CS460 Report: On the Expressive Power of Geometric GNN

Akankshya Nayak
School of Biological Sciences
NISER

Soumya Dasgupta
School of Physical Sciences
NISER

Abstract

The major constraint of the expressivity of the standard Graphical Neural Networks (GNN) through Weisfeiler-Leman (WL) isomorphic test is their inaplicability to geometric graphs. In order to get a better understanding of the expressive power of GNNs with respect to geometric graphs that are embedded within the Euclidean space, Chaitanya et al. [10] had proposed a geometric version of Weisfeiler-Leman test. In their study, the expressivity was focused on prevalent reference frame-free architectures like Invariant, Cartesian Equivariant and Spherical Equivariant. Here, we aim to involve certain reference frame-based geometric GNNs and use some synthetic experiments to measure the expressivity in this new class of architecture. Through this, we aim to get a deeper theoretical and empirical understanding of the expressivity and theoretical limits of reference frame-based geometric GNNs.

1 Introduction

Graphical Neural Networks (GNNs) can be simply defined as topologies of neural networks that operate on graphs. One of the key features that is taken into consideration in order to evaluate the expressivity of GNNs is their ability to distinguish between different types of graphs. This is often done by testing their performance in the graph isomorphism problem to determine whether the given two graphs show topological equivalence.

Research into the expressive capabilities of GNNs has primarily relied on the Weisfeiler-Lehman (WL) test [12]. However, this test is not suitable for analyzing geometric graphs that are embedded in Euclidean space, like those found in biomolecules, materials, and other physical systems. To address this limitation, Chaitainya et al. [10] propose a geometric version of the WL test, known as the GWL test, which explores the expressive power of geometric GNNs while accounting for physical symmetries such as permutations, rotations, reflections, and translations. GWL unravels how key design choices influence geometric GNN expressivity:

1. Invariant layers have limited expressivity as they fail to distinguish one-hop identical geometric graphs
2. Equivariant layers distinguish a larger class of graphs by propagating geometric information beyond local neighbourhoods
3. Higher order tensors and scalarization enable maximally powerful geometric GNNs
4. GWL's discrimination-based perspective is equivalent to universal approximation

However, the aforementioned study is based on reference frame free algorithms like SchNet, DimeNet, PaiNN, E(n)-GNN etc. (Figure-1) We aim to look along an emerging direction of reference frame based geometric GNN models like ComENet and ClofNet. Here we aim to add reference frame based models and use the synthetic experiments on studying their expressivity.

Preprint. Under review.

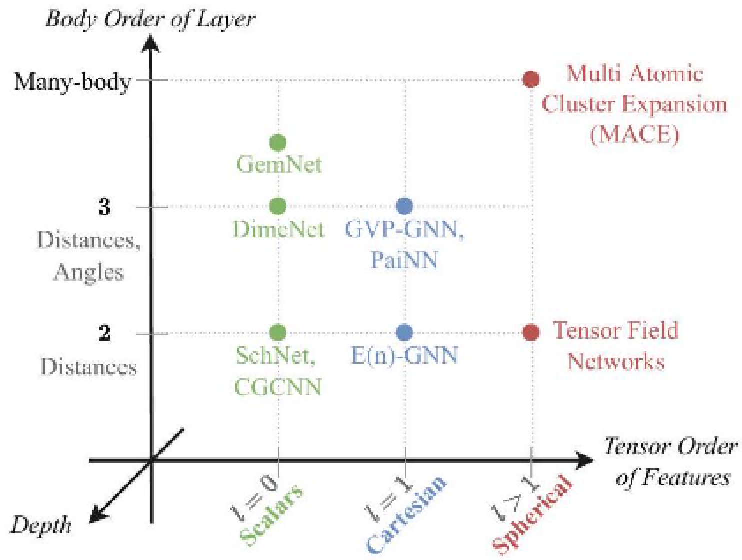

Figure 1: Axes of geometric GNN expressivity: (1) Scalarisation body order: increasing body order of scalarisation builds expressive local neighbourhood descriptors; (2) Tensor order: higher order spherical tensors determine the relative orientation of neighbourhoods; and (3) Depth: deep equivariant layers propagate geometric information beyond local neighbourhoods [10]

2 Related Work

Literature suggests that standard GNNs are at most expressive as the WL algorithm [12][19][13]. k-WL hierarchy attempts in generalizing the WL algorithm for classifying k-tuples of vertices[7][1]. Researchers have also proposed a new hierarchy building on high-order subgraphs within neighborhood aggregation (N-WL) instead of k-WL [18]. However, recent theoretical studies on geometric graph neural networks and their universal properties have indicated that architectures such as TFN, GemNet, and GVP-GNN can serve as universal approximators of continuous, $\mathfrak{G}$-equivariant, or $\mathfrak{G}$-invariant multiset functions on point clouds, which are essentially fully connected graphs [5][16][11][8].

3 Baseline Algorithms

Before going into the Geometric Weisfeller-Leman Test(GWL), we first briefly discuss about the WL test.

WL test It checks whether two graphs are isomorphic or not. In the zeroth iteration, WL assigns a colour $c_i^{(0)} \in C$ where C is a countable space of colours. Nodes having the same features are given the same colour. In the subsequent iterations the colours of each node are updated in the following manner:

$$\boldsymbol{c}_i^{(t)} := \mathrm{HASH}\left(\boldsymbol{c}_i^{(t-1)}, \left\{\left\{\boldsymbol{c}_j^{(t-1)} \mid j \in \mathcal{N}_i\right\}\right\}\right) \tag{1}$$

where HASH is an injective map which gives a unique colour to each input.The test terminates when the partition of the nodes induced by the colours becomes stable. Given two graphs $\mathcal{G}$ and $\mathcal{H}$, if there exists some iteration t for which $\left\{\left\{c_i^{(t)} \mid i \in \mathcal{V}(\mathcal{G})\right\}\right\} \left\{\left\{c_i^{(t)} \mid i \in \mathcal{V}(\mathcal{H})\right\}\right\}$, then the graphs are not isomorphic. Otherwise, the WL test is inconclusive, and we say it cannot distinguish the two graphs. This WL test has its own limitations. However we focus on the limitations which are related to our project. One such limitation is its inapplicability for Geometric graphs.

Isomorphism for Geometric graphs Two geometric graphs $\mathcal{G}$ and $\mathcal{H}$ are geometrically isomorphic if there exists an attributed graph isomorphism b such that the geometric attributes are equivalent, up to global group actions $Q_{\mathfrak{g}} \in \mathfrak{G}$ and $\overrightarrow{\boldsymbol{t}} \in T(d)$

$$\left(\boldsymbol{s}_i^{(\mathcal{G})}, \overrightarrow{\boldsymbol{v}}_i^{(\mathcal{G})}, \overrightarrow{\boldsymbol{x}}_i^{(\mathcal{G})}\right) = \left(\boldsymbol{s}_{b(i)}^{(\mathcal{H})}, \boldsymbol{Q}_{\mathfrak{g}} \overrightarrow{\boldsymbol{v}}_{b(i)}^{(\mathcal{H})}, \boldsymbol{Q}_{\mathfrak{g}}\left(\overrightarrow{\boldsymbol{x}}_{b(i)}^{(\mathcal{H})} + \overrightarrow{\boldsymbol{t}}\right)\right) \quad \text{for all } i \in \mathcal{V}(\mathcal{G})$$

where $\mathfrak{G}$ is a Lie group and $Q_{\mathfrak{g}}$ denotes orthogonal transformations. $T(d)$ denotes translation group. In simple words, two geomtric graphs are isomorphic if the two graphs super-impose on each other after some rotations and translations.

Geomtric Weisfeller Leman Test : At the zeroth iteration, every node in the graph $i \in \mathcal{V}$ is assigned a node color $c_i \in C$ depending on its scalar features s_i and a $\boldsymbol{g}_i$ containing the geometric or the vector information associated to it:

$$c_i^{(0)} := \mathrm{HASH}\left(\boldsymbol{s}_i\right), \quad \boldsymbol{g}_i^{(0)} := \left(c_i^{(0)}, \overrightarrow{\boldsymbol{v}}_i\right)$$

At every tth iteration the geometric features are updated:

$$\boldsymbol{g}_i^{(t)} := \left(\left(c_i^{(t-1)}, \boldsymbol{g}_i^{(t-1)}\right), \left\{\left\{\left(c_j^{(t-1)}, \boldsymbol{g}_j^{(t-1)}, \overrightarrow{\boldsymbol{x}}_{ij}\right) \mid j \in \mathcal{N}_i\right\}\right\}\right)$$

The colours are then changed in the following manner

$$c_i^{(0)} := \text{I-HASH}^t(\boldsymbol{g}_i^{(t)})$$

The procedure terminates when the partitions of the nodes induced by the colours do not change from the previous iteration. Finally, given two geometric graphs $\mathcal{G}$ and $\mathcal{H}$, if there exists some iteration t for which $\left\{\left\{c_i^{(t)} \mid i \in \mathcal{V}(\mathcal{G})\right\}\right\} \neq \left\{\left\{c_i^{(t)} \mid i \in \mathcal{V}(\mathcal{H})\right\}\right\}$, then GWL deems the two graphs as being geometrically non-isomorphic. Otherwise, we say the test cannot distinguish the two graphs.

Invariant GWL Here the geometric information are not passed beyond the local neighbourhood and the colour in each iteration is uploaded as

$$c_i^{(t)} := \mathrm{I-HASH}\left(\left(c_i^{(t-1)}, \overrightarrow{\boldsymbol{v}}_i\right), \left\{\left\{\left(c_j^{(t-1)}, \overrightarrow{\boldsymbol{v}}_j, \overrightarrow{\boldsymbol{x}}_{ij}\right) \mid j \in \mathcal{N}_i\right\}\right\}\right).$$

Characterising the expressive power of geometric GNNs Chaitanya et al. [1] have suggested the following theorem and proposition that tells us about the expressive power of gGNN(Proofs are given in [1]).

Theorem 1 Any pair of geometric graphs distinguishable by a $\mathfrak{G}$-equivariant GNN is also distinguishable by GWL.
With a sufficient number of iterations, the output of $\mathfrak{G}$-equivariant GNNs can be equivalent to GWL if certain conditions are met regarding the aggregate, update and readout functions.

Proposition 2 $\mathfrak{G}$-equivariant GNNs have the same expressive power as GWL if the following conditions hold: (1) The aggregation AGG is an injective, $\mathfrak{G}$-equivariant multiset function. (2) The scalar part of the update UPD_s is a $\mathfrak{G}$-orbit injective, $\mathfrak{G}$-invariant multiset function. (3) The vector part of the update UPD_v is an injective, $\mathfrak{G}$-equivariant multiset function. (4) The graph-level readout f is an injective multiset function. The AGG function denotes the aggregate function in GNNs. The UPD_s and UPD_v are respectively scalar and vector update functions in GNNs. Similar statement can be made for $\mathfrak{G}$-invariant GNN and IGWL

4 Experiments

Chaitanya et. al in [10] have suggested few synthetic experiments to characterise the expressive power of a gGNN w.r.t their depth and tensor order.

Experiment 1: Depth and non local geometric properties In this experiment, we check how well a particular gGNN passes geometric information beyond its local neighbourhood w.r.t the number of layers in it. Here we train $\mathfrak{G}$-equivariant and $\mathfrak{G}$-invariant GNNs on two k chains graph in each run. Each pair of k chain graphs contain $k+2$ nodes with k nodes arranged in a line and are different by the orientation of the two end points. In every run the number of layers are increased. Since the k chain graphs are $\left(\left\lfloor \frac{k}{2} \right\rfloor + 1\right)$ hop different so theoretically GWL requires only $\left(\left\lfloor \frac{k}{2} \right\rfloor + 1\right)$ iterations to distinguish them.[1]. The results are given in Table 1.

(k=4 chains) GNN Layer	Number of Layers [k/2]	[k/2]+1=3	[k/2]+2	[k/2]+3	[k/2]+4
IGWL	**50%**	**50%**	**50%**	**50%**	**50%**
SchNet	50.00±0.00	50.00±0.0	50.00±0.0	50.00±0.0	50.00±0.0
DimeNet	50.00±0.00	50.00±0.0	50.00±0.0	50.00±0.0	50.00±0.0
GWL	**50%**	**100%**	**100%**	**100%**	**100%**
E-GNN	50.00±0.00	50.00±0.0	50.00±0.0	50.00±0.0	100.00±0.0
GVP-GNN	50.00±0.00	100.00±0.0	100.00±0.0	100.00±0.0	100.00±0.0
TFN	50.00±0.00	50.00±0.0	50.00±0.0	80.0±24.5	85.0±22.9
MACE	50.00±0.00	90.0±20.0	90.0±20.0	95.0±15.0	95.0±15.0

Table 1: Results for experiment on depth

GNN Layer	Rotational Symmetry 2-fold	3-fold	5-fold	10-fold
E-GNN (L=1)	50.00±0.0	50.00±0.0	50.00±0.0	50.00±0.0
GVP-GNN (L=1)	50.00±0.0	50.00±0.0	50.00±0.0	50.00±0.0
TFN/MACE (L=1)	50.00±0.0	50.00±0.0	50.00±0.0	50.00±0.0
TFN/MACE (L=2)	100.00±0.0	50.00±0.0	50.00±0.0	50.00±0.0
TFN/MACE (L=3)	100.00±0.0	100.00±0.0	50.00±0.0	50.00±0.0
TFN/MACE (L=5)	100.00±0.0	100.00±0.0	100.00±0.0	50.00±0.0
TFN/MACE (L=10)	100.00±0.0	100.00±0.0	100.00±0.0	100.00±0.0

Table 2: Result for experiment on tensor order. L denotes the tensor order taken

Experiment 2: Higher order tensors and rotational symmetry Here we check how rotational symmetries interact with tensor order. An L-fold symmetric structure does not change when rotated by an angle $\frac{2\pi}{L}$ about a point in 2D and an axis in 3D. Two distinct rotated versions of each L-fold symmetric structure are taken and single layer $\mathfrak{G}$-equivariant GNNs are trained on them to classify them. The results are given in Table 2.

Dataset There is no use of dataset in our algorithms. We train each of our model on the said graphs and check how well they classify. The models which we use are given in the references: SchNet[15], DimeNet[6], E-GNN[14], GVP-GNN[9], TFN[3] and MACE[2].

GitHub link Github link for code.

Parameters The hyperparameters mentioned here are as per [1]. The scalar feature channels for SchNet, DimeNet and E-GNN are set to 128. The scalar/vector/tensor channels for GVP-GNN, TFN, MACE are set to 64. TFN and MACE use order $L = 2$ tensors by default. MACE uses local body order 4 by default. All models are trained using Adam optimiser for 100 epochs, with an initial learning rate 1e-4, which we reduce by a factor 0.9 and patience of 25 epochs when the performance plateaus. All results are averaged across 10 random seeds.

5 Plans

The GNN models tested here are all reference frame free GNN. We are aiming at including reference frame based GNN models like ComENet[17] and ClofNet[4] into the code and test their expressivity.

References

[1] Laszlo Babai and Ludik Kucera. Canonical labelling of graphs in linear average time. In *20th Annual Symposium on Foundations of Computer Science (sfcs 1979)*, pages 39–46, 1979.

[2] Ilyes Batatia, Dávid Péter Kovács, Gregor N. C. Simm, Christoph Ortner, and Gábor Csányi. Mace: Higher order equivariant message passing neural networks for fast and accurate force fields, 2022.

[3] Qian Chen, Xingjian Dong, Guowei Tu, Dong Wang, Baoxuan Zhao, and Zhike Peng. Tfn: An interpretable neural network with time-frequency transform embedded for intelligent fault diagnosis, 2022.

[4] Weitao Du, He Zhang, Yuanqi Du, Qi Meng, Wei Chen, Bin Shao, and Tie-Yan Liu. Se(3) equivariant graph neural networks with complete local frames, 2021.

[5] Nadav Dym and Haggai Maron. On the universality of rotation equivariant point cloud networks, 2020.

[6] Johannes Gasteiger, Janek Groß, and Stephan Günnemann. Directional message passing for molecular graphs, 2020.

[7] Martin Grohe. *Descriptive Complexity, Canonisation, and Definable Graph Structure Theory*. Cambridge University Press, August 2017.

[8] Bowen Jing, Stephan Eismann, Patricia Suriana, Raphael J. L. Townshend, and Ron Dror. Learning from protein structure with geometric vector perceptrons, 2020.

[9] Bowen Jing, Stephan Eismann, Patricia Suriana, Raphael J. L. Townshend, and Ron Dror. Learning from protein structure with geometric vector perceptrons, 2020.

[10] Chaitanya K. Joshi, Cristian Bodnar, Simon V. Mathis, Taco Cohen, and Pietro Liò. On the expressive power of geometric graph neural networks, 2023.

[11] Johannes Klicpera, Florian Becker, and Stephan Gunnemann. Gemnet: Universal directional graph neural networks for molecules. *ArXiv*, abs/2106.08903, 2021.

[12] Adrien Leman. The reduction of a graph to canonical form and the algebra which appears therein. 2018.

[13] Christopher Morris, Martin Ritzert, Matthias Fey, William L. Hamilton, Jan Eric Lenssen, Gaurav Rattan, and Martin Grohe. Weisfeiler and leman go neural: Higher-order graph neural networks, 2018.

[14] Victor Garcia Satorras, Emiel Hoogeboom, and Max Welling. E(n) equivariant graph neural networks, 2021.

[15] Kristof T. Schütt, Pieter-Jan Kindermans, Huziel E. Sauceda, Stefan Chmiela, Alexandre Tkatchenko, and Klaus-Robert Müller. Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. 2017.

[16] Soledad Villar, David W Hogg, Kate Storey-Fisher, Weichi Yao, and Ben Blum-Smith. Scalars are universal: Equivariant machine learning, structured like classical physics. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan, editors, *Advances in Neural Information Processing Systems*, volume 34, pages 28848–28863. Curran Associates, Inc., 2021.

[17] Limei Wang, Yi Liu, Yuchao Lin, Haoran Liu, and Shuiwang Ji. Comenet: Towards complete and efficient message passing for 3d molecular graphs, 2022.

[18] Qing Wang, Dillon Ze Chen, Asiri Wijesinghe, Shouheng Li, and Muhammad Farhan. $\mathscr{N}$-WL: A new hierarchy of expressivity for graph neural networks. In *The Eleventh International Conference on Learning Representations*, 2023.

[19] Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How powerful are graph neural networks?, 2018.

Akankshya Nayak Paper Check

ORIGINALITY REPORT

5%
SIMILARITY INDEX

PRIMARY SOURCES

1	www.researchgate.net Internet	67 words — 3%
2	Michael O. Albertson. "Automorphisms and Distinguishing Numbers ofGeometric Cliques", Discrete & Computational Geometry, 06/2008 Crossref	16 words — 1%
3	www.ncbi.nlm.nih.gov Internet	9 words — < 1%
4	www.freepatentsonline.com Internet	8 words — < 1%

EXCLUDE QUOTES	ON	EXCLUDE SOURCES	OFF
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	OFF