Conformal Prediction Sets with Limited False Positives

Adam Fisch, Tal Schuster, Tommi Jaakkola, and Regina Barzilay

Confident set-valued predictions

- Conformal prediction identifies a small *set* of promising output candidates.
- This set is guaranteed to contain the correct answer with high probability.

Conformal prediction framework

- Given n exchangeable examples $(X_i, Y_i) \in \mathcal{X} \times \mathcal{Y}$ and a desired significance level ϵ, for a new input X_{n+1}, return a **set of predictions** $C_\epsilon(X_{n+1}) \subseteq \mathcal{Y}$.

- A predictor is **valid** if $C_\epsilon(X_{n+1})$ covers the correct label Y_{n+1} w.p. at least $1 - \epsilon$:

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

→ {**Golden retriever**, Labrador}

The catch: guaranteed coverage doesn't come for free

- A classifier is only **efficient** if the output set is small, $\mathbb{E}[\,|C_\epsilon(X)|\,] \ll |\mathcal{Y}|$.

- To meet the desired coverage, output sets may be forced to include **false positives** that can't be otherwise ruled out (with high confidence).

- **This is problematic if having too many false positives has substantial cost.**

Our goal: Can we trade guarantees on *coverage* for guarantees on *false positives*?

- Controlling accuracy depends on how hard the task is—hard tasks have low efficiency.

- Can we at least guarantee *actionable* predictions with low amounts of noise?

Conformal prediction sets with limited false positives

- Proposal: change the setting to a **constrained optimization problem**.
- We want to maximize accuracy, but respect a false positive budget.
- Work in a generalized multi-label setting where input $X \in \mathcal{X}$ is associated with a *true positive set* $Z \subseteq \mathcal{Y}$, where Z is a set containing any number of correct labels (or none!).

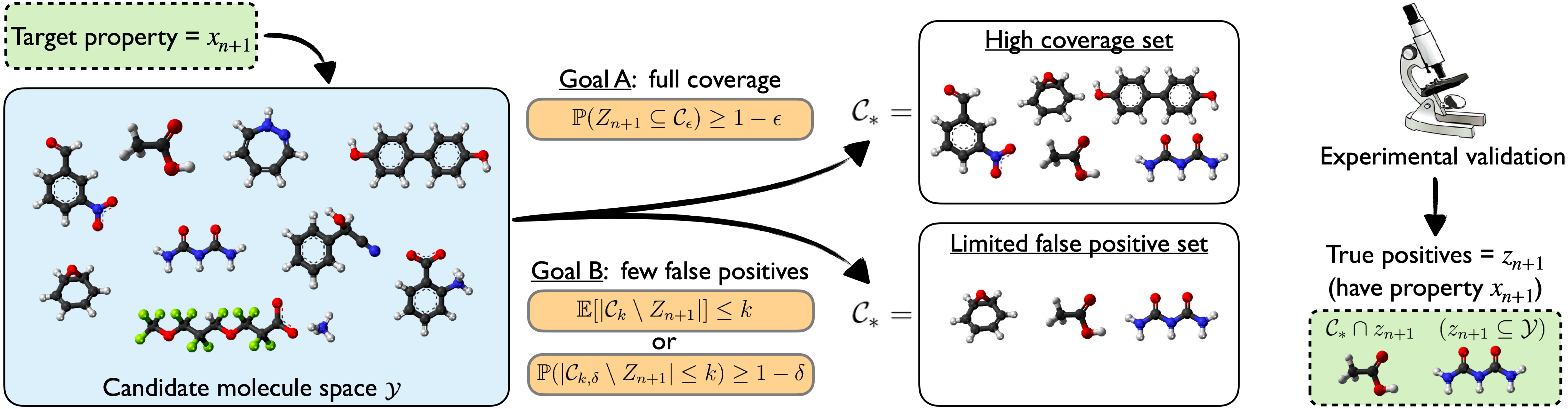
Control in expectation:

$$\text{maximize } \mathbb{E}\left[\frac{|\mathcal{C}_k(X_{n+1}) \cap Z_{n+1}|}{\max(|Z_{n+1}|, 1)}\right] \text{ s.t. } \mathbb{E}\Big[|\mathcal{C}_k(X_{n+1}) \setminus Z_{n+1}|\Big] \leq k$$

Control in probability:

$$\text{maximize } \mathbb{E}\left[\frac{|\mathcal{C}_{k,\delta}(X_{n+1}) \cap Z_{n+1}|}{\max(|Z_{n+1}|, 1)}\right] \text{ s.t. } \mathbb{P}\Big(|\mathcal{C}_{k,\delta}(X_{n+1}) \setminus Z_{n+1}| \leq k\Big) \geq 1 - \delta$$

Application to experimental design (in-silico screening)



- In-silico screening uses computational tools to identify drugs with desired properties.
- For a given property, many such drugs may exist, or none.
- Any candidate that is flagged via in-silico screening must be validated experimentally.
- Limiting false positives is critical when balancing an experimental budget.

An oracle set predictor

- Imagine an *oracle* with access to $P_{Z|Y}$, the true conditional distribution of $Z \mid X$.

- This would be able to exactly maximize our goal:

$$\mathcal{C}_k^{\text{oracle}}(x) := \underset{\mathcal{S} \in 2^{\mathcal{Y}}}{\operatorname{argmax}} \left\{\mathbb{E}[\mathrm{TPP}(Z, \mathcal{S}) \mid x] \colon \mathbb{E}[\mathrm{FP}(Z, \mathcal{S}) \mid x] \leq k\right\}$$

$$\mathcal{C}_{k,\delta}^{\text{oracle}}(x) := \underset{\mathcal{S} \in 2^{\mathcal{Y}}}{\operatorname{argmax}} \left\{\mathbb{E}[\mathrm{TPP}(Z, \mathcal{S}) \mid x] \colon \mathbb{P}(\mathrm{FP}(Z, \mathcal{S}) \mid x] > k) < \delta\right\}$$

- But of course, this isn't practical (as we don't know $P_{Z|X}$).

A calibrated approximate set predictor

- Using training data, we *learn* a model to directly predict the # FP in a candidate set $S \mid X = x$.

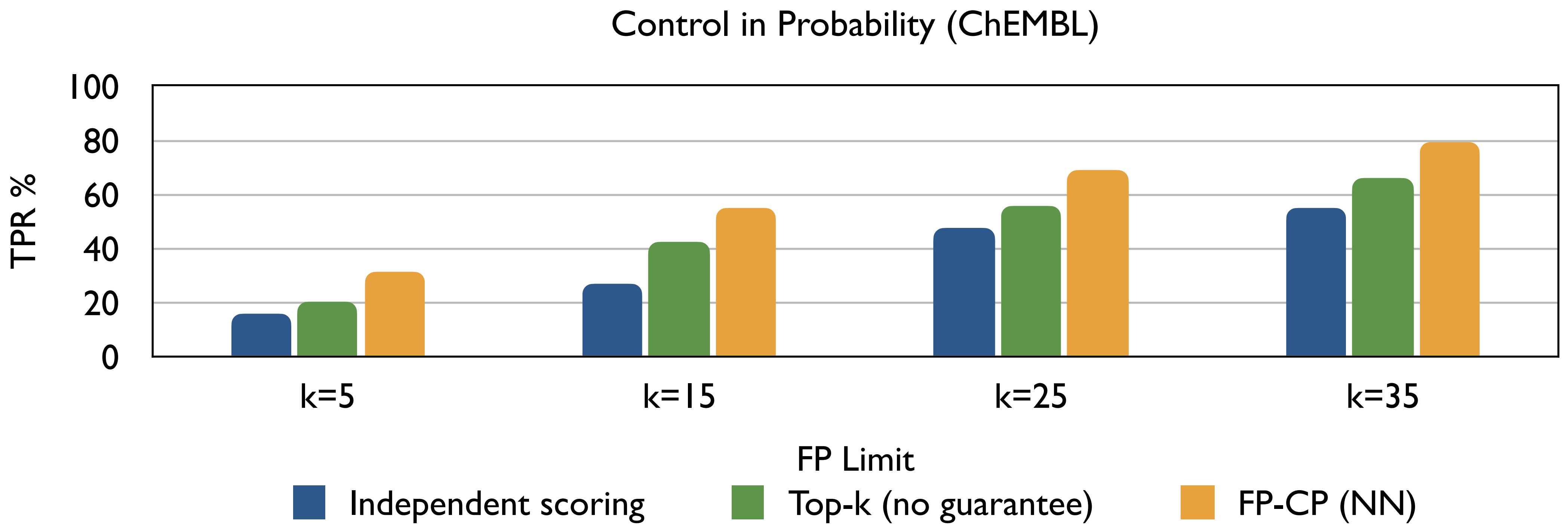
- We first greedily consider a subset of candidate sets that are formed by individually ranking labels $y \in \mathcal{Y}$ using some auxiliary model, such as a model of $p_\theta(y \in Z \mid x)$.

- As we consider each progressively larger set (top 1, top 2, $\ldots$), we try to directly *predict* the number of false positives (like a confidence score!) using a DeepSet NN.

- To maximize coverage of true positives, we take the *largest* candidate set whose predicted FP score is below a threshold that we calibrate to guarantee our desired type of FP control.

Overview of results

- We show that our calibration procedure can be used to guarantee false positive control.
- Empirically, we also show that we can still achieve high true positive rates with low # FPs.
- Our DeepSet model is effective, and leads to both better conditional error and higher TPR.



Summary

- Conformal prediction grants theoretical coverage guarantees.
- But naive application of conformal prediction can sometimes yield disappointing results in practice, if the output sets are simply too large to know what to do with!
- Our method (1) offers control over the number of false positives, (2) still empirically achieves strong true discovery rates in most cases, and (3) is simple to calibrate and implement.

Questions?

Come talk to us at our poster!