CATALYSIS CENTER FOR ENERGY INNOVATION
AN ENERGY FRONTIER RESEARCH CENTER

NEXTorch: A Design and Bayesian Optimization Toolkit for Chemical Sciences and Engineering

Yifan Wang

Dec 2023 Vlab Workshop

U.S. DEPARTMENT OF ENERGY
Office of Science

An Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Basic Sciences

About Me

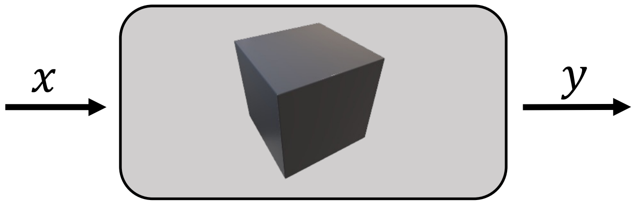
- Graduated from Vlachos Group Dec 2021
- ML Research Scientist at Meta
- Based in San Francisco
- Develop models and tools for Meta's data labeling and RLHF platform
- LinkedIn: https://www.linkedin.com/in/wangyifan411/

Global Optimization of Blackbox Functions

Blackbox functions [1]

- Expensive computer model or laboratory experiments
- Unknown explicit model form
- Multi-dimensional

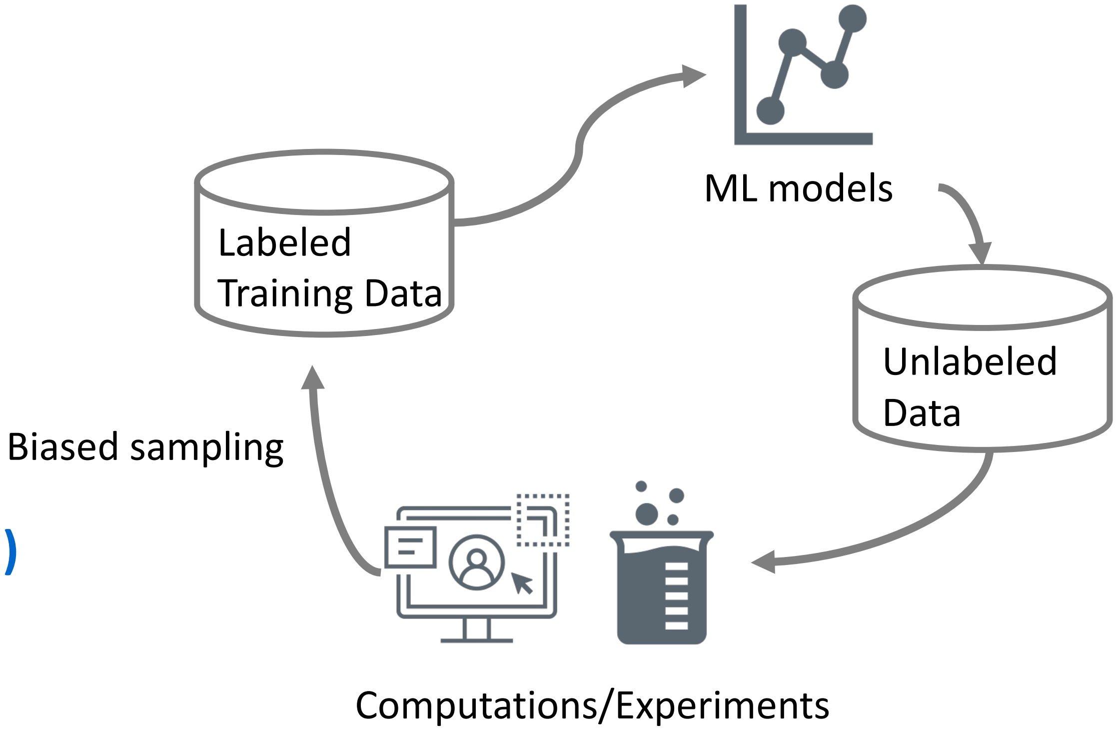
$f(x)$



$$\max_{x \in A} f(x)$$

Active learning [2]

- An algorithm "learning" from data, proposing next experiments, and improving prediction accuracy with fewer training data or lower cost

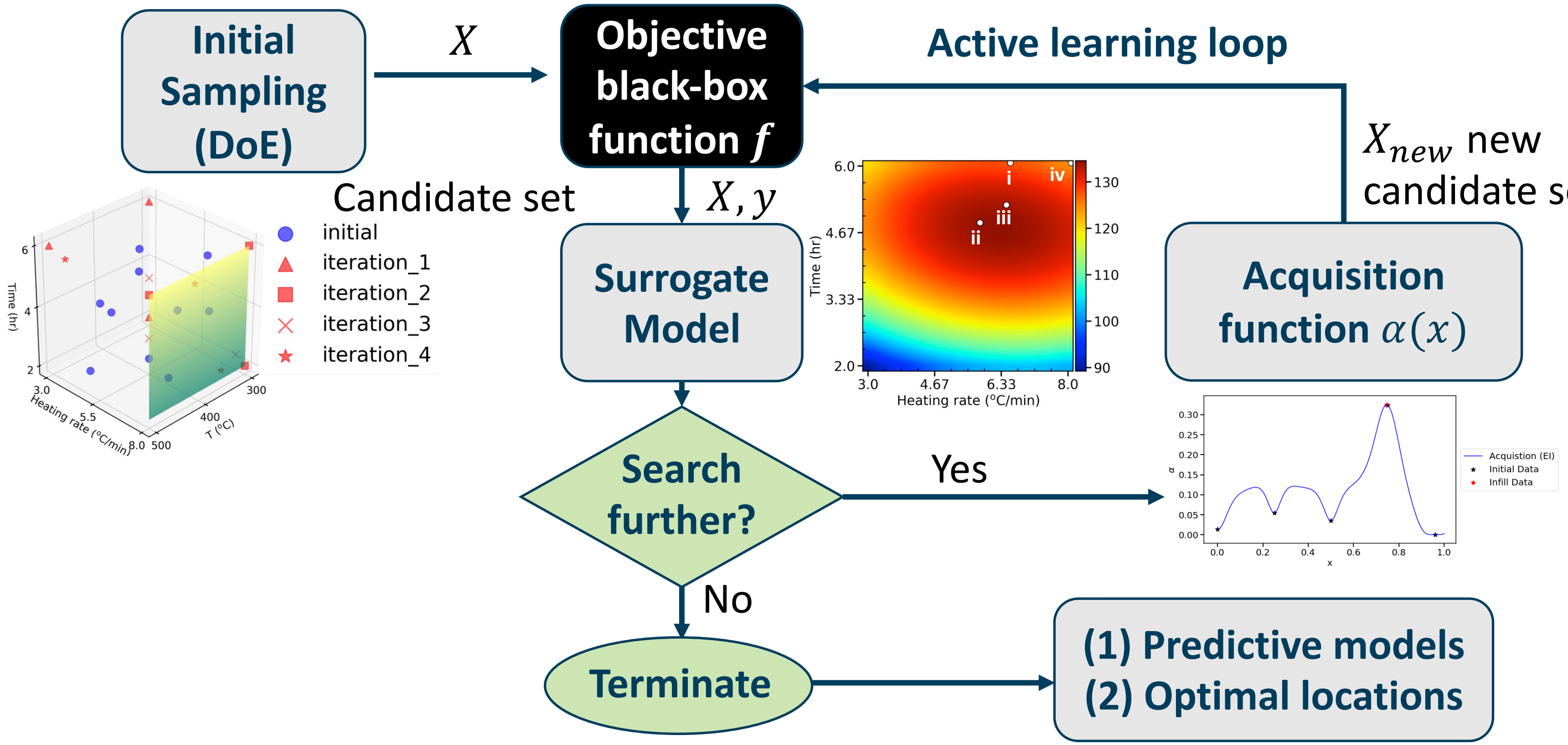


Use active learning to reduce experimental (computational) cost and improve accuracy of the surrogate model

[1] D.R. Jones, M. Schonlau, and W. J. Welch, J. Glob. Optim. **13**, 455 (1998).
[2] Settles, B. *Active Learning Literature Survey*. *Active Learning Literature Survey* (2009).

CATALYSIS CENTER FOR ENERGY INNOVATION

Bayesian Optimization (BO)



Bayesian Statistics

Posterior Data Prior

$$P(f|D) \propto P(D|f)P(f)$$

Thomas Bayes

- Initial sampling can be generated through design of experiments (DoE)
- The surrogate model is typically a Gaussian Process (GP)
- Next experiment points are generated by acquisition functions (exploration vs. exploitation)

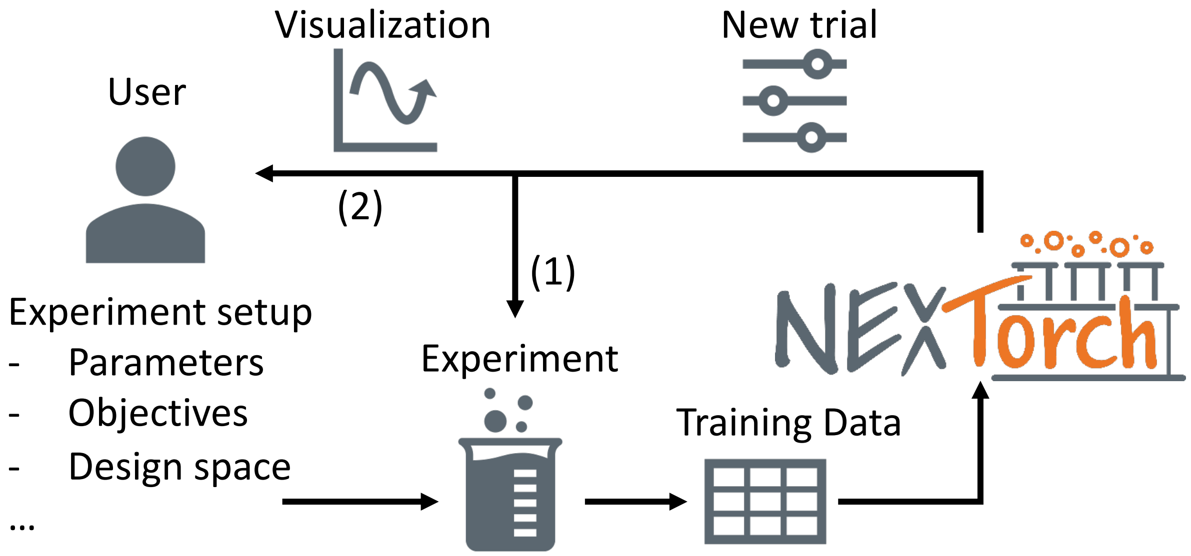
[1] P.I. Frazier, 1 (2018). http://arxiv.org/abs/1807.02811

NEXTorch

Toolkit for Design of Experiments + Bayesian Optimization

Workflow

(1) Automated (2) Human-in-the-loop optimization



- Install via pip `pip install nextorch`
- Key dependencies

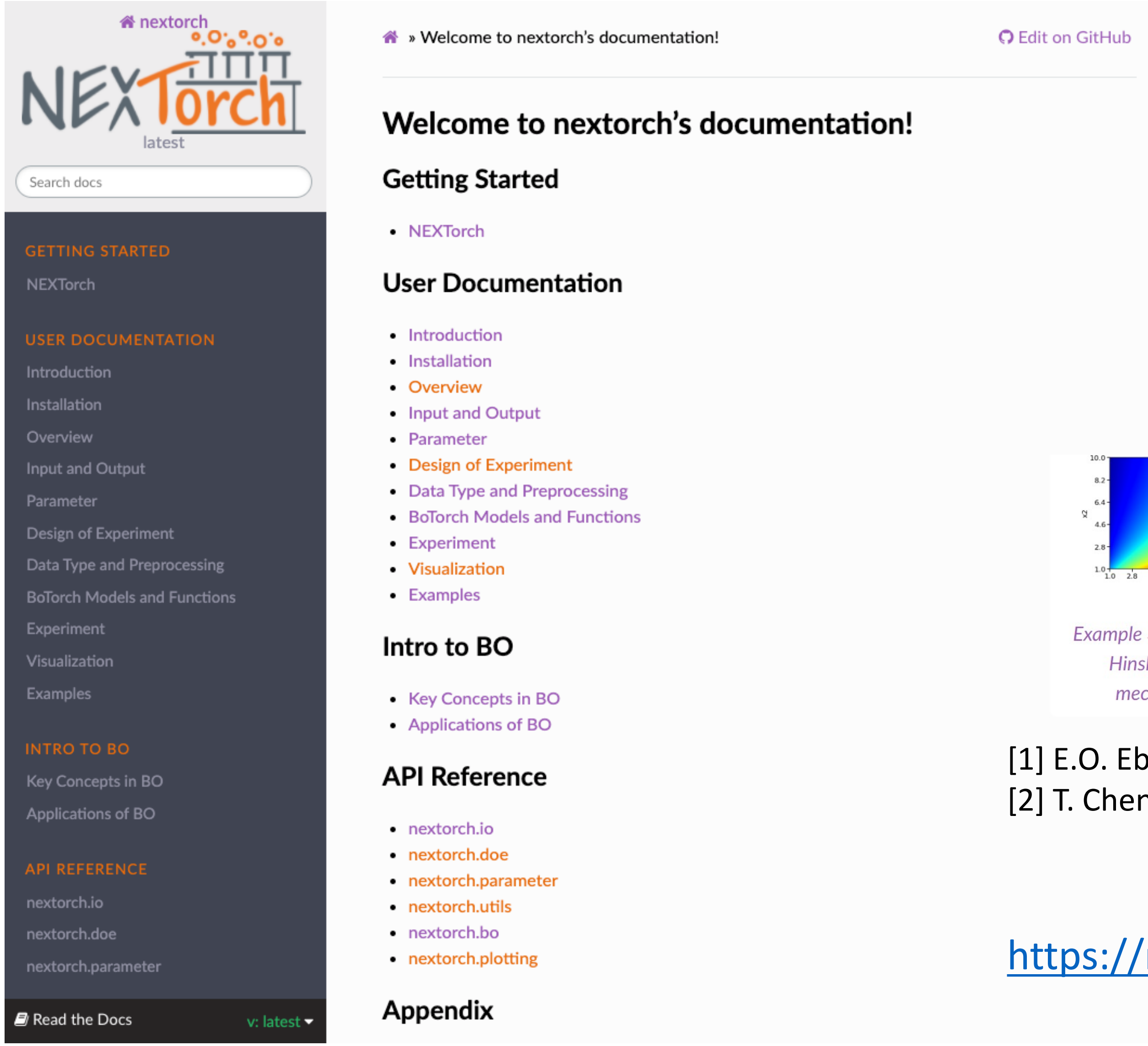
python pyDOE PyTorch BoTorch

https://botorch.org/

[1] **Y. Wang**, T. Chen, and D.G. Vlachos, J. Chem. Inf. Model. 61, 5312–5319 (2021).
GitHub: https://github.com/VlachosGroup/nextorch
Documentation: https://nextorch.readthedocs.io/en/latest/index.html

CATALYSIS CENTER FOR ENERGY INNOVATION

Online Documentation Page


nextorch

latest

Search docs

GETTING STARTED

NEXTorch

USER DOCUMENTATION

Introduction
Installation
Overview
Input and Output
Parameter
Design of Experiment
Data Type and Preprocessing
BoTorch Models and Functions
Experiment
Visualization
Examples

INTRO TO BO

Key Concepts in BO
Applications of BO

API REFERENCE

nextorch.io
nextorch.doe
nextorch.parameter

Read the Docs v: latest

» Welcome to nextorch's documentation! Edit on GitHub

Welcome to nextorch's documentation!

Getting Started

- NEXTorch

User Documentation

- Introduction
- Installation
- Overview
- Input and Output
- Parameter
- Design of Experiment
- Data Type and Preprocessing
- BoTorch Models and Functions
- Experiment
- Visualization
- Examples

Intro to BO

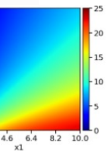
- Key Concepts in BO
- Applications of BO

API Reference

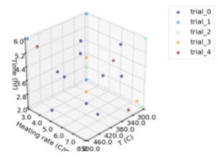
- nextorch.io
- nextorch.doe
- nextorch.parameter
- nextorch.utils
- nextorch.bo
- nextorch.plotting

Appendix

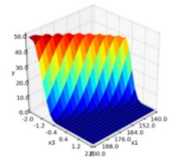
- NEXTorch modules and functions
- Tutorials with code examples
- Introduction to BO theory
- BO applications in literature



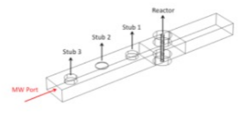
Example 3 - Langmuir Hinshelwood mechanism



Example 4 - Nitrogen-doped carbon catalysts



Example 5 - Plug flow reactor yield



Example 8 - Stub tuner of the microwave cavity

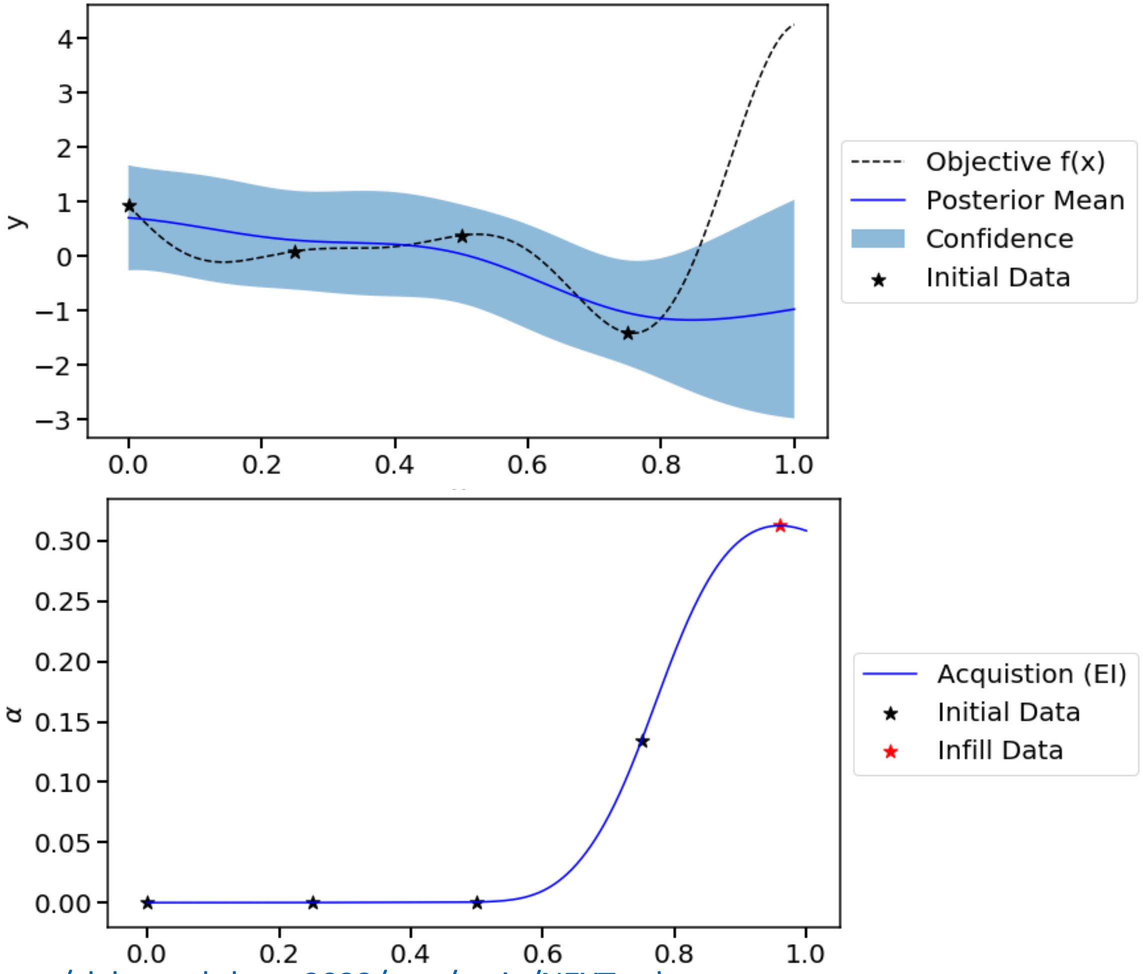
[1] E.O. Ebikade, et al, React. Chem. Eng. (2020).
[2] T. Chen, et al, Ind. Eng. Chem. Res. 59, 10418 (2020).

https://nextorch.readthedocs.io/en/latest/index.html

CATALYSIS CENTER FOR ENERGY INNOVATION

BO Examples in 1D

- Find the minima of $f(x) = (6x-2)^2 sin(12x-4); x \in [0,1]$
- Starting from $x = 0, 0.25, 0.5, 0.75$



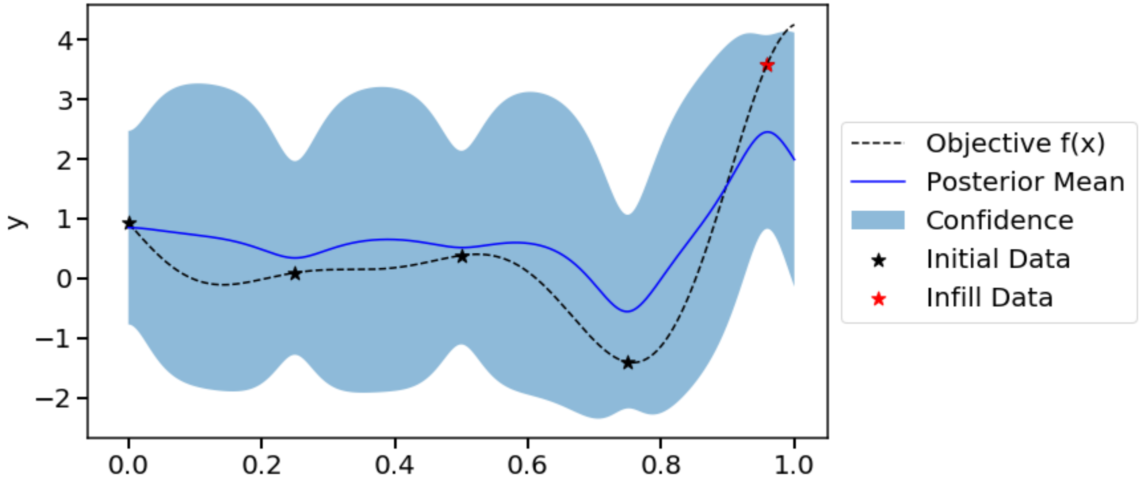
Demo instructions: https://github.com/VlachosGroup/vlab_workshop_2023/tree/main/NEXTorch

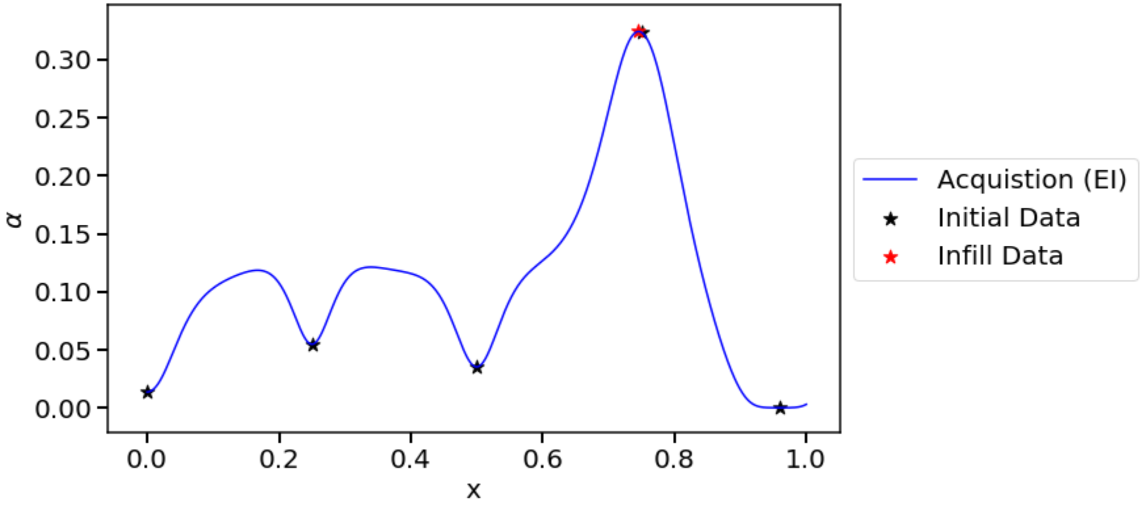
Notebook location: https://github.com/VlachosGroup/nextorch/blob/main/examples/notebooks/01_simple_1d.ipynb

CATALYSIS CENTER FOR ENERGY INNOVATION

BO Examples in 1D

- Find the minima of $f(x) = (6x - 2)^2 sin(12x - 4); x \in [0,1]$
- Iteration 1, acquisition function – expected improvement (EI)

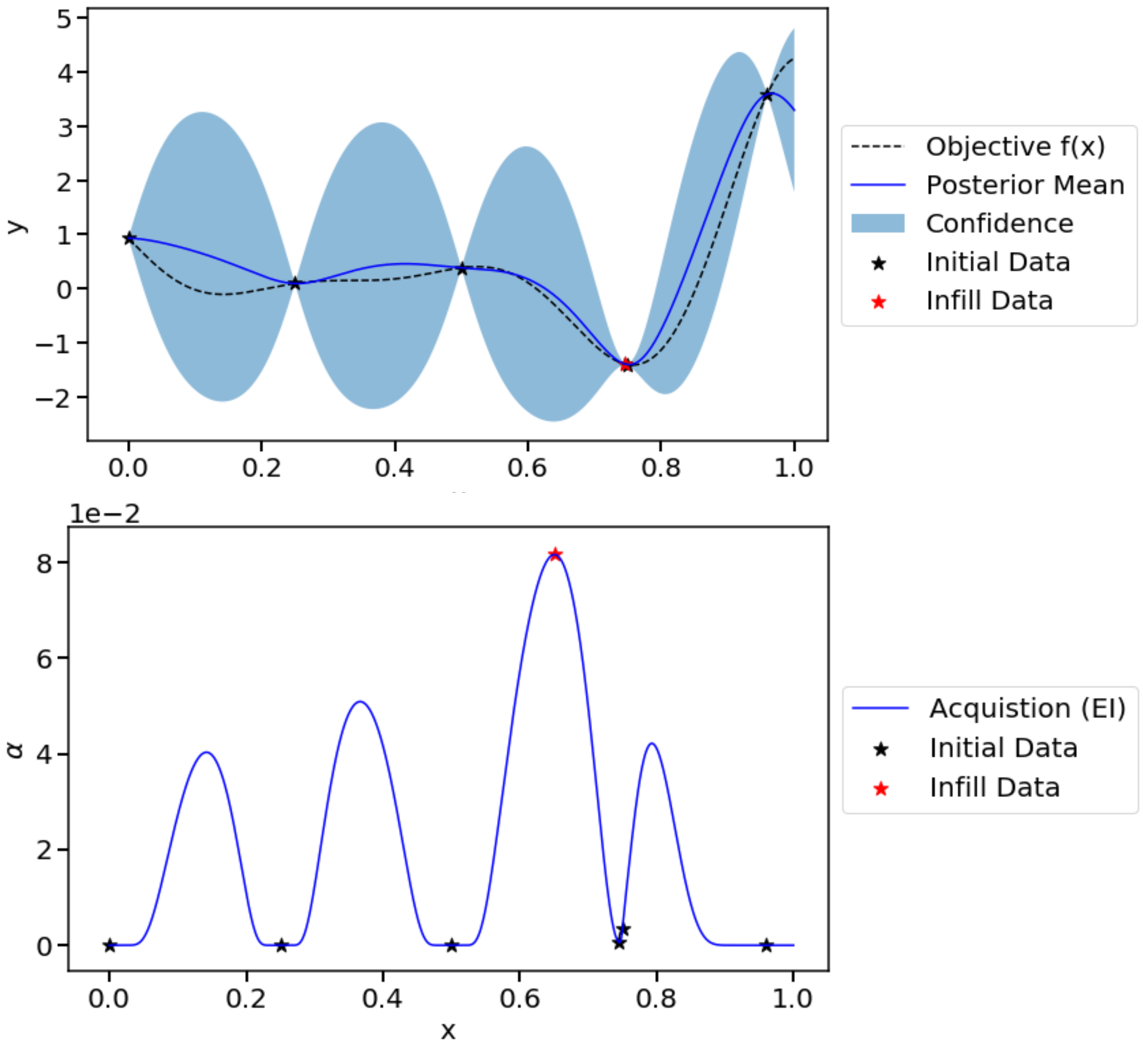




CATALYSIS CENTER FOR ENERGY INNOVATION

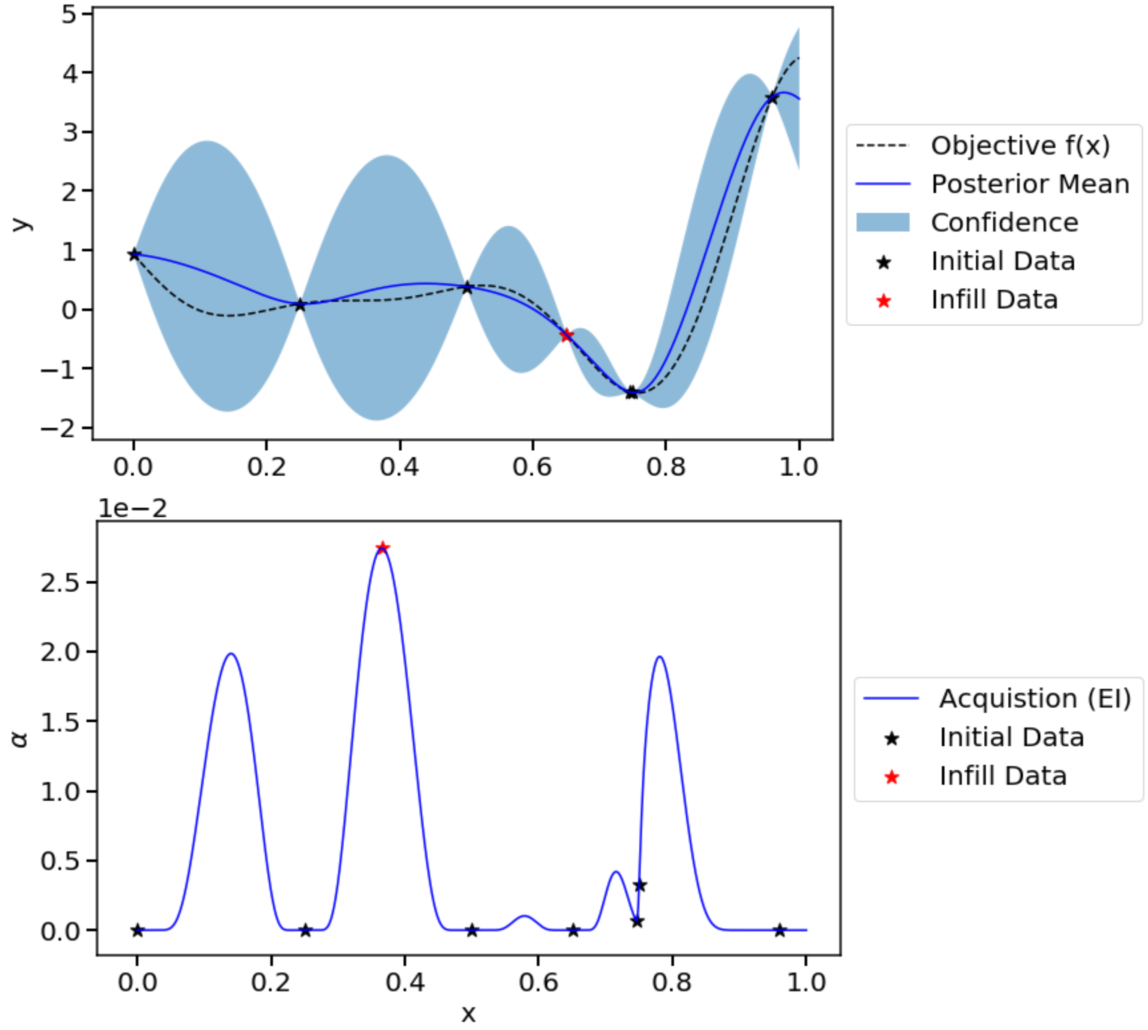
BO Examples in 1D

- Find the minima of $f(x) = (6x - 2)^2 sin(12x - 4); x \in [0,1]$
- Iteration 2



BO Examples in 1D

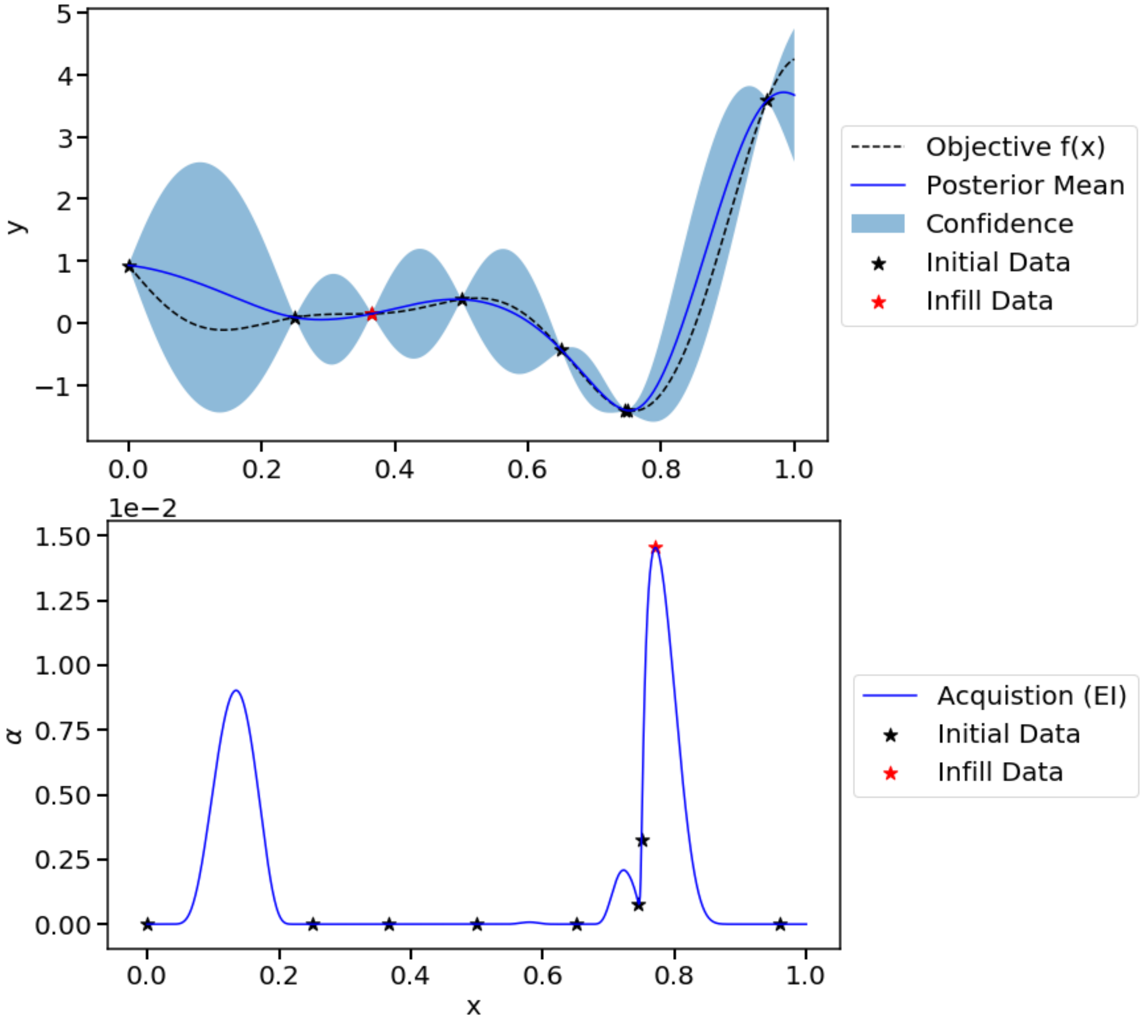
- Find the minima of $f(x) = (6x-2)^2 sin(12x-4); x \in [0,1]$
- Iteration 3



CATALYSIS CENTER FOR ENERGY INNOVATION

BO Examples in 1D

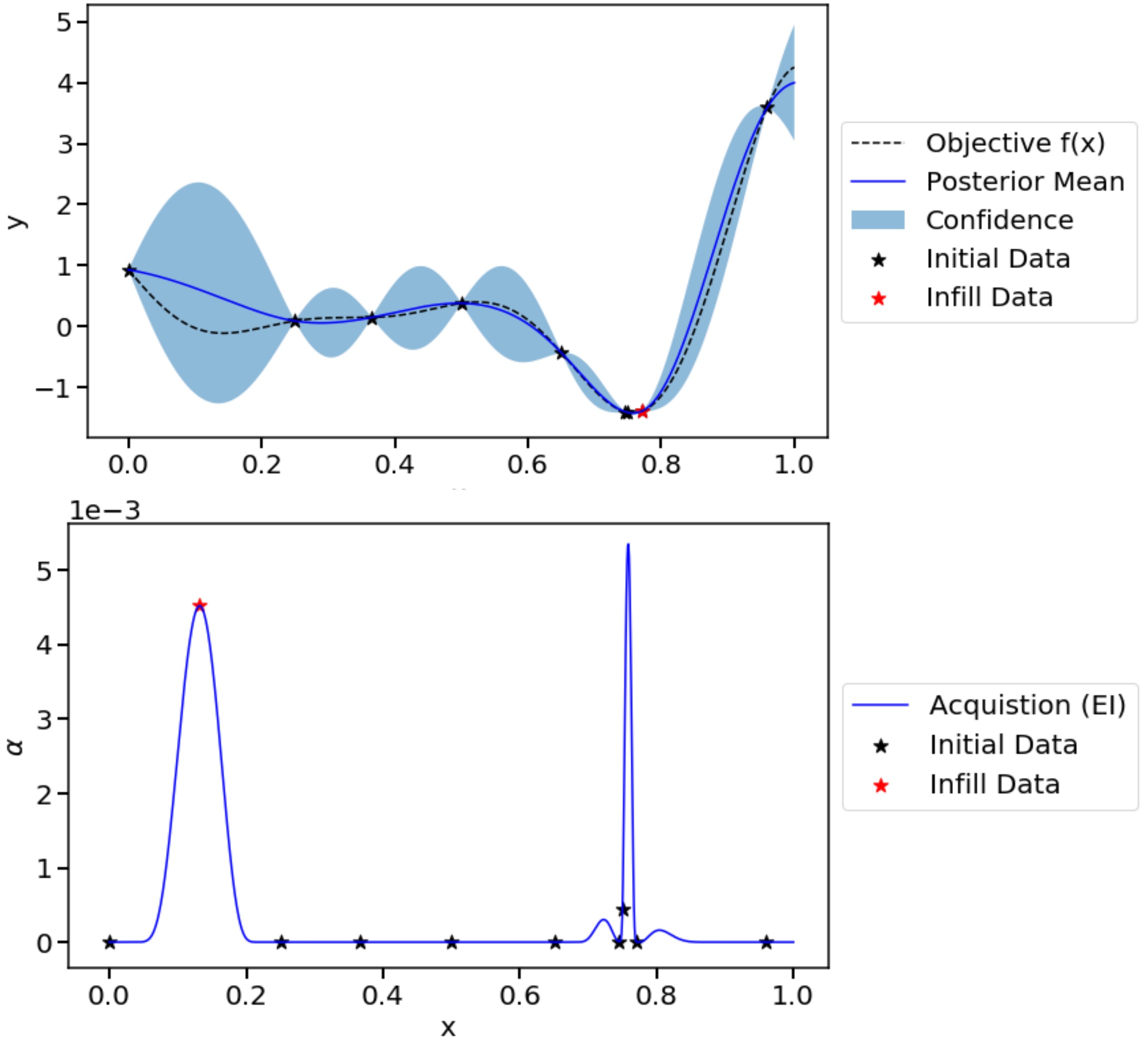
- Find the minima of $f(x) = (6x-2)^2 sin(12x-4); x \in [0,1]$
- Iteration 4



CATALYSIS CENTER FOR ENERGY INNOVATION

BO Examples in 1D

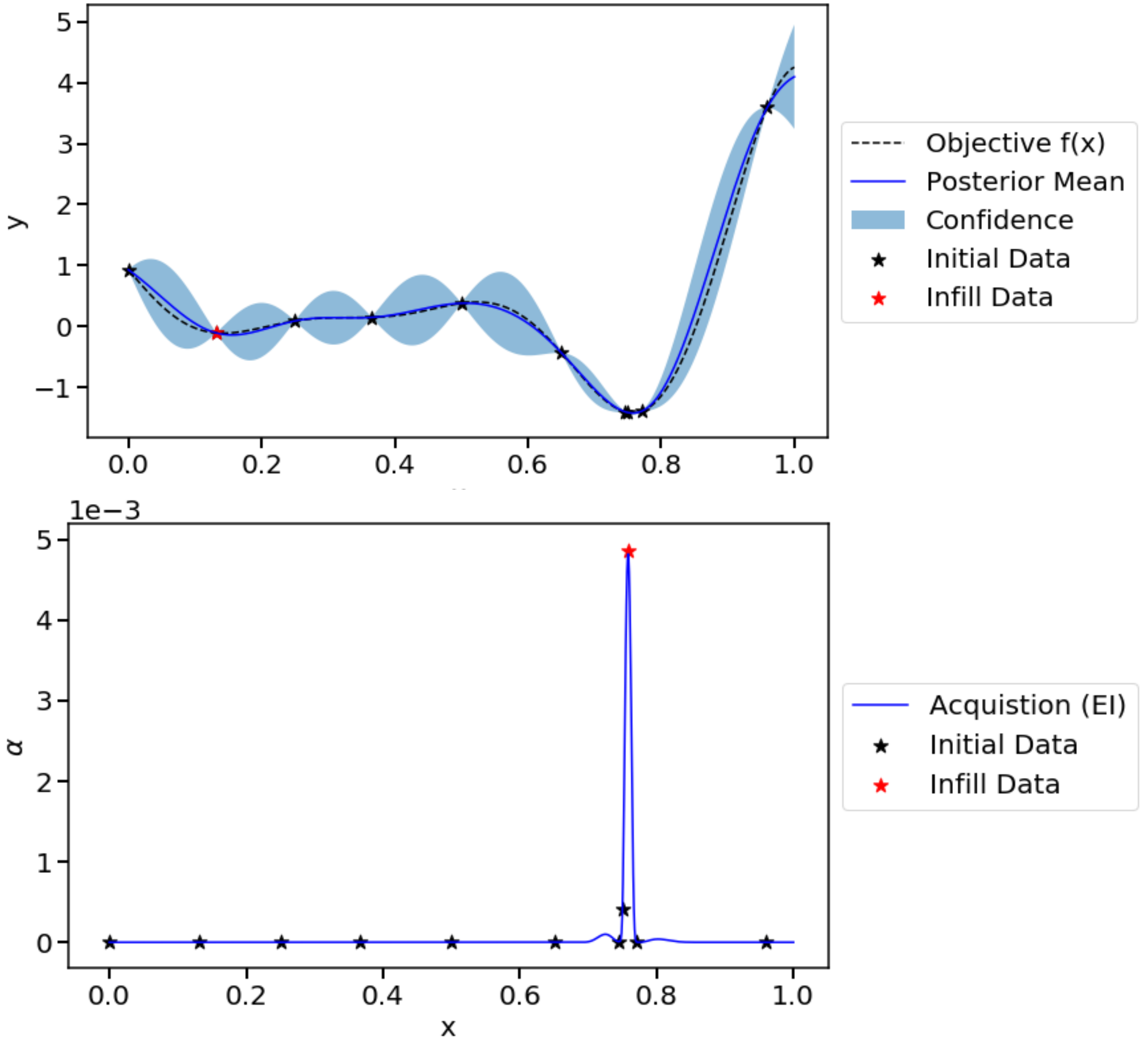
- Find the minima of $f(x) = (6x-2)^2 sin(12x-4); x \in [0,1]$
- Iteration 5



CATALYSIS CENTER FOR ENERGY INNOVATION

BO Examples in 1D

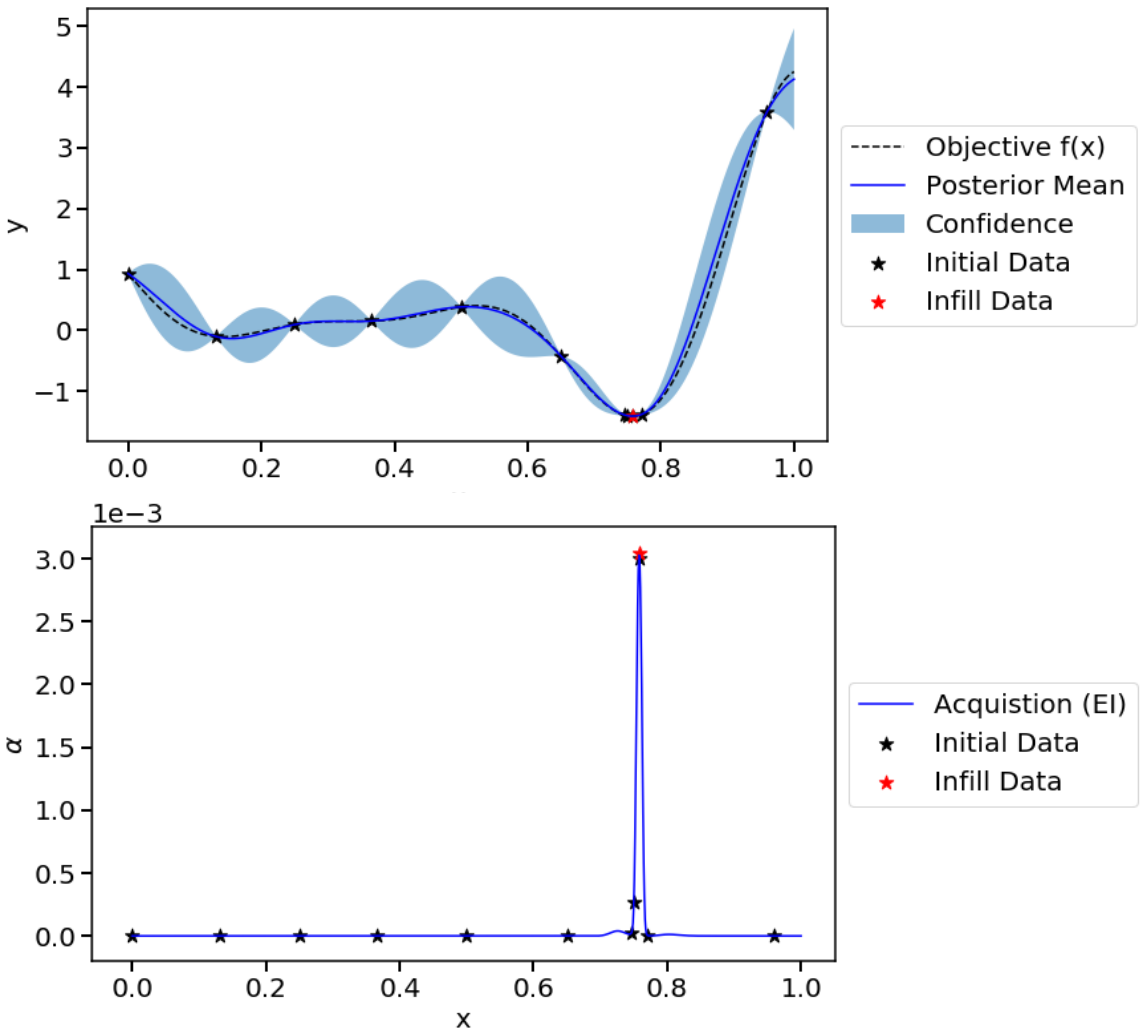
- Find the minima of $f(x) = (6x-2)^2 sin(12x-4); x \in [0,1]$
- Iteration 6



CATALYSIS CENTER FOR
ENERGY INNOVATION

BO Examples in 1D

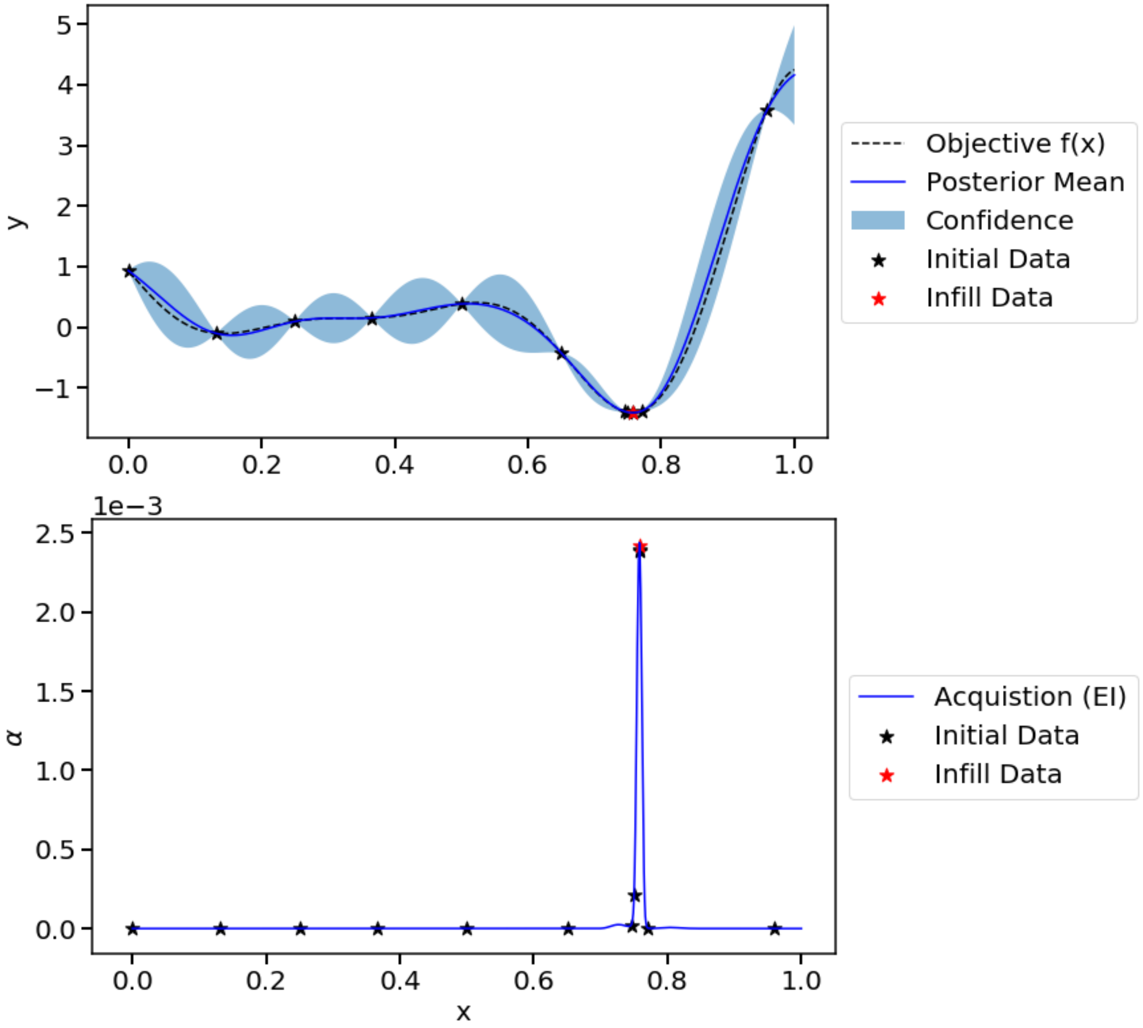
- Find the minima of $f(x) = (6x-2)^2 sin(12x-4); x \in [0,1]$
- Iteration 7



CATALYSIS CENTER FOR ENERGY INNOVATION

BO Examples in 1D

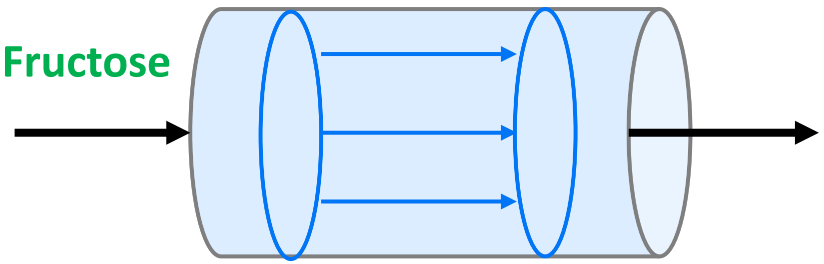
- Find the minima of $f(x) = (6x-2)^2 sin(12x-4); x \in [0,1]$
- Iteration 8



CATALYSIS CENTER FOR ENERGY INNOVATION

Case Study – HMF Yield Optimization

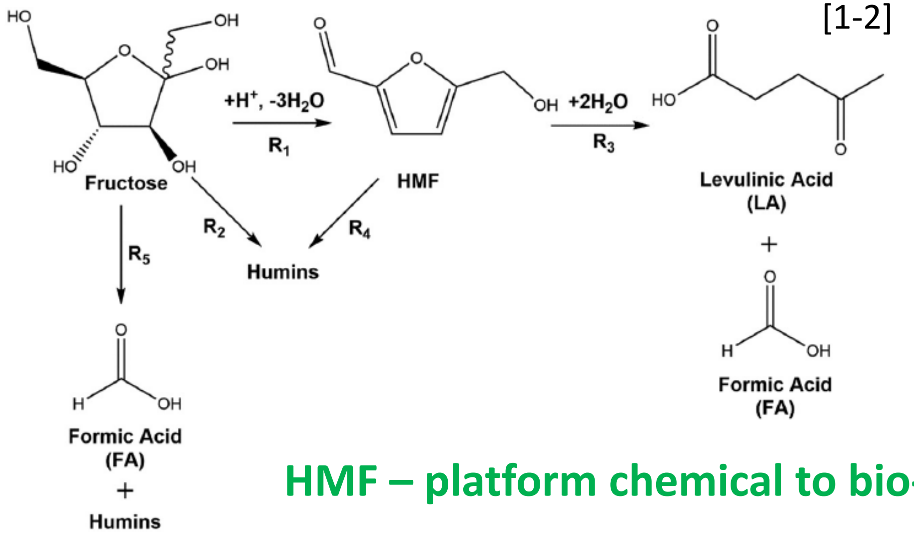
Fructose



$$u\frac{dC_i}{dx} = R_i$$

HMF
LA,
FA,
Fructose,
Humins
…

[1-2]



HMF – platform chemical to bio-products

- Goal: maximize the HMF yield (Y) to improve productivity and reduce downstream costs.
- Three key input parameters ($\boldsymbol{X}$) and bounds:
  - X_1, **Temperature** – (140 – 200 °C)
  - X_2, **pH** – (0 – 1)
  - X_3, **Residence time** – (0.01 – 100 min), sampled in log space (-2,2)

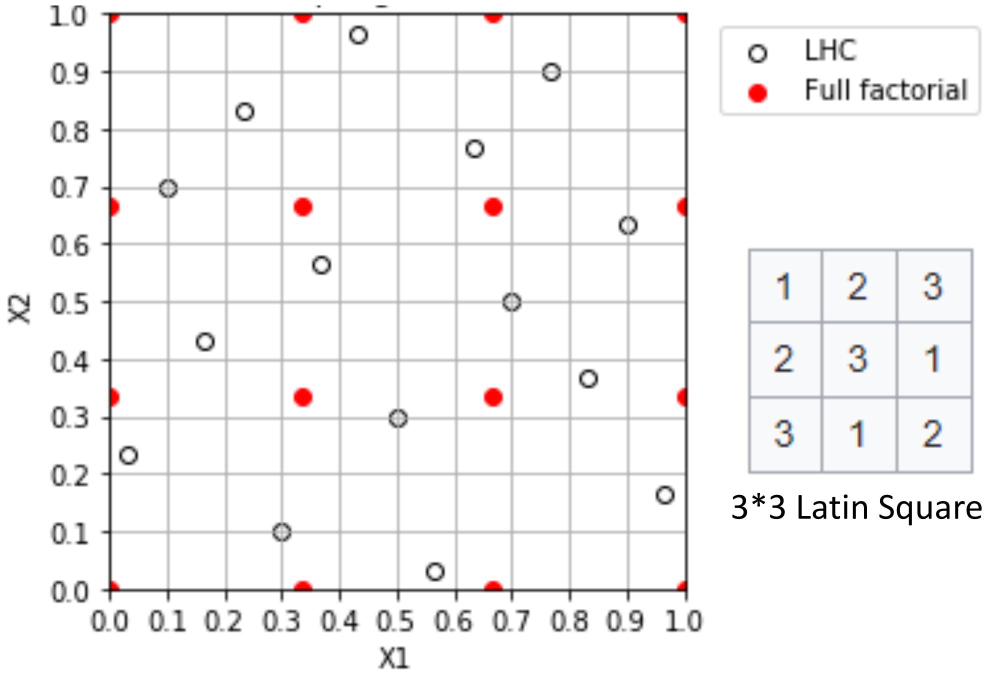
[1] T.D. Swift et al., *ACS Catal.* 4, **2014**, 259.
[2] Desir, P.; Saha, B.; Vlachos, D. G. Energy Environ. Sci. 2019.
Model available in Python and MATLAB: https://github.com/VlachosGroup/Fructose-HMF-Model

Optimized Latin Hypercube design (LHS)

Initial Sampling plan:

- Monte Carlo sampling
- Maximize the distance between points
- Maximize the information gain in the sampling space

2D example (level = 4, 16 samples)



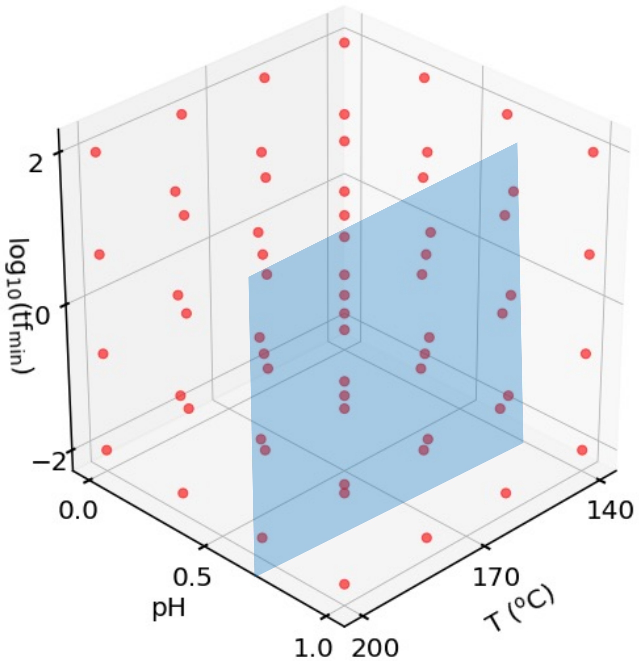
1	2	3
2	3	1
3	1	2

3*3 Latin Square

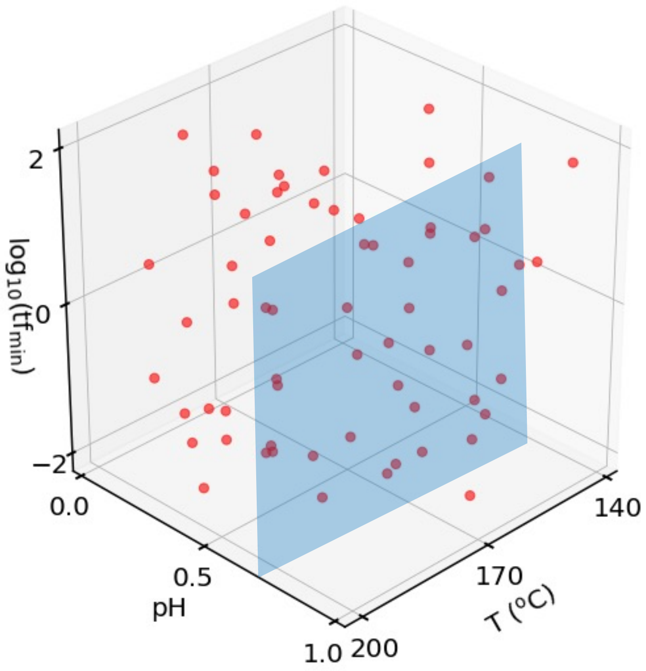
CATALYSIS CENTER FOR ENERGY INNOVATION

Sampling Plans

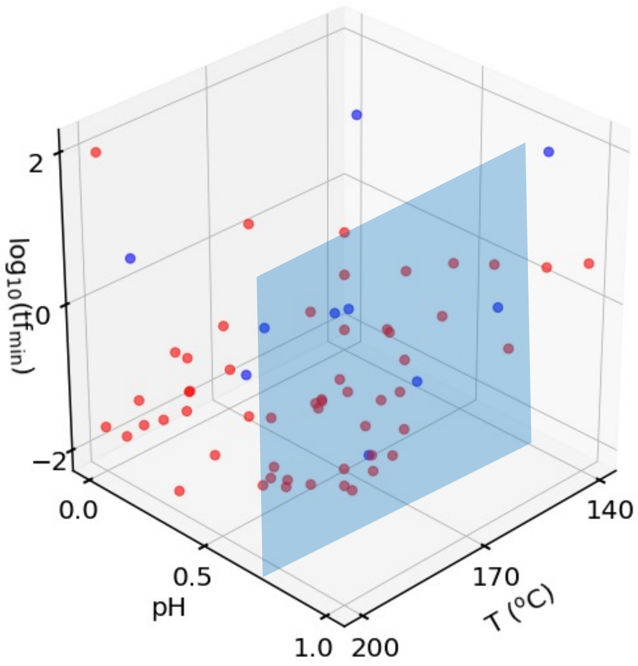
Design 1
Full factorial (level = 4)
64 samples



Design 2
64 random samples

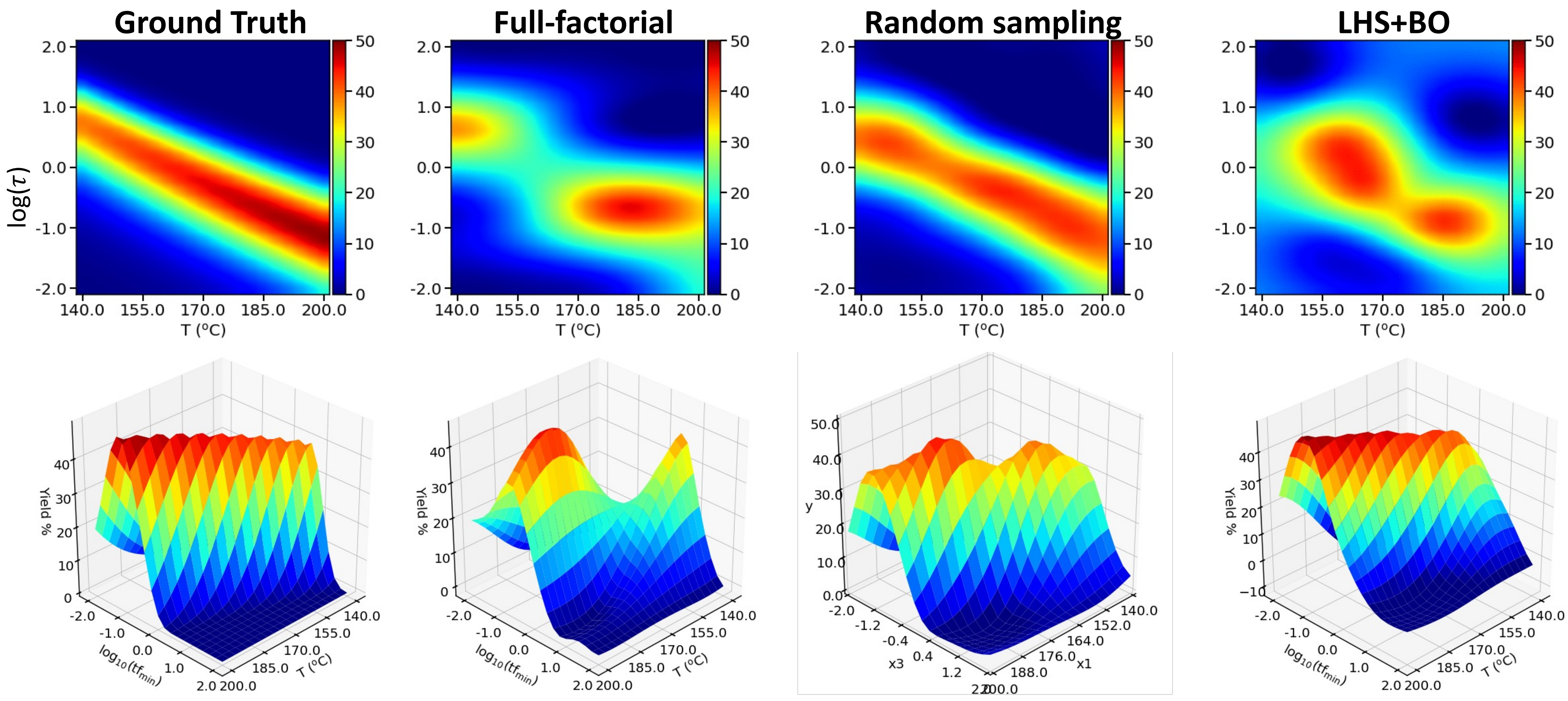


Design 3
10 samples from Latin hypercube (LHS)
54 samples from 54 BO loops,
Acquisition function – EI



- LHS is an efficient space-filling, Monte Carlo sampling method
- We compare response surfaces of HMF yield at **pH=0.7** with varying temperature and residence time

Surrogate Model Performance



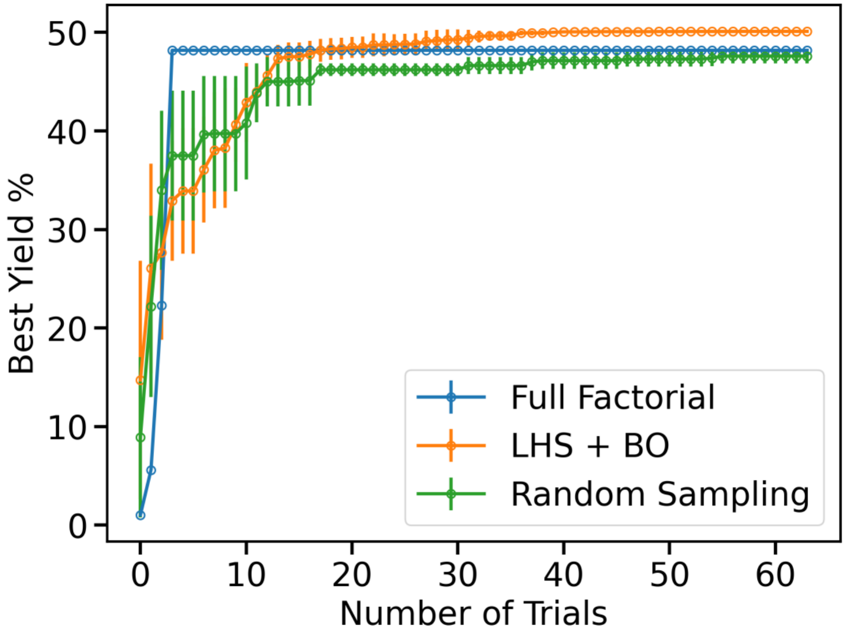
- LHS+BO produces more accurate surrogate models

CATALYSIS CENTER FOR ENERGY INNOVATION

Single-, Multi-Objective Optimization

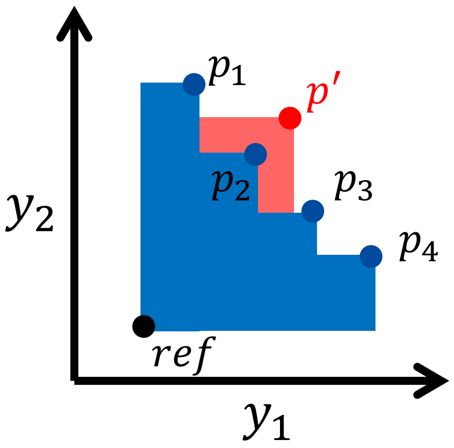
Maximize HMF Yield

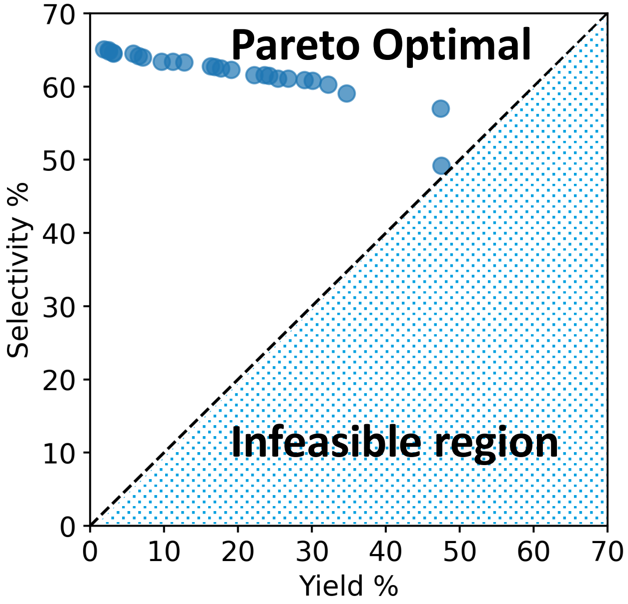
- Optimal condition: Temperature – 200 °C
  pH – 0.705
  Residence time – 0.076 min (4.56 s)



Co-maximize HMF Yield and Selectivity

- HMF Yield = Fructose Conversion × HMF Selectivity
- Fructose Conversion ≤ 100 %



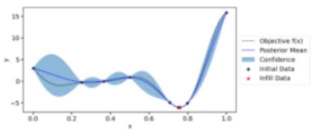


- LHS+BO locates a higher optimal value compared to others
- The runtime of core BO functions completes in seconds per iteration on a laptop CPU
- NEXTorch requires little code and reduces the time or materials for **computations** or **lab experiments**

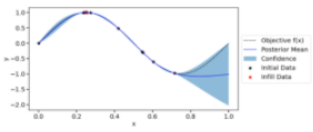
CATALYSIS CENTER FOR ENERGY INNOVATION

More Examples

Basic API Usage

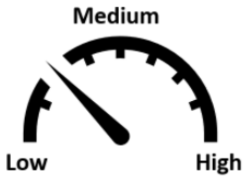


Example 1 - Simple 1d nonlinear function



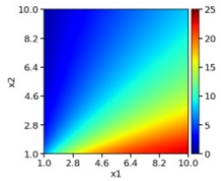
Example 2 - Sin(x) 1d function

Mixed Type Parameters

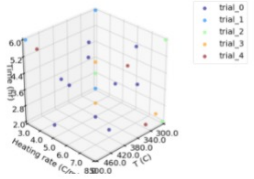


Example 10 - Plug flow reactor yield with mixed type inputs

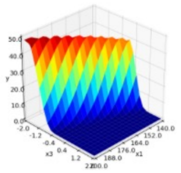
Applications in Reaction Engineering



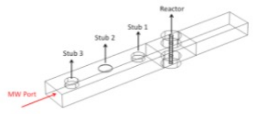
Example 3 - Langmuir Hinshelwood mechanism



Example 4 - Nitrogen-doped carbon catalysts

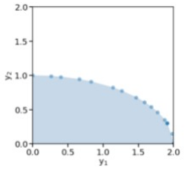


Example 5 - Plug flow reactor yield

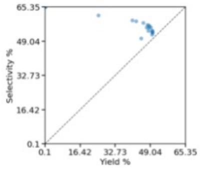


Example 8 - Stub tuner of the microwave cavity

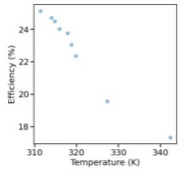
Multi-Objective Optimization(MOO)



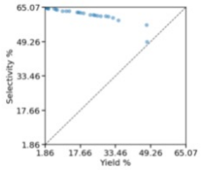
Example 6 - Multi-objective optimization for an ellipse function



Example 7 - Multi-objective optimization for plug flow reactor



Example 9 - Multi-objective optimization for Microwave operating conditions



Example 11 - Multi-objective optimization for plug flow reactor

Documentation: https://nextorch.readthedocs.io/en/latest/index.html