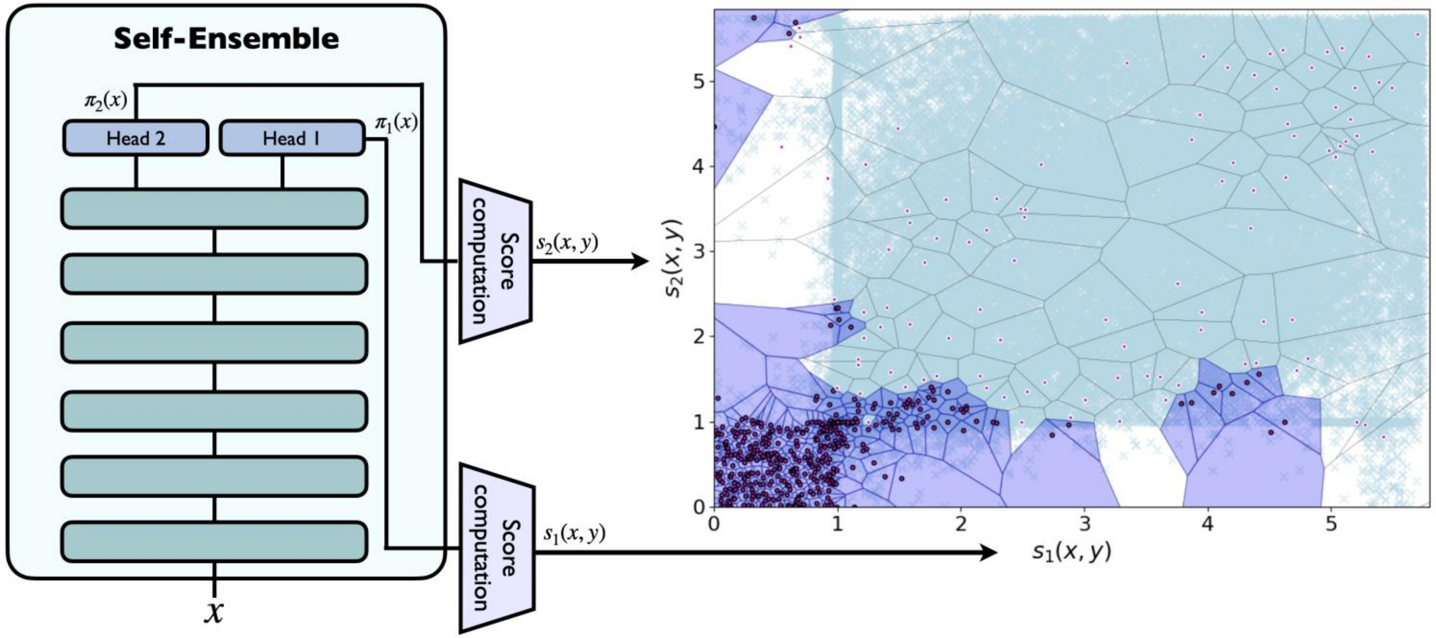
Multi-Dimensional Conformal Prediction



Yam Tawachi & Bracha Laufer-Goldshtein

School of Electrical Engineering, Tel Aviv University

ICLR 2025

School of Electrical Engineering & Computer Engineering, ECE
The Iby and Aladar Fleischman Faculty of Engineering
Tel Aviv University

Conformal Prediction

Standard Image classification

→ **Black Box Model** → {Lion}

+ Conformal perdition

→ **Black Box Model** → **CP Block** → {Lion, Cat, Tiger}

Contains true label with high-probability:

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

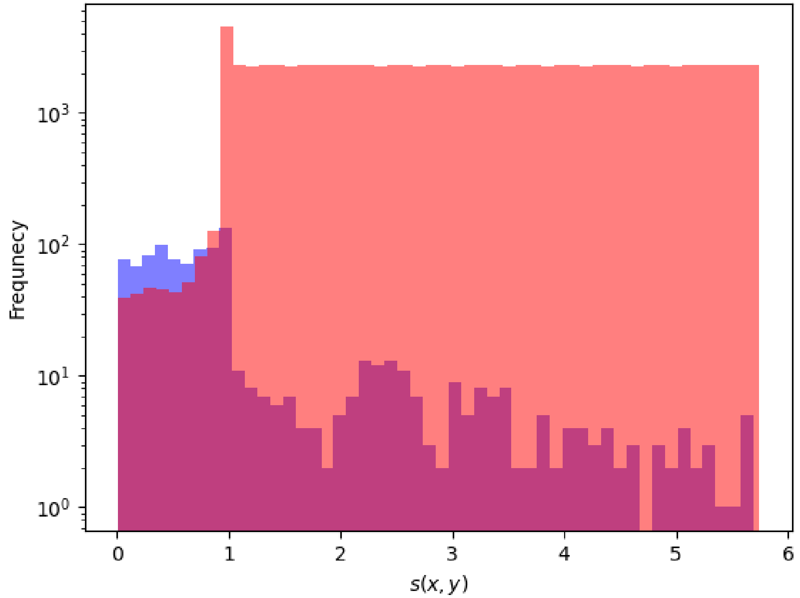
Important when deploying models in safety-critical applications (healthcare, autonomous-driving, finance..)

Motivation

- **Conformal prediction effectiveness** depends on how uncertainty is quantified.
- **Traditional methods** use a **single-dimensional conformity score.**

- False Scores
- True Scores

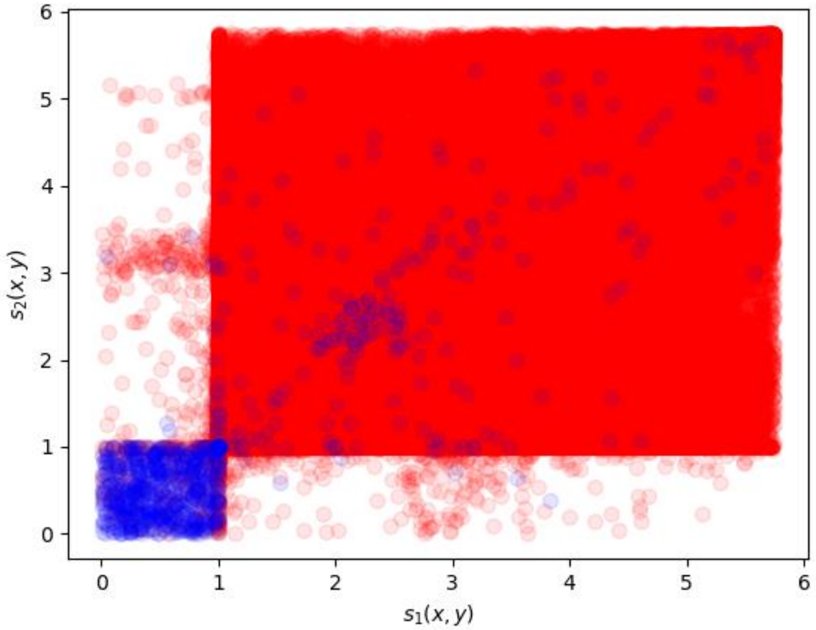
Single Dimension



✗ A single score may **fail to capture uncertainty**.

✗ **Limited separation** between true and false predictions.

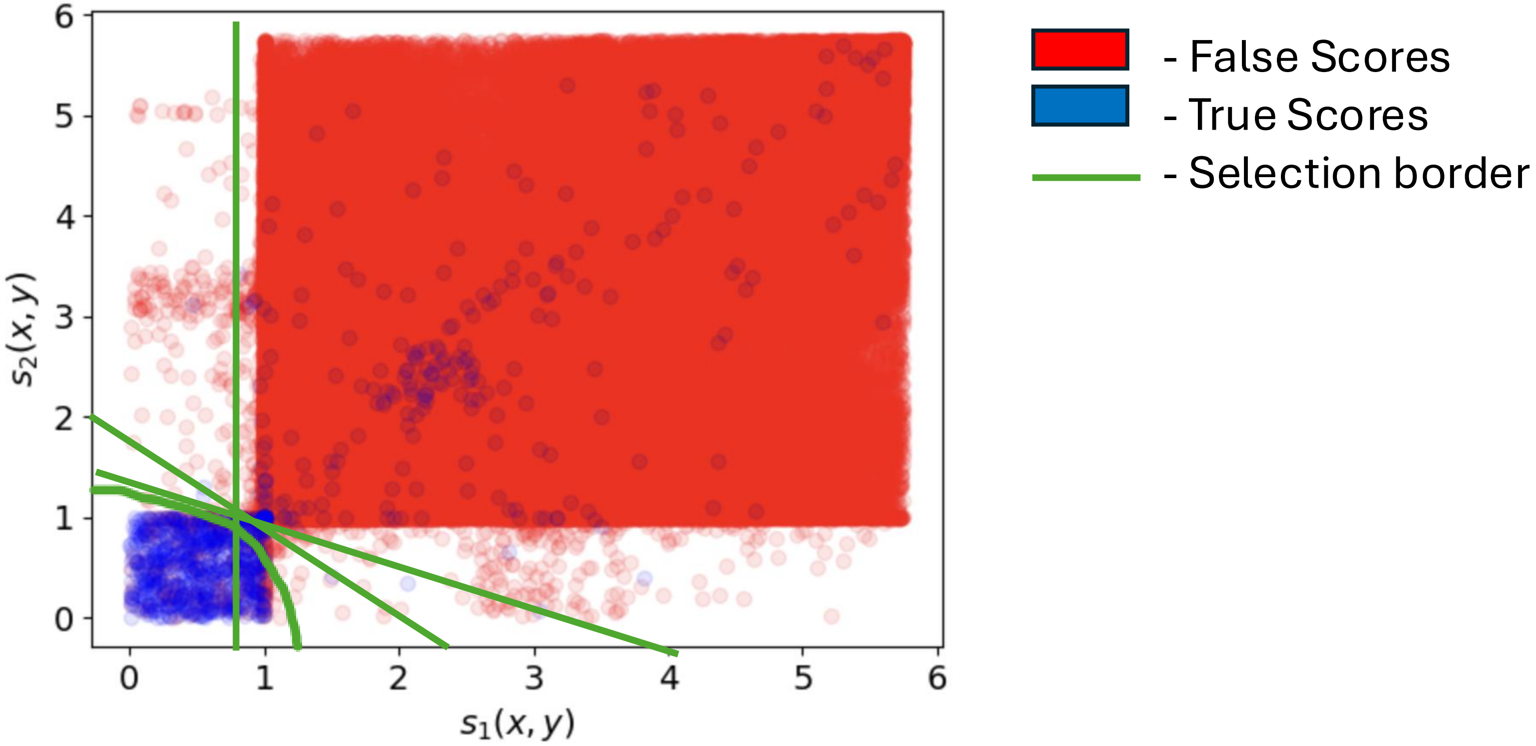
Multi-Dimension



✓ Multiple scores present **complimentary viewpoints on uncertainty**.

✓ In higher dimensions we can obtain **better separation** between true and false predictions.

Challenges in High Dimensions



Numerous ways of splitting the space to obtain $1 - \alpha$ coverage.

Structured selection regions (e.g. weighted average) may be **suboptimal**

New optimization is required for each α level.

Proposed Method: Multi Dimensional CP

Partition

Partition the space into cells centered around each calibration score

Scoring

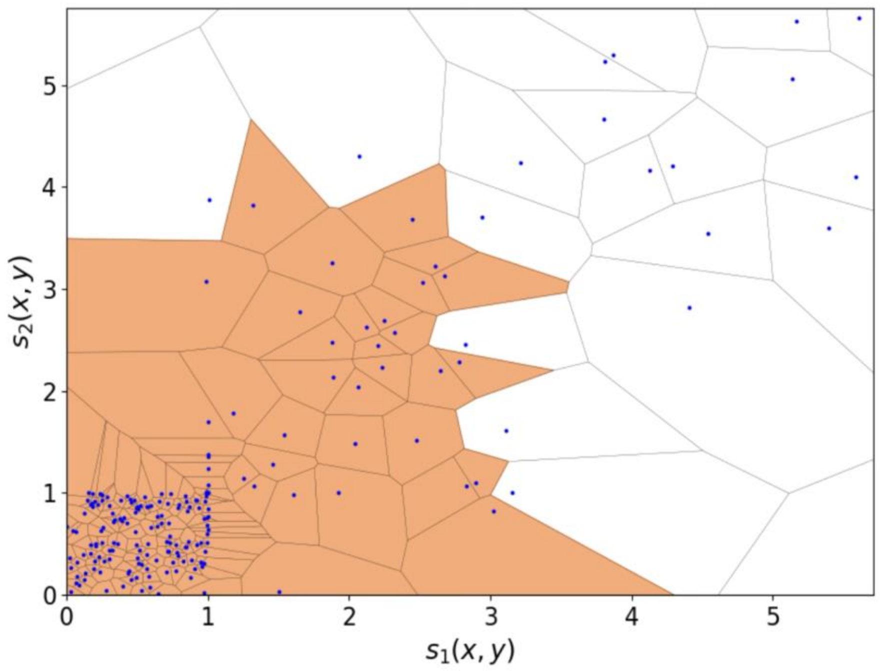
Compute cell scores: $D_i = \frac{F_i + T_i}{T_i}$, F_i - false scores, T_i - true scores

Ranking

Rank the cells $D_{(i)} \leq D_{(2)} \leq \cdots \leq D_{(k)}$

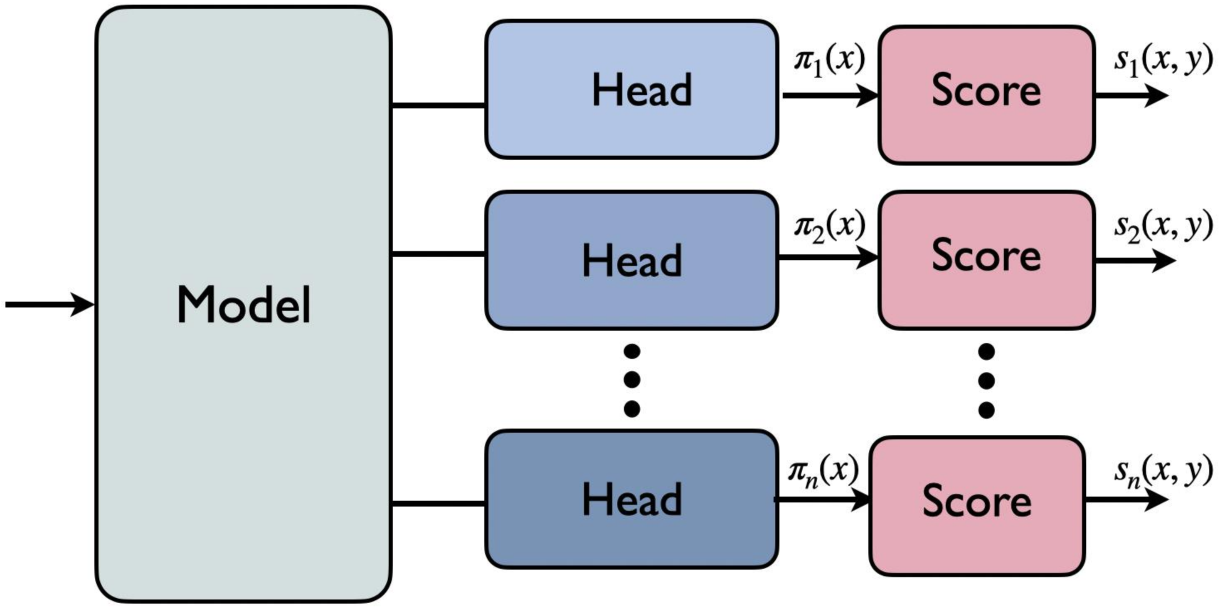
Selection

Select the region $\mathcal{C}_{\text{in}}^{\eta^*} = \cup_{i=1}^{\eta^*} \mathcal{C}_{(i)}$, ensuring $(1-\alpha)$ coverage.



Constructing a Multi Dimensional Score

- We propose a new method for obtaining **diverse scores** with **low cost**.
- We attach **multiple classification heads** $\{\pi_i(x)\}_{i=1}^{n}$.
- We add a **regularization term** to increase diversity among heads: $\frac{\beta}{n(n-1)}\sum_{i=1}^{n}\sum_{i\neq j}\mathrm{sim}(\pi_i(x),\pi_j(x))$

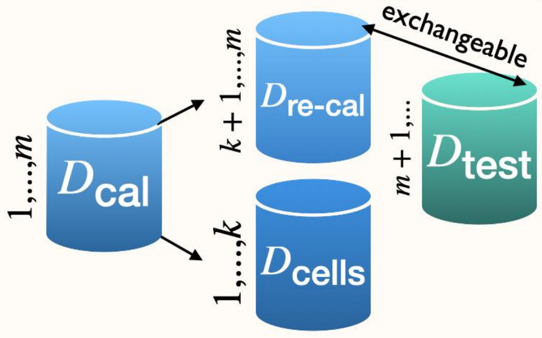


Theoretical analysis

Coverage Guarantees

Let $\mathscr{D}_{\text{cells}}$ and $\mathscr{D}_{\text{re-cal}}$ be two **disjoint subsets**,
and $\{(X_i, Y_i)\}_{i=k+1}^{m+1}$ is **exchangeable**.
For any multi-dimensional score $\mathbf{s} : \mathscr{X} \times \mathscr{Y} \to \mathscr{S} \subseteq \mathbb{R}^n$ and $\alpha \in (0,1)$,
the prediction set $\Gamma_{\eta^*}(X_{m+1}) = \left\{ y \in \mathscr{Y} \mid \mathbf{s}(X_{m+1}, y) \in \mathscr{C}_{\text{in}}^{\eta^*} \right\}$ satisfies:

$$\mathbb{P}\left(Y_{m+1} \in \Gamma_{\eta^*}(X_{m+1}) \right) \geq 1 - \alpha$$



Efficiency

It can be shown that our selection procedure is equivalent to solving:

$$\text{argmin}_{I \subseteq 2^k} \mathbb{E}\left[\text{size}(\Gamma^I(X))\right]$$ Set-size minimization

$$\text{s.t. } \mathbb{E}\left[\mathbf{1}\{Y \in \Gamma^I(X)\}\right] \geq 1 - \alpha$$ Coverage constraint

where $\Gamma^I(x) = \{y \in \mathscr{Y} \mid \mathbf{s}(x, y) \in \cup_{i \in I} \mathscr{C}_i\}$ is a **union of cells**.

Results

Coverage & Set-size

Baselines

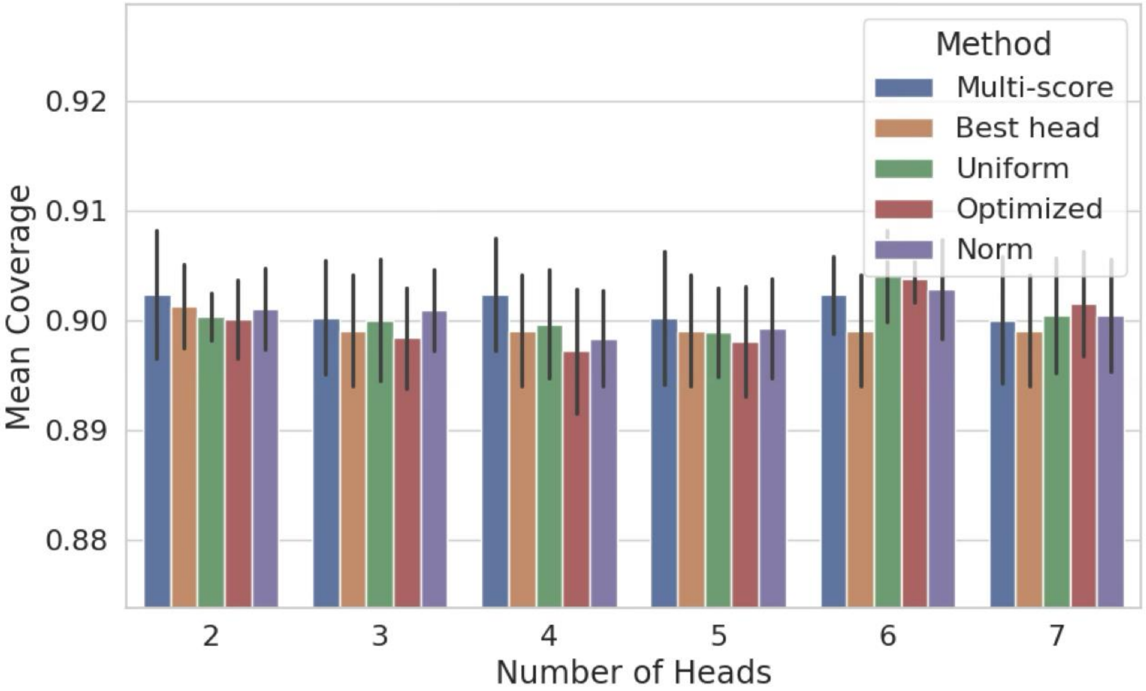
Best Head - single best head

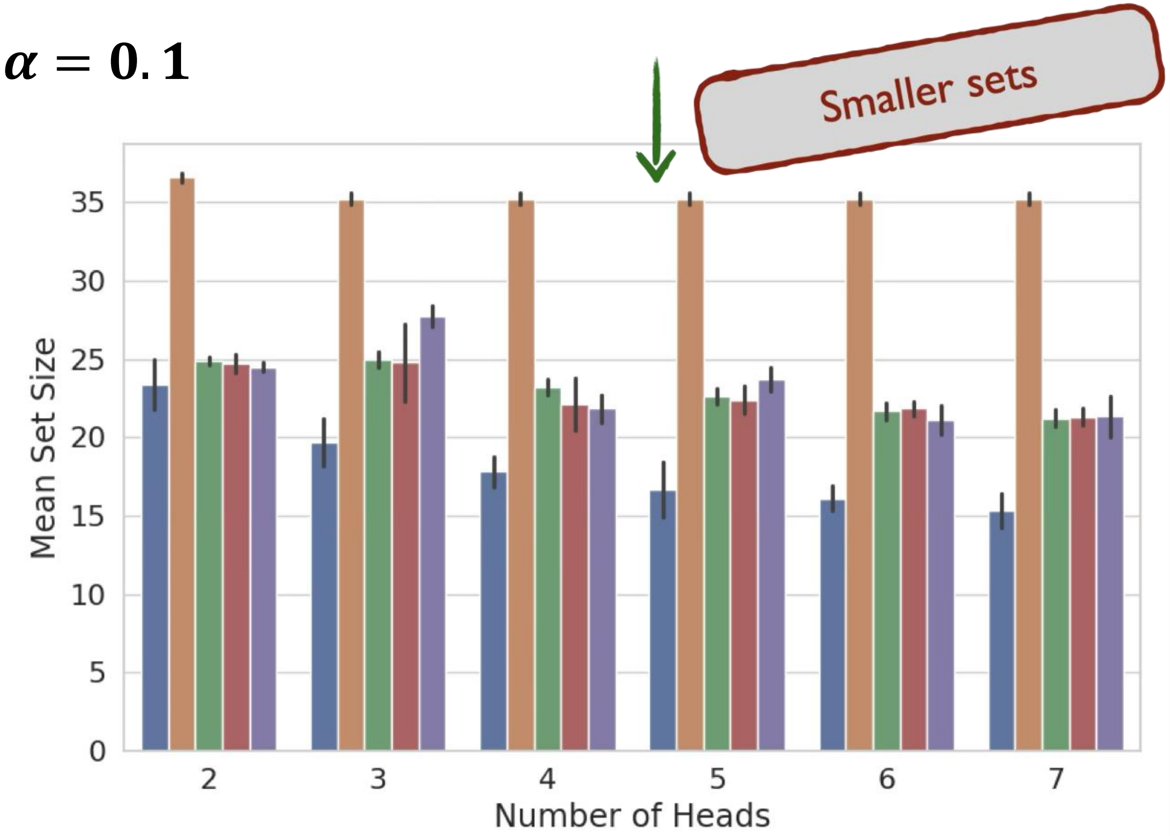
Uniform – uniform avg.

Optimized –optimized weighted avg.

Norm - circle-shaped region

CIFAR-100, $\alpha = 0.1$



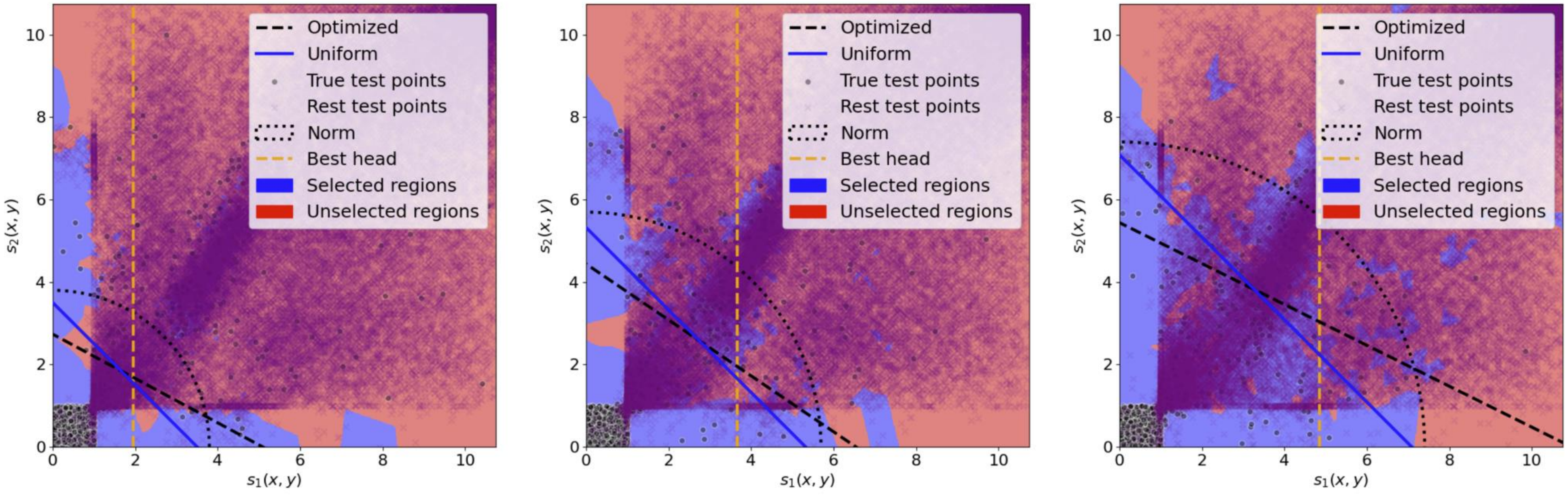


Results

selected region

Showing:

- The region selected by our method in blue.
- The decision boundaries for the baselines.
- True test points: Green circles
- Incorrect labels: Purple x-marks.

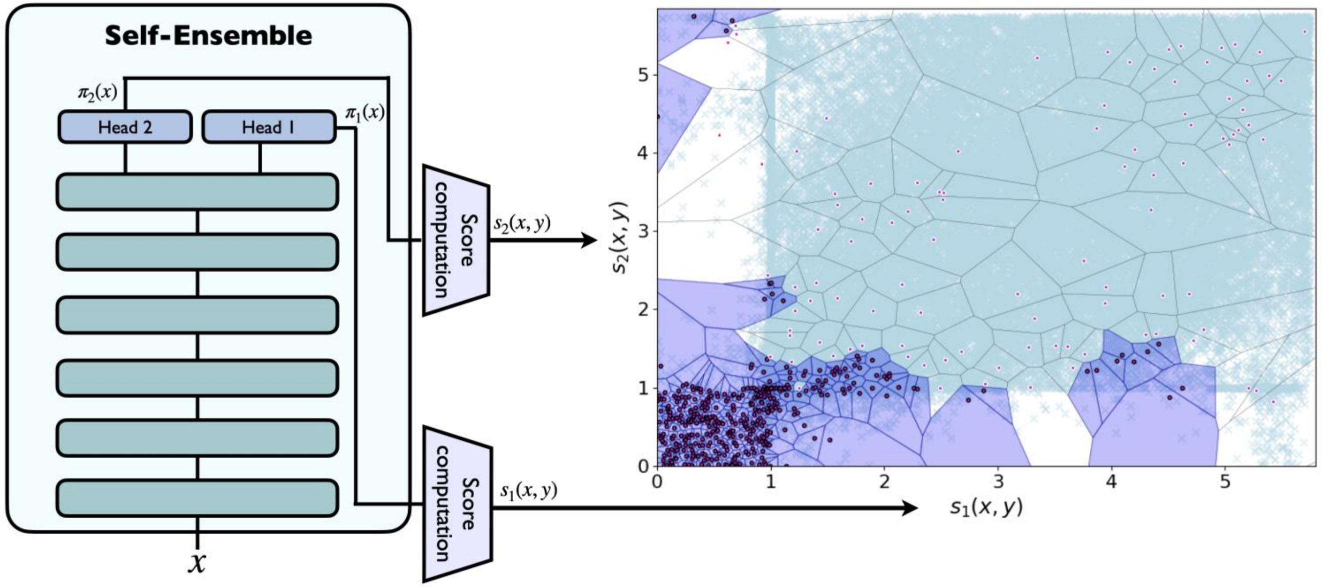


(a) $\alpha = 0.2$ (b) $\alpha = 0.1$ (c) $\alpha = 0.05$

Our method effectively **focuses on regions with fewer false labels**, while the baselines are constrained by their fixed structure, including areas with a high density of false labels.

Summary

- Introduce an extension of conformal prediction to higher dimensions
- Provides smaller and more informative prediction sets
- Requires no additional optimization
- Offers flexible, finite-sample coverage guarantees



Visit our Github page & Colab demo - https://github.com/yamtawa/Multi-CP

Or check out the paper for more details - *MULTI-DIMENSIONAL CONFORMAL PREDICTION*