# A General Recipe for Likelihood-free Bayesian Optimization
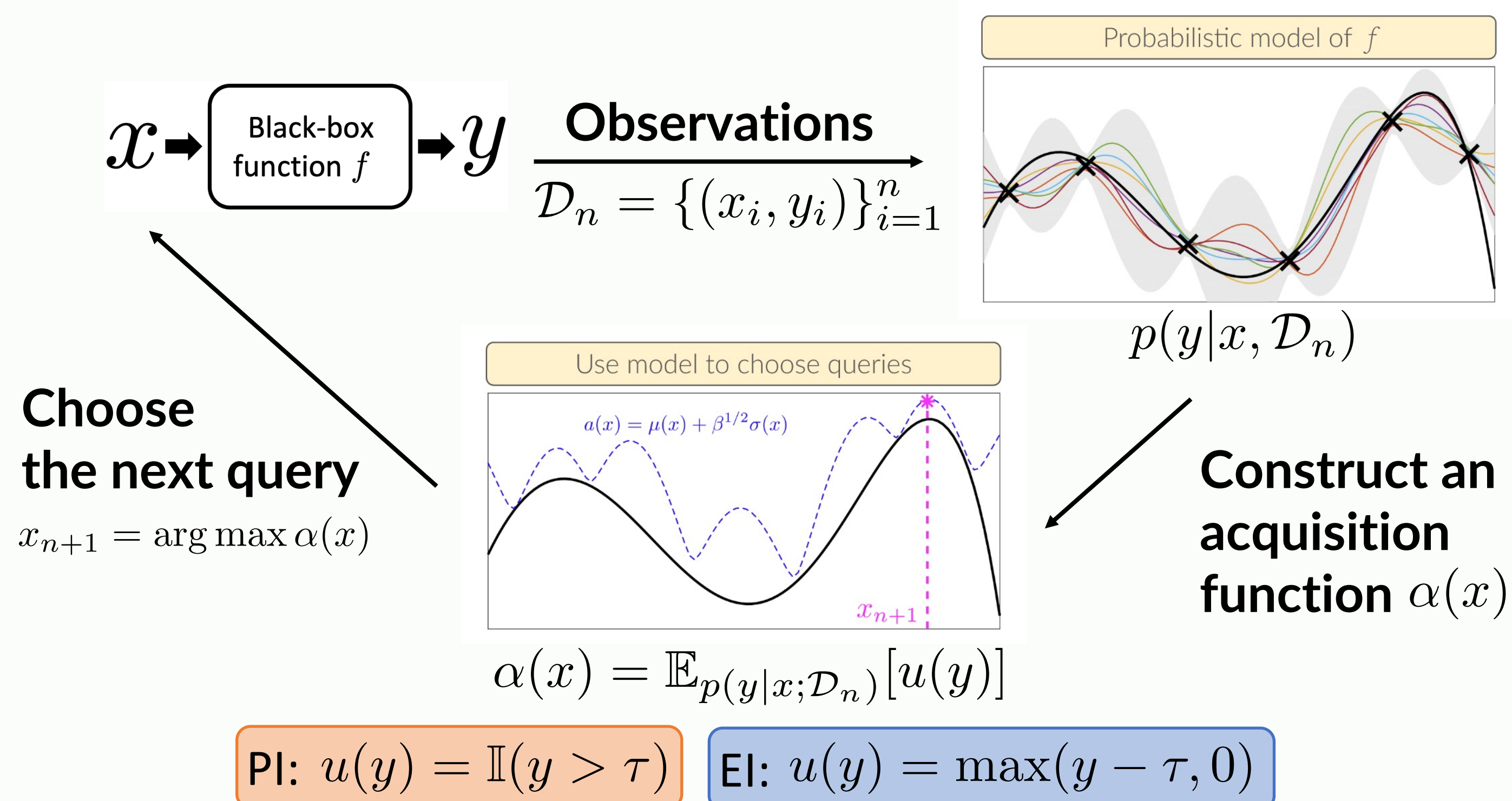
Jiaming Song*, Lantao Yu*, Willie Neiswanger, Stefano Ermon

Stanford ARTIFICIAL INTELLIGENCE

## BO for Global Optimization of Black-box Functions

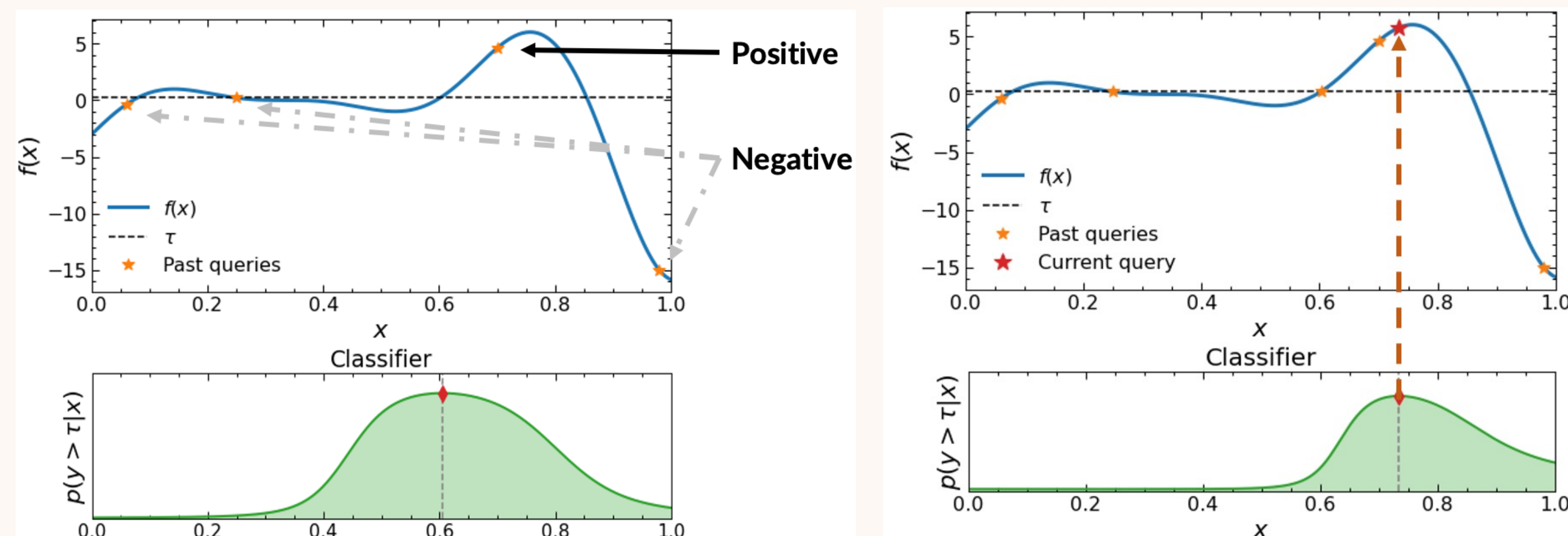**Function evaluations are expensive, e.g. neural architecture search and material discovery.**

$x \rightarrow$ [Black-box function $f$] $\rightarrow y$ $\xrightarrow{\textbf{Observations}}$ $\mathcal{D}_n = \{(x_i, y_i)\}_{i=1}^n$

Probabilistic model of $f$

$p(y|x, \mathcal{D}_n)$

**Choose the next query**

$x_{n+1} = \arg\max \alpha(x)$

Use model to choose queries

$a(x) = \mu(x) + \beta^{1/2}\sigma(x)$

$x_{n+1}$

**Construct an acquisition function $\alpha(x)$**

$$\alpha(x) = \mathbb{E}_{p(y|x;\mathcal{D}_n)}[u(y)]$$

PI: $u(y) = \mathbb{I}(y > \tau)$    EI: $u(y) = \max(y - \tau, 0)$

## Likelihood-free Bayesian Optimization

Reweight the "positive" queries by their utility values!

Black-box function

Weights based on EI utility

Black-box function

Weights based on EI utility
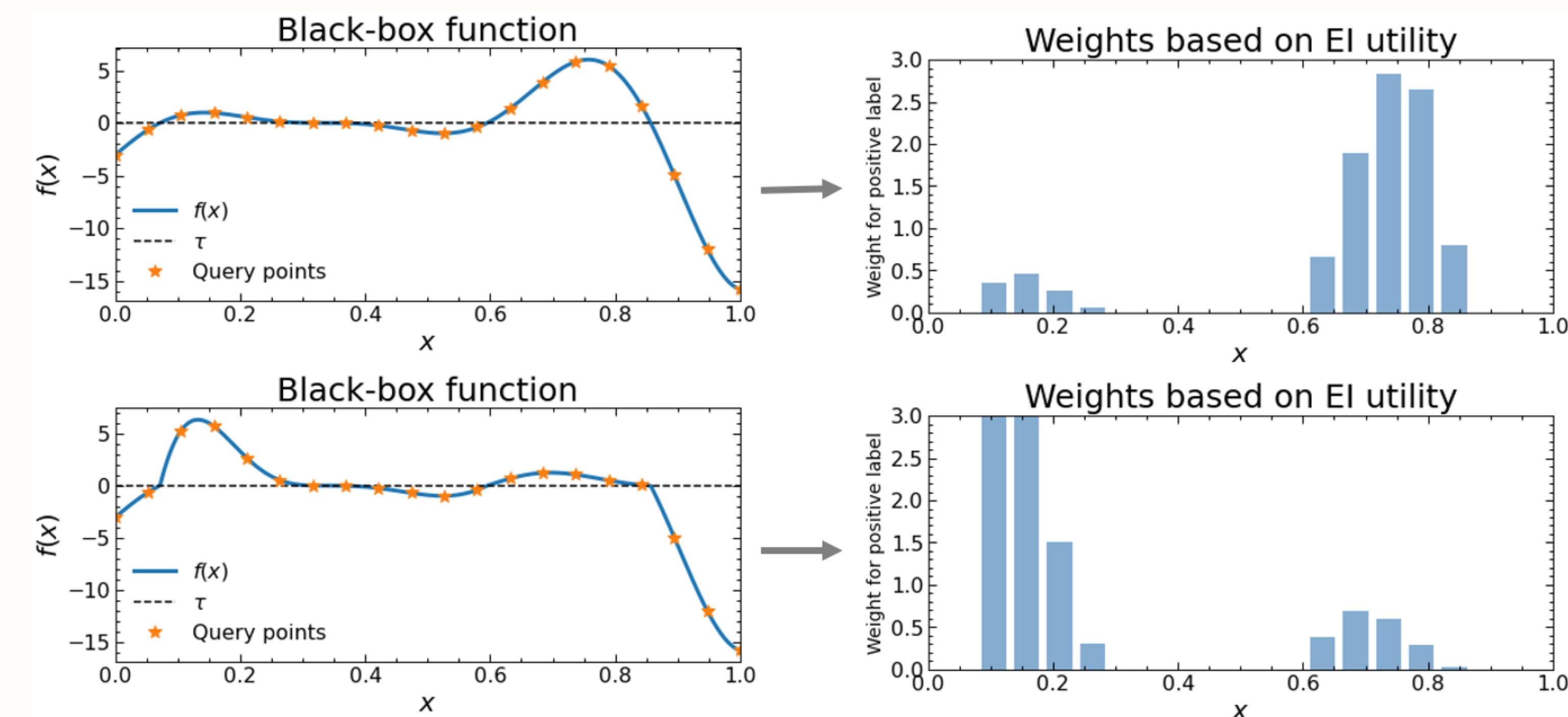
Theorem: LFBO converges to desired acquisition function asymptotically.

## Bayesian Optimization via Classification (BORE)

Positive

Negative

$f(x)$, $\tau$, Past queries

$p(y > \tau|x)$ — Classifier

$f(x)$, $\tau$, Past queries, Current query

$p(y > \tau|x)$ — Classifier

**Train a classifier to predict whether a point is above the threshold.**

**Then use the classifier to select the next query.**

This approach does not care a point is over the threshold by how much.  $\Longrightarrow$  Only equivalent to PI

## Experiments: Hyperparameter Tuning & Neural Architecture Search

Protein

Parkinsons

Slice

Naval

immediate regret

# of queries

Random
SMAC
TPE
GP
BORE (MLP)
LFBO (MLP)
BORE (XGB)
LFBO (XGB)
BORE (RF)
LFBO (RF)