Max-Quantile Grouped Infinite-Arm Bandits

Ivan Lau (Rice)
Yan Hao Ling (NUS)
Mayank Shrivastava (UIUC)
Jonathan Scarlett (NUS)

ALT 2023

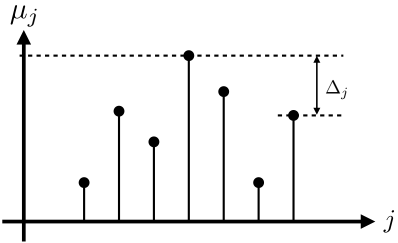
Best-Arm Identification

- **Typical best-arm identification problem:** (e.g., [Even-Dar *et al.*, 2002])
  - n arms with means $\mu_1, \ldots, \mu_n$
  - Sub-Gaussian rewards (e.g., Gaussian, Bernoulli)
  - After T rounds (possibly random), guess the best arm $\hat{j} \in \{1, \ldots, n\}$
  - Typical result: $\Pr[\text{error}] \le \delta$ with

$$\mathbb{E}[T] = \text{const.} \times \sum_j \frac{1}{\Delta_j^2} \log \frac{n}{\delta}$$

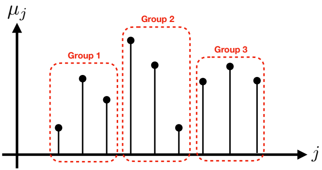
where $\Delta_j = \mu^* - \mu_j$

  - Replace Δ_j^2 by $\max\{\Delta^2, \Delta_j^2\}$ if only Δ-optimality is required



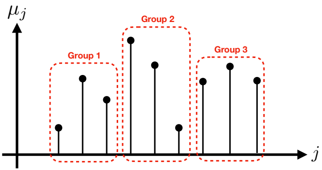
Variants of Best-Arm Identification

- **Grouped bandit problems:** (e.g., [Gabillon *et al.*, NeurIPS 2011])
  - Arms arranged into groups (possibly overlapping)
  - Typical goals: Best arm in each group / group whose worst arm is highest / ...

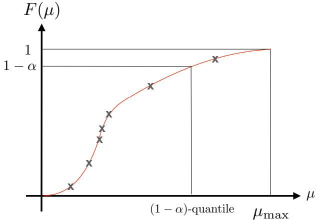


Variants of Best-Arm Identification

- **Grouped bandit problems:** (e.g., [Gabillon *et al.*, NeurIPS 2011])
  - Arms arranged into groups (possibly overlapping)
  - Typical goals: Best arm in each group / group whose worst arm is highest / ...

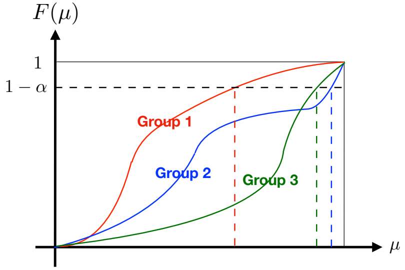


- **Infinite-arm bandit problems:** (e.g., [Aziz *et al.*, ALT 2018])
  - When a new arm is requested, its mean is drawn from a reservoir distribution F_X
  - Typical goal: Identify an arm in the top-α proportion



Our Problem: Max-Quantile Grouped Infinite-Arm Bandits

- **Max-quantile grouped infinite-arm bandits:**
  - Multiple infinite-arm "groups", each with its own reservoir distribution
  - Decision-making procedure:
    - In each round, choose a group and an arm from it to pull
    - New arms from any reservoir distribution can be requested at any time
  - Goal: Identify the group with the highest $(1-\alpha)$-quantile (e.g., highest median), with probability at least $1-\delta$ and as few (avg.) pulls as possible
  - Motivation: Finding the "best" among large populations (e.g., the one with the highest median click-through rate)

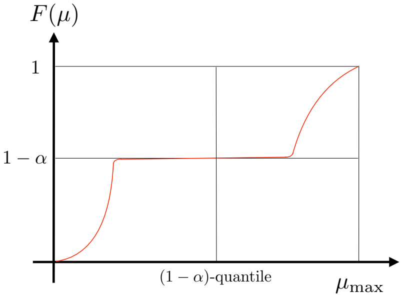


Relaxed Recovery Guarantee

- **Δ-relaxation in standard best-arm problems:** Instead of insisting on the best arm, just require one within Δ of the optimum

- **Our (ϵ, Δ)-relaxation:** Seek a group such that

$$F_G^{-1}(1-\alpha+\epsilon) \geq \max_{G' \in \mathcal{G}} F_{G'}^{-1}(1-\alpha-\epsilon) - \Delta$$

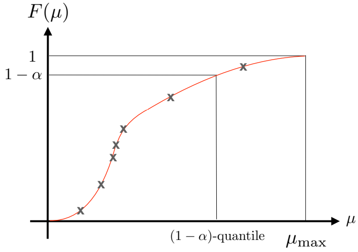


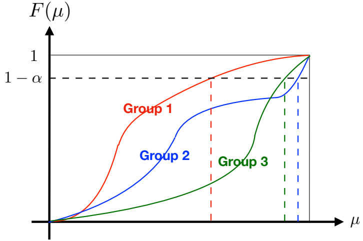
(Note: Introducing just Δ or just ϵ isn't enough – the problem becomes arbitrarily hard if <u>either</u> of them are set to zero)

Algorithm: High-Level Outline

- We propose a 2-step algorithm along similar lines to [Aziz *et al.*, ALT 2018]

- **Step 1:** Request N arms from each group, where $N \propto \frac{1}{\epsilon^2}$ so that "empirical quantiles" are ϵ-accurate





- **Step 2:** Run an elimination-based finite-arm best-quantile identification algorithm with accuracy parameter Δ (most closely related to [Wang *et al.*, AAAI 2022])

Analysis of Finite-Arm Part

Finite-arm subroutine: We consider an elimination algorithm that pulls every non-eliminated arm in every group, and then eliminates/terminates as follows.

- **Elimination rule 1: Arm definitely above (or below) its group's quantile**
  - e.g., ($\alpha = \frac{1}{2}$) $\mu_1 \in [0.1, 0.3]$, $\mu_2 \in [0.4, 0.6]$, $\mu_3 \in [0.5, 0.7]$

- **Elimination rule 2: Group definitely suboptimal**
  - e.g., ($\alpha = \frac{1}{2}$) half of group 1's LCBs exceed half of group 2's UCBs

Analysis of Finite-Arm Part

Finite-arm subroutine: We consider an elimination algorithm that pulls every non-eliminated arm in every group, and then eliminates/terminates as follows.

- **Elimination rule 1: Arm definitely above (or below) its group's quantile**
  - e.g., ($\alpha = \frac{1}{2}$) $\mu_1 \in [0.1, 0.3]$, $\mu_2 \in [0.4, 0.6]$, $\mu_3 \in [0.5, 0.7]$
- **Elimination rule 2: Group definitely suboptimal**
  - e.g., ($\alpha = \frac{1}{2}$) half of group 1's LCBs exceed half of group 2's UCBs
- **Termination rule 1: Optimal group found**
  - i.e., all other groups are eliminated
- **Termination rule 2: All remaining groups are good enough**
  - i.e., we can certify that they are all (ϵ, Δ)-optimal

Analysis of Finite-Arm Part

Finite-arm subroutine: We consider an elimination algorithm that pulls every non-eliminated arm in every group, and then eliminates/terminates as follows.

- **Elimination rule 1: Arm definitely above (or below) its group's quantile**
  - e.g., ($\alpha = \frac{1}{2}$) $\mu_1 \in [0.1, 0.3]$, $\mu_2 \in [0.4, 0.6]$, $\mu_3 \in [0.5, 0.7]$
- **Elimination rule 2: Group definitely suboptimal**
  - e.g., ($\alpha = \frac{1}{2}$) half of group 1's LCBs exceed half of group 2's UCBs
- **Termination rule 1: Optimal group found**
  - i.e., all other groups are eliminated
- **Termination rule 2: All remaining groups are good enough**
  - i.e., we can certify that they are all (ϵ, Δ)-optimal
- **Number of arm pulls:** $\mathbb{E}[T] \le c \sum_H \sum_j \frac{1}{\Delta_{H,j}^2} \log(\cdot)$ w/ groups H and arms j

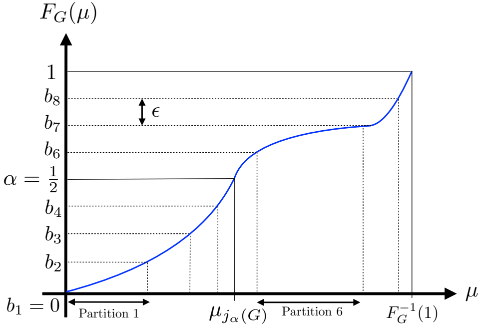
Analysis of Infinite-Arm Part

- For the overall infinite-arm problem, the preceding gaps $\Delta_{H,j}$ are random because they depend on the random reservoir distribution samples

Analysis of Infinite-Arm Part

- For the overall infinite-arm problem, the preceding gaps $\Delta_{H,j}$ are random because they depend on the random reservoir distribution samples

- Idea: Partition the reservoir distribution into regions such that $\Delta_{H,j}$ is roughly the same for all arms in a given region $\implies$ can characterize $\Delta_{H,j}$ w.h.p.
  - Analogous to [Aziz *et al.*, ALT 2018], but our problem demands finer regions and a greater number of arms to be requested
  - To ge the desired guarantee, we use ϵ-spacing in the partitioning and take $N \propto \frac{1}{\epsilon^2}$ arms from each group



Multi-Step Improvement

- **Limitation:** We pull $\frac{1}{\epsilon^2}$ arms from every group, even those that could be certified as suboptimal using much fewer arms

- **Multi-step refinement:** For epochs indexed by $k = 1, 2, \dots$
  - Request $N_k \propto \frac{1}{\epsilon_k^2}$ arms from each group
  - Run the finite-arm algorithm with parameter Δ_k
  - Eliminate all groups certified as suboptimal
  - Terminate when 1 group remains, or when final (ϵ, Δ) reached

Here (ϵ_k, Δ_k) follows a decreasing pattern (e.g., halving)

- **Regret improvement:** Each group's #pulls is dictated by the smallest ϵ_k before elimination, rather than a common choice of ϵ

Lower Bound

- Our upper bound is instance-dependent, but also gives the minimax scaling

$$T \leq \tilde{O}\left(\frac{|\mathcal{G}|}{\Delta^2 \epsilon^2}\right)$$

Theorem (Lower Bound)

For any algorithm guaranteeing (ϵ, Δ)-optimality on all instances, there must exist an instance such that

$$\mathbb{E}[T] \geq \Omega\left(\frac{|\mathcal{G}|}{\Delta^2 \epsilon^2}\right). \tag{1}$$

That is, the worst-case upper bound is tight up to logarithmic factors.

Lower Bound

- Our upper bound is instance-dependent, but also gives the minimax scaling

$$T \le \tilde{O}\left(\frac{|\mathcal{G}|}{\Delta^2 \epsilon^2} \right)$$

Theorem (Lower Bound)

For any algorithm guaranteeing (ϵ, Δ)-optimality on all instances, there must exist an instance such that

$$\mathbb{E}[T] \ge \Omega\left(\frac{|\mathcal{G}|}{\Delta^2 \epsilon^2} \right). \tag{1}$$

That is, the worst-case upper bound is tight up to logarithmic factors.

- **Proof idea:** (for $\alpha = \frac{1}{2}$)
  - Two types of arms – "good" (mean $\frac{1+\Delta}{2}$) and "bad" (mean $\frac{1-\Delta}{2}$)
  - Optimal group has $\frac{1+\epsilon}{2}$ fraction of good arms, the rest have $\frac{1-\epsilon}{2}$ fraction
  - For each group, need to observe $\Omega\left(\frac{1}{\epsilon^2}\right)$ arms; for each arm, need $\Omega\left(\frac{1}{\Delta^2}\right)$ pulls

Conclusion

Summary:

- Instance-dependent analysis of two-step alg. (sample arms + finite-arm alg.)
- Multi-step improvement
- Near-matching lower bound for worst-case instances

Possible future work:

- Instance-dependent lower bounds
- Better understanding when upper bound is/isn't near-optimal
- Other group properties beyond quantile