Universal Graph Contrastive Learning with a Novel Laplacian Perturbation

Taewook Ko[1] **Yoonhyuk Choi**[1] **Chong-Kwon Kim**[2]

[1]Department of Computer Science and Engineering, Seoul National University
[2]Research Institute of Energy AI, Korea Institute of Energy Technology

Abstract

Graph Contrastive Learning (GCL) is an effective method for discovering meaningful patterns in graph data. By evaluating diverse augmentations of the graph, GCL learns discriminative representations and provides a flexible and scalable mechanism for various graph mining tasks. This paper proposes a novel contrastive learning framework by introducing Laplacian perturbation. The proposed framework offers a distinct advantage by employing an indirect perturbation method, which provides a more stable approach while maintaining the perturbation effects. Moreover, it exhibits a wide range of applicability by not being restricted to specific graph types. We demonstrate that a spectral graph convolution based on the Laplacian successfully extracts representations from diverse graph types. Our extensive experiments on a variety of real-world datasets, covering multiple graph types, show that the proposed model outperforms state-of-the-art baselines in both node classification and link sign prediction tasks.

1 INTRODUCTION

Contrastive learning, which maximizes the mutual information between representations of augmented data, is a widely used method that helps learn discriminative representations from limited labeled data or self-supervised environments [Chen et al., 2020a, He et al., 2020]. Recently, a growing number of studies have successfully demonstrated the effectiveness of contrastive learning in graph data. These studies, known as Graph Contrastive Learning (GCL), make diverse graph views through several graph perturbation methods such as deleting or adding edges/nodes [You et al., 2020, Zeng and Xie, 2021], random-walk sampling [Qiu et al., 2020], or masking attributes [Zhu et al., 2021]. On the other hand, JOAO [You et al., 2021] and AD-GCL [Suresh et al., 2021] proposed novel graph encoders for contrastive learning. SimGRACE [Xia et al., 2022] added Gaussian noise to graph encoders.

Majority of prior GCL techniques are designed for specific graph types. Many are confined to unsigned undirected graphs, severely limiting their application domain. Several algorithms that expand the applicability of GCL have been proposed [Tong et al., 2021, Shu et al., 2021]. However, these expanded schemes focused on a single graph type. For example, DiGCL [Tong et al., 2021] and SGCL [Shu et al., 2021] are customized for directed and signed directed graphs, respectively. Although their approaches are innovative, DiGCL is adequate for dense directed graphs only and SGCL may fail to capture valuable network semantics emanated from the balance theory [Heider, 1946, Holland and Leinhardt, 1971].

This paper proposes a novel comprehensive Graph Contrastive Learning algorithm called **UGCL**[1] (Universal GCL). We devise a new data augmentation scheme based on magnetic Laplacian perturbation. UGCL claims to have wide applicability ranging from simple unsigned undirected graphs to signed directed graphs. The magnetic Laplacian [Shubin, 1994, Olgiati, 2017, Colin de Verdière, 2013], initially studied in quantum physics, recently has been applied in graph studies due to its Hermitian properties. Cucuringu et al. [2020], Cloninger [2017] used the magnetic Laplacian for clustering in directed graphs. Zhang et al. [2021] introduced directed graph convolution based on a magnetic Laplacian. Recent studies [Ko et al., 2023, Fiorini et al., 2022, He et al., 2022] expanded the idea of magnetic Laplacian to signed graphs by defining signed magnetic Laplacian.

The signed magnetic Laplacian captures the structure of the graph by encoding both edge signs and directions with complex values and phases. By adjusting the phase value as a parameter, the Laplacian can be modified, leading us

[1]The source code is available at `https://github.com/twko05/UGCL.git`.

Accepted for the 39th Conference on Uncertainty in Artificial Intelligence (UAI 2023).

to introduce a Laplacian perturbation method. To the best of our knowledge, this is the first attempt to apply perturbation to the magnetic Laplacian matrix. Unlike the structural perturbation methods that induce abrupt changes to graph structures, phase perturbation enjoys the flexibility of fine adjustments while maintaining the overall graph structure.

To perform graph convolutions on these augments, we introduce a magnetic Laplacian-based spectral convolution layer, inspired by the principles of graph signal processing [Defferrard et al., 2016, Hammond et al., 2011]. Finally, we maximize the mutual information between the augmented representations by reducing the node-level contrastive loss [Zhu et al., 2021, Shu et al., 2021]. We conducted experiments on real-world graph data of various types to demonstrate the effectiveness of our method. The proposed framework outperformed state-of-the-art baseline in node classification and link sign prediction tasks.

The main contributions of this paper are as follows.

- This paper introduces a novel Universal Graph Contrastive Learning (**UGCL**).
- To the best of our knowledge, this is the first attempt to introduce magnetic Laplacian perturbation.
- The magnetic Laplacian perturbation enables fine adjustments while keeping the graph topology.
- The proposed framework has wide applicability and can be applied to various graph types.
- UGCL demonstrates enhanced graph representation learning on diverse real-world graphs.

2 RELATED WORK

2.1 GRAPH CONTRASTIVE LEARNING

SimCLR [Chen et al., 2020a] and MoCo [He et al., 2020] introduced contrastive learning methods for image classification, which have achieved significant success. These works have stimulated numerous subsequent studies aiming to enhance learning efficiency through self-supervised approaches. DGI [Velickovic et al., 2019] and GMI [Peng et al., 2020] are two pioneering attempts that brought the contrastive learning mechanism to graph analysis by measuring the mutual information between graphs and node representations. InfoGraph [Sun et al., 2019] introduced patch-level representation, while GCC [Qiu et al., 2020] utilized random-walk sampling to create positive and negative samples.

These early studies proved the feasibility of constrastive learning in graph representation learning and ignited immense research investigations leading to a plethora of clever graph augmentation techniques [You et al., 2020, Zeng and Xie, 2021, Zhu et al., 2021, You et al., 2021] and graph encoding schemes [Shu et al., 2021, Xia et al., 2022, Suresh et al., 2021, Xia et al., 2022]. These advancements have led to various application studies including graph clustering [Pan and Kang, 2021, Zhong et al., 2021], node embedding [Zhu et al., 2021], DDI (Drug Drug Interaction) prediction [Wang et al., 2021, Li et al., 2022], and recommendation [Lin et al., 2022].

However, most existing GCL techniques are limited their use to unsigned undirected graphs. To broaden the scope of the GCL technique, DiGCL [Tong et al., 2021] introduced a contrastive learning for directed graphs through perturbing the teleport probability of a transition matrix. SGCL [Shu et al., 2021] proposed a signed directed contrastive learning by randomly altering the signs and directions of edges. However, both approaches have limitations and are restricted in their applicability to all graph types. In light of this, this study introduces a novel generalized graph contrastive learning framework with Laplacian matrix perturbation.

2.2 HERMITIAN ADJACENCY MATRIX AND MAGNETIC LAPLACIAN

The traditional adjacency matrix, which utilizes 0s and 1s to encode graph connectivity, possesses the advantageous property of symmetry, allowing for efficient spectral analyses of graphs. However, when it comes to dealing with signed or directed graphs, binary encoding alone proves insufficient for effective representation. Several studies have proposed to use Hermitian adjacency matrices as an alternative to the traditional adjacency matrix for directed graphs. Liu and Li [2015], Guo and Mohar [2017] encode graph edges using a Hermitian matrix form, which is equal to its conjugate transpose. Bidirectional edges are encoded as 1, while two types of unidirectional edges are represented by imaginary numbers (i and -i). The plus and minus signs to the imaginary number distinguish the edge directions. Mohar [2020], Cucuringu et al. [2020] used complex numbers with various phases to improve the interpretability of directed adjacency matrices.

With the Hermitian matrix, we can define a magnetic Laplacian. The magnetic Laplacian, originally used in quantum mechanics for the analyses of charged particles under magnetic flux [Shubin, 1994, Olgiati, 2017, Lieb and Loss, 1993, Colin de Verdière, 2013, Fanuel et al., 2018], has recently emerged as a flexible and powerful tool for directed graph analyses. Many algorithms based on the Hermitian matrix were introduced including graph clustering [F. de Resende and F. Costa, 2020, Cloninger, 2017], community detection [Fanuel et al., 2017], and graph representation learning [Furutani et al., 2019]. MagNet [Zhang et al., 2021] proved the PSD (Positive Semi-Definite) property of the proposed directed magnetic Laplacian and introduced a spectral graph convolution for directed graphs. Recently, several studies [Ko et al., 2023, Fiorini et al., 2022, He et al., 2022, Singh

and Chen, 2022] proposed signed magnetic Laplacians to apply the idea of magnetic Laplacian to signed graphs.

3 PROBLEM FORMULATION

Let $\mathcal{G} = (V, \mathcal{E})$ be a graph where V is a set of nodes and $\mathcal{E}$ is a set of directed edges. As a weighted graph, we use a sign matrix $\mathbf{S}$ to denote the sign of edges. The value of $\mathbf{S}(u, v)$ is set to 1 if there exists a positive directed edge from node u to node v, or -1 if there is a negative directed edge. Each node has one of three possible relationships with other nodes: none, positive, negative. This results in nine possible relationships for each node pair, as shown in Figure 1(b). The goal of this paper is to discover the latent features of nodes as a low-dimensional embedding vector $z_u \in \mathbb{R}^d$ as:

$$f(\mathcal{G}) = \mathbf{Z}, \tag{1}$$

where $\mathbf{Z} \in \mathbb{R}^{|V| \times d}$ is a node embedding matrix. Note that we describe the proposed algorithm assuming signed directed graphs, the most generic form of graphs. However, the proposed scheme can be applied to any types of graphs with a straightforward modification.

4 MAGNETIC LAPLACIAN

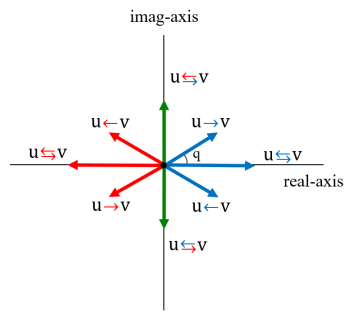
A graph Laplacian ($\mathbf{L} = \mathbf{D} - \mathbf{A}$), where $\mathbf{D}$ is a diagonal degree matrix, is a powerful tool to represent graph structure. They not only have the positive semi-definite (PSD) property but also have non-negative eigenvalues with associated orthonormal eigenvectors when the graphs are unsigned and undirected. GCN [Kipf and Welling, 2016] and ChebyNet [Defferrard et al., 2016] proposed spectral graph convolution techniques based on those properties. However, in the case of signed or directed graphs, the presence of complex eigenvalues of the graph Laplacian makes it challenging to satisfy the conditions for spectral convolution in the graph Fourier transform.

Recent studies [Ko et al., 2023, Fiorini et al., 2022] employed novel signed directed magnetic Laplacian matrices in representing the structure of signed directed graphs. The magnetic Laplacian satisfies the PSD property and is amenable to spectral graph analyses. First of all, we define a complex Hermitian adjacency matrix.

$$\mathbf{H}^q = \mathbf{A}_s \odot \mathbf{P}^q, \tag{2}$$

where $\mathbf{A}_s := \frac{1}{2}(\mathbf{A} + \mathbf{A}^\intercal)$ is a symmetrized adjacency matrix, and $\mathbf{P}^q$ is a phase matrix with complex numbers. $\odot$ is an element-wise multiplication operation. The definition of the phase matrix is as:

$$\mathbf{P}^q(u, v) := \frac{\exp(i\Theta^q(u, v)) + \exp(i\overline{\Theta}^q(u, v))}{\|\exp(i\Theta^q(u, v)) + \exp(i\overline{\Theta}^q(u, v))\| + \epsilon}. \tag{3}$$



(a)

Edge type	$\mathbf{H}^q(u, v)$
u→v	$\frac{1}{2}(\cos q + i\sin q)$
u←v	$\frac{1}{2}(\cos q - i\sin q)$
u←v	$-\frac{1}{2}(\cos q - i\sin q)$
u→v	$-\frac{1}{2}(\cos q + i\sin q)$
u⇆v	$\cos q$
u⇆v	$-\cos q$
u⇆v	$i\sin q$
u⇆v	$-i\sin q$
u v	0

(b)

Figure 1: Encoding values of a signed directed graph via complex Hermitian adjacency matrix. The parameter q controls the phase angle. The blue and red arrows indicate positive and negative edges, respectively.

$$\Theta^q(u, v) = \begin{cases} q & \text{if } \mathbf{S}(u, v) = 1 \\ \pi + q & \text{if } \mathbf{S}(u, v) = -1 \\ i\,\infty & \text{if } \mathbf{S}(u, v) = 0 \end{cases}$$

$$\overline{\Theta}^q(u, v) = \begin{cases} -q & \text{if } \mathbf{S}(v, u) = 1 \\ \pi - q & \text{if } \mathbf{S}(v, u) = -1 \\ i\,\infty & \text{if } \mathbf{S}(v, u) = 0 \end{cases}$$

where $q \in [0, \pi/2]$ is a parameter that controls the encoding phases. The effect of q is described in detail in Section 6. The symmetrized adjacency matrix encodes the node connectivity, while the phase matrix encodes directions and signs of edges with different phase values. Figure 1 illustrates the edge encoding of the defined Hermitian adjacency matrix. They uniquely encode all the nine node pair relationships of signed-directed graphs. Each relation has a distinct phase and magnitude combination. We can see that $\mathbf{H}^q$ is a complex numbered skew-symmetric form, a complex Hermitian matrix. We then define signed directed magnetic Laplacian with the Hermitian adjacency matrix as follows:

$$\mathbf{L}_U^q := \mathbf{D_s} - \mathbf{H}^q = \mathbf{D_s} - \mathbf{A}_s \odot \mathbf{P}^q, \tag{4}$$

$$\mathbf{L}_N^q := \mathbf{I} - (\mathbf{D_s}^{-\frac{1}{2}} \mathbf{A}_s \mathbf{D_s}^{-\frac{1}{2}}) \odot \mathbf{P}^q, \tag{5}$$

where $\mathbf{D}_s$ is a symmetric degree matrix.

$$\mathbf{D}_s(u, v) = \begin{cases} \sum_{w \in V} \mathbf{A}_s(u, w) & \text{if } u = v \\ 0 & \text{if } u \neq v. \end{cases}$$

$\mathbf{L}_U^q$ and $\mathbf{L}_N^q$ are unnormalized and normalized signed directed magnetic Laplacians, respectively.

Theorem 1. ***The unnormalized and normalized signed directed magnetic Laplacians are positive semi-definite.***

$$x^T \boldsymbol{L}_N^q x \geq 0 \quad x \in \mathbb{R}^n.$$

Theorem 2. ***The eigenvalues of the normalized magnetic Laplacians are in the range of [0, 2].***

Proofs of Theorems are reported in Supplementary Material. By Theorem 1, the Laplacians are diagonalizable by a spectral decomposition. For example, the normalized Laplacian is diagonalized as:

$$\mathbf{L}_N^q = \mathbf{U}\Lambda\mathbf{U}^\dagger. \tag{6}$$

Each column of the matrix $\mathbf{U}$ is eigenvector $\mathbf{u}_k$ and $\mathbf{U}^\dagger$ is a conjugate transpose of $\mathbf{U}$. Λ is a diagonal matrix where the elements are k-th eigenvalues $\Lambda_{k,k} = \lambda_k$. The eigenvalues and eigenvectors contain the structural information of the signed directed graph. We leverage this matrix to define spectral graph convolution.

5 MODEL FRAMEWORK

We propose two perturbation stages: structure perturbation and Laplacian perturbation. The structure perturbation directly perturbs the graph data, resulting in a significant perturbation effect. However, it can hinder the convergence of learning. In contrast, the Laplacian perturbation indirectly influences the graph data. It introduces limited alterations while still maintaining effective perturbation effects.

5.1 STRUCTURE PERTURBATION

Structure perturbation consists of two types of perturbation: edge sign perturbation and edge direction perturbation. In edge sign perturbation, we randomly alter the signs of edges in a given graph. For instance, we sample $p\%$ of positive edges and change their signs to negative, and do the same to negative edges. Similarly, we sample $r\%$ of edges and reverse their directions. Random edge perturbation yields a perturbed graph view, $\widetilde{\mathcal{G}} = (V, \tilde{\mathcal{E}}, \tilde{\mathbf{S}})$. Figure 2 shows an example of structure perturbation.

A moderate degree of perturbation can allow the model to learn robust representations from noisy real-world data and uncover relationships that were previously hidden. This can ultimately enhance the model's generalization performance. Many existing studies [You et al., 2020, Zeng and Xie, 2021, Qiu et al., 2020, Zhu et al., 2021, You et al., 2021] solely depend on this kind of sturucture perturbation. However, excessive perturbation can lead to loss of graph information. The significance of the information in edge signs and directions is acknowledged by the balance and status theories [Heider, 1946, Holland and Leinhardt, 1971]. A single change in edge sign or direction can lead to violations of balance or status theories in all triads associated with that edge. Even perturbing a few edges can result in a catastrophic disruption to the semantics of the graph. Therefore, we introduce a novel Laplacian perturbation that allows robust alterations to the original graphs.

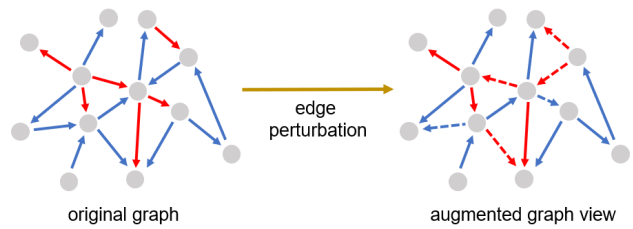


Figure 2: Structure perturbation of a signed directed graph. Dashed edges indicate perturbed edges.

5.2 LAPLACIAN PERTURBATION

The magnetic Laplacian can be modified by adjusting the parameter q. By varying the q value, we can obtain different magnetic Laplacian matrices from a given graph structure. With this property, we implement Laplacian perturbation through q variation, selecting q values from 0 to 0.5π. Despite variations of the Laplacian matrix caused by the q, the underlying graph structure remains unaltered. This technique of Laplacian perturbation is an effective way to augment graphs without distorting the original data. The meaning of Laplacian perturbation is analyzed in Section 6.

The overall perturbation procedure of UGCL consists of two steps; structure perturbation and Laplacian perturbation. We create two different perturbed graph views, $\widetilde{\mathcal{G}}_1$ and $\widetilde{\mathcal{G}}_2$ by applying edge-based structure perturbation to the original graph. Then, we obtain the perturbed signed directed magnetic Laplacians, $\widetilde{\mathbf{L}}_1^{q_1}$ and $\widetilde{\mathbf{L}}_2^{q_2}$, from each of the two structurally perturbed graph views with sampled q values.

5.3 GRAPH ENCODER

5.3.1 Spectral Convolution via Magnetic Laplacian

We define a graph encoder with the perturbed Laplacian matrix. The signed directed magnetic Laplacian $\mathbf{L}^q$, is diagonalizable with eigenvector matrix $\mathbf{U}$, and diagonal eigenvalue matrix Λ thanks to its PSD property. Several studies on graph convolution [Defferrard et al., 2016, Hammond et al., 2011] have utilized the eigenvectors as the discrete Fourier modes in graph signal processing. The transformation of graph signals is performed through the graph Fourier transform, $\hat{\mathbf{x}} = \mathbf{U}^\dagger\mathbf{x}$. A spectral convolution operation of the graph signal is described as:

$$\mathbf{g}_\theta * \mathbf{x} = \mathbf{U}\mathbf{g}_\theta\mathbf{U}^\dagger\mathbf{x}, \tag{7}$$

where $\mathbf{g}_\theta = diag(\theta)$ is a trainable filter. For efficient calculation, Hammond et al. [2011] proposed a truncated Chebyshev polynomial expansion of the filter by:

$$\mathbf{g}_{\theta'}(\mathbf{\Lambda}) \approx \sum_{k=0}^{K} \theta'_k T_k(\overline{\mathbf{\Lambda}}). \tag{8}$$

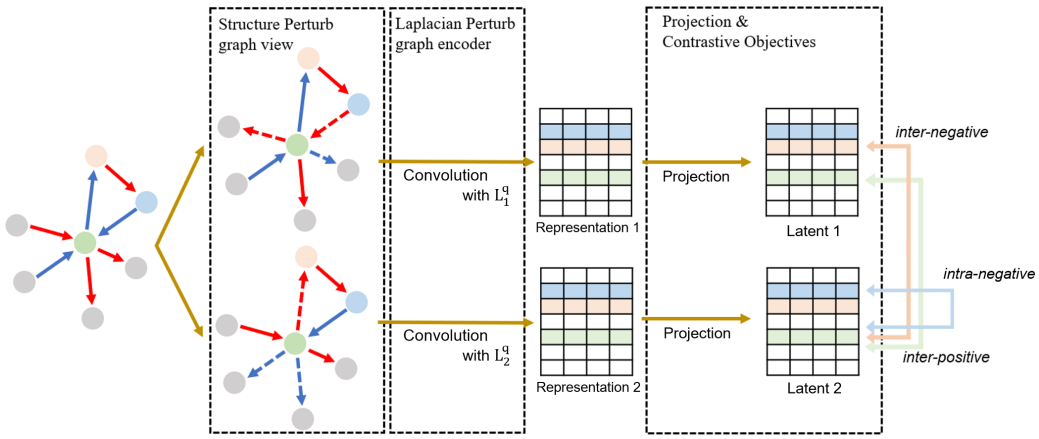


Figure 3: Model overview. There are two structurally perturbed graph views and get perturbed Laplacians from them. Defined graph encoder with the perturbed Laplacians. Contrastive objectives are calculated after projection heads.

Here, $T_0(x) = 1, T_1(x) = x$, and $T_k = 2xT_{k-1}(x) + T_{k-2}(x)$ for $k \geq 2$. θ'_k are Chebyshev coefficients, and $\overline{\mathbf{\Lambda}} = \frac{2}{\lambda_{max}}\mathbf{\Lambda} - \mathbf{I}$ is a normalized eigenvalue matrix where λ_{max} is the largest eigenvalue. Equation (7) becomes a simplified form of spectral graph convolution as

$$\mathbf{g}_{\theta'} * \mathbf{x} = \sum_{k=0}^{K} \theta'_k T_k(\overline{\mathbf{L}})x, \tag{9}$$

where $\overline{\mathbf{L}} = \frac{2}{\lambda_{max}}\mathbf{L} - \mathbf{I}$ analogous to $\overline{\mathbf{\Lambda}}$.

5.3.2 Spectral Convolution Layer

We define the spectral convolution layer with the approximated spectral convolution operation Equation (9). We set the maximum polynomial order K as 1, and assume $\lambda_{max} = 2$ to make it practical. Like GCN [Kipf and Welling, 2016], we set $\theta = \theta'_0 = -\theta'_1$. Then we have approximated convolution layer as:

$$\mathbf{g}_{\theta'} * \mathbf{x} \approx \theta(\mathbf{I} + (\mathbf{D}_s^{-\frac{1}{2}}\mathbf{A}_s\mathbf{D}_s^{-\frac{1}{2}}) \odot \mathbf{P}^q)\mathbf{x}. \tag{10}$$

By the following renormalization trick:

$$\mathbf{I} + (\mathbf{D}_s^{-\frac{1}{2}}\mathbf{A}_s\mathbf{D}_s^{-\frac{1}{2}}) \odot \mathbf{P}^q \rightarrow \overline{\mathbf{D}}_s^{-\frac{1}{2}}\overline{\mathbf{A}}_s\overline{\mathbf{D}}_s^{-\frac{1}{2}} \odot \mathbf{P}^q, \tag{11}$$

where, $\overline{\mathbf{A}}_\mathbf{s} = \mathbf{A}_s + \mathbf{I}$ and $\overline{\mathbf{D}}_\mathbf{s}(i,i) = \sum_j \overline{\mathbf{A}}_\mathbf{s}(i,j)$. The spectral convolution layer is defined as:

$$\mathbf{X}^{l+1} = (\overline{\mathbf{D}}_s^{-\frac{1}{2}}\overline{\mathbf{A}}_s\overline{\mathbf{D}}_s^{-\frac{1}{2}} \odot \mathbf{P}^q)\mathbf{X}^l\mathbf{W}. \tag{12}$$

$\mathbf{X}^{l+1} \in \mathbb{R}^{|V|\times F}$ is the convoluted graph signals or representations after the l-th layer. $\mathbf{W} \in \mathbb{R}^{C\times F}$ is a learnable matrix. C and F are the numbers of input and output channels, respectively. The renormalization trick prevents gradient vanishing and exploding problems.

5.3.3 Graph Encoder

As the output of the convolution layer has both real and imaginary values, we apply an unwinding operation to concatenate the features in the common domain.

$$\mathbf{X}^{(L)}_{\text{unwind}} = [\text{real}(\mathbf{X}^{(L)})||\text{imag}(\mathbf{X}^{(L)}) \otimes (-i)]. \tag{13}$$

A fully connected layer after unwinding finally yields the node representations as

$$\mathbf{Z} = \sigma(\mathbf{X}^{(L)}_{\text{unwind}}\mathbf{W}^{L+1} + \mathbf{B}^{(L+1)}). \tag{14}$$

$\mathbf{Z} \in \mathbb{R}^{|V|\times D}$ is an augmented node representation. We apply a projection head in advocate of [Jacovi et al., 2021, Chen et al., 2020b]. A non-linear transformation $g(\cdot)$ maps the representations to another latent space that can enhance the discriminative power of contrastive learning. $\mathbf{M}$ is a projected latent of the augmented representation.

$$\mathbf{M} = g(\mathbf{Z}). \tag{15}$$

5.4 CONTRASTIVE OBJECTIVE

5.4.1 Inter-view Loss

The contrastive objective aims to align the latent of the same node while differentiating that of other nodes. Two identical nodes from different graph views are considered as an inter-positive pair, while other node pairs are considered inter-negative pairs. For example, a node u from $\widetilde{\mathcal{G}}_1$ and the same node u from $\widetilde{\mathcal{G}}_2$ are the inter-positive pair. On the other hand, other nodes $\{v \in \text{V}; v \neq u\}$ from $\widetilde{\mathcal{G}}_2$ are the inter-negative pair with the node u of $\widetilde{\mathcal{G}}_1$. Even though the nodes in the inter-positive pair come from different graph views, they are the same nodes. Therefore, we aim to maximize the agreement of positive pair latent, $\mathbf{m}_1^u$ and $\mathbf{m}_2^u$. For the same reason, we minimize the agreement of negative pair

latent, $\mathbf{m}_1^u$ and $\mathbf{m}_2^v$. The goal of the inter-view objective is to maximize the similarity of positive pairs and minimize the similarity of negative pairs.

$$\mathcal{L}_{inter} = \frac{1}{|V|} \sum_{u \in V} \log \frac{\exp((\mathbf{m}_1^u \cdot \mathbf{m}_2^u)/\tau)}{\sum_{v \in V} \exp((\mathbf{m}_1^u \cdot \mathbf{m}_2^v)/\tau)} \tag{16}$$

5.4.2 Intra-view Loss

While the inter-view loss compares the latent representations of nodes between two distinct graph views, the intra-view loss calculates the discriminative loss within a single graph view. It is essential to differentiate the latent representations of all nodes from each other, as each node possesses unique characteristics. The objective is to make the latent of all nodes being distinctive. The intra-loss is defined as:

$$\mathcal{L}_{intra} = \frac{1}{K} \sum_{k=1}^{K} \frac{1}{|V|} \sum_{u \in V} \log \frac{1}{\sum_{v \in V, u \neq v} \exp((\mathbf{m}_k^u \cdot \mathbf{m}_k^v)/\tau)}, \tag{17}$$

where k indicates the graph view index. The contrastive loss is the sum of the inter- and intra-view loss functions.

$$\mathcal{L}_{contrastive} = \mathcal{L}_{inter} + \mathcal{L}_{intra} \tag{18}$$

5.5 PREDICTION

The augmented two graph views are the input to graph encoders and they make two node representations, $\mathbf{Z}_1$ and $\mathbf{Z}_2$. The representations are concatenated and fed into the output layer that produces the final node embedding as,

$$\mathbf{R} = \sigma([\mathbf{Z}_1 \parallel \mathbf{Z}_2]\mathbf{W}^{out} + \mathbf{B}^{out}). \tag{19}$$

The final embedding is utilized for downstream tasks such as predicting the edge sign from u to v. The prediction layer is defined as:

$$\hat{y}_{u,v} = \sigma([\mathbf{r}_u || \mathbf{r}_v]\mathbf{W}^{pred} + \mathbf{B}^{pred}). \tag{20}$$

For semi-supervised learning, the proposed model is trained by the following objective function with weight parameter α.

$$\mathcal{L} = \alpha \times \mathcal{L}_{contrastive} + \mathcal{L}_{label}. \tag{21}$$

6 THEORETICAL ANALYSIS

6.1 MEANING OF THE q VALUE

The phases of signed directed magnetic Laplacian are controlled by the parameter q. The q value affects the sensitivity to the sign and direction information by determining the phase angle between the real and imaginary axes. When the q is small, the phase difference between the two reverse edges is small. The Laplacian places less emphasis on directional information. In an extreme case of $q = 0$, direction information is ignored and becomes an undirected model. On the contrary, a large q value also decreases the validity of the encoding. When $q = \pi/2$, a positive edge from node u to node v, and a negative edge from node v to node u are encoded to the same value. Figure 4 describes the effect of q value.

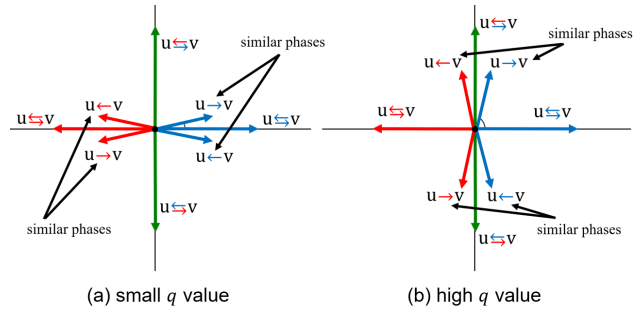


Figure 4: Effect of q value.

6.2 ANALYSIS OF THE LAPLACIAN PERTURBATION

Direct perturbations to nodes, edges, or attributes of the graph can lead to significant differences between the perturbed graph view and the original graph. In contrast, the proposed Laplacian perturbation, achieved through the variation of the q value, indirectly influences the graph data. Analyzing the exact nature of this perturbation and how it differs from the original graph is challenging. To address this issue, we analyze the impact of Laplacian perturbation on the graph information by quantifying graph entropy.

Von Neumann entropy is a widely used form to quantify graph entropy. Ye et al. [2014] introduced the Von Neumann entropy of a directed graph with Laplacian matrix as:

$$H(\mathcal{G}_D) = \frac{\mathrm{Tr}[\mathrm{L}]}{|V|} - \frac{\mathrm{Tr}[\mathrm{L}^2]}{|V|}.$$

Leveraging the Von Neumann entropy of directed graphs, we derive Theorem 3 which states that the entropy of a signed directed graph is less than or equal to the sum of the entropy of a positive edge graph and a negative edge graph.

Theorem 3. ***Von Neumann Entropy of a Signed Directed Graph***

$$H(\mathcal{G}) \leq H(\mathcal{G}_D^+) + H(\mathcal{G}_D^-).$$

$\mathcal{G}$ is a signed directed graph. $\mathcal{G}_D^+$ and $\mathcal{G}_D^-$ are the graph with positive edges and the graph with negative edges, respectively. The Von Neumann entropy of a signed directed graph satisfies the upper bound. This theorem is proved in Supplementary Material. Then, we quantify the effect of Laplacian

		Signed Convolution			Contrastive Learning				Proposed		
Dataset	Metric	SGCN	SDGNN	SDGCN	DiGCL	GCA	SimGRACE	SGCL	UGCL-S	UGCL-L	UGCL
Bitcoin-Alpha	AUC	0.782	0.835	0.858	0.814	0.838	0.823	0.849	**0.896**	0.883	0.886
	Macro-F1	0.668	0.683	0.723	0.653	0.671	0.657	0.712	0.740	0.744	**0.754**
	Micro-F1	0.899	0.909	0.923	0.907	0.913	0.919	0.923	0.947	0.942	**0.949**
	Binary-F1	0.941	0.947	0.958	0.950	0.953	0.957	0.959	**0.973**	0.969	0.971
Bitcoin-OTC	AUC	0.832	0.879	0.887	0.852	0.868	0.859	0.893	**0.914**	0.902	0.910
	Macro-F1	0.710	0.751	0.773	0.725	0.743	0.725	0.781	**0.803**	0.796	0.802
	Micro-F1	0.886	0.902	0.911	0.904	0.907	0.906	0.920	0.935	0.930	**0.937**
	Binary-F1	0.924	0.938	0.950	0.948	0.948	0.948	0.956	0.964	0.962	**0.965**
Epinions	AUC	0.848	0.914	0.939	0.839	0.911	0.913	0.876	0.941	**0.943**	0.942
	Macro-F1	0.741	0.831	0.850	0.726	0.814	0.812	0.798	0.861	**0.865**	0.863
	Micro-F1	0.893	0.912	0.925	0.887	0.913	0.915	0.909	0.934	**0.936**	**0.936**
	Binary-F1	0.937	0.944	0.956	0.936	0.950	0.951	0.948	0.962	**0.963**	**0.963**
Slashdot	AUC	0.740	0.849	0.886	0.813	0.870	0.865	0.783	0.900	0.891	**0.902**
	Macro-F1	0.688	0.729	0.780	0.667	0.750	0.745	0.683	**0.792**	0.785	0.789
	Micro-F1	0.786	0.823	0.855	0.813	0.842	0.833	0.811	**0.864**	0.859	0.863
	Binary-F1	0.869	0.889	0.908	0.887	0.902	0.895	0.884	**0.915**	0.911	0.914

Table 1: Link sign prediction performance. **Bold** and underline indicate the best and the second performance respectively. The performances are the average score of ten experiments with different seed sets.

perturbation through the change in the Von Neumann entropy by q value variation.

Definition 1. ***Perturbation Error*** *Given a perturbation term Δq, we define the perturbation error of the Von Neumann entropy caused by Laplacian perturbation.*

$$\Delta H(\mathcal{G}, q, \Delta q) = H(\mathcal{G}, q) - H(\mathcal{G}, q + \Delta q).$$

Theorem 4. ***Perturbation Error of a Signed Directed Graph***

$$\Delta H(\mathcal{G}_D^+, q, \Delta q), \Delta H(\mathcal{G}_D^-, q, \Delta q) \leq \Delta H(\mathcal{G}, q, \Delta q),$$
$$\Delta H(\mathcal{G}, q, \Delta q) \leq \Delta H(\mathcal{G}_D^+, q, \Delta q) + \Delta H(\mathcal{G}_D^-, q, \Delta q).$$

Signed directed perturbation error is described with lower and upper bounds. And we can notice that the graph entropy varies by the q value variation even though the degree matrix and adjacency matrix are fixed. This perturbation error caused by Laplacian perturbation provides contrastive information in various magnitudes for the encoder. It helps the encoder to focus more on the graph structure rather than just learning from the supervised learning. The proof of Theorem 4 is in the Supplementary Material.

In this subsection, we investigate the impact of Laplacian perturbation on graphs and infer the resulting variations. This analysis demonstrates the effectiveness of Laplacian perturbation as an alternative method that avoids direct perturbation of the graph structure. By preserving the graph's structural integrity while inducing meaningful perturbations, Laplacian perturbation proves to be highly valuable across various applications.

Dataset	# nodes	# pos edges	# neg edges	ratio (%)
Bitcoin-Alpha	3,783	22,650	1,536	93.7
Bitcoin-OTC	5,881	32,029	3,563	90.0
Epinions	131,828	717,667	123,705	85.3
Slashdot	82,144	425,072	124,130	77.4

Table 2: Dataset statistics.

7 EXPERIMENTS

To demonstrate the effectiveness and universality of our model, we evaluate it on various graph datasets. We conduct node classification tasks on unsigned undirected graphs and link sign prediction tasks on signed directed graphs. Through the experiments, we aim to address the following research questions:

- **RQ1**: Does UGCL have better representation learning ability compared to state-of-the-art baselines?
- **RQ2**: How does UGCL demonstrate its universality across different types of graphs?
- **RQ3**: How do the perturbation methods of UGCL affect the performance?
- **RQ4**: How do the model components of UGCL affect the performance?

7.1 LINK SIGN PREDICTION (RQ1)

7.1.1 Baselines

We implemented seven baselines to compare the model performance. There are three signed graph convolution models and four constative learning models.

- **SGCN** [Derr et al., 2018] defines (un)balanced path based on the balanced theory for neighbor aggregation.
- **SDGNN** [Huang et al., 2021] proposed four weight matrices to aggregate neighbor features by edge types.
- **SDGCN** [Ko et al., 2023] proposed a magnetic Laplacian to overcome the limitation of graph Laplacian.
- **GraphCL** [You et al., 2020] randomly perturbs graph structures by dropping or adding edges and nodes.
- **GCA** [Zhu et al., 2021] proposed score-based graph augmentation and node-level contrastive objective.
- **SimGRACE** [Xia et al., 2022] introduced a graph encoder perturbation rather than graph augmentation to overcome the cumbersome augmentation search.
- **SGCL** [Shu et al., 2021] is a GCL for signed directed graphs, which perturbs the edge sign and directions.

7.1.2 Experimental Result

Table 1 summarizes the link sign prediction results. The results are the average of ten independent experiments. There are two variants of UGCL; UGCL-S is a model with structure perturbation, and UGCL-L is a model with Laplacian perturbation. The results show that UGCL and its variants always perform the best on all datasets and in all metrics. We can infer that the proposed model learns the node representation properly from the signed directed graphs.

SGCL or SDGCN demonstrates the second-best performance, following the proposed model. SDGCN achieves good performance by leveraging its signed directed spectral convolution, which effectively utilizes the sign and direction of edges. However, its performance is limited in the context of semi-supervised learning. On the other hand, SGCL is a contrastive learning model specifically designed for signed graphs, and its strong performance on most datasets validates the effectiveness of contrastive learning. However, SGCL exhibits poor performance on the Epinions and Slashdot datasets. Although the structure perturbation methods of SGCL and the proposed model share similarities, the proposed model mitigates the drawbacks of structure perturbation by incorporating Laplacian augmentation. Furthermore, UGCL benefits from an advanced spectral graph encoder compared to the simple GNN encoder utilized in SGCL.

7.2 UNIVERSALITY OF UGCL (RQ2)

The signed directed magnetic Laplacian has the capability to represent various graph types. For instance, in the case of vanilla graphs where all edges are assumed to be bidirectional and positive ($\mathbf{S} = 1$), the encoding of edges is done using 0, 1, resulting in a Laplacian equivalent to the traditional graph Laplacian. When dealing with unsigned directed graphs, the edges are encoded using 0, 1, $\frac{1}{2}(\cos q + i\sin q)$, $\frac{1}{2}(\cos q - i\sin q)$, which resembles the Laplacian definition in MagNet [Zhang et al., 2021]. In essence, the traditional graph Laplacian and the Laplacian employed in MagNet can be viewed as special cases of our encoding. Similarly, the signed undirected graph can be handled as well. Consequently, the signed directed magnetic Laplacian proves to be applicable to all graph types, and the Laplacian perturbation method exhibits wide-ranging utility.

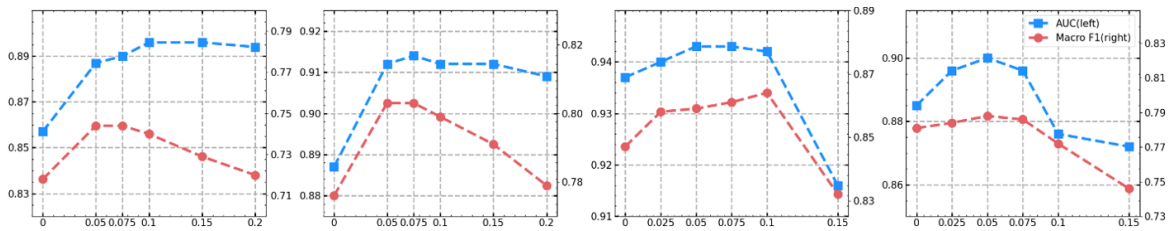
To demonstrate the wide applicability of our approach, we conducted a node classification task on different graph types. However, due to space limitations, we have provided the results in the Supplementary Material. It is worth noting that the proposed UGCL model performs well even on both undirected graphs and directed graphs, further highlighting its versatility and effectiveness across different graph types.

7.3 PERTURBATION ANALYSIS (RQ3)

Figure 5 illustrates the performance variation with respect to the edge perturbation ratio, where Laplacian perturbation is not utilized in this particular experiment. The results indicate that the performance improves initially with a small perturbation ratio, but starts to decline when the ratio exceeds 0.1. This observation highlights the effectiveness of structure perturbation when applied in moderation, while also emphasizing the potential risks associated with excessive perturbation.

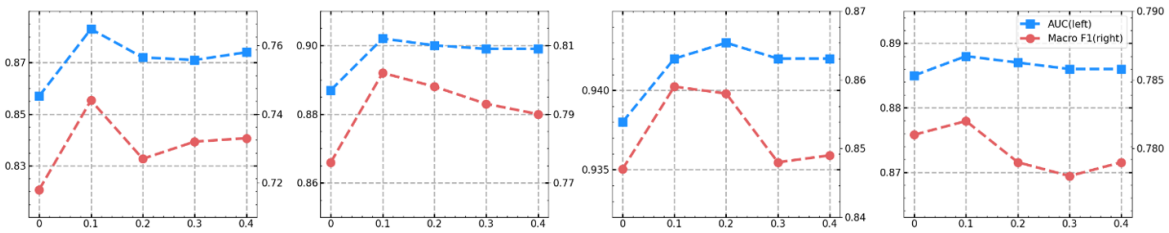
Figure 6 displays the performance variation in response to Laplacian perturbation, where structure perturbation is not employed in this specific experiment. We set $q = \pi/4$ and introduce Gaussian noise with zero mean and varying standard deviations. The x-axis represents the standard deviation of the Gaussian noise. Similar to structure perturbation, performances are low when the standard deviation is zero, indicating no Laplacian perturbation. The results show an initial increase in performance followed by a decrease. Compared to structure perturbation, Laplacian perturbation demonstrates lower sensitivity to the perturbation ratio. This suggests that Laplacian perturbation can provide more stable and robust perturbation effects.

While graph augmentations are essential in contrastive learning, excessively large-scale perturbations can be detrimental to the training process. Structure perturbation provides direct contrastive information, but it runs the risk of compro-



(a) BitcoinAlpha (b) BitcoinOTC (c) Epinions (d) Slashdot

Figure 5: Structure perturbation analysis. The x-axis indicates the perturbing ratio.



(a) BitcoinAlpha (b) BitcoinOTC (c) Epinions (d) Slashdot

Figure 6: Laplacian perturbation analysis. The x-axis indicates noise variance.

mising the network semantics. On the other hand, Laplacian perturbation offers indirect perturbations to the graph data while maintaining effectiveness. In UGCL, these two augmentation techniques are combined to achieve efficient and stable graph augmentation, striking a balance between informative perturbations and preserving the integrity of the network semantics.

7.4 ABLATION STUDY (RQ4)

We check the effects of UGCL components though ablation studies. Table 3 shows the F1 scores and w/o struct, w/o Lapla, and w/o aug are the variation of perturbation methods. Especially, w/o aug, which turns off both perturbations, shows the lowest performance. As we expected, the model leverages the advantage of contrastive learning. It shows that perturbation are important in our model. w/o contrast is a model with $\alpha = 0$ and the model uses label loss only. Note that it does not mean that the model does not utilize the benefits of contrastive learning. Even though the contrastive loss weight is zero, label loss is calculated with augmented representations of graph views. w/o proj is a model without a projection head. It is confirmed that the projection head is useful for robust contrastive learning [Jacovi et al., 2021, Chen et al., 2020b].

	Bitcoin-Alpha	Bitcoin-OTC	Epinions	Slashdot
UGCL	**0.949**	**0.937**	**0.936**	0.863
w/o struct	0.942	0.930	**0.936**	0.859
w/o Lapla	0.947	0.935	0.934	**0.864**
w/o aug	0.919	0.913	0.920	0.853
w/o contrast	0.940	0.931	0.929	0.855
w/o proj	0.942	0.932	0.931	0.858

Table 3: The results of ablation study.

8 CONCLUSION

This paper proposes UGCL, a graph contrastive learning framework. It incorporates two levels of perturbation, structure and Laplacian. Structure perturbation involves modifying the signs and directions of random edges. Although this perturbation may lead to the loss of graph information, it enhances the noise robustness. Laplacian perturbation changes the phase parameter q during each training iteration. It does not directly impact the graph structure but rather influences the magnetic Laplacian. The efficacy of Laplacian perturbation is verified through both theoretical analysis and empirical experiments. By utilizing the perturbed Laplacian, we define a spectral graph encoder. The proposed framework demonstrates its wide applicability to all graph types. Through extensive evaluations on diverse real-world graphs, the proposed framework consistently demonstrates superior performance compared to other existing approaches.

Acknowledgements

This work was supported by the KENTECH Research Grant (202200019A), IITP (No.2021-0-02068, 2023-RS-2022-00156287) grant funded by the Korean government (MSIT), and BK21 Four Intelligence Computing (4199990214639) funded by the National Research Foundation of Korea (NRF)

References

Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pages 1597–1607. PMLR, 2020a.

Ting Chen, Simon Kornblith, Kevin Swersky, Mohammad Norouzi, and Geoffrey E Hinton. Big self-supervised models are strong semi-supervised learners. *Advances in neural information processing systems*, 33:22243–22255, 2020b.

Alexander Cloninger. A note on markov normalized magnetic eigenmaps. *Applied and Computational Harmonic Analysis*, 43(2):370–380, 2017.

Yves Colin de Verdière. Magnetic interpretation of the nodal defect on graphs. *Analysis & PDE*, 6(5):1235–1242, 2013.

Mihai Cucuringu, Huan Li, He Sun, and Luca Zanetti. Hermitian matrices for clustering directed graphs: insights and applications. In *International Conference on Artificial Intelligence and Statistics*, pages 983–992. PMLR, 2020.

Michaël Defferrard, Xavier Bresson, and Pierre Vandergheynst. Convolutional neural networks on graphs with fast localized spectral filtering. *Advances in neural information processing systems*, 29, 2016.

Tyler Derr, Yao Ma, and Jiliang Tang. Signed graph convolutional networks. In *2018 IEEE International Conference on Data Mining (ICDM)*, pages 929–934. IEEE, 2018.

Bruno Messias F. de Resende and Luciano da F. Costa. Characterization and comparison of large directed networks through the spectra of the magnetic laplacian. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 30(7): 073141, 2020.

Michaël Fanuel, Carlos M Alaiz, and Johan AK Suykens. Magnetic eigenmaps for community detection in directed networks. *Physical Review E*, 95(2):022302, 2017.

Michaël Fanuel, Carlos M Alaíz, Ángela Fernández, and Johan AK Suykens. Magnetic eigenmaps for the visualization of directed networks. *Applied and Computational Harmonic Analysis*, 44(1):189–199, 2018.

Stefano Fiorini, Stefano Coniglio, Michele Ciavotta, and Enza Messina. Sigmanet: One laplacian to rule them all. *arXiv preprint arXiv:2205.13459*, 2022.

Satoshi Furutani, Toshiki Shibahara, Mitsuaki Akiyama, Kunio Hato, and Masaki Aida. Graph signal processing for directed graphs based on the hermitian laplacian. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 447–463. Springer, 2019.

Krystal Guo and Bojan Mohar. Hermitian adjacency matrix of digraphs and mixed graphs. *Journal of Graph Theory*, 85(1):217–248, 2017.

David K Hammond, Pierre Vandergheynst, and Rémi Gribonval. Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis*, 30(2): 129–150, 2011.

Kaiming He, Haoqi Fan, Yuxin Wu, Saining Xie, and Ross Girshick. Momentum contrast for unsupervised visual representation learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9729–9738, 2020.

Yixuan He, Michael Perlmutter, Gesine Reinert, and Mihai Cucuringu. Msgnn: A spectral graph neural network based on a novel magnetic signed laplacian. In *Learning on Graphs Conference*, pages 40–1. PMLR, 2022.

Fritz Heider. Attitudes and cognitive organization. *The Journal of psychology*, 21(1):107–112, 1946.

Paul W Holland and Samuel Leinhardt. Transitivity in structural models of small groups. *Comparative group studies*, 2(2):107–124, 1971.

Junjie Huang, Huawei Shen, Liang Hou, and Xueqi Cheng. Sdgnn: Learning node representation for signed directed networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 196–203, 2021.

Alon Jacovi, Swabha Swayamdipta, Shauli Ravfogel, Yanai Elazar, Yejin Choi, and Yoav Goldberg. Contrastive explanations for model interpretability. *arXiv preprint arXiv:2103.01378*, 2021.

Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.

Taewook Ko, Yoonhyuk Choi, and Chong-Kwon Kim. A spectral graph convolution for signed directed graphs via magnetic laplacian. *Neural Networks*, 2023.

Shuangli Li, Jingbo Zhou, Tong Xu, Dejing Dou, and Hui Xiong. Geomgcl: geometric graph contrastive learning for molecular property prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 4541–4549, 2022.

Elliott H Lieb and Michael Loss. Fluxes, laplacians, and kasteleyn's theorem. In *Statistical Mechanics*, pages 457–483. Springer, 1993.

Zihan Lin, Changxin Tian, Yupeng Hou, and Wayne Xin Zhao. Improving graph collaborative filtering with neighborhood-enriched contrastive learning. In *Proceedings of the ACM Web Conference 2022*, pages 2320–2329, 2022.

Jianxi Liu and Xueliang Li. Hermitian-adjacency matrices and hermitian energies of mixed graphs. *Linear Algebra and its Applications*, 466:182–207, 2015.

Bojan Mohar. A new kind of hermitian matrices for digraphs. *Linear Algebra and its Applications*, 584:343–352, 2020.

Alessandro Olgiati. Remarks on the derivation of gross-pitaevskii equation with magnetic laplacian. In *Advances in Quantum Mechanics*, pages 257–266. Springer, 2017.

Erlin Pan and Zhao Kang. Multi-view contrastive graph clustering. *Advances in neural information processing systems*, 34:2148–2159, 2021.

Zhen Peng, Wenbing Huang, Minnan Luo, Qinghua Zheng, Yu Rong, Tingyang Xu, and Junzhou Huang. Graph representation learning via graphical mutual information maximization. In *Proceedings of The Web Conference 2020*, pages 259–270, 2020.

Jiezhong Qiu, Qibin Chen, Yuxiao Dong, Jing Zhang, Hongxia Yang, Ming Ding, Kuansan Wang, and Jie Tang. Gcc: Graph contrastive coding for graph neural network pre-training. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1150–1160, 2020.

Lin Shu, Erxin Du, Yaomin Chang, Chuan Chen, Zibin Zheng, Xingxing Xing, and Shaofeng Shen. Sgcl: Contrastive representation learning for signed graphs. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, pages 1671–1680, 2021.

MA Shubin. Discrete magnetic laplacian. *Communications in mathematical physics*, 164(2):259–275, 1994.

Rahul Singh and Yongxin Chen. Signed graph neural networks: A frequency perspective. *arXiv preprint arXiv:2208.07323*, 2022.

Fan-Yun Sun, Jordan Hoffmann, Vikas Verma, and Jian Tang. Infograph: Unsupervised and semi-supervised graph-level representation learning via mutual information maximization. *arXiv preprint arXiv:1908.01000*, 2019.

Susheel Suresh, Pan Li, Cong Hao, and Jennifer Neville. Adversarial graph augmentation to improve graph contrastive learning. *Advances in Neural Information Processing Systems*, 34:15920–15933, 2021.

Zekun Tong, Yuxuan Liang, Henghui Ding, Yongxing Dai, Xinke Li, and Changhu Wang. Directed graph contrastive learning. *Advances in Neural Information Processing Systems*, 34:19580–19593, 2021.

Petar Velickovic, William Fedus, William L Hamilton, Pietro Liò, Yoshua Bengio, and R Devon Hjelm. Deep graph infomax. *ICLR (Poster)*, 2(3):4, 2019.

Yingheng Wang, Yaosen Min, Xin Chen, and Ji Wu. Multi-view graph contrastive representation learning for drug-drug interaction prediction. In *Proceedings of the Web Conference 2021*, pages 2921–2933, 2021.

Jun Xia, Lirong Wu, Jintao Chen, Bozhen Hu, and Stan Z Li. Simgrace: A simple framework for graph contrastive learning without data augmentation. In *Proceedings of the ACM Web Conference 2022*, pages 1070–1079, 2022.

Cheng Ye, Richard C Wilson, César H Comin, Luciano da F Costa, and Edwin R Hancock. Approximate von neumann entropy for directed graphs. *Physical Review E*, 89(5):052804, 2014.

Yuning You, Tianlong Chen, Yongduo Sui, Ting Chen, Zhangyang Wang, and Yang Shen. Graph contrastive learning with augmentations. *Advances in Neural Information Processing Systems*, 33:5812–5823, 2020.

Yuning You, Tianlong Chen, Yang Shen, and Zhangyang Wang. Graph contrastive learning automated. In *International Conference on Machine Learning*, pages 12121–12132. PMLR, 2021.

Jiaqi Zeng and Pengtao Xie. Contrastive self-supervised learning for graph classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 10824–10832, 2021.

Xitong Zhang, Yixuan He, Nathan Brugnone, Michael Perlmutter, and Matthew Hirn. Magnet: A neural network for directed graphs. *Advances in Neural Information Processing Systems*, 34:27003–27015, 2021.

Huasong Zhong, Jianlong Wu, Chong Chen, Jianqiang Huang, Minghua Deng, Liqiang Nie, Zhouchen Lin, and Xian-Sheng Hua. Graph contrastive clustering. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9224–9233, 2021.

Yanqiao Zhu, Yichen Xu, Feng Yu, Qiang Liu, Shu Wu, and Liang Wang. Graph contrastive learning with adaptive augmentation. In *Proceedings of the Web Conference 2021*, pages 2069–2080, 2021.