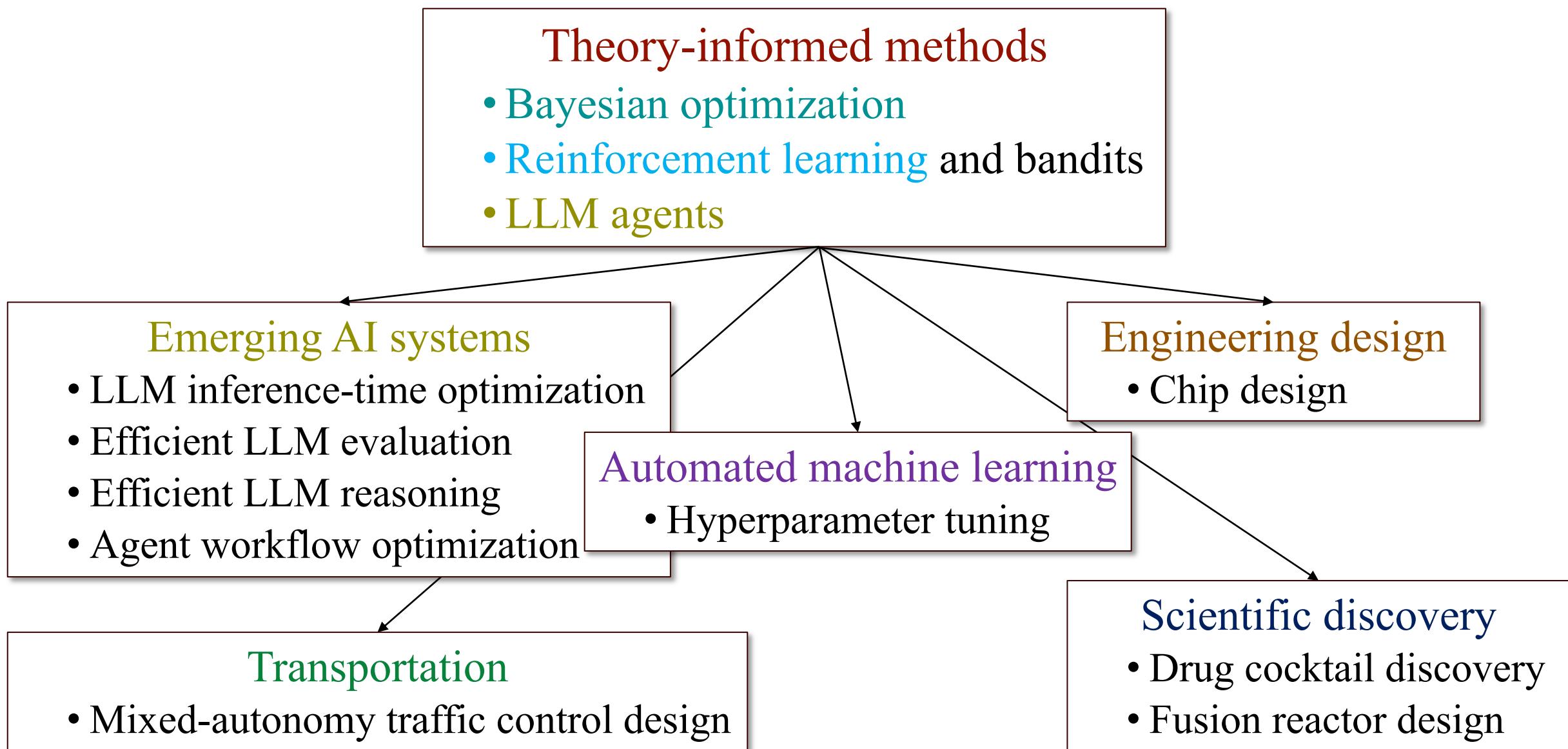


Cost-Aware Bayesian Optimization with Adaptive Stopping via the Pandora's Box Gittins Index

Qian Xie 谢倩 (Cornell ORIE)

Job Talk

Research Vision: Principled Efficient Decision-Making



Research Overview: ML–OR Decision-Making

- Data-efficient Black-Box Optimization (Recent)
 - Bayesian optimization via Gittins indices for hyperparameter optimization
[NeurIPS'24, ICML'26 (under review), OR (in prep)]



Research Overview: ML–OR Decision-Making

- Data-efficient Black-Box Optimization (Recent)

- Bayesian optimization via Gittins indices for hyperparameter optimization
[NeurIPS'24, ICML'26 (under review), OR (in prep)]
- LLM agent for mixed-autonomy traffic control design
[NeurIPS'25 LAW workshop]
- Scalable reinforcement learning for rectangle packing (chip design)
[DAC'25 (under review), RLC'26 (in prep)]



Research Overview: ML–OR Decision-Making

- Data-efficient Black-Box Optimization (Recent)
 - Bayesian optimization via Gittins indices for hyperparameter optimization [NeurIPS'24, ICML'26 (under review), OR (in prep)]
 - LLM agent for mixed-autonomy traffic control design [NeurIPS'25 LAW workshop]
 - Scalable reinforcement learning for rectangle packing (chip design) [DAC'25 (under review), RLC'26 (in prep)]
- Interactive Online Decision-making (Earlier)
 - Online recommendation (bandits) [ICML'23 & OR (major revision)]
 - Online resource allocation (MDP & stochastic game) [Automatica (2024)]

Research Overview: ML–OR Decision-Making

- Data-efficient Black-Box Optimization (Recent)

- Bayesian optimization via Gittins indices for hyperparameter optimization
[NeurIPS'24, ICML'26 (under review), OR (in prep)]
- LLM agent for mixed-autonomy traffic control design
[NeurIPS'25 LAW workshop]
- Scalable reinforcement learning for rectangle packing (macro placement)
[DAC'25 (under review), RLC'26 (in prep)]

This talk's focus

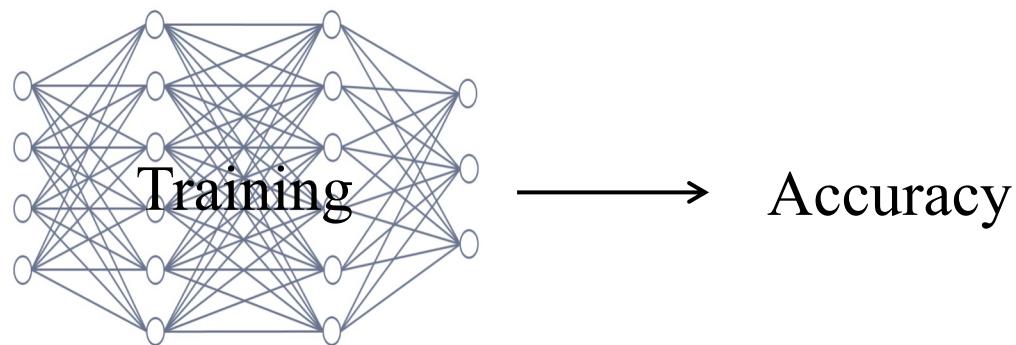
OR for ML

ML for OR

Motivation: World of Optimization under Uncertainty

ML model training:

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



Accuracy

Motivation: World of Optimization under Uncertainty

ML model training:

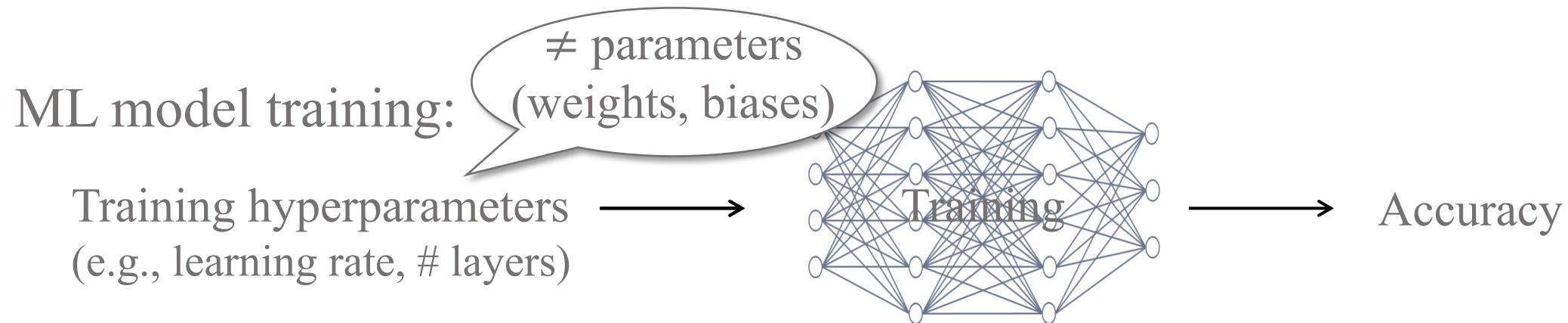
Training hyperparameters
(e.g., learning rate, # layers)

\neq parameters
(weights, biases)

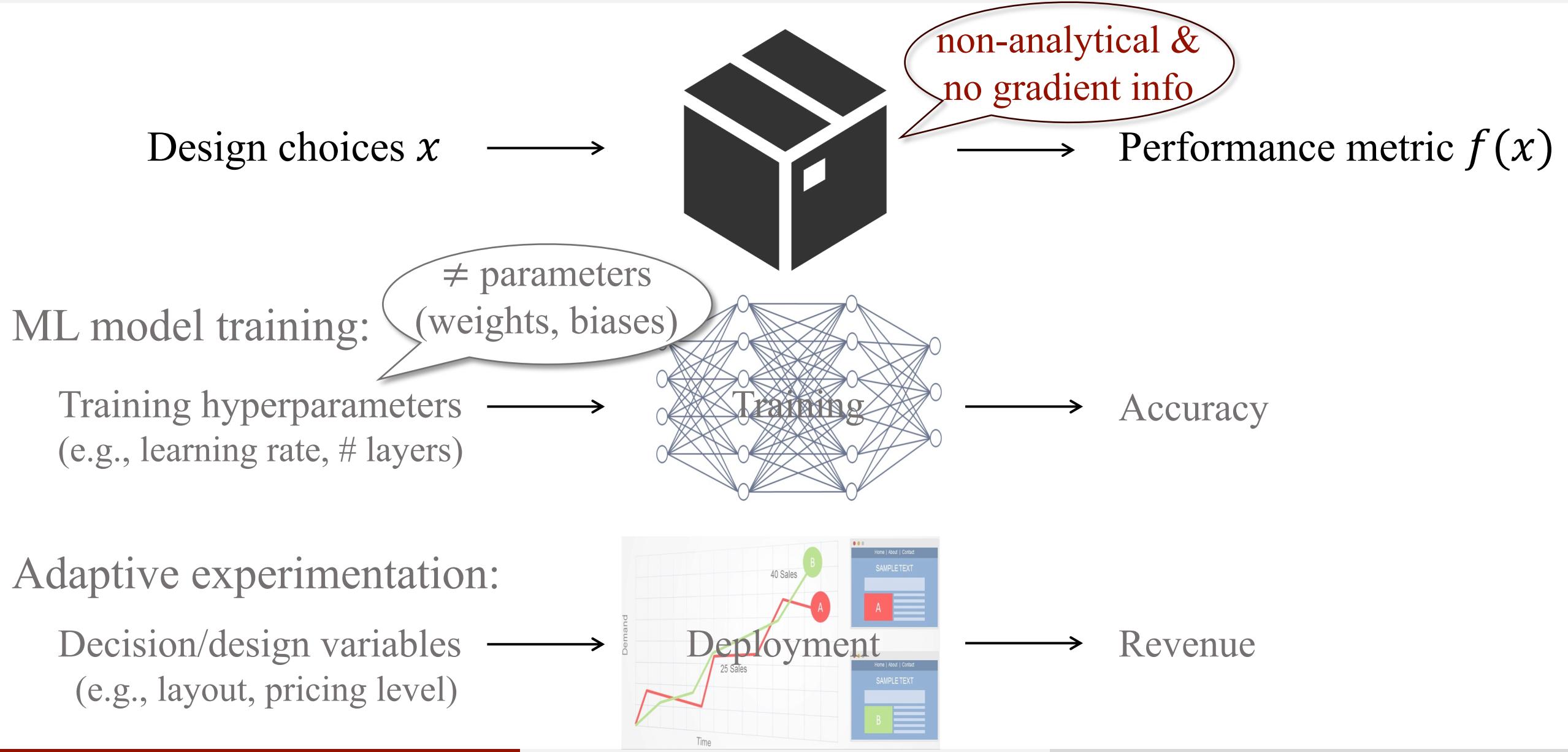


Accuracy

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

Observed outcome $f(x)$

ML model training:

\neq parameters
(weights, biases)

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



\longrightarrow Accuracy

Adaptive experimentation:

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



\longrightarrow Revenue

Background: Black-Box Optimization

Black-box optimization:

(gradient-based methods not applicable)

Input x →



expensive-to-evaluate

Observed outcome $f(x)$

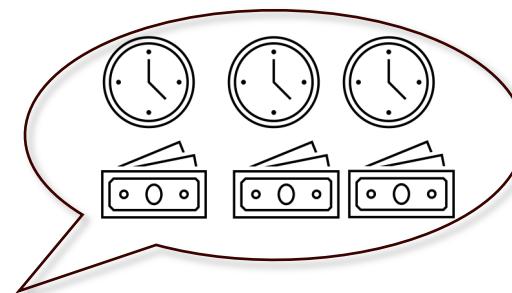
ML model training:

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



Training time

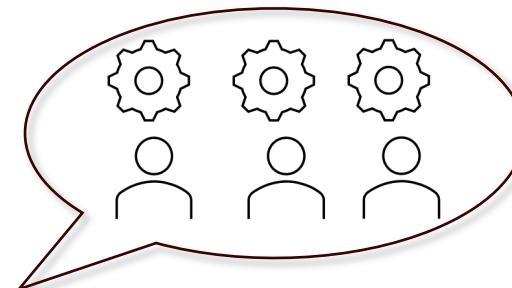
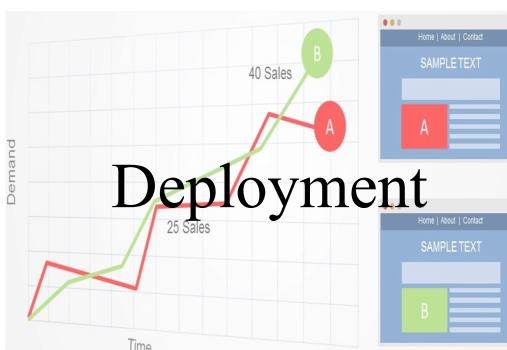
Compute credits



Accuracy

Adaptive experimentation:

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



Operational cost

User experience

Revenue

How to do black-box optimization efficiently?

Black-box optimization:

(gradient-based methods not applicable)

Input x →

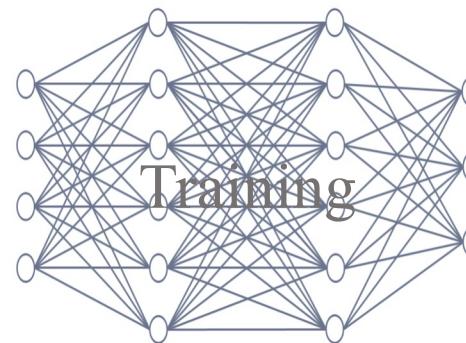


expensive-to-evaluate

Observed outcome $f(x)$

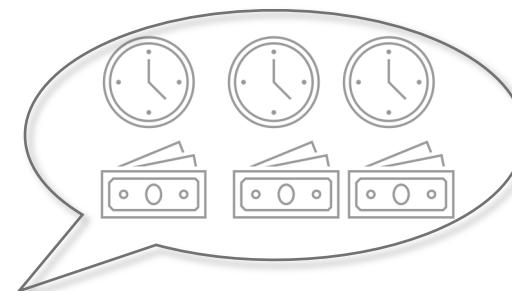
ML model training:

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



Training time

Compute credits



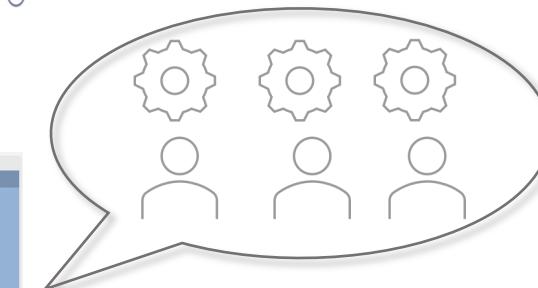
Accuracy

Adaptive experimentation:

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



Qian Xie (Cornell ORIE)



Operational cost

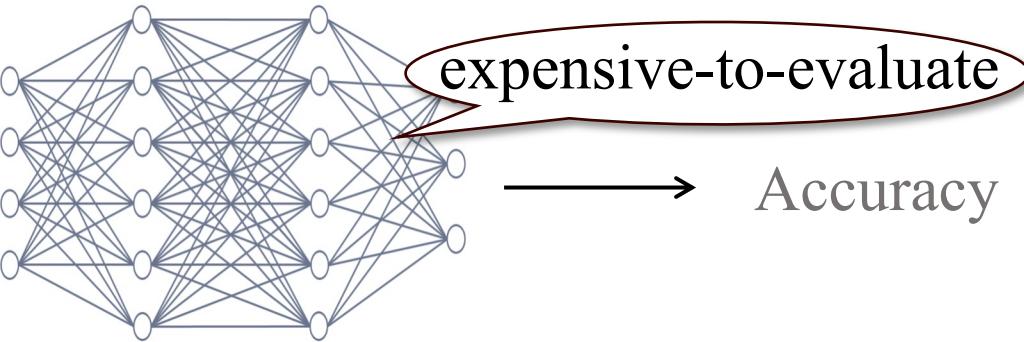
User experience

Revenue

Naïve Non-Adaptive Approach: Grid Search

ML model training:

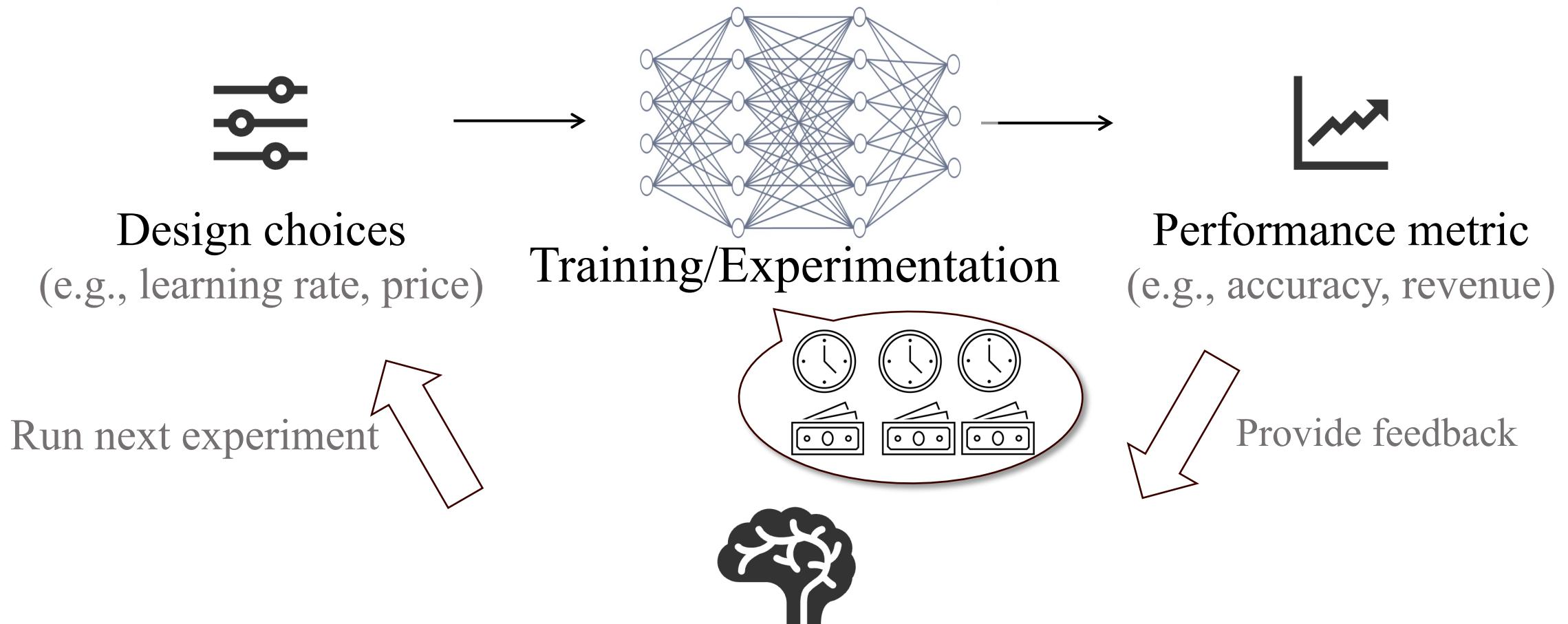
Training hyperparameters → Accuracy



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

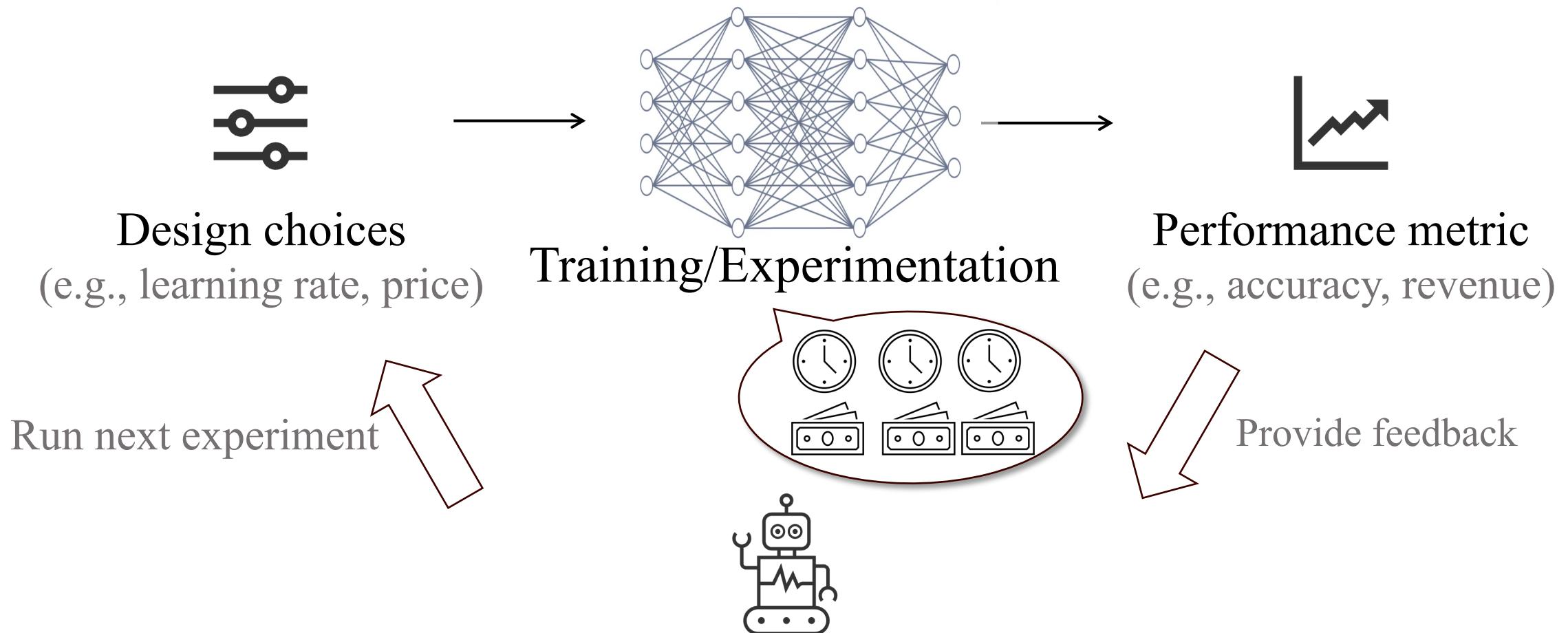
40,000,000
combinations!

Naïve Adaptive Approach: Manual Tuning



Experience-based human decision rule
(What to try next, when to stop)

Data-Driven Adaptive Approaches



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

Existing Umbrellas of Black-Box Optimization

Naïve approaches:

- Grid search
- Random search
- Manual tuning

Data-driven approaches:

- Local search
- Evolutionary algorithms
- Bayesian optimization
- Reinforcement learning
- LLM agent

New Methods for Black-Box Optimization

Naïve approaches:

- Grid search
- Random search
- Manual tuning

Data-driven approaches:

- Local search
- Evolutionary algorithms
- 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

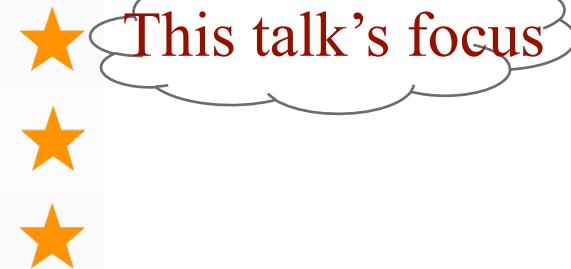
New Methods for Black-Box Optimization

Naïve approaches:

- Grid search
- Random search
- Manual tuning

Data-driven approaches:

- Local search
- Evolutionary algorithms
- Bayesian optimization
- Reinforcement learning
- LLM agent

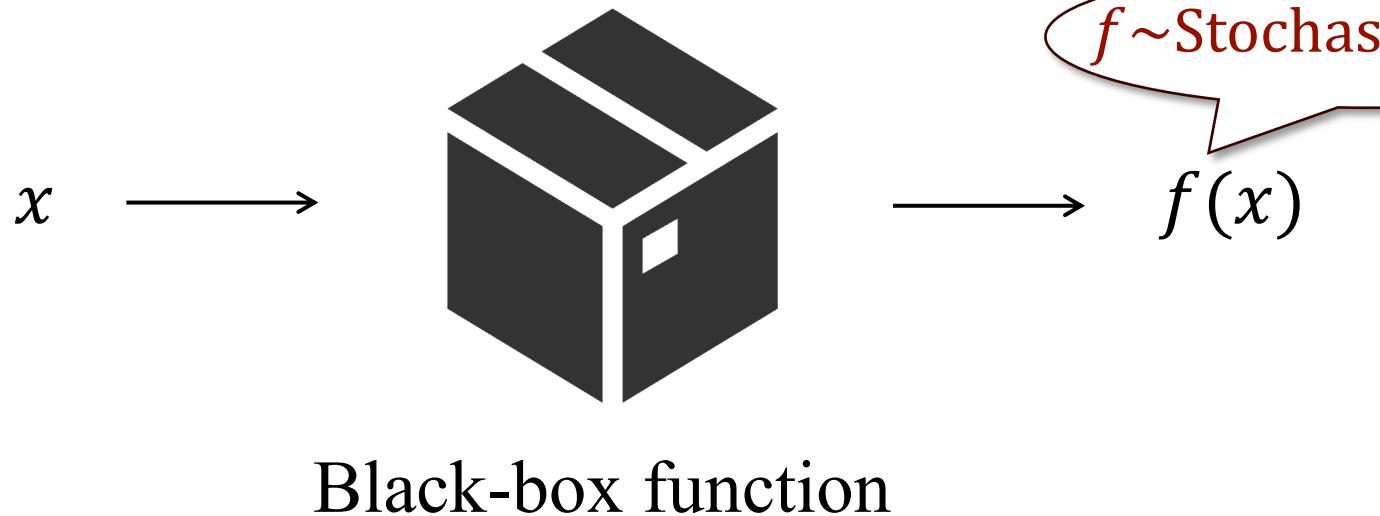


Contributions of new methods proposed in my work:

1. Novel connection to related decision problems
2. Principled decision rules
3. Competitive empirical performance

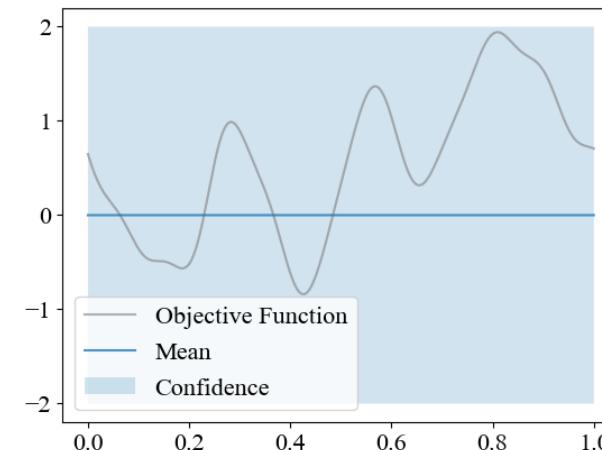
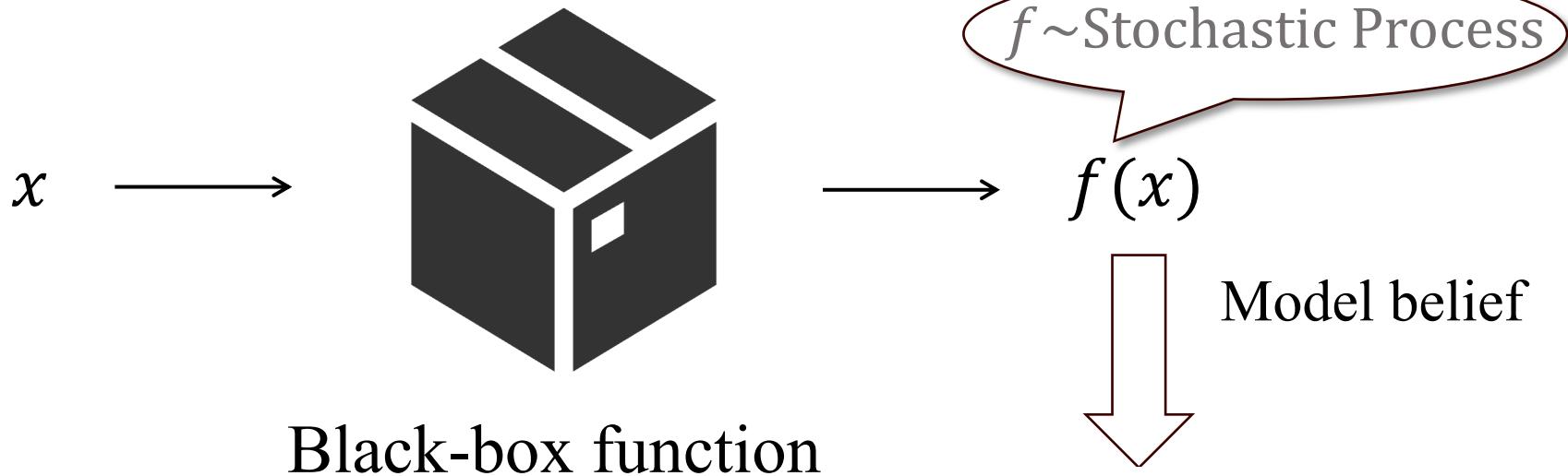
★ New methods under this umbrella

Bayesian Optimization



Bayesian Optimization

Time 0

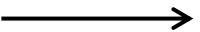


Probabilistic model
(e.g., Gaussian process)

Bayesian Optimization

Time t

x_1, \dots, x_t

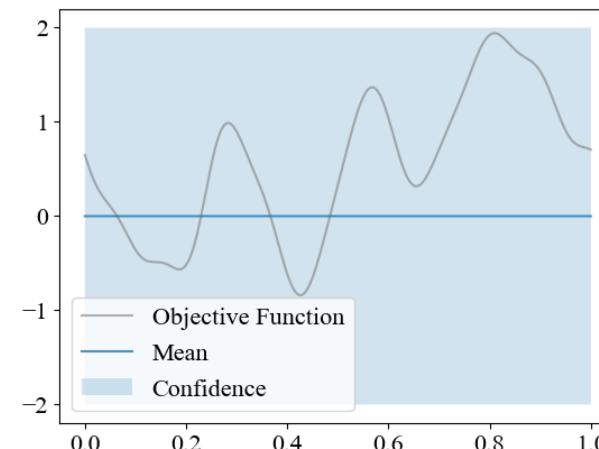


Black-box function

$f \sim \text{Stochastic Process}$

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

Model belief



Probabilistic model
(e.g., Gaussian process)

Bayesian Optimization

Time t

x_1, \dots, x_t \longrightarrow

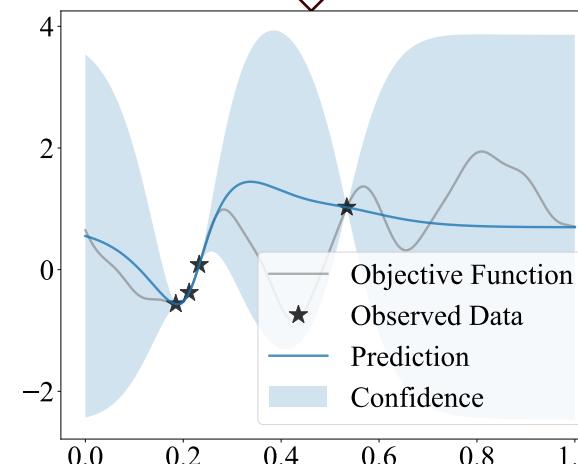


Black-box function

$f \sim \text{Stochastic Process}$

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

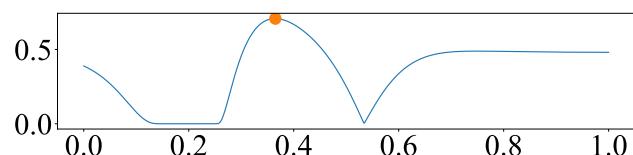
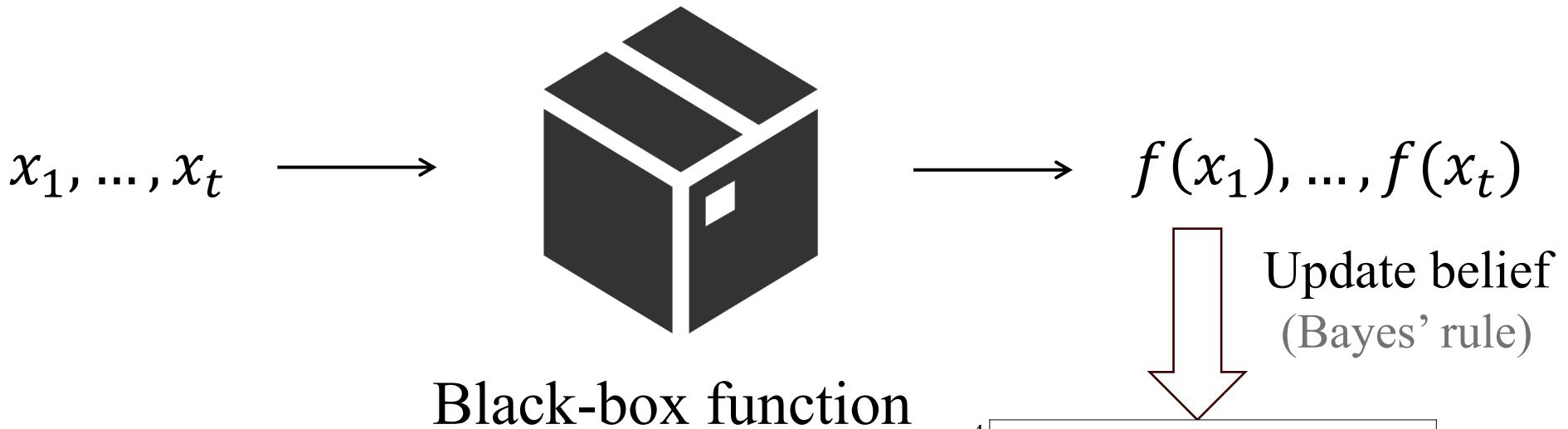
Update belief
(Bayes' rule)



Probabilistic model
(e.g., Gaussian process)

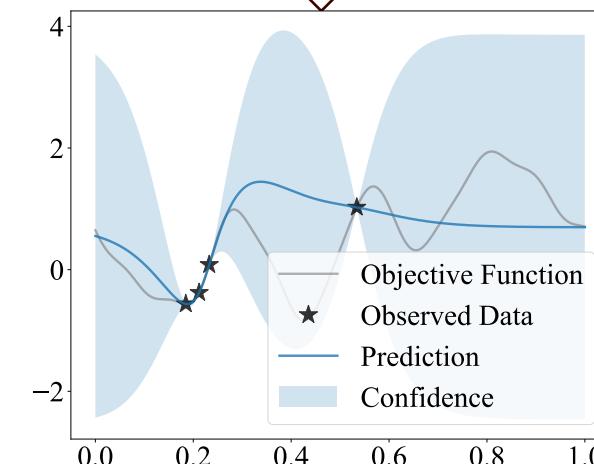
Bayesian Optimization

Time t



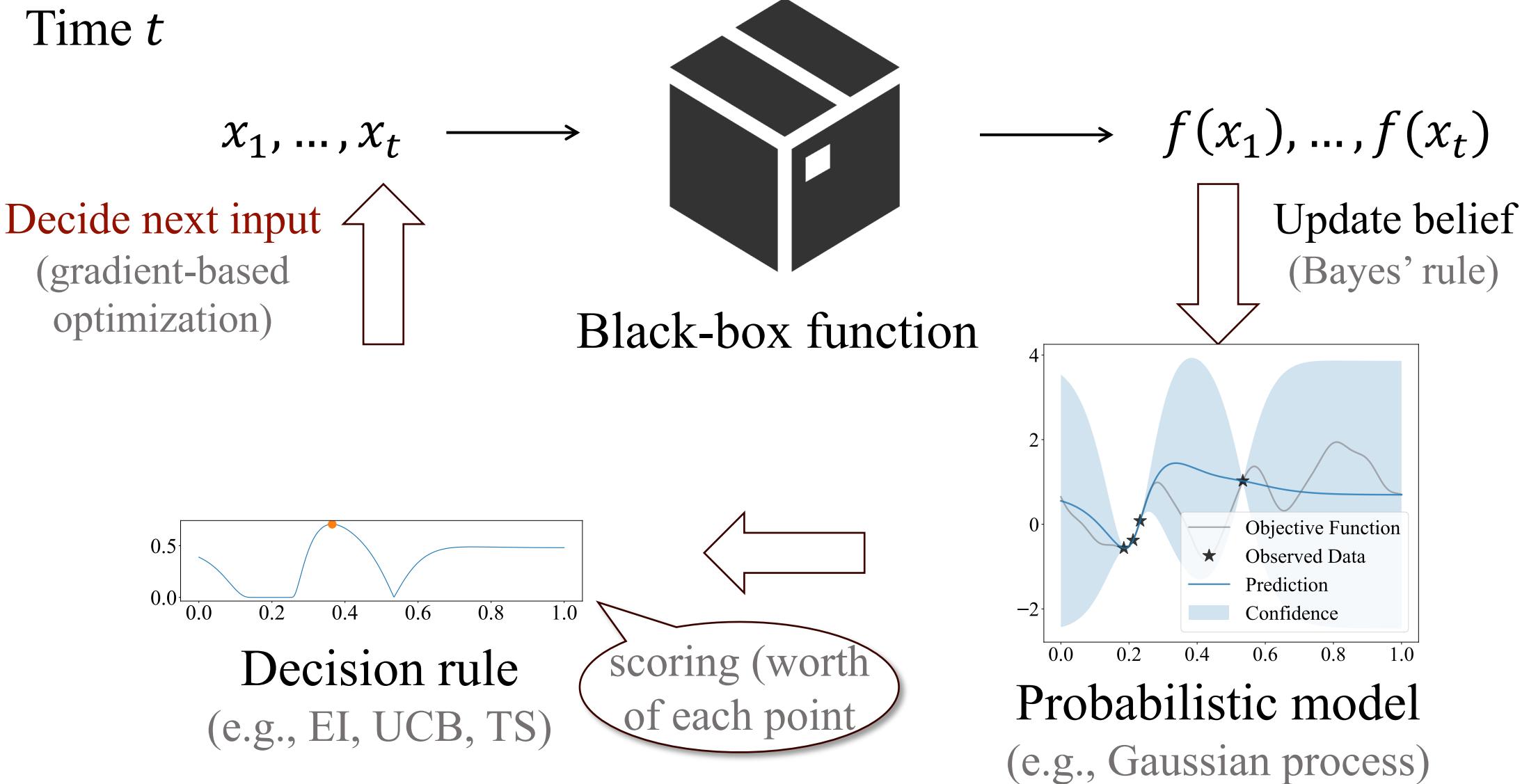
Decision rule
(e.g., EI, UCB, TS)

scoring (worth
of each point)

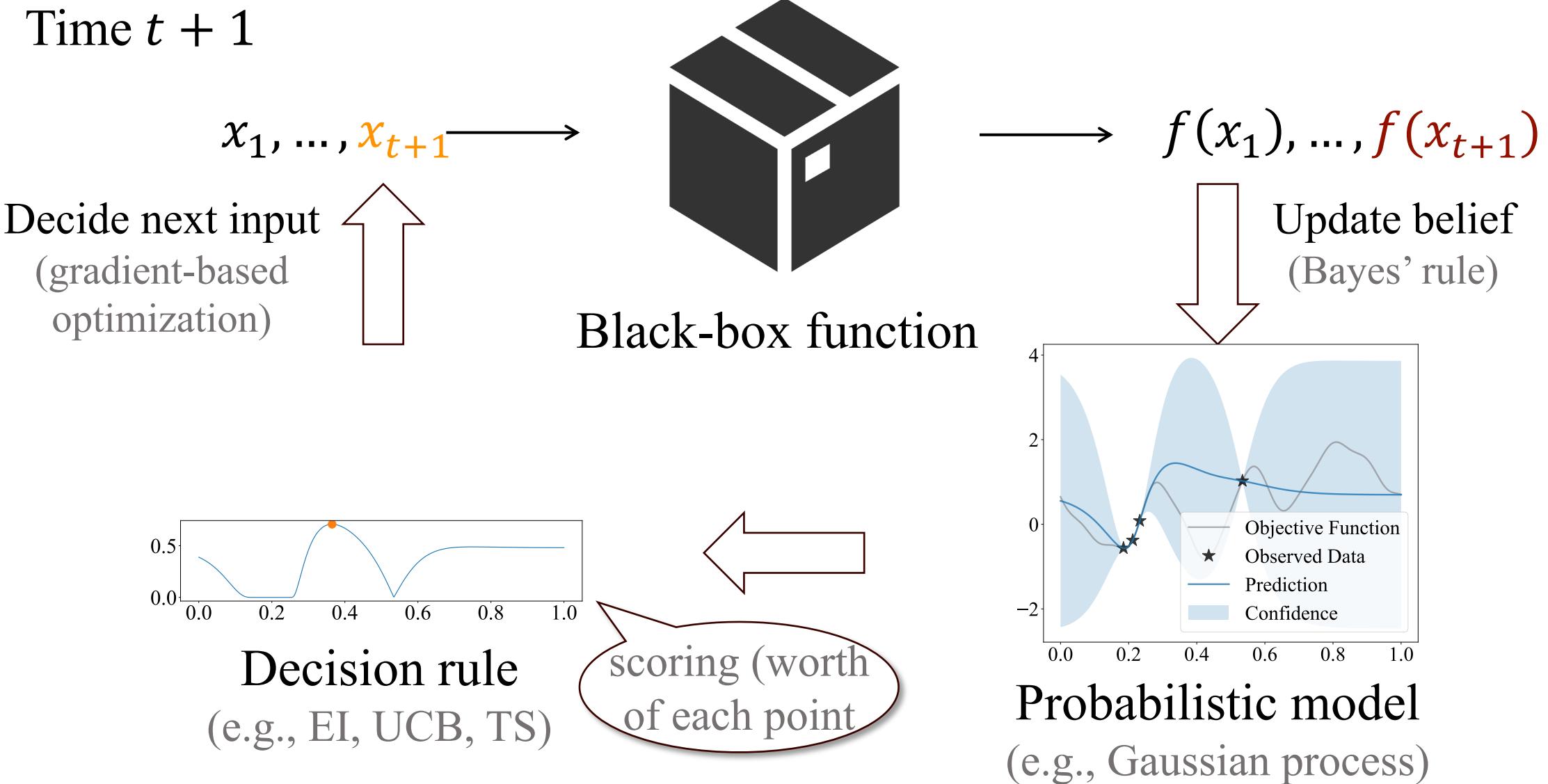


Probabilistic model
(e.g., Gaussian process)

Bayesian Optimization



Bayesian Optimization



Bayesian Optimization

Time $t + 1$

Decide next input
(gradient-based
optimization)

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

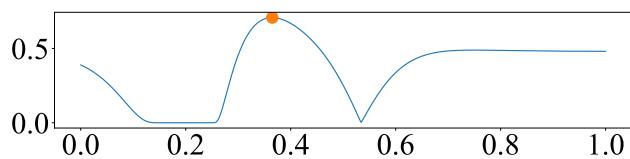


$f(x_1), \dots, f(x_{t+1})$

Update belief
(Bayes' rule)

Black-box function

My focus



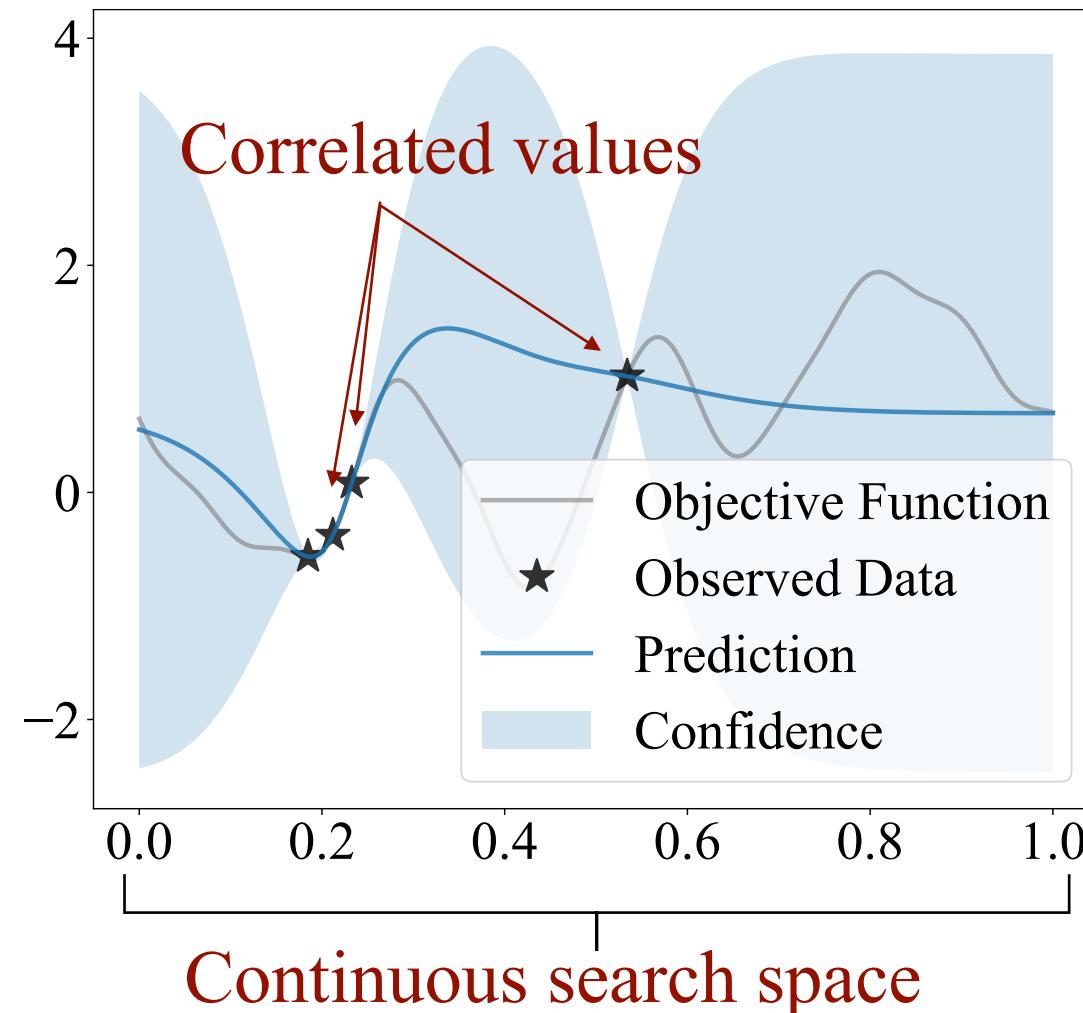
Decision rule
(e.g., EI, UCB, TS)

scoring (worth
of each point)



Probabilistic model
(e.g., Gaussian process)

Challenges in Decision Rule Design



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

Popular Decision Rule: Expected Improvement

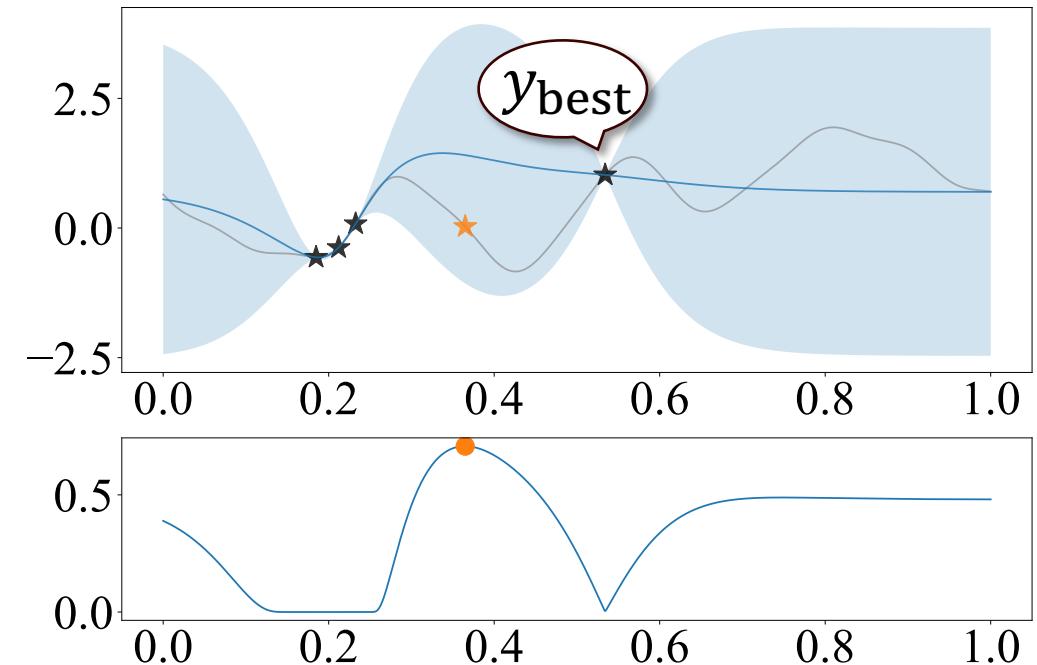
$$EI(x) = \mathbb{E}[\max(f(x) - y_{\text{best}}, 0) \mid x_1, \dots, x_t]$$

current best observed data D
"improvement"

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

posterior distribution

One-step approximation to MDP



Expected improvement $EI(x)$

Popular Decision Rule: Expected Improvement

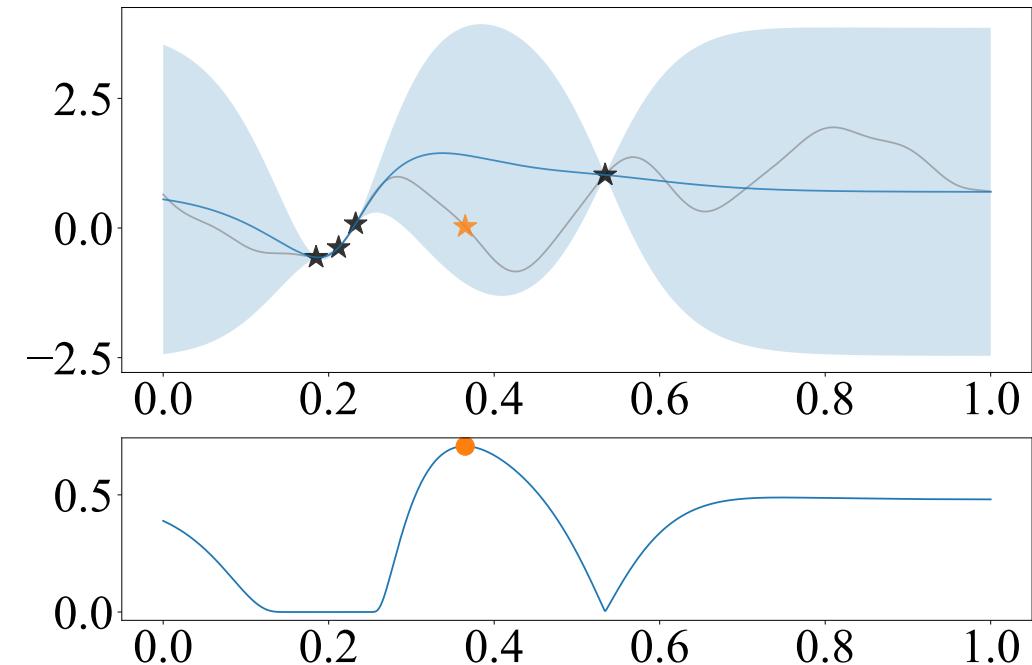
$$EI(x) = \mathbb{E}[\max(f(x) - y_{\text{best}}, 0) \mid x_1, \dots, x_t]$$

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

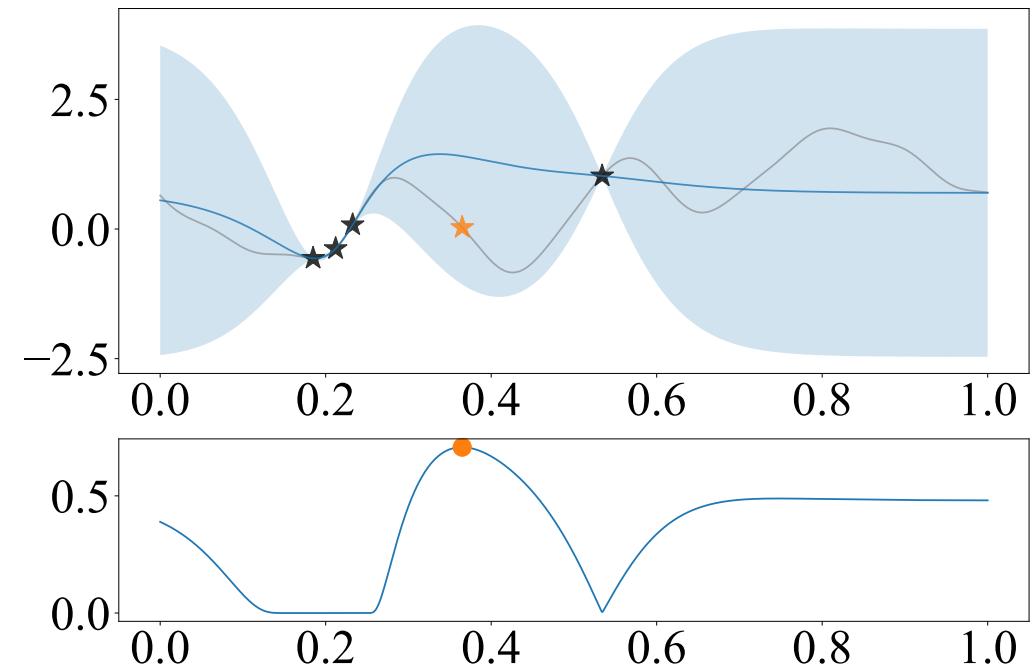


Expected improvement $EI(x)$

Improvement-based
design principle

Existing Design Principles

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

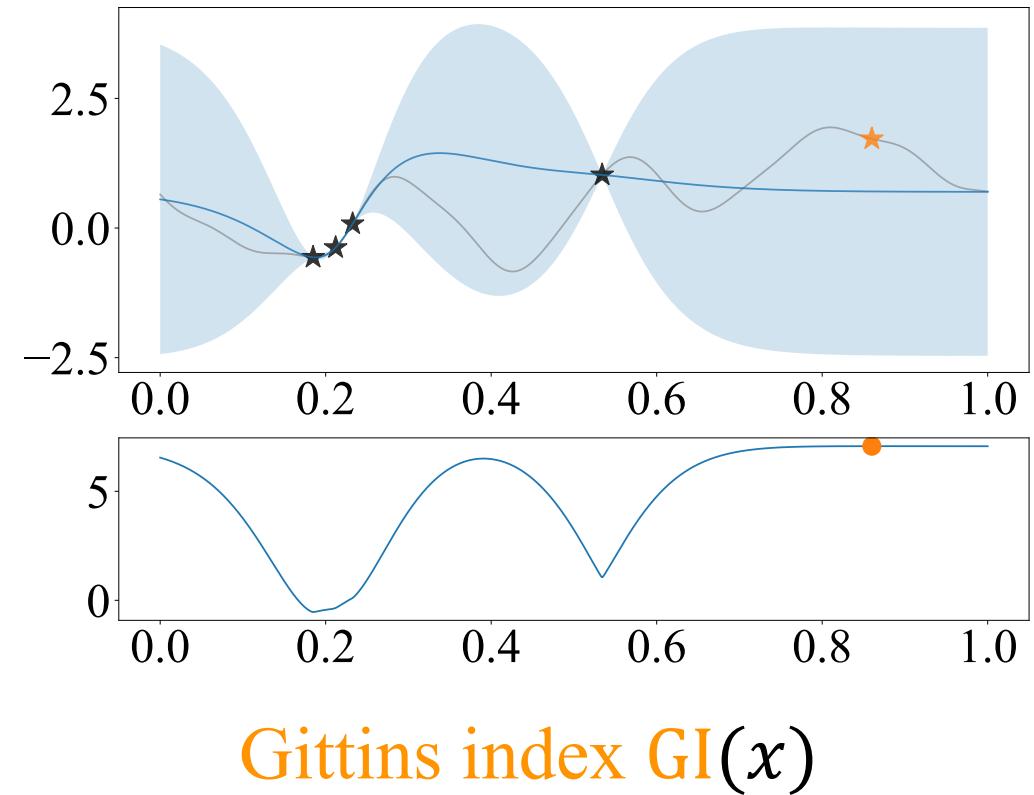


Expected improvement $EI(x)$

Improvement-based
design principle

New Design Principle: Gittins Index

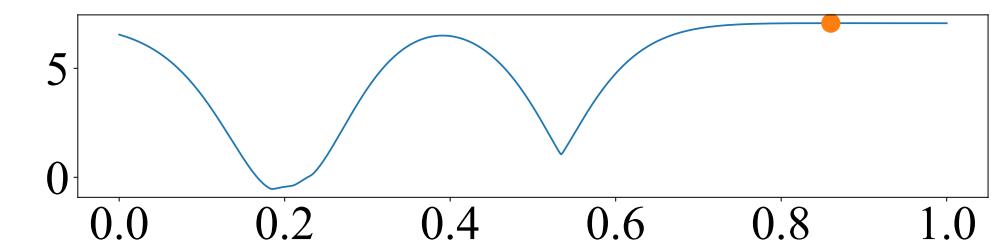
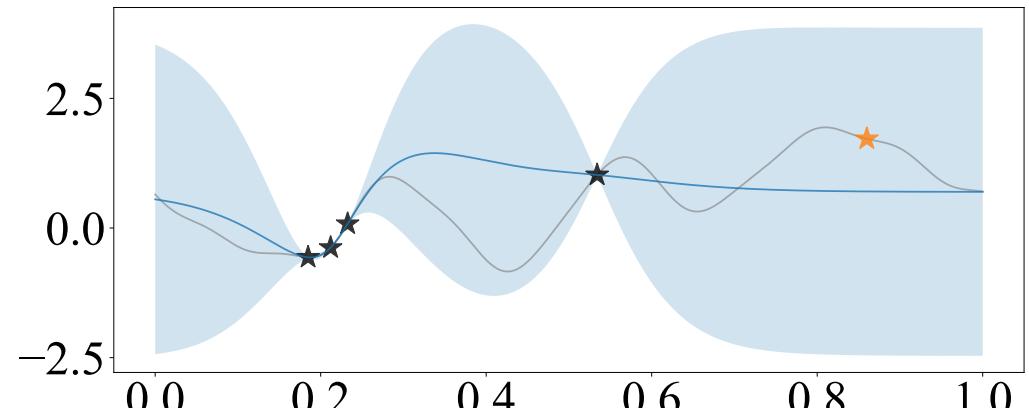
- Improvement-based (e.g., EI)
- Entropy-based
- Confidence bounds (UCB/LCB)
- Thompson sampling (TS)
- **Gittins Index**



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**

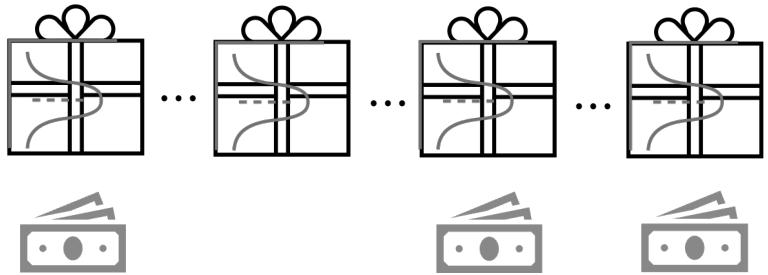


Gittins index $GI(x)$

? Why another principle?

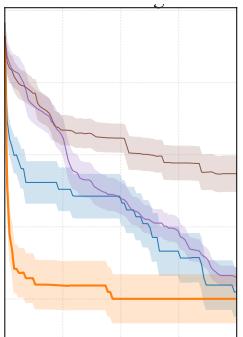
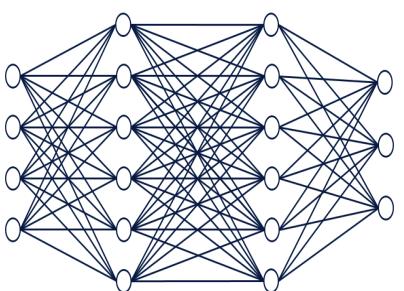
Our Contribution: Gittins Index Principle

Novel connection



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

Competitive empirical performance



Interests from practitioners (e.g., Meta)

Principled decision rules

- Varying evaluation costs
- Adaptive stopping time

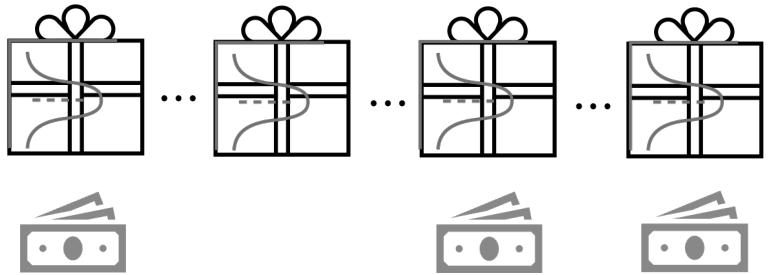
Unified framework for selection and stopping

Future potential

- Best-prompt identification
- Adaptive response sampling
- Chain-of-thought selection
- Application to **efficient LLM**

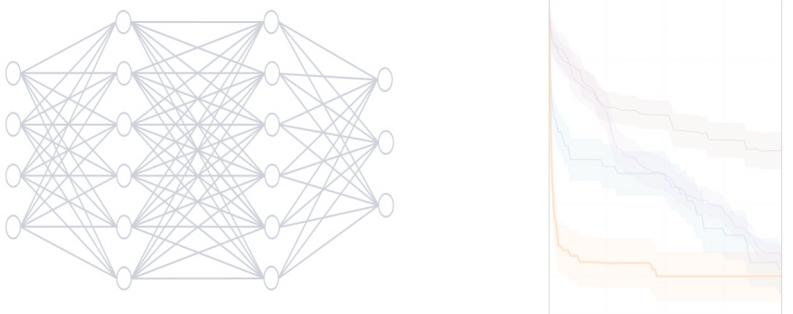
Our Contribution: Gittins Index Principle

Novel connection



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

Competitive empirical performance



Interests from practitioners (e.g., Meta)

Principled decision rules



Unified framework for cost-aware selection and stopping

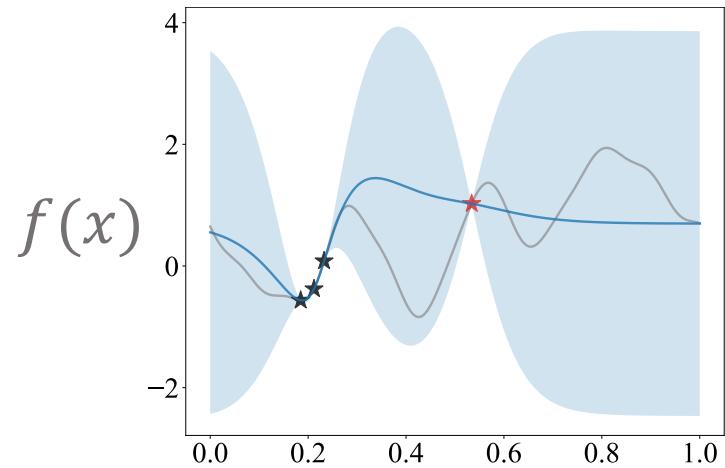
Future potential
Best-prompt identification



Chain-of-thought selection

Application to efficient LLM

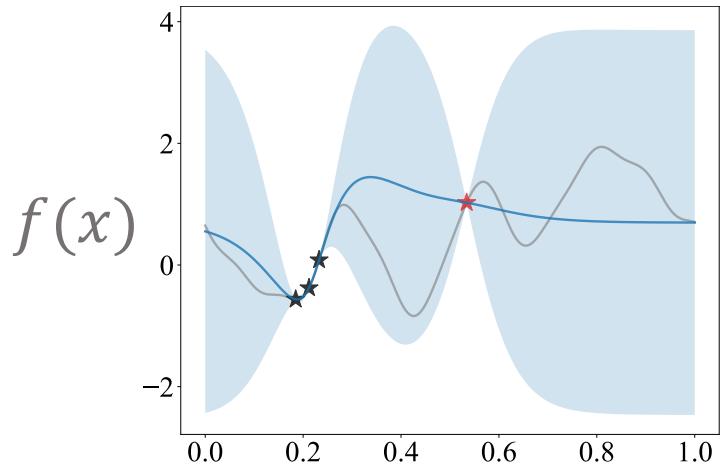
Bayesian Optimization



Continuous search space

Correlated function values

Bayesian Optimization



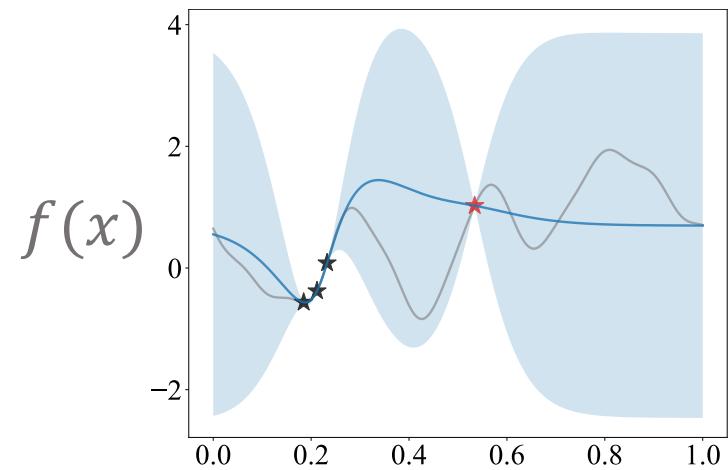
Continuous search space \Rightarrow

Discrete

Correlated function values \Rightarrow

Independent

Bayesian Optimization

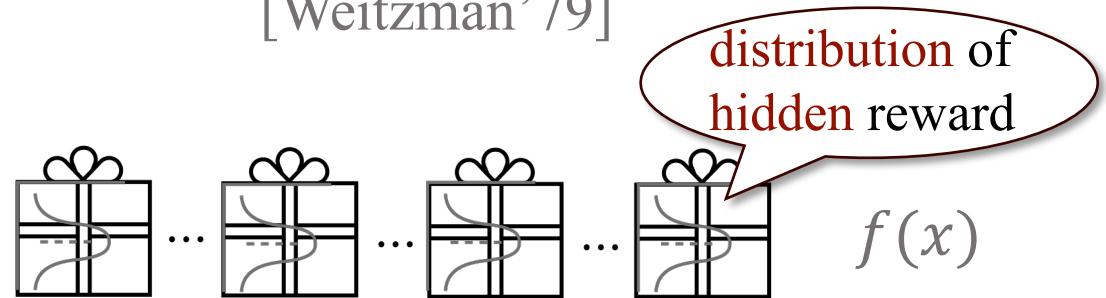


Continuous search space

Correlated function values

Pandora's Box

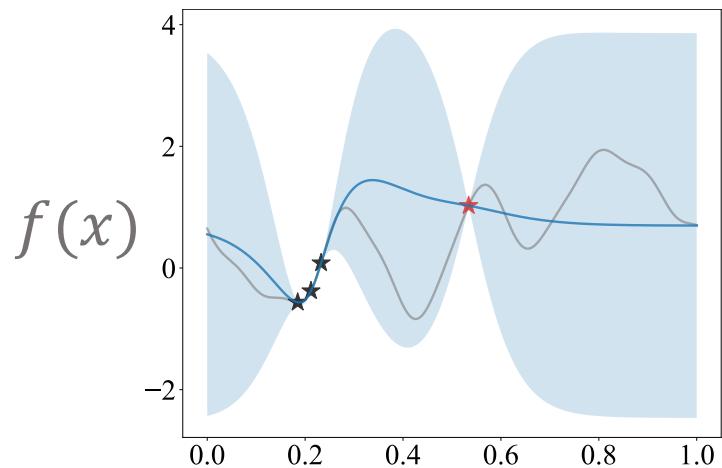
[Weitzman'79]



Discrete

Independent

Bayesian Optimization

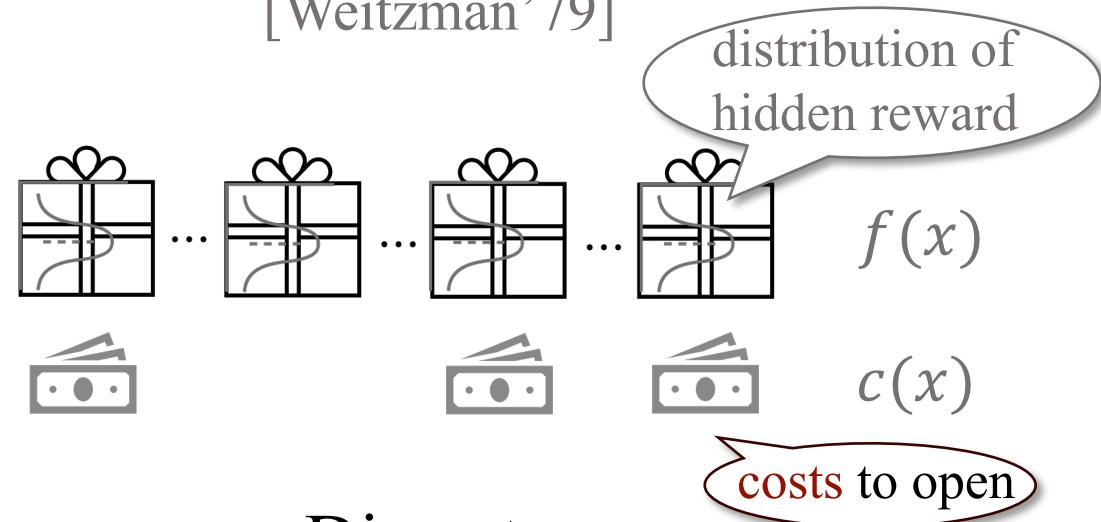


Continuous search space

Correlated function values

Pandora's Box

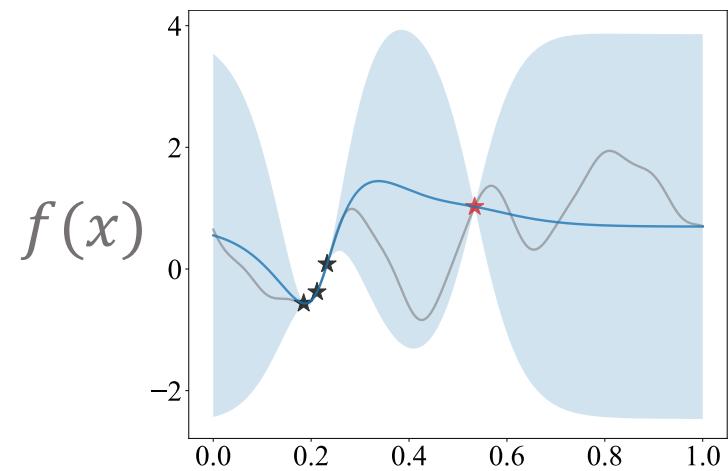
[Weitzman'79]



Discrete

Independent

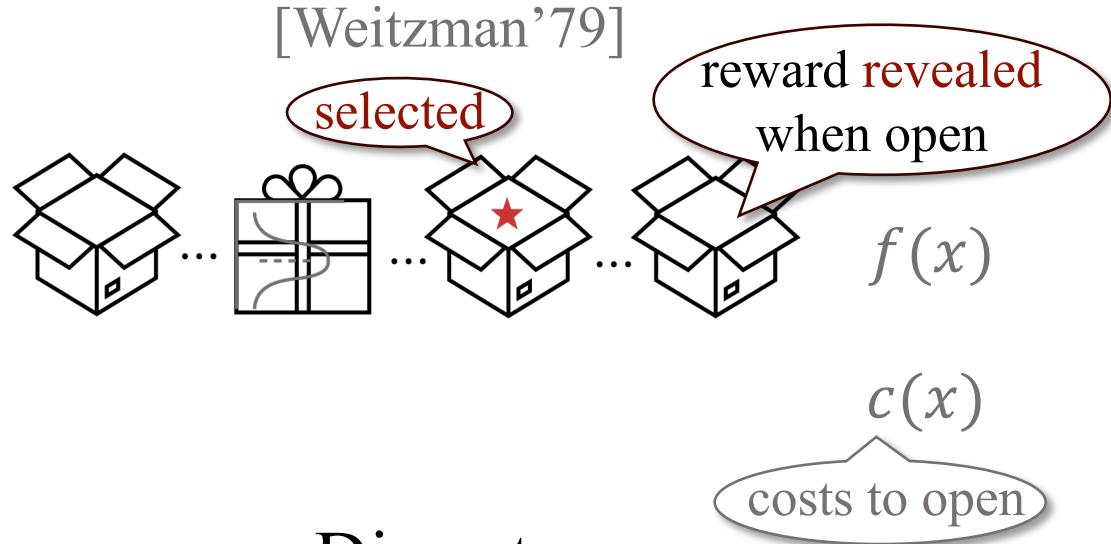
Bayesian Optimization



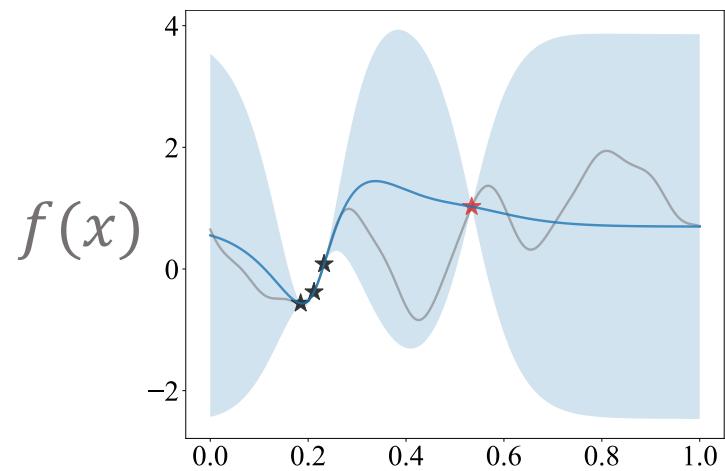
Continuous search space

Correlated function values

Pandora's Box



Bayesian Optimization

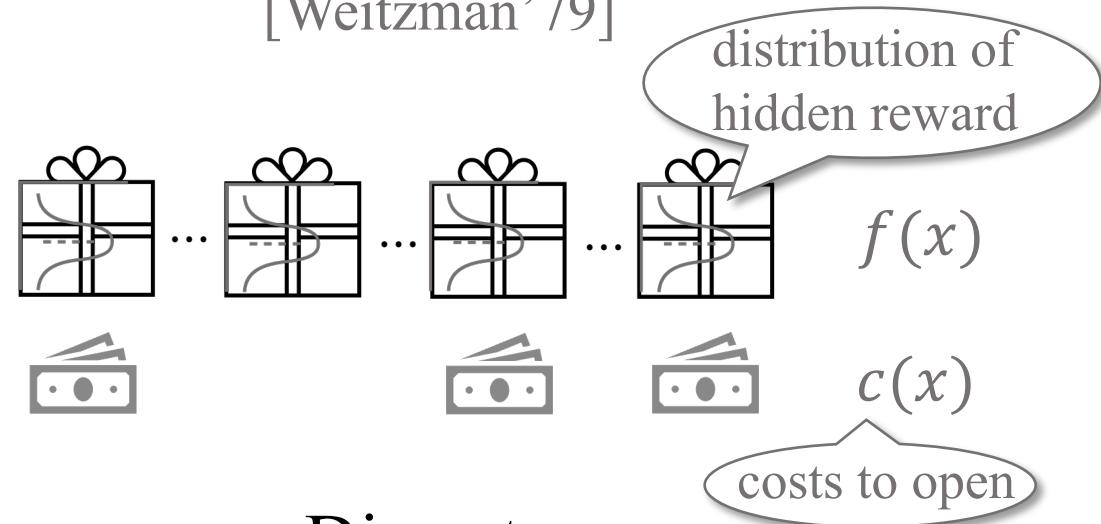


Continuous search space

Correlated function values

Pandora's Box

[Weitzman'79]

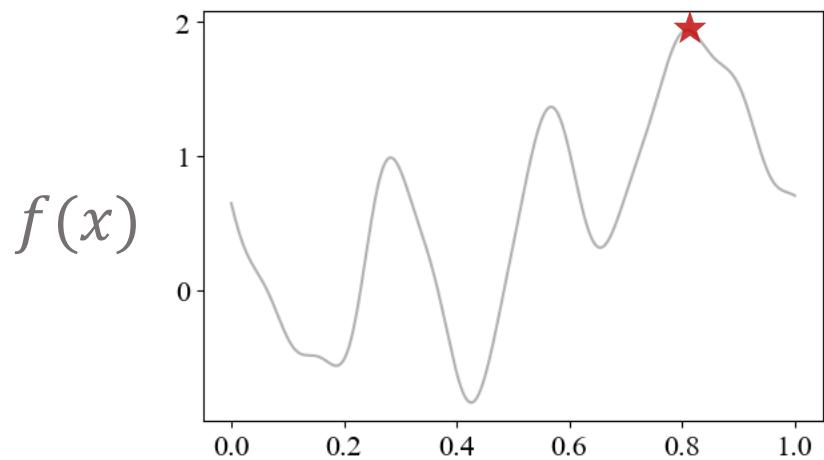


Discrete

Independent

Optimal policy: Gittins index

Bayesian Optimization

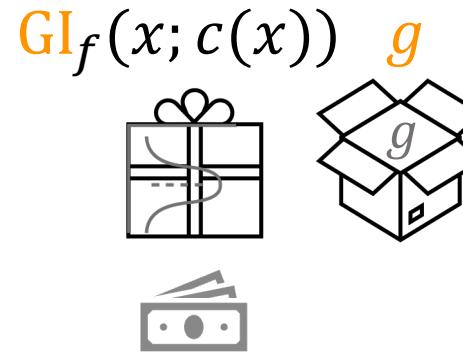


Continuous

Correlated

Pandora's Box

[Weitzman'79]



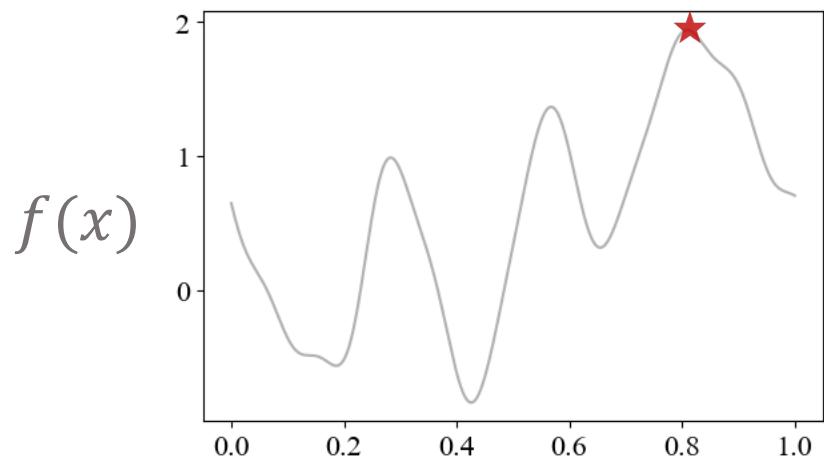
Discrete

Independent

Optimal policy: Gittins index

Step 1: Assign each box a Gittins index

Bayesian Optimization

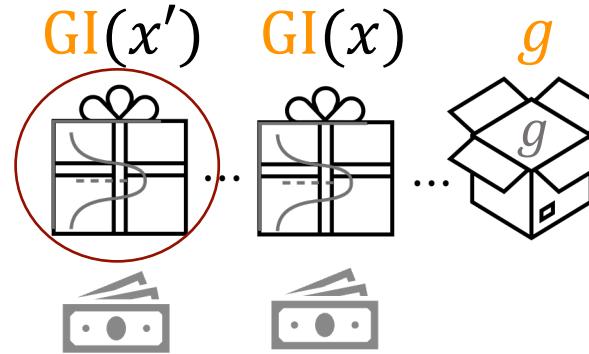


Continuous

Correlated

Pandora's Box

[Weitzman'79]



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

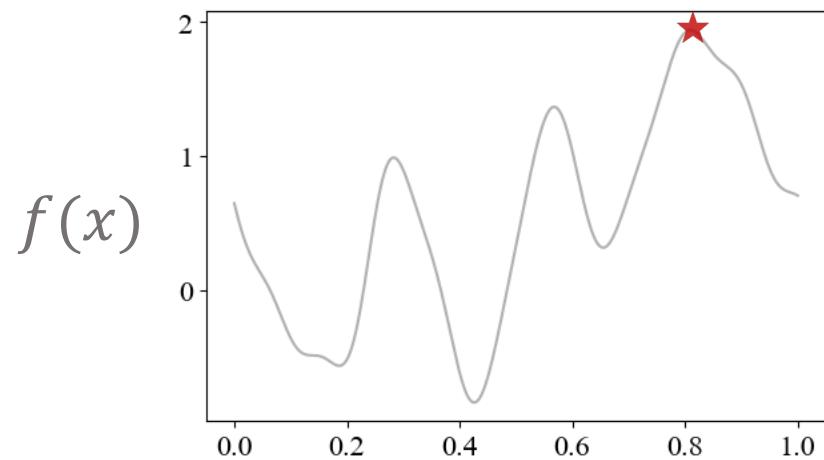
- *Closed*: open it
- *Opened*: select & stop

Discrete

Independent

Optimal policy: **Gittins index**

Bayesian Optimization

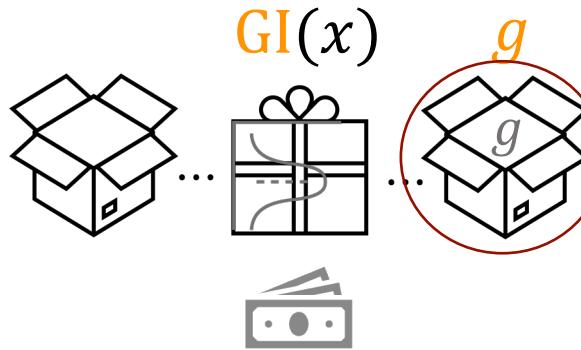


Continuous

Correlated

Pandora's Box

[Weitzman'79]



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

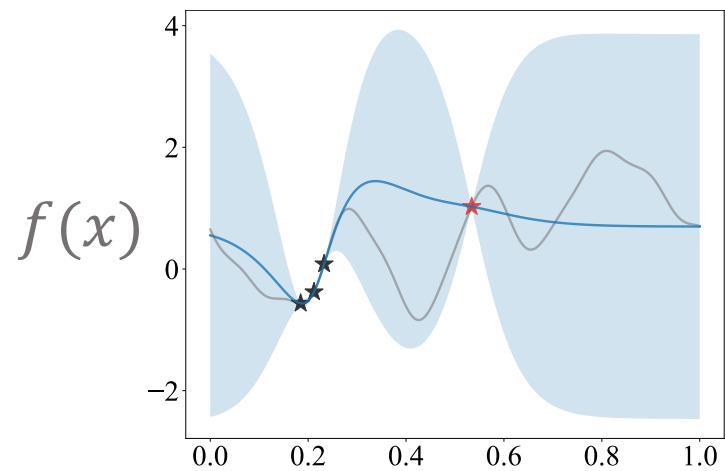
- *Closed*: open it
- *Opened*: select & stop

Discrete

Independent

Optimal policy: **Gittins index**

Bayesian Optimization

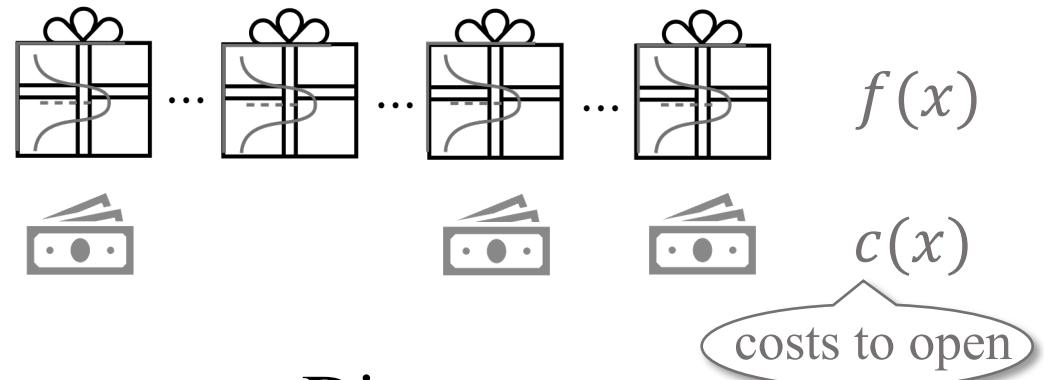


Continuous search space

Correlated function values

Pandora's Box

[Weitzman'79]



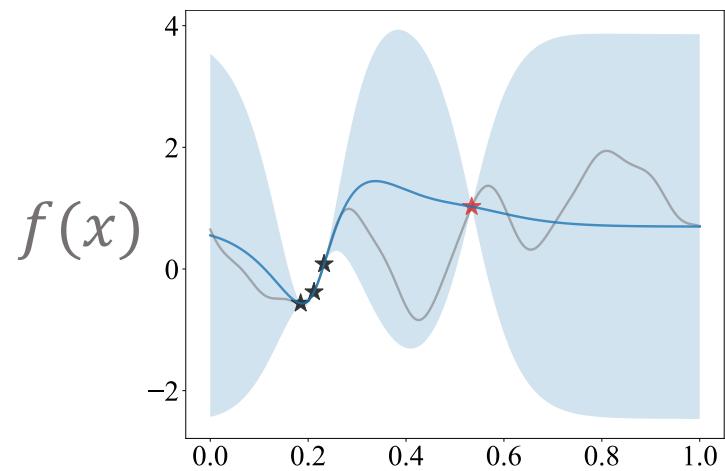
Discrete

Independent

How to translate?

Optimal policy: Gittins index

Bayesian Optimization

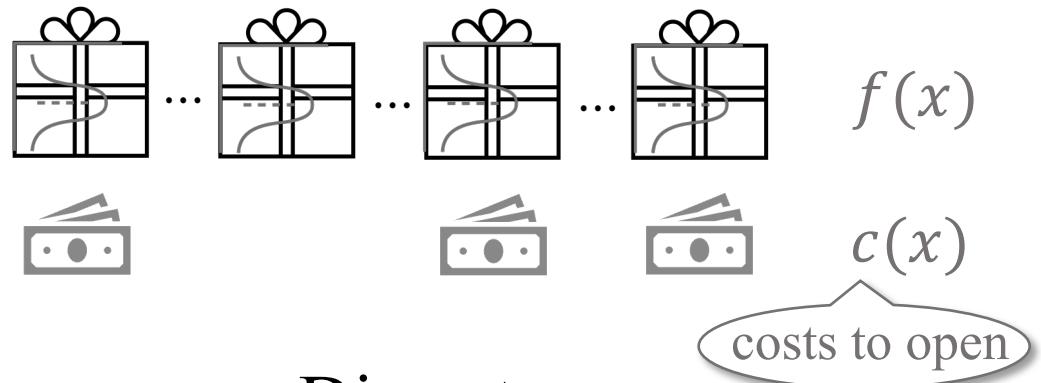


Continuous search space

Correlated function values

Pandora's Box

[Weitzman'79]



Discrete

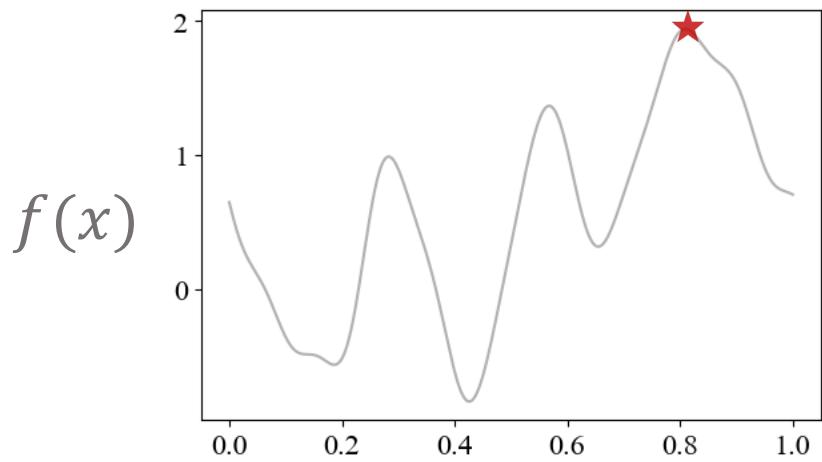
Independent

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

incorporate posterior
take continuum limit

New!

Bayesian Optimization



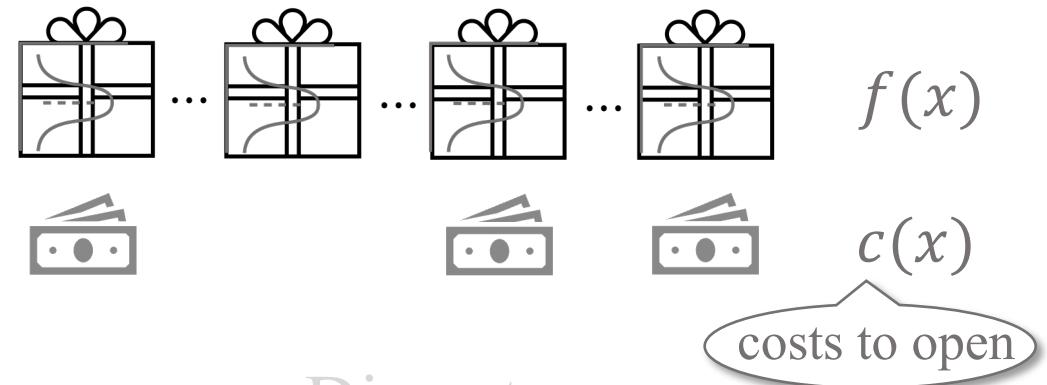
Continuous

Correlated

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

Pandora's Box

[Weitzman'79]



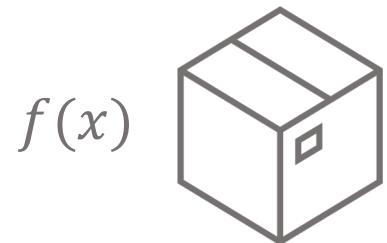
Discrete

Independent

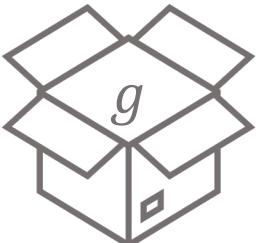
incorporate posterior
take continuum limit
 \Leftarrow Optimal policy: $\text{GI}_f(x; c(x))$

Intuition

Exploration



Exploitation



vs.

Open closed box

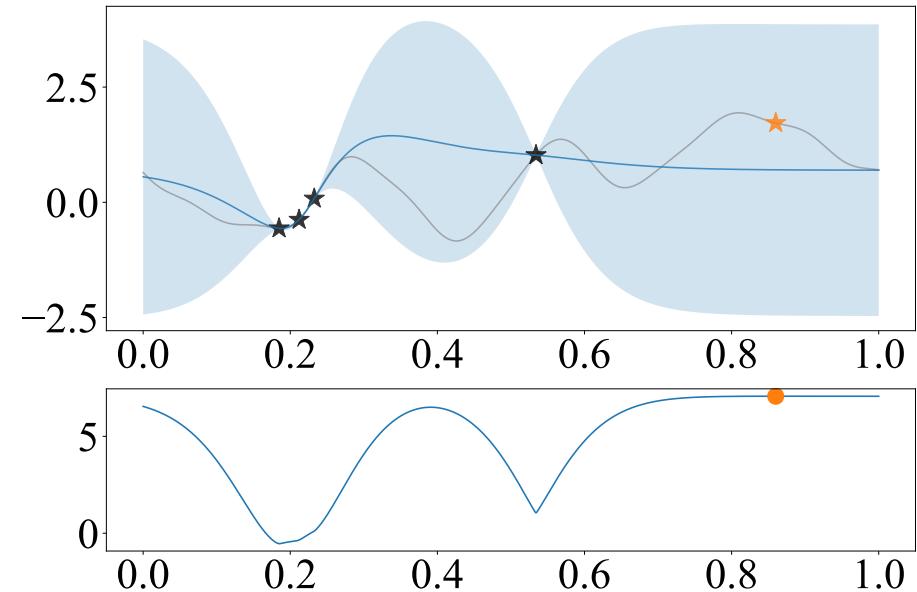
Take opened box

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

$$g$$

Should one open box? Depend on g !

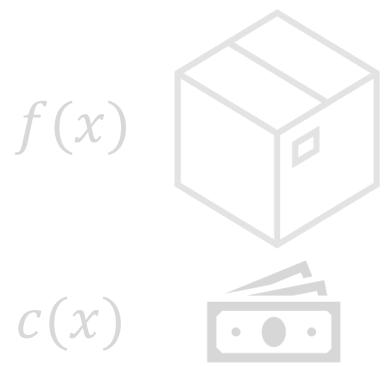
Gittins Index



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

Intuition

Exploration



vs.

Exploitation



Open closed box

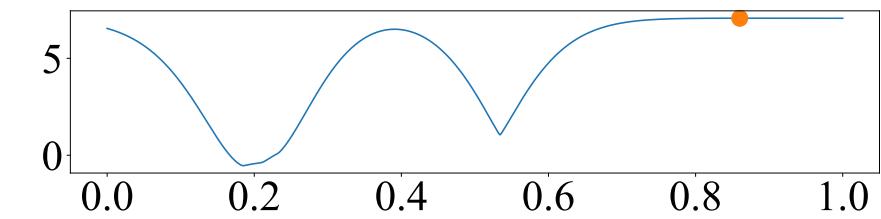
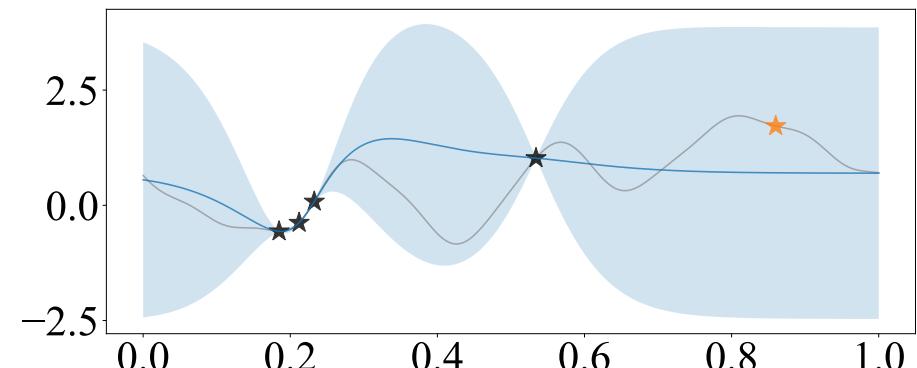
Take opened box

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

$$g$$

Should one open box? Depend on g !

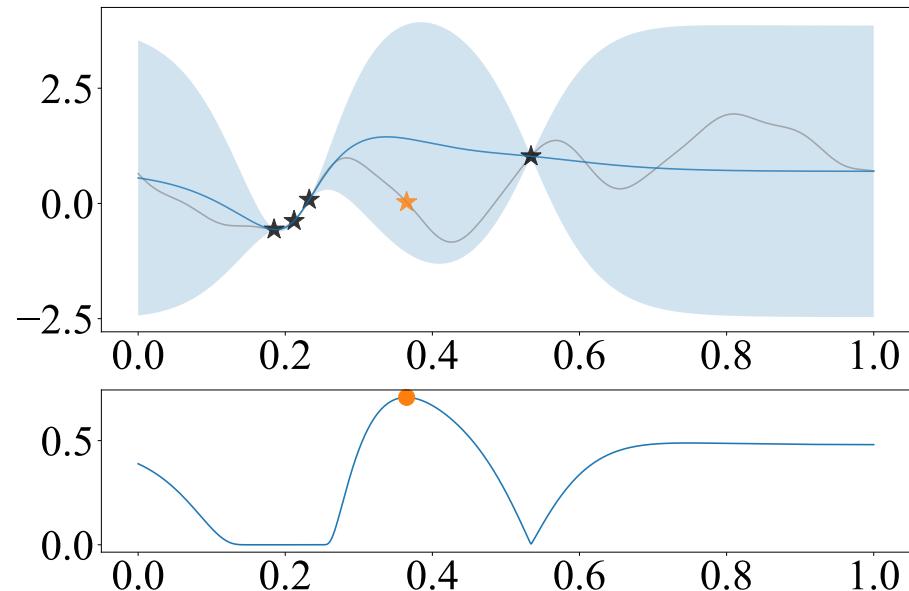
Gittins Index



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

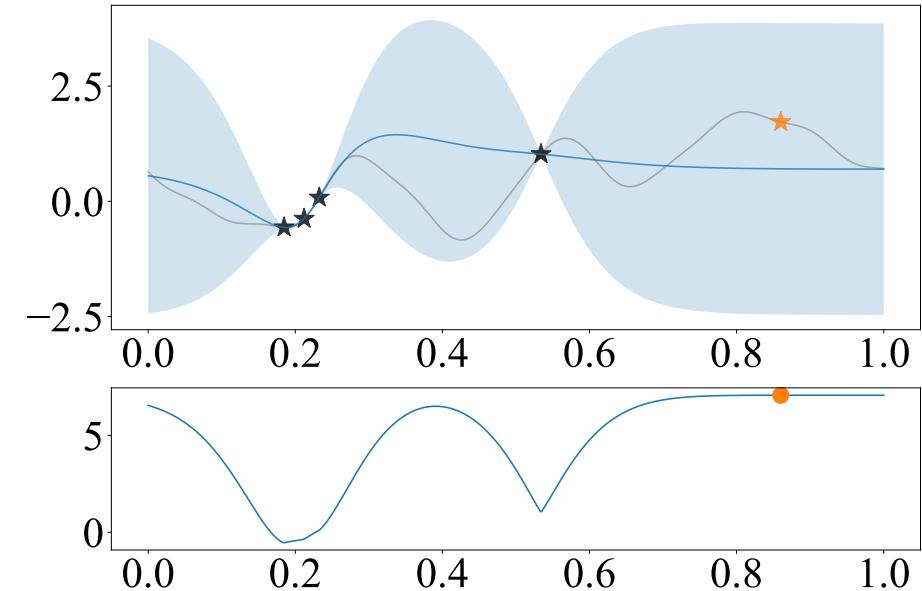
$$\text{EI}_{f|D}(x; g)$$

Expected Improvement

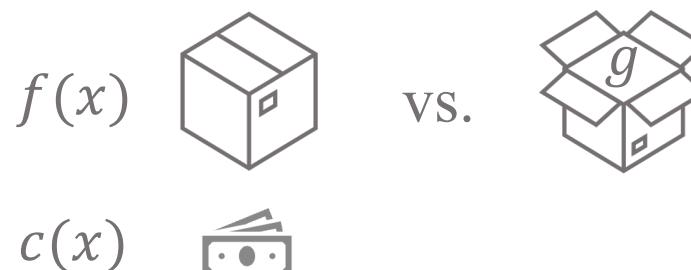


$$\text{EI}_{f|D}(x) := \mathbb{E}[\max(f(x) - y_{\text{best}}, 0) \mid D]$$

Gittins Index

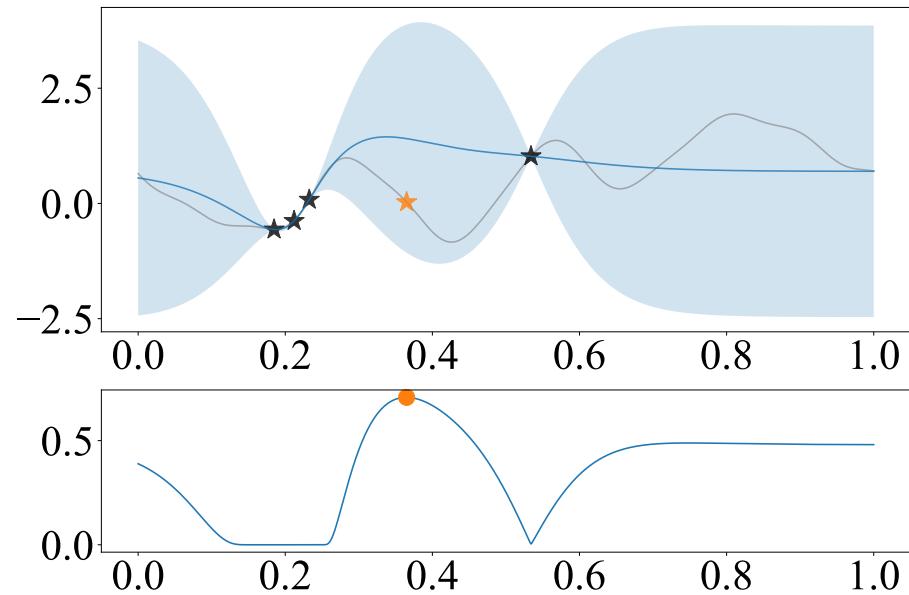


$$\begin{aligned}\text{GI}_{f|D}(x; c) &:= \text{solution } g \text{ s.t. } \text{EI}_{f|D}(x; g) = c(x) \\ \text{where } \text{EI}_{f|D}(x; g) &:= \mathbb{E}[\max(f(x) - g, 0) \mid D]\end{aligned}$$



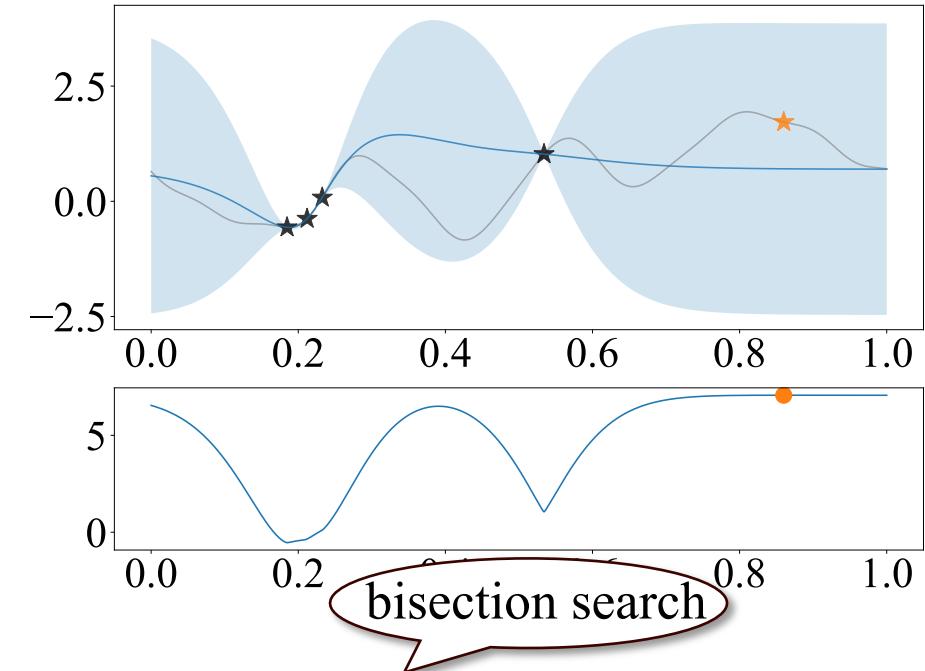
Exploration Exploitation

Expected Improvement



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

Gittins Index



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

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

analytical expression

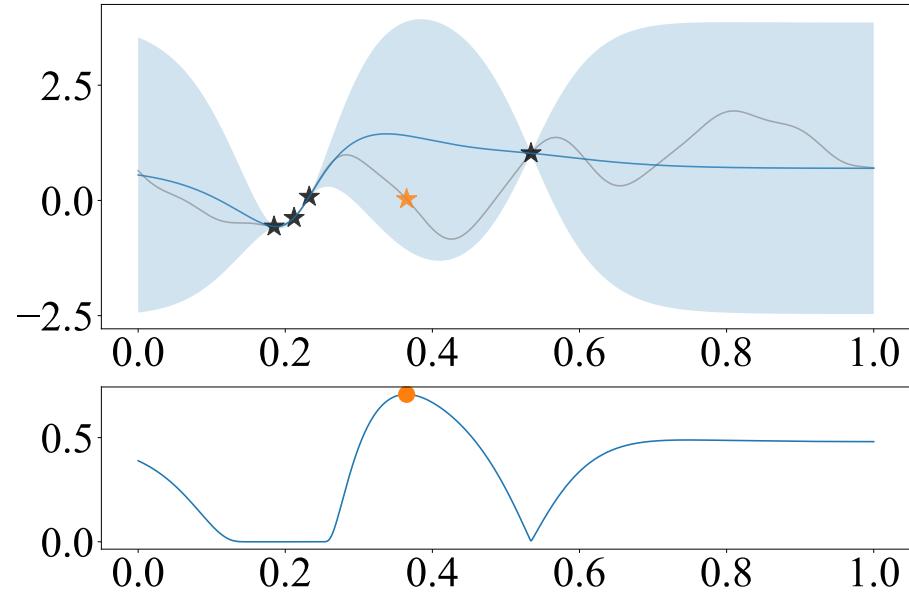


Easy-to-compute decision rule!



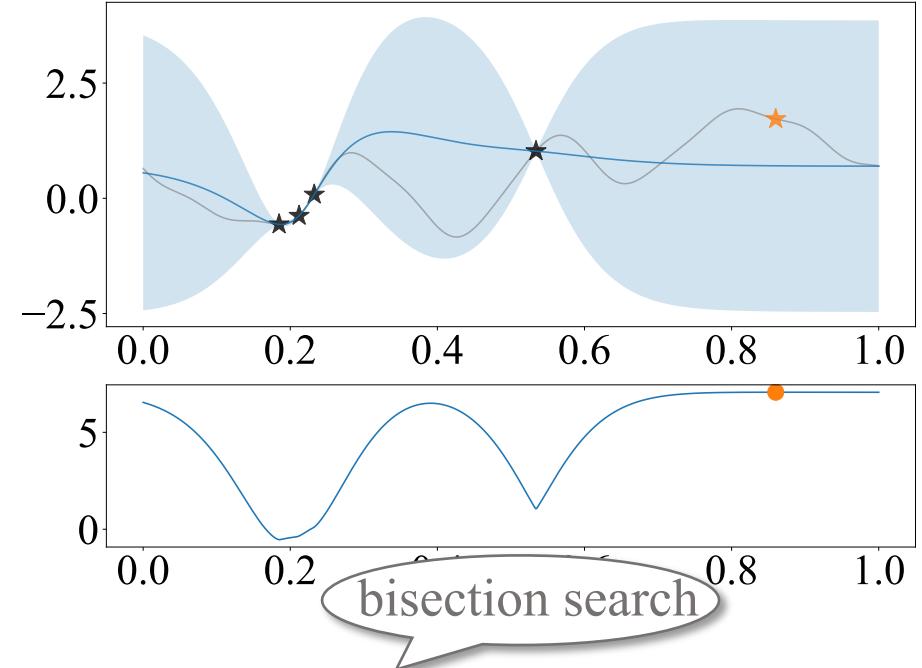
"Cost-aware Bayesian Optimization via the Pandora's Box Gittins Index." NeurIPS'24.

Expected Improvement



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

Gittins Index



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

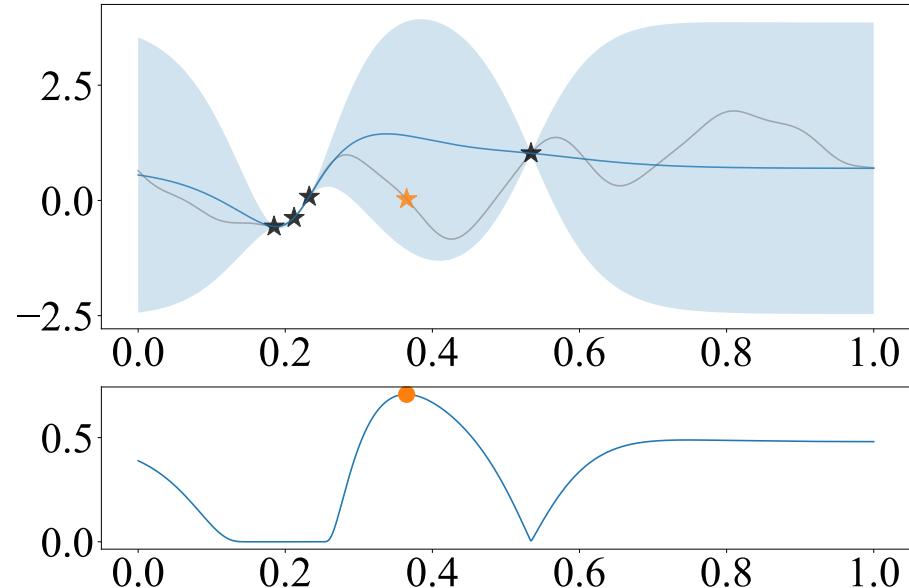
where $\text{EI}_{f|D}(x; g) := \mathbb{E}[\max(f(x) - g, 0) | D]$

Google DeepMind

FunBO: Discovering new acquisition functions for
Bayesian Optimization with FunSearch

hard to discover GI

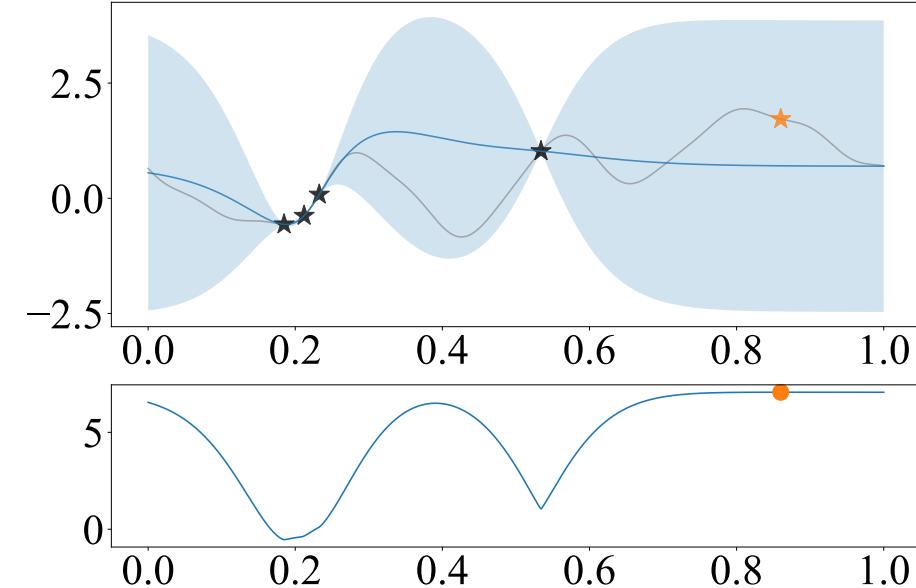
Expected Improvement



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

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

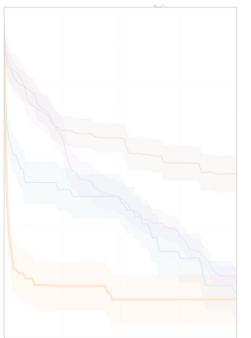
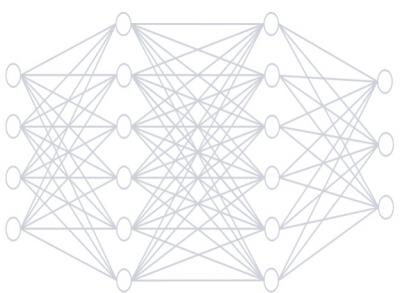
Our Contribution: Gittins Index Principle

Novel connection



Link to Pandora's Box problem
& Gittins index theory

Competitive empirical performance



Interests from practitioners (e.g., Meta)

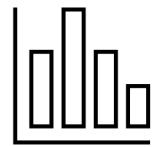
Principled decision rules

- Varying evaluation costs
- Adaptive stopping time

Unified framework for selection and stopping

- Future potential
- Best-prompt identification
- Adaptive response sampling
- Chain-of-thought selection
- Application to efficient LLM

Under-explored Factors for Better Decisions



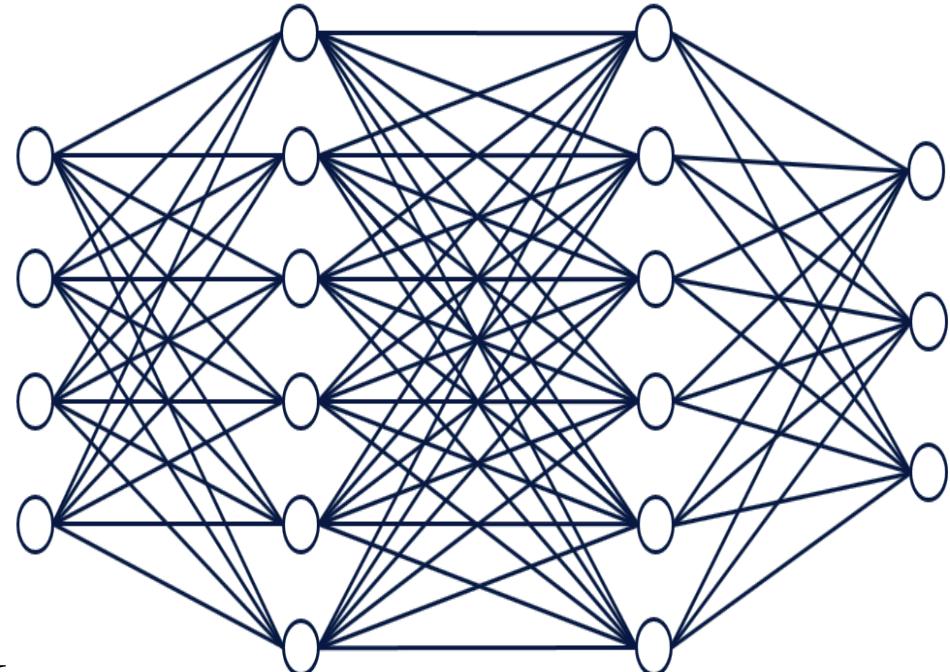
Varying evaluation costs



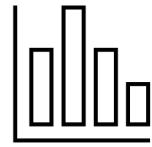
Smart stopping time



Observable multi-stage feedback



How does existing principle incorporate them?



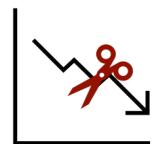
Varying evaluation costs

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

[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,

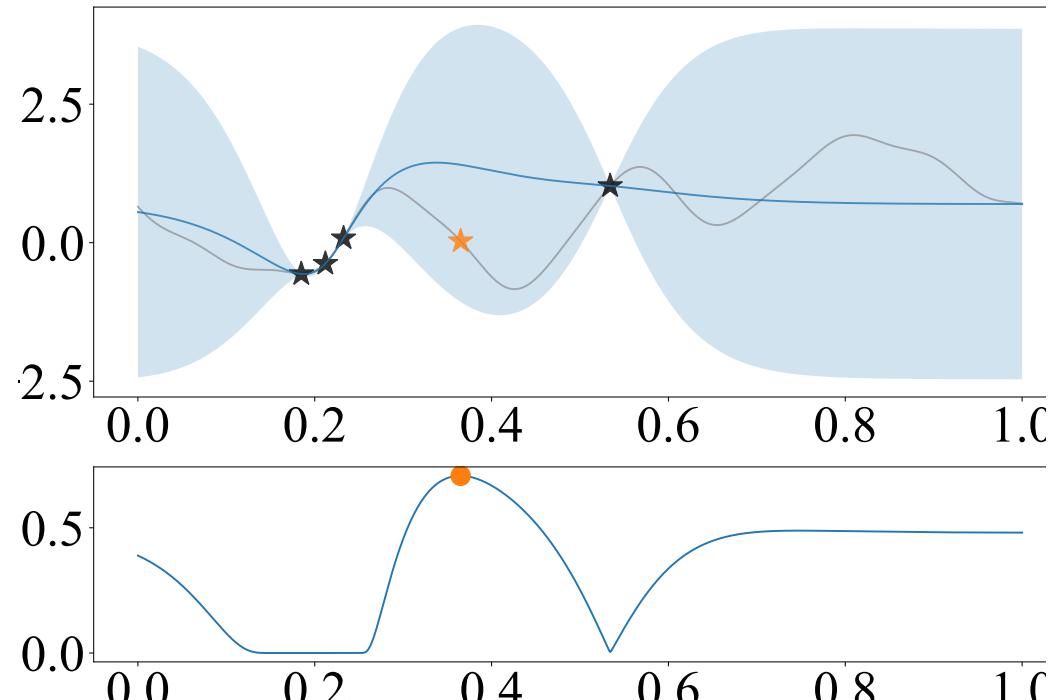
Ishibashi et al.'23]

Which threshold?



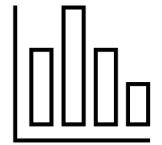
Observable multi-stage feedback

?



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

Under-explored Factors for Better Decisions



Varying evaluation costs



Smart stopping time



Observable multi-stage feedback

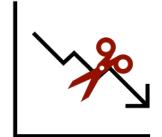


New design principle:
Gittins index

Why Gittins index?



Varying evaluation costs



Smart stopping time

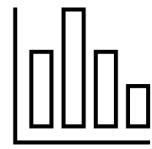


Observable multi-stage feedback

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

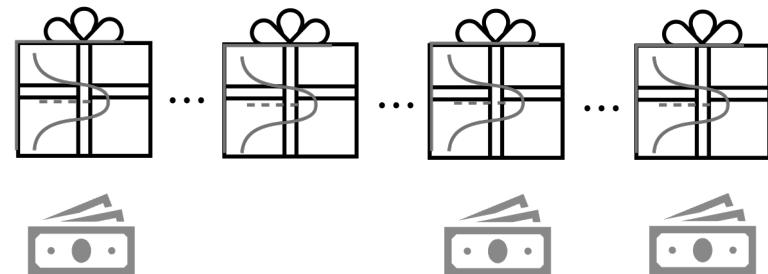
Features in **Pandora's box**



Observable multi-stage feedback

New design principle:
Gittins index

Optimal in related sequential
decision problems

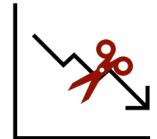


Why Gittins index?



Varying evaluation costs

Features in Pandora's box



Smart stopping time

Features in Pandora's box

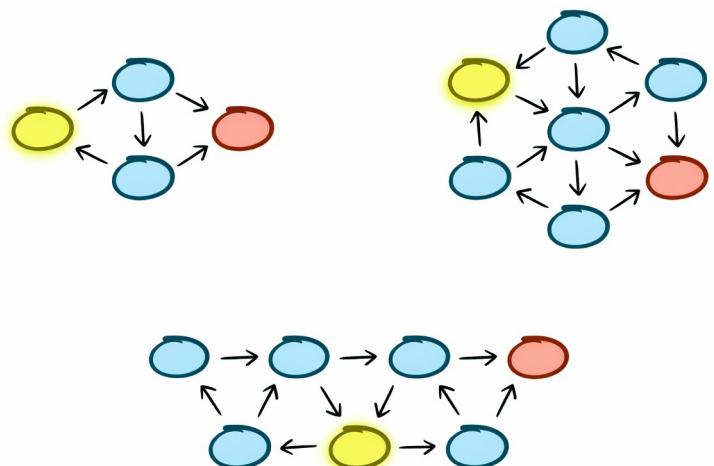


Observable multi-stage feedback

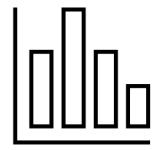
Features in **Markov chain selection**

New design principle:
Gittins index

Optimal in related sequential
decision problems

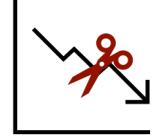


Why Gittins index?



Varying evaluation costs

Features in **Pandora's box**



Smart stopping time

Features in **Pandora's box**



Observable multi-stage feedback

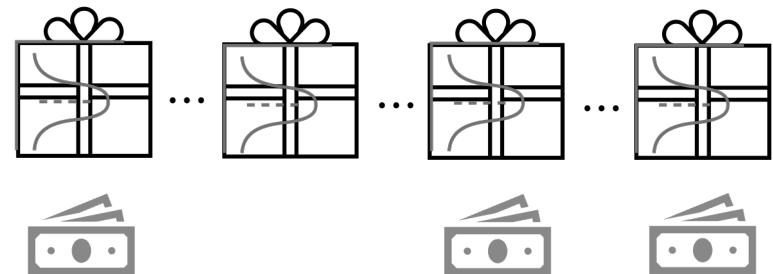
Features in Markov chain selection



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

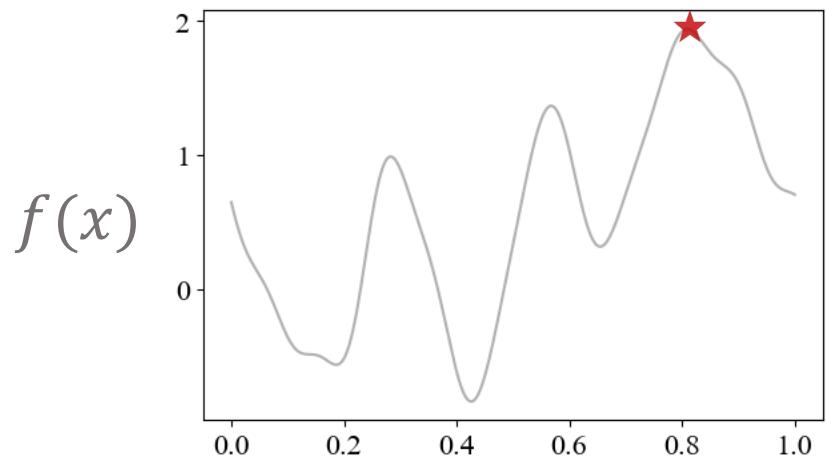
New design principle:
Gittins index

Optimal in related sequential
decision problems



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

Bayesian Optimization



Continuous

Correlated

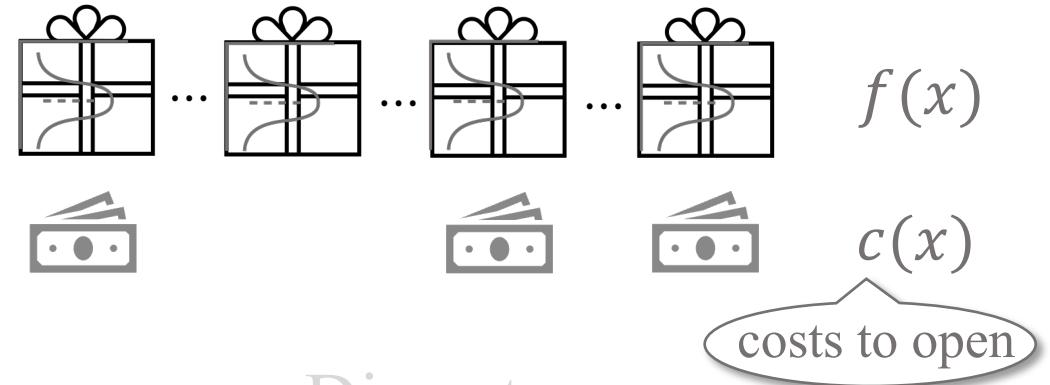
Cost-unaware

Fixed-iteration

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

Pandora's Box

[Weitzman'79]



Discrete

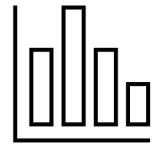
Independent

Cost-aware

Flexible-stopping

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

Expected Improvement vs Gittins Index



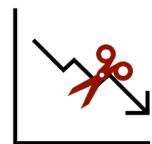
Varying evaluation costs

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

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.

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:

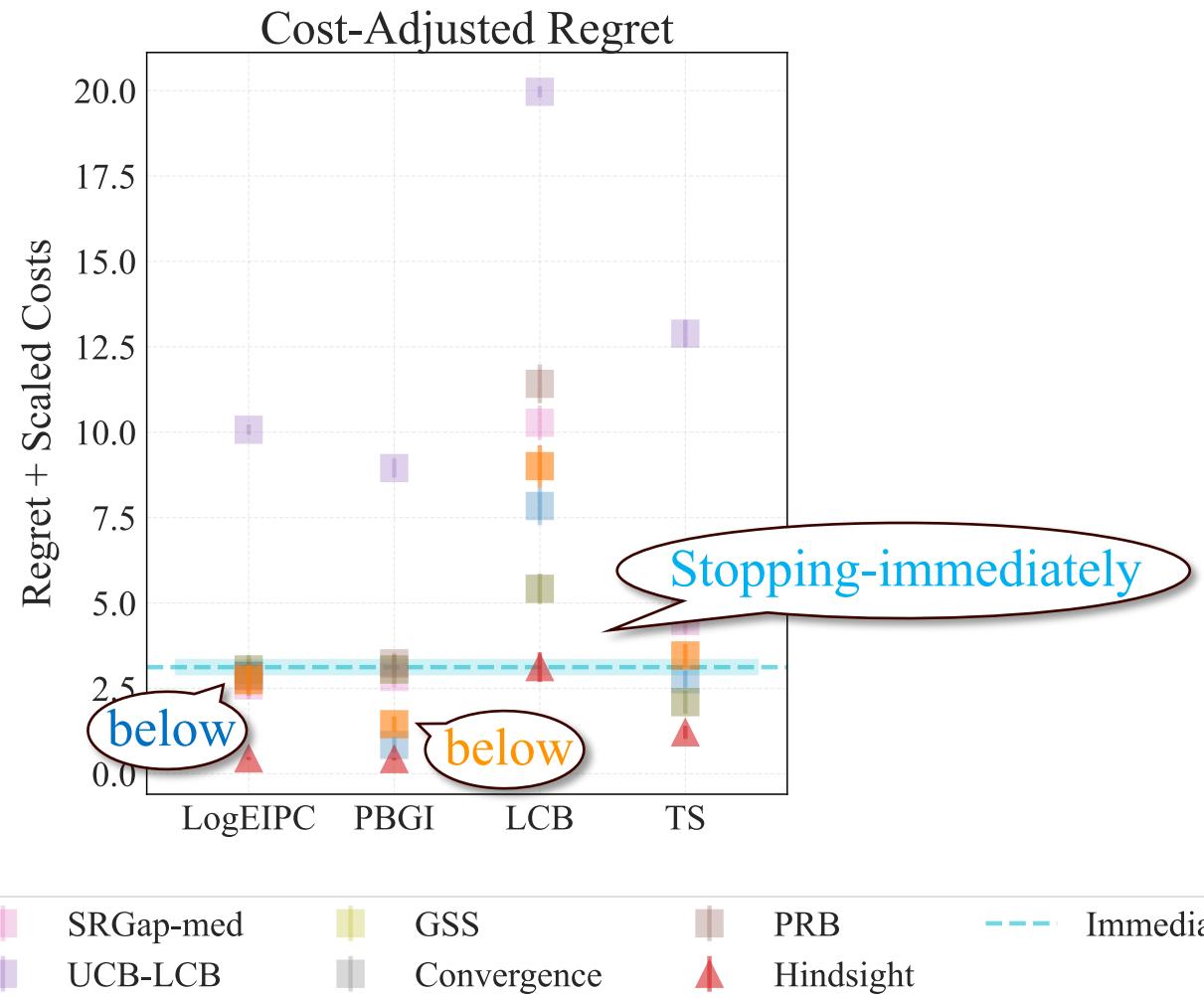
- Matches the **best achievable performance in the worst case** (evaluations are all very costly).
- Avoids **over-spending** — a property many cost-unaware stopping rules lack.

New

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.

Our Contribution: Gittins Index Principle

Novel connection



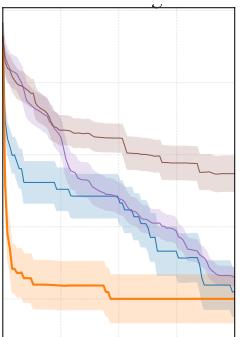
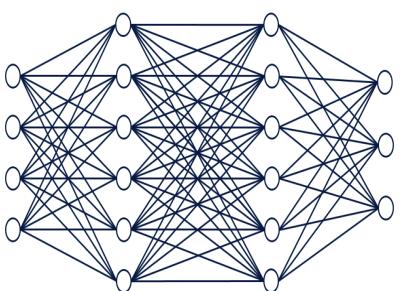
Link to Pandora's Box problem
& Gittins index theory

Principled decision rules



Unified framework for
selection and stopping

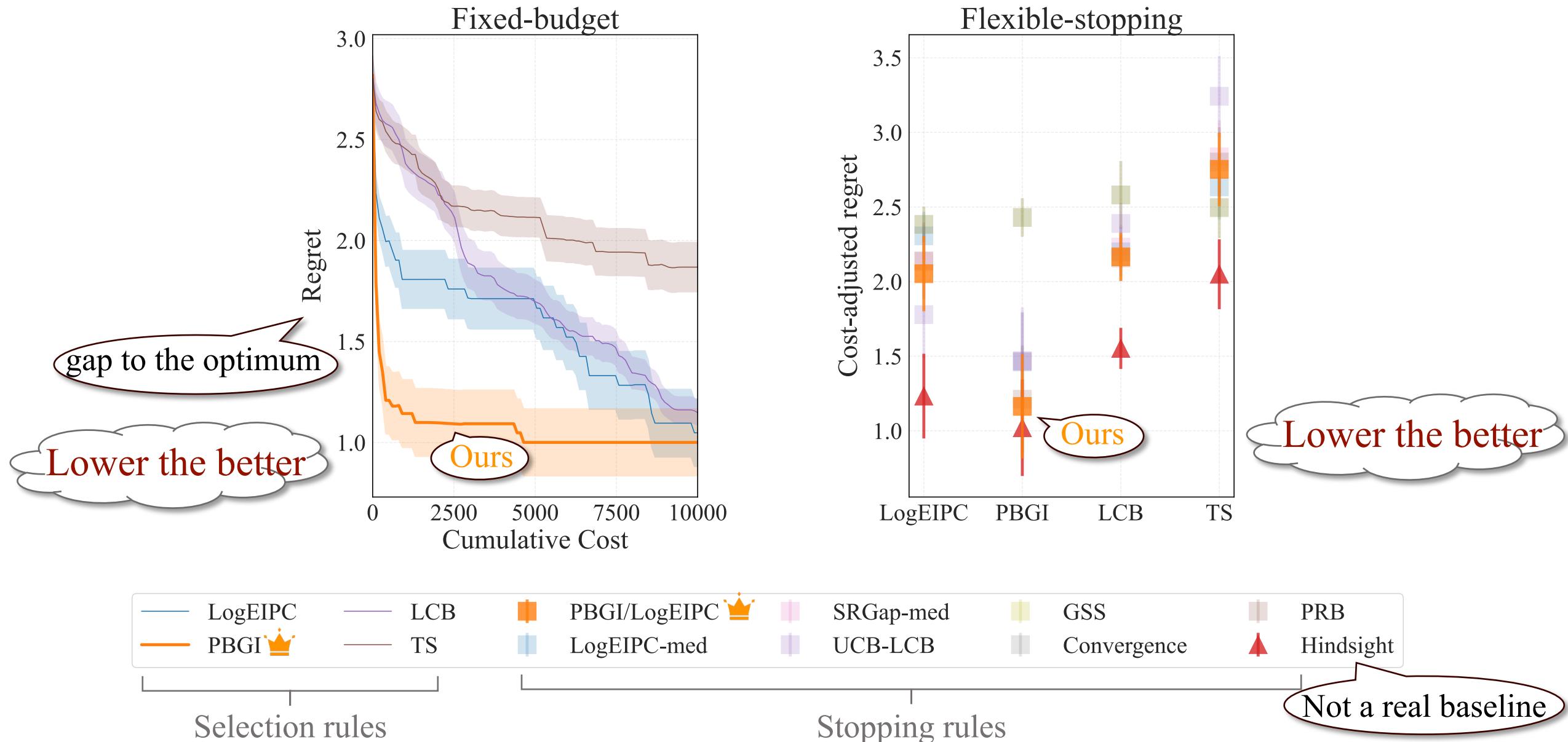
Competitive empirical performance



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



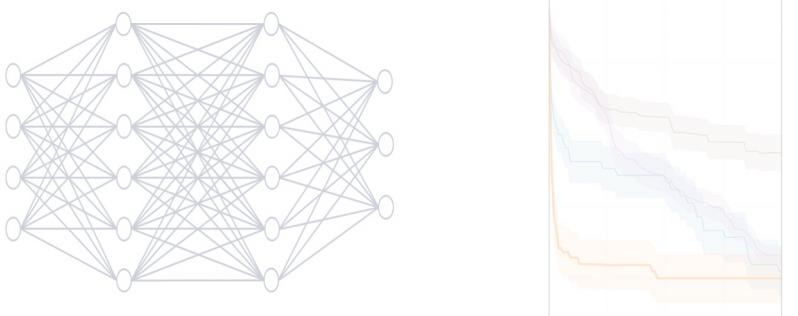
Our Contribution: Gittins Index Principle

Novel connection



Link to Pandora's Box problem
& Gittins index theory

Competitive empirical performance



Interests from practitioners (e.g., Meta)

Principled decision rules

- Varying evaluation costs
- Adaptive stopping time

Unified framework for cost-aware selection and stopping

Future potential

- Best-prompt identification
- Adaptive response sampling
- Chain-of-thought selection
- Application to efficient LLM

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.