

Experiment I: Arranging Nodes with Given Metric Matrices

- We pass the node features through 8 MGNN layers, with $\alpha = 0.05$ and $\beta = 0.5$. After each round of propagation, we visualize the results.

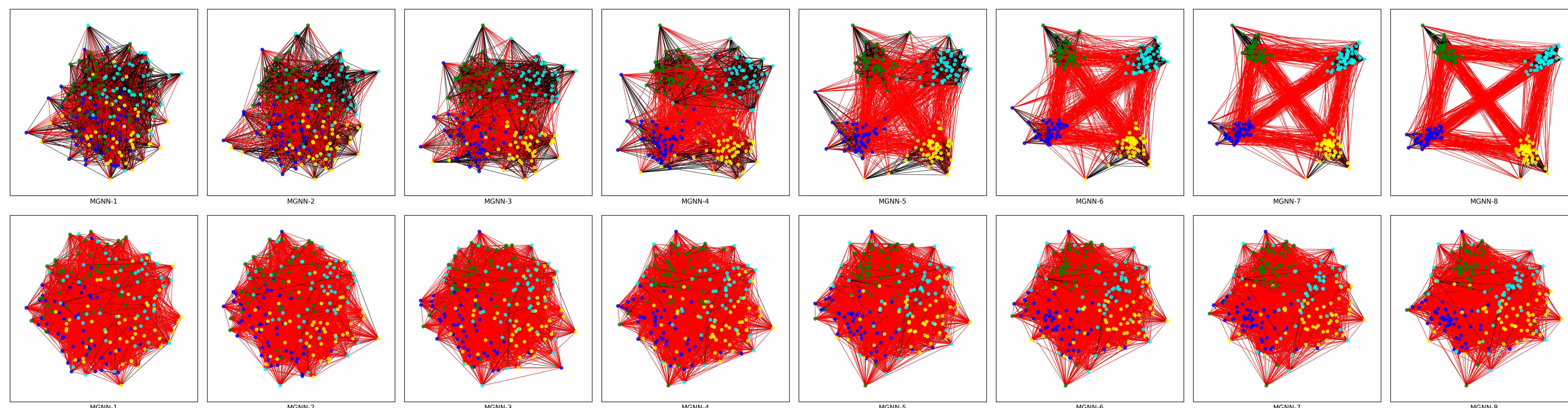


Figure 1. Visualization results of MGNN propagation layers.

- Based on the results, we can confirm that our MGNN model aims to separate the four blocks.
- The presence of homophilic edges adds convexity to our objective function, which is easier to optimize.
- Additionally, we provide visualization results obtained from the SGC layers and APPNP layers.

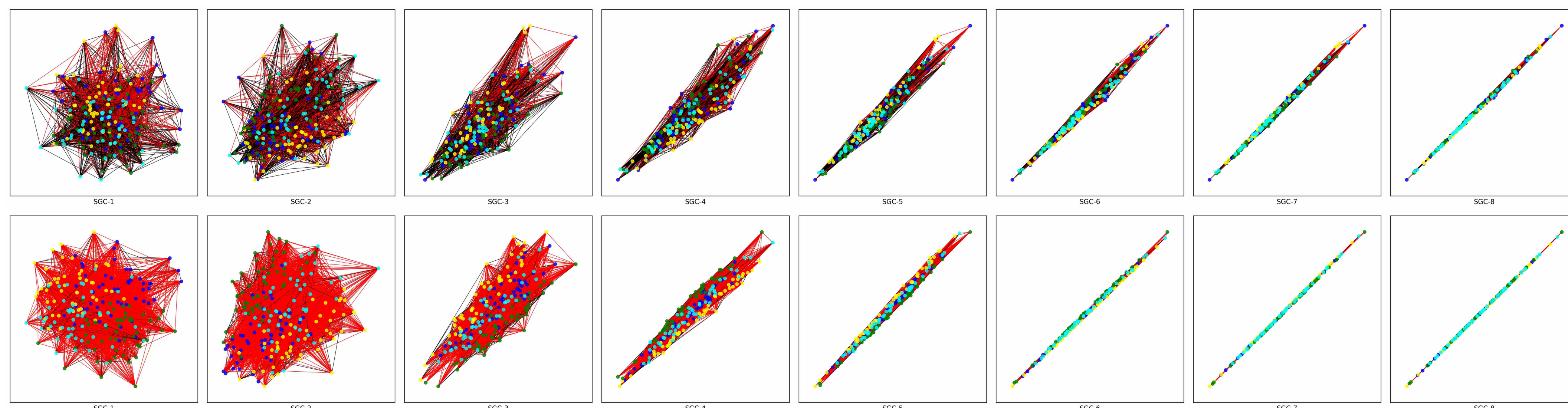


Figure 2. Visualization results of SGC propagation layers.

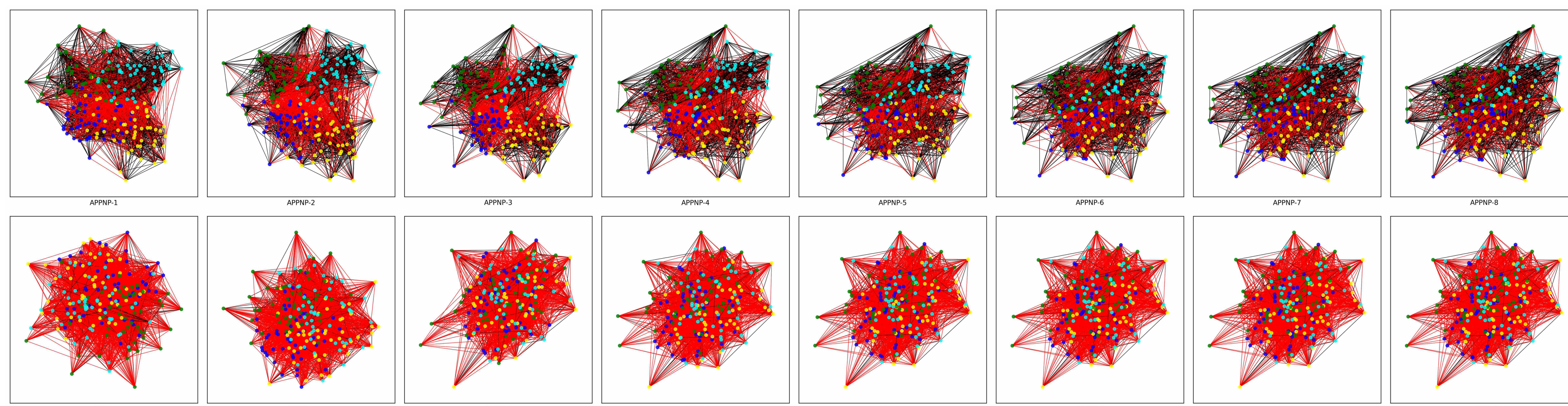


Figure 3. Visualization results of APPNP propagation layers.

Experiment II: Supervised Node Classification

- For all datasets except ogbn-arxiv, we generate 10 random splits with train/valid/test ratio of 60%/20%/20%. For the ogbn-arxiv dataset, we generate 10 random splits using the `get_idx_split()` function provided by the official package.
- We then tune the hyper-parameters 100 times for each model on each dataset. (Please refer to our paper for the detailed ranges of hyperparameters.)
- We train each model for a maximum of 1500 epochs, with early stopping set to 100, on each dataset and report the results after hyper-tuning.

Table 1. The results (accuracy with standard deviation) of the supervised node classification experiments. Boldface results indicate the best model on each dataset, and underlined results are second best models.

	Non-Graph		Spectral			Spatial (Non-Spectral)					MGNN	
	Linear	MLP	GCN	SGC	APPNP	PointNet	GAT	GIN	GCNII	pGNN		LINKX
Cora	76.09 \pm 1.55	76.11 \pm 1.58	88.25 \pm 1.09	88.34 \pm 1.41	89.30 \pm 1.45	84.43 \pm 1.94	88.71 \pm 0.89	85.76 \pm 1.18	88.52 \pm 1.40	88.78 \pm 1.20	82.82 \pm 1.96	89.04 \pm 1.20
CiteSeer	71.11 \pm 1.81	73.25 \pm 1.16	77.10 \pm 1.12	<u>77.47</u> \pm 1.42	76.80 \pm 1.10	72.83 \pm 1.38	76.44 \pm 1.19	72.83 \pm 1.47	77.05 \pm 0.78	77.28 \pm 0.88	71.79 \pm 1.55	78.08 \pm 1.50
PubMed	87.06 \pm 0.67	88.21 \pm 0.46	89.07 \pm 0.42	<u>87.59</u> \pm 0.54	89.75 \pm 0.47	89.18 \pm 0.38	87.77 \pm 0.67	87.15 \pm 0.54	<u>90.24</u> \pm 0.55	89.79 \pm 0.38	86.77 \pm 0.59	90.37 \pm 0.47
CoraFull	60.55 \pm 0.76	61.61 \pm 0.56	71.86 \pm 0.75	71.86 \pm 0.70	71.88 \pm 0.75	63.32 \pm 1.02	70.65 \pm 0.86	67.11 \pm 0.46	<u>71.09</u> \pm 0.72	<u>72.36</u> \pm 0.58	63.57 \pm 0.55	72.42 \pm 0.69
CS	94.51 \pm 0.32	94.89 \pm 0.23	93.77 \pm 0.37	94.10 \pm 0.40	95.91 \pm 0.23	93.13 \pm 0.42	93.25 \pm 0.31	91.81 \pm 0.34	<u>95.98</u> \pm 0.22	95.83 \pm 0.23	95.06 \pm 0.24	96.01 \pm 0.16
Physics	95.92 \pm 0.15	96.01 \pm 0.26	96.46 \pm 0.25	OOM	<u>97.14</u> \pm 0.21	96.37 \pm 0.27	96.47 \pm 0.21	94.66 \pm 0.24	97.10 \pm 0.21	96.93 \pm 0.10	96.87 \pm 0.16	97.15 \pm 0.15
Cornell	78.65 \pm 2.82	75.95 \pm 4.43	50.27 \pm 6.86	53.51 \pm 4.32	77.84 \pm 6.49	71.08 \pm 5.93	50.27 \pm 7.66	49.46 \pm 7.46	76.76 \pm 7.76	77.30 \pm 8.22	71.08 \pm 12.09	81.89 \pm 6.29
Texas	81.35 \pm 4.90	83.78 \pm 7.55	61.08 \pm 8.65	56.22 \pm 6.37	86.76 \pm 1.46	82.16 \pm 6.30	61.08 \pm 5.57	64.05 \pm 4.20	<u>88.65</u> \pm 4.95	85.14 \pm 5.16	84.32 \pm 5.24	90.00 \pm 2.72
Wisconsin	86.20 \pm 3.63	88.80 \pm 2.40	55.80 \pm 5.83	58.00 \pm 4.29	86.00 \pm 3.10	81.60 \pm 3.98	56.40 \pm 5.99	57.20 \pm 6.82	88.20 \pm 4.14	85.40 \pm 3.58	82.00 \pm 3.10	88.40 \pm 3.44
Chameleon	50.77 \pm 2.07	50.75 \pm 2.13	<u>69.21</u> \pm 2.08	67.12 \pm 2.17	67.85 \pm 2.65	63.76 \pm 2.52	66.97 \pm 2.45	46.73 \pm 13.51	67.56 \pm 1.18	69.19 \pm 1.57	67.47 \pm 1.62	72.37 \pm 2.25
Squirrel	35.76 \pm 1.05	35.79 \pm 1.65	55.43 \pm 2.05	52.18 \pm 1.49	54.60 \pm 1.88	47.39 \pm 8.31	<u>55.68</u> \pm 2.81	20.96 \pm 2.08	53.88 \pm 2.77	51.61 \pm 1.28	57.86 \pm 1.17	54.45 \pm 1.85
Actor	36.16 \pm 0.75	37.67 \pm 1.60	30.42 \pm 1.58	30.14 \pm 1.18	36.98 \pm 1.28	36.41 \pm 1.23	29.22 \pm 0.94	26.09 \pm 1.75	<u>37.75</u> \pm 1.23	36.47 \pm 1.07	35.00 \pm 2.11	38.36 \pm 1.46
WikiCS	78.74 \pm 0.57	79.85 \pm 0.69	84.02 \pm 0.61	83.47 \pm 0.83	84.88 \pm 0.55	84.09 \pm 0.86	83.82 \pm 0.73	66.08 \pm 22.77	<u>85.09</u> \pm 0.71	84.41 \pm 0.46	84.13 \pm 0.56	85.09 \pm 0.59
ogbn-arxiv	52.80 \pm 0.19	53.81 \pm 0.23	70.48 \pm 0.18	68.77 \pm 0.06	70.50 \pm 0.11	69.89 \pm 0.59	<u>71.04</u> \pm 0.27	66.60 \pm 0.53	71.48 \pm 0.21	68.50 \pm 0.17	59.28 \pm 2.22	70.73 \pm 0.17
Average Rank	9.21	7.86	6.79	7.69	3.50	7.86	7.57	10.50	3.50	4.36	7.29	1.57

Experiment III: Graph Regression

- We utilized the publicly available train/validation/test split of the ZINC-subset dataset.
- For all models, we set the number of hidden units to 64. We limit the training of each model to a maximum of 500 epochs. A batch size of 512 is utilized, and early stopping is implemented after 50 epochs on each dataset.
- We conduct hyperparameter tuning for each model on each dataset and repeat this process 100 times. (Please refer to our paper for the detailed ranges of hyperparameters.)
- To enhance reliability, we repeat the whole process five times and calculate the mean MAE (Mean Absolute Error) and standard deviation for each model.

Table 2. The results of the graph regression experiments.

GCN	GAT	GIN	PointNet	MGNN
0.6143 \pm 0.0200	0.6458 \pm 0.0647	0.4410 \pm 0.0065	0.5293 \pm 0.0138	<u>0.4751</u> \pm 0.0112