

GraphCroc: Cross-Correlation Autoencoder for Graph Structural Reconstruction

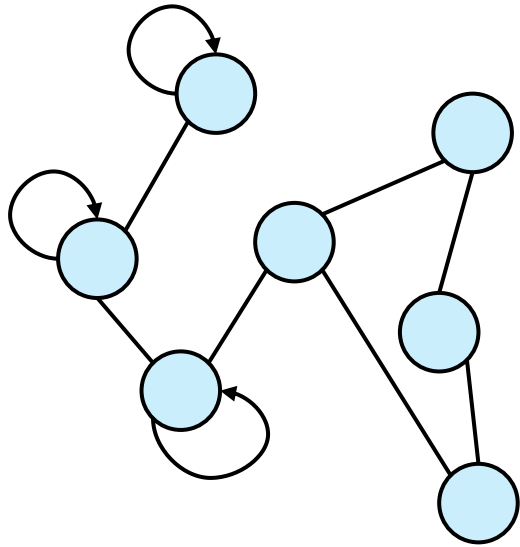
Shijin Duan*, Ruyi Ding*, Jiaying He, Aidong Adam Ding, Yunsi Fei, Xiaolin Xu

Northeastern University

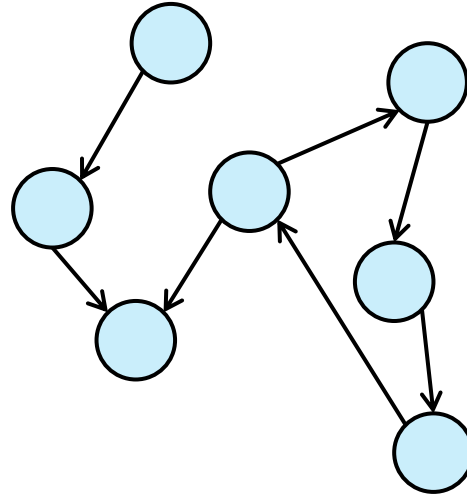
{duan.s, ding.ruy, he.jiayi, a.ding, y.fei, x.xu}@northeastern.edu



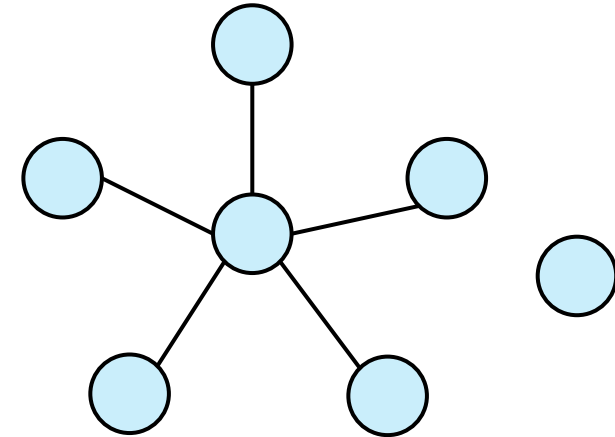
Graph Data



Undirected
Asymmetric
(with self-loop)

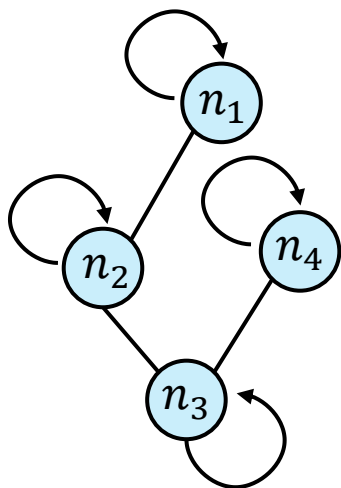


Directed
Asymmetric



Symmetric
with islands

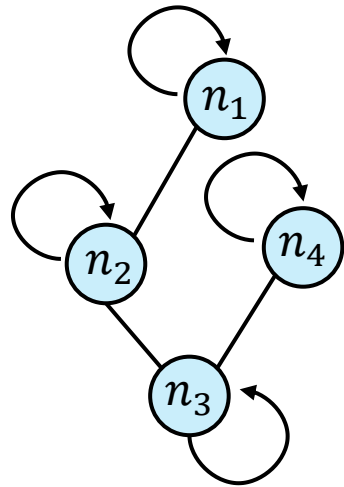
Adjacency Matrix:



$G = \{V, E\}$

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Graph Structural Reconstruction



$$G = \{V, E\}$$

Adjacency Matrix:

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Self-correlation:

$$\text{encoder: } Z = \Phi(Z|G) = f(X, A)$$

$$\text{decoder: } \tilde{A} = \Theta(A|Z) = \text{sigmoid}(ZZ^T)$$

$$\tilde{A} = \begin{bmatrix} 0.9 & 0.8 & 0.2 & 0.2 \\ 0.8 & 0.6 & 0.6 & 0.1 \\ 0.2 & 0.6 & 0.7 & 0.9 \\ 0.2 & 0.1 & 0.9 & 0.8 \end{bmatrix}$$

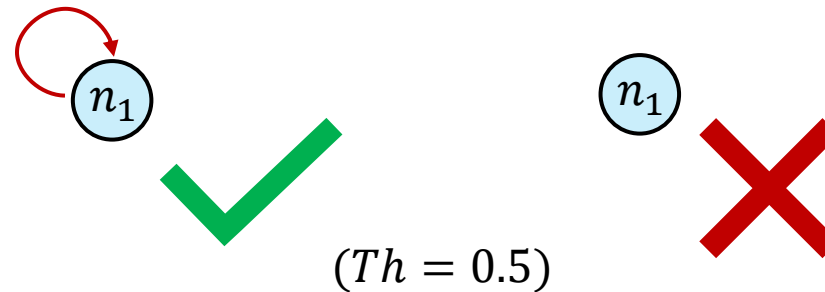
Self-correlation:

$$\text{encoder: } Z = \Phi(Z|G) = f(X, A)$$

$$\text{decoder: } \tilde{A} = \Theta(A|Z) = \text{sigmoid}(ZZ^T)$$

Issues in self-correlation representation:

- 1. Islands:** $\tilde{A}_{i,i} = \text{sigmoid}(z_i z_i) > 0.5$, because $z_i^2 > 0$



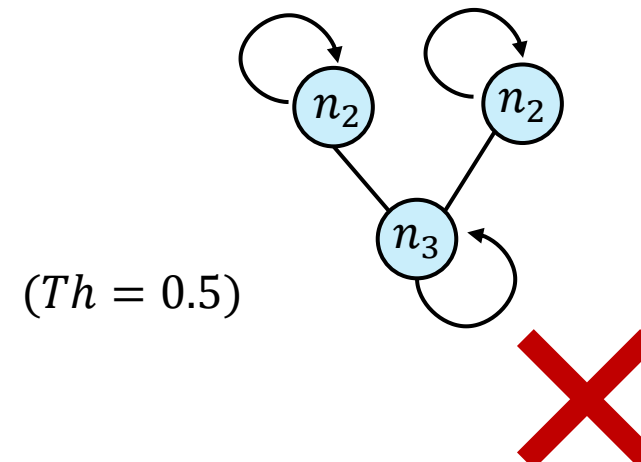
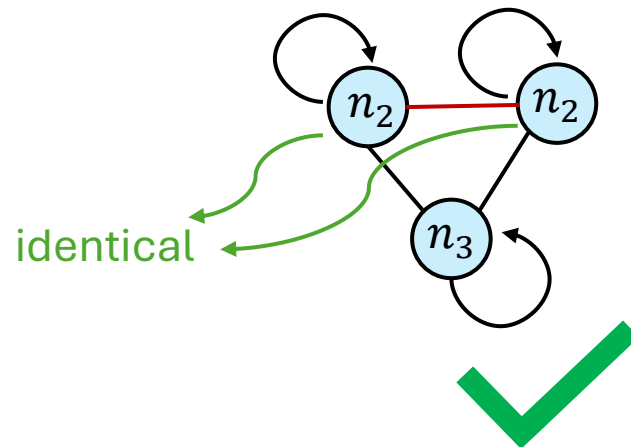
Self-correlation:

$$\text{encoder: } Z = \Phi(Z|G) = f(X, A)$$

$$\text{decoder: } \tilde{A} = \Theta(A|Z) = \text{sigmoid}(ZZ^T)$$

Issues in self-correlation representation:

2. **Symmetric structure:** $\tilde{A}_{i,j} = \text{sigmoid}(z_i z_j) > 0.5$



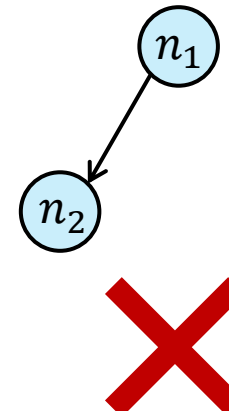
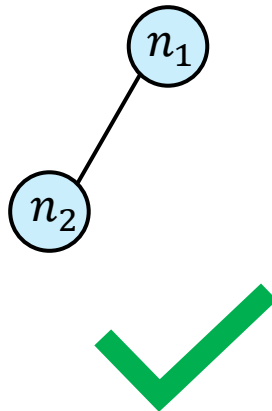
Self-correlation:

$$\text{encoder: } Z = \Phi(Z|G) = f(X, A)$$

$$\text{decoder: } \tilde{A} = \Theta(A|Z) = \text{sigmoid}(ZZ^T)$$

Issues in self-correlation representation:

3. Directed graph: $\tilde{A}_{i,j} = \tilde{A}_{j,i}$



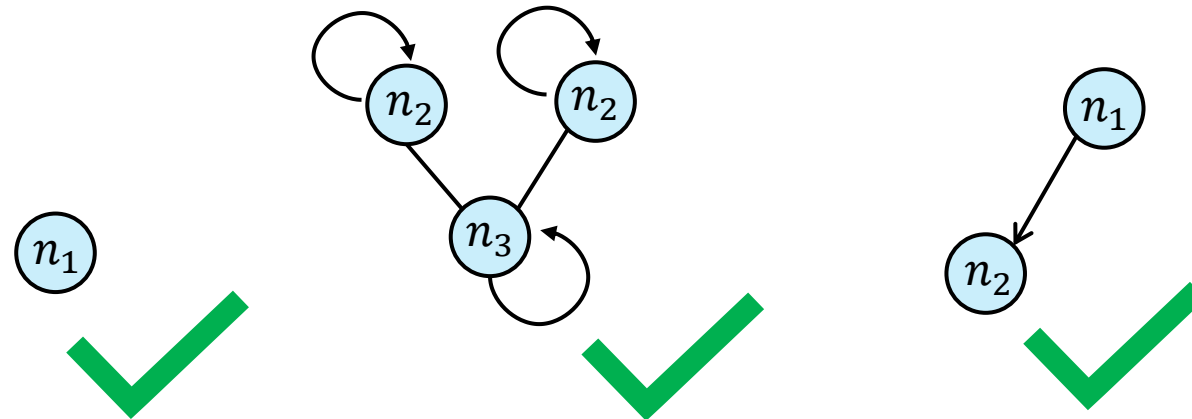
Graph Structural Reconstruction

Cross-correlation:

$$\text{encoder: } Z' = \Phi(Z'|G) = f(X, A)$$

$$\text{decoder: } \tilde{A} = \text{sigmoid}(PQ^T), \\ P = g_1(Z', \{A', h'\}), Q = g_2(Z', \{A', h'\})$$

Asymmetry between P and Q :



Reconstruction Visualization on Special Struc.

cross-correlation

self-correlation



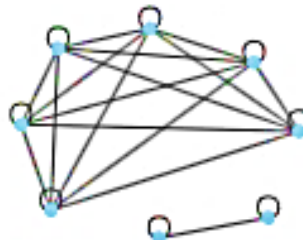
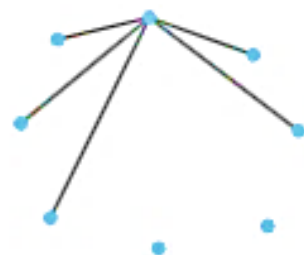
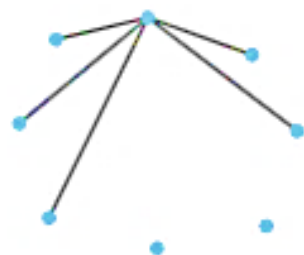
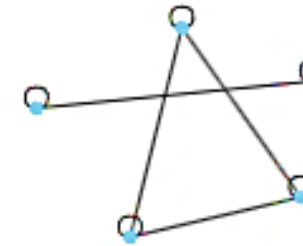
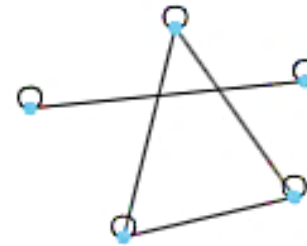
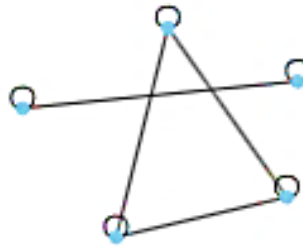
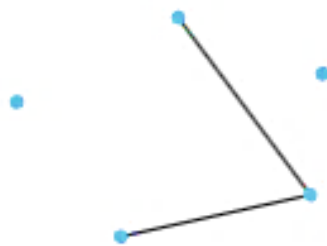
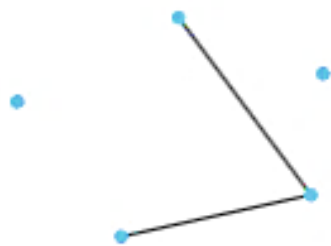
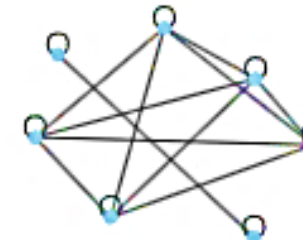
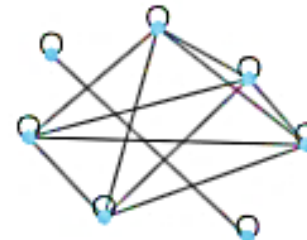
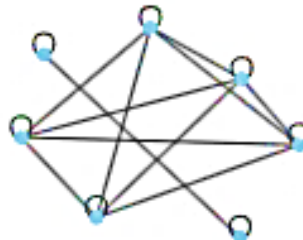
Ground Truth

GraphCroc

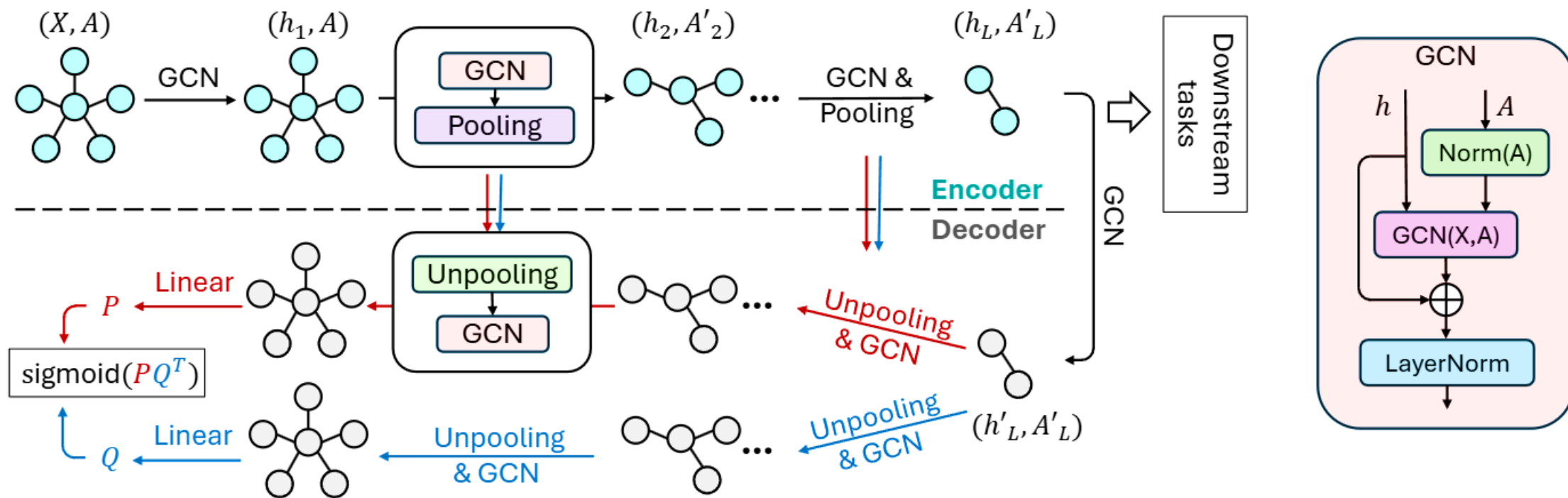
GraphCroc(SC)

GAE

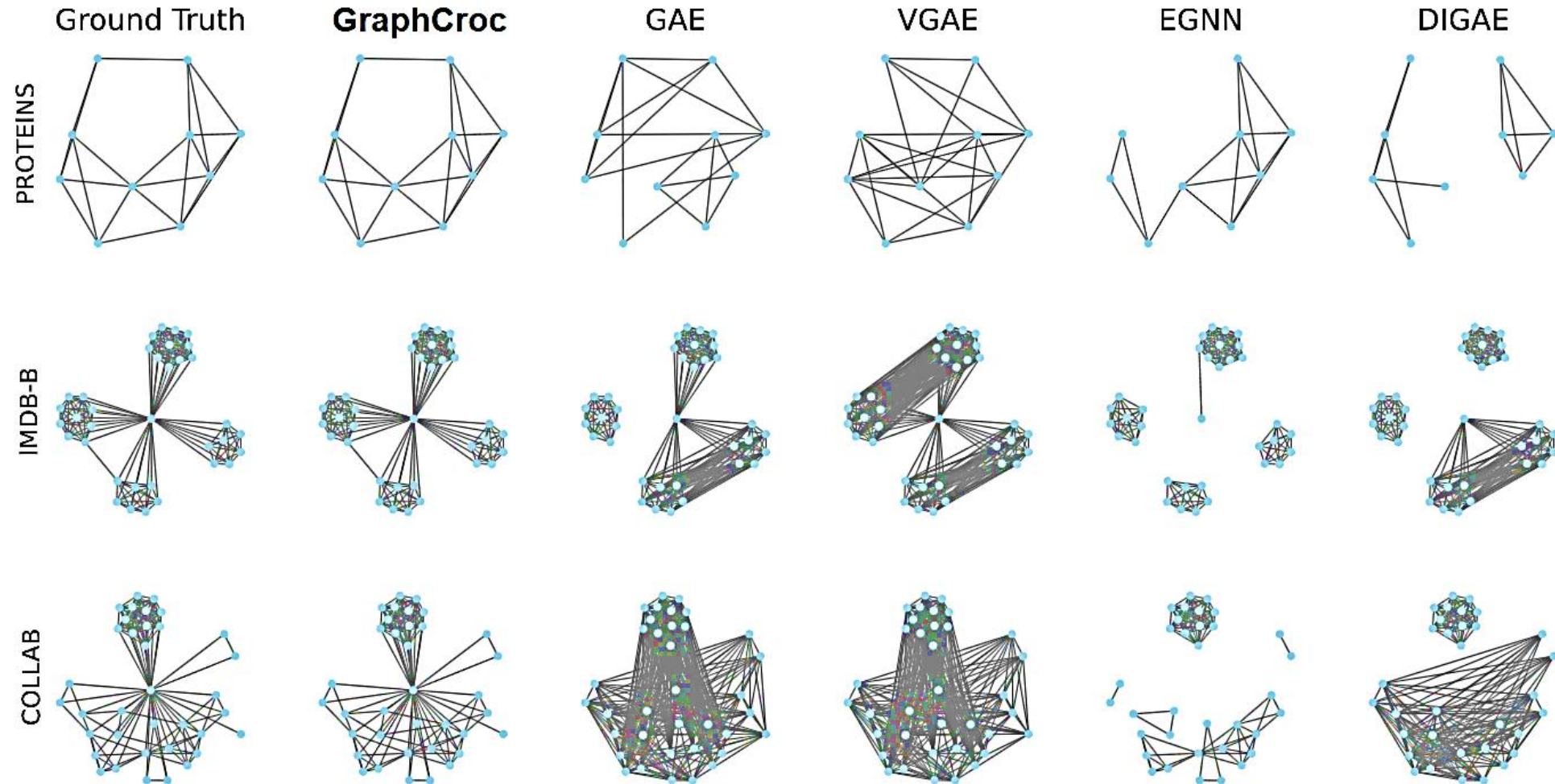
VGAE



GraphCroc Based on Cross-Correlation



More Visualization



Reconstruction Results – AUC Score

Table: AUC Score of reconstructing the adjacency matrix

| | Self-Correlation | | | | Cross-Correlation | |
|----------|------------------|--------|---------------|----------------------|-------------------|------------------|
| | GAE | VGAE | EGNN | GraphCroc(SC) | DiGAE | GraphCroc |
| PROTEINS | 0.4750 | 0.4764 | 0.9608 | <u>0.9781</u> | 0.7577 | 0.9958 |
| IMDB-B | 0.7556 | 0.7105 | 0.9873 | <u>0.9892</u> | 0.7500 | 0.9992 |
| Collab | 0.7885 | 0.7946 | <u>0.9947</u> | 0.9926 | 0.7973 | 0.9989 |
| PPI | 0.6330 | 0.6239 | –† | <u>0.9764</u> | 0.8364 | 0.9831 |
| QM9 | 0.5376 | 0.4852 | <u>0.9984</u> | <u>0.9967</u> | 0.7791 | 0.9987 |

Reconstruction Results – WL Test

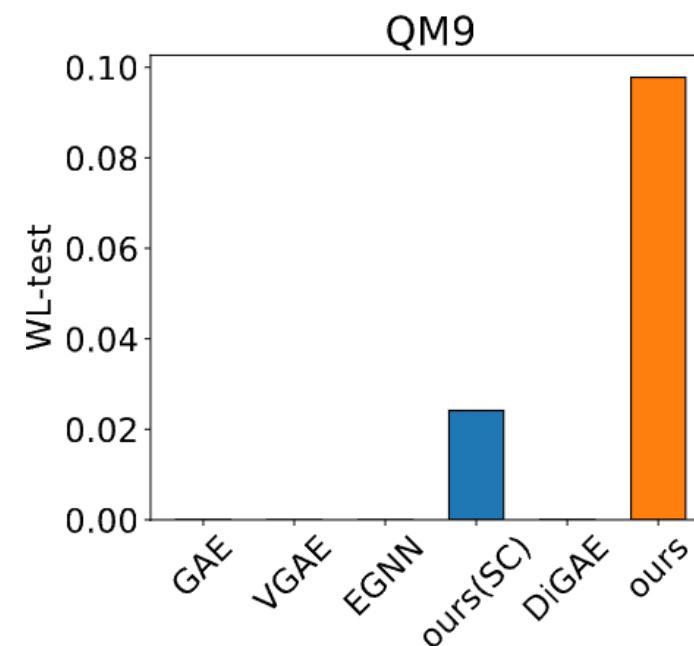
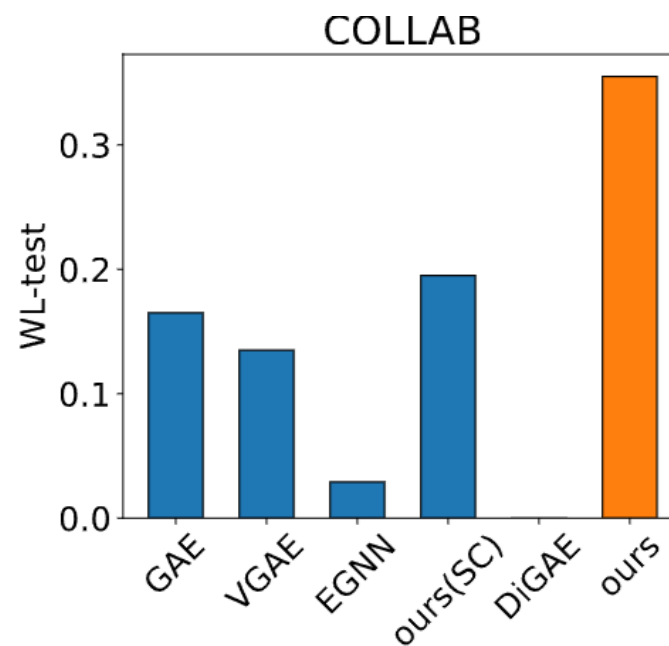
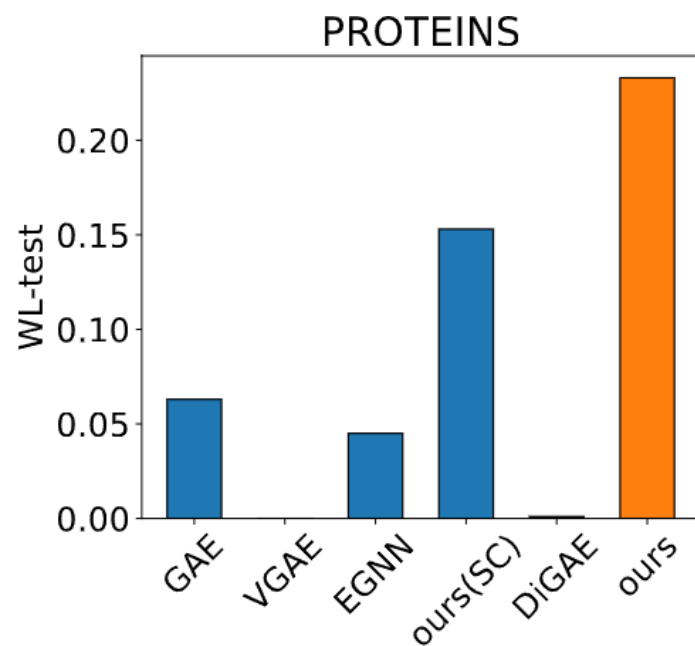


Table: Graph classification accuracy (%) on different tasks

| | Infograph | GraphCL | InfoGCL | GraphMAE | S2GAE | StructMAE | ours (10-epoch) | ours (100-epoch) |
|----------|-----------|---------|---------|----------|--------------|-----------|---------------------------|----------------------------|
| PROTEINS | 74.44 | 74.39 | – | 75.30 | <u>76.37</u> | 75.97 | 73.99 \pm 1.32 | 79.09 \pm 1.63 |
| IMDB-B | 73.03 | 71.14 | 75.10 | 75.52 | <u>75.76</u> | 75.52 | 76.69 \pm 1.02 | 78.75 \pm 1.35 |
| COLLAB | 70.65 | 71.36 | 80.00 | 80.32 | 81.02 | 80.53 | <u>81.70</u> \pm 0.54 | 82.40 \pm 0.20 |

Thank You
