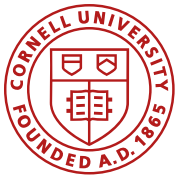
INFORMS Tutorial

Co-presented with Ziv Scully

The Gittins Index: A Design Principle for Decision-making Under Uncertainty

Part II of III



Cornell University

Alexander Terenin

https://avt.im/ · @avt_im

Announcement: I'm on the job market!

Research interests: decision-making under uncertainty

This tutorial: Gittins indices for Bayesian optimization

Talk for job market showcase: Bayesian algorithms for adversarial online learning

Both slides are available on my website!

The Gittins Index: A Design Principle for Decision-making Under Uncertainty

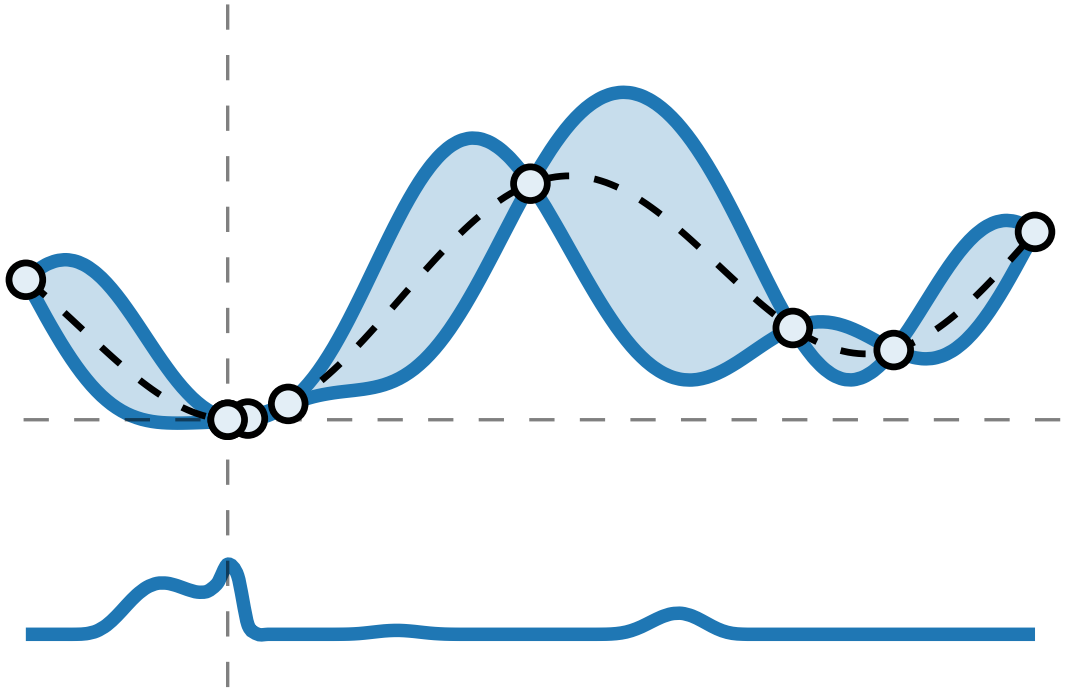
Part I: Introducing Gittins Indices via Pandora's Box

Part II: Gittins Indices for Bayesian Optimization

Part III: Tail Scheduling

Decision-making Under Uncertainty
Bayesian Optimization

Bayesian Optimization



Automatic explore-exploit tradeoff

Performance impact

Nando de Freitas 
@NandoDF

I agree with this thread of @avt_im.

Prior to the match with Lee Sedol, we tuned the latest AlphaGo agent with Bayesian Optimization and this improved its win-rate from 50% to 66.5% in self-play games. This tuned version was deployed in the final match. See arxiv.org/abs/1812.06855 for details.

At the time, we didn't publicise this as it was one of the secret ingredients, but we definitely benefited from being open minded, embracing many approaches, and ultimately testing inasmuch as possible.

The GPs of Bayesian Optimisation will likely be superseded, as in the works of @yutianc et al, but the ideas will continue being useful.

> 
> **Alexander Terenin** @avt_im · May 16
> Replying to @avt_im
> So let's learn from the scientific mistakes of the past, and not broadly dismiss *any* area of machine learning.
>
> Show more

4:30 PM · May 16, 2024 · **30.3K** Views

Bayesian Optimization

Goal: optimize unknown function f in as few evaluations as possible

1. Build posterior $f \mid y$ using data $(x_1, f(x_1)), .., (x_t, f(x_t))$
2. Choose

$$x_{t+1} = \arg\max_{x \in \mathcal{X}} \alpha_{f|y}(x)$$

using the acquisition function $\alpha_{f|y}$, built from the posterior

Useful property: separation of modeling from decision-making

Motivating application: hyperparameter tuning

Algorithms have hyperparameters! Neural network training:

- x: number of layers, layer width, learning rate, ..
- $f(x)$: test accuracy of trained model
- $c(x)$: total training compute

Goal: maximize test accuracy under expected compute budget

$$\max_{t=1,..,T} f(x_t) \quad \text{subject to} \quad \mathbb{E} \sum_{t=1}^{T} c(x_t) \leq B$$

Expected improvement per unit cost

Cost-aware baseline: expected improvement per unit cost

$$\alpha_t^{\mathrm{EIPC}}(x) = \frac{\mathrm{EI}_{f|y_1,..,y_t}(x; \max_{1\le\tau\le t} y_\tau)}{c(x)} \qquad \mathrm{EI}_\psi(x; y) = \mathbb{E}\max(0, \psi(x) - y)$$

Often strong in practice, but can perform arbitrarily-badly

- Issue: high-cost high-variance points (Astudillo et al., 2021)

Derivation: one-step approximation to intractable dynamic program

What if I told you this dynamic program can sometimes be solved exactly?

Expected improvement: time-based simplification
Gittins index: *space-based* simplification

Cost-aware Bayesian Optimization: a simplified setting

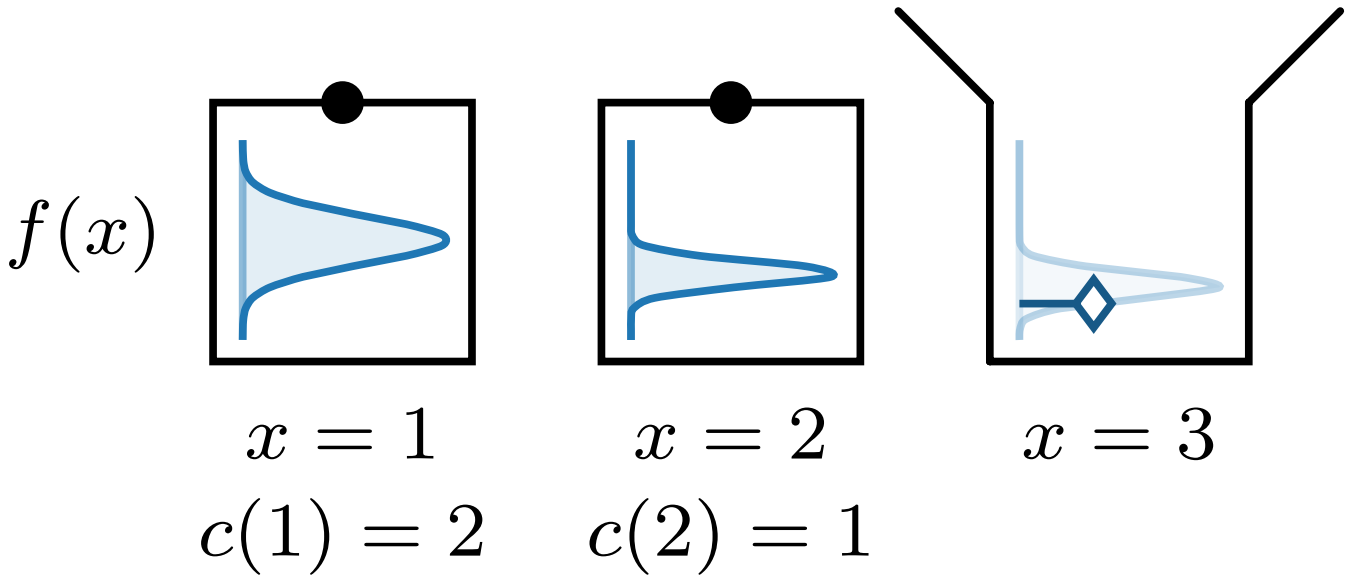
Assumptions:

- Cost-per-sample problem: algorithm decides when to stop
- Reward once stopped: best observed point (simple regret)
- Distribution over objective functions is known
- X is discrete, $f(x_i)$ and $f(x_j)$ for $x_i \neq x_j$ are independent

These are restrictive! But they lead to an interesting, general solution

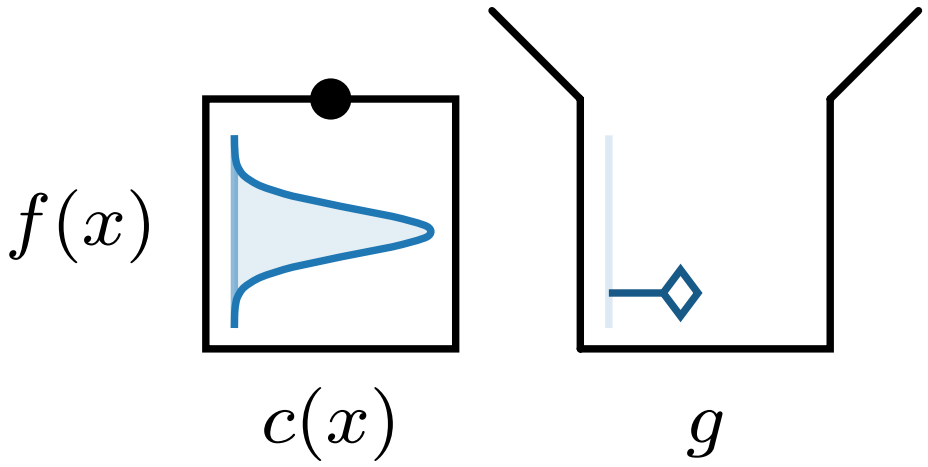
This is just Pandora's Box!

Whether to open Pandora's Box?



Solving Pandora's Box

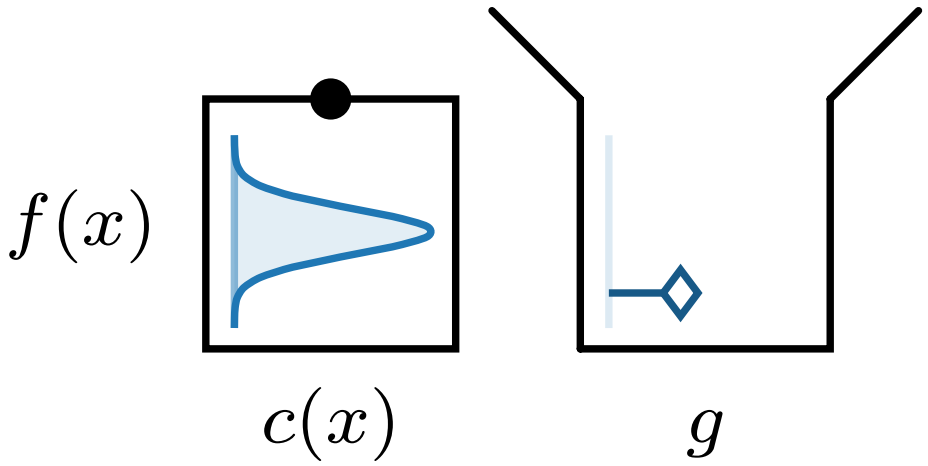
Consider: one closed vs. one open box



Should we open the closed box? *Maybe!*

Depends on costs c, reward distribution f, and value of open box g

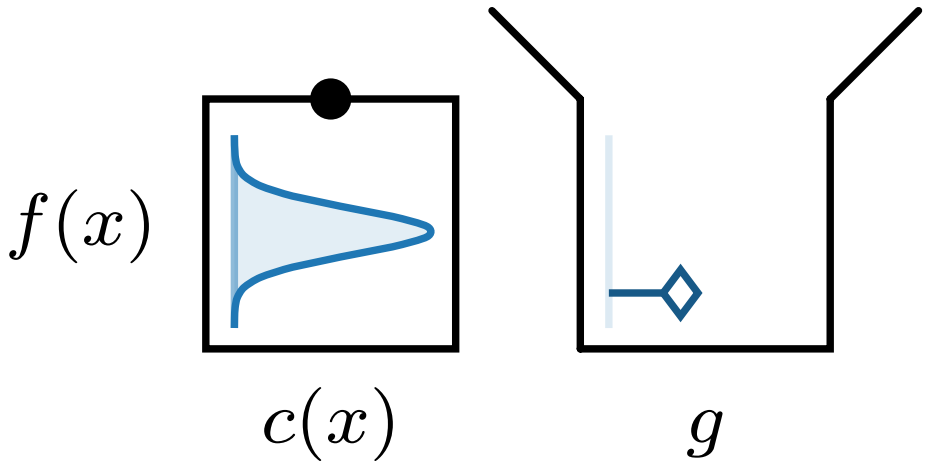
Consider: one closed vs. one open box



One closed vs. open box: Markov decision process

Optimal policy: open if $\mathrm{EI}_f(x; g) > c(x)$

Consider: one closed vs. one open box

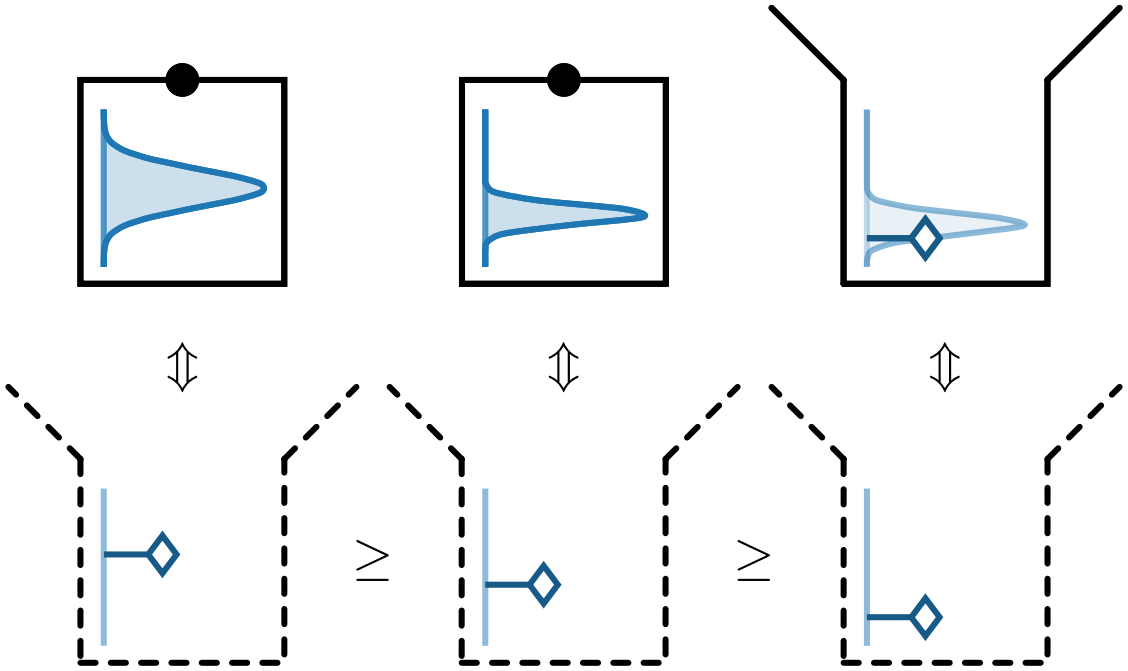


Consider how optimal policy changes as a function of g

If both opening and not opening are co-optimal: g is a *fair value*

Define: $\alpha_t^\star(x) = g$ where g solves $\mathrm{EI}_f(x; g) = c(x)$

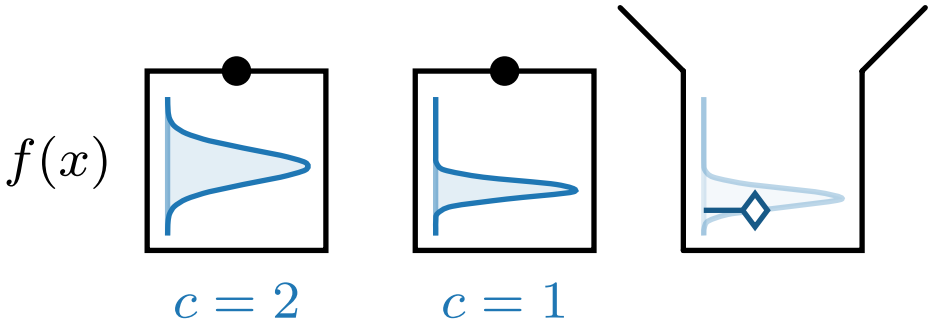
Back to many boxes

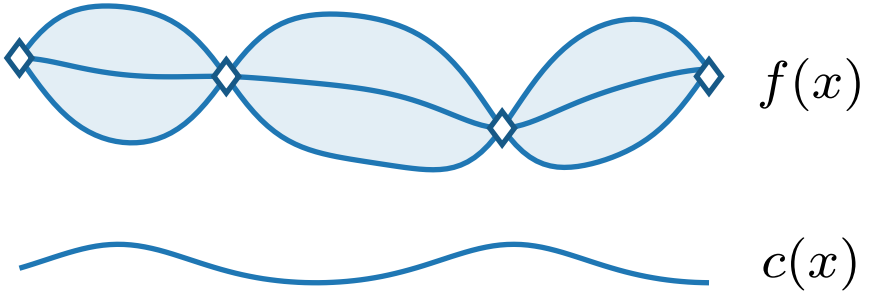


Theorem (Weitzman, 1979). This policy is optimal in expectation.

Caveat! Optimality theorems are fragile. *Definitions are not!*

Cost-aware Bayesian Optimization: Correlated Pandora's Box?





Difference so far: expected budget constraint

Expected Budget-constrained vs. Cost-per-sample

Gittins index $\alpha^\star$: optimal for cost-per-sample problem

- What about expected budget-constrained problem?

Theorem (Aminian et al., 2024; Xie et al., 2024). Assume the expected budget constraint is feasible and active. Then there exists a $\lambda > 0$ and a tie-breaking rule such that the policy defined by maximizing the Gittins index acquisition function $\alpha^\star(\cdot)$, defined using costs $\lambda c(x)$, is Bayesian-optimal.

Proof idea: Lagrangian duality

Pandora's Box Gittins Index for Bayesian Optimization

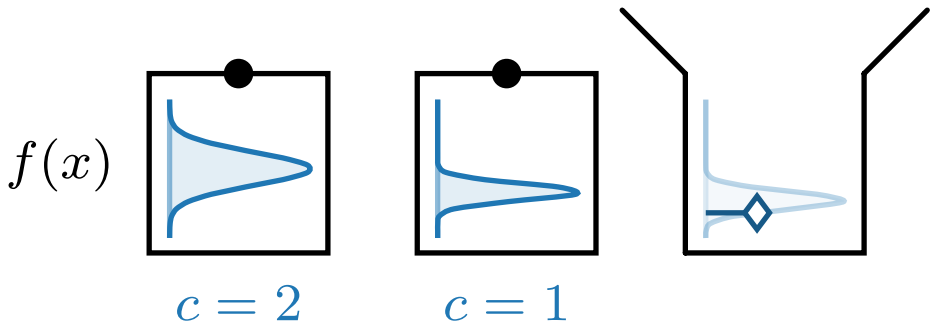
Bayesian optimization: posterior distribution is correlated

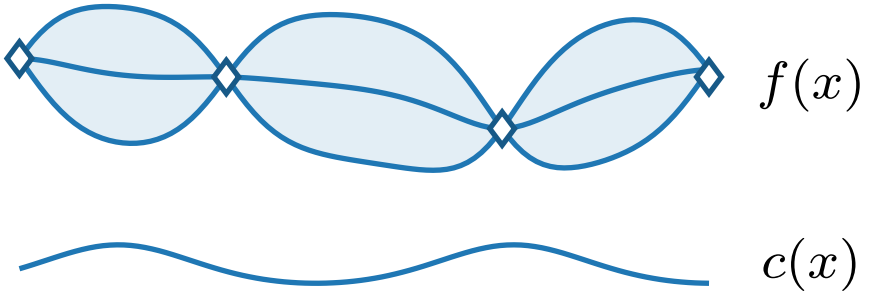
Define *Pandora's Box Gittins Index* acquisition function:

$$\alpha^{\text{PBGI}}(x) = g \quad \text{where } g \text{ solves} \quad \text{EI}_{f|y}(x; g) = \underbrace{\lambda} c(x)$$

Lagrange multiplier from budget constraint

Correlations: incorporated into acquisition function via the posterior

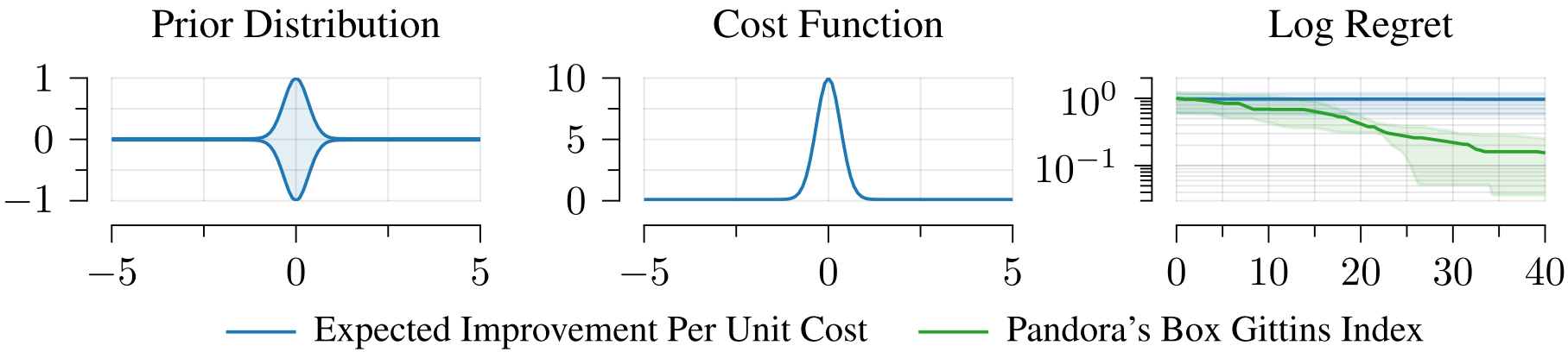




Does it work?

Not an optimal policy. But, recall the caveat: *maybe a strong one?*

An initial sanity-check

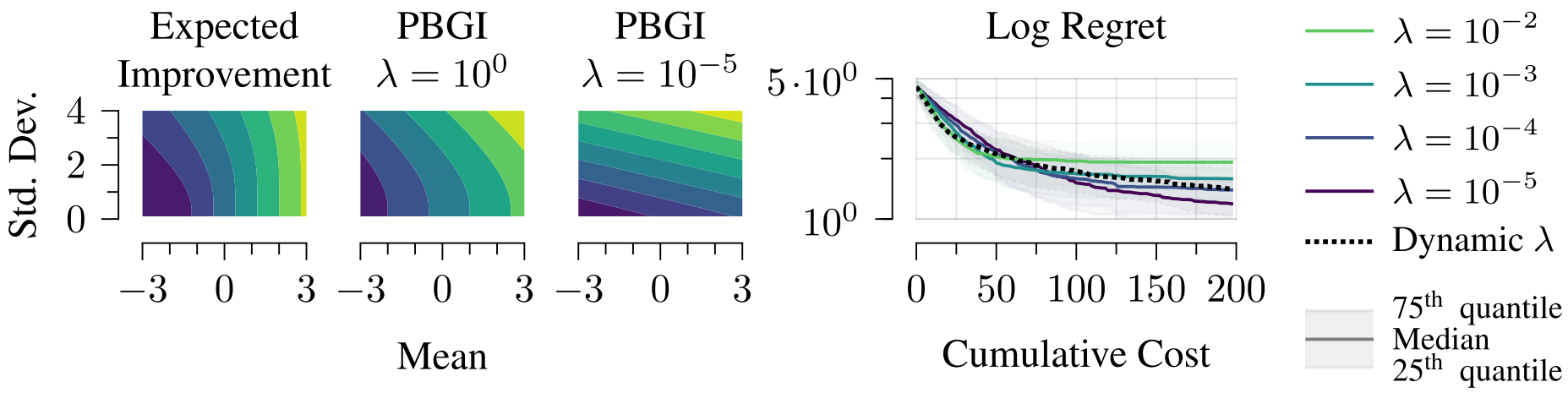


Performance counterexample for EIPC baseline:

- Random objective with high-variance high-cost region

Pandora's Box Gittins Index: performs well

Setting the hyperparameters: what does λ do?

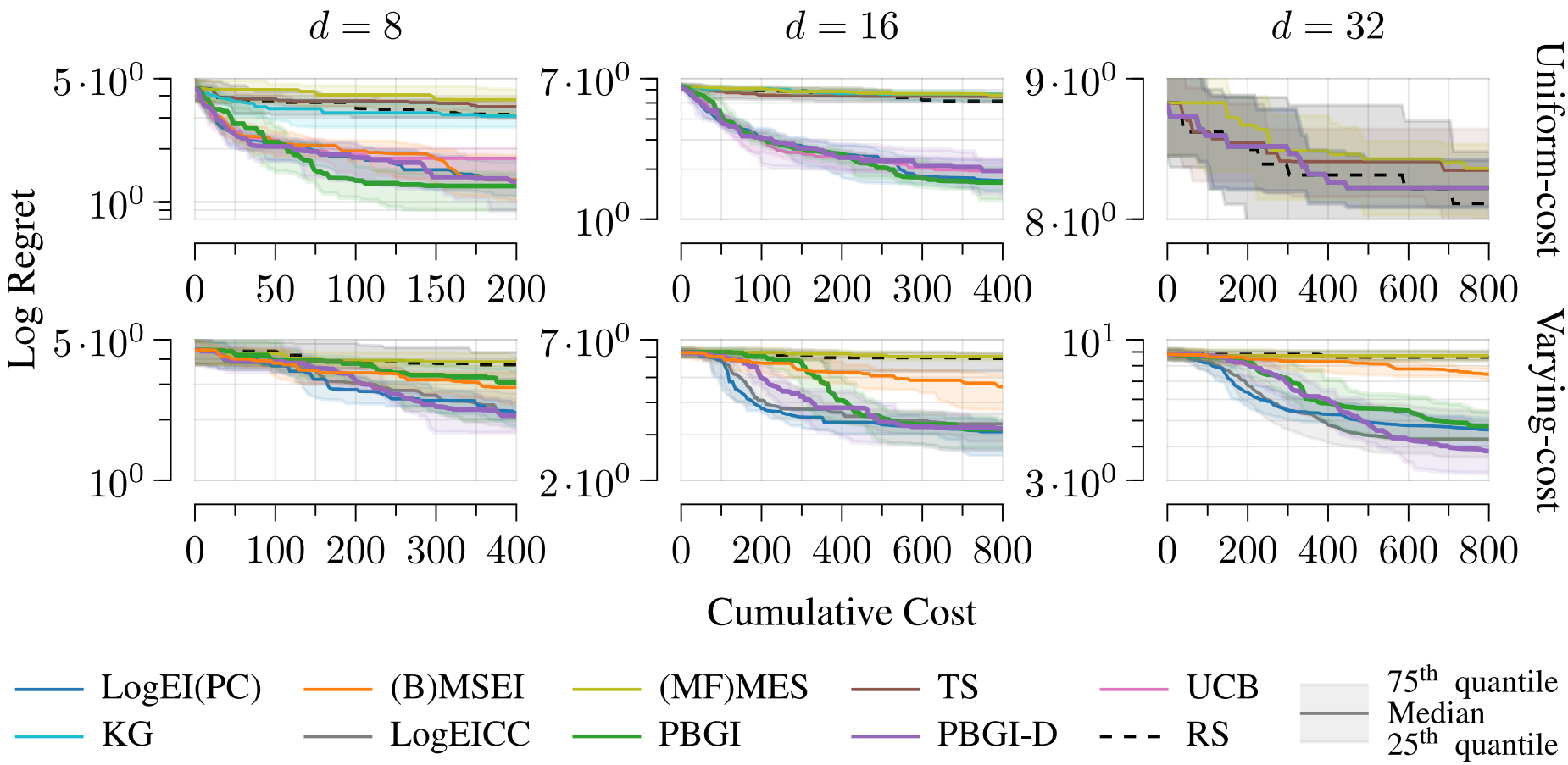


Controls risk-averse vs. risk-seeking behavior

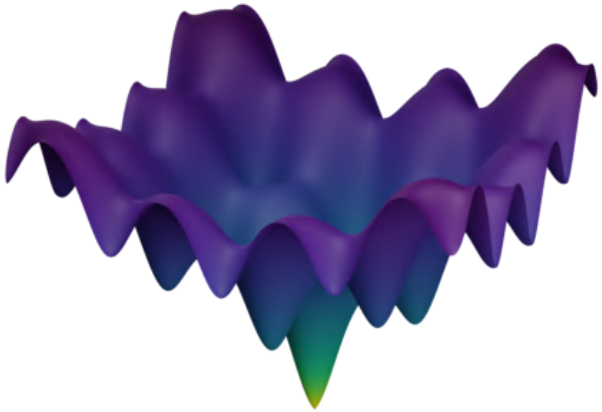
Optimal tuning of λ: depends on the expected budget

Limit as $\lambda \to 0$: converges to UCB with automatic learning rate

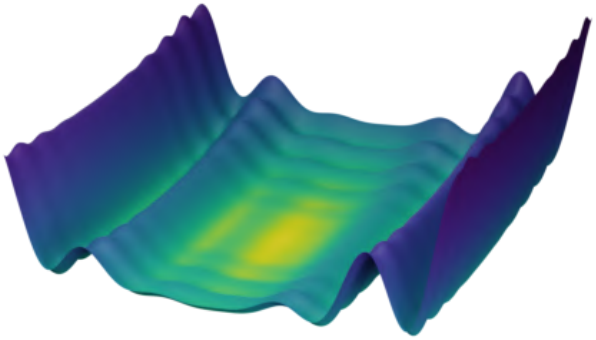
Objective functions: sampled from the prior



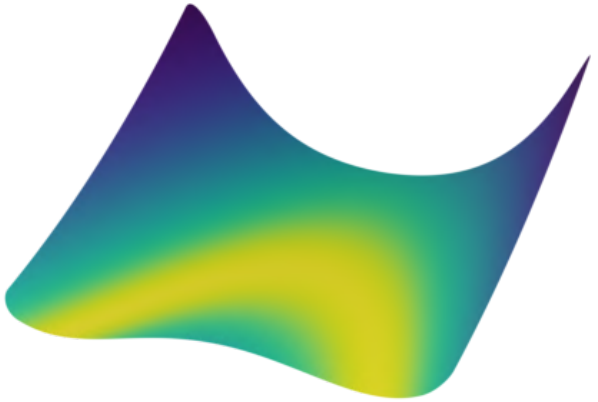
Synthetic benchmark functions



Ackley function

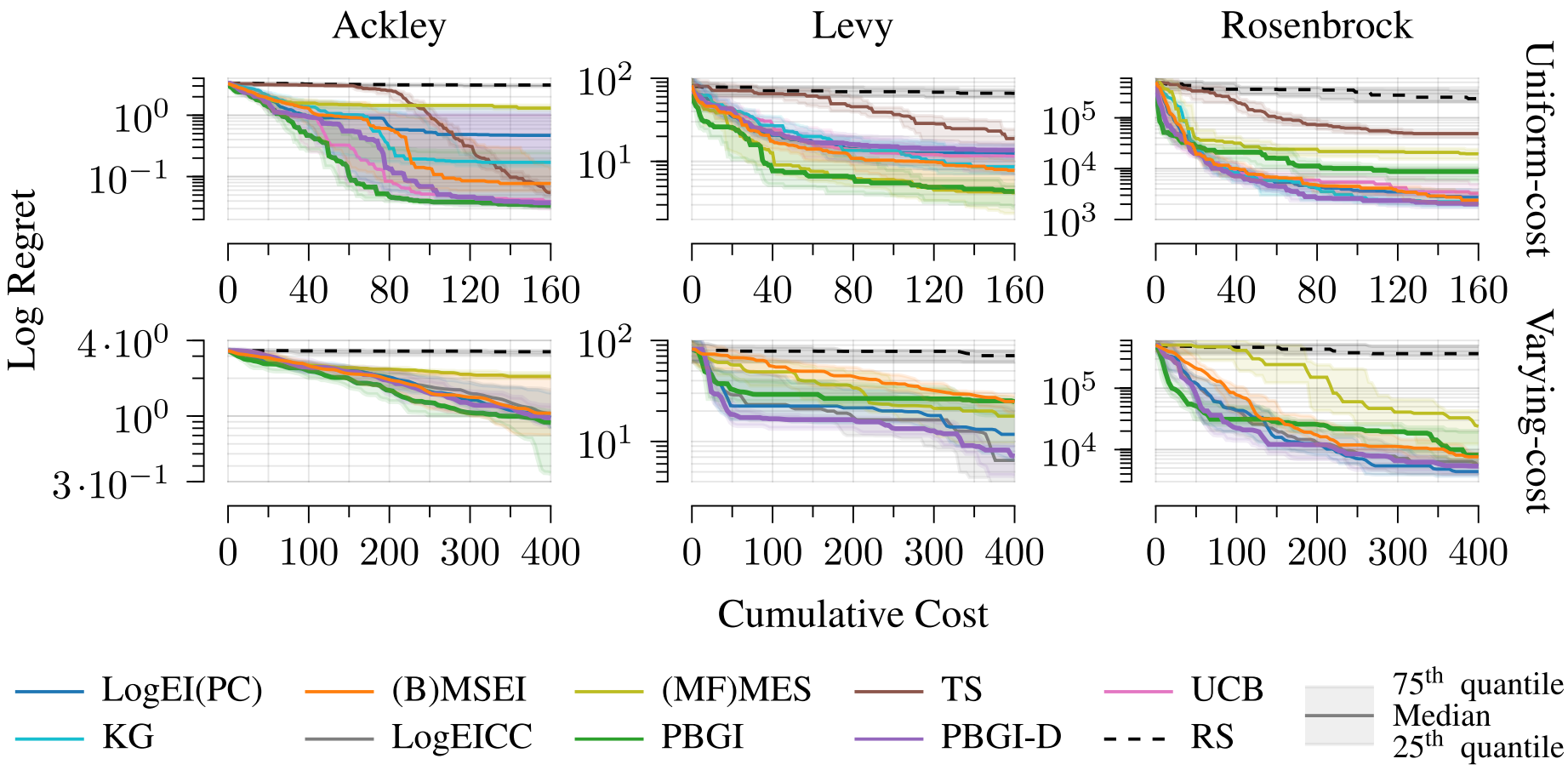


Levy function

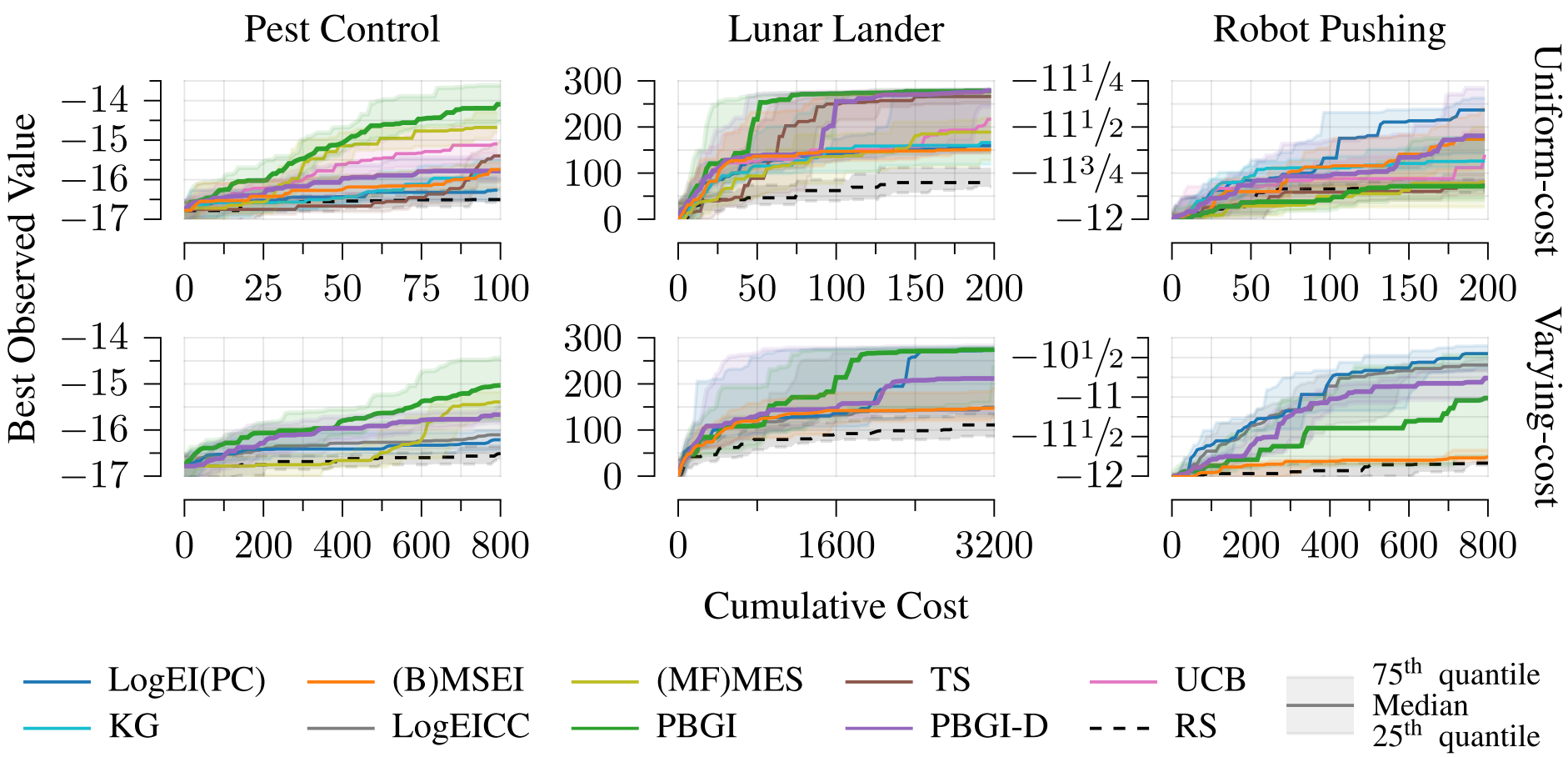


Rosenbrock function

Synthetic benchmark functions



Empirical objectives



Conclusion: Gittins indices can perform well,
even where they are not optimal

Back to our motivating application

Neural network training: proceeds over time

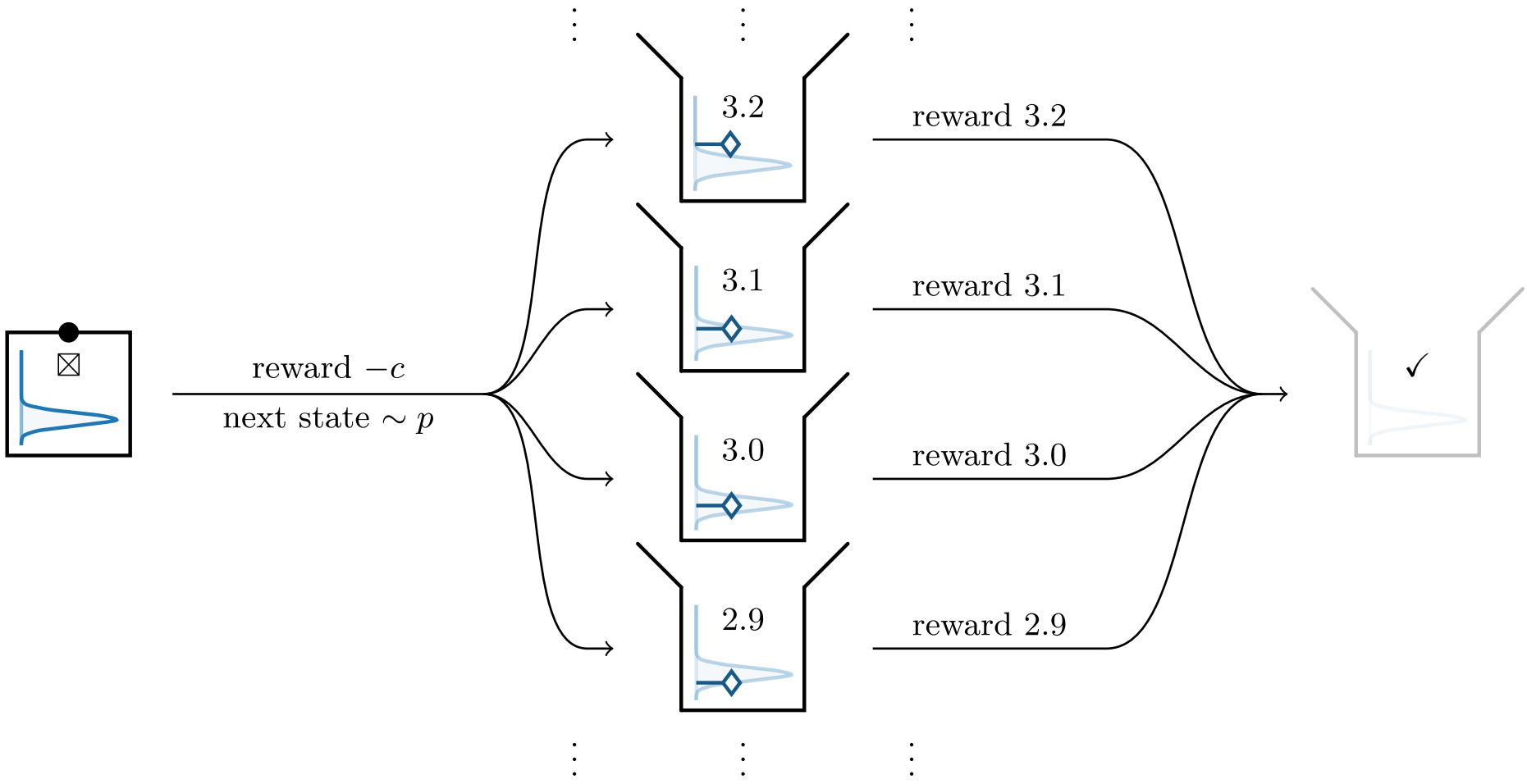
- Can sometimes predict final test loss from early iterations
- Why finish a training run we know will turn out bad?

New goal: maximize test accuracy, *allowing for early stopping*

$$\max_{k=1,..,K} f(x_{k,T_k}) \quad \text{subject to} \quad \mathbb{E}\sum_{k=1}^{K}\sum_{t=1}^{T_k} c(x_{t,k}) \leq B$$

Cost-aware *freeze-thaw* Bayesian optimization: this setting is *novel*!

Pandora's Box: what's this an example of?



Transient Markov chain: closed box $\rightsquigarrow$ open box $\rightsquigarrow$ selected box

Pandora's Box as a Markov chain (with rewards)

State space: $S = \{\boxtimes\} \cup \mathbb{R} \cup \{\checkmark\}$, with terminal states $\partial S = \{\checkmark\}$

- Closed box: $\boxtimes$
- Open box: $v_i \in \mathbb{R}$
- Selected box: $\checkmark$

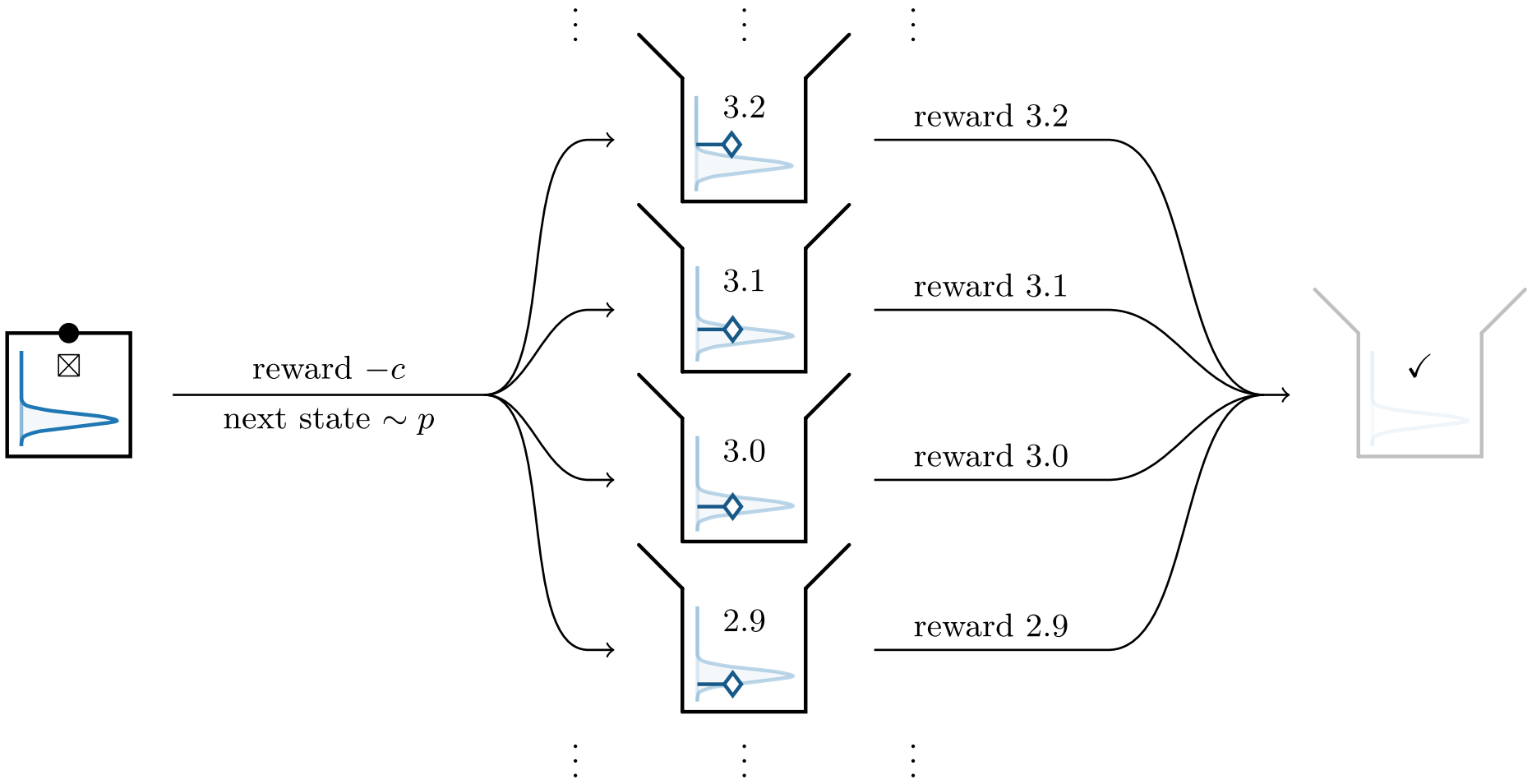
Transition kernel: $p : S \to \mathcal{P}(S)$

- Jump from $\boxtimes$ to $v_i \in \mathbb{R}$ according to the box's reward distribution
- Jump from $v_i \in \mathbb{R}$ to absorbing terminal state $\checkmark$ deterministically

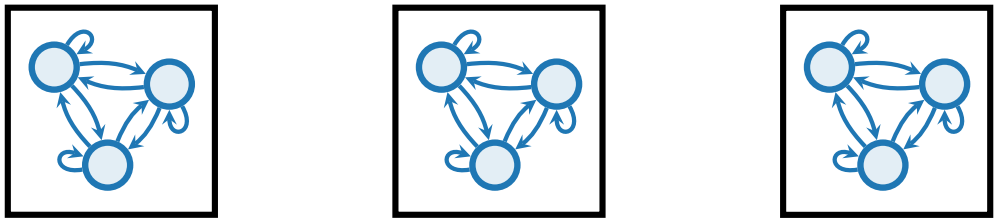
Reward function: $r : S \setminus \partial S \to \mathbb{R}$

- $r(\boxtimes) = -c$
- $r(v_i) = v_i$

Pandora's Box as a transient Markov chain



From multiple Pandora's Boxes to multiple Markov chains



Markov Chain Selection

Definition. Given mutually independent Markov chains $(S_i, \partial S_i, p_i, r_i)$ for $i = 1, .., n$, define a Markov decision process:

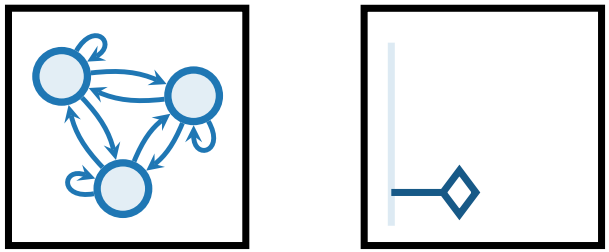
1. State space: $S_{\text{MCS}} = \{(s_1, .., s_n) : \forall i, s_i \in S_i\}$.
2. Terminal states: $\partial S_{\text{MCS}} = \{(s_1, .., s_n) \in S : \exists i, s_i \in \partial S_i\}$, which *can be empty.*
3. Action space: $A_{\text{MCS}} = \{1, .., n\}$.
4. Reward function: $r_{\text{MCS}}(s, a) = r_a(s_a)$.
5. Transition kernel: given $(s_1, .., s_n)$ and action a, replace s_a with $s'_a \sim p(\cdot \mid s_a)$.
6. Discount factor: $\gamma \in (0, 1]$, i.e. allowed but *not required*.

Unifies Pandora's Box with discounted bandits, various queues, ...

Discounted case: reduces to *undiscounted case*

The Local MDP

Pandora's Box: solved by considering one closed and one open box

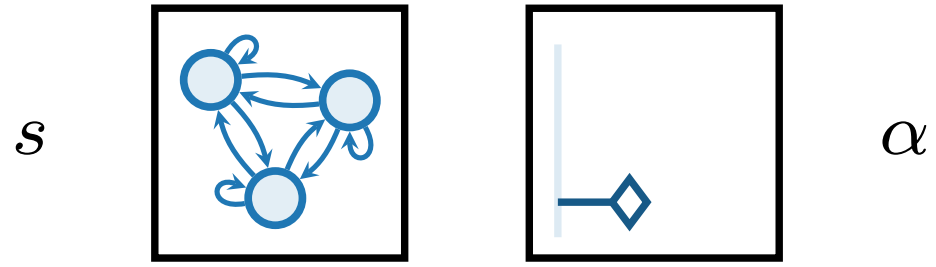


Let's generalize this!

The Local MDP

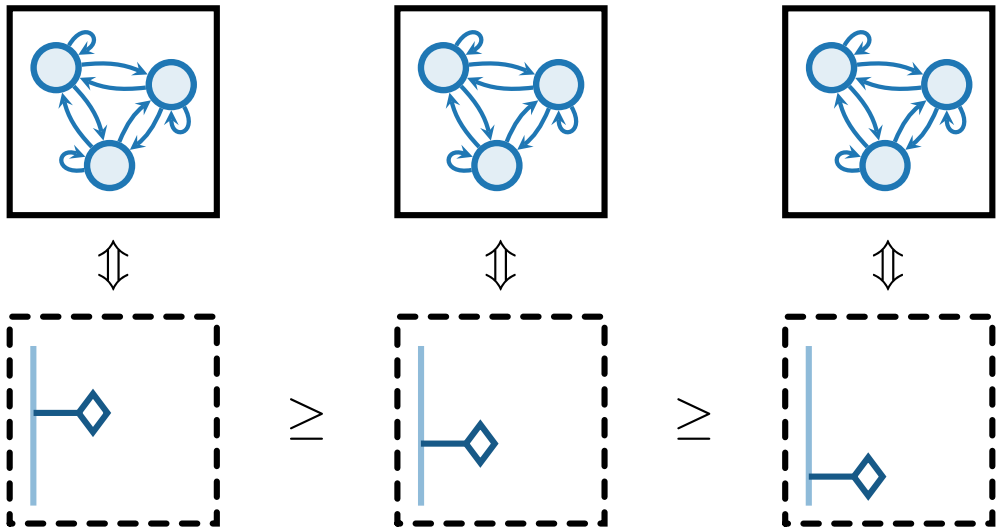
Definition. Given a Markov chain $(S, \partial S, r, p)$, an *alternative option* $\alpha \in \mathbb{R}$, and an initial state $s \in S$, define the (s, α)-*local MDP:*

1. State space: $S_{\text{loc}} = S \cup \{\checkmark\}$.
2. Terminal states: $\partial S_{\text{loc}} = \{\checkmark\}$.
3. Action spaces: $A_{\text{loc}} = \{\square, \triangleright\}$, called *stop* and *go*.
4. Reward function: $r_{\text{loc}}(s, \square) = \alpha$, $r_{\text{loc}}(s, \triangleright) = r(s)$, and $r_{\text{loc}}(\checkmark, \cdot) = 0$.
5. Transition kernel: if $a = \triangleright$, then let $s' \sim p(\cdot \mid s)$, otherwise $a = \square$ so let $s' = \checkmark$.



The Gittins Index of a Local MDP

Definition. Given a Markov chain $(S, \partial S, r, p)$, define its *Gittins index* $G : S \to \mathbb{R} \cup \{\infty\}$ to be the unique number g for which $\triangleright$ and $\square$ are co-optimal actions in the (α, s)-local MDP (or ∞ if no such g exists).



Theorem. In MCS, choosing the Markov chain of maximal Gittins index is optimal.

Caveat! *Optimality is fragile.* Definitions are not!

Gittins indices for advanced variants of Bayesian optimization

Example goal: maximize test accuracy, *allowing for early stopping*

$$\max_{k=1,..,K} f(x_{k,T_k}) \quad \text{subject to} \quad \mathbb{E}\sum_{k=1}^{K}\sum_{t=1}^{T_k} c(x_{t,k}) \leq B$$

Admits a well-defined Gittins index:

- Discrete simplifying case: special case of MCS – Gittins is optimal!
- Gaussian process: correlations – not optimal. *But maybe strong?*

Open challenges: non-discrete numerics,
regret analysis, novel applications

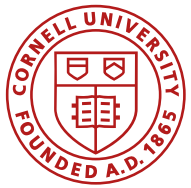
Up next: Part III

HTTPS://AVT.IM/ · @AVT_IM

Q. Xie, R. Astudillo, P. Frazier, Z. Scully, and A. Terenin. Cost-aware Bayesian optimization via the Pandora's Box Gittins index. *NeurIPS*, 2024. **INFORMS Data Mining Best Paper Finalist.**

Q. Xie, L. Cai, A. Terenin, P. Frazier, and Z. Scully. Cost-aware Stopping for Bayesian Optimization. *In review at NeurIPS*, 2025.

Z. Scully and A. Terenin. Gittins Index: A Design Principle for Decision-making Under Uncertainty. *INFORMS Tutorials in Operations Research*, 2025.



Cornell University