Conformal prediction for the design problem

Clara Wong-Fannjiang
UC Berkeley

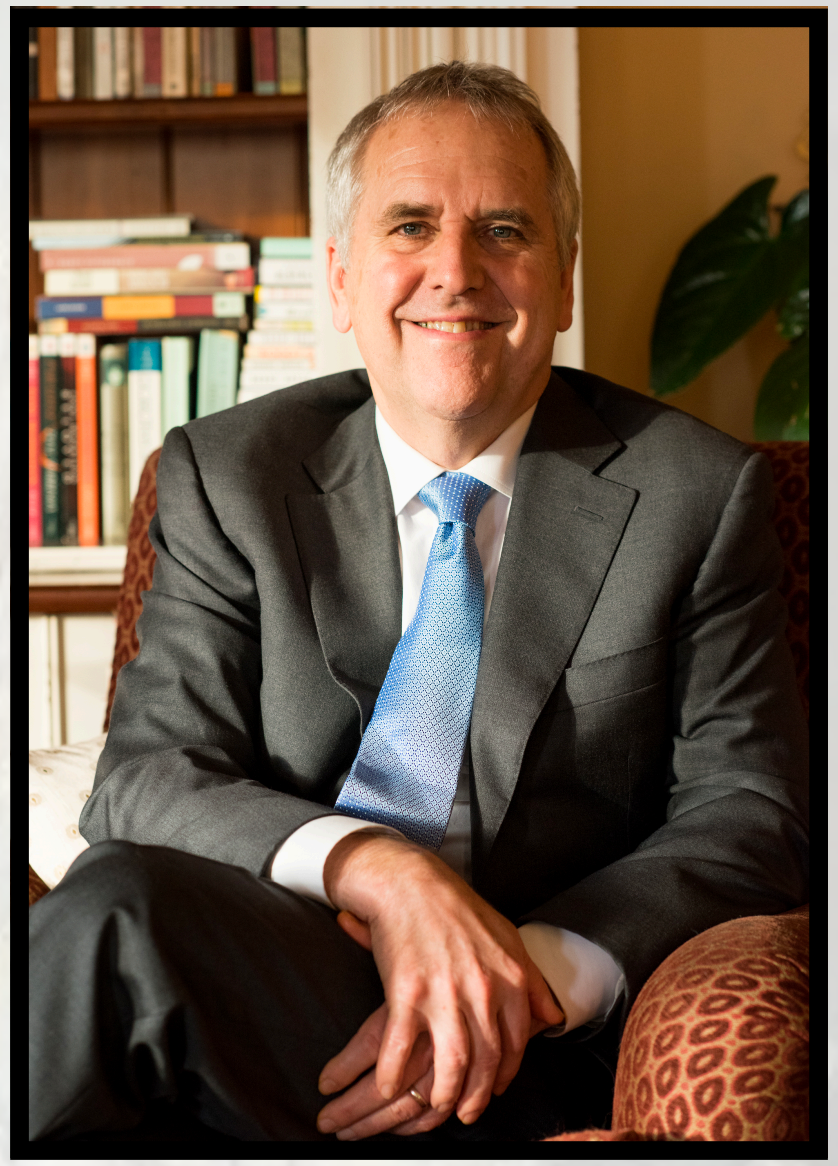
Conformal prediction for the design problem

arXiv:2202.03613, in submission

Stephen Bates

Anastasios N. Angelopoulos

Jennifer Listgarten



Michael I. Jordan

Roadmap

Single-shot design

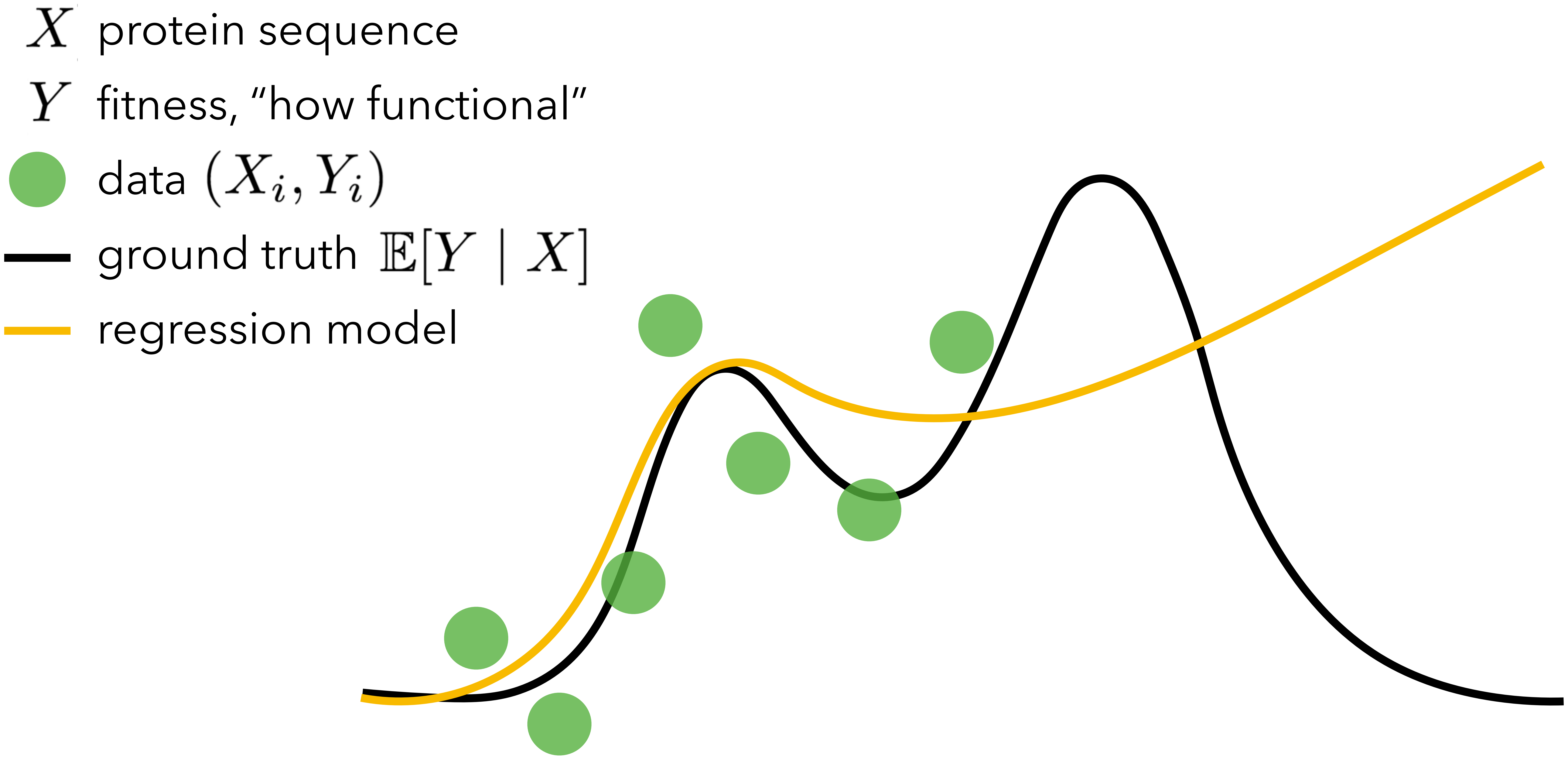
Feedback covariate shift

Intro to conformal prediction

Generalization to feedback covariate shift

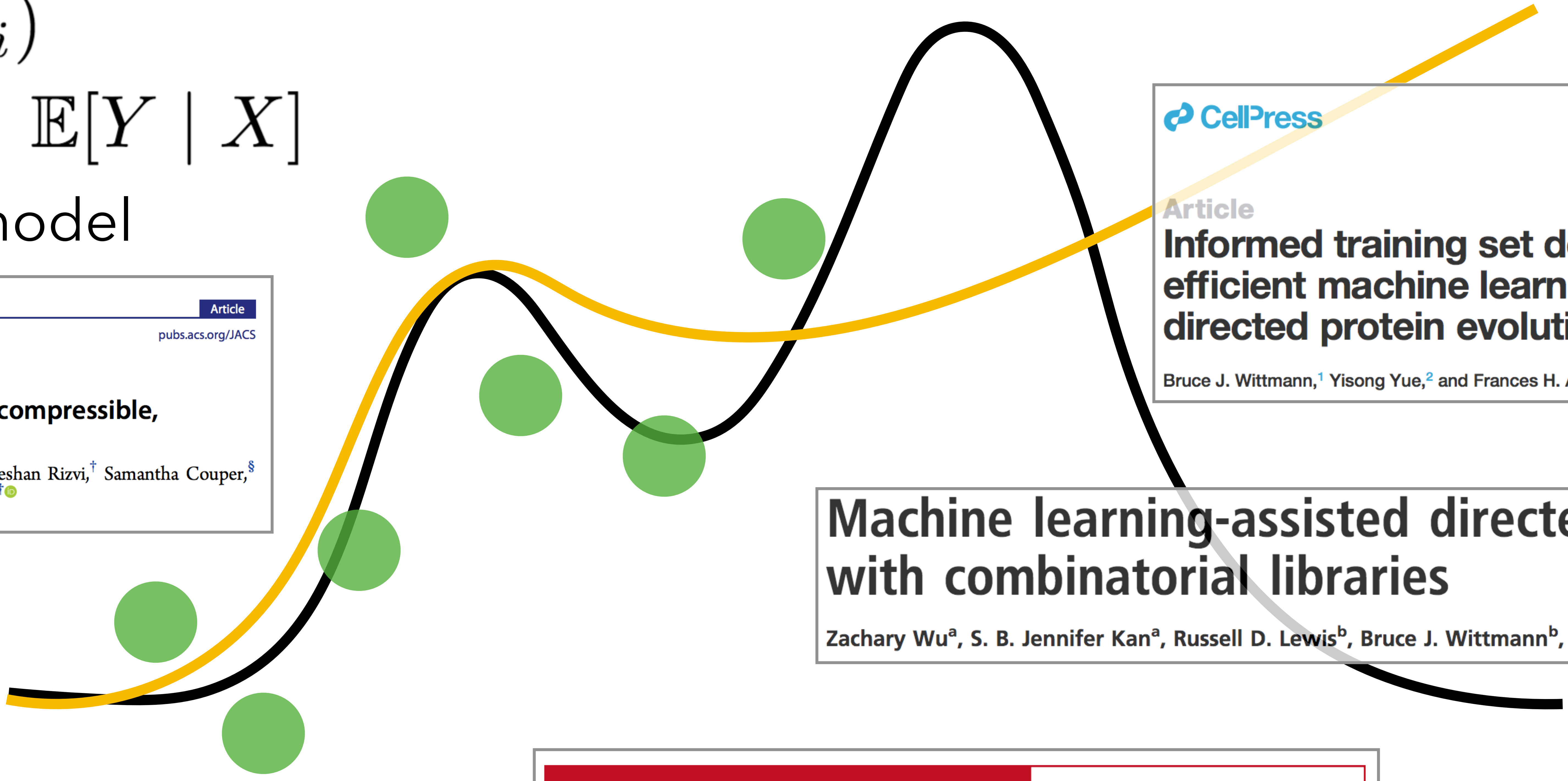
Demonstration with
protein design experiments

Single-shot design



Single-shot design

- X protein sequence
- Y fitness, "how functional"
- data (X_i, Y_i)
- ground truth $\mathbb{E}[Y \mid X]$
- regression model



nature methods — ARTICLES
https://doi.org/10.1038/s41592-021-01100-y

Low-*N* protein engineering with data-efficient deep learning

Surojit Biswas[1,2,6], Grigory Khimulya[3,6], Ethan C. Alley[4,6], Kevin M. Esvelt[4] and George M. Church[1,5]

CellPress — Cell Systems

Article

Informed training set design enables efficient machine learning-assisted directed protein evolution

Bruce J. Wittmann,[1] Yisong Yue,[2] and Frances H. Arnold[1,3,4,*]

JACS — JOURNAL OF THE AMERICAN CHEMICAL SOCIETY — Article

Cite This: J. Am. Chem. Soc. 2018, 140, 9844–9853 — pubs.acs.org/JACS

Machine Learning Directed Search for Ultraincompressible, Superhard Materials

Aria Mansouri Tehrani,[†,⊥] Anton O. Oliynyk,[†,⊥] Marcus Parry,[‡] Zeshan Rizvi,[†] Samantha Couper,[§] Feng Lin,[§] Lowell Miyagi,[§] Taylor D. Sparks,[‡] and Jakoah Brgoch[*,†]

Machine learning-assisted directed protein evolution with combinatorial libraries

Zachary Wu[a], S. B. Jennifer Kan[a], Russell D. Lewis[b], Bruce J. Wittmann[b], and Frances H. Arnold[a,b,1]

ACS central science — Research Article

Cite This: ACS Cent. Sci. 2018, 4, 268–276

Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules

Rafael Gómez-Bombarelli,[†,#] Jennifer N. Wei,[‡,#] David Duvenaud,[¶,#] José Miguel Hernández-Lobato,[§,#] Benjamín Sánchez-Lengeling,[‡] Dennis Sheberla,[‡] Jorge Aguilera-Iparraguirre,[†] Timothy D. Hirzel,[†] Ryan P. Adams,[∇,∥] and Alán Aspuru-Guzik[*,‡,⊥]

ARTICLES — nature methods
https://doi.org/10.1038/s41592-019-0583-8

Machine learning-guided channelrhodopsin engineering enables minimally invasive optogenetics

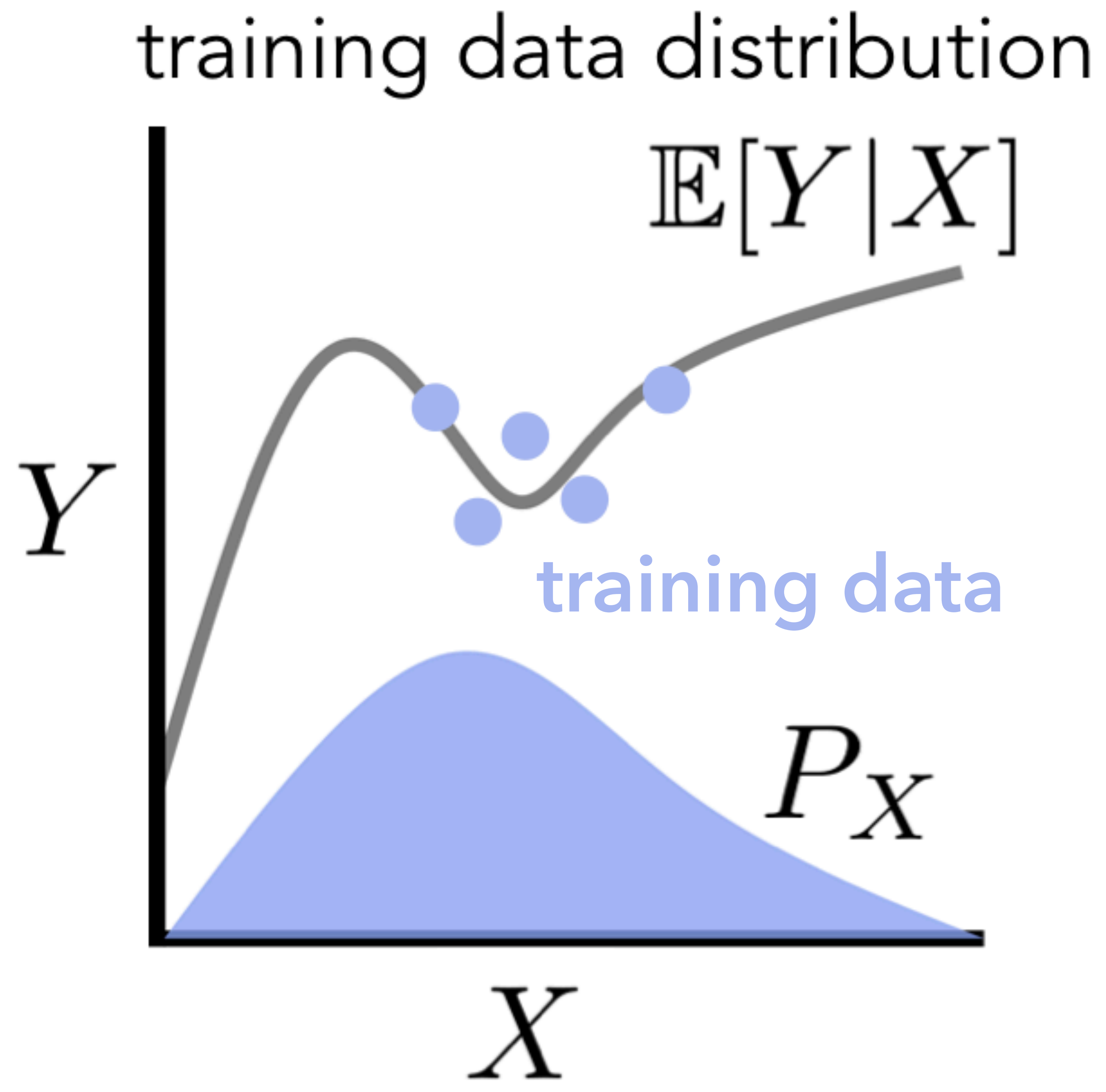
Claire N. Bedbrook[1], Kevin K. Yang[2,3], J. Elliott Robinson[1,3], Elisha D. Mackey[1], Viviana Gradinaru[1*] and Frances H. Arnold[1,2*]

Feedback covariate shift (FCS)

A distribution shift where the training and designed data are **statistically dependent**, because the training data are used to choose the distribution of designed inputs.

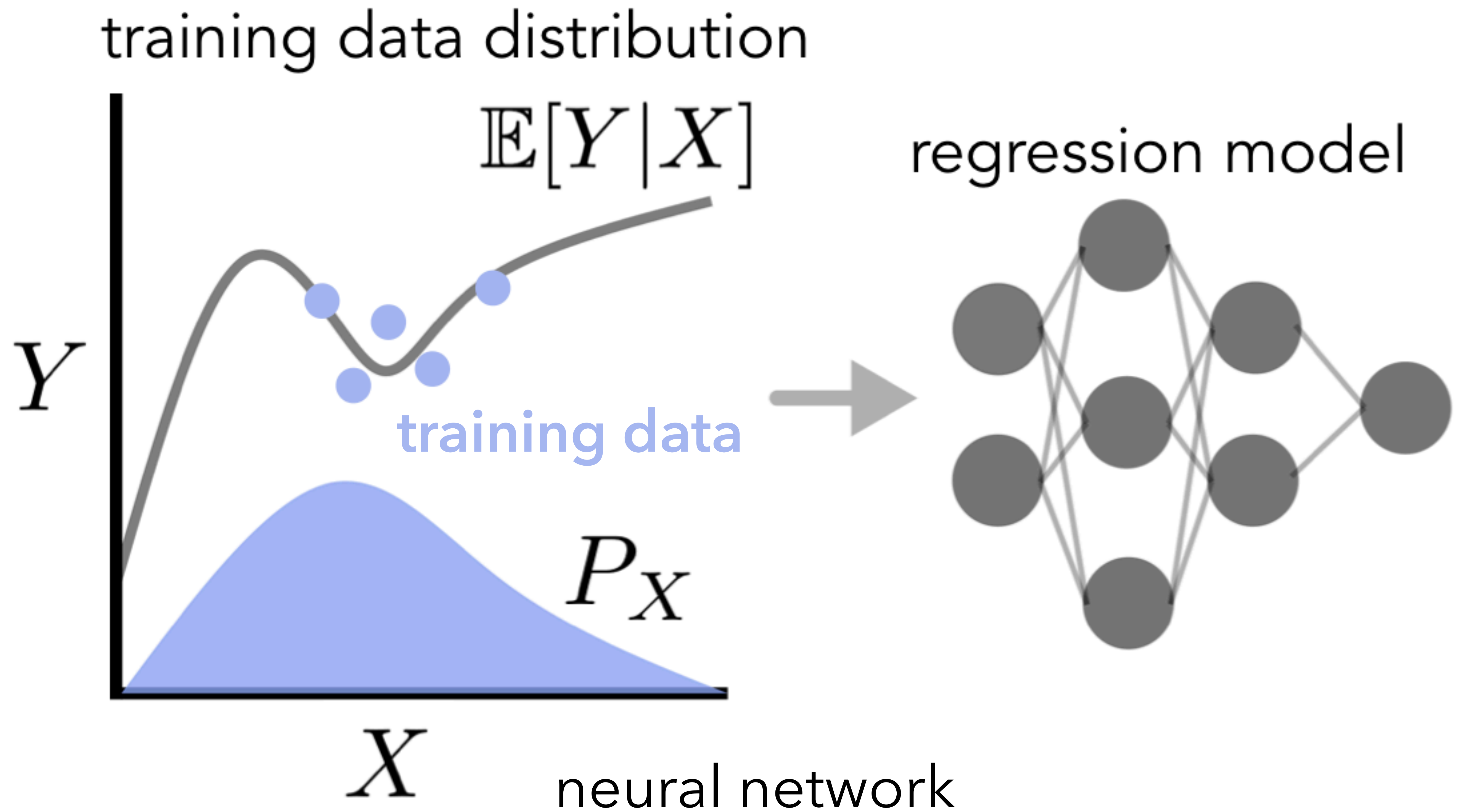
Feedback covariate shift (FCS)

A distribution shift where the training and designed data are **statistically dependent**, because the training data are used to choose the distribution of designed inputs.



Feedback covariate shift (FCS)

A distribution shift where the training and designed data are **statistically dependent**, because the training data are used to choose the distribution of designed inputs.



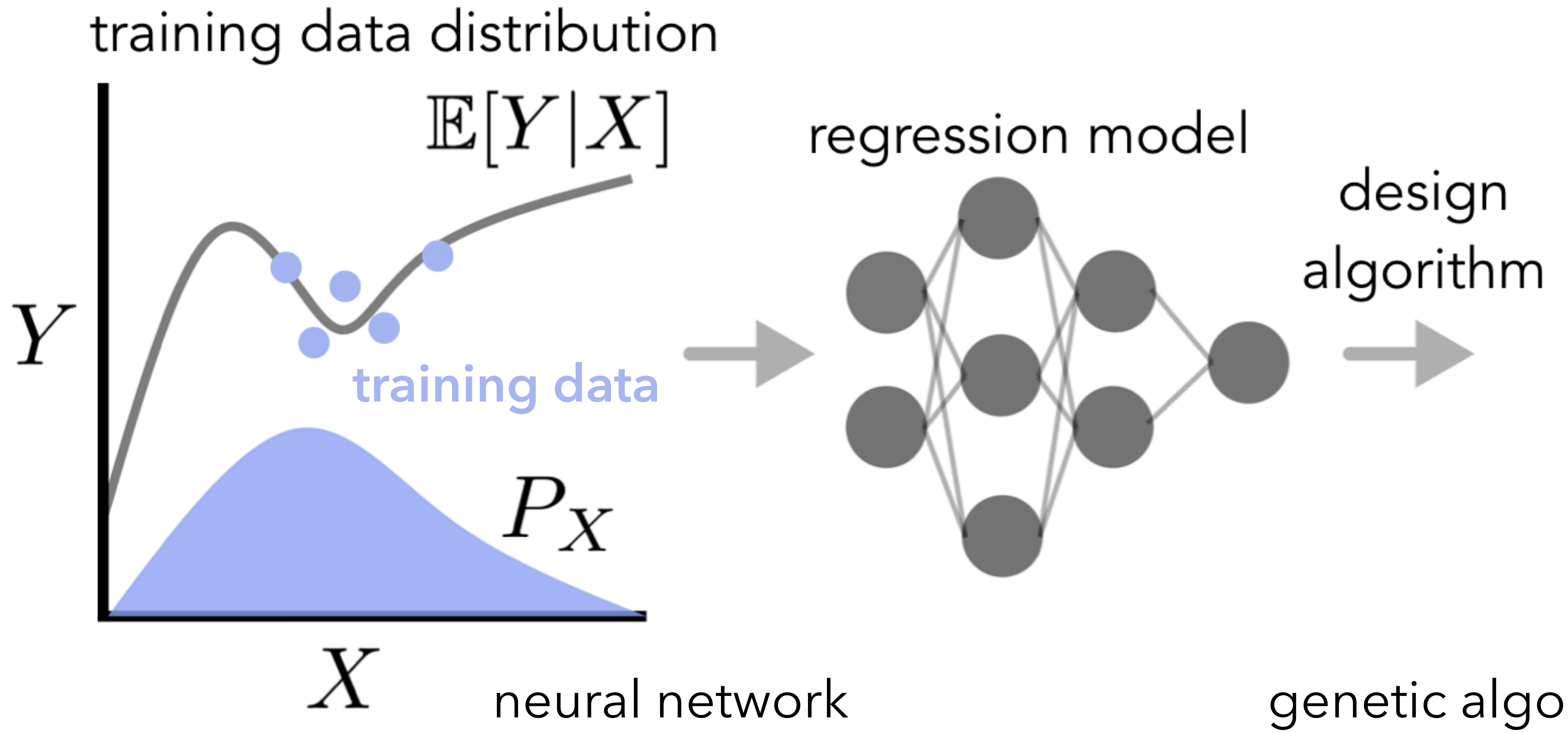
neural network

Gaussian process regression

gradient-boosted trees

Feedback covariate shift (FCS)

A distribution shift where the training and designed data are **statistically dependent**, because the training data are used to choose the distribution of designed inputs.

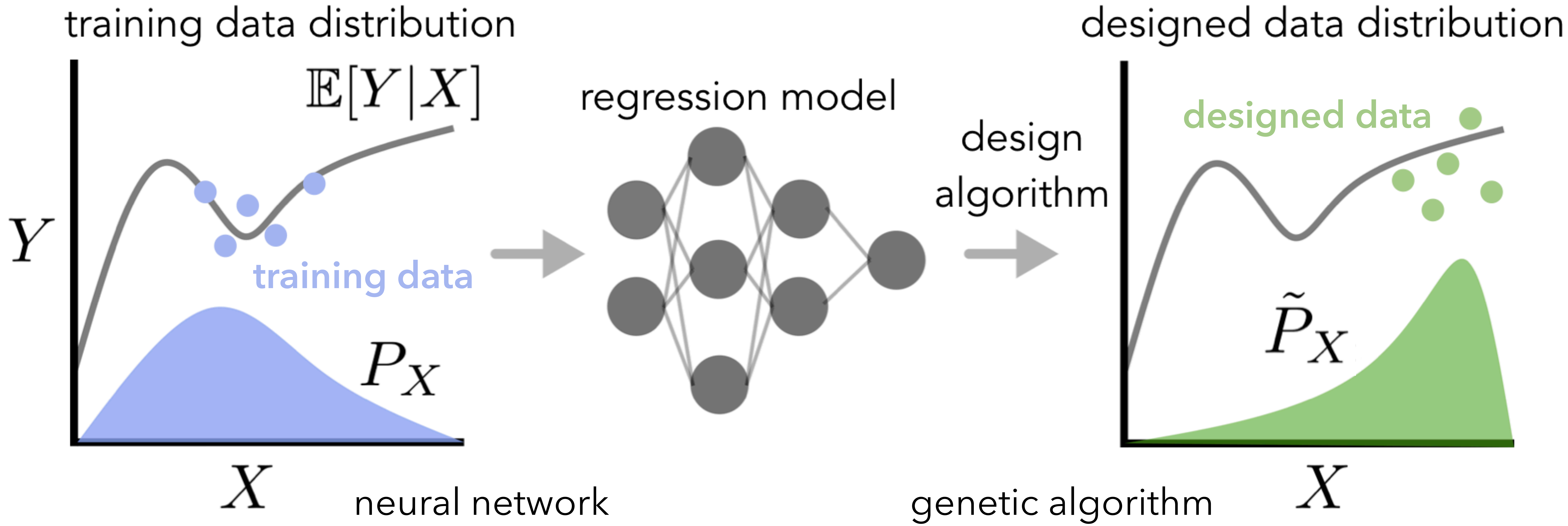


neural network
Gaussian process regression
gradient-boosted trees

genetic algorithm
gradient-based algorithm
conditional generative model

Feedback covariate shift (FCS)

A distribution shift where the training and designed data are **statistically dependent**, because the training data are used to choose the distribution of designed inputs.



neural network
Gaussian process regression
gradient-boosted trees

genetic algorithm
gradient-based algorithm
conditional generative model

Roadmap

Single-shot design

Feedback covariate shift

Intro to conformal prediction

Generalization to feedback covariate shift

Demonstration with
protein design experiments

Introduction to conformal prediction

$(X_1, Y_1), \ldots, (X_n, Y_n), (X_{n+1}, Y_{n+1})$ **exchangeable** training and test data points

Introduction to conformal prediction

$(X_1, Y_1), \ldots, (X_n, Y_n), (X_{n+1}, Y_{n+1})$ **exchangeable** training and test data points

Train regression model μ on training data.

Goal: given a test input, X_{n+1}, construct a confidence set, $C_\alpha(X_{n+1})$, that gives **coverage**:

$$\mathbb{P}(Y_{n+1} \in C_\alpha(X_{n+1})) \geq 1 - \alpha$$

for any user-specified **miscoverage level**, α.

Introduction to conformal prediction

$(X_1, Y_1), \ldots, (X_n, Y_n), (X_{n+1}, Y_{n+1})$ **exchangeable** training and test data points

Train regression model μ on training data.

Goal: given a test input, X_{n+1}, construct a confidence set, $C_\alpha(X_{n+1})$, that gives **coverage**:

$$\mathbb{P}(Y_{n+1} \in C_\alpha(X_{n+1})) \geq 1 - \alpha$$

for any user-specified **miscoverage level**, α.

- no assumptions on the model μ or on $P(Y \mid X)$
- finite-sample guarantee that holds for any amount of training data, n
- exchangeability is the key assumption!
- "**full**" **conformal prediction**

Introduction to conformal prediction

Intuition behind full conformal prediction: include all real values, y, such that the **candidate test point**, (X_{n+1}, y), looks sufficiently similar to training data as quantified by a score.

Score function $S((X,Y),D)$: user-specified function that quantifies how similar data point (X,Y) is to a multiset of data points D. Smaller score means more similar.

Introduction to conformal prediction

Intuition behind full conformal prediction: include all real values, y, such that the **candidate test point**, (X_{n+1}, y), looks sufficiently similar to training data as quantified by a score.

Score function $S((X,Y),D)$: user-specified function that quantifies how similar data point (X,Y) is to a multiset of data points D. Smaller score means more similar.

- representative example is model residual, $S((X,Y),D) = |Y - \mu_D(X)|$
- variance of predictions from ensemble of models
- variance of predictions for small, random perturbations of input
- …or any other heuristic notion of model's uncertainty

Introduction to conformal prediction

Intuition behind full conformal prediction: include all real values, y, such that the **candidate test point**, (X_{n+1}, y), looks sufficiently similar to training data as quantified by a score.

Score function $S((X,Y),D)$: user-specified function that quantifies how similar data point (X,Y) is to a multiset of data points D. Smaller score means more similar.

- representative example is model residual, $S((X,Y),D) = |Y - \mu_D(X)|$
- variance of predictions from ensemble of models
- variance of predictions for small, random perturbations of input
- …or any other heuristic notion of model's uncertainty

heuristic uncertainty → conformal prediction → statistically valid uncertainty

Angelopoulos & Bates (2021) arxiv:2107.07511

Introduction to conformal prediction

$$
\begin{aligned}
(X_1, Y_1)\\
(X_2, Y_2)\\
\dots\\
(X_n, Y_n)\\
(X_{n+1}, y)
\end{aligned}
$$

Introduction to conformal prediction

$$(X_1, Y_1)$$

$$(X_2, Y_2)$$

$$\dots$$

$$(X_n, Y_n)$$

$$(X_{n+1}, y)$$

Introduction to conformal prediction

(X_1, Y_1)

(X_2, Y_2)

$\dots$

(X_n, Y_n)

(X_{n+1}, y)

→ train regression model $\mu_{-1,y}$

Introduction to conformal prediction

$\mu_{-i,y}(X_i)$: regression model trained on all but i^{th} training point + candidate test point

(X_1, Y_1)

(X_2, Y_2)

$\ldots$

(X_n, Y_n)

(X_{n+1}, y)

$\longrightarrow$ train regression model $\mu_{-1,y}$

Introduction to conformal prediction

$\mu_{-i,y}(X_i)$: regression model trained on all but i^{th} training point + candidate test point

$$(X_1, Y_1) \xrightarrow{\text{score}} |Y_1 - \mu_{-1,y}(X_1)|$$

$$(X_2, Y_2)$$

$$\dots$$

$$(X_n, Y_n)$$

$$(X_{n+1}, y)$$

$\rightarrow$ train regression model $\mu_{-1,y}$

Introduction to conformal prediction

$\mu_{-i,y}(X_i)$: regression model trained on all but i^{th} training point + candidate test point

$$(X_1, Y_1) \xrightarrow{\text{score}} |Y_1 - \mu_{-1,y}(X_1)|$$

(X_2, Y_2)

$\dots$

(X_n, Y_n)

(X_{n+1}, y)

Introduction to conformal prediction

$\mu_{-i,y}(X_i)$: regression model trained on all but i^{th} training point + candidate test point

(X_1, Y_1) $\xrightarrow{\text{score}}$ $|Y_1 - \mu_{-1,y}(X_1)|$

(X_2, Y_2)

$\ldots$

(X_n, Y_n) → train regression model $\mu_{-2,y}$

(X_{n+1}, y)

Introduction to conformal prediction

$\mu_{-i,y}(X_i)$: regression model trained on all but i^{th} training point + candidate test point

$$(X_1, Y_1) \xrightarrow{\text{score}} |Y_1 - \mu_{-1,y}(X_1)|$$

$$(X_2, Y_2) \longrightarrow |Y_2 - \mu_{-2,y}(X_2)|$$

$$\dots$$

$$(X_n, Y_n)$$

$$(X_{n+1}, y)$$

train regression model $\mu_{-2,y}$

Introduction to conformal prediction

$\mu_{-i,y}(X_i)$: regression model trained on all but i^{th} training point + candidate test point

$$
\begin{array}{ccc}
(X_1, Y_1) & \xrightarrow{\text{score}} & |Y_1 - \mu_{-1,y}(X_1)| \\
(X_2, Y_2) & \longrightarrow & |Y_2 - \mu_{-2,y}(X_2)| \\
\dots & & \\
(X_n, Y_n) & \longrightarrow & |Y_n - \mu_{-n,y}(X_n)| \\
(X_{n+1}, y) & &
\end{array}
$$

Introduction to conformal prediction

$\mu_{-i,y}(X_i)$: regression model trained on all but i^{th} training point + candidate test point

$$(X_1, Y_1) \xrightarrow{\text{score}} |Y_1 - \mu_{-1,y}(X_1)|$$

$$(X_2, Y_2) \longrightarrow |Y_2 - \mu_{-2,y}(X_2)|$$

$$\ldots \longrightarrow \text{train regression model } \mu_{1:n}$$

$$(X_n, Y_n) \longrightarrow |Y_n - \mu_{-n,y}(X_n)|$$

$$(X_{n+1}, y)$$

Introduction to conformal prediction

$\mu_{-i,y}(X_i)$: regression model trained on all but i^{th} training point + candidate test point

$$
\begin{array}{rcl}
(X_1, Y_1) & \xrightarrow{\text{score}} & |Y_1 - \mu_{-1,y}(X_1)| \\
(X_2, Y_2) & \longrightarrow & |Y_2 - \mu_{-2,y}(X_2)| \\
\ldots & \longrightarrow & \text{train regression model } \mu_{1:n} \\
(X_n, Y_n) & \longrightarrow & |Y_n - \mu_{-n,y}(X_n)| \\
(X_{n+1}, y) & \longrightarrow & |y - \mu_{1:n}(X_{n+1})|
\end{array}
$$

Introduction to conformal prediction

$\mu_{-i,y}(X_i)$: regression model trained on all but i^{th} training point + candidate test point

$$
\begin{array}{ccc}
(X_1, Y_1) & \xrightarrow{\text{score}} & |Y_1 - \mu_{-1,y}(X_1)| \\
(X_2, Y_2) & \longrightarrow & |Y_2 - \mu_{-2,y}(X_2)| \\
\dots & & \\
(X_n, Y_n) & \longrightarrow & |Y_n - \mu_{-n,y}(X_n)| \\
(X_{n+1}, y) & \longrightarrow & |y - \mu_{1:n}(X_{n+1})|
\end{array}
$$

Introduction to conformal prediction

$$S_i(y) = |Y_i - \mu_{-i,y}(X_i)|$$

$\mu_{-i,y}(X_i)$: regression model trained on all but *i*th training point + candidate test point

$$\begin{aligned}
(X_1, Y_1) &\xrightarrow{\text{score}} |Y_1 - \mu_{-1,y}(X_1)| = S_1(y) \\
(X_2, Y_2) &\longrightarrow |Y_2 - \mu_{-2,y}(X_2)| = S_2(y) \\
&\cdots \\
(X_n, Y_n) &\longrightarrow |Y_n - \mu_{-n,y}(X_n)| = S_n(y) \\
(X_{n+1}, y) &\longrightarrow |y - \mu_{1:n}(X_{n+1})| = S_{n+1}(y)
\end{aligned}$$

Introduction to conformal prediction

$$S_i(y) = |Y_i - \mu_{-i,y}(X_i)|$$

$\mu_{-i,y}(X_i)$: regression model trained on all but i^{th} training point + candidate test point

$$\begin{array}{ccl}
(X_1, Y_1) & \xrightarrow{\text{score}} & |Y_1 - \mu_{-1,y}(X_1)| = S_1(y) \\
(X_2, Y_2) & \longrightarrow & |Y_2 - \mu_{-2,y}(X_2)| = S_2(y) \\
\dots & & \\
(X_n, Y_n) & \longrightarrow & |Y_n - \mu_{-n,y}(X_n)| = S_n(y) \\
(X_{n+1}, y) & \longrightarrow & |y - \mu_{1:n}(X_{n+1})| = S_{n+1}(y)
\end{array}$$

Introduction to conformal prediction

$$S_i(y) = |Y_i - \mu_{-i,y}(X_i)|$$

$\mu_{-i,y}(X_i)$: regression model trained on all but i^{th} training point + candidate test point

$$\begin{aligned}
(X_1, Y_1) &\xrightarrow{\text{score}} |Y_1 - \mu_{-1,y}(X_1)| = S_1(y) \\
(X_2, Y_2) &\longrightarrow |Y_2 - \mu_{-2,y}(X_2)| = S_2(y) \\
&\ldots \\
(X_n, Y_n) &\longrightarrow |Y_n - \mu_{-n,y}(X_n)| = S_n(y) \\
(X_{n+1}, y) &\longrightarrow |y - \mu_{1:n}(X_{n+1})| = S_{n+1}(y)
\end{aligned}$$

$S_{(\lceil (1-\alpha)(n+1) \rceil)}(y)$ $\qquad$ $S_{(\lceil 0.6 \cdot (n+1) \rceil)} = S_{(4)}(y)$

Introduction to conformal prediction

$$S_i(y) = |Y_i - \mu_{-i,y}(X_i)|$$

$\mu_{-i,y}(X_i)$: regression model trained on all but i^{th} training point + candidate test point

$$\begin{array}{ccc}
(X_1, Y_1) & \xrightarrow{\text{score}} & |Y_1 - \mu_{-1,y}(X_1)| = S_1(y) \\
(X_2, Y_2) & \longrightarrow & |Y_2 - \mu_{-2,y}(X_2)| = S_2(y) \\
\ldots & & \\
(X_n, Y_n) & \longrightarrow & |Y_n - \mu_{-n,y}(X_n)| = S_n(y) \\
(X_{n+1}, y) & \longrightarrow & |y - \mu_{1:n}(X_{n+1})| = S_{n+1}(y)
\end{array}$$

$$S_{n+1}(y) \leq S_{(\lceil (1-\alpha)(n+1) \rceil)}(y)$$

$$S_{(\lceil 0.6 \cdot (n+1) \rceil)} = S_{(4)}(y)$$

Introduction to conformal prediction

$S_i(y) = |Y_i - \mu_{-i,y}(X_i)|$

$\mu_{-i,y}(X_i)$: regression model trained on all but i^{th} training point + candidate test point

$$\begin{array}{ccc}
(X_1, Y_1) & \xrightarrow{\text{score}} & |Y_1 - \mu_{-1,y}(X_1)| = S_1(y) \\
(X_2, Y_2) & \longrightarrow & |Y_2 - \mu_{-2,y}(X_2)| = S_2(y) \\
\dots & & \\
(X_n, Y_n) & \longrightarrow & |Y_n - \mu_{-n,y}(X_n)| = S_n(y) \\
(X_{n+1}, y) & \longrightarrow & |y - \mu_{1:n}(X_{n+1})| = S_{n+1}(y)
\end{array}$$

If $S_{n+1}(y) \leq S_{(\lceil (1-\alpha)(n+1) \rceil)}(y)$, include the candidate label y in $C_\alpha(X_{n+1})$.

Introduction to conformal prediction

$$S_i(y) = |Y_i - \mu_{-i,y}(X_i)|$$

$\mu_{-i,y}(X_i)$: regression model trained on all but i^{th} training point + candidate test point

$$\begin{aligned}
(X_1, Y_1) &\xrightarrow{\text{score}} |Y_1 - \mu_{-1,y}(X_1)| = S_1(y) \\
(X_2, Y_2) &\longrightarrow |Y_2 - \mu_{-2,y}(X_2)| = S_2(y) \\
&\cdots \\
(X_n, Y_n) &\longrightarrow |Y_n - \mu_{-n,y}(X_n)| = S_n(y) \\
(X_{n+1}, y) &\longrightarrow |y - \mu_{1:n}(X_{n+1})| = S_{n+1}(y)
\end{aligned}$$

If $S_{n+1}(y) \leq S_{(\lceil (1-\alpha)(n+1) \rceil)}(y)$, include the candidate label y in $C_\alpha(X_{n+1})$.

$$C_\alpha(X_{n+1}) = \{y \in \mathbb{R} : S_{n+1}(y) \leq S_{(\lceil (1-\alpha)(n+1) \rceil)}(y)\}$$

Introduction to conformal prediction

$$C_\alpha(X_{n+1}) = \{y \in \mathbb{R} : S_{n+1}(y) \leq S_{(\lceil(1-\alpha)(n+1)\rceil)}(y)\}$$

If the data are exchangeable, this confidence set achieves **coverage**:

$$\mathbb{P}(Y_{n+1} \in C_\alpha(X_{n+1})) \geq 1 - \alpha$$

Vovk et al. (2005), Vovk et al. (2009) *Ann. Stat.*

Introduction to conformal prediction

$S_1, \ldots, S_n$ exchangeable random variables

$S_{(\lceil (1-\alpha)(n+1) \rceil)} = \lceil (1-\alpha)(n+1) \rceil$ smallest variable

Introduction to conformal prediction

$S_1, \ldots, S_n, S_{n+1}$ exchangeable random variables

$S_{(\lceil(1-\alpha)(n+1)\rceil)} = \lceil(1-\alpha)(n+1)\rceil$ smallest variable

$$\mathbb{P}(S_{n+1} \leq S_{(\lceil(1-\alpha)(n+1)\rceil)}) \geq 1 - \alpha$$

Introduction to conformal prediction

$S_1, \ldots, S_n, S_{n+1}$ exchangeable random variables

$S_{(\lceil (1-\alpha)(n+1) \rceil)} = \lceil (1-\alpha)(n+1) \rceil$ smallest variable

$$\mathbb{P}(S_{n+1} \leq S_{(\lceil (1-\alpha)(n+1) \rceil)}) \geq 1 - \alpha$$

Proof. Since $S_1, \ldots, S_n, S_{n+1}$ are exchangeable, the rank of S_{n+1} is uniformly distributed on $\{1, \ldots, n+1\}$.

$$\mathbb{P}(\pi(S_{n+1}) \leq \lceil (1-\alpha)(n+1) \rceil) \geq 1 - \alpha$$

Introduction to conformal prediction

$$C_\alpha(X_{n+1}) = \{y \in \mathbb{R} : S_{n+1}(y) \leq S_{(\lceil (1-\alpha)(n+1) \rceil)}(y)\}$$

If the data are exchangeable, this confidence set achieves **coverage**:

$$\mathbb{P}(Y_{n+1} \in C_\alpha(X_{n+1})) \geq 1 - \alpha$$

Introduction to conformal prediction

$$C_\alpha(X_{n+1}) = \{y \in \mathbb{R} : S_{n+1}(y) \leq S_{(\lceil(1-\alpha)(n+1)\rceil)}(y)\}$$

If the data are exchangeable, this confidence set achieves **coverage**:

$$\mathbb{P}(Y_{n+1} \in C_\alpha(X_{n+1})) \geq 1 - \alpha$$

Proof:

$$\begin{aligned} &(X_1, Y_1) \\ &\cdots \\ &(X_n, Y_n) \\ &(X_{n+1}, Y_{n+1}) \end{aligned}$$

exchangeable data

Introduction to conformal prediction

$$C_\alpha(X_{n+1}) = \{y \in \mathbb{R} : S_{n+1}(y) \leq S_{(\lceil(1-\alpha)(n+1)\rceil)}(y)\}$$

If the data are exchangeable, this confidence set achieves **coverage**:

$$\mathbb{P}(Y_{n+1} \in C_\alpha(X_{n+1})) \geq 1 - \alpha$$

Proof:

$$\begin{array}{rl} (X_1, Y_1) & S_1(Y_{n+1}) \\ \cdots & \cdots \\ (X_n, Y_n) & S_n(Y_{n+1}) \\ (X_{n+1}, Y_{n+1}) & S_{n+1}(Y_{n+1}) \end{array}$$

exchangeable data $\Longrightarrow$ exchangeable scores

Introduction to conformal prediction

$$C_\alpha(X_{n+1}) = \{y \in \mathbb{R} : S_{n+1}(y) \leq S_{(\lceil (1-\alpha)(n+1) \rceil)}(y)\}$$

If the data are exchangeable, this confidence set achieves **coverage**:

$$\mathbb{P}(Y_{n+1} \in C_\alpha(X_{n+1})) \geq 1 - \alpha$$

Proof:

$$\begin{array}{cc}
(X_1, Y_1) & S_1(Y_{n+1}) \\
\cdots & \cdots \\
(X_n, Y_n) & S_n(Y_{n+1}) \\
(X_{n+1}, Y_{n+1}) & S_{n+1}(Y_{n+1})
\end{array}$$

exchangeable data $\Longrightarrow$ exchangeable scores

$$\mathbb{P}(S_{n+1}(Y_{n+1}) \leq S_{(\lceil (1-\alpha)(n+1) \rceil)}(Y_{n+1})) \geq 1 - \alpha$$

Introduction to conformal prediction

$$C_\alpha(X_{n+1}) = \{y \in \mathbb{R} : S_{n+1}(y) \leq S_{(\lceil (1-\alpha)(n+1) \rceil)}(y)\}$$

If the data are exchangeable, this confidence set achieves **coverage**:

$$\mathbb{P}(Y_{n+1} \in C_\alpha(X_{n+1})) \geq 1 - \alpha$$

Proof:

$$\begin{array}{rl} (X_1, Y_1) & S_1(Y_{n+1}) \\ \cdots & \cdots \\ (X_n, Y_n) & S_n(Y_{n+1}) \\ (X_{n+1}, Y_{n+1}) & S_{n+1}(Y_{n+1}) \end{array}$$

exchangeable data $\Longrightarrow$ exchangeable scores

$$\mathbb{P}(S_{n+1}(Y_{n+1}) \leq S_{(\lceil (1-\alpha)(n+1) \rceil)}(Y_{n+1})) \geq 1 - \alpha$$

Introduction to conformal prediction

$$C_\alpha(X_{n+1}) = \left\{ y \in \mathbb{R} : S_{n+1}(y) \leq \textsc{Quantile}_{1-\alpha}\left(\frac{1}{n+1} \sum_{i=1}^{n+1} \delta_{S_i(y)} \right) \right\}$$

If the data are exchangeable, this confidence set achieves **coverage**:

$$\mathbb{P}(Y_{n+1} \in C_\alpha(X_{n+1})) \geq 1 - \alpha$$

Proof:

$$\begin{array}{cc} (X_1, Y_1) & S_1(Y_{n+1}) \\ \cdots & \cdots \\ (X_n, Y_n) & S_n(Y_{n+1}) \\ (X_{n+1}, Y_{n+1}) & S_{n+1}(Y_{n+1}) \end{array}$$

exchangeable data $\Longrightarrow$ exchangeable scores

$$\mathbb{P}(S_{n+1}(Y_{n+1}) \leq S_{(\lceil (1-\alpha)(n+1) \rceil)}(Y_{n+1})) \geq 1 - \alpha$$

Introduction to conformal prediction

Intuition for exchangeable data: include all real values, y, such that the **candidate test point**, (X_{n+1}, y), looks sufficiently similar to training data as quantified by a score.

$$C_\alpha(X_{n+1}) = \left\{ y \in \mathbb{R} : \overset{\text{score of candidate test point}}{S_{n+1}(y)} \leq \textsc{Quantile}_{1-\alpha}\left(\underset{\text{equal weights}}{\frac{1}{n+1}} \sum_{i=1}^{n+1} \delta_{\overset{\text{scores of training + candidate test points}}{S_i(y)}} \right) \right\}$$

Conformal prediction for feedback covariate shift

Intuition for data under FCS: include all real values, y, such that the **candidate test point**, (X_{n+1}, y), looks sufficiently similar to the **weighted** training data as quantified by a score.

designed input
(*e.g.*, designed protein)

scores of training + candidate test points

score of candidate test point

$$C_\alpha(X_{n+1}) = \left\{ y \in \mathbb{R} : S_{n+1}(y) \leq \textsc{Quantile}_{1-\alpha}\left(\sum_{i=1}^{n+1} w_i(y) \cdot \delta_{S_i(y)} \right) \right\}$$

weights that take into account that the data are **statistically dependent** through FCS

Fannjiang et al. (2022), arXiv:2202.03613

Conformal prediction for feedback covariate shift

Intuition for data under FCS: include all real values, y, such that the **candidate test point**, (X_{n+1}, y), looks sufficiently similar to the **weighted** training data as quantified by a score.

designed input
(*e.g.*, designed protein)

scores of training + candidate test points

score of candidate test point

$$C_\alpha(X_{n+1}) = \left\{ y \in \mathbb{R} : S_{n+1}(y) \leq \textsc{Quantile}_{1-\alpha}\left(\sum_{i=1}^{n+1} w_i(y) \cdot \delta_{S_i(y)} \right) \right\}$$

$$w_i(y) \propto \frac{\tilde{p}_{X;Z_{-i} \cup \{(X_{n+1}, y)\}}(X_i)}{p_X(X_i)}$$

weights that take into account that the data are **statistically dependent** through FCS

training input distribution

input distribution induced by regression model trained on $Z_{-i} \cup \{(X_{n+1}, y)\}$

Fannjiang et al. (2022), arXiv:2202.03613

Conformal prediction for feedback covariate shift

Intuition for data under FCS: include all real values, y, such that the **candidate test point**, (X_{n+1}, y), looks sufficiently similar to the **weighted** training data as quantified by a score.

designed input
(*e.g.*, designed protein)

scores of training + candidate test points

score of candidate test point

$$C_\alpha(X_{n+1}) = \left\{ y \in \mathbb{R} : S_{n+1}(y) \le \textsc{Quantile}_{1-\alpha} \left(\sum_{i=1}^{n+1} w_i(y) \cdot \delta_{S_i(y)} \right) \right\}$$

$$w_i(y) \propto \frac{\tilde{p}_{X; Z_{-i} \cup \{(X_{n+1}, y)\}}(X_i)}{p_X(X_i)}$$

weights that take into account that the data are **statistically dependent** through FCS

training input distribution

input distribution induced by regression model trained on

$$Z_{-i} \cup \{(X_{n+1}, y)\} = \{(X_1, Y_1), \ldots, (X_n, Y_n), (X_{n+1}, y)\} \setminus \{(X_i, Y_i)\}$$

Fannjiang et al. (2022), arXiv:2202.03613

Conformal prediction for feedback covariate shift

Intuition for data under FCS: include all real values, y, such that the **candidate test point**, (X_{n+1}, y), looks sufficiently similar to the **weighted** training data as quantified by a score.

designed input
(*e.g.*, designed protein)

scores of training + candidate test points

score of candidate test point

$$C_\alpha(X_{n+1}) = \left\{ y \in \mathbb{R} : S_{n+1}(y) \leq \textsc{Quantile}_{1-\alpha}\left(\sum_{i=1}^{n+1} w_i(y) \cdot \delta_{S_i(y)} \right) \right\}$$

weights that take into account that the data are **statistically dependent** through FCS

This set achieves coverage for data under feedback covariate shift:

$$\mathbb{P}(Y_{n+1} \in C_\alpha(X_{n+1})) \geq 1 - \alpha$$

Fannjiang et al. (2022), arXiv:2202.03613

Roadmap

Single-shot design

Feedback covariate shift

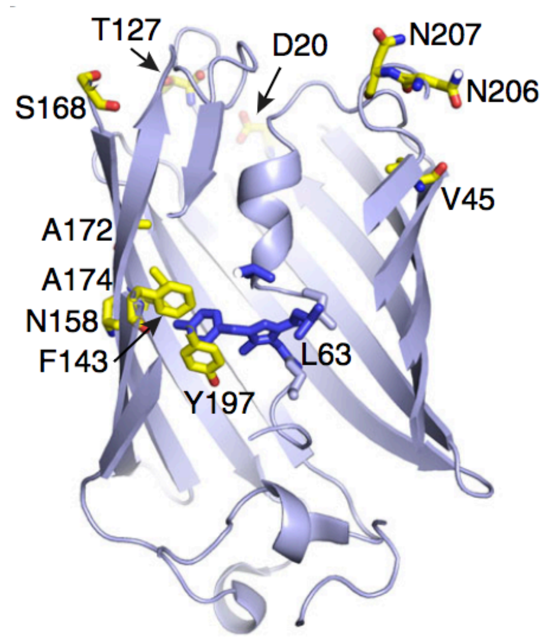
Intro to conformal prediction

Generalization to feedback covariate shift

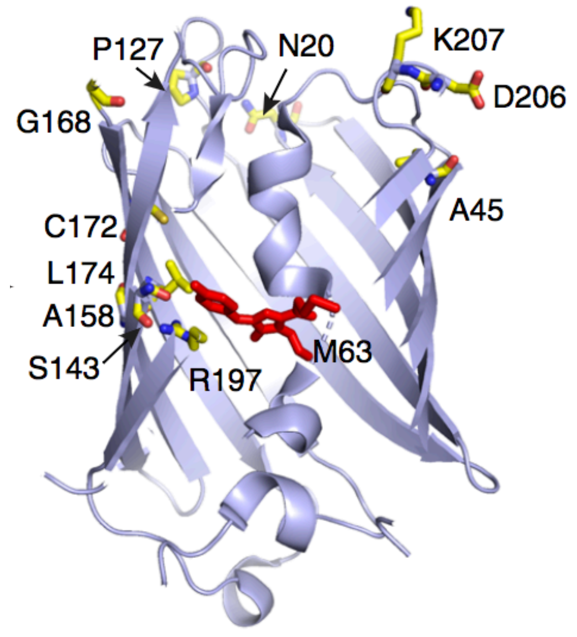
Demonstration with protein design experiments

Protein design experiments

Goal: design brighter blue fluorescent and red fluorescent proteins



example blue fluorescent protein



example red fluorescent protein

- "Complete landscape" data set: blue/red brightness of all 2^{13} variants that differ at 13 sites
- Training data: sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ uniformly from the 2^{13} variants
- Regression model: ridge regression model with singleton and pairwise features
- Design algorithm: sample sequence, X_{n+1}, from

regression model
fit to *n* training points

$$\tilde{p}_X(X_{n+1}) \propto \exp(\lambda \cdot \boxed{\mu_{1:n}}(X_{n+1}))$$

Data from Poelwijk *et al.* (2019), *Nat. Commun.*
Design algorithm from Biswas *et al.* (2021), *Nat. Methods*

Uncertainty quantification can guide design algorithm selection

Design algorithm: sample X_{n+1} from $\tilde{p}_X(X_{n+1}) \propto \exp(\lambda \cdot \mu_{1:n}(X_{n+1}))$

λ: inverse temperature $\mu_{1:n}$: trained regression model

Greater λ means both (i) higher predicted fluorescence and (ii) higher predictive uncertainty.

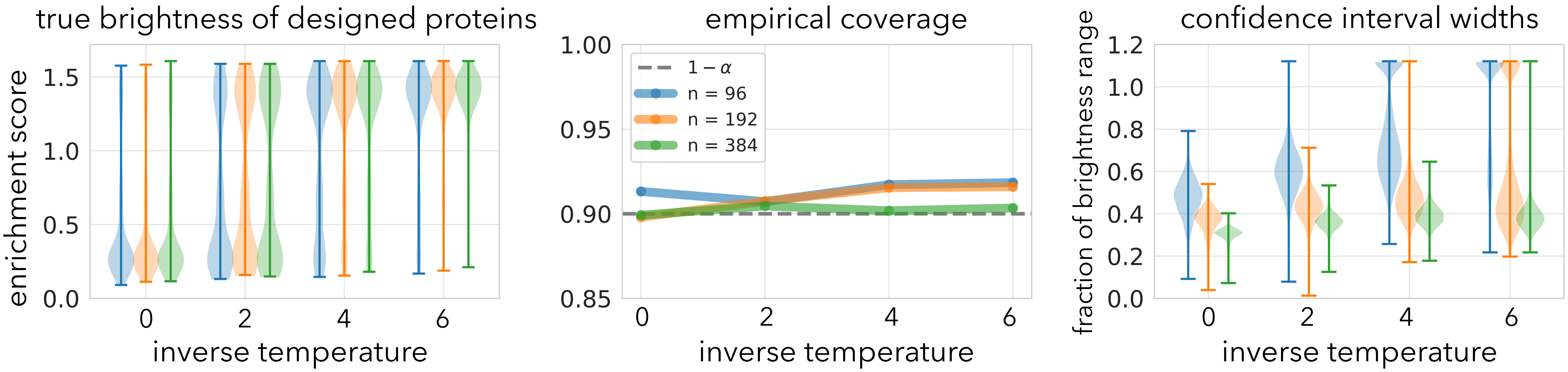
Data from Poelwijk *et al.* (2019), *Nat. Commun.*
Design algorithm from Biswas *et al.* (2021), *Nat. Methods*

Uncertainty quantification can guide design algorithm selection

Design algorithm: sample X_{n+1} from $\tilde{p}_X(X_{n+1}) \propto \exp(\lambda \cdot \mu_{1:n}(X_{n+1}))$

inverse temperature   trained regression model

Greater λ means both (i) higher predicted fluorescence and (ii) higher predictive uncertainty.



Data from Poelwijk *et al.* (2019), *Nat. Commun.*
Design algorithm from Biswas *et al.* (2021), *Nat. Methods*

Uncertainty quantification can guide design algorithm selection

Design algorithm: sample X_{n+1} from $\tilde{p}_X(X_{n+1}) \propto \exp(\lambda \cdot \mu_{1:n}(X_{n+1}))$

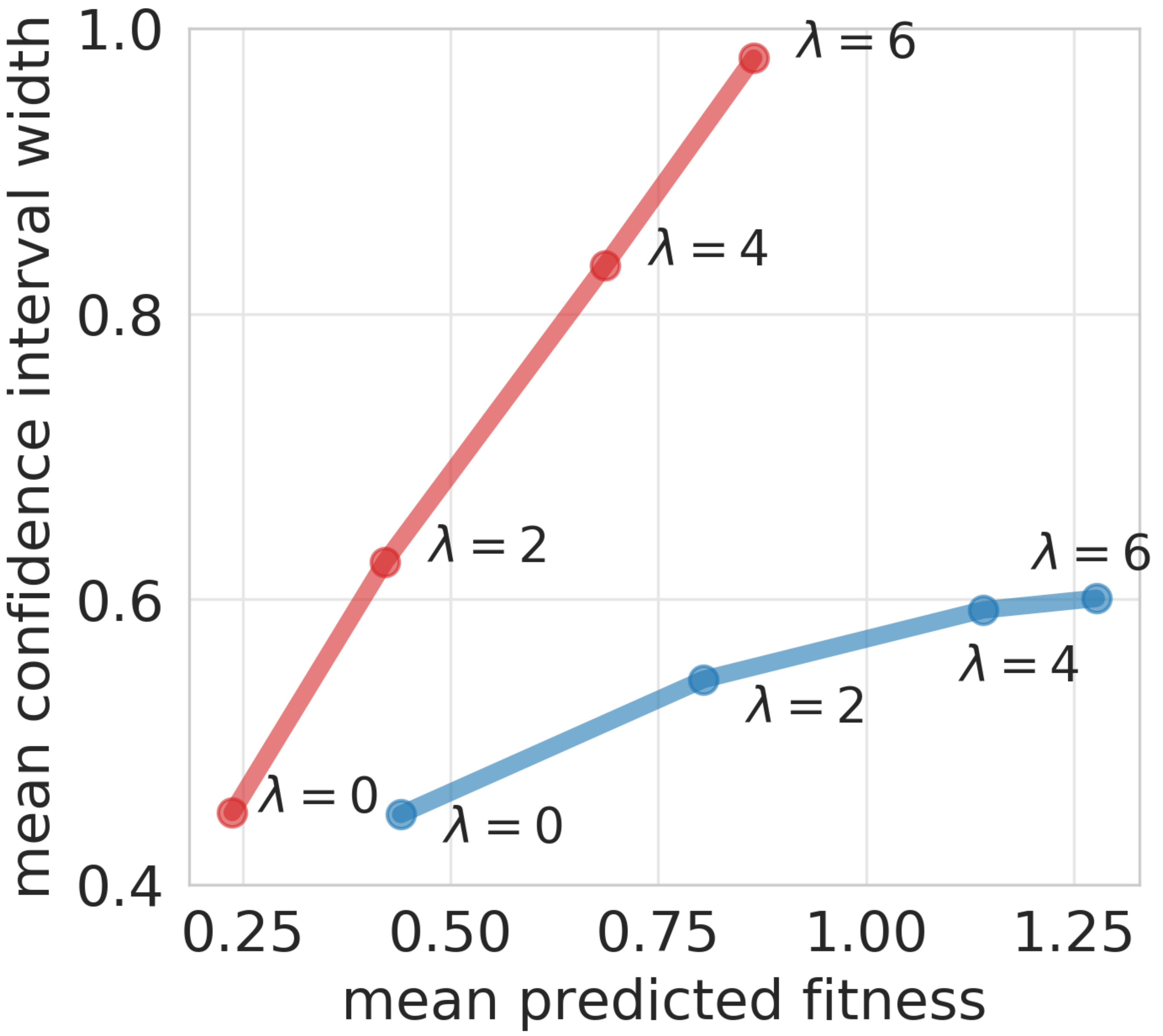
inverse temperature   trained regression model

Greater λ means both (i) higher predicted fluorescence and (ii) higher predictive uncertainty.

How should we set λ? We can use confidence interval width to navigate trade-off between (i) and (ii).

More details and examples:

Fannjiang, Bates, Angelopoulos, Listgarten, & Jordan. "Conformal prediction for the design problem", arXiv:2202.03613, in submission.



Data from Poelwijk *et al.* (2019), *Nat. Commun.*
Design algorithm from Biswas *et al.* (2021), *Nat. Methods*

thank you!

NSF

BAIR

BERKELEY ARTIFICIAL INTELLIGENCE RESEARCH

Berkeley
UNIVERSITY OF CALIFORNIA