THE UNIVERSITY OF CHICAGO

Bayesian Experimental Design

in the Physical Sciences

Yuxin Chen

IPAM Workshop IV @ UCLA | Nov 22, 2019

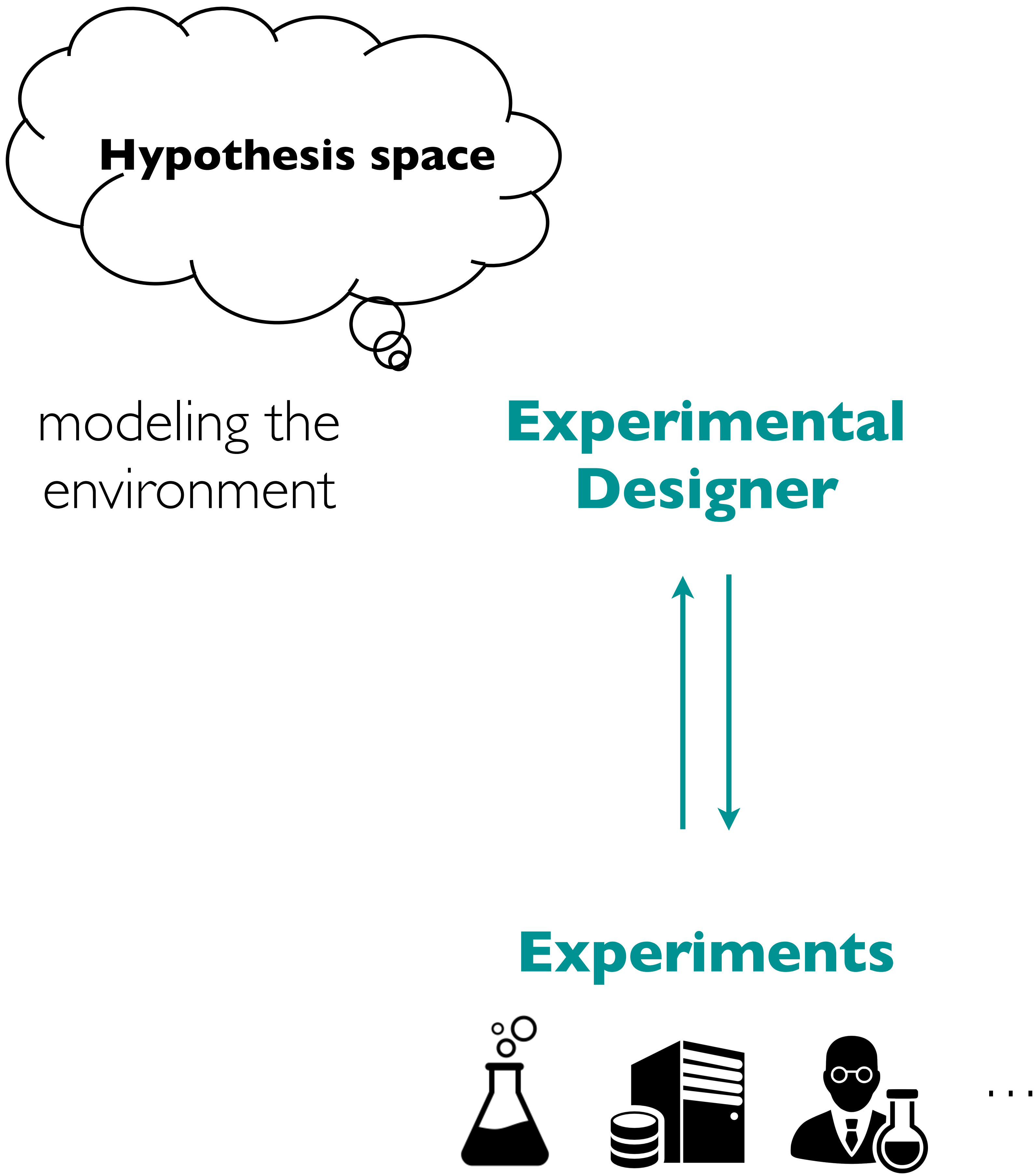
Experimental Designer

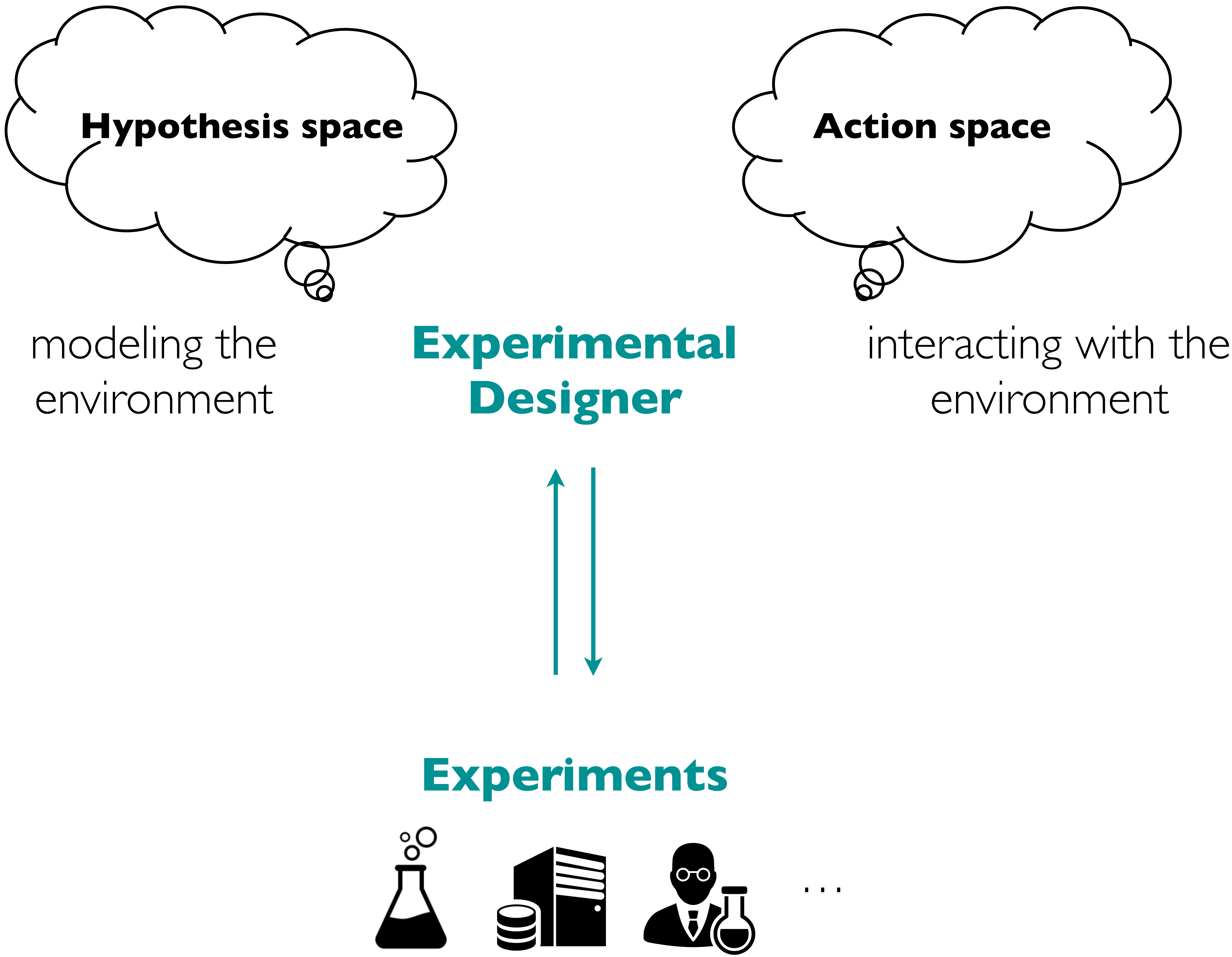
Experiments

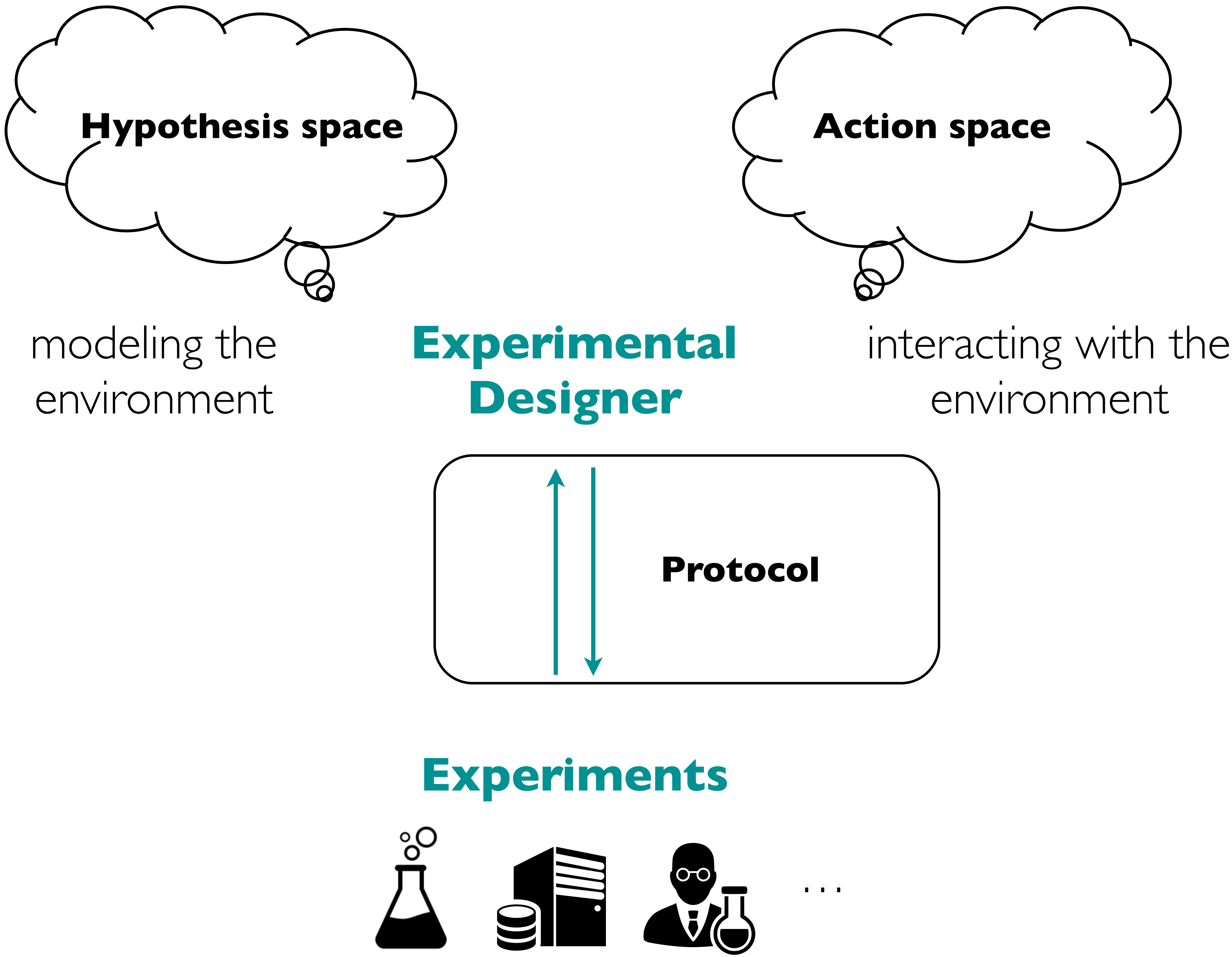
Experimental Designer

Experiments

. . .


Hypothesis space
modeling the environment
Experimental Designer
Experiments
. . .


Hypothesis space
Action space
modeling the environment
Experimental Designer
interacting with the environment
Experiments
...

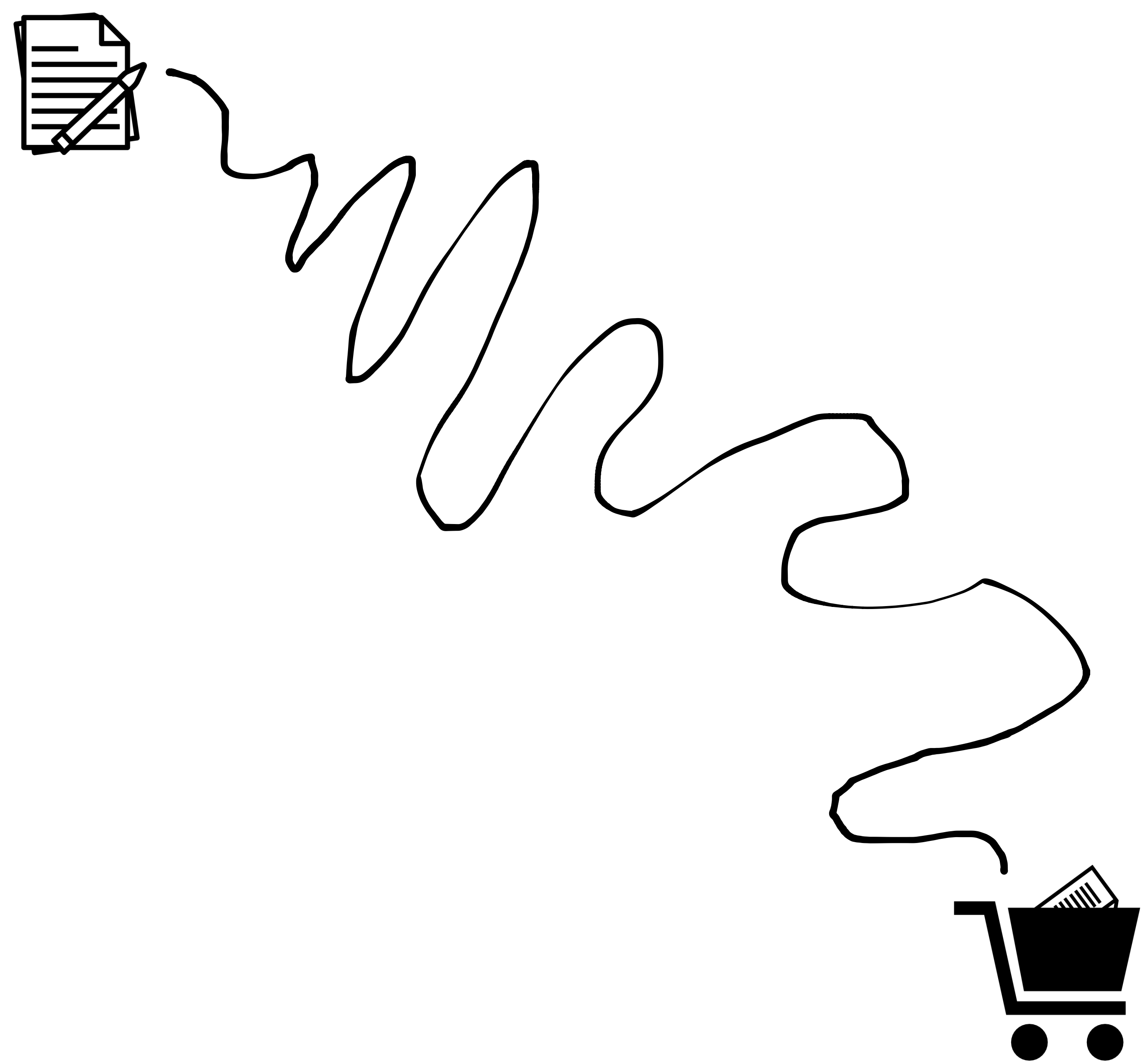

Hypothesis space
Action space
modeling the environment
Experimental Designer
interacting with the environment
Protocol
Experiments
. . .

Hypothesis space

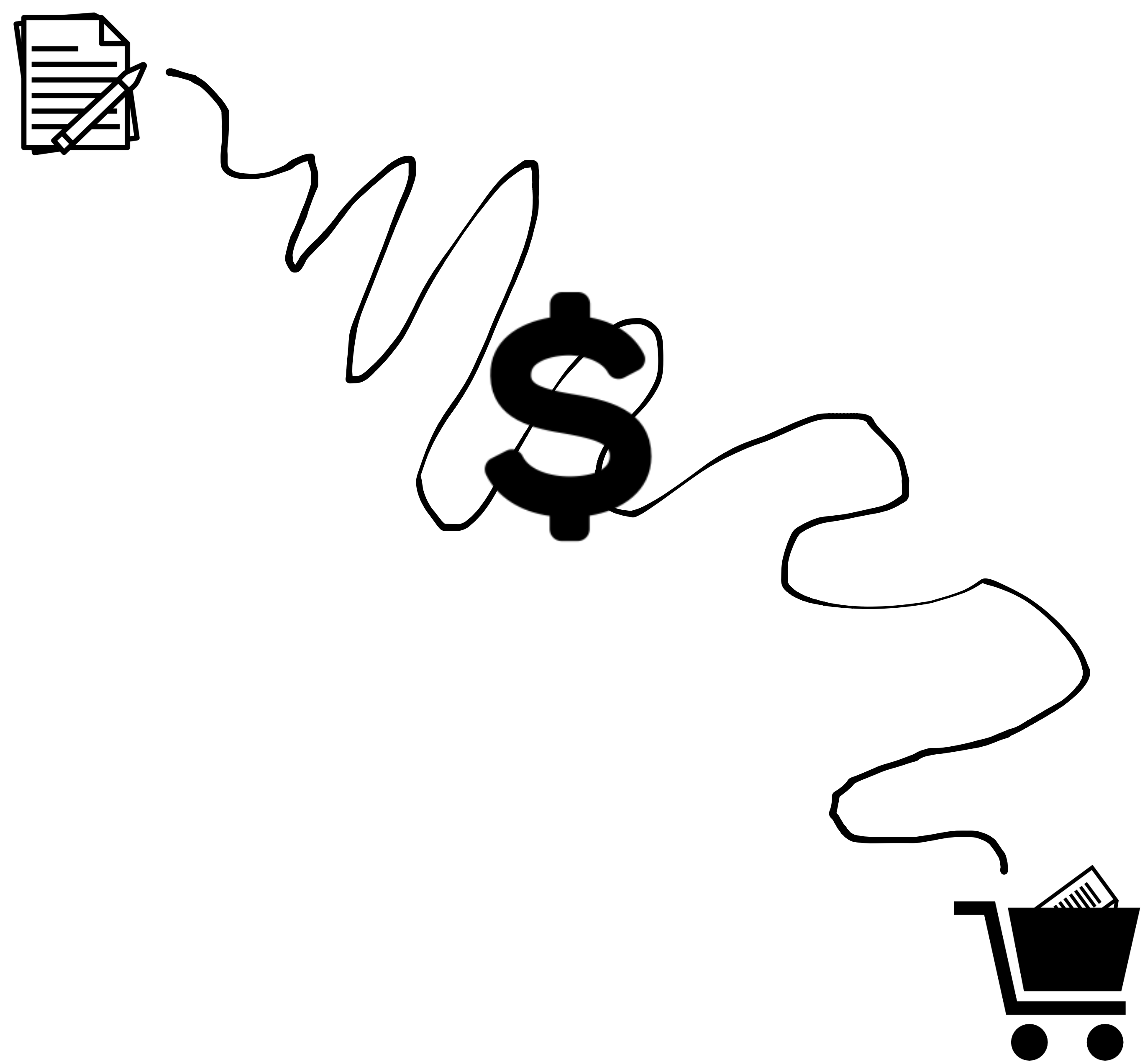
Hypothesis space

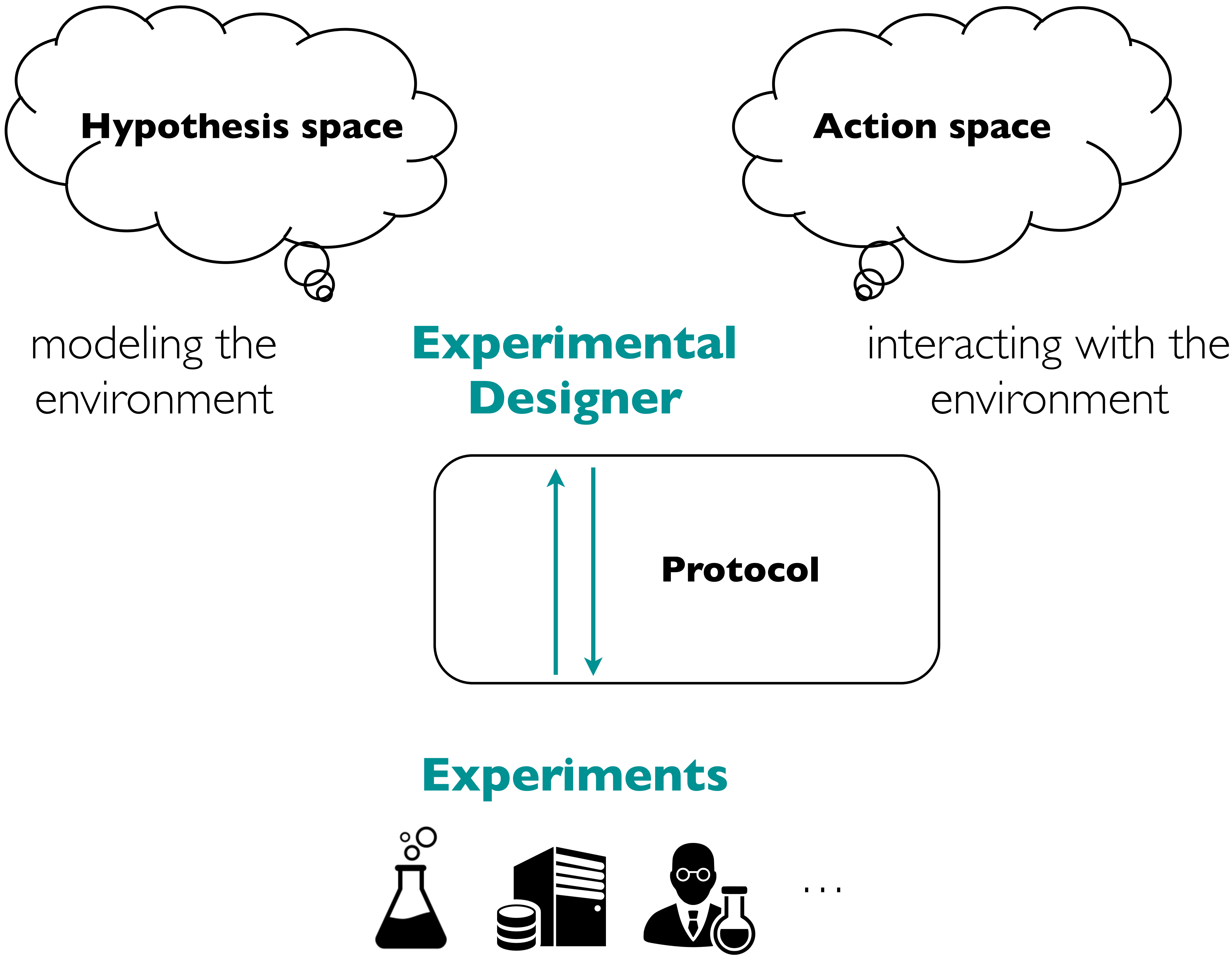
Hypothesis space

Hypothesis space

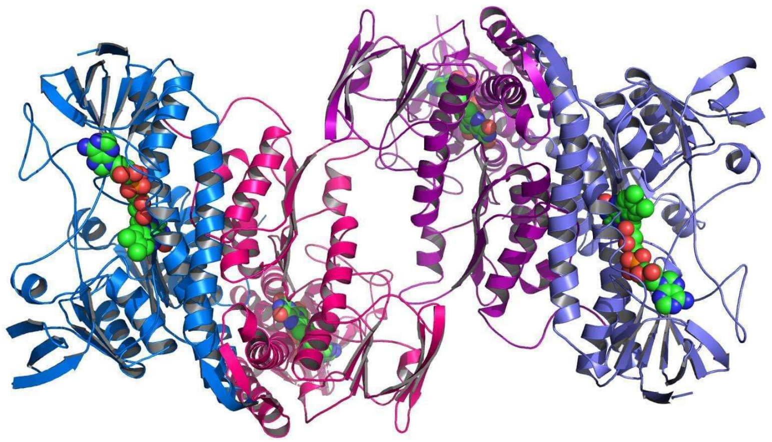


Hypothesis space



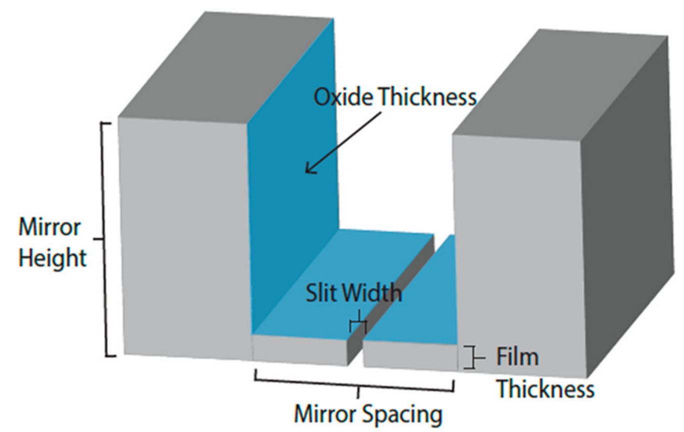

Hypothesis space
Action space
modeling the environment
Experimental Designer
interacting with the environment
Protocol
Experiments
...

Applications of Exp. Design in the Physical Sciences



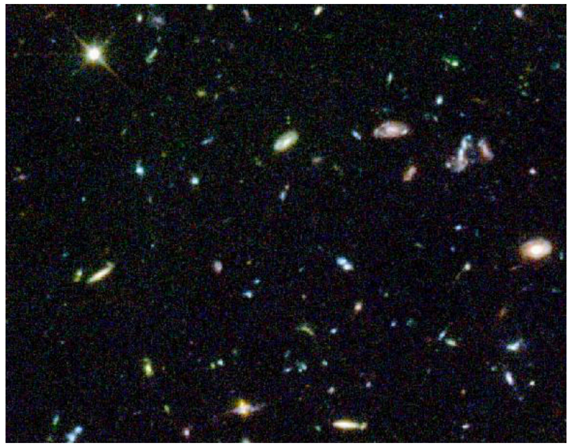
Protein engineering

image credit @ creativebiomart

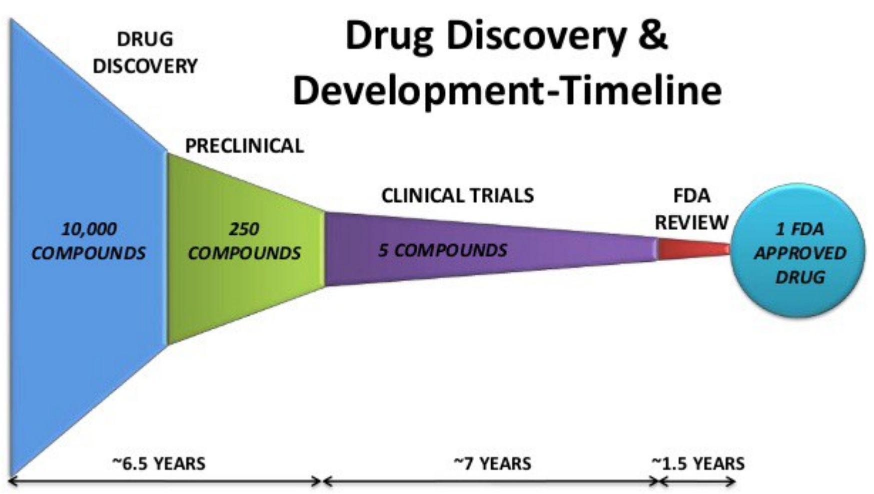


Nanophotonics

image credit @ phys.org

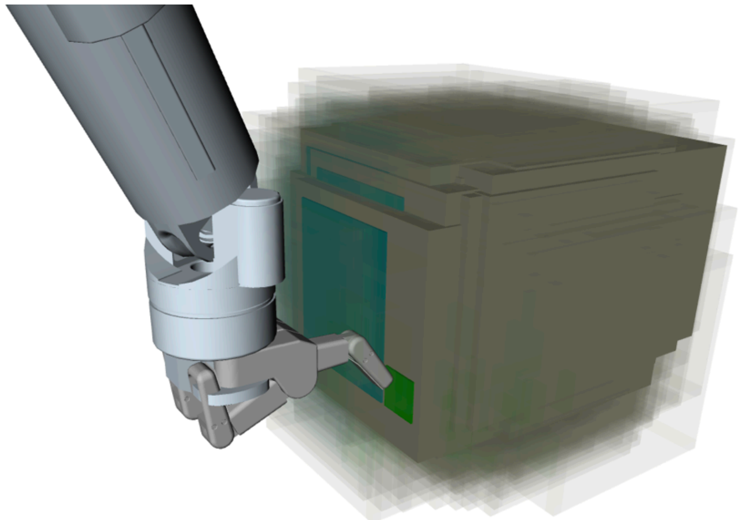


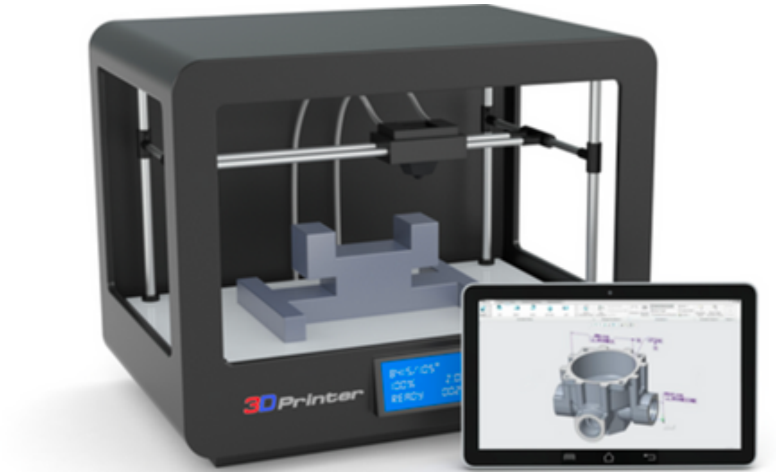
Cosmic survey



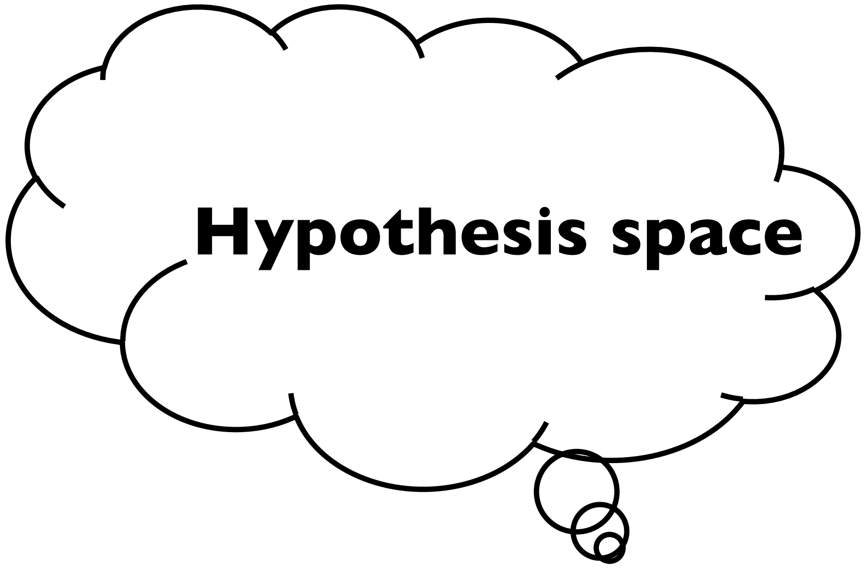
Drug discovery

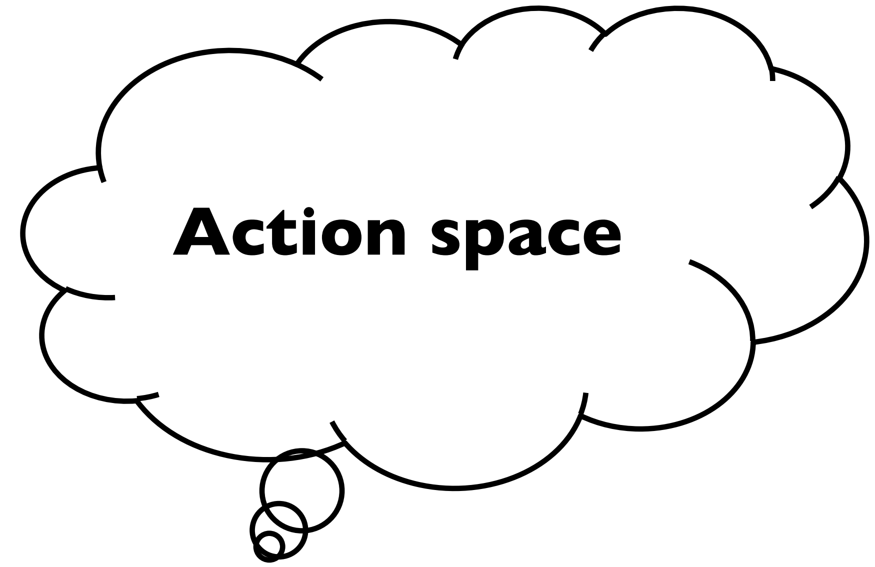
from Slideshare





Control (policy design and calibration) • • •





Bayesian Experimental Designer

Experiments

Hypothesis space

Action space

Bayesian Experimental Designer

Input $\boldsymbol{x} \in \boldsymbol{X}$

Experiments

 . . .

Hypothesis space

Action space

Bayesian Experimental Designer

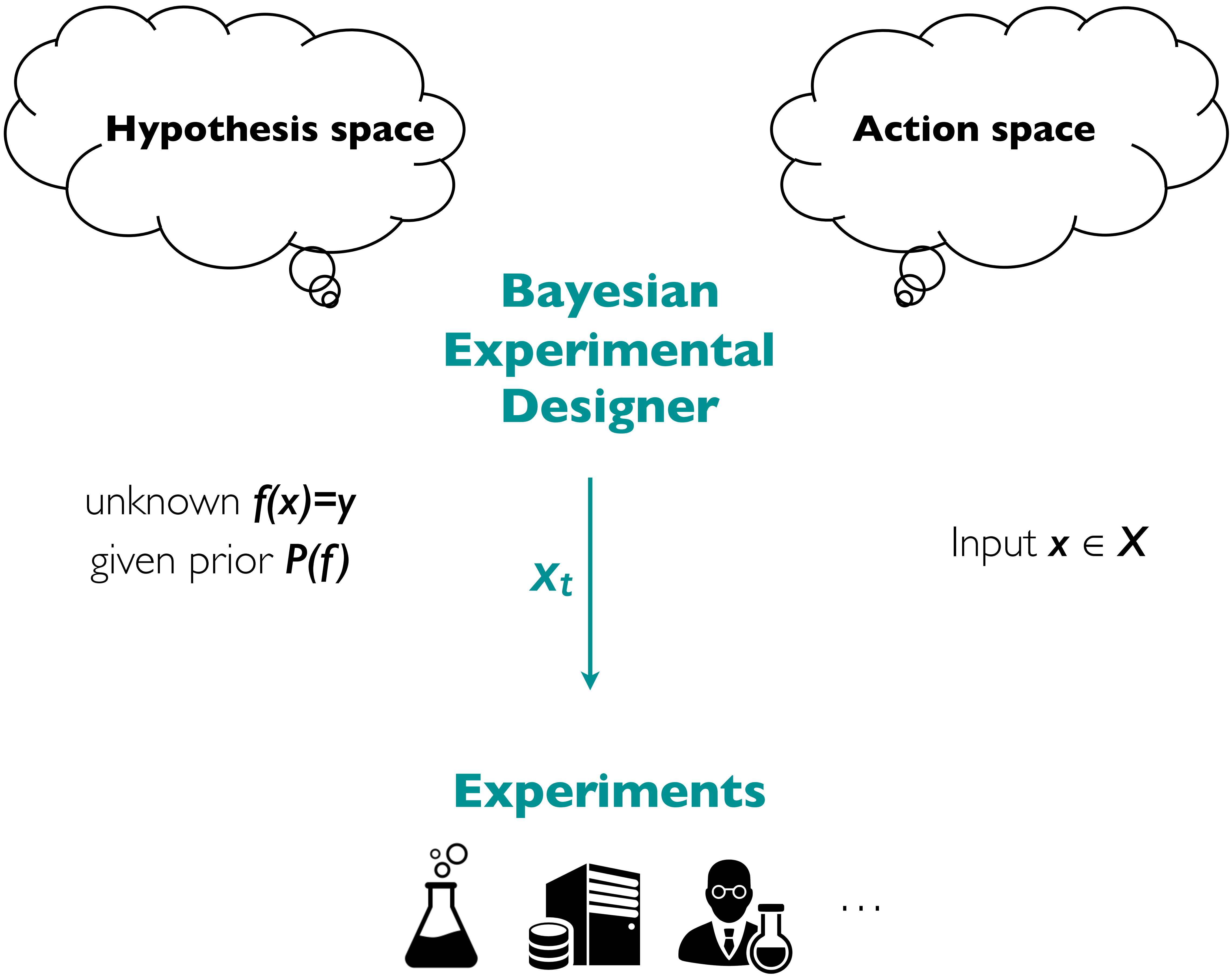
unknown $f(x)=y$

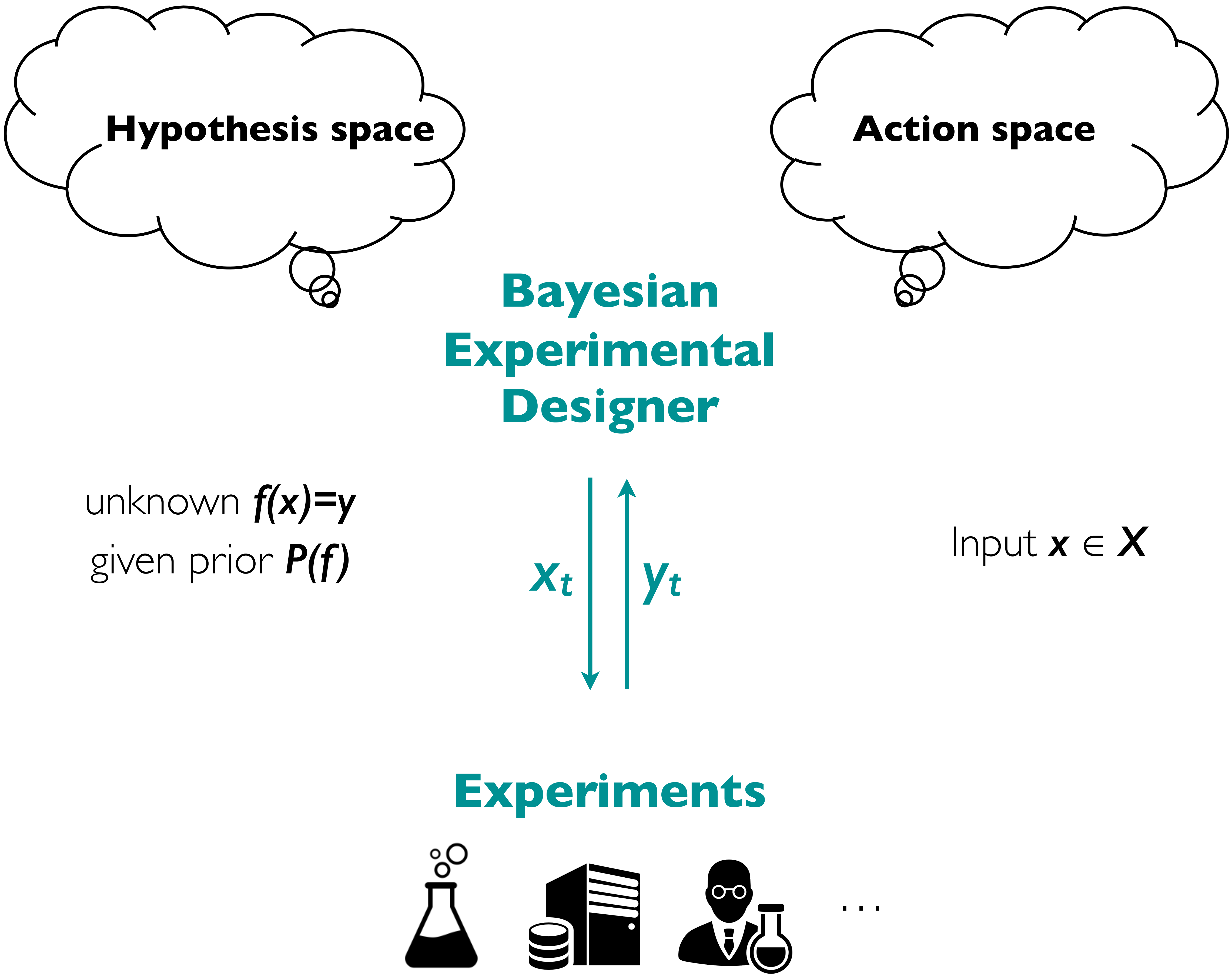
given prior $P(f)$

Input $x \in X$

Experiments

...


Hypothesis space
Action space
Bayesian Experimental Designer
unknown $f(x)=y$
given prior $P(f)$
x_t
Input $x \in X$
Experiments
...


Hypothesis space
Action space
Bayesian Experimental Designer
unknown f(x)=y
given prior P(f)
x_t
y_t
Input x ∈ X
Experiments
. . .

Hypothesis space

Action space

Bayesian Experimental Designer

unknown $f(x)=y$

given prior $P(f)$

maintain $P(f \mid obs)$

x_t y_t

Input $x \in X$

Experiments

...

Optimize a Blackbox Function

Modeling performance as a function of parameters

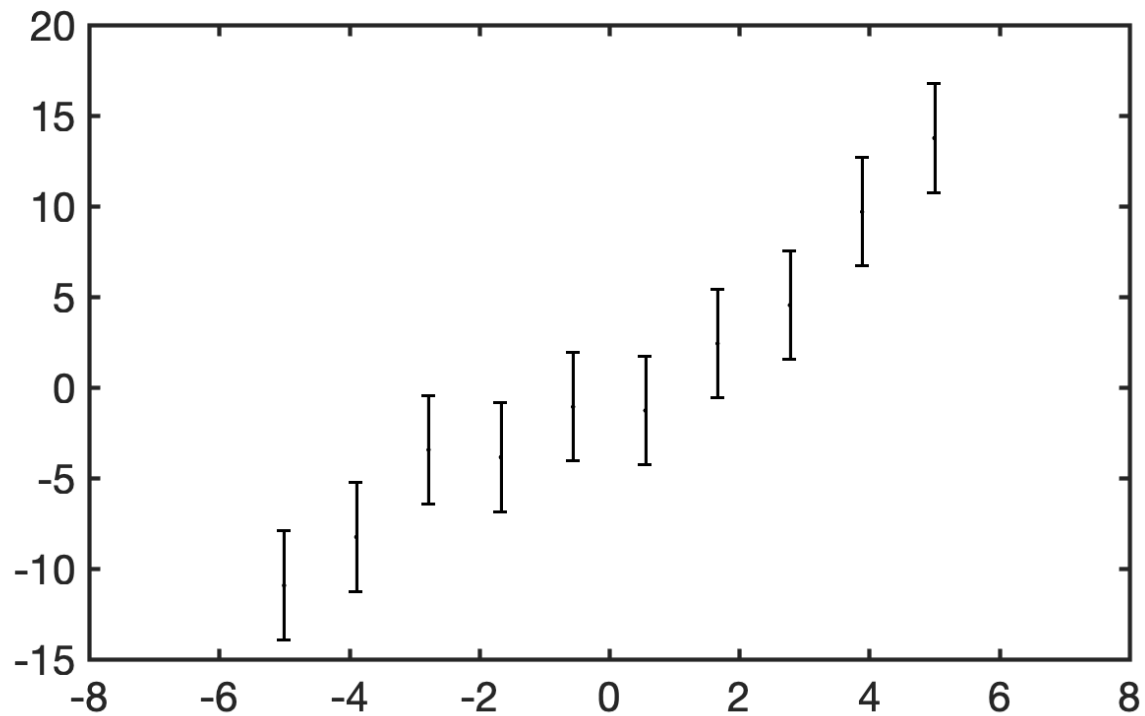
- E.g., robotics/control, material science

Modeling structured/sequential data

- E.g., biochemical engineering

…

$$x^* = \arg\max_x f(x)$$



A Primer on Bayesian Optimization

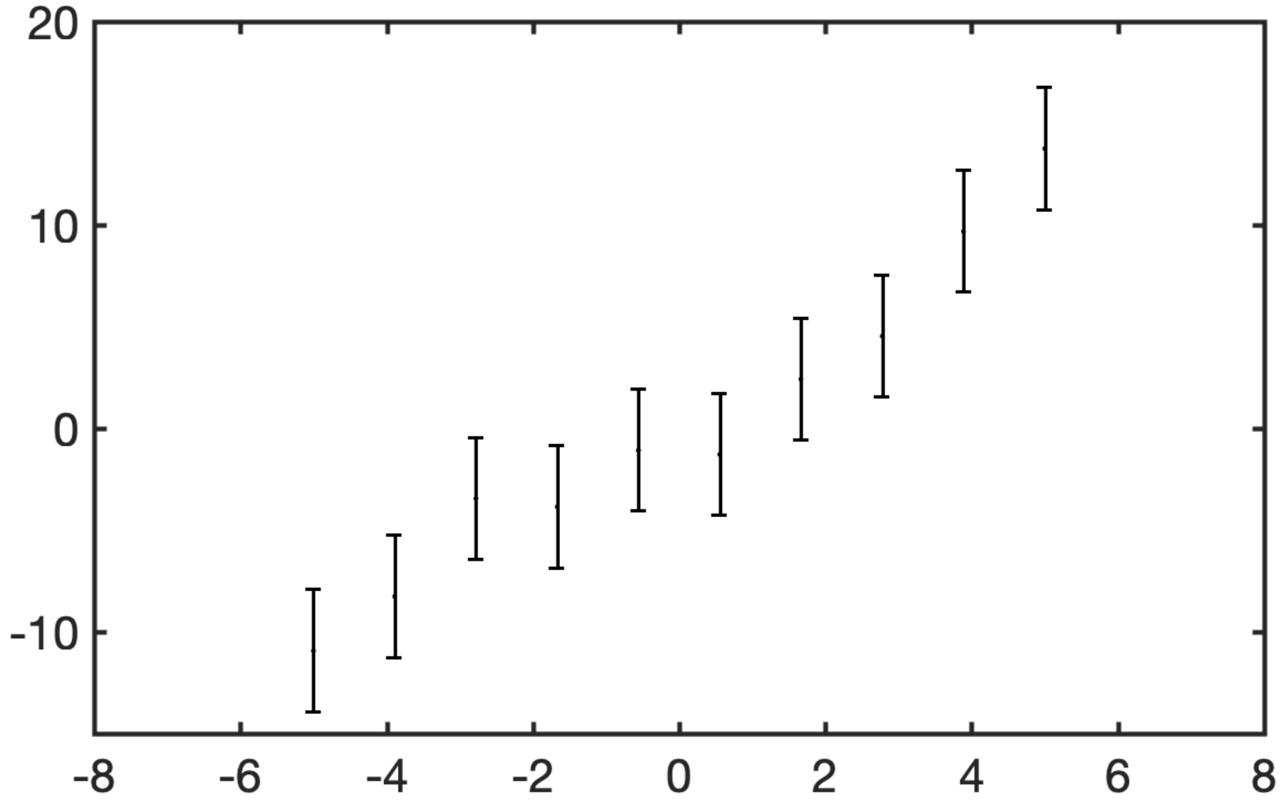
Bayesian Inference

$$f_X = [f_{x_1}, \ldots, f_{x_N}]$$

$$p(f_X \mid y) = \frac{p(y \mid f_X)\, p(f_X)}{p(y)}$$

What can we do with this?

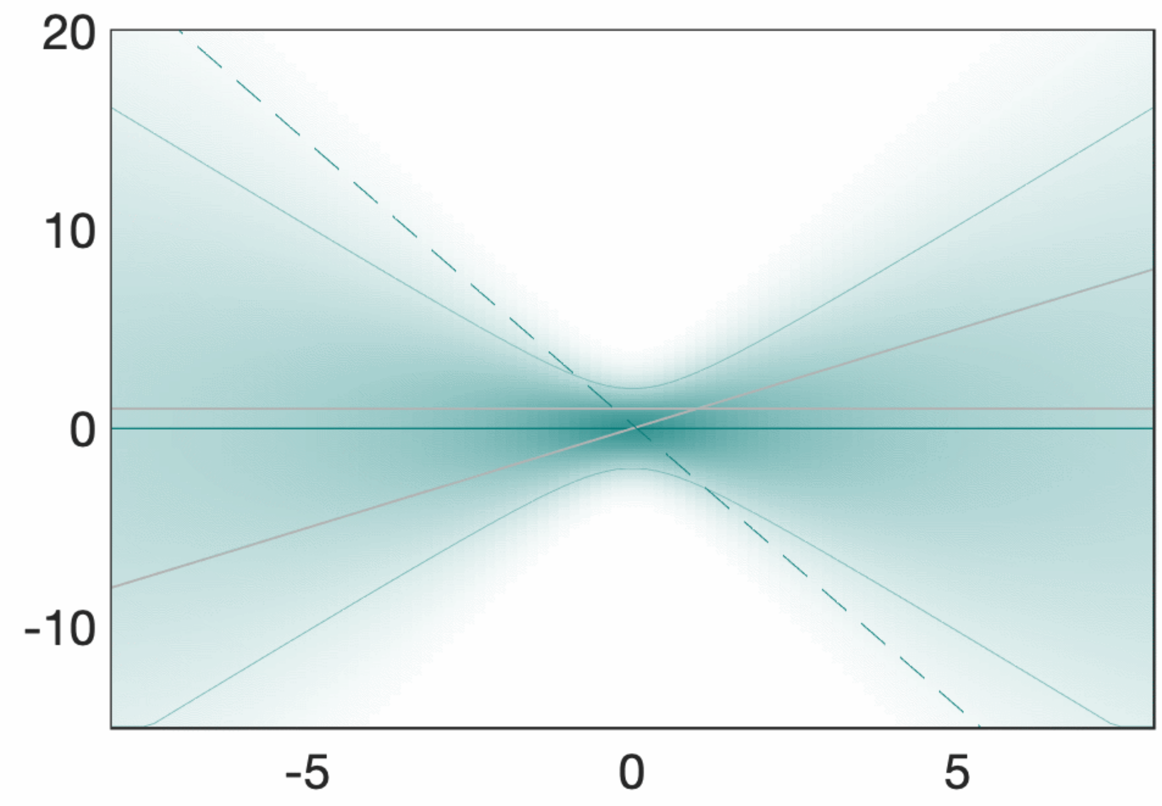
Given $y, p(y \mid f)$, what's f ?



A Prior over Linear Functions

$$f(x) = w_1 + w_2 x = \phi_x^\top w \qquad p(w) = \mathcal{N}(w; \mu, \Sigma)$$

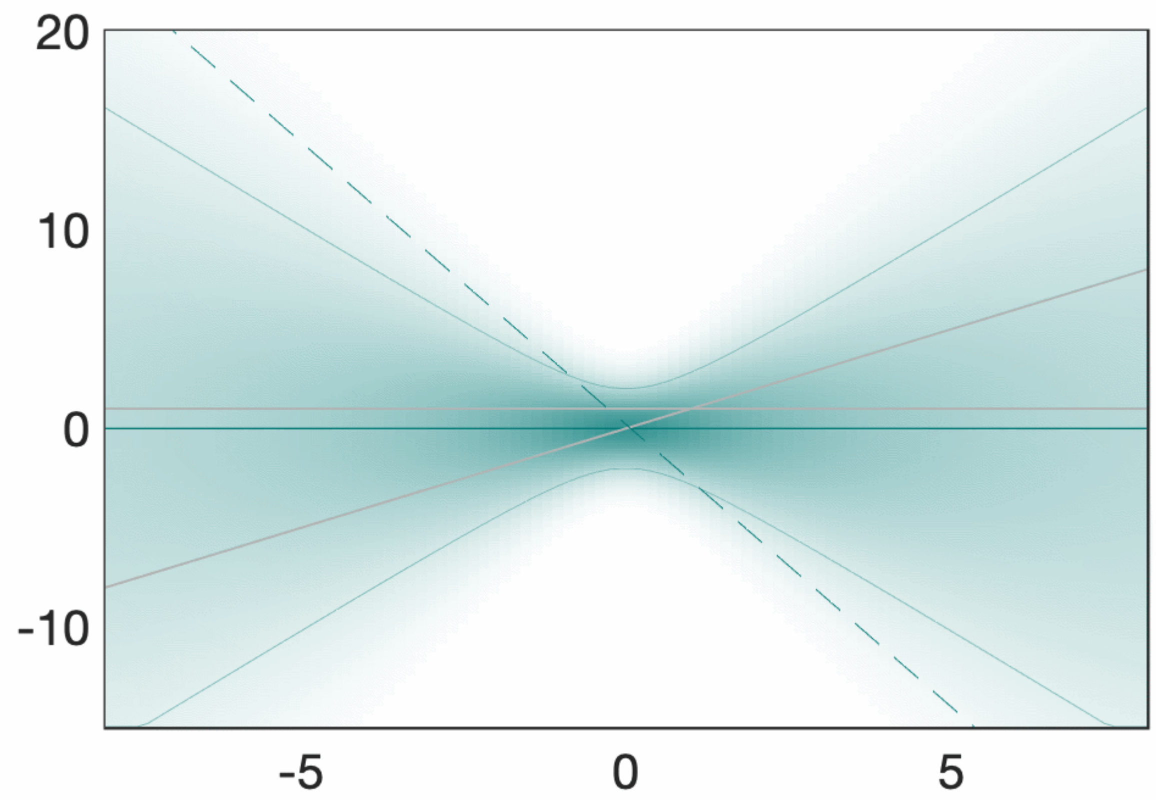
$$\phi = \begin{pmatrix} 1 \\ x \end{pmatrix} \qquad p(f) = \mathcal{N}(f; \phi(x)^\top \mu, \phi_x^\top \Sigma \phi_x)$$



A Prior over Linear Functions

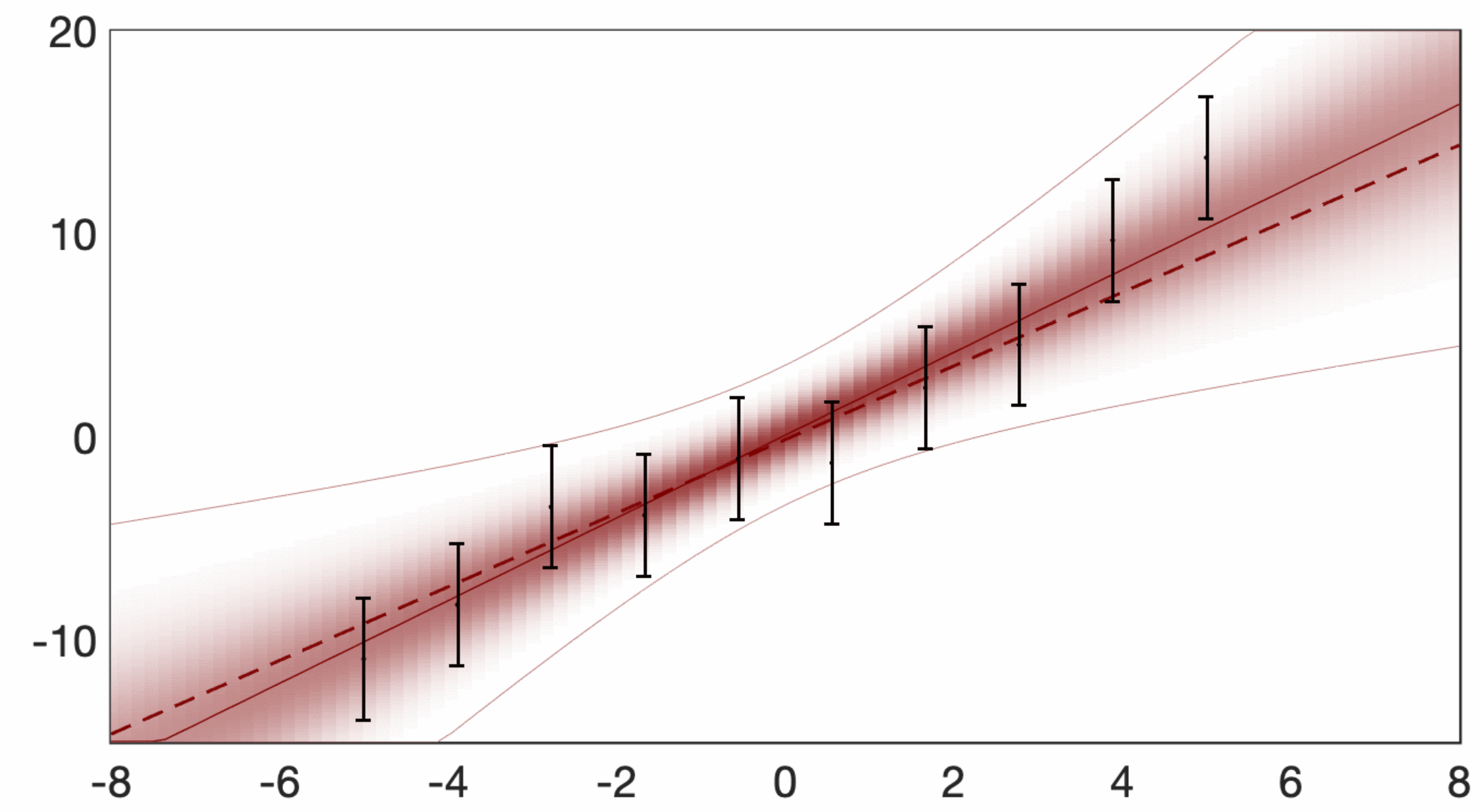
$$f(x) = w_1 + w_2 x = \phi_x^\top w \qquad p(w) = \mathcal{N}(w; \mu, \Sigma)$$

$$\phi = \begin{pmatrix} 1 \\ x \end{pmatrix} \qquad p(f) = \mathcal{N}(f; \phi(x)^\top \mu, \phi_x^\top \Sigma \phi_x)$$



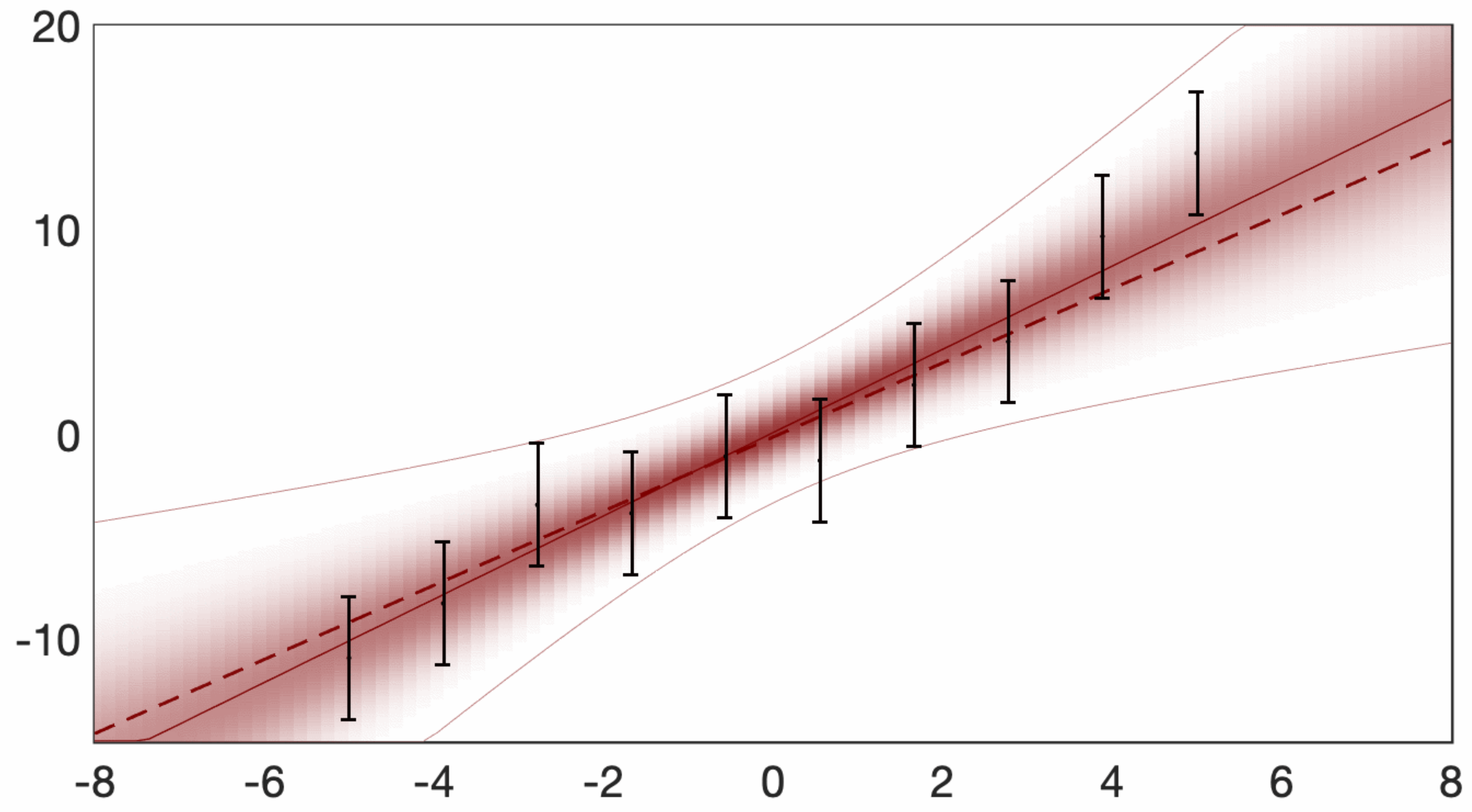
$$p(y \mid w, \phi_X) = \mathcal{N}(y; \phi_x^\top \mu, \sigma^2 I)$$

$$p(w \mid y, \phi_X) = \mathcal{N}(w; \mu + \Sigma\phi_X(\phi_X{}^\top \Sigma\phi_X + \sigma^2 I)^{-1}(y - \phi_X^\top \mu),$$
$$\Sigma - \Sigma\phi_X(\phi_X{}^\top \Sigma\phi_X + \sigma^2 I)^{-1}\phi_X{}^\top \Sigma\phi_x)$$



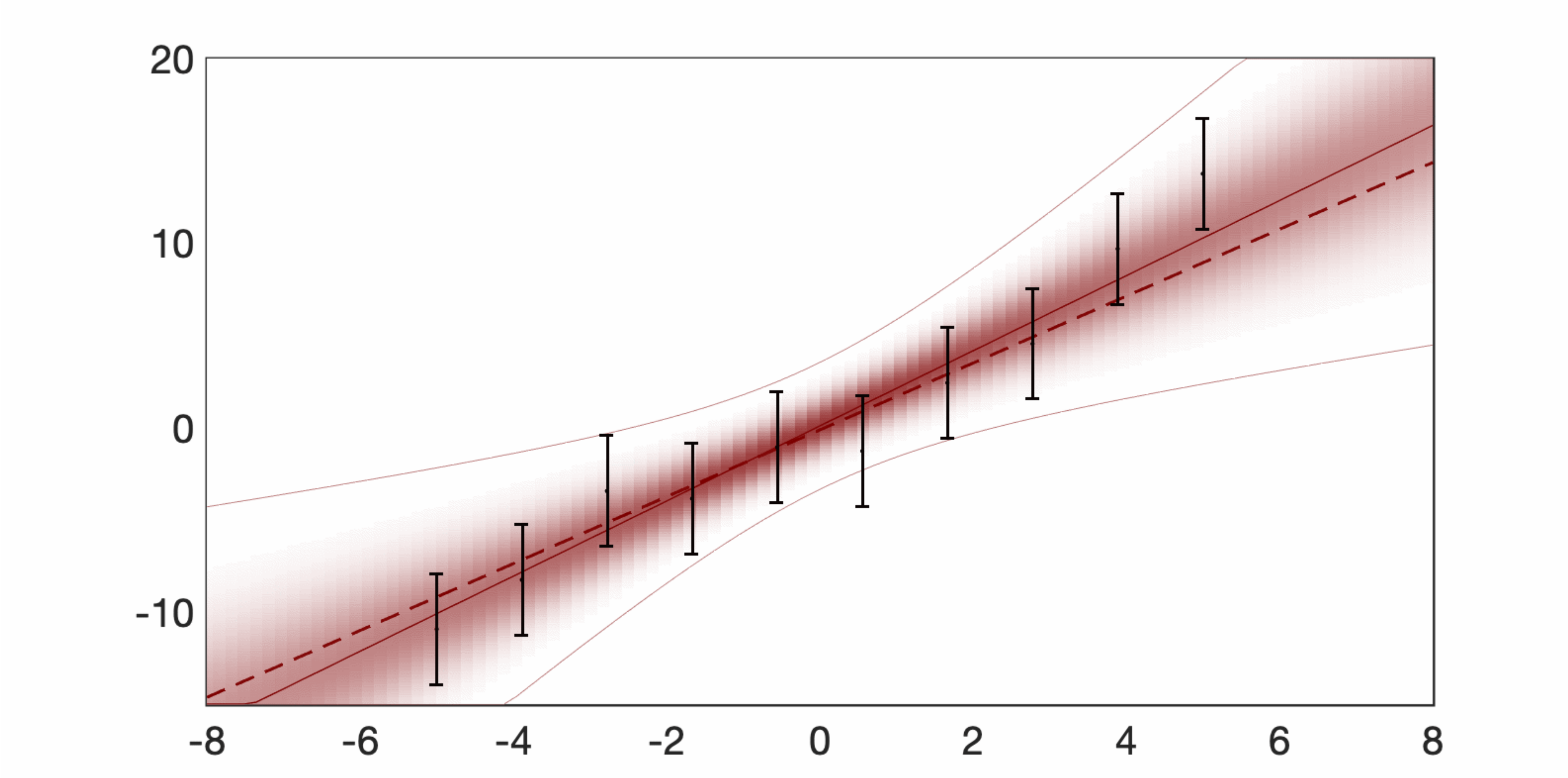
$$p(y \mid w, \phi_X) = \mathcal{N}(y; \phi_x^\top \mu, \sigma^2 I)$$

$$\begin{aligned} p(w \mid y, \phi_X) = \mathcal{N}(w; &\mu + \Sigma\phi_X(\phi_X{}^\top \Sigma \phi_X + \sigma^2 I)^{-1}(y - \phi_X^\top \mu), \\ &\Sigma - \Sigma\phi_X(\phi_X{}^\top \Sigma \phi_X + \sigma^2 I)^{-1} \phi_X{}^\top \Sigma \phi_x) \end{aligned}$$



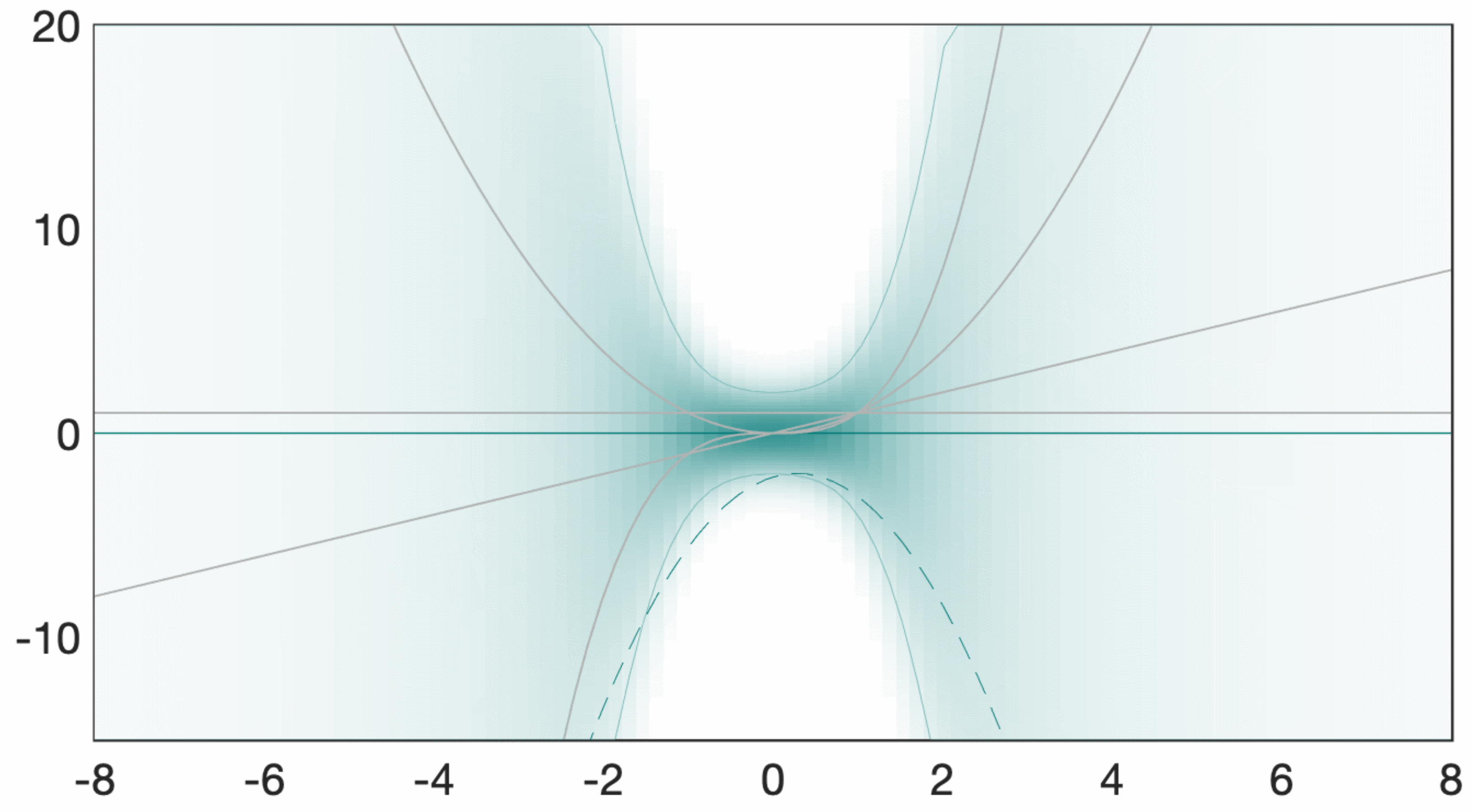
$$p(y \mid w, \phi_X) = \mathcal{N}(y; \phi_x^\top \mu, \sigma^2 I)$$

$$p(f_x \mid y, \phi_X) = \mathcal{N}(f_x; \phi_x{}^\top \mu + \phi_x{}^\top \Sigma \phi_X (\phi_X{}^\top \Sigma \phi_X + \sigma^2 I)^{-1} (y - \phi_X^\top \mu),$$
$$\phi_x{}^\top \Sigma \phi_x - \phi_x{}^\top \Sigma \phi_X (\phi_X{}^\top \Sigma \phi_X + \sigma^2 I)^{-1} \phi_X{}^\top \Sigma \phi_x)$$



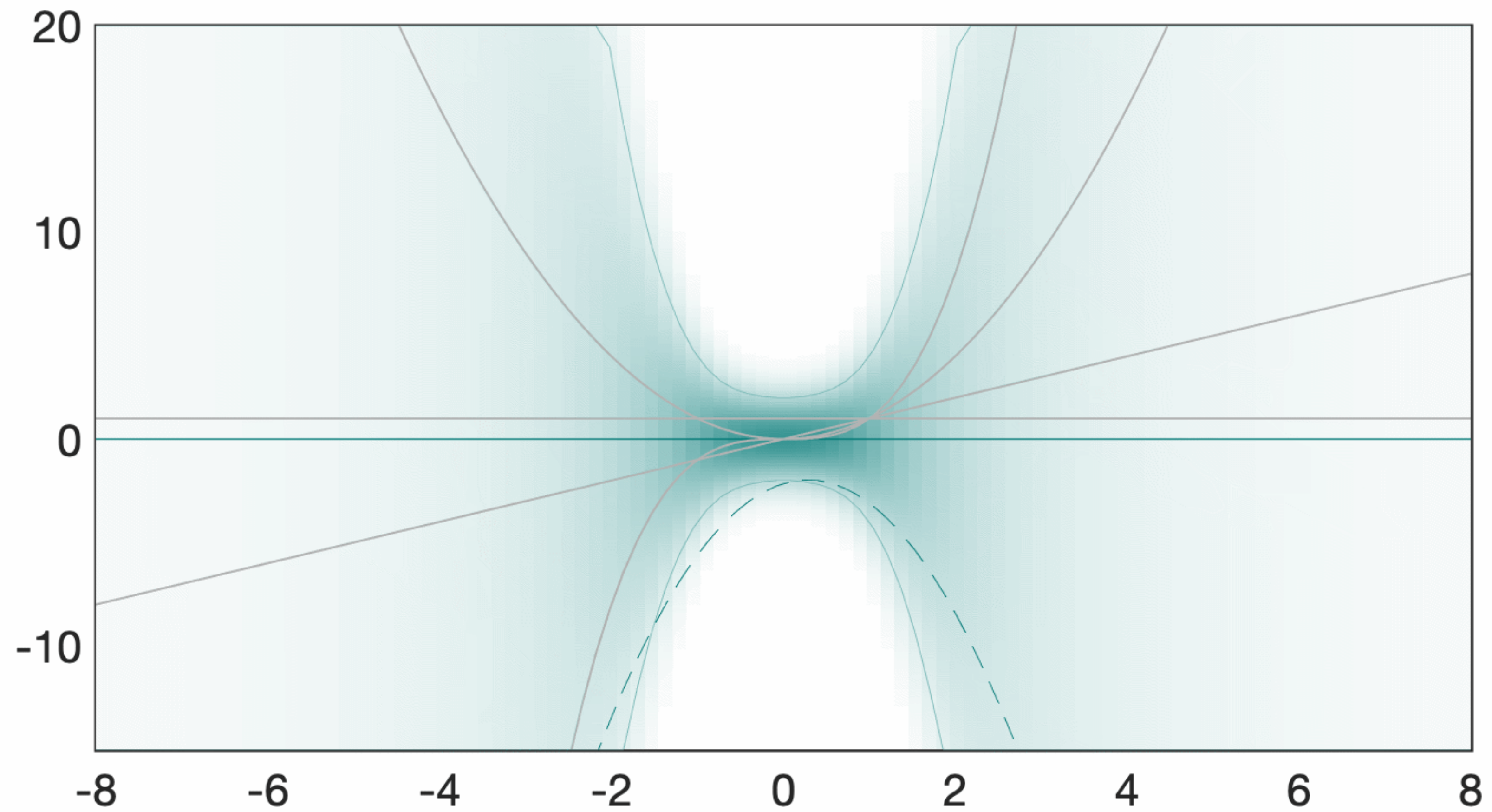
Cubic Regression

$$f(x) = \phi(x)^\top w \qquad \phi(x) = (1, x, x^2, x^3)$$



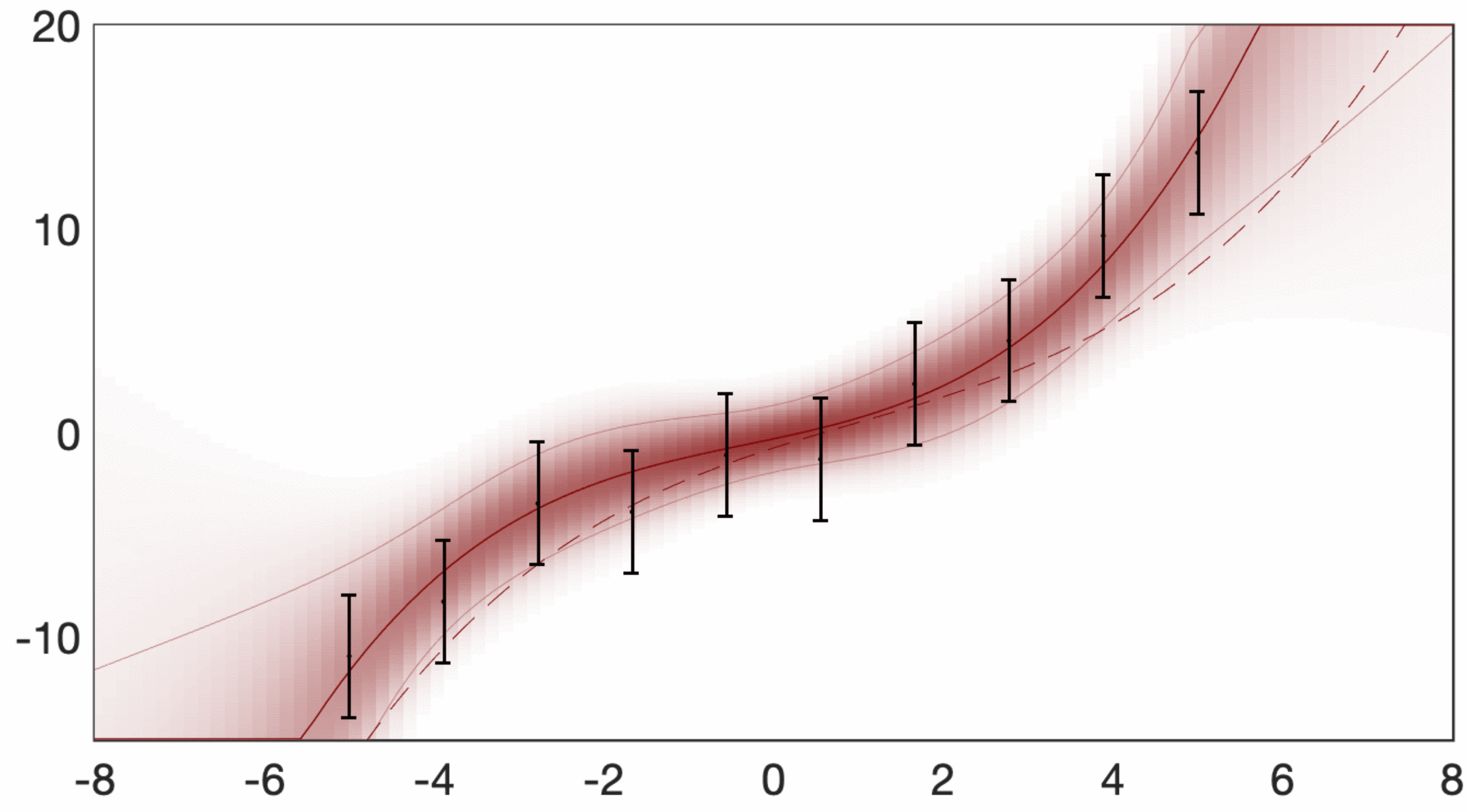
Cubic Regression

$$f(x) = \phi(x)^\top w \qquad \phi(x) = (1, x, x^2, x^3)$$



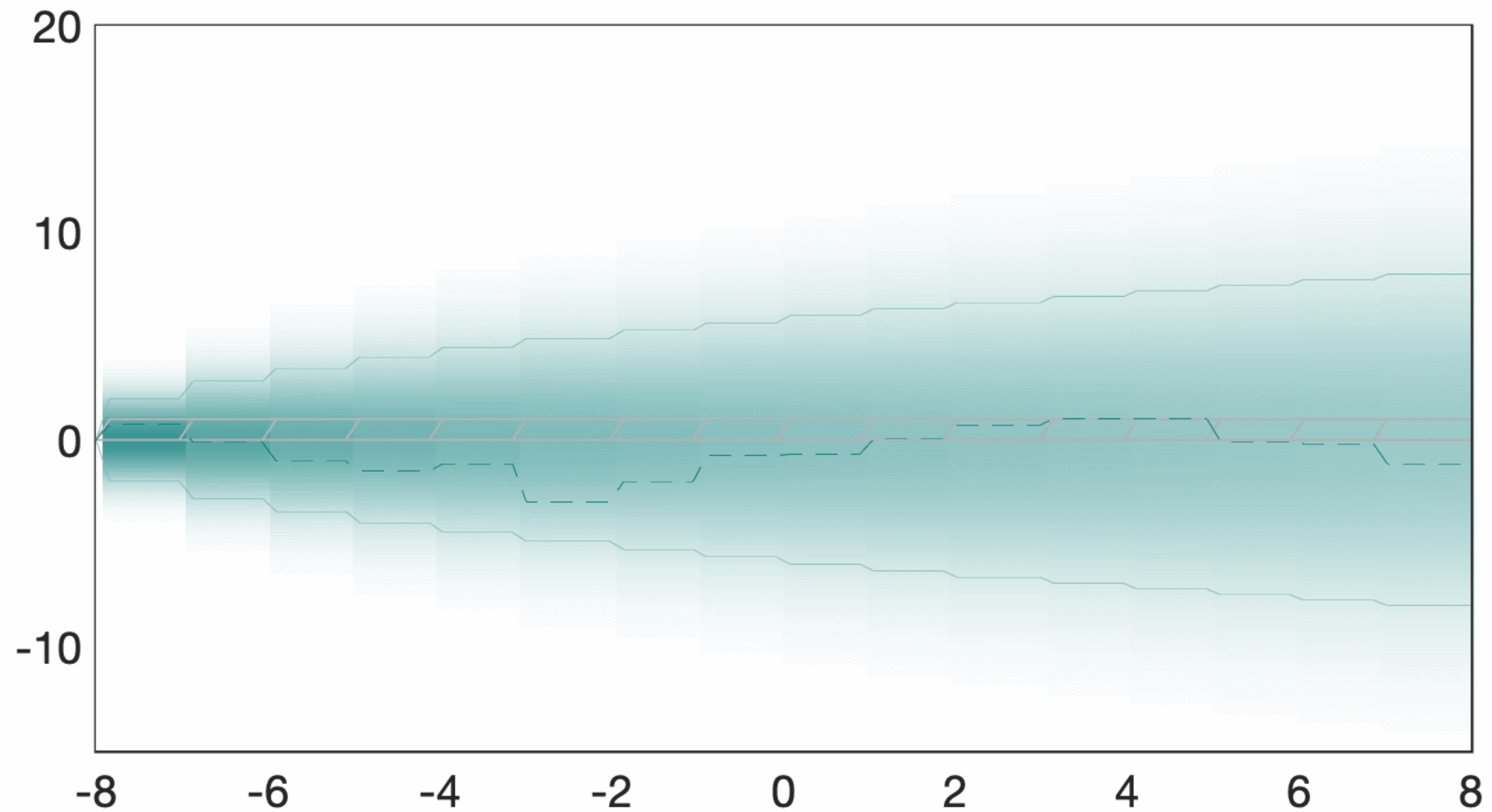
Cubic Regression

$$f(x) = \phi(x)^\top w \qquad \phi(x) = \left(1, x, x^2, x^3\right)$$



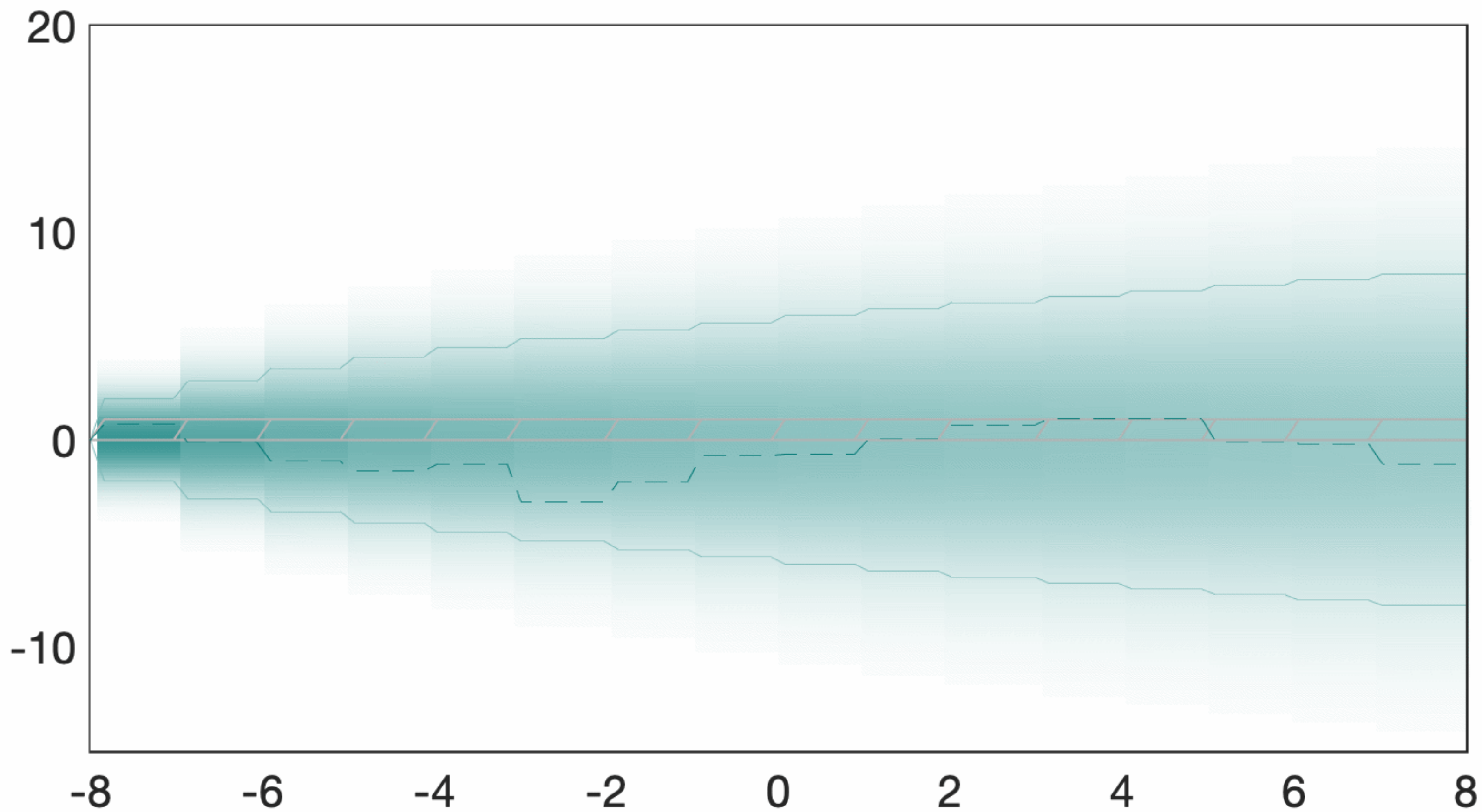
Step Regression

$$\phi(x) = (\theta(x-8), \theta(8-x), \theta(x-7), \theta(7-x), \ldots)^\top$$



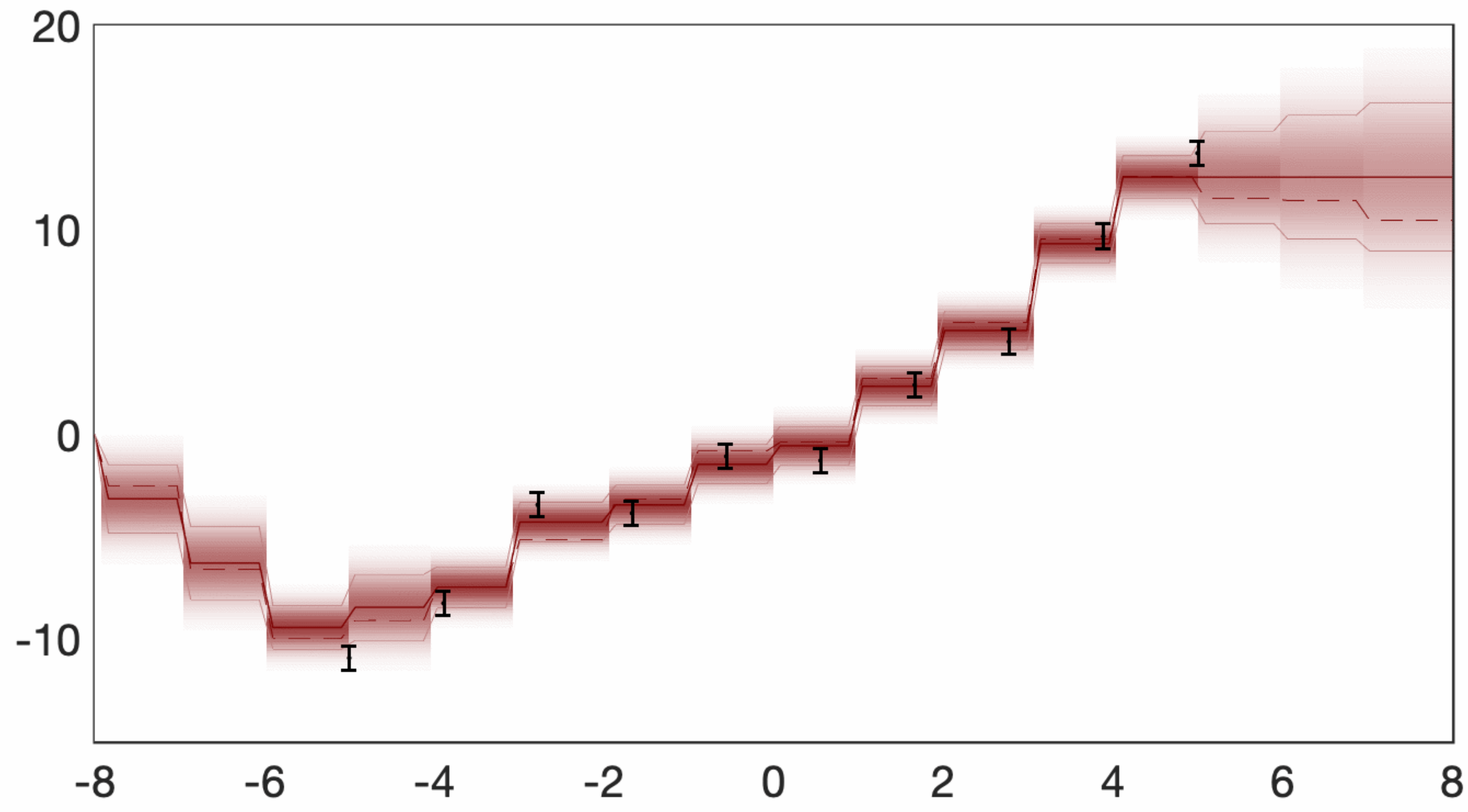
Step Regression

$$\phi(x) = (\theta(x-8), \theta(8-x), \theta(x-7), \theta(7-x), \dots)^\top$$



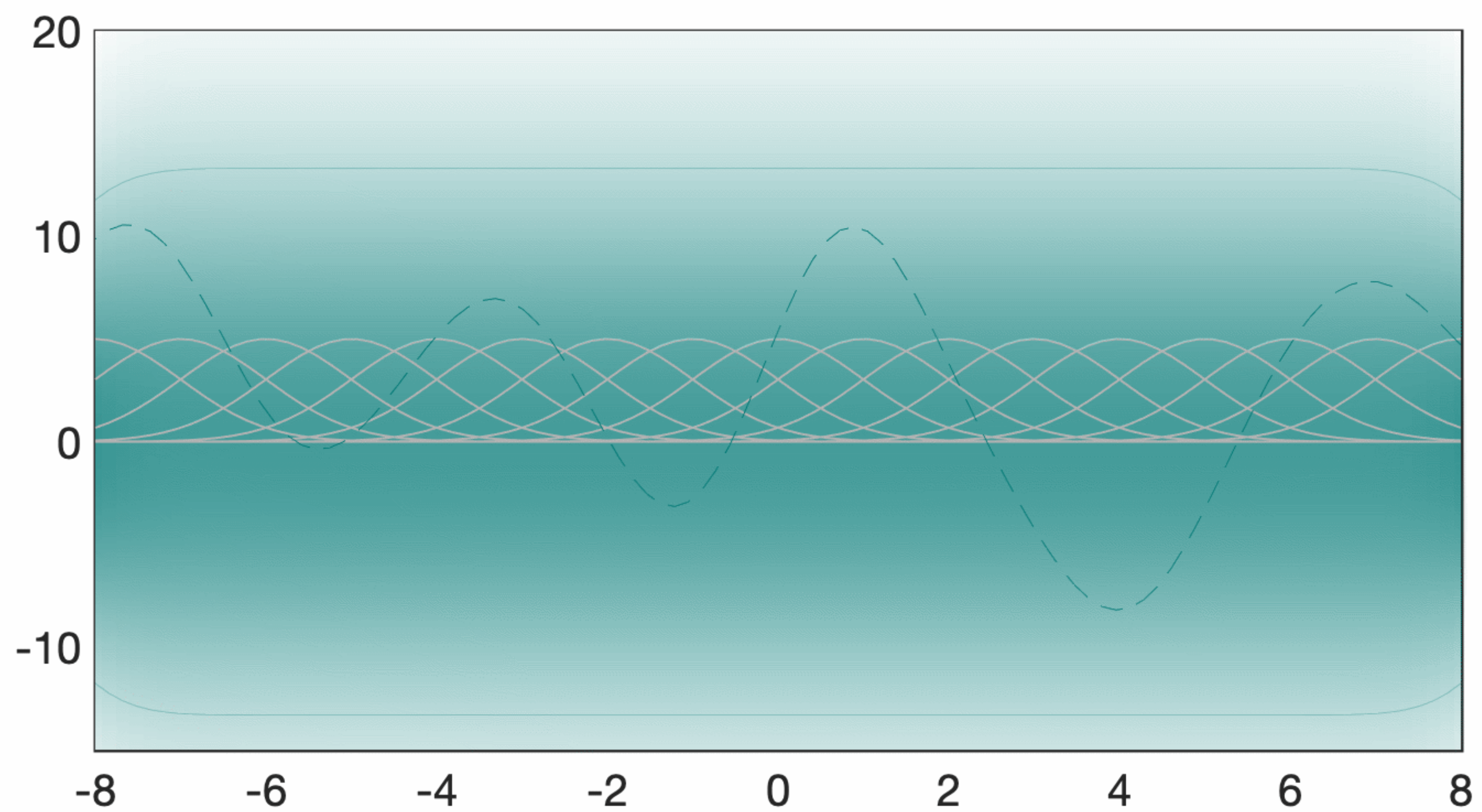
Step Regression

$$\phi(x) = (\theta(x-8), \theta(8-x), \theta(x-7), \theta(7-x), \dots)^\top$$



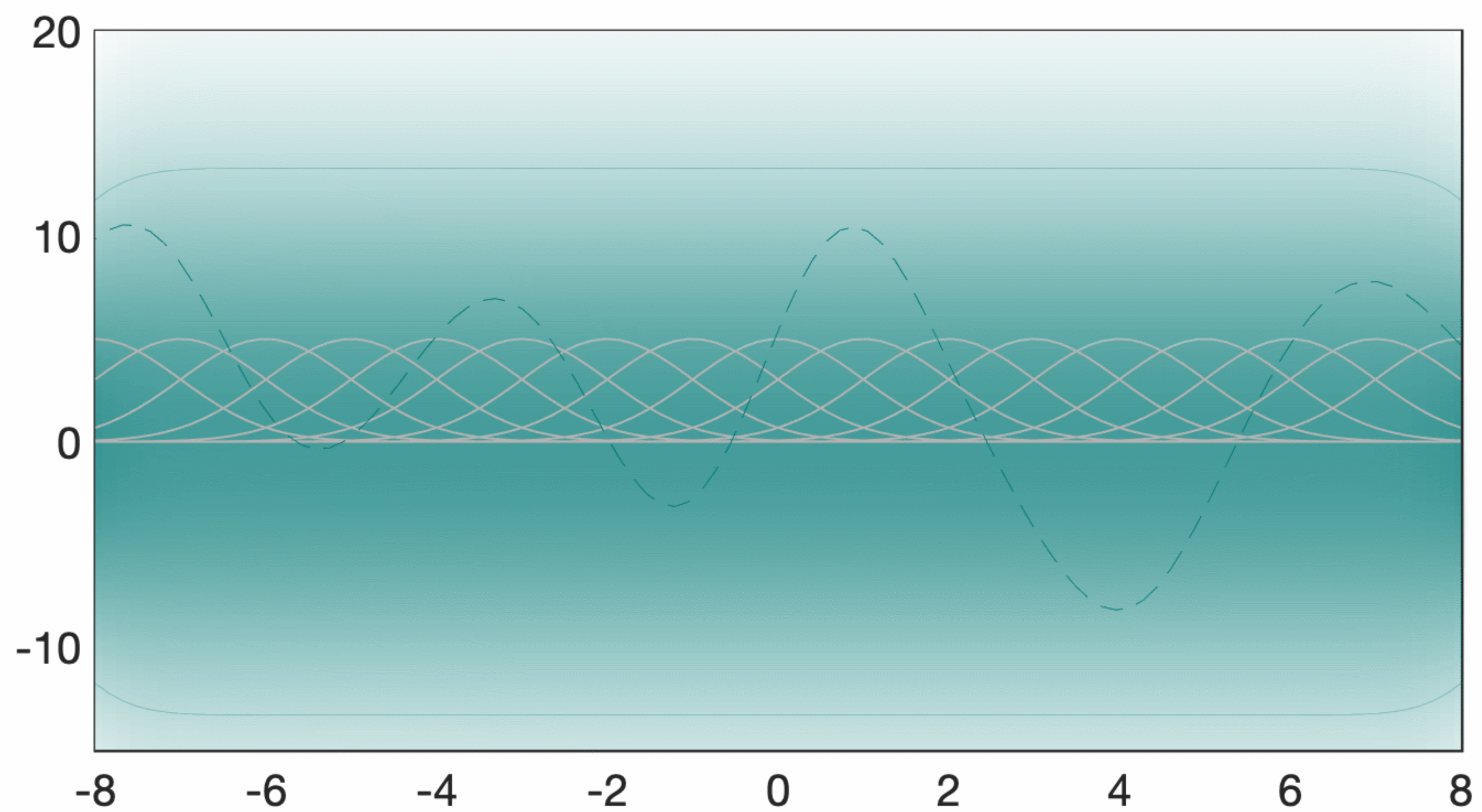
Bell Curve Regression

$$\phi(x) = \left(e^{-\frac{1}{2}(x-8)^2}, e^{-\frac{1}{2}(x-7)^2}, e^{-\frac{1}{2}(x-6)^2}, \ldots\right)^\top$$



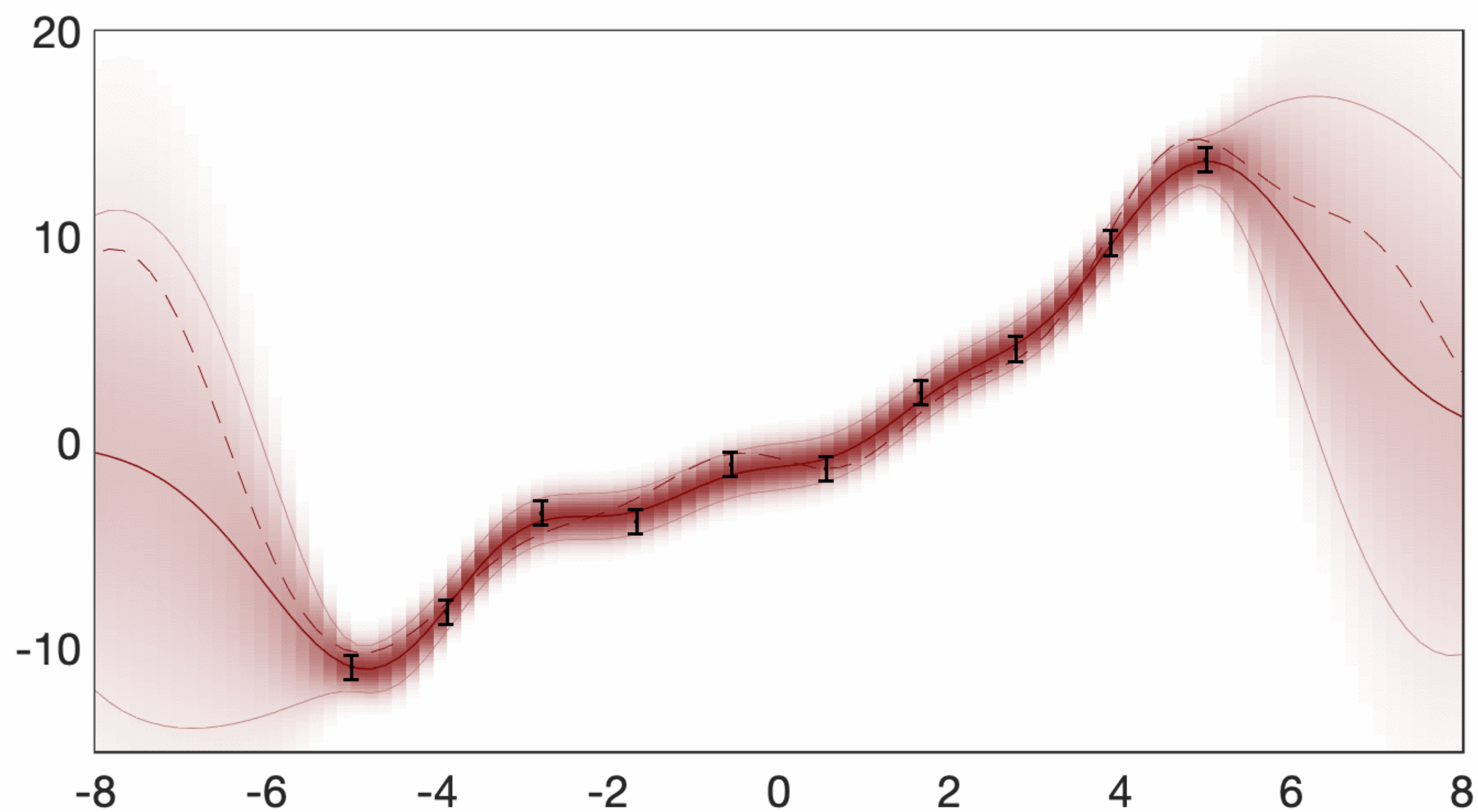
Bell Curve Regression

$$\phi(x) = \left(e^{-\frac{1}{2}(x-8)^2}, e^{-\frac{1}{2}(x-7)^2}, e^{-\frac{1}{2}(x-6)^2}, \ldots\right)^\top$$



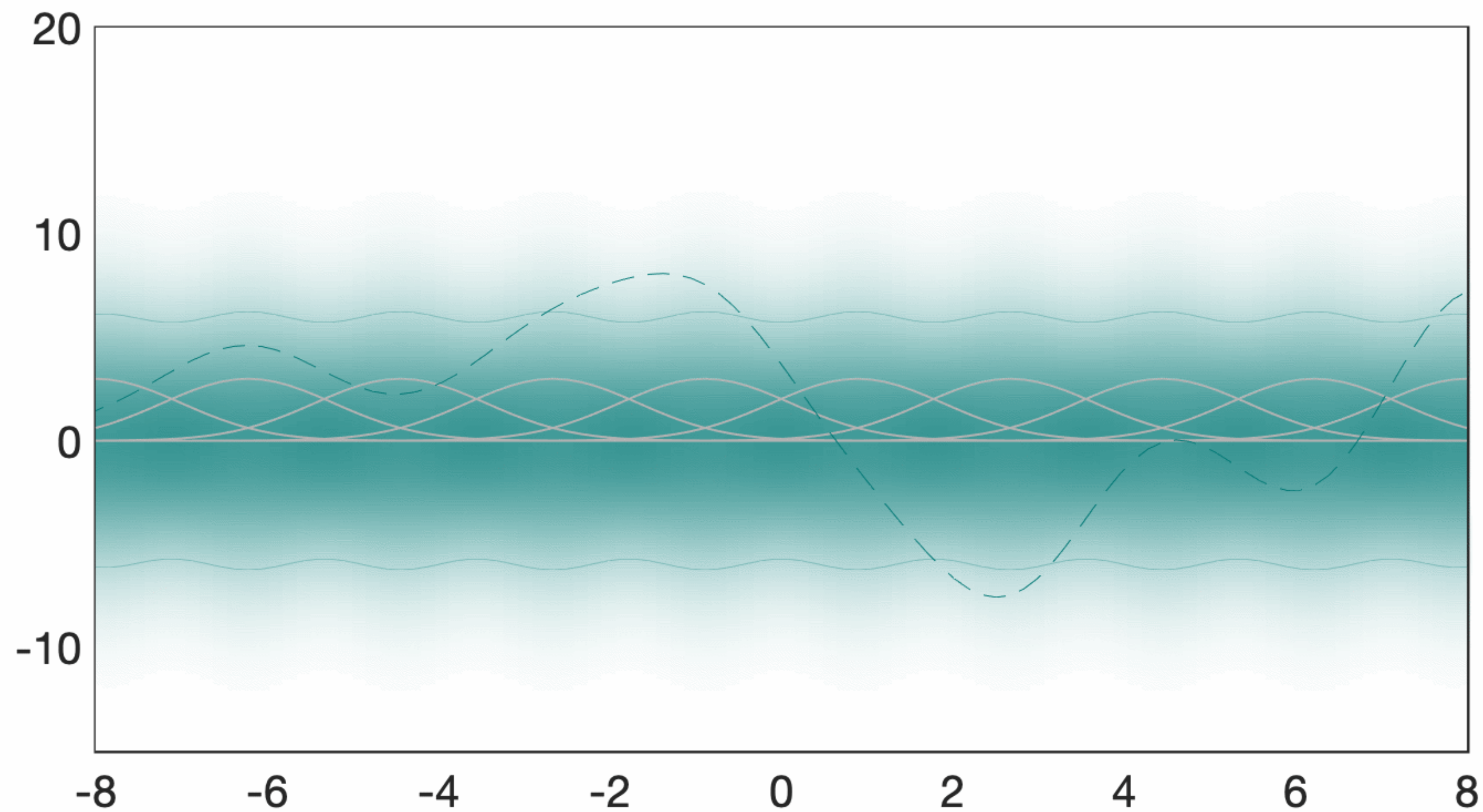
Bell Curve Regression

$$\phi(x) = \left(e^{-\frac{1}{2}(x-8)^2}, e^{-\frac{1}{2}(x-7)^2}, e^{-\frac{1}{2}(x-6)^2}, \dots\right)^\top$$



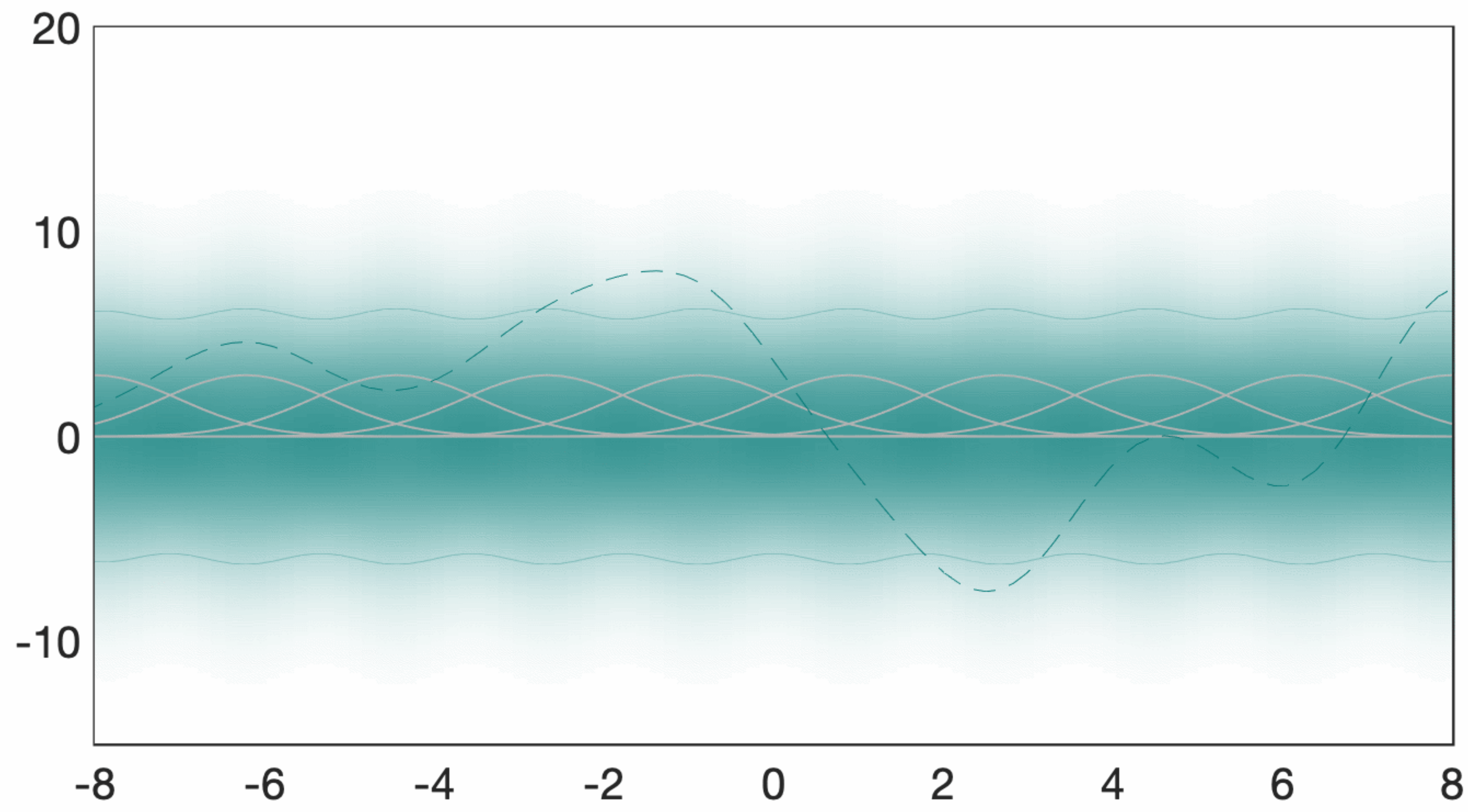
Exponentiated Squares

$$\phi(x) = \overbrace{\left(e^{-\frac{1}{2}(x-8)^2}, \ldots, e^{-\frac{1}{2}(x+8)^2}\right)}^{10}$$



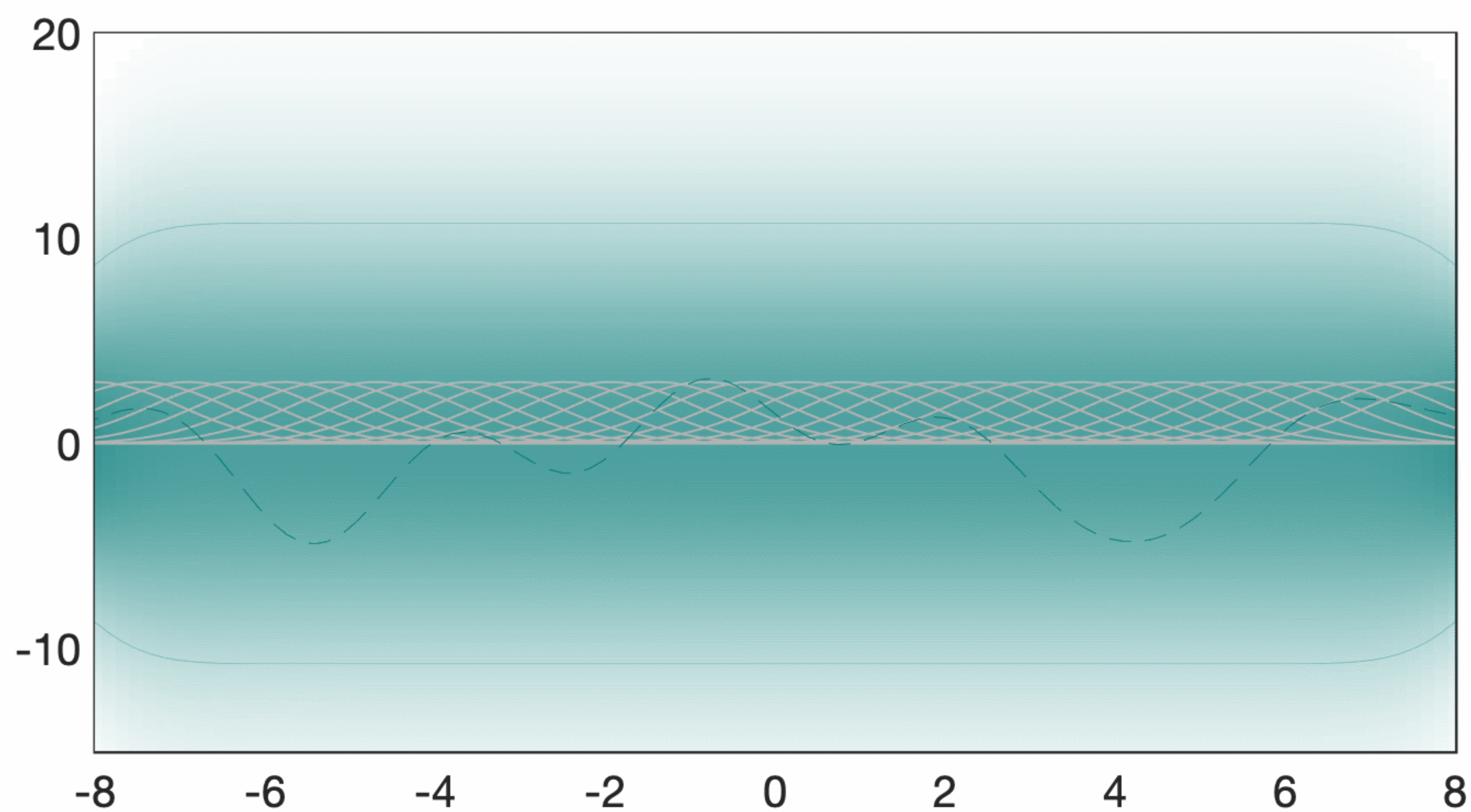
Exponentiated Squares

$$\phi(x) = \overbrace{\left(e^{-\frac{1}{2}(x-8)^2}, \ldots, e^{-\frac{1}{2}(x+8)^2}\right)}^{10}$$



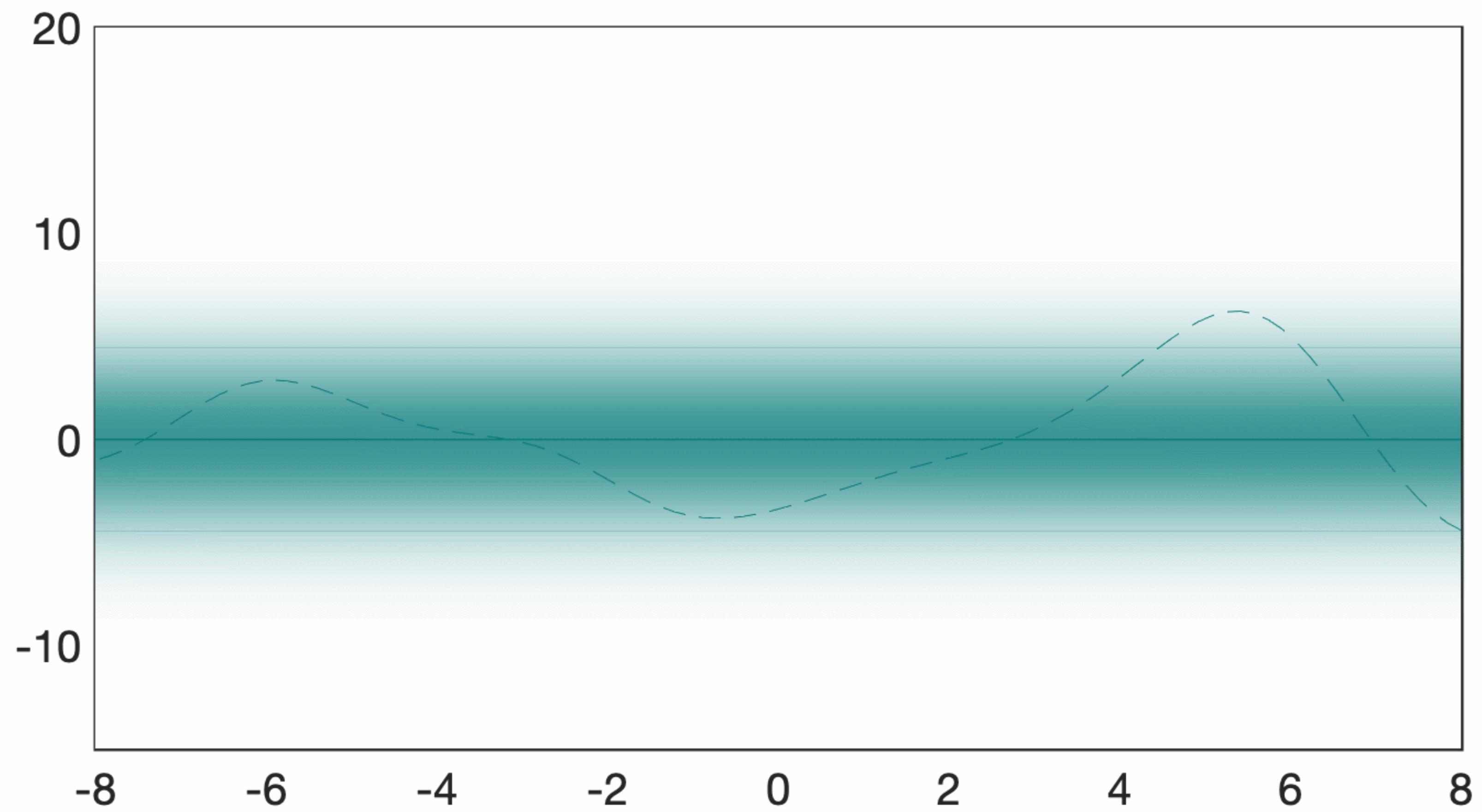
Exponentiated Squares

$$\phi(x) = \overbrace{\left(e^{-\frac{1}{2}(x-8)^2}, \ldots, e^{-\frac{1}{2}(x+8)^2}\right)}^{30}$$



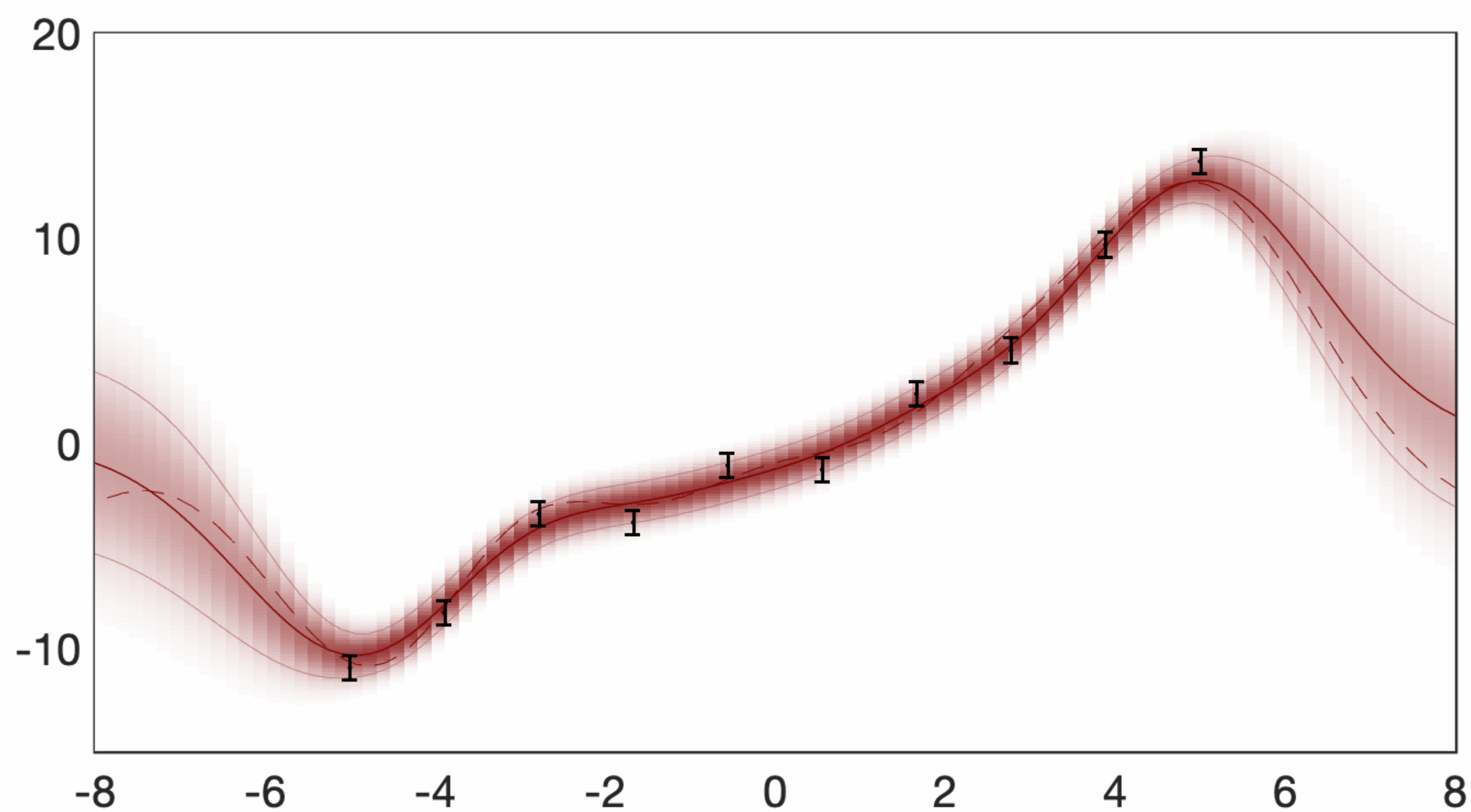
Exponentiated Squares

$$k(x_i, x_j) = 5\exp\left(-\frac{(x_i - x_j)^2}{4}\right)$$



Exponentiated Squares

$$k(x_i, x_j) = 5\exp\left(-\frac{(x_i - x_j)^2}{4}\right)$$



Gaussian Process

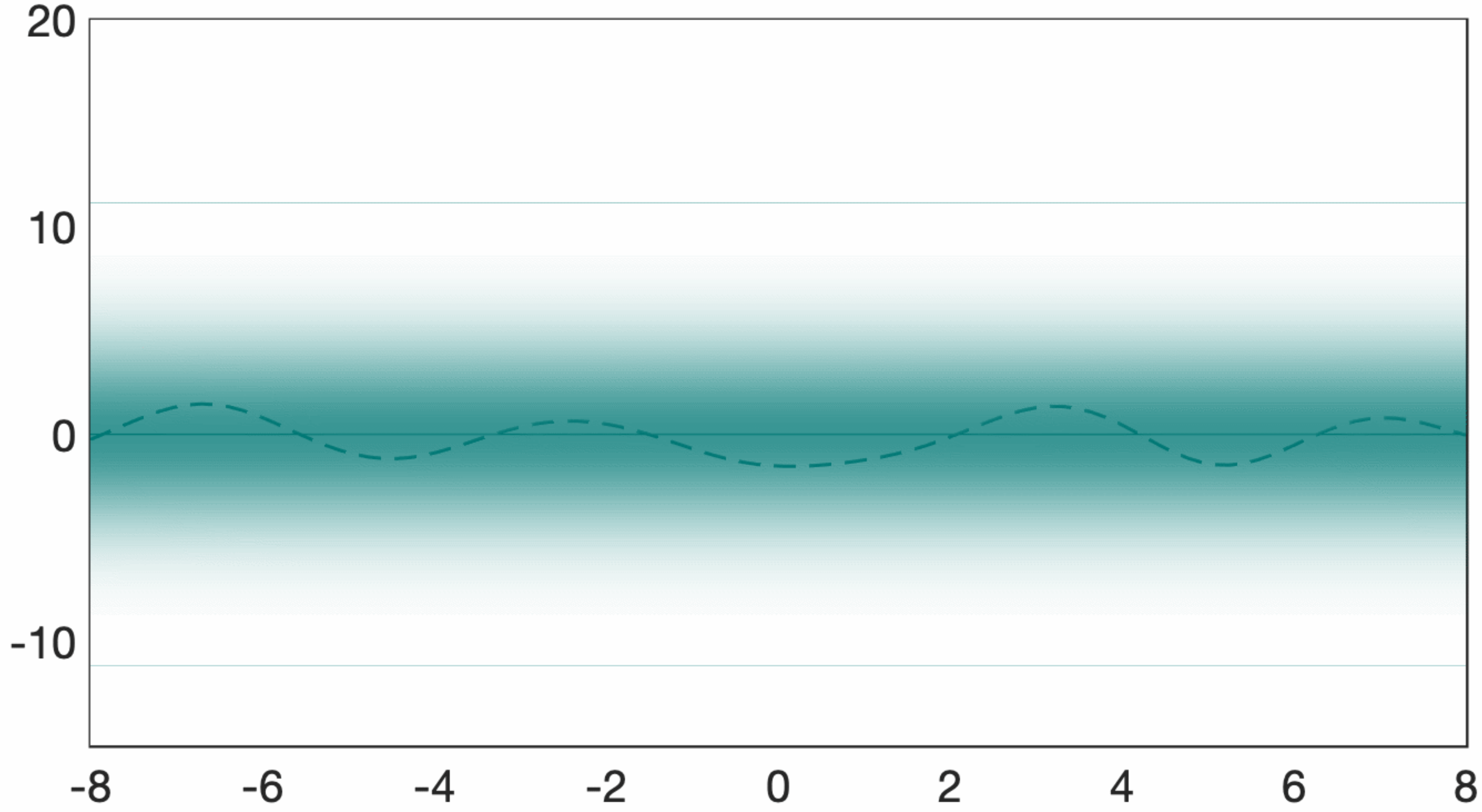
A Gaussian process

$$p(f) = \mathcal{GP}(f; \mu, k)$$

is a probability distribution over the function *f*, such that every finite restriction to function values $f_X = [f_{x_1}, \ldots, f_{x_N}]$ is a Gaussian distribution $p(f_X) = N(f_X; \mu_X, k_{XX})$.

Gaussian Process Posterior

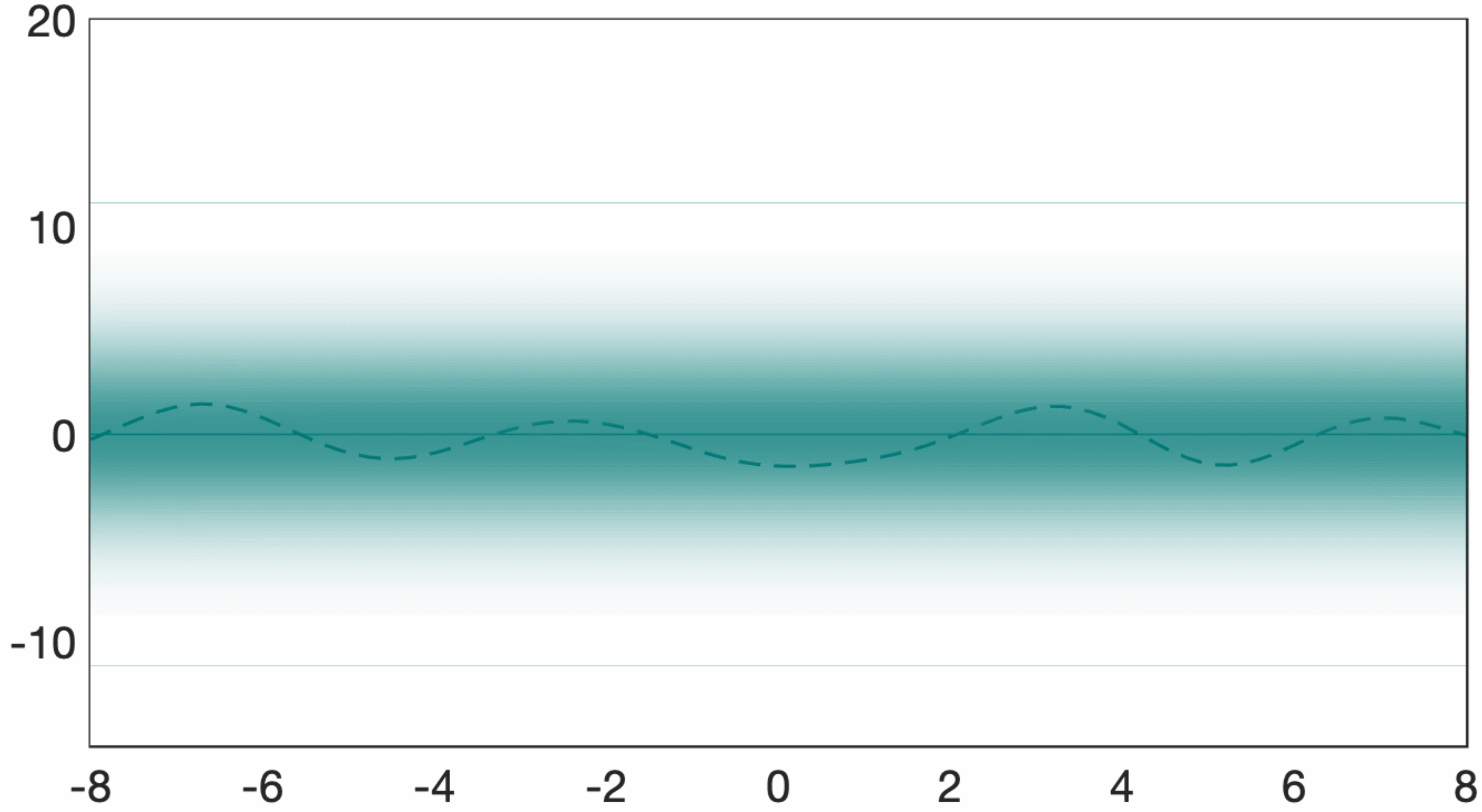
$$p(f_x \mid y, X) = \mathcal{N}(f_x; \mu_x + k_{Xx}^\top (k_{XX} + \sigma^2 I)^{-1}(y - \mu_x),$$
$$k_{xx} - k_{Xx}^\top (k_{XX} + \sigma^2 I)^{-1} k_{Xx})$$



$$p(f_x \mid y, \phi_X) = \mathcal{N}(f_x; m_x + k(x, X)^\top (k(X, X) + \sigma^2 I)^{-1}(y - m_X),$$
$$k(x, x) - k(x, X)^\top (k(X, X) + \sigma^2 I)^{-1} k(X, x))$$

Gaussian Process Posterior

$$p(f_x \mid y, X) = \mathcal{N}(f_x; \mu_x + k_{Xx}^\top (k_{XX} + \sigma^2 I)^{-1}(y - \mu_x),$$
$$k_{xx} - k_{Xx}^\top (k_{XX} + \sigma^2 I)^{-1} k_{Xx})$$



$$p(f_x \mid y, \phi_X) = \mathcal{N}(f_x; m_x + k(x, X)^\top (k(X, X) + \sigma^2 I)^{-1}(y - m_X),$$
$$k(x, x) - k(x, X)^\top (k(X, X) + \sigma^2 I)^{-1} k(X, x))$$

Gaussian Regression

$$p(f_X \mid y) = \frac{p(y \mid f_X)\,p(f_X)}{p(y)} = \frac{\overbrace{\mathcal{N}(y; f_X, \sigma^2 I)}^{\text{likelihood/loss}}\,\overbrace{\mathcal{N}(f_X; m_X, k_{XX})}^{\text{prior/regularizer}}}{\mathcal{N}(y; m_X, k_{XX} + \sigma^2 I)}$$

Gaussian Regression

likelihood/loss prior/regularizer

$$p(f_X \mid y) = \frac{p(y \mid f_X)p(f_X)}{p(y)} = \frac{\mathcal{N}(y; f_X, \sigma^2 I)\mathcal{N}(f_X; m_X, k_{XX})}{\mathcal{N}(y; m_X, k_{XX} + \sigma^2 I)}$$

$$\begin{aligned} -2 \log p(f \mid y) &= (y - f_X)^\top \sigma^{-2} I (y - f_X) + (f_X - m_X)^\top k_{XX}^{-1}(f_X - m_X) + \text{const.} \\ &= \sigma^{-2}\|y - f_X\|_I^2 + \|f_X - m_X\|_{k_{XX}}^2 + \text{const.} \end{aligned}$$

Gaussian Regression

$$p(f_X \mid y) = \frac{p(y \mid f_X)p(f_X)}{p(y)} = \frac{\overset{\text{likelihood/loss}}{\mathcal{N}(y; f_X, \sigma^2 I)}\,\overset{\text{prior/regularizer}}{\mathcal{N}(f_X; m_X, k_{XX})}}{\mathcal{N}(y; m_X, k_{XX} + \sigma^2 I)}$$

$$\begin{aligned}
-2\log p(f \mid y) &= (y - f_X)^\top \sigma^{-2} I (y - f_X) + (f_X - m_X)^\top k_{XX}^{-1} (f_X - m_X) + \text{const.} \\
&= \sigma^{-2} \|y - f_X\|_I^2 + \|f_X - m_X\|_{k_{XX}}^2 + \text{const.}
\end{aligned}$$

The GP posterior mean is the **regularized least-squares** estimate (a.k.a. kernel ridge regression)

Gaussian Process **Bandit Optimization**

- Given: Unlabeled data *U*
- For t = 1, … ,*T*
  - Select *x* from *U*
  - Query label *y* of *x*
  - Add (*x*,*y*) to labeled set *S*
  - Train *f*(*x*) on *S*
- Goal:
  - Find *x* that maximize *f(x)*
  - Minimize "regret":

$$Tf(x^*) - (f(x_1) + f(x_2) + \cdots + f(x_T))$$

Best experiment only

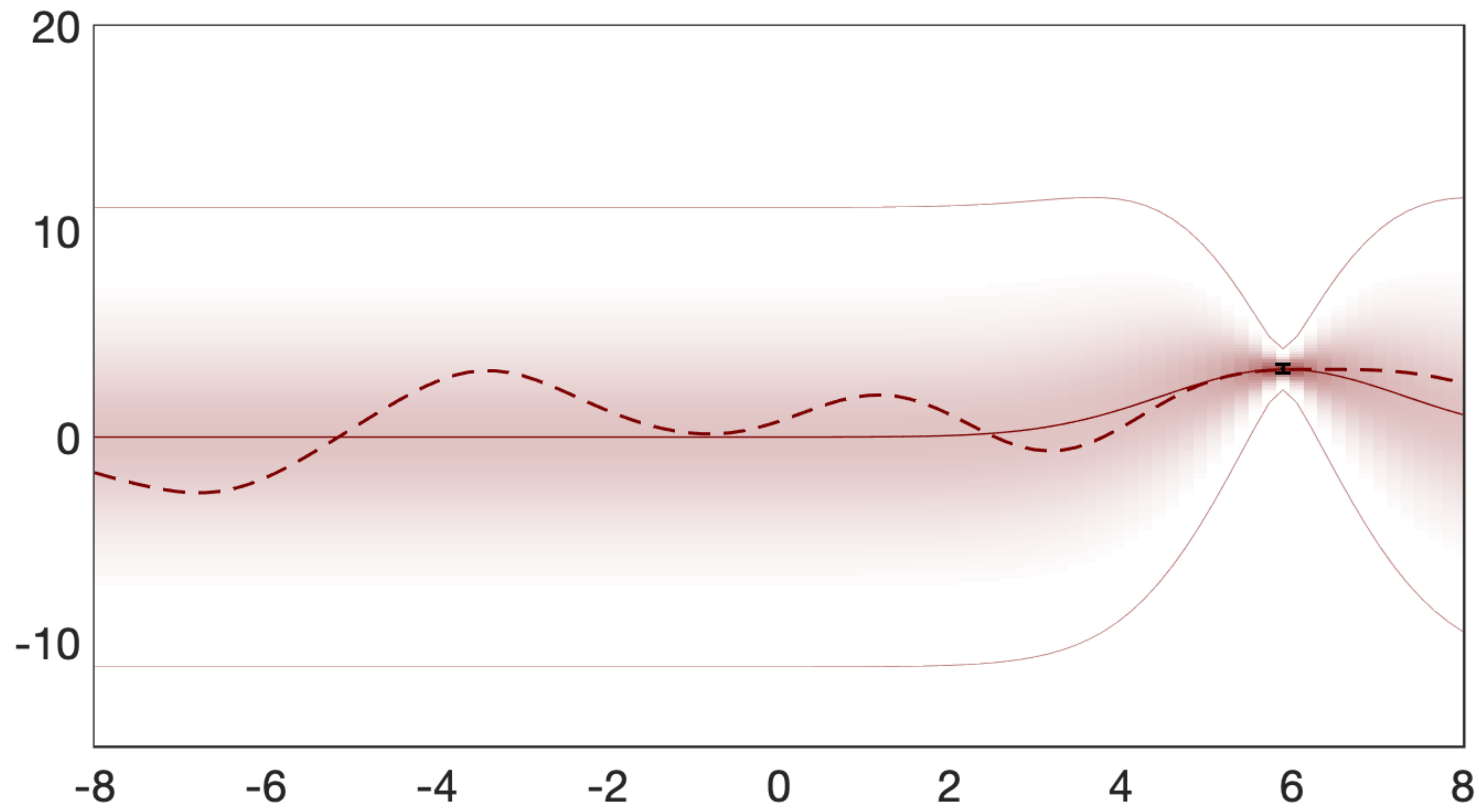
Each experiment matters

Thompson Sampling (TS) [Agrawal et al. '12]

$$x_{t+1} = \arg\max_x \tilde{f}(x) \qquad \tilde{f} \sim p(f \mid y)$$

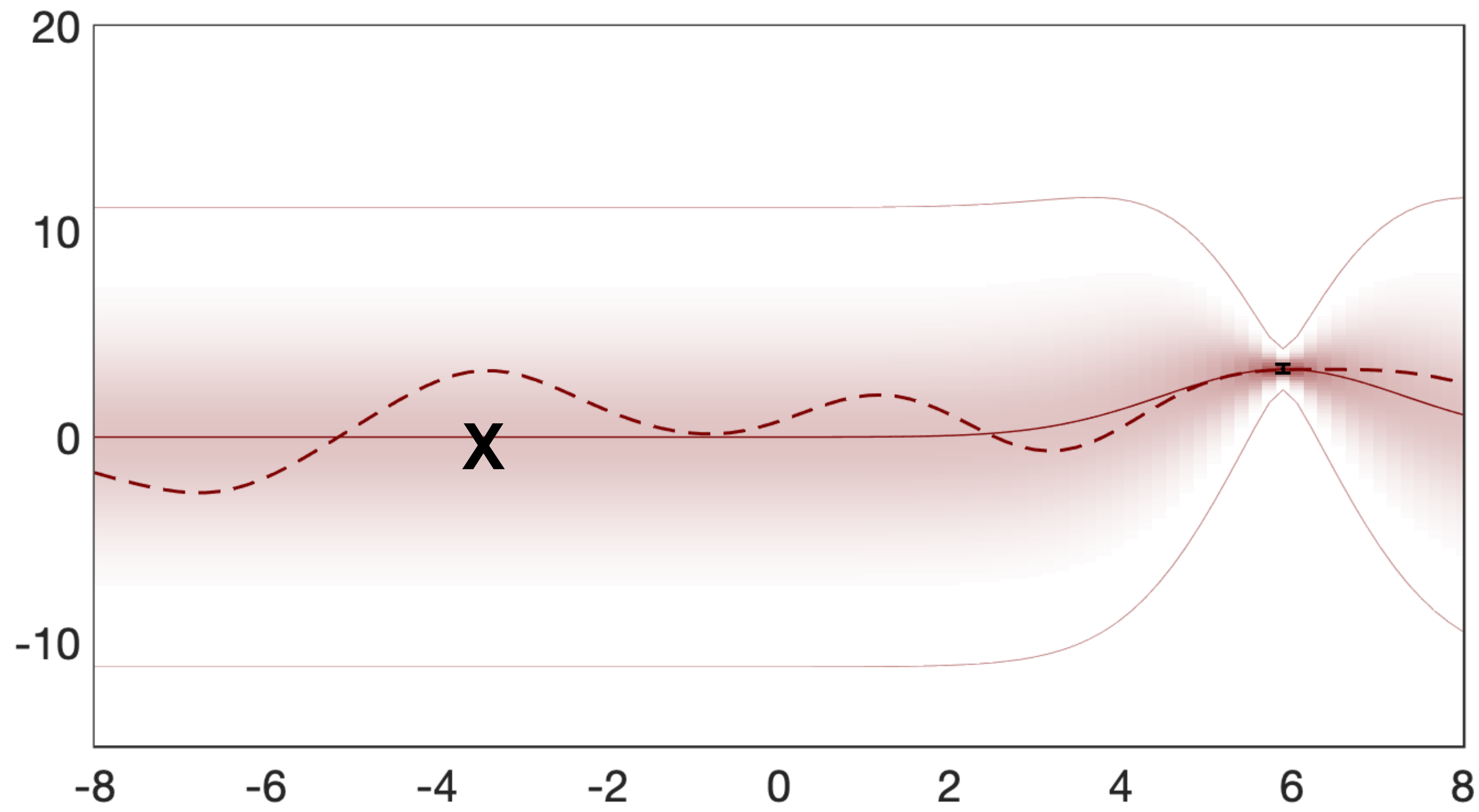
Thompson Sampling (TS) [Agrawal et al. '12]

$$x_{t+1} = \arg\max_x \tilde{f}(x) \qquad \tilde{f} \sim p(f \mid y)$$



Thompson Sampling (TS) [Agrawal et al. '12]

$$x_{t+1} = \arg\max_{x} \tilde{f}(x) \qquad \tilde{f} \sim p(f \mid y)$$



GP Upper Confidence Bound [Thompson '33; Srinivas et al. '09]

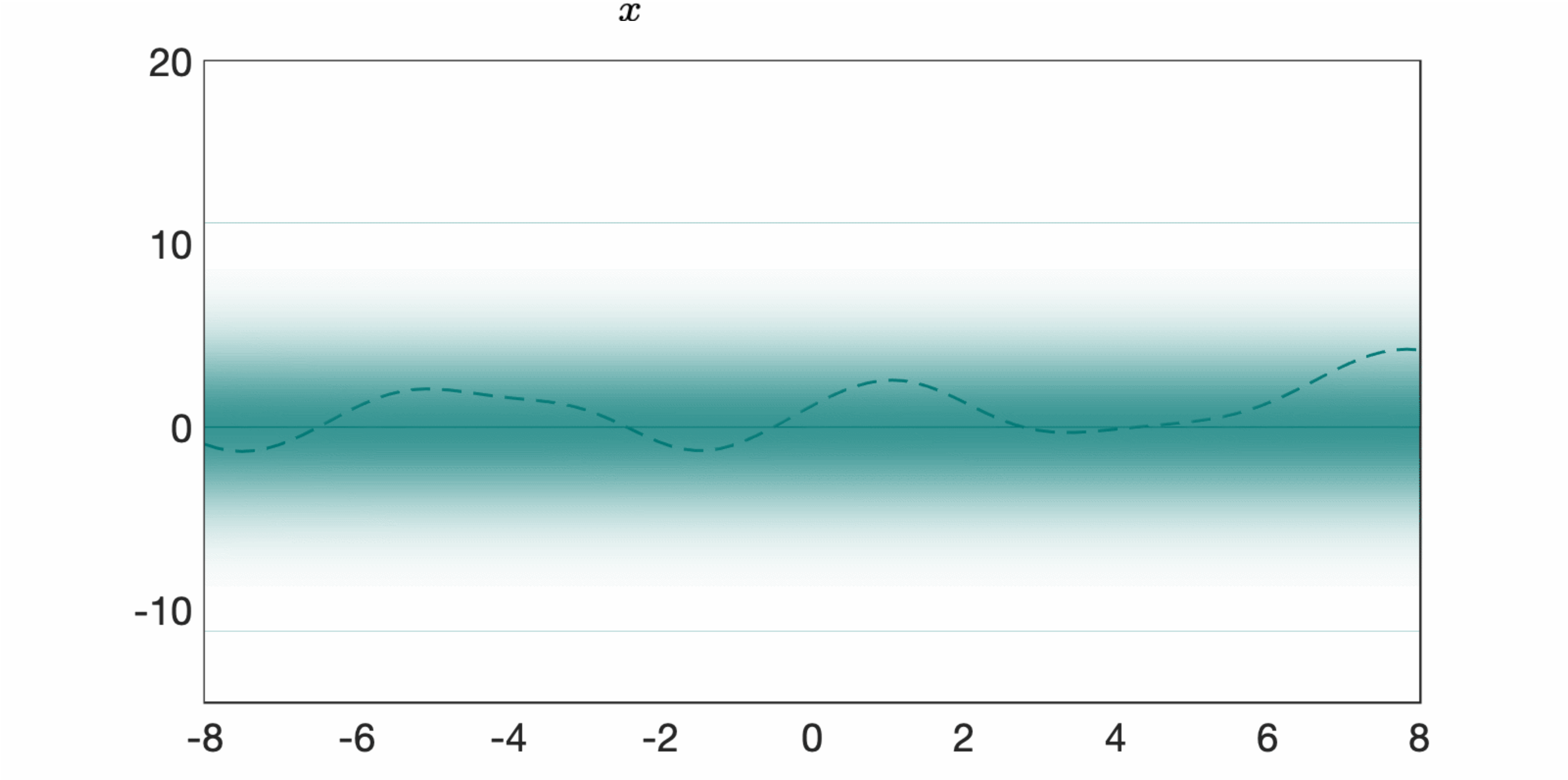
"Optimism under the face of uncertainty"

$$x_{t+1} = \arg\max_x \mu_t(x) + \beta_t \sqrt{k_t(x,x)}$$

GP Upper Confidence Bound [Thompson '33; Srinivas et al. '09]

"Optimism under the face of uncertainty"

$$x_{t+1} = \arg\max_x \mu_t(x) + \beta_t \sqrt{k_t(x,x)}$$



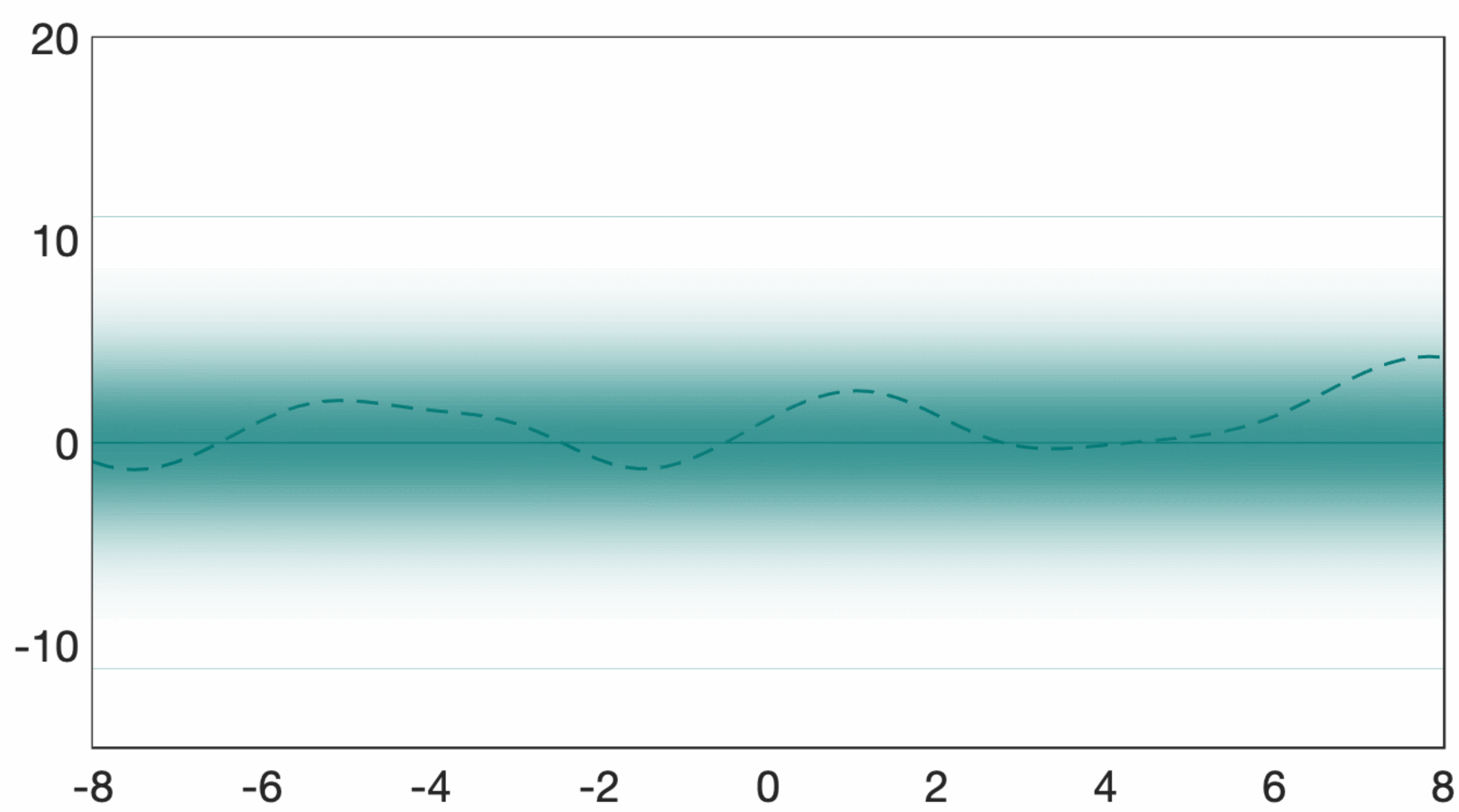
Algorithmic Questions

Analyze convergence to $f(x^*)$?

Corrupted or indirect measurements?

Efficiently search combinatorial design spaces?

Incorporate domain knowledge such as physics?



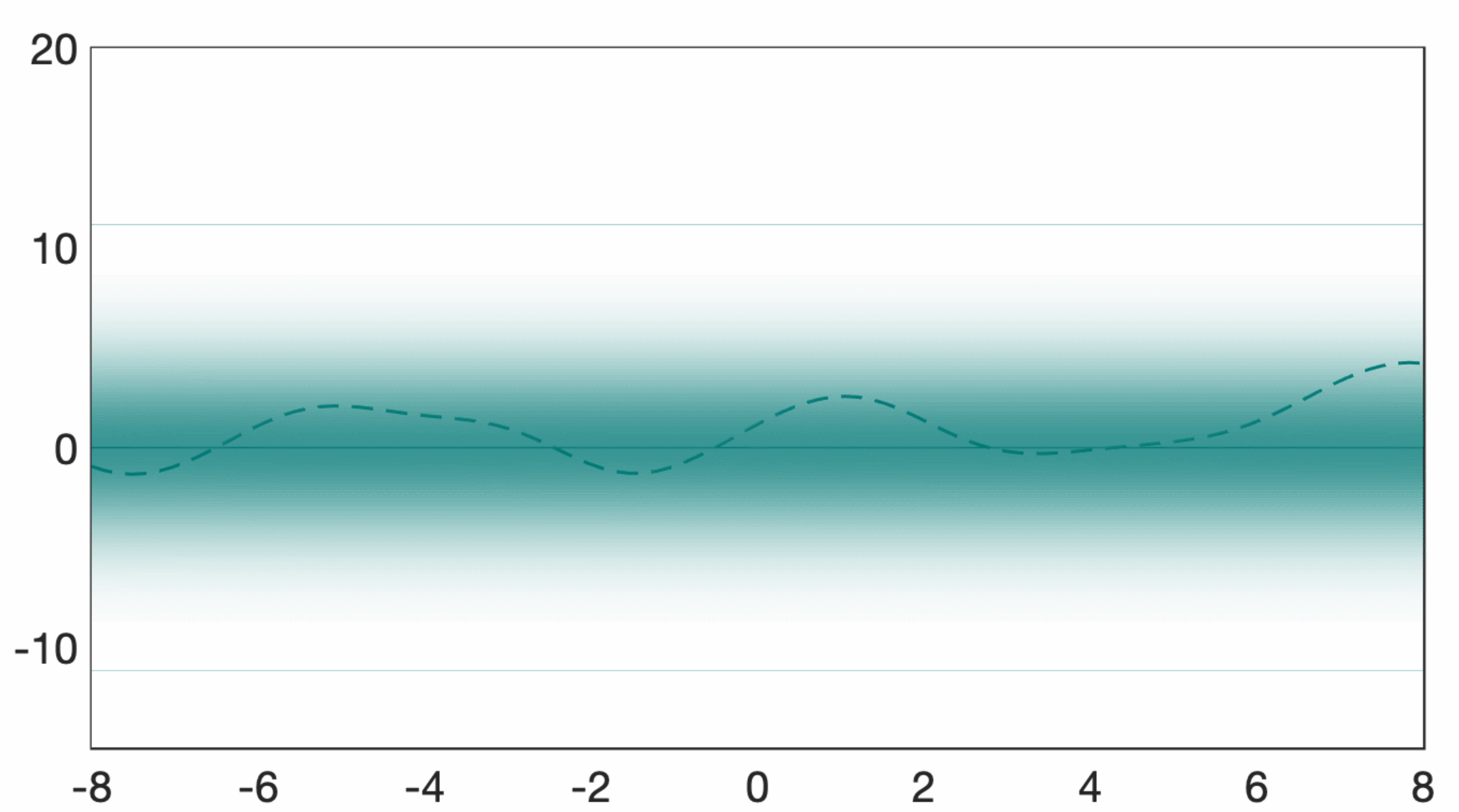
Algorithmic Questions

Analyze convergence to f(x*)?

Corrupted or indirect measurements?

Efficiently search combinatorial design spaces?

Incorporate domain knowledge such as physics?



A General Framework for Multi-fidelity Bayesian Optimization with Gaussian Processes
J. Song, Y. Chen, Y. Yue. AISTATS'19.

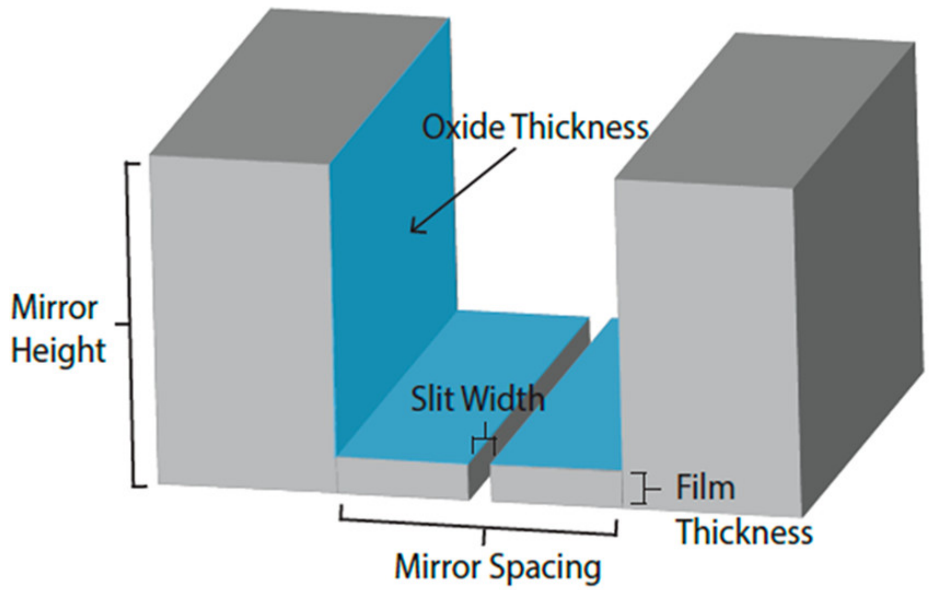
Optimizing Photonic Nanostructures via Multi-fidelity Gaussian Processes
J. Song, Y. S. Tokpanov, Y. Chen, D. Fleischman, K. T. Fountaine, H. A. Atwater, Y. Yue.
NeurIPS Workshop on Machine Learning for Molecules and Materials, 2018

Corrupted / Indirect Measurements: Multi-fidelity Bayesian Optimization

A General Framework for Multi-fidelity Bayesian Optimization with Gaussian Processes
J. Song, Y. Chen, Y. Yue. AISTATS'19.

Optimizing Photonic Nanostructures via Multi-fidelity Gaussian Processes
J. Song, Y. S. Tokpanov, Y. Chen, D. Fleischman, K. T. Fountaine, H. A. Atwater, Y. Yue.
NeurIPS Workshop on Machine Learning for Molecules and Materials, 2018

Nano-photonics Structure Design

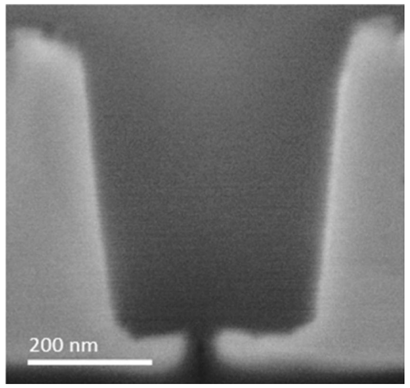


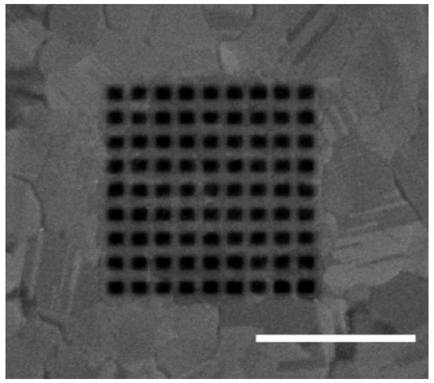
Jialin
Song

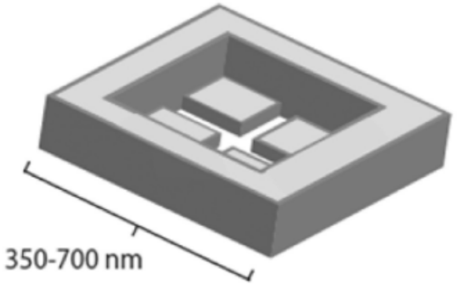
Yury
Tokpanov

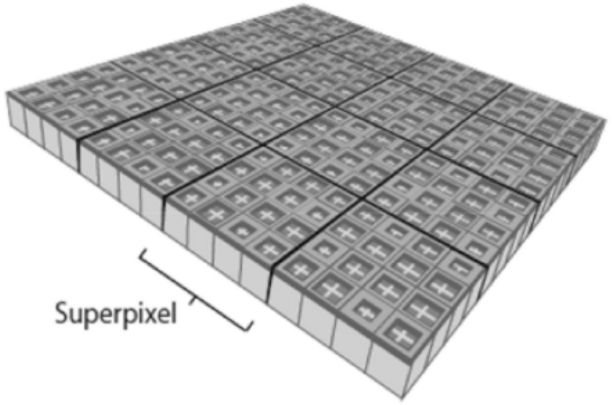
Harry
Atwater

Yisong
Yue



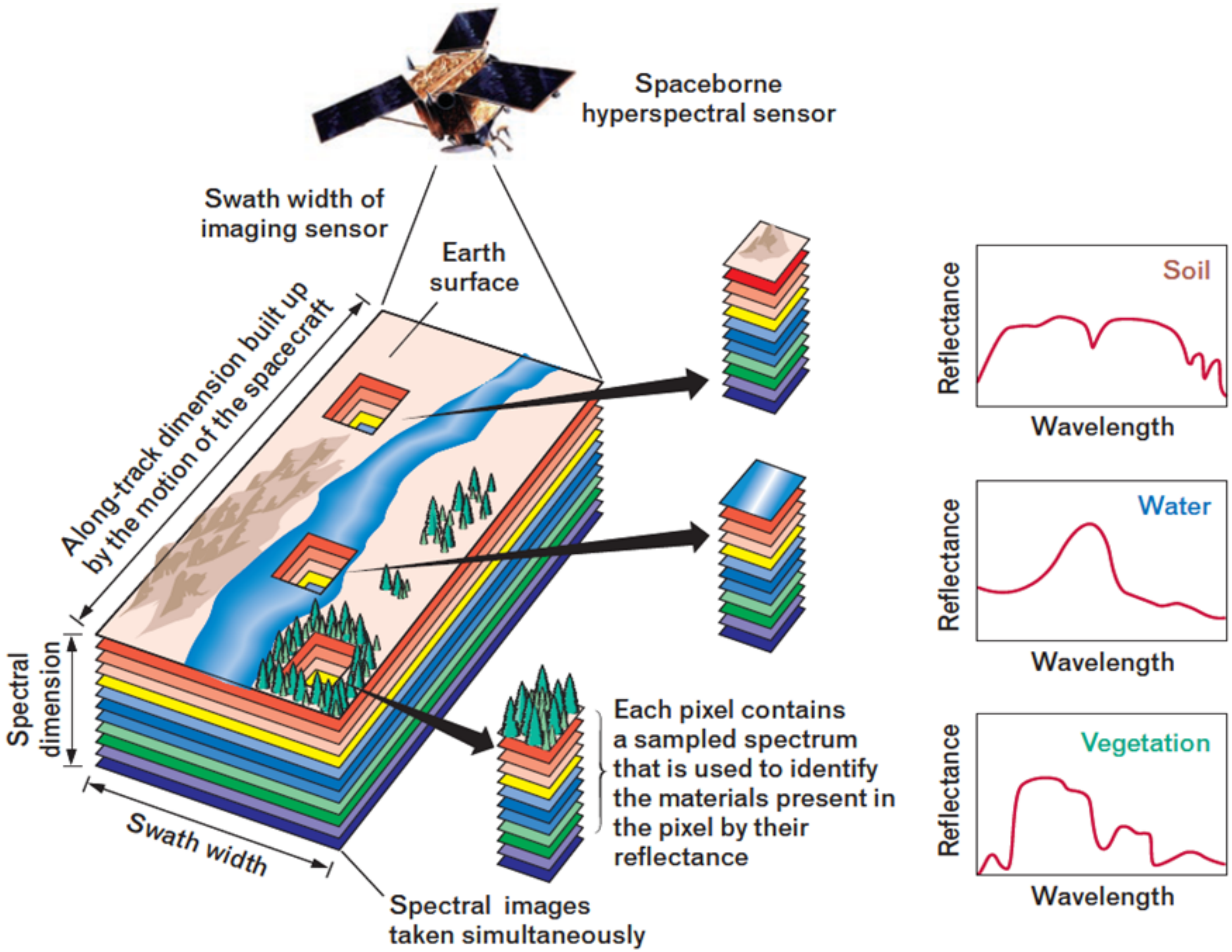






Fleischman et al.: https://doi.org/10.1021/acsphotonics.8b01634

Hyperspectral Imaging



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

Fitness Function (Figure of Merit)

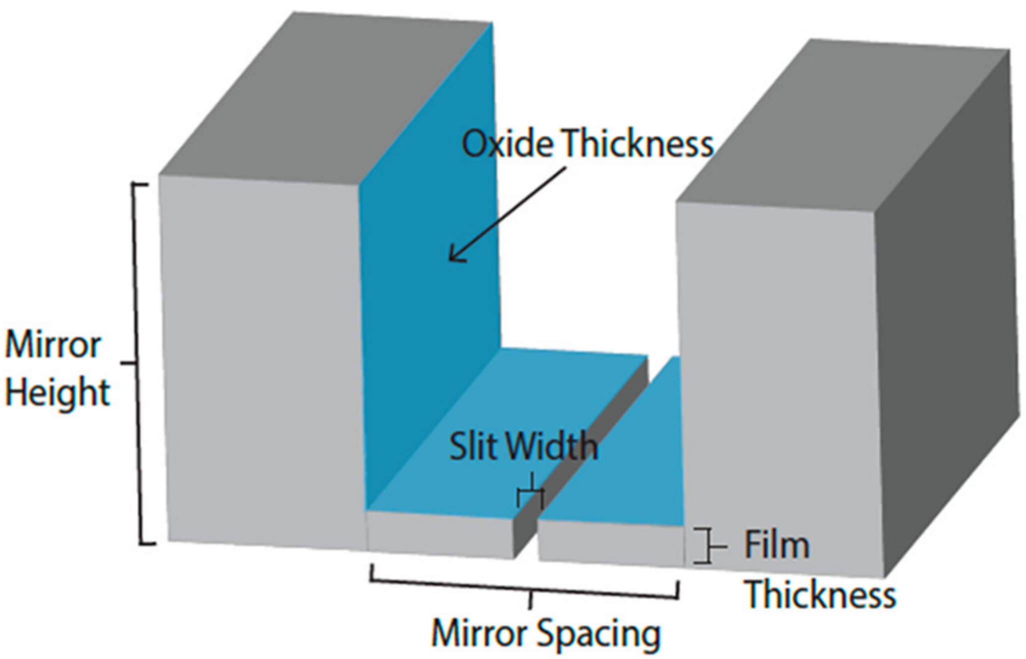
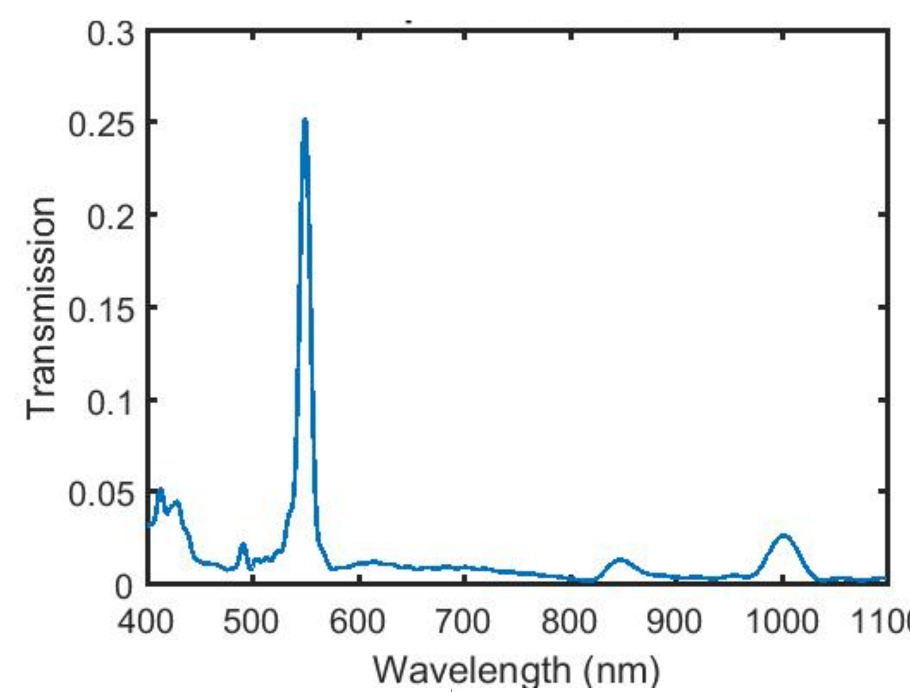


Image Credit: Yury Tokpanov

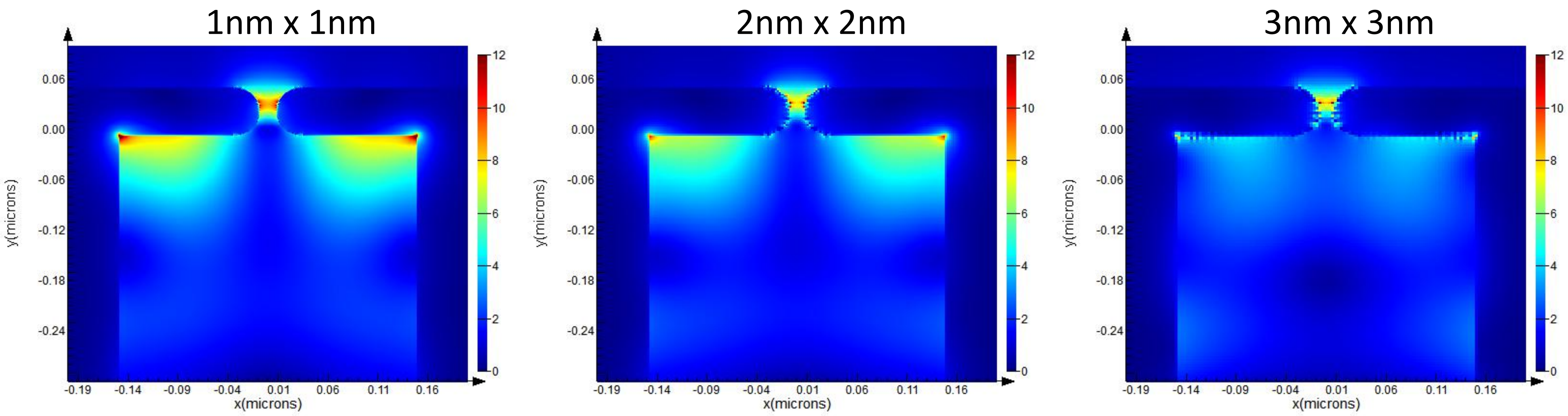
Maxwell's equations FDTD solver



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

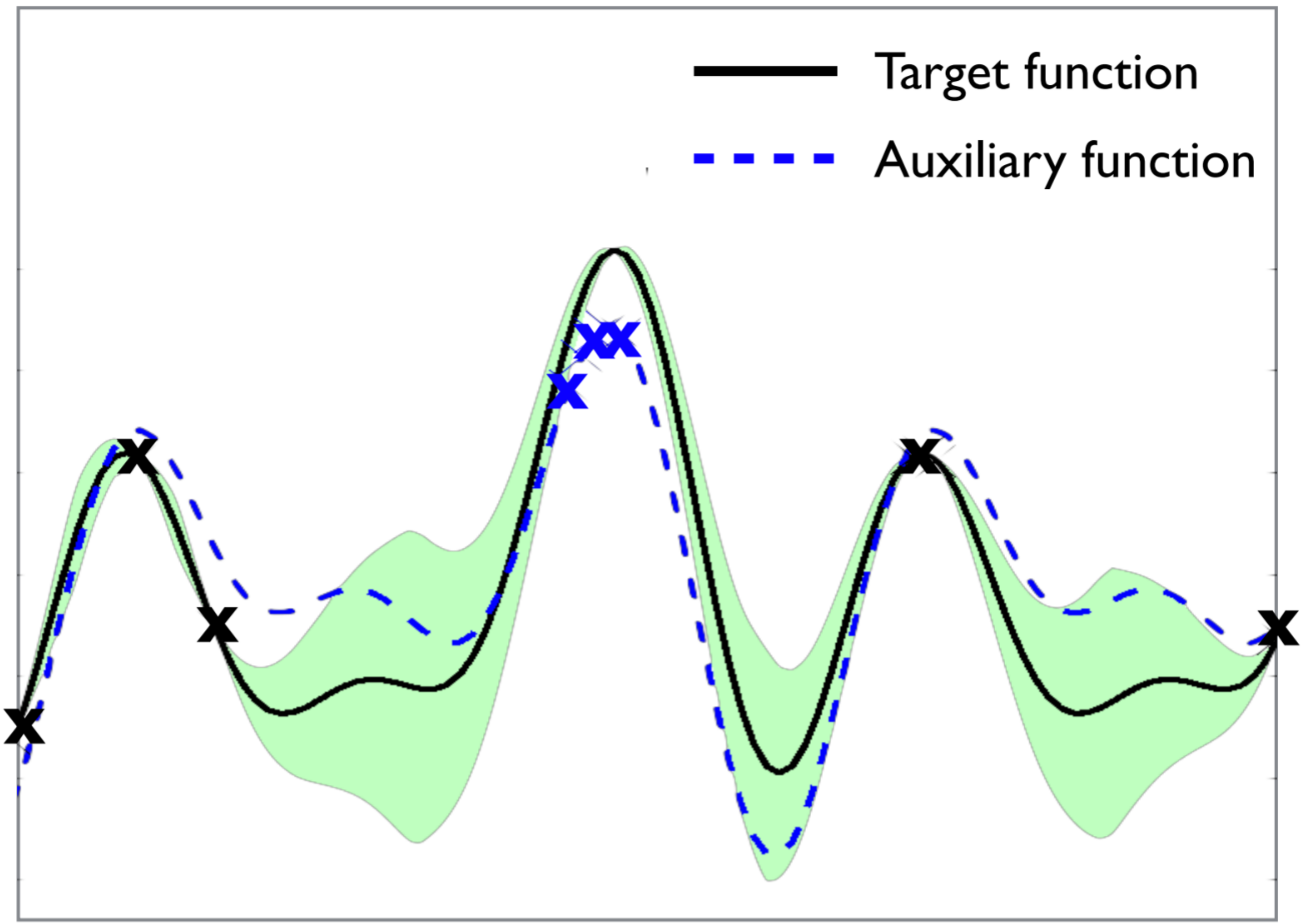
Multi-Fidelity Simulations

- 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

Algorithmic Insights

Jialin Song

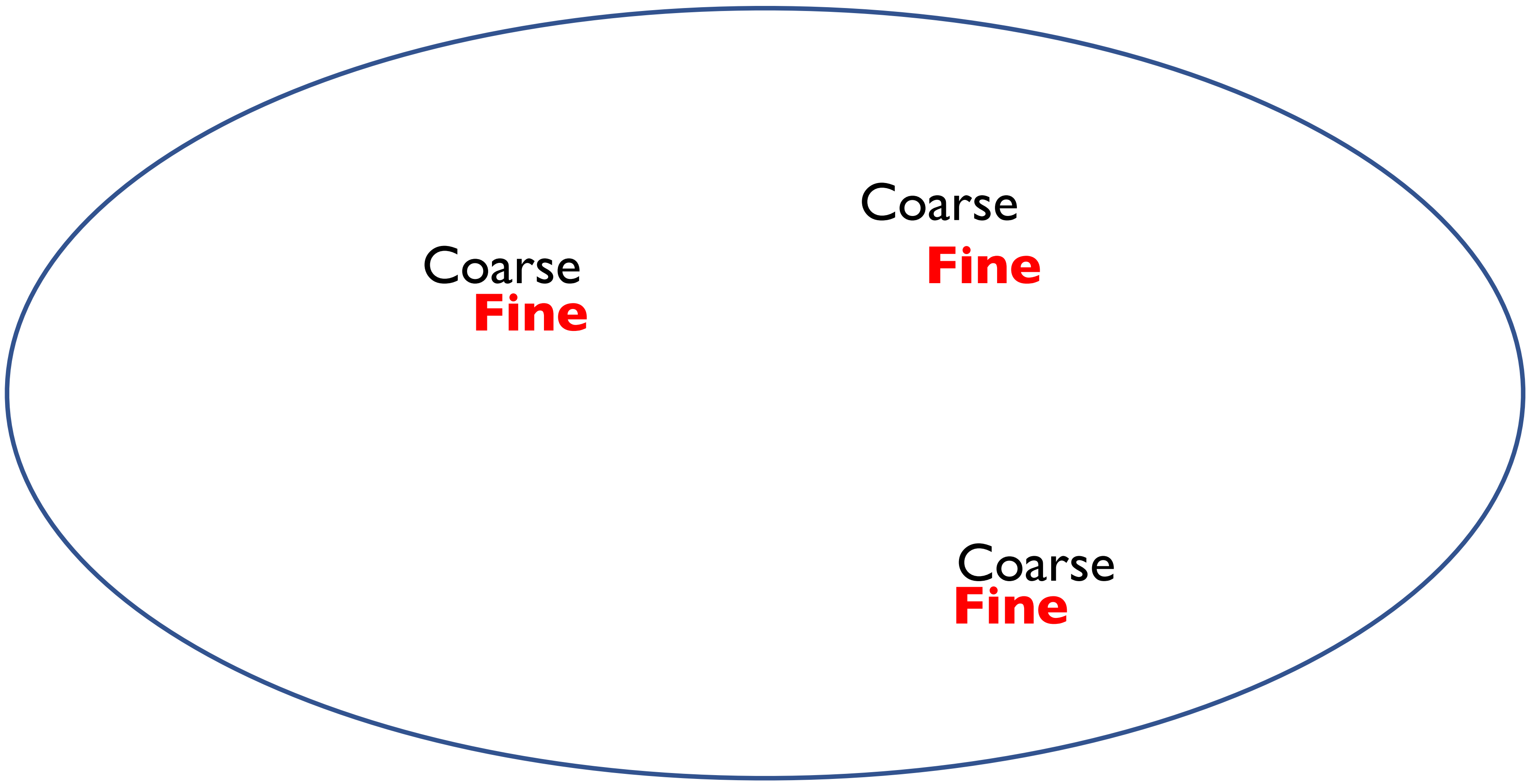
Yisong Yue

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

Algorithmic Insights

Jialin Song

Yisong Yue

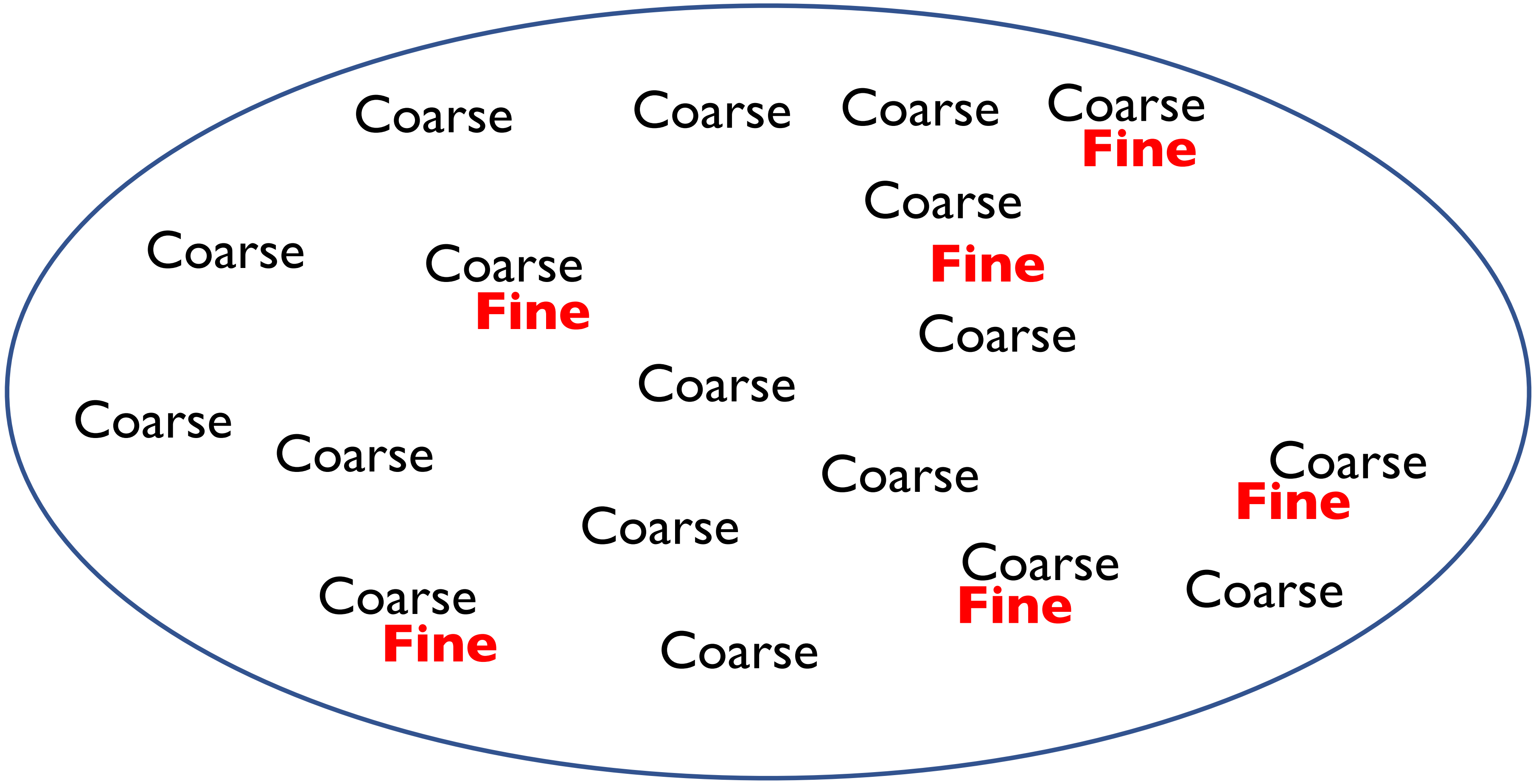


A General Framework for Multi-fidelity Bayesian Optimization with Gaussian Processes, Jialin Song et al., AISTATS 2019

Algorithmic Insights

Jialin
Song

Yisong
Yue

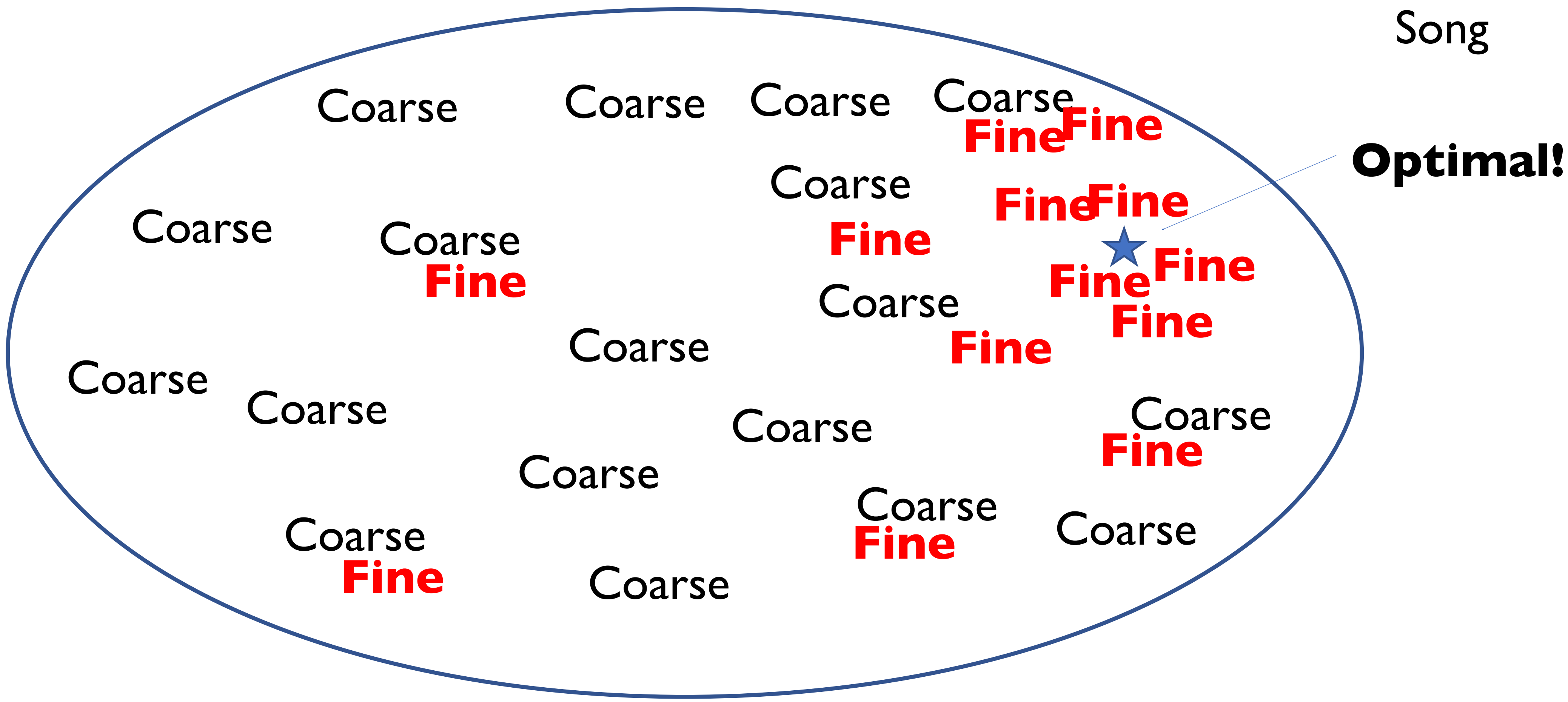


A General Framework for Multi-fidelity Bayesian Optimization with Gaussian Processes, Jialin Song et al., AISTATS 2019

Algorithmic Insights


Jialin Song


Yisong Yue



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

Algorithmic Insights


Jialin Song   Yisong Yue

Coarse   Coarse   Coarse   Coarse

Optimal!

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**

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

MF-MI-Greedy: The Multi-fidelity Optimization Algorithm

Def. [Multi-fidelity cumulative regret]

optimal cumulative reward - $\sum_j$ (reward of action at step j)

Multi-fidelity Mutual Information Greedy Optimization (MF-MI-Greedy)

Input Budget Λ; cost for each fidelity; joint (GP) prior.

Output Optimizer of the target function

Start $S \leftarrow \varnothing$

Loop $L \leftarrow$ Explore lower fidelity ***w. information gain per unit-cost***

$x^* \leftarrow$ Select target fidelity ***using single-fidelity optimization***

$S \leftarrow S \cup L \cup \{x^*\}$;

Until *budget exhausted*

MF-MI-Greedy: The Multi-fidelity Optimization Algorithm

Def. [Multi-fidelity cumulative regret]

optimal cumulative reward - $\sum_j$ (reward of action at step j)

Multi-fidelity Mutual Information Greedy Optimization (MF-MI-Greedy)

Input Budget Λ; cost for each fidelity; joint (GP) prior.

Output Optimizer of the target function

Start $S \leftarrow \varnothing$

Loop $L \leftarrow$ Explore lower fidelity ***w. information gain per unit-cost***

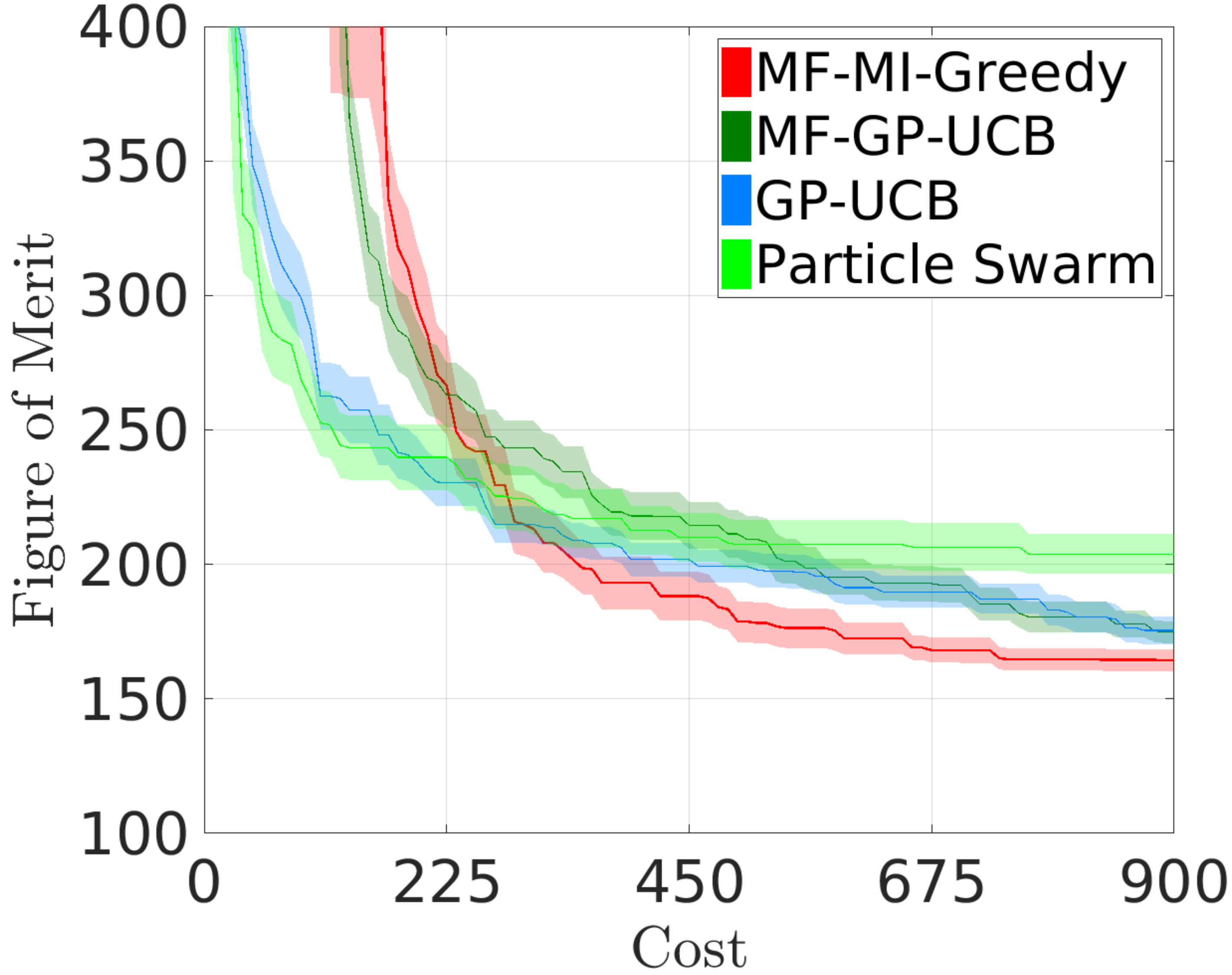
$x^* \leftarrow$ Select target fidelity ***using single-fidelity optimization***

$S \leftarrow S \cup L \cup \{x^*\}$;

Until *budget exhausted*

Thm. The cumulative regret of MF-MI-Greedy is $C_1\sqrt{\Lambda\gamma_t} + C_2\gamma_L \cdot o\left(\sqrt{\Lambda}\right)$.

Results: Optimizing Photonic Nanostructures



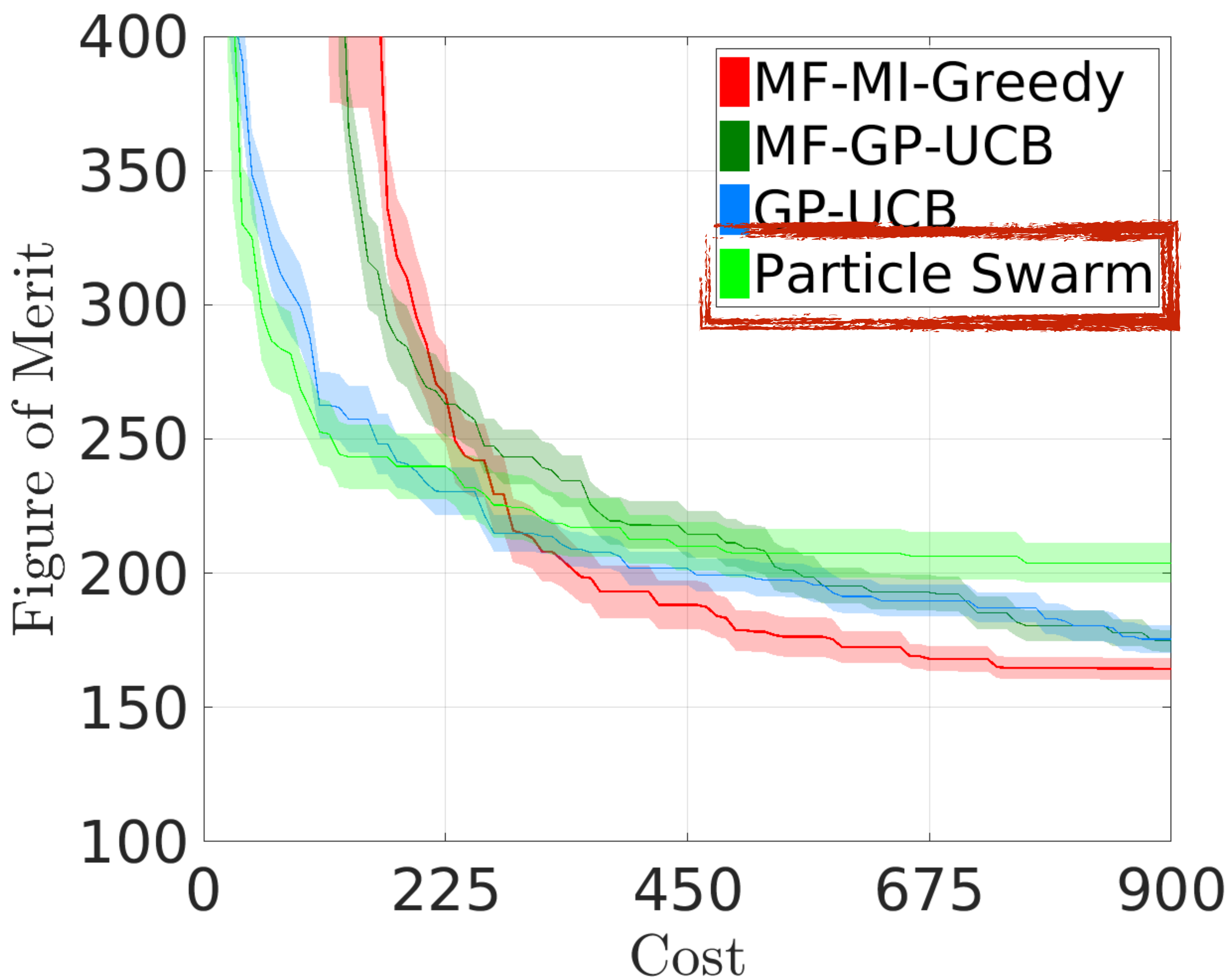
multi-fidelity baseline

single-fidelity baseline

Optimizing Figure of Merit (FOM) with 3 fidelities

[Song, Tokpanov, Chen, Fleischman, Fountaine, Atwater, Yue, 2018]

Results: Optimizing Photonic Nanostructures



multi-fidelity baseline

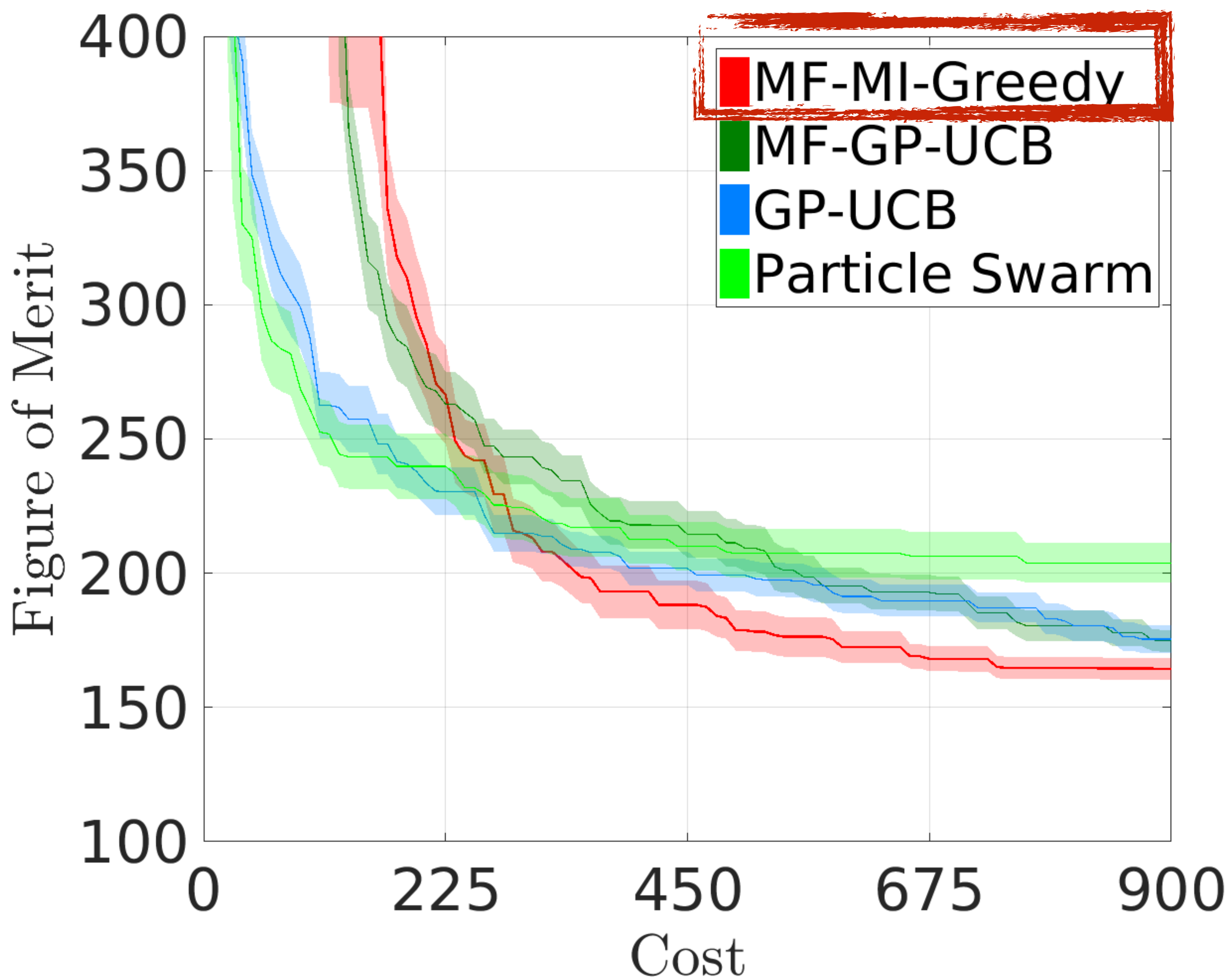
single-fidelity baseline

standard approach in nano-photonics

Optimizing Figure of Merit (FOM) with 3 fidelities

[Song, Tokpanov, Chen, Fleischman, Fountaine, Atwater, Yue, 2018]

Results: Optimizing Photonic Nanostructures



ours
multi-fidelity baseline
single-fidelity baseline
standard approach in nano-photonics

Balances different costs

State-of-the-art performance

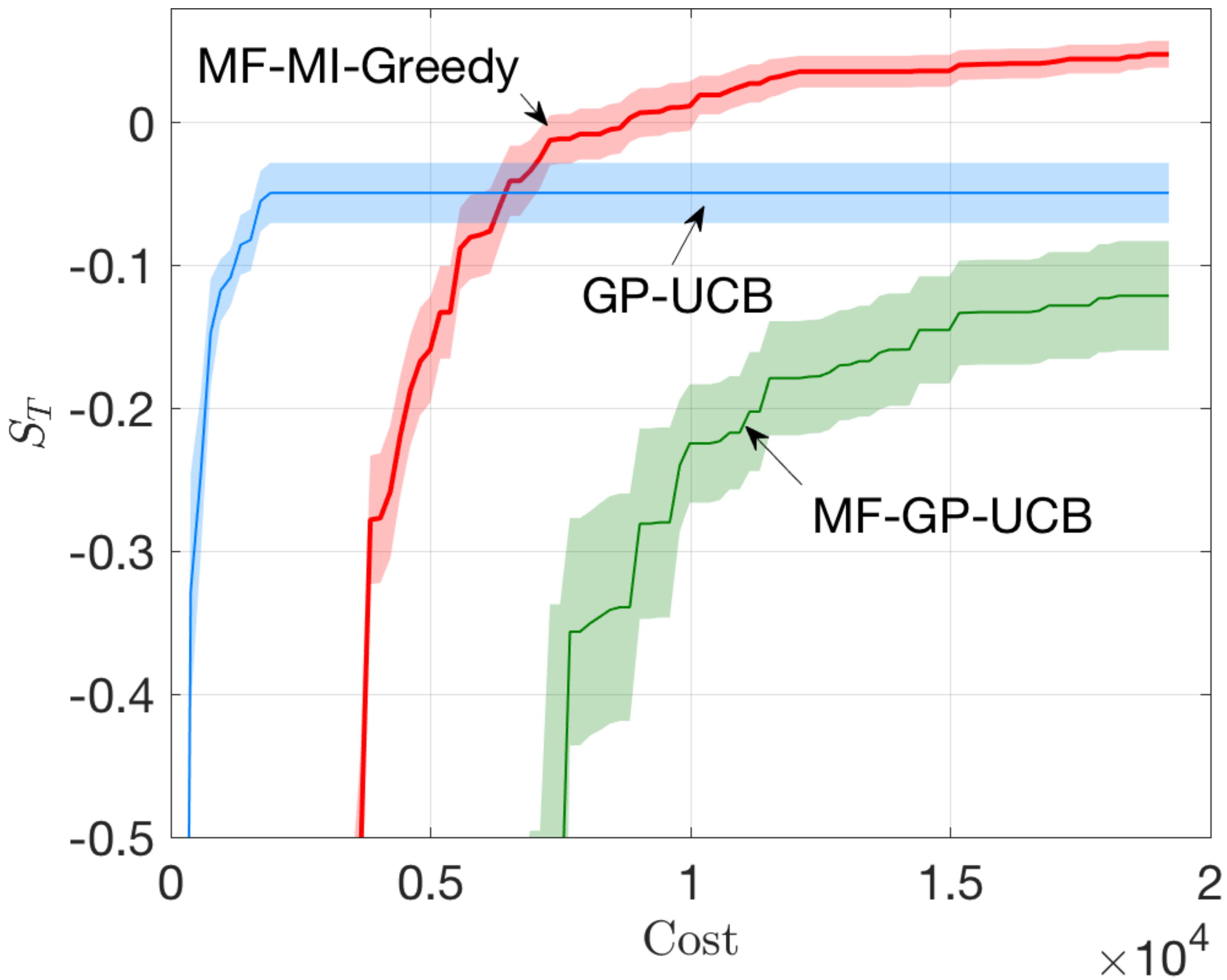
Optimizing Figure of Merit (FOM) with 3 fidelities

[Song, Tokpanov, Chen, Fleischman, Fountaine, Atwater, Yue, 2018]

Results: Cosmological Experimental Design

Max. Likelihood Inference on three cosmological parameters:

(1) the Hubble constant
(2) the dark matter fraction
(3) the dark energy fraction



ours

single-fidelity baseline

multi-fidelity baseline

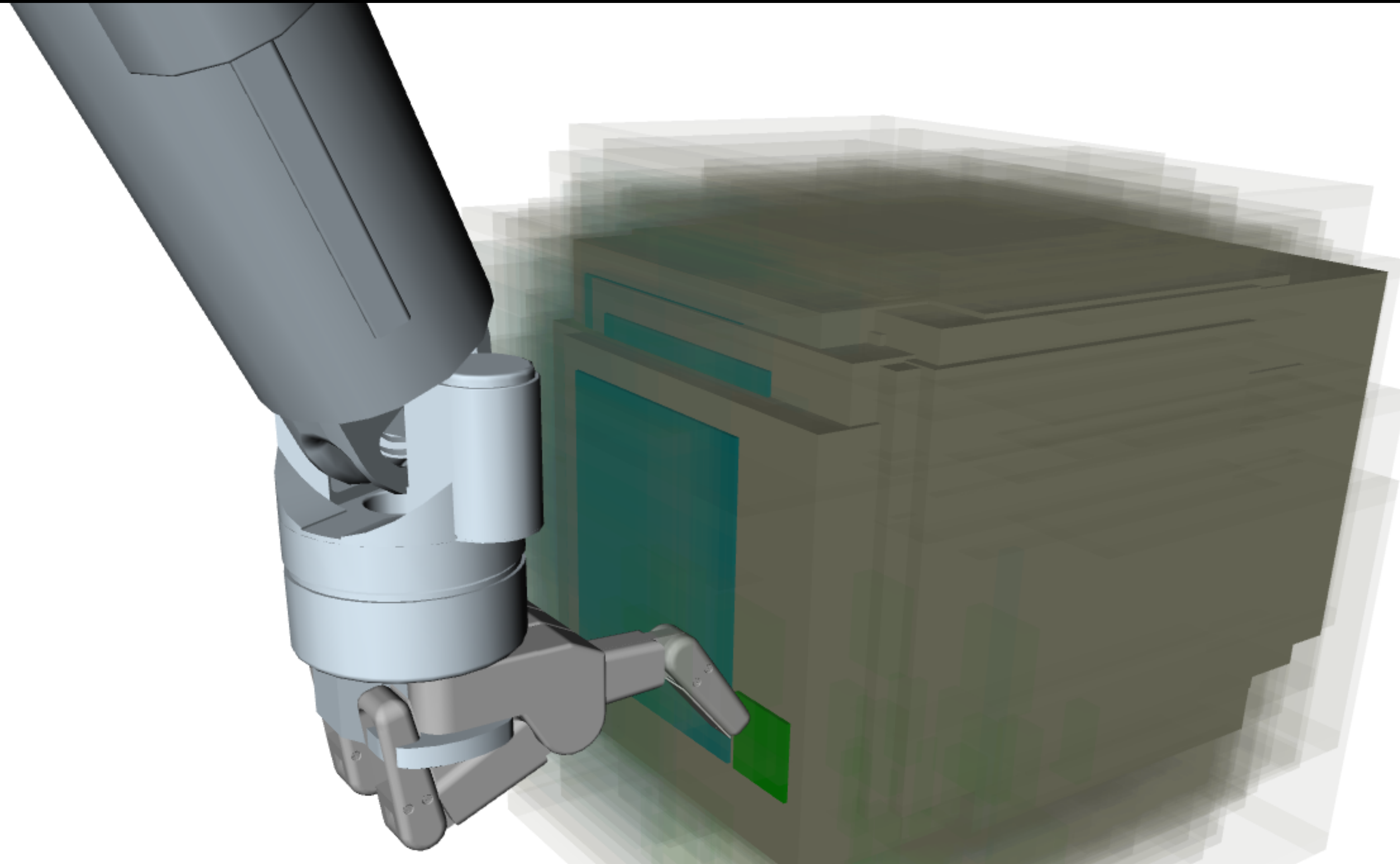
A General Framework for Multi-fidelity Bayesian Optimization with Gaussian Processes, Jialin Song et al., AISTATS 2019

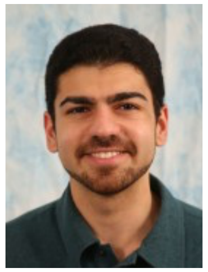
Automated Cosmic Experimental Design (ongoing)

Brian
Nord

https://news.fnal.gov/2019/09/survey-delivers-on-dark-energy-with-multiple-probes/

Aside: Bayesian Active Learning for Decision Making





Shervin Javdani

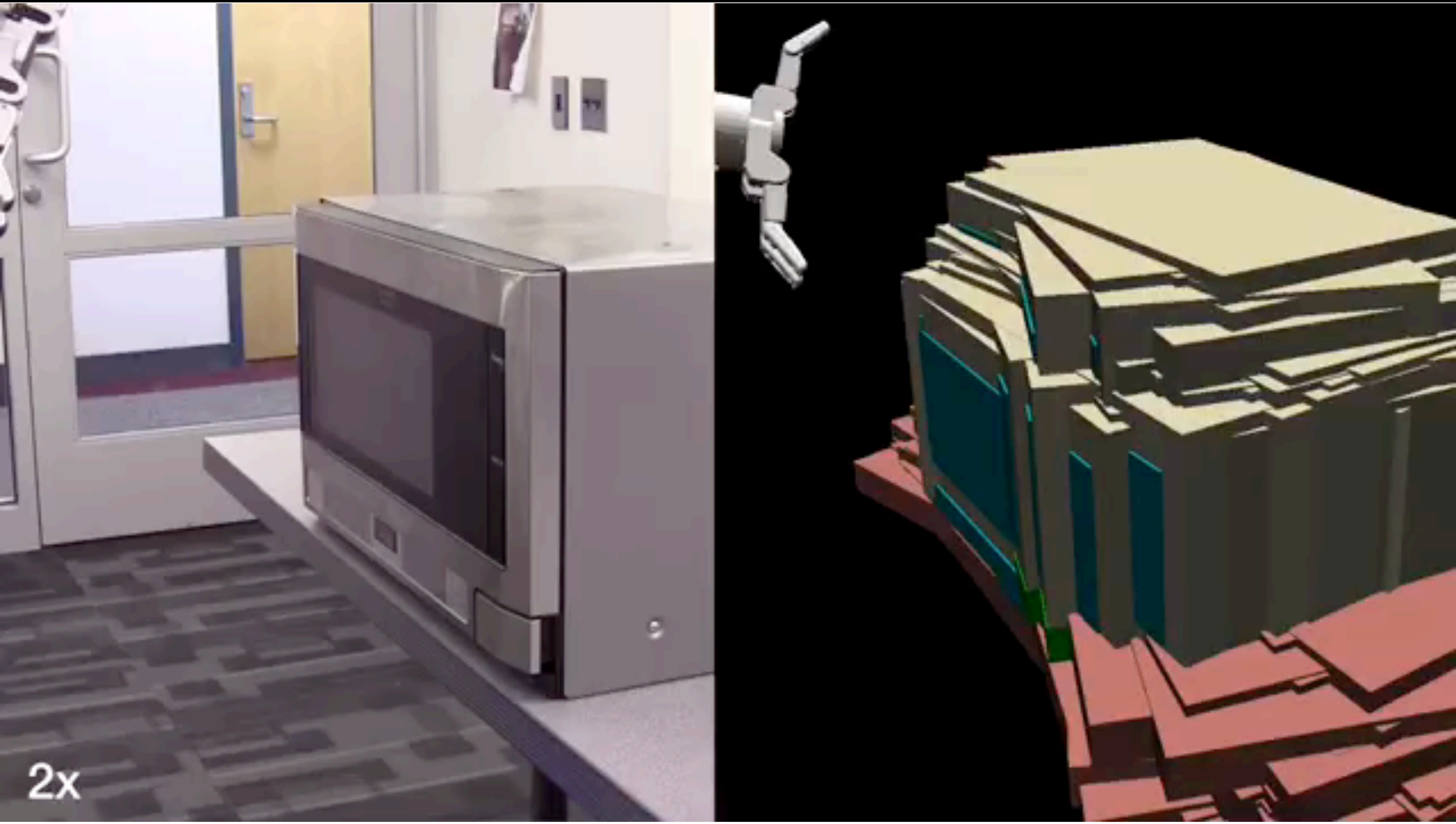
Andreas Krause

Drew Bagnell

Siddhartha Srinivasa

AISTATS'14
AAAI'15

Aside: Bayesian Active Learning for Decision Making



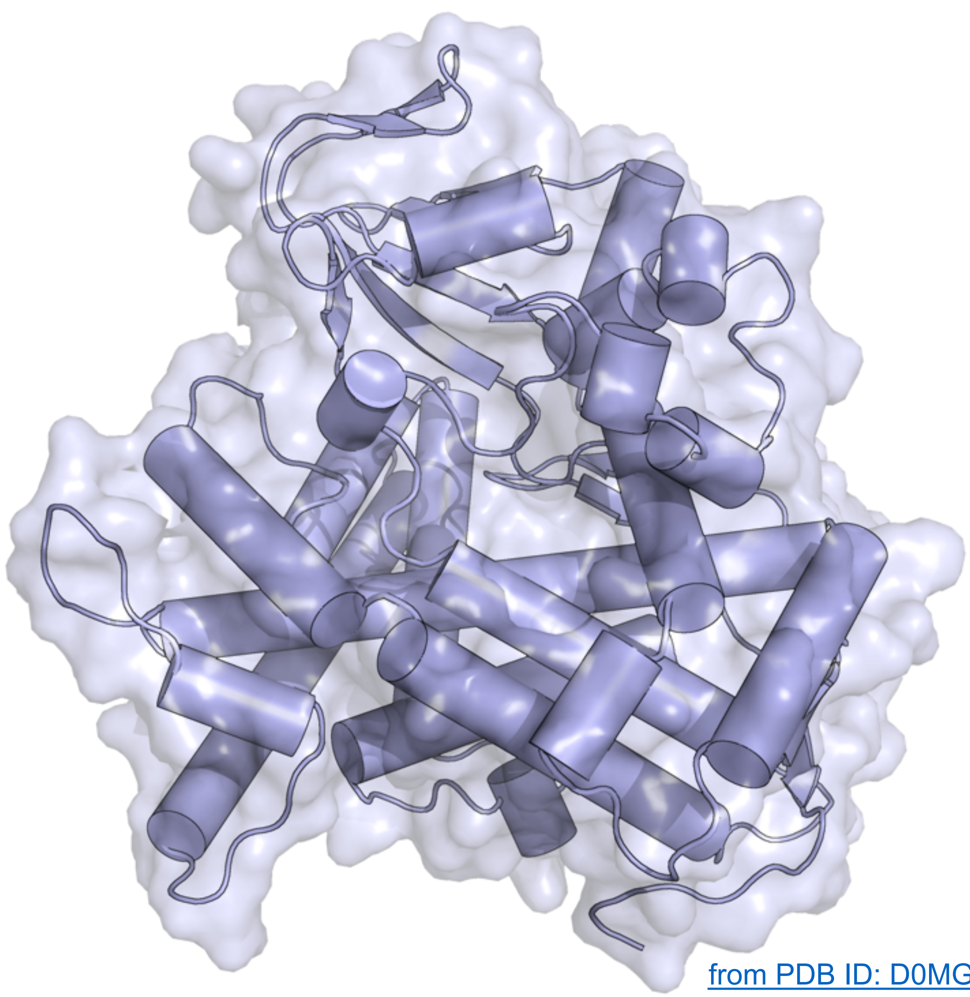
Batched Stochastic Bayesian Optimization via Combinatorial Constraints Design

K. Yang, Y. Chen, A. Lee, Y. Yue, AISTATS'19

How to deal with combinatorial action space?

Batched Stochastic Bayesian Optimization via Combinatorial Constraints Design
K. Yang, Y. Chen, A. Lee, Y. Yue, AISTATS'19

Protein Engineering



from PDB ID: D0MGT2

Image Credit: Zach Wu
(Arnold Group at Caltech)

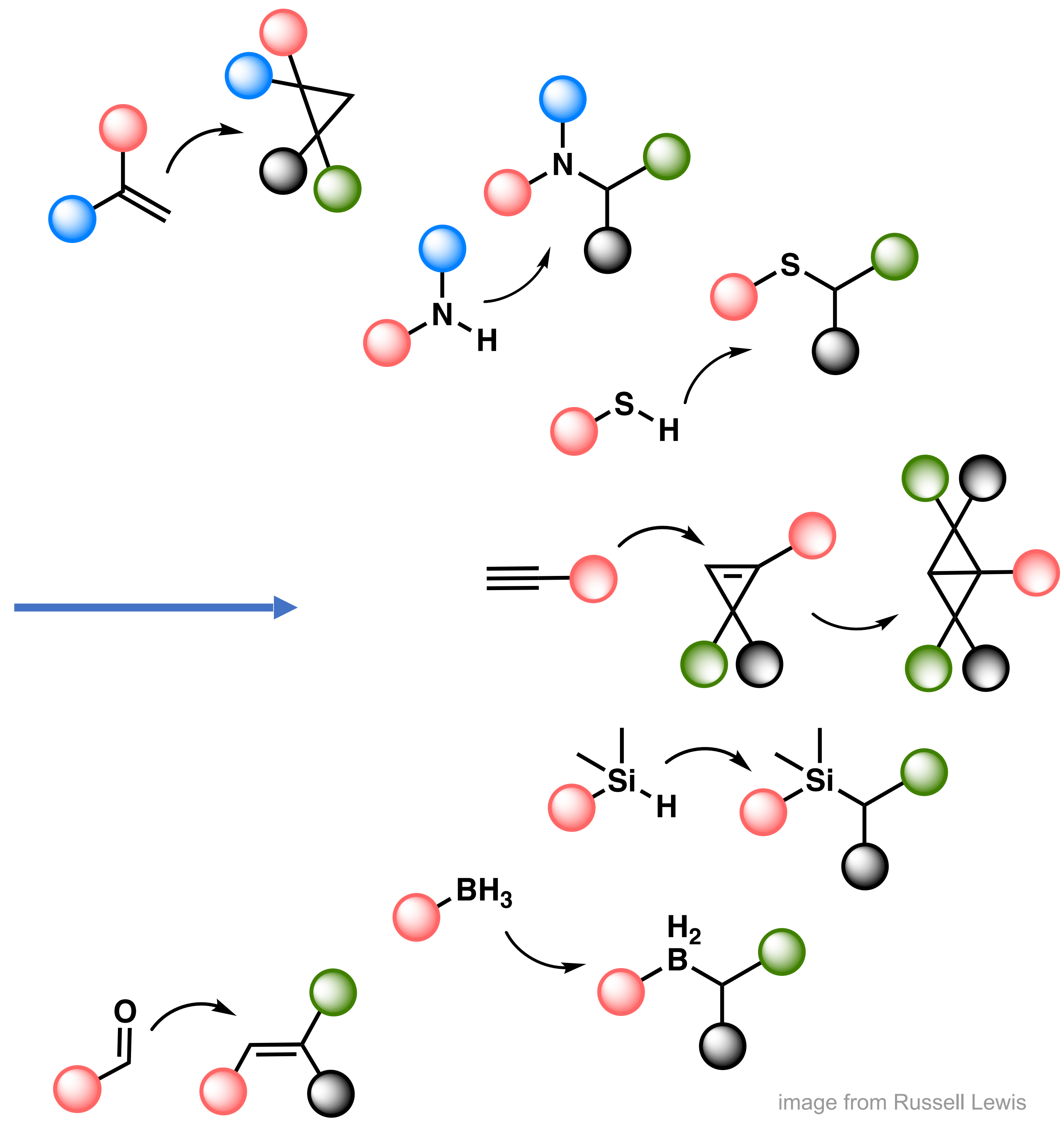
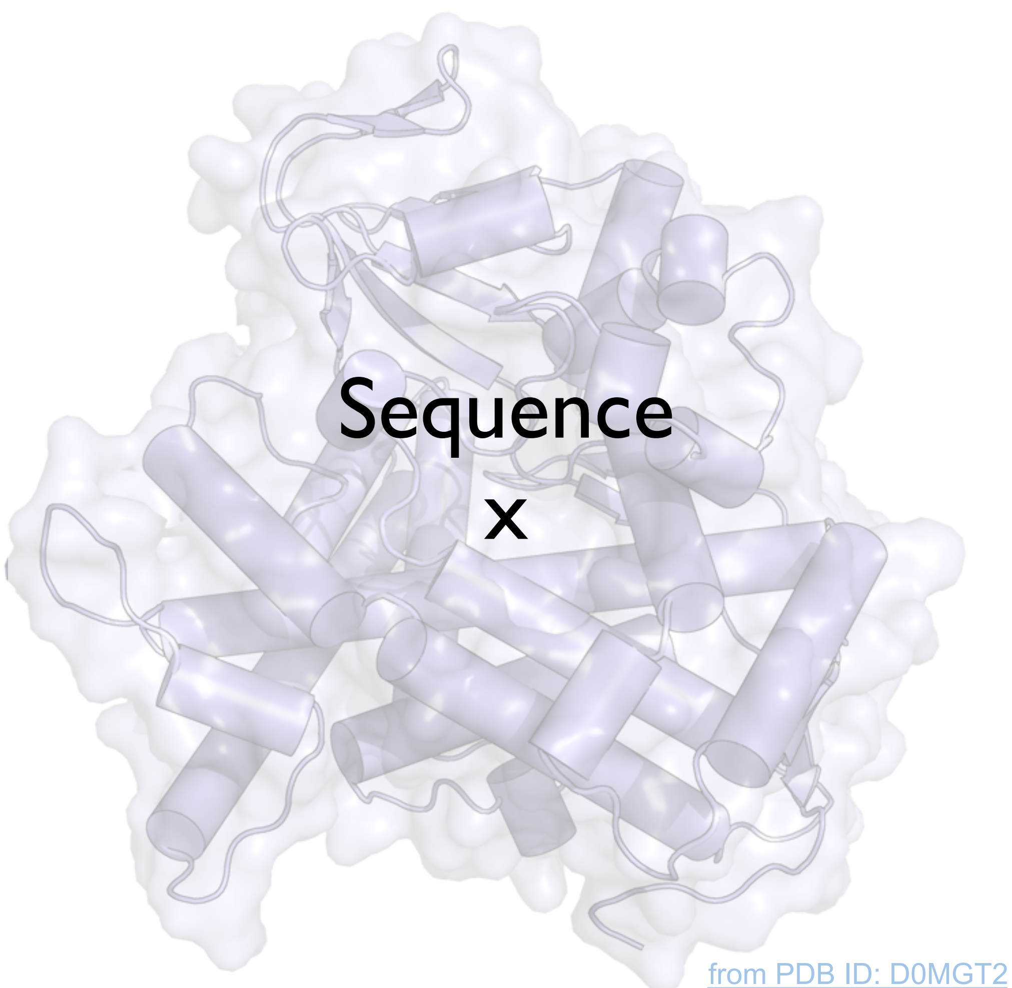


image from Russell Lewis

Protein Engineering



Sequence
x

Encodes

Function
f(x)

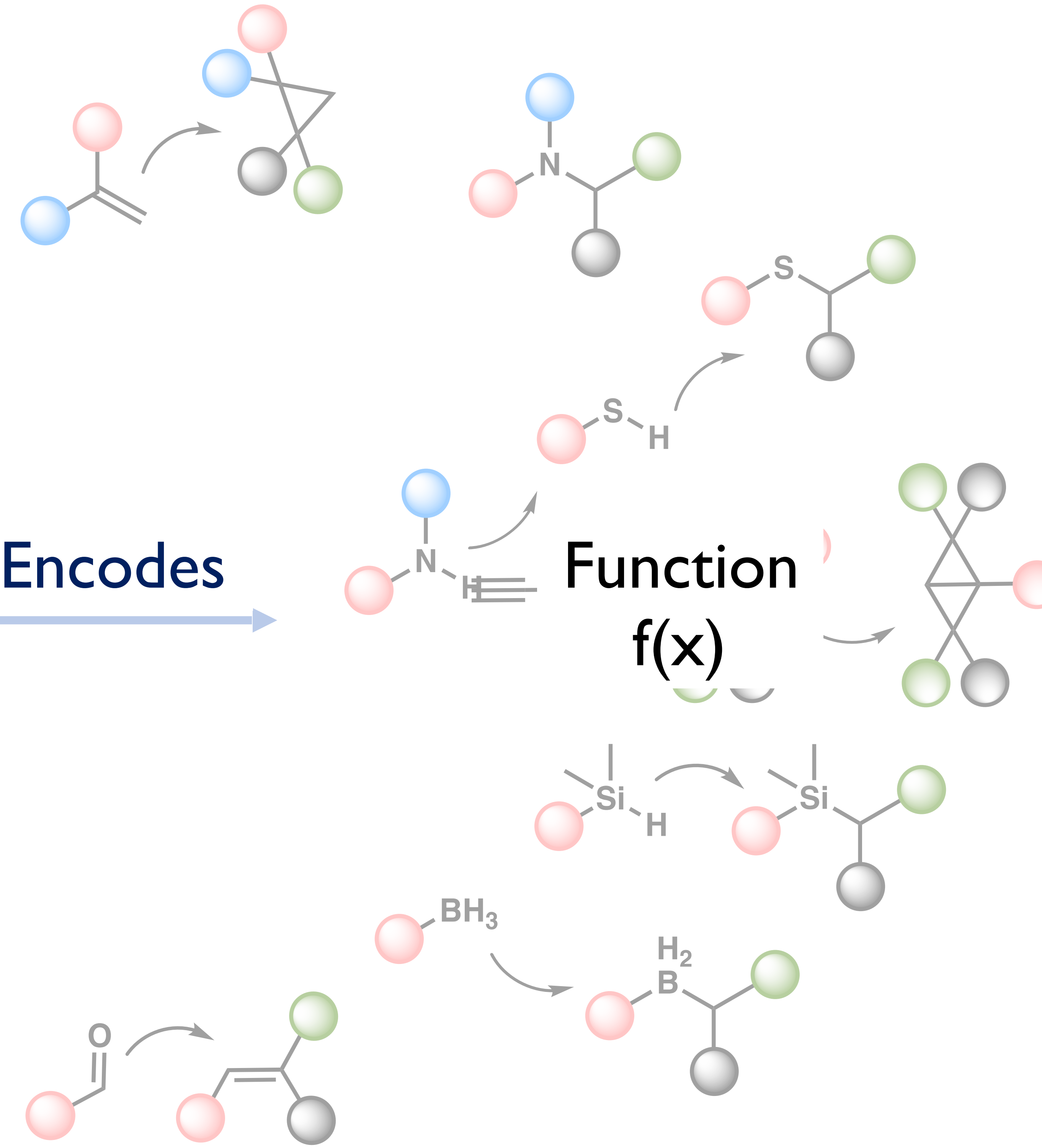
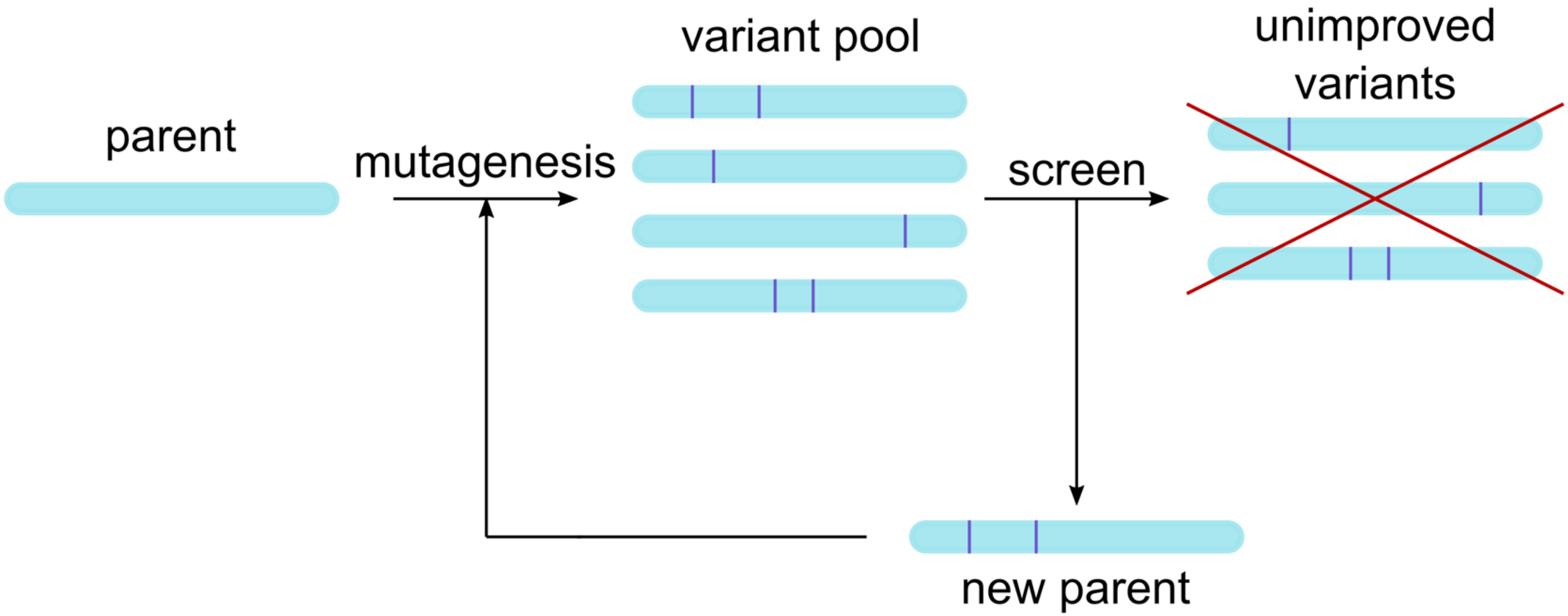


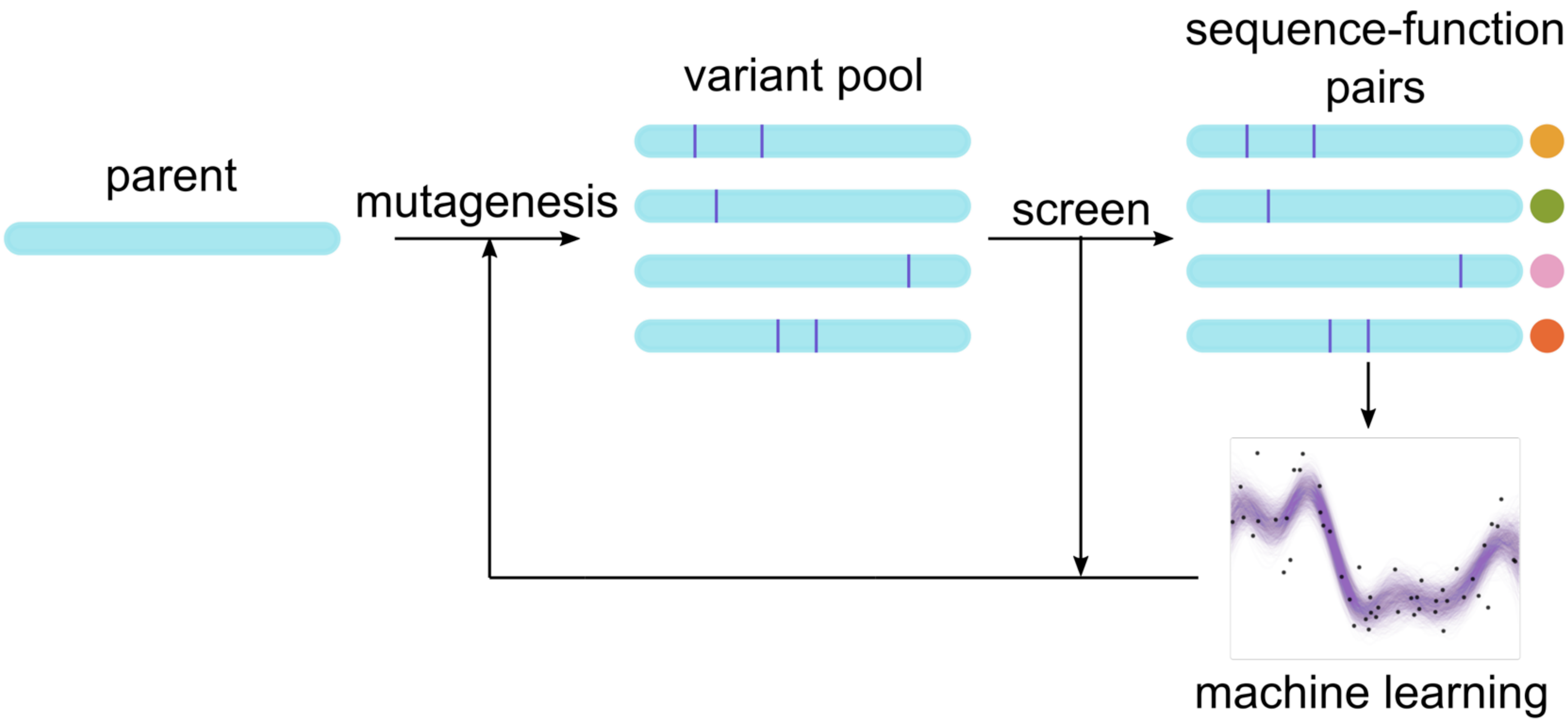
Image Credit: Zach Wu
(Arnold Group at Caltech)

Directed Evolution



Yang, Wu, Arnold (2018) *bioRxiv*

Directed Evolution + AI



[Yang, Wu, Arnold (2018) *bioRxiv*]

Batched Stochastic Bayesian Optimization

[Yang, Chen, Lee, Yue. AISTATS'19]

Kevin Yang

Yisong Yue

Frances Arnold

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

Batched Stochastic Bayesian Optimization

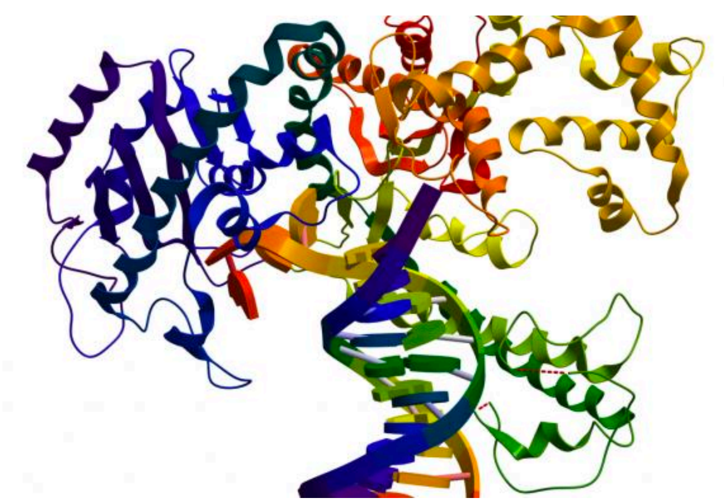
[Yang, Chen, Lee, Yue. AISTATS'19]


Kevin Yang


Yisong Yue


Frances Arnold

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



Batched Stochastic Bayesian Optimization

[Yang, Chen, Lee, Yue. AISTATS'19]

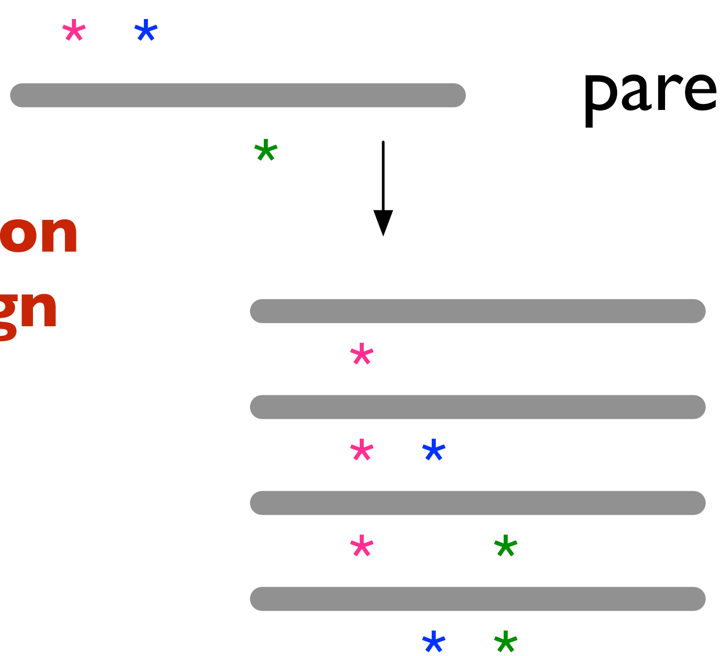
Kevin Yang

Yisong Yue

Frances Arnold

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

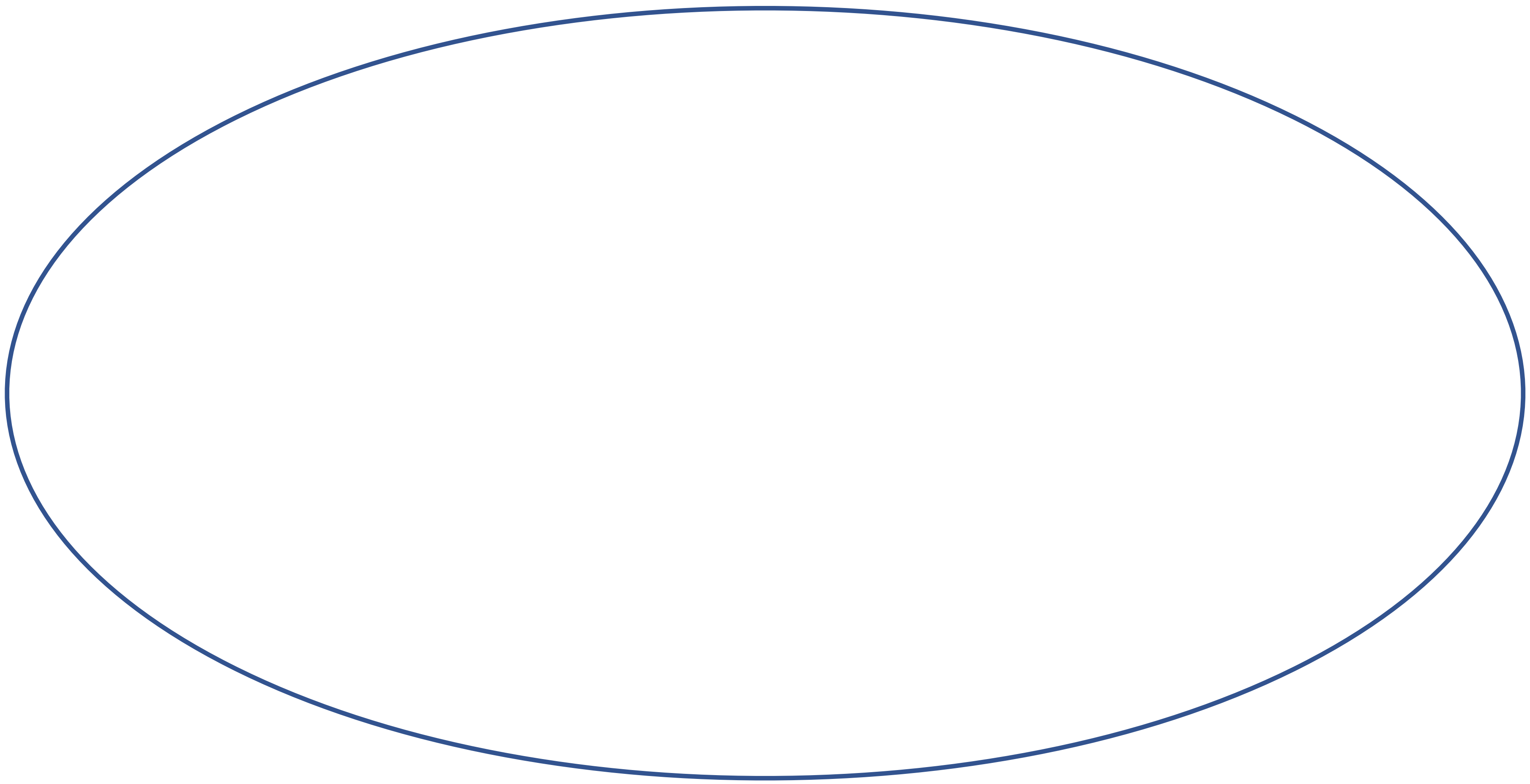
site-saturation library design



parent seq

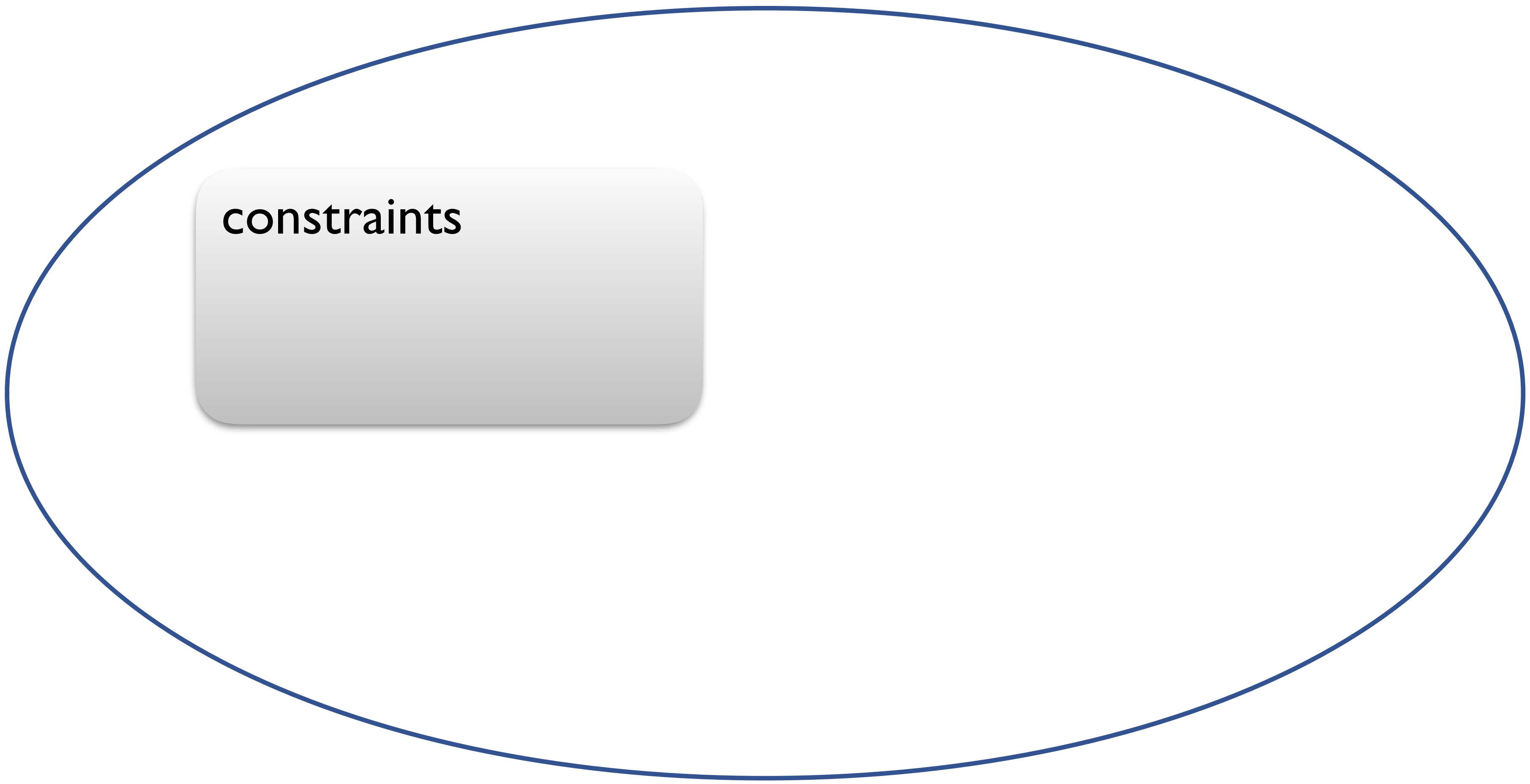
site-saturation library

Algorithmic Insights



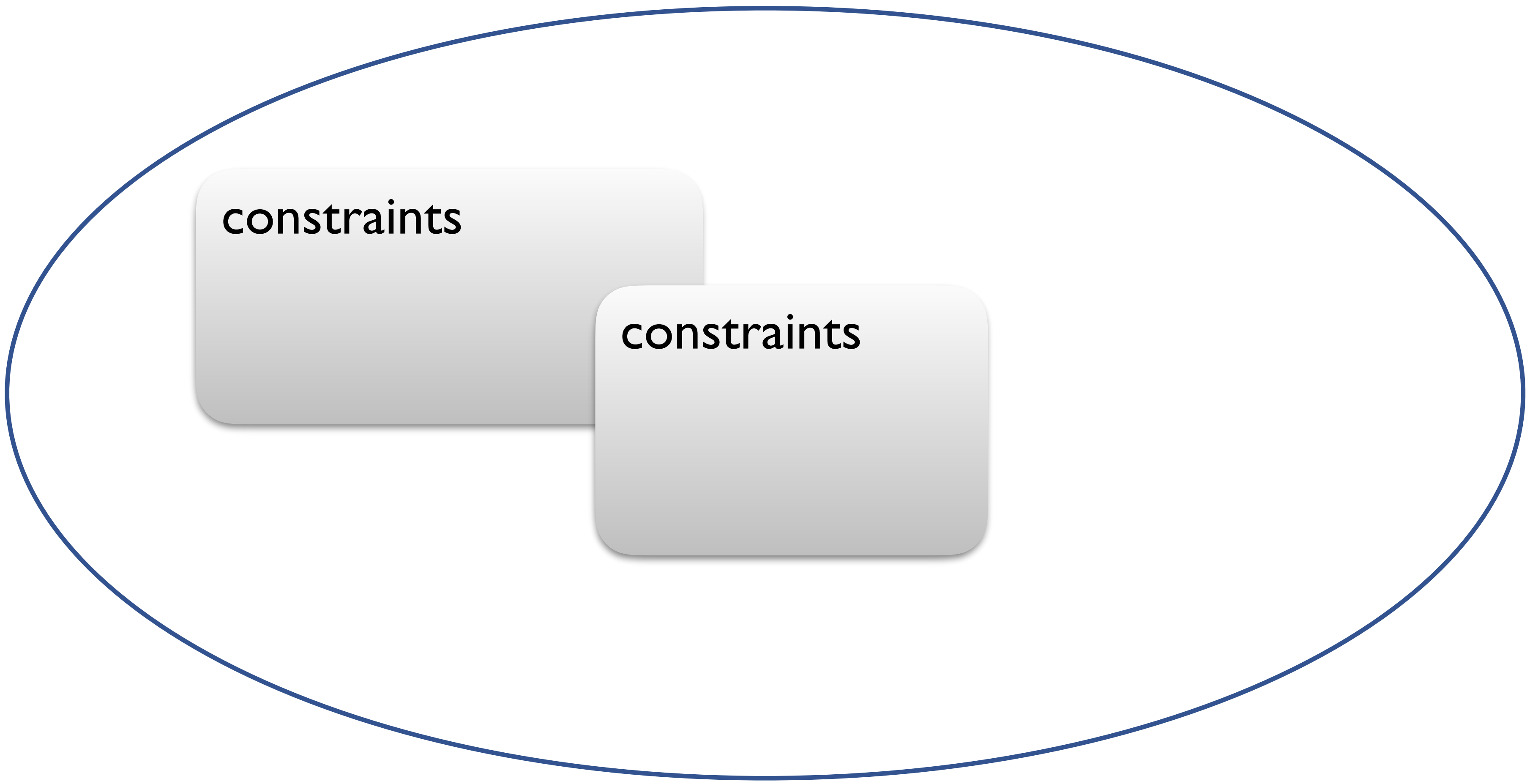
Batched Stochastic Bayesian Optimization via Combinatorial Constraints Design Kevin Yang et al, AISTATS'19

Algorithmic Insights



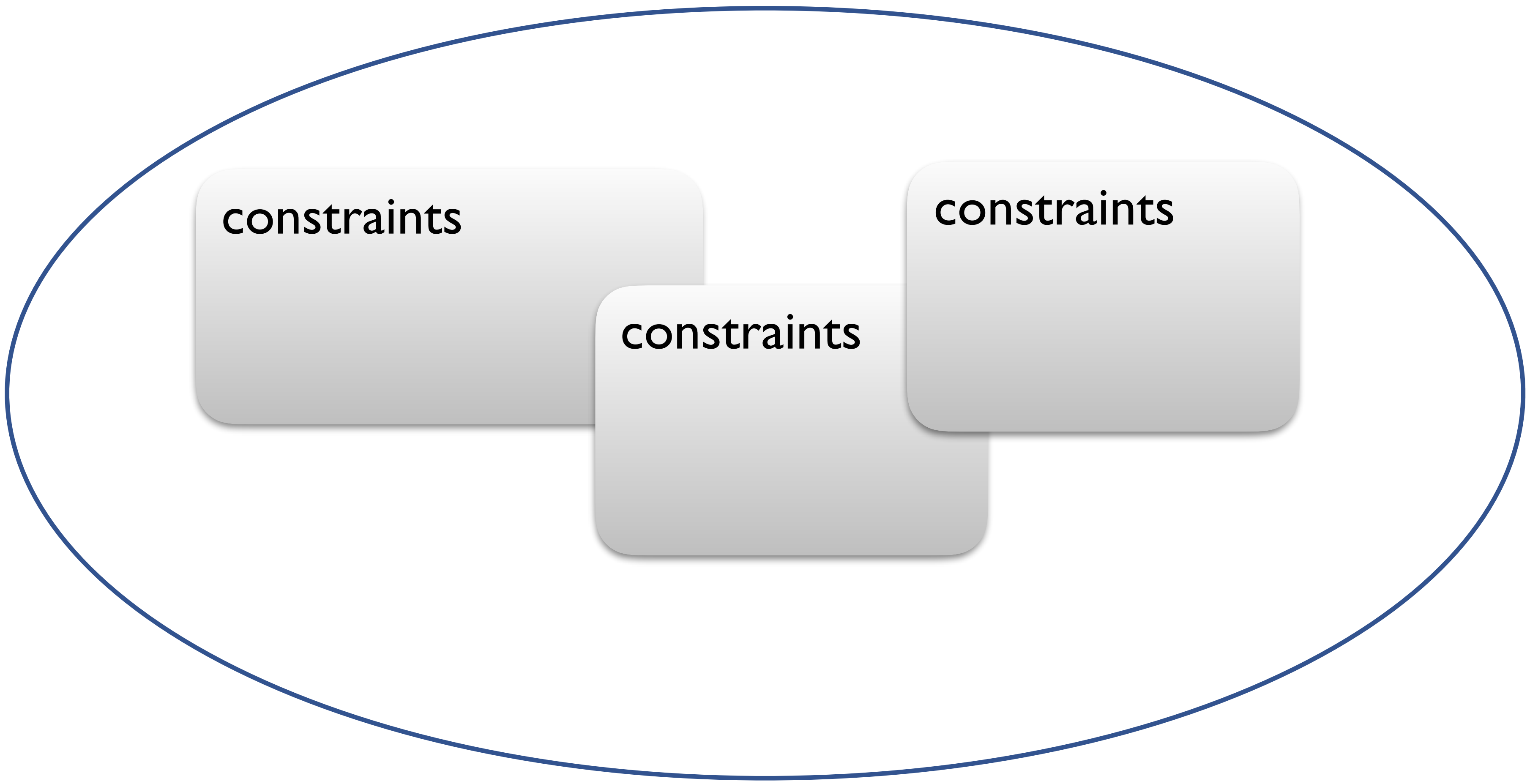
Batched Stochastic Bayesian Optimization via Combinatorial Constraints Design Kevin Yang et al, AISTATS'19

Algorithmic Insights



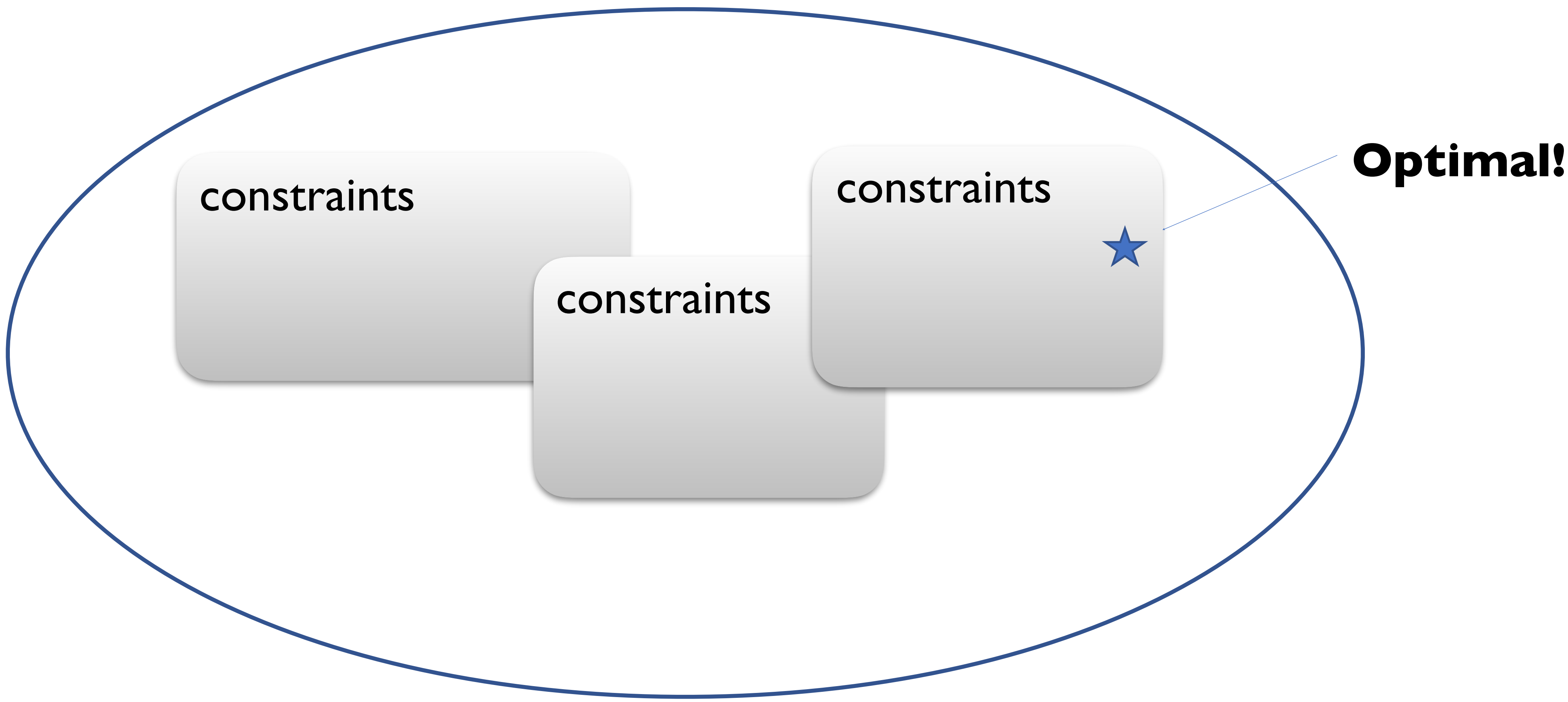
Batched Stochastic Bayesian Optimization via Combinatorial Constraints Design Kevin Yang et al, AISTATS'19

Algorithmic Insights



Batched Stochastic Bayesian Optimization via Combinatorial Constraints Design Kevin Yang et al, AISTATS'19

Algorithmic Insights



Batched Stochastic Bayesian Optimization via Combinatorial Constraints Design Kevin Yang et al, AISTATS'19

Algorithmic Insights

Maximize Expected Improvement

- Define utility as the # of "good" sequences
- Solve a combinatorial optimization problem
- Represent the set function as the difference of two submodular functions (DS)

Provably convergence guarantees

Optimal!

Batched Stochastic Bayesian Optimization via Combinatorial Constraints Design Kevin Yang et al, AISTATS'19

Batched Stochastic Bayesian Optimization

[Yang, Chen, Lee, Yue. AISTATS'19]

Online Batched Constraints Design vs DS Optimization (Online-DSOpt)

Input Set of constraints *C*; budget on each batch *n*; joint (GP) prior.

Output Optimizer of the target function

Start $A \leftarrow \varnothing$

Loop $M \leftarrow$ Computer expected reward for all candidate sequences

$S \leftarrow$ *DS-Opt(M, C, n)* */* Select locally optimal constraints set using **D**ifference-of-**S**ubmodular functions **Opt**imization */*

$A \leftarrow A \cup B(S)$

Until *budget exhausted*

Batched Stochastic Bayesian Optimization

[Yang, Chen, Lee, Yue. AISTATS'19]

Online Batched Constraints Design vs DS Optimization (Online-DSOpt)

Input Set of constraints *C*; budget on each batch *n*; joint (GP) prior.

Output Optimizer of the target function

Start $A \leftarrow \varnothing$

Loop $M \leftarrow$ Computer expected reward for all candidate sequences

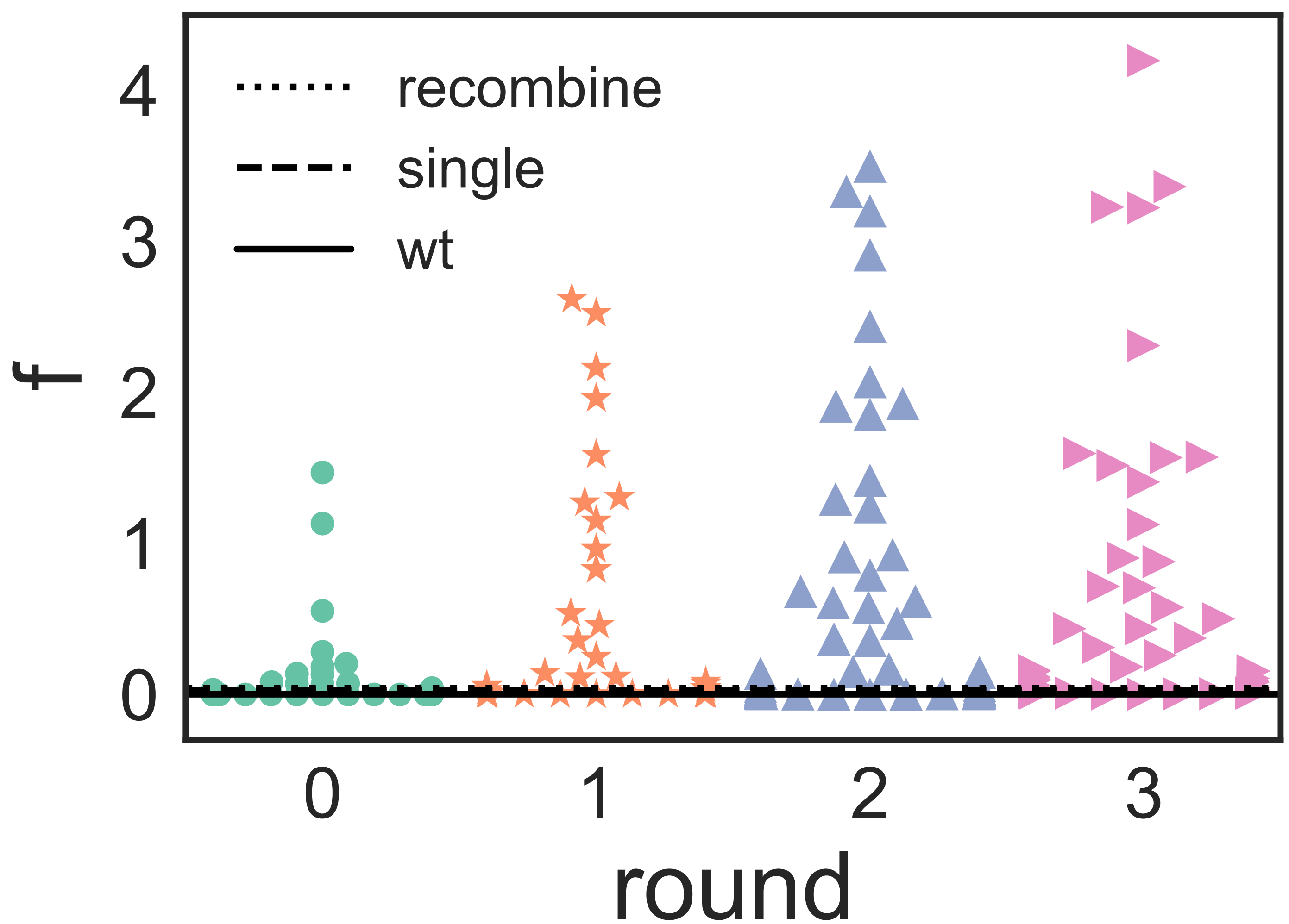
$S \leftarrow$ *DS-Opt(M, C, n)* */* Select locally optimal constraints set using **D**ifference-of-**S**ubmodular functions **Opt**imization */*

$A \leftarrow A \cup B(S)$

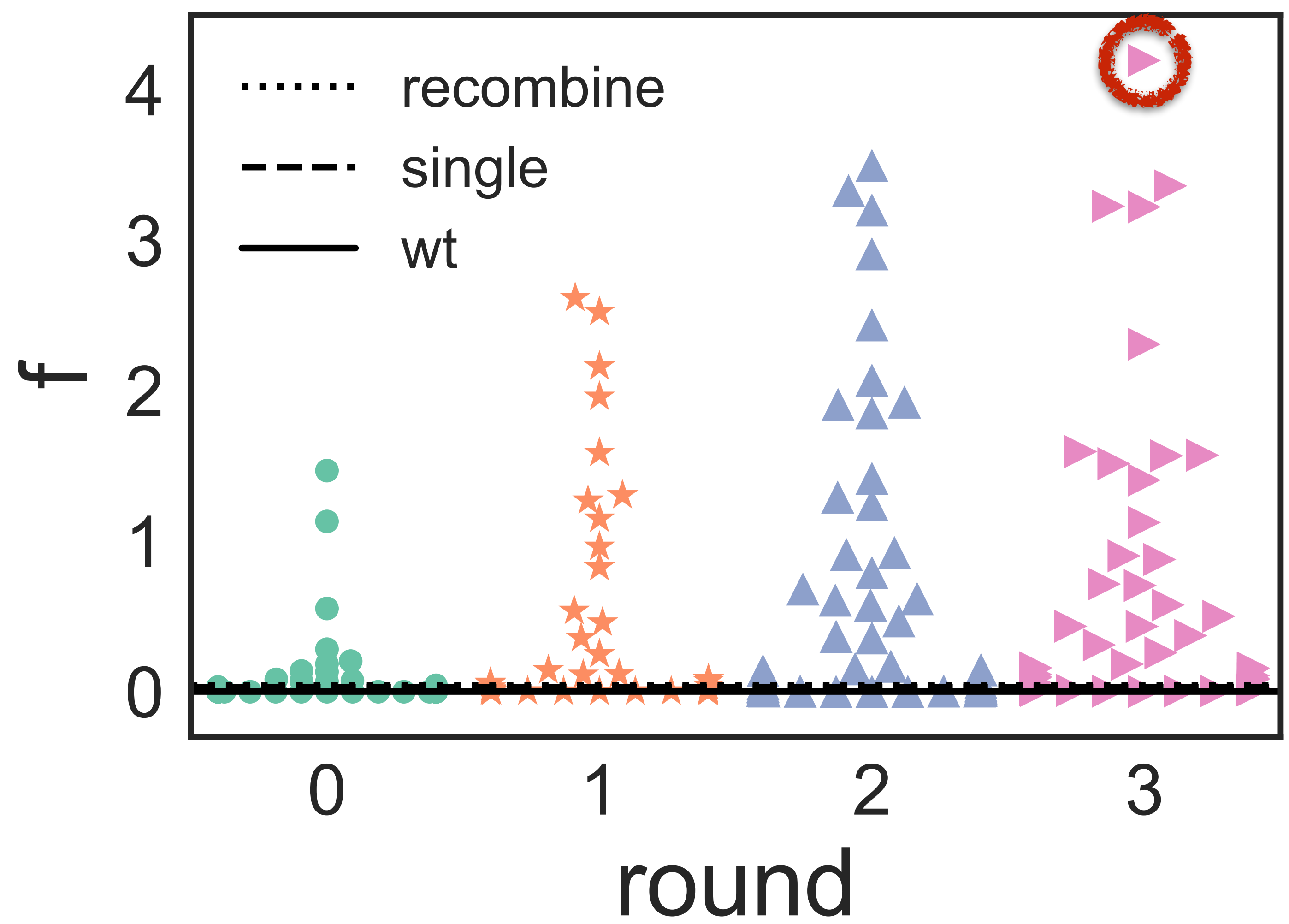
Until *budget exhausted*

Thm. DS-Opt converges to a local optima in polynomial time.

Results: Site-saturation mutagenesis



Results: Site-saturation mutagenesis



Aside: 3D Printer Calibration (ongoing)

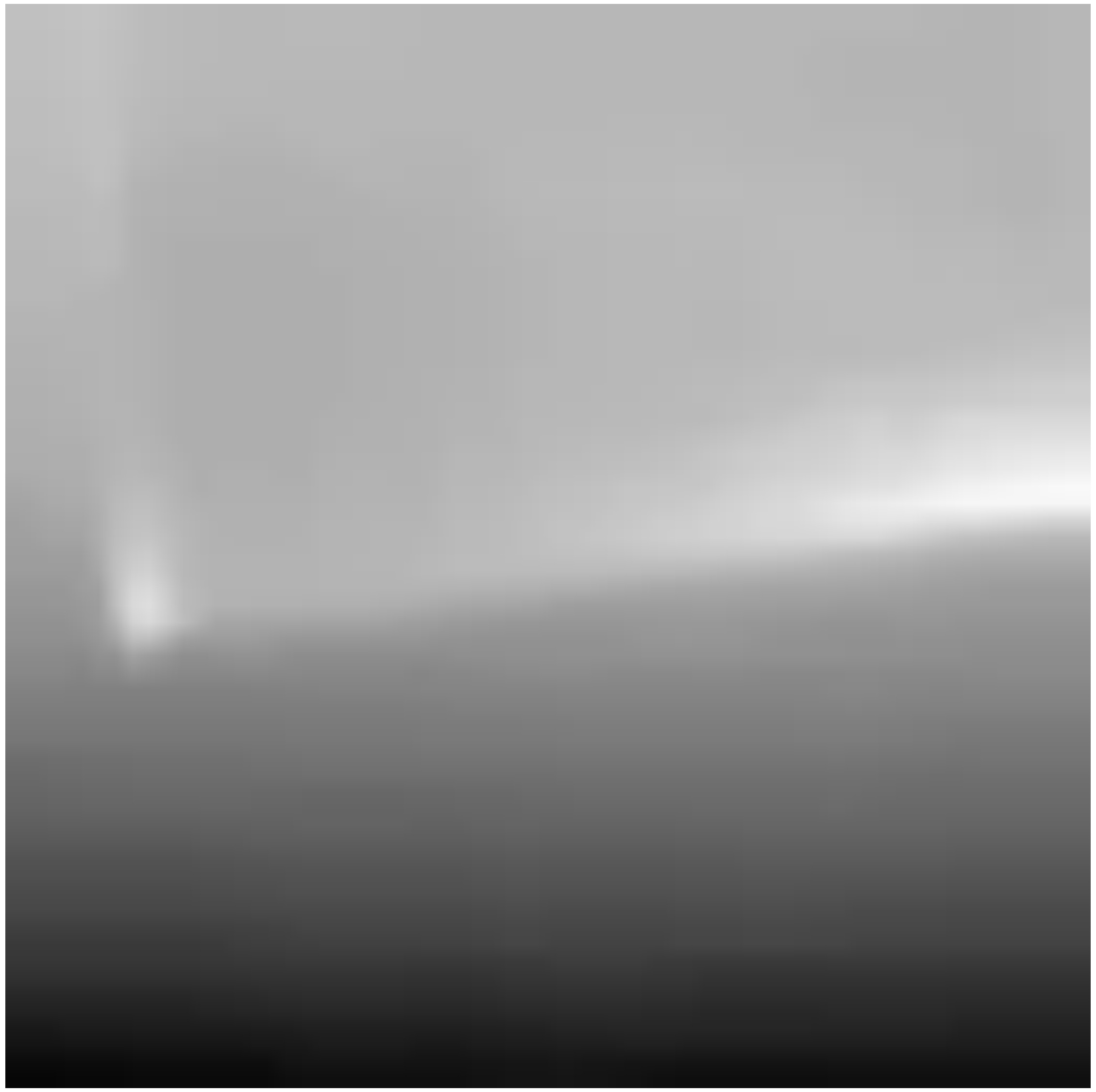
Yisong Yue

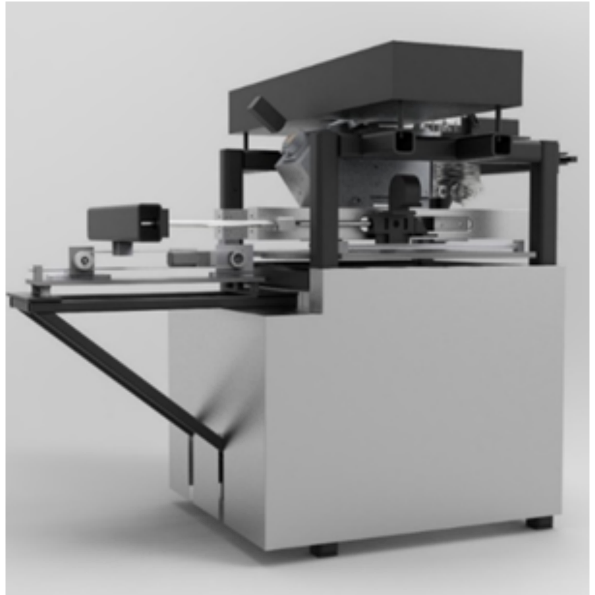
Ufuk Topcu

Alex Nettekoven

Steven Carr

Aside: 3D Printer Calibration (ongoing)





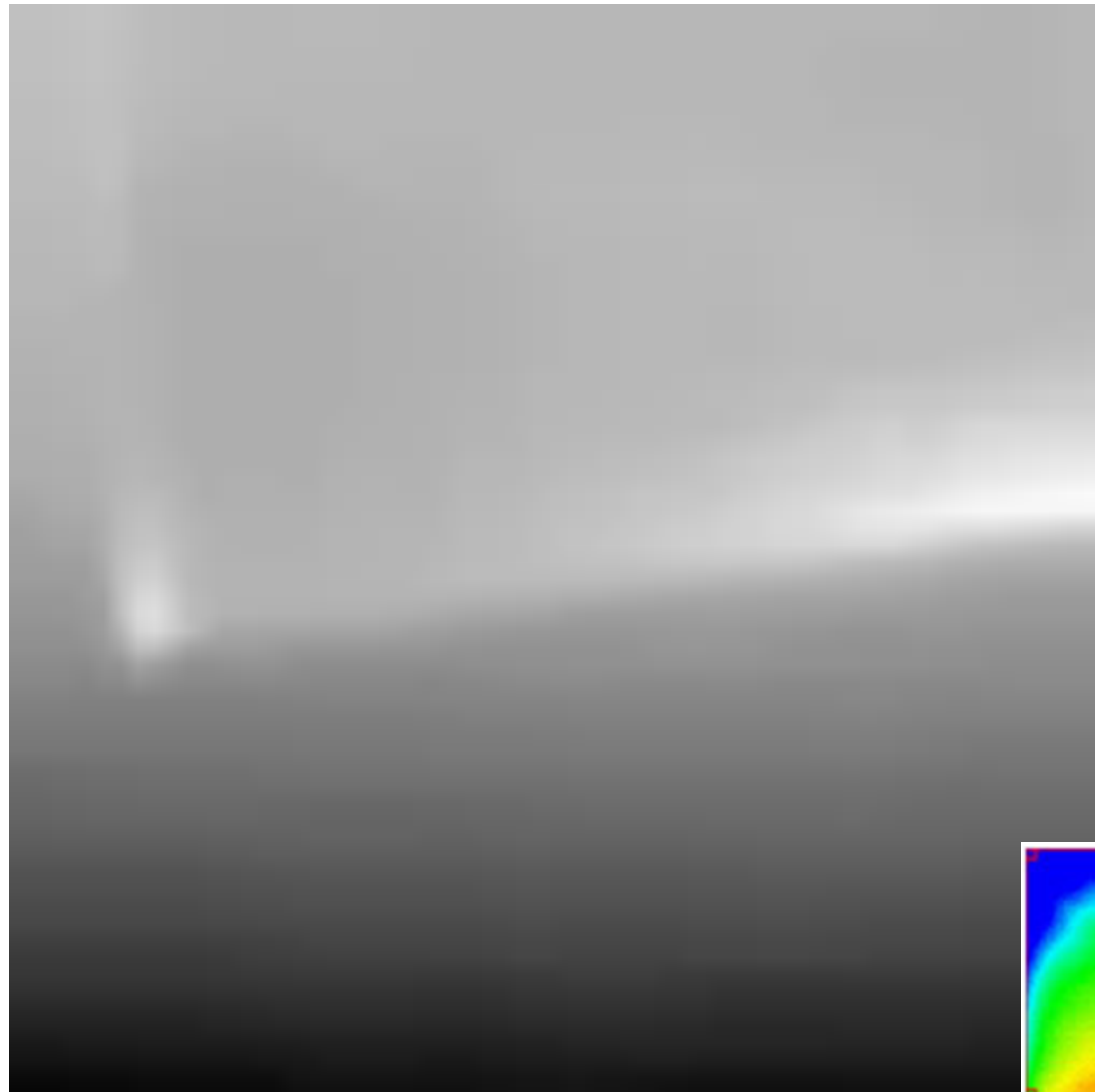
Yisong Yue

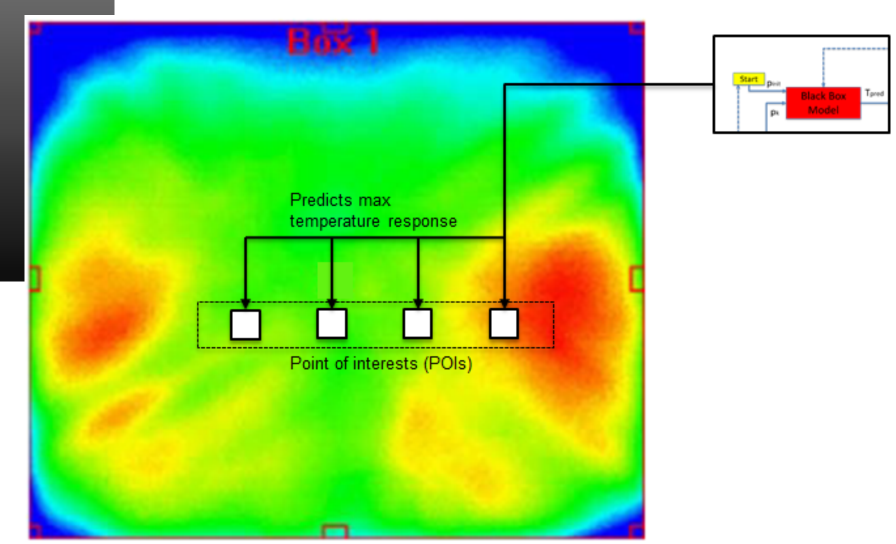
Ufuk Topcu

Alex Nettekoven

Steven Carr

Aside: 3D Printer Calibration (ongoing)





[image credit: Advanced manufacturing and design lab at UT Austin]

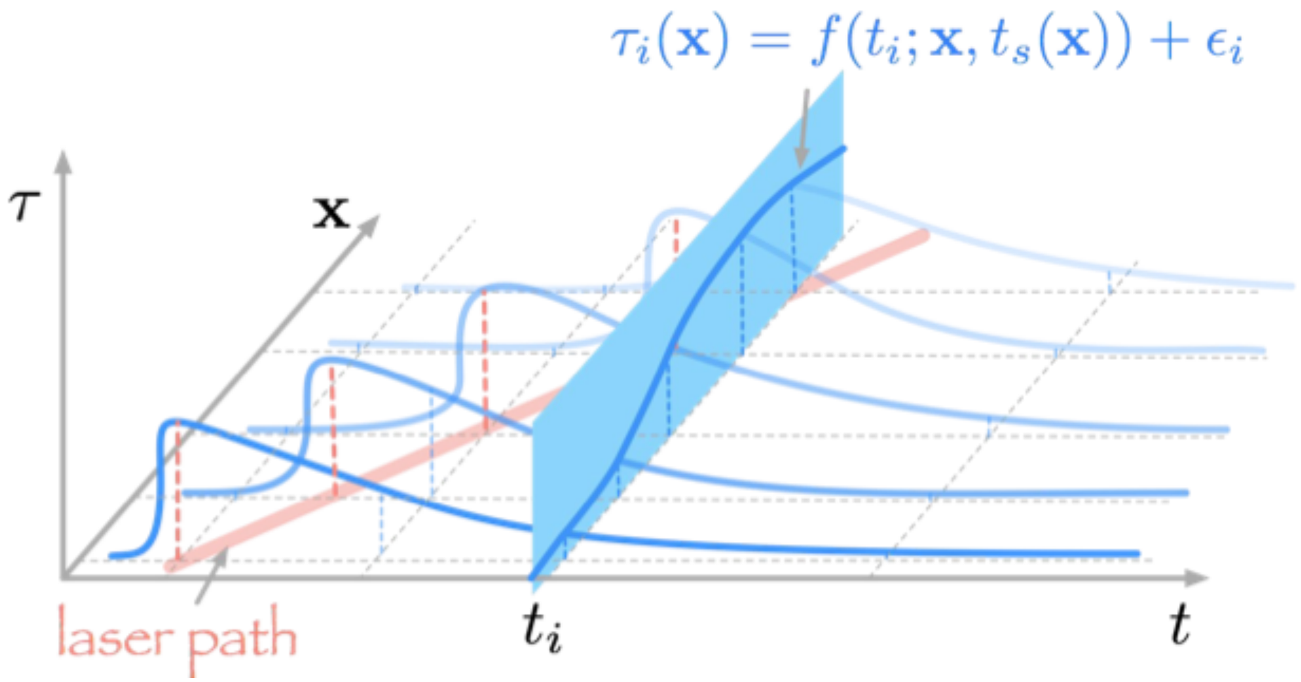

Yisong Yue

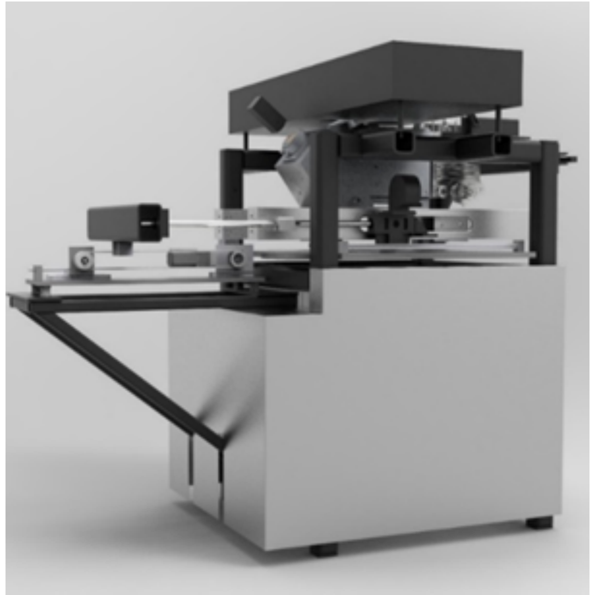

Ufuk Topcu

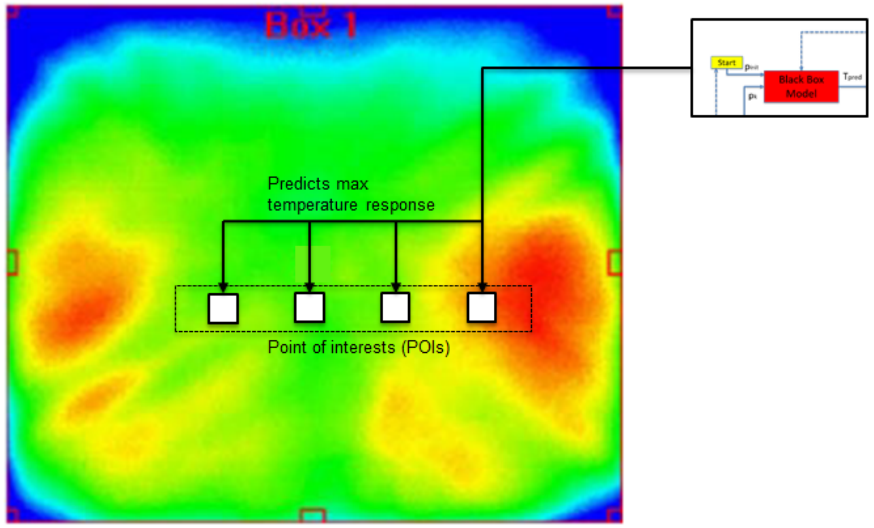

Alex Nettekoven


Steven Carr

Aside: 3D Printer Calibration (ongoing)







[image credit: Advanced manufacturing and design lab at UT Austin]

Yisong Yue

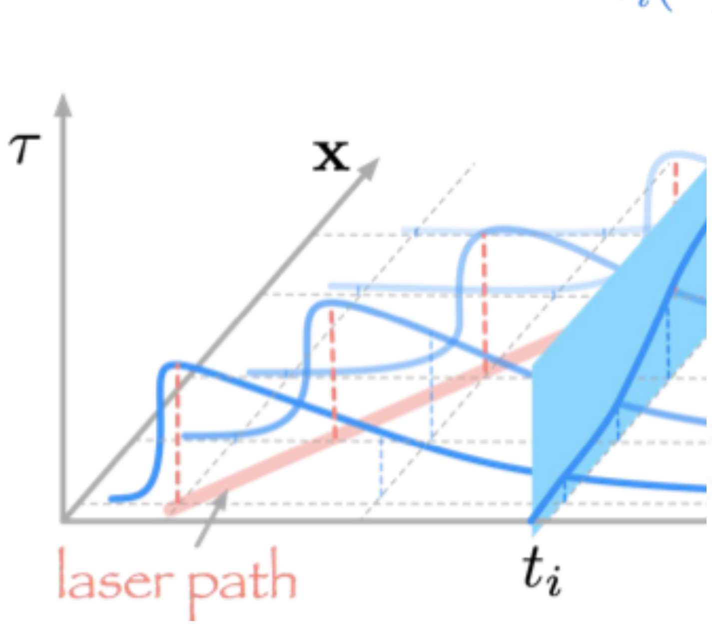
Ufuk Topcu

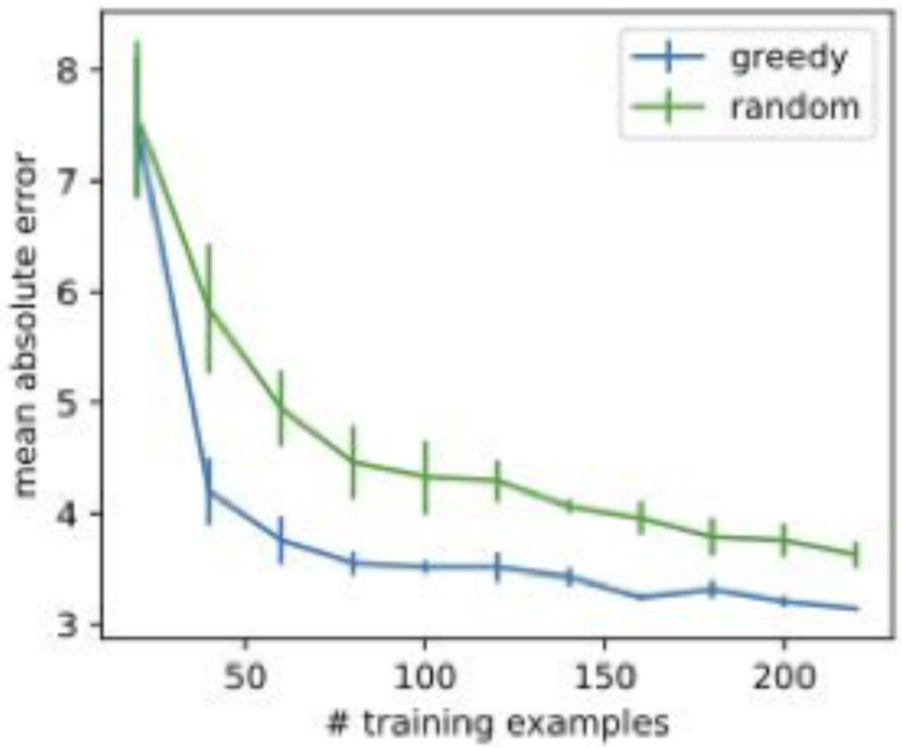
Alex Nettekoven

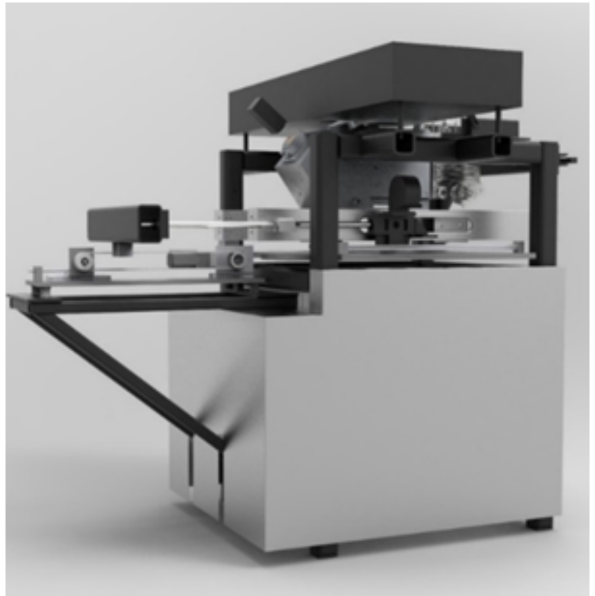
Steven Carr

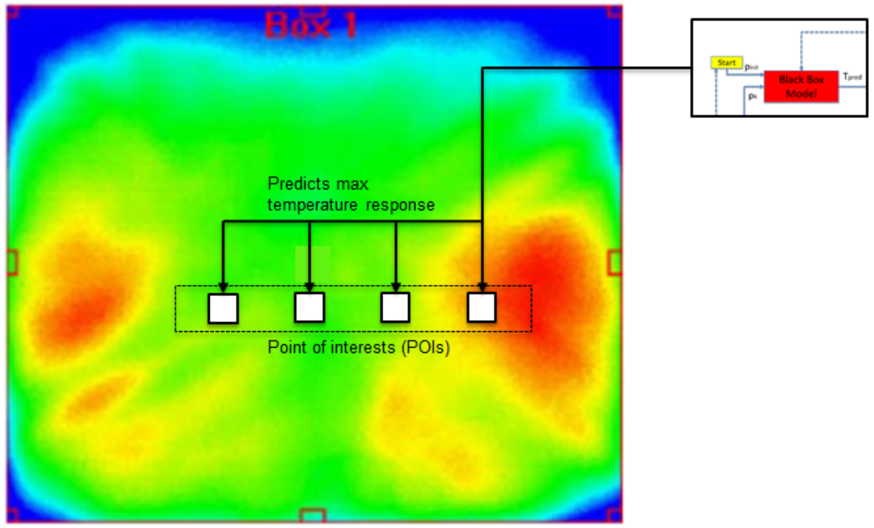
Aside: 3D Printer Calibration (ongoing)

$$\tau_i(\mathbf{x}) = f(t_i; \mathbf{x}, t_s(\mathbf{x})) + \epsilon_i$$









[image credit: Advanced manufacturing and design lab at UT Austin]


Yisong Yue


Ufuk Topcu


Alex Nettekoven

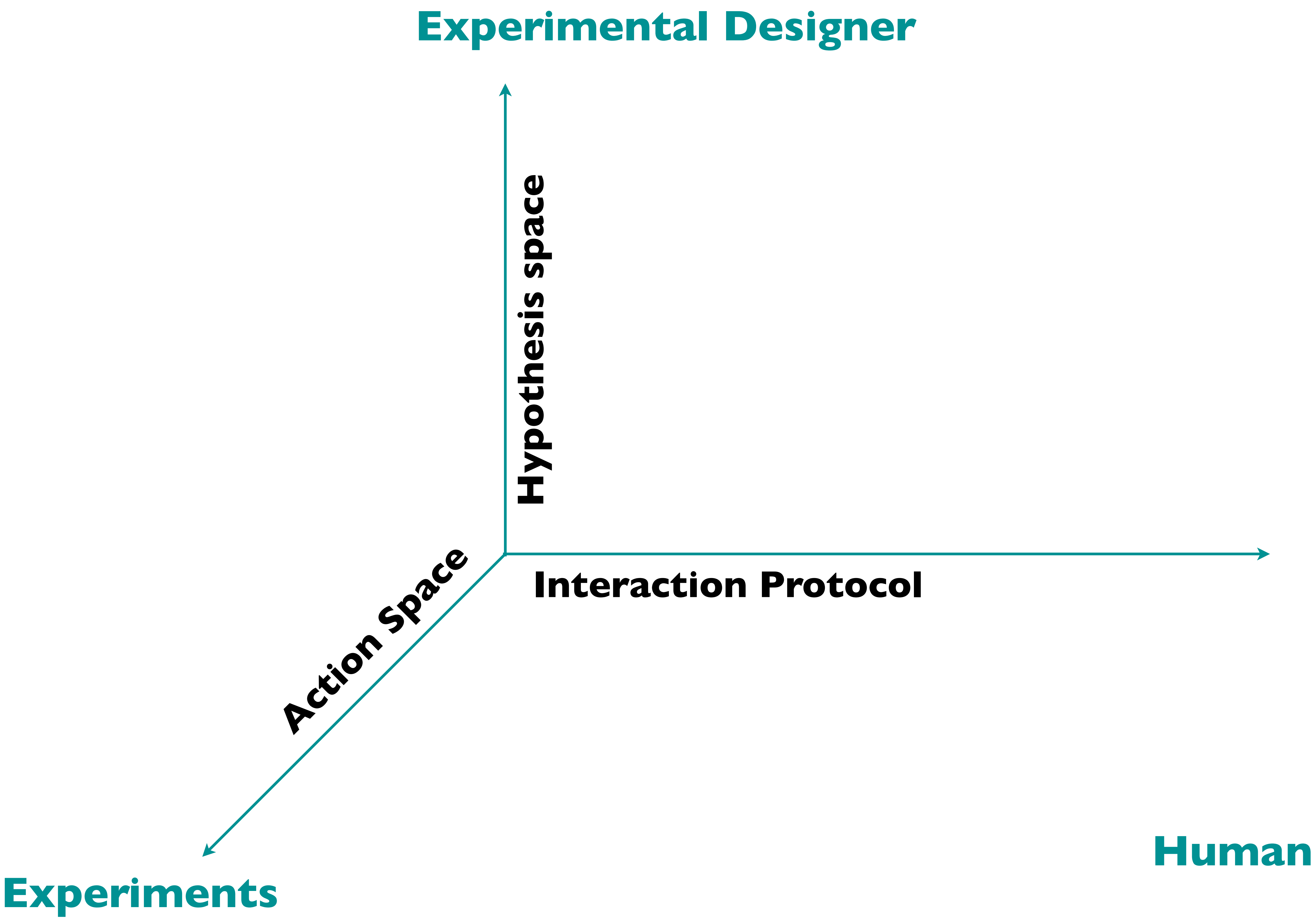

Steven Carr

Summary

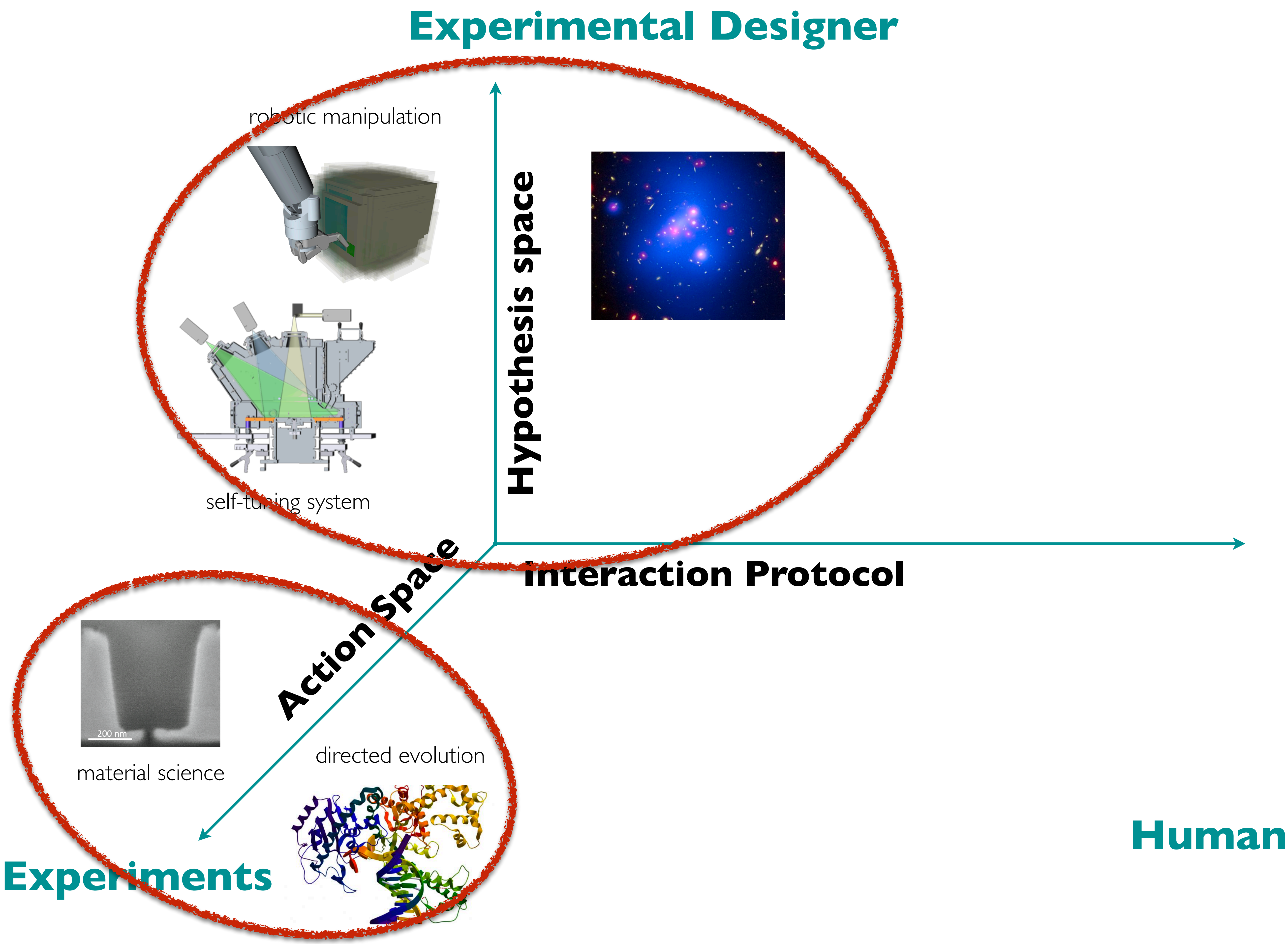
Experimental Designer

Human

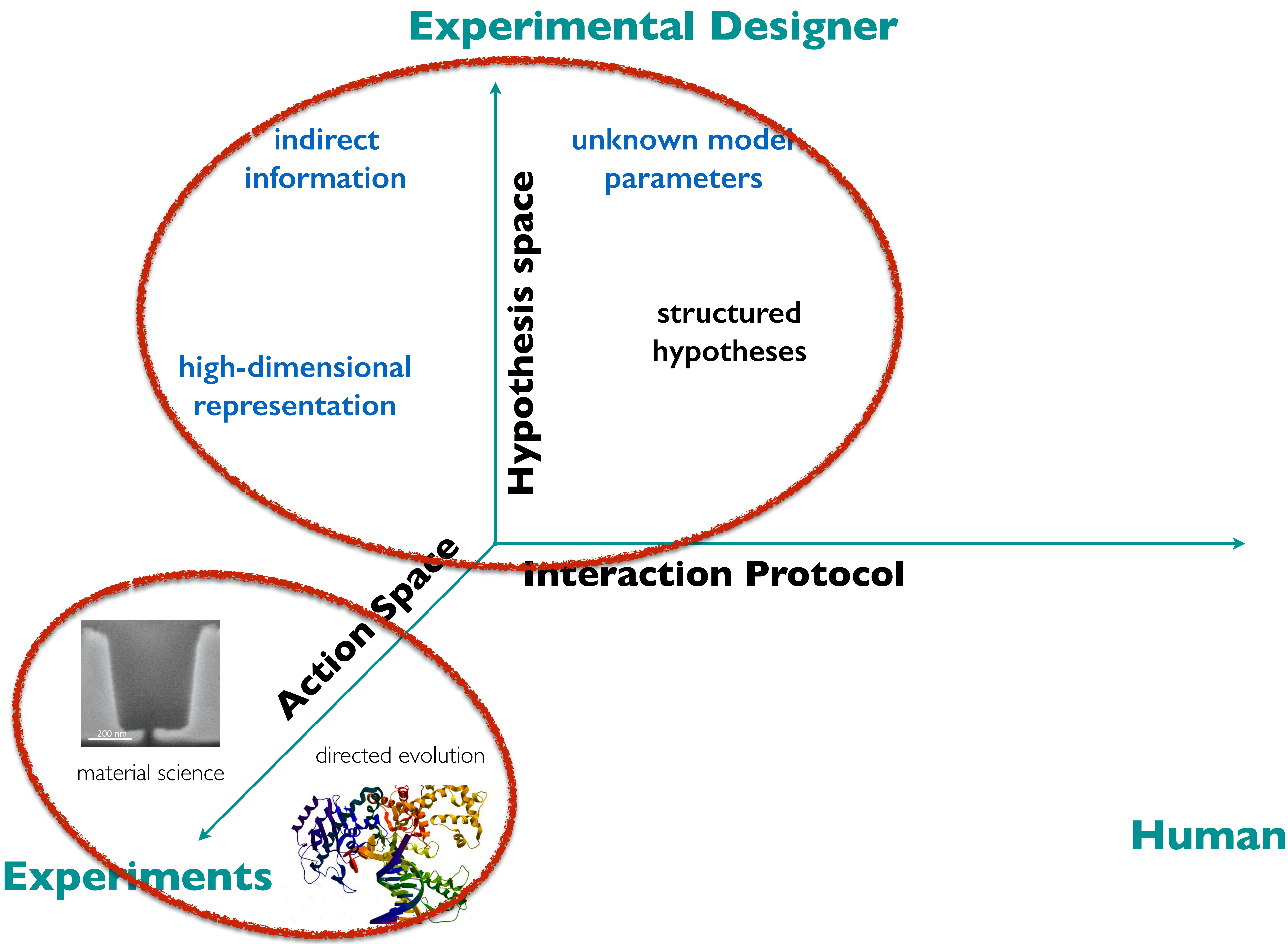
Experiments



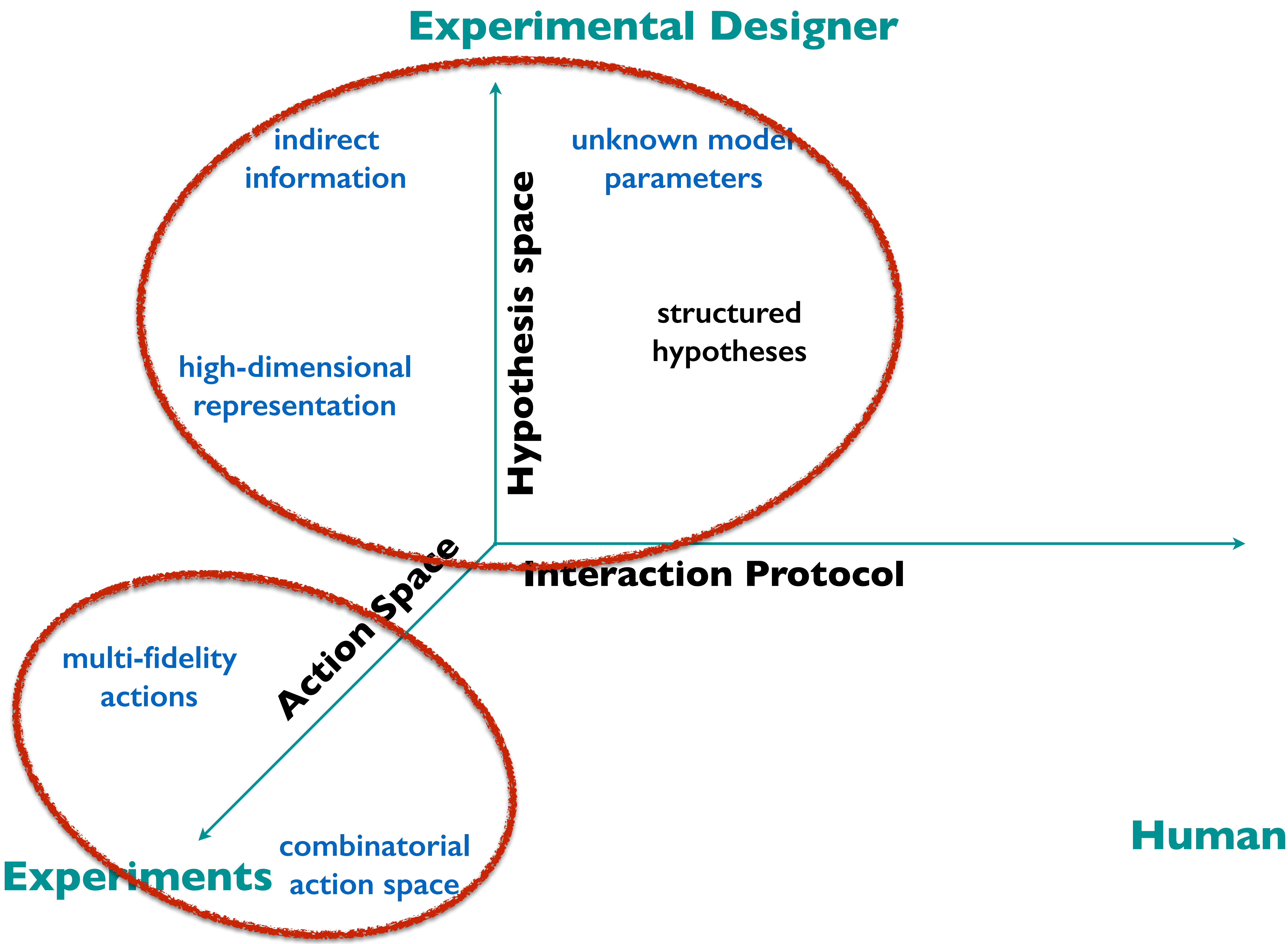
Experimental Designer
Hypothesis space
Interaction Protocol
Action Space
Human
Experiments

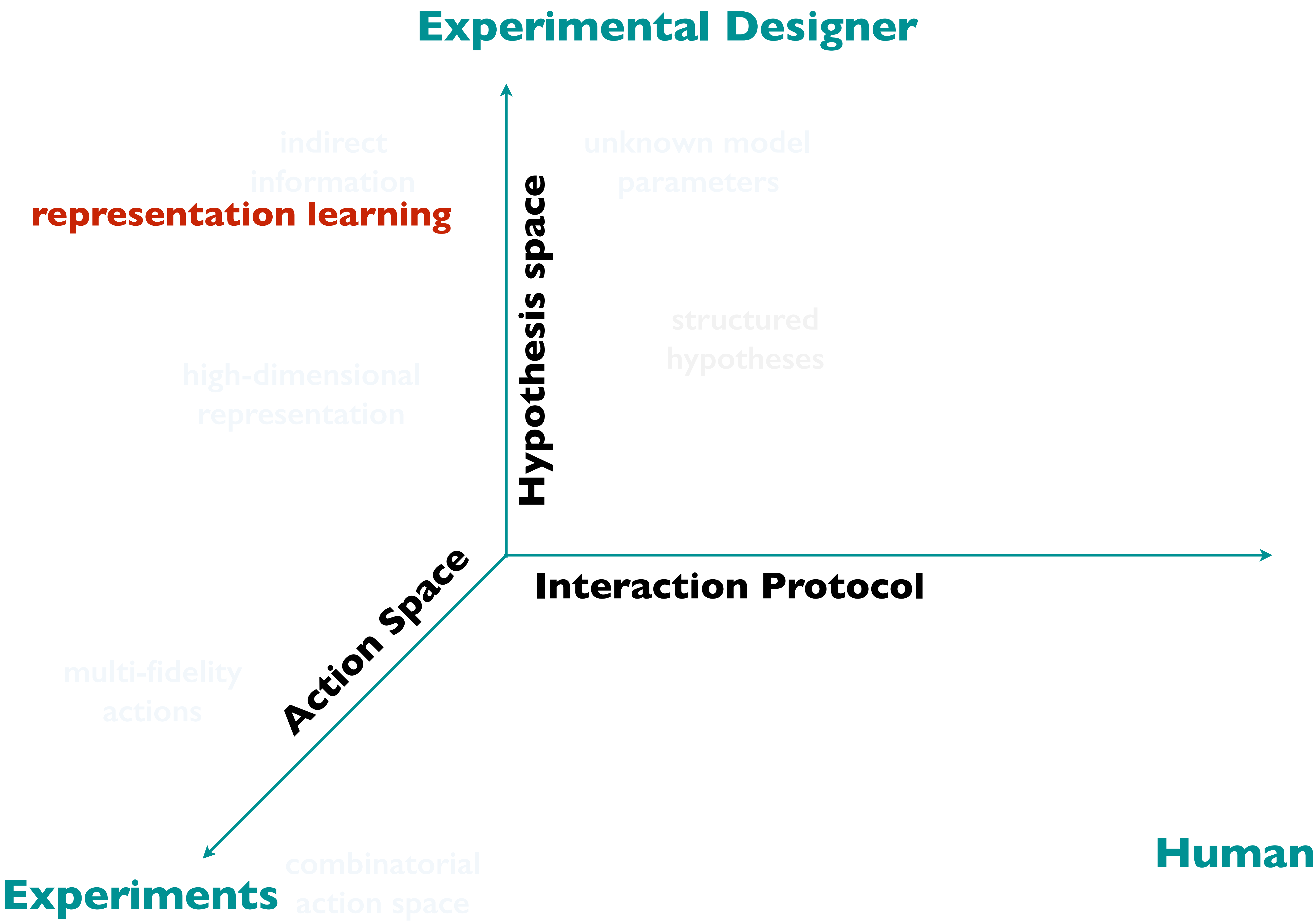


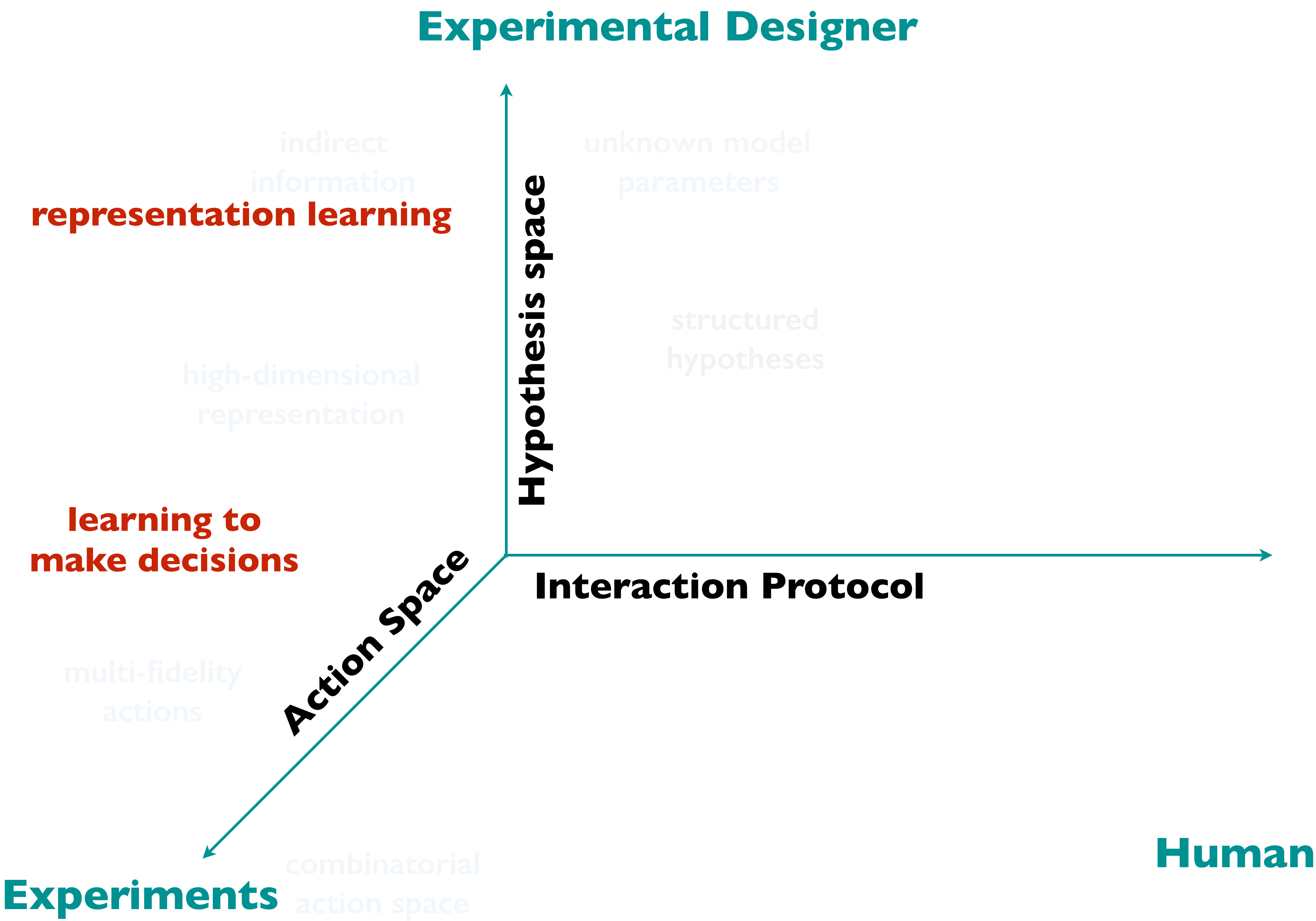
Experimental Designer
robotic manipulation
self-tuning system
Hypothesis space
Interaction Protocol
Action Space
200 nm
material science
directed evolution
Experiments
Human

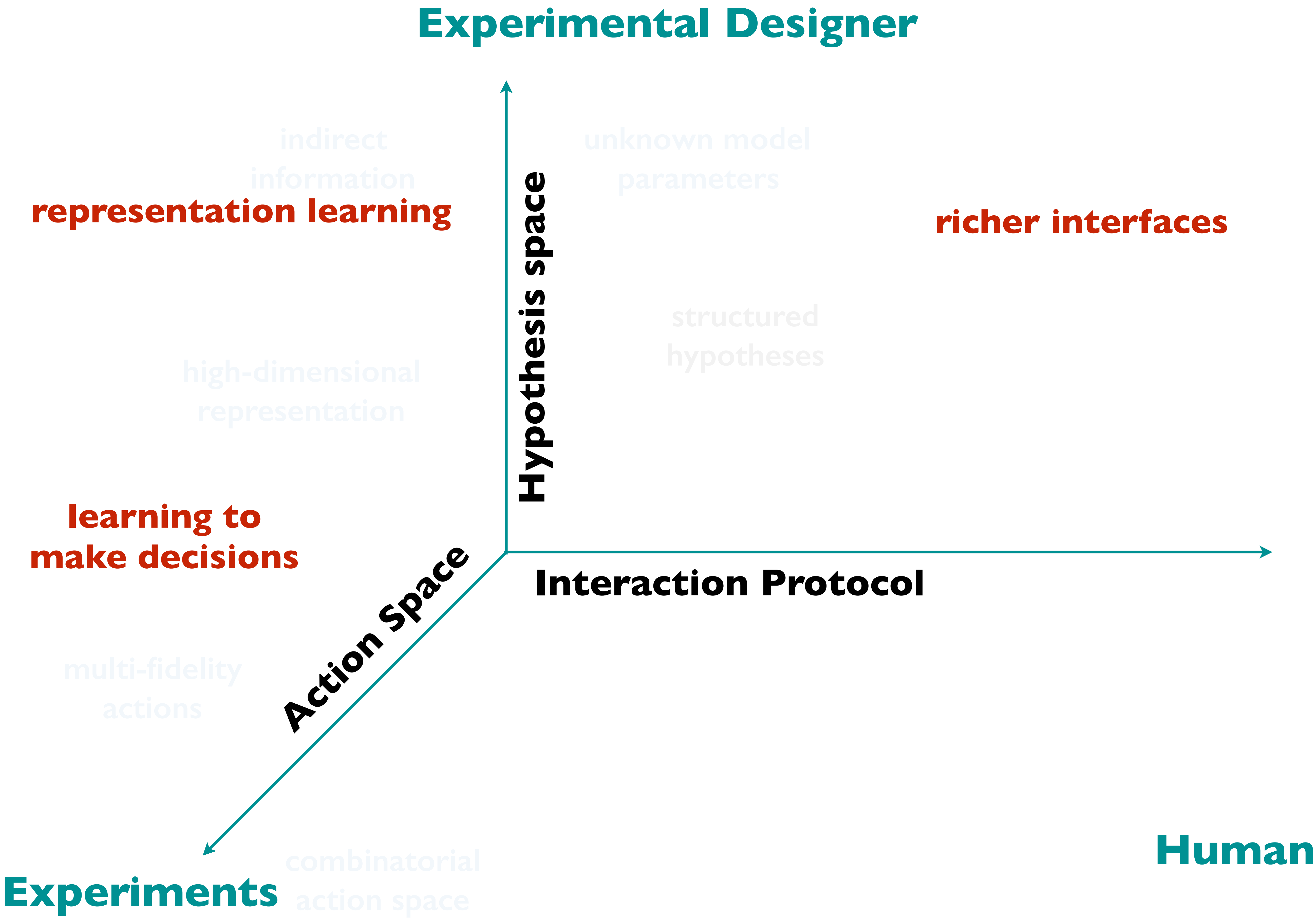


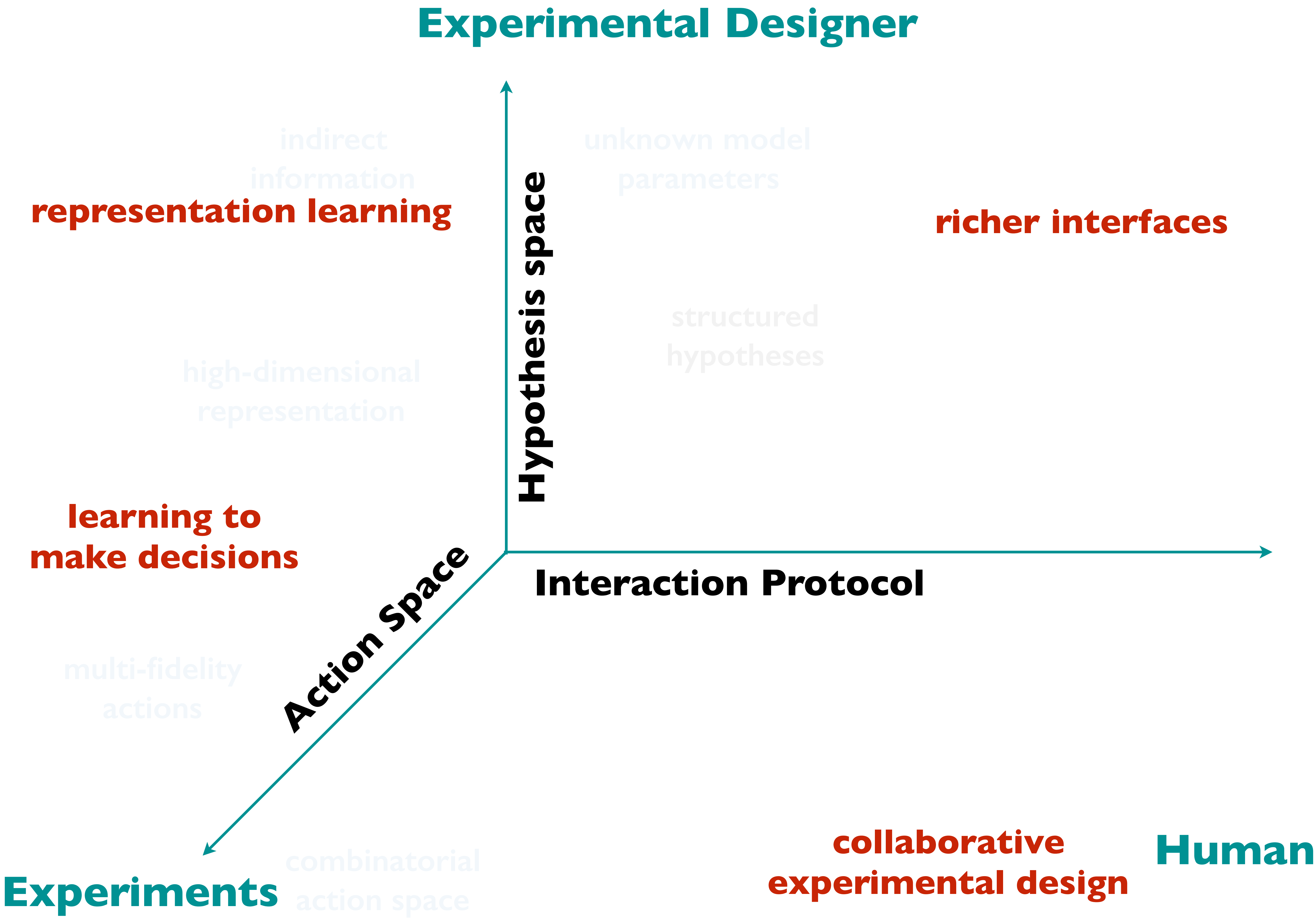
Experimental Designer
indirect information
unknown model parameters
Hypothesis space
structured hypotheses
high-dimensional representation
Action Space
Interaction Protocol
200 nm
material science
directed evolution
Experiments
Human


Experimental Designer
indirect information
unknown model parameters
Hypothesis space
structured hypotheses
high-dimensional representation
Action Space
Interaction Protocol
multi-fidelity actions
combinatorial action space
Experiments
Human


Experimental Designer
indirect information
unknown model parameters
representation learning
Hypothesis space
structured hypotheses
high-dimensional representation
Interaction Protocol
Action Space
multi-fidelity actions
Human
combinatorial action space
Experiments


Experimental Designer
indirect
information
unknown model
parameters
representation learning
Hypothesis space
structured
hypotheses
high-dimensional
representation
learning to
make decisions
Action Space
Interaction Protocol
multi-fidelity
actions
Human
combinatorial
action space
Experiments

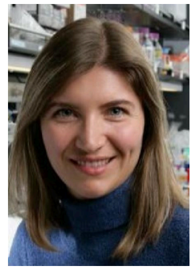

Experimental Designer
indirect information
unknown model parameters
representation learning
richer interfaces
Hypothesis space
structured hypotheses
high-dimensional representation
learning to make decisions
Interaction Protocol
Action Space
multi-fidelity actions
combinatorial action space
Human
Experiments


Experimental Designer
indirect information
unknown model parameters
representation learning
richer interfaces
Hypothesis space
structured hypotheses
high-dimensional representation
learning to make decisions
Interaction Protocol
Action Space
multi-fidelity actions
collaborative experimental design
Human
Experiments
combinatorial action space

Acknowledgments


Yisong Yue


Harry Atwater

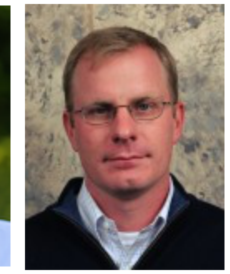

Viviana Gradinaru


Frances Arnold


Ufuk Topcu

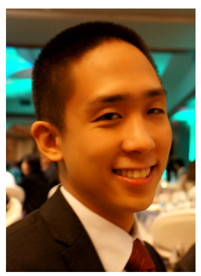

Brian Nord


Andreas Krause

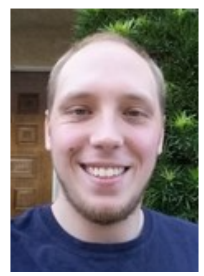

Drew Bagnell


Siddhartha Srinivasa

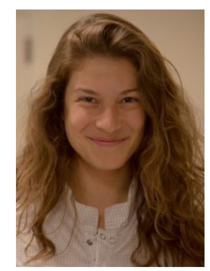

Jialin Song


Kevin Yang


Yury Tokpanov


David Brown


Kate Fountaine


Dagny Fleischman


Shervin Javdani

Alex Nettekoven

Steven Carr

erc

NSF

NORTHROP GRUMMAN

FNSNF
SWISS NATIONAL SCIENCE FOUNDATION

PIMCO

Center for Data and Computing
AT THE UNIVERSITY OF CHICAGO