Bandit Pareto Set Identification in a Multi-Output Linear Model

Anonymous Author(s)
Affiliation
Address
email

Abstract

We study the Pareto Set Identification (PSI) problem in a structured multi-output linear bandit model. In this setting each arm is associated a feature vector belonging to $\mathbb{R}^h$ and its mean vector in $\mathbb{R}^d$ linearly depends on this feature vector througxh a common unknown matrix $\Theta \in \mathbb{R}^{h\times d}$. The goal is to identity the set of non-dominated arms by adaptively collecting samples from the arms. We introduce and analyze the first optimal design-based algorithms for PSI, providing nearly optimal guarantees in both the fixed-budget and the fixed-confidence settings. Notably, we show that the difficulty of these tasks mainly depends on the sub-optimality gaps of h arms only. Our theoretical results are supported by an extensive benchmark on synthetic and real-world datasets.

1 Introduction

A multi-armed bandit is a stochastic game where an agent faces K distributions (or arms) whose means are unknown to her. When the distributions are scalar-valued, the agent faces two main tasks: regret minimization and pure exploration. In the former, the agent aims at maximizing the sum of observations collected along its trajectory [Lattimore and Szepesvári, 2020]. In pure exploration the agent has to solve a stochastic optimization problem after some steps of exploration and it does not suffer any loss during exploration [Bubeck and Munos, 2008]. Examples of pure exploration tasks include best arm identification in which the goal is to find the arm with largest mean [Audibert and Bubeck, 2010], thresholding bandit [Locatelli et al., 2016] or combinatorial bandits [Chen et al., 2014], to name a few.

In this paper, we are interested in the less common setting where the rewards are $\mathbb{R}^d$-valued, with $d > 1$. Different pure exploration tasks have been considered in this context, e.g. finding the set of feasible arms, i.e. arms whose mean satisfy some constraints [Katz-Samuels and Scott, 2018], or a feasible arm maximizing a linear combination of the different criteria [Katz-Samuels and Scott, 2019, Faizal and Nair, 2022]. Finding appropriate constraints is not always possible in practical problems and our focus is on the identification of the Pareto set, that is the set of arms whose means are not uniformly dominated by that of any other arm, a setting first studied by [Auer et al., 2016]. We note that a regret minimization counterpart of this problem has been considered by [Drugan and Nowe, 2013].

Pareto set identification can be relevant in many real-world problems where there are multiple, possibly conflicting objectives to optimize simultaneously. Examples include monitoring the energy consumption and runtime of different algorithms (see our use case in Section 5), or identifying a set of interesting vaccine by observing different immunogenicity criteria (antibodies, cellular response, that are not always correlated, as exemplified by Kone et al. [2023]). In both cases, there could be many arms with a few descriptor of the different arms (e.g. vaccine technology, doses, injection

Submitted to 17th European Workshop on Reinforcement Learning (EWRL 2024). Do not distribute.

times). By incorporating such arm features in the model we expect to reduce substantially the number of samples needed to identify the Pareto set.

In this work, we incorporate some structure in the PSI identification problem through a multi-output linear model, formally described in Section 2. In this model, each of the K arms whose means are in $\mathbb{R}^d$ is described by a feature vector in $\mathbb{R}^h$, $h > 1$. We propose the GEGE algorithm, which combines a G-optimal design exploration mechanism with an accept/reject mechanism based on the estimation of some notion of sub-optimality gap. GEGE can be instantiated in both the fixed-budget setting (given at most T samples, output a guess of the Pareto set minimizing the error probability) and the fixed-confidence setting (minimize the number of sample used so as to guarantee an error probability smaller than some prescribed δ). Through a unified analysis, we show that in both cases the sample complexity of GEGE, that is the number of samples needed to guarantee a certain probability of error, scales only with the h smallest sub-optimality gaps. This yields a reduction in sample complexity due to the structural assumption. Finally, we empirically evaluate our algorithms with extensive synthetic and real-world data-sets, and compare their performance with other state-of-the-art algorithms.

Related work When $d = 1$ and the feature vectors are the canonical basis of $\mathbb{R}^K$, PSI coincides with the best arm identification problem, that has been extensively studied in the literature both in the fixed-budget [Audibert and Bubeck, 2010, Karnin et al., 2013, Carpentier and Locatelli, 2016] and the fixed-confidence settings Kalyanakrishnan et al. [2012], Jamieson et al. [2014]. For sub-Gaussian distributions, the sample complexity is known to be essentially characterized (up to a $\log(K)$ factor in the fixed-budget setting) by a sum over the K arms of the inverse squared value of their *sub-optimality gap*, which is their distance to the (unique) optimal arm. In the fixed-confidence setting and for Gaussian distributions there are even algorithms matching the minimal sample complexity when δ goes to zero, which takes a more complex, non-explicit form (e.g., Garivier and Kaufmann [2016], You et al. [2023]).

Still when $d = 1$ but for general features in $\mathbb{R}^h$, our model coincides with the well-studied linear bandit model (with finitely many arms), in which the best arm identification task has also received some attention. It was first studied by Soare et al. [2014] in the fixed-confidence setting who established the link with optimal designs of experiments [Pukelsheim, 2006] showing that the minimal sample complexity can be expressed as an optimal (XY) design. The authors proposed the first elimination algorithms where in each round the surviving arms are pulled according to some optimal designs and obtained a sample complexity scaling in $(h/\Delta_{\min}^2)\log(1/\delta)$ where $\Delta_{\min}$ is the smallest gap in the model. Tao et al. [2018] further proposed an elimination algorithm using a novel estimator of the regression parameter based on a G-optimal design, with an improved sample complexity in $\sum_{i=1}^{h} \Delta_{(i)}^{-2} \log(1/\delta)$ where $\Delta_{(1)} \leq \cdots \leq \Delta_{(h)}$ are the h smallest gaps. This bound improves upon the complexity of the un-structured setting when $K \gg h$. Some algorithms even match the minimal sample complexity either in the asymptotic regime $\delta \to 0$ [Degenne et al., 2020, Jedra and Proutiere, 2020] or within multiplicative factors Fiez et al. [2019]. Some adaptive algorithms such as LinGapE Xu et al. [2018] are also very effective in practice, but without provably improving over un-structured algorithms in all instances.

The fixed-budget setting has been studied by Azizi et al. [2022], Yang and Tan [2022] who propose algorithms based on Sequential Halving Karnin et al. [2013] where in each round the active arms are sampled according to a G-optimal design. The best guarantees are those obtained by Yang and Tan [2022] who show that a budget T of order $\log_2(h) \sum_{i=1}^{h} \Delta_{(i)}^{-2} \log(1/\delta)$ is sufficient to get an error smaller than δ. Katz-Samuels et al. [2020] propose an elimination algorithm that can be instantiated both in the fixed confidence and fixed budget settings, and is close in spirit to our algorithm. However, unlike prior work, their optimal design aims at minimizing a new complexity measure called the Gaussian width that may better characterize the non asymptotic regime of the error. Extending this notion, or that of minimal (asymptotic) sample complexity to linear PSI is challenging due to the complex structure of the set of alternative models with a different Pareto set. In this work, our focus is on obtaining refined gap-based guarantees for the structured PSI problem.

When $d > 1$, the PSI identification problem has been mostly studied in the unstructured setting ($h = K$, canonical basis features). Auer et al. [2016] introduced some appropriate (non-trivial) notion of sub-optimality gaps for the PSI problem, which we recall in the next section. They proposed an elimination-based fixed-confidence algorithm whose sample complexity scales in $\sum_{i=1}^{K} \Delta_i^{-2} \log(1/\delta)$, which is proved to be near-optimal. A fully sequential algorithm with some slightly smaller bound was later given by Kone et al. [2023], who can further address different relaxations of the PSI problem. Kone et al. [2024] proposed the first fixed-budget PSI algorithm: a generic round-based elimination algorithm that estimates the sub-optimality gaps of Auer et al. [2016] and discard and classify some arms at the end of each round, with a sample complexity in $\sum_{i=1}^K \Delta_i^{-2} \log(K) \log(1/\delta)$.

The multi-output linear setting that we consider in this paper was first studied by Lu et al. [2019] from the Pareto regret minimization perspective. This model may also be viewed as a special case of the multi-ouput kernel regression model considered by Zuluaga et al. [2016] when a linear kernel is chosen. This work provide guarantees for approximate identification of the Pareto set, scaling with the information gain. Choosing appropriately the approximation parameter in ε-PAL as a function of the smallest gap $\Delta_{\min}$ yields a fixed-confidence PSI algorithm with sample complexity of order $(h^2/\Delta_{\min}^2)\log(1/\delta)$. More recently, the preliminary work of Kim et al. [2023] proposed an extension of the fixed-confidence algorithm of Auer et al. [2016] with a robust estimator to simultaneously minimize the Pareto regret and identify the Pareto set. Their claimed sample complexity bound is in $(h/\Delta_{\min}^2)\log(1/\delta)$.

For the fixed-confidence variant of GEGE we prove an improved sample complexity bounds in which $(h/\Delta_{\min}^2)$ is replaced by the sum $\sum_{i=1}^h \Delta_{(i)}^{-2}$. Moreover, to the best of our knowledge the fixed-budget variant of GEGE is the first algorithm for fixed-budget PSI in a multi-output linear bandit model, and enjoys a similar sample complexity. Our experiments confirm these good theoretical properties, and illustrate the impact of the structural assumption.

2 Setting

We formalize the linear PSI problem. Let $d, h \in \mathbb{N}^\star$ and $K \geq 2$. $\nu_1, \dots, \nu_K$ are distributions over $\mathbb{R}^d$ with means (resp.) $\mu_1, \dots, \mu_K \in \mathbb{R}^d$. We assume there are known feature vectors $x_1, \dots, x_K \in \mathbb{R}^h$ associated to each arm and an unknown matrix $\Theta \in \mathbb{R}^{h\times d}$ such that for any arm k, $\mu_k = \Theta^\intercal x_k$. Let $\mathcal{X} := (x_1 \dots x_K)^\intercal$ and $[K] = \{1, \dots, K\}$. The Pareto set is defined as $\mathcal{S}^\star = \{i \in [K] : \nexists j \in [K]\backslash\{i\} : \mu_i \preceq \mu_j\}$ in the sense of the following (Pareto) dominance relationship.

Definition 1. *For any two arms $i, j \in [K]$, i is weakly dominated by j if for any $c \in \{1, \dots, d\}$, $\mu_i(c) \leq \mu_j(c)$. An arm i is dominated by j ($\mu_i \preceq \mu_j$ or simply $i \preceq j$) if i is weakly dominated by j and there exists $c \in \{1, \dots, d\}$ such that $\mu_i(c) < \mu_j(c)$. An arm i is strictly dominated by j ($\mu_i \prec \mu_j$ or simply $i \prec j$) if for any $c \in \{1, \dots, d\}$, $\mu_i(c) < \mu_j(c)$.*

In each round t, an agent chooses an action a_t from $[K]$ and observes a response $y_t = \Theta^\intercal x_{a_t} + \eta_t$ where $(\eta_s)_{s\leq t}$ are *i.i.d* centered vectors in $\mathbb{R}^d$ whose marginal distributions are σ-subgaussian.[1] In this stochastic game, the goal of the agent is to identify the Pareto set $\mathcal{S}^\star$. In the fixed-confidence setting, given $\delta \in (0, 1)$, the agent collects samples up to a (random) stopping time τ and outputs a guess $\widehat{S}_\tau$ that should satisfy $\mathbb{P}(\mathcal{S}^\star \neq \widehat{S}_\tau) \leq \delta$ while minimizing τ (either with high-probability or in expectation). In the fixed-budget setting, the agent should output a set $\widehat{S}_T$ after T (fixed) rounds and minimize $e_T := \mathbb{P}(\widehat{S}_T \neq \mathcal{S}^\star)$.

We following notation is used throughout the paper. $\boldsymbol{\Delta}_n$ is the probability simplex of $\mathbb{R}^n$ and if $A \in \mathbb{R}^{n\times n}$ is positive semidefinite, for $x \in \mathbb{R}^n$, $\|x\|_A^2 = x^\intercal A x$ and $x(i)$ denotes its i-th component.

2.1 Complexity Measures for Pareto Set Identification

Choosing the features vectors to be the canonical basis of $\mathbb{R}^K$ and $\Theta = (\mu_1, \dots, \mu_K)$, we recover the unstructured multi-dimensional bandit model, in which the complexity of Pareto set identification is known to depend on some notion of sub-optimality gaps, first introduced by Auer et al. [2016]. These gaps can be expressed with the quantities

$$\mathrm{m}(i,j) := \min_{c\in[d]} [\mu_j(c) - \mu_i(c)] \text{ and } \mathrm{M}(i,j) := -\,\mathrm{m}(i,j).$$

We can observe that $\mathrm{m}(i, j) > 0$ iff $i \prec j$ and represents the amount by which j dominates i when positive. Similarly $\mathrm{M}(i, j) > 0$ iff $i \nprec j$ and when positive represents the quantity that should be added component-wise to j for it to dominate i. The sub-optimality gap Δ_i measures the difficulty to

[1] A centered random variable X is σ- subgaussian if for any $\lambda \in \mathbb{R}, \log \mathbb{E}[\exp(\lambda X)] \leq \lambda^2\sigma^2/2$.

classify arm i as optimal or sub-optimal and can be written (Lemma 1 of Kone et al. [2024])

$$\Delta_i := \begin{cases} \Delta_i^\star := \max_{j\in[K]} \mathrm{m}(i,j) & \text{if } i \notin \mathcal{S}^\star \\ \delta_i^\star & \text{else,} \end{cases} \tag{1}$$

where $\delta_i^\star := \min_{j\neq i}[\mathrm{M}(i,j) \wedge (\mathrm{M}(j,i)_+ + (\Delta_j^\star)_+)]$. For a sub-optimal arm i, Δ_i is the smallest quantity by which μ_i should be increased to make i non dominated. For an optimal arm i, Δ_i is the minimum between some notion of distance to the other optimal arms, $\min_{j\in\mathcal{S}^\star\setminus\{i\}}[\mathrm{M}(i,j)\wedge\mathrm{M}(j,i)]$ and the smallest margin to the sub-optimal arms $\min_{j\notin\mathcal{S}^\star}[\mathrm{M}(j,i)_+ + (\Delta_j^\star)_+]$. These quantities are illustrated Appendix G. We assume without loss of generality that $\Delta_1 \leq \dots \leq \Delta_K$ and we recall the quantities $H_1 = \sum_{i=1}^K \Delta_i^{-2}$ and $H_2 := \max_{i\in[K]} i\Delta_i^{-2}$ which have been used to measure the difficulty of Pareto set identification respectively in fixed-confidence [Auer et al., 2016] and fixed-budget [Kone et al., 2024] settings. In this work we introduce two analogue quantities for linear PSI namely

$$H_{1,\text{lin}} = \sum_{i=1}^h \frac{1}{\Delta_i^2} \quad \text{and} \quad H_{2,\text{lin}} := \max_{i\in[h]} \frac{i}{\Delta_i^2} \tag{2}$$

and we will show that the hardness of linear PSI can be characterized by $H_{1,\text{lin}}$ and $H_{2,\text{lin}}$ respectively in the fixed-confidence and fixed-budget regimes. These complexity measures are smaller than H_1 and H_2 respectively as they only feature the h smallest gaps. In order to obtain this reduction in complexity, it is crucial to estimate the underlying parameter $\Theta \in \mathbb{R}^{h\times d}$ instead of the K mean vectors.

2.2 Least Square Estimation and Optimal Designs

Given n arm choices in the model, $a_1,\dots,a_n$, we define $X_n := (x_{a_1}\dots x_{a_n})^\intercal \in \mathbb{R}^{n\times h}$ and we denote by $Y_n := (y_1 \dots y_n)^\intercal \in \mathbb{R}^{n\times d}$ the matrix gathering the vector of responses collected. We define the information matrix as $V_n := X_n^\intercal X_n = \sum_{i=1}^K T_n(i) x_i x_i^\intercal \in \mathbb{R}^{h\times h}$ where $T_i(n)$ denotes the number of observations from arm i among the n samples. More generally, given $\omega \in \mathbb{R}^K$, we define $V^\omega := \sum_{i=1}^K \omega(i) x_i x_i^\intercal$.

The multi-output regression model can be written in matrix form as $Y_n = X_n\Theta + H_n$ where $H_n = (\eta_1\dots\eta_n)^\intercal$ is the noise matrix. The least-square estimate $\widehat{\Theta}_n$ of the matrix Θ is defined as the matrix minimizing the least-square error $\mathrm{Err}_n(A) := \|X_nA - Y_n\|_\mathrm{F}^2$. Computing the gradient of the loss yields $V_n\widehat{\Theta}_n = X_n^\top Y_n$. If the matrix V_n is non-singular, the least-square estimator can be written

$$\widehat{\Theta}_n = V_n^{-1} X_n^\intercal Y_n.$$

In the course of our elimination algorithm, we will compute least-square estimates based on observation from a restricted number of arms, and we will face the case in which V_n is singular. In this case, different choices have been made in prior work on linear bandits: Alieva et al. [2021] defines a custom "pseudo-inverse" while Yang and Tan [2022] define new contexts $\widetilde{x}_i$ that are projections of the x_i onto a sub-space of dimension $\mathrm{rank}(\mathcal{X}_S)$ where $\mathcal{X}_S := (x_i : i \in S)^\intercal$ and S is the set of arms that are active. We adopt an approach close to the latter which is described below. Let the singular-value decomposition of $(\mathcal{X}_S)^\intercal$ be $USV^\intercal$ where U, V are orthogonal matrices and $B := (u_1,\dots,u_m)$ is formed with the first m columns of U where $m = \mathrm{rank}(\mathcal{X}_S)$. We then define

$$V_n^\dagger := B(B^\intercal V_n B)^{-1}B^\intercal \quad \text{and} \quad \widehat{\Theta}_n = V_n^\dagger X_n^\intercal Y_n. \tag{3}$$

The following result addresses the statistical uncertainty of this estimator.

Lemma 1. *If the noise η_t has covariance $\Sigma \in \mathbb{R}^{d\times d}$ and $a_1,\dots,a_n$ are deterministically chosen then for any $x_i \in \{x_{a_1},\dots,x_{a_n}\}$,* $\mathrm{Cov}(\widehat{\Theta}_n^\intercal x_i) = \|x_i\|^2_{V_n^\dagger}\Sigma$.

Therefore, estimating all arms'mean uniformly efficiently amounts to pull $\{a_1,\dots,a_n\}$ to minimize $\max_{i\in S}\|x_i\|^2_{V_n^\dagger}$. The continuous relaxation of this problem is equivalent to computing an allocation

$$\omega_S^\star \in \operatorname*{argmin}_{\omega\in\boldsymbol{\Delta}_{|S|}} \max_{i\in S} \|\widetilde{x}_i\|^2_{(\widetilde{V}^w)^{-1}} \tag{4}$$

where $\widetilde{x}_i := B^\intercal x_i$, $\widetilde{V}^\omega := \sum_{i\in S}\omega(s_i)\widetilde{x}_i\widetilde{x}_i^\intercal$ and $i \mapsto s_i$ maps S to $\{1,\dots,|S|\}$. (4) is a G-optimal design over the features $(B^\intercal x_i, i\in S)$ and it can be interpreted as a distribution over S that yields a uniform estimation of the mean responses for (3). This is formalized in Appendix H.

3 Optimal design algorithms for linear PSI

Our elimination algorithms operate in rounds. They progressively eliminate a portion of arms and classify them as optimal or sub-optimal based on empirical estimation of their gaps. In each round, a sampling budget is allocated among the surviving arms based on a G-optimal design.

3.1 Optimal Designs and Gap Estimation

At round r, we denote by A_r the set of arms that are still active. To estimate the means and henceforth the gaps, we first compute an estimate of the matrix $\widehat{\Theta}_r$. This estimate is obtained by carefully sampling the arms using the integral rounding of a G-optimal design.

Algorithm 1: OptEstimator(S, N, κ)

Input: Subset $S \subset [K]$, sample size N, precision κ

Compute the transformed features $\widetilde{\mathcal{X}}_S = (B^\intercal x_i, i \in S)$ with B as defined in Section 2.2

Compute a G-optimal design $w_S^\star$ over the set $\widetilde{\mathcal{X}}$

Pull $(a_1, \dots, a_N) \leftarrow \texttt{ROUND}(N, \widetilde{\mathcal{X}}_S, \omega_S^\star, \kappa)$ and collect responses $y_1, \dots, y_N$

Compute $V_N^\dagger$ as in Eq. (3) and compute the OLS estimator on the samples collected

$$\widehat{\Theta} \leftarrow V_N^\dagger \sum_{t=1}^{N} x_{a_t}^\intercal y_t$$

return: $\widehat{\Theta}$

Algorithm 1 takes as input a set of arms S, a budget N and chooses some N arms to pull (with repetitions) based on an integer rounding of $w_S^\star$, a continuous G-optimal design over the set $\{\widetilde{x}_i, i \in S\}$ of (transformed) features associated to that arms. Several rounding procedures have been proposed in the literature and we use that of Allen-Zhu et al. [2017], henceforth referred to as ROUND. In Appendix H, we show that $\texttt{ROUND}(N, \widetilde{\mathcal{X}}_S, w_S^\star, \kappa)$ outputs a sequence of arms $a_1, \dots, a_N \in S$ such that $\max_{i\in S} ||x_i||^2_{V_N^\dagger} \leq (1+6\kappa)\frac{F_S(w_S^\star)}{N}$, where $F_S(w_S^\star)$ is the optimal value of (4). Using the Kiefer-Wolfowitz theorem [Kiefer and Wolfowitz, 1960], we further prove that $F_S(w_S^\star) = h_S$, the dimension of span$(\{x_i, i \in S\})$. This observation is crucial to prove the following concentration result, at the heart of our analysis.

Lemma 2. *Let $S \subset [K]$, $\kappa \in (0, 1/3]$ and $N \geq 5h_S/\kappa^2$ where $h_S = \dim(\text{span}(\{x_i : i \in S\}))$. The output $\widehat{\Theta}$ of* OptEstimator*$(S,\ N,\ \kappa)$ satisfies for all $\varepsilon > 0$ and $i \in S$*

$$\mathbb{P}\left(\|(\Theta - \widehat{\Theta})^\intercal x_i\|_\infty \geq \varepsilon\right) \leq 2d \exp\left(-\frac{N\varepsilon^2}{2(1+6\kappa)\sigma^2 h_S}\right).$$

Once the parameter $\widehat{\Theta}_r$ has been obtained as an output of Algorithm 1 with $S = A_r$ and an appropriate value of the budget N, we compute estimates of the mean vectors as $\widehat{\mu}_{i,r} := \widehat{\Theta}_r^\intercal x_i$ and the empirical Pareto set of active arms,

$$S_r := \{i \in A_r : \nexists j \in A_r : \widehat{\mu}_{i,r} \prec \widehat{\mu}_{j,r}\}.$$

In both the fixed-confidence and fixed-budget settings, at round r, after collecting new samples from the surviving arms, GEGE discards a fraction of the arms based on the empirical estimation of their gaps. We first introduce the empirical quantities used to compute the gaps:

$$\text{M}(i, j; r) := \max_{c\in[d]}[\widehat{\mu}_{i,r}(c) - \widehat{\mu}_{j,r}(c)] \quad \text{and} \quad \text{m}(i, j; r) := \min_{c\in[d]}[\widehat{\mu}_{j,r}(c) - \widehat{\mu}_{i,r}(c)].$$

We define for any arm $i \in A_r$,

$$\widehat{\Delta}_{i,r} := \begin{cases} \widehat{\Delta}^\star_{i,r} := \max_{j\in A_r} \text{m}(i, j; r) & \text{if } i \in A_r \backslash S_r \\ \widehat{\delta}^\star_{i,r} := \min_{j\in A_r\backslash\{i\}}[\text{M}(i, j; r) \wedge (\text{M}(j, i; r)_+ + (\widehat{\Delta}^\star_{i,r})_+)] & \text{if } i \in S_r \end{cases} \tag{5}$$

the empirical estimates of the gaps introduced earlier. Differently from BAI, as the size of the Pareto set is unknown, we need an accept/reject mechanism to classify any discarded arm, described in details in the next sections for the fixed budget and fixed-confidence versions.

Final output In both cases, letting A_r be the set of active arms and B_r be the set of arms already classified as optimal at the beginning of round r, GEGE outputs $B_{\tau+1} \cup A_{\tau+1}$ as the candidate Pareto optimal set, where τ denotes the final round. And $A_{\tau+1}$ contains at most one arm.

3.2 Fixed-budget algorithm

Algorithm 2, operates over $\lceil \log_2(h) \rceil$ rounds, with an equal budget of $T/\lceil \log_2(h) \rceil$ allocated per round. By construction $|A_{\lceil \log_2(h) \rceil+1}|= 1$. At the end of round r, the $\lceil h/2^r \rceil$ arms with the smallest empirical gaps are kept active while the remaining arms are discarded and classified as Pareto optimal (added to B_{r+1}) if they are empirically optimal (belonging to set S_r) and deemed sub-optimal otherwise. If a tie occurs, we break it to eliminate arms that are empirically sub-optimal. This is crucial to prove the guarantees on the algorithm, as sketched in Section 4.

Algorithm 2: GEGE: G-optimal Empirical Gap Elimination [fixed-budget]

Input: budget T
Initialize: let $A_1 \leftarrow [K], B_1 \leftarrow \emptyset, D_1 \leftarrow \emptyset$
for $r = 1$ ***to*** $\lceil \log_2(h) \rceil$ **do**
  Compute $\widehat{\Theta}_r \leftarrow \mathrm{OptEstimator}(A_r, T/\log_2(h), 1/3)$
  Compute S_r the empirical Pareto set and the empirical gaps $\widehat{\Delta}_{i,r}$ with Eq.(5)
  Compute A_{r+1} the set of $\left\lceil \frac{h}{2^r} \right\rceil$ arms in A_r with the smallest empirical gaps `// ties broken by keeping arms of` S_r
  Update $B_{r+1} \leftarrow B_r \cup \{S_r \cap (A_r \backslash A_{r+1})\}$ and $D_{r+1} \leftarrow D_r \cup \{(A_r \backslash A_{r+1}) \backslash S_r\}$
return: $B_{\lceil \log_2(h) \rceil+1} \bigcup A_{\lceil \log_2(h) \rceil+1}$

Theorem 1. *The probability of error of Algorithm 2 run with budget $T \geq 45h \log_2 h$ is at most*

$$\exp\left(-\frac{T}{1200\sigma^2 H_{2,lin}\lceil \log_2 h \rceil} + \log C(h, d, K)\right)$$

where $C(h, d, K) = 2d\left(K + \frac{h}{2} + \lceil \log_2 h \rceil\right)$.

To the best of our knowledge GEGE is the first algorithm with theoretical guarantees for fixed-budget linear PSI. Our result shows that in this setting, the probability of error scales only with the first h gaps. Kone et al. [2024] proposed EGE-SH, an algorithm for fixed-budget PSI in the unstructured setting whose probability of error is essentially upper-bounded by

$$\exp\left(-\frac{T}{288\sigma^2 H_2 \log_2 K} + \log(2d(K-1)|\mathcal{S}^\star|\log_2 K)\right).$$

Therefore, GEGE largely improves upon EGE-SH when $K \gg h$. Moreover, when $K = h$ and $x_1, \dots, x_K$ is the canonical $\mathbb{R}^h$-basis, both algorithms coincide, thus, GEGE can be seen as a generalization of EGE-SH.

We state below a lower bound for linear PSI in the fixed-budget setting, showing that GEGE is optimal in the worse case, up to constants and a $\log_2(h)$ factor.

Theorem 2. *Let $\mathbb{W}_H$ be the set of instances with complexity $H_{2,lin}$ at most H. For any budget T, letting $\widehat{S}_T^{\mathcal{A}}$ be the output of algorithm $\mathcal{A}$, it holds that*

$$\min_{\mathcal{A}} \max_{\nu \in \mathbb{W}_H} \mathbb{P}_\nu(\widehat{S}_T^{\mathcal{A}} \neq \mathcal{S}^\star(\nu)) \geq \frac{1}{4}\exp\left(-\frac{2T}{H\sigma^2}\right).$$

3.3 Fixed-confidence algorithm

At round r, Algorithm 3, allocates a budget t_r to compute an estimator $\widehat{\Theta}_r$ of Θ^* by calling Algorithm 1. t_r is computed so that through $\widehat{\Theta}_r$, the mean of each arm is estimated with precision $\varepsilon_r/4$ with probability larger than $1 - \delta_r$ (using Lemma 2). Then, the empirical Pareto set S_r, of the active arms is computed and the empirical gaps are updated following (5). At the end of round r, empirically optimal arms (those in S_r) whose empirical gap is larger than ε_r are discarded and classified as optimal (added to B_{r+1}). Empirically sub-optimal arms whose empirical gap is larger than $\varepsilon_r/2$ are also discarded and classified as sub-optimal (added to D_{r+1}).

Algorithm 3: GEGE: G-optimal Empirical Gap Elimination [fixed-confidence]

Initialize: $A_1 \leftarrow [K]$, $B_1 \leftarrow \emptyset$, $D_1 \leftarrow \emptyset$, $r \leftarrow 1$
while $|A_r| > 1$ **do**
  Let $\varepsilon_r \leftarrow 1/(2 \cdot 2^r)$ and $\delta_r \leftarrow 6\delta/\pi^2 r^2$ and $h_r \leftarrow \dim(\mathrm{span}(\{x_i : i \in A_r\}))$
  Update $t_r := \left\lceil \frac{32(1+3\varepsilon_r)\sigma^2 h_r}{\varepsilon_r^2} \log(\frac{|A_r|d}{2\delta_r}) \right\rceil$
  Compute $\widehat{\Theta}_r \leftarrow \mathrm{OptEstimator}(A_r, t_r, \varepsilon_r)$
  Compute S_r and the empirical gaps $\widehat{\Delta}_{i,r}$ with Eq. (5)
  Update $B_{r+1} \leftarrow B_r \cup \{i \in S_r : \widehat{\Delta}_{i,r} \geq \varepsilon_r\}$ and $D_{r+1} \leftarrow D_r \cup \{i \in A_r \backslash S_r : \widehat{\Delta}_{i,r} \geq \varepsilon_r/2\}$
  Update $A_{r+1} \leftarrow A_r \backslash (D_{r+1} \cup B_{r+1})$
  $r \leftarrow r + 1$
return: $B_r \cup A_r$

Theorem 3. *The following statement holds with probability at least $1 - \delta$: Algorithm 3 identifies the Pareto set using at most*

$$\log_2(2/\Delta_1) + \sum_{i=2}^{h} \frac{64\sigma^2}{\Delta_i^2} \log\left(\frac{Kd}{\delta} \log_2\left(\frac{2}{\Delta_i}\right)\right)$$

samples and $\lceil \log_2(1/\Delta_1) \rceil$ rounds.

This result shows that complexity of Algorithm 3 scales only with the first h gaps. In particular, when $K \gg h$ using our algorithm substantially reduces the sample complexity of PSI. In Table 1, we compare the sample complexity of GEGE to that of existing fixed-confidence PSI algorithms, showing that GEGE enjoys stronger guarantees than its competitors. We emphasize that both Kim et al. [2023] and Zuluaga et al. [2016] use uniform sampling and do not exploit an optimal design which prevents them from reaching the guarantees given in Theorem 3.

Table 1: Sample complexity up to constant multiplicative terms for different algorithms for PSI in the fixed-confidence setting.

Algorithm	Upper-bound on τ_δ	Linear PSI
Zuluaga et al. [2016]	$\left(\frac{h^2}{\Delta_{\min}^2}\right) \log^3\left(\frac{dK}{\delta}\right)$	✔
Kone et al. [2023]	$\sum_{i=1}^{K} \frac{1}{\Delta_i^2} \log(\frac{12Kd}{\delta} \log(\frac{1}{\Delta_i}))$	✘
Kim et al. [2023]	$\frac{h}{\Delta_{\min}^2} \log(\frac{d(h \vee K)}{\delta \Delta_{\min}^2})$	✔
GEGE (**Ours**)	$\sum_{i=1}^{h} \frac{1}{\Delta_i^2} \log(\frac{Kd}{\delta} \log_2(\frac{2}{\Delta_i}))$	✔

We state a lower bound showing that our algorithm is essentially minimax optimal for linear PSI.

Theorem 4. *For any $K, d, h \in \mathbb{N}$, there exists a set $\mathcal{B}(K, d, h)$ of linear PSI instances s.t for $\nu \in \mathcal{B}(K, d, h)$ and for any δ-correct algorithm for PSI, with probability at least $1 - \delta$,*

$$\tau_\delta^{\mathcal{A}} = \Omega\left(H_{1,lin}(\nu) \log(\delta^{-1})\right).$$

Remark 1. *When $K = h$ and $x_1, \ldots, x_K$ forms the canonical $\mathbb{R}^h$ basis we recover the classical PSI problem. We note that unlike its fixed-budget version, GEGE does not coincide with an existing PSI identification algorithm. Instead, it matches the optimal guarantees of Kone et al. [2023] while needing only $\lceil \log(1/\Delta_1) \rceil$ rounds of adaptivity, which is the first fixed-confidence PSI algorithm having this property. Such a batched algorithm may be desirable in some applications e.g. in clinical trials where measuring different biological indicators of efficacy can take time.*

4 A unified analysis of GEGE

Before sketching our proof strategy, we highlight a key property of PSI that makes the analysis differ from classical BAI settings. Let a be a (Pareto) sub-optimal arm. From (1), there exits $a^\star \in \mathcal{S}^\star$

such that $\Delta_a = \mathrm{m}(a, a^\star)$ and importantly, $a^\star$ could be the unique arm dominating a. Therefore, discarding $a^\star$ before a may result in the latter appearing as optimal in the remaining rounds, thus leading to mis-identification of the Pareto set.

To avoid this, an elimination algorithm for PSI should guarantee that if a sub-optimal arm a is active, then $a^\star$ is also active. We introduce the following event

$$\mathcal{P}_r := \{\forall\ s \le r : \forall i \in (\mathcal{S}^\star)^\mathsf{c},\ i \in A_s \Rightarrow i^\star \in A_r\}.$$

An important aspect of our proofs is to control the occurrence of $\mathcal{P}_\infty$ (by convention, if $\mathcal{P}_t$ holds and $A_s = \emptyset$ for any $s \ge t$ then $\mathcal{P}_\infty$ holds). The first step of the proof is to show that when $\mathcal{P}_r$ holds, we can control the deviations of the empirical gaps. We now define for $\eta > 0$, the good event

$$\mathcal{E}^r(\eta) = \left\{\forall\ i, j \in A_r : \|(\widehat{\Theta}_r - \Theta)^\intercal (x_i - x_j)\|_\infty \le \eta\right\}. \tag{6}$$

Letting $n_r = |A_r|$ and λ a constant to be specified, we introduce $\mathcal{E}_{\text{fb}}^\lambda := \cap_{r=1}^{\lceil \log_2(h) \rceil} \mathcal{E}^r(\lambda \Delta_{n_{r+1}+1})$ and $\mathcal{E}_{\text{fc}} := \cap_{r=1}^{\infty} \mathcal{E}^r(\varepsilon_r/2)$. We then prove by concentration and induction the following key result.

Proposition 1. *Let $\lambda \in (0, 1/5)$ and assume $\mathcal{E}_{\text{fc}}$ (resp. $\mathcal{E}_{\text{fb}}^\lambda$ in fixed-budget) holds. Then at any round r, $\mathcal{P}_r$ holds and for all arm $i \in A_r$,*

$$\widehat{\Delta}_{i,r} - \Delta_i \ge \begin{cases} -\eta_r & \text{if } i \in \mathcal{S}^\star \\ -\eta_r/2 & \text{else,} \end{cases} \qquad \text{where } \eta_r = \begin{cases} 2\lambda \Delta_{n_{r+1}+1} & \text{(fixed-budget)} \\ \varepsilon_r & \text{(fixed-confidence).} \end{cases}$$

Building on this result, we show that the recommendation of Algorithm 2 is correct on $\mathcal{E}_{\text{fb}}^\lambda$, so its probability of error is upper-bounded by $\inf_{\lambda \in (0,1/5)} \mathbb{P}(\mathcal{E}_{\text{fb}}^\lambda)$. We conclude the proof of Theorem 1 by upper bounding this probability (see Appendix D).

Similarly, using Proposition 1 we prove the correctness of Algorithm 3 on $\mathcal{E}_{\text{fc}}$: at any round r, $B_r \subset \mathcal{S}^\star$ and $D_r \subset (\mathcal{S}^\star)^\mathsf{c}$. To upper bound its sample complexity we need an additional result to control the size of A_r.

Lemma 3. *The following holds for Algorithm 3 on $\mathcal{E}_{\text{fc}}$: for all $p \in [K]$, after $\lceil \log(1/\Delta_p) \rceil$ rounds it remains less than p active arms. In particular, GEGE stops after at most $\lceil \log(1/\Delta_1) \rceil$ rounds.*

To get the sample complexity bound of Theorem 3 some extra arguments are needed. We sketch some elements below (the full proof is given in Appendix E.3). Assume $\mathcal{E}_{\text{fc}}$ holds and let τ_δ be the sample complexity of Algorithm 3. Lemma 3 yields $\tau_\delta \le \sum_{r=1}^{\lceil \log(1/\Delta_1) \rceil} \Omega(h_r/\varepsilon_r^2)$ with $h_r \le |A_r|$.

Using Lemma 3, we introduce "checkpoints rounds" between which we control $|A_r|$ and thus h_r. Let the sequence $(\alpha_s)_{s \ge 0}$ defined as $\alpha_0 = 0$ and $\alpha_s = \lceil \log_2(1/\Delta_{\lfloor h/2^s \rfloor}) \rceil$, for $s \ge 1$. Simple calculation yields $\alpha_{\lfloor \log_2(h) \rfloor} = \lceil \log_2(1/\Delta_1) \rceil$ and $\{1, \dots, \lceil \log_2(1/\Delta_1) \rceil\} = \cup_{s=1}^{\lfloor \log_2(h) \rfloor} [\![\alpha_{s-1}, \alpha_s]\!]$. Therefore

$$\tau_\delta \le \sum_{s=1}^{\lfloor \log_2(h) \rfloor} \sum_{r=\alpha_{s-1}+1}^{\alpha_s} \Omega(|A_r|/\varepsilon_r^2).$$

Now by Lemma 3, for $r > \alpha_s$, $|A_r| \le \lfloor h/2^s \rfloor$, so essentially $\tau_\delta \le \sum_{s=1}^{\lfloor \log_2(h) \rfloor} \Omega(4^{\alpha_s} \lfloor h/2^s \rfloor)$. Carefully re-indexing this sum and addressing some few more technicalities we obtain the result in Theorem 3. Showing that $\mathbb{P}(\mathcal{E}_{\text{fc}}) \ge 1 - \delta$ using Lemma 2 completes the proof.

5 Experiments

We evaluate GEGE on real-world and synthetic instances. In the fixed-budget setting we compare against EGE-SH and EGE-SR [Kone et al., 2024], two algorithms for unstructured PSI in fixed-budget setting, and a uniform sampling baseline. In the fixed-confidence setting we compare to APE [Kone et al., 2023], a fully adaptive algorithm for unstructured PSI and PAL [Zuluaga et al., 2013], an algorithm that uses Gaussian process modeling for PSI, instantiated with a linear kernel.

5.1 Experimental protocol

We describe below the datasets in our experiments and we detail our experimental setup.

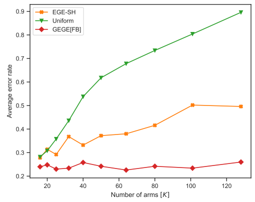

Figure 1: Average misidentification rate w.r.t K on the synthetic dataset

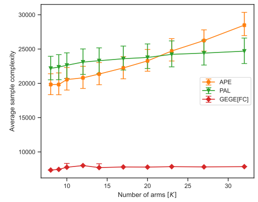

Figure 2: Average sample complexity w.r.t K in the synthetic experiment

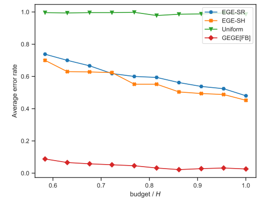

Figure 3: Average misidentification rate w.r.t T on NoC experiment

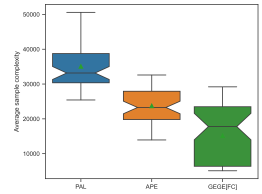

Figure 4: Empirical distribution of the sample complexity on the NoC dataset

Synthetic instances We fix features $x_1, \dots, x_h$ and Θ common to the instances described below. For any $K \geq h$ we define a linear PSI instance ν_K augmented with arms $x_{h+1}, \dots, x_K$ chosen so that arms $1, \dots, h$ have the same lowest gaps in ν_K. This implies that the complexity terms $H_{1,\text{lin}}$ and $H_{2,\text{lin}}$ are equal on such instances, irrespective of the number of arms. We set $h = 8, d = 2$.

Real-world dataset NoC [Almer et al., 2011] is a bi-objective optimization dataset for hardware design. The goal is to optimize $d = 2$ performance criteria: energy consumption and runtime of the implementation of a Network on Chip (NoC). The dataset contains $K = 259$ implementations, each of them described by $h = 4$ features.

On each instance, we report, for different algorithms, the empirical error probability (fixed-budget) and empirical distribution of the sample complexity (fixed-confidence), averaged over 500 seeded runs. We set $\delta = 0.01$ for the fixed-confidence experiments and $T = H_{2,\text{lin}}$ for fixed-budget.

5.2 Summary of the results

By Theorem 1 and 3, on the synthetic instance with K arms the sample complexity of GEGE should be a constant plus a $\log(K)$ term. This is coherent with what we observe: Fig.1 shows that the probability of error of GEGE merely increases with K whereas for EGE-SH/SR it grows much faster. Similarly, on Fig.2, the sample complexity of GEGE does not significantly increase with K, unlike that of APE. Therefore, GEGE only suffers a small cost for the number of arms.

For the real-world scenario, GEGE significantly outperforms its competitors in both settings. Fig.4 shows that it uses significantly fewer samples to identify the Pareto set compared to both APE and PAL. Fig.3 shows that the probability of misidentification of GEGE is reduced by up to 0.5 compared to EGE-SH. Moreover, it is worth noting that EGE-SH requires $T \geq K \log_2(K) \approx 2000$ (for NoC) to run on this instance while GEGE only needs $T \geq \log_2(h)$.

We reported runtimes around 10 seconds for single runs on instances with up to $K = 500, d = 8$ (cf Table 2 in Appendix I.1). The time and memory complexity of is addressed in Appendix I.1 and additional details about the implementation are provided. Appendix I.2 contains additional experimental results on a real-world multi-criteria optimization problem with $K = 768$ arms.

6 Conclusion and remarks

We have proposed the first algorithms for PSI in a multi-output linear bandit model that are guaranteed to outperform their un-structured counterparts. They leverage optimal design approaches to estimate the means vector and some sub-optimality gaps for PSI. In the fixed-budget setting GEGE is the first algorithm with nearly optimal guarantees for linear PSI. In the fixed-confidence setting, GEGE provably outperforms its competitors both in theory and in our experiments. It is also the first fixed-confidence PSI algorithm using a limited number of batches.

While the sample complexity of GEGE features a complexity term depending only on h gaps we still have $\log(K)$ terms due to union bounds. Katz-Samuels et al. [2020] showed that such union bounds can be avoided in linear BAI by using results from supremum of empirical processes. Further work could investigate if these observations would apply in linear PSI. In the alternative situation where $h \gg K$ for example in a RKHS, following the work of Camilleri et al. [2021], we could investigate how to extend this optimal design approach to PSI with high dimensional features.

References

A. Alieva, A. Cutkosky, and A. Das. Robust pure exploration in linear bandits with limited budget. In *Proceedings of the 38th International Conference on Machine Learning*, Proceedings of Machine Learning Research. PMLR, 2021.

Z. Allen-Zhu, Y. Li, A. Singh, and Y. Wang. Near-optimal design of experiments via regret minimization. In *Proceedings of the 34th International Conference on Machine Learning*. JMLR.org, 2017.

O. Almer, N. Topham, and B. Franke. A learning-based approach to the automated design of mpsoc networks. 2011.

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

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

M. Azizi, B. Kveton, and M. Ghavamzadeh. Fixed-budget best-arm identification in structured bandits. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*. International Joint Conferences on Artificial Intelligence Organization, 2022.

S. Bubeck and G. Munos, R.and Stoltz. Pure exploration for multi-armed bandit problems, 2008.

R. Camilleri, K. Jamieson, and J. Katz-Samuels. High-dimensional experimental design and kernel bandits. In *Proceedings of the 38th International Conference on Machine Learning*, Proceedings of Machine Learning Research. PMLR, 2021.

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

S. Chen, T. Lin, I. King, M. R. Lyu, and W. Chen. Combinatorial pure exploration of multi-armed bandits. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2014.

R. Degenne, P. Ménard, X. Shang, and M. Valko. Gamification of pure exploration for linear bandits. In *International Conference on Machine Learning (ICML)*, 2020.

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

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

T. Fiez, L. Jain, K. G. Jamieson, and L. Ratliff. Sequential experimental design for transductive linear bandits. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2019.

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

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

Y. Jedra and A. Proutiere. Optimal best-arm identification in linear bandits. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2020.

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

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

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

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

J. Katz-Samuels, L. Jain, z. karnin, and K. G. Jamieson. An empirical process approach to the union bound: Practical algorithms for combinatorial and linear bandits. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2020.

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

J. Kiefer and J. Wolfowitz. The equivalence of two extremum problems. *Canadian Journal of Mathematics*, 1960.

W. Kim, G. Iyengar, and A. Zeevi. Pareto front identification with regret minimization, 2023.

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

C. Kone, E. Kaufmann, and L. Richert. Bandit Pareto set identification: the fixed budget setting. In *Proceedings of The 27th International Conference on Artificial Intelligence and Statistics*, Proceedings of Machine Learning Research. PMLR, 2024.

T. Lattimore and C. Szepesvári. *Bandit Algorithms*. Cambridge University Press, 2020.

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

S. Lu, G. Wang, Y. Hu, and L. Zhang. Multi-objective generalized linear bandits. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*. AAAI Press, 2019.

F. Pukelsheim. *Optimal Design of Experiments (Classics in Applied Mathematics) (Classics in Applied Mathematics, 50)*. Society for Industrial and Applied Mathematics, 2006.

M. Soare, A. Lazaric, and R. Munos. Best-arm identification in linear bandits. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2014.

C. Tao, S. Blanco, and Y. Zhou. Best arm identification in linear bandits with linear dimension dependency. In *Proceedings of the 35th International Conference on Machine Learning*, Proceedings of Machine Learning Research. PMLR, 2018.

A. Tsanas and A. Xifara. Energy efficiency. UCI Machine Learning Repository, 2012.

L. Xu, J. Honda, and M. Sugiyama. A fully adaptive algorithm for pure exploration in linear bandits. In *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*. PMLR, 2018.

J. Yang and V. Tan. Minimax optimal fixed-budget best arm identification in linear bandits. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2022.

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

M. Zuluaga, G. Sergent, A. Krause, and M. Püschel. Active learning for multi-objective optimization. In *Proceedings of the 30th International Conference on Machine Learning*. PMLR, 2013.

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

A Outline

In section C, we prove Proposition 1, which is a crucial result to prove the guarantees of GEGE in fixed-confidence and fixed-budget settings. Section D proves the fixed-budget guarantees of GEGE, in particular Theorem 1. In section E we prove the fixed-confidence guarantees of GEGE by proving Theorem 3. Section F contains some ingredient concentration lemmas that are used in our proofs. In section G we analyze the lower bounds in both fixed-confidence and fixed-budget settings. In section H we analyze the properties of Algorithm 1 by using some results on G-optimal design. Finally section I contains additional experimental results and the detailed experimental setup.

B Notation

We introduce some additional notation used in the following sections.

In the subsequent sections, r will always denote a round of GEGE which should be clear from the context. We then denote by A_r active arms at round r and by $\widehat{\Theta}_r$ the empirical estimate of Θ at round r, computed by a call to Algorithm 1. By convention we let $\max_{\emptyset} = -\infty$.

For any sub-optimal arm i there exists a Pareto-optimal arm $i^\star$ (not necessarily unique) such that $\Delta_i = \mathrm{m}(i, i^\star)$. More generally given a sub-optimal i we denote by $i^\star$ any arm of $\operatorname{argmax}_{j\in\mathcal{S}^\star} \mathrm{m}(i,j)$.

At a round r we let

$$\mathcal{P}_r := \{\forall\, s \in \{1,\dots,r\},\ \forall\, i \in A_s, i \in (\mathcal{S}^\star)^{\mathsf{c}} \cap A_s \Rightarrow i^\star \in A_s\} \tag{7}$$

and $\mathcal{P} = \mathcal{P}_\infty$. In particular if for some τ, $\mathcal{P}_\tau$ is true and $A_{\tau+1} = \emptyset$ then we say that $\mathcal{P}$ holds.

C Proof of Proposition 1

We first recall the result.

Proposition 1. *Let $\lambda \in (0, 1/5)$ and assume $\mathcal{E}_{\text{fc}}$ (resp. $\mathcal{E}_{\text{fb}}^\lambda$ in fixed-budget) holds. Then at any round r, $\mathcal{P}_r$ holds and for all arm $i \in A_r$,*

$$\widehat{\Delta}_{i,r} - \Delta_i \geq \begin{cases} -\eta_r & \text{if } i \in \mathcal{S}^\star \\ -\eta_r/2 & \text{else,} \end{cases} \qquad \text{where } \eta_r = \begin{cases} 2\lambda\Delta_{n_{r+1}+1} & \text{(fixed-budget)} \\ \varepsilon_r & \text{(fixed-confidence).} \end{cases}$$

In both the fixed-budget and fixed-confidence setting, the proof proceeds by induction on the round r. Before presenting the inductive argument separately in each case, we establish in Appendix C.1 an important result that is used in both cases (Lemma 7): if $\mathcal{P}_r$ holds at some round r then, the empirical gaps computed at this round are good estimators of the true PSI gaps.

To establish this first result, we need the following intermediate lemmas, proved in Appendix F.

Lemma 4. *At any round r and for any arms $i, j \in A_r$ it holds that*

$$|\mathrm{M}(i,j;r) - \mathrm{M}(i,j)| \leq \|(\widehat{\Theta}_r - \Theta)^\intercal (x_i - x_j)\|_\infty \text{ and}$$
$$|\mathrm{m}(i,j;r) - \mathrm{m}(i,j)| \leq \|(\widehat{\Theta}_r - \Theta)^\intercal (x_i - x_j)\|_\infty.$$

Lemma 5. *At any round 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 $\Delta_i^\star < \|(\widehat{\Theta}_r - \Theta)^\intercal (x_i - x_{i^\star})\|_\infty$.*

C.1 Deviations of the gaps when $\mathcal{P}_r$ holds

In this part, we control the deviations of the empirical gaps when proposition $\mathcal{P}_r$ holds.

Lemma 6. *Assume that the proposition $\mathcal{P}_r$ holds at some round r. Then for any arm $i \in A_r$ it holds that*

$$\left|(\widehat{\Delta}_{i,r}^\star)_+ - (\Delta_i^\star)_+\right| \leq \left|\widehat{\Delta}_{i,r}^\star - \Delta_i^\star\right| \leq \gamma_{i,r}$$

where $\gamma_{i,r} := \max_{j\in A_r} \|(\widehat{\Theta}_r - \Theta)^\intercal (x_i - x_j)\|_\infty$.

Proof. This inequality is a direct consequence of Lemma 4 and the relation $|x_+ - y_+| \le |x - y|)$ which holds for any $x, y \in \mathbb{R}$. Note that for a Pareto-optimal arm i we trivially have $(\Delta_i^\star)^+ = 0 = (\max_{j\in A_r} \mathrm{m}(i,j))_+$. And for a sub-optimal arm $i \in A_r$, as $i^\star \in A_r$ (from proposition $\mathcal{P}_r$) we have $\Delta_i^\star = \mathrm{m}(i, i^\star) = \max_{j\in A_r} \mathrm{m}(i,j)$. Thus for any arm $i \in A_r$ we have

$$\begin{aligned}\left|(\widehat{\Delta}_{i,r}^\star)_+ - (\Delta_i^\star)_+\right| &= \left|(\max_{j\in A_r} \mathrm{m}(i,j;r))_+ - (\max_{j\in A_r} \mathrm{m}(i,j))_+\right|, \\ &\le \left|(\max_{j\in A_r} \mathrm{m}(i,j;r)) - (\max_{j\in A_r} \mathrm{m}(i,j))\right|, \\ &\le \max_{j\in A_r} |\mathrm{m}(i,j;r) - \mathrm{m}(i,j)|, \\ &\le \max_{j\in A_r} \left\|(\widehat{\Theta}_r - \Theta)^\intercal (x_i - x_j)\right\|_\infty = \gamma_{i,r},\end{aligned}$$

where the last inequality follows from Lemma 4. $\square$

Lemma 7. *If the proposition $\mathcal{P}_r$ holds at some round r then for any arm $i \in A_r$,*

$$\widehat{\Delta}_{i,r} - \Delta_i \ge \begin{cases} -2\gamma_r & \text{if } i \in \mathcal{S}^\star, \\ -\gamma_{i,r} & \text{else,} \end{cases}$$

where $\gamma_{i,r} := \max_{j\in A_r} \|(\widehat{\Theta}_r - \Theta)^\intercal (x_i - x_j)\|_\infty$ and $\gamma_r := \max_{i\in A_r} \gamma_{i,r}$.

Proof. We first prove the result a sub-optimal arm i. From the proposition $\mathcal{P}_r$, as $i \in A_r$ we have $i^\star \in A_r$ so $\Delta_i = \max_{j\in A_r} \mathrm{m}(i,j)$ and we recall that

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

Note that by reverse triangle we have for any arm $i \in A_r$ (sub-optimal or not)

$$\begin{aligned}\left|\left(\max_{j\in A_r} \mathrm{m}(i,j;r)\right) - \left(\max_{j\in A_r} \mathrm{m}(i,j)\right)\right| &\le \max_{j\in A_r} |\mathrm{m}(i,j;r) - \mathrm{m}(i,j)|, & (9)\\ &\le \max_{j\in A_r} \|(\widehat{\Theta}_r - \Theta)^T (x_i - x_j)\|_\infty = \gamma_{i,r}. & (10)\end{aligned}$$

where the last inequality follows from Lemma 4. If i a sub-optimal arm ($i \notin \mathcal{S}^\star$) then as $\Delta_i = \Delta_i^\star$, it follows

$$\widehat{\Delta}_{i,r} - \Delta_i \ge \widehat{\Delta}_{i,r}^\star - \Delta_i^\star$$

therefore

$$\begin{aligned}\widehat{\Delta}_{i,r} - \Delta_i &\ge -|\widehat{\Delta}_{i,r}^\star - \Delta_i^\star| \\ &= -|(\max_{j\in A_r} \mathrm{m}(i,j;r)) - (\max_{j\in A_r} \mathrm{m}(i,j))| \\ &\ge -\gamma_{i,r} \quad \text{(see (10))}\end{aligned}$$

Now we assume i is a Pareto-optimal arm ($i \in \mathcal{S}^\star$) so that now

$$\Delta_i = \delta_i^\star.$$

Combining with Eq. (54) yields

$$\widehat{\Delta}_{i,r} - \Delta_{i,r} \ge \widehat{\delta}_{i,r}^\star - \delta_{i,r}^\star,$$

where we recall that

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

and

$$\delta_i^\star := \min_{j\in[K]\setminus\{i\}} [\mathrm{M}(i,j) \wedge (\mathrm{M}(j,i)_+ + (\Delta_j^\star)_+)].$$

As for any $x, y \in \mathbb{R}$ we have $|x^+ - y^+| \leq |x - y|$, the following holds for any $i, j \in A_r$

$$\begin{aligned} |\mathrm{M}(j,i;r)^+ - \mathrm{M}(j,i)^+| &\leq |\mathrm{M}(j,i;r) - \mathrm{M}(j,i)| \quad (11) \\ &\leq \gamma_{j,r}. \quad (12) \end{aligned}$$

From Lemma 6 we have for any $j \in A_r$

$$(\widehat{\Delta}^\star_{j,r})_+ - (\Delta^\star_j)_+ \geq -\gamma_{j,r}. \quad (13)$$

Combining (12) and (13) yields for any $j \in A_r$

$$\mathrm{M}(j,i;r)_+ + (\widehat{\Delta}^\star_{j,r})_+ \geq \mathrm{M}(j,i)_+ + (\Delta^\star_j)_+ - 2\gamma_{j,r}, \quad (14)$$

which in addition to $\mathrm{M}(j,i;r) \geq \mathrm{M}(j,i) - \gamma_{j,r}$ yields

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

for any arm $j \in A_r$. Thus taking the min over A_r yields

$$\begin{aligned} \widehat{\delta}^\star_{i,r} &= \min_{j \in A_r \backslash \{i\}} [\mathrm{M}(i,j;r) \wedge (\mathrm{M}(j,i;r)_+ + (\widehat{\Delta}^\star_{j,r})_+)] \\ &\geq \min_{j \in A_r \backslash \{i\}} [\mathrm{M}(i,j) \wedge (\mathrm{M}(j,i)_+ + (\Delta^\star_j)_+)] - 2\gamma_r, \\ &\geq \min_{j \in [K] \backslash \{i\}} [\mathrm{M}(i,j) \wedge (\mathrm{M}(j,i)_+ + (\Delta^\star_j)_+)] - 2\gamma_r, \\ &= \delta^\star_i - 2\gamma_r \end{aligned}$$

which concludes the proof the proof of this lemma. □

Building on this result, we show that $\mathcal{P}_\infty$ holds in the fixed-confidence and fixed-budget settings.

C.2 Fixed-budget setting

We recall the definition of the good event for any $\lambda > 0$.

$$\mathcal{E}^{r,\lambda}_{\text{fb}} = \left\{ \forall\, i, j \in A_r : \|(\widehat{\Theta}_r - \Theta)^\intercal (x_i - x_j)\|_\infty \leq \lambda \Delta_{n_{r+1}+1} \right\}$$

and $\mathcal{E}^\lambda_{\text{fb}} := \cap_{r=1}^{\lceil \log_2(h) \rceil} \mathcal{E}^{r,\lambda}_{\text{fb}}$. We prove that proposition $\mathcal{P}_\infty$ holds on the event $\mathcal{E}^\lambda_{\text{fb}}$ for some any $\lambda \in (0, 1/5)$.

Lemma 8. *The proposition holds $\mathcal{P}_\infty$ on the event $\mathcal{E}^\lambda_{\text{fb}}$ for any $\lambda \in (0, 1/5)$: at any round $r \in \{1, \dots, \lceil \log_2 h \rceil + 1\}$ and for any arm $i \in A_r \cap (\mathcal{S}^\star)^\mathsf{c}$, $i^\star \in A_r$.*

Proof. We prove $\mathcal{P}_\infty$ by induction on the round r. In the sequel we assume $\mathcal{E}^\lambda_{\text{fb}}$ holds. We also assume $\mathcal{P}_r$ is true until some round r. As $\mathcal{E}^\lambda_{\text{fb}}$ holds, we have by application of Lemma 7: for any arm $i \in A_r$,

$$\widehat{\Delta}_{i,r} - \Delta_i \geq \begin{cases} -2\lambda \Delta_{n_{r+1}+1} & \text{if } i \in \mathcal{S}^\star \\ -\lambda \Delta_{n_{r+1}+1} & \text{else.} \end{cases} \quad (15)$$

We shall prove that if a Pareto-optimal arm i is discarded at the end of round r then there exists no arm sub-optimal $j \in A_{r+1}$ such that $j^\star = i$. Since i is removed and $|A_{r+1}| = n_{r+1}$ there exists $k_r \in A_{r+1} \cup \{i\}$ such that

$$\Delta_{k_r} \geq \Delta_{n_{r+1}+1}. \quad (16)$$

If i is empirically sub-optimal then as it is discarded we have

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

for any arm $k \in A_{r+1}$. So $\widehat{\Delta}^\star_{i,r} \geq \widehat{\Delta}_{k_r,r}$ thus using (15) and (16) it comes that

$$\begin{aligned} \max_{q \in A_r \backslash \{i\}} \mathrm{m}(i,q) &\geq \Delta_{n_{r+1}+1} - 3\lambda \Delta_{n_{r+1}+1} \\ &= (1 - 3\lambda)\Delta_{n_{r+1}+1} \end{aligned}$$

and the latter inequality is not possible for $\lambda < 1/3$ as the LHS of the inequality is negative as i is a Pareto-optimal arm.

Next we assume that i is empirically optimal. We claim that j is not dominated by i. To see this, first note that as $j \in A_{r+1}$ we have

$$\widehat{\Delta}_{i,r} \geq \widehat{\Delta}_{j,r} \tag{17}$$

so that as i is empirically optimal, if j was empirically dominated by i we would have

$$\widehat{\Delta}_{i,r} \leq \mathrm{M}(j,i;r)_+ + (\widehat{\Delta}^\star_{j,r})_+ = \widehat{\Delta}_{j,r}. \tag{18}$$

Combining (17) and (18) yield $\widehat{\Delta}_{i,r} = \widehat{\Delta}_{j,r}$, i is empirically optimal and j is empirically sub-optimal. However our breaking rule ensures that this case cannot occur. Therefore j is not dominated by i. But, by assumption, j is such that $j^\star = i$ and we have proved that i does not empirically dominate j so by Lemma 5

$$\Delta_j \leq \|(\widehat{\Theta}_r - \Theta)^\intercal (x_j - x_i)\|_\infty$$

which on the event $\mathcal{E}_{\text{fb}}$ yields

$$\Delta_j \leq \lambda \Delta_{n_{r+1}+1}. \tag{19}$$

On the other side, as i is discarded as an empirically optimal arm we have

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

for any arm $k \in A_{r+1}$. Since $k_r \in A_{r+1} \cup \{i\}$ it comes $\widehat{\delta}^\star_{i,r} \geq \widehat{\Delta}_{k_r,r}$ thus using (15) and (16) yields

$$\mathrm{M}(j,i)_+ + \Delta_j \geq \Delta_{n_{r+1}+1} - 4\lambda \Delta_{n_{r+1}+1}$$

which further combined with (19) yields

$$\mathrm{M}(j,i)_+ \geq (1 - 5\lambda)\Delta_{n_{r+1}+1}.$$

However, as $j^\star = i$ we have $\mathrm{M}(j,i)_+ = 0$ so the latter inequality is not possible as long as $\lambda < 1/5$. Put together, we have proved proved that if $\mathcal{P}_r$ holds then for any Pareto-optimal arm i which is removed at the end of round r, there does not exist an arm $j \in A_{r+1}$ such that $j^\star = i$. So $\mathcal{P}_{r+1}$ holds. Finally noting that $\mathcal{P}_r$ trivially holds for $r = 1$ we conclude that $\mathcal{P}_\infty$ holds on the event $\mathcal{E}^\lambda_{\text{fb}}$ for any $\lambda < 1/5$. □

Combining this result with Lemma 7 and assuming $\mathcal{E}^\lambda_{\text{fb}}$ holds then yields at any round $r \in \{1, \dots, \lceil \log_2 h \rceil\}$ and for any arm $i \in A_r$:

$$\widehat{\Delta}_{i,r} - \Delta_i \geq \begin{cases} -2\lambda \Delta_{n_{r+1}+1} & \text{if } i \in \mathcal{S}^\star \\ -\lambda \Delta_{n_{r+1}+1} & \text{else,} \end{cases} \tag{20}$$

which proves Proposition 1 in the fixed-budget setting.

C.3 Fixed-confidence setting

We recall below the good events we study in the fixed-confidence setting:

$$\mathcal{E}^r_{\text{fc}} = \left\{ \forall\, i,j \in A_r : \|(\widehat{\Theta}_r - \Theta)^\intercal (x_i - x_j)\|_\infty \leq \varepsilon_r/2 \right\}$$

and $\mathcal{E}_{\text{fc}} := \cap_{r=1}^\infty \mathcal{E}^r_{\text{fc}}$.

Lemma 9. *The proposition $\mathcal{P}_\infty$ holds on the event $\mathcal{E}_{fc}$: at any round r for any arm $i \in A_r \cap (\mathcal{S}^\star)^{\mathsf{c}}$, $i^\star \in A_r$.*

Proof of Lemma 9. We prove the proposition by induction on the round r. Note that the proposition $\mathcal{P}_r$ trivially holds for $r = 1$. Assume the property holds until the beginning of some round r. Let $i \in \mathcal{S}^\star$ be an optimal arm and assume i is discarded at the end of round r. We will prove that there exists no sub-optimal arm $j \in A_{r+1}$ such that $j^\star = i$. Recall that when i is discarded, we have either $i \in S_r$ (empirically optimal) or $i \notin S_r$ (empirically sub-optimal). We analyze both cases below. If $i \notin S_r$ then it holds that

$$\widehat{\Delta}_{i,r} \geq \varepsilon_r/2,$$

then, as $i \notin S_r$ it follows that $\widehat{\Delta}_{i,r} = \widehat{\Delta}^\star_i := \max_{j\in A_r\setminus\{i\}} \mathrm{m}(i,j;r)$, so

$$\max_{j\in A_r\setminus\{i\}} \mathrm{m}(i,j;r) \geq \varepsilon_r/2$$

which using Lemma 4 and assuming event $\mathcal{E}^r_{\text{fc}}$ holds would yield

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

The latter inequality is not possible as $i \in \mathcal{S}^\star$ is a Pareto-optimal arm. Therefore, on $\mathcal{E}^r_{\text{fc}}$, when $i \in \mathcal{S}^\star$ is discarded we have $i \in S_r$.

Next, we analyze the case $i \in S_r$: that is i is discarded and classified as optimal. In this case it follows from the definition of $\widehat{\Delta}_{i,r}$ that

$$\min_{j\in A_r\setminus\{i\}} [\mathrm{M}(j,i;r)_+ + (\widehat{\Delta}^\star_{j,r})_+] \geq \varepsilon_r. \tag{21}$$

Let $j \in A_{r+1} \cap (\mathcal{S}^\star)^c$ be such that $j^\star = i$. If j is empirically optimal then $(\widehat{\Delta}^\star_{j,r})_+ = 0$ thus $\mathrm{M}(j,i;r)_+ \geq \varepsilon_r$. On the contrary, if j is empirically sub-optimal then because it has not been removed at the end of round r it holds that

$$\widehat{\Delta}^\star_{j,r} < \varepsilon_r/2,$$

which combined with (21) yields $\mathrm{M}(j,i;r)_+ > \varepsilon_r/2$. Thus, in both cases we have $\mathrm{M}(j,i;r)_+ > \varepsilon_r/2$ which using Lemma 4 and assuming event $\mathcal{E}^r_{\text{fc}}$ would imply that

$$\mathrm{M}(j,i)_+ > 0,$$

which is impossible as, by assumption $j^\star = i$, so j is dominated by i.

Put together with what precedes, on $\mathcal{E}_{\text{fc}}$, if $\mathcal{P}_r$ holds then $\mathcal{P}_{r+1}$ holds. Since the property trivially holds for $r = 1$ we have proved that the property $\mathcal{P}_r$ holds at any round when $\mathcal{E}_{\text{fc}}$ holds. □

Combining this result with Lemma 7 proves that, on the event $\mathcal{E}_{\text{fc}}$, for any round r and for any arm $i \in A_r$

$$\widehat{\Delta}_{i,r} - \Delta_i \geq \begin{cases} -\varepsilon_r & \text{if } i \in \mathcal{S}^\star \\ -\varepsilon_r/2 & \text{else,} \end{cases} \tag{22}$$

which proves Proposition 1 in the fixed-confidence setting.

D Upper bound on the probability of error

In this section, we prove the theoretical guarantees of GEGE in the fixed-budget setting. We prove Theorem 1 and some ingredient lemmas.

Theorem 1. *The probability of error of Algorithm 2 run with budget $T \geq 45h\log_2 h$ is at most*

$$\exp\left(-\frac{T}{1200\sigma^2 H_{2,lin}\lceil\log_2 h\rceil} + \log C(h,d,K)\right)$$

where $C(h,d,K) = 2d\left(K + \frac{h}{2} + \lceil\log_2 h\rceil\right)$.

Proof of Theorem 1. We first prove the correctness of GEGE on the event $\mathcal{E}^\lambda_{\text{fb}}$ for some λ small enough. Let us assume $\mathcal{E}^\lambda_{\text{fb}}$ holds which by Proposition 1 implies that $\mathcal{P}_\infty$ holds and at round r, we have for any arm $i \in A_r$

$$\widehat{\Delta}_{i,r} - \Delta_i \geq \begin{cases} -2\lambda\Delta_{n_{r+1}+1} & \text{if } i \in \mathcal{S}^\star \\ -\lambda\Delta_{n_{r+1}+1} & \text{else.} \end{cases} \tag{23}$$

We recall the definition of the good event for any $\lambda > 0$,

$$\mathcal{E}^{r,\lambda}_{\text{fb}} = \left\{\forall\, i,j \in A_r : \|(\widehat{\Theta}_r - \Theta)^\intercal (x_i - x_j)\|_\infty \leq \lambda\Delta_{n_{r+1}+1}\right\}$$

and $\mathcal{E}_{\text{fb}} := \cap_{r=1}^{\lceil \log_2(h) \rceil} \mathcal{E}_{\text{fb}}^{r,\lambda}$. Applying Lemma 4 on this event then yields for all arms $i, j \in A_r$,

$$|\mathrm{M}(i,j;r) - \mathrm{M}(i,j)| \leq \lambda \Delta_{n_{r+1}+1} \text{ and} \tag{24}$$
$$|\mathrm{m}(i,j;r) - \mathrm{m}(i,j)| \leq \lambda \Delta_{n_{r+1}+1}. \tag{25}$$

Let i be an arm discarded at the end of round r. Since i is discarded and $|A_{r+1}| = n_{r+1}$ there exists $k_r \in A_{r+1} \cup \{i\}$ such that

$$\Delta_{k_r} \geq \Delta_{n_{r+1}+1}. \tag{26}$$

If $i \notin S_r$ that is i is empirically sub-optimal then

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

then, recalling that

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

and further applying (23) to k_r and using (25) yields

$$\max_{j \in A_r \setminus \{i\}} \mathrm{m}(i,j) \geq (1 - 3\lambda) \Delta_{n_{r+1}+1}$$

which for $\lambda < 1/3$ implies that $\max_{j \in A_r} \mathrm{m}(i,j) > 0$, that is there exists $j \in A_r$ such that $\mu_i \prec \mu_j$ so i is a sub-optimal arm.

Next, assume $i \in S_r$ (i.e i is empirically Pareto-optimal). In this case we have $\widehat{\Delta}_{i,r} = \widehat{\delta}^\star_{i,r} \geq \widehat{\Delta}_{k_r,r}$. We recall that

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

Applying (23) to k_r and using (24), it follows that

$$\min_{j \in A_r \setminus \{i\}} \mathrm{M}(i,j) \geq (1 - 3\lambda) \Delta_{n_{r+1}+1}.$$

Thus, for $\lambda < 1/3$, we have $\min_{j \in A_r \setminus \{i\}} \mathrm{M}(i,j) > 0$. Therefore, no active arm at round r dominates i which together with proposition $\mathcal{P}_\infty$ yields that i is a Pareto-optimal arm (otherwise, we would have $i^\star \in A_r$ that dominates i).

All put together, we have proved that for any $\lambda < 1/5$ (we need $\lambda < 1/5$ for $\mathcal{P}_\infty$ to hold), Algorithm 2 does not make any error on the event $\mathcal{E}_{\text{fb}}^\lambda$. It then follows that the probability of error of GEGE is at most

$$\inf_{\lambda \in (0,1/5)} \mathbb{P}\left((\mathcal{E}_{\text{fb}}^\lambda)^c\right) \tag{27}$$

Now we upper-bound Eq. (27) which will conclude the proof. Let $\lambda \in (0, 1/5)$ be fixed. We have by union bound

$$\begin{aligned} \mathbb{P}\left((\mathcal{E}_{\text{fb}}^\lambda)^c\right) &\leq \sum_{r=1}^{\lceil \log_2 h \rceil} \mathbb{E}\left[\mathbb{P}\left((\mathcal{E}_{\text{fb}}^{r,\lambda})^c | A_r\right)\right] \\ &\leq \sum_{r=1}^{\lceil \log_2 h \rceil} \mathbb{E}\left[\sum_{i \in A_r} \mathbb{P}(\|(\widehat{\Theta}_r - \Theta)^\intercal x_i\|_\infty > \frac{1}{2}\lambda \Delta_{n_{r+1}+1} | A_r)\right] \end{aligned}$$

Note that for i fixed, we can use Lemma 2 with $\kappa = 1/3$ and the conditions of this theorem are satisfied as the budget per phase is $T/\log_2(h) \geq 45h$ (recall from the theorem that GEGE is run with $T \geq 45h \log_2(h)$). Thus applying this theorem yields

$$\begin{aligned} \mathbb{P}\left((\mathcal{E}_{\text{fb}}^\lambda)^\mathsf{c}\right) &\leq 2d \sum_{r=1}^{\lceil \log_2 h \rceil} n_r \mathbb{E}\left[\exp\left(-\frac{\lambda^2 \Delta^2_{n_{r+1}+1} T}{24\sigma^2 h_r \log_2 h\rceil}\right)\right] \\ &\leq 2d \sum_{r=1}^{\lceil \log_2 h \rceil} n_r \exp\left(-\frac{\lambda^2 T \Delta^2_{n_{r+1}+1}}{24\sigma^2 \min(h, n_r) \lceil \log_2 h \rceil}\right), \quad \text{as } h_r \leq \min(n_r, h) \end{aligned}$$

Then, note that

$$\begin{aligned}\frac{\Delta^2_{n_{r+1}+1}}{\min(h,n_r)} &= \frac{\Delta^2_{\lceil h/2^r\rceil+1}}{\lceil h/2^{r-1}\rceil}\\ &= \frac{\Delta^2_{\lceil h/2^r\rceil+1}}{\lceil h/2^r\rceil+1}\frac{\lceil h/2^r\rceil+1}{\lceil h/2^{r-1}\rceil}\\ &\geq \frac{\Delta^2_{\lceil h/2^r\rceil+1}}{\lceil h/2^r\rceil+1}\frac{h/2^r+1}{h/2^{r-1}+1}\\ &\geq \frac{\Delta^2_{\lceil h/2^r\rceil+1}}{\lceil h/2^r\rceil+1}\frac{1}{2},\end{aligned}$$

which follows as $(x+1)/(2x+1) \geq \frac{1}{2}$ for $x \geq 1$. Therefore,

$$\begin{aligned}\frac{\Delta^2_{n_{r+1}+1}}{\min(h,n_r)} &\geq \frac{1}{2}\frac{\Delta^2_{\lceil h/2^r\rceil+1}}{\lceil h/2^r\rceil+1}\\ &\geq \frac{1}{2H_{2,\text{lin}}}.\end{aligned}$$

Finally,

$$\begin{aligned}\mathbb{P}\left((\mathcal{E}^\lambda_{\text{fb}})^c\right) &\leq 2\exp\left(-\frac{\lambda^2 T}{48\sigma^2 H_{2,\text{lin}}\lceil\log_2 h\rceil}+\log(d)\right)\sum_{r=1}^{\lceil\log_2 h\rceil} n_r\\ &\leq 2\left(K+\frac{h}{2}+\lceil\log_2 h\rceil\right)\exp\left(-\frac{\lambda^2 T}{48\sigma^2 H_{2,\text{lin}}\lceil\log_2 h\rceil}+\log(d)\right)\end{aligned}$$

Finally it follows that

$$\inf_{\lambda\in(0,1/5)}\mathbb{P}\left((\mathcal{E}^\lambda_{\text{fb}})^c\right) \leq 2\left(K+\frac{h}{2}+\lceil\log_2 h\rceil\right)\exp\left(-\frac{T}{1200\sigma^2 H_{2,\text{lin}}\lceil\log_2 h\rceil}+\log(d)\right),$$

which concludes the proof. □

E Upper bound on the sample complexity

We prove the theoretical guarantees in the fixed-confidence setting. We prove the correctness of Algorithm 3 and we prove the sample complexity bound of Theorem 3 and some key lemmas. We first prove the correctness of the fixed-confidence variant of GEGE.

E.1 Proof of the correctness

We need to prove that the final recommendation of Algorithm 3 is correct: that is we should show that : at any round r, $B_r \subset \mathcal{S}^\star$ and $D_r \subset (\mathcal{S}^\star)^\mathsf{c}$.

Lemma 10. *On the event $\mathcal{E}_{\text{fc}}$, Algorithm 3 identifies the correct Pareto set.*

Proof of Lemma 10. In this part let τ denotes the stopping time of Algorithm 3. We assume $\mathcal{E}_{\text{fc}}$ holds. Using Proposition 1 : for any round $r \leq \tau$ for any (Pareto) sub-optimal $i \in A_r$ we have $i^\star \in A_r$. We then prove the correctness of the algorithm as follows. Let i be an arm that is removed at the end of some round r. Assume $i \in S_r$ then, as i is discarded and empirically optimal we have $\widehat{\Delta}_{i,r} = \widehat{\delta}^\star_i \geq \varepsilon_r$. In particular, it holds that

$$\min_{j\in A_r\setminus\{i\}} \mathrm{M}(i,j;r) \geq \varepsilon_r$$

which using Lemma 4 on the event $\mathcal{E}_{\text{fc}}$ yields

$$\min_{j\in A_r\setminus\{i\}} \mathrm{M}(i,j) > \varepsilon_r/2 > 0,$$

that is no active arm dominates i. Put together with proposition $\mathcal{P}_\infty$ (cf Lemma 9) the latter inequality yields $i \in \mathcal{S}^\star$. Now assume we have $i \notin S_r$: i is discarded and it is empirically sub-optimal. Then

$$\widehat{\Delta}_{i,r} = \max_{j \in A_r} \mathrm{m}(i, j; r) \geq \varepsilon_r/2,$$

so using Lemma 4 again on event $\mathcal{E}_{\text{fc}}$ it follows that there exists $j \in A_r$ such that $\mathrm{m}(i, j) > 0$: that is $i \notin \mathcal{S}^\star$. Put together, we have proved that if $\mathcal{E}_{\text{fc}}$ holds then for any arm i discarded at some round r,

$$i \in B_{r+1} \iff i \in \mathcal{S}^\star.$$

Note that if A_τ is non-empty then it contains a single arm and because $\mathcal{P}_\infty$ holds, this arm is also Pareto optimal. □

Thus, Algorithm 3 is correct on $\mathcal{E}_{\text{fc}}$. Before proving Theorem 3 we need Lemma 3 to control the size of the active set A_r in the fixed-confidence setting.

E.2 Controlling the size of the active set

We prove the following result that controls the size of the active set.

Lemma 3. *The following holds for Algorithm 3 on $\mathcal{E}_{\text{fc}}$: for all $p \in [K]$, after $\lceil \log(1/\Delta_p) \rceil$ rounds it remains less than p active arms. In particular, GEGE stops after at most $\lceil \log(1/\Delta_1) \rceil$ rounds.*

Proof of Lemma 3. By Lemma 9 we on the event $\mathcal{E}_{\text{fc}}$: for any round r and for any arm $i \in A_r$,

$$\widehat{\Delta}_{i,r} - \Delta_i \geq \begin{cases} -\varepsilon_r & \text{if } i \in \mathcal{S}^\star \\ -\varepsilon_r/2 & \text{else.} \end{cases}$$

Then let $p \in [K]$ and let assume an arm $i \in \{p, \dots, K\}$ is still active at round $r = \lceil \log_2(1/\Delta_p) \rceil$. We have $\widehat{\Delta}_{i,r} \geq \Delta_i - \varepsilon_r$ with $\varepsilon_r = 1/2^{r+1}$ and $\Delta_i \geq \Delta_p$ which combined with $\widehat{\Delta}_{i,r} \geq \Delta_i - \varepsilon_r$ yields

$$\widehat{\Delta}_{i,r} \geq \Delta_p - \varepsilon_r. \tag{28}$$

As $r = \lceil \log_2(1/\Delta_p)$ it holds that $2\varepsilon_r \leq \Delta_p$ so Eq. (28) yields $\widehat{\Delta}_{i,r} \geq \varepsilon_r$ thus i will be discarded at the end of round r that is any arm $i \in \{p, \dots, K\}$ will be discarded at the end of round $\lceil \log_2(1/\Delta_p) \rceil$. □

We now prove the main lemma on the sample complexity of GEGE in the fixed-confidence setting.

E.3 Proof of Theorem 3

We provide an upper bound on the sample complexity of the algorithm.

Theorem 3. *The following statement holds with probability at least $1 - \delta$: Algorithm 3 identifies the Pareto set using at most*

$$\log_2(2/\Delta_1) + \sum_{i=2}^{h} \frac{64\sigma^2}{\Delta_i^2} \log\left(\frac{Kd}{\delta} \log_2\left(\frac{2}{\Delta_i}\right)\right)$$

samples and $\lceil \log_2(1/\Delta_1) \rceil$ rounds.

Proof. We assume $\mathcal{E}_{\text{fc}}$ holds. The correctness of Algorithm 3 is then proven in Lemma 10 and Lemma 3 upper-bounds the number of rounds before termination. It remains to bound the sample complexity of the algorithm on $\mathcal{E}_{\text{fc}}$ and compute $\mathbb{P}(\mathcal{E}_{\text{fc}})$ to conclude.

By Lemma 3 an upper-bound on $|A_r|$ for some specific rounds. Interestingly we can bound the sample complexity between consecutive "checkpoints rounds”. In what follows, we rewrite the complexity as a sum of number of pulls between these intermediate "checkpoints rounds”. Let us introduce the sequence $\{\alpha_s : s \geq 0\}$ defined as $\alpha_0 = 0$ and for any $s \geq 1$, $\alpha_s = \lceil \log_2(1/\Delta_{\lfloor h/2^s \rfloor}) \rceil$. We assume *w.l.o.g* that the sequence is increasing. Simple calculation shows that $\alpha_{\lfloor \log_2(h) \rfloor} = \lceil \log_2(1/\Delta_1) \rceil$ and

$$\{1, \dots, \lceil \log_2(1/\Delta_1) \rceil\} = \bigcup_{s=1}^{\lfloor \log_2(h) \rfloor} [\![\alpha_{s-1}, \alpha_s]\!]. \tag{29}$$

Letting

$$T_r = \frac{32(1+3\varepsilon_r)\sigma^2 h_r}{\varepsilon_r^2} \log\left(\frac{dn_r}{\delta_r}\right),$$

where $n_r = |A_r|$ and $t_r = \lceil T_r \rceil$, so $t_r \le T_r + 1$. Using (29) then leads to

$$\begin{aligned}\sum_{r=1}^{\lceil \log_2(1/\Delta_1)\rceil} T_r &= \sum_{s=0}^{\lfloor \log_2(h)\rfloor -1} \sum_{r=\alpha_s+1}^{\alpha_{s+1}} T_r \\ &=: \sum_{s=0}^{\lfloor \log_2(h)\rfloor -1} N_s\end{aligned}$$

where $N_s = \sum_{r=\alpha_s+1}^{\alpha_{s+1}} T_r$ is "the number of arms pulls” between round (α_s+1) and α_{s+1}.

Next we bound the term N_s for $s \in \{0, \dots, \lfloor \log_2(h)\rfloor - 1\}$. We recall that $h_r \le \min(h, n_r)$ as, $n_r = |A_r|$ is the number of active arms at round r and h_r is the dimension of the space spanned by the features of the active arms. Using Lemma 3 on $\mathcal{E}_{\text{fc}}$, it holds that for $r \ge \alpha_s + 1$

$$n_r \le \begin{cases} K & \text{if } s = 0 \\ \lfloor h/2^s \rfloor & \text{if } s \ge 1 \end{cases} \tag{30}$$

Therefore for $s \in \{0, \dots, \lfloor \log_2(h)\rfloor - 1\}$ and for any $r \ge \alpha_s + 1$, we simply have $\min(h, n_r) \le \lfloor h/2^s \rfloor$, so $h_r \le \lfloor h/2^s \rfloor$. It then follows that

$$\begin{aligned}N_s &= \sum_{r=\alpha_s+1}^{\alpha_{s+1}} T_r && (31)\\ &\le 64\sigma^2 \lfloor h/2^s \rfloor \log\left(\frac{Kd}{\delta_{\alpha_{s+1}}}\right) \sum_{r=\alpha_s+1}^{\alpha_{s+1}} \frac{1}{\varepsilon_r^2} && (32)\\ &= 64\sigma^2 \lfloor h/2^s \rfloor \log\left(\frac{Kd}{\delta_{\alpha_{s+1}}}\right) \sum_{r=\alpha_s+1}^{\alpha_{s+1}} 4^r && (33)\\ &\le 64\sigma^2 \lfloor h/2^s \rfloor \log\left(\frac{Kd}{\delta_{\alpha_s+1}}\right) \sum_{r=1}^{\alpha_{s+1}} 4^r && (34)\\ &= \frac{64\sigma^2 \lfloor h/2^s \rfloor}{3} \log\left(\frac{Kd}{\delta_{\alpha_{s+1}}}\right) (4^{\alpha_{s+1}} - 1) && (35)\end{aligned}$$

then further using that

$$\alpha_s \ge \begin{cases} \log_2(1/\Delta_{\lfloor h/2^s \rfloor}) & \text{if } s \ge 1 \\ 0 & \text{if } s = 0 \end{cases}$$

yields

$$4^{\alpha_{s+1}} \le \frac{1}{\Delta^2_{\lfloor h/2^{s+1} \rfloor}}$$

which combined with (35) yields

$$N_s \le \frac{64\sigma^2 \lfloor h/2^s \rfloor}{3\Delta^2_{\lfloor h/2^{s+1} \rfloor}} \log\left(\frac{Kd}{\delta_{\alpha_s+1}}\right). \tag{36}$$

We can now bound $N = \sum_s N_s$ in terms of the sub-optimality gaps:

$$\begin{aligned}N &= \sum_{s=0}^{\lfloor \log_2 h \rfloor -1} N_s && (37)\\ &\le \frac{64\sigma^2}{3} \sum_{s=0}^{\lfloor \log_2 h \rfloor -1} \frac{\lfloor h/2^s \rfloor}{\Delta^2_{\lfloor h/2^{s+1} \rfloor}} \log\left(\frac{\pi^2 K d \lceil \log_2(1/\Delta_{\lfloor h/2^{s+1} \rfloor}) \rceil^2}{6\delta}\right) && (38)\end{aligned}$$

then we note that the mapping

$$u \mapsto \frac{1}{\Delta_u^2} \log\left(\frac{\pi^2 K d \lceil \log_2(1/\Delta_u) \rceil^2}{6\delta}\right)$$

is non-increasing and it is easy to check that

$$\lfloor h/2^s \rfloor - \lceil \lfloor h/2^s \rfloor / 2 \rceil + 1 \geq \frac{1}{2} \lfloor h/2^s \rfloor$$

therefore

$$\frac{\lfloor h/2^s \rfloor}{\Delta^2_{\lfloor h/2^s \rfloor}} \log\left(\frac{\pi^2 K d \lceil \log_2(1/\Delta_{\lfloor h/2^s \rfloor}) \rceil^2}{12\delta}\right) \leq 2 \sum_{u=\lceil \lfloor h/2^s \rfloor /2 \rceil}^{\lfloor h/2^s \rfloor} \frac{1}{\Delta_u^2} \log\left(\frac{\pi^2 K(K-1) d \lceil \log_2(1/\Delta_u) \rceil^2}{6\delta}\right) \tag{39}$$

Combining (38) and (39) yields

$$N \leq \frac{128}{3}\sigma^2 \sum_{s=1}^{\lfloor \log_2 h \rfloor} \sum_{u=\lceil \lfloor h/2^s \rfloor /2 \rceil}^{\lfloor h/2^s \rfloor} \frac{1}{\Delta_u^2} \log\left(\frac{\pi^2 K d \lceil \log_2(1/\Delta_u) \rceil^2}{6\delta}\right) \tag{40}$$

Now let us introduce for any s, the set of integers $\mathcal{I}_s = [\![\lceil \lfloor h/2^s \rfloor /2 \rceil, \lfloor h/2^s \rfloor]\!]$. We have

$$\bigcup_{s=1}^{\lfloor \log_2 h \rfloor} \mathcal{I}_s \subset \{2, \dots, h\}.$$

We show that for any $p, q \in \{1, \dots, \lfloor \log_2(h) \rfloor\}$ if $|p - q| \geq 2$ then $\mathcal{I}_p \cap \mathcal{I}_q = \emptyset$. Assuming $p \leq q$ we claim that

$$\lfloor h/2^{p+2} \rfloor < \lceil \lfloor h/2^p \rfloor /2 \rceil \tag{41}$$

Assume otherwise then $\lfloor h/2^{p+2} \rfloor \geq \lceil \lfloor h/2^p \rfloor /2 \rceil \geq \lfloor h/2^p \rfloor /2$ so

$$h/2^{p+1} \geq \lfloor h/2^p \rfloor$$

which is impossible since for any $p \in \{0, \dots, \lfloor \log_2(h) \rfloor - 1\}$, $h/2^p \geq 1$. Therefore we have proved (41) and for any $q \geq p+2$ it holds that

$$\lfloor h/2^q \rfloor \leq \lfloor h/2^{p+2} \rfloor < \lceil \lfloor h/2^p \rfloor /2 \rceil$$

thus $\mathcal{I}_q \cap \mathcal{I}_p = \emptyset$ and for any $i \in \{2, \dots, h\}$, i belongs to no more than 2 of the subsets $\mathcal{I}_1, \dots \mathcal{I}_{\lfloor \log_2 h \rfloor}$ so it comes that

$$N \leq \frac{128}{3}\sigma^2 \sum_{s=1}^{\lfloor \log_2 h \rfloor} \sum_{u=\lceil \lfloor h/2^s \rfloor /2 \rceil}^{\lfloor h/2^s \rfloor} \frac{1}{\Delta_u^2} \log\left(\frac{\pi^2 K d \lceil \log_2(1/\Delta_u) \rceil^2}{6\delta}\right) \tag{42}$$

$$\leq \frac{128}{3}\sigma^2 \sum_{i=2}^{h} \frac{1}{\Delta_i^2} \log\left(\frac{\pi^2 K d \lceil \log_2(1/\Delta_i) \rceil^2}{6\delta}\right) \tag{43}$$

$$\leq \frac{128}{3}\sigma^2 \sum_{i=2}^{h} \frac{1}{\Delta_i^2} \log\left(\frac{\pi^2 K d \log_2(2/\Delta_i)^2}{6\delta}\right) \tag{44}$$

$$\leq \frac{128}{3}\sigma^2 \sum_{i=2}^{h} \frac{1}{\Delta_i^2} \log\left(\frac{Kd}{\delta} \log_2\left(\frac{2}{\Delta_i}\right)\right) \tag{45}$$

Then, from Lemma 9 it holds that with probability at least $1 - \delta$ the sample complexity N_δ of GEGE is upper-bounded as

$$\log_2(2/\Delta_1) + \sum_{i=2}^{h} \frac{64\sigma^2}{\Delta_i^2} \log\left(\frac{Kd}{\delta} \log_2\left(\frac{2}{\Delta_i}\right)\right).$$

□

Therefore, we have shown the sample complexity bound and the correctness on $\mathcal{E}_{\text{fc}}$. Thus proving that $\mathbb{P}(\mathcal{E}_{\text{fc}}) \geq 1 - \delta$ will conclude the proof.

E.4 Probability of the good event $\mathcal{E}_{\text{fc}}$.

At round r,

$$\mathbb{P}\left((\mathcal{E}_{\text{fc}}^r)^c \mid A_r\right) \leq \sum_{i \in A_r} \mathbb{P}\left(\|(\widehat{\Theta}_r - \Theta)^\intercal x_i\|_\infty > \varepsilon_r/4 | A_r\right)$$

Then, note that at round r, Algorithm 3 calls OptEstimator with precision $\varepsilon_r/2$ and budget t_r and by design we have $t_r \geq 20 h_r/\varepsilon_r^2$, so using Lemma 2, it follows

$$\begin{aligned}\mathbb{P}\left((\mathcal{E}_{\text{fc}}^r)^c \mid A_r\right) &\leq 2d \exp\left(-\frac{t_r \varepsilon_r^2}{32(1+3\varepsilon_r)\sigma^2 h_r}\right) \\ &\leq \delta_r/|A_r|\end{aligned}$$

which follows by plugging in the value of t_r. Therefore, by union bound over A_r and r it holds that $\mathbb{P}\left(\mathcal{E}_{\text{fc}}\right) \geq 1 - \sum_{r\geq} \delta_r \geq 1 - \delta$. This conludes the proof of Theorem 3.

F Concentration results

In this section we prove some concentration inequalities that are essential to the proofs of others results.

Lemma 4. *At any round r and for any arms $i, j \in A_r$ it holds that*

$$\begin{aligned}&|\mathrm{M}(i,j;r) - \mathrm{M}(i,j)| \leq \|(\widehat{\Theta}_r - \Theta)^\intercal (x_i - x_j)\|_\infty \textit{ and} \\ &|\mathrm{m}(i,j;r) - \mathrm{m}(i,j)| \leq \|(\widehat{\Theta}_r - \Theta)^\intercal (x_i - x_j)\|_\infty.\end{aligned}$$

Proof. We have

$$\begin{aligned}|\mathrm{M}(i,j;r) - \mathrm{M}(i,j)| &= \left|\max_c \left[\widehat{\mu}_{i,r}(c) - \widehat{\mu}_{j,r}(c)\right] - \max_c \left[\mu_i(c) - \mu_j(c)\right]\right|, \\ &\overset{(i)}{\leq} \max_c \left|(\widehat{\mu}_{i,r}(c) - \widehat{\mu}_{j,r}(c)) - (\mu_i(c) - \mu_j(c))\right|, \\ &= \|(\widehat{\mu}_{i,r} - \widehat{\mu}_{j,r}) - (\mu_i - \mu_j)\|_\infty, \\ &= \|(\widehat{\Theta}_r - \Theta)^\intercal (x_i - x_j)\|_\infty.\end{aligned}$$

where (i) follows from reverse triangle inequality. The second part of the lemma is a direct consequence of the relation $\mathrm{M}(i,j) = -\mathrm{m}(i,j)$ as well as $\mathrm{M}(i,j;r) = -\mathrm{m}(i,j;r)$ that holds for any pair of arms i, j. □

Lemma 5. *At any round 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 $\Delta_i^\star < \|(\widehat{\Theta}_r - \Theta)^\intercal (x_i - x_{i^\star})\|_\infty$.*

Proof. Since $i^\star$ does not empirically dominate i it holds that $\mathrm{M}(i,i^\star;r) > 0$ so $\mathrm{M}(i,i^\star;r) - \mathrm{M}(i,i^\star) > -\mathrm{M}(i,i^\star)$. Then noting that

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

yields $\mathrm{M}(i,i^\star;r) - \mathrm{M}(i,i^\star) > \Delta_i$. Therefore

$$\begin{aligned}\Delta_i = \Delta_i^\star &< \mathrm{M}(i,i^\star;r) - \mathrm{M}(i,i^\star) \\ &\leq \|(\widehat{\Theta}_r - \Theta)^\intercal (x_i - x_{i^\star})\|_\infty,\end{aligned}$$

where the last inequality is a consequence of Lemma 4. □

We recall the following lemma from the main paper.

Lemma 1. *If the noise η_t has covariance $\Sigma \in \mathbb{R}^{d\times d}$ and $a_1, \ldots, a_n$ are deterministically chosen then for any $x_i \in \{x_{a_1}, \ldots, x_{a_n}\}$, $\mathrm{Cov}(\widehat{\Theta}_n^\intercal x_i) = \|x_i\|^2_{V_n^\dagger}\Sigma$.*

We actually prove a stronger statement that is stated below.

Lemma 11. *If the noise η_t has covariance $\Sigma \in \mathbb{R}^{d\times d}$ and $a_1, \dots, a_N$ are deterministically. Assuming the set of active arms is $x_1, \dots, x_K$ then for any $x \in$ span$(\{x_1, \dots, x_K\})$,* $\mathrm{Cov}(\widehat{\Theta}_N^\intercal x) = \|x\|^2_{V_N^\dagger}\Sigma$.

Proof. In what follows we let $E := \mathrm{span}(\{x_1, \dots, x_K\})$ be the space spanned the vectors $x_1, \dots x_K$. As the columns of B forms an orthogonal basis of E, $P = B(B^\intercal B)^{-1}B^\intercal = BB^\intercal$ is a matrix that project onto E. Therefore, for any $x \in E$

$$\Theta^\intercal x = \Theta^\intercal BB^\intercal x = (B^\intercal\Theta)^\intercal B^\intercal x.$$

Thus recalling that $X_N = (x_{a_1}, \dots, x_{a_N})^\intercal$ it holds that $X_N\Theta = (X_N B)(B^\intercal\Theta)$. Rewriting the solution of the least squares leads to

$$\begin{aligned}
\widehat{\Theta}_N &= B(B^\intercal V_N B)^{-1}B^\intercal X_N^\intercal(X_N\Theta + H_N)\\
&= B(B^\intercal V_N B)^{-1}B^\intercal X_N^\intercal(X_N\Theta) + V_N^\dagger X_N^\intercal H_N\\
&= B(B^\intercal V_N B)^{-1}B^\intercal X_N^\intercal(X_N B)(B^\intercal\Theta) + V_N^\dagger X_N^\intercal H_N\\
&= B(B^\intercal V_N B)^{-1}(B^\intercal V_N B)(B^\intercal\Theta) + V_N^\dagger X_N^\intercal H_N\\
&= BB^\intercal\Theta + V_N^\dagger X_N^\intercal H_N
\end{aligned}$$

then for any $x \in E$, as $BB^\intercal x = x$ it follows that

$$\begin{aligned}
\widehat{\Theta}_N^\intercal x &= \Theta^\intercal BB^\intercal x + (V_N^\dagger X_N^\intercal H_N)^\intercal x\\
&= \Theta^\intercal x + (V_N^\dagger X_N^\intercal H_N)^\intercal x
\end{aligned}$$

thus we have for $x \in E$,

$$(\widehat{\Theta}_N - \Theta)^\intercal x = (V_N^\dagger X_N^\intercal H_N)^\intercal x. \tag{46}$$

Computing the covariance follows as

$$\begin{aligned}
\mathrm{Cov}((\widehat{\Theta}_N - \Theta)^\intercal x) &= \mathbb{E}\left[(V_N^\dagger X_N^\intercal H_N)^\intercal x x^\intercal (V_N^\dagger X_N^\intercal H_N)\right] &(47)\\
&= \mathbb{E}\left[H_N^\intercal \tilde{x}\tilde{x}^\intercal H_N\right] &(48)
\end{aligned}$$

where $\tilde{x} := X_N V_N^\dagger x$. Letting $h_i^\intercal$ denotes the i-th row of $H_N^\intercal$, for each i, j

$$\begin{aligned}
\mathbb{E}[h_i^\intercal \tilde{x}\tilde{x}^\intercal h_j] &= \tilde{x}^\intercal \mathbb{E}[h_i h_j^\intercal]x &(49)\\
&= \tilde{x}^\intercal \sigma_{i,j}\tilde{x} &(50)
\end{aligned}$$

where $\Sigma := (\sigma_{r,s})_{r,s\le d}$ and the last line follows since for any $t, t' \le N$ by independence of successive observations we have $\mathbb{E}[h_i(t)h_j(t')] = \delta^{\mathrm{kro}}_{t,t'}\sigma_{i,j}$. Combining Eq. (50) with Eq. (48) yields

$$\mathrm{Cov}((\widehat{\Theta}_N - \Theta)^\intercal x) = \Sigma\tilde{x}^\intercal\tilde{x}$$

then further noting that

$$\begin{aligned}
\tilde{x}^\intercal\tilde{x} &= x^\intercal V_N^\dagger X_N^\intercal X_N V_N^\dagger x\\
&= x^\intercal B(B^\intercal V_N B)^{-1}B^\intercal V_N B(B^\intercal V_N B)^{-1}B^\intercal x\\
&= x^\intercal V_N^\dagger x = \|x\|^2_{V_N^\dagger}
\end{aligned}$$

concludes the proof. □

The following results is proven in Appendix H.

Lemma 2. *Let $S \subset [K]$, $\kappa \in (0, 1/3]$ and $N \ge 5h_S/\kappa^2$ where $h_S = \dim(\mathrm{span}(\{x_i : i \in S\}))$. The output $\widehat{\Theta}$ of* OptEstimator*(S, N, κ) satisfies for all $\varepsilon > 0$ and $i \in S$*

$$\mathbb{P}\left(\|(\Theta - \widehat{\Theta})^\intercal x_i\|_\infty \ge \varepsilon\right) \le 2d\exp\left(-\frac{N\varepsilon^2}{2(1+6\kappa)\sigma^2 h_S}\right).$$

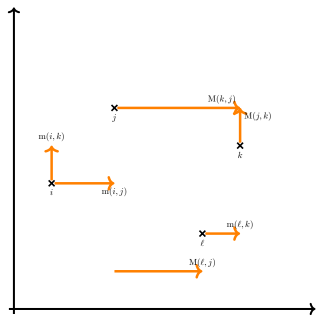

Figure 5: PSI gaps and distances

G Lower Bounds

Before proving the lower bounds, we illustrate the PSI and the quantities M, m on Fig.5

We note that, in this instance $\Delta_i = \mathrm{m}(i,j)$ and by increasing i by Δ_i on 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 $\delta_j^- = \min(\mathrm{M}(\ell, j) + \Delta_\ell, \Delta_i)$.

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 now prove minimax lower bounds in both fixed-confidence and fixed-budget settings. We recall the lower-bound below for un-structured PSI in the fixed confidence setting.

Theorem 5 (Theorem 17 of Auer et al. [2016])**.** *For any set of operating points $\mu_i \in [1/4, 3/4]^d$, $i = 1, \dots, K$, there exist distributions $\mathcal{D}_i$ such that with probability at least $1-\delta$, any δ-correct algorithm for PSI requires at least*

$$\Omega\left(\sum_{i=1}^K \frac{1}{\widetilde{\Delta}_i^2} \log(\delta^{-1})\right)$$

samples to identify the Pareto set. Where for any sub-optimal arm $\widetilde{\Delta}_i = \Delta_i$ and for an optimal arm $\widetilde{\Delta}_i = \delta_i^+$.

In particular, there exist instances where $\Delta_i = \delta_i^+$ for any Pareto-optimal arm i. Thus, this result shows that H_1 is a good proxy to measure the complexity of PSI in the fixed-confidence setting. The proof of this result is based on the celebrated change of distribution technique (see e.g Kaufmann et al. [2016]) which given the instance $\nu := (\nu_1, \dots, \nu_K)$ shifts the mean of ν_i for an arm i while keeping the others fixed constant. However in linear PSI the arms' means are correlated through Θ. So that in general Theorem 5 does not directly apply to linear PSI. We recall below our lower-bound for linear PSI in the fixed-confidence setting.

Theorem 4. *For any $K, d, h \in \mathbb{N}$, there exists a set $\mathcal{B}(K,d,h)$ of linear PSI instances s.t for $\nu \in \mathcal{B}(K,d,h)$ and for any δ-correct algorithm for PSI, with probability at least $1-\delta$,*

$$\tau_\delta^{\mathcal{A}} = \Omega\left(H_{1,\mathit{lin}}(\nu) \log(\delta^{-1})\right).$$

Proof of Theorem 4. The idea of the proof is to transform an unstructured bandit instance into a linear PSI instance. Let ν be a bandit instance with $K \geq 2$ arms and dimension $d \geq 1$ and with means $\mu_1, \dots, \mu_K \in [0,1]^d$. Let $e_1, \dots e_h$ denote the canonical basis of $\mathbb{R}^h$. We define a linear PSI instance ν_{lin} with features

$$x_i = \begin{cases} e_i & \text{if } i \leq h \\ \mathbf{0} & \text{else.} \end{cases}$$

We assume that the learner knows that $\mu_i \in [0,1]^d$ for any arm i. We claim that with this information an "efficient" algorithm for PSI should not pull arms from $\{h+1,\dots,K\}$. To see this, first note that these arms will be sub-optimal so $\mathcal{S}^\star \subset [h]$. Moreover, even if an arm $i \in \{h+1,\dots,K\}$ dominates another arm $j \in \{1,\dots,h\}$, as j is not Pareto-optimal there exits another arm $j^\star \in \mathcal{S}^\star \subset \{1,\dots,h\}$ which dominates j with a larger margin, so is "cheaper" to pull. Therefore the complexity of ν_{lin} reduces to the complexity of a linear bandit $\tilde{\nu}_{\text{lin}}$ with only h arms. As the features in $x_1,\dots,x_h$ forms the canonical $\mathbb{R}^h$ basis, $\tilde{\nu}_{\text{lin}}$ reduces to an un-structured bandit instance with (un-correlated) means $\tilde{\mu}_i = \Theta^\intercal x_i$, $i = 1,\dots,h$. Therefore, by choosing $\mu_1,\dots,\mu_h \in [1/4, 3/4]^d$, we can apply Theorem 5 to $\tilde{\nu}_{\text{lin}}$. □

Actually in the result stated above we have proved that this bound holds for a class of instances $\mathcal{B}(K,d,h)$ of and not just a single fixed instance .

For the fixed-budget setting Kone et al. [2024] proved a lower-bound for a class of instances. We recall their result below after introducing some notation. Their lower-bound applies to class of instances $\mathcal{B}$ defined as follows. $\mathcal{B}$ contains the instances such that each sub-optimal arm i is only dominated by a Pareto-optimal arm denoted by $i^\star$ and that for each optimal arm j there exists a unique sub-optimal arm which is dominated by j, denoted by $\underline{j}$. Moreover for any instance in $\mathcal{B}$ the authors require its Pareto-optimal arms not to be close to the sub-optimal arms they don't dominate: for any sub-optimal arm i and Pareto-optimal arm j such that $\mu_i \nprec \mu_j$,

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

Let $\nu := (\nu_1,\dots,\nu_K)$ be an unstructured instance whose means belongs to $\mathcal{B}$ and with isotropic multi-variate normal arms $\nu_i \sim \mathcal{N}(\mu_i, \sigma^2 I)$. For every $i \in [K]$, define the alternative instance $\nu^{(i)} := (\nu_1, \dots \nu_i^{(i)}, \dots, \nu_K)$ in which *only* the mean of arm i is shifted:

$$\mu_i^{(i)} := \begin{cases} \mu_i - 2\Delta_i \tilde{e}_{d_{\underline{i}}} & \text{if } i \in \mathcal{S}^\star(\nu), \\ \mu_i + 2\Delta_i \tilde{e}_{d_i} & \text{else,} \end{cases} \tag{51}$$

where $\tilde{e}_1,\dots,\tilde{e}_d$ denotes the canonical basis of $\mathbb{R}^d$ and for any arm i, $d_i := \operatorname{argmin}_{c\in[d]}[\mu_{i^\star}(c) - \mu_i(c)]$. Defining $\nu^{(0)} := \nu$, the theorem below holds.

Theorem 6 (Theorem 5 of Kone et al. [2024])**.** *Let $\nu = (\nu_1,\dots,\nu_K)$ be an instance in $\mathcal{B}$ with means $\mu := (\mu_1 \dots \mu_K)^\intercal$ and $\nu_i \sim \mathcal{N}(\mu_i, \sigma^2 I)$. For any algorithm $\mathcal{A}$, there exists $i \in \{0,\dots,K\}$ such that $H(\nu^{(i)}) \leq H(\nu)$ and the probability of error $\mathcal{A}$ on $\nu^{(i)}$ is at least*

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

As explained above for the fixed-confidence setting. The proof of this lower bound also uses the change of distribution lemma. In the instances $\nu^{(i)}$ introduced above, it is crucial that only the mean of arm i has changed w.r.t $\nu^{(0)}$. Therefore, Theorem 6 does not apply to general instances in linear PSI. We recall our lower-bound for linear PSI in the fixed-budget.

Theorem 2. *Let $\mathbb{W}_H$ be the set of instances with complexity $H_{2,lin}$ at most H. For any budget T, letting $\widehat{S}_T^{\mathcal{A}}$ be the output of algorithm $\mathcal{A}$, it holds that*

$$\min_{\mathcal{A}} \max_{\nu\in\mathbb{W}_H} \mathbb{P}_\nu(\widehat{S}_T^{\mathcal{A}} \neq \mathcal{S}^\star(\nu)) \geq \frac{1}{4}\exp\left(-\frac{2T}{H\sigma^2}\right).$$

Proof of Theorem 2. Let H be fixed and recall that $\mathbb{W}_H : \{\nu_{\text{lin}} : H_{2,\text{lin}}(\nu) \leq H\}$ is the set of linear PSI instances with complexity less than H. The proof of Theorem 2 follows similar lines to Theorem 4. Let ν be an un-structured bandit instance with $K \geq 2$ arms, dimension $d \geq 1$, with means $\mu_1,\dots,\mu_K \in [0,1]^d$ and such that $H_2(\nu) \leq H$. We construct a linear PSI instance ν_{lin} from an unstructured multi-dimensional instance ν by setting $x_i := e_i$ for any $i \leq h$ and for $i > h$, $x_i = \mathbf{0}$ where $e_1,\dots,e_h$ is the canonical $\mathbb{R}^h$-basis. We also assume that the agent knows that $\mu_i \in [0,1]^d$ for any arm i. For ν_{lin} the arms $\{h+1,\dots,K\}$ are necessarily sub-optimal so $\mathcal{S}^\star \subset [h]$ thus to identify the Pareto set and efficient algorithm should reduce to pull arms in $\{1,\dots,h\}$. Indeed, as explained in the proof of Theorem 4 even if an arm $i \in \{h+1,\dots,K\}$ dominates another arm $j \in \{1,\dots,h\}$, as j is not Pareto-optimal there exits another arm $j^\star \in \mathcal{S}^\star \subset \{1,\dots,h\}$ which is "cheaper" to pull as it dominates j with a larger margin. ν_{lin} reduces to a linear bandit $\tilde\nu_{\text{lin}}$ with only h arms and since the features $x_1,\dots,x_h$ forms the canonical basis of $\mathbb{R}^h$, $\tilde\nu_{\text{lin}}$ is an un-structured bandit instance with (un-correlated) means $\tilde\mu_i = \Theta^\intercal x_i$, $i=1,\dots,h$. Therefore, by choosing $\tilde\nu := (\nu_1,\dots,\nu_h)$ that belongs to $\mathcal{B}$, we can apply Theorem 6 which yields

$$\max_{i\in\{0,\dots,K\}} \mathbb{P}_{\tilde\nu^{(i)}}(S_T^{\mathcal{A}} \neq \mathcal{S}^\star(\tilde\nu^{(i)})) \geq \frac{1}{4}\exp\left(-\frac{2T}{H\sigma^2}\right)$$

where by construction $\tilde\nu^{(i)}$ (see construction above) is also a linear PSI instance. Then further noting that $H \geq H_2(\nu) \geq H_2(\tilde\nu)$ and by Theorem 6 for any $i \leq h$ $H_{2,\text{lin}}(\tilde\nu) \geq H_2(\tilde\nu^{(i)})$. Then recalling that ν_{lin} is equivalent to $\tilde\nu$ it comes

$$\min_{\mathcal{A}}\max_{\nu\in\mathbb{W}_H} \mathbb{P}_\nu(S_T^{\mathcal{A}} \neq \mathcal{S}^\star(\nu)) \geq \frac{1}{4}\exp\left(-\frac{2T}{H\sigma^2}\right),$$

which is the claimed result. □

H Computing and rounding a G-optimal design

In this section, we discuss the results related to the G-design and the rounding. In what follows let $S \subset [K]$ be a set of arms. To ease notation we re-index the arms of S by assuming $S := \{1,\dots,|S|\}$. Let N be the allocation budget (the total number of pulls of arms in S) and $\kappa \in (0,1/3]$ the parameter of the rounding algorithm to be fixed. $h_S = \dim(\text{span}(\{x_i : i \in S\}))$ is the dimension of the space spanned by the covariates of S. $\mathcal{X}_S := (x_i, i\in S)^\intercal$ and $B_S := (u_1,\dots,u_m)$ is the matrix formed with the first $m = h_S = \text{rank}(S)$ columns of U, the matrix of left singular vectors of $\mathcal{X}_S^\intercal$ obtained by singular value decomposition. We recall that for N pulls of arms in $[S]$, letting $T_i(N)$ be number of samples collected from arm i,

$$V_N^\dagger := B_S(B_S^\intercal V_N B_S)^{-1}B_S^\intercal \quad \text{and} \quad V_N := \sum_{i=1}^K T_i(N)x_i x_i^\intercal. \tag{52}$$

As from Lemma 1 the statistical uncertainty from estimating the mean of arm i scales with $\|x_i\|_{V_N^\dagger}$, a call to $\text{OptEstimator}(S, N, \kappa)$ is meant to estimate the hidden parameter Θ by collecting N samples from arms in S according to an approximation of the solution of the following problem (ordinal G-optimal design):

$$\begin{aligned} &\operatorname*{argmin}_{s\in[0,\dots,N]^{|S|}} \max_{i\in S} \quad \|x_i\|_{(V^s)^\dagger} \\ &\qquad\qquad \text{s.t.} \quad \sum_{i\in S} s(i) = N\,. \end{aligned} \tag{53}$$

Finding such an optimal design with integer values is a NP-hard problem [Allen-Zhu et al., 2017]. Instead, its continuous relaxation (obtained by normalizing by N), amounts to finding an allocation ω that minimizes

$$\max_{i\in S}(B_S^\intercal x_i)^\intercal \left(\sum_{i\in S}\omega(i)B_S^\intercal x_i x_i^\intercal B_S\right)^{-1} B_S^\intercal x_i, \tag{54}$$

which reduces to compute a G-optimal allocation on the covariates $B_S^\intercal x_i, i\in S$:

$$w_S^\star \in \operatorname*{argmin}_{\omega\in\mathbf{\Delta}_{|S|}} \max_{i\in S} \|\widetilde{x}_i\|^2_{(\widetilde{V}^\omega)^{-1}}, \text{ and } \widetilde{V}^\omega := \sum_{i\in S}\omega(i)\widetilde{x}_i\widetilde{x}_i^\intercal. \tag{55}$$

This is a convex optimization problem on the probability simplex of $\mathbb{R}^{|S|}$. Efficient solvers have been used in the literature for linear BAI and experiment design optimization see (e.g Fiez et al. [2019], Soare et al. [2014]). In this work, we follow Allen-Zhu et al. [2017] and we recommend an entropic mirror descent algorithm to solve Eq. (55), which is recalled as Algorithm 4 for the sake of completeness.

Then, computing an integer allocation whose value is close to the optimal value of Eq. (55) can be done in different ways. Tao et al. [2018] and Camilleri et al. [2021] use a stochastic rounding: they sample N arms from S following the distribution $\omega_S^\star$ and use a novel estimator different from the least-squares estimate. Yang and Tan [2022], Azizi et al. [2022] use floors and ceilings of $N\omega_S^\star$. Although practical, it is known that the value of such rounded allocations can deviate a lot from the optimal value of Eq. (53) [Tao et al., 2018].

Algorithm 4: Entropic mirror descent algorithm for computing $w_S^\star$ Tao et al. [2018]

Input: A set of arms S and covariates $(\widetilde{x}_i, i \in S)$, tolerance ε and Lipschitz constant L_f
Initialize: $t \leftarrow 1$ and $w^{(1)} \leftarrow (1/|S|, \dots, 1/|S|)$
while $|\max_{i\in S} \widetilde{x}_i^\intercal (\widetilde{V}^{w^{(t)}})^{-1}\widetilde{x}_i - h_S| \geq \varepsilon$ **do**
  set $\eta_t \leftarrow \frac{\sqrt{2\ln N}}{L_f}\frac{1}{\sqrt{t}}$
  Compute gradient $g_i^{(t)} \leftarrow \mathrm{Tr}\left(\widetilde{V}\left(w^{(t)}\right)^{-1}\left(\widetilde{x}_i\widetilde{x}_i^T\right)\right)$
  Update $w_i^{(t+1)} \leftarrow \frac{w_i^{(t)}\exp\left(\eta_t g_i^{(t)}\right)}{\sum_{i=1}^N w_i^{(t)}\exp\left(\eta_t g_i^{(t)}\right)}$
  $t \leftarrow t+1$
return: $w^{(t)}$

Allen-Zhu et al. [2017] proposed an efficient rounding procedure that guarantees that the value of the returned integer allocation is within a small factor of the optimal value of Eq. (55). Before recalling their result we introduce the notation $F_S(s) := \max_{i\in S}\|x_i\|^2_{(V^s)^\dagger}$.

We recall the celebrated Kiefer–Wolfowitz equivalence theorem below.

Theorem 7 (Restatement of Kiefer and Wolfowitz [1960])**.** *Let covariates $\{x_i : i \in S\} \subset \mathbb{R}^h$ and for any $\omega \in \mathbf{\Delta}_{|S|}$ define $V^\omega = \sum_{i\in S}\omega(i)x_ix_i^\intercal$ and when V^ω is non-singular $f(x;\omega) := x^\intercal(V^\omega)^{-1}x$. The following two extremum problems:*

a) ω maximizing det(V^ω)

b) ω minimizing $\max_{i\in S} f(x_i;\omega)$

are equivalent and a sufficient condition to satisfy Eq. (b) is $\max_{i\in S} f(x_i,\omega) = h$, which is satisfied when the covariates $\{x_i : i \in S\}$ span $\mathbb{R}^h$.

Theorem 8 (reformulated; rounding of Allen-Zhu et al. [2017])**.** *Suppose $\kappa \in (0, 1/3]$ and $N \geq 5h_S/\kappa^2$. Let $\omega_S^\star = \mathrm{argmin}_{\omega\in\mathbf{\Delta}_S} F_S(\omega)$. Then, there exists an algorithm that outputs an integer allocation $s^\star$ satisfying*

$$s^\star \in \boldsymbol{\mathcal{D}}_{S,N} \quad \text{and} \quad F_S(s^\star) \leq (1+6\kappa)\frac{F_S(\omega_S^\star)}{N}$$

where $\boldsymbol{\mathcal{D}}_{S,N} := \{s \in \{0,\dots,N\}^{|S|} : \sum_{i\in S} s(i) = N\}$. This algorithm runs in time complexity $\widetilde{O}\left(N|S|\widetilde{h}^2\right)$.

We refer to a call to this algorithm as $\texttt{ROUND}(N, \{\widetilde{x}_i, i \in S\}, \omega_S^\star, \kappa)$. It returns an integer allocation $s^\star = (s^*(1), \dots, s^*(|S|))$ from which we can immediately deduce a list of arms to pull (the first arm in S replicated $s^*(1)$ times, the second replicated $s^*(2)$ times, etc.).

Simple arguments from linear algebra show that the h_S columns of B_S form a basis of span($\{x_i : i \in S\}$), hence $\{B_S^\intercal x_i : i \in S\}$ spans $\mathbb{R}^{h_S}$. Using Theorem 7 applied to the covariates $\{B_S^\intercal x_i : i \in S\}$ yields

$$F_S(\omega_S^\star) = h_S$$

and thus the integer allocation $s^\star$ output by $\texttt{ROUND}(N, \{\widetilde{x}_i, i \in S\}, \omega_S^\star, \kappa)$ satisfies for $N \geq 5h_S/\kappa^2$,

$$F(s^\star) \leq (1+6\kappa)\frac{h_S}{N},$$

which is stated below.

Lemma 12. *Let $S \subset [K]$, $\kappa \in (0, 1/3]$ and $N \geq 5h_S/\kappa^2$ where h_S = dim(span($\{x_i : i \in S\}$)). The allocation $\{T_i(N) : i \in S\}$ computed by* OptEstimator*(S, N, κ) to estimate Θ satisfies*

$$\max_{i\in S}\|x_i\|^2_{V_N^\dagger} \leq (1+6\kappa)\frac{h_S}{N}.$$

Building on this result, we derive the following concentration result.

Lemma 2. *Let $S \subset [K]$, $\kappa \in (0, 1/3]$ and $N \geq 5h_S/\kappa^2$ where $h_S = \dim(\mathrm{span}(\{x_i : i \in S\}))$. The output $\widehat{\Theta}$ of* $\mathrm{OptEstimator}$*(S, N, κ) satisfies for all $\varepsilon > 0$ and $i \in S$*

$$\mathbb{P}\left(\|(\Theta - \widehat{\Theta})^\intercal x_i\|_\infty \geq \varepsilon\right) \leq 2d \exp\left(-\frac{N\varepsilon^2}{2(1+6\kappa)\sigma^2 h_S}\right).$$

Proof of Lemma 2. We recall that by assumption the vector noise has σ-sub-gaussian marginals. From the proof of Lemma 11 it is easy to see that for any $i \in S$, the marginals of $(\Theta - \widehat{\Theta})x_i$ are $\sigma\|X_N^\intercal V_N^\dagger x_i\|_2$-sub-gaussian. Then direct calculations shows that

$$\begin{aligned}
\|X_N^\intercal V_N^\dagger x_i\|_2^2 &= x_i^\intercal V_N^\dagger V_N V_N^\dagger x_i \\
&= x_i^\intercal \left(B_S(B_S^\intercal V_N B_S)^{-1} B_S^\intercal\right) V_N \left(B_S(B_S^\intercal V_N B_S)^{-1} B_S^\intercal\right) x_i \\
&= x_i^\intercal B_S(B_S^\intercal V_N B_S)^{-1} B_S^\intercal x_i \\
&= x_i^\intercal V_N^\dagger x_i = \|x_i\|^2_{V_N^\dagger}.
\end{aligned}$$

Therefore, by concentration of sub-gaussian variables (see e.g Lattimore and Szepesvári [2020]) we have for i fixed,

$$\begin{aligned}
\mathbb{P}(\|(\Theta - \widehat{\Theta})^\intercal x_i\|_\infty \geq \varepsilon) &\leq 2d \exp\left(-\frac{\varepsilon^2}{2\sigma^2\|x_i\|^2_{V_N^\dagger}}\right) \\
&\leq 2d \exp\left(-\frac{\varepsilon^2}{2\sigma^2 \max_{k\in S}\|x_k\|^2_{V_N^\dagger}}\right)
\end{aligned}$$

then the G-optimal design and the rounding (Lemma 12) ensure that

$$\max_{k \in S} \|x_k\|^2_{V_N^\dagger} \leq (1 + 6\kappa)h_S/N.$$

Therefore

$$\mathbb{P}\left(\|(\Theta - \widehat{\Theta})^\intercal x_i\|_\infty \geq \varepsilon\right) \leq 2d \exp\left(-\frac{N\varepsilon^2}{2(1+6\kappa)\sigma^2 h_S}\right).$$

□

I Implementation details and additional experiments

In this section we detail our experimental setup and provide additional experimental results.

I.1 Complexity and setup

Time and memory complexity The main computational cost of GEGE (excepting calls to $\mathrm{OptEstimator}$) is the computation of the empirical gaps. Which requires to compute $\mathrm{M}(i, j; r)$ for any tuple (i, j) of active arms and to temporarily store them. Computing the gaps results in a total $\mathcal{O}(K^2 d)$ time complexity and $\mathcal{O}(K^2)$ memory complexity. Note that for the memory allocation we can maintain the same arrays for the whole execution of the algorithm thus only cheap memory allocations are made after initialization. The overall computational complexity is reasonable as GEGE is an elimination algorithm the computational cost reduces after rounds and we have proven that no more than $\lceil \log_2(1/\Delta_1)\rceil$ rounds are required in the fixed-confidence regime and only $\lceil \log_2(h)\rceil$ rounds in the fixed-budget setting. For this reason the computational complexity of a call to $\mathrm{OptEstimator}$ has a limited impact in practice. We report below the average runtime on a personal computer with an ARM CPU 8GB RAM and 256GB SSD storage. The values are averaged over 50 runs.

Table 2: Runtime of GEGE recorded different instances.

$[K, h, d]$	GEGE[$\delta = 0.1$]	GEGE[$T = 500$]
$[10, 2, 2]$	6ms	217ms
$[50, 8, 2]$	7ms	464ms
$[100, 8, 4]$	545ms	791ms
$[200, 8, 8]$	768ms	1139ms
$[500, 8, 8]$	1013ms	2425ms

Setup We have implemented the algorithms mainly in `python3` and `C++`. For each experiment, the value reported (sample complexity or probability of error) are averaged over 500 runs. For the experiments on synthetic instances we generate and instance satisfying the conditions reported in the main by first choosing the h vectors by hand (and thus Θ) then the remaining arms are generated by sampling and normalizing some features from $\mathcal{U}([0, 1]^h)$ to satisfy the contraints. For the real-world datasets we normalize the features and (when mentioned) we use a least square to estimate a regression parameter $\widehat{\Theta}$ or we use the dataset as such (mis-specified setting). PAL is run with same confidence bonus used in Zuluaga et al. [2016] (which are tuned empirically) and for APE we follow Kone et al. [2023] and we use their confidence bonuses on pair of arms, which was already suggested by Auer et al. [2016].

I.2 Additional experiments

We provide additional experiments on synthetic and real-world datasets. GEGE is evaluated both in the fixed-confidence and fixed-budget regimes.

Multi-objective optimization of energy efficiency We use the energy efficiency dataset of Tsanas and Xifara [2012]. This dataset is made for buildings energy performance optimization. The efficiency of each building is characterized by $d = 2$ quantities: the cooling load and the heating load. The heating load is the amount of energy that should be brought to maintain a building in an acceptable temperature and the cooling load is the amount of energy that should be extracted from a building to sustain a temperature in an acceptable range. Ideally both heating and cooling loads should be low for energy efficiency and they are characterized by different factors like glazing area and the orientation of the building, amongst other parameters. Tsanas and Xifara [2012] reported the simulated heating and cooling loads of $K = 768$ buildings together with $h = 8$ features characterizing each building including surface, roof and wall areas, the relative compactness, overall hight etc. The dataset was primarily made for multivariate regression but we use it for linear PSI as the goal is to optimize simultaneously heating and cooling loads which in general (and in this case), results into a Pareto front of 3 arms. We evaluate Algorithm 2 with a budget $T = 10000$ and in the fixed-confidence we set $\delta = 0.1$ for Algorithm 3. We report the results average over 500 runs on Fig.6 and Fig.7. In the fixed-confidence experiment, "Racing" is the algorithm of Auer et al. [2016] for unstructured PSI.

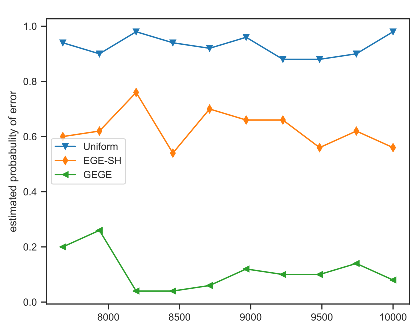

Figure 6: Average probability of error on the energy efficiency dataset.

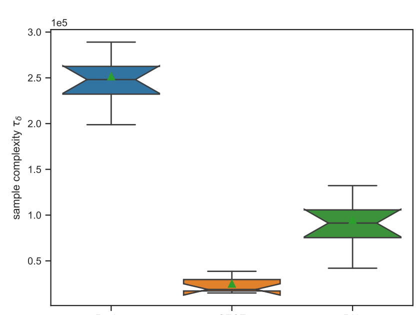

Figure 7: Sample complexity distribution on the energy efficiency dataset.

We observe that in both fixed-confidence and fixed-budget, GEGE largely outperforms its competitors. It worth noting in the fixed-budget setting, as $K = 768$, Uniform Allocation requires $T \geq 768$ to be run correctly while EGE-SH requires $T \geq 7360$. On the contrary GEGE just requires $T \geq h = 8$ which is negligible w.r.t $K = 768$. Moreover we observed that its probability of error is reasonable even for a budget $T < K$.