PURE EXPLORATION FOR THE INFINITELY-ARMED BANDIT MODELS IN FIXED-CONFIDENCE AND FIXED-BUDGET SETTINGS

Maryam Aziz

Advisor: Javed Aslam

Committee: Emilie Kaufmann, Byron Wallace, Jonathan Ullman

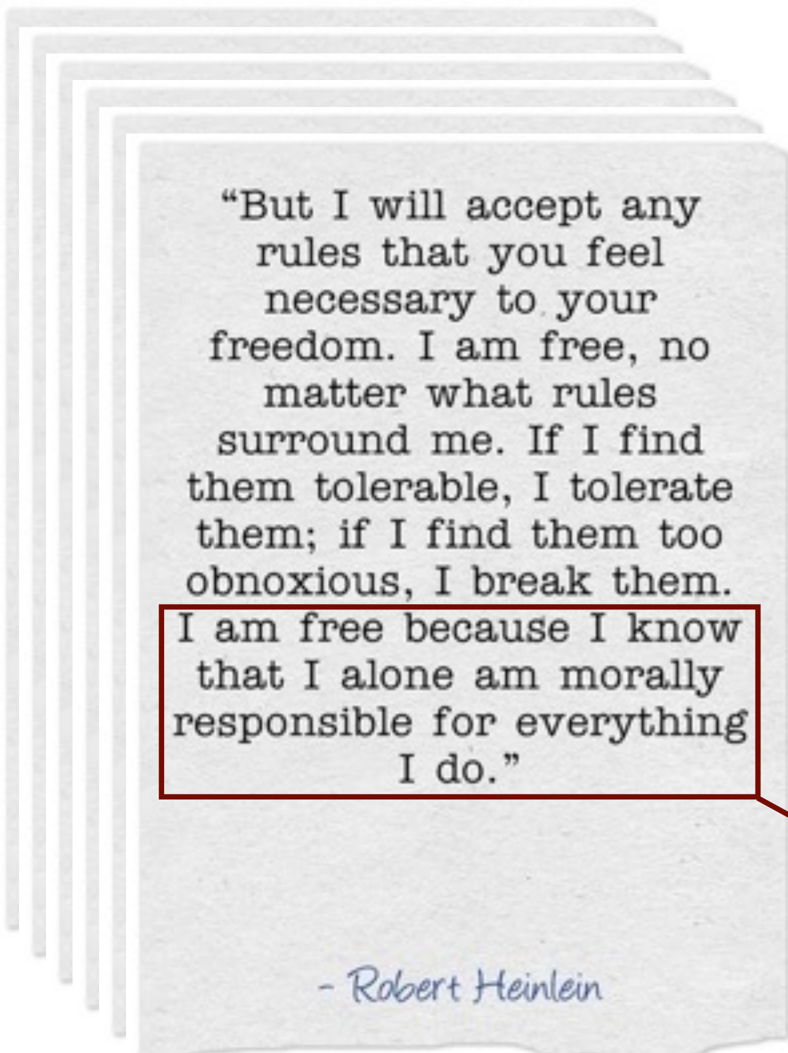
Drug Reservoir

Coin Reservoir

DRUG DISCOVERY

Given a finite number of tests and an effectively infinite number of drugs/chemical compounds, can I find an effective drug?

Skip-gram Reservoir



Coin Reservoir

(morally, free, responsible)

TEXT CLASSIFICATION

Given a finite number of tests and an effectively infinite number of skip-grams, can I find a skip-gram that accurately classifies whether a quote (document) belongs to Robert Heinlein?

CLASSIC MULTI-ARMED BANDIT SETTING

- finite number of one-arm slot machines
- fixed budget, e.g. time
- maximize your wins

e.g.Thompson, 1933; Wang et al., 2009; Bonald and Proutiere, 2013; David and Shimkin, 2014

VARIANTS OF THE MULTI-ARMED BANDIT PROBLEM

- cumulative regret vs. simple regret
- fixed budget vs. fixed confidence
- finite number of coins vs. infinite number of coins
- lower bounds vs. algorithms/upper bounds

Carpentier and Valko, 2015

VARIANTS OF THE MULTI-ARMED BANDIT PROBLEM

- cumulative regret vs. simple regret
- fixed budget vs. fixed confidence
- finite number of coins vs. infinite number of coins
- lower bounds vs. algorithms/upper bounds

Carpentier and Valko, 2015

OUR (α, δ)-FRAMEWORK

Where $\alpha \in (0,1)$ and $\delta \in (0,1)$ are parameters to model

α is the top fraction of coins

δ is the probability of the algorithm returning a "bad" coin

In the fixed confidence setting, fix α and δ and minimize T

In the fixed budget setting, fix T and minimize α and δ

OUR PROPOSED WORK

Pure Exploration for the Infinitely Armed Bandit Models		
	Lower Bound	**Algorithm/Upper Bound**
Fixed Confidence	Completed Work	Completed Work
Fixed Budget	Future Work	Current Work

Applications

OUR CONTRIBUTION IN RELATION TO THE FINITE MAB LITERATURE FOR PURE EXPLORATION

Finite Lower Bound $\xrightarrow{\text{Generalize}}$ Infinite Lower Bound

Finite Upper Bound $\xleftarrow{\text{Reduce}}$ Infinite Upper Bound

OUR PROPOSED WORK

Pure Exploration for the Infinitely Armed Bandit Models		
	Lower Bound	Algorithm/Upper Bound
Fixed Confidence	Completed Work	Completed Work
Fixed Budget	Future Work	Current Work

Applications

FINDING THE BEST COIN

Efficiently!

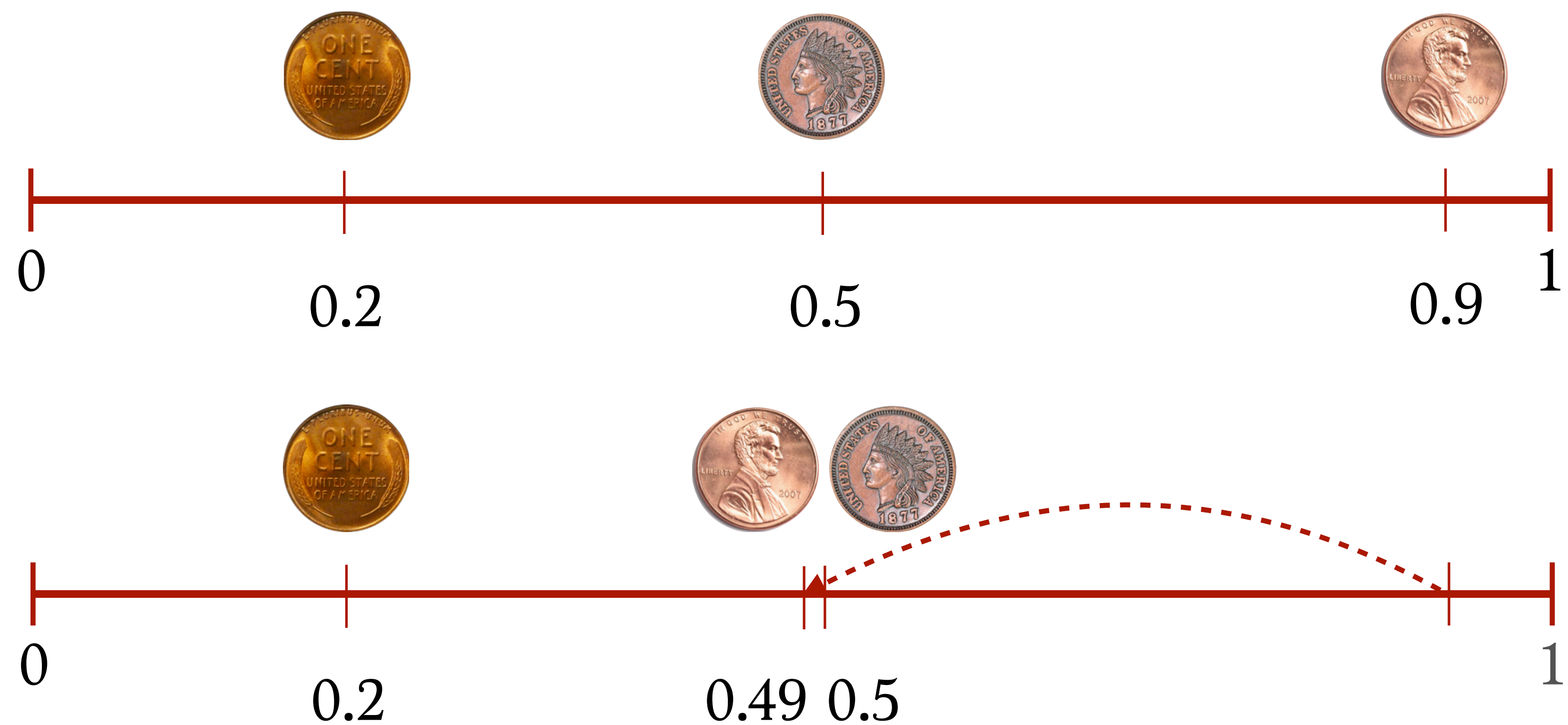
WHEN IS IT DIFFICULT?

Difficulty	Coin 1	Coin 2	Coin 3
Easy	0.9	0.4	0.1
Hard	0.9	0.89	0.1

CHANGE OF DISTRIBUTION FOR THREE COINS

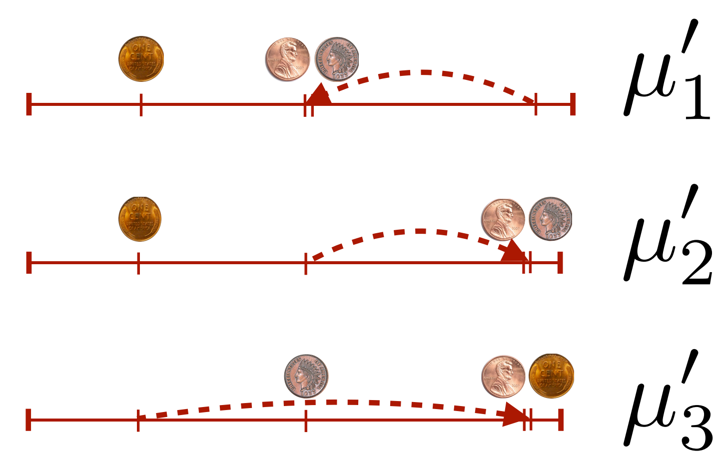
Input case 1:

Input case 2:



The lower bound on the problem is based on the *smallest distance* that changes the answer.

Complexity of the problem: $H := \sum_{a=1}^{3} \frac{1}{kl(\mu_a|\mu'_a)}$



e.g. Mannor, S. et. al. 2004; Kaufmann et al., 2016

DRUG DISCOVERY LOWER BOUND IN THE FIXED CONFIDENCE SETTING

How many tests would any agent need to find an effective drug with high confidence?

↓

How many total coin flips would any algorithm need to find a coin whose bias is in top-α fraction with probability $1-\delta$?

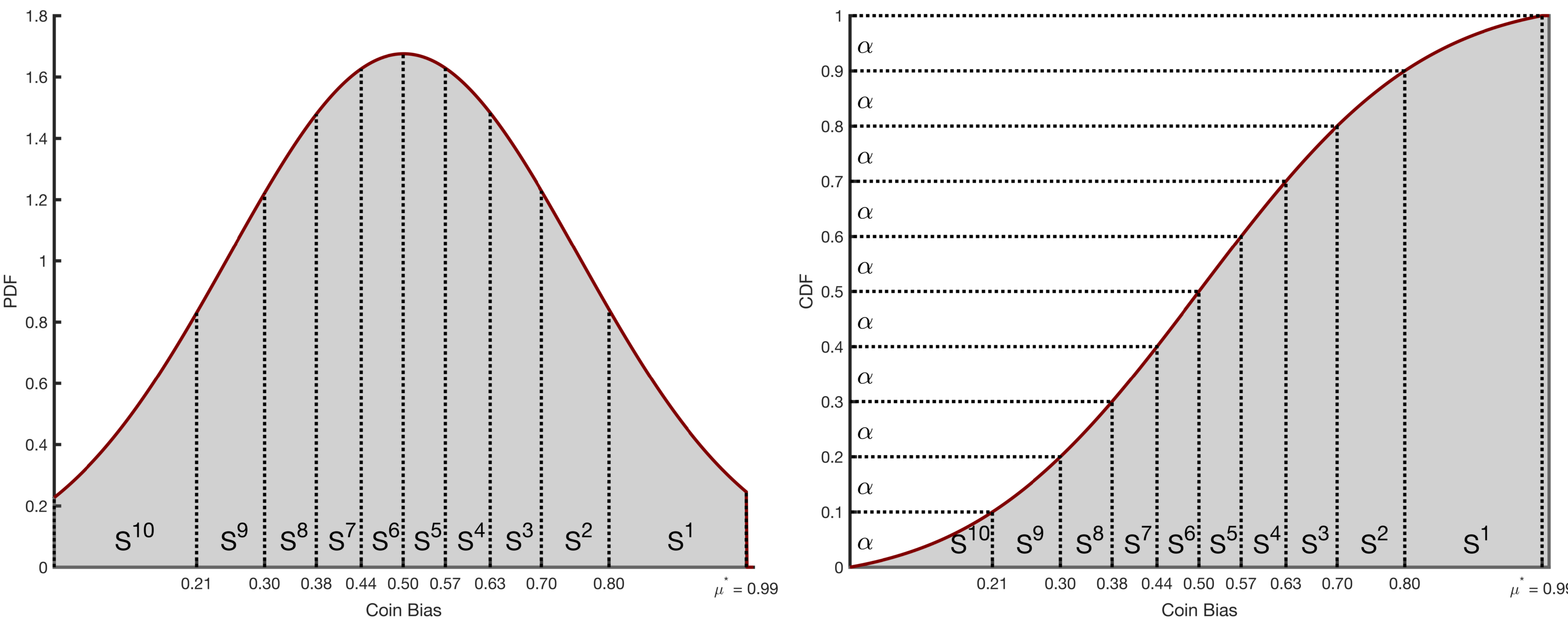
OUR CHANGE OF DISTRIBUTION FOR MANY COINS, USING SUBSET OF COINS

Coin Reservoir

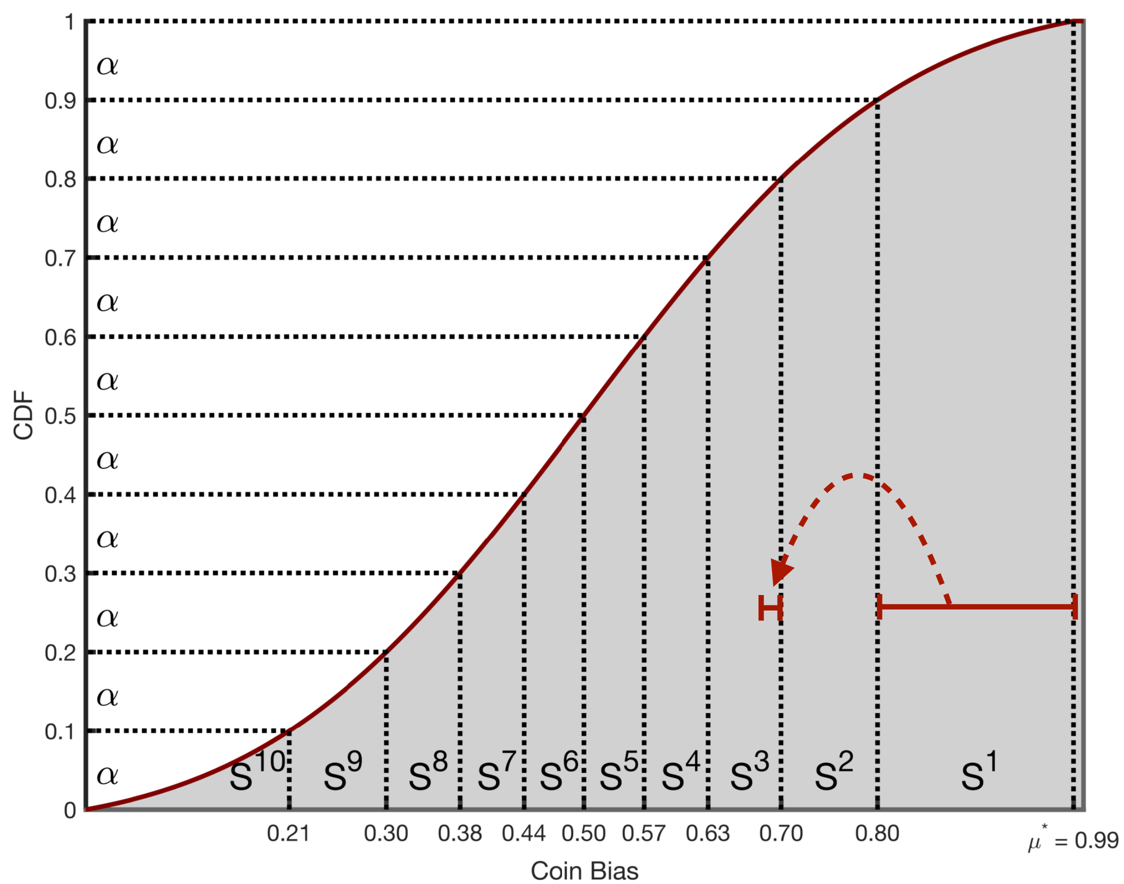
The lower bound on the problem is based on the *smallest distance* that changes the answer.

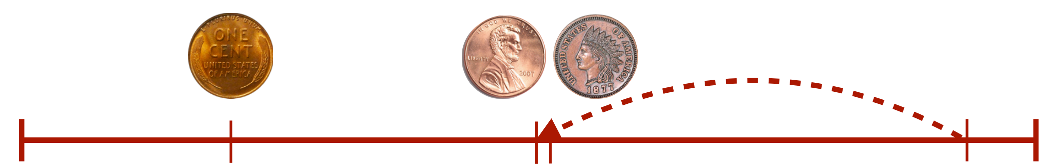
RESERVOIR SUBSETS WHEN $\alpha = 0.1$

Reservoir Distribution



OUR LOWER BOUND IS BASED ON CHANGE OF DISTRIBUTION





Form of our Lower Bound: $E[T] \geq H * \log(\frac{1}{\delta})$

Our lower bound is based on the *smallest distance* that changes the answer.

OUR PROPOSED WORK

Pure Exploration for the Infinitely Armed Bandit Models		
	Lower Bound	**Algorithm/Upper Bound**
Fixed Confidence	Completed Work	Completed Work
Fixed Budget	Future Work	Current Work

Applications

DRUG DISCOVERY UPPER BOUND IN THE FIXED CONFIDENCE SETTING

Can we build an agent that needs as few tests as possible (matching our lower bound) to find an effective drug with high confidence?

↓

Can we build an algorithm that needs as few total coin flips as possible (matching our lower bound) to find a coin whose bias is in top-α fraction with probability $1-\delta$?

α-KL-LUCB (OUR PROPOSED ALGORITHM)

- a two-phase algorithm

  1. draw n coins s.t. $n = \frac{1}{\alpha} \log(\frac{1}{\delta})$
  2. run KL-LUCB [*], an existing algorithm, as sub-routine

- an upper bound on sample complexity which is within log factor of our lower bound: $E[T] \leq H * (\log(\frac{1}{\delta}))^2$

Recall our lower bound: $E[T] \geq H * \log(\frac{1}{\delta})$

[*] Kaufmann, E. and Kalyanakrishnan, S. 2013

α-KL-LUCB RESULTS

Table 1: Performance of α-KL-LUCB.
(Mean of 100 runs.)

Reservoir	α	δ	% Success	Regret	T
Beta(1,1)	0.05	0.10	99	0.010	43K
Beta(1,1)	0.10	0.05	100	0.012	12K
Beta(1,1)	0.10	0.10	99	0.018	11K
Beta(1,10)	0.05	0.10	99	0.557	32K
Beta(1,10)	0.10	0.05	100	0.589	25K
Beta(1,10)	0.10	0.10	98	0.597	20K
Beta(1,2)	0.05	0.10	99	0.052	9K
Beta(1,2)	0.10	0.05	100	0.080	6K
Beta(1,2)	0.10	0.10	100	0.079	5K
Beta(1,3)	0.05	0.10	99	0.139	10K
Beta(1,3)	0.10	0.05	100	0.192	8K
Beta(1,3)	0.10	0.10	99	0.189	6K
Beta(2,1)	0.10	0.05	100	0.008	44K
Beta(2,1)	0.10	0.10	100	0.009	39K

OUR PROPOSED WORK

Pure Exploration for the Infinitely Armed Bandit Models		
	Lower Bound	**Algorithm/Upper Bound**
Fixed Confidence	Completed Work	Completed Work
Fixed Budget	Future Work	Current Work

Applications

DRUG DISCOVERY GOAL IN THE FIXED BUDGET SETTING

Given a fixed number of trials (budget) what's the most effective drug one can find with as high confidence as possible?

↓

Given a fixed number of total coin flips $\boldsymbol{T}$ what's the smallest top-$\boldsymbol{\alpha}$ and the smallest probability of error $\boldsymbol{\delta}$ one can guarantee?

SEQUENTIAL-HALVING FOR THE FINITE CASE

Total flips: 240

flips each 10 times, 80 flips

flips each 20 times, 80 flips

flips each 40 times, 80 flips

best coin

Karnin, Z., Koren, T., and Somekh, O. 2013.

SEQUENTIAL-HALVING FOR THE INFINITE CASE

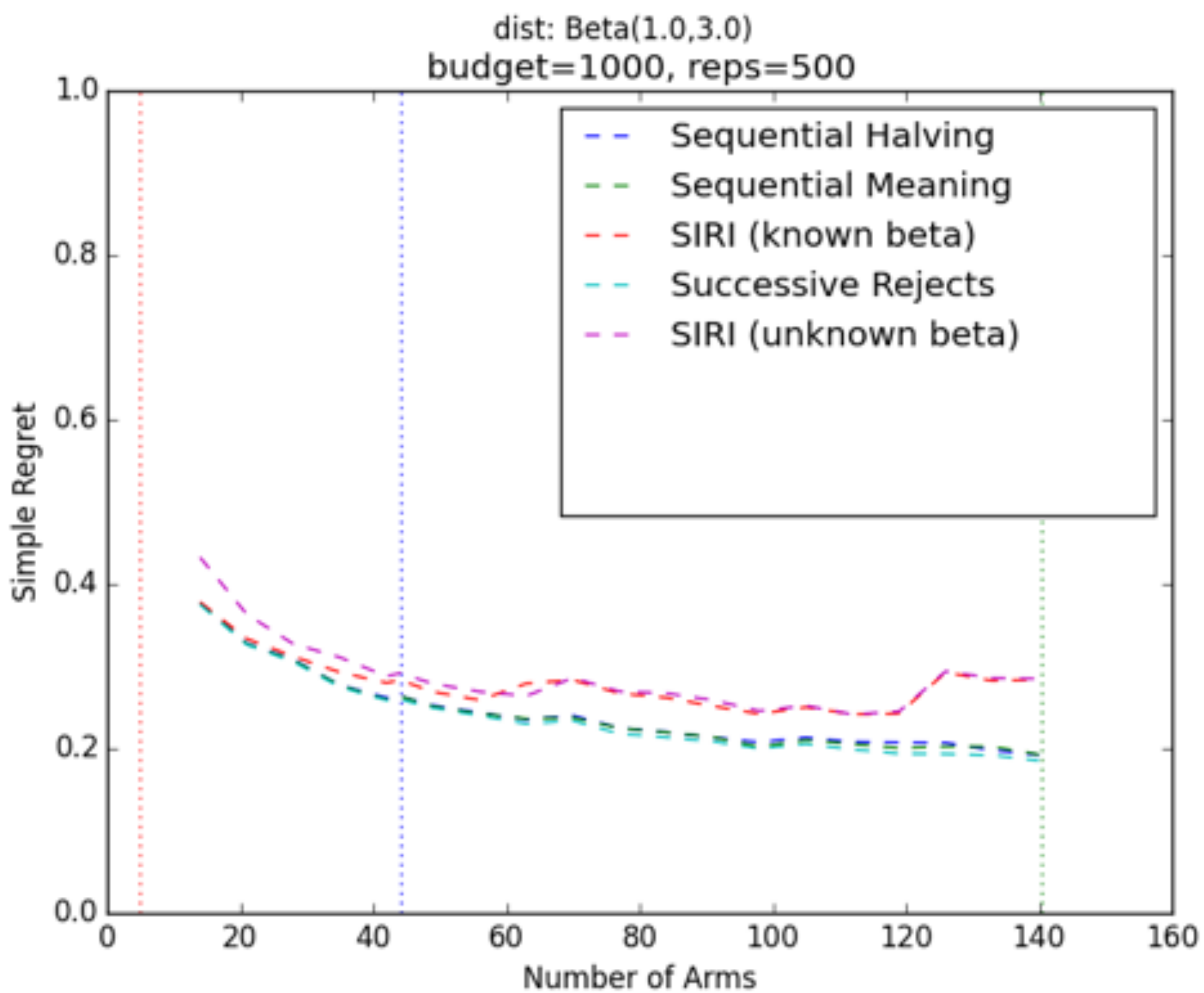
- a two-phase algorithm

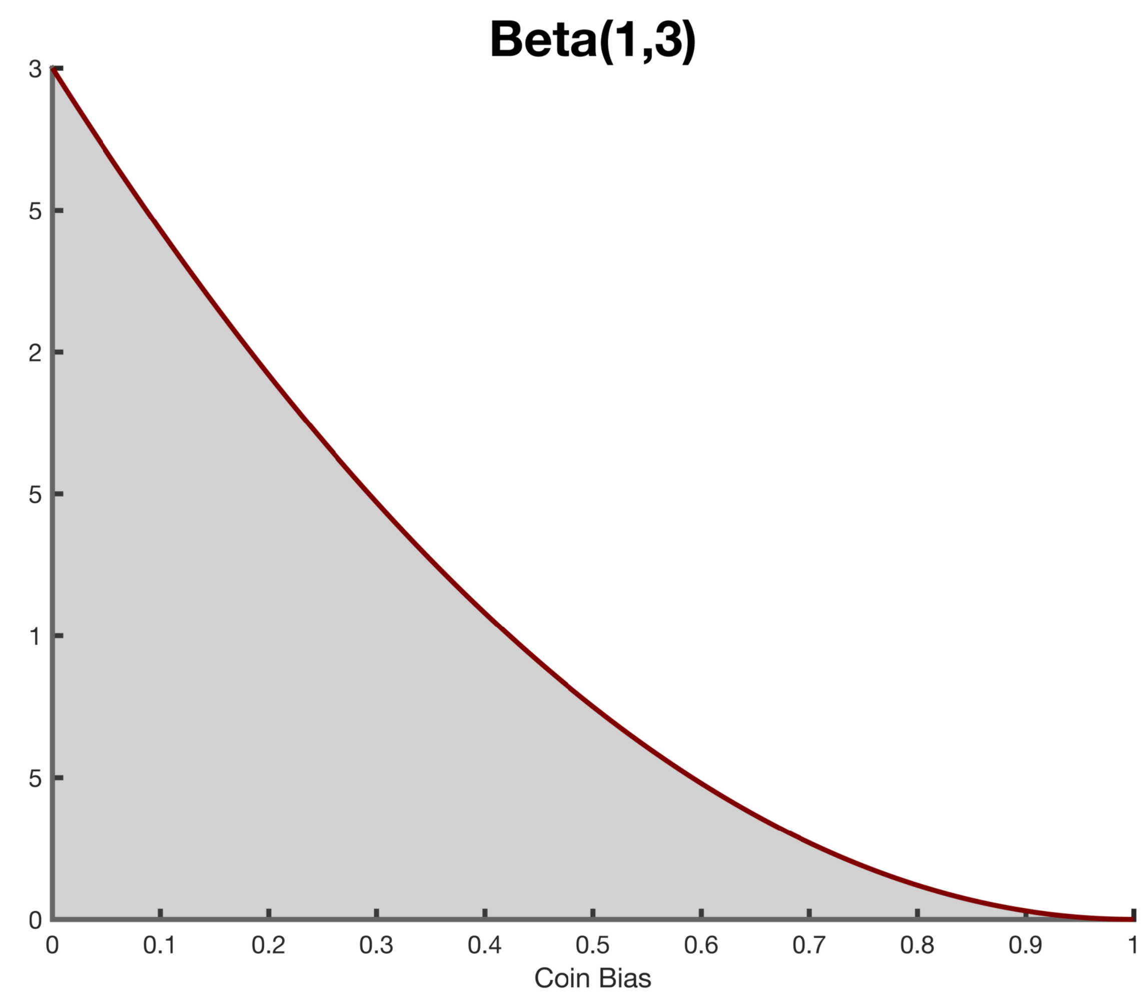
  1. draw n coins s.t. budget $T = n\log(n)$
  2. run Sequential Halving as sub-routine

- an upper bound on simple regret

[*] Karnin, Z., Koren, T., and Somekh, O. 2013.

BUDGET ALGORITHM RESULTS FOR RESERVOIR BETA(1,3)



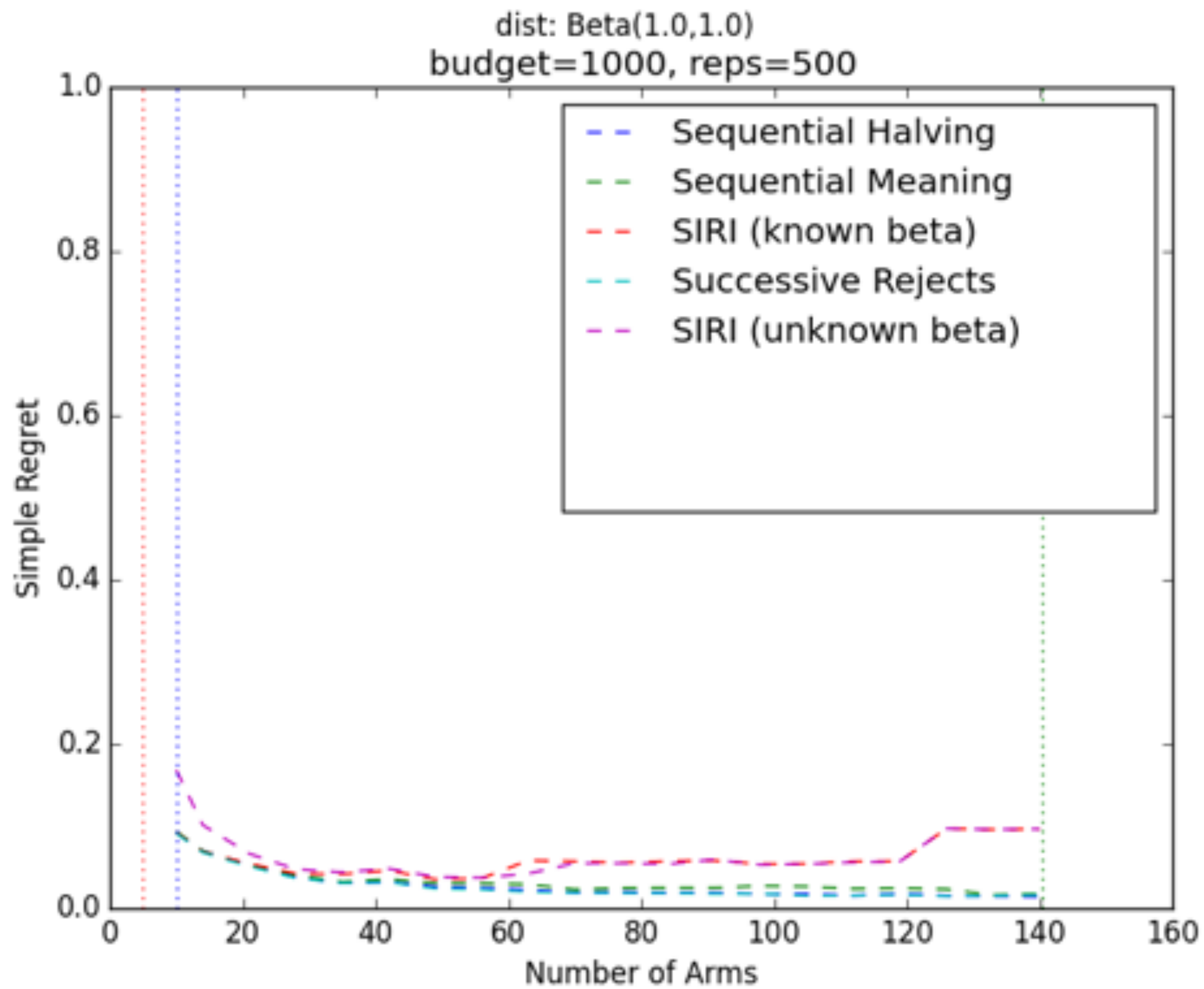


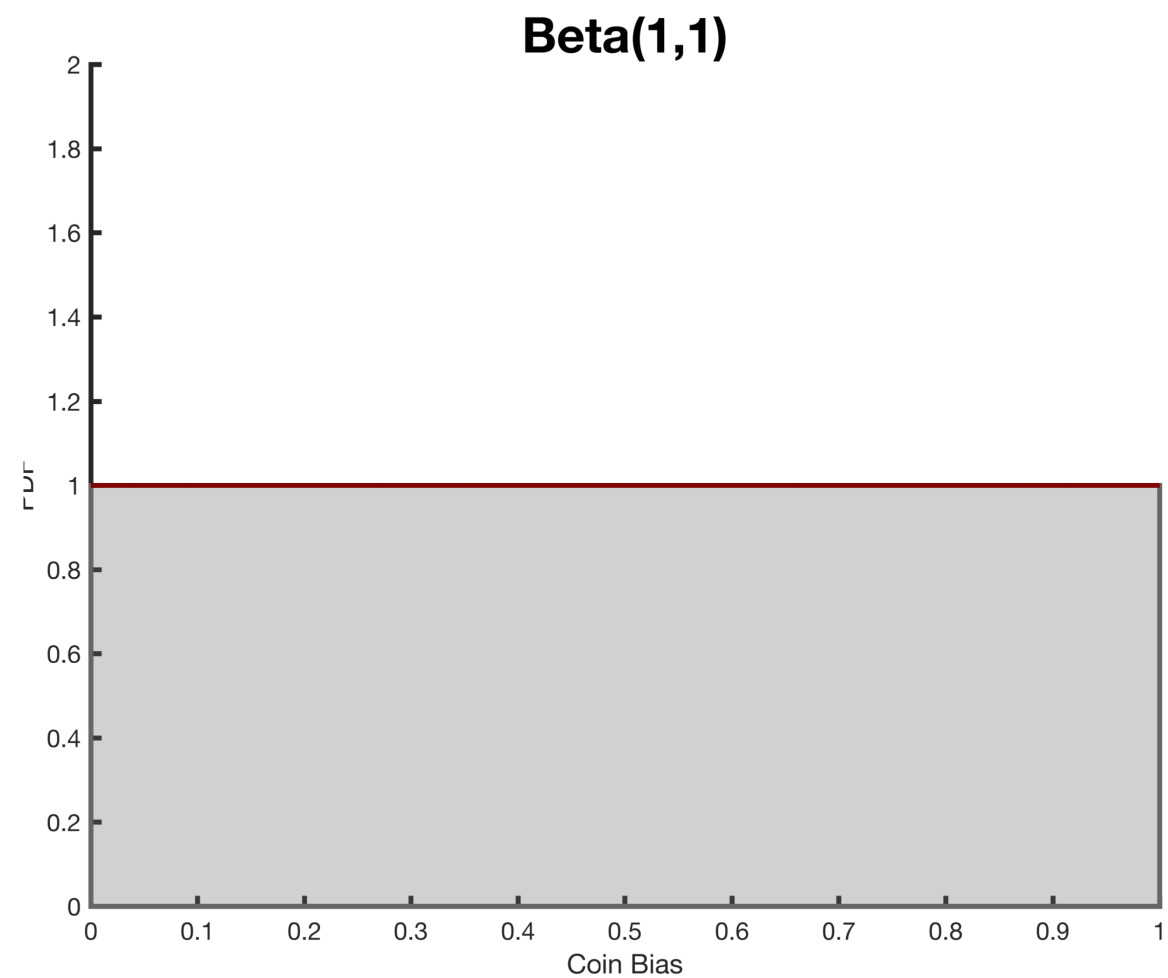
Optimal number of coins for SH: $n = 140$

Original number of coins for SIRI (known beta): $n \simeq 45$

Carpentier, A. and Valko, M. 2015

BUDGET ALGORITHM RESULTS FOR RESERVOIR BETA(1,1)





Optimal number of coins for SH: $n = 140$

Original number of coins for SIRI (known beta): $n \simeq 10$

OUR PROPOSED WORK IN THE FIXED CONFIDENCE SETTING

Pure Exploration for the Infinitely Armed Bandit Models		
	Lower Bound	**Algorithm/Upper Bound**
Fixed Confidence	Completed Work	Completed Work
Fixed Budget	Future Work	Current Work

A fixed confidence algorithm for the infinitely armed bandit problem that matches our lower bound (ideally a generalized algorithm that also works for the finitely armed bandit model), or an improved the lower bound. **Proposed. Finish by May 2017.**

OUR PROPOSED WORK IN THE FIXED BUDGET SETTING

Pure Exploration for the Infinitely Armed Bandit Models		
	Lower Bound	**Algorithm/Upper Bound**
Fixed Confidence	Completed Work	Completed Work
Fixed Budget	Future Work	Current Work

A fixed budget top-α arm identification algorithm based on the Sequential Halving algorithm and a simple regret upper bound analysis of it. **Proposed. Finish by May 2017.**

Ideally, a fixed budget simple regret lower bound. **Proposed. Finish by May 2017.**

OUR PROPOSED WORK FOR APPLICATIONS

Pure Exploration for the Infinitely Armed Bandit Model		
	Lower Bound	**Algorithm/Upper Bound**
Fixed Confidence	Completed Work	Completed Work
Fixed Budget	Future Work	Current Work

Applications

ML, e.g., Boosting, applications using our (α, δ)-framework. **Proposed. Resume by April 2017. Finish by August 2017.**

One or more applications modeled using our (α, δ)-framework. **Proposed. Finish by January 2018.**

OUR PROPOSED WORK FOR APPLICATIONS

The dose-finding application. **Finish by April 2017.**

Thesis defense. **By March 2018.**

MORE DETAILS

You can find the proposal document here:

https://mazesweb.wordpress.com/proposal/