

Geometric deep learning on the sphere

Spherical CNNs and scattering networks

Jason McEwen

www.jasonmcewen.org

Kagenova Limited

Mullard Space Science Laboratory (MSSL), UCL

In collaboration with:

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

March 2022

Outline

1. Symmetry in deep learning
2. Spherical CNNs
3. Efficient generalised spherical CNNs
4. Scattering networks on the sphere

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 \longleftrightarrow Deep Learning

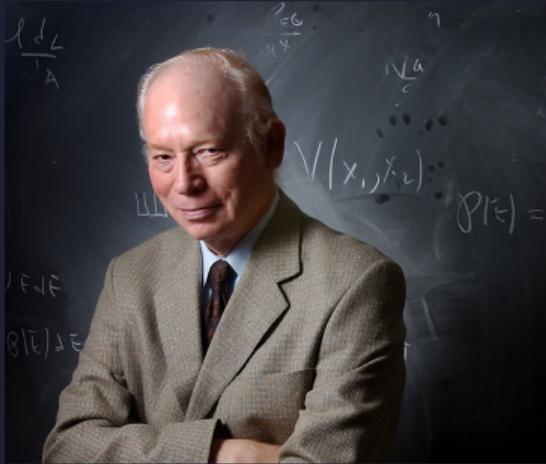
Physics \longleftrightarrow Deep Learning

Here we focus on integrating physics \rightarrow deep learning
(in other works focus on reverse: physics \leftarrow deep learning).

Physics \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

Symmetry

Mirror symmetry



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

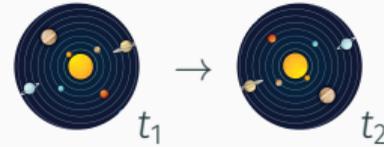
In physics we typically consider **continuous symmetries**, where system is symmetric (invariant) to continuous transformation.



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 \Leftrightarrow conservation of **momentum**
- **Rotational** symmetry \Leftrightarrow conservation of **angular momentum**
- **Time translational** symmetry \Leftrightarrow 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

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

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

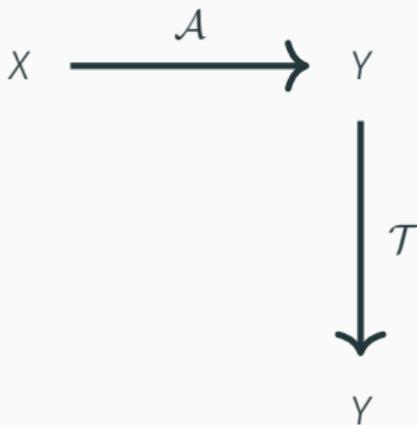
Equivariance

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

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

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

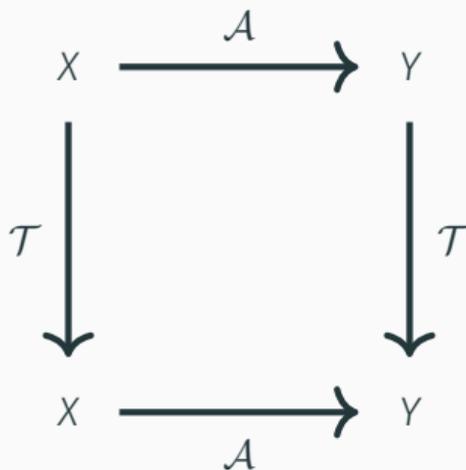
Equivariance

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

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

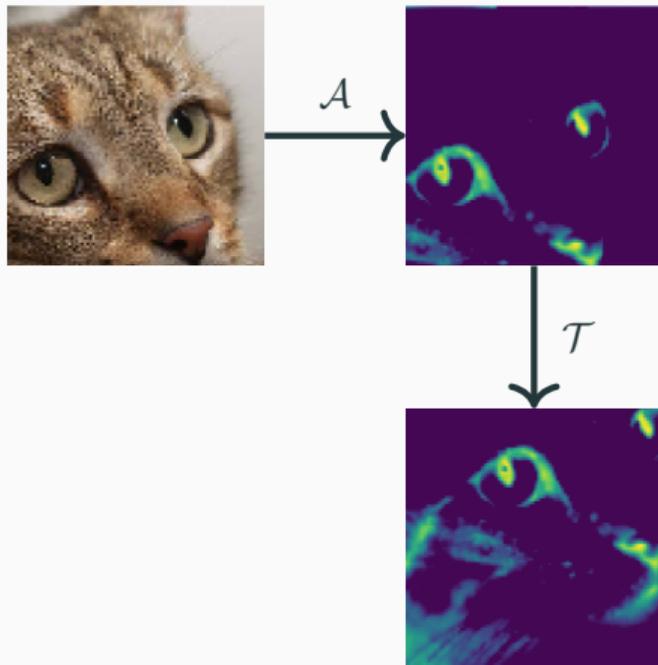
for all possible signals f .

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



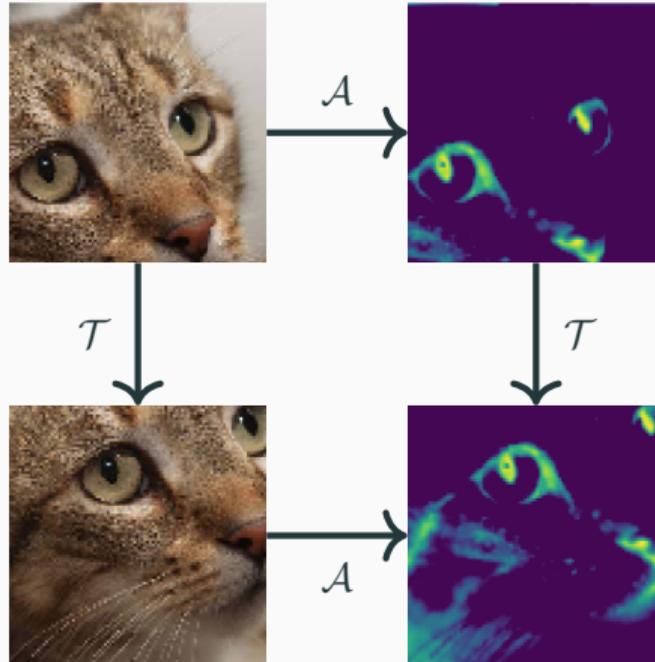
Planar (Euclidean) CNNs exhibit translational equivariance

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



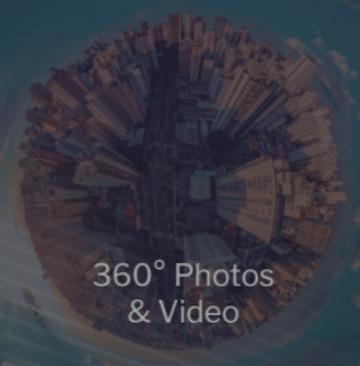
Still a cat

Spherical CNNs

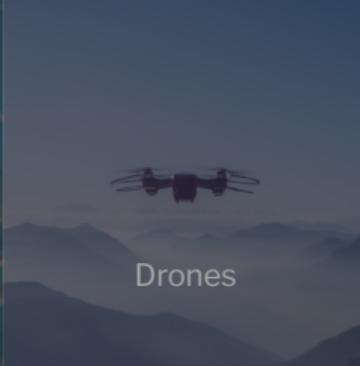
Data on the sphere is prevalent

Data on the sphere is prevalent

Encode symmetries of the sphere and rotations



360° Photos
& Video



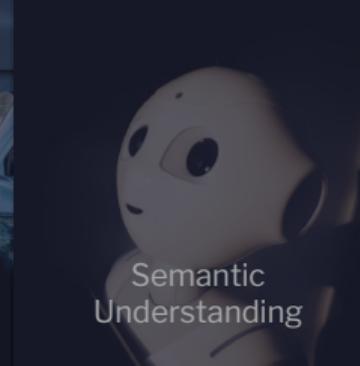
Drones



Extended Reality
(VR / AR / MR)



Autonomous
Vehicles

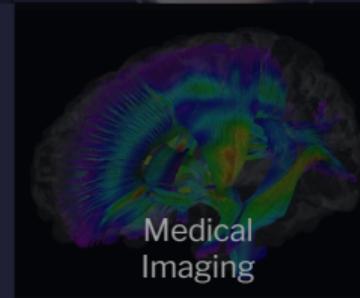


Semantic
Understanding

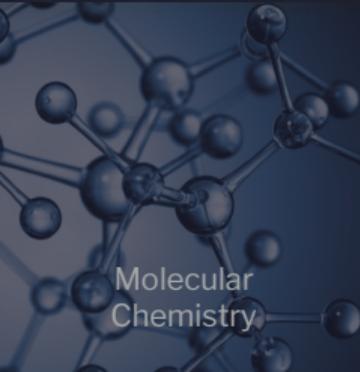


Surveillance &
Monitoring

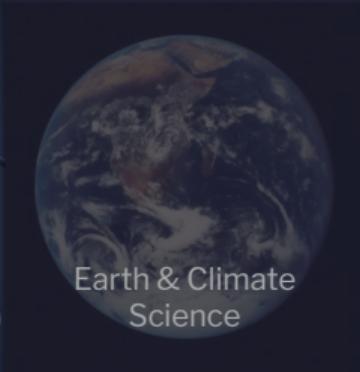
Data on the sphere arises
in many applications



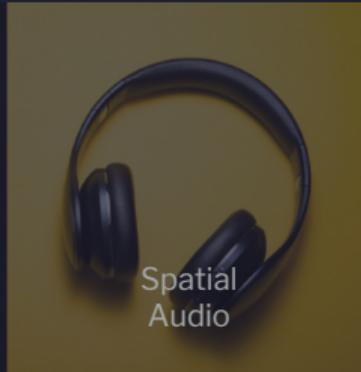
Medical
Imaging



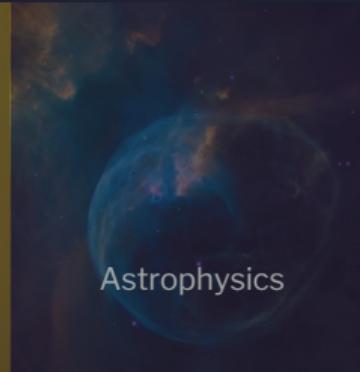
Molecular
Chemistry



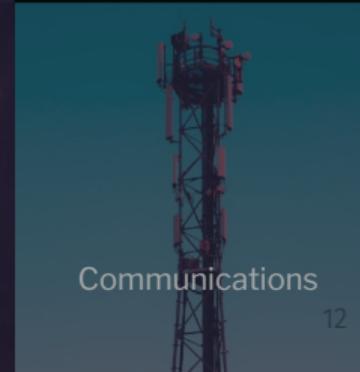
Earth & Climate
Science



Spatial
Audio



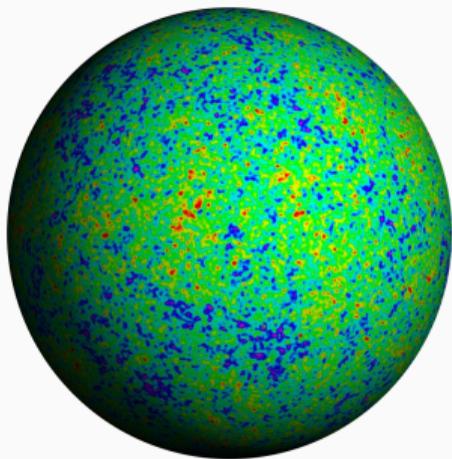
Astrophysics



Communications

Cosmology and virtual reality

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



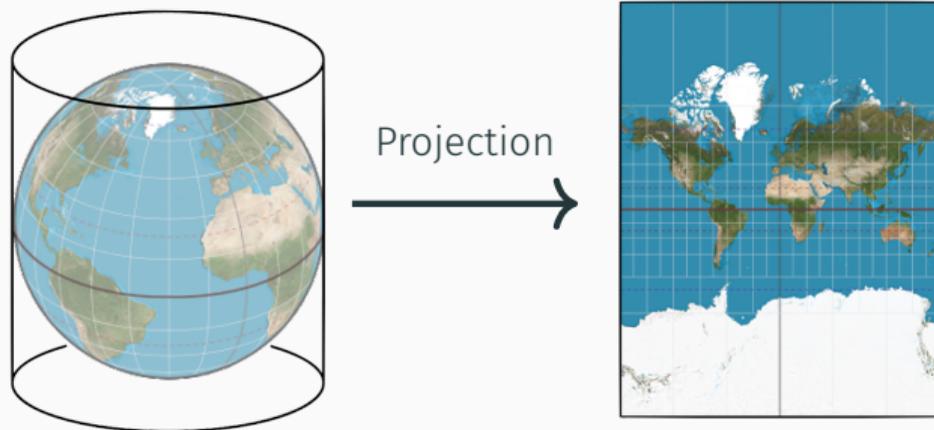
Cosmic microwave background



360° virtual reality

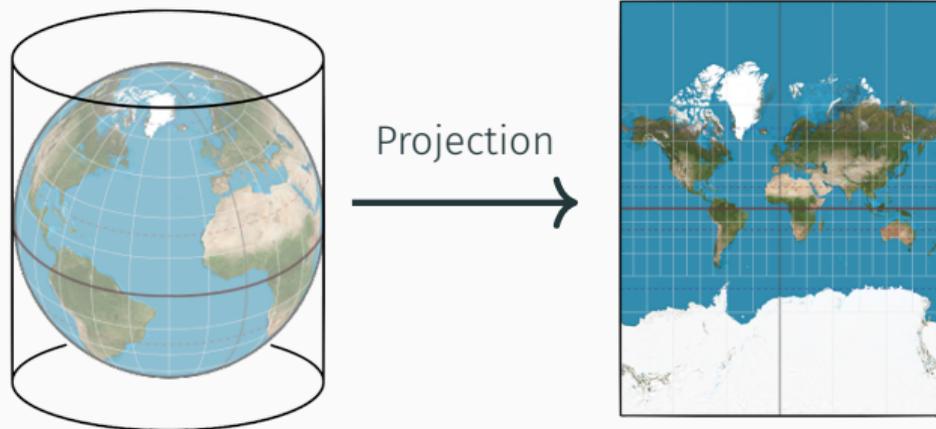
Why not standard (Euclidean) deep learning approaches?

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.



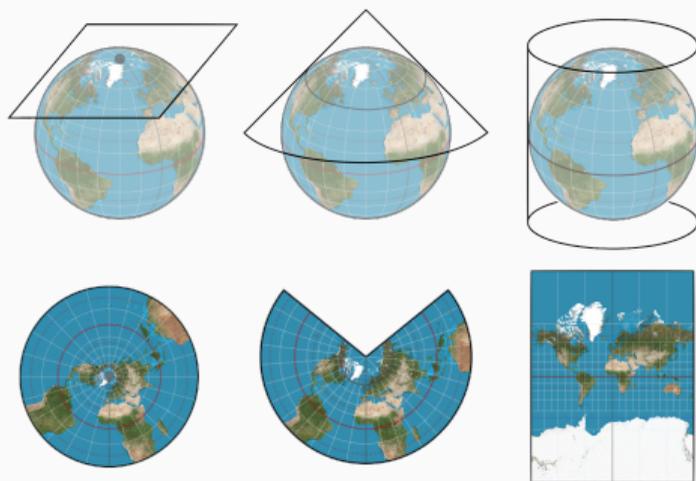
Greenland appears to be a similar size to Africa in the projected planar map, whereas it is over 10 times smaller.

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.

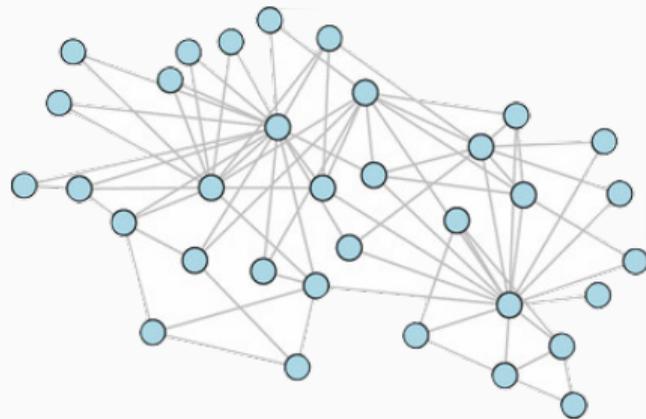


No projection of the sphere to the plane can preserve both shapes and areas
⇒ distortions are unavoidable.

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

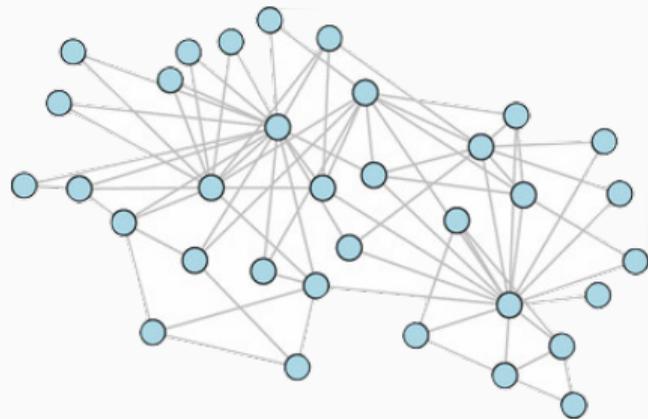
Why not graph-based geometric deep learning?

Could construct graph representation of sphere and apply graph CNNs.



Why not graph-based geometric deep learning?

Could construct graph representation of sphere and apply graph CNNs.

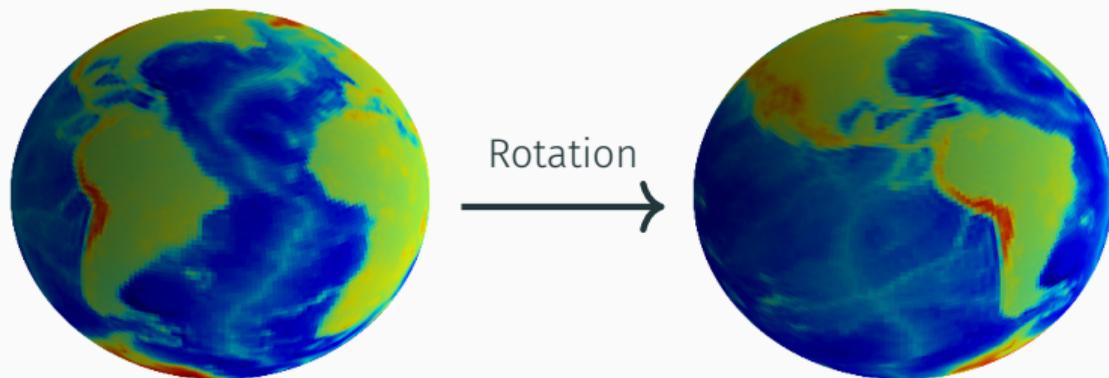


Again, **breaks symmetries and geometric properties** of sphere.

Cannot capture rotational equivariance.

Rotational equivariance

On the sphere, the analog of translations are rotations.



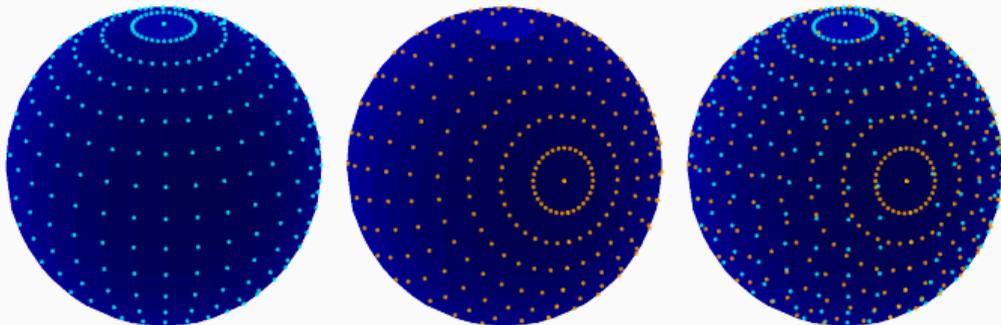
Would like spherical CNNs to exhibit rotational equivariance.

(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).

⇒ Not possible to discretise sphere in a manner that is invariant to rotations.

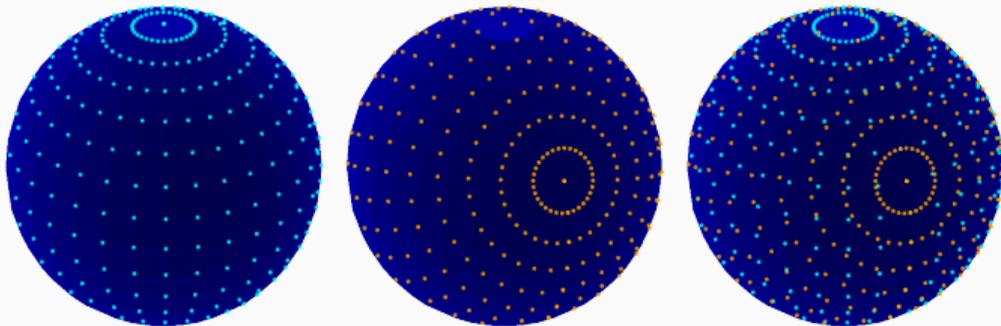


Capturing strict equivariance with operations defined directly in discretised (pixel) space **not possible** due to structure of sphere.

Capturing rotational equivariance in spherical CNNs

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

⇒ Not possible to discretise sphere in a manner that is invariant to rotations.

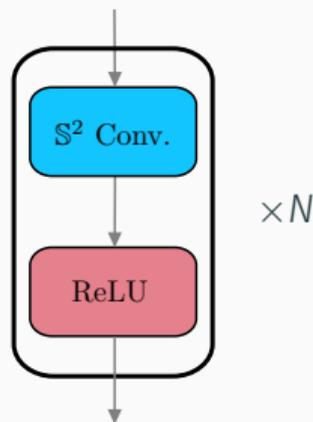


Capturing strict equivariance with operations defined directly in discretised (pixel) space **not possible** due to structure of sphere.

Instead, consider **Fourier approach** → access to underlying continuous representations.

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).



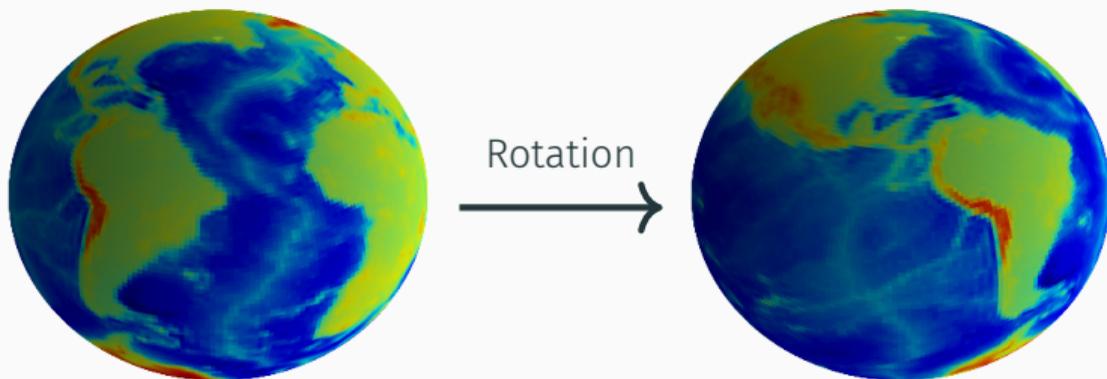
(Alternative, real space constructions have also been developed but do not exhibit rotational equivariance so not considered further; e.g. Boomsma & Frelsen 2017, Jiang et al. 2019, Perraudin et al. 2019.)

Rotation of signals

Rotation of signals in spatial domain

A signal $f \in L^2(\Omega)$ on the sphere ($\Omega = \mathbb{S}^2$) or rotation group ($\Omega = SO(3)$) can be rotated by $\rho \in SO(3)$ by

$$Rf(\omega) = f(\rho^{-1}\omega), \quad \text{for } \omega \in \Omega.$$



Convolution of signals

Convolution of signals in spatial domain

Convolution of two signals $f, \psi \in L^2(\Omega)$ is given by

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

where $d\mu(\omega)$ denotes the Haar measure on Ω and \cdot^* complex conjugation.

Convolution of signals

Convolution of signals in spatial domain

Convolution of two signals $f, \psi \in L^2(\Omega)$ is given by

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

where $d\mu(\omega)$ denotes the Haar measure on Ω and \cdot^* complex conjugation.

Convolution of signals

Convolution of signals in spatial domain

Convolution of two signals $f, \psi \in L^2(\Omega)$ is given by

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

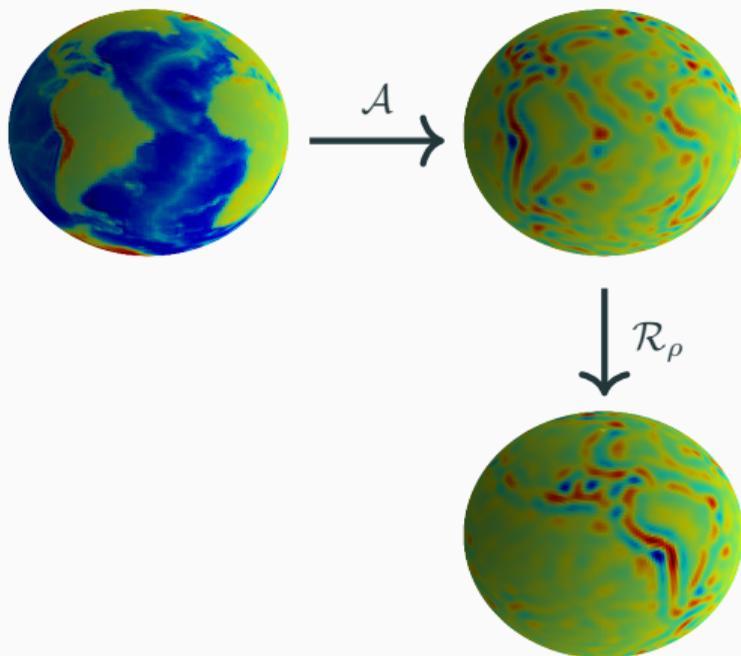
where $d\mu(\omega)$ denotes the Haar measure on Ω and \cdot^* complex conjugation.

Since no regular discretization of the sphere, compute in Fourier space to ensure equivariant.

Convolution is rotationally equivariant

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

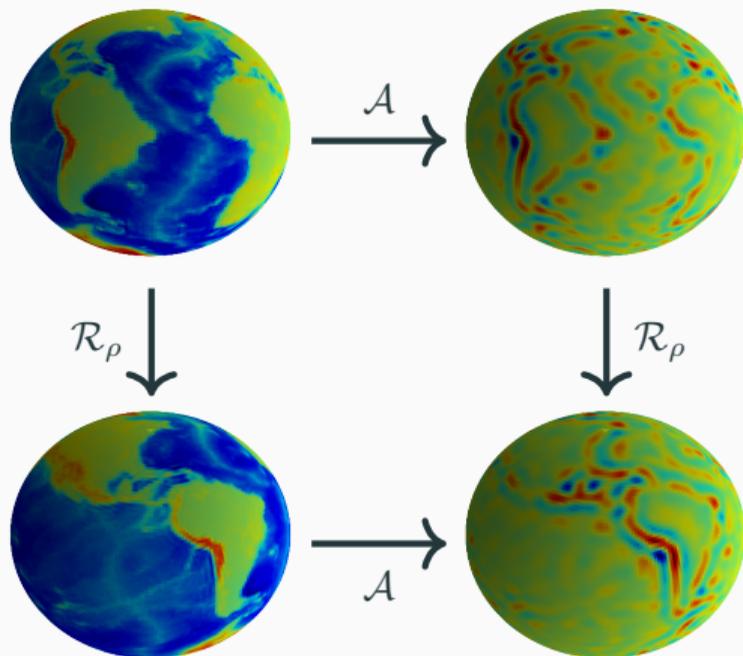
$$((Rf) \star \psi)(\rho') = (\mathcal{R}_\rho(f \star \psi))(\rho').$$



Convolution is rotationally equivariant

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

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



Pointwise activation

While **pointwise activations** are rotationally equivariant in the continuous limit, they are **not equivariant in practice** when applied to discretised signals (since regular discretisation of sphere does not exist).

Pointwise activation

While **pointwise activations** are rotationally equivariant in the continuous limit, they are **not equivariant in practice** when applied to discretised signals (since regular discretisation of sphere does not exist).

Equivariance errors

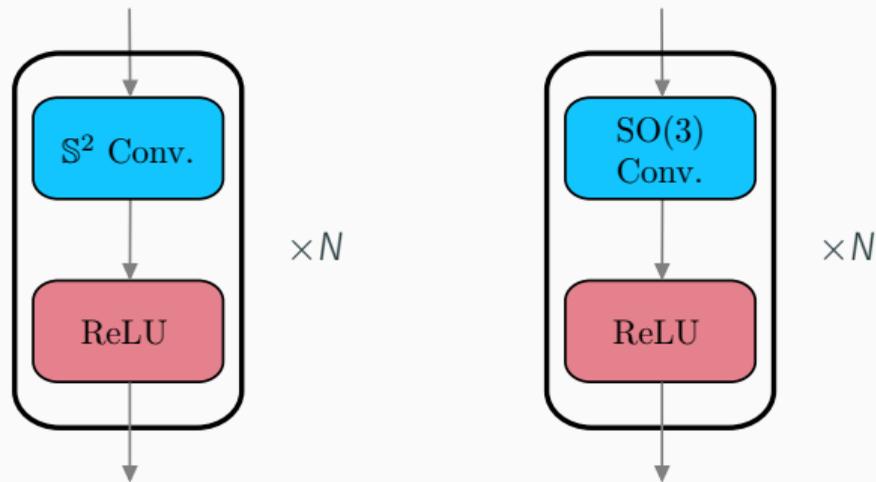
Layer	Mean Relative Error*
\mathbb{S}^2 to \mathbb{S}^2 conv.	4.4×10^{-7}
\mathbb{S}^2 to $SO(3)$ conv.	5.3×10^{-7}
$SO(3)$ to $SO(3)$ conv.	9.3×10^{-7}
\mathbb{S}^2 ReLU	3.4×10^{-1}
$SO(3)$ ReLU	3.7×10^{-1}

* Floating point precision.

Spherical CNNs

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

Despite imperfect equivariance, find empirically that such models maintain a reasonable degree of equivariance and generally perform well.



Efficient generalised spherical CNNs



Group theory is the mathematical study of symmetry.



Since we're concerned with rotational symmetry, leverage the machinery from the study of angular momentum in quantum mechanics.

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)

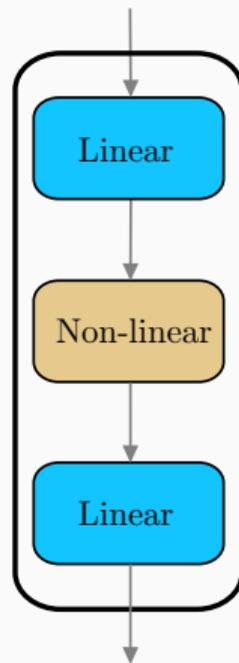
$$\mathcal{A}^{(s)} = (\mathcal{L}_1, \mathcal{N}, \mathcal{L}_2),$$

such that

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

where

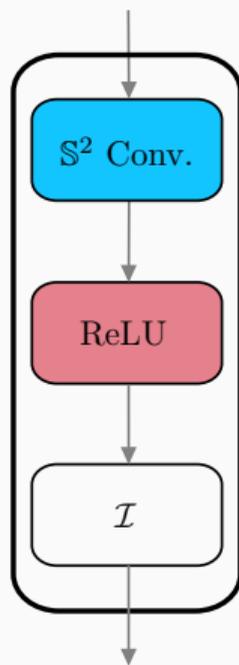
- $\mathcal{L}_1, \mathcal{L}_2 : \mathcal{F}_L \rightarrow \mathcal{F}_L$ are **spherical convolution** operators,
- $\mathcal{N} : \mathcal{F}_L \rightarrow \mathcal{F}_L$ is a **non-linear, spherical activation** operator.



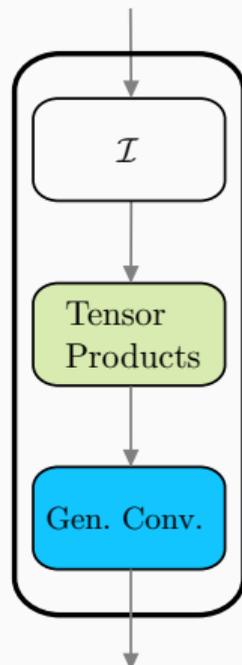
Generalised spherical CNNs

- 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.

Existing spherical CNN layers are **highly computationally costly**, particularly those non-linear layers that satisfy strict rotational equivariance.



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

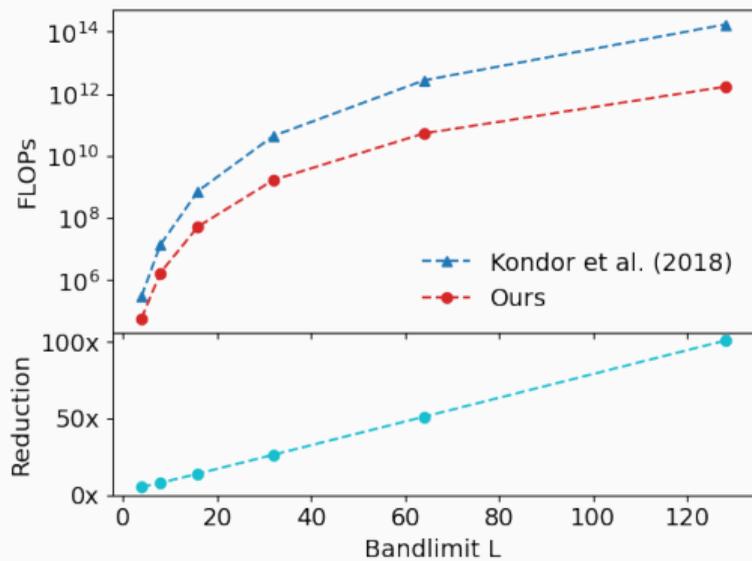


Kondor et al. (2018)

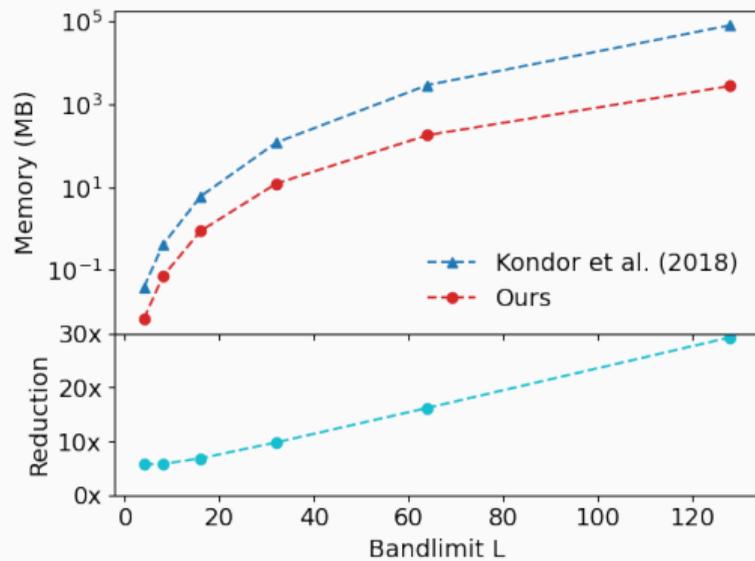
Contributions to improve efficiency

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

Rotational equivariance

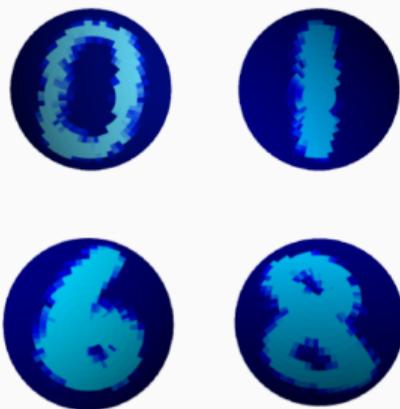
Equivariance errors

Layer	Mean Relative Error*
Tensor-product activation → Generalized convolution	5.0×10^{-7}
\mathbb{S}^2 ReLU	3.4×10^{-1}
SO(3) ReLU	3.7×10^{-1}

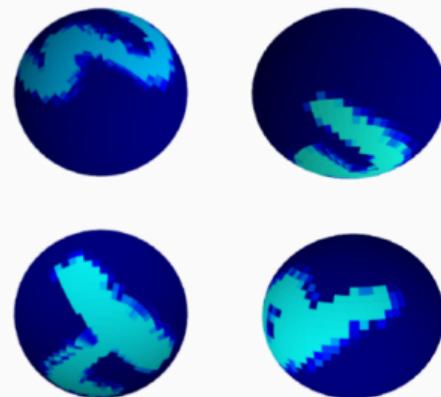
* Floating point precision.

Spherical MNIST: problem

Canonical benchmark problem of classifying MNIST digits projects onto the sphere.

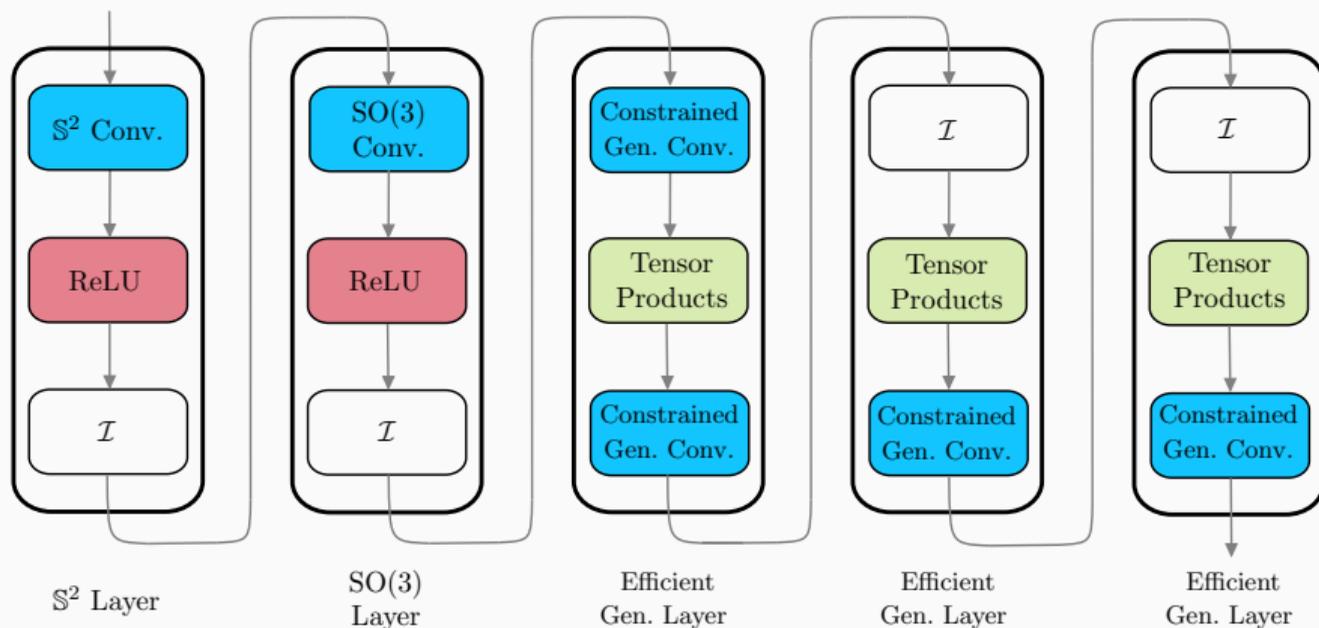


Non-rotated (NR)



Rotated (R)

Spherical MNIST: architecture



Spherical MNIST: results

Test accuracy for spherical MNIST digits classification problem

	NR/NR	R/R	NR/R	Params
Planar CNN	99.32			58k
Cohen et al. 2018	95.59			58k
Kondor et al. 2018	96.40			286k
Esteves et al. 2018	99.37			58k
Ours (MST)	99.35			58k
Ours (RMST)	99.29			57k

Spherical MNIST: results

Test accuracy for spherical MNIST digits classification problem

	NR/NR	R/R	NR/R	Params
Planar CNN	99.32	90.74		58k
Cohen et al. 2018	95.59	94.62		58k
Kondor et al. 2018	96.40	96.60		286k
Esteves et al. 2018	99.37	99.37		58k
Ours (MST)	99.35	99.38		58k
Ours (RMST)	99.29	99.17		57k

Spherical MNIST: results

Test accuracy for spherical MNIST digits classification problem

	NR/NR	R/R	NR/R	Params
Planar CNN	99.32	90.74	11.36	58k
Cohen et al. 2018	95.59	94.62		58k
Kondor et al. 2018	96.40	96.60		286k
Esteves et al. 2018	99.37	99.37		58k
Ours (MST)	99.35	99.38		58k
Ours (RMST)	99.29	99.17		57k

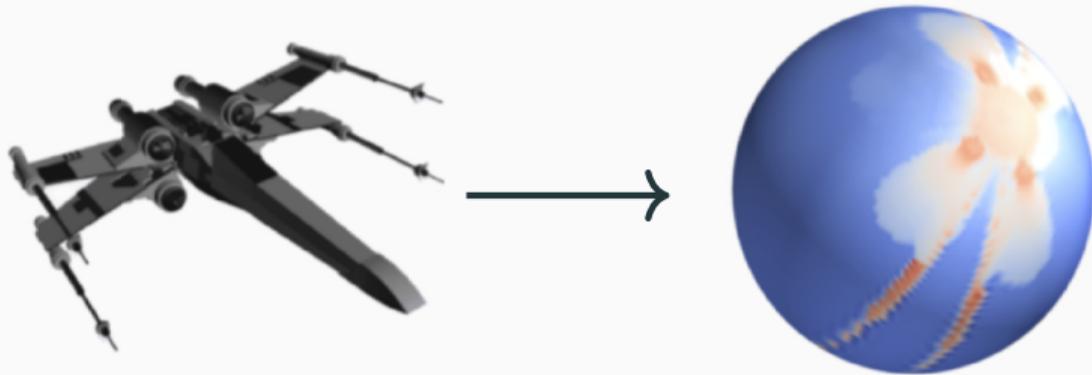
Spherical MNIST: results

Test accuracy for spherical MNIST digits classification problem

	NR/NR	R/R	NR/R	Params
Planar CNN	99.32	90.74	11.36	58k
Cohen et al. 2018	95.59	94.62	93.40	58k
Kondor et al. 2018	96.40	96.60	96.00	286k
Esteves et al. 2018	99.37	99.37	99.08	58k
Ours (MST)	99.35	99.38	99.34	58k
Ours (RMST)	99.29	99.17	99.18	57k

3D shape classification: problem

Classify 3D meshes and perform shape retrieval.



[Image credit: Esteves et al. 2018]

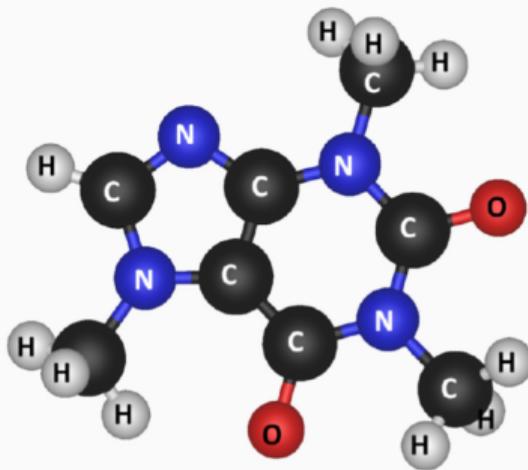
3D shape classification: results

SHREC'17 object retrieval competition metrics (perturbed micro-all)

	P@N	R@N	F1@N	mAP	NDCG	Params
Kondor et al. 2018	0.707	0.722	0.701	0.683	0.756	>1M
Cohen et al. 2018	0.701	0.711	0.699	0.676	0.756	1.4M
Esteves et al. 2018	0.717	0.737	-	0.685	-	500k
Ours	0.719	0.710	0.708	0.679	0.758	250k

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

Scattering networks on the sphere

Despite the efficient generalized approach discussed
rotationally equivariant spherical CNNs are not scalable to high-resolution data

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)

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)

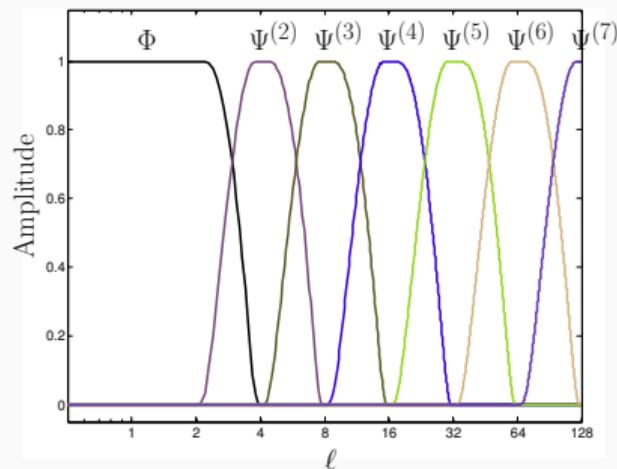
⇒ **Scattering networks on the sphere** (McEwen et al. 2022; arXiv:2102.02828)

Wavelets on the sphere

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

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

Scaling function $\phi \in L^2(\mathbb{S}^2)$ 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 $\psi_j \in L^2(\mathbb{S}^2)$ capture high-frequency signal content at scale j .

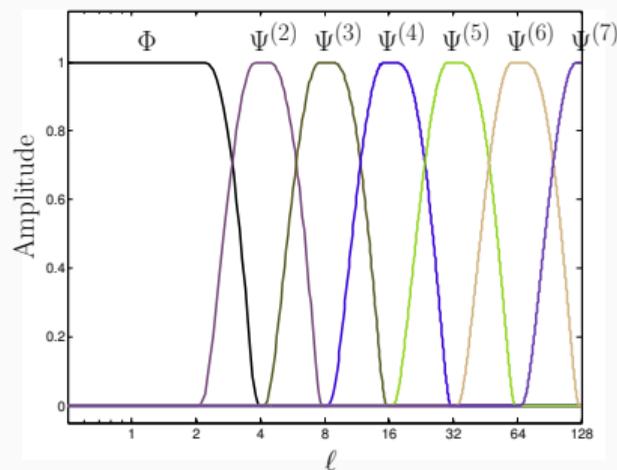
Scaling function $\phi \in L^2(\mathbb{S}^2)$ captures low-frequency content.

Spherical wavelet transform given by

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

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

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

Spherical wavelet transform given by

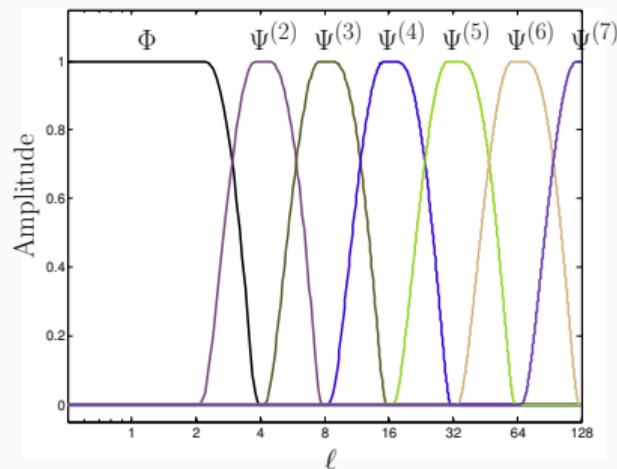
$$W_j(\omega) = (f \star \psi_j)(\omega) = \int_{\mathbb{S}^2} d\mu(\omega') f(\omega') (R_\omega \psi_j)^*(\omega').$$

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:

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

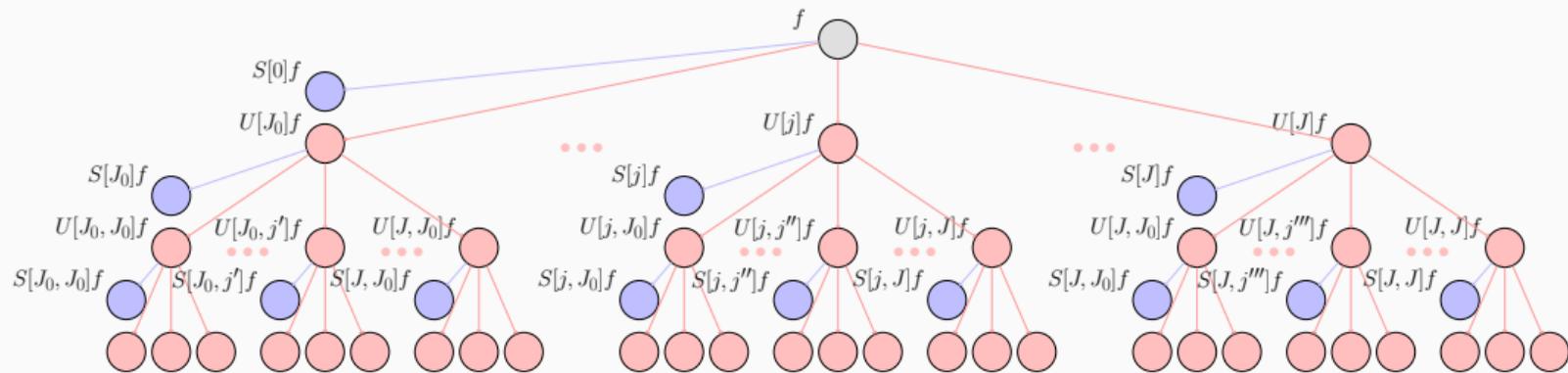
for the path $p = (j_1, j_2, \dots, j_d)$ with depth d .

Scattering coefficients:

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

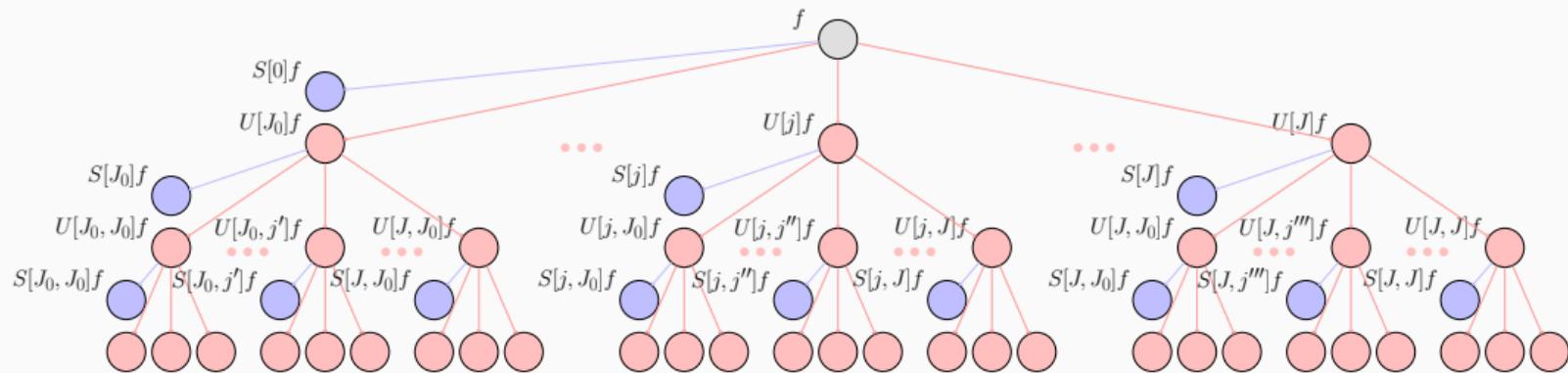
Scattering networks on the sphere

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



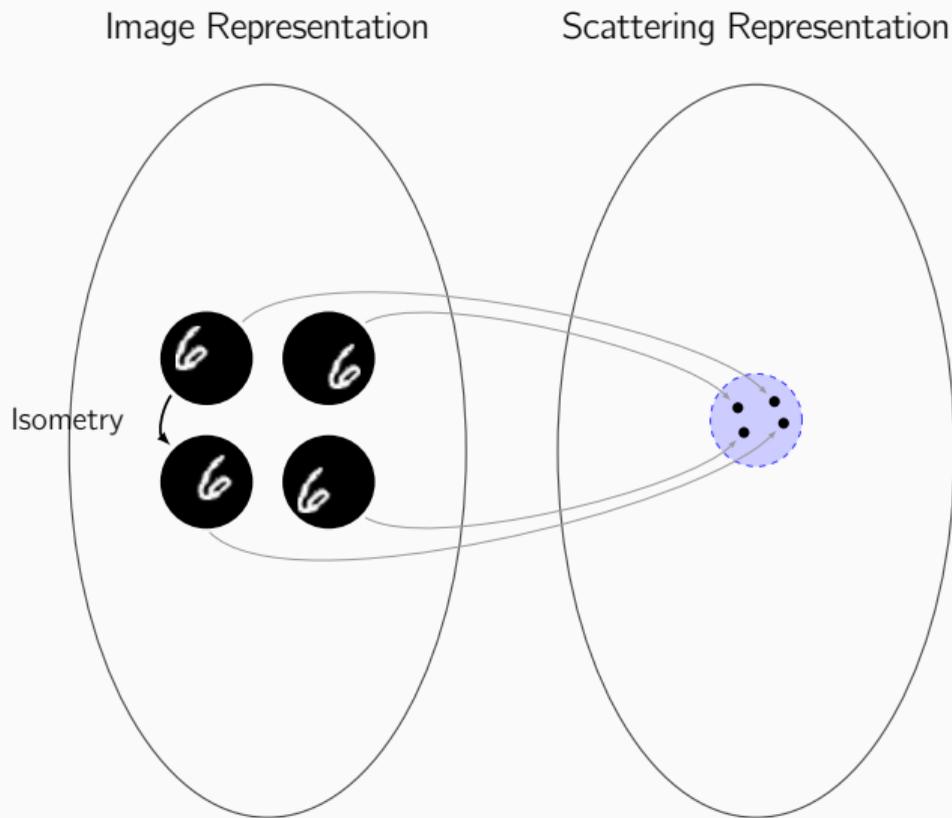
Scattering networks on the sphere

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



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

Isometric invariance



Theorem (Isometric Invariance)

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

$$\|\mathcal{S}_{\mathbb{P}_D} f - \mathcal{S}_{\mathbb{P}_D} V_{\zeta} f\|_2 \leq CL^{5/2}(D+1)^{1/2} \lambda^0 \|\zeta\|_{\infty} \|f\|_2.$$

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

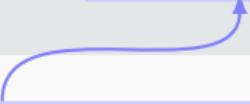
Isometric invariance

Theorem (Isometric Invariance)

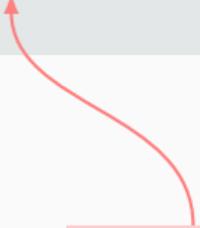
Let $\zeta \in \text{Isom}(\mathbb{S}^2)$, then there exists a constant C such that for all $f \in L^2(\mathbb{S}^2)$,

$$\|\mathcal{S}_{\mathbb{P}_D} f - \mathcal{S}_{\mathbb{P}_D} V_{\zeta} f\|_2 \leq CL^{5/2}(D+1)^{1/2} \lambda^0 \|\zeta\|_{\infty} \|f\|_2.$$

Difference in representation

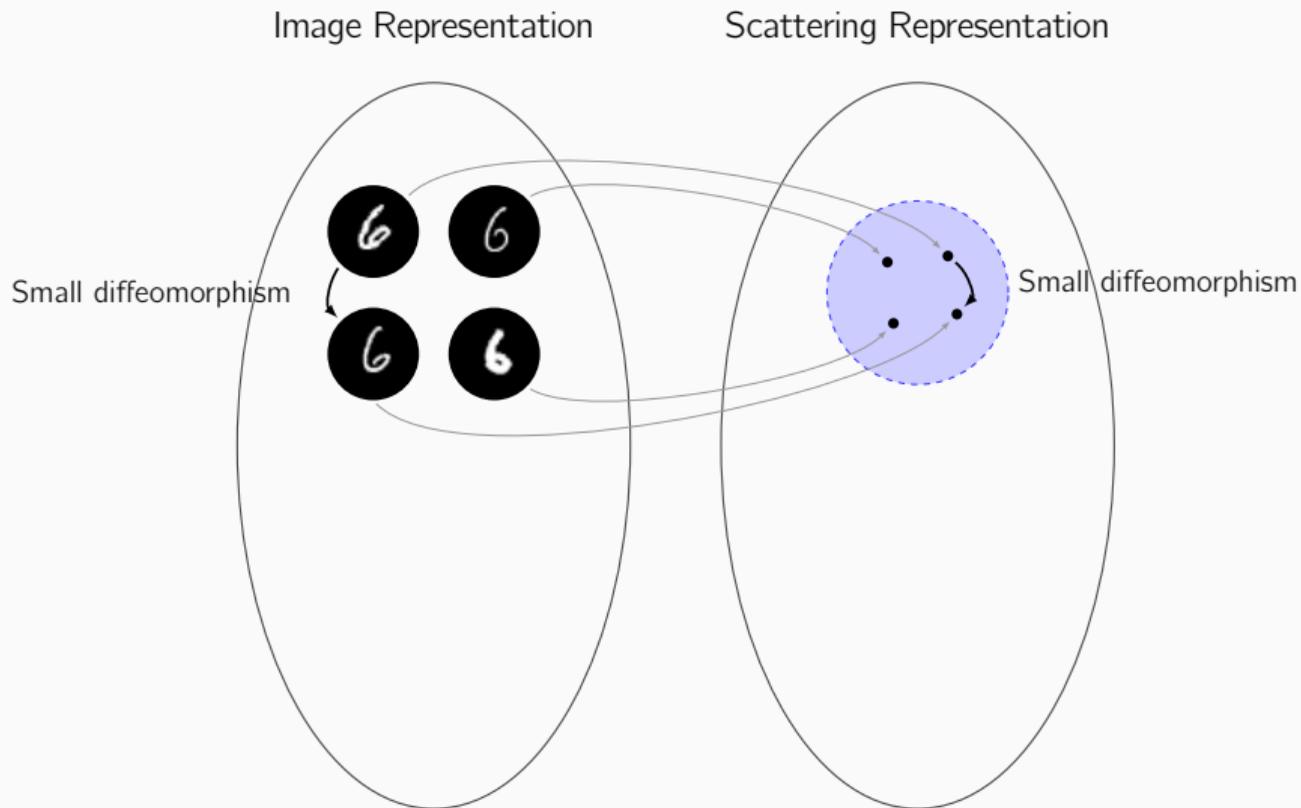


Scattering network representation is invariant to isometries up to a scale.

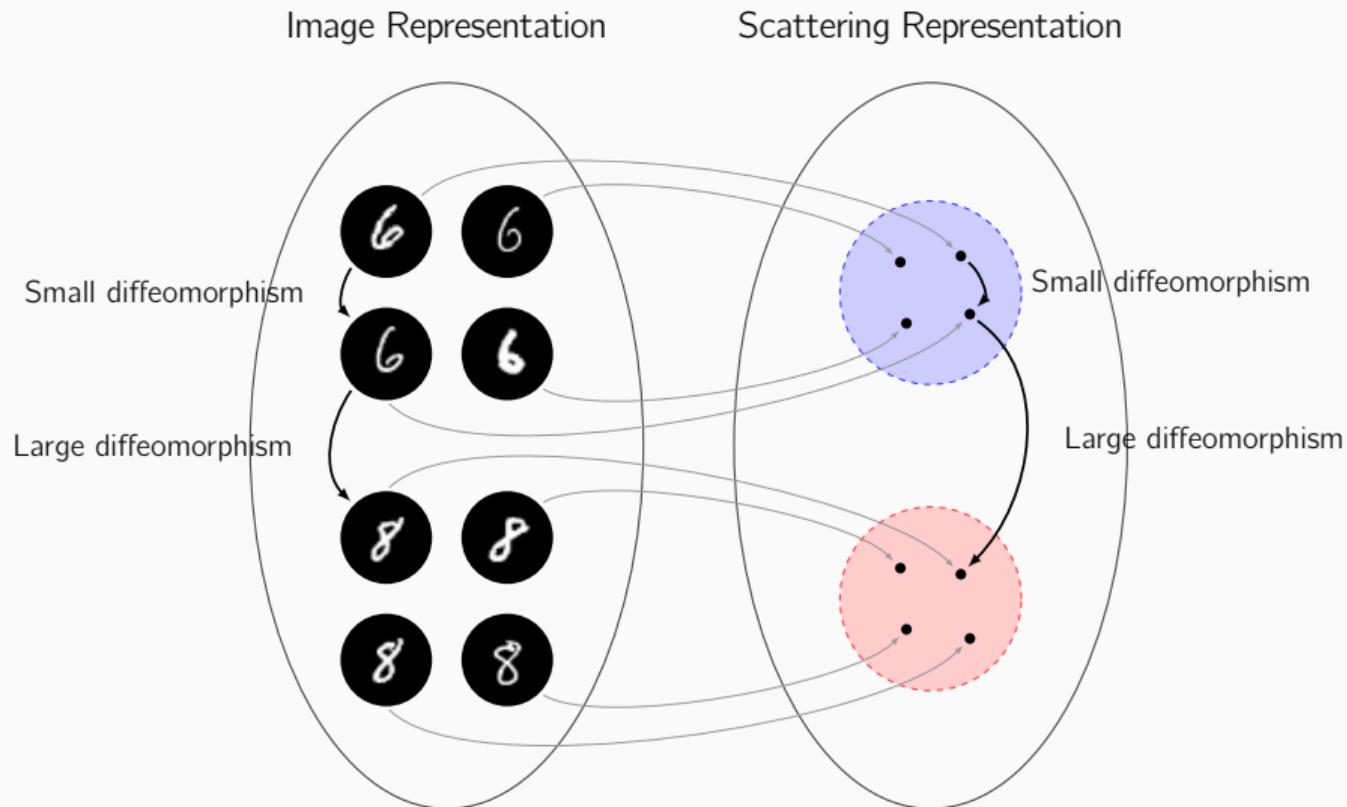


(**Proof:** Follows by straightforward extension of proof of Perlmutter et al. 2020.)

Stability to diffeomorphisms



Stability to diffeomorphisms



Stability to diffeomorphisms

Theorem (Stability to Diffeomorphisms)

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

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

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

Stability to diffeomorphisms

Theorem (Stability to Diffeomorphisms)

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

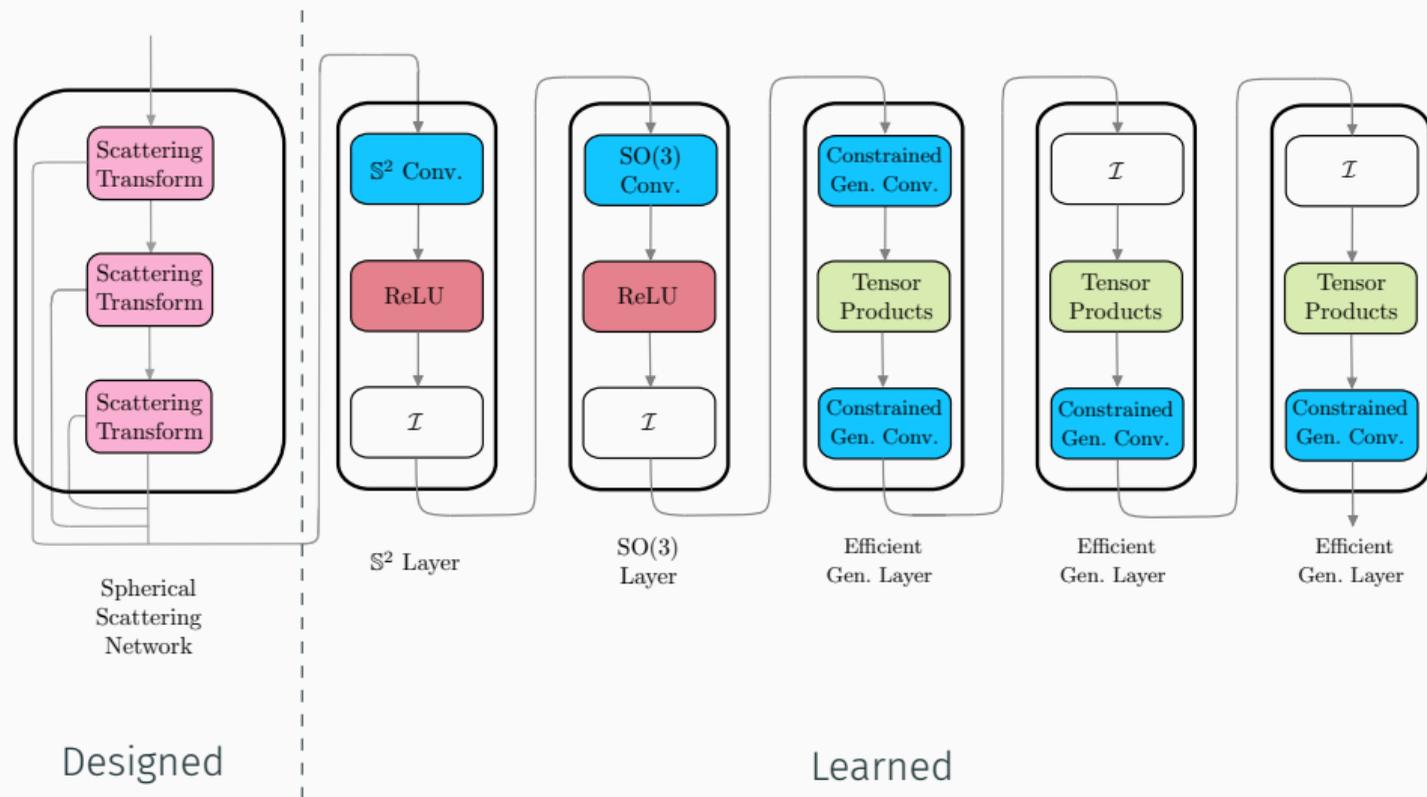
$$\|\mathcal{S}_{\mathbb{P}_D} f - \mathcal{S}_{\mathbb{P}_D} V_{\zeta} f\|_2 \leq C L^2 [L^2 \|\zeta_2\|_{\infty} + L^{1/2} (D+1)^{1/2} \lambda^{j_0} \|\zeta_1\|_{\infty}] \|f\|_2.$$

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

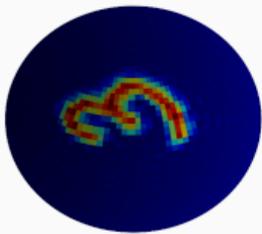


Rotational equivariance

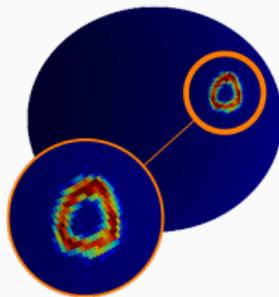
Path depth d	Equivariance Error		
	Minimum	Median	Maximum
0	0.00%	0.00%	0.00%
1	0.01%	0.05%	0.24%
2	0.18%	1.01%	5.36%
3	0.56%	3.47%	10.68%

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

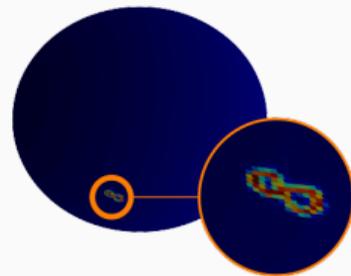
Spherical MNIST at varying resolution



$L = 64$



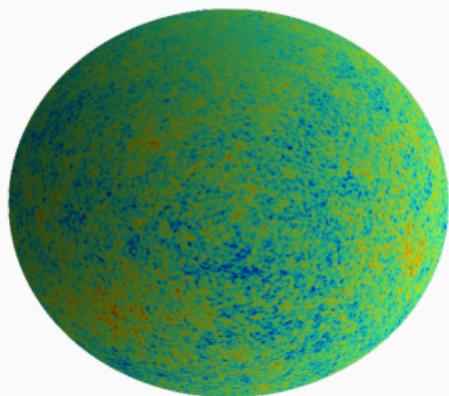
$L = 128$



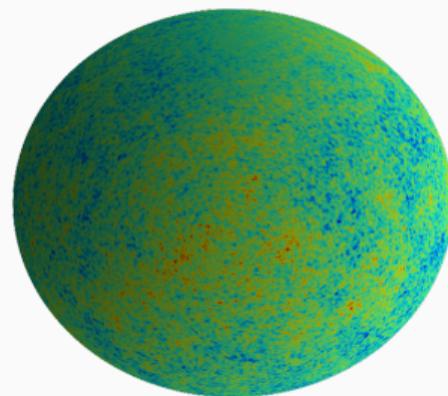
$L = 256$

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

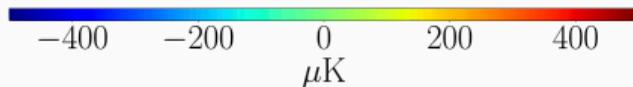
Gaussianity of the cosmic microwave background



Gaussian



Non-Gaussian



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

Summary

- Importance of encoding **equivariance to symmetry transforms** in order to capture fundamental physical understanding of generative process.

Summary

- Importance of encoding **equivariance to symmetry transforms** in order to capture fundamental physical understanding of generative process.
- Need for geometric deep learning techniques **constructed natively on manifolds**, such as the sphere.

Summary

- Importance of encoding **equivariance to symmetry transforms** in order to capture fundamental physical understanding of generative process.
- Need for geometric deep learning techniques **constructed natively on manifolds**, such as the sphere.
- Reviewed **spherical CNNs constructions**, with a focus on rotational equivariance (Cohen et al. 2018, Esteves et al. 2018, Kondor et al. 2018).

Summary

- Importance of encoding **equivariance to symmetry transforms** in order to capture fundamental physical understanding of generative process.
- Need for geometric deep learning techniques **constructed natively on manifolds**, such as the sphere.
- Reviewed **spherical CNNs constructions**, with a focus on rotational equivariance (Cohen et al. 2018, Esteves et al. 2018, Kondor et al. 2018).
- **Efficient generalised spherical CNNs** (Cobb et al. 2021; arXiv:2010.11661)
- **Scattering networks on the sphere for scalable and rotationally equivariant spherical CNNs** (McEwen et al. 2022; arXiv:2102.02828)

Questions?

Extra slides

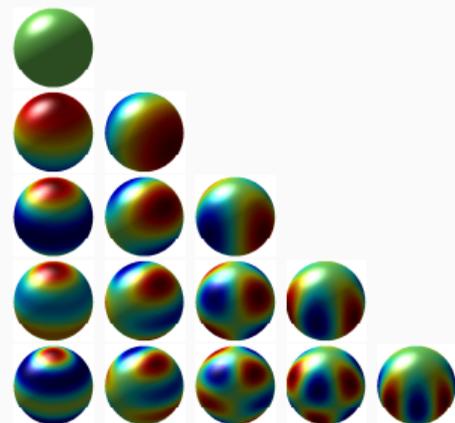
Fourier transforms on the sphere and rotation group

Fourier transform on sphere: spherical harmonic transform

A signal $f \in L^2(\mathbb{S}^2)$ can be decomposed into its harmonic representation by

$$f(\omega) = \sum_{\ell m} \hat{f}_m^\ell Y_m^\ell(\omega), \quad \text{for } \omega \in \mathbb{S}^2.$$

where $\hat{f}_m^\ell = \langle f, Y_m^\ell \rangle$ (also denoted \hat{f}^ℓ).



Spherical harmonics

Fourier transform on rotation group: Wigner transform

A signal $f \in L^2(\text{SO}(3))$ can be decomposed into its harmonic representation by

$$f(\rho) = \sum_{\ell m} \frac{2\ell + 1}{8\pi} \hat{f}_{mn}^\ell D_{mn}^{\ell*}(\rho), \quad \text{for } \rho \in \text{SO}(3).$$

where $\hat{f}_{mn}^\ell = \langle f, D_{mn}^{\ell*} \rangle$ (also denoted \hat{f}^ℓ).

Harmonic computations

Rotation of signals in harmonic domain

The rotation $f \mapsto Rf$ of a signal $f \in L^2(\Omega)$ can be described in harmonic space by

$$\hat{f}^\ell \mapsto D^\ell(\rho) \hat{f}^\ell.$$

Convolution of signals in harmonic domain

Convolution of two signals $f, \psi \in L^2(\Omega)$ can be described in harmonic space by

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

By computing convolutions in harmonic space, discretisation effects are eliminated.

Furthermore, fast harmonic transform algorithms can be leveraged.

Generalised signals

Consider generalised signal representations and convolutions of Kondor et al. (2018).

Generalised signals

Consider set of variable length vectors of the form

$$f = \{\hat{f}_t^\ell \in \mathbb{C}^{2\ell+1} : \ell = 0, \dots, L-1; t = 1, \dots, \tau_f^\ell\},$$

for t -th fragment of degree ℓ . Let \mathcal{F}_L be the space of all such sets of variable length vectors, with type $\tau_f = (\tau_f^0, \dots, \tau_f^{L-1})$ unconstrained.

Includes signals on the sphere and rotation group as special cases:

- $\tau_f^\ell = 1$ for signals on the sphere
- $\tau_f^\ell = 2\ell + 1$ for signals on the rotation group

Rotation of generalised signals

Rotation of generalised signals

The rotation $f \mapsto Rf$ of a signal $f \in \mathcal{F}_L$ can be described by

$$\hat{f}_t^\ell \mapsto D^\ell(\rho) \hat{f}_t^\ell.$$

We may therefore extend the usual notion of rotational equivariance to \mathcal{F}_L .

Convolution of generalised signals

Convolution of generalised signals

Generalised convolution of a signal $f \in \mathcal{F}_L$ with a filter ψ is given by

$$(f * \psi)_t^\ell = \sum_{t'=1}^{\tau_f^\ell} \hat{f}_{t'}^\ell \hat{\psi}_{t,t'}^\ell,$$

for a filter $\psi = \{\hat{\psi}^\ell \in \mathbb{C}^{\tau_f^\ell \times \tau_{(f*\psi)}^\ell} : \ell = 0, \dots, L-1\}$.

Do not force the filter ψ to occupy the same domain as the signal f , allowing control over the type $\tau_{(f*\psi)}$ of transformed signal.

Provides **generalised rotationally equivariant linear operator**.

Non-linear transforms of generalised signals

How introduce non-linearity in an equivariant manner?

Non-linear transforms of generalised signals

How introduce non-linearity in an equivariant manner?

Consider irreducible representations of the rotation group $SO(3)$ and leverage the decomposability of the **tensor product** between these representations (Thomas et al. 2018, Kondor et al. 2018).

Decomposition of tensor product representations

Representation theory is concerned with the representation of abstract algebraic structures, e.g. groups, by linear transformations.

Consider tensor product of representation spaces (generalisation of outer product).

$D^\ell : \text{SO}(3) \rightarrow \text{GL}(\mathbb{C}^{2\ell+1})$ is an irreducible group representation of $\text{SO}(3)$ on $\mathbb{C}^{2\ell+1}$ (since it is a group homomorphism from $\text{SO}(3)$ to the general linear group $\text{GL}(\mathbb{C}^{2\ell+1})$).

Decomposition of tensor product representations

Tensor-product group representation $D^{\ell_1} \otimes D^{\ell_2}$ is defined such that

$$(D^{\ell_1} \otimes D^{\ell_2})(\rho) = D^{\ell_1}(\rho) \otimes D^{\ell_2}(\rho),$$

which is **not irreducible**.

Recover irreducible representation through change of basis such that $(D^{\ell_1} \otimes D^{\ell_2})(\rho)$ is block diagonal, where for each ℓ there is a block equal to $D^\ell(\rho)$.

Decomposition of tensor product representations

Change of basis for $\hat{u}^{\ell_1} \otimes \hat{v}^{\ell_2} \in \mathbb{C}^{2\ell_1+1} \otimes \mathbb{C}^{2\ell_2+1}$ to recover an irreducible representation is

$$(\hat{u}^{\ell_1} \otimes \hat{v}^{\ell_2})_m^\ell = \sum_{m_1=-\ell_1}^{\ell_1} \sum_{m_2=-\ell_2}^{\ell_2} C_{m_1, m_2, m}^{\ell_1, \ell_2, \ell} \hat{u}_{m_1}^{\ell_1} \hat{v}_{m_2}^{\ell_2},$$

where $C_{m_1, m_2, m}^{\ell_1, \ell_2, \ell} \in \mathbb{C}$ denote Clebsch-Gordan coefficients.

Why is the tensor product decomposition useful?

Given two fragments \hat{f}^{ℓ_1} and \hat{f}^{ℓ_2} , then

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

is **non-linear** in f and **rotationally equivariant** (used shorthand notation for Glebsch-Gordan decomposition).

Non-linearly transforming generalised signals

Simply compute $(C^{l_1, l_2, \ell})^\top (\hat{f}^{l_1} \otimes \hat{f}^{l_2})$ for all pairs of input fragments and collect them into a generalised signal (Kondor et al. 2018).

Tensor-product based activation of generalised signals

A generalised signal $f \in \mathcal{F}_L$ may be equivariantly and non-linearly transformed by an operator $\mathcal{N}_\otimes : \mathcal{F}_L \rightarrow \mathcal{F}_L$ defined as

$$\mathcal{N}_\otimes(f) = \{(C^{l_1, l_2, \ell})^\top (\hat{f}_{t_1}^{l_1} \otimes \hat{f}_{t_2}^{l_2}) : \ell = 0, \dots, L - 1; (l_1, l_2) \in \mathbb{P}_L^\ell; t_1 = 0, \dots, \tau_f^{l_1}; t_2 = 0, \dots, \tau_f^{l_2}\},$$

where for each degree the set

$$\mathbb{P}_L^\ell = \{(l_1, l_2) \in \{0, \dots, L - 1\}^2 : |l_1 - l_2| \leq \ell \leq l_1 + l_2\}$$

is defined in order to avoid the computation of trivially equivariant all-zero fragments.

Computational cost of strictly equivariant layers

For strictly equivariant layers the non-linear transformation is **prohibitively costly**.

Computing $g = \mathcal{N}_{\otimes}(f)$ is

$$\mathcal{O}(C^2L^5),$$

where C is **representational capacity** and L **spatial resolution** (bandlimit):

- $\mathcal{O}(C^2L^3)$ fragments,
- cost of computing each fragment is $\mathcal{O}(L^2)$.

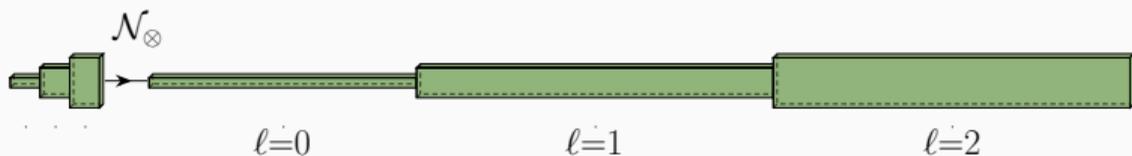
Contributions to improve efficiency

1. Channel-wise structure
2. Constrained generalized convolutions
3. Optimized degree mixing sets
4. Efficient sampling theory on the sphere and rotation group

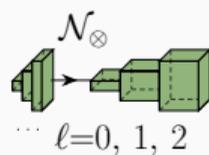
Channel-wise structure

Split generalized signals in K channels and apply a tensor-product activation to each channel separately.

Representational capacity then controlled through **linear dependence** on channels K , **rather than quadratic dependence** (on generalized harmonic representation type τ_f).



Prior approach to applying a tensor-product based non-linear operator



Ours (Cobb et al. 2021)

Constrained generalized convolutions

Under new multi-channel structure, decompose the generalized convolution into **three separate constrained linear operators**:

1. **Uniform convolution**: linear projection uniformly across channels to project down onto the desired type (interpreted as learned extension of tensor-product activations to undo expansion of representation space).
2. **Channel-wise convolution**: linear combinations of the fragments within each channel.
3. **Cross-channel convolution**: linear combinations to learn new features.

Computational and parameter **efficiency significantly improved**.

Optimized degree mixing sets

Non-linear operators must perform degree mixing (equivariant linear operators cannot mix information corresponding to different degrees).

But, it is **not necessary** to compute all possible tensor-product based fragments.

Degree mixing set \mathbb{P}_L^ℓ :

$$\mathbb{P}_L^\ell = \{(\ell_1, \ell_2) \in \{0, \dots, L-1\}^2 : |\ell_1 - \ell_2| \leq \ell \leq \ell_1 + \ell_2\}.$$

Consider subsets of \mathbb{P}_L^ℓ that scale better than $\mathcal{O}(L^2)$.

Optimized degree mixing sets

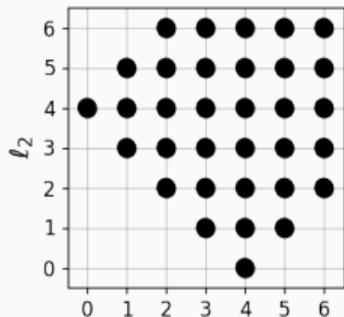
Consider the graph $G_L^\ell = (\mathbb{N}_L, \mathbb{P}_L^\ell)$ with nodes $\mathbb{N}_L = \{0, \dots, L - 1\}$ and edges \mathbb{P}_L^ℓ .

- Some notion of relationship between ℓ_1 and ℓ_2 is captured if there exists a path between the two nodes in G_L^ℓ .
- Select smallest subgraph such that all relationships are preserved \Rightarrow **minimum spanning tree** (MST). Weight edges by computational cost to minimise overall cost.
- Consider **logarithmic subsampling** (reduced MST).

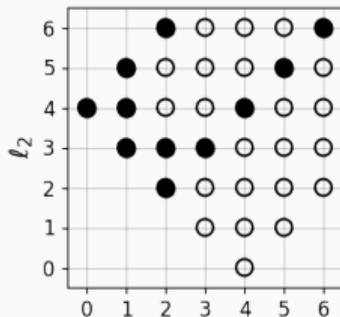
Computational complexity significantly reduced from $\mathcal{O}(L^5)$ to $\mathcal{O}(L^3 \log L)$, where L denotes resolution (bandlimit).

Optimised degree mixing sets

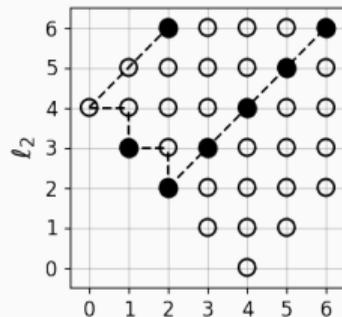
Visualization of the degree mixing set \mathbb{P}_L^ℓ for $L = 7$ and $\ell = 4$.



Full \mathbb{P}_L^ℓ set of size $\mathcal{O}(L^2)$



MST subset of size $\mathcal{O}(L)$



RMST subset of size $\mathcal{O}(\log L)$

Efficient sampling theory and fast harmonic transforms

Adopt **efficient sampling theory** and **fast algorithms** to compute harmonic transforms on the sphere and rotation group.

Leverage to access underlying continuous signal representations, **avoiding discretization artifacts**, and **compute fast convolutions**.

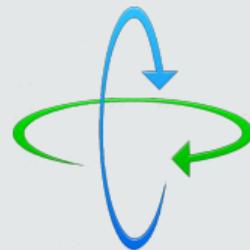
Novel sampling theorem on sphere (McEwen & Wiaux 2011)



SSHT: Spin spherical harmonic transforms

www.spinsht.org

Novel sampling theorem on rotation group (McEwen et al. 2015)



SO3: Fast Wigner transforms on rotation group

www.sothree.org