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

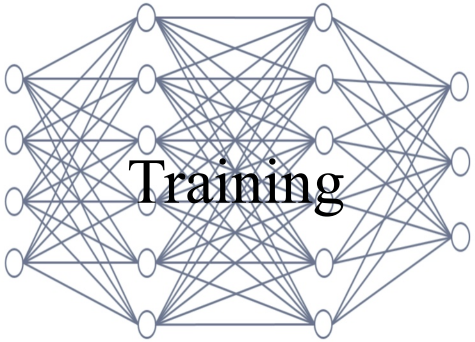
Qian Xie 谢倩 (Cornell ORIE)

TTAP Job Talk

Motivation: World of Optimization under Uncertainty

ML model training:

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



⟶ Accuracy

Motivation: World of Optimization under Uncertainty

ML model training:

≠ parameters
(weights, biases)

Training hyperparameters ⟶

(e.g., learning rate, # layers)

⟶ Accuracy

Motivation: World of Optimization under Uncertainty

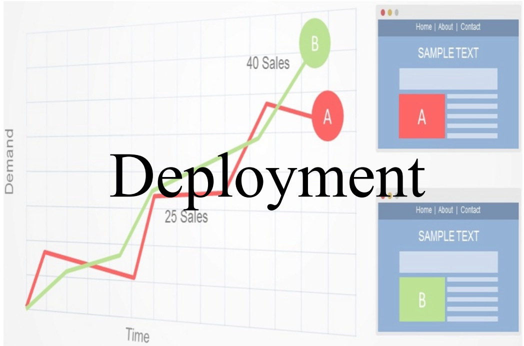
ML model training:

≠ parameters
(weights, biases)

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

Adaptive experimentation:

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



Motivation: World of Optimization under Uncertainty

Design choices x ⟶

non-analytical & no gradient info

⟶ Performance metric $f(x)$

ML model training:

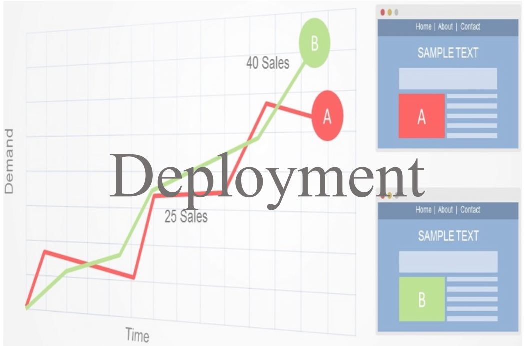
≠ parameters (weights, biases)

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)

non-analytical & no gradient info

Input x $\longrightarrow$

$\longrightarrow$ Observed outcome $f(x)$

ML model training:

$\neq$ parameters (weights, biases)

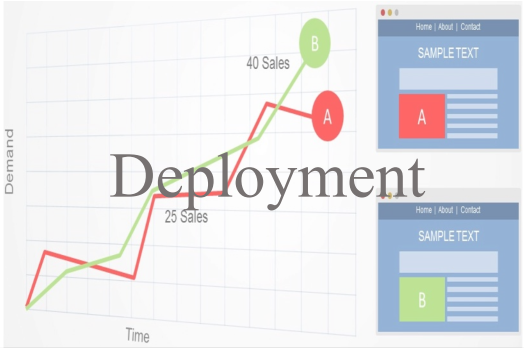
Training hyperparameters (e.g., learning rate, # layers) $\longrightarrow$

Training

$\longrightarrow$ Accuracy

Adaptive experimentation:

Decision/design variables (e.g., layout, pricing level) $\longrightarrow$



Deployment

$\longrightarrow$ Revenue

Background: Black-Box Optimization

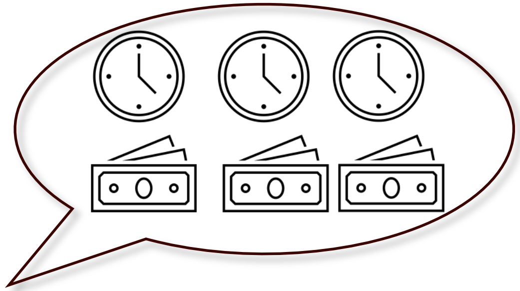
Black-box optimization:

(gradient-based methods not applicable)

Input x ⟶

expensive-to-evaluate

⟶ Observed outcome $f(x)$

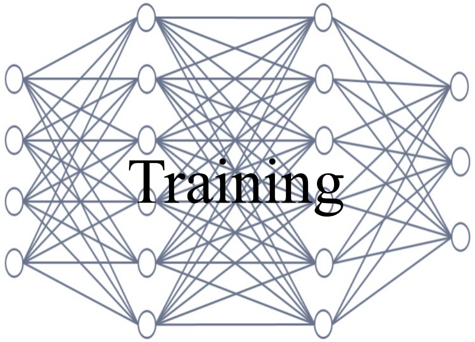


Training time

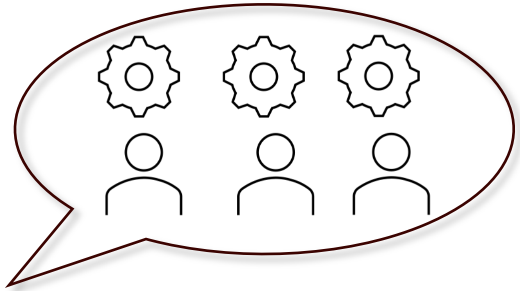
Compute credits

ML model training:

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



⟶ Accuracy

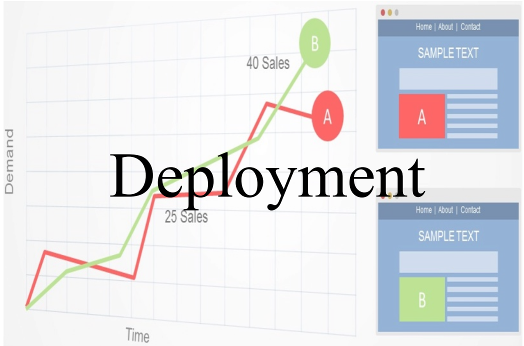


Operational cost

User experience

Adaptive experimentation:

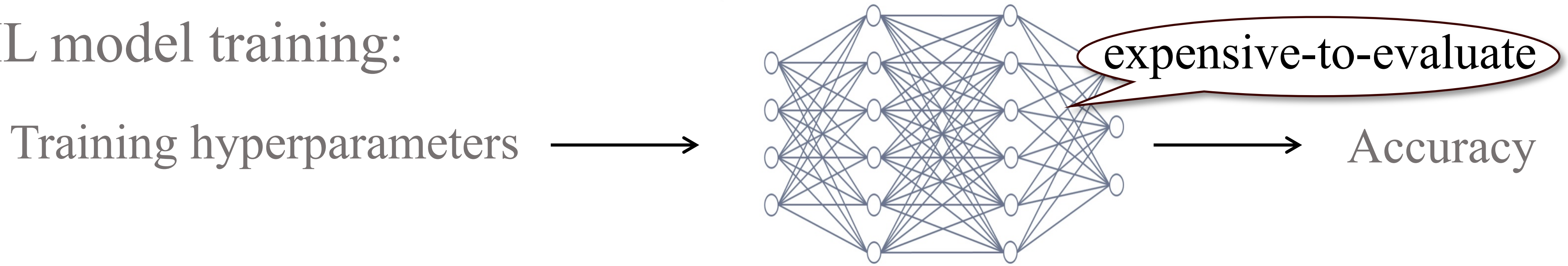
Decision/design variables (e.g., layout, pricing level) ⟶



⟶ Revenue

Naïve (Non-Adaptive) Approach: Grid Search

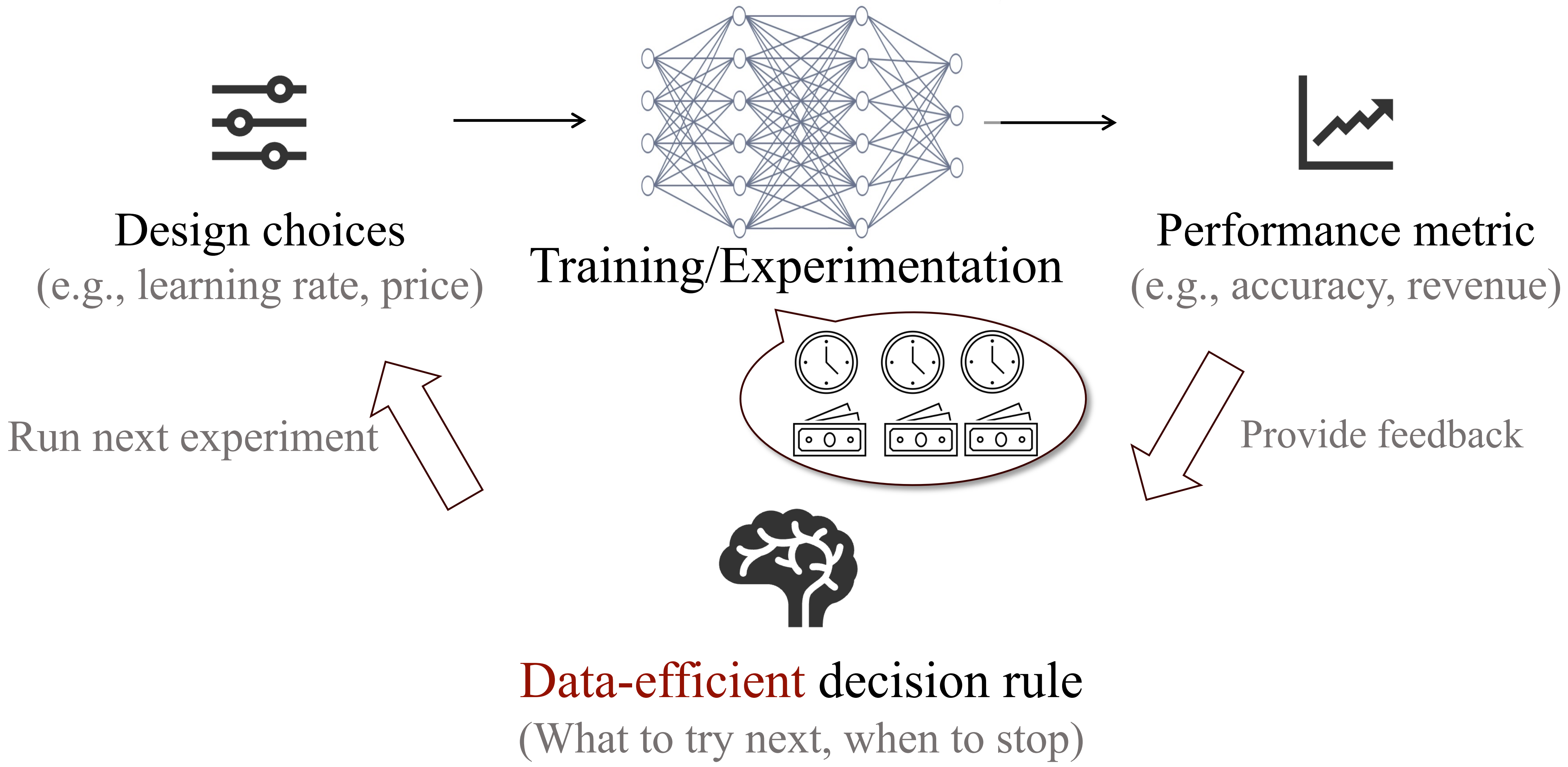
ML model training:



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!

Data-Driven (Adaptive) Approach



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-based agent

New Methods for Black-Box Optimization

Naïve approaches:	Data-driven approaches:
• Grid search	• Local search
• Random search	• Evolutionary algorithms
• Manual tuning	• Bayesian optimization ★ This talk's focus
	• Reinforcement learning ★
	• LLM-based agent ★

Contributions of new methods proposed in my work:

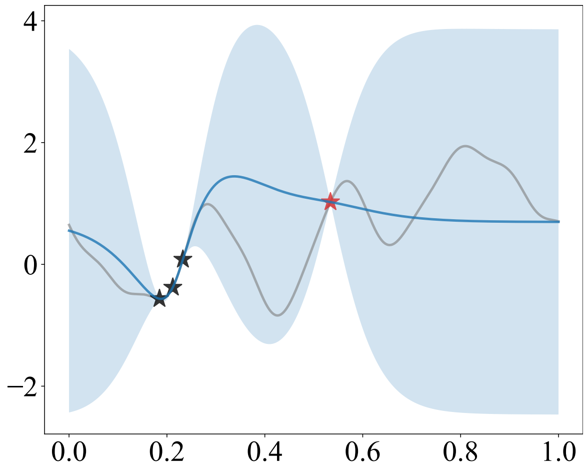
1. Theory-informed decision rules

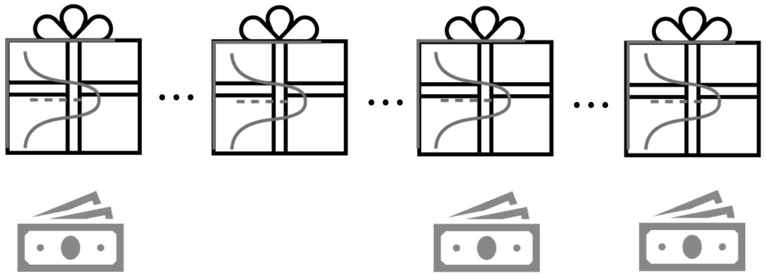
2. Competitive empirical performance

3. Principled use of information for better decisions

★ New methods under this umbrella

Part I (Recent):
Bayesian Optimization via Gittins Index
Design Principle





Bayesian Optimization

$f \sim$ Stochastic Process

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

Black-box function

Bayesian Optimization

Time 0

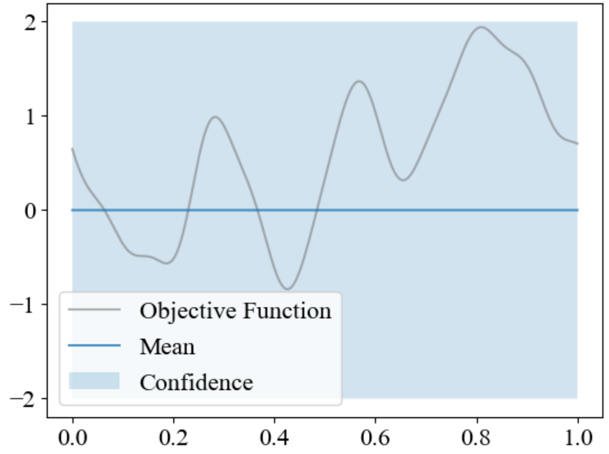
$f \sim$Stochastic Process

x $\longrightarrow$

$\longrightarrow$ $f(x)$

Black-box function

Model belief



Probabilistic model

(e.g., Gaussian process)

Bayesian Optimization

Time t

$x_1, \ldots, x_t \longrightarrow$

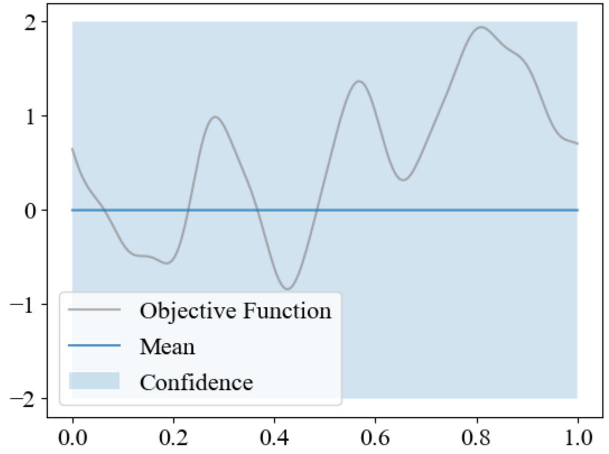
Black-box function

$\longrightarrow$

$f \sim$Stochastic Process

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

Model belief



Probabilistic model
(e.g., Gaussian process)

Bayesian Optimization

Time t

$x_1, \dots, x_t$ $\longrightarrow$

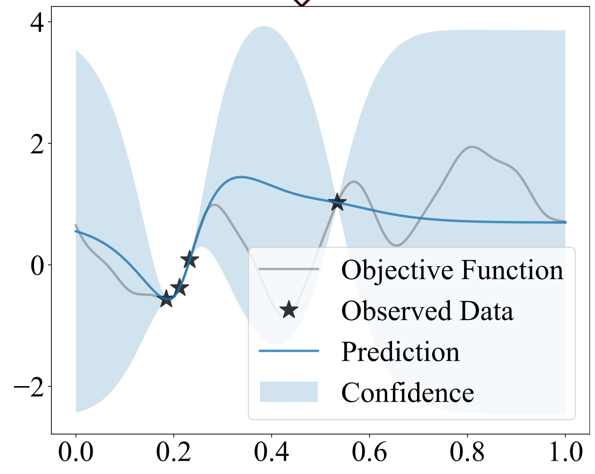
Black-box function

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

$f \sim$Stochastic Process

Update belief

(Bayes' rule)



Probabilistic model

(e.g., Gaussian process)

Bayesian Optimization

Time t

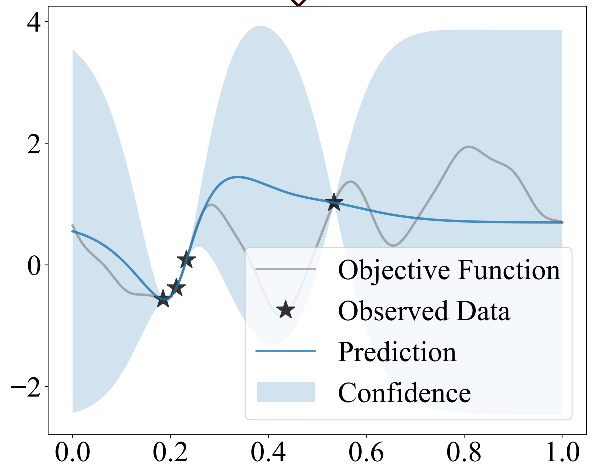
$x_1, \ldots, x_t \longrightarrow$

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

Black-box function

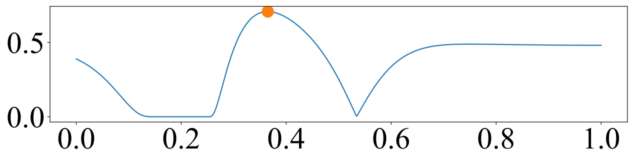
Update belief

(Bayes' rule)



Probabilistic model

(e.g., Gaussian process)



scoring (worth of each point

Decision rule

(e.g., EI, UCB, TS)

Bayesian Optimization

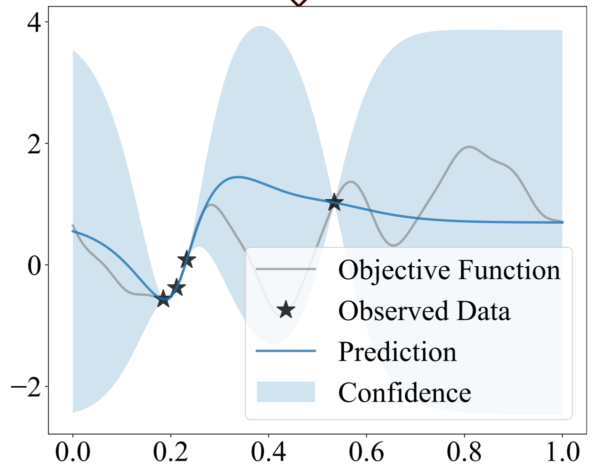
Time t

$x_1, \ldots, x_t \longrightarrow$

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

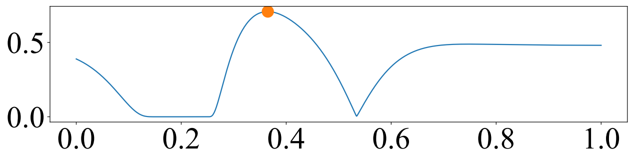
Black-box function

Update belief
(Bayes' rule)



Probabilistic model
(e.g., Gaussian process)

scoring (worth of each point



Decision rule
(e.g., EI, UCB, TS)

Decide next input
(gradient-based optimization)

Bayesian Optimization

Time $t + 1$

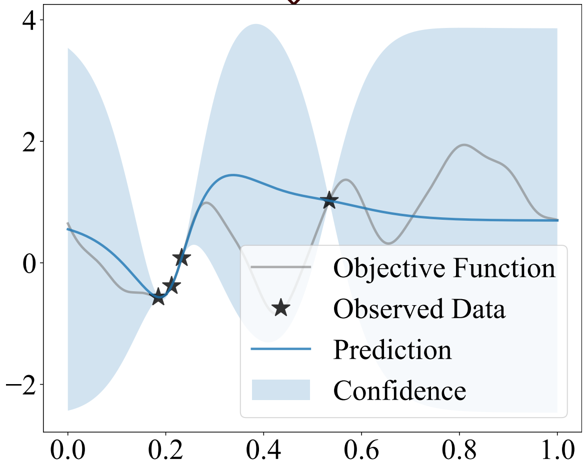
$x_1, \dots, x_{t+1} \longrightarrow$

Black-box function

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

Update belief

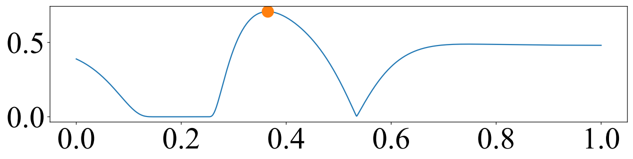
(Bayes' rule)



Probabilistic model

(e.g., Gaussian process)

scoring (worth of each point



Decision rule

(e.g., EI, UCB, TS)

Decide next input

(gradient-based optimization)

Bayesian Optimization

Time $t + 1$

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

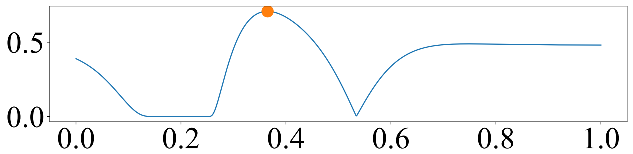
$f(x_1), \dots, f(x_{t+1})$

Black-box function

Decide next input (gradient-based optimization)

Update belief (Bayes' rule)

Probabilistic model (e.g., Gaussian process)

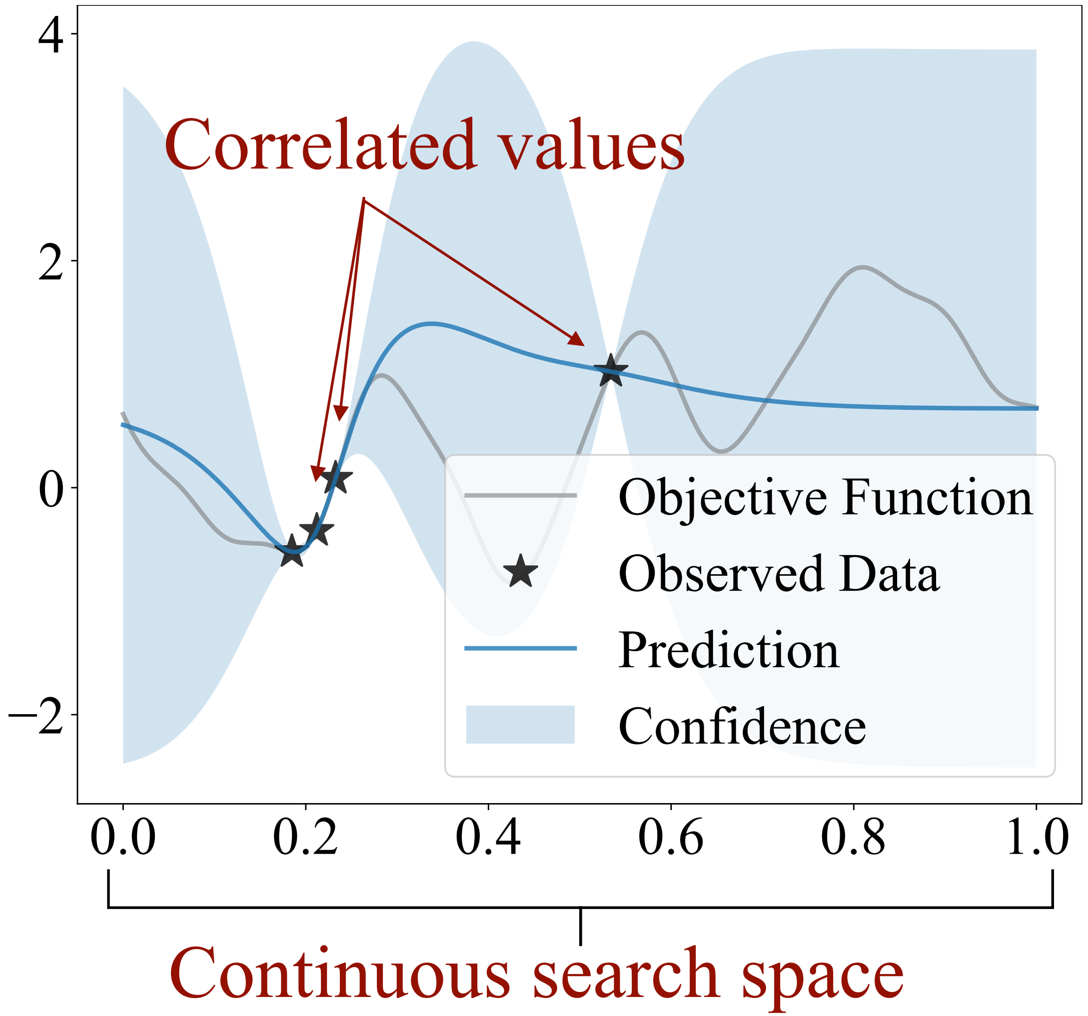


My focus

Decision rule (e.g., EI, UCB, TS)

scoring (worth of each point

Challenges in Decision Rule Design



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

Popular Decision Rule: Expected Improvement

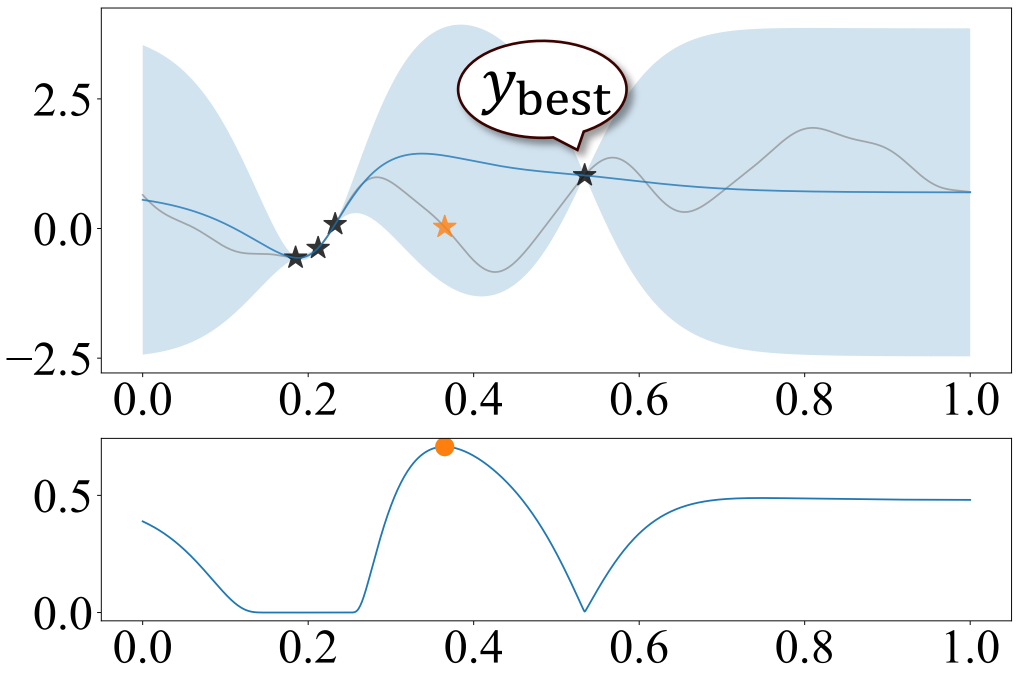
current best observed

data D

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

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

posterior distribution



One-step approximation to MDP

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

Popular Decision Rule: Expected Improvement

current best observed

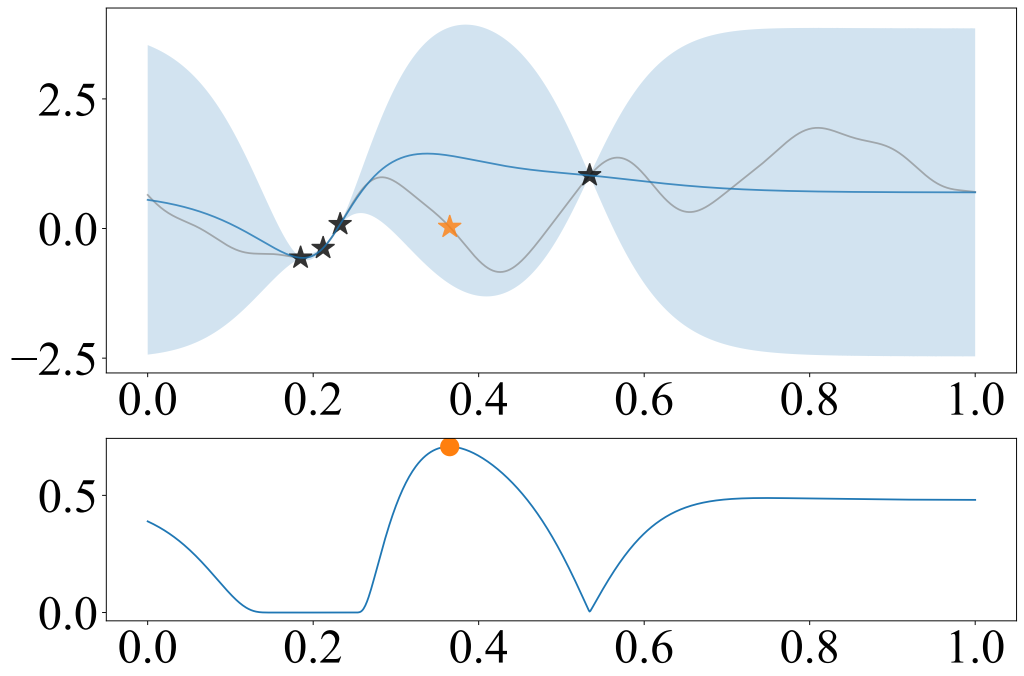
data D

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

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

posterior distribution

One-step approximation to MDP

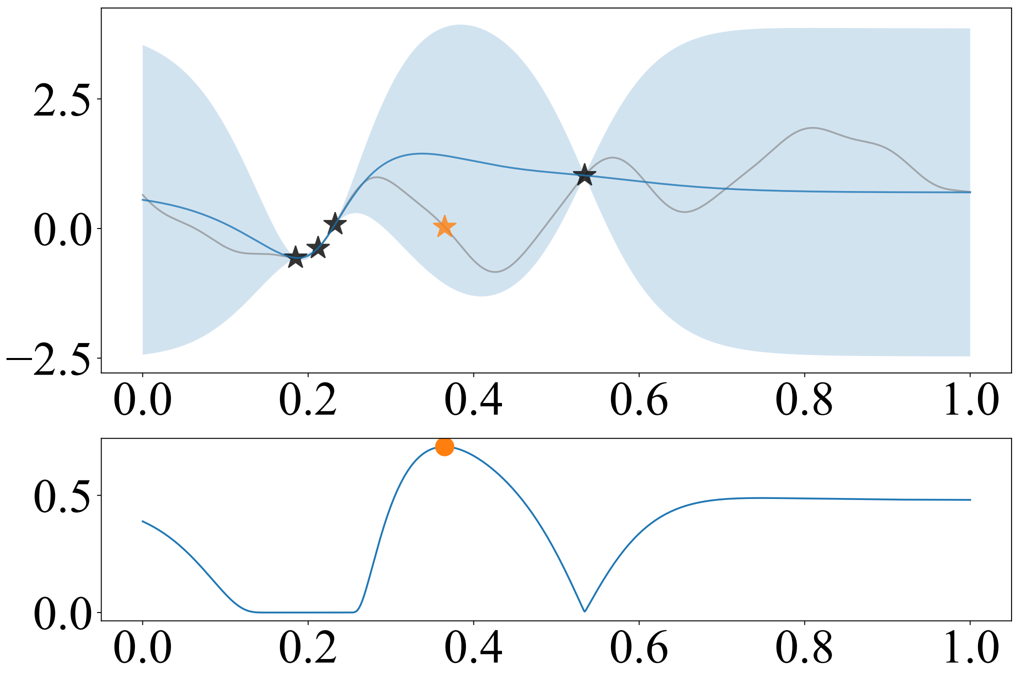


Expected improvement $\mathrm{EI}(x)$

Improvement-based design principle

Existing Design Principles

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

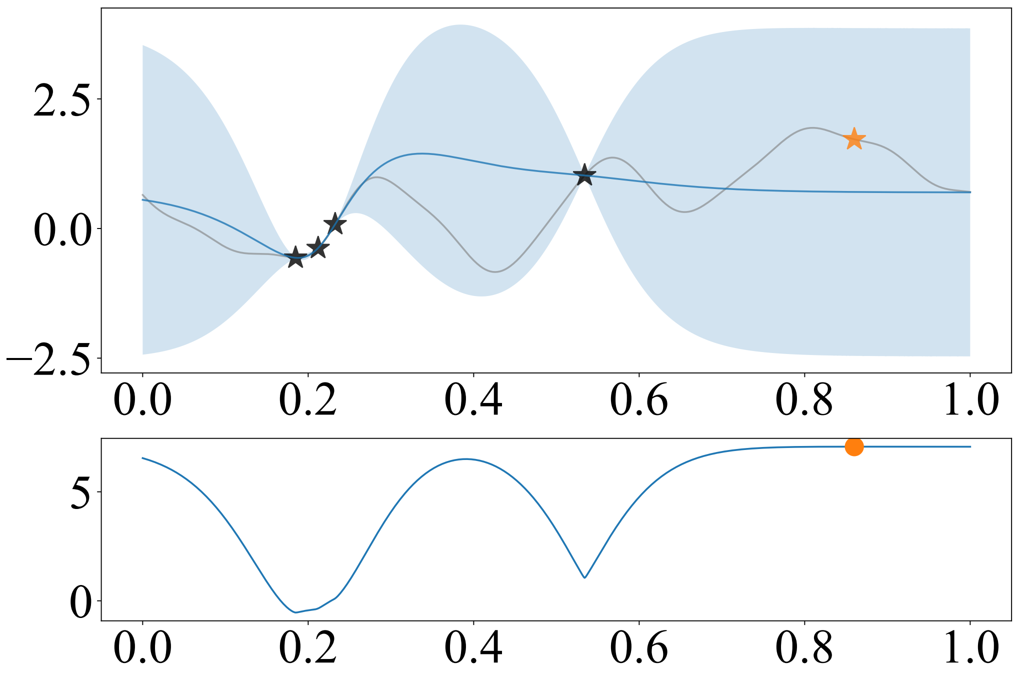


Expected improvement $\mathrm{EI}(x)$

Improvement-based design principle

New Design Principle: Gittins Index

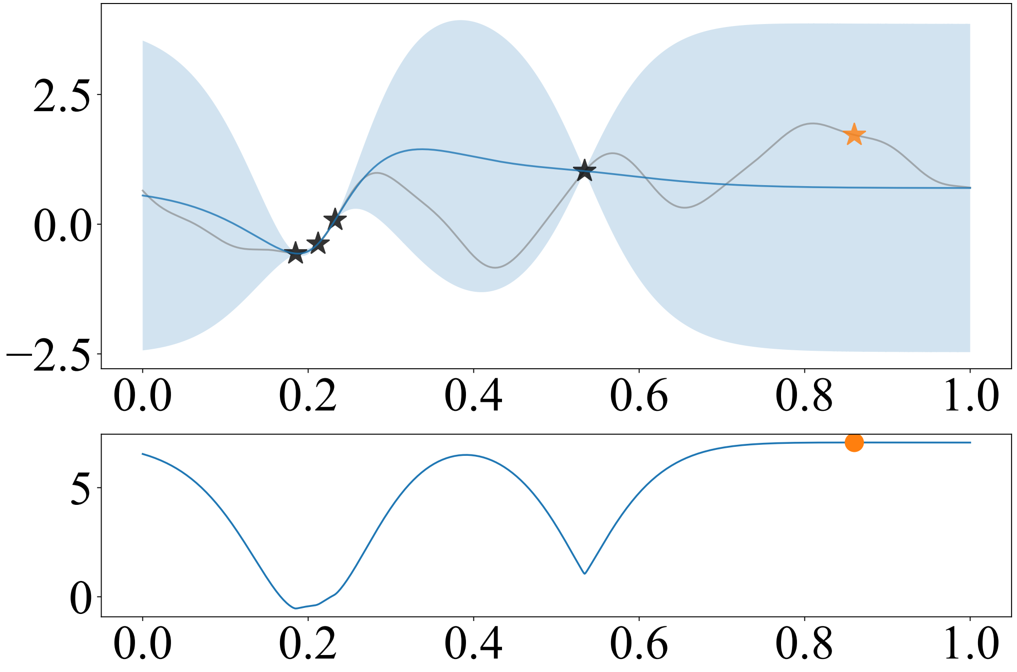
- Improvement-based (e.g., EI)
- Entropy-based
- Confidence bounds (UCB/LCB)
- Thompson sampling (TS)
- Gittins Index



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

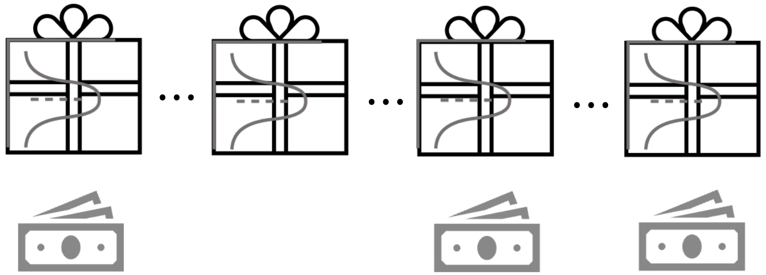


Gittins index $\mathrm{GI}(x)$

? Why another principle?

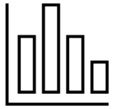
Our Contribution: Gittins Index Principle

Novel connection



Link to Pandora's Box problem
& Gittins index theory

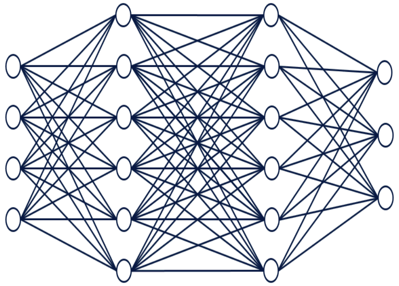
Principled decision rules

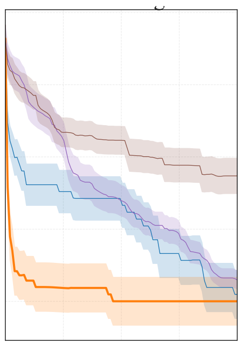
 Varying evaluation costs

 Adaptive stopping time

Unified framework for
selection and stopping

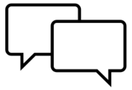
Competitive empirical performance

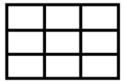


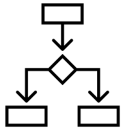


Interests from practitioners (e.g., Meta)

Future potential

 Adaptive response sampling

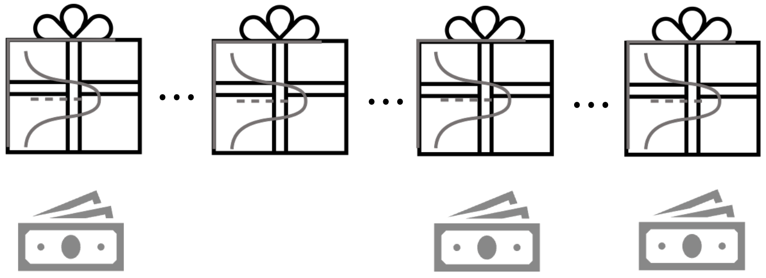
 Efficient LLM evaluation

 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

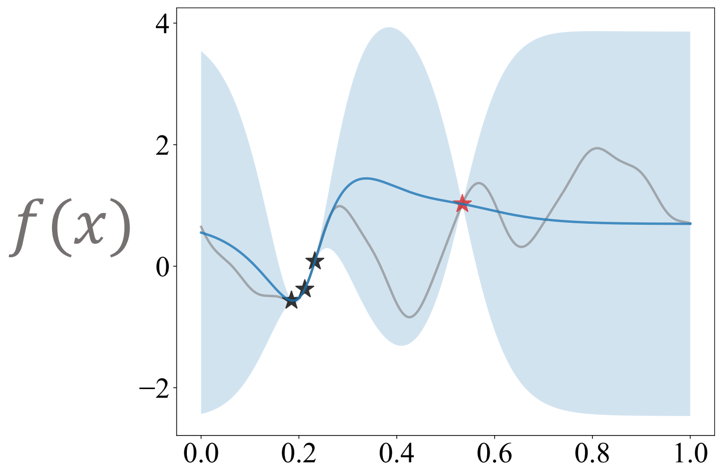
Adaptive response sampling

Efficient LLM evaluation

Chain-of-thought selection

Application to efficient LLM

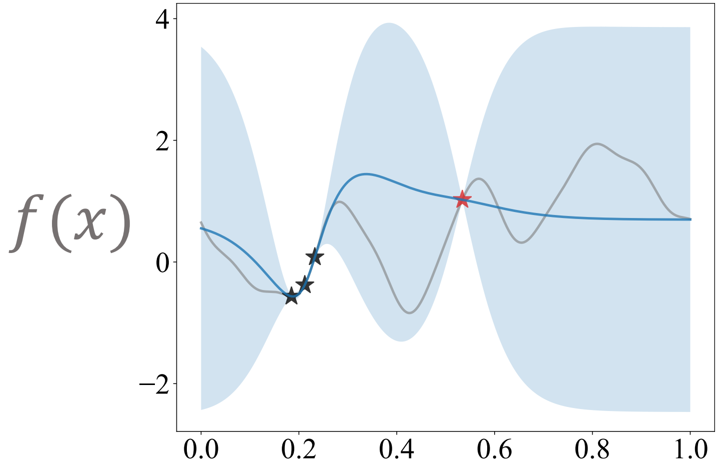
Bayesian Optimization



Continuous search space

Correlated function values

Bayesian Optimization



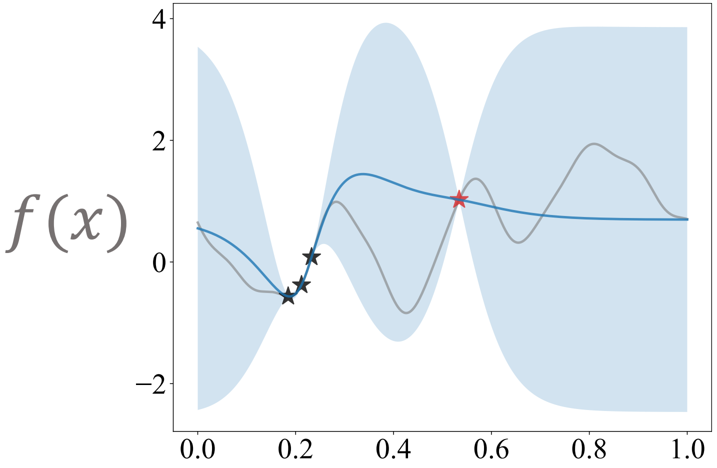
Continuous search space ⟹ Discrete

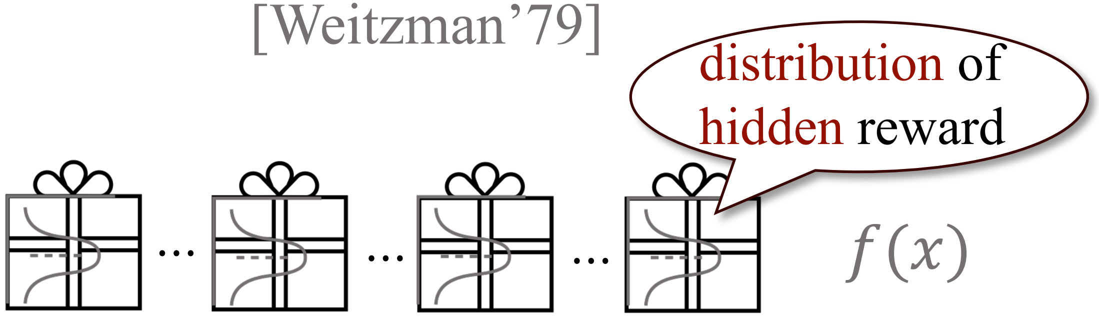
Correlated function values ⟹ Independent

Bayesian Optimization

Pandora's Box

[Weitzman'79]





Continuous search space

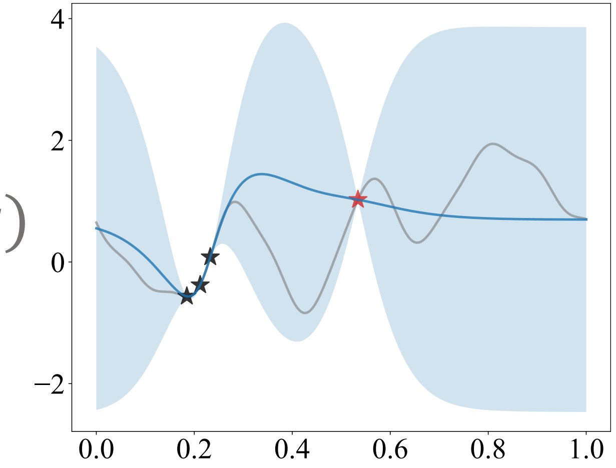
Discrete

Correlated function values

Independent

Bayesian Optimization

$f(x)$



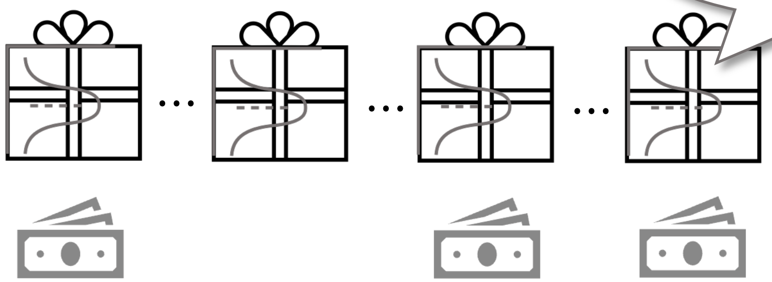
Continuous search space

Correlated function values

Pandora's Box

[Weitzman'79]

distribution of hidden reward



$f(x)$

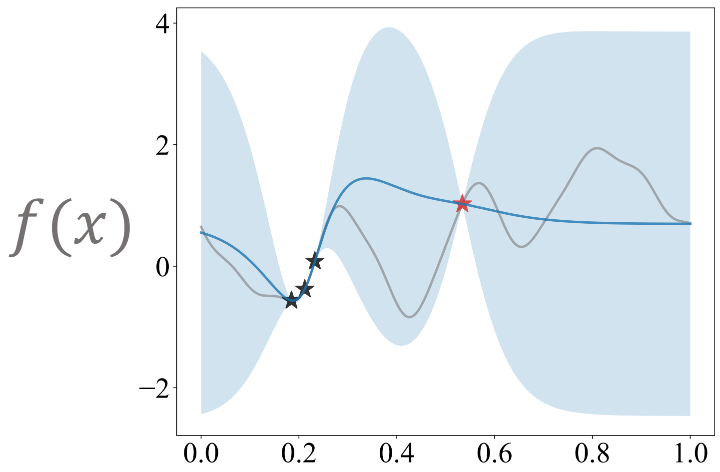
$c(x)$

costs to open

Discrete

Independent

Bayesian Optimization

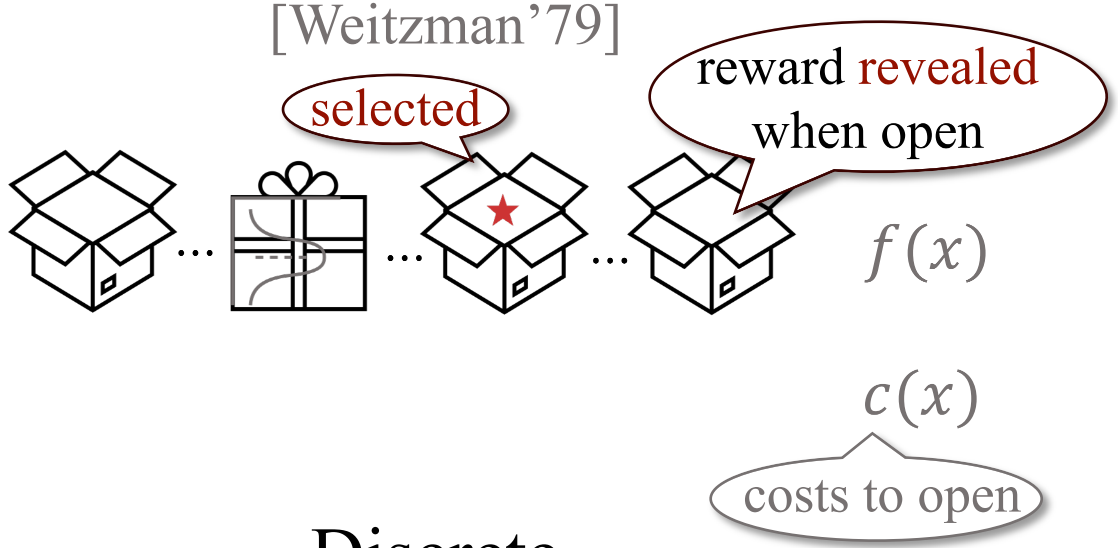


Continuous search space

Correlated function values

Pandora's Box

[Weitzman'79]



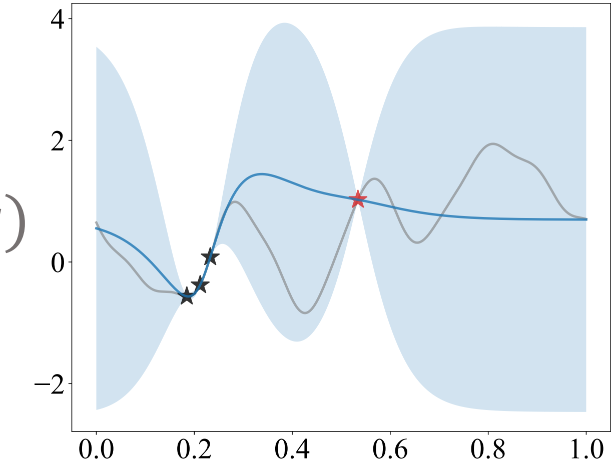
Discrete

Independent

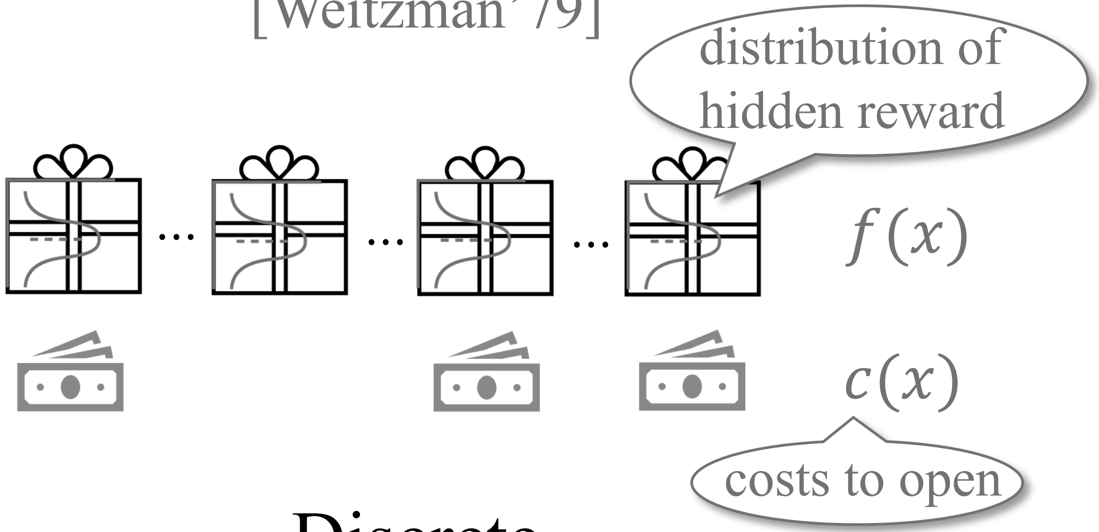
Bayesian Optimization

Pandora's Box

[Weitzman'79]



$f(x)$

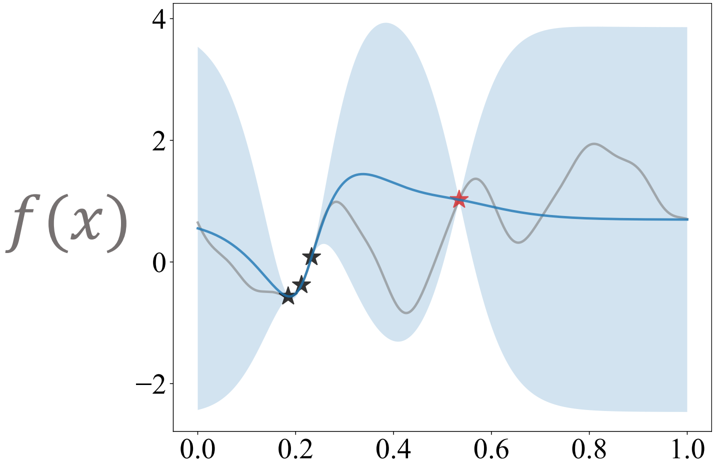


Continuous search space | Discrete

Correlated function values | Independent

Optimal policy: Gittins index

Bayesian Optimization

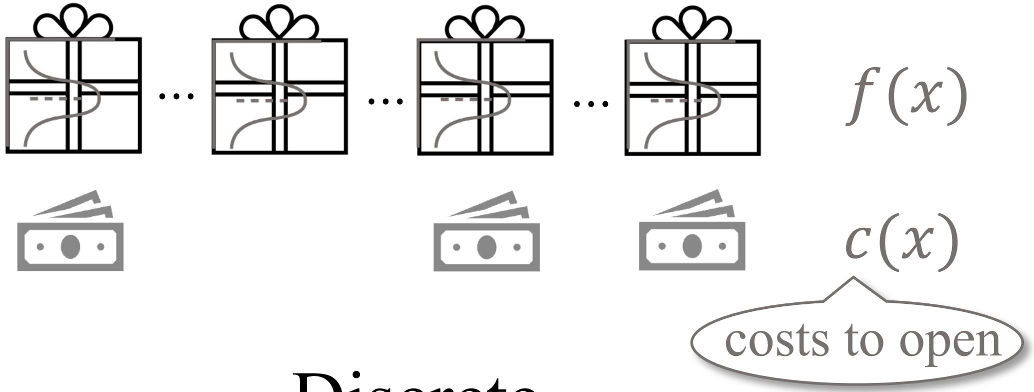


Continuous search space

Correlated function values

Pandora's Box

[Weitzman'79]



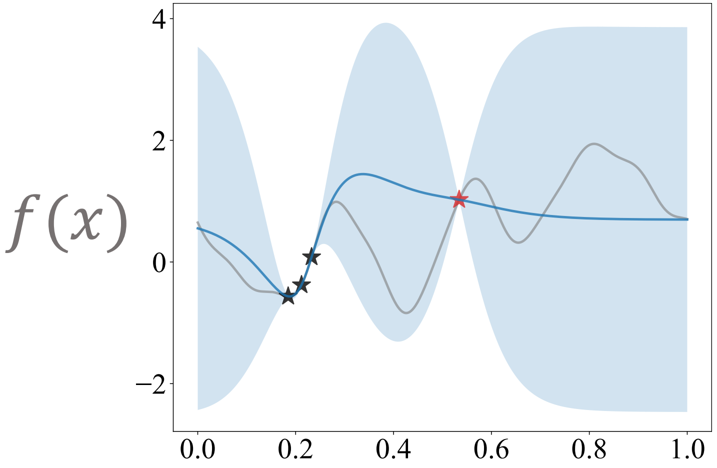
Discrete

Independent

How to translate?

$\Longleftarrow$ Optimal policy: Gittins index

Bayesian Optimization

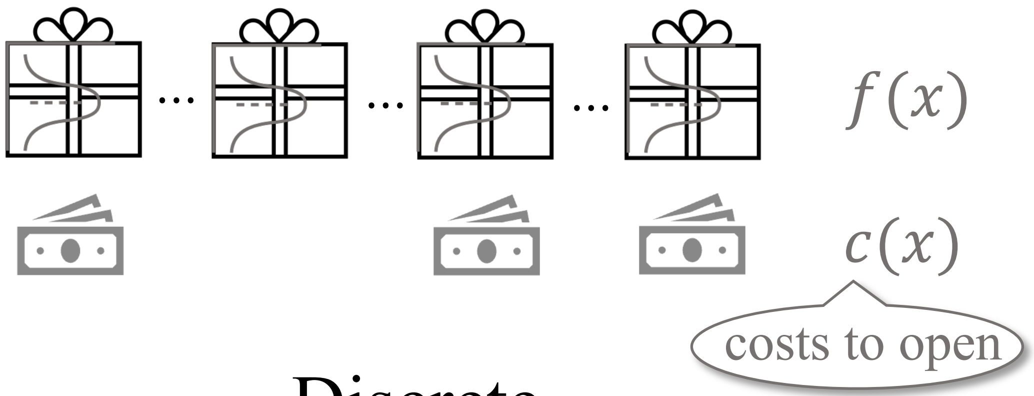


Continuous search space

Correlated function values

Pandora's Box

[Weitzman'79]



Discrete

Independent

New!

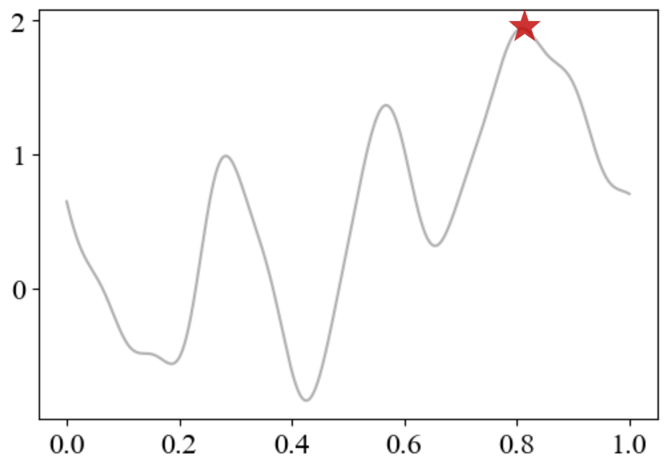
incorporate posterior

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

take continuum limit

Bayesian Optimization

$f(x)$

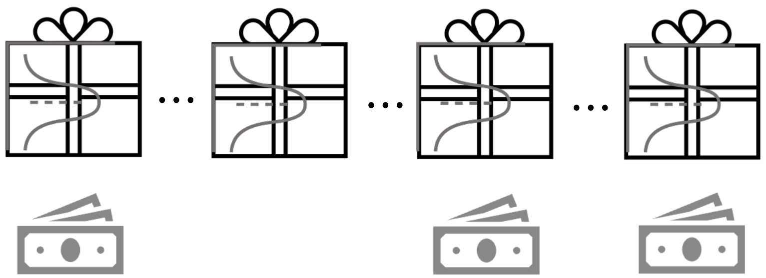


Continuous

Correlated

Pandora's Box

[Weitzman'79]



$f(x)$

$c(x)$

costs to open

Discrete

Independent

incorporate posterior

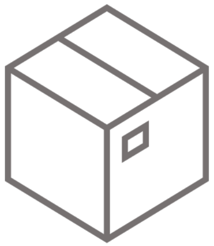
Our policy: $\mathrm{GI}_{f|D}(x; c(x))$ $\xLeftarrow{}$ Optimal policy: $\mathrm{GI}_f(x; c(x))$

take continuum limit

How to compute?

Intuition

Exploration Exploitation

$f(x)$  vs. 

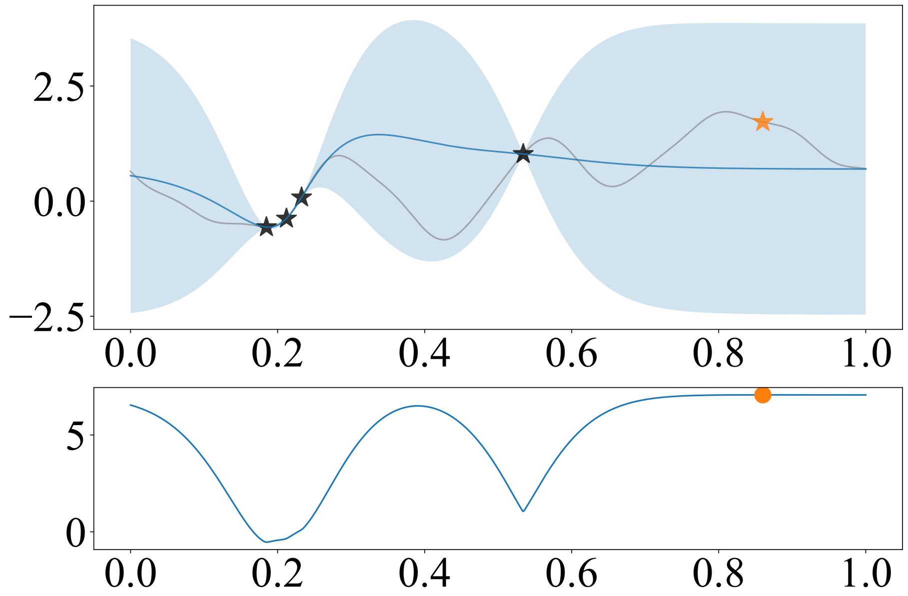
$c(x)$ 

Open closed box | Take opened box

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

Should one open box? Depend on g!

Gittins Index



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

Intuition

Exploration | Exploitation

$f(x)$ vs.

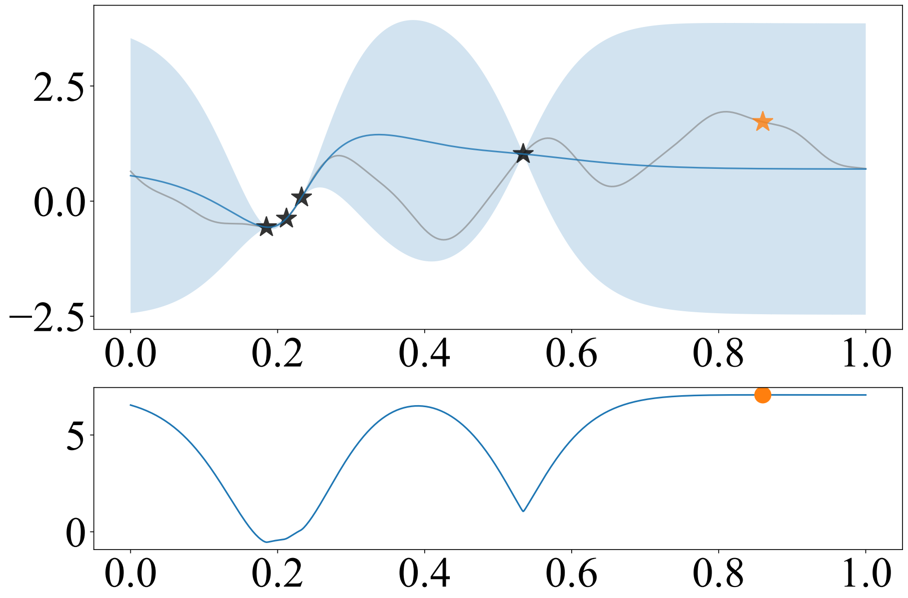
$c(x)$

Open closed box | Take opened box

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

Should one open box? Depend on g!

Gittins Index



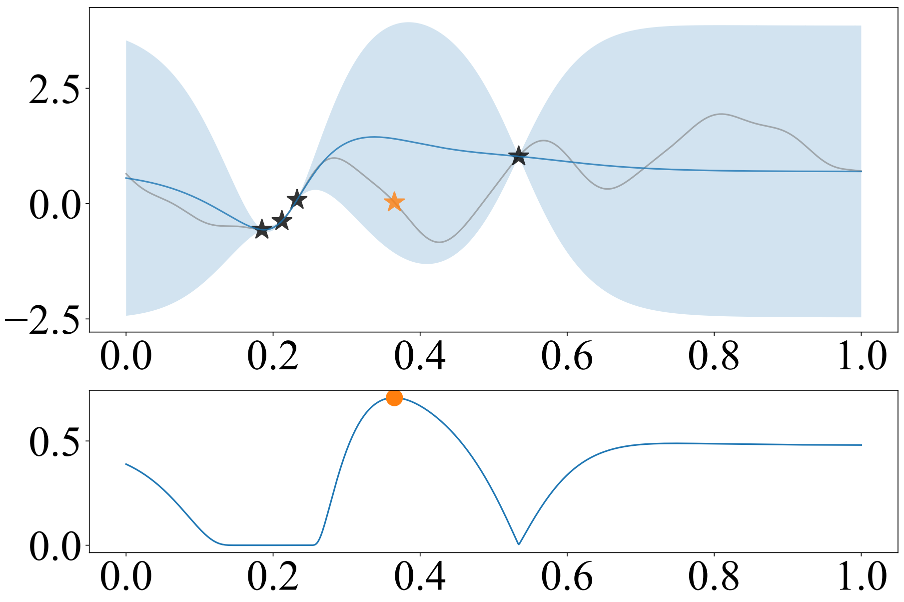
$$\mathrm{GI}_{f|D}(x; c) := \text{solution } g \text{ s.t.}$$

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

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

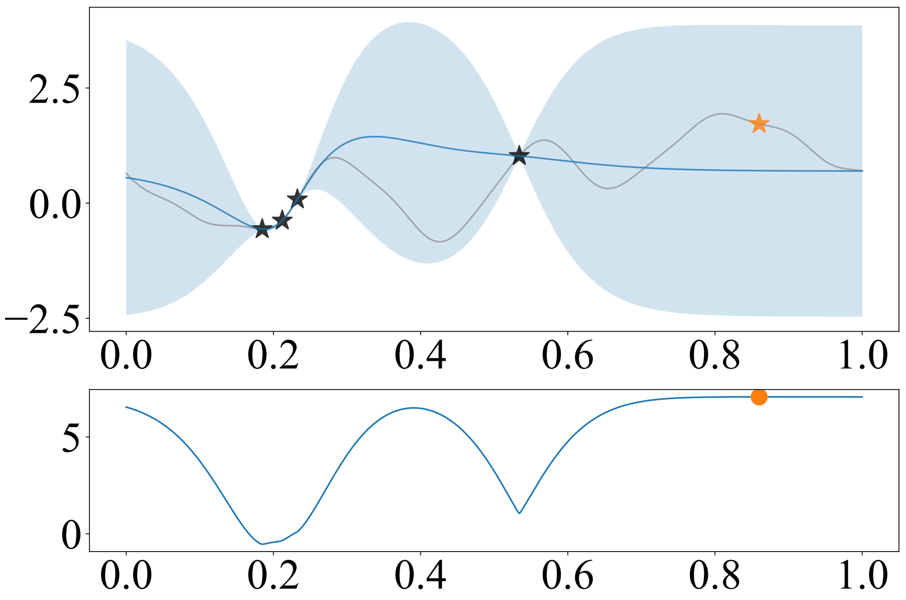
$$\Leftrightarrow \underbrace{\mathbb{E}[\max(f(x) - g, 0) \mid D]}_{\mathrm{EI}_{f|D}(x;\, g)} = c(x)$$

Expected Improvement

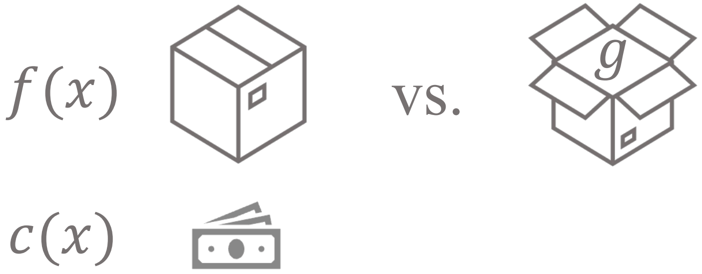


$$\mathrm{EI}_{f|D}(x) := \mathbb{E}[\max(f(x) - y_{\text{best}}, 0) \mid D]$$

Gittins Index

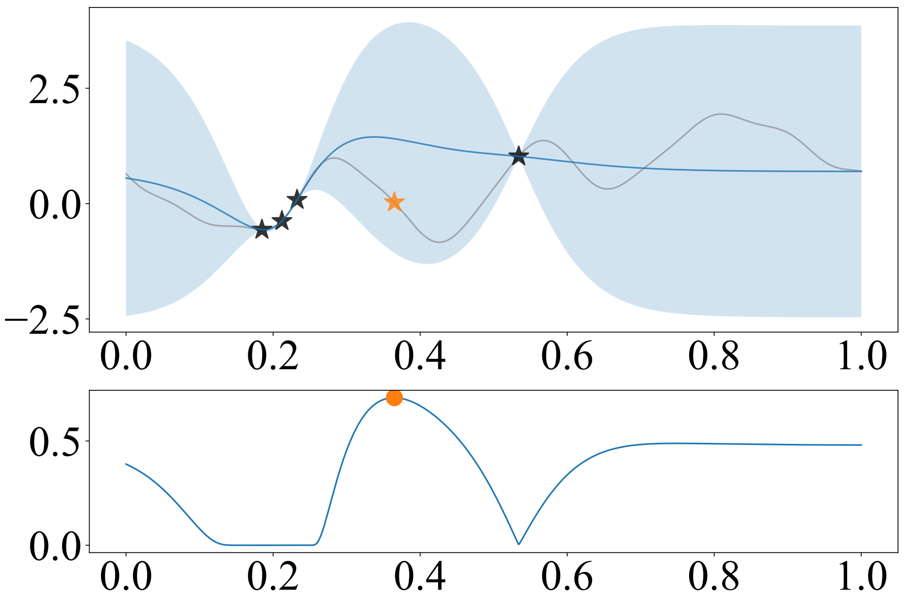


$$\mathrm{GI}_{f|D}(x; c) := \text{solution } g \text{ s.t. } \mathrm{EI}_{f|D}(x; g) = c(x)$$
$$\text{where } \mathrm{EI}_{f|D}(x; g) := \mathbb{E}[\max(f(x) - g, 0) \mid D]$$

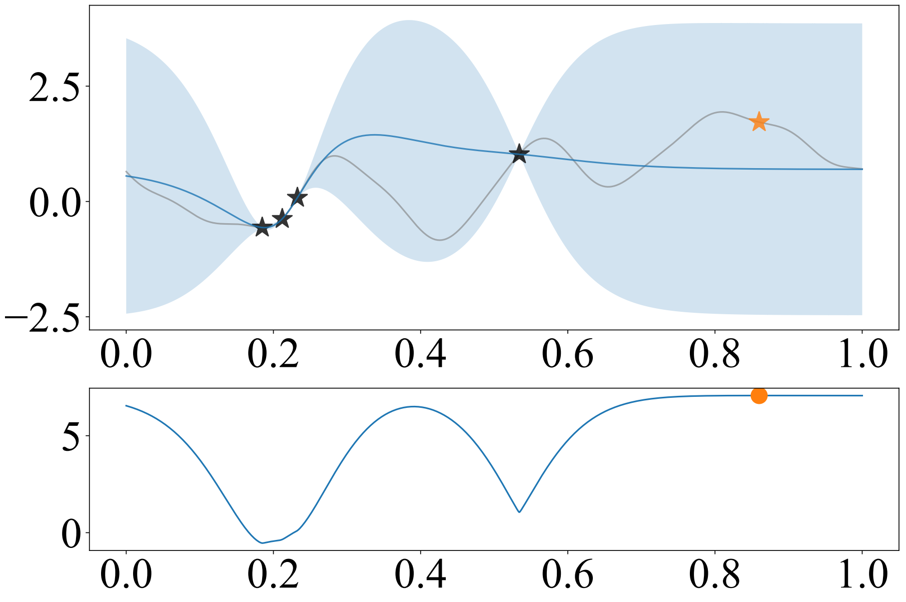


Exploration Exploitation

Expected Improvement



Gittins Index



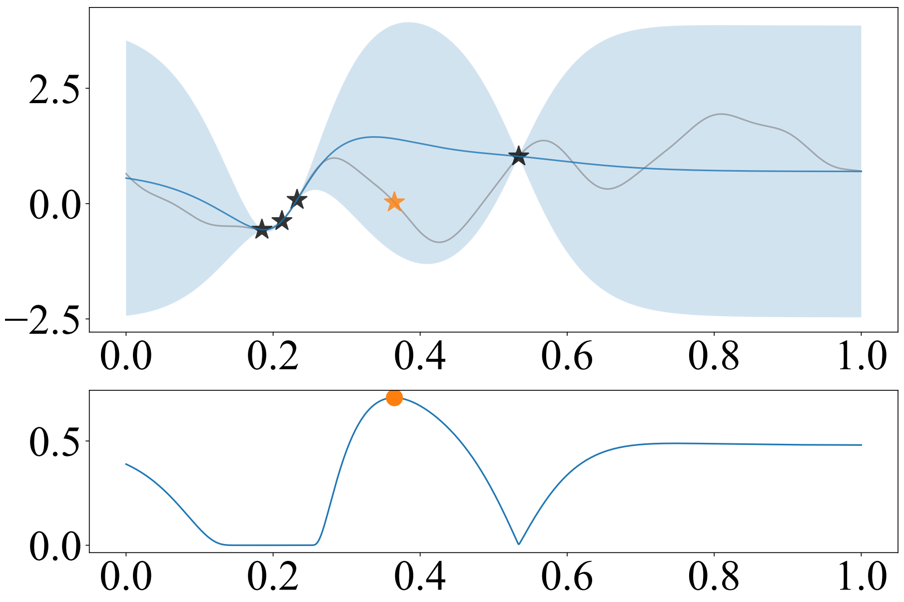
$\mathrm{EI}_{f|D}(x) := \mathbb{E}[\max(f(x) - y_{\mathrm{best}}, 0) \mid D]$

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

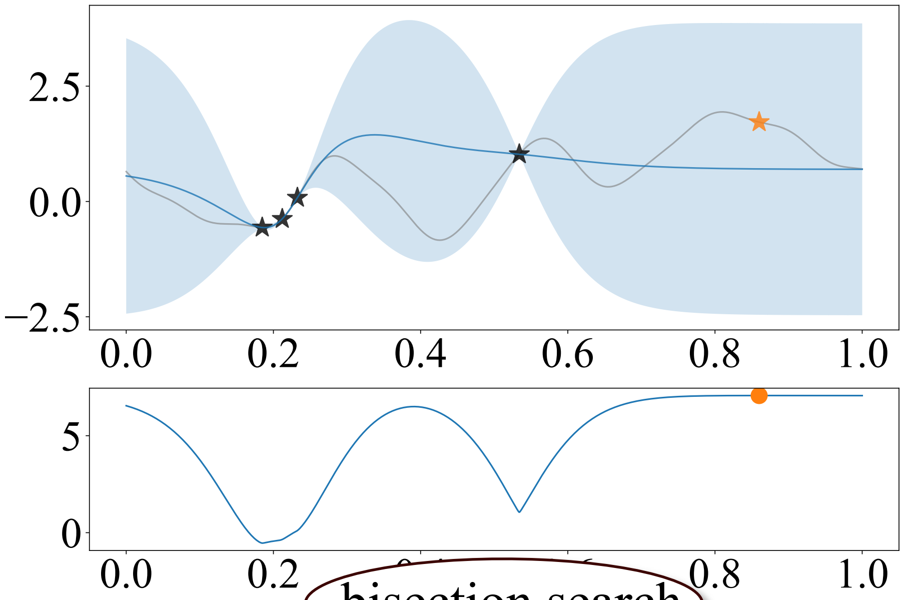
Temporal simplification to MDP
(One-step)

Spatial simplification to MDP

Expected Improvement



Gittins Index



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

bisection search

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

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

analytical expression

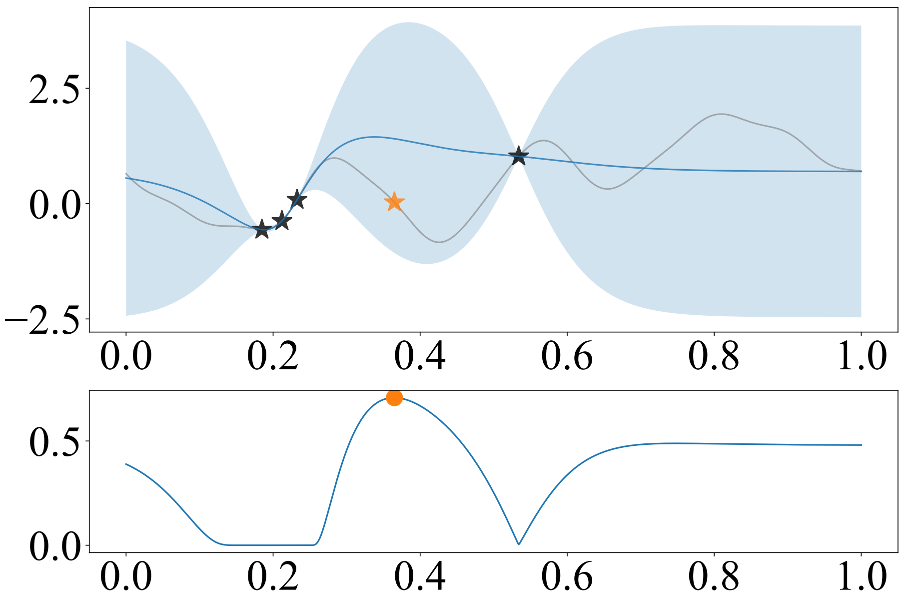
Temporal simplification to MDP

Spatial simplification to MDP

Both are principled and easy-to-compute!

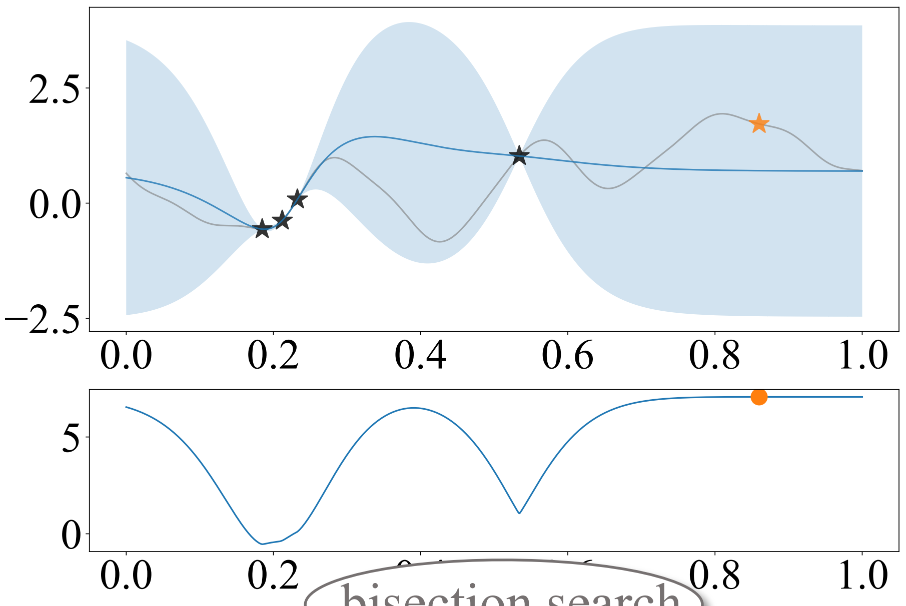
"Cost-aware Bayesian Optimization via the Pandora's Box Gittins Index." NeurIPS'24.

Expected Improvement



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

Gittins Index



bisection search

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

Google DeepMind

FunBO: Discovering new acquisition functions for Bayesian Optimization with FunSearch

hard to discover GI

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

Adaptive response sampling

Efficient LLM evaluation

Chain-of-thought selection

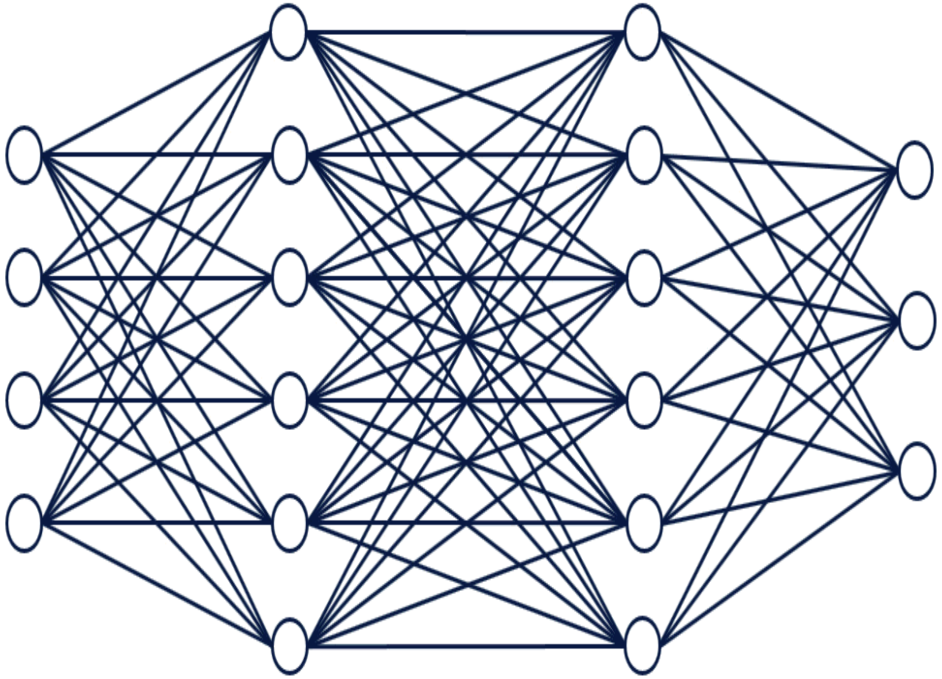
Application to efficient LLM

Under-explored Information 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

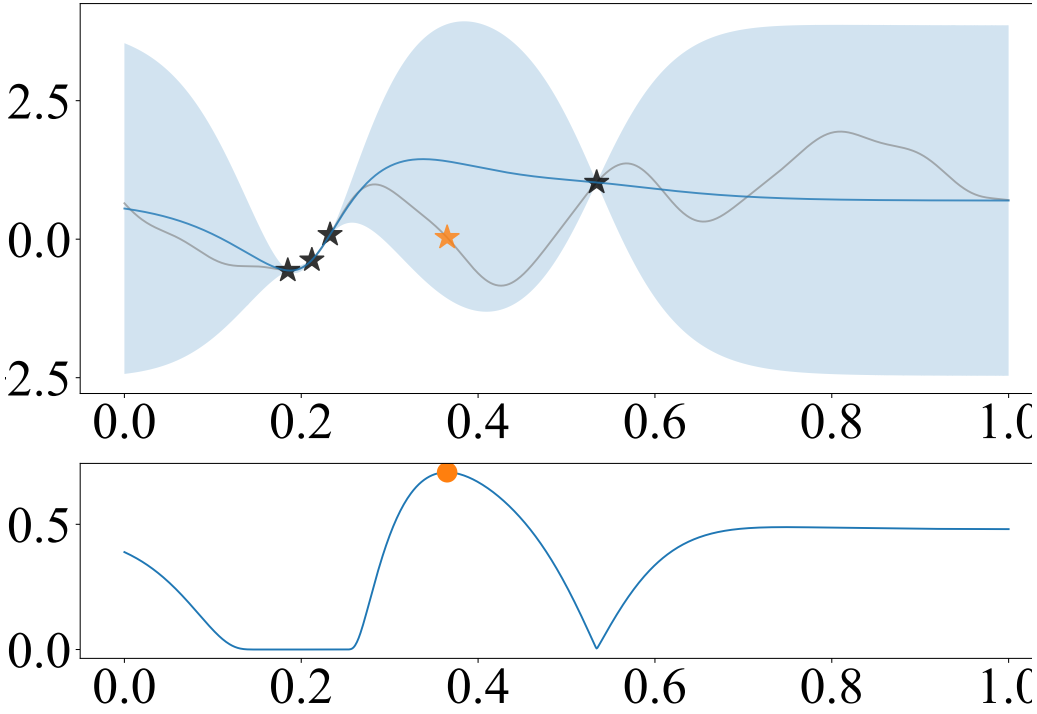
$$\tau\text{: } \mathrm{EI}(x_\tau) \leq \theta$$

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

Which threshold?

Observable multi-stage feedback

?



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

Under-explored Information 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

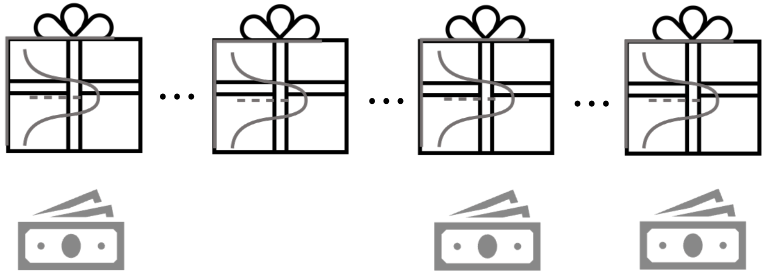
Smart stopping time

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

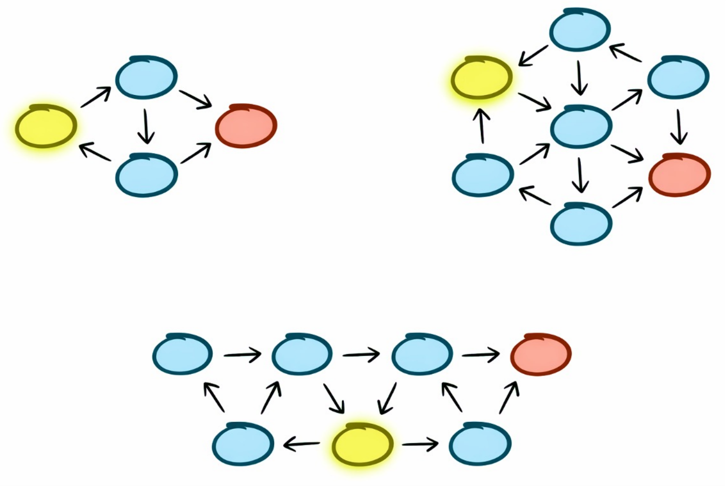
Features in Pandora's box

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

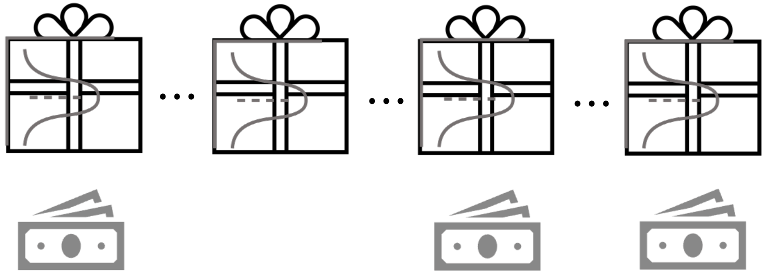
Features in Pandora's box

Observable multi-stage feedback

Features in Markov chain selection

New design principle: Gittins index

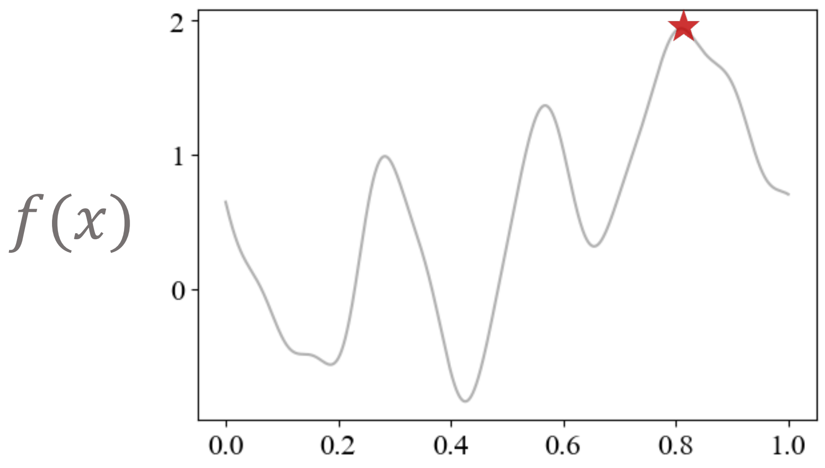
Optimal in related sequential decision problems



"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

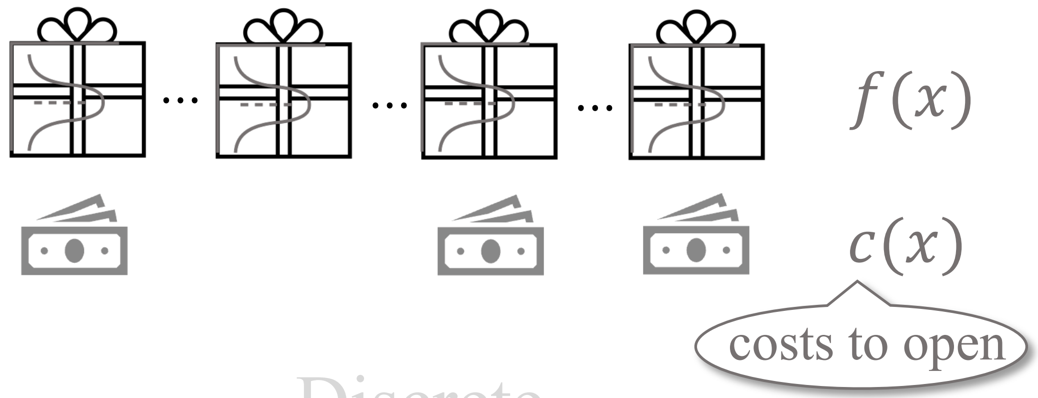
Cost-unaware

Fixed-iteration

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

Pandora's Box

[Weitzman'79]



Discrete

Independent

Cost-aware

Flexible-stopping

Optimal policy: $\mathrm{GI}_f(x;c)$

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

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

Which threshold?

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

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

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

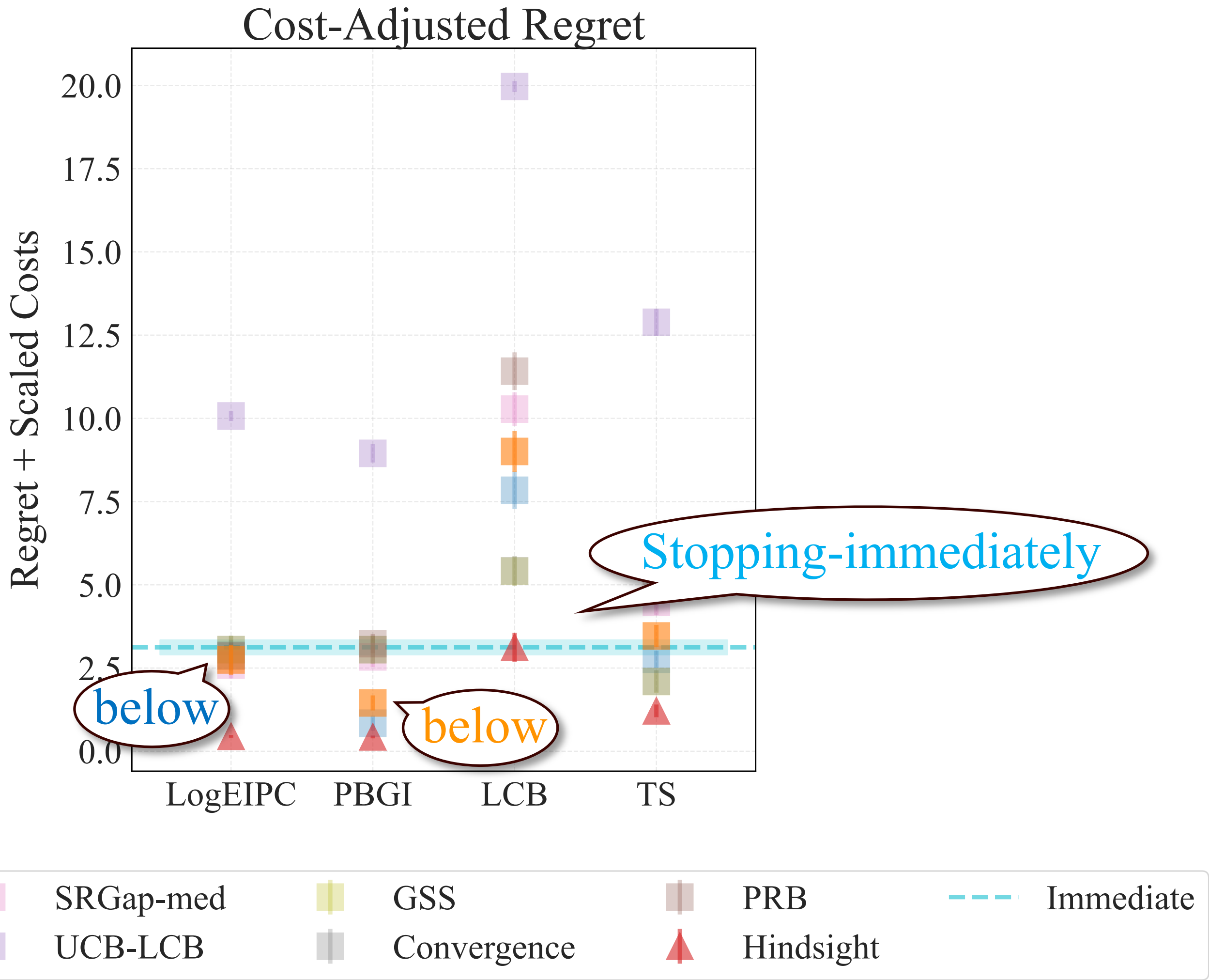
Implication:

- 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

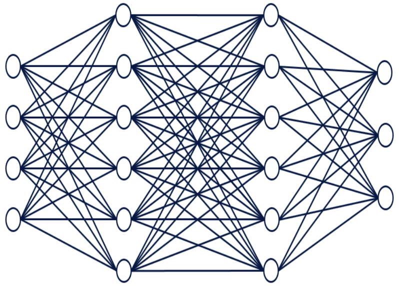
Principled decision rules

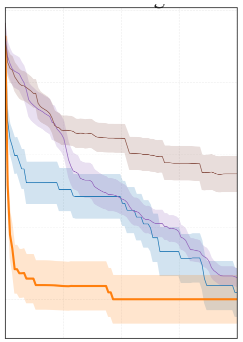
Varying evaluation costs

Adaptive stopping time

Unified framework for
selection and stopping

Competitive empirical performance





Interests from practitioners (e.g., Meta)

Future potential

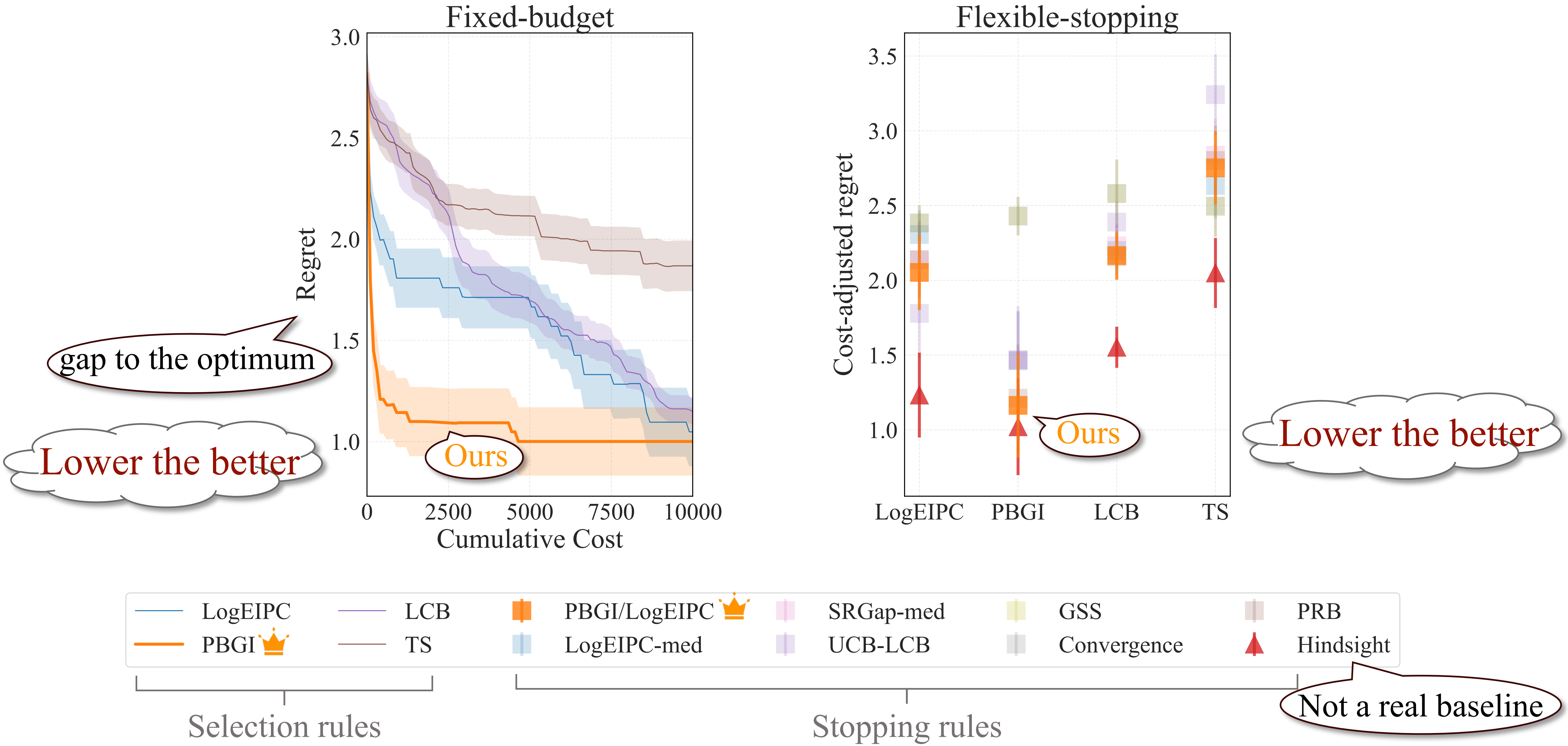
Adaptive response sampling

Efficient LLM evaluation

Chain-of-thought selection

Application to efficient LLM

Gittins Index vs Baselines on AutoML Benchmark



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

Adaptive response sampling

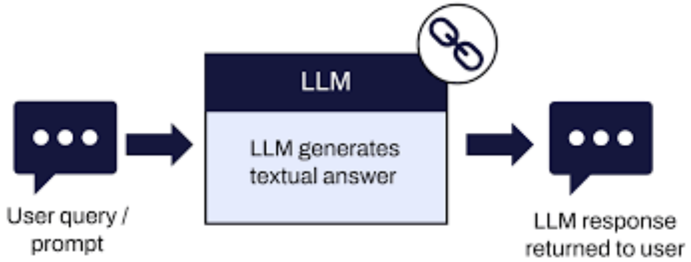
Efficient LLM evaluation

Chain-of-thought selection

Application to efficient LLM

Black-Box Optimization in Emerging AI (LLM)

LLM inference time optimization:

LLM response  Reward

Method: meta-learning + Gittins

Black-Box Optimization in Emerging AI (LLM)

LLM inference-time optimization:

LLM response → → Reward

Method: meta-learning + Gittins

LLM evaluation (best-prompt identification):

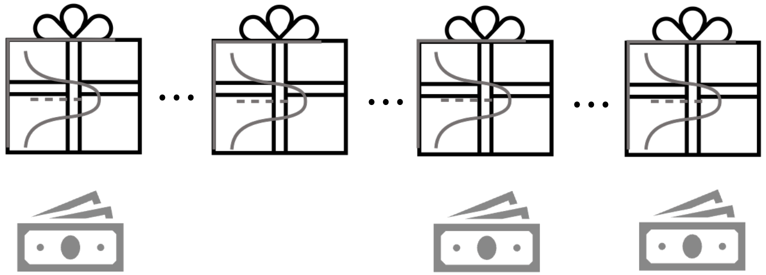
Prompt template for each LLM →

	Zero-shot	Few-shot	CoT	RAG	Revise
ChatGPT (GPT-4.1)	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Claude 3.5 Sonnet	★★★★★	★★★★★	★★★★★		★★★★★
Gemini 1.5 Pro	★★★★★	★★★	?	★★★★	★★★★★
deepseek	?	?	★★★★	?	★★★★
Llama 3.1-70B	?	?	★★★★	?	★★★
Mistral Large	?	?	★★★	?	★

→ Average performance of (LLM, prompt) pairs

Method: BayesOpt + Gittins

Novel connection



Link to Pandora's Box problem
& Gittins index theory

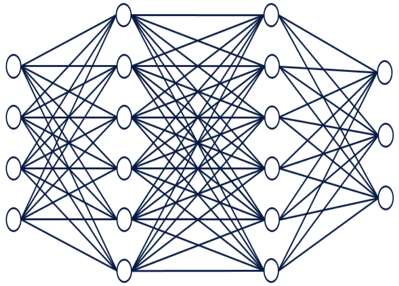
Principled decision rules

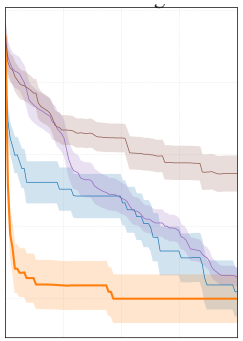
Varying evaluation costs

Adaptive stopping time

Unified framework for
selection and stopping

Competitive empirical performance





Interests from practitioners (e.g., Meta)

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

Future potential

Adaptive response sampling

Efficient LLM evaluation

Chain-of-thought selection

Application to efficient LLM

"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.