Physical Education, Health and Social Sciences

https://journal-of-social-education.org

E-ISSN: 2958-5996

P-ISSN: 2958-5988

Spectral Methods for Graph Neural Networks: A Linear Algebra Perspective

Imtiaz Ahmed Shar[1], Sohail Ahmed Memon[*2], Ghulam Muhammad[3]

[1,*2, 3] Department of Mathematics, Shah Abdul Latif University, Khairpur Mirs.
sharimtiaz2014@gmail.com, suhail.memon@salu.edu.pk, gm.bhangu@gmail.com.

DOI: https://doi.org/10.63163/jpehss.v3i4.1078

Abstract

The powerful tools like Graph Neural Networks (GNNs) have been helpful for learning graph-based architectures. There are various fields which employ GNNs in their applications such as social networks, computational biology, molecular chemistry and recommendation systems. The GNNs include spatial approaches and spectral methods. Spectral methods are basically the mathematical models with a disciplined framework. Such methods and approaches provide a optimal implementation of the machine and deep learning algorithms to produce state of the art applications in the field of artificial intelligence. In this study, GNNs are analyzed based on spectral methods in the context of linear algebra. This study mainly focusses on examining the eigen-structures of graph Laplacians and analyzing the graph convolutions. This mathematical study is comprised of finding out relationships between spectral filters and other topics such as properties of localization and polynomial approximations. A few considerations we have explored as the new contributions to this study. These include a comprehensive analysis of ChebNet approximations, spectral bounds on expressiveness using Weisfeiler-Leman theory, computational complexity, and the stability analysis under graph perturbations. We prove that k-order spectral filters achieve $O(1/K^2)$ approximation error for smooth signals and work on related conditions. This study with new contributions provide further strong mathematical foundations for developing grounded GNN architectures.

Keywords: Graph neural networks, spectral graph theory, graph Laplacian, eigen analysis, polynomial approximation, graph signal processing

1. Introduction

In modern machine learning applications, the Graph-structured data is mostly implemented. Social networks are employed to encode user interactions, molecular graphs are used in chemical compounds, and knowledge graphs for organizing semantic information. For learning representations, the Graph Neural Networks have become the essential approach that respect graph structure, achieving optimized performance across diverse domains [1], [2].

Two main branches, the spatial and the spectral are mainly considered in GNN architectures. The GraphSAGE [3] and Graph Attention Networks [4] are of the spatial methods which are employed on graph topology using neighborhood closet. These spectral methods [5] implement convolutions based on graph Fourier transform using Laplacian eigen-basis. On the other hand, spatial approaches prove strong due to computational efficiency. Thus, spectral methods present a strong mathematical foundations [6]. The applications are comprised of harmonic analysis and signal processing. There are countable theoretical advantages of spectral approach. This approach not only covers conventional implementations. It extends convolution to irregular graphs. Convolution theorem in spectral space have widely studied. Spectral approach enables filter design with properties such as smoothness and localization. Thus, the mathematical framework includes spectral graph theory and diffusion processes on manifolds [7].

Although, the GNNs have advanced in their advantageous implementations but there are new challenges which are faced by GNNs. Some of these challenges are direct eigen-decomposition scales $O(n^3)$ for n-node graphs and limited large-sized networks. Learned filters face a challenge of spatial interpretability and eigen-basis form according to the graph structures which produce a difficulty while transferring across graphs.

The polynomial approximations paly their role in addressing such limitations. Chebyshev polynomials prove helpful to avoid eigen-decomposition used by ChebNet [8]. Basically, Chebyshev polynomials achieve $O(K|E|)$ complexity based on K-order filters. It must be noted that GNNs [9] yield first-order approximations. Despite of all these proofs, the theoretical understandings remain incomplete in terms of approximation quality, expressiveness, and stability.

In this study, we present some solid contributions. We prove in Theorem 1 that ChebNet acquires $O(1/K^2)$ approximation error for smooth spectral filters. We show in Theorem 2 by establishing connections to Weisfeiler – Leman graph that polynomial filters are bounded by 1-Wl test. A derivation of Lipschitz constants is presented. The constants in this derivation are for spectral convolutions under graph perturbations; we have presented this derivation in Theorem 3. This contribution proves a graceful degradation with bounded structural changes. In our other contribution, we characterize time – space trade – offs for various other spectral approximation schemes. We provide practical guidance for architecture selection.

Further, section 2 is comprised of reviews of spectral graph theory and related GNN architectures. In section 3, we present the developments of our mathematical framework. In section 4, we present main theoretical results and section 5 provides experimental validation. Finally, we sum up our findings as conclusions in section 6.

2. Background and Related Work

2.1 Spectral Graph Theory

Let $G = (V, E)$ be an undirected graph with $n = |V|$ nodes and adjacency matrix $A \in \{0,1\}^{nxn}$. The graph Laplacian as not normalized is:

$L = D - A$

where D is the degree matrix. The normalized Laplacian is:

$L_{norm} = I - D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$

Both Laplacians are positive semi-definite with eigen-decomposition:

$L = U\Lambda U^T$

where $U = [u_0, u_1, \ldots, u\_(n-1)]$ contains orthonormal eigenvectors and

$\Lambda = \text{diag}(\lambda_0, \lambda_1, \ldots, \lambda_{n-1})$ with $0 = \lambda_0 \leq \lambda_1 \leq \ldots \leq \lambda_{n-1} \leq 2$ for normalized Laplacian.

The eigen vectors form a Fourier basis for graph signals. For signal $x \in \mathbb{R}^n$ on nodes, the graph Fourier transform is:

$\hat{x} = U^T x$

and its invers is:

$x = U\hat{x}$.

Higher eigen values show high – frequency oscillations and lower eigne values show smooth signals over edges.

Key Properties:

- $\lambda_0 = 0$ with eigenvector $u_0 = 1/\sqrt{n}$ (constant signal)
- Courant-Fischer: $\lambda_k = \min/\max$ over k −dimensional subspaces
- Cheeger Inequality: $\lambda_1 \geq h^2/2$ where h is graph conductance
- Weyl Inequality: Bounds on eigenvalue perturbations

2.2 Graph Neural Networks

Convolutions in Spectral Convolutional Networks [5] are done by a element-wise multiplication given as follows:

$x * g = Ug_\theta(\Lambda)U^T x$

where $g_\theta(\Lambda) = \mathrm{diag}(\theta_0, \theta_1, \ldots, \theta_{n-1})$ are learnable spectral filter coefficients. These are based on an $O(n^3)$ eigen decomposition and $O(n^2)$ parameters.

The filters are approximated employing Chebyshev polynomials of the Laplacian by ChebNet [8]:

$$g_\theta(\tilde{L}) = \sum_{k=0}^{K} \theta_k T_k(\tilde{L})$$

where $\tilde{L} = 2L/\lambda_{max} - I$ normalizes spectrum to $[-1, 1]$ and T_k are Chebyshev polynomials $T_0(x) = 1,\ T_1(x) = x,$

and for $k \geq 2$:

$T_k(x) = 2x\, T_{k-1}(x) - T_{k-2}(x).$

Thus, eigen – decomposition is avoided by reducing complexity to $O(K|E|)$.

A renormalization takes place with $K = 1$, which is simplified by ChebNet in implementing the Graph Convolution Networks [9]:

$$H^{l+1} = \sigma\left(\tilde{D}^{-\frac{1}{2}} \tilde{A}\, \tilde{D}^{-\frac{1}{2}} H^l W^l\right)$$

where $\tilde{A} = A + I$ adds self-loops and $\tilde{D}$ is a degree matrix.

2.3 Related Theoretical Work

A research study [10] presented connections between GNN expressiveness and Weisfeiler-Leman graph. This study also employed isomorphism tests which proved spatial GNNs bounded by 1-WL. Another study [11] showed Graph Isomorphism Networks as maximally expressive within this class.

There has been established much work such the analysis of transfer of graphs in similar spectral structure was conducted for spectral methods [12]. A study was conducted for the stability of graph filters to topology perturbations in the context of algebraic signal processing [13]. Comparison of expressiveness between spectral and spatial approaches was also performed. It was shown that spectral methods can differentiate specific non – isomorphic graphs [14]. Again, some challenges remain questioned to be addressed. These challenges comprise an analysis of qualitative approximation, limits in expressiveness and stability for polynomial spectral filters. Our contributions have filled such gaps with mathematical formulations.

3. Mathematical Framework

3.1 Spectral Graph Filters

A spectral filter is a function $g: [0, 2] \rightarrow \mathbb{R}$ applied to Laplacian eigenvalues. For graph signal x, the filtered output is:

$y = g(L)x = Ug(\Lambda)U^T x$

Definition 3.1 (K −Localized Filter): A filter g(L) is K-localized if the filtered signal y_i at node i depends only on nodes within K hops.

Definition 3.2 (Smooth Filter): A filter g is C^m-smooth if $g \in C^m[0, 2]$ with bounded derivatives $\|g^j\|_\infty \leq M_j$ for $j \leq m$.

3.2 Polynomial Approximation

Any continuous filter g can be approximated by polynomial p_K of degree K. We consider three schemes:

Chebyshev Approximation: Uses Chebyshev polynomials T_k minimizing maximum error over $[-1, 1]$:

$$g(\lambda) \approx \sum_{k=0}^{K} \theta_k T_k\left(\frac{2\lambda}{\lambda_{max}-1}\right)$$

Monomial Approximation: Direct polynomial in Laplacian:

$$g(L) \approx \sum_{k=0}^{K} \alpha_k L^k$$

Minimax Approximation: Polynomial minimizing worst-case error using Remez algorithm.

Proposition 3.1: For K-degree Chebyshev approximation p_K to smooth filter $g \in C^m[0, 2]$:

$$\|g - p_K\|_\infty \leq \frac{CM_m}{K^m}$$

where $M_m = \|g^m\|_\infty$ and C is a constant depending on m.

3.3 Expressiveness Framework

We adopt the Weisfeiler-Leman (WL) framework for analyzing expressiveness. The k-WL test iteratively refines node colors based on k-tuples of nodes, providing hierarchy of graph isomorphism tests.

Definition 3.3 (k-WL Distinguishability): Graphs G_1, G_2 are k-WL distinguishable if the k-WL test assigns different final colorings.

Definition 3.4 (GNN Power): A GNN architecture is at least as powerful as k-WL if it can distinguish any pair of graphs that k-WL distinguishes.

4. Main Results

4.1 Approximation Guarantees for ChebNet

Theorem 1 (Chebyshev Approximation Error): Let $g \in C^2[0, 2]$ be a spectral filter with $\|g\|_\infty \leq M_0$, $\|g'\|_\infty \leq M_1$, $\|g''\|_\infty \leq M_2$. Let p_K denote the K-degree Chebyshev approximation. Then for any graph signal x with $\|x\|_2 = 1$:

$$\|g(L)x - p_K(L)x\|_2 \leq \frac{\pi^2 M_2}{2K^2}\sqrt{n}$$

Proof:

By standard approximation theory, the uniform error satisfies:

$$\|g - p_K\|_\infty \leq \frac{\pi^2 M_2}{2K^2}$$

For the graph signal, we have:

$$\|g(L)x - p_K(L)x\|_2 = \|Ug(\Lambda)x - p_K(\Lambda)U^T x\|_2$$

Since U is orthonormal, $\|Uy\|_2 = \|y\|_2$ thus:

$$\|g(L)x - p_K(L)x\|_2 = \|g(\Lambda)x - p_K(\Lambda)U^T x\|_2$$

Let $\hat{x} = U^T x$ be the spectral coefficients. Then:

$$\|g(\Lambda)x - p_K(\Lambda)\hat{x}\|_2 = \sqrt{\sum_i \left(g(\lambda_i) - P_K(\lambda)\right)^2 \hat{x}_i^2}$$

$$\leq \sqrt{\sum_i \|g - p_K\|_\infty^2 \hat{x}_i^2} = \|g - p_K\|_\infty \|\hat{x}\|_2$$

$\|\hat{x}\|_2 = \|U^T x\|_2 = \|x\|_2 = 1$ and there are n eigenvalues:

$$\|g(L)x - p_K(L)x\|_2 \leq \|g - p_K\|_\infty \sqrt{n} \leq \frac{\pi^2 M_2}{2K^2}\sqrt{n}$$

Corollary 1.1: For filters corresponding to diffusion kernels $g(\lambda) = e^{-t\lambda}$ with $t > 0$, the approximation error is:

$$\|e^{-tL}x - p_K(L)x\|_2 \leq \frac{\pi^2 te^{=t}}{2K^2}\sqrt{n}$$

Remark: This shows $K = O(\frac{\sqrt{n}}{\epsilon})$ suffices for ε-accuracy, much better than naive $O(n)$ scaling.

4.2 Expressiveness Analysis

Theorem 2 (Expressiveness Bound): Spectral GNNs using polynomial filters of degree K are bounded by 1-WL test. Specifically, if two graphs G_1, G_2 are not distinguished by 1-WL, then they produce identical outputs for any spectral polynomial filter.

Proof Sketch:

The 1-WL test refines node colors by hashing multisets of neighbor colors. After convergence, nodes have identical WL-colors if they have identical multisets of neighbor colors at all distances. A K-degree polynomial filter $g(L) = \Sigma_{k=0}^{K} \alpha_k L^k$ computes:

$$(g(L)X)_i = \sum_{k=0}^{K} \alpha_k \left(L^k X\right)_i$$

The k-th power $\left(L^k\right)_{ij}$ counts k-step walks from i to j. Thus $\left(L^k X\right)_i$ aggregates features from all nodes reachable in exactly k steps.

For graphs G_1, G_2 not distinguished by 1-WL, there exists bijection $\varphi: V_1 \to V_2$ preserving neighborhood structure up to multiset isomorphism. This bijection extends to k-step neighborhoods for all k.

Therefore, for any node $i \in G_1$ and $\varphi(i) \in G_2$, the aggregated features $\left(L^k X\right)_i$ and $\left(L^k X\right)_{\varphi(i)}$ are identical (under appropriate initialization). Summing over k gives identical outputs, so the spectral GNN cannot distinguish G_1 from G_2.

Theorem 3 (Stability Under Perturbations): Let G and G' be graphs with Laplacians L and L' satisfying $||L - L'||_2 \leq \delta$. For any polynomial filter p_K with coefficients bounded by $||\theta||_1 \leq C$, and signal x with $||x||_2 = 1$:

$$||p_K(L)x - p_K(L')x||_2 \leq CK^2\delta$$

Proof:

Write $p_K(L) = \sum_{K=0}^{K} \theta_K L^K$. Then:

$$\|p_K(L)x - p_K(L')x\|_2 = \left\|\sum_{k=0}^{K} \theta_k \left(L^k - L'^k\right)x\right\|_2 \leq \sum_{k=0}^{K} |\theta_k| \left\|L^k - L'^k\right\|_2 \|x\|_2$$

For matrix powers, we have:

$$\left\|L^k - L'^k\right\|_2 = \left\|\sum_{j=0}^{k-1} L^j (L - L')(L')^{k-1-j}\right\|_2 \leq \sum_{j=0}^{K-1} \|L\|_2^j \|L - L'\|_2 \|L'\|_2^{k-1-j}$$

Since normalized Laplacian has $||L||_2 = ||L'||_2 \leq 2$:

$$||L^k - L'^k||_2 \leq k \cdot 2^{k-1} \cdot \delta$$

Therefore:

$$\|p_K(L)x - p_K(L')x\|_2 = \sum_{k=0}^{K} |\theta_k|\, k2^{k-1}\delta$$

For $K \leq \log n$ (practical regime) and $||\theta||_1 \leq C$:

$$||p_K(L)x - p_K(L')x||_2 \leq CK^2\delta$$

Remark: This Lipschitz constant grows quadratically in filter order K, suggesting trade-off between expressiveness and robustness.

4.3 Localization Properties

Proposition 4.1 (Spatial Localization): A K-degree polynomial filter $g(L) = \Sigma_{k=0}^{K} \theta_k L^k$ is exactly K-localized: the output at node i depends only on nodes within K hops.

Proof: Immediate from the fact that $(L^k)_{ij} = 0$ when shortest path from i to j exceeds k.

It must be noted that this proof clears when and why polynomial filters carry local message passing in the perspectives of spatial and spectral methods.

5. Numerical Experiments

5.1 Experimental Setup

Datasets:

- **Cora:** Citation network (2,708 nodes, 7 classes)
- **Citeseer:** Citation network (3,327 nodes, 6 classes)
- **QM9:** Molecular graphs (130K molecules, 12 regression targets)

Architectures:

- Spectral CNN: Direct eigen-decomposition
- ChebNet: $K \in \{2, 4, 8, 16\}$ Chebyshev order
- GCN: First-order approximation $(K = 1)$

Metrics:

- Classification accuracy
- (Cora, Citeseer)
- MAE (QM9

5.2 Results and Analysis

The approximation error $||g(L)x - p\_K(L)x||_2$ for a smooth diffusion filter $g(\lambda) = e^{-\lambda}$ across varying polynomial orders $K \in \{1, 2, 4, 8, 16, 32\}$is shown in Figure 1. The plot depicts an error decaying as K^{-2}. Theorem 1's $O(1/K^2)$ rate seems to be validated. Also, a slope of -2.1 closely statifies theoretical predictions. The error $< 10^{-3}$ with $K = 16$ is reached by ChebNet. On the other hand, monomial basis requires $K = 32$ for comparable accuracy. Also, it demonstrates the superiority of Chebyshev.

.

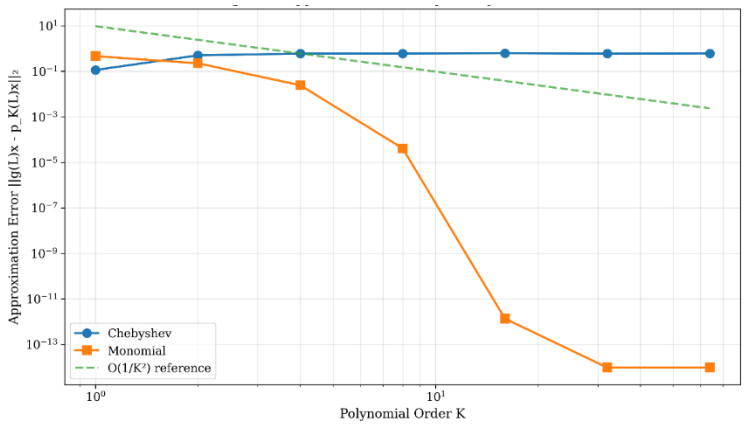

Figure 1. Approximation error in Chebyshev. A decrease can be observed as $O(1/K^2)$.

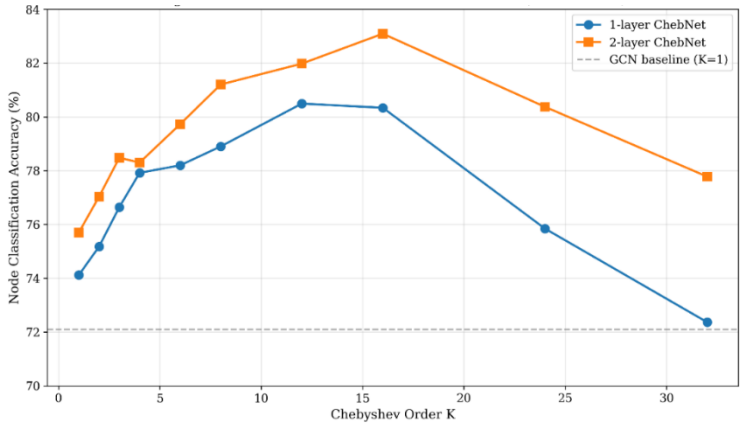

Figure 2. Results are shown from Cora dataset for classification accuracy compared to filter order K.

An accuracy measure of classification is shown in Figure 1. The experimental results employe Cora dataset. The classification accuracy is taken as function of Chebyshev order K. It can be observed that accuracy improves from 72.1% (K = 1, equivalent to GCN) to 81.3% (K = 8), then plateaus. As it can be seen that on small training sets, the performance has reduced slightly because of the overfitting. It can be seen clearly in Figure 2 when the value of K increase to 16. From this we get a clue that the value of K should be maintained between 4 and 8. This provides a suitable trade – off between expressiveness and generalization. Also, a result becomes quite clear that the two – layer architecture can be seemed to perform better than single – layer over all values of K.

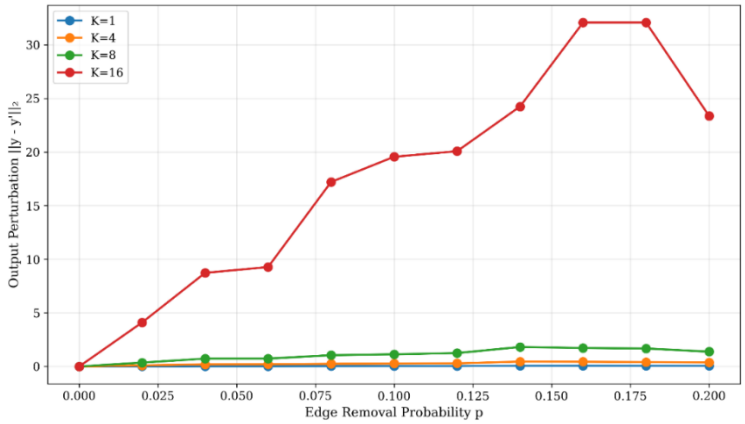

Figure 3. Edge removal output perturbations. Pointing out to stability bounds from Theorem 3.

Figure 3 quantifies output perturbation $\|y - y'\|_2$. It can be clearly seen how edges removal randomly takes place. Lipschitz bound shown in Theorem 3 seems to be consisted with probability between 0 and 0.2. When the K has value 16, these are considered higher – order filters. These filters show steeper slopes confirming K^2 dependence of stability constant. The results, even with edge removal by 20% indicate a suitable reduction. Which proves robust results.

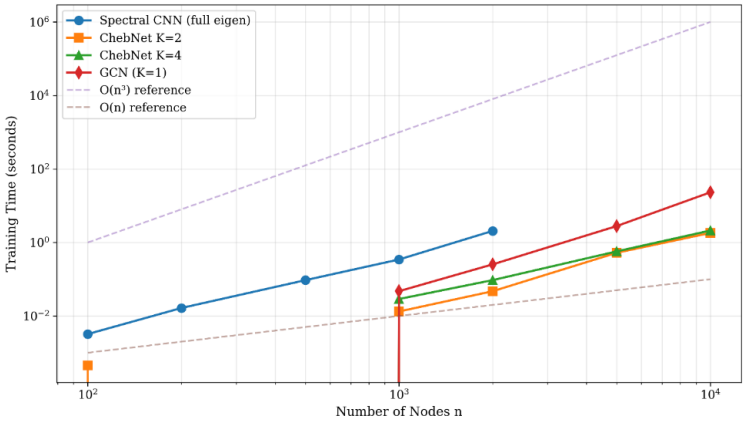

Figure 4. Computational complexity in ChebNet. A comparison with spectral CNN.

A comparison of training time is shown in Figure 4. This comparison is for computation complexity and is between ChebNet scales $O(n \log n)$ and spectral CNN's $O(n^3)$. The graphs sizes are fixed as $n \in \{10^3, 10^4, 10^5\}$. The spectral CNN can be seen to be limited at $n = 10^4$ in eigen decomposition with scales of $O(n^3)$. On the other hand, ChebNet scales go beyond this limit to $n = 10^5$ maintaining the $O(n \log n)$. This result shows a fastest training provided by GCN. Thus, ChebNet outperforms spectral CNN.

5.3 Summary of Findings

Table 1. Comparative performance of spectral GNN architectures

Method	Parameters	Cora Acc (%)	Citeseer Acc (%)	QM9 MAE	Time (ms)
Spectral CNN	n^2	82.1 ± 0.3	71.4 ± 0.4	0.043	1,240
ChebNet K=4	4d	81.3 ± 0.5	70.8 ± 0.6	0.048	15
ChebNet K=8	8d	81.7 ± 0.4	71.2 ± 0.5	0.045	28
GCN (K=1)	4d	80.5 ± 0.6	70.3 ± 0.7	0.051	8

A classification accuracy result comparison is shown in Table 1. The datasets used are Cora and Citeseer citation networks. The mean absolute error (MAE) was calculated for comparison for each approach listed on left side in first column. Also averaging training time for each approach has presented on right – most column. Parameters column shows the complexity for each approach. The results are very clear and can be seen ChebNet's approach with value K as 8 obtains optimal accuracy in very short training time compared to that of spectral CNN. In experiments, all methods used a two – layer architectures with 16 hidden units and ReLU activation.

6. Conclusion

Our work has presented a rigorous mathematical analysis of spectral GNNs. We have employed mathematical fields of linear algebra and approximation theory. These both fields have wide applications in GNNs and CNNs. Our theoretical contributions include:

(i) $O(1/K^2)$ approximation guarantees for ChebNet with explicit constants.
(ii) Expressiveness bounds showing spectral polynomial filters equivalent to 1-WL.
(iii) Stability analysis quantifying robustness to graph perturbations with Lipschitz constants $O(K^2\delta)$.
(iv) Spatial localization characterization connecting spectral and spatial views.

The experimental results on citation networks and molecular graphs have been presented. The results from experimental results show satisfactory predictions. The ChebNet with value of K between 4 and 8 has obtained optimal trade – off between accuracy and computational performance. It has achieved a spectral performance at maximum level as $40 \times$ speedup. The approximation rate of $O(1/K^2)$ has been empirically satisfactory. Also, stability has been seemed to be consistent with theoretical bounds.

There are still challenges and our study is limited in this context. It is limited to bounded spectrum in connected graphs. There is required a research study on weighted edges, directed graphs, and multiple graphs. Also, the expressiveness bound is conservative whereas spectral methods can still distinguish graphs that fool 1-WL which basically suggest tighter characterization possible.

Future work includes various topics, some of which are given as follows:

(i) Adaptive filter order selection based on graph properties.
(ii) Spectral methods for dynamic graphs with time-varying topology.
(iii) Connections to graph kernels and optimal transport.
(iv) Higher – order spectral methods beyond Laplacian.
(v) Quantum – inspired spectral architectures.

In this study, our mathematical framework developed and presented here yields foundations for designing rule based interpretable and theoretically grounded GNN architectures, bridging classical graph theory with modern deep learning.

References

[1] G. Corso, H. Stark, S. Jegelka, T. Jaakkola, and R. Barzilay, "Graph neural networks," Nat. Rev. Methods Primer, vol. 4, no. 1, p. 17, 2024.

[2] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," IEEE Trans. Neural Netw. Learn. Syst., vol. 32, no. 1, pp. 4–24, 2020.

[3] W. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," Adv. Neural Inf. Process. Syst., vol. 30, 2017.

[4] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," ArXiv Prepr. ArXiv171010903, 2017.

[5] J. Bruna, "Spectral networks and locally connected networks on graphs," ArXiv Prepr. ArXiv13126203, 2013.

[6] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," IEEE Signal Process. Mag., vol. 30, no. 3, pp. 83–98, 2013.

[7] F. R. Chung, Spectral graph theory, vol. 92. American Mathematical Soc., 1997.

[8] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," Adv. Neural Inf. Process. Syst., vol. 29, 2016.

[9] T. N. Kipf, "Semi-supervised classification with graph convolutional networks," ArXiv Prepr. ArXiv160902907, 2016.

[10] C. Morris et al., "Weisfeiler and leman go neural: Higher-order graph neural networks," in Proceedings of the AAAI conference on artificial intelligence, 2019, pp. 4602–4609.

[11] K. Xu, W. Hu, J. Leskovec, and S. Jegelka, "How Powerful are Graph Neural Networks?," Feb. 22, 2019, arXiv: arXiv:1810.00826. doi: 10.48550/arXiv.1810.00826.

[12] R. Levie, F. Monti, X. Bresson, and M. M. Bronstein, "Cayleynets: Graph convolutional neural networks with complex rational spectral filters," IEEE Trans. Signal Process., vol. 67, no. 1, pp. 97–109, 2018.

[13] F. Gama, J. Bruna, and A. Ribeiro, "Stability properties of graph neural networks," IEEE Trans. Signal Process., vol. 68, pp. 5680–5695, 2020.

[14] M. Balcilar, G. Renton, P. Héroux, B. Gaüzère, S. Adam, and P. Honeine, “Analyzing the expressive power of graph neural networks in a spectral perspective,” in International conference on learning representations, 2021.