Restructuring Graph for Higher Homophily via Adaptive Spectral Clustering

- **Abstract**: Although the ability to handle less-homophilic graphs is restricted, classical GNNs still stand out in several nice properties such as efficiency, simplicity, and explainability. In this work, we propose a novel graph restructuring method that can be integrated into any type of GNNs, including classical GNNs, to leverage the benefits of existing GNNs while alleviating their limitations. Our method learns to cluster nodes using eigenvectors beyond spectral clustering. We also proposed a new density-aware homophilic metric to better reflect the homophily of a graph.

Shouheng Li[1,3], Dongwoo Kim[2], Qing Wang[1]

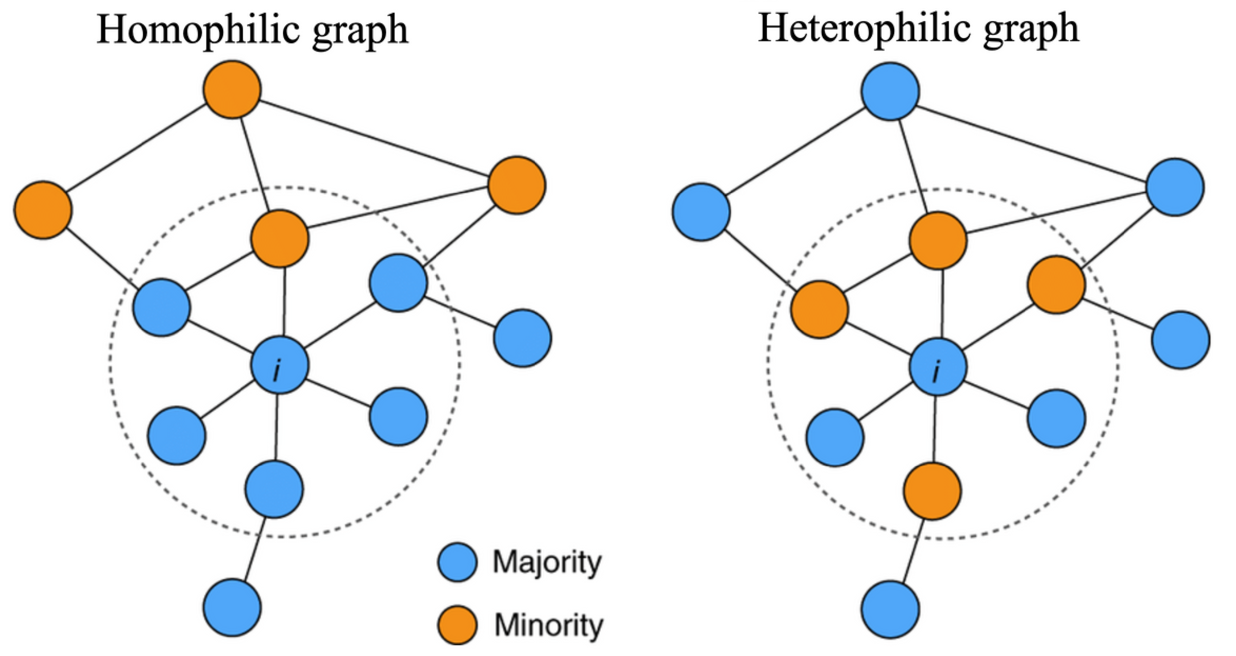
shouheng.li@anu.edu.au

[1]School of Computing, Australian National University

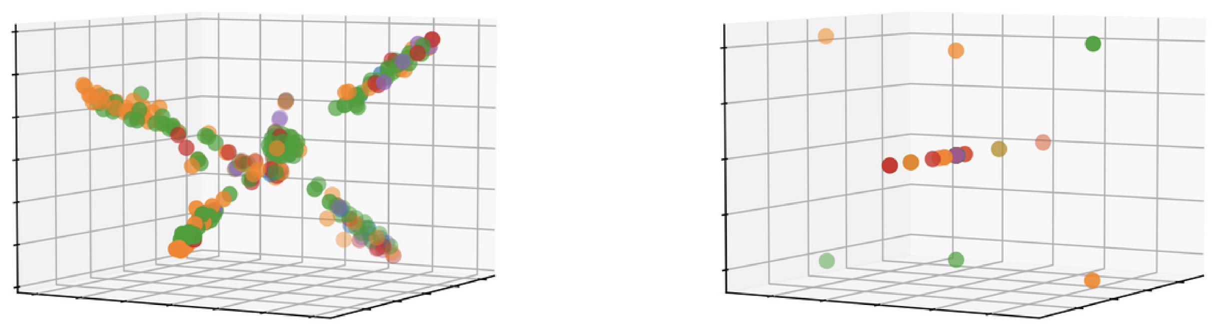
[2]CSE & GSAI, POSTECH

[3]Data61 CSIRO

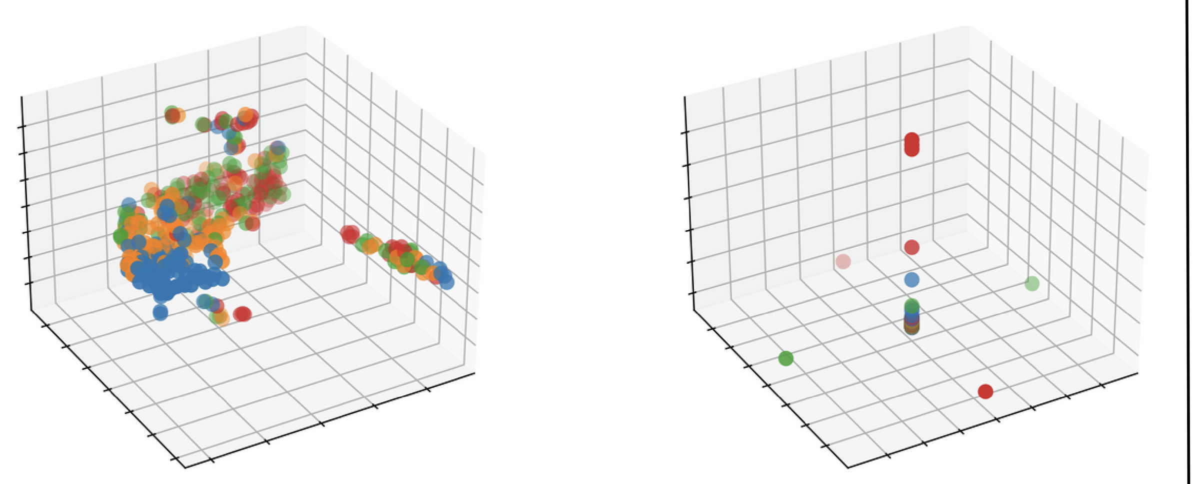
Graphs



Spectral clustering



a. WISCONSIN: T-SNE with the leading 5 eigenvectors (left); visualized using the 22th, 44th and 206th eigenvectors (right).



b. EUROPE AIRPORT: T-SNE with the leading 5 eigenvectors (left); visualized using the 366th, 382th and 3rd eigenvectors (right).

Density-aware homophily metric

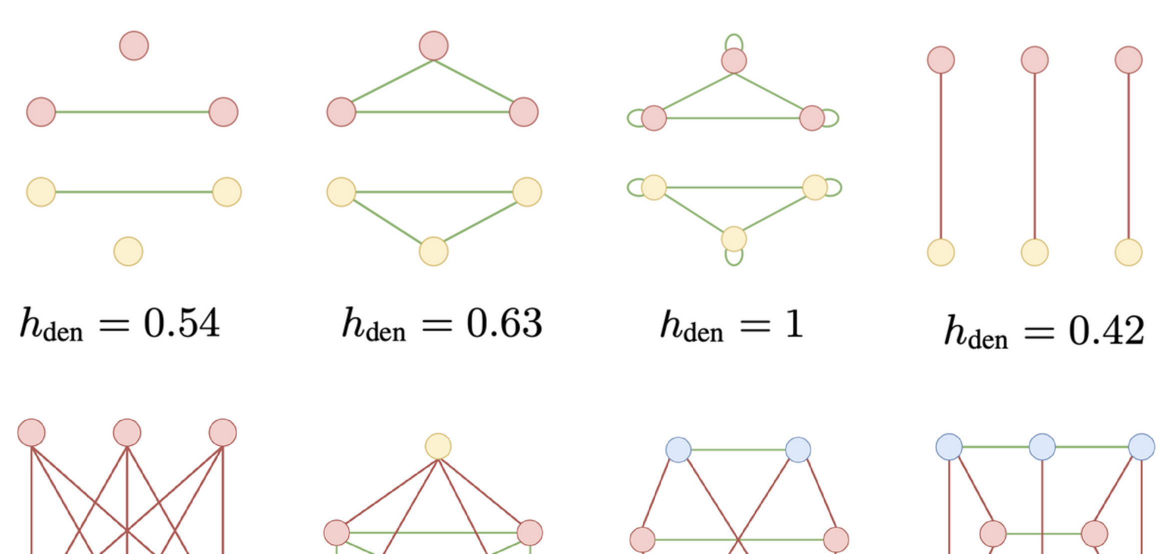
$$d_k = \frac{|(u,v) \in E : k_u = k_v = k|}{|Y_k||Y_k|},$$
$$\bar{d}_k = \max\{d_{kj} : j = 0, ..., K-1; j \neq k\}$$

$$\hat{h}_{\text{den}} = \min\{d_k - \bar{d}_k\}_{k=0}^{K-1}$$

Intra-class edge density

Inter-class edge density

$$h_{\text{den}} = \frac{1 + \hat{h}_{\text{den}}}{2}$$



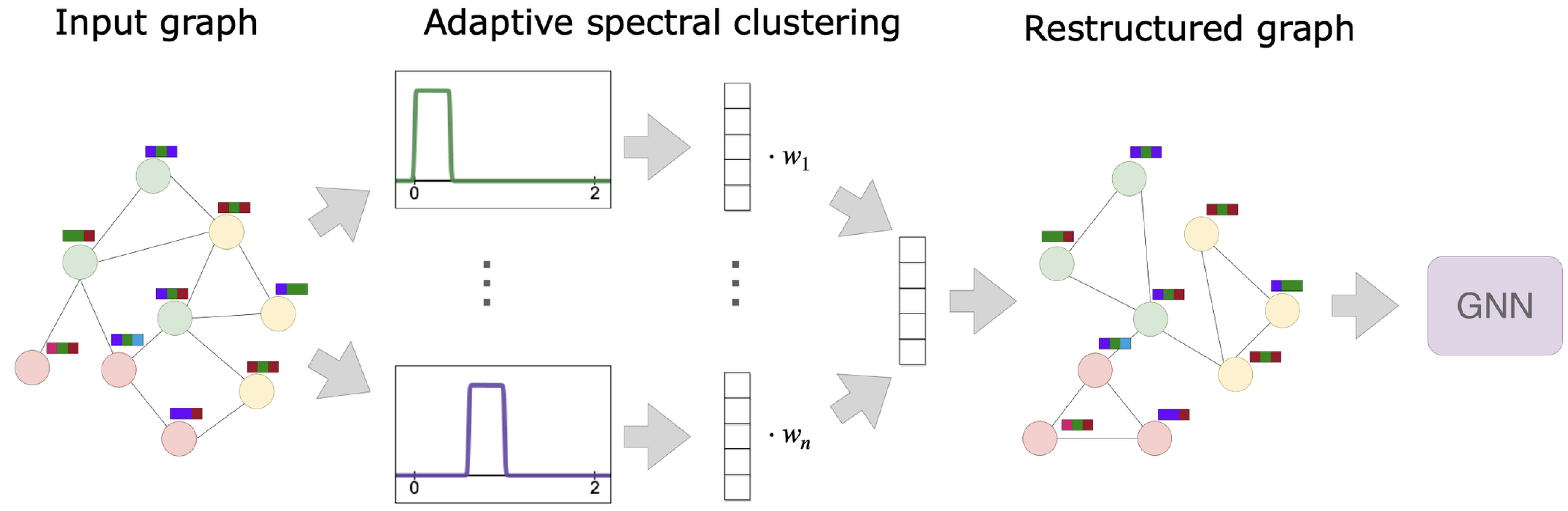
$h_{\text{den}} = 0.54$ $h_{\text{den}} = 0.63$ $h_{\text{den}} = 1$ $h_{\text{den}} = 0.42$

$h_{\text{den}} = 0$ $h_{\text{den}} = 0.15$ $h_{\text{den}} = 0.42$ $h_{\text{den}} = 0.5$

Selected references

- Klicpera, J.; Weißenberger, S.; and Günnemann, S. Diffusion Improves Graph Learning. NeurIPS 2019
- Zhu, J., Yan, Y., Zhao, L., Heimann, M., Akoglu, L., Koutra, D.: Beyond homophily in graph neural networks: Current limitations and effective designs. NeurIPS 2020
- Tremblay, N.; Puy, G.; Borgnat, P.; Gribonval, R.; and Vandergheynst, P. 2016. Accelerated spectral clustering using graph filtering of random signals. ICASSP 2016
- Balcilar, M.; Renton, G.; Héroux, P.; Gaüzère, B.; Adam, S.; and Honeine, P. Analyzing the Expressive Power of Graph Neural Networks in a Spectral Perspective. ICLR 2021

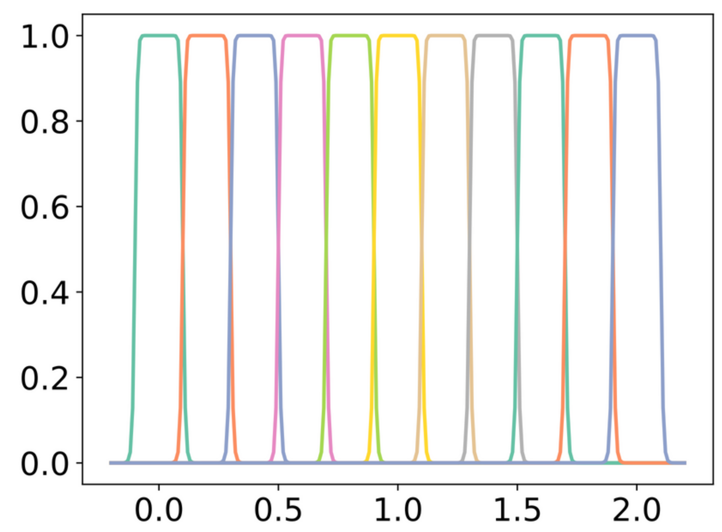
Restructure a graph



Adaptive spectral clustering

Rectangular band-pass filters

$$U\hat{g}_{s,a}(\Lambda)U^H x = \frac{1}{s^{2m}}\left(\left(\frac{L - aI}{2+\hat{\epsilon}}\right)^{2m} + \frac{I}{s^{2m}}\right)^{-1}$$



Spectral clustering as spectral filtering

Laplacian matrix

Node feature

Rectangular band-pass filter

Random node feature

$$\Gamma_{s,a} = \hat{g}_{s,a}(L)(R \frown X)$$

Random projection (Johnson-Lindenstrauss lemma)

Concatenation

Learnable function

$$\Gamma = (\Gamma_{s_1,a_1} \frown \Gamma_{s_2,a_2} \frown ...)$$

$$H = \Theta(\Gamma).$$

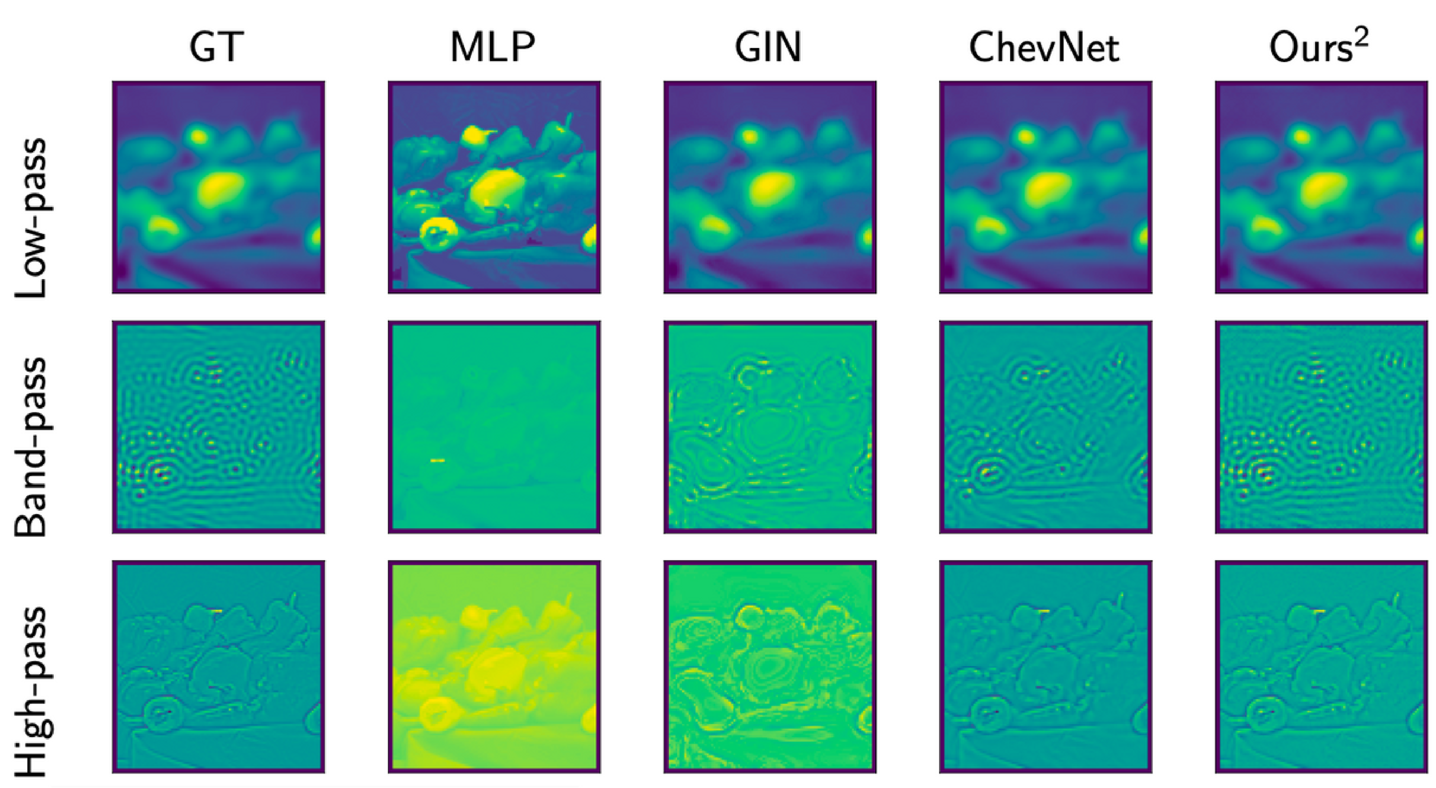
Triplet loss function

$$\mathcal{L}(\Theta) = \sum_{\substack{i,j \in V_Y \\ k \in \mathcal{N}_Y(i) \\ y_i = y_j}} \left[||H_{i\cdot} - H_{j\cdot}||^2 - ||H_{i\cdot} - H_{k\cdot}||^2 + \epsilon\right]_+$$

positive pairs nagative pairs

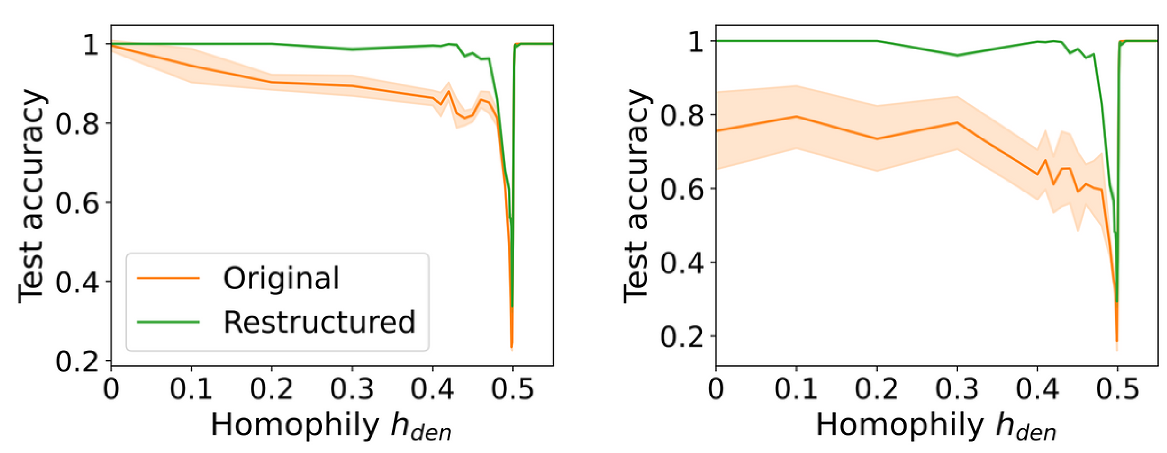
Experimental results

Spectral expressivity



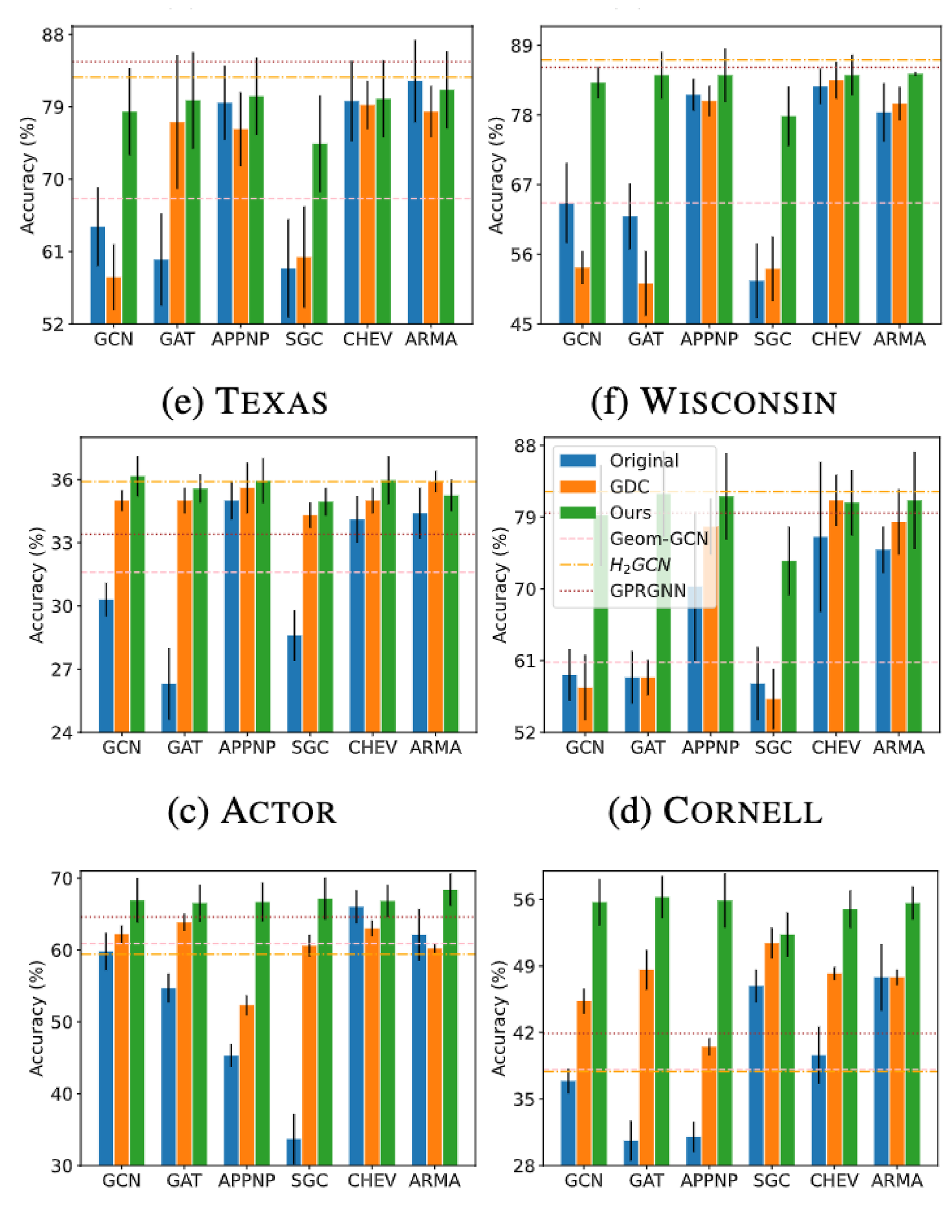
[2]uses 20× less parameters than ChevNet

Synthetic heterophilic graphs



(a) GCN performance (b) SGC performance

Node classification



(e) TEXAS (f) WISCONSIN

(c) ACTOR (d) CORNELL

(a) CHAMELEON (b) SQUIRREL

- The model can best recover images filtered using low-, band- and high-pass functions.
- The restructured real-world graphs have higher homophily which improves GNNs performance by an average 25%.
- The method yields high accuracy on synthetic heterophilic graphs.