# Corrected Max-Rank for Finite-Sample Conformal Prediction

Filip Schlembach     Evgueni Smirnov     Mark H. M. Winands

Department of Advanced Computing Sciences
Maastricht University

15th Symposium on Conformal and Probabilistic Prediction with Applications
September 2026

# Download Slides

research.schlembach.be

# Outline

# Max-Rank (in 4 easy steps)

# What is Max-Rank?

`Max-Rank` [2] is a method to construct hyperrectangular conformal prediction regions for multivariate regression.

# Inductive Conformal Prediction and Multivariate Regression

We are in the inductive conformal prediction setting.

Let $\mathcal{D}_{cal} = \{\mathbf{x}^{(i)}, \mathbf{y}^{(i)}\}_{i=1}^{n}$ be a calibration set with true labels $\mathbf{y}^{(i)} \in \mathbb{R}^m$,
and $\hat{\mathbf{y}}^{(i)} = \hat{\mathbf{f}}\left(\mathbf{x}^{(i)}\right)$ the associated predicted labels with $\hat{\mathbf{f}}$ an underlying model.

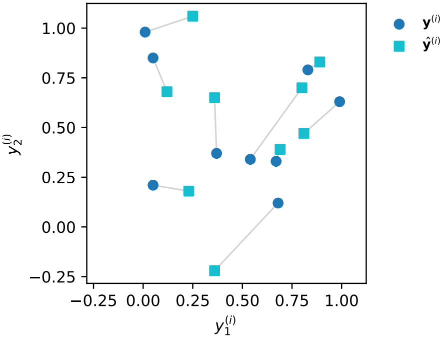



Figure: 2D example of $\mathbf{y}^{(i)}$ and $\hat{\mathbf{y}}^{(i)}$ in the label space

# Max-Rank Step 1: "Internal" Nonconformity Scores

Calculate the "internal" dimension-wise nonconformity scores $\boldsymbol{\beta}^{(i)} = (\beta_1^{(i)}, \ldots, \beta_m^{(i)})$ for all elements of the calibration dataset $\mathcal{D}_{cal}$.

In the example $\beta_j^{(i)} = \left| y_j^{(i)} - \hat{y}_j^{(i)} \right|$.

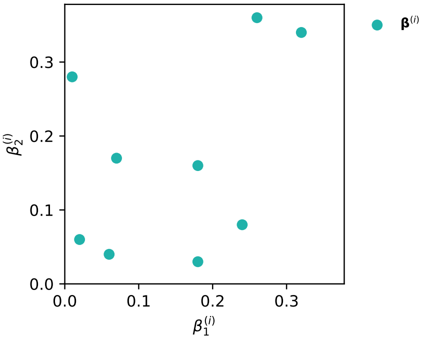



Figure: 2D example of $\boldsymbol{\beta}^{(i)}$ in the score space

# Max-Rank Step 2: Ranks

For all "internal" scores $\beta_j^{(i)}$, determine their rank in dimension $j$

$$\rho_j^{(i)} = \operatorname{rank}\left(\beta_j^{(i)}, \left\{\beta_j^{(k)}\right\}_{k=1}^{n}\right)$$
$$= \left|\left\{k \in \{1, \ldots, n\} : \beta_j^{(k)} \leqslant \beta_j^{(i)}\right\}\right|.$$

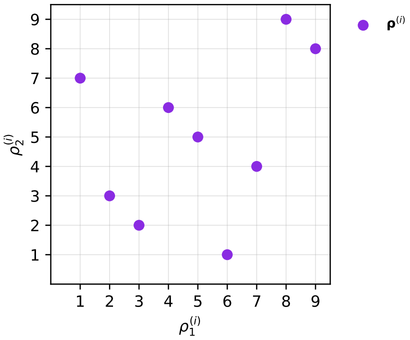



Figure: 2D example of $\rho_j^{(i)}$ in the rank space

# Max-Rank Step 3: Maximum Ranks

For each $i \in \{1, \ldots, n\}$ compute the maximum rank

$$\alpha^{(i)} = \max_{j=1}^{m} \left( \rho_j^{(i)} \right)$$

across all dimensions $j$.

The maximum ranks $\alpha^{(i)}$ **act** like nonconformity scores in the classic inductive conformal prediction sense, however, they are **not**.

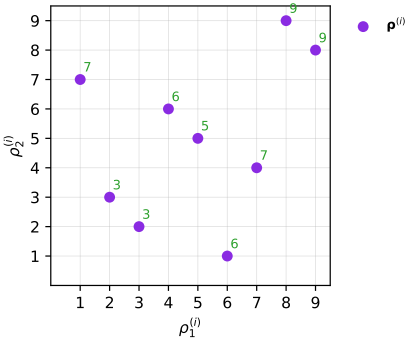



Figure: 2D example of the maximum ranks $\alpha^{(i)}$ in the rank space

# Max-Rank Step 4: Maximum Rank Threshold

For a new test instance $\mathbf{x}^{(n+1)}$ construct the prediction region

$$\Gamma^{\epsilon}\left(\mathbf{x}^{(n+1)}\right) = \left\{\mathbf{y} \in \mathcal{Y} : \alpha^{(n+1)}(\mathbf{y}) \leqslant q\right\} \quad (1)$$

where $q$ is the $\lceil(1-\epsilon)(n+1)\rceil$-smallest value of the maximum ranks $\{\alpha^{(i)}\}_{i=1}^{n}$.

Eq. (1) provides an inclusion rule on the maximum rank of a candidate label that results in valid prediction regions [1, 2].

In the example $q = 7$.

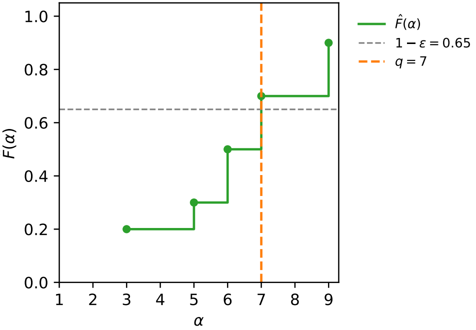



Figure: Example empirical CDF of the maximum ranks $\alpha^{(i)}$.

# Rank Space Prediction Region

The inclusion rule of Eq. (1) is equivalent to

$$\Gamma^{\epsilon}\left(\mathbf{x}^{(n+1)}\right) = \Big\{\mathbf{y} \in \mathcal{Y} : \rho_j^{(n+1)}(\mathbf{y}) \leqslant q \quad \forall j \in \{1, \ldots, m\}\Big\} \qquad (2)$$

in the rank space.

In the example $q = 7$.

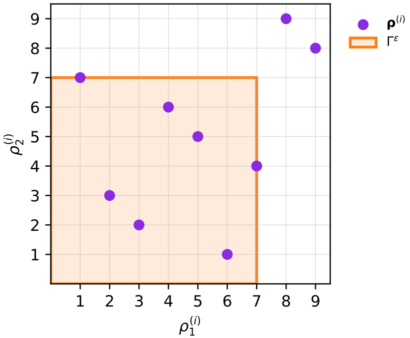



Figure: Prediction region $\Gamma^{\epsilon}$ in the rank space

# Score Space Prediction Region

Timans et al. [2] claim the maximum rank inclusion rule of Eq. (1) translates to

$$\tilde{\Gamma}^{\epsilon}\left(\mathbf{x}^{(n+1)}\right) = \left\{\mathbf{y} \in \mathcal{Y} : \beta_j^{(n+1)}(\mathbf{y}) \leqslant \beta_j^{(\pi_j^{(q)})} \quad \forall j \in \{1, \dots, m\}\right\} \tag{3}$$

in the score space, where $\pi_j$ sorts the internal nonconformity scores $\left\{\beta_j^{(i)}\right\}_{i=1}^{n}$ of dimension $j$ in ascending order.

In the example $(\beta_1^{(\pi_1^{(q)})}, \beta_2^{(\pi_2^{(q)})}) = (0.24,\ 0.28)$.

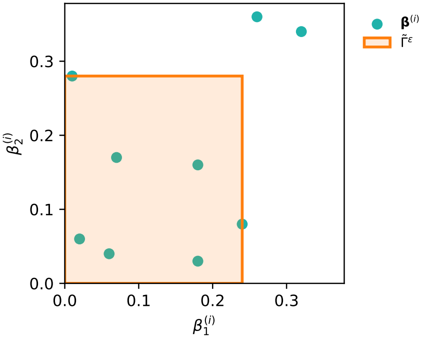



Figure: Prediction region $\tilde{\Gamma}^{\epsilon}$ of Timans et al. in the score space

# Label Space Prediction Region

Using the dimension-wise absolute residuals as the "internal" nonconformity scores, Eq. (3) results in

$$\tilde{\Gamma}^{\epsilon}\left(\mathbf{x}^{(n+1)}\right) = \bigtimes_{j=1}^{m} \left[\hat{y}_j^{(n+1)} - \beta_j^{(\pi_j^{(q)})}, \hat{y}_j^{(n+1)} + \beta_j^{(\pi_j^{(q)})}\right]. \quad (4)$$

$\Rightarrow$ `Max-Rank` is computationally efficient, easy to implement, and produces prediction regions that are less conservative than `Bonferroni CP` on average.

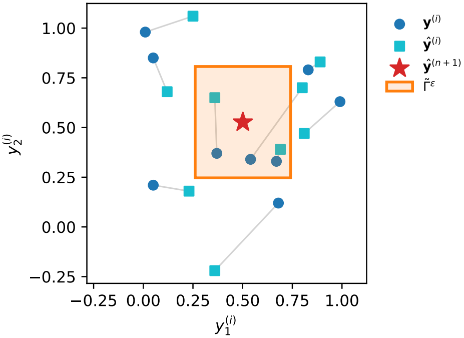



Figure: Prediction region $\tilde{\Gamma}^{\epsilon}$ of Timans et al. in the label space

# Max-Rank Inclusion Rule on Different Levels

$$\Gamma^{\epsilon}\left(\mathbf{x}^{(n+1)}\right) = \left\{\mathbf{y} \in \mathcal{Y} : \alpha^{(n+1)}(\mathbf{y}) \leqslant q\right\}$$ Eq. (1): maximum rank rule

$$\Gamma^{\epsilon}\left(\mathbf{x}^{(n+1)}\right) = \left\{\mathbf{y} \in \mathcal{Y} : \rho_j^{(n+1)}(\mathbf{y}) \leqslant q \ \forall j \in \{1, \dots, m\}\right\}$$ Eq. (2): rank space rule

$$\tilde{\Gamma}^{\epsilon}\left(\mathbf{x}^{(n+1)}\right) = \left\{\mathbf{y} \in \mathcal{Y} : \beta_j^{(n+1)}(\mathbf{y}) \leqslant \beta_j^{\left(\pi_j^{(q)}\right)} \ \forall j \in \{1, \dots, m\}\right\}$$ Eq. (3): score space rule

$$\tilde{\Gamma}^{\epsilon}\left(\mathbf{x}^{(n+1)}\right) = \bigtimes_{j=1}^{m}\left[\hat{y}_j^{(n+1)} - \beta_j^{\left(\pi_j^{(q)}\right)}, \hat{y}_j^{(n+1)} + \beta_j^{\left(\pi_j^{(q)}\right)}\right]$$ Eq. (4): label space rule

# Rank-to-Score Conversion Errors

# Candidate Label in the Label Space

Select a new candidate label $\bar{\mathbf{y}}$ for $\mathbf{x}^{(n+1)}$ that falls just outside the prediction region, i.e.

$$\bar{\mathbf{y}} \notin \tilde{\Gamma}^{\epsilon}\left(\mathbf{x}^{(n+1)}\right) = \bigtimes_{j=1}^{m} \left[\hat{y}_j^{(n+1)} - \beta_j^{(\pi_j^{(q)})}, \hat{y}_j^{(n+1)} + \beta_j^{(\pi_j^{(q)})}\right].$$

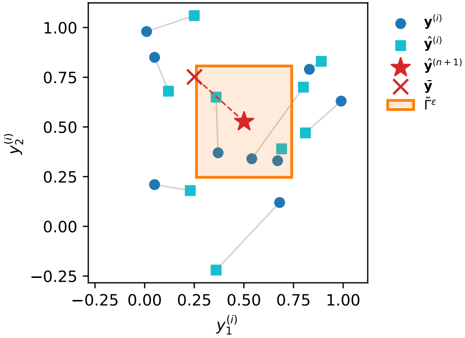



Figure: Candidate label $\bar{\mathbf{y}}$ outside the prediction region $\tilde{\Gamma}^{\epsilon}$ of Timans et al. in the label space

# Candidate Label in the Score Space

According to the score space inclusion rule, $\bar{\mathbf{y}}$ should not be part of the prediction region, i.e.

$$\bar{\mathbf{y}} \notin \tilde{\Gamma}^{\epsilon}\left(\mathbf{x}^{(n+1)}\right) = \left\{\mathbf{y} \in \mathcal{Y} : \beta_j^{(n+1)}(\mathbf{y}) \leqslant \beta_j^{(\pi_j^{(q)})} \quad \forall j \in \{1, \ldots, m\}\right\}$$

with $\beta_j^{(i)} = \left|y_j^{(i)} - \hat{y}_j^{(i)}\right|$,

$(\beta_1^{(\pi_1^{(q)})}, \beta_2^{(\pi_2^{(q)})}) = (0.24,\ 0.28)$, and

$\boldsymbol{\beta}^{(n+1)}(\bar{\mathbf{y}}) = (0.25,\ 0.225)$.

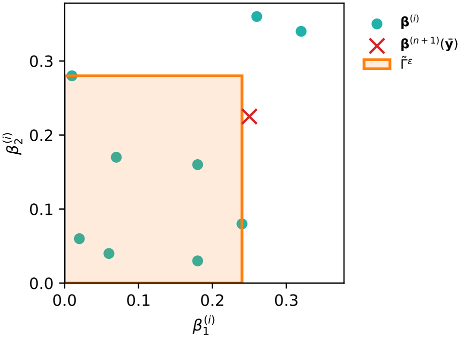



Figure: Candidate label "internal" scores $\boldsymbol{\beta}^{(n+1)}(\bar{\mathbf{y}})$ outside the prediction region $\tilde{\Gamma}^{\epsilon}$ of Timans et al. in the score space

# Candidate Label in the Rank Space

However, according to the rank space inclusion rule, $\bar{\mathbf{y}}$ should be part of the prediction region, i.e.

$$\bar{\mathbf{y}} \in \Gamma^{\epsilon}\left(\mathbf{x}^{(n+1)}\right) = \left\{\mathbf{y} \in \mathcal{Y} : \rho_j^{(n+1)}(\mathbf{y}) \leqslant q \ \forall j \in \{1, \ldots, m\}\right\},$$

with ranks

$$\rho_j^{(i)} = \left|\left\{k \in \{1, \ldots, n\} : \beta_j^{(k)} \leqslant \beta_j^{(i)}\right\}\right|,$$

$\rho^{(n+1)}(\bar{\mathbf{y}}) = (7,\ 6)$, and threshold $q = 7$.

$\Rightarrow$ **Label space and score space rules are not equivalent to the rank space rule.**

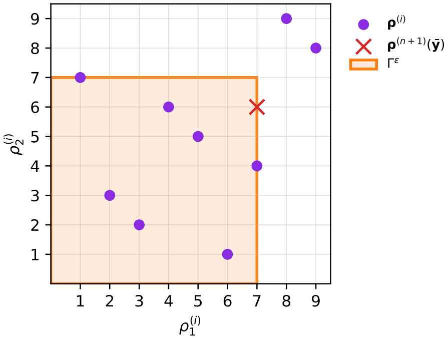



Figure: Candidate label ranks $\rho^{(n+1)}(\bar{\mathbf{y}})$ inside the prediction region $\Gamma^{\epsilon}$ in the rank space

# Maximum Rank of the Candidate Label

According to the maximum rank inclusion rule, $\bar{\mathbf{y}}$ should also be part of the prediction region, i.e.

$$\bar{\mathbf{y}} \in \Gamma^{\epsilon}\left(\mathbf{x}^{(n+1)}\right) = \left\{\mathbf{y} \in \mathcal{Y} : \alpha^{(n+1)}(\mathbf{y}) \leqslant q\right\},$$

with $\alpha^{(n+1)}(\bar{\mathbf{y}}) = \max(\{7, 6\}) = 7 \leqslant 7 = q$.

$\Rightarrow$ **Label space and score space rules are not equivalent to the rank space and maximum rank rules.**

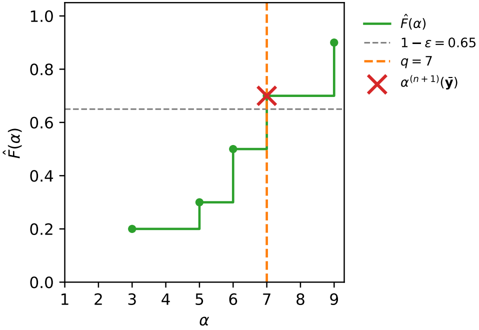
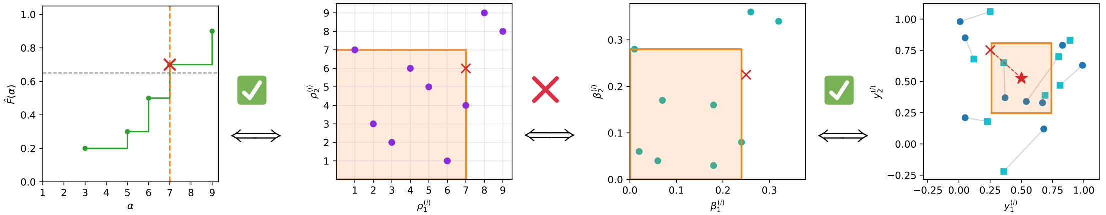



Figure: Maximum rank of the candidate label $\alpha^{(n+1)}(\bar{\mathbf{y}})$ and the empirical CDF of maximum ranks.

# Equivalence of Timans et al.'s Inclusion Rules




Maximum Ranks | Rank Space | Score Space | Label Space

# Corrected Max-Rank Implementation

# Corrected Score Space Prediction Region

The rank space rule requires

$$\rho_j^{(n+1)}(\bar{\mathbf{y}}) \leqslant q$$

$$\Leftrightarrow \left|\left\{k \in \{1, \dots, n\} : \beta_j^{(k)} \leqslant \beta_j^{(n+1)}(\mathbf{y})\right\}\right| \leqslant q$$

$$\Leftrightarrow \beta_j^{(n+1)}(\mathbf{y}) < \beta_j^{(\pi_j^{(q+1)})},$$

with $\beta_j^{(\pi_j^{(q+1)})} = +\infty$ if $q + 1 > n$.

Hence, the corrected score space rule is

$$\Gamma^\epsilon\left(\mathbf{x}^{(n+1)}\right) = \Big\{\mathbf{y} \in \mathcal{Y} : \beta_j^{(n+1)}(\mathbf{y}) \leqslant \beta_j^{(\pi_j^{(q+1)})} \quad \forall j \in \{1, \dots, m\}\Big\}. \tag{5}$$

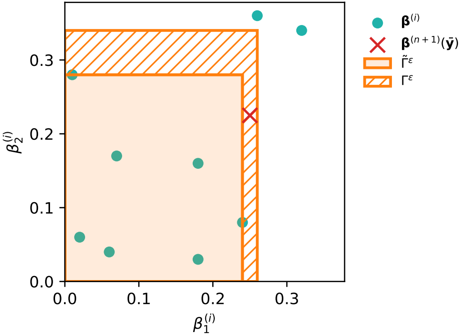



Figure: Corrected Max-Rank prediction region $\Gamma^\epsilon$ in the score space

# Corrected Label Space Prediction Region

Using the dimension-wise absolute residuals as the "internal" nonconformity scores, Eq. (5) results in

$$\Gamma^{\epsilon}\left(\mathbf{x}^{(n+1)}\right) = \bigtimes_{j=1}^{m} \left[\hat{y}_j^{(n+1)} - \beta_j^{(\pi_j^{(q+1)})}, \hat{y}_j^{(n+1)} + \beta_j^{(\pi_j^{(q+1)})}\right], \quad (6)$$

with $\beta_j^{(\pi_j^{(q+1)})} = +\infty$ if $q+1 > n$.

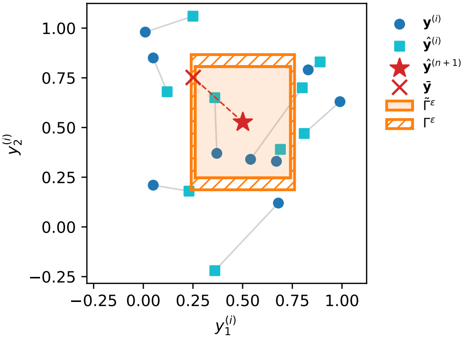
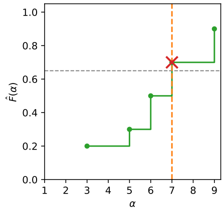


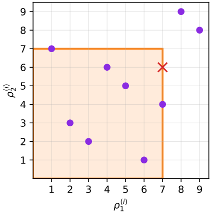

Figure: Corrected `Max-Rank` prediction region $\Gamma^{\epsilon}$ in the label space

# Equivalence of the Corrected Inclusion Rules



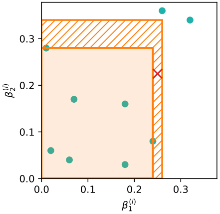

Maximum Ranks

$\iff$



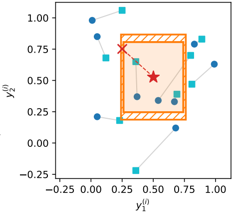

Rank Space

$\iff$




Score Space

$\iff$




Label Space

# Corrected Max-Rank Inclusion Rule on Different Levels

$$\Gamma^{\epsilon}\left(\mathbf{x}^{(n+1)}\right) = \left\{\mathbf{y} \in \mathcal{Y} : \alpha^{(n+1)}(\mathbf{y}) \leqslant q\right\}$$ Eq. (1): maximum rank rule

$$\Gamma^{\epsilon}\left(\mathbf{x}^{(n+1)}\right) = \left\{\mathbf{y} \in \mathcal{Y} : \rho_j^{(n+1)}(\mathbf{y}) \leqslant q \ \forall j \in \{1, \ldots, m\}\right\}$$ Eq. (2): rank space rule

$$\Gamma^{\epsilon}\left(\mathbf{x}^{(n+1)}\right) = \left\{\mathbf{y} \in \mathcal{Y} : \beta_j^{(n+1)}(\mathbf{y}) \leqslant \beta_j^{(\pi_j^{(q+1)})} \ \forall j \in \{1, \ldots, m\}\right\}$$ Eq. (5): corrected score space rule

$$\Gamma^{\epsilon}\left(\mathbf{x}^{(n+1)}\right) = \times_{j=1}^{m}\left[\hat{y}_j^{(n+1)} - \beta_j^{(\pi_j^{(q+1)})}, \hat{y}_j^{(n+1)} + \beta_j^{(\pi_j^{(q+1)})}\right]$$ Eq. (6): corrected label space rule

with $\beta_j^{(\pi_j^{(q+1)})} = +\infty$ if $q + 1 > n$.

# Experiments

# Experimental Setup

Draw $2n$ values $(\hat{\mathbf{y}}^{(i)} - \mathbf{y}^{(i)}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\sigma})$ from an $m$-dimensional multivariate normal distribution, where $\boldsymbol{\sigma}$ is an $m \times m$ matrix with values equalling $1$ on the main diagonal and $\sigma$ otherwise.

The absolute residuals serve as the "internal" nonconformity score function, such that $\beta_j^{(i)} = |\hat{y}_j^{(i)} - y_j^{(i)}|$.

All experiments are repeated $1000$ times and the results are averaged.

# Experimental Results I: Example Configuration in Detail

- Calibration set size
  $n = 25$
- Number of $\mathcal{Y}$ dimensions
  $m = 4$
- Correlation coefficient
  $\sigma = 0.5$

`Corrected Max-Rank` is valid for every $\epsilon \in \{0.005, 0.010, \ldots, 0.995\}$.

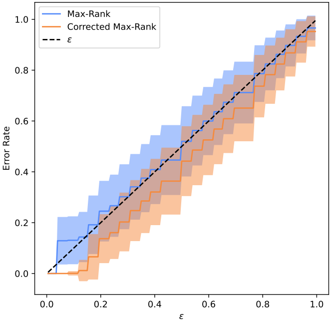



Figure: Mean error rate

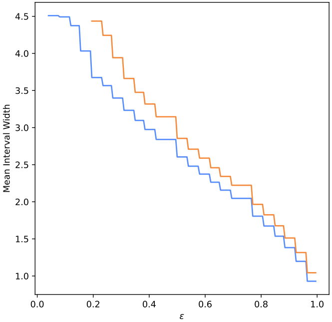



Figure: Mean interval width

# Metrics

$$\text{Mean excess error rate} = \frac{1}{|E|}\sum_{\epsilon \in E}\left(\underbrace{\frac{1}{|\mathcal{D}_{test}|}\left(\sum_{(\mathbf{x}^{(i)},\mathbf{y}^{(i)})\in\mathcal{D}_{test}} \mathbb{1}\left(\mathbf{y}^{(i)} \notin \Gamma^{\epsilon}\left(\mathbf{x}^{(i)}\right)\right)\right)}_{\text{empirical error rate for significance level } \epsilon} - \epsilon\right)$$

$$\text{Mean relative interval width difference} = \frac{1}{|E|}\sum_{\substack{\epsilon \in E \\ \epsilon \geqslant N}}\left(\underbrace{\frac{1}{|\mathcal{D}_{test}|}\left(\sum_{(\mathbf{x}^{(i)},\mathbf{y}^{(i)})\in\mathcal{D}_{test}} \sum_{j=1}^{m} \frac{\left|\Gamma_j^{\epsilon}\left(\mathbf{x}^{(i)}\right)\right|}{\left|\tilde{\Gamma}_j^{\epsilon}\left(\mathbf{x}^{(i)}\right)\right|}\right)}_{\text{relative interval width for significance level } \epsilon} - 1\right)$$

where $E = \{0.005, 0.010, \ldots, 0.995\}$ is the set of tested significance levels, and $N = (m+1)/(n+1)$ ensures no infinite intervals are encountered.

# Experimental Results II: Varying Correlation

- Calibration set size
  $n = 200$
- Number of $\mathcal{Y}$ dimensions
  $m = 4$
- Correlation coefficient
  $\sigma \in \{0, 0.05, 0.10, \ldots, 1.0\}$

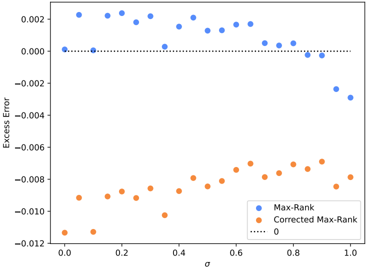



Figure: Mean excess error rate

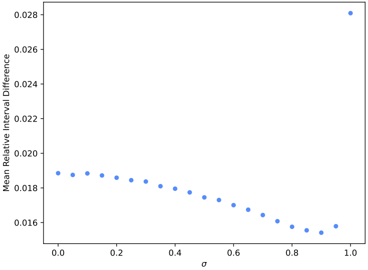



Figure: Mean relative interval width difference

# Experimental Results III: Varying Number of Dimensions

- Calibration set size
  $n = 200$
- Number of $\mathcal{Y}$ dimensions
  $m \in \{2, 3, \ldots, 12\}$
- Correlation coefficient
  $\sigma = 0.5$

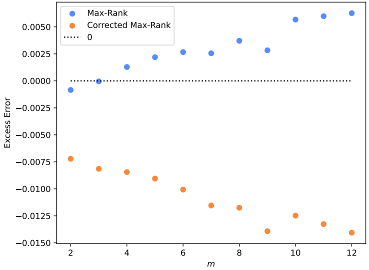



Figure: Mean excess error rate

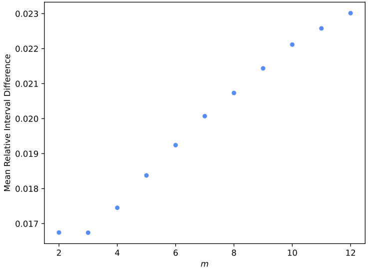



Figure: Mean relative interval width difference

# Experimental Results IV: Varying Calibration Set Size

- Calibration set size $n \in \{10, 25, 50, 100, 200, 400, 800, 1600, 3200, 6400\}$
- Number of $\mathcal{Y}$ dimensions $m = 4$
- Correlation coefficient $\sigma = 0.5$

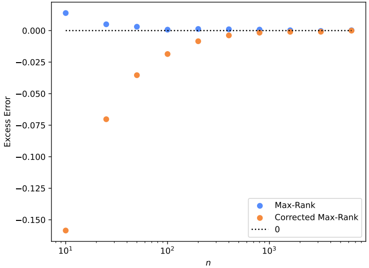



Figure: Mean excess error rate

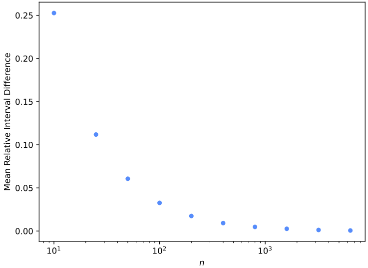



Figure: Mean relative interval width difference

# Conclusion

# Conclusion

- The `Max-Rank` rank space rule produces valid hyperrectangular prediction regions, is generally more efficient than `Bonferroni CP`, and is cheap to compute.
- The original score space implementation of `Max-Rank` [2]
  - translates the rank threshold $q$ into the $q$th smallest calibration score, and
  - does not implement the convention for infinite prediction regions.
- We propose the `Corrected Max-Rank` implementation
  - that uses the $(q+1)$st smallest calibration score, and
  - implements the infinite prediction region convention.
- The `Corrected Max-Rank` implementation restores the finite-sample validity guarantee at the cost of slightly larger prediction regions.

Thank you! Questions?

# Bibliography I


[1] Filip Schlembach, Evgueni Smirnov, and Mark H. M. Winands. "Corrected Max-Rank for Finite-Sample Conformal Prediction". In: *Proceedings of the Fifteenth Symposium on Conformal and Probabilistic Prediction with Applications*. Ed. by Ernst Ahlberg et al. Vol. 329. Proceedings of Machine Learning Research. PMLR, Sept. 2026.

[2] Alexander Timans et al. *Max-Rank: Efficient Multiple Testing for Conformal Prediction*. _eprint: 2311.10900. Mar. 2025. URL: https://arxiv.org/abs/2311.10900.