Scattering Networks on the Sphere for Scalable and Rotationally Equivariant Spherical CNNs

arXiv:2102.02828

Jason D. McEwen*, Chris Wallis & Augustine Mavor-Parker

* Corresponding author: jason.mcewen@kagenova.com, www.jasonmcewen.org

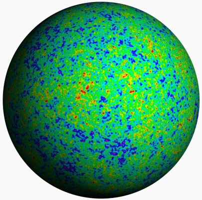
Kagenova Limited

www.kagenova.com

International Conference on Learning Representations (ICLR) 2022

Data on the sphere

Whenever observe over angles, recover data on 2D sphere (or 3D rotation group).



Cosmic microwave background

360° virtual reality

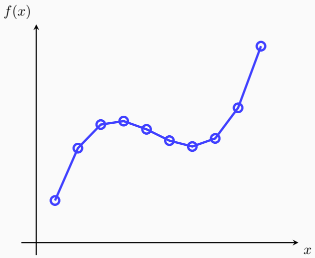
Construct CNNs natively on the sphere and encode rotational equivariance.

Problem: Rotationally equivariant spherical CNNs are not scalable

Categories of spherical CNN constructions

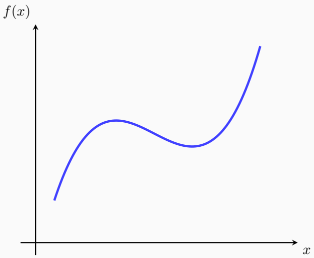
Discrete representations

- – Not rotationally equivariant
- + Scalable



Continuous representations

- + Rotationally equivariant
- – Not scalable



Solution: hybrid networks

Efficient generalized spherical CNN framework of Cobb et al. 2021 advocates hybrid networks, with different spherical layers leveraged alongside each other.

(Building on equivariant spherical CNNs of Cohen et al. 2018, Esteves et al. 2018, Kondor et al. 2018.)

Solution: hybrid networks

Efficient generalized spherical CNN framework of Cobb et al. 2021 advocates hybrid networks, with different spherical layers leveraged alongside each other.

(Building on equivariant spherical CNNs of Cohen et al. 2018, Esteves et al. 2018, Kondor et al. 2018.)

Introduce new initial layer, with following properties:

1. Scalable
2. Allow subsequent layers to operate at low-resolution (i.e. mixes information to low frequencies)
3. Rotationally equivariant
4. Stable and locally invariant representation (i.e. effective representation space)

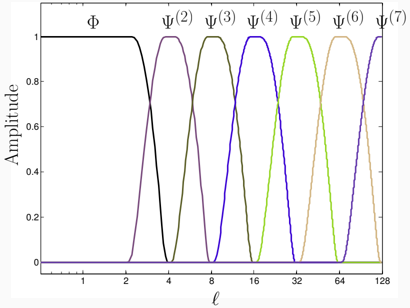
Scattering networks on the sphere

Wavelets on the sphere

Adopt scale-discretized wavelets on the sphere (e.g. McEwen et al. 2018, McEwen et al. 2015).

Wavelets $\psi_j \in \mathrm{L}^2(\mathbb{S}^2)$ capture high-frequency signal content at scale *j*.

Scaling function $\phi \in \mathrm{L}^2(\mathbb{S}^2)$ captures low-frequency content.



Tiling of harmonic line.

Wavelets on the sphere

Adopt scale-discretized wavelets on the sphere (e.g. McEwen et al. 2018, McEwen et al. 2015).

Wavelets $\psi_j \in \mathrm{L}^2(\mathbb{S}^2)$ capture high-frequency signal content at scale *j*.

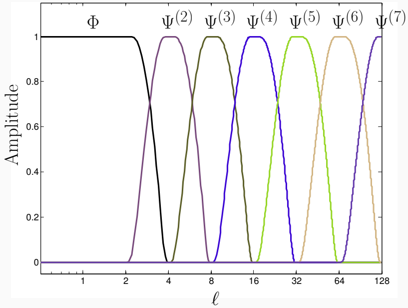
Scaling function $\phi \in \mathrm{L}^2(\mathbb{S}^2)$ captures low-frequency content.

Spherical wavelet transform given by

$$W_j(\omega) = (f \star \psi_j)(\omega) = \int_{\mathbb{S}^2} \mathrm{d}\mu(\omega') f(\omega') (R_\omega \psi_j)^*(\omega').$$

Spherical convolution

Rotated wavelet



Tiling of harmonic line.

Wavelets on the sphere

Adopt scale-discretized wavelets on the sphere (e.g. McEwen et al. 2018, McEwen et al. 2015).

Wavelets $\psi_j \in \mathrm{L}^2(\mathbb{S}^2)$ capture high-frequency signal content at scale *j*.

Scaling function $\phi \in \mathrm{L}^2(\mathbb{S}^2)$ captures low-frequency content.

Spherical wavelet transform given by

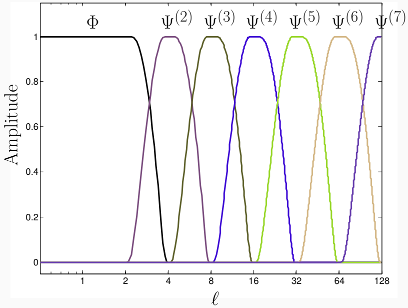
$$W_j(\omega) = (f \star \psi_j)(\omega) = \int_{\mathbb{S}^2} \mathrm{d}\mu(\omega') f(\omega') (R_\omega \psi_j)^*(\omega').$$

Spherical convolution

Rotated wavelet

Scalable since fast algorithms available

(e.g. McEwen et al. 2007, 2013, 2015)



Tiling of harmonic line.

Scattering transform on the sphere

Scattering on the sphere follows by direct analogue of Euclidian construction (Mallat 2012).

Scattering transform on the sphere

Scattering on the sphere follows by direct analogue of Euclidian construction (Mallat 2012).

Spherical scattering propagator for scale j:

$$U[j]f = |f \star \psi_j|.$$

Modulus function is adopted for the activation function since non-expansive. Acts to mix signal content to low frequencies.

Scattering transform on the sphere

Scattering on the sphere follows by direct analogue of Euclidian construction (Mallat 2012).

Spherical scattering propagator for scale j:

$$U[j]f = |f \star \psi_j|.$$

Modulus function is adopted for the activation function since non-expansive. Acts to mix signal content to low frequencies.

Spherical cascade of propagators:

$$U[p]f = |||f \star \psi_{j_1}| \star \psi_{j_2}| \dots \star \psi_{j_d}|,$$

for the path $p = (j_1, j_2, \dots, j_d)$ with depth d.

Scattering transform on the sphere

Scattering on the sphere follows by direct analogue of Euclidian construction (Mallat 2012).

Spherical scattering propagator for scale j:

$$U[j]f = |f \star \psi_j|.$$

Modulus function is adopted for the activation function since non-expansive. Acts to mix signal content to low frequencies.

Spherical cascade of propagators:

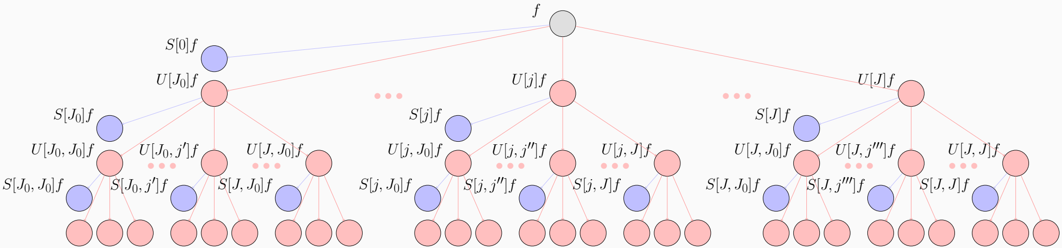
$$U[p]f = |||f \star \psi_{j_1}| \star \psi_{j_2}| \dots \star \psi_{j_d}|,$$

for the path $p = (j_1, j_2, \dots, j_d)$ with depth d.

Scattering coefficients:

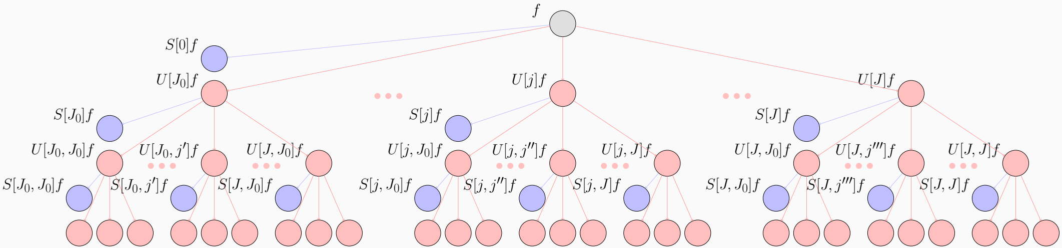
$$S[p]f = |||f \star \psi_{j_1}| \star \psi_{j_2}| \dots \star \psi_{j_d}| \star \phi.$$

Spherical scattering network is collection of scattering transforms for a number of paths: $\mathcal{S}_{\mathbb{P}} f = \{S[p]f : p \in \mathbb{P}\}$, where the general path set $\mathbb{P}$ denotes the infinite set of all possible paths $\mathbb{P} = \{p = (j_1, j_2, \dots, j_d) : J_0 \leq j_i \leq J, 1 \leq i \leq d,\ d \in \mathbb{N}_0\}$.



Scattering networks on the sphere

Spherical scattering network is collection of scattering transforms for a number of paths: $\mathcal{S}_{\mathbb{P}} f = \{S[p]f : p \in \mathbb{P}\}$, where the general path set $\mathbb{P}$ denotes the infinite set of all possible paths $\mathbb{P} = \{p = (j_1, j_2, \ldots, j_d) : J_0 \leq j_i \leq J, 1 \leq i \leq d,\ d \in \mathbb{N}_0\}$.



Scattering networks are rotationally equivariant (since the spherical wavelet transform and modulus operator are rotationally equivariant).

Isometric invariance and stability to diffeomorphisms

Theorem (Isometric Invariance)

Let $\zeta \in \mathrm{Isom}(\mathbb{S}^2)$*, then there exists a constant* C *such that for all* $f \in \mathrm{L}^2(\mathbb{S}^2)$*,*

$$\|\mathcal{S}_{\mathbb{P}_D} f - \mathcal{S}_{\mathbb{P}_D} V_\zeta f\|_2 \leq C L^{5/2} (D+1)^{1/2} \lambda^{J_0} \|\zeta\|_\infty \|f\|_2 .$$

Theorem (Stability to Diffeomorphisms)

Let $\zeta \in \mathrm{Diff}(\mathbb{S}^2)$*. If* $\zeta = \zeta_1 \circ \zeta_2$ *for some isometry* $\zeta_1 \in \mathrm{Isom}(\mathbb{S}^2)$ *and diffeomorphism* $\zeta_2 \in \mathrm{Diff}(\mathbb{S}^2)$*, then there exists a constant* C *such that for all* $f \in \mathrm{L}^2(\mathbb{S}^2)$*,*

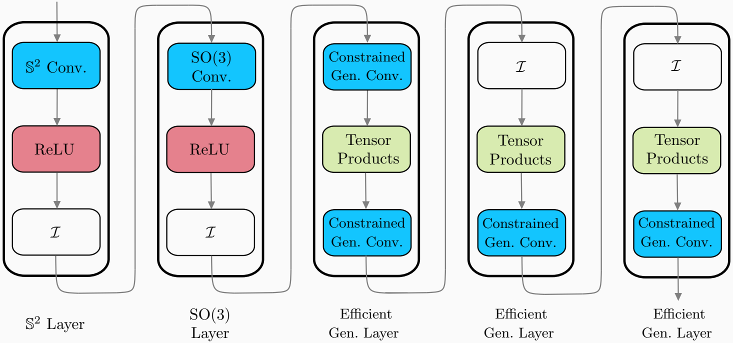
$$\|\mathcal{S}_{\mathbb{P}_D} f - \mathcal{S}_{\mathbb{P}_D} V_\zeta f\|_2 \leq C L^2 \left[L^2 \|\zeta_2\|_\infty + L^{1/2} (D+1)^{1/2} \lambda^{J_0} \|\zeta_1\|_\infty \right] \|f\|_2 .$$

(**Proof:** Follows by straightforward extension of proof of Perlmutter et al. 2020.)

Scalable and rotationally equivariant spherical CNNs

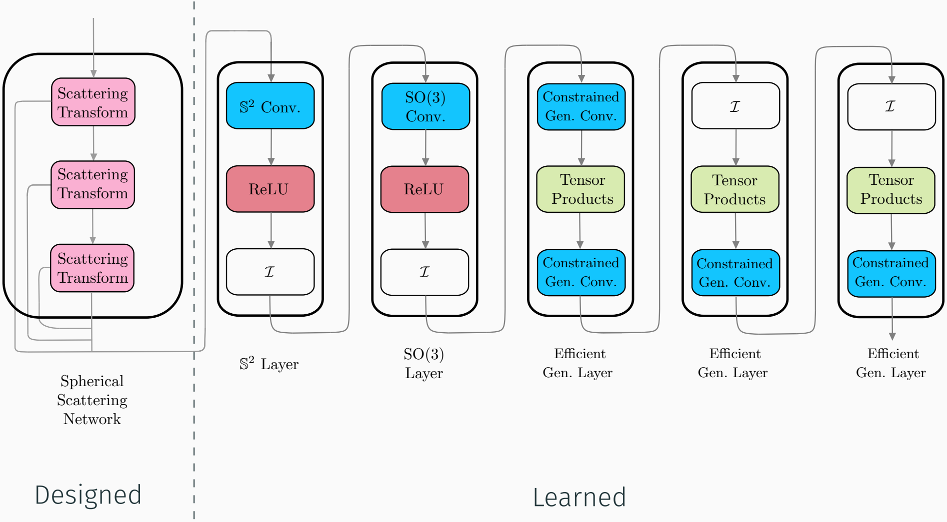
Hybrid spherical CNNs

Generalized spherical CNN framework efficient but remains computationally costly (Cobb et al. 2021).



Embrace hybrid approach and advocate spherical scattering network as initial layer.

Scalable and rotationally equivariant spherical CNNs



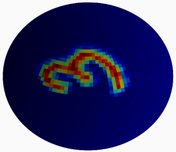
Numerical experiments

Rotational equivariance

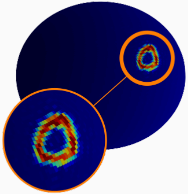
Path depth d	Equivariance Error		
	Minimum	Median	Maximum
0	0.00%	0.00%	0.00%
1	0.01%	0.05%	0.24%
2	0.18%	1.01%	5.36%
3	0.56%	3.47%	10.68%

Equivariance errors are considerably smaller that the spherical ReLU (which has error $\sim$ 35%).

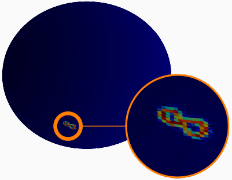
Spherical MNIST at varying resolution



$L = 64$



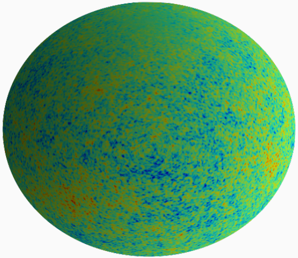
$L = 128$



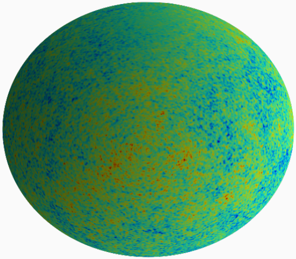
$L = 256$

L	Digit Size	Accuracy (NR/R) no scattering	Accuracy (NR/R) scattering
64	82.2°	88.66	97.22
128	42.5°	51.71	76.81
256	21.4°	17.23	59.48

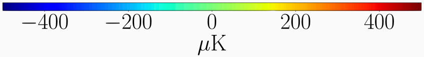
Gaussianity of the cosmic microwave background



Gaussian



Non-Gaussian



At $L = 1024$ ($\sim$2 million pixels), we achieve classification accuracy of:
53% without scattering network versus 95% with scattering network.

Summary

Scattering Networks on the Sphere for Scalable and Rotationally Equivariant Spherical CNNs
(McEwen et al. 2022; arXiv:2102.02828)

Spherical scattering networks:

1. Scalable (leverage fast wavelet transforms)
2. Allow subsequent layers to operate at low-resolution (i.e. mixes information to low frequencies)
3. Rotationally equivariant
4. Stable and locally invariant representation (i.e. effective representation space)

\+ Designed and not learned (can be considered as preprocessing with low storage requirements)

Can scale rotationally equivariant spherical CNNs to high-resolution data typical of many practical applications, with many tens of megapixels and beyond.

Code available on request at https://kagenova.com/products/fourpiAI/ or contact jason.mcewen@kagenova.com.