On Kernelized Multi-Armed Bandits with *Constraints*

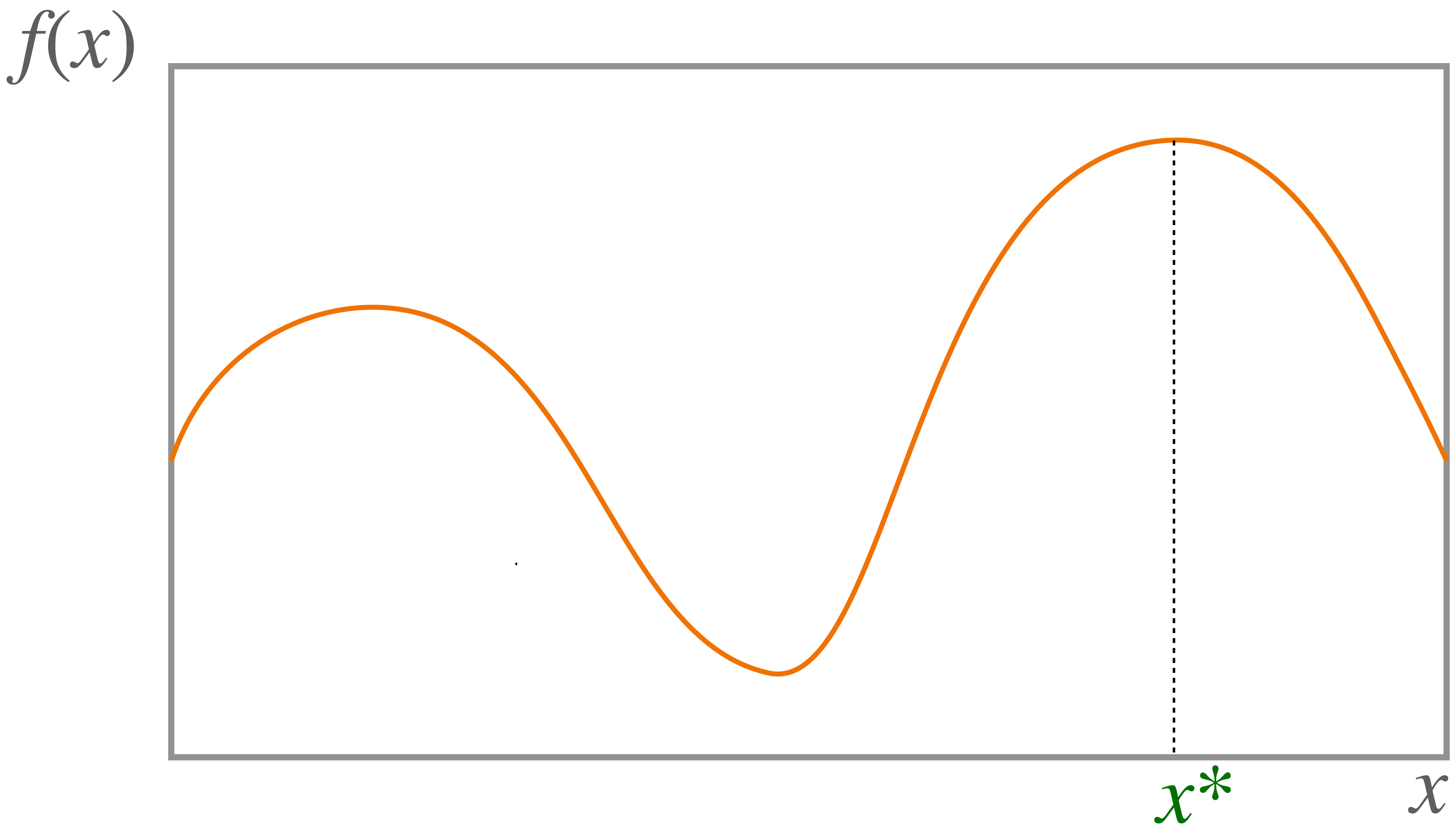
Xingyu Zhou, **Bo Ji**

Wayne State University, Virginia Tech

NEURAL INFORMATION PROCESSING SYSTEMS

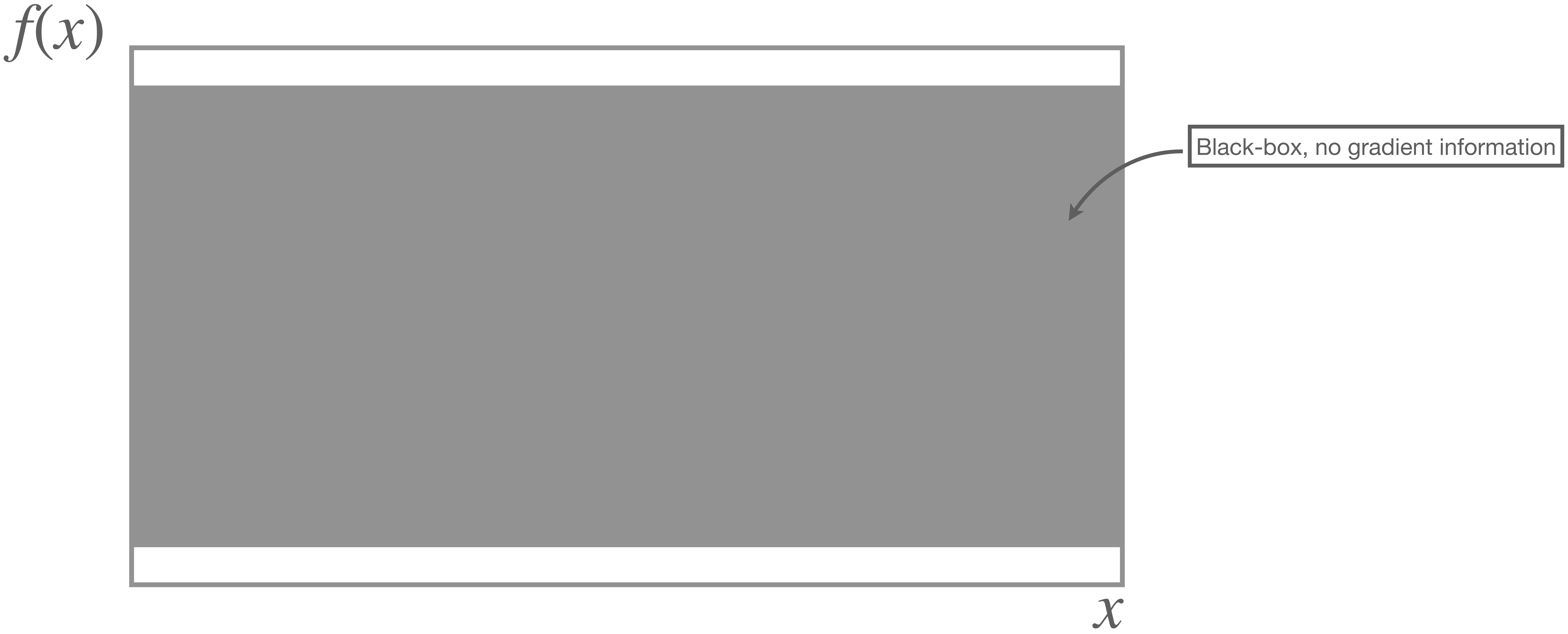
Sequential Decision Making

Black-box optimization



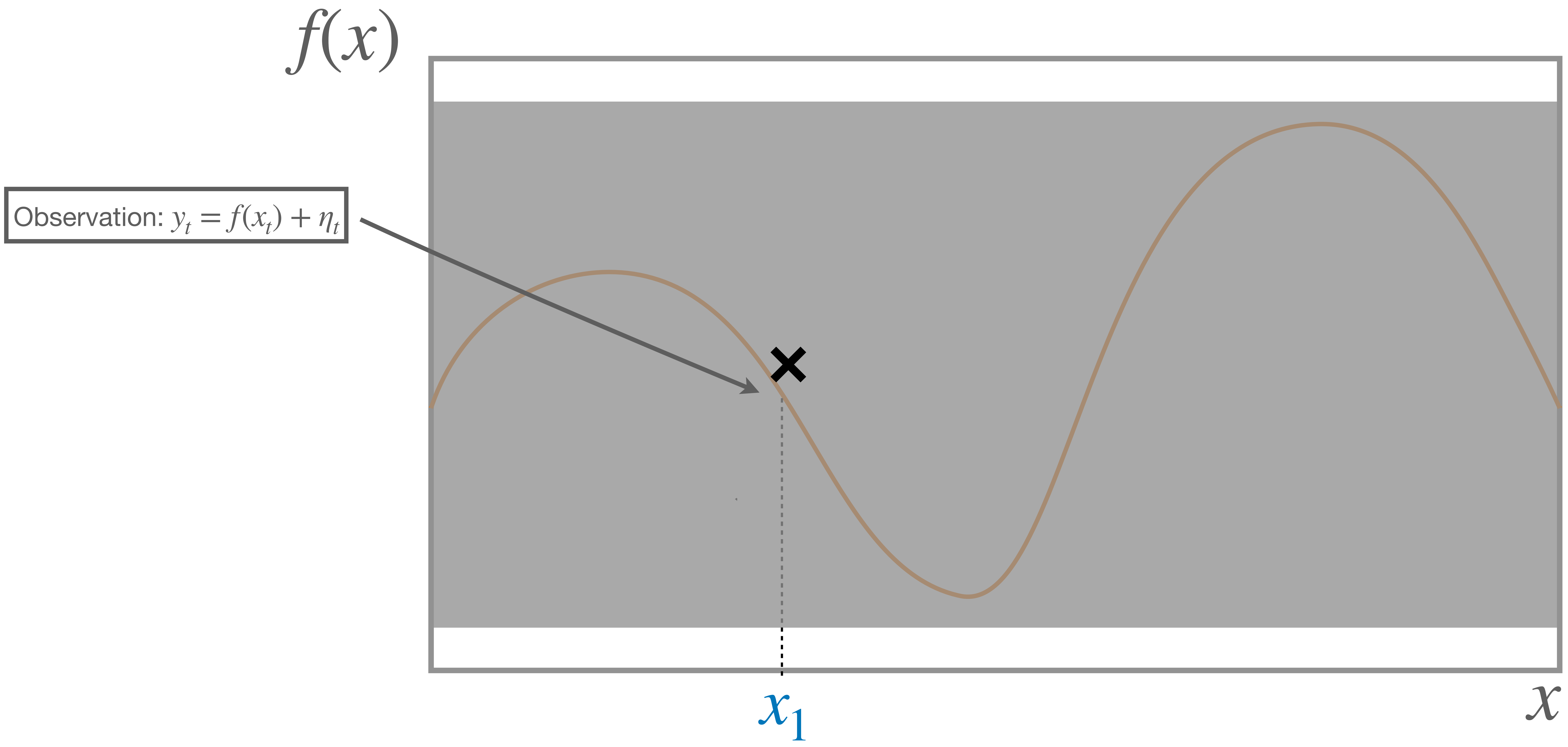
Sequential Decision Making

Black-box optimization



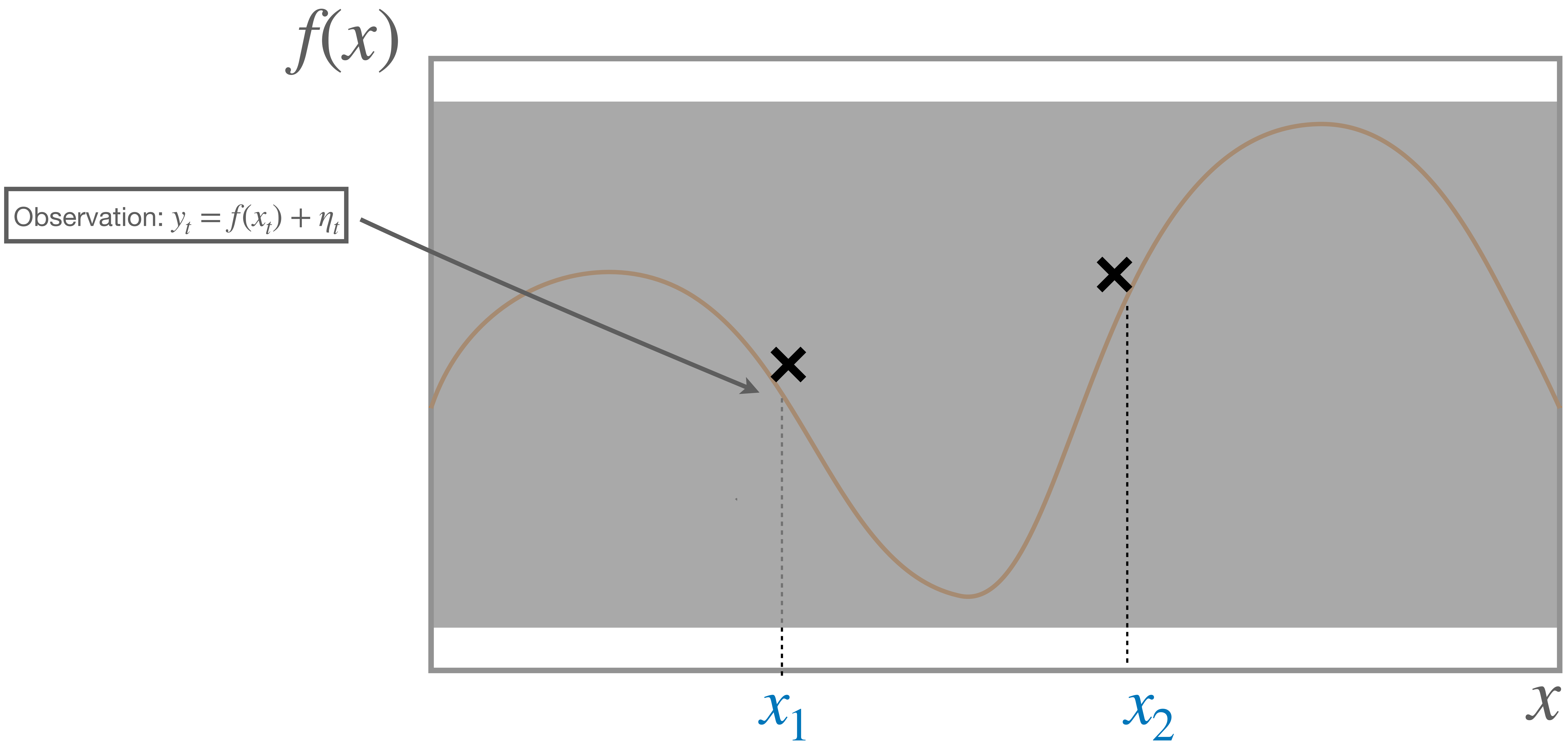
Sequential Decision Making

Black-box optimization



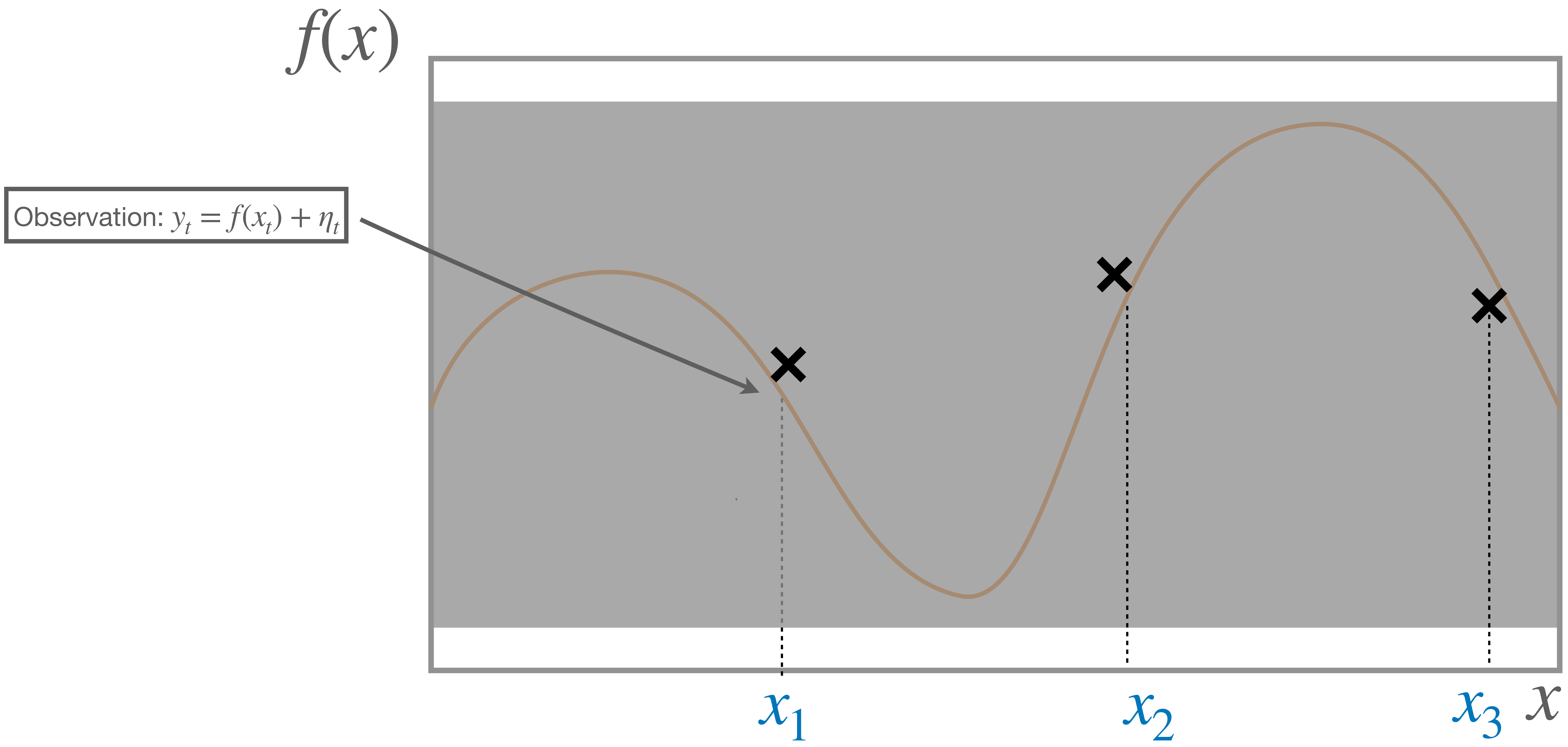
Sequential Decision Making

Black-box optimization



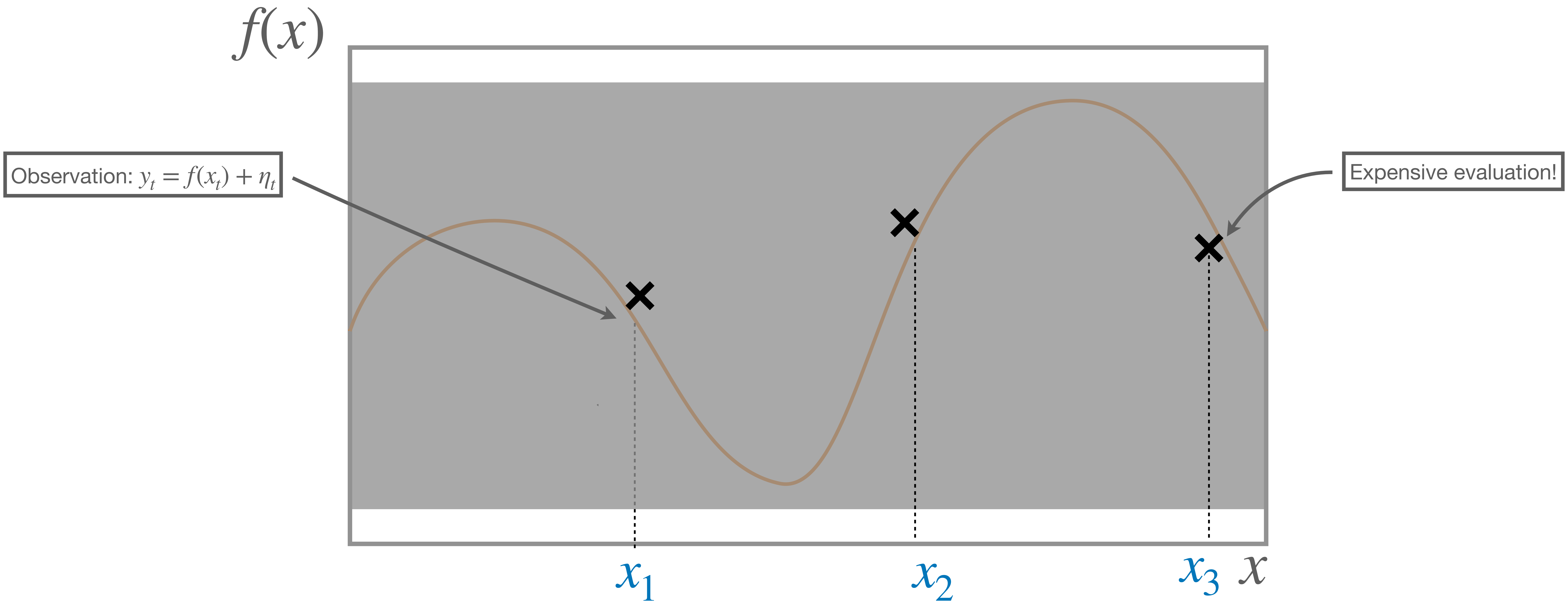
Sequential Decision Making

Black-box optimization



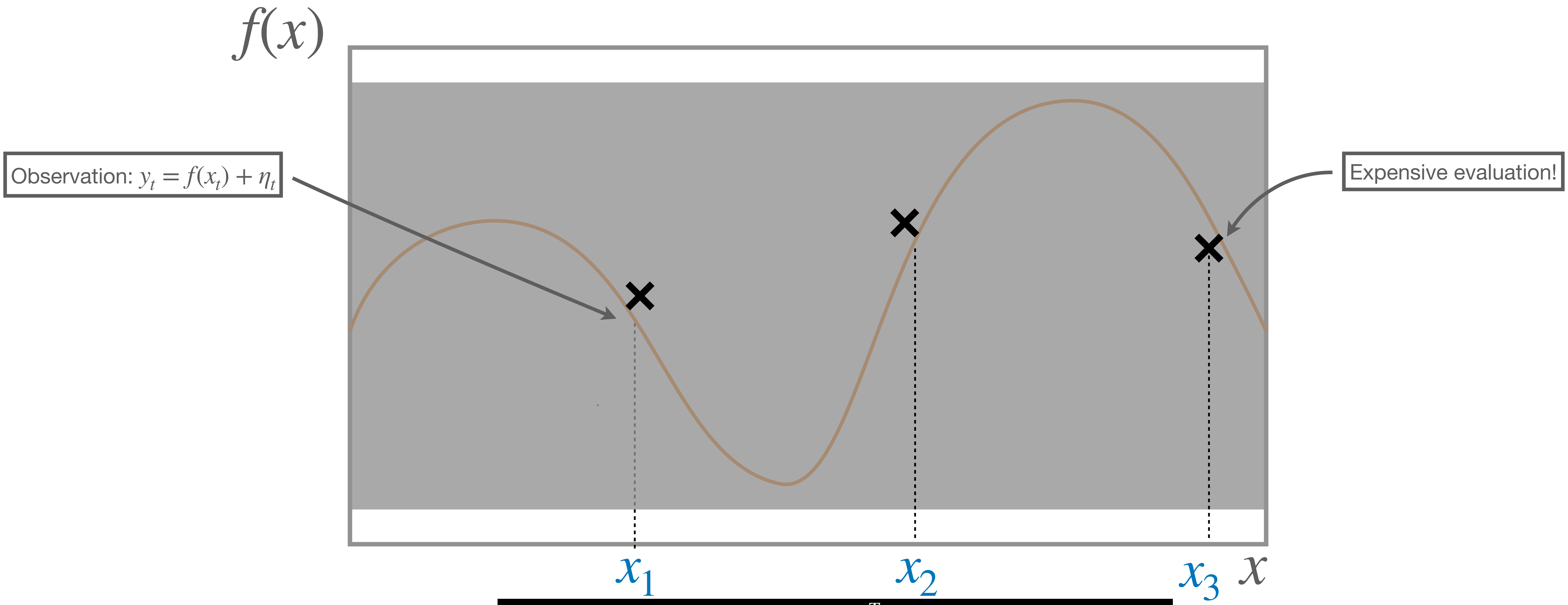
Sequential Decision Making

Black-box optimization



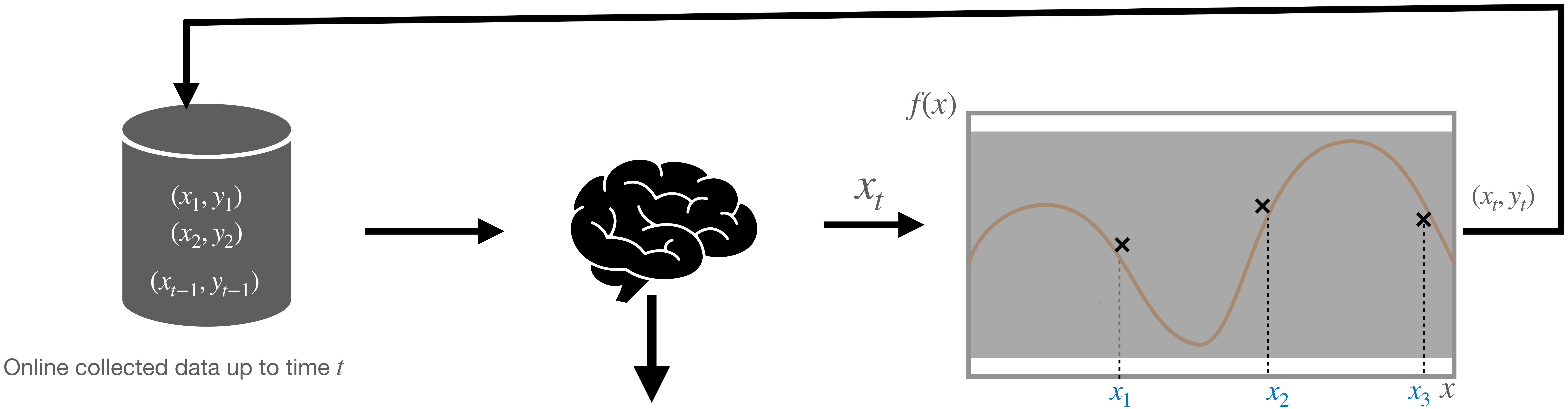
Sequential Decision Making

Black-box optimization

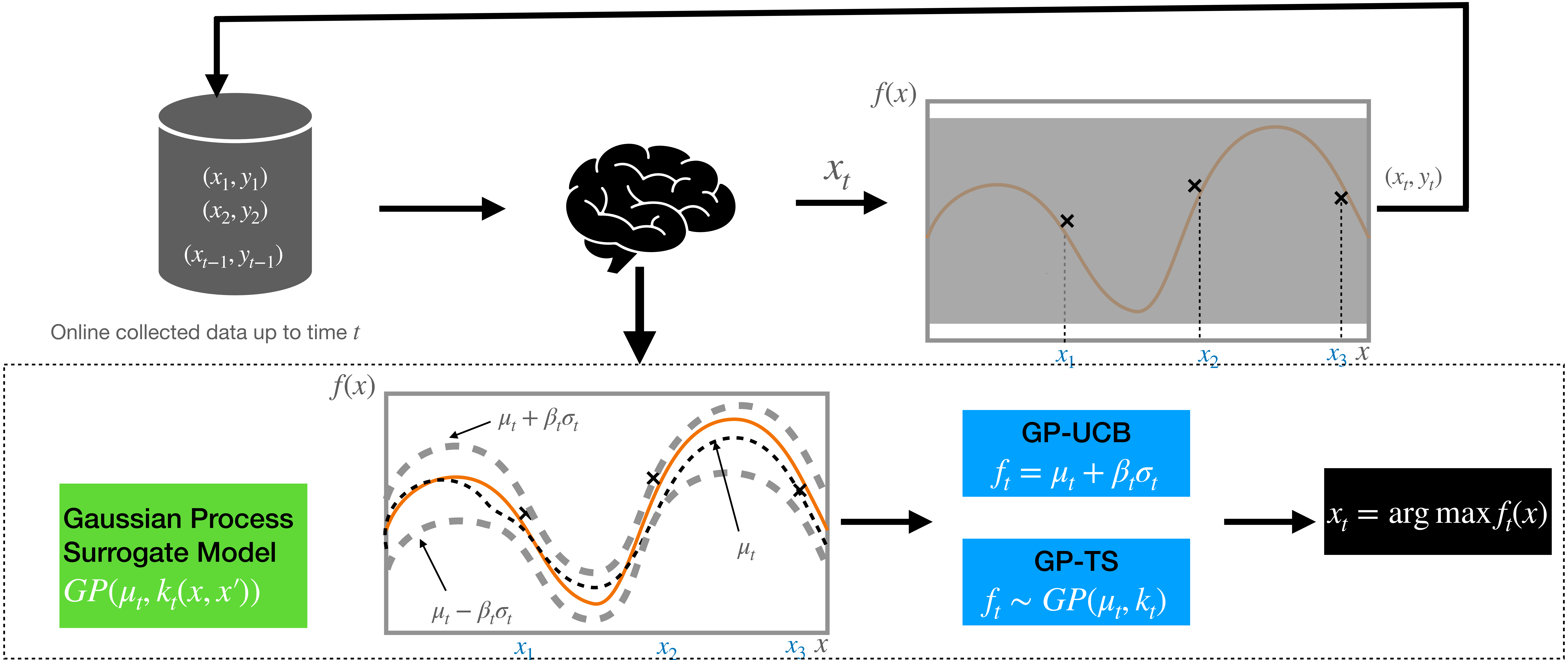


Minimize $R(T) = \sum_{t=1}^{T} f(x^*) - f(x_t)$

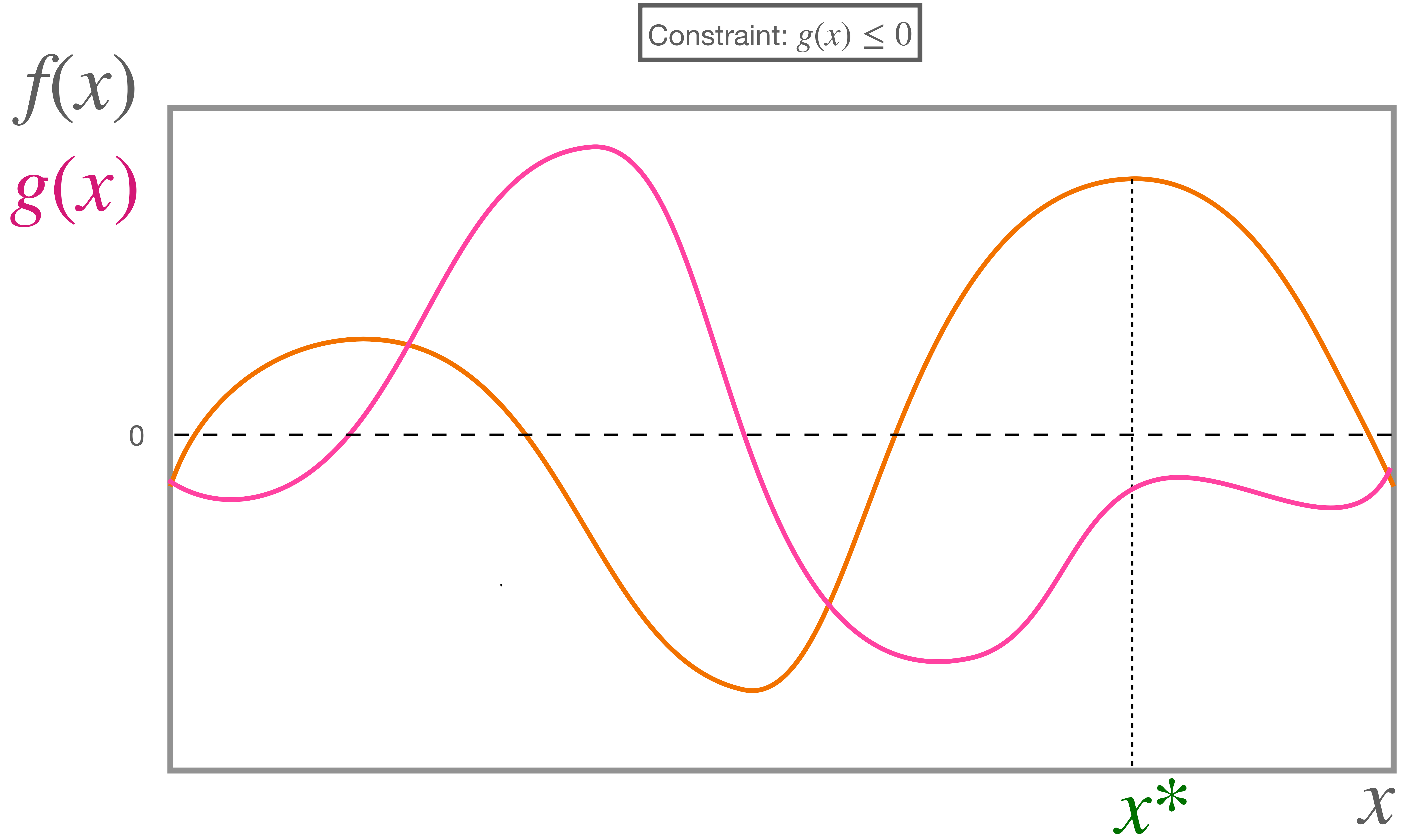
General Algorithm Design



General Algorithm Design



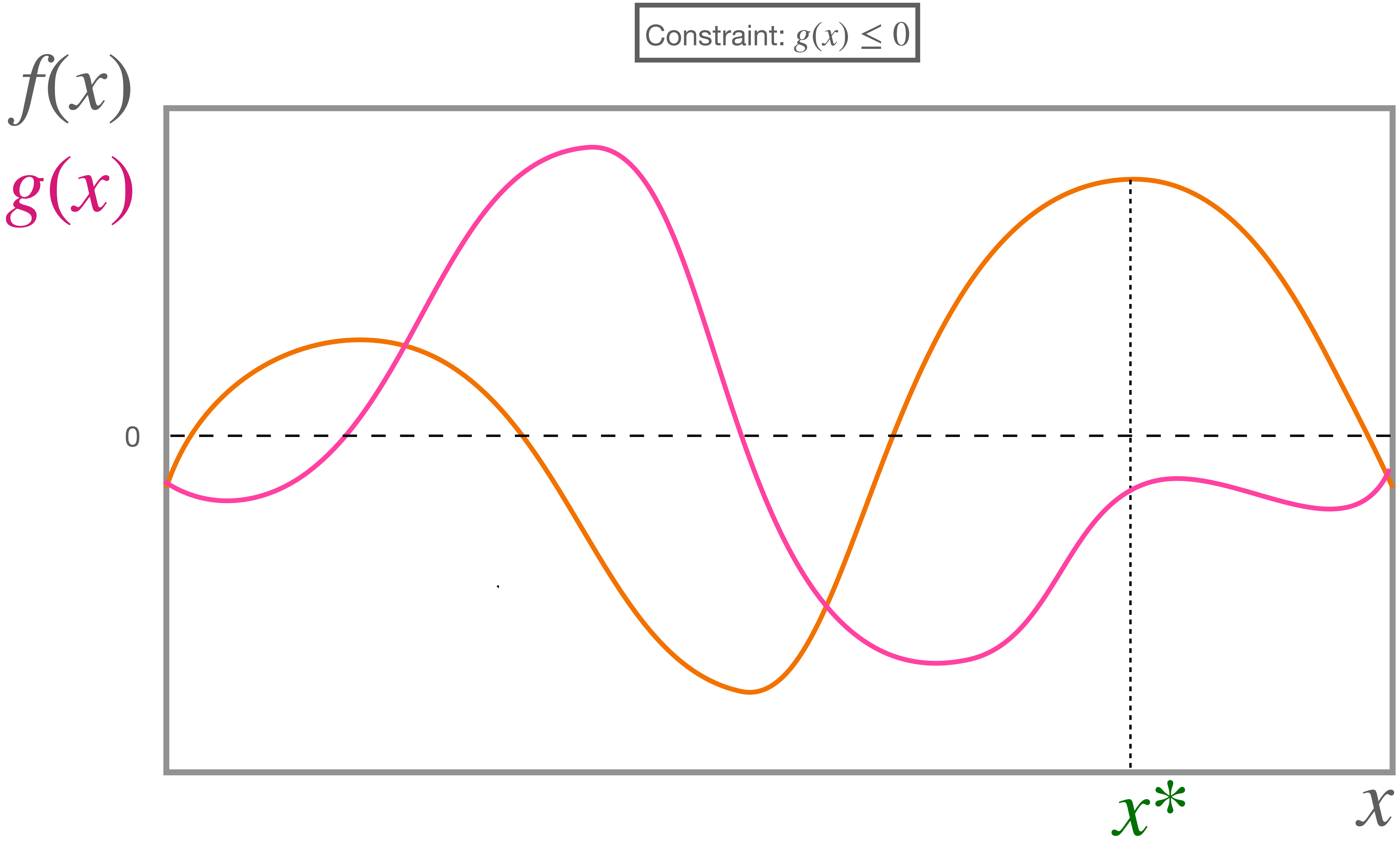
Constrained Kernelized Bandits (CKB)



Previous work with theoretical guarantees

- [SGBK'15, SBY'18, AAT'20]
- Hard constraints: each x_t satisfies constraint w.h.p
- Hence, additional computation is required
- Moreover, only GP-UCB is considered

Constrained Kernelized Bandits (CKB)



Previous work with theoretical guarantees

- [SGBK'15, SBY'18, AAT'20]
- Hard constraints: each x_t satisfies constraint w.h.p
- Hence, additional computation is required
- Moreover, only GP-UCB is considered

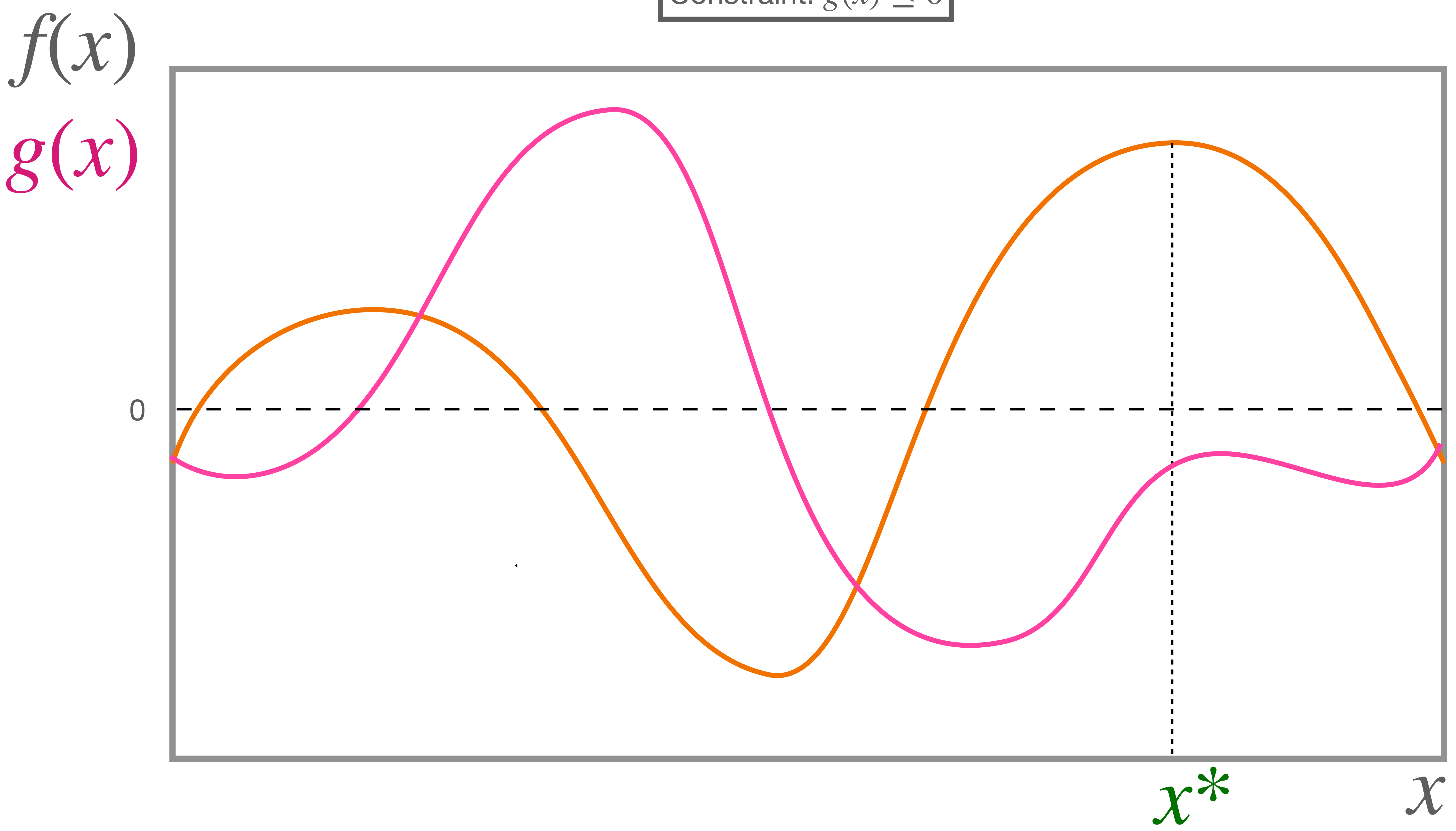
Motivations

- Practical soft constraints (e.g., energy)
- Maintain the same computation as before
- Other exploration strategy, e.g., GP-TS

Constrained Kernelized Bandits (CKB)

Soft Constraints

Constraint: $g(x) \leq 0$



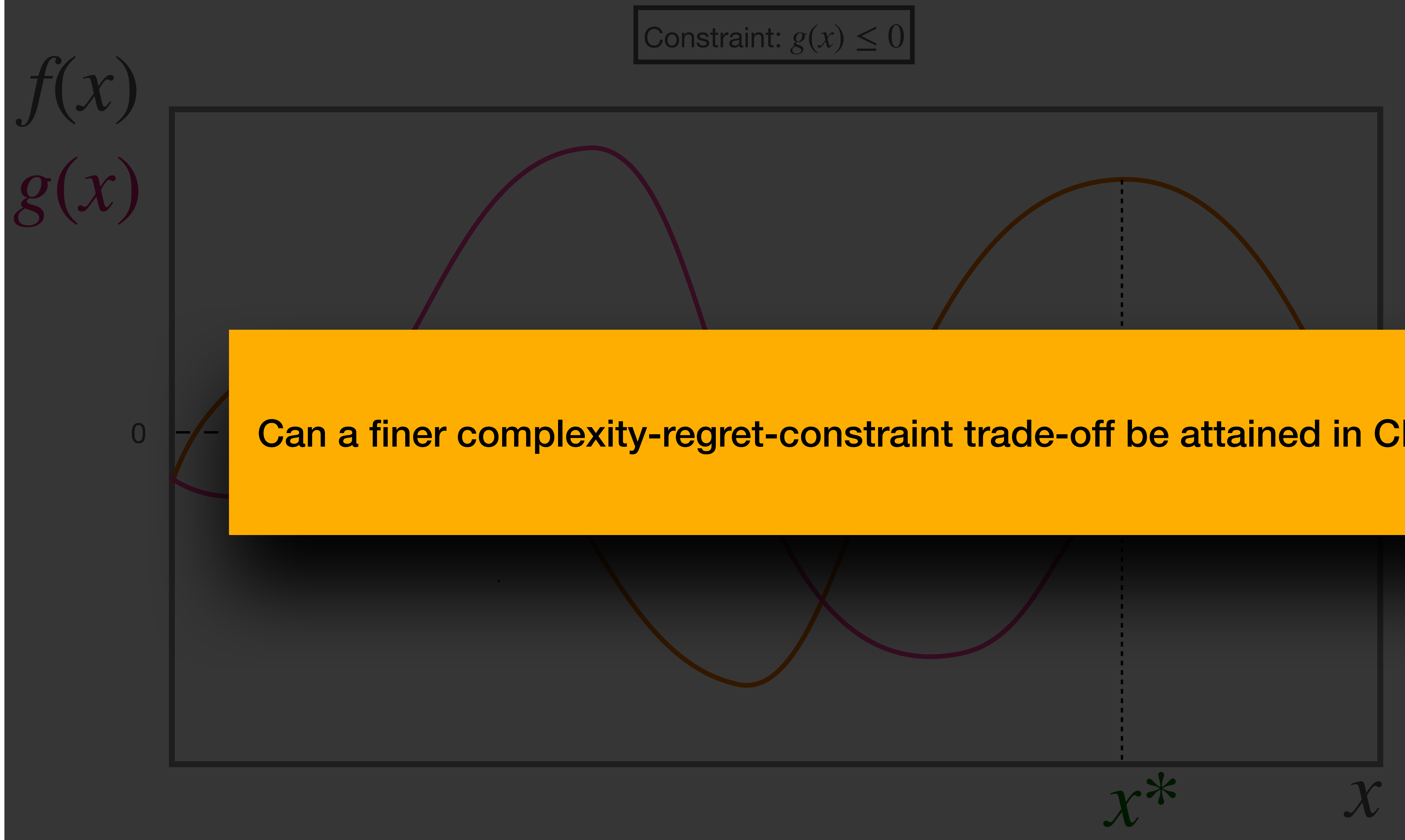
Minimize $R(T) = \sum_{t=1}^{T} f(x^*) - f(x_t)$

$V(T) = \sum_{t=1}^{T} g(x_t)$

Cumulative constraint violation

Constrained Kernelized Bandits (CKB)

Soft Constraints



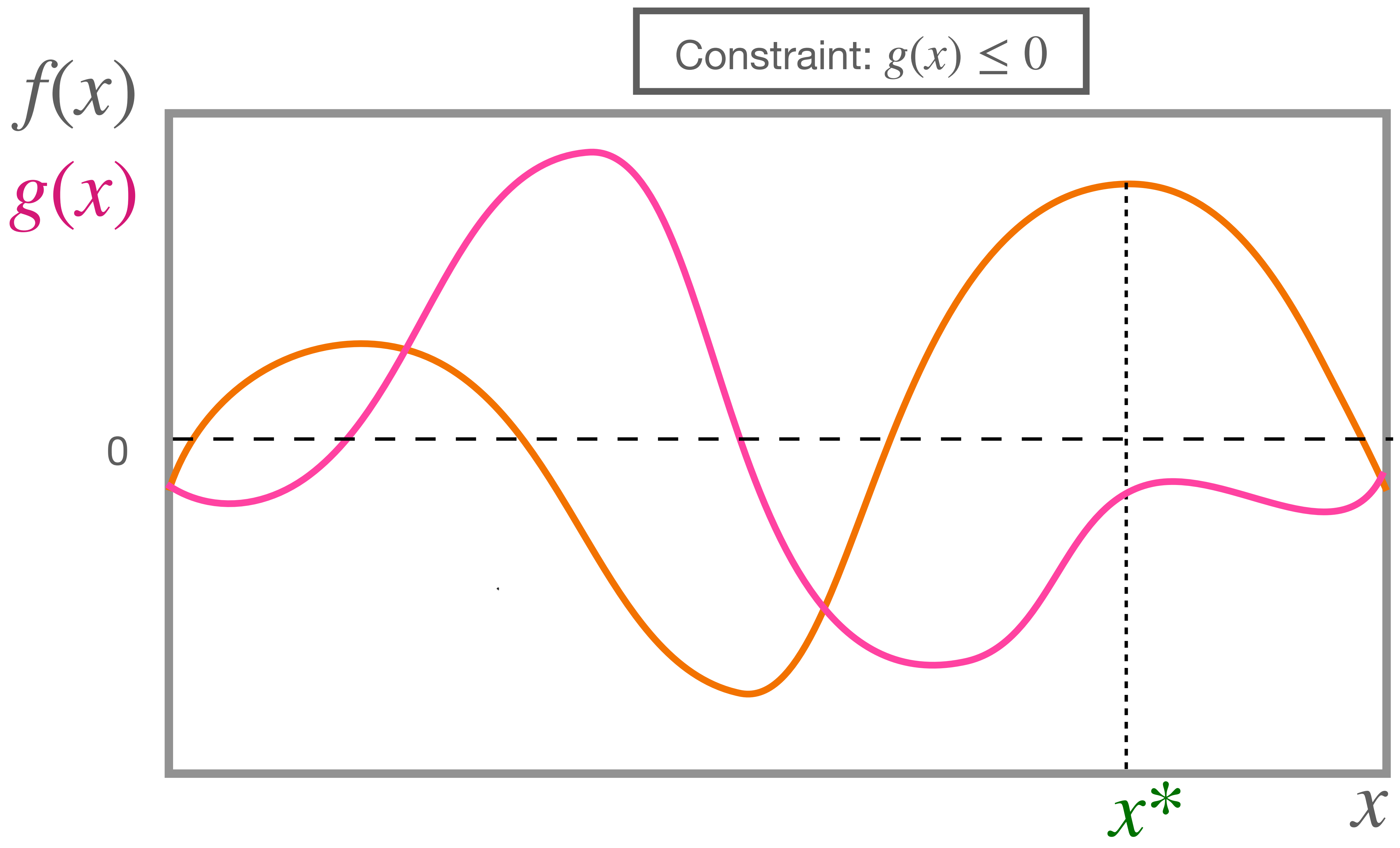
Minimize $R(T) = \sum_{t=1}^{T} f(x^*) - f(x_t)$

Cumulative constraint violation

Can a finer complexity-regret-constraint trade-off be attained in CKB under general explorations?

Contribution

Main Results



Minimize

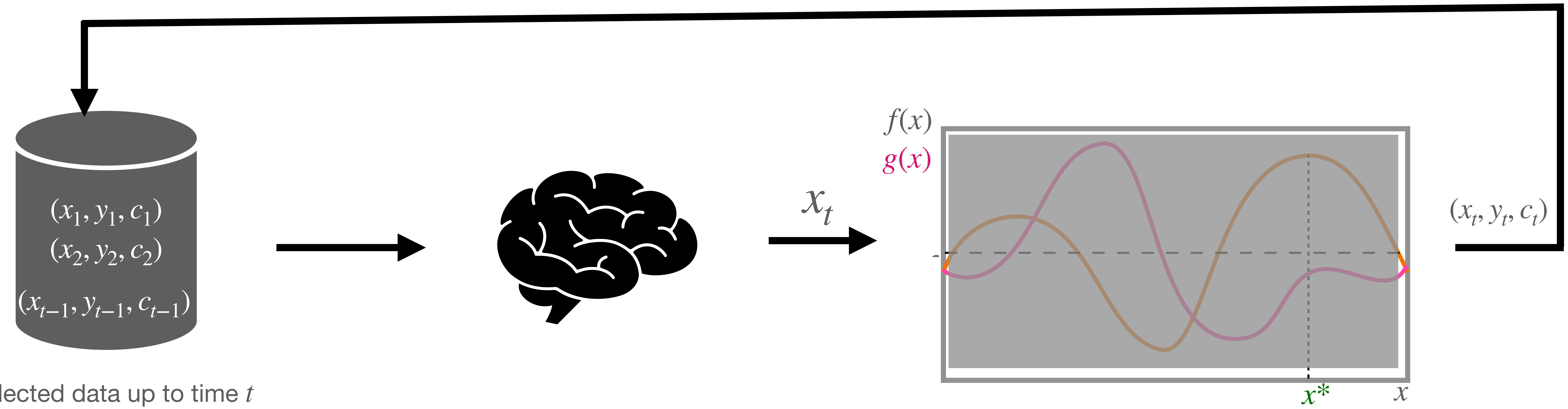
$$R(T) = \sum_{t=1}^{T} f(x^*) - f(x_t)$$

$$V(T) = \sum_{t=1}^{T} g(x_t)$$

1. Propose a generic CKB algorithm based on primal-dual optimization with $\tilde{O}(\gamma_T \sqrt{T})$ regret and zero constraint violation
2. This algorithm is compatible with GP-UCB, GP-TS, RandGP-UCB, and many more…
3. An extensive evaluations on both synthetic and real-world data
4. The first detailed discussion on two common techniques for constrained bandits

General Algorithm Design

Our CKB algorithm



Online collected data up to time t

General Algorithm Design

Our CKB algorithm

General Algorithm Design

Our CKB algorithm

A Generic Performance Bound

Many strategies satisfy our sufficient condition (let $h = f, g$)
GP-UCB: $h_t = \mu_t + \beta_t \sigma_t$
GP-TS: $h_t \sim GP(\mu_t, k_t)$
RandGP-UCB: $h_t = \mu_t + Z_t \sigma_t, Z_t \sim \mathcal{N}(0, \beta_t^2)$

Theorem

Suppose f_t and g_t satisfy the sufficient condition, then we have

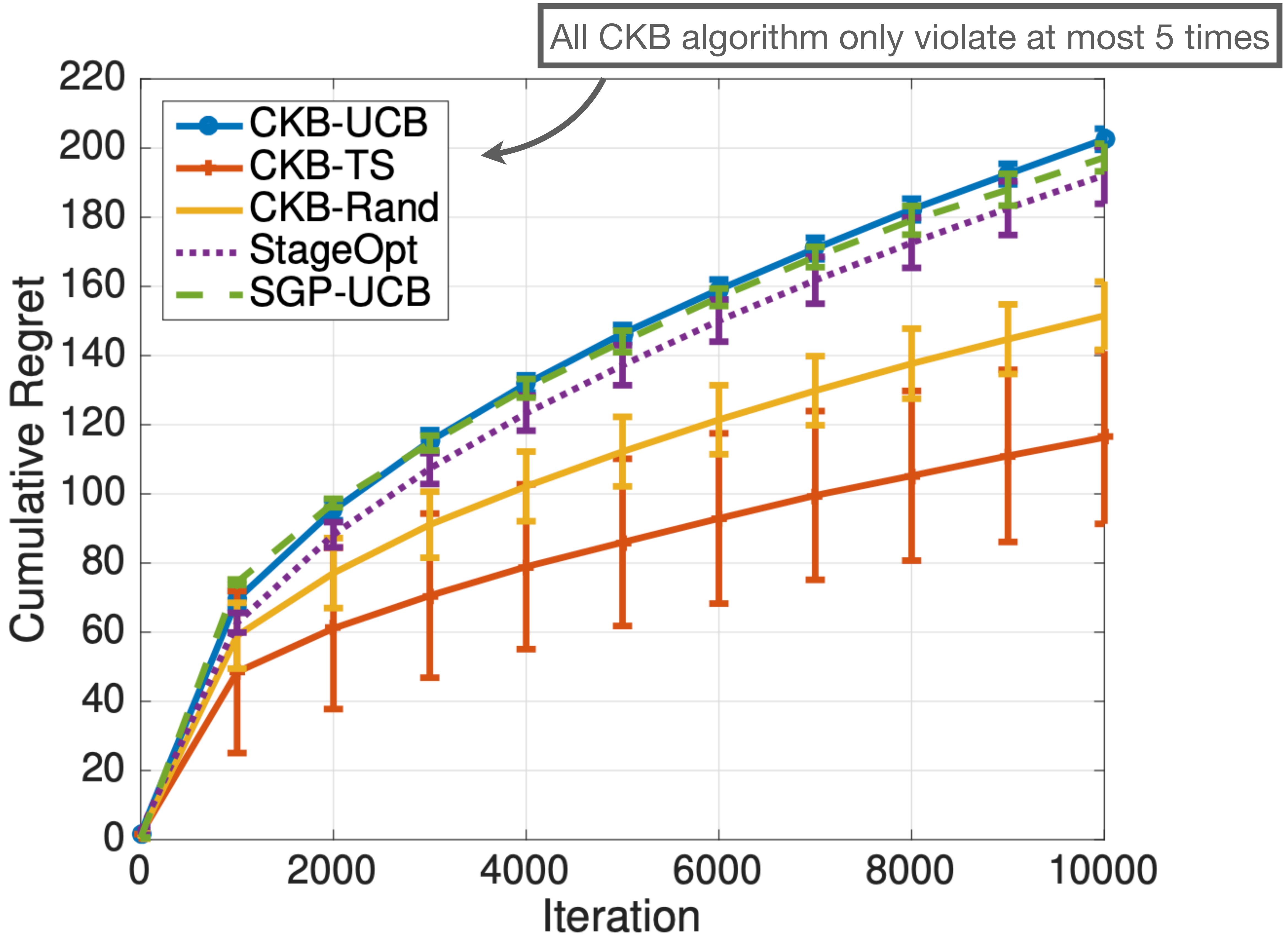
$$R(T) = \tilde{O}\left(\gamma_T \sqrt{T}\right) \qquad V(T) = 0, \text{ for a sufficient large } T$$

Maximum information gain
Linear kernel: $O(d \ln T)$
SE kernel: $O((\ln T)^{d+1})$

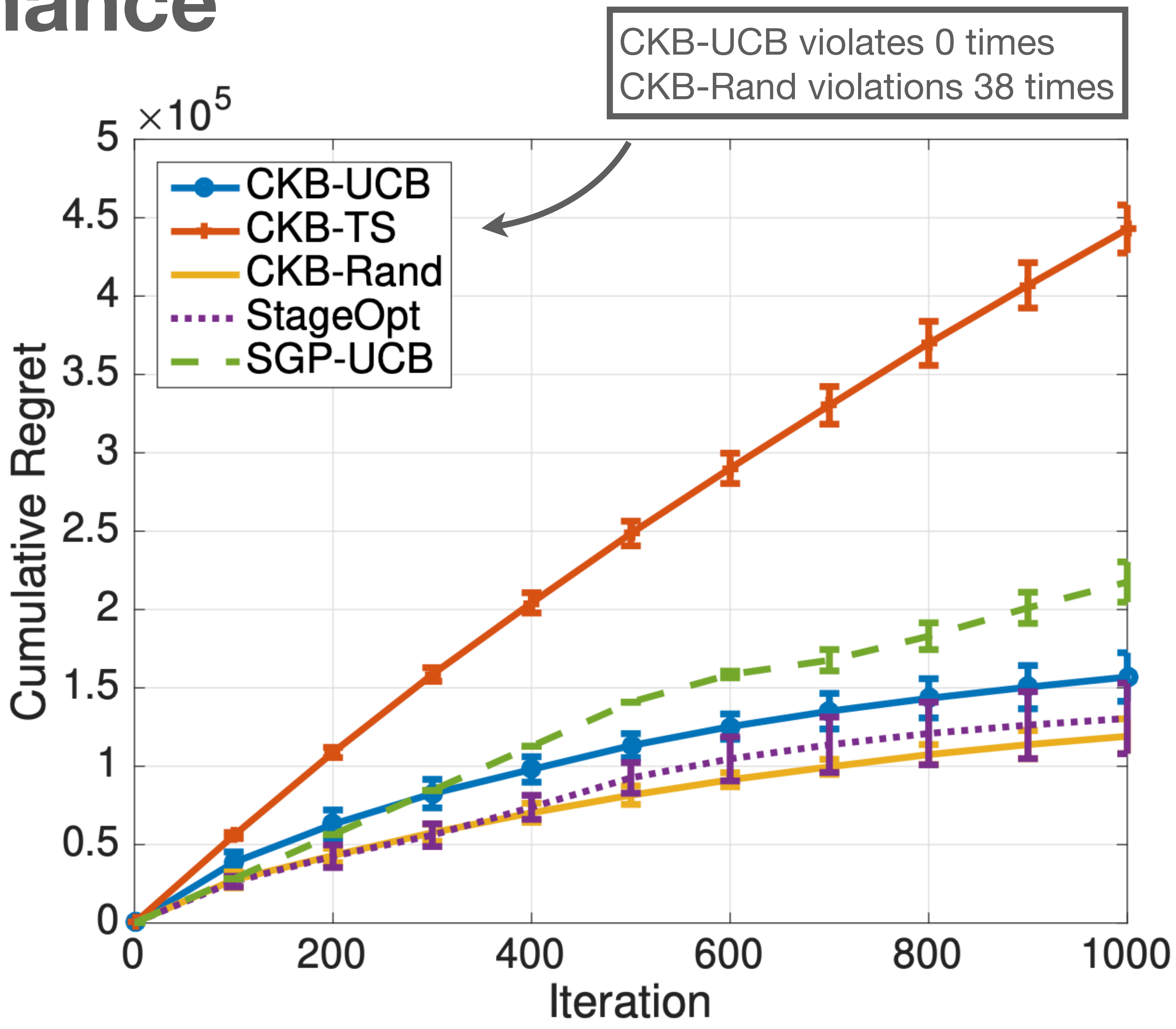
For small T, we also have $V(T) = \tilde{O}\left(\gamma_T \sqrt{T}\right)$

Evaluations

Regret / constraint violation performance



synthetic data

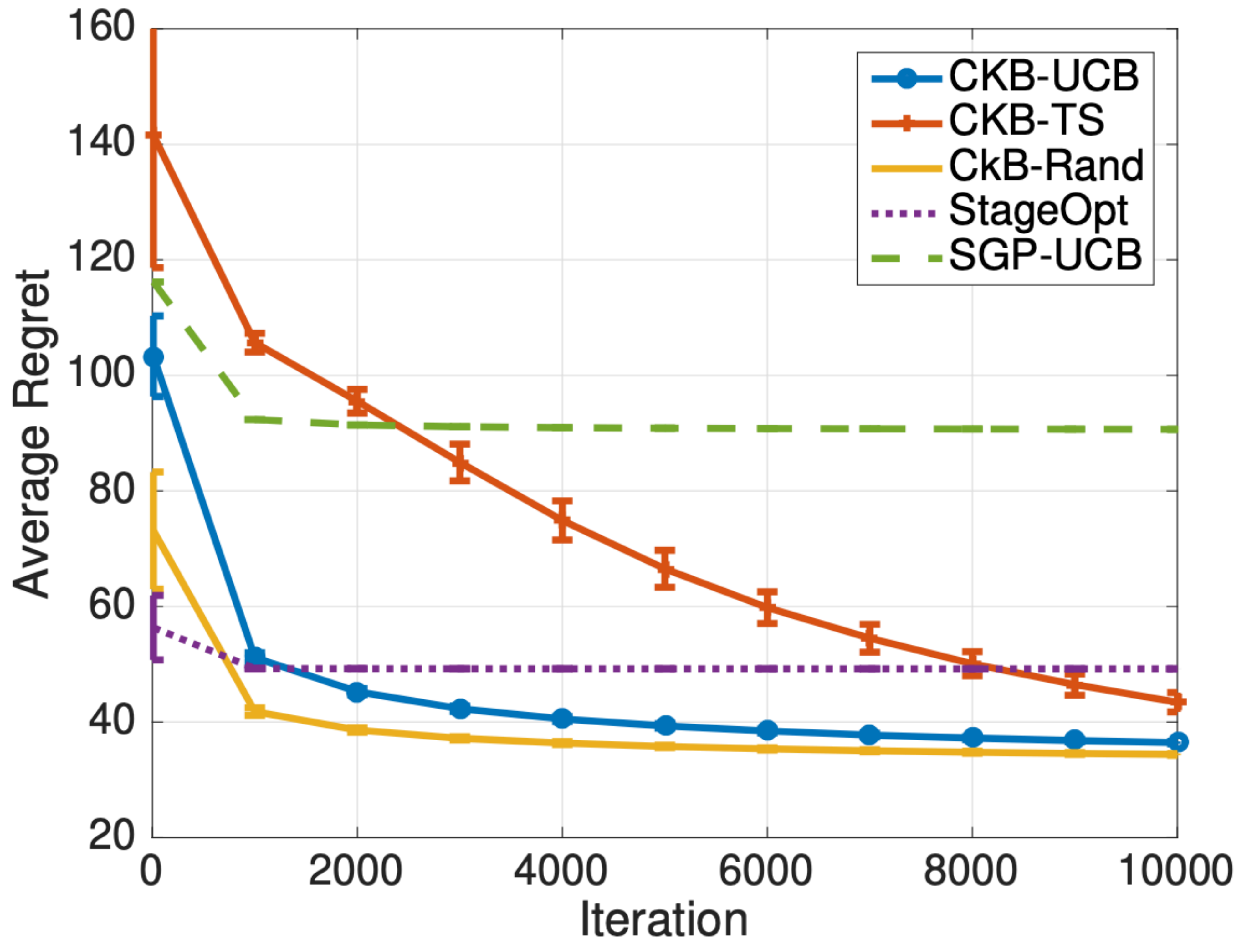


real-world data

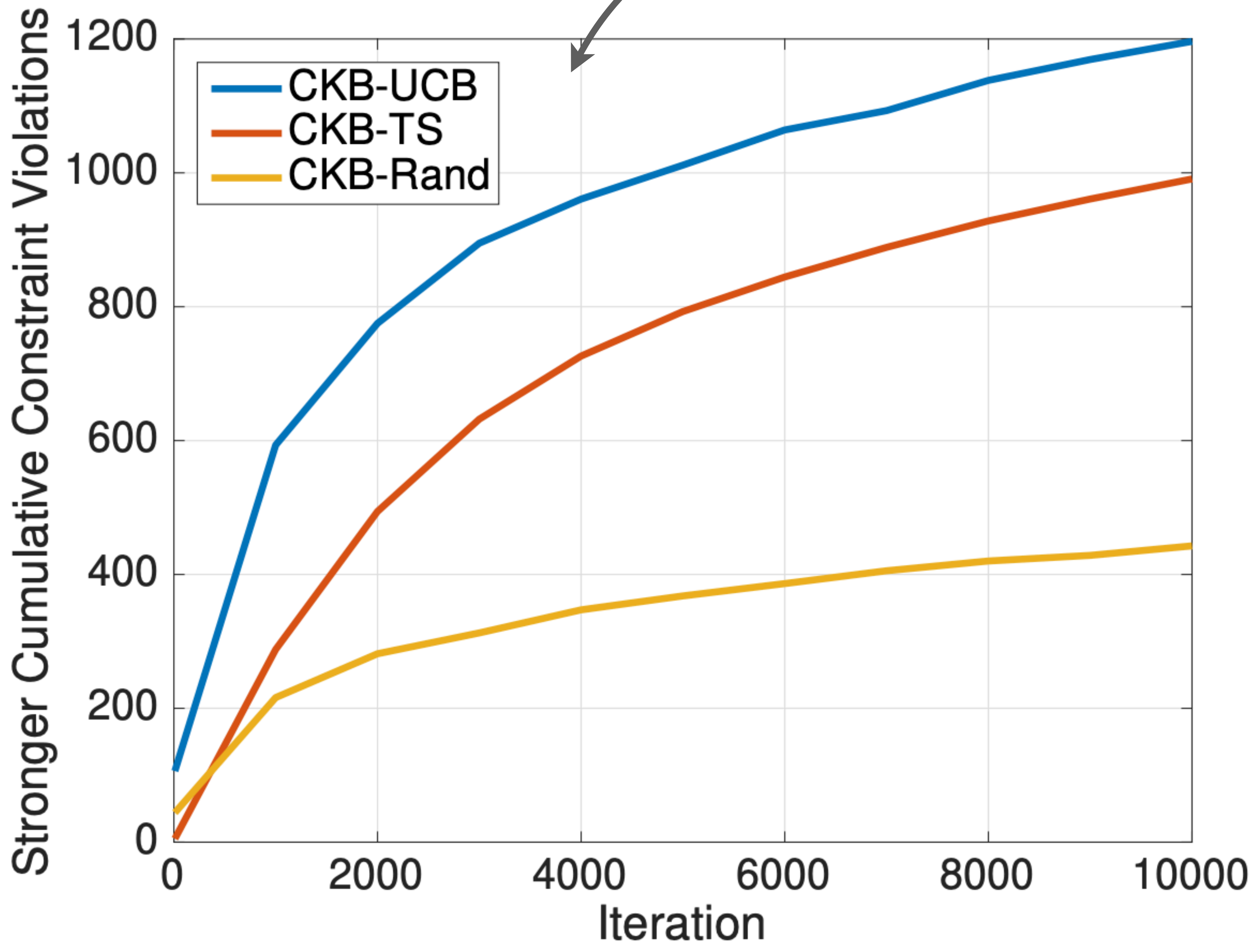
Evaluations - Heavy-tailed

Regret / constraint violation performance

We plot the stronger metric:
$V_+(T) = \sum_t \max(0, g(x_t))$

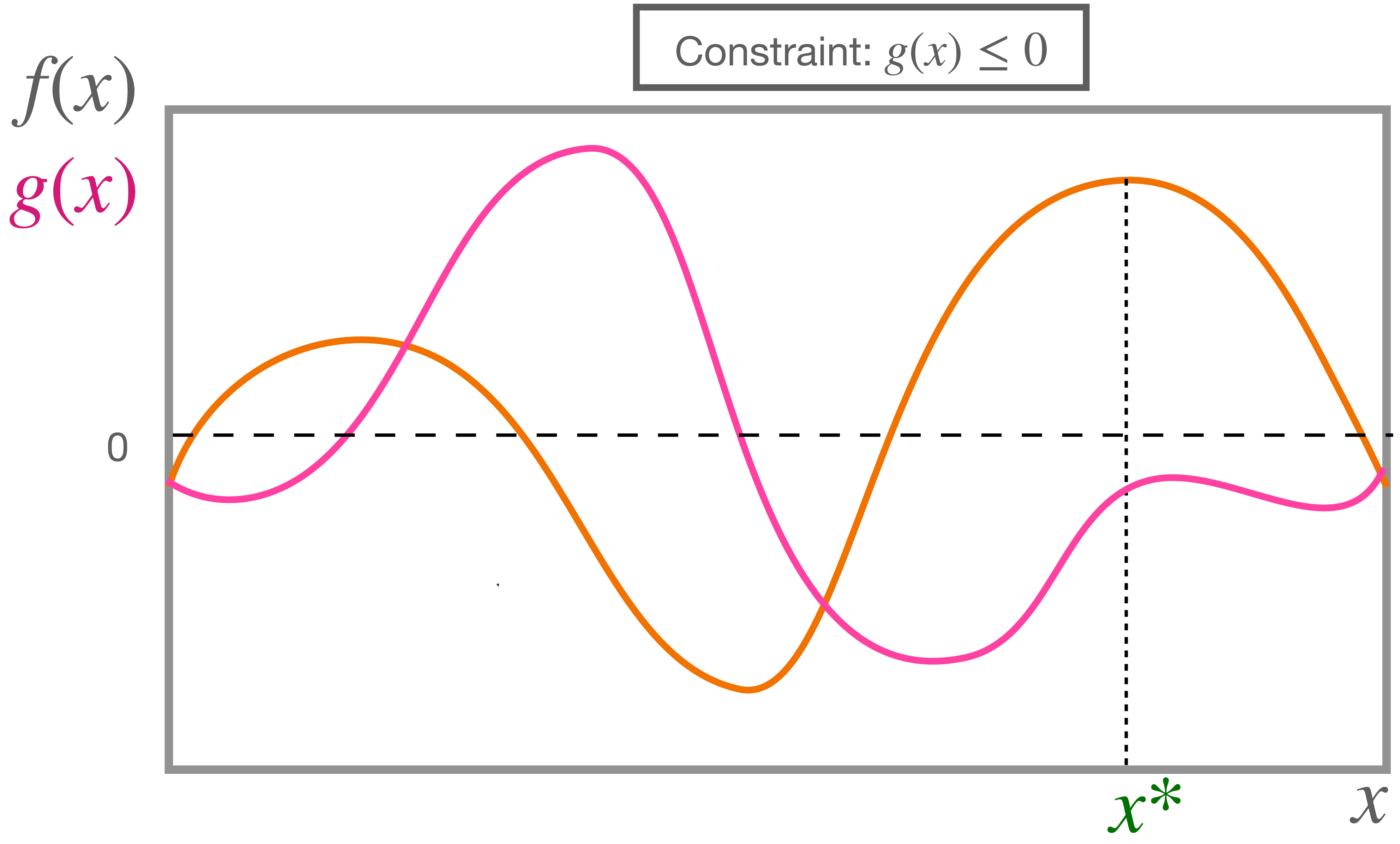


Heavy-tailed financial data (Regret)



Heavy-tailed financial data (Constraint)

Conclusion



Minimize

$$R(T) = \sum_{t=1}^{T} f(x^*) - f(x_t)$$

$$V(T) = \sum_{t=1}^{T} g(x_t)$$

1. Propose a generic CKB algorithm based on primal-dual optimization with $\tilde{O}(\gamma_T \sqrt{T})$ regret and zero constraint violation
2. This algorithm is compatible with GP-UCB, GP-TS, RandGP-UCB, and many more…
3. An extensive evaluations on both synthetic and real-world data
4. The first detailed discussion on two common techniques for constrained bandits

Thank you!