Minimax Optimal Fixed-Budget Best Arm Identification (BAI) in Linear Bandits.

Vincent Tan (Dept of Maths,
Dept of ECE, NUS)

Joint work with Junwen Yang (IORA, NUS)

To be presented at NeurIPS 2022.

<u>Setup</u>: Arm set $[K]=\{1,\dots,K\}$ finite

Each $i\in[K]$ is assoc. to a vector $a(i)\in\mathbb{R}^d$

Arm vectors $\underbrace{\{a(1),a(2),\dots,a(K)\}}_{}$ $\subset\mathbb{R}^d$ known

At each $t\in\mathbb{N}$, agent chooses $A_t\in[K]$, observes

$$X_t=\langle\theta^*,a(A_t)\rangle+\eta_t \qquad \text{(Linear bandit)}$$

$\theta^*\in\mathbb{R}^d$: unknown

zero-mean 1-subG

$\mathcal{E}$: set of instances defined above $\{a(1),\dots,a(K),\theta^*\}$

<u>Fixed-Budget</u> # arm selections $T\in\mathbb{N}$ fixed

Online alg $\pi=(\pi_t)_{t=1}^T$ depends only on previous arm pulls rewards.

$\pi_t(a_t\,|\,A_1,X_1,A_2,X_2,\dots,A_{t-1},X_{t-1})$

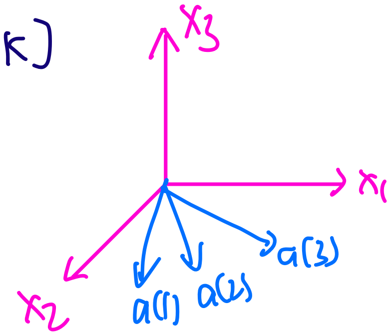
Goal: Design π that achieve

$$\min_{\pi} \Pr\Big(\underset{\in [K]}{\hat{i}_{\text{out}}} \neq \underbrace{\underset{j\in[K]}{\arg\max} \langle \theta^*, a(j)\rangle}_{\text{unique best arm}}\Big)$$

Observation: If $a(1), \dots, a(K)$ do not span $\mathbb{R}^d$, $\exists B \in \mathbb{R}^{d\times d'}$ s.t.

$$a'(j) = B^T a(j) \quad , \quad j \in [K]$$

$\{a'(j)\}_{j\in[K]}$ span $\mathbb{R}^{d'}$



B: calculated via a Gram-Schmidt.

OLS: $A_1, \dots, A_n \in [K]$ observe $X_1, \dots, X_n \in \mathbb{R}$

$a(A_1), \dots, a(A_n)$ span $\mathbb{R}^d$

$$\hat{\theta} = V^{-1} \sum_{t=1}^{n} a(A_t) X_t$$

$V = \sum_{t=1}^{n} a(A_t) a(A_t)^T$ invertible

Thm: If $A_1, \dots, A_n$ are chosen deterministically, then $\forall b \in \mathbb{R}^d$, $\delta > 0$,

$$\Pr\left(\langle \hat{\theta} - \theta^*, b \rangle \geq \sqrt{2\|b\|^2_{V^{-1}} \log\left(\tfrac{1}{\delta}\right)} \right) \leq \delta$$

$\|b\|^2_{V^{-1}} = b^T V^{-1} b$

Inspired by this

Aim: Find a prob. distn $\pi : \{a(i) : i \in [K]\} \to [0,1]$ that min

$$\min_{\pi} g(\pi) = \max_{i\in[K]} \|a(i)\|^2_{V(\pi)^{-1}}$$

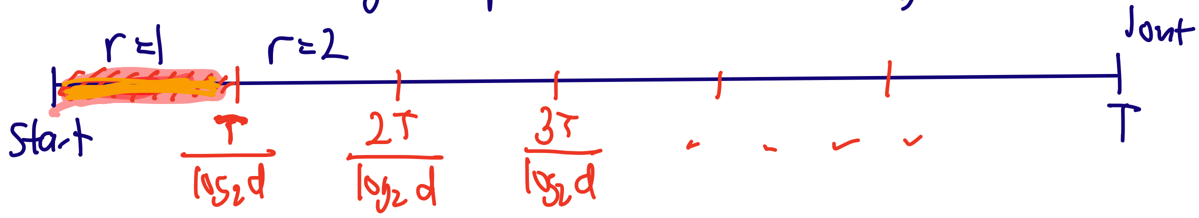
$$V(\pi) = \sum_{i=1}^{K} \pi(a(i))\, a(i)\, a(i)^T$$

G-optimal design

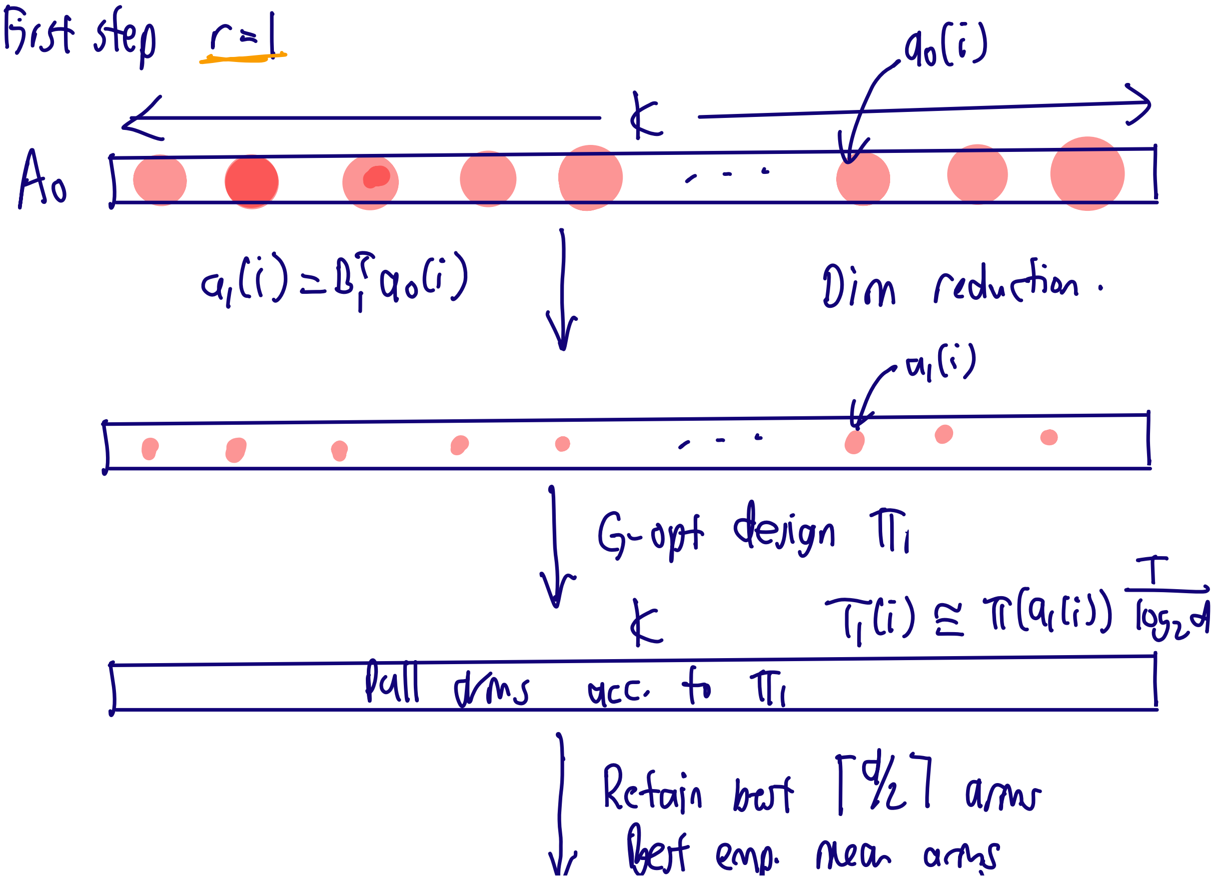
$g(\pi)$: related to confidence bds in the linear BAI prob.

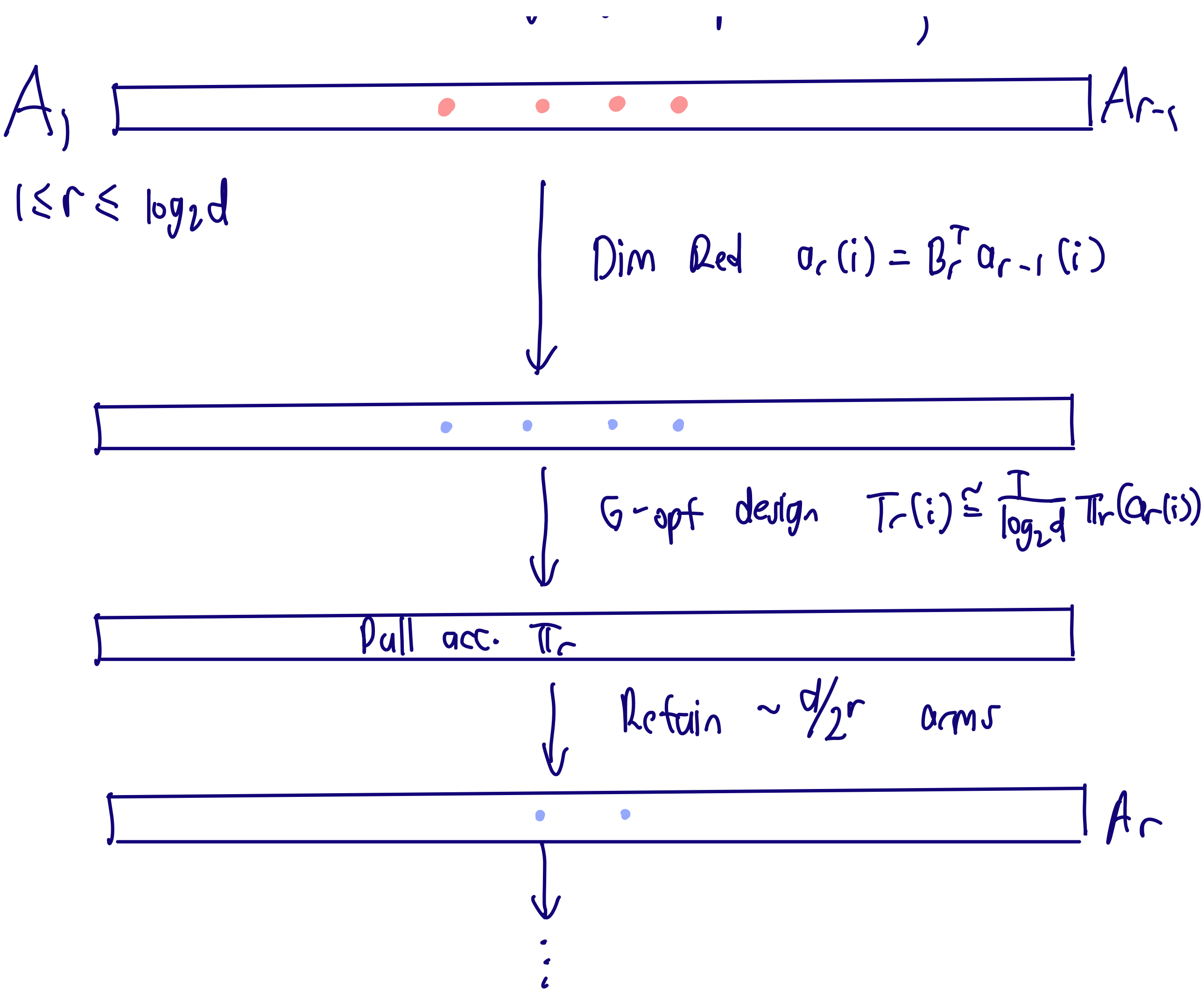
Algorithm: Optimal Design-Based Linear BAI (OD-LinBAI)

$T \in \mathbb{N}$: fixed $\quad \log_2 d$ phases. $\quad (d \ll K)$



First step $r=1$





Theoretical Guarantee for OD-LinBAI.

$p(i) = \langle \theta^*, a(i) \rangle \qquad p(1) > p(2) \ge \cdots \ge p(K)$

Define: $\Delta_i = p(1) - p(i)$, $i \in [K] \setminus \{1\}$

$\Delta_1 = \Delta_2. \qquad \Delta_i \ge 0$

Hardness Quantities $H_{1,\text{lin}} \triangleq \sum_{i=1}^{d} \Delta_i^{-2}$, $H_{2,\text{lin}} = \max_{i=1,\dots,d} \frac{i}{\Delta_i^2}$.

$H_{2,\text{lin}} \le H_{1,\text{lin}}$

Thm: $\Pr\left(i_{\text{out}}^{\text{OD-LinBAI}} \ne 1\right) \lesssim \exp\left(- \frac{T}{32 \cdot H_{2,\text{lin}} \cdot \log_2 d}\right)$

Comparison with SOTA

BayesGap (Hoffman 2014) $\Rightarrow$ not param-free

Peace (Katz-Samuel et al. NeurIPS 2020)

i) not minimax opt.

ii) not param-free

$$\sim \exp\left(-\Omega\left(\frac{T}{H_2 \log^2 d}\right)\right)$$

Linear Exploration (Alieva et al. ICML 2021)

$$\sim \exp\left(-\Omega\left(\frac{T}{\tilde{H}_2 \log_2 K}\right)\right)$$

GSE (Azizi et al. IJCAI 2022)

$$\sim \exp\left(-\Omega\left(-\frac{T\Delta_i^2}{d \log_2 K}\right)\right)$$

Lower Bd

$\forall v \in \mathcal{E}$ $\quad H_{1,\text{lin}} = H_{1,\text{lin}}(v)$

$\mathcal{E}(c) = \{v \in \mathcal{E} : H_{1,\text{lin}}(v) \leq c\}$ Hardness bounded instances.

Thm: For all T, c suff. large,

$$\inf_{\pi} \sup_{v \in \mathcal{E}(c)} \Pr\left(i_{\text{out}}^{\pi} \neq \underset{\text{best arm}}{1}\right) \exp\left(\frac{2700\,T}{H_{1,\text{lin}}(c) \log_2 d}\right) \geq \frac{1}{6}$$

Rmk: Compare upper & lower bds.

OD-LinBAI: $\forall v \in \mathcal{E}$ all instances

$$\Pr\left(i_{\text{out}}^{\text{OD-LinBAI}} \neq 1\right) \leq \exp\left(-\Omega\left(\frac{T}{H_{2,\text{lin}} \cdot \log d}\right)\right)$$

Lower bd:

← specific instance

$$\Pr(i_{out}^{\pi} \neq 1) \geq \exp\left(-O\left(\frac{T}{H_{1,lin} \cdot \log d}\right)\right)$$

$H_{2,lin} \leq H_{1,lin}$ $\Rightarrow$ OD-LinBAI is minimax optimal.

$\forall v \in \mathcal{E}$,

$$\Pr(i_{out}^{OD\text{-}LinBAI} \neq 1) \leq \exp\left(-\Omega\left(\frac{T}{H_{2,lin}(v) \cdot \log d}\right)\right)$$

$\forall \pi$ s.t. $\exists v \in \mathcal{E}(c)$,

$$\Pr(i_{out}^{\pi} \neq 1) \geq \exp\left(-O\left(\frac{T}{H_{1,lin}(v) \cdot \log_2 d}\right)\right)$$

For $\pi =$ OD-LinBAI, $\exists v \in \mathcal{E}(c, \pi)$ s.t.

$$\Pr(i_{out}^{\pi} \neq 1) \geq \exp\left(-O\left(\frac{T}{H_{1,lin}(v) \cdot \log d}\right)\right)$$

This instance v must be s.t. $H_{1,lin}(v)$ is of the same order as $H_{2,lin}(v)$.

$$H_{1,lin}(v) = \Theta(H_{2,lin}(v)).$$

$$\forall \pi \quad \sup_{v \in \mathcal{E}(c)} \Pr(i_{out}^{\pi} \neq 1) \exp\left(\frac{2700T}{H_{1,lin}(v) \cdot \log_2 d}\right) \geq \frac{1}{6}$$

$\forall \pi$ $\exists v \in \mathcal{E}(c)$ s.t.

$$\Pr(i_{out}^{\pi} \neq 1) \exp\left(\frac{2700T}{H_{1,lin}(v) \cdot \log_2 d}\right) \geq \frac{1}{6}$$