[1] Johns Hopkins University

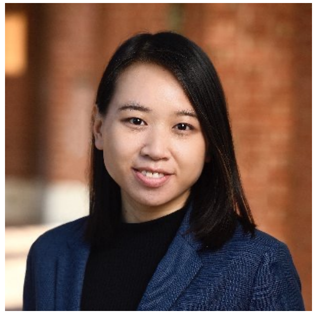
ICML
International Conference
On Machine Learning

[2] Genentech
A Member of the Roche Group

[2] Prescient Design
A Genentech Accelerator

Conformal Validity Guarantees Exist for Any Data Distribution (and How to Find Them)



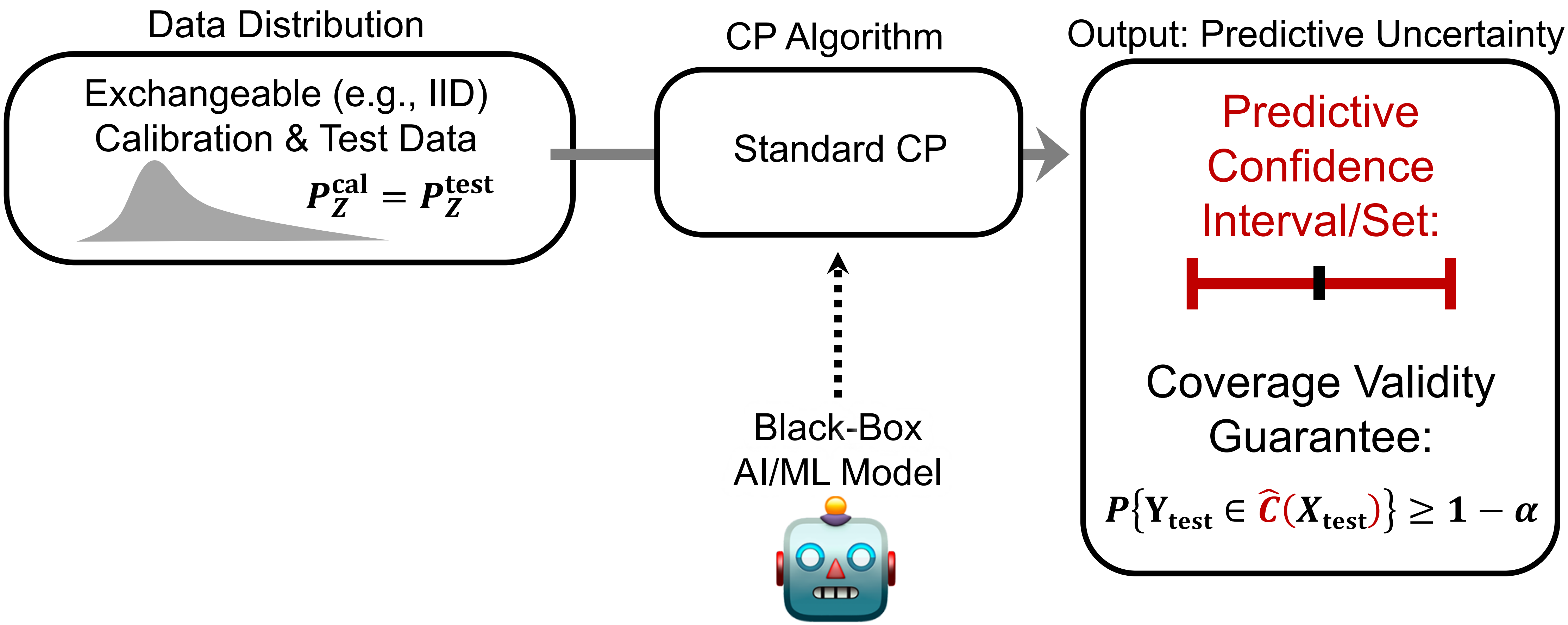



Drew Prinster*[1] Samuel Stanton*[2] Anqi Liu[1] Suchi Saria[1]

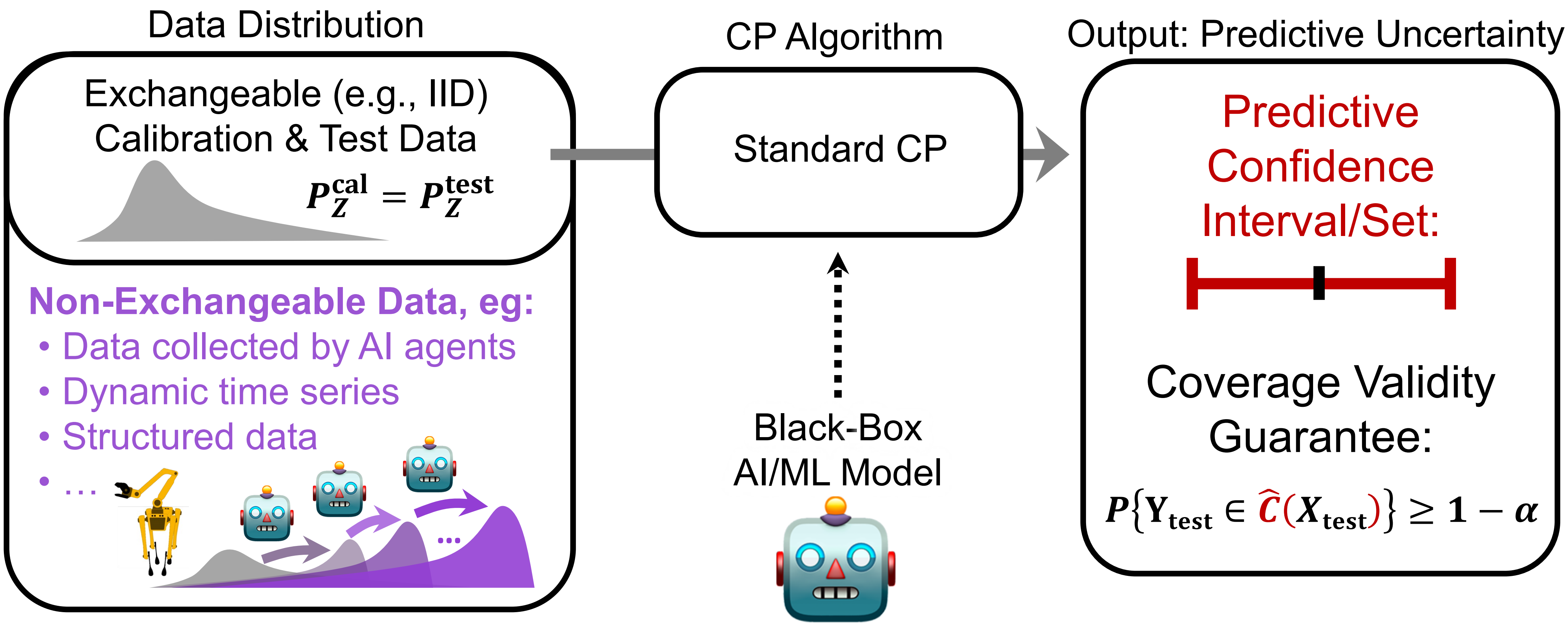
Roadmap

- **Introduction:** AI Uncertainty Quantification via Conformal Prediction
- **Key Background**
- **Theory and Method Contributions**
- **Experiments**
- **Discussion**

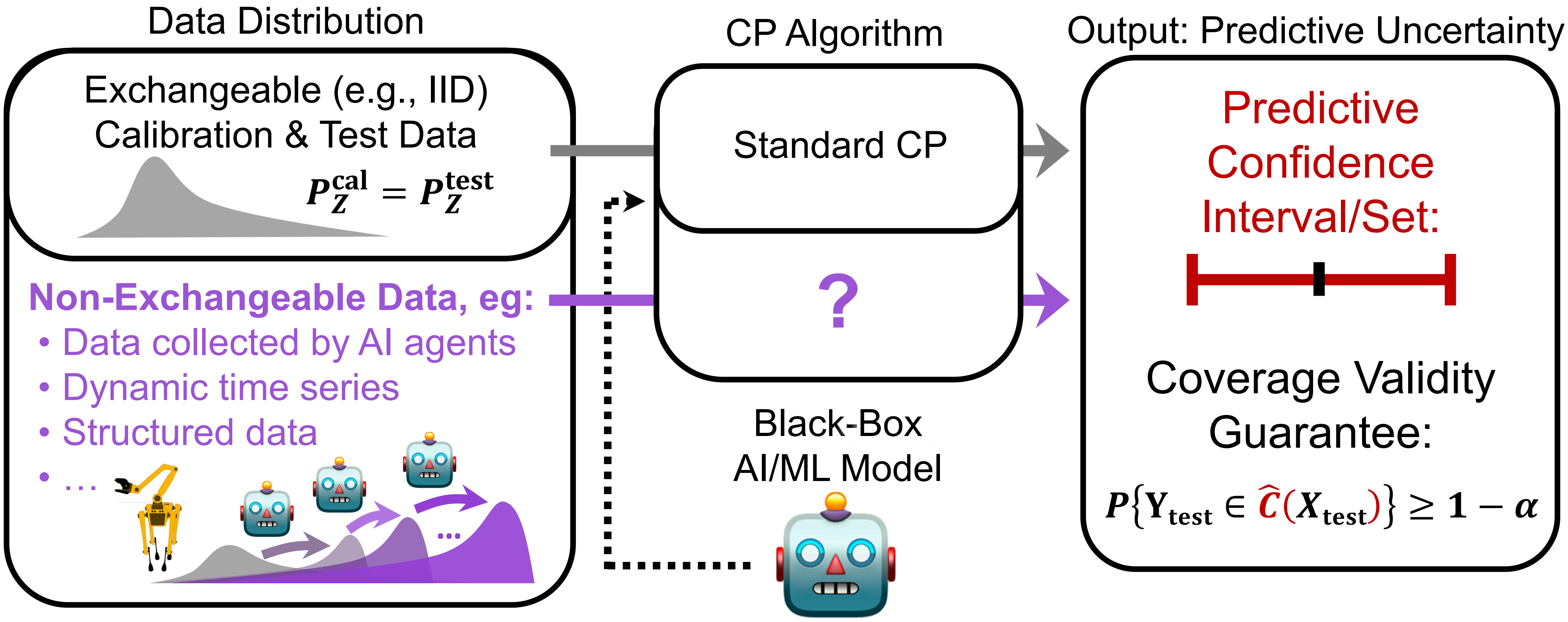
Introduction: AI Uncertainty Quantification via Conformal Prediction (CP)



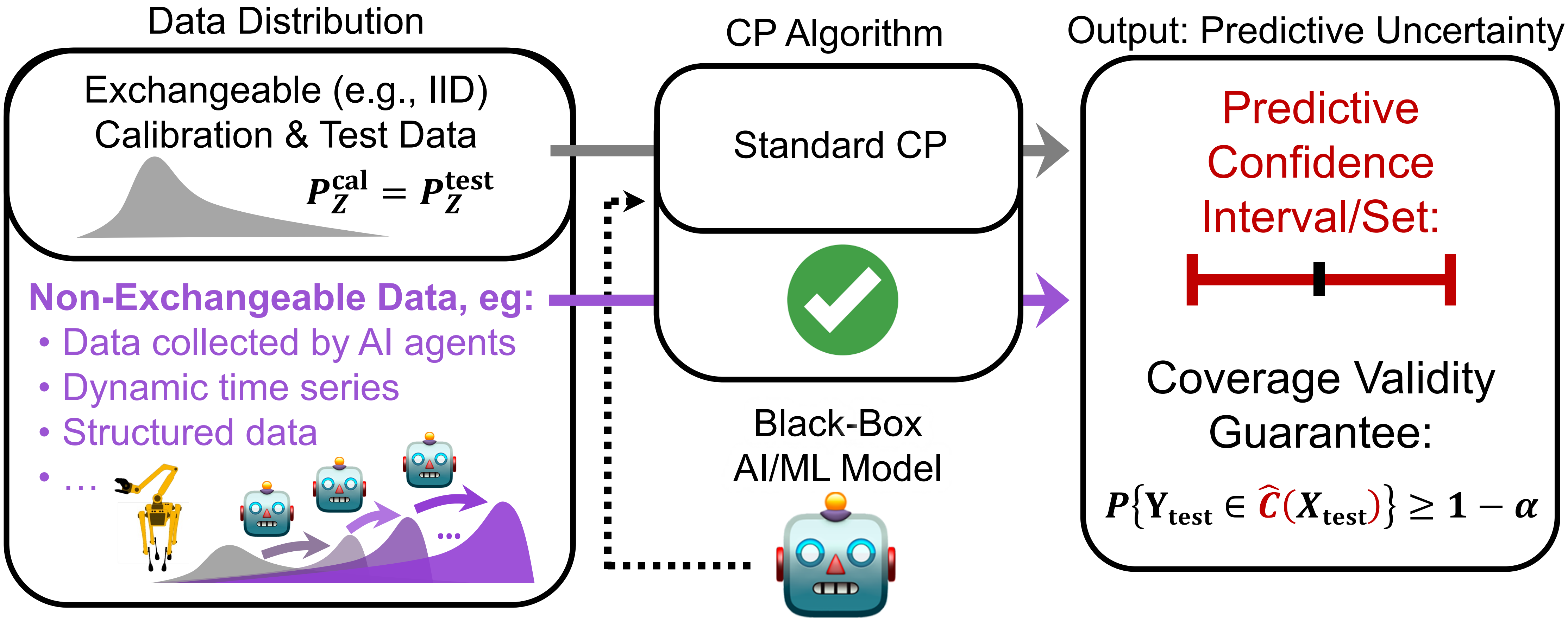
Introduction: AI Uncertainty Quantification via Conformal Prediction (CP)



Introduction: AI Uncertainty Quantification via Conformal Prediction (CP)

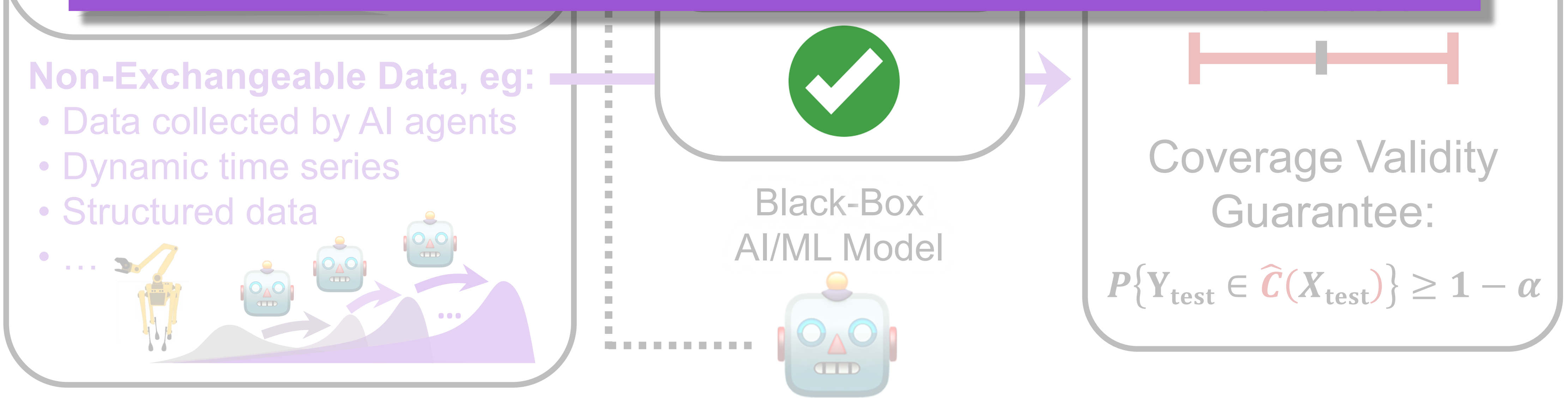


Introduction: AI Uncertainty Quantification via Conformal Prediction (CP)



Introduction: AI Uncertainty Quantification via Conformal Prediction (CP)

- Theory: We prove that **conformal prediction** (with coverage validity guarantees) **can theoretically extend to *any***, potentially **non-exchangeable** joint **data distribution**. This builds on analysis from Tibshirani et al. (2019).



Introduction: AI Uncertainty Quantification via Conformal Prediction (CP)

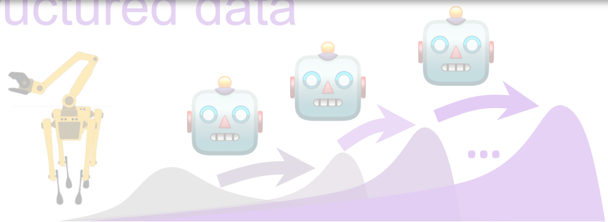
Data Distribution

CP Algorithm

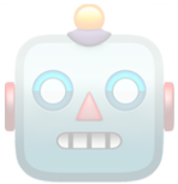
Output: Predictive Uncertainty

- Theory: We prove that **conformal prediction** (with coverage validity guarantees) **can theoretically extend to *any***, potentially **non-exchangeable** joint **data distribution**. This builds on analysis from Tibshirani et al. (2019).
- Practical:
    1. We outline **“how to find”** valid CP guarantees for any data distribution.

- Structured data
- …



Black-Box AI/ML Model

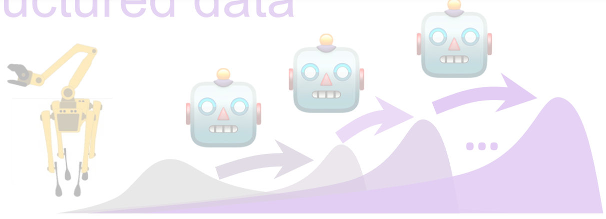


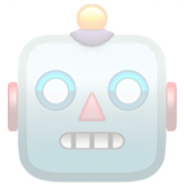
Guarantee:

$$P\{Y_{\text{test}} \in \widehat{C}(X_{\text{test}})\} \geq 1 - \alpha$$

Introduction: AI Uncertainty Quantification via Conformal Prediction (CP)

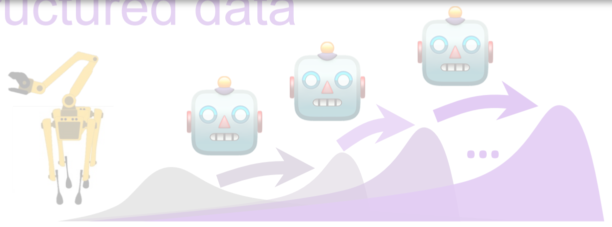
- Theory: We prove that **conformal prediction** (with coverage validity guarantees) **can theoretically extend to *any***, potentially **non-exchangeable** joint **data distribution**. This builds on analysis from Tibshirani et al. (2019).
- Practical:
  1. We outline **"how to find"** valid CP guarantees for any data distribution.
  2. We derive practical CP algorithms for **AI agent feedback-loop shifts**.

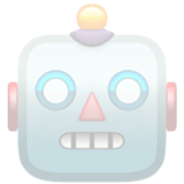




Introduction: AI Uncertainty Quantification via Conformal Prediction (CP)

- Theory: We prove that **conformal prediction** (with coverage validity guarantees) **can theoretically extend to *any***, potentially **non-exchangeable** joint **data distribution**. This builds on analysis from Tibshirani et al. (2019).
- Practical:
  1. We outline **"how to find"** valid CP guarantees for any data distribution.
  2. We derive practical CP algorithms for **AI agent feedback-loop shifts**.
  3. Our extensions enable **adaptive AI exploration** with **sharp intervals**.





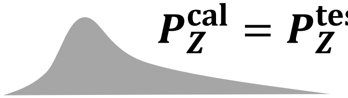
Roadmap

- **Introduction**
- **Key Background:** Weighted Conformal Prediction for Covariate Shifts
- **Theory and Method Contributions**
- **Experiments**
- **Discussion**

Background: Weighted CP for Covariate Shifts

Vovk et al.
(2005)

Exchangeable Data
≈ No Shift



$P_Z^{\text{cal}} = P_Z^{\text{test}}$

E.g., IID

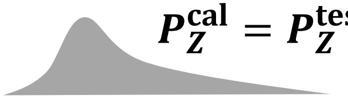
P_Z invariant.

Background: Weighted CP for Covariate Shifts

Vovk et al. (2005)

Exchangeable Data

≈ No Shift



$P_Z^{\text{cal}} = P_Z^{\text{test}}$

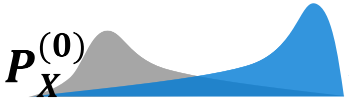
E.g., IID

P_Z invariant.

Tibshirani et al. (2019)

Weighted Exchangeable Data

≈ Independent Shifts



$P_X^{(0)}$ $P_X^{(1)}$

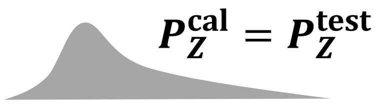
E.g., Standard Covariate Shift (SCS)

$P_{Y|X}$ invariant.

Background: Weighted CP for Covariate Shifts

Vovk et al. (2005)

Exchangeable Data ≈ No Shift



E.g., IID

P_Z invariant.

Tibshirani et al. (2019)

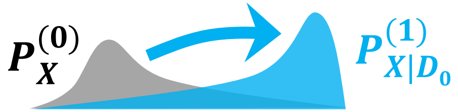
Weighted Exchangeable Data ≈ Independent Shifts

E.g., Standard Covariate Shift (SCS)

$P_{Y|X}$ invariant.

Fannjiang et al. (2022)

Pseudo-Exchangeable Data ≈ Limited *Dependent* Shifts



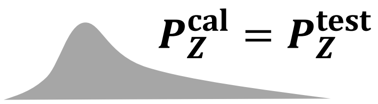
E.g., ***Feedback*** Covariate Shift (FCS)

$P_{Y|X}$ invariant.

Background: Weighted CP for Covariate Shifts

Vovk et al. (2005)

Exchangeable Data ≈ No Shift



E.g., IID

P_Z invariant.

Tibshirani et al. (2019)

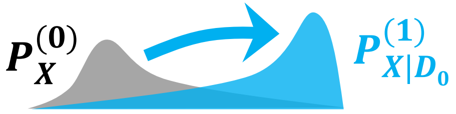
Weighted Exchangeable Data ≈ Independent Shifts

E.g., Standard Covariate Shift (SCS)

$P_{Y|X}$ invariant.

Fannjiang et al. (2022)

Pseudo-Exchangeable Data ≈ Limited *Dependent* Shifts

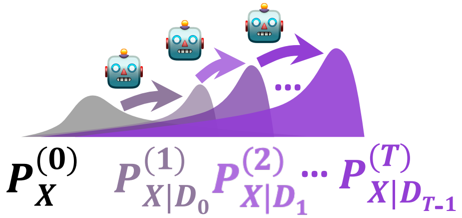


E.g., ***Feedback*** Covariate Shift (FCS)

$P_{Y|X}$ invariant.

Our Paper

Any Data Distribution ≈ Any Shifts



E.g., ***Multistep*** Feedback Covariate Shift (MFCS)*

*MFCS is similar to a setting in Nair & Janson (2023)

$P_{Y|X}$ invariant.

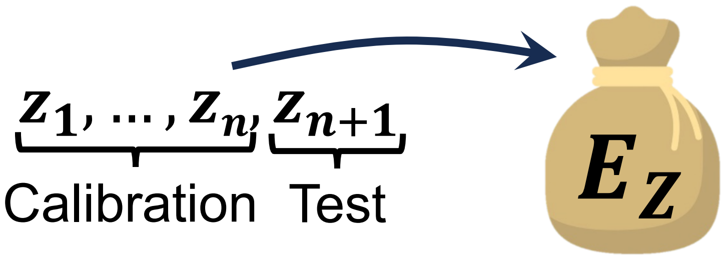
Roadmap

- Introduction
- Key Background
- **Theory and Method Contributions:**
  - Revisiting Tibshirani et al. (2019)'s Alternate CP Proof
  - Key Insight
  - Main Result: Conformal Validity Guarantees Exist for Any Data Distribution
  - "How to Find Them"
- Experiments
- Discussion

Revisiting Tibshirani et al. (2019)'s Proof

- Collect *Bag* (e.g., Set) of Data/Scores:
  Condition on event $\{\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{n+1}\} = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{n+1}\}$

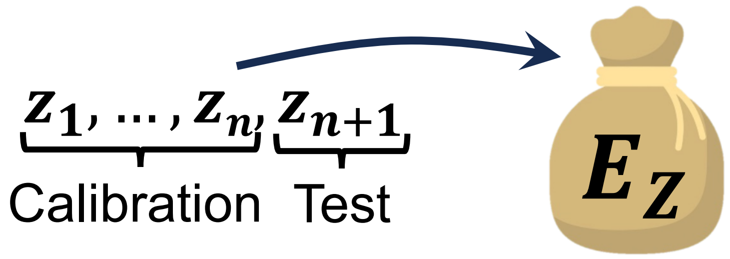
  Note: We know $\boldsymbol{Z}_i$ takes a value in $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{n+1}\}$ but *not which* one (same for scores).



Revisiting Tibshirani et al. (2019)'s Proof

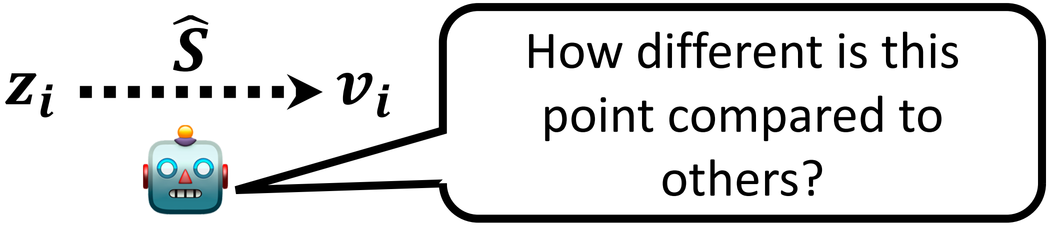
- Collect *Bag* (e.g., Set) of Data/Scores:
  Condition on event $\{\boldsymbol{Z}_1, \dots, \boldsymbol{Z}_{n+1}\} = \{\boldsymbol{z}_1, \dots, \boldsymbol{z}_{n+1}\}$

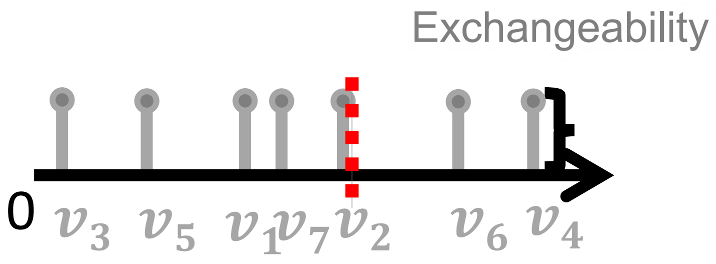
  Note: We know $\boldsymbol{Z}_i$ takes a value in $\{\boldsymbol{z}_1, \dots, \boldsymbol{z}_{n+1}\}$ but *not which* one (same for scores).



- Score Datapoints:
  Compute "nonconformity" scores.

  E.g., Residual scores: $\widehat{\boldsymbol{S}}(x, y) = |y - \widehat{\boldsymbol{\mu}}(x)|$

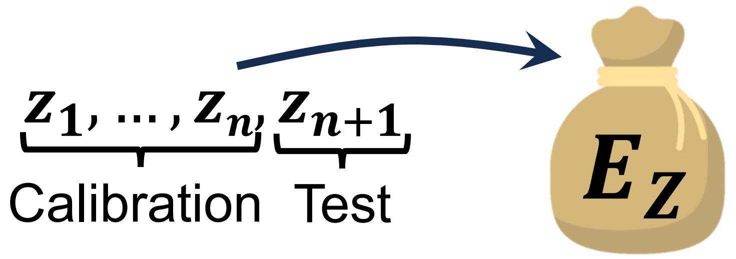




Revisiting Tibshirani et al. (2019)'s Proof

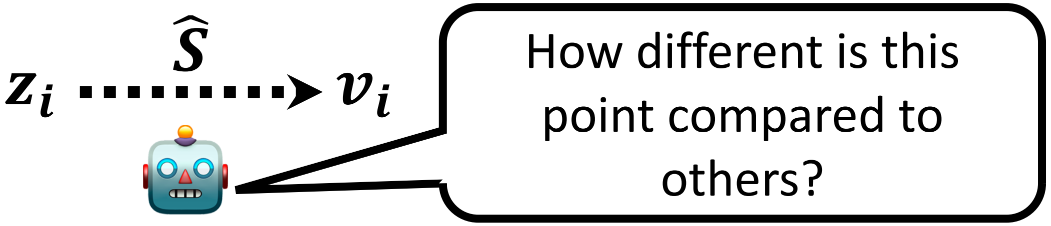
- <u>Collect *Bag* (e.g., Set) of Data/Scores:</u>
  Condition on event $\{\boldsymbol{Z_1}, \ldots, \boldsymbol{Z_{n+1}}\} = \{\boldsymbol{z_1}, \ldots, \boldsymbol{z_{n+1}}\}$

  Note: We know $\boldsymbol{Z_i}$ takes a value in $\{\boldsymbol{z_1}, \ldots, \boldsymbol{z_{n+1}}\}$ but *not which* one (same for scores).



- <u>Score Datapoints:</u>
  Compute "nonconformity" scores.

  E.g., Residual scores: $\widehat{\boldsymbol{S}}(x, y) = |y - \widehat{\boldsymbol{\mu}}(x)|$



- <u>Compute CP Weights: Examine Probability $\boldsymbol{V_{n+1}} = \boldsymbol{v_i}$ & Simplify with (Weighted) Exchangeability:</u>

$$\mathbb{P}\{V_{n+1} = v_i \mid \boldsymbol{E_Z}\} \overset{*}{=} \frac{\sum_{\sigma:\sigma(n+1)=i} f(z_{\sigma(1)}, \ldots, z_{\sigma(n+1)})}{\sum_{\sigma} f(z_{\sigma(1)}, \ldots, z_{\sigma(n+1)})} \quad \textbf{(1)}$$

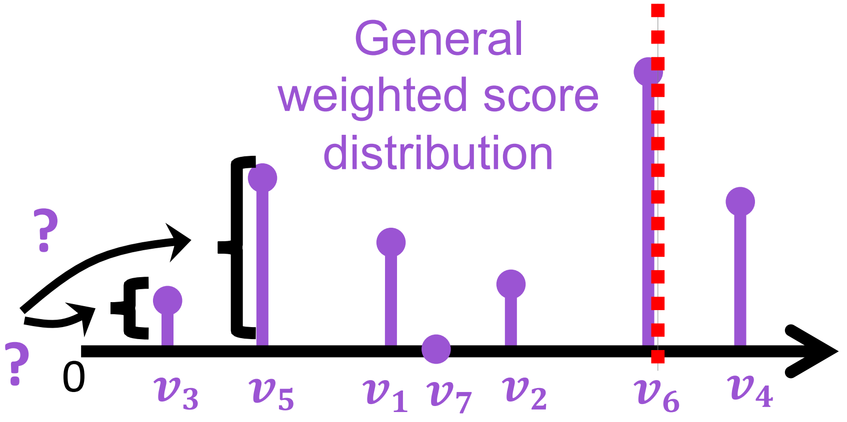
Key Insight: Exchangeability Conditions are *Practical*, not Theoretically Necessary

For Intuition: We can derive Eq. (1) without any assumptions on the joint PDF $\boldsymbol{f}$:

$$\mathbb{P}\{V_{n+1} = v_i \mid E_Z\} = \ ?$$

*

$$= \frac{\sum_{\sigma:\sigma(n+1)=i} f\big(z_{\sigma(1)}, \dots, z_{\sigma(n+1)}\big)}{\sum_{\sigma} f\big(z_{\sigma(1)}, \dots, z_{\sigma(n+1)}\big)}$$

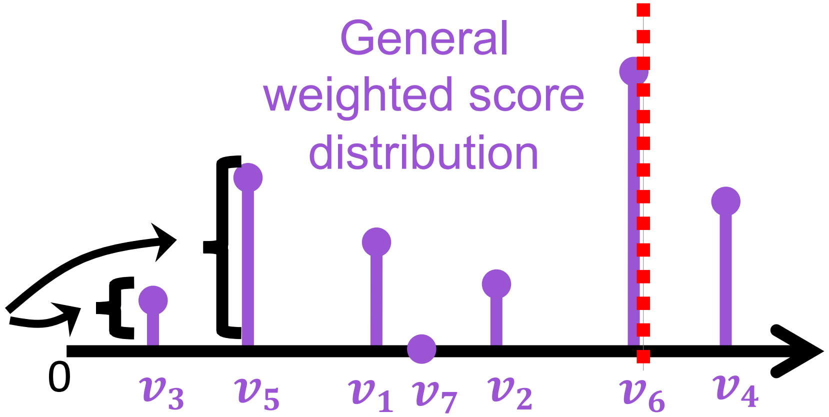


Key Insight: Exchangeability Conditions are *Practical*, not Theoretically Necessary

<u>For Intuition:</u> We can derive Eq. (1) without any assumptions on the joint PDF $\boldsymbol{f}$:

$$\mathbb{P}\{V_{n+1} = v_i \mid E_Z\} = \mathbb{P}\{Z_{n+1} = z_i \mid E_Z\}$$

*



Key Insight: Exchangeability Conditions are *Practical*, not Theoretically Necessary

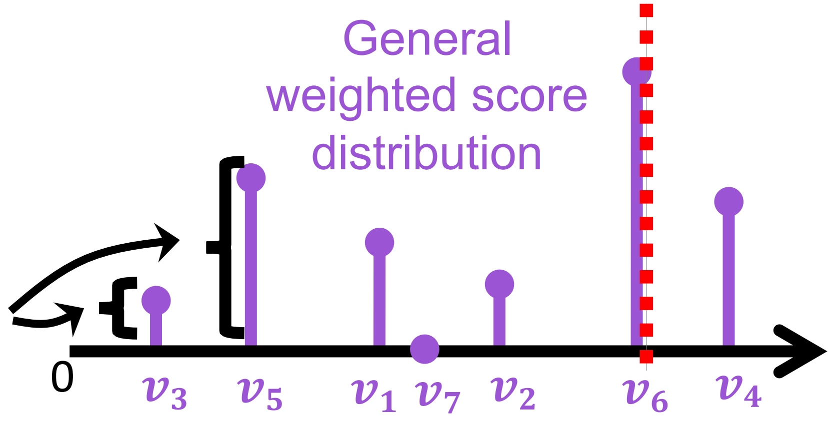
For Intuition: We can derive Eq. (1) without any assumptions on the joint PDF $\boldsymbol{f}$:

$$\mathbb{P}\{V_{n+1} = v_i \mid E_Z\} = \mathbb{P}\{Z_{n+1} = z_i \mid E_Z\}$$

Conditional probability def.

$$= \frac{p\{Z_{n+1} = z_i, E_Z\}}{p\{E_Z\}}$$

*



Key Insight: Exchangeability Conditions are *Practical*, not Theoretically Necessary

<u>For Intuition:</u> We can derive Eq. (1) without any assumptions on the joint PDF $\boldsymbol{f}$:

$$\mathbb{P}\{V_{n+1} = v_i \mid E_Z\} = \mathbb{P}\{Z_{n+1} = z_i \mid E_Z\}$$

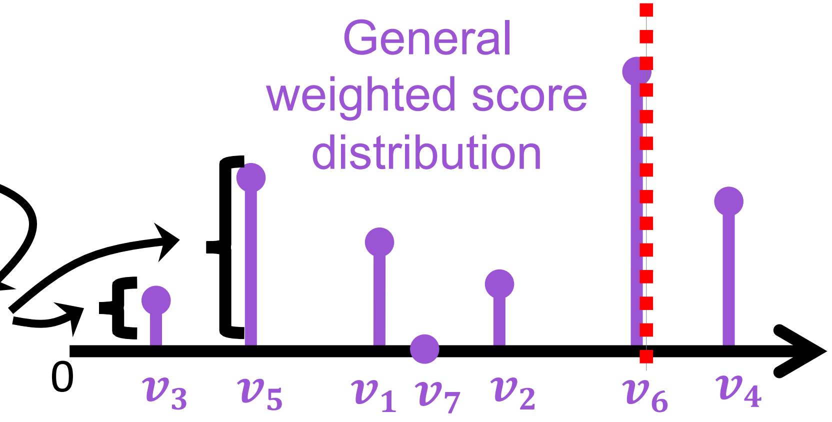
Conditional probability def.

$$= \frac{p\{Z_{n+1} = z_i, E_Z\}}{p\{E_Z\}}$$

LOTP over σ

*

$$= \frac{\sum_{\sigma:\sigma(n+1)=i} f\big(z_{\sigma(1)}, \dots, z_{\sigma(n+1)}\big)}{\sum_{\sigma} f\big(z_{\sigma(1)}, \dots, z_{\sigma(n+1)}\big)}$$



Key Insight: Exchangeability Conditions are *Practical*, not Theoretically Necessary

<u>For Intuition:</u> We can derive Eq. (1) without any assumptions on the joint PDF $\boldsymbol{f}$:

$$\mathbb{P}\{V_{n+1} = v_i \mid E_Z\} = \mathbb{P}\{Z_{n+1} = z_i \mid E_Z\}$$

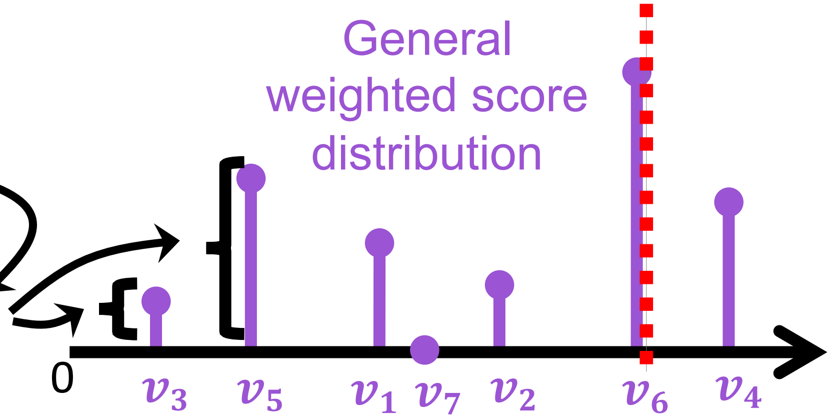
Conditional probability def.

$$= \frac{p\{Z_{n+1} = z_i, E_Z\}}{p\{E_Z\}}$$

LOTP over σ

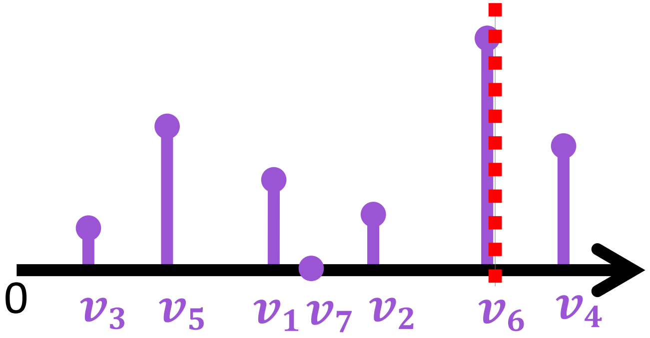
*

$$= \frac{\sum_{\sigma:\sigma(n+1)=i} f(z_{\sigma(1)}, \ldots, z_{\sigma(n+1)})}{\sum_{\sigma} f(z_{\sigma(1)}, \ldots, z_{\sigma(n+1)})}$$



The *only role* of exchangeability assumptions is to simplify these general weights into a tractable form!

Main Theorem: Conformal Validity Guarantees Exist for Any Data Distribution



General Conformal Prediction Set:

$$\widehat{C}_n(x) = \left\{ y \in \mathbb{R} : V_{n+1}^{(x,y)} \leq Q_{1-\alpha}\left(\sum_{i=1}^{n} \underbrace{\mathbb{P}_{n+1}\{Z_i | E_z\}}_{\text{general weights for calibration data}} \delta_{V_i^{(x,y)}} + \underbrace{\mathbb{P}_{n+1}\{Z_{n+1} | E_z\}}_{\text{general weight for test point}} \delta_{\infty} \right) \right\}$$

general weights for calibration data

general weight for test point

Satisfies:

$$\mathbb{P}\{Y_{n+1} \in \widehat{C}_n(X_{n+1})\} \geq 1 - \alpha.$$

“How to Find” CP Validity Guarantees

1. List assumptions on f (if any). E.g., for MFCS, $\boldsymbol{X}$ changes depending on past but $\boldsymbol{Y} \mid \boldsymbol{X}$ does not.

“How to Find” CP Validity Guarantees

1. List assumptions on ***f*** (if any). E.g., for MFCS, ***X*** changes depending on past but ***Y* | *X*** does not.

2. Factorize **f**: Factorize ***f*** into “dynamic” and “invariant” factors (w.r.t. permutations), using standard probability rules and Step 1 assumptions. E.g., for MFCS:

$$f(z_1, \ldots, z_{n+t}) = \prod_{j=1}^{n+t} \Big[\underbrace{p(x_j \mid z_1, \ldots, z_{j-1})}_{\text{Time-dependent factors}} \Big] \cdot \underbrace{\prod_{j=1}^{n+t} \Big[p(y_j | x_j) \Big]}_{\text{Time-invariant factor}}$$

“How to Find” CP Validity Guarantees

1. <u>List assumptions on ***f*** (if any).</u> E.g., for MFCS, ***X*** changes depending on past but ***Y* | *X*** does not.

2. <u>Factorize **f**:</u> Factorize ***f*** into “dynamic” and “invariant” factors (w.r.t. permutations), using standard probability rules and Step 1 assumptions. E.g., for MFCS:

$$f(z_1, \ldots, z_{n+t}) = \underbrace{\prod_{j=1}^{n+t} \Big[\underbrace{p(x_j \mid z_1, \ldots, z_{j-1})}_{\text{Time-dependent factors}} \Big]}\cdot \underbrace{\prod_{j=1}^{n+t} \big[p(y_j | x_j) \big]}_{\text{Time-invariant factor}}$$

3. <u>Compute or estimate weights:</u> Plug factorized ***f*** from Step 2 into Eq. (1). E.g., for MFCS:

$$\mathbb{P}\{Z_{n+t} = z_i \mid E_Z^{(t)}\} = \frac{\sum_{\sigma:\sigma(n+t)=i} f(z_{\sigma(1)}, \ldots, z_{\sigma(n+t)})}{\sum_{\sigma} f(z_{\sigma(1)}, \ldots, z_{\sigma(n+t)})} = \frac{\sum_{\sigma:\sigma(n+t)=i} \prod_{j=1}^{n+t} p(x_{\sigma(j)} \mid z_{\sigma(1)}, \ldots, z_{\sigma(j-1)})}{\sum_{\sigma} \prod_{j=1}^{n+t} p(x_{\sigma(j)} \mid z_{\sigma(1)}, \ldots, z_{\sigma(j-1)})}$$

"How to Find" CP Validity Guarantees

1. List assumptions on f (if any). E.g., for MFCS, $\boldsymbol{X}$ changes depending on past but $\boldsymbol{Y} \mid \boldsymbol{X}$ does not.

2. Factorize f: Factorize f into "dynamic" and "invariant" factors (w.r.t. permutations), using standard probability rules and Step 1 assumptions. E.g., for MFCS:

$$f(z_1, \ldots, z_{n+t}) = \prod_{j=1}^{n+t} \Big[\underbrace{p(x_j \mid z_1, \ldots, z_{j-1})}_{\text{Time-dependent factors}} \Big] \cdot \underbrace{\prod_{j=1}^{n+t} \Big[p(y_j | x_j) \Big]}_{\text{Time-invariant factor}}$$

3. Compute or estimate weights: Plug factorized f from Step 2 into Eq. (1). E.g., for MFCS:

$$\mathbb{P}\{Z_{n+t} = z_i \mid E_Z^{(t)}\} = \frac{\sum_{\sigma:\sigma(n+t)=i} f(z_{\sigma(1)}, \ldots, z_{\sigma(n+t)})}{\sum_{\sigma} f(z_{\sigma(1)}, \ldots, z_{\sigma(n+t)})} = \frac{\sum_{\sigma:\sigma(n+t)=i} \prod_{j=1}^{n+t} p(x_{\sigma(j)} \mid z_{\sigma(1)}, \ldots, z_{\sigma(j-1)})}{\sum_{\sigma} \prod_{j=1}^{n+t} p(x_{\sigma(j)} \mid z_{\sigma(1)}, \ldots, z_{\sigma(j-1)})}$$

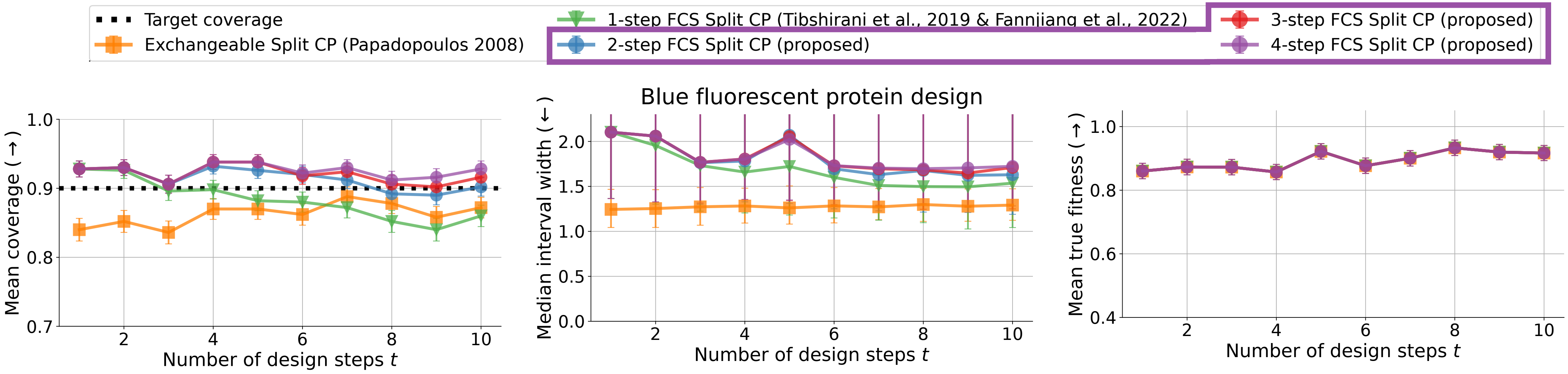
$$\approx \frac{\sum_{\sigma:\sigma(n+t)=i} \prod_{j=n+t+1-d}^{n+t} p(x_{\sigma(j)} \mid z_{\sigma(1)}, \ldots, z_{\sigma(j-1)})}{\sum_{\sigma} \prod_{j=n+t+1-d}^{n+t} p(x_{\sigma(j)} \mid z_{\sigma(1)}, \ldots, z_{\sigma(j-1)})}$$

Roadmap

- **Introduction**
- **Key Background**
- **Theory and Method Contributions**
- **Experiments:**
  - Black-Box Optimization (Multi-Round Synthetic Protein Design)
  - Adaptive AI Exploration with Sharp Intervals
- **Discussion**

Multi-Round Protein Design Experiments

We induce MFCS by an ML agent actively selecting each point with query functions $p(x \mid Z_{\text{train}}^{(t)}) \propto \exp(\lambda \cdot u_t(x))$, for utility function u_t.

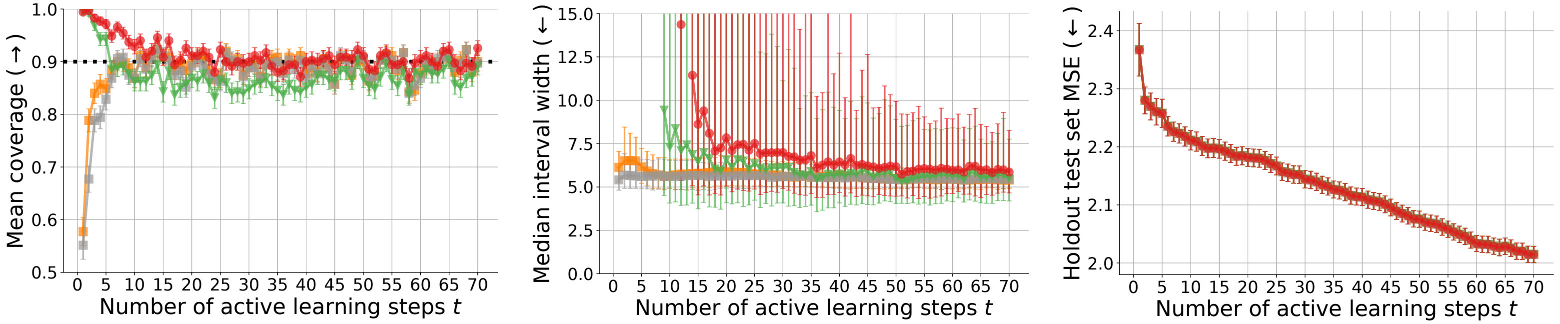


Our proposed Split CP MFCS methods allow for **complex models like neural networks** and **maintain coverage** even at later design steps ***t***.

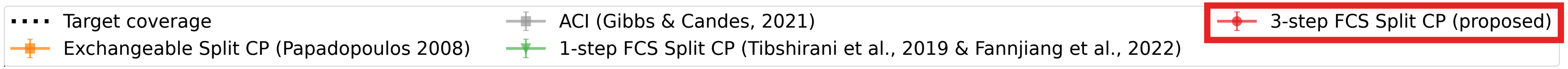
Active Learning Experiments

- Target coverage
- Exchangeable Split CP (Papadopoulos 2008)
- ACI (Gibbs & Candes, 2021)
- 1-step FCS Split CP (Tibshirani et al., 2019 & Fannjiang et al., 2022)
- 3-step FCS Split CP (proposed)

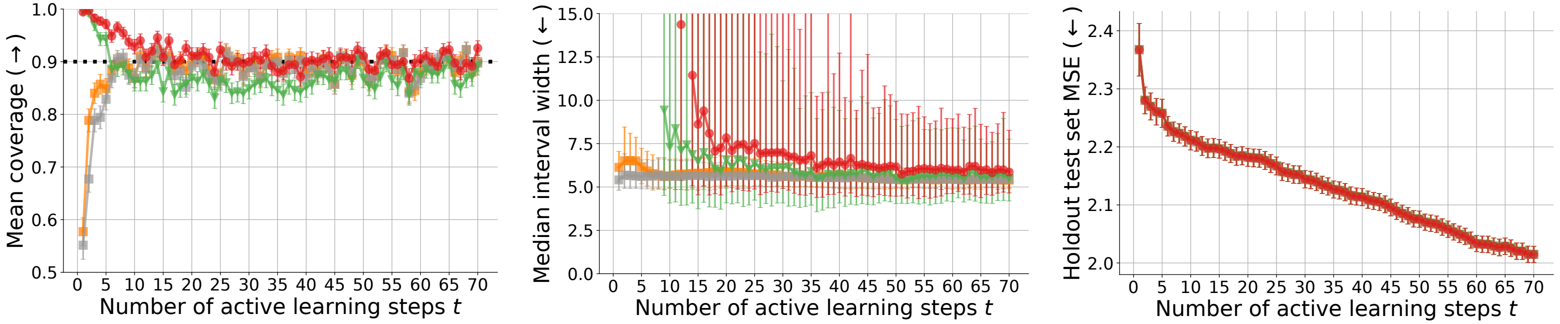
Unbounded Query Function (meps data)



Adaptive Exploration with Sharp Intervals

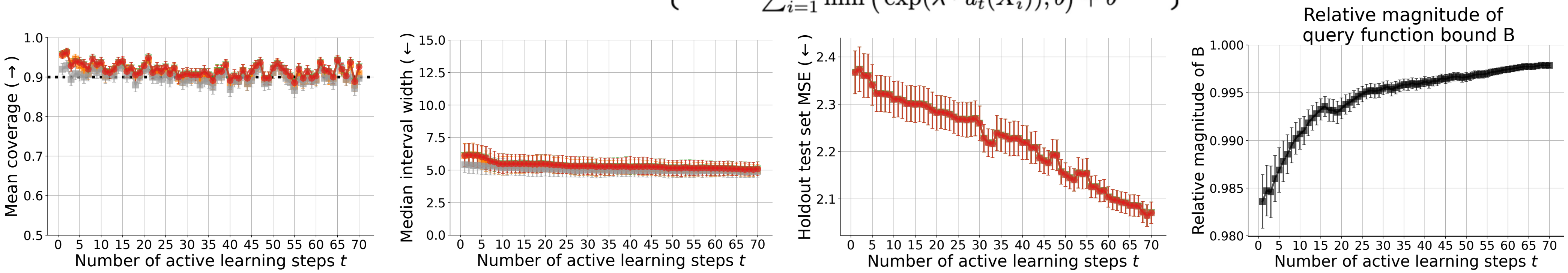


Unbounded Query Function (meps data)



Proposed Bounded Query Function (meps data)

$$p(x \mid Z_1, \ldots, Z_{j-1}) \propto \min\left(\exp(\lambda \cdot u_t(x)), B\right), \quad \text{where } B = \sup\left\{ b > 0 : \frac{b}{\sum_{i=1}^{n} \min\left(\exp(\lambda \cdot u_t(X_i)), b\right) + b} < \alpha \right\}$$

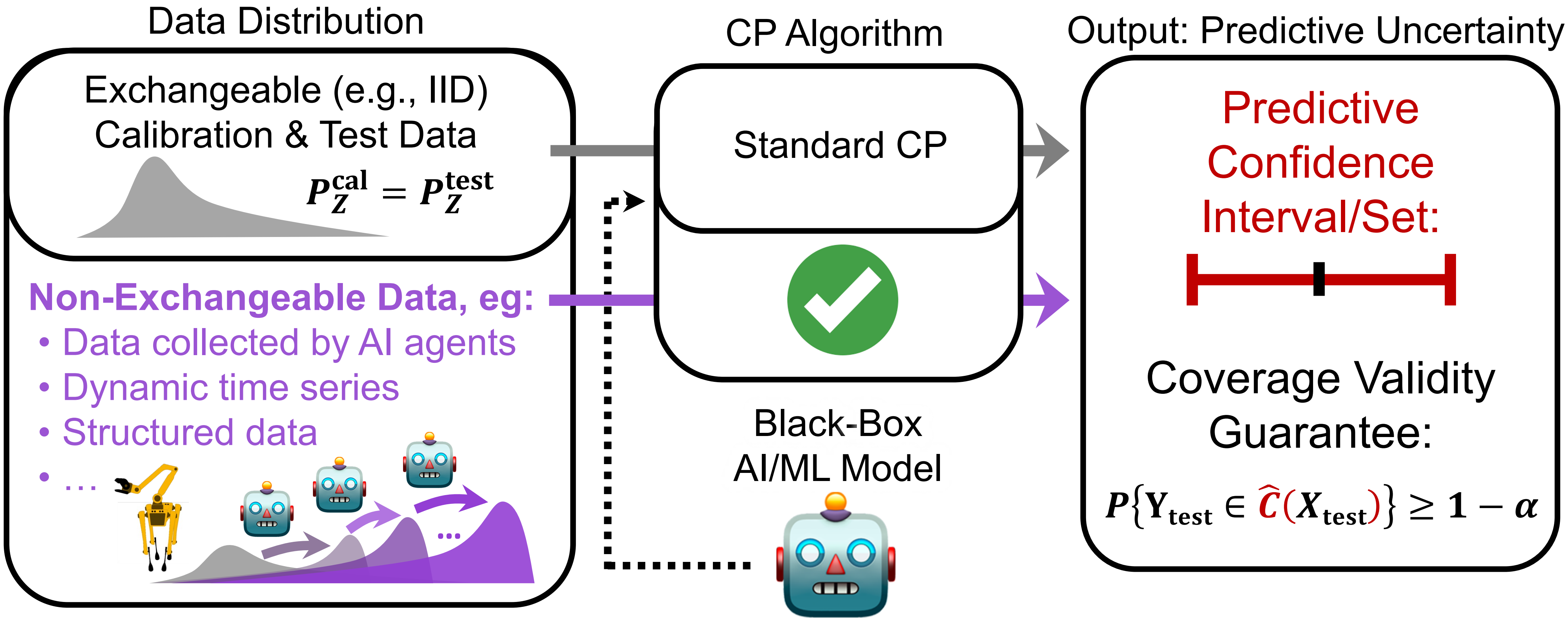


The **bounded AI/ML agent initially "explores slowly,"** until it has seen enough data!

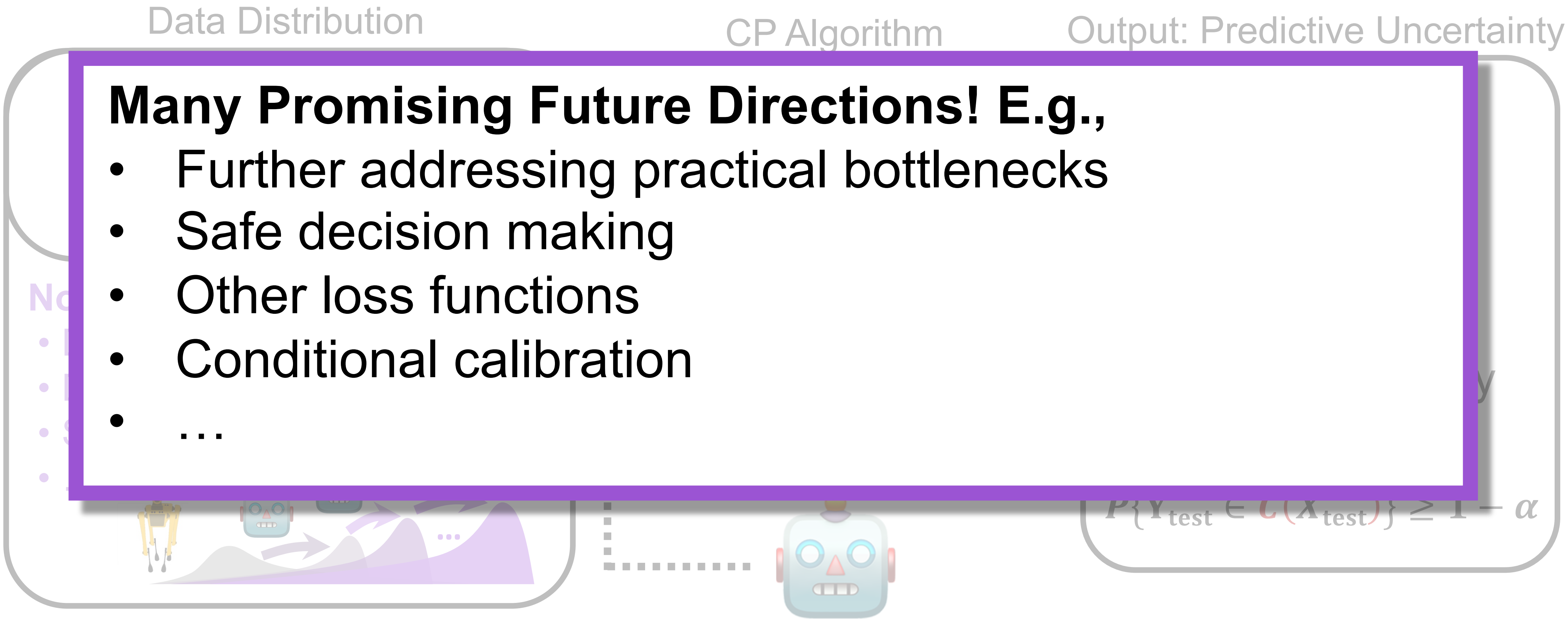
Roadmap

- **Introduction**
- **Key Background**
- **Theory and Method Contributions**
- **Experiments:**
- **Discussion**

Recap and Future Directions



Recap and Future Directions



Many Promising Future Directions! E.g.,

- Further addressing practical bottlenecks
- Safe decision making
- Other loss functions
- Conditional calibration
- …

References

Fannjiang, C., Bates, S., Angelopoulos, A. N., Listgarten, J., & Jordan, M. I. (2022). Conformal prediction under feedback covariate shift for biomolecular design. *PNAS*.

Nair, Y., & Janson, L. (2023). Randomization tests for adaptively collected data. *arXiv*.

Papadopoulos, H., Proedrou, K., Vovk, V., & Gammerman, A. (2002). Inductive confidence machines for regression. *ECML.*

Poelwijk, et al. (2019). Learning the pattern of epistasis linking genotype and phenotype in a protein. *Nature communications*.

Prinster, D., Liu, A., & Saria, S. (2022). JAWS: Auditing predictive uncertainty under covariate shift. *NeurIPS*.

Prinster, D., Saria, S., & Liu, A. (2023). JAWS-x: Addressing efficiency bottlenecks of conformal prediction under standard and feedback covariate shift. *ICML*.

Stanton, S., Maddox, W., & Wilson, A. G. (2023, April). Bayesian optimization with conformal prediction sets. *AIStats*.

Tibshirani, R. J., Foygel Barber, R., Candes, E., & Ramdas, A. (2019). Conformal prediction under covariate shift. *NeurIPS*, 32

Vovk, V., Gammerman, A., & Shafer, G. (2005). *Algorithmic learning in a random world*. Springer Science & Business Media.

Thank you!

NSF

Paper QR

GORDON AND BETTY MOORE FOUNDATION