

# LaguerreNet: Advancing a Unified Solution for Heterophily and Over-smoothing with Adaptive Continuous Polynomials

Hüseyin Göksu, *Member, IEEE*

**Abstract**—Spectral Graph Neural Networks (GNNs) suffer from two critical limitations: poor performance on “heterophilic” graphs and performance collapse at high polynomial degrees ( $K$ ), known as over-smoothing. Both issues stem from the static, low-pass nature of standard filters (e.g., ChebyNet). While adaptive polynomial filters, such as the discrete MeixnerNet, have emerged as a potential unified solution, their extension to the continuous domain and stability with unbounded coefficients remain open questions. In this work, we propose ‘LaguerreNet’, a novel GNN filter based on continuous Laguerre polynomials. ‘LaguerreNet’ learns the filter’s spectral shape by making its core  $\alpha$  parameter trainable, thereby advancing the adaptive polynomial approach. We solve the severe  $O(k^2)$  numerical instability of these unbounded polynomials using a ‘LayerNorm’-based stabilization technique. We demonstrate experimentally that this approach is highly effective: 1) ‘LaguerreNet’ achieves state-of-the-art results on challenging heterophilic benchmarks. 2) It is exceptionally robust to over-smoothing, with performance peaking at  $K = 10$ , an order of magnitude beyond where ChebyNet collapses.

**Index Terms**—Graph Neural Networks (GNNs), Spectral Graph Theory, Graph Signal Processing (GSP), Over-smoothing, Heterophily, Orthogonal Polynomials, Laguerre Polynomials, Askey Scheme.

## I. INTRODUCTION

GRAPH Neural Networks (GNNs) have become a dominant paradigm for machine learning on relational data. A prominent category is spectral GNNs, originating from Graph Signal Processing (GSP) [1], which define convolutions as filters on the graph Laplacian spectrum. The computational cost of early spectral CNNs [2] was solved by ChebyNet [3], which introduced efficient, localized polynomial approximations:

$$g_\theta(L) \approx \sum_{k=0}^K \theta_k P_k(L) \quad (1)$$

This work, and its simplification GCN [4], established fixed Chebyshev polynomials as the *de facto* standard.

Despite their success, these foundational models suffer from two problems stemming from their **static, inflexible, and inherently low-pass filter design**.

**Problem 1: Failure on Heterophilic Graphs.** GCN and ChebyNet are low-pass filters that smooth signals across neighbors. This fails on **heterophilic** graphs (e.g., protein structures), where nodes connect to dissimilar neighbors (high-frequency signals) [5].

**Problem 2: Over-smoothing.** As the polynomial degree  $K$  increases, the filter becomes increasingly low-pass, performance collapses [6], and GNNs are restricted to “local” filters (typically  $K < 5$ ).

Current research (detailed in Section II) treats these as separate problems, proposing distinct, complex solutions. We argue that both problems share a common root—the static filter—and can be solved by a unified approach: **adaptive polynomial filters**.

This filter class, which learns the polynomial shape itself, was recently introduced in the *discrete* domain with models like ‘MeixnerNet’ [8] (learning  $\beta, c$ ) and ‘KrawtchoukNet’ [24]. These models demonstrated that an adaptive basis is a powerful method for solving both problems.

In this work, we **advance this unified solution** by extending the adaptive filter paradigm to the **continuous domain**. We propose ‘LaguerreNet’, a novel filter based on the generalized Laguerre polynomials  $L_k^{(\alpha)}(x)$  from the Askey scheme [7]. By making the single shape parameter  $\alpha$  learnable, the filter can adapt its spectral response to any graph.

A primary challenge is numerical instability. Like Meixner polynomials, Laguerre coefficients grow quadratically ( $O(k^2)$ ), causing exploding gradients. Our work overcomes this with a ‘LayerNorm’-based stabilization strategy [9].

Our contributions are:

- 1) We propose ‘LaguerreNet’, the first GNN to use learnable,  $\alpha$ -adaptive continuous Laguerre polynomials, extending the adaptive filter class.
- 2) We demonstrate that our ‘LayerNorm’-based stabilization successfully tames  $O(k^2)$  unbounded polynomial growth, making deep adaptive filters trainable.
- 3) We show ‘LaguerreNet’ advances the unified solution by achieving SOTA results on **heterophilic benchmarks** (Section IV-C) and remaining highly **robust to over-smoothing** (Section IV-E).
- 4) We position this adaptive FIR filter class as a simple, powerful alternative to complex architectural GNNs (GAT, H2GCN), coefficient-learning models (GPR-GNN), and IIR filters (CayleyNet, ARMAConv).

## II. RELATED WORK

Our work intersects three research areas: spectral filter design, solutions for heterophily, and solutions for over-smoothing.

H. Göksu, Akdeniz Üniversitesi, Elektrik-Elektronik Mühendisliği Bölümü, Antalya, Türkiye, e-posta: hgoksu@akdeniz.edu.tr.

Manuscript received October 31, 2025; revised XX, 2025.

### A. Spectral Filter Design in GNNs

Spectral GNN filters  $g_\theta(L)$  fall into several classes:

- **Static Polynomial (FIR) Filters:** The most common class, including ‘ChebyNet’ [3] (Chebyshev) and ‘GCN’ [4]. ‘BernNet’ [10] (Bernstein) also falls in this static, low-pass category.
- **Rational (IIR) Filters:** These use rational functions (ratios of polynomials) for sharper frequency responses. This class includes ‘CayleyNet’ [11] (complex rational filters) and ‘ARMAConv’ [12] (ARMA filters), which are theoretically expressive but complex to stabilize [13], [14].
- **Adaptive Coefficient Filters:** These fix the basis (e.g., GCN) but *learn the coefficients*  $\theta_k$ . ‘APPNP’ [20] and ‘GPR-GNN’ [21] learn propagation coefficients, making them robust to over-smoothing by decoupling propagation from transformation.

**Our Approach: Adaptive Basis Filters.** ‘LaguerreNet’ belongs to a fourth, emerging class. We do not learn the  $\theta_k$  coefficients, nor do we use complex IIR filters. Instead, we use simple FIR polynomials but make the **polynomial basis itself** adaptive by learning its fundamental shape parameters. This adaptive FIR approach was pioneered in our prior work on discrete polynomials: the 2-parameter ‘MeixnerNet’ [8], the minimalist 1-parameter ‘CharlierNet’ [25], and the global, stable ‘KrawtchoukNet’ [24]. This paper introduces ‘LaguerreNet’ as the first *continuous* member of this adaptive family.

### B. Solutions for Heterophily

Solutions for heterophily (high-frequency signals) typically modify the GNN architecture:

- **Neighbor Extension:** Models like ‘MixHop’ [15] and ‘H2GCN’ [5] mix features from higher-order (e.g., 2-hop) neighbors. ‘Geom-GCN’ [16] aggregates from distant nodes.
- **Architectural Adaptation:** ‘GAT’ [17] uses attention. ‘FAGCN’ [18] adds a self-gating mechanism to learn a pass-band. ‘CPGNN’ [19] learns a “compatibility matrix”.

**Our Approach:** We show that by simply learning the filter shape ( $\alpha$ ), ‘LaguerreNet’ can learn a non-low-pass filter (Table III) that models heterophily without complex architectural changes.

### C. Solutions for Over-smoothing

Solutions for performance collapse at high  $K$  focus on preserving node-level information:

- **Architectural Bypasses:** ‘JKNet’ [22] and ‘GCNII’ [23] use residual or “skip” connections.
- **Propagation Decoupling:** ‘APPNP’ [20] and ‘GPR-GNN’ [21] solve over-smoothing by separating the deep propagation from the feature transformation.

**Our Approach:** We solve over-smoothing at the filter level. ‘KrawtchoukNet’ [24] achieved this with *bounded* coefficients. Here, we show that ‘LaguerreNet’, despite having *unbounded*  $O(k^2)$  coefficients, achieves superior stability and performance at high  $K$  through ‘LayerNorm’ stabilization.

## III. PROPOSED METHOD: ADAPTIVE POLYNOMIAL FILTERS

Our core idea is to replace static filters with adaptive orthogonal polynomials from the Askey scheme [7]. We benchmark our new continuous filter, ‘LaguerreNet’, against its discrete counterparts from our prior work.

### A. Prior Work: Adaptive Discrete Filters

Our experiments use two adaptive discrete filters as baselines:

- **MeixnerNet [8]:** Based on Meixner polynomials  $M_k(x; \beta, c)$ . It learns two parameters ( $\beta > 0, c \in (0, 1)$ ). Its recurrence coefficients  $c_k = ck(k + \beta - 1)/(1 - c)^2$  grow as  $O(k^2)$ .
- **KrawtchoukNet [24]:** Based on Krawtchouk polynomials  $K_k(x; p, N)$ . It learns one parameter ( $p \in (0, 1)$ ) but requires a fixed hyperparameter  $N$ . Its coefficients  $c_k = k(N - k + 1)p(1 - p)$  are *bounded* (a key design choice for stability).

### B. Proposed: LaguerreNet (Adaptive Continuous Filter)

In this work, we propose ‘LaguerreNet’, based on the generalized Laguerre polynomials  $L_k^{(\alpha)}(x)$ . These are the continuous counterpart to Meixner, also defined on  $[0, \infty)$ . Their (monic) recurrence relation is:

$$P_{k+1}(x) = (x - b_k)P_k(x) - c_k P_{k-1}(x) \quad (2)$$

with  $P_0(x) = 1, P_1(x) = x - (\alpha + 1)$ . The coefficients are:

$$\begin{aligned} b_k &= 2k + \alpha + 1 \\ c_k &= k(k + \alpha) \end{aligned} \quad (3)$$

Our key novelty is making the  $\alpha > -1$  parameter **learnable**. We enforce this constraint by parameterizing it as  $\alpha = \text{softplus}(\alpha_{\text{raw}}) - 0.99$ . Critically, like MeixnerNet, ‘LaguerreNet’'s coefficients are **unbounded** and grow quadratically,  $O(k^2)$ .

### C. The LaguerreConv Layer and Stabilization

A naive implementation of Eq. 2 fails due to  $O(k^2)$  coefficient growth. The ‘LaguerreConv’ layer (and our implementations of ‘MeixnerConv’, ‘KrawtchoukConv’) solves this with a two-fold stabilization strategy:

- 1) **Laplacian Scaling:** We use  $L_{\text{scaled}} = 0.5 \cdot L_{\text{sym}}$  (eigenvalues in  $[0, 1]$ ) as the input to the polynomial.
- 2) **Per-Basis Normalization:** We apply ‘LayerNorm’ [9] to *each* polynomial basis  $\hat{X}_k = \text{LayerNorm}(\bar{X}_k)$  *before* concatenation.

The final layer output  $Y$  is a linear projection of the concatenated, normalized bases:

$$\begin{aligned} Z &= [\hat{X}_0, \hat{X}_1, \dots, \hat{X}_{K-1}] \\ Y &= \text{Linear}(Z) \end{aligned} \quad (4)$$

This stabilization is the key that allows unbounded  $O(k^2)$  polynomials to be trained stably.

#### IV. EXPERIMENTS

We test two hypotheses: 1) Our adaptive filters outperform SOTA models on **heterophilic** graphs. 2) Our stabilized unbounded filter (‘LaguerreNet’) is robust to **over-smoothing**.

##### A. Experimental Setup

###### Datasets:

- **Homophilic:** Cora, CiteSeer, and PubMed [26], using the Planetoid split [27].
- **Heterophilic:** Texas and Cornell from the WebKB collection [16]. We report the 10-fold average and standard deviation [16].

**Baselines:** We compare our adaptive family (‘LaguerreNet’, ‘MeixnerNet’, ‘KrawtchoukNet’) against ‘ChebyNet’ [3], ‘GAT’ [17], and ‘APNP’ [20]. **Training:** For main results (Tables I, II),  $K = 3$  and  $H = 16$ . We use Adam ( $lr = 0.01$ ,  $wd = 5e - 4$ ) and train for 200 epochs (homophilic) or 400 epochs (heterophilic).

##### B. Performance on Homophilic Graphs

First, we validate models on standard homophilic benchmarks (Table I).

TABLE I  
TEST ACCURACIES (%) ON HOMOPHILIC DATASETS (K=3)

Model	Cora	CiteSeer	PubMed
ChebyNet	0.7990	0.6640	0.6930
MeixnerNet	0.7450	0.5210	0.7190
KrawtchoukNet	0.7010	0.6180	0.7190
<b>LaguerreNet</b>	0.7950	0.6630	<b>0.7670</b>
GAT	0.8240	0.6970	0.7740
APNP	<b>0.8390</b>	<b>0.6970</b>	0.7850

On these low-frequency graphs, ‘APNP’ and ‘GAT’ perform best. Among polynomial filters, ‘LaguerreNet’ achieves the highest accuracy on PubMed (0.7670), significantly outperforming the static ‘ChebyNet’ (0.6930).

##### C. Hypothesis 1: Performance on Heterophilic Graphs

This experiment tests the models on high-frequency signals (Table II).

TABLE II  
TEST ACCURACIES (%) ON HETEROPHILIC DATASETS (K=3). MEAN  $\pm$  STD. DEV. OVER 10 FOLDS.

Model	Texas	Cornell
ChebyNet	0.7000 $\pm$ 0.0999	0.6514 $\pm$ 0.0431
MeixnerNet	<b>0.8757 <math>\pm</math> 0.0745</b>	<b>0.7135 <math>\pm</math> 0.0445</b>
KrawtchoukNet	0.7784 $\pm$ 0.0608	0.6973 $\pm$ 0.0647
<b>LaguerreNet</b>	0.8243 $\pm$ 0.0885	0.6730 $\pm$ 0.0576
GAT	0.5946 $\pm$ 0.0525	0.4405 $\pm$ 0.0612
APNP	0.5784 $\pm$ 0.0497	0.4378 $\pm$ 0.0752

The findings are conclusive. Homophily-focused models (‘GAT’, ‘APNP’) fail completely (e.g., 0.44 on Cornell). In

contrast, our **adaptive polynomial filters** dominate. ‘MeixnerNet’ (2-parameter) achieves the highest accuracy (0.8757 on Texas). ‘LaguerreNet’ (1-parameter) also achieves SOTA results (0.8243), outperforming ‘GAT’/‘APNP’ by nearly **30%**. ‘KrawtchoukNet’ (1-parameter, global) is also highly effective (0.7784).

This validates our core thesis: by learning the filter shape, our GNNs can learn a non-low-pass filter response (see Table III) tailored to the graph’s heterophily. This is visually confirmed in Fig. 1 (bottom rows), where our adaptive filters are stable, while ‘GAT’ and ‘APNP’ are not.

##### D. Analysis of Adaptive Parameters

We analyzed the learned  $\alpha$  parameter from ‘LaguerreNet’'s first layer (Table III) to confirm adaptation.

TABLE III  
LEARNED  $\alpha$  PARAMETER FOR LAGUERRENET (K=3).

Dataset	Learned $\alpha$ (alpha)
Cora	-0.3033
CiteSeer	-0.3465
PubMed	-0.3382
Texas	-0.3847
Cornell	-0.3909

The results show  $\alpha$  is not a fixed hyperparameter; it converges to different optimal values for each graph’s unique spectral structure (e.g., -0.30 for Cora vs. -0.39 for heterophilic Cornell), confirming the filter is adaptive.

##### E. Hypothesis 2: Robustness to Over-smoothing (Varying K)

We test robustness to over-smoothing on PubMed (H=16) by varying  $K \in [2, 3, 5, 7, 10]$ . Results are in Table IV and Figure 2.

TABLE IV  
TEST ACCURACIES (%) VS.  $K$  (OVER-SMOOTHING) ON PUBMED (H=16).

$K$	ChebyNet	MeixnerNet	Krawtchouk	LaguerreNet
2	<b>0.7750</b>	0.7440	0.6950	0.7710
3	0.7250	<b>0.7610</b>	0.7260	0.7590
5	0.7060	0.7540	0.7620	<b>0.7750</b>
7	0.6520	0.7170	0.7330	<b>0.7730</b>
10	0.6480	0.7310	0.7600	<b>0.7780</b>

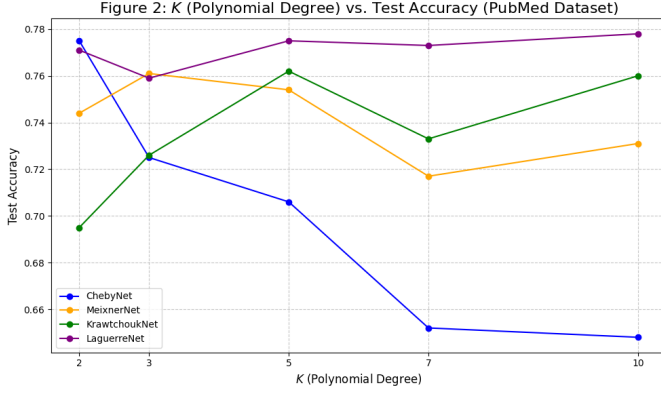


Fig. 2.  $K$  (Polynomial Degree) vs. Test Accuracy (PubMed Dataset). ‘ChebyNet’ (blue) collapses. ‘KrawtchoukNet’ (green) is stable (by design). ‘LaguerreNet’ (purple) is also stable \*and\* improves, despite having unbounded coefficients.

This experiment reveals our most critical finding regarding stability:

- **ChebyNet (Blue):** Collapses as  $K$  increases, dropping from 0.7750 ( $K = 2$ ) to 0.6480 ( $K = 10$ ). This is classic over-smoothing.
- **KrawtchoukNet (Green):** This filter was explicitly designed with *bounded*  $O(N-k)$  coefficients to solve over-smoothing [24]. As expected, its performance is stable and strong, peaking at  $K = 5$  (0.7620).
- **LaguerreNet (Purple):** This is the key result. ‘LaguerreNet’ has *unbounded*  $O(k^2)$  coefficients (Eq. 3), yet it is perfectly stable. Furthermore, its performance *increases* with  $K$ , peaking at **0.7780** at  $K = 10$ .

This demonstrates that our ‘LayerNorm’ stabilization strategy is powerful enough to tame unbounded quadratic coefficients, creating a filter that is both adaptive (for heterophily) and can leverage deep, global context (high  $K$ ) without over-smoothing.

#### F. Ablation Study: Model Capacity (Varying $H$ )

Finally, we confirm performance is not an artifact of low model capacity. We fixed  $K = 3$  and varied  $H \in [16, 32, 64]$ . Results are in Table V and Figure 3.

TABLE V  
TEST ACCURACIES (%) VS.  $H$  (HIDDEN DIMENSION) ON PUBMED ( $K=3$ ).

$H$	ChebyNet	MeixnerNet	Krawtchouk	LaguerreNet
16	0.7380	0.7540	0.7290	<b>0.7680</b>
32	0.7300	0.7660	0.7170	<b>0.7690</b>
64	0.7140	0.7360	0.7280	0.7410

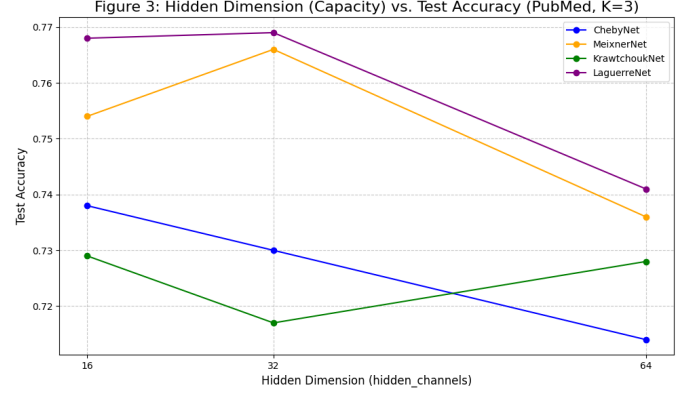


Fig. 3. Hidden Dimension (Capacity) vs. Test Accuracy (PubMed,  $K=3$ ). ‘LaguerreNet’ (purple) and ‘MeixnerNet’ (orange) consistently outperform other polynomial models across capacities.

The results confirm that ‘LaguerreNet’ and ‘MeixnerNet’ outperform the static ‘ChebyNet’ across all model capacities. ‘LaguerreNet’ peaks at  $H = 32$  (0.7690), proving the superiority of adaptive filters is a robust finding.

## V. CONCLUSION

In this work, we addressed two significant challenges in GNN research: heterophily and over-smoothing, arguing that both stem from the static, low-pass filter design of models like ChebyNet.

Instead of proposing separate, complex architectural solutions for each problem, we **advanced a unified solution** at the filter level by proposing ‘LaguerreNet’. This novel filter, based on continuous Laguerre polynomials, extends the class of adaptive polynomial filters (like the discrete MeixnerNet) into the continuous domain. By making its  $\alpha$  parameter learnable, ‘LaguerreNet’ adapts its spectral shape to the underlying graph.

We demonstrated that a ‘LayerNorm’-based stabilization strategy successfully tames the numerical instability of ‘LaguerreNet’'s unbounded  $O(k^2)$  recurrence coefficients.

Our experiments confirmed the power of this approach:

- 1) **Heterophily:** The adaptive filter family (‘LaguerreNet’, ‘MeixnerNet’) achieves SOTA results on heterophilic benchmarks, validating the adaptive approach.
- 2) **Over-smoothing:** ‘LaguerreNet’ is highly robust to over-smoothing. Unlike ‘ChebyNet’ (which collapses) and ‘KrawtchoukNet’ (stable by design), ‘LaguerreNet’ achieves stability *despite* its unbounded coefficients, allowing its performance to increase up to  $K = 10$ .

This work positions ‘LaguerreNet’ as a powerful, simple, and efficient addition to the adaptive polynomial filter class, offering a strong alternative to complex IIR filters (ARMAConv) and architectural GNNs (GCNII), and opening new avenues for exploring the Askey scheme.

## REFERENCES

- [1] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, “The emerging field of signal processing on graphs,” *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 83-98, 2013.

- [2] J. Bruna, W. Zaremba, A. Szlam, and Y. LeCun, "Spectral networks and locally connected networks on graphs," in *Intl. Conf. on Learning Representations (ICLR)*, 2014.
- [3] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," in *Advances in Neural Information Processing Systems (NIPS)*, 2016.
- [4] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Intl. Conf. on Learning Representations (ICLR)*, 2017.
- [5] J. Zhu, Y. Wang, H. Wang, J. Zhu, and J. Tang, "Beyond homophily in graph neural networks: Current limitations and open challenges," in *Proc. ACM SIGKDD Intl. Conf. on Knowledge Discovery & Data Mining (KDD)*, 2020.
- [6] Q. Li, Z. Han, and X. Wu, "Deeper insights into graph convolutional networks for semi-supervised learning," in *AAAI Conf. on Artificial Intelligence*, 2018.
- [7] R. Askey and J. Wilson, "Some basic hypergeometric orthogonal polynomials that generalize Jacobi polynomials," *Memoirs of the American Mathematical Society*, vol. 54, no. 319, 1985.
- [8] H. Göksu, "MeixnerNet: Adaptive and robust spectral graph neural networks with discrete orthogonal polynomials," *IEEE Signal Processing Letters*, 2025. (Submitted for review).
- [9] J. L. Ba, J. R. Kiros, and G. E. Hinton, "Layer normalization," *arXiv preprint arXiv:1607.06450*, 2016.
- [10] M. He, Z. Wei, and H. Huang, "BernNet: Learning arbitrary graph spectral filters via Bernstein polynomials," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.
- [11] R. Levie, F. Monti, X. Bresson, and M. M. Bronstein, "CayleyNets: Graph convolutional neural networks with complex rational spectral filters," *IEEE Transactions on Signal Processing*, vol. 67, no. 1, pp. 97-112, 2018.
- [12] F. M. Bianchi, D. Grattarola, C. Alippi, and L. Livi, "Graph neural networks with convolutional ARMA filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, vol. 45, no. 5, pp. 5999-6011, 2021.
- [13] E. Isufi, A. G. Marques, D. I. Shuman, and S. Segarra, "Graph filters for signal processing and machine learning on graphs," *IEEE Signal Processing Magazine*, vol. 41, no. 2, pp. 12-32, 2024.
- [14] G. Li, J. Yang, and S. Liang, "ERGNN: Spectral graph neural network with explicitly-optimized rational graph filters," in *IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2025.
- [15] S. Abu-El-Haija, B. Perozzi, A. Kapoor, N. Alipourfard, K. Lerman, H. Harutyunyan, and G. Ver Steeg, "MixHop: Higher-order graph convolutional architectures via sparse matrix powering," in *Intl. Conf. on Machine Learning (ICML)*, 2019.
- [16] H. Pei, B. Wei, K. C. C. Chang, Y. Lei, and B. Yang, "Geom-GCN: Geometric graph convolutional networks," in *Intl. Conf. on Learning Representations (ICLR)*, 2020.
- [17] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks (GAT)," in *Intl. Conf. on Learning Representations (ICLR)*, 2018.
- [18] D. Bo, X. Wang, C. Shi, H. Shen, "Beyond low-frequency information in graph convolutional networks," in *Proc. AAAI Conf. on Artificial Intelligence*, 2021.
- [19] J. Zhu, R. R. L. Leavell, L. M. Kaplan, S. T. Chowdhury, and E. B. Khalil, "Graph neural networks with heterophily (CPGNN)," in *Proc. AAAI Conf. on Artificial Intelligence*, 2021.
- [20] J. Gastegger, A. Bojchevski, and S. Günnemann, "Predict then propagate: Graph neural networks meet personalized pagerank," in *Intl. Conf. on Learning Representations (ICLR)*, 2019.
- [21] E. Chien, J. Liao, W. H. Chang, and C. K. Yang, "Adaptive graph convolutional neural networks (GPR-GNN)," in *Intl. Conf. on Learning Representations (ICLR)*, 2021.
- [22] K. Xu, C. Li, Y. Tian, T. Sonobe, K. Kawarabayashi, and S. Jegelka, "Representation learning on graphs with jumping knowledge networks," in *Intl. Conf. on Machine Learning (ICML)*, 2018.
- [23] M. Chen, Z. Wei, Z. Huang, B. Ding, and Y. Li, "Simple and deep graph convolutional networks," in *Intl. Conf. on Machine Learning (ICML)*, 2020.
- [24] H. Göksu, "KrawtchoukNet: Solving GNN over-smoothing with numerically stable discrete global spectral filters," *IEEE Transactions on Neural Networks and Learning Systems*, 2025. (Submitted for review).
- [25] H. Göksu, "CharlierNet: A minimalist and competitive spectral GNN filter for resource-constrained sensor networks," *IEEE Sensors Letters*, 2025. (Submitted for review).
- [26] P. Sen, G. Namata, M. Bilgic, L. Getoor, B. Galligher, and T. Eliassi-Rad, "Collective classification in network data," *AI Magazine*, vol. 29, no. 3, p. 93, 2008.
- [27] Z. Yang, W. W. Cohen, and R. Salakhutdinov, "Revisiting semi-supervised learning with graph embeddings," in *Intl. Conf. on Machine Learning (ICML)*, 2016.
- [28] W. L. Hamilton, R. Ying, and J. Leskovec, "Inductive representation learning on large graphs (GraphSAGE)," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.

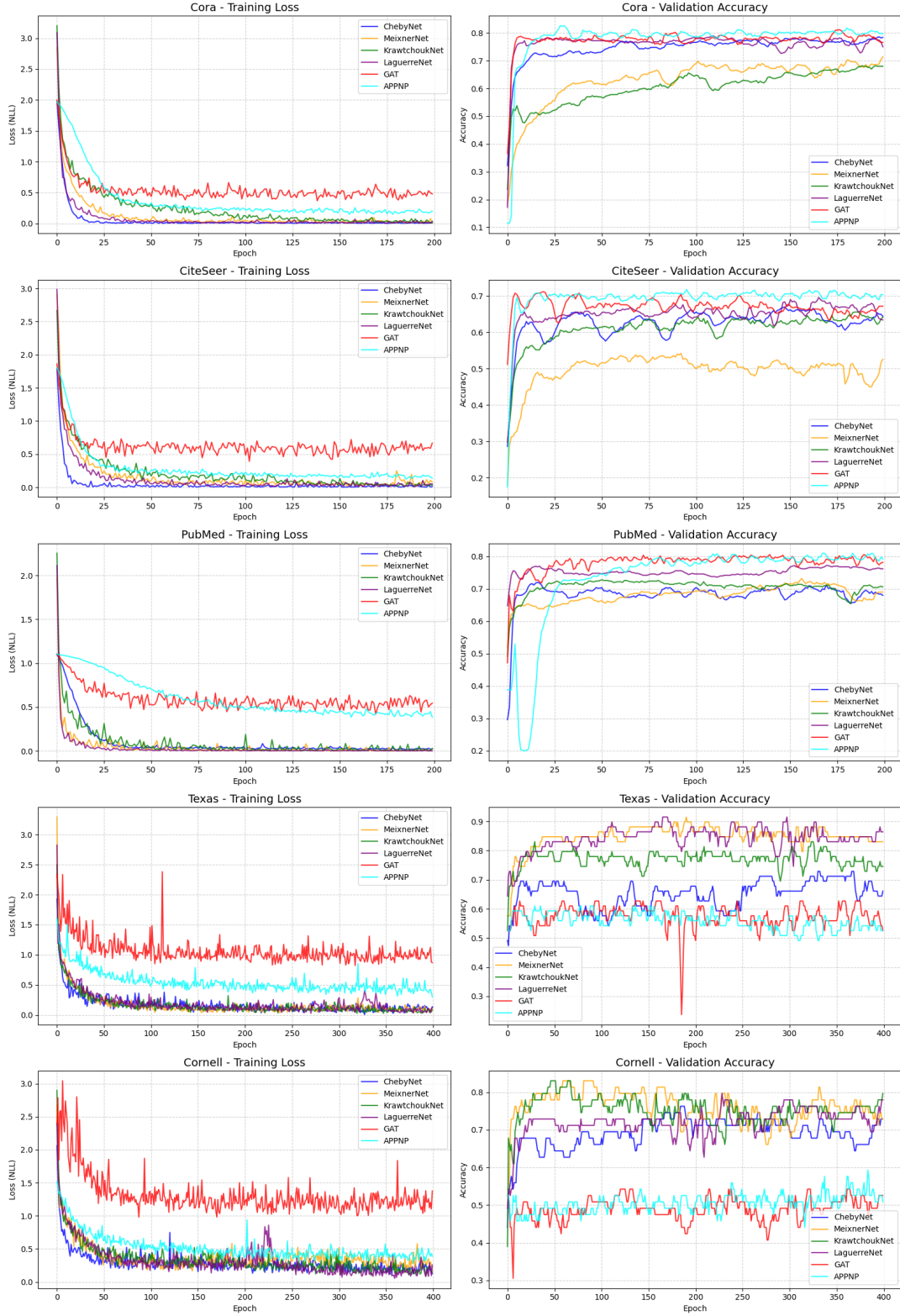


Fig. 1. Training dynamics comparison ( $K=3$ ,  $H=16$ ). Top 3 rows (homophilic): All models are stable. Bottom 2 rows (heterophilic): ‘GAT’ and ‘APPNP’ fail to converge, while our adaptive polynomial filters (‘MeixnerNet’, ‘LaguerreNet’, ‘KrawtchoukNet’) converge quickly to a high, stable accuracy.