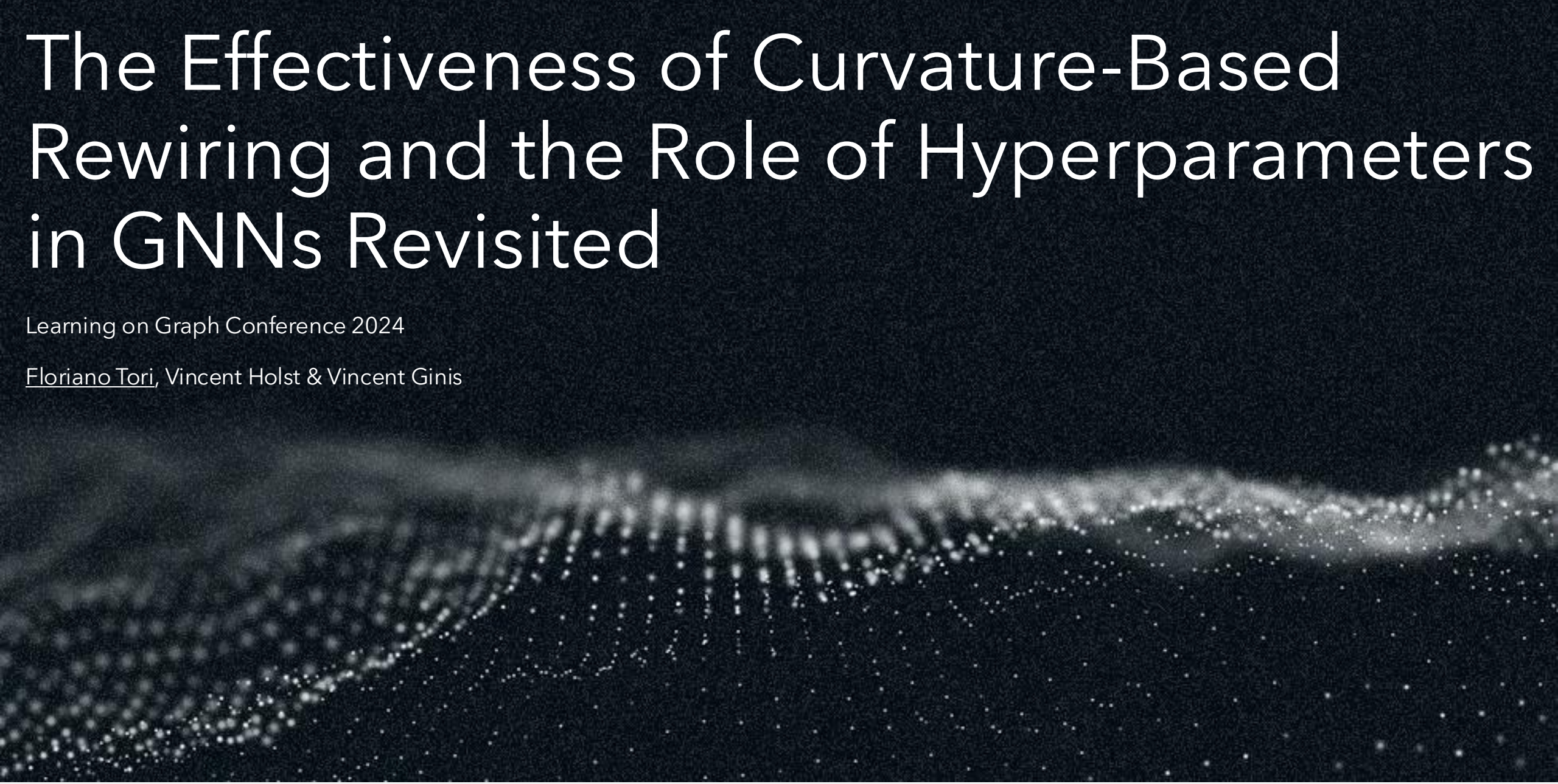
The Effectiveness of Curvature-Based Rewiring and the Role of Hyperparameters in GNNs Revisited

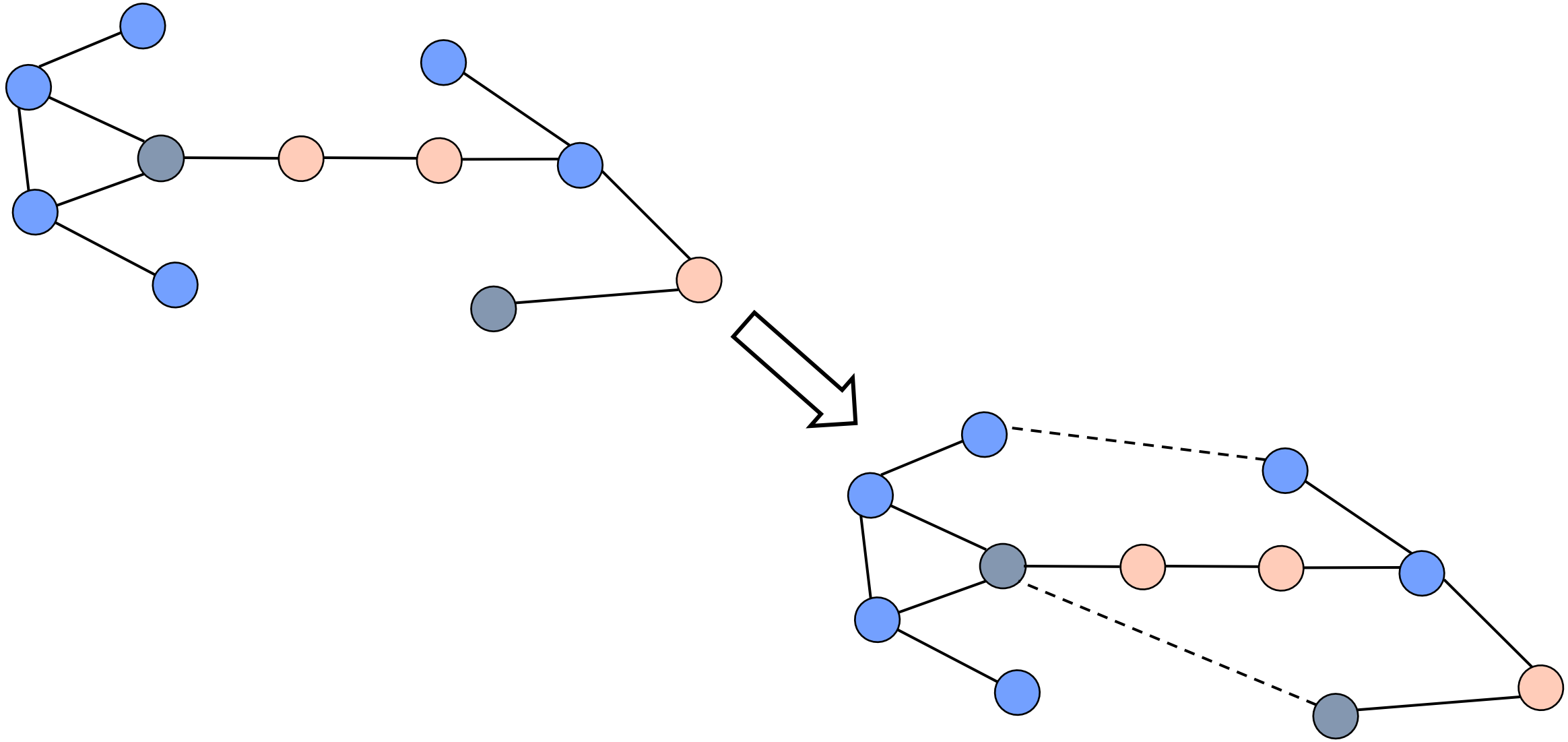


Learning on Graph Conference 2024

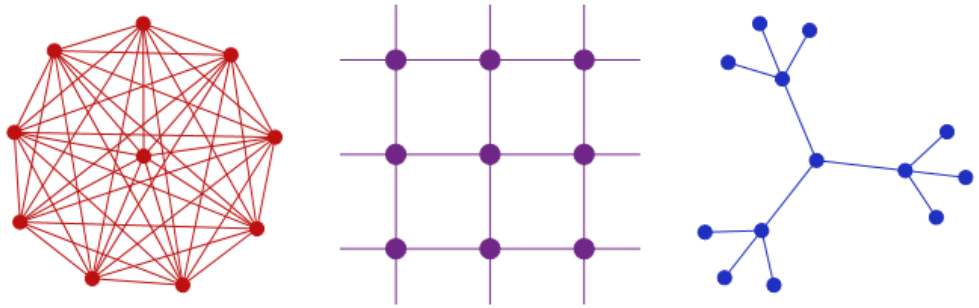
Floriano Tori, Vincent Holst & Vincent Ginis

VUB DATA ANALYTICS LABORATORY RESEARCH GROUP

Altering the graph structure is a standard approach **to alleviate** message passing GNNs from **oversquashing**



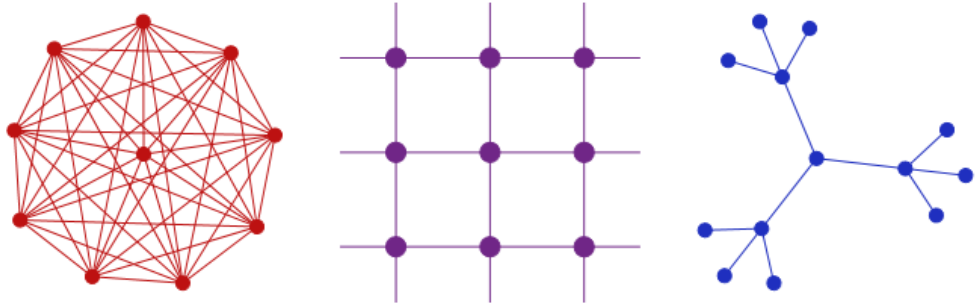
For this purpose, **discrete graph curvature notions** are used to **detect** and rewire around **bottlenecks.**



(a) Clique (> 0) (b) Grid ($= 0$) (c) Tree (< 0)

Topping, Jake, et al. "Understanding over-squashing and bottlenecks on graphs via curvature." International Conference on Learning Representations

For this purpose, **discrete graph curvature notions** are used to **detect** and rewire around **bottlenecks.**



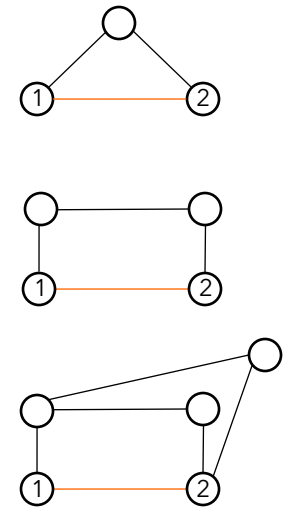
(a) Clique (> 0) (b) Grid ($= 0$) (c) Tree (< 0)

Balanced Forman Curvature (BFc) ~

$\#_{\triangle}$

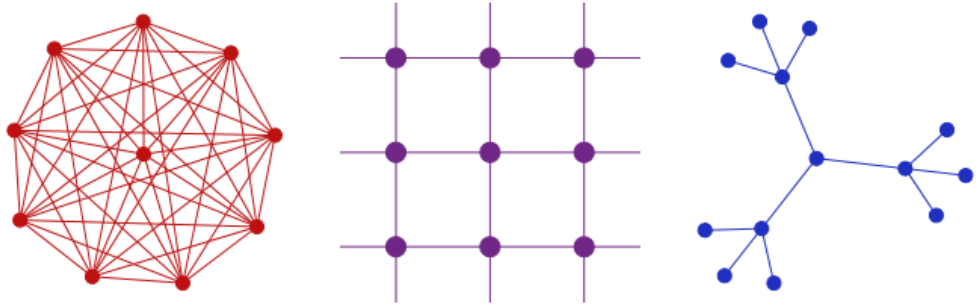
$\#_{\blacksquare}$

γ_{max}



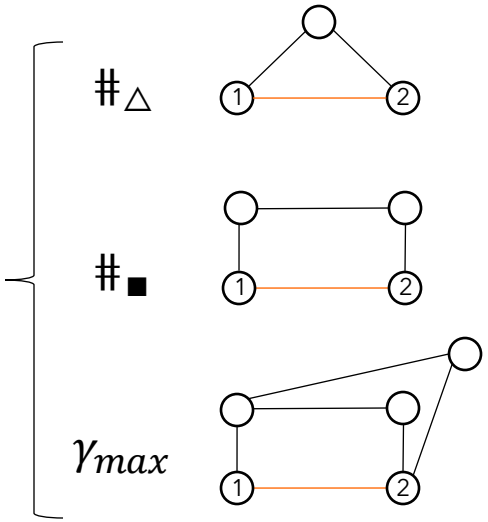
Topping, Jake, et al. "Understanding over-squashing and bottlenecks on graphs via curvature." International Conference on Learning Representations

For this purpose, **discrete graph curvature notions** are used to **detect** and rewire around **bottlenecks.**



(a) Clique (> 0) (b) Grid ($= 0$) (c) Tree (< 0)

Balanced Forman Curvature (BFc) ~



Oversquashing Theorem [Topping *et al.* 2022]

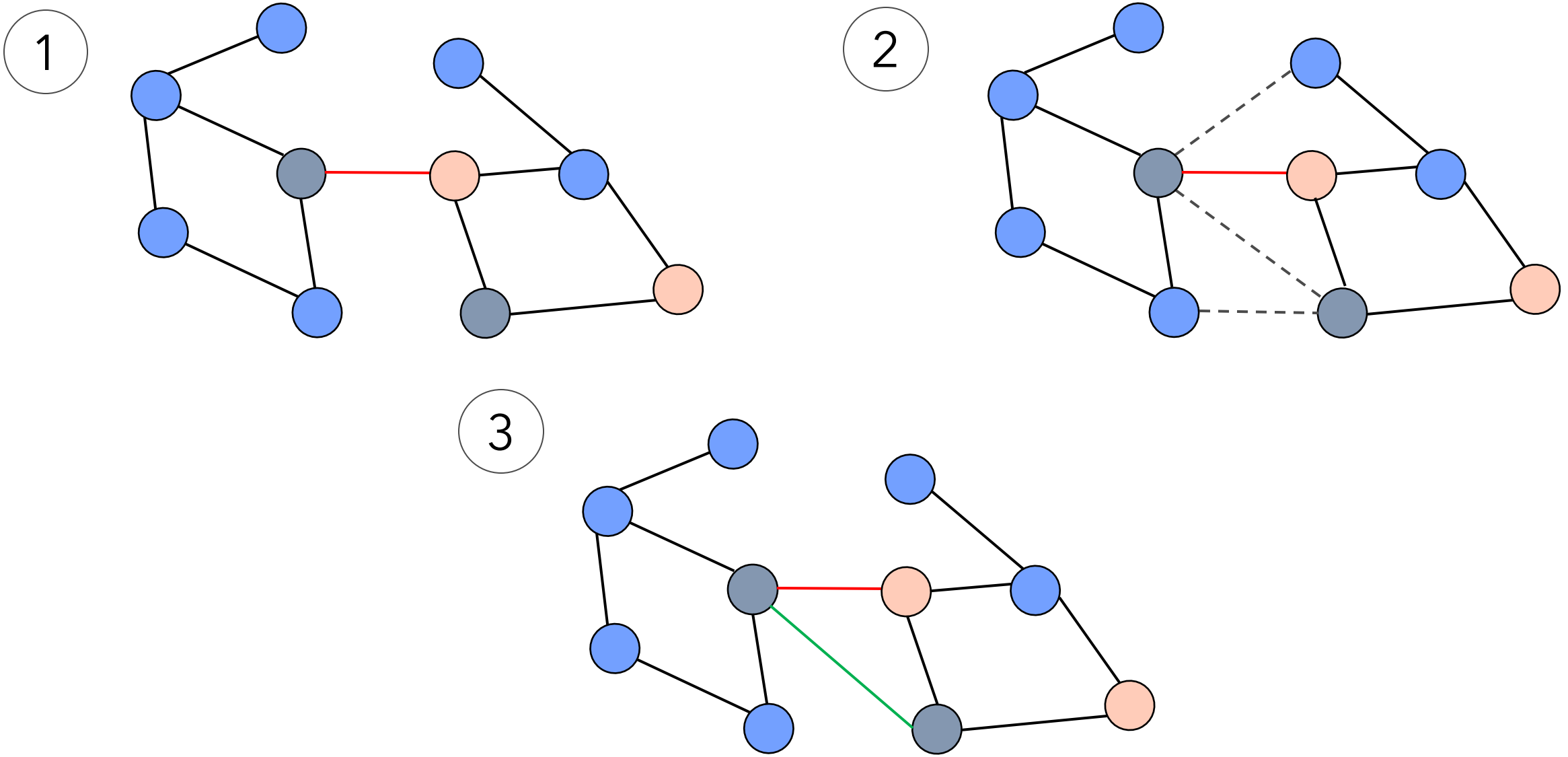
Consider a MPNN creating embedding vectors h_i^l at layer l. Suppose we have an edge $i \sim j$ for which $BFc(i,j) < -2 + \delta$. Then we can bound the Jacobian of message passing as

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

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

Topping, Jake, et al. "Understanding over-squashing and bottlenecks on graphs via curvature." International Conference on Learning Representations

During **Stochastic Discrete Ricci Flow (SDRF)** negatively curved edges are rewired around in order to reduce their curvature



Topping, Jake, et al. "Understanding over-squashing and bottlenecks on graphs via curvature." International Conference on Learning Representations

The oversquashing theorem also contains **a condition** in order for **edges** to be **identified as bottlenecks** which is **not checked** during rewiring.

Oversquashing Theorem [Topping *et al.* 2022]

Consider a MPNN creating embedding vectors h_i^l at layer l. Suppose we have an edge $i \sim j$ for which $BFc(i,j) < -2 + \delta$ and for which the following holds:

$$\delta < \frac{1}{\sqrt{\max\{d_i, d_j\}}} \text{ and } \delta < \frac{1}{\gamma_{max}}$$

Then we can bound the Jacobian of message passing as

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

The **benchmark datasets** of SDRF **do not contain** enough **edges that satisfy the conditions** of the oversquashing theorem.

Table : Percentage of edges that satisfy condition 2 during SDRF rewiring.

Dataset	Edges rewired	Condition 2 (%)
Texas	*89*	*0 (0 %)*

The **benchmark datasets** of SDRF **do not contain** enough **edges that satisfy the conditions** of the oversquashing theorem.

Table : Percentage of edges that satisfy condition 2(b) during SDRF rewiring.

Dataset	Edges rewired	Condition 2 (%)	Condition 2b (%)
Texas	*89*	*0 (0 %)*	*6 (6.7 %)*

Suppose we have an edge $i \sim j$ for which $BFc(i,j) < -2 + \delta$ and for which the following holds:

Condition 2

$$\delta < \frac{1}{\sqrt{\max\{d_i, d_j\}}} \text{ and } \delta < \frac{1}{\gamma_{max}}$$

Condition 2b

$$\delta < \frac{1}{\#_{\triangle}} \text{ and } \delta < \frac{1}{\gamma_{max}}$$

This is **not a saturation type effect**, as edges that do not satisfy the condition are selected **during the entire rewiring process.**

Table : Percentage of edges that satisfy condition 2(b) during SDRF rewiring.

Dataset	Edges rewired	Condition 2 (%)	Condition 2b (%)
Texas	*89*	*0 (0 %)*	*6 (6.7 %)*

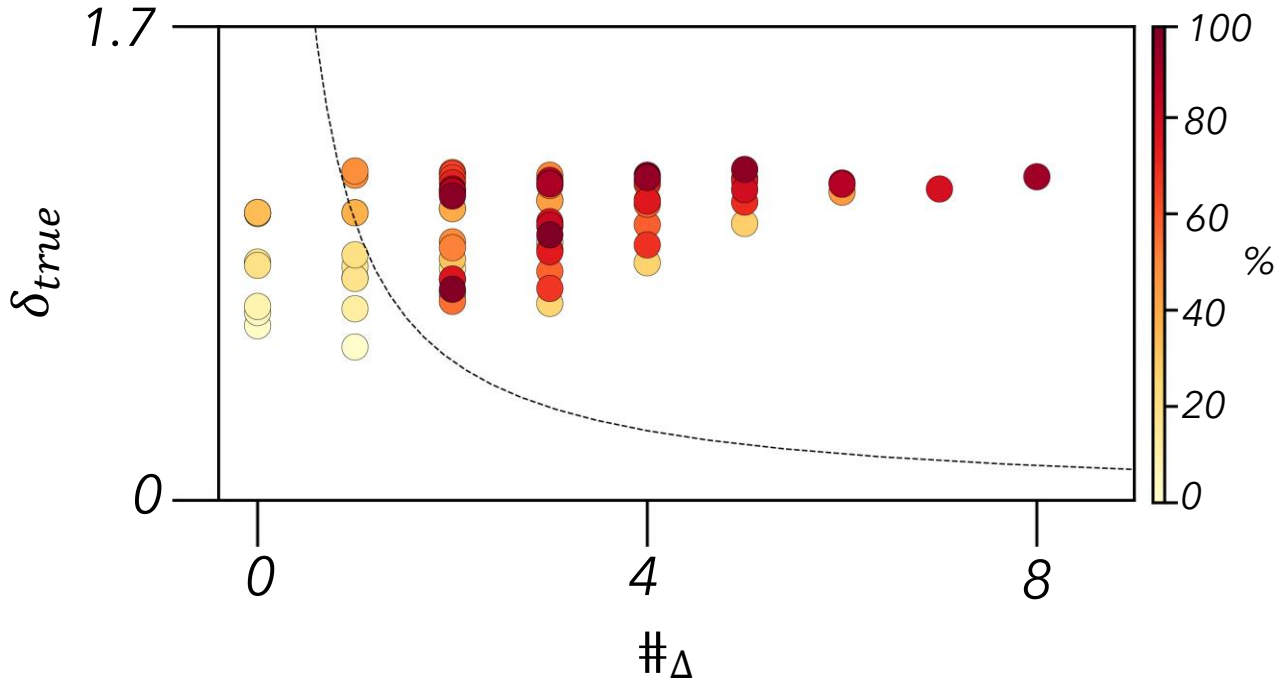


Fig. :Visual representation of the selected edges during SDRF rewiring that do not satisfy condition 2b. Colour indicates at which % of total rewiring the edge was selected. Dotted line indicates $1/\#_\Delta$

This comes from the distributions of **curvature values**, as there are **not enough edges** very **close to** -2.

Table : Percentage of edges that satisfy condition 2(b) during SDRF rewiring.

Dataset	Edges rewired	Condition 2 (%)	Condition 2b (%)
Texas	*89*	*0 (0 %)*	*6 (6.7 %)*

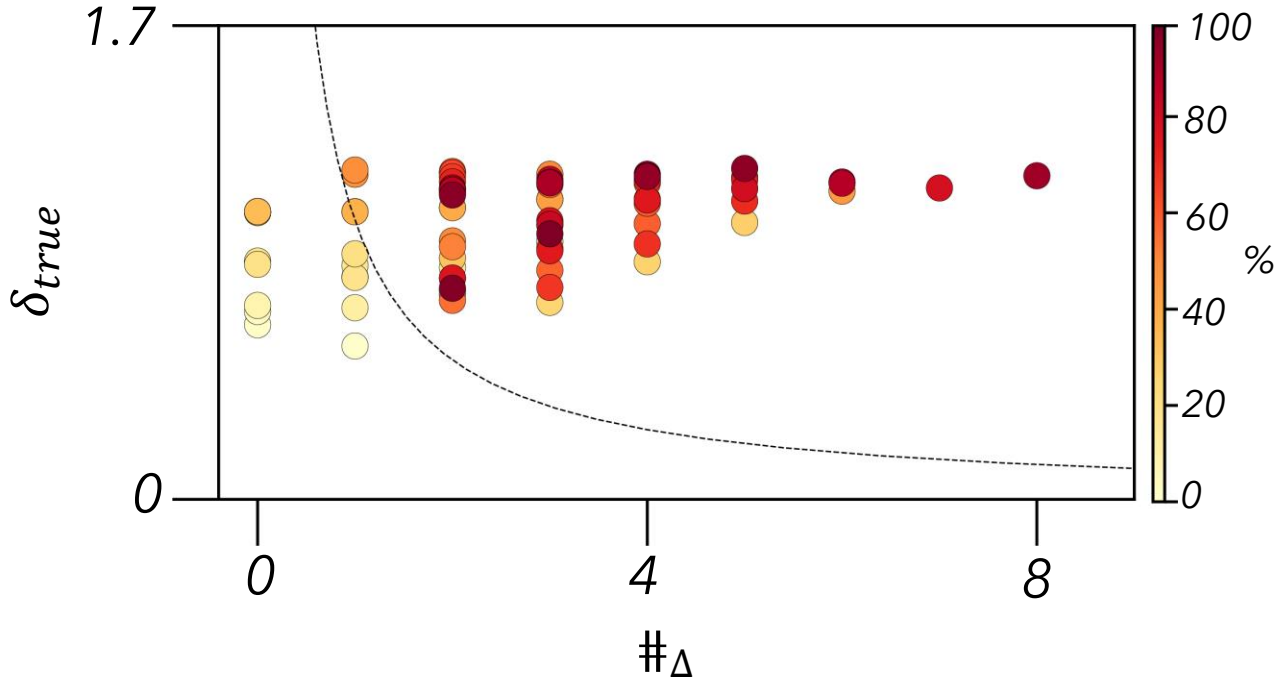


Fig. :Visual representation of the selected edges during SDRF rewiring that do not satisfy condition 2b. Colour indicates at which % of total rewiring the edge was selected. Dotted line indicates $1/\#_\Delta$

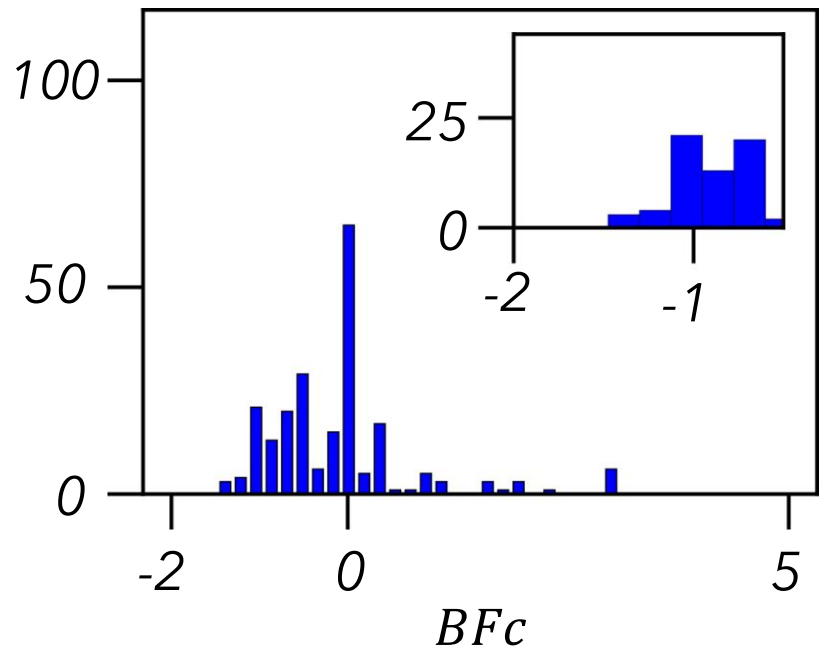


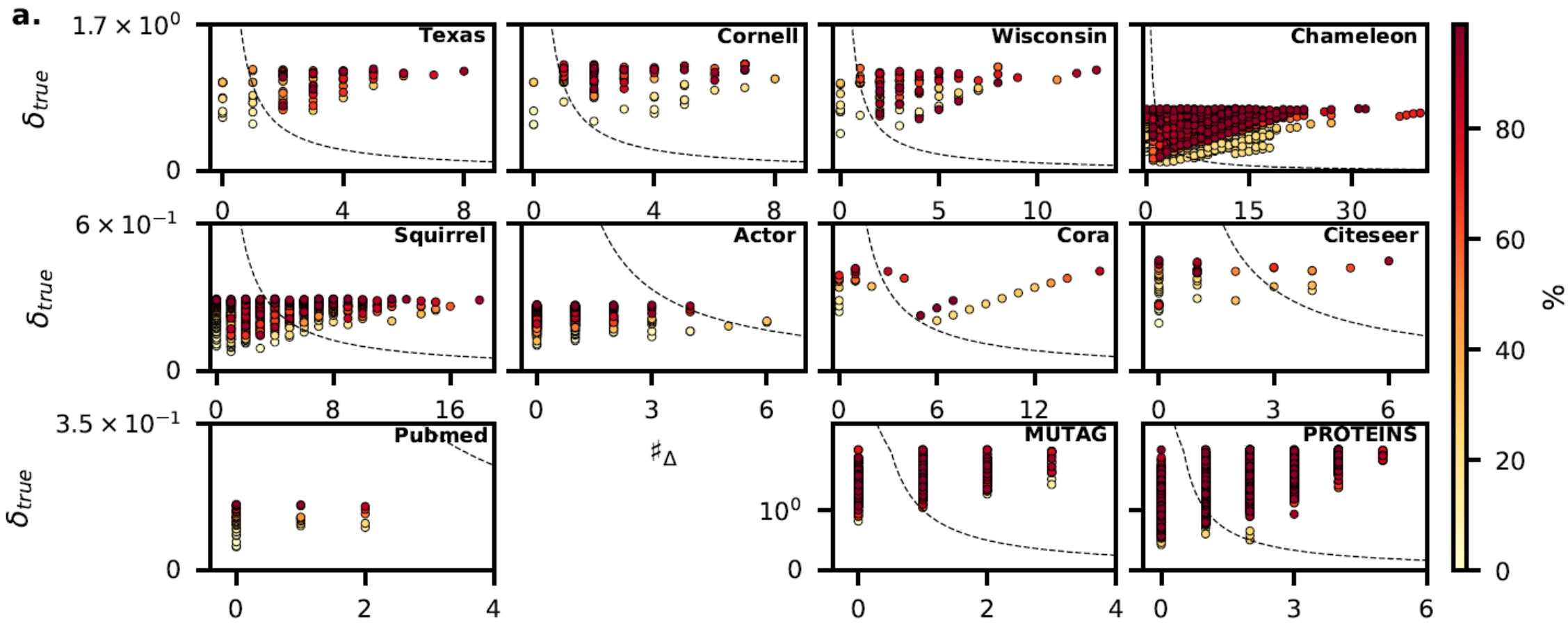
Fig. : Values of curvatures for all edges in the Texas dataset.

The results from this analysis **are similar** over all datasets previously used for **evaluating rewiring methods**.

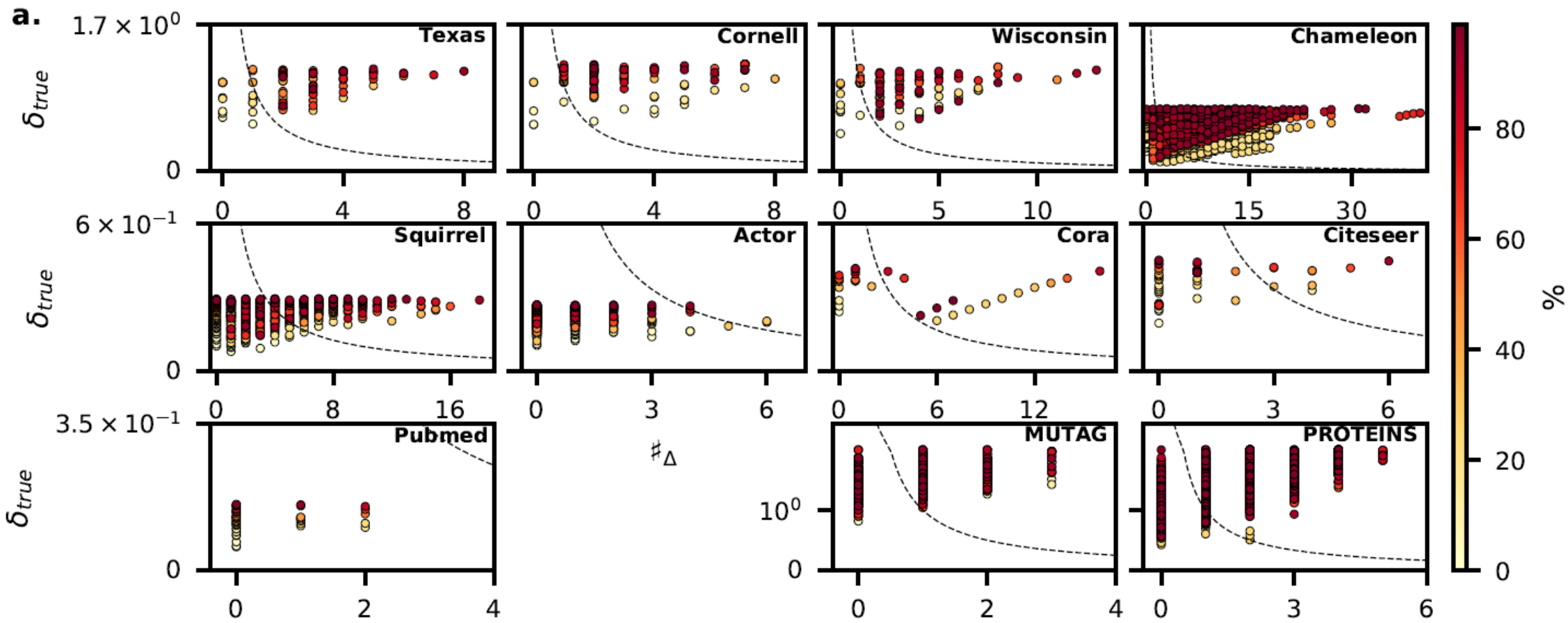
Table : Percentage of edges that satisfy condition 2(b) for all datasets during SDRF rewiring.

Dataset	Edges rewired	Condition 2 (%)	Condition 2b (%)
Texas	*89*	*0 (0 %)*	*6 (6.7 %)*
Cornell	*126*	*0 (0 %)*	*15 (11.90 %)*
Wisconsin	*136*	*0 (0 %)*	*11 (8.09 %)*
Chameleon	*2441*	*4 (0.16 %)*	*141 (5.78 %)*
Actor	*1000*	*11 (1.1 %)*	*237 (23.70 %)*
Squirrel	*787*	*0 (0 %)*	*34 (4.32 %)*
Cora	*100*	*0 (0 %)*	*68 (68.0 %)*
Citeseer	*84*	*0 (0 %)*	*24 (28.57 %)*
Pubmed	*166*	*25 (16.06 %)*	*116 (69.88 %)*
MUTAG	*3497*	*0 (0 %)*	*1128 (35.16 %)*
PROTEINS	*50936*	*0 (0 %)*	*5944 (11.67 %)*

The results from this analysis **are similar** over all datasets previously used for **evaluating rewiring methods**.



The results from this analysis **are similar** over all datasets previously used for **evaluating rewiring methods**.



At this point we face **a contradiction** with reported accuracies **in previous works.**

To **analyse the performance** differences, we performed **hyperparameter sweeps** for different curvature definitions.

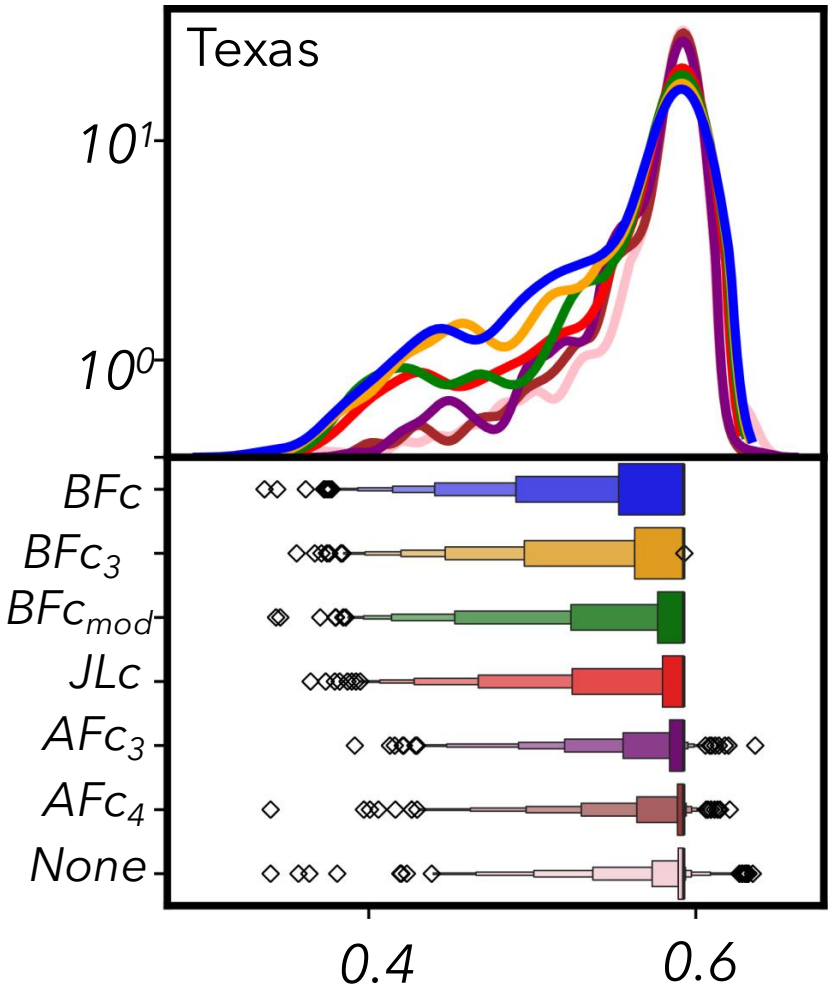


Fig. : Distribution of the mean test accuracy over the entire hyperparameter sweep

Jürgen Jost and Shiping Liu. Ollivier's ricci curvature, local clustering and curvature-dimension inequalities on graphs. *Discrete & Computational Geometry*, 51(2):300–322, 2014
Fesser, Lukas, and Melanie Weber. "Mitigating over-smoothing and over-squashing using augmentations of Forman-Ricci curvature." *Learning on Graphs Conference*. PMLR, 2024

To **analyse the performance** differences, we performed **hyperparameter sweeps** for different curvature definitions.

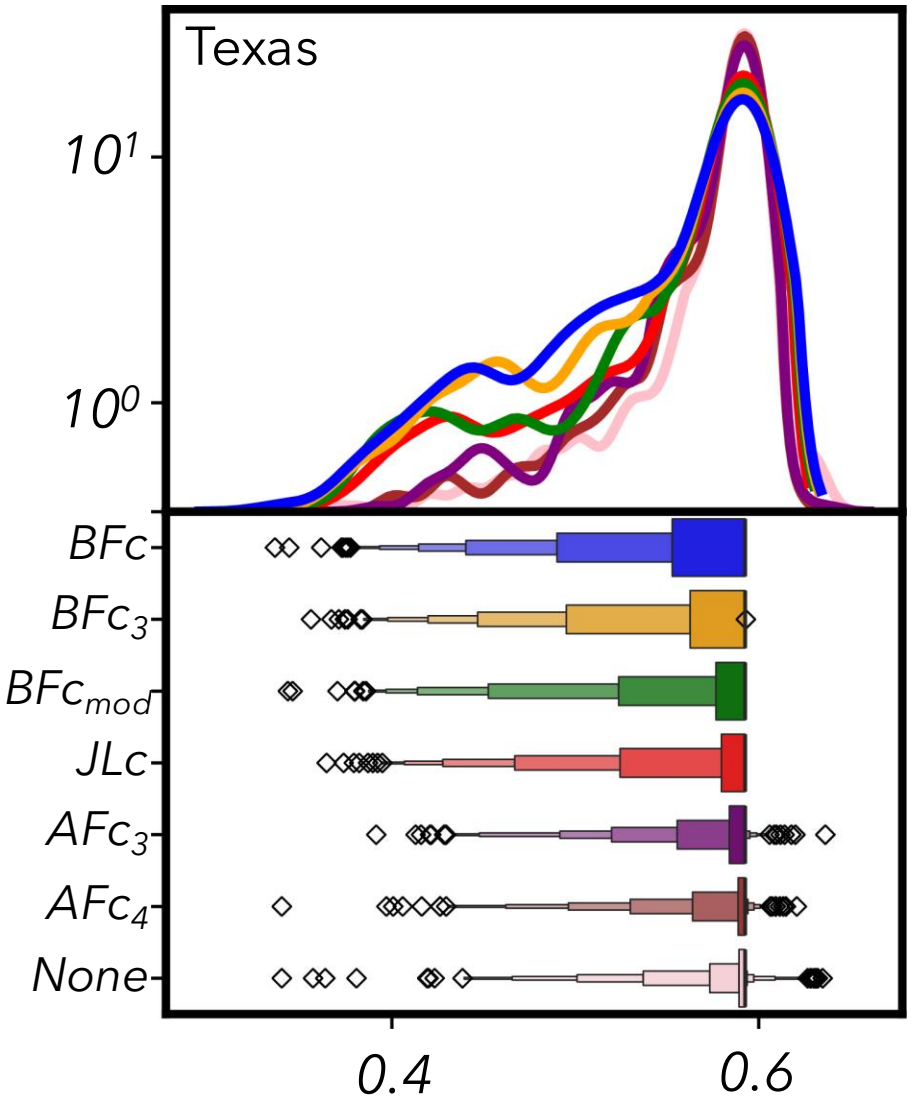


Fig. : Distribution of the mean test accuracy over the entire hyperparameter sweep

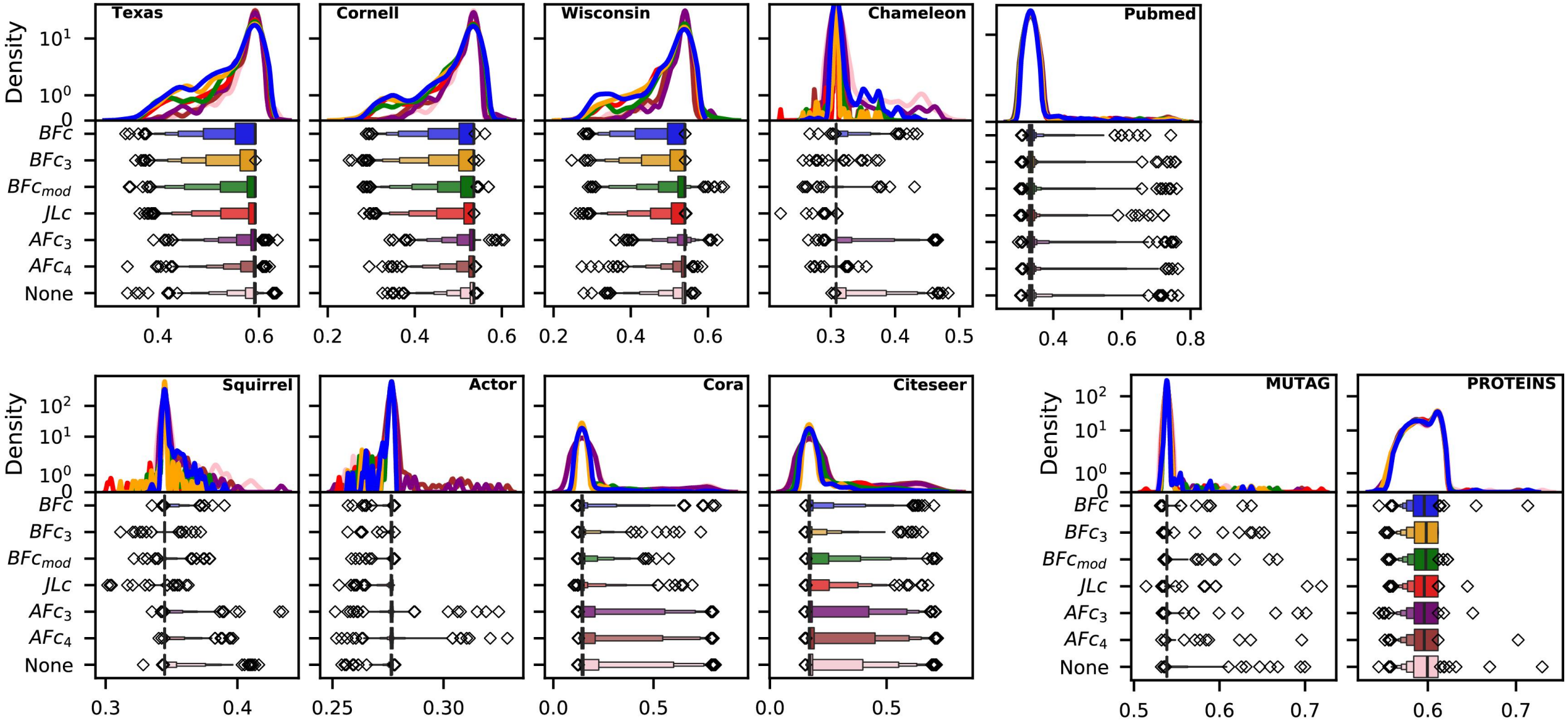
Table : Average mean test accuracy of the top 10% hyperparameter configurations

	Texas
BFc	59.26 ± 0.00
BFc_3	59.26 ± 0.00
BFc_{mod}	59.26 ± 0.00
JLc	59.26 ± 0.00
AFc_3	59.58 ± 0.52
AFc_4	59.79 ± 0.54
None	59.95 ± 1.15

Table : Top mean test accuracy results

	BFc	BFc_3	BFc_{mod}	*JLc*	AFc_3	AFc_4	*None*
Texas	*59.26*	*59.30*	*59.26*	*59.26*	*63.63*	*62.07*	*63.48*

The hyperparameter sweeps for the **different datasets** also show that **performances** after rewiring are **not systematic improvements**.

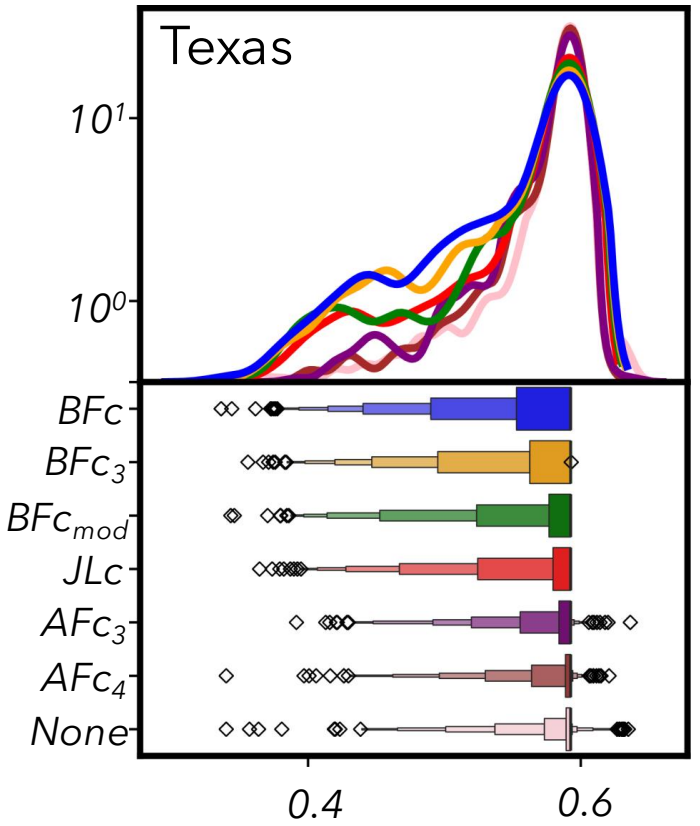


Takeaways & Future work

Dataset	Edges rewired	Condition 2 (%)
Texas	*89*	*0 (0 %)*

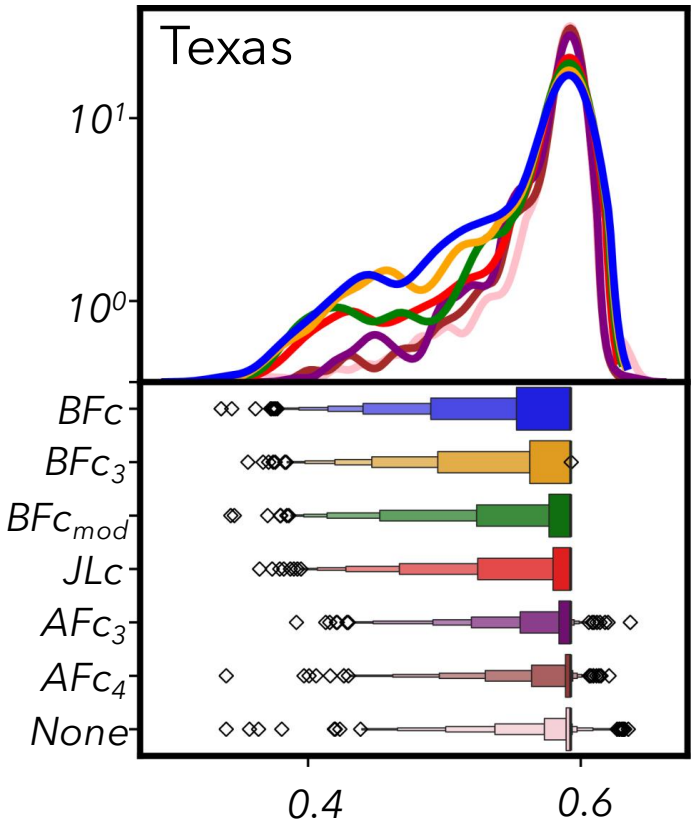
Takeaways & Future work

Dataset	Edges rewired	Condition 2 (%)
Texas	*89*	*0 (0 %)*



Takeaways & Future work

Dataset	Edges rewired	Condition 2 (%)
Texas	*89*	*0 (0 %)*



Published in Transactions on Machine Learning Research (05/2024)

Where Did the Gap Go?
Reassessing the Long-Range Graph Benchmark

Published as a conference paper at ICLR 2023

Leave Graphs Alone: Addressing Over-Squashing without Rewiring

A CRITICAL LOOK AT THE EVALUATION OF GNNS UNDER HETEROPHILY: ARE WE REALLY MAKING PROGRESS?

Tönshoff, Jan, et al. "Where Did the Gap Go? Reassessing the Long-Range Graph Benchmark." *Transactions on Machine Learning Research*
Tortorella, Domenico, and Alessio Micheli. "Leave Graphs Alone: Addressing Over-Squashing without Rewiring." *The First Learning on Graphs Conference.*
Platonov, Oleg, et al. "A critical look at the evaluation of GNNs under heterophily: Are we really making progress?." *The 11th International Conference on Learning Representations.*

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

Floriano Tori
Data Analytics Laboratory
Vrije Universiteit Brussel
Floriano.Tori@vub.be

Vincent Holst
Data Analytics Laboratory
Vrije Universiteit Brussel
Vincent.Thorge.Holst@vub.be

Vincent Ginis
Data Analytics Laboratory
Vrije Universiteit Brussel
SEAS, Harvard University
Vincent.Ginis@vub.be

https://github.com/FloTori/Revisiting-Graph-Rewiring

Come visit at Poster Session 2!

DATA ANALYTICS LABORATORY RESEARCH GROUP

Harvard John A. Paulson School of Engineering and Applied Sciences