Cost-Aware Bayesian Optimization with Adaptive Stopping via the Pandora's Box Gittins Index

Qian Xie 谢倩 (Cornell ORIE)

Seminar @CityUHK DAO

Motivation: World of Optimization under Uncertainty

≠ parameters
(weights, biases)

ML model training:

Training hyperparameters ⟶
(e.g., learning rate, # layers)

⟶ Accuracy

Goal: find hyperparameters that achieve high accuracy

Motivation: World of Optimization under Uncertainty

ML model training:

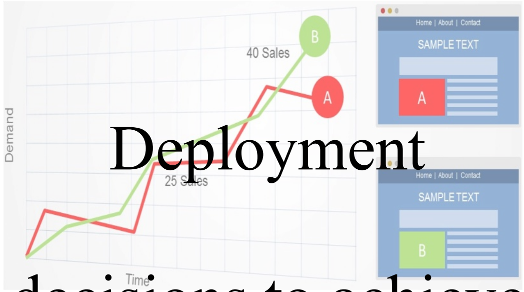
≠ parameters
(weights, biases)

Training hyperparameters (e.g., learning rate, # layers) ⟶ Training ⟶ Accuracy

Goal: find hyperparameters that achieve high accuracy

Adaptive experimentation:

Decision/design variables (e.g., layout, pricing level) ⟶ Deployment ⟶ Revenue



Goal: make decisions to achieve high revenue

Motivation: World of Optimization under Uncertainty

Design choices x ⟶  ⟶ Performance metric $f(x)$

non-analytical &
no gradient info

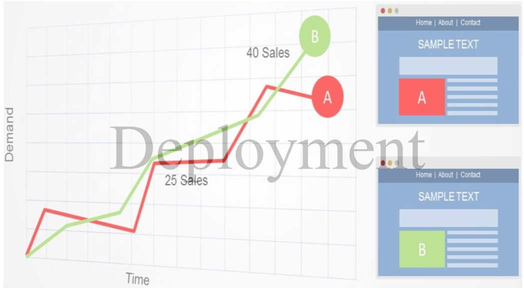
Goal: $\max_{x \in \mathcal{X}} f(x)$

ML model training:

Training hyperparameters
(e.g., learning rate, # layers) ⟶ 

⟶ Accuracy

Adaptive experimentation:

Decision/design variables
(e.g., layout, pricing level) ⟶ 

⟶ Revenue

Motivation: World of Optimization under Uncertainty

Black-box optimization:

(gradient-based methods not applicable)

Input x →  → Observed outcome $f(x)$

non-analytical & no gradient info

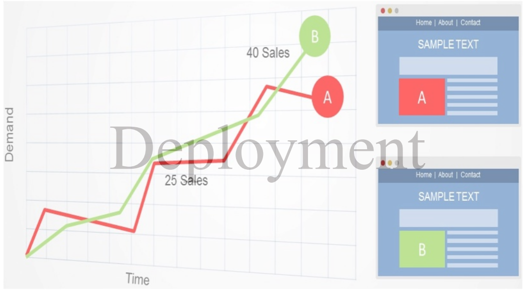
Goal: $\max_{x \in \mathcal{X}} f(x)$

ML model training:

Training hyperparameters (e.g., learning rate, # layers) → 

→ Accuracy

Adaptive experimentation:

Decision/design variables (e.g., layout, pricing level) → 

→ Revenue

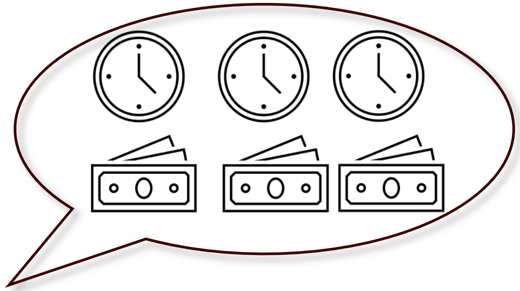
Background: Black-Box Optimization

Black-box optimization:

Input x ⟶  ⟶ Observed outcome $f(x)$

expensive-to-evaluate

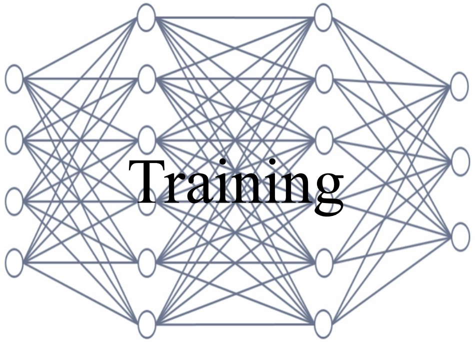
Goal: $\max_{x \in \mathcal{X}} f(x)$

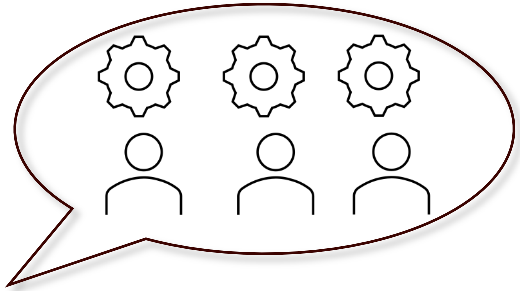


Training time

Compute credits

ML model training:

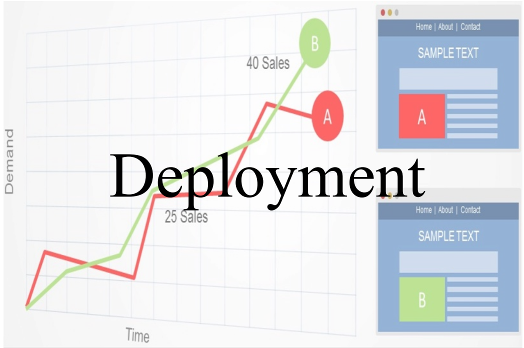
Training hyperparameters (e.g., learning rate, # layers) ⟶  ⟶ Accuracy



Operational cost

User experience

Adaptive experimentation:

Decision/design variables (e.g., layout, pricing level) ⟶  ⟶ Revenue

Background: Black-Box Optimization

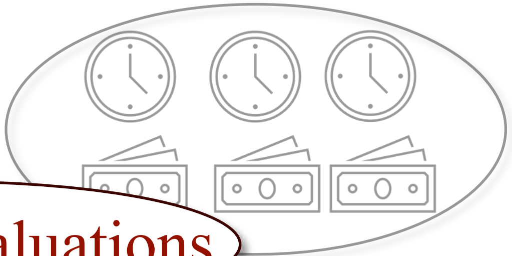
Black-box optimization:

Input x ⟶  ⟶ Observed outcome $f(x)$

expensive-to-evaluate

Goal: $\mathbb{E}\ \max_{t=1,2,\ldots,T} f(x_t)$

limited #evaluations



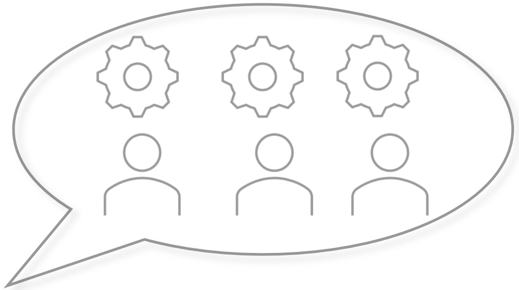
Training time

Compute credits

ML model training:

Training hyperparameters (e.g., learning rate, # layers) ⟶ 

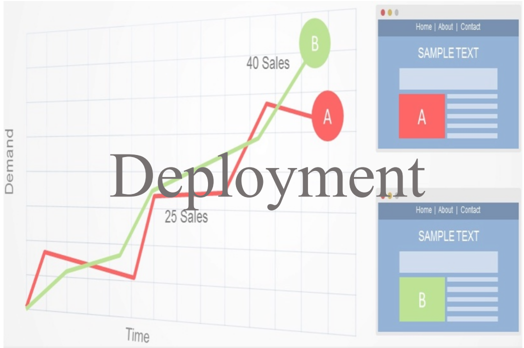
⟶ Accuracy



Operational cost

User experience

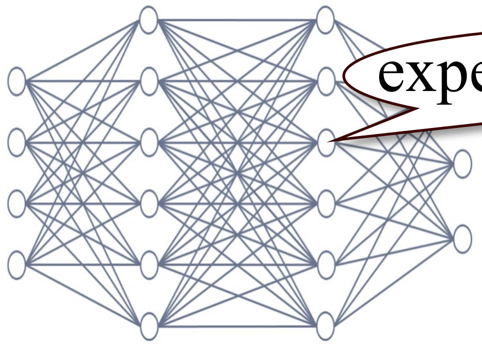
Adaptive experimentation:

Decision/design variables (e.g., layout, pricing level) ⟶ 

⟶ Revenue

Naïve Non-Adaptive Approach: Grid Search

ML model training:

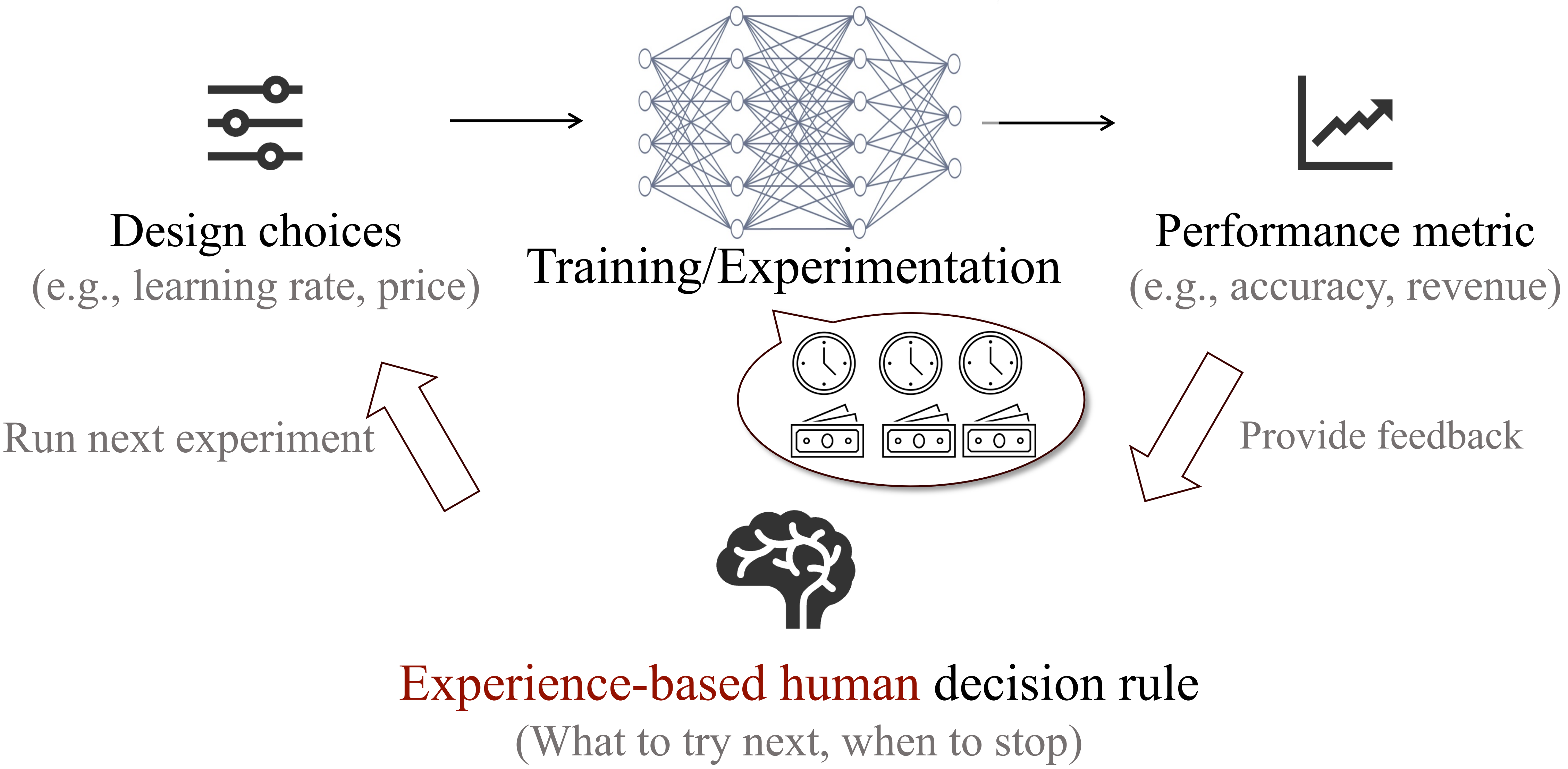
Training hyperparameters →  → Accuracy

Goal: find hyperparameters that achieve high accuracy

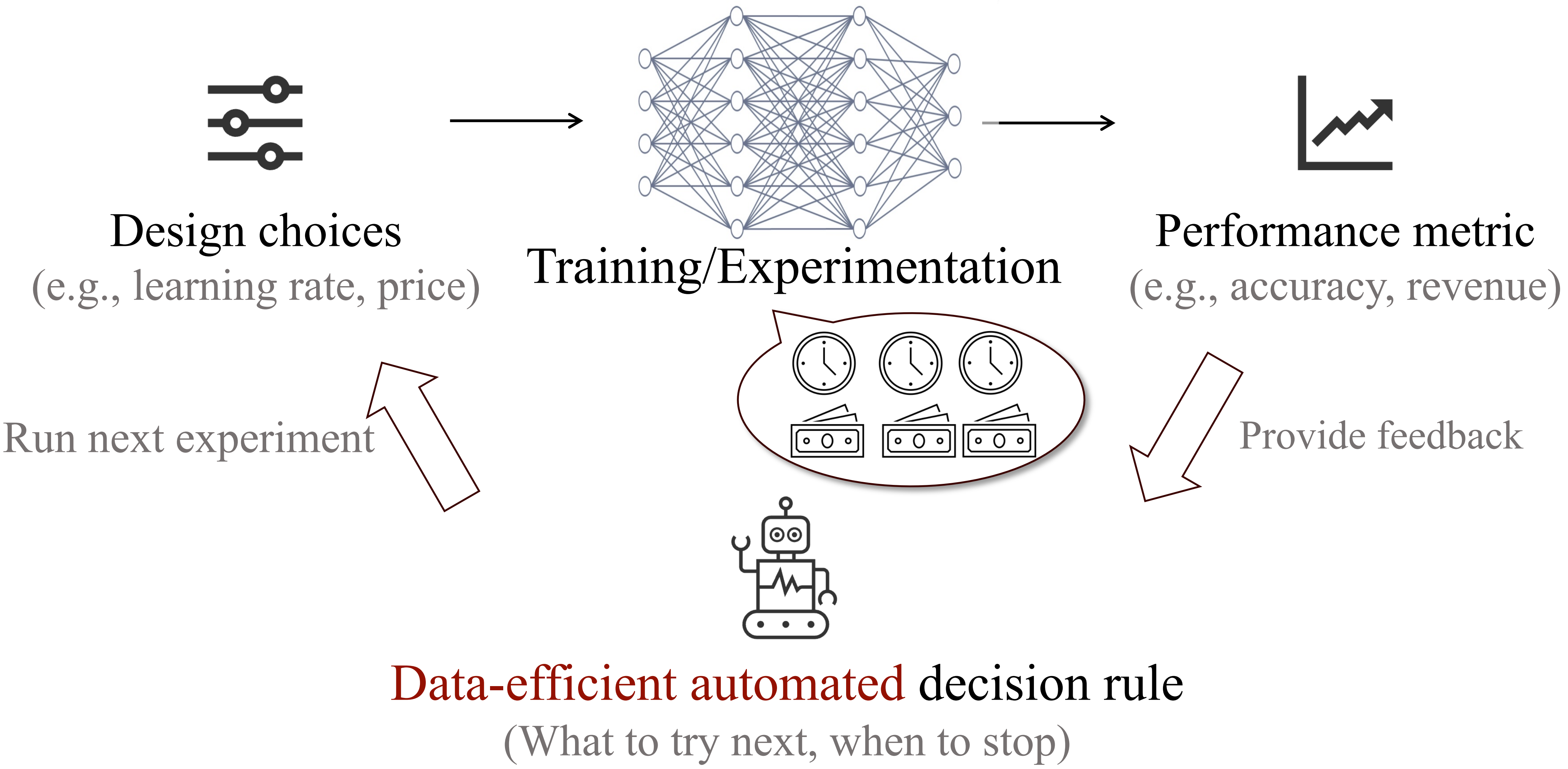
Training hyperparameter	Range	Number of Options
Batch size	[16, 512]	10
Learning rate	[1e-4, 1e-1]	10
Momentum	[0.1, 0.99]	10
Weight decay	[1e-5, 1e-1]	10
Number of layers	{1, 2, 3, 4}	4
Max units per layer	[64, 1024]	10
Dropout	[0.0, 1.0]	10

40,000,000 combinations!

Naïve Adaptive Approach: Manual Tuning



Data-Driven Adaptive Approaches



Existing Umbrellas of Black-Box Optimization

Naïve approaches:	Data-driven approaches:
• Grid search	• Local search
• Random search	• Evolutionary algorithms
• Manual tuning	• Bayesian optimization
	• Reinforcement learning
	• LLM agent

New Methods for Black-Box Optimization

Naïve approaches:	Data-driven approaches:
• Grid search	• Local search
• Random search	• Evolutionary algorithms
• Manual tuning	• Bayesian optimization ★
	• Reinforcement learning ★
	• LLM agent ★

Contributions of new methods proposed in my work:

1. Novel connection to related decision problems

2. Principled decision rules

3. Competitive empirical performance

★ New methods under this umbrella

New Methods for Black-Box Optimization

Naïve approaches:	Data-driven approaches:
• Grid search	• Local search
• Random search	• Evolutionary algorithms
• Manual tuning	• Bayesian optimization ★
	• Reinforcement learning ★
	• LLM agent ★

This talk's focus

Contributions of new methods proposed in my work:

1. Novel connection to related decision problems

2. Principled decision rules

3. Competitive empirical performance

★ New methods under this umbrella

Bayesian Optimization

$f \sim$ Stochastic Process

$x \longrightarrow$  $\longrightarrow f(x)$

Black-box function

Goal: $\mathbb{E} \max_{t=1,2,\ldots,T} f(x_t)$

Bayesian Optimization

Time 0

$x \longrightarrow$

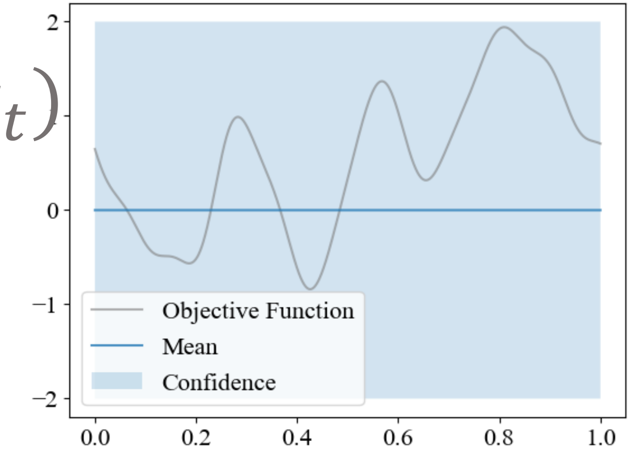
$\longrightarrow f(x)$

$f \sim$Stochastic Process

Model belief

Black-box function

Goal: $\mathbb{E} \max\limits_{t=1,2,\dots,T} f(x_t)$



Probabilistic model

(e.g., Gaussian process)

Bayesian Optimization

Time t

$x_1, \ldots, x_t \longrightarrow$

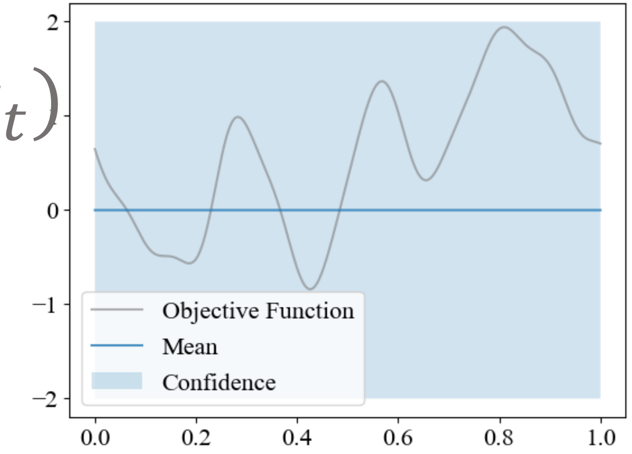
$\longrightarrow f(x_1), \ldots, f(x_t)$

$f \sim$Stochastic Process

Model belief

Black-box function

Goal: $\mathbb{E} \max_{t=1,2,\ldots,T} f(x_t)$



Probabilistic model

(e.g., Gaussian process)

Bayesian Optimization

Time t

$x_1, \dots, x_t \longrightarrow$

Black-box function

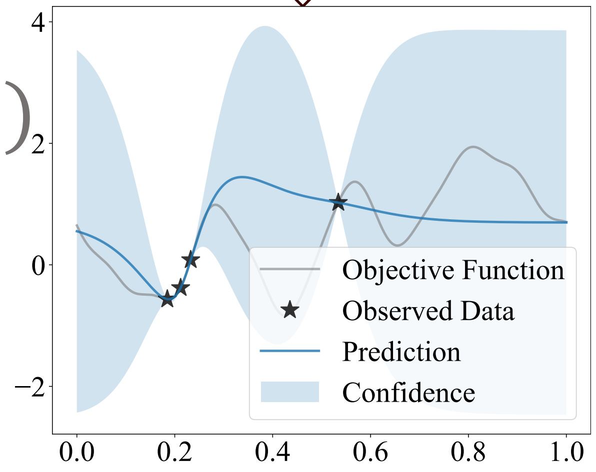
$\longrightarrow f(x_1), \dots, f(x_t)$

$f \sim$Stochastic Process

Update belief

(Bayes' rule)

Goal: $\mathbb{E} \max_{t=1,2,\dots,T} f(x_t)$



Probabilistic model

(e.g., Gaussian process)

Bayesian Optimization

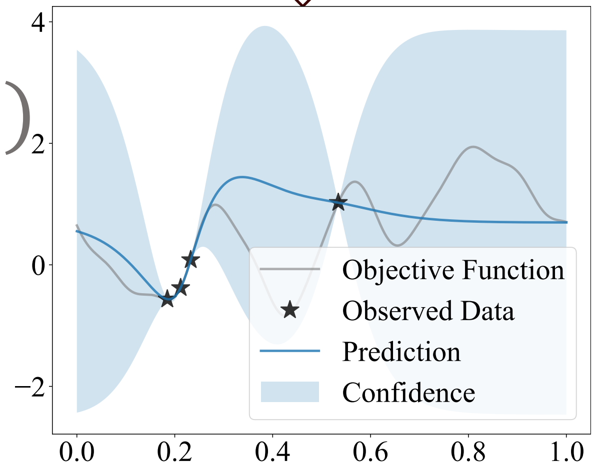
Time t

$x_1, \dots, x_t \longrightarrow$

$\longrightarrow f(x_1), \dots, f(x_t)$

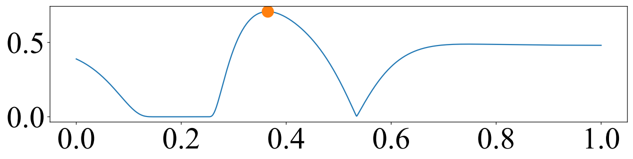
Black-box function

Update belief
(Bayes' rule)



Probabilistic model
(e.g., Gaussian process)

Goal: $\mathbb{E} \max_{t=1,2,\dots,T} f(x_t)$



scoring (worth of each point

Decision rule
(What to evaluate next)

Bayesian Optimization

Time t

$x_1, \dots, x_t$

$f(x_1), \dots, f(x_t)$

Decide next input

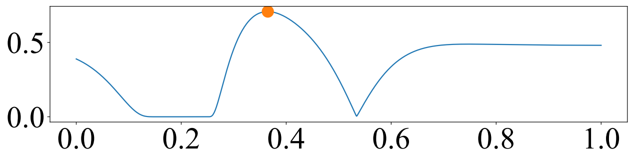
(gradient-based optimization)

Black-box function

Update belief

(Bayes' rule)

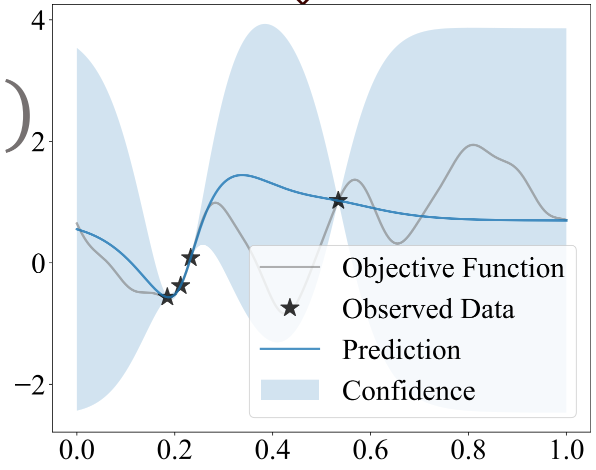
Goal: $\mathbb{E} \max_{t=1,2,\dots,T} f(x_t)$



Decision rule

(What to evaluate next)

scoring (worth of each point



Probabilistic model

(e.g., Gaussian process)

Bayesian Optimization

Time $t + 1$

$x_1, \dots, x_{t+1} \longrightarrow$

$\longrightarrow f(x_1), \dots, f(x_{t+1})$

Decide next input

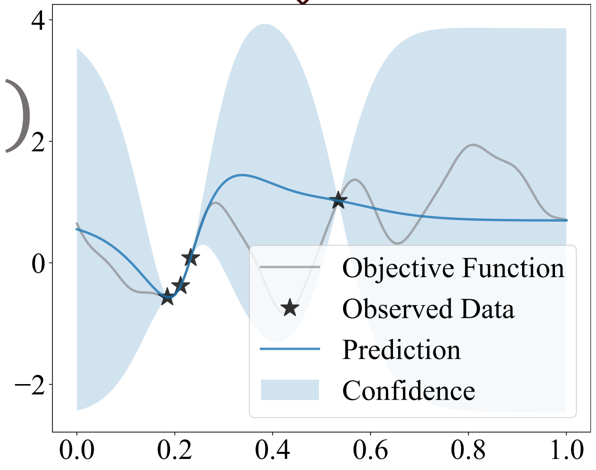
(gradient-based optimization)

Update belief

(Bayes' rule)

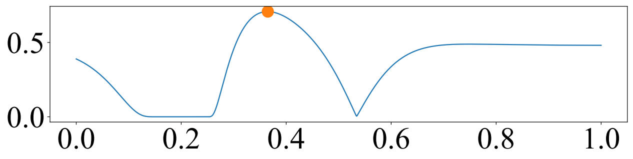
Black-box function

Goal: $\mathbb{E} \max_{t=1,2,\dots,T} f(x_t)$



Probabilistic model

(e.g., Gaussian process)

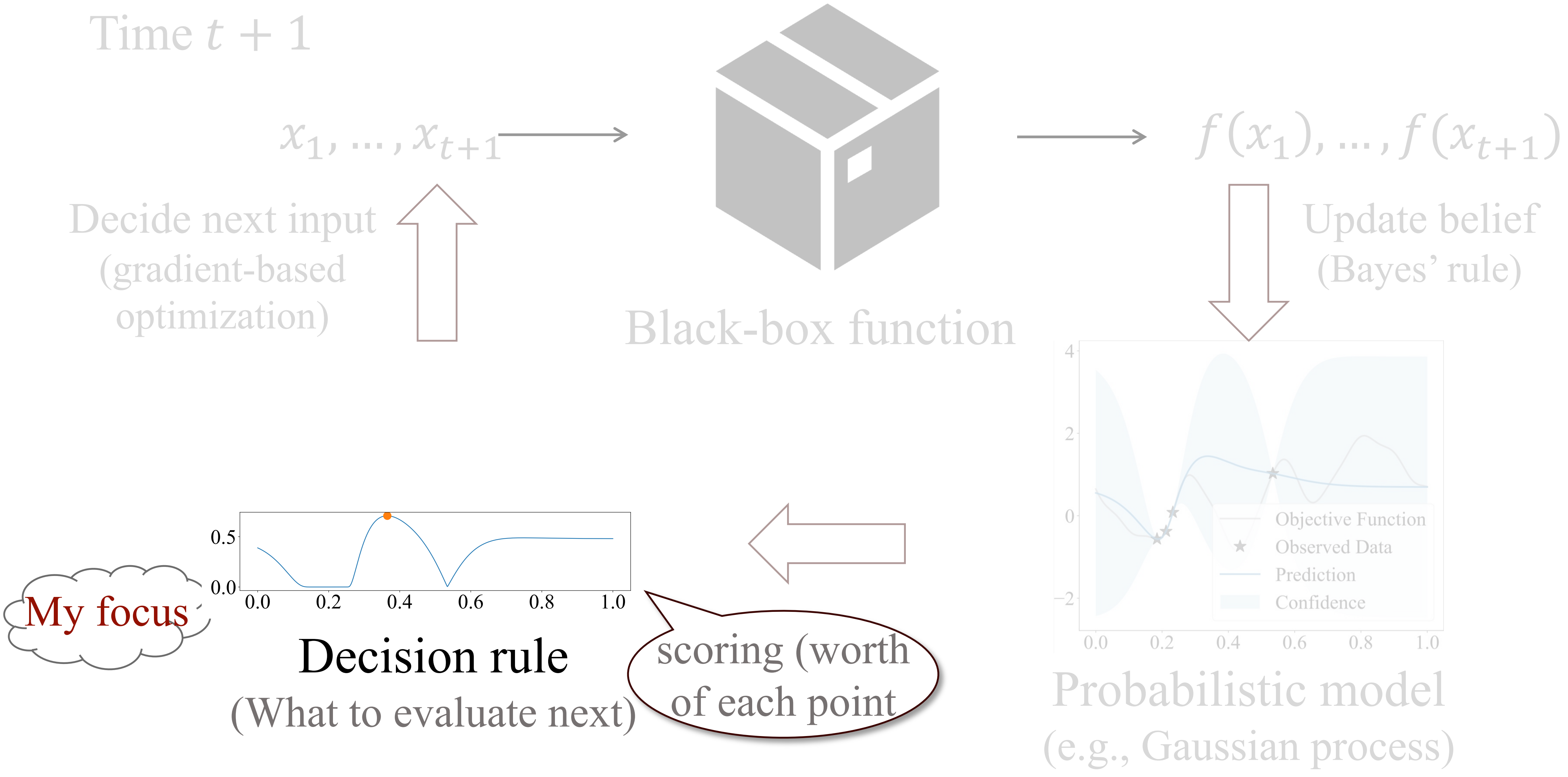


Decision rule

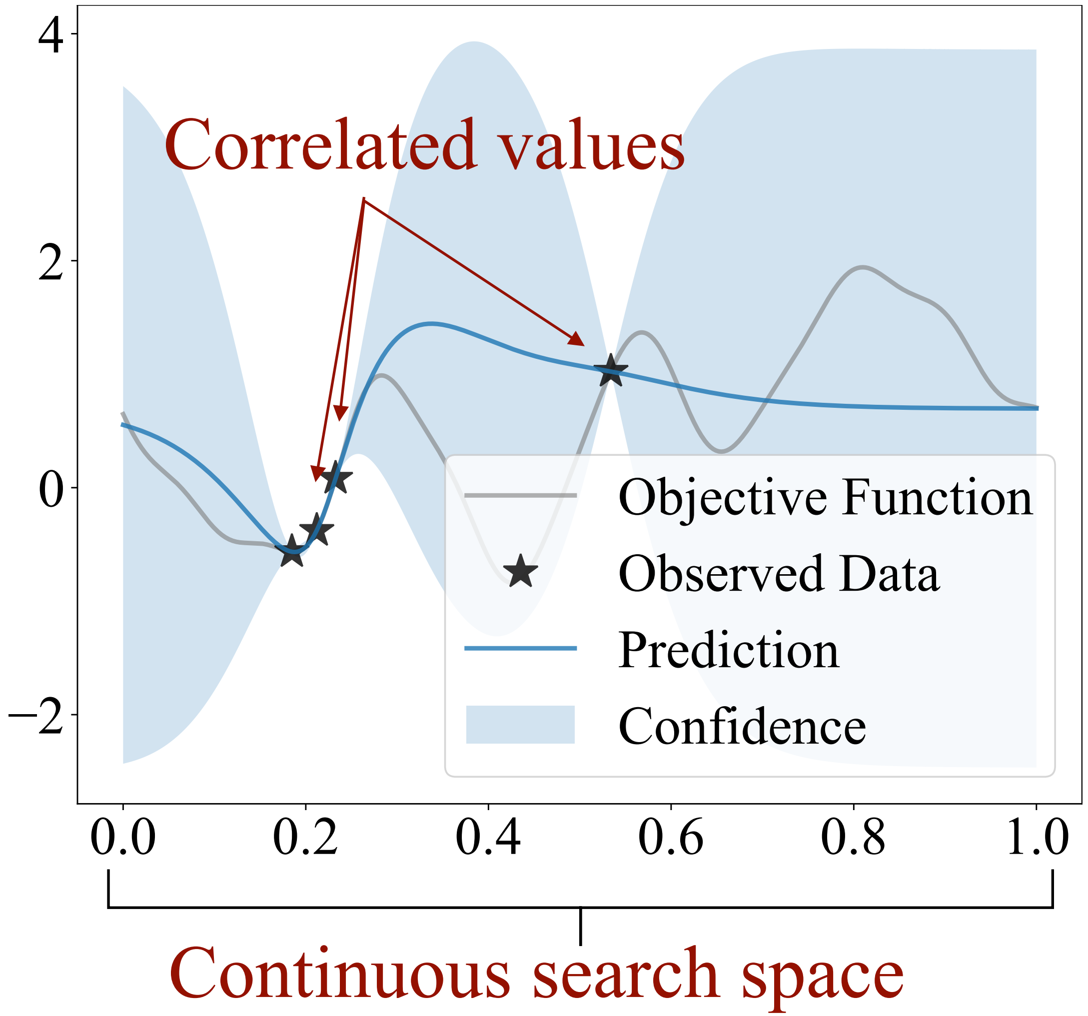
(What to evaluate next)

scoring (worth of each point

Bayesian Optimization



Challenges in Decision Rule Design



Correlation & continuity $\Longrightarrow$ Intractable MDP $\Longrightarrow$ Optimal policy unknown

Popular Decision Rule: Expected Improvement

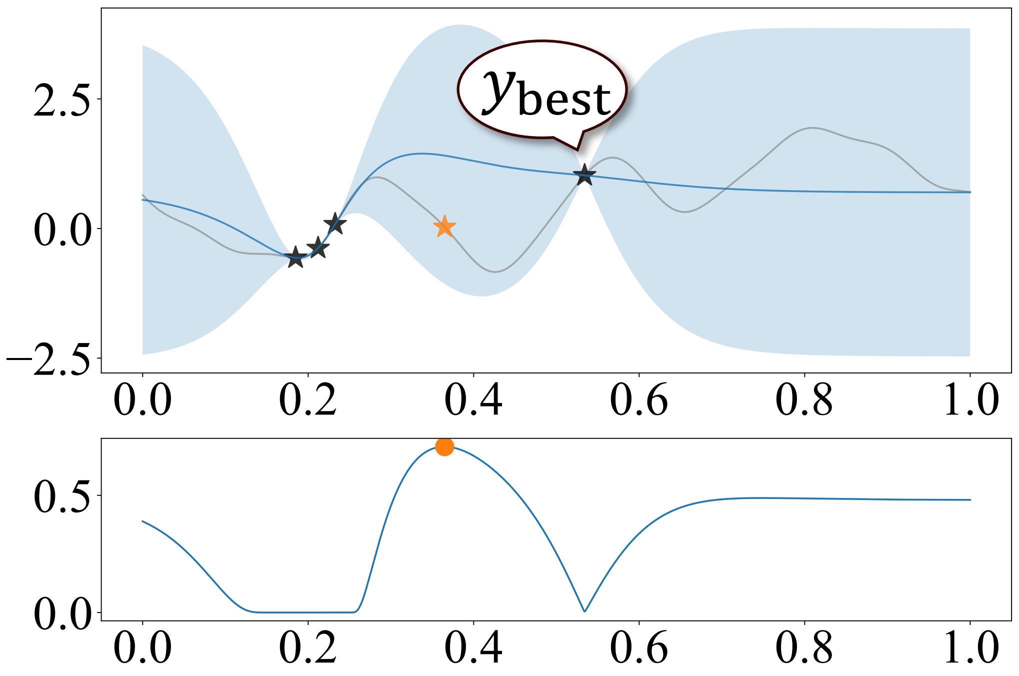
current best observed

data D

$$\mathrm{EI}(x) = \mathbb{E}[\underbrace{\max(f(x) - y_{\text{best}}, 0)}_{\text{“improvement”}} \mid x_1, \ldots, x_t]$$

$$x_{t+1} = \max_x \mathrm{EI}_{f|D}(x)$$

posterior distribution



One-step approximation to MDP

Expected improvement $\mathrm{EI}(x)$

Popular Decision Rule: Expected Improvement

$$\mathrm{EI}(x) = \mathbb{E}[\underbrace{\max(f(x) - y_{\text{best}}, 0)}_{\text{“improvement”}} \mid x_1, \ldots, x_t]$$

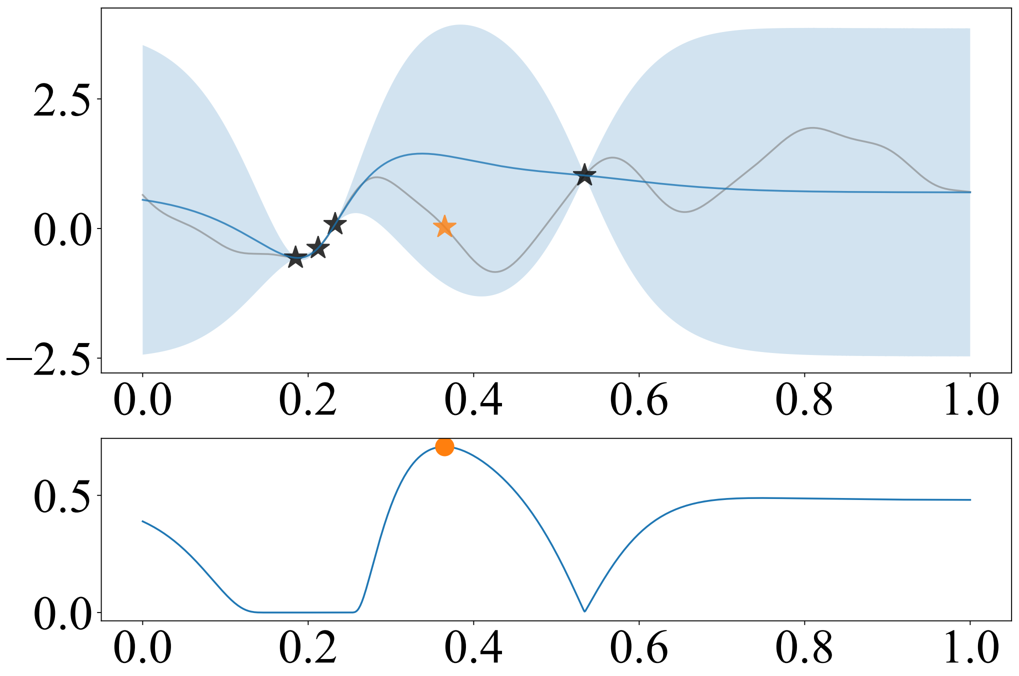
current best observed

data D

$$x_{t+1} = \max_x \mathrm{EI}_{f|D}(x; y_{\text{best}})$$

posterior distribution

One-step approximation to MDP

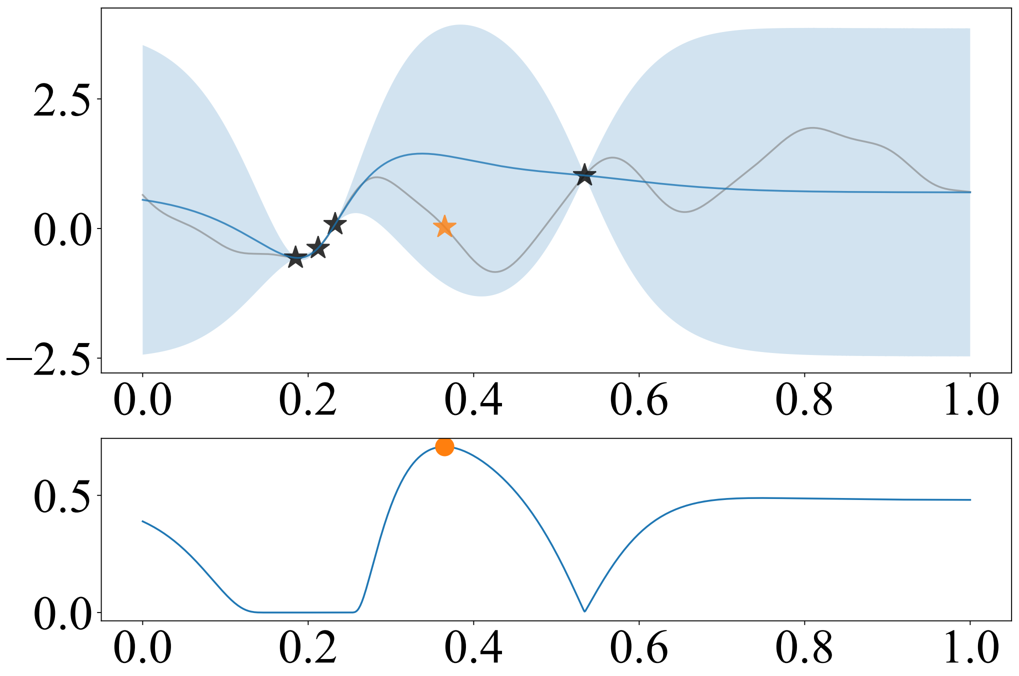


Expected improvement $\mathrm{EI}(x)$

Improvement-based design principle

Existing Design Principles

- Improvement-based (e.g., EI)
- Entropy-based
- Confidence bounds (UCB/LCB)
- Thompson sampling (TS)

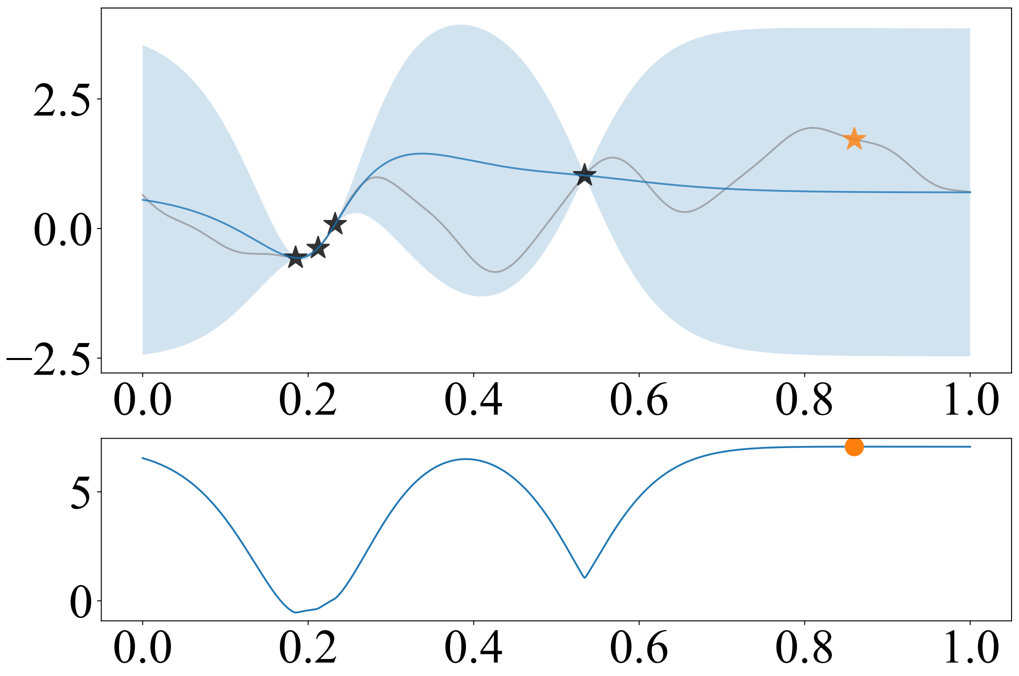


Expected improvement $\mathrm{EI}(x)$

Improvement-based design principle

New Design Principle: Gittins Index

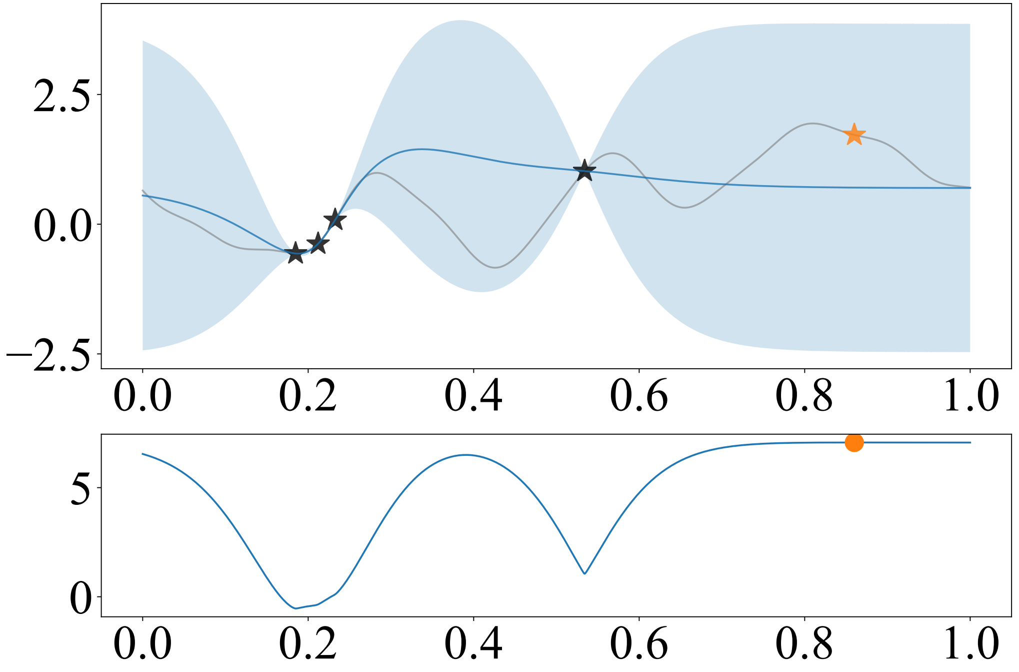
- Improvement-based (e.g., EI)
- Entropy-based
- Confidence bounds (UCB/LCB)
- Thompson sampling (TS)
- Gittins Index



Gittins index $\mathrm{GI}(x)$

New Design Principle: Gittins Index

- Improvement-based (e.g., EI)
- Entropy-based
- Confidence bounds (UCB/LCB)
- Thompson sampling (TS)
- Gittins Index

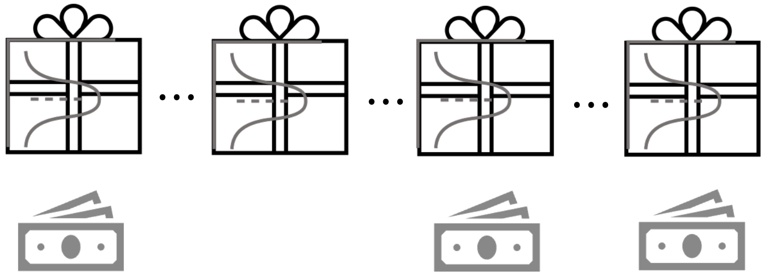


Gittins index $\mathrm{GI}(x)$

? Why another principle?

Our Contribution: Gittins Index Principle

Novel connection



Link to Pandora's Box problem
& Gittins index theory

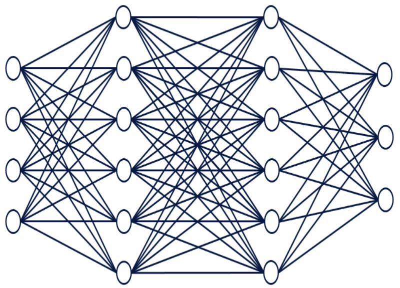
Principled decision rules

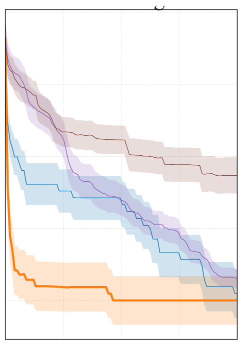
Varying evaluation costs

Adaptive stopping time

Unified framework for
selection and stopping

Competitive empirical performance





Interests from practitioners (e.g., Meta)

Future potential

Best-prompt identification

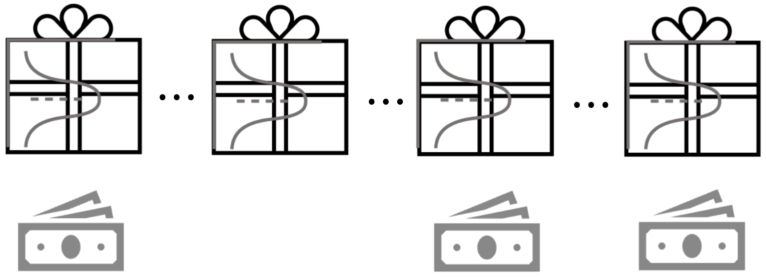
Adaptive response sampling

Chain-of-thought selection

Application to efficient LLM

Our Contribution: Gittins Index Principle

Novel connection



Link to Pandora's Box problem
& Gittins index theory

Principled decision rules

Varying evaluation costs

Adaptive stopping time

Unified framework for cost-aware selection and stopping

Competitive empirical performance

Interests from practitioners (e.g., Meta)

Future potential

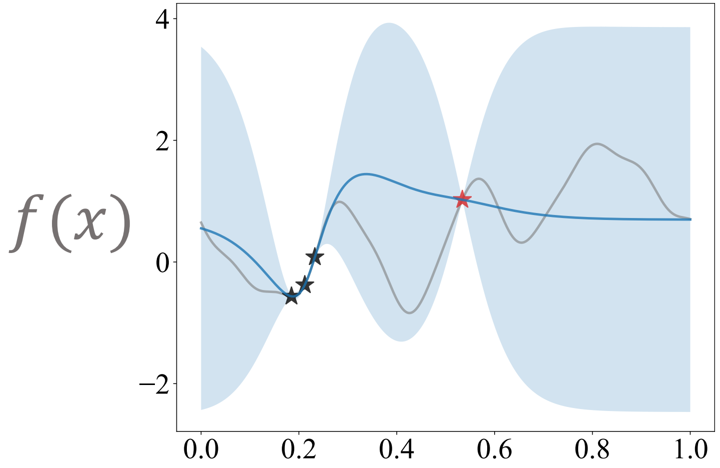
Best-prompt identification

Adaptive response sampling

Chain-of-thought selection

Application to efficient LLM

Bayesian Optimization

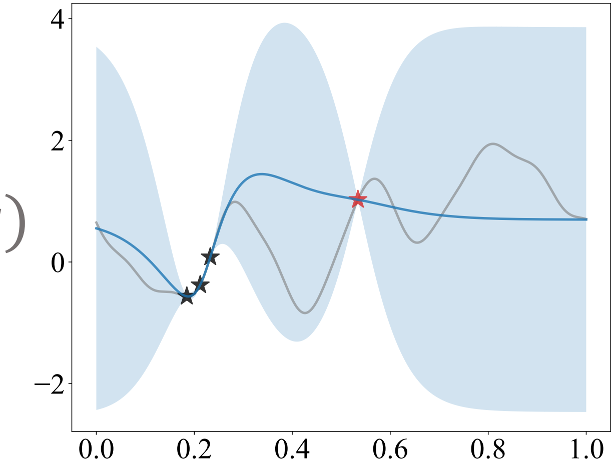


Continuous search space

Correlated function values

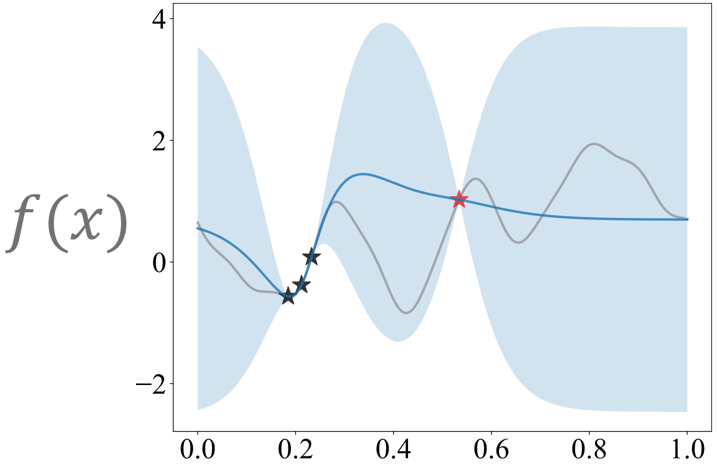
Bayesian Optimization

$f(x)$



Continuous search space	$\Longrightarrow$	Discrete
Correlated function values	$\Longrightarrow$	Independent

Bayesian Optimization

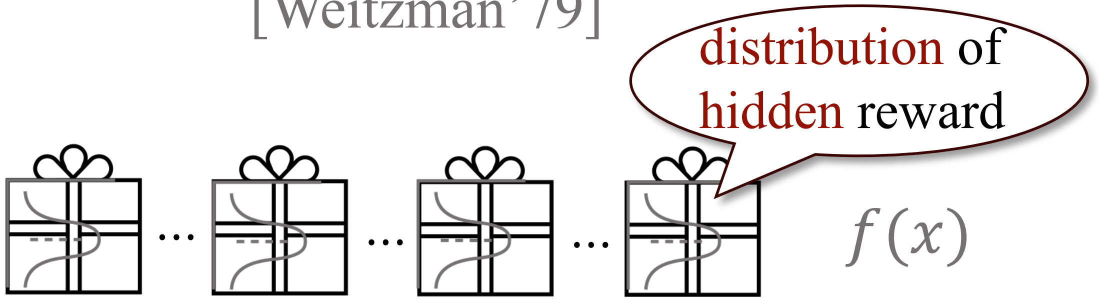


Continuous search space

Correlated function values

Pandora's Box

[Weitzman'79]

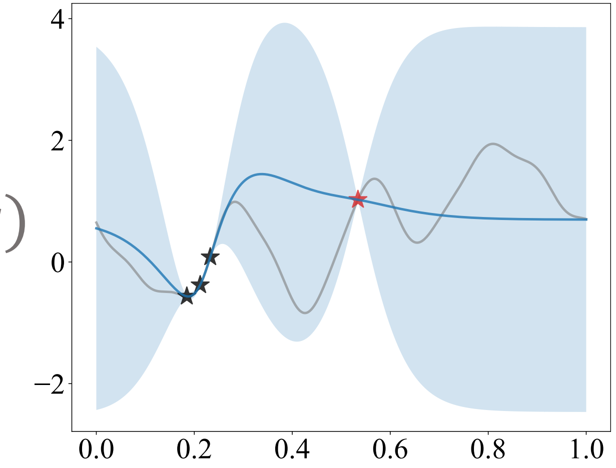


Discrete

Independent

Bayesian Optimization

$f(x)$



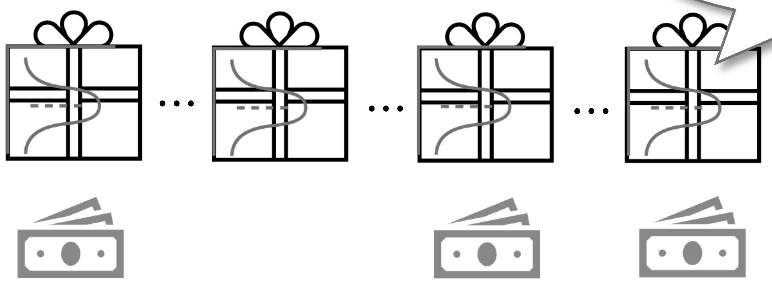
Continuous search space

Correlated function values

Pandora's Box

[Weitzman'79]

distribution of hidden reward



$f(x)$

$c(x)$

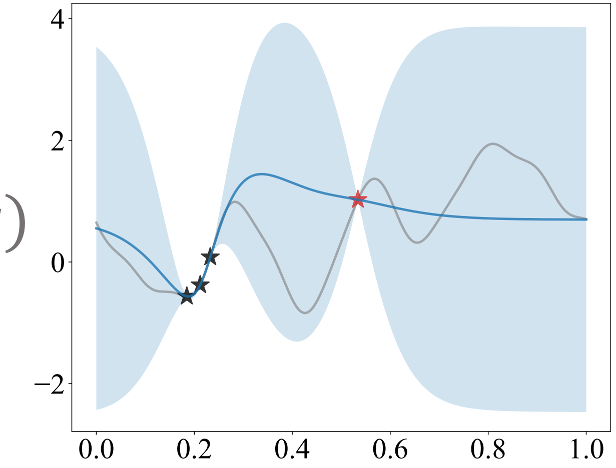
costs to open

Discrete

Independent

Bayesian Optimization

$f(x)$



Continuous search space

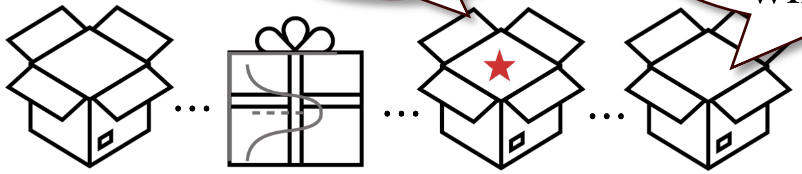
Correlated function values

Pandora's Box

[Weitzman'79]

selected

reward revealed when open



$f(x)$

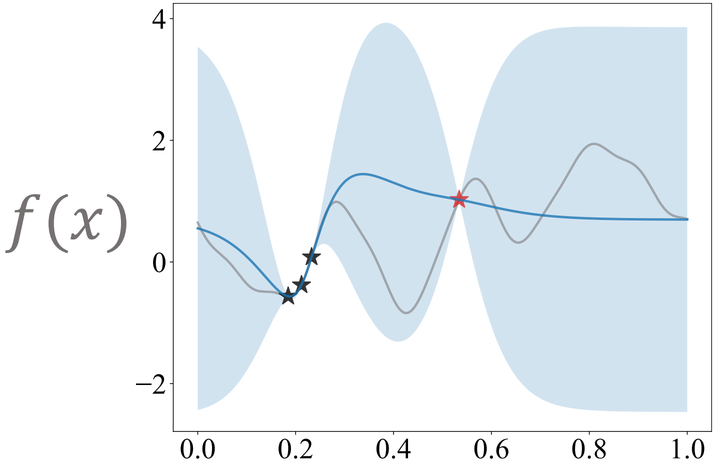
$c(x)$

costs to open

Discrete

Independent

Bayesian Optimization

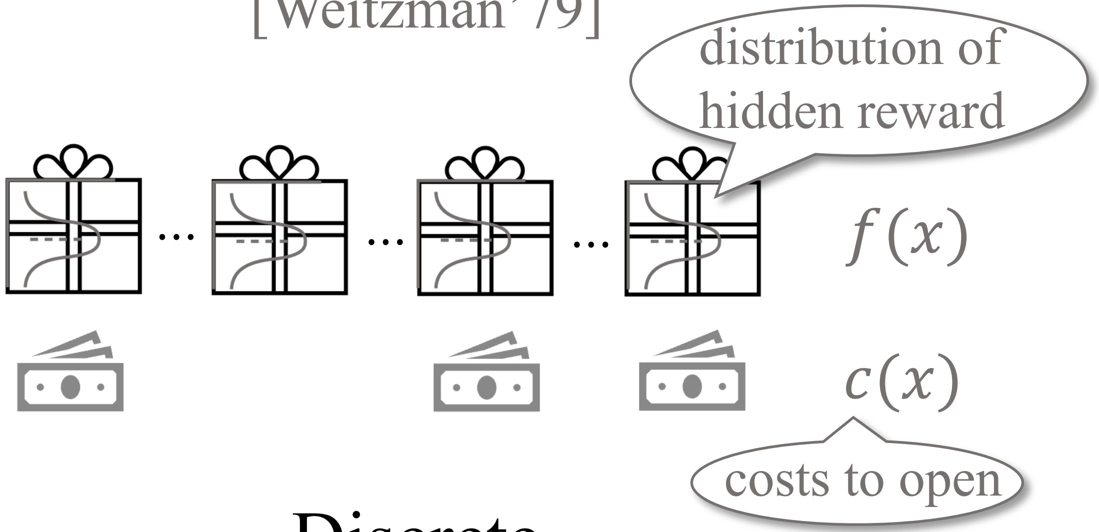


Continuous search space

Correlated function values

Pandora's Box

[Weitzman'79]

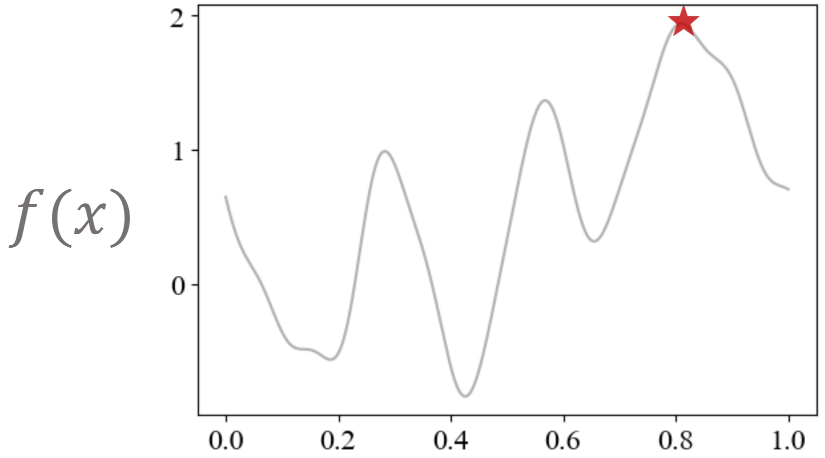


Discrete

Independent

Optimal policy: Gittins index

Bayesian Optimization



Continuous

Correlated

Pandora's Box

[Weitzman'79]

$\mathrm{GI}_f(x; c(x))$ g

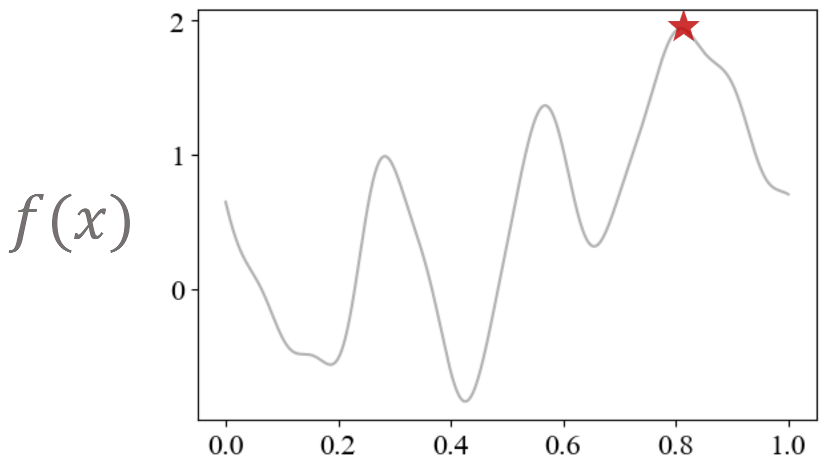
Step 1: Assign each box a Gittins index

Discrete

Independent

Optimal policy: Gittins index

Bayesian Optimization

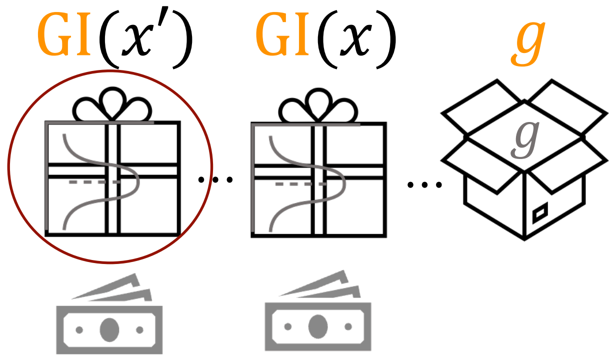


Continuous

Correlated

Pandora's Box

[Weitzman'79]



Step 2: Act on the box with the highest index

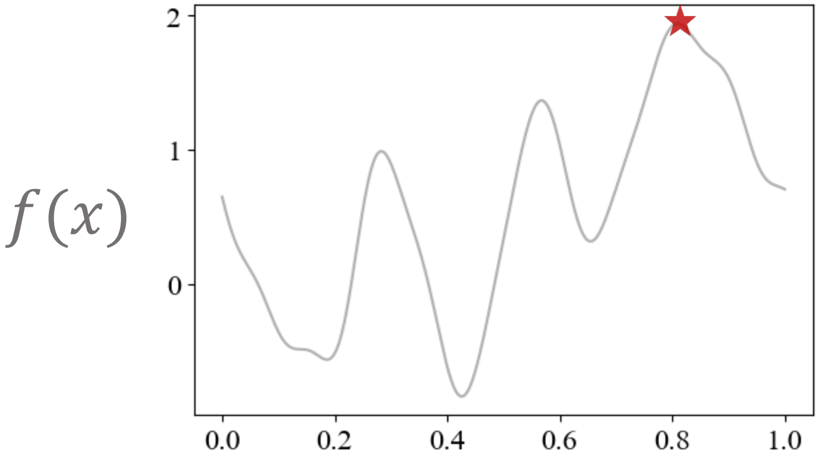
- *Closed*: open it
- *Opened*: select & stop

Discrete

Independent

Optimal policy: Gittins index

Bayesian Optimization

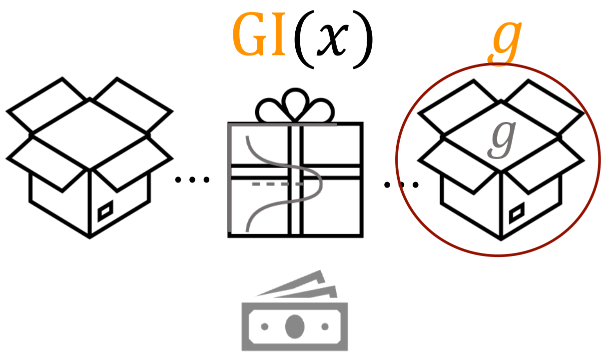


Continuous

Correlated

Pandora's Box

[Weitzman'79]



Step 2': Act on the box with the highest index

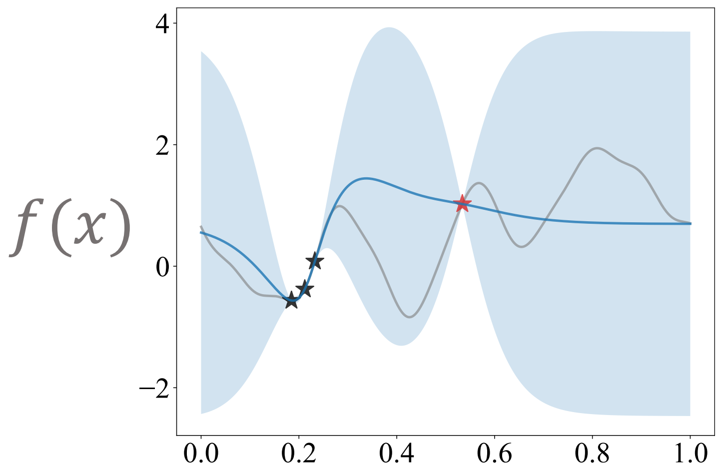
- *Closed*: open it
- *Opened*: select & stop

Discrete

Independent

Optimal policy: Gittins index

Bayesian Optimization

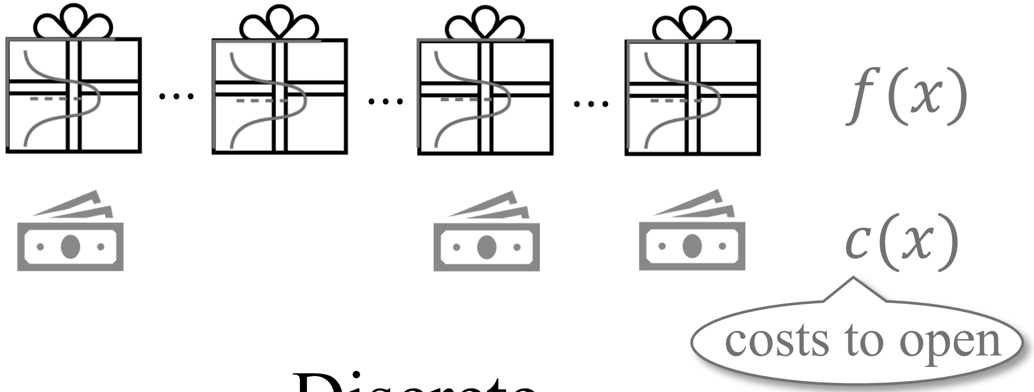


Continuous search space

Correlated function values

Pandora's Box

[Weitzman'79]



Discrete

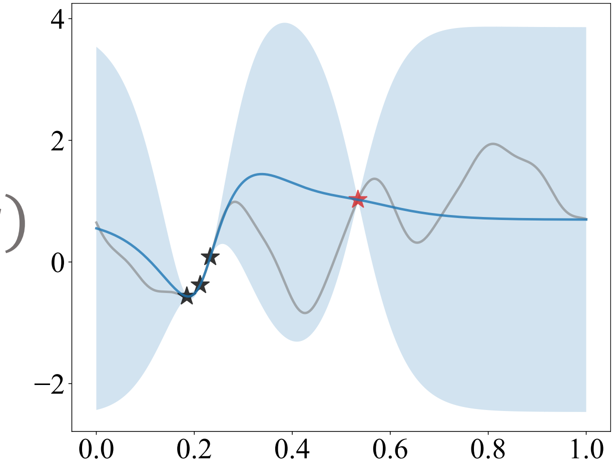
Independent

How to translate?

$\Longleftarrow$ Optimal policy: Gittins index

Bayesian Optimization

$f(x)$

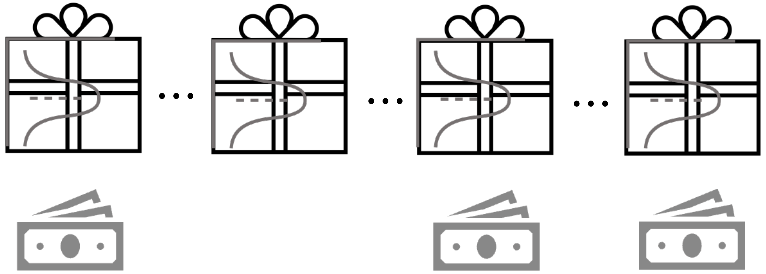


Continuous search space

Correlated function values

Pandora's Box

[Weitzman'79]



$f(x)$

$c(x)$

costs to open

Discrete

Independent

New!

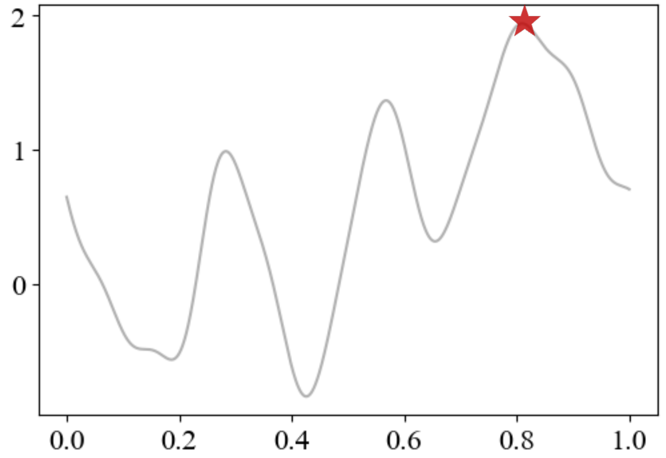
incorporate posterior

Our policy: $\mathrm{GI}_{f|D}(x;c)$ $\Longleftarrow$ Optimal policy: $\mathrm{GI}_f(x;c)$

take continuum limit

Bayesian Optimization

$f(x)$

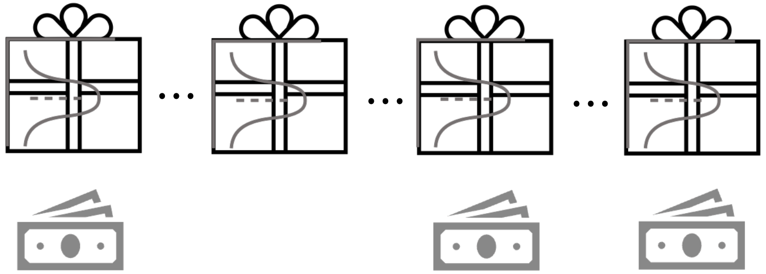


Continuous

Correlated

Pandora's Box

[Weitzman'79]



$f(x)$

$c(x)$

costs to open

Discrete

Independent

incorporate posterior

Our policy: $\mathrm{GI}_{f|D}(x; c(x))$ $\Longleftarrow$ Optimal policy: $\mathrm{GI}_f(x; c(x))$

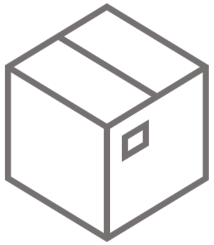
take continuum limit

How to compute?

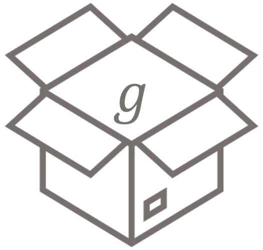
Intuition

Exploration vs. Exploitation

$f(x)$



vs.



$c(x)$

Open closed box — Take opened box

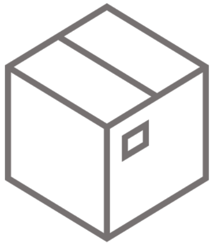
$\mathbb{E}[\max(f(x), g)] - c(x)$ — g

Should one open box? Depend on g!

Intuition

Exploration | Exploitation

$f(x)$



vs.

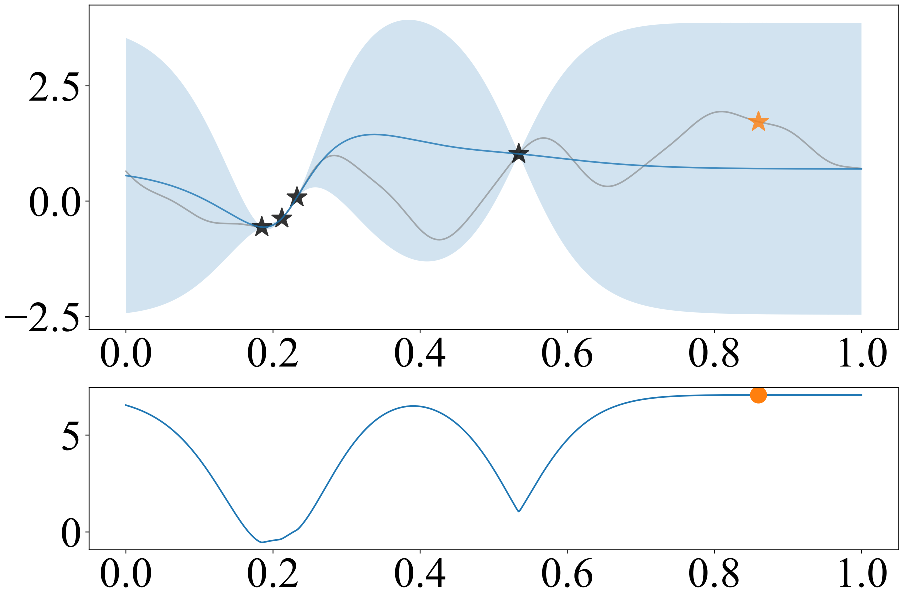
$c(x)$

Open closed box | Take opened box

$\mathbb{E}[\max(f(x), g)] - c(x)$ | g

Should one open box? Depend on g!

Gittins Index



$\mathrm{GI}_{f|D}(x;c) := $ solution g s.t.

$\mathbb{E}[\max(f(x), g) \mid D] - c(x) = g$

Intuition

Exploration | Exploitation

$f(x)$ vs.

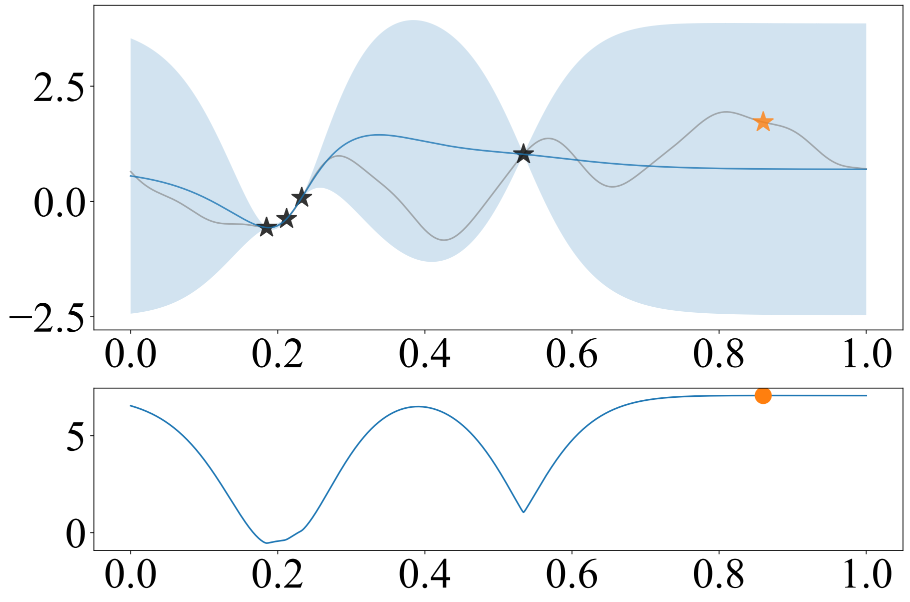
$c(x)$

Open closed box | Take opened box

$\mathbb{E}[\max(f(x), g)] - c(x)$ | g

Should one open box? Depend on g!

Gittins Index



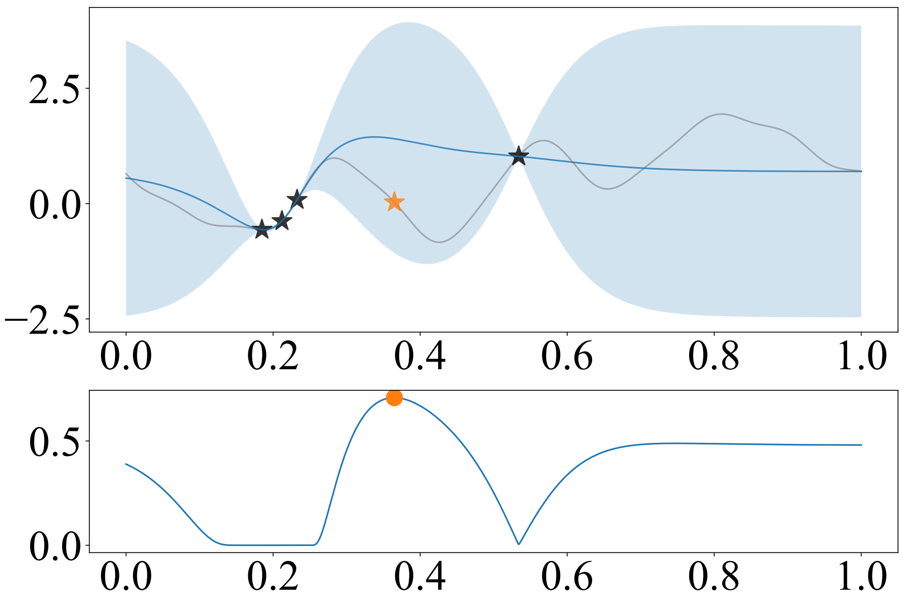
$$\mathrm{GI}_{f|D}(x; c) := \text{solution } g \text{ s.t.}$$

$$\mathbb{E}[\max(f(x), g) \mid D] - c(x) = g$$

$$\Leftrightarrow \mathbb{E}[\max(f(x) - g, g - g) \mid D] - c(x) = 0$$

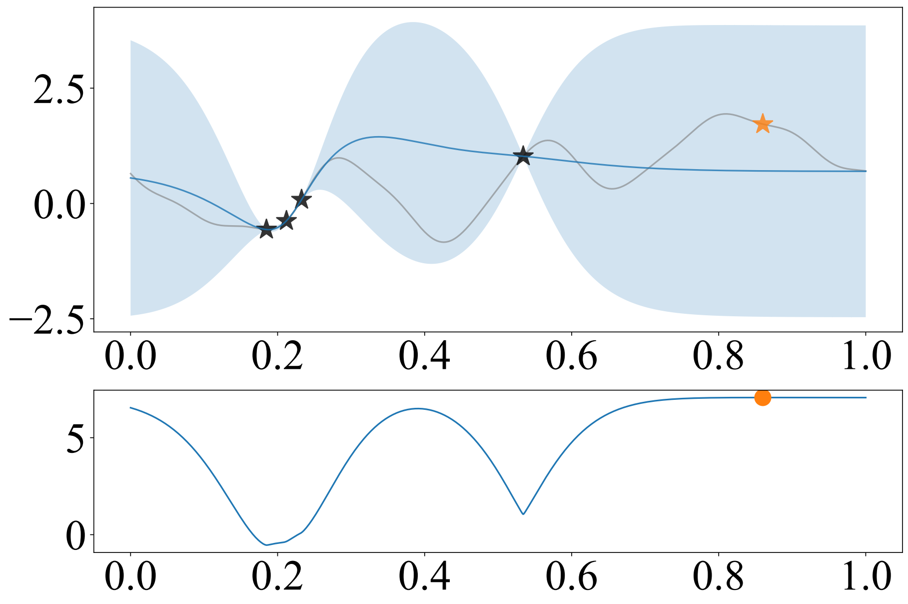
$$\Leftrightarrow \underbrace{\mathbb{E}[\max(f(x) - g, 0) \mid D]}_{\mathrm{EI}_{f|D}(x; g)} = c(x)$$

Expected Improvement

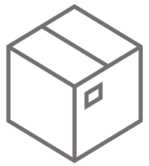


$\mathrm{EI}_{f|D}(x) := \mathbb{E}[\max(f(x) - y_{\text{best}}, 0) \mid D]$

Gittins Index



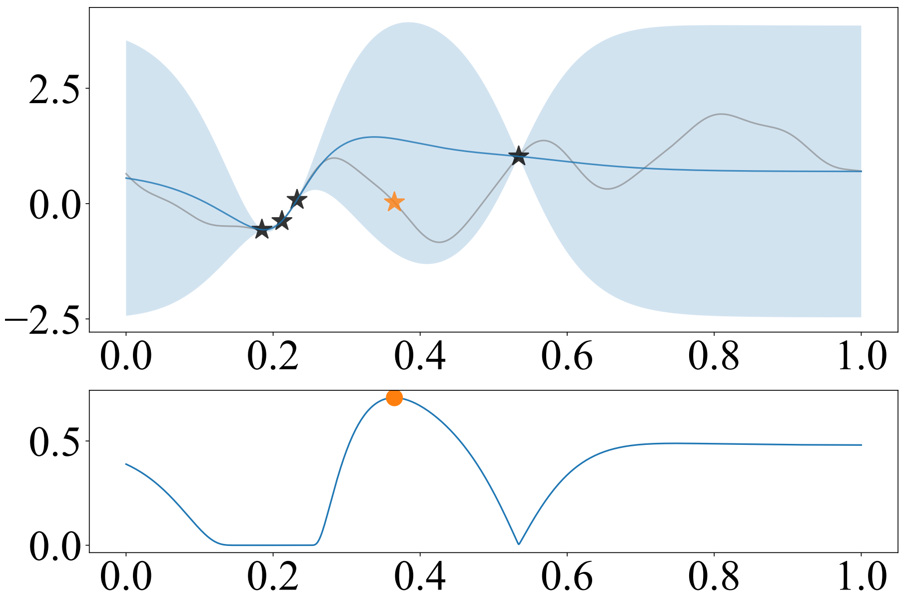
$\mathrm{GI}_{f|D}(x; c) :=$ solution g s.t. $\mathrm{EI}_{f|D}(x; g) = c(x)$
where $\mathrm{EI}_{f|D}(x; g) := \mathbb{E}[\max(f(x) - g, 0) \mid D]$

$f(x)$  vs. 

$c(x)$ 

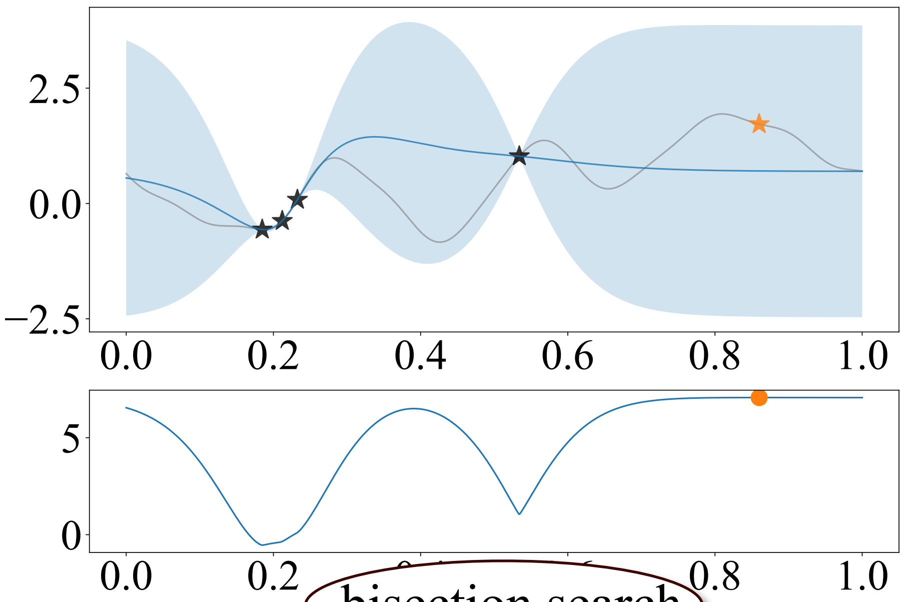
Exploration Exploitation

Expected Improvement



$$\mathrm{EI}_{f|D}(x) := \mathbb{E}[\max(f(x) - y_{\text{best}}, 0) \mid D]$$

Gittins Index



bisection search

$$\mathrm{GI}_{f|D}(x; c) := \text{solution } g \text{ s.t. } \mathrm{EI}_{f|D}(x; g) = c(x)$$

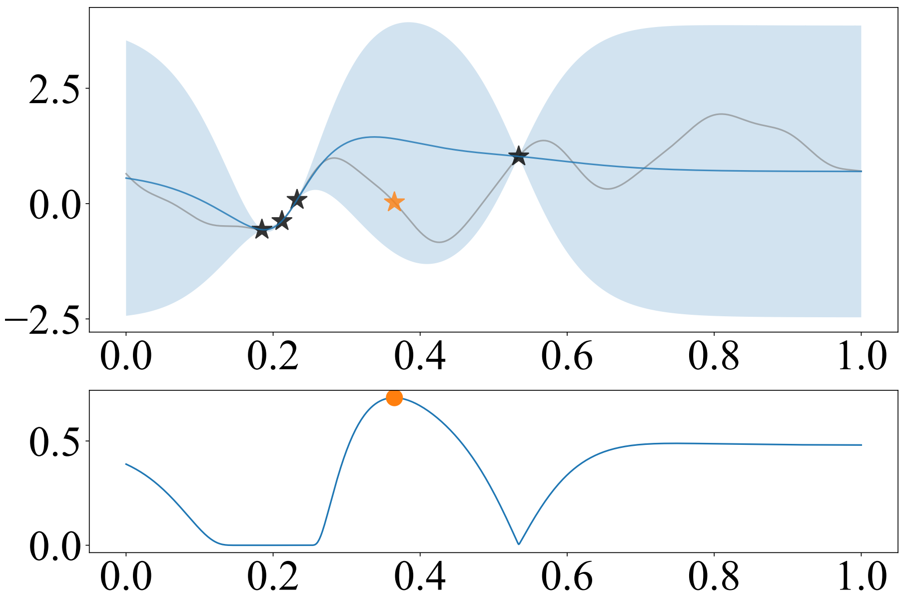
$$\text{where } \mathrm{EI}_{f|D}(x; g) := \mathbb{E}[\max(f(x) - g, 0) \mid D]$$

analytical expression

Easy-to-compute decision rule!

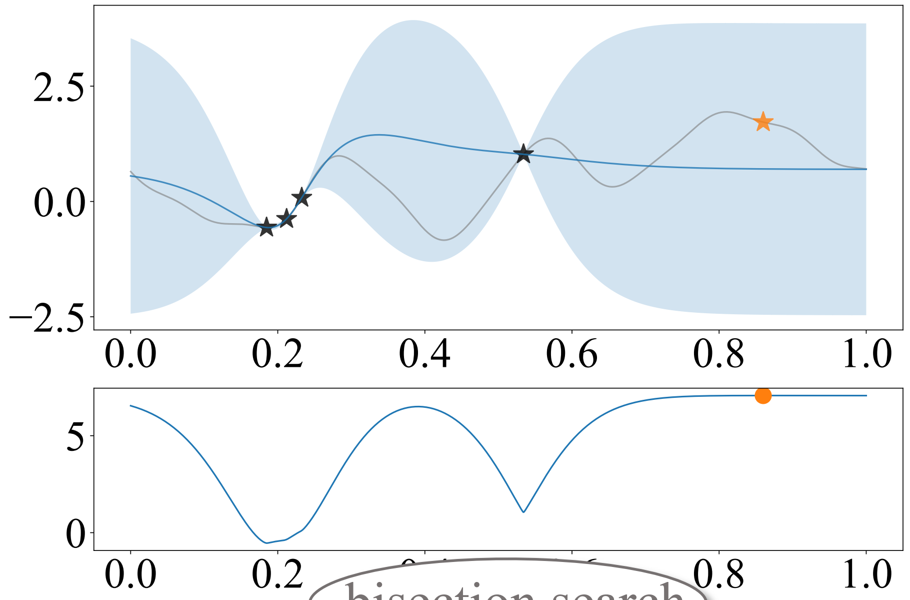
"Cost-aware Bayesian Optimization via the Pandora's Box Gittins Index." NeurIPS'24.

Expected Improvement



$$\mathrm{EI}_{f|D}(x) := \mathbb{E}[\max(f(x) - y_{\mathrm{best}}, 0) \mid D]$$

Gittins Index



bisection search

$$\mathrm{GI}_{f|D}(x; c) := \text{solution } g \text{ s.t. } \mathrm{EI}_{f|D}(x; g) = c(x)$$

$$\text{where } \mathrm{EI}_{f|D}(x; g) := \mathbb{E}[\max(f(x) - g, 0) \mid D]$$

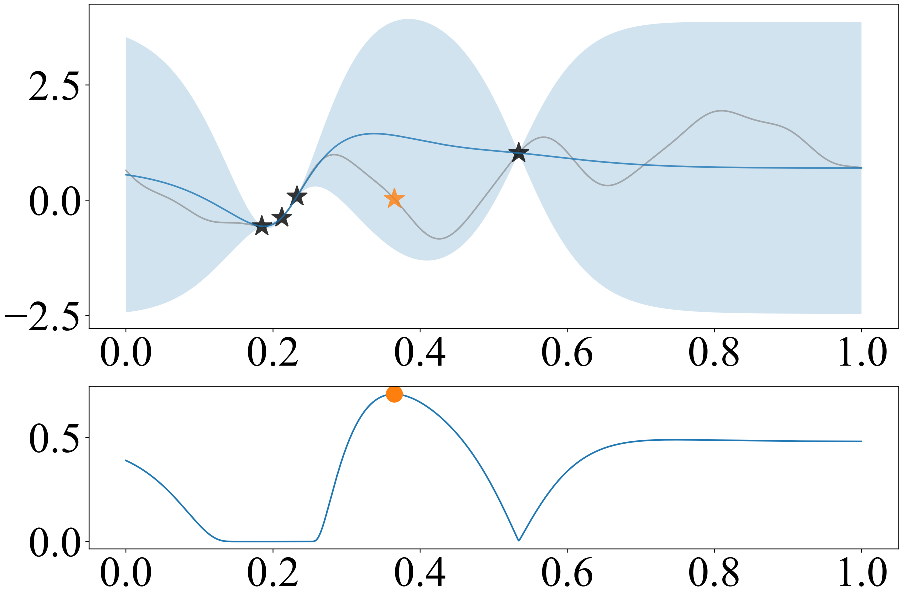
Google DeepMind

FunBO: Discovering new acquisition functions for Bayesian Optimization with FunSearch

LLM-driven evolutionary search method

GI is not easy to discover!

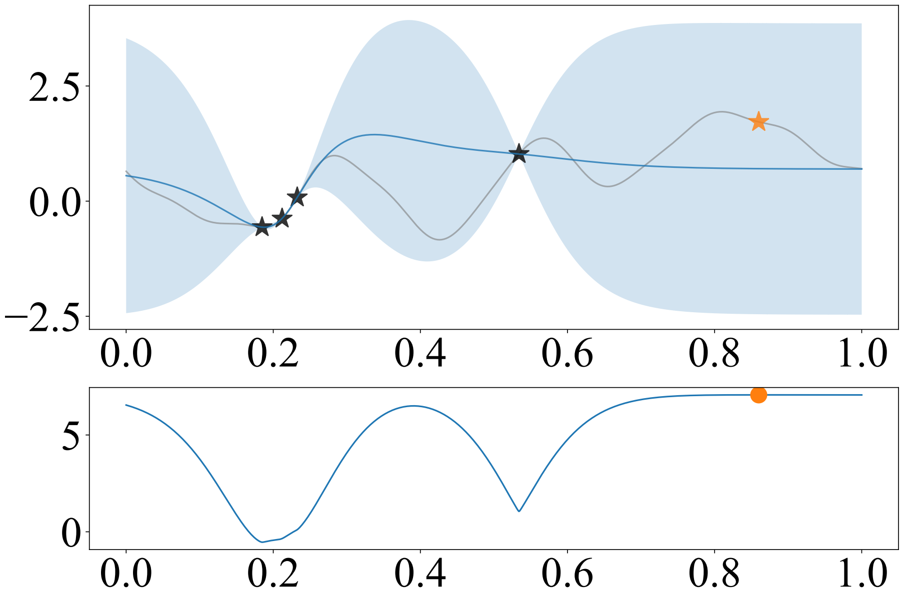
Expected Improvement



$\mathrm{EI}_{f|D}(x) := \mathbb{E}[\max(f(x) - y_{\text{best}}, 0) \mid D]$

Temporal simplification to MDP
(One-step)

Gittins Index



$\mathrm{GI}_{f|D}(x; c) :=$ solution g s.t. $\mathrm{EI}_{f|D}(x; g) = c(x)$
where $\mathrm{EI}_{f|D}(x; g) := \mathbb{E}[\max(f(x) - g, 0) \mid D]$

Spatial simplification to MDP

Our Contribution: Gittins Index Principle

Novel connection

Link to Pandora's Box problem
& Gittins index theory

Principled decision rules

Varying evaluation costs

Adaptive stopping time

Unified framework for
selection and stopping

Competitive empirical performance

Interests from practitioners (e.g., Meta)

Future potential

Best-prompt identification

Adaptive response sampling

Chain-of-thought selection

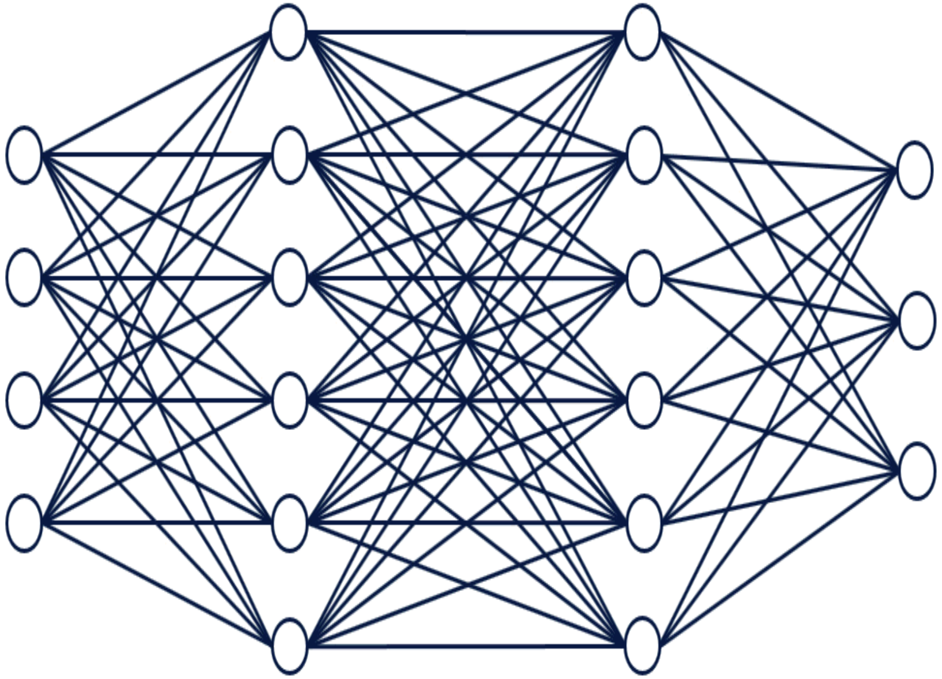
Application to efficient LLM

Under-explored Factors for Better Decisions

Varying evaluation costs

Smart stopping time

Observable multi-stage feedback



How does existing principle incorporate them?

Varying evaluation costs

$$\mathrm{EIPC}(x;c) = \mathrm{EI}(x)/c(x)$$

[Snoek et al.'12]

Arbitrarily bad

[Astudillo et al.'21]

Smart stopping time

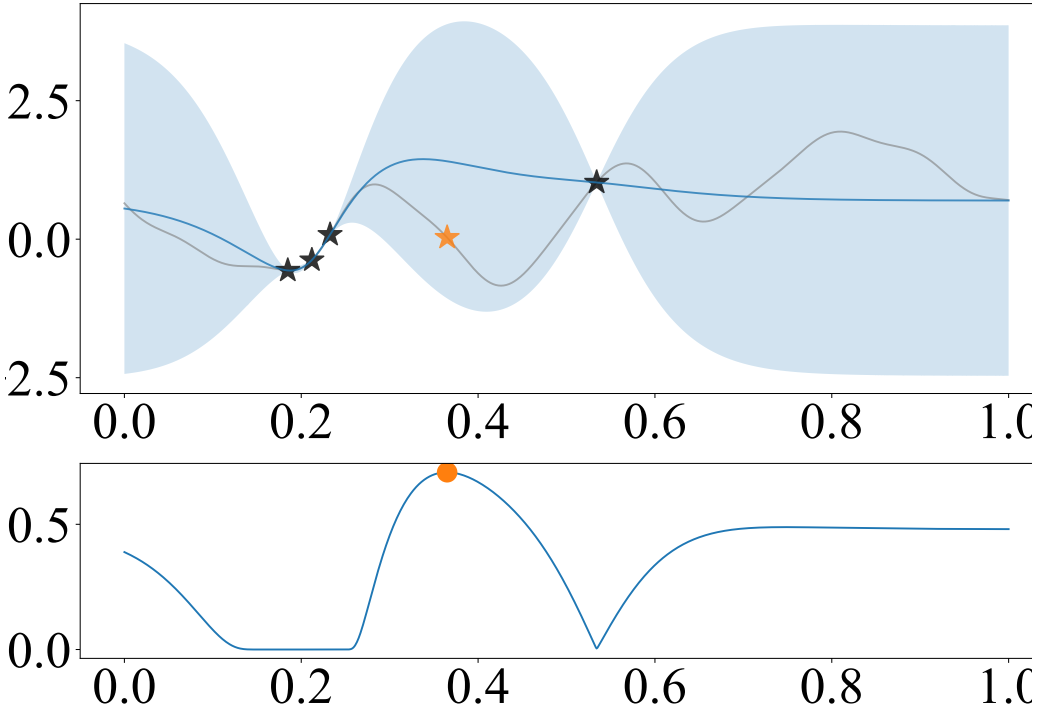
$$\tau:\ \mathrm{EI}(x_\tau) \le \theta$$

[Locatelli'97,
Nguyen et al.'17,
Ishibashi et al.'23]

Which threshold?

Observable multi-stage feedback

?



Expected improvement $\mathrm{EI}(x)$

Under-explored Factors for Better Decisions

Varying evaluation costs

Smart stopping time

Observable multi-stage feedback

New design principle:
Gittins index

Why Gittins index?

Varying evaluation costs

Smart stopping time

Observable multi-stage feedback

New design principle:
Gittins index

Optimal in related sequential decision problems

Why Gittins index?

Varying evaluation costs

Features in Pandora's box

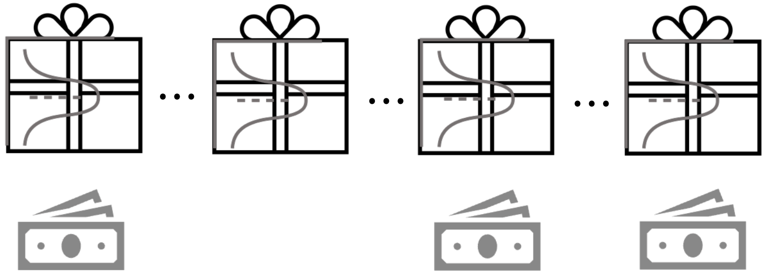
Smart stopping time

Features in Pandora's box

Observable multi-stage feedback

New design principle:
Gittins index

Optimal in related sequential decision problems



Why Gittins index?

Varying evaluation costs

Features in Pandora’s box

Smart stopping time

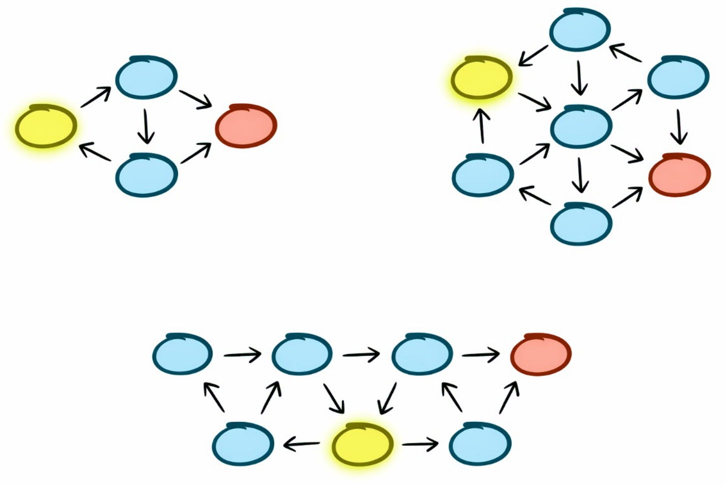
Features in Pandora’s box

Observable multi-stage feedback

Features in Markov chain selection

New design principle:
Gittins index

Optimal in related sequential decision problems



Why Gittins index?

Varying evaluation costs

Features in Pandora's box

This talk's focus

Smart stopping time

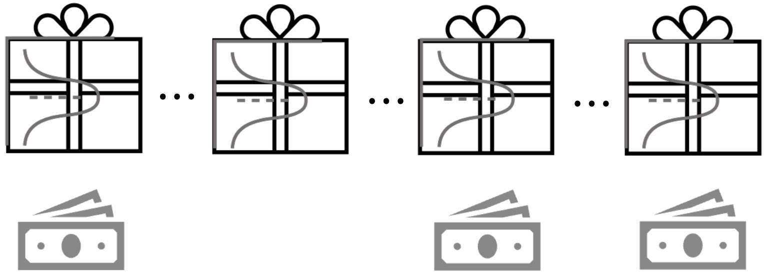
Features in Pandora's box

Observable multi-stage feedback

Features in Markov chain selection

New design principle:
Gittins index

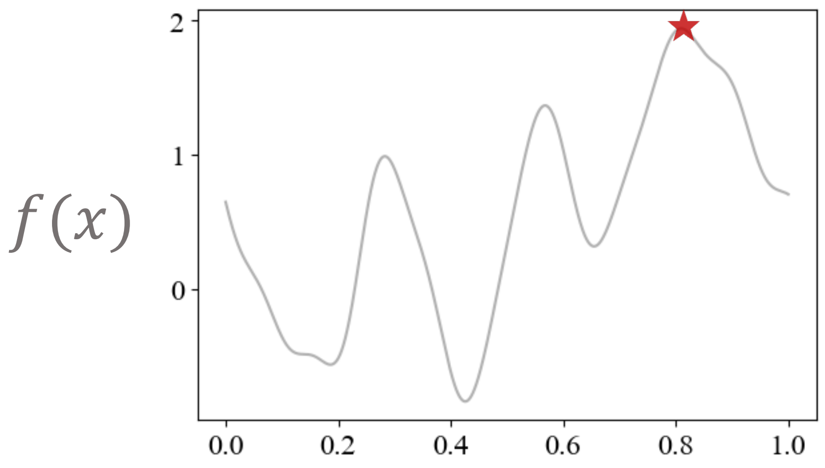
Optimal in related sequential decision problems



"Cost-aware Bayesian Optimization via the Pandora's Box Gittins Index." NeurIPS'24.

"Cost-aware Stopping for Bayesian Optimization." Under review.

Bayesian Optimization



Continuous

Correlated

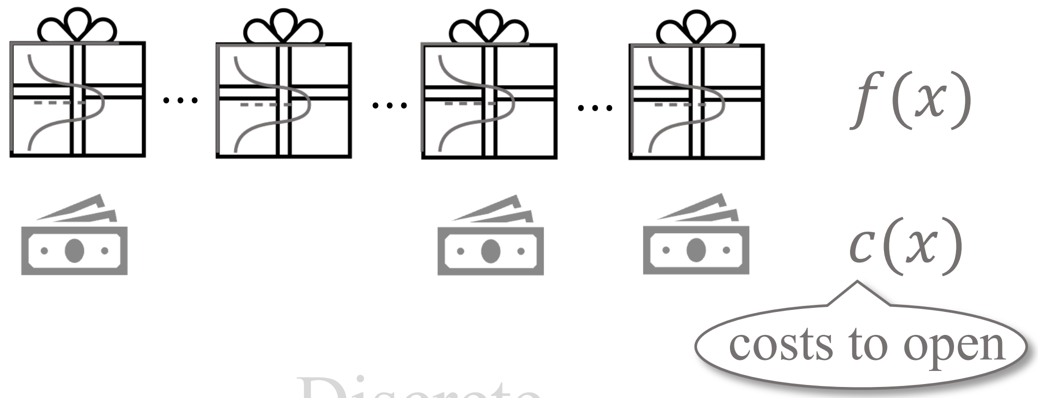
Cost-unaware

Fixed-iteration

$$\mathbb{E} \max_{t=1,2,\ldots,T} f(x_t)$$

Pandora's Box

[Weitzman'79]



Discrete

Independent

Cost-aware

Flexible-stopping

$$\mathbb{E} \max_{t=1,2,\ldots,T} f(x_t) - \mathbb{E} \sum_{t=1}^{T} c(x_t)$$

Expected Improvement vs Gittins Index

Varying evaluation costs

$\mathrm{EIPC}(x; c) = \mathrm{EI}(x)/c(x)$

Arbitrarily bad

$\mathrm{GI}(x; c) := \text{solution } g \text{ s.t. } \mathrm{EI}(x; g) = c(x)$

naturally incorporates costs

Smart stopping time

$\tau\colon \mathrm{EI}(x_\tau) \le \theta$

Which threshold?

$\tau\colon \mathrm{GI}(x_\tau; c) \le y_{\mathrm{best}}$

$\Leftrightarrow \tau\colon \mathrm{EIPC}(x_\tau; c) \le 1$

derived shared stopping rule

"Cost-aware Bayesian Optimization via the Pandora's Box Gittins Index." NeurIPS'24.

"Cost-aware Stopping for Bayesian Optimization." Under review.

Theoretical Guarantee and Empirical Validation

Theorem (Safeguard Guarantee)

$\mathbb{E}[R(\text{ours; PBGI})] \leq R[\text{stopping immediately}]$

or LogEIPC

cost-adjusted regret

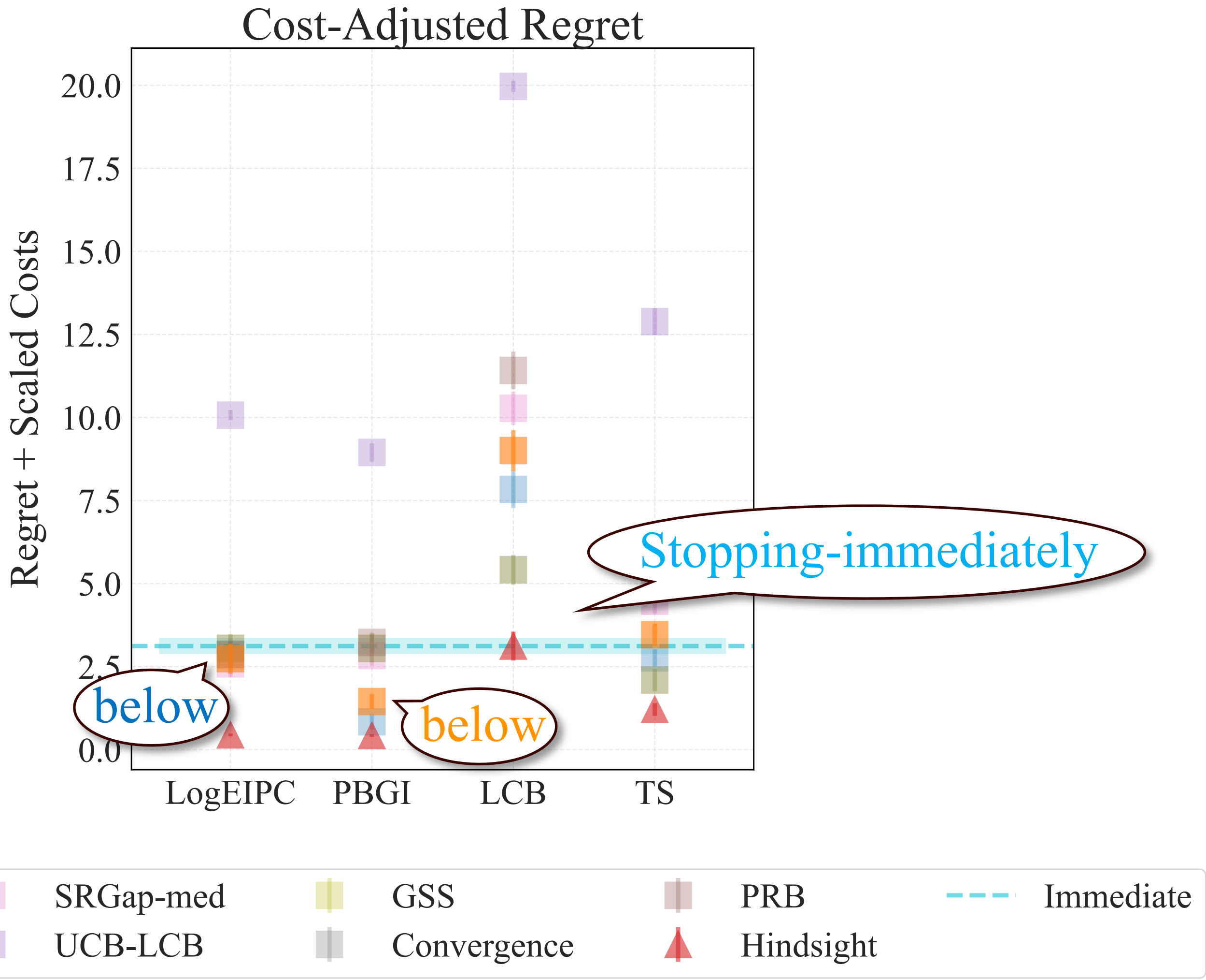
Implication:

• Matches the best achievable performance in the worst case (evaluations are all very costly).

• Avoids over-spending — a property many cost-unaware stopping rules lack. New

Proof idea: For all $t < \tau$, $\mathrm{EI}(x_{t+1}) \geq c(x_{t+1})$.

stopping time



"Cost-aware Stopping for Bayesian Optimization." Under review.

Our Contribution: Gittins Index Principle

Novel connection

Link to Pandora's Box problem
& Gittins index theory

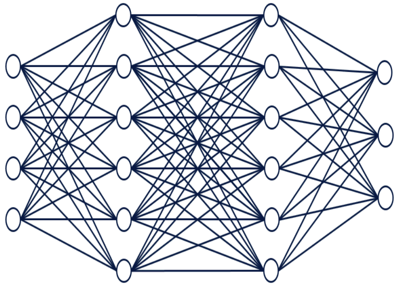
Principled decision rules

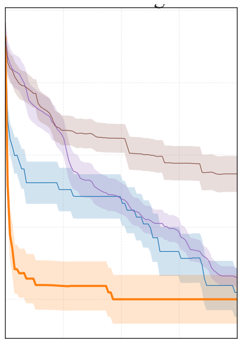
Varying evaluation costs

Adaptive stopping time

Unified framework for
selection and stopping

Competitive empirical performance





Interests from practitioners (e.g., Meta)

Future potential

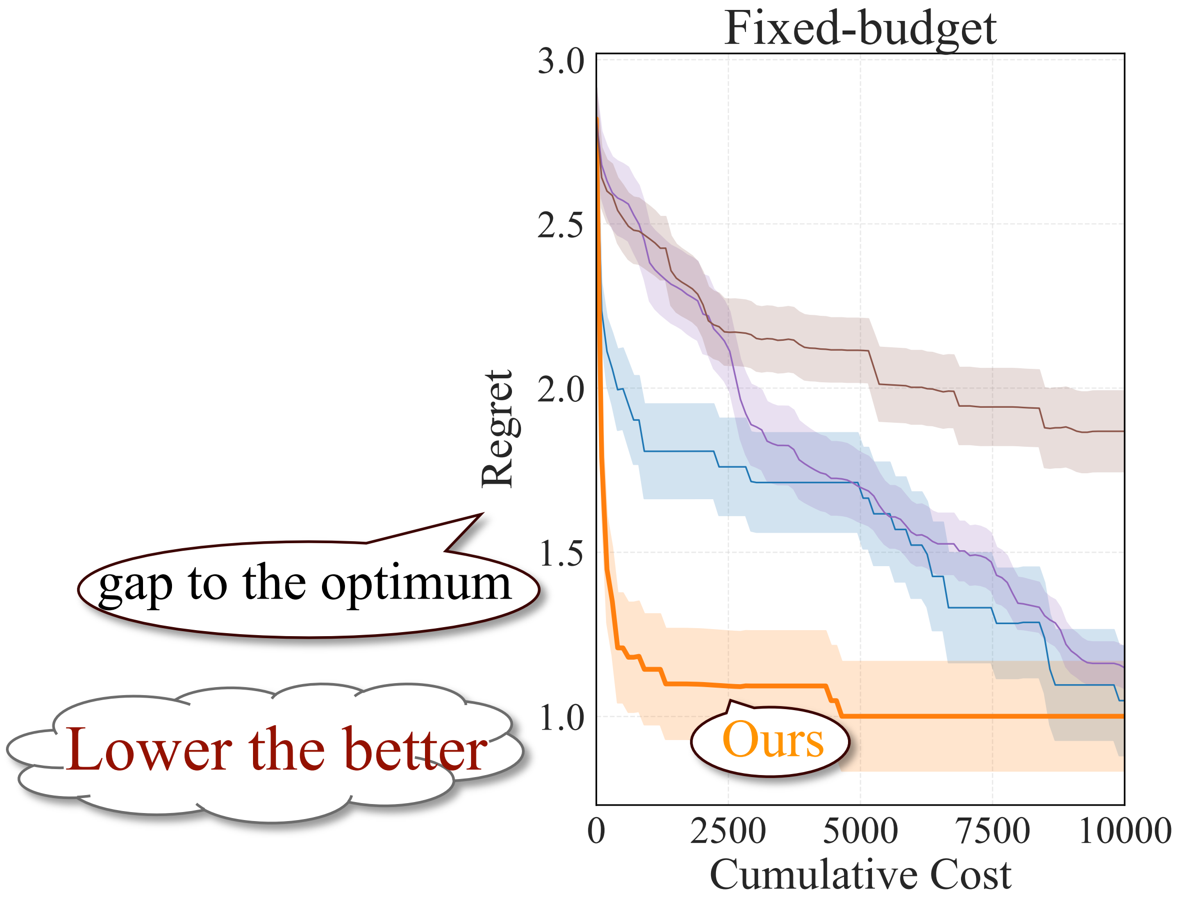
Best-prompt identification

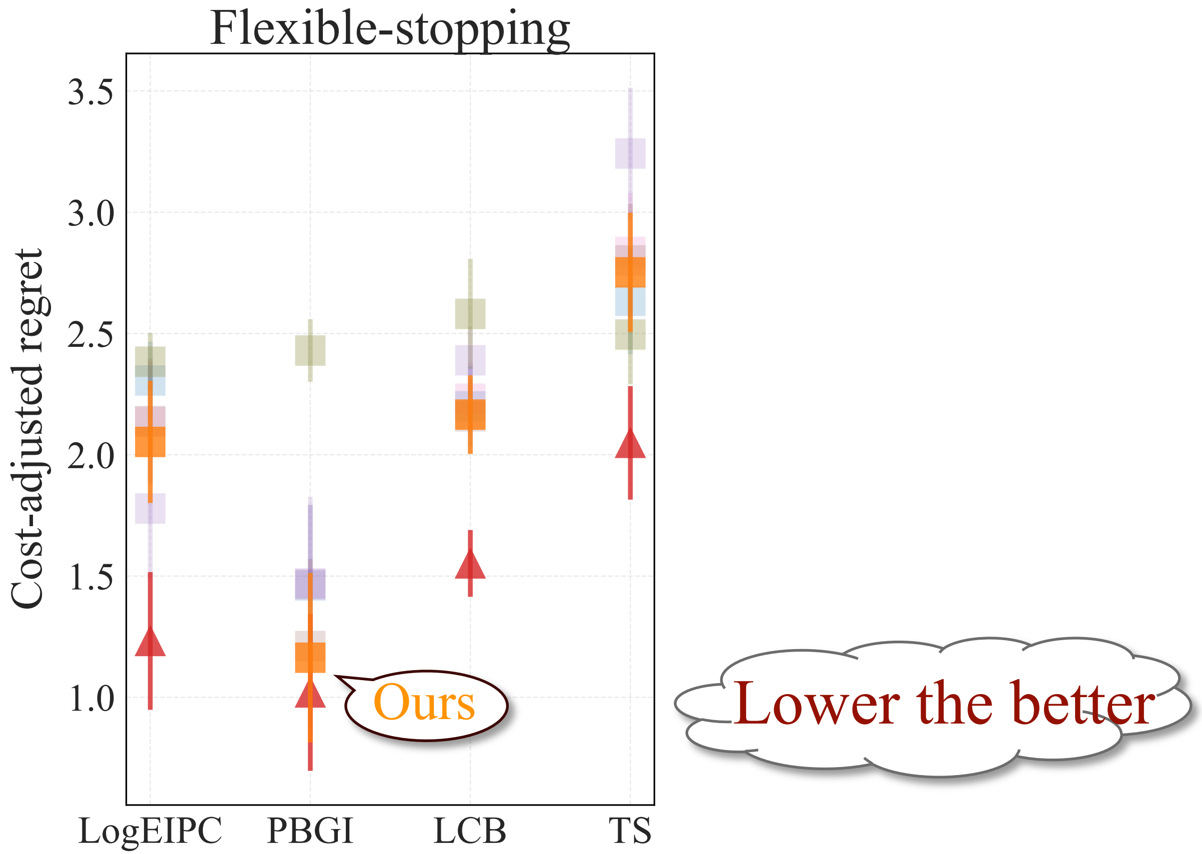
Adaptive response sampling

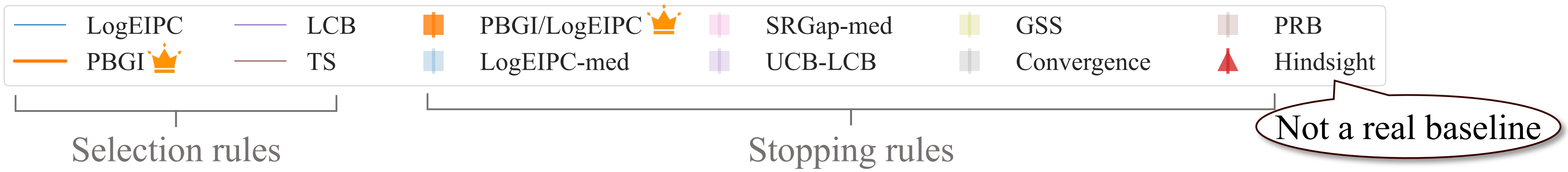
Chain-of-thought selection

Application to efficient LLM

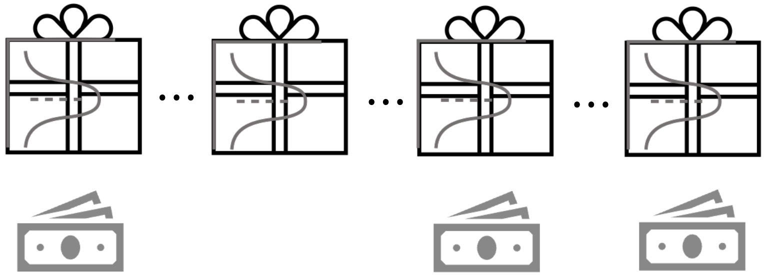
Gittins Index vs Baselines on AutoML Benchmark







Novel connection



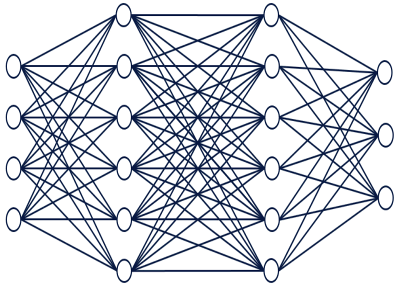
Link to Pandora's Box problem & Gittins index theory

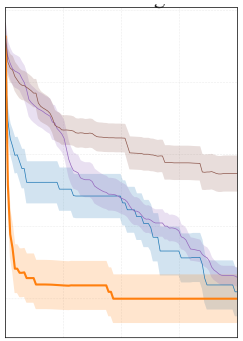
Principled decision rules

- Varying evaluation costs
- Adaptive stopping time

Unified framework for selection and stopping

Competitive empirical performance





Interests from practitioners (e.g., Meta)

Future potential

- Best-prompt identification
- Adaptive response sampling
- Chain-of-thought selection

"Cost-aware Bayesian Optimization via the Pandora's Box Gittins Index." NeurIPS'24.

"Cost-aware Stopping for Bayesian Optimization." Under review.

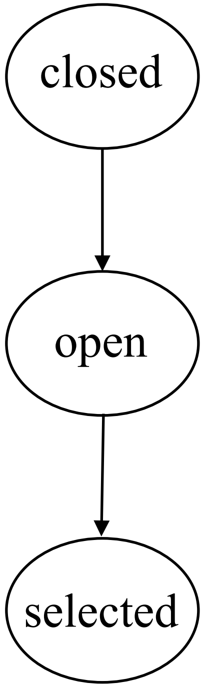
Find my papers on arXiv!

"Cost-aware Bayesian Optimization via the Pandora's Box Gittins Index.” NeurIPS’24.

"Cost-aware Stopping for Bayesian Optimization.” Under review.

Markovian Bandits

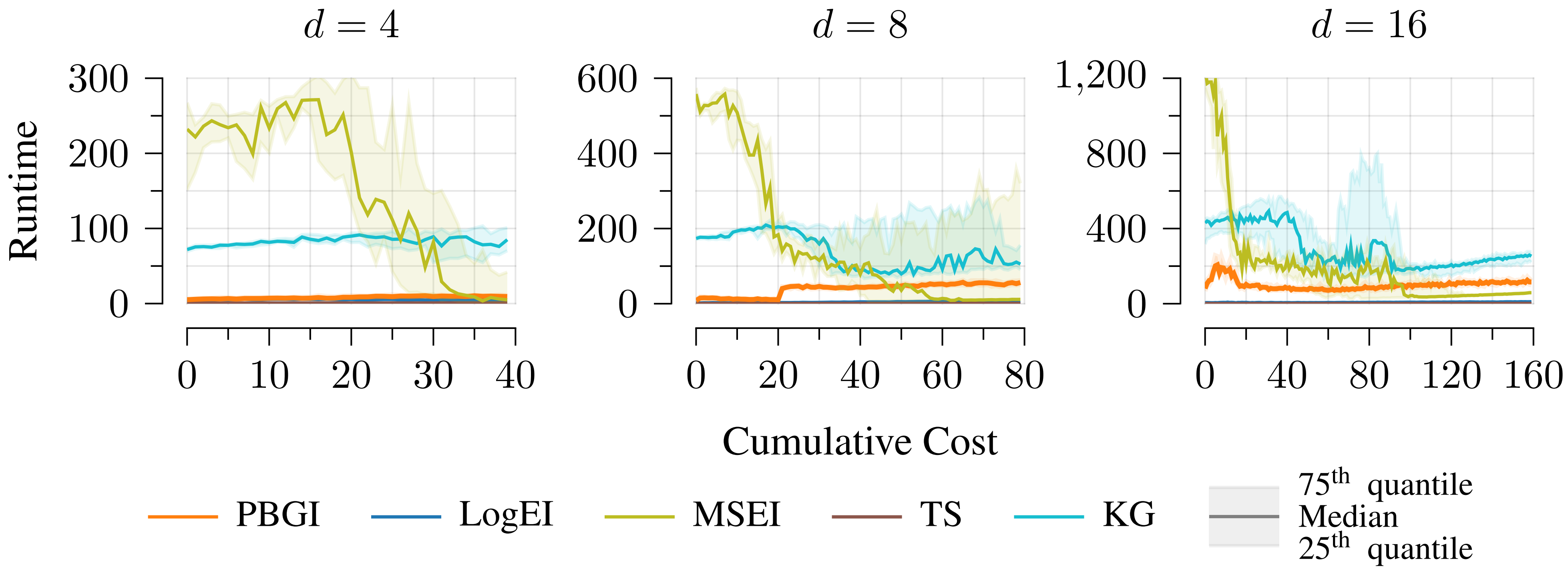
[Gittins'79]



Pandora's Box

[Weitzman'79]

Timing Experiment: Gittins Index vs Baselines



PBGI is computationally-efficient!