Valid Conformal Prediction for Dynamic GNNs

Ed Davis[1] Ian Gallagher[2] Daniel John Lawson[1] Patrick Rubin-Delanchy[3]

[1]University of Bristol, U.K. [2]The University of Melbourne, Australia, [3]School of Mathematics and Maxwell Institute for Mathematical Sciences, University of Edinburgh, U.K.

{edward.davis, dan.lawson}@bristol.ac.uk, ian.gallagher@unimelb.edu.au, prd@ed.ac.uk

Link to full paper

Summary of Contributions

This work presents a novel method for uncertainty quantification in dynamic graph neural networks (GNNs), contributing to the broader field of trustworthy AI.

In this work, we propose to use **unfolded GNNs** (UGNNs) in discrete-time dynamic network settings. Compared to the standard “block“ GNN approach, we find two main advantages of using UGNN.

1. **UGNN has higher predictive performance.**
   - In transductive scenarios, gains are minor (e.g. **92% vs 86%**). In semi-inductive scenarios, gains are major (e.g. **92% vs 12%**).
2. **Conformal prediction on UGNN is valid in more settings.**
   - We show that UGNN and existing GNNs are valid in transductive cases.
   - However, only UGNN is valid in semi-inductive settings, as long as there is no data drift.

Considered Regimes

We consider the problem of using a GNN to classify nodes within a discrete-time dynamic network. Below, we illustrate the *transductive* and *semi-inductive* regimes, which define how each node can be assigned to training, validation, calibration or testing masks. Training and validation nodes are used for GNN training and best model selection, calibration nodes for split conformal prediction, and testing nodes for performance metrics. Note that we use the term *semi*-inductive as the GNN will still see all the future graphs in the series, however, it will not be able to see the label of future nodes.

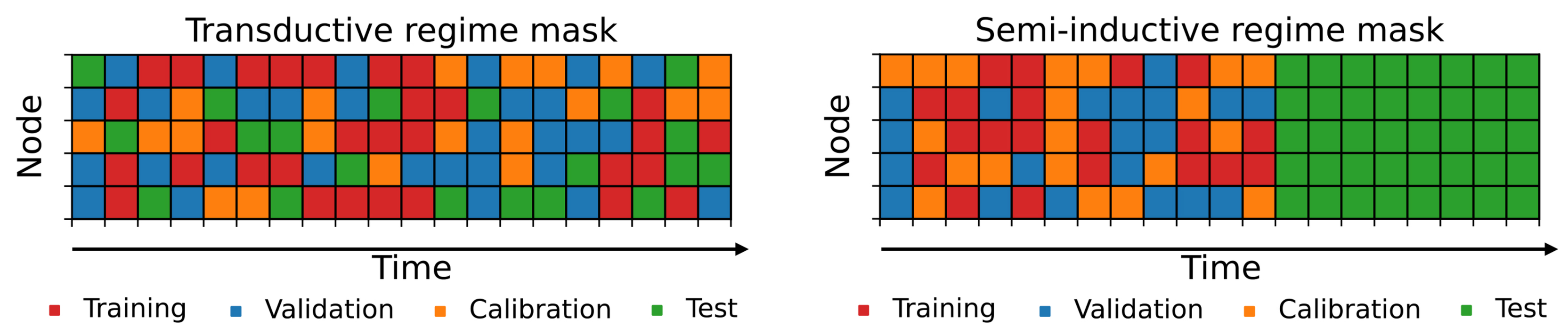


Fig. 1: An illustration of how nodes are allocated to masks under two regimes. In the transductive regime, all time points are exchangeable in terms of training/validation/calibration/testing split. In the semi-inductive regime, a future period is reserved for testing.

Unfolded GNN

We propose to use a *dilated unfolding* [1] to structure the input of multiple networks, as opposed to the established *block diagonal* approach [2]. A UGNN is simply a GNN that is given a set of graphs in a dilated unfolding structure.

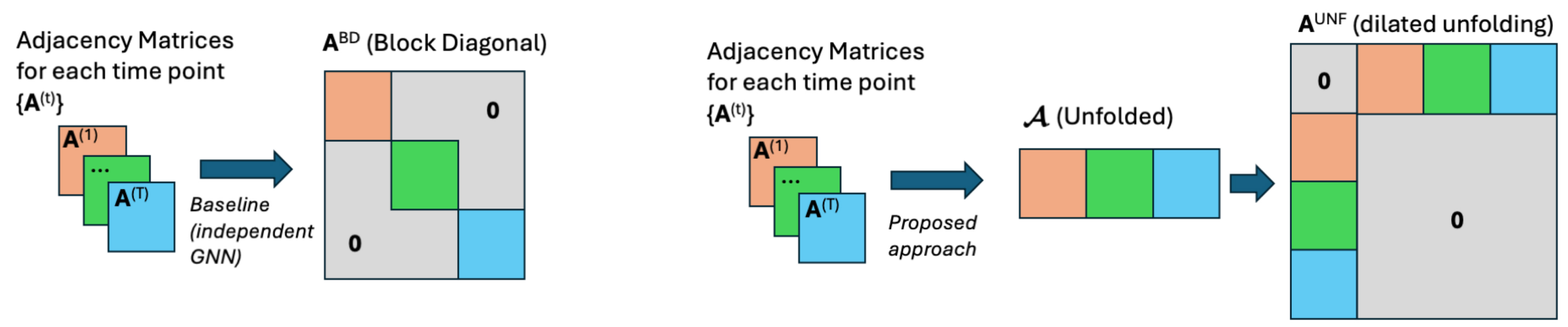


Fig. 2: The baseline (current practice) approach treats adjacency matrices as independent and can be viewed as padding a ‘block-diagonal’ matrix with zeroes. Unfolding instead column concatenates which links nodes to themselves over time. Dilation results in a square symmetric matrix.

We show that UGNN embeddings possess desirable exchangeability properties, which prevent the de-alignment between subsequent embedding time points seen in conventional GNNs. This is beneficial for two reasons.

1. The alignment of embeddings over time improves predictive performance, particularly when predicting into the future (semi-inductive regime).
2. The *exchangeability* of embeddings over time allows for the valid application of conformal prediction in the semi-inductive regime.

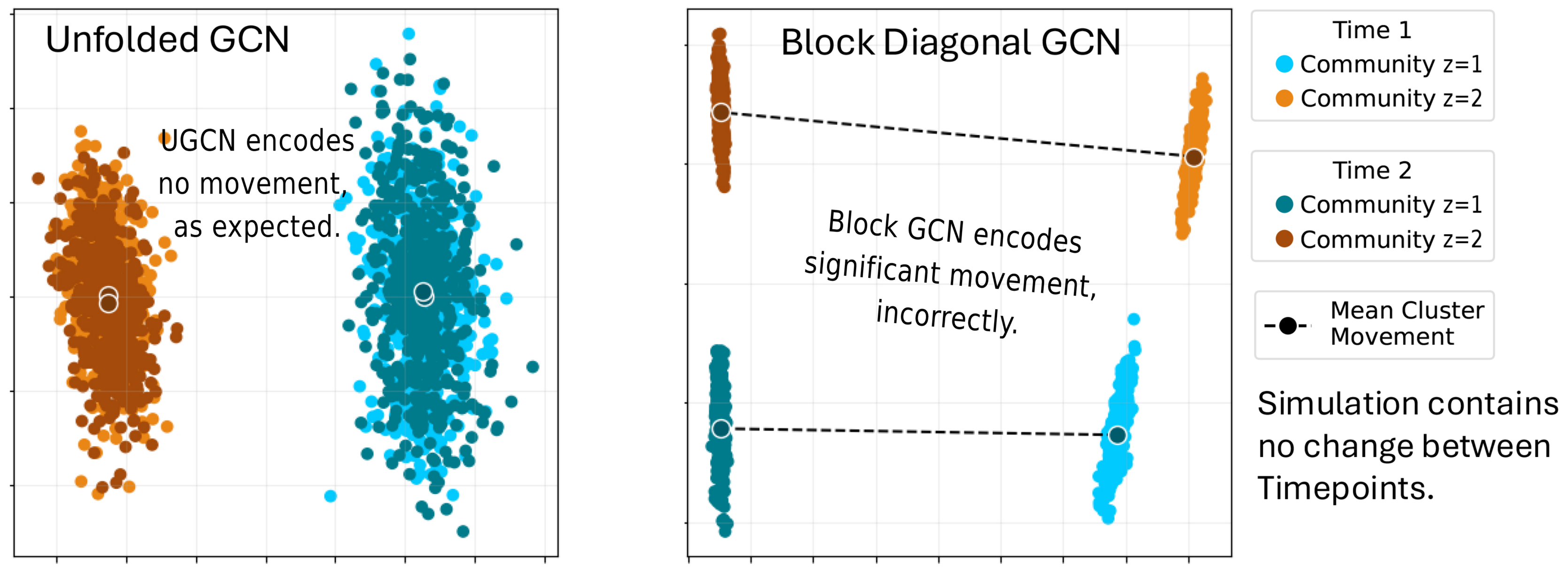
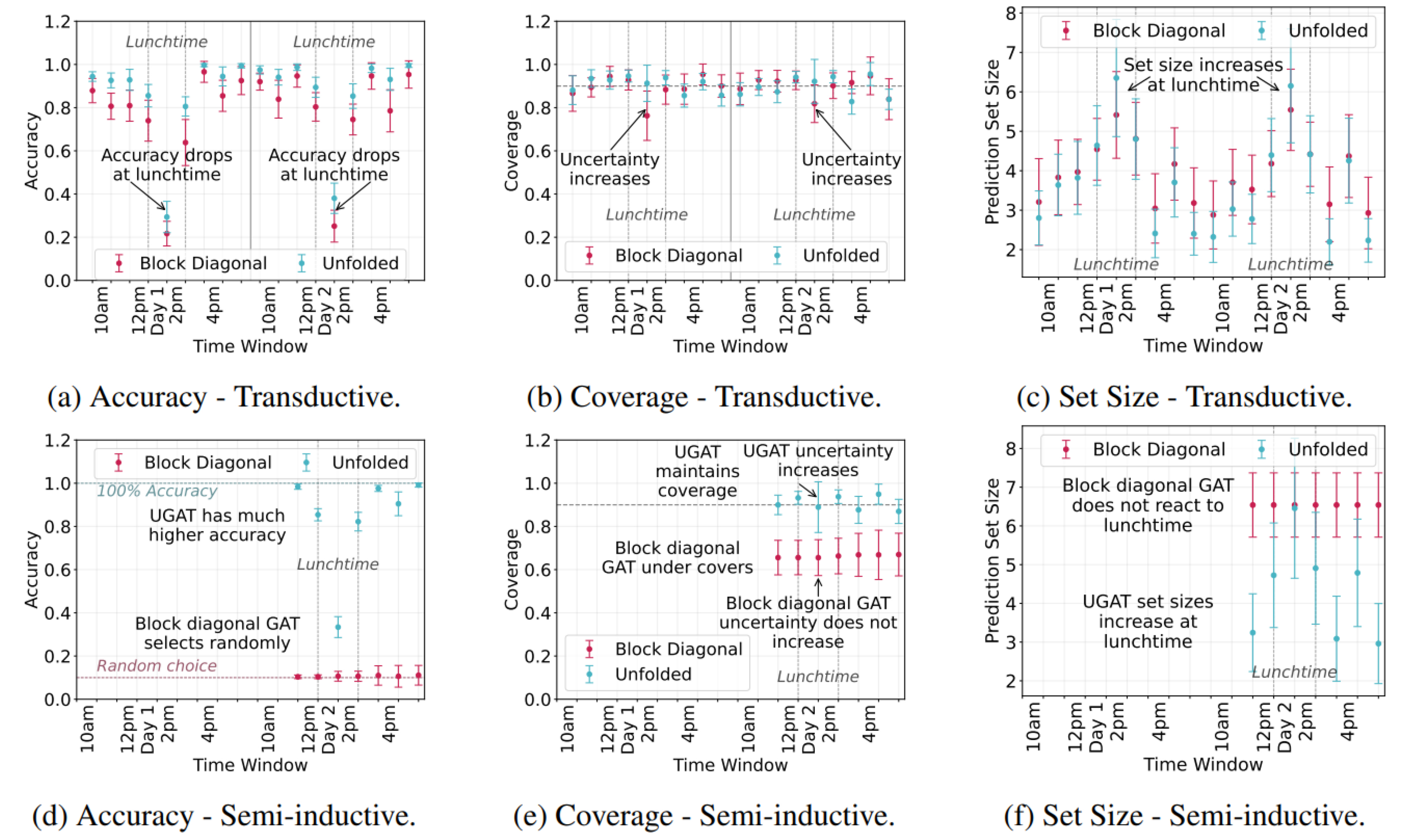


Fig. 3: GCN embeddings of an i.i.d. stochastic block model. We plot the output layer (pre-softmax) of each GCN, trained using transductive masks, and reduce to 2D using PCA. Block diagonal GCN appears to encode a significant change over time despite there being none. The embedding from UGCN is exchangeable over time, as should be expected.



(a) Accuracy - Transductive. (b) Coverage - Transductive. (c) Set Size - Transductive.

(d) Accuracy - Semi-inductive. (e) Coverage - Semi-inductive. (f) Set Size - Semi-inductive.

Fig. 4: Performance metrics for each time window of the school dataset for unfolded GAT and block diagonal GAT. The prediction task gets more difficult at lunchtime, as shown by the drop in accuracy of both methods in the transductive case. UGAT has marginally better performance in the transductive case and significantly better performance in the semi-inductive case. Prediction set sizes increase at lunchtime, with only UGAT set sizes reacting in the semi-inductive case. Both methods maintain target coverage in the transductive case, with uncertainty increasing at the more difficult lunchtime window. UGAT also maintains target coverage in the semi-inductive case, while block GAT under-covers.

Real Data Example: School Interaction Network

To demonstrate practical power, we compare the performance of UGNN vs the standard approach on a real dataset. This dataset is a dynamic interaction network between people in a primary school over two days. The task is to predict the class of each person, something that is easy during class hours but gets more difficult at lunchtime as students are outside for breaktime.
Both models were valid in the transductive regime, however only UGNN was valid in the semi-inductive regime.

Methods	Accuracy		Coverage		Avg. Set Size	
	Trans.	Semi-ind.	Trans.	Semi-ind.	Trans.	Semi-ind.
Block GCN	0.856 ± 0.011	0.116 ± 0.011	$\mathbf{0.901 \pm 0.012}$	0.812 ± 0.033	4.542 ± 0.167	8.079 ± 0.188
UGCN	$\mathbf{0.924 \pm 0.009}$	$\mathbf{0.915 \pm 0.013}$	$\mathbf{0.901 \pm 0.012}$	$\mathbf{0.924 \pm 0.013}$	$\mathbf{2.763 \pm 0.311}$	$\mathbf{3.037 \pm 0.251}$
Block GAT	0.807 ± 0.016	0.107 ± 0.024	$\mathbf{0.901 \pm 0.012}$	0.662 ± 0.084	3.863 ± 0.813	6.540 ± 0.830
UGAT	$\mathbf{0.896 \pm 0.016}$	$\mathbf{0.868 \pm 0.017}$	$\mathbf{0.901 \pm 0.012}$	$\mathbf{0.909 \pm 0.021}$	$\mathbf{3.552 \pm 0.756}$	$\mathbf{4.185 \pm 1.228}$

Table 1: Results on a real data example for two GNNs (GCN or GAT) under two representations (block diagonal or unfolding), in both the transductive and semi-inductive regimes. Coverage is targeted to ≥ 0.9. Bold values indicate ideal performance.

The goal of conformal prediction is to provide a notion of uncertainty to an otherwise black-box model. As a problem becomes more difficult, this should be reflected in a larger prediction set size to maintain target coverage.
In the transductive regime, the conformal set size for both models increases during the lunchtime period, which correctly reflects the increased difficulty of the problem. However, in the semi-inductive regime, only UGAT can display this behaviour, with block diagonal GAT returning a constant set size for this period.

Related Work and Contact

Other unfolded embedding methods

Personal website

Acknowledgements

Ed Davis gratefully acknowledges support by LV= Insurance (Allianz Group) and the Centre for Doctoral Training in Computational Statistics and Data Science (Compass, EPSRC Grant number EP/S023569/1). Compass is funded by United Kingdom Research and Innovation (UKRI) through the Engineering and Physical Sciences Research Council (EPSRC), https://www.ukri.org/ councils/epsrc. This work was carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol - http://www.bris.ac.uk/acrc/. PRD gratefully acknowledges support from the NeST programme grant, EPSRC grant number EP/X002195/1.

References

[1] Ed Davis et al. “A Simple and Powerful Framework for Stable Dynamic Network Embedding”. In: *arXiv preprint arXiv:2311.09251* (2023).

[2] Matthias Fey and Jan E. Lenssen. “Fast Graph Representation Learning with PyTorch Geometric”. In: *ICLR Workshop on Representation Learning on Graphs and Manifolds.* 2019.