GSP workshop 2026, Madrid, Spain

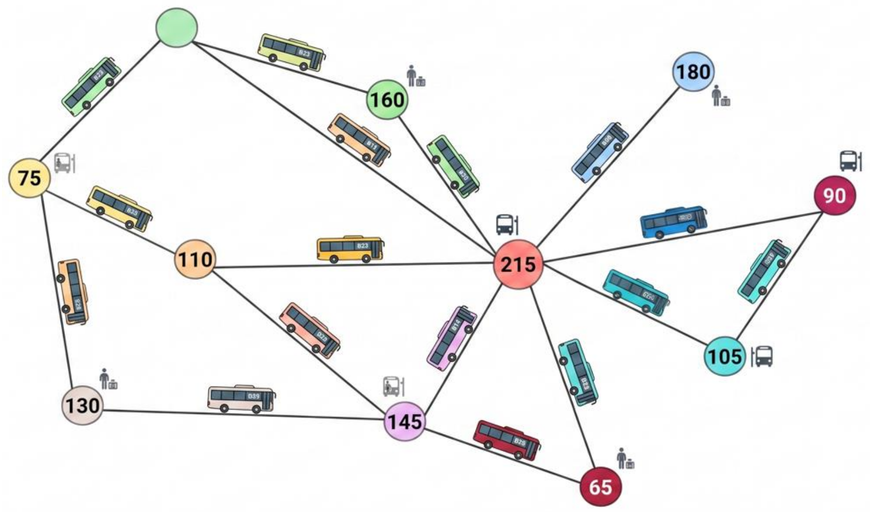
CONFORMAL INFERENCE FOR TIME SERIES OVER GRAPHS

Sonakshi Dua, Gonzalo Mateos, and **Sundeep Prabhakar Chepuri**
Indian Institute of Science, Bengaluru, India
University of Rochester, USA

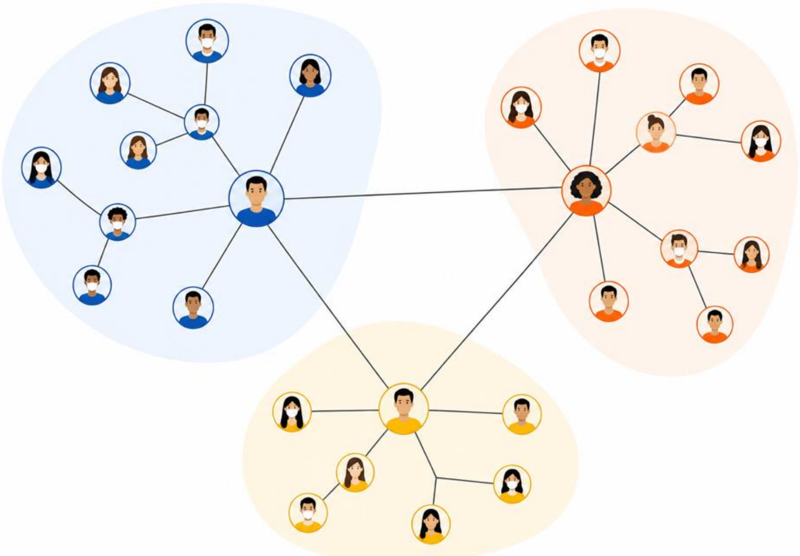
Presented at ICASSP, May 2026, Barcelona, Spain

INDIAN INSTITUTE OF SCIENCE
भारतीय विज्ञान संस्थान

UNIVERSITY of ROCHESTER

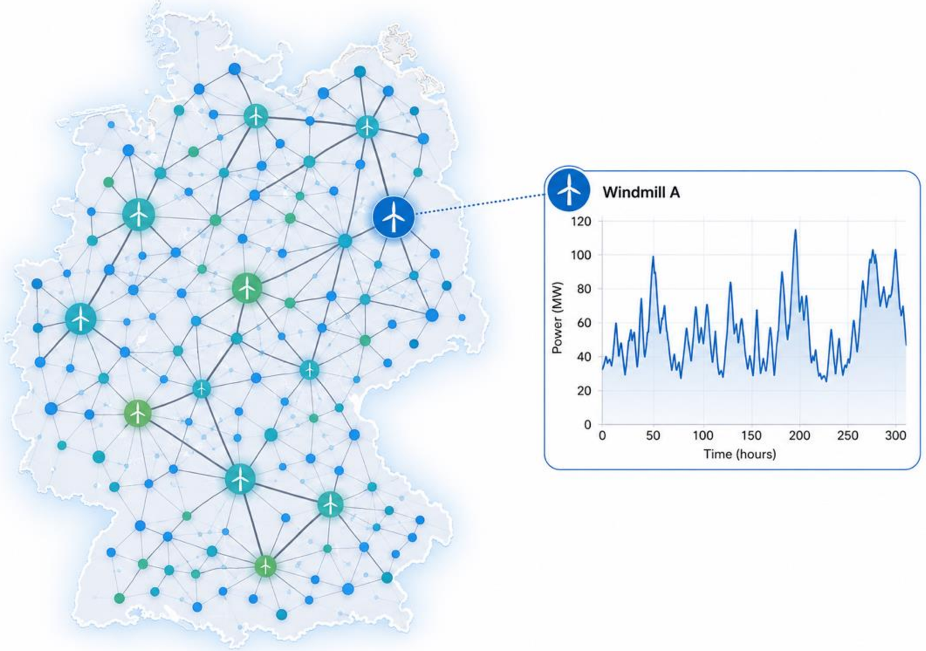


Passenger footfall at different bus stations



Evolving epidemic network

Temperature over different cities

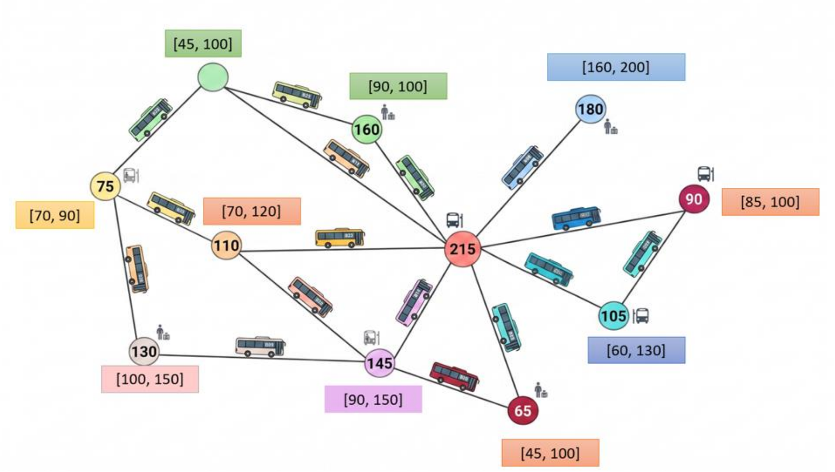


Hourly energy output of windmills

Going beyond point predictions

Consider a supervised learning problem with features $\mathbf{x}_i$ and labels $y_i \in \mathcal{Y}$

Prediction set examples on *Imagenet* *[1]*



Prediction intervals for regression

Distribution-free wrapper around **pre-trained neural model** with **finite-sample coverage guarantee**

$$\mathbb{P}(y \in \mathcal{C}(\mathbf{x})) \geq 1 - \alpha$$

with $\mathcal{C}(\mathbf{x})$ **as small as possible (efficiency).** Assume **exchangeability** between calibration and test data.

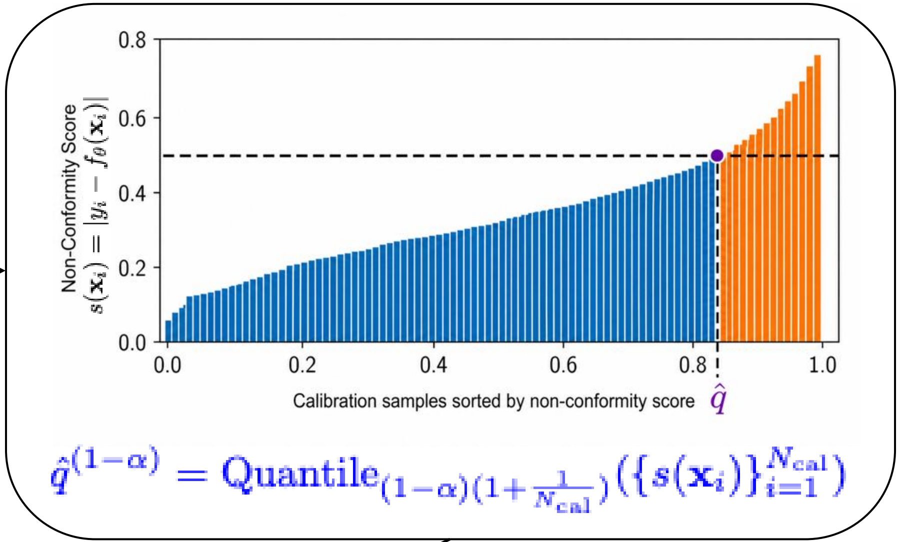
[1] Angelopoulos, A. N. & Bates, S. (2023). Conformal prediction: A gentle introduction. Foundations and Trends in Machine Learning, 16(4), 494–591.

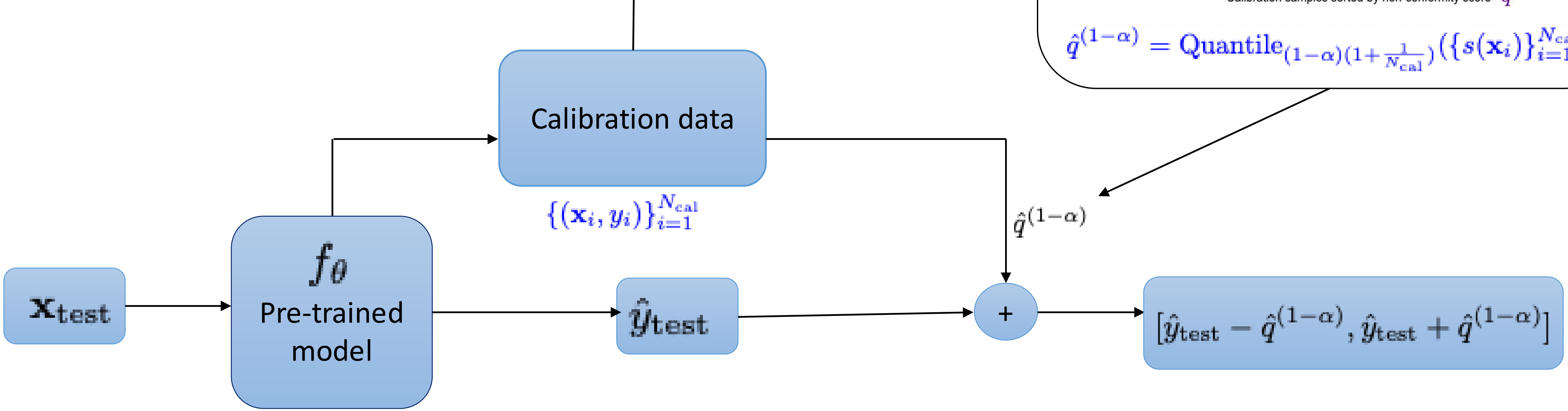
Conformal Prediction

Exchangeability : Random variables $W_1, \dots, W_n$ for $n \geq 1$ are said to be exchangeable if

$$(W_1, \dots, W_n) \stackrel{d}{=} (W_{\pi(1)}, \dots, W_{\pi(n)}),$$

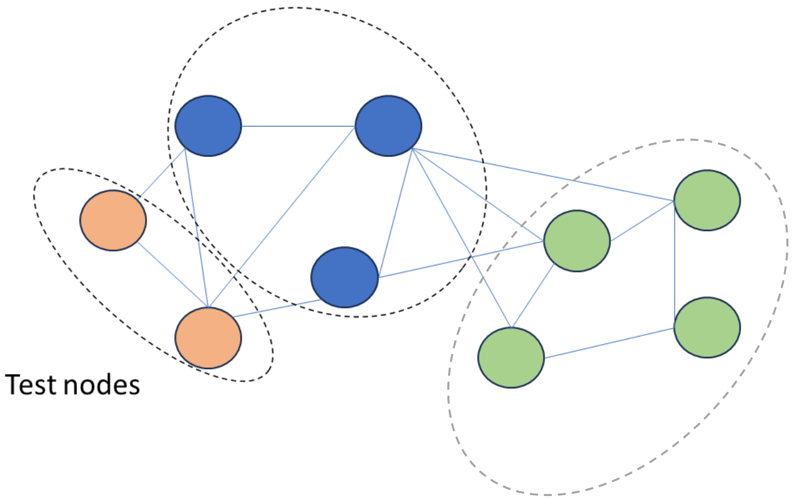
for any permutation π.



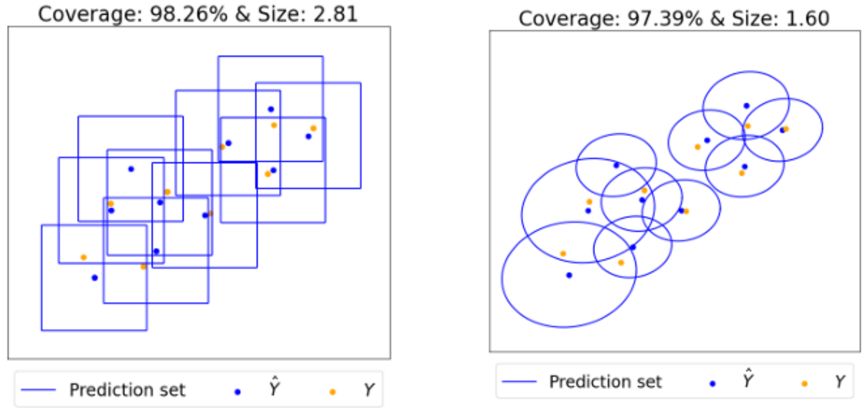


State-of-the-art CP methods on graphs and multivariate time-series

- CP over static graphs : DAPS *[2]*, CF-GNN *[3]*



- Multidimensional time series: Multidim SPCI *[4]*



Coordinate wise set

Ellipsoidal set *[4]*

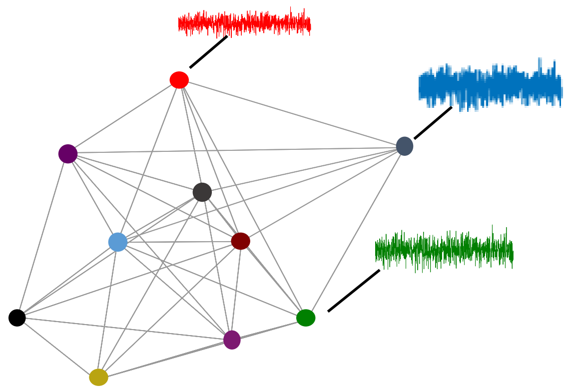
[2] H. Zargarbashi, S., Antonelli, S. & Bojchevski, A.. (2023). Conformal Prediction Sets for Graph Neural Networks. Proceedings of the 40th International Conference on Machine Learning.

[3] Huang, K., Jin, Y., Candes, E., & Leskovec, J. (2023). Uncertainty quantification over graph with conformalized graph neural networks. *Advances in Neural Information Processing Systems*, *36*, 26699-26721.

[4] Xu, C., Jiang, H., & Xie, Y. (2024). Conformal prediction for multi-dimensional time series by ellipsoidal sets. *arXiv preprint arXiv:2403.03850*.

Problem statement

- Consider a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with N nodes



- Graph data $\mathbf{x}_t \in \mathbb{R}^N$ or $\mathcal{G}_t = (\mathbf{x}_t, \mathcal{G})$ and target $\mathbf{y}_t \in \mathbb{R}^N$
- Assume access to a pre-trained neural model $f_\theta(\cdot)$

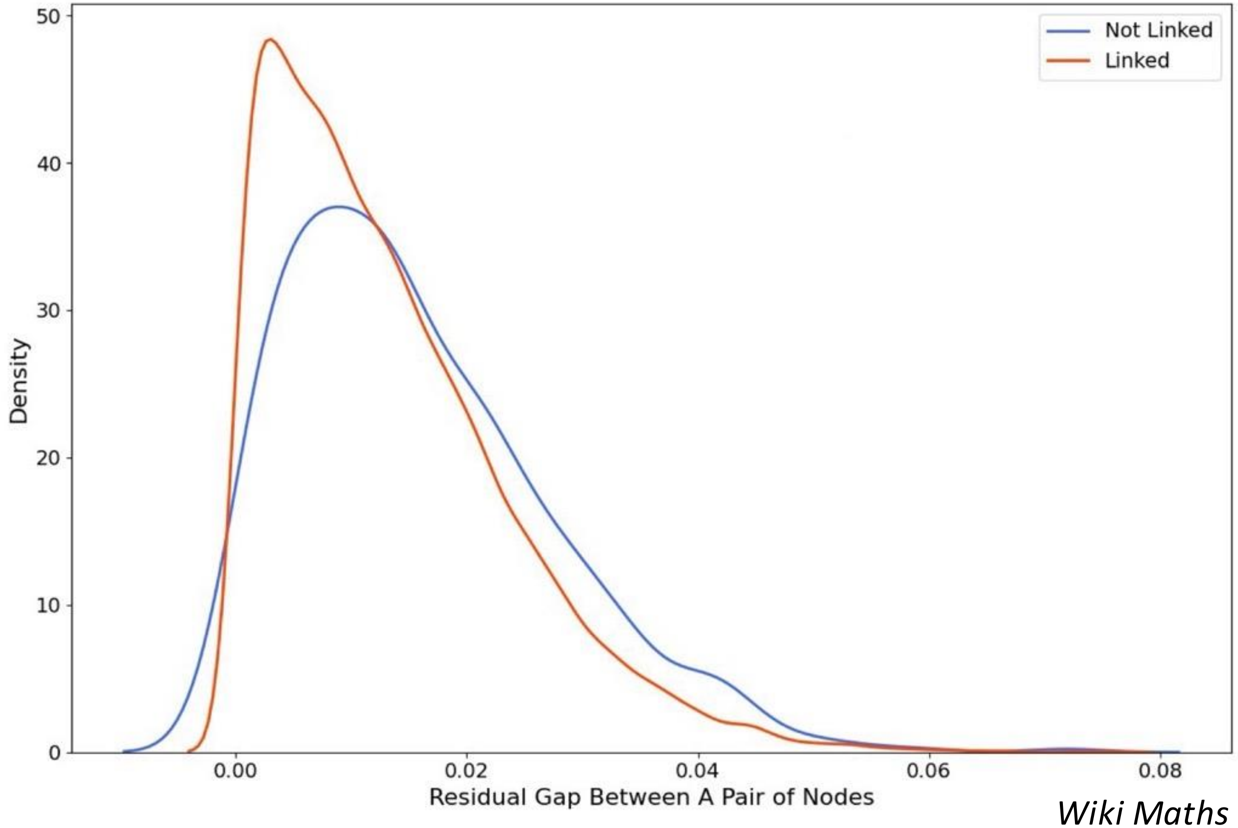
$$\hat{\mathbf{y}}_t = \hat{\mathbf{x}}_{t+1} = f_\theta(\mathcal{G}_t)$$

Estimate an efficient prediction set $\mathcal{C}_t(\mathcal{G}_t)$ such that

$$\mathbb{P}(\mathbf{y}_t \in \mathcal{C}_t(\mathcal{G}_t)) \geq 1 - \alpha \quad \forall t$$

Step 1: non-conformity scores

- Compute residuals as non-conformity scores: $\varepsilon_t = |\mathbf{y}_t - f_\theta(\mathcal{G}_t)| \in \mathbb{R}^N$
  (Residuals are viewed as graph signals)



Wiki Maths dataset

Residuals are smooth over the graph!

Step 2: Graph aware non-conformity scores

- Compute graph-aware non-conformity scores using a *graph filter*

$$\mathbf{e_t} = \mathbf{H}\boldsymbol{\varepsilon}_\mathbf{t}$$
$$= [(1-\tau)\mathbf{I} + \tau(\mathbf{D}^{-1}\mathbf{A}]$$

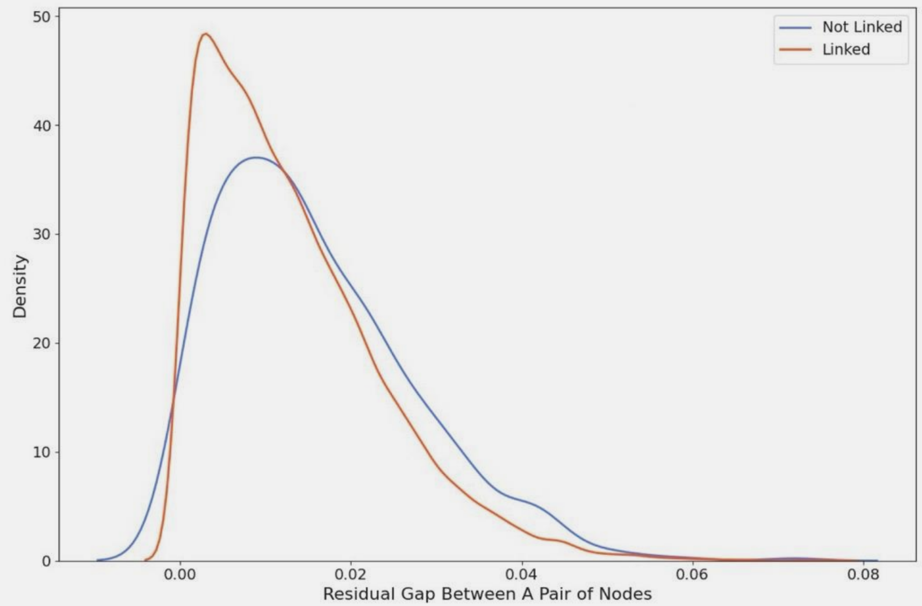
hyperparameter

- Convert to a scalar non-conformity score via the *squared Mahalanobis distance*

$$s_t = (\mathbf{e}_t - \bar{e})^\top \boldsymbol{\Sigma}_t^{-1}(\mathbf{e}_t - \bar{e})$$

$$\boldsymbol{\Sigma}_t = \mathbb{E}[(\mathbf{e}_t - \bar{e})(\mathbf{e}_t - \bar{e})^\top]$$

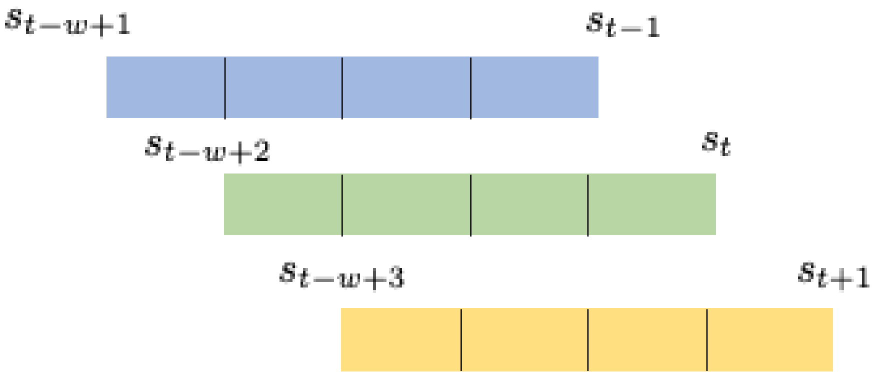
In practice, sample mean and covariance are used



Since the residuals were already smooth, graph filtering ***denoises*** those scores that were leading to larger uncertainty.

Step 3: ellipsoidal prediction sets

- Pick a window size w



Quantile regression $\longrightarrow$ $\hat{q}_t^{(1-\alpha)} = \text{Quantile}_{(1-\alpha)(1+\frac{1}{w})}(\{s_i\}_{i=t-w+1}^{t-1})$

Quantile regression $\longrightarrow$ $\hat{q}_{t+2}^{(1-\alpha)} = \text{Quantile}_{(1-\alpha)(1+\frac{1}{w})}(\{s_i\}_{i=t-w+2}^{t})$

Quantile regression $\longrightarrow$ $\hat{q}_{t+2}^{(1-\alpha)} = \text{Quantile}_{(1-\alpha)(1+\frac{1}{w})}(\{s_i\}_{i=t-w+3}^{t+1})$

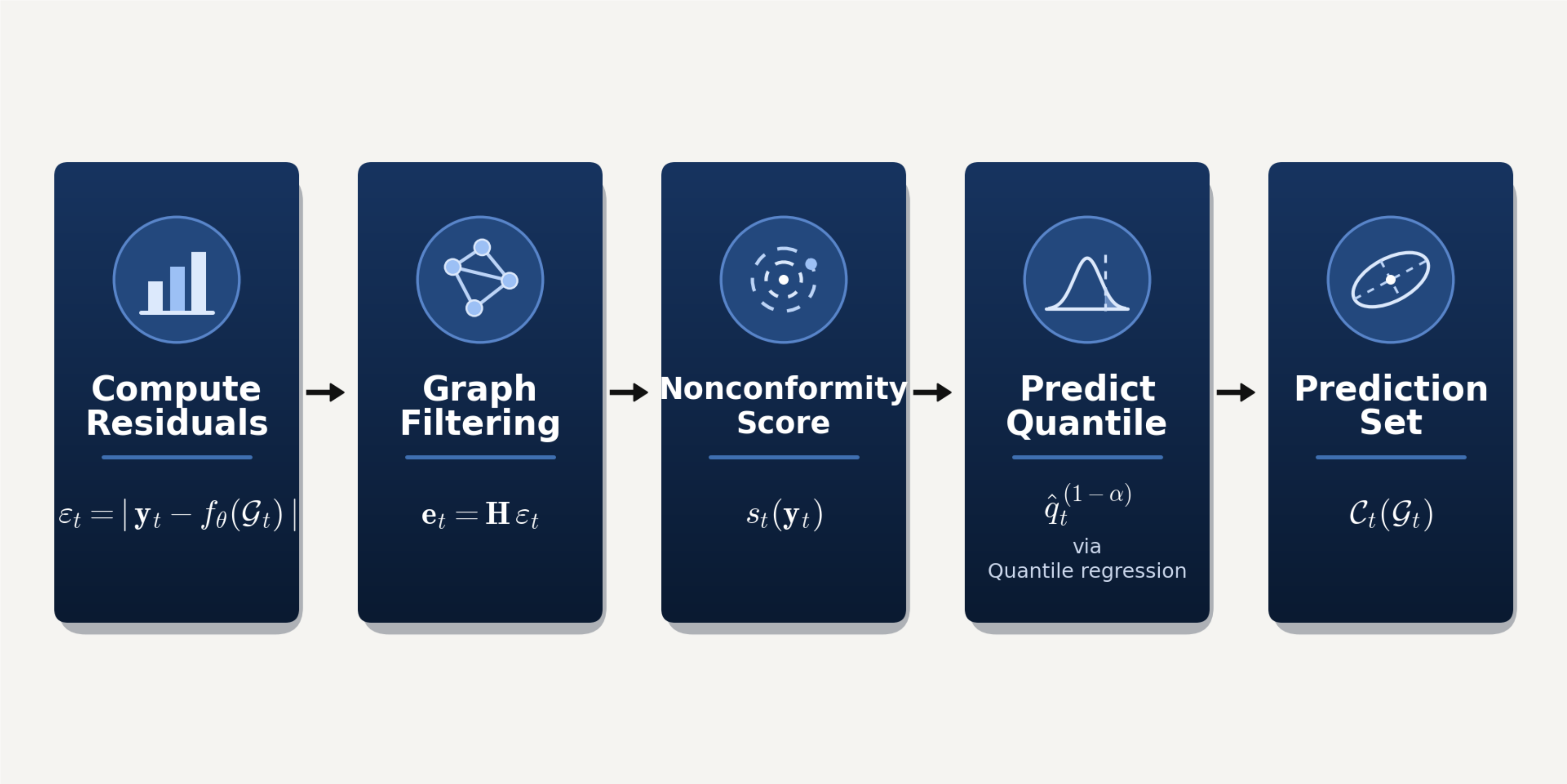
- Construct an ellipsoid set interval around the point prediction from the base neural model

$$\begin{aligned}\mathcal{C}_t(\mathcal{G}_t) &= \{\mathbf{y}_t | s_t \le \hat{q}_t^{(1-\alpha)}\} \\ &= f_\theta(\mathcal{G}_t) + \mathcal{B}\left(\sqrt{\hat{q}_t^{(1-\alpha)}}, \bar{\mathbf{e}}_t, \mathbf{\Sigma}_t\right)\end{aligned}$$

prediction set keeps every point whose squared distance is $\le \hat{q}_t^{(1-\alpha)}$

$\mathcal{B}(r, \mathbf{c}, \mathbf{\Sigma}) = \{\mathbf{x} | (\mathbf{x} - \mathbf{c}^\top \mathbf{\Sigma}^{-1}(\mathbf{x} - \mathbf{c}\}$

The pipeline



Theoretical results

Ellipsoidal Volume Shrinkage

If $\frac{\hat{q}_t^{(1-\alpha)}}{\hat{q}_{\mathcal{G},t}^{(1-\alpha)}} \approx 1$ then $\mathrm{Vol}(\mathcal{B}_\mathcal{G}) \leq \mathrm{e}^{-\eta\tau}\mathrm{Vol}(\mathcal{B})$, for some $\eta > 0$, where τ is the graph filter coefficient

Proof sketch: $$\log\left[\frac{\mathrm{Vol}(\mathcal{B}_G)}{\mathrm{Vol}(\mathcal{B})}\right] = \underbrace{\log\det(\mathbf{H})}_{\leq -\tau\eta} + \frac{N}{2}\left(\log\left(\frac{\hat{q}_{t,\mathcal{G}}^{(1-\alpha)}}{\hat{q}_t^{(1-\alpha)}}\right)\right)$$

Coverage Guarantees

Assume that the true covariance matrix $\boldsymbol{\Sigma}_t$ is known and positive definite with minimum eigenvalue at least $\lambda > 0$, the filtered residuals are i.i.d. over time, and the CDF of the true non-conformity score is Lipschitz. Then we have

$$\left|\mathbb{P}\left(\mathbf{y}_{t+1} \in \mathcal{C}_{\mathcal{G}_t}\right) - (1-\alpha)\right| \leq 12\sqrt{\frac{\log(16T)}{T}} + L\left(\frac{\delta_T}{\sqrt{\lambda}} + \delta_T\right).$$

where T is the training data size, δ_T is a bound on the residual error, and L depends on the Lipschitz constant.

Experimental results

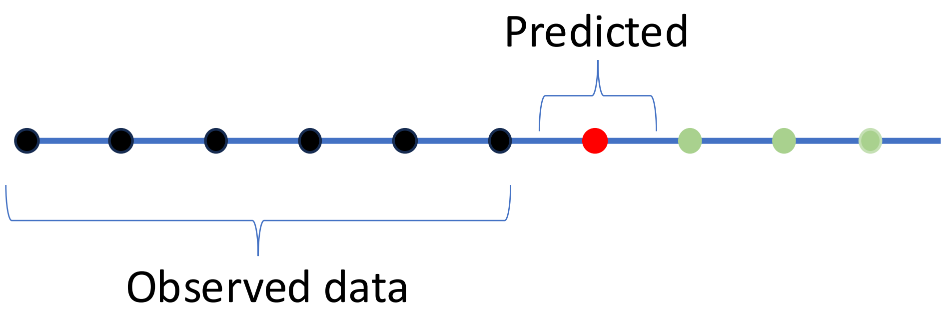
$\alpha = 0.1$

Method	Metric	Wiki Maths	MontevideoBus	Chickenpox Hungary
Graph-agnostic	Coverage	0.903 ± 0.006	0.91 ± 0.0122	0.89 ± 0.0167
	Volume	$5.19 \times 10^3 \pm 2010.298$	$3.09 \times 10^3 \pm 247.325$	$2.74 \times 10^2 \pm 14.842$
Graph-aware (proposed)	Coverage	0.897 ± 0.010	0.912 ± 0.008	0.89 ± 0.018
	Volume	$\mathbf{1.46 \times 10^3 \pm 135.769}$	$\mathbf{1.56 \times 10^3 \pm 880.327}$	$\mathbf{1.25 \times 10^2 \pm 70.851}$

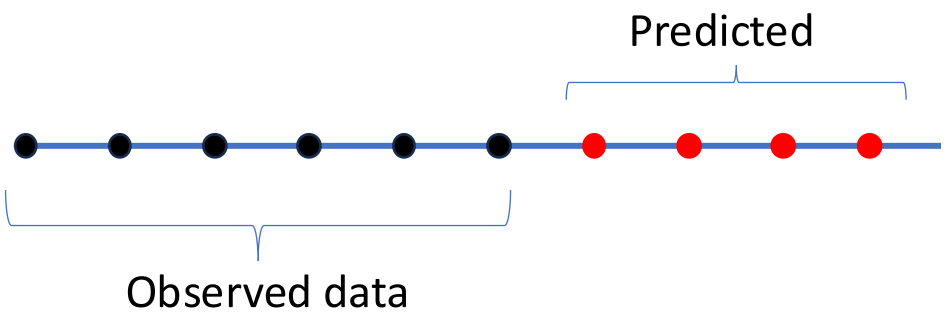
$\alpha = 0.05$

Method	Metric	Wiki Maths	MontevideoBus	Chickenpox Hungary
Graph-agnostic	Coverage	0.954 ± 0.005	0.948 ± 0.004	0.916 ± 0.011
	Volume	$8.51 \times 10^3 \pm 1458.303$	$1.406 \times 10^4 \pm 205.985$	$1.6 \times 10^2 \pm 14.153$
Graph-aware (proposed)	Coverage	0.952 ± 0.004	0.952 ± 0.008	0.924 ± 0.005
	Volume	$\mathbf{2.04 \times 10^3 \pm 238.579}$	$\mathbf{2.7 \times 10^3 \pm 112.472}$	$\mathbf{1.29 \times 10^2 \pm 20.144}$

Experimental results



One-step prediction (r = 1)

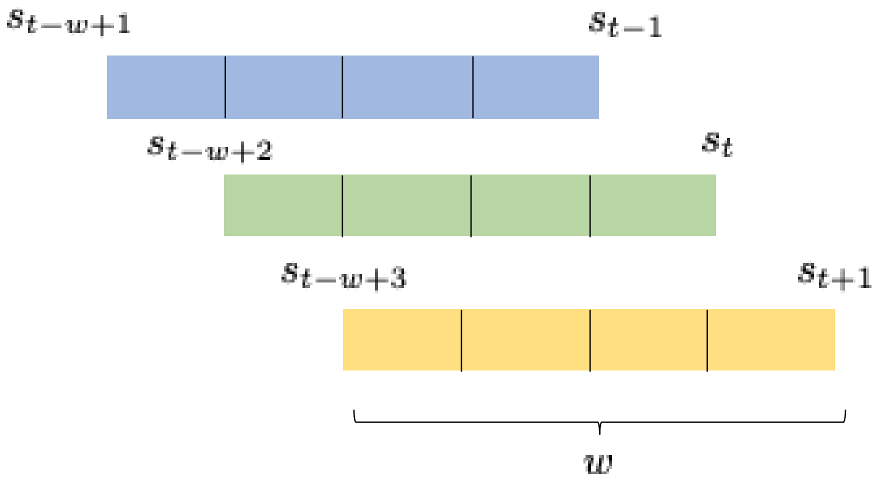


Multi-step prediction (r = 4)

Wiki Maths Dataset	$r = 1$		$r = 5$		$r = 10$	
Method	Coverage	Volume	Coverage	Volume	Coverage	Volume
Graph-agnostic	0.903 ± 0.006	$5.19 \times 10^3 \pm 2010.29$	0.889 ± 0.012	$1.20 \times 10^4 \pm 754.91$	0.875 ± 0.008	$9.40 \times 10^3 \pm 5218.74$
Graph-aware (proposed)	0.897 ± 0.010	$\mathbf{1.46 \times 10^3 \pm 135.77}$	0.885 ± 0.0129	$\mathbf{2.40 \times 10^3 \pm 694.09}$	0.867 ± 0.002	$\mathbf{3.50 \times 10^3 \pm 595.01}$

Results over one-step prediction and multi-step prediction (r being the number of steps)

Experimental results



Wiki Maths Dataset	$w = 10$		$w = 50$		$w = 100$	
Method	Coverage	Volume	Coverage	Volume	Coverage	Volume
Graph-agnostic	$0.903 \pm .006$	$5.19 \times 10^3 \pm 2010.298$	0.892 ± 0.005	$4.80 \times 10^3 \pm 1135.54$	0.885 ± 0.014	$2.85 \times 10^3 \pm 1260.33$
Graph-aware (proposed)	0.897 ± 0.01	$\mathbf{1.46 \times 10^3 \pm 135.769}$	0.896 ± 0.009	$\mathbf{1.27 \times 10^3 \pm 182.67}$	0.886 ± 0.002	$\mathbf{1.22 \times 10^3 \pm 345.89}$

Results over different window lengths to predict the quantiles

Conclusions

- Conformal prediction for graph-structured data
  - ✓ Refine nonconformity scores using a graph filter
- Derived theoretical guarantees on volume shrinkage
  - ✓ Leads to smaller prediction sets while maintaining coverage

Thank You!

https://ece.iisc.ac.in/~spchepuri/