Carnegie Mellon University

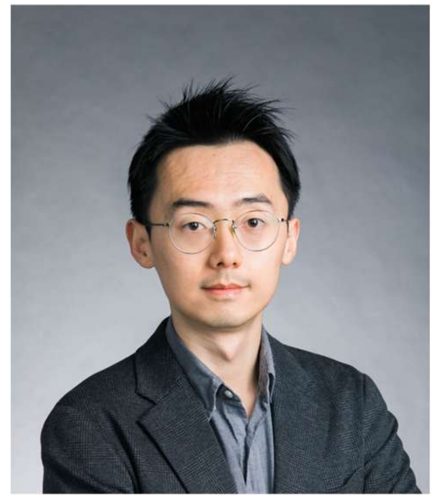
Conformalized Decision Risk Assessment

Wenbin Zhou

Co-Authors

Agni Orfanoudaki

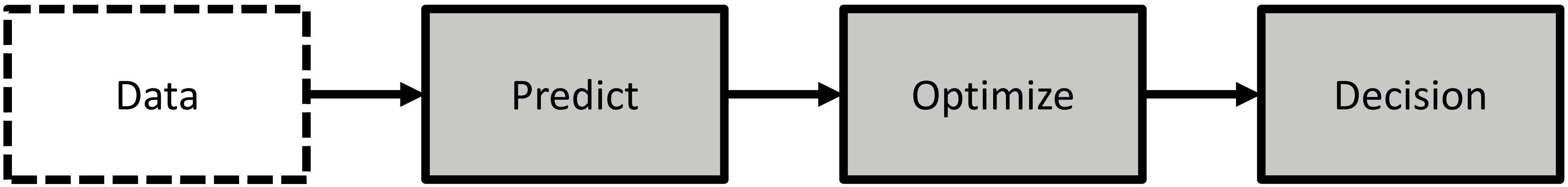
Associate Professor,
University of Oxford



Woody Zhu

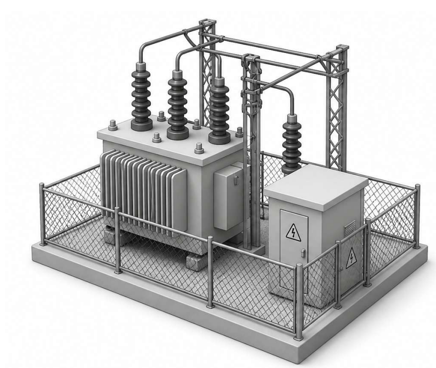
Assistant Professor,
Carnegie Mellon University

Introduction: Predict-Then-Optimize (PTO)

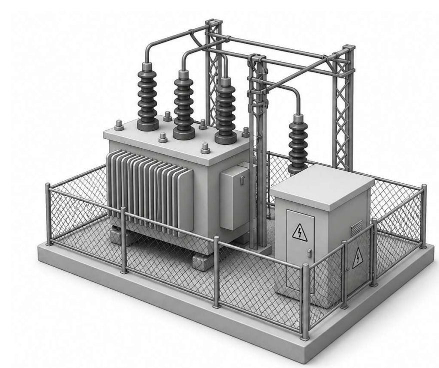


Example problem: power grid infrastructure planning

Two outdated substations:

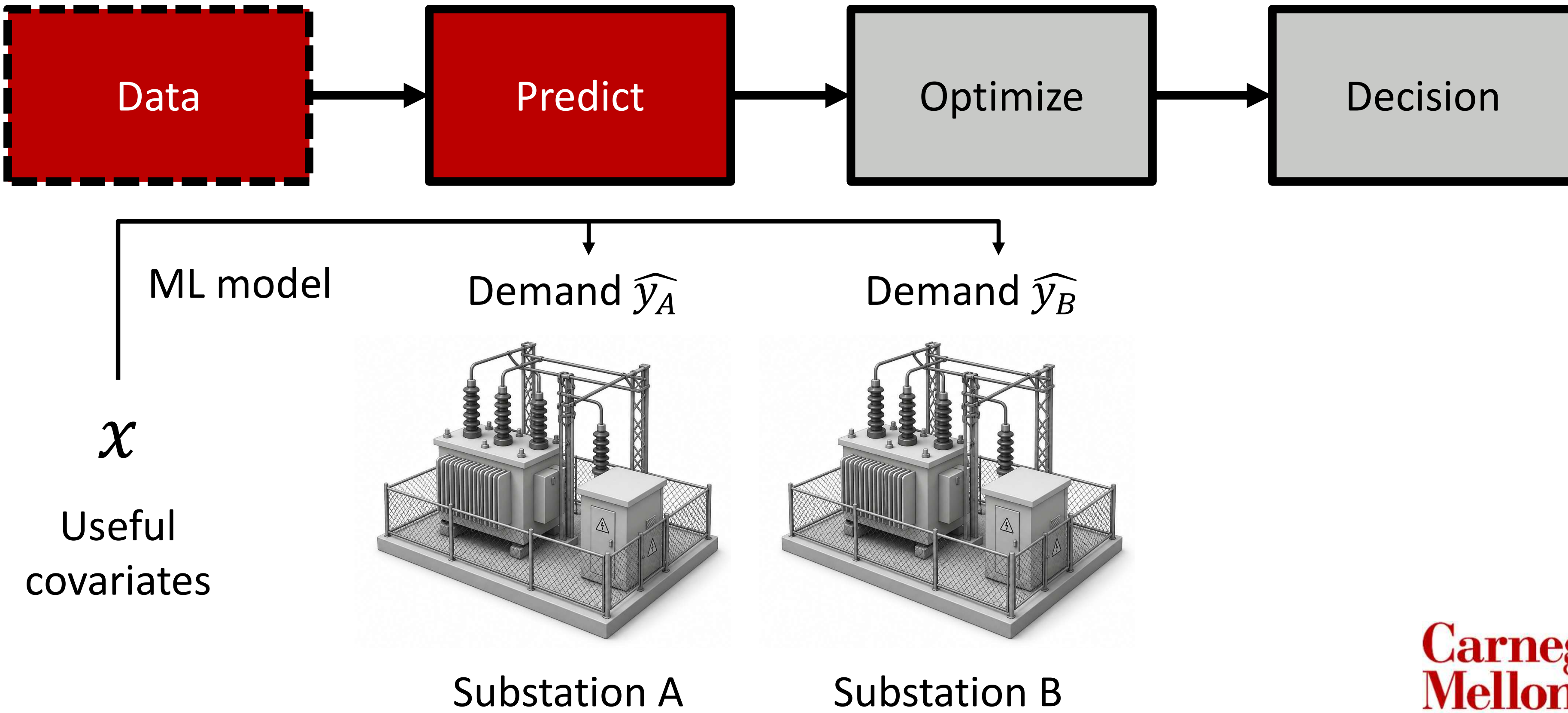


Substation A



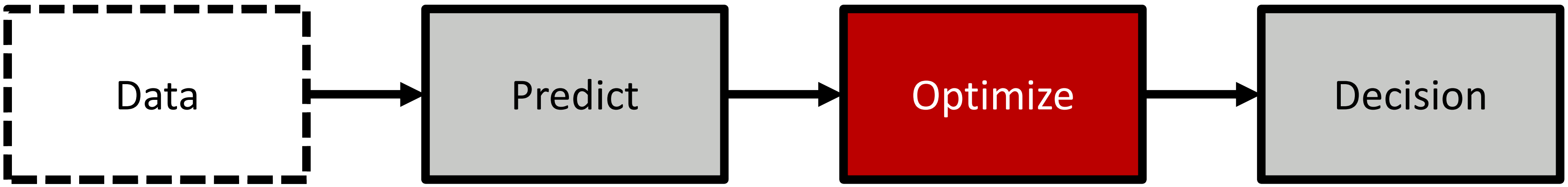
Substation B

Introduction: Predict-Then-Optimize (PTO)



Carnegie Mellon University

Introduction: Predict-Then-Optimize (PTO)



$$\min_{z \in \mathcal{Z}} f(z; \hat{y})$$

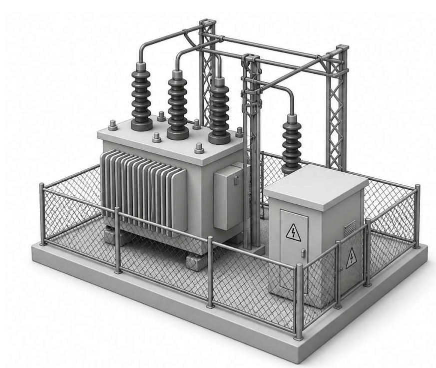
Minimize the loss, which is a function of budget allocation and estimated future demand

($\mathcal{Z}$ is a bounded feasible region, representing limited budget)

Introduction: Predict-Then-Optimize (PTO)

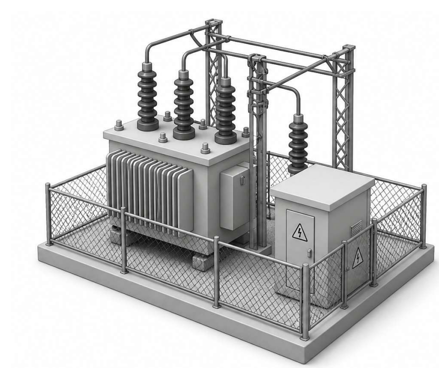
Data → Predict → Optimize → **Decision**

+$100



Substation A

+$200



Substation B

Introduction: Rethinking PTO

While PTO has been around for decades, it suffers from two core limitations:

Domain expert

"Liability issue"
"Interpretability"
"AI hallucination"

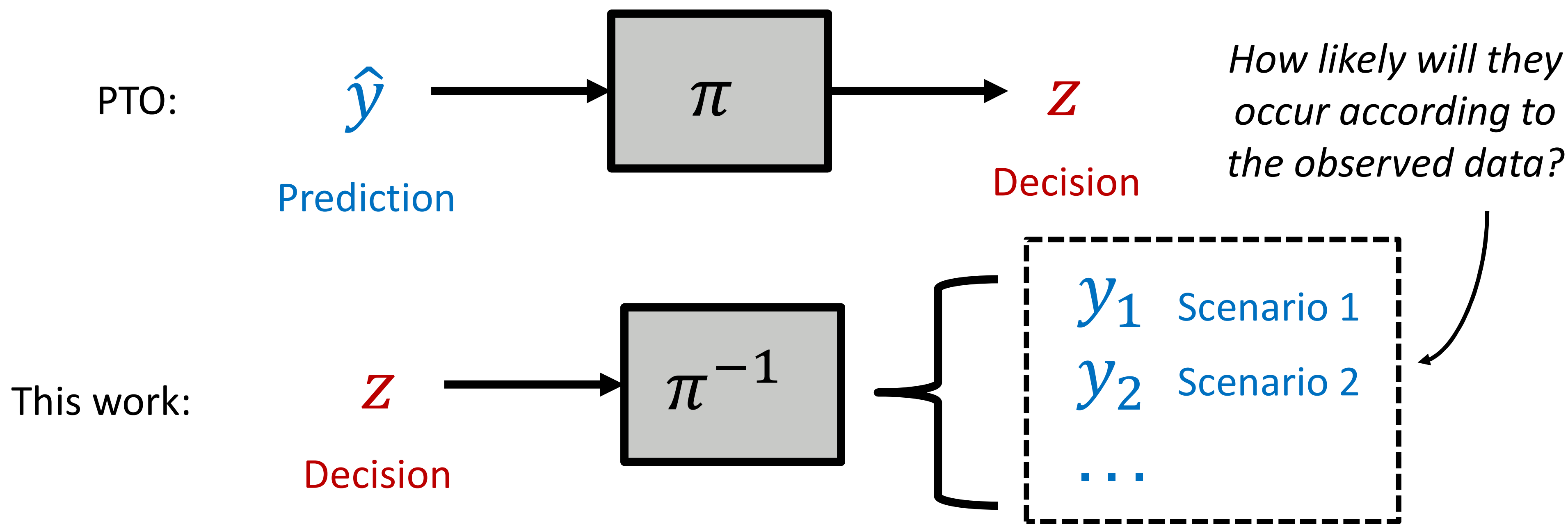
→ PTO prescribed decision are rarely directly adopted in practice

I have my own insights on the problem and I know what I want to do.

→ Want algorithms to help assess the risk of decisions they hope to take.

Q: Can we develop an algorithm that help people assess decisions they propose?

Overview: Decision Risk Assessment Framework



Step 1: Find π^{-1}

Step 2: Evaluate probability of

Problem Setup

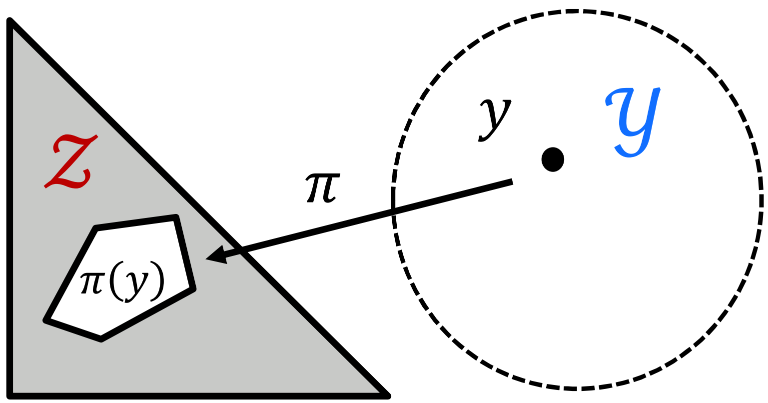
Definition: Decision Policy

$$\pi(y) \coloneqq \arg\min_{z \in \mathcal{Z}} f(z; y)$$

Decision policy defines a mapping:

$$\pi: \mathcal{Y} \mapsto 2^{\mathcal{Z}}$$

Ex. given future demand scenario, what are the best budget allocation plan?



Problem Setup

Definition: Decision Risk

Given fixed z and random variable Y, decision risk is defined as

$$\mathbb{P}\{z \in \pi(Y)\}$$

Ex. Give some proposal budget allocation plan, what is the probability it is optimal w.r.t. the random, unrealized future demand?

Definition: Decision Risk Assessment

Estimate $\hat{\alpha}$ (potentially dependent on z) such that

$$\mathbb{P}\{z \in \pi(Y)\} \geq 1 - \hat{\alpha}$$

Background: Inverse Optimization

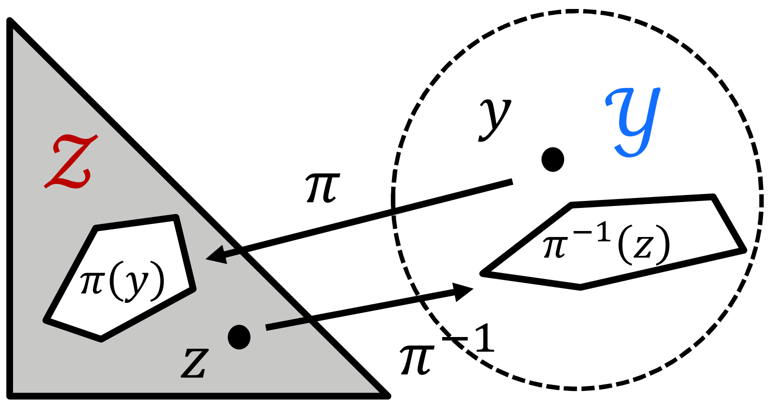
Definition: Inverse Feasible Region (Chan et al. 2020)

$$\pi^{-1}(z) = \{y \in \mathcal{Y}: f(z, y) = \min_{z' \in \mathcal{Z}} f(z', y)\}$$

Inverse feasible region also defines a mapping:

$$\pi^{-1}: \mathcal{Z} \mapsto 2^{\mathcal{Y}}$$

Ex. Find all scenarios of future demand that would make a given budget allocation plan optimal



Step 1: Inverse Problem Reformulation

Lemma: Objective reformulation

$$\mathbb{P}\{z \in \pi(Y)\} \equiv \mathbb{P}\{Y \in \pi^{-1}(z)\}$$

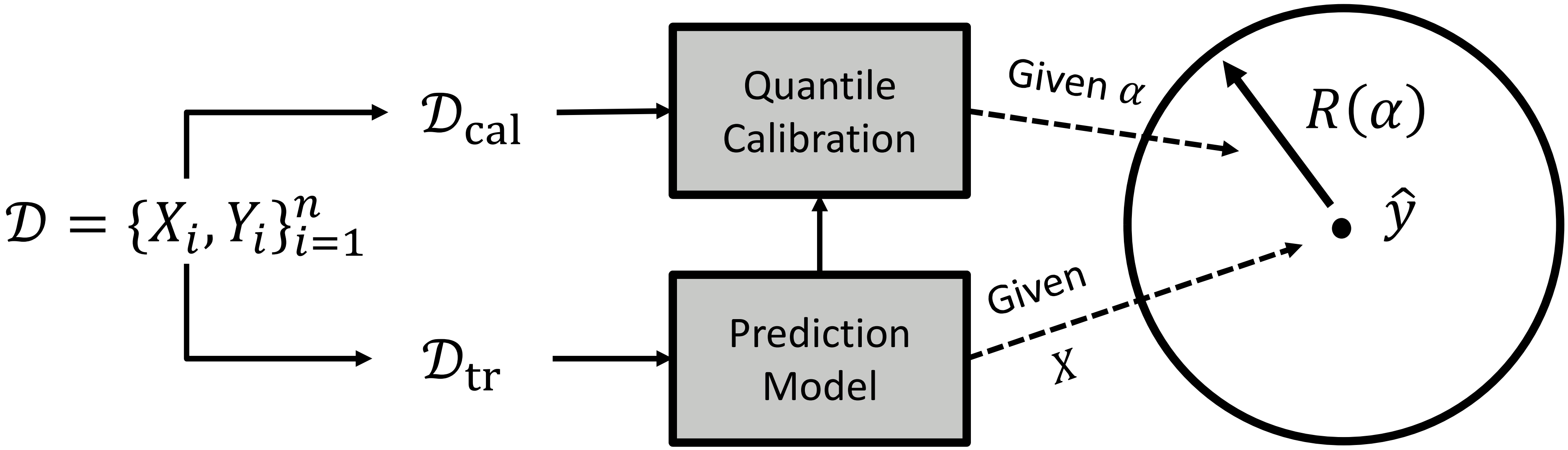
Decision risk ⟶ How likely would Y fall in $\pi^{-1}(z)$

Deterministic variable ∈ Random set ⟶ Random variable ∈ Deterministic set

Converted to a standard problem of uncertainty quantification, which has been widely studied in statistics & machine learning

Background: Conformal Prediction

Goal: construct valid prediction set from data.

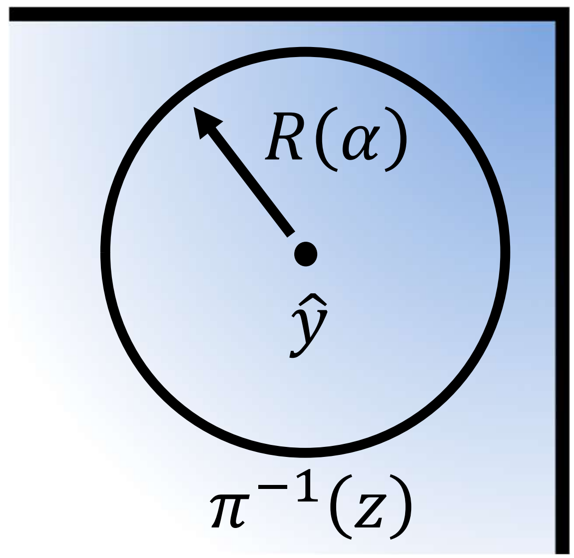


Validity: $\mathbb{P}\{Y \in \mathcal{C}(X;\alpha)\} \geq 1-\alpha$

Carnegie Mellon University

Intuitions For Step 2

If our configuration somehow gives:



Is a natural estimate for decision risk!

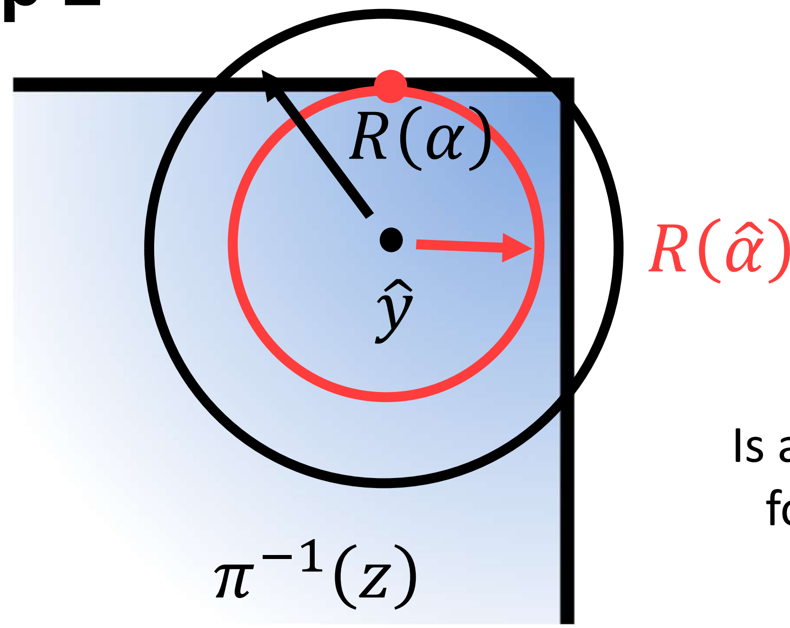
$$\mathbb{P}\{Y \in \pi^{-1}(z)\} \geq \mathbb{P}\{Y \in \mathcal{C}(X;\alpha)\} \geq 1-\alpha$$

Monotonicity

Validity (previous slide)

Intuitions For Step 2

It is also possible that our configuration gives:



Is again an estimate for decision risk!

$$\mathbb{P}\{Y \in \pi^{-1}(z)\} \geq \mathbb{P}\{Y \in \mathcal{C}(X; \hat{\alpha})\} \geq 1 - \hat{\alpha}$$

Monotonicity

Validity (previous slide)

Step 2: Conformalization

Step 2 can be described by a constrained optimization problem:

$$\min_{\alpha \in [0,1]} \alpha \quad \text{s.t.} \quad \mathcal{C}(X;\alpha) \subseteq \pi^{-1}(z)$$

Maximize the size of the set

Ensure the set is contained within $\pi^{-1}(z)$

The solution $\hat{\alpha}$ should supposedly meet our objective:

$$\mathbb{P}\{Y \in \pi^{-1}(z)\} \geq 1 - \hat{\alpha}$$

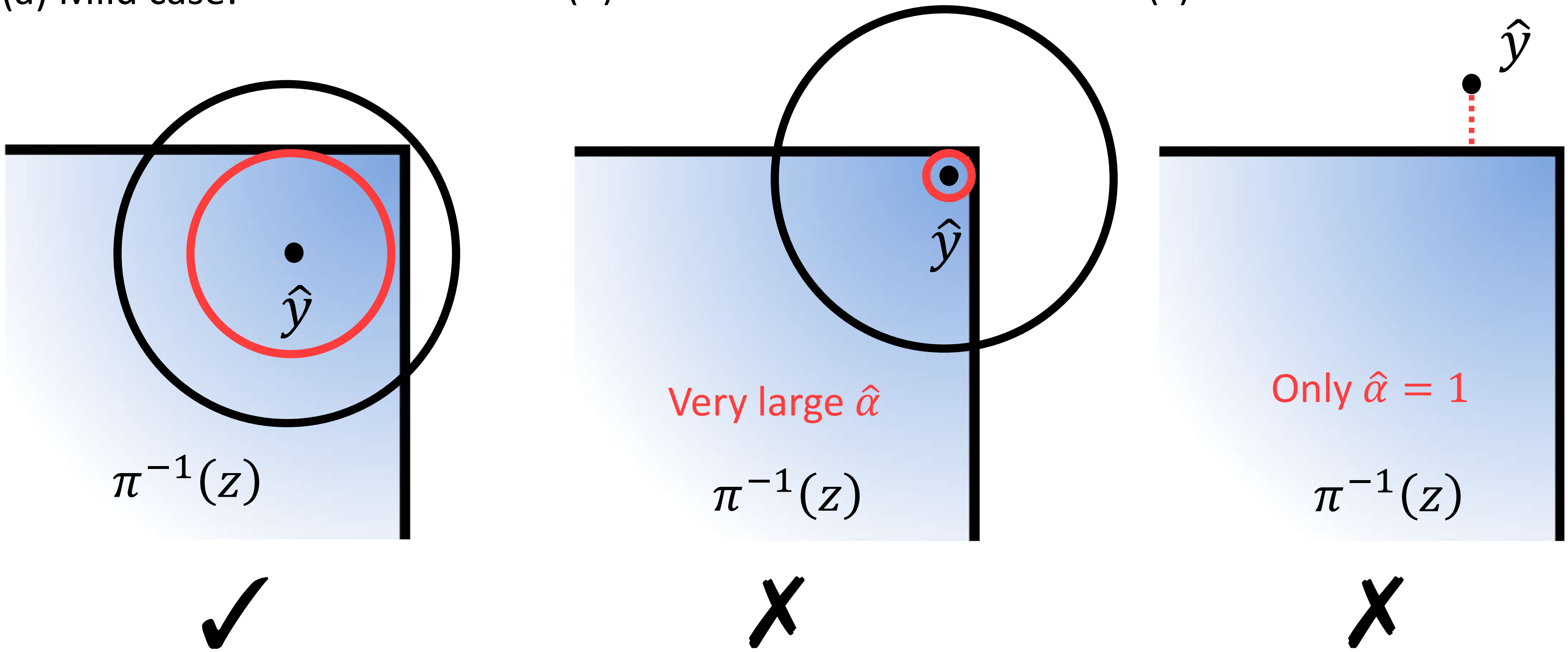
However, is that the end of the story?

Challenge 1: Over-conservatism

(a) Mild case:

(b) Severe case:

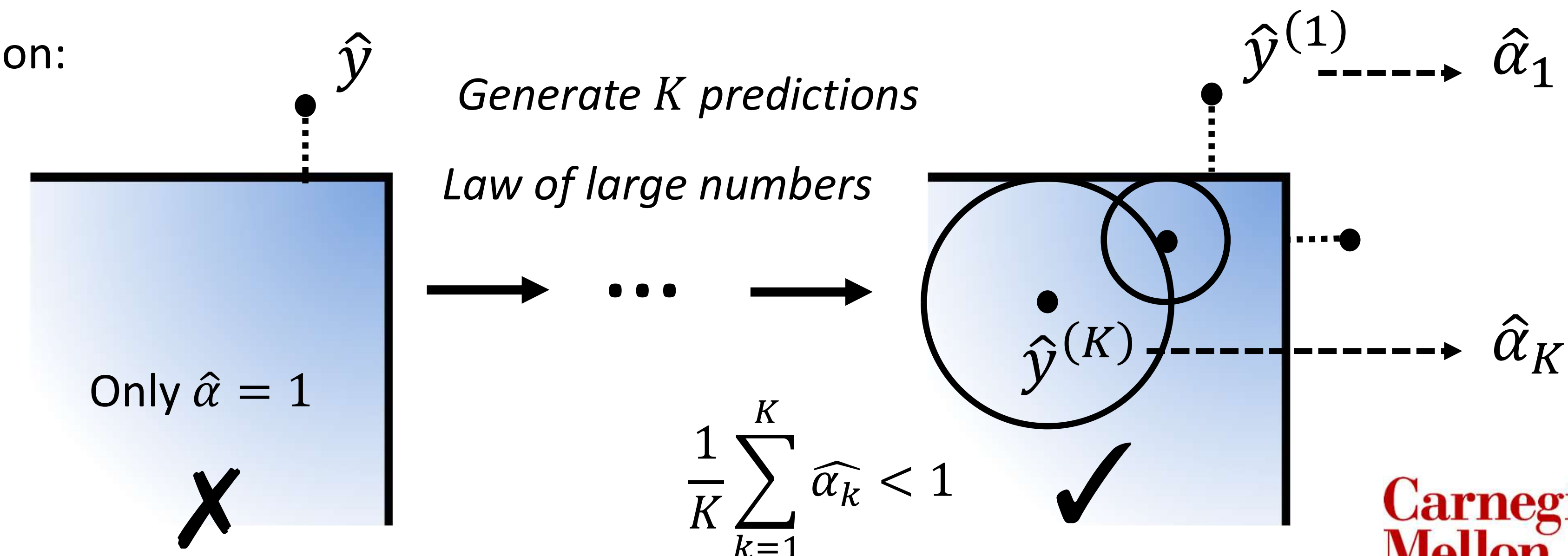
(c) Severest case:



Solution: Generate Multiple Prediction Sets

Idea: Replace **Prediction Model** with **Sampling Model** in conformal prediction

Intuition:



Theoretical Foundation For This Idea

Lemma (Informal): Weighted Sample Average Estimator

Sampling Model :

$$\hat{\alpha} \propto \frac{1}{K} \sum_{k=1}^{K} w_k \cdot \mathbb{1}\{\hat{y}^{(k)} \notin \pi^{-1}(z)\}$$

Conformalized weight

Indicator of whether the generated prediction is not within $\pi^{-1}(z)$

Prediction Model : Special case with $K = 1$ (i.e., one shot estimator)

Challenge 2: Post-hoc Invalidity

Recall the following inequality previously when talking about the intuition of Step 2:

$$\mathbb{P}\{Y \in \pi^{-1}(z)\} \geq \mathbb{P}\{Y \in \mathcal{C}(X; \hat{\alpha})\} \geq 1 - \hat{\alpha}$$

In theory, this does not hold for $\hat{\alpha}$!

Lemma (Informal): Potential Invalidity

$$\mathbb{P}\{Y \in \mathcal{C}(X; \hat{\alpha})\} \geq 1 - \hat{\alpha} - \Delta(P)$$

The reason: $\hat{\alpha}$ becomes data-dependent though post-hoc tuning, whereas standard conformal prediction validity only holds when α is a prespecified constant

Solution: *e*-statistics

Idea: Replace **Quantile Calibration** with ***e*-value Calibration** in conformal prediction

$$\hat{Q}\left(\frac{\lfloor (n+1)(1-\alpha) \rfloor}{n}\right)$$

$$\frac{\sum_{i=1}^{n} || \hat{Y}_i - Y_i ||_2}{\alpha(n+1) - 1}$$

Intuition: Convert testing to betting

Theorem (Informal): Post-hoc Validity

e-value Calibration : We recover exact validity without the $\Delta(P)$ term

Computation: End-To-End Algorithm

Theorem (Informal): End-To-End Optimization Cleaned-Up Form

$$\hat{\alpha}(z) = \sup_{y \in \mathcal{Y}} R^{-1}(||y - \hat{y}||_2) \quad \text{s.t.} \quad f(z; y) > \min_{z' \in \mathcal{Z}} f(z'; y) + \epsilon$$

Closed-form Solution for Linear Programming

$$\hat{\alpha}(z) = \max_{v \in \mathcal{V}} R^{-1}\left(\frac{|\langle \hat{y}, z - v\rangle - \epsilon|}{||z - v||_2}\right)$$

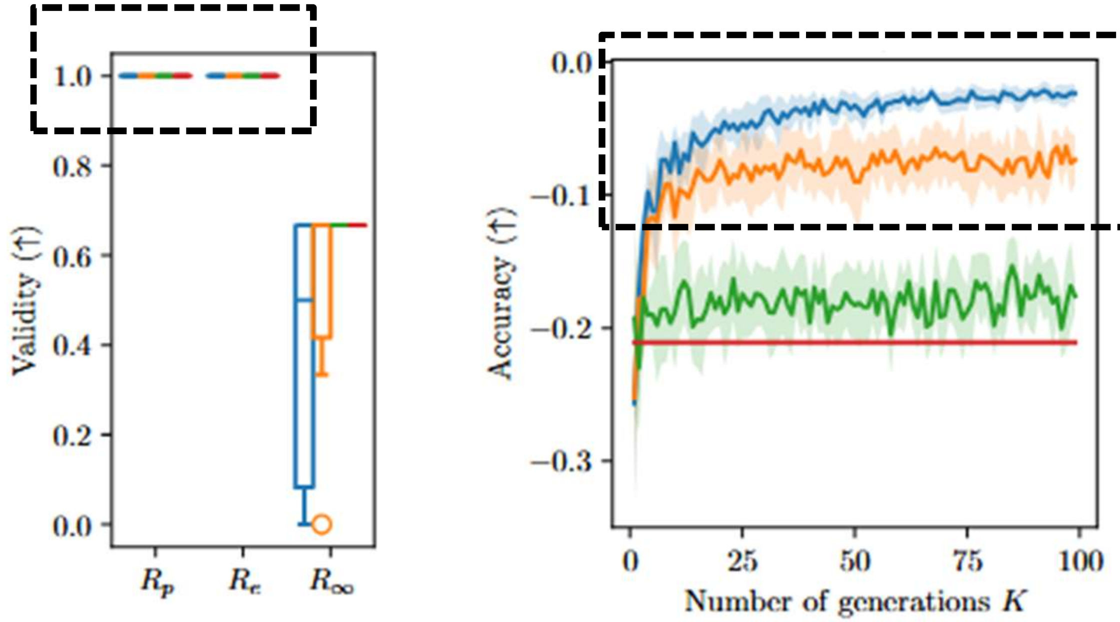
Other General Convex Problems

Proposed an alternating optimization algorithm…

The framework can be feasibly executed for many types of optimizations.

Experiments: Estimation Quality Analysis

100% valid risk estimate for all variants



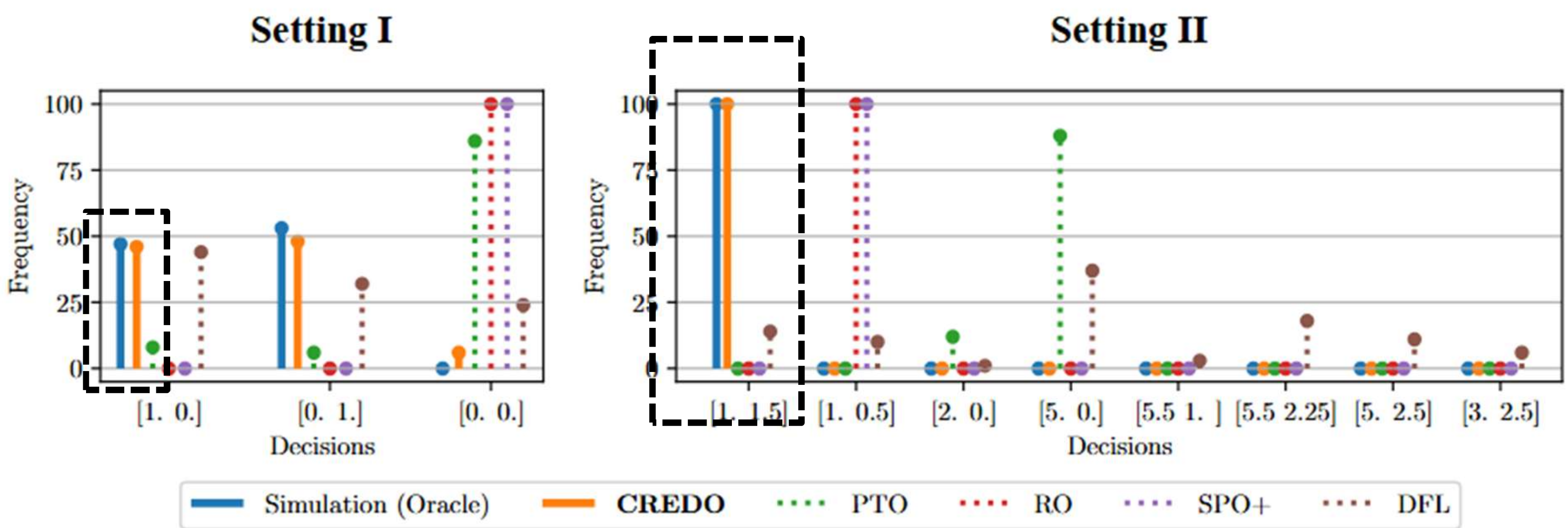
Approximation error between our method and oracle → 0

Figure: Evaluation results. Different colors represent different ablation variants.

(1) The risk estimate is both valid and accurate; (2) Generative Design helps mitigate overconservativeness.

Carnegie Mellon University

Experiments: Decision Prescription Analysis



Our method chooses same decision as the oracle

Summary

Objective: We study the problem of conservatively estimating decision risk

Procedure: (1) Inverse Problem Reformulation, (2) Conformalization

Refinements:
- Prediction model → Sampling models
- Quantile calibration → *e*-statistics calibration

Theories: (1) Weighted sample average, (2) post-hoc validity, (3) end-to-end optimization

Email: wenbinz2@andrew.cmu.edu

Link to this paper

Link to my homepage

Appendix

Table 3 Evaluated metrics for different optimization procedures solving (13b) across different settings.

	LP Setting I			LP Setting II			QP			SOCP			IP		
	Obj	Vio	Err	Obj	Vio	Err	Obj	Vio	Err	Obj	Vio	Err	Obj	Vio	Err
GD	<u>0.44</u>	0.40	0.45	<u>0.06</u>	<u>0.07</u>	<u>0.05</u>	0.60	0.23	0.50	1.53	0.13	1.31	0.22	**0.00**	0.19
BF	0.66	**0.00**	<u>0.43</u>	0.11	**0.00**	0.10	<u>0.52</u>	**0.00**	<u>0.34</u>	<u>0.66</u>	**0.00**	<u>0.44</u>	<u>0.11</u>	**0.00**	<u>0.09</u>
RS	0.89	**0.00**	0.66	0.12	**0.00**	0.11	0.70	**0.00**	0.52	0.91	**0.00**	0.69	0.12	**0.00**	0.10
RG	5.21	<u>0.20</u>	4.98	1.35	0.17	1.34	4.81	0.17	4.63	5.21	0.17	4.99	1.35	**0.00**	1.32
CREDO	**0.23**	**0.00**	**0.00**	**0.01**	**0.00**	**0.00**	**0.10**	<u>0.10</u>	**0.08**	**0.14**	<u>0.03</u>	**0.08**	**0.00**	**0.00**	**0.02**

Table 1 **Evaluated metrics for different risk estimation methods across different optimization settings.**

	LP Setting I		LP Setting II		QP ($\epsilon = 0.1$)		SOCP ($\epsilon = 0.2$)		IP ($\epsilon = 0.3$)	
	Validity (↑)	MAE (↓)	Validity (↑)	MAE (↓)	Validity (↑)	MAE (↓)	Validity (↑)	MAE (↓)	Validity (↑)	MAE (↓)
SA	0.53 ± 0.50	$\mathbf{0.04 \pm 0.03}$	0.56 ± 0.44	$\mathbf{0.03 \pm 0.02}$	0.43 ± 0.48	$\mathbf{0.04 \pm 0.03}$	0.47 ± 0.48	$\mathbf{0.03 \pm 0.03}$	0.41 ± 0.46	$\mathbf{0.04 \pm 0.03}$
LR	0.50 ± 0.48	0.06 ± 0.04	0.59 ± 0.39	0.03 ± 0.02	0.47 ± 0.50	0.05 ± 0.04	0.50 ± 0.45	0.06 ± 0.05	0.45 ± 0.46	0.07 ± 0.04
NN	0.50 ± 0.45	0.10 ± 0.09	0.39 ± 0.38	0.05 ± 0.03	0.63 ± 0.46	0.08 ± 0.05	0.40 ± 0.48	0.08 ± 0.06	0.48 ± 0.49	0.09 ± 0.06
QE	0.00 ± 0.00	0.62 ± 0.13	0.89 ± 0.16	0.15 ± 0.06	0.03 ± 0.10	0.60 ± 0.10	0.00 ± 0.00	0.56 ± 0.00	0.14 ± 0.04	0.53 ± 0.06
CP	$\mathbf{1.00 \pm 0.00}$	0.31 ± 0.00	$\mathbf{1.00 \pm 0.00}$	0.12 ± 0.00	$\mathbf{1.00 \pm 0.00}$	0.37 ± 0.00	$\mathbf{1.00 \pm 0.00}$	0.43 ± 0.01	$\mathbf{1.00 \pm 0.00}$	0.31 ± 0.00
CREDO (p)	$\mathbf{1.00 \pm 0.00}$	0.27 ± 0.01	$\mathbf{1.00 \pm 0.00}$	0.11 ± 0.00	$\mathbf{1.00 \pm 0.00}$	0.16 ± 0.03	$\mathbf{1.00 \pm 0.00}$	0.38 ± 0.02	$\mathbf{1.00 \pm 0.00}$	0.25 ± 0.02
CREDO (e)	$\mathbf{1.00 \pm 0.00}$	0.31 ± 0.00	$\mathbf{1.00 \pm 0.00}$	0.12 ± 0.00	$\mathbf{1.00 \pm 0.00}$	0.18 ± 0.04	$\mathbf{1.00 \pm 0.00}$	0.44 ± 0.00	$\mathbf{1.00 \pm 0.00}$	0.31 ± 0.00
CREDO (∞)	0.50 ± 0.50	$\underline{0.05 \pm 0.03}$	0.53 ± 0.43	$\underline{0.03 \pm 0.02}$	0.60 ± 0.48	$\underline{0.05 \pm 0.03}$	0.53 ± 0.49	$\underline{0.05 \pm 0.03}$	0.47 ± 0.47	$\underline{0.05 \pm 0.04}$

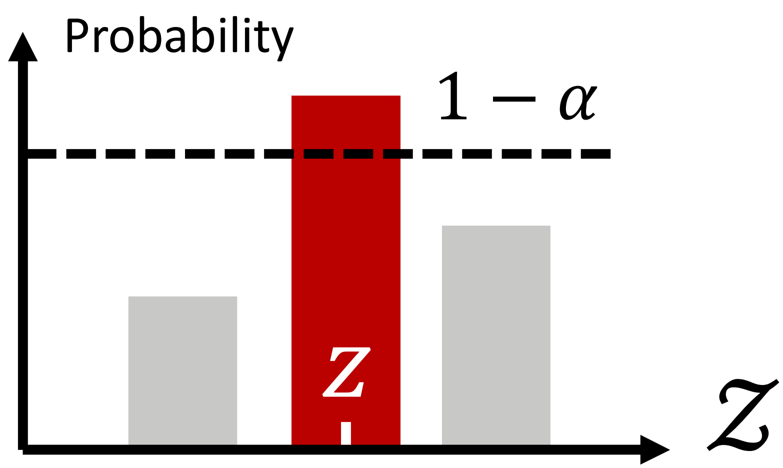
Table 2 **Evaluated empirical confidence ranking (↓) for different methods across different optimization settings.**

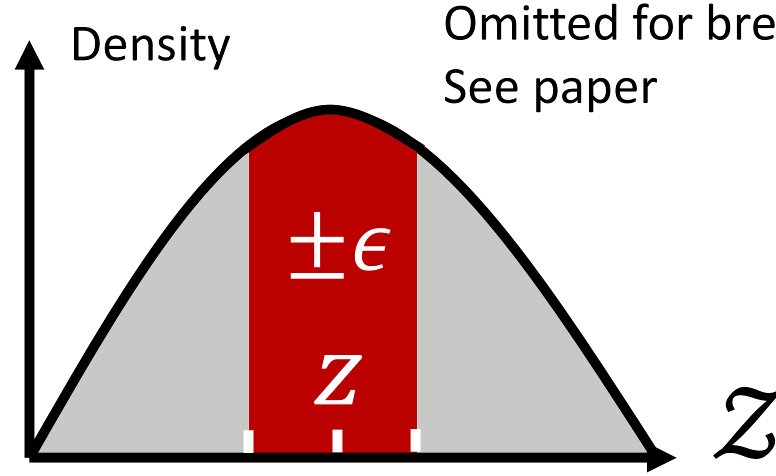
Method	Setting I			Setting II			Real Data
	$\sigma = 0.1$	$\sigma = 1$	$\sigma = 10$	$\sigma = 0.1$	$\sigma = 1$	$\sigma = 10$	
PTO	**1.00 ± 0.00**	2.76 ± 0.59	2.24 ± 0.79	3.55 ± 0.50	<u>3.36 ± 0.48</u>	2.04 ± 1.65	<u>1.75 ± 1.69</u>
RO	**1.00 ± 0.00**	2.98 ± 0.14	3.00 ± 0.00	4.99 ± 0.10	6.00 ± 0.00	3.98 ± 0.80	3.00 ± 1.29
SPO+	**1.00 ± 0.00**	2.68 ± 0.65	2.02 ± 0.82	3.95 ± 1.20	4.67 ± 1.56	3.56 ± 1.50	2.67 ± 1.43
DFL	2.44 ± 0.64	1.83 ± 0.81	2.06 ± 0.79	3.60 ± 1.52	3.96 ± 2.07	3.66 ± 2.48	1.92 ± 1.04
CREDO (1-GMM)	1.94 ± 0.87	1.56 ± 0.54	<u>1.49 ± 0.50</u>	3.74 ± 0.98	3.94 ± 1.37	2.02 ± 1.41	1.92 ± 1.04
CREDO (3-GMM)	1.75 ± 0.77	**1.61 ± 0.56**	**1.48 ± 0.52**	1.05 ± 0.22	**1.00 ± 0.00**	2.03 ± 0.96	**1.75 ± 0.92**
CREDO (5-GMM)	1.89 ± 0.87	1.65 ± 0.62	1.54 ± 0.52	<u>1.03 ± 0.17</u>	**1.00 ± 0.00**	<u>1.92 ± 0.89</u>	1.92 ± 1.04
CREDO (VAE)	<u>1.01 ± 0.10</u>	<u>1.61 ± 0.58</u>	1.77 ± 0.71	**1.00 ± 0.00**	**1.00 ± 0.00**	**1.06 ± 0.24**	1.92 ± 1.04

Problem Setup

Objective: Decision Risk Assessment

$$\mathbb{P}\{z \in \pi^{-1}(Y)\} \geq 1 - \alpha$$

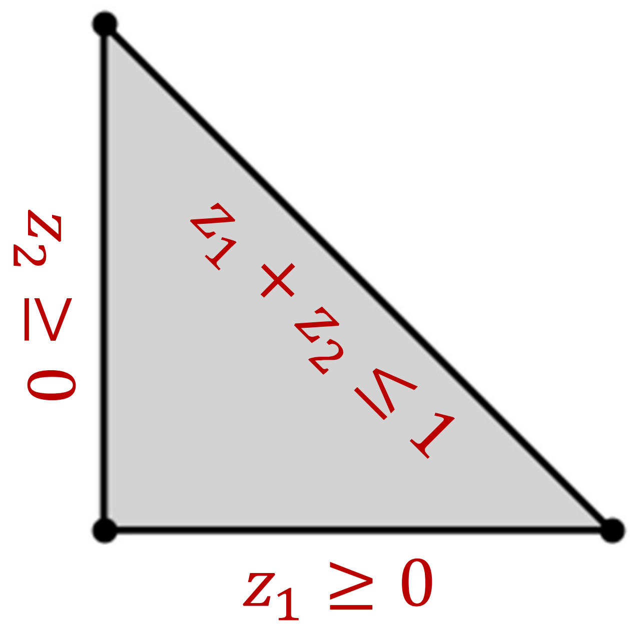




Omitted for brevity.
See paper

Running Example: Linear Programming

$$\min_{z} y^{\top} z \quad \text{s.t.} \quad Az \leq b$$



$$A \in \mathbb{R}^{3 \times 2}$$

Number of constraints

Spatial dimension

This is a 2D (triangular) shape!

Example: Linear Programming Inverse Optimization

$$\min_z y^\top z \ \text{ s.t. } Az \leq b$$