Geometric deep learning on the sphere

Scalable and equivariant spherical CNNs

Jason McEwen

www.jasonmcewen.org

Kagenova Limited

Mullard Space Science Laboratory (MSSL), UCL

In collaboration with:

Jeremy Ocampo · Matthew Price · Oliver Cobb · Chris Wallis · Augustine Mavor-Parker · Augustin Marignier · Mayeul d'Avezac

September 2022

Outline

Mentimeter

Give your input at `https://www.menti.com/puiqjn97i9`.

Or go to `https://www.menti.com` and enter voting code: 80 20 22 0.

Symmetry in deep learning

Physics and deep learning

Physics

Understanding the world by
modelling from first principles
for generative models and inference.

Deep Learning

Understanding the world by
learning informative representations
for generative models and inference.

Physics and deep learning

Physics

Understanding the world by
modelling from first principles
for generative models and inference.

Hard!

Deep Learning

Understanding the world by
learning informative representations
for generative models and inference.

Hard!

Physics and deep learning

Physics $\longleftrightarrow$ Deep Learning

Physics and deep learning

Physics $\longleftrightarrow$ Deep Learning

Here we focus on integrating physics $\rightarrow$ deep learning

(in other works focus on reverse: physics $\leftarrow$ deep learning).

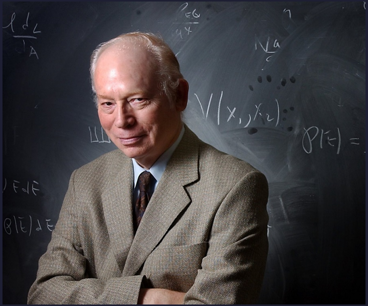
Physics and deep learning

Physics $\longleftrightarrow$ Deep Learning

Here we focus on integrating physics $\rightarrow$ deep learning

(in other works focus on reverse: physics $\leftarrow$ deep learning).

As we will see, this key factor driving the deep learning revolution.



"Symmetry: key to nature's secrets."

— Steven Weinberg

Mirror symmetry

Mirror symmetry

Symmetry

Mirror symmetry

Rotational symmetry

Symmetry

Mirror symmetry

Rotational symmetry

Symmetry

Mirror symmetry

Rotational symmetry

Symmetry

Mirror symmetry

Rotational symmetry

Symmetry

Mirror symmetry

Rotational symmetry

Symmetry (invariance) to continuous transformation

In physics we typically consider **continuous symmetries**, where system is symmetric (invariant) to continuous transformation.

Spatial translation

Symmetry (invariance) to continuous transformation

In physics we typically consider **continuous symmetries**, where system is symmetric (invariant) to continuous transformation.

Spatial translation

Symmetry (invariance) to continuous transformation

In physics we typically consider **continuous symmetries**, where system is symmetric (invariant) to continuous transformation.

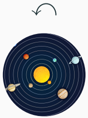
Spatial translation

Rotation

Symmetry (invariance) to continuous transformation

In physics we typically consider **continuous symmetries**, where system is symmetric (invariant) to continuous transformation.

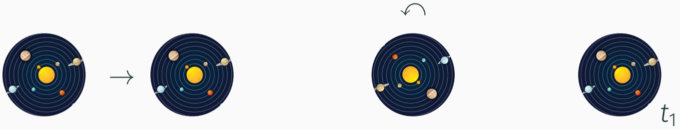
Spatial translation



Rotation

Symmetry (invariance) to continuous transformation

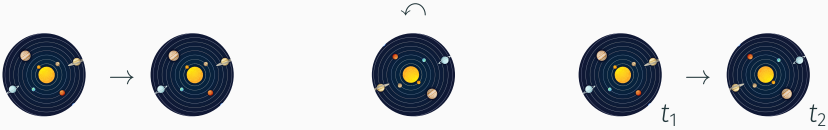
In physics we typically consider **continuous symmetries**, where system is symmetric (invariant) to continuous transformation.



Spatial translation | Rotation | Time translation

Symmetry (invariance) to continuous transformation

In physics we typically consider **continuous symmetries**, where system is symmetric (invariant) to continuous transformation.



Spatial translation Rotation Time translation

Noether's theorem

Noether's theorem

For every continuous symmetry of the universe, there exists a conserved quantity.

Emmy Noether

Noether's theorem

Noether's theorem

For every continuous symmetry of the universe, there exists a conserved quantity.

Symmetries at the heart of physics:

- **Translational** symmetry ⇔ conservation of **momentum**
- **Rotational** symmetry ⇔ conservation of **angular momentum**
- **Time translational** symmetry ⇔ conservation of **energy**

Emmy Noether

Symmetry is the foundation underlying the fundamental laws of physics.

Symmetry in deep learning

Encoding symmetry in deep learning models captures fundamental properties about the underlying nature of our world.

Symmetry in deep learning

Encoding symmetry in deep learning models captures fundamental properties about the underlying nature of our world.

Key factor driving the deep learning revolution, with the advent of CNNs.

- CNNs resulted in a step-change in performance.
- Convolutional structure of CNNs capture translational symmetry (i.e. translational equivariance).

Equivariance

Equivariance

An operator $\mathcal{A}$ is equivariant to a transformation $\mathcal{T}$ if

$$\mathcal{T}\big(\mathcal{A}(f)\big) = \mathcal{A}\big(\mathcal{T}(f)\big)$$

for all possible signals f.

Transforming the signal after application of the operator, is equivalent to transformation of the signal first, followed by application of the operator.

Equivariance

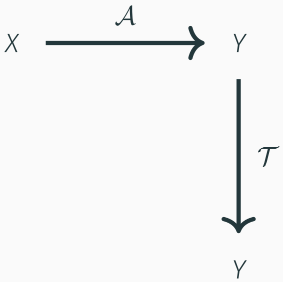
Equivariance

An operator $\mathcal{A}$ is equivariant to a transformation $\mathcal{T}$ if

$$\mathcal{T}\big(\mathcal{A}(f)\big) = \mathcal{A}\big(\mathcal{T}(f)\big)$$

for all possible signals f.

Transforming the signal after application of the operator, is equivalent to transformation of the signal first, followed by application of the operator.



Equivariance

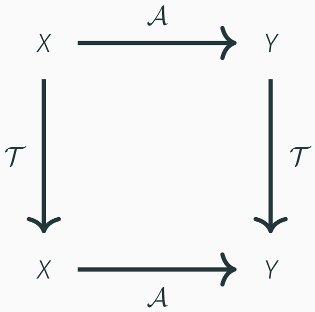
Equivariance

An operator $\mathcal{A}$ is equivariant to a transformation $\mathcal{T}$ if

$$\mathcal{T}\big(\mathcal{A}(f)\big) = \mathcal{A}\big(\mathcal{T}(f)\big)$$

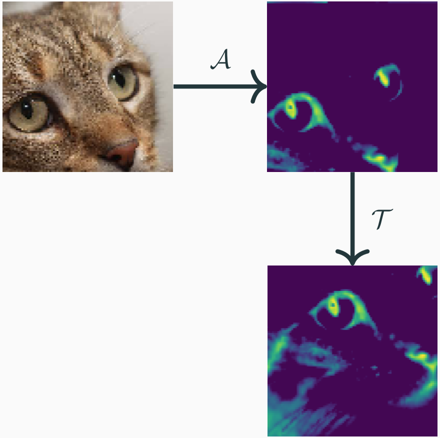
for all possible signals f.

Transforming the signal after application of the operator, is equivalent to transformation of the signal first, followed by application of the operator.



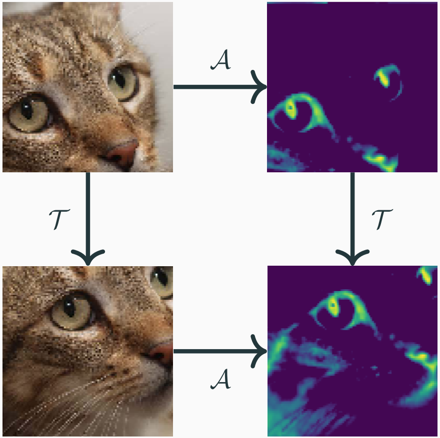
Planar (Euclidean) CNNs exhibit translational equivariance

Planar (Euclidean) convolution is translationally equivariant.



Planar (Euclidean) CNNs exhibit translational equivariance

Planar (Euclidean) convolution is translationally equivariant.



Importance of equivariance

Imposing inductive biases in deep learning models, such as equivariance to symmetry transformations, allows models to be learned in a more principled and effective manner.

Capture fundamental physical understanding of generative process.

Importance of equivariance

In some sense, equivariance to a transformation means a pattern need only be learnt once, and may then be recognised in all transformed scenarios.

Importance of equivariance

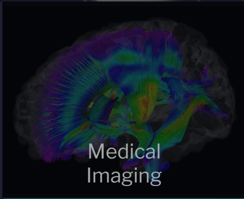
In some sense, equivariance to a transformation means a pattern need only be learnt once, and may then be recognised in all transformed scenarios.

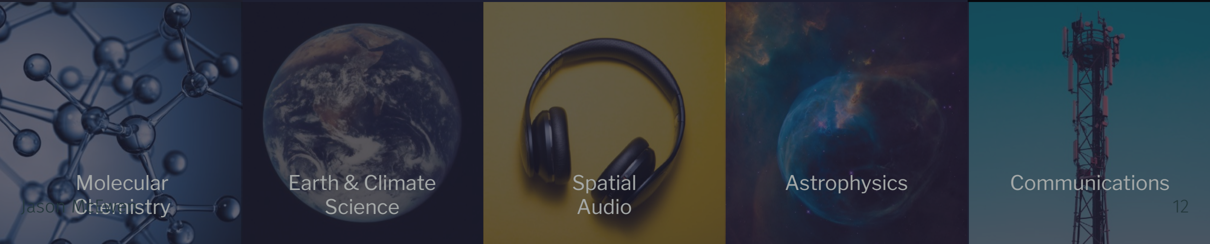
Cat

Still a cat

Geometric deep learning on the sphere

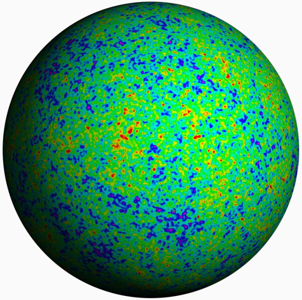
Data on the sphere arises in many applications





Cosmology and virtual reality

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

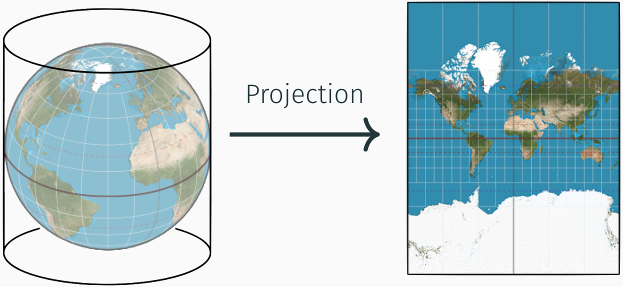


Cosmic microwave background

360° virtual reality

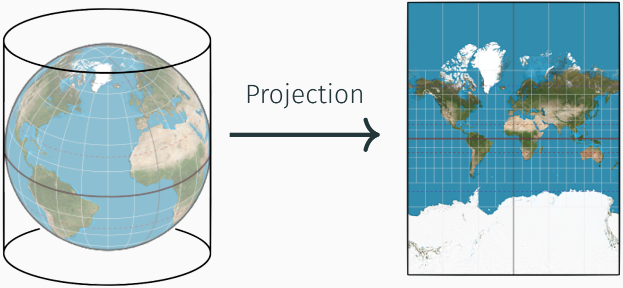
Why not standard (Euclidean) deep learning approaches?

Could project sphere to plane and then apply standard planar CNNs.



Why not standard (Euclidean) deep learning approaches?

Could project sphere to plane and then apply standard planar CNNs.



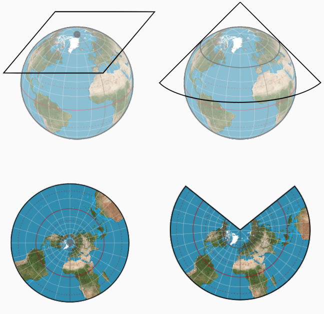
Greenland appears to be a similar size to Africa in the projected planar map, whereas it is over 10 times smaller.

Why not standard (Euclidean) deep learning approaches?

Projection breaks symmetries and geometric properties of sphere.

Why not standard (Euclidean) deep learning approaches?

Projection breaks symmetries and geometric properties of sphere.



No projection of the sphere to the plane can preserve both shapes and areas
$\Rightarrow$ distortions are unavoidable.

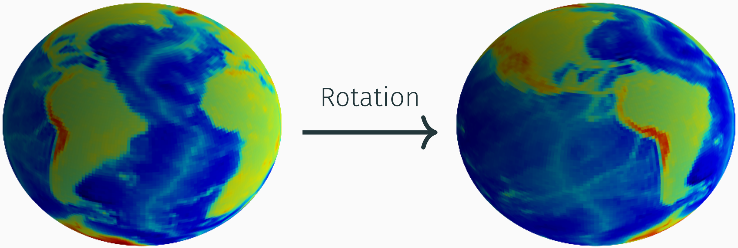
(Formally: a conformal, area-preserving projection does not exist.)

Goals of geometric deep learning on the sphere

1. **Capture geometry and symmetry** of the sphere (rotational equivariance)
2. **Computationally scalable** to support high-resolution data

Rotational equivariance

On the sphere, the analog of translations are rotations.



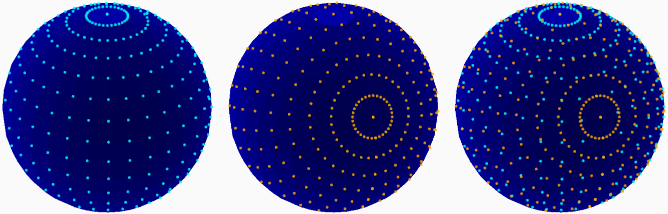
Would like spherical CNNs to exhibit rotational equivariance.

(Just as planar CNNs exhibit translational equivariance.)

Capturing rotational equivariance in spherical CNNs

Well-known that regular discretisation of the sphere does not exist (e.g. Tegmark 1996).

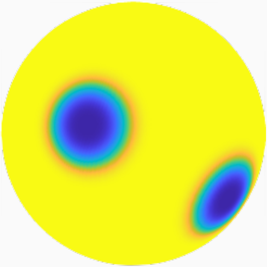
$\Rightarrow$ Not possible to discretise sphere in a manner that is invariant to rotations.



Capturing strict equivariance with operations defined directly in discretised (pixel) space **not possible** due to structure of sphere.

Categorisation of spherical CNNs frameworks

Continuous

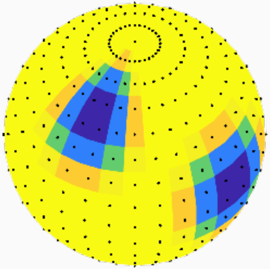


✔ Equivariant

✖ Not Scalable

(Cohen et al. 2018; Esteves et al. 2018; Kondor et al. 2018; Cobb et al. 2021; McEwen et al. 2022)

Discrete



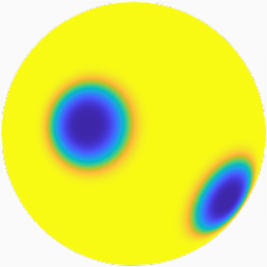
✖ Not Equivariant

✔ Scalable

(Jiang et al. 2019, Zhang et al. 2019, Perraudin et al. 2019, Cohen et al. 2019)

Categorisation of spherical CNNs frameworks

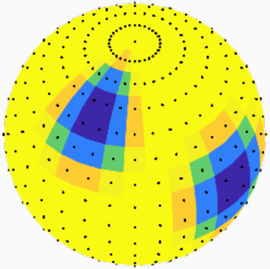
Continuous



✔ Equivariant
✖ Not Scalable

(Cohen et al. 2018; Esteves et al. 2018; Kondor et al. 2018; Cobb et al. 2021; McEwen et al. 2022)

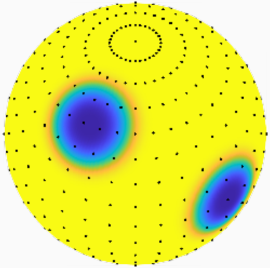
Discrete



✖ Not Equivariant
✔ Scalable

(Jiang et al. 2019, Zhang et al. 2019, Perraudin et al. 2019, Cohen et al. 2019)

Discrete-Continuous (DISCO)



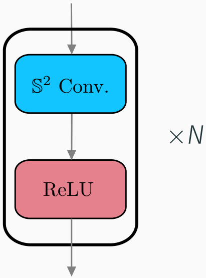
✔ Equivariant
✔ Scalable

(Ocampo, Price & McEwen, in prep.)

Continuous spherical CNNs

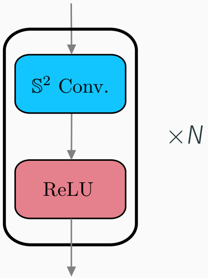
Spherical CNN

Spherical CNNs constructed by analog of Euclidean CNNs but using convolution on the sphere and with pointwise non-linear activations functions, e.g. ReLU (Cohen et al. 2018; Esteves et al. 2018).



Spherical CNN

Spherical CNNs constructed by analog of Euclidean CNNs but using convolution on the sphere and with pointwise non-linear activations functions, e.g. ReLU (Cohen et al. 2018; Esteves et al. 2018).



Consider Fourier representation → access to underlying continuous representations.

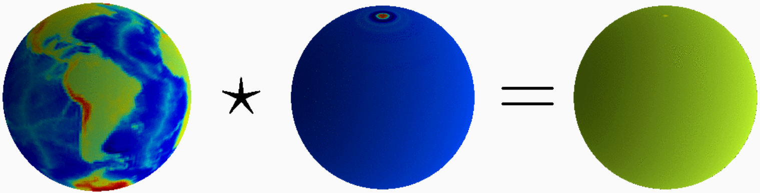
Convolution of signals on the sphere

Convolution of signals in spatial domain

Convolution of two signals $f, \psi \in L^2(\mathbb{S}^2)$ is given by

$$(f \star \psi)(\rho) = \langle f, R\psi \rangle = \int_{\mathbb{S}^2} \mathrm{d}\mu(\omega)\, f(\omega)\, \psi^*(\rho^{-1}\omega), \qquad \text{for } \omega \in \mathbb{S}^2, \rho \in \mathrm{SO}(3),$$

where $d\mu(\omega)$ denotes the Haar measure on $\mathbb{S}^2$ and $\cdot^*$ complex conjugation.



Convolution of signals on the sphere

Since sphere is compact manifold, Fourier space is discrete and **sampling theorems** can be leveraged to compute Fourier representations exactly for bandlimited signals (e.g. McEwen & Wiaux 2011).

$\Rightarrow$ Provides access to underlying continuous signals and symmetries of sphere.

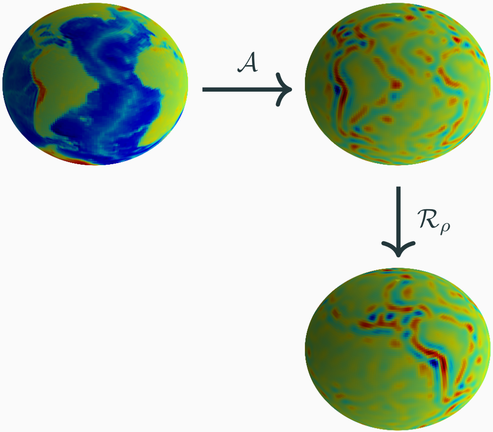
Convolution of signals in harmonic domain

Convolution of two signals $f, \psi \in L^2(\mathbb{S}^2)$ can be computed as a product in harmonic space:

$$\widehat{(f \star \psi)}^{\ell} = \hat{f}^{\ell} \, \hat{\psi}^{\ell *} .$$

Convolution is rotationally equivariant (when computed in harmonic domain):

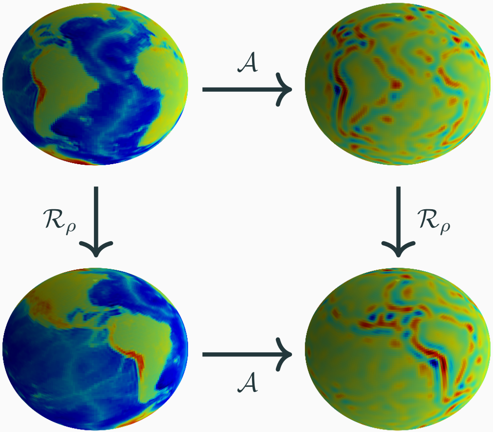
$$((\mathcal{R}_\rho f) \star \psi)(\rho') = (\mathcal{R}_\rho (f \star \psi))(\rho').$$



Convolution is rotationally equivariant

Convolution is rotationally equivariant (when computed in harmonic domain):

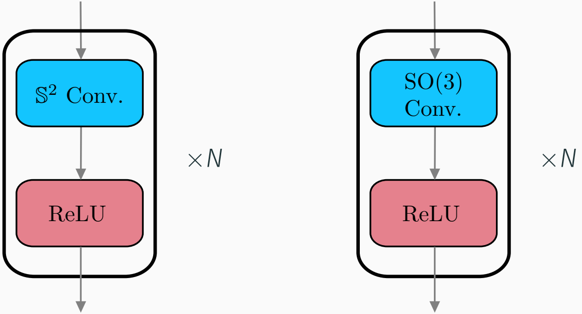
$$((\mathcal{R}_\rho f) \star \psi)(\rho') = (\mathcal{R}_\rho (f \star \psi))(\rho').$$



Spherical CNNs

Approach taken by Cohen et al. (2018) and Esteves et al. (2018).

Compute non-linear activation pixel-wise in spatial domain.



Harmonic space tensor product activations

Given two harmonic fragments $\hat{f}^{\ell_1}$ and $\hat{f}^{\ell_2}$, then

$$(C^{\ell_1,\ell_2,\ell})^\top(\hat{f}^{\ell_1} \otimes \hat{f}^{\ell_2}),$$

where $C^{\ell_1,\ell_2,\ell}$ are Clebsch-Gordan coefficients, which is non-linear in f and rotationally equivariant (Kondor et al. 2018).

Efficient generalized spherical CNNs

Consider the s-th layer of a generalized spherical CNN to take the form of a triple (Cobb et al. 2021; arXiv:2010.11661)

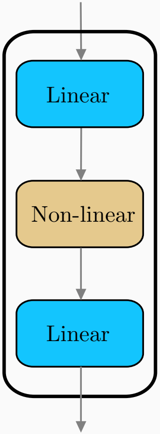
$$\mathcal{A}^{(s)} = (\mathcal{L}_1, \mathcal{N}, \mathcal{L}_2),$$

such that

$$\mathcal{A}^{(s)}(f^{(s-1)}) = \mathcal{L}_2(\mathcal{N}(\mathcal{L}_1(f^{(s-1)}))),$$

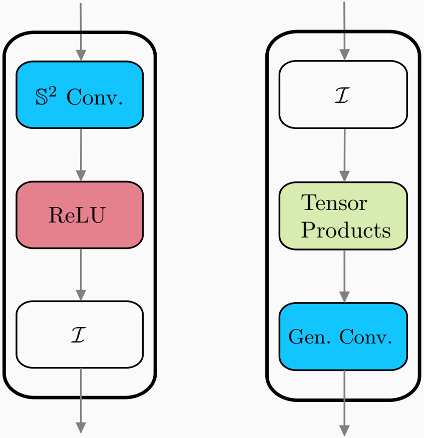
where

- $\mathcal{L}_1, \mathcal{L}_2 : \mathcal{F}_L \rightarrow \mathcal{F}_L$ are spherical convolution operators,
- $\mathcal{N} : \mathcal{F}_L \rightarrow \mathcal{F}_L$ is a non-linear, spherical activation operator.



Efficient generalised spherical CNNs

- Build on other **influential equivariant spherical CNN** constructions:
    - Cohen et al. (2018)
    - Esteves et al. (2018)
    - Kondor et al. (2018)
- Encompass other frameworks as special cases.
- General framework supports hybrids models.
- Significant efficiency improvements.



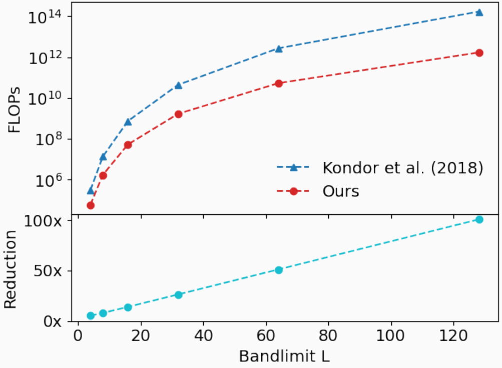
Cohen et al. (2018), Esteves et al. (2018)

Kondor et al. (2018)

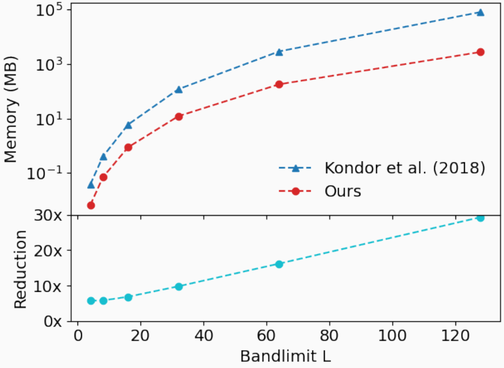
Contributions to improve efficiency

1. Channel-wise structure
2. Constrained generalized convolutions
3. Optimized degree mixing sets
4. Efficient sampling theory on the sphere and rotation group
   (McEwen & Wiaux 2011; McEwen et al. 2015)

Computational cost and memory requirements



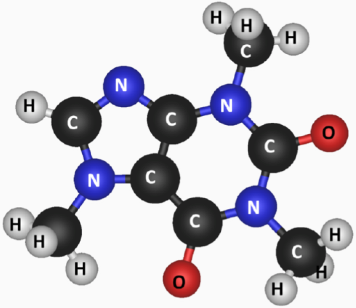
Computational cost



Memory requirements

Atomization energy prediction: problem

Predict atomization energy of molecule give the atom charges and positions.



Atomization energy prediction: results

Test root mean squared (RMS) error for QM7 regression problem

	RMS	Params
Montavon et al. 2012	5.96	-
Cohen et al. 2018	8.47	1.4M
Kondor et al. 2018	7.97	>1.1M
Ours (MST)	**3.16**	337k
Ours (RMST)	3.46	**335k**

Despite the efficient generalized approach

such equivariant spherical CNNs are limited to low-resolution data.

Scattering networks on the sphere

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

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)

$\Rightarrow$ Scattering networks on the sphere (McEwen et al. 2022; arXiv:2102.02828)

Scattering transform on the sphere

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

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

Scattering transform on the sphere

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

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

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).

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

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).

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

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:

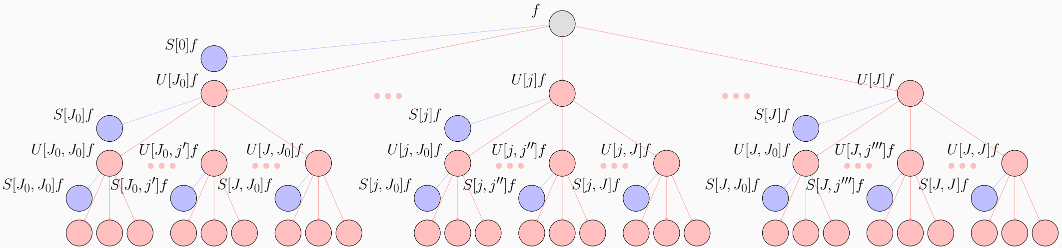
$$U[p]f = |||f \star \psi_{j_1}| \star \psi_{j_2}| \dots \star \psi_{j_d}|,$$

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

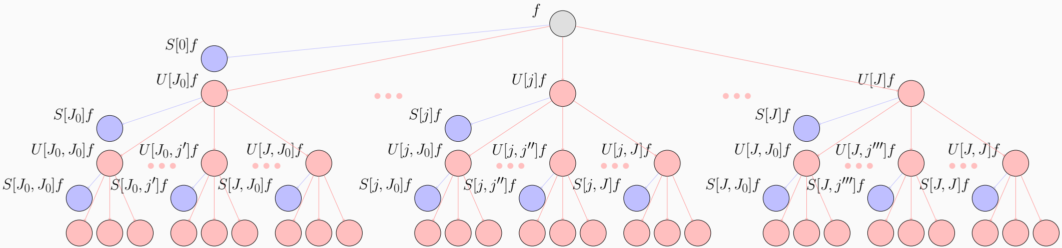
Scattering coefficients:

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

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 $\mathbb{P} = \{p = (j_1, j_2, \ldots, j_d) : J_0 \leq j_i \leq J, 1 \leq i \leq d, \; d \in \mathbb{N}_0\}$.

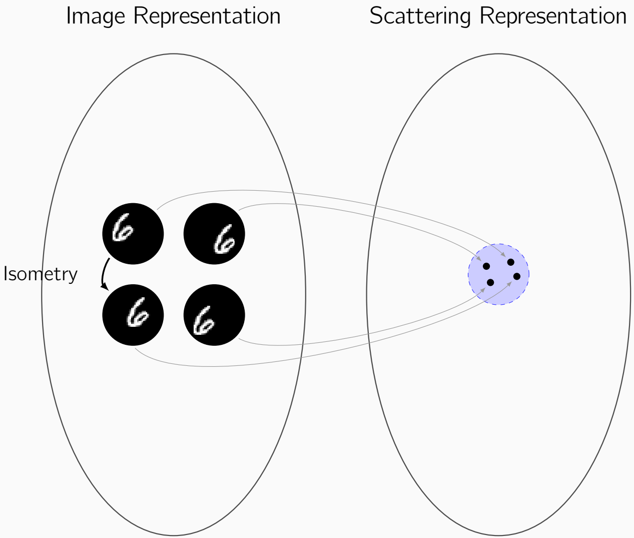


Spherical scattering network is collection of scattering transforms for a number of paths: $\mathcal{S}_{\mathbb{P}} f = \{S[p]f : p \in \mathbb{P}\}$, where the general path set $\mathbb{P}$ denotes the infinite set of all possible paths $\mathbb{P} = \{p = (j_1, j_2, \ldots, j_d) : J_0 \leq j_i \leq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



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

Isometric invariance



Isometric invariance

Theorem (Isometric Invariance)

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

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

Difference in representation

Scattering network representation is invariant to isometries up to a scale.

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

Isometric invariance

Theorem (Isometric Invariance)

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

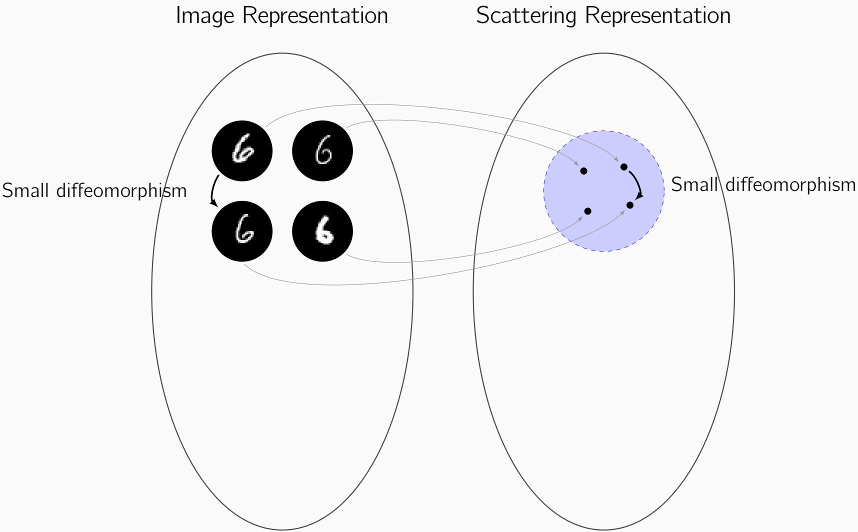
$$\|\mathcal{S}_{\mathbb{P}_D} f - \mathcal{S}_{\mathbb{P}_D} V_\zeta f\|_2 \leq C L^{5/2} (D+1)^{1/2} \lambda^{J_0} \|\zeta\|_\infty \|f\|_2 .$$

Difference in representation

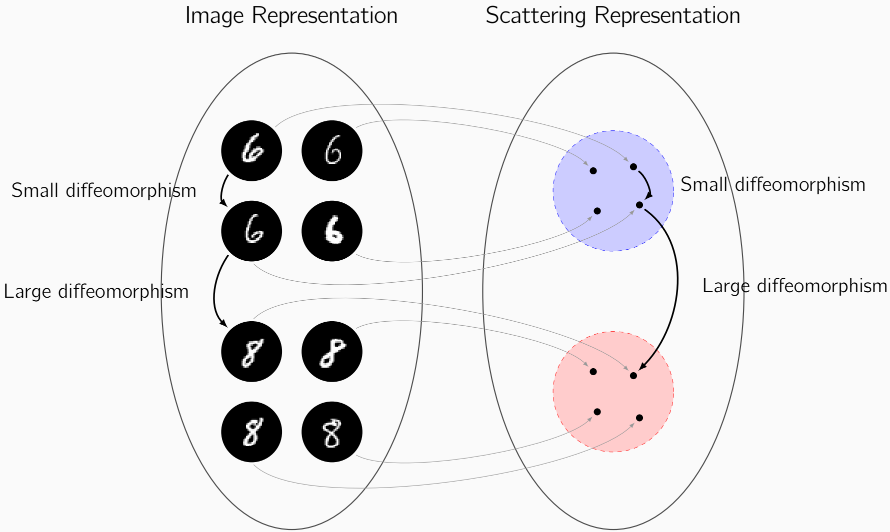
Scattering network representation is invariant to isometries up to a scale.

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

Stability to diffeomorphisms



Stability to diffeomorphisms



Stability to diffeomorphisms

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* $f \in \mathrm{L}^2(\mathbb{S}^2)$,

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

Difference in representation

Scattering network representation is stable to small diffeomorphisms about isometry.

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

Stability to diffeomorphisms

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* $f \in \mathrm{L}^2(\mathbb{S}^2)$,

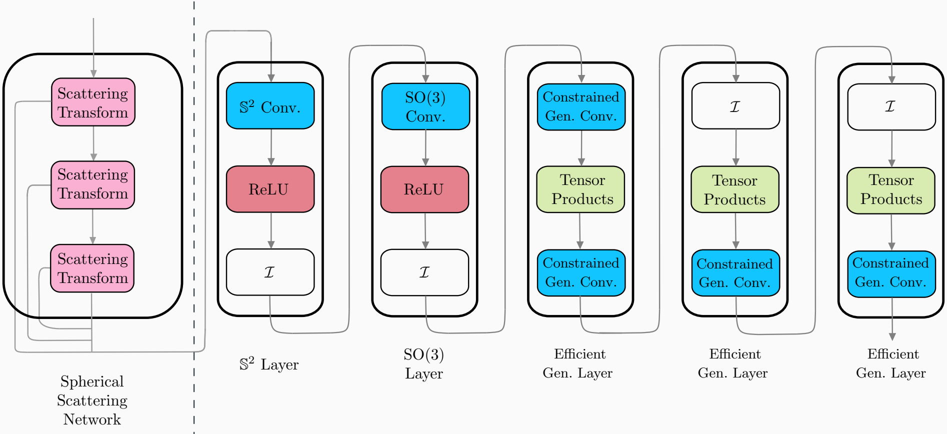
$$\|\mathcal{S}_{\mathbb{P}_D} f - \mathcal{S}_{\mathbb{P}_D} V_\zeta f\|_2 \leq C L^2 \left[L^2 \, \|\zeta_2\|_\infty + L^{1/2} (D+1)^{1/2} \lambda^{J_0} \, \|\zeta_1\|_\infty \right] \|f\|_2 .$$

Difference in representation

Scattering network representation is stable to small diffeomorphisms about isometry.

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

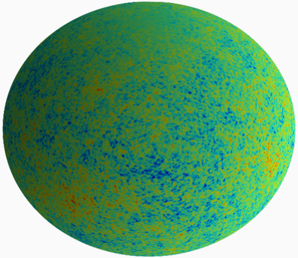
Scalable and rotationally equivariant spherical CNNs



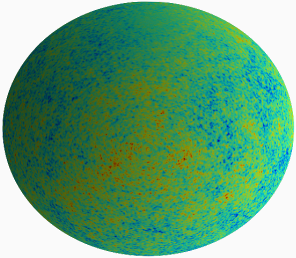
Designed

Learned

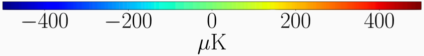
Gaussianity of the cosmic microwave background



Gaussian



Non-Gaussian



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

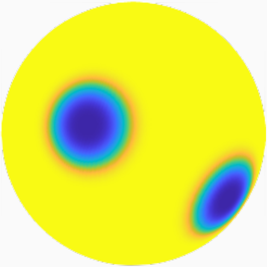
While spherical scattering networks help to scale to high-resolution input data,

high-resolution outputs for dense predictions are not supported.

Discrete-continuous spherical CNNs

Categorisation of spherical CNNs frameworks

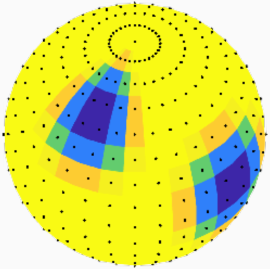
Continuous



✔ Equivariant
✖ Not Scalable

(Cohen et al. 2018; Esteves et al. 2018; Kondor et al. 2018; Cobb et al. 2021; McEwen et al. 2022)

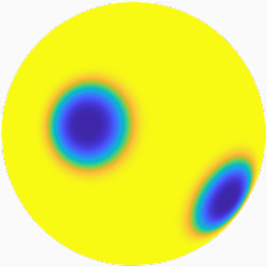
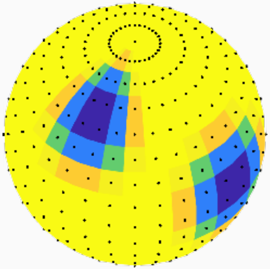
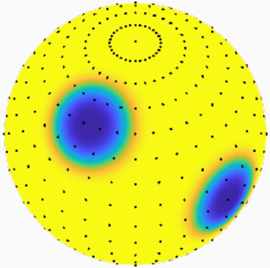
Discrete



✖ Not Equivariant
✔ Scalable

(Jiang et al. 2019, Zhang et al. 2019, Perraudin et al. 2019, Cohen et al. 2019)

Categorisation of spherical CNNs frameworks

Continuous	Discrete	Discrete-Continuous (DISCO)
		
✓ Equivariant ✗ Not Scalable	✗ Not Equivariant ✓ Scalable	✓ Equivariant ✓ Scalable
(Cohen et al. 2018; Esteves et al. 2018; Kondor et al. 2018; Cobb et al. 2021; McEwen et al. 2022)	(Jiang et al. 2019, Zhang et al. 2019, Perraudin et al. 2019, Cohen et al. 2019)	(Ocampo, Price & McEwen, in prep.)

Discrete-continuous (DISCO) spherical convolution

Follows by a careful hybrid representation of the spherical convolution:

- some components left continuous, to facilitate accurate rotational equivariance;
- while other components are discretized, to yield scalable computation.

Discrete-continuous (DISCO) spherical convolution

Follows by a careful hybrid representation of the spherical convolution:

- some components left continuous, to facilitate accurate rotational equivariance;
- while other components are discretized, to yield scalable computation.

DISCO spherical convolution

Spherical convolution can be carefully approximated by the DISCO representation

$$(f \star \psi)(R) = \int_{\mathbb{S}^2} f(\omega)\psi(R^{-1}\omega)\mathrm{d}\omega \approx \sum_i f[\omega_i]\psi(R^{-1}\omega_i)\delta\omega_i,$$

where, for now, we consider 3D rotations $R \in \mathrm{SO}(3)$.

Restricting rotations to SO(3)/SO(2)

While the DISCO spherical convolution is already efficient, we seek further computational savings by reducing the space of rotations to SO(3)/SO(2).

Rotations restricted to the quotient space may be written $R = Z(\alpha)Y(\beta) \in \text{SO}(3)/\text{SO}(2)$.

Restricting rotations to SO(3)/SO(2)

While the DISCO spherical convolution is already efficient, we seek further computational savings by reducing the space of rotations to SO(3)/SO(2).

Rotations restricted to the quotient space may be written $R = Z(\alpha)Y(\beta) \in \mathrm{SO}(3)/\mathrm{SO}(2)$.

Denote corresponding SO(3)/SO(2) convolution by $f \circledast \psi$ to avoid confusion.

Restricting rotations to SO(3)/SO(2)

While the DISCO spherical convolution is already efficient, we seek further computational savings by reducing the space of rotations to SO(3)/SO(2).

Rotations restricted to the quotient space may be written $R = Z(\alpha)Y(\beta) \in \text{SO}(3)/\text{SO}(2)$.

Denote corresponding SO(3)/SO(2) convolution by $f \circledast \psi$ to avoid confusion.

Analogous to Euclidean planar CNNs, where filters are translated across the image but are *not* rotated in the plane.

Restricting rotations to SO(3)/SO(2)

While the DISCO spherical convolution is already efficient, we seek further computational savings by reducing the space of rotations to SO(3)/SO(2).

Rotations restricted to the quotient space may be written $R = Z(\alpha)Y(\beta) \in \mathrm{SO}(3)/\mathrm{SO}(2)$.

Denote corresponding SO(3)/SO(2) convolution by $f \circledast \psi$ to avoid confusion.

Analogous to Euclidean planar CNNs, where filters are translated across the image but are *not* rotated in the plane.

However, as the space SO(3)/SO(2) is **not a group**, when restricting rotations in this manner important differences to the usual setting arise since we no longer have a group convolution.

SO(3) rotational equivariance

DISCO spherical convolution $f \star \psi$ for rotations $Q, R \in \mathrm{SO}(3)$ satisfies SO(3) rotational equivariance:

$$
\begin{aligned}
((\mathcal{Q}f) \star \psi)(R) &\approx \sum_i (\mathcal{Q}f)[\omega_i] \psi(R^{-1}\omega_i) \delta\omega_i \\
&= \sum_i f[\omega_i] \psi((Q^{-1}R)^{-1}\omega_i) \delta\omega_i \\
&\overset{(**)}{\approx} (f \star \psi)(Q^{-1}R) \\
&= (\mathcal{Q}(f \star \psi))(R).
\end{aligned}
$$

SO(3) rotational equivariance

DISCO spherical convolution $f \star \psi$ for rotations $Q, R \in \mathrm{SO}(3)$ satisfies SO(3) rotational equivariance:

$$
\begin{aligned}
((\mathcal{Q}f) \star \psi)(R) &\approx \sum_i (\mathcal{Q}f)[\omega_i]\psi(R^{-1}\omega_i)\delta\omega_i \\
&= \sum_i f[\omega_i]\psi((Q^{-1}R)^{-1}\omega_i)\delta\omega_i \\
&\overset{(**)}{\approx} (f \star \psi)(Q^{-1}R) \\
&= (\mathcal{Q}(f \star \psi))(R).
\end{aligned}
$$

Note that step $(**)$ only holds since SO(3) exhibits a group structure and so $Q^{-1}R \in \mathrm{SO}(3)$.

Asymptotic SO(3)/SO(2) rotational equivariance

DISCO spherical convolution $f \circledast \psi$ for rotations $Q, R \in \mathrm{SO}(3)/\mathrm{SO}(2)$ **does not satisfy SO(3) or SO(3)/SO(2) rotational equivariance** (since SO(3)/SO(2) is not a group).

Asymptotic SO(3)/SO(2) rotational equivariance

DISCO spherical convolution $f \circledast \psi$ for rotations $Q, R \in \mathrm{SO}(3)/\mathrm{SO}(2)$ **does not satisfy SO(3) or SO(3)/SO(2) rotational equivariance** (since SO(3)/SO(2) is not a group).

But DISCO spherical convolution $f \circledast \psi$ does satisfy asymptotic SO(3)/SO(2) equivariance.

Recover asymptotic SO(3)/SO(2) rotational equivariance as $\beta \to 0$, for $Q = Z(\alpha)Y(\beta)$.

Asymptotic SO(3)/SO(2) rotational equivariance

DISCO spherical convolution $f \circledast \psi$ for rotations $Q, R \in \text{SO}(3)/\text{SO}(2)$ **does not satisfy SO(3) or SO(3)/SO(2) rotational equivariance** (since SO(3)/SO(2) is not a group).

But DISCO spherical convolution $f \circledast \psi$ does satisfy asymptotic SO(3)/SO(2) equivariance.

Recover asymptotic SO(3)/SO(2) rotational equivariance as $\beta \to 0$, for $Q = Z(\alpha)Y(\beta)$.

Asymptotic SO(3)/SO(2) equivariance of significant practical use since content in spherical signals often orientated and similar content appears at similar latitudes, particularly for 360° panoramic photos and video.

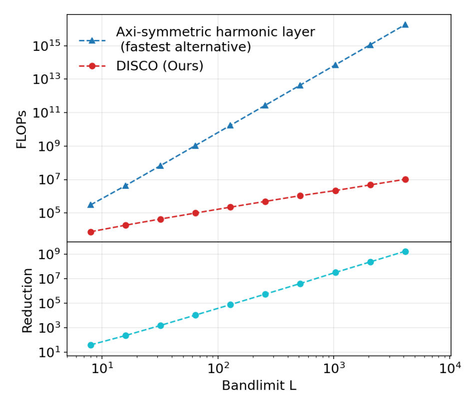
Computationally scalable DISCO spherical convolution

DISCO convolution affords a computationally scalable implementation.

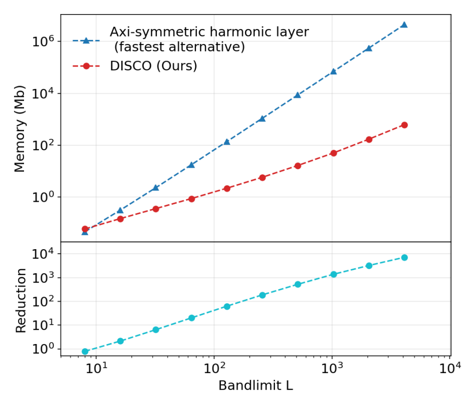
1. Spare tensor representation.
2. Memory compression.
3. Custom sparse gradients.

Linear scaling in number of pixels on the sphere $O(N) = O(L^2)$ for both computational cost and memory usage.

Computational cost and memory requirements



Computational cost



Memory requirements

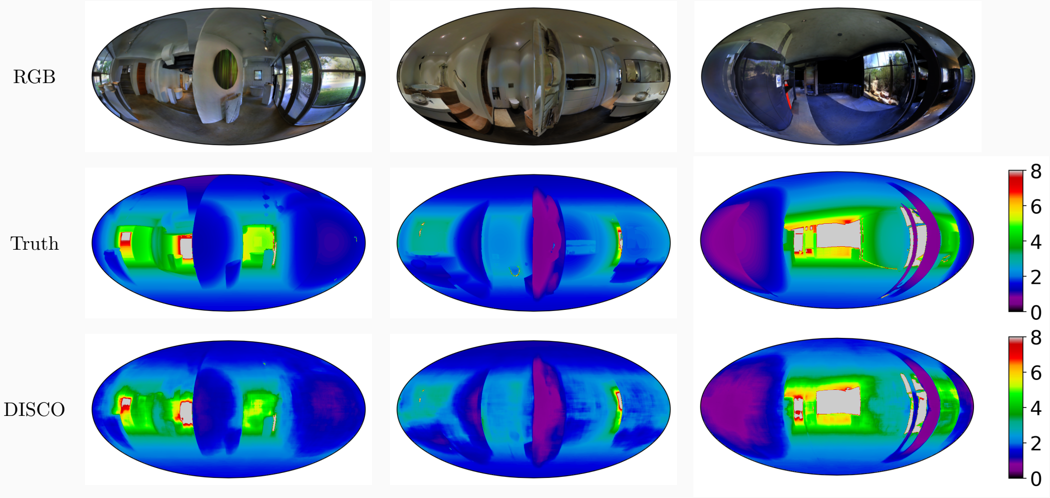
For 4k spherical image, 10^9 saving in computational cost and 10^4 saving in memory usage.

DISCO spherical CNNs exhibit excellent rotational equivariance properties and

are computationally scalable

supporting high-resolution input and output data for dense-prediction tasks.

Depth estimation for Pano3D

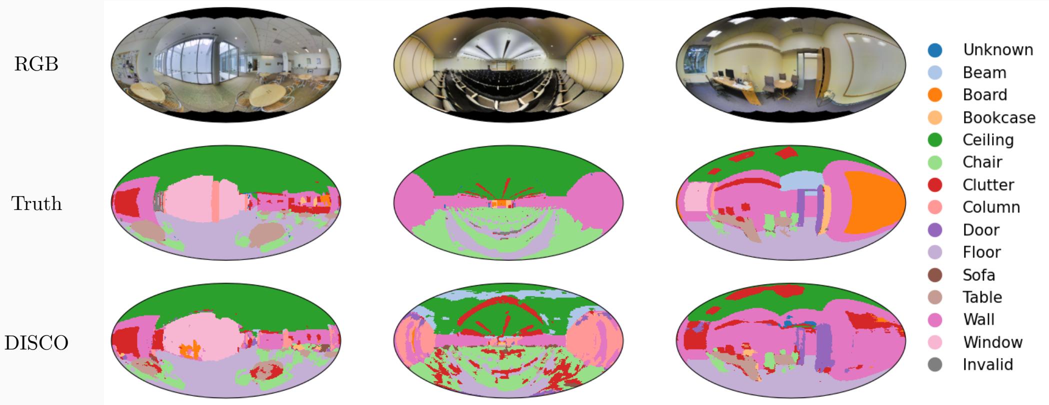


Example predictions for depth estimation of Pano3D data (depth plotted in meters).

Depth estimation for Pano3D

Model	Parameters	Depth Error Metrics				Depth Accuracy Metrics			
		*w*RMSE	*w*RMSLE	*w*AbsRel	*w*SqRel	$\delta^{\mathrm{ico}}_{1.05}$	$\delta^{\mathrm{ico}}_{1.1}$	$\delta^{\mathrm{ico}}_{1.25}$	$\delta^{\mathrm{ico}}_{1.25^2}$
Planar UNet	27M	**0.4520**	**0.1300**	0.1147	**0.0811**	36.68%	60.59%	88.31%	96.96%
DISCO-Directional (Ours)	**658k**	0.5063	0.1695	**0.1109**	0.0852	**38.32%**	**62.12%**	**88.65%**	**97.29%**

Semantic segmentation for 2D3Ds dataset

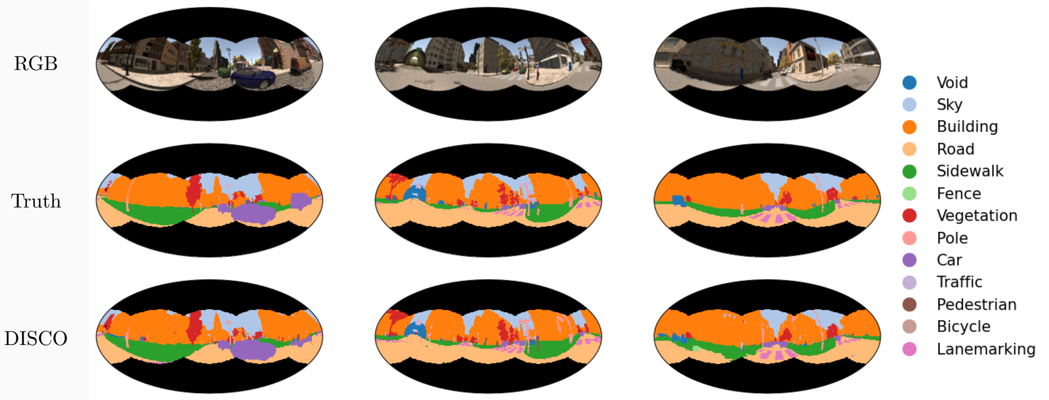


Example predictions for semantic segmentation of 2D3DS data.

Semantic segmentation for 2D3Ds dataset

Model	mIoU	mAcc
Planar UNet	35.9	50.8
UGSCNN	38.3	54.7
GaugeNet	39.4	55.9
HexRUNet	43.3	58.6
SWSCNNs	43.4	58.7
CubeNet	45.0	62.5
MöbiusConv	43.3	60.9
DISCO-Axisymmetric (Ours)	39.7	54.1
DISCO-Directional-Separable (Ours)	43.9	60.9
DISCO-Directional (Ours)	45.2	61.5
DISCO-Directional-Aug (Ours)	**45.7**	**62.7**

Semantic segmentation for Omni-SYNTHIA dataset



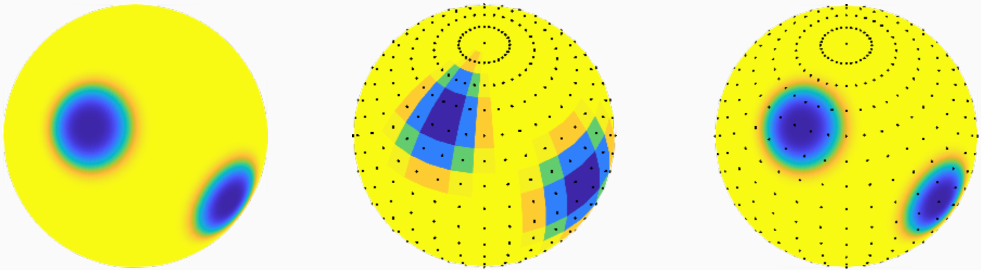
Example predictions for semantic segmentation of Omni-SYNTHIA data.

Semantic segmentation for Omni-SYNTHIA dataset

Model	mIoU	mAcc
Planar UNet	44.6	52.6
UGSCNN	37.6	48.9
HexUNet	48.3	57.1
DISCO-Directional-Separable (Ours)	48.3	59.3
DISCO-Directional-Separable-Aug (Ours)	**49.2**	**63.7**

Summary

Continuous	Discrete	Discrete-Continuous (DISCO)



Continuous	Discrete	Discrete-Continuous (DISCO)
✔ Equivariant	✖ Not Equivariant	✔ Equivariant
✖ Not Scalable	✔ Scalable	✔ Scalable
(Cohen et al. 2018; Esteves et al. 2018; Kondor et al. 2018; Cobb et al. 2021; McEwen et al. 2022)	(Jiang et al. 2019, Zhang et al. 2019, Perraudin et al. 2019, Cohen et al. 2019)	(Ocampo, Price & McEwen, in prep.)