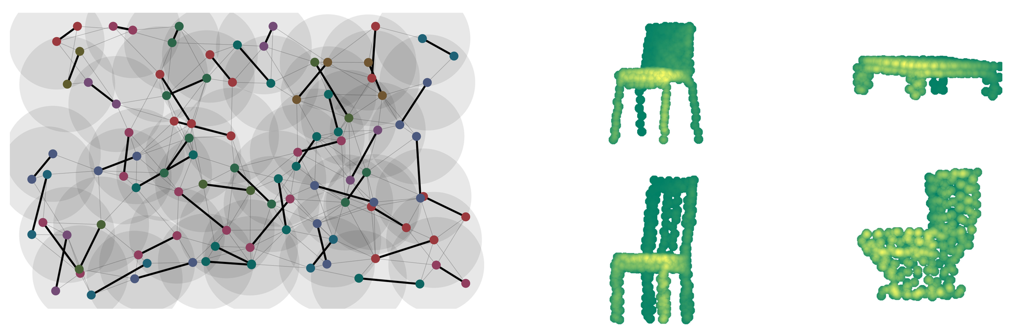
Manifolds Filters and Neural Networks: Geometric Graph Signal Processing in the Limit

Zhiyang Wang

Dept. of Electrical & Systems Engineering, University of Pennsylvania

Geometric Graphs and Graph Signals

- Signals on geometric graphs appear in several application domains
  - ⇒ Wireless communication networks, 3D point clouds, Climate data



- We develop a limit theory of signal processing (SP) on geometric graphs
  - ⇒ Geometric graphs converge (or are sampled from) Manifolds
  - ⇒ Convergence. Stability. Wireless Networks. Vector Fields

Manifold Convolutional Filters

- Manifold $\mathcal{M} \subset \mathbb{R}^N$ is d-dimensional with Laplace-Beltrami (LB) operator $\mathcal{L}$
- A Manifold filter with coefficients $\tilde{h}$ is defined by the input-output relationship

$$g(x) = \int_0^\infty \tilde{h}(t)\, e^{-t\mathcal{L}} f(x) \mathrm{d}t \;=\; \mathbf{h}(\mathcal{L})\, f(x).$$

- Discretizing a manifold filter yields a graph filter with shift operator $e^{-T_s \mathbf{L}_n}$

$$\mathbf{g} = \sum_{k=0}^{K_t-1} \tilde{h}(kT_s)\, e^{-kT_s \mathbf{L}_n} \mathbf{f} \approx \sum_{k=0}^{K_t-1} \tilde{h}(kT_s)\, (\mathbf{I} - T_s \mathbf{L}_n)^k\, \mathbf{f}$$

- Recover standard convolutions if we make the particular choice $\mathcal{L} = \mathrm{d}/\mathrm{d}x$

$$g(x) = \int_0^\infty \tilde{h}(t)\, e^{-t\mathrm{d}/\mathrm{d}x} f(x)\, \mathrm{d}t = \int_0^\infty \tilde{h}(t)\, f(x-t)\, \mathrm{d}t$$

- Manifold convolutions generalize standard (time) and graph convolutions

Spectral Representation of Manifold Convolutional Filters

- LB operator admits discrete spectral decomposition $\Rightarrow \mathcal{L}f = \sum_{i=1}^{\infty} \lambda_i \langle f, \phi_i \rangle \phi_i$
- Manifold Fourier Transform of f is the set of projections $\Rightarrow [\hat{f}]_i = \langle f, \phi_i \rangle$
- Frequency response of filter h is $\Rightarrow \hat{h}(\lambda) = \int_0^\infty \tilde{h}(t) e^{-t\lambda} \mathrm{d}t$

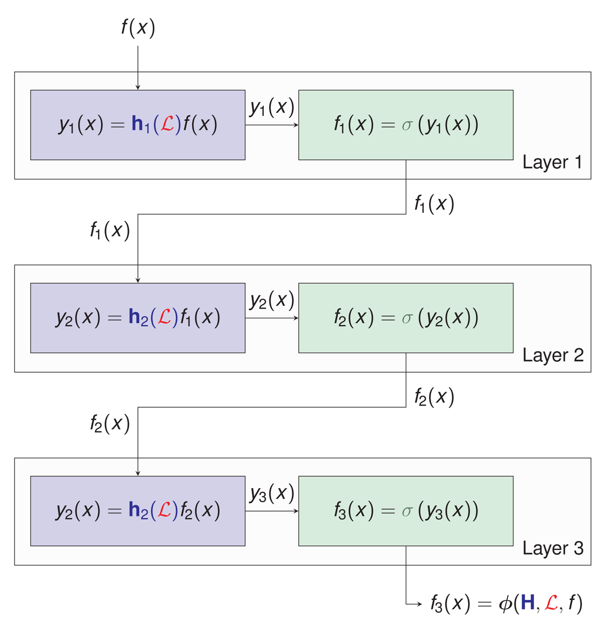
Theorem (Manifold Filters in the Manifold Spectral Domain)

Manifold filters are pointwise in the spectral domain $\Rightarrow [\hat{g}]_i = \hat{h}(\lambda_i)[\hat{f}]_i$

- Manifold filters are easy to study in the manifold frequency (spectral) domain

Manifold Neural Networks (MNNs)

- A MNN is a cascade of L layers
- Each of the layers is composed of
  - ⇒ Manifold convolutions $\mathbf{h}(\mathcal{L})$
  - ⇒ Pointwise nonlinearities σ
- Group learnable coefficients in $\mathbf{H}$
- Write MNN as map $y = \Phi(\mathbf{H}, \mathcal{L}, f)$



Transferability of Geometric Graph Neural Networks

- Geometric graph filters and GNNs converge to their manifold counterparts
  - ⇒ Enables transferability of geometric GNNs from small to large graphs
- Sample the manifold at $\{x_i\}_{i=1}^n$. Construct graph Laplacian of $\mathbf{G}_n$ with edges

$$w_{ij} = K_\xi\left(\frac{\|x_i - x_j\|^2}{\xi}\right)$$

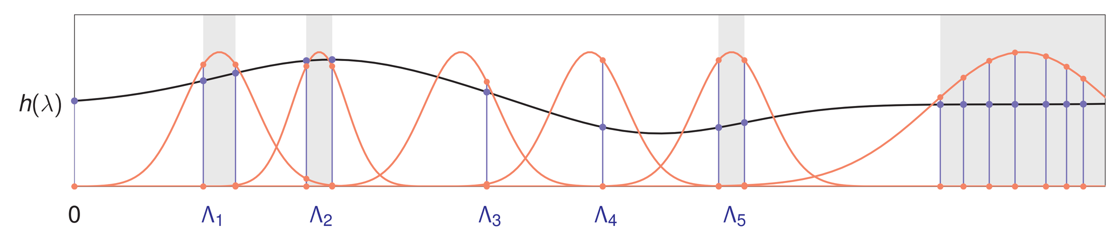
- Geometric graph filter is defined by replacing with graph Laplacians $\mathbf{L}_n$

$$\mathbf{g} = \int_0^\infty \tilde{h}(t) e^{-t\mathbf{L}_n} \mathrm{d}t \mathbf{f} = \mathbf{h}(\mathbf{L}_n)\mathbf{f}, \qquad [\mathbf{f}]_i = f(x_i)$$

- Geometric graph neural networks on $\mathbf{G}_n \Rightarrow \Phi(\mathbf{H}, \mathbf{L}_n, \mathbf{f})$

Lipschitz and Frequency Difference Threshold (FDT) Filters

- A filter is A_h-Lipschitz if its frequency response $\hat{h}(\lambda)$ is A_h-Lipschitz
- Partition spectrum such that λ_i and λ_j are in different partitions if $|\lambda_i - \lambda_j| \geq \alpha$
- A filter is α-FDT if $|\hat{h}(\lambda_i) - \hat{h}(\lambda_j)| \leq \delta_D$ for all λ_i, λ_j in the same partition



- Does not discriminate frequency components associated to close eigenvalues

Convergence of Geometric GNNs to MNNs

Theorem (Convergence of Geometric GNNs)

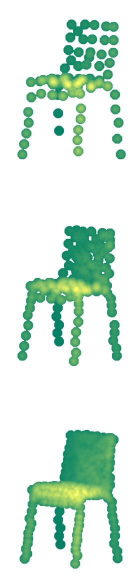
If an L-layer MNN $\Phi(\mathbf{H}, \mathcal{L}, \cdot)$ on $\mathcal{M}$ and GNN $\Phi(\mathbf{H}, \mathbf{L}_n, \cdot)$ on $\mathbf{G}_n$ have normalized Lipschitz nonlinearities, it holds in high probability that

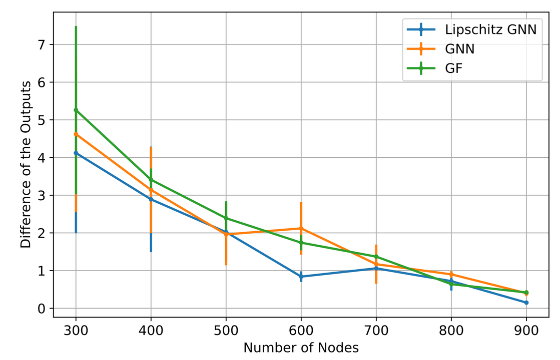
$$\left\| \Phi(\mathbf{H}, \mathbf{L}_n', \mathbf{P}_n f) - \mathbf{P}_n \Phi(\mathbf{H}, \mathcal{L}, f) \right\|_{L^2(\mathbf{G}_n)} \leq O\left[\left(\frac{N}{\alpha} + A_h\right)\sqrt{\xi}\right] + O\left(\frac{\log(n)}{n}\right)$$

with filters that are α-FDT with $\delta_D \leq O(\sqrt{\xi}/\alpha)$ and A_h-Lipschitz continuous.

- The properties of large GNNs can be analyzed via MNN as their limit
- The error bounds show trade-off between discriminability and approximation

Training through Transferability on Point Clouds





	Graph Filters	GNN	Lipschitz GNN
$n = 300$	$21.15\% \pm 3.48\%$	$9.35\% \pm 2.46\%$	$7.63\% \pm 3.36\%$
$n = 500$	$18.09\% \pm 6.28\%$	$7.80\% \pm 3.50\%$	$7.54\% \pm 4.01\%$
$n = 700$	$17.31\% \pm 6.59\%$	$8.16\% \pm 2.95\%$	$7.97\% \pm 2.45\%$
$n = 900$	$15.58\% \pm 4.54\%$	$7.20\% \pm 3.77\%$	$6.68\% \pm 3.94\%$

Z. Wang, L. Ruiz, and A. Ribeiro. "Geometric Graph Filters and Neural Networks: Limit Properties and Discriminability Trade-offs." arXiv preprint arXiv:2305.18467 (2023).

Manifold Deformations as Operator Perturbations

- Stability to deformations is a distinguishable property of CNNs
- Stability of MNNs to deformations can be generalized to GNNs and CNNs
  - ⇒ Consider manifold signal f and a deformation $\tau(x)$ over the manifold

$$p(x) = \mathcal{L}' f(x) = \mathcal{L} g(x) = \mathcal{L} f(\tau(x))$$

  - ⇒ Translate manifold signal perturbations as LB operator perturbations

Theorem (Manifold deformations)

Let the deformation $\tau(x) : \mathcal{M} \to \mathcal{M}$ satisfies $\mathrm{dist}(x, \tau(x)) = \epsilon$ and $J(\tau_*) = I + \Delta$ with $\|\Delta\|_F = \epsilon$. If the gradient field is smooth, it holds that

$$\mathcal{L} - \mathcal{L}' = \mathbf{E}\mathcal{L} + \mathcal{A},$$

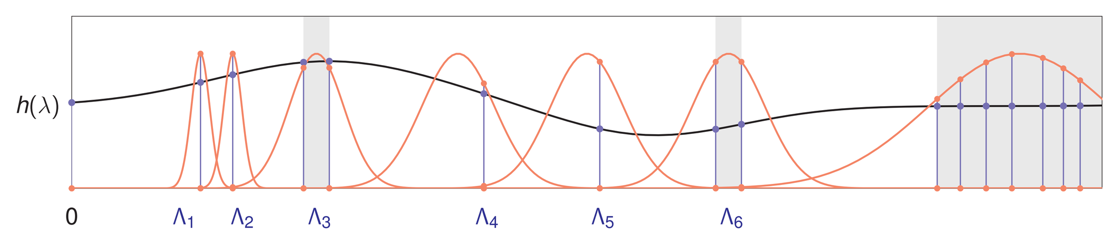
where $\mathbf{E}$ and $\mathcal{A}$ satisfy $\|\mathbf{E}\| = O(\epsilon)$ and $\|\mathcal{A}\|_{op} = O(\epsilon)$.

Integral Lipschitz and Frequency Ratio Threshold (FRT) Filters

- A filter is B_h-Integral Lipschitz if its frequency response satisfies

$$|\hat{h}(a) - \hat{h}(b)| \leq \frac{B_h |a-b|}{(a+b)/2}, \quad \text{for all } a, b \in (0, \infty)$$

- Partition spectrum such that λ_i and λ_j are in different partitions if $\left|\frac{\lambda_i}{\lambda_j} - 1\right| \geq \gamma$
- A filter is γ-FRT if $|\hat{h}(\lambda_i) - \hat{h}(\lambda_j)| \leq \delta_R$ for all λ_i, λ_j in the same partition



- Discriminate frequency components that are relatively far from each other

Stability of Manifold Neural Networks

Theorem (Stability of MNNs to deformations)

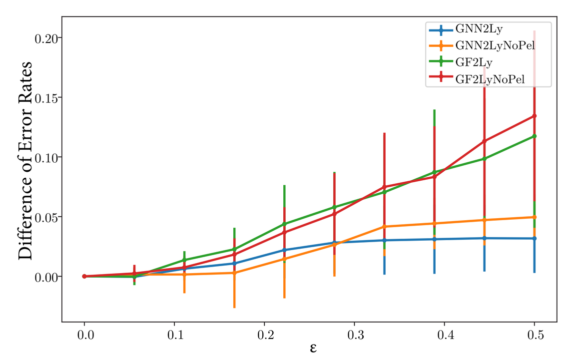
An L-layer MNN $\Phi(\mathbf{H}, \mathcal{L}, f)$ have normalized Lipschitz continuous nonlinearities. Let $\mathcal{L}'$ be the deformed LB operator with $\max\{\alpha, 2, |\gamma/1 - \gamma|\} \gg \epsilon$, then

$$\left\| \Phi(\mathbf{H}, \mathcal{L}, f) - \Phi(\mathbf{H}, \mathcal{L}', f) \right\|_{L^2(\mathcal{M})} \leq O\left[\left(\frac{N}{\alpha} + A_h + \frac{M}{\gamma} + B_h\right)\epsilon\right] \|f\|_{L^2(\mathcal{M})}$$

if the manifold filters are α-FDT with $\delta_D \leq O(\epsilon/\alpha)$, γ-FRT with $\delta_R \leq O(\epsilon/\gamma)$, A_h-Lipschitz continuous and B_h-integral Lipschitz continuous.

- The difference bound shows a trade-off between stability and discriminability

Verifications of Stability under Perturbations

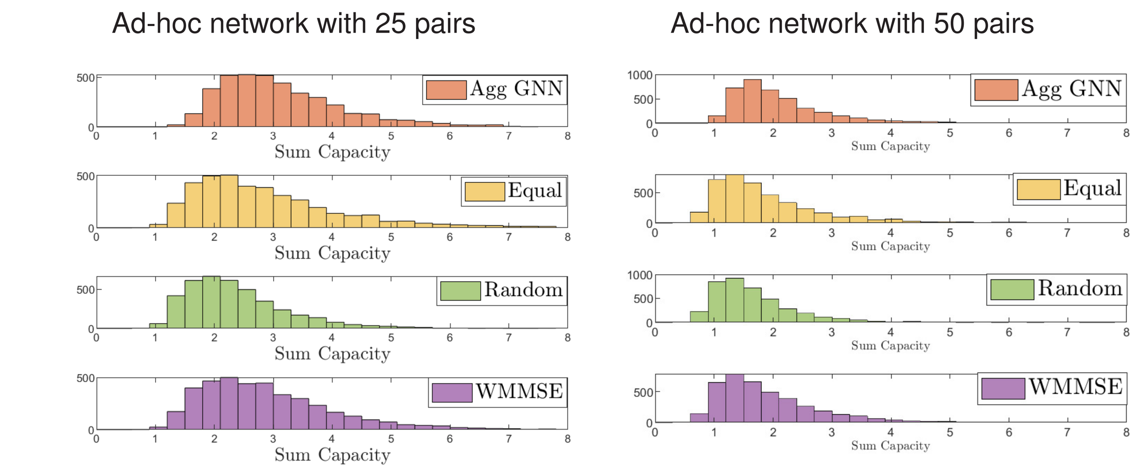


Architecture	$\epsilon = 0.2$	0.4
GNN2Ly	$7.37\% \pm 1.43\%$	$7.71\% \pm 3.96\%$
GF2Ly	$13.76\% \pm 6.82\%$	$13.54\% \pm 7.16\%$
Architecture	$\epsilon = 0.6$	0.8
GNN2Ly	$8.04\% \pm 2.83\%$	$11.01\% \pm 6.33\%$
GF2Ly	$14.76\% \pm 5.67\%$	$16.04\% \pm 6.34\%$

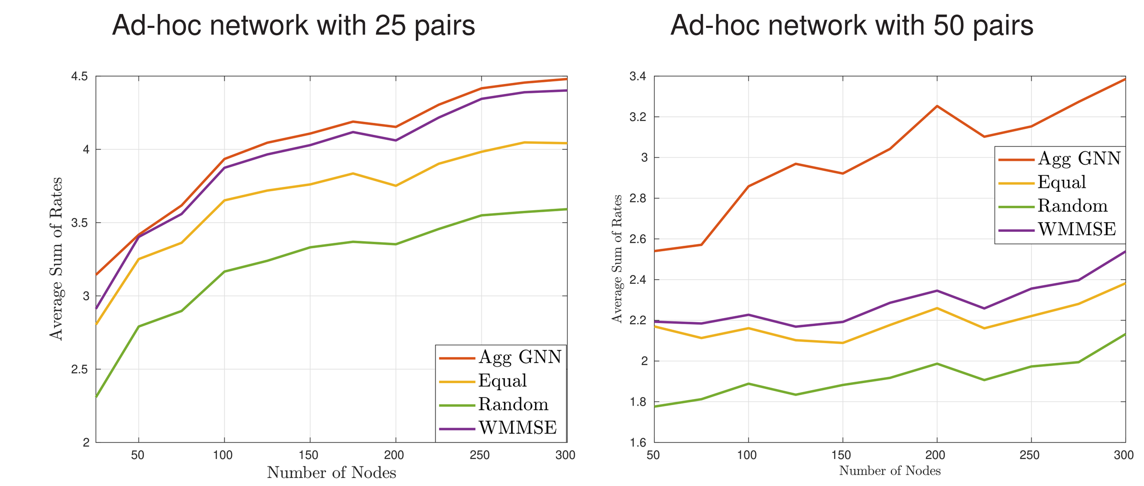
Z. Wang, L. Ruiz, and A. Ribeiro. "Stability to Deformations of Manifold Filters and Manifold Neural Networks." arXiv preprint arXiv:2106.03725 (2021).

Large-scale Wireless Power Allocation

- We test the trained GNN in other ad-hoc networks of fixed size and density
  - ⇒ The GNN remains optimal across permutations of ad-hoc networks



- We test in other networks of increasing size and fixed density
  - ⇒ The GNN transfers to larger ad-hoc networks with no need of retraining



Z. Wang, M. Eisen, and A. Ribeiro. "Learning decentralized wireless resource allocations with graph neural networks." IEEE Transactions on Signal Processing 70 (2022): 1850-1863.

Tangent Bundle Neural Networks

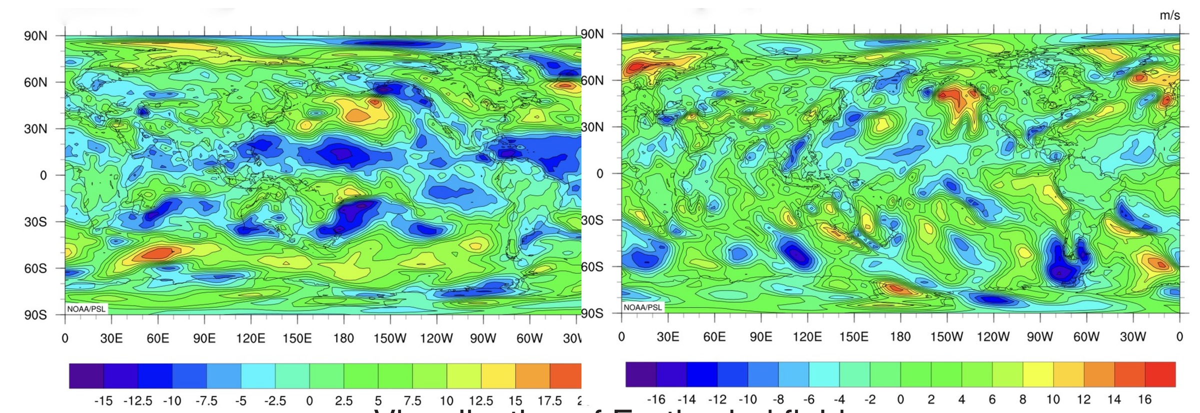
- MNNs process scalar signals over the manifold w/o covering vector fields
- We define Tangent Bundle convolution with the Connection Laplacian Δ
- The tangent bundle filter with impulse response $\tilde{h} : \mathbb{R}^+ \to \mathbb{R}$ is given by

$$\mathcal{G}(x) = \int_0^\infty \tilde{h}(t) e^{t\Delta} \mathcal{F}(x) \mathrm{d}t = \mathbf{h}(\Delta)\mathcal{F}(x).$$

- Connection Laplacian has spectral decomposition $\Delta\mathcal{F} = -\sum_{i=1}^{\infty} \lambda_i \langle \mathcal{F}, \phi_i \rangle \phi_i$
- Tangent bundle Fourier Transform is the projections $\Rightarrow \left[\hat{\mathcal{F}}\right]_i = \langle \mathcal{F}, \phi_i \rangle$
- Frequency response of filter h is $\Rightarrow \hat{h}(\lambda) = \int_0^\infty \tilde{h}(t) e^{-t\lambda} \mathrm{d}t$

Theorem (Tangent bundle Filters in the Spectral Domain)

Tangent bundle filters are pointwise in the spectral domain $\left[\hat{\mathcal{G}}\right]_i = \hat{h}(\lambda_i)\left[\hat{\mathcal{F}}\right]_i$



Visualization of Earth wind field

C. Battiloro, Z. Wang, H. Riess, P. Di Lorenzo and A. Ribeiro. "Tangent Bundle Convolutional Learning: from Manifolds to Cellular Sheaves and Back" arXiv preprint arXiv:2303.11323 (2023).