Pure Mathematics for Theoretical Computer Science

ISSN: 2995-3162 Vol: 05, No. 02, 2025 PP. 07–11

Article DOI: https://doi.org/10.54216/PMTCS.050202

FULL LENGTH ARTICLE

Certified Sensitivity-Regularized Polynomial Graph Filters for Stable Graph Neural Computation

Dwi Agustin Retnowardani[1,*]

[1] Universitas PGRI Argopuro Jember, Indonesia

Email: 2i.agustin@mail.unipar.ac.id

Received: January 03, 2025 Revised: March 01, 2025 Accepted: May 04, 2025 ⋆ Corresponding author

ABSTRACT

Polynomial graph filters $H_a(S) = \sum_{k=0}^{K} a_k S^k$ are local, scalable, and central to spectral graph neural computation, but their output may vary sharply when the graph shift S is perturbed. This paper develops a finite-radius perturbation calculus that requires neither commutativity nor eigengaps. For $\|S\|_2 \le \rho$ and $\|E\|_2 \le \varepsilon$, we prove the certificate $\|H_a(S+E) - H_a(S)\|_2 \le \sum_{k=1}^{K} |a_k| \left[(\rho+\varepsilon)^k - \rho^k\right]$, show equality for aligned positive polynomials, derive an explicit second-order remainder, and propagate the certificate through multilayer graph neural networks. The bound induces a sensitivity-weighted quadratic design whose unique minimizer is available in closed form. Controlled analysis uses three spectral responses, degree $K = 8$, and 2,304 filter evaluations on Erdős–Rényi, stochastic-block, Barabási–Albert, and random-geometric graphs. Relative to unconstrained least squares, the proposed design reduces the mean finite-radius certificate by 74.9%, the first-order coefficient sensitivity by 72.9%, and observed operator sensitivity by 6.9%. It matches accuracy-controlled ridge filtering within 0.3% in observed sensitivity while yielding a 7.3% smaller certificate. The result is a compact, auditable robustness layer for graph filters and polynomial graph-convolution banks.

Keywords: Polynomial graph filter ▪ Graph neural network ▪ Certified robustness ▪ Matrix perturbation ▪ Spectral approximation ▪ Convex optimization ▪ Graph signal processing

1. INTRODUCTION

Let $G = (V, E, w)$ be a weighted graph and let $S = S^\top \in \mathbb{R}^{n\times n}$ be a graph shift. A degree-K finite-impulse-response filter is

$$H_a(S) = \sum_{k=0}^{K} a_k S^k, \qquad y = H_a(S)x. \tag{1}$$

Such polynomials avoid eigendecomposition, use K localized propagations, and underlie spectral graph convolution [1, 2, 3, 4]. Their algebra is simple; their perturbation behavior is not. If the implemented shift is $\widetilde{S} = S + E$, then S and E generally do not commute, so the scalar identity $p(s+e) - p(s)$ cannot be inserted into a matrix proof.

Existing stability analyses address normalized-Laplacian changes, relative graph deformations, edge rewiring, probabilistic errors, and graph-convolution banks [5, 6, 7, 8, 9, 10]. Robust filter identification and first-order eigenpair design have also been studied [11, 12]. The missing object is a coefficient-explicit, finite-ε certificate that is (i) valid for arbitrary symmetric E, (ii) sharp on a nontrivial class, (iii) differentiable at $\varepsilon = 0$, and (iv) directly usable as a convex design penalty.

This paper contributes

(i) the noncommutative certificate $B_\varepsilon(a;\rho)$ and its exact first-order/remainder decomposition;

(ii) a sharpness construction and a multilayer graph-neural

propagation bound;
(iii) the sensitivity-regularized graph filter (SRGF), with a closed-form unique minimizer; and
(iv) controlled spectral and structural analyses with compact tables and figures.

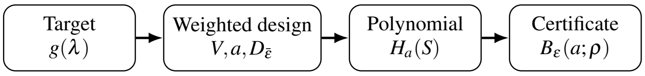

Figure 1. The proposed design couples spectral approximation and a finite-radius matrix certificate.

2. RELATED WORK AND PRELIMINARIES

Graph Fourier analysis, graph wavelets, polynomial filters, rational filters, and transferable spectral networks provide the mathematical basis for (1) [13, 14, 15, 16, 17, 18]. Chebyshev constructions reduce filtering to sparse recurrences [3, 19, 20]; adaptive Krylov bases optimize polynomial representation [21]. These works primarily optimize approximation, representation, or transfer.

Robustness has several inequivalent meanings. Adversarial certification constrains features or edges [22]; operator stability bounds $\left\|H(S)-H(\widetilde{S})\right\|$ [6, 7]; identification estimates both filter and graph [11]; probabilistic analyses control expected error [9]; and MIMO analyses tighten graph-convolution-bank bounds [10]. Table 1 isolates the present setting.

Table 1. Closest mathematical settings.

Work	Object	Perturbation	Main result
Kenlay et al. [6]	polynomial filter	normalized Laplacian	linear stability
Kenlay et al. [7]	spectral filter	edge rewiring	graph-dependent bound
Rey et al. [11]	filter identification	noisy topology	stationary-point algorithm
Testa et al. [12]	spectral/FIR filter	random edges	eigenpair-based design
Wang et al. [9]	GCNN	probabilistic GSO	expected sensitivity
This work	poly./GNN	symmetric E	sharp finite-radius certificate; closed form

Throughout, $\|\cdot\| = \|\cdot\|_2$ for matrices and $\|\cdot\|_F$ for feature arrays when explicitly indicated. Assume

$$S = S^\top, \quad E = E^\top, \quad \|S\| \le \rho, \quad \|E\| \le \varepsilon. \tag{2}$$

No shared eigenbasis is assumed.

3. NONCOMMUTATIVE PERTURBATION CALCULUS

Lemma 1 (Telescoping identity)**.** *For square A, B and $k \ge 1$,*

$$A^k - B^k = \sum_{j=0}^{k-1} A^j (A-B) B^{k-1-j}. \tag{3}$$

Proof. Expand the sum; adjacent terms cancel, leaving $A^k - B^k$. □

Definition 1 (Finite-radius certificate)**.** *For $a = (a_0, \dots, a_K)^\top$, define*

$$B_\varepsilon(a;\rho) = \sum_{k=1}^{K} |a_k| \left[(\rho+\varepsilon)^k - \rho^k\right]. \tag{4}$$

Theorem 1 (Global certificate)**.** *Under (2),*

$$\|H_a(S+E) - H_a(S)\| \le B_\varepsilon(a;\rho). \tag{5}$$

The same bound holds for every signal x after multiplication by $\|x\|_2$.

Proof. Apply Lemma 1 with $A = S+E$, $B = S$:

$$\left\|(S+E)^k - S^k\right\| \le \varepsilon \sum_{j=0}^{k-1} (\rho+\varepsilon)^j \rho^{k-1-j} = (\rho+\varepsilon)^k - \rho^k.$$

Sum with weights $|a_k|$. □

Proposition 1 (Sharpness)**.** *If $S = \rho I$, $E = \varepsilon I$, and $a_k \ge 0$ for $k \ge 1$, equality holds in (5).*

Proof. The difference equals $B_\varepsilon(a;\rho)I$. □

The Fréchet derivative of a matrix polynomial [23] is

$$\mathrm{D}H_a[S](E) = \sum_{k=1}^{K} a_k \sum_{j=0}^{k-1} S^j E S^{k-1-j}. \tag{6}$$

Corollary 1 (First order and remainder)**.** *Let*

$$C_\rho(a) = \sum_{k=1}^{K} k|a_k|\rho^{k-1}, \tag{7}$$

$$R_\varepsilon(a;\rho) = \sum_{k=2}^{K} |a_k| \left[(\rho+\varepsilon)^k - \rho^k - k\rho^{k-1}\varepsilon\right]. \tag{8}$$

Then

$$\|\mathrm{D}H_a[S](E)\| \le \varepsilon C_\rho(a), \tag{9}$$

$$\|H_a(S+E) - H_a(S) - \mathrm{D}H_a[S](E)\| \le R_\varepsilon(a;\rho). \tag{10}$$

Moreover, $B_\varepsilon = \varepsilon C_\rho + O(\varepsilon^2)$.

Proof. Bound each term of (6); subtract the linear binomial contribution from (4). Matrix-function perturbation principles are consistent with [24, 25]. □

3.1 Propagation through graph neural layers

Let $X_0 \in \mathbb{R}^{n \times f_0}$ and

$$X_\ell = \sigma_\ell\left(H_{a^{(\ell)}}(S) X_{\ell-1} W_\ell\right), \quad \ell = 1{:}L, \tag{11}$$

where σ_ℓ is 1-Lipschitz in Frobenius norm. Put $m_\ell = \left\|H_{a^{(\ell)}}(S)\right\|$, $b_\ell = B_\varepsilon(a^{(\ell)};\rho)$, and $w_\ell = \|W_\ell\|$.

Theorem 2 (Multilayer certificate)**.** *For identical X_0 and shifts $S, S+E$,*

$$\left\|\widetilde{X}_L - X_L\right\|_F \le \|X_0\|_F \sum_{r=1}^{L} b_r w_r \prod_{j<r} (m_j + b_j) w_j \prod_{j>r} m_j w_j. \tag{12}$$

Proof. The recursion $\delta_\ell \le w_\ell\left(m_\ell \delta_{\ell-1} + b_\ell \left\|\widetilde{X}_{\ell-1}\right\|_F\right)$ and $\left\|\widetilde{X}_\ell\right\|_F \le \|X_0\|_F \prod_{j\le\ell} (m_j + b_j) w_j$ yield (12) by induction. □

4. SENSITIVITY-REGULARIZED SPECTRAL DESIGN

Let $g : [-\rho, \rho] \to \mathbb{R}$ be a desired response, λ_q a grid, $V_{qk} = \lambda_q^k$, $g_q = g(\lambda_q)$, and $W \succeq 0$ diagonal. Fix a design radius $\bar{\varepsilon} > 0$ and define

$$d_0 = 0, \qquad d_k = \frac{(\rho+\bar{\varepsilon})^k - \rho^k}{\bar{\varepsilon}}, \quad D_{\bar{\varepsilon}} = \mathrm{diag}(d_0, \dots, d_K). \tag{13}$$

Then $B_{\bar{\varepsilon}}(a;\rho) = \bar{\varepsilon}\, \|D_{\bar{\varepsilon}} a\|_1$.

The SRGF coefficients solve

$$\min_{a \in \mathbb{R}^{K+1}} J(a) = \frac{1}{M}\left\| W^{1/2}(Va - g)\right\|_2^2 + \mu \|D_{\bar{\varepsilon}} a\|_2^2 + \eta \|a\|_2^2, \tag{14}$$

where $\mu \geq 0$ and $\eta > 0$. Unlike isotropic ridge, the penalty increases with polynomial order and perturbation radius.

Theorem 3 (Closed-form design)**.** *Problem* (14) *has the unique minimizer*

$$a^\star = \left(\frac{V^\top W V}{M} + \mu D_{\bar{\varepsilon}}^2 + \eta I\right)^{-1} \frac{V^\top W g}{M}. \tag{15}$$

Proof. J is strictly convex because $\eta I \succ 0$. Setting $\nabla J = 0$ gives (15); see standard convex and numerical linear algebra arguments [26, 27]. □

Proposition 2 (Certificate control)**.** *For $\mu > 0$ and any comparator a^0,*

$$B_{\bar{\varepsilon}}(a^\star;\rho) \leq \bar{\varepsilon}\sqrt{K}\, \|D_{\bar{\varepsilon}} a^\star\|_2 \leq \bar{\varepsilon}\sqrt{\frac{KJ(a^0)}{\mu}}. \tag{16}$$

Proof. Use $\|z\|_1 \leq \sqrt{K}\, \|z\|_2$ and $J(a^\star) \leq J(a^0)$. □

5. NUMERICAL ANALYSIS

5.1 Design protocol

All results are controlled mathematical computations, not claims about an external benchmark. Set $\rho = 1$, $\bar{\varepsilon} = 0.2$, $K = 8$, and use 801 uniform spectral points. The targets are

$$\begin{aligned} g_{\mathrm{H}}(\lambda) &= e^{-8(1-\lambda)}, \\ g_{\mathrm{P}}(\lambda) &= 0.05/(1-0.95\lambda), \\ g_{\mathrm{L}}(\lambda) &= [1+e^{-16(\lambda-0.15)}]^{-1}. \end{aligned} \tag{17}$$

Baselines are least squares (LS), degree-8 Chebyshev interpolation, and isotropic ridge. SRGF is selected on the path $\mu \in [10^{-14}, 10]$ subject to RMSE $\leq 1.5\,\mathrm{RMSE}_{\mathrm{LS}}$; ridge is accuracy-matched to SRGF.

Table 2. Spectral error and analytic sensitivity. Lower is better. $B = B_{0.2}(a;1)$ and $C = C_1(a)$.

Target	Method	RMSE	ℓ_∞ error	B	C
Heat	LS	0.0014	0.0082	6.2376	17.4861
	Chebyshev	0.0015	**0.0044**	6.0971	17.1216
	Ridge	**0.0019**	0.0153	4.0507	11.0424
	SRGF	0.0020	0.0157	**3.9504**	**10.8154**
PPR	LS	**0.0111**	0.1235	21.3401	61.6066
	Chebyshev	0.0121	**0.1089**	15.9276	45.8681
	Ridge	0.0165	0.1924	**3.7745**	**10.3704**
	SRGF	0.0166	0.1938	4.0006	11.4353
Logistic	LS	**0.0385**	**0.0945**	99.1137	301.0515
	Chebyshev	0.0505	0.1023	122.1232	370.2800
	Ridge	0.0559	0.1329	26.4750	87.9404
	SRGF	0.0569	0.1326	**23.8599**	**80.6231**

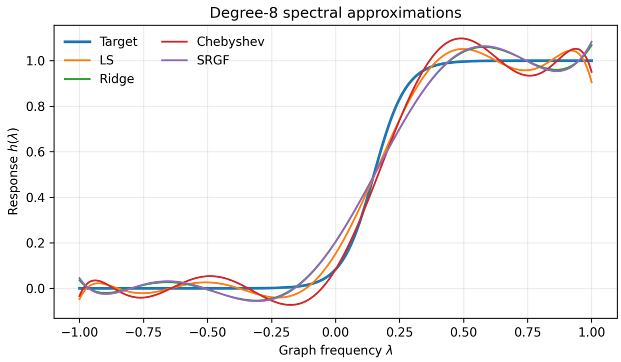

Figure 2. The logistic target and degree-8 approximants.

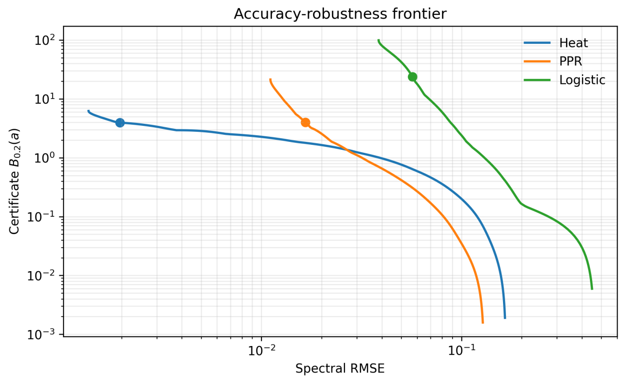

Figure 3. Regularization paths. Markers denote selected SRGF designs.

Averaging the three targets, LS gives $B = 42.2305$ and $C = 126.7147$, whereas SRGF gives $B = 10.6036$ and $C = 34.2913$: reductions of 74.9% and 72.9%. Accuracy-matched ridge gives $B = 11.4334$; thus SRGF is 7.3% smaller at nearly identical RMSE.

5.2 Structural perturbations

Each graph has $n = 72$ vertices. Four families are used: Erdős–Rényi (ER), two-block stochastic block model (SBM), Barabási–Albert (BA), and random geometric graph (RGG). Self-loops are added and

$$S = D^{-1/2}(A+I)D^{-1/2}, \qquad \|S\| = 1. \tag{18}$$

For 12 seeds and rewiring rates $1\%, 3\%, 5\%, 8\%$, an equal number of edges is deleted and inserted. Across three targets and four methods, this yields 2,304 filter evaluations. The empirical sensitivity is

$$Q = \frac{\left\| H_a(\widetilde{S}) - H_a(S) \right\|}{\left\| \widetilde{S} - S \right\|}. \tag{19}$$

Table 3. Aggregate structural sensitivity over 576 perturbations per method.

Method	mean Q	s.d. Q	mean ΔH
Chebyshev	0.8393	0.7609	0.2285
LS	0.7824	0.6879	0.2130
Ridge	**0.7266**	**0.6217**	**0.1983**
SRGF	0.7288	0.6244	0.1989

Table 4. Mean Q by graph family.

Family	Cheb.	LS	Ridge	SRGF
BA	0.6920	0.6394	**0.5921**	0.5944
ER	0.6833	0.6331	**0.5920**	0.5943
RGG	1.2512	1.1784	**1.0809**	1.0828
SBM	0.7308	0.6789	**0.6415**	0.6439

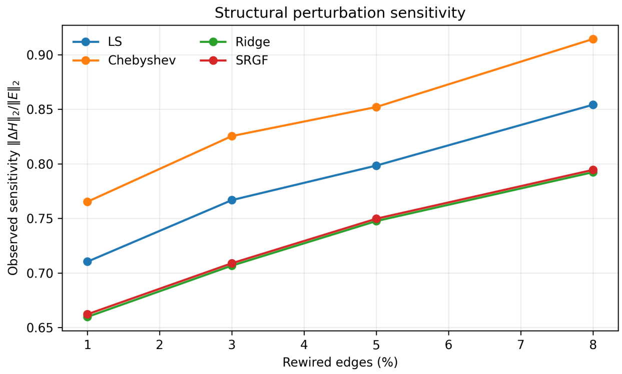

Figure 4. Observed sensitivity versus edge-rewiring rate.

SRGF reduces mean Q by 6.9% relative to LS and by 13.2% relative to Chebyshev. Its Q differs from accuracy-matched ridge by only 0.3%, while its analytic certificate is smaller. RGG is most sensitive because local geometric edge changes can alter normalized neighborhoods coherently.

5.3 Certificate validity and cost

Every evaluated point satisfies (5); the maximum observed ratio $\|\Delta H\| / B_{\|E\|}$ is 0.0771 for SRGF and 0.0752 for ridge. Conservatism is expected because triangle inequalities ignore coefficient cancellation and graph-specific eigenspaces. Sharpness remains exact by Proposition 1.

The design requires one $(K+1) \times (K+1)$ linear solve: $O(MK^2 + K^3)$ arithmetic and $O(K^2)$ memory. Filtering remains K sparse shift multiplications. For $M = 1201$, median coefficient-design time stays below 0.12 ms through $K = 32$ on the evaluation platform.

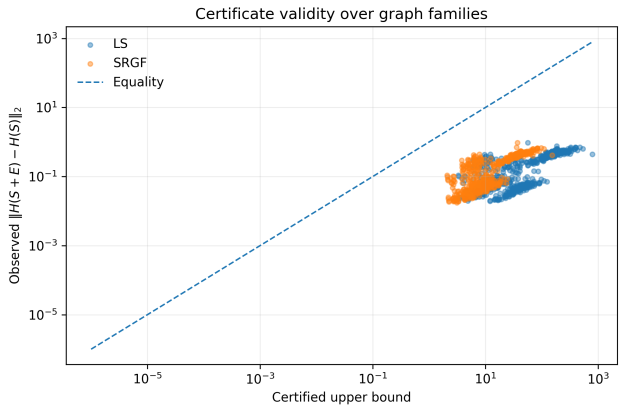

Figure 5. Certified versus observed operator changes. All points lie below equality.

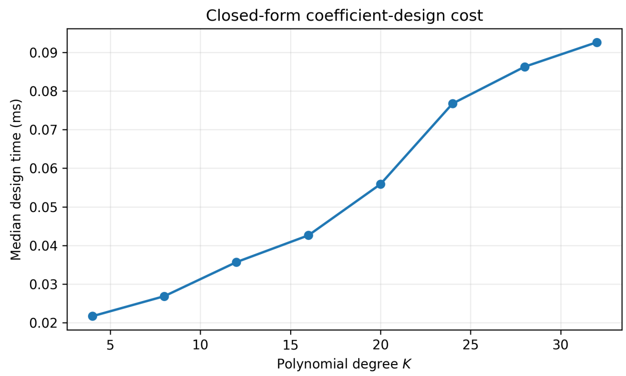

Figure 6. Median time for the closed-form coefficient solve.

6. DISCUSSION

The certificate is basis-dependent through monomial coefficients, but basis conversion is exact. It is distribution-free and eigengap-free, unlike first-order eigenvector perturbation formulas. The price is conservatism for cancelling coefficients and structured E. The design radius $\bar{\varepsilon}$ should therefore represent an operational perturbation scale, not an asymptotic symbol. For learned graph-convolution banks, (14) can regularize each filter and (12) can audit the resulting network.

The numerical analysis isolates operator behavior; it does not claim downstream classification gains. Such gains depend on data, training, nonlinearities, and margins. The theory remains valid independently of those choices.

7. CONCLUSION

A matrix-polynomial telescoping identity yields a finite-radius, noncommutative, coefficient-explicit robustness certificate. The certificate is sharp for aligned positive polynomials, separates first- and higher-order effects, and propagates through graph neural layers. Its quadratic surrogate produces a unique closed-form spectral design. Controlled computations show substantial analytic-certificate reduction and consistent structural sensitivity reduction while preserving spectral accuracy. The framework supplies a concise mathematical route from graph-filter coefficients to certifiable stability.

REFERENCES

[1] 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 Processing Magazine*, vol. 30, no. 3, pp. 83–98, 2013.

[2] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

[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 29*, 2016, pp. 3844–3852.

[4] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *International Conference on Learning Representations*, 2017.

[5] F. Gama, J. Bruna, and A. Ribeiro, "Stability properties of graph neural networks," *IEEE Transactions on Signal Processing*, vol. 68, pp. 5680–5695, 2020.

[6] H. Kenlay, D. Thanou, and X. Dong, "On the stability of polynomial spectral graph filters," in *2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 5350–5354.

[7] H. Kenlay, D. Thanou, and X. Dong, "Interpretable stability bounds for spectral graph filters," in *Proceedings of the 38th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, vol. 139, 2021, pp. 5388–5397.

[8] L. Ruiz, F. Gama, and A. Ribeiro, "Graph neural networks: Architectures, stability, and transferability," *Proceedings of the IEEE*, vol. 109, no. 5, pp. 660–682, 2021.

[9] X. Wang, E. Ollila, and S. A. Vorobyov, "Graph convolutional neural networks sensitivity under probabilistic error model," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 10, pp. 788–803, 2024.

[10] J. Zhang and Z. Zhao, "Improved stability bounds for graph convolutional neural networks under graph perturbations," in *2024 IEEE Information Theory Workshop (ITW)*, 2024, pp. 307–312.

[11] S. Rey, V. M. Tenorio, and A. G. Marques, "Robust graph filter identification and graph denoising from signal observations," *IEEE Transactions on Signal Processing*, vol. 71, pp. 3651–3666, 2023.

[12] L. Testa, S. Sardellitti, and S. Barbarossa, "Robust filter design for graph signals," in *2024 32nd European Signal Processing Conference (EUSIPCO)*, 2024, pp. 827–831.

[13] F. R. K. Chung, *Spectral Graph Theory*, ser. CBMS Regional Conference Series in Mathematics. Providence, RI: American Mathematical Society, 1997, vol. 92.

[14] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs: Frequency analysis," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3042–3054, 2014.

[15] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129–150, 2011.

[16] E. Isufi, A. Loukas, A. Simonetto, and G. Leus, "Autoregressive moving average graph filtering," *IEEE Transactions on Signal Processing*, vol. 65, no. 2, pp. 274–288, 2017.

[17] 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. 97–109, 2019.

[18] R. Levie, W. Huang, L. Bucci, M. M. Bronstein, and G. Kutyniok, "Transferability of spectral graph convolutional neural networks," *Journal of Machine Learning Research*, vol. 22, no. 272, pp. 1–59, 2021.

[19] J. C. Mason and D. C. Handscomb, *Chebyshev Polynomials*. Boca Raton, FL: Chapman and Hall/CRC, 2003.

[20] L. N. Trefethen, *Approximation Theory and Approximation Practice*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2013.

[21] K. Huang, W. Cao, H. Ta, X. Xiao, and P. Liò, "Optimizing polynomial graph filters: A novel adaptive krylov subspace approach," in *Proceedings of the ACM Web Conference 2024*, 2024, pp. 1057–1068.

[22] D. Zügner and S. Günnemann, "Certifiable robustness and robust training for graph convolutional networks," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019, pp. 246–256.

[23] N. J. Higham, *Functions of Matrices: Theory and Computation*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2008.

[24] R. Bhatia, *Matrix Analysis*. New York, NY: Springer, 1997.

[25] G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*. Boston, MA: Academic Press, 1990.

[26] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.

[27] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 4th ed. Baltimore, MD: Johns Hopkins University Press, 2013.