Caltech

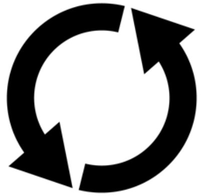
AI for Adaptive Experiment Design

Yisong Yue

Goal-Oriented Experiment Design

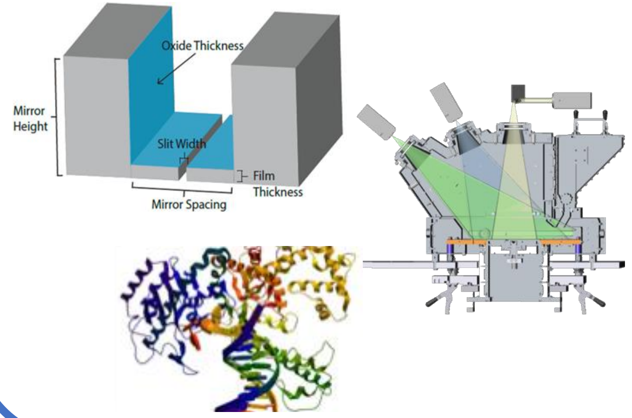
Experiment Designer

Measurement



Experiment

Experiment Platform

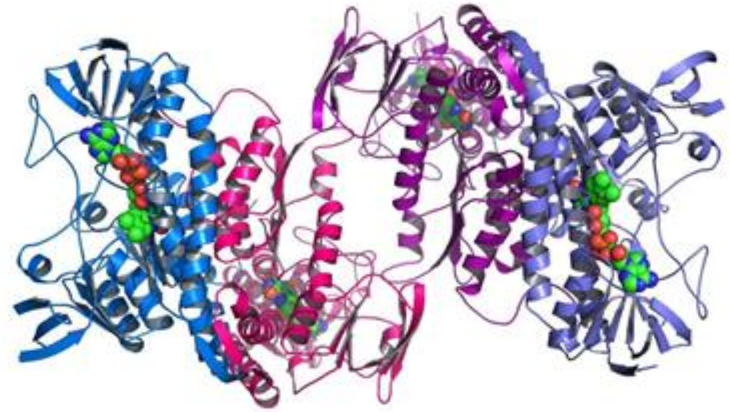


- Iterative & adaptive
- Utility maximizing (find best outcome)

Many Applications (Goal-Oriented)

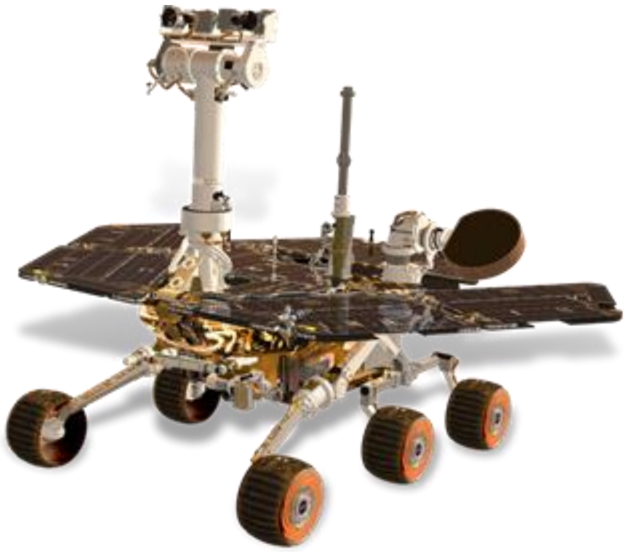
Robotics & Control

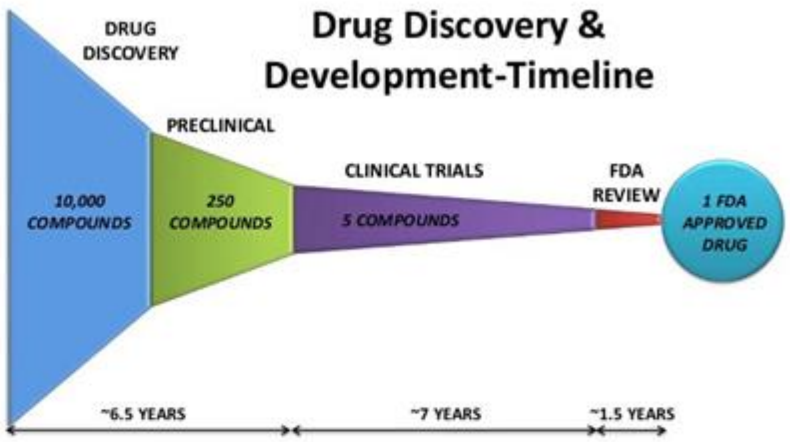
image credit @ mwfarmandfield



Protein Engineering

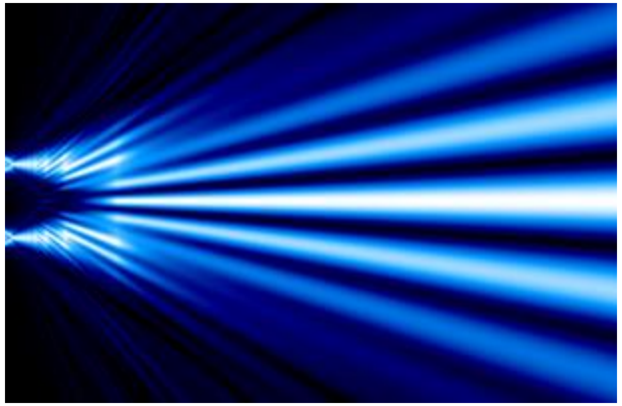
image credit @ creativebiomart





Drug Discovery

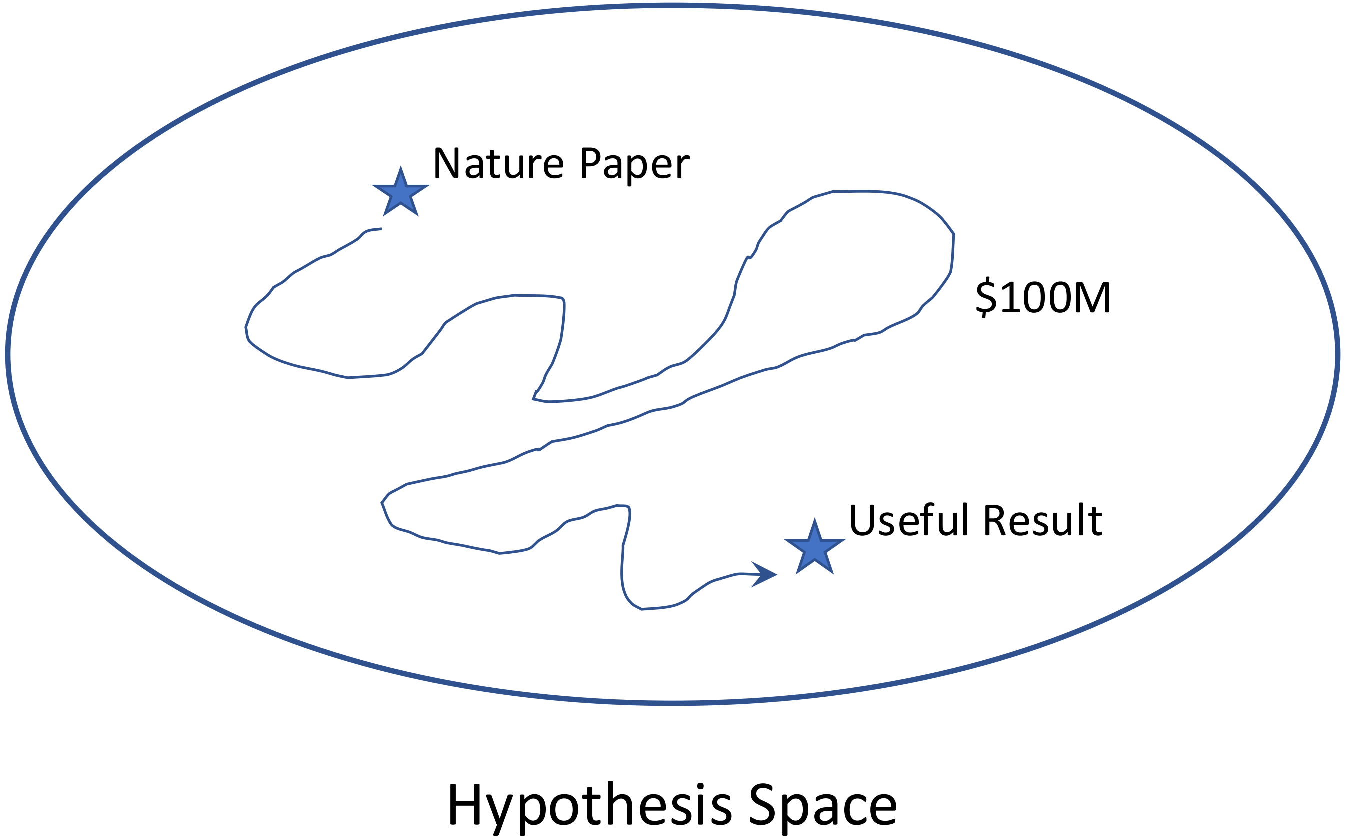
from Slideshare



Material Science

image credit @ phys.org

AI for Goal-Oriented Experiment Design



Batch Supervised Learning

Data: X

Target Signal: Y

$f(x) \approx y$

Data Collected Up Front!

Experiment Design as Interactive Learning

- Collect data on the fly
    - Not available a priori

- Limited budget on data collection

- How to choose?

Three Modes of Interactive Learning

Active Learning

- **Goal:** Discover truth
- E.g., model of world
- Maximize accuracy

(Bayesian) Optimization

- **Goal:** Best single prediction
- E.g., best protein
- Maximize final utility

Multi-Armed Bandits

- **Goal:** Utility over time
- E.g., recommender systems
- Maximize utility over time

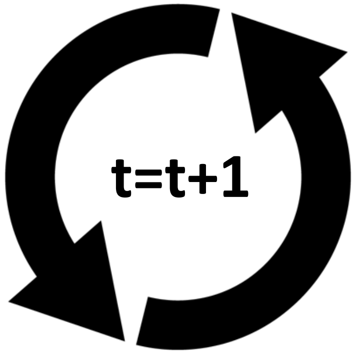
Focus of Talk

Learning Setup

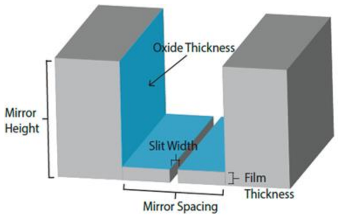
Given: input space X
Unknown: fitness F(x)=y
Maintain: posterior P(F|D) (D=measurements)

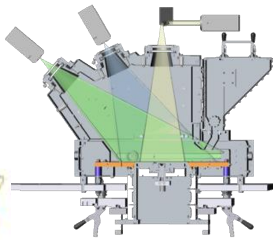
Measure y_t

Experiment Designer



Experiment Platform





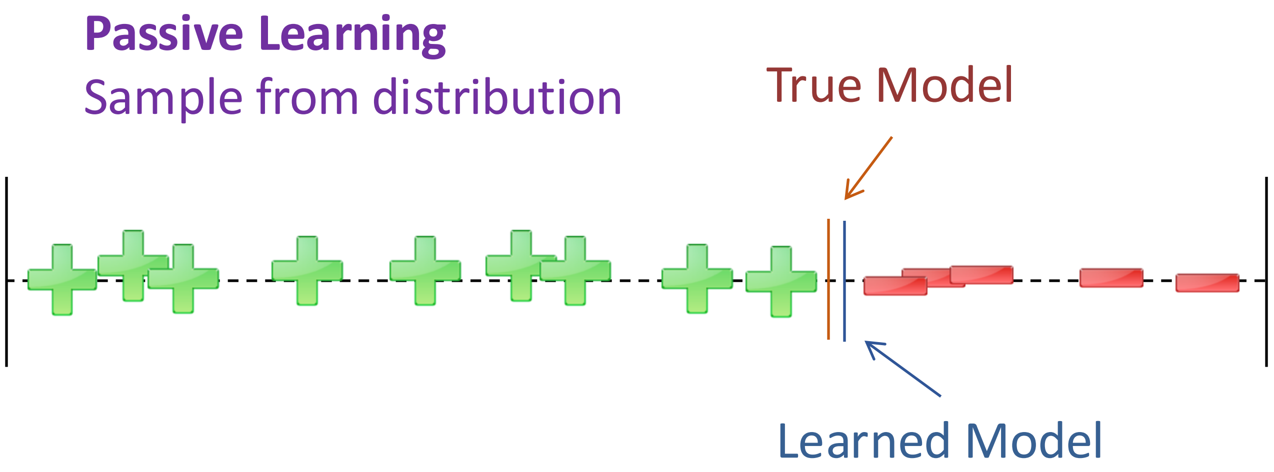
Choose x_t

Add (x_t, y_t) to D
Update posterior P(F|D)

Upper Confidence Bound: $\text{argmax}_x\, \mu(x) + \beta\sigma(x)$
Posterior Sampling: $\text{argmax}_x\, f(x)$, $f \sim P(F|D)$

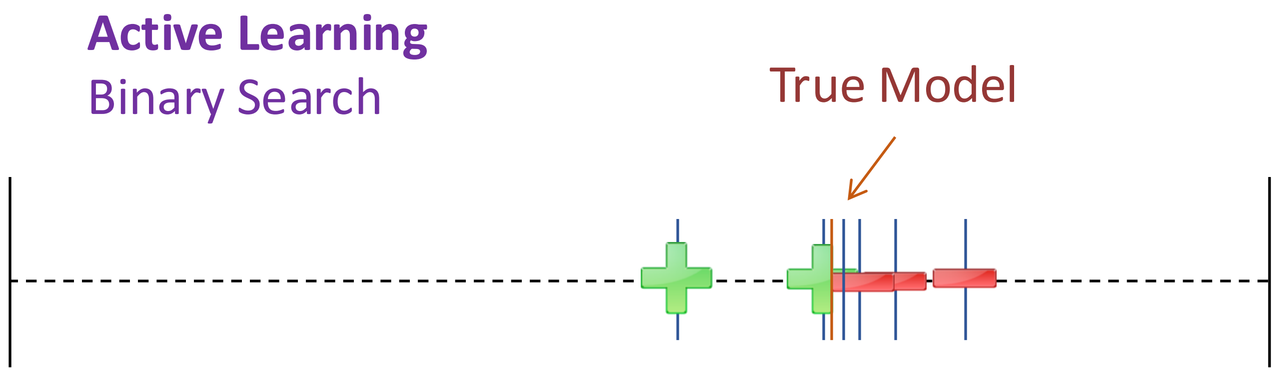
Active Learning Simple Example

- 1 feature
- Learn threshold function



Active Learning Simple Example

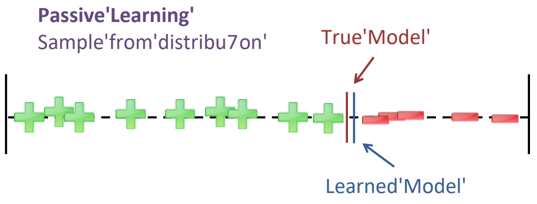
- 1 feature
- Learn threshold function



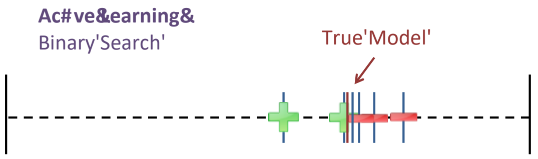
Comparison with Passive/Batch Learning

- # samples to be within ε of true model

- Passive Learning: $O\left(\frac{1}{\varepsilon}\right)$

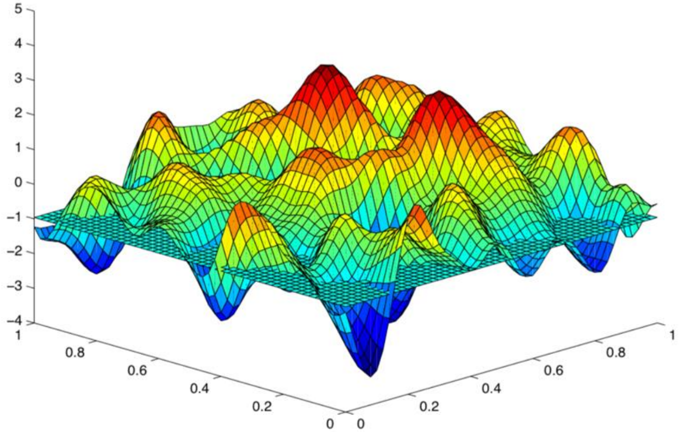


- Active Learning: $O\left(\log \frac{1}{\varepsilon}\right)$



(Bayesian) Optimization Example

X = space of proteins
F(x) = fitness landscape



Frances Arnold

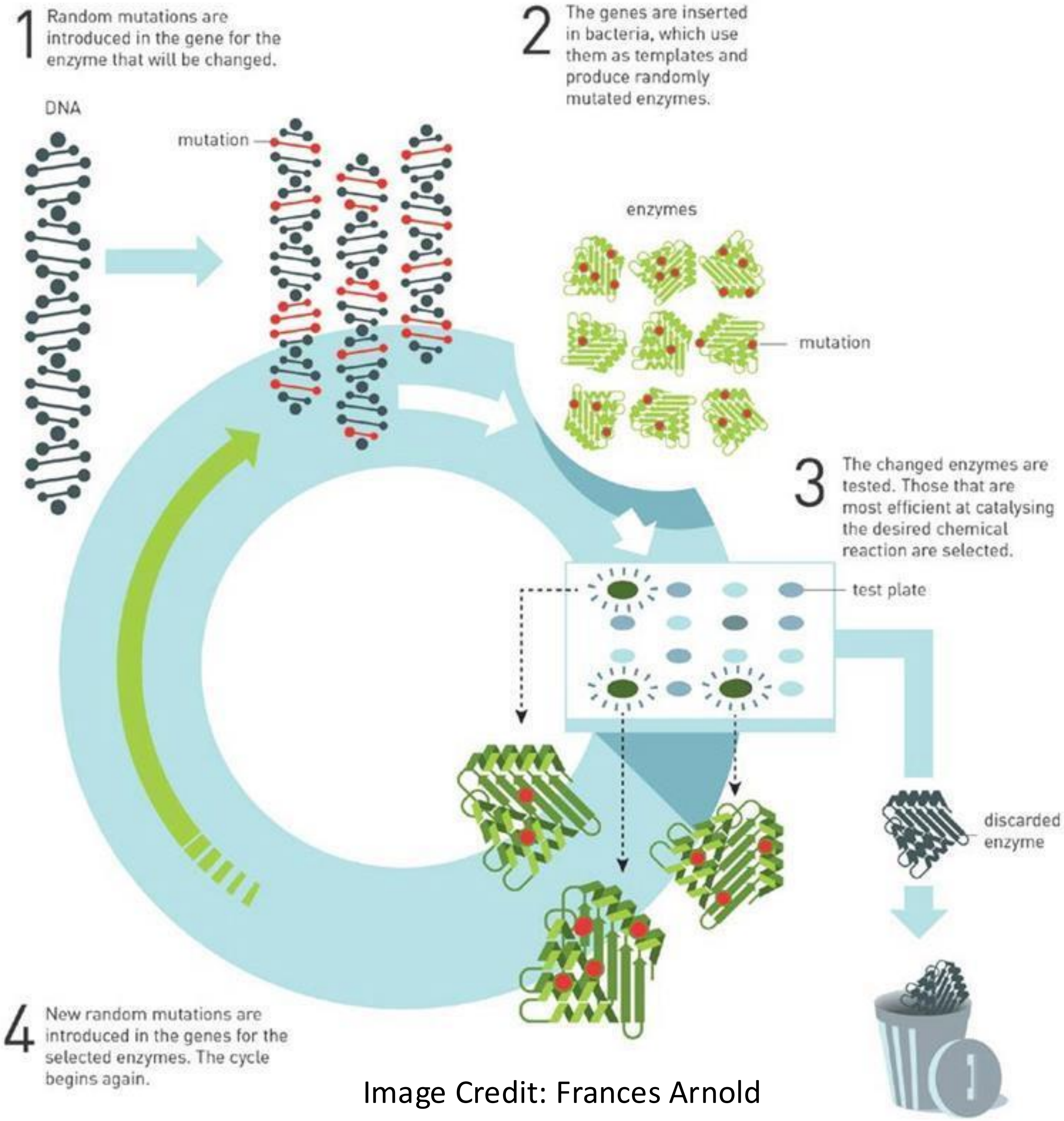
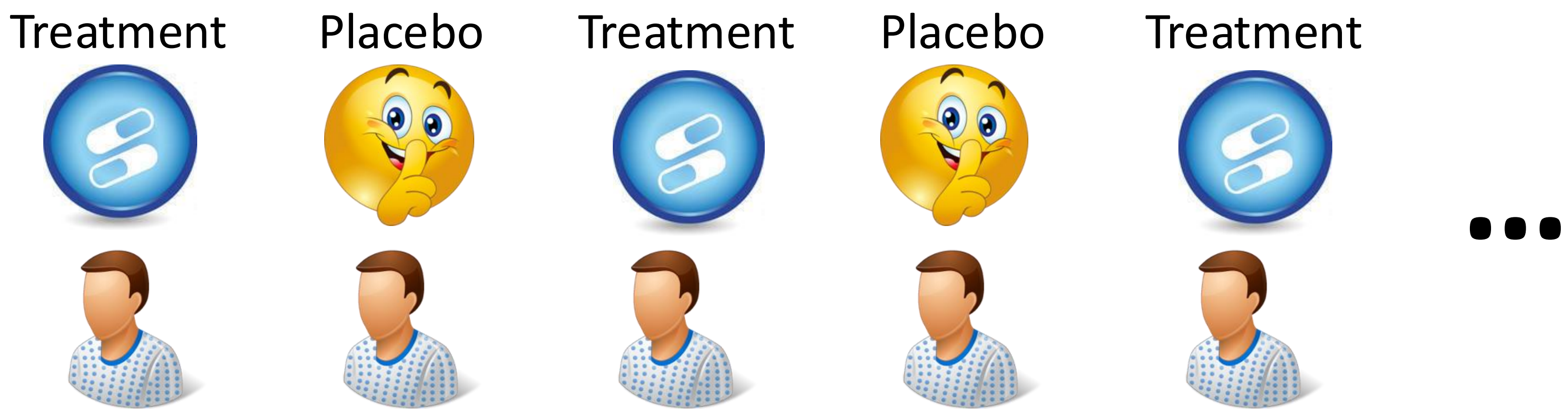


Image Credit: Frances Arnold

Bandits Example

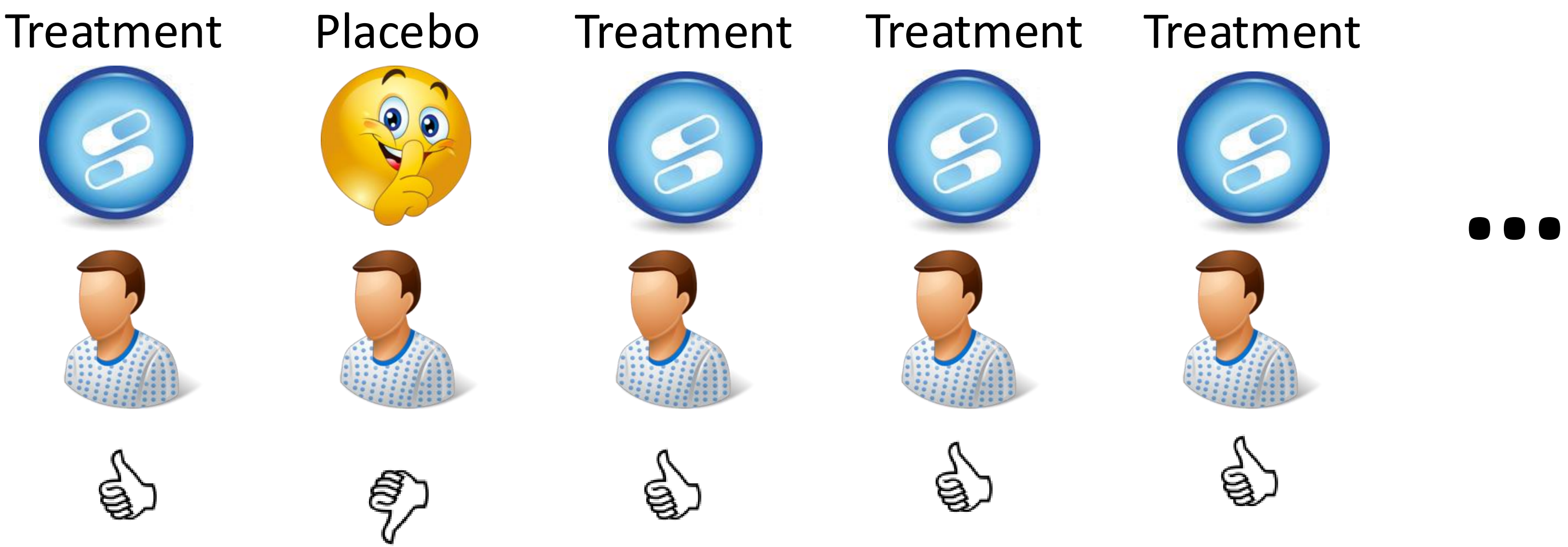
- How to split trials to collect information
- **Static Experimental Design**
  - Standard practice
  - (pre-planned)



http://en.wikipedia.org/wiki/Design_of_experiments

Bandits Example

- Adapt experiments based on outcomes



Each Experiment Matters ("Cumulative Regret")

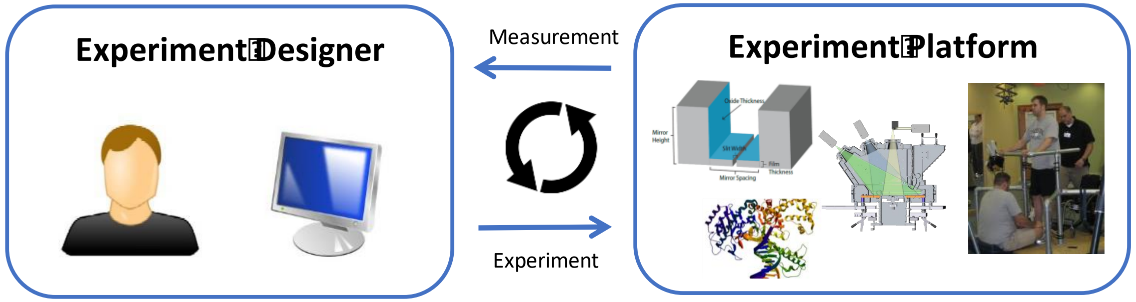
Monica Almeida/The New York Times, left

Two Cousins, Two Paths Thomas McLaughlin, left, was given a promising experimental drug to treat his lethal skin cancer in a medical trial; Brandon Ryan had to go without it.

http://www.nytimes.com/2010/09/19/health/research/19trial.html

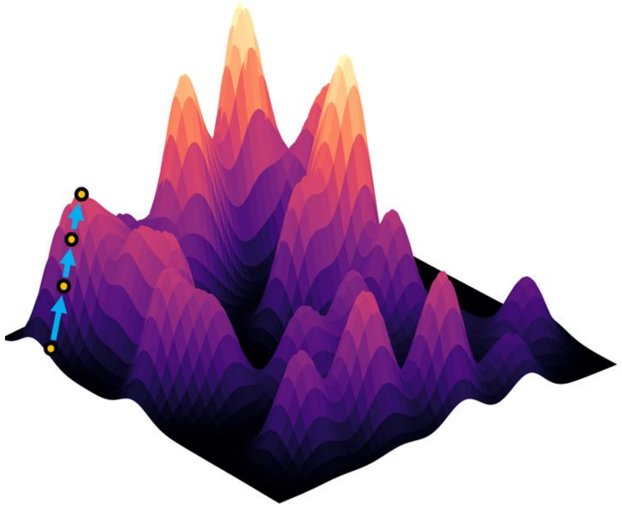
Comparison
(Active Learning, Optimization, Bandits)

- Similarities:
  - Same interaction protocol
  - Query F(x)
  - “Sequential experimental design”

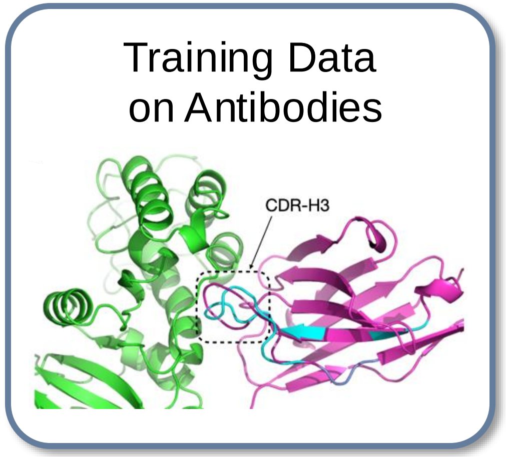


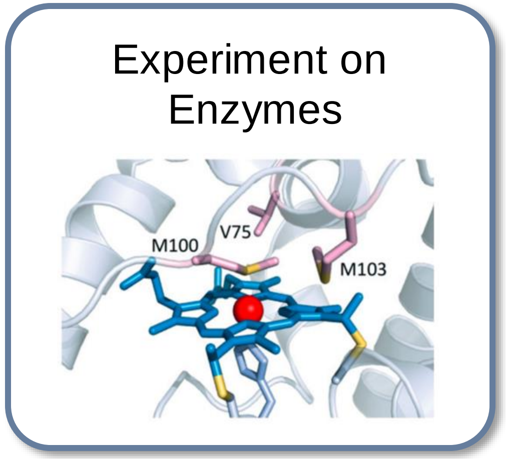
- Active Learning: learn F(x) as accurately as possible — **Discovering “truth”**
- (Bayesian) Optimization: find maximizer F(x*) — **Best result only**
- Bandits: maximize $\sum_i F(x_i)$ — **Each experiment matters**
  “Regret Minimization”

Challenges

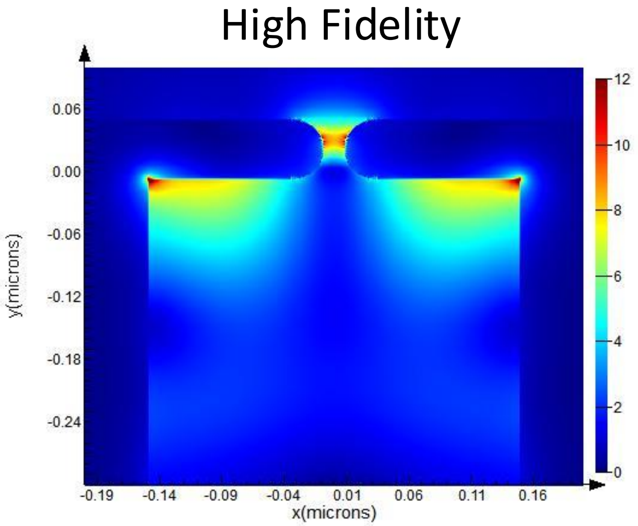


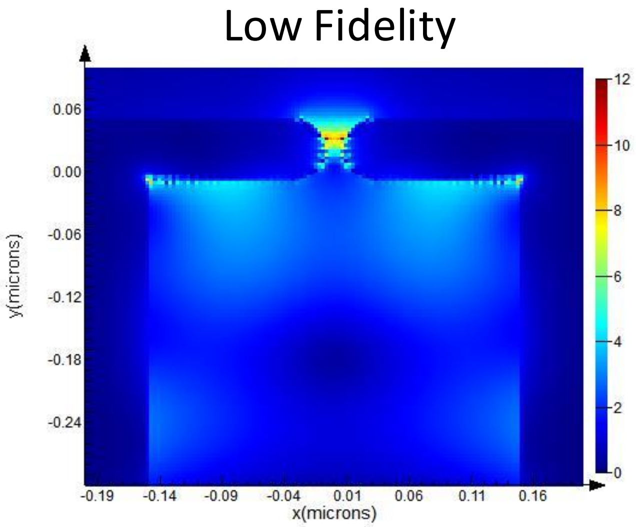
Volatile Utility Landscapes



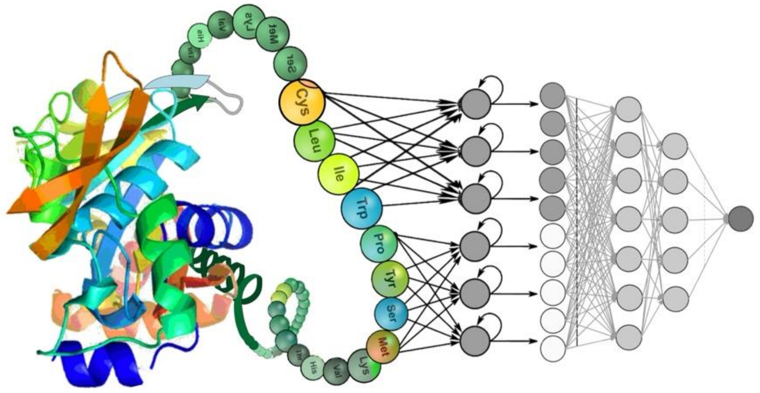


Distribution Shift





Imperfect & Heterogeneous Experiments

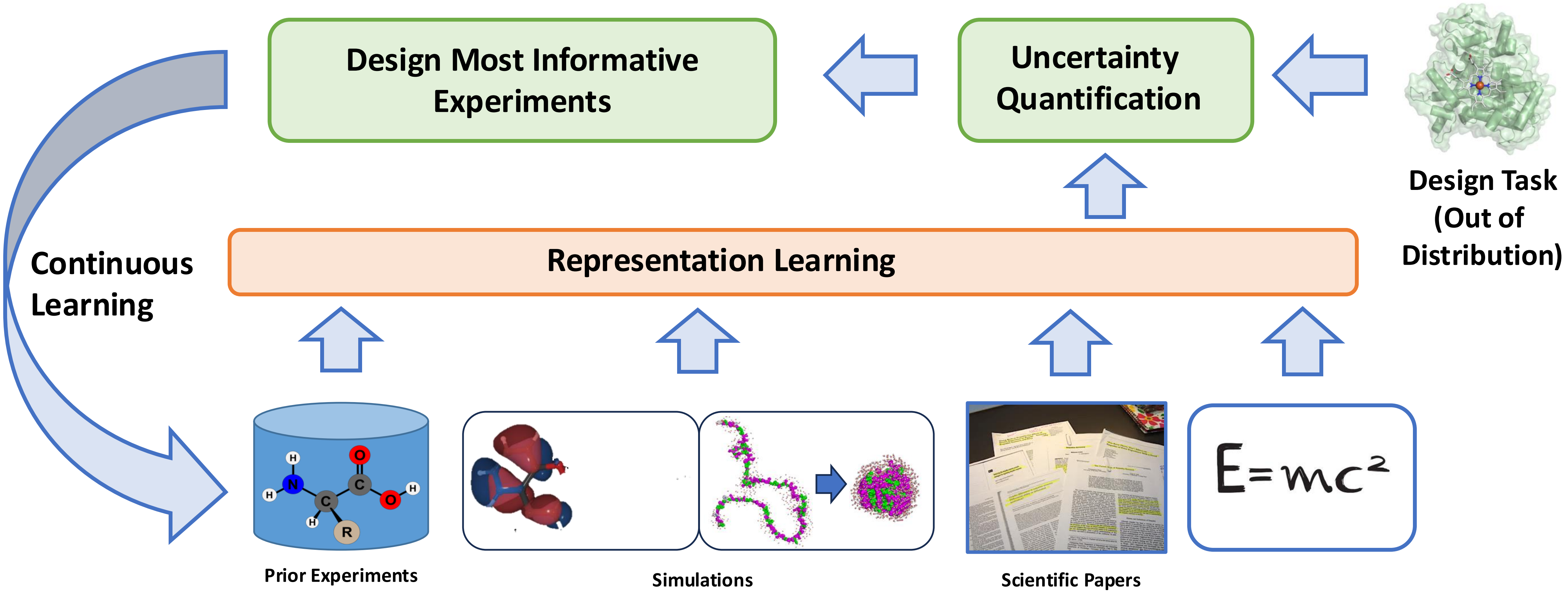


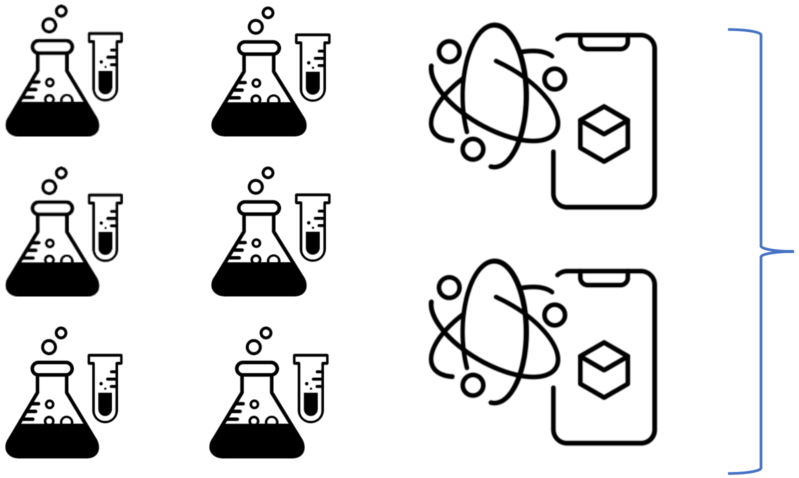
High Dimensionality

https://phys.org/news/2022-04-deep-decode-functional-properties-proteins.html

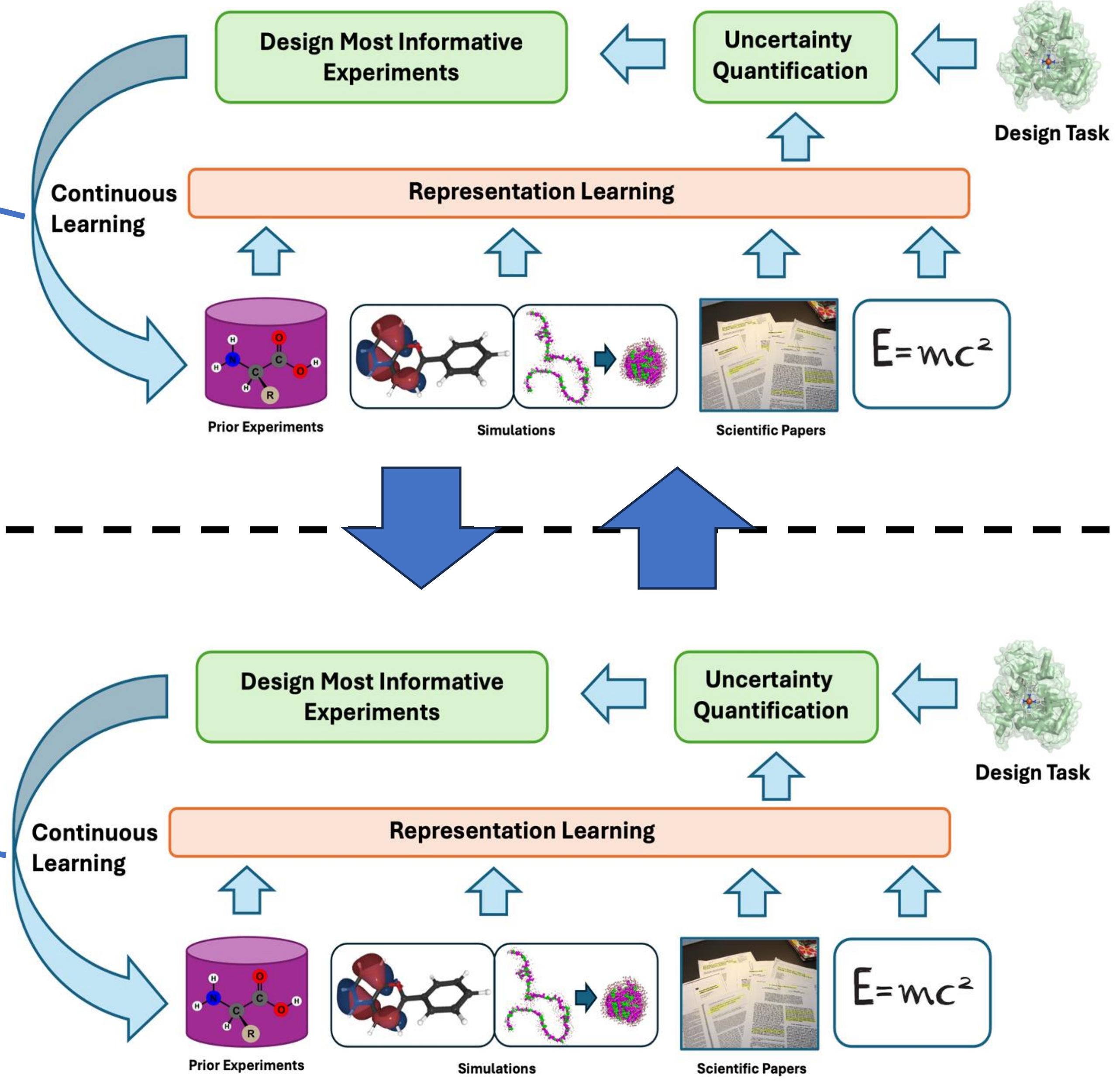
Multi-Objective (Safety & Efficacy)

Real Applications are Complicated!





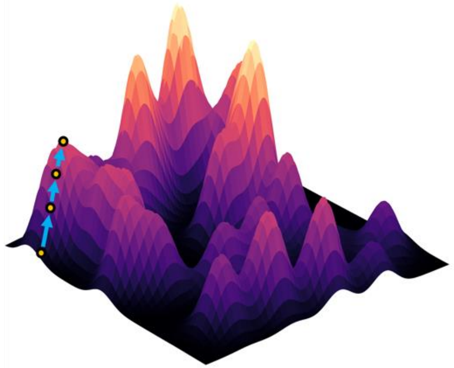
Large & Heterogenous Batched Experiments



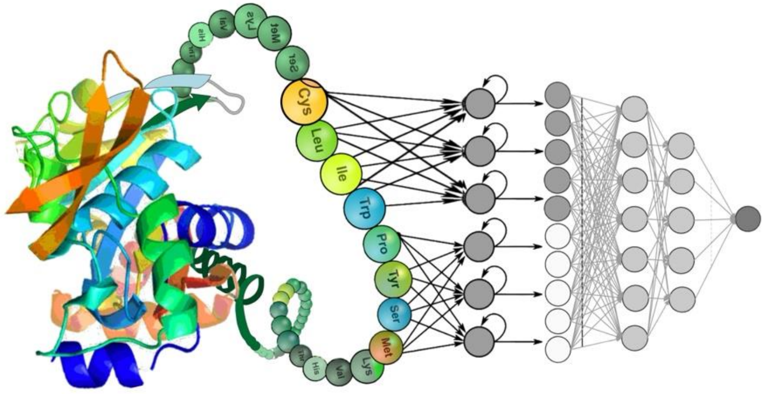
New constraints & complexity

Icon image src: https://uxwing.com/lab-icon/ https://thenounproject.com/icon/physics-simulator-3127977/

Basic Recipe

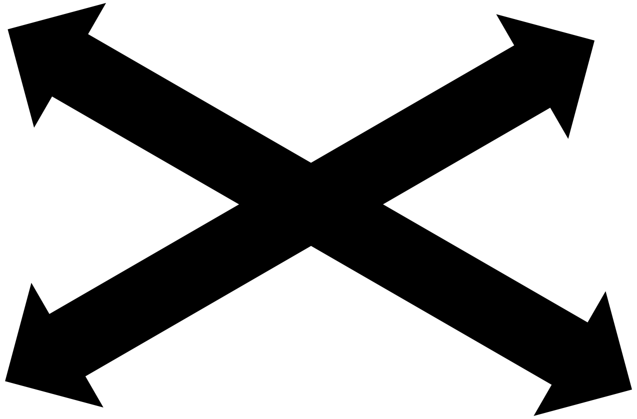


Choose Objective
(can be multiple)

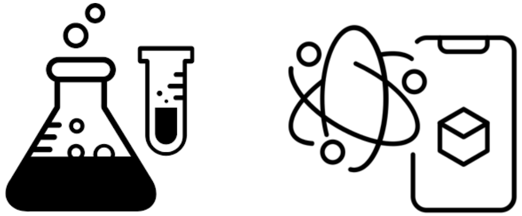


Choose Representation

https://phys.org/news/2022-04-deep-decode-functional-properties-proteins.html



Interdependent!

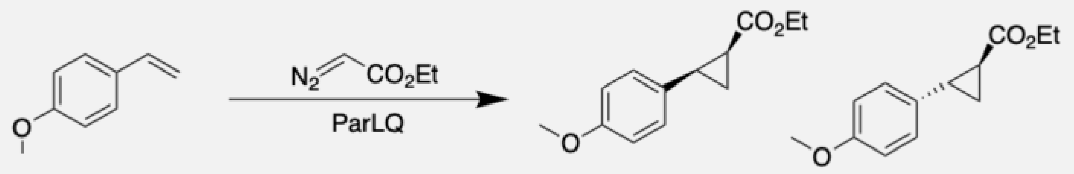


Choose Experiment Platforms

Choose
Algorithm

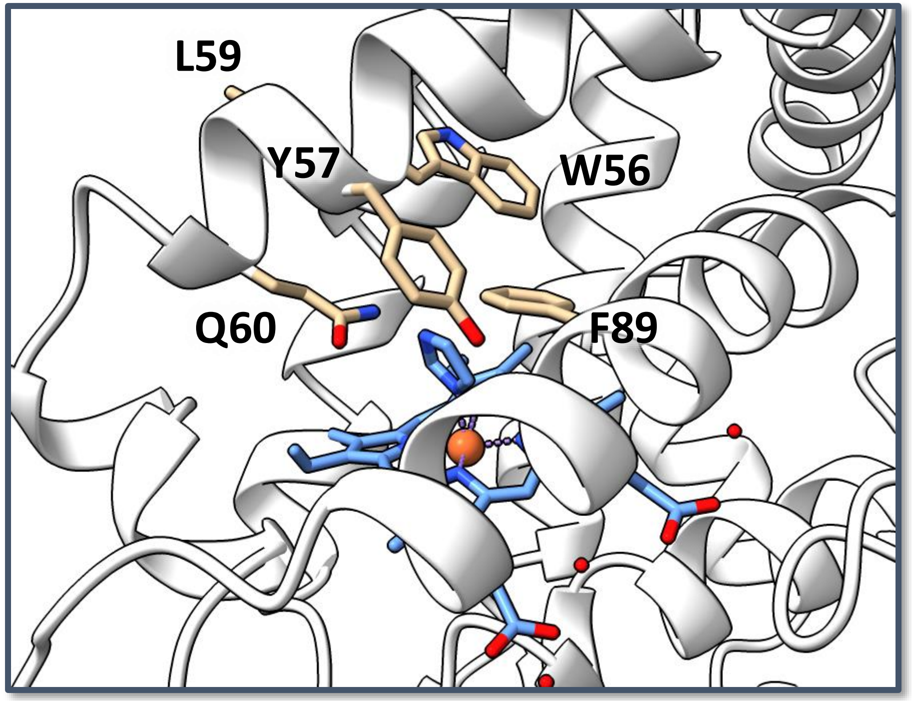
Icon image src: https://uxwing.com/lab-icon/ https://thenounproject.com/icon/physics-simulator-3127977/

Example: Engineering Active Site of Protoglobin



1
cis product

2
trans product



***Par*Pgb-LQ**

Jason Yang

Design Problem:
find the optimal set of mutations on 5 sites
for high yield and selectivity for **1**, the *cis* product

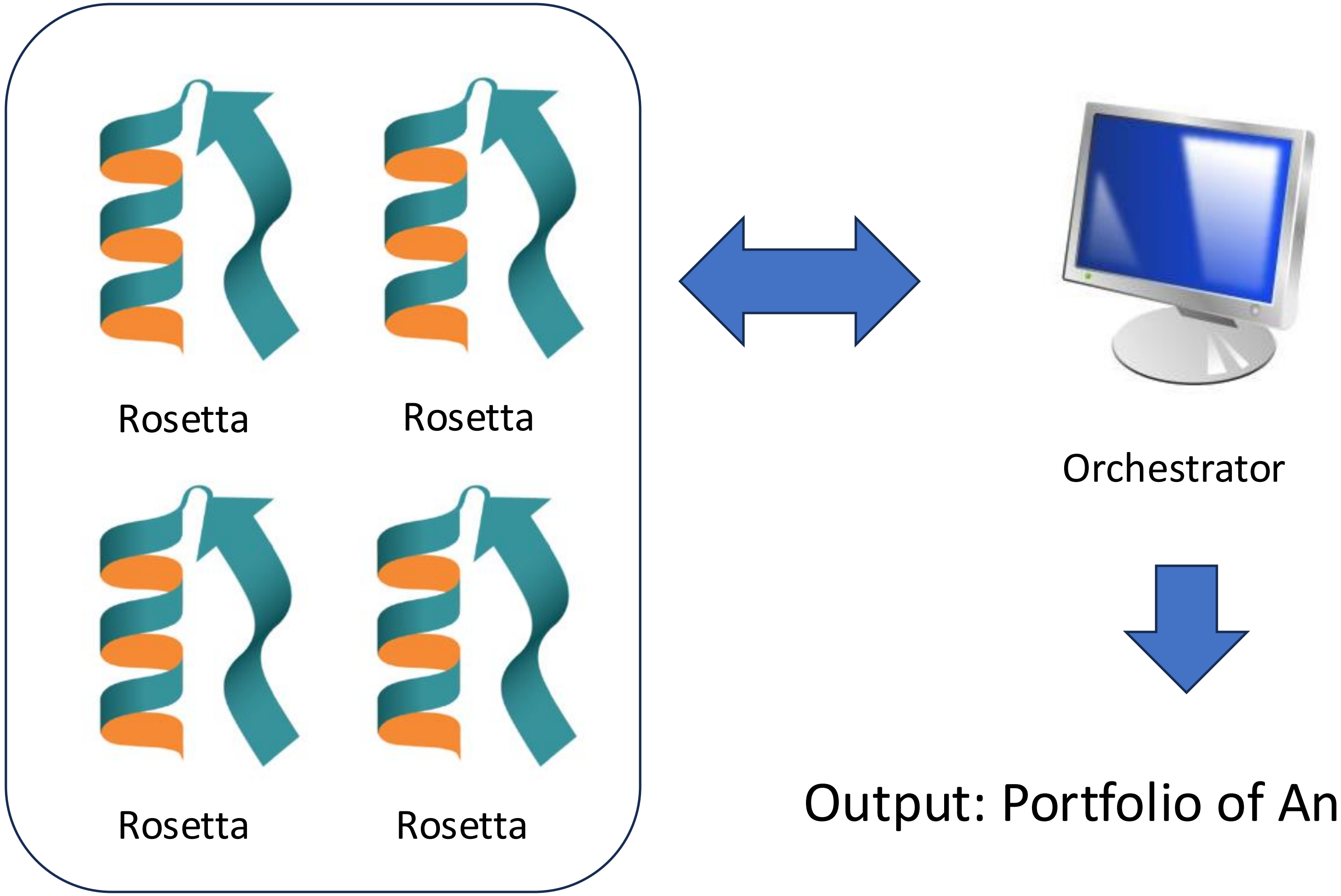
Design space:
20^5 = 3.2 million variants

Example: Computational Antibody Design

Cloris Cheng

Raul Astudillo

Tom Desautels



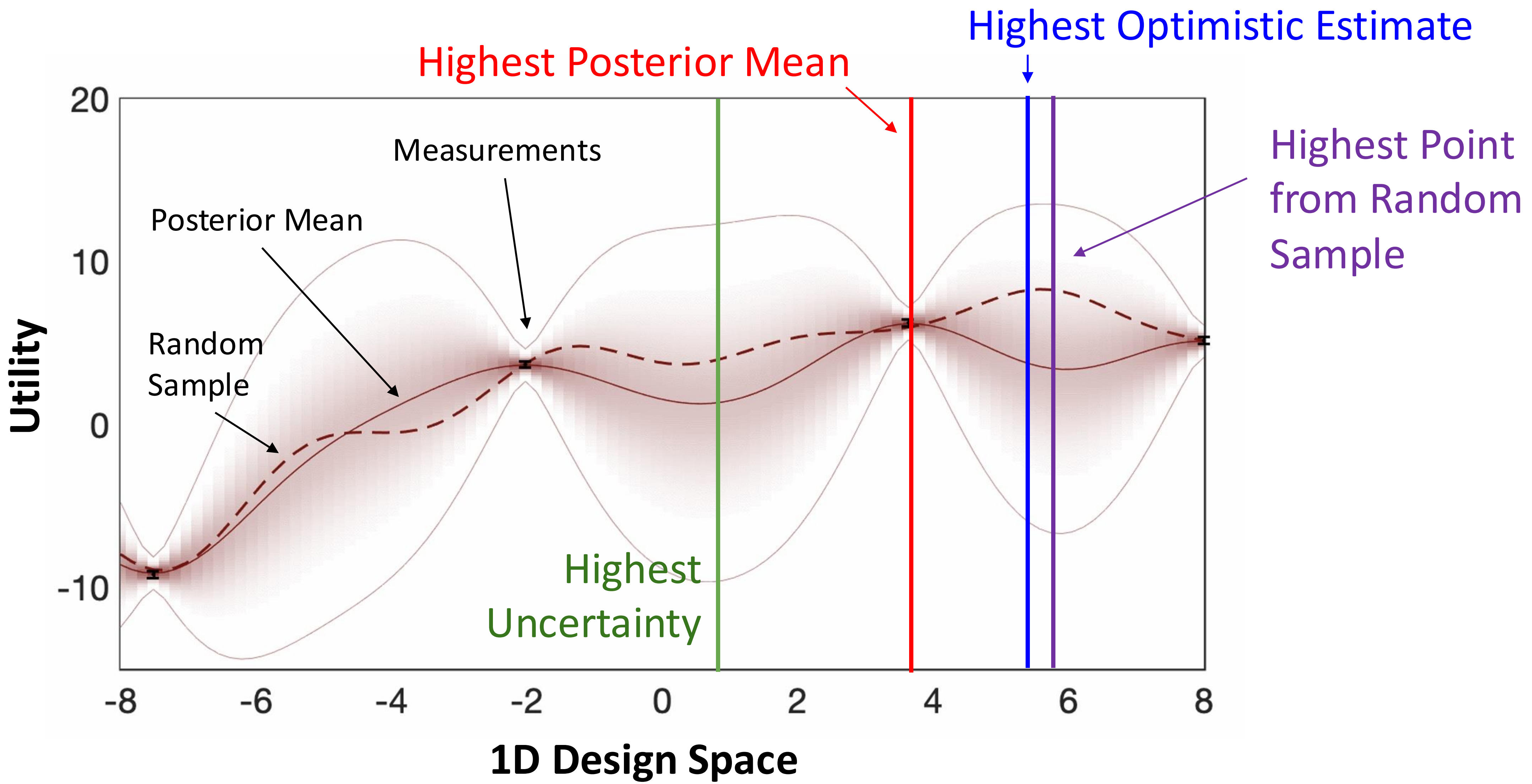
Output: Portfolio of Antibodies

Research Questions

- How do we scale in a principled way?
- How do we quantify uncertainty?
- How can we design methods for larger scope?
- How can we develop full-stack algorithmic exploration?

Work in Progress: Advice Welcome!

Thought Experiment: What point should we choose next?

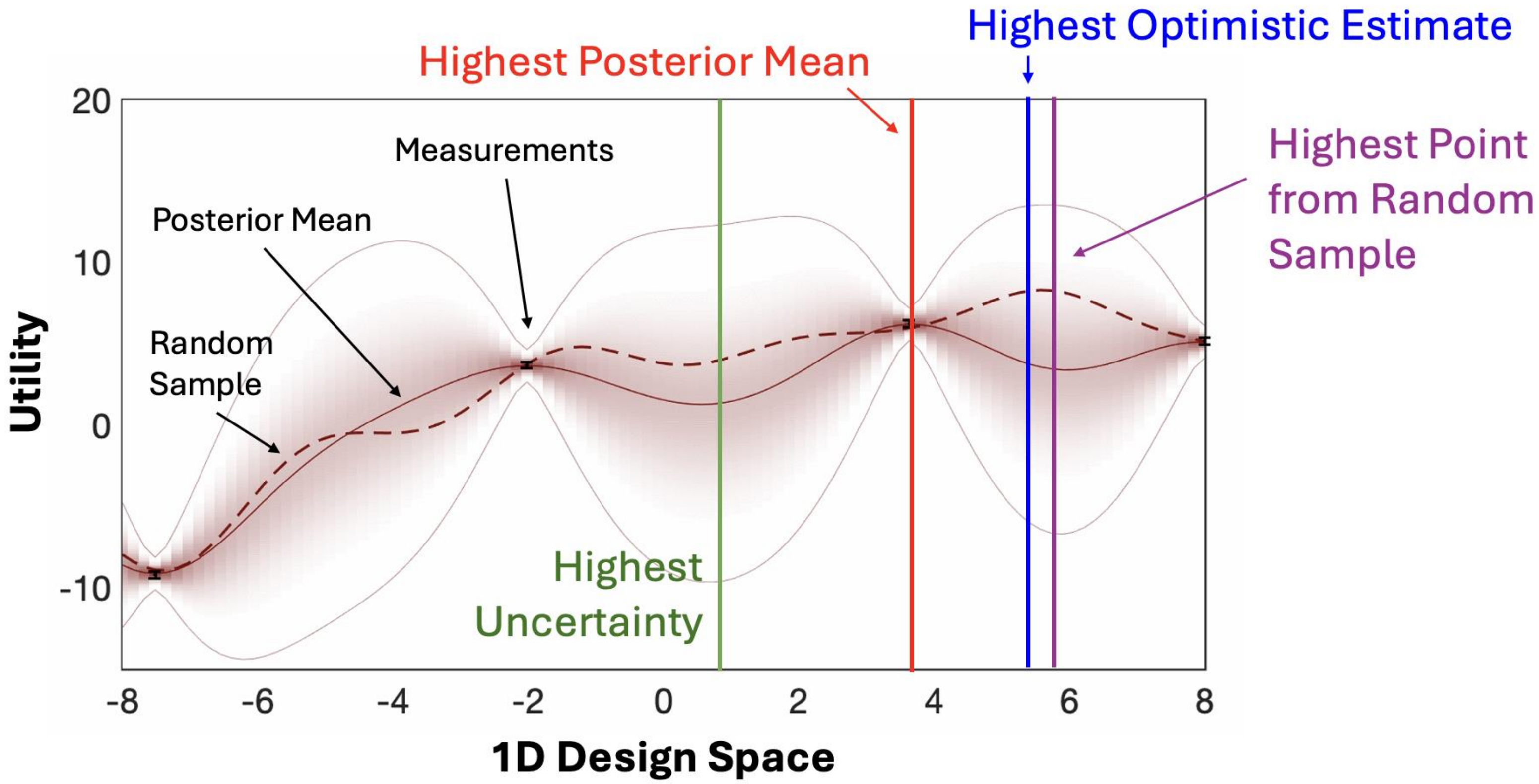


Value of Information

- **Informal Definition:** how much information (in expectation) the measurement tells us about our desired goal

- Desired goal is task-dependent:
  - Best solution
  - Pareto frontier
  - Most accurate model

- **Decision Policy:** choose experiment with highest VoI



Measures of VoI:

- Posterior Mean: won't learn that much
- Highest Uncertainty: learn useless information
- Optimistic: Either good solution or pruned design space
- Highest from Sample: similar to optimistic in expectation

Upper Confidence Bound

Posterior Sampling

Well understood in classic settings!
(convergence guarantees, applications)

Simple models (Gaussian Processes)
Low dimensional
Simple objectives
Requires calibrated uncertainty!

Measures of VoI:

- Posterior Mean: won't learn that much
- Highest Uncertainty: learn useless information
- Optimistic: Either good solution or pruned design space
- Highest from Sample: similar to optimistic in expectation

Upper Confidence Bound

Posterior Sampling

Example Application: Treating Lower Spine Injuries

10 mm

Medtronic
human
array

49 mm

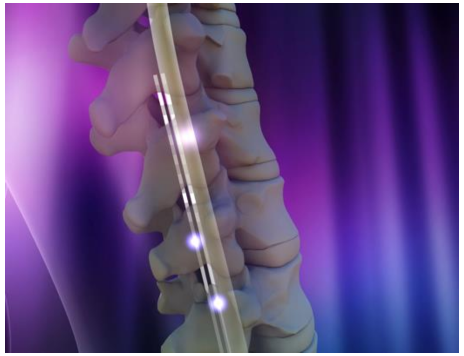


Image source:
williamcapicottomd.com

SCI Patient

Yanan
Sui

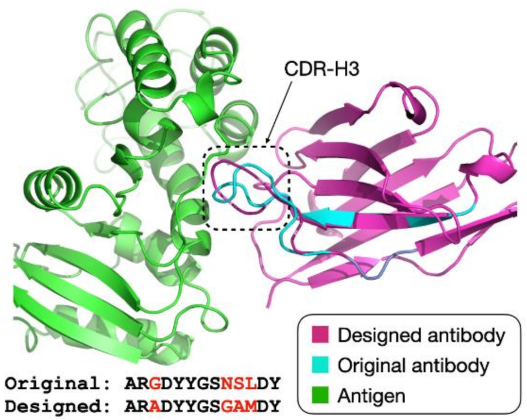
Joel
Burdick

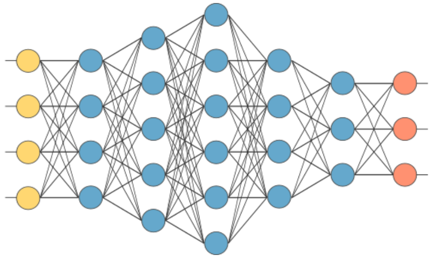
Each patient is unique

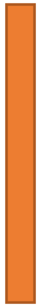
10^9 possible configurations!

Clinical Experiments

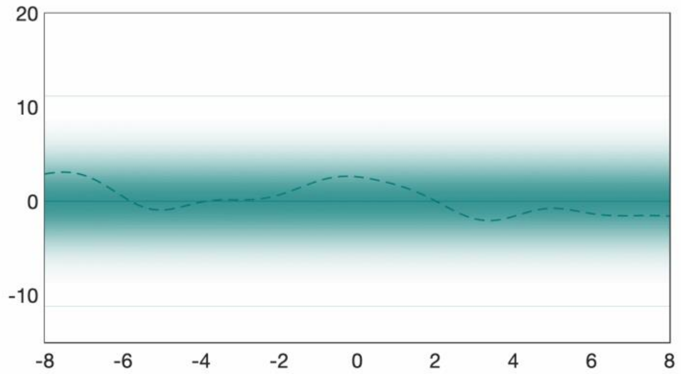
Rest of This Talk



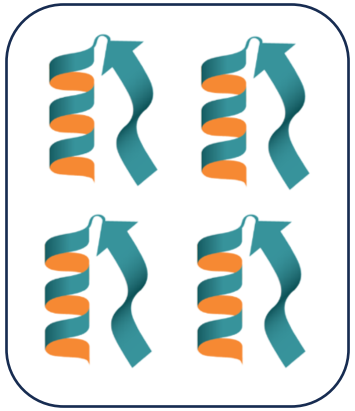




Embedding



Uncertainty Quantification
& Decision Making



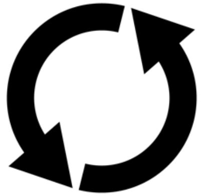
Orchestrator

Output:
Portfolio of Antibodies

Goal: Find Optimal Design

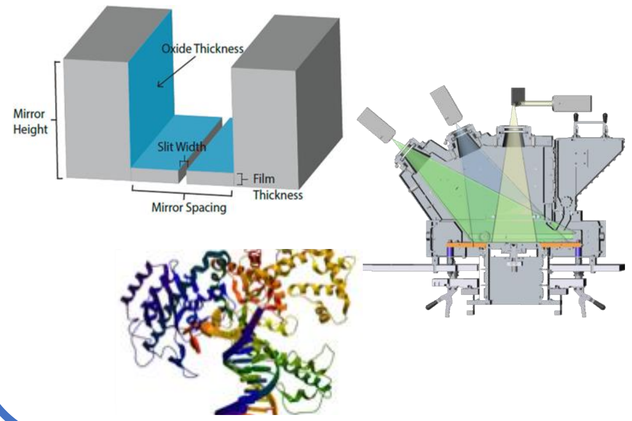
Experiment Designer

Measurement



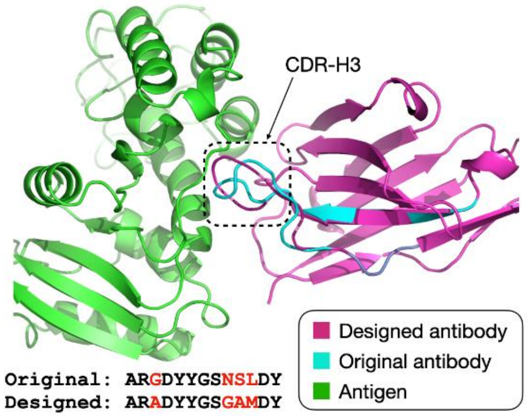
Experiment

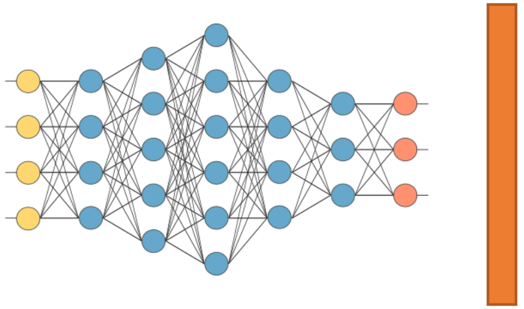
Experiment Platform

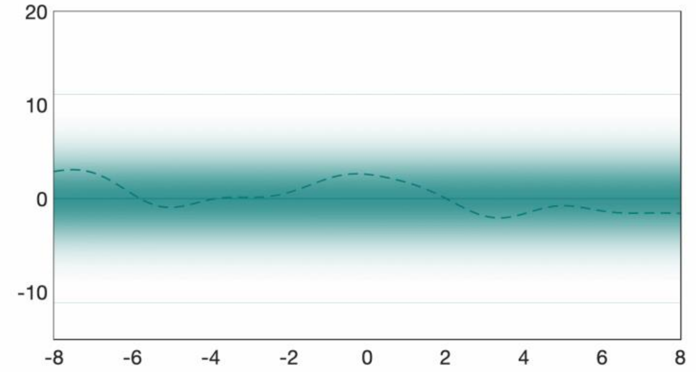


- Use **deep learning** for modeling & uncertainty quantification
- Design problem **out of distribution** (w.r.t. prior data)

Option 1: Deep Kernel Learning







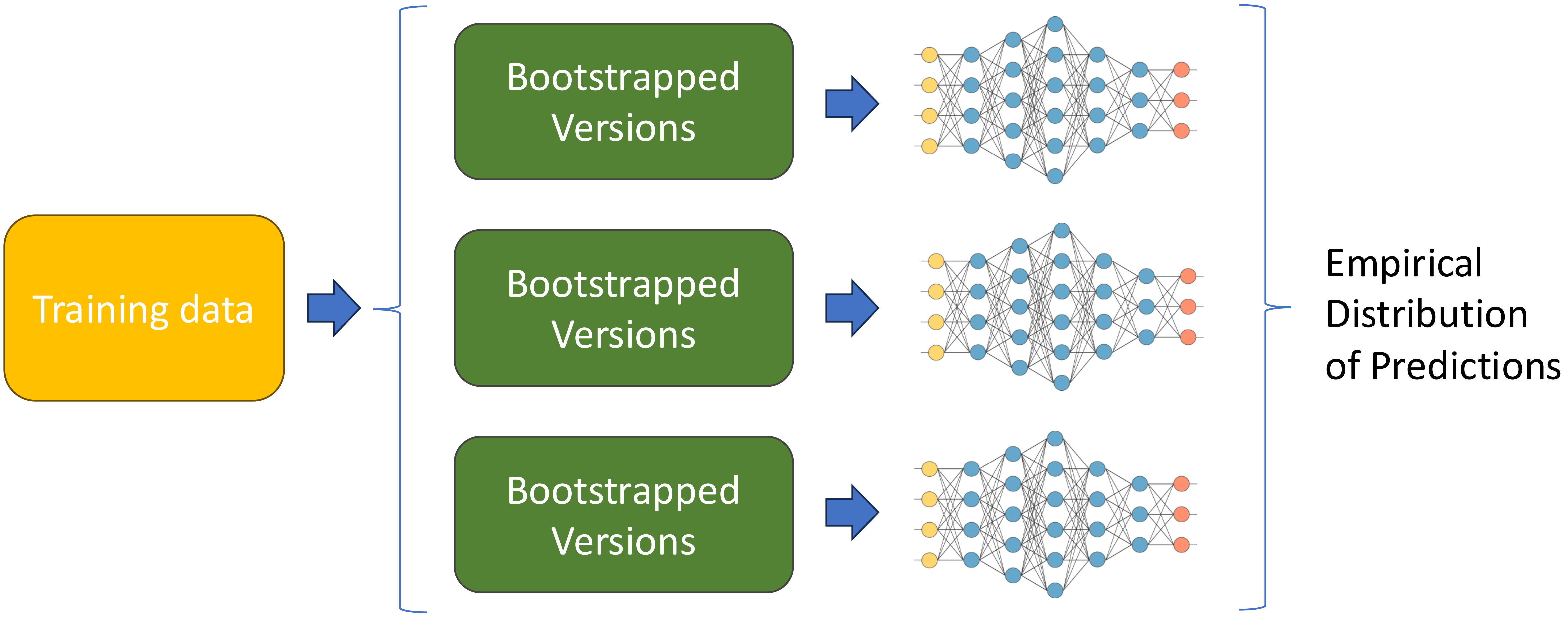
Deep Kernel
(previously hand crafted)

Gaussian Process

- Input to NN can be any existing representation
- Learn DK from scratch during experimentation
  - Update NN after each round of measurements

https://arxiv.org/abs/1511.02222

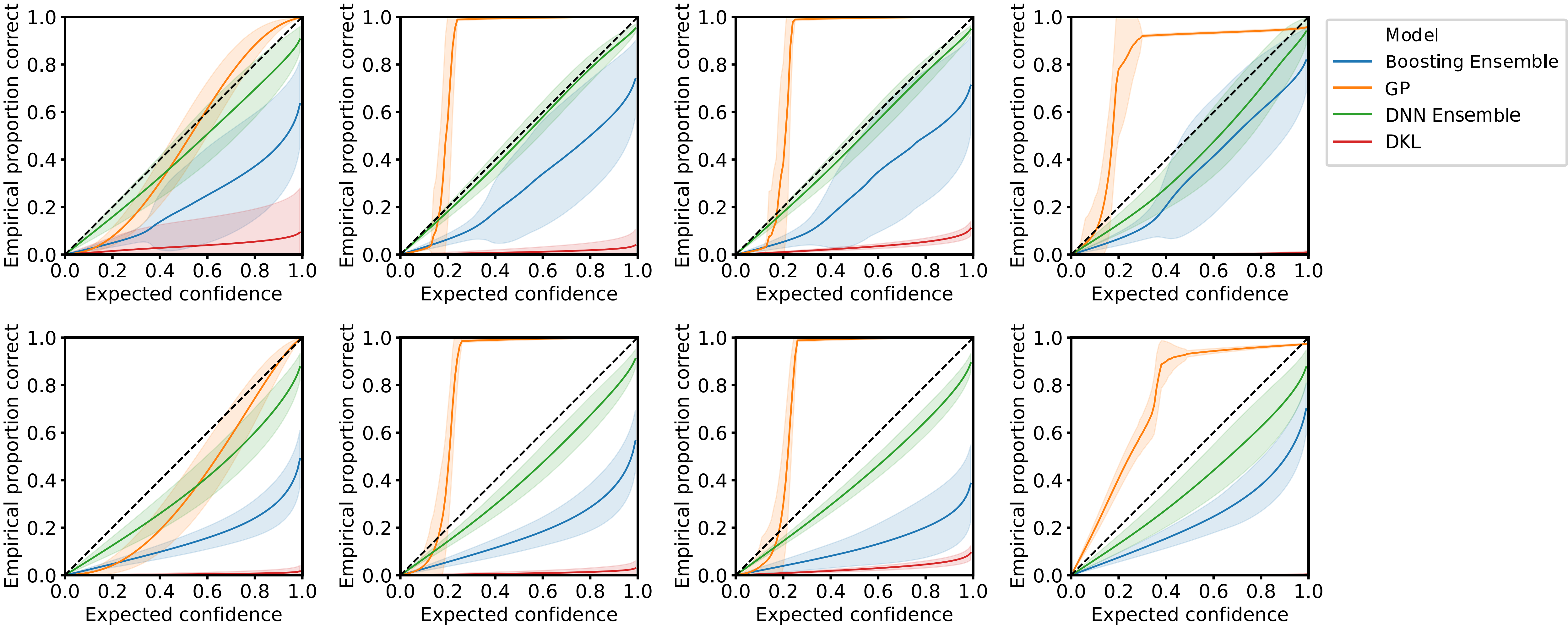
Option 2: Frequentist Ensembles



Comparison

Active Learning-Assisted Directed Evolution, Yang, et al., *(under review)*

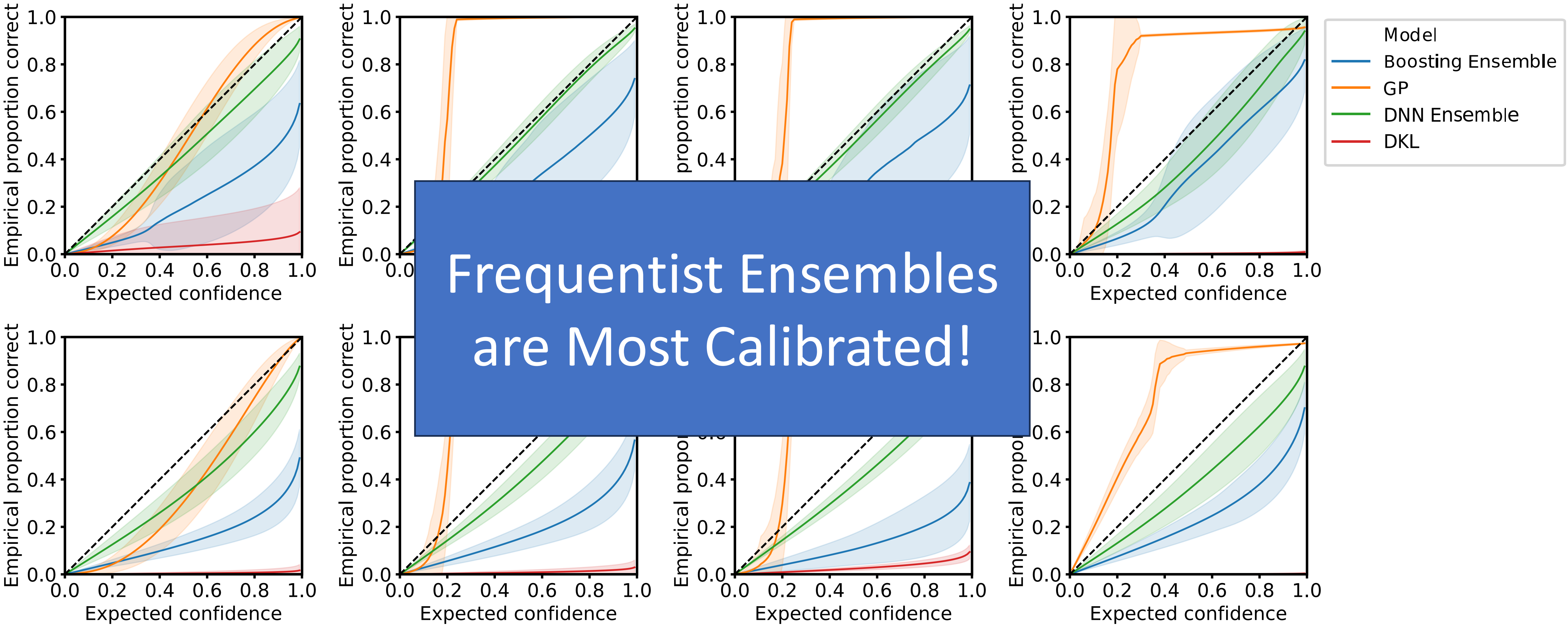
Jason
Yang



Comparison

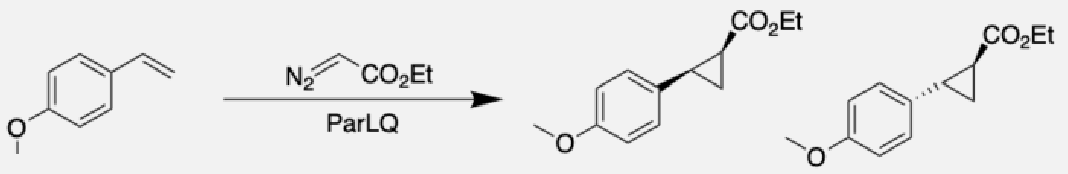
Active Learning-Assisted Directed Evolution, Yang, et al., *(under review)*

Jason
Yang



Frequentist Ensembles
are Most Calibrated!

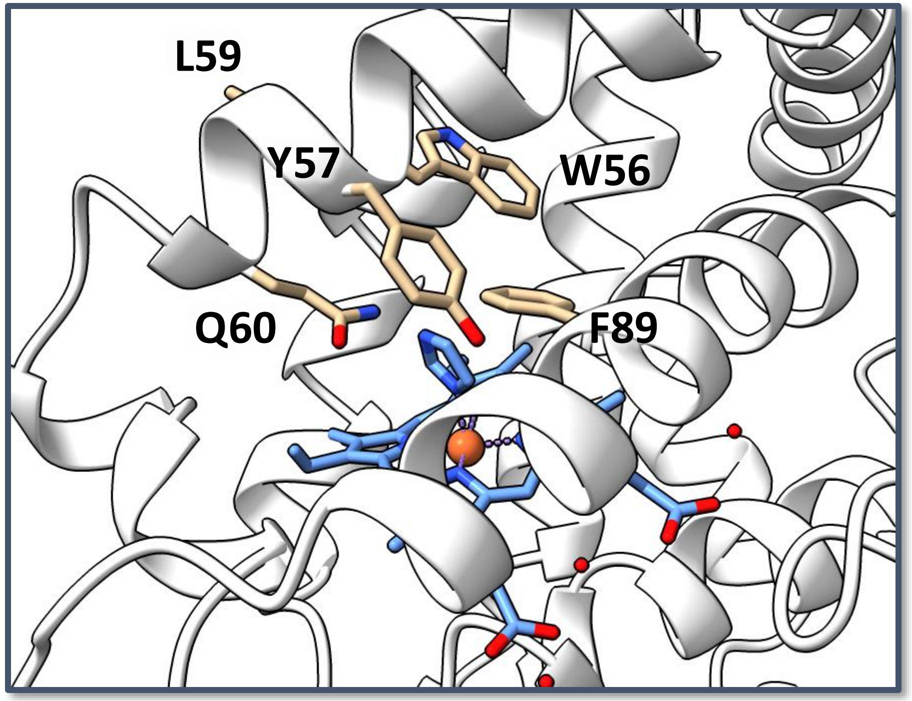
Recall: Engineering Active Site of Protoglobin



1 *cis* product

2 *trans* product

ELEGEN



***Par*Pgb-LQ**

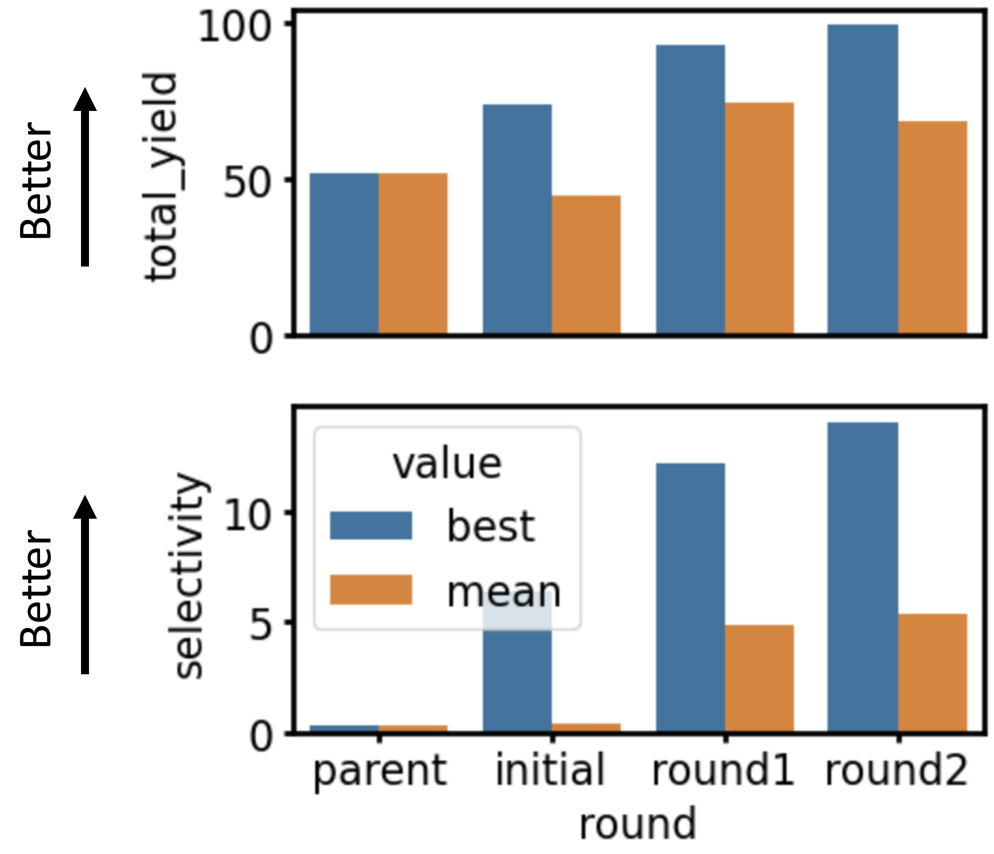
Jason Yang

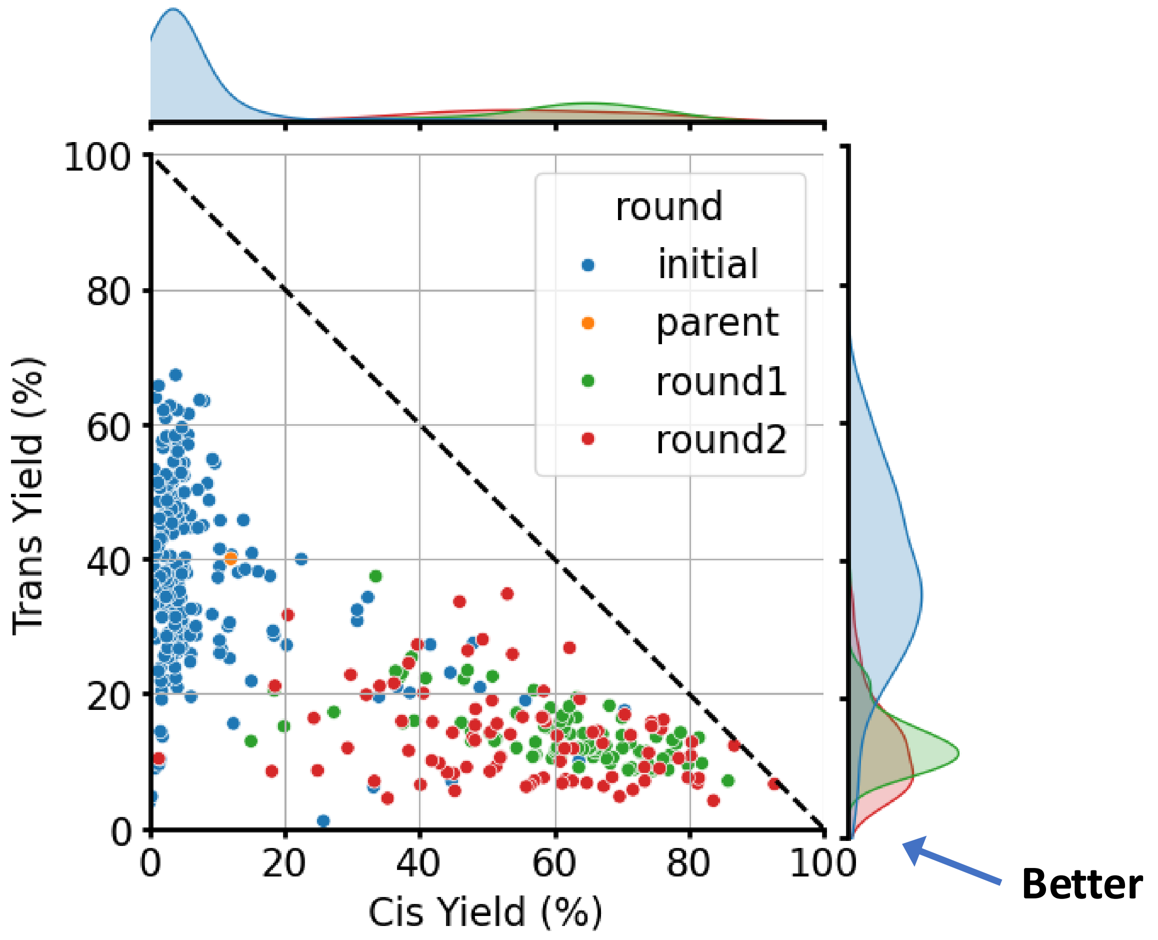
Design Problem:
find the optimal set of mutations on 5 sites
for high yield and selectivity for **1**, the *cis* product

Design space:
20^5 = 3.2 million variants

Results

- 96 experiments per round
- Posterior sampling
- Standard directed evolution ineffective (highly epistatic)

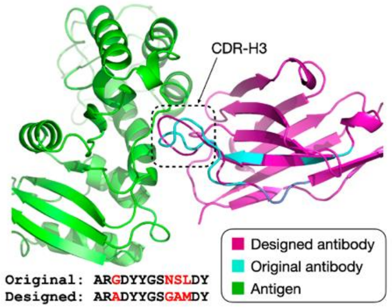


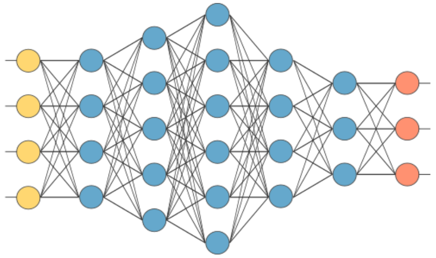


Active Learning-Assisted Directed Evolution, Yang, et al., (in preparation)

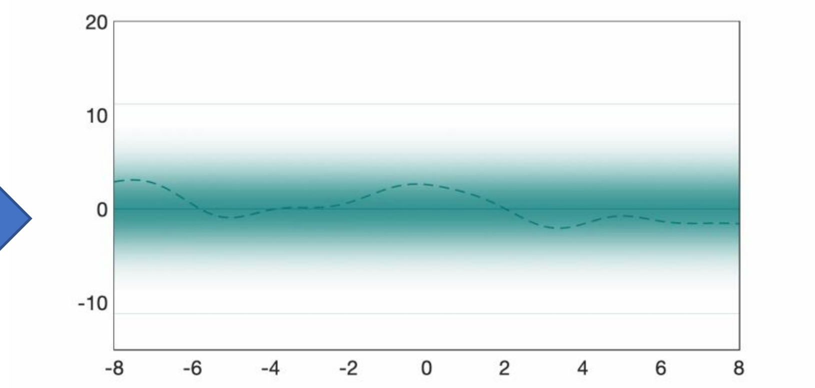
Deep Kernel Learning Revisited

(Let’s be fully Bayesian!)





P(DK)



P(GP|DK)

Given Data S: P(DK|S) P(GP|DK,S)

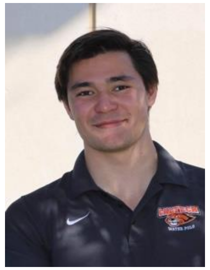
Normal DKL: Point estimate of DK

Bayesian Optimization with Bayesian Deep Kernel Learning

Bowden, Yeh, Astudillo, et al. *(in preparation)*

- Key challenge: computing posterior P(DK|S)
  - P(DK) = Prior over NN weights
  - P(DK|S) = Posterior of NN weights

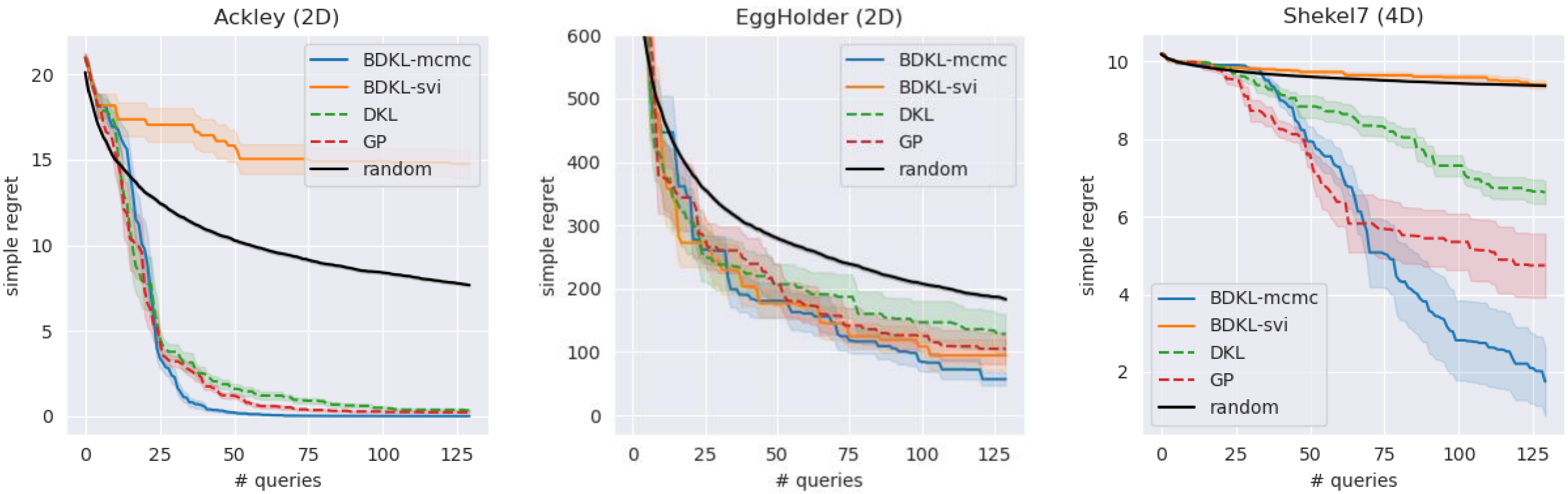
- Two Options:
  - Stochastic Variational Inference (fast, inaccurate)
  - MCMC (slow, more accurate)


James Bowden


Chris Yeh

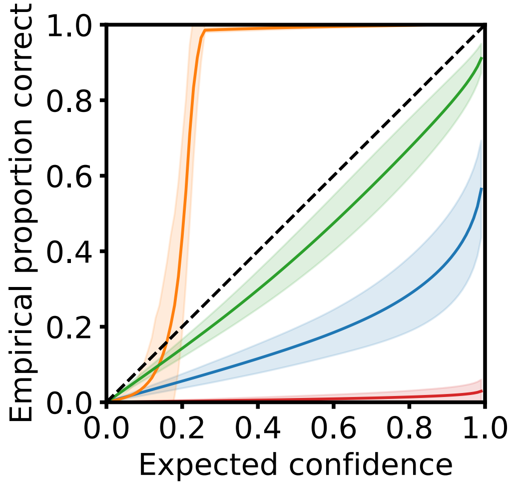

Raul Astudillo

Preliminary Results

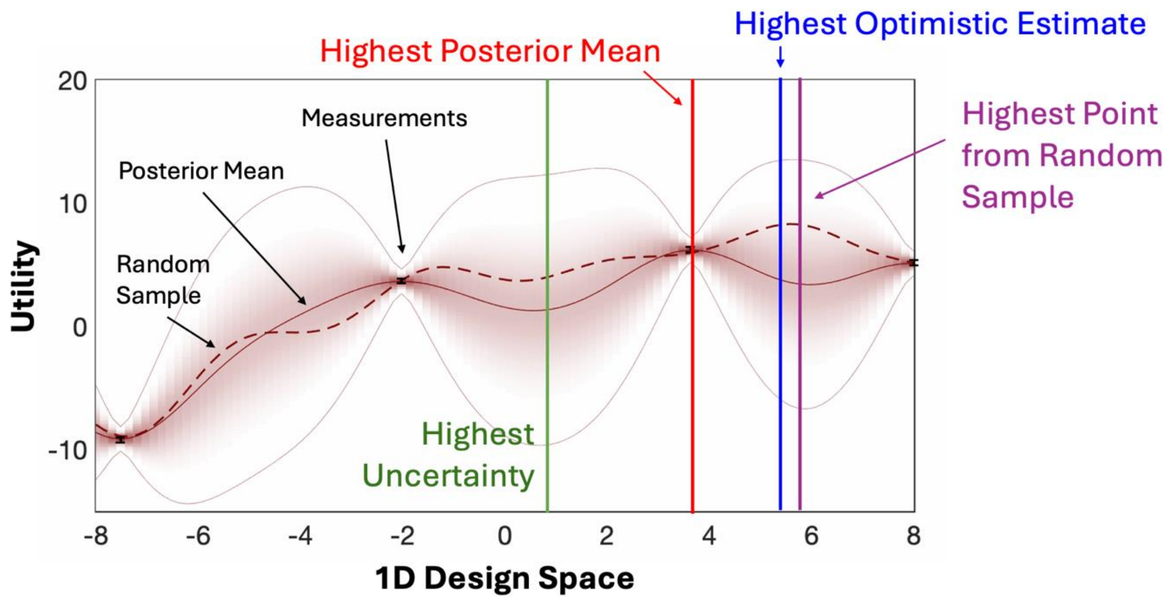


Note: implementation highly non-trivial. Will release code soon!

Story so Far



Calibration is Hard!

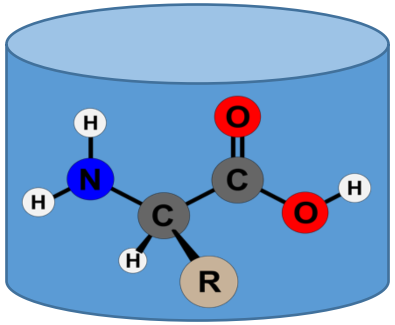


Value of Information requires extra reasoning

How to compute VoI at scale (or side-step it)?

- Computational
- Statistical Calibration

Aside #1: Directly Learn Value of Information



Prior Experiments

Simulate Policy

Experiments

$(x_1, y_1, x_2, y_2, x_3, y_3, \ldots)$

Measurements

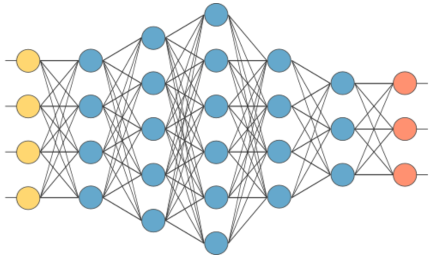
Update Policy

Measure VoI
(post hoc)

Training Data:
Input: $(x_{<i}, y_{<i}, x_i)$
Target: VoI of y_i

Regression

VoI Model



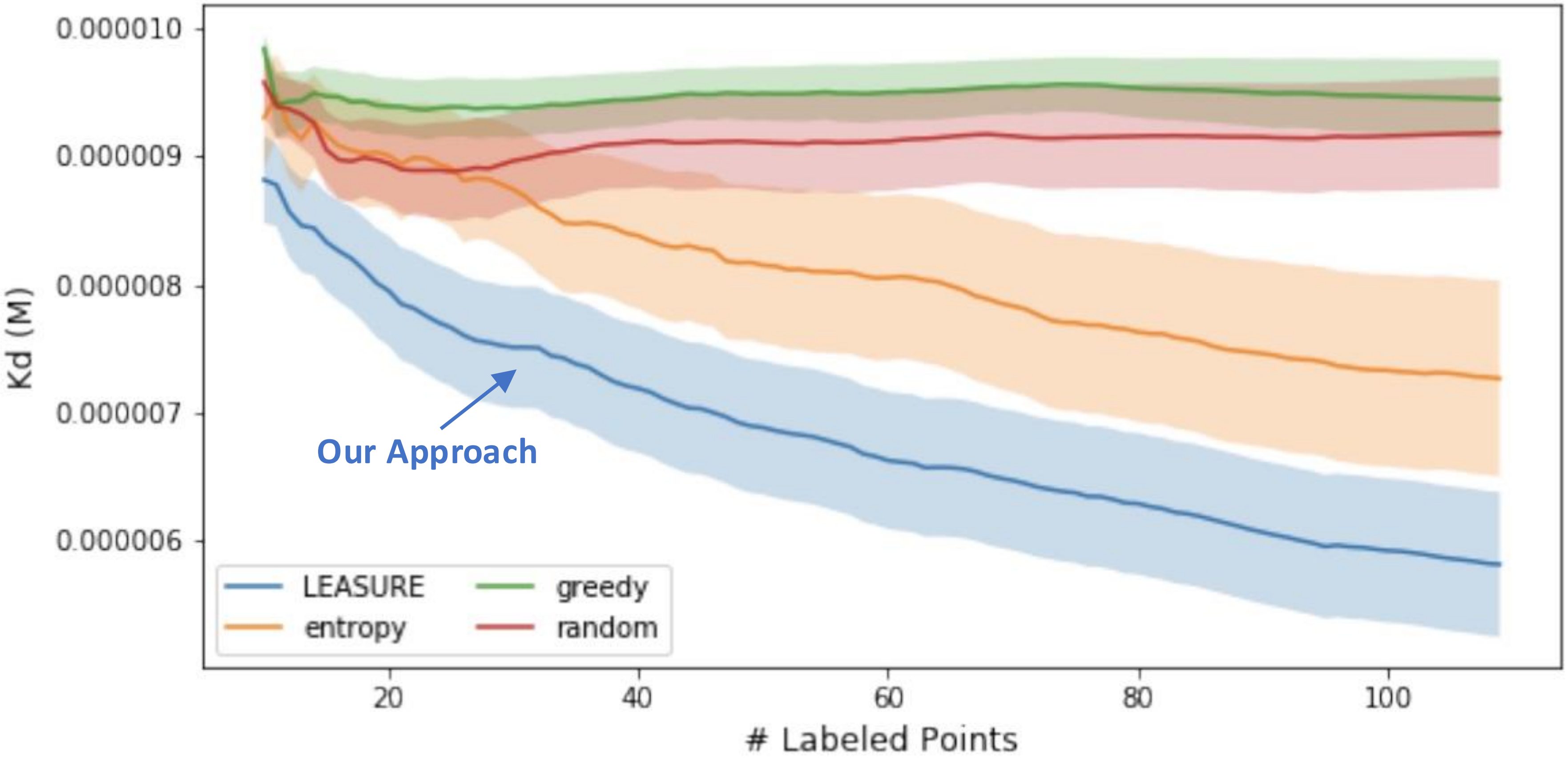
Key Insight: regularize to "look like" submodular

Ayya
Alieva

Learning to Make Decisions via Submodular Regularization, Ayya Alieva, et al., ICLR 2021

Preliminary Results

(Train VoI model on one antibody task, deploy on new task)

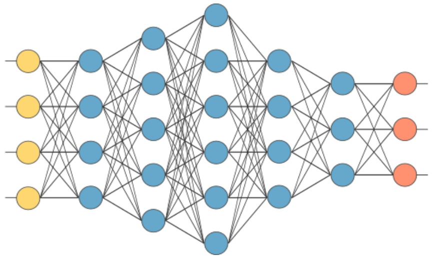


Propensity for protein binding to break apart

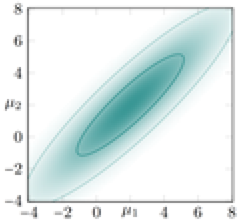
BETTER

Learning to Make Decisions via Submodular Regularization, Ayya Alieva, et al., ICLR 2021

Aside #2: Decision-Aware Uncertainty Calibration



Conformal Calibration



Uncertainty-Aware Decision Policy

$$z^\star(x) = \arg\min_{z} \max_{\hat{y} \in \Omega(\mathrm{x})} \; f(x, \hat{y}, z)$$
$$\text{s.t.} \quad g(x, \hat{y}, z) \le 0$$

Backprop

Backprop

Backprop

Deploy / Simulate

Reward

Chris Yeh

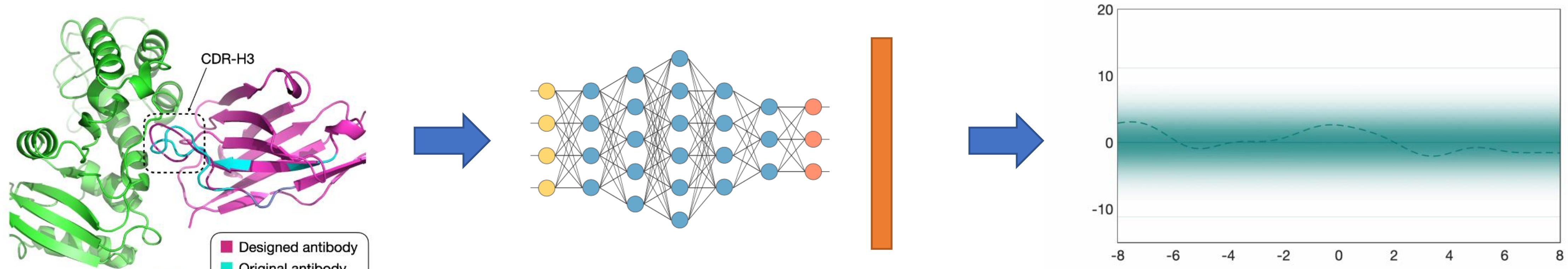
Nico Christianson

End-to-End Conformal Calibration for Optimization Under Uncertainty, Yeh, Christianson, et al., *arXiv*

Preliminary Results (battery storage)

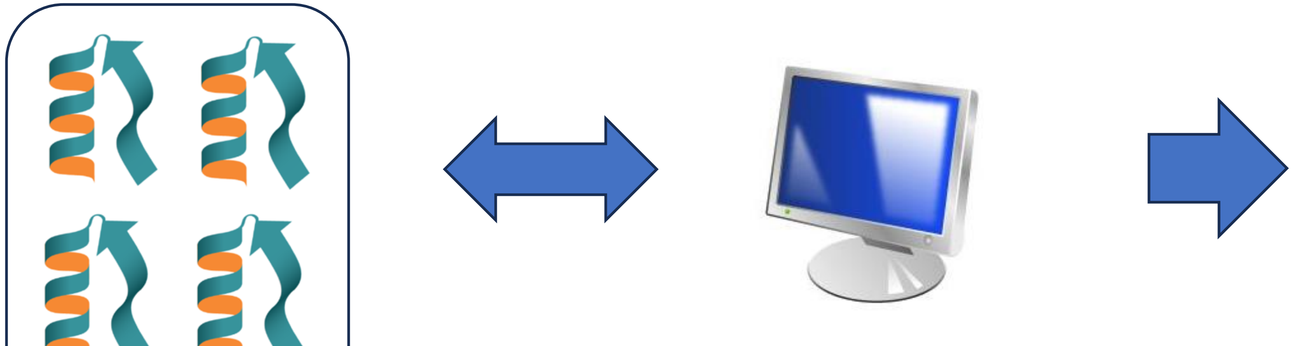
End-to-End Conformal Calibration for Optimization Under Uncertainty, Yeh, Christianson, et al., *arXiv*

Rest of This Talk



Embedding

Uncertainty Quantification
& Decision Making



Orchestrator

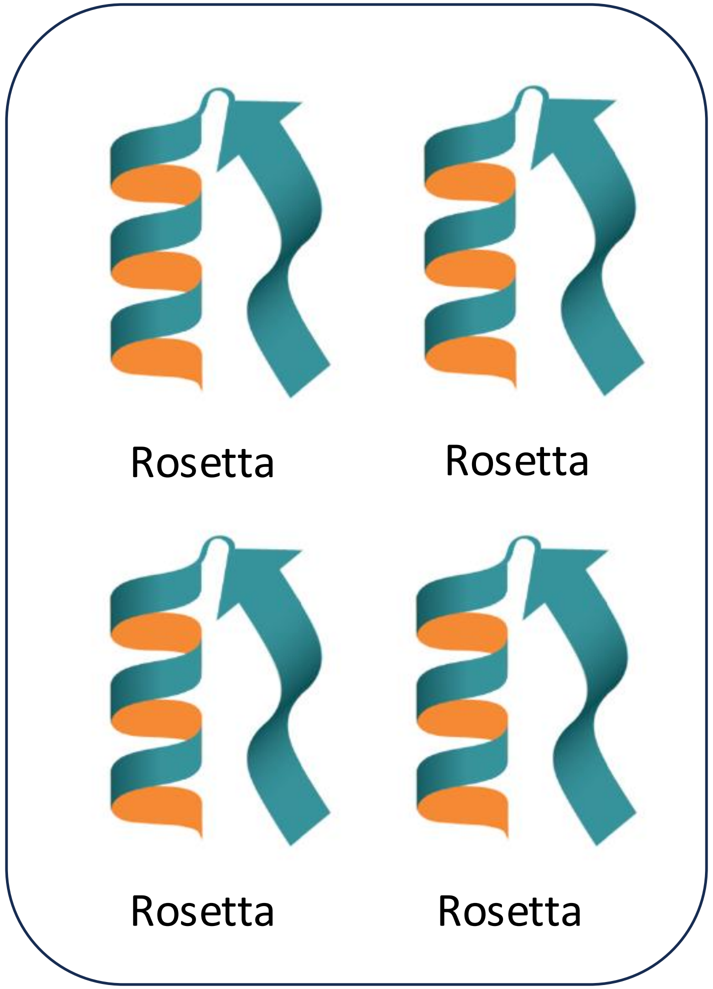
Output:
Portfolio of Antibodies

Recall: Computational Antibody Design

Cloris Cheng

Raul Astudillo

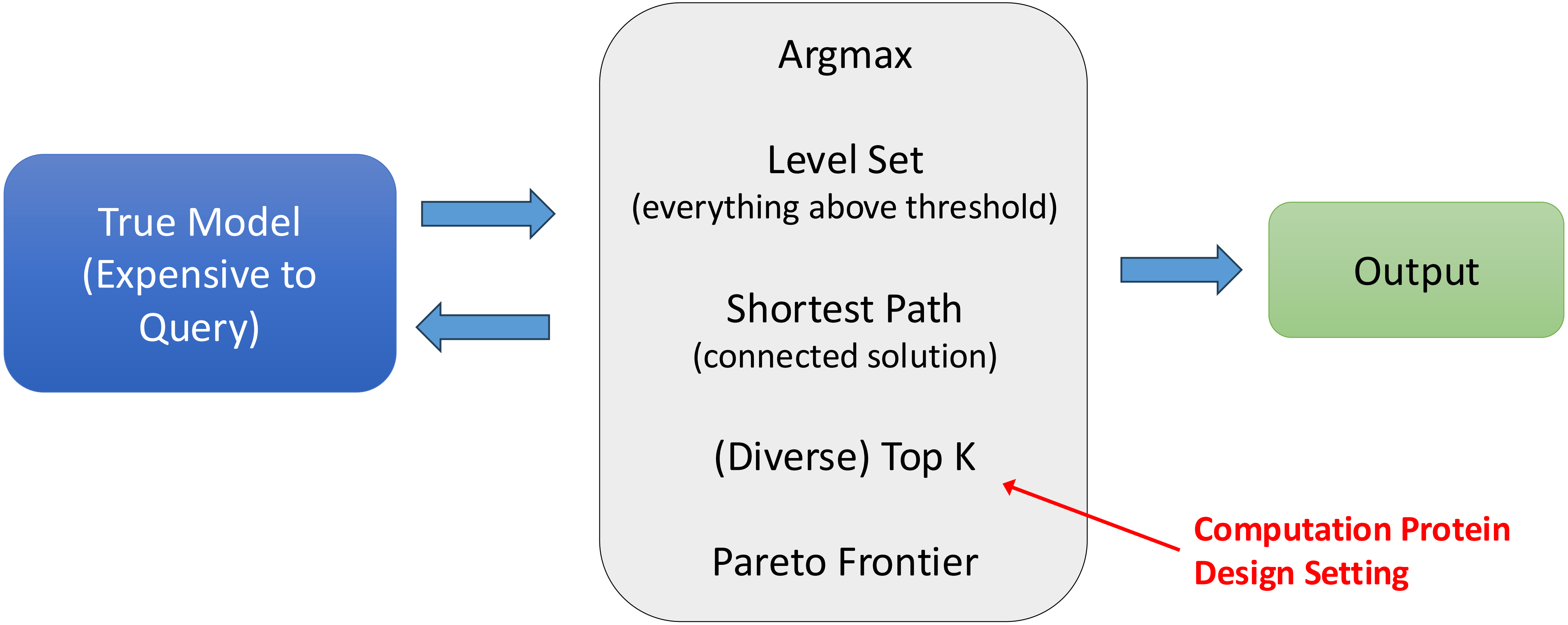
Tom Desautels



Orchestrator

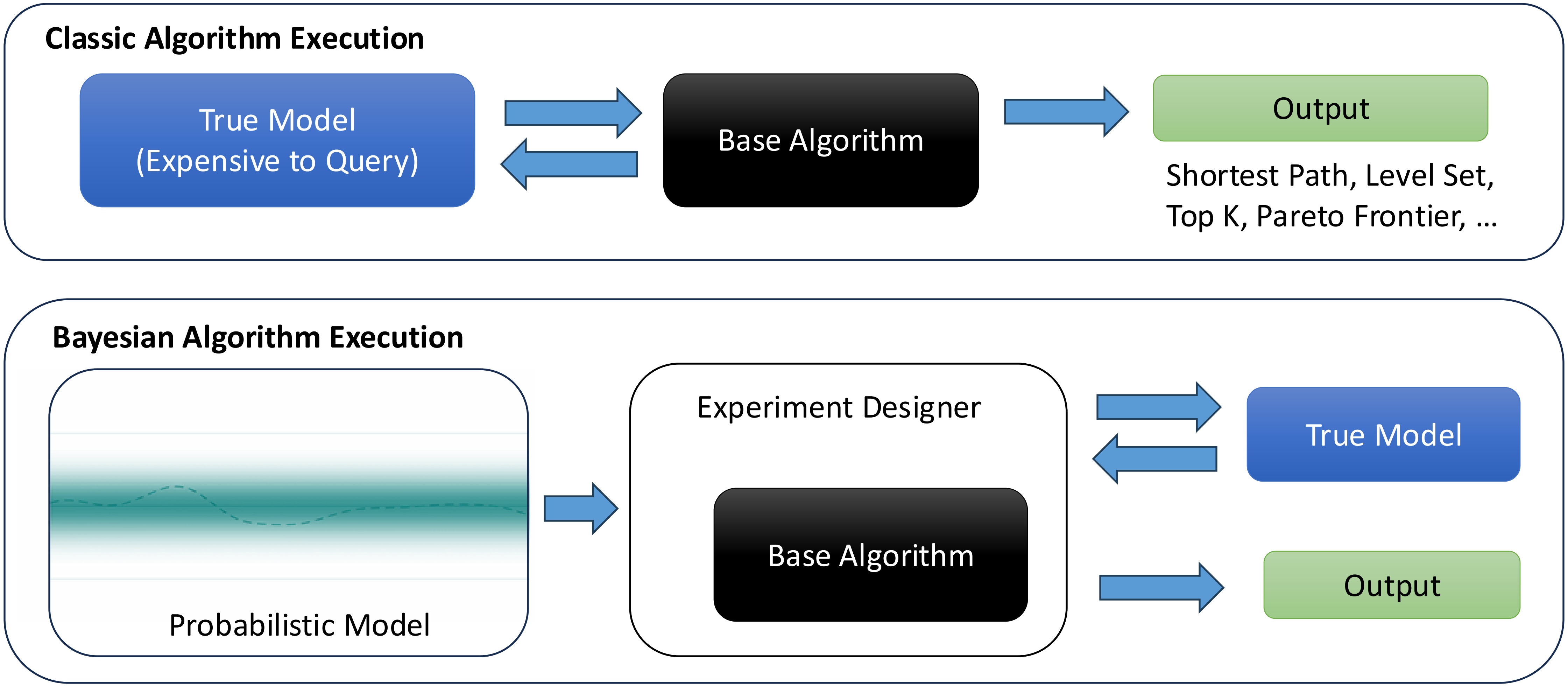
Output: Portfolio of Antibodies

Complex Objectives

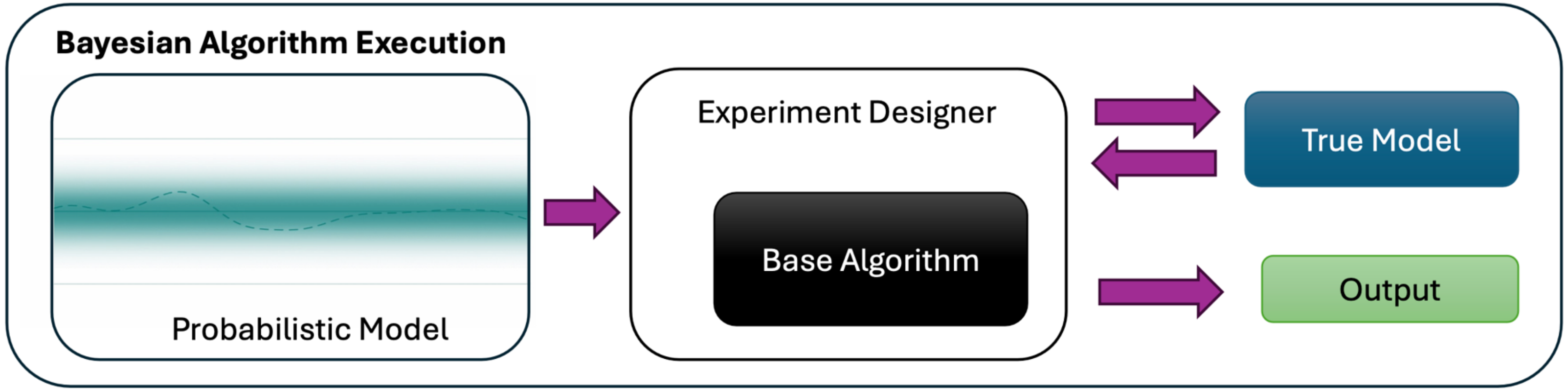


Problem Setup: Bayesian Algorithm Execution

https://arxiv.org/abs/2104.09460



Revisiting Value of Information



- **Informal Definition:** how much information (in expectation) the measurement tells us about our desired goal

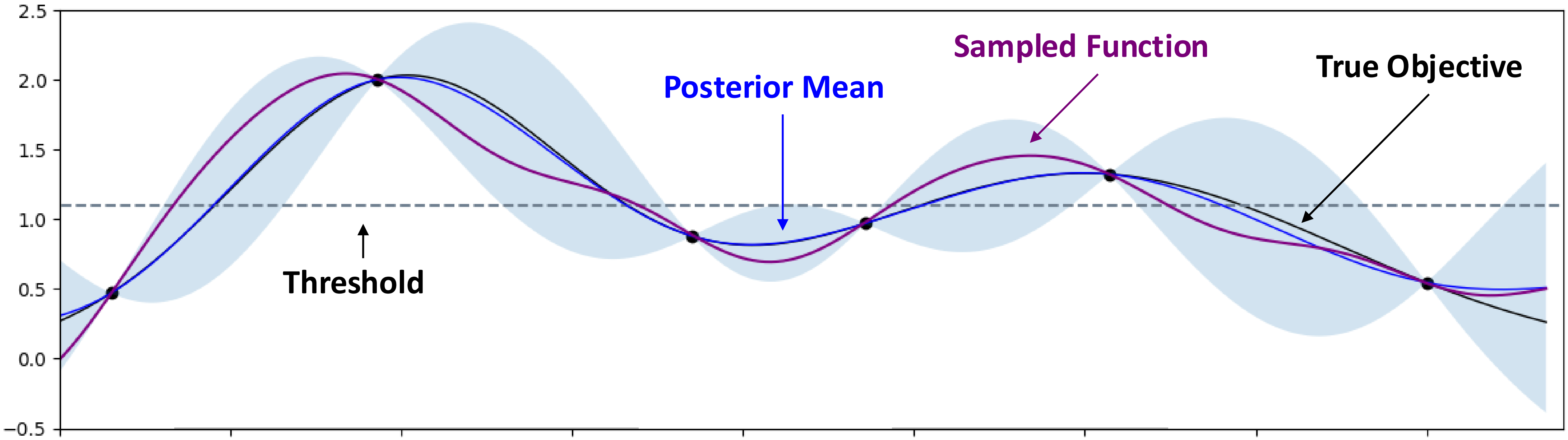
- How to quantify VoI for best output of Base Algorithm?
  - **Direct formulation is intractable!**

Practical Bayesian Algorithm Execution via Posterior Sampling

Cheng*, Astudillo*, Desautels, Yue, NeurIPS 2024

Goal: Identify All Points Above Threshold With Few Queries of True Objective

Example for Level Set in 1D Space



Cloris Cheng

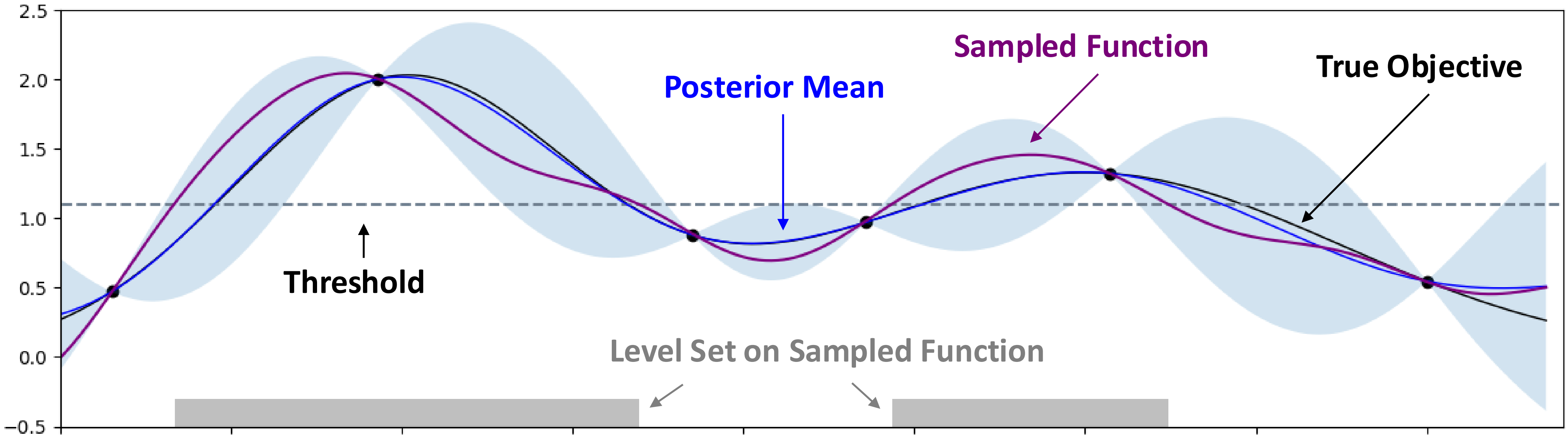
Raul Astudillo

Practical Bayesian Algorithm Execution via Posterior Sampling

Cheng*, Astudillo*, Desautels, Yue, NeurIPS 2024

Goal: Identify All Points Above Threshold With Few Queries of True Objective

Example for Level Set in 1D Space



Cloris Cheng

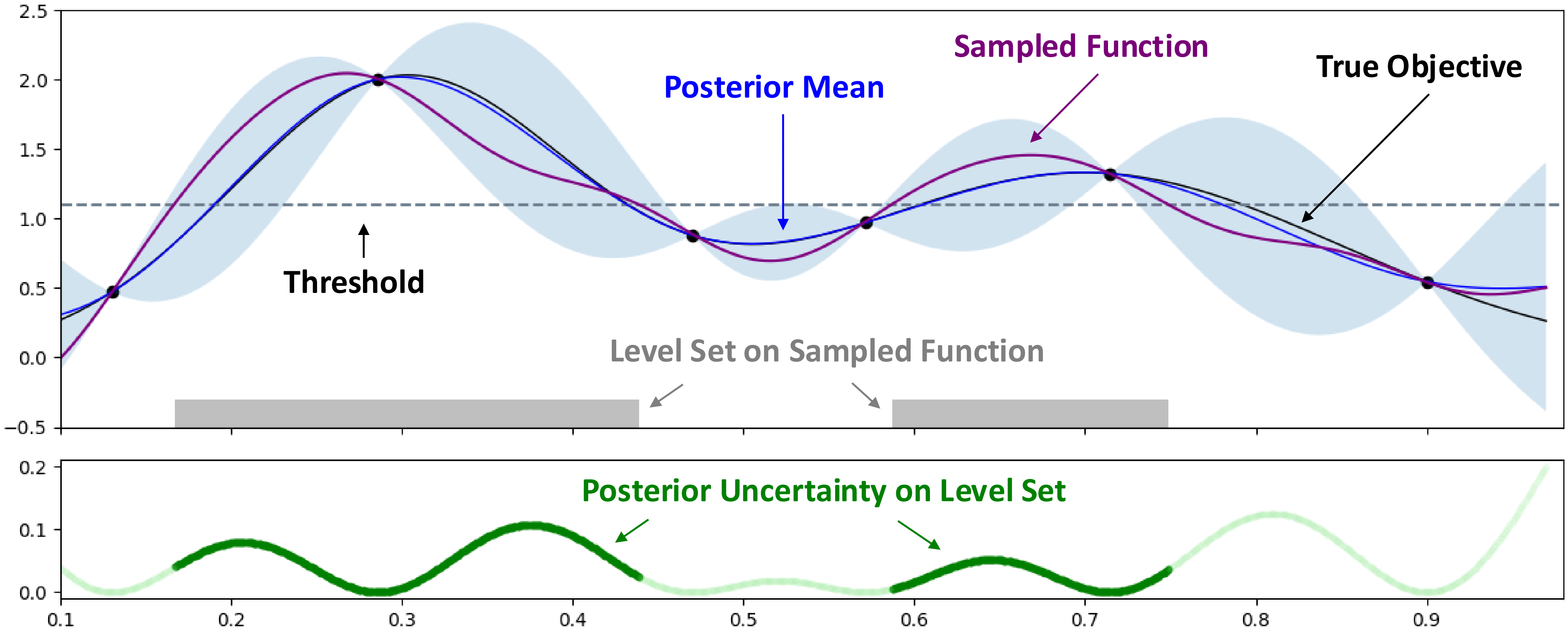
Raul Astudillo

Practical Bayesian Algorithm Execution via Posterior Sampling

Cheng*, Astudillo*, Desautels, Yue, NeurIPS 2024

Goal: Identify All Points Above Threshold With Few Queries of True Objective

Example for Level Set in 1D Space



Cloris Cheng

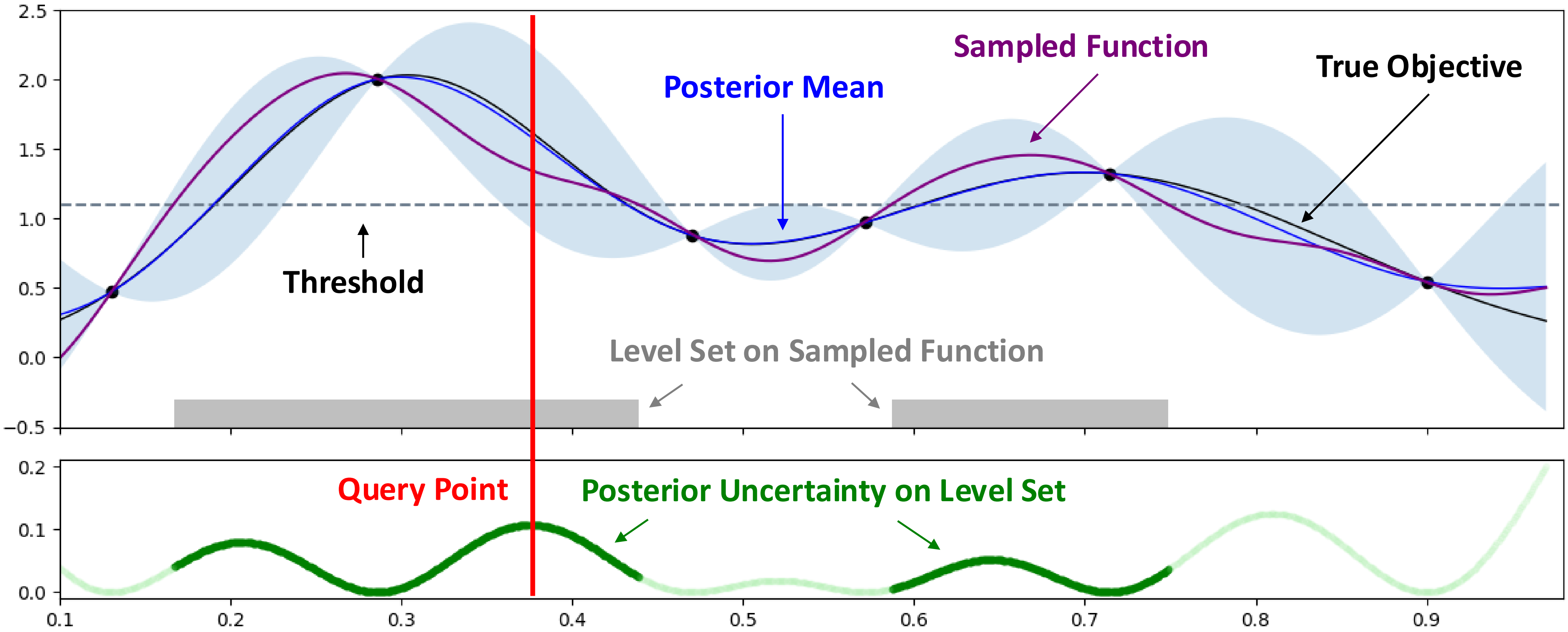
Raul Astudillo

Practical Bayesian Algorithm Execution via Posterior Sampling

Cheng*, Astudillo*, Desautels, Yue, NeurIPS 2024

Goal: Identify All Points Above Threshold With Few Queries of True Objective

Example for Level Set in 1D Space

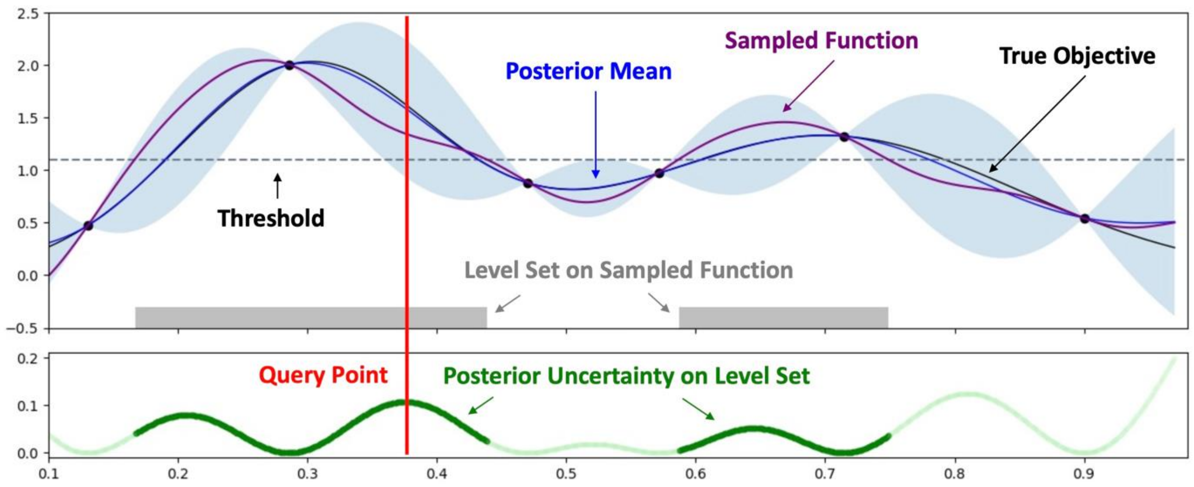


Cloris Cheng

Raul Astudillo

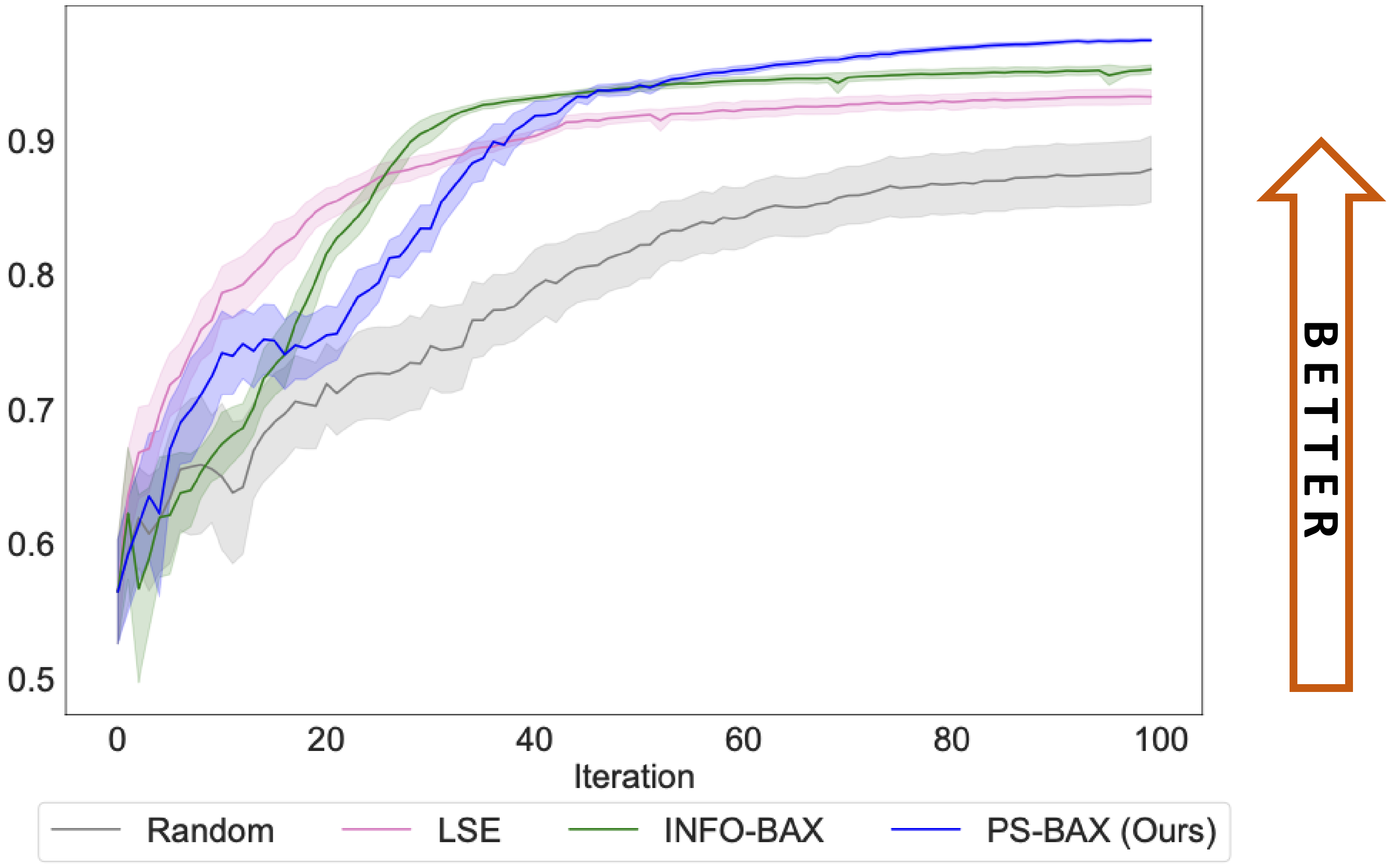
Benefits of PS-BAX

- Computationally efficient
  - 15x-400x faster than baseline
- Straightforward to implement
  - Batch parallel experiments
- Asymptotically consistent



Practical Bayesian Algorithm Execution via Posterior Sampling,
Cheng*, Astudillo*, Desautels, Yue, NeuIPS 2024

Level Set Experiments



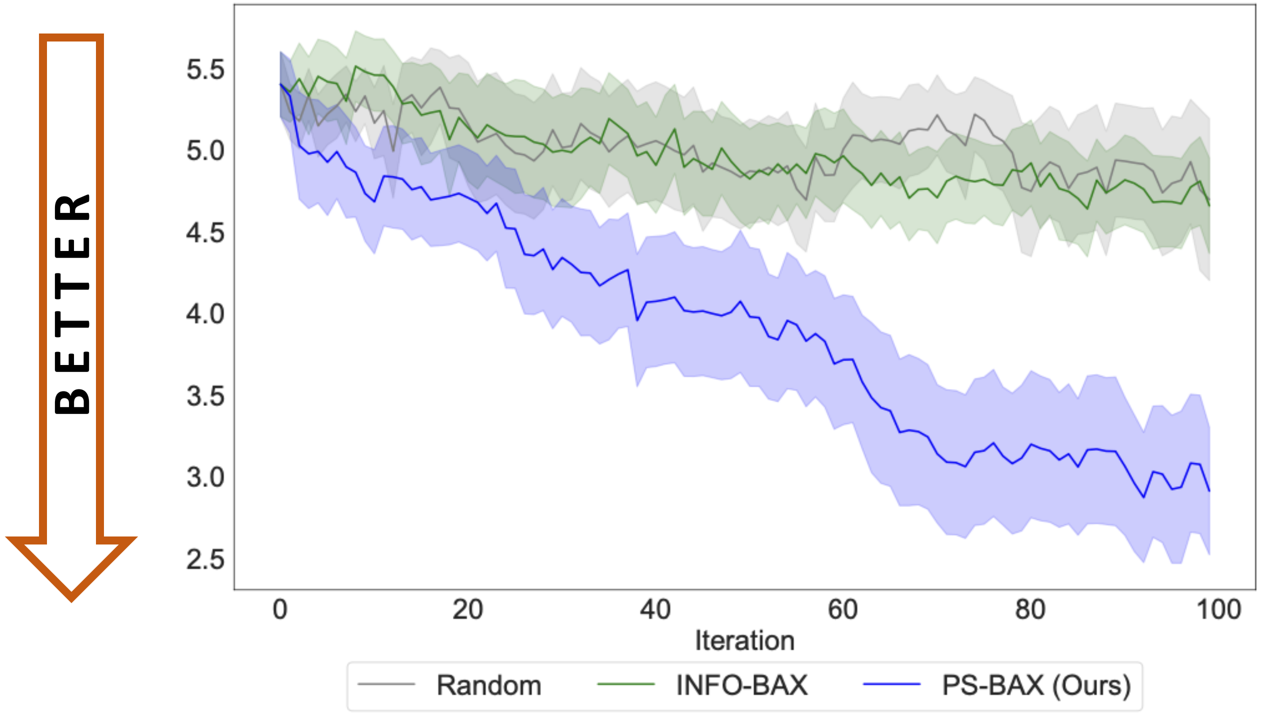
PS-BAX is ~400x faster than INFO-BAX

Drug Discovery Application (DiscoBAX)

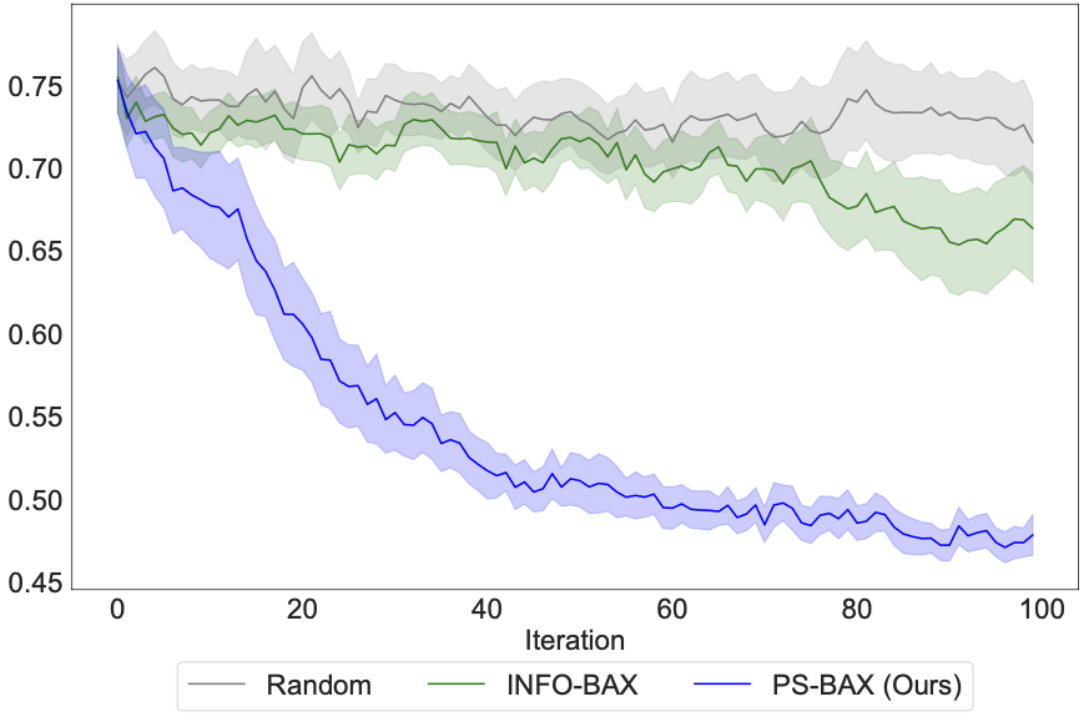
Choose K=10 interventions (reward = best intervention)

Dataset from Project Achilles (https://www.biorxiv.org/content/10.1101/720243)

Tau Protein Assay



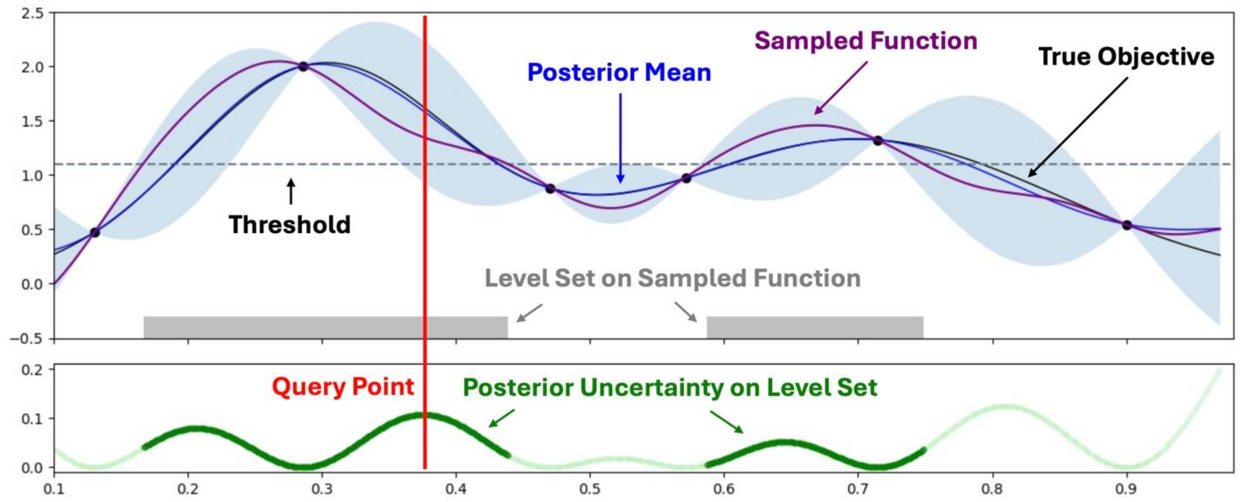
Intergeron γ Assay



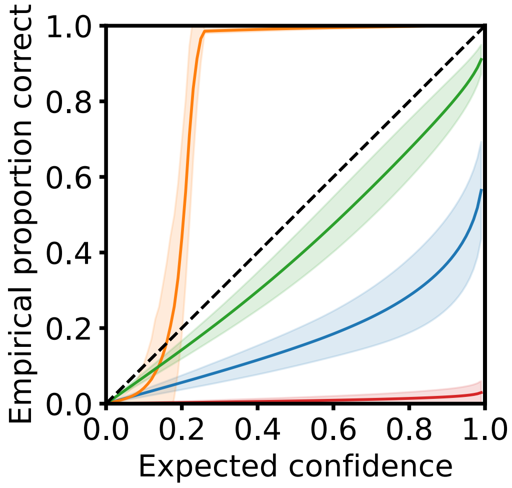
PS-BAX is ~30x faster than INFO-BAX

https://arxiv.org/abs/2312.04064

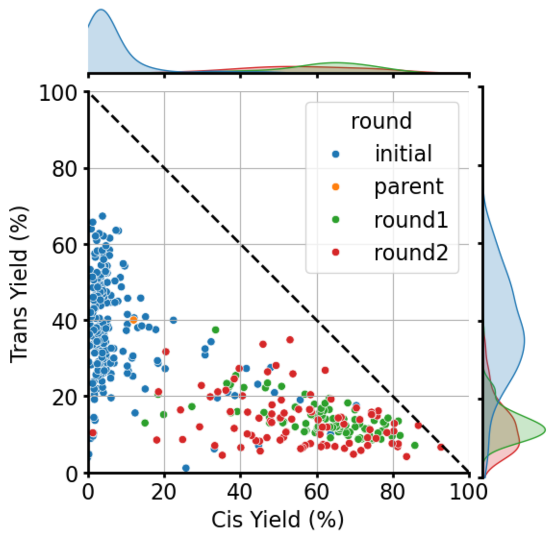
Summary of Results



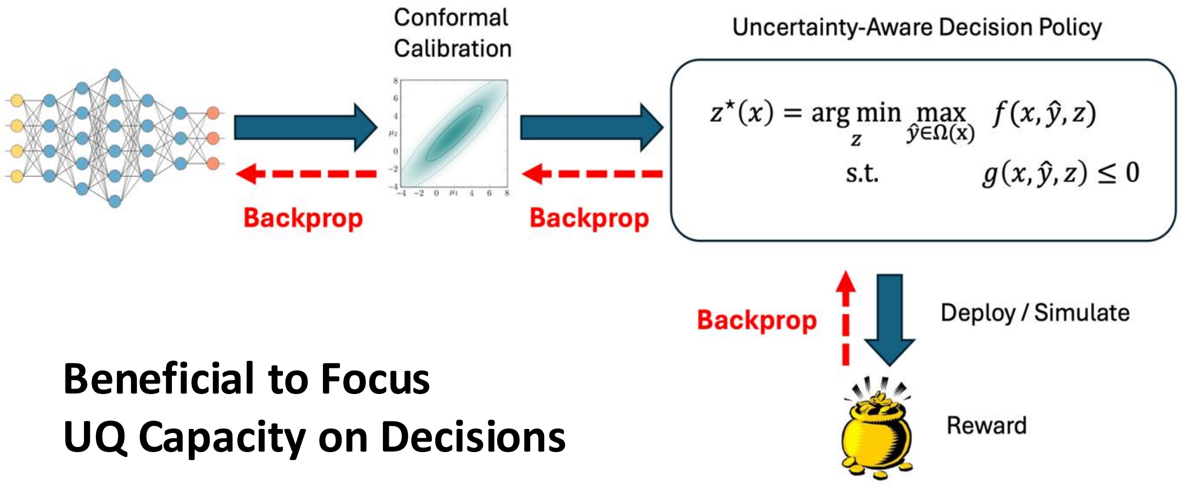
Progress on Scaling to Complex Settings



Calibration is Hard!



Work on Real Applications



Beneficial to Focus UQ Capacity on Decisions


Jason Yang


Chris Yeh


Cloris Cheng


Raul Astudillo


James Bowden


Ayya Alieva

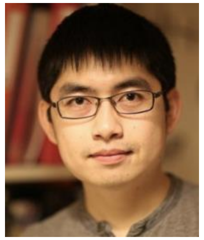

Nico Christianson

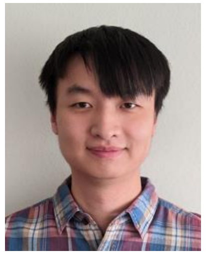

Steve Mayo


Aiden Aceves


Tom Desautels


Frances Arnold


Yuxin Chen


Jialin Song


Adam Wierman

ELEGEN Lawrence Livermore National Laboratory U.S. DEPARTMENT OF ENERGY

- **Active Learning-Assisted Directed Evolution,** Yang, et al., *bioRxiv*
- **Bayesian Optimization with Bayesian Deep Kernel Learning,** Bowden, Yeh, Astudillo, et al. *(in preparation)*
- **Practical Bayesian Algorithm Execution via Posterior Sampling,** Cheng*, Astudillo*, et al., NeurIPS 2024
- **End-to-End Conformal Calibration for Optimization Under Uncertainty**, Yeh, Christianson, et al., *arXiv*
- **Learning to Make Decisions via Submodular Regularization**, Alieva, et al., ICLR 2021