

Rebuttal

Table 1: Node Classification Results with Conformal Baselines (Coverage \uparrow / Inefficiency \downarrow)

Dataset	CF-GNN [1]		DAPS [2]		RR-GNN (Ours)		Cluster-RR-GNN (Ours)	
Model	Cover	Ineff	Cover	Ineff	Cover	Ineff	Cover	Ineff
Cora								
GraphSAGE	0.9456 \pm 0.0569	1.6284 \pm 0.0483	0.9453 \pm 0.0535	1.8025 \pm 0.0421	0.9460 \pm 0.0542	1.6100 \pm 0.0415	0.9463 \pm 0.0509	1.6076 \pm 0.0397
SGC	0.9461 \pm 0.0603	1.6633 \pm 0.04415	0.9452 \pm 0.0538	1.7856 \pm 0.0426	0.9462 \pm 0.0581	1.6297 \pm 0.0428	0.9468 \pm 0.0662	1.6017 \pm 0.0465
GCN	0.9473 \pm 0.0556	1.6344 \pm 0.0418	0.9435 \pm 0.053	1.7120 \pm 0.0354	0.9432 \pm 0.0573	1.6251 \pm 0.0367	0.9476 \pm 0.0732	1.6315 \pm 0.0303
GAT	0.9464 \pm 0.0702	1.6278 \pm 0.0334	0.9480 \pm 0.065	1.7052 \pm 0.0384	0.9475 \pm 0.0624	1.6146 \pm 0.0351	0.9491 \pm 0.0539	1.6254 \pm 0.0396
DBLP								
GraphSAGE	0.9501 \pm 0.0523	1.5723 \pm 0.0683	0.9500 \pm 0.0420	1.6436 \pm 0.0627	0.9499 \pm 0.0531	1.5351 \pm 0.0473	0.9503 \pm 0.0510	1.5607 \pm 0.0487
SGC	0.9451 \pm 0.0617	1.5274 \pm 0.0416	0.9427 \pm 0.0526	1.6020 \pm 0.0317	0.9462 \pm 0.0528	1.4286 \pm 0.0541	0.9443 \pm 0.0462	1.3921 \pm 0.0624
GCN	0.9473 \pm 0.0596	1.5644 \pm 0.0733	0.9458 \pm 0.0565	1.6384 \pm 0.0703	0.9458 \pm 0.0702	1.5512 \pm 0.0295	0.9430 \pm 0.0713	1.5491 \pm 0.0278
GAT	0.9467 \pm 0.0717	1.5729 \pm 0.0463	0.9455 \pm 0.0685	1.6493 \pm 0.0455	0.9485 \pm 0.0589	1.5725 \pm 0.0349	0.9491 \pm 0.0539	1.5720 \pm 0.0322
CiteSeer								
GraphSAGE	0.9528 \pm 0.0203	1.1680 \pm 0.0439	0.9501 \pm 0.0195	1.3425 \pm 0.0412	0.9538 \pm 0.0853	1.1621 \pm 0.0552	0.9540 \pm 0.0926	1.1679 \pm 0.0605
SGC	0.9525 \pm 0.0257	1.1827 \pm 0.0552	0.9513 \pm 0.0245	1.3578 \pm 0.0525	0.9579 \pm 0.0536	1.1782 \pm 0.0415	0.9594 \pm 0.0582	1.1898 \pm 0.0399
GCN	0.9496 \pm 0.0392	1.2310 \pm 0.0332	0.9520 \pm 0.036	1.4026 \pm 0.0327	0.9512 \pm 0.0358	1.2189 \pm 0.0276	0.9518 \pm 0.0373	1.2153 \pm 0.0290
GAT	0.9508 \pm 0.0309	1.2396 \pm 0.0416	0.9513 \pm 0.0291	1.4152 \pm 0.0393	0.9535 \pm 0.0447	1.2085 \pm 0.0361	0.9548 \pm 0.0491	1.2020 \pm 0.0392

Coverage (\uparrow): Empirical coverage rate (target: $1 - \alpha = 0.95$)

Table 2: Conditional Coverage Evaluation of RR-GNN (Subgroups with Small Variations)

Condition Type	Subgroup	$\alpha = 0.1$	$\alpha = 0.2$	$\alpha = 0.3$
3*Cluster-Conditional	Cluster 1	0.9023	0.8041	0.7124
	Cluster 2	0.9015	0.8034	0.7085
	Cluster 3	0.8967	0.7945	0.6845
4*Class-Conditional	Class A	0.9018	0.8075	0.7064
	Class B	0.8986	0.7914	0.6947
	Class C	0.9047	0.8035	0.7016
	Class D	0.8935	0.7942	0.6964

Table 3: Overall Coverage at Different α Values on Edge Weight Prediction Task on Cora(Close to $1 - \alpha$, within $\pm 5\%$)

α Value	Ineff
0.05	1.6315
0.10	1.5576
0.15	1.5528
0.20	1.2468
0.25	1.1178
0.30	1.0373

Table 4: Coverage Across Different Set Sizes (Close to $1 - \alpha$, within $\pm 1\%$)

Set Size	Node Count	Coverage ($\alpha = 0.1$)
Small (2DatasetTwitter)	81,306	0.8957
Medium (5: Dataset: CiteSeer)	3,327	0.9036
Large (≥ 10 : Dataset: OGBN-Arxiv)	169,343	0.9014

Table 5: Performance(AUC) Comparison of Graph Transformer Models with RR Enhancement on dataset Mol-HIV

Model	Method	Cora	CiteSeer	PubMed	OGB-Arxiv
2*Graphormer	Original	0.763 ± 0.012	0.691 ± 0.015	0.792 ± 0.008	0.718 ± 0.005
	+ RR	0.781 ± 0.011	0.705 ± 0.013	0.803 ± 0.007	0.729 ± 0.004
2*Graphormer with Spatial Encoding	Original	0.772 ± 0.010	0.702 ± 0.014	0.801 ± 0.007	0.725 ± 0.005
	+ RR	0.789 ± 0.009	0.715 ± 0.012	0.812 ± 0.006	0.736 ± 0.004
2*Graphormer with Graph Structure	Original	0.781 ± 0.011	0.712 ± 0.013	0.808 ± 0.007	0.732 ± 0.004
	+ RR	0.796 ± 0.010	0.724 ± 0.011	0.819 ± 0.006	0.742 ± 0.003

Table 6: Performance of Graph Transformer Networks (GT) with RR Enhancement

Node Classification	F1 Score	ACM	DBLP	IMDB
2*Base GT	Original	0.912 ± 0.014	0.938 ± 0.034	0.609 ± 0.023
	+ RR	0.923 ± 0.012	0.942 ± 0.025	0.724 ± 0.036