I. The Supplementary Material of Light Single-view Graph Contrastive Learning

A. Datasets for node classification

In this article, we use 9 datasets for node classification, i.e., three homophilic datasets: Computers, Pubmed, and Photo, six heterophilic datasets: Chameleon, Squirrel, Film, Cornell, Texas as well as large-scale dataset obgn-arxiv. Pubmed is a node classification dataset with one citation network for each paper, and nodes have bag-of-words document representations. Amazon co-purchase graphs include Photo and Computers. Nodes represent goods, edges indicate whether two things are usually purchased together, and node features are a bag-of-words product review presentation. Chameleon and Squirrel are two Wikipedia page networks that have entries with mutual ties. Article nouns are node features. Cornell and Texas are computer science department networks. Web pages are nodes, edges are hyperlinks, and node characteristics are TF-IDF vectors of their contents.

B. Datasets for graph classification

Bioinformatics datasets. MUTAG is a nitro compounds dataset including 188 samples and is divided into 2 classes. PTC-MR consists of compounds labeled according to carcinogenicity on rodents with 19 node labels.

Social network datasets. IMDB-BINARY (IMDB-B for short) and IMDB-MULTI (IMDB-M for short) are movie collaboration datasets, where every graph belongs to a kind of movie genre which is also to be predicted. The IMDB-BINARY dataset contains two types: Action and Romance. While the IMDB-MULTI dataset is constructed from Comedy, Romance, and Sci-Fi genres. Each node represents an actor/actress. If two actors appear in the same movie, then they are connected by an edge.

We compare our model with the baselines with using the same learning rate and weight decay, and other important parameters are listed in the following two tables. specifically, we list the url of code projects and important super-parameters in the following Table. 1 and Table. 2. With referring to the raw project, we also set the important parameters of the benchmark methods for fair comparison.

TABLE I
DETAILED URL ABOUT THE EXPERIMENTS FOR CLASSIFICATION TASK.

model	url
DeepWalk/Node2vec	https://github.com/shenweichen/GraphEmbedding
COLES	https://github.com/allenhaozhu/COLES
DGI	https://github.com/PetarV-/DGI
MVGRL	https://github.com/kavehhassani/mvgrl
GRACE	https://github.com/CRIPAC-DIG/GRACE
LightGCL	https://github.com/HKUDS/LightGCL/tree/main
Sub2vec	https://github.com/bijayaVT/sub2vec/tree/master
Graph2vec	https://github.com/benedekrozemberczki/graph2vec
InfoGraph	https://github.com/sunfanyunn/InfoGraph
MVGRL	https://github.com/kavehhassani/mvgrl
GCN/GIN/GAT/GrapSage	https://github.com/weihua916/powerful-gnns

TABLE II
IMPORTANT PARAMETERS FOR BASELINES.

Model	important parameters
DeepWalk	walk_length=10, num_walks=80, window_size=5
Node2vec	walk_length=10, num_walks=80, p=0.25, q=4,use_rejection_sampling=0
COLES	hidden=256, $L = 8, \lambda = 0.8$, $sample\_size = 2$
DGI	hidden=256, $L = 4$, nonlinearity = 'prelu'
MVGRL	hidden=256, $L = 4$, nonlinearity = 'prelu'
GRACE	hidden=256, $L = 4$, nonlinearity = 'relu'
LightGCL	hidden=256, $L = 4$, 150 is used for SVD
Sub2vec	dimension=256
Graph2vec	dimension=256
InfoGraph	hidden=256, batch=64
MVGRL	hidden=256, batch=64