Online Locally Differentially Private Conformal Prediction via Binary Inquiries

Qiangqiang Zhang

joint with Chenfei Gu, Xinwei Feng, Jinhan Xie, Ting Li

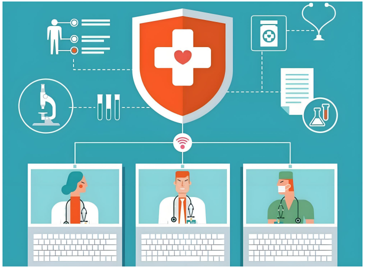
Shandong University

November 6, 2025

Outline

Background and Motivation

The Era of Streaming Data



Motivation

Conformal prediction (Vovk et al., 2005) is a general framework for constructing prediction sets $\widehat{C}$ with

- Finite-sample coverage guarantee (exact)
- For any data distribution (distribution-free)
- For any predictive model (model-free)

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

Two main challenges:

- Streaming data & distribution shift
- Privacy leakage

Related work

Online Conformal Prediction

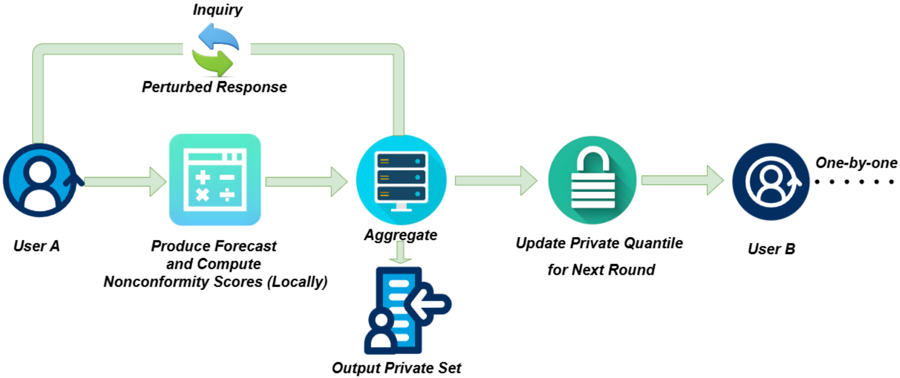
- Adaptive conformal inference (ACI), first link between online convex optimization and CP (Gibbs and Candes, 2021).
- A parameter-free that adaptively builds upon ACI (Zaffran et al., 2022; Podkopaev et al., 2024).
- An alternative expert selection scheme for choosing the stepsize in ACI (Gibbs and Candès, 2024).

Privacy-Preserving Conformal Prediction

- Private prediction sets via exponential mechanism (Angelopoulos et al., 2022).
- Private prediction sets within FL that addresses label shift between agents (Plassier et al., 2023).
- Two complementary CP under LDP (i) label perturbation via k-ary randomized response, and (ii) binary indicators for privatized conformity scores (Penso et al., 2025).

Gap: distribution-free, privacy-preserving uncertainty quantification in streaming data.

Our framework



We propose an online CP framework under Local Differential Privacy(LDP) for streaming environments.

1. Algorithm: one-pass, constant time and space complexity ($\mathcal{O}(1)$ per instance).
2. Privacy-preserving: either noise-based or randomized response.

Preliminaries

Review of CP

Key idea: CP converts any heuristic notion of uncertainty into a rigorous uncertainty measure.

Procedure:

1. Identify a heuristic notion of uncertainty using the pre-trained model.
2. Define a non-conformity score function $s(x, y)$ (larger $\Rightarrow$ worse fit).
3. Use calibration data to compute quantile $\widehat{q}$.
4. Prediction set:

$$\widehat{C}(X_{\text{test}}) = \{y : s(X_{\text{test}}, y) \leq \widehat{q}\}$$

.

Coverage Guarantee:

$$1 - \alpha \ \leq \ \mathbb{P}\big(Y_{\text{test}} \in \widehat{C}(X_{\text{test}})\big) \ \leq \ 1 - \alpha + \tfrac{1}{n+1}.$$

Validity requires: i.i.d. data and an independent calibration set.

Setup

- At time point $t \geq 1$ with t observations arrive sequentially:

$$\{(X_t, Y_t)\}_{t\geq 1}, \quad X_t \in \mathcal{X} \subset \mathbb{R}^d, \; Y_t \in \mathcal{Y} \subset \mathbb{R}.$$

- Non-conformity score function: $S_t : \mathcal{X} \times \mathcal{Y} \to [0, D]$.
- For example, a common choice for regression is the absolute residual:

$$S_t(X_t, Y_t) = |Y_t - \widehat{f}_t(X_t)|,$$

where $\widehat{f}_t : \mathcal{X} \to \mathbb{R}$ is the predictive model trained in an online manner.

- Goal: Construct a private prediction set $\widehat{C}_t$ at each time t such that the long-run empirical coverage converges to a target level $1 - \alpha$:

$$\lim_{T\to\infty} \left| \frac{1}{T} \sum_{t=1}^{T} \mathbb{I}\big(Y_t \in \widehat{C}_t(X_t)\big) - (1 - \alpha) \right| = 0. \tag{1}$$

Differential Privacy

Central Differential Privacy (Dwork et al., 2006)

Let $\mathcal{X}$ be the sample space for an individual data, a randomized algorithm M: $\mathcal{X}^n \to \mathcal{R}$ is (ϵ, δ)-DP if and only if for every pair of adjacent datasets X, $X' \subset \mathcal{X}^n$ and for any measurable event $E \subseteq \mathcal{R}$, the inequality below holds:

$$\Pr(M(X) \in E) \leq e^{\epsilon} \cdot \Pr(M(X') \in E) + \delta.$$

When $\delta = 0$, then M is called pure ϵ-DP.

Local Differential Privacy (Xiong et al., 2020)

A randomized algorithm $M : \mathcal{X} \to \mathcal{R}$ satisfies (ϵ, δ)-LDP if, for any pair of inputs $x, x' \in \mathcal{X}$ and for any measurable event $E \subseteq \mathcal{R}$, we have

$$\Pr(M(x) \in E) \leq e^{\epsilon} \cdot \Pr(M(x') \in E) + \delta.$$

When $\delta = 0$, this reduces to pure ϵ-LDP.

Methodology

Quantile Estimation

- Coverage event: $\{Y_t \in \widehat{C}_t(X_t)\} \Leftrightarrow \{S_t \leq q_t\}$, where $q_{t+1} = q_t + \eta_t(\mathbb{I}(Y_t \notin \widehat{C}_t(X_t)) - \alpha)$.
- Target property:

$$\lim_{T\to\infty} \left| \frac{1}{T} \sum_{t=1}^{T} \mathbb{I}\{S_t \leq q_t\} - (1-\alpha) \right| = 0.$$

- Reduces to learn $(1-\alpha)$-quantile of scores $(S_t)_{t\geq 1}$.
- q_t can also be interpreted as an online (sub)gradient descent algorithm on the pinball loss:

$$\ell_{1-\alpha}(q_t, S_t) = (\mathbb{I}\{q_t \geq S_t\} - (1-\alpha))(q_t - S_t).$$

- Subgradient:

$$\partial \ell_{1-\alpha}(q, S_t) = \begin{cases} \mathbb{I}\{S_t \leq q\} - (1-\alpha), & q \neq S_t, \\ [-(1-\alpha), \alpha], & q = S_t. \end{cases}$$

Private Quantile Estimator

- Binary-response estimator: $\mathbb{I}\{S_t \leq q\}$, a binary (yes/no) structure.

Algorithm Locally Randomized Binary Response (LRBR, Liu et al. (2023))

Input: local nonconformity score S, private quantile q, response rate r
$u \sim \text{Bernoulli}(r)$
$v \sim \text{Bernoulli}(0.5)$
if $u = 1$ **then**
  return $1_{\{q > S\}}$
end
else
  return v
end

Randomizer and Conversion Table

Theorem

Algorithm LRBR is an $(\epsilon, 0)$-randomizer with

$$\epsilon = \log\left(\frac{1+r}{1-r}\right).$$

r	ϵ	r	ϵ
0	0.00	0.5	1.10
0.05	0.10	0.55	1.24
0.1	0.20	0.6	1.39
0.15	0.30	0.65	1.55
0.2	0.40	0.7	1.73
0.25	0.51	0.75	1.95
0.3	0.62	0.8	2.20
0.35	0.73	0.85	2.51
0.4	0.85	0.9	2.94
0.45	0.97	0.95	3.66

Conversion table between r and ϵ

Coin betting framework (Orabona and Pál, 2016):

- **Intuition**: Learning as repeated betting — a gambler adapts bets based on feedback while tracking wealth.
- Wealth update:

$$W_t = W_0 + \sum_{i=1}^{t} \lambda_i W_{i-1} c_i$$

  where a fraction $\lambda_t \in [-1, 1]$ of wealth W_{t-1} is bet on outcome $c_t \in [-1, 1]$.
- Adaptive betting fraction:

$$\lambda_{t+1} = \frac{1}{t+1} \sum_{i=1}^{t} c_i \quad \text{(KT estimator; Krichevsky and Trofimov (1981))}\,.$$

- Privatized quantile estimation: Feedback $c_t := -\widehat{g}_t$; signed bet

$$w_{t+1} = \lambda_{t+1} W_t$$

  serves as the updated privatized quantile estimate q_{t+1}.

Continued

Motivated by LRBR, we define an Adjusted Feedback:

$$g_t = \begin{cases} 1 - (r(1-\alpha) + 0.5(1-r)), & \text{if } L = 1 \\ -(r(1-\alpha) + 0.5(1-r)), & \text{otherwise} \end{cases} \tag{2}$$

Observe:

$$\mathbb{E}[g_t] = r \cdot (\mathbb{I}\{q_t > S_t\} - (1-\alpha)),$$

where r is the response rate.

Algorithm Private Online Conformal Prediction

Input: Data stream $\{(X_t, Y_t)\}_{t\geq 1}$; Response rate $r_t \in (0,1)$; Miscoverage level $\alpha \in (0,1)$; Initialize $W_0 = 1$, $\lambda_1 = 0$, $q_1 = 0$

for $t = 1, 2, \ldots$ **do**

- Predict $\widehat{Y}_t \leftarrow \widehat{f}_t(X_t)$
- **Output:** the private prediction interval: $\widehat{C}_t := [\widehat{Y}_t - q_t, \ \widehat{Y}_t + q_t]$;
- Nonconformity score: $S_t = |Y_t - \widehat{Y}_t|$
- Inquiry LRBR response: $L \leftarrow \mathrm{LRBR}(S_t, q_t, r_t)$
- **if** $L = 1$ **then**
  - $g_t \leftarrow 1 - (r_t(1-\alpha) + 0.5(1-r_t))$
- **else**
  - $g_t \leftarrow -(r_t(1-\alpha) + 0.5(1-r_t))$
- **end**
- Update: $W_t \leftarrow W_{t-1} - g_t q_t$
- Update: $\lambda_{t+1} \leftarrow \frac{t}{t+1}\lambda_t - \frac{1}{t+1} g_t$
- Update quantile estimate: $q_{t+1} \leftarrow \lambda_{t+1} W_t$

end

Theoretical Properties

Privacy Guarantee

Theorem

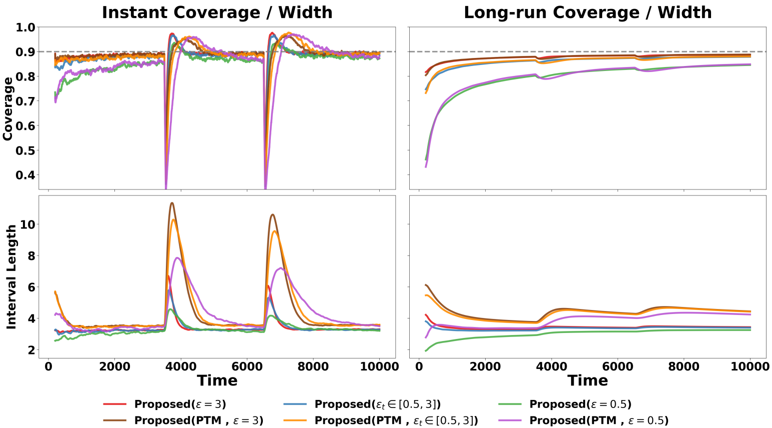
After t updates, Algorithm satisfies $(\max_{1\le j\le t}\epsilon_j, 0)$*-LDP, where* $t \ge 1$.

Theorem

Suppose the predictive model itself satisfies $(\max_{1\le j\le t}\varepsilon_j, \max_{1\le j\le t}\zeta_j)$*-LDP. Then the entire pipeline, mapping private individual data to the final prediction interval* $\widehat{C}_t$*, inherits LDP guarantees. Specifically, Algorithm ensures* $(\max_{1\le j\le t}(\epsilon_j+\varepsilon_j), \max_{1\le j\le t}\zeta_j)$*-LDP, where* $t \ge 1$.

Continued

- Fixed $\epsilon = 0.5$ vs. $\epsilon = 3$ vs. random $\epsilon_t \in [0.5, 3]$.
- Result: Coverage and width behave similarly — privacy bound depends only on $\max_t \epsilon_t$.



Long-run Coverage Guarantee

Theorem

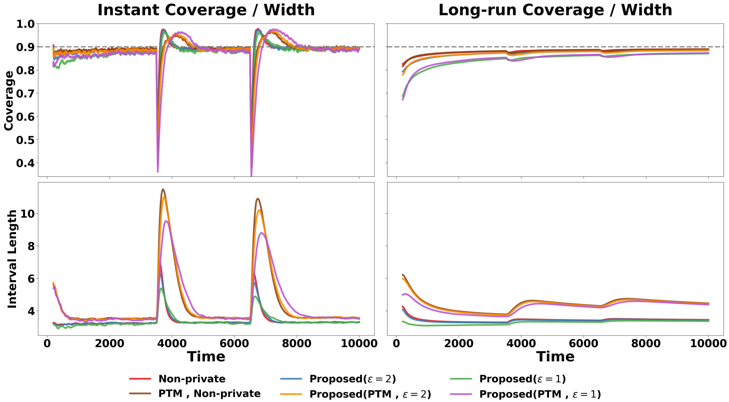
Let the target miscoverage level be fixed at $\alpha \in (0, 1/2)$. Assume the non-conformity scores are bounded, $S_t \in [0, D]$ for all $t \geq 1$, with $D > 0$. Then the procedure in Algorithm satisfies

$$\lim_{T\to\infty} \left| \frac{1}{T} \sum_{t=1}^{T} \mathbb{I}\Big(Y_t \in \widehat{C}_t(X_t)\Big) - (1-\alpha) \right| = 0.$$

Continued

Comparison: private (PTMs) vs non-private trained models.

- Coverage: Long-run coverage ≈ unaffected.
- Interval width: PTMs $\downarrow$ accuracy $\Rightarrow$ larger S_t $\Rightarrow$ wider intervals.



Numerical Studies

Simulation Setup: regression problem

Generate data as follows:

$$x_t \sim \mathcal{N}(0, I_5), \quad y_t = x_t^\top \beta_t + \varepsilon_t, \quad t = 1, \ldots, 10{,}000,$$

with $\beta_t \in \mathbb{R}^5$ and Gaussian noise ε_t.

Consider three scenarios:

1. Abrupt shifts:

$$\beta^{(1)} = (1, 2, 1, 0, 0),\ \beta^{(2)} = (0, -1, -2, -1, 0),\ \beta^{(3)} = (0, 0, 1, 2, 1).$$

   - Case A: homoskedastic $\varepsilon_t \sim \mathcal{N}(0, 1)$
   - Case B: heteroskedastic $\varepsilon_t = x_{t,1}^2 \eta_t,\ \eta_t \sim \mathcal{N}(0, 1)$

2. Smooth shifts:

$$\beta_t = (1 - \alpha_t)\beta_{\text{start}} + \alpha_t \beta_{\text{end}}, \quad \alpha_t = \frac{t-1}{n-1},$$

with $\beta_{\text{start}} = (1, 2, 1, 0, 0)$, $\beta_{\text{end}} = (0, 0, 1, 2, 1)$. Noise: homoskedastic $\varepsilon_t \sim \mathcal{N}(0, 1)$ (Case C).

Simulation Results

Table: Coverages and widths for the proposed method, DPCP (Angelopoulos et al., 2022) and DtACI (Gibbs and Candès, 2024).

Method	no-privacy		$\epsilon = 3$		$\epsilon = 1$		$\epsilon = 0.5$	
	Coverage	Width	Coverage	Width	Coverage	Width	Coverage	Width
Case A								
Proposed	0.890 (0.03)	3.43 (0.30)	0.889 (0.30)	3.42 (0.40)	0.875 (1.00)	3.36 (1.20)	0.853 (2.10)	3.28 (2.70)
DPCP	0.900 (0.51)	8.29 (1.02)	0.904 (0.52)	8.40 (1.07)	0.911 (0.61)	8.60 (1.48)	0.922 (0.90)	8.95 (2.96)
DtACI	0.897 (0.09)	3.47 (0.30)	*	*	*	*	*	*
Case B								
Proposed	0.890 (0.05)	4.46 (0.96)	0.889 (0.26)	4.44 (1.30)	0.874 (0.90)	4.22 (2.91)	0.850 (1.82)	3.83 (5.06)
DPCP	0.900 (0.54)	9.00 (1.45)	0.904 (0.53)	9.14 (1.56)	0.911 (0.60)	9.40 (2.19)	0.923 (0.93)	9.89 (4.55)
DtACI	0.899 (0.07)	4.76 (1.02)	*	*	*	*	*	*
Case C								
Proposed	0.890 (0.03)	3.26 (0.30)	0.889 (0.27)	3.26 (0.37)	0.875 (1.06)	3.21 (1.03)	0.852 (1.92)	3.11 (1.89)
DPCP	0.900 (0.50)	4.43 (0.63)	0.904 (0.51)	4.49 (0.60)	0.911 (0.61)	4.60 (0.88)	0.920 (1.01)	4.80 (1.99)
DtACI	0.899 (0.07)	3.34 (0.27)	*	*	*	*	*	*

Case A

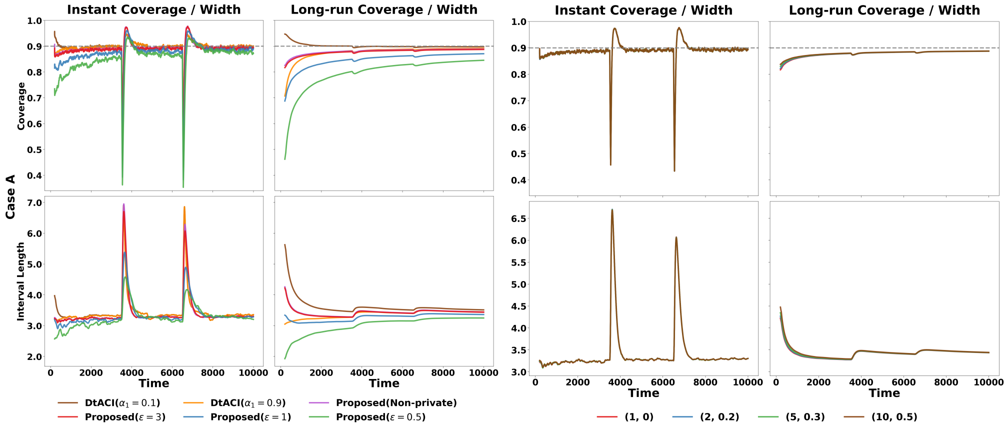
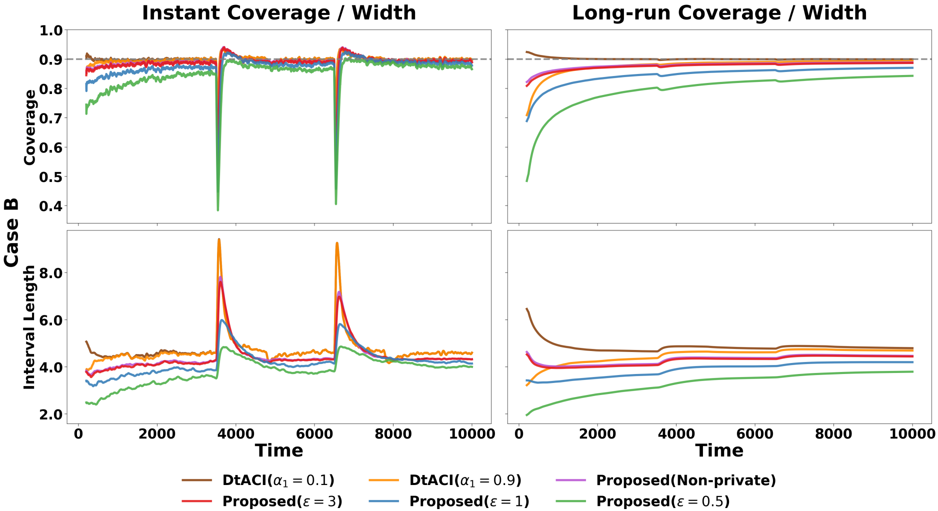
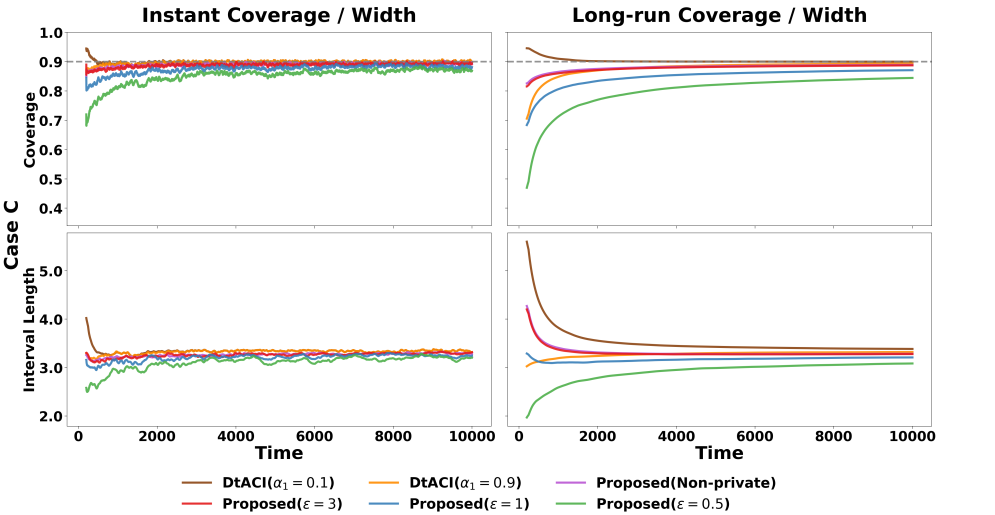


Figure: Case A. Left panels: comparison of methods by coverage, interval width, long-run coverage, and long-run width. Right panels: impact of initialization settings (W_0, λ_1) for the proposed method with $\epsilon = 3$.

Case B


Instant Coverage / Width
Long-run Coverage / Width
Case B
Coverage
Interval Length
1.0
0.9
0.8
0.7
0.6
0.5
0.4
8.0
6.0
4.0
2.0
0
2000
4000
6000
8000
10000
Time
DtACI($\alpha_1 = 0.1$)
DtACI($\alpha_1 = 0.9$)
Proposed(Non-private)
Proposed($\varepsilon = 3$)
Proposed($\varepsilon = 1$)
Proposed($\varepsilon = 0.5$)

Case C

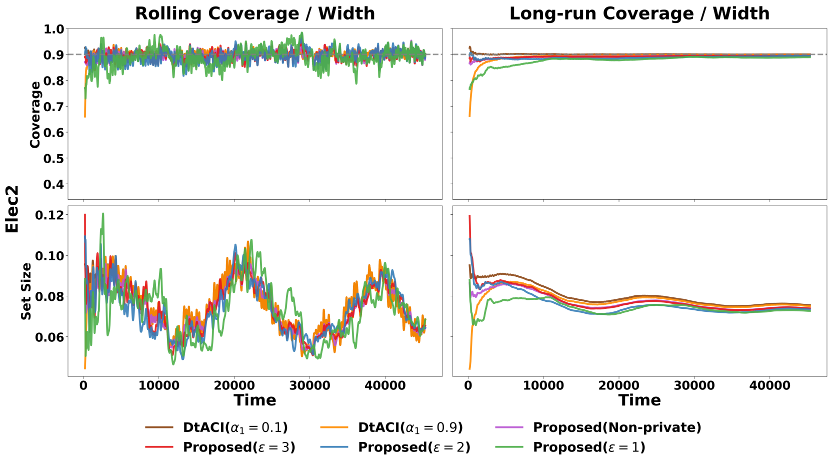


Instant Coverage / Width
Long-run Coverage / Width
Case C
Coverage
Interval Length
1.0
0.9
0.8
0.7
0.6
0.5
0.4
5.0
4.0
3.0
2.0
0
2000
4000
6000
8000
10000
Time
DtACI($\alpha_1 = 0.1$)
DtACI($\alpha_1 = 0.9$)
Proposed(Non-private)
Proposed($\varepsilon = 3$)
Proposed($\varepsilon = 1$)
Proposed($\varepsilon = 0.5$)

ELEC2 Dataset

- Electricity market dataset (ELEC2) (Harries et al., 1999):
  - Collected from the New South Wales Electricity Market 1996-1998.
  - 45,312 instances, each a 30-min interval (48 per day).
  - Potential factors: day of week, time, NSW demand, Victoria demand, scheduled transfer, etc.
- Goal: Forecasting electricity demand in New South Wales.
- Base predictor: AR(3) (Angelopoulos et al., 2023).

Elec 2



Rolling coverage (window size w): $\widehat{\mathrm{Cov}}^{(w)}(t) = \frac{1}{w} \sum_{i=t-w+1}^{t} \mathbb{I}\{Y_i \in \widehat{C}_i(X_i)\},\ t \geq w.$

Simulation Setup: classification problem

Generate data as follows:

$$x_t \sim \mathcal{N}(0, I_p), \quad \mathbb{P}(y_t = k \mid x_t) = \frac{\exp(\langle \beta_t^{(k)}, x_t \rangle)}{\sum_{j=0}^{K-1} \exp(\langle \beta_t^{(j)}, x_t \rangle)},$$

with linearly evolving coefficients

$$\beta_t^{(k)} = (1 - \alpha_t)\beta_{\text{start}}^{(k)} + \alpha_t \beta_{\text{end}}^{(k)}, \quad \alpha_t = \tfrac{t-1}{T-1}.$$

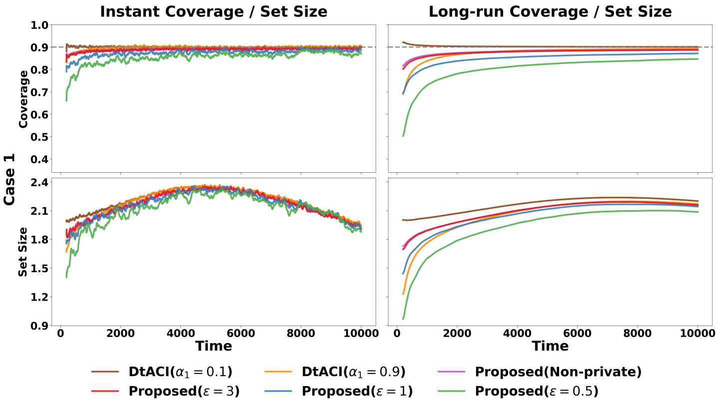
- Case 1: Smooth Drift $(K = 3, p = 3)$ Classes 0/1 drift symmetrically on feature 1; class 2 fixed. $\beta_{\text{start}}^{(0)} = [-1, 0, 0]$, $\beta_{\text{end}}^{(0)} = [1, 0, 0]$ $\beta_{\text{start}}^{(1)} = [1, 0, 0]$, $\beta_{\text{end}}^{(1)} = [-1, 0, 0]$ $\beta^{(2)} = [0, 0, 1]$
- Case 2: Amplified Drift $(K = 3, p = 3)$ Same as Case 1, but drift magnitude doubled. $\beta_{\text{start}}^{(0)} = [-2, 0, 0]$, $\beta_{\text{end}}^{(0)} = [2, 0, 0]$ $\beta_{\text{start}}^{(1)} = [2, 0, 0]$, $\beta_{\text{end}}^{(1)} = [-2, 0, 0]$ $\beta^{(2)} = [0, 0, 2]$
- Case 3: Class Emergence $(K = 4, p = 5)$ New class 3 gradually emerges via feature 5. $\beta^{(0)} = [2, 0, 0, 0, 0]$, $\beta^{(1)} = [-2, 0, 0, 0, 0]$ $\beta^{(2)} = [0, 0, 2, 0, 0]$ $\beta_{\text{start}}^{(3)} = [0, 0, 0, 0, 0]$, $\beta_{\text{end}}^{(3)} = [0, 0, 0, 0, 4]$

Simulation Results

Table: Coverages and widths for the proposed method, DPCP (Angelopoulos et al., 2022) and DtACI (Gibbs and Candès, 2024).

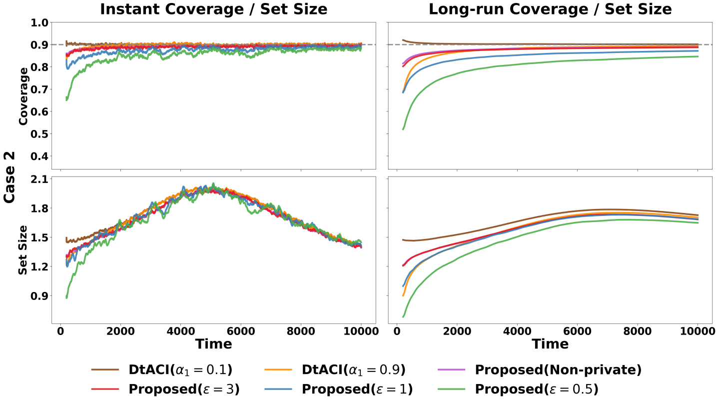
Method	no-privacy		$\epsilon = 3$		$\epsilon = 1$		$\epsilon = 0.5$	
	Coverage	size	Coverage	size	Coverage	size	Coverage	size
Case 1								
Proposed	0.890 (0.02)	2.17 (0.16)	0.889 (0.24)	2.17 (0.20)	0.875 (1.05)	2.16 (0.55)	0.854 (2.09)	2.11 (1.02)
DPCP	0.900 (0.42)	2.19 (0.14)	0.902 (0.41)	2.20 (0.15)	0.905 (0.44)	2.22 (0.17)	0.911 (0.53)	2.25 (0.24)
DtACI	0.901 (0.08)	2.20 (0.15)	*	*	*	*	*	*
Case 2								
Proposed	0.890 (0.03)	1.69 (0.16)	0.889 (0.25)	1.69 (0.22)	0.875 (1.04)	1.70 (0.62)	0.853 (2.03)	1.67 (1.16)
DPCP	0.901 (0.42)	1.72 (0.16)	0.902 (0.42)	1.73 (0.16)	0.906 (0.44)	1.74 (0.18)	0.911 (0.52)	1.77 (0.23)
DtACI	0.901 (0.08)	1.73 (0.16)	*	*	*	*	*	*
Case 3								
Proposed	0.890 (0.02)	1.71 (0.19)	0.889 (0.23)	1.72 (0.25)	0.875 (1.04)	1.77 (0.94)	0.852 (1.86)	1.75 (1.62)
DPCP	0.900 (0.42)	1.74 (0.19)	0.901 (0.42)	1.75 (0.19)	0.905 (0.45)	1.78 (0.22)	0.910 (0.54)	1.81 (0.32)
DtACI	0.900 (0.08)	1.75 (0.19)	*	*	*	*	*	*

Case 1



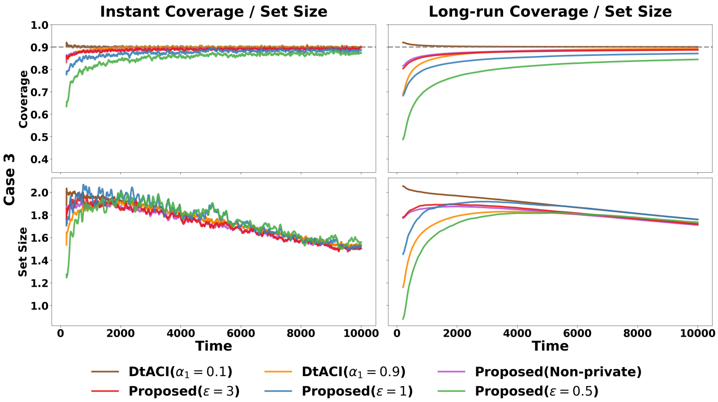
Case 1: comparison of different methods in terms of coverage, set size, long-run coverage, and long-run set size (averaged over 200 runs; first 200 points excluded).

Case 2



Case 2: comparison of different methods in terms of coverage, set size, long-run coverage, and long-run set size (averaged over 200 runs; first 200 points excluded).

Case 3



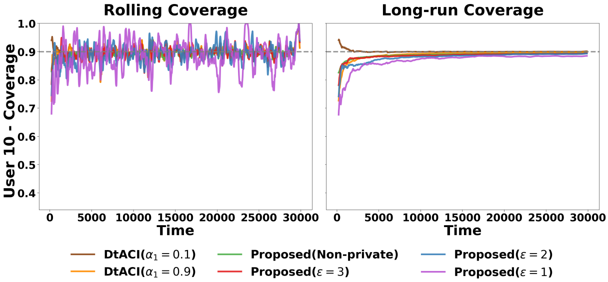
Case 3: comparison of different methods in terms of coverage, set size, long-run coverage, and long-run set size (averaged over 200 runs; first 200 points excluded).

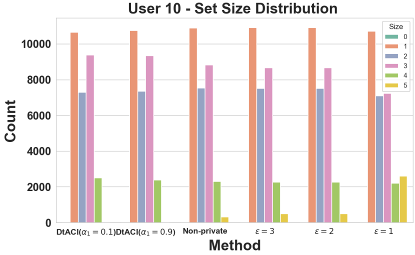
WISDM Dataset

- Accelerometer and gyroscope time-series dataset (WISDM):
  - Collected from smartphone and smartwatch sensors.
  - 51 subjects, each performing 18 activities (3 minutes each) (Kwapisz et al., 2011).
- Focus on two representative users:
  - User 10: 5 activity types
  - User 14: 4 activity types

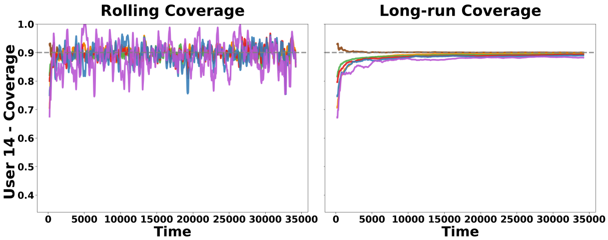
  $\Rightarrow$ two independent classification tasks.
- Goal: smartphone-based activity recognition from streaming sensor data.
- Base predictor: XGBoost classifier $\widehat{f}_t$.

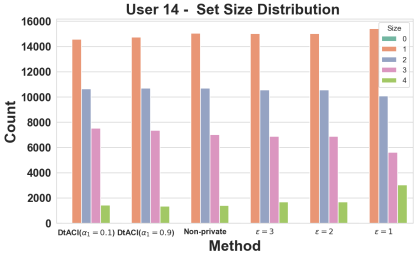
User 10: Coverage and Set Size





User 14: Coverage and Set Size





Closing remarks

Future work:

- Non-asymptotic analysis
- Multivariate conformal prediction
- Federated learning

References I

Angelopoulos, A., Candes, E., and Tibshirani, R. J. (2023). Conformal pid control for time series prediction. *Advances in neural information processing systems*, 36:23047–23074.

Angelopoulos, A. N., Bates, S., Zrnic, T., and Jordan, M. I. (2022). Private prediction sets. *Harvard Data Science Review*, 4(2).

Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography: Third Theory of Cryptography Conference, TCC 2006, New York, NY, USA, March 4-7, 2006. Proceedings 3*, pages 265–284. Springer.

Gibbs, I. and Candes, E. (2021). Adaptive conformal inference under distribution shift. *Advances in Neural Information Processing Systems*, 34:1660–1672.

Gibbs, I. and Candès, E. J. (2024). Conformal inference for online prediction with arbitrary distribution shifts. *Journal of Machine Learning Research*, 25(162):1–36.

Harries, M., of New South Wales. School of Computer Science, U., and Engineering (1999). *Splice-2 Comparative Evaluation: Electricity Pricing*. PANDORA electronic collection. University of New South Wales, School of Computer Science and Engineering.

References II

Krichevsky, R. and Trofimov, V. (1981). The performance of universal encoding. *IEEE Transactions on Information Theory*, 27(2):199–207.

Kwapisz, J. R., Weiss, G. M., and Moore, S. A. (2011). Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter*, 12(2):74–82.

Liu, Y., Hu, Q., Ding, L., and Kong, L. (2023). Online local differential private quantile inference via self-normalization. In *International Conference on Machine Learning*, pages 21698–21714. PMLR.

Orabona, F. and Pál, D. (2016). Coin betting and parameter-free online learning. *Advances in Neural Information Processing Systems*, 29.

Penso, C., Mahpud, B., Goldberger, J., and Sheffet, O. (2025). Privacy-preserving conformal prediction under local differential privacy. *arXiv preprint arXiv:2505.15721*.

Plassier, V., Makni, M., Rubashevskii, A., Moulines, E., and Panov, M. (2023). Conformal prediction for federated uncertainty quantification under label shift. In *International Conference on Machine Learning*, pages 27907–27947. PMLR.

Podkopaev, A., Xu, D., and Lee, K.-c. (2024). Adaptive conformal inference by betting. *arXiv preprint arXiv:2412.19318*.

References III

Vovk, V., Gammerman, A., and Shafer, G. (2005). *Algorithmic Learning in a Random World*. Springer US.

Xiong, X., Liu, S., Li, D., Cai, Z., and Niu, X. (2020). A comprehensive survey on local differential privacy. *Security and Communication Networks*, 2020(1):8829523.

Zaffran, M., Féron, O., Goude, Y., Josse, J., and Dieuleveut, A. (2022). Adaptive conformal predictions for time series. In *International Conference on Machine Learning*, pages 25834–25866. PMLR.

Thanks for your attention!