ICML
International Conference
On Machine Learning

A New Perspective on the Effects of Spectrum in Graph Neural Networks

Mingqi Yang, Yanming Shen, Rui Li, Heng Qi, Qiang Zhang, Baocai Yin

Dalian University of Technology, Peng Cheng Laboratory, China

Agenda

- The effects of graph spectrum in existing GNN architectures
- A new architecture inspired by the effects of graph spectrum
- Results

The effects of graph spectrum in existing GNN architectures

A unified formulation of Existing GNNs

$$H = \sigma\left(p_\gamma(S) f_\Theta(H)\right)$$

- S: Graph matrix, e.g. (normalized) Adjacency, Laplacian, etc
- H: Hidden features/signals
- f_Θ: MLP with parameters Θ
- p_γ: Polynomial with coefficients γ
- σ: Nonlinear function

Table 1. A summary of p_γ in general graph convolutions.

	GCN	SGC	APPNP	GCNII	GDC	SSGC	GPR	ChebyNet	CayleNet	BernNet
Poly-basis	General	General	Residual	Residual	General	General	General	Chebyshev	Cayle	Bernstein
Poly-coefficient	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed	Learnable	Fixed	Learnable	Learnable

Correlation analysis

Graph convolution on a single signal $\boldsymbol{h}$

Before convolution

$$\cos(\langle \boldsymbol{h}, \boldsymbol{p}_i \rangle) = \frac{\alpha_i}{\sqrt{\sum_{j=1}^{n} \alpha_j^2}}$$

After convolution

$$\cos(\langle \boxed{\mathcal{S}}\boldsymbol{h}, \boldsymbol{p}_i \rangle) = \frac{\alpha_i \boxed{\lambda_i}}{\sqrt{\sum_{j=1}^{n} \alpha_j^2 \boxed{\lambda_j^2}}}$$

$\boldsymbol{p}_i$ the i-th eigenvector of $\mathcal{S} = p_\gamma(S)$, $\boldsymbol{h} \in f_\Theta(H)$, $\alpha_i = \boldsymbol{h}^\top \boldsymbol{p}_i$.
The convolution on the signal $\boldsymbol{h}$ is represented as $\mathcal{S}\boldsymbol{h}$.

Stacking multiple graph convolutions

- $\lim_{k\to\infty} |\cos(\langle \mathcal{S}^k \boldsymbol{h}, \boldsymbol{p}_1 \rangle)| = \lim_{k\to\infty} |\cos(\langle \mathcal{S}^k \boldsymbol{h}, \mathcal{S}^k \boldsymbol{h}' \rangle)| = 1$
- $|\cos(\langle \mathcal{S}^{k+1} \boldsymbol{h}, \boldsymbol{p}_1 \rangle)| \geqslant |\cos(\langle \mathcal{S}^k \boldsymbol{h}, \boldsymbol{p}_1 \rangle)|$ and $|\cos(\langle \mathcal{S}^{k+1} \boldsymbol{h}, \boldsymbol{p}_n \rangle)| \leqslant |\cos(\langle \mathcal{S}^k \boldsymbol{h}, \boldsymbol{p}_n \rangle)|$

Spectral smoothness determines the correlations among signals

Correlation issue restricts developing deep models as well as powerful filters.

Oversmoothing analysis VS correlation analysis

		Convergence analysis	
		Oversmoothing analysis based on Markov process	Correlation analysis (Ours)
Considered metric		Signal similarity	Signal cosine similarity
Requirements	Connected	Necessary	Unnecessary
	Normalized	Necessary	Unnecessary
	Deep enough	Necessary	Unnecessary
Convergence property	Monotonous	-	Yes

A new architecture inspired by the effects of graph spectrum

Correlation-free architecture

Correlation-free architecture

$$H = \sigma\left(p_{\gamma}(S) f_{\Theta}(H)\right)$$

Channel-shared filter (Existing)

$$H = f_{\Psi}\left(\left[p_{\Gamma_1}(S) f_{\Theta_1}(H), \ldots, p_{\Gamma_{d'}}(S) f_{\Theta_{d'}}(H)\right]\right)$$

Channel-independent filter (Ours)

Look into the *i*-th filter

$$g_{\Gamma_i} = \sum_{j=0}^{k} \Gamma_{i,j} \boldsymbol{\lambda}^j = \left(\boldsymbol{\lambda}^1, \boldsymbol{\lambda}^2, \ldots, \boldsymbol{\lambda}^k\right) \times \Gamma_i$$

Coefficients of the *i*-th filter

The columns of Vandermonde matrix serve as filter bases

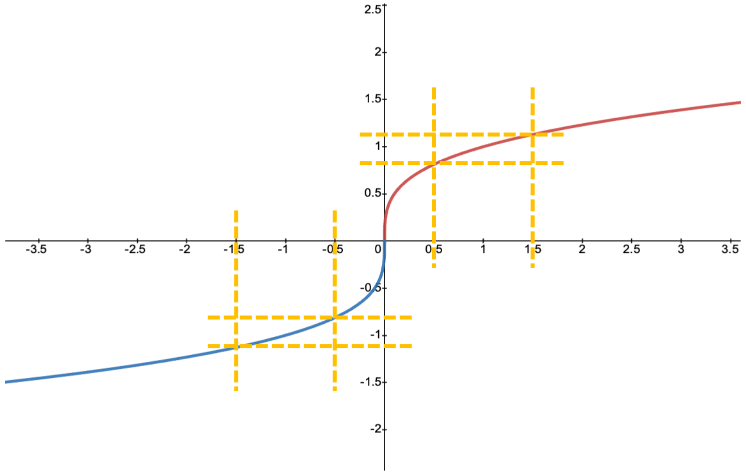
Improve polynomial filters with optimized bases

Eigendecomposition based

$$\tilde{A}_\rho = P\mathrm{diag}(f_\rho(\lambda_i))P^\top,$$

$$f_\rho(\lambda) = \begin{cases} -(-\lambda)^\rho, \lambda < 0 \\ \lambda^\rho, \lambda \geqslant 0 \end{cases},$$

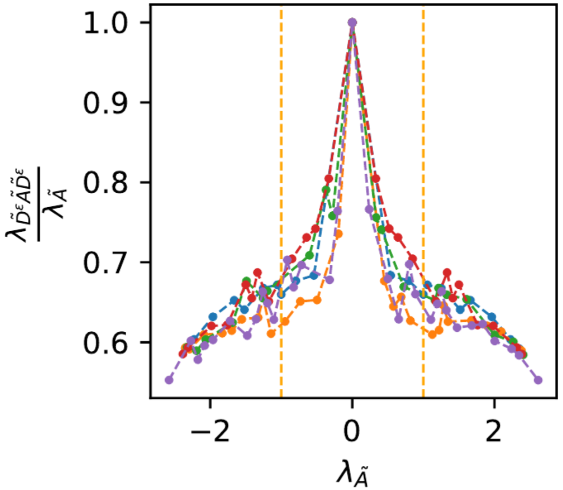
$$\rho \in (0,1)$$



Generalized graph normalization

$$\tilde{D}^\epsilon \tilde{A} \tilde{D}^\epsilon = (D + \eta I)^\epsilon (A + \eta I)(D + \eta I)^\epsilon,$$

$$\epsilon \in [-0.5, 0]$$



Results

Ablation study: filter complexity

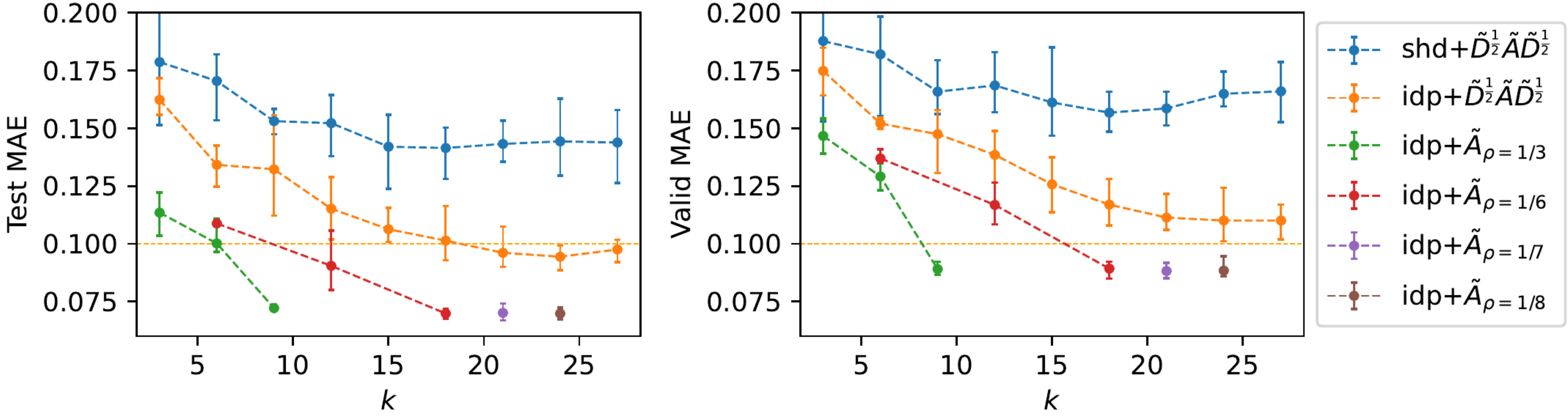


Figure 4. Ablation study results on ZINC with different number of bases k.

Ablation study: model depth

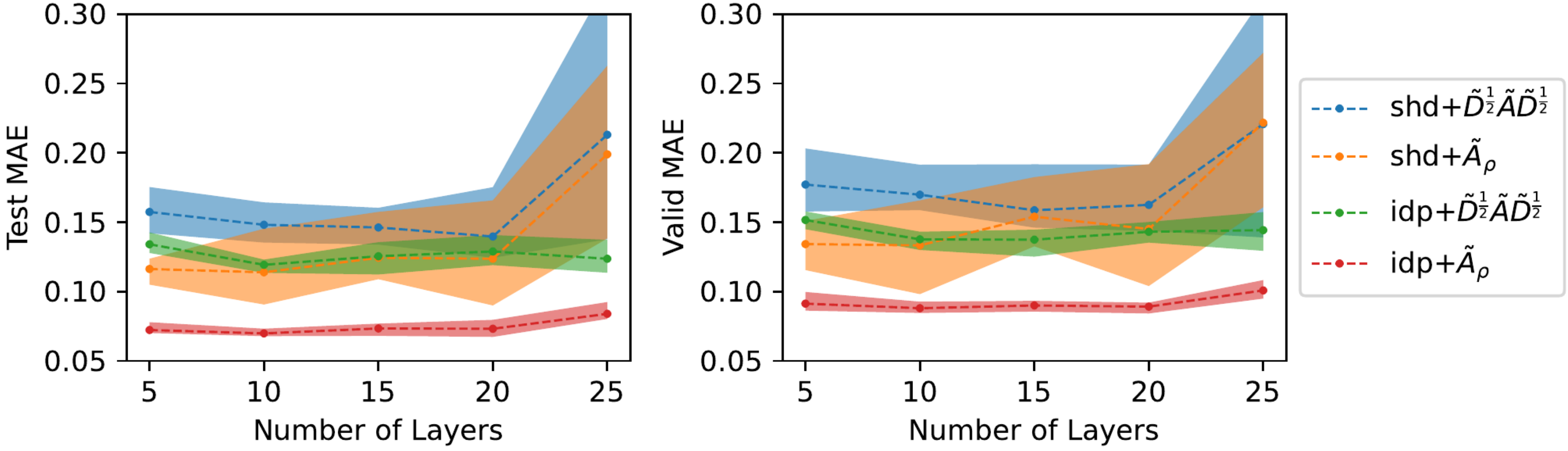


Figure 5. Ablation study results on ZINC with different number of layers.

Summary

- Correlation analysis
  - We show that in existing architectures, the unsmooth spectrum results in the correlation issue, which acts as the obstacle to developing deep models as well as applying more powerful graph filters.

- Propose the correlation-free architecture
  - decouples the correlation issue from filter design.

- Improve the approximation abilities of polynomial filters
  - We show that the spectral characteristics of bases also hinder the approximation abilities of polynomial filters and address it by altering the graph's spectrum.

Thank you for listening!

https://github.com/qslim/gnn-spectrum