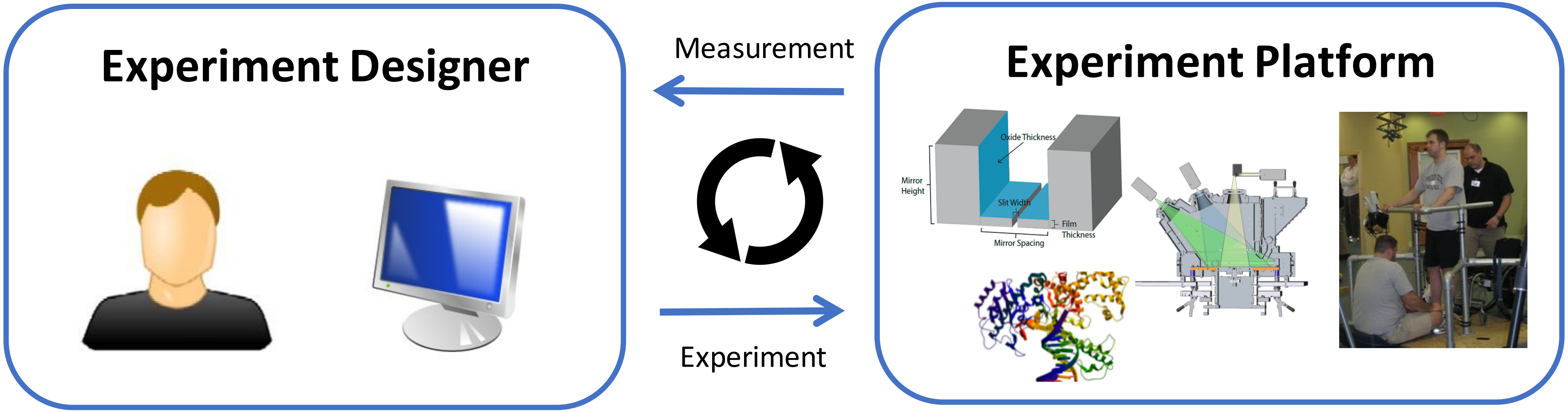
Real-World Bayesian Optimization

Yisong Yue

Caltech

Goal-Oriented Experiment Design

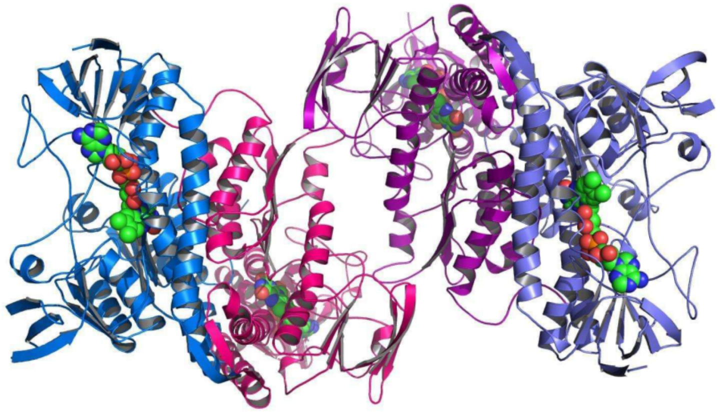


- Iterative & adaptive
- Utility maximizing (find best outcome)

Many Applications (Goal-Oriented)

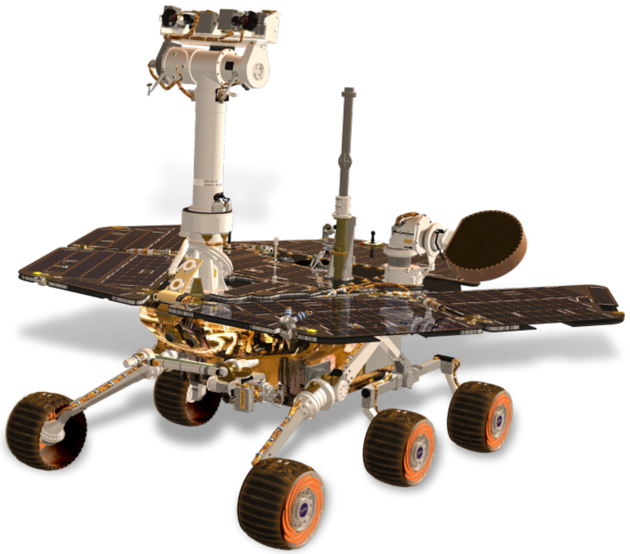
Robotics & Control

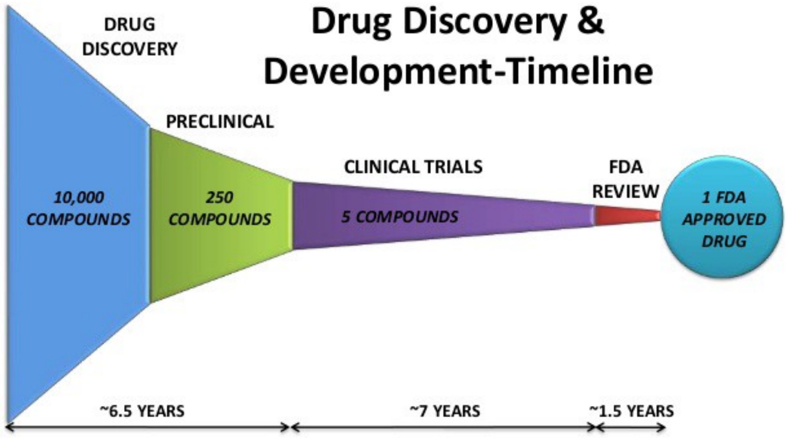
image credit @ mwfarmandfield



Protein Engineering

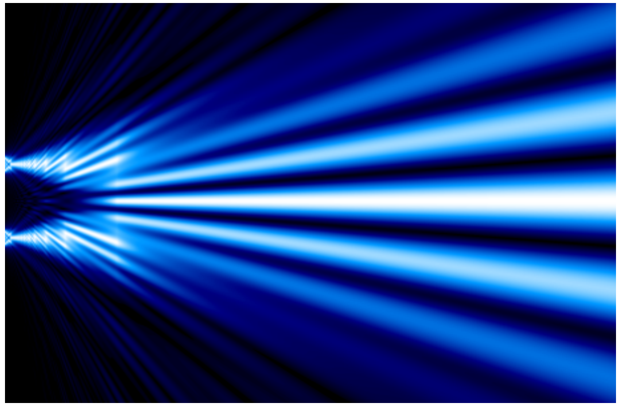
image credit @ creativebiomart





Drug Discovery

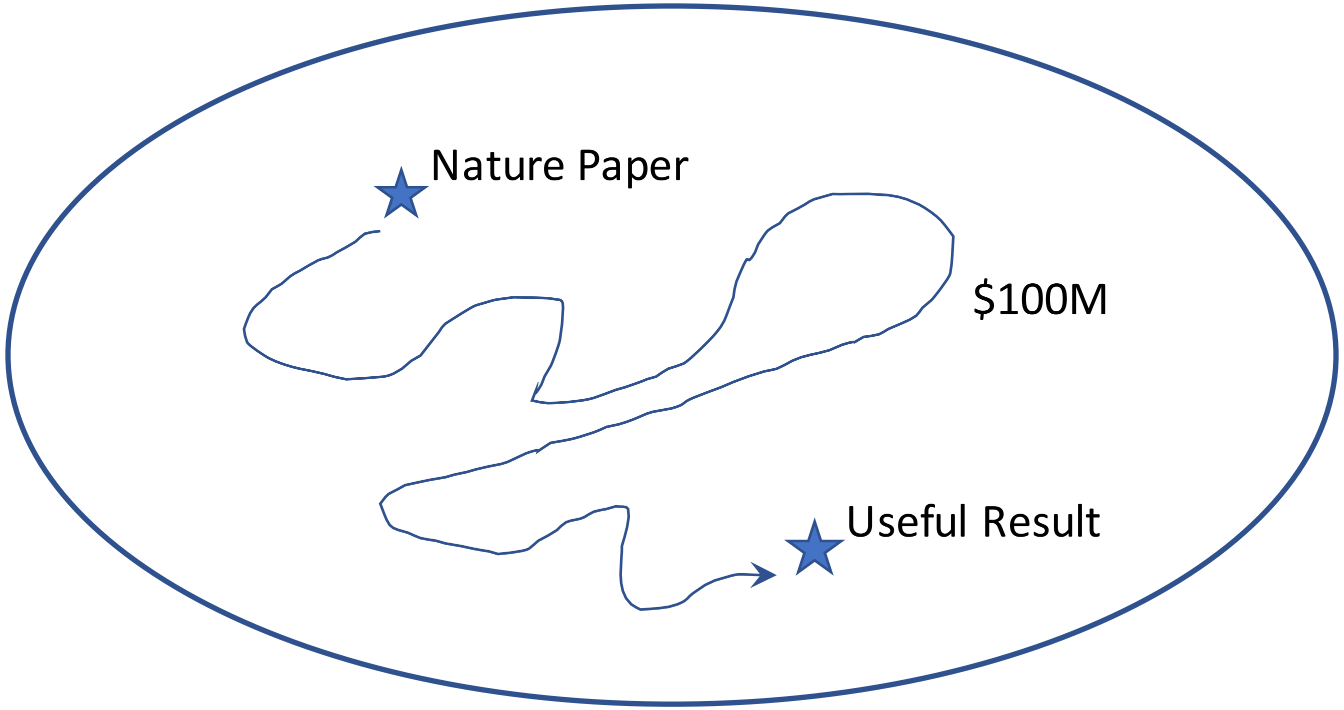
from Slideshare



Material Science

image credit @ phys.org

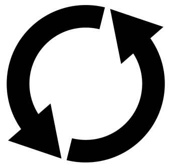
AI for Goal-Oriented Experiment Design



Hypothesis Space

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

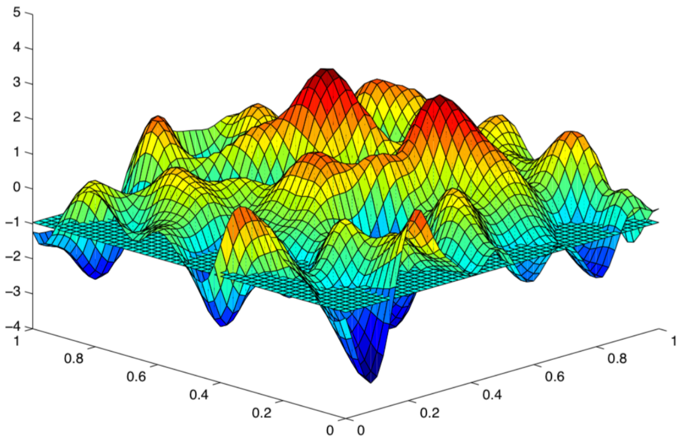
- **Goal:** Utility over time
- E.g., recommender systems
- Maximize utility over time

Focus of Talk

Bayesian Optimization Example

X = space of proteins

F(x) = fitness landscape



Frances Arnold

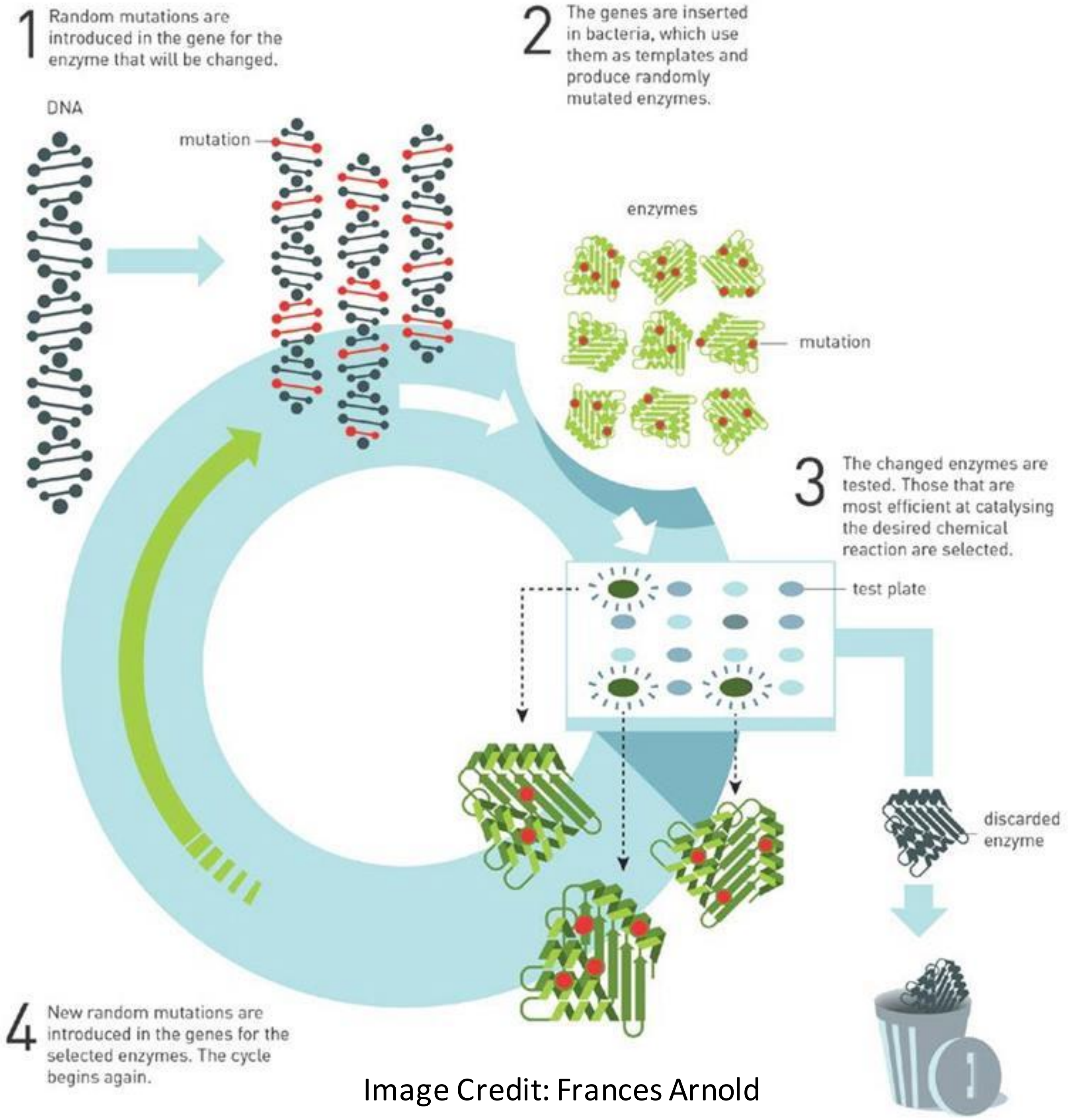


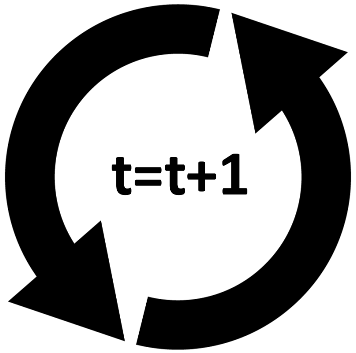
Image Credit: Frances Arnold

Learning Setup

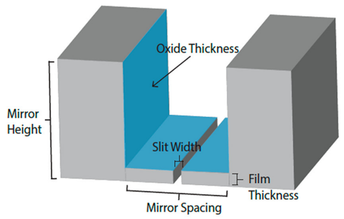
Given: input space X
Unknown: fitness F(x)=y
Maintain: posterior P(F|D) (D=measurements)

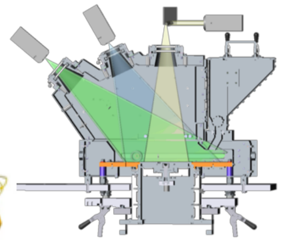
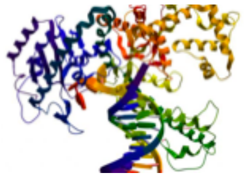
Measure y_t

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: $\text{argmax}_x\, f(x)$, $f \sim P(F|D)$

Algorithmic & Theoretical Questions
(see papers for details)

- Analyze convergence to F(x*)?
- Guarantee side constraints (e.g., safety)?
- Corrupted or indirect measurements?
- Efficiently search combinatorial design spaces?
- Incorporate domain knowledge such as physics?

Real-World Bayesian Optimization

Safety, Preference

Multi-Fidelity

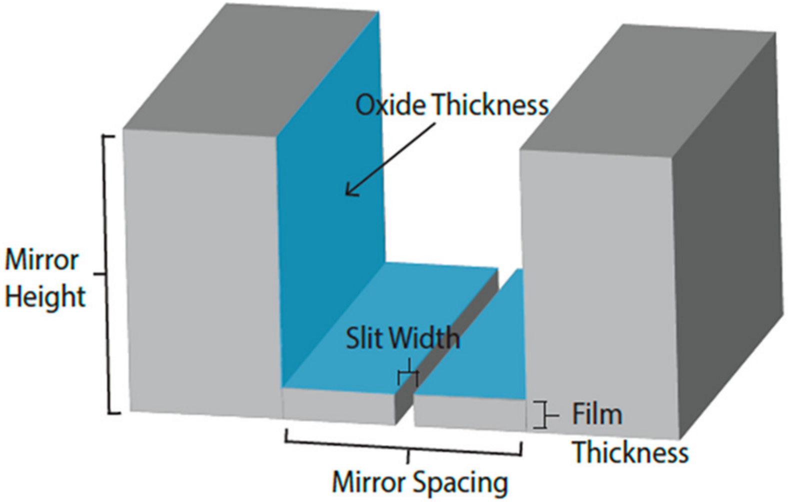
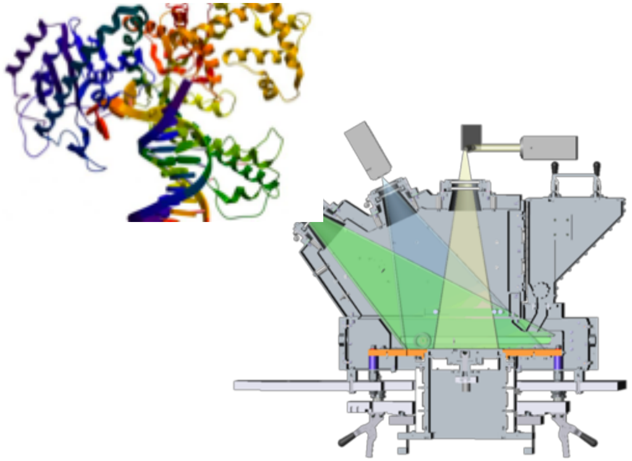


Photo credit: Yury Tokpanov

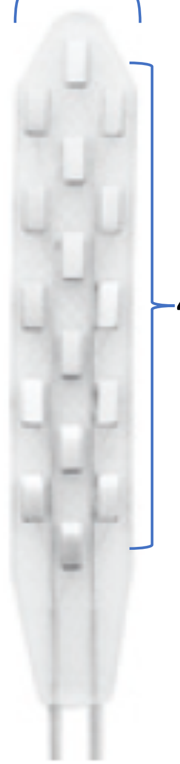
Combinatorial, Physics, …



Treating Lower Spine Injuries

10 mm

Medtronic human array



49 mm

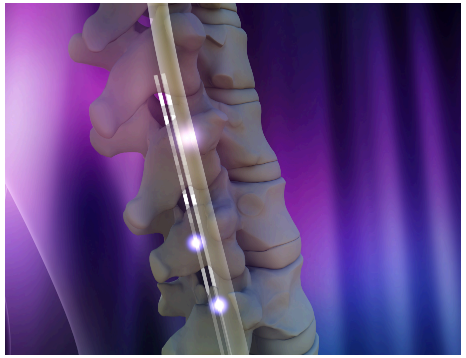
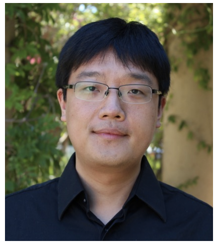


Image source: williamcapicottomd.com

SCI Patient



Yanan Sui

Joel Burdick

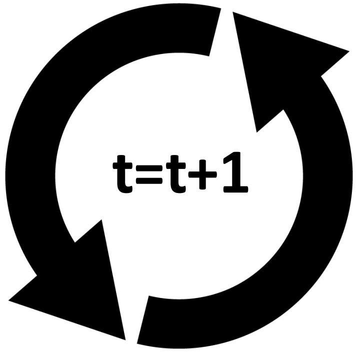
Each patient is unique

10^9 possible configurations!

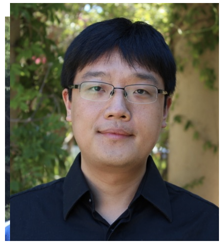
Learning Setup

Update posterior
P(F|D)

Receive Response y_t



Apply Stimulus x_t



Yanan
Sui

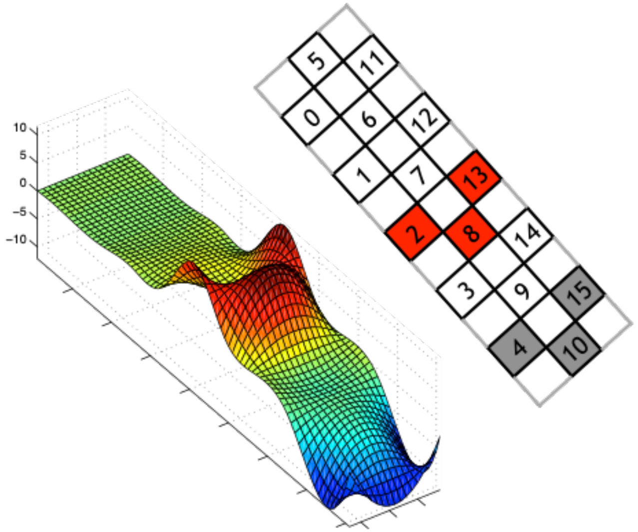
Joel
Burdick

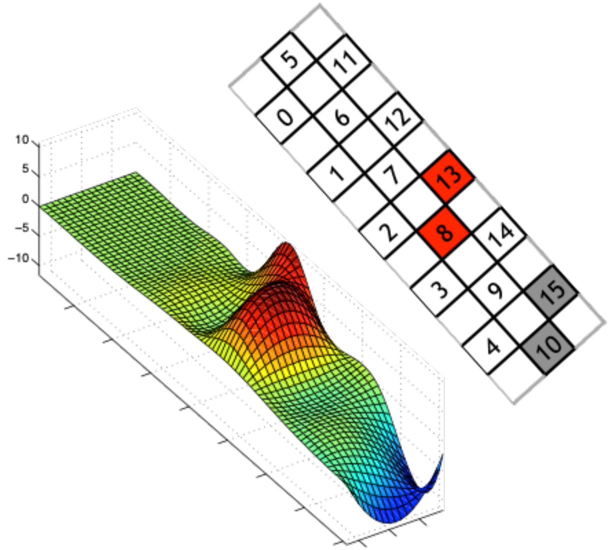
Electrode Array

SCI Patient

Challenges

- Many actions
  - 10^6 to 10^9
- Measuring utility difficult
- Safety





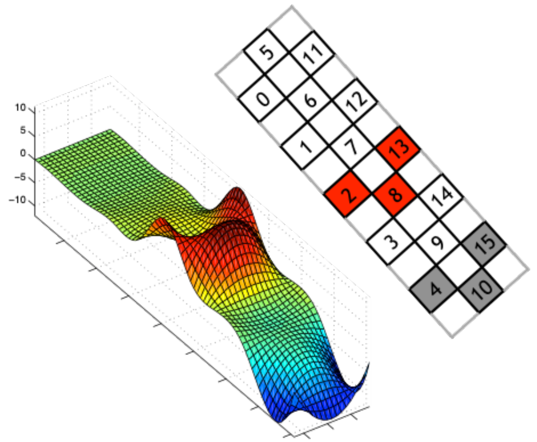
Modeling Correlations: Gaussian Processes

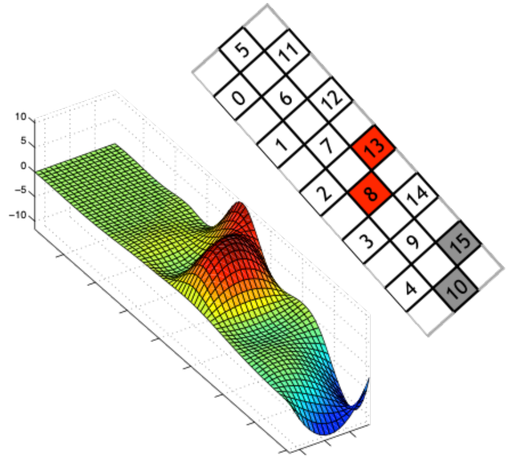
- Defined by $GP(\mu, k)$ ← **Covariance "kernel"**

  ↑ **Mean function**

- Sample a function: $f \sim GP(\mu, k)$
  - Expected value of $f(x)$ is $\mu(x)$
  - Correlation of $f(x_1)$ & $f(x_2)$ is $k(x_1, x_2)$

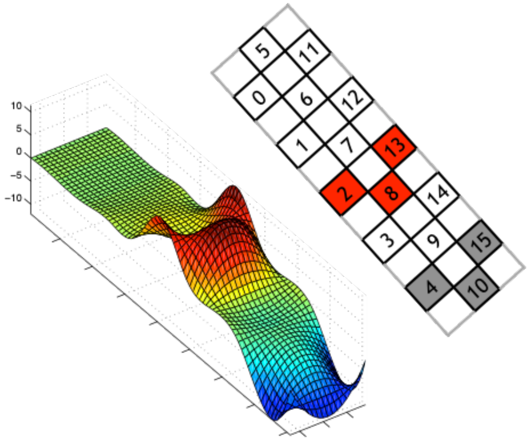
- Finite input domain: (e.g., 10 choices of x)
  - Reduces to multivariate Gaussian distribution



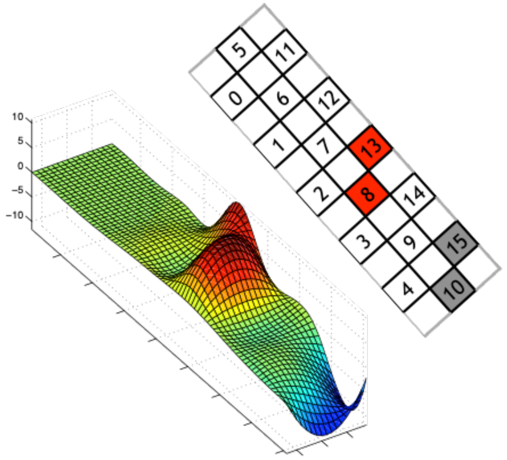


Benefits of Gaussian Processes

- Reason about uncertainty
  - What is the spread of outcomes for f(x)?

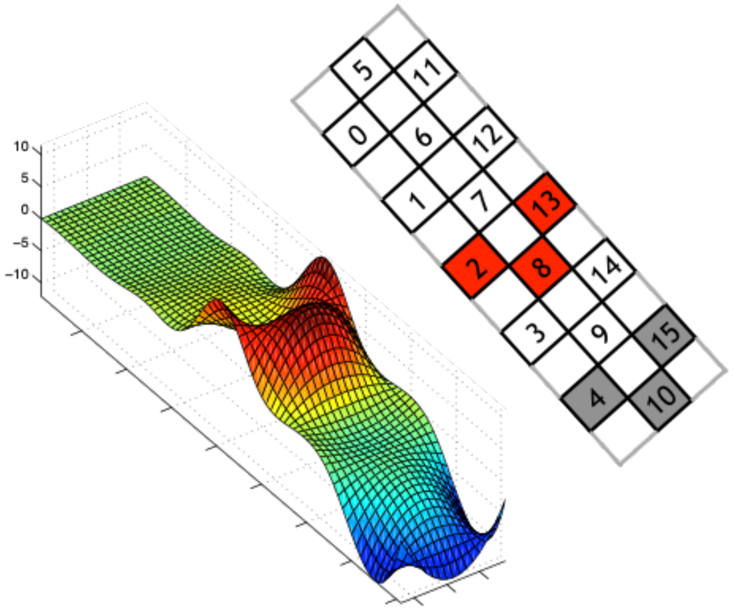


- Correlations over input space
  - Measuring $f(x_1)$ gives information on $f(x_2)$



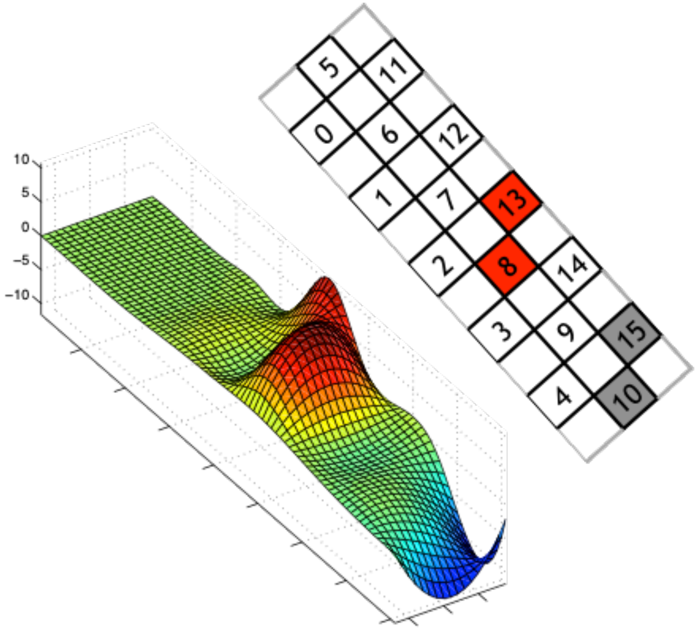
- Work with domain experts to build correlations

Measurements via Preference Feedback



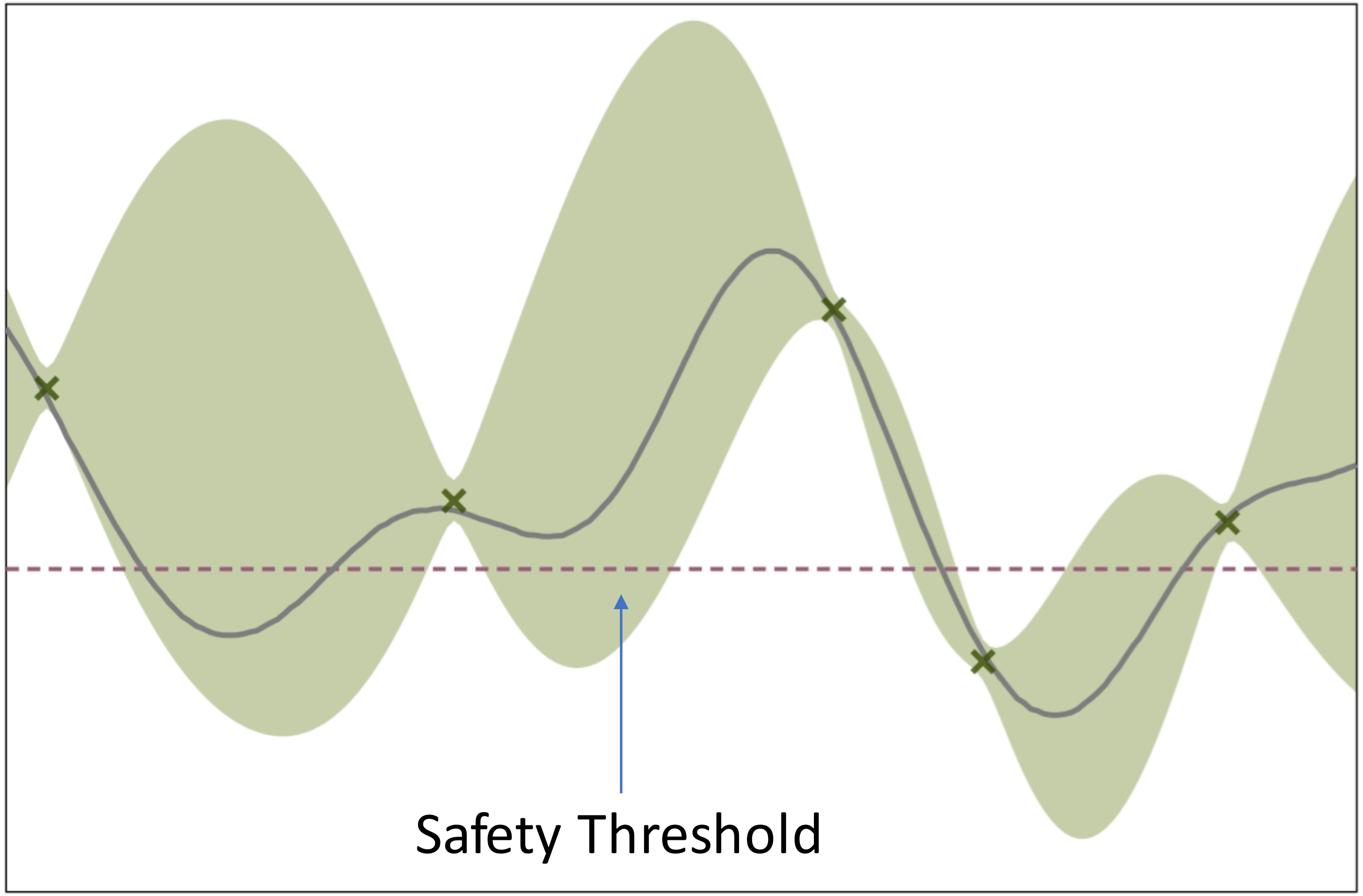
?

$\succ$



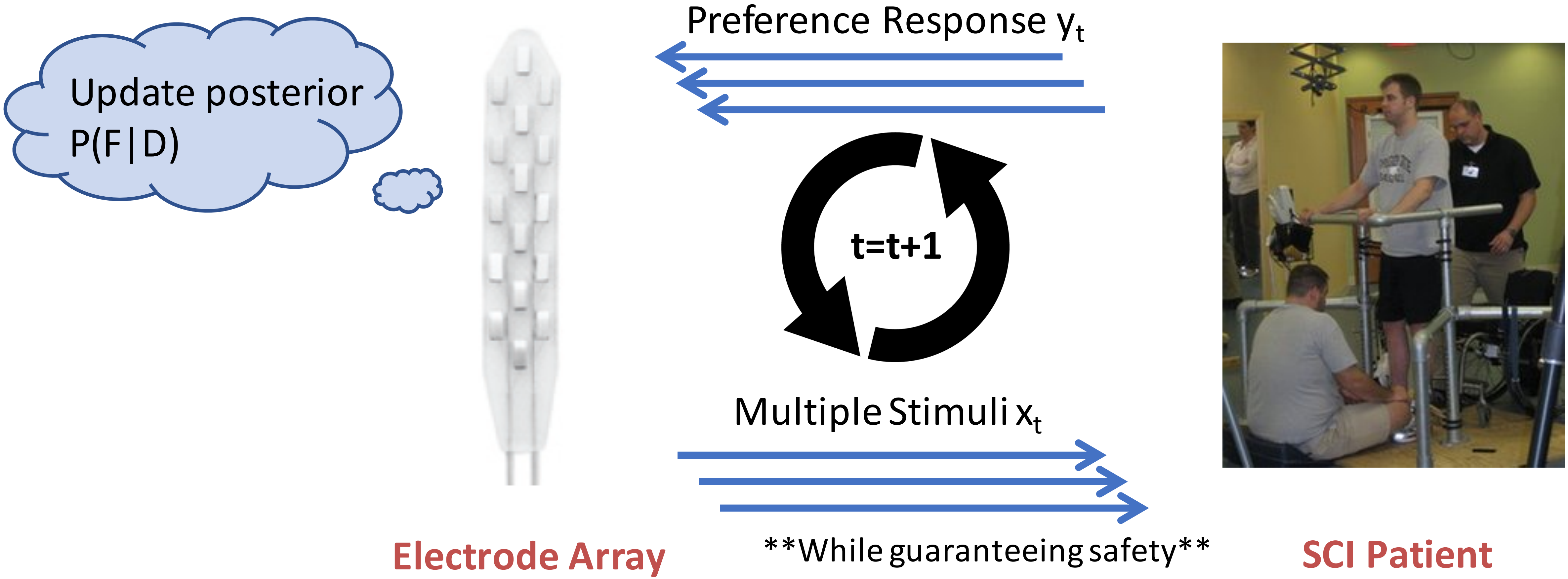
Multi-dueling Bandits with Dependent Arms, Sui, Zhuang, Burdick & Yue, UAI 2017
Correlational Dueling Bandits with Application to Clinical Treatment in Large Decision Spaces, Sui, Yue & Burdick, IJCAI 2017

Gaussian Process Safety Model



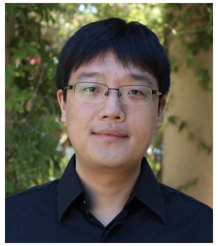
Stagewise Safe Bayesian Optimization with Gaussian Processes, Sui, Zhuang, Burdick & Yue, ICML 2018

Full Learning Setup

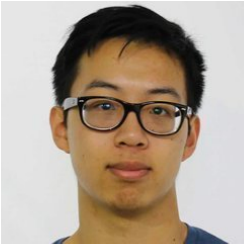


Multi-dueling Bandits with Dependent Arms, Sui, Zhuang, Burdick & Yue, UAI 2017
Correlational Dueling Bandits with Application to Clinical Treatment in Large Decision Spaces, Sui, Yue & Burdick, IJCAI 2017
Stagewise Safe Bayesian Optimization with Gaussian Processes, Sui, Zhuang, Burdick & Yue, ICML 2018

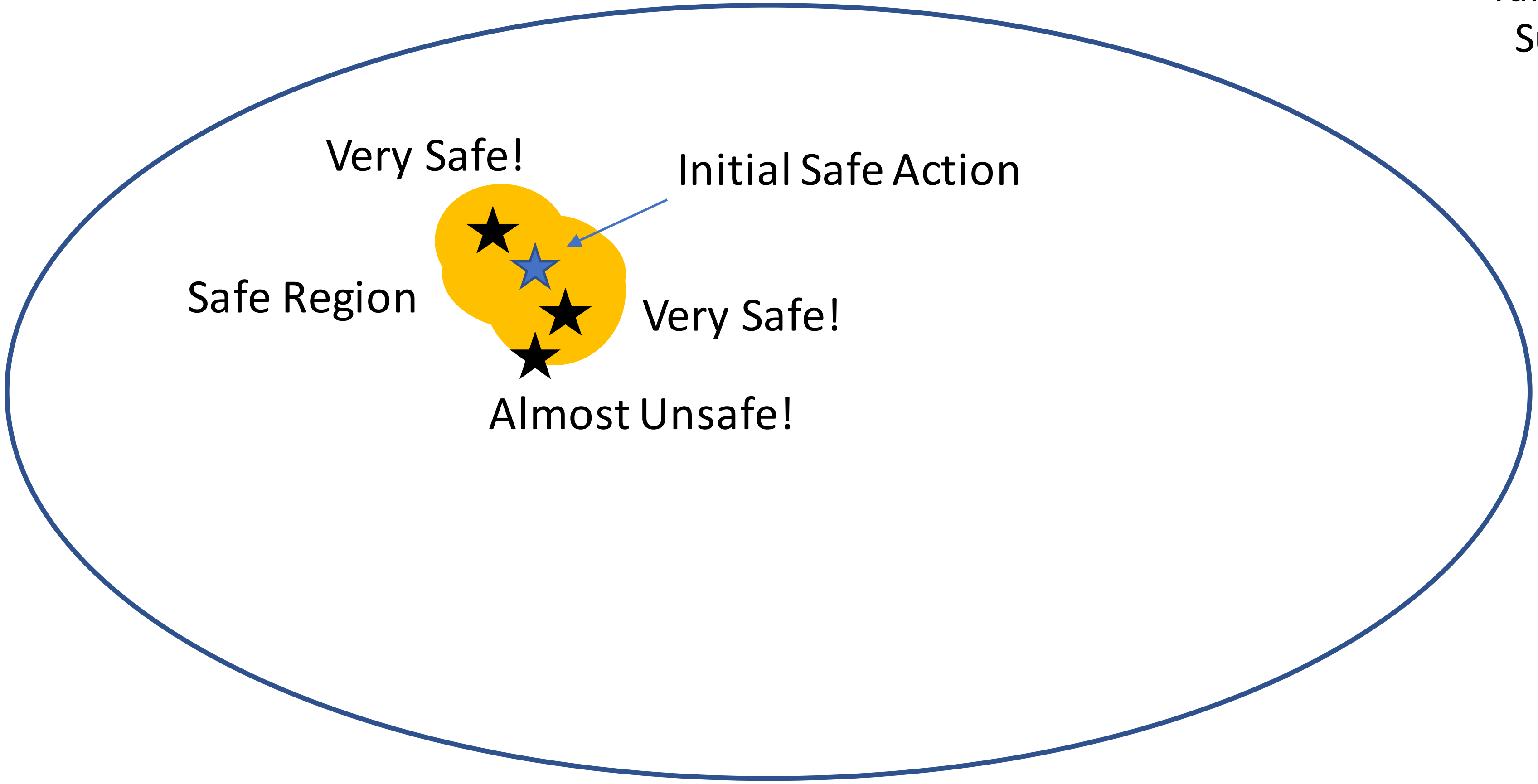
Algorithmic Insights



Yanan Sui



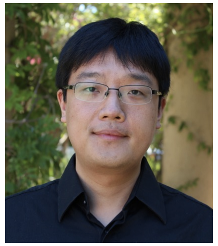
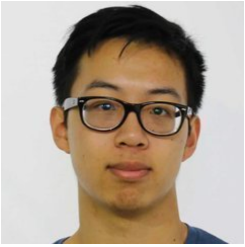
Vincent Zhuang



Hypothesis Space

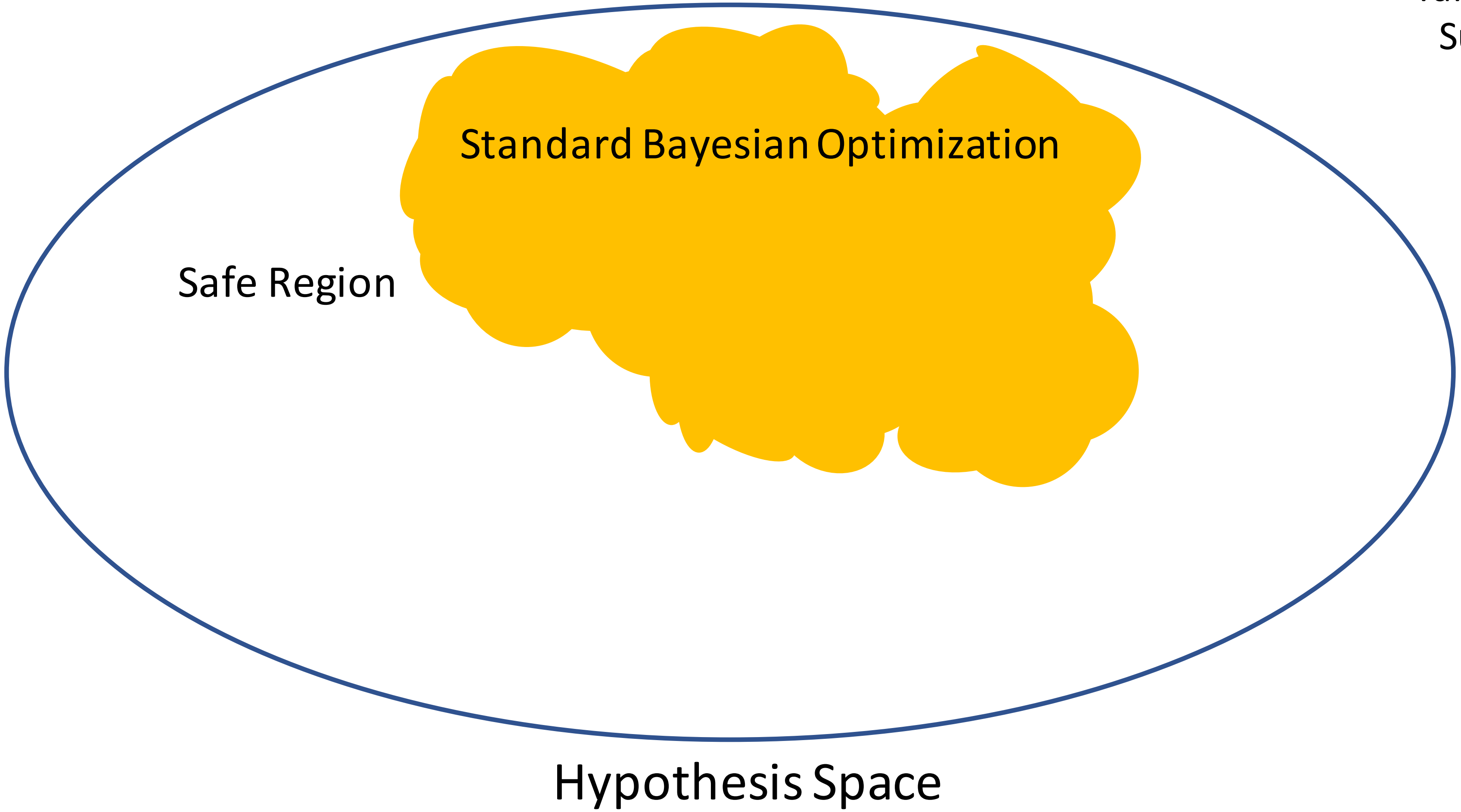
Stagewise Safe Bayesian Optimization with Gaussian Processes, Sui, Zhuang, Burdick & Yue, ICML 2018

Algorithmic Insights

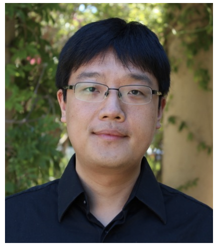
Yanan Sui

Vincent Zhuang

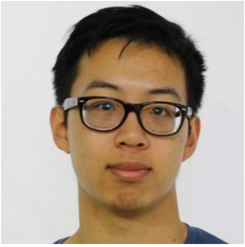


Stagewise Safe Bayesian Optimization with Gaussian Processes, Sui, Zhuang, Burdick & Yue, ICML 2018

Algorithmic Insights



Yanan Sui



Vincent Zhuang

Standard Bayesian Optimization

First Maximize Safety Region

- Optimistic in the face of uncertainty
- Identify reachable safety region
- Approximately maximal w/ convergence guarantees

Meta-Framework

- Leverage any Bayesian Optimization alg.
- Inherit guarantees

Stagewise Safe Bayesian Optimization with Gaussian Processes, Sui, Zhuang, Burdick & Yue, ICML 2018

Clinical Experiments

Real-World Bayesian Optimization

Safety, Preference

Multi-Fidelity

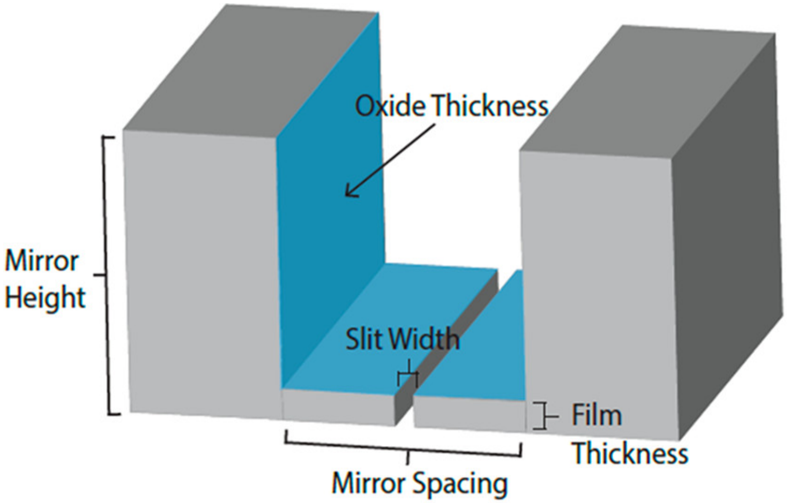
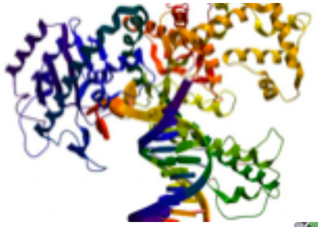
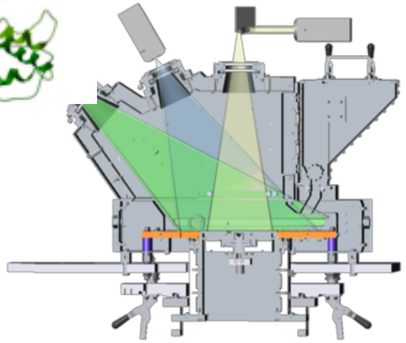


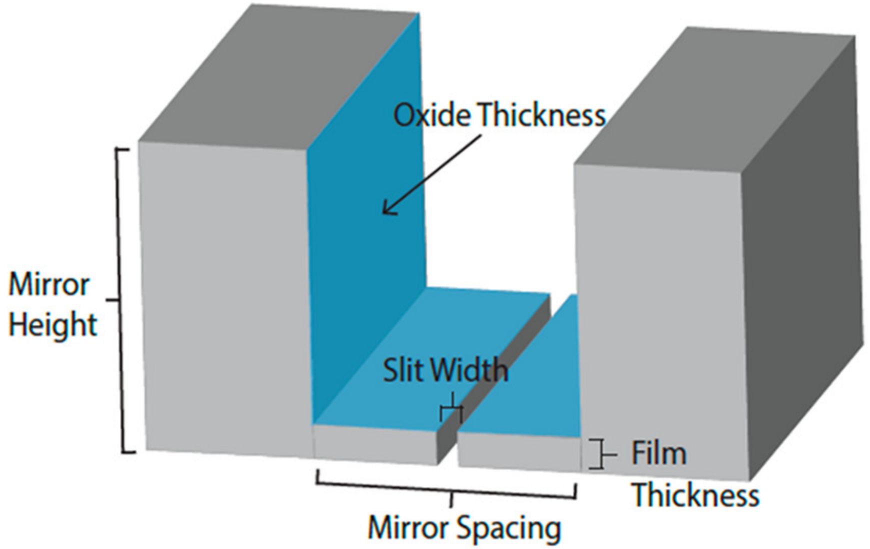
Photo credit: Yury Tokpanov

Combinatorial, Physics, ...





Nano-photonics Structure Design

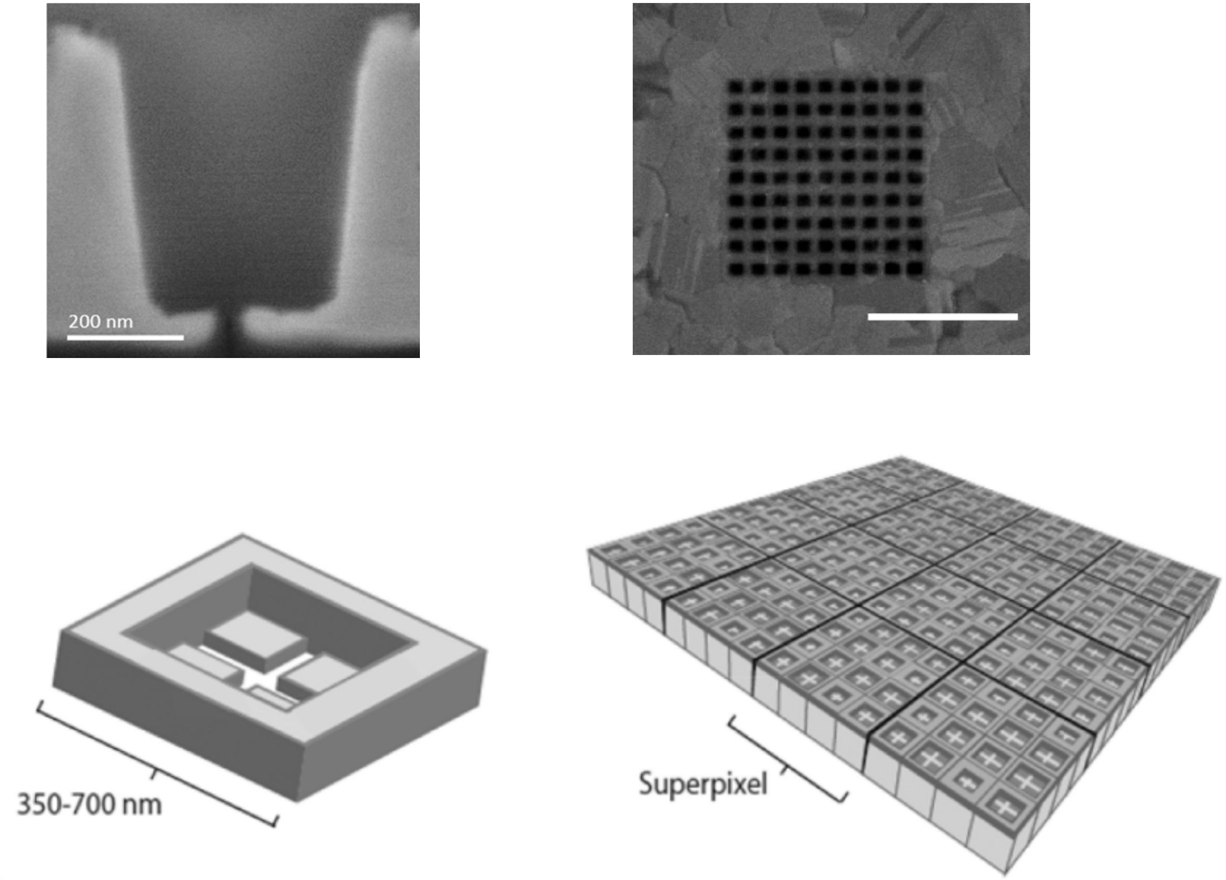



Jialin Song


Yuxin Chen

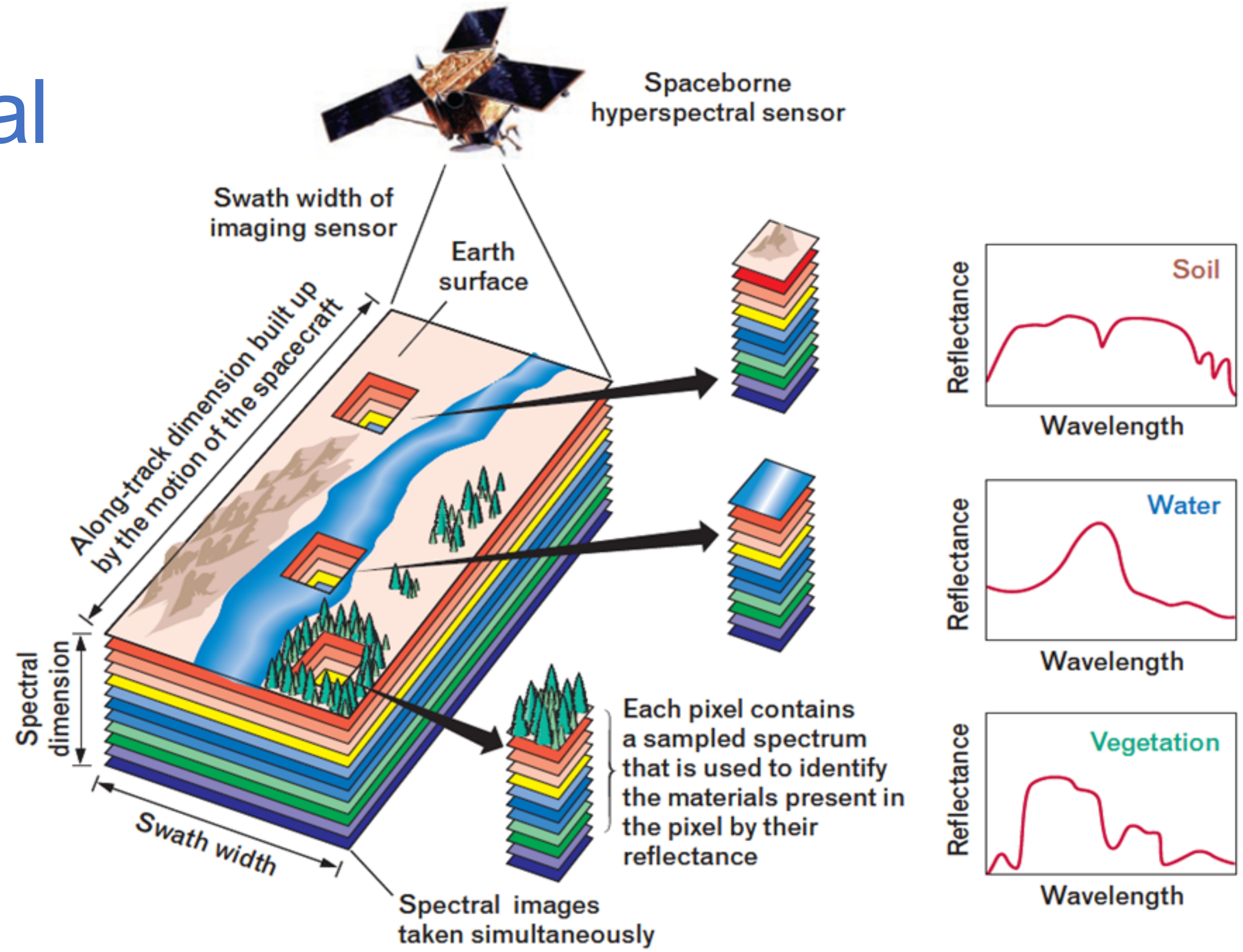

Yury Tokpanov


Harry Atwater



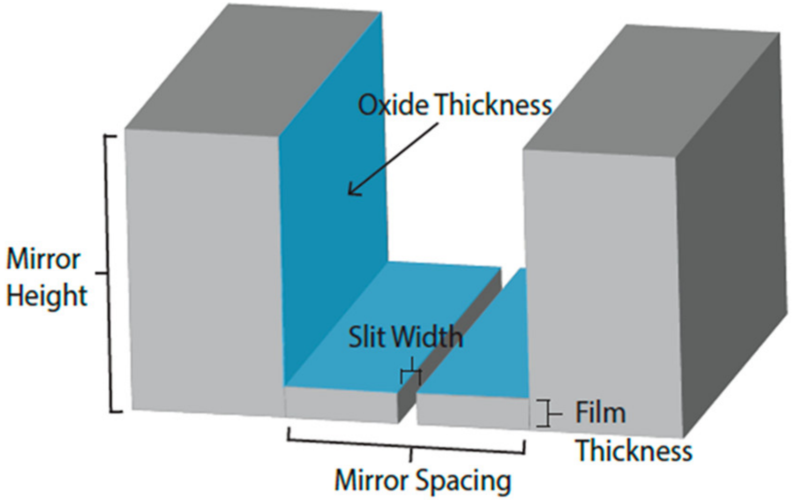
Fleischman et al.: https://doi.org/10.1021/acsphotonics.8b01634

Hyperspectral Imaging

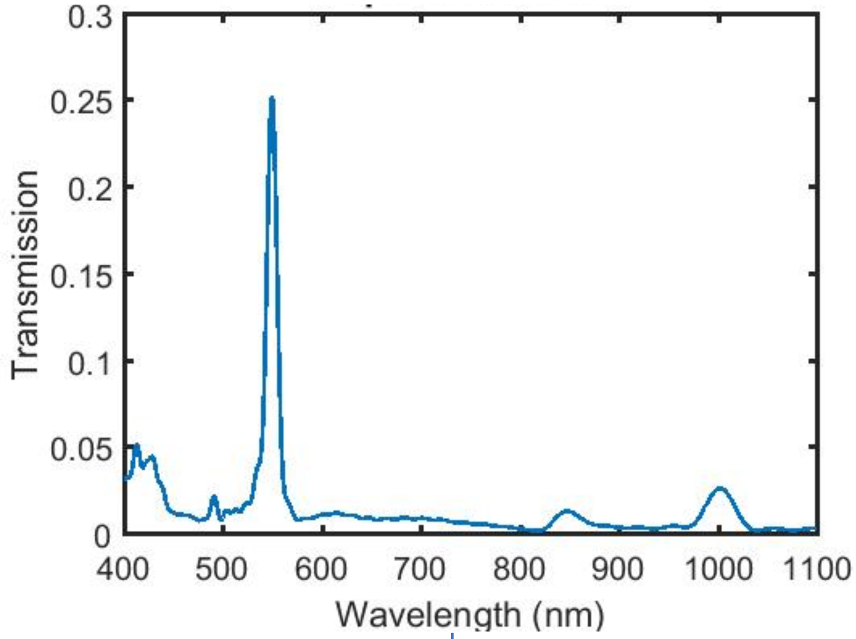


https://doi.org/10.1016/B978-0-444-63638-6.00006-1

Fitness Function (Figure of Merit)



Maxwell's equations
FDTD solver

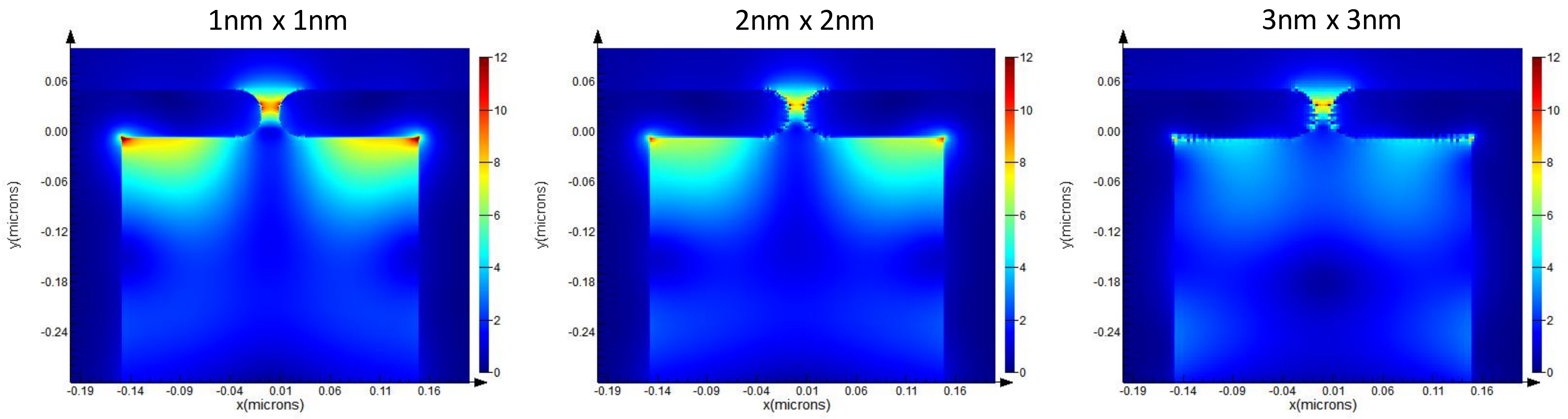


Fitness Function: $FOM = \frac{\Delta wav}{c.wav} + \frac{c.peak}{peak} + \frac{noise}{c.noise} + \frac{FWHM}{c.FWHM}$

Image Credit: Yury Tokpanov

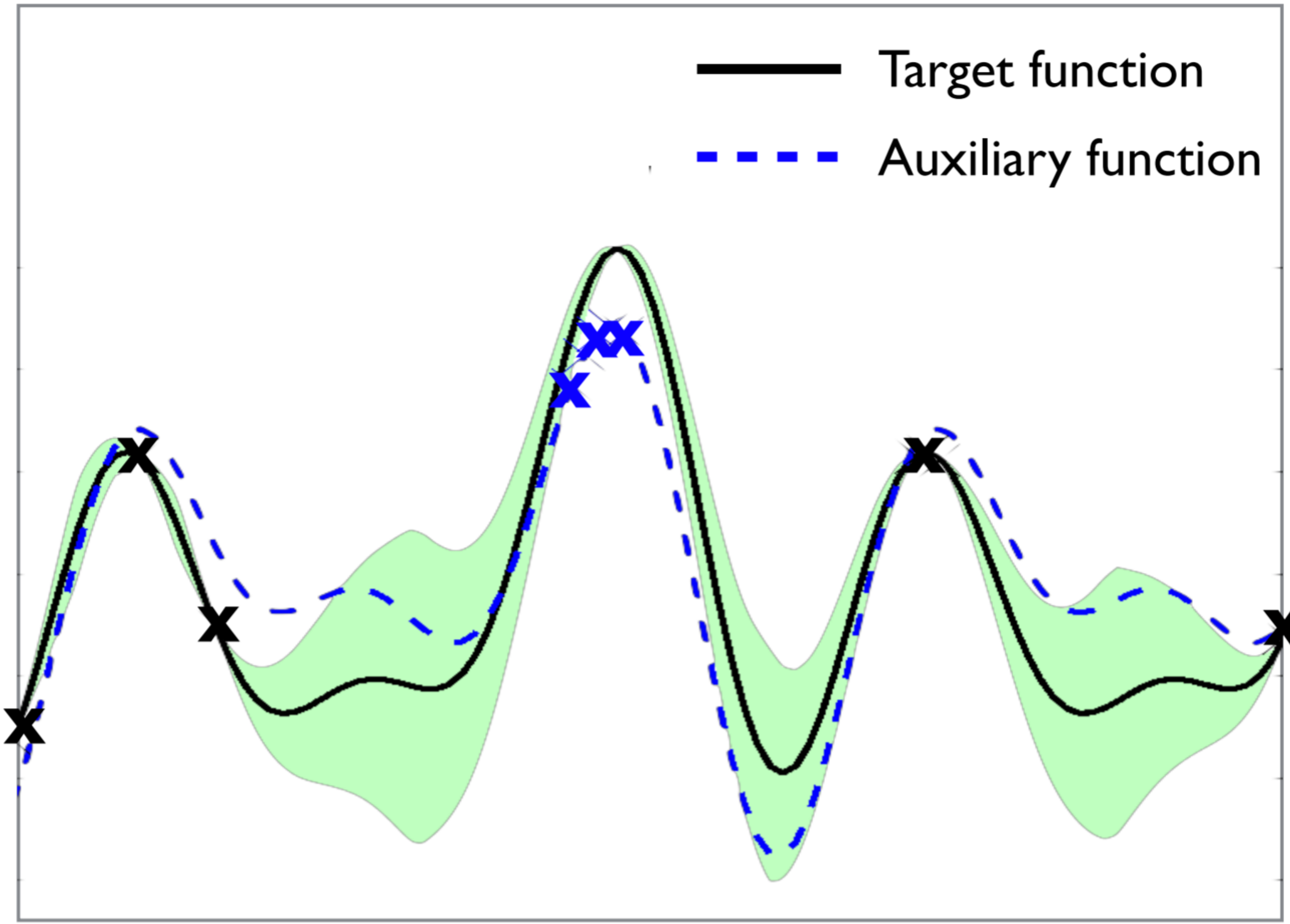
Multi-Fidelity Simulations

- Solve Maxwell’s equations
- Fidelity depends on temporal and spatial resolution
- **Do we need to accurately simulate bad structures?**

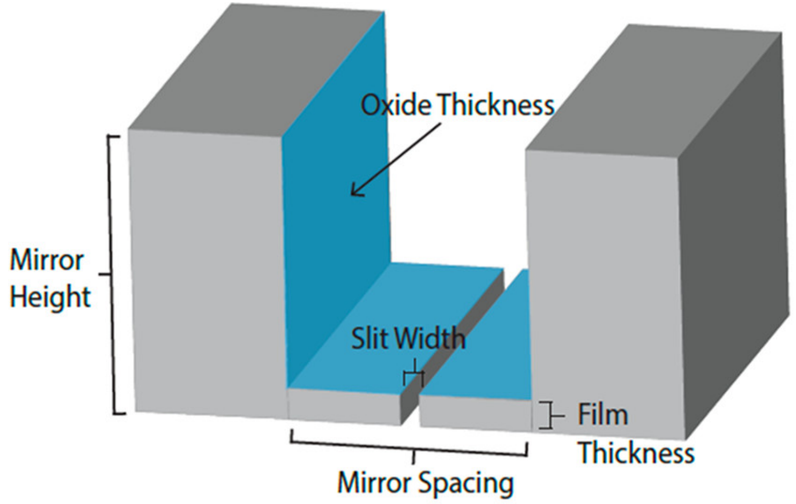


Electric field profiles at 550nm for different mesh sizes

Image Credit: Yury Tokpanov

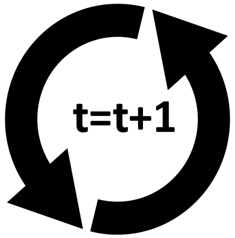

Target function
Auxiliary function

Learning Setup

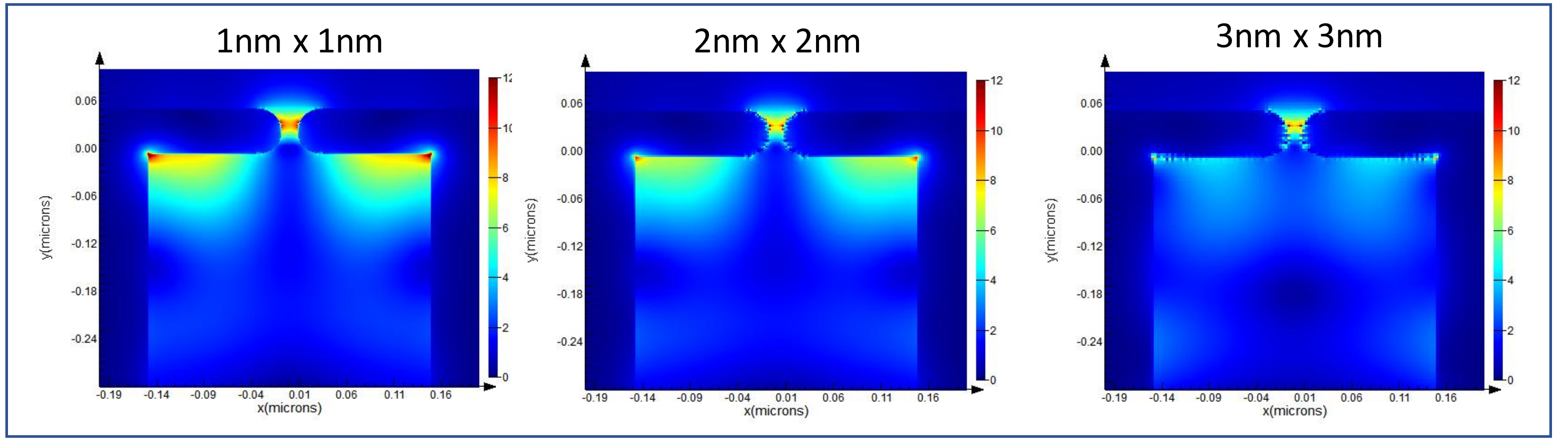


Update posterior P(F|D)

Choose Design x_t & Fidelity Level ℓ_t



Measure Response y_t



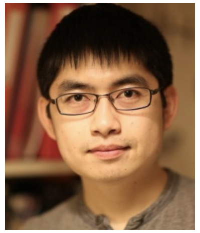
Jialin Song

Yuxin Chen

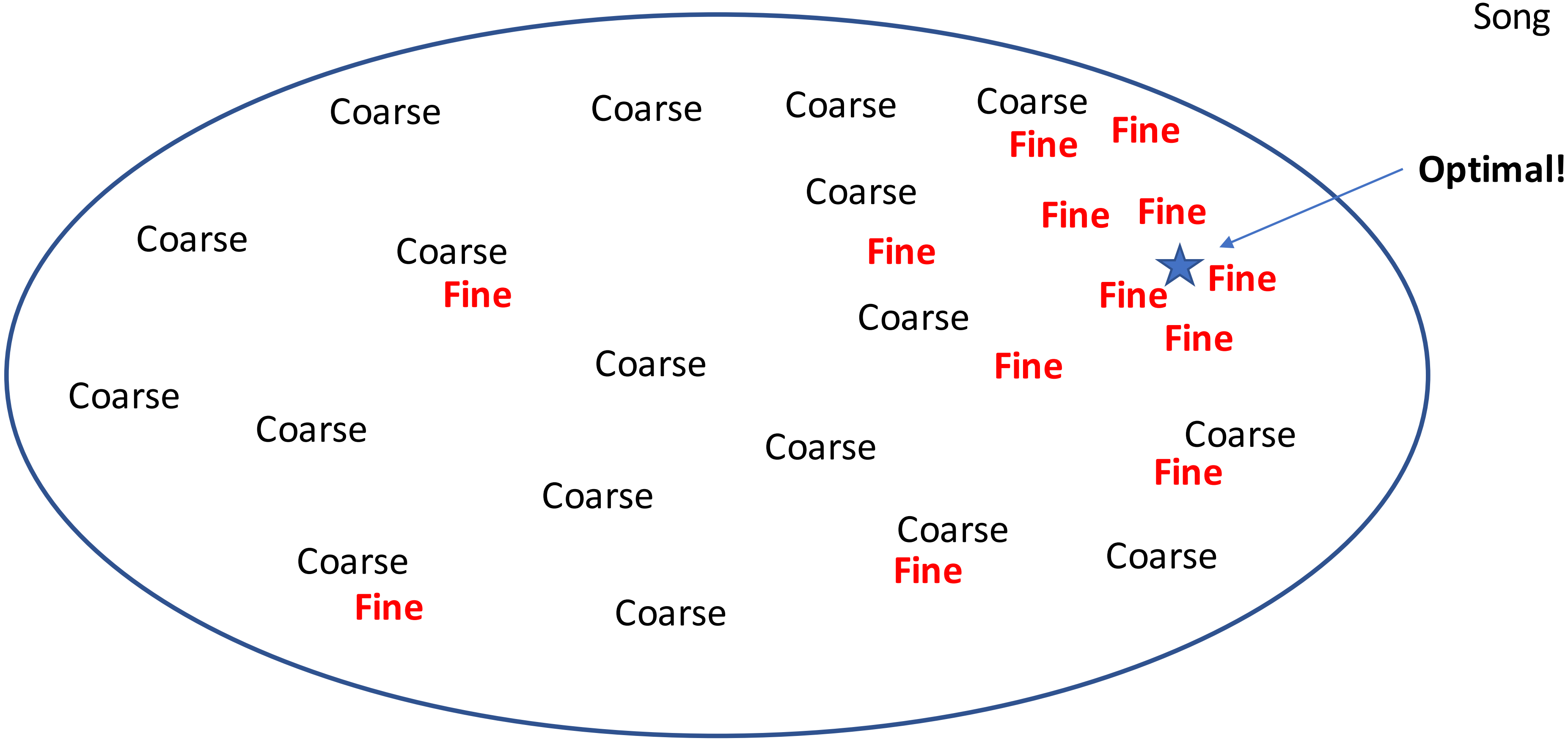
A General Framework for Multi-fidelity Bayesian Optimization with Gaussian Processes, Jialin Song et al., AISTATS 2019

Algorithmic Insights

Jialin Song



Yuxin Chen

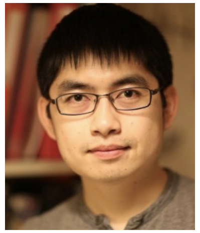


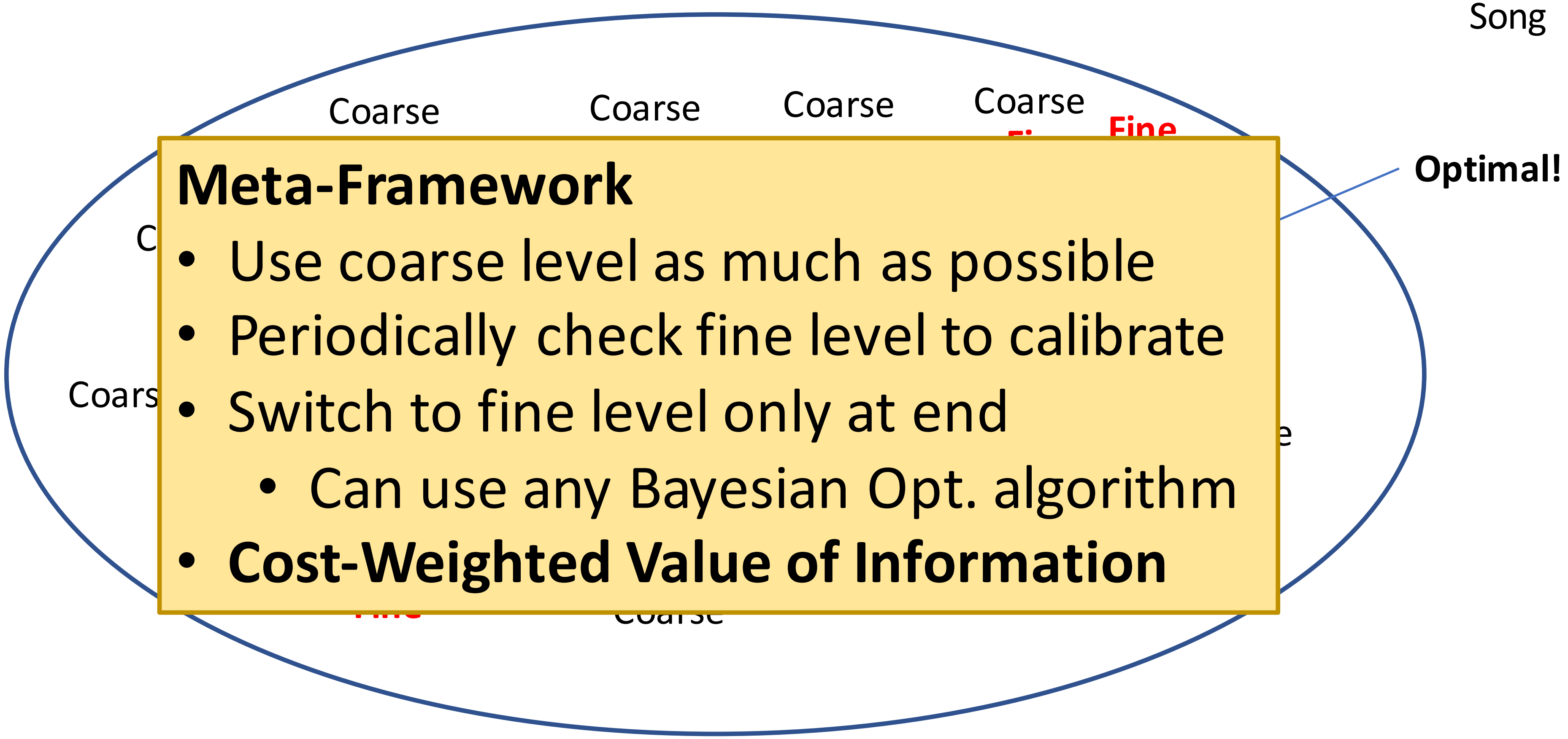
Hypothesis Space

A General Framework for Multi-fidelity Bayesian Optimization with Gaussian Processes, Jialin Song et al., AISTATS 2019

Algorithmic Insights


Jialin Song


Yuxin Chen



Meta-Framework

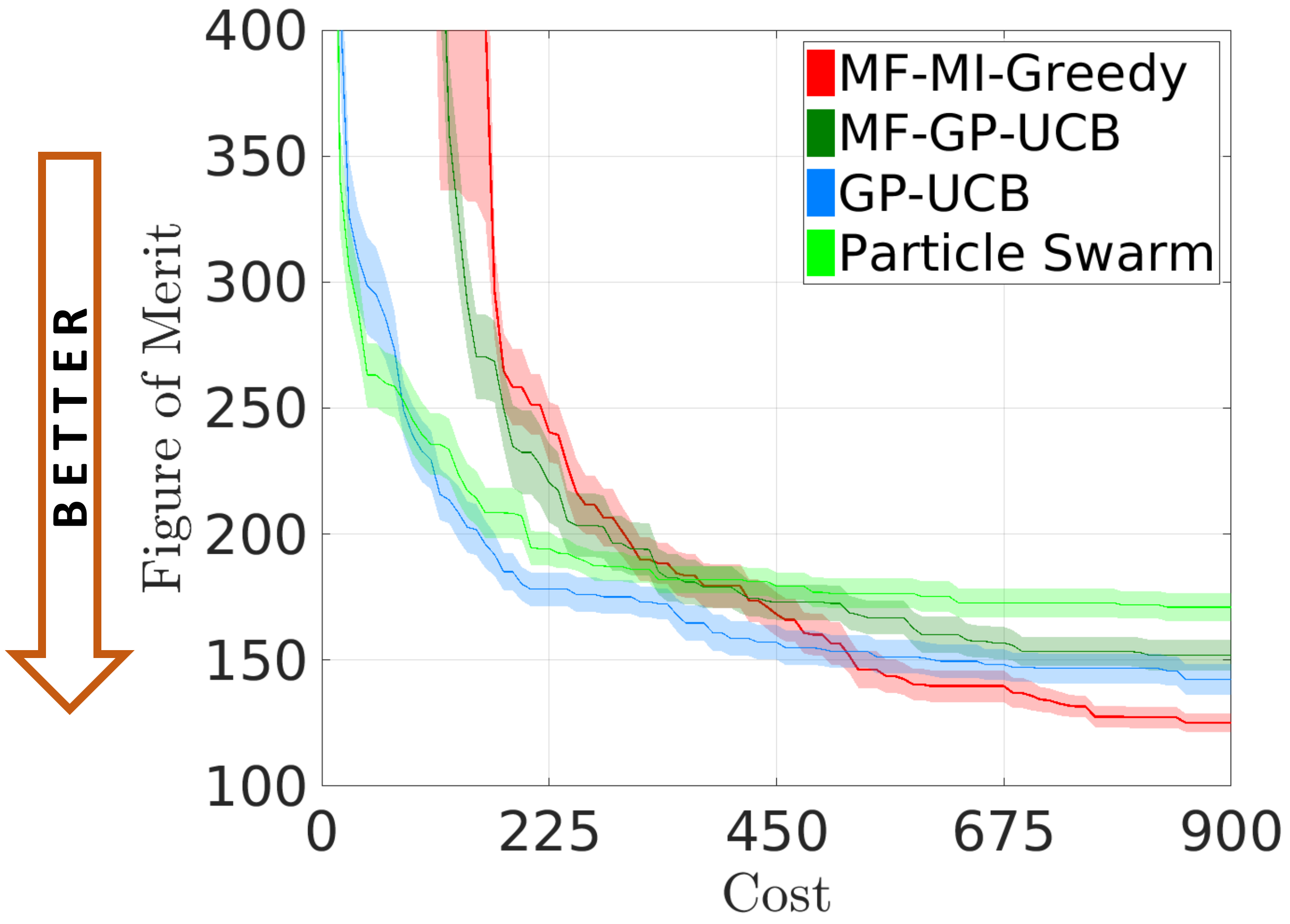
- Use coarse level as much as possible
- Periodically check fine level to calibrate
- Switch to fine level only at end
  - Can use any Bayesian Opt. algorithm
- **Cost-Weighted Value of Information**

Hypothesis Space

A General Framework for Multi-fidelity Bayesian Optimization with Gaussian Processes, Jialin Song et al., AISTATS 2019

Results

- 3 fidelities
- Balances different costs
- State-of-the-art performance



Optimizing Photonic Nanostructures via Multi-fidelity Gaussian Processes
Song, Tokpanov, Chen, Fleischman, Fountaine, Atwater, Yue, 2018

Real-World Bayesian Optimization

Safety, Preference

Multi-Fidelity

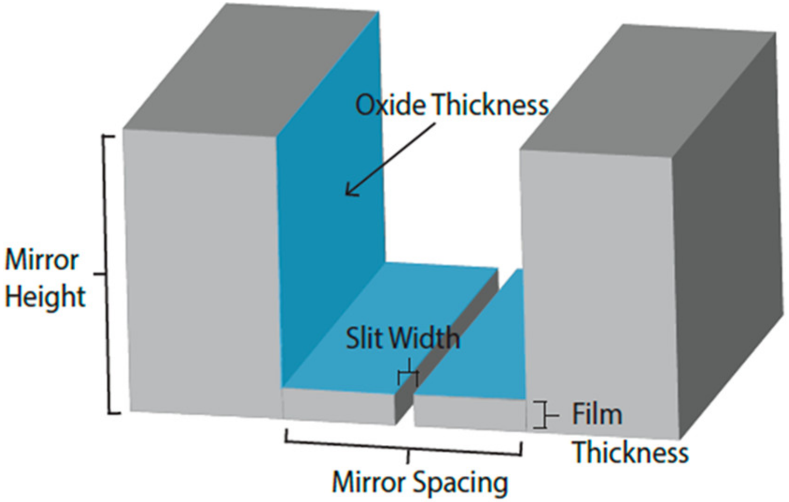
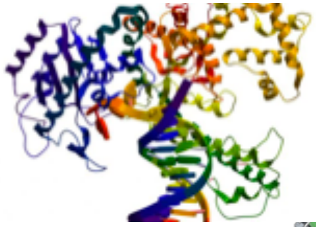
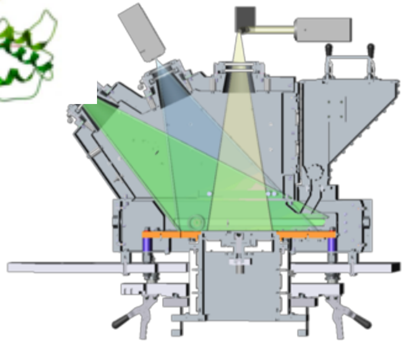


Photo credit: Yury Tokpanov

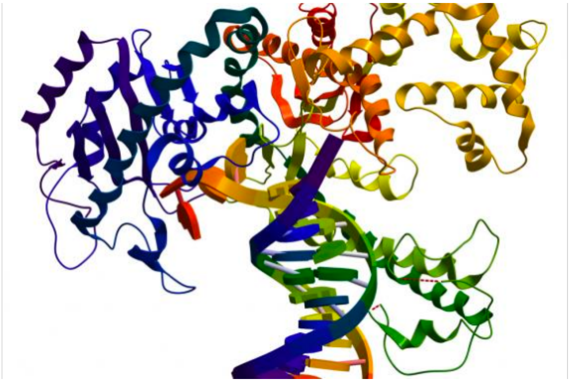
Combinatorial, Physics, …

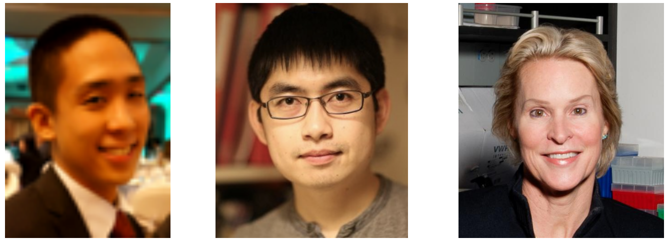




Batched Stochastic Bayesian Optimization

- Start with initial amino acid sequence
- Choose with sites to mutate
- Mutations are probabilistic
- **Combinatorial structure in experiment design**

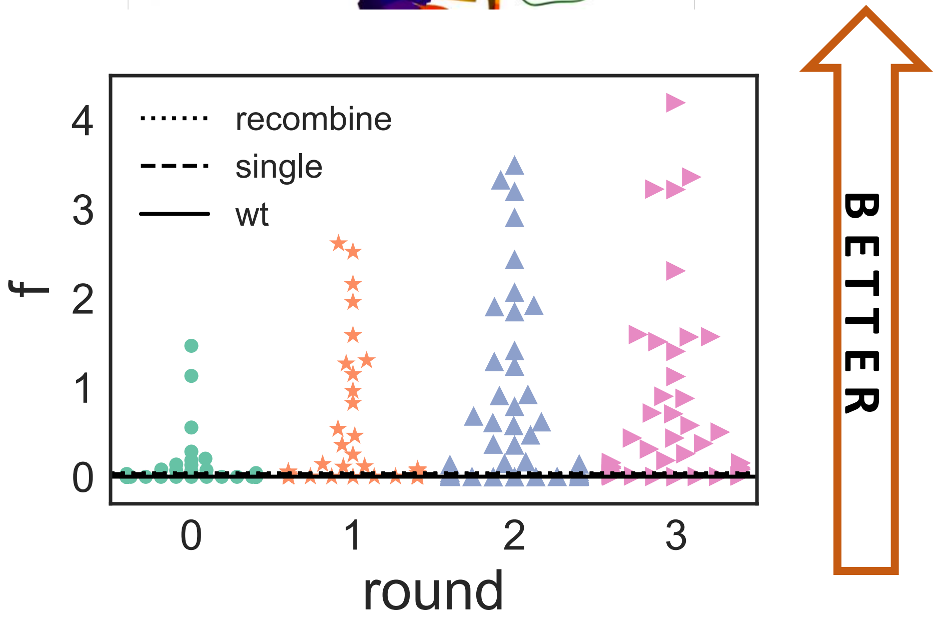




Kevin Yang

Yuxin Chen

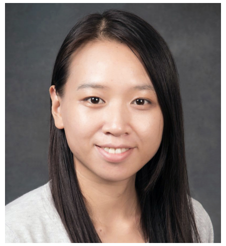
Frances Arnold



Batched Stochastic Bayesian Optimization, Yang, Chen, Lee, Yue, AISTATS 2019

Any Many More…

(physics, non-Bayesian)



Angie Liu

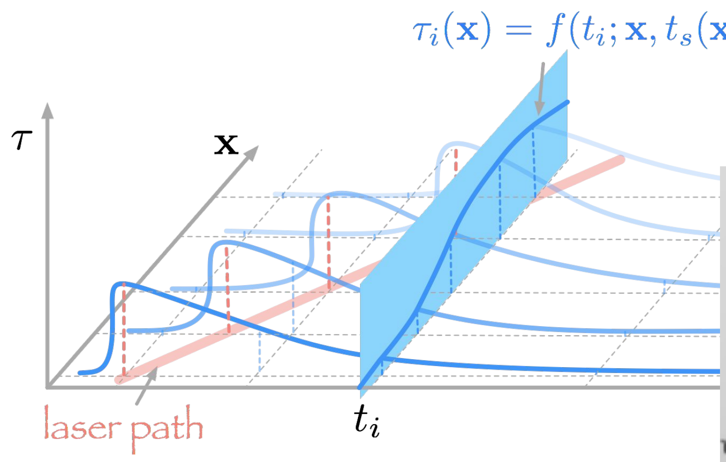
Guanya Shi

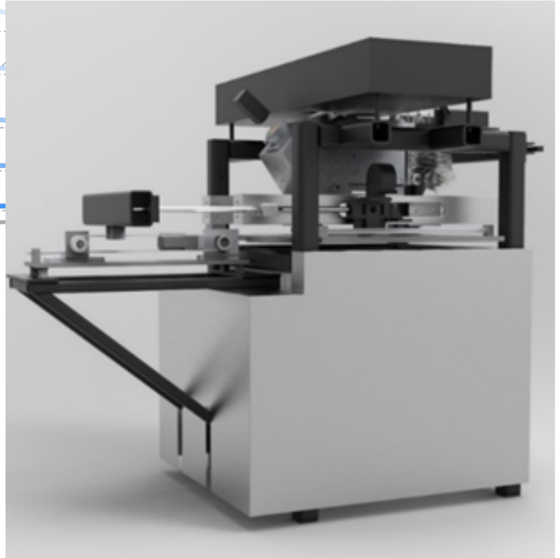
Yuxin Chen

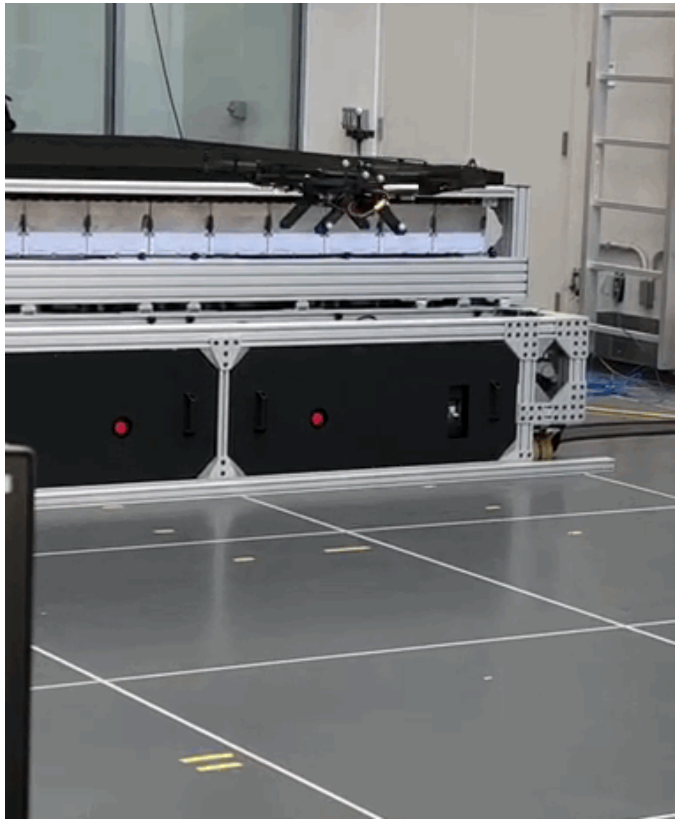
Ufuk Topcu

…

Interactive Controller Calibration



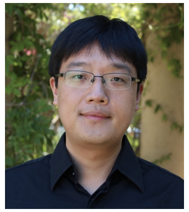




Robust Regression for Safe Exploration in Control, Angie Liu, Guanya Shi, et al., arxiv

Any Many More…
(subjective & dynamical)

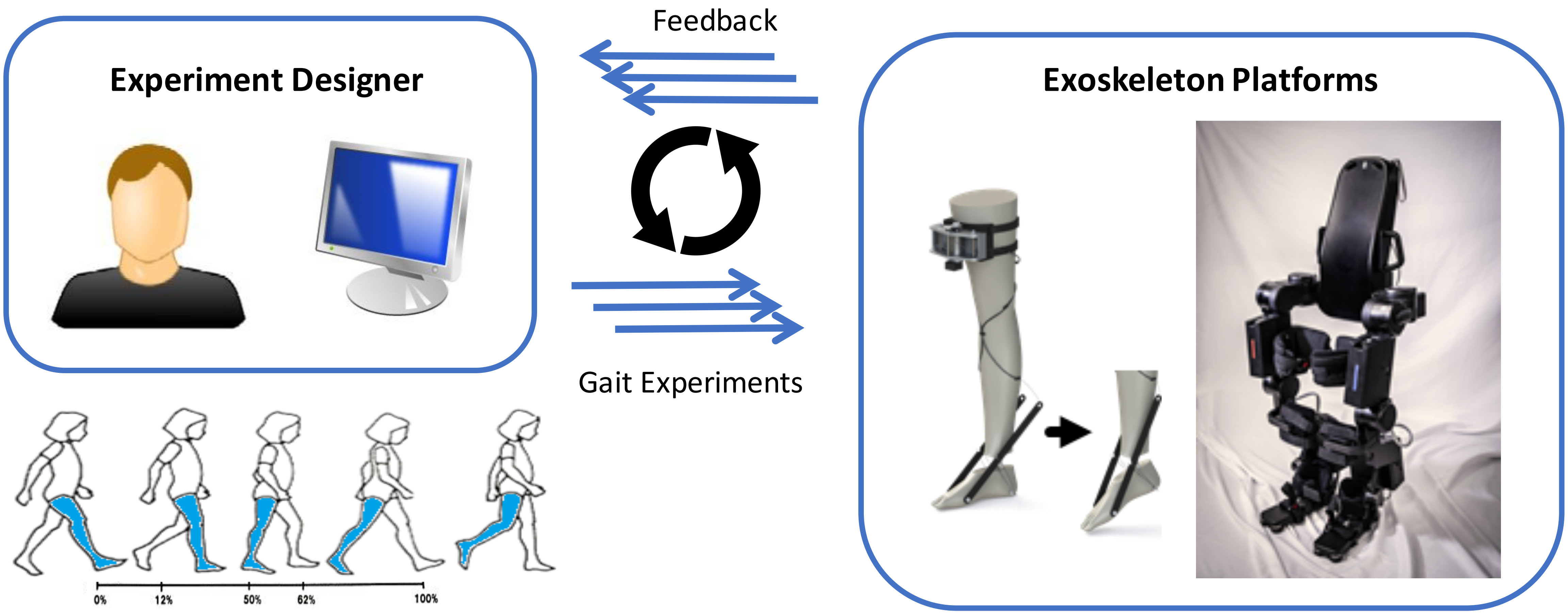

Ellen Novoseller


Yanan Sui


Maegan Tucker


Claudia Kann

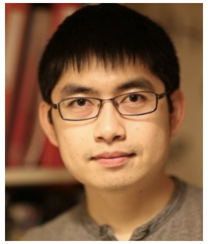
…



Dueling Posterior Sampling for Preference-Based Reinforcement Learning, Ellen Novoseller et al., arxiv

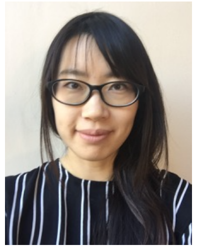
Any Many More…

(human cognitive factors)

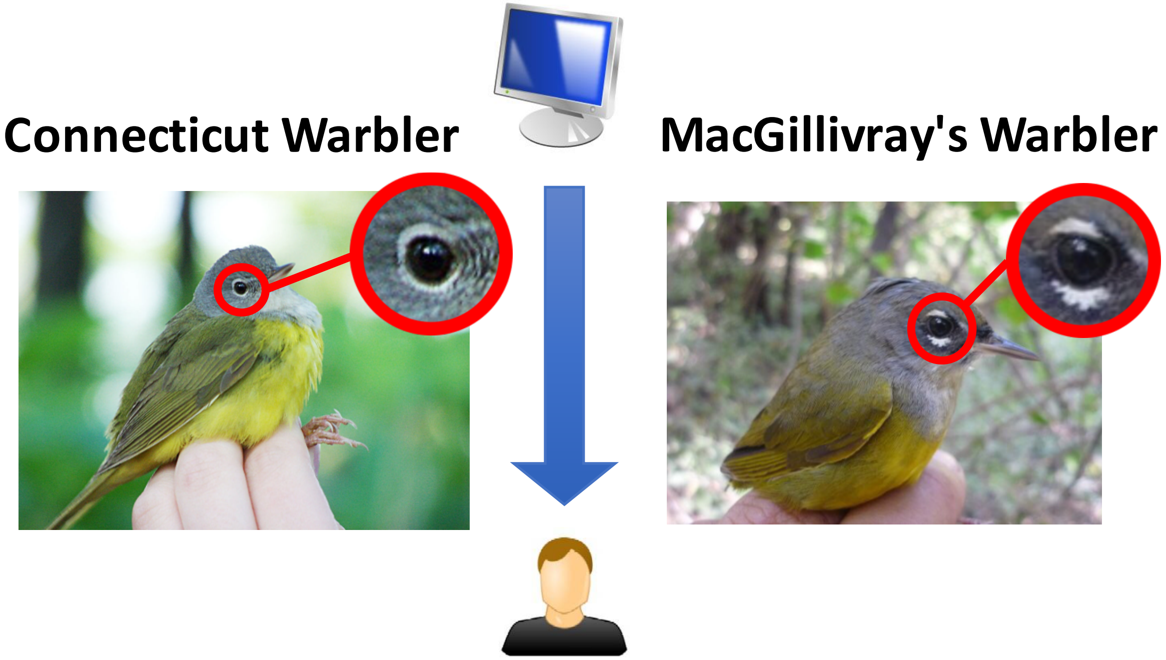

Yuxin Chen


Oisin Mac Aodha

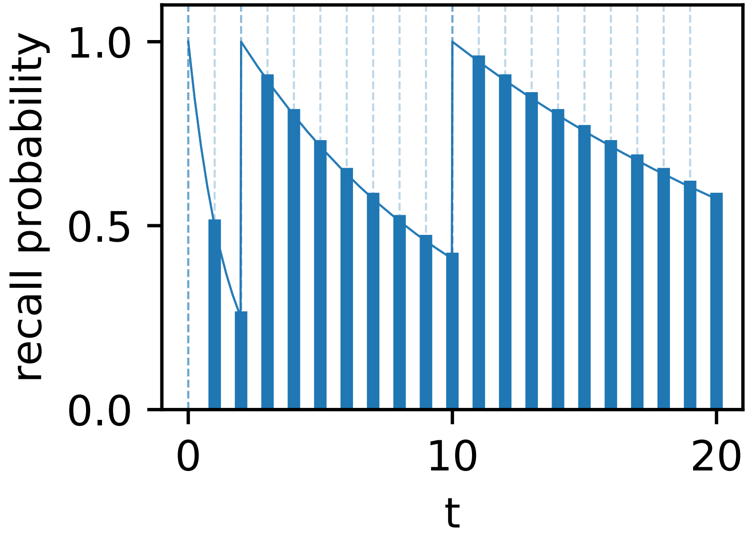

Anette Hunziker


Shihan Su

…



Interpretable Teaching



Teaching Forgetful Learners

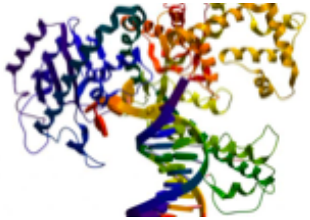
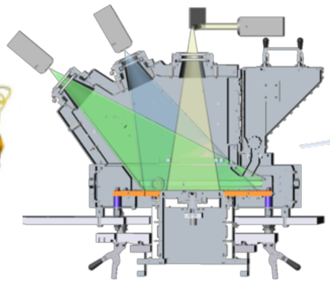
Near-Optimal Machine Teaching via Explanatory Teaching Sets, Yuxin Chen, Oisin Mac Aodha, et al., AISTATS 2018
Teaching Categories to Human Learners with Visual Explanations, Oisin Mac Aodha, et al., CVPR 2018
Teaching Multiple Concepts to Forgetful Learners, Anette Hunziker, Yuxin Chen, et al., arxiv

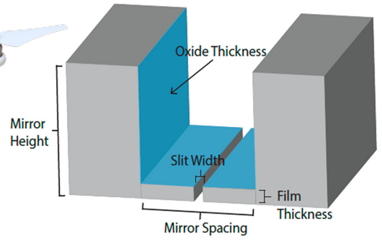
AI for Adaptive Experiment Design

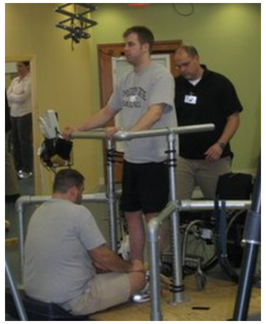
- **Experimental Platforms Increasingly Automated**
  - Motivates using Active Learning / Bayesian Optimization / Bandits

- **Real-World Considerations**
  - Indirect measurements
    - Preference feedback
    - Multi-fidelity
  - Constraints
    - Safety
    - Physical constraints
  - Domain Knowledge
    - Dynamics
    - Human factors

- **Cool Applications!**





Yanan Sui, Ellen Novoseller, Jialin Song, Kevin Yang, Yuxin Chen, Yury Tokpanov, Maegan Tucker, Claudia Kann, Vincent Zhuang, Angie Liu, Guanya Shi, Oisin Mac Aodha, Anette Hunziker, Shihan Su

Alycia Lee, Kate Fountaine, Dagny Fleischman, Ufuk Topcu, Frances Arnold, Harry Atwater, Joel Burdick, Anima Anandkumar, Soon-Jo Chung, Aaron Ames, Adish Singla, Pietro Perona, Manuel Gomez Rodriguez, Andreas Krause

Multi-dueling Bandits with Dependent Arms, Yanan Sui et al., UAI 2017
Correlational Dueling Bandits with Application to Clinical Treatment in Large Decision Spaces, Yanan Sui et al., IJCAI 2017
Dueling Posterior Sampling for Preference-Based Reinforcement Learning, Ellen Novoseller et al., arxiv
Stagewise Safe Bayesian Optimization with Gaussian Processes, Yanan Sui et al., ICML 2018
A General Framework for Multi-fidelity Bayesian Optimization with Gaussian Processes, Jialin Song et al., AISTATS 2019
Optimizing Photonic Nanostructures via Multi-fidelity Gaussian Processes, Jialin Song et al., NeurIPS Workshop on Machine Learning for Molecules and Materials, 2018
Batched Stochastic Bayesian Optimization, Kevin Yang et al., AISTATS 2019
Robust Regression for Safe Exploration in Control, Angie Liu, Guanya Shi, et al., arxiv
Near-Optimal Machine Teaching via Explanatory Teaching Sets, Yuxin Chen, Oisin Mac Aodha, et al., AISTATS 2018
Teaching Categories to Human Learners with Visual Explanations, Oisin Mac Aodha, et al., CVPR 2018
Teaching Multiple Concepts to Forgetful Learners, Anette Hunziker, Yuxin Chen, et al., arxiv