ICASSP 2026 - 2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 979-8-3315-6701-9/26/$31.00 ©2026 IEEE | DOI: 10.1109/ICASSP55912.2026.11460649

DATA-DRIVEN GRAPH FILTERS VIA ADAPTIVE SPECTRAL SHAPING

Dylan Sandfelder[⋆] *Mihai Cucuringu*[†⋆] *Xiaowen Dong*[⋆]

[⋆] University of Oxford, Department of Engineering Science
[†] University of California, Los Angeles, Department of Mathematics

ABSTRACT

We introduce Adaptive Spectral Shaping, a data-driven framework for graph filtering that learns a reusable baseline spectral kernel and modulates it with a small set of Gaussian factors. The resulting multi-peak, multi-scale responses allocate energy to heterogeneous regions of the Laplacian spectrum while remaining interpretable via explicit centers and bandwidths. To scale, we implement filters with Chebyshev polynomial expansions, avoiding eigendecompositions. We further propose Transferable Adaptive Spectral Shaping (TASS): the baseline kernel is learned on source graphs and, on a target graph, kept fixed while only the shaping parameters are adapted, enabling few-shot transfer under matched compute. Across controlled synthetic benchmarks spanning graph families and signal regimes, Adaptive Spectral Shaping reduces reconstruction error relative to fixed-prototype wavelets and learned linear banks, and TASS yields consistent positive transfer. The framework provides compact spectral modules that plug into graph signal processing pipelines and graph neural networks, combining scalability, interpretability, and cross-graph generalization.

***Index Terms*—** Graph Signal Processing, Data-driven Filter Design, Spectral Graph Wavelets

1 Introduction

Graph Signal Processing (GSP) provides a principled framework for analyzing data supported on irregular domains such as networks and point clouds [1, 2]. Within this framework, spectral graph wavelets have become a central tool for multi-scale analysis, enabling localized filtering across graph frequencies via a prototype bandpass function and its dilations [3]. However, many real-world graph signals—from traffic and sensing systems to social and biological networks—exhibit heterogeneous, often multi-peak spectral structure that violates the implicit assumption of a fixed bandpass shape shared across the entire spectrum and across graphs. In such cases, classic wavelet designs may either underfit (by failing to place energy at where the signal lives) or require ad-hoc, hand-tuned banks that sacrifice interpretability and scalability [4].

Recent data-driven approaches begin to address these limitations by learning kernels in the graph spectral domain, including Gaussian-process-based constructions and spectral-kernel learning [5, 6, 7]. Alternatively, attempts have been made to apply similar concepts to tools like spectral graph neural networks [8, 9], or learn node-specific filter weights which can better model local heterogeneity in graphs [10]. Despite this, several challenges remain. First, many learned designs still rely on a single global prototype (or a small, fixed family) whose shape is not free to adapt locally across the spectrum, limiting their ability to capture multi-modal spectral phenomena. Second, practical deployments must avoid repetitive eigendecompositions during training and inference; otherwise, computational costs quickly become prohibitive as graph size grows. Finally, graph neural network designs often require large amounts of training data and suffer from limited interpretability.

To address these challenges, we introduce a *data-driven adaptive spectral shaping* framework that learns a parametric baseline kernel and modulates it with learnable Gaussian shaping factors that target specific spectral regions. By summing multiple such shaped components, our model forms a *multi-scale*, *multi-peak* filter family that can flexibly conform to heterogeneous spectra while remaining interpretable: each component has an explicit center and bandwidth that reveal where (and how narrowly) the filter allocates attention. To ensure scalability, we implement filtering through polynomial approximations of the Laplacian operator, thereby avoiding eigendecomposition at training time while retaining accurate spectral behavior [3]. Qualitatively, the learned responses capture the increasing complexity of the target spectra with additional scales, as illustrated in Fig. 2, and decompose into intuitive spectral parts (Fig. 3).

Beyond fitting a single graph, we further develop a *transfer* variant entitled *Transferable Adaptive Spectral Shaping (TASS)*, in which a shared base kernel is learned on source graphs and only the per-graph shaping parameters (centers, bandwidths, amplitudes) are adapted to a new target graph. This separates what is *universal* (the baseline spectral template) from what is *graph-specific* (where the spectrum must be emphasized), enabling few-shot adaptation and faster convergence on the target domain. Although the transferability of spectral filters between graphs has already been shown to be effective in the case of graph convolutional networks, TASS constitutes, to our knowledge, the first method that generalizes this finding to learnable graph filters [11]. In our controlled synthetic benchmarks spanning multiple graph topologies and signal regimes, TASS consistently improves reconstruction error on the target relative to training from scratch with comparable compute, and it provides clear, component-wise interpretations of how spectral emphasis changes across domains.

The main contributions of our work are as follows.

- We propose an *adaptive spectral shaping* family for graph filters that learns locally tuned, multi-peak spectral responses with an interpretable center/width parameterization.
- We propose *TASS*, a transfer-learning procedure that freezes a shared base kernel and adapts only lightweight shaping parameters on the target, enabling few-shot, cross-graph generalization.
- We provide a scalable implementation based on polynomial filtering of the Laplacian, eliminating expensive eigendecompositions during training and inference.
- On synthetic benchmarks with heterogeneous spectra, our approach achieves lower reconstruction error than fixed-prototype baselines and yields faithful component-wise decompositions that match the ground-truth spectral structure.

Authorized licensed use limited to: Bodleian Libraries of the University of Oxford. Downloaded on May 30,2026 at 10:59:46 UTC from IEEE Xplore. Restrictions apply.

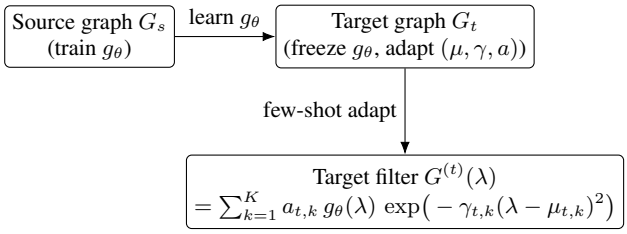


Fig. 1: **TASS pipeline.** Learn a reusable baseline kernel g_θ on a source; on the target, freeze g_θ and adapt only (μ, γ, a) for K components to localize spectral emphasis.

2 Proposed Model

Setup and notation. Let $G = (V, E)$ be a weighted, undirected graph with $|V| = N$ and symmetric adjacency matrix $\mathbf{W} \in \mathbb{R}^{N \times N}$. We write $\mathbf{L}$ for the (combinatorial or normalized) Laplacian with eigendecomposition $\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\top$, where $\mathbf{\Lambda} = \mathrm{diag}(\lambda_0, \ldots, \lambda_{N-1})$ and $0 = \lambda_0 \leq \cdots \leq \lambda_{N-1} = \lambda_{\max}$.

A spectral filter $g : \mathbb{R}_{\geq 0} \to \mathbb{R}$ acts on a signal $\mathbf{x} \in \mathbb{R}^N$ as:

$$g(\mathbf{L})\mathbf{x} = \mathbf{U}\,\mathrm{diag}\big(g(\lambda_0), \ldots, g(\lambda_{N-1})\big)\,\mathbf{U}^\top\mathbf{x}. \tag{1}$$

Classical spectral graph wavelets instantiate g via dilations of a fixed prototype bandpass [3], which can be too rigid when signals exhibit heterogeneous or multi-peak spectra [1].

Adaptive Spectral Shaping. We learn a *baseline* kernel $g_\theta(\lambda)$ that maps $[0, \lambda_{\max}]$ to $\mathbb{R}$. In our design, we choose to model $g_\theta(\lambda)$ as a small multi-layer perception (MLP), which would allow more fine control of its properties using regularization. Furthermore, to target specific spectral regions, we modulate this baseline with learnable Gaussian *shaping* factors:

$$\phi_{\theta,\gamma,\mu}(\lambda) = g_\theta(\lambda)\, \exp\big(-\gamma\,(\lambda - \mu)^2\big),\ \gamma \geq 0,\ \mu \in [0, \lambda_{\max}], \tag{2}$$

where μ and γ control the center frequency and bandwidth respectively. The nonparametric shape of g_θ captures broad trends and is able to adapt to elaborate multi-modal spectra, while the Gaussian parameters μ and γ focus energy where it is needed.

Multi-scale, multi-peak bank. To represent complex, multi-modal spectra, we sum K shaped components with learnable amplitudes $\{a_k\}_{k=1}^K$:

$$G(\lambda) = \sum_{k=1}^{K} a_k\, g_\theta(\lambda)\, \exp\big(-\gamma_k\,(\lambda - \mu_k)^2\big). \tag{3}$$

Note that each component is interpretable (center/width), and increasing K increases expressivity (cf. Fig. 2–3).

Transferable Adaptive Spectral Shaping (TASS). For a collection of graphs $\{\mathbf{L}^{(m)}\}_{m=1}^M$ (possibly with different spectra), we *share* the baseline g_θ across domains and adapt only light per-graph shaping parameters:

$$G^{(m)}(\lambda) = \sum_{k=1}^{K} a_{m,k}\, g_\theta(\lambda)\, \exp\big(-\gamma_{m,k}\,(\lambda - \mu_{m,k})^2\big). \tag{4}$$

This design decouples universal spectral structure (learned in g_θ) from graph-specific emphasis (via $\mu_{m,k}, \gamma_{m,k}, a_{m,k}$), enabling efficient adaptation to a new graph by tuning a small parameter set. To achieve this, we learn g_θ on a chosen source (graph family + signal regime), then initialize the target with those source parameters and adapt (μ, γ, a) while freezing g_θ. We evaluate both zero-/few-shot adaptation (small target sample sizes) and matched-compute comparisons against training from scratch. The TASS pipeline is shown in Fig. 1.

Fast polynomial filtering. Direct use of (1) would require eigendecomposition. Following [3], we approximate each spectral kernel by a Chebyshev expansion on the scaled Laplacian $\widetilde{\mathbf{L}} = \frac{2}{\lambda_{\max}}\mathbf{L} - \mathbf{I}$:

$$\begin{aligned} G(\mathbf{L})\mathbf{x} &\approx \sum_{r=0}^{R} c_r\, T_r(\widetilde{\mathbf{L}})\,\mathbf{x}, \\ T_0(\mathbf{z}) &= \mathbf{I}, \\ T_1(\mathbf{z}) &= \mathbf{z}, \\ T_{r+1}(\mathbf{z}) &= 2\mathbf{z}T_r(\mathbf{z}) - T_{r-1}(\mathbf{z}), \end{aligned} \tag{5}$$

with coefficients $\{c_r\}$ determined by projecting $G(\lambda)$ onto Chebyshev polynomials on [-1, 1]. This yields $\mathcal{O}(R|E|)$ time and $\mathcal{O}(N)$ memory per component, avoiding eigenvectors and scaling to large graphs [8]. Jackson-Chebyshev damping can be used to reduce Gibbs oscillations [12].

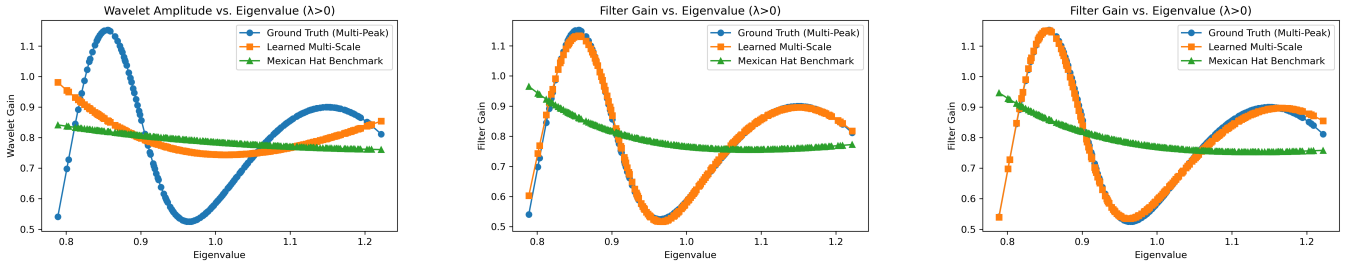
Learning objective. Given training pairs $\{(\mathbf{x}_i^{(m)}, \mathbf{y}_i^{(m)})\}$ on graphs $m = 1, \ldots, M$, we minimize:

$$\begin{aligned} \mathcal{L}(\theta, \mu, \gamma, a) = \sum_{m=1}^{M} \sum_{i} \big| G^{(m)}(\mathbf{L}^{(m)})\,\mathbf{x}_i^{(m)} - \mathbf{y}_i^{(m)} \big|_2^2 \\ + \alpha\, \mathcal{R}_{\mathrm{smooth}}(g_\theta) \\ + \beta\, \mathcal{R}_{\mathrm{shape}}(\mu, \gamma, a), \end{aligned} \tag{6}$$

where $\mathcal{R}_{\mathrm{smooth}}$ encourages a smooth baseline (e.g., finite-difference penalties on g_θ over a spectral grid), and $\mathcal{R}_{\mathrm{shape}}$ regularizes the amplitudes and keeps $\mu \in [0, \lambda_{\max}]$, $\gamma \geq 0$ (enforced in practice via parameterizations such as $\gamma = \mathrm{softplus}(\cdot)$). For TASS, g_θ is first learned on source graphs and then $\{\mu_{m,k}, \gamma_{m,k}, a_{m,k}\}$ are adapted on a target graph with g_θ frozen.

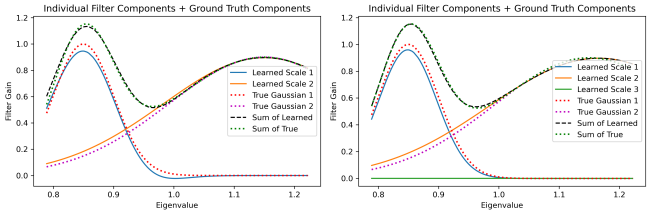
Discussion and relation to prior work. Adaptive Spectral Shaping generalizes fixed-prototype wavelet banks by letting the passband *shape* be learned (via g_θ) and locally *focused* (via μ and γ), while retaining the scalability of polynomial filtering [3]. Our data-driven method contrasts with other methods that, while spectrum-adapted, are nevertheless hand-crafted [13]. Other promising graph filter frameworks with flexible frequency responses that have been used alongside graph neural network approaches include autoregressive moving average graph filtering (ARMA), which has also been used to develop ARMA-based graph neural network layers with improved robustness relative to polynomial filters, CayleyNets, which achieve sharp, band-selective spectral responses using Cayley filters, and BernNets, which achieve near-arbitrary spectral responses using Bernstein polynomial bases [14, 15, 16, 17, 18]. However, these graph neural networks generally require more data to train with limited interpretability. Finally, compared to learning a single global spectral kernel [5, 6], our shaped, multi-component construction captures multi-peak phenomena and remains interpretable: the learned centers and widths directly reveal which spectral regions matter and at what scale. The ability to adapt its shaping parameters to each graph renders our method more adaptive than other interpretable methods like windowed Fourier analysis on graphs [19, 20].

Authorized licensed use limited to: Bodleian Libraries of the University of Oxford. Downloaded on May 30,2026 at 10:59:46 UTC from IEEE Xplore. Restrictions apply.



(a) Single component filters (b) Two component filters (c) Three component filters

Fig. 2: Learned filter adaptations with increasing number of components K: (a) single component, (b) two component, and (c) three component filters.



(a) Two-scale filter components (b) Three-scale filter components

Fig. 3: Learned filter decompositions compared against ground truth components.

3 Experiments

We evaluate Adaptive Spectral Shaping in two regimes: (i) *single-graph filter reconstruction* and (ii) *transfer across graphs and signal regimes* via the proposed TASS, where a shared base kernel is adapted to a new target graph by tuning only lightweight shaping parameters.

3.1 Experimental setup

Graphs and signals. For controlled studies, we generate synthetic graphs and graph signals. Unless noted otherwise, we consider small graphs (N=32) to allow exact spectral filtering via eigendecomposition; we use larger graphs when discussing scalability with polynomial filtering. Graph families include Erdős–Rényi, Barabási–Albert, Watts–Strogatz, 2-D grids, and stochastic block models. We construct ground-truth multi-peak spectral responses $G^\star(\lambda)$ as normalized sums of Gaussians on $[0, \lambda_{\max}]$. Supervision pairs are produced by filtering i.i.d. Gaussian node signals $\mathbf{x}_i$ through $G^\star(\mathbf{L})$ to obtain $\mathbf{y}_i = G^\star(\mathbf{L})\,\mathbf{x}_i$. To probe different GSP regimes we also include smooth signals (low-pass envelopes), spatially localized bumps, band-limited signals, and diffusion-like processes.

Models and baselines. *Adaptive Spectral Shaping (ours)* uses a small MLP $g_\theta(\lambda)$ as the baseline kernel and $K \in \{1, 2, 3\}$ Gaussian shaping components with learnable centers, bandwidths, and amplitudes $\{(\mu_k, \gamma_k, a_k)\}_{k=1}^{K}$. *TASS (ours)* first learns g_θ on a source domain; on the target, g_θ is frozen and only $\{(\mu_k, \gamma_k, a_k)\}$ are adapted. Baselines include (i) classic fixed-prototype graph wavelets (e.g., Mexican hat) with K scales, and (ii) *learned* linear combinations of those fixed atoms with free amplitudes (matched capacity).

Training and implementation details. We optimize mean-squared error with AdamW (lr 10^{-3}, batch 8). The baseline g_θ is regularized for smoothness on a uniform spectral grid; $\gamma_k \geq 0$ is enforced via softplus. For TASS we jointly train on the source, then adapt (μ, γ, a) on the target with g_θ frozen. On larger graphs, filtering is implemented via a Chebyshev expansion on the scaled Laplacian, with degree chosen to bound approximation error for the learned passbands.

Metrics. We report (i) vertex-domain reconstruction error,

$$\mathrm{MSE} = \tfrac{1}{S}\sum_{i=1}^{S} \big\| \hat{\mathbf{y}}_i - \mathbf{y}_i \big\|_2^2,$$

and (ii) a discrete spectral discrepancy over the graph eigenvalues,

$$\mathcal{E}_{\mathrm{spec}} = \tfrac{1}{N}\sum_{j=0}^{N-1} \big|\, G(\lambda_j) - G^\star(\lambda_j) \,\big|^2.$$

For transfer experiments we also measure (iii) the fractional improvement from pre- to post-adaptation on the target,

$$\mathrm{Imp}_{\to\mathrm{adapt}} = \frac{\mathrm{MSE}_{\mathrm{before}} - \mathrm{MSE}_{\mathrm{after}}}{\mathrm{MSE}_{\mathrm{before}}}.$$

3.2 Single-graph filter reconstruction

This setting isolates the benefit of adaptive shaping when spectra are heterogeneous or multi-peak. Fig. 2 shows how increasing K allows the learned response to allocate energy to multiple spectral regions; Fig. 3 decomposes the learned response into components that align with ground-truth peaks.

Findings. Across one-, two-, and three-peak targets, we consistently observe:

- **Adaptive single-scale helps.** Even K=1 improves MSE over fixed-prototype wavelets when the optimal passband drifts across graphs.
- **Multi-peak tracking with K>1.** As K grows, components specialize to distinct bands, reducing both MSE and $\mathcal{E}_{\mathrm{spec}}$ relative to matched-capacity fixed banks.
- **Interpretability.** Component centers/widths provide an intuitive explanation of where the filter allocates attention; recovered components match the ground-truth peaks (cf. Fig. 3).

Adaptive Spectral Shaping reliably outperforms fixed-prototype banks on heterogeneous spectra, with interpretable multi-peak decompositions (Fig. 2–3).

Authorized licensed use limited to: Bodleian Libraries of the University of Oxford. Downloaded on May 30,2026 at 10:59:46 UTC from IEEE Xplore. Restrictions apply.

Generator	Source/Target Relation	K	Num. Exp.	Class Size	$\text{Imp}_{\rightarrow\text{adapt}}\downarrow$
Graph Structure	Same	4	10	157	**-0.073 ± 0.006**
	Different	4	10	444	-0.062 ± 0.006
Graph Signal	Same	4	10	101	-0.060 ± 0.008
	Different	4	10	500	**-0.066 ± 0.008**

Table 1: Transfer across graphs and signal regimes. We pretrain g_θ on a source, then adapt only (μ, γ, a) on the target (TASS).

Spectral dist.	Degree corr.	Clustering sim.	Path len. sim.	Density sim.	Signal corr.	Spectral sim.	Moment sim.
0.035	-0.124	-0.019	0.007	0.002	0.025	0.051	-0.103

Table 2: Pearson correlations between distance/similarity metrics and normalized transfer improvement.

3.3 Transfer across graphs and signal regimes (TASS)

We now assess whether a *shared* baseline kernel g_θ learned on a source can be adapted efficiently to a new target graph by tuning only (μ, γ, a).

Results. The main results of our transfer experiment can be seen in Table 1 and Table 2.

- **Positive transfer and rapid adaptation.** TASS reliably reduces target error after adaptation (positive $\text{Imp}_{\rightarrow\text{adapt}}$) and typically surpasses from-scratch training under matched compute (positive Δ_{T}).
- **When is transfer strongest?** Benefits are largest when source and target share structural characteristics (e.g., both source and target graphs are from the same generator) and when Laplacian spectra are not drastically different. With simultaneous shifts in both structure and signal type, TASS remains competitive, with smaller but still positive gains. Graph signal generators, on the other hand, do not impact transfer performance as much.
- **Few-shot efficiency.** Adapting (μ, γ, a) requires few target pairs: most of the capacity resides in the shared g_θ, so target MSE drops sharply with very limited data and saturates quickly.
- **Structural property correlations.** By measuring the correlations of various properties of both the graphs and the graph signals from the sources to the targets, we can see how graph structure has a much larger impact on transfer performance than choice of signal.

TASS delivers positive, few-shot transfer across graphs by freezing a shared g_θ and adapting only lightweight shaping parameters, yielding faster convergence and competitive or superior error under matched compute.

3.4 Ablations and practical notes

Number of scales (K). Increasing K monotonically improves fit on multi-peak targets up to the target peak count; beyond that, gains diminish and components split peaks with similar aggregate performance.

Polynomial degree. With Chebyshev filtering on larger graphs, moderate degrees suffice when passbands are not extremely narrow. Approximation degree can be tuned to the learned bandwidths for efficient inference.

Runtime/memory. Exact filtering (small N) is used for diagnosis; polynomial filtering scales as $\mathcal{O}(R|E|)$ per component with linear memory in N, making Adaptive Spectral Shaping and TASS viable as a spectral layer in larger pipelines.

4 Conclusion & Future Work

We presented *Adaptive Spectral Shaping*, a data-driven graph filtering framework that learns a reusable baseline kernel $g_\theta(\lambda)$ and modulates it with a small set of Gaussian factors to capture heterogeneous, multi-peak spectra. The Chebyshev implementation scales to larger graphs without eigendecompositions. We further introduced *TASS*, which transfers g_θ across graphs and adapts only lightweight per-graph parameters. Across synthetic studies, the approach reduces reconstruction error relative to fixed-prototype baselines, offers interpretable component-wise explanations, and yields positive few-shot transfer with graceful degradation under structural perturbations.

Limitations. Our evaluation focuses on synthetic ground-truth filters and small/medium graphs; real data may introduce additional confounders. Objectives are reconstruction-centric, and practical choices for polynomial degree/conditioning merit a more systematic treatment at scale.

Future directions.

- **Task-driven use and GNN integration:** Train with downstream losses (forecasting, detection, label propagation) and deploy as a spectral layer within larger GNN pipelines.
- **Local adaptivity & theory:** Move from graph-level to node/community-aware shaping with sparsity/smoothness priors; study stability, localization, and sample-complexity guarantees where polynomial spectral filters have shown promise [21].
- **Transfer at scale:** Multi-source pretraining of g_θ, similarity-based target selection to avoid negative transfer, and few-shot/continual adaptation via hypernetworks or meta-learning.
- **Computation & uncertainty:** Degree-adaptive polynomials, Krylov/Lanczos variants for very large graphs, and Bayesian priors or bandpass/monotonicity constraints to quantify and control spectral uncertainty.

Overall, adaptive and transferable spectral shaping offers a compact, interpretable alternative to fixed wavelet banks and black-box filters, and is well-suited as a drop-in spectral module for scalable GSP/GNN systems.

Authorized licensed use limited to: Bodleian Libraries of the University of Oxford. Downloaded on May 30,2026 at 10:59:46 UTC from IEEE Xplore. Restrictions apply.

5 References

[1] David I Shuman, Sunil K Narang, Pascal Frossard, Antonio Ortega, and Pierre 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] Antonio Ortega, Pascal Frossard, Jelena Kovačević, José MF Moura, and Pierre Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

[3] David K Hammond, Pierre Vandergheynst, and Rémi Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129–150, 2011.

[4] Nicolas Tremblay and Pierre Borgnat, "Graph wavelets for multiscale community mining," *IEEE Transactions on Signal Processing*, vol. 62, no. 20, pp. 5227–5239, 2014.

[5] Yin-Cong Zhi, Yin Cheng Ng, and Xiaowen Dong, "Gaussian processes on graphs via spectral kernel learning," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 9, pp. 304–314, 2023.

[6] Felix Opolka, Yin-Cong Zhi, Pietro Lio, and Xiaowen Dong, "Adaptive gaussian processes on graphs via spectral graph wavelets," in *International Conference on Artificial Intelligence and Statistics*. PMLR, 2022, pp. 4818–4834.

[7] Henry Kenlay, Dorina Thanou, and Xiaowen Dong, "Interpretable stability bounds for spectral graph filters," in *International conference on machine learning*. PMLR, 2021, pp. 5388–5397.

[8] Michaël Defferrard, Xavier Bresson, and Pierre Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," *Advances in neural information processing systems*, vol. 29, 2016.

[9] Fernando Gama, Alejandro Ribeiro, and Joan Bruna, "Diffusion scattering transforms on graphs," *arXiv preprint arXiv:1806.08829*, 2018.

[10] S. A. Tailor, F. L. Opolka, P. Lio, and N. D. Lane, "Do we need anisotropic graph neural networks?," in *International Conference on Learning Representations (ICLR)*, 2022.

[11] Ron Levie, Wei Huang, Lorenzo Bucci, Michael Bronstein, and Gitta Kutyniok, "Transferability of spectral graph convolutional neural networks," *Journal of Machine Learning Research*, vol. 22, no. 272, pp. 1–59, 2021.

[12] Alexander Weiße, Gerhard Wellein, Andreas Alvermann, and Holger Fehske, "The kernel polynomial method," *Reviews of modern physics*, vol. 78, no. 1, pp. 275–306, 2006.

[13] David I Shuman, Christoph Wiesmeyr, Nicki Holighaus, and Pierre Vandergheynst, "Spectrum-adapted tight graph wavelet and vertex-frequency frames," *IEEE Transactions on Signal Processing*, vol. 63, no. 16, pp. 4223–4235, 2015.

[14] Elvin Isufi, Andreas Loukas, Andrea Simonetto, and Geert Leus, "Autoregressive moving average graph filtering," *IEEE Transactions on Signal Processing*, vol. 65, no. 2, pp. 274–288, 2016.

[15] Giuseppe Patanè, "Fourier-based and rational graph filters for spectral processing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 6, pp. 7063–7074, 2022.

[16] Filippo Maria Bianchi, Daniele Grattarola, Lorenzo Livi, and Cesare Alippi, "Graph neural networks with convolutional arma filters," *IEEE transactions on pattern analysis and machine intelligence*, vol. 44, no. 7, pp. 3496–3507, 2021.

[17] Ron Levie, Federico Monti, Xavier Bresson, and Michael M Bronstein, "Cayleynets: Graph convolutional neural networks with complex rational spectral filters," *IEEE Transactions on Signal Processing*, vol. 67, no. 1, pp. 97–109, 2018.

[18] Mingguo He, Zhewei Wei, Hongteng Xu, et al., "Bernnet: Learning arbitrary graph spectral filters via bernstein approximation," *Advances in neural information processing systems*, vol. 34, pp. 14239–14251, 2021.

[19] David I Shuman, Benjamin Ricaud, and Pierre Vandergheynst, "Vertex-frequency analysis on graphs," *Applied and Computational Harmonic Analysis*, vol. 40, no. 2, pp. 260–291, 2016.

[20] David I Shuman, Benjamin Ricaud, and Pierre Vandergheynst, "A windowed graph fourier transform," in *2012 IEEE Statistical Signal Processing Workshop (SSP)*. Ieee, 2012, pp. 133–136.

[21] Henry Kenlay, Dorina Thanou, and Xiaowen Dong, "On the stability of polynomial spectral graph filters," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 5350–5354.

Authorized licensed use limited to: Bodleian Libraries of the University of Oxford. Downloaded on May 30,2026 at 10:59:46 UTC from IEEE Xplore. Restrictions apply.