Collaborative Learning with Limited Interaction: Tight Bounds for Distributed Exploration in Multi-Armed Bandits

Chao Tao, Qin Zhang

IUB

Yuan Zhou

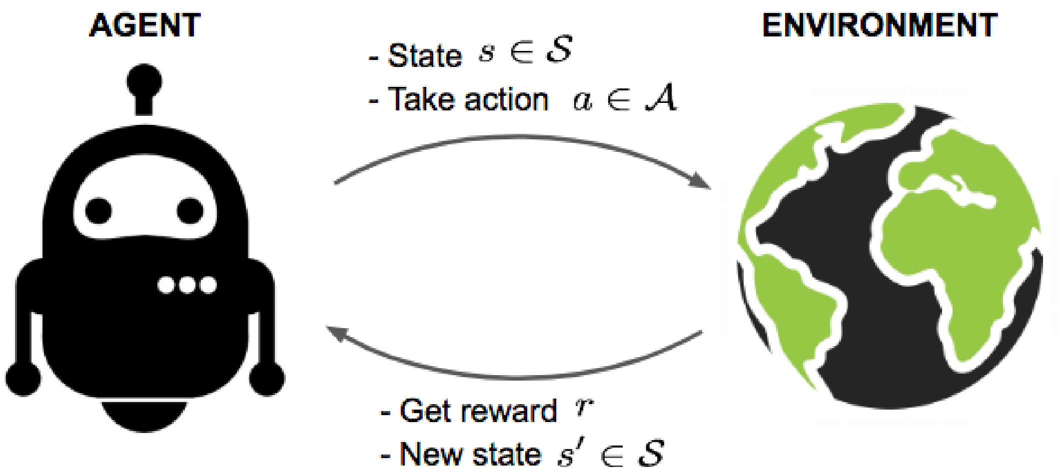
UIUC

Nov. 10, 2019

FOCS 2019

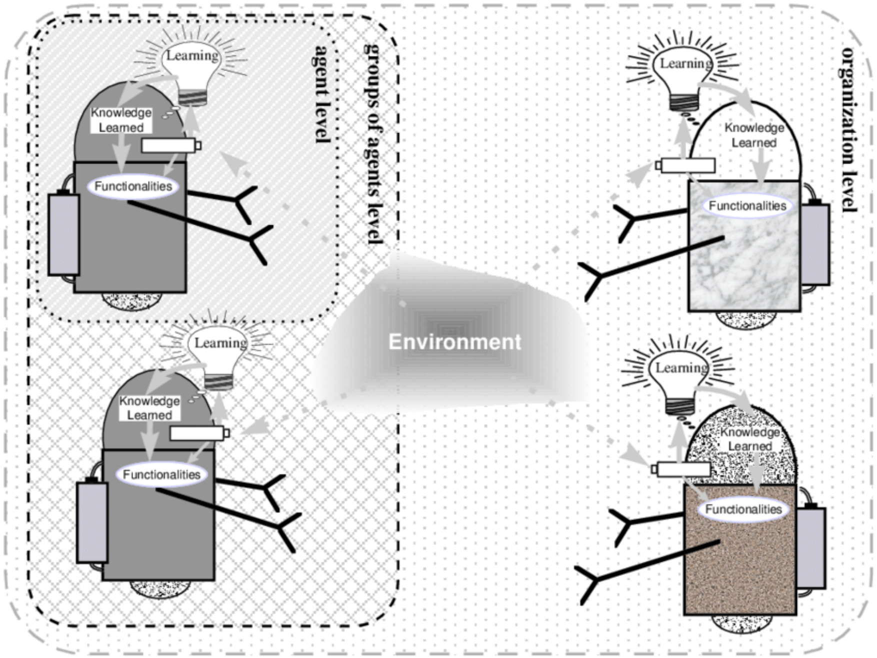
Collaborative Learning

- One of the most important tasks in machine learning is to make learning scalable.



Collaborative Learning

- One of the most important tasks in machine learning is to make learning scalable.
- A natural way to speed up the learning process is to introduce multiple agents



Collaborative Learning with Limited Collaboration

- Interaction between agents can be expensive.

Collaborative Learning with Limited Collaboration

- Interaction between agents can be expensive.

  - Time: network bandwidth/latency, protocol handshaking
  - Energy: e.g., robots exploring in the deep sea and on Mars

Collaborative Learning with Limited Collaboration

- Interaction between agents can be expensive.

  - Time: network bandwidth/latency, protocol handshaking
  - Energy: e.g., robots exploring in the deep sea and on Mars

- Interested in **tradeoffs** between #rounds of interaction and the "speedup" of collaborative learning (to be defined shortly)

Best Arm Identification in Multi-Armed Bandits

- n alternative arms (randomly permuted), where the i-th arm is associated with an unknown reward distribution μ_i with support on $[0, 1]$
- Want to identify the arm with the largest mean
- Tries to identify the best arm by a sequence of arm pulls; each pull on the i-th arm gives an *i.i.d.* sample from μ_i
- Goal (centralized setting): minimize total #arm-pulls

Best Arm Identification (cont.)

Assume each arm pull takes one time step

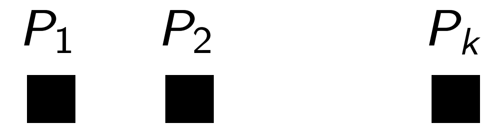
- **Fixed-time best arm**: Given a time budget T, identify the best arm with the smallest error probability

- **Fixed-confidence best arm**: Given an error probability δ, identify the best arm with error probability at most δ using the smallest amount of time

We consider both in this paper

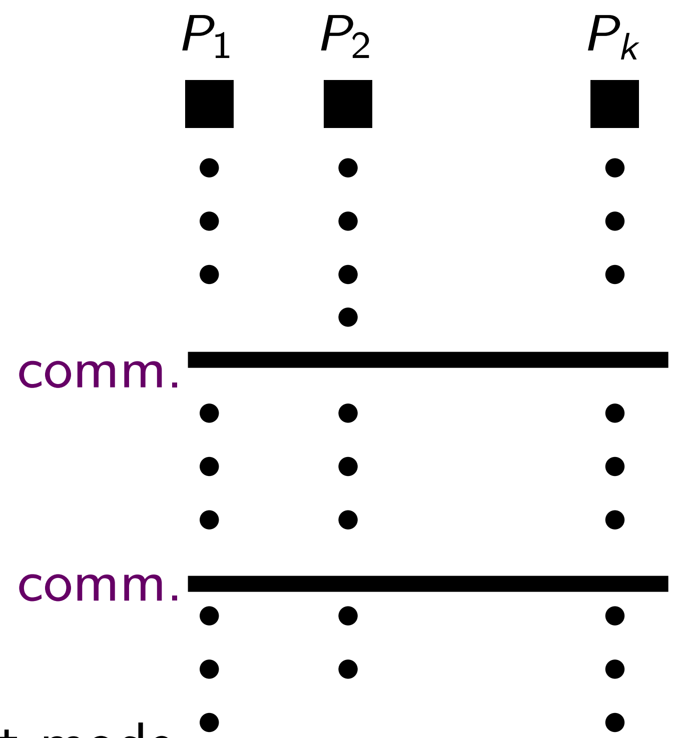
Collaborative Best Arm Identification

- n alternative arms. K agents.
  Learning proceeds in rounds.



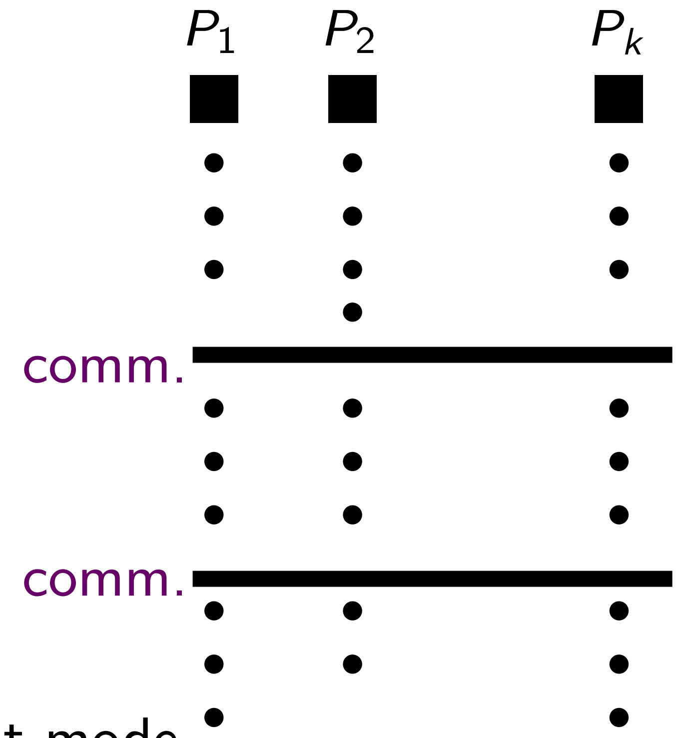
Collaborative Best Arm Identification

- n alternative arms. K agents. Learning proceeds in rounds.
- Each agent at any time, based on outcomes of all previous pulls, all msgs received, and randomness of the algo, takes one of the followings
  - makes the next pull
  - requests a comm. step and enters the wait mode
  - terminates and outputs the answer.



Collaborative Best Arm Identification

- n alternative arms. K agents. Learning proceeds in rounds.
- Each agent at any time, based on outcomes of all previous pulls, all msgs received, and randomness of the algo, takes one of the followings
  - makes the next pull
  - requests a comm. step and enters the wait mode
  - terminates and outputs the answer.
- A comm. step starts if all non-terminated agents are in the wait mode. After that agents start a new round of arm pulls



Collaborative Best Arm Identification (cont.)

- At the end, all agents need to output the *same* best arm

Collaborative Best Arm Identification (cont.)

- At the end, all agents need to output the *same* best arm
- Try to minimize
  - number of rounds R;
  - running time $T = \sum_{r \in [R]} t_r$,
    where t_r is the #time steps in the r-th round

Collaborative Best Arm Identification (cont.)

- At the end, all agents need to output the *same* best arm

- Try to minimize
  - number of rounds R;
  - running time $T = \sum_{r \in [R]} t_r$,
    where t_r is the #time steps in the r-th round

- Total cost of the algorithm: a weighted sum of R and T.
  Call for the best round-time tradeoffs

Speedup

$T_{\mathcal{A}}(I, \delta)$: expected time needed for $\mathcal{A}$ to succeed on I with probability at least $(1 - \delta)$.

- **Speedup** (of collaborative learning algorithms)

$$\beta_{\mathcal{A}}(T) = \inf_{\text{centralized } \mathcal{O}} \; \inf_{\text{instance } I} \; \inf_{\substack{\delta \in (0,1/3]: \\ T_{\mathcal{O}}(I,\delta) \leq T}} \frac{T_{\mathcal{O}}(I, \delta)}{T_{\mathcal{A}}(I, \delta)}$$

Speedup

$T_{\mathcal{A}}(I, \delta)$: expected time needed for $\mathcal{A}$ to succeed on I with probability at least $(1 - \delta)$.

- **Speedup** (of collaborative learning algorithms)

$$\beta_{\mathcal{A}}(T) = \inf_{\text{centralized } \mathcal{O}} \inf_{\text{instance } I} \inf_{\substack{\delta \in (0,1/3]: \\ T_{\mathcal{O}}(I,\delta) \leq T}} \frac{T_{\mathcal{O}}(I, \delta)}{T_{\mathcal{A}}(I, \delta)} \quad \frac{T(\text{best cen})}{T(\mathcal{A})}$$

Speedup

$T_{\mathcal{A}}(I, \delta)$: expected time needed for $\mathcal{A}$ to succeed on I with probability at least $(1 - \delta)$.

- **Speedup** (of collaborative learning algorithms)

$$\beta_{\mathcal{A}}(T) = \inf_{\text{centralized } \mathcal{O}} \inf_{\text{instance } I} \inf_{\substack{\delta \in (0,1/3]: \\ T_{\mathcal{O}}(I,\delta) \le T}} \frac{T_{\mathcal{O}}(I, \delta)}{T_{\mathcal{A}}(I, \delta)} \quad \frac{T(\text{best cen})}{T(\mathcal{A})}$$

– Our upper bound slowly degrades (in log) as T grows

Speedup

$T_{\mathcal{A}}(I, \delta)$: expected time needed for $\mathcal{A}$ to succeed on I with probability at least $(1 - \delta)$.

- **Speedup** (of collaborative learning algorithms)

$$\beta_{\mathcal{A}}(T) = \inf_{\text{centralized } \mathcal{O}} \inf_{\text{instance } I} \inf_{\substack{\delta \in (0,1/3]: \\ T_{\mathcal{O}}(I,\delta) \leq T}} \frac{T_{\mathcal{O}}(I, \delta)}{T_{\mathcal{A}}(I, \delta)}$$

– Our upper bound slowly degrades (in log) as T grows

- $\beta_{K,R}(T) = \sup_{\mathcal{A}} \beta_{\mathcal{A}}(T)$

  where sup is taken over all R-round algorithms $\mathcal{A}$ for the collaborative learning model with K agents

Our Goal

Find the best round-speedup tradeoffs

Clearly there is a tradeoff between R and $\beta_{K,R}$:

- When $R = 1$ (i.e., *no* communication step), each agent needs to solve the problem by itself, and thus $\beta_{K,1} \leq 1$.
- When R increases, $\beta_{K,R}$ may increase.
- On the other hand we always have $\beta_{K,R} \leq K$.

Previous and Our Results

problem	number of rounds[4]	$\beta_{K,R}(T)$	UB/LB	ref.
fixed-time	1	1	–	trivial
	2	$\tilde{\Omega}(\sqrt{K})$	UB	[21]
	2	$\tilde{O}(\sqrt{K})$	LB	[21]
	R	$\tilde{\Omega}(K^{\frac{R-1}{R}})$	UB	**new**
	$\Omega\left(\frac{\ln \tilde{K}}{\ln\ln \tilde{K} + \ln \frac{K}{\beta}}\right)$ when $\beta \in [K/\tilde{K}^{0.1}, K]$	β	LB	**new**
fixed-confidence	R	$\tilde{\Omega}\left((\Delta_{\min})^{\frac{2}{R-1}} K\right)$	UB	[21]
	$\Omega\left(\min\left\{\frac{\ln \frac{1}{\Delta_{\min}}}{\ln\left(1+\frac{K(\ln K)^2}{\beta}\right)+\ln\ln \frac{1}{\Delta_{\min}}}, \sqrt{\frac{\beta}{(\ln K)^3}}\right\}\right)$	β	LB	**new**

[21]: Hillel et al. NIPS 2013; $\Delta_{\min}$ = mean of best arm - mean of 2nd best arm

Previous and Our Results

problem	number of rounds[4]	$\beta_{K,R}(T)$	UB/LB	ref.
fixed-time	1	1	–	trivial
	2	$\tilde{\Omega}(\sqrt{K})$	UB	[21]
	2	$\tilde{O}(\sqrt{K})$	LB	[21]
	$\ln K$	$\tilde{\Omega}(K)$	UB	**new**
	$\Omega(\ln K / \ln\ln K)$	$K / \ln^{O(1)} K$	LB	**new**
fixed-confidence	$\ln \frac{1}{\Delta_{\min}}$	$\tilde{\Omega}(K)$	UB	[21]
	$\Omega\left(\ln \frac{1}{\Delta_{\min}} / (\ln\ln K + \ln\ln \frac{1}{\Delta_{\min}})\right)$	$K / \ln^{O(1)} K$	LB	**new**

[21]: Hillel et al. NIPS 2013; $\Delta_{\min}$ = mean of best arm - mean of 2nd best arm

Previous and Our Results

problem	number of rounds[4]	$\beta_{K,R}(T)$	UB/LB	ref.
fixed-time	1	1	–	trivial
	2	$\tilde{\Omega}(\sqrt{K})$	UB	[21]
	2	$\tilde{O}(\sqrt{K})$	LB	[21]
	$\ln K$	$\tilde{\Omega}(K)$	UB	**new**
	$\Omega(\ln K / \ln \ln K)$	$K / \ln^{O(1)} K$	LB	**new**
fixed-confidence	$\ln \frac{1}{\Delta_{\min}}$	$\tilde{\Omega}(K)$	UB	[21]
	$\Omega\left(\ln \frac{1}{\Delta_{\min}} / (\ln \ln K + \ln \ln \frac{1}{\Delta_{\min}})\right)$	$K / \ln^{O(1)} K$	LB	**new**

[21]: Hillel et al. NIPS 2013; $\Delta_{\min}$ = mean of best arm - mean of 2nd best arm

- Almost tight round-speedup tradeoffs for fixed-time. Today's focus (LB)
- Almost tight round-speedup tradeoffs for fixed-confidence.
- A separation for two problems.

Previous and Our Results

problem	number of rounds[4]	$\beta_{K,R}(T)$	UB/LB	ref.
fixed-time	1	1	–	trivial
	2	$\tilde{\Omega}(\sqrt{K})$	UB	[21]
	2	$\tilde{O}(\sqrt{K})$	LB	[21]
	$\ln K$	$\tilde{\Omega}(K)$	UB	**new**
	$\Omega(\ln K / \ln \ln K)$	$K / \ln^{O(1)} K$	LB	**new**
fixed-confidence	$\ln \frac{1}{\Delta_{\min}}$	$\tilde{\Omega}(K)$	UB	[21]
	$\Omega\left(\ln \frac{1}{\Delta_{\min}} / (\ln \ln K + \ln \ln \frac{1}{\Delta_{\min}})\right)$	$K / \ln^{O(1)} K$	LB	**new**

[21]: Hillel et al. NIPS 2013; $\Delta_{\min}$ = mean of best arm - mean of 2nd best arm

- Almost tight round-speedup tradeoffs for fixed-time. Today's focus (LB)
- Almost tight round-speedup tradeoffs for fixed-confidence.
- A separation for two problems.
- A generalization of the round-elimination technique. Today
- A new technique for instance-dependent round complexity.

Lower Bound: Fixed-Time

Round Elimination: A Technique for Round LB

- $\exists$ an r-round algorithm with error prob. δ_r and time budget T on an input distribution σ_r,

  $\Rightarrow$

  $\exists$ an $(r-1)$-round algorithm with error prob. $\delta_{r-1} (> \delta_r)$ and time budget T on an input distribution σ_{r-1}.

- There is *no* 0-round algorithm with error prob. $\delta_0 \ll 1$ on a nontrivial input distribution σ_0.

Round Elimination: A Technique for Round LB

- $\exists$ an r-round algorithm with error prob. δ_r and time budget T on an input distribution σ_r,

  $\Rightarrow$

  $\exists$ an $(r-1)$-round algorithm with error prob. $\delta_{r-1} (> \delta_r)$ and time budget T on an input distribution σ_{r-1}.

- There is *no* 0-round algorithm with error prob. $\delta_0 \ll 1$ on a nontrivial input distribution σ_0.

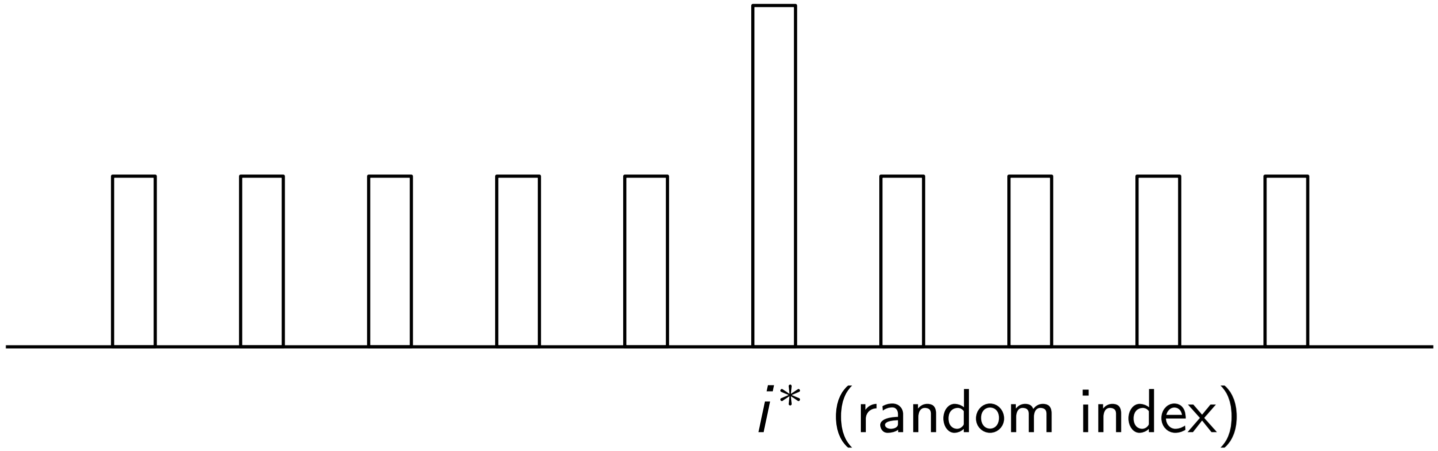
$\Rightarrow$ Any algo with time budget T and error prob. 0.01 needs at least r rounds of comm.

Previous Use of Round Elimination

- Agarwal et al. (COLT'17) used *round elimination* to prove an $\Omega(\log^* n)$ for best arm identification under time budget $T = \tilde{O}\left(\frac{n}{\Delta_{\min}^2}/K\right)$ for non-adaptive algos
  - Translated into our collaborative learning setting
  - Non-adaptive algos: all arm pulls should be determined at the beginning of each round

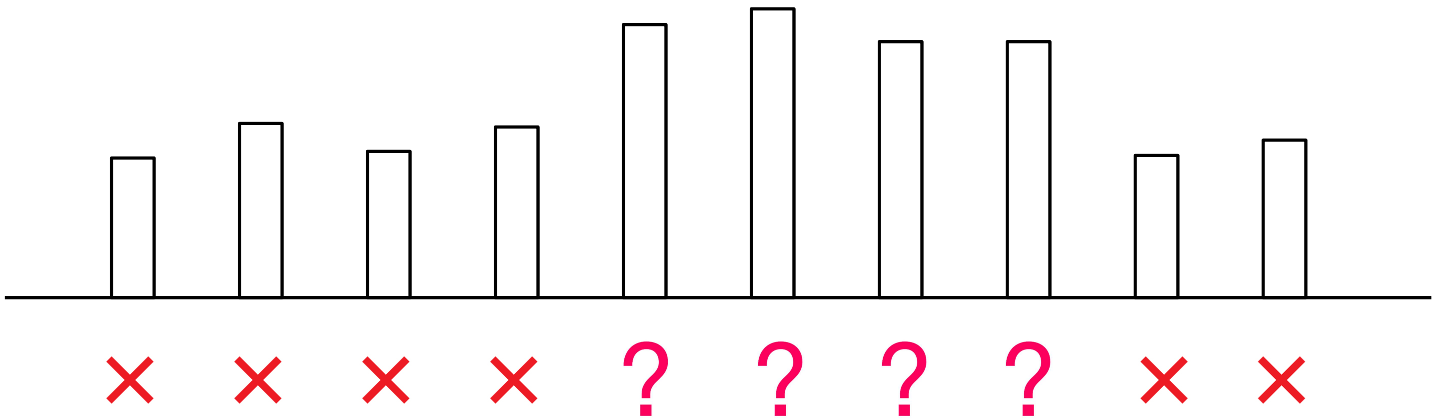
Previous Use of Round Elimination

- Agarwal et al. (COLT'17) used *round elimination* to prove an $\Omega(\log^* n)$ for best arm identification under time budget $T = \tilde{O}\left(\frac{n}{\Delta_{\min}^2}/K\right)$ for non-adaptive algos
  - Translated into our collaborative learning setting
  - Non-adaptive algos: all arm pulls should be determined at the beginning of each round
- **"One-spike" distribution**: a random single arm with mean $\frac{1}{2}$, and $(n-1)$ arms with mean $\left(\frac{1}{2} - \Delta_{\min}\right)$.



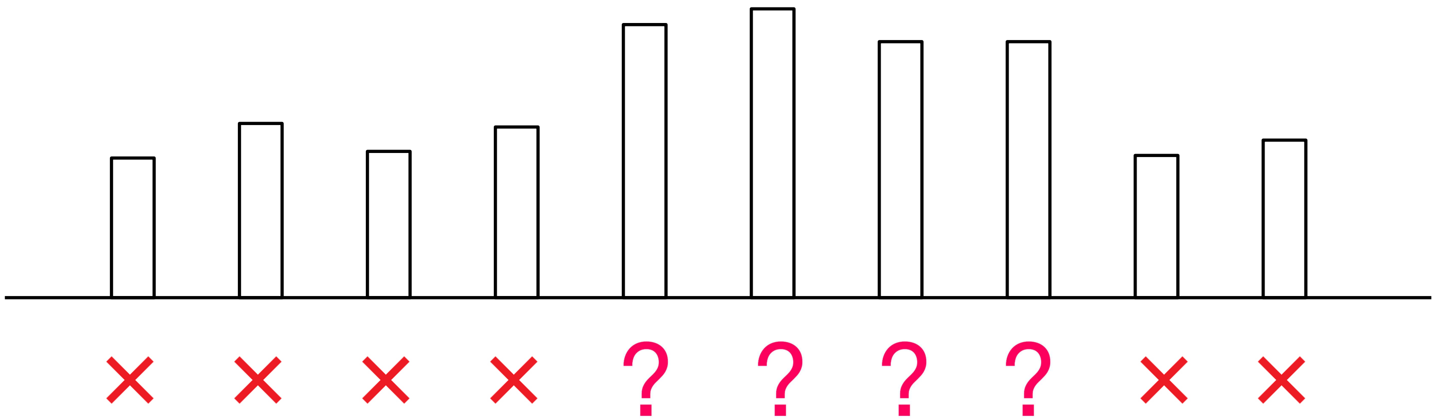
Previous Use of Round Elimination (Cont.)

- **Basic argument** (of COLT'17): If we do not make enough pulls in the first round, then conditioned on the pull outcomes, the index of the best arm is still quite uncertain



Previous Use of Round Elimination (Cont.)

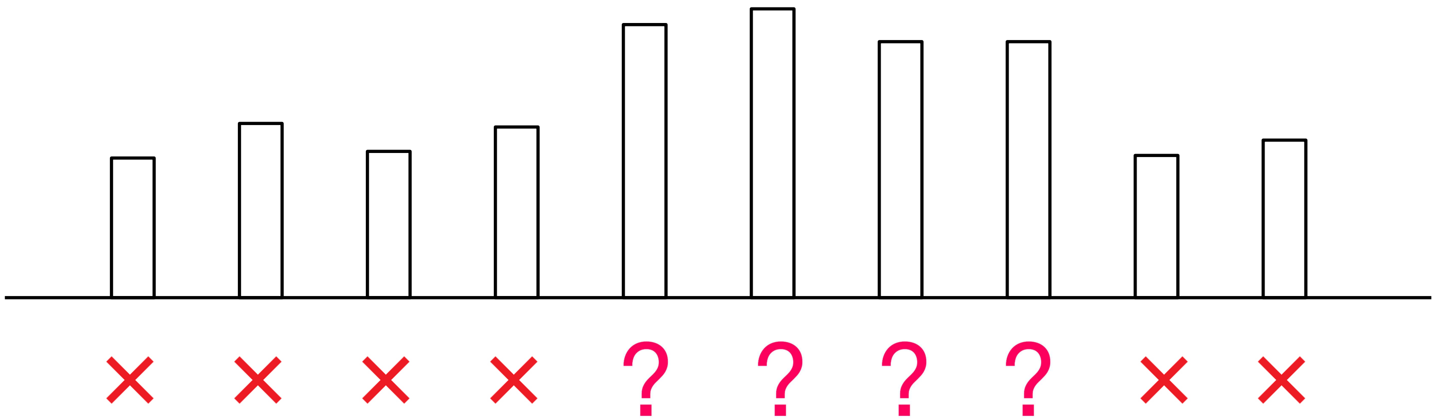
- **Basic argument** (of COLT'17): If we do not make enough pulls in the first round, then conditioned on the pull outcomes, the index of the best arm is still quite uncertain



- More precisely, the posterior distribution of the index of the best arm can be written as a convex combination of a set of distributions, each of which has a large support size ($\geq \log n$) and is close to the uniform distribution

Previous Use of Round Elimination (Cont.)

- **Basic argument** (of COLT'17): If we do not make enough pulls in the first round, then conditioned on the pull outcomes, the index of the best arm is still quite uncertain



- More precisely, the posterior distribution of the index of the best arm can be written as a convex combination of a set of distributions, each of which has a large support size ($\geq \log n$) and is close to the uniform distribution

  $\Rightarrow$ an $\Omega(\log^* n)$ LB

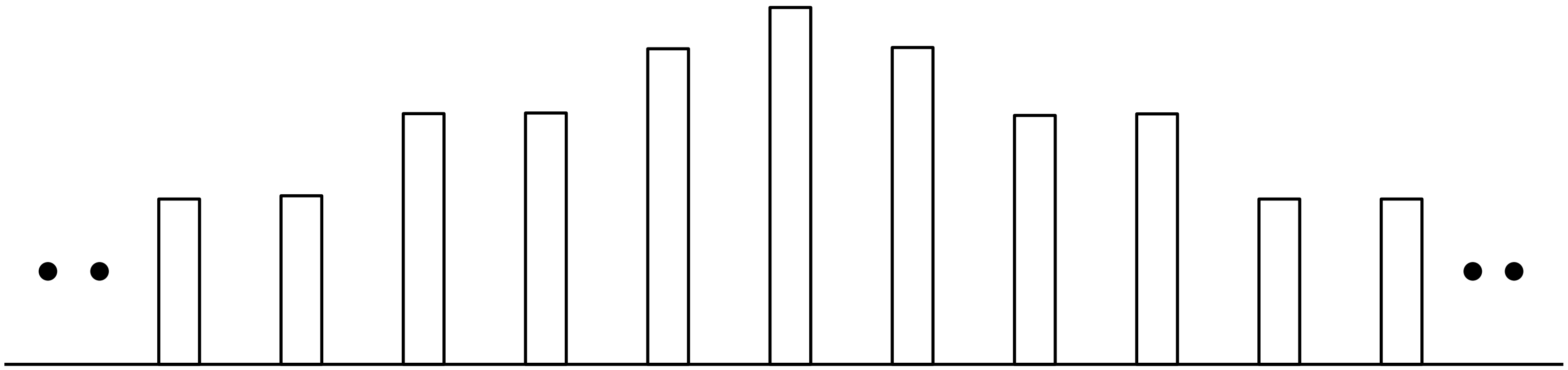
The Challenge

- We want to prove a logarithmic round lower bound.
- We need to restrict the time budget within a better bound $\tilde{O}(H/K) = \tilde{O}\left(\sum_{i=2}^{n} \frac{1}{\Delta_i^2}/K\right)$

  ($\Delta_i =$ mean of the best arm - mean of the i-th best arm in the input)

The Challenge

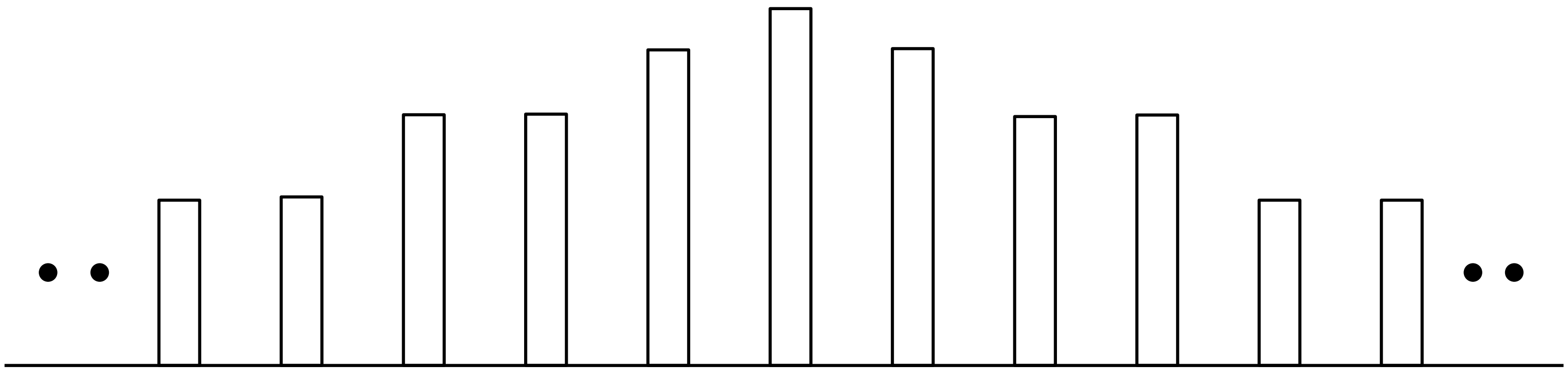
- We want to prove a logarithmic round lower bound.
- We need to restrict the time budget within a better bound $\tilde{O}(H/K) = \tilde{O}\left(\sum_{i=2}^{n} \frac{1}{\Delta_i^2}/K\right)$
  (Δ_i = mean of the best arm - mean of the i-th best arm in the input)
- **“Pyramid-like” distribution**: Roughly speaking, we take $n/2$ random arms and assign them with mean $(1/2 - 1/4)$, $n/4$ random arms with mean $(1/2 - 1/8)$, and $n/8$ random arms with mean $(1/2 - 1/16)$, ...



The Challenge (Cont.)

Technical challenge (if want to follow COLT'17):

Not clear how to decompose the posterior distribution of the means of arms into a convex combination of a set of distributions, each of which is close to the same pyramid-like distribution.



New Idea: Generalized Round Elimination

- $\exists$ r-round algorithm with error prob. δ_r and time budget T on *any* distribution in distribution class $\mathcal{D}_r$

  $\Rightarrow$

  $\exists$ $(r-1)$-round algorithm with error prob. $\delta_{r-1}(> \delta_r)$ and time budget T on *any* distribution in distribution class $\mathcal{D}_{r-1}$

- There is *no* 0-round algorithm with error prob. $\delta_0 \ll 1$ on any input distribution in $\mathcal{D}_0$

New Idea: Generalized Round Elimination

- $\exists$ r-round algorithm with error prob. δ_r and time budget T on *any* distribution in distribution class $\mathcal{D}_r$

  $\Rightarrow$

  $\exists$ $(r-1)$-round algorithm with error prob. $\delta_{r-1} (> \delta_r)$ and time budget T on *any* distribution in distribution class $\mathcal{D}_{r-1}$
- There is *no* 0-round algorithm with error prob. $\delta_0 \ll 1$ on any input distribution in $\mathcal{D}_0$

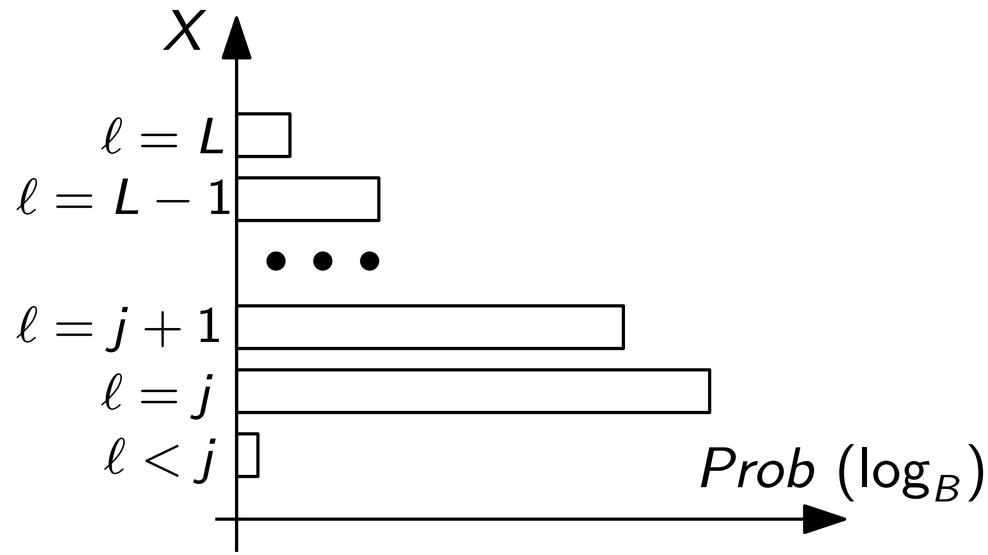
Advantage: we do *not* need to show that the posterior distribution ν' of $\nu \in \mathcal{D}_r$ is close to a particular distribution, but only that $\nu' \in \mathcal{D}_{r-1}$.

Hard Input Distribution Classes

Let $\alpha \in [1, n^{0.2}]$ be a parameter, $B = \gamma = \alpha \log^{10} n$,
$L = \log n / (\log\log n + \log \alpha)$, $\rho = \log^3 n$.

Define $\mathcal{D}_j$ to be the class of distributions μ with support $\{B^{-1}, \ldots, B^{-(j-1)}, B^{-j}, \ldots, B^{-L}\}$, such that if $X \sim \mu$, then

1. For any $\ell = j, \ldots, L$, $\Pr[X = B^{-\ell}] = \lambda_j \cdot B^{-2\ell} \cdot \left(1 \pm \rho^{-\ell+j-1}\right)$, where λ_j is a normalization factor
2. $\Pr\left[(X = B^{-1}) \vee \cdots \vee (X = B^{-(j-1)})\right] \leq n^{-9}$, $(j \geq 2)$



Hard Input Distribution Classes

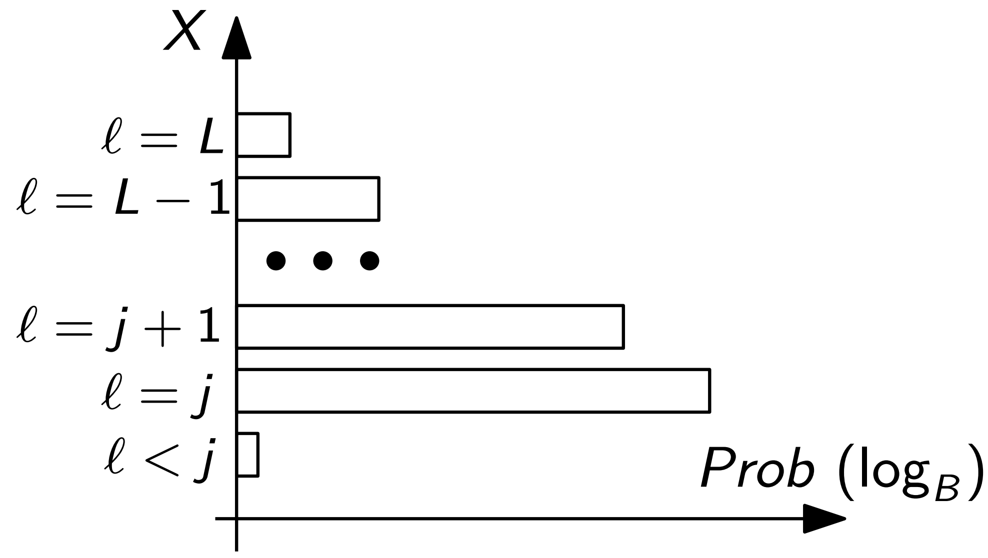
Let $\alpha \in [1, n^{0.2}]$ be a parameter, $B = \gamma = \alpha \log^{10} n$,
$L = \log n / (\log\log n + \log \alpha)$, $\rho = \log^3 n$.

Define $\mathcal{D}_j$ to be the class of distributions μ with support $\{B^{-1}, \ldots, B^{-(j-1)}, B^{-j}, \ldots, B^{-L}\}$, such that if $X \sim \mu$, then

1. For any $\ell = j, \ldots, L$, $\Pr[X = B^{-\ell}] = \lambda_j \cdot B^{-2\ell} \cdot \left(1 \pm \rho^{-\ell+j-1}\right)$, where λ_j is a normalization factor
2. $\Pr\left[(X = B^{-1}) \vee \cdots \vee (X = B^{-(j-1)})\right] \leq n^{-9}$, $(j \geq 2)$

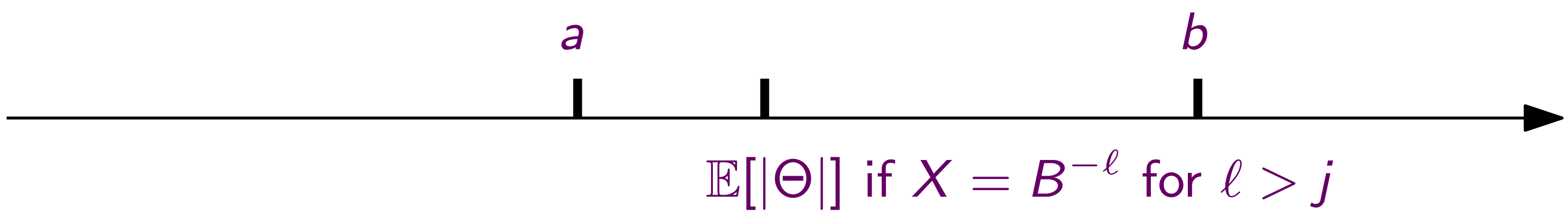
Arms i.i.d. with mean $\frac{1}{2} - X$

Try to embed the pyramid distribution into each arm



Hard Input Distribution Classes (cont.)

$a = \left(\frac{1}{2} - B^{-(j+1)}\right)\gamma B^{2j} - \sqrt{10\gamma \ln n}B^{j}$, $b = \frac{\gamma B^{2j}}{2} + B^{j+0.6}$



Key property of the distribution class:

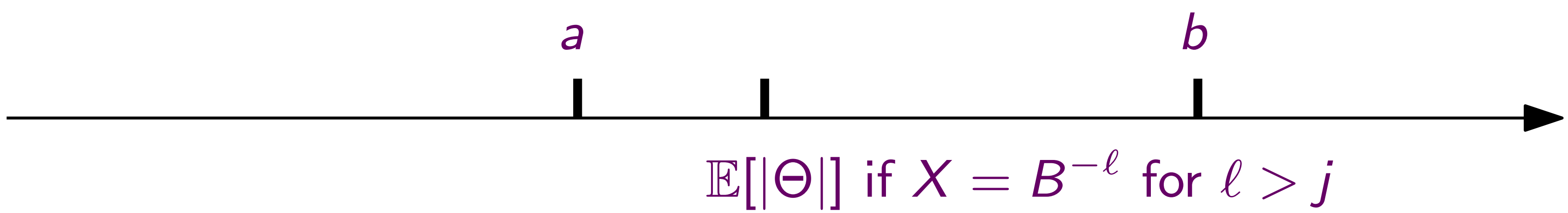
Consider an arm with mean $\left(\frac{1}{2} - X\right)$ where $X \sim \mu \in \mathcal{D}_j$ for some $j \in [L-1]$. We pull the arm γB^{2j} times.

Let $\Theta = (\Theta_1, \Theta_2, \ldots, \Theta_{\gamma B^{2j}})$ be the pull outcomes, and let $|\Theta| = \sum_{i \in [\gamma B^{2j}]} \Theta_i$.

If $|\Theta| \notin [a, b]$, then publish the arm.

Hard Input Distribution Classes (cont.)

$a = \left(\frac{1}{2} - B^{-(j+1)}\right)\gamma B^{2j} - \sqrt{10\gamma \ln n} B^j$, $b = \frac{\gamma B^{2j}}{2} + B^{j+0.6}$



Key property of the distribution class:

Consider an arm with mean $\left(\frac{1}{2} - X\right)$ where $X \sim \mu \in \mathcal{D}_j$ for some $j \in [L-1]$. We pull the arm γB^{2j} times.

Let $\Theta = (\Theta_1, \Theta_2, \ldots, \Theta_{\gamma B^{2j}})$ be the pull outcomes, and let $|\Theta| = \sum_{i \in [\gamma B^{2j}]} \Theta_i$.

If $|\Theta| \notin [a, b]$, then publish the arm.

Let ν be the posterior distribution of X after observing Θ.
If the arm is not published, then we must have $\nu \in \mathcal{D}_{j+1}$.

Lower Bound for Non-Adaptive Algorithms

Theorem 1. Any (K/α)-speedup non-adaptive algorithm for the fixed-time best arm identification problem in the collaborative learning model with K agents needs $\Omega(L) = \Omega(\ln n/(\ln\ln n + \ln\alpha))$ rounds.

Lower Bound for Non-Adaptive Algorithms

Theorem 1. Any (K/α)-speedup non-adaptive algorithm for the fixed-time best arm identification problem in the collaborative learning model with K agents needs $\Omega(L) = \Omega(\ln n/(\ln \ln n + \ln \alpha))$ rounds.

Round reduction. For any $j \leq \frac{L}{2} - 1$,

$\exists$ r-round (K/α)-speedup non-adaptive algorithm with error prob. δ on any input distribution in $(\mathcal{D}_j)^{n_j}$ for any $n_j \in I_j$. $(I_j = ((1 \pm \frac{1}{L})B^{-2})^{j-1})$

$\Rightarrow$

$\exists$ $(r-1)$-round (K/α)-speedup non-adaptive algorithm with error prob. $\delta + o\left(\frac{1}{L}\right)$ on any input distribution in $(\mathcal{D}_{j+1})^{n_{j+1}}$ for any $n_{j+1} \in I_{j+1}$

Lower Bound for Non-Adaptive Algorithms

Theorem 1. Any (K/α)-speedup non-adaptive algorithm for the fixed-time best arm identification problem in the collaborative learning model with K agents needs $\Omega(L) = \Omega(\ln n/(\ln\ln n + \ln\alpha))$ rounds.

Round reduction. For any $j \le \frac{L}{2} - 1$,

$\exists$ r-round (K/α)-speedup non-adaptive algorithm with error prob. δ on any input distribution in $(\mathcal{D}_j)^{n_j}$ for any $n_j \in I_j$. $(I_j = ((1 \pm \frac{1}{L})B^{-2})^{j-1})$

$\Rightarrow$

$\exists$ $(r-1)$-round (K/α)-speedup non-adaptive algorithm with error prob. $\delta + o\left(\frac{1}{L}\right)$ on any input distribution in $(\mathcal{D}_{j+1})^{n_{j+1}}$ for any $n_{j+1} \in I_{j+1}$

Base Case: Any 0-round algorithm must have error 0.99 on any distribution in $(\mathcal{D}_{\frac{L}{2}})^{n_{\frac{L}{2}}}$ $(\forall\ n_{\frac{L}{2}} \in I_{\frac{L}{2}})$.

Proof Idea for Round Reduction

Let S be the set of arms which will be pulled more than γB^{2j} times (note: we are considering non-adaptive algos)

Proof Idea for Round Reduction

Let S be the set of arms which will be pulled more than γB^{2j} times (note: we are considering non-adaptive algos)

Algorithm Augmentation (for j-th round)

1. Publish all arms in S.
2. For the rest arms $z \in [n_j] \backslash S$, keep pulling them until #pulls reaches γB^{2j}. Let $\Theta_z = (\Theta_{z,1}, \ldots, \Theta_{z,\gamma B^{2j}})$ be the γB^{2j} pull outcomes.
   If $|\Theta_z| \notin [a, b]$, we publish the arm.
3. If #unpublished arms is not in the range of I_{j+1}, or there is a published arm with mean $\left(\frac{1}{2} - B^{-L}\right)$, then we return "error".

$\Rightarrow$ (by key property of $\mathcal{D}_j$) resulting posterior distribution on unpublished arms in $(\mathcal{D}_{j+1})^{n_{j+1}}$ ($n_{j+1} \in I_{j+1}$)

Proof Idea for Round Reduction

Let S be the set of arms which will be pulled more than γB^{2j} times (note: we are considering non-adaptive algos)

Algorithm Augmentation (for j-th round)

1. Publish all arms in S.
2. For the rest arms $z \in [n_j] \setminus S$, keep pulling them until #pulls reaches γB^{2j}. Let $\Theta_z = (\Theta_{z,1}, \ldots, \Theta_{z,\gamma B^{2j}})$ be the γB^{2j} pull outcomes. If $|\Theta_z| \notin [a, b]$, we publish the arm.
3. If #unpublished arms is not in the range of I_{j+1}, or there is a published arm with mean $\left(\frac{1}{2} - B^{-L}\right)$, then we return "error".

$\Rightarrow$ (by key property of $\mathcal{D}_j$) resulting posterior distribution on unpublished arms in $(\mathcal{D}_{j+1})^{n_{j+1}}$ $(n_{j+1} \in I_{j+1})$

- Steps 1&2 only help the algorithm $\Rightarrow$ a stronger lower bound.
- Extra error by Step 3 is small; counted in $o(\frac{1}{L})$ in the induction.

Lower Bound for Adaptive Algorithms

Theorem 2. Any (K/α)-speedup (adaptive) algorithm for the fixed-time best arm identification problem in the collaborative learning model with K agents needs $\Omega(\ln K/(\ln \ln K + \ln \alpha))$ rounds.

Intuition: When the number of arms n is smaller than $\#$agents K, adaptive pulls do not have much advantage against non-adaptive pulls in each round.

- Prove by a coupling-like argument: Compare the behavior of an adaptive algorithm with that of a non-adaptive one.

Other main results:

1. An almost matching upper bound for the fixed-time case
2. An almost tight lower bound for the fixed-confidence case

Concluding Remarks and Future Work

- A systematic study of the best arm identification problem in the setting of collaborative learning with limited interaction

- Almost tight round-speedup tradeoffs for both fixed-time and fixed-confidence settings.

- New techniques for proving round lower bounds for multi-agent collaborative learning

Concluding Remarks and Future Work

- A systematic study of the best arm identification problem in the setting of collaborative learning with limited interaction
- Almost tight round-speedup tradeoffs for both fixed-time and fixed-confidence settings.
- New techniques for proving round lower bounds for multi-agent collaborative learning
- New direction: comm.-efficient collaborative learning. Many open problems: regrets (bandits), general reinforcement learning, etc.

Thank you!
Questions?

Upper Bound: Fixed-Time

Algorithm with Constant Error Probability

When $T = \tilde{\Theta}(HK^{-\frac{R-1}{R}})$, the algo succeeds w.pr. 0.99

Phase 1 : Eliminate most of the suboptimal arms and leave at most K candidates.

– Randomly partition the n arms to K agents.

– Each agent runs a centralized algo for $T/2$ time, outputs the best arm if terminates, '$\perp$' otherwise

Phase 2 : Run R rounds, the goal of the r-th round is to reduce #candidates to $K^{\frac{R-1}{R}}$.

In each round:

– Each agent spends $T/(2R)$ time uniformly on K arms.

– Eliminate arms whose empirical means smaller than (top empirical mean $-\ \epsilon(K, R, T, \#\text{candidates})$)

Algorithm for General Parameter Settings

- **Goal**: When $T \gg HK^{-\frac{R-1}{R}}$, the error diminishes exponentially in T

Algorithm for General Parameter Settings

- **Goal**: When $T \gg HK^{-\frac{R-1}{R}}$, the error diminishes exponentially in T

- **Challenge** 1: We do not know the instance dependent parameter H

Algorithm for General Parameter Settings

- **Goal**: When $T \gg HK^{-\frac{R-1}{R}}$, the error diminishes exponentially in T
- **Challenge** 1: We do not know the instance dependent parameter H
- **Idea** 1: Guess H using the *doubling* method

Algorithm for General Parameter Settings

- **Goal**: When $T \gg HK^{-\frac{R-1}{R}}$, the error diminishes exponentially in T
- **Challenge** 1: We do not know the instance dependent parameter H
- **Idea** 1: Guess H using the *doubling* method
- **Challenge** 2: When $T \ll HK^{-\frac{R-1}{R}}$, centralized algo may consistently return the same suboptimal arm (there is no guarantee).

Algorithm for General Parameter Settings

- **Goal**: When $T \gg HK^{-\frac{R-1}{R}}$, the error diminishes exponentially in T
- **Challenge** 1: We do not know the instance dependent parameter H
- **Idea** 1: Guess H using the *doubling* method
- **Challenge** 2: When $T \ll HK^{-\frac{R-1}{R}}$, centralized algo may consistently return the same suboptimal arm (there is no guarantee).
- **Idea** 2: Instead of fixing time budget of the first phase to $\frac{T}{2}$, choose a random time budget in $\{\frac{T}{2}, \frac{T}{200}\}$

Lower Bound: Fixed-Confidence

The *SignID* Problem

SignID: There is one Bernoulli arm with mean $(\frac{1}{2} + \Delta)$

Goal: Make min #pulls on the arm and decide whether $\Delta > 0$ or $\Delta < 0$. Let $I(\Delta)$ denote the input instance.

Say $\mathcal{A}$ is β-fast for the instance $I(\Delta)$, if

$$\Pr_{I(\Delta)}\left[\mathcal{A} \text{ succeeds within } \Delta^{-2}/\beta \text{ time}\right] \geq 2/3.$$

The *SignID* Problem

SignID: There is one Bernoulli arm with mean $(\frac{1}{2} + \Delta)$

Goal: Make min #pulls on the arm and decide whether $\Delta > 0$ or $\Delta < 0$. Let $I(\Delta)$ denote the input instance.

Say $\mathcal{A}$ is β-fast for the instance $I(\Delta)$, if

$$\Pr_{I(\Delta)}\left[\mathcal{A} \text{ succeeds within } \Delta^{-2}/\beta \text{ time}\right] \geq 2/3.$$

- A β-speedup best arm identification algorithm $\Rightarrow$ an $\Omega(\beta)$-fast algorithm for *SignID*

Main Theorem for *SignID*

Theorem. Let $\Delta^* \in (0, 1/8)$. If $\mathcal{A}$ is a $(1/K^5)$-error β-fast algorithm for every *SignID* problem instance $I(\Delta)$ where $|\Delta| \in [\Delta^*, 1/8)$, then there exists $\Delta^\flat \geq \Delta^*$ such that

$$\Pr_{I(\Delta^\flat)}\left[\mathcal{A} \text{ uses } \Omega\left(\frac{\ln(1/\Delta^*)}{\ln(1+K/\beta) + \ln\ln(1/\Delta^*)}\right) \text{ rounds}\right] \geq \frac{1}{2}.$$

Main Theorem for *SignID*

Theorem. Let $\Delta^* \in (0, 1/8)$. If $\mathcal{A}$ is a $(1/K^5)$-error β-fast algorithm for every *SignID* problem instance $I(\Delta)$ where $|\Delta| \in [\Delta^*, 1/8)$, then there exists $\Delta^\flat \geq \Delta^*$ such that

$$\Pr_{I(\Delta^\flat)}\left[\mathcal{A} \text{ uses } \Omega\left(\frac{\ln(1/\Delta^*)}{\ln(1+K/\beta) + \ln\ln(1/\Delta^*)}\right) \text{ rounds}\right] \geq \frac{1}{2}.$$

Prove using two lemmas alternatively (next slide)

- Progress lemma
- Distribution exchange lemma

The Progress Lemma

$\mathcal{E}(\alpha, T)$: $\mathcal{A}$ uses at least α rounds and at most T time before the end of the α-th round.

$\mathcal{E}^*(\alpha, T)$: $\mathcal{A}$ uses at least $(\alpha + 1)$ rounds and at most T time before the end of the α-th round.

Progress Lemma. For any $\Delta \leq 1/8$, $\alpha \geq 0$, $q \geq 1$, if $\Pr_{I(\Delta)}[\mathcal{E}(\alpha, \Delta^{-2}/(Kq))] \geq 1/2$, then

$$\Pr_{I(\Delta)}[\mathcal{E}^*(\alpha, \Delta^{-2}/(Kq))] \geq \Pr_{I(\Delta)}[\mathcal{E}(\alpha, \Delta^{-2}/(Kq))] - \delta(K, q)$$

The Progress Lemma

$\mathcal{E}(\alpha, T)$: $\mathcal{A}$ uses at least α rounds and at most T time before the end of the α-th round.

$\mathcal{E}^*(\alpha, T)$: $\mathcal{A}$ uses at least $(\alpha + 1)$ rounds and at most T time before the end of the α-th round.

Progress Lemma. For any $\Delta \leq 1/8$, $\alpha \geq 0$, $q \geq 1$, if $\Pr_{I(\Delta)}[\mathcal{E}(\alpha, \Delta^{-2}/(Kq))] \geq 1/2$, then

$$\Pr_{I(\Delta)}[\mathcal{E}^*(\alpha, \Delta^{-2}/(Kq))] \geq \Pr_{I(\Delta)}[\mathcal{E}(\alpha, \Delta^{-2}/(Kq))] - \delta(K, q)$$

Intuition. If $\mathcal{A}$ can only use $\Delta^{-2}/(Kq) \times K = \Delta^{-2}/q$ pulls for a large enough q in one round , then we cannot tell $I(\Delta)$ from $I(-\Delta)$.

The Distribution Exchange Lemma

$\mathcal{E}(\alpha, T)$: $\mathcal{A}$ uses at least α rounds and at most T time before the end of the α-th round.

$\mathcal{E}^*(\alpha, T)$: $\mathcal{A}$ uses at least $(\alpha + 1)$ rounds and at most T time before the end of the α-th round.

Distribution Exchange Lemma. For any $\Delta \leq 1/8$, $\alpha \geq 0$, $q \geq 100$, $\zeta \geq 1$,

$$\Pr_{I(\Delta/\zeta)}[\mathcal{E}(\alpha + 1, \Delta^{-2}/(Kq) + \Delta^{-2}/\beta)]$$

$$\geq \Pr_{I(\Delta)}[\mathcal{E}^*(\alpha, \Delta^{-2}/(Kq))] - \delta'(K, q, \beta)$$

The Distribution Exchange Lemma

$\mathcal{E}(\alpha, T)$: $\mathcal{A}$ uses at least α rounds and at most T time before the end of the α-th round.

$\mathcal{E}^*(\alpha, T)$: $\mathcal{A}$ uses at least $(\alpha + 1)$ rounds and at most T time before the end of the α-th round.

Distribution Exchange Lemma. For any $\Delta \le 1/8$, $\alpha \ge 0$, $q \ge 100$, $\zeta \ge 1$,

$$\Pr_{I(\Delta/\zeta)}[\mathcal{E}(\alpha+1, \Delta^{-2}/(Kq) + \Delta^{-2}/\beta)] \ge \Pr_{I(\Delta)}[\mathcal{E}^*(\alpha, \Delta^{-2}/(Kq))] - \delta'(K, q, \beta)$$

Intuition. For instance $I(\Delta)$, since $\mathcal{A}$ is a β-fast algorithm, each agent uses at most Δ^{-2}/β pulls during the $(\alpha+1)$-st round, and only sees at most $(\Delta^{-2}/(Kq) + \Delta^{-2}/\beta)$ pull outcomes before the next communication, which is insufficient to tell between $I(\Delta)$ and $I(\Delta/\zeta)$.

A Technical Lemma

Cannot simply bound the statistical distance of induced by Δ and Δ/ζ. Need the following technical lemma.

Technical Lemma. Suppose $0 \le \Delta' \le \Delta \le 1/8$.
For any positive integer $m = \Delta^{-2}/\xi$ where $\xi \ge 100$.
$\mathcal{D} = \mathcal{B}(1/2 + \Delta)^{\otimes m}$, $\mathcal{D}' = \mathcal{B}(1/2 + \Delta')^{\otimes m}$

Let $\mathcal{X}$ be any probability distribution with sample space X. For any event $A \subseteq \{0,1\}^m \times X$ such that $\Pr_{\mathcal{D}\otimes\mathcal{X}}[A] \le \gamma$, we have that

$$\Pr_{\mathcal{D}'\otimes\mathcal{X}}[A] \le \gamma \cdot \exp\left(5\sqrt{(3\ln Q)/\xi}\right) + 1/Q^6,$$

holds for all $Q \ge \xi$.

Put Together

Progress Lemma. For any $\Delta \leq 1/8$, $\alpha \geq 0$, $q \geq 1$, if $\Pr_{I(\Delta)}[\mathcal{E}(\alpha, \Delta^{-2}/(Kq))] \geq 1/2$, then

$$\Pr_{I(\Delta)}[\mathcal{E}^*(\alpha, \Delta^{-2}/(Kq))] \geq \Pr_{I(\Delta)}[\mathcal{E}(\alpha, \Delta^{-2}/(Kq))] - \delta(K, q)$$

Distribution Exchange Lemma. For any $\Delta \leq 1/8$, $\alpha \geq 0$, $q \geq 100$, $\zeta \geq 1$,

$$\Pr_{I(\Delta/\zeta)}[\mathcal{E}(\alpha + 1, \Delta^{-2}/(Kq) + \Delta^{-2}/\beta)]$$
$$\geq \Pr_{I(\Delta)}[\mathcal{E}^*(\alpha, \Delta^{-2}/(Kq))] - \delta'(K, q, \beta)$$

Put Together

Progress Lemma. For any $\Delta \leq 1/8$, $\alpha \geq 0$, $q \geq 1$, if $\Pr_{I(\Delta)}[\mathcal{E}(\alpha, \Delta^{-2}/(Kq))] \geq 1/2$, then

$$\Pr_{I(\Delta)}[\mathcal{E}^*(\alpha, \Delta^{-2}/(Kq))] \geq \Pr_{I(\Delta)}[\mathcal{E}(\alpha, \Delta^{-2}/(Kq))] - \delta(K, q)$$

Distribution Exchange Lemma. For any $\Delta \leq 1/8$, $\alpha \geq 0$, $q \geq 100$, $\zeta \geq 1$,

$$\Pr_{I(\Delta/\zeta)}[\mathcal{E}(\alpha + 1, \Delta^{-2}/(Kq) + \Delta^{-2}/\beta)] \geq \Pr_{I(\Delta)}[\mathcal{E}^*(\alpha, \Delta^{-2}/(Kq))] - \delta'(K, q, \beta)$$

Set $\zeta = \sqrt{1 + (Kq)/\beta}$ to connect the two lemmas: $\Delta^{-2}/(Kq) + \Delta^{-2}/\beta = (\Delta/\zeta)^{-2}/(Kq)$