Universal Graph Contrastive Learning with a Novel Laplacian Perturbation

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 based on Laplacian perturbation. The framework employs a complex Hermitian adjacency matrix which enables the application of a novel graph Laplacian. Unlike existing GCL methods, the proposed framework is not restricted to specific graph types and enjoys a wider range of applicability. We demonstrate that a spectral graph convolution based on the Laplacian extracts the representations of diverse graph types successfully. 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] adds 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** (Universal GCL). We devised 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 the magnetic Laplacian. Recent studies [Ko, 2022, 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 encodes the signs and directions of the edges using complex phase values. We propose to modify the phase values of the original Laplacian, which led us 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 struc-

Submitted to the 39th Conference on Uncertainty in Artificial Intelligence (UAI 2023). **To be used for reviewing only**.

tural 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.

- We introduce a novel Universal Graph Contrastive Learning (**UGCL**) technique that can be applied to any types of graphs.
- To the best of our knowledge, this is the first attempt to introduce the magnetic Laplacian perturbation method. Laplacian perturbation enables fine adjustments while keeping the topology of graphs.
- The proposed model shows improved performance in experiments on various real-world graphs.

2 RELATED WORK

2.1 GRAPH CONTRASTIVE LEARNING

SimCLR [Chen et al., 2020a] and MoCo [He et al., 2020] proposed contrastive learning for image classification achieved great success, triggering many follow-up studies that improved 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].

Most existing GCL techniques are limited to unsigned undirected graphs. To expand the range of applications, 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], which deals with signed directed graphs by randomly altering the signs and directions of edges, is another attempt to broaden the scope of the GCL technique.

2.2 HERMITIAN ADJACENCY MATRIX AND MAGNETIC LAPLACIAN

The traditional adjacency matrix uses 0 and 1 to encode the connectivity in graphs. The matrix has a desirable property of symmetricity that makes the graphs amenable to efficient spectral analyses. However, the symmetricity is lost when edges are directed. Several studies have proposed to use the Hermitian adjacency matrices as an alternative to the traditional adjacency matrices 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 magnetic Laplacian and introduced a spectral graph convolution technique for directed graphs. Recently, several researchers [Ko, 2022, 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 graph analyses.

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 $\mathbf{S}(u, v)$ takes 1 or -1 if there is a positive or negative directed edge from node u to node v, respectively. Each node has one of three possible relationships with other nodes: none, positive, or negative. This results in nine possible relationships for each node pair, as shown in Fig. 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, when a graph is signed or directed, the graph Laplacian becomes asymmetric. This leads to complex eigenvalues, making it difficult to satisfy the conditions for spectral convolution in the graph Fourier transform.

Recent studies [Ko, 2022, 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 as:

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

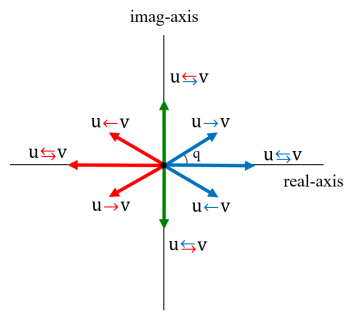
where $\mathbf{A}_s := \frac{1}{2}(\mathbf{A} + \mathbf{A}^\intercal)$ is a symmetrized adjacency matrix, and $\mathbf{P}^q$ is a phase matrix with complex numbers. $\odot$ is an element-wise multiplication operation. The definition of the phase matrix is as:

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

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

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

where $q \in [0, \pi/2]$ is a parameter that controls the encoding phases. The effect of q is described in detail in Sec. 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. The encoding scheme distinctively depicts all the possible node pair relationships that can appear in signed directed graphs. Each relation type has a unique phase and magnitude combination. We then define signed directed magnetic Laplacian with the Hermitian adjacency matrix as follows:



(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. q controls the phase angle.

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

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

$\mathbf{L}_U^q$ 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 real 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 consists of $|V|$ 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

5.1 STRUCTURE PERTURBATION

Edge-based 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, $\widetilde{\mathcal{G}} = (V, \widetilde{\mathcal{E}}, \widetilde{\mathbf{S}})$. Fig. 2 shows an example of structure perturbation. Many existing studies on GCL [You et al., 2020, Zeng and Xie, 2021, Qiu et al., 2020, Zhu et al., 2021, You et al., 2021] solely depend on sturucture perturbation.

The significance of the information in edge signs and directions is acknowledged by the balance and status theories [Heider, 1946, Holland and Leinhardt, 1971]. The random perturbation may disturb vital triad patterns. A single edge sign or direction change incurs violations of balance or status theories in all triads involved with the edge. Because the clustering coefficients of most real networks are much larger than those of random networks, perturbation of a few edges can cause catastrophic disturbance to the graph semantics. Of course, carefully designed moderate degree of perturbation can allow the model to learn robust representations from noisy data and uncover features that were previously hidden. On the other hand, the excessive perturbation can lead to loss of inherent graph information and to dismal performance. Therefore, we introduce a novel Laplacian perturbation that allows robust alterations to the original graphs.

5.2 LAPLACIAN PERTURBATION

The parameter q, which determines the encoding phases (See Sec. 4), is the core of our Laplacian perturbation. First, we randomly select edges for perturbation. Then for each selected edge, we change q of the edge by adding a random Gaussian noise. Note that contrary to structural perturbation, the Laplacian perturbation value is continuously allowing refined manipulation. Despite the Laplacian matrix being changed by the q variation, the underlying graph topology remains unaltered if proper restrictions are levied to the added Gaussian noise. 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 Sec. 6.

The overall perturbation procedure of UGCL consists of two steps; edge based 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, $\widetilde{\mathbf{L}}_1^q$ and $\widetilde{\mathbf{L}}_2^q$, from each of the two structurally perturbed graph views.

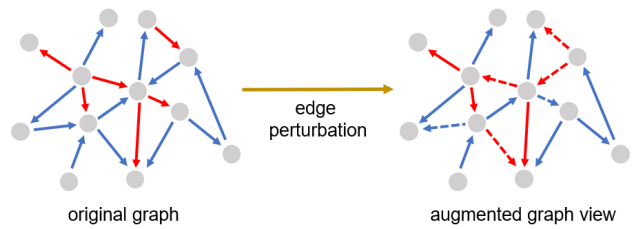


Figure 2: Structure perturbation of a signed directed graph. Dashed edges indicate perturbed edges.

5.3 GRAPH ENCODER

5.3.1 Spectral Convolution via Magnetic Laplacian

Now we define a graph encoder with a perturbed Laplacian matrix. The signed directed magnetic Laplacian $\mathbf{L}^q$, is diagonalizable with eigenvector matrix $\mathbf{U}$, and diagonal eigenvalue matrix Λ. 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:

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

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. Eq. 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 (Eq. 9). We set the maximum polynomial order K as 1, and assume $\lambda_{max} = 2$

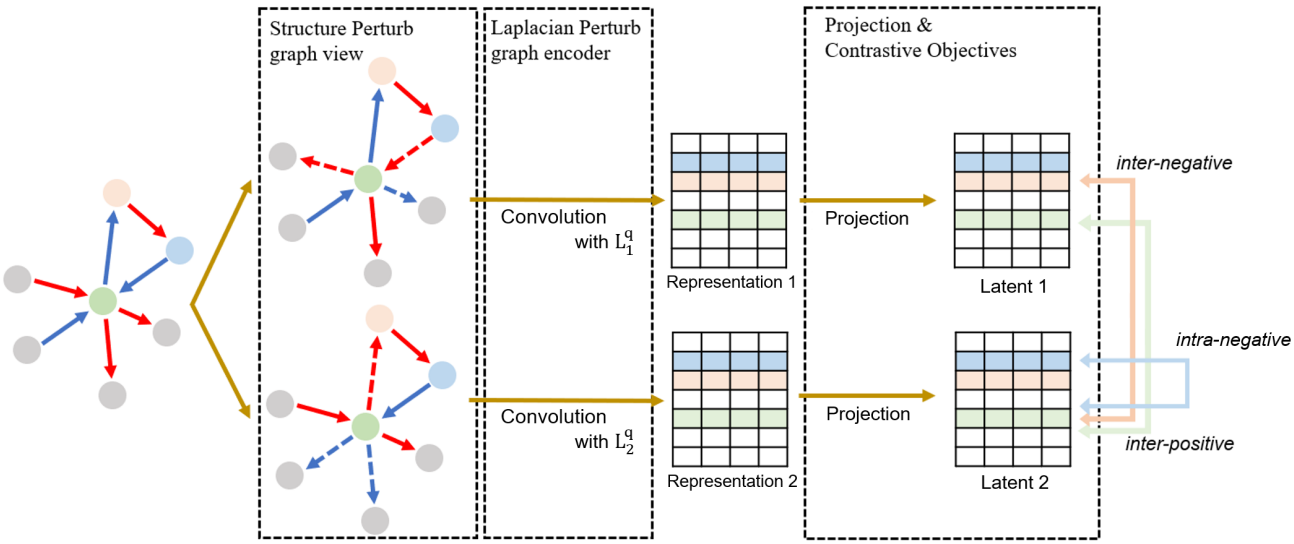


Figure 3: Model overview. There are two structure perturbed graph views and get perturbed Laplacians from them. Defined graph encoder with the perturbed Laplacians.

to make it practical. We prove that eigenvalues of magnetic Laplacians are between [0, 2] (Refer to Appendix). Like GCN [Kipf and Welling, 2016], we set $\theta = \theta'_0 = -\theta'_1$. Then we have approximated the convolution layer as:

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

By the following renormalization trick:

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

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:

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

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

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

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

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

$\mathbf{Z} \in \mathbb{R}^{|V|\times D}$ 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. $\mathbf{M}$ is a projected latent of the augmented representation.

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

5.4 CONTRASTIVE OBJECTIVE

The loss in GCL has two components; inter-view loss and intra-view loss.

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 an inter-positive pair. On the other hand, a node $\{v \in \text{V}; v \neq u\}$ from $\widetilde{\mathcal{G}}_2$ and a node u from $\widetilde{\mathcal{G}}_1$ form an inter-negative pair. We aim to minimize the disagreement of inter-positive pair latent, $\mathbf{m}_1^u$ and $\mathbf{m}_2^u$ while maximizing the disagreement of inter-negative pair latent, $\mathbf{m}_1^u$ and $\mathbf{m}_2^v$.

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

5.4.2 Intra-view Loss

The intra-view loss is similar to the inter-view loss. While the inter-view loss deals with a pair of nodes from two

different graph views, the intra-view loss calculates the discriminative loss within a single graph view. The objective is to make the latent of all nodes being distinctive. The intra-loss is defined as,

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

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

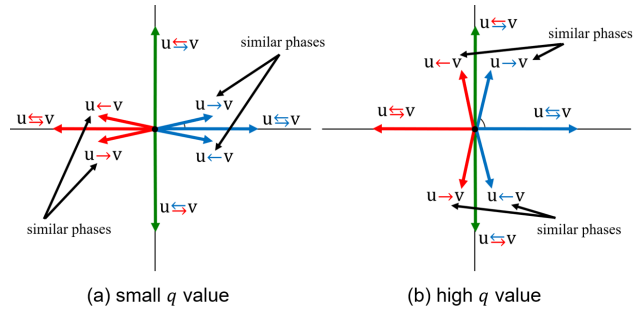
6.1 MEANING OF THE q VALUE

The phases of signed directed magnetic Laplacian are controlled by the parameter q which determines the sensitivity of the sign and direction information. When q is small, the phase difference between two reverse edges is small (see Fig. 4(a)), and the Laplacian places less emphasis on directional information. In an extreme case of $q = 0$, directional information is ignored and becomes an undirected model. On the contrary, a large q 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 (see Fig. 4(b)).

6.2 MEANING OF THE LAPLACIAN PERTURBATION

Graph information can be quantified by the Von Neumann entropy. Ye et al. [2014] defined the Von Neumann entropy of a directed graph with Laplacian matrix, L, as:

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



(a) small q value (b) high q value

Figure 4: Effect of the q value.

We derive a theorem 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*** *The Von Neumann entropy of a signed directed graph satisfies the following bound.*

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

where $\mathcal{G}$, $\mathcal{G}_D^+$, and $\mathcal{G}_D^-$ are the original graph, the graph with positive edges only, and the graph with negative edges only, respectively.

Then, we quantify the effect of Laplacian 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*** *Signed directed perturbation error is described with lower and upper bounds.*

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

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

Note that the graph entropy is perturbed by the q value variation even though the degree matrix and adjacency matrix are fixed. Moreover, it is approximately quantified via the Von Neumann entropy. The entropy error resulting from Laplacian perturbation provides varying degrees of contrastive information that can be used by the encoder to learn representations while preserving the underlying graph structure. This enables the encoder to place greater emphasis on the graph structure. The proofs of Theorem 3 and 4 are included in the Supplementary Material.

		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.

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. For details on the datasets, baselines, task settings, and reproducible information, please refer to the Supplemental Material. 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**: Is UGCL universal?
- **RQ3**: How does the perturbation affect the performance?

7.1 LINK SIGN PREDICTION (RQ1)

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 augmentation only, and UGCL-L is a model with Laplacian augmentation only. 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 shows the second best performance next to the proposed model. SDGCN achieves good performance thanks to its signed directed spectral convolution which can fully utilize the sign and direction information. SGCL is a contrastive model specifically designed for signed graphs, and its good performance on most datasets indicates the effectiveness of contrastive learning. However, on Epinions and Slashdot, SGCL shows poor performance. Though the structure perturbation methods of SGCL and the proposed model are similar, the proposed model reduces the disadvantage of structure perturbation by adding Laplacian augmentation. Furthermore, UGCL enjoys an advanced spectral graph encoder. We conduct a node classification task also, and its results are reported in the Supplement Material.

7.2 UNIVERSALITY OF UGCL (RQ2)

The signed directed magnetic Laplacian can represent any graph types. For example, edges of vanilla graphs are encoded with $\{0, 1\}$ assuming all edges are bidirectional and positive ($\mathbf{S} = 1$). Then the Laplacian is the same as the traditional graph Laplacian. Similarly, if a graph is unsigned directed, its edges are encoded with $\{0, 1, \frac{1}{2}(\cos q + i\sin q), \frac{1}{2}(\cos q - i\sin q)\}$. It is the same form of Laplacian defined at MagNet [Zhang et al., 2021]. In other words, the traditional graph Laplacian and the Laplacian from MagNet are the special cases of our encoding. Similarly, we can handle the signed undirected graph. Thus, the signed directed magnetic Laplacian handles all graph types and the Laplacian perturbation method also enjoys wide applicability.

7.3 PERTURBATION ANALYSIS (RQ3)

Here, we analyze the effect of structure and Laplacian augmentation. Fig. 5 shows that the performance varies according to the edge perturbing ratio. Laplacian perturbation is not adopted in this experiment. From no perturbation, the performances increase with a small perturbing ratio and

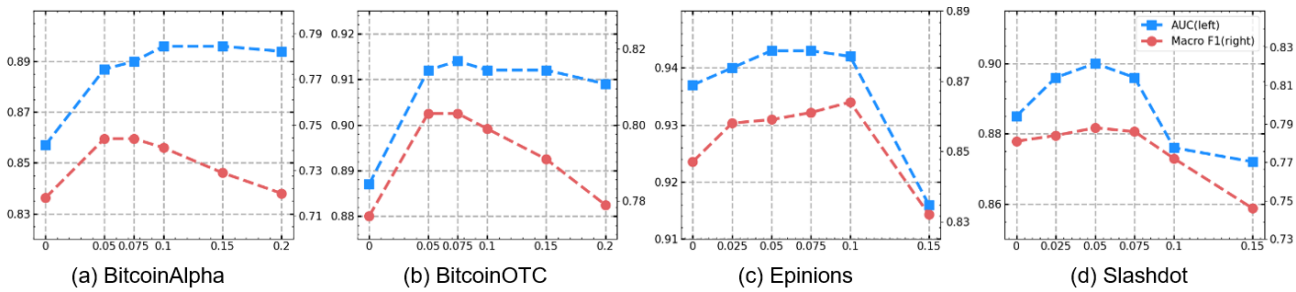


Figure 5: Structure perturbing analysis. The x-axis indicates the perturbing ratio.

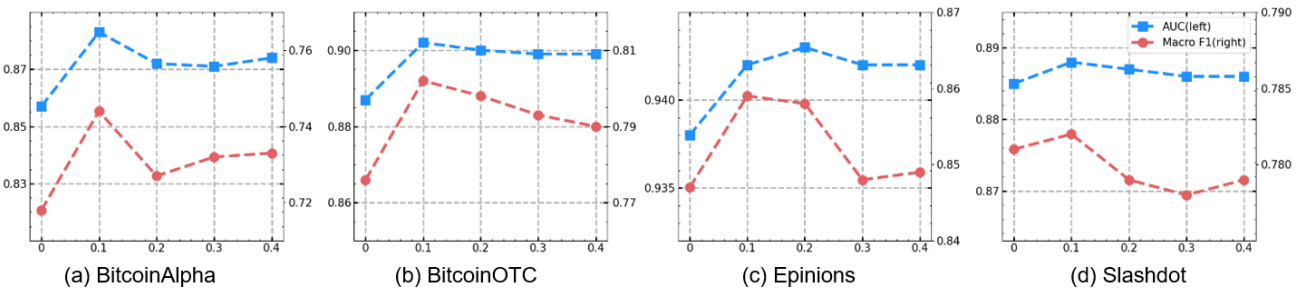


Figure 6: Laplacian perturbing analysis. The x-axis indicates noise variance.

decrease if the ratio is larger than 0.1.

Fig. 6 shows the performances of the Laplacian augmentation only. We set $q = \pi/4$ and add Gaussian noises with zero mean and different standard variations. The x-axis is the standard deviation of the Gaussian noise. We manipulate the Gaussian noises such that $0 \leq q \leq \pi/2$ after perturbation. Similar to structure perturbation, zero standard deviations show low performances. Zero standard deviation means there is no Laplacian perturbation. From no perturbation, the performances increase and then decrease. Though graph augmentations are compulsory in contrastive learning, large scale perturbation is detrimental to training. Structure perturbation gives direct contrastive information while it may lose the network semantics. Laplacian perturbation gives indirect perturbation of graph data but is still effective. UGCL combines these two augmentations for efficient and stable graph augmentation.

8 ABLATION STUDY

We check the effects of UGCL components through ablation studies. Table 2 shows the F1 scores of UGCL, UGCL-S, UGCL-L, and no augmentation. As expected, no-augmentation shows the lowest performance, showing that augmentations are important in our model. The full model performed best on all datasets except the Slashdot dataset. Even though UGCL-L performs best on the Slashdot dataset, its performance gain over UGCL is minimal. The result indicates that the full model which utilizes both structure augmentation and Laplacian augmentation generally provides satisfactory performance on all datasets.

	Bitcoin-Alpha	Bitcoin-OTC	Epinions	Slashdot
UGCL	**0.949**	**0.937**	**0.936**	0.863
w/o structure aug	0.942	0.930	**0.936**	0.859
w/o Laplacian aug	0.947	0.935	0.934	**0.864**
w/o augmentation	0.919	0.913	0.920	0.853

Table 2: The results of ablation study.

9 CONCLUSION

This paper proposed UGCL, a graph contrastive learning framework for universal graphs. To make augmented node representations, we introduced two levels of the perturbation process. Structure perturbation changes the signs and directions of random edges. Even though it may lose some vital information of the original graphs, we check that it makes the model noise-robust. Laplacian perturbation changes the phase parameter q for every training iteration. It does not directly influence the graph structure but perturbs the Laplacian matrix. With the perturbed Laplacian, we defined a spectral convolution-based graph encoder. Then we used the inter- and intra-view contrastive objectives with the augmented node representations. The proposed model was evaluated on several real-world datasets and its performance is compared with that of multiple baselines. The experimental results show that our proposed model has superior performance to others. UGCL and its variants ranked at the top for every dataset and metric.

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.

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.

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. A graph convolution for signed directed graphs. *arXiv preprint arXiv:2208.11511*, 2022.

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

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.

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.

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.

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.