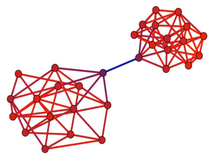
The Effectiveness of Curvature-Based Rewiring and the Role of Hyperparameters in GNNs Revisited

Floriano Tori[1] , Vincent Holst[1] & Vincent Ginis[1,2]

1. Background

- **Message passing** in Graph Neural Networks can be limited by the topology of the graph. This happens when information is lost during propagation due to being **oversquashed**.
- Discrete graph curvature, such as ***Balanced Forman curvature (BFc),*** is supposed to detect bottleneck between two nodes.



- The **oversquashing theorem** (right) says that bottlenecks are edges with very low *BFc* (close to ***minimal value - 2***).
- **Rewiring** around these edges could be beneficial for information propagation.

2. Contributions

- In our work however, we show that edges selected during the rewiring process in real-world datasets **do not satisfy *condition 2*.** This questions their identification as **bottlenecks** and the need for rewiring.
- We find that results of accuracy improvements due to rewiring can be attributed to **outliers** originating from hyperparameter sweeps, instead of consistent improvements.

3. Experiments

Dataset: Squirrel

a. Look at the edges selected during the rewiring process

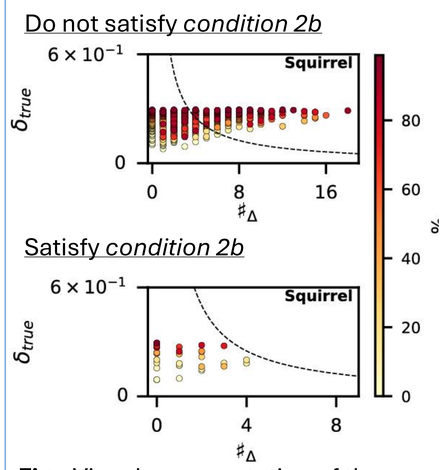


Fig.: Visual representation of the selected edges during SDRF rewiring. Colour indicates at which % of total rewiring the edge was selected.

b. Analyse distribution of accuracies obtained from hyperparameter sweeps

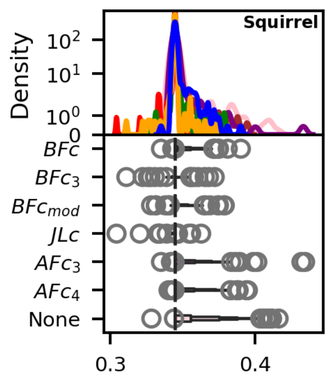


Fig.: We look at multiple definitions discrete curvatures defined on graphs to assess their performance with respect to no rewiring (None).

c. Assess the convergence of the distribution

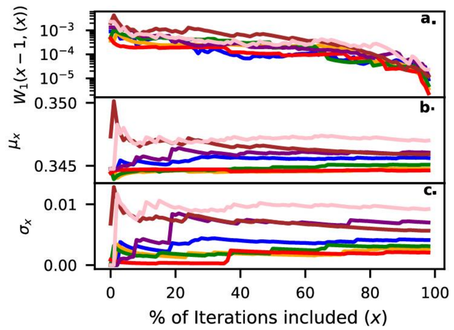


Fig.: a. Wasserstein distance of two subsequent distributions. **b., c.** Mean and standard deviation when including x% of the total iterations.

Dataset: General

Our evaluation on 11 datasets shows that this is not limited to specific choices, but present in all benchmark datasets for GNNs.

Dataset	Edges rewired	Condition 2 (%)
Texas	*89*	*6 (6.7 %)*
Cornell	*126*	*15 (11.90 %)*
Wisconsin	*136*	*11 (8.09 %)*
Chameleon	*2441*	*141 (5.78 %)*
Actor	*1000*	*237 (23.70 %)*
Squirrel	*787*	*34 (4.32 %)*
Cora	*100*	*68 (68.0 %)*
Citeseer	*84*	*24 (28.57 %)*
Pubmed	*166*	*116 (69.88 %)*
MUTAG	*3497*	*1128 (35.16 %)*
PROTEINS	*50936*	*5944 (11.67 %)*

Table: Percentage of edges that satisfy condition 2b for all datasets during SDRF rewiring.

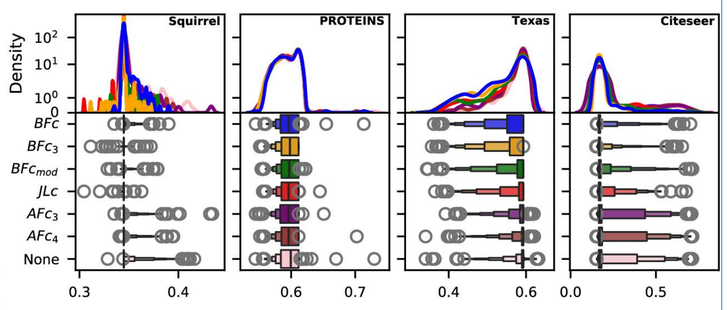


Fig.: Distribution of test accuracies over hyperparameter sweeps.

Oversquashing Theorem
[Topping *et al.* 2022]

Consider an MPNN and let $i \sim j$ be an edge.

If a $\delta > 0$ exists such that $\boldsymbol{BFc(i,j) < -2 + \delta}$ and for which the following holds:

$$\delta < \frac{1}{\sqrt{max\{d_i,d_j\}}} \text{ and } \delta < \frac{1}{\gamma_{max}} \textbf{ (condition 2*)}$$

Here γ_{max} represent the degeneracy factor of 4-cycles. Then we can bound the Jacobian of message passing as

$$\frac{1}{|Q_j|}\sum_{k \in Q_j}\left|\frac{\partial h_k^{(\ell_0+2)}}{\partial h_i^{(\ell_0)}}\right| < C\delta^{\frac{1}{4}}$$

where $Q_j \subset S_2(i)$ satisfying $|Q_j| > \delta^{-1}$.

** We note that a softer version of this condition is sufficient, where $\delta < \frac{1}{\sqrt{max\{d_i,d_j\}}}$ is replaced by $\delta < 1/\#_\triangle$ (**condition 2b**).*

4. Take Aways

- Our work re-evaluates **curvature-based rewiring** and puts into question its effectiveness on improving GNNs. It highlights the importance to closely check theoretical results and experiments.
- By analysing results from **hyperparameter sweeps** as distributions we show the influence of outliers on evaluating new methods. We advocate therefore to consider these distributions during evaluation of future techniques.

[1] Data Analytics Laboratory, Vrije Universiteit Brussel ; [2]School of Engineering and Applied Sciences, Harvard University

floriano.tori@vub.be