)^2|\mathcal{S}^\star| D \exp\left(-\frac{T-K}{144\sigma^2 H_2(\nu)\overline{\log}(K)}\right),$$

and for EGE-SH, $e_T^{SH}(\nu)$ *is upper-bounded by*

$$2(K-1)\lceil\log_2(K)\rceil|\mathcal{S}^\star| D \exp\left(-\frac{T}{288\sigma^2 H_2(\nu)\lceil\log_2(K)\rceil}\right).$$

Algorithm	Error probability
EGE-SR	$K^2\lvert\mathcal{S}^\star\rvert D\exp(-T/(H_2(\nu)\log K))$
EGE-SH	$K\log(K)\lvert\mathcal{S}^\star\rvert D\exp(-T/(2H_2(\nu)\log K))$
APE-FB*	$K\log(T)D\exp(-T/H(\nu))$
UA	$K\lvert\mathcal{S}^\star\rvert D\exp\left(-T/(K\Delta_{(1)}^{-2})\right)$

Table 1: Upper bounds on $e_T(\nu)$ for different algorithms (up to constants). *APE-FB tuned with $H(\nu)$.

The complexity measure $H_2(\nu)$ featured in our error exponent satisfies $H_2(\nu) \le H(\nu) \le H_2(\nu)\log(2K)$ as proved by Audibert and Bubeck (2010). For BAI ($D = 1$), we essentially recover the existing guarantees for SR and SH, whose error bounds also feature $H_2(\nu)$, up to constant factors inside the exponential and an extra multiplicative K factor for Sequential Halving. Still, to our knowledge this is the first analysis of the variant of SH that does not discard samples between rounds, which often performs (much) better in practice.

In the general case ($D \ge 1$) we remark that the bounds obtained for EGE-SR and EGE-SH are hard to compare: the latter has an improved polynomial dependence ($K\log_2(K)$ instead of K^2) but a worse constant inside the exponential. As we shall see in the experiments, both algorithms have actually pretty close performance (like in BAI, see Karnin et al. (2013)). Moreover, they both outperform a simple baseline using Uniform Allocation (UA) and recommending the Pareto set of the empirical means. We remark that Theorem 1 yields an upper bound on the error probability of this strategy (by choosing $R = 1$, $t_1 = T/K$ and $\lambda_2 = 0$, for which $\widetilde{T}^{1,\boldsymbol{t},\boldsymbol{\lambda}}(\nu) = n_1\Delta^2_{(\lambda_2+1)} = T/(K\Delta_{(1)}^{-2})$), which we add to our summary in Table 1. This bound can be much worse than that for EGE-SR/SH when the gaps are distinct.

4.1 Lower Bound

We present in this section a lower-bound for some class of instances. We define $\mathcal{B}$ to be the set of means $\boldsymbol{\Theta} \in \mathbb{R}^{K\times D}$ such that each sub-optimal arm i is only dominated by a single arm, denoted by $i^\star$ (that has to belong to $\mathcal{S}^\star$) and that for each optimal arm j there exists a unique sub-optimal arm which is dominated by j, denoted by $\underline{j}$. We further assume that optimal arms are not too close to arms they don't dominate: for any sub-optimal arm i and optimal arm j such that $\boldsymbol{\theta}_i \not\prec \boldsymbol{\theta}_j$,

$$\mathrm{M}(i,j) \ge 3\max(\Delta_i, \Delta_{\underline{j}}).$$

Let $\nu := (\nu_1, \dots, \nu_K)$ be an instance whose means $\boldsymbol{\Theta} \in \mathcal{B}$ and such that $\nu_i \sim \mathcal{N}(\boldsymbol{\theta}_i, \sigma^2 I)$. For every $i \in [K]$ we define the alternative instance $\nu^{(i)} :=$

$(\nu_1, \dots \nu_i^{(i)}, \dots, \nu_K)$ in which only the mean of arm i is modified to:

$$\boldsymbol{\theta}_i^{(i)} := \begin{cases} \boldsymbol{\theta}_i - 2\Delta_i e_{d_{\underline{i}}} & \text{if } i \in \mathcal{S}^\star(\nu), \\ \boldsymbol{\theta}_i + 2\Delta_i e_{d_i} & \text{else,} \end{cases} \tag{6}$$

where $e_1, \dots, e_D$ denotes the canonical basis of $\mathbb{R}^D$ and $d_i := \text{argmin}_d[\theta_{i^\star}^d - \theta_i^d]$. With $\nu^{(0)} := \nu$, we prove the following.

Theorem 2. *Let* $\boldsymbol{\Theta} := (\boldsymbol{\theta}_1 \dots \boldsymbol{\theta}_K)^\intercal \in \mathcal{B}$ *and* $\nu = (\nu_1, \dots, \nu_K)$ *where* $\nu_i \sim \mathcal{N}(\boldsymbol{\theta}_i, \sigma^2 I)$. *For any algorithm* $\mathcal{A}$, *there exists* $i \in \{0, \dots, K\}$ *such that*

$$e_T^{\mathcal{A}}(\nu^{(i)}) \geq \frac{1}{4} \exp\left(-\frac{2T}{\sigma^2 H(\nu^{(i)})}\right).$$

In particular, there exists some instances $\widetilde{\nu} \in \mathcal{B}$ such that $e_T^{\mathcal{A}}(\tilde{\nu}) \geq \frac{1}{4}\exp\left(-2T/(\sigma^2 H(\widetilde{\nu}))\right)$. On such instances, the decay rate of EGE-SR and EGE-SH is optimal up to constants and $\log(K)$ factors, and that of APE-FB is optimal up to constant factors, when the complexity is known. In Appendix C we prove a lower bound that holds for a larger class of instances.

4.2 Sketch of proof of Theorem 1

We define for any arms i, j and round r the events

$$\xi_{i,j,r} := \left\{ \left\| (\widehat{\boldsymbol{\theta}}_{i,n_r} - \widehat{\boldsymbol{\theta}}_{j,n_r}) - (\boldsymbol{\theta}_i - \boldsymbol{\theta}_j) \right\|_\infty \leq c\Delta_{(\lambda_{r+1}+1)} \right\}$$

$$\mathcal{E}_c^1 := \bigcap_{r\in[R]} \bigcap_{i\in\mathcal{S}^\star} \bigcap_{j\in[K]} \xi_{i,j,r} \text{ for any } c > 0.$$

We shall prove that there exists some $c > 0$ such that EGE does not make any error on the event $\mathcal{E}_c^1$. That is, no sub-optimal arm is added to B_r and no optimal arm is added to D_r, in any round r, and the possibly remaining arm in A_{R+1} is an optimal arm.

To do so, an important step is to justify that any sub-optimal arm should be de-activated before the optimal arm that dominates it the most. More formally, for any sub-optimal arm i, we let $i^\star \in \text{argmax}_{j\in\mathcal{S}^\star} \text{m}(i, j)$, which by definition is such that $\Delta_i = \text{m}(i, i^\star)$. For a sub-optimal arm i, we know that $i^\star \in \mathcal{S}^\star$ always exists. More importantly, $i^\star$ could be the only arm dominating i. Therefore it is crucial to ensure that i is no longer active before discarding $i^\star$, otherwise i could appear as optimal w.r.t the remaining active arms. Another challenge is to properly estimate the gaps of the active arms. For this purpose we novelly prove Lemma 1 using geometrical properties of the Pareto set (cf Appendix E for its full proof). Using this lemma and defining

$$\mathcal{P}_r := \{\forall\, i \;\notin \mathcal{S}_\star,\; i \in A_r \Rightarrow i^\star \in A_r\},$$

we first prove the following concentration result.

Lemma 2. *Assume that* $\mathcal{E}_c^1$ *holds. Let* $r \in [R]$ *such that* $\mathcal{P}_r$ *holds. Then, for any sub-optimal arm* $i \in A_r$,

$$|\widehat{\Delta}_{i,r}^\star - \Delta_i^\star| \leq 2c\Delta_{(\lambda_{r+1}+1)}$$

and for any optimal arm $i \in A_r$,

$$\widehat{\delta}_{i,r}^\star \geq \Delta_i - 2c\Delta_{(\lambda_{r+1}+1)}.$$

This result then permits to prove by induction that $\mathcal{P}_r$ holds in any round r, when c is small enough.

Lemma 3. *Let* $c < 1/6$. *On the event* $\mathcal{E}_c^1$, *for any* $r \in [R+1]$, $\mathcal{P}_r$ *holds. In particular, for any sub-optimal arm* i, $i^\star$ *cannot be deactivated before* i.

A first consequence is that if A_{R+1} contains one arm (i.e. $\lambda_{R+1} = 1$) and $\mathcal{E}_c^1$ holds, then it is an optimal arm.

Note that in BAI (i.e PSI with $D = 1$), for any sub-optimal arm i, $i^\star$ is the unique best arm and Lemma 3 is sufficient to ensure the correctness of EGE. But in the general case we need to ensure that no optimal arm is rejected and no sub-optimal arm is accepted. This is done by using Lemma 2 and Lemma 3.

Lemma 4. *Let* $c < 1/6$. *On the event* $\mathcal{E}_c^1$, *the recommendation of EGE is correct i.e* $\widehat{S}_T = \mathcal{S}^\star$.

Theorem 1 then follows by upper-bounding the probability of $\bar{\mathcal{E}}_c^1$ using Hoeffding's inequality.

5 RELAXING PSI

In this section we explain how EGE-SR can be slighlty adapted to tackle the "at most k optimal arms" relaxation (or PSI-k) first introduced by Kone et al. (2023) in the fixed confidence setting.

In this problem the goal is to return a subset $\widehat{S}_T$ of $\mathcal{S}^\star$ of size k , or $\mathcal{S}^\star$ itself if its size is smaller than k. In the fixed budget setting, we define the k-relaxed expected loss as $e_{T,k}(\nu) := \mathbb{E}_\nu[\mathcal{L}(\widehat{S}_T, k)]$ where

$$\mathcal{L}(\widehat{S}, k) := \begin{cases} \mathbb{I}\{\widehat{S} \subset \mathcal{S}^\star\} & \text{if } \;|\widehat{S}| = k, \\ \mathbb{I}\{\widehat{S} = \mathcal{S}^\star\} & \text{else.} \end{cases}$$

To minimize $e_{T,k}(\nu)$ for any parameter k and budget T, we propose EGE-SR-k, a variant of EGE-SR which may stop at some round $r < K - 1$ if $|B_{r+1}| = k$ and recommend B_{r+1} (see pseudocode in algorithm 2).

In Appendix A.3, we prove an upper-bound on the expected loss $e_{T,k}(\nu)$. To introduce it, we define $\omega_{(k)}$ to be the k-th largest gap among the optimal arms : $\omega_{(k)} := \max_{i\in\mathcal{S}^\star}^k \Delta_i$ with $\omega_{(k)} = 0$ if $|\mathcal{S}^\star| < k$. Our bound features the complexity measure $H_2^{(k)}(\nu) :=$

Algorithm 2: EGE-SR-k

Result: At most k Pareto optimal arms
initialize: $k \in [K]$, Set $A_1 = [K]$, $B_1 = \emptyset$, $D_1 = \emptyset$
for $r = 1, 2, \dots, K-1$ **do**
  Collect t_r samples from each arm $a \in A_r$
  Compute S_r the empirical Pareto set
  Choose $i_r \in \text{argmax}_{i \in A_r} \widehat{\Delta}_{i,r}$ `// ties broken in favor of arms in` S_r
  $A_{r+1} \leftarrow A_r \backslash \{i_r\}$;
  $B_{r+1} \leftarrow B_r \cup \{S_r \cap \{i_r\}\}$;
  $D_{r+1} \leftarrow D_r \cup \{\{i_r\} \cap (A_r \backslash S_r)\}$;
  if $|B_{r+1}| = k$ **then**
    break and return $\widehat{S}_T = B_{r+1}$;
return : $\widehat{S}_T = B_K \cup A_K$

$\max_{i \in [K]} i(\Delta_{(i)}^{(k)})^{-2}$ with the k-relaxed gaps

$$\Delta_i^{(k)} := \begin{cases} \max(\Delta_i, \omega_{(k)}) & \text{if} \quad i \in \mathcal{S}^\star \\ \Delta_i & \text{else.} \end{cases} \tag{7}$$

Theorem 3. *Let $k \in [K]$. EGE-SR-k satisfies*

$$e_{T,k}(\nu) \leq 2(K-1)^2 |\mathcal{S}^\star| D \exp\left(-\frac{T-K}{144\sigma^2 H_2^{(k)}(\nu)\overline{\log}(K)}\right).$$

This result is particularly insightful when there are many optimal arms and some of them are easy to identify as such (large gaps). Indeed when $|\mathcal{S}^\star| \approx K$ and $k \ll |\mathcal{S}^\star|$, $H_2^{(k)}(\nu)$ can be an order of magnitude smaller than $H_2(\nu)$. We also note that when $k > |\mathcal{S}^\star|$ (then PSI-k reduces to PSI), $H_2^{(k)}(\nu) = H_2(\nu)$ and we recover the result of Corollary 1.1.

The stopping time of EGE-SR-k is

$$\tau := \inf\{r : |B_{r+1}| = k\} \wedge (K-1).$$

Letting N_τ denote the total number of samples used at termination, we upper-bound $\mathbb{E}_\nu[\tau]$ and $\mathbb{E}_\nu[N_\tau]$ in Appendix A.3, showing that when the budget is large, we essentially have $\mathbb{E}_\nu[\tau] \leq q$ and $\mathbb{E}_\nu[N_\tau] \leq N_q$, with $q := K - |\mathcal{S}^\star| + k$. Intuitively it suggests that in the worst case the $(K - |\mathcal{S}^\star|)$ sub-optimal arms are discarded before k optimal arms are accepted as they might be needed to dominate some sub-optimal arms. Moreover, we remark that q can be way smaller than $(K-1)$. For instance when $[K] = \mathcal{S}^\star$, we have $q = k$.

6 EXPERIMENTAL STUDY

We evaluate our algorithms on synthetic and real-world tasks. We compare to Uniform Allocation (UA) and APE-FB for three parameters of the form $a_c = c\frac{25}{36}\frac{T-K}{H(\nu)}$, $c \in \{1/10, 1, 10\}$. Our guarantees on APE-FB are only valid for $c \leq 1$ and are optimal for $c = 1$. We consider a heuristic version of APE-FB which adaptively estimates the hardness $H(\nu)$. We refer to this algorithm as APE-FB-ADAPT. We run the experiments 4000 times with different seeds and we report the $\log_{10}$ of the average mis-identification rate.

6.1 Real-world Datasets

COV-BOOST (Munro et al., 2021) is a phase II vaccine clinical trial conducted on 2883 participants to measure the immunogenicity of different Covid-19 vaccines as a booster (third dose). Combining the first two doses received and the third dose investigated in the trial, there were $K = 20$ arms and the authors reported the sample mean response and confidence intervals (based on log-normal assumption of the data) of each arm to a bunch of immunogenicity indicators. Kone et al. (2023) further extracted and processed the responses to 3 indicators (cellular response, anti-spike IgG and NT_{50}) to generate a multivariate normal bandit with $K = 20, D = 3$ and a diagonal covariance matrix. We use their dataset in our simulations with a total budget of $T = 2500$.

Hardware Design We use the SNW (Sorting Network Width) dataset (Zuluaga et al., 2012) to generate a bandit model. The dataset is made of 206 different sorting networks (devices used to sort items). The authors reported area and throughput of each network when synthetized on a FPGA (Field-Programmable Gate Array). The area is the number of FPGA slices (resources units) used during execution and the throughput is the number of samples treated per second. The goal is then to optimize both objectives simultaneously (reduced area and large throughput). The simulation process is costly and the measures may slightly vary for the same network due to randomness in the circuits. We simulate a Gaussian bandit with $K = 206, D = 2$ with the SNW dataset. We use a total budget of $T = 5000$.

6.2 Synthetic Benchmark

We run the algorithms on different synthetic instances. For each of them, we compute the complexity $H(\nu)$ and following Audibert and Bubeck (2010); Karnin et al. (2013) we set the total budget to $T = H(\nu)$.

Experiment 1: Arms on a convex Pareto set. $K = 60, D = 2$. We choose $x_1, \dots, x_{20}$ equally spaced in $[0.55, 0.95]$ and $i = 1, \dots, 10$, $\boldsymbol{\theta}_i := (x_i^2, 1/(4x_i^2))^\intercal$. $\boldsymbol{\theta}_{11}, \dots, \boldsymbol{\theta}_{60}$ are chosen from $\{(x,y) \in [0.1, 0.8]^2 : xy \leq \frac{1}{5}\}$.

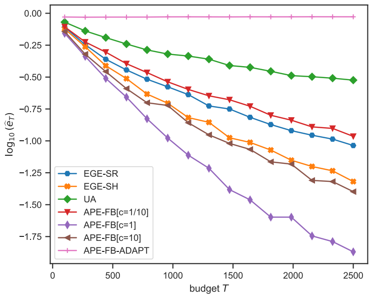

Figure 1: Application 1: COV-BOOST trial

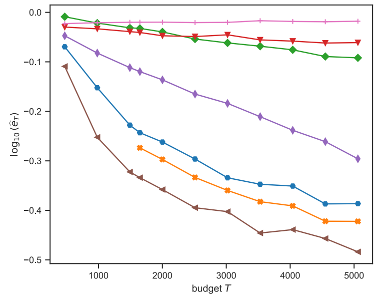

Figure 2: Application 2: Sorting Networks dataset.

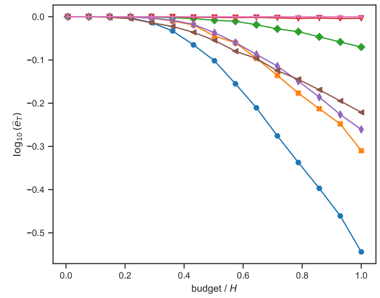

Figure 3: Arms on a convex Pareto set.

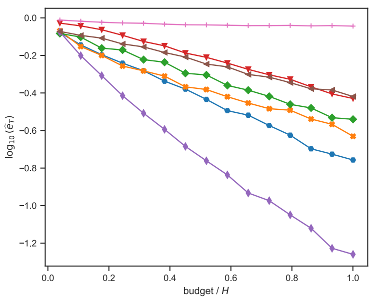

Figure 4: Each sub-optimal i is only dominated by $i^\star$.

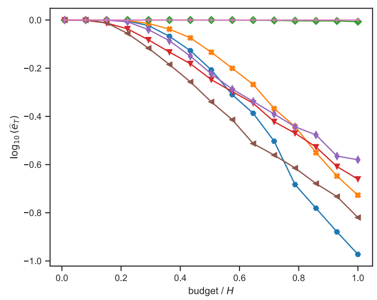

Figure 5: $K = 200$ arms on the unit circle.

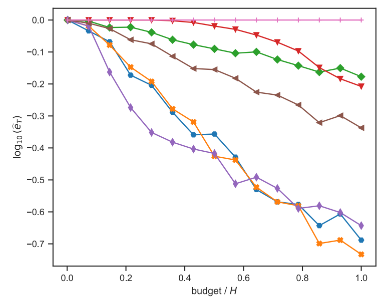

Figure 6: High dimension ($D = 10$) with 2 group of arms.

Experiment 2: Group of sub-optimal arms where each sub-optimal arm i has a unique $i^\star$. $K = 10, D = 2, |\mathcal{S}^\star| = 2$. For each sub-optimal arm i, there is a unique j such that $\boldsymbol{\theta}_i \prec \boldsymbol{\theta}_j$. We choose $\boldsymbol{\theta}_1 := (0.4, 0.75)^\intercal, \boldsymbol{\theta}_2 := (0.75, 0.4)^\intercal$ and for $i = 1, \dots, 4$, $\boldsymbol{\theta}_{2i+1} := (0.45 + 0.2^i, 0.35 - 0.2^i)^\intercal$, $\boldsymbol{\theta}_{2i+2} := (0.10 + 0.20^i, 0.70 - 0.20^i)^\intercal$.

Experiment 3: Many arms on the unit circle. $K = 200, D = 2$ and generate an isotropic multivariate normal instance with $\sigma = 1/4$. We choose $\beta_1, \dots, \beta_{20}$ evenly spaced in $[\pi/12, \pi/2 - \pi/12]$ and $\beta_{21}, \dots, \beta_{200}$ evenly spaced in $[\pi/2 + \pi/6, 2\pi - \pi/6]$. For each arm $i, \boldsymbol{\theta}_i := (\cos(\beta_i), \sin(\beta_i))^\intercal$.

Experiment 4: High dimension ($D = 10$) We set $K = 50$ and generate $(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_{30}) \sim \mathcal{U}\left([0.2, 0.45]^{10}\right)^{\otimes 30}$ and $(\boldsymbol{\theta}_{31}, \dots, \boldsymbol{\theta}_{50}) \sim \mathcal{U}\left([0.55, 0.75]^{10}\right)^{\otimes 20}$

6.3 Results

In all experiments, the uniform allocation (UA) baseline is largely outperformed by both EGE-SR and EGE-SH (with no clear ordering between the two algorithms). This is particularly the case when there are many arms and for complex Pareto sets. By estimating the hardness to classify each arm, EGE eventually allocates more samples to arms that are difficult to classify, leading to a smaller error probability. APE-FB is a good competitor to our EGE algorithms however it is not robust to the hyper-parameter a, which requires the knowledge of $H(\nu)$ to be properly tuned. Our proposed heuristic that estimates the complexity online fails dramatically.

In Appendix G.1 we discuss the computational complexity of EGE and we detail the implementation setup. In Appendix G.2 we report additional experimental results on synthetic datasets. We illustrate the PSI-k relaxation on the SNW dataset, showing that EGE-SR-k can efficiently identify a subset of the Pareto set.

7 CONCLUSION

We proposed the first algorithms for Pareto set identification in the fixed-budget setting.

Our generic algorithm EGE can be coupled with different allocation and arm elimination schemes. We proved that for two instantiations, EGE-SR and EGE-SH, the probability of error decays exponentially fast with the budget, with an exponent that is unimprovable for some bandit instances. We conducted experiments showing that these algorithms consistently outperform a uniform sampling baseline and are competitive with an oracle algorithm that knows the complexity of the underlying instance.

Extensions of this work could consider variance-aware estimates of the gaps or a Bayesian setting where prior information on the means could be used to improve performance (Atsidakou et al., 2023). Finally, just like Successive Rejects or Sequential Halving, EGE algorithms heavily relies on the knowledge of T for their tuning. In future work, we are interested in proposing anytime algorithms, that can have a small error probability for any budget T. We may take inspiration from some anytime algorithms recently proposed for simpler pure exploration tasks (Katz-Samuels and Scott, 2018; Jourdan et al., 2023).

Acknowledgements

Cyrille Kone is funded by an Inria/Inserm PhD grant. Emilie Kaufmann acknowledges the support of the French National Research Agency under the projects BOLD (ANR-19-CE23-0026-04) and FATE (ANR22-CE23-0016-01).

References

Ararat, C. and Tekin, C. (2023). Vector optimization with stochastic bandit feedback. In *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, pages 2165–2190. PMLR.

Atsidakou, A., Katariya, S., Sanghavi, S., and Kveton, B. (2023). Bayesian fixed-budget best-arm identification. *arXiv preprint, eprint:2211.08572*.

Audibert, J.-Y. and Bubeck, S. (2010). Best arm identification in multi-armed bandits. In *COLT - 23th Conference on Learning Theory - 2010*.

Auer, P., Chiang, C.-K., Ortner, R., and Drugan, M.-M. (2016). Pareto front identification from stochastic bandit feedback. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, pages 939–947. PMLR.

Bubeck, S., Wang, T., and Viswanathan, N. (2013). Multiple identifications in multi-armed bandits. In *Proceedings of the 30th International Conference on Machine Learning*, pages 258–265. PMLR.

Carpentier, A. and Locatelli, A. (2016). Tight (lower) bounds for the fixed budget best arm identification bandit problem. In *29th Annual Conference on Learning Theory*, Proceedings of Machine Learning Research, pages 590–604. PMLR.

Daulton, S., Balandat, M., and Bakshy, E. (2020). Differentiable expected hypervolume improvement for parallel multi-objective bayesian optimization. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*. Curran Associates Inc.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE Transactions on Evolutionary Computation*, pages 182–197.

Degenne, R. (2023). On the existence of a complexity in fixed budget bandit identification. In *Proceedings of Thirty Sixth Conference on Learning Theory*, Proceedings of Machine Learning Research, pages 1131–1154. PMLR.

Drugan, M.-M. and Nowe, A. (2013). Designing multi-objective multi-armed bandits algorithms: A study. In *The 2013 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8.

Even-Dar, E., Mannor, S., and Mansour, Y. (2002). Pac bounds for multi-armed bandit and markov decision processes. In *Annual Conference Computational Learning Theory*.

Faizal, F. Z. and Nair, J. (2022). Constrained pure exploration multi-armed bandits with a fixed budget.

Gabillon, V., Ghavamzadeh, M., and Lazaric, A. (2012). Best arm identification: A unified approach to fixed budget and fixed confidence. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc.

Garivier, A. and Kaufmann, E. (2016). Optimal best arm identification with fixed confidence. In *29th Annual Conference on Learning Theory*, pages 998–1027. PMLR.

Jamieson, K., Malloy, M., Nowak, R., and Bubeck, S. (2014). lil' ucb : An optimal exploration algorithm for multi-armed bandits. In *Proceedings of The 27th Conference on Learning Theory*, pages 423–439. PMLR.

Jourdan, M., Degenne, R., and Kaufmann, E. (2023). An ε-best-arm identification algorithm for fixed-confidence and beyond. *Thirty-Seventh Conference on Neural Information Processing Systems*.

Kalyanakrishnan, S., Tewari, A., Auer, P., and Stone, P. (2012). PAC subset selection in stochastic multi-armed bandits. In *Proceedings of the 29th International Coference on International Conference on Machine Learning*, pages 227–234. Omnipress.

Karnin, Z., Koren, T., and Somekh, O. (2013). Almost optimal exploration in multi-armed bandits. In *Proceedings of the 30th International Conference on International Conference on Machine Learning*. JMLR.

Karpov, N. and Zhang, Q. (2022). Collaborative best arm identification with limited communication on non-iid data. *arXiv preprint, eprint:2207.08015*.

Katz-Samuels, J. and Scott, C. (2018). Feasible arm identification. In *Proceedings of the 35th International Conference on Machine Learning*, pages 2535–2543. PMLR.

Katz-Samuels, J. and Scott, C. (2019). Top feasible arm identification. In *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, pages 1593–1601. PMLR.

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

Knowles, J. (2006). Parego: a hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE Transactions on Evolutionary Computation*, pages 50–66.

Kone, C., Kaufmann, E., and Richert, L. (2023). Adaptive algorithms for relaxed pareto set identification. In *Thirty-seventh Conference on Neural Information Processing Systems*.

Lattimore, T. and CSzepesvri, C. (2020). *Bandit Algorithms*. Cambridge University Press.

Locatelli, A., Gutzeit, M., and Carpentier, A. (2016). An optimal algorithm for the thresholding bandit problem. In *Proceedings of The 33rd International Conference on Machine Learning*, Proceedings of Machine Learning Research, pages 1690–1698. PMLR.

Munro, A.-P.-S., Janani, L., Cornelius, V., and et al. (2021). Safety and immunogenicity of seven COVID-19 vaccines as a third dose (booster) following two doses of ChAdOx1 nCov-19 or BNT162b2 in the UK (COV-BOOST): a blinded, multicentre, randomised, controlled, phase 2 trial. *The Lancet*, 398(10318):2258–2276.

You, W., Qin, C., Wang, Z., and Yang, S. (2023). Information-directed selection for top-two algorithms. In *Proceedings of Thirty Sixth Conference on Learning Theory*, Proceedings of Machine Learning Research, pages 2850–2851. PMLR.

Zuluaga, M., Krause, A., and Püschel, M. (2016). e-pal: An active learning approach to the multi-objective optimization problem. *Journal of Machine Learning Research*, 17(104):1–32.

Zuluaga, M., Milder, P., and Pschel, M. (2012). Computer generation of streaming sorting networks. In *DAC Design Automation Conference 2012*, pages 1241–1249.

Zuluaga, M., Sergent, G., Krause, A., and Pschel, M. (2013). Active learning for multi-objective optimization. In *Proceedings of the 30th International Conference on Machine Learning*, volume 28, pages 462–470. PMLR.

Checklist

1. For all models and algorithms presented, check if you include:
   (a) A clear description of the mathematical setting, assumptions, algorithm, and/or model. [Yes]: Section 2, 3, 4, 5
   (b) An analysis of the properties and complexity (time, space, sample size) of any algorithm. [Yes]: Section 3, 4, 5 and supplemental material
   (c) (Optional) Anonymized source code, with specification of all dependencies, including external libraries. [Not Applicable]

2. For any theoretical claim, check if you include:
   (a) Statements of the full set of assumptions of all theoretical results. [Yes]: Section 3, 4, 5 and supplemental material
   (b) Complete proofs of all theoretical results. [Yes]: Supplemental material
   (c) Clear explanations of any assumptions. [Yes]

3. For all figures and tables that present empirical results, check if you include:
   (a) The code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL). [Yes]
   (b) All the training details (e.g., data splits, hyperparameters, how they were chosen). [Not Applicable]
   (c) A clear definition of the specific measure or statistics and error bars (e.g., with respect to the random seed after running experiments multiple times). [Yes]
   (d) A description of the computing infrastructure used. (e.g., type of GPUs, internal cluster, or cloud provider). [Not Applicable]

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets, check if you include:
   (a) Citations of the creator If your work uses existing assets. [Yes]
   (b) The license information of the assets, if applicable. [Not Applicable]
   (c) New assets either in the supplemental material or as a URL, if applicable. [Not Applicable]
   (d) Information about consent from data providers/curators. [Not Applicable]
   (e) Discussion of sensible content if applicable, e.g., personally identifiable information or offensive content. [Not Applicable]

5. If you used crowdsourcing or conducted research with human subjects, check if you include:
   (a) The full text of instructions given to participants and screenshots. [Not Applicable]
   (b) Descriptions of potential participant risks, with links to Institutional Review Board (IRB) approvals if applicable. [Not Applicable]
   (c) The estimated hourly wage paid to participants and the total amount spent on participant compensation. [Not Applicable]

A ANALYSIS OF EMPIRICAL GAP ELIMINATION

A.1 Proof of Theorem 1

We state below an important lemma used in the proof of Theorem 1. For each sub-optimal arm i, we define $i^\star$ to be an arbitrary element in $\text{argmax}_{k\in\mathcal{S}^\star} \text{m}(i,k)$, which by definition yields $\Delta_i^\star = \text{m}(i,i^\star)$. Introducing the property

$$\mathcal{P}_r := \{\forall i \notin \mathcal{S}^\star, i\in A_r \Rightarrow i^\star \in A_r\},$$

the first step of the proof consists in proving that $\mathcal{P}_r$ holds for all r. To do so, we introduce several intermediate results whose proofs are gathered in Appendix B.

The first one controls the deviation of the empirical quantities $\text{M}(i,j;r)$ and $\text{m}(i,j;r)$ from their actual values.

Lemma 5. *On the event $\mathcal{E}_c^1$, for all $r\in[R]$ and $i,j\in A_r$, if $i\in\mathcal{S}^\star$ or $j\in\mathcal{S}^\star$ then*

$$|\text{M}(i,j;r) - \text{M}(i,j)| \leq c\Delta_{(\lambda_{r+1}+1)} \text{ and}$$
$$|\text{m}(i,j;r) - \text{m}(i,j)| \leq c\Delta_{(\lambda_{r+1}+1)}.$$

The second one builds on it and uses Lemma 1 to prove that the empirical gaps cannot be too far from their true gaps, if $\mathcal{P}_r$ holds. It is complementary to Lemma 2 stated in the main paper.

Lemma 6. *Assume that $\mathcal{E}_1^c$ holds and let $r\in[R]$ such that $\mathcal{P}_r$ holds. Then for any $i\in A_r$*

$$\widehat{\Delta}_{i,r} - \Delta_i \geq \begin{cases} -2c\Delta_{(\lambda_{r+1}+1)} & \text{if } i\in\mathcal{S}^\star, \\ -c\Delta_{(\lambda_{r+1}+1)} & \text{else.}\end{cases}$$

Lemma 7. *Let $c>0$ and assume $\mathcal{E}_c^1$ holds. At round $r\in[R]$, for any sub-optimal arm $i\in A_r$, if $i^\star\in A_r$ and $i^\star$ does not empirically dominate i then $\Delta_i^\star \leq c\Delta_{(\lambda_{r+1}+1)}$.*

We are now ready to prove the following key result.

Lemma 3. *Let $c<1/6$. On the event $\mathcal{E}_c^1$, for any $r\in[R+1]$, $\mathcal{P}_r$ holds. In particular, for any sub-optimal arm i, $i^\star$ cannot be deactivated before i.*

Proof. We assume that the event $\mathcal{E}_c^1$ holds and we prove the result by induction on r.

$\mathcal{P}_r$ trivially holds for $r=1$ as all arms are active. Let $r\geq 1$ such that $\cap_{s=1}^r \mathcal{P}_s$ holds. We shall prove that for any $i\in A_{r+1}\cap(\mathcal{S}^\star)^\mathsf{c}$, $i^\star$ cannot be deactivated at the end of round r. A first observation is that the empirical gap $\widehat{\Delta}_{i,r}$ of each arm i satisfies

$$\widehat{\Delta}_{i,r} = \max\left(\widehat{\Delta}_{i,r}^\star, \widehat{\delta}_{i,r}^\star\right) \tag{8}$$

which follows from the fact that when $i\in S_r$, $\widehat{\Delta}_{i,r}^\star < 0 \leq \widehat{\delta}_{i,r}^\star$ and when $i\in A_r\backslash S_r$, $\widehat{\delta}_{i,r}^\star < 0 \leq \widehat{\Delta}_{i,r}^\star$. Using Lemma 6 and the inductive hypothesis permits to prove that

$$\forall i\in A_r, \quad \widehat{\Delta}_{i,r} \geq \Delta_i - 2c\Delta_{(\lambda_{r+1}+1)}. \tag{9}$$

To prove that $\mathcal{P}_{r+1}$ holds, we proceed by contradiction and assume that there exists a sub-optimal arm $i\in A_r$ (and therefore $i^\star\in A_r$ by $\mathcal{P}_r$) such that $i\in A_{r+1}$ and $i^\star\in A_r\backslash A_{r+1}$.

A first observation is that as there are λ_{r+1} arms in A_{r+1} and $i^\star$ is deactivated at the end of round r, there exists an arm $a_r\in A_{r+1}\cup\{i^\star\}$ such that $\Delta_{a_r}\geq\Delta_{(\lambda_{r+1}+1)}$ and $\widehat{\Delta}_{i^\star}\geq\widehat{\Delta}_{a_r}$. We now consider two cases depending on whether $i^\star$ is empirically optimal or empirically sub-optimal.

Case 1: arm $i^\star\notin S_r$ i.e. $i^\star$ is empirically sub-optimal. We have

$$\max_{j\in A_r\backslash\{i^\star\}} \text{m}(i^\star,j;r) := \widehat{\Delta}_{i^\star,r}^\star = \widehat{\Delta}_{i^\star,r} \geq \widehat{\Delta}_{a_r,r}$$

Using Lemma 5 (on the LHS) and Equation (9) (on the RHS) yields

$$\begin{aligned}\max_{j\in A_r\setminus\{i^\star\}} \mathrm{m}(i^\star, j) &\geq \Delta_{a_r} - 3c\Delta_{(\lambda_{r+1}+1)}, \\ &\geq (1-3c)\Delta_{(\lambda_{r+1}+1)}\ ,\end{aligned}$$

where the last inequality follows since $\Delta_{a_r} \geq \Delta_{(\lambda_{r+1}+1)}$. As $i^\star$ is an optimal arm, the LHS of the previous inequality is negative. So it follows that $0 \geq (1-3c)\Delta_{(\lambda_{r+1}+1)}$ which yields a contradiction if $3c < 1$.

Case 2: arm $i^\star \in S_r$ i.e. $i^\star$ is empirically optimal. We first prove that $i^\star$ does not empirically dominate i. Indeed, if i were dominated by $i^\star$, we would have $i \notin S_r$, so $\widehat{\Delta}_{i,r} = \widehat{\Delta}^\star_{i,r} > 0$ and since $i^\star \in S_r$,

$$\begin{aligned}\widehat{\Delta}_{i^\star,r} = \widehat{\delta}^\star_{i^\star,r} &\leq (\mathrm{M}(i, i^\star; r))^+ + (\widehat{\Delta}^\star_{i,r})^+, &(10)\\ &= 0 + \widehat{\Delta}_{i,r}. &(11)\end{aligned}$$

Recalling that $i \in A_{r+1}$ we also have $\widehat{\Delta}_{i^\star,r} \geq \widehat{\Delta}_{i,r}$, which combined with Equation (11) yields

$$\widehat{\Delta}_{i^\star,r} = \widehat{\Delta}_{i,r}. \tag{12}$$

However, the tie-breaking rule of EGE ensures that Equation (12) is not possible since $i^\star \in S_r$ and it is deactivated while $i \in A_{r+1} \cap (S_r)^\mathrm{c}$ (in case of an equality in the gaps, empirically sub-optimal arms are removed). Therefore i is not empirically dominated by $i^\star$. Hence, by Lemma 7

$$\Delta^\star_i \leq c\Delta_{(\lambda_{r+1}+1)}. \tag{13}$$

Moreover, since $a_r \in A_{r+1} \cup \{i^\star\}$, using the definition of $\widehat{\Delta}_{i^\star,r}$ yields

$$\mathrm{M}(i, i^\star; r)^+ + (\widehat{\Delta}^\star_{i,r})^+ \geq \widehat{\Delta}_{a_r,r}$$

Using Lemma 5, Lemma 6 applied to i and Equation (9) applied to a_r, we obtain

$$\begin{aligned}\mathrm{M}(i, i^\star)^+ + \Delta^\star_i &\geq \Delta_{a_r} - 5c\Delta_{(\lambda_{r+1}+1)} \\ &\geq (1-5c)\Delta_{(\lambda_{r+1}+1)}\end{aligned}$$

so

$$\Delta^\star_i \geq (1-5c)\Delta_{(\lambda_{r+1}+1)}. \tag{14}$$

Combining this inequality with Equation (13) yields a contradiction if $6c < 1$. Therefore $\mathcal{P}_{r+1}$ holds and we have proved the claimed result by induction on r. □

Before proving Theorem 1 we need Lemma 8 to upper-bound $\mathbb{P}(\overline{(\mathcal{E}^1_c)})$ for any $c > 0$.

Lemma 8. *For any $c > 0$,*

$$\mathbb{P}(\overline{(\mathcal{E}^1_c)}) \leq 2|\mathcal{S}^\star|(K-1)DR\exp\left(-\frac{c^2\widetilde{T}^{R,\boldsymbol{t},\boldsymbol{\lambda}}(\nu)}{4\sigma^2}\right).$$

We can now prove Theorem 1.

Theorem 1. *Let ν be a bandit with marginally σ-subgaussian arms. Then Empirical Gap Elimination with parameters $R, \boldsymbol{\lambda}$ and $\boldsymbol{t}$ satisfies*

$$e_T^{EGE}(\nu) \leq 2(K-1)|\mathcal{S}^\star|RD\exp\left(-\frac{\widetilde{T}^{R,\boldsymbol{t},\boldsymbol{\lambda}}(\nu)}{144\sigma^2}\right).$$

Proof. We show that on $\mathcal{E}_c^1$, every arm is well classified so the recommended set is the true Pareto optimal set. First by Lemma 3 we know that on $\mathcal{E}_c^1$, the property $(i \in A_r \cap (\mathcal{S}^\star)^{\mathsf{c}} \Rightarrow i^\star \in A_r)$ holds for every $r \in [R+1]$. As in the proof of Lemma 3 (see Equation (9)), this permits to prove in Lemma 6

$$\forall r \in [R], \forall i \in A_r, \quad \widehat{\Delta}_{i,r} \geq \Delta_i - 2c\Delta_{(\lambda_{r+1}+1)} \tag{15}$$

We now establish that at the end of every round $r \in [R]$ no mis-classification can occur. That is for every arm $i \in A_r \backslash A_{r+1}$:

a) if $i \notin S_r$ (that is, i is added to D_r) then $i \notin \mathcal{S}\star$,

b) if $i \in S_r$ (that is, i is added to B_r) then $i \in \mathcal{S}^\star$.

Let $i \in A_r \backslash A_{r+1}$. Since $\lambda_{r+1} = |A_{r+1}|$ should remain active at the end of round r, if i is removed at the end of the round, there exists $a_r \in A_{r+1} \cup \{i\}$ such that $\Delta_{a_r} \geq \Delta_{(\lambda_{r+1}+1)}$ and a_r has a smaller empirical gap than i.

Case 1: $i \notin S_r$ i.e. i is empirically sub-optimal. As $a_r \in A_{r+1} \cup \{i\}$ has a smaller empirical gap than i, we have

$$\widehat{\Delta}^\star_{i,r} \quad \geq \quad \widehat{\Delta}_{a_r,r},$$

Assume by contradiction $i \in \mathcal{S}^\star$. Using Lemma 5 and Equation (15) yields

$$\max_{j \in A_r \backslash \{i\}} \mathrm{m}(i,j) \quad \geq \quad \Delta_{a_r} - 3c\Delta_{(\lambda_{r+1}+1)} \geq (1-3c)\Delta_{(\lambda_{r+1}+1)},$$

When $3c < 1$, the RHS of Equation 16 is positive, so this inequality yields

$$\max_{j \in A_r \backslash \{i\}} \mathrm{m}(i,j) > 0,$$

that is there exists $j \in A_r$ such that

$$\mathrm{m}(i,j) > 0$$

so there exists j such that $\boldsymbol{\theta}_i \prec \boldsymbol{\theta}_j$, which contradicts the assumption $i \in \mathcal{S}^\star$. Therefore, $i \notin \mathcal{S}^\star$: i is a sub-optimal arm.

Case 2: $i \in S_r$ i.e. i is empirically optimal. Since a_r has a larger empirical gap than i and $i \in S_r$, we have

$$\min_{j \in A_r \backslash \{i\}} \mathrm{M}(i,j;r) \geq \widehat{\delta}_i^\star \geq \widehat{\Delta}_{a_r,r} \tag{16}$$

Assume by contradiction that i is sub-optimal. By Lemma 3 (for $c < 1/6$), $i^\star \in A_r$. Combining with (16) yields

$$\mathrm{M}(i, i^\star; r) \geq \widehat{\Delta}_{a_r,r}.$$

Further using Lemma 5 and Equation (15) yields

$$\mathrm{M}(i, i^\star) \geq (1-3c)\Delta_{(\lambda_{r+1}+1)}.$$

Then by taking $6c < 1$, the RHS of the inequality is positive so

$$\mathrm{M}(i, i^\star) > 0, \tag{17}$$

which is not possible as $\boldsymbol{\theta}_i \prec \boldsymbol{\theta}_{i^\star}$. Therefore, $i \in \mathcal{S}^\star$: it is an optimal arm.

This proves that on the event $\mathcal{E}_c^1$, $B_R \subseteq \mathcal{S}^\star$ and $D_R \subseteq (\mathcal{S}^\star)^{\mathsf{c}}$. Moreover, if there is a remaining active arm in A_{R+1} (which happens if and only if $\lambda_{R+1} = 1$), this arm has to be optimal by Lemma 3. Therefore, on $\mathcal{E}_c^1$, the set recommended by EGE is $\widehat{S}_T = B_R \cup A_{R+1} = \mathcal{S}^\star$.

As a consequence, for any $0 < x < 1/6$, letting $c_x = 1/6 - x$ we have

$$e_T^{\mathrm{EGE}}(\nu) \leq \mathbb{P}(\overline{(\mathcal{E}_{c_x}^1)}),$$

which by Lemma 8 yields

$$e_T^{\mathrm{EGE}}(\nu) \leq 2|\mathcal{S}^\star|(K-1)DR \exp\left(-\frac{c_x^2 \widetilde{T}^{R,\boldsymbol{t},\boldsymbol{\lambda}}}{4\sigma^2}\right).$$

We conclude by letting $x \to 0$. □

A.2 Proof of Corollary 1.1

Corollary 1.1. *Let $T \geq K$ and ν be a bandit with σ-subgaussian marginals. Then* EGE-SR *satisfies*

$$e_T^{SR}(\nu) \leq 2(K-1)^2 |\mathcal{S}^\star| D \exp\left(-\frac{T-K}{144\sigma^2 H_2(\nu)\overline{\log}(K)}\right),$$

and for EGE-SH, $e_T^{SH}(\nu)$ *is upper-bounded by*

$$2(K-1)\lceil\log_2(K)\rceil |\mathcal{S}^\star| D \exp\left(-\frac{T}{288\sigma^2 H_2(\nu)\lceil\log_2(K)\rceil}\right).$$

Proof. The result follows from Theorem 1 and the expression of the arm schedule and allocation vector. For EGE-SR, we have $R = K-1$, $\lambda_r^{\mathrm{SR}} = K+1-r$ and $t_r^{\mathrm{SR}} = n_r^{\mathrm{SR}} - n_{r-1}^{\mathrm{SR}}$ where $n_r^{\mathrm{SR}} = \left\lceil \frac{1}{\overline{\log}(K)} \frac{T-K}{K+1-r} \right\rceil$, which yields

$$\begin{aligned}
\widetilde{T}^{R, \boldsymbol{t}^{\mathrm{SR}}, \boldsymbol{\lambda}^{\mathrm{SR}}}(\nu) \quad &:= \min_{r\in[K-1]} n_r^{\mathrm{SR}} \Delta^2_{(\lambda_{r+1}^{\mathrm{SR}}+1)}, \\
&\geq \min_{r\in[K-1]} \frac{\Delta^2_{(K+1-r)}}{\overline{\log}(K)} \frac{T-K}{K+1-r}, \\
&= \frac{T-K}{\overline{\log}(K)} \frac{1}{\max_{r\in\{2,\dots K\}} r\Delta^{-2}_{(r)}} \\
&= \frac{T-K}{\overline{\log}(K)} \frac{1}{H_2(\nu)}.
\end{aligned}$$

For EGE-SH we have $R = \lceil\log_2(K)\rceil$, $n_r^{\mathrm{SH}} := \sum_{s=1}^r t_s^{\mathrm{SH}} \geq \frac{T}{\lambda_r^{\mathrm{SH}}\lceil\log_2(K)\rceil}$ and $\lambda_r^{\mathrm{SH}} \leq 2(\lambda_{r+1}^{\mathrm{SH}}+1)$. Then

$$\begin{aligned}
\widetilde{T}^{R, \boldsymbol{t}^{\mathrm{SH}}, \boldsymbol{\lambda}^{\mathrm{SH}}}(\nu) &:= \min_{r\in\{1,\dots,\lceil\log_2(K)\rceil\}} n_r^{\mathrm{SH}} \Delta^2_{(\lambda_{r+1}^{\mathrm{SH}}+1)}, \\
&\geq \min_{r\in\{1,\dots,\lceil\log_2(K)\rceil\}} \frac{\Delta^2_{(\lambda_{r+1}^{\mathrm{SH}}+1)}}{\lambda_{r+1}^{\mathrm{SH}}+1} \times \frac{T}{2\lceil\log_2(K)\rceil}, \\
&= \frac{T/(2\lceil\log_2(K)\rceil)}{\max_{r\in|\lceil\log_2(K)\rceil|}(\lambda_{r+1}^{\mathrm{SH}}+1)\Delta^{-2}_{(\lambda_{r+1}^{\mathrm{SH}}+1)}} \\
&\geq \frac{T}{2H_2(\nu)\lceil\log_2(K)\rceil}.
\end{aligned}$$

□

A.3 Analysis of EGE-SR-k

In this section, we analyze EGE-SR-k and bound its stopping time.

We define for any $c > 0$ and $t \in [K-1]$:

$$\mathcal{E}_c^2(t) := \bigcap_{r\in[t]} \bigcap_{i\in\mathcal{S}^\star} \bigcap_{j\in[K]} \left\{ \left\| (\widehat{\boldsymbol{\theta}}_{i,n_r} - \widehat{\boldsymbol{\theta}}_{j,n_r}) - (\boldsymbol{\theta}_i - \boldsymbol{\theta}_j) \right\|_\infty \leq c\Delta^{(k)}_{(K+1-r)} \right\}$$

and in particular we abuse notation and define

$$\mathcal{E}_c^2 := \mathcal{E}_c^2(K-1).$$

We define $\tilde{k} = \min(k, |\mathcal{S}^\star|)$.

For any round r, let $\alpha(r) = |B_r|$ denote the number of arms so far identified as optimal at the beginning of round r. We denote these arms by $a_1, \dots, a_{\alpha(r)}$. We say that the algorithm has made an error before round r if $B_r \cap (\mathcal{S}^\star)^\mathsf{c} \neq \emptyset$ or $D_r \cap \mathcal{S}^\star \neq \emptyset$. We will show that on $\mathcal{E}_c^2(t)$, the algorithm does not make any error until the end of round τ_k^t such that $\alpha(\tau_k^t + 1) = \tilde{k}$.

Lemma 9. *Let $c < 1/6, t \in [R]$ and $k \leq |\mathcal{S}^\star|$. Let $\tau_k^t := \min\{r \in [t] : \alpha(r+1) = \tilde{k}\} \wedge t$. On the event $\mathcal{E}_c^2(t)$, the algorithm makes no error until the end round τ_k^t and for any $r \leq \tau_k^t + 1$, if $j \in A_r$ is a sub-optimal arm then $j^\star \in A_r$. In particular, on the event $\mathcal{E}_c^2(t)$, $\{a_1, \dots, a_{\alpha(\tau_k^t+1)}\} \subset \mathcal{S}^\star$.*

Said otherwise, this lemma states that the first arms that will be declared as optimal by EGE-SR-k will be actually optimal if $\mathcal{E}_c^2(t)$ holds for c small enough.

Proof of Lemma 9. In the sequel we assume that $\mathcal{E}_c^2(t)$ holds. We prove the correctness by induction on the round r. Let $r \in [t]$ and let $B_r = \left\{a_1, \dots, a_{\alpha(r)}\right\}$ denote the arms so far identified as optimal. Let $\mathcal{H}_r$ be the property "for any sub-optimal arm $i \in A_r$, $i^\star \in A_r$ and no error occurred so far".

$\mathcal{H}_r$ trivially holds for $r = 1$ and $\alpha(1) = 0$. We now assume that it holds until the beginning of round r and that $\alpha(r) < \tilde{k}$. We will show that the arm i_r de-activated at the end of round r is well classified and that for any sub-optimal arm $j \in A_{r+1}$, $j^\star \in A_{r+1}$.

Lemma 10. *If $\mathcal{H}_r$ holds at round r and $\alpha(r) < \tilde{k}$, there exists $a \in A_r$ such that $\Delta_a^{(k)} = \Delta_a$ and $\Delta_a^{(k)} \geq \Delta_{(K+1-r)}^{(k)}$.*

Proof of Lemma 10. At the beginning of round r, it remains $K - r + 1$ active arms so there exists $a \in A_r$ such that $\Delta_a^{(k)} \geq \Delta_{(K+1-r)}^{(k)}$. If a is sub-optimal then $\Delta_a^{(k)} = \Delta_a$. Otherwise, if a is an optimal arm, since $\alpha(r) < \tilde{k}$ and no error has occurred so for (by assumption) then there exists one of the optimal arms $a' \in A_r$ (one of those with the $\tilde{k}$ largest gaps) such that $\Delta_{a'}^{(k)} = \Delta_{a'}$ and $\Delta_{(K+1-r)}^{(k)} \leq \Delta_a^{(k)} \leq \Delta_{a'}^{(k)}$. ☐

Lemma 5 still holds for the event $\mathcal{E}_c^2(t)$ with the modified gaps introduced earlier. We state the following lemma which is similar to Lemma 2.

Lemma 11. *Assume that the event $\mathcal{E}_c^2(t)$ holds. Let $r \in [t]$ and assume that for* any *sub-optimal arm $j \in A_r$, $j^\star \in A_r$. Then, for any sub-optimal arm $i \in A_r$,*

$$|\widehat{\Delta}_{i,r}^\star - \Delta_i^\star| \leq 2c\Delta_{(K+1-r)}^{(k)}$$

and for any optimal arm $i \in A_r$,

$$\widehat{\delta}_{i,r}^\star \geq \Delta_i - 2c\Delta_{(K+1-r)}^{(k)}.$$

As already noted in the proof of Lemma 3 and Lemma 6, it is simple to see that the empirical gap $\widehat{\Delta}_{i,r}$ of each arm i satisfies

$$\widehat{\Delta}_{i,r} = \max\left(\widehat{\Delta}_{i,r}^\star, \widehat{\delta}_{i,r}^\star\right). \tag{18}$$

It follows from the fact that when $i \in S_r$, $\widehat{\Delta}_{i,r}^\star < 0 \leq \widehat{\delta}_{i,r}^\star$ and when $i \in A_r \backslash S_r$, $\widehat{\delta}_{i,r}^\star < 0 \leq \widehat{\Delta}_{i,r}^\star$.

Using Lemma 11 and the inductive hypothesis $\mathcal{H}_r$ permits to prove that

$$\forall i \in A_r, \quad \widehat{\Delta}_{i,r} \geq \Delta_i - 2c\Delta_{(K+1-k)}^{(k)} \tag{19}$$

We split the proof in three steps to prove that $\mathcal{H}_{r+1}$ holds.

Step 1: i_r is well classified. Let i_r be empirically sub-optimal ($i_r \notin S_r$) and assume $i_r \in \mathcal{S}^\star$. Since i_r is removed at the end of round r, by the inductive assumption and using Equation 19 and Lemma 5 (adapted) we have

$$\max_{j \in A_r \backslash \{i_r\}} \mathrm{m}(i_r, j) \geq \Delta_i - 3c\Delta_{(K+1-r)}^{(k)}, \ \forall i \in A_r. \tag{20}$$

By Lemma 10, there exists $a \in A_r$ such that $\Delta_a = \Delta_a^{(k)} \geq \Delta_{(K+1-r)}^{(k)}$. Applying (20) to this arm a yields

$$\max_{j \in A_r \backslash \{i_r\}} \mathrm{m}(i_r, j) \geq (1 - 3c)\Delta_{(K+1-r)}^{(k)}. \tag{21}$$

Recalling that $c < 1/6$ we see that the LHS of (21) is positive, so there exists j such that

$$\boldsymbol{\theta}_{i_r} \prec \boldsymbol{\theta}_j$$

which contradicts the assumption $i_r \in \mathcal{S}^\star$, so $i_r \in [K]\backslash\mathcal{S}^\star$. Now if i_r is empirically optimal, i.e $i_r \in S_r$ assume i_r is a sub-optimal arm that is $i_r \notin \mathcal{S}^\star$. By the hypothesis $\mathcal{H}_r$, $i_r^\star \in A_r$ and by definition of $\widehat{\Delta}_{i_r,r}$ and since i_r is removed, we have (Lemma 5 and Equation 19)

$$\mathrm{M}(i_r, i_r^\star) \geq \Delta_i - 3c\Delta^{(k)}_{(K+1-r)} \quad \forall i \in A_r. \tag{22}$$

Using Lemma 10 as before, there exists $a \in A_r$ such that $\Delta^{(k)}_{(K+1-r)} \leq \Delta^{(k)}_a = \Delta_a$. So the LHS of (22) is positive for $c < 1/6$. That is

$$\mathrm{M}(i_r, i_r^\star) > 0,$$

which is impossible as $\boldsymbol{\theta}_{i_r} \prec \boldsymbol{\theta}_{i_r^\star}$. This concludes the proof of this part: i_r is well classified.

Step 2: If $i_r \in \mathcal{S}^\star$ is an optimal arm, then no active arm is "optimally" dominated by i_r:

$$\forall i \in A_{r+1} \cap (\mathcal{S}^\star)^\mathsf{c}, i^\star \neq i_r$$

To prove this, assume $i_r \in \mathcal{S}^\star$. Note that we can deduce from "Step 1" that $i_r \in S_r$. Then by contradiction, let $i \in A_{r+1} \cap (\mathcal{S}^\star)^\mathsf{c}$ such that $i^\star = i_r$. We claim that i is not empirically dominated by i_r. Indeed, if i were dominated by i_r, we would have $i \notin S_r$ (i.e empirically sub-optimal) so

$$\widehat{\Delta}_{i,r} = \widehat{\Delta}^\star_{i,r} > 0$$

and since $i_r \in S_r$, by defintion of $\widehat{\delta}^\star_{i_r,r}$

$$\begin{aligned} \widehat{\Delta}_{i_r,r} &\leq (\mathrm{M}(i, i_r; r))^+ + (\widehat{\Delta}^\star_i) \\ &= 0 + \widehat{\Delta}_{i,r}, \end{aligned}$$

and further noting that $\widehat{\Delta}_{i_r,r} \geq \widehat{\Delta}_{i,r}$, we would have $\widehat{\Delta}_{i_r,r} = \widehat{\Delta}_{i,r}$. However, our tie-breaking ensures that this inequality is impossible: since $i_r \in S_r, i \notin S_r$ and i_r is deactivated we have $\widehat{\Delta}_{i_r,r} > \widehat{\Delta}_{i,r}$. Therefore i is not empirically dominated by i_r. Moreover, since $i_r = i^\star$ (by assumption), Lemma 7 (it trivially holds with the modified gaps) yields

$$\Delta^\star_i \leq c\Delta^{(k)}_{(K+1-r)}. \tag{23}$$

Recalling that i_r is deactivated, we have for any arm $j \in A_r$,

$$(\mathrm{M}(i, i_r; r))^+ + (\widehat{\Delta}^\star_{i,r})^+ \geq \widehat{\Delta}_{j,r} \tag{24}$$

On the other side, there exists $a \in A_r$ such that $\Delta^{(k)}_{(K+1-r)} \leq \Delta^{(k)}_a = \Delta_a$ (Lemma 10). Applying Equation (19) and Lemma 5 and taking $j = a$ in Equation 24 yields

$$(\mathrm{M}(i, i_r))^+ + (\Delta^\star_i)^+ \geq (1 - 5c)\Delta^{(k)}_{(K+1-r)}$$

then as $i_r = i^\star$, so $i \prec i_r$, the latter yields

$$\Delta^\star_i \geq (1 - 5c)\Delta^{(k)}_{(K+1-r)}. \tag{25}$$

Both (23) and (25) cannot hold when $c < 1/6$. So for any $i \in A_{r+1}$

$$i^\star \neq i_r.$$

Which achieves the proof of "Step 2". Put together, "Step 1&2" prove that $\mathcal{H}_{r+1}$ holds.

Step 3: Conclusion

We have proved that if $\mathcal{H}_r$ holds and $\alpha(r) < \tilde{k}$ then $\mathcal{H}_{r+1}$ holds. Note that if $i_r \in S_r$ then $\alpha(r+1) = \alpha(r) + 1$ otherwise $\alpha(r+1) = \alpha(r)$. Therefore, no error occurs until the end of round $\min(r', t)$ where r' is such that $\alpha(r'+1) = \tilde{k}$. As a consequence $B_{\tau_k^t+1} = \{a_1, \dots, a_{\alpha(\tau_k^t+1)}\} \subset \mathcal{S}^\star$. And by the proof of "Step 2", if $j \in A_{\tau_p^t}$ is a sub-optimal arm, then $j^\star \in A_{\tau_p^t+1}$. ☐

We can now prove the main theorem of this section.

Theorem 3. *Let $k \in [K]$. EGE-SR-k satisfies*

$$e_{T,k}(\nu) \leq 2(K-1)^2 |\mathcal{S}^\star| D \exp\left(-\frac{T-K}{144\sigma^2 H_2^{(k)}(\nu)\overline{\log}(K)}\right).$$

Proof. The proof is a direct consequence of Lemma 9 and Hoeffding's inequality. Indeed, Lemma 9 yields that EGE-SR-k is correct on the event $\mathcal{E}_c^2$ for $c < 1/6$. Therefore, for any $0 < c < 1/6$

$$e_{T,k}(\nu) \leq \mathbb{P}(\overline{(\mathcal{E}_c^2)}).$$

Letting $0 < c < 1/6$ fixed, by union bound and Hoeffing's inequality (as in the proof of Lemma 8) it follows that

$$\mathbb{P}(\overline{(\mathcal{E}_c^2)}) \leq 2(K-1)^2 |\mathcal{S}^\star| D \exp\left(-c^2 \frac{T-K}{4\sigma^2 H_2^{(k)} \overline{\log}(K)(\nu)}\right)$$

therefore,

$$e_{T,k}(\nu) \leq 2(K-1)^2 |\mathcal{S}^\star| D \exp\left(-c^2 \frac{T-K}{4\sigma^2 H_2^{(k)}(\nu) \overline{\log}(K)}\right)$$

and taking the limit to 1/6 (as it holds for any $0 < c < 1/6$) yields the expected result. ☐

The theorem below bounds the expected stopping time and the number of samples used at stopping.

Theorem 4. *Fix $k < |\mathcal{S}^\star|$ and let $q := K - |\mathcal{S}^\star| + k$. Then*

$$\begin{aligned}
\mathbb{E}[\tau] &\leq q + 2(K-1)|\mathcal{S}^\star|(K-q-1)qD \exp\left(-\frac{T-K}{144\sigma^2 H_2^{(k)}(\nu)\overline{\log}(K)}\right) \quad \textit{and} \\
\mathbb{E}[N_\tau] &\leq N_q + 2(K-1)|\mathcal{S}^\star|(K-q-1)qDT \exp\left(-\frac{T-K}{144\sigma^2 H_2^{(k)}(\nu)\overline{\log}(K)}\right).
\end{aligned}$$

This result suggests that for this relaxed problem, the algorithm might not need to use the whole budget, in particular when T is large. For example, consider a setting $[K] = \mathcal{S}^\star$ then $q = k$ and we roughly use N_k samples which can be way smaller than N_{K-1}.

Proof. The idea is to show that if the algorithm has not stopped after round q then some high probability event must not hold. Let $c < 1/6$ be fixed. We have

$$\begin{aligned}
\mathbb{E}[\tau] &\leq q + \mathbb{E}[\tau \mathbb{I}\{\tau > q\}], \\
&\leq q + \sum_{s=q+1}^{K-1} \mathbb{P}(\tau \geq s).
\end{aligned}$$

We claim that for any $s > q$,

$$\{\tau \geq s\} \subset \overline{(\mathcal{E}_c^2(q))}.$$

Indeed,

$$\begin{aligned}\tau \geq s &\implies \tau > q \\ &\implies \alpha(q+1) < k,\end{aligned}$$

However, by Lemma 9, if $\mathcal{E}_c^2(q)$ holds and $\alpha(q+1) < k$ then no error has occured until the end of round q, therefore $D_q = (\mathcal{S}^\star)^\mathsf{c}$ and $|B_q| = k$, which is not possible as $\alpha(q+1) < k$. So $\tau \geq s \implies (\mathcal{E}_c^2(q)$ does not hold). Then

$$\begin{aligned}\mathbb{E}[\tau] &\leq q + \sum_{s=q+1}^{K-1} \mathbb{P}(\overline{(\mathcal{E}_c(q))}), && (26)\\ &\leq q + \sum_{s=q+1}^{K-1} 2(K-1)|\mathcal{S}^\star| qD \exp\left(-c^2 \frac{T-K}{4\sigma^2 H_2^{(k)}(\nu)\overline{\log}(K)}\right), && (27)\\ &\leq q + 2(K-1)|\mathcal{S}^\star|(K-q-1)qD \exp\left(-c^2 \frac{T-K}{4\sigma^2 H_2^{(k)}(\nu)\overline{\log}(K)}\right). && (28)\end{aligned}$$

Similarly, we have

$$\mathbb{E}[N_\tau] \leq N_q + (N_\tau \mathbb{I}_{\tau > q}).$$

Then,

$$\begin{aligned}\mathbb{E}[N_\tau] &\leq N_q + \sum_{s=q+1}^{K-1} N_s \mathbb{P}(\tau \geq s)\\ &\leq N_q + \sum_{s=q+1}^{K-1} N_s \mathbb{P}(\overline{(\mathcal{E}_c^2(q))})\\ &\leq N_q + 2(K-1)|\mathcal{S}^\star| qD \exp\left(-c^2 \frac{T-K}{4\sigma^2 H_2^{(k)}(\nu)\overline{\log}(K)}\right) \sum_{s=q+1}^{K-1} N_s.\end{aligned}$$

Simple algebra yields for any $r \in [K-1]$,

$$\begin{aligned}N_r &= (K-r)n_r + \sum_{s=1}^{r} n_s,\\ &\leq \frac{T-K}{\overline{\log}(K)} + (K-r) + r + \frac{T-K}{\overline{\log}(K)} \sum_{s=1}^{r} \frac{1}{K+1-s}\\ &= \frac{T-K}{\overline{\log}(K)}\left(1 + \sum_{s=K+1-r}^{K} s^{-1}\right) + K\\ &\leq T.\end{aligned}$$

Therefore,

$$\mathbb{E}[N_\tau] \leq N_q + 2(K-1)|\mathcal{S}^\star|(K-q-1)qDT \exp\left(-c^2 \frac{T-K}{4\sigma^2 H_2^{(k)}(\nu)\overline{\log}(K)}\right). \quad (29)$$

As (28) and (29) holds for any $c < 1/6$ taking the limit for a sequence $c_x \to 1/6$ yields the expected constants in the exponent, which achieves the proof. □

B CONCENTRATION RESULTS

In this section we prove the concentration results used in Appendix A. We recall the definition of the good event:

$$\mathcal{E}_c^1 \quad := \quad \bigcap_{r\in[R]} \bigcap_{i\in\mathcal{S}^\star} \bigcap_{j\in[K]} \left\{ \left\| (\widehat{\boldsymbol{\theta}}_{i,n_r} - \widehat{\boldsymbol{\theta}}_{j,n_r}) - (\boldsymbol{\theta}_i - \boldsymbol{\theta}_j) \right\|_\infty \leq c\Delta_{(\lambda_{r+1}+1)} \right\} .$$

B.1 Proof of Lemma 5

Lemma 5. *On the event $\mathcal{E}_c^1$, for all $r \in [R]$ and $i, j \in A_r$, if $i \in \mathcal{S}^\star$ or $j \in \mathcal{S}^\star$ then*

$$|\mathrm{M}(i,j;r) - \mathrm{M}(i,j)| \leq c\Delta_{(\lambda_{r+1}+1)} \text{ and}$$
$$|\mathrm{m}(i,j;r) - \mathrm{m}(i,j)| \leq c\Delta_{(\lambda_{r+1}+1)}.$$

Proof. Assume $i \in \mathcal{S}^\star$ or $j \in \mathcal{S}^\star$. We have

$$\begin{aligned}
|\mathrm{M}(i,j;r) - \mathrm{M}(i,j)| &= \left| \max_d \left[\widehat{\theta}_{i,n_r}^d - \widehat{\theta}_{j,n_r}^d \right] - \max_d \left[\theta_i^d - \theta_j^d \right] \right|, \\
&\stackrel{(a)}{\leq} \max_d \left| (\widehat{\theta}_{i,n_r}^d - \widehat{\theta}_{j,n_r}^d) - (\theta_i^d - \theta_j^d) \right|, \\
&= \left\| (\widehat{\boldsymbol{\theta}}_{i,n_r} - \widehat{\boldsymbol{\theta}}_{j,n_r}) - (\boldsymbol{\theta}_i - \boldsymbol{\theta}_j) \right\|_\infty, \\
&\stackrel{(b)}{\leq} c\Delta_{(\lambda_{r+1}+1)}.
\end{aligned}$$

where (a) follows by reverse triangle inequality and (b) holds on the event $\mathcal{E}_c^1$. The second part of the lemma follows from

$$|\mathrm{m}(i,j;r) - \mathrm{m}(i,j)| \quad = \quad |-\mathrm{M}(i,j;r) + \mathrm{M}(i,j)|,$$

as $\mathrm{M}(i,j) = -\,\mathrm{m}(i,j)$ and $\mathrm{M}(i,j;r) = -\,\mathrm{m}(i,j;r)$. □

B.2 Proof of Lemma 6

We recall the definition of the property

$$\mathcal{P}_r = \{\forall i \notin \mathcal{S}^\star, i \in A_r \Rightarrow i^\star \in A_r\}.$$

and start by proving a first intermediate result.

Lemma 12. *Assume that $\mathcal{E}_1^c$ holds and let $r \in [R]$ such that $\mathcal{P}_r$ holds. Then for any $i \in A_r$,*

$$(\widehat{\Delta}_{i,r}^\star)^+ - (\Delta_i^\star)^+ \leq 2c\Delta_{(\lambda_{r+1}+1)} \quad \text{ and } \quad (\widehat{\Delta}_{i,r}^\star)^+ - (\Delta_i^\star)^+ \geq -c\Delta_{(\lambda_{r+1}+1)}.$$

Proof. We define the max of an empty set to be $-\infty$. We first analyze the case $i \in \mathcal{S}^\star$. When $i \in \mathcal{S}^\star$, we have

$$(\Delta_i^\star)^+ = 0 = \left(\max_{j\in A_r\setminus\{i\}} \mathrm{m}(i,j) \right)^+,$$

which yields

$$\begin{aligned}
|(\widehat{\Delta}_{i,r}^\star)^+ - (\Delta_i^\star)^+| &= \left| \left(\max_{j\in A_r\setminus\{i\}} \mathrm{m}(i,j;r) \right)^+ - \left(\max_{j\in A_r\setminus\{i\}} \mathrm{m}(i,j) \right)^+ \right|, \\
&\leq \left| \left(\max_{j\in A_r\setminus\{i\}} \mathrm{m}(i,j;r) \right) - \left(\max_{j\in A_r\setminus\{i\}} \mathrm{m}(i,j) \right) \right| \quad (\text{since } |x^+ - y^+| \leq |x-y|), \\
&\leq \max_{j\in A_r\setminus\{i\}} |\mathrm{m}(i,j;r) - \mathrm{m}(i,j)|, \\
&\leq c\Delta_{(\lambda_{r+1}+1)},
\end{aligned}$$

where the last inequality follows from Lemma 5. We now assume that i is a sub-optimal arm. Since $\mathcal{P}_r$ holds at round r, $i^\star \in A_r$ and

$$\Delta_i^\star = \max_{j \in A_r \setminus \{i\}} \mathrm{m}(i, j).$$

Let $\hat{i} \in \mathrm{argmax}_{j \in A_r \setminus \{i\}} \mathrm{m}(i, j; r)$ then

$$\begin{aligned} (\widehat{\Delta}_{i,r}^\star)^+ - (\Delta_i^\star)^+ &= (\mathrm{m}(i, \hat{i}; r))^+ - (\mathrm{m}(i, i^\star))^+, &(30)\\ &\geq (\mathrm{m}(i, i^\star; r))^+ - (\mathrm{m}(i, i^\star))^+ \quad (\text{since } i^\star \in A_r). &(31) \end{aligned}$$

We further note that

$$\begin{aligned} \left|(\mathrm{m}(i, i^\star; r))^+ - (\mathrm{m}(i, i^\star))^+\right| &\leq |\mathrm{m}(i, i^\star; r) - \mathrm{m}(i, i^\star)|, \\ &\leq c\Delta_{(\lambda_{r+1}+1)}, \end{aligned}$$

which follows from Lemma 5. Combining the latter inequality with (31) yields

$$(\widehat{\Delta}_{i,r}^\star)^+ - (\Delta_i^\star)^+ \geq -c\Delta_{(\lambda_{r+1}+1)}. \tag{32}$$

We also have

$$\begin{aligned} (\widehat{\Delta}_{i,r}^\star)^+ - (\Delta_i^\star)^+ &\leq (\mathrm{m}(i, \hat{i}; r))^+ - (\mathrm{m}(i, \hat{i}))^+, \\ &\leq \left|\mathrm{m}(i, \hat{i}; r) - \mathrm{m}(i, \hat{i})\right|, \\ &= \left|\mathrm{M}(i, \hat{i}) - \mathrm{M}(i, \hat{i}; r)\right|, \\ &\leq \left\|(\boldsymbol{\theta}_i - \boldsymbol{\theta}_{\hat{i}}) - (\widehat{\boldsymbol{\theta}}_{i,n_r} - \widehat{\boldsymbol{\theta}}_{\hat{i},n_r})\right\|_\infty, \\ &= \left\|\left((\boldsymbol{\theta}_i - \boldsymbol{\theta}_{i^\star}) - \left(\widehat{\boldsymbol{\theta}}_{i,n_r} - \widehat{\boldsymbol{\theta}}_{i^\star,n_r}\right)\right) + \left((\boldsymbol{\theta}_{i^\star} - \boldsymbol{\theta}_{\hat{i}}) - \left(\widehat{\boldsymbol{\theta}}_{i^\star,n_r} - \widehat{\boldsymbol{\theta}}_{\hat{i},n_r}\right)\right)\right\|_\infty, \\ &\leq 2c\Delta_{(\lambda_{r+1}+1)}, \end{aligned}$$

where the last inequality follows from the triangle inequality and Lemma 5. This proves the lemma. □

We then prove Lemma 2 which can be viewed as a "symmetric" version of Lemma 6.

Lemma 2. *Assume that $\mathcal{E}_c^1$ holds. Let $r \in [R]$ such that $\mathcal{P}_r$ holds. Then, for any sub-optimal arm $i \in A_r$,*

$$|\widehat{\Delta}_{i,r}^\star - \Delta_i^\star| \leq 2c\Delta_{(\lambda_{r+1}+1)}$$

and for any optimal arm $i \in A_r$,

$$\widehat{\delta}_{i,r}^\star \geq \Delta_i - 2c\Delta_{(\lambda_{r+1}+1)}.$$

Proof. Let $i \in A_r \cap (\mathcal{S}^\star)^\mathsf{c}$. By assumption, $i^\star \in A_r$, so

$$\Delta_i := \Delta_i^\star = \max_{j \in A_r \setminus \{i\}} \mathrm{m}(i, j),$$

then we have

$$\begin{aligned} |\widehat{\Delta}_{i,r}^\star - \Delta_i| &= \left|\left(\max_{j \in A_r \setminus \{i\}} \mathrm{m}(i, j; r)\right) - \left(\max_{j \in A_r \setminus \{i\}} \mathrm{m}(i, j)\right)\right|, \\ &\overset{(a)}{\leq} \max_{j \in A_r \setminus \{i\}} |\mathrm{m}(i, j; r) - \mathrm{m}(i, j)|, \\ &\leq \max_{j \in A_r \setminus \{i\}} |\mathrm{M}(i, j) - \mathrm{M}(i, j; r)| \\ &, \\ &\leq \max_{j \in A_r \setminus \{i\}} \left\|(\boldsymbol{\theta}_i - \boldsymbol{\theta}_j) - (\widehat{\boldsymbol{\theta}}_{i,n_r} - \widehat{\boldsymbol{\theta}}_{j,n_r})\right\|_\infty \\ &= \max_{j \in A_r \setminus \{i\}} \left\|\left((\boldsymbol{\theta}_i - \boldsymbol{\theta}_{i^\star}) - \left(\widehat{\boldsymbol{\theta}}_{i,n_r} - \widehat{\boldsymbol{\theta}}_{i^\star,n_r}\right)\right) + \left((\boldsymbol{\theta}_{i^\star} - \boldsymbol{\theta}_j) - \left(\widehat{\boldsymbol{\theta}}_{i^\star,n_r} - \widehat{\boldsymbol{\theta}}_{j,n_r}\right)\right)\right\|_\infty, \\ &\overset{(b)}{\leq} 2c\Delta_{(\lambda_{r+1}+1)}, \end{aligned}$$

where (a) follows by reverse triangle inequality and (b) follows by triangle inequality and Lemma 5. So we have proved the first statement of the lemma.

Before proving the second statement of the lemma, recall that

$$\widehat{\delta}^\star_{i,r} = \min_{j \in A_r \setminus \{i\}} [\mathrm{M}(i,j;r) \ \wedge \ (\mathrm{M}(j,i;r)^+ + (\widehat{\Delta}^\star_{j,r})^+)].$$

Let $i \in A_r \cap \mathcal{S}^\star$. For any $j \in A_r$ we have

$$\begin{aligned} |\mathrm{M}(j,i;r)^+ - \mathrm{M}(j,i)^+| &\leq |\mathrm{M}(j,i;r) - \mathrm{M}(j,i)|, \\ &\leq c\Delta_{(\lambda_{r+1}+1)} \quad \text{(Lemma 5)}, \end{aligned}$$

and similarly

$$|\mathrm{M}(i,j;r) - \mathrm{M}(i,j)| \leq c\Delta_{(\lambda_{r+1}+1)}$$

which put together yields

$$\mathrm{M}(j,i;r)^+ \geq \mathrm{M}(j,i)^+ - c\Delta_{(\lambda_{r+1}+1)} \quad \text{and} \tag{33}$$

$$\mathrm{M}(i,j;r) \geq \mathrm{M}(i,j) - c\Delta_{(\lambda_{r+1}+1)}. \tag{34}$$

Letting $j \in A_r$, we have by Lemma 12

$$(\widehat{\Delta}^\star_{j,r})^+ - (\Delta^\star_j)^+ \geq -c\Delta_{(\lambda_{r+1}+1)} \tag{35}$$

Put together, (35) and (33) yields

$$\mathrm{M}(j,i;r)^+ + (\widehat{\Delta}^\star_{j,r})^+ \geq \mathrm{M}(j,i)^+ + (\Delta^\star_j)^+ - 2c\Delta_{(\lambda_{r+1}+1)}$$

for any $j \in A_r$. Which combined with (34) yields

$$[\mathrm{M}(i,j;r) \wedge (\mathrm{M}(j,i;r)^+ + (\widehat{\Delta}^\star_{j,r})^+)] \geq [\mathrm{M}(i,j) \wedge (\mathrm{M}(j,i)^+ + (\Delta^\star_j)^+)] - 2c\Delta_{(\lambda_{r+1}+1)}$$

for any $j \in A_r$. Therefore,

$$\begin{aligned} \widehat{\delta}^\star_{i,r} &\geq \left(\min_{j \in A_r \setminus \{i\}} [\mathrm{M}(i,j) \wedge (\mathrm{M}(j,i)^+ + (\Delta^\star_j)^+)] \right) - 2c\Delta_{(\lambda_{r+1}+1)}, \\ &\geq \left(\min_{j \in [K] \setminus \{i\}} [\mathrm{M}(i,j) \wedge (\mathrm{M}(j,i)^+ + (\Delta^\star_j)^+)] \right) - 2c\Delta_{(\lambda_{r+1}+1)}, \\ &= \delta^\star_i - 2c\Delta_{(\lambda_{r+1}+1)}. \end{aligned}$$

□

Finally, we prove Lemma 6.

Lemma 6. *Assume that $\mathcal{E}_1^c$ holds and let $r \in [R]$ such that $\mathcal{P}_r$ holds. Then for any $i \in A_r$*

$$\widehat{\Delta}_{i,r} - \Delta_i \geq \begin{cases} -2c\Delta_{(\lambda_{r+1}+1)} & \text{if } i \in \mathcal{S}^\star, \\ -c\Delta_{(\lambda_{r+1}+1)} & \text{else.} \end{cases}$$

Proof. Let $i \in A_r \cap (\mathcal{S}^\star)^\mathsf{c}$. Since $i \in (\mathcal{S}^\star)^\mathsf{c}$, $\Delta_i = \Delta^\star_i$. If $i \notin S_r$ then $\widehat{\Delta}_{i,r} = \widehat{\Delta}^\star_{i,r}$. If $i \in S_r$ then $\widehat{\Delta}_{i,r} = \widehat{\delta}^\star_{i,r}$ and

$$\widehat{\Delta}^\star_{i,r} \leq 0 \leq \widehat{\delta}^\star_{i,r}$$

which follows since $i \in S_r$. Therefore in both cases

$$\begin{aligned} \widehat{\Delta}_{i,r} - \Delta_i &\geq \widehat{\Delta}^\star_{i,r} - \Delta^\star_i \\ &= \max_j \mathrm{m}(i,j;r) - \mathrm{m}(i,i_\star) \\ &\geq \mathrm{m}(i,i_\star;r) - \mathrm{m}(i,i_\star) \\ &\geq -c\Delta_{(\lambda_{r+1}+1)} \end{aligned}$$

where the last inequality uses Lemma 5.

Similarly, let $i \in A_r \cap \mathcal{S}^\star$. We have $\Delta_i = \delta_i^\star$ (Lemma 1) and $\widehat{\Delta}_{i,r} = \widehat{\delta}^\star_{i,r}$ if $i \in S_r$. If $i \in A_r \backslash S_r$ (empirically sub-optimal) then

$$\widehat{\Delta}_{i,r} = \widehat{\Delta}^\star_{i,r} \geq 0$$

and there exists $j \in A_r$ such that $\widehat{\boldsymbol{\theta}}_{i,n_r} \prec \widehat{\boldsymbol{\theta}}_{j,n_r}$ so for this arm j, $\mathrm{M}(i,j;r) < 0$. Therefore $\widehat{\delta}^\star_{i,r} < 0$ and

$$\widehat{\delta}^\star_{i,r} < 0 \leq \widehat{\Delta}_{i,r} = \widehat{\Delta}^\star_{i,r}.$$

Put together we have in all cases

$$\widehat{\Delta}_{i,r} - \Delta_i \geq \widehat{\delta}^\star_{i,r} - \Delta_i \geq -2c\Delta_{(\lambda_{r+1}+1)},$$

where the last inequality follows from Lemma 2. □

B.3 Proof of Lemma 7

Lemma 7. *Let $c > 0$ and assume $\mathcal{E}_c^1$ holds. At round $r \in [R]$, for any sub-optimal arm $i \in A_r$, if $i^\star \in A_r$ and $i^\star$ does not empirically dominate i then $\Delta_i^\star \leq c\Delta_{(\lambda_{r+1}+1)}$.*

Proof. We have by definition

$$\begin{aligned}
\widehat{\boldsymbol{\theta}}_{i,n_r} \nprec \widehat{\boldsymbol{\theta}}_{i^\star,n_r} \quad &\implies \quad \exists\, d\, : \widehat{\theta}^d_{i,n_r} > \widehat{\theta}^d_{i^\star,n_r} \\
&\implies \quad \exists\, d\, : (\widehat{\theta}^d_{i,n_r} - \theta_i^d) - (\widehat{\theta}^d_{i^\star,n_r} - \theta^d_{i^\star}) > \theta^d_{i^\star} - \theta^d_i \geq \mathrm{m}(i, i^\star),
\end{aligned}$$

so

$$\left\| (\widehat{\boldsymbol{\theta}}_{i,n_r} - \widehat{\boldsymbol{\theta}}_{i^\star,n_r}) - (\boldsymbol{\theta}_i - \boldsymbol{\theta}_{i^\star}) \right\|_\infty \geq \Delta_i^\star,$$

which, on the event $\mathcal{E}_c^1$ is only possible if

$$\Delta_i^\star \leq c\Delta_{(\lambda_{r+1}+1)}.$$

□

B.4 Proof of Lemma 8

Lemma 8. *For any $c > 0$,*

$$\mathbb{P}(\overline{(\mathcal{E}_c^1)}) \leq 2|\mathcal{S}^\star|(K-1)DR\exp\left(-\frac{c^2\widetilde{T}^{R,\boldsymbol{t},\boldsymbol{\lambda}}(\nu)}{4\sigma^2}\right).$$

Proof. We have

$$\begin{aligned}
\mathbb{P}(\overline{(\mathcal{E}_c^1)}) \quad &\leq \quad \sum_{r=1}^R \sum_{i\in\mathcal{S}^\star}\sum_{j\neq i}\sum_{d=1}^D \mathbb{P}\left(|(\widehat{\theta}^d_{i,n_r} - \widehat{\theta}^d_{j,n_r}) - (\theta^d_i - \theta^d_j)| > c\Delta_{\lambda_{r+1}+1}\right), \\
&\leq \quad 2|\mathcal{S}^\star|(K-1)D\sum_r \exp\left(-\frac{c^2 n_r \Delta^2_{\lambda_{r+1}+1}}{4\sigma^2}\right) \quad \text{(Hoedding's inequality)}, \\
&\leq \quad 2|\mathcal{S}^\star|(K-1)DR\exp\left(-\frac{c^2\widetilde{T}^{R,\boldsymbol{t},\boldsymbol{\lambda}}(\nu)}{4\sigma^2}\right).
\end{aligned}$$

□

C LOWER BOUND

In this section, we state our lower bounds for fixed-budget PSI that hold for some class of instances.

Illustration of the sub-optimality gaps Before presenting of our lower bounds, we give additional details on the sub-optimality gaps. Recall that for a sub-optimal arm we have

$$\Delta_i := \Delta_i^\star := \max_{j\in\mathcal{S}^\star} \mathrm{m}(i,j), \tag{36}$$

and for an optimal arm $i \in \mathcal{S}^\star$,

$$\Delta_i := \min(\delta_i^+, \delta_i^-) \tag{37}$$

where

$$\begin{aligned}\delta_i^+ &:= \min_{j\in\mathcal{S}^\star\setminus\{i\}} \min(\mathrm{M}(i,j), \mathrm{M}(j,i)) \ , \\ \delta_i^- &:= \min_{j\in[K]\setminus\mathcal{S}^\star} [(\mathrm{M}(j,i))^+ + \Delta_j]\end{aligned}$$

with the convention $\min_\emptyset = +\infty$. First note that when $D = 1$ we recover the classical definition of sub-optimal gaps in single-objective bandit (Kaufmann et al., 2016; Audibert and Bubeck, 2010). Indeed, in this case there is a single optimal arm denoted $a_\star$ and $\mathcal{S}^\star = \{a_\star\}$ then as any arm $a \neq a_\star$ is sub-optimal it comes

$$\begin{aligned}\Delta_a &= \Delta_a^\star = \mathrm{m}(a, a_\star) \\ &= \theta_{a_\star} - \theta_a\end{aligned}$$

and as $\delta_{a_\star}^+ = \infty$ we have

$$\begin{aligned}\Delta_{a_\star} &= \delta_{a_\star}^- = \min_{j\neq a_\star}[(\mathrm{M}(j,a_\star))^+ + \Delta_j] \\ &= \min_{j\neq a_\star} \Delta_j\end{aligned}$$

where the last line follows since every $j \neq a_\star$ is dominated by $a_\star$. Now we illustrate these gaps in higher-dimension. Following Appendix A of Auer et al. (2016) we illustrate below the sub-optimality gaps in dimension $D = 2$.

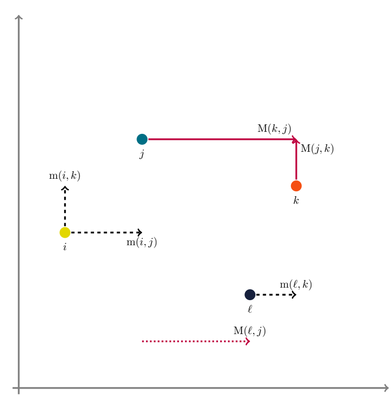

Figure 7: Illustration of the sub-optimality gaps. Plain lines represent "distances" between Pareto optimal arms and dashed lines are for margins from sub-optimal to optimal arms.

In this instance $\Delta_i = \mathrm{m}(i,j)$ and it is easy to see that if i is increased by Δ_i in both x and y axes it will become non-dominated. We also have $\Delta_\ell = \mathrm{m}(\ell,j)$. As ℓ is only dominated by j, if is it translated by $\mathrm{m}(\ell,j)$ on the x-axis it will become Pareto optimal. For Pareto-optimal arms k,j, $\delta_k^+ = \delta_j^+ = \mathrm{M}(j,k)$. As k dominates both i and ℓ its margin to sub-optimal arms is $\delta_k^- = \min(\Delta_i, \Delta_\ell)$ and we have $\delta_j^- = \min(\mathrm{M}(\ell,j) + \Delta_\ell, \Delta_i)$. Finally, observe that for both j,k, $\Delta_j = \Delta_k = \mathrm{M}(j,k)$. If k is translated by $\mathrm{M}(j,k)$ on the y-axis it will dominate j. Similarly, if j is translated by $-\mathrm{M}(j,k)$ on the y-axis, it will be dominated by k.

We will now prove the lower bound. The idea is to build alternative instances where the Pareto set is changed and the complexity is less than that of the base instance. To obtain a different Pareto set we can either make optimal an arm that was sub-optimal or make sub-optimal an arm that was optimal. However in doing so, we need to take care not to decrease the gap of any arm, in particular that of the arm that is shifted, otherwise the alternative instance could be harder. This is challenging as the definition of the gaps are completely different for optimal and sub-optimal arms. In the sequel we restrict to some class of instances.

We define $\mathcal{B}$ to be the set of means $\boldsymbol{\Theta} \in \mathbb{R}^{K\times D}$ such that each sub-optimal arm i is only dominated by a single arm, denoted by $i^\star$ (that has to belong to $\mathcal{S}^\star$) and that for each optimal arm j there exists a unique sub-optimal arm which is dominated by j, denoted by $\underline{j}$. We further assume that optimal arms are not too close to arms they don't dominate: for any sub-optimal arm i and optimal arm j such that $\boldsymbol{\theta}_i \nprec \boldsymbol{\theta}_j$,

$$\mathrm{M}(i,j) \geq 3\max(\Delta_i, \Delta_{\underline{j}}).$$

An instance ν belongs to $\mathcal{B}$ if its means $\boldsymbol{\Theta} = (\boldsymbol{\theta}_1 \dots \boldsymbol{\theta}_K)^\intercal$ belongs to $\mathcal{B}$.

Theorem 2. *Let $\boldsymbol{\Theta} := (\boldsymbol{\theta}_1 \dots \boldsymbol{\theta}_K)^\intercal \in \mathcal{B}$ and $\nu = (\nu_1, \dots, \nu_K)$ where $\nu_i \sim \mathcal{N}(\boldsymbol{\theta}_i, \sigma^2 I)$. For any algorithm $\mathcal{A}$, there exists $i \in \{0, \dots, K\}$ such that*

$$e_T^{\mathcal{A}}(\nu^{(i)}) \geq \frac{1}{4}\exp\left(-\frac{2T}{\sigma^2 H(\nu^{(i)})}\right).$$

Proof. $\nu_1, \dots, \nu_K$ are distributions over $\mathbb{R}^D$ parameterized by their means (resp.) $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_K$. We let $\nu := (\nu_1, \dots, \nu_K)$ and $\boldsymbol{\Theta}^{(0)} := (\boldsymbol{\theta}_1 \dots \boldsymbol{\theta}_K)^\intercal \in \mathbb{R}^{K\times D}$. We denote by $\mathcal{S}^\star(\nu)$ the Pareto optimal set of the bandit instance ν. For all $i \in [K]$, we define $\boldsymbol{\Theta}^{(i)} := (\boldsymbol{\theta}_1^{(i)} \; \dots \; \boldsymbol{\theta}_K^{(i)})^\intercal \in \mathbb{R}^{K\times D}$ (with $\boldsymbol{\theta}^{(i)}$ to be defined later) and we let $\nu^{(i)}$ denote the bandit with parameter $\boldsymbol{\Theta}^{(i)}$. Let $\mathcal{A}$ denote an algorithm for Pareto set identification in the fixed-budget setting. We further assume that for any $i \in [K]$, $\boldsymbol{\Theta}_{(i)}$ differs from $\boldsymbol{\Theta}^{(0)}$ only in line i:

$$\forall a \neq i, \; \boldsymbol{\theta}_a^{(i)} = \boldsymbol{\theta}_a, \tag{38}$$

so $\nu^{(i)}$ and ν only differs in arm i.

We state the following lemma which is a rewriting of Lemma 16 of Kaufmann et al. (2016).

Lemma 13 (Kaufmann et al. (2016))**.** *Let ν and ν' be two bandit instances such that $\mathcal{S}^\star(\nu) \neq \mathcal{S}^\star(\nu')$. Then*

$$\max\left(e_T^{\mathcal{A}}(\nu), e_T^{\mathcal{A}}(\nu')\right) \geq \frac{1}{4}\exp\left(-\sum_{a=1}^K \mathbb{E}_\nu[T_a(T)]\,\mathrm{KL}(\nu_i, \nu_i')\right)$$

For any $i \in [K]$,

$$\boldsymbol{\theta}_i^{(i)} := \begin{cases} \boldsymbol{\theta}_i + 2\Delta_i e_{d_i} & \text{if } i \notin \mathcal{S}^\star(\nu), \\ \boldsymbol{\theta}_i - 2\Delta_i e_{d_{\underline{i}}} & \text{else.} \end{cases} \tag{39}$$

Let for any i, $\Omega(i) := \{j : \boldsymbol{\theta}_i \prec \boldsymbol{\theta}_j\}$ the set of arms that dominate i. We recall that for a sub-optimal arm i,

$$d_i \in \operatorname*{argmin}_d \left[\theta_{i^\star}^d - \theta_i^d\right].$$

We now justify that for any $i \in [K], \mathcal{S}^\star(\nu^{(i)}) \neq \mathcal{S}^\star(\nu)$, by considering two cases:

- If $i \notin \mathcal{S}^\star(\nu)$, as $\Omega(i) = \{i^\star\}$ and i has been increased to be larger than $i^\star$ on d_i, we have that $i \in \mathcal{S}^\star(\nu^{(i)})$.
- If $i \in \mathcal{S}^\star(\nu)$, since there exists $\underline{i}$ such that $\Omega(\underline{i}) = \{i\}$ and $\Delta_i = \Delta_{\underline{i}}^\star$, the shifting ensures that $\boldsymbol{\theta}_{\underline{i}}^{(i)} \nprec \boldsymbol{\theta}_i^{(i)}$ it comes that $\underline{i} \in \mathcal{S}^\star(\nu^{(i)})$.

We remark that in both cases, we have either $\boldsymbol{\theta}_i^{(i)} = \boldsymbol{\theta}_i + 2\Delta_i e_{d_i}$ or $\boldsymbol{\theta}_i^{(i)} = \boldsymbol{\theta}_i - 2\Delta_i e_{d_{\underline{i}}}$. Recalling that for multivariate Gaussian distribution with the same covariance matrix, $\mathrm{KL}\left(\mathcal{N}(\boldsymbol{\theta}, \Sigma), \mathcal{N}(\boldsymbol{\theta}', \Sigma)\right) = \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\theta}')^\intercal \Sigma^{-1}(\boldsymbol{\theta} - \boldsymbol{\theta}')$, we get for all i

$$\mathrm{KL}(\nu_i, \nu_i^{(i)}) \quad = \quad 2\Delta_i^2/\sigma^2.$$

By applying Lemma 13, we have for any $i \in [K]$,

$$\begin{aligned}\max(e_T(\nu), e_T(\nu^{(i)})) &\geq \frac{1}{4}\exp\left(-\mathbb{E}_\nu[T_i(T)]\,\mathrm{KL}(\nu_i, \nu_i^i)\right)\\ &\geq \frac{1}{4}\exp\left(-\frac{2\mathbb{E}_\nu[T_i(T)]\Delta_i^2}{\sigma^2}\right)\end{aligned}$$

We conclude by using the pigeonhole principle: there must exist $i \in [K]$ such that $\mathbb{E}_\nu[T_i(T)] \leq \frac{T}{H(\nu)\Delta_i^2}$. Therefore, there exists $i \in [K]$ such that

$$\max(e_T(\nu), e_T(\nu^{(i)})) \geq \frac{1}{4}\exp\left(-\frac{2T}{\sigma^2 H(\nu)}\right).$$

It remains to show that for any $i \in [K], H(\nu^{(i)}) = H(\nu)$. We introduce the following notation for any arms p, q

$$\begin{aligned}\mathrm{M}^i(p,q) &:= \max_d[(\theta_p^i)^d - (\theta_q^i)^d],\\ \mathrm{m}^i(p,q) &:= -\mathrm{M}^i(p,q),\end{aligned}$$

which are the quantities M and m computed in the bandit $\nu^{(i)}$. We denote by $\mathcal{D}(\nu) := (\mathcal{S}^\star(\nu))^\mathsf{c}$ the set of non-optimal arms in the bandit ν. Letting i be fixed we denote by $\Delta_k^{(i)}$ the sub-optimality gap of arm k in the bandit $\nu^{(i)}$. $\delta_k^{-,(i)}$, $\delta_k^{+,(i)}$ and $\Delta_k^{\star,(i)}$ are respectively the equivalent of δ_k^-, δ_k^+ and $\Delta_k^\star$ computed in the bandit $\nu^{(i)}$. We will prove that

$$\forall\, k \in [K],\ \Delta_k^{(i)} = \Delta_k. \tag{40}$$

We recall the assumptions on the class $\mathcal{B}$:

$$\forall i \in \mathcal{D}(\nu),\ \exists!\text{ arm } \ i^\star \text{ such that } \boldsymbol{\theta}_i \prec \boldsymbol{\theta}_{i^\star}, \tag{41}$$

$$\forall i \in \mathcal{S}^\star(\nu),\ \exists!\ \text{ arm } \underline{i} \text{ such that } \boldsymbol{\theta}_{\underline{i}} \prec \boldsymbol{\theta}_i, \tag{42}$$

$$\forall i \in \mathcal{D}(\nu),\ \forall j \in \mathcal{S}^\star(\nu) \text{ such that } \boldsymbol{\theta}_i \not\prec \boldsymbol{\theta}_j,\ \mathrm{M}(i,j) \geq 3\max(\Delta_{\underline{j}}, \Delta_i). \tag{43}$$

As for any optimal arm i there exists $\underline{i}$ which is only dominated by i, Assumption (42) and (43) yield

$$\forall i, j \in \mathcal{S}^\star(\nu),\ i \neq j,\ \mathrm{M}(i,j) \geq 3\max(\Delta_{\underline{i}}, \Delta_{\underline{j}}) \tag{44}$$

In the sequel, we fix $i \in [K]$ and we prove (40) by analyzing first the case $i \notin \mathcal{S}^\star(\nu)$ then the case $i \in \mathcal{S}^\star(\nu)$.

Step 1: Assume $i \notin \mathcal{S}^\star(\nu)$.

We recall that $\boldsymbol{\theta}_i^{(i)} = \boldsymbol{\theta}_i + 2\Delta_i e_{d_i}$ and for any $j \neq i$, $\boldsymbol{\theta}_i^{(i)} = \boldsymbol{\theta}_i$. Then

$$\forall p \neq i,\ \forall q \neq i,\ \mathrm{M}^i(p,q) = \mathrm{M}(p,q) \text{ and } \mathrm{m}^i(p,q) = \mathrm{m}(p,q). \tag{45}$$

Since only arm i has changed (and increased in one coordinate) from ν to $\nu^{(i)}$, if $j \in \mathcal{D}(\nu)\setminus\{i\}$ then $j \in \mathcal{D}(\nu^{(i)})$. We will prove that for any $j \in \mathcal{S}^\star(\nu)$,

$$\boldsymbol{\theta}_j = \boldsymbol{\theta}_j^{(j)} \not\prec \boldsymbol{\theta}_i^{(i)}. \tag{46}$$

By contradiction, assume an arm j exists such that (46) holds. Then $\mathrm{M}(j,i) \leq 2\Delta_i$ and

$$\mathrm{M}(j, i^\star) \leq \mathrm{M}(j,i) \leq 2\Delta_i,$$

which is not possible by (44) as $j, i^\star \in \mathcal{S}^\star(\nu)$. Therefore, (46) does not hold and we deduce:

$$\mathcal{D}(\nu^{(i)}) = \mathcal{D}(\nu)\setminus\{i\} \quad \text{ and } \quad \mathcal{S}^\star(\nu^{(i)}) = \mathcal{S}^\star(\nu)\cup\{i\}.$$

Moreover, we will show that

$$\forall j \in \mathcal{D}(\nu^{(i)}), \Delta_j^{\star,(i)} = \Delta_j^\star.$$

Indeed, as for any $j \in \mathcal{D}(\nu^{(i)})$, $j, j^\star$ have not been shifted, we have

$$\Delta_j^{\star,(i)} \geq \Delta_j^\star$$

and if the inequality is strict then, since only i has changed,

$$\Delta_j^{\star,(i)} = \mathrm{m}^i(j,i) > \Delta_j^\star,$$

which yields (as only d_i has been increased)

$$\mathrm{m}(j,i) > -\Delta_j^\star,$$

so

$$\mathrm{M}(j,i) < \Delta_j^\star,$$

which combined with $\mathrm{M}(j,i^\star) \leq \mathrm{M}(j,i)$ yields

$$\mathrm{M}(j,i^\star) < \Delta_j^\star$$

which is not possible by assumption (43). In short, we have proved that for any $j \in \mathcal{D}(\nu^{(i)})$,

$$\Delta_j^{\star,(i)} = \Delta_j^\star.$$

Let $j \in \mathcal{S}^\star(\nu^{(i)})\backslash\{i\}$. For any $k \neq i$

$$\mathrm{M}^i(j,k) = \mathrm{M}(j,k) \text{ and } \mathrm{M}^i(k,j) = \mathrm{M}(k,j).$$

We have

$$\begin{aligned}
\mathrm{M}^i(i^\star, i) &= \left(\max_{d\neq d_i}\left[\theta_{i^\star}^d - \theta_i^d\right]\right) \vee \left(\theta_{i^\star}^{d_i} - \theta_i^{d_i} - 2\left(\theta_{i^\star}^{d_i} - \theta_i^{d_i}\right)\right), \\
&= \left(\max_{d\neq d_i}\left[\theta_{i^\star}^d - \theta_i^d\right]\right) \vee \left(\theta_i^{d_i} - \theta_{i^\star}^{d_i}\right), \\
&\overset{(a)}{=} \max_{d\neq d_i}\left[\theta_{i^\star}^d - \theta_i^d\right], \\
&\overset{(b)}{=} \mathrm{M}(i^\star, i), \\
&\overset{(c)}{\geq} \Delta_i^\star = \Delta_{i^\star} \qquad (47)
\end{aligned}$$

where (a) follows since $(\theta_i^{d_i} - \theta_{i^\star}^{d_i}) < 0$ while $\mathrm{M}^i(i^\star, i) > 0$, (b) follows from the definition of d_i as the argmin and (c) follows from the inequality $\mathrm{M}(p,q) \geq \mathrm{m}(q,p)$ for any p, q.

Additionally for $j \in \mathcal{S}^\star(\nu^{(i)})\backslash\{i, i^\star\}$,

$$\begin{aligned}
\mathrm{M}^i(j,i) &\geq \mathrm{M}(j,i) - 2\Delta_i, &(48)\\
&\geq \mathrm{M}(j,i^\star) - 2\Delta_i, &(49)\\
&\geq \frac{1}{3}\mathrm{M}(j,i^\star), \quad (\text{since } \mathrm{M}(j,i^\star) \geq 3\Delta_i \text{ by } (44)) &(50)\\
&\geq \max(\Delta_j, \Delta_i) \geq \Delta_j. &(51)
\end{aligned}$$

Combining (51) with (47) yields for $j \in \mathcal{S}^\star(\nu^{(i)})\backslash\{i\}$,

$$\mathrm{M}^i(j,i) \geq \Delta_j. \qquad (52)$$

We now compute $\mathrm{M}^i(i,j)$. Direct calculation yields

$$\mathrm{M}^i(i,i^\star) = \Delta_i = \Delta_{i^\star},$$

and if $j \neq i^\star$,

$$\begin{aligned}
\mathrm{M}^i(i,j) &\geq \mathrm{M}(i,j), &(53)\\
&\geq \max(\Delta_j, \Delta_i) \quad (\text{from } (43)). &(54)
\end{aligned}$$

With what precedes we state the following for any $j \in \mathcal{S}^\star(\nu^{(i)})\backslash\{i\}$:

a) $\min_{k\in\mathcal{S}^\star(\nu^{(i)})\backslash\{j\}} \mathrm{M}^i(j,k) = \left(\min_{k\in\mathcal{S}^\star(\nu)\backslash\{j\}} \mathrm{M}(j,k)\right) \wedge \mathrm{M}^i(j,i)$,

b) $\min_{k\in\mathcal{S}^\star(\nu^{(i)})\backslash\{j\}} \mathrm{M}^i(k,j) = \left(\min_{k\in\mathcal{S}^\star(\nu)\backslash\{j\}} \mathrm{M}(k,j)\right) \wedge \mathrm{M}^i(i,j)$, and

c) $\min_{k\in\mathcal{D}(\nu^{(i)})} \left[(\mathrm{M}^i(k,j))^+ + \Delta_k^{\star,(i)}\right] = \left(\min_{k\in\mathcal{D}(\nu)\backslash\{i\}} [\mathrm{M}(k,j)^+ + \Delta_k^\star]\right)$.

Which combined with $\delta_j^+ \geq \Delta_j$ and (54) and (52) yields

$$\delta_j^{+,(i)} = \min(\delta_j^+, \min(\mathrm{M}^i(i,j), \mathrm{M}^i(j,i))) \geq \Delta_j,$$

and for $j \neq i^\star$,

$$\delta_j^{-,(i)} = \Delta_j$$

As a consequence, for $j \neq i^\star$, $\Delta_j^{(i)} = \Delta_j$. For $j = i^\star$, $\mathrm{M}^i(i,i^\star) = \Delta_i = \Delta_{i^\star}$ and by (47)

$$\mathrm{M}^i(i^\star, i) \quad \geq \quad \Delta_i^\star = \Delta_{i^\star}.$$

Therefore, $\min(\mathrm{M}^i(i,i^\star), \mathrm{M}^i(i^\star,i)) = \Delta_{i^\star}$ and as $\delta_{i^\star}^+ \geq \Delta_{i^\star}$, we have

$$\delta_{i^\star}^{+,(i)} = \min(\delta_{i^\star}^+, \min(\mathrm{M}^i(i,i^\star), \mathrm{M}^i(i^\star,i))) = \Delta_{i^\star},$$

and

$$\delta_{i^\star}^{-,(i)} > \Delta_{i^\star}, \quad \text{(from (43))}$$

so

$$\Delta_{i^\star}^{(i)} = \Delta_{i^\star}$$

To sum up, so far we have proved that for any $j \neq i$,

$$\Delta_j^{(i)} = \Delta_j.$$

Proving that $\Delta_i^{(i)} = \Delta_i$ will conclude the proof for Step 1.

Recall that

$$\Delta_i^{(i)} := \left(\min_{j\in\mathcal{S}^\star(\nu)} \min(\mathrm{M}^i(i,j), \mathrm{M}^i(j,i))\right) \wedge \left(\min_{j\in\mathcal{D}(\nu)\backslash\{i\}} [(\mathrm{M}^i(j,i))^+ + \Delta_j^{\star,(i)}]\right),$$

with $\min(\mathrm{M}^i(i,i^\star), \mathrm{M}^i(i^\star,i)) = \Delta_i$ and by (54) and (51):

$$\forall j \in \mathcal{S}^\star(\nu) \setminus \{i^\star\}, \min(\mathrm{M}^i(i,j), \mathrm{M}^i(j,i)) \geq \Delta_i.$$

Moreover, for any $j \in \mathcal{D}(\nu)\backslash\{i\}$,

$$\mathrm{M}^i(j,i) \geq \mathrm{M}(j,i^\star) - 2\Delta_i \overset{(43)}{\geq} \Delta_i$$

Therefore, $\delta_i^{-,(i)} \geq \Delta_i$ and $\delta_i^{+,(i)} = \Delta_i$. We conclude that

$$\Delta_i^{(i)} = \Delta_i.$$

which achieves the proof of Step 1: If $i \notin \mathcal{S}^\star(\nu)$

$$H(\nu^{(i)}) = H(\nu).$$

Step 2: Assume $i \in \mathcal{S}^\star(\nu)$.

We have $\boldsymbol{\theta}_i^{(i)} = \boldsymbol{\theta}_i - 2\Delta_i e_{d_{\underline{i}}}$ and for any $j \neq i$, $\boldsymbol{\theta}_j^{(i)} = \boldsymbol{\theta}_j$. First, note that $\underline{i} \in \mathcal{S}^\star(\nu^{(i)})$. This is immediate as i is the only arm such that $\boldsymbol{\theta}_{\underline{i}} \prec \boldsymbol{\theta}_i$ and due to the shifting, $\boldsymbol{\theta}_{\underline{i}}^{(i)} \not\prec \boldsymbol{\theta}_i^{(i)}$. We claim that $i \in \mathcal{S}^\star(\nu^{(i)})$. Otherwise, there exists $j \in \mathcal{S}^\star(\nu)$ such that $\mathrm{M}(i,j) \leq 2\Delta_i$ which is not possible by (44). For any $j \in \mathcal{D}(\nu) \setminus \{\underline{i}\}$, j and $j^\star$ have not been shifted, so $j \in \mathcal{D}(\nu^{(i)})$ and similarly to Step 1, we have

$$\mathcal{S}^\star(\nu^{(i)}) = \mathcal{S}^\star(\nu) \cup \{\underline{i}\} \text{ and } \mathcal{D}(\nu^{(i)}) = \mathcal{D}(\nu)\backslash\{\underline{i}\},$$

and for any $j \in \mathcal{D}(\nu^{(i)})$,

$$\Delta_j^{\star,(i)} = \Delta_j^\star.$$

Let $j \in \mathcal{S}^\star(\nu^{(i)})\backslash\{i, \underline{i}\}$. We have,

$$\delta_j^{+,(i)} = \left(\min_{k \in \mathcal{S}^\star(\nu)\backslash\{j,i\}} [\min(\mathrm{M}(k,j), \mathrm{M}(j,k))]\right) \wedge \underbrace{(\min(\mathrm{M}(j,\underline{i}), \mathrm{M}(\underline{i},j)))}_{\alpha_j} \wedge \underbrace{(\min(\mathrm{M}^i(j,i), \mathrm{M}^i(i,j)))}_{\beta_j}.$$

We have

$$\begin{aligned} \mathrm{M}^i(j,i) &\geq \mathrm{M}(j,i) \\ &\geq 3\max(\Delta_i, \Delta_j) \quad \text{(from (44))}, \end{aligned}$$

and

$$\begin{aligned} \mathrm{M}^i(i,j) &\geq \mathrm{M}(i,j) - 2\Delta_i \\ &\geq \max(\Delta_j, \Delta_i) \quad \text{(from (44))}. \end{aligned}$$

It follows that

$$\beta_j \geq \max(\Delta_j, \Delta_i). \tag{55}$$

On the other side,

$$\begin{aligned} \mathrm{M}^i(j,\underline{i}) &= \mathrm{M}(j,\underline{i}), \\ &\geq \mathrm{M}(j,i) \quad (\text{as } \boldsymbol{\theta}_{\underline{i}} \prec \boldsymbol{\theta}_i), \\ &\geq 3\max(\Delta_j, \Delta_i) \quad \text{(from (44))} \end{aligned}$$

and

$$\mathrm{M}^i(\underline{i},j) = \mathrm{M}(\underline{i},j) \geq \max(\Delta_j, \Delta_{\underline{i}}) \quad \text{(from (43))}$$

Therefore,

$$\alpha_j \geq \max(\Delta_j, \Delta_{\underline{i}}), \tag{56}$$

which yields

$$\delta_j^{+,(i)} \geq \Delta_j.$$

As $\underline{j} \in \mathcal{D}(\nu^{(i)})$, we check that

$$\delta_j^{-,(i)} = \min_{k \in \mathcal{D}(\nu)\backslash\{\underline{i}\}} \left[\mathrm{M}(k,j)^+ + \Delta_k^\star\right] = \Delta_{\underline{j}}^\star = \Delta_j.$$

Put together, we have for any $j \in \mathcal{S}^\star(\nu^{(i)})\backslash\{i, \underline{i}\}$,

$$\Delta_j^{(i)} := \min(\delta_j^{+,(i)}, \delta_j^{-,(i)}) = \Delta_j.$$

We also easily check that

$$\delta_i^{-,(i)} = \min_{k \in \mathcal{D}(\nu)\backslash\{\underline{i}\}} \left[\mathrm{M}(k,i)^+ + \Delta_k^\star\right] > \Delta_{\underline{i}} = \Delta_i \tag{57}$$

and

$$\delta_i^{+,(i)} = \left(\min_{k \in \mathcal{S}^\star(\nu) \setminus \{i\}} \left[\min(\mathrm{M}^i(k,i), \mathrm{M}^i(i,k)) \right] \right) \wedge \underbrace{\left(\min(\mathrm{M}^i(i,\underline{i}), \mathrm{M}^i(\underline{i},i)) \right)}_{\beta_{\underline{i}}}. \tag{58}$$

We further note that $\mathrm{M}^i(\underline{i}, i) = \Delta_i$ and similarly to (47),

$$\begin{aligned}
\mathrm{M}^i(i,\underline{i}) &= \left(\max_{d \neq d_{\underline{i}}} \left[\theta_i^d - \theta_{\underline{i}}^d \right] \right) \vee \left(\theta_i^{d_{\underline{i}}} - \theta_{\underline{i}}^{d_{\underline{i}}} - 2\left(\theta_i^{d_{\underline{i}}} - \theta_{\underline{i}}^{d_{\underline{i}}} \right) \right), \\
&= \left(\max_{d \neq d_{\underline{i}}} \left[\theta_i^d - \theta_{\underline{i}}^d \right] \right) \vee \left(\theta_{\underline{i}}^{d_{\underline{i}}} - \theta_i^{d_{\underline{i}}} \right), \\
&\overset{(a)}{=} \max_{d \neq d_{\underline{i}}} \left[\theta_i^d - \theta_{\underline{i}}^d \right], \\
&\overset{(b)}{=} \mathrm{M}(i,\underline{i}), \\
&\overset{(c)}{\geq} \Delta_{\underline{i}} = \Delta_i
\end{aligned}$$

where (a) follows since $(\theta_{\underline{i}}^{d_{\underline{i}}} - \theta_i^{d_{\underline{i}}}) < 0$ while $\mathrm{M}^i(i,\underline{i}) > 0$, (b) follows from the definition of $d_{\underline{i}}$ as the argmin and (c) follows from the inequality $\mathrm{M}(p,q) \geq \mathrm{m}(q,p)$ for any p, q. Therefore we have $\beta_{\underline{i}} = \Delta_i$ and from (55), (56) and (57),

$$\delta_i^{+,(i)} = \Delta_i \text{ and } \delta_i^{-,(i)} > \Delta_i$$

so that finally

$$\Delta_i^{(i)} = \Delta_i.$$

It remains to compute $\Delta_{\underline{i}}^{(i)}$. We have

$$\begin{aligned}
\delta_{\underline{i}}^{+,(i)} &= \left(\min_{k \in \mathcal{S}^\star(\nu) \setminus \{\underline{i}, i\}} \left[\min(\mathrm{M}(k,\underline{i}), \mathrm{M}(\underline{i},k)) \right] \right) \wedge \underbrace{\left(\min(\mathrm{M}^i(\underline{i},i), \mathrm{M}^i(i,\underline{i})) \right)}_{\beta_{\underline{i}}}, \\
&= \left(\min_{k \in \mathcal{S}^\star(\nu) \setminus \{\underline{i}, i\}} \alpha_k \right) \wedge \beta_{\underline{i}}, \\
&= \Delta_{\underline{i}},
\end{aligned}$$

which follows from (56) and the fact that $\beta_{\underline{i}} = \Delta_{\underline{i}}$. On the other side,

$$\delta_{\underline{i}}^{-,(i)} = \min_{k \in \mathcal{D}(\nu) \setminus \{\underline{i}\}} \left[\mathrm{M}(k,\underline{i})^+ + \Delta_k^\star \right] \geq 3\Delta_{\underline{i}} \quad \text{(from (43))}.$$

Therefore,

$$\Delta_{\underline{i}}^{(i)} = \Delta_{\underline{i}},$$

which concludes the proof. □

The conditions of Theorem 2 include a large class of instances. In particular, for any value of D, on can find instances in $\mathcal{B}$. We illustrate in Fig.8 to Fig.12 below an instance $\nu \in \mathcal{B}$ with 4 arms in dimension 2, and the 4 corresponding alternative instances $\nu^{(i)}$. Similar examples could be given for any dimension and any number of arms.

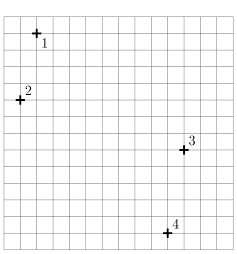

Figure 8: Instance $\nu \in \mathcal{B}$

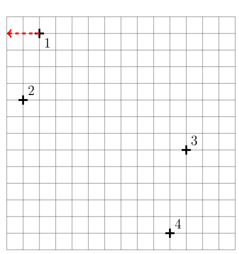

Figure 9: Instance $\nu^{(1)}$

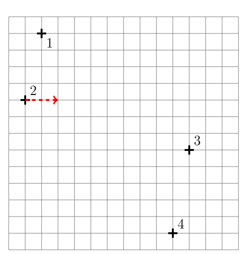

Figure 10: Instance $\nu^{(2)}$

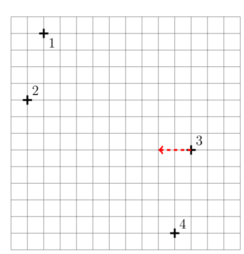

Figure 11: Instance $\nu^{(3)}$

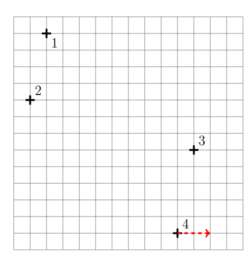

Figure 12: Instance $\nu^{(4)}$

We propose below a larger class $\mathcal{B}'$. Before stating our result we recall some notation. We recall that $e_1, \dots, e_D$ is the canonical basis of $\mathbb{R}^D$. We identify a bandit ν with its means $\boldsymbol{\Theta} := (\boldsymbol{\theta}_1 \dots \boldsymbol{\theta}_K)^\intercal \in \mathbb{R}^{K\times D}$. For any arm i, let $\Omega(i) := \{j \in [K] : \boldsymbol{\theta}_i \prec \boldsymbol{\theta}_j\}$ the set of arms that dominate i and $\Omega^\star(i) := \{j \in [K] : \mathrm{m}(i,j) = \Delta_i^\star\}$ the arms that "maximally" dominate i. Let $\Pi(i) := \{i \in \mathcal{D}(\nu) : \Omega^\star(i) = \{i\}\}$ the set of arms that are "only maximally dominated by i". For a bandit ν we let $\mathcal{D}(\nu)$ or simply $\mathcal{D}$ denote the set of sub-optimal arms. Let $\mathcal{B}'$ be the set of instances such that :

$$\text{(1)} \quad \forall i \in \mathcal{D}(\nu) \text{ there exists } d_i \text{ such that } \boldsymbol{\theta}_i + \Delta_i e_{d_i} \text{ is Pareto optimal w.r.t } \{\boldsymbol{\theta}_j : j \in [K] \setminus \{i\}\}, \tag{59}$$

$$\text{(2)} \quad \forall i \in \mathcal{S}^\star(\nu) \text{ there exists } \underline{i} \in \mathcal{D} \text{ such that } \Pi(i) = \{\underline{i}\} \text{ and } \Omega(\underline{i}) = \{i\}, \tag{60}$$

$$\text{(3)} \quad \forall i,j \in \mathcal{S}^\star(\nu), \mathrm{M}(i,j) \geq 3\max(\Delta_{\underline{i}}, \Delta_{\underline{j}}), \tag{61}$$

$$\text{(4)} \quad \forall i,j, \text{ if } \boldsymbol{\theta}_i \not\prec \boldsymbol{\theta}_j \text{ then } \mathrm{M}(i,j) \geq 3\max(\Delta_i, \Delta_j) \tag{62}$$

Let $\nu := (\nu_1, \dots, \nu_K)$ be an instance whose means $\boldsymbol{\Theta} \in \mathcal{B}$ and such that $\nu_i \sim \mathcal{N}(\boldsymbol{\theta}_i, \sigma^2 I)$. For every $i \in [K]$ we define the alternative instance $\nu^{(i)} := (\nu_1, \dots \nu_i^{(i)}, \dots, \nu_K)$ in which only the mean of arm i is modified to:

$$\boldsymbol{\theta}_i^{(i)} := \begin{cases} \boldsymbol{\theta}_i - 2\Delta_i e_{d_{\underline{i}}} & \text{if } i \in \mathcal{S}^\star(\nu), \\ \boldsymbol{\theta}_i + 2\Delta_i e_{d_i} & \text{else,} \end{cases} \tag{63}$$

where $e_1, \dots, e_D$ denotes the canonical basis of $\mathbb{R}^D$ and $d_i := \mathrm{argmin}_d[\theta_{i^\star}^d - \theta_i^d]$. With $\nu^{(0)} := \nu$, we prove the following.

Theorem 5. *Let $\boldsymbol{\Theta} := (\boldsymbol{\theta}_1 \dots \boldsymbol{\theta}_K)^\intercal \in \mathcal{B}'$ and $\nu = (\nu_1, \dots, \nu_K)$ where $\nu_i \sim \mathcal{N}(\boldsymbol{\theta}_i, \sigma^2 I)$. For any algorithm $\mathcal{A}$, there exists $i \in \{0, \dots, K\}$ such that $H(\nu^{(i)}) \leq H(\nu)$ and*

$$e_T^{\mathcal{A}}(\nu^{(i)}) \geq \frac{1}{4}\exp\left(-\frac{2T}{\sigma^2 H(\nu^{(i)})}\right).$$

Proof. The proof is similar to the proof of Theorem 2. In particular we use the notation introduced therein: the superscript i means that the quantity is computed on the means of bandit $\nu^{(i)}$ i.e $\{\boldsymbol{\theta}_i^{(1)}, \dots, \boldsymbol{\theta}^{(K)}\}$. Proceeding similarly to to the proof of Theorem 2 we have : there exists $i \in [K]$ such that

$$\max(e_T(\nu), e_T(\nu^{(i)})) \geq \frac{1}{4}\exp\left(-\frac{2T}{\sigma^2 H(\nu)}\right).$$

It remains to show that for each alternative bandit $\nu^{(i)}$, $\mathcal{S}^\star(\nu^{(i)}) \neq \mathcal{S}^\star(\nu)$ and $H(\nu^{(i)}) \leq H(\nu)$. To do so we analyze first an instance $\nu^{(i)}$ for $i \in \mathcal{D}(\nu)$.

Step 1: Assume $i \notin \mathcal{S}^\star(\nu)$.

As only arm i has changed in $\nu^{(i)}$ w.r.t ν and $\boldsymbol{\theta}_i^{(i)} = \boldsymbol{\theta}_i + 2\Delta_i e_{d_i}$, letting $j \in \mathcal{D}(\nu) \backslash \{i\}$ j and $j^\star$ has not changed, thus $j \in \mathcal{D}(\nu^{(i)})$ and

$$\Delta_j^{(i)} \geq \Delta_j \tag{64}$$

Let $j \in \mathcal{S}^\star(\nu)$, we claim that $j \in \mathcal{S}^\star(\nu^{(i)})$. Indeed, if $j \notin \mathcal{S}^\star(\nu^{(i)})$ then

$$\boldsymbol{\theta}_j \prec \boldsymbol{\theta}_i + 2\Delta_i e_{d_i},$$

so $\mathrm{M}(j,i) \leq 2\Delta_i$. However $\mathrm{M}(j,i^\star) \leq \mathrm{M}(j,i)$ which yields

$$\mathrm{M}(j,i^\star) \leq 2\Delta_i,$$

which is not possible by (61). Therefore $j \in \mathcal{S}^\star(\nu^{(i)})$ and we will show that $\Delta_j^{(i)} \geq \Delta_j$. Direct calculation yields

$$\delta_j^{+,(i)} = \min(\delta_j^+, \min(\mathrm{M}^i(i,j), \mathrm{M}^i(j,i))) \tag{65}$$

and we have

$$\mathrm{M}^i(i,j) = \left(\max_{d\neq d_i}[\theta_i^d - \theta_j^d]\right) \vee \left(\theta_i^{d_i} - \theta_j^{d_i} + 2\Delta_i\right). \tag{66}$$

If $\boldsymbol{\theta}_i \prec \boldsymbol{\theta}_j$ then since $\boldsymbol{\theta}_i + \Delta_i e_{d_i}$ is non-dominated w.r.t $\{\boldsymbol{\theta}_k : k \in [K]\backslash\{i\}\}$ it follows that

$$\theta_i^{d_i} + \Delta_i \geq \theta_j^{d_i},$$

so

$$\left(\theta_i^{d_i} - \theta_j^{d_i} + 2\Delta_i\right) \geq \Delta_i,$$

which put back in (66) yields $\mathrm{M}^i(i,j) \geq \Delta_i$. And, as $\boldsymbol{\theta}_i \prec \boldsymbol{\theta}_j$, $\Delta_j \leq \Delta_i$ so

$$\mathrm{M}^i(i,j) \geq \Delta_j.$$

In case $\boldsymbol{\theta}_i \not\prec \boldsymbol{\theta}_j$, since $\mathrm{M}^i(i,j) \geq \mathrm{M}(i,j)$ and

$$\mathrm{M}(i,j) \geq \Delta_i$$

by condition (62) (as $\boldsymbol{\theta}_i \not\prec \boldsymbol{\theta}_j$ and $j \in \mathcal{S}^\star(\nu)$). Therefore in all cases

$$\mathrm{M}^i(i,j) \geq \max(\Delta_j, \Delta_i). \tag{67}$$

Similarly to (47) $\mathrm{M}^i(i^\star, i) \geq \Delta_{i^\star}$ and for $j \neq i^\star$

$$\begin{aligned}
\mathrm{M}^i(j,i) &\geq \mathrm{M}(j,i) - 2\Delta_i, &(68)\\
&\geq \mathrm{M}(j,i^\star) - 2\Delta_i, &(69)\\
&\geq \frac{1}{3}\mathrm{M}(j,i^\star) \quad (61) &(70)\\
&\geq \max(\Delta_{\underline{j}}, \Delta_i) \geq \Delta_j. &(71)
\end{aligned}$$

Combining (65), (67) and (71) yields

$$\delta_j^{+,(i)} \geq \Delta_j$$

On the other side,

$$\begin{aligned}
\delta_j^{-,(i)} &:= \min_{k\in\mathcal{D}(\nu^{(i)})}\left[\mathrm{M}^i(k,j)^+ + \Delta_k^{(i)}\right],\\
&= \min_{k\in\mathcal{D}(\nu)\backslash\{i\}}\left[\mathrm{M}(k,j)^+ + \Delta_k^{(i)}\right],\\
&\geq \min_{k\in\mathcal{D}(\nu)\backslash\{i\}}\left[\mathrm{M}(k,j)^+ + \Delta_k\right],\\
&\geq \delta_j^- \geq \Delta_j
\end{aligned}$$

Therefore, for any $j \in \mathcal{S}^\star(\nu)$,

$$\Delta_j^{(i)} = \min(\delta_j^{+,(i)}, \delta_j^{-,(i)}) \geq \Delta_j. \tag{72}$$

So far we have proved that for any $j \neq i$,

$$\Delta_j^{(i)} \geq \Delta_j.$$

It remains to compute $\Delta_i^{(i)}$. Recall that $i \in \mathcal{D}(\nu)$ but $i \in \mathcal{S}^\star(\nu^{(i)})$ so the expression of its gap has completely changed and we have to check that it does not decrease. We have

$$\Delta_i^{(i)} = \min(\delta_i^{+,(i)}, \delta_i^{-,(i)})$$

where

$$\begin{aligned}\delta_i^{+,(i)} &:= \min_{k\in\mathcal{S}^\star(\nu^{(i)})\backslash\{i\}} \min(\mathrm{M}^i(i,j), \mathrm{M}^i(j,i)),\\ &= \min_{k\in\mathcal{S}^\star(\nu)} \min(\mathrm{M}^i(i,j), \mathrm{M}^i(j,i)),\end{aligned}$$

and by (67) and (71) $\min(\mathrm{M}^i(i,k), \mathrm{M}^i(k,i)) \geq \Delta_i$ for any $k\in\mathcal{S}^\star(\nu)$ so

$$\delta_i^{+,(i)} \geq \Delta_i. \tag{73}$$

Let $k\in\mathcal{D}(\nu)\backslash\{i\}$. If $\boldsymbol{\theta}_k \prec \boldsymbol{\theta}_i$ then $\Delta_i \leq \Delta_k$ and

$$(\mathrm{M}^i(k,i))^+ + \Delta_k^{(i)} \geq \Delta_k^{(i)} \geq \Delta_k$$

so $(\mathrm{M}^i(k,i))^+ + \Delta_k^{(i)} \geq \Delta_i$. If $\boldsymbol{\theta}_k \not\prec \boldsymbol{\theta}_i$ we have by (62) $\mathrm{M}(k,i)\geq 3\Delta_i$ and $\mathrm{M}^i(k,i)\geq \mathrm{M}(k,i) - 2\Delta_i$ so

$$(\mathrm{M}^i(k,i))^+ + \Delta_k^{(i)} \geq \Delta_i.$$

Therefore

$$\delta_i^{-,(i)} := \min_{k\in\mathcal{D}(\nu^{(i)})} \left[(\mathrm{M}^i(k,i))^+ + \Delta_k^{(i)}\right], \tag{74}$$

$$\begin{aligned}\delta_i^{-,(i)} &= \min_{k\in\mathcal{D}(\nu)\backslash\{i\}} \left[(\mathrm{M}^i(k,i))^+ + \Delta_k^{(i)}\right], &(75)\\ &\geq \Delta_i. &(76)\end{aligned}$$

Combining (73) and (76) yields

$$\Delta_i^{(i)} \geq \Delta_i.$$

Put together with what precedes yields for any arm k

$$\Delta_k^{(i)} \geq \Delta_k,$$

therefore

$$H(\nu^{(i)}) \leq H(\nu).$$

Step 2: Assume $i\in\mathcal{S}^\star(\nu)$.

We have $\underline{i}\in\Pi(i)$. If $j\in\mathcal{D}(\nu)\cap(\Pi(i))^\mathsf{c}$, as j and $j^\star$ have not been modified from ν to $\nu^{(i)}$ and the only change has decreased i in one coordinate, we have $j\in\mathcal{D}(\nu^{(i)})$ and

$$\Delta_j^{(i)} = \Delta_i.$$

Direct calculation shows that : $\mathcal{S}^\star(\nu^{(i)}) = \mathcal{S}^\star(\nu)\cup\{\underline{i}\}$ and $\mathcal{D}(\nu^{(i)}) = \mathcal{D}(\nu)\backslash\{\underline{i}\}$. Since $\Pi(i) = \{\underline{i}\}$, for any $k\in\mathcal{D}(\nu)\backslash\{\underline{i}\}$,

$$\Delta_k^{(i)} = \Delta_k.$$

The rest of the proof is similar to "Step 2" in the proof of Theorem 2 but we recall the main arguments to make this part self-contained. For any $j\in\mathcal{S}^\star(\nu^{(i)})\backslash\{i,\underline{i}\}$. We have,

$$\delta_j^{+,(i)} = \left(\min_{k\in\mathcal{S}^\star(\nu)\backslash\{j,i\}} [\min(\mathrm{M}(k,j),\mathrm{M}(j,k))]\right) \wedge \underbrace{(\min(\mathrm{M}(j,\underline{i}),\mathrm{M}(\underline{i},j)))}_{\alpha_j} \wedge \underbrace{(\min(\mathrm{M}^i(j,i),\mathrm{M}^i(i,j)))}_{\beta_j}.$$

As $\boldsymbol{\theta}_{\underline{i}} \not\prec \boldsymbol{\theta}_j$ and $\boldsymbol{\theta}_j \not\prec \boldsymbol{\theta}_{\underline{i}}$,

$$\alpha_j \geq \Delta_j. \quad \text{(by (62))}$$

Similarly

$$\mathrm{M}^i(j,i) \geq \mathrm{M}(j,i) \geq \max(\Delta_j,\Delta_i) \tag{77}$$

where the last inequality follows since $i, j \in \mathcal{S}^\star(\nu)$ and from (61). On the other side ,

$$\begin{aligned} \mathrm{M}^i(i,j) &\geq \mathrm{M}(i,j) - 2\Delta_i, &(78)\\ &\geq \frac{1}{3}\mathrm{M}(i,j) \quad ((61)), &(79)\\ &\geq \max(\Delta_i, \Delta_j). &(80)\end{aligned}$$

Further noting that

$$\left(\min_{k\in\mathcal{S}^\star(\nu)\backslash\{j,i\}}\left[\min(\mathrm{M}(k,j),\mathrm{M}(j,k))\right]\right) \geq \Delta_j$$

yields

$$\delta_j^{+,(i)} \geq \Delta_j. \tag{81}$$

Moreover, for any $k \in \mathcal{D}(\nu^{(i)}) = \mathcal{D}(\nu)\backslash\{\underline{i}\}$,

$$(\mathrm{M}(k,j))^+ + \Delta_k^{(i)} = \mathrm{M}(k,j)^+ + \Delta_k \geq \Delta_j.$$

Therefore, for any $j \in \mathcal{S}^\star(\nu^{(i)})\backslash\{i, \underline{i}\}$

$$\Delta_j^{(i)} := \min(\delta_j^{+,(i)}, \delta_j^{-,(i)}) \geq \Delta_j. \tag{82}$$

To summarize, we have proved that for any $j \in [K]\backslash\{i, \underline{i}\}$,

$$\Delta_j^{(i)} \geq \Delta_j.$$

We now analyze $\Delta_i^{(i)}$.

We have

$$\delta_i^{+,(i)} = \left(\min_{k\in\mathcal{S}^\star(\nu)\backslash\{i\}}\left[\min(\mathrm{M}^i(k,i),\mathrm{M}^i(i,k))\right]\right) \wedge \underbrace{(\min(\mathrm{M}(i,\underline{i}),\mathrm{M}(\underline{i},i)))}_{\beta_{\underline{i}}}. \tag{83}$$

We further note by direct calculation $\mathrm{M}^i(\underline{i}, i) = \Delta_i$ and $\mathrm{M}^i(i, \underline{i}) \geq \Delta_i$. Combining with (77) and (80) yields

$$\delta_i^{+,(i)} \geq \Delta_i.$$

Let $k \in \mathcal{D}(\nu)\backslash\{\underline{i}\}$. If $\boldsymbol{\theta}_k \prec \boldsymbol{\theta}_i$ then $\Delta_i \leq \Delta_k$ and

$$(\mathrm{M}^i(k,i))^+ + \Delta_k^{(i)} \geq \Delta_k^{(i)} \geq \Delta_k$$

so

$$(\mathrm{M}^i(k,i))^+ + \Delta_k^{(i)} \geq \Delta_i.$$

Similarly, if $\boldsymbol{\theta}_k \not\prec \boldsymbol{\theta}_i$ as in "Step 1" we prove

$$(\mathrm{M}^i(k,i))^+ + \Delta_k^{(i)} \geq \Delta_i.$$

Thus $\delta_i^{-,(i)} \geq \Delta_i$, hence

$$\Delta_i^{(i)} \geq \Delta_i. \tag{84}$$

The computation of $\Delta_{\underline{i}}^{(i)}$ proceeds identically to that of $\Delta_i^{(i)}$ to prove that

$$\Delta_{\underline{i}}^{(i)} \geq \Delta_{\underline{i}}.$$

Therefore, for any $k \in [K]$,

$$\Delta_k^{(i)} \geq \Delta_k,$$

which yields $H(\nu^{(i)}) \leq H(\nu)$.

□

D WHEN THE COMPLEXITY TERM H IS KNOWN

In this section we analyze APE-FB which has been used in the experiments reported in the main paper. This algorithm is similar to UCB-E (Audibert and Bubeck, 2010) and UGapEb (Gabillon et al., 2012). The idea is to use the sampling rule APE of Kone et al. (2023) (designed for fixed-confidence PSI) for T rounds and analyze the theoretical guarantees of the resulting algorithm. We recall that at each round t, $T_i(t)$ is the number of times arm i has been pulled up to time t and $\widehat{\boldsymbol{\theta}}_i(t)$ is the empirical estimate of $\boldsymbol{\theta}_i$ based on the $T_i(t)$ samples collected so far and $S(t)$ is the empirical Pareto set at time t. For any arms i, j,

$$\begin{aligned} \mathrm{M}(i,j;t) &:= \max_d \left[\widehat{\theta}_i^d(t) - \widehat{\theta}_j^d(t)\right] \text{ and} \\ \mathrm{m}(i,j;t) &:= \min_d \left[\widehat{\theta}_j^d(t) - \widehat{\theta}_i^d(t)\right] = -\mathrm{M}(i,j;t). \end{aligned}$$

Let $a \geq 0$ and define for any arm i,

$$\beta_i(t) := \frac{2}{5}\sqrt{\frac{a}{T_i(t)}}.$$

Let

$$\begin{aligned} Z_1(t) &:= \min_{i\in S(t)} \min_{j\in S(t)\setminus\{i\}} \left[\mathrm{M}(i,j;t) - \beta_i(t) - \beta_j(t)\right], \\ Z_2(t) &:= \min_{i\in S(t)^c} \max_{j\neq i} \left[\mathrm{m}(i,j;t) - \beta_i(t) - \beta_j(t)\right], \end{aligned}$$

and we use by convention $\min_\emptyset = \infty$. Letting

$$\mathrm{OPT}(t) := \{i \in [K] : \forall j \neq i, \mathrm{M}(i,j;t) - \beta_i(t) - \beta_j(t) > 0\},$$

we recall that APE(Kone et al., 2023) defines two arms b_t and c_t as follows:

$$b_t := \operatorname*{argmax}_{i\notin \mathrm{OPT}(t)} \min_{j\neq i} \left[\mathrm{M}(i,j;t) + \beta_i(t) + \beta_j(t)\right], \tag{85}$$

and when $\mathrm{OPT}(t) = [K]$ (which could often occur when $\mathcal{S}^\star = [K]$), define

$$b_t := \operatorname*{argmin}_{i\in[K]} \min_{j\neq i} \left[\mathrm{M}(i,j;t) - \beta_i(t) - \beta_j(t)\right].$$

Finally, we define

$$c_t := \operatorname*{argmin}_{j\neq b_t} \left[\mathrm{M}(b_t,j;t) - \beta_j(t)\right], \tag{86}$$

and a_t, the least explored among b_t and c_t. Then at each round we pull a_t.

Algorithm 3: APE-FB

Result: Pareto set
Data: parameter $a \geq 0$
initialize: pull each arm once and set $T_i(K) = 1$ for each arm i
for $t = K+1, \dots, T$ **do**
  Compute b_t (85) and c_t (86)
  Sample arm a_t the least explored among b_t and c_t
return : $\widehat{S}_T = S(T)$

We below state an upper-bound on the probability of error of APE-FB.

Theorem 6. *Let $T \geq K$ and ν a bandit with σ-subgaussian marginals . Let $0 \leq a \leq \frac{25}{36}\frac{T-K}{H(\nu)}$. The probability of error of APE-FB run with parameter a satisfies*

$$e_T^{APE}(\nu) \leq 2(1+\log(T))DK \exp\left(-\frac{a}{100\sigma^2}\right).$$

For $a = \frac{25}{36}\frac{T-K}{H(\nu)}$, we have

$$e_T^{\mathrm{APE}}(\nu) \leq 2(1+\log(T))DK\exp\left(-\frac{T-K}{144\sigma^2 H(\nu)}\right).$$

Compared to EGE-SH and EGE-SR, the exponent in the upper-bound is larger by an order $\log(K)$ but outside of the exponent there is an extra multiplicative $\log(T)$ term due to time uniform concentration. As this term grows very slowly with T it shouldn't be too high w.r.t $K|\mathcal{S}^\star|$ or $\log(K)|\mathcal{S}^\star|$ and the exponential term becomes quickly predominant when T is large. Thus APE-FB improves upon both EGE-SR and EGE-SH but requires a to be tuned optimally. This is an important limitation as in practice $H(\nu)$ is not known and the results reported in the main suggest that the performance of APE-FB heavily rely on this proper tuning. On the contrary, both EGE-SR and EGE-SH are parameter-free.

Proof. For any arm i and any component d, let $I_i^d := \left[\theta_i^d - \frac{1}{2}\beta_i(t); \theta_i^d + \frac{1}{2}\beta_i(t)\right]$. We define the event

$$\mathcal{E} := \left\{\forall i \in [K],\ \forall t \leq T,\ \forall d \in [D],\ \widehat{\theta}_i^d(t) \in I_i^d\right\}.$$

We assume the event $\mathcal{E}$ holds and we show that then $S(T) = \mathcal{S}^\star$. Let us introduce

$$\tau := \inf\left\{K \leq t \leq T : Z_1(t) > 0 \text{ and } Z_2(t) > 0\right\}.$$

Note that from the definition, τ can be infinite as the inf of an empty set but we will further show that indeed $\tau \leq T$. We claim that on the event $\mathcal{E}$, for any $\tau \leq t \leq T$, $S(t) = \mathcal{S}^\star$. We have at time $t = \tau$, $Z_1(t) \geq 0$ and $Z_2(t) \geq 0$. Letting $i \in S(t)$, for any $j \neq i$, there exists d_j such that

$$\widehat{\theta}_j^{d_j}(\tau) + \beta_j(\tau) < \widehat{\theta}_i^{d_j}(\tau) - \beta_i(\tau),$$

that is

$$\theta_j^{d_j} + \frac{1}{2}\beta_j(\tau) < \theta_i^{d_j} - \frac{1}{2}\beta_i(\tau),$$

which yields (as β is a decreasing function) for any $t \geq \tau$,

$$\theta_j^{d_j} + \frac{1}{2}\beta_j(t) < \theta_i^{d_j} - \frac{1}{2}\beta_i(t),$$

therefore, for any $t \geq \tau$, $\widehat{\theta}_j^{d_j}(t) < \widehat{\theta}_i^{d_j}(t)$. Said otherwise, on the event $\mathcal{E}$, if $i \in S(\tau)$ then for any $t \geq \tau$, $i \in S(t)$. Using an identical reasoning for any arm $i \in S(\tau)^\mathsf{c}$ also yields that on the event $\mathcal{E}$, if $i \in S(\tau)^\mathsf{c}$ then for any $t \geq \tau$, $i \in S(t)^\mathsf{c}$. At this point, we have proved that under the event $\mathcal{E}$, if $\tau \leq T$ then $S(T) = \mathcal{S}^\star$.

Showing that on the event $\mathcal{E}$, $\tau \leq T$ will conclude the proof. We proceed by contradiction and we assume $\tau > T$. First, note that $\tau > T$ imply that for any $t \leq T$, $\mathrm{OPT}(t) \neq [K]$ (otherwise we would have $Z_1(t) > 0$ and since $S(t)^\mathsf{c}$ would be empty, $Z_2(t) = \infty$). Therefore, the two candidate arms b_t and c_t are always defined as in Kone et al. (2023).

Introducing for any pair of arms i, j

$$U_{i,j}^d(t) = \theta_i^d - \theta_j^d + \frac{1}{2}\beta_i(t) + \frac{1}{2}\beta_j(t) \text{ and } L_{i,j}^d(t) = \theta_i^d - \theta_j^d - \frac{1}{2}\beta_i(t) - \frac{1}{2}\beta_j(t),$$

we deduce the following result from Lemma 1 of Kone et al. (2023).

Lemma 14. *On the event $\mathcal{E}$, at any round t and for any pair i, j we have*

$$|\mathrm{M}(i,j) - \mathrm{M}(i,j;t)| \leq \frac{1}{2}\left(\beta_i(t) + \beta_j(t)\right) \text{ and } |\mathrm{m}(i,j) - \mathrm{m}(i,j;t)| \leq \frac{1}{2}\left(\beta_i(t) + \beta_j(t)\right).$$

We state below a lemma adapted from Kone et al. (2023) which is important to upper-bound the sub-optimality gap of each arm by the confidence bonuses. The following lemma's proof is essentially identical to Lemma 4 of Kone et al. (2023).

Lemma 15. *Assume $\mathcal{E}$ holds. For any $t < \tau$, $\Delta_{a_t} < 3\beta_{a_t}(t)$.*

Since by assumption $\tau > T$, by Lemma 15 we have for any $t \leq T$,

$$\Delta_{a_t} < 3\beta_{a_t}(t). \tag{87}$$

For any arm i, let t_i be the last time i was pulled. We have $T_i(T) = 1 + T_i(t_i)$. Since i has been pulled at time t_i and $t_i < \tau$ we have $a_{t_i} = i$ and

$$\Delta_i < 3\beta_i(t_i),$$

that is

$$T_i(T) - 1 < \frac{36}{25}\frac{1}{\Delta_i^2 a},$$

therefore

$$T - K = \sum_{i=1}^{K}(T_i(T) - 1) < \frac{36}{25}\frac{a}{H(\nu)},$$

which is impossible for $a \leq \frac{25}{36}\frac{T-K}{H(\nu)}$. All put together, we have proved by contradiction that on the event $\mathcal{E}$, $\tau \leq T$ and by what precedes, the recommendation of APE-FB is correct : $S(T) = S(\tau) = \mathcal{S}^\star$. The upper-bound on $e_T^{\text{APE}}(\nu)$ then follows by upper-bounding $\mathbb{P}(\bar{\mathcal{E}})$ with a direct application of Hoeffding's maximal inequality (see e.g section A.2 of Locatelli et al. (2016)). □

Remark 2. *For $D = 1$, APE-FB slightly differs from UCB-E Audibert and Bubeck (2010) as it does not sample the arm with maximum upper confidence bound but the least sampled among the two arms with the largest upper-confidence bound. Moreover, in case $D = 1$, Theorem 6 recovers the result of Audibert and Bubeck (2010) (Theorem 1 therein) with a slightly different algorithm.*

E TECHNICAL LEMMAS

In this section we prove the lemma that allow to remove the explicit dependency on $\mathcal{S}^\star$ in the expression of the sub-optimality gaps. We use the following lemma which is taken from Kone et al. (2023).

Lemma 16 (Lemma 10 of Kone et al. (2023))**.** *For any sub-optimal arm a, there exists a Pareto optimal arm $a^\star$ such that $\boldsymbol{\theta}_a \prec \boldsymbol{\theta}_{a^\star}$ and $\Delta_a = \mathrm{m}(a, a^\star) > 0$. Moreover, For any $i \in [K] \setminus \mathcal{S}^\star$, $j \in \mathcal{S}^\star$*

i) $\max_{k \in \mathcal{S}^\star} \mathrm{m}(i,k) = \max_{k\in[K]} \mathrm{m}(i,k)$,

ii) If $i \in \operatorname{argmin}_{k\in[K]\setminus\{j\}} \mathrm{M}(j,k)$ then j is the unique arm such that $\boldsymbol{\theta}_i \prec \boldsymbol{\theta}_j$.

We now prove Lemma 1.

Lemma 1. *For any arm $i \in [K]$,*

$$\Delta_i = \begin{cases} \Delta_i^\star = \max_{j\in[K]} \mathrm{m}(i,j) & \text{if } i \notin \mathcal{S}^\star \\ \delta_i^\star & \text{if } i \in \mathcal{S}^\star \end{cases},$$

where $\delta_i^\star := \min_{j\neq i}[\mathrm{M}(i,j) \wedge (\mathrm{M}(j,i)^+ + (\Delta_j^\star)^+)]$.

Proof. For sub-optimal arms, the result follows from Lemma 16. It remains to prove the equality for sub-optimal arms. By definition, for an optimal arm i, we have

$$\Delta_i = \min(\delta_i^+, \delta_i^-),$$

where

$$\delta_i^+ := \min_{j\in\mathcal{S}^\star\setminus\{i\}} \min(\mathrm{M}(i,j), \mathrm{M}(j,i)) \text{ and } \quad \delta_i^- := \min_{j\in[K]\setminus\mathcal{S}^\star} \left[(\mathrm{M}(j,i))^+ + \Delta_j^\star\right].$$

For any optimal arm i, $\Delta_i^\star \leq 0$ (by direct calculation), so introducing

$$\delta_i^{-'} := \min(\delta_i^-, \min_{j\in\mathcal{S}^\star\setminus\{i\}} \mathrm{M}(j,i)),$$

we have

$$\delta_i^{-'} = \min_{j\neq i}[\mathrm{M}(j,i)^+ + (\Delta_j^\star)^+]$$

Then, if

$$\min_{j\in\mathcal{S}^\star\setminus\{i\}} \mathrm{M}(i,j) = \min_{j\neq i} \mathrm{M}(i,j), \tag{88}$$

holds, the result simply follows as for any optimal arm i,

$$\begin{aligned} \Delta_i = \min(\delta_i^+, \delta_i^-) &= \min(\min_{j\in\mathcal{S}^\star\setminus\{i\}} \mathrm{M}(i,j), \delta_i^{-'}), \\ &= \min(\min_{j\neq i} \mathrm{M}(i,j), \delta_i^{-'}), \\ &= \min_{j\neq i}[\mathrm{M}(i,j) \wedge (\mathrm{M}(j,i)^+ + (\Delta_j^\star)^+)], \\ &= \delta_i^\star. \end{aligned}$$

In the sequel, assume (88) does not hold, that is assume

$$\min_{j\in\mathcal{S}^\star\setminus\{i\}} \mathrm{M}(i,j) > \min_{j\neq i} \mathrm{M}(i,j).$$

From Lemma 16 we know that in this case, there exists a sub-optimal arm k such that i is the unique arm dominating k and

$$\Delta_k^\star = \mathrm{m}(k,i) \quad \text{and} \quad \min_{j\neq i} \mathrm{M}(i,j) = \mathrm{M}(i,k).$$

Thus, we have

$$
\begin{aligned}
\min_{j \neq i} \mathrm{M}(i,j) &= \mathrm{M}(i,k), && (89)\\
&\geq \mathrm{m}(k,i) = \Delta_k^\star, && (90)\\
&\geq \delta_i^{-'} \quad (\text{since } i \text{ dominates } k, \ \mathrm{M}(k,i)^+ = 0). && (91)
\end{aligned}
$$

On the other side, as $\boldsymbol{\theta}_k \prec \boldsymbol{\theta}_i$, using the definition of Δ_i in particular δ_i^- directly yields

$$
\begin{aligned}
\Delta_i &\leq \Delta_k^\star = \mathrm{m}(k,i) && (92)\\
&\leq \mathrm{M}(i,k) = \min_{j \neq i} \mathrm{M}(i,j) && (93)\\
&< \min_{j \in \mathcal{S}^\star \setminus \{i\}} \mathrm{M}(i,j). && (94)
\end{aligned}
$$

We recall that

$$
\Delta_i = \min(\delta_i^{-'}, \min_{j \in \mathcal{S}^\star \setminus \{i\}} \mathrm{M}(i,j)),
$$

which combined with (94) yields

$$
\Delta_i = \delta_i^{-'}
$$

and further combining with (91) yields

$$
\begin{aligned}
\Delta_i = \delta_i^{-'} &= \min(\min_{j \neq i} \mathrm{M}(i,j), \delta_i^{-'})\\
&= \delta_i^\star,
\end{aligned}
$$

which concludes the proof. □

F GEOMETRIC R-ROUND ALLOCATION

In this section, we mention an alternative allocation scheme, which could be coupled with EGE and we derive an upper-bound on the mis-identification probability of the resulting algorithm. Karpov and Zhang (2022) proposed an R-round allocation for fixed-budget BAI. Their allocation scheme follows a geometric grid. In our notation, their allocation is as follows: given $R \in \mathbb{N}^\star$,

$$\alpha_0 = 0 \text{ and } \forall r \in \{1, \dots, R\}, \alpha_r = \left\lfloor \frac{T}{R} \cdot \frac{K^{r/R}}{K^{1+1/R}} \right\rfloor$$

and at round r, $t_r := \alpha_r - \alpha_{r-1}$. The arm elimination schedule is

$$\forall r \in [R+1], \lambda_r = \left\lfloor \frac{K}{K^{(r-1)/R}} \right\rfloor .$$

Direct calculation shows that as defined, $\boldsymbol{\lambda}$ and $\boldsymbol{t}$ satisfy the conditions (3) and (4). We denote by EGE-GG the specialization of EGE for the $\boldsymbol{\lambda}, \boldsymbol{t}$ aforementioned. We deduce the following result from Theorem 1.

Proposition 1. *Let $R \in \mathbb{N}^\star$. For any σ-subgaussian multi-variate bandit, the mis-identification probability of EGE-GG satisfies*

$$e_T^{GG}(\nu) \leq 2(K-1)|\mathcal{S}^\star| RD \exp\left(-\frac{T}{288\sigma^2 R K^{1/R} H_2(\nu)}\right).$$

Proof. We assume $T \geq 2RK$. We have

$$\begin{aligned}
\widetilde{T}^{R,\boldsymbol{t}^{\mathrm{GG}},\boldsymbol{\lambda}^{\mathrm{GG}}}(\nu) &:= \min_{r\in[R]} \left(\sum_{s=1}^{r} t_s\right) \Delta^2_{(\lambda_{r+1}+1)}, \\
&= \min_{r\in[R]} [\alpha_r \Delta^2_{(\lambda_{r+1}+1)}], \\
&= \min_{r\in[R]} \left[\alpha_r (\lambda_{r+1}+1) \frac{\Delta^2_{(\lambda_{r+1}+1)}}{(\lambda_{r+1}+1)}\right], \\
&\geq \frac{T}{2RK^{1/R}} \frac{1}{\max_{r\in[R]}[(\lambda_{r+1}+1)\Delta^{-2}_{(\lambda_{r+1}+1)}]} \\
&\geq \frac{T}{2RK^{1/R}} \frac{1}{H_2(\nu)}.
\end{aligned}$$

The results follows by Theorem 1. □

Similar derivation can be done for any allocation scheme that satisfy (3) and (4) resulting in different instanciations of EGE.

G IMPLEMENTATION DETAILS AND ADDITIONAL EXPERIMENTS

In this section, we give additional details about the experiments and we report additional experimental results.

G.1 Implementation details and computational complexity

We discuss implementation details and give additional information on the datasets for reproducibility.

G.1.1 Implementation details

Setup We have implemented the algorithms mainly in `C++17` interfaced with `python 3.10` through the `cython3` package. The experiments are run on an ARM64 8GB RAM/8 core/256GB disk storage computer.

Datasets : The COV-BOOST datasets can be found in Appendix I of Kone et al. (2023). A link to download the SNW dataset is given in Zuluaga et al. (2013). We plot below 2 of the 4 synthetic instances used in the main paper. The means of the instance in the fourth experiment in dimension 10 are given as a `numpy` data file in the supplementary material.

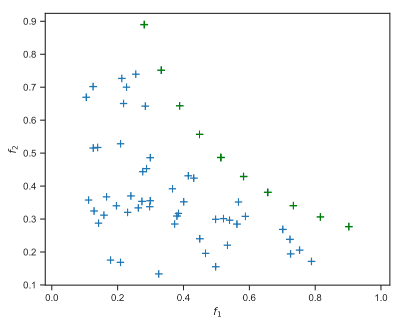

Figure 13: Synthetic Experiment 1: Group of arms on a convex Pareto set.

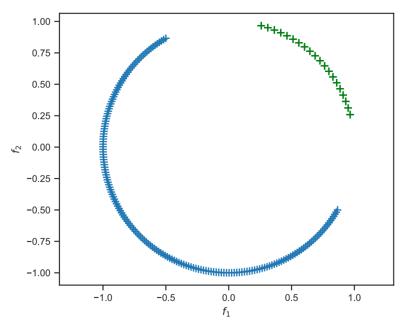

Figure 14: Synthetic Experiment 3: $K = 200$ arms on the unit circle.

	K	D	$\lvert\mathcal{S}^\star\rvert$
COV-BOOST	20	3	2
SNW	206	2	5
Exp. 1	60	2	10
Exp. 2	10	2	2
Exp. 3	200	2	20
Exp. 4	50	10	18

Table 2: Summary of the bandit instances used in the main paper.

We summarize in Table 2 the instances used in the main paper.

G.1.2 Computational complexity

Time and memory complexity The time and memory complexity of our implementations are dominated by the computation and the storage of $\mathrm{M}(i,j;r)$ and $\mathrm{m}(i,j;r)$ at each round r for all the pairs of active arms. The time complexity of and R-rounds implementation of EGE is $\mathcal{O}(RK^2D)$ and the memory complexity is $\mathcal{O}(K^2)$.

Runtime and scaling We report below the average runtime of our algorithms for the experiments reported in the main paper. We report on Fig.16 and Fig.16 the runtime versus the number of arms for EGE-SR/SH with varying K and $T = 1000$. We average the runtime over 2000 runs. We observe that both algorithms can be used for applications on large scale datasets with a reasonable runtime.

G.2 Additional experiments

We report results of additional experiments in particular for the k-relaxation. We try different values of k and we compute the loss or mis-identification error averaged over the trials. We plot the average loss versus the budget for each value of k. For additional experiments on PSI, the protocol is as described in the main paper.

G.2.1 Experiments on PSI-k

We report the experimental result of EGE-SR-k on two bandit instances with a large number of arms. For each bandit instance we test different values of k and we report $\log_{10}$ of the loss averaged over 4000 independent trials.

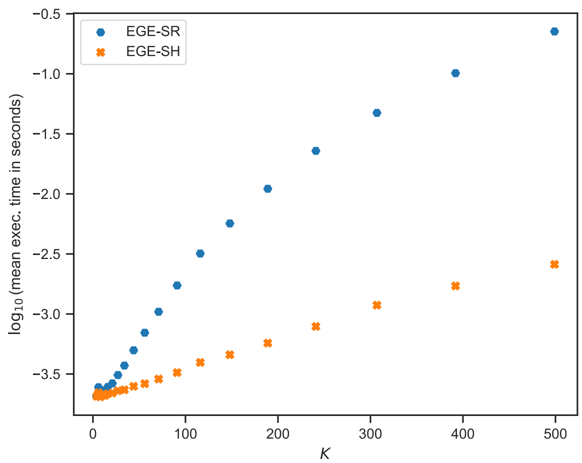

Figure 15: $\log_{10}$ of the mean execution time in seconds averaged over 2000 trials for an instance with $D = 2$.

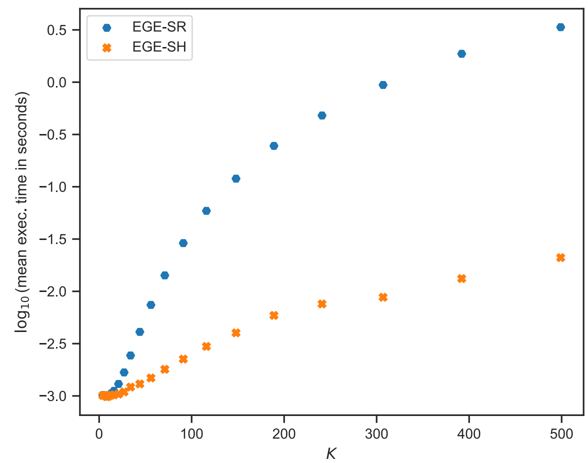

Figure 16: $\log_{10}$ of the mean execution time in seconds averaged over 2000 trials for an instance with $D = 20$.

Recall that given k and T the PSI-k loss is

$$\mathcal{L}(\widehat{S}_T, k) := \begin{cases} \mathbb{I}\{\widehat{S}_T \subset \mathcal{S}^\star\} & \text{if} \quad |\widehat{S}_T| = k, \\ \mathbb{I}\{\widehat{S}_T = \mathcal{S}^\star\} & \text{else.} \end{cases}$$

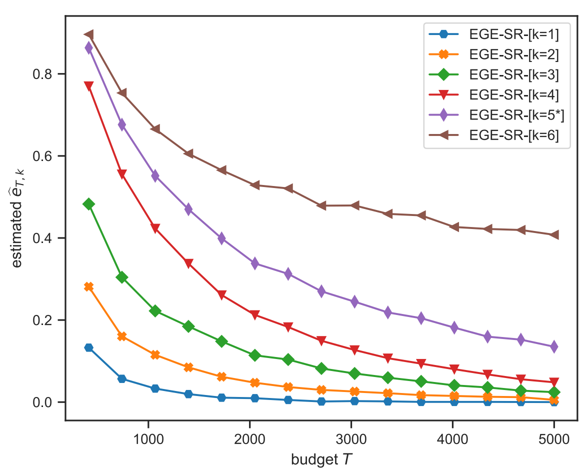

Figure 17: Estimated PSI-k loss for different values of k on the SNW dataset.

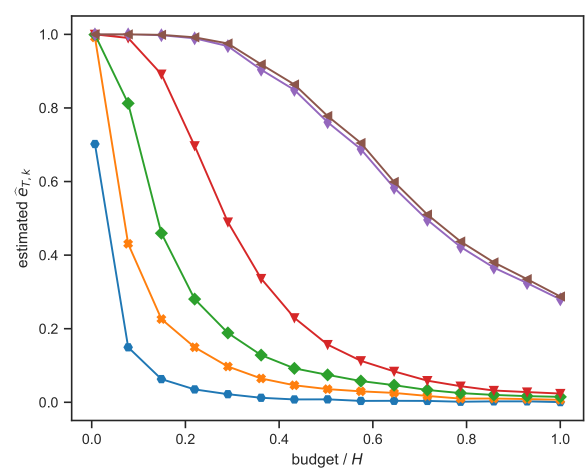

Figure 18: Estimated PSI-k loss for different values of k on Exp.1 (convex Pareto set)

We observe (Fig. 17 and Fig.18) that the loss (the probability of error of EGE-SR-k) decreases exponentially fast with the budget. The asterisk indicates the true size of the Pareto set of the instance. We remark that the smaller k the smaller the loss, which is expected as the quantity $H_2(\nu)^{(k)}$ (Theorem 3) increases with k. We also note that for $k = |\mathcal{S}^\star|$, the loss can be smaller than when $k > |\mathcal{S}^\star|$. An equivalent remark (in terms of sample complexity) has been made by Kone et al. (2023) for the PSI-k problem in fixed-confidence. Indeed for $k = |\mathcal{S}^\star|$ the PSI-k problem is not exactly PSI. As in PSI-k we stop as soon as k optimal arms have been found, in PSI, on the same run, we would continue until all the arms have been classified; thus we could make mistakes on theses additional steps. Finally, note that the expected loss of PSI-k is the same as PSI for any $k > |\mathcal{S}^\star|$.

Hyper-volume metric : The hyper-volume metric (see e.g Daulton et al. (2020)) measures the region dominated by a set of points. Given a reference point $\boldsymbol{r}$ and a set $S \subset \mathbb{R}^D$ the hyper-volume of S is

$$\mathrm{HV}(S) := \lambda\left(\bigcup_{\boldsymbol{x}\in S}[\boldsymbol{r}, \boldsymbol{x}]\right),$$

where λ is the Lebesgue measure on $\mathbb{R}^D$ and $[\boldsymbol{r}, \boldsymbol{x}]$ is the hyper-rectangle $[r^1, x^1] \times \cdots \times [r^D, x^D]$. If $S' \subset S$ then $\mathrm{HV}(S) \geq \mathrm{HV}(S')$. More importantly, the hyper-volume of S is equal to the hyper-volume of its Pareto set. We use this indicator to measure the quality of the set returned by EGE-SR-k. We report the average value of

$$\alpha_{\mathrm{HV}} := \frac{\mathrm{HV}(\widehat{S}_T)}{\mathrm{HV}(\mathcal{S}^\star)}$$

over the trials for different values of k and T fixed. The larger this hyper-volume fraction the more a set $\widehat{S}_T \subset \mathcal{S}^\star$ covers $\mathcal{S}^\star$ in term of dominated region and thus is representative of $\mathcal{S}^\star$. Note that $S = [K]$ also has maximal hyper-volume indicator, therefore we need to couple α_{HV} with $e_{T,k}(\nu)$ (reported in Fig.2 in the main and Fig.19) to properly interpret the quality of the returned set.

We observe on Fig.19 that the estimated value of $\mathbb{E}[\alpha_{\mathrm{HV}}]$ is nearly 0.85 for $k = 1$ and it naturally increases with k. In Fig.20, for $k = 1$, the arm returned by EGE-SR-k covers nearly half of the region dominated by the Pareto set. These observations suggest that the arms with the highest hyper-volume contribution are easier to identify as optimal hence they will be the first added to B_r.

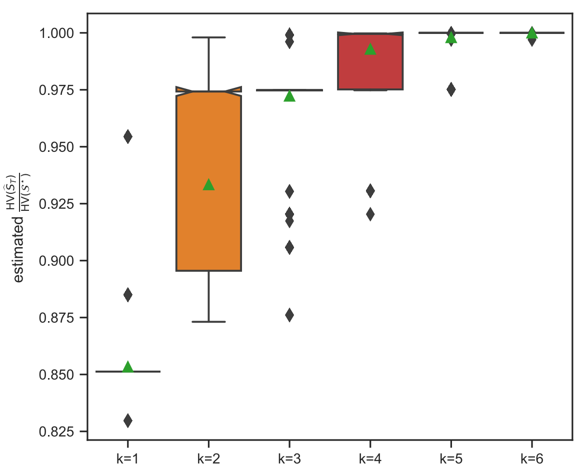

Figure 19: Hyper-volume fraction of the returned set for $T = 5000$ on the SNW dataset.

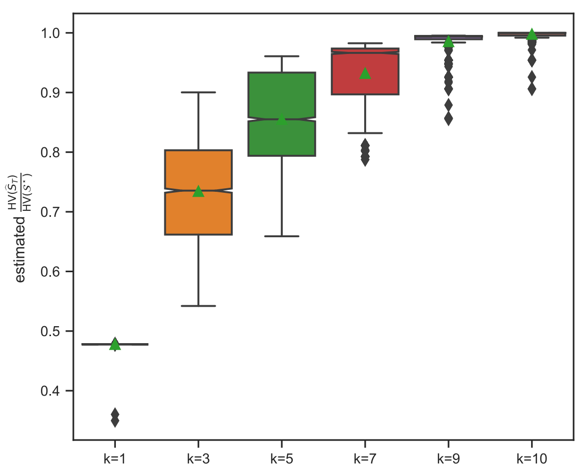

Figure 20: Hyper-volume fraction of the returned set for $T = H$ on Exp.1 (convex Pareto set)

G.2.2 Additional experiments on PSI

We report additional experimental results on 4 synthetic instances which complements our results reported in the main paper. The new instances are described below.

Experiment 5: 2 clusters of arms. We set $K = 20$ and we choose $(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_{10}) \sim \mathcal{U}\left([0.2, 0.4]^2\right)^{\otimes 10}$ and $(\boldsymbol{\theta}_{11}, \dots, \boldsymbol{\theta}_{20}) \sim \mathcal{U}\left([0.5, 0.7]^2\right)^{\otimes 10}$. There are 4 optimal arms for this instance.

Experiment 6: All the arms are optimal $K = 10, D = 2$ and for any arm i, $\theta_i^1 = 0.75 - 0.65^i, \theta_i^2 = 0.25 + 0.65^i$

Experiment 7: All the arms have the same sub-optimality gap. $K = 22, D = 2$ and we choose $\boldsymbol{\Theta}$ to have $\Delta_1 = \cdots = \Delta_{22}$. Unlike single-objective bandit, such instances can be generated with all arms being different. We choose for $i = 1, \dots 8$, $\boldsymbol{\theta}_i := (0.3 + c_i, 0.8 - c_i)^\intercal$ and $c_i = (i-1) * c$. For $i = 9, \dots, 15$, $\boldsymbol{\theta}_i := (0.25 + c_{i-8}, 0.7 - c_{i-8})^\intercal$ and for $i = 16, \dots, 22$, $\boldsymbol{\theta}_i := \boldsymbol{\theta}_{i-7} - (0, -0.05)^\intercal$.

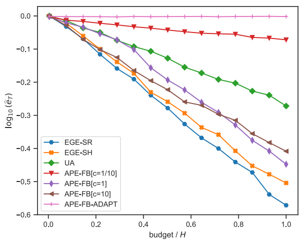

Figure 21: Synthetic Experiment 5: 2 clusters of arms

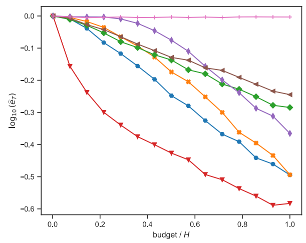

Figure 22: Synthetic Experiment 6: All the arms are optimal.

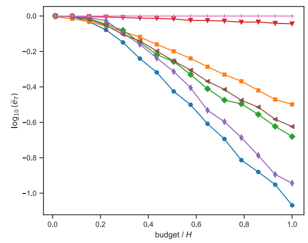

Figure 23: Synthetic Experiment 7: The same gap: $\Delta_i = \Delta$ for each arm

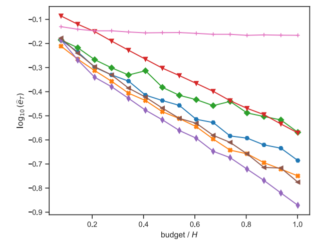

Figure 24: Synthetic Experiment 8: Geometric progression

Experiment 8: Geometric progression with a single optimal arm. We set $K = 5, D = 2$ and for any $i \in \{1, \dots, 5\}$, $\theta_i^1 = \theta_i^2 = 0.75 - 0.25^i$. For this instance there is a unique optimal arm.

The experiments (Fig.21 to Fig.24) confirm the superior performance of EGE on general instance types. We note however that on experiment 7, where all the arms have the same sub-optimality gaps EGE-SH is slightly outperformed by Uniform allocation but both are largely outperformed by EGE-SR and APE-FB. This is not surprising as on such instances, the exponential decay rate of the mis-identification error of uniform allocation is smaller than that of EGE-SH and the "aggressive" geometrical allocation of SH makes it allocate way less samples to arms that are discarded in early stages. Indeed when $\Delta_1 = \dots = \Delta_K$ we have $K\Delta_{(1)}^{-2} = H_1(\nu) = H_2(\nu)$ thus the result of Tab.1 becomes

$$e_T^{\text{UA}}(\nu) \leq 2(K-1)|\mathcal{S}^\star| D \exp\left(-\frac{T}{144\sigma^2 H(\nu)}\right),$$

which is even tighter than APE-FB when $\log(T)$ is larger than $|\mathcal{S}^\star|$. For BAI (PSI with $D = 1$), it is known (Audibert and Bubeck, 2010) that on instances where the gaps are all the same, the uniform allocation is optimal. For example, by taking $\theta_1 = 0.7$ and $\theta_i = 0.4$ for $i = 2, \dots, 8$ and a Bernoulli bandit, the gaps are all the same on this instance and one can observe empirically that the Uniform allocation slightly outperforms Sequential Halving on this instance.