

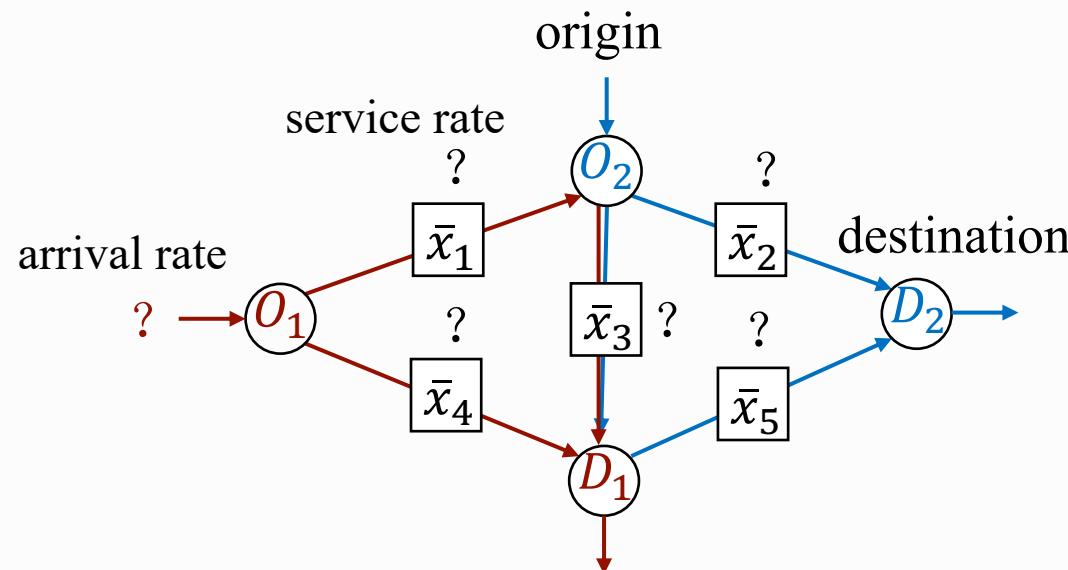
Data-Efficient Decision-making under Uncertainty: AutoML and Beyond

Qian Xie 谢倩 (Cornell ORIE)

Job talk for engineering school audiences

Research Overview: Decision-making under Uncertainty

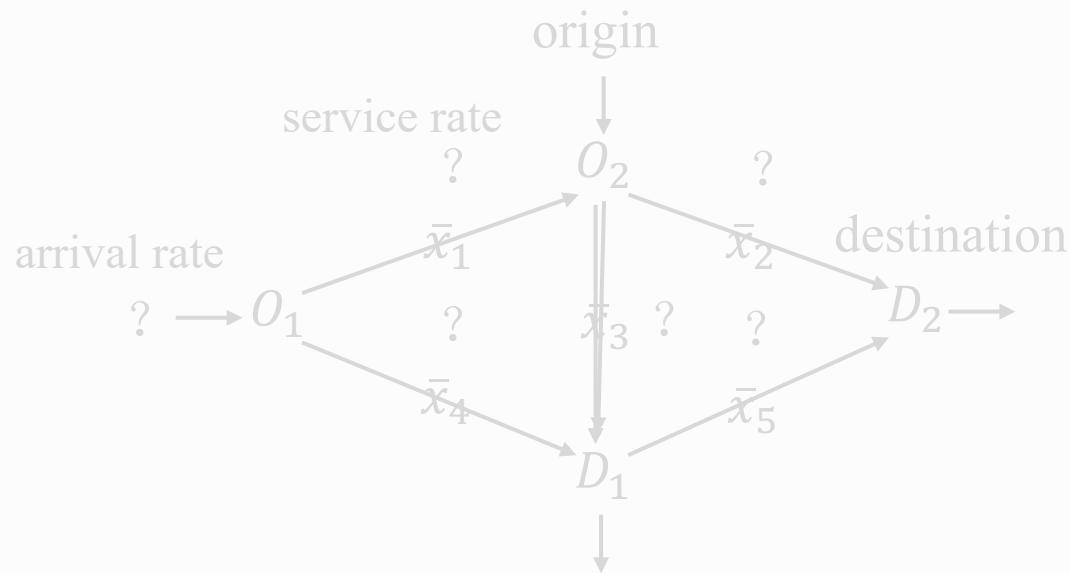
- Earlier: stochastic control (actions chosen to regulate system)
 - Unknown system rates (robust/model-free control) [IEEE TCNS]



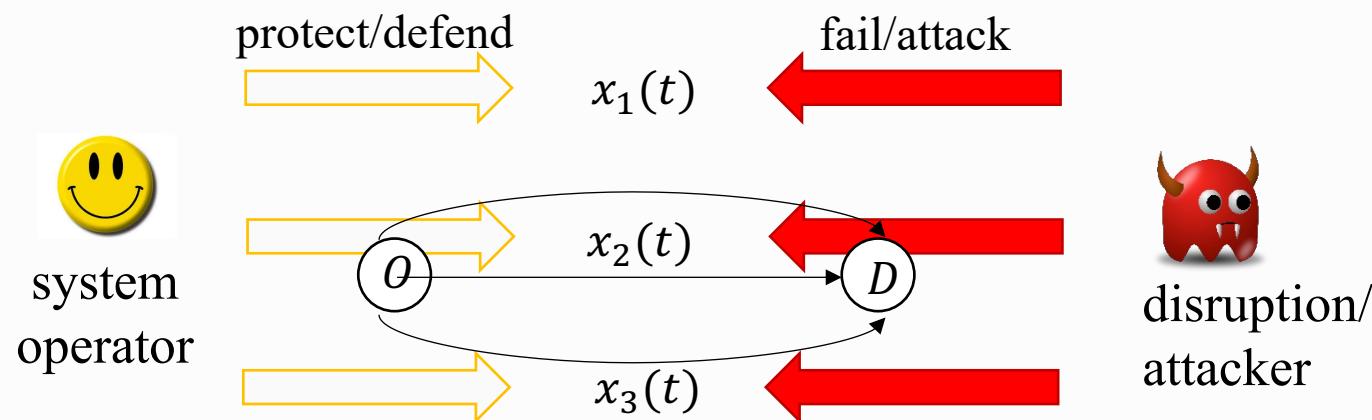
“Stabilizing Queuing Networks with Model Data-Independent Control.” IEEE TCNS.

Research Overview: Decision-making under Uncertainty

- Earlier: stochastic control (actions chosen to regulate system)
 - Unknown system rates (robust/model-free control) [IEEE TCNS]
 - Random failure/attack realizations (optimal & robust control) [Automatica]



“Stabilizing Queuing Networks with Model Data-Independent Control.” IEEE TCNS.



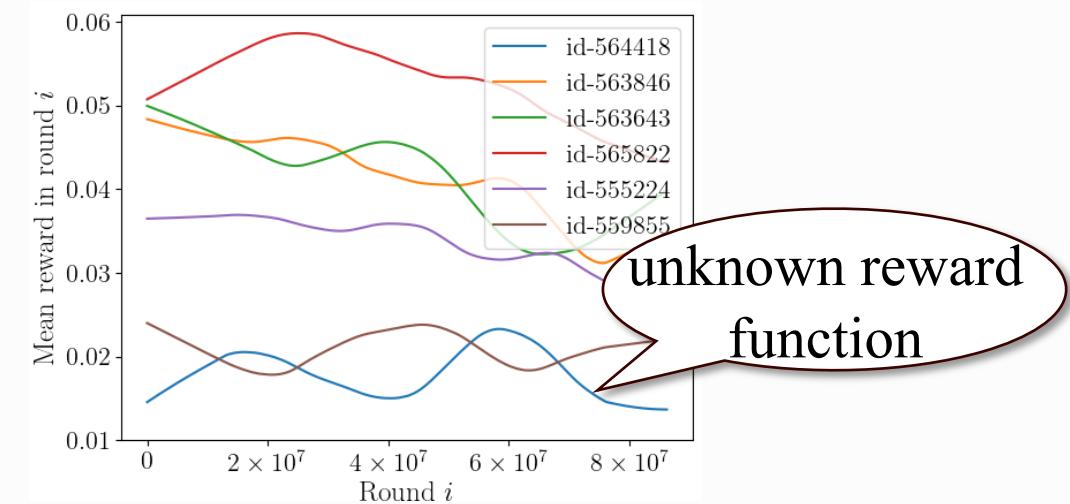
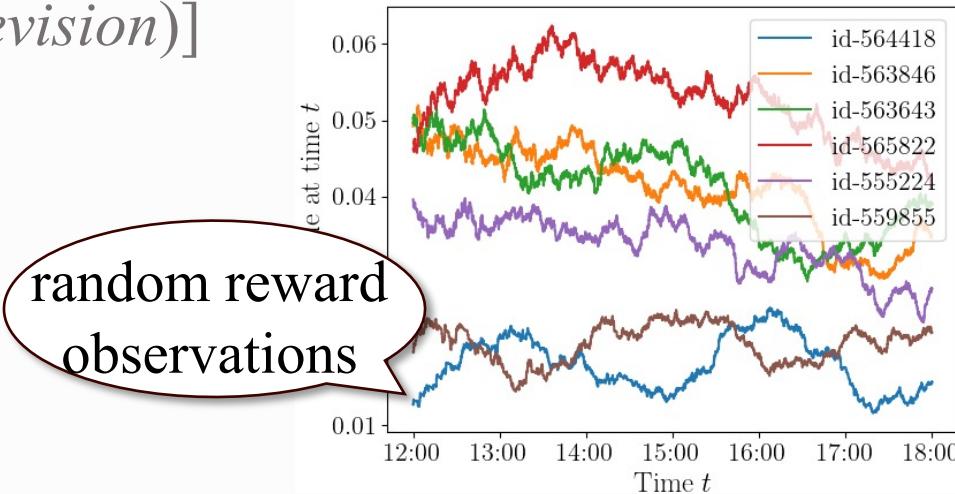
“Cost-aware Defense for Parallel Server Systems against Reliability and Security Failures.” Automatica.

Research Overview: Decision-making under Uncertainty

- Earlier: stochastic control (actions chosen to regulate system)
 - Unknown system rates (robust/model-free control) [IEEE TCNS]
 - Random failure/attack realizations (optimal & robust control) [Automatica]
- Recent: **learning-based decision-making** (actions chosen to reduce uncertainty due to limited feedback)

Research Overview: Decision-making under Uncertainty

- Earlier: stochastic control (actions chosen to regulate system)
 - Unknown system rates (robust/model-free control) [IEEE TCNS]
 - Random failure/attack realizations (optimal & robust control) [Automatica]
- Recent: learning-based decision-making (actions chosen to reduce uncertainty due to limited feedback)
 - Unknown reward function; random observations (bandits) [ICML'23 + OR (*major revision*)]



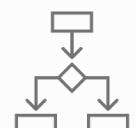
“Smooth Nonstationary Bandits.” ICML’23

Research Overview: Decision-making under Uncertainty

- Recent: learning-based decision-making (actions chosen to reduce uncertainty due to limited feedback)
 - Unknown reward function; random observations (bandits) [ICML'23 + OR]
 - Unknown objective functions (**black-box** optimization) with
 - █ varying evaluation costs [NeurIPS'24 + INFORMS DM Paper Finalist]
 - █ adaptive stopping time [Under review + AutoML'25 (non-archival)]
 - █ multi-stage feedback [To be submitted]
 - █ (i) multi-source environment information [NeurIPS'25 LAW]



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



"Bayesian-optimal Decision-making under Cost-aware Multi-stage Feedback via the Gittins Index."



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



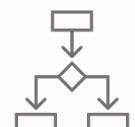
"LLM-Driven Composite NAS for Multi-Source RL State Encoding." NeurIPS'25 LAW.

Research Overview: Decision-making under Uncertainty

- Recent: learning-based decision-making (actions chosen to reduce uncertainty due to limited feedback)
 - Unknown reward function; random observations (bandits) [ICML'23 + ORI]
 - Unknown objective functions (black-box optimization) with This talk's focus
 - varying evaluation costs [NeurIPS'24 + INFORMS DM Paper Finalist]
 - adaptive stopping time [Under review + AutoML'25 (non-archival)]
 - multi-stage feedback [To be submitted]
 - (i) multi-source environment information [NeurIPS'25 LAW]



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



"Bayesian-optimal Decision-making under Cost-aware Multi-stage Feedback via the Gittins Index."



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



"LLM-Driven Composite NAS for Multi-Source RL State Encoding." NeurIPS'25 LAW.

Motivation: World of Optimization under Uncertainty

ML model training:

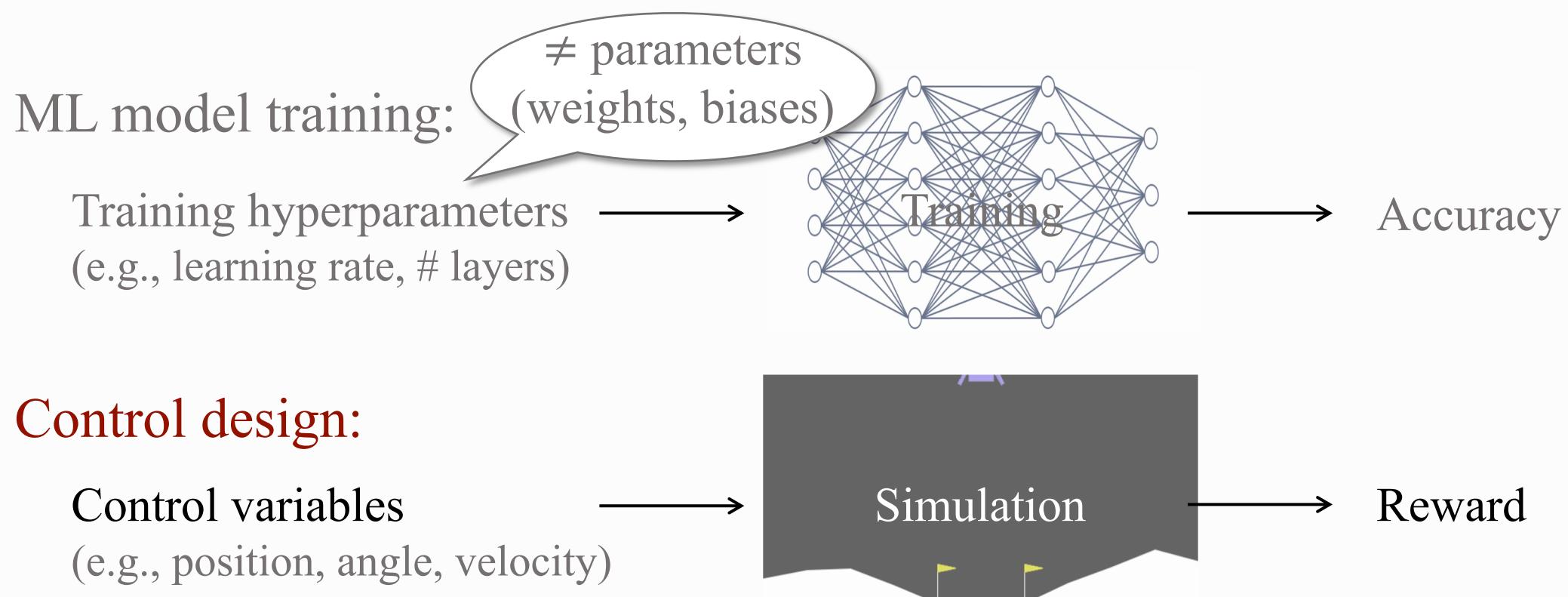
Training hyperparameters
(e.g., learning rate, # layers)

≠ parameters
(weights, biases)

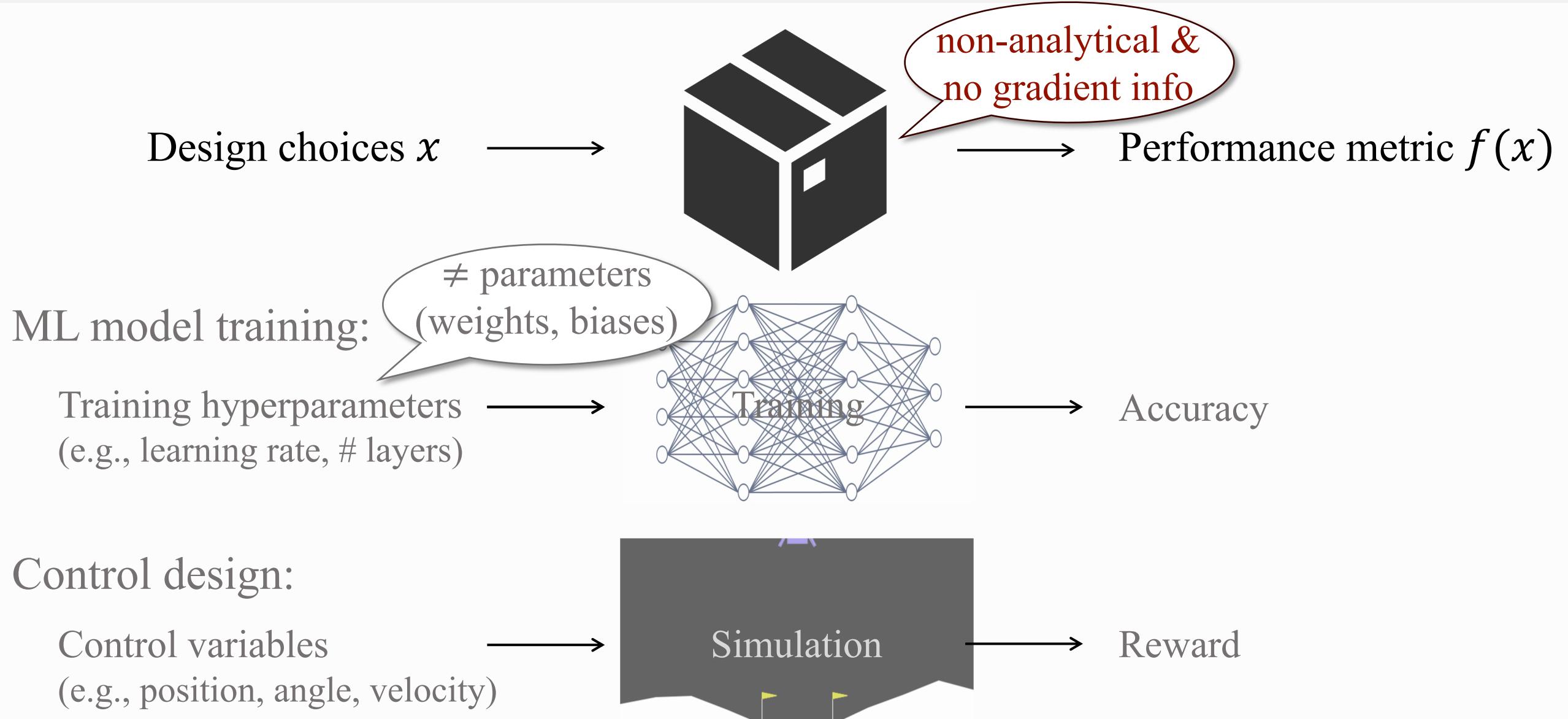


→ Accuracy

Motivation: World of Optimization under Uncertainty



Motivation: World of Optimization under Uncertainty



Motivation: World of Optimization under Uncertainty

Black-box optimization:

(gradient-based methods not applicable)

Input x \longrightarrow



non-analytical &
no gradient info

\longrightarrow Observed outcome $f(x)$

ML model training:

\neq parameters

(weights, biases)

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



\longrightarrow Accuracy

Background: Black-Box Optimization

Black-box optimization:

(gradient-based methods not applicable)

Input x \longrightarrow



expensive-to-evaluate

\longrightarrow Observed outcome $f(x)$

ML model training:

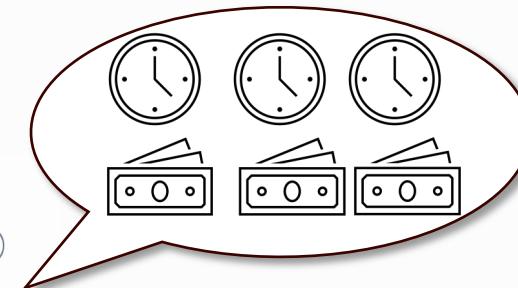
Training hyperparameters \longrightarrow
(e.g., learning rate, # layers)



Training time

Compute credits

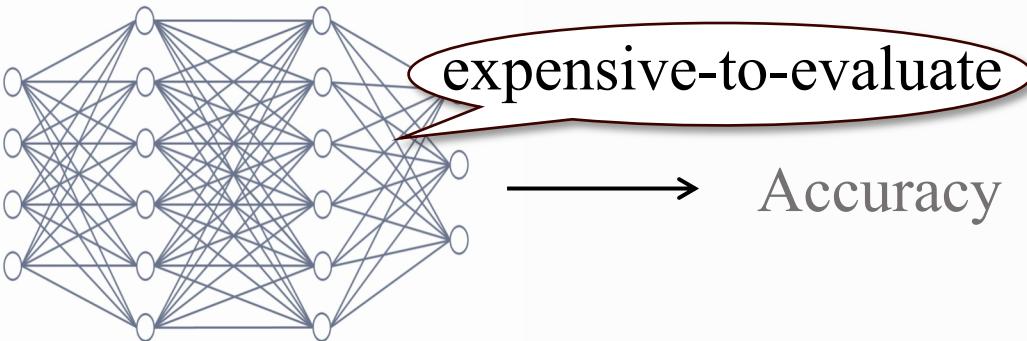
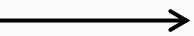
\longrightarrow Accuracy



Naïve (Non-Adaptive) Approach: Grid Search

ML model training:

Training hyperparameters



Accuracy

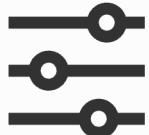
Training hyperparameter	Range	Number of Options
Batch size	[16, 512]	10
Learning rate	[1e-4, 1e-1]	10
Momentum	[0.1, 0.99]	10
Weight decay	[1e-5, 1e-1]	10
Number of layers	{1, 2, 3, 4}	4
Max units per layer	[64, 1024]	10
Dropout	[0.0, 1.0]	10

40,000,000
combinations!

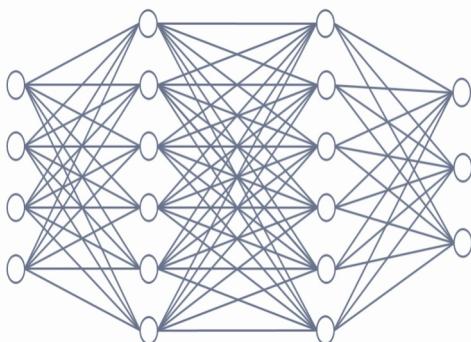
Data-Driven (Adaptive) Approach

Automated machine learning:

(AutoML)

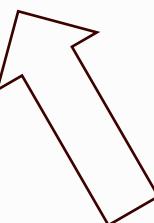


Hyperparameters

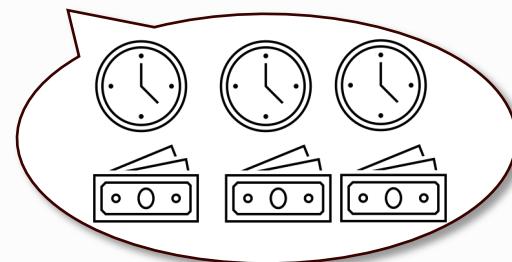


Performance metric
(e.g., accuracy)

Run next evaluation



Training Pipeline



Provide feedback



Data-efficient decision rule
(What to try next, when to stop)

Existing Umbrellas of Black-Box Optimization

Naïve (non-adaptive) approaches:

- Grid search
- Random search
- Manual tuning

Data-driven (adaptive) approaches:

- Local search (e.g., simulated annealing)
- Evolutionary algorithms (e.g., genetic algo)
- Bayesian optimization (e.g., EI, UCB, TS)
- Reinforcement learning (e.g., PPO, ENAS)
- LLM-based search agent (e.g., GENIUS)

Existing Umbrellas of Black-Box Optimization

Naïve (non-adaptive) approaches:

- Grid search
- Random search
- Manual tuning

Data-driven (adaptive) approaches:

- Local search (e.g., simulated annealing)
- Evolutionary algorithms (e.g., genetic algo)
- Bayesian optimization (e.g., EI, UCB, TS)
- Reinforcement learning (e.g., PPO, ENAS)
- LLM-based search agent (e.g., GENIUS)



Limitations in practice:

1. Limited principled guidance (e.g., naïve)
2. Often data-inefficient (e.g., naïve, local search, evolutionary algo)
3. No or ad-hoc incorporation of additional side info (most)
(e.g., varying eval costs)

Overview of Contributions Across My Work

Naïve (non-adaptive) approaches:

- Grid search
- Random search
- Manual tuning

Data-driven (adaptive) approaches:

- Local search
- Evolutionary algorithms

- Bayesian optimization  Part I
- Reinforcement learning  Part II
- LLM-based search agent 



Contributions of methods in my work:

1. Principled guidance
2. Competitive empirical performance
3. Principled incorporation of additional side info



New methods under this umbrella

Outline

Part I (Recent):

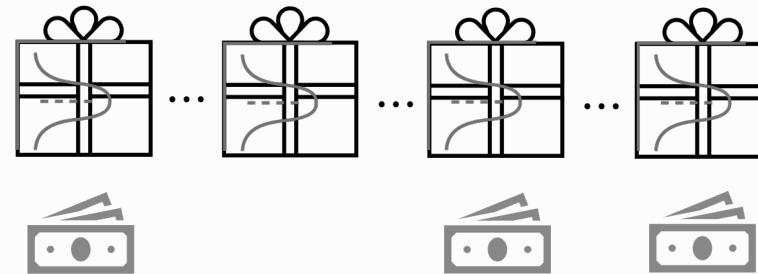
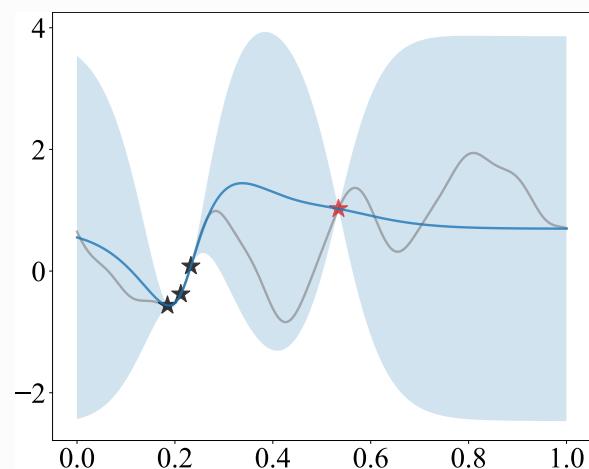
Bayesian Optimization via Gittins Index Design Principle

Part II (Ongoing):

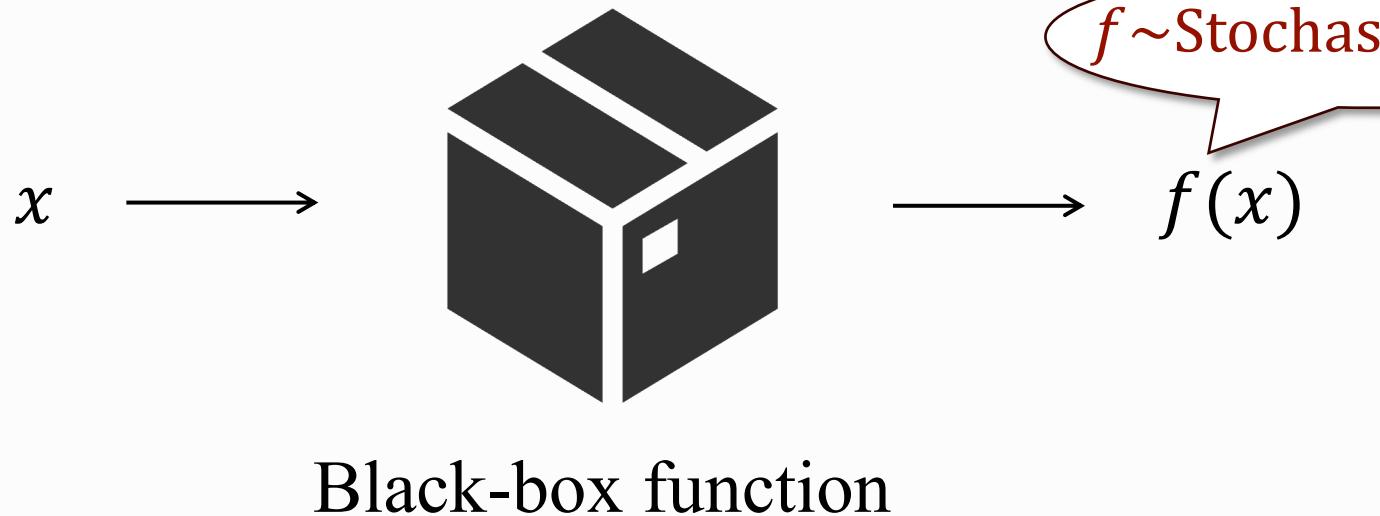
Black-box Optimization Beyond AutoML

Reinforcement Learning · *Engineering Design* · *Scientific Discovery* · *LLM*

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

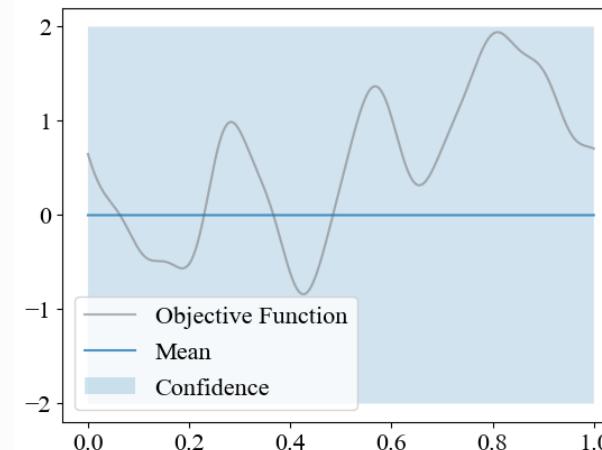
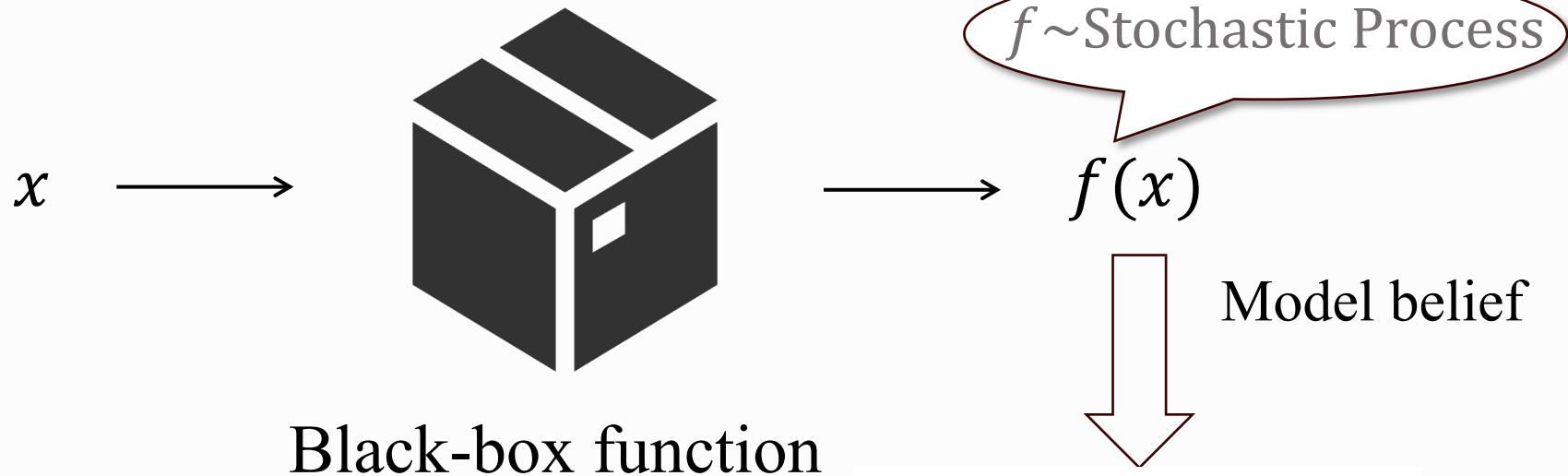


Bayesian Optimization



Bayesian Optimization

Time 0



Probabilistic model
(e.g., Gaussian process)

Bayesian Optimization

Time t

x_1, \dots, x_t

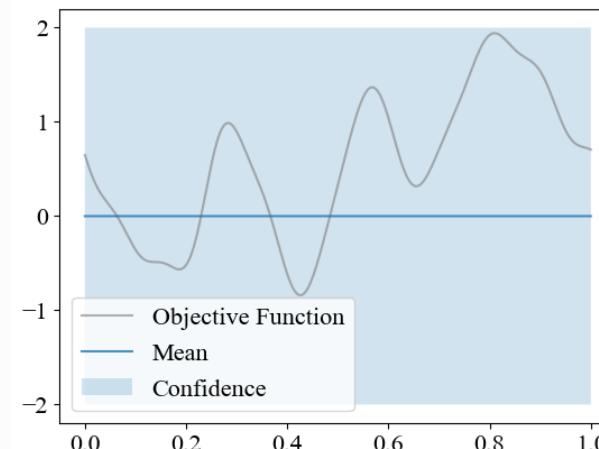


Black-box function

$f \sim \text{Stochastic Process}$

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

Model belief



Probabilistic model
(e.g., Gaussian process)

Bayesian Optimization

Time t

x_1, \dots, x_t \longrightarrow

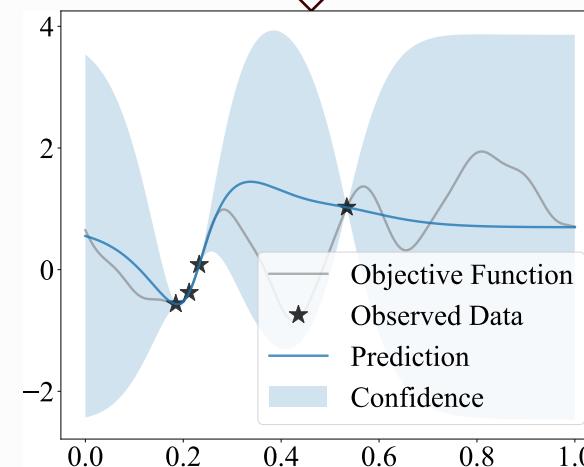


Black-box function

$f \sim \text{Stochastic Process}$

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

Update belief
(Bayes' rule)



Probabilistic model
(e.g., Gaussian process)

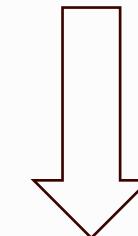
Bayesian Optimization

Time t

x_1, \dots, x_t

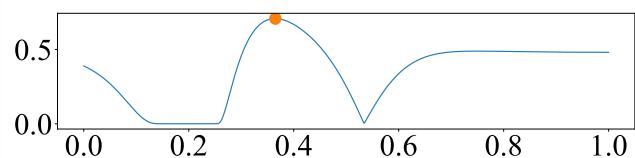


$f(x_1), \dots, f(x_t)$

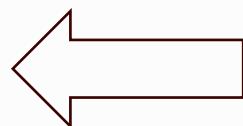


Update belief
(Bayes' rule)

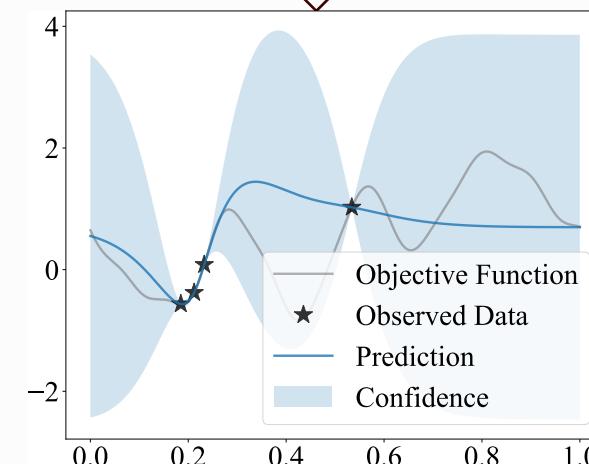
Black-box function



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

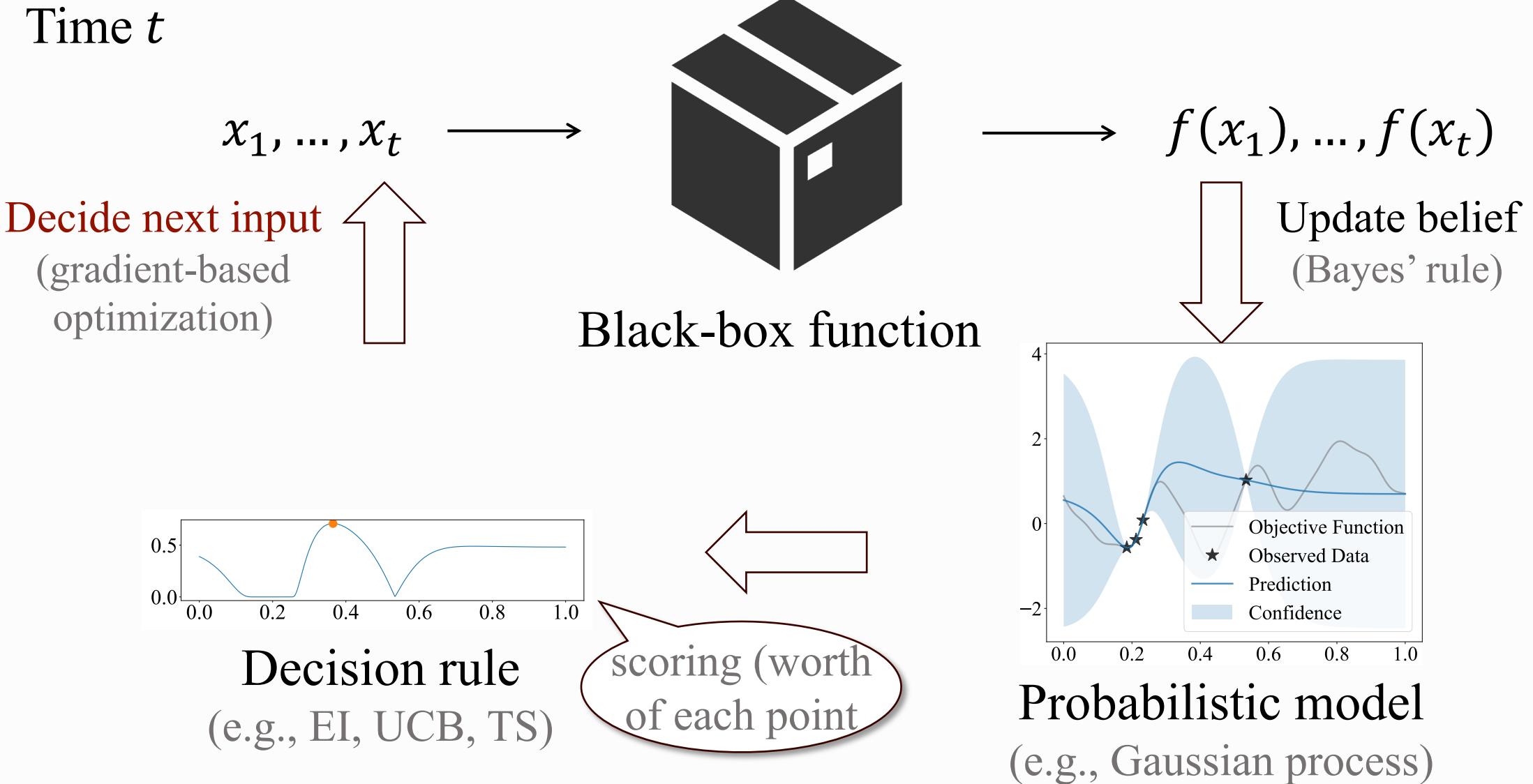


scoring (worth
of each point)



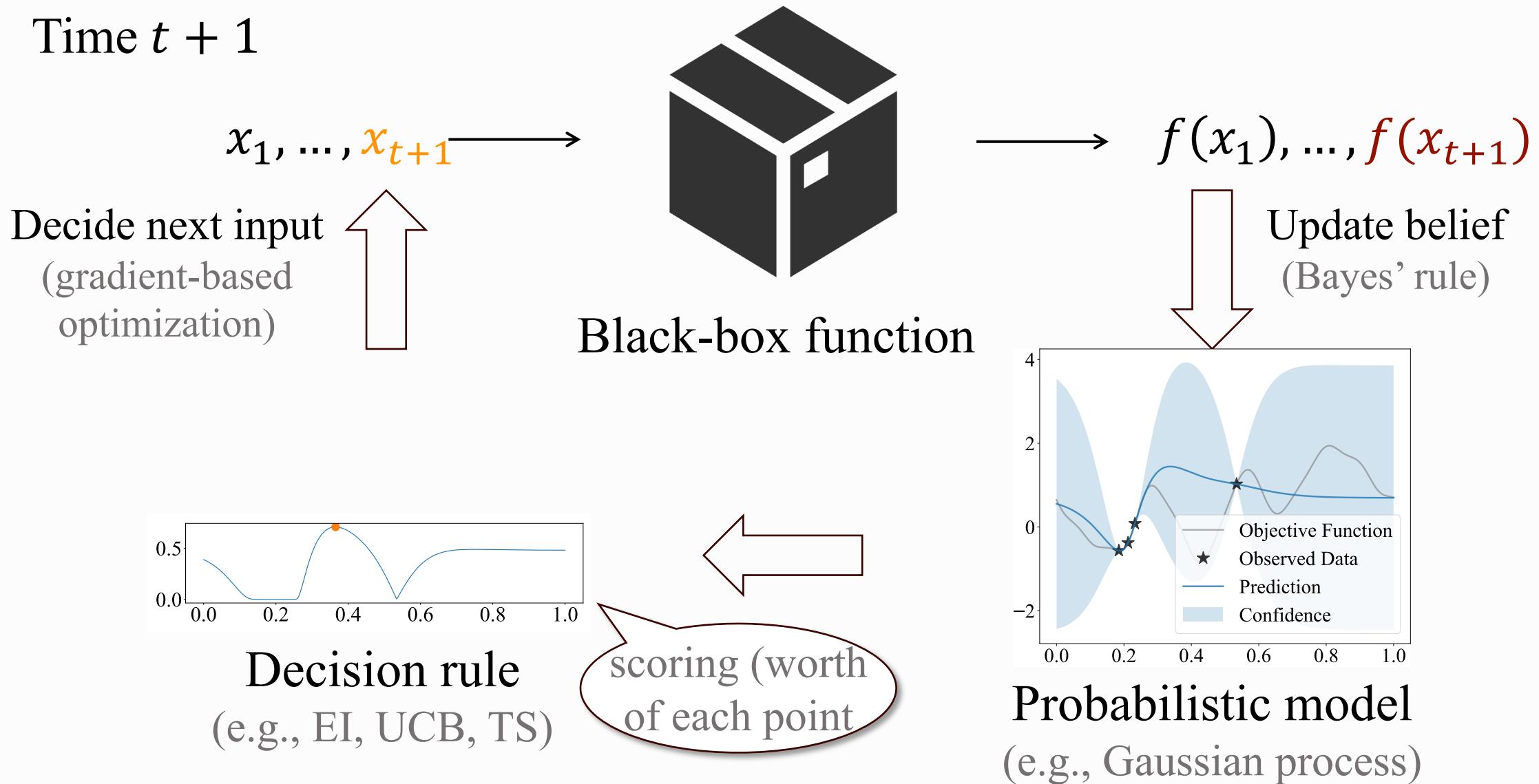
Probabilistic model
(e.g., Gaussian process)

Bayesian Optimization



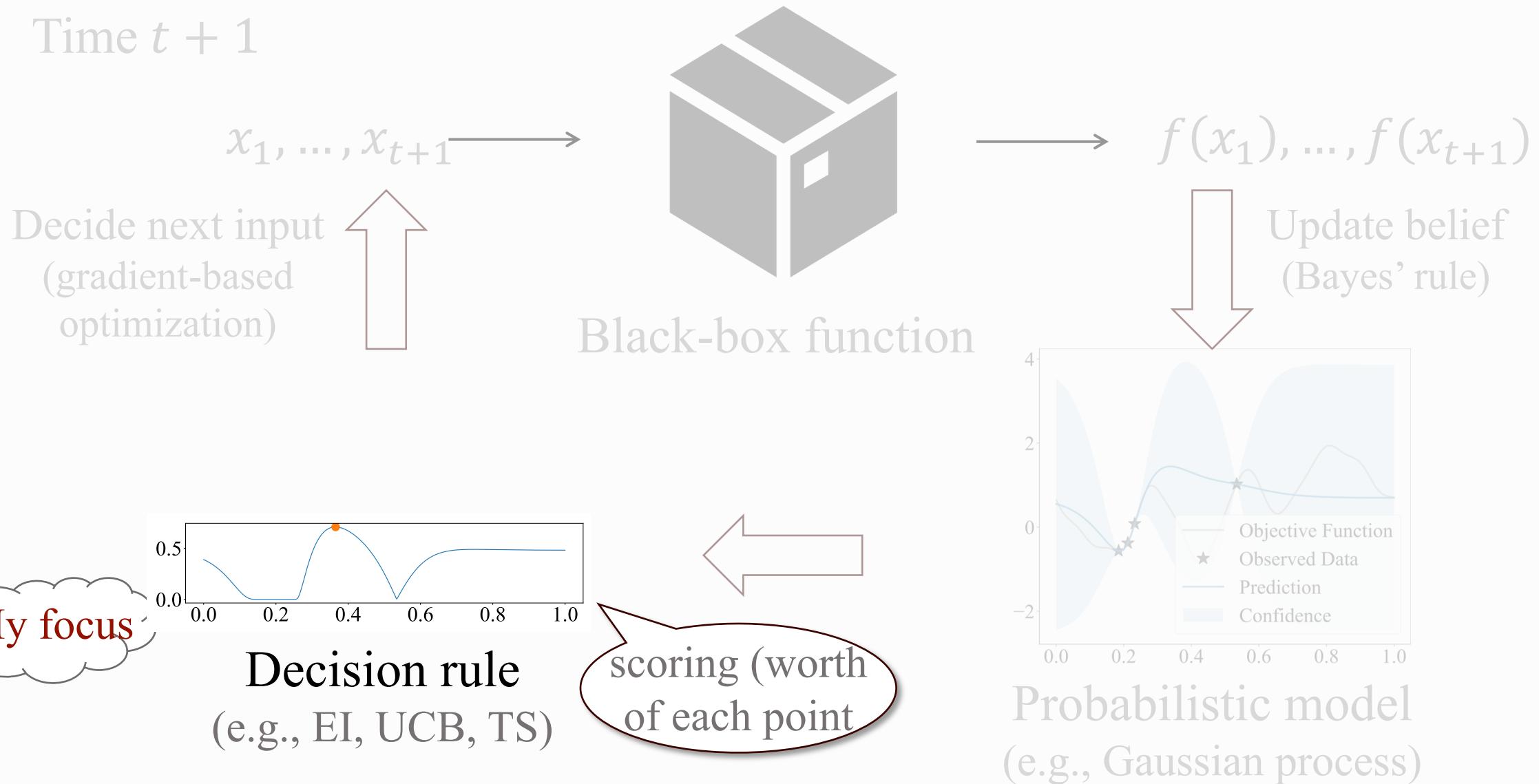
Bayesian Optimization

Time $t + 1$

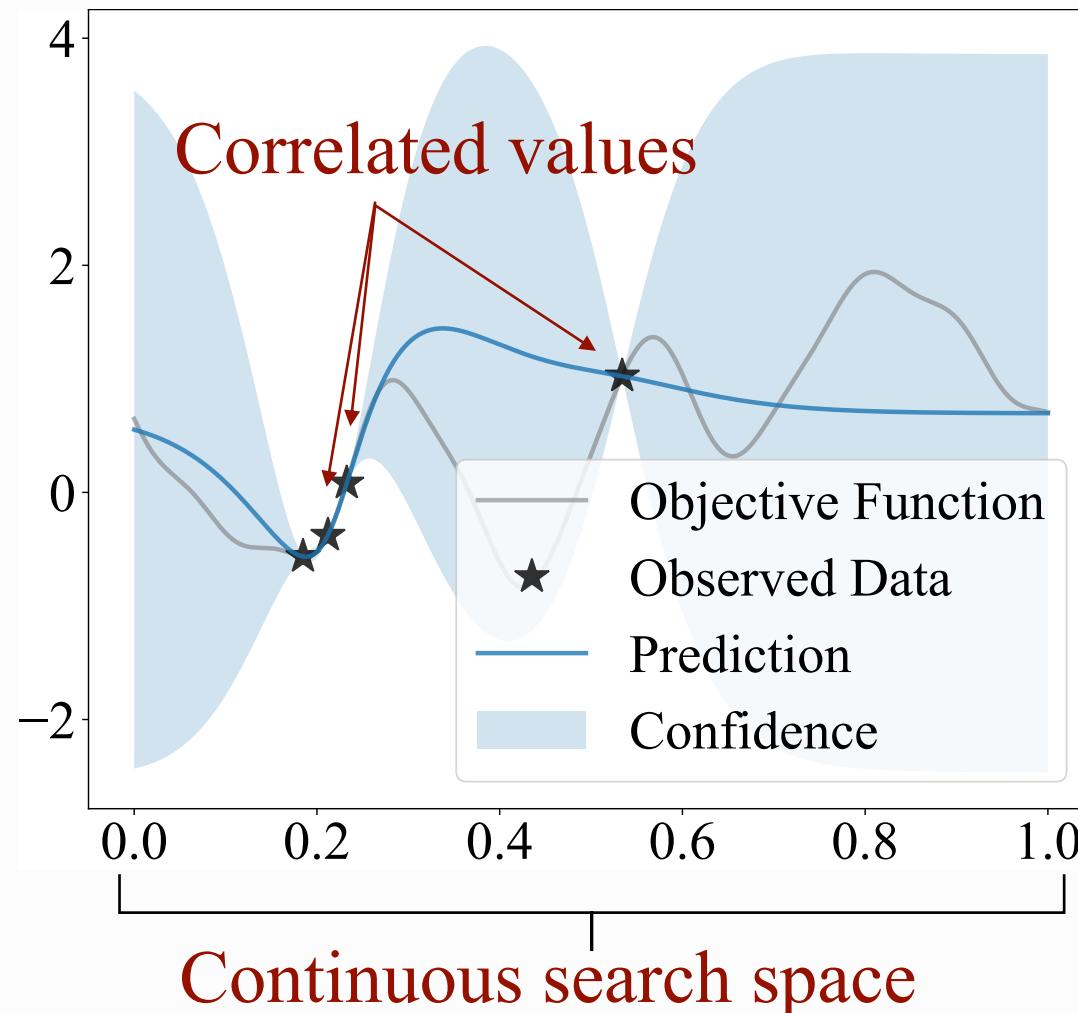


Bayesian Optimization

Time $t + 1$



Challenges in Decision Rule Design



Correlation & continuity \Rightarrow Intractable MDP \Rightarrow Optimal policy unknown

Popular Decision Rule: Expected Improvement

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

“improvement”

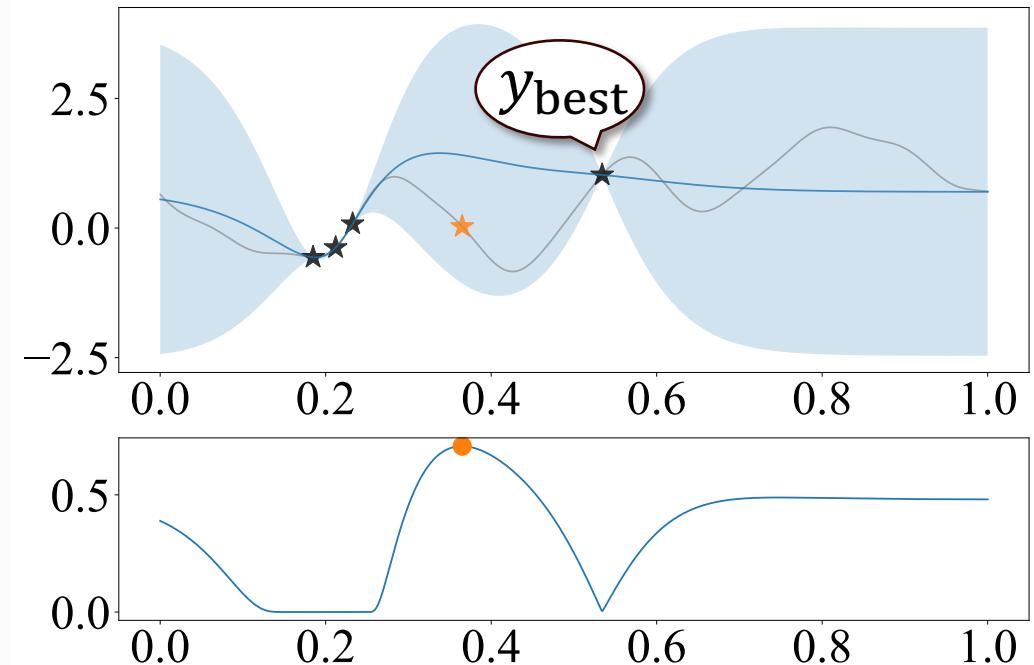
current best observed

data D

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

posterior distribution

One-step approximation to MDP



Expected improvement $EI(x)$

Popular Decision Rule: Expected Improvement

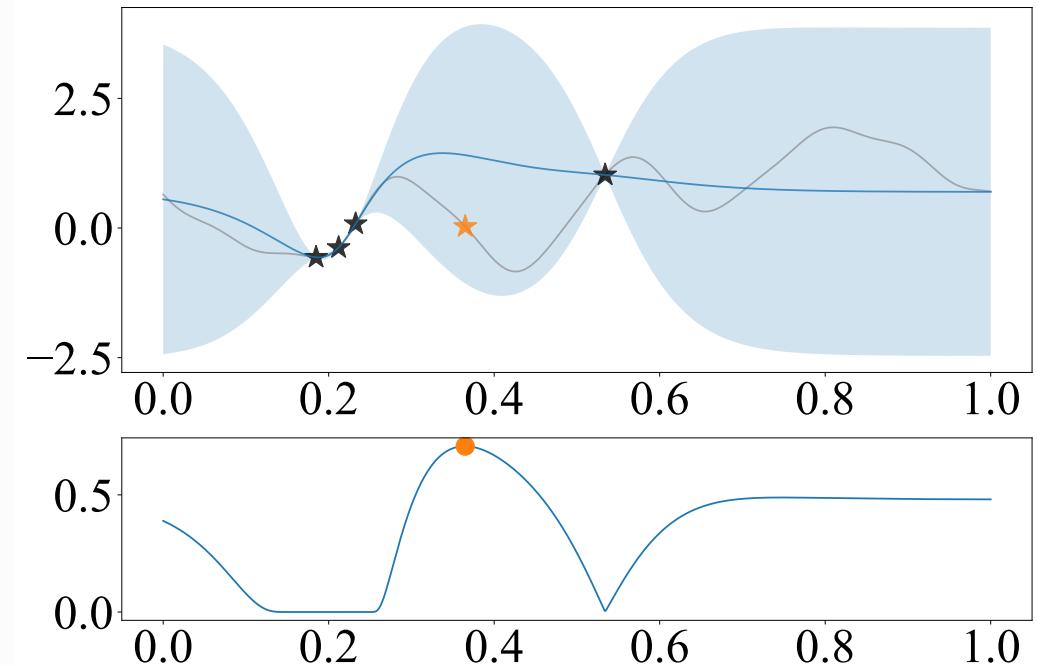
$$EI(x) = \mathbb{E}[\max(f(x) - y_{\text{best}}, 0) \mid x_1, \dots, x_t]$$

current best observed
data D
“improvement”

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

posterior distribution

One-step approximation to MDP

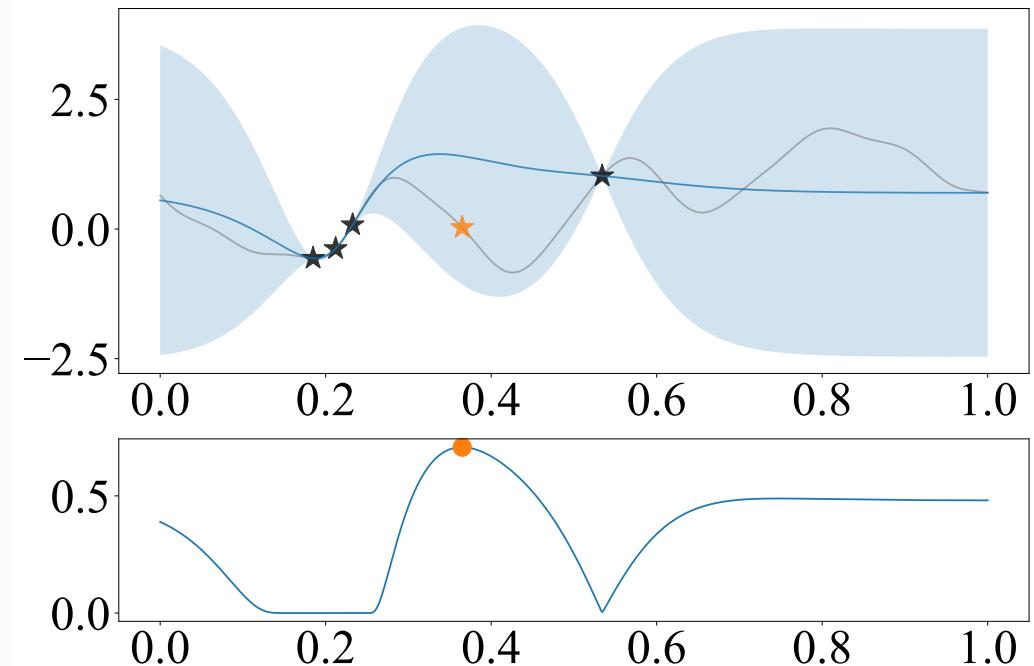


Expected improvement $EI(x)$

Improvement-based
design principle

Existing Design Principles

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

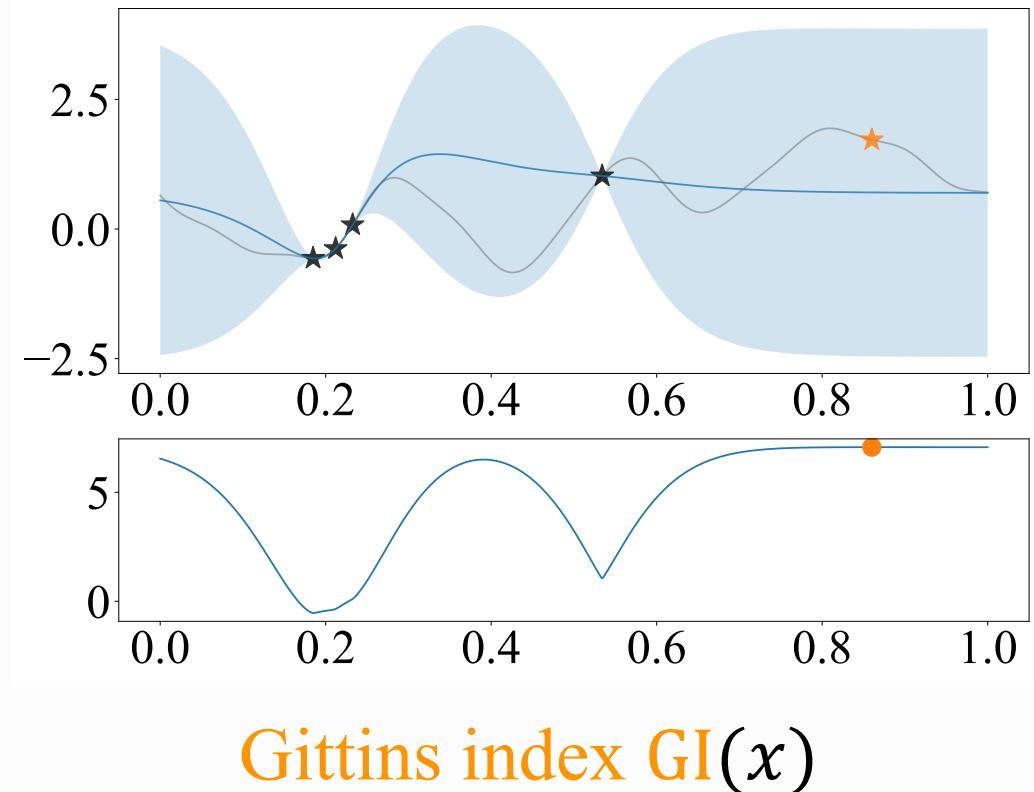


Expected improvement $EI(x)$

Improvement-based
design principle

New Design Principle: Gittins Index

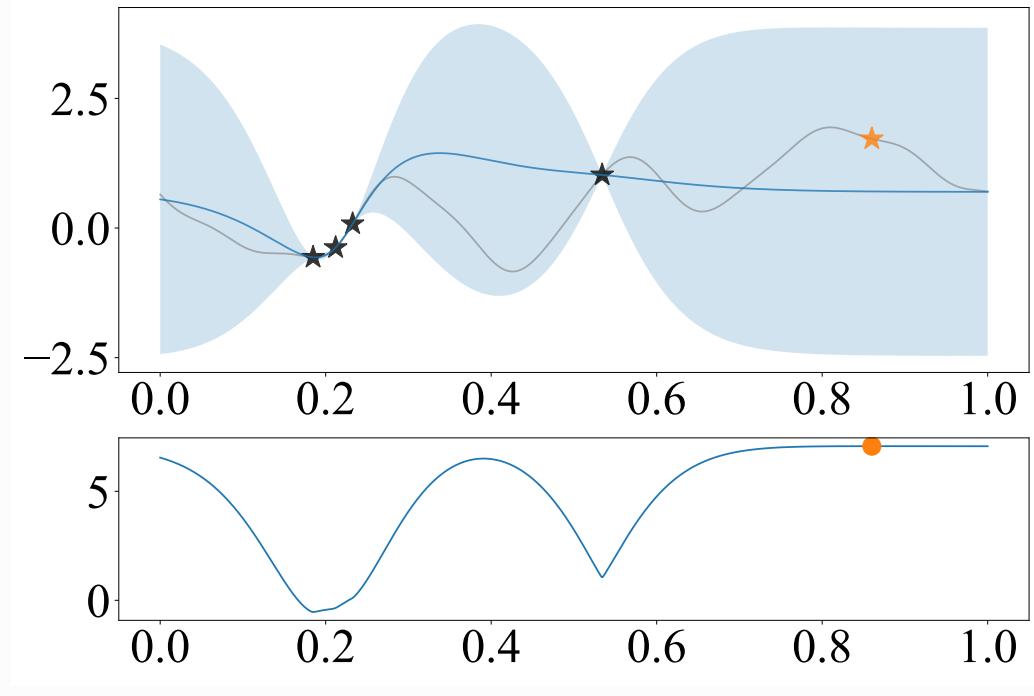
- Improvement-based (e.g., EI)
- Entropy-based
- Confidence bounds (UCB/LCB)
- Thompson sampling (TS)
- **Gittins Index**



Gittins index $GI(x)$

New Design Principle: Gittins Index

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



Gittins index $GI(x)$

? Why another principle?

Our Contribution: Gittins Index Principle

Joint work with Ziv Scully and Alexander Terenin et al.

1. Principled easy-to-compute policy class
2. Natural incorporation of side info and flexibility
3. Competitive performance on benchmarks
4. Theoretical guarantees

Our Contribution: Gittins Index Principle

Joint work with Ziv Scully and Alexander Terenin et al.

1. Principled easy-to-compute policy class

EI
✓

2. Natural incorporation of side info and flexibility

✗

3. Competitive performance on benchmarks

✗

4. Theoretical guarantees

✗

Our Contribution: Gittins Index Principle

Joint work with Ziv Scully and Alexander Terenin et al.

1. Principled easy-to-compute policy class

EI
✓

2. Natural incorporation of side info and flexibility

✗

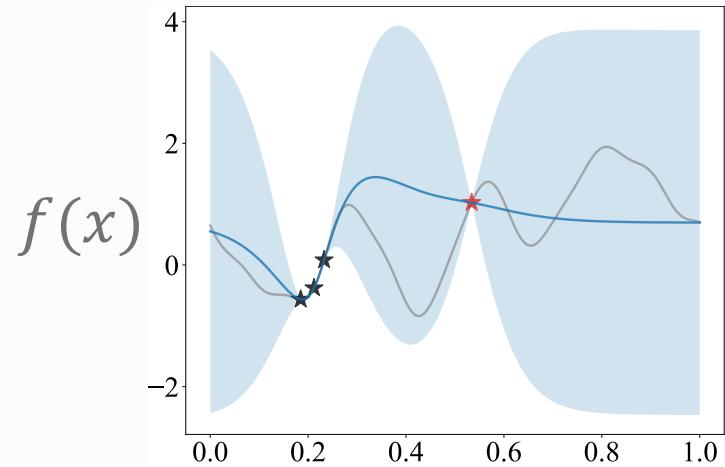
3. Competitive performance on benchmarks

✗

4. Theoretical guarantees

✗

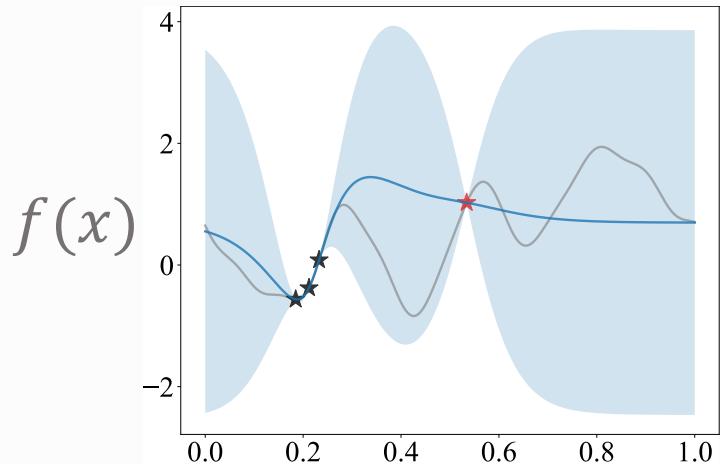
Bayesian Optimization



Continuous search space

Correlated function values

Bayesian Optimization



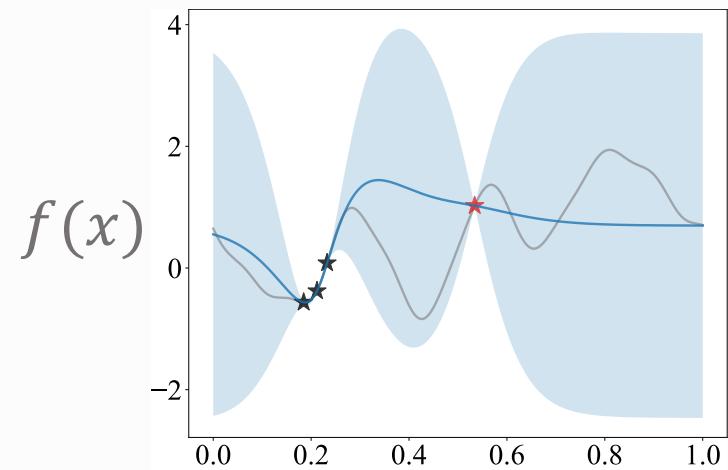
Continuous search space \Rightarrow

Discrete

Correlated function values \Rightarrow

Independent

Bayesian Optimization



Continuous search space

Correlated function values

Pandora's Box

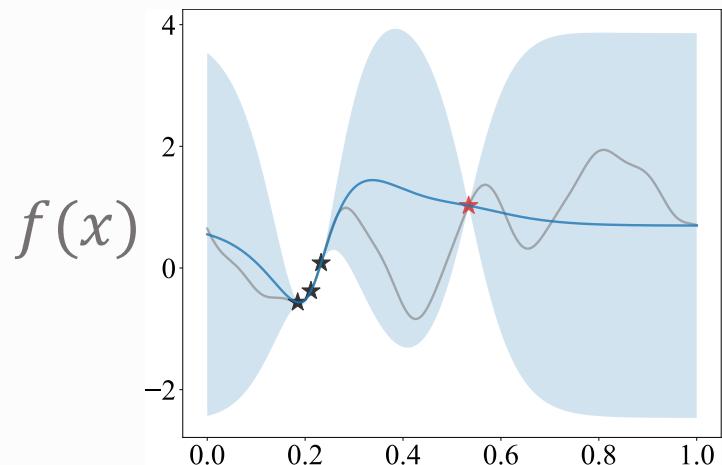
[Weitzman'79]



Discrete

Independent

Bayesian Optimization

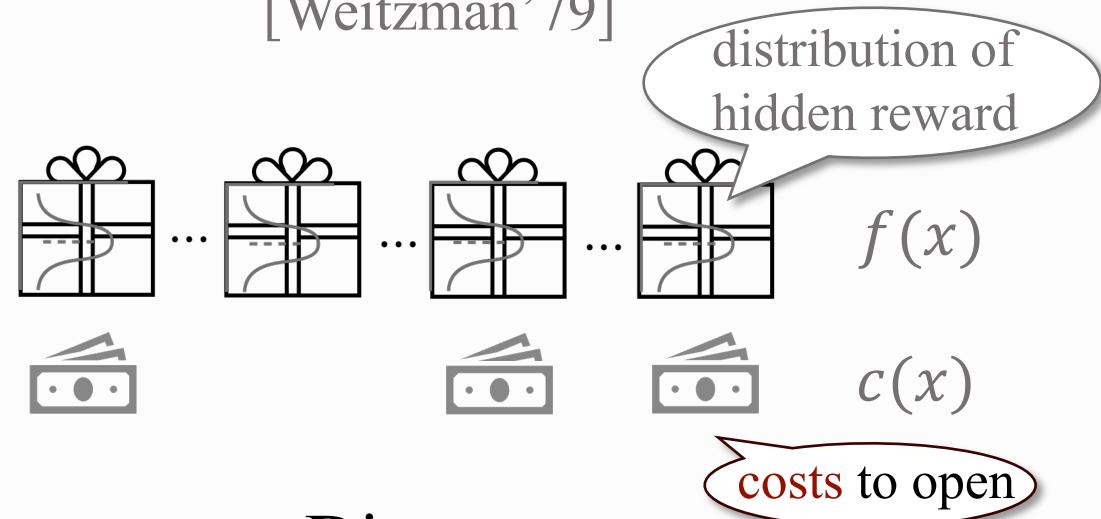


Continuous search space

Correlated function values

Pandora's Box

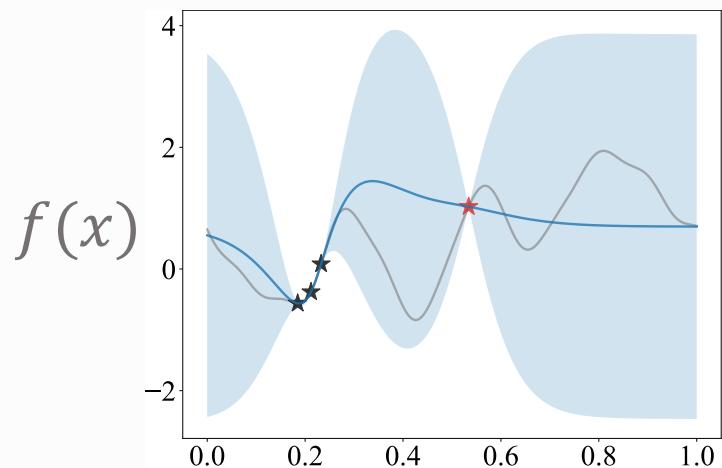
[Weitzman'79]



Discrete

Independent

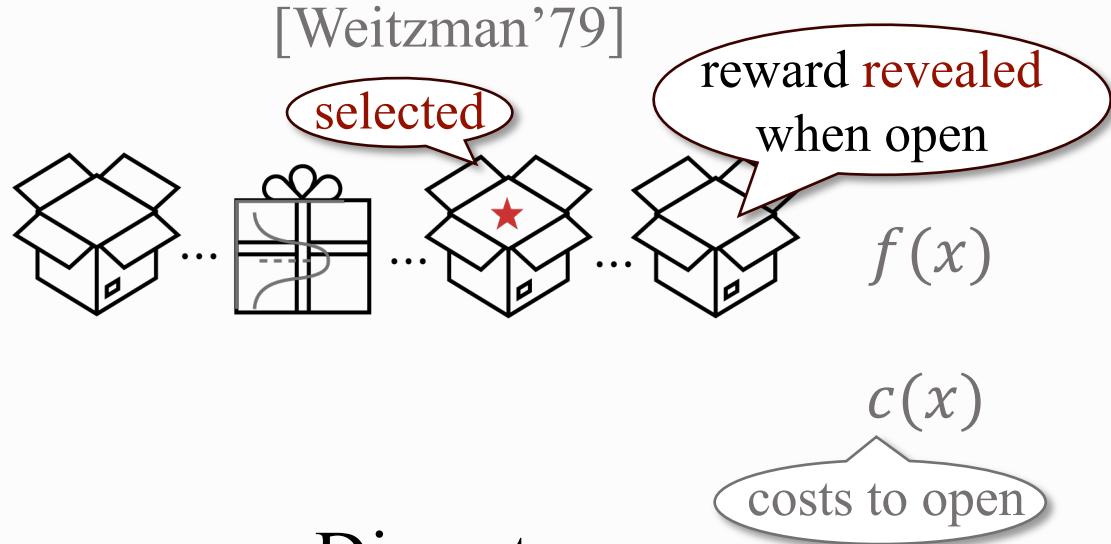
Bayesian Optimization



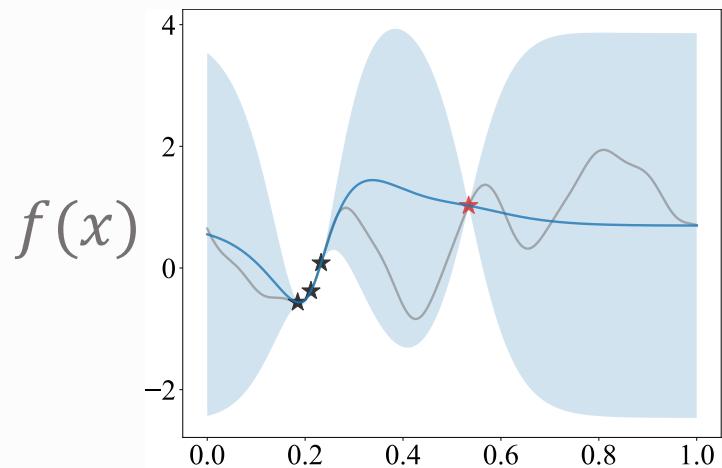
Continuous search space

Correlated function values

Pandora's Box



Bayesian Optimization

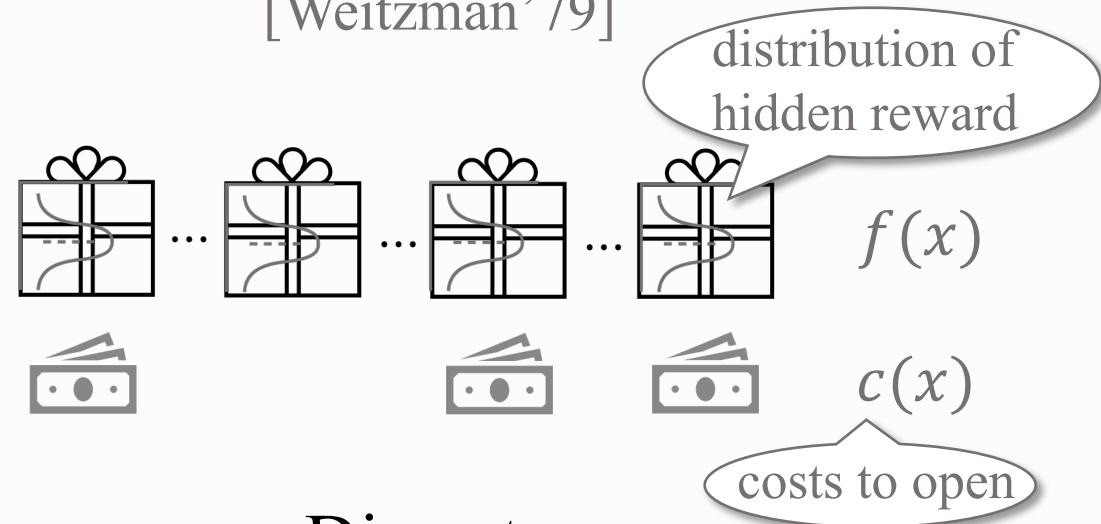


Continuous search space

Correlated function values

Pandora's Box

[Weitzman'79]

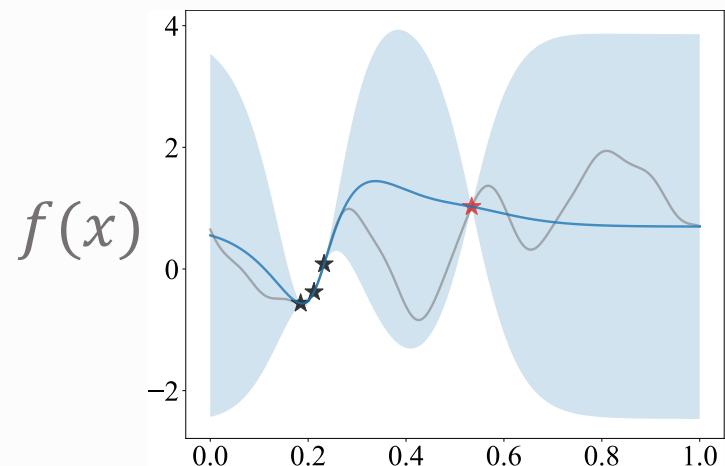


Discrete

Independent

Optimal policy: Gittins index

Bayesian Optimization

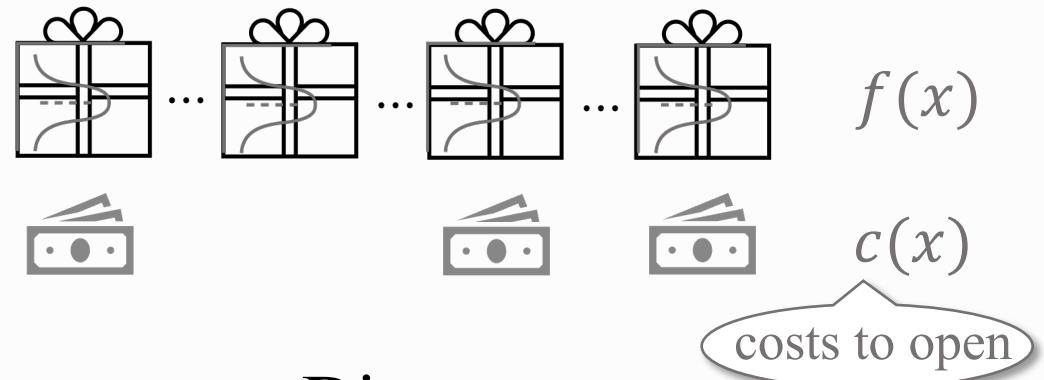


Continuous search space

Correlated function values

Pandora's Box

[Weitzman'79]



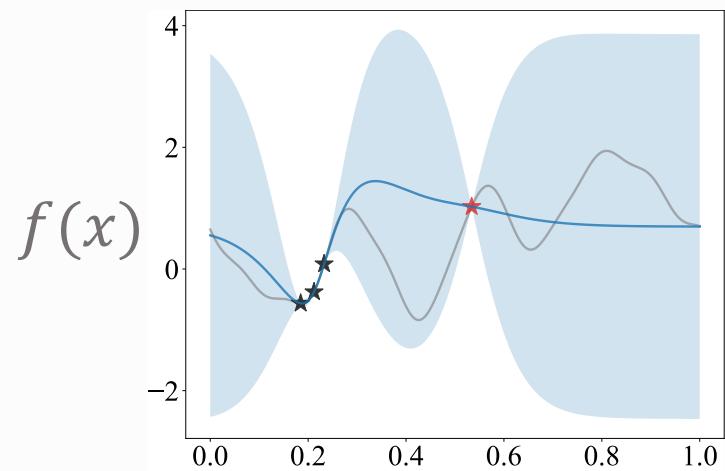
Discrete

Independent

How to translate?

Optimal policy: Gittins index

Bayesian Optimization

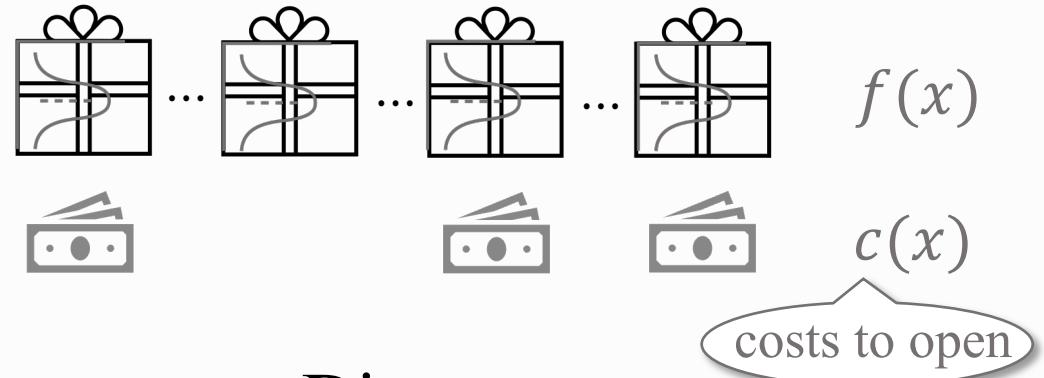


Continuous search space

Correlated function values

Pandora's Box

[Weitzman'79]

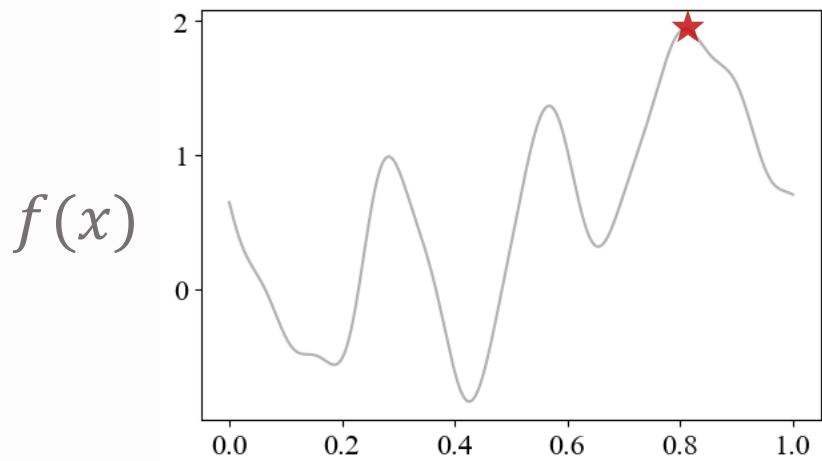


Discrete

Independent

Our policy: $\text{GI}_{f|D}(x; c)$ \leftarrow Optimal policy: $\text{GI}_f(x; c)$
incorporate posterior
take continuum limit

Bayesian Optimization



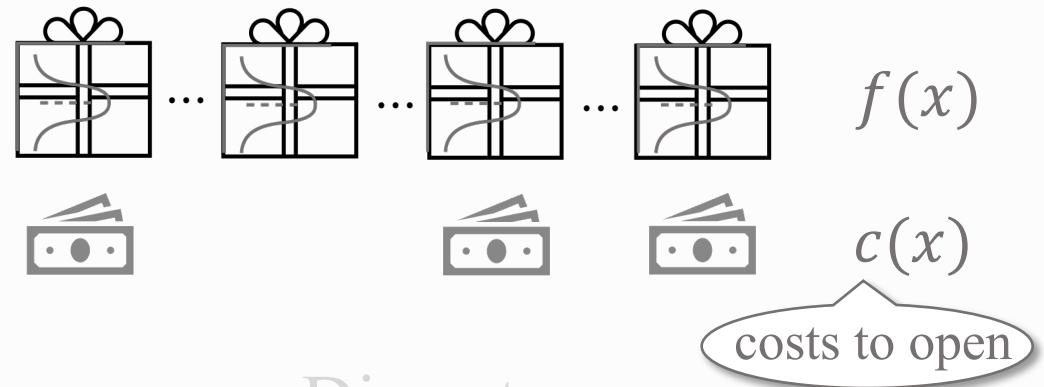
Continuous

Correlated

Our policy: $\text{GI}_{f|D}(x; c(x))$
How to compute?

Pandora's Box

[Weitzman'79]



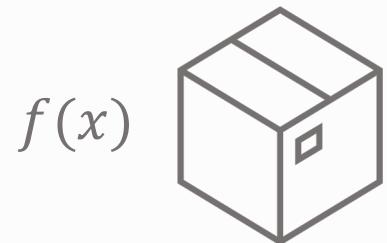
Discrete

Independent

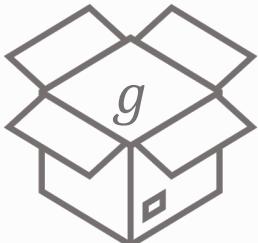
incorporate posterior
take continuum limit
Optimal policy: $\text{GI}_f(x; c(x))$

Intuition

Exploration



Exploitation



vs.

Open closed box

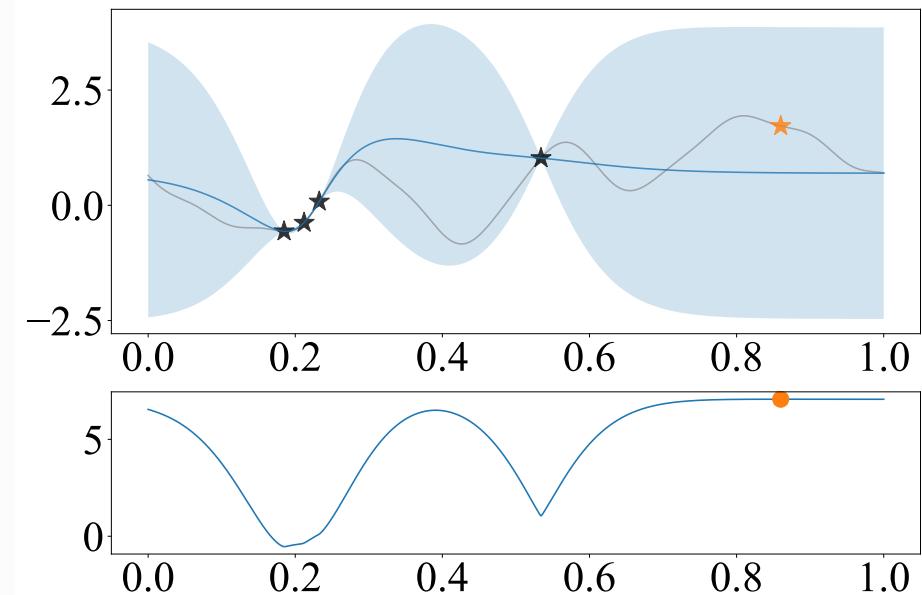
Take opened box

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

$$g$$

Should one open box? Depend on g !

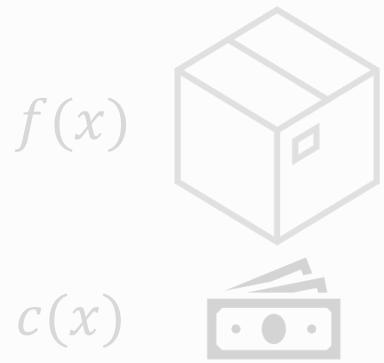
Gittins Index



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

Intuition

Exploration



vs.

Exploitation



Open closed box

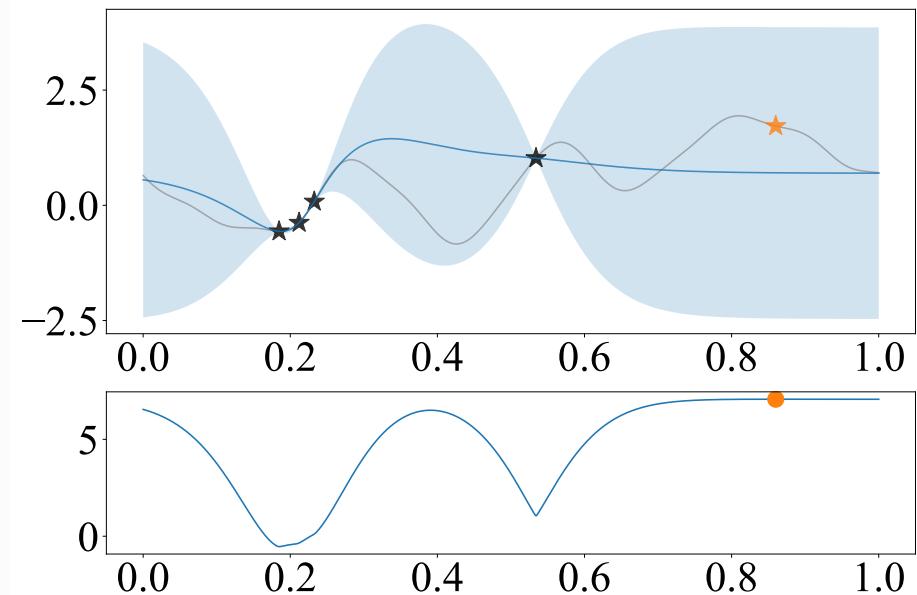
Take opened box

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

$$g$$

Should one open box? Depend on g !

Gittins Index

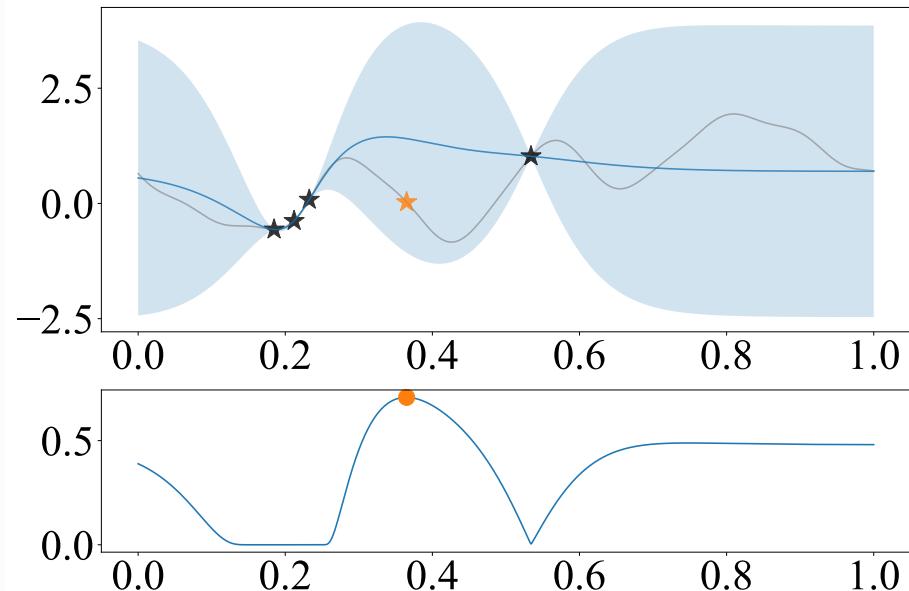


$$\begin{aligned}
 \text{GI}_{f|D}(x; c) &:= \text{solution } g \text{ s.t.} \\
 \mathbb{E}[\max(f(x), g) | D] - c(x) &= g \\
 \Leftrightarrow \mathbb{E}[\max(f(x) - g, g - g) | D] - c(x) &= 0 \\
 \Leftrightarrow \mathbb{E}[\max(f(x) - g, 0) | D] &= c(x)
 \end{aligned}$$

\sqcup

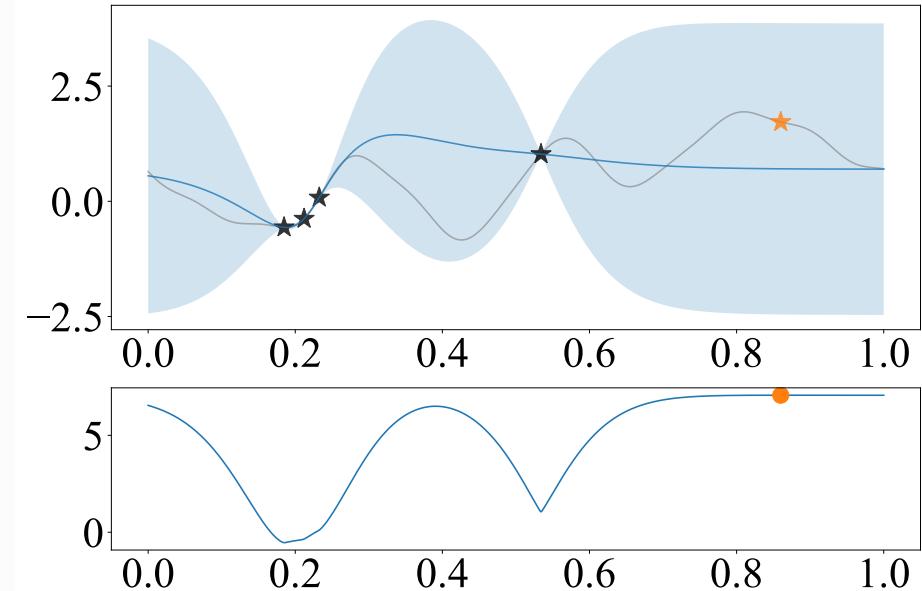
$\text{EI}_{f|D}(x; g)$

Expected Improvement

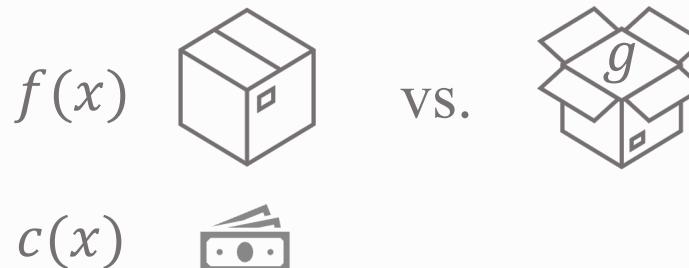


$$\text{EI}_{f|D}(x) := \mathbb{E}[\max(f(x) - y_{\text{best}}, 0) \mid D]$$

Gittins Index

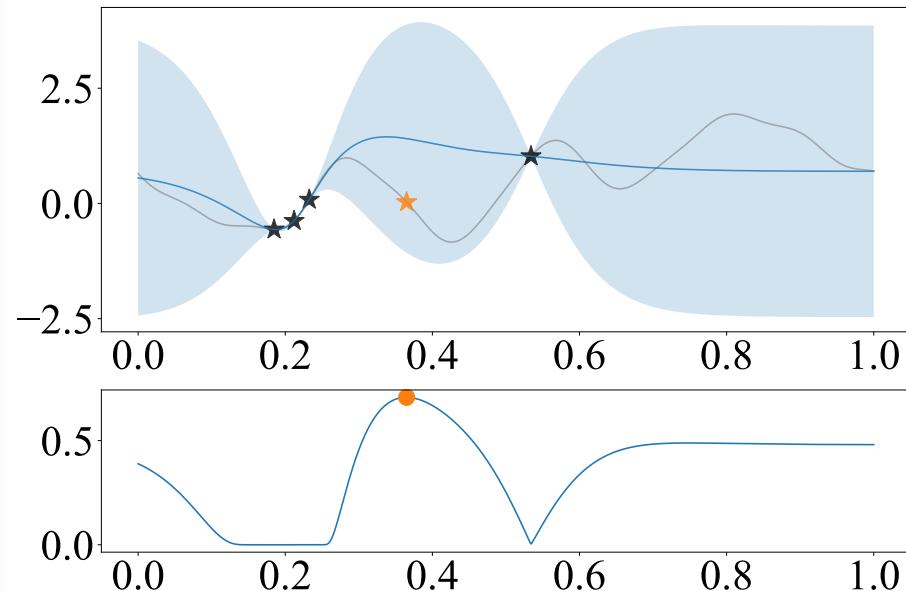


$$\begin{aligned} \text{GI}_{f|D}(x; c) &:= \text{solution } g \text{ s.t. } \text{EI}_{f|D}(x; g) = c(x) \\ \text{where } \text{EI}_{f|D}(x; g) &:= \mathbb{E}[\max(f(x) - g, 0) \mid D] \end{aligned}$$



Exploration Exploitation

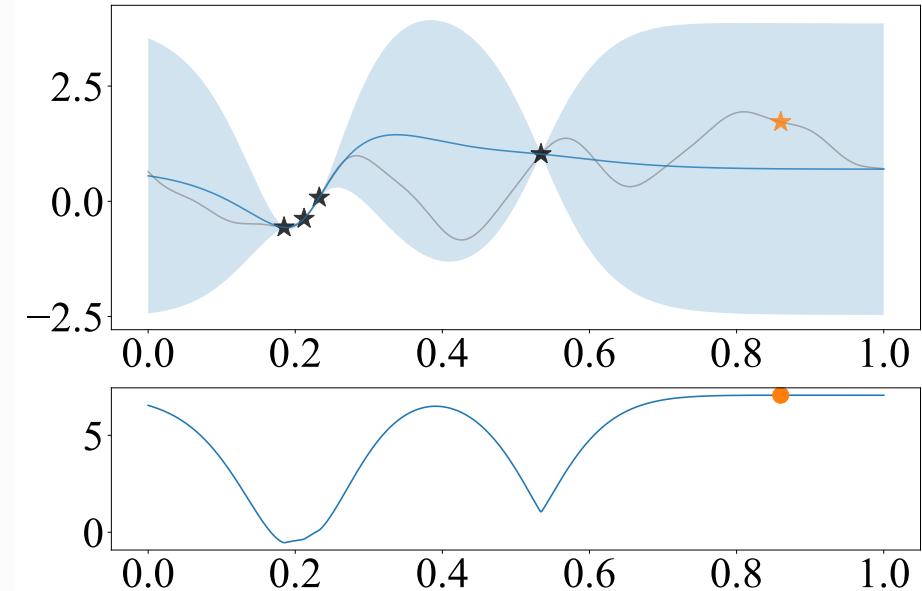
Expected Improvement



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

Temporal simplification to MDP
(One-step)

Gittins Index

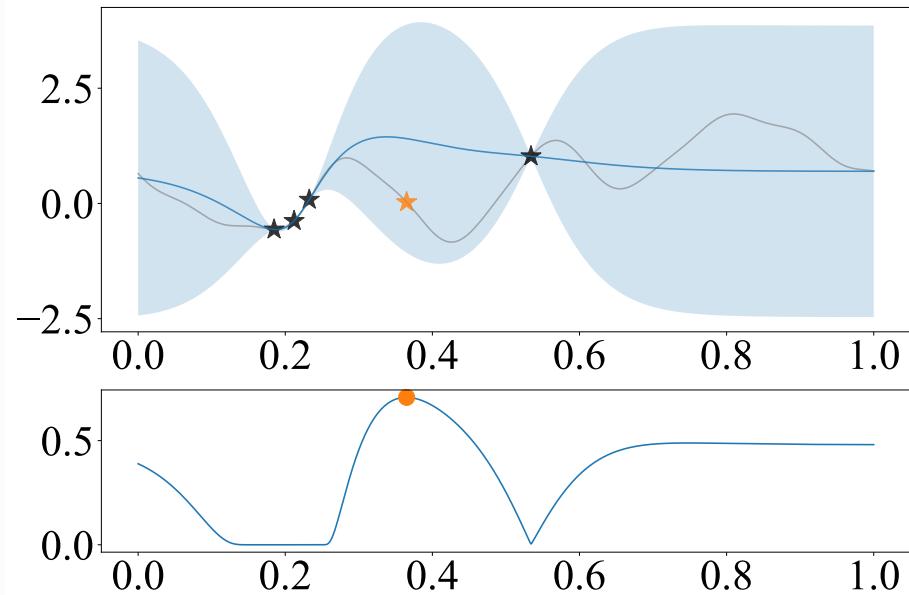


$$\text{GI}_{f|D}(x; c) := \text{solution } g \text{ s.t. } \text{EI}_{f|D}(x; g) = c(x)$$

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

Spatial simplification to MDP

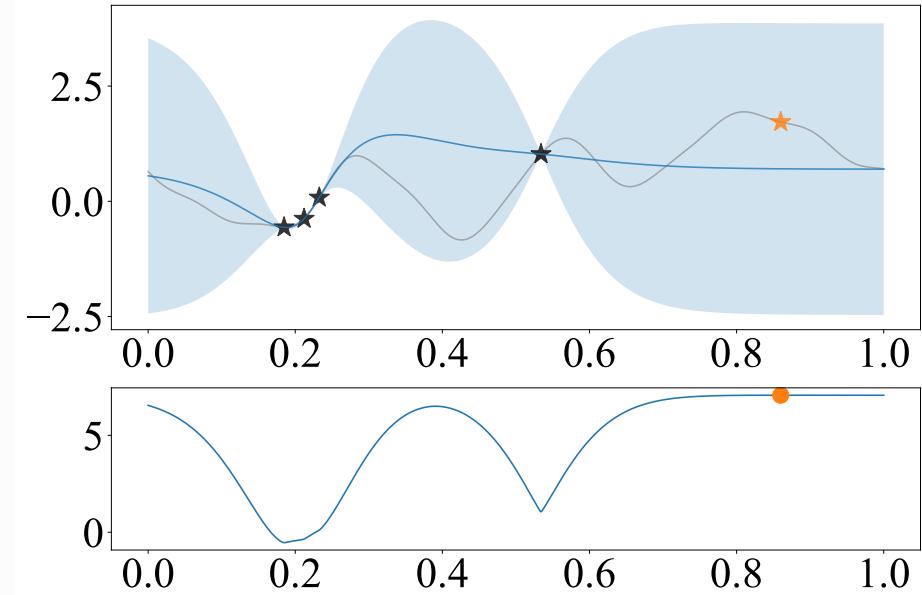
Expected Improvement



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

Temporal simplification to MDP
(One-step)

Gittins Index

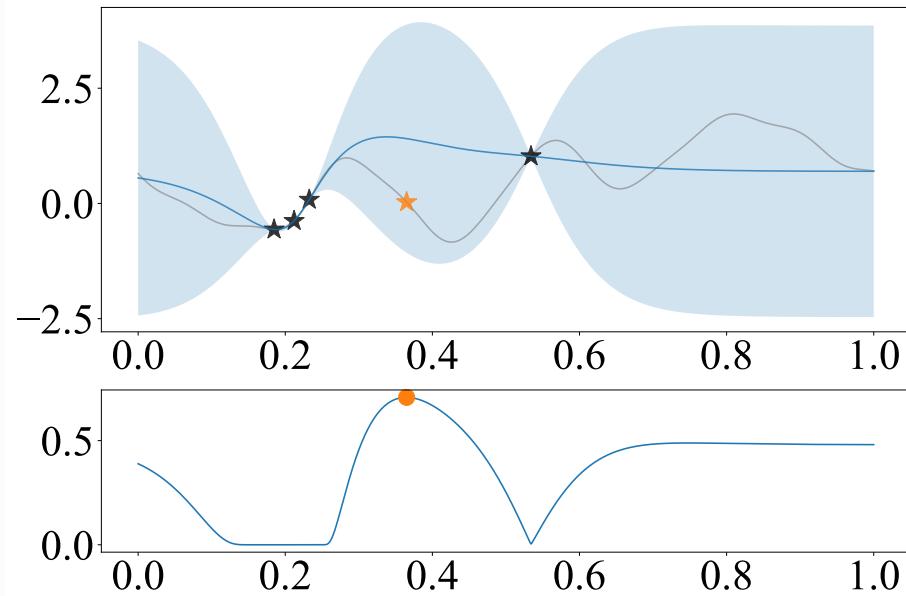


$$GI_{f|D}(x; c) := \text{solution } g \text{ s.t. } EI_{f|D}(x; g) = c(x)$$

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

Spatial simplification to MDP

Expected Improvement

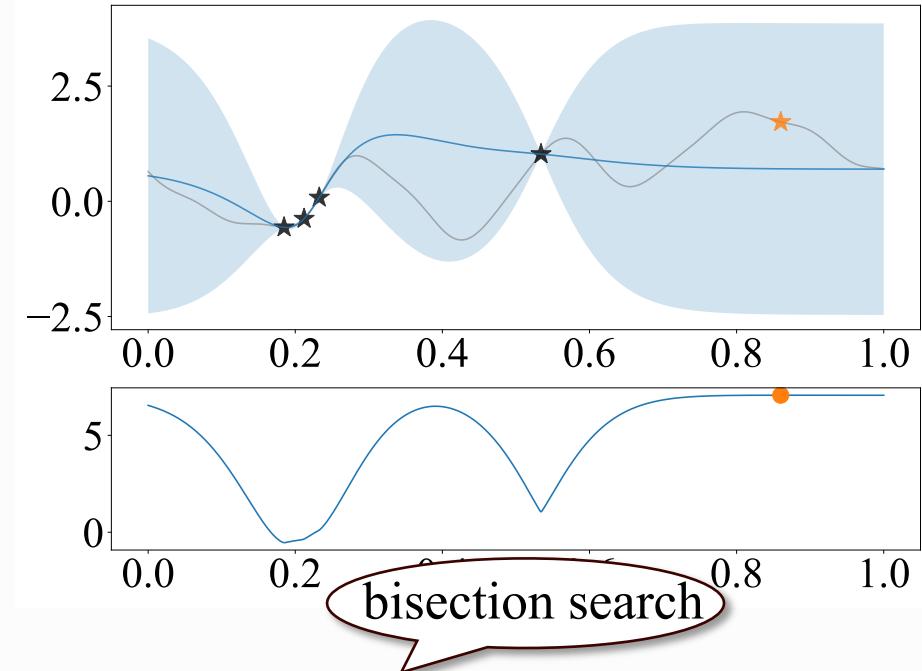


$$EI_{f|D}(x) := \mathbb{E}[\max(f(x) - y_{\text{best}}, 0) \mid D]$$

Temporal simplification to MDP

Both are **principled** and **easy-to-compute!**

Gittins Index



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

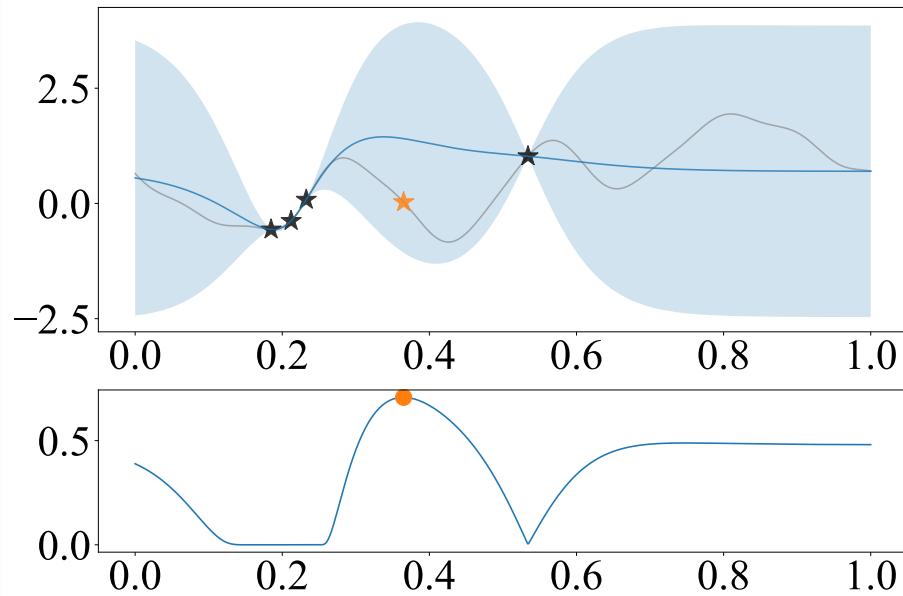
where $EI_{f|D}(x; g) := \mathbb{E}[\max(f(x) - g, 0) \mid D]$

Spatial simplification to MDP



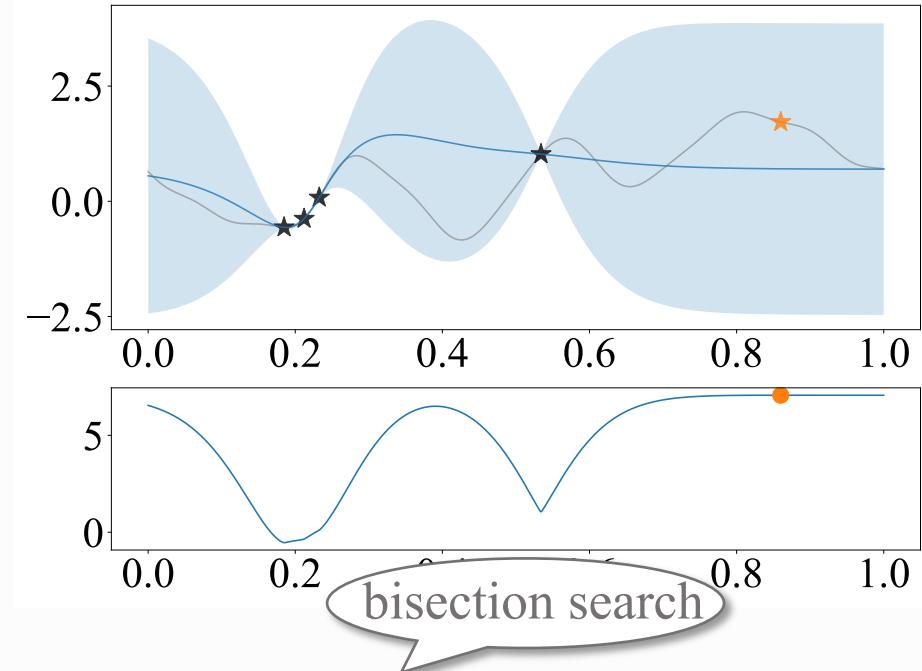
"Cost-aware Bayesian Optimization via the Pandora's Box Gittins Index." NeurIPS'24.

Expected Improvement



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

Gittins Index



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

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

Google DeepMind

FunBO: Discovering new acquisition functions for
Bayesian Optimization with FunSearch

hard to discover GI

Our Contribution: Gittins Index Principle

Joint work with Ziv Scully and Alexander Terenin et al.

1. Principled easy-to-compute policy class

EI
✓

2. Natural incorporation of side info and flexibility

✗

3. Competitive performance on benchmarks

✗

4. Theoretical guarantees

✗

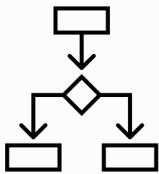
Under-explored Side Info and Flexibility



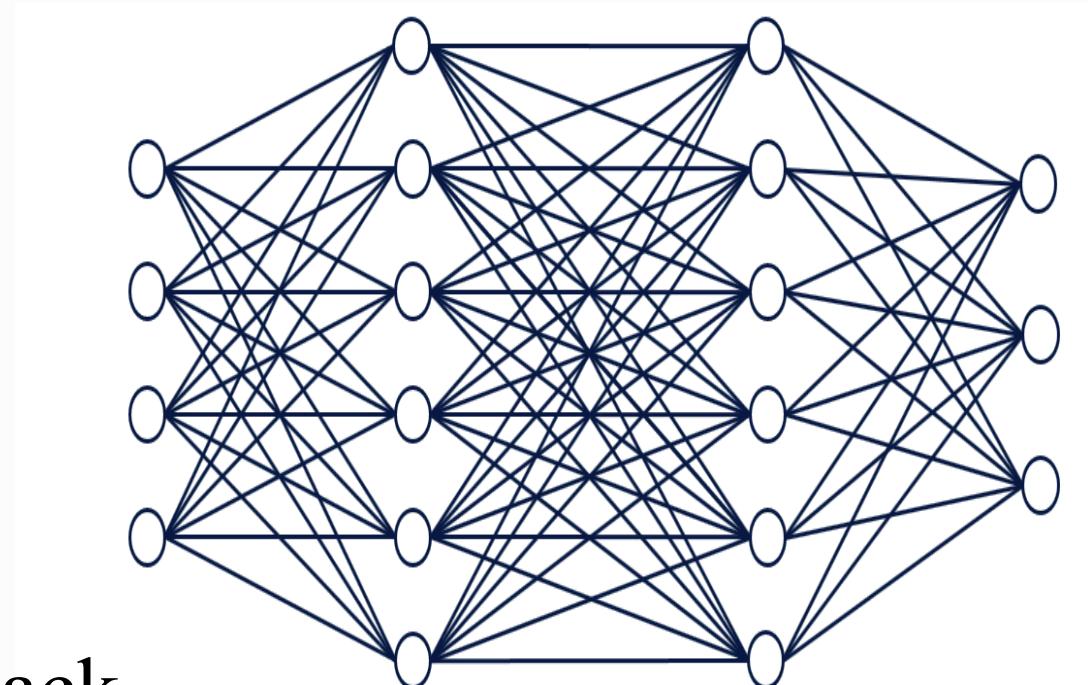
Varying evaluation costs



Smart stopping time



Observable multi-stage feedback



How does existing principle incorporate them?



Varying evaluation costs

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

[Snoek et al.'12]

Arbitrarily bad

[Astudillo et al.'21]



Smart stopping time

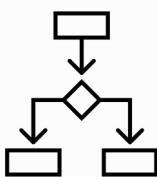
$$\tau: \text{EI}(x_\tau) \leq \theta$$

[Locatelli'97,

Nguyen et al.'17,

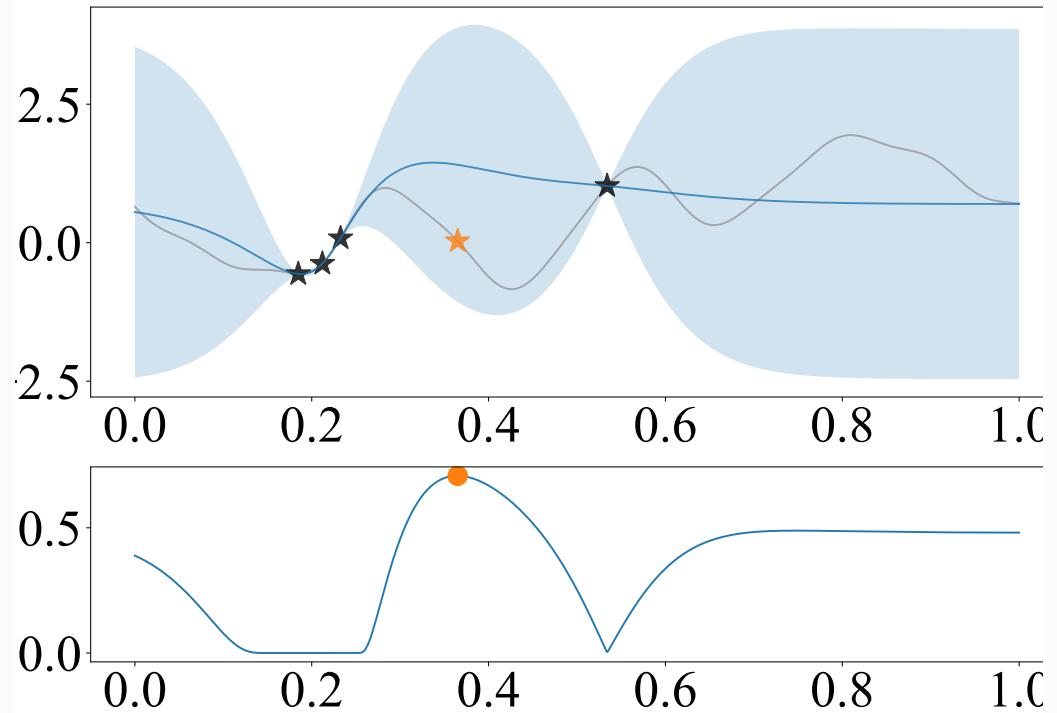
Ishibashi et al.'23]

Which threshold?



Observable multi-stage feedback

?



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

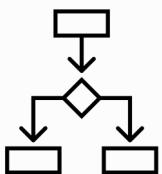
Under-explored Side Info and Flexibility



Varying evaluation costs



Smart stopping time



Observable multi-stage feedback



New design principle:
Gittins index

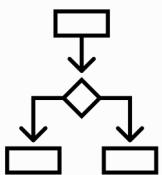
Why Gittins index?



Varying evaluation costs



Smart stopping time



Observable multi-stage feedback

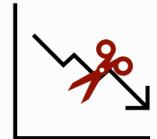
New design principle:
Gittins index

Optimal in related sequential
decision problems

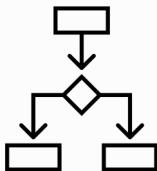
Why Gittins index?



Varying evaluation costs
Features in **Pandora's box**



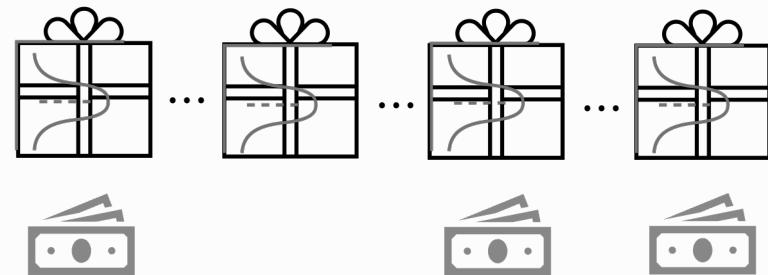
Smart stopping time
Features in **Pandora's box**



Observable multi-stage feedback

New design principle:
Gittins index

Optimal in related sequential
decision problems



Why Gittins index?



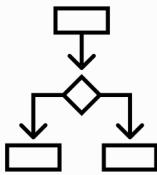
Varying evaluation costs

Features in Pandora's box



Smart stopping time

Features in Pandora's box

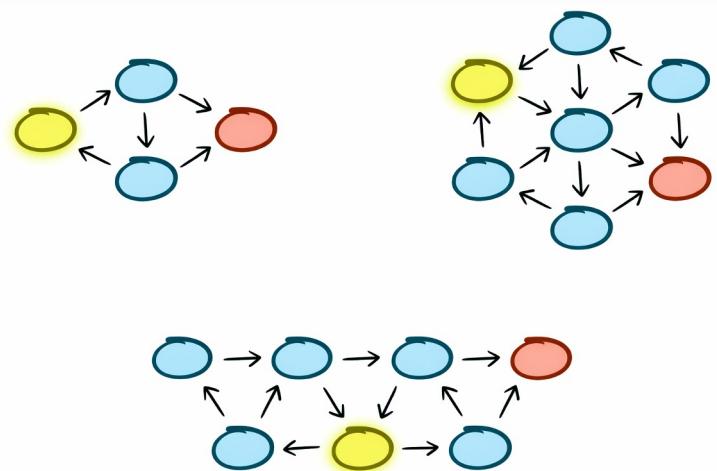


Observable multi-stage feedback

Features in Markov chain selection

New design principle:
Gittins index

Optimal in related sequential
decision problems



Why Gittins index?



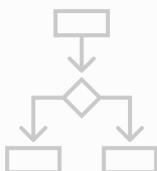
Varying evaluation costs

Features in **Pandora's box**



Smart stopping time

Features in **Pandora's box**



Observable multi-stage feedback

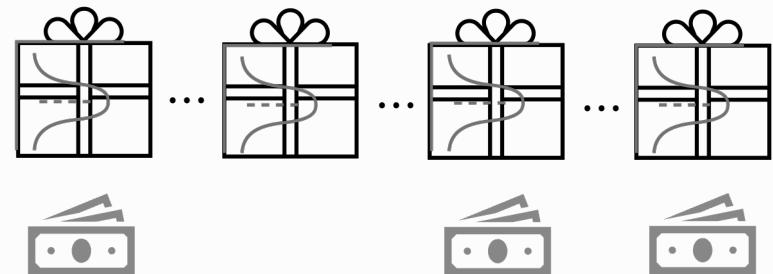
Features in Markov chain selection



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

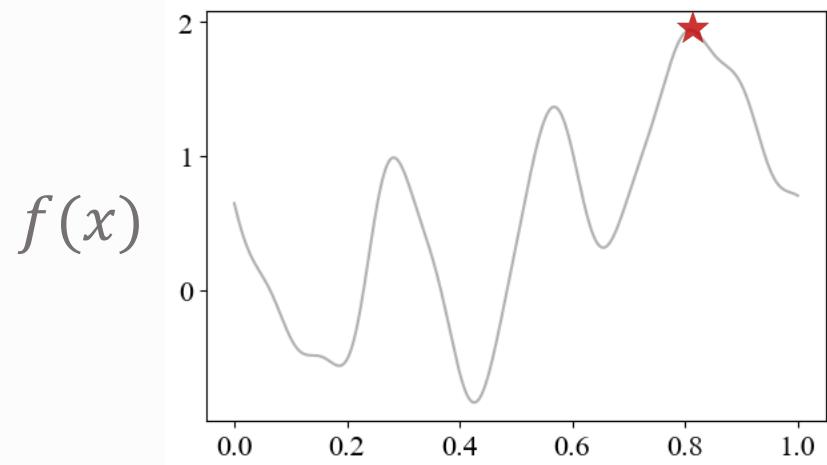
New design principle:
Gittins index

Optimal in related sequential
decision problems



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

Bayesian Optimization



Continuous

Correlated

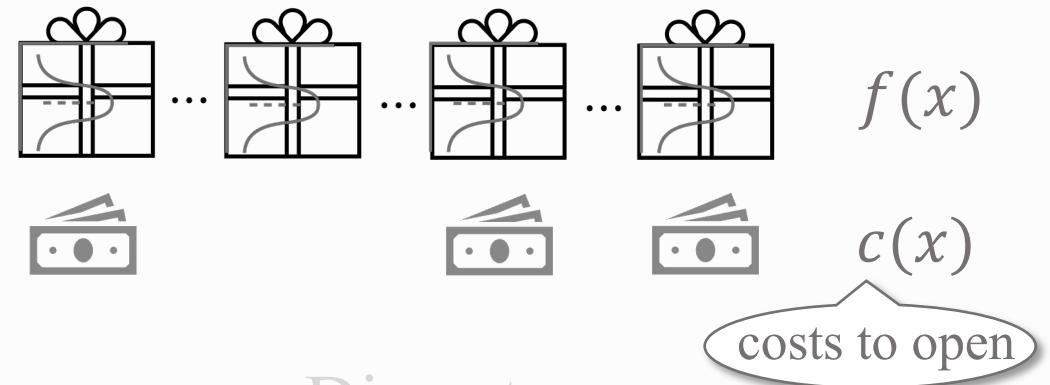
Fixed-iteration

Cost-unaware

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

Pandora's Box

[Weitzman'79]



Discrete

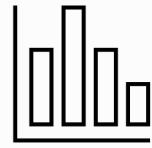
Independent

Flexible-stopping

Cost-aware

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

Expected Improvement vs Gittins Index



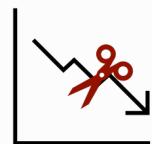
Varying evaluation costs

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

Arbitrarily bad

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

naturally incorporates costs



Smart stopping time

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

Which threshold?

$$\tau: \text{GI}(x_\tau; c) \leq y_{\text{best}}$$

$$\Leftrightarrow \tau: \text{EIPC}(x_\tau; c) \leq 1$$

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.

Our Contribution: Gittins Index Principle

Joint work with Ziv Scully and Alexander Terenin et al.

1. Principled easy-to-compute policy class

EI
✓

2. Natural incorporation of side info and flexibility

✗

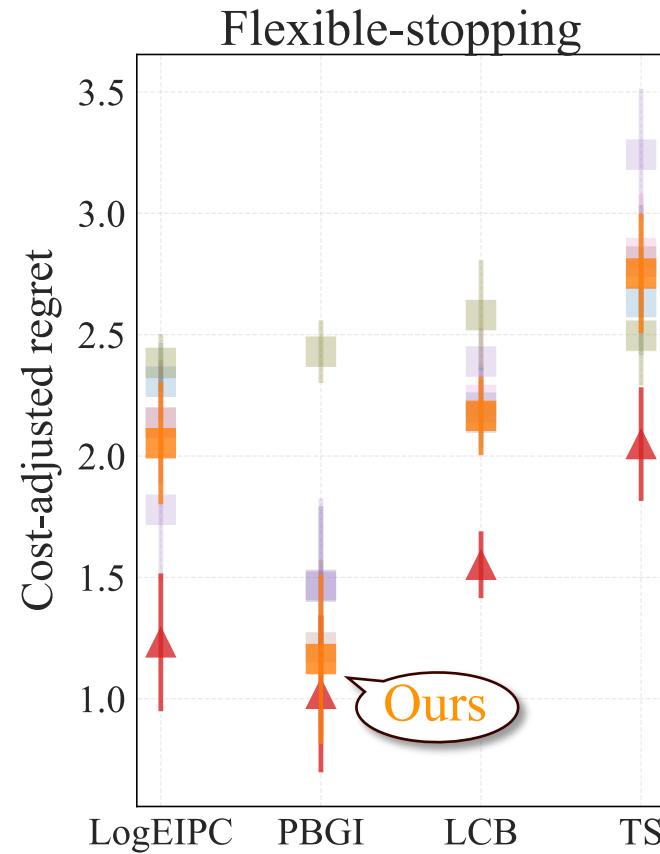
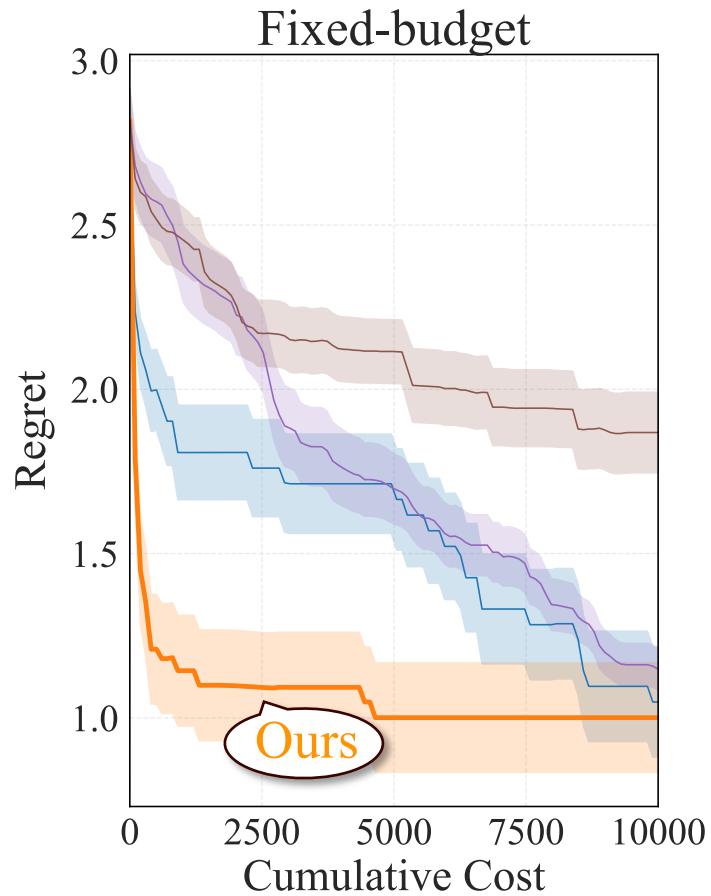
3. Competitive performance on benchmarks

✗

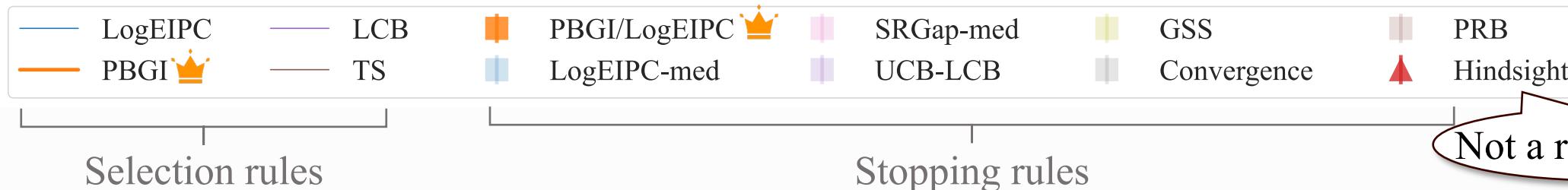
4. Theoretical guarantees

✗

Gittins Index vs Baselines on AutoML Benchmark



Lower the better



Our Contribution: Gittins Index Principle

Joint work with Ziv Scully and Alexander Terenin et al.

1. Principled easy-to compute policy class

EI
✓

2. Natural incorporation of side info and flexibility

✗

3. Competitive performance on benchmarks

✗

4. Theoretical guarantees

✗

Theoretical Guarantee and Empirical Validation

Theorem (Safeguard Guarantee)

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

or LogEIPC

cost-adjusted regret

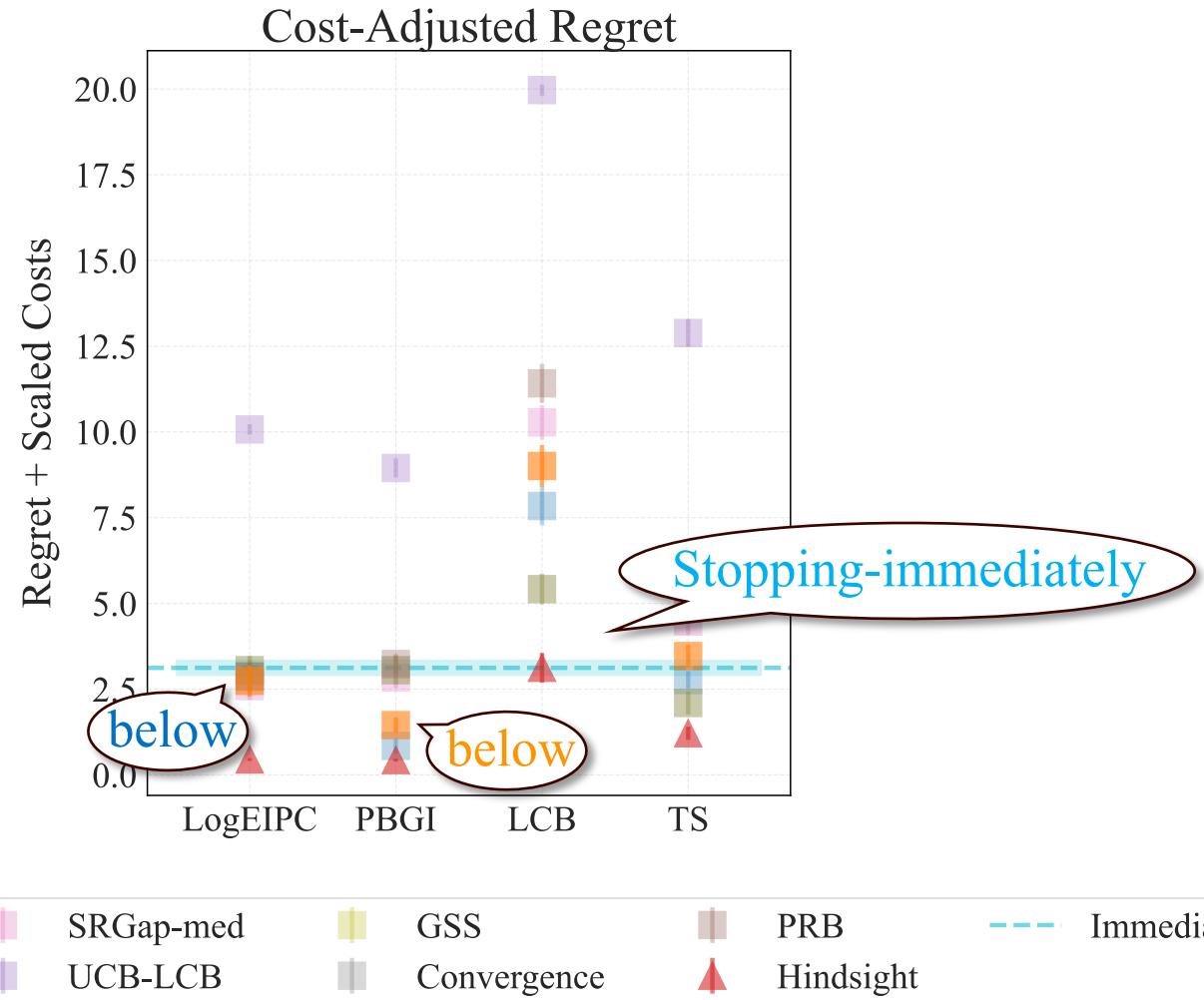
Implication:

- 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.

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

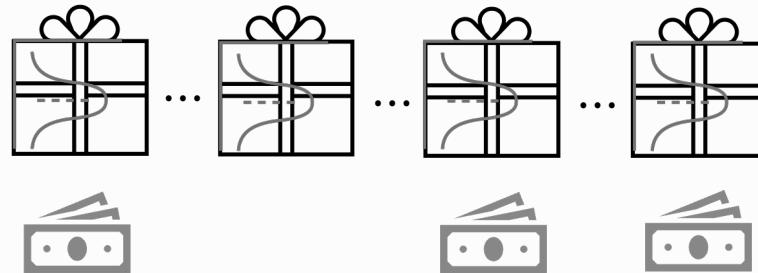
stopping time

PBGI/LogEIPC
LogEIPC-med



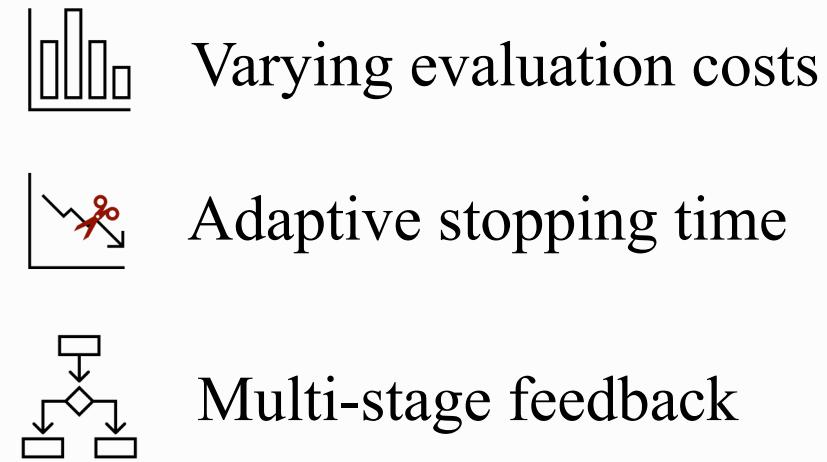
"Cost-aware Stopping for Bayesian Optimization." Under review.

Principled method design

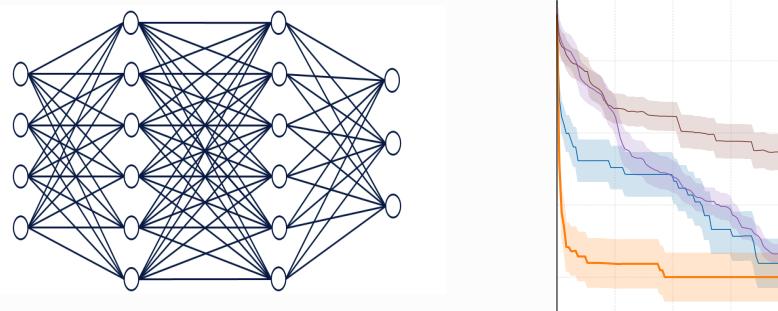


Link to **Pandora's Box** problem
& **Gittins index** theory

Natural incorporation of side info



Competitive empirical performance



Interests from practitioners (e.g., Meta)

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

Theoretical guarantees

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

Ongoing: regret bound

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