

DRUG						
Model	COD↓	CAD↓	CHD↓	CLD↓	IND↑	KL↓
GCN <sub>KL</sub>	0.220±.025	9.209±.740	0.245±.031	1.963±.134	0.676±.029	0.484±.075
HAN <sub>KL</sub>	0.279±.023	10.084±.823	0.268±.027	2.155±.148	0.632±.025	0.579±.055
SeHGNN <sub>KL</sub>	0.286±.018	10.178±.781	0.267±.020	2.166±.143	0.640±.016	0.600±.059
GLDL	0.199±.064	<b>8.887</b> ±.838	0.232±.044	<b>1.89</b> ±.17	0.699±.049	0.409±.105
HINormer	0.365±.028	10.563±.571	0.308±.029	2.227±.099	0.575±.023	0.742±.102
HGDL	<b>0.168</b> ±.019	9.179±.574	<b>0.217</b> ±.017	1.957±.114	<b>0.710</b> ±.020	<b>0.392</b> ±.044
HGDL <sub>-transformer</sub>	0.199±.014	9.371±.679	0.235±.017	2.004±.137	0.687±.021	0.492±.059
HGDL <sub>-t</sub>	0.212±.023	9.510±.602	0.240±.018	2.029±.110	0.671±.020	0.462±.050
HGDL <sub>ed</sub>	0.204±.026	9.602±.882	0.239±.028	2.040±.162	0.681±.030	0.574±.085

DBLP						
Model	COD↓	CAD↓	CHD↓	CLD↓	IND↑	KL↓
GCN <sub>KL</sub>	0.031±.004	2.852±.009	0.091±.006	1.647±.002	0.908±.006	0.114±.011
HAN <sub>KL</sub>	0.025±.002	2.819±.012	0.071±.004	1.633±.007	0.929±.004	0.082±.008
SeHGNN <sub>KL</sub>	0.086±.140	2.887±.170	0.155±.208	<b>1.624</b> ±.049	0.842±.214	0.252±.397
GLDL	0.019±.001	2.8±.009	<b>0.056</b> ±.002	1.633±.004	0.943±.003	0.06±.003
HINormer	0.053±.008	2.936±.013	0.146±.019	1.669±.008	0.853±.019	0.22±.025
HGDL	<b>0.019</b> ±.002	<b>2.796</b> ±.014	0.057±.005	1.633±.005	<b>0.943</b> ±.005	<b>0.057</b> ±.011
HGDL <sub>-transformer</sub>	0.025±.002	2.828±.004	0.074±.005	1.642±.001	0.925±.005	0.090±.008
HGDL <sub>-t</sub>	0.020±.002	2.808±.013	0.062±.005	1.637±.005	0.937±.005	0.070±.011
HGDL <sub>ed</sub>	0.023±.001	2.819±.005	0.070±.003	1.639±.003	0.929±.003	0.082±.006

Table 1: Experimental results of HINormer and GLDL on the Drug and DBLP datasets, where HINormer is a recent heterogenous graph learning baseline and GLDL is for LDL in homogenous graphs.

100 epochs (s)	DRUG (894)	URBAN (604)	ACM (5810)	YELP (3001)	DBLP (4057)
GCN <sub>KL</sub>	2.51	2.77	10.97	<b>1.8569</b>	<b>23</b>
HAN <sub>KL</sub>	1.73	0.979	75	12.53	23.63
SeHGNN <sub>KL</sub>	2.0592	1.16	75.25	12.45	23.45
HGDL <sub>KL</sub>	<b>1.19</b>	<b>1.05</b>	<b>8.81</b>	8.75	26.57

Table 2: Runtime results for 100 training epochs. The numbers of target nodes are labeled beside their corresponding dataset names.

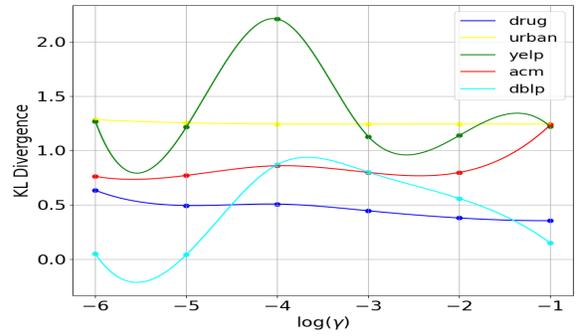


Figure 1: Sensitive analysis of  $\gamma$ , ranging from 0, 1e-5, 1e-4, 1e-3, 1e-2, and 0.1, converted to the values of  $\log(\gamma)$  as -6, -5, -4, -3, -2, and -1 along the X-axis.

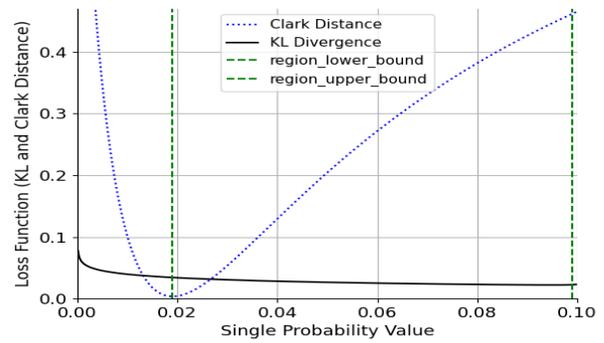


Figure 2: KL and CLD tradeoff function example. The estimated probability distribution is  $[x_1, x_2, 0.9]$  and true probability distribution is  $[0.05, 0.05, 0.9]$ , with  $x_1 + x_2 = 0.1$ . Horizontal axis is the  $x_1$  value and vertical axis is the loss for both CLD and KL divergence. Green dashed lines cover the tradeoff region where the CLD loss monotonically increases and KL-divergence decreases.

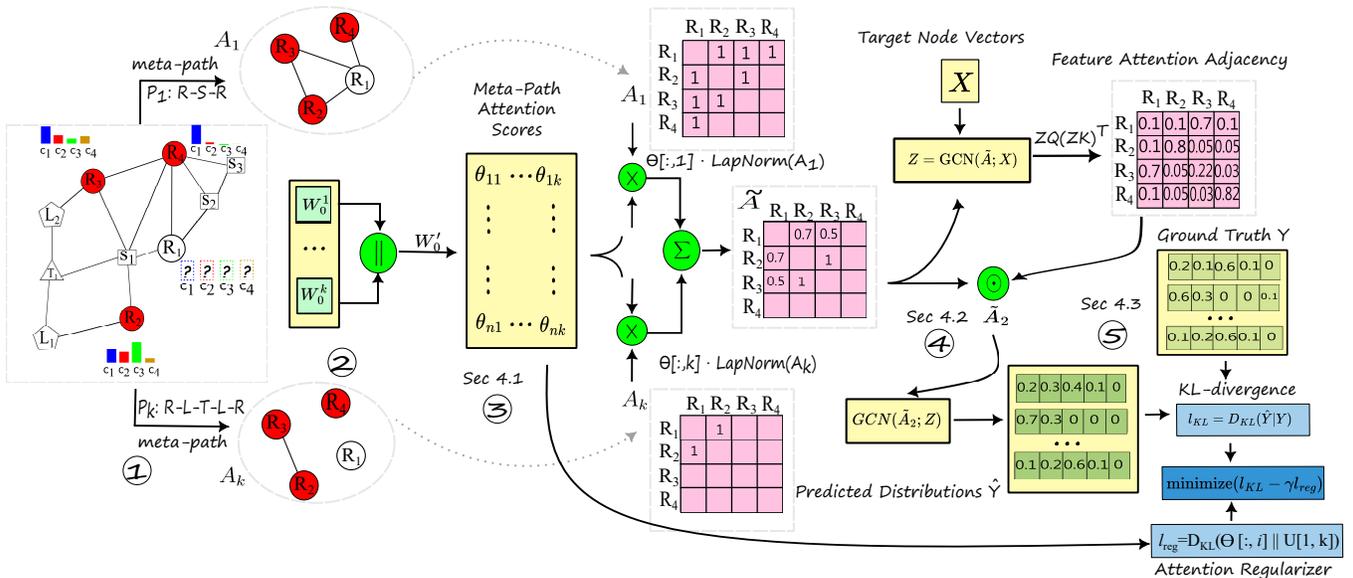


Figure 3: The revised illustration of HGDL framework, which shall replace Figure 2 in the original manuscript.